

ĐẠI HỌC ĐÀ NẴNG
TRƯỜNG ĐẠI HỌC KINH TẾ
-----oOo-----

PHƯƠNG PHÁP NGHIÊN CỨU TRONG KINH DOANH

Lê Văn Huy, Trương Trần Trâm Anh

**Nhà xuất bản Tài chính
2024 (Tái bản)**

UNIVERSITY OF DANANG
THE UNIVERSITY OF ECONOMICS
-----o0o-----

BUSINESS RESEARCH METHODS

Huy L. V, Anh T. T. T.

Financial Publishing
2024

LỜI MỞ ĐẦU

Nghiên cứu là tìm kiếm kiến thức mới. Nghiên cứu khoa học là cách thức con người tìm hiểu các hiện tượng khoa học một cách có hệ thống; là quá trình áp dụng các ý tưởng, nguyên lý và phương pháp khoa học để tìm ra các kiến thức mới nhằm giải thích hay dự báo các sự vật hiện tượng trong thế giới khách quan. Để thực hiện nghiên cứu một cách khoa học, người nghiên cứu không những phải biết các quy trình nghiên cứu mà còn phải nắm vững các triết lý, phương pháp nghiên cứu và sử dụng các công cụ phân tích một cách thích hợp.

Qua quá trình tham gia giảng dạy môn phương pháp nghiên cứu khoa học, nhận thấy việc thực hiện những đề tài nghiên cứu khoa học và luận án tốt nghiệp của sinh viên đại học và học viên cao học gặp nhiều khó khăn do nguồn tài liệu tiếng Việt để trang bị cho học viên nhận thức, phương pháp, công cụ, quy trình trong thực hiện nghiên cứu còn hạn chế. Nhiều giáo trình chỉ giới thiệu những vấn đề lý thuyết về nghiên cứu khoa học, một số giáo trình chỉ giới thiệu phương pháp phân tích làm cho các học viên khó khăn trong việc tiếp cận xuyên suốt lý thuyết và kỹ thuật phân tích. Giáo trình này ra đời với mong muốn cung cấp cho học viên những vấn đề cơ bản về lý thuyết nghiên cứu khoa học, đồng thời, giới thiệu quy trình, phương pháp ứng dụng phần mềm phân tích dữ liệu để thực hiện nghiên cứu.

Bố cục của cuốn sách gồm 10 chương, chương 1 giới thiệu tổng quan về nghiên cứu khoa học sẽ giúp người đọc nắm được “ngôn ngữ” dùng trong nghiên cứu và tiến trình thực hiện một nghiên cứu khoa học. Chương 2 với những ví dụ sinh động từ quá trình nghiên cứu của tác giả sẽ gợi mở cho người đọc cách phát hiện ý tưởng nghiên cứu và xác định vấn đề nghiên cứu. Chương 3 mô tả bức tranh tổng thể về cách mà đề tài nghiên cứu được thực hiện trên cơ sở thiết kế nghiên cứu, chương 4 giới thiệu các phương pháp và công cụ thực hiện nghiên cứu định tính trong nghiên cứu khoa học, chương 5 và 6 trình bày cách thức thực hiện đo lường khái niệm nghiên cứu, thiết kế công cụ thu thập dữ liệu, chọn mẫu trong nghiên cứu và tổ chức thu thập dữ liệu. Các chương từ 7, 8, 9 và 10 giúp bạn đọc làm quen với phần mềm SPSS và thực hành các công cụ phân tích dữ liệu. Dữ liệu sử dụng trong phân tích được thiết kế thông qua một dự án nghiên cứu của tác giả, từ việc xây dựng công cụ thu thập dữ liệu, tổ chức thu thập và tiến hành phân tích, người đọc sẽ nhận thấy việc sử dụng nhiều công cụ để phân tích trên một nguồn dữ liệu, điều này giúp cho người đọc có những góc nhìn khác nhau trong việc sử dụng các công cụ phân tích và khai phá dữ liệu.

Các công cụ phân tích dữ liệu được hướng dẫn bao gồm kiểm định T-test, ANOVA, tương quan và hồi quy đơn, phân tích nhân tố khám phá (EFA), đánh giá độ tin cậy của thang đo (Cronbach Alpha), hồi quy bội (Multi - Regression), MANOVA, phân tích tách biệt (Discriminant Analysis) và Binary Logistic, phân tích nhóm (cluster analysis) và kiểm định phi tham số.

Ý tưởng và nội dung của cuốn sách được hình thành và hoàn thiện dựa trên các nguồn tài liệu trong và ngoài nước, đặc biệt ảnh hưởng bởi tư tưởng về phương pháp nghiên cứu trong các cuốn sách của Cooper và Schindler (2011); Nguyễn Đình Thọ (2011); Hair, Black, Babin và Anderson (2009); Saunders, Lewis và Thornhill (2007); Kothari (2004); Creswell (2003)...., kinh

nghiệm nghiên cứu và giảng dạy của tác giả với sự đóng góp của các đồng nghiệp trong và ngoài Trường Đại học Kinh tế - Đại học Đà Nẵng.

Tập thể tác giả xin trong trọng cảm ơn các đồng nghiệp trong và ngoài Trường Đại học Kinh tế - Đại học Đà Nẵng, các bạn học viên cao học và đặc biệt là Giáo sư Trương Bá Thanh đã động viên, khuyến khích cho tác giả trong quá trình hình thành ý tưởng và viết cuốn sách.

Xin dành những tình cảm sâu sắc, lời cảm ơn đến những người thân trong gia đình đã chia sẻ, động viên và dành thời gian để tác giả có thể chuyên tâm trong việc phát triển cuốn sách.

Xin tri ân Giáo sư Naoufel Daghfous (Trường Đại học Québec vùng Montréal – Canada), Giáo sư Jean Francois Lemoine, Frantz Rowe (Trường Đại học Nantes - Pháp) vì sự thông hiểu, những kiến thức khoa học và sự giúp đỡ vô hạn của Quý Giáo sư trong cung cấp những kinh nghiệm trong thực hiện nghiên cứu khoa học, đặc biệt đã cung cấp những tài liệu và ý kiến quý báu cho việc xây dựng nội dung cuốn sách.

Tuy nhiên, trong quá trình thực hiện không tránh khỏi những thiếu sót, tác giả kính mong nhận được những góp ý, phê bình của bạn đọc để cuốn sách được hoàn thiện hơn trong những lần tái bản sau.

Tập thể tác giả:

**Lê Văn Huy
Trương Trần Trâm Anh**

CHƯƠNG

1

TỔNG QUAN VỀ PHƯƠNG PHÁP NGHIÊN CỨU

(INTRODUCTION TO RESEARCH METHODS)

ĐỊNH NGHĨA NGHIÊN CỨU VÀ NGHIÊN CỨU KINH DOANH

Nghiên cứu được xem là một môn khoa học, là một hệ thống tìm kiếm thông tin thích hợp trên một chủ đề cụ thể nhằm phát hiện các kiến thức. *Nghiên cứu là quá trình thu thập, phân tích dữ liệu một cách có hệ thống nhằm khám phá các vấn đề liên quan (Kothari, 2004).* Như vậy, nghiên cứu là quá trình thu thập và phân tích dữ liệu. Tuy nhiên, quá trình này cần thực hiện một cách có hệ thống dựa trên nền tảng những mối quan hệ logic nhằm khám phá, phát hiện vấn đề. Nghiên cứu cần phải dựa trên dữ liệu được thu thập một cách có hệ thống; được phân tích, diễn giải đúng kỹ thuật và có mục đích rõ ràng. Để làm được điều đó, nghiên cứu cần giải thích một cách rõ ràng phương pháp để thu thập dữ liệu, giải thích tại sao những kết quả đạt được lại có ý nghĩa cùng những hạn chế liên quan nhằm trả lời cho một hoặc nhiều câu hỏi đưa ra. Mục đích của nghiên cứu, theo Kothari (2004) là khám phá những câu trả lời cho các câu hỏi thông qua thực hiện qui trình mang tính khoa học hay tìm ra sự thật ẩn dấu hoặc sự thật chưa được khám phá. Mặc dù mỗi nghiên cứu có mục đích của riêng mình, nhưng tựu trung mục đích của nghiên cứu thường là:

- Để hiểu hơn về hiện tượng thông qua việc có được những hiểu biết mới về hiện tượng (gọi là nghiên cứu khám phá).
- Để mô tả chính xác những đặc điểm về một cá nhân, tình huống hoặc một nhóm cụ thể (gọi là nghiên cứu mô tả).
- Để xác định tần suất sự việc, sự kiện xảy ra hoặc tần suất xảy ra những sự việc/ sự kiện/ đặc điểm này liên quan đến tần suất xảy ra những sự việc, sự kiện khác (gọi là nghiên cứu chẩn đoán (diagnostic research study)).
- Để kiểm định một giả thuyết của mối quan hệ nhân quả giữa những biến số (nghiên cứu có mục tiêu này gọi là nghiên cứu kiểm định giả thuyết (hypothesis-testing research studies))

Sử dụng định nghĩa trên về nghiên cứu, *“nghiên cứu trong kinh doanh (business research) là quá trình điều tra một cách có hệ thống nhằm cung cấp thông tin hướng dẫn ra các quyết định quản trị, đây là quá trình hoạch định, điều tra/ tìm kiếm, phân tích và phổ biến những dữ liệu, thông tin có ý nghĩa cho người ra quyết định một cách linh hoạt và phù hợp để tối đa hóa hiệu suất, năng lực của tổ chức” (Cooper & Schindler, 2011).*

Tùy theo mục đích, cấp độ và hình thức nghiên cứu, người ta có thể phân nghiên cứu thành nghiên cứu cơ bản hoặc nghiên cứu ứng dụng; nghiên cứu khám phá, mô tả hay nhân quả; nghiên cứu định tính hay nghiên cứu định lượng. Cooper và Schindler (2011) cho rằng, mục đích của nghiên cứu trong kinh doanh là:

- Xác định và định nghĩa các cơ hội và vấn đề;
- Xác định, kiểm soát và điều chỉnh chiến lược;
- Xác định, kiểm soát và điều chỉnh các hoạt động tác nghiệp;
- Nâng cao hiểu biết những lĩnh vực quản trị khác nhau.

PHƯƠNG PHÁP NGHIÊN CỨU VÀ PHƯƠNG PHÁP LUẬN NGHIÊN CỨU

Để hiểu hơn về các nội dung được trình bày ở các chương sau, cần phải hiểu và phân biệt phương pháp nghiên cứu (research methods) và phương pháp luận nghiên cứu (research methodology). *Phương pháp nghiên cứu có thể được hiểu là tất cả những phương pháp/ kỹ thuật được sử dụng cho việc nghiên cứu (Saunders, Lewis, & Thornhill, 2007).* Đôi khi, cũng có một sự phân biệt giữa kỹ thuật nghiên cứu (research techniques) và phương pháp nghiên cứu (research methods). Kỹ thuật nghiên cứu (research techniques) đề cập đến hành vi và các công cụ mà chúng ta sử dụng trong việc thực hiện các hoạt động nghiên cứu chẳng hạn như quan sát, ghi chép dữ liệu, xử lý dữ liệu và các hoạt động tương tự. Phương pháp nghiên cứu đề cập đến các công cụ được sử dụng trong việc lựa chọn và xây dựng kỹ thuật nghiên cứu. Trên thực tế, hai thuật ngữ có thể được hoán đổi cho nhau và khi chúng ta nói về những phương pháp nghiên cứu cũng tức là bao gồm cả các kỹ thuật nghiên cứu trong phạm vi của nó. Phương pháp nghiên cứu có thể phân thành ba nhóm sau đây:

- Nhóm gồm những phương pháp có liên quan đến thu thập dữ liệu;
- Nhóm gồm những kỹ thuật thống kê được sử dụng cho việc thiết lập mối quan hệ giữa các dữ liệu và những vấn đề chưa biết;
- Nhóm gồm những phương pháp được sử dụng để đánh giá tính chính xác của các kết quả thu được.

Theo Kothari (2004), *phương pháp luận nghiên cứu là một cách giải quyết có hệ thống vấn đề nghiên cứu, nó có thể được xem như là một ngành khoa học nghiên cứu cách thức thực hiện nghiên cứu một cách khoa học.* Thật cần thiết cho nhà nghiên cứu không những chỉ biết các phương pháp/ kỹ thuật nghiên cứu mà còn biết cả phương pháp luận. Các nhà nghiên cứu không những cần biết làm thế nào để phát triển các chỉ số hoặc việc kiểm tra nhất định, làm thế nào để tính toán giá trị trung bình, phương sai, giá trị t-student, Fisher..., làm thế nào để áp dụng các kỹ thuật nghiên cứu cụ thể mà còn cần phải biết những phương pháp hoặc kỹ thuật này có liên quan với nhau hay không, chúng có ý nghĩa như thế nào và tại sao sử dụng để phân tích trong các trường hợp đó.

Các nhà nghiên cứu cũng cần phải hiểu những giả thuyết trên cơ sở những kỹ thuật khác nhau và họ cần phải biết các tiêu chí có thể quyết định các kỹ thuật. Tất cả điều này có nghĩa là nhà nghiên cứu cần thiết phải thiết kế phương pháp luận cho vấn đề của mình. Người nghiên cứu trước khi thực hiện cần phải biết và giải thích tại sao phải nghiên cứu vấn đề đó, để nghiên cứu cần phải sử dụng các phương pháp nào...

Từ những điều đã nêu ở trên, có thể nói rằng phương pháp luận nghiên cứu có nhiều khía cạnh và phương pháp nghiên cứu cấu thành phương pháp luận nghiên cứu. Theo Kothari (2004), phương pháp luận nghiên cứu được hiểu là khoa học của việc nghiên cứu hay làm thế

nào nghiên cứu được thực hiện một cách khoa học. Người nghiên cứu không những phải biết các phương pháp nghiên cứu mà còn phải nắm vững phương pháp luận nghiên cứu. Người nghiên cứu không những cần biết làm thế nào để thực hiện các kiểm định, làm thế nào để tính toán số trung bình, trung vị, độ lệch chuẩn, làm thế nào để thực hiện một phương pháp nghiên cứu nào đấy mà họ còn phải biết **phương pháp nào phù hợp, phương pháp nào không, ý nghĩa của phương pháp**. Nhà nghiên cứu cũng phải hiểu những giả định/ giả thiết (assumptions) cho các kỹ thuật khác nhau và cần phải biết tiêu chuẩn, theo đó họ có thể quyết định kỹ thuật và qui trình nghiên cứu nào phù hợp với một vấn đề nghiên cứu cụ thể. Tóm lại, phương pháp luận nghiên cứu là những **lý luận/ logic**, dựa vào đó người nghiên cứu có thể lựa chọn phương pháp/ kỹ thuật nghiên cứu và giải thích vì sao phương pháp/ kỹ thuật nghiên cứu đó được lựa chọn. Phạm vi của phương pháp luận nghiên cứu rộng hơn phương pháp nghiên cứu bởi vì khi chúng ta bàn đến phương pháp luận nghiên cứu chúng ta không chỉ nói đến phương pháp nghiên cứu mà còn xem xét logic đằng sau những phương pháp nghiên cứu mà chúng ta sử dụng và giải thích tại sao chúng ta lại sử dụng phương pháp nghiên cứu này mà không phải là phương pháp nghiên cứu khác, qua đó, chính bản thân người nghiên cứu có thể đánh giá được nghiên cứu của mình. Phương pháp luận nghiên cứu trả lời các câu hỏi:

- Tại sao nghiên cứu lại được tiến hành?
- Vấn đề nghiên cứu này được định nghĩa như thế nào?
- Tại sao và bằng cách nào những giả thuyết này được đưa ra?
- Dữ liệu gì được thu thập? Phương pháp nào dùng để thu thập dữ liệu?
- Tại sao kỹ thuật phân tích dữ liệu này được sử dụng?

NGHIÊN CỨU KHOA HỌC VÀ PHƯƠNG PHÁP KHOA HỌC

Theo Babbie (1986), *nghiên cứu khoa học là cách thức con người tìm hiểu các hiện tượng khoa học một cách có hệ thống*. Trên thực tế, *nghiên cứu khoa học là quá trình áp dụng các ý tưởng, nguyên lý và phương pháp khoa học để tìm ra các kiến thức mới nhằm giải thích hay dự báo các sự vật hiện tượng trong thế giới khách quan*.

Mục tiêu của nghiên cứu khoa học là để tìm ra câu trả lời cho những câu hỏi, tìm ra sự thật được ẩn đi và chưa được phát hiện thông qua việc áp dụng các thủ tục khoa học. Theo Gill và Johnson (2002), để hiểu biết một sự việc, có hai cách là (1) chấp nhận: cách thức con người hiểu biết sự việc thông qua việc thừa nhận các nghiên cứu hay kinh nghiệm của người khác hoặc (2) nghiên cứu: cách thức con người tìm hiểu sự việc thông qua việc thực hiện các nghiên cứu.

Phần này sẽ trình bày các quan điểm về phương pháp nghiên cứu và phương pháp khoa học theo quan điểm của Kothari (2004), theo đó phương pháp nghiên cứu và phương pháp khoa học liên quan chặt chẽ với nhau.

Như chúng ta đã đề cập, phương pháp nghiên cứu là một cách giải quyết có hệ thống vấn đề nghiên cứu, nó có thể được xem như là một ngành khoa học nghiên cứu cách thức thực hiện nghiên cứu một cách khoa học. Thông thường, nhà nghiên cứu rất quan tâm đến kết quả cụ

thể, sự lặp lại của các kết quả và mức độ phù hợp của chúng trong những tình huống phức tạp và tổng quát hơn.

Trong khi đó, phương pháp khoa học là một và giống nhau trong các ngành (khoa học). Phương pháp đó là phương pháp của tất cả kiến thức được đào tạo một cách hợp lý, thống nhất của tất cả các ngành khoa học. Nó là việc theo đuổi sự thật được xác định bởi những cân nhắc hợp lý, lý tưởng của khoa học là đạt được một hệ thống mối quan hệ qua lại của các nhân tố. Vì vậy, phương pháp khoa học cố gắng để đạt được lý tưởng này bằng thử nghiệm, quan sát, lập luận hợp lý được mặc nhiên chấp nhận và sự kết hợp của ba điều này trong các sự cân đối khác nhau. Phương pháp khoa học được mặc nhiên chấp nhận trong các trường hợp sau:

- Tính duy lý (rationalism): Các nhận xét, giải thích hay kết luận phải dựa trên những suy luận logic.
- Thực chứng (empiricism): Các nhận xét, giải thích hay kết luận rút ra phải được dựa trên các quan sát thực tiễn.

Tuy nhiên, không phải khi nào kiến thức đều mặc nhiên chấp nhận mà phải thực hiện nghiên cứu, như đã đề cập, việc thực hiện nghiên cứu phải thực hiện thông qua một phương pháp khoa học, sử dụng các công cụ hợp lý nhằm thu thập, phân tích và diễn giải dữ liệu theo đúng bản chất của sự vật, hiện tượng.

CÁC THUẬT NGỮ VỀ PHƯƠNG PHÁP NGHIÊN CỨU

Khi thực hiện nghiên cứu, chúng ta mong muốn hiểu, giải thích và dự đoán các hiện tượng. Giả sử chúng ta mong muốn đánh giá sự hài lòng của nhân viên đối với tổ chức, chúng ta phải hiểu và đồng ý về ý nghĩa của các khái niệm như sự hài lòng của nhân viên hoặc muốn giải thích tại sao khách hàng mua dầu gội đầu X-men, cần phải hiểu hành vi của khách hàng khi mua dầu gội đầu và hành vi mua của khách hàng với sản phẩm X-men. Nói cách khác, chúng ta chỉ ra ý nghĩa của những hiện tượng mà mình muốn nghiên cứu, giải thích cách chúng ta sẽ đánh giá, quan sát hiện tượng. Trong nghiên cứu, có nhiều thuật ngữ được sử dụng như là ngôn ngữ để giải thích về hiện tượng nghiên cứu. Trong các tài liệu tiếng Anh, chúng ta có thể bắt gặp các từ khác nhau như concepts, construct, definition..., tuy nhiên, diễn đạt bằng tiếng Việt vẫn chưa cho sự thống nhất nhất định. Do vậy, để dễ dàng hơn trong việc tiếp cận, trong khuôn khổ quyển sách này, chúng ta có thể tạm hiểu những thuật ngữ được sử dụng để tạo nên cơ sở khoa học cho việc triển khai thực hiện nghiên cứu:

Khái niệm (Concepts)

Khái niệm

Để hiểu và truyền thông về những đối tượng, sự kiện, chúng ta cần phải dựa vào một nền tảng, những điểm chung được hiểu và chấp nhận. Khái niệm (concepts) hay còn gọi là khái niệm lý thuyết phục vụ mục đích này.

“Một khái niệm lý thuyết là một tập hợp những ý nghĩa và đặc điểm được chấp nhận rộng rãi liên quan đến những sự kiện, đối tượng, điều kiện, tình huống và hành vi cụ thể”.

“A generally accepted collection of meanings or characteristics associated with certain events, objects, conditions, situations and behavior” (Cooper & Schindler, 2011).

“Phân loại những sự kiện hay đối tượng có những đặc điểm chung vượt ra ngoài khuôn khổ của những quan sát đơn lẻ tạo nên khái niệm lý thuyết”.

“Classifying and categorizing objects or events that have common characteristics beyond any single observation create concepts” (Cooper & Schindler, 2011).

“Khái niệm là sự diễn tả trừu tượng sự vật/ hiện tượng bằng cách tổng quát hóa từ các đặc điểm cụ thể” (Kerlinger, 1986).

“.... expresses an abstraction formed by generalization from particulars” (Kerlinger, 1986).

Khi bạn nghĩ đến máy tính hay xe đạp, những gì hiện ra trong tâm trí bạn không phải là một ví dụ đơn lẻ mà là tập hợp những liên tưởng về máy tính và xe đạp, từ đó bạn rút ra một tập hợp những đặc điểm được xác định cụ thể. Chúng ta trừu tượng hóa những ý nghĩa này từ kinh nghiệm của chúng ta và sử dụng từ ngữ như các “nhãn (tên)” dán vào chúng. Ví dụ, chúng ta thấy một người đàn ông đi ngang qua và nhận thấy rằng anh ta đang chạy, đi bộ, bò hay nhảy, những chuyển động này chính là các khái niệm.

Nguồn của khái niệm lý thuyết

Theo Cooper và Schindler (2011), những khái niệm được sử dụng rộng rãi và thường xuyên được phát triển theo thời gian thông qua việc sử dụng ngôn ngữ chung. Chúng ta có được những khái niệm này thông qua các kinh nghiệm cá nhân. Những khái niệm thông dụng đóng góp phần lớn vào việc giao tiếp và kể cả trong nghiên cứu, nhưng chúng ta thường rơi vào khó khăn khi đối mặt với một khái niệm không phổ biến hoặc một ý tưởng mới. Một cách để giải quyết vấn đề này là mượn từ những ngôn ngữ khoa học khác. Ví dụ như khái niệm “lực hấp dẫn” (gravity) mượn từ khoa học vật lý để sử dụng trong marketing nhằm giải thích tại sao người tiêu dùng lại mua sắm những thương hiệu cụ thể nào đó hay biểu đồ nhận thức (perceptual map) là một phương pháp sử dụng các điểm, khoảng cách được trình bày trong một không gian nhằm giúp cho nhà nghiên cứu hiểu khoảng cách giữa các thương hiệu, các thuộc tính để từ đó có thể biết được vị trí của các thương hiệu trong tâm trí của khách hàng. Thỉnh thoảng chúng ta cần phát triển những “nhãn mới” (new label) cho các khái niệm mới. Khi đó chúng ta phải bắt đầu phát triển các từ ngữ chuyên môn (specialized jargon/ terminology).

Tầm quan trọng đối với nghiên cứu

Sự thành công của nghiên cứu dựa vào việc (1) chúng ta đã khái quát hóa lý thuyết, mô hình lý thuyết rõ ràng như thế nào và (2) những người khác hiểu khái niệm để chúng ta dùng ở mức độ nào. Ví dụ, khi chúng ta điều tra về lòng trung thành khách hàng, những câu hỏi mà chúng ta dùng cần phải đề cập đến thái độ của khách hàng.

Thái độ là khái niệm trừu tượng, vì vậy chúng ta phải đo lường thái độ bằng việc sử dụng các khái niệm được chọn lựa. Thử thách ở đây là phát triển các khái niệm mà những người khác có thể hiểu rõ. Ví dụ, chúng ta hỏi các đáp viên để ước tính tổng thu nhập của gia đình họ. Đây có thể là một khái niệm đơn giản và rõ ràng, nhưng chúng ta có thể sẽ nhận được những câu trả lời khác nhau và rối rắm trừ khi chúng ta giới hạn lại khái niệm “thu nhập” bằng cách cụ thể hóa: thời gian như hàng tuần, hàng tháng hoặc hàng năm; trước hay sau thuế thu nhập; của người chủ gia đình hay tất cả các thành viên trong gia đình; chỉ tính lương hay các lợi nhuận

khác.... Do vậy, đối với những khái niệm mang tính trừu tượng cần phải có một cách hiểu khái niệm và phát triển thang đo nhằm đo lường chính xác đối tượng cần nghiên cứu.

Khái niệm nghiên cứu (Constructs)

Khái niệm nghiên cứu (constructs) có thể đơn giản hoặc trừu tượng, mức độ ở đó khái niệm lý thuyết có hoặc không có những đặc điểm khách quan để liên tưởng đến. Ví dụ, máy tính là một khái niệm khách quan, chúng ta có thể chỉ một máy tính và chúng ta có các hình ảnh về đặc điểm của tất cả các máy tính trong tâm trí (như màn hình, bộ nhớ, bàn phím, con chuột...). Tuy nhiên, một khái niệm trừu tượng như tính cách (personality), thái độ (attitude), động cơ (motivation) thì khó hơn nhiều để hình dung. Những khái niệm mang tính trừu tượng cao được sử dụng trong nghiên cứu thường được gọi là khái niệm nghiên cứu (constructs).

“Khái niệm nghiên cứu là một hình ảnh hay một ý tưởng trừu tượng được tạo ra cho một nghiên cứu cụ thể và/ hoặc cho mục đích xây dựng lý thuyết”.

“An image or idea specially invented for a given research or theory building purpose” (Cooper & Schindler, 2011).

“Chúng ta xây dựng khái niệm nghiên cứu (constructs) bằng cách kết hợp những khái niệm lý thuyết (concepts) đơn giản và cụ thể hơn, đặc biệt khi ý tưởng hoặc hình ảnh mà chúng ta dự định biểu đạt không phải là đối tượng cho việc quan sát trực tiếp”.

“We build construct by combining the simpler concepts, especially when the idea or image we intend to convey is not directly subject to observation” (Cooper & Schindler, 2011).

Ví dụ nhà nghiên cứu muốn mô tả chất lượng dịch vụ kỹ thuật của các mạng điện thoại di động và chất lượng dịch vụ kỹ thuật không thể “quan sát” trực tiếp được, do vậy, người nghiên cứu phải tìm ra 5 chỉ báo cho phép “quan sát được” như sau:

- | | |
|---|------------------------------------|
| Chất lượng dịch vụ kỹ thuật của mạng điện thoại di động | 1. Kết nối ngay lần gọi đầu tiên |
| | 2. Thông suốt trong suốt cuộc gọi |
| | 3. Kết nối âm thanh rõ ràng |
| | 4. Vùng phủ sóng rộng |
| | 5. Thực hiện được vào giờ cao điểm |

Rõ ràng rằng, chúng ta có thể “quan sát được” việc có hay không kết nối được ngay lần gọi đầu tiên, thông suốt trong suốt quá trình gọi, âm thanh rõ ràng... Như vậy, “kết nối ngay lần gọi đầu tiên”, “thông suốt cuộc gọi”, “kết nối âm thanh rõ ràng”... là những khái niệm trong nhóm các khái niệm hình thành nên khái niệm nghiên cứu (construct) chất lượng dịch vụ kỹ thuật các mạng điện thoại di động. Như vậy, chất lượng dịch vụ kỹ thuật là một khái niệm nghiên cứu (construct) được sử dụng để trình bày tập hợp các ý nghĩa được biểu diễn bởi năm khái niệm trên.

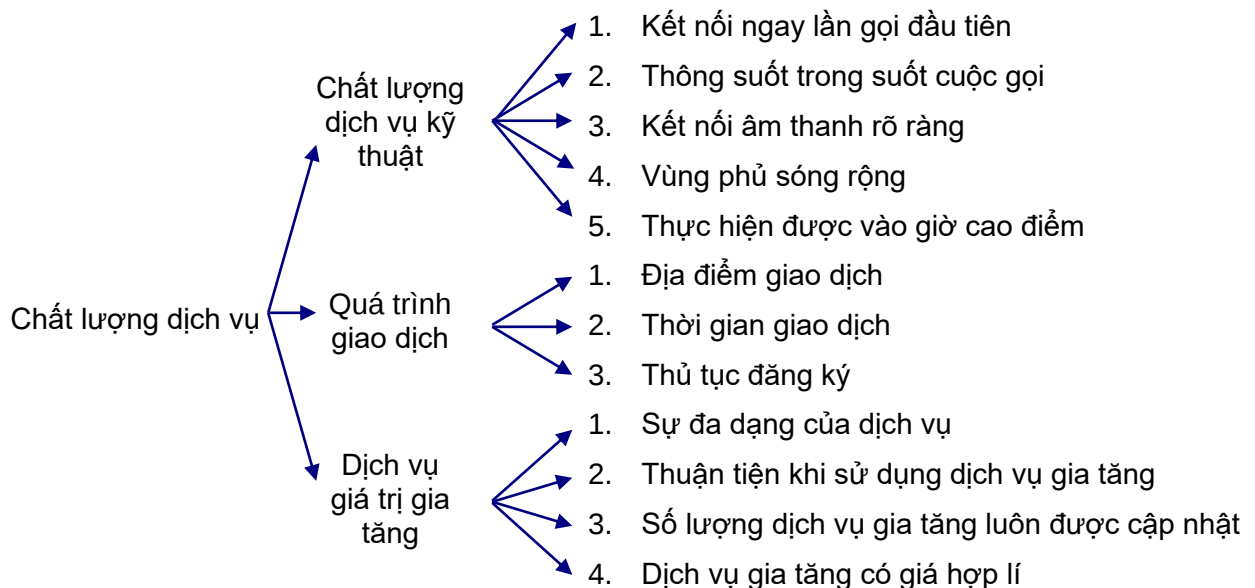
Một khái niệm nghiên cứu có thể đo lường bởi một hoặc nhiều biến tiềm ẩn (latent variables), biến tiềm ẩn là biến được kết nối với các biến quan sát (observed variables, indicators, items...) như động cơ, thái độ của cán bộ công nhân viên hoặc khách hàng. Các biến quan sát được sử dụng hệ thống các con số để biểu thị mức độ theo một quy tắc xác định.

Một khái niệm nghiên cứu (construct) có thể bao gồm tập biến quan sát (ví dụ khái niệm nghiên cứu chất lượng dịch vụ bao gồm tập 5 biến quan sát là kết nối ngay lần gọi đầu tiên, thông suốt trong suốt cuộc gọi, kết nối âm thanh rõ ràng, vùng phủ sóng rộng và thực hiện được vào giờ cao điểm và khi đó, năm biến quan sát này được gọi là thang đo của khái niệm nghiên cứu (construct). Lúc này, khái niệm chất lượng dịch vụ được gọi là khái niệm đơn hướng hay khái niệm bậc nhất.

Ở cấp cao hơn, nhà nghiên cứu nhận thấy ba khái niệm nghiên cứu (ba thành phần) (1) *chất lượng dịch vụ kỹ thuật*, (2) *quá trình giao dịch* (3) *dịch vụ giá trị gia tăng* có liên hệ với nhau và tạo thành khái niệm nghiên cứu có thể được tạm gọi là “**chất lượng dịch vụ**” (service quality). Khi đó, khái niệm chất lượng dịch vụ là khái niệm đa hướng (mutidimensional) hay khái niệm bậc cao.

Như vậy, có nhiều khái niệm rất khó quan sát và đo lường, một khái niệm có thể bao gồm nhiều khái niệm khác nhau mà trong số đó có rất nhiều khái niệm mang tính trừu tượng cao. Những nhà nghiên cứu thường xem những khái niệm nghiên cứu đó là “các khái niệm nghiên cứu mang tính giả thuyết/giả định” (hypothetical constructs) bởi vì chúng chỉ có thể được suy luận từ dữ liệu thu thập được; vì vậy chúng được giả định là tồn tại nhưng phải chờ đến khi được kiểm tra/kiểm định để xem thực sự chúng bao gồm những gì. Nếu nghiên cứu chỉ ra rằng các khái niệm lý thuyết và khái niệm nghiên cứu trong ví dụ này có quan hệ với nhau và nếu các liên kết giữa chúng được chấp nhận, thì chúng ta có được điểm khởi đầu của lược đồ khái niệm (conceptual scheme).

Ở dạng hình ảnh, lược đồ khái niệm biểu diễn tên của các khái niệm, mối quan hệ giữa các khái niệm cũng như tập biến (quan sát) của khái niệm đó.



Định nghĩa (Definitions)

Định nghĩa hay còn gọi là định nghĩa đưa vào nghiên cứu (operational definitions) là định nghĩa được phát biểu dưới dạng những tiêu chuẩn cụ thể để kiểm định và đo lường. Những thuật ngữ được sử dụng trong định nghĩa phải liên quan đến tiêu chuẩn đo lường (nhà nghiên cứu phải đếm, đo lường được những thông tin mà họ thu thập). Cho dù đối tượng được định nghĩa

mang tính vật chất (ví dụ: một lon sữa) hoặc mang tính trừu tượng cao (ví dụ: động cơ thăng tiến), định nghĩa đưa vào nghiên cứu phải cụ thể hóa những đặc điểm của đối tượng và chỉ rõ nhà nghiên cứu sẽ quan sát những đặc điểm này như thế nào. Những đặc điểm cụ thể phải rất rõ ràng để khi bất kỳ nhà nghiên cứu nào sử dụng cũng có thể hiểu cùng một ý nghĩa. Chẳng hạn như sinh viên đại học có thể được phân loại theo lớp. Có thể sẽ không có vấn đề gì khi hiểu các thuật ngữ như sinh viên năm nhất, sinh viên năm hai. Nhưng có thể sẽ không đơn giản nếu chúng ta phải xác định sinh viên nào ở lớp nào. Để làm điều này, chúng ta cần định nghĩa đưa vào nghiên cứu (operational definitions). Định nghĩa đưa vào nghiên cứu có thể thay đổi dựa vào mục đích của nhà nghiên cứu và cách mà nhà nghiên cứu chọn để đo lường.

Ví dụ, đối với những khái niệm nghiên cứu mang tính trừu tượng như “lòng trung thành khách hàng” (customer loyalty), chúng ta có thể hình dung ra ý nghĩa của khái niệm nhưng để đo lường nó trên đối tượng khách hàng thì không đơn giản. Chúng ta có thể phát triển những câu hỏi về hành vi, thái độ hay chúng ta sử dụng thang đo (scale) đã được phát triển và kiểm định bởi các nhà nghiên cứu khác và chính thang đo này định nghĩa khái niệm nghiên cứu (construct).

Mục đích chính của việc định nghĩa khái niệm nghiên cứu là để cung cấp sự thông hiểu và đo lường các khái niệm, những định nghĩa này sẽ luôn được sử dụng để phát triển các mối quan hệ trong các giả thuyết và lý thuyết.

Ví dụ: Các khái niệm nghiên cứu và định nghĩa như:

Khái niệm nghiên cứu	Định nghĩa
Nhận thức thương hiệu	Phần trăm đáp viên nghe đến thương hiệu, nhận thức có thể được trợ giúp hoặc không được trợ giúp.
Thái độ đối với thương hiệu	Số lượng đáp viên và mức độ cảm nhận tích cực hoặc tiêu cực đối với một thương hiệu cụ thể.
Ý định mua	Số người dự định mua một sản phẩm, dịch vụ cụ thể trong một khoảng thời gian nhất định.
Tầm quan trọng của các yếu tố	Những yếu tố và mức độ ảnh hưởng của các yếu tố đến lựa chọn mua.
Các đặc điểm nhân khẩu học	Tuổi, giới tính, nghề nghiệp, thu nhập,... của đối tượng
Lòng trung thành thương hiệu	Đáp viên đã mua/sử dụng sản phẩm bao nhiêu lần.
Đặc điểm tâm lý	Khách hàng nghĩ và cư xử như thế nào

Nguồn: Theo (Hair, Bush, & Ortineau, 2002)

Biến số nghiên cứu (Variables)

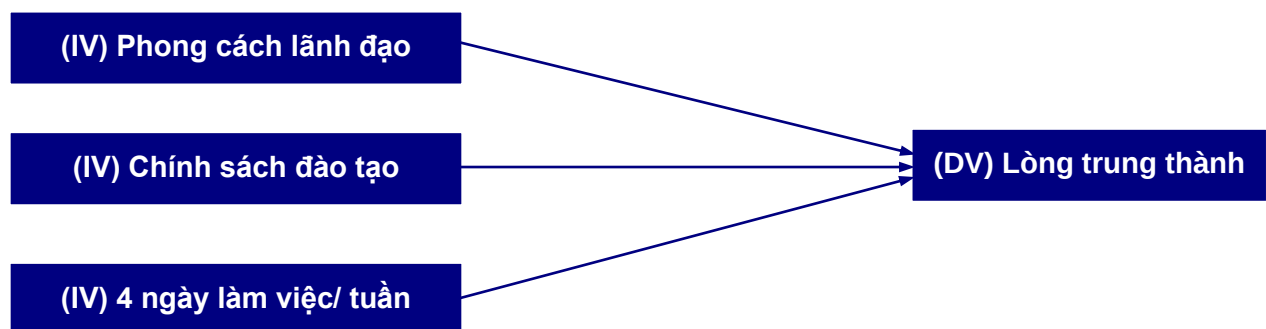
Trong thực tế, thuật ngữ “biến số” được dùng như từ đồng nghĩa của *khái niệm nghiên cứu* (construct) hay các đặc tính được nghiên cứu. Trong bối cảnh này, một biến số là một biểu tượng của sự kiện, hành động, đặc điểm, đặc tính hoặc tính chất có thể đo lường được và ta có thể gán các giá trị vào biến số. Trong quá trình thực hiện, nhà nghiên cứu cần phải phân biệt các loại biến như biến độc lập, biến phụ thuộc, biến điều tiết, biến ngoại vi..., mỗi loại biến cho

những đặc điểm khác nhau, do vậy, trong thiết kế nghiên cứu và phân tích cần phải chú ý để tránh những sai sót.

Biến số độc lập (independent variables - IV) và biến số phụ thuộc (dependent variables - DV)

Các nhà nghiên cứu rất quan tâm đến mối quan hệ giữa các biến số. Ví dụ, liệu phiếu mua hàng giảm giá của một tờ báo (biến số độc lập) có ảnh hưởng đến số lượng đặt báo (biến số phụ thuộc) hoặc các phong cách giao tiếp của người bán hàng có ảnh hưởng đến khả năng duy trì mối quan hệ với khách hàng của anh/ cô ta? Có nhiều tài liệu người ta sử dụng thuật ngữ *biến số dự đoán (predictor variable)* thay cho **biến số độc lập (independent variable - IV)**. Biến số độc lập có tác động/ ảnh hưởng đến biến số phụ thuộc và lưu ý rằng có một vài biến số độc lập cùng ảnh hưởng lên một biến phụ thuộc nhưng những biến số độc lập này có thể liên hệ với nhau và do vậy, giữa chúng có một mối quan hệ nào đó hay nói cách khác chúng không còn độc lập với nhau nữa. Tương tự, thuật ngữ *biến số tiêu chuẩn (criterion variable)* được sử dụng như từ đồng nghĩa của **biến số phụ thuộc (dependent variable - DV)**. Biến số này được đo lường, dự đoán hay được xem xét/ theo dõi và được kỳ vọng bị ảnh hưởng bởi sự thay đổi của biến số độc lập. Hình 1 mô tả mối quan hệ giữa các biến độc lập (phong cách lãnh đạo, chính sách đào tạo, số ngày làm việc/ tuần) và biến phụ thuộc (lòng trung thành của nhân viên đối với tổ chức).

Hình 1. Mối quan hệ giữa biến độc lập (IV) và biến phụ thuộc (DV)



Trong mỗi mối quan hệ, có ít nhất một biến số độc lập (IV) và một biến số phụ thuộc (DV). Chúng ta thường giả định rằng biến độc lập (IV) ảnh hưởng/ tác động đến biến phụ thuộc (DV). Tuy nhiên, cần lưu ý rằng mặc dù rất dễ dàng để nói rằng biến độc lập tác động/ ảnh hưởng đến biến phụ thuộc, nhưng việc chỉ ra rằng mối quan hệ giữa biến độc lập và biến phụ thuộc là mối quan hệ nhân quả là việc không dễ dàng. Trong **Hình 1**, mối quan hệ này được biểu diễn qua mũi tên chỉ từ biến độc lập đến biến phụ thuộc. Đối với những mối quan hệ đơn giản, những biến số khác được xem không có sự ảnh hưởng và được bỏ qua.

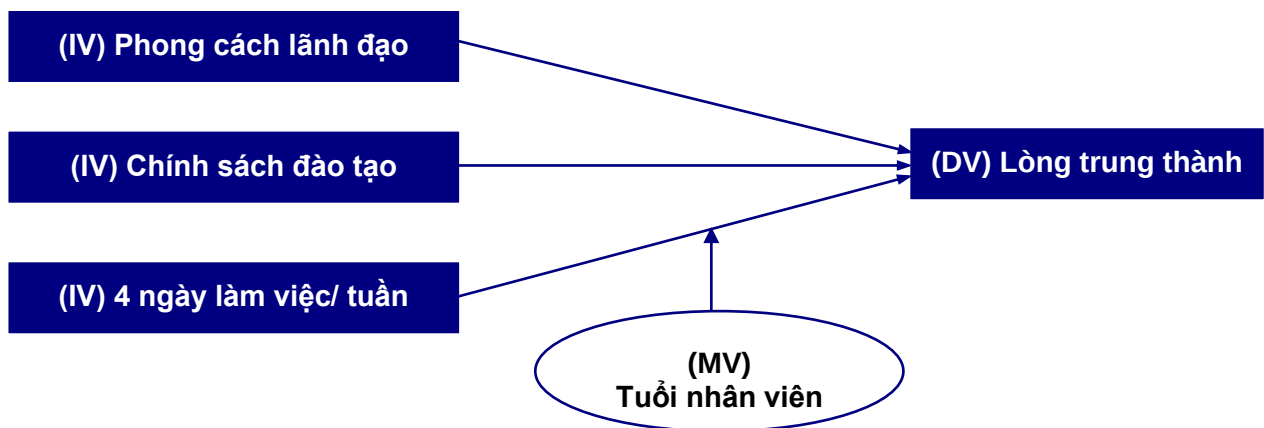
Biến điều tiết (moderating/ interaction variables)

Trong thực tế nghiên cứu, mối quan hệ một - một đơn giản cần được đặt trong những điều kiện hoặc phải được xem xét với sự có mặt của những biến số khác. Thông thường, chúng ta sử dụng một loại biến mang tính giải thích gọi là biến điều tiết (moderating variable - MV). Theo Cooper và Schindler (2011), biến điều tiết là một biến độc lập cấp hai được đưa vào bởi vì nó có sự tác động lớn hoặc tác động ngẫu nhiên (contingent) đến mối quan hệ ban đầu giữa biến

độc lập và biến phụ thuộc. Mũi tên từ biến điều tiết đến mũi tên giữa biến độc lập và biến phụ thuộc trong **Hình 2** chỉ ra sự khác nhau giữa sự tác động trực tiếp của biến độc lập lên biến phụ thuộc và biến điều tiết tác động lên mối quan hệ giữa biến độc lập và biến phụ thuộc. Ví dụ, có thể giả định như sau:

Việc đưa qui định một tuần làm 4 ngày (IV) dẫn đến sự hài lòng (DV) cao hơn, đặc biệt đối với những nhân viên trẻ (MV).

Hình 2. Mối quan hệ giữa biến độc lập (IV) và biến phụ thuộc (DV) có biến điều tiết (MV)



Trong trường hợp này, có một sự khác nhau trong mối quan hệ giữa bốn ngày làm việc một tuần và lòng trung thành của nhân viên đối với tổ chức giữa những nhân viên có độ tuổi khác nhau. Vì thế, sau khi đưa qui định tuần làm việc 4 ngày, lòng trung thành tăng thêm của những nhân viên trẻ lớn hơn của những nhân viên lớn tuổi. Cần chú ý rằng tác động của biến điều tiết là sự tăng thêm xảy ra đồng thời của việc đưa qui định tuần làm việc 4 ngày và tuổi của nhân viên. Ví dụ, giả định rằng lòng trung thành của một nhân viên trẻ là 12% nhiều hơn so với một nhân viên lớn tuổi và lòng trung thành của nhân viên khi làm việc 4 ngày trong tuần là 6% cao hơn những người làm việc 5 ngày trong tuần. Nếu lòng trung thành của nhân viên trẻ làm việc 4 ngày một tuần là 18% cao hơn lòng trung thành của nhân viên lớn tuổi làm việc 5 ngày một tuần, thì không có tác động điều tiết vì 18% là tổng của tác động chính đến từ biến độc lập. Tác động điều tiết sẽ xảy ra nếu lòng trung thành của nhân viên trẻ làm việc 4 ngày một tuần trên 18%, cao hơn lòng trung thành của nhân viên lớn tuổi làm việc 5 ngày một tuần.

Việc xem một biến là biến độc lập hay biến điều tiết sẽ phụ thuộc vào giả định/ giả thuyết nghiên cứu, nếu chúng ta muốn nghiên cứu sự tác động của số ngày làm việc lên lòng trung thành thì số ngày làm việc sẽ là biến độc lập, nếu chúng ta tập trung vào mối quan hệ giữa độ tuổi của nhân viên và lòng trung thành thì biến độc lập sẽ là độ tuổi của nhân viên.

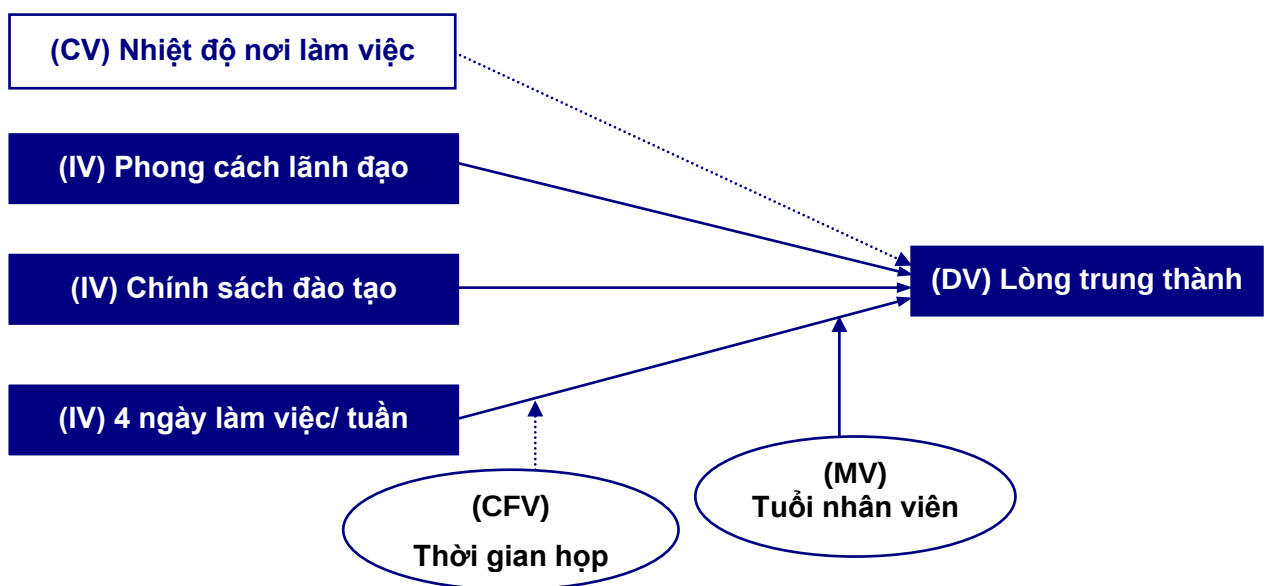
Biến ngoại vi (extraneous variables - EV)

Có vô số biến ngoại vi (EV) tồn tại có thể ảnh hưởng đến mối quan hệ được xem xét. Một số có thể được xem như biến độc lập (IV) hoặc biến điều tiết (MV) nhưng hầu như phải được giả định hoặc bị loại ra khỏi nghiên cứu. May mắn là phần lớn những biến số sẽ ít ảnh hưởng hoặc không ảnh hưởng trong tình huống nghiên cứu. Hầu hết có thể được loại bỏ một cách an toàn bởi vì sự tác động của những biến này diễn ra một cách ngẫu nhiên hoặc rất ít. Những biến khác có thể ảnh hưởng đến biến phụ thuộc nhưng sự ảnh hưởng của những biến này không phải là vấn đề chính mà chúng ta nghiên cứu. Tuy nhiên, chúng ta muốn kiểm tra liệu kết quả

của nghiên cứu có bị ảnh hưởng bởi chúng. Vì vậy, chúng ta đưa những biến này vào nghiên cứu như là những biến kiểm soát/ biến điều khiển (**control variables - CV**) để đảm bảo rằng kết quả của chúng ta không thiên lệch (biased) khi không đưa chúng vào. Lấy ví dụ lại sự tác động của số ngày làm việc lên lòng trung thành, chúng ta có thể nghĩ rằng điều kiện về thời tiết, nhiệt độ tại nơi làm việc, thuế thu nhập tăng, số ngày họp, sự kiện vị chủ tịch thành phố mới nhậm chức và hàng ngàn sự kiện, điều kiện tương tự có thể có ít nhiều ảnh hưởng đến số ngày làm việc và lòng trung thành.

Biến ngoại vi (EV) có thể là biến số gây nhiễu (**confounding variables - CFV**) cho mối quan hệ giữa biến độc lập và biến phụ thuộc (IV – DV). Bạn có thể cho rằng loại hình công việc có thể có ảnh hưởng đến sự tác động của số ngày làm việc lên lòng trung thành. Vì thế bạn có thể đưa thời gian tham dự cuộc họp bàn về công việc như là một biến gây nhiễu (CFV). Trong ví dụ hiệu suất làm việc, chúng ta cố gắng kiểm soát loại hình công việc bằng cách nghiên cứu sự tác động của bốn ngày làm việc một tuần lên lòng trung thành với công việc tham gia họp ở mức độ khác. Trong **Hình 3**, thời tiết, nhiệt độ tại nơi làm việc có thể là biến ngoại vi, đường đứt nét chỉ ra rằng chúng ta có thể đưa nó vào nghiên cứu vì nó có thể ảnh hưởng đến biến độc lập và biến gây nhiễu trong ví dụ này là loại hình công việc.

Hình 3. Mối quan hệ giữa biến độc lập (IV) và biến phụ thuộc (DV) có biến điều tiết (MV) và biến ngoại vi (CV, CFV)



Biến can thiệp (Intervening variables - IVV)

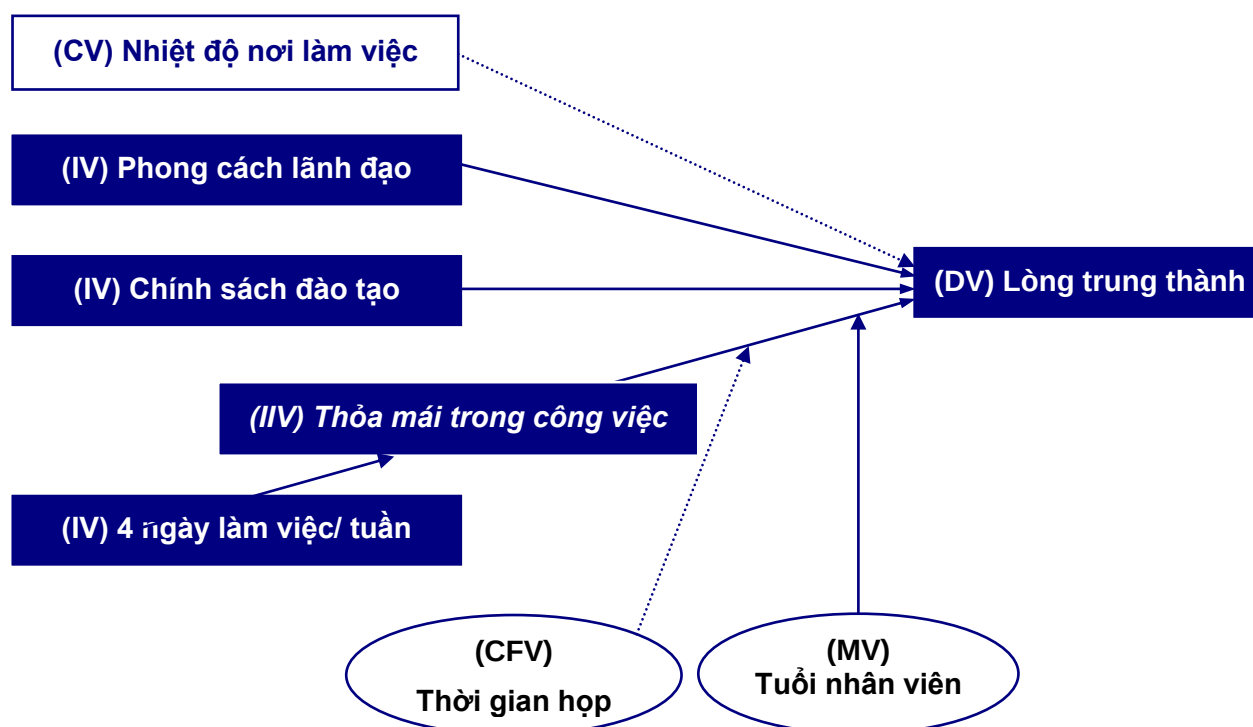
Các biến số được đưa vào nghiên cứu là những biến số cụ thể và đo lường được, tuy nhiên, một biến số có thể không hoàn toàn thỏa mãn điều kiện đó. Vì vậy, trong khi chúng ta công nhận rằng bốn ngày làm việc một tuần có thể đem lại lòng trung thành cao hơn, chúng ta có thể suy luận rằng đó không phải là toàn bộ câu chuyện. Số ngày làm việc có thể ảnh hưởng đến một số biến can thiệp (intervening variables - IVV), sau đó những biến này tác động đến hiệu suất làm việc. IVV là biến thông qua nó, biến độc lập và biến điều tiết tác động đến biến phụ thuộc. IVV có thể được định nghĩa như một yếu tố mà theo lý thuyết có tác động đến biến phụ thuộc nhưng không thể quan sát được hoặc chưa thể đo lường được; tác động của nó, có thể được suy ra từ tác động của biến độc lập và biến điều tiết lên hiện tượng được nghiên cứu.

Trong ví dụ hiệu suất làm việc, biến số can thiệp có thể là sự thỏa mái công việc, chẳng hạn như có một giả thuyết như sau:

Việc đưa qui định tuần làm việc bốn ngày sẽ dẫn đến việc tăng hiệu suất làm việc thông qua tăng sự thoải mái đối với công việc.

Ở đây chúng ta giả định rằng bốn ngày làm việc trong tuần tăng lòng trung thành đối với tổ chức; tương tự, chúng ta có thể giả định rằng tham dự cuộc họp nội bộ có mối quan hệ nghịch đến tính chất thường ngày/ lặp lại của công việc.

Hình 4. Mối quan hệ giữa biến độc lập (IV) và biến phụ thuộc (DV) có biến điều tiết (MV), biến can thiệp (IIV) và biến ngoại vi (CV, CFV)



LÝ THUYẾT KHOA HỌC

Theo Kerlinger (1986), Saunders và cộng sự (2007), lý thuyết khoa học là “một tập hợp những khái niệm, định nghĩa và giả thuyết trình bày có hệ thống thông qua mối quan hệ giữa các khái niệm nhằm giải thích và dự báo các hiện tượng”.

Theo Saunders và cộng sự (2007), lý thuyết khoa học phải bao gồm các khái niệm và mối quan hệ giữa các khái niệm (được gọi là giả thuyết), lý thuyết khoa học phải biểu hiện các vấn đề thực tại, có thể giải thích và dự báo các hiện tượng. Như vậy, để tạo ra lý thuyết khoa học hoặc thực hiện việc nghiên cứu khoa học, nhà nghiên cứu có thể thực hiện bằng cách (1) xem xét và hình thành các khái niệm cũng như thiết kế mối quan hệ giữa các khái niệm đó, (2) kiểm định các mối quan hệ hoặc (3) thực hiện cả hai công việc này.

Một lý thuyết khoa học cần phải trả lời một cách rõ ràng câu hỏi nghiên cứu trong một phạm vi xác định, phải dựa trên những lý thuyết nền nhằm hình thành một cơ sở lý luận vững chắc, xác định phạm vi của khái niệm và mối quan hệ giữa chúng, đồng thời bình luận những đóng góp

của lý thuyết (kết quả nghiên cứu) trong lý thuyết khoa học nhằm gia tăng giá trị của nghiên cứu.

TRIẾT LÝ NGHIÊN CỨU

Theo Saunders và cộng sự (2007), để thực hiện việc nghiên cứu khoa học, cần xem xét triết lý nghiên cứu trên cơ sở xem xét cách thức tư duy về nhận thức học, bản thể học và thuyết giá trị.

Nhận thức luận (epistemology)

Nhận thức luận liên quan đến điều gì tạo ra kiến thức để có thể chấp nhận trong lĩnh vực nghiên cứu, nó giải thích nguồn gốc, bản chất của lý thuyết khoa học. Lý thuyết này được giải thích bởi chủ nghĩa thực chứng (positivism), chủ nghĩa hiện thực (realism) hoặc diễn giải luận (interpretivism).

Chủ nghĩa thực chứng (positivism) dựa trên nguyên tắc sử dụng lý thuyết hiện tại để phát triển các giả thuyết, những giả thuyết này sẽ được kiểm định và xác nhận, toàn phần hoặc từng phần hoặc phản bác. Điều này dẫn đến việc phát triển tiếp theo sau và có thể được kiểm định bởi các lý thuyết tiếp theo.

Chủ nghĩa hiện thực (realism) là điều mà giác quan thể hiện cho chúng ta như thực thể chính là chân lý. Lý thuyết về chủ nghĩa hiện thực là có một thực thể hoàn toàn độc lập với tâm trí con người, nó đối nghịch với chủ nghĩa lý tưởng khi lý thuyết quan niệm chỉ có trong tâm trí. Nó cho phép nhà nghiên cứu thu thập dữ liệu để tìm hiểu về những dữ liệu này. Theo quan điểm này, những điều mà chúng ta trải nghiệm, những điều mà chúng ta thấy đúng là những hình ảnh của sự vật trong thế giới thực.

Diễn giải luận (interpretivism) là nhận thức ủng hộ quan điểm nhà nghiên cứu cần hiểu sự khác nhau giữa những con người trong vai trò các tác nhân xã hội và điều này cho phép khẳng định việc nghiên cứu con người khác với việc nghiên cứu các sự vật.

Bản thể học (ontology)

Bản thể học liên quan đến bản chất của hiện thực, hai khía cạnh của bản thể học liên quan đến chủ nghĩa khách quan (objectivism) và chủ nghĩa chủ quan (subjectivism). Chủ nghĩa khách quan (objectivism) mô tả quan điểm các thực thể xã hội tồn tại trong thực tại bên ngoài các tác nhân xã hội, liên quan tới sự tồn tại của thực thể. Chủ nghĩa chủ quan (subjectivism) quan niệm các hiện tượng xã hội được tạo ra từ các nhận thức và các hoạt động tiếp theo của những tác nhân xã hội liên quan với sự tồn tại của các thực thể.

Trong khi đó, chủ nghĩa thực dụng (pragmatism) cho rằng yếu tố quan trọng nhất của triết lý nghiên cứu là chọn câu hỏi nghiên cứu để trả lời những câu hỏi cụ thể và nếu câu hỏi nghiên cứu không rõ ràng thì nên chọn chấp nhận triết lý thực chứng hoặc diễn giải.

Thuyết giá trị (axiology)

Thuyết giá trị nghiên cứu những phán đoán về giá trị và đó là quá trình tìm hiểu xã hội mà chúng ta quan tâm. Thuyết giá trị cho rằng giá trị của con người là lý lẽ hướng dẫn toàn bộ hoạt

động của con người và trong suốt quá trình nghiên cứu, nhà nghiên cứu cần phải chứng minh giá trị hay độ tin cậy của nghiên cứu.

Ngoài ra, thuyết giá trị còn nhấn mạnh liên quan đến tính trung thực của bản thân, nó nhấn mạnh việc nhà nghiên cứu cần phải thực hiện trung thực nghiên cứu (đạo đức trong nghiên cứu) và diễn giải kết quả nghiên cứu theo đúng với giá trị thực tại.

CÁCH TIẾP CẬN VÀ TRƯỜNG PHÁI NGHIÊN CỨU

Tiếp cận nghiên cứu

Theo Kothari (2004), có hai phương pháp tiếp cận nghiên cứu chính đó là phương pháp tiếp cận định tính và định lượng:

Cách tiếp cận định tính

Cách tiếp cận định tính liên quan đến các nhận định chủ quan về thái độ, ý kiến và hành vi. Nghiên cứu theo cách tiếp cận này giúp cho nhà nghiên cứu hiểu sâu hơn về vấn đề cần nghiên cứu. Theo Thọ (2011), nghiên cứu định tính thường được dùng để xây dựng lý thuyết khoa học. Các công cụ được sử dụng trong nghiên cứu định tính là phỏng vấn nhóm (focus groups), phỏng vấn sâu (in-depth interview) và kỹ thuật ánh xạ (projective technique).

Cách tiếp cận định lượng

Liên quan đến việc xử lý, rút trích ra các dữ liệu ở dạng số lượng/ định lượng trên dữ liệu bằng các công cụ phân tích định lượng. Cách tiếp cận này có thể được phân ra ba loại: tiếp cận suy luận (inferential), cách tiếp cận thực nghiệm (experimental) và cách tiếp cận mô phỏng (simulation). Mục tiêu của cách tiếp cận suy luận là hình thành nên cơ sở dữ liệu từ đó có thể suy ra đặc điểm hoặc mối quan hệ trên tổng thể. Điều này có nghĩa là với cách tiếp cận này, chúng ta sử dụng nghiên cứu điều tra (survey research) trên mẫu đại diện cho tổng thể và những kết luận về đặc điểm, mối quan hệ trên mẫu có thể được suy ra cho tổng thể mà mẫu đại diện. Cách tiếp cận thực nghiệm được phân loại dựa vào mức độ kiểm soát lớn đối với môi trường nghiên cứu. Với cách tiếp cận thực nghiệm, một vài biến số nghiên cứu được thay đổi, điều chỉnh để quan sát tác động của những biến này trên các biến số nghiên cứu khác. Cách tiếp cận mô phỏng liên quan đến việc hình thành nên một môi trường nhân tạo trong đó những thông tin và dữ liệu được tạo ra.

Trường phái nghiên cứu

Diễn giải (deduction): Kiểm định lý thuyết

Cách tiếp cận diễn giải bao gồm việc sử dụng các tài liệu, dữ liệu để phát triển lý thuyết và các giả thuyết, đồng thời thiết kế chiến lược nghiên cứu để kiểm định các giả thuyết. Trong phương pháp tư duy diễn giải, kết luận được rút ra từ các lí do/ cơ sở được chứng minh là đúng và có giá trị (đi từ cái chung tới cái riêng). Nếu một trong các lí do, cơ sở này không có giá trị thì kết luận rút ra không có nghĩa. Mối quan hệ giữa lí do/ chứng cứ và kết luận trong phương pháp tư duy diễn giải mạnh hơn trong phương pháp tư duy qui nạp.

Phương pháp tư duy diễn giải được áp dụng trong qui trình kiểm định lý thuyết khoa học (Thọ, 2011). Khi gặp một vấn đề cần nghiên cứu, người nghiên cứu tìm kiếm lí thuyết phù hợp để giải quyết vấn đề. Những lí thuyết có sẵn phù hợp với vấn đề nghiên cứu sẽ được tổng hợp để xây

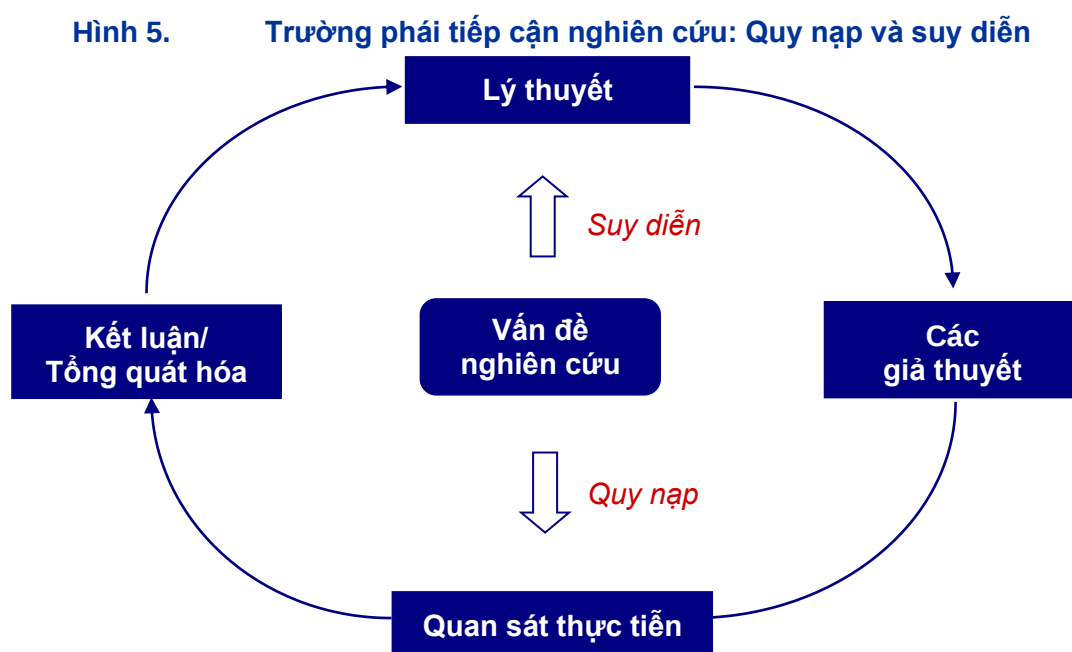
dựng mô hình, giả thuyết nghiên cứu, phát triển các thang đo cho khái niệm nghiên cứu. Sau đó, nhà nghiên cứu sẽ dựa trên mô hình này để thiết kế nghiên cứu, thu thập dữ liệu. Những dữ liệu thu thập được sẽ dùng để kiểm định mô hình lý thuyết được xây dựng. Kết quả của nghiên cứu theo phương pháp tư duy diễn giải là việc chấp nhận hay bác bỏ giả thuyết, mô hình nghiên cứu.

Phương pháp diễn dịch có một số đặc điểm sau:

- Tìm kiếm để giải thích quan hệ nhân quả giữa các biến: Sau thời gian nghiên cứu, nhà nghiên cứu phát hiện ra rằng những nhân viên có thâm niên làm việc trong công ty có lương trung bình khá cao trong khi đó những nhân viên trẻ, mới vào làm việc có lương khá thấp. Do vậy, có thể phát triển giả thuyết tồn tại mối quan hệ giữa thâm niên công tác và lương của nhân viên hay nhân viên có thâm niên càng nhiều thì lương càng cao.
- Thường sử dụng dữ liệu định lượng để kiểm định, cần chú ý rằng điều này không có nghĩa là cách tiếp cận diễn dịch không sử dụng dữ liệu định tính.
- Thực tiễn hóa hay diễn đạt giả thuyết theo thuật ngữ thực tiễn.
- Tổng quát hóa nghĩa là kết quả nghiên cứu có thể suy rộng lên tổng thể hoặc một quy mô lớn hơn.

Quy nạp (induction): Xây dựng lý thuyết

Theo (Cooper & Schindler, 2011): “Theo trường phái quy nạp, kết luận được rút ra từ một hoặc nhiều sự kiện chứng cứ từ thực tế (đi từ cái riêng tới cái chung). Kết luận giải thích sự kiện và sự kiện góp phần tạo ra kết luận, một kết luận có thể chỉ là một giải thích cho các sự kiện và có thể có nhiều kết luận cho các sự kiện xảy ra hay nói cách khác, hay nói một cách cụ thể, kết quả của quy nạp là các giả thuyết”.



Phương pháp quy nạp là phương pháp mà người nghiên cứu phải thu thập, khai thác dữ liệu và phát triển những lý thuyết từ chúng và sau đó liên hệ với lý thuyết nghiên cứu hay nói khác hơn: lý thuyết đến từ dữ liệu. Nghiên cứu theo cách tiếp cận quy nạp đặc biệt liên quan với bối

cảnh xảy ra các sự kiện, vì vậy, nghiên cứu một mẫu nhỏ các chủ thể sẽ thích hợp hơn mẫu lớn như cách tiếp cận diễn giải.

Phương pháp tư duy qui nạp được áp dụng để xây dựng lý thuyết khoa học (Thọ, 2011), đối với một hiện tượng cụ thể nhưng chưa có lý thuyết giải thích, người nghiên cứu lấy nó làm xuất phát điểm để xây dựng một lý thuyết mới để giải thích hiện tượng đó. Sau khi quyết định xây dựng lý thuyết khoa học, người nghiên cứu tiến hành nghiên cứu lý thuyết nền, phương pháp tình huống, quan sát, thảo luận nhóm, phỏng vấn sâu (những phương pháp này sẽ được đề cập sâu hơn trong phần nghiên cứu định tính) để thu thập dữ liệu từ thực tế. Sau đó, dựa vào những dữ liệu thu thập cùng với các lý thuyết hiện có, nhà nghiên cứu xây dựng các giả thuyết và mô hình; và đây cũng chính là kết quả của nghiên cứu theo phương pháp tư duy qui nạp.

Như vậy, sơ lược quy trình nghiên cứu qui nạp và suy diễn (định tính và định lượng) có thể tổng hợp tại hình sau:



Hỗn hợp

Phương pháp hỗn hợp được sử dụng để xây dựng và kiểm định lý thuyết khoa học. Phương pháp này được sử dụng khi vấn đề nghiên cứu có tính mới, lý thuyết hiện có chưa giải quyết được, người nghiên cứu vừa có mục tiêu phải xây dựng lý thuyết mới và vừa có mục tiêu kiểm định lý thuyết được xây dựng bằng các dữ liệu thực tế.

Cách tiếp cận theo các phương pháp hỗn hợp là cách tiếp cận trong đó nhà nghiên cứu có khuynh hướng đặt những lời khẳng định tri thức trên những cơ sở thực dụng. Cách tiếp cận này sử dụng các chiến lược điều tra bao gồm việc thu thập dữ liệu hoặc là đồng thời hoặc là theo trình tự để hiểu được tốt nhất vấn đề được nghiên cứu. Việc thu thập dữ liệu cũng đòi hỏi tập hợp cả thông tin bằng số (ví dụ, trên các công cụ) cũng như thông tin bằng văn bản (ví dụ, trong các cuộc phỏng vấn) thế nào để có cơ sở dữ liệu cuối cùng thể hiện cả thông tin định lượng lẫn thông tin định tính.

PHÂN LOẠI NGHIÊN CỨU

Theo (Kothari, 2004), có thể phân loại nghiên cứu thành:

- Nghiên cứu ứng dụng (applied research) và nghiên cứu cơ bản (fundamental research): Nghiên cứu ứng dụng nhằm tìm ra giải pháp cho vấn đề, cơ hội mà một tổ chức, doanh nghiệp đang đối diện, trong khi đó nghiên cứu cơ bản nhằm hình thành nên một lý thuyết

hoặc mở rộng, phát triển kiến thức. Các nghiên cứu quan tâm đến các hành vi của con người và có mục tiêu tổng quát hóa các hành vi đó là ví dụ của nghiên cứu cơ bản. Các nghiên cứu để xác định các xu hướng về kinh tế, xã hội và chính trị cụ thể có thể ảnh hưởng đến một tổ chức, doanh nghiệp nhất định là ví dụ của nghiên cứu ứng dụng. Vì vậy, có thể cho rằng nghiên cứu ứng dụng là nghiên cứu hướng đến việc tìm thông tin, giải pháp cho một vấn đề, cơ hội cụ thể còn nghiên cứu cơ bản hướng đến tìm những thông tin, kiến thức bổ sung vào các kiến thức đã có.

- Nghiên cứu định lượng (quantitative) và định tính (qualitative): Nghiên cứu định lượng dựa vào việc đo lường số lượng. Nghiên cứu định lượng được áp dụng đối với các hiện tượng có thể được diễn tả bằng số lượng. Nghiên cứu định tính quan tâm đến tính chất của hiện tượng. Ví dụ, nếu chúng ta quan tâm đến việc điều tra những lí do cho những hành vi nhất định của người tiêu dùng (tại sao người tiêu dùng nghĩ hoặc làm những việc cụ thể nào đó), chúng ta đang nói đến nghiên cứu về động cơ (motivation research), một loại nghiên cứu quan trọng trong nghiên cứu định tính. Loại nghiên cứu này nhằm khám phá các động cơ, mong muốn đằng sau hành vi cụ thể qua việc sử dụng phỏng vấn sâu hay phỏng vấn sâu (in-depth Interview).
- Nghiên cứu về mặt nhận thức (conceptual) và nghiên cứu kinh nghiệm (empirical): Nghiên cứu nhận thức liên quan đến những ý tưởng trừu tượng hoặc lý thuyết, loại nghiên cứu này được sử dụng bởi những nhà nghiên cứu mong muốn phát triển những khái niệm mới hoặc diễn giải lại những khái niệm đã có. Nghiên cứu kinh nghiệm chủ yếu dựa vào kinh nghiệm và quan sát mà ít quan tâm đến hệ thống lý thuyết. Đây là nghiên cứu dựa vào dữ liệu, kết quả nghiên cứu được giải thích bởi quan sát hoặc thực nghiệm.

DỮ LIỆU TRONG NGHIÊN CỨU

Nguồn dữ liệu thứ cấp

Dữ liệu thứ cấp là loại dữ liệu được thu thập sẵn, đã công bố nên dễ thu thập, ít tốn thời gian, tiền bạc trong quá trình thu thập.

Các loại dữ liệu thứ cấp

Dữ liệu thứ cấp có đặc điểm là chỉ cung cấp các thông tin mô tả tình hình, chỉ rõ qui mô của hiện tượng chứ chưa thể hiện được bản chất hoặc các mối liên hệ bên trong của hiện tượng nghiên cứu. Vì dữ liệu thứ cấp, dù thu thập từ bên trong hoặc bên ngoài doanh nghiệp, nó cũng là những thông tin đã được công bố nên thiếu tính cập nhật, đôi khi thiếu chính xác và không đầy đủ. Tuy nhiên, dữ liệu thứ cấp cũng đóng một vai trò quan trọng trong nghiên cứu do các lý do:

- Các dữ liệu thứ cấp có thể giúp người quyết định đưa ra giải pháp để giải quyết vấn đề trong những trường hợp thực hiện những nghiên cứu mà các dữ liệu thứ cấp là phù hợp mà không cần thiết phải có các dữ liệu sơ cấp, ví dụ như các nghiên cứu thăm dò hoặc nghiên cứu mô tả.
- Ngay cả khi dữ liệu thứ cấp không giúp ích cho việc ra quyết định thì nó vẫn rất quan trọng vì nó giúp xác định và hình thành các giả thiết về các giải pháp cho vấn đề. Nó là cơ sở để

hoạch định việc thu thập các dữ liệu sơ cấp; cũng như được sử dụng để xác định tổng thể chọn mẫu và thực hiện chọn mẫu để thu thập dữ liệu sơ cấp.

Dữ liệu thứ cấp bên trong

Khi tìm kiếm dữ liệu thứ cấp nên bắt đầu từ các nguồn bên trong tổ chức. Hầu hết các tổ chức đều có những nguồn thông tin rất phong phú, vì vậy có những dữ liệu có thể sử dụng ngay lập tức. Chẳng hạn như dữ liệu về doanh thu bán hàng và chi phí bán hàng hay các chi phí khác sẽ được cung cấp đầy đủ thông qua các bảng báo cáo thu nhập của doanh nghiệp. Những thông tin khác có thể tìm kiếm lâu hơn nhưng thật sự không khó khăn khi thu thập loại dữ liệu này. Có hai thuận lợi chính khi sử dụng dữ liệu thứ cấp bên trong doanh nghiệp là thu thập được một cách dễ dàng và có thể không tốn kém chi phí. Để tạo ra cơ sở dữ liệu thứ cấp bên trong, doanh nghiệp cần tổ chức cơ sở dữ liệu. Đó là việc sử dụng máy tính để nắm bắt và theo dõi các hồ sơ doanh nghiệp, khách hàng.... Thông tin thứ cấp này phục vụ như một nền tảng cho các chương trình nghiên cứu hoặc như là nguồn thông tin nội bộ liên quan đến hành vi khách hàng trong nhiều doanh nghiệp.

Dữ liệu thứ cấp bên ngoài

Những nguồn dữ liệu thứ cấp bên ngoài là các tài liệu đã được xuất bản có được từ các nghiệp đoàn, chính phủ, chính quyền địa phương, các tổ chức phi chính phủ (NGO), các hiệp hội thương mại, các tổ chức chuyên môn, các ấn phẩm thương mại, các tổ chức nghiên cứu chuyên nghiệp... Sự phát triển của mạng thông tin toàn cầu đã tạo nên một nguồn dữ liệu vô cùng phong phú và đa dạng, đó là các dữ liệu thu thập từ internet. Trong thực tế, có rất nhiều dữ liệu thứ cấp có thể sử dụng được và có thể tìm kiếm từ nhiều nguồn khác nhau, điều quan trọng là phải phân loại nguồn dữ liệu để có một phương thức tìm kiếm thích hợp.

Các tiêu chuẩn đánh giá dữ liệu thứ cấp

Tính cụ thể

Dữ liệu thứ cấp phải đảm bảo tính cụ thể, có nghĩa nó phải rõ ràng, phù hợp mục tiêu nghiên cứu, có thể hỗ trợ cho việc phân tích nhận diện vấn đề hay mô tả vấn đề nghiên cứu. Tính cụ thể còn đòi hỏi sự rõ ràng về nguồn dữ liệu thu thập cũng như hiệu quả của dữ liệu (so sánh lợi ích của dữ liệu với chi phí thu thập).

Tính chính xác của dữ liệu

Người nghiên cứu phải xác định dữ liệu có đủ chính xác phục vụ cho mục đích nghiên cứu hay không, dữ liệu thứ cấp có thể có sai số (hay không chính xác), điều này phụ thuộc vào nguồn cung cấp dữ liệu. Vì vậy, uy tín của nhà cung cấp và độ tin cậy của nguồn dữ liệu những tiêu chuẩn cần xem xét khi thu thập dữ liệu thứ cấp.

Tính thời sự

Do một phần kết quả nghiên cứu sẽ cung cấp thông tin cho việc ra quyết định, làm cơ sở cho việc thực hiện các nghiên cứu tiếp theo nên nghiên cứu đòi hỏi dữ liệu có tính chất thời sự (dữ liệu mới) vì giá trị của dữ liệu sẽ bị giảm qua thời gian. Đó cũng là lý do vì sao các nghiên cứu luôn cập nhật thông tin định kỳ, tạo ra nguồn thông tin có giá trị cao.

Mục đích của dữ liệu được thu thập

Dữ liệu thu thập nhằm đáp ứng cho một số mục tiêu nghiên cứu đã xác định và giải đáp câu hỏi "Dữ liệu cần được thu thập để làm gì?". Dữ liệu được thu thập phục vụ cho một mục tiêu nghiên cứu cụ thể, vì vậy các dữ liệu có thể phù hợp cho mục tiêu nghiên cứu này nhưng lại có thể không phù hợp trong trường hợp khác.

Nguồn dữ liệu sơ cấp

Các dữ liệu sơ cấp được thu thập trực tiếp từ đối tượng nghiên cứu, có thể là người tiêu dùng, nhóm người tiêu dùng... Nó còn được gọi là các dữ liệu gốc, chưa được xử lý. Vì vậy, các dữ liệu sơ cấp giúp người nghiên cứu đi sâu vào đối tượng nghiên cứu, tìm hiểu động cơ của khách hàng, phát hiện các quan hệ trong đối tượng nghiên cứu. Dữ liệu sơ cấp được thu thập trực tiếp nên độ chính xác khá cao, đảm bảo tính cập nhật nhưng lại mất thời gian và tốn kém chi phí để thu thập. Dữ liệu sơ cấp có thể thu thập từ việc quan sát, ghi chép hoặc tiếp xúc trực tiếp với đối tượng điều tra; cũng có thể sử dụng các phương pháp thử nghiệm để thu thập dữ liệu sơ cấp.

TIẾN TRÌNH NGHIÊN CỨU

Theo (Kothari, 2004), tiến trình nghiên cứu gồm những giai đoạn sau:

- (1) Xác định vấn đề nghiên cứu (research problem);
- (2) Tổng quan/ tổng kết lý thuyết (review literature);
 - Tổng quan/ hợp tuyển về khái niệm, khái niệm nghiên cứu và lý thuyết liên quan;
 - Tổng quan/ hợp tuyển về phương pháp, kết quả của các nghiên cứu có trước liên quan đến vấn đề nghiên cứu.
- (3) Phát triển giả thuyết;
- (4) Thiết kế nghiên cứu (bao gồm cả thiết kế chọn mẫu);
- (5) Thu thập và chuẩn bị dữ liệu;
- (6) Phân tích dữ liệu và diễn giải;
- (7) Báo cáo kết quả.

Theo (Kothari, 2004), tiến trình nghiên cứu gồm các hoạt động ảnh hưởng mật thiết lên nhau chứ không phải theo một qui trình cứng nhắc. Mặc dù tiến trình nghiên cứu nêu trên là một hướng dẫn về qui trình hữu ích, nhưng không phải nhất nhất phải làm xong giai đoạn này rồi mới bắt đầu làm bước kế tiếp mà ở những giai đoạn trước, người nghiên cứu đã phải dự đoán và đặt ra yêu cầu cho những giai đoạn tiếp theo. Việc thực hiện các bước cần phải thận trọng, tránh xảy ra tình trạng khi đã hoàn thành các bước sau mới nhận thấy những sai sót ở bước trước.

Giai đoạn 1: Xác định vấn đề nghiên cứu/ mục tiêu nghiên cứu

Vấn đề nghiên cứu xuất phát từ những vấn đề gặp phải trong thực tế mà chưa tìm ra lời giải thích thích hợp từ các lý thuyết đã có hoặc những "lỗ hổng" (gap) trong các lý thuyết hiện tại.

Người nghiên cứu phải làm quen với các lý thuyết, tài liệu liên quan đến vấn đề nghiên cứu được lựa chọn, từ đó đối chiếu với những vấn đề mang tính chất lý thuyết để tìm ra hướng nghiên cứu. Giai đoạn này sẽ được giới thiệu chi tiết tại chương 2.

Giai đoạn 2: Tổng quan/ hợp tuyển lý thuyết

Tổng quan/ hợp tuyển lý thuyết là cơ sở để xác định các vấn đề nghiên cứu dựa trên các “lỗ hổng” nghiên cứu về các lý thuyết khoa học và phương pháp. Việc tổng hợp lý thuyết sẽ giúp cho nhà nghiên cứu xác định được vấn đề nghiên cứu (nhằm đóng góp cho lý thuyết khoa học các vấn đề về các lý thuyết và phương pháp luận) và làm cơ sở cho việc bình luận kết quả sau khi có kết quả phân tích.

Giai đoạn 3: Phát triển giả thuyết

Sau khi thực hiện tổng quan về tài liệu, người nghiên cứu phát triển các giả thuyết nghiên cứu có thể kiểm chứng được (working hypothesis/ hypotheses). Giả thuyết được phát biểu để được kiểm chứng tính logic và thực chứng, vì thế việc phát biểu giả thuyết nghiên cứu đặc biệt quan trọng khi nó là trọng tâm của vấn đề nghiên cứu. Giả thuyết nghiên cứu ảnh hưởng đến cách mà các kiểm định được thực hiện trong phân tích dữ liệu và yêu cầu về đặc điểm và chất lượng dữ liệu dùng để phân tích. Giả thuyết nghiên cứu cần phải cụ thể, rõ ràng và liên quan trực tiếp đến vấn đề nghiên cứu. Vai trò của giả thuyết nghiên cứu là hạn định/giới hạn lại phạm vi nghiên cứu và hướng dẫn người nghiên cứu đi đúng đường. Giả thuyết còn giúp xác định loại dữ liệu và các loại phương pháp sử dụng để phân tích dữ liệu.

Giai đoạn 4: Thiết kế nghiên cứu

Xác định cách tiếp cận/ phương pháp nghiên cứu

Khi vấn đề nghiên cứu đã được xác định rõ ràng, thiết kế nghiên cứu chỉ rõ những bước cần thực hiện để thu thập được những dữ liệu cần thiết với sự tính toán về nguồn lực. Việc thiết kế nghiên cứu như thế nào chủ yếu phụ thuộc vào mục đích nghiên cứu. Mục đích nghiên cứu có thể chia thành 4 nhóm sau: (1) Khám phá/ thăm dò (exploratory), (2) Mô tả (description), (3) Dự đoán (diagnosis) và (4) Thực nghiệm (experimentation). Nếu mục đích nghiên cứu là thăm dò, một thiết kế nghiên cứu linh hoạt cho phép người nghiên cứu xem xét nhiều khía cạnh của vấn đề là phù hợp. Nhưng nếu mục đích nghiên cứu là phải cung cấp sự mô tả chính xác về một tình huống hoặc mối quan hệ giữa các biến số, thiết kế nghiên cứu phù hợp phải tối thiểu hóa những sai lệch (bias) và tối đa hóa độ tin cậy của dữ liệu được thu thập và phân tích. Ngoài ra, còn có những thiết kế nghiên cứu khác như thực nghiệm để kiểm định giả thuyết và kiểm định giả thuyết không thực nghiệm. Thiết kế thực nghiệm có thể là thiết kế không chính thức (chẳng hạn như trước và sau mà không kiểm soát, chỉ kiểm soát sau đó, kiểm soát trước và sau) hoặc là thiết kế chính thức (như thiết kế hoàn toàn ngẫu nhiên, thiết kế mô hình thừa số đơn giản hoặc phức tạp). Người nghiên cứu phải chọn cho mình một thiết kế nghiên cứu phù hợp. Để chuẩn bị một thiết kế phù hợp với mục đích nghiên cứu, cần phải xem xét những điều sau:

- Phương pháp thu thập thông tin;
- Sự sẵn sàng và kỹ năng của những người tham gia nghiên cứu/những người nghiên cứu;
- Giải thích cho việc lựa chọn phương pháp thu thập thông tin;

- Thời gian dành cho nghiên cứu;
- Nguồn tài chính việc nghiên cứu.

Xác định thiết kế chọn mẫu

Nhà nghiên cứu muốn rằng tất cả những thông tin thu thập được có thể áp dụng trên tổng thể mục tiêu. Tuy nhiên, với sự hạn chế về nguồn lực, một mẫu đại diện cho tổng thể mục tiêu phải được lựa chọn thay vì nghiên cứu trên cả tổng thể. Người nghiên cứu phải xác định được cách thức chọn mẫu hay xác định được thiết kế chọn mẫu phù hợp. Mẫu có thể là mẫu xác suất hoặc mẫu không xác suất. Với mẫu xác suất, người nghiên cứu biết được xác suất được đưa vào mẫu của mỗi đối tượng nghiên cứu. Trong khi đó, người nghiên cứu không biết được xác suất này với mẫu không xác suất. Mẫu xác suất dựa trên chọn mẫu ngẫu nhiên đơn giản, chọn mẫu hệ thống, chọn mẫu ngẫu nhiên phân tầng, chọn mẫu theo cụm, chọn mẫu nhiều giai đoạn. Trong khi đó, mẫu không xác suất dựa trên chọn mẫu thuận tiện, chọn mẫu phán đoán và chọn mẫu kiểm tra tỉ lệ.

Xác định phương pháp thu thập dữ liệu

Trong thực tế, những thông tin có sẵn có thể không phù hợp hoặc không có giá trị, vì thế việc thu thập dữ liệu sơ cấp là cần thiết. Tùy thuộc vào nguồn lực và mục tiêu nghiên cứu mà nhà nghiên cứu chọn một phương pháp thu thập dữ liệu phù hợp. Dữ liệu sơ cấp có thể được thông qua phỏng vấn, quan sát, điều tra hoặc thực nghiệm. Đối với phương pháp điều tra, có nhiều cách thức được sử dụng như: phỏng vấn trực tiếp cá nhân bằng bản câu hỏi có cấu trúc (personal interviews), phỏng vấn qua điện thoại (telephone interviews), gửi bản câu hỏi qua bưu điện hoặc gửi trực tuyến đến e-mail. Người nghiên cứu cần xác định phương pháp lựa chọn thu thập dữ liệu phù hợp dựa vào bản chất của nghiên cứu, qui mô nghiên cứu, nguồn lực về tài chính, thời gian và mức độ chính xác mong muốn trong dữ liệu thu thập.

Xác định công cụ phân tích dữ liệu

Trong thiết kế nghiên cứu, người nghiên cứu cũng lên kế hoạch sử dụng các công cụ phân tích dữ liệu nào để phục vụ mục tiêu nghiên cứu của mình. Công cụ phân tích được sử dụng dựa vào câu hỏi nghiên cứu, loại dữ liệu, các giả định về phân phối của dữ liệu,... Có những công cụ phân tích dành cho nghiên cứu định tính và cũng có những công cụ phân tích đặc thù dùng cho nghiên cứu định lượng.

Giai đoạn 5: Tổ chức thu thập

Thu thập dữ liệu cũng là giai đoạn quan trọng trong tiến trình nghiên cứu, thu thập dữ liệu đúng cách quyết định độ tin cậy và phù hợp của dữ liệu. Có nhiều cách để thu thập dữ liệu như quan sát, phỏng vấn (trực tiếp, qua thư, qua điện thoại, qua Internet...). Mỗi phương pháp có những đặc điểm và ưu, nhược điểm khác nhau. Nếu phương pháp điều tra với công cụ thu thập dữ liệu được thiết kế (ví dụ, bản câu hỏi có cấu trúc), dữ liệu có thể được xử lý bằng phần mềm. Trong trường hợp này, câu hỏi và các câu trả lời cụ thể được mã hóa. Nếu dữ liệu được thu thập thông qua phỏng vấn cá nhân trực tiếp, việc sắp xếp cuộc phỏng vấn và đào tạo cho những người phỏng vấn phải được xem xét. Việc kiểm tra hiện trường (field checks) là cần thiết để đảm bảo rằng người phỏng vấn thực hiện công việc của họ trung thực và hữu hiệu.

Giai đoạn 6: Chuẩn bị dữ liệu và phân tích dữ liệu

Công việc của bước này bao gồm việc (1) chuẩn bị dữ liệu, (2) mã hóa dữ liệu, (3) kiểm tra và hiệu chỉnh dữ liệu (nếu cần thiết), (4) nhập liệu và (5) xử lý và phân tích dữ liệu để đáp ứng mục tiêu nghiên cứu. Việc phân tích dữ liệu đòi hỏi một những bước liên quan mật thiết lẫn nhau như phân biệt các loại cấp độ thang đo, sử dụng phân tích mô tả hay phân tích mối quan hệ... tùy theo dữ liệu được sử dụng.

Giai đoạn 7: Báo cáo kết quả

Sau khi phân tích dữ liệu cần tiến hành viết báo cáo và trình bày các vấn đề đã được nghiên cứu, những kết luận được trình bày cô đọng và logic theo các chuẩn đã được quy định. Đặc biệt, việc trình bày tài liệu tham khảo cần được phải được trình bày đầy đủ trong nội dung của bài viết và tài liệu tham khảo. Lý do của việc ghi nguồn tham khảo là:

- Ghi một lời cảm ơn, tri ân đến các tác giả khi nhà nghiên cứu đã kể đến hay trích dẫn công việc của họ;
- Tăng hàm lượng khoa học của nghiên cứu: một công trình khoa học tham khảo nhiều tài liệu, các tài liệu tham khảo có hàm lượng khoa học cao sẽ gia tăng hàm lượng khoa học trong báo cáo của nhà nghiên cứu;
- Thông tin tài liệu tham khảo phải đầy đủ để người đọc (nếu muốn) có thể truy xuất được tài liệu.

Mỗi tạp chí, viện nghiên cứu, trường đại học có một số quy định khác nhau về cách thức trình bày một báo cáo và tài liệu tham khảo. Tuy nhiên, dù quy định thế nào, trong phần báo cáo kết quả, việc trình bày báo cáo và tài liệu tham khảo cần phải theo một chuẩn nhất định. Trong quá trình viết báo cáo, các nhà khoa học thường sử dụng phần mềm Endnote để quản lý và trình bày tài liệu tham khảo, một số tạp chí sử dụng cách trình bày APA 6th, Vancouver,... để trình bày. Tuy nhiên, ở Việt Nam do có những đặc thù nhất định nên thông thường, các tạp chí, viện nghiên cứu thường có những quy định riêng. Sau đây là cách ghi trích dẫn và trình bày tài liệu tham khảo khá phổ biến của một số báo cáo khoa học:

1. Trích dẫn trực tiếp:

- Nếu có một người là tác giả: Tên tác giả và năm xuất bản (năm xuất bản để trong ngoặc).
Ví dụ: Chất lượng dịch vụ, theo Parasuramane (1985), “là sự cảm nhận.....”
- Nếu đồng tác giả (2 người): Tên 2 tác giả và năm xuất bản (năm xuất bản để trong ngoặc).
Ví dụ: Huy và Hoàng (2012) cho rằng: “Việt Nam cần phải xây dựng chiến lược phát triển kinh tế Miền Trung và Tây Nguyên”
- Nếu đồng tác giả (3 người):
 - + Trích dẫn lần đầu: Tên tất cả các tác giả (năm xuất bản để trong ngoặc)
 - + Trích dẫn lần thứ 2 trở đi: Tên tác giả đầu tiên “và cộng sự” năm xuất bản (năm xuất bản để trong ngoặc).

Ví dụ: Huy, Hà, Hiền và Hồng (2000) cho rằng: “Việt Nam cần phải xây dựng chiến lược phát triển kinh tế Miền Trung và Tây Nguyên” cho trích dẫn lần đầu và Huy và cộng sự (2000) cho rằng: “Việt Nam cần phải xây dựng chiến lược phát triển kinh tế Miền Trung và Tây Nguyên” cho trích dẫn các lần sau.

2. Trích dẫn gián tiếp:

- Có một công trình:

- Nếu có 1 tác giả: (Tên tác giả, năm xuất bản)
- Nếu có 2 tác giả: (Tên tác giả 1 và tên tác giả 2, năm xuất bản)
- Nếu có nhiều tác giả:
 - + Lần đầu: (Ghi hết tên tác giả, năm xuất bản);
 - + Lần thứ (Tên tác giả 1 và cộng sự, năm xuất bản)

Ví dụ: Công nghiệp Việt Nam đóng vai trò quan trọng trong nền kinh tế Việt Nam (Kenichi Ohno, 2005) hoặc (Kenichi Ohno và cộng sự, 2005).

- Khi có nhiều tác giả thì bạn phải sắp xếp theo thứ tự theo năm tăng dần, sau năm là dấu chấm phẩy (;).

Ví dụ: (Handy, 1979; Johnson, 1992; Brown, 1999)

3. Trích dẫn của các trích dẫn:

- Khi bạn trích dẫn ý tưởng của một tác giả, mà ý tưởng này được trích dẫn từ một tác giả khác thì bạn phải ghi là: (Parasuraman, 1987, trích trong Huy, 2008)
- Chúng ta phải trích dẫn như vậy bởi vì chúng ta không đọc bản gốc của Parasuraman mà đọc của Huy (2008).

Cách trình bày tài liệu tham khảo:

1. Tài liệu tham khảo được xếp riêng theo từng ngôn ngữ (Việt, Anh, Pháp, Đức, Nga, Trung, Nhật...). Các tài liệu bằng tiếng nước ngoài phải giữ nguyên văn, không phiên âm, không dịch, kể cả tài liệu bằng tiếng Trung Quốc, Nhật... (đối với các tài liệu bằng ngôn ngữ còn ít người biết có thể thêm phần dịch tiếng Việt đi kèm theo mỗi tài liệu).
2. Tài liệu tham khảo xếp theo thứ tự ABC họ tên tác giả luận án theo thông lệ của từng nước:
 - Tác giả là người nước ngoài: Xếp thứ tự ABC theo họ;
 - Tác giả là người Việt Nam: Xếp thứ tự ABC theo tên nhưng vẫn giữ nguyên thứ tự thông thường của tên người Việt Nam, không đảo tên lên trước họ;
 - Tài liệu không có tên tác giả thì xếp theo thứ tự ABC từ đầu của tên cơ quan ban hành báo cáo hay ấn phẩm, Ví dụ: Tổng cục Thống kê xếp vào vần T, Bộ Giáo dục và Đào tạo xếp vào vần B.
3. Tài liệu tham khảo là sách, luận án, báo cáo phải ghi đầy đủ các thông tin sau:
 - Tên các tác giả hoặc cơ quan ban hành (không có dấu ngăn cách)

- (năm xuất bản), (đặt trong ngoặc đơn, dấu phẩy sau ngoặc đơn)
 - *Tên sách, luận án hoặc báo cáo, (in nghiêng, dấu phẩy cuối tên)*
 - Nhà xuất bản, (dấu phẩy cuối tên nhà xuất bản)
 - Nơi xuất bản. (dấu chấm kết thúc tài liệu tham khảo)
4. Tài liệu tham khảo là bài báo trong tạp chí, bài trong một cuốn sách... ghi đầy đủ các thông tin sau:
- Tên tác giả (không có dấu ngăn cách)
 - (Năm công bố), (đặt trong ngoặc đơn, dấu phẩy sau ngoặc đơn)
 - Tên bài báo, (không in nghiêng, dấu phẩy cuối tên)
 - *Tên tạp chí hoặc tên sách, (in nghiêng, dấu phẩy cuối tên)*
 - Tập (không có dấu ngăn cách)
 - (Số), (đặt trong ngoặc đơn, dấu phẩy sau ngoặc đơn)
 - Các số trang. (gạch ngang giữa hai chữ số, dấu chấm kết thúc)
5. Tài liệu từ các website thì phải ghi rõ địa chỉ truy xuất đến nội dung trích dẫn và ngày đăng xuất nội dung.

Cần chú ý những chi tiết về trình bày nêu trên. Nếu tài liệu dài hơn một dòng thì nên trình bày sao cho từ dòng thứ hai lùi vào so với dòng thứ nhất 1 cm để phần tài liệu tham khảo được rõ ràng và dễ theo dõi. Ví dụ về cách trình bày trang tài liệu tham khảo:

TÀI LIỆU THAM KHẢO

Tiếng Anh

1. Auger, P., BarNir, A. & Gallagher, J. M. (2003), Strategic orientation, competition, and Internet– Based electronic commerce, *Information Technology and Management*, Apr-Jul, 4(2), 139-164.
2. Chieochan, O., Lindley, D. & Dunn, T. (2000), Factors affecting the use information technology in Thai agricultural cooperatives: a work in progress, *The Electronic Journal on Information Systems in Developing Countries*, University of Hong Kong, 2 (1), 1-15.
3. Daniel, E., Wilson, H. & Myers, A. (2002), Adoption of e-commerce by SMEs in the UK, *International Small Business Journal*, London, Aug, 20(3), 253 – 270.
4. ECRC (Electronic Commerce Resource Center) (2002), E-commerce in Thailand: statistical and analysis in each business sector, Bangkok, Thailand: Se-Education.
5. Grover, V. and Goslar, M. D. (1993), The initiation, adoption, and implementation of telecommunications technologies in U. S. organizations, *Journal of Management Information Systems*, Summer, 10(1), 141-164.

6. Ling, C. Y., (2001), Model of factors influences on electronic commerce adoption and diffusion in small & medium sized enterprises, *ECIS Doctoral Consortium*, 24-26 June, AIS region 2 (Europe, Africa, Middle-East).

Tiếng Việt:

7. Lê Văn Hà (2005), Mô hình các nhân tố tác động đến hội nhập thương mại điện tử ở các doanh nghiệp Việt Nam: Một nghiên cứu thực tiễn, *Tạp chí nghiên cứu kinh tế, Viện nghiên cứu kinh tế*, 323, tháng 4, 72-78.
8. Nguyễn Hoàn Thiên và Đinh Thành Trung (2005), Mô hình phát triển thương mại điện tử, *Tạp chí nghiên cứu kinh tế, Viện nghiên cứu kinh tế*, 323, tháng 4, 72-78.
9. Nguyễn Hoàn Thiên, Đinh Thành Trung và Đặng Thành Tâm (2005), Mô hình phát triển thương mại điện tử, *Tạp chí nghiên cứu kinh tế, Viện nghiên cứu kinh tế*, 323, tháng 4, 72-78.

Website

1. Kelly. C, (1997), David Kolb, The Theory of Experiential Learning and ESL” (online) <http://iteslj.org/Articles/Kelly-Experiential/> (ngày truy xuất 28/02/2009).
2. Nguyễn Văn A (2000), “Chất lượng hay không chất lượng” (online) (<http://www.tienphong.vn/Tianyon/Index.aspx>) ngày truy xuất 28/04/2009.

TIÊU CHUẨN CỦA MỘT NGHIÊN CỨU TỐT

Như đã đề cập ở trên, kết quả nghiên cứu là nhằm cung cấp thông tin cho việc ra quyết định, hình thành lý thuyết và cung cấp kiến thức cho xã hội, do vậy, việc thực hiện nghiên cứu cần phải đảm bảo những yêu cầu để đảm bảo một nghiên cứu tốt. Một nghiên cứu được đánh giá là tốt nếu đảm bảo các vấn đề sau:

- Mục đích, mục tiêu nghiên cứu được xác định rõ ràng;
- Quá trình nghiên cứu, thiết kế nghiên cứu được hoạch định một cách chi tiết;
- Những giới hạn nghiên cứu được trình bày rõ ràng;
- Đáp ứng tốt các tiêu chuẩn đạo đức của nhà nghiên cứu;
- Các phân tích phù hợp với nhu cầu của người ra quyết định;
- Kết quả trình bày một cách rõ ràng, rành mạch, không mơ hồ;
- Các kết luận được chứng minh, bình luận với các nghiên cứu trước có nền tảng và có cơ sở vững chắc.

TÓM TẮT CHƯƠNG 1

Chương 1 giới thiệu định nghĩa nghiên cứu và nghiên cứu trong kinh doanh, so sánh giữa phương pháp nghiên cứu và phương pháp luận nghiên cứu; phương pháp nghiên cứu và phương pháp khoa học.

Theo (Cooper & Schindler, 2011), “nghiên cứu trong kinh doanh (business research) là quá trình điều tra một cách có hệ thống nhằm cung cấp thông tin hướng dẫn ra các quyết định quản trị, đây là quá trình hoạch định, điều tra/ tìm kiếm, phân tích và phổ biến những dữ liệu, thông tin có ý nghĩa cho người ra quyết định một cách linh động và phù hợp để tối đa hóa hiệu suất/ năng lực của tổ chức”

Thông thường, cần phân biệt giữa phương pháp nghiên cứu và phương pháp luận nghiên cứu. Phương pháp nghiên cứu có thể được hiểu là tất cả những phương pháp, kỹ thuật được sử dụng cho việc nghiên cứu (Saunders et al., 2007). Theo (Kothari, 2004), phương pháp luận nghiên cứu là một cách giải quyết có hệ thống vấn đề nghiên cứu, nó có thể được xem như là một ngành khoa học nghiên cứu cách thức thực hiện nghiên cứu một cách khoa học.

Đồng thời, chương này cung giới thiệu các thuật ngữ trong nghiên cứu như khái niệm lý thuyết (concepts), khái niệm nghiên cứu (constructs), định nghĩa đưa vào nghiên cứu (operational definitions) và biến số (variables) nhằm giúp cho người nghiên cứu hiểu rõ hơn về cách khái niệm, dễ dàng cho việc hình thành vấn đề nghiên cứu và thiết kế nghiên.

Một khái niệm lý thuyết (concepts) là một tập hợp những ý nghĩa và đặc điểm được chấp nhận rộng rãi liên quan đến những sự kiện, đối tượng, điều kiện, tình huống và hành vi cụ thể, trong khi đó, khái niệm nghiên cứu (constructs) là một hình ảnh hay một ý tưởng trừu tượng được tạo ra cho một nghiên cứu cụ thể và/ hoặc cho mục đích xây dựng lý thuyết.

Chương 1 cũng đã giới thiệu sơ lược tiến trình nghiên cứu gồm 7 bước: Xác định vấn đề nghiên cứu/ mục tiêu nghiên cứu; tổng quan/ hợp tuyển lý thuyết; phát triển giả thuyết; thiết kế nghiên cứu; tổ chức thu thập; chuẩn bị dữ liệu và phân tích dữ liệu; báo cáo kết quả. Nội dung cụ thể của các bước sẽ được giới thiệu ở các chương tiếp theo.

CHƯƠNG

2

XÁC ĐỊNH VẤN ĐỀ NGHIÊN CỨU (RESEARCH PROBLEM DEFINITION)

Như đã trình bày ở chương 1, Kothari (2004) đã giới thiệu tiến trình nghiên cứu gồm 7 giai đoạn gồm (1) xác định vấn đề nghiên cứu (research problem), (2) tổng quan/ hợp tuyển lý thuyết (review literature), (3) phát triển giả thuyết, (4) thiết kế nghiên cứu (bao gồm cả thiết kế chọn mẫu), (5) thu thập và chuẩn bị dữ liệu, (6) phân tích dữ liệu và diễn giải, báo cáo kết quả.

Chương này sẽ trình bày bước đầu tiên trong tiến trình nghiên cứu là xác định vấn đề nghiên cứu. Trên thực tế thực hiện nghiên cứu (định tính hoặc định lượng), để thực hiện bước này, nhà nghiên cứu phải thực hiện phân tích, hợp tuyển lý thuyết. Do vậy, chương này trình bày phương pháp xác định vấn đề nghiên cứu trong mối quan hệ giới thiệu tổng quan/ hợp tuyển lý thuyết nghiên cứu và phát biểu các giả thuyết.

VẤN ĐỀ NGHIÊN CỨU

Theo Kothari (2004), nghiên cứu một vấn đề đề cập đến một số khó khăn trong kinh nghiệm của nhà nghiên cứu trong bối cảnh của lý thuyết hay tình huống thực tế và muốn thu được một giải pháp cho chúng. Các thành phần của một vấn đề nghiên cứu luôn đặt trong những vấn đề sau:

- Những khó khăn của cá nhân hoặc tổ chức đặt ra cần phải nghiên cứu;
- Những mục tiêu cần phải đạt được;
- Những phương pháp có thể thực hiện nhằm đạt được mục tiêu mong muốn. Thông thường, cần có ít nhất hai phương pháp trở lên để nhà nghiên cứu có những lựa chọn được cho là tối ưu để thực nghiên cứu và đó cũng chính là những phương án thay thế có thể xảy ra trong quá trình nghiên cứu;
- Những nghi ngờ trong suy nghĩ của nhà nghiên cứu đến các phương án thay thế. Điều này có nghĩa là nhà nghiên cứu phải trả lời các câu hỏi liên quan đến hiệu quả tương đối của các phương án thay thế có thể xảy ra;
- Phải gắn liền với những khó khăn trong môi trường, khung cảnh nghiên cứu.

Một vấn đề nghiên cứu là sự đòi hỏi nhà nghiên cứu tìm ra giải pháp tốt nhất cho vấn đề được đưa ra. Cụ thể, cần tìm ra tiến trình hành động, mục tiêu có thể đạt được một cách tối ưu trong hoàn cảnh của môi trường được đưa ra.

Có hai loại vấn đề nghiên cứu: loại vấn đề liên quan đến bản chất của hiện tượng và loại vấn đề liên quan đến mối quan hệ giữa các biến số nghiên cứu. Dù là loại vấn đề nào thì khi xác định vấn đề nghiên cứu, cần chú ý những vấn đề sau:

- Chủ đề đã được thực hiện quá nhiều thường không nên chọn vì chúng ta sẽ rất khó khăn trong việc tìm ra điểm mới và những đóng góp cho nghiên cứu;
- Chủ đề gây sự tranh cãi không nên trở thành sự lựa chọn của nhà nghiên cứu vì rất khó khăn trong việc tổng hợp các lý thuyết liên quan, xác định vấn đề nghiên cứu và thuyết phục sự đồng thuận;
- Nên tránh những vấn đề quá hẹp hoặc quá mơ hồ: Chủ đề nghiên cứu nên quen thuộc (nhưng dĩ nhiên không lựa chọn những vấn đề mà quá nhiều người đã thực hiện như đã đề cập ở trên) và khả thi để có thể có nguồn tài liệu để nghiên cứu. Thông thường, trong các trường Đại học hoặc Viện nghiên cứu có các thư viện điện tử, trong đó đăng tải các công trình đã được công bố, nhà nghiên cứu có thể theo dõi những lý thuyết đã thực hiện nghiên cứu, ưu điểm và hạn chế của nghiên cứu về lý thuyết cũng như phương pháp để làm cơ sở cho việc lựa chọn vấn đề nghiên cứu;
- Khả năng về trình độ, chi phí về thời gian, tiền bạc để thực hiện nghiên cứu: Một nghiên cứu có thể rất thời sự nhưng đòi hỏi một thời gian nghiên cứu lâu dài hoặc chi phí lớn, vượt quá khả năng của nhà nghiên cứu thì không nên lựa chọn để thực hiện.

Tầm quan trọng của việc xác định vấn đề nghiên cứu

Mặc dù mỗi bước trong tiến trình nghiên cứu đều quan trọng nhưng bước xác định vấn đề luôn là quan trọng nhất bởi vì chỉ khi nào vấn đề nghiên cứu đã được xác định rõ ràng thì nghiên cứu mới được thiết kế và tiến hành một cách hợp lý. Tuy nhiên, việc xác định vấn đề không chính xác là nguyên nhân dẫn đến thất bại của nghiên cứu. Xác định vấn đề nghiên cứu có khi rất đơn giản như trong trường hợp có những nghiên cứu sơ bộ, nghiên cứu định tính đã công bố các lý thuyết, nhà nghiên cứu chỉ cần thực hiện nghiên cứu thực nghiệm nhằm kiểm định các mối quan hệ giữa các khái niệm trong mô hình hoặc nghiên cứu lặp lại hoàn toàn (thực hiện lặp lại nghiên cứu đi trước trong một khung gian mới...). Nhưng cũng không hiếm trường hợp người nghiên cứu phải đương đầu với những tình huống mới mà ở đó dường như chưa có một phác họa trước nên trong những tình huống như vậy, việc xác định vấn đề nghiên cứu sẽ trở nên hết sức khó khăn.

Đặc biệt, đối với những nghiên cứu ứng dụng, việc xác định vấn đề nghiên cứu phải được khởi đầu với việc nhận thức vấn đề hoặc cơ hội trong kinh doanh và xác định vấn đề mà nhà quản trị cần phải giải quyết từ đó chuyển thành vấn đề nghiên cứu.

Ví dụ: Nhà quản trị muốn thâm nhập thị trường X cho sản phẩm A [**đây là vấn đề quản trị**], để nhà quản trị ra quyết định này, nhà nghiên cứu cần phải cung cấp các thông tin liên quan đến tiềm năng thị trường, hành vi của khách hàng tại thị trường X đối với sản phẩm A... [**đây là vấn đề cần nghiên cứu**].

Như vậy, xác định vấn đề nghiên cứu rõ ràng là điều kiện tiên quyết cho bất kỳ nghiên cứu và là bước quan trọng nhất và chỉ cần chi tiết về vấn đề nghiên cứu là chúng ta có thể thực hiện các thiết kế nghiên cứu và thực hiện trôi chảy tất cả các bước liên quan trong quá trình thực hiện nghiên cứu. Việc xác định vấn đề nghiên cứu có thể xuất phát từ lý thuyết (ví dụ như những hạn chế về lý thuyết và phương pháp của nghiên cứu trước là vấn đề nghiên cứu của nghiên cứu sau); những yêu cầu từ thực tế cuộc sống cần phải có những kiến thức để điều

chỉnh các hoạt động và đôi khi, từ kinh nghiệm thông qua những trải nghiệm của chính nhà nghiên cứu.

Xác định vấn đề nghiên cứu

Vấn đề nghiên cứu xuất phát từ những vấn đề gặp phải trong thực tế mà chưa tìm ra lời giải thích hợp từ các lý thuyết đã có hoặc những “lỗ hổng” (gap) trong các lý thuyết hiện tại. Người nghiên cứu phải làm quen với các lý thuyết, tài liệu liên quan đến vấn đề nghiên cứu được lựa chọn. Nghiên cứu cần tổng hợp, xem xét (review) hai loại tài liệu, (1) tài liệu liên quan đến các khái niệm lý thuyết (concepts) và (2) những tài liệu là những nghiên cứu thực chứng (empirical literature) được thực hiện trước đây liên quan đến vấn đề nghiên cứu đề xuất. Về cơ bản, kết quả của việc tóm tắt tổng quan tài liệu là kiến thức về những dữ liệu, tài liệu cần cho mục đích thực hiện nghiên cứu, từ đó giúp người nghiên cứu có thể cụ thể hóa vấn đề nghiên cứu của mình trong một bối cảnh có ý nghĩa. Sau đó, người nghiên cứu phát biểu vấn đề nghiên cứu càng cụ thể càng tốt, bước này là bước quan trọng nhất trong tiến trình nghiên cứu. Vấn đề nghiên cứu cần phải được phát biểu một cách không mập mờ (unambiguous) để phân biệt dữ liệu phù hợp và những dữ liệu không phù hợp. Mục tiêu nghiên cứu quyết định dữ liệu nào cần phải thu thập, đặc tính phù hợp của dữ liệu, những mối quan hệ cần khám phá, lựa chọn công cụ phân tích sẽ được sử dụng, lựa chọn công cụ dùng trong việc khám phá các mối quan hệ này và hình thức báo cáo cuối cùng của nghiên cứu. Bên cạnh đó, các vấn đề thực tế tại “thị trường” cũng đóng vai trò quan trọng trong việc xác định vấn đề nghiên cứu, chính những vấn đề thực tế diễn ra là vấn đề cần phải nghiên cứu.

Vấn đề đặt ra là làm thế nào để xác định một vấn đề nghiên cứu rõ ràng? Đây là vấn đề rất khó khăn nhưng cần phải giải quyết một cách thông minh để tránh những rắc rối gặp phải trong một hoạt động nghiên cứu. Cách tiếp cận thông thường là nhà nghiên cứu nên tự mình đưa ra một câu hỏi (hoặc nếu đây là một vấn đề liên quan đến tổ chức thì cá nhân hoặc tập thể trong tổ chức có thể đặt ra câu hỏi cho nhà nghiên cứu), đây được xem là **ý tưởng nghiên cứu** và từ đó biến các ý tưởng nghiên cứu thành vấn đề nghiên cứu.

Xác định vấn đề nghiên cứu đúng đắn và rõ ràng là một phần quan trọng của nghiên cứu và không nên được thực hiện một cách vội vàng. Tuy nhiên, trong thực tế việc này thường xuyên bị bỏ qua và gây ra nhiều vấn đề sau đó. Do đó, vấn đề nghiên cứu nên được xác định một cách có hệ thống, đưa ra điểm tựa cho tất cả các bước liên quan.

Kothari (2004), Saunders et al. (2007) đã giới thiệu một số kỹ thuật đối với việc xác định vấn đề nghiên cứu gồm:

- Khảo sát điểm mạnh của nhà nghiên cứu và sự hiểu biết về bản chất của vấn đề: Để thực hiện một nghiên cứu tốt, nhà nghiên cứu cần phải xem xét điểm mạnh của cá nhân, chúng sẽ giúp cho việc tiếp cận và phát hiện ý tưởng một cách nhanh chóng. Đồng thời, cần hiểu rõ nguồn gốc và bản chất vấn đề. Nhà nghiên cứu có thể thảo luận với những người đầu tiên xây dựng nó, tìm ra nguồn gốc vấn đề như thế nào và với những mục tiêu gì trong quan điểm của họ. Nếu nhà nghiên cứu đã tự mình chỉ rõ vấn đề, nên xem xét lại tất cả một lần nữa những điểm này. Dĩ nhiên, việc có hiểu biết tốt hơn về bản chất của vấn đề liên quan,

nhà nghiên cứu có thể dễ dàng tham gia vào cuộc thảo luận với những người có kiến thức tốt về các vấn đề quan tâm hay là những vấn đề tương tự khác của vấn đề nghiên cứu.

- Xem xét lại và phát hiện vấn đề nghiên cứu từ các nghiên cứu có sẵn (tổng hợp, hợp tuyển lý thuyết): Tất cả những lý thuyết sẵn có liên quan đến vấn đề sắp tới cần thiết phải được khảo sát. Điều này có nghĩa là nhà nghiên cứu phải quen thuộc với các lý thuyết có liên quan trong lĩnh vực này, các báo cáo và bản ghi chép cũng như các lý thuyết liên quan khác. Nhà nghiên cứu phải dành đủ thời gian xem xét các nghiên cứu đã được thực hiện trong các vấn đề liên quan. Điều này được thực hiện để tìm hiểu dữ liệu và các tài liệu khác, có sẵn cho mục đích nghiên cứu. Việc biết những dữ liệu có sẵn thường giúp thu hẹp bản thân vấn đề cũng như kỹ thuật mà có thể được sử dụng. Điều này cũng giúp nhà nghiên cứu biết được nếu có “lỗ hổng” trong lý thuyết hoặc liệu các lý thuyết hiện hành áp dụng đối với các vấn đề được nghiên cứu có mâu thuẫn với nhau?. Các phương pháp được sử dụng trong nghiên cứu trước có hợp lý hay không? Tất cả những điều trên đây sẽ cho phép nhà nghiên cứu có những bước tiến mới trong phát triển kiến thức, ví dụ, nhà nghiên cứu có thể phát triển thêm dựa trên những lý thuyết nền hiện có. Thực hiện nghiên cứu những vấn đề liên quan rất hữu ích cho việc cho biết loại khó khăn có thể gặp phải trong nghiên cứu hiện tại cũng như những thiếu sót có thể có trong phân tích. Việc tổng hợp những nghiên cứu như vậy cũng có thể mang lại những gợi ý hữu ích và cả những **cách tiếp cận mới cho vấn đề hiện tại**.

Các tài liệu sử dụng cho nghiên cứu có thể là các bài viết trên những tạp chí học thuật và chuyên sâu, các loại báo cáo, các loại sách. Đặc biệt, các bài báo khoa học (articles) trên các tạp chí học thuật (journal) có vai trò đặc biệt quan trọng, nhà nghiên cứu có thể đọc phần tổng quan, hợp tuyển lý thuyết (literature review) và phương pháp nghiên cứu trong các bài viết đánh giá (review articles) để nhận thức những kiến thức trong chủ đề và định hướng những lĩnh vực cần thiết để tiến hành xa hơn.

- Phát triển ý tưởng thông qua thảo luận với các chuyên gia/ điều tra kinh nghiệm (experience survey): Thảo luận liên quan đến một vấn đề thường cho ra những thông tin hữu ích. Do vậy, nhà nghiên cứu phải thảo luận vấn đề với những đồng sự và những người có đủ kinh nghiệm trong cùng lĩnh vực hoặc trong vấn đề tương tự. Những người có giàu kinh nghiệm giúp làm sáng tỏ cho các nhà nghiên cứu về các khía cạnh khác nhau của nghiên cứu đề xuất, những lời khuyên và bình luận của họ thường là vô giá với các nhà nghiên cứu. Thông thường, thông tin từ các chuyên gia có được là do trao đổi chứ không phải bằng bản câu hỏi chính thức, tuy nhiên để gia tăng hiệu quả khi trao đổi người nghiên cứu nên chuẩn bị một số nội dung cần thiết trước cuộc trao đổi.
- Xem xét dữ liệu thứ cấp: Dữ liệu thứ cấp là các dữ liệu có sẵn, được thu thập cho một mục đích nào đó ngoài dữ liệu điều tra trực tiếp về vấn đề nghiên cứu (dữ liệu sơ cấp). Dữ liệu thứ cấp bao gồm những thông tin được cung cấp từ các doanh nghiệp hoặc từ các tổ chức chính phủ, từ những doanh nghiệp nghiên cứu công bố trên các phương tiện thông tin đại chúng hay những cơ sở dữ liệu được công bố, lưu trữ tại chính doanh nghiệp... Phân tích dữ liệu thứ cấp là bước cần thiết giúp phác thảo và xác định vấn đề nghiên cứu, điều mà nếu dựa vào các dữ liệu sơ cấp, dù mức độ chi tiết và cụ thể cao hơn, cũng không thể đạt

được. Xem xét dữ liệu thứ cấp giúp cho việc xác định có thật sự cần thiết thực hiện nghiên cứu bởi vì có thể có những dữ liệu thứ cấp đã làm rõ vấn đề nghiên cứu. Đồng thời, qua đó có thể giúp cho việc định hướng những dữ liệu, thông tin cần phải thu thập trong tương lai.

- Tập kích não (brainstorming) có thể được sử dụng để hình thành và chất lọc ý tưởng trên các bước thực hiện:
 - Xác định vấn đề của chính người tham gia;
 - Yêu cầu những đề xuất liên quan đến vấn đề;
 - Ghi lại tất cả các đề xuất liên quan đến vấn đề trên cơ sở mọi đề xuất sẽ không bị phê bình hay đánh giá; tất cả những đề xuất dù kỳ lạ cũng phải được ghi lại và ghi lại càng nhiều đề xuất càng tốt;
 - Đánh giá lại tất cả những đề xuất và hàm ý của nó;
 - Phân tích danh sách các đề xuất và xác định đề xuất nào là ý tưởng nghiên cứu hấp dẫn đối với nhà nghiên cứu.
- Chọn lọc các ý tưởng, vấn đề nghiên cứu thông qua kỹ thuật Delphi: Kỹ thuật Delphi là kỹ thuật chất lọc các ý tưởng, kỹ thuật này bao gồm việc sử dụng một nhóm người tham gia (có hiểu biết và quan tâm đến chủ đề nghiên cứu) để hình thành và lựa chọn ý tưởng, vấn đề nghiên cứu. Cách thức thực hiện kỹ thuật này như sau:
 - Tóm tắt cho các thành viên của nhóm về ý tưởng, vấn đề nghiên cứu;
 - Phổ biến và khuyến khích các thành viên nhóm làm sáng tỏ và tìm thêm thông tin phù hợp;
 - Yêu cầu các thành viên hình thành một cách độc lập 3-5 ý tưởng cụ thể dựa trên ý tưởng đã được mô tả;
 - Thu thập ý tưởng theo dạng chưa biên tập và chưa có tác giả, phân phát cho tất cả các thành viên;
 - Lặp lại bước 2 đến 4 để các thành viên bình luận về ý tưởng nghiên cứu và xem lại những đóng góp của họ liên quan điều những người khác nói;
 - Tiếp tục quy trình cho đến khi đạt được sự đồng thuận.

Việc xác định những câu hỏi nghiên cứu luôn là một khó khăn, giống như việc xác định vấn đề nghiên cứu, câu hỏi phải đủ hàm ý để hình thành dạng nghiên cứu nhằm đáp ứng những vấn đề mong đợi. Câu hỏi nghiên cứu không nên thiết kế quá khó (khó khăn trong việc giải quyết vấn đề sau này), tránh bằng mọi giá đưa ra câu hỏi phải mang lại kiến thức mới (vì có những nghiên cứu lặp lại sẽ phù hợp trong việc ứng dụng nó vào bối cảnh nghiên cứu). Tốt nhất, khi có ý tưởng, vấn đề nghiên cứu, chúng ta sẽ chia nhỏ những câu hỏi từ những ý tưởng, vấn đề ban đầu thành những phần bớt dần mức độ phức tạp và khó hiểu.

Mục tiêu nghiên cứu nên có xu hướng đưa đến tính cụ thể hơn so với những câu hỏi nghiên cứu, các nhà nghiên cứu thường chấp nhận những mục tiêu nghiên cứu khi có chứng cứ cảm

nhận rõ ràng về mục đích và phương pháp của người nghiên cứu. Cụ thể hơn, mục tiêu nghiên cứu phải đáp ứng những vấn đề (theo nguyên tắc SMART):

- Cụ thể (Specific): Mong đợi cụ thể, chính xác vấn đề gì từ nghiên cứu?
- Có thể đo lường được (Measureable): Mục tiêu có thể đo lường được không?, phương pháp nào có thể sử dụng để đo lường?
- Có thể đạt được (Achievable): Các mục tiêu đưa ra có thể đạt được trong bối cảnh nghiên cứu không?
- Thực tế (Realistic): Có thể có thời gian, kinh phí, khả năng để hoàn thành mục tiêu trong thời hạn cho phép không?
- Đúng lúc (Timely): Thời gian hoàn thành là bao nhiêu?, có thể chấp nhận khoảng thời gian đó không?

Ví dụ: Xác định vấn đề và mục tiêu nghiên cứu¹

Việt Nam gia nhập Tổ chức Thương mại thế giới (WTO) đã và đang mở ra cho kinh tế Việt Nam nói chung và ngành ngân hàng nói riêng những vận hội và thách thức mới. Các ngân hàng trong nước sẽ có cơ hội mở rộng hợp tác kinh doanh với cộng đồng doanh nghiệp trong và ngoài nước, có thể tiếp thu nguồn vốn và những kinh nghiệm trong việc quản lý và điều hành hoạt động kinh doanh trong môi trường kinh doanh năng động. Những điều khoản trong việc đàm phán yêu cầu chúng ta phải chấp nhận mở cửa hơn nữa các dịch vụ ngân hàng. Khi đó, các ngân hàng nước ngoài, các tổ chức tín dụng dưới nhiều hình thức khác nhau sẽ có nhiều cơ hội hơn trong việc xâm nhập thị trường Việt Nam. Theo cam kết giữa Việt Nam và các nước thành viên, cho đến nay (năm 2010), các ngân hàng nước ngoài sẽ được phép thực hiện hầu hết các dịch vụ ngân hàng như một ngân hàng trong nước (trừ dịch vụ tư vấn và cung cấp thông tin ngân hàng) **[Tình hình chung kinh tế Việt Nam và vấn đề đặt ra cho ngành ngân hàng]**

Một trong những mối quan tâm hàng đầu là việc mở cửa thị trường dẫn tới cạnh tranh gay gắt giữa các ngân hàng trong và ngoài nước trong khi phần lớn ngân hàng Việt Nam có ít vốn, công nghệ và khả năng cạnh tranh để có thể đứng vững được trên thị trường. Thực tế cho thấy khi Việt Nam gia nhập WTO, nhiều ngân hàng nước ngoài có tiềm lực mạnh mẽ đã và đang nhanh chóng đầu tư vào Việt Nam. Do vậy, các ngân hàng trong nước sẽ nhường thị phần tiền vay và tiền gửi cho các ngân hàng nước ngoài. Theo báo cáo khảo sát của Chương trình Phát triển Liên hợp quốc năm 2006 cho rằng trong tương lai, có khoảng 45% khách hàng sẽ chuyển sang vay vốn của ngân hàng nước ngoài thay vì của ngân hàng trong nước, 50% chọn dịch vụ ngân hàng nước ngoài thay thế và 50% còn lại chọn ngân hàng nước ngoài để gửi tiền, đặc biệt là ngoại tệ... Đây là một trong những thách thức rất lớn đặt ra cho các ngân hàng trong nước trong việc đảm bảo và nâng cao thị phần của doanh nghiệp **[Biện luận khó khăn trong lĩnh vực ngân hàng]**.

¹. Tài liệu phục vụ nghiên cứu: Lê Văn Huy (2010), Thiết lập chỉ số hài lòng khách hàng (CSI) trong lĩnh vực ngân hàng: Nghiên cứu thực tiễn tại Thành phố Đà Nẵng, *Đề tài Khoa học và Công nghệ cấp Bộ, Mã số B2008-ĐN04-31*.

Trong cuộc sống, con người luôn ý thức việc tồn tại sự khác biệt giữa trạng thái thực tế và ước muốn về vật chất cũng như tinh thần, sự khác biệt sẽ tạo ra động cơ và thôi thúc con người dẫn đến hành động tiêu dùng và mong muốn có sự thỏa mãn về sản phẩm đó. Kinh nghiệm tại các nước phát triển, khi cường độ cạnh tranh càng cao thì doanh nghiệp càng phải nghiên cứu và nâng cao khả năng hài lòng nhằm gia tăng sự trung thành của khách hàng đối với doanh nghiệp. Tại các nước phát triển, nhiều nghiên cứu đã tập trung nghiên cứu các nhân tố tiền tố của sự hài lòng khách hàng và khẳng định chất lượng dịch vụ là một trong những biến số quan trọng quyết định sự hài lòng của khách hàng, từ đó xây dựng chỉ số hài lòng khách hàng (CSI – customer satisfaction index) làm cơ sở cho công tác quản trị tại các ngân hàng. Tại Việt Nam nói chung và trong lĩnh vực ngân hàng nói riêng không nhiều những nghiên cứu thực nghiệm liên quan đến sự hài lòng của khách hàng, đặc biệt là các nghiên cứu cơ bản **[lỗ hổng nghiên cứu]**. Do vậy, trong điều kiện nghiên cứu lí thuyết và ứng dụng chỉ số hài lòng khách hàng (CSI) tại Việt Nam không nhiều như đã đề cập, nghiên cứu **“Thiết lập chỉ số hài lòng khách hàng dựa trên chất lượng dịch vụ”** phần nào giải quyết vấn đề trên.

Mục tiêu nghiên cứu

Mục tiêu của đề tài gồm các vấn đề chủ yếu là:

- Xây dựng thang đo chất lượng dịch vụ làm định hướng xây dựng chỉ số hài lòng khách hàng;
- Nghiên cứu mối quan hệ giữa chất lượng dịch vụ và sự hài lòng khách hàng;
- Thiết lập chỉ số hài lòng khách hàng trong lĩnh vực ngân hàng.
- Một số gợi ý cho nhà quản lý trong việc sử dụng chỉ số hài lòng trong hoạch định chiến lược kinh doanh.

Câu hỏi nghiên cứu

- Cấu thành chất lượng dịch vụ là gì?
- Có tồn tại hay không mối quan hệ giữa các yếu tố thành phần chất lượng dịch vụ và sự hài lòng của khách hàng?
- Đây là chỉ số hài lòng khách hàng trong lĩnh vực ngân hàng tại Việt Nam?

TỔNG KẾT, HỢP TUYÊN LÝ THUYẾT (LITERATURE REVIEW)

Trong nghiên cứu, người ta phân thành nghiên cứu ứng dụng và nghiên cứu cơ bản. Trong đó, nghiên cứu cơ bản bao gồm nghiên cứu cơ bản nguyên thủy (nghiên cứu hoàn toàn mới) và nghiên cứu lặp lại gồm (1) lặp lại hoàn toàn; (2) lặp lại gần như hoàn toàn (về khái niệm, về mô hình nghiên cứu) nhưng trong một không gian hoặc một giới hạn nghiên cứu khác (ví dụ như trong một môi trường văn hóa khác); (3) lặp lại trong nhiều không gian, ngữ cảnh khác nhau và (4) lặp lại có điều chỉnh (thêm, bớt một vài khái niệm hoặc điều chỉnh, bổ sung một số mối quan hệ).

Thông thường, sinh viên và học viên cao học thường thực hiện các nghiên cứu ứng dụng (như marketing cho sản phẩm xăm lớp ô tô tại Công ty Cao su...) trên cơ sở ứng dụng lý thuyết marketing triển khai chiến lược, chính sách cho công ty hoặc nghiên cứu cơ bản dạng lặp lại.

Để thực hiện tốt nghiên cứu, nhà nghiên cứu cần thực hiện giới thiệu tổng quan về nghiên cứu, đồng thời hợp tuyển lý thuyết. Như đã đề cập ở trên, việc tổng quan, hợp tuyển về lý thuyết bao gồm:

- Tổng quan, hợp tuyển về khái niệm, khái niệm nghiên cứu và lý thuyết liên quan;
- Tổng quan, hợp tuyển về phương pháp, kết quả của các nghiên cứu có trước liên quan đến vấn đề nghiên cứu. Phần này có nhà nghiên cứu có thể phân tích, bình luận về lý thuyết, phương pháp, kết quả nghiên cứu để có thể lựa chọn vấn đề nghiên cứu trong tương lai.

Một khi vấn đề nghiên cứu đã được xác định, một bản tóm tắt về vấn đề nghiên cứu phải được viết ra. Đồng thời, người nghiên cứu thực hiện điều tra lý thuyết (extensive literature survey) liên quan đến vấn đề nghiên cứu đã được phát biểu. Bước xác định vấn đề phải phụ thuộc vào nghiên cứu lý thuyết nhằm xem xét (1) lý thuyết nền, (2) các lỗ hổng nghiên cứu (về lý thuyết cũng như về phương pháp nghiên cứu), những vấn đề này chính là cơ sở để xác định vấn đề nghiên cứu.

Với mục đích này, việc đầu tiên là tìm kiếm những bài báo nghiên cứu trên những tạp chí khoa học, những báo cáo chính phủ, báo cáo về ngành, tổng kết hội nghị cũng là những nguồn tham khảo đáng tin cậy phụ thuộc vào bản chất của vấn đề nghiên cứu. Đặc biệt, trong một nghiên cứu cơ bản, các bài báo khoa học sẽ giúp cho nhà nghiên cứu xác định và tổng hợp các ý thuyết nền; kết quả nghiên cứu giúp cho việc vận dụng, đối chiếu, bình luận kết quả; phương pháp nghiên cứu giúp cho việc thực hiện lại phương pháp nghiên cứu trong một môi trường mới (nếu phương pháp nghiên cứu phù hợp) hoặc bình luận để áp dụng một phương pháp nghiên cứu mới.

Trong quá trình tổng hợp, hợp tuyển lý thuyết cần chú ý đến các định nghĩa, khái niệm, thang đo, mối quan hệ và hướng quan hệ giữa các khái niệm (biến số) làm cơ sở cho việc đưa ra các bình luận, biện luận, so sánh trong phần xây dựng lý thuyết cũng như bình luận kết quả sau này.

Ví dụ: Giới thiệu tổng quan/ hợp tuyển lý thuyết về Hội nhập thương mại điện tử (e-commerce) tại các doanh nghiệp vừa và nhỏ tại Việt Nam²

Thương mại điện tử (TMĐT) ngày càng khẳng định được vai trò của mình trong việc đưa lại cho DN những thông tin phong phú về thị trường và đối tác, giảm chi phí, rút ngắn chu kỳ sản xuất, dễ dàng tạo dựng và củng cố quan hệ khách hàng. Ở Việt Nam, TMĐT đã có quá trình hình thành hơn 10 năm qua và được nhận định là đang bắt đầu bước vào giai đoạn phát triển mạnh. Bên cạnh đó, TMĐT ngày càng chứng minh được tính ưu việt của mình bằng việc cho phép tiến hành các thương vụ ở mọi lúc, mọi nơi một cách thuận tiện, thời

². Tài liệu phục vụ nghiên cứu: Lê Văn Huy (2005), Nhân tố tác động đến hội nhập thương mại điện tử trong các doanh nghiệp Việt Nam: Một nghiên cứu thực tiễn, *Tạp chí Nghiên cứu kinh tế*, Viện Nghiên cứu kinh tế, 323, 72-78.

gian giao dịch có thể lên tới 24 giờ mỗi ngày cả 7 ngày trong tuần. Thông qua mạng Internet, các giao dịch điện tử được tiến hành trên phạm vi toàn cầu. Với lợi thế này một công ty nhỏ cũng có khả năng như một công ty xuyên quốc gia. Để có được một mạng lưới khách hàng trên thế giới, các công ty xuyên quốc gia đã phải đầu tư rất nhiều tiền của và thời gian nhưng nay các công ty đa quốc gia này đang phải chịu cạnh tranh với hàng vạn công ty nhỏ đã xuất hiện trong mắt các khách hàng trên toàn thế giới nhờ có TMĐT. Đối với Việt Nam, trong điều kiện phát triển kinh tế đất nước và hội nhập nền kinh tế toàn cầu, để đảm bảo sự tồn tại và phát triển, đòi hỏi các doanh nghiệp phải sử dụng những kĩ thuật công nghệ mới, đặc biệt là ứng dụng công nghệ thông tin, Internet và thương mại điện tử (TMĐT) trong quá trình hoạt động kinh doanh của doanh nghiệp **[Giới thiệu để dẫn nhập vấn đề]**.

Khái niệm về “thương mại điện tử” dựa trên cách tiếp cận của hai thuật ngữ “thương mại” và “điện tử”, có nhiều cách tiếp cận khái niệm thương mại điện tử **[Sau đây giới thiệu các khái niệm về TMĐT]**.

Theo Luật mẫu về Thương mại điện tử UNCITRAL (1996), TMĐT được hiểu theo nghĩa rộng: *“Thuật ngữ Thương mại cần được diễn giải theo nghĩa rộng để bao quát các vấn đề phát sinh từ mọi quan hệ mang tính chất thương mại dù có hay không có hợp đồng. Các quan hệ mang tính thương mại bao gồm bất cứ giao dịch thương mại nào về cung cấp hoặc trao đổi hàng hóa hoặc dịch vụ; thỏa thuận phân phối; đại diện hoặc đại lý thương mại, ủy thác hoa hồng; cho thuê dài hạn; xây dựng các công trình; tư vấn; kỹ thuật công trình; đầu tư; cấp vốn; ngân hàng; bảo hiểm; thỏa thuận khai thác hoặc tô nhượng; liên doanh các hình thức khác về hợp tác công nghiệp hoặc kinh doanh; chuyên chở hàng hóa hay hành khách bằng đường biển, đường không, đường sắt hoặc đường bộ”*.

Ủy ban Châu Âu (1998) nêu khái niệm: *“TMĐT được hiểu là việc thực hiện hoạt động kinh doanh qua các phương tiện điện tử. Nó dựa trên việc xử lý và truyền dữ liệu điện tử dưới dạng văn bản, âm thanh, hình ảnh”*. Trong khi đó, theo Tổ chức Thương mại thế giới (WTO): *“TMĐT bao gồm việc sản xuất, quảng cáo, bán hàng và phân phối sản phẩm được mua bán và thanh toán trên mạng internet, nhưng được giao nhận một cách hữu hình cả các sản phẩm được giao nhận cũng như những thông tin số hóa thông qua mạng internet”*.

Trong nghiên cứu này sẽ sử dụng định nghĩa TMĐT của Efraim Turban (2004): *“TMĐT là quy trình mua, bán, chuyển giao hoặc trao đổi sản phẩm, dịch vụ và/hoặc thông tin thông qua mạng máy tính, bao gồm Internet”*.

Phần lớn, những nghiên cứu về các nhân tố ảnh hưởng đến việc hội nhập TMĐT được phát triển trên cơ sở các nhân tố ảnh hưởng đến hội nhập công nghệ mới như: công nghệ thông tin (IT), hệ thống thông tin (IS), Internet. Việc hội nhập công nghệ mới ở một doanh nghiệp được thể hiện theo ba giai đoạn: ý định (ý tưởng) ban đầu, hội nhập và ứng dụng (Pierre & Delbecq, 1977). Theo đánh giá của các chuyên gia thuộc Bộ Thương Mại (Vụ châu Á – Thái Bình Dương) thì Việt Nam đang ở bước đầu tiên của giai đoạn thứ nhất của quá trình hội nhập. Trong điều kiện tại Việt nam trong việc hội nhập TMĐT, chúng ta có thể xem xét và phân tích việc hội nhập TMDT của các doanh nghiệp trên hai khía cạnh: *giai đoạn (thời kì) hội nhập và mức độ hội nhập*.

- Giai đoạn (thời kì) hội nhập của doanh nghiệp được phản ánh bởi việc doanh nghiệp đó đang trong giai đoạn áp dụng, có hoặc không có ý định triển khai trong tương lai những hoạt động thương mại điện tử như thư điện tử (e-mail), trao đổi dữ liệu điện tử (EDI), thanh toán điện tử (electronic payment), giao gửi số hóa (digital delivery of content)) trong hoạt động kinh doanh của doanh nghiệp (Teo & Tan, 1998; Thong, 1999; Truong & Rao, 2002). Do đó có thể sắp xếp một doanh nghiệp ở trong giai đoạn *đang áp dụng* (doanh nghiệp đó đang áp dụng ít nhất một hoạt động thương mại điện tử trong quá trình kinh doanh của mình), *doanh nghiệp thăm dò* (doanh nghiệp dự kiến sẽ áp dụng ít nhất một hoạt động trong 5 năm tới) và *doanh nghiệp đi sau* (không có ý định áp dụng hoặc sẽ áp dụng nhưng sau 5 năm nữa).
- Mức độ hội nhập TMĐT trong các doanh nghiệp được phản ánh bởi mức độ ứng dụng và triển khai các hoạt động TMĐT trong *doanh nghiệp đang áp dụng TMĐT*. **[Giới thiệu, về mức độ hội nhập hay sử dụng TMĐT]**. *Mức độ hội nhập TMĐT* trong các doanh nghiệp được xem xét thông qua *mức độ ứng dụng* các hoạt động TMĐT trong doanh nghiệp như thư điện tử (e-mail), trao đổi dữ liệu điện tử (EDI), xây dựng trang web để giới thiệu thông tin về doanh nghiệp, thanh toán điện tử (electronic payment), giao gửi số hóa (digital delivery of content), giao dịch trực tuyến (e-transaction)... Trong nghiên cứu này, chúng tôi sắp xếp các doanh nghiệp, tùy theo tình hình thực tế, có thể hoặc ở trong giai đoạn *đang ứng dụng* (doanh nghiệp đó đang ứng dụng ít nhất một hoạt động thương mại điện tử trong quá trình kinh doanh của mình), hoặc trong giai đoạn *thăm dò* (doanh nghiệp dự kiến sẽ ứng dụng ít nhất một hoạt động trong 5 năm tới) hoặc *không áp dụng, lạc hậu* (không có ý định áp dụng hoặc sẽ ứng dụng nhưng sau 5 năm nữa).

Chúng ta nhận thấy rằng, mỗi doanh nghiệp, dù trong giai đoạn nào và với mức độ hội nhập nào cũng đều chịu sự tác động bởi các yếu tố bên trong và bên ngoài doanh nghiệp, những nhân tố tác động đến hội nhập TMĐT tại doanh nghiệp được sắp xếp thành:

Biểu 1. Biểu tóm tắt (sơ lược) các nghiên cứu thực nghiệm liên quan đến hội nhập TMĐT tại các doanh nghiệp vừa và nhỏ

[Thực hiện biểu này để làm cơ sở cho việc hợp tuyển các nhân tố tác động/ ảnh hưởng đến hội nhập thương mại điện tử trong các DNVVN tại các quốc gia trên thế giới]

Tác giả	Đối tượng	Loại hình	Yếu tố ảnh hưởng chính
(Ling, 2001)	TMĐT lý thuyết	DNNVV Singapore	- Tổ chức: quy mô doanh nghiệp, sự ủng hộ của lãnh đạo, cấu trúc và văn hóa DN - Đổi mới: Lợi ích liên quan, sự phức tạp, sự tương hợp - Truyền thông: Nguồn thông tin, chất lượng và số lượng kênh thông tin - Yếu tố ngành: cường độ cạnh tranh, sức ép của các trung gian - Chính phủ: sự ủng hộ của chính phủ, hạ tầng công nghệ thông tin quốc gia, sự khác

			biệt về văn hóa
Kendall, Tung, Chua, Ng, and Tan (2001)	TMĐT thực nghiệm	DNNVV Singapore	- Nhận thức được lợi ích liên quan - Nhận thức sự tương thích - Nhận thức sự phức tạp
Rashid, Al-Qirim (2005)	TMĐT lý thuyết	DNNVV New Zeland	- Nhân tố kỹ thuật : lợi ích liên quan, sự tương hợp, sự phức tạp - Nhân tố tổ chức: Quy mô DN, chất lượng của kỹ thuật thông tin, cường độ thông tin, sự ủng hộ của lãnh đạo - Nhân tố môi trường : Sức ép cạnh tranh, yêu cầu của nhà cung cấp và người mua, chính sách của chính phủ, vai trò của chính phủ - Nhân tố cá nhân : Sự đổi mới của lãnh đạo, kiến thức về kỹ thuật và hệ thống thông tin của lãnh đạo
Lertwongsatien and Wongpinunwatana (2003)	TMĐT thực nghiệm	DNNVV Thái Lan	- Nhân tố tổ chức: Quy mô doanh nghiệp, sự ủng hộ của lãnh đạo, sự tồn tại của phòng kỹ thuật thông tin - Nhân tố kỹ thuật: Lợi ích liên quan, sự tương hợp - Nhân tố môi trường
Seyal and Rahman (2003)	TMĐT thực nghiệm	DNNVV Brunây	- Vai trò của nhà lãnh đạo : Thái độ và sự trợ giúp đối với việc hội nhập - Yếu tố thuộc về tổ chức : Quy mô doanh nghiệp, loại hình kinh doanh - Yếu tố hội nhập : Lợi ích liên quan, sự tương hợp, sự phức tạp
Teo, Chan và Parker (2004)	TMĐT	DNNVV	- Yếu tố về công nghệ: Nhận thức về lợi thế tương đối, khả năng tương thích, sự phức tạp - Yếu tố về tổ chức : Kiến thức và kinh nghiệm về TMĐT, thái độ của nhà quản lý đối với TMĐT - Yếu tố môi trường : thay đổi tác nhân bên ngoài, áp lực từ đối thủ cạnh tranh, áp lực từ đối tác kinh doanh
Laosethakul K. (2006)	TMĐT	DNNVV Thailand	- Hành vi xã hội & nền văn hóa dân tộc - Cơ sở hạ tầng

Huy and Filiatrault (2006)	TMĐT	DNNVV Việt Nam	- Quy mô và nguồn lực DN - Định hướng chiến lược - Hiểu biết về TMĐT của nhân viên - Thái độ của nhà quản lý đối với đổi mới - Hỗ trợ của Chính phủ - Cường độ cạnh tranh - Hạ tầng công nghệ quốc gia...
Sutanonpaiboon và Pearson (2006)	TMĐT		- Định hướng kinh doanh - Môi trường - TMĐT dễ sử dụng cho khách hàng - TMĐT hữu dụng cho khách hàng - Sự sẵn sàng của tổ chức
Princely Ifinedo (2008)	TMĐT	DNNVV Canada	- Tính sẵn sàng của DN - Sự hỗ trợ của nhà quản lý
Ramsey, Ibbotson và Mccole (2008)	TMĐT		- Năng lực của TMĐT - Cơ hội công nhận công nghệ - Định hướng khách hàng - Độ nhạy của cạnh tranh - Nhận thức tính khả thi của CNTT - Kỹ năng phát triển thiết bị điện tử
Wilson, Daniel và Davies (2008)	TMĐT		- Lãnh đạo hỗ trợ - Nhà quản lý biết về lợi ích kinh doanh - Có kỹ năng CNTT - Sự sẵn sàng giúp đỡ - Ưu tiên của TMĐT cho DN - Nhận thức nguy cơ - Yêu cầu của khách hàng
Ghorishi M. (2009)	TMĐT	DNNVV Iran	- Tính hữu ích của nhận thức - Áp lực chính trị - Tính sẵn sàng về tổ chức - Tính tương hợp - Thái độ của nhà quản lý - Áp lực kinh tế xã hội

[Bảng tổng hợp một số nghiên cứu thực nghiệm liên quan đến các nhân tố tác động đến hội nhập TMĐT tại các doanh nghiệp vừa và nhỏ, thông thường được tổng hợp trong Phụ lục, làm cơ sở cho tổng hợp lý thuyết]

Các yếu tố thuộc về tổ chức:

Khi hội nhập thương mại điện tử, một trong những nhóm yếu tố đầu tiên cần phải xem xét là những yếu tố về tổ chức, các tác giả đã khẳng định những yếu tố: *đặc điểm sản phẩm*, *quy mô doanh nghiệp* (lượng nhân viên, quy mô thị trường), *loại hình kinh doanh* (Thong & Yap, 1995), *định hướng chiến lược* về hội nhập công nghệ thông tin và TMĐT của doanh nghiệp (Auger, BarNir, & Gallagher, 2003), *những hiểu biết về TMĐT của nhân viên* (Thong, 1999), *những nguồn lực* (về tài chính, nhân sự, cơ sở hạ tầng về CNTT, khả năng duy trì hoạt động TMĐT...), và *văn hóa* của doanh nghiệp (Chieochan, Lindley, & Dunn, 2002) có tác động trực tiếp đến hội nhập công nghệ mới và TMĐT tại các doanh nghiệp. Thực tế ở Việt Nam, một trong những khó khăn lớn nhất của các doanh nghiệp là việc thiếu hụt lực lượng có tay nghề để đáp ứng nhu cầu của công nghệ cao và thiếu vốn đầu tư để xây dựng hạ tầng công nghệ. Mặt khác, chính các doanh nghiệp chưa chịu hoặc chưa có định hướng chiến lược về TMĐT cũng như chưa biết cách triển khai dự án TMĐT tại chính doanh nghiệp, chưa cung cấp cho nhân viên một sự nhận thức đúng đắn về tầm quan trọng của việc ứng dụng công nghệ mới và đặc biệt là TMĐT trong doanh nghiệp. Theo báo cáo của Bộ bưu chính viễn thông (8-2004), hiện nay, trên 85% doanh nghiệp đã có sử dụng mạng máy tính, trong hầu hết các doanh nghiệp đã sử dụng máy tính, 14% có mạng diện rộng (WAN), đây là một điều kiện thuận lợi cho việc phát triển TMĐT tại doanh nghiệp, tuy nhiên, để triển khai được dự án TMĐT, cần phải có một giải pháp đồng bộ từ phía doanh nghiệp về hạ tầng cơ sở về công nghệ thông tin cũng như việc huấn luyện cho nhân viên những kiến thức về kinh doanh trên mạng để cho họ có thể làm việc trong môi trường kinh doanh TMĐT một cách tốt nhất.

Các yếu tố liên quan đến đặc điểm của người lãnh đạo:

Để cho việc hội nhập công nghệ mới nói chung và TMĐT nói riêng được thực hiện nhanh chóng trong doanh nghiệp, một vấn đề đặt ra là đòi hỏi *người lãnh đạo* phải có những *nhận thức* và *kiến thức* nhất định về vai trò của TMĐT đối với doanh nghiệp (Chieochan et al., 2002; Thong & Yap, 1995) từ đó họ sẽ có những *thái độ tích cực* (Seyal & Rahman, 2003; Thong, 1999) đối với việc xúc tiến thực hiện và ứng dụng TMĐT trong hoạt động kinh doanh của doanh nghiệp. Theo VietnamNet, "TMĐT đang... "đi bộ" ở Việt Nam", một nguyên nhân lớn nhất và quan trọng nhất là do nhận thức của xã hội về TMĐT và đặc biệt là nhận thức của chính bản thân người lãnh đạo doanh nghiệp đối với loại hình này". Trong thời gian tới đây, TMĐT là việc đương nhiên Việt Nam phải thực hiện trong tiến trình hội nhập kinh tế khu vực và thế giới – nhất là khi Việt Nam đã là thành viên của AFTA, tiến tới gia nhập WTO. Việc cam kết gia nhập e-ASEAN, e-APEC cũng là một động lực cho chính phủ Việt Nam cũng như chính các doanh nghiệp nhận thức rõ ràng hơn vai trò và lợi thế trong việc phát triển TMĐT tại doanh nghiệp, và Việt Nam cũng sẽ và phải tận dụng được những lợi thế về công nghệ mới của những quốc gia phát triển. Cũng theo VietNamNet, hiện ở Việt Nam có một số doanh nghiệp mà lãnh đạo đang nỗ lực biến TMĐT trở thành công cụ chính trong hoạt động kinh doanh của mình như VEC, VASC, VDC, VNET... hay sử dụng TMĐT để tiến hành hoạt động xúc tiến thương mại, hỗ trợ các doanh nghiệp Việt Nam thâm nhập thị trường Australia, Mỹ, Châu Âu như công ty VBD, GOL (www.goodsonlines.com). Với sự

hỗ trợ của các công ty xúc tiến thương mại điện tử, các sở khoa học và công nghệ của các tỉnh, thành phố đã mở những lớp bồi dưỡng cho các lãnh đạo doanh nghiệp nhận thức về tầm quan trọng của TMĐT đối với doanh nghiệp cũng như cách thức khảo sát và triển khai dự án TMĐT, điều đó giúp cho lãnh đạo của các doanh nghiệp có những hiểu biết nhất định và tạo điều kiện cho họ có một thái độ tích cực hơn đối với việc hội nhập. Mặt khác, để thực hiện TMĐT, ban lãnh đạo cũng cần phải nhận thức và chấp nhận những cải tiến quy trình quản lí, cải tiến bộ máy, thay đổi văn hóa làm việc trong chính doanh nghiệp để đáp ứng với những thay đổi trong quá trình kinh doanh.

Các yếu tố bên ngoài (môi trường):

Hạ tầng công nghệ thông tin, những chính sách vĩ mô của chính phủ và *sự trợ giúp của các doanh nghiệp lớn* đã hội nhập TMĐT đóng một vai trò cực kì quan trọng trong tiến trình hội nhập của các doanh nghiệp (Ling, 2001; Tan & Teo, 2000), việc chuẩn bị những cơ sở về server (máy chủ) cho mỗi khu vực, đường truyền (theo đường điện thoại, cable) giúp cho việc truy cập Internet của các doanh nghiệp một cách nhanh chóng. Bên cạnh đó, việc chậm xúc tiến và xây dựng những văn bản chấp nhận chữ kí điện tử, chứng nhận điện tử, thanh toán điện tử, luật thương mại điện tử cũng như những tồn tại trong vấn đề bảo mật làm cho doanh nghiệp “ngần ngại” trong việc áp dụng. Xây dựng một hệ thống văn bản pháp quy liên quan đến thương mại điện tử, chữ kí và thanh toán điện tử để điều chỉnh các vấn đề phát sinh trong giao dịch điện tử, theo đó, chính phủ là người chịu trách nhiệm chính trong việc triển khai chiến lược và xây dựng cơ sở hạ tầng cho việc phát triển, xây dựng những văn bản quy định về việc áp dụng luật pháp và các quy định quốc tế về ứng xử trong TMĐT với các giao dịch quốc tế là một vấn đề cấp thiết cần đặt ra. Bên cạnh đó, chính phủ cũng cần chuẩn bị các điều kiện thích hợp cho việc cam kết thực hiện các chính sách của WTO về tiến trình tự do hóa thương mại, trong đó tự do hóa ngành công nghệ truyền thông, cạnh tranh độc quyền để đảm bảo các điều kiện kĩ thuật và tương quan giá cả hợp lí so với các nước trong khu vực và thế giới, khuyến khích nghiên cứu và triển khai các dịch vụ viễn thông cho các thành phần kinh tế, lập các kế hoạch tổng thể cho ứng dụng kĩ thuật TMĐT trong các doanh nghiệp. Một trong những yếu tố môi trường tác động đến việc hội nhập là *sức ép từ chính những khách hàng và nhà cung cấp* của doanh nghiệp (Ling, 2001; Tan & Teo, 2000). Khách hàng và nhà cung cấp có quyền yêu cầu đối tác của mình ứng dụng TMĐT để giúp họ thuận tiện hơn trong việc tìm kiếm thông tin, mua hàng và thanh toán.

Các yếu tố về đổi mới công nghệ:

Việc *nhận thức những lợi thế* của việc ứng dụng TMĐT trong doanh nghiệp là một tiền đề cơ bản giúp cho việc hội nhập được thực hiện nhanh chóng (Limthongchai & Speece, 2003; Seyal & Rahman, 2003). Lãnh đạo doanh nghiệp cần phải nhận thức được rằng TMĐT không chỉ mở ra những cơ hội kinh doanh mới, những sản phẩm và dịch vụ mới, những ngành nghề kinh doanh mới mà bản thân nó thực sự là một phương thức kinh doanh mới: Phương thức kinh doanh điện tử. Nó chuyển hóa các chức năng kinh doanh từ nghiên cứu thị trường, sản xuất sản phẩm đến bán hàng, dịch vụ sau bán từ phương thức kinh doanh truyền thống đến phương thức kinh doanh điện tử. Tuy nhiên, các nhà lãnh đạo cũng cần nhận thức được những *phức tạp* (Grover & Goslar, 1993; Seyal & Rahman, 2003) của việc

ứng dụng TMĐT trong doanh nghiệp và *sự tương hợp hay thích hợp* (Grover & Goslar, 1993; Teo & Tan, 1998) của phương thức kinh doanh này với hạ tầng công nghệ thông tin và văn hóa của doanh nghiệp. Chính việc quá thận trọng trong việc duy trì tính ổn định tại doanh nghiệp mà nhiều nhà lãnh đạo chưa có những quyết định xúc tiến việc ứng dụng TMĐT tại doanh nghiệp của mình. Để giải quyết vấn đề này, chính phủ cần phải có những chương trình đào tạo và bồi dưỡng cán bộ những hiểu biết về TMĐT cũng như những vấn đề kiểm soát những sự thay đổi khi doanh nghiệp khi ứng dụng TMĐT.

Biểu 2. Tóm tắt các chiều hướng tác động của nhân tố ảnh hưởng đến hội nhập TMĐT tại các doanh nghiệp vừa và nhỏ³

Biến số																	
		EC	EC	EC	EC	BE	EC	EC	EC	EC	EC	EC	EC	EC	EDI	EC	Mail
1	Nhận thức về TMĐT của CBCNV					ns				+					+		
2	Nguồn lực của doanh nghiệp		ns		+		+	+		+					+	+	
3	Định hướng chiến lược về TMĐT		+														
4	Quy mô doanh nghiệp			ns		ns	NS					+			+		+
5	Nhu cầu hội nhập của doanh nghiệp					+											
		To and Ngai (2007)	Valaseca-Requena, Torrent-Sellens, Messeguer-Artola, and Rodriguez-Ardura (2007)	Chuang, Nakatani, Chen, and Huang (2007)	Al-Qirim (2005)	Jeon, Han, and Lee (2006)	To and Ngai (2007)	(Al-Qirim, 2005; Hong & Zhu, 2006)	Looi (2005)	Molla and Licker (2005)	Dholakia and Kshetri (2004)	Seyal, Awais, Shamail, and Abbas (2004)	Seyal and Rahman (2003)	Scupola (2003)	Chwelos, Benbasat, and Dexter (2001)	Kendall et al. (2001)	Premkumar and Ramamurthy (1995)
																	Premkumar and Ramamurthy (1995)
																	Premkumar, Ramamurthy, and Nilakanta (1994)
																	Thong (1999)

³. Tài liệu phục vụ nghiên cứu: Lê Văn Huy (2005), Nhân tố tác động đến hội nhập thương mại điện tử trong các doanh nghiệp Việt Nam: Một nghiên cứu thực tiễn, *Tạp chí Nghiên cứu kinh tế, Viện Nghiên cứu kinh tế*, 323, 72-78.

6	Lĩnh vực kinh doanh			+		+													
7	Thái độ đối với đổi mới của nhà quản lý				+	+						ns					+	+	ns
8	Hiểu biết về CNTT và TMĐT của nhà quản lý				+	+													+
9	Sức ép cạnh tranh	+	ns		+	ns	+		+		+			+	+		+	+	ns
10	Sự hỗ trợ của ngành								+										
11	Sự hỗ trợ của chính phủ					+			+		+		+						
12	Hạ tầng quốc gia về CNTT								+		+								
13	Hành vi/ nhu cầu giao dịch TMĐT của khách hàng				+			+						+		+	+	+	
14	Cơ hội	+			+	+	+		+			ns	+		+	-	+	+	+
15	Sự thuận tiện	+			+	ns						+	+			-	+	ns	ns
16	Sự phức tạp					ns		ns				ns	ns			ns	+	ns	ns
17	Rủi ro							+	+								+		

Hệ thống ký hiệu (+):

EC: Thương mại điện tử (TMĐT);

BE: Kinh doanh điện tử;

(+): Quan hệ đồng biến;

(-) : Quan hệ nghịch biến;

(ns) : Không ý nghĩa.

[Biểu 3 sẽ tóm tắt kết quả nghiên cứu thực nghiệm các nhân tố tác động đến hội nhập thương mại điện tử tại các nước (đang phát triển trên thế giới), trong tổng hợp này chỉ ra hướng tác động làm cơ sở cho việc viết phần hợp tuyển lý thuyết và so sánh/ bình luận kết quả nghiên cứu ở phần sau]

[Biểu tổng hợp này thực sự cần thiết khi bình luận đóng góp nghiên cứu ở chương sau, nhà nghiên cứu có thể so sánh kết quả nghiên cứu của mình với các tác giả nghiên cứu trước, kết quả nghiên cứu (của tác giả) đã ủng hộ kết quả nghiên cứu của ai? Phản bác kết quả nghiên cứu của ai? Bình luận lí do... → Đây cũng là một trong những đóng góp về lý thuyết và thực tiễn của nghiên cứu].

Ví dụ 2: Nghiên cứu mối quan hệ giữa chất lượng dịch vụ và sự hài lòng của khách hàng tại các khách sạn 4-5 sao⁴ (phân liên quan đến thang đo)

[Ví dụ này liên quan đến việc tổng hợp các biến số quan sát hay chỉ báo (observed variables, items, indicators) nhằm làm cơ sở cho việc phát triển thang đo]

Chất lượng dịch vụ

Xây dựng và cải tiến chất lượng dịch vụ được xem như là một vấn đề hết sức quan trọng trong kinh doanh nói chung và kinh doanh khách sạn nói riêng nhằm gia tăng sự hài lòng và trung thành của khách hàng đối với doanh nghiệp. Chất lượng dịch vụ là nhân tố tạo sự khác biệt bên trong của việc kinh doanh khách sạn, cung cấp các dịch vụ đồng nhất nhằm đảm bảo chất lượng chung của khách hàng theo một tiêu chuẩn nhất định (hạng sao). Nhiều nghiên cứu xem xét các nhu cầu, đặc điểm của các du khách và cho rằng hầu hết du khách sẽ quan tâm đến các thuộc tính khách sạn khi đưa ra quyết định lựa chọn khách sạn như vị trí, giá phòng, an ninh, chất lượng dịch vụ và danh tiếng của khách sạn hoặc chuỗi khách sạn (Ananth và cộng sự, 1992; Atkinson, 1988; Clow và cộng sự, 1994; Le Blanc và Nguyen năm 1996; Mc Cleary và cộng sự, 1993; Rivers và cộng sự, 1991; Weaver và Heung, 1993; Wilensky và Buttle, 1998) **[Giới thiệu lý thuyết nền và các nghiên cứu liên quan]**. Trong đó, Atkinson (1998) cho rằng sự sạch sẽ của khách sạn, an toàn và an ninh, sự lịch thiệp và trợ giúp của nhân viên là các thuộc tính mà du khách chú ý trong việc lựa chọn khách sạn, Le Blanc và Nguyen (1996) thì khẳng định 5 yếu tố của khách sạn là môi trường vật chất, hình ảnh khách sạn, nhân viên phục vụ, chất lượng dịch vụ và khả năng tiếp cận là những tín hiệu trong việc lựa chọn khách sạn của khách hàng **[Giới thiệu sơ lược về chất lượng dịch vụ]**

Sự hài lòng của khách hàng

Theo Oliver (1997), sự hài lòng là phản ứng đáp lại của người tiêu dùng, Barsky (1995) định nghĩa sự hài lòng của khách hàng được xác định bằng cách xác định các cảm nhận của khách hàng về chất lượng. Nhìn chung, chất lượng dịch vụ thúc đẩy sự hài lòng của khách hàng, kích thích ý định quay trở lại và giới thiệu cho người khác về dịch vụ đã sử dụng (Hackl và Westlund, 2000).

Như vậy, trên cơ sở kết quả nghiên cứu hệ thống các mối quan hệ tuyến tính xuất phát từ những biến số khởi tạo dựa trên các nghiên cứu của Heung (2000), Ray Chu (2002), Thanika Devi Fuwaheer (2004) và ý kiến tham khảo của các nhân viên, quản lý khách sạn tại các khách sạn 4-5 sao ở Đà Nẵng, nghiên cứu đề nghị mô hình về các nhân tố như: nhân viên (Employees), phòng nghỉ (Room) của khách sạn, đồ ăn/ thức uống (Food and Beverage), các dịch vụ giá trị gia tăng (Value added service), giá cảm nhận (Perceived price), sự tiện lợi (Convenience), an ninh (Security) của khách sạn tác động đến sự hài lòng. Thang đo được sử dụng cho các khái niệm trong nghiên cứu được giới thiệu:

⁴. Tài liệu phục vụ nghiên cứu: Lê Văn Huy, Nguyễn Duy Quang (2011), Nghiên cứu các nhân tố tác động đến sự hài lòng của khách du lịch quốc tế đối với khách sạn 4-5 sao: Nghiên cứu thực tiễn tại Green Plaza – Đà Nẵng. *Tạp chí Phát triển kinh tế*.

Biểu 3. Thang đo chất lượng dịch vụ trong lĩnh vực khách sạn

TT	Factors (Biến tiềm ẩn)	Chỉ báo	Chỉ báo	Tác giả	Thang đo
1	Employees (Nhân viên)	E1	Nhân viên lễ tân đăng ký nhận phòng/ trả phòng cho khách nhanh chóng	Ray Chu	Likert 5
2		E2	Nhân viên khách sạn lịch sự	Ray Chu	Likert 5
3		E3	Nhân viên khách sạn thân thiện/nhiệt tình	Vincent C.S. Heung	Likert 5
4		E4	Nhân viên khách sạn ăn mặc lịch sự	Ray Chu	Likert 5
5		E5	Nhân viên khách sạn cung cấp dịch vụ đúng như cam kết (như chất lượng đã quy định)	Ray Chu	Likert 5
6		E6	Nhân viên khách sạn có thể giao tiếp với khách	Ray Chu	Likert 5
7		E7	Nhân viên khách sạn hiểu các yêu cầu của khách	Ray Chu	Likert 5
8	Room (Phòng nghỉ)	R1	Phòng nghỉ yên tĩnh	Ray Chu	Likert 5
9		R2	Phòng nghỉ sạch sẽ/vệ sinh	Ray Chu	Likert 5
10		R3	Gường/nệm/gối thỏa mái	Ray Chu	Likert 5
11		R4	Nhiệt độ phòng có chất lượng cao	Ray Chu	Likert 5
12		R5	Khách sạn có phòng tắm sạch sẽ, thoải mái	Thanika Devi Juwaheer	Likert 5
13		R6	Thiết bị nghe (nhìn) trong phòng (như TV, điện thoại,...) tốt	Vincent C.S. Heung	Likert 5
14		R7	Thiết bị khác trong phòng đảm bảo hoạt động tốt, đáng tin cậy (như: tủ lạnh, máy sấy tóc, điều hòa nhiệt độ, kết sắt...)	Al-Rousan, M.Ramzi, Badaruddin Mohamed	Likert 5
15	Food and beverage (Đồ ăn/thức uống)	FB1	Khách sạn đã cung cấp miễn phí (trà, cà phê, nước uống, ấm pha trà,...) trong phòng nghỉ	Thanika Devi Juwaheer	Likert 5
16		FB2	Đồ ăn/thức uống đa dạng	Ray Chu	Likert 5
17		FB3	Đồ ăn/thức uống hợp chất lượng	Vincent C.S. Heung/ Ray Chu	Likert 5
18		FB4	Thức ăn/đồ uống hợp vệ sinh	Thanika Devi Juwaheer	Likert 5
19		FB5	Khách sạn cung cấp đầy đủ thực đơn về chế độ ăn kiêng cho khách	Thanika Devi Juwaheer	Likert 5
20		FB6	Khách sạn cung cấp thực đơn đảm bảo sức khỏe	Vincent C.S. Heung	Likert 5
21	Value Added services (Các dịch vụ gia tăng)	VA1	Khu vui chơi giải trí trong khách sạn có chất lượng tốt	Al-Rousan, M.Ramzi, Badaruddin Mohamed	Likert 5

22	dịch vụ giá trị gia tăng)	VA2	Cuộc gọi báo thức là đáng tin cậy	Ray Chu	Likert 5
23		VA3	Dịch vụ giặt/ủi đảm bảo tốt	Ray Chu	Likert 5
24		VA4	Dịch vụ buồng phòng phục vụ tốt	Ray Chu	Likert 5
25		VA5	Các dụng cụ thể dục đảm bảo về tiêu chuẩn (như: máy chạy bộ, tạ, dụng cụ bơi lội,...)	Vincent C.S. Heung	Likert 5
26		VA6	Bàn thông tin cung cấp đầy đủ các thông tin theo nhu cầu của khách	Ray Chu	Likert 5
27	Percived price (Giá cảm nhận)	P1	Giá phòng nghỉ hợp lý	Ray Chu	Likert 5
28		P2	Đồ ăn/thức uống phù hợp với giá	Ray Chu	Likert 5
29		P3	Giá các dịch vụ khách phù hợp	Hesham Z.Al-Sabbahy, Yuksel Ekinci, Michael Riley	Likert 5
33	Convenience (Sự tiện ích)	C1	Khách sạn có hồ bơi hoạt động tốt	Lau Pei Mey, Abdolali Khatibi, Akbar and David Yong Gun Fie	Likert 5
34		C2	Các tiện ích chăm sóc cá nhân luôn sẵn có (như: massage, tắm hơi và câu lạc bộ sức khỏe,...)	Vincent C.S. Heung	Likert 5
35		C3	Vị trí của khách sạn thuận lợi, dễ dàng di chuyển	Vincent C.S. Heung	Likert 5
30	Security (An ninh)	S1	Khách sạn cung cấp môi trường thoải mái và an toàn	Ray Chu	Likert 5
31		S2	An ninh cá nhân là đáng tin cậy	Ray Chu	Likert 5
32		S3	Chuông báo cháy lớn là đáng tin cậy	Ray Chu	Likert 5
36	Sự hài lòng và trung thành của khách hàng (Customer loyalty and satisfaction)	CS1	Bạn hạnh phúc với các dịch vụ của khách sạn	Al-Rousan, M.Ramzi, Badaruddin Mohamed	Likert 5
37		CS2	Các kỳ vọng của bạn từ một khách sạn đã được đáp ứng	Al-Rousan, M.Ramzi, Badaruddin Mohamed	Likert 5
38		CS3	Các dịch vụ của khách sạn tốt hơn những nơi khác		Likert 5
39		CL4	Bạn hài lòng về chất lượng dịch vụ đã được cung ứng tại khách sạn	D.Selcen O.Aykac, Serkan Aydin, Metin Ates, Ayse Tansel Cetin	Likert 5
40		CL5	Bạn sẽ lưu trú tại khách sạn này khi có dịp quay lại Đà Nẵng		Likert 5
41		CL6	Bạn sẽ giới thiệu cho người thân, bạn bè về chất lượng dịch vụ của khách sạn này		Likert 5

[Đây là file các chỉ báo (items) của các khái niệm, để tiện cho việc điều chỉ và sử dụng thiết kế công cụ thu thập dữ liệu và định nghĩa các biến trong quá trình phân tích dữ liệu, nhà nghiên cứu nên trình bày nội dung của phần này trên file Excel]

PHÁT TRIỂN GIẢ THUYẾT

Sau khi thực hiện tổng quan về tài liệu, người nghiên cứu phát triển các giả thuyết nghiên cứu có thể kiểm chứng được (working hypothesis/ hypotheses). Giả thuyết được phát biểu để được kiểm chứng tính logic và thực chứng vì thế việc phát biểu giả thuyết nghiên cứu đặc biệt quan trọng khi nó là trọng tâm của vấn đề nghiên cứu. Giả thuyết nghiên cứu ảnh hưởng đến cách mà các kiểm định được thực hiện trong phân tích dữ liệu và yêu cầu về đặc điểm và chất lượng dữ liệu dùng để phân tích. Giả thuyết nghiên cứu cần phải cụ thể, rõ ràng và liên quan trực tiếp đến vấn đề nghiên cứu. Vai trò của giả thuyết nghiên cứu là giới hạn phạm vi nghiên cứu và hướng dẫn người nghiên cứu đi đúng đường. Giả thuyết còn giúp xác định loại dữ liệu và các loại phương pháp sử dụng để phân tích dữ liệu.

Những cách tiếp cận được sử dụng để phát triển giả thuyết:

- Thảo luận với đồng nghiệp và chuyên gia về vấn đề, xuất xứ của vấn đề, và những mục tiêu trong khi tìm kiếm giải pháp
- Kiểm tra dữ liệu và lưu trữ (nếu có) để phát hiện ra những xu hướng, sự khác thường hoặc những “manh mối” khác
- Đọc các nghiên cứu tương tự trong lĩnh vực đang xem xét hoặc các nghiên cứu cho vấn đề tương tự
- Điều tra cá nhân thăm dò (exploratory personal investigation) liên quan đến những phỏng vấn tại hiện trường trong phạm vi giới hạn với những bên, cá nhân có những điều kiện, mà theo nhà nghiên cứu, sẽ cung cấp một sự thấu hiểu tốt hơn về những khía cạnh thực tế của vấn đề.

Những giả thuyết là kết quả của việc suy nghĩ thấu đáo về chủ đề nghiên cứu, kiểm tra những dữ liệu có sẵn, và các tài liệu bao gồm những nghiên cứu liên quan và sự tư vấn của các chuyên gia và những bên liên quan.

Để phát biểu được giả, cần phải hiểu giả thuyết lý thuyết (propositions) và giả thuyết nghiên cứu (hypothesis).

Giả thuyết lý thuyết

Giả thuyết lý thuyết (propositions) là nhận định về hiện tượng được quan sát (khái niệm lý thuyết) mà có thể được đánh giá hoặc đúng hoặc sai. Khi giả thuyết lý thuyết được phát biểu để kiểm định bằng các dữ liệu thu thập được từ thực tế thì nó được gọi là giả thuyết nghiên cứu (hypothesis). Giả thuyết nghiên cứu là một tuyên bố về mối quan hệ giữa hai hoặc nhiều biến số, và nó mang tính thử nghiệm và phỏng đoán.

Các loại giả thuyết

Thông thường, có 3 cách phát triển giả thuyết nghiên cứu gồm:

Giả thuyết mô tả:

Giả thuyết mô tả có thể ở dạng một câu phát biểu hoặc có thể để nguyên ở dạng câu hỏi nghiên cứu.

Ví dụ:

Ở dạng câu hỏi nghiên cứu	Ở dạng giả thuyết mô tả
Thị phần của sản phẩm X ở Đà Nẵng là bao nhiêu?	Ở Đà Nẵng, thị phần sản phẩm X ở mức 15%.
Liệu người tiêu dùng ở Quảng Nam có thích trà túi có hương vị cam?	70% người tiêu dùng ở Quảng Nam thích trà túi có hương vị cam.

Giả thuyết tương quan

Giả thuyết tương quan diễn tả mối quan hệ giữa hai biến trong đó không ám chỉ sự thay đổi ở biến này gây ra sự thay đổi ở biến còn lại. Ví dụ:

- Phụ nữ trẻ tuổi (dưới 35 tuổi) mua trà túi Lipton ít hơn phụ nữ lớn tuổi (35 tuổi trở lên).
- Người tiêu dùng ở Thành phố Hồ Chí Minh thích trà túi có hương vị cam hơn người tiêu dùng ở Hà Nội.

Giả thuyết nhân quả

Giả thuyết nhân quả chỉ ra rằng sự tồn tại hay thay đổi ở biến này gây ra hoặc dẫn đến sự thay đổi ở biến kia, biến gây ra thay đổi gọi là biến độc lập, biến chịu ảnh hưởng gọi là biến phụ thuộc.

Ví dụ:

- Sự tăng lên trong thu nhập của hộ gia đình (biến độc lập) dẫn đến sự tăng lên trong tiết kiệm (biến phụ thuộc).
- Lòng trung thành đối với siêu thị Big C (biến độc lập) tăng xác suất mua sản phẩm mang thương hiệu Big C (biến phụ thuộc).

Những nguyên tắc hướng dẫn việc viết các câu hỏi và các giả thuyết trong nghiên cứu định lượng:

- Các biến được sử dụng trong các câu hỏi nghiên cứu hay các giả thuyết thường được giới hạn trong ba cách tiếp cận căn bản: (1) so sánh các nhóm, đối với một biến độc lập, để xem tác động của biến độc lập đối với một biến phụ thuộc; (2) thiết lập mối quan hệ giữa một hay nhiều hơn một biến độc lập với một biến phụ thuộc và (3) mô tả sự tác động của biến độc lập, biến trung gian hay biến phụ thuộc.
- Biến độc lập và biến phụ thuộc phải được đo lường riêng biệt.
- Rất nhiều nghiên cứu thường nêu cả câu hỏi nghiên cứu và giả thuyết nghiên cứu, tuy nhiên, chỉ viết câu hỏi nghiên cứu hoặc viết giả thuyết chứ không nên viết cả hai.

- **Mối quan hệ giữa các biến số: quan hệ tương quan và quan hệ nhân quả**

Mối quan hệ tương quan giữa hai biến số có thể được xem xét trong thiết kế nghiên cứu mô tả. Ở đây, nhà nghiên cứu quan tâm liệu hai biến số có thay đổi cùng nhau hay không.

Ví dụ:

- H_0 : Không có sự tương quan giữa điểm GMAT và kết quả học tập ở bậc cao học trên tổng thể.
- H_1 : Có sự tương quan giữa điểm GMAT và kết quả học tập ở bậc cao học trên tổng thể.

Mối quan hệ nhân quả như trình bày ở phần trên (nghiên cứu nhân quả) giữa hai biến số được phát biểu trong giả thuyết như sau:

- H_0 : Điểm GMAT không ảnh hưởng đến kết quả học tập ở bậc cao học.
- H_1 : Điểm GMAT có ảnh hưởng đến kết quả học tập ở bậc cao học.

- **Hướng của giả thuyết (direction of hypotheses)**

Khi nhà nghiên cứu không có thông tin để xác định hướng của mối quan hệ/đặc điểm của tổng thể được quan sát, giả thuyết được phát biểu mà không đề cập đến hướng của mối quan hệ (two-tailed hypotheses).

Ví dụ 1:

- H_0 : Không có sự khác biệt trong số lượng mua sản phẩm A trước và sau khi sử dụng quảng cáo mới trên tổng thể.
- H_1 : Có sự khác biệt trong số lượng mua sản phẩm A trước và sau khi sử dụng quảng cáo mới trên tổng thể.

Nếu nhà nghiên cứu có được thông tin về hướng của mối quan hệ (cùng chiều hay trái chiều, đồng biến hay nghịch biến), giả thuyết bao gồm hướng của mối quan hệ.

Ví dụ 2:

- H_0 : Không có sự tăng lên trong số lượng mua sản phẩm A so với trước khi sử dụng quảng cáo mới.
- H_1 : Có sự tăng lên trong số lượng mua sản phẩm A so với trước khi sử dụng quảng cáo mới.

Ví dụ 3:

- H_0 : Không tồn tại sự tương quan đồng biến giữa điểm thi GMAT và kết quả học tập ở bậc cao học.
- H_1 : Tồn tại sự tương quan đồng biến giữa điểm thi GMAT và kết quả học tập ở bậc

cao học.

Ví dụ 4:

- H_0 : Không có sự tương quan giữa tần suất nghỉ học và kết quả học tập.
- H_1 : Tồn tại sự tương quan nghịch biến giữa tần suất nghỉ học và kết quả học tập.

- Có hai hình thức phát biểu giả thuyết:

- Phát biểu giả thuyết không (null hypothesis) và giả thuyết đối (alternative hypothesis): Giả thuyết không (H_0) thường tiên đoán rằng **không** tồn tại mối quan hệ hoặc **không** tồn tại sự khác biệt giữa các biến trên tổng thể. Giả thuyết đối (H_1) thể hiện **có** tồn tại mối quan hệ giữa hai biến trên tổng thể.

Ví dụ:

- H_0 : Không tồn tại mối quan hệ giữa bản chất công việc với sự hài lòng của nhân viên đối với tổ chức trên tổng thể.
- H_1 : Tồn tại mối quan hệ giữa bản chất công việc với sự hài lòng của nhân viên đối với tổ chức trên tổng thể.

Hoặc:

- H_0 : Lương trung bình của nam cán bộ công nhân viên bằng với lương trung bình của nữ trên tổng thể.
- H_1 : Lương trung bình của nam cán bộ công nhân viên khác với lương trung bình của nữ trên tổng thể.
- Chỉ phát biểu giả thuyết thay thế: Hình thức này thường được sử dụng trình bày trong các bài báo, hay các luận án. Nhà nghiên cứu đưa ra tiên đoán về kết quả kỳ vọng cho tổng thể của nghiên cứu. Những tiên đoán này thường xuất phát từ tài liệu và những công trình nghiên cứu trước đó về đề tài đang xét mà tài liệu và những nghiên cứu đó gợi ý kết quả tiềm năng mà nhà nghiên cứu có thể kỳ vọng.

Ví dụ:

- Tồn tại mối quan hệ giữa bản chất công việc với sự hài lòng của nhân viên đối với tổ chức trên tổng thể.
- Tồn tại mối quan hệ đồng biến giữa sự hài lòng và lòng trung thành của nhân viên đối với tổ chức trên tổng thể.

Đối với phân tích mối quan hệ hay phân tích nhân quả, thông qua kết quả nghiên cứu lý thuyết, nhà nghiên cứu thường hình thành **mô hình nghiên cứu lý thuyết**. Mô hình nghiên cứu lý thuyết bao thể hiện các khái niệm (biến số) và mối quan hệ giữa các khái niệm (dựa vào đó để phát biểu các giả thuyết).

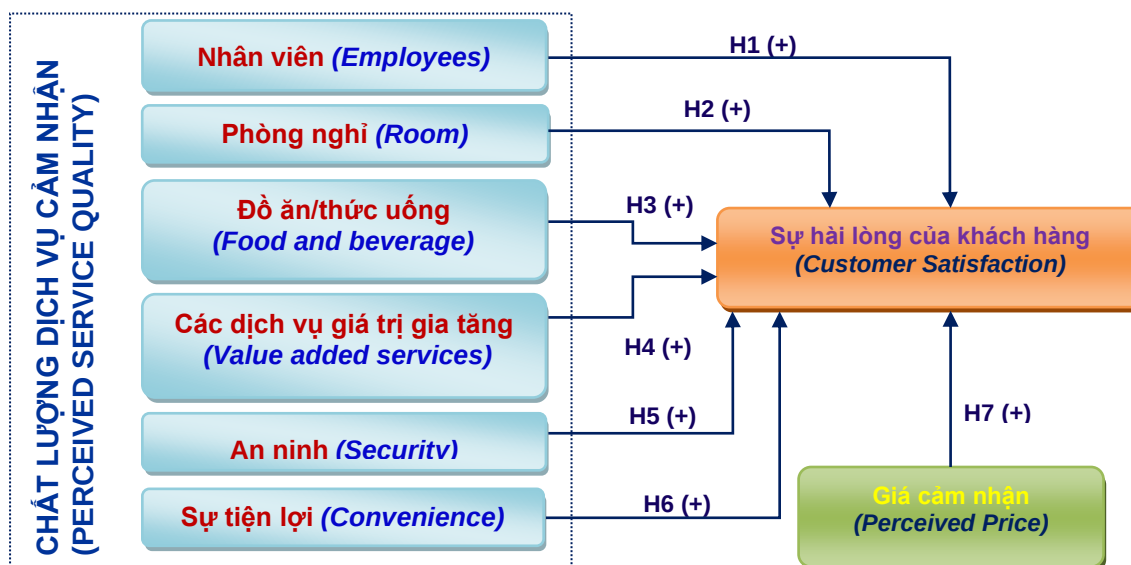
Thông thường, một nghiên cứu có thể chỉ phát biểu các giả thuyết hoặc vừa hình thành mô hình

nghiên cứu lý thuyết (để gia tăng tính trực quan) và trên cơ sở đó, phát biểu các giả thuyết liên quan.

Ví dụ: Mô hình lý thuyết nghiên cứu mối quan hệ giữa chất lượng dịch vụ và sự hài lòng của khách hàng lịch quốc tế đến khách sạn 4-5 sao⁵

Trên cơ sở kết quả nghiên cứu hệ thống các mối quan hệ tuyến tính xuất phát từ những biến số khởi tạo dựa trên các nghiên cứu của Heung (2000), Ray Chu (2002), Thanika Devi Fuwaheer (2004) và nghiên cứu định tính, nghiên cứu đề nghị mô hình về các nhân tố như: nhân viên (Employees), phòng nghỉ (Room) của khách sạn, đồ ăn/Thức uống (Food and Beverage), các dịch vụ gia trị gia tăng (Value added service), giá cảm nhận (Perceived price), sự tiện lợi (Convenience), an ninh (Security) của khách sạn tác động đến sự hài lòng. Trên cơ sở các biến số đo lường, mô hình nghiên cứu được đề nghị như sau:

Hình 7. Mô hình nghiên cứu mối quan hệ giữa chất lượng dịch vụ và sự hài lòng của khách du lịch quốc tế đến khách sạn 4-5 sao



Giả thuyết H1: Tồn tại mối quan hệ đồng biến giữa cảm nhận của khách hàng về cung cách phục vụ của nhân viên khách sạn và mức độ hài lòng của khách hàng đối với dịch vụ đó;

Giả thuyết H2: Tồn tại mối quan hệ đồng biến giữa cảm nhận của khách hàng về chất lượng của phòng nghỉ và mức độ hài lòng của khách hàng đối với dịch vụ đó;

Giả thuyết H3: Cảm nhận của khách hàng về đồ ăn/thức uống càng cao thì mức độ hài lòng của họ đối với dịch vụ đó càng cao;

Giả thuyết H4: Tồn tại mối quan hệ giữa cảm nhận của khách hàng về các dịch vụ giá trị và mức độ hài lòng của họ đối với các dịch vụ đó;...

⁵. Tài liệu phục vụ nghiên cứu: Lê Văn Huy, Nguyễn Duy Quang (2011), Nghiên cứu các nhân tố tác động đến sự hài lòng của khách du lịch quốc tế đối với khách sạn 4-5 sao: Nghiên cứu thực tiễn tại Green Plaza – Đà Nẵng. *Tạp chí Phát triển kinh tế*.

Trên cơ sở các giả thuyết đã phát biểu, nhiệm vụ của nhà nghiên cứu trong giai đoạn tiếp theo là thu thập dữ liệu để kiểm định các giả thuyết được phát biểu. Như vậy, có thể nói giả thuyết sẽ thể hiện mối quan hệ giữa các khái niệm trong lý thuyết khoa học và định hướng cho việc lựa chọn các kỹ thuật nhằm kiểm định mối quan hệ trên.

TÓM TẮT CHƯƠNG 2

Trong chương 1 đã giới thiệu tiến trình nghiên cứu khoa học, chương 2 giới thiệu bước đầu tiên trong tiến trình nghiên cứu là xác định vấn đề nghiên cứu trong mối quan hệ với bước 2 (giới thiệu tổng quan/ hợp tuyển lý thuyết) và phát triển giả thuyết.

Có hai loại vấn đề nghiên cứu: loại vấn đề liên quan đến bản chất của hiện tượng và những loại vấn đề liên quan đến mối quan hệ giữa các biến số nghiên cứu. Ban đầu vấn đề nghiên cứu có thể được phát biểu chung chung.

Vấn đề nghiên cứu xuất phát từ những vấn đề gặp phải trong thực tế mà chưa tìm ra lời giải thích thích hợp từ các lý thuyết đã có hoặc những “lỗ hổng” (gap) trong các lý thuyết hiện tại. Người nghiên cứu phải làm quen với các lý thuyết, tài liệu liên quan đến vấn đề nghiên cứu được lựa chọn. Người nghiên cứu cần tổng hợp/ xem xét (review) hai loại tài liệu, (1) tài liệu liên quan đến các khái niệm (concepts), lý thuyết và (2) những tài liệu là những nghiên cứu thực chứng (empirical literature) được thực hiện trước đây liên quan đến vấn đề nghiên cứu đề xuất.

Một số kỹ thuật để xác định vấn đề nghiên cứu gồm (1) Khảo sát điểm mạnh của nhà nghiên cứu và sự hiểu biết về bản chất của vấn đề, (2) Xem xét lại và phát hiện vấn đề nghiên cứu từ các nghiên cứu có sẵn, (3) Phát triển ý tưởng thông qua thảo luận, (4) Xem xét dữ liệu thứ cấp và (5) Tập kích não (brainstorming).

CHƯƠNG

3

THIẾT KẾ NGHIÊN CỨU (RESEARCH DESIGN)

Trong chương 2 đã giới thiệu những vấn đề cơ bản liên quan đến nghiên cứu khoa học, so sánh sự khác nhau giữa kỹ thuật và phương pháp nghiên cứu, các triết lý và cách tiếp cận nghiên cứu. Trong chương 3 sẽ cung cấp định nghĩa, cách thức phân loại và kỹ thuật thiết kế nghiên cứu.

ĐỊNH NGHĨA THIẾT KẾ NGHIÊN CỨU

Trong nghiên cứu, vấn đề khó khăn đặt ra là việc xác định được vấn đề nghiên cứu cần phải xác định một cách rõ ràng, các quyết định liên quan đến cái gì, ở đâu, khi nào, bao nhiêu, bởi cái gì liên quan tới một cuộc điều tra hoặc một phác thảo nghiên cứu tạo thành một thiết kế nghiên cứu. Thiết kế nghiên cứu là sắp xếp các điều kiện cho việc thu thập và phân tích dữ liệu theo một cách thức nhằm kết nối thích hợp với mục đích nghiên cứu (Kothari, 2004). Trong thực tế, thiết kế nghiên cứu là khái niệm cấu trúc trong tiến hành nghiên cứu; nó cấu thành kế hoạch chi tiết cho việc thu thập, đo lường và phân tích dữ liệu.

Thiết kế nghiên cứu bao gồm một phác thảo những gì người nghiên cứu sẽ làm từ việc đưa ra cách thức xây dựng công cụ thu thập dữ liệu, phương pháp thu thập, thiết kế mẫu, mô hình phân tích, giả thuyết, tổ chức thu thập dữ liệu cho đến phương pháp phân tích dữ liệu. Theo Cooper và Schindler (2011), thiết kế nghiên cứu là bản kế hoạch (blueprint) cho thấy cách nhà nghiên cứu trả lời các câu hỏi nghiên cứu như thế nào thông qua:

- Xác định vấn đề nghiên cứu thông trên cơ sở phương pháp tiếp cận nghiên cứu (thăm dò, định tính hay định lượng);
- Đo lường và xác định thang đo đối với các khái niệm nghiên cứu và biến số nghiên cứu;
- Lựa chọn phương pháp chọn mẫu và qui mô mẫu;
- Lựa chọn công cụ và phương pháp phân tích dữ liệu.

Những yếu tố quan trọng trong thiết kế nghiên cứu:

- Luôn dựa trên câu hỏi nghiên cứu;
- Là hướng dẫn cho việc chọn nguồn thông tin và loại thông tin;
- Là “khung” (framework) cho việc cụ thể hóa mối quan hệ giữa các biến trong nghiên cứu;
- Là bản hướng dẫn về qui trình từ phát triển giả thuyết đến phân tích dữ liệu.

SỰ CẦN THIẾT CỦA THIẾT KẾ NGHIÊN CỨU

Việc thiết kế nghiên cứu là cần thiết bởi nó tạo điều kiện cho sự khởi đầu thuận lợi của các hoạt động nghiên cứu khác nhau, làm nghiên cứu đạt hiệu quả cao nhất có thể mà không tiêu tốn nhiều thời gian, công sức và tiền bạc nhưng vẫn thu được lượng thông tin tối đa. Cũng như việc xây dựng một công trình, chúng ta cần một bản thiết kế tốt được xây dựng bởi một nhà thiết kế kinh nghiệm. Thiết kế nghiên cứu phác thảo các phương pháp được thông qua trước cho việc thu thập dữ liệu liên quan và các kỹ thuật được sử dụng trong phân tích của nhà nghiên cứu, theo sát mục tiêu nghiên cứu, sẵn sàng về nhân lực, thời gian và chi phí. Sự chuẩn bị của thiết kế nghiên cứu nên được thực hiện một cách cẩn thận vì bất kì lỗi nào cũng có thể gây ảnh hưởng đến toàn bộ dự án. Thiết kế nghiên cứu, trong thực tế, có ảnh hưởng rất lớn đến độ tin cậy của các kết quả và tạo thành nền tảng vững chắc cho toàn bộ nghiên cứu. Thiết kế nghiên cứu giúp cho nhà nghiên cứu và các thành viên tham gia biết được mục tiêu, định hướng được các công việc cần phải thực hiện. Một cách cụ thể, thiết kế nghiên cứu cần phải quyết định:

- Nghiên cứu về cái gì?
- Tại sao nghiên cứu được thực hiện?
- Loại dữ liệu nào được yêu cầu?
- Dữ liệu yêu cầu được tìm thấy ở đâu?
- Nghiên cứu bao gồm giai đoạn thời gian nào?
- Thiết kế mẫu sẽ là gì?
- Kỹ thuật thu thập dữ liệu nào được sử dụng?
- Dữ liệu sẽ được phân tích bằng cách nào?
- Nghiên cứu sẽ được ứng dụng ở đâu?
- Báo cáo được chuẩn bị sẽ theo kiểu nào?

Điều bắt buộc là một bản thiết kế hiệu quả và thích hợp phải được chuẩn bị trước khi bắt đầu hoạt động nghiên cứu, thiết kế giúp người nghiên cứu tổ chức các ý tưởng của mình theo một hình thức mà theo đó họ dễ dàng tìm ra chỗ sai sót và không hợp lý, thiết kế nghiên cứu có thể được đưa cho người khác để lấy ý kiến và thu nhập những ý kiến góp ý để giảm bớt những sai sót và tạo sự đồng thuận trong tập thể nghiên cứu

CÁC ĐẶC ĐIỂM CỦA MỘT THIẾT KẾ TỐT

Một thiết kế tốt thường được đặc trưng bởi các đặc tính như tính linh động, thích ứng, hiệu quả, kinh tế... Đặc biệt, một thiết kế giảm thiểu độ sai lệch và tối đa hóa độ tin cậy của dữ liệu được thu thập và phân tích thì được xem là một thiết kế tốt. Để hạn chế những sai sót có thể xảy ra khi thực hiện chính thức, nhà nghiên cứu cần thiết kế thử nghiệm trên mẫu nhỏ. Một thiết kế nghiên cứu thích hợp cho một vấn đề nghiên cứu cụ thể, thường liên quan đến việc xem xét các yếu tố sau:

- Các phương tiện thu được thông tin;
- Sự sẵn có các kỹ năng của nhà nghiên cứu và nhân viên của mình (nếu có);
- Sự khách quan của vấn đề được nghiên cứu;
- Bản chất của vấn đề được nghiên cứu;
- Sự sẵn sàng về thời gian và chi phí cho công việc nghiên cứu.

Nếu công trình nghiên cứu là một nghiên cứu thăm dò, cần nhấn mạnh việc phát hiện ý tưởng, thiết kế nghiên cứu thích hợp nhất phải đủ linh hoạt để cho phép xem xét nhiều khía cạnh khác nhau của một hiện tượng. Khi mục đích của nghiên cứu là mô tả, cần mô tả chính xác về một tình huống hoặc một mối quan hệ giữa các biến và khi đó, một thiết kế nghiên cứu giảm thiểu sai lệch và tối đa hóa độ tin cậy trong việc thu thập dữ liệu được coi là một thiết kế tốt. Đối với thiết kế nghiên cứu nhân quả, việc kiểm tra mối quan hệ nhân quả giữa các biến đòi hỏi một thiết kế cho phép kết luận về quan hệ nhân quả trong điều kiện nghiên cứu, đặc biệt chú ý đến mối quan hệ giữa các biến số (đảm bảo mối quan hệ nhân quả) và các điều kiện thử nghiệm để đảm bảo độ tin cậy của kết quả nghiên cứu.

PHÂN LOẠI THIẾT KẾ NGHIÊN CỨU

Cooper và Schindler (2011) sử dụng những vấn đề sau của thiết kế nghiên cứu để phân loại các thiết kế nghiên cứu:

Căn cứ vào mức độ cụ thể của câu hỏi nghiên cứu

Một nghiên cứu có thể được xem là khám phá (exploratory) hay chính thức (formal). Điểm phân biệt chính giữa hai loại nghiên cứu là mức độ có cấu trúc và cụ thể của câu hỏi nghiên cứu. **Nghiên cứu khám phá** có khuynh hướng đáp ứng mục tiêu định hướng được những hoạt động nghiên cứu sau này. Mục đích của nghiên cứu khám phá là phát triển được các giả thuyết nghiên cứu hoặc các câu hỏi nghiên cứu cho những nghiên cứu sâu hơn. **Nghiên cứu chính thức** bắt đầu với mô tả cơ cấu liên quan (nghiên cứu mô tả) hoặc hình thành các giả thuyết (câu hỏi nghiên cứu) và liên quan đến qui trình, nguồn dữ liệu cụ thể nhằm mục đích kiểm định các giả thuyết hoặc trả lời các câu hỏi nghiên cứu (nghiên cứu nhân quả).

Căn cứ vào phương pháp thu thập dữ liệu

Sự phân biệt giữa qui trình theo dõi (monitoring) và truyền đạt/truyền thông (communication) để thu thập dữ liệu. Theo dõi bao gồm những nghiên cứu trong đó người nghiên cứu tìm hiểu/ điều tra những hoạt động của đối tượng mà không đòi hỏi phản ứng nào từ đối tượng. Ví dụ: theo dõi/ quan sát để đếm số người đi bộ qua đường đúng làn đường cho người đi bộ tại một ngã tư. Trong trường hợp này, người nghiên cứu phải ghi lại thông tin từ quan sát.

Trong các nghiên cứu sử dụng phương pháp truyền đạt (communication) để thu thập dữ liệu, người nghiên cứu hỏi và thu thập câu trả lời qua các hình thức như phỏng vấn trực tiếp qua điện thoại, đáp viên tự trả lời bản câu hỏi.

Căn cứ vào năng lực ảnh hưởng đến các biến số của người nghiên cứu

Ở khía cạnh này, thiết kế nghiên cứu được phân biệt giữa thực nghiệm (experimental design) và dựa trên thiết kế dựa trên thực tế đã có (ex post facto design). Trong thực nghiệm, người nghiên cứu kiểm soát hoặc điều chỉnh các biến số trong nghiên cứu. Thiết kế nghiên cứu thực nghiệm thích hợp khi người nghiên cứu muốn khám phá liệu có mối quan hệ nhân quả giữa các biến (những biến này có gây ảnh hưởng lên những biến khác). Thiết kế nghiên cứu thực nghiệm cung cấp những chứng cứ “xác thực/ hùng hồn” cho các giả thuyết về nhân quả. Trong thiết kế dựa trên thực tế đã có (ex post facto design), người nghiên cứu không kiểm soát các biến số hay điều chỉnh các biến số, họ chỉ có thể trình bày/ báo cáo những gì đã hoặc đang xảy ra.

Căn cứ vào mục đích của nghiên cứu

Tùy theo mục đích nghiên cứu, các thiết kế nghiên cứu được phân theo các nhóm sau: báo cáo (reporting), mô tả (descriptive), giải thích về nhân quả (causal-explanatory) và dự đoán về nhân quả (causal-predictive).

Nghiên cứu báo cáo (reporting study) cung cấp bản tổng hợp dữ liệu, thông thường điều chỉnh (recast) dữ liệu để hiểu hoặc đưa ra các con số thống kê cho việc so sánh. Ví dụ: Tổng hợp doanh số bán hàng tại thị trường Thành phố Hồ Chí Minh và Hà Nội.

Nếu nghiên cứu quan tâm đến các câu trả lời cho câu hỏi ai?, cái gì?, ở đâu?, khi nào? hoặc bao nhiêu?, thì nghiên cứu đó là nghiên cứu mô tả. Ví dụ, nghiên cứu mô tả về giới tính, độ tuổi, bộ phận làm việc của các bộ công nhân viên, có bao nhiêu người đã nhận được danh hiệu thi đua trong thời gian qua hoặc những người nào ủng hộ việc ứng dụng phương pháp quản lý mới trong doanh nghiệp...

Nếu nghiên cứu quan tâm đến việc trả lời câu hỏi - một biến số sẽ tạo ra thay đổi trên một biến số khác như thế nào (*how one variable produces changes in another*) - đó là nghiên cứu giải thích về nhân quả (**causal-explanatory**). Trong nghiên cứu giải thích về nhân quả, người nghiên cứu cố gắng giải thích mối quan hệ giữa các biến. Ví dụ tại sao nam giới thường ít trung thành với doanh nghiệp hơn nữ giới; phong cách lãnh đạo ảnh hưởng như thế nào đến sự hài lòng của nhân viên đối với tổ chức; việc thay đổi chiến lược quản lý ảnh hưởng như thế nào đến thái độ làm việc của nhân viên...

Nghiên cứu dự đoán về nhân quả (**causal-predictive**) dự đoán hiệu ứng/ tác động lên một biến số bằng cách điều chỉnh/ thay đổi biến số khác trong khi các biến số còn lại không đổi. Ví dụ, người nghiên cứu sử dụng nghiên cứu dự đoán về nhân quả bằng cách thay đổi chiến lược làm việc để xem xét sự thay đổi như thế nào trong thái độ, cách thức làm việc của nhân viên.

Căn cứ vào phương diện thời gian

Nghiên cứu tại một thời điểm (cross - sectional studies) đối với các nghiên cứu được thực hiện và vẽ một bức tranh tổng thể tại một thời điểm nhất định.

Nghiên cứu dài hạn (longitudinal studies) là nghiên cứu được lặp đi lặp lại trong một thời gian

dài. Thuận lợi của nghiên cứu dài hạn là có thể theo dõi và điều chỉnh theo những thay đổi theo thời gian.

Căn cứ vào phạm vi đề tài

Nghiên cứu thống kê (statistical studies) được thiết kế vì bề rộng hơn là chiều sâu của nghiên cứu. Người nghiên cứu cố gắng nắm những đặc điểm của tổng thể bằng cách suy từ những đặc điểm của mẫu. Những giả thuyết được kiểm định một cách định lượng. Suy rộng các kết quả nghiên cứu dựa trên tính đại diện của mẫu và giá trị của thiết kế.

Nghiên cứu tình huống (case studies) nhấn mạnh vào phân tích bối cảnh đầy đủ của một số ít các sự kiện/hiện tượng hoặc điều kiện và mối quan hệ qua lại giữa chúng. Nghiên cứu tình huống được sử dụng để cung cấp sự am hiểu về vấn đề, cách giải quyết vấn đề, đánh giá, và hình thành chiến lược. Nghiên cứu tình huống còn giúp người nghiên cứu có cái nhìn phản biện đối với các lý thuyết hiện có và hình thành những giả thuyết mới.

Căn cứ vào môi trường nghiên cứu

Thiết kế nghiên cứu còn khác nhau ở phương diện môi trường trong đó nghiên cứu được thực hiện. Có ba loại chính sau: các nghiên cứu được thực hiện trong môi trường thực tế (field conditions) hoặc các nghiên cứu được thực hiện trong các điều kiện có điều chỉnh (**laboratory conditions**), và các nghiên cứu được thực hiện trong môi trường mô phỏng trong đó có những điều kiện giống thực tế (**simulation**).

Căn cứ vào nhận thức của người tham gia

Nhận thức của người tham gia ảnh hưởng đến kết quả của nghiên cứu. Trong nhiều nghiên cứu, tính hữu dụng của thiết kế nghiên cứu giảm đi khi người tham gia nhận thức rằng nghiên cứu đang được thực hiện. Ví dụ nghiên cứu quan sát những người mua hàng trong các cửa hàng bán lẻ/siêu thị (mystery shopper) để hiểu hành vi của khách hàng và đánh giá cách trưng bày trong cửa hàng đòi hỏi những khách hàng được quan sát không hề biết rằng mình đang tham gia vào nghiên cứu.

THIẾT KẾ NGHIÊN CỨU: NGHIÊN CỨU THĂM DÒ, MÔ TẢ VÀ NHÂN QUẢ

Như đã giới thiệu ở trên, có nhiều cách tiếp cận nhằm phân loại thiết kế nghiên cứu và tùy vào mỗi tình huống nhất định mà người nghiên cứu lựa chọn cách tiếp cận nghiên cứu phù hợp. Mỗi cách tiếp cận nghiên cứu sẽ hướng dẫn những bước sau của thiết kế nghiên cứu như nguồn dữ liệu, đo lường, chọn mẫu và công cụ để phân tích dữ liệu.

Trong phần này sẽ giới thiệu 3 cách thiết kế nghiên cứu dựa vào mục tiêu của nghiên cứu là (1) nhận diện và xác định vấn đề; (2) mô tả vấn đề đã được xác định và (3) phát hiện mối quan hệ giữa các biến số trong vấn đề nghiên cứu để đề ra giải pháp giải quyết. Phần này sẽ giới thiệu ba cách tiếp cận nghiên cứu để đáp ứng 3 mục tiêu đó là: nghiên cứu thăm dò/ khám phá, nghiên cứu mô tả và nghiên cứu nhân quả.

Nghiên cứu thăm dò/ khám phá

Nghiên cứu thăm dò đặc biệt hữu ích khi người nghiên cứu chưa có ý tưởng rõ ràng về vấn đề mà họ gặp phải hoặc vấn đề nghiên cứu. Thông qua nghiên cứu thăm dò, người nghiên cứu phát triển các khái niệm (concepts) rõ ràng hơn, thiết lập các ưu tiên, phát triển các định nghĩa đưa vào nghiên cứu (operational definitions) và “cải thiện” thiết kế nghiên cứu cuối cùng. Nghiên cứu thăm dò cũng sẽ tiết kiệm thời gian và nguồn lực cho nhà nghiên cứu, đặc biệt trong các lĩnh vực nghiên cứu còn mới, khi các biến số nghiên cứu quan trọng, vẫn chưa được hiểu và xác định rõ. Điều này có nghĩa rằng chưa có gì nhiều được viết ra về đề tài hay tổng thể đang được nghiên cứu và nhà nghiên cứu tìm cách lắng nghe những người tham gia trong nghiên cứu và xây dựng sự hiểu biết dựa trên các ý tưởng của họ. Bên cạnh đó, thông qua nghiên cứu thăm dò, nhà nghiên cứu còn ước tính được tính thực tế của nghiên cứu chính thức (formal) sau này. Kết quả của nghiên cứu thăm dò là các khái niệm, các câu hỏi nghiên cứu và các giả thuyết nghiên cứu. Các nghiên cứu định lượng và định tính có thể được sử dụng trong nghiên cứu thăm dò, tuy nhiên nghiên cứu định tính thường được sử dụng phổ biến hơn. Các mục đích cụ thể của nghiên cứu này như sau:

- Hình thành và định nghĩa vấn đề thật chính xác;
- Đưa ra những phương án hành động;
- Kiểm định các giả thuyết;
- Rút gọn, bóc tách những nhân tố quan trọng và các mối quan hệ để nghiên cứu xa hơn;
- Đạt được sự hiểu biết sâu hơn trong tiếp cận vấn đề;
- Thiết lập thứ tự ưu tiên cho nghiên cứu sâu hơn.

Trong nghiên cứu định tính, chưa có sự phân biệt một cách hoàn chỉnh về các loại phương pháp và công cụ (Marshall và Rossman, 1999). (Thọ, 2011) chia làm hai nhóm: nhóm phương pháp và nhóm công cụ. Nhóm phương pháp gồm có: phương pháp lý thuyết nền (grounded theory), phương pháp tình huống (case studies) và phương pháp ethnography. Nhóm công cụ gồm các công cụ thu thập dữ liệu trong nghiên cứu định tính: phỏng vấn sâu (in-depth interview), thảo luận nhóm mục tiêu (focus group), quan sát (observations). Lưu ý, quan sát có thể dùng như công cụ thu thập dữ liệu định lượng và một số “công cụ đôi khi được xem như là phương pháp”. Ví dụ, Kitzinger (1994) sử dụng thảo luận nhóm như là một phương pháp nghiên cứu, một số tác giả khác thì cho rằng “thu thập dữ liệu bằng các tình huống cụ thể” và xem phương pháp tình huống (case studies) như một công cụ.

Theo (Kothari, 2004), để thực hiện nghiên cứu thăm dò, có thể thực hiện thông qua: (a) khảo sát lý thuyết, (b) khảo sát kinh nghiệm và (c) phân tích các ví dụ về tác nhân kích thích.

- Khảo sát lý thuyết là phương pháp hiệu quả và đơn giản nhất để xây dựng các vấn đề nghiên cứu một cách chính xác hoặc phát triển giả thuyết. Giả thuyết nêu ra bởi các nhà nghiên cứu trước đó có thể được xem xét lại tính hữu dụng của chúng và được đánh giá như là cơ sở cho nghiên cứu sâu hơn. Nó cũng có thể được xem xét liệu có hay không khi đưa ra giả thuyết mới. Bằng cách này, người nghiên cứu nên xem xét và xây dựng dựa trên

những nghiên cứu đã được thực hiện bởi những nhà nghiên cứu khác, nhưng trong trường hợp giả thuyết chưa được xác định, nhiệm vụ của ta là xem xét các tài liệu sẵn có để đề xuất những giả thuyết có liên quan từ cơ sở lý luận trên.

Việc xem xét lại tài liệu cung cấp một khuôn khổ để xác lập tầm quan trọng của công trình nghiên cứu đang xét cũng như một chuẩn mực để so sánh các kết quả của một công trình nghiên cứu với các kết quả tìm thấy khác. Nhà nghiên cứu ngoài việc đưa vào ở phần lý thuyết nhằm chứng minh tính hợp lý của các khái niệm (dựa trên lý thuyết nền), tính liên kết của các lý thuyết còn có thể đưa tài liệu liên quan vào phần cuối của công trình nghiên cứu, ở đó tài liệu được sử dụng để so sánh và đối chiếu với các kết quả (hay các chủ đề hay các phạm trù) mới nổi lên từ công trình nghiên cứu này.

- Khảo sát kinh nghiệm nghĩa là khảo sát người đã trải nghiệm thực tế với vấn đề được nghiên cứu. Các đối tượng của một khảo sát như vậy để nhận được cái nhìn sâu sắc về mối quan hệ giữa các biến và các ý tưởng mới liên quan đến vấn đề nghiên cứu. Đối với người khảo sát có năng lực và có thể đóng góp ý tưởng mới có thể được lựa chọn cẩn thận để đảm bảo tính đại diện của các loại trải nghiệm khác nhau. Người trả lời được lựa chọn để có thể phỏng vấn bởi điều tra viên. Người nghiên cứu phải chuẩn bị một lịch trình phỏng vấn cho việc hỏi thông tin có hệ thống nhưng cuộc phỏng vấn phải đảm bảo tính linh hoạt trong nên có thể vấn đề và câu hỏi mà điều tra viên đã không chuẩn bị trước. Việc phỏng vấn thu thập trải nghiệm thường kéo dài vài giờ, do vậy, có thể gửi trước một bản sao các câu hỏi để người trả lời suy nghĩ trước về toàn bộ vấn đề và họ có thể đóng góp hiệu quả hơn. Vì vậy, một khảo sát kinh nghiệm có thể cho phép người nghiên cứu xác định vấn đề chính xác hơn và giúp xây dựng giả thuyết nghiên cứu.
- Phân tích các ví dụ “hiểu rõ tác nhân kích thích” cũng là một phương pháp hiệu quả để đề xuất các giả thuyết cho nghiên cứu. Nó đặc biệt thích hợp trong các lĩnh vực có ít kinh nghiệm để đưa ra các chỉ dẫn. Phương pháp này bao gồm các nghiên cứu chuyên sâu, các ví dụ được lựa chọn của hiện tượng trong vấn đề nghiên cứu quan tâm. Với mục đích này các tài liệu sẵn có, nếu có, có thể được kiểm tra, có thể thực hiện phỏng vấn phi cấu trúc hoặc thông qua 1 số phương pháp khác. Thái độ của điều tra viên, cường độ của nghiên cứu và khả năng của người nghiên cứu sẽ đưa ra thông tin đa dạng tập hợp vào một giải thích thống nhất là những đặc điểm chính làm cho phương pháp này phù hợp để đưa ra các hiểu biết.

Nghiên cứu mô tả

Khác với nghiên cứu thăm dò, nghiên cứu mô tả là những nghiên cứu chính thức và có cấu trúc hơn với những giả thuyết được phát biểu rõ ràng. Sử dụng nghiên cứu mô tả nhằm đạt được những mục tiêu sau:

- Mô tả về hiện tượng hoặc những đặc điểm/ tính chất liên quan đến tổng thể nghiên cứu (trả lời các câu hỏi như ai?, cái gì?, khi nào?, ở đâu? và bao nhiêu? về chủ đề);
- Ước tính tỉ lệ tổng thể (proportions of a population) có những đặc điểm/ tính chất này;

- Khám phá mối liên hệ giữa những biến số.

Có nhiều cấp độ trong nghiên cứu mô tả, ở cấp đơn giản nhất, nghiên cứu mô tả liên quan đến những câu hỏi/ giả thuyết đơn biến, trong đó người nghiên cứu hỏi hoặc phát biểu về qui mô, hình thức, phân phối hoặc sự tồn tại của biến số. Ví dụ, nghiên cứu mô tả về giới tính, nghề nghiệp, độ tuổi, thời gian làm việc của nhân viên. Ở cấp cao hơn, nghiên cứu mô tả nhằm khám phá mối liên hệ giữa những biến số hay gọi là nghiên cứu tương quan (correlational study). Trong nghiên cứu mô tả, các phương pháp thu thập dữ liệu được dùng như là điều tra (survey), phỏng vấn và quan sát.

Một thiết kế mô tả yêu cầu phải cụ thể và rõ ràng bằng cách đặt câu hỏi với 5 W gồm WHO (ai?), WHAT (cái gì?), WHEN (Khi nào?), WHERE (ở đâu?) và WAY (cách nào?).

Các bước cần thực hiện trong nghiên cứu mô tả:

- Xác định đối tượng nghiên cứu: Cần thực hiện một cách thận trọng để chắc chắn rằng dữ liệu thu thập được là phù hợp, nghiên cứu có thể cung cấp thông tin không như mong muốn.
- Phương pháp thu thập thông tin: Có nhiều phương pháp thu thập thông tin trong nghiên cứu mô tả như phương pháp quan sát, phỏng vấn, xem hồ sơ... với những ưu và nhược điểm của mỗi loại, việc lựa chọn phương pháp nào tùy thuộc vào mỗi nhà nghiên cứu sao cho phù hợp với các mục đích nghiên cứu. Nhà nghiên cứu cũng có thể sử dụng một hay nhiều phương pháp trong cùng một nghiên cứu sao cho thu thập thông tin đầy đủ, đảm bảo tránh sự sai lệch và tối đa hóa độ tin cậy. Các câu hỏi phải được chuẩn bị tốt và rõ ràng, người thực hiện quan sát, phỏng vấn phải được hướng dẫn kỹ trước khi thực hiện nghiên cứu
- Mẫu nghiên cứu phải đủ lớn và mang tính đại diện để đảm bảo cho việc suy rộng cho tổng thể dựa trên phân tích mẫu.
- Để dữ liệu không bị sai lệch khi thu thập, điều cần thiết là giám sát kỹ lưỡng nhân viên khi họ thu thập và ghi lại thông tin. Kiểm tra có thể tiến hành để đảm bảo dữ liệu nhân viên thu thập thực hiện nhiệm vụ của họ một cách trung thực và không có sự định kiến. Khi dữ liệu được thu thập sẽ được kiểm tra về tính toàn diện, hiểu được, kiên định và tin cậy.
- Các dữ liệu thu thập phải được xử lý và phân tích, các thủ tục xử lý và phân tích cần được lên kế hoạch chi tiết trước khi công việc thực tế được bắt đầu nhằm đảm bảo việc sử dụng các phương pháp phân tích phù hợp với các thang đo lường đã sử dụng. Phương pháp phân tích thích hợp sẽ gia tăng niềm tin của người phân tích, giúp cho việc diễn giải dữ liệu rõ ràng hơn.

Nghiên cứu nhân quả (kiểm định giả thuyết - hypothesis-testing research studies)

Trong nghiên cứu mô tả tập trung là rõ 5 W như đã giới thiệu ở trên, nghiên cứu nhân quả ngoài việc làm rõ 5 W còn tập trung trả lời câu hỏi chính là tại sao (WHY?) và giải thích hoặc dự đoán sự ảnh hưởng của một biến này lên một biến khác. Nghiên cứu quan hệ nhân quả được sử dụng để tìm ra những bằng chứng của quan hệ nhân quả và phương pháp chính của nghiên

cứu nhân quả là thí nghiệm. Mục đích của nghiên cứu nhân quả là:

- Tìm ra những nhân tố nào là nhân tố nguyên nhân (các biến độc lập) và nhân tố nào là kết quả (biến phụ thuộc/ biến tiêu chuẩn).
- Xác định bản chất của mối quan hệ giữa các nhân tố nguyên nhân và kết quả phục vụ mục đích dự báo.

Nhà nghiên cứu nên hiểu bản chất của mối quan hệ nhân quả, mối quan hệ nhân quả giữa A và B được hiểu là A “sản sinh ra” B hoặc A làm B xảy ra. Để kiểm định giả thuyết cho mối quan hệ nhân quả, những bằng chứng sau cần phải thu thập:

- Mối quan hệ tương quan giữa A và B:
 - A và B có xảy ra/ thay đổi cùng nhau như giả thuyết đưa ra không?
 - Khi không có A, có phải B cũng không xuất hiện không?
 - Khi có ít hơn hoặc nhiều hơn A, B có ít hơn hoặc nhiều hơn không?
- Thứ tự thời gian xảy ra của A và B theo hướng như trong giả thuyết:
 - A có xảy ra trước B không?
- Không có những nguyên nhân khác gây ra B:
 - B có thay đổi cùng với các biến số khác như thay đổi cùng với A hay không?

(Kothari, 2004) cho rằng có hai loại nghiên cứu nhân quả: nghiên cứu kiểm định giả thuyết thực nghiệm (experimental hypothesis-testing research) trong đó biến độc lập được kiểm soát, và nghiên cứu kiểm định giả thuyết không thực nghiệm (non-experimental hypothesis-testing research). Ví dụ, giả sử người nghiên cứu muốn nghiên cứu liệu trí thông minh có ảnh hưởng đến khả năng đọc của sinh viên không. Để thực hiện mục tiêu này, người nghiên cứu chọn ngẫu nhiên 100 sinh viên để kiểm tra trí thông minh và khả năng đọc của họ bằng cách tính hệ số tương quan (coefficient of correlation) giữa hai điểm số. Đây là ví dụ của nghiên cứu kiểm định giả thuyết không dùng thực nghiệm bởi vì biến số độc lập ở đây là sự thông minh/ trí thông minh không được kiểm soát. Nhưng nếu người nghiên cứu chọn ngẫu nhiên 100 sinh viên và chia 100 sinh viên này làm hai nhóm: một nhóm vẫn theo học chương trình bình thường và một nhóm theo học thêm môn học kỹ năng đọc sách. Sau một thời gian, người nghiên cứu kiểm tra kỹ năng đọc của sinh viên và so sánh giữa hai nhóm. Ở ví dụ này, thiết kế nghiên cứu này thuộc dạng nghiên cứu kiểm định giả thuyết thực nghiệm vì biến độc lập ở đây là chương trình học được thay đổi/điều chỉnh.

Phương pháp thu thập dữ liệu trong nghiên cứu nhân quả có thể là điều tra (đối với nghiên cứu nhân quả không kiểm soát biến độc lập hoặc thực nghiệm (đối với nghiên cứu nhân quả có điều chỉnh / kiểm soát biến độc lập).

Sau đây là một ví dụ về thiết kế nghiên cứu nhân quả:

- Xác định mục tiêu và phạm vi nghiên cứu: Bao gồm việc giới thiệu về (1) phát biểu mục tiêu

ngiên cứu (research objectives), (2) xác định loại hình đề tài và phương pháp nghiên cứu và (3) phạm vi, ràng buộc, tiền đề nghiên cứu.

- Giới thiệu sơ lược cơ sở lý thuyết và phương pháp hình thành mô hình nghiên cứu: Giới thiệu, hợp tuyển các lý thuyết nghiên cứu và phương pháp xây dựng mô hình nghiên cứu (gồm các khái niệm nào và mối quan hệ giữa các khái niệm).
- Xác định nguồn cung cấp thông tin: Dữ liệu thứ cấp, dữ liệu sơ cấp hay mô hình thử nghiệm.
- Phương pháp đo lường và thu thập thông tin: Các loại thang đo được sử dụng, cách thức xây dựng công cụ thu thập dữ liệu (biểu mẫu quan sát hoặc bản câu hỏi).
- Phương pháp thu thập thông tin: Quan sát, phỏng vấn (trực tiếp, thư tín, qua điện thoại, Internet...), giới thiệu thời gian và địa điểm thực hiện việc thu thập thông tin.
- Mẫu và thiết kế mẫu: Không gian mẫu, khung mẫu, cỡ mẫu, phương pháp lấy mẫu.
- Phân tích dữ liệu: Công cụ phân tích dữ liệu và phương pháp phân tích dữ liệu sẽ sử dụng.
- Kế hoạch nhân lực, thời gian và ngân sách.

Mối quan hệ giữa nghiên cứu thăm dò, mô tả và nhân quả

Mặc dù có ba loại nghiên cứu được mô tả trong thiết kế nghiên cứu: nghiên cứu thăm dò, nghiên cứu mô tả và nghiên cứu nhân quả nhưng sự khác biệt giữa chúng chỉ có tính chất tương đối. Một dự án nghiên cứu có thể liên quan đến các loại nghiên cứu khác nhau nếu dự án có nhiều mục đích khác nhau. Sự kết hợp các nghiên cứu trên như thế nào là tùy thuộc vào vấn đề nghiên cứu. Sau đây là một số hướng dẫn chung cho sự kết hợp các nghiên cứu:

- Nghiên cứu thăm dò được tiến hành khi nhà nghiên cứu biết quá ít về tình hình có liên quan đến vấn đề nghiên cứu nhằm xác định chính xác hơn vấn đề cần được nghiên cứu, các câu hỏi nghiên cứu hoặc các giả thuyết cần được xây dựng.
- Nghiên cứu thăm dò là bước đầu tiên trong mô hình thiết kế nghiên cứu chung, tiếp đến là nghiên cứu mô tả và nghiên cứu nhân quả.
- Không nhất thiết khi bắt đầu một dự án nghiên cứu phải là nghiên cứu thăm dò, việc thực hiện vấn đề nghiên cứu thăm dò hay không phụ thuộc vào mức độ xác định vấn đề nghiên cứu và khả năng chắc chắn trong cách tiếp cận vấn đề nghiên cứu của nhà nghiên cứu. Một nghiên cứu có thể bắt đầu bằng nghiên cứu mô tả hoặc nghiên cứu nhân quả khi mục tiêu, câu hỏi nghiên cứu rõ ràng, các khái niệm, biến số cần được đo lường đã được biết.

MỘT SỐ VẤN ĐỀ CẦN THIẾT TRONG THIẾT KẾ NGHIÊN CỨU

Một trong những cách tiếp cận thường thấy trong nghiên cứu khoa học của sinh viên đại học và cao học là cách tiếp cận nghiên cứu định lượng, vì vậy, trong phần này sẽ giới thiệu một số vấn đề cần thiết nhằm giúp cho việc thiết kế nghiên cứu của tốt hơn:

Xem xét lý thuyết (nền) hay tổng hợp lý thuyết

Như đã trình bày ở trên, theo Creswell (2003), trong nghiên cứu định lượng, tài liệu thường được sử dụng ở phần đầu của một công trình nghiên cứu để giới thiệu một vấn đề, khi đó nhà nghiên cứu sử dụng tài liệu theo phép suy diễn như là một khuôn khổ cho việc thiết lập các câu hỏi nghiên cứu và các giả thuyết. Đồng thời, lý thuyết cũng đưa vào phần cuối của một công trình nghiên cứu, trong đó nhà nghiên cứu so sánh các kết quả của công trình nghiên cứu của mình với các kết quả trong tài liệu nhằm xem xét sự đóng góp của nghiên cứu đối với lý thuyết khoa học đã tồn tại.

Phần xem xét lý thuyết cho phép tổng hợp, phân tích về các định nghĩa, các khái niệm, lý thuyết đã có hoặc các phương pháp đã được sử dụng, từ đó có thể phê bình để có thể nhận thức được “lỗ hổng” nghiên cứu hoặc cho phép người đọc thấy được căn cứ của nghiên cứu và sự cần thiết của vấn đề nghiên cứu.

Phần xem xét tài liệu có thể dưới một số hình thức khác nhau nhưng hầu như chưa có sự đồng thuận về hình thức nào được ưa thích hơn. Creswell (2003) giới thiệu đề nghị của Cooper (1984) khi cho rằng có một số hình thức sau:

- Xem xét lại lý thuyết có thể mang tính hợp tuyển (integrative) với việc các nhà nghiên cứu tóm tắt các chủ đề tổng quát trong tài liệu, cách thực hiện này thông dụng trong các luận án tiến sĩ;
- Xem xét lý thuyết (theoretical review) khi nhà nghiên cứu tập trung vào việc tổng hợp lý thuyết hiện có liên quan đến vấn đề đang được nghiên cứu. Hình thức này xuất hiện trong các bài báo đăng trên tạp san hay tạp chí trong đó tác giả kết hợp lý thuyết nói trên vào trong phần giới thiệu công trình nghiên cứu.
- Xem xét về phương pháp luận (methodological review) trong đó nhà nghiên cứu tập trung vào các phương pháp và các định nghĩa. Những phần xem xét lại này có thể cung cấp không chỉ một bản tóm tắt về các công trình nghiên cứu đã có mà còn một bản phê bình thực sự về những điểm mạnh và những điểm yếu của các phần trình bày về phương pháp.

Câu hỏi nghiên cứu và giả thuyết

Theo Creswell (2003), trong các công trình nghiên cứu định lượng, các nhà điều tra sử dụng các câu hỏi nghiên cứu và các giả thuyết để định hình và tập trung một cách chuyên biệt mục đích của công trình nghiên cứu. Các câu hỏi nghiên cứu là những lời phát biểu nghi vấn hay những câu hỏi mà nhà nghiên cứu cố gắng trả lời. Mặt khác, các giả thuyết là những điều tiên đoán mà nhà nghiên cứu đưa ra về những mối quan hệ giữa các biến.

Các phương pháp thu thập dữ liệu sơ cấp

Có hai cách tiếp cận để thu thập dữ liệu sơ cấp là quan sát và truyền thông, truyền đạt (communications). Tùy vào các câu hỏi nghiên cứu, các loại dữ liệu, và cách tiếp cận nghiên cứu (thăm dò, mô tả, nhân quả, thực nghiệm) mà nhà nghiên cứu chọn ra một phương pháp phù hợp.

Đối với phương pháp quan sát, có thể dùng phương pháp này trong cả thiết kế nghiên cứu định

tính và định lượng. Phương pháp điều tra thuộc cách tiếp cận truyền đạt (communication) được sử dụng phổ biến trong nghiên cứu mô tả. Phương pháp thực nghiệm là phương pháp thu thập dữ liệu sơ cấp dùng trong nghiên cứu nhân quả.

Các phương pháp phân tích dữ liệu

Ở giai đoạn thiết kế nghiên cứu, nhà nghiên cứu còn phải xác định cho mình các phương pháp phân tích dữ liệu phù hợp dựa trên câu hỏi nghiên cứu, giả thuyết nghiên cứu và cách tiếp cận nghiên cứu được áp dụng. Đối với nghiên cứu thăm dò, các phần mềm về phân tích nội dung (content analysis) được sử dụng, thông thường, trong nghiên cứu thăm dò người ta thường sử dụng các phần mềm Nvivo, HyperRESEARCH để nghiên cứu lý thuyết (nền) hoặc SPSS Text Analysis for Surveys... để phân tích kết quả khảo sát nghiên cứu thảo luận nhóm mục tiêu (focus groups) hoặc nghiên cứu chuyên sâu (In-depth interviews).

Đối với nghiên cứu mô tả, các phương pháp thống kê mô tả được sử dụng để mô tả về đối tượng/hiện tượng, đặc điểm, tính chất của đối tượng/hiện tượng (ví dụ như tính số trung bình, ước tính cho tổng thể, kiểm định mối quan hệ tương quan,...) có thể sử dụng các phần mềm đơn giản như Excel hoặc các phần mềm chuyên dụng hơn như EVIEW, SPSS, STATA... Đối với nghiên cứu nhân quả, các phương pháp thống kê thường sử dụng để kiểm định sự khác biệt, sự ảnh hưởng của biến số này lên biến số khác hay mối quan hệ giữa các biến số trong mô hình. Các phần mềm thường sử dụng trong việc phân tích dữ liệu của thiết kế nghiên cứu này thường là SPSS, AMOS, LISREL...

TÓM TẮT CHƯƠNG 3

Trong nghiên cứu, vấn đề khó khăn đặt ra là vấn đề nghiên cứu cần phải được xác định một cách rõ ràng, các quyết định liên quan đến cái gì, ở đâu, khi nào, bao nhiêu, bởi cái gì liên quan tới một cuộc điều tra hoặc một phác thảo nghiên cứu tạo thành một thiết kế nghiên cứu.

Thiết kế nghiên cứu bao gồm một phác thảo những gì người nghiên cứu sẽ làm từ việc đưa ra cách thức xây dựng công cụ thu thập dữ liệu, phương pháp thu thập, thiết kế mẫu, mô hình phân tích, giả thuyết, tổ chức thu thập dữ liệu cho đến phương pháp phân tích dữ liệu.

Có nhiều cách phân loại thiết kế nghiên cứu, đặc biệt, nhà nghiên cứu thường dựa trên mục tiêu của nghiên cứu là (1) nhận diện và xác định vấn đề; (2) mô tả vấn đề đã được xác định và (3) phát hiện mối quan hệ giữa các biến số trong vấn đề nghiên cứu để đề ra giải pháp giải quyết để phân thiết kế nghiên cứu thành thiết kế nghiên cứu thăm dò/ khám phá, nghiên cứu mô tả và nghiên cứu nhân quả.

Thông thường, nghiên cứu thăm dò là bước đầu tiên trong mô hình thiết kế nghiên cứu chung, tiếp đến là nghiên cứu mô tả và nghiên cứu nhân quả. Tuy nhiên, không nhất thiết khi bắt đầu một dự án nghiên cứu phải là nghiên cứu thăm dò, việc thực hiện vấn đề nghiên cứu thăm dò hay không phụ thuộc vào mức độ rõ ràng và cụ thể của vấn đề nghiên cứu và khả năng chắc chắn trong cách tiếp cận vấn đề nghiên cứu của nhà nghiên cứu.

CHƯƠNG

4

NGHIÊN CỨU ĐỊNH TÍNH (QUALITATIVE RESEARCH)

NGHIÊN CỨU ĐỊNH TÍNH VÀ NGHIÊN CỨU ĐỊNH LƯỢNG

Nghiên cứu định tính (Qualitative Research):

Trong nghiên cứu, nghiên cứu định tính cung cấp sự hiểu biết sâu sắc về vấn đề nghiên cứu, phương pháp này thường được dùng để xây dựng lý thuyết khoa học dựa vào quy trình quy nạp (Thọ, 2011). Phương pháp định tính dựa vào quy trình quy nạp và thường liên quan đến việc phân tích và diễn giải dữ liệu nhằm mục đích khám phá quy luật của hiện tượng khoa học.

Nghiên cứu định lượng (Quantitative Research):

Theo Thọ (2011), khác với nghiên cứu định tính, nghiên cứu định lượng nhằm vào mục đích thu thập dữ liệu để kiểm định các lý thuyết khoa học được suy diễn từ lý thuyết đã có. Thông thường, bất kỳ khi nào một vấn đề nghiên cứu được hình thành thì nghiên cứu định lượng phải đi trước nghiên cứu định tính và đôi khi, nghiên cứu định tính cũng được thực hiện để giải thích kết quả đạt được của nghiên cứu định lượng.

Các phương pháp thu thập thông tin khác nhau đem lại thông tin khác nhau. Do vậy, trước khi quyết định sử dụng phương pháp nào cần phải xác định loại thông tin nào cần thiết nhất cho mục đích nghiên cứu. Các phương pháp nghiên cứu định tính và nghiên cứu định lượng có thể kết hợp để bổ sung lẫn cho nhau. Nghiên cứu định tính có thể hỗ trợ cho nghiên cứu định lượng với mục đích xác định chủ đề phù hợp, nghiên cứu định lượng có thể hỗ trợ cho nghiên cứu định tính bằng cách khái quát hóa các phát hiện ra một mẫu lớn hơn hay nhận biết các nhóm cần nghiên cứu sâu hoặc nghiên cứu định tính có thể giúp giải thích các mối quan hệ giữa các biến số được phát hiện trong các nghiên cứu định lượng. Trong một nghiên cứu định tính cũng có thể có nghiên cứu định lượng và ngược lại.

MỘT SỐ VẤN ĐỀ CHẤT LƯỢNG DỮ LIỆU VÀ CHUẨN BỊ THU THẬP DỮ LIỆU

Chất lượng dữ liệu

Theo Saunders và các cộng sự (2007), chất lượng dữ liệu liên quan đến độ tin cậy, các dạng sai lệch, giá trị và tính tổng quát hóa. Nhà nghiên cứu cần phải xem xét liệu những nhà nghiên cứu khác có khám phá thông tin tương tự hay không, có tồn tại những sai lệch liên quan đến người phỏng vấn (như những sai lệch do giọng điệu, cử chỉ, hành vi của người phỏng vấn), sai lệch do người được phỏng vấn.

Địa điểm nghiên cứu

Địa điểm tiến hành phỏng vấn sẽ ảnh hưởng đến chất lượng dữ liệu thu thập, do vậy, cần phải lựa chọn điểm điểm thuận tiện cho những người tham gia, tạo ra sự thoải mái đối với người tham gia phỏng vấn. Đặc biệt, những yếu tố khung cảnh, sự nhận diện hoặc tiếng ồn cũng có thể ảnh hưởng đến chất lượng.

Giao tiếp trước cuộc phỏng vấn

Đối với những người chưa quen biết, những giao tiếp sơ lược trước cuộc phỏng vấn có tác động đáng kể đến kết quả của cuộc phỏng vấn, đồng thời, tạo ra sự tin nhiệm của người được phỏng vấn đối với người nghiên cứu. Tất cả những thông tin liên quan đến lý thuyết mà nhà nghiên cứu xem xét, kinh nghiệm của người nghiên cứu, những hiểu biết từ thực tế, những thảo luận đối với sinh viên, đồng nghiệp, giáo viên hướng dẫn cũng có thể tổng hợp và nảy sinh ý tưởng cho nghiên cứu cũng có thể đưa ra như là sự tranh thủ ý kiến của những người tham gia.

Phương pháp hỏi

Phương pháp hỏi giúp giảm phạm vi sai lệch trong suốt quá trình phỏng vấn và tăng độ tin cậy của thông tin thu được. Câu hỏi của người nghiên cứu cần được diễn đạt rõ ràng, với quan điểm để người được phỏng vấn có thể hiểu và tránh thiên lệch. Easterby-Smith và cộng sự (2002) cho rằng, việc sử dụng câu hỏi mở có thể giúp tránh được các sai lệch trong nghiên cứu.

Người nghiên cứu

Đặc điểm, hành vi, văn hóa của người thực hiện phỏng vấn ảnh hưởng đến sự sai lệch của thông tin. Việc chăm chú lắng nghe, phương pháp ghi chép giúp người phỏng vấn có được những ý kiến nhận xét xác đáng và có ý nghĩa đối với nghiên cứu và cũng khám phá điều này với những người được phỏng vấn.

Các loại câu hỏi được thiết lập

Câu hỏi mở

Câu hỏi mở cho phép khuyến khích người được phỏng vấn cung cấp thông tin toàn diện, mở rộng và dễ dàng bộc lộ thái độ, quan điểm của mình. Ví dụ, “Vì sao những phương pháp quản lý hiện tại giúp cho công việc của bạn hoàn thành tốt”, “tại sao cách quản lý hiện tại không khuyến khích say mê công việc của nhân viên?”...

Câu hỏi thăm dò

Câu hỏi thăm dò để khám phá những hồi đáp có ý nghĩa hoặc tìm kiếm một sự giải thích (rõ hơn) đối với đề tài nghiên cứu. Ví dụ: “làm thế nào để khuyến khích đóng góp của nhân viên?”, “ngoài những điều trên còn có điều gì nữa không?”, “hay nói rõ hơn là?”.

Câu hỏi cụ thể và câu hỏi đóng

Những câu hỏi này tương tự với các loại câu hỏi được sử dụng phỏng vấn có cấu trúc, nó được

dùng để thu thập những thông tin rõ ràng, cụ thể hoặc để xác nhận một sự hiện diện hoặc một ý kiến.

PHƯƠNG PHÁP VÀ KỸ THUẬT NGHIÊN CỨU ĐỊNH TÍNH

Phương pháp

Grounded theory (GT)

Theo Glaser và Strauss (1967), mục đích của grounded theory (GT) là để “khám phá ra một lý thuyết”. Cũng theo hai tác giả này, GT được định nghĩa là “sự khám phá ra lý thuyết từ các dữ liệu được thu thập một cách có hệ thống từ các nghiên cứu xã hội” (Glaser và Strauss 1967). Charmaz (1995, 2002) xác định một số các yếu tố/ đặc điểm mà GT có:

- Sự thực hiện đồng thời việc thu thập và phân tích dữ liệu;
- Tạo ra các mã phân tích (analytic codes) và các phạm trù (categories) được phát triển từ dữ liệu mà không phải từ các lý thuyết dự kiến trước (pre-existing conceptualisations/ theoretical sensitivity);
- Khám phá các quá trình xã hội cơ bản (basic social processes) trong dữ liệu;
- Dùng phương pháp qui nạp xây dựng các phạm trù trừu tượng (abstract categories);
- Chọn mẫu lý thuyết (theoretical sampling) để điều chỉnh các phạm trù;
- Ghi chú dữ liệu (writing analytical memos) là giai đoạn giữa mã hóa và viết;
- Tích hợp các phạm trù vào khung/mô hình lý thuyết (theoretical framework);

Khi thực hiện GT, chúng ta có ý tưởng:

- Bắt đầu một nghiên cứu như thế nào (xác định lĩnh vực yêu thích, tránh dự kiến trước về lý thuyết (theoretical preconceptions) và sử dụng nhạy cảm về lý thuyết (theoretical sensitivity));
- Làm thế nào để thực hiện nghiên cứu (thông qua qui trình thống kê và chiến lược chọn mẫu);
- Dừng lại như thế nào (khi mức độ bão hòa (theoretical saturation) về lý thuyết xảy ra).

Các phương pháp thu thập dữ liệu trong GT bao gồm phỏng vấn sâu (câu hỏi trong phỏng vấn sâu có thể được điều chỉnh để phản ánh các lý thuyết xuất hiện), các phương pháp quan sát, thảo luận nhóm (focus group). Đối với nhạy cảm về lý thuyết (theoretical sensitivity):

- Người nghiên cứu sẽ trở nên mải mê trong dữ liệu và cố gắng thấu hiểu những người tham gia nhìn nhận cái gì là quan trọng;
- Thực hiện đồng thời thu thập và phân tích dữ liệu cho phép người nghiên cứu trở nên nhạy cảm với dữ liệu;
- Bắt đầu với càng ít những ý tưởng được định sẵn càng tốt.

Những lý thuyết hiện tại và những kiến thức, kinh nghiệm đã có của người nghiên cứu có thể được sử dụng để phát triển các phạm trù (categories) nhưng những phạm trù không nên bị ép buộc phải phù hợp với lý thuyết và những kiến thức, kinh nghiệm không nên được sử dụng để tạo ra các phạm trù.

Trong phương pháp này, vai trò của tổng quan tài liệu (literature review)

- Để có được sự nhạy cảm về lý thuyết, nhà nghiên cứu phải bắt đầu với càng ít những ý tưởng đã xác định sẵn, đặc biệt là các giả thuyết càng tốt vì thế họ có thể nhạy cảm với dữ liệu càng nhiều càng tốt.
- Nhà nghiên cứu phải có một đầu óc cởi mở chứ không phải là một “cái đầu rỗng” (Dey 1999)
- Đây là cách mà các kiến thức có trước được sử dụng để gây sự khác biệt, được sử dụng để hỗ trợ/cung cấp thông tin cho việc phân tích dữ liệu.
- Tài liệu có thể được sử dụng như dữ liệu và thường xuyên được so sánh với các phạm trù trừu tượng xuất hiện để tích hợp trong lý thuyết (Glaser 1992).

Nghiên cứu tình huống (Case studies)

Nghiên cứu tình huống

Nghiên cứu tình huống có thể dựa trên các chứng cứ định lượng lẫn định tính. Nghiên cứu tình huống, theo Yin (1994), là phương pháp điều tra/nghiên cứu một hiện tượng trong bối cảnh thực tế. Đặc biệt khi ranh giới giữa hiện tượng và bối cảnh không có sự phân chia rõ ràng. Một nghiên cứu tình huống liên quan đến việc phát triển kiến thức sâu và chi tiết về một tình huống đơn lẻ, hoặc một số những tình huống liên quan với nhau. Phương pháp thu thập dữ liệu trong nghiên cứu tình huống có thể có những phương pháp sau: phỏng vấn, quan sát, phân tích tài liệu và điều tra.

Sử dụng nghiên cứu tình huống

Nghiên cứu tình huống phù hợp trong nghiên cứu thăm dò hoặc để xem xét lại một lý thuyết đang tồn tại. Vấn đề đặt ra là khi nào nghiên cứu tình huống được sử dụng, nghiên cứu tình huống sử dụng khi:

- Khi cần trả lời câu hỏi làm thế nào và tại sao?;
- Khi người nghiên cứu ít có thể kiểm soát lên các sự kiện;
- Khi tập trung nhiều vào các sự kiện hiện tại hơn là các sự kiện, hiện tượng lịch sử;
- Khi có nhiều biến số cần nghiên cứu.

Khái quát hóa trong nghiên cứu tình

Nghiên cứu tình huống không đại diện cho một mẫu đại diện và nó không thể sử dụng để suy ra tổng thể (statistical generation). Tuy nhiên, nghiên cứu tình huống có thể được suy ra/khái quát

hóa thành các giả thuyết lý thuyết (theoretical propositions).

Thiết kế nghiên cứu trong nghiên cứu tình huống

- Xác định câu hỏi nghiên cứu;
- Phát triển giả thuyết lý thuyết;
- Lựa chọn tình huống;
- Xác định đơn vị phân tích;
- Lựa chọn phương pháp thu thập dữ liệu;
- Liên kết một cách logic giữa dữ liệu và giả thuyết lý thuyết;
- Tiêu chuẩn để diễn giải kết quả.

Ethnography

Nguồn gốc của nghiên cứu dân tộc học là nghiên cứu nhân chủng học với trọng tâm là nghiên cứu các xã hội ở qui mô nhỏ để hiểu được thế giới quan, hành vi của con người trong xã hội đó. Mục tiêu của nghiên cứu dân tộc học là thấu hiểu hệ thống các ý nghĩa chung được chia sẻ giữa những người dân trong xã hội được nghiên cứu, hay còn gọi là văn hóa. Trong nghiên cứu kinh doanh, đặc biệt là trong lĩnh vực marketing và hành vi người tiêu dùng, phương pháp nghiên cứu dân tộc học là một trong những phương pháp được sử dụng khá phổ biến (ví dụ như những nghiên cứu của Arnould và Wallendorf, 1994; và Holt, 1997). Pettigrew (2000) lập luận rằng việc sử dụng các kỹ thuật nghiên cứu dân tộc học có thể giải thích việc tiêu dùng một cách hiệu quả. Một trong những giả định cho việc nghiên cứu người tiêu dùng theo phương pháp dân tộc học là những ý nghĩa xã hội gắn liền với việc sở hữu vật chất có thể được xem như là những tín hiệu về văn hóa (Douglas và Isherwood, 1979, trích trong Pettigrew, 2000). Cũng vì lí do trên mà những nhà nghiên cứu sử dụng phương pháp này để “*giải thích các hình thức khác nhau của hành vi tiêu dùng mà không cần phải tham chiếu trực tiếp đến những khía cạnh văn hóa của hành vi đó*” (Arnould và Wallendorf, 1994, trích trong Pettigrew, 2000).

Đặc điểm của nghiên cứu dân tộc học

Theo Arnould và Wallendorf (1994, trang 485), những đặc điểm của một nghiên cứu dân tộc học định hướng thị trường gồm có:

- Việc thu thập dữ liệu được thực hiện trong bối cảnh tự nhiên, không được dàn xếp hoặc sắp đặt
- Việc sử dụng quan sát tham gia (participant observation): liên quan đến việc tham gia lâu dài và mang tính trải nghiệm của nhà nghiên cứu trong một bối cảnh văn hóa cụ thể.
- Độ tin cậy của nghiên cứu dân tộc học được đánh giá qua ba tiêu chuẩn: sự minh bạch trong lý thuyết, sự trình bày rõ ràng về lộ trình của nhà nghiên cứu thông qua thu thập

dữ liệu, và tính đến mối quan hệ giữa diễn giải nghiên cứu dân tộc học và các bằng chứng ghi lại từ hiện trường (Sanjek, 1990, trích trong Arnould và Wallendorf, 1994)

- Việc tích hợp nhiều nguồn dữ liệu là điều cần thiết.

Trong nghiên cứu, dựa vào cách tiếp cận nghiên cứu người ta phân thành phương pháp tiếp cận trực tiếp và phương pháp tiếp cận gián tiếp. Phương pháp tiếp cận trực tiếp gồm (1) thảo luận nhóm – Focus groups và (2) phỏng vấn chuyên sâu – In-depth Interview và (3) quan sát – Observation. Phương pháp tiếp cận gián tiếp hay còn gọi là kỹ thuật ánh xạ (Projective technique) gồm (1) kỹ thuật liên tưởng – Association techniques, (2) kỹ thuật hoàn chỉnh – Completion techniques, (3) kỹ thuật diễn giải – Construction techniques và (4) kỹ thuật diễn cảm – Expressive techniques.

Phương pháp tiếp cận trực tiếp

Thảo luận nhóm (focus groups)

Thảo luận nhóm (focus groups) được định nghĩa là việc thu thập dữ liệu được thông qua hình thức thảo luận giữa các đối tượng nghiên cứu dưới sự hướng dẫn của người điều khiển. Người điều khiển có thể là nhà nghiên cứu hoặc một người đã được tập huấn theo hướng không chính thức nhưng rất linh hoạt, có nhiệm vụ hướng dẫn thảo luận nhóm. Mục đích của kỹ thuật này nhằm đạt được những hiểu biết sâu sắc vấn đề nghiên cứu bằng cách lắng nghe một nhóm người được chọn ra từ một nhóm khách hàng hoặc chuyên gia mục tiêu phù hợp với những vấn đề mà nhà nghiên cứu đang quan tâm. Thảo luận nhóm là một kỹ thuật nghiên cứu định tính quan trọng và thường được sử dụng phổ biến trong thực tế nghiên cứu (Morgan, 1996).

Kỹ thuật phỏng vấn nhóm thường thực hiện trên nhóm khoảng từ 4 - 12 người (tùy theo chủ đề có phức tạp hay đơn giản, chủ đề càng phức tạp thì số lượng người tham gia càng ít), những người này nên có những đặc điểm tương đồng về nhân khẩu và điều kiện kinh tế xã hội nhằm tránh những mâu thuẫn hoặc phân cấp giữa các thành viên trong, các cá thành viên không quen biết nhau, thời gian thảo luận có thể kéo dài từ 1 - 3 giờ nhưng thảo luận từ 1,5 - 2 giờ là tốt nhất. Trong quá trình thực hiện, người nghiên cứu có thể dùng các thiết bị hỗ trợ như video hoặc máy ghi âm, tuy nhiên, cần thiết kế một cách bí mật (để tránh sự nhận diện) hoặc thông báo trước cho người tham gia để hạn chế những tác động của các thiết bị hỗ trợ. Người điều khiển cần phải có khả năng dẫn dắt chương trình, phân tích, tổng hợp dữ liệu có liên quan nên cần phải có tính tình thân thiện, thoải mái nhằm gia tăng sự chân thành của người tham gia, có hiểu biết, linh hoạt để ứng phó với những tình huống xảy ra trong thực tế thảo luận.

Trong quá trình thảo luận, người điều khiển có thể đưa ra những câu hỏi nhằm khuyến sự tham gia thảo luận của người tham gia, ví dụ, “bạn thường dùng loại sản phẩm nào?”, “khi mua sản phẩm tại cửa hàng, bạn thường chú ý đến đặc điểm nào của sản phẩm?”, “khi mua hàng, bạn cần những thông tin tư vấn gì từ phía người bán?”...

Nhà nghiên cứu thường sử dụng phương pháp thảo luận nhóm bởi vì có thể (1) kết hợp nhiều người sẽ thu thập được một lượng lớn thông tin; (2) kích thích các thành viên tham gia trả lời từ những mức độ chung đến những nội dung cụ thể; (3) người trả lời cảm thấy thoải mái và hài

lòng cho ý kiến bởi vì những cảm nghĩ của người tham gia cũng tương tự các thành viên khác trong nhóm; (4) câu trả lời không phải đòi hỏi quá lớn cho nên người tham gia trả lời rất tự nhiên, cho các câu trả lời chính xác; (5) các ý kiến có thể khách quan của nhóm chứ không phải từ cá nhân.

Những bất lợi của thảo luận nhóm được tổng kết thông qua nguyên tắc 5Ms sau:

- Ứng dụng sai/ lạm dụng (Misuse): Việc sử dụng phòng vấn nhóm ứng dụng sai hoặc lạm dụng khi xem xét kết quả như là một kết luận hơn là một sự thăm dò để phát hiện vấn đề.
- Đánh giá sai (Misjudge): Kết quả của thảo luận nhóm dễ bị đánh giá sai hơn các kỹ thuật phỏng vấn khác vì thành kiến của người nghiên cứu.
- Điều khiển (Moderation): Rất khó để tìm người điều khiển thỏa mãn các yêu cầu, kỹ năng để dẫn dắt thảo luận nhóm thảo luận nhóm và chất lượng nghiên cứu phụ thuộc rất nhiều vào người điều khiển.
- Lộn xộn (Messy): Kết quả thảo luận rất khó phân tích và tổng hợp vì dữ liệu khá lộn xộn.
- Không đại diện (Misrepresentation): Kết quả của thảo luận nhóm khó đại diện cho tổng thể vì phỏng vấn chỉ một nhóm.

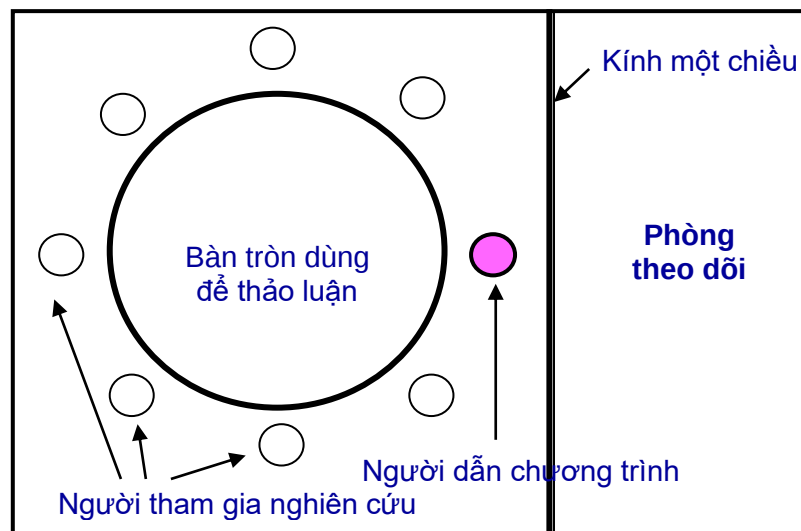
Một số hình thức khác của thảo luận nhóm gồm:

- Nhóm thảo luận hai chiều: điều này cho phép một nhóm mục tiêu lắng nghe hoặc học hỏi một nhóm khác có liên hệ (liên hệ) tương ứng
- Nhóm thảo luận song đôi: là nhóm phỏng vấn được tiến hành bởi hai người điều khiển, một người đảm bảo dẫn dắt buổi thảo luận (về hình thức) và người kia đảm bảo về nội dung của thảo luận.
- Nhóm thảo luận tay đôi: là nhóm phỏng vấn có hai người điều khiển với vị trí ngược nhau về các vấn đề được thảo luận, cho phép nhà nghiên cứu khai thác cả hai mặt của các vấn đề thảo luận.
- Nhóm kết hợp người điều khiển và người trả lời: nhóm thảo luận cho phép người điều khiển nhờ người tham gia thảo luận đóng vai trò người điều khiển (tạm thời) nhằm nâng cao sự linh hoạt của nhóm.
- Nhóm khách hàng tham gia: khách hàng được mời thành lập nhóm thảo luận.
- Nhóm thảo luận nhỏ (mini): những nhóm này bao gồm người điều khiển và từ 4 -5 người được phỏng vấn tham gia, cách này được sử dụng khi các vấn đề quan tâm đòi hỏi cần phải bao quát hơn trong nhóm từ 8 đến 12 người.
- Nhóm thảo luận bằng điện thoại: dùng điện thoại để thảo luận các vấn đề quan tâm giữa nhà nghiên cứu và nhóm.

Nội dung cần phải thực hiện đối với thảo luận nhóm gồm:

- Xác định mục tiêu của nghiên cứu và vấn đề nghiên cứu;
- Xác định mục tiêu cụ thể của nghiên cứu định tính;
- Thiết kế câu hỏi thảo luận nhóm;
- Thiết lập bản câu hỏi;
- Phát triển đề cương của người điều khiển (moderator's outline);
- Tiến hành phỏng vấn nhóm;
- Nhập và phân tích dữ liệu;
- Tóm tắt kết luận và lập kế hoạch nghiên cứu cho nghiên cứu tiếp theo.

Hình 8. Không gian thiết kế nghiên cứu định tính



Để đảm bảo cho quá trình nghiên cứu định tính, thông thường không gian nghiên cứu định tính được thiết đặt như sau gồm phòng phỏng vấn có bố trí bàn tròn hoặc ô van, người dẫn chương trình ngồi xoay lưng lại với kính một chiều. Trong phòng theo dõi có thể bố trí người để theo dõi song song hoặc máy quay để ghi lại các hoạt động liên, giữa phòng phỏng vấn và phòng theo dõi được bố trí kính một chiều để các đối tượng nghiên cứu không nhận thấy được việc theo dõi, tránh những tác động tâm lý đối với người tham gia nghiên cứu.

Phòng vấn sâu (In-depth interview)

Phòng vấn chuyên sâu là phương pháp phỏng vấn cá nhân, trực tiếp và không chính thức. Phương pháp này thường thích hợp trong các trường hợp (1) chủ đề nghiên cứu mang tính cá nhân cao; (2) khó mời các chuyên gia phỏng vấn cùng một lúc do đặc điểm của người tham gia phỏng vấn; (3) do muốn đào sâu vấn đề nghiên cứu. Đặc biệt, phương pháp phỏng vấn này người trả lời được hỏi về các khía cạnh niềm tin, thái độ và cảm nghĩ về chủ đề nghiên cứu dưới sự điều khiển của người phỏng vấn có kỹ năng cao.

Do chỉ có hai người phỏng vấn trực tiếp (one – to – one, face – to – face) giữa người phỏng vấn và người được phỏng vấn nên thời gian khoảng từ 30 phút đến 1 giờ. Người phỏng vấn thường

bắt đầu từ những câu hỏi tổng quát, sau đó vào các câu hỏi chuyên sâu, cụ thể bằng các câu hỏi “tại sao?”, “như thế nào?”...

Phòng vấn cá nhân không thể hiểu biết sâu sắc bản chất bên trong của vấn đề nghiên cứu so với phỏng vấn nhóm nhưng biết chính xác câu trả lời của riêng từng người được phỏng vấn. Phòng vấn cá nhân còn thực hiện trong bầu không khí trao đổi thông tin hoàn toàn tự do và hoàn toàn không có bất kỳ một áp lực mang tính xã hội như trong phỏng vấn nhóm.

Tuy nhiên, phỏng vấn cá nhân cũng có nhiều nhược điểm, chẳng hạn như để tìm được người phỏng vấn có kỹ năng thì rất khó khăn và tốn kém. Thiếu cấu trúc có thể tạo ra kết quả dễ xúc cảm đối với ảnh hưởng của người phỏng vấn. Chất lượng và sự hoàn hảo của kết quả rất phụ thuộc vào kỹ năng của phỏng vấn. Dữ liệu thu thập được khó phân tích và tổng hợp hơn phỏng vấn nhóm thậm chí còn rất phức tạp vì chi phí phỏng vấn cao nên thường có cỡ mẫu nhỏ, điều này thể hiện tính đại diện thấp.

Quan sát (Observation)

Quan sát là phương pháp ghi lại có kiểm soát các biến cố hoặc tác phong của con người qua cảm nhận nơi mình đang sống hay hành động, con người có thể ghi nhận và lượng định các sự kiện bên ngoài. Quan sát gồm hai hành động của con người: nghe nhìn để cảm nhận và lượng định. Con người có thể quan sát trực tiếp bằng tai, mắt để nghe, nhìn hay bằng phương tiện cơ giới như máy quay, camera... (Giới và cộng sự, 2006). Chương trình quan sát có thể được xác định chi tiết gồm việc trả lời các câu hỏi: ai?, cái gì?, khi nào?, ở đâu? và bằng cách nào? Trong khoa học, người ta chia quan sát thành hai loại, quan sát tham gia và quan sát có tính cấu trúc.

Quan sát tham gia (participant observation)

Quan sát tham gia là loại quan sát mà nhà nghiên cứu tham gia hoàn toàn vào hoạt động của những chủ thể nghiên cứu và trở thành thành viên trong nhóm, điều này giúp nhà nghiên cứu chia sẻ được kinh nghiệm không những thông qua quan sát mà cảm nhận về sự vật, hiện tượng (Gill và Johnson, 2002). Gill và Johnson (2002) đã phân quan sát tham gia thành 4 loại khi căn cứ vào vai trò mà một người quan sát tham gia có thể là (1) người tham gia hoàn toàn, (2) người quan sát hoàn toàn, (3) tham gia chủ động để quan sát và (4) tham gia thụ động để quan sát.

- Người tham gia hoàn toàn (complete participant): người quan sát tham gia hoàn toàn vào nhóm như là một thành viên, không tiết lộ mục đích (thật) cho các thành viên trong nhóm.

Ví dụ: Bạn muốn nghiên cứu hành vi mua hàng của các khách hàng tại một siêu thị, bạn có thể tham gia với bạn bè để đi mua hàng nhằm quan sát số lượng mua, các thức tìm kiếm thông tin trong siêu thị, những “điểm dừng chân” trong siêu thị để xem và mua hàng...

- Người quan sát hoàn toàn (complete observer): người quan sát không tham gia vào hoạt động của nhóm và cũng không tiết lộ mục đích quan sát.
- Người tham gia chủ động để quan sát (participant as observer): Trong vai trò là người quan sát tham gia nhóm và thông báo mục đích nghiên cứu cho các thành viên.

- Người tham gia thụ động để quan sát (observer as participant): người quan sát không phải là thành viên (là “khán giả”) và không thông báo mục đích quan sát.

Quan sát có cấu trúc

Quan sát có cấu trúc là phương pháp quan sát có tính hệ thống và có cấu trúc định sẵn.

Thủ tục tiếp cận gián tiếp

Kỹ thuật liên tưởng (Association Technique)

Kỹ thuật liên tưởng là kỹ thuật trong đó người được phỏng vấn trình bày ý kiến với sự kích thích và được hỏi bằng bảng liệt kê các từ để trả lời từng từ một mà từ đó sẽ rất gợi nhớ. Sau khi thực hiện phỏng vấn, nhà nghiên cứu xử lý bằng cách xem xét và nổi bật điều mà những người phỏng vấn quan tâm (thường là những từ đồng nghĩa lặp lại).

Kỹ thuật hoàn chỉnh (Completion technique)

Kỹ thuật hoàn chỉnh là kỹ thuật đòi hỏi người đọc phỏng vấn hoàn chỉnh những tình huống chưa kết thúc các vấn đề quan tâm. Thông thường, kỹ thuật hoàn chỉnh được sử dụng để hoàn thành câu hoặc một câu chuyện và họ sẽ điền thêm vào bất kỳ nội dung gì mà họ chọn.

Ví dụ: Tôi không thích công việc đó vì:.....
 Phương pháp quản lý đang áp dụng có hạn chế là:.....
 Công việc của người làm công tác văn phòng là:.....

Kỹ thuật dựng hình (Construction Technique)

Kỹ thuật dựng hình là Theo kỹ thuật này, người được hỏi được cho xem một số tình huống gợi mở nào đó, sau đó đề nghị họ viết lại câu chuyện hay phác họa diễn tả vấn đề cần nghiên cứu. Kỹ thuật này đòi hỏi người đọc phỏng vấn trình bày câu trả lời theo hình thức của một câu chuyện, một mẫu đàm thoại hay mô tả. Kỹ thuật này bao gồm hai hình thức: diễn giải qua tranh ảnh và đặt lời chú giải cho phim hoạt hình.

- Kỹ thuật diễn giải hình ảnh (Picture Response Technique): Người đọc phỏng vấn xem các tranh ảnh hoặc sản phẩm cụ thể và trả lời theo dạng mô tả lại theo cách suy nghĩ của người được phỏng vấn.
- Kỹ thuật kiểm tra bằng phim hoạt hình (Cartoon Test): Sau khi xem phim, người được phỏng vấn trả lời mẫu đàm thoại trong phim hoạt hình. Kỹ thuật này giúp nhà nghiên cứu hiểu được cảm nghĩ, sự tin tưởng và thái độ của người được phỏng vấn đối với vấn đề đang được quan tâm.

Kỹ thuật diễn cảm (Expressive Technique)

Người được phỏng vấn trong kỹ thuật này trình bày câu trả lời dưới hình thức kể hay quan sát và trả lời những câu hỏi có liên quan đến cảm nghĩ và thái độ của người khác đối với vấn đề nghiên cứu.

Những thuận lợi và khó khăn của thủ tục tiếp cận gián tiếp:

- Thuận lợi: Kỹ thuật tiếp cận gián tiếp có thể đa ra các câu trả lời đối với các câu hỏi còn thiếu thông tin, có thể tăng tính hiệu lực của các câu trả lời bằng cách phân biệt các mục đích nghiên cứu nhất là khi các vấn đề thuộc về cá nhân, cảm xúc, niềm tin, thái độ hay các tiêu chuẩn xã hội.
- Khó khăn: Đòi hỏi phỏng vấn viên và những người tổng hợp, phân tích thông tin phải có trình độ, có kỹ năng và được huấn luyện. Phương pháp này tốn khá nhiều chi phí nhưng cũng dễ tạo ra sai lệch trong quá trình phân tích dữ liệu.

Ví dụ: Dàn bài thảo luận nhóm mục tiêu (focus groups)⁶

Xin chào Quý Anh, Chị !

Tôi là Tôi đang tiến hành nghiên cứu về quyết định mua bảo hiểm tự nguyện xe ô tô tại..... Trước tiên tôi xin cảm ơn các Anh/ Chị đã dành thời gian tham gia với tôi. Tôi rất hân hạnh đón tiếp và thảo luận với các Anh/ Chị về chủ đề này. Và cũng xin các Anh/ Chị chú ý, không có quan điểm nào đúng hay sai mà từng ý kiến của Anh/ Chị đều hữu ích và là nguồn dữ liệu quý giá đối với đề tài nghiên cứu của tôi. Kính mong được nhận sự giúp đỡ của Anh/ Chị.

1. Khi tham gia giao thông đường bộ, theo Anh/ Chị yếu tố nào ảnh hưởng đến nhận thức sự rủi ro? Vì sao?
2. Anh/ Chị có biết đến bảo hiểm tự nguyện xe ô tô không?
3. Anh/ Chị hiểu thế nào về bảo hiểm tự nguyện xe ô tô ?
4. Theo quan điểm của Anh/ Chị , nhân tố nào ảnh hưởng đến ý định mua bảo hiểm tự nguyện xe ô tô? Vì sao?
5. Gọi ý những yếu tố (trong mô hình giả thuyết nghiên cứu) thì những yếu tố nào là quan trọng và những yếu tố nào là không quan trọng? Vì sao?
6. Theo quan điểm của Anh/ Chị, nhân tố nào ảnh hưởng đến quyết định mua bảo hiểm tự nguyện xe ô tô? Vì sao?
7. Những ý kiến khác của Anh/ Chị về bảo hiểm tự nguyện xe ô tô ?

Xin chân thành cảm ơn sự giúp đỡ của Quý Anh/ Chị.

TÓM TẮT CHƯƠNG 4

⁶. Dàn bài thảo luận nhóm (focus groups) cho nghiên cứu: Lê Văn Huy, Lê Thị Hương Giang (2010). Các nhân tố ảnh hưởng đến quyết định mua bảo hiểm tự nguyện xe ô tô. *Tạp chí Phát triển kinh tế*. 240. 25-31.

Phương pháp định tính dựa vào quy trình quy nạp và thường liên quan đến việc phân tích và diễn giải dữ liệu nhằm mục đích khám phá quy luật của hiện tượng khoa học, nghiên cứu định tính cung cấp sự hiểu biết sâu sắc về vấn đề nghiên cứu, phương pháp này thường được dùng để xây dựng lý thuyết khoa học dựa vào quy trình quy nạp.

Trong nghiên cứu định tính, những phương pháp thường dùng là Grounded theory, nghiên cứu tình huống (case studies) và dân tộc học (Ethnography) bao gồm các phương pháp tiếp cận trực tiếp gồm (1) thảo luận nhóm – Focus groups và (2) phỏng vấn chuyên sâu – In-depth Interview và (3) quan sát - Observation. Thủ tục tiếp cận gián tiếp hay còn gọi là kỹ thuật ánh xạ (Projective technique) gồm (1) kỹ thuật liên tưởng – Association techniques, (2) kỹ thuật hoàn chỉnh – Completion techniques, (3) kỹ thuật diễn giải – Construction techniques và (4) kỹ thuật diễn cảm – Expressive techniques.

CHƯƠNG

5

PHÁT TRIỂN KHÁI NIỆM VÀ THANG ĐO LƯỜNG

(CONSTRUCT DEVELOPMENT AND SCALE MEASUREMENT)

ĐO LƯỜNG

Định nghĩa đo lường

Trong cuộc sống hằng ngày, có rất nhiều vấn đề cần phải đo lường như cần phải biết số cân nặng, chiều cao, số năm học của các học sinh, sinh viên; số năm công tác, chức vụ, lương của cán bộ công nhân viên trong doanh nghiệp... Đây là các vấn đề khá dễ dàng đo lường và lượng định như để xác định cân nặng có thể dùng cân, xác định chiều cao có thể dùng thước... Tuy nhiên, có những vấn đề rất khó để đo lường như sự cảm nhận của cán bộ công nhân viên đối với phong cách của nhà quản lý; thái độ của khách hàng đối với một sản phẩm, cửa hàng, siêu thị... bởi vì đây là những vấn đề mang tính định tính. Để đo lường, theo Kothari (2004), “là quá trình gắn những con số với các đặc tính của các sự vật, hiện tượng nghiên cứu theo các quy tắc đã được xác định để có thể đánh giá, so sánh và phân tích”.

Ý nghĩa của đo lường

- Trong quá trình điều hành hoạt động, nhà quản lý, hoạch định chính sách luôn cần những thông tin trong việc ra quyết định và đo lường là cơ sở để cung cấp các tin tức có ý nghĩa giúp cho việc ra quyết định đó.
- Nhờ đo lường mà các đặc tính của sự vật được biến thành dạng mà nhà nghiên cứu có thể phân tích được, các đặc tính khác nhau đó giúp ta phân biệt các sự vật với nhau. Những đặc tính của một cá nhân và rất nhiều những hiện tượng khác đều là những quan tâm của người nghiên cứu và cần được đo lường, đánh giá. Có những đặc tính của sự vật là định lượng như chiều cao, cân nặng của một người nào đó, nhưng có nhiều đặc tính chỉ ở dạng định tính như mức độ nhận biết, thái độ của người tiêu dùng về một nhãn hiệu hàng hóa... Các cố gắng để gắn các con số cho các đặc tính sự vật là hết sức quan trọng vì các phân tích, tính toán và thống kê chỉ có thể thực hiện bằng các con số (Giới và cộng sự, 2006).

Các lỗi trong đo lường

Đo lường có thể chính xác và rõ ràng trong nghiên cứu lý tưởng. Tuy nhiên, trên thực tế, việc đo lường có thể có một số lỗi do các nguyên nhân sau:

- Cách thức trả lời: Thái độ, tâm lý của người trả lời có ảnh hưởng rất lớn đến chất lượng của thông tin thu thập, khi tiếp xúc với đối tượng phỏng vấn, thái độ miễn cưỡng, một tâm lý mệt mỏi, chán nản, lo lắng hoặc sự thiếu hiểu biết thông tin về vấn đề nghiên cứu... có thể giới

hạn khả năng của người trả lời để trả lời chính xác và đầy đủ thông tin.

- Không gian thu thập dữ liệu: Việc lựa chọn không gian phỏng vấn sẽ ảnh hưởng đến việc đo lường. Với những thông tin cần thu thập mang tính riêng tư hoặc có liên quan đến một đối tượng nào đó thì việc có một người khác hiện diện trong không gian thu thập dữ liệu sẽ có thể khiến cho người cung cấp thông tin bóp méo câu trả lời. Ví dụ: người nghiên cứu muốn xem xét lòng trung thành của nhân viên đối với tổ chức, việc lựa chọn địa điểm phỏng vấn là nơi làm việc của người được phỏng vấn có thể ảnh hưởng đến kết quả nghiên cứu, người trả lời không muốn cung cấp thông tin chính xác vì sợ bị nhận diện. Trong một số trường hợp, mặc dù đã giấu tên nhưng nếu người trả lời cảm thấy rằng việc giấu tên không được bảo đảm, họ có thể miễn cưỡng thể hiện ý kiến, cảm xúc...
- Người phỏng vấn: Người phỏng vấn có thể bóp méo câu trả lời bằng cách viết lại, sắp xếp lại các câu trả lời có thể do những nguyên nhân khách quan hoặc chủ quan. Ví dụ: người đi phỏng vấn trực tiếp không có khả năng trong việc tập hợp những ý kiến của người trả lời thông qua các câu hỏi mở hoặc không chú ý để ghi đúng câu trả lời của những câu hỏi đóng. Đôi khi, vì những nguyên nhân chủ quan mà người phỏng vấn cố tình thay đổi thông tin cung cấp bởi người được phỏng vấn. Bên cạnh đó, hành vi, phong cách, ngoại hình của người phỏng vấn có thể gây thiện cảm hoặc phản cảm đối với người trả lời.
- Phương tiện: Việc thiết kế công cụ đo lường phù hợp với mục tiêu nghiên cứu, các từ ngữ, câu hỏi được thiết kế dễ hiểu, sắp xếp câu hỏi thuận tiện cho việc trả lời sẽ giúp cho việc hạn chế những lỗi trong quá trình.

PHÁT TRIỂN KHÁI NIỆM NGHIÊN CỨU

Như đã đề cập trong chương 1, theo Kerlinger (1986), một lý thuyết khoa học là “tập những khái niệm, định nghĩa và giả thuyết trình bày có hệ thống thông qua các mối quan hệ giữa các khái niệm nhằm mục đích giải thích và dự báo các hiện tượng khoa học”. Có những khái niệm chúng ta có thể đo lường trực tiếp được, tuy nhiên, nhiều khái niệm không thể đo lường trực tiếp được mà phải thông qua một tập biến quan sát (observed variables, items, indicators...) và khi đó, biến để đại diện cho khái niệm đó gọi là biến tiềm ẩn (latent variables) và tập biến quan sát đó gọi là thang đo cho biến tiềm ẩn. Mỗi biến tiềm ẩn bao gồm hoặc tập những biến quan sát (observed variables, items, indicators...) hoặc tập các thành phần mà mỗi thành phần gồm tập những biến quan sát.

Để đo lường được khái niệm thể hiện bởi các biến tiềm ẩn, chúng ta chỉ đo lường tập các biến quan sát (thang đo) và từ đó dùng các công cụ để tính toán biến tiềm ẩn. Trong quá trình đo lường, để đảm bảo độ tin cậy và giá trị của thang đo, nhà nghiên cứu cần phải kiểm soát và hạn chế những sai số. Một số sai số cần phải xem xét và kiểm soát nhằm nâng cao độ tin cậy và giá trị của đo lường, theo Churchill (1979), có hai loại sai số đó là sai số hệ thống và sai số ngẫu nhiên được thể hiện như sau:

$$DL_{QS} = DL_T + SS_{HT} + SS_{NN}$$

Trong đó: DL_{QS} : Điểm (số) của quan sát;
 DL_T : Điểm đánh giá thực tế của mỗi đặc điểm;
 SS_{HT} : Sai số hệ thống;
 SS_{NN} : Sai số ngẫu nhiên.

Tổng sai số trong đo lường bao gồm sai số hệ thống (SS_{HT}) và sai số ngẫu nhiên (SS_{NN}). Sai số hệ thống là sai số cố định và nó sẽ giống nhau cho mỗi lần đánh giá, sai số ngẫu nhiên sẽ làm thay đổi các điểm số đánh giá trong mỗi lần đánh giá. Việc phân biệt sự khác nhau giữa sai số hệ thống và sai số ngẫu nhiên là rất quan trọng vì qua đó chúng ta có thể hiểu khả năng tin cậy và hiệu quả đánh giá.

Một thang đo cần phải đảm bảo những đặc điểm sau:

Tính dễ trả lời

Theo Giới và cộng sự (2006), tính dễ trả lời là tiêu chuẩn cần quan tâm khi phỏng vấn vấn những người sẽ cung cấp dữ liệu, nếu người được hỏi từ chối vì khó trả lời hoặc đưa ra những nhận định sai lệch về những thông tin cần thiết ở họ do cách đặt câu hỏi không phù hợp thì công việc nghiên cứu sẽ hết sức khó khăn; nhưng nếu câu hỏi quá dễ thì không thể kích thích sự hưng phấn của người trả lời.

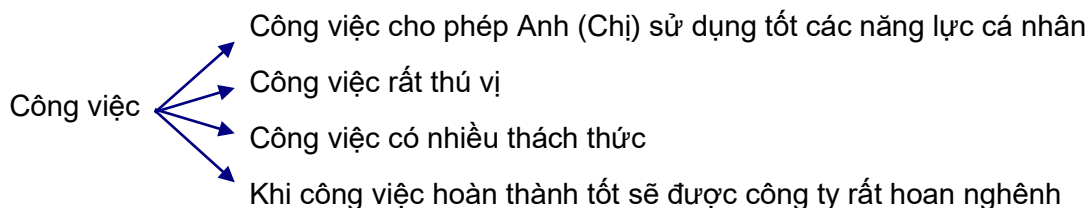
Tính đa dạng

Là tính chất của kết quả đo lường có thể được sử dụng cho nhiều mục đích như cùng để giải thích, hỗ trợ cho tiêu chuẩn giá trị của kết quả và suy đoán những ý nghĩa từ những kết quả đo lường thu thập được.

Hướng (dimensionality)

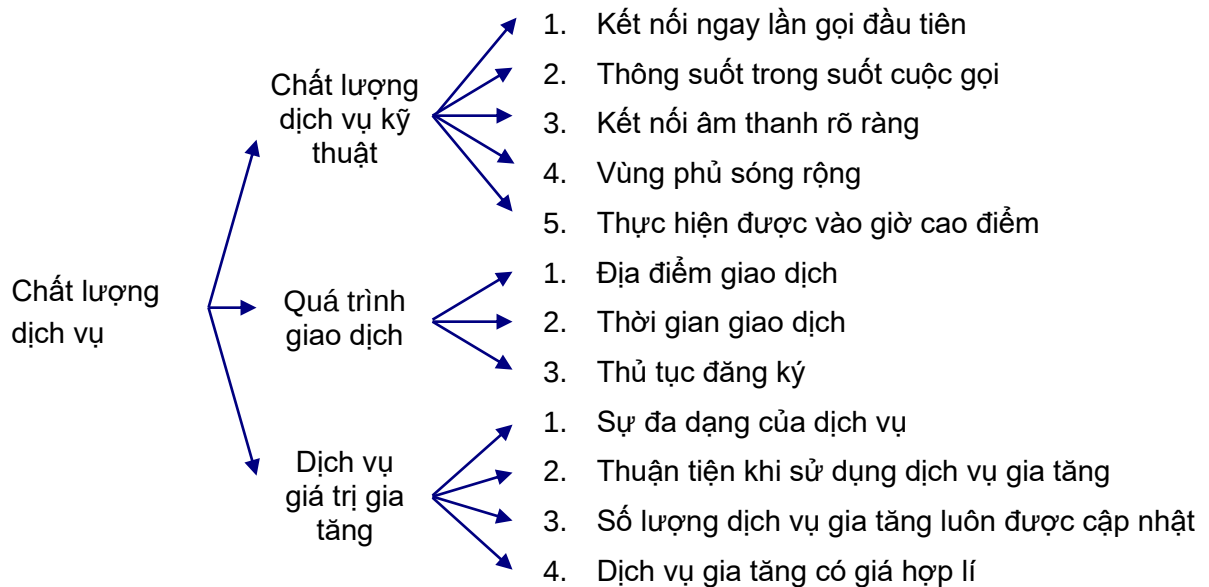
Một thang đo có thể đơn hướng (unidimensionality) tức có thể dùng một tập biến quan sát để đo lường hoặc đa hướng (multidimensionality) tức khái niệm bao gồm nhiều thành phần, nhà nghiên cứu không thể đo lường khái niệm bằng tập biến quan sát mà thông qua các thành phần, mỗi thành phần có thể bao gồm một số thành phần khác và được đo lường bởi tập biến quan sát.

Ví dụ: Thang đo công việc là một thang đo đơn hướng (unidimensionality) bao gồm tập các 4 biến quan sát (chỉ báo).



Thang đo chất lượng dịch vụ là thang đo đa hướng (multidimensionality) bao gồm 3 thành phần, mỗi thành phần bao gồm tập các biến quan sát (chỉ báo).

Ví dụ: Giả sử thang đo chất lượng dịch vụ gồm 3 thành phần (1) chất lượng dịch vụ kỹ thuật, (2) quá trình giao dịch và (3) dịch vụ giá trị gia tăng. Mỗi thành phần gồm tập các biến quan sát (chỉ báo).



Độ tin cậy (Reliability)

Một thang đo lường cung cấp những kết quả nhất quán qua những lần đo khác nhau được coi là có độ tin cậy. Những thang đo lường, những thiết bị dụng cụ đo lường có độ tin cậy cung cấp những đo lường ổn định ở những thời điểm khác nhau với những điều kiện tương đương. Vấn đề cơ bản đối với độ tin cậy là “nếu người nghiên cứu đo lường một hiện tượng lặp đi lặp lại với cùng một dụng cụ đo lường, người nghiên cứu có được kết quả tương tự hay không?”. Nếu câu trả lời là tương tự thì thang đo lường là có độ tin cậy. Độ tin cậy liên quan với mức độ liên kết những kết quả của một thang đo khi người ta lặp lại các lần đo lường. Vì thế, đo lường đảm bảo độ tin cậy là cách loại trừ sai số ngẫu nhiên, cung cấp được dữ liệu tin cậy cho việc ra quyết định và dĩ nhiên, những thông tin từ các dữ liệu có đảm bảo tin cậy mới có thể làm căn cứ cho việc ra quyết định được. Một số phương pháp để đánh giá độ tin cậy gồm:

- Độ tin cậy dựa vào kiểm tra lại các biến (test - retest): trong cách tiếp cận này người trả lời nhận dạng một bộ các nội dung thang đo để đánh giá hai lần trong điều kiện tương đối giống nhau, khoảng cách giữa hai lần đánh giá nên cách nhau từ 2 đến 4 tuần. Mức độ giống nhau giữa hai lần đánh giá sẽ được xác định bằng hệ số tương quan. Hệ số tương quan càng cao chứng tỏ khả năng tin cậy càng lớn. Tuy nhiên, phương pháp này cũng tồn tại nhiều nhược điểm, chẳng hạn như càng kéo dài thời gian giữa hai lần đánh giá thì khả năng tin cậy càng thấp.
- Độ tin cậy dựa vào các hình thức khác nhau của thang đo: phương pháp này được xây dựng dựa vào hai hình thức tương đương của thang đo. Một người trả lời cho cả hai hình thức trên ở hai thời điểm khác nhau (tốt nhất cách nhau từ 2 đến 4 tuần cho cả hai lần đánh giá). Khả năng tin cậy cũng được đánh giá bằng hệ số tương quan. Nếu hệ số tương quan

thấp, điều này phản ánh hoặc khả năng tin cậy thấp hoặc là hai hình thức của thang đo không tương đương.

- Độ tin cậy dựa vào điểm cố định: phương pháp này được sử dụng để đánh giá khả năng tin cậy của thang đo khi có nhiều nội dung được đánh giá có điểm tổng cộng chung cố định. Với thang đo loại này, mỗi nội dung đo lường ở một lĩnh vực nào đó của toàn bộ thang đo, mỗi nội dung đánh giá nên thể hiện một đặc điểm nào đó.

Phương pháp đơn giản nhất của điểm cố định là độ tin cậy được chia làm đôi, tất cả nội dung trong thang đo được chia làm hai phần, mỗi phần có tổng cộng nửa số điểm và giữa chúng phải có tương quan với nhau. Nếu có sự tương quan cao giữa hai nửa nội dung thì chứng tỏ khả năng tin cậy cao, phương pháp này được giới thiệu cụ thể trong phần kiểm định hệ số tin cậy Cronbach Alpha ở nội dung hướng dẫn SPSS.

Giá trị (Validity)

Vấn đề cuối cùng của một công cụ đo lường tốt là giá trị của đo lường, những đo lường trong nghiên cứu phải có giá trị, nếu không thì sẽ không khoa học và thông tin không có ý nghĩa hoặc có thể là ảnh hưởng không tốt cho vấn đề ra quyết định. Khi một thang đo có giá trị hoàn hảo thì không có khác biệt giữa điểm số đo lường và điểm số thực tế hay không có bất kỳ một sai số nào của đo lường ($DL_{QS} = DL_T$, $SS_{HT} = 0$, $SS_{NN} = 0$).

Một thang đo gọi là có giá trị nếu thang đo đo đáp ứng 5 loại giá trị sau: (1) giá trị nội dung (content validity); (2) giá trị hội tụ (convergent validity); (3) giá trị phân biệt (discriminant validity); (4) giá trị liên hệ lý thuyết (nomological validity) và (5) giá trị tiêu chuẩn (criterion validity) (Bagozzi, 1980).

- Giá trị nội dung (content validity) bao gồm một nội dung đánh giá có tính chất hệ thống, đại diện của một thang đo đối với nhiệm vụ đo lường hay thang đo bao phủ đầy đủ nội dung của khái niệm. Về mặt bản chất, nếu chỉ đánh giá về nội dung thì không đo lường hết giá trị của một thang đo.
- Giá trị hội tụ (convergent validity): Giá trị hội tụ là phạm vi mà thang đo có quan hệ chặt chẽ với các đo lường khác có cùng cấu trúc hay nói cách khác, nếu chúng ta đo lường một khái niệm qua hai lần đo thì số đo của hai lần phải tương quan chặt chẽ với nhau.
- Giá trị phân biệt (discriminant validity): Hai thang đo lường hai khái niệm khác nhau phải khác biệt nhau. Khi đó, hai khái niệm phân biệt thì hệ số tương quan giữa hai khái niệm phải khác 0.
- Giá trị liên hệ lý thuyết (nomological validity): Theo (Thọ, 2011), giá trị liên hệ lý thuyết nói lên mối quan hệ của khái niệm với khái niệm khác trong hệ lý thuyết đang nghiên cứu, vì vậy, chúng ta kiểm định giá trị liên hệ lý thuyết cùng với kiểm định mô hình lý thuyết nghĩa là xem xét mối quan hệ của nó với các khái niệm khác trong mô hình.
- Giá trị tiêu chuẩn (criterion validity): Theo Bollen (1989), Saunders và cộng sự (2007) thì giá trị tiêu chuẩn thể hiện mức độ liên kết của khái niệm đang nghiên cứu với một khái niệm

khác đóng vai trò là biến tiêu chuẩn để đánh giá. Dựa vào yếu tố thời gian, hiệu quả tiêu chuẩn có thể có hai dạng: hiệu quả đồng thời và hiệu quả dự báo.

- Hiệu quả đồng thời được đánh giá khi dữ liệu trên thang đo đang được đánh giá và dữ liệu trên các biến tiêu chuẩn được thu nhập có cùng một thời gian.
- Hiệu quả dự báo được đánh giá khi nhà nghiên cứu thu nhập dữ liệu trên thang đo ở một thời điểm nào đó trong hiện tại và dữ liệu trên các biến tiêu chuẩn được thu nhập vào một thời điểm trong tương lai.

QUY TRÌNH XÂY DỰNG THANG ĐO

Trong nghiên cứu khoa học, có 3 cách để có một thang đo sử dụng trong nghiên cứu:

- Sử dụng thang đo đã có: Nhà nghiên cứu sử dụng nguyên thang đo do các nhà nghiên cứu trước đã xây dựng.
- Sử dụng thang đo đã có nhưng có bổ sung, sửa chữa cho phù hợp với không gian nghiên cứu như nhà nghiên cứu có thể sử dụng thang đo nghiên cứu về chất lượng dịch vụ trong lĩnh vực ngân hàng của một nhà nghiên cứu trước thực hiện ở Thái Lan để ứng dụng triển khai nghiên cứu trong điều kiện các Ngân hàng ở Việt Nam. Trên cơ sở thang đo ở Thái Lan, nhà nghiên cứu sử dụng phương pháp nghiên cứu định tính để xem xét, bổ sung, sửa chữa (nếu có) để thang đo đó phù hợp triển khai nghiên cứu trong bối cảnh tại các ngân hàng ở Việt Nam. Dù sử dụng nguyên thang đo đã có hay thang đo đã có nhưng có bổ sung sửa chữa thì nhà nghiên cứu cũng cần phải kiểm định độ tin cậy và giá trị của thang đo trong bối cảnh nghiên cứu đó, tối thiểu, trong trường hợp này cần phải thực hiện phân tích nhân tố khám phá (EFA) và tính nhất quán nội tại của thang đo (Cronbach Alpha).
- Xây dựng thang đo hoàn toàn mới: Đối với một thang đo hoàn toàn mới, việc triển khai xây dựng cần phải đảm bảo quy trình xây dựng thang đo, một quy trình xây dựng thang đo thường được sử dụng trong các nghiên cứu là quy trình Churchill (1979).

Phần này sẽ giới thiệu mô hình sơ lược quy trình xây dựng thang đo của Churchill (1979) trên cơ sở 8 bước, đây là quy trình được rất nhiều nghiên cứu ứng dụng nhằm phát triển một thang đo:

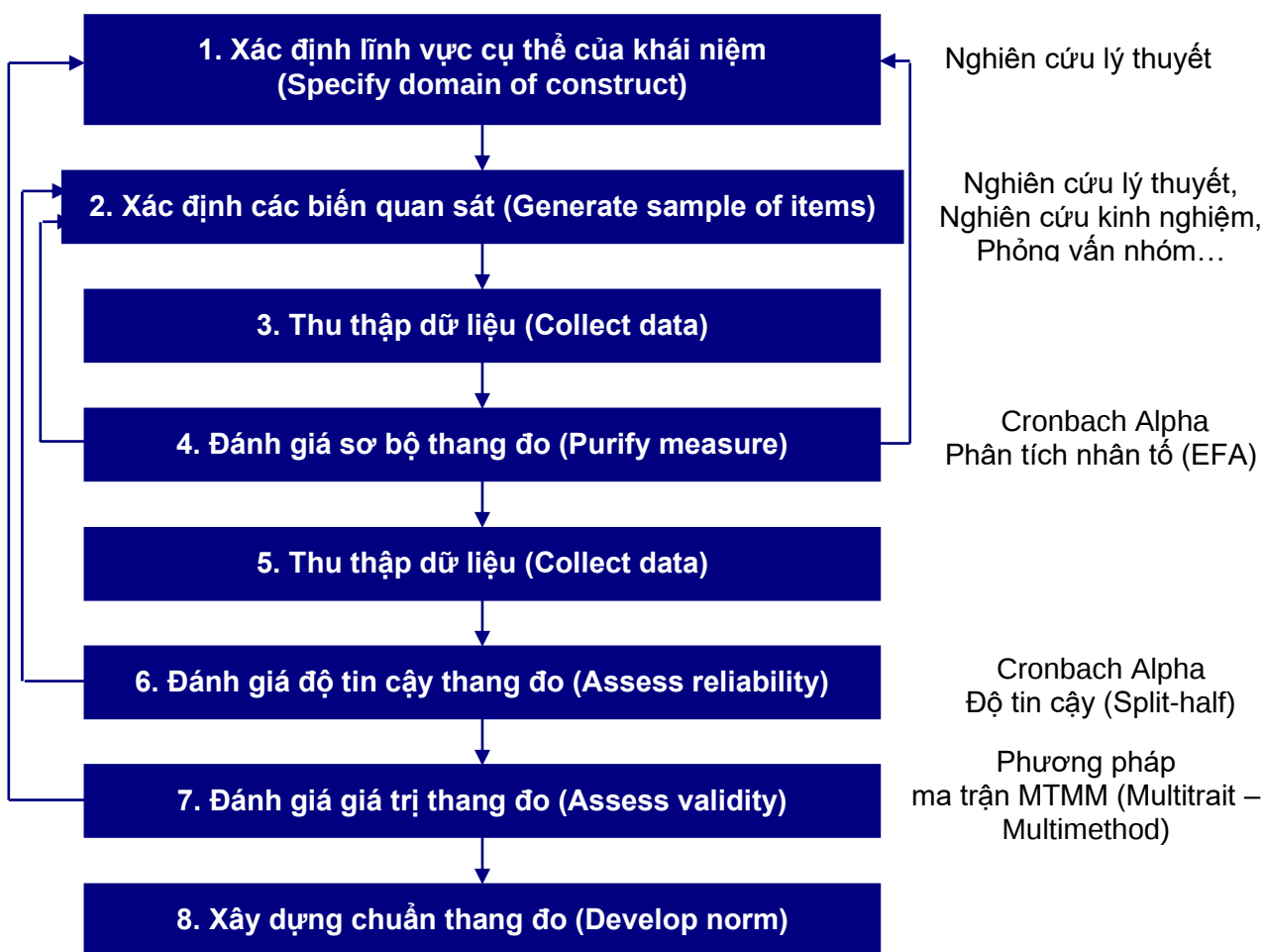
Xác định lĩnh vực và biến quan sát của khái niệm

Việc xác định lĩnh vực của khái niệm đóng vai trò quan trọng và quyết định trong việc xây dựng thang đo, một khái niệm được lựa chọn nghiên cứu cần phải có cơ sở, thông thường thông qua nghiên cứu lý thuyết, nghiên cứu định tính, dữ liệu thứ cấp... Đồng thời, nghiên cứu cũng cần chỉ rõ khái niệm được sử dụng dựa trên lý thuyết nền nào bởi vì lý thuyết nền là cơ sở để xác định nội dung của khái niệm. Trong một số trường hợp, khái niệm dựa trên lý thuyết nền này cũng có thể có nội dung hoặc giới hạn khác với lý thuyết nền của nghiên cứu khác.

Đồng thời, nghiên cứu kinh nghiệm thông qua việc trao đổi với những nhà nghiên cứu có kinh nghiệm khác sẽ giúp cho các nhà nghiên cứu kiểm tra, sửa chữa hoặc bổ sung các biến quan sát để có thể bao phủ nội dung của khái niệm. Sau đó, phỏng vấn nhóm sẽ giúp một lần nữa

hoàn chỉnh các biến quan sát cho một khái niệm, từ đó có thể phát triển bản câu hỏi để thực hiện thu thập dữ liệu nhằm đánh giá sơ bộ thang đo.

Hình 9. Quy trình xây dựng thang đo lường (Churchill, 1979)



Thu thập dữ liệu và đánh giá sơ bộ thang đo

Kết quả nghiên cứu ở bước 2 giúp cho các nhà nghiên cứu phát triển công cụ thu thập dữ liệu và thực hiện việc thu thập dữ liệu trên một mẫu (nhỏ), dữ liệu thu thập trên mẫu này dùng để đánh giá sơ bộ định lượng thang đo. Trước tiên, cần đánh giá tính nhất quán nội tại (internal consistency) bằng hệ số Cronbach Alpha (được hướng dẫn cụ thể trong phần SPSS) nhằm loại các biến “rác” (garbage items) và phân tích nhân tố khám phá (EFA) để kiểm định tính đơn hướng, giá trị hội tụ và giá trị phân biệt của thang đo (Churchill, 1979). Kết quả của giai đoạn này sẽ làm cơ sở cho việc phát triển bản câu hỏi để thu thập dữ liệu chính thức.

Thu thập dữ liệu, đánh giá độ tin cậy và giá trị thang đo

Thu thập dữ liệu chính thức cần thực hiện trên một mẫu lớn nhằm kiểm định thang đo bằng cách tiếp tục thực hiện phân tích nhân tố khám phá (EFA) và Cronbach Alpha. Chú ý, trong quy trình của Churchill (1979) không khẳng định bước này nên thực hiện trước hay sau, tuy nhiên, theo kinh nghiệm của tác giả, phân tích nhân tố khám phá (EFA) nên thực hiện trước và kiểm định Cronbach Alpha nên thực hiện đối với các chỉ báo tương ứng với kết quả của các thành

phần trong phân tích nhân tố.

Churchill (1979) đề nghị dùng ma trận MTMM (Multitrait – Multimethod) để kiểm định giá trị của thang đo, tuy nhiên, có thể sử dụng phân tích nhân tố khẳng định (Confirmation Factor Analysis - CFA) để thực hiện.

Phát triển chuẩn thang đo

Sau khi thực hiện việc phát triển khái niệm, xác định các chỉ báo làm cơ sở cho việc nghiên cứu định tính, nghiên cứu sơ bộ định lượng và nghiên cứu chính thức định lượng, các chỉ báo còn lại (sau khi đã loại bỏ những chỉ báo không phù hợp trong các kiểm định giá trị và độ tin cậy) sẽ được sử dụng để phát triển chuẩn thang đo và đây là thang đo chính thức cho các khái niệm.

Chú ý: Theo kinh nghiệm của tác giả, trong một số nghiên cứu sử dụng thang đo cũ, việc kiểm định thang đo lường nên được thực hiện với phân tích nhân tố khám phá trước, sau đó, kết quả của nhân tố khám phá được sử dụng để kiểm tra tính nhất quán nội tại Cronbach Alpha.

CẤP ĐỘ THANG ĐO

Trong nghiên cứu, để đo lường các sự vật hiện tượng, nhà nghiên cứu có thể thực hiện bằng cách sử dụng các cấp độ thang đo. Có 4 loại cấp độ thang đo là biểu danh (nominal scales), thứ tự (ordinal scales), khoảng cách (interval scales) và tỷ lệ (ratio scales).

Cấp độ thang đo biểu danh (nominal scale)

Cấp độ thang đo biểu danh (hay định danh) là loại thang đo cơ bản nhất trong 4 cấp độ thiết kế thang đo, trong cấp độ này, cấp độ thang đo sử dụng các con số đánh dấu (mã số) để phân loại đối tượng hoặc sử dụng như ký hiệu để phân biệt và nhận dạng đối tượng. Cấp độ thang đo biểu danh không có ý nghĩa về mặt lượng mặc dù nó được ký hiệu bằng các con số.

Trong nghiên cứu, cấp độ thang đo biểu danh được sử dụng cho việc nhận dạng, xếp loại người được phỏng vấn (giới tính, nghề nghiệp, tôn giáo, nơi làm việc...); tên sản phẩm, dịch vụ...

Ví dụ:

1. Anh (Chị) đang công tác tại bộ phận:

☐ Kinh doanh

☐ Kỹ thuật

☐ Văn phòng

☐ Kế toán

☐ Khác, cụ thể là:

2. Giới tính:

☐ Nam

☐ Nữ

3. Anh (Chị) đang công tác tại:

☐ Miền Bắc

☐ Miền Trung

☐ Miền Nam

Một số câu hỏi thường gặp với cấp độ thang đo định danh:

Ví dụ 1: Câu hỏi cho một lựa chọn là câu hỏi mà người trả lời chỉ chọn một phương án mà họ cho là đúng nhất.

a. Anh (Chị) đang công tác tại bộ phận:

- ☐ Kinh doanh ☐ Kỹ thuật ☐ Văn phòng
☐ Kế toán ☐ Khác, cụ thể là:

b. Anh (Chị) thường xuyên sử dụng loại nước uống nào dưới đây (chỉ chọn một loại nước uống mà Anh (Chị) thường dùng):

- ☐ Coca-cola ☐ Pepsi ☐ 7up
☐ Trà xanh không độ ☐ Nước khoáng.

Ví dụ 2: Câu hỏi cho nhiều lựa chọn là câu hỏi mà người trả lời có thể một hoặc nhiều phương án thích hợp:

Hằng tháng, Anh (Chị) đọc loại báo nào dưới đây (có thể chọn nhiều loại báo):

- ☐ Hà Nội mới
☐ Lao động
☐ Tiền phong
☐ Thanh niên
☐ Tuổi trẻ

Cấp độ thang đo thứ tự (ordinal scales)

Cấp độ thang đo thứ tự cung cấp thông tin về mối quan hệ thứ tự, so sánh giữa các đối tượng, cấp độ của thang đo này bao gồm cả thông tin về sự biểu danh và xếp hạng theo thứ tự. Nó cho phép xác định một đặc tính của một sự vật này có hơn một sự vật khác hay không nhưng không cho phép chỉ ra mức độ sự khác biệt này. Cấp độ thang đo này cũng không có ý nghĩa về mặt lượng (không cho biết nhiều hay ít hơn bao nhiêu; chỉ cho biết cấp độ chênh lệch). Ví dụ như xếp hạng về chất lượng sản phẩm, phân loại các hiện tượng. Thang đo thứ tự được dùng rất phổ biến trong nghiên cứu để đo lường thái độ, ý kiến, quan điểm, nhận thức và sở thích.

Theo Giới và cộng sự (2006), cũng như thang định danh, các con số trong thang thứ tự được gán một cách quy ước. Nhưng vì các hạng mục tự nó có trật tự nên hệ thống số phải có trật tự. Người nghiên cứu có thể biến đổi một thang thứ tự mà không làm thay đổi tính chất của nó. Không nhất thiết phải gán các số từ 1 đến 5 để đạt được một sự sắp xếp về sự ưa thích cho 5 cửa hàng nói trên, người nghiên cứu có thể gán các con số 10, 25, 30, 45 và 50, và việc này là hoàn toàn hợp lệ, tuy nhiên, người nghiên cứu cần phải chú ý là nếu sự vật hiện tượng nào được gán bởi số 10 thì sẽ được ưa thích ít hơn so với được gán số 25, 30... nhưng nó không có nghĩa là 25 sẽ hơn 10 là 15 đơn vị và cũng không có ý nghĩa là 25 gấp 10 là 2,5 lần. Người nghiên cứu cần chú ý và xem xét đặc tính này để so sánh với loại thang đo tỷ lệ được đề cập ở phần sau. Cũng giống như thang định danh, các phép toán số học thông dụng như: cộng, trừ, nhân, chia không thể áp dụng trong thang điểm thứ tự. Thống kê một biến thường dùng cho thang điểm biểu danh là phân tích tần suất, phần trăm, trung vị trong khi đó thống kê hai biến là những phương pháp thống kê dựa trên các quan hệ thứ tự. Sau đây là một số ví dụ cho cấp độ thang đo thứ tự:

Ví dụ 1: Xin Anh (Chị) cho biết mức độ hiểu biết về các dịch vụ cung cấp trên chuyến bay của Vietnam Airlines:

- ☐ Hoàn toàn hiểu về dịch vụ,
- ☐ Hiểu về dịch vụ,
- ☐ Khá hiểu biết,
- ☐ Không hiểu biết,
- ☐ Hoàn toàn không hiểu biết.

Ví dụ 2: Câu hỏi bắt buộc sắp xếp thứ tự

Một người nghiên cứu đang muốn thăm dò sự ưa thích của khách hàng về 6 cửa hàng mà họ đang xem xét ở ví dụ trên bằng cách đề nghị người trả lời xếp hạng ưa thích của họ đối với các cửa hàng đó theo thứ tự ưa thích nhất thì người trả lời sẽ xếp thứ 1 tiếp theo là thứ 2, 3, 4 và 6 cho từng cửa hàng. Cụ thể:

Xin Anh (Chị) sắp xếp thứ tự ưu tiên từ 1 đến 6 với 1 là cửa hàng mà Anh (Chị) thích nhất và 6 là cửa hàng Anh (Chị) không thích:

- Cửa hàng A _____
- Cửa hàng B _____
- Cửa hàng C _____
- Cửa hàng D _____
- Cửa hàng E _____
- Cửa hàng F _____

Ví dụ 3: Câu hỏi so sánh từng cặp

Trong các yếu tố dưới đây, xin Anh (Chị) đánh số 1 vào ô (☐) với thuộc tính mà anh chị cho là quan trọng hơn theo từng cặp:

- | | | | |
|-----------|--------------------------|-----------|--------------------------|
| Giá | ☐ | Độ bền | ☐ |
| Giá | ☐ | Kiểu dáng | ☐ |
| Giá | ☐ | Màu sắc | ☐ |
| Độ bền | ☐ | Màu sắc | ☐ |
| Kiểu dáng | ☐ | Màu sắc | ☐ |

Cấp độ thang đo khoảng cách (interval scales)

Cấp độ thang đo khoảng cách là cấp độ thang đo cũng có thể dùng để xếp hạng các đối tượng nghiên cứu nhưng khoảng cách bằng nhau trên thang đo đại diện cho khoảng cách bằng nhau trong đặc điểm của đối tượng. Một thang đo khoảng chứa đựng tất cả thông tin trong thang đo thứ tự nhưng nó cũng cho phép so sánh sự khác biệt giữa các đối tượng. Hơn nữa, trong thang đo khoảng, có một khoảng cách cố định hoặc bằng nhau giữa các giá trị thuộc về thang đo. Tuy nhiên, trong thang đo khoảng vị trí của điểm "0" thì không cố định.

Ví dụ 1: Xin vui lòng cho biết số lần trong tháng Anh (Chị) giao dịch với ngân hàng A:

- ☐ Không lần
- ☐ Từ 1 đến 3 lần
- ☐ Từ 4 đến 6 lần

☐ Từ 7 đến 9 lần ☐ Từ 10 đến 12 lần ☐ Trên 12 lần

Ví dụ 2: Anh (Chị) sống tại địa chỉ này khoảng:

☐ 1 đến 5 năm ☐ Từ 5 đến 10 năm ☐ Từ 10 đến 15 năm
☐ Từ 15 đến 20 năm ☐ Trên 12 năm

Ví dụ 3: Thu nhập hàng tháng của Anh (Chị) khoảng:

☐ Từ 0 đến 3 triệu ☐ Từ 3 đến 6 triệu ☐ Từ 6 đến 9 triệu
☐ Từ 9 đến 12 triệu ☐ Trên 12 triệu

Ví dụ 4: Xin vui lòng cho biết ý kiến của các Anh (Chị) với (1) là rất không đồng ý và (7) là rất đồng ý các nhận định được cho dưới đây:

	1	2	3	4	5	6	7
Công việc cho phép Anh (Chị) sử dụng tốt các năng lực cá nhân	☐	☐	☐	☐	☐	☐	☐

Trong ví dụ trên, sự khác nhau giữa 1 và 2 bằng sự khác nhau giữa 3 và 4, và dĩ nhiên sự khác nhau giữa 2 và 4 bằng 2 lần sự khác nhau giữa 1 và 2.

Cấp độ thang đo tỷ lệ (ratio scales)

Cấp độ thang đo tỷ lệ có tất cả các đặc điểm của thang định danh, thang thứ tự và thang khoảng và ngoài ra nó còn có điểm 0 (zero) cố định. Do vậy, với thang điểm này người nghiên cứu có thể xác định, xếp hạng thứ tự, so sánh các khoảng cách hay những sự khác biệt và cho phép tính toán tỷ lệ giữa các giá trị của thang đo. Người nghiên cứu có thể nói đến các khái niệm gấp đôi, một nửa.... trong thang đo này.

Vì vậy, trong thang đo tỷ lệ chúng ta có thể nhận dạng hoặc phân loại đối tượng, xếp hạng đối tượng và so sánh sự khác biệt. Thang đo tỷ lệ không chỉ cho biết sự khác biệt giữa 1 và 3 bằng với sự khác biệt giữa 2 và 4 mà còn cho biết 4 gấp 2 lần 2, hoặc một người có thời gian công tác là 10 năm sẽ gấp 2 lần với một người có thời gian công tác là 2 năm.

Trong nghiên cứu, cấp độ thang đo tỷ lệ thường dùng để đo lường chiều cao, trọng lượng, tuổi, thu nhập của các cá nhân, mức bán, doanh số của doanh nghiệp hoặc mức giá mà người tiêu dùng sẵn sàng trả cho sản phẩm. Cấp độ thang đo tỷ lệ là loại thang đo có thể áp dụng tất cả các phương pháp thống kê. Những thuận lợi của việc lượng hóa này có thể được bổ sung bởi các dữ liệu do 3 loại thang điểm trên, do vậy cần cân nhắc để lựa chọn loại thang điểm thích hợp và sử dụng chúng.

Ví dụ 1: Xin khoanh tròn số người trong gia đình Anh (Chị):

0 1 2 3 4 5 6 7 (nếu trên 7, cụ thể là người).

Ví dụ 2: Số ngày mà Anh (Chị) làm việc tại công ty trong một tuần là ngày.

Ví dụ 3: Tuổi của Anh (Chị) là..... (tuổi).

Ví dụ 4: Chi tiêu trung bình của gia đình Anh (Chị) khoảng triệu đồng.

Ví dụ 5: Câu hỏi sử dụng thang có tổng số không đổi

Xin cho biết Anh (Chị) bố trí thời gian cho một ngày:

Học bài	_____
Chuẩn bị bữa ăn	_____
Đọc truyện	_____
Xem phim	_____
Ngủ	_____
Công việc khác	_____
Tổng:	24 (giờ)

Chú ý, hai cấp độ thang đo biểu danh và thứ tự được gọi là cấp độ thang đo định tính (hay cấp độ thấp), hai cấp độ thang đo khoảng và tỷ lệ gọi là cấp độ thang đo định lượng (cấp độ mạnh). Cấp độ thang đo định lượng (cấp độ mạnh) thường cho nhiều thông tin hơn cấp độ nghiên cứu định tính bởi (cấp độ thấp) vì câu hỏi sử dụng cấp độ thang đo định lượng có thể chuyển về cấp độ thang đo định tính nhưng không xảy ra trường hợp ngược lại. Vì vậy, nhà nghiên cứu cần phải cân nhắc khi sử dụng cấp độ thang đo để thiết kế các câu hỏi sao cho phù hợp với thông tin cần thu thập để nghiên cứu, tránh trường hợp đã sử dụng cấp độ thang đo định tính để thiết kế và thu thập dữ liệu, sau khi đã thu thập dữ liệu xong mới nhận ra rằng dữ liệu thu thập bằng thang đo định tính không thể đáp ứng đủ thông tin của quá trình nghiên cứu.

Ví dụ 1: Xin vui lòng cho biết ý kiến của các Anh (Chị) với (1) là rất không đồng ý và (7) là rất đồng ý nhận định được cho dưới đây:

	1	2	3	4	5	6	7
Công việc cho phép anh/chị sử dụng tốt các năng lực cá nhân	☐	☐	☐	☐	☐	☐	☐

Nhà nghiên cứu đã sử dụng cấp độ thang đo khoảng (cấp độ thang đo định lượng – cấp độ mạnh), với thang đo này trong quá trình phân tích, nếu có nhu cầu, có thể chuyển thành cấp độ thang đo định tính bằng cách mã hóa lại với 1 đến 3 là không đồng ý (0) và 4 đến 7 là đồng ý (1). Ngược lại, nếu câu hỏi thiết kế 1 là đồng ý và 0 là không đồng ý, khi có nhu cầu, nhà nghiên cứu không thể chuyển lại thành thang đo cấp định lượng từ 1 đến 7 được.

Ví dụ 2: Tuổi của Anh (Chị) là..... (tuổi).

Với câu hỏi này, đây là cấp độ thang đo tỷ lệ, tuy nhiên, nếu có nhu cầu, nhà nghiên cứu có thể chuyển từ cấp độ thang đo khoảng, ví dụ:

- | | | |
|---|---|---|
| ☐ Nhỏ hơn 20 tuổi | ☐ Từ 20 đến 40 năm | ☐ Từ 40 đến 15 năm |
| ☐ Từ 15 đến 20 năm | ☐ Trên 12 năm | |

MỘT SỐ VẤN ĐỀ KHI THIẾT KẾ THANG ĐO LƯỜNG

Vấn đề khi phát triển khái niệm nghiên cứu

Khái niệm nghiên cứu phải có định nghĩa đưa vào nghiên cứu rõ ràng, phải được xác định là khái niệm nghiên cứu đơn hướng hay đa hướng. Cần tránh trường hợp diễn tả hai hướng

(dimension) của một khái niệm nghiên cứu như thể chúng là một. Ví dụ, khi nghiên cứu về quan niệm của khách hàng về chất lượng dịch vụ, đừng cố gắng kết hợp năng lực công nghệ của nhà cung cấp dịch vụ và năng lực dự đoán thành một thành phần/hướng (dimension). Giá trị của khái niệm nghiên cứu phải được đánh giá trước khi tiến hành phát triển thang đo, nhà nghiên cứu cần phải thực hiện đánh giá trước khi thực hiện việc phân tích để đảm bảo giá trị của dữ liệu thu thập.

Vấn đề liên quan đến thang đo

Sử dụng từ ngữ rõ ràng, tránh mập mờ trong các câu hỏi. Chỉ dẫn rõ ràng, tránh dùng những từ, những cụm từ mang tính dẫn dắt (leading word) trong các câu hỏi. Các câu hỏi liên quan đến nhiều đặc tính/chỉ báo, đảm bảo rằng các chỉ báo được phát biểu một chiều.

Số lượng mục lựa chọn

Việc xác định số mục lựa chọn cho cấp độ thang đo được thiết kế trong mỗi câu hỏi tùy thuộc vào mục tiêu nghiên cứu và cách thức phân tích. Đơn giản nhất là chỉ để hai mục đối nhau (đồng ý và không đồng ý; đúng và sai; có và không). Tuy nhiên, loại thiết kế rất khó cho việc làm rõ thông tin thu thập, hoàn toàn là những thông tin định tính vì ngoài thông tin đồng ý, không đồng ý; nhà nghiên cứu không biết thông tin biểu thị được “mức độ” đồng ý hoặc không đồng ý. Nếu sử dụng nhiều khoản mục sẽ cho phép người được hỏi có nhiều sự lựa chọn rộng rãi và có thể phân tích mức độ khác biệt trong sự trả lời hơn là chỉ có hai khoản mục. Tuy nhiên, nếu nhiều mục quá lại sinh rắc rối cho việc lựa chọn và các nhà nghiên cứu có kinh nghiệm cho rằng câu hỏi có 5 hoặc 6 mục trả lời là phù hợp hơn cả.

Cân bằng 2 “cực” trả lời

Một vấn đề được đặt ra là một câu hỏi nên hay không nên có những mục trả lời thuận lợi (positive) và bất thuận lợi (negative) trong mỗi câu hỏi là ngang nhau? Ý kiến này được nêu lên nhằm mục đích tránh việc đặt nặng sự trả lời nghiêng về “cực” này hay “cực” kia. Kinh nghiệm cho thấy, nên có sự cân bằng hai cực để tránh sự thiên lệch trong thông tin thu thập.

Ví dụ: Xin Anh (Chị) cho biết đánh giá của mình về phương pháp quản lý của Trưởng phòng Kinh doanh:

☐ Tuyệt ☐ Rất tốt ☐ Tốt ☐ Khá ☐ Trung bình

Khi đó, nếu người được hỏi không thích hoặc cho rằng phương pháp quản lý của Trưởng phòng Kinh doanh là không tốt thì họ không biết phải trả lời thể nào trên thang điểm đánh giá này và rõ ràng thông tin thu được có sự bất cân xứng và hoàn toàn thiên lệch về thuận lợi (positive) đối với đối tượng được nghiên cứu.

Số các mục trả lời là chẵn hay lẻ

Tùy thuộc vào mục tiêu nghiên cứu, người nghiên cứu có thể quyết định số mục trả lời là chẵn hay lẻ. Nếu số mục trả lời lẻ thì người trả lời dễ tiến tới thái độ trung dung bằng cách lựa chọn mục trả lời ở giữa và nếu số mục trả lời chẵn thì người được hỏi bắt buộc phải biểu lộ thái độ của mình, ít nhất cũng là ở mức độ nào đó. Trong thực tế, qua các kết quả thử nghiệm người

ta nhận thấy không có sự sai biệt đáng kể nếu dùng số mục trả lời chẵn hoặc lẻ nếu thang đo được thiết kế tốt.

Sự lai ghép giữa cấp độ thang đo khoảng cách và thứ tự (ordinally interval hybrid scales)

Trong một số trường hợp thực hiện nghiên cứu cần phải thu thập những vấn đề mang tính thái độ, hành vi, ý định của đối tượng, thang đo cần sử dụng là cấp độ thang đo thứ tự (ordinal) nhưng được giả định là cấp độ thang đo khoảng nhằm thuận tiện cho việc sử dụng những phân tích thống kê chuyên sâu. Như vậy, thang đo khoảng cách thứ tự (ordinally interval scale) là việc chuyển “giả” cấp độ thang đo thứ tự vào cấp độ thang đo khoảng của nhà nghiên cứu. Một nhà nghiên cứu quan tâm đến việc nghiên cứu thái độ đơn đơn lẻ mà phối hợp đo lường thái độ tổng quát với một nhóm các vấn đề tài một thời điểm. Để thực hiện điều này, nhà nghiên cứu phải có thể thực hiện cùng một lúc nhiều mục hỏi với cùng một phần mục đánh giá, sau đó có thể tính “trung bình” các câu trả lời theo các nhóm chủ đề. Trong trường hợp này, khi thiết kế và phân tích, nhà phân tích sử dụng hoàn toàn giống cấp độ thang đo khoảng chứ không phải là thang đo thứ tự. Một ví dụ được chuyển từ thang đo cấp độ thang đo thứ tự sang thang đo lai ghép giữa khoảng và thứ tự và sau đây giới thiệu câu hỏi sử dụng cấp độ thang đo thứ tự (ví dụ 1) sang ví dụ 2, 3 là cấp độ lai ghép khoảng cách thứ tự (ordinally interval scale).

Ví dụ 1: Xin cho biết ý kiến bằng cách đánh (✓) vào ô mà Anh (Chị) cho là đúng nhất với các phát biểu dưới đây:

Phát biểu		Hoàn toàn không đồng ý	Không đồng ý	Không ý kiến	Đồng ý	Hoàn toàn đồng ý
1.	Anh (Chị) được trả lương cao					
2.	Anh (Chị) thường được tăng lương					
3.	Anh (Chị) có thể sống hoàn toàn dựa vào thu nhập từ công ty					
4.	Tiền lương tương xứng với kết quả làm việc					
5.	Tiền lương, thu nhập được trả công bằng					

Để hình thành thang đo lai ghép giữa khoảng và thứ tự, nhà nghiên cứu cần phải gán “điểm” 1, 2, 3, 4, 5 cho mỗi cấp độ phương án trả lời, khi đó, hoàn toàn không đồng ý gán với 1, không đồng ý gán với 2, không ý kiến gán với 3, đồng ý gán với 4 và hoàn toàn đồng ý gán với 5 và câu hỏi được thiết kế như sau:

Ví dụ 2: Xin cho biết ý kiến bằng cách đánh khoanh tròn vào số mà Anh (Chị) cho là đúng nhất với các phát biểu dưới đây:

Phát biểu

	Hoàn toàn không đồng ý	Không đồng ý	Không ý kiến	Đồng ý	Hoàn toàn đồng ý
1. Anh (Chị) được trả lương cao	1	2	3	4	5
2. Anh (Chị) thường được tăng lương	1	2	3	4	5
3. Anh (Chị) có thể sống hoàn toàn dựa vào thu nhập từ công ty	1	2	3	4	5
4. Tiền lương tương xứng với kết quả làm việc	1	2	3	4	5
5. Tiền lương, thu nhập được trả công bằng	1	2	3	4	5

Đôi khi, người ta chỉ thiết kết phần mô tả chỉ ở 2 cực, ví dụ với (1) là rất không đồng ý hoặc hoàn toàn không đồng ý và (5) là rất đồng ý hoặc hoàn toàn không đồng ý

Ví dụ 3: Xin vui lòng cho biết ý kiến của các Anh (Chị) bằng cách đánh dấu (✓) vào các ô vuông tương ứng với (1) là rất không đồng ý và (5) là rất đồng ý các nhận định được cho dưới đây:

	1	2	3	4	5
1. Anh (Chị) được trả lương cao	☐	☐	☐	☐	☐
2. Anh (Chị) thường được tăng lương	☐	☐	☐	☐	☐
3. Anh (Chị) có thể sống hoàn toàn dựa vào thu nhập từ công ty	☐	☐	☐	☐	☐
4. Tiền lương tương xứng với kết quả làm việc	☐	☐	☐	☐	☐
5. Tiền lương, thu nhập được trả công bằng	☐	☐	☐	☐	☐

Câu hỏi nhảy quãng (skip questions)

Câu hỏi nhảy quãng có thể xuất hiện ở bất kỳ vị trí nào trong bản câu hỏi nếu câu hỏi hoặc một tập hợp các câu hỏi tiếp theo chỉ được trả lời bởi những đáp viên có đủ điều kiện cần thiết trước đó. Cần tránh các câu hỏi nhảy quãng này đến mức tối đa, nếu câu hỏi nhảy quãng cần thiết thì phải hướng dẫn rõ ràng cho đáp viên trả lời.

Ví dụ: Bạn đã từng đi máy bay của Vietnam Airlines chưa?

- ☐ Đã → Chuyển sang câu số 5
☐ Chưa → Chuyển sang câu số 9.

Thang đo thái độ (attitude scale measurement)

Như đã giới thiệu ở chương trước, có những vấn đề rất dễ đo lường như độ cao, cân nặng, số năm đào tạo... nhưng có những vấn đề đo lường rất phức tạp như thái độ hay suy nghĩ của con người. Phần này sẽ giới thiệu thang đo lường thái độ (attitude scale measurement) để thu thập

những dữ liệu về thái độ, cảm xúc và ý định về hành vi của con người.

Bản chất của thái độ

Theo Hair và cộng sự (2002):

“Thái độ là một khuynh hướng phản ứng một cách tích cực hay tiêu cực đối với một đối tượng, ý tưởng hoặc một tập hợp thông tin cụ thể.”

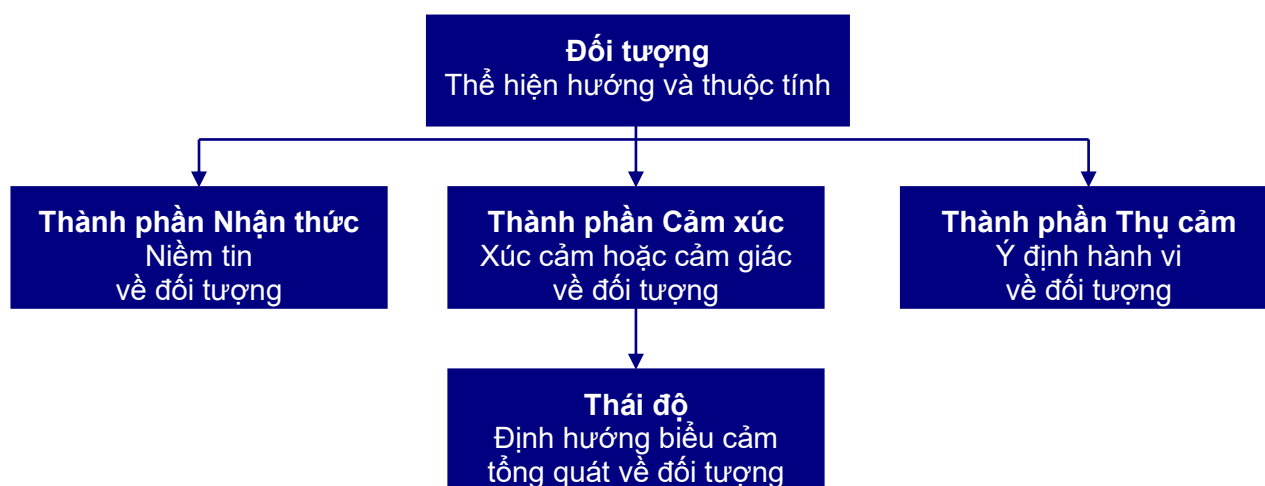
“Attitude: A learned predisposition to react in some consistent positive or negative way to a given object, idea, or set of information”.

Cũng theo Hair và cộng sự (2002), thái độ không phải là một khái niệm nghiên cứu dễ dàng đo lường. Cấu trúc thật sự của thái độ nằm trong tâm trí của cá nhân có thái độ đó. Hiện nay, có hai cách tiếp cận về cấu trúc thái độ đó là: cách tiếp cận ba thành phần (trilogy approach) và cách tiếp cận tâm lí (affect global approach).

Cách tiếp cận ba thành phần (trilogy approach)

Cách tiếp cận này cho rằng sự thấu hiểu về thái độ của một người đối với một đối tượng, một người, hoặc một hiện tượng đòi hỏi sự hiểu biết về thành phần nhận thức (cognitive), cảm xúc (affective) và thụ cảm (conative). Chính sự tích hợp của ba thành phần trên cho phép một người hình thành thái độ về một đối tượng nào đó.

Hình 10. Ba thành phần của thái độ



Nhận thức (cognitive)

Nhận thức biểu thị niềm tin, quan niệm và kiến thức về một đối tượng cụ thể và các đặc tính của nó, những khía cạnh này là những thành phần chính và là kết quả của việc học hỏi. Ví dụ, nhìn lại việc công ty ABC thực hiện trách nhiệm xã hội của doanh nghiệp mình (CSR). Người tiêu dùng có những niềm tin, quan niệm khác nhau về sự việc này, những niềm tin, quan niệm này sẽ phát triển và thay đổi khi họ có nhiều thông tin hơn về sự việc. Một số những niềm tin mà người tiêu dùng có thể có:

- Công ty ABC kinh doanh góp phần thể hiện trách nhiệm xã hội.

- Công ty ABC thể hiện trách nhiệm xã hội của doanh nghiệp sẽ đóng góp cho sự phát triển của cộng đồng.

Cảm xúc (affective)

Thành phần cảm xúc biểu thị những cảm xúc, tình cảm đối với một đối tượng cụ thể. Ví dụ, khi một người xem xét việc mua một chiếc xe ô tô có thể xác định một số yếu tố quan trọng liên quan đến xe để dựa vào đó đưa ra lựa chọn. Thành phần cảm xúc trong thái độ cho phép người đó quyết định yếu tố nào trong các yếu tố đã được xác định (ví dụ: sức mạnh động cơ, hệ thống an toàn bảo đảm an toàn, tiết kiệm xăng,...) là yếu tố quan trọng hơn (hoặc ít quan trọng hơn) liên quan đến một chiếc xe cụ thể. Cảm xúc đối với một đối tượng cụ thể dựa vào một tập hợp những niềm tin được tạo lập về đối tượng đó.

Thụ cảm (conative)

Thành phần thụ cảm của thái độ liên quan đến ý định hành vi hoặc hành vi thật sự của người tiêu dùng. Thành phần này của thái độ là kết quả có thể quan sát được hình thành bởi sự tương tác giữa thành phần nhận thức (niềm tin) và thành phần cảm xúc (sức mạnh cảm xúc của những niềm tin đó) liên quan đến một đối tượng cụ thể.

Điểm chính về cách tiếp cận ba thành phần là việc đo lường thái độ một cách toàn diện không thể đạt được bằng việc sử dụng một thang đo một chỉ báo (single-item scale) hoặc nhiều chỉ báo (multiple-item scale) mà đòi hỏi việc phát triển mô hình đa thành phần. Để đo lường thái độ, người nghiên cứu phải thu thập những loại dữ liệu (nhận thức, cảm xúc và thụ cảm) về đối tượng và những đặc tính của nó và sau đó, thông qua quá trình phát triển mô hình, hình thành nên điểm thái độ tổng hợp.

Có hai mô hình được sử dụng phổ biến là thái độ đối với đối tượng (attitude-toward-object) và thái độ đối với hành vi (attitude-toward-behavior). Sau đây là ví dụ về sự khác nhau giữa nhận thức, cảm xúc và sự thụ cảm:

- Nhận thức: Tôi nghĩ dùng Internet để liên lạc hiệu quả hơn dùng điện thoại (cố định).
- Cảm xúc: Tôi không thích dùng điện thoại (cố định).
- Thụ cảm: Tôi dự định dùng Internet nhiều hơn trong quá trình thực hiện công việc.
 - Mô hình thái độ đối với đối tượng (attitude-toward-object)

Ở mô hình này, để đo lường thái độ của người tiêu dùng đối với một đối tượng cụ thể, người nghiên cứu đo lường nhận thức (niềm tin) và cảm xúc của họ đối với đối tượng đó. Ví dụ: Thang đo lường sử dụng để đo lường thái độ đối với xe ô tô Honda Civic.

Đo lường nhận thức (niềm tin)

Xin vui lòng nghĩ đến những trải nghiệm mà Anh (Chị) có khi lái xe Honda Civic và cho chúng tôi biết ý kiến của Anh (Chị) về mỗi yếu tố sau. Đối với mỗi yếu tố, vui lòng khoanh tròn con số diễn tả suy nghĩ của Anh (Chị) về những yếu tố sau liên quan đến xe Honda Civic. Đối với yếu tố nào không phù hợp với đánh giá của Anh (Chị) (do chưa trải nghiệm

qua hoặc không có thông tin), vui lòng khoanh tròn số 0 có nghĩa là không thích hợp trong trường hợp của Anh (Chị) (N/A)

	Thực sự kém	Kém	Trung bình	Tốt	Rất tốt	Không áp dụng (N/A)
Độ êm khi lái	1	2	3	4	5	0
Tiếng máy nổ	1	2	3	4	5	0
Tiết kiệm xăng	1	2	3	4	5	0
Độ an toàn	1	2	3	4	5	0

Đo lường cảm xúc (mức độ quan trọng)

Vui lòng cho điểm vào chỗ trống từ 1-6 theo mức độ mà Anh (Chị) cảm thấy quan trọng đối với mỗi đặc tính sau khi chọn mua một chiếc xe ô tô:

Không quan trọng	Hơi quan trọng	Phần nào quan trọng	Nhìn chung là quan trọng	Chắc chắn là quan trọng	Cực kỳ quan trọng
(1)	(2)	(3)	(4)	(5)	(6)
Độ êm khi lái xe					
Tiếng máy nổ					
Tiết kiệm xăng					
Độ an toàn					

- Mô hình thái độ đối với hành vi (attitude-toward-behavior)

Với mô hình này, người nghiên cứu mong muốn biết được hành vi hoặc ý định hành vi của người tiêu dùng đối với một đối tượng cụ thể hơn là biết được bản thân thái độ. Điểm quan trọng ở đây là những niềm tin định hướng thái độ được đo lường và nhấn mạnh việc đo lường những đánh giá cảm xúc về kết quả hành vi. Cách tiếp cận này có thể giúp người nghiên cứu hiểu được tại sao người tiêu dùng lại có những hành vi hiện tại đối với đối tượng.

Ví dụ: Thu thập dữ liệu về nhận thức, cảm xúc của khách hàng liên quan đến việc đánh giá thái độ của đại lý về việc mua một chiếc ô tô. Người nghiên cứu đo lường mong đợi của người tiêu dùng về khả năng thỏa mãn những nhu cầu/mong muốn của khách hàng của Honda Civic.

Liên quan đến mong đợi của bạn về một chiếc ô tô mới, hãy cho biết những ý kiến của Anh (Chị) về các yếu tố dưới đây liên quan đến xe Honda Civic. Đối với mỗi yếu tố, vui lòng khoanh tròn số điểm diễn tả đúng nhất mức độ đồng ý hoặc không đồng ý của Anh (Chị)

về việc mua một chiếc ô tô Honda Civic sẽ đáp ứng yếu tố đó. Đối với các yếu tố mà của Anh (Chị) không có thông tin, hoặc chưa trải nghiệm, vui lòng khoanh tròn số 0 (N/A).

Mua một chiếc Honda Civic sẽ....	Hoàn toàn đồng ý	Nhìn chung là đồng ý	Hơi đồng ý	Hơi không đồng ý	Nhìn chung là không đồng ý	Hoàn toàn không đồng ý	Không áp dụng (N/A)
cho tôi một chiếc xe tiết kiệm xăng	1	2	3	4	5	6	0
cho phép tôi lái xe một cách thoải mái	1	2	3	4	5	6	0
cho tôi một chiếc xe có chất lượng mà tôi tìm kiếm	1	2	3	4	5	6	0

Cách tiếp cận tâm lí (affect global approach)

Cách tiếp cận này cho rằng thái độ không gì hơn là sự thể hiện cảm xúc thích hay không thích đối với một đối tượng cụ thể. Sử dụng cách tiếp cận này, người nghiên cứu chỉ biết được cảm xúc chung của người tiêu dùng/đáp viên đối với đối tượng (ví dụ: thích/ không thích, hài lòng/ không hài lòng,...) mà không biết niềm tin nào dựa vào đó hình thành những thái độ mà họ sở hữu.

Các thang đo thường sử dụng đo lường thái độ

Thang Likert

Thang Likert là một dạng thang đánh giá được sử dụng khá phổ biến trong nghiên cứu. Theo thang đo này, những người trả lời phải biểu thị một mức độ đồng ý hoặc không đồng ý với chuỗi các đề nghị (chuỗi các phát biểu) được trình bày.

Thang đo Likert thường dùng để đo lường các chỉ báo của biến tiềm ẩn và giá trị của biến tiềm ẩn thường được tính bằng trung bình cộng giá trị của các chỉ báo (items) trong trường hợp thang đo đơn hướng (unidimensionally).

Một thang điểm Likert thường gồm 2 phần, phần khoản mục và phần đánh giá. Phần khoản mục liên quan đến ý kiến, thái độ về các đặc tính một sản phẩm, sự kiện cần đánh giá. Phần đánh giá là một danh sách đặc tính trả lời.

Thông thường, các khoảng mục đánh giá được thiết kế từ 5 đến 9 hạng trả lời (mục lựa chọn), đi từ “hoàn toàn không đồng ý” đến “hoàn toàn đồng ý” hoặc từ “rất không quan trọng” đến “rất quan trọng”... và người trả lời sẽ chọn một hạng (mục trả lời) trong các mục được liệt kê. Trong trường hợp này, thang đo Likert sẽ là thang khoảng cách thứ tự (ordinally interval scale) như đã giới thiệu ở phần trên.

Ví dụ: Nghiên cứu về chất lượng dịch vụ của mạng điện thoại di động A:

	Hoàn toàn không đồng ý	Không đồng ý	Không ý kiến	Đồng ý	Hoàn toàn đồng ý
Khi Anh (Chị) sử dụng mạng điện thoại viễn thông A, Anh (Chị) :					
1. Kết nối ngay lần gọi đầu tiên	1	2	3	4	5
2. Thông suốt trong suốt cuộc gọi	1	2	3	4	5
3. Kết nối âm thanh rõ ràng	1	2	3	4	5
4. Vùng phủ sóng rộng	1	2	3	4	5
5. Thực hiện được vào giờ cao điểm	1	2	3	4	5

Thang đo có ngữ nghĩa đối nghịch nhau (semantic differential)

Thang đo có ngữ nghĩa đối nghịch nhau là một thang đo thể hiện thành hai cực và hai cực của nó sử dụng các từ ngữ có ý nghĩa đối nghịch nhau. Người trả lời đánh giá đối tượng trên một số các phát biểu bằng thang đo mà hai cực được xác định bằng cặp tính từ hoặc câu có ý nghĩa đối nghịch nhau như “lạnh” và “nóng”, “đắt” và “rẻ”... Tính từ hoặc thành ngữ phủ định khi thì đặt ở bên trái, khi thì ở bên phải. Điều này là nhằm mục đích tránh tình huống người trả lời chỉ đánh dấu vào một bên mà không đọc đầy đủ các phát biểu, theo Giới và cộng sự (2006), các trạng thái có ý nghĩa tốt không hoàn toàn ở phía bên trái hoặc bên phải nên tránh cho người được hỏi không bị hiệu ứng “lóa” và không chỉ chú ý đến một phía.

Ví dụ: Hãy cho biết cảm nhận của Anh (Chị) về thủ tục một cửa tại Quận A bằng cách khoanh tròn số tương ứng với sự lựa chọn của Anh (Chị):

Người tiếp nhận thân thiện	1	2	3	4	5	Người tiếp nhận cáu gắt
Phục vụ nhanh chóng	1	2	3	4	5	Phục vụ chậm
Thiếu hướng dẫn quy trình	1	2	3	4	5	Hướng dẫn quy trình đầy đủ
Thông tin rõ ràng	1	2	3	4	5	Thông tin thiếu rõ ràng

Dữ liệu đạt được từ thang điểm này được phân tích bằng cách tính các giá trị trung bình điểm số của tất cả những người được hỏi đối với mỗi yếu tố đánh giá cho mỗi đối tượng và dùng đồ thị định dạng để biểu thị nhằm so sánh nhận thức hoặc thái độ của họ đối với từng yếu tố đánh giá riêng biệt của hai hay nhiều đối tượng nghiên cứu này. Mặt khác khi người nghiên cứu muốn xác định một sự ưa thích tổng thể, điểm số được xác định bằng cách cộng tất cả điểm số đánh giá các yếu tố riêng rẽ (chú ý các cực để trái nhau thì khi nhập liệu phải mã hóa lại (recode) cho thuận chiều).

Thang điểm Stapel

Theo thang điểm này, Stapel chỉ sử dụng một tính từ cho mỗi hạng mục đưa ra để đánh giá và sự đánh giá được diễn tả bằng một dãy số từ dương (+) đến âm (-) hoặc ngược lại. Như vậy, thang điểm Stapel là một sự phát triển của thang điểm có ý nghĩa đối nghịch. Một tính từ duy nhất được sử dụng, nhiều khi nó được đặt giữa thang điểm. Thông thường nó được thiết kế như thang điểm 10 sắp xếp từ + 5 đến -5. Như vậy trong thang điểm các con số và dấu đo lường cả hướng, cường độ của thái độ và dữ liệu đạt được từ thang điểm Stapel cũng được phân tích như dữ liệu trong thang điểm có ý nghĩa đối lập nhau. Ưu điểm của thang điểm này là cho phép nhà nghiên cứu tránh được khó khăn phải tìm những cặp tính từ đối lập nhau.

Ví dụ: Hãy cho biết cảm nhận của Anh (Chị) về thủ tục một cửa tại Quận A bằng cách đánh dấu (✓) vào ô (□) tương ứng với sự lựa chọn của Anh (Chị):

	-2	-1	0	+1	+2
Người tiếp nhận thân thiện	☐	☐	☐	☐	☐
Phục vụ nhanh chóng	☐	☐	☐	☐	☐
Thiếu hướng dẫn quy trình	☐	☐	☐	☐	☐
Thông tin rõ ràng	☐	☐	☐	☐	☐

THIẾT KẾ CÔNG CỤ THU THẬP DỮ LIỆU

Phần trên đã giới thiệu các thang đo và quy trình xây dựng thang đo, trong phần này sẽ tập trung giới thiệu cách thức thiết kế những công cụ nhằm thu thập và ghi chép dữ liệu. Trong cả hai trường hợp thu thập dữ liệu bằng quan sát và phỏng vấn, việc sử dụng các biểu mẫu ghi chép là rất cần thiết, các biểu mẫu này chính là các bản câu hỏi. Theo Giới và cộng sự (2006), bản câu hỏi là một trong những công cụ để thu thập dữ liệu, nó bao hàm một tập hợp các câu hỏi và câu trả lời theo một logic nhất định. Do vậy, bản câu hỏi là một tiến trình được chính thức hóa nhằm thu thập, ghi chép lại những thông tin xác đáng và được chỉ định rõ với sự chính xác và hoàn hảo tương đối. Mục tiêu của bản câu hỏi là:

- Chuyển đổi những thông tin cần thu thập vào các câu hỏi cụ thể để dễ dàng trả lời;
- Thúc đẩy, động viên đáp viên cung cấp thông tin một cách chính xác, lời cuốn đáp viên nhiệt tâm với cuộc phỏng vấn;
- Giảm thiểu những sai sót xảy ra do đáp viên trả lời không chính xác hoặc khi câu trả lời của họ không thể ghi nhận được.

Để đảm bảo dữ liệu thu thập được đáp ứng mục tiêu nghiên cứu và vấn đề nghiên cứu, việc xây dựng bản câu hỏi cần thông qua các bước trên cơ sở 4 giai đoạn sau:

- Giai đoạn 1: Xác định dữ liệu cần thu thập và phương pháp thu thập dữ liệu gồm các bước (1) xác định dữ liệu cần thiết phải thu thập và (2) xác định phương pháp thu thập dữ liệu;
- Giai đoạn 2: Phát triển câu hỏi gồm (1) xác định nội dung câu hỏi; (2) quyết định dạng câu hỏi và câu trả lời; (3) xác định các từ ngữ trong bản câu hỏi;

- Giai đoạn 3: Phát triển và điều chỉnh bản câu hỏi gồm (1) xác định cấu trúc bản câu hỏi và sắp xếp thứ tự các câu hỏi; (2) xác định đặc tính vật lý và hoàn chỉnh bản câu hỏi;
- Giai đoạn 4: Điều tra thử, kiểm tra và sửa chữa.

Giai đoạn 1: Xác định dữ liệu cần thu thập và phương pháp thu thập dữ liệu

Xác định dữ liệu cần thu thập

Để bảo đảm thông tin thu thập có thể bao phủ được các vấn đề được đưa ra, nhà nghiên cứu cần phải tập hợp một biểu thông tin (theo kinh nghiệm, nhà nghiên cứu nên liệt kê trên Excel để dễ dàng kiểm soát và phát triển bản câu hỏi sau này). Trong đó liệt kê những khái niệm, những chỉ báo... thật sự cần thiết để đo lường. Qua biểu thông tin này, cần phải cân nhắc những câu hỏi, phát biểu, từ ngữ... phù hợp với mục tiêu, đối tượng nghiên cứu. Cần cân nhắc số lượng vấn đề, thông tin cần phải thu thập. Nếu những nội dung quá xa với mục tiêu nghiên cứu thì nên không đưa vào bản câu hỏi để tiết kiệm thời gian và chi phí trong quá trình thu thập thông tin. Trong giai đoạn này cũng cần dự tính xem những cấp độ thang đo nào dùng để đo lường các biến số, từ đó xem xét nên dùng loại kỹ thuật phân tích nào để mang lại ý nghĩa cho dữ liệu để làm cơ sở cho việc thiết kế câu hỏi ở phần sau.

Xác định phương pháp thu thập dữ liệu

Những phương pháp thu thập dữ liệu sơ cấp bao gồm phỏng vấn cá nhân trực tiếp (face – to – face, one – to – one), cho đáp viên tự trả lời, phỏng vấn qua thư, điện thoại, mạng Internet (qua website hoặc qua e-mail). Mỗi phương pháp phỏng có ưu và nhược điểm khác nhau, tùy theo tính chất của mỗi nghiên cứu mà chọn lựa phương pháp cho phù hợp. Phương pháp thu thập dữ liệu ảnh hưởng đến tiến trình thiết kế bản câu hỏi. Theo Giới và cộng sự (2006), cần lưu ý một số vấn đề sau:

- Phỏng vấn trực tiếp: là hình thức phỏng vấn hoàn thiện nhất trong việc trao đổi trực tiếp giữa người trả lời và phỏng vấn viên. Khi gặp trực tiếp đối tượng được phỏng vấn, phỏng vấn viên có thể giải thích một cách sinh động các nội dung câu hỏi bằng lời hoặc hình ảnh minh họa (showcard), có thể thiết đặt mối quan hệ nhằm kích thích mức độ quan tâm của người trả lời để sẵn lòng trả lời bản câu hỏi. Trong trường hợp này, những vấn đề mà tính chất kỹ thuật, những chủ đề khó, những câu hỏi dài và phức tạp cũng có thể được sử dụng. Tuy nhiên, đây là phương pháp tốn nhiều thời gian, tiền bạc vì chi phí cho phỏng vấn viên.
- Điều tra bằng thư tín: là hình thức gửi bản câu hỏi qua bưu điện cho người trả lời tự trả lời theo hình thức truyền thống hoặc không thông qua bưu điện đối với thư khảo sát (gửi trực tiếp hoặc người trả lời tự lấy ở những địa điểm nhất định). Do vậy, những câu hỏi trong bản câu hỏi phải đơn giản và những chỉ dẫn cho người trả lời phải hết sức chi tiết, rõ ràng.
- Phỏng vấn qua điện thoại: là hình thức phỏng vấn giao tiếp bằng lời, phỏng vấn viên dùng điện thoại gọi cho đáp viên. Trong trường hợp này, đáp viên trả lời các câu hỏi của người phỏng vấn thông qua điện thoại mà không thấy được người hỏi và bản câu hỏi. Hình thức này cho phép phỏng vấn viên giải thích các câu hỏi phức tạp cho người trả lời nhằm đáp

ứng nội dung phỏng vấn. Tuy nhiên, trong những trường hợp những vấn đề nghiên cứu phức tạp, mang tính kỹ thuật cao, người phỏng vấn không thể sử dụng các catalogue, showcard minh họa cho quá trình phỏng vấn.

- Điều tra qua mạng Internet: Một hình thức mà hiện nay trên thế giới thường dùng là hình thức phỏng vấn bằng thư điện tử (email) hoặc phỏng vấn trực tiếp qua trang website có thiết kế bản câu hỏi nhằm phỏng vấn mạng Internet. Cũng như phương pháp phỏng vấn qua thư, người nghiên cứu có thể hỏi những câu hỏi có cấu trúc phức tạp, cũng có thể gửi kèm với catalogue và showcard về sản phẩm. Tuy nhiên nhược điểm của phương pháp này là tỉ lệ trả lời không cao.

Giai đoạn 2: Phát triển câu hỏi

Ở giai đoạn này, người nghiên cứu thiết kế những câu hỏi cụ thể trong đó tập trung vào cách đặt câu hỏi và chiến lược trả lời (liên quan đến các cấp độ thang đo và kế hoạch phân tích dữ liệu). Thứ tự, hình thức và cách đặt câu hỏi, phần giới thiệu, hướng dẫn trả lời, phần chuyển tiếp và phần kết thúc của một bản câu hỏi được thiết kế nhằm thu được những dữ liệu theo yêu cầu đặt ra. Các câu hỏi cần:

- Khuyến khích đáp viên cung cấp những câu trả lời chính xác/ trung thực;
- Khuyến khích đáp viên cung cấp lượng thông tin phù hợp;
- Không để cho đáp viên từ chối trả lời các câu hỏi cụ thể nào đó;
- Không để cho đáp viên không hoàn thành bản câu hỏi;
- Để lại cho đáp viên ấn tượng tốt đẹp khi tham gia vào điều tra.

Xác định nội dung câu hỏi

Các câu hỏi trong bản câu hỏi được thiết kế dựa trên các câu hỏi nghiên cứu đã được xác định. Theo Cooper and Schindler (2011), cần có sự kết nối giữa câu hỏi nghiên cứu và câu hỏi trong bản câu hỏi (câu hỏi đo lường), các câu hỏi trong bản câu hỏi phải chuyển tải nội dung muốn hỏi (hay thông tin yêu cầu đạt được) vào trong các câu hỏi; sau quá trình phỏng vấn, đòi hỏi người nghiên cứu phải có thông tin theo câu hỏi, mục tiêu nghiên cứu đưa ra. Do vậy, trong quá trình thiết kế nghiên cứu, câu hỏi và mục tiêu nghiên cứu càng rõ ràng thì việc xác định các nội dung cần phải thu thập dữ liệu càng thuận tiện.

Nhà nghiên cứu cần xem xét cấp độ thang đo được lựa chọn (biểu danh, thứ tự, khoảng, tỉ lệ) để có thể thể hiện thông tin muốn hỏi và thực hiện các phương pháp phân tích dữ liệu yêu cầu. Như đã đề cập ở phần trên, việc lựa chọn cấp độ thang đo để xây dựng các câu hỏi sẽ ảnh hưởng rất lớn đến việc sử dụng các kỹ thuật phân tích. Ví dụ, vấn đề nghiên cứu cần những dữ liệu định lượng thì các câu hỏi phải sử dụng cấp độ thang đo khoảng cách hoặc tỷ lệ. Những thang đo cấp độ định lượng có thể chuyển sang cấp độ thang đo định tính và không có trường hợp ngược lại. Vì vậy, đây là bước cần thực hiện thận trọng, nếu không sẽ thu thập dữ liệu không đúng với yêu cầu đặt ra.

Trong quá trình xây dựng, nhà nghiên cứu cần tự trả lời các câu hỏi sau đây:

- Câu hỏi này có nên đưa vào hay không (nó có đáp ứng được mục tiêu của nghiên cứu không)?
- Độ bao phủ của câu hỏi cho vấn đề nghiên cứu đã thích hợp chưa?
- Có nên đặt nhiều câu hỏi nhỏ để thay cho một câu hỏi lớn nào đó không?
- Đáp viên có thể hiểu đúng câu hỏi để trả lời không?
- Đáp viên có sẵn sàng trả lời câu hỏi không?

Quyết định dạng câu hỏi và câu trả lời

Có hai dạng câu hỏi chính sau:

- Câu hỏi mở là dạng câu hỏi mà trong đó câu hỏi được cấu trúc còn câu trả lời thì không. Người trả lời có thể trả lời với bất cứ thông tin nào và bất cứ câu nào được coi là thích hợp. Người phỏng vấn sẽ có nhiệm vụ viết lại chính xác những gì có thể thu thập được. Những thuận lợi của câu hỏi tự do trả lời là cho phép người nghiên cứu thu được những câu trả lời bất ngờ, không dự liệu trước; người trả lời có thể bộc lộ rõ ràng hơn quan điểm của mình về một vấn đề nào đó, mà không bị gò bó bởi nội dung câu hỏi; giảm bớt sự thất vọng của người trả lời so với câu hỏi đóng là không có cơ hội phát biểu ý kiến, chỉ lựa chọn trong tình huống có sẵn.

Những khó khăn khi sử dụng câu hỏi tự do trả lời là có thể gặp khó khăn để hiểu người trả lời khi họ diễn đạt kém; khó mã hóa và phân tích; phụ thuộc vào sự ghi chép của người phỏng vấn nên có thể bỏ qua những chi tiết quan trọng nhưng không ghi chép lại vì cho rằng không cần thiết. Dạng câu hỏi này ít được dùng trong trường hợp phỏng vấn bằng thư tín vì tâm lý người được hỏi thường ngại viết hơn là nói, đôi khi mất thời gian vì người trả lời nói lan man.

- Câu hỏi đóng là dạng câu hỏi mà cả câu hỏi và câu trả lời được cấu trúc. Gồm các dạng câu hỏi sau:

- ***Câu hỏi lưỡng phân (dichotomous)***

Câu hỏi lưỡng phân là câu hỏi có hai lựa chọn, trong hầu hết các trường hợp, hai lựa chọn của câu hỏi thường đối nghịch nhau. Người trả lời chỉ có thể chọn một trong hai câu trả lời như “có hoặc không”, “đúng hoặc sai”..., loại câu hỏi này thường dùng trong trường hợp để phân loại đối tượng.

Ưu điểm của loại câu hỏi này là thiết kế nhanh chóng và dễ dàng, rất tiện lợi trong những câu hỏi có nhiều chi tiết; dễ dàng cho người trả lời; thuận tiện trong xử lý, tính toán và phân tích.

Nhược điểm của nó là cung cấp không đủ thông tin chi tiết và bắt buộc người trả lời lựa chọn cho dù họ có thể chưa chắc chắn lắm khi chọn câu trả lời.

- *Câu hỏi nhiều lựa chọn (multiple responses)*

Câu hỏi mà người thiết kế phải liệt kê tất cả các phương án có thể có, người được hỏi sẽ đánh dấu một hay nhiều loại trả lời được liệt kê.

Ví dụ: Hằng tháng, Anh (Chị) đọc loại báo nào dưới đây:

- ☐ Hà Nội mới
- ☐ Lao động
- ☐ Tiền phong
- ☐ Thanh niên
- ☐ Tuổi trẻ

Như vậy, có hai vấn đề hay gặp trong khi sử dụng câu hỏi nhiều lựa chọn: (1) liệu các phương án mà người thiết kế liệt kê đã bao phủ hết các phương án có thể trả lời, khi một hoặc nhiều lựa chọn nên đưa vào tập hợp câu trả lời (lựa chọn) nhưng không được đưa vào và (2) có thể đáp viên muốn đưa ra câu trả lời khác với những lựa chọn đã được thiết kế. Trong trường hợp này, câu hỏi có thể thêm lựa chọn là “khác” để đáp viên điền câu trả lời vào. Khi đó, câu hỏi có thể thiết kế như sau:

Ví dụ: Hằng tháng, Anh (Chị) đọc loại báo nào dưới đây:

- ☐ Hà Nội mới
- ☐ Lao động
- ☐ Tiền phong
- ☐ Thanh niên
- ☐ Tuổi trẻ
- ☐ Khác. Cụ thể là:

- *Câu hỏi đánh dấu tình hướng theo danh sách*

Về cấu trúc, nó tương tự như câu hỏi nhiều lựa chọn, các phương án trả lời được liệt kê, tuy nhiên khác biệt là người được hỏi sẽ đánh dấu một phương án (tình hướng) mà họ cho là đúng nhất.

Ví dụ: Hằng tháng, Anh (Chị) thường đọc loại báo nào nhất trong các loại báo dưới đây:

- ☐ Hà Nội mới
- ☐ Lao động
- ☐ Tiền phong
- ☐ Thanh niên
- ☐ Tuổi trẻ

- *Câu hỏi xếp hạng thứ tự (Ranking question)*

Với nhu cầu xem xét thứ tự của các lựa chọn là quan trọng, các câu hỏi xếp hạng nên được áp dụng.

Ví dụ: Xin Anh (Chị) sắp xếp thứ tự ưu tiên về các thuộc tính khi mở thẻ ATM tại một ngân

hàng với 1 là yếu tố quan trọng nhất và (7) là yếu tố ít quan trọng nhất:

Uy tín ngân hàng	
Mạng lưới chi nhánh	
Mạng lưới chi nhánh dịch vụ thẻ	
Chiến lược marketing	
Công nghệ sử dụng	
Chất lượng dịch vụ thẻ	
Khả năng kết nối	

Xác định các từ ngữ trong bản câu hỏi

Một trong những lí do mà đáp viên không hiểu hoặc hiểu nhầm câu hỏi là do cách dùng từ trong câu hỏi. Đáp viên có thể sẽ không hiểu được các từ ngữ chuyên môn hoặc các câu hỏi dài và phức tạp. Do vậy, để đảm bảo đáp viên hiểu được câu hỏi, những vấn đề sau đây cần được lưu ý:

- Những từ thông dụng nên được sử dụng trong một bản câu hỏi và chúng nên phù hợp với trình độ học vấn trung bình của đáp viên (tổng thể nghiên cứu). Vì vậy, khi sử dụng từ ngữ, nhà nghiên cứu nên biết trình độ học vấn trung bình của các đối tượng nghiên cứu, nếu như trình độ học vấn thấp thì nhà nghiên cứu không nên sử dụng những từ ngữ mang tính kỹ thuật và một số từ ngữ chuyên môn phức tạp khác.
- Câu hỏi có gồm những từ chỉ có một nghĩa: Những từ ngữ được sử dụng cho bản câu hỏi có thể có một ý nghĩa duy nhất đối với đáp viên. Một số từ ngữ khi sử dụng có thể làm nảy sinh nhiều nghĩa khác nhau như "thường thường"; "thỉnh thoảng"; "đôi khi "; "luôn luôn",... không nên sử dụng trong quá trình thiết kế.
- Tránh câu hỏi mang tính mang tính hướng dẫn hoặc sai lệch, mang ý nghĩ thiên vị (biased wording).

Ví dụ: Bạn nghĩ rằng công ty A có nên thực hiện các chương trình quảng cáo khi công ty đang có những cáo buộc liên quan đến xả nước thải làm ảnh hưởng đến môi trường không?

- ☐ Không
- ☐ Có
- ☐ Không biết

Câu hỏi này thường làm cho đáp viên có câu trả lời là không vì có những thông tin định hướng thiên lệch (xả nước thải làm ảnh hưởng đến môi trường không).

- Tránh những câu hỏi mang tính ngấm ngầm:

Ví dụ: Bạn có muốn sử dụng đồ mới không?

Trong câu hỏi này từ đồ mới không thể hiểu được hoặc nếu hiểu được thì có nghĩa là tất cả mọi người đều mong muốn sử dụng sản phẩm, dịch vụ mới.

- Tránh những câu hỏi mang tính khái quát hóa hoặc ước đoán: Một câu hỏi nên được định ra

một cách cụ thể, không nên chung chung, câu hỏi không nên diễn đạt quá khái quát hay phải suy tính mới có thể trả lời chính xác.

Giai đoạn 3: Phát triển và điều chỉnh bản câu hỏi

Xác định cấu trúc bản câu hỏi và sắp xếp thứ tự câu hỏi

Theo Giơri và cộng sự (2006), người ta có thể chia các câu hỏi thành 5 loại và tạo thành 5 phần chính trong bản câu hỏi theo chức năng của chúng đóng góp vào sự thành công của cuộc phỏng vấn gồm:

- Phần mở đầu hoặc câu hỏi hướng dẫn: Có tác dụng mở đầu cuộc phỏng vấn một cách thuận tiện, khởi đầu cho chuỗi những câu trả lời và gây thiện cảm với người được phỏng vấn.
- Câu hỏi gạn lọc: Có tác dụng chỉ rõ đối tượng cần được phỏng vấn, tránh phải phỏng vấn những người không có kiến thức về vấn đề đang điều tra.
- Câu hỏi hâm nóng: Có tác dụng gợi nhớ thông tin và tập trung vào chủ đề nghiên cứu, tránh các cảm xúc đột ngột khi đi vào chủ đề quá nhanh, người trả lời có thể chưa tạo được hứng thú và hồi tưởng thông tin kịp.
- Các câu hỏi đặc thù: Là những câu hỏi đi vào chủ đề nghiên cứu: mức độ thường xuyên khi sử dụng sản phẩm? địa điểm mua? cảm giác về chất lượng? động cơ sử dụng?...
- Các câu hỏi phụ: Được sử dụng để thu nhận thêm các thông tin về đặc điểm nhân khẩu của người trả lời (giới tính, tuổi tác, nghề nghiệp...).

Trong khi thiết kế, cần sắp xếp thứ tự các câu hỏi sao cho đáp viên có thể dễ dàng trả lời nhất, những câu hỏi khó hoặc những câu hỏi mang tính cá nhân nên được đặt sau cùng. Những câu hỏi được hỏi trước có thể làm ảnh hưởng đến nội dung trả lời cho những câu hỏi theo sau. Do vậy, những câu hỏi chung nên được đặt trước những câu hỏi cụ thể, có như vậy mới tránh được những sai lệch từ các câu hỏi cụ thể đến các câu hỏi chung. Tất cả những câu hỏi có liên quan đến một chủ đề riêng biệt nào đó nên được hỏi trước khi bắt đầu một chủ đề mới. Khi chuyển đổi chủ đề, những câu hỏi chuyển tiếp nên được đặt ra để giúp cho đáp viên chuyển đổi suy nghĩ của họ dễ dàng.

Các câu hỏi nên được sắp xếp theo tâm lí của đáp viên vì thứ tự của câu hỏi có thể khuyến khích hoặc làm đáp viên không muốn trả lời. Các hướng dẫn sau đây cần được lưu ý:

- Các câu hỏi phải tạo động lực cho đáp viên, đặt những câu hỏi thú vị ngay ở phần đầu, các câu hỏi về đặc điểm nhân khẩu (mà không phải là câu hỏi lọc) đưa vào phần cuối.
- Đưa các câu hỏi có thể khiến đáp viên dừng trả lời ở phần cuối của bản câu hỏi.
- Thứ tự của các câu hỏi nên từ đơn giản đến phức tạp, các câu hỏi đòi hỏi đáp viên phải suy nghĩ nên đưa vào phía cuối.
- Sử dụng các câu chuyển giữa các chủ đề/nhóm câu hỏi.

Các vấn đề khi thiết kế công cụ thu thập dữ liệu thường nằm ở từ ngữ dùng cho câu hỏi, nội dung câu hỏi và trình tự của câu hỏi. Người nghiên cứu có thể cải thiện chất lượng của cuộc điều tra bằng cách:

- Xây dựng mối quan hệ với đáp viên: Xây dựng mối quan hệ với đáp viên có thể đem đến cho đáp viên sự thích thú khi trả lời phỏng vấn hoặc trả lời bản câu hỏi. Nếu đáp viên cảm thấy phỏng vấn viên đáng tin cậy, thông tin của họ được bảo mật hoặc cuộc điều tra này đang đóng góp vào một vấn đề quan trọng thì động lực để trả lời của đáp viên sẽ cao hơn. Nếu giữa phỏng vấn viên và đáp viên có mối quan hệ tốt, đáp viên còn có thể chia sẻ những suy nghĩ, ý kiến của mình về những chủ đề có thể mang tính nhạy cảm và riêng tư.
- Thiết kế tiến trình hỏi, trình tự câu hỏi: Tiến trình hỏi hoặc trình tự câu hỏi ảnh hưởng đến sự trung thực trong khi trả lời câu hỏi của đáp viên, đặc biệt trong những chủ đề nhạy cảm và mang tính riêng tư. Vì thế, người nghiên cứu cần thiết kế trình tự các câu hỏi sao cho dần dần đưa đáp viên từ các câu hỏi “an toàn” đến các câu hỏi nhạy cảm hơn.
- Khám phá những chiến lược trả lời khác: Khi thiết kế các câu hỏi, người nghiên cứu nên cố gắng phát triển các phiên bản khác nhau của một câu hỏi (bao gồm những phát biểu cùng chiều tích cực hay cùng chiều tiêu cực, có lựa chọn trung lập, không có lựa chọn trung lập). Những kiểu lựa chọn trả lời này sẽ ảnh hưởng đến việc hoàn thành bản câu hỏi của đáp viên, đặc biệt trong các chủ đề nhạy cảm. Ví dụ, khi hỏi về thu nhập của đáp viên, đưa ra các lựa chọn thu nhập nằm trong các khoảng khác nhau thì đáp viên sẽ dễ hoàn thành hơn việc yêu cầu họ điền chính xác con số thu nhập của mình.

Xác định các đặc tính vật lý và hoàn chỉnh bản câu hỏi

- Xem xét hình dạng bản câu hỏi, loại font chữ, cỡ chữ... và trình bày bản câu hỏi sao cho ngắn gọn và rõ ràng. Bản câu hỏi càng ngắn gọn (nhưng phải đảm bảo những nội dung cần thiết phải thu thập) sẽ tạo cảm giác dễ dàng chấp nhận tham gia hơn đối với người trả lời.
- Chất lượng giấy, chất lượng in ấn... phải đảm bảo để tạo thiện cảm và lôi cuốn người trả lời tham gia vào cuộc phỏng vấn (có một số trường hợp, nếu chúng ta in bản câu hỏi trên giấy màu thì cũng có thể gia tăng tỉ lệ trả lời).
- Nếu dùng câu hỏi mở thì nên chừa khoảng trống đủ để người được hỏi ghi câu trả lời và diễn đạt ý kiến của mình.
- Khi nhảy quãng câu hỏi trên bản câu hỏi thì phải chú thích rõ ràng để giúp cho người trả lời có thể biết bước tiếp theo họ cần làm gì.

Ví dụ: Anh (Chị) đã từng sử dụng dịch vụ Internet Banking chưa?

- | | |
|--|----------------------------------|
| ☐ Đã từng sử dụng | ⇒ Xin chuyển sang trả lời câu 12 |
| ☐ Chưa từng sử dụng | ⇒ Xin trả lời tiếp câu 6 |

Giai đoạn 4: Điều tra thử, kiểm tra và sửa chữa

Trước khi dùng bản câu hỏi để điều tra, bản câu hỏi phải được kiểm tra và đánh giá qua việc

test thử. Những lí do cho việc test thử từng câu hỏi, bản câu hỏi và thời gian biểu cho việc phỏng vấn là:

- Khám phá ra các cách khác nhau để nâng cao sự thích thú của đáp viên khi tham gia điều tra;
- Tăng xác suất đáp viên duy trì được sự quan tâm và tập trung cho đến lúc hoàn thành điều tra/phỏng vấn/trả lời;
- Xác định được những vấn đề liên quan đến nội dung câu hỏi, từ ngữ sử dụng và trình tự câu hỏi;
- Xác định được những nhóm câu hỏi mục tiêu mà phỏng vấn viên phải được đào tạo để hỏi;
- Tìm ra cách cải thiện chất lượng dữ liệu.

Test thử bản câu hỏi có thể không giải quyết được hết những lỗi, sai sót tuy nhiên nó rất cần thiết để giảm thiểu những lỗi sai trong quá trình thu thập dữ liệu. Khi thiết kế, việc chỉnh sửa có thể phải lặp lại nhiều lần (5 lần hoặc hơn) để đảm bảo chất lượng trong việc thu thập thông tin nghiên cứu.

Người trả lời được chọn trong điều tra thử phải nằm trong tổng thể nghiên cứu của cuộc điều tra chính thức để bảo đảm tính không sai biệt của hai cuộc điều tra của cùng một nội dung nghiên cứu. Chỉnh sửa câu hỏi, loại bỏ hay bổ sung thêm câu hỏi được thực hiện trong suốt quá trình điều tra thử. Sau khi điều tra thử, dữ liệu sẽ được mã hóa và lập danh bạ mã để tiết kiệm tổng thời gian của dự án nghiên cứu. Đồng thời, xử lí và phân tích dữ liệu trong điều tra thử giúp cho người nghiên cứu sử dụng tốt tất cả các dữ liệu được thu thập từ bản câu hỏi.

TÓM TẮT CHƯƠNG 5

Đo lường được định nghĩa là quá trình gán những con số đối với những đặc tính của các sự vật nghiên cứu theo các nguyên tắc đã được xác định để có thể đánh giá, so sánh và phân tích. Có bốn loại cấp độ thang đo là cấp độ thang đo biểu danh, thứ tự, khoảng và tỷ lệ.

Cấp độ thang đo biểu danh là thang đo có cấp độ đo lường đơn giản nhất vì người ta chỉ sử dụng các số để xác định, nêu danh các đối tượng. Cấp độ thang đo thứ tự thường dùng các con số biểu thị vị trí tương đối giữa các đối tượng nhưng không cho biết khoảng khác biệt giữa chúng. Thang đo khoảng cách cho phép so sánh những khác biệt giữa các đối tượng nhưng không có một điểm gốc "0". Cấp độ thang đo cao nhất của đo lường được thể hiện bằng thang đo tỷ lệ, trong đó điểm zero là cố định.

Để đo lường thái độ của con người, người ta thường ứng dụng thang đo khoảng để triển khai các thang đo đặc thù gồm thang Likert, thang có ngữ nghĩa đối lập nhau và thang điểm Stapel.

Có 3 cách để đánh giá độ tin cậy của một đo lường là đo lường lặp lại (test - retest), sử dụng dụng cụ đo lường tương đương và đo lường tính nhất quán nội tại. Tuy nhiên, một công cụ đo lường tốt là phải có giá trị. Khi một thang đo có giá trị hoàn hảo thì không có khác biệt giữa điểm số đo lường và điểm số thực tế hay không có bất kỳ một sai số nào của đo lường. Giữa độ

tin cậy và giá trị của thang đo lường có mối quan hệ chặt chẽ với nhau. Một thang đo lường trước hết phải đảm bảo độ tin cậy thì mới có giá trị, tuy nhiên, đảm bảo độ tin cậy chưa hẳn đã có giá trị nếu tồn tại sai số hệ thống.

Các cấp độ thang đo dùng để xây dựng các công cụ thu thập dữ liệu như bản câu hỏi hoặc biểu mẫu quan sát. Bản câu hỏi phải đảm bảo chuyển những thông tin cần thiết thu thập thành những câu hỏi cụ thể mà người được hỏi có thể trả lời với một sự hợp tác cao và giảm được sai số do trả lời; người đi phỏng vấn dễ thực hiện việc phỏng vấn và tiên liệu trước được những yêu cầu trong việc xử lý dữ liệu.

Không có một tiêu chuẩn cố định nào để xây dựng một bản câu hỏi nhưng có những bước tiếp cận giúp cho người lập bản câu hỏi có thể thiết kế một bản câu hỏi đạt được hiệu quả thông tin cao. Các giai đoạn xây dựng bản câu hỏi gồm:

- Giai đoạn 1: Xác định dữ liệu cần thu thập và phương pháp thu thập dữ liệu gồm các bước (1) xác định dữ liệu cần thiết phải thu thập và (2) xác định phương pháp thu thập dữ liệu;
- Giai đoạn 2: Phát triển câu hỏi gồm (1) xác định nội dung câu hỏi; (2) quyết định dạng câu hỏi và câu trả lời; (3) xác định các từ ngữ trong bản câu hỏi;
- Giai đoạn 3: Phát triển và điều chỉnh bản câu hỏi gồm (1) xác định cấu trúc bản câu hỏi và sắp xếp thứ tự các câu hỏi; (2) xác định đặc tính vật lý và hoàn chỉnh bản câu hỏi;
- Giai đoạn 4: Điều tra thử, kiểm tra và sửa chữa.

BẢN CÂU HỎI

LÒNG TRUNG THÀNH CỦA NHÂN VIÊN ĐỐI VỚI TỔ CHỨC

Xin chào Quý Anh, Chị!

Tôi tên là, hiện là Trong khuôn khổ chương trình, tôi thực hiện đề tài nghiên cứu về *Lòng trung thành của nhân viên đối với tổ chức*. Để hoàn thành đề tài, tôi rất mong nhận được sự quan tâm giúp đỡ của Quý Anh (Chị) trong việc tham gia trả lời bản câu hỏi này.

Trước khi bắt đầu trả lời, mong Quý Anh (Chị) đọc những chú ý dưới đây:

Trả lời tất cả các câu hỏi (theo những chỉ dẫn trong bản câu hỏi).

Tất cả những thông tin mà Quý Anh (Chị) cung cấp trong bản câu hỏi, chúng tôi chỉ sử dụng cho mục đích nghiên cứu của đề tài, ngoài ra, chúng tôi hoàn toàn không sử dụng cho mục đích khác.

Bản câu hỏi bao gồm trang

Xin chân thành cảm ơn sự giúp đỡ của Quý Anh (Chị),

Nguyễn Văn A

Người phỏng vấn :

Ngày, giờ phỏng vấn:

PHẦN A Ý KIẾN VỀ CÁC NHÂN TỐ ẢNH HƯỞNG ĐẾN LÒNG TRUNG THÀNH

1. Xin vui lòng cho biết ý kiến của các Anh (Chị) với (1) là rất không đồng ý và (7) là rất đồng ý các nhận định được cho dưới đây:

	1	2	3	4	5	6	7
1. Công việc cho phép Anh (Chị) sử dụng tốt các năng lực cá nhân	☐	☐	☐	☐	☐	☐	☐
2. Công việc rất thú vị	☐	☐	☐	☐	☐	☐	☐
3. Công việc có nhiều thách thức	☐	☐	☐	☐	☐	☐	☐
4. Khi công việc hoàn thành tốt sẽ được công ty rất hoan nghênh	☐	☐	☐	☐	☐	☐	☐
5. Anh (Chị) được trả lương cao	☐	☐	☐	☐	☐	☐	☐
6. Anh (Chị) thường được tăng lương	☐	☐	☐	☐	☐	☐	☐
7. Anh (Chị) có thể sống hoàn toàn dựa vào thu nhập từ công ty	☐	☐	☐	☐	☐	☐	☐
8. Tiền lương tương xứng với kết quả làm việc	☐	☐	☐	☐	☐	☐	☐
9. Tiền lương, thu nhập được trả công bằng	☐	☐	☐	☐	☐	☐	☐
10. Công ty có chế độ phúc lợi tốt	☐	☐	☐	☐	☐	☐	☐
11. Công ty thực hiện chế độ bảo hiểm xã hội tốt	☐	☐	☐	☐	☐	☐	☐
12. Công ty thực hiện chế độ bảo hiểm y tế tốt	☐	☐	☐	☐	☐	☐	☐
13. Anh (Chị) hài lòng với chế độ tiền thưởng của công ty	☐	☐	☐	☐	☐	☐	☐
14. Công việc không bị áp lực cao	☐	☐	☐	☐	☐	☐	☐
15. Công việc không đòi hỏi thường xuyên phải làm ngoài giờ	☐	☐	☐	☐	☐	☐	☐
16. Trang thiết bị nơi làm việc an toàn, sạch sẽ	☐	☐	☐	☐	☐	☐	☐

17. Công việc ổn định, không phải lo lắng về mất việc làm	☐	☐	☐	☐	☐	☐	☐
18. Đồng nghiệp của Anh (Chị) thoải mái và dễ chịu	☐	☐	☐	☐	☐	☐	☐
19. Anh (Chị) và các đồng nghiệp phối hợp làm việc tốt	☐	☐	☐	☐	☐	☐	☐
20. Những người mà Anh (Chị) làm việc với thường giúp đỡ lẫn nhau	☐	☐	☐	☐	☐	☐	☐
21. Anh (Chị) cảm thấy có nhiều động lực trau dồi chuyên môn khi được làm việc với các đồng nghiệp của mình	☐	☐	☐	☐	☐	☐	☐
22. Anh (Chị) được biết những điều kiện để được thăng tiến	☐	☐	☐	☐	☐	☐	☐
23. Công ty tạo cho Anh (Chị) nhiều cơ hội thăng tiến	☐	☐	☐	☐	☐	☐	☐
24. Anh (Chị) được cung cấp kiến thức/ kỹ năng cần thiết cho công việc	☐	☐	☐	☐	☐	☐	☐
25. Công ty tạo cho Anh (Chị) nhiều cơ hội phát triển cá nhân	☐	☐	☐	☐	☐	☐	☐
26. Cấp trên hỏi ý kiến Anh (Chị) khi có vấn đề liên quan đến công việc của anh/chị	☐	☐	☐	☐	☐	☐	☐
27. Cấp trên khuyến khích cấp dưới tham gia vào những quyết định quan trọng	☐	☐	☐	☐	☐	☐	☐
28. Nhân viên được sự hỗ trợ của cấp trên	☐	☐	☐	☐	☐	☐	☐
29. Công ty hoạt động có hiệu quả	☐	☐	☐	☐	☐	☐	☐
30. Nhân viên được tôn trọng và tin cậy trong công việc	☐	☐	☐	☐	☐	☐	☐
31. Lãnh đạo có tác phong lịch sự, hòa nhã	☐	☐	☐	☐	☐	☐	☐
32. Nhân viên được đối xử công bằng, không phân biệt	☐	☐	☐	☐	☐	☐	☐
33. Anh (Chị) có ý định ở lại lâu dài cùng công ty	☐	☐	☐	☐	☐	☐	☐
34. Nếu có nơi khác có lời đề nghị lương bổng tương đối hấp dẫn hơn, Anh (Chị) vẫn sẽ ở lại cùng công ty	☐	☐	☐	☐	☐	☐	☐
35. Về nhiều phương diện, Anh (Chị) coi công ty là mái nhà thứ hai của mình	☐	☐	☐	☐	☐	☐	☐

2. Quý Anh (Chị) đã được nhà quản trị trình bày phương thức kinh doanh mới, xin cho biết ý kiến của Quý Anh (Chị) về mức độ tự chủ trước và sau khi thử nghiệm phương thức lãnh đạo mới. Xin vui lòng cho biết ý kiến của các Anh (Chị) với (1) là hoàn toàn không đồng ý và (7) là hoàn toàn đồng ý các nhận định được cho dưới đây:

	1	2	3	4	5	6	7
1. Phương thức quản lý hiện tại trao quyền tự chủ cho nhân viên	☐	☐	☐	☐	☐	☐	☐
2. Phương thức quản lý mới trao quyền tự chủ cho nhân viên	☐	☐	☐	☐	☐	☐	☐
3. Anh (Chị) coi công ty là mái nhà thứ hai của mình	☐	☐	☐	☐	☐	☐	☐

3. Trong phương thức quản lý mới:

a. Anh (Chị) mong muốn áp dụng chiến lược:

- ☐ Định hướng vào công việc
- ☐ Định hướng vào nhân viên
- ☐ Định hướng vào quy trình

b. Anh (Chị) mong muốn cho phép mức độ linh hoạt:

- ☐ Cao
- ☐ Trung bình
- ☐ Thấp

PHẦN THÔNG TIN VỀ ĐÁP VIÊN

Xin vui lòng cho biết ý kiến của Anh (Chị) với mỗi câu hỏi dưới đây:

1. Anh (Chị) đang công tác tại bộ phận:
☐ Kinh doanh ☐ Kỹ thuật ☐ Văn phòng
☐ Kế toán ☐ Khác, cụ thể là:
2. Vị trí (chức vụ) công tác của Anh (chị) trong doanh nghiệp là:
☐ Nhân viên các bộ phận
☐ Cán bộ phụ trách ở các bộ phận
☐ Trưởng phòng (ban) hoặc tương đương
3. Giới tính:
☐ Nam ☐ Nữ
4. Trình độ chuyên môn, nghiệp vụ:
☐ Trung cấp ☐ Cao đẳng ☐ Đại học ☐ Sau đại học
5. Anh (Chị) đã công tác tại công ty năm.
6. Độ tuổi của Anh (Chị) là tuổi.
7. Anh (Chị) đang công tác tại:
☐ Miền Bắc ☐ Miền Trung ☐ Miền Nam
8. Anh (Chị) đang được quy hoạch ở đội ngũ kế cận:
☐ Loại 1 ☐ Loại 2 ☐ Loại 3 ☐ Loại 4
9. Anh (Chị) đã nhận được danh hiệu thi đua của Công ty từ khi công tác đến nay:
☐ Có
☐ Không
10. Trung bình hằng tháng, lương và thưởng của Anh (Chị) là:
Lương:đồng /tháng
Thưởng:đồng /tháng
11. Chi tiêu hằng tháng:
☐ Dưới 3 triệu ☐ Từ 3 đến 5
☐ Từ 5 đến 7 ☐ Từ 7 trở lên
12. Hằng tháng, Anh (Chị) đọc loại báo nào dưới đây (có thể chọn nhiều loại báo):
☐ Hà Nội mới ☐ Lao động ☐ Tiền phong
☐ Thanh niên ☐ Tuổi trẻ
13. Trong 3 môn thể thao đề nghị, Anh (Chị) thích triển khai phong trào thể thao:
☐ Bóng đá
☐ Cầu lông
☐ Tennis

Nếu Anh (Chị) có những ý kiến đóng góp thêm thông tin trong bản câu hỏi, hãy ghi vào khoảng trống dưới đây:

Xin cảm ơn sự cộng tác của các Anh (Chị)!

CHƯƠNG

6

CHỌN MẪU VÀ TỔ CHỨC THU THẬP DỮ LIỆU

(SAMPLING AND DATA COLLECTION)

CHỌN MẪU TRONG NGHIÊN CỨU

Khi thực hiện nghiên cứu trên thực tế, thông thường nhà nghiên cứu không thể thực hiện nghiên cứu trên tổng thể do nhiều nguyên nhân như tổng thể quá lớn. Bên cạnh đó, có một số nghiên cứu phải dùng chính sản phẩm để thực hiện, điều đó không cho phép khi tổng thể quá lớn. Do vậy, cần phải chọn mẫu để thực hiện nghiên cứu. Chương này sẽ giới thiệu mẫu và phương pháp chọn mẫu, một số vấn đề cần phải thực hiện trong quá trình thu thập dữ liệu cũng được trình bày trong chương này.

Một số định nghĩa

Phần tử (elements)

Một phần tử là một đơn vị trong đó thông tin về nó được thu thập và làm cơ sở cho việc phân tích, thông thường trong lấy mẫu nghiên cứu, những phần tử là con người, tuy vậy cũng có những loại phần tử khác như là: gia đình, cửa hàng hoặc doanh nghiệp.

Tổng thể (populations)

Tất cả các quan sát, các phần tử trong bất kỳ lĩnh vực của cuộc điều tra tạo thành một “tổng thể”, một liệt kê đầy đủ của tất cả các phần tử trong “tổng thể” được biết đến như một yêu cầu điều tra tổng thể. Một tổng thể là sự tập hợp các phần tử, có 2 loại tổng thể:

- Tổng thể chủ đích (target population): Là một tổng thể được yêu cầu bởi đặc trưng thông tin cần nghiên cứu.
- Tổng thể lấy mẫu (sampling population): Là một tổng thể thực tế được chọn trên yêu cầu thông tin cần nghiên cứu.

Mẫu (sample)

Khi nghiên cứu thực địa được thực hiện trong thực tế, cân nhắc về thời gian và chi phí hầu như luôn dẫn đến một sự lựa chọn số người trả lời nghĩa là, lựa chọn chỉ một số quan sát nhất định, người trả lời được lựa chọn càng đại diện cho tổng thể càng tốt, mục đích để tạo ra một mặt cắt ngang thu nhỏ biểu thị cho tổng thể. Người trả lời được lựa chọn gọi (theo kỹ thuật) là một “mẫu” và thủ tục lựa chọn gọi là “kỹ thuật lấy mẫu” và khảo sát thực hiện được biết đến như là “điều tra mẫu”. Về mặt toán học, quy mô tổng thể là N và nếu một phần của kích thước n của tổng thể này được chọn theo một vài quy tắc để nghiên cứu một số đặc tính của tổng thể, nhóm

bao gồm các đơn vị này được gọi là “mẫu”. Nhà nghiên cứu phải chuẩn bị một thiết kế mẫu cho nghiên cứu của mình, nghĩa là phải lập kế hoạch một mẫu nên được lựa chọn như thế nào và kích cỡ một mẫu sẽ là bao nhiêu.

Tại sao phải chọn mẫu

- Việc thực hiện một cuộc điều tra luôn liên quan đến chi phí bỏ ra của doanh nghiệp hoặc của nhà nghiên cứu như chi phí, thời gian.... khi mẫu lớn. Một nhà nghiên cứu không thể thực hiện nghiên cứu trên cả tổng thể hoặc trên mẫu quá lớn vì sẽ phát sinh chi phí và tốn thời gian (đặc biệt là thực hiện trong điều kiện hạn hẹp về chi phí và thời gian). Bên cạnh đó, cơ hội chỉ xảy ra một lần và vào một thời điểm nhất định, khi chúng ta thực hiện nghiên cứu trên tổng thể, đôi khi thời gian quá lớn và làm bỏ mất cơ hội.
- Trong một số trường hợp, chỉ tiêu nghiên cứu có nội dung phức tạp, không có điều kiện điều tra ở diện rộng. Nhờ đó, kết quả điều tra thu được trên mẫu sẽ phản ánh được nhiều mặt, cho phép nghiên cứu các mối quan hệ cần thiết của hiện tượng nghiên cứu.
- Đồng thời, khi mẫu quá lớn, việc tiến hành điều tra toàn bộ tổng thể vẫn không thể nâng cao độ chính xác của thông tin trong khi lại tốn kém chi phí và mất nhiều thời gian. Việc nghiên cứu trên mẫu có thể làm giảm sai số phi chọn mẫu (sai số do cân, đong, đo, đếm, khai báo, ghi chép...), các sai số này luôn luôn tồn tại và ảnh hưởng không nhỏ đến chất lượng số liệu thống kê, nhất là các chỉ tiêu có nội dung phức tạp, việc tiếp cận để thu thập số liệu khó khăn, tốn nhiều thời gian trong quá trình phỏng vấn, ghi chép và đặc biệt hơn là đối với các chỉ tiêu điều tra không có sẵn thông tin mà đòi hỏi phải hỏi tường để nhớ lại thì nghiên cứu trên một mẫu quá lớn hay tổng thể không thích hợp.
- Trong những tình huống mà việc kiểm tra, đo lường có thể phá hủy phần tử (ví dụ khi nghiên cứu kiểm tra chất lượng lon sữa thì cần phải phá vỡ lon sữa) thì việc lấy mẫu hoặc một số các hiện tượng kinh tế - xã hội, môi trường,... không thể tiến hành theo phương pháp điều tra toàn bộ.

Hạn chế của điều tra chọn mẫu

- Do việc chọn mẫu nên nghiên cứu chỉ tiến hành thu thập số liệu trên một số đơn vị, sau đó dùng kết quả để suy rộng cho toàn bộ tổng thể chung nên kết quả luôn tồn tại *sai số chọn mẫu*. Sai số chọn mẫu phụ thuộc vào độ đồng đều của chỉ tiêu nghiên cứu, vào cỡ mẫu và phương pháp tổ chức chọn mẫu.
- Kết quả của nghiên cứu trong mẫu có thể không thể tiến hành phân nhỏ theo mọi phạm vi và tiêu thức nghiên cứu như điều tra toàn bộ mà chỉ thực hiện được ở mức độ nhất định tùy thuộc vào cỡ mẫu, phương pháp tổ chức chọn mẫu và độ đồng đều giữa các đơn vị theo các chỉ tiêu được điều tra.

Kỹ thuật chọn mẫu

Có thể chia kỹ thuật chọn mẫu thành hai loại:

- Chọn mẫu xác suất hay gọi là lấy mẫu đại diện (representative): Đây là phương pháp mà xác

suất của mỗi phần tử được chọn từ tổng thể đã biết và thường như nhau cho tất cả các trường hợp.

- Chọn mẫu phi xác suất hay gọi là lấy mẫu phán đoán (judgemental): Xác suất mỗi phần tử được chọn từ tổng thể chưa được biết.

Kỹ thuật chọn mẫu xác suất

Kỹ thuật chọn mẫu xác suất có thể chia thành 4 giai đoạn:

- Xác định khung chọn mẫu (sampling frame): Khung chọn mẫu là danh sách đầy đủ mọi phần tử trong tổng thể mà từ đó chúng ta lấy mẫu, với việc xác định khung chọn mẫu sẽ giúp chúng ta xác định tổng thể cần tổng quát hóa.
- Quyết định kích cỡ lấy mẫu: Sự tổng quát hóa từ mẫu lên tổng thể luôn gắn với một sai số nhất định, kích cỡ mẫu càng lớn thì khả năng sai số càng thấp và ngược lại. Tuy nhiên, việc chọn một mẫu lớn luôn đi kèm với chi phí về thời gian và tiền bạc. Việc quyết định cỡ lấy mẫu phụ thuộc vào (1) độ tin cậy, (2) sai số có thể chấp nhận được, (3) loại (kỹ thuật) phân tích dự định sử dụng và (4) độ lớn của tổng thể.

Bên cạnh đó, cần phải cân nhắc tỷ lệ hồi đáp (trả lời) trong quá trình nghiên cứu, nếu việc dự đoán tỷ lệ không hồi đáp cao, kích cỡ lấy mẫu phải cao để đảm bảo độ tin cậy của nghiên cứu. Tỷ lệ không hồi đáp có thể do từ chối hồi đáp, không đủ điều kiện hồi đáp, không có khả năng xác định địa chỉ hoặc địa chỉ người hồi đáp được xác định nhưng không thể tiếp cận. Khi đó, tỷ lệ hồi đáp thực sự sẽ là:

$$\text{Tỷ lệ hồi đáp} = \frac{\text{Tổng số người hồi đáp}}{\text{thực sự (f) Tổng số người trong mẫu - Số không đủ điều kiện - Số không tiếp cận được}}$$

Trong quá trình nghiên cứu, cần phải ước lượng tỷ lệ hồi đáp thực sự để tính toán kích cỡ mẫu cần lựa chọn để thực hiện nghiên cứu (ví dụ 55%). Giả sử nhà nghiên cứu muốn thực hiện với một mẫu tối thiểu n_{tt} là 500 để đảm bảo cho kỹ thuật sử dụng phân tích. Khi đó, cỡ mẫu cần thiết lựa chọn để thực hiện (điều tra) tối thiểu n là:

$$\text{Cỡ mẫu thực hiện điều tra (n)} = \frac{n_{tt} * 100\%}{f} = \frac{500 * 100}{55\%} = 909$$

Tỷ lệ hồi đáp có ảnh hưởng rất lớn đến chất lượng của nghiên cứu, việc không trả lời có thể do là bản câu hỏi không đến được nơi nhận hoặc người nhận được bản câu hỏi nhưng từ chối trả lời. Chính vì thế, các nỗ lực nhằm gia tăng tỷ lệ trả lời luôn là vấn đề quan trọng để giảm bớt sai số do không trả lời. Sau đây là một số giải pháp nhằm cố gắng gia tăng tỷ lệ trả lời:

- Nên gửi thư trước để thông báo cho người phỏng vấn và đề nghị họ hợp tác.
- Đối với hình thức phỏng vấn bằng thư tín, khi gửi thư, cần lưu ý hình thức và chất lượng bao thư, tem.
- Động viên bằng tặng phẩm hoặc tiền sẽ làm gia tăng tỷ lệ trả lời.

- Cần đảm bảo việc dấu tên hay giữ bí mật trong những trường hợp liên quan đến vấn đề riêng tư..
 - Huấn luyện, tuyển chọn nhân viên phỏng vấn kỹ lưỡng sẽ góp phần tăng hiệu quả tiếp xúc.
- Lựa chọn kỹ thuật chọn mẫu và lựa chọn mẫu: Có năm kỹ thuật có thể được sử dụng trong chọn mẫu xác suất:

Chọn mẫu ngẫu nhiên đơn giản (simple random)

Chọn mẫu ngẫu nhiên đơn giản hay còn gọi là chọn mẫu ngẫu nhiên bao gồm việc lựa chọn mẫu một cách ngẫu nhiên từ khung mẫu bằng các bảng số ngẫu nhiên sao cho mỗi đơn vị lấy mẫu trong cấu trúc có một cơ hội hiện diện trong mẫu bằng nhau. Để thực hiện phương pháp lấy mẫu này, cần đánh số với mỗi phần tử trong khung mẫu ứng với một số duy nhất từ 1 đến n và tiến hành lựa chọn các phần tử tương ứng với các con số cho đến khi đạt được kích cỡ mẫu thực hiện điều tra.

Chọn mẫu ngẫu nhiên có hai loại: lấy mẫu ngẫu nhiên thay thế hoặc không thay thế. Trong lấy mẫu ngẫu nhiên thay thế là hình thức lấy mẫu mà phần tử được đưa trở lại khung mẫu sau khi phần tử đó đã được chọn, mẫu không thay thế là hình thức không đưa phần tử trở lại vào khung mẫu. Trong thực tế nghiên cứu marketing, lấy mẫu ngẫu nhiên không thay thế được sử dụng chủ yếu và người ta thường dùng máy tính để hình thành các con số ngẫu nhiên, giúp cho việc lựa chọn các phần tử. Phương pháp chọn mẫu ngẫu nhiên đơn giản được áp dụng có kết quả khi tổng thể nghiên cứu không phân tán quá rộng về mặt địa lý; các phần tử trong tổng thể có khá nhiều sự đồng nhất về đặc điểm muốn nghiên cứu.

Theo Giới và cộng sự (2006), ưu điểm của phương pháp chọn mẫu ngẫu nhiên đơn giản là dễ hiểu, dễ thực hiện; trung bình mẫu là một sự tính toán khách quan của trung bình tổng thể nghiên cứu; phương pháp tính toán đơn giản, dễ dàng. Nhược điểm của phương pháp chọn mẫu ngẫu nhiên là (1) trong nhiều trường hợp, sự biến thiên của tổng thể nghiên cứu rất rời rạc và không theo quy tắc, thì lấy mẫu ngẫu nhiên không được dùng đến vì nó kém chính xác; mẫu có thể không mang tính đại diện, hoặc bị lệch; (2) để lựa chọn các phần tử, cần phải đánh dấu và lập danh sách toàn bộ tổng thể để sử dụng bảng số ngẫu nhiên, bốc thăm, quay số,... công việc này khó thực hiện được khi tổng thể là quá lớn và (3) mẫu được chọn có thể bị phân tán, do vậy tốn kém chi phí và khó khăn trong đi lại khi thu thập dữ liệu.

Chọn mẫu có hệ thống (systematic)

Chọn mẫu có hệ thống với sự bắt đầu ngẫu nhiên là một phương pháp chọn mẫu được tiến hành bằng cách lấy từng đơn vị thứ k từ một tổng thể nghiên cứu có thứ tự. Đơn vị đầu tiên được chọn một cách ngẫu nhiên, k được gọi là khoảng cách lấy mẫu, số nghịch đảo $1/k$ là tỷ lệ lấy mẫu. Phương pháp chọn mẫu có hệ thống thường được áp dụng khi thứ tự của các đơn vị lấy mẫu là ngẫu nhiên.

Ưu điểm của phương pháp chọn mẫu có hệ thống là chính xác, dễ tiếp cận, thích hợp cho

mọi kích cỡ, chi phí thấp. Nhược điểm của phương pháp chọn mẫu này là có thể một mẫu được lấy chỉ bao gồm những đơn vị có cùng một dạng và cần thiết phải có danh sách các đơn vị lấy mẫu theo thứ tự.

Chọn mẫu ngẫu nhiên phân tầng (stratified random)

Kỹ thuật chọn mẫu ngẫu nhiên phân tầng là kỹ thuật phân chia tổng thể thành hai hay nhiều tầng quan trọng và có ý nghĩa, dựa trên một hay một số thuộc tính. Việc chia tổng thể thành một loạt các tầng có nghĩa là mẫu sẽ có tính đại diện hơn và lúc đó có thể chắc chắn rằng mỗi tầng được đại diện theo tỷ lệ trong mẫu. Tùy theo đặc điểm nghiên cứu, tổng thể có thể được phân tầng theo nhiều tiêu thức khác nhau; và có thể phân tầng một cấp (một tiêu thức) hoặc nhiều cấp (nhiều tiêu thức); và khi chọn mẫu ngẫu nhiên phân tầng có thể theo tỷ lệ (tỷ lệ mẫu tương ứng với tỷ lệ tổng thể) hoặc không theo tỷ lệ.

Ưu điểm của phương pháp chọn mẫu ngẫu nhiên phân tầng: (1) sự phân nhóm có thể làm gia tăng mức độ chính xác của việc đánh giá các đặc điểm tổng thể nghiên cứu; (2) thực hiện thuận tiện, phân tích số liệu khá toàn diện. Nhược điểm của phương pháp này là cần phải lập danh sách các đơn vị lấy mẫu theo từng nhóm; tốn kém chi phí đi lại, đặc biệt khi tổng thể nghiên cứu trải rộng trên một vùng địa lý rộng lớn. Với những ưu điểm và nhược điểm trên, phương pháp chọn mẫu ngẫu nhiên phân tầng thường được áp dụng khi tổng thể nghiên cứu có sự phân bố của đặc điểm nghiên cứu rất rời rạc, hay tập trung trên những điểm nhỏ bị phân tán của tổng thể.

Chọn mẫu theo cụm (cluster)

Chọn mẫu theo cụm nhìn có vẻ tương tự như phương pháp chọn mẫu phân tầng vì cần chia tổng thể thành những nhóm riêng biệt trước khi lấy mẫu. Tuy nhiên, theo Giới và cộng sự (2006), các cụm được tạo nên bởi những phần tử dị biệt, không đồng nhất, miễn sao mỗi nhóm sẽ là đặc trưng của tổng thể nghiên cứu. Ví dụ, nghiên cứu về sinh viên trong một trường đại học, thay vì chọn các phần tử là sinh viên theo kích thước mẫu, có thể chọn đơn vị lấy mẫu là lớp; do vậy không cần phải lập danh sách sinh viên, mà lập danh sách các lớp. Khi thực hiện điều tra, thì tất cả sinh viên trong một lớp được chọn đều được tiếp xúc.

Ưu điểm của phương pháp chọn mẫu theo cụm là không cần thiết phải xây dựng một danh sách tất cả các phần tử trong tổng thể nghiên cứu mà cấu trúc đối với lấy mẫu theo cụm là một danh sách các cụm. Ngay cả khi danh sách các phần tử đã có sẵn, việc lấy mẫu theo cụm vẫn ít tốn kém hơn về chi phí. Nhược điểm của phương pháp này là ở chỗ trong thực tế, lấy mẫu theo cụm không hiệu quả bằng lấy mẫu ngẫu nhiên hay phân tầng. Chẳng hạn, những hộ gần kề nhau thường có đặc điểm tương tự nhau hơn những hộ riêng biệt, điều này sẽ ảnh hưởng đến tính đại diện của mẫu, và được thể hiện qua sai số chọn mẫu tăng.

Chọn mẫu nhiều giai đoạn (multi-stage)

Chọn mẫu phân tầng là sự phát triển của kỹ thuật lấy mẫu theo cụm, bao gồm một chuỗi lấy mẫu theo cụm, mỗi lần lấy mẫu có dạng kỹ thuật lấy mẫu ngẫu nhiên. Phương pháp này thường được sử dụng để khắc phục những vấn đề của một tổng thể phân tán về địa lý,

khi việc tiếp xúc trực tiếp là cần thiết hay việc xây dựng một khung mẫu cho một khu vực địa lý rộng lớn tốn kém và mất nhiều thời gian.

- Kiểm tra tính đại diện của mẫu: Thông thường, việc kiểm tra tính đại diện của mẫu bằng cách so sánh dữ liệu thu thập được từ mẫu với dữ liệu từ nguồn khác của tổng thể.

Kỹ thuật chọn mẫu phi xác suất

Chọn mẫu theo hạn mức (quota)

Chọn mẫu theo hạn mức là hình thức chọn mẫu mang tính phi ngẫu nhiên và thường được sử dụng cho các cuộc khảo sát bằng phỏng vấn. Kỹ thuật này thực hiện bằng cách (1) chia tổng thể thành những nhóm cụ thể, (2) tính hạn mức cho mỗi nhóm, (3) giao hạn mức cho mỗi người phỏng vấn và (4) tổng hợp dữ liệu thu thập bởi những người phỏng vấn để cung cấp một mẫu đầy đủ. Ưu điểm của phương pháp lấy mẫu theo hạn mức là ít tốn kém và có thể hình thành nhanh chóng nên thường được ứng dụng cho những tổng thể lớn.

Chọn mẫu theo mục đích (purpose)

Chọn mẫu theo mục đích hay chọn mẫu phán đoán cho phép sử dụng phán đoán để lựa chọn các phần tử. Theo phương pháp chọn mẫu phán đoán, những phần tử lấy mẫu được chọn dựa vào điều mà nhà chuyên môn suy nghĩ có thể thỏa mãn một tiêu chuẩn nào đó. Có hai hình thức lựa chọn phán đoán: lấy mẫu theo dư luận và phán đoán thống kê.

Chọn mẫu tích lũy nhanh (snowball)

Kỹ thuật lấy mẫu mở rộng dần thường được sử dụng khi khó xác định các thành viên của tổng thể mong muốn. Theo phương pháp này, những đơn vị lấy mẫu (hay phần tử) ban đầu được lựa chọn bằng cách sử dụng các phương pháp xác suất, nhưng những đơn vị bổ sung tiếp đó được xác định từ thông tin được cung cấp bởi các đơn vị lấy mẫu ban đầu. Tuy nhiên cách chọn mẫu mở rộng dần (snowball) có thể có sai lệch vì những người được giới thiệu ra thường có một số đặc điểm tương đồng về nhân khẩu học hay tâm lý, sở thích.

Lấy mẫu tự chọn (self-selection)

Lấy mẫu tự chọn sử dụng khi nghiên cứu cho phép mỗi phần tử (thường là các cá nhân) xác định mong muốn được tham gia vào nghiên cứu. Vì vậy, cần công bố nhu cầu cần các phần tử trên các phương tiện truyền thông và tập hợp các phần tử tham gia để thu thập dữ liệu.

Chọn mẫu thuận tiện

Theo cách chọn mẫu này, người nghiên cứu chọn ra các đơn vị lấy mẫu dựa vào “sự thuận tiện” hay “tính dễ tiếp cận”. Với phương pháp chọn mẫu thuận tiện, rất khó xác định tính đại diện của mẫu. Sự lựa chọn các đơn vị mẫu mang tính chủ quan của người nghiên cứu vì thế độ chính xác và độ tin cậy không cao, ít được sử dụng rộng rãi. Sau khi đã xác định mẫu, với công cụ đã được thiết kế, nhà nghiên cứu sẽ tổ chức thu thập dữ liệu nhằm đảm bảo yêu cầu và mục tiêu nghiên cứu.

TỔ CHỨC THU THẬP DỮ LIỆU

Sau khi hoàn tất việc lập kế hoạch thu thập dữ liệu và chọn mẫu nghiên cứu, nhà nghiên cứu sẽ tiến hành thu thập dữ liệu. Trong quá trình thu thập cần chú ý đến việc (1) giảm thiểu những sai sót và (2) tối đa hóa thông tin thu được. Đặc biệt, nghiên cứu luôn phải cân nhắc nhằm đảm bảo vấn đề (1) thời gian, (2) chi phí luôn hạn chế và (3) môi trường thu thập rộng lớn và luôn thay đổi làm ảnh hưởng đến động cơ hành động, mức độ tích cực của người phỏng vấn... và có thể làm giảm tính chính xác của những thông tin thu thập được. Có nhiều phương pháp để thu thập dữ liệu như:

- Quan sát,
- Phỏng vấn cá nhân trực tiếp,
- Phỏng vấn bằng điện thoại,
- Phỏng vấn bằng thư tín,
- Phỏng vấn qua mạng Internet,
- Phỏng vấn qua mạng nội bộ.

Mỗi phương pháp đều có những ưu nhược điểm nhất định, vì vậy, khi thiết kế nghiên cứu cần phải xác định phương pháp thu thập dữ liệu là phương pháp nào để thực hiện việc thiết kế công cụ thu thập dữ liệu và phác thảo quy trình thực hiện việc thu thập dữ liệu phù hợp. Một số sai sót có thể xảy ra làm ảnh hưởng đến chất lượng thông tin, tỷ lệ hồi đáp... và cần chú ý khi thực hiện việc tổ chức thu thập dữ liệu:

- Do lựa chọn đối tượng: Do việc tiếp xúc không đúng người trong phỏng vấn qua điện thoại hay nhầm địa chỉ trong phỏng vấn qua thư...
- Do thực hiện phỏng vấn không đầy đủ: Do người phỏng vấn cố tình bỏ bớt bộ phận người phải được phỏng vấn trong phỏng vấn trực tiếp, không có người nhận trong phỏng vấn qua thư tín hoặc không có điện thoại, Internet...
- Do không trả lời: Khi người được phỏng vấn không muốn cung cấp thông tin do phỏng vấn ở nơi không thích hợp, điều tra những vấn đề mang tính riêng tư, tế nhị nhưng thể kế công cụ thu thập không tốt, người trả lời sợ nhận diện.
- Do giao tiếp không tốt hoặc thiết lập bầu không khí không tốt giữa người phỏng vấn và người được phỏng vấn.
- Do sai sót trong lúc ghi chép.
- Do giả mạo của người phỏng vấn...

Để cho việc thu thập dữ liệu được tốt, các nhà nghiên cứu cần phải quản trị tốt việc tổ chức thu thập dữ liệu tại hiện trường từ việc chuẩn bị các công cụ thu thập dữ liệu, tổ chức việc khảo sát thử, tuyển chọn người phỏng vấn, hướng dẫn, giám sát và kiểm tra để kịp thời phát hiện, điều chỉnh, thực hiện lại việc thu thập dữ liệu ngay tại thời điểm và hiện trường mong muốn thu thập

dữ liệu để đảm bảo thời gian và chi phí trong quá trình thực hiện.

TÓM TẮT CHƯƠNG 6

Việc thực hiện điều tra toàn bộ tổng thể mục tiêu rất khó thực hiện vì những giới hạn về ngân sách và thời gian và qui mô lớn của tổng thể. Do đó, chọn mẫu được thực hiện dựa trên cơ sở cho rằng nếu chọn mẫu có tính đại diện cho tổng thể thì những thông tin thu được trên mẫu sẽ có thể suy rộng ra tổng thể. Để có thể tiến hành chọn mẫu cần trả lời các câu hỏi (1) tổng thể mục tiêu là những ai; (2) liệu có biết được khung lấy mẫu không; (3) nên lựa chọn phương pháp lấy mẫu phù hợp nào; và (4) qui mô mẫu bao nhiêu là cần thiết.

Có hai phương pháp lấy mẫu là phương pháp lấy mẫu xác suất và phương pháp lấy mẫu phi xác suất. Các kỹ thuật trong phương pháp lấy mẫu xác suất được dựa trên sự lựa chọn ngẫu nhiên, trong đó xác suất mà mỗi phần tử được chọn vào mẫu là được biết trước. Kỹ thuật đơn giản nhất là chọn mẫu ngẫu nhiên đơn giản, tiếp theo là chọn mẫu ngẫu nhiên phân tầng, chọn mẫu theo hệ thống, chọn mẫu theo cụm và chọn mẫu nhiều giai đoạn. Tuy nhiên, phương pháp lấy mẫu xác suất là khó thực hiện vì qui mô tổng thể mục tiêu thường không được biết. Do vậy, phương pháp lấy mẫu phi xác suất được sử dụng khá rộng rãi vì tính tiện lợi của nó. Các kỹ thuật lấy mẫu xác suất bao gồm chọn mẫu theo hạn mức, theo mục đích, tích lũy nhanh, tự chọn và thuận tiện.

Quá trình thu thập dữ liệu đóng vai trò quan trọng trong nghiên cứu vì ở giai đoạn này, thông tin phải đảm bảo sự khách quan, trung thực và sai sót được tối thiểu hóa. Chương này đã nhắc lại các phương pháp thu thập dữ liệu và chú ý về những sai sót khi thực hiện thu thập dữ liệu như: (1) sai sót do lựa chọn đối tượng; (2) sai sót do thực hiện phỏng vấn không đầy đủ; (3) sai sót do không trả lời; (4) sai sót do giao tiếp không tốt; (5) sai sót do lúc ghi chép; (6) sai sót do người phỏng vấn giả mạo.

7

PHẦN MỀM ỨNG DỤNG TRONG PHÂN TÍCH DỮ LIỆU: PHẦN MỀM SPSS

Ngày nay, việc ứng dụng công nghệ thông tin để phân tích dữ liệu trong nghiên cứu khoa học là hết sức phổ biến, có một số phần mềm được sử dụng để phân tích dữ liệu trong nghiên cứu khoa học như EVIEW, STATA, SPSS... mỗi loại đều có những ưu nhược điểm nhất định. Do vậy, trước khi thực hiện nghiên cứu cần xác định phần mềm nào được sử dụng trong quá trình phân tích để đạt được hiệu quả cao nhất. Quyển sách này sẽ giới thiệu ứng dụng phần mềm SPSS (*Statistical Product and Service Solutions* trước đây là *Statistical Package for Social Sciences*) để phân tích dữ liệu. Ưu điểm của phần mềm này là tính đa năng và linh hoạt trong việc lập các bảng, sử dụng các mô hình phân tích đồng thời loại bỏ một số công đoạn (bước) không cần thiết mà một số phần mềm khác gặp phải.

Trong nhiều trường hợp, nhà phân tích dữ liệu có thể là nhà nghiên cứu hoặc không phải là nhà nghiên cứu (hay là nhà phân tích), để đảm bảo tính chính xác của quá trình phân tích, nhà nghiên cứu (hoặc phân tích) cần phải lập kế hoạch phân tích, đạt thực hiện được điều đó, nhà phân tích cần phải:

- Nắm vững mục tiêu nghiên cứu;
- Nắm vững các cấp độ thang đo và đặc điểm của cấp độ thang đo sử dụng trong việc thiết kế công cụ thu thập dữ liệu (bản câu hỏi...);
- Cấu trúc của công cụ thu thập dữ liệu;
- Liệt kê các công cụ, phương pháp thực hiện;
- Xem xét thời gian, chi phí... để hoàn thành việc phân tích.
- Thực hiện việc làm sạch dữ liệu và mã hóa một cách khoa học.

Trên cơ sở xác định bản câu hỏi và kế hoạch phân tích dữ liệu, quá trình chuẩn bị dữ liệu, nhập liệu và phân tích có thể thông qua một số công đoạn như sau:

KIỂM TRA CÔNG CỤ THU THẬP DỮ LIỆU

Sau khi tiến hành thu thập dữ liệu phải kiểm tra các dữ liệu để bảo đảm chúng có ý nghĩa, tức là có giá trị đối với việc xử lý và phân tích, việc làm cho dữ liệu có giá trị tiến hành theo hai bước:

- Tiến hành xem xét một cách kỹ lưỡng các phương pháp và các công cụ đã được sử dụng để thu thập dữ liệu (tức kiểm tra các công cụ dùng để thu thập dữ liệu)
- Tiến hành nghiên cứu kỹ các bản câu hỏi đã được phỏng vấn và những chỉ dẫn về thủ tục phỏng vấn để phát hiện ra những nguyên nhân dẫn đến các sai sót. Vì vậy, lời khuyên đối với người tổ chức thu thập dữ liệu là cần phải kiểm tra các sai sót ngay sau khi thu thập, nếu có xảy ra, cần tiến hành phỏng vấn lại ngay (nếu thời gian và chi phí cho phép).

HIỆU CHỈNH DỮ LIỆU

Do những nguyên nhân khách quan và chủ quan, quá trình thu thập dữ liệu dù được chuẩn bị chu đáo vẫn còn có thể tồn tại những sai sót, vì vậy cần phải hiệu chỉnh (data editing) để dữ liệu có ý nghĩa và tăng chất lượng đối với quá trình nghiên cứu. Hiệu chỉnh dữ liệu là sửa chữa các sai sót về ghi chép hoặc ngôn từ phát hiện được qua kiểm tra. Trong quá trình hiệu chỉnh, cần sửa chữa những sai sót phổ biến như: những cuộc phỏng vấn giả tạo do người đi phỏng vấn “tự đưa ra” như câu trả lời không đầy đủ (là những câu trả lời không rõ ý hoặc bỏ nửa chừng); những câu trả lời thiếu nhất quán (không thống nhất trong ý kiến trả lời giữa phần trước và phần sau); những câu trả lời không thích hợp; những câu trả lời không đọc được...

Việc hiệu chỉnh dữ liệu có thể (1) hiệu chỉnh tại hiện trường hoặc (2) hiệu chỉnh tại nơi xử lí. Có 3 cách được sử dụng để hiệu chỉnh các dữ liệu "xấu" là:

- Quay trở lại người đi phỏng vấn hay người trả lời câu hỏi để làm sáng tỏ vấn đề. Điều này sẽ làm nảy sinh hai vấn đề:
 - + Tăng thời gian và chi phí của dự án nghiên cứu: Có nhiều dự án nghiên cứu, khi thực hiện việc nghiên cứu lại làm nảy sinh nhiều chi phí và thời gian (ví dụ những nghiên cứu khó tiếp xúc đối tượng, phải chờ đợi thời gian dài mới có thể thu thập được...). Theo kinh nghiệm, người nghiên cứu có thể không cần tìm cách thu thập lại dữ liệu nếu tỉ lệ các câu trả lời nghi vấn tương đối nhỏ và/ hoặc quy mô của mẫu tương đối lớn (tỉ lệ các câu hỏi nghi vấn nhỏ hơn 20% và mẫu lớn hơn 500).
 - + Dữ liệu thu thập lần sau không thống nhất những dữ liệu đã được thu thập trong cuộc phỏng vấn đầu tiên do các cá nhân có thể không nhớ thông tin cần thiết cũng như có thể do sử dụng phương pháp khác và điều này ảnh hưởng rất lớn đến kết quả của câu trả lời (liên quan đến độ tin cậy của cuộc điều tra).
- Suy luận từ những câu trả lời khác của người trả lời: Để tiến hành cách tiếp cận này, người hiệu chỉnh phỏng đoán từ các dữ liệu khác để thay thế người được phỏng vấn trả lời các câu hỏi còn thiếu nhưng đây là cách làm đầy rủi ro, khó có thể minh định được các quy luật để suy luận các câu trả lời. Do đó, để an toàn khi hiệu chỉnh dữ liệu, người nghiên cứu cần hết sức thận trọng với phương pháp này và không nên suy luận một câu trả lời trừ phi biết tương đối chắc chắn về ý định của người trả lời.
- Loại toàn bộ câu trả lời: Trong trường hợp quy mô của mẫu tương đối lớn, người hiệu chỉnh có thể loại bỏ toàn bộ các câu trả lời nếu thông tin thiếu nhất quán và người hiệu chỉnh không thể giải quyết vấn đề thiếu nhất quán đó trong các dữ liệu được thu thập từ các đối tượng phỏng vấn. Đây là cách thực hiện rất dễ dàng, tuy nhiên, khuyết điểm trong cách tiếp cận này là sự thiên vị trong kết quả nếu những người trả lời thiếu nhất quán đó bị loại ra khỏi cuộc nghiên cứu, khi đó kết quả đạt được sẽ bị lệch nếu ý kiến những người trả lời bị loại này khác với những người còn được giữ lại trong mẫu điều tra.

Do vậy, người giám sát quá trình phỏng vấn đóng vai trò rất quan trọng trong việc kiểm tra:

- Sự hoàn thiện của các bản câu hỏi;
- Tính hợp lí giữa các câu hỏi và câu trả lời;

- Tính rõ ràng của các câu trả lời;
- Tính nghiêm túc trong quá trình phỏng vấn của phỏng vấn viên.

MÃ HÓA DỮ LIỆU

Mã hóa dữ liệu là việc quyết định chọn các mã số cho các câu hỏi và các phương án trả lời, có hai hình thức mã hóa là mã hóa trước (precoding) và mã hóa sau (postcoding):

- Mã hóa trước là hình thức tiến hành lựa chọn các mã số cho các câu hỏi và phương án trả lời từ khi thiết kế bản câu hỏi và in ngay các mã số lên bản câu hỏi. Hình thức mã hóa này thích hợp khi thiết kế câu hỏi sử dụng hầu hết các câu hỏi dạng câu hỏi đóng. Ví dụ:

1. Anh (Chị) đang công tác tại bộ phận:	Code
☐ Kinh doanh	1
☐ Kỹ thuật	2
☐ Kế toán	3
☐ Văn phòng	4
☐ Khác, cụ thể là:	5
2. Giới tính:	
☐ Nữ	0
☐ Nam	1
3. Khi sử dụng sản phẩm A, Anh (Chị) cảm thấy:	
☐ Rất không thích	1
☐ Không thích	2
☐ Vừa thích vừa không	3
☐ Thích	4
☐ Rất thích	5

- Mã hóa sau là hình thức tiến hành lựa chọn các mã số cho các câu hỏi và phương án trả lời sau khi thiết kế bản câu hỏi. Sau khi mã hóa, cần kiểm tra danh bạ mã hóa để đảm bảo tính chính xác của bộ mã hóa, đồng thời có thể làm bảng danh mục cho nhiều người nhập liệu sau này.

SẮP XẾP VÀ ĐÁNH SỐ THỨ TỰ BẢN CÂU HỎI

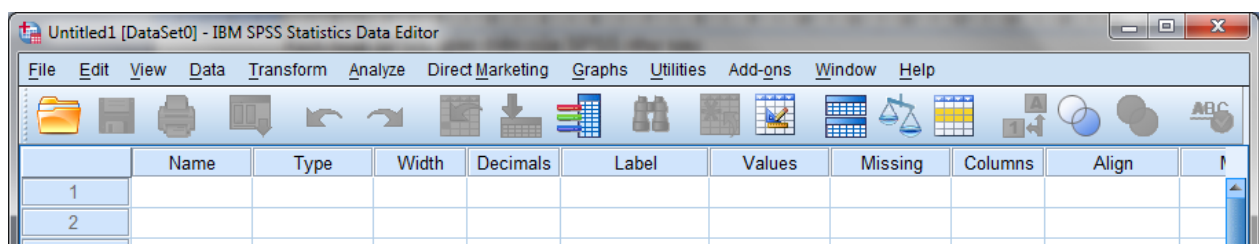
Để thuận tiện cho việc nhập liệu, phân tích, kiểm soát và điều chỉnh các lỗi (nếu có) do quá trình nhập liệu, người phân tích cần chú ý sắp xếp công cụ thu thập dữ liệu theo tiêu chí mong muốn của người phân tích và đánh số thứ tự bản câu hỏi. Khi nhập liệu, cần yêu cầu người trực tiếp nhập phải đảm bảo đúng số thứ tự trên bản câu hỏi và bản ghi trong phần mềm SPSS để thuận tiện cho việc kiểm soát dữ liệu khi có nhu cầu.

NHẬP LIỆU TRÊN SPSS

Trước khi tiến hành phân tích, cần cài đặt phần mềm SPSS vào máy tính, phần mềm kích hoạt chương trình SPSS sẽ được lưu trong phần **Program Files**.

Giao diện nhập liệu

Trong phần **All Programs**, kích hoạt phần mềm **SPSS** và giao diện của **SPSS** như sau:



Trong đó:

- **Name:** Tên biến (khi đặt trên biến cần phải dễ dàng nhận biết, không có kí tự trống hoặc kí tự đặc biệt).
- **Type:** Kiểu của bộ mã hóa.
- **Width:** Độ lớn của bộ mã.
- **Decimal:** Số phần dư.
- **Labels:** Nhãn của biến, trong phần này nhà phân tích sẽ đặt tên nhãn có ý nghĩa sao cho dễ dàng nhận ra biến số (tương ứng với tên của nhãn).
- **Values:** Giá trị mã hóa (value) tương ứng với nhãn giá trị (value label) của nó.
- **Missing:** Ký hiệu câu trả lời đúng ra phải trả lời nhưng bị bỏ qua (lỗi), chú ý là giá trị này phải có nét đặc thù riêng biệt so với giá trị khác để phân biệt trong quá trình tính toán.
- **Column:** Thiết đặt độ lớn của cột mang tên biến và vị trí nhập liệu của biến này.
- **Measure:** Cấp độ thang đo lường, trên cơ sở 4 cấp độ thang đo lường đã giới thiệu ở chương trước gồm cấp độ thang đo biểu danh, thứ tự, khoảng cách và tỉ lệ, với phần mềm SPSS sẽ phân ra thành 3 thang đo gồm biểu danh (nominal), thứ tự (ordinal) và scale (gồm cấp độ thang đo khoảng cách và tỉ lệ).

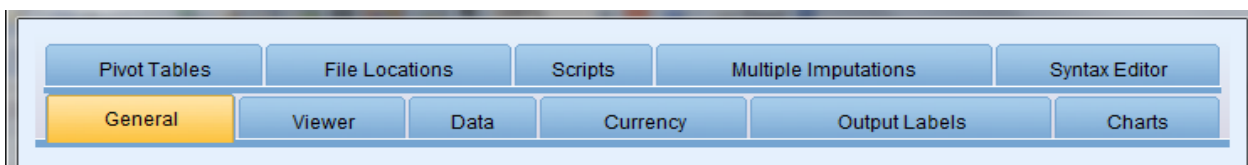
Đồng thời, nhà nghiên cứu cũng quy định thang đo định tính gồm các biến được đo lường bởi cấp độ thang đo *biểu danh* hoặc *thứ tự*, thang đo định lượng được đo lường bởi cấp độ thang *khoảng cách* hoặc *tỉ lệ* (trong SPSS là scale).

Đối với các nghiên cứu ở Việt Nam, quá trình nhập liệu và viết báo cáo đều bằng Tiếng Việt nên để thuận lợi cho việc soạn thảo báo cáo sau này, trước khi định nghĩa các biến cần phải cài đặt phần đánh Tiếng Việt (bằng Unicode) trên phần mềm SPSS. Chú ý, chỉ có phần mềm từ phiên bản 16 trở lên mới cài đặt phần đánh Tiếng Việt trực tiếp trên SPSS, các phần mềm cũ hơn phải sử dụng những file ngôn ngữ bổ sung.

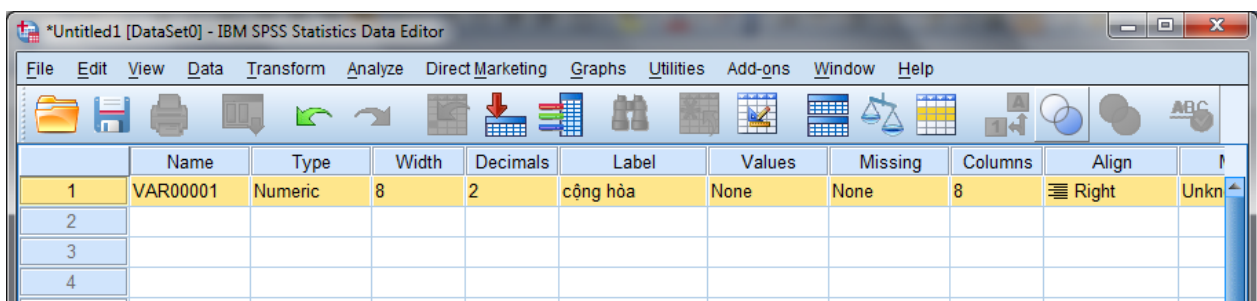
Để tiến hành cài đặt Tiếng Việt Unicode, cần đóng tất cả các cửa sổ liên quan đến SPSS sau đó kích vào **All Programs** để mở cửa sổ SPSS (chú ý không thực hiện bất kỳ một thao tác nào trước khi thực hiện cài đặt Unicode).

Trên thanh trình đơn:

- Kích **Edit/ Options** cửa sổ **Option** sẽ mở ra
- Tại thẻ **General**, kích vào **Unicode (universal character set)** trong khung **Character Encoding for Data and Syntax** phía cuối, bên trái màn hình.
- Kích vào **OK** để hoàn tất việc thiết đặt kiểu đánh Tiếng Việt.



Kích vào biểu tượng **Untitled1** ở phía dưới trang màn hình để mở cửa sổ **Untitled1**, ở bản ghi đầu tiên dưới phần **Label**, đánh chữ Tiếng Việt có dấu (ví dụ: cộng hòa) và kích **Enter**, nếu phần mềm SPSS chấp nhận (con trỏ chuyển sang ô tiếp theo) thì việc cài đặt kiểu gõ Tiếng Việt bằng Unicode đã được hoàn tất, nếu không, đóng hết tất cả các cửa sổ liên quan đến SPSS và thực hiện lại từ đầu theo như quy trình đã hướng dẫn.



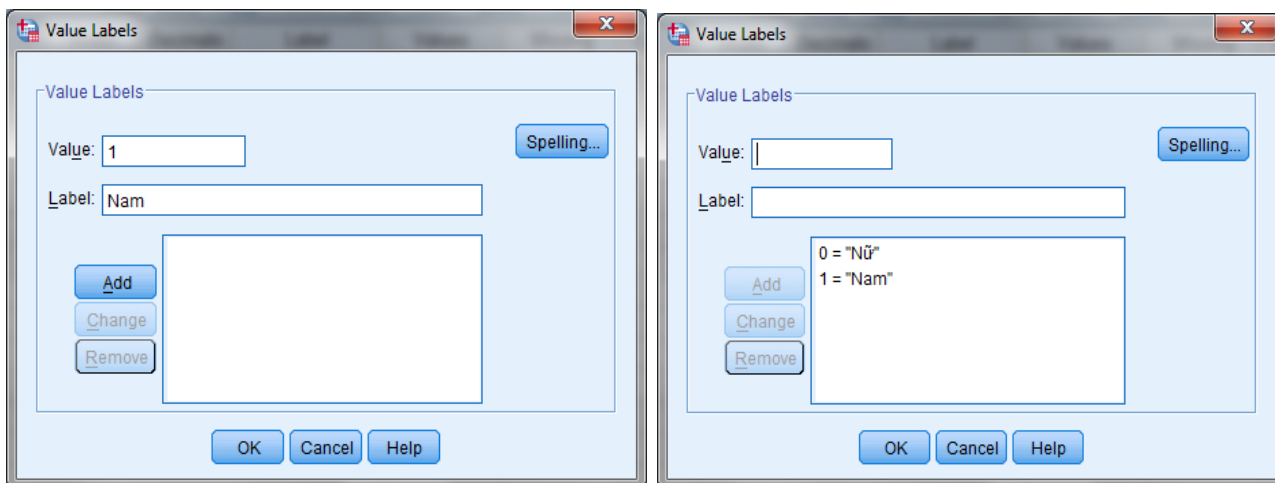
Định nghĩa biến

Sau khi thực hiện việc thiết đặt cách đánh Tiếng Việt trong SPSS hoàn tất, trước khi tiến hành nhập dữ liệu vào phần mềm, cần tiến hành định nghĩa các biến.

Ví dụ 1: Định nghĩa biến Giới tính

Biến giới tính theo cách thức mã hóa được đặt tên là GIOITINH và được mã hoá với 1 = Nam và 0 = Nữ. Trong ô đầu tiên (Name), gõ tên biến là GIOITINH. Tại ô tương ứng với **Label**, đặt

tên nhân là Giới tính. Loại mã hóa (Type) là numeric, không có phần thập phân (Decimals = 0).
 Kích vào ô tương ứng với **Values**, cửa sổ **Value Label** xuất hiện, tại ô Value, nhập 1 tương ứng với **Label** là Nam, nhấn **Add**, tương tự nhập liệu cho Nữ = 0, nhấn **Add** và **OK**. Biến giới tính nhận giá trị 0, 1 là biến định tính với cấp độ thang đo biểu danh nên chọn **Nominal** ở **Measure**.



Chú ý: Các biến liên tục (nhận thang đo tỷ lệ) không cần định nghĩa các giá trị của biến bởi vì mỗi giá trị trong các giá trị trả lời của phỏng vấn viên là một giá trị của biến.

Thực hiện tương tự đối với các biến trong danh mục mã hóa và cần chú ý mỗi biến tương ứng với thang đo gì ở được thiết lập ở phần Measure.

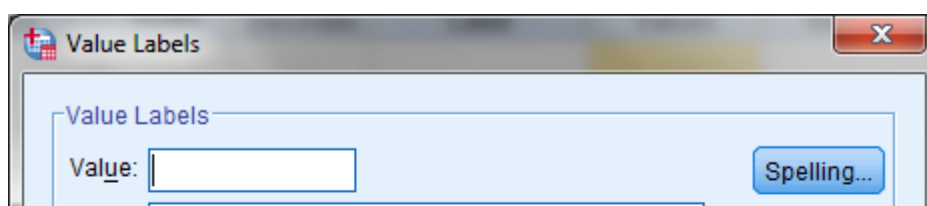
Ví dụ 2: Định nghĩa biến được đo lường như sau

	Hoàn toàn không đồng ý	Không đồng ý	Không ý kiến	Đồng ý	Hoàn toàn đồng ý
Khi Anh (Chị) sử dụng mạng điện thoại viễn thông A, Anh (Chị) :					
1. Kết nối ngay lần gọi đầu tiên	1	2	3	4	5
2. Thông suốt trong suốt cuộc gọi	1	2	3	4	5
3. Kết nối âm thanh rõ ràng	1	2	3	4	5
4. Vùng phủ sóng rộng	1	2	3	4	5
5. Thực hiện được vào giờ cao điểm	1	2	3	4	5

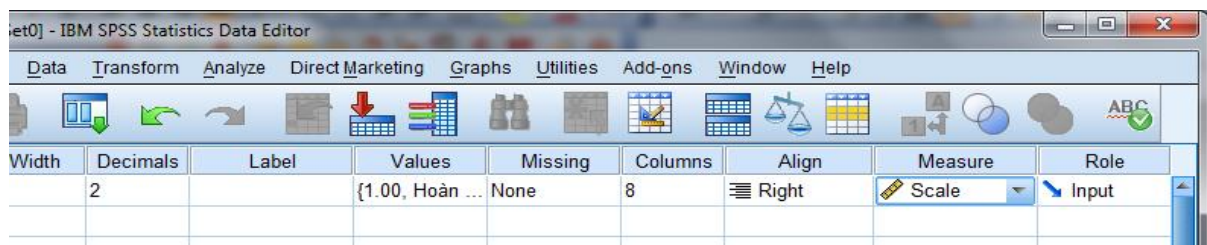
Đây là 6 chỉ báo đo lường chất lượng dịch vụ, do vậy có thể đặt trên biến là CLDV1, CLDV2,..., CLDV6 với các giá trị (1) là hoàn toàn không đồng ý, (2) không đồng ý, (3) không ý kiến, (4) đồng ý và (5) hoàn toàn đồng ý.

Để việc định nghĩa biến trên SPSS nhanh và thuận tiện, có thể tiến hành như sau:

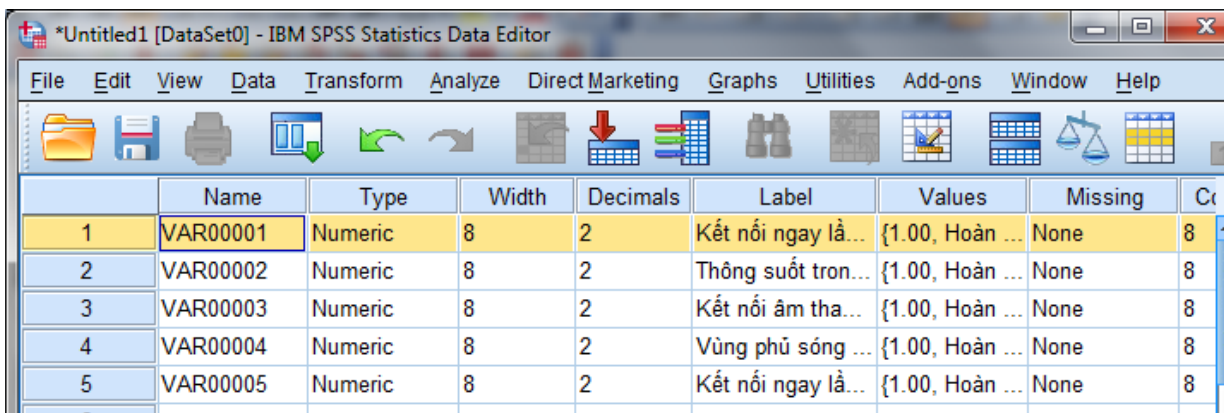
- Trước hết, kích vào dưới ô **Value**, định nghĩa các giá trị của các biến như sau:



- Tại thang đo **Measure**, chọn Scale (vì đây là thang đo khoảng kiểu thứ tự)



- Kích vào số thứ tự ở bên trái màn hình, chép (copy) biến vừa định nghĩa thành n biến (tương ứng với n chỉ báo), sau đó, chép nội dung chỉ báo vào ô Label, lần lượt thay đổi các tên biến vào ô **Name** sẽ hoàn thành việc định nghĩa các biến gồm nhiều chỉ báo.



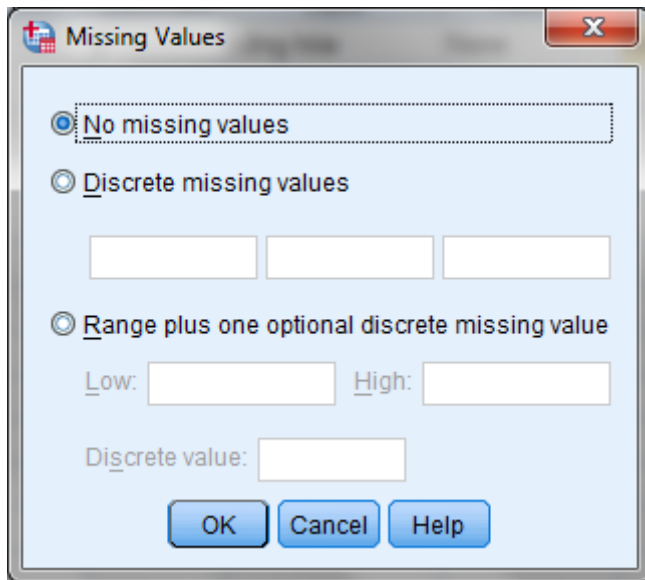
Nhập giá trị khuyết (missing)

Trong quá trình phỏng vấn, có những câu hỏi mà đúng ra được được phỏng vấn phải trả lời câu hỏi đó, tuy nhiên, do một số nguyên nhân, người được phỏng vấn bỏ qua một hoặc vài câu hỏi (hoặc câu trả lời) gọi là giá trị khuyết (missing).

Để đảm bảo thông tin trong quá trình phân tích, chúng ta cần phải định nghĩa những giá trị này như sau:

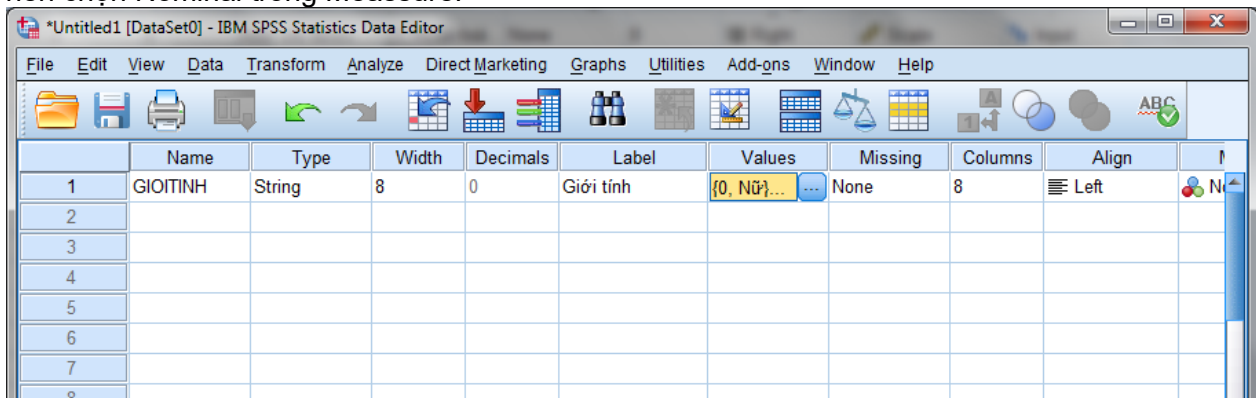
- Nhấn **Missing** - Hộp hội thoại **Missing Values** xuất hiện.

- Nhấn **Discrete missing values**, đặt các trị missing values vào các ô trống, trị được nhập tại các ô trống sẽ đại diện cho những giá trị khuyết (chú ý: trị của giá trị khuyết phải không trùng với các giá trị được dùng để mã hóa biến để tránh sự nhầm lẫn trong quá trình nhập liệu).
- Chúng ta có thể định nghĩa các giá trị khuyết theo một khoảng giá trị nào đó bằng các nhấn và nhập liệu vào **Range plus one optional discrete missing value**.
- Tất cả các giá trị khuyết sẽ không tham gia vào quá trình phân tích.



Chọn cấp độ thang đo tương ứng (Measure)

Chọn các loại thang đo tương ứng trong SPSS gồm **Nominal** (biểu danh), **Ordinal** (thứ tự) và **Scale** (khoảng cách và tỉ lệ). Ví dụ: Cấp độ thang đo của biến GIOITINH là thang đo biểu danh nên chọn Nominal trong Measure.



Chèn một biến mới hoặc bản ghi mới

- Chèn thêm một biến, nhấn **Edit/ Insert Variable**;
- Chèn thêm một bản ghi, nhấn **Data/ Insert Cases**;
- Tìm đến bản ghi cần thiết: **Go to Case**.

Kiểm tra giá trị nhập

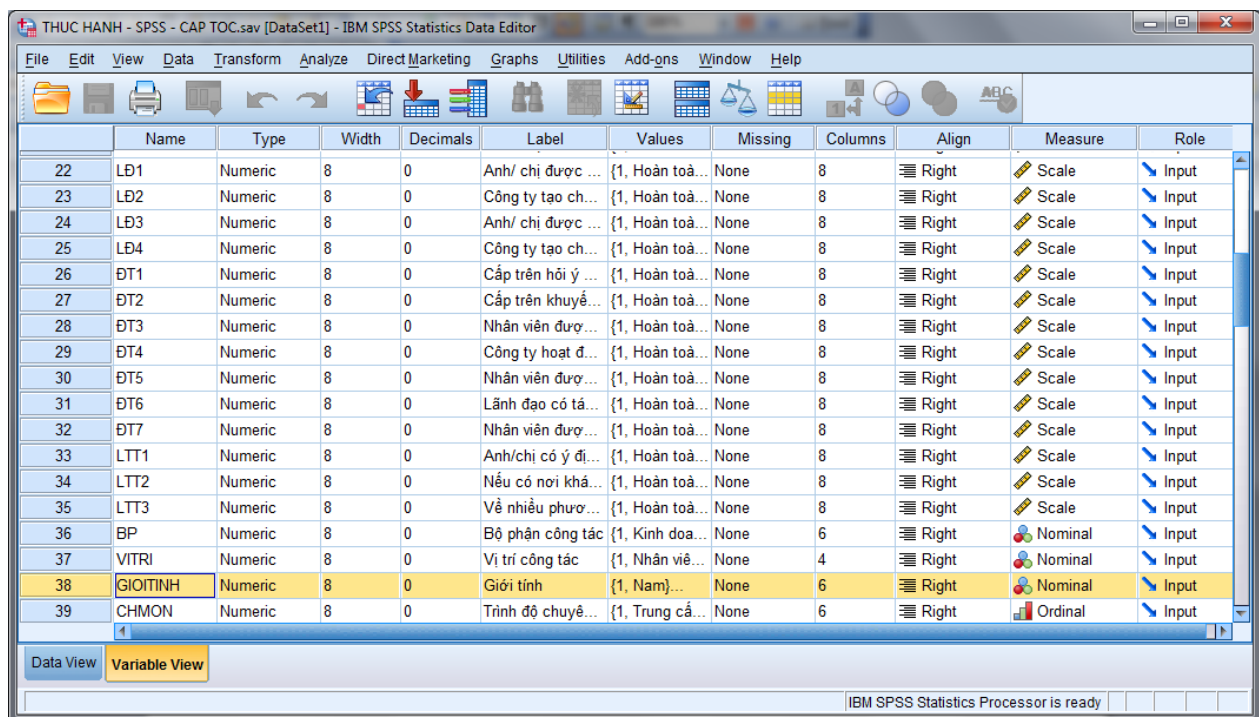
- Nhấn toàn bộ giá trị: **Nhấn View/ Value Labels**

- Kiểm tra một biến nào đó: **Utilities/ Variables**
- Kiểm tra bộ mã hoá **Utilities/ File Info**, với bộ mã hoá này, ta có thể kiểm tra lại một lần nữa công việc định nghĩa các biến với danh bạ mã hóa mà chúng ta thiết lập trước khi định nghĩa biến.

Nhập dữ liệu

Để bắt đầu nhập liệu, tại phía trái cuối màn hình, nhấn vào **Data View** và tiến hành nhập liệu các giá trị tương ứng, cần chú ý là tương ứng với số thứ tự trên thanh trượt phía trái màn hình là số thứ tự của bản câu hỏi nhập liệu tương ứng.

Sau khi định nghĩa các biến và được lưu trữ tại file THUC HANH SPSS có dạng như sau:



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
22	LB1	Numeric	8	0	Anh/ chỉ được ...	{1, Hoàn toà...	None	8	Right	Scale	Input
23	LB2	Numeric	8	0	Công ty tạo ch...	{1, Hoàn toà...	None	8	Right	Scale	Input
24	LB3	Numeric	8	0	Anh/ chỉ được ...	{1, Hoàn toà...	None	8	Right	Scale	Input
25	LB4	Numeric	8	0	Công ty tạo ch...	{1, Hoàn toà...	None	8	Right	Scale	Input
26	ĐT1	Numeric	8	0	Cấp trên hỏi ý ...	{1, Hoàn toà...	None	8	Right	Scale	Input
27	ĐT2	Numeric	8	0	Cấp trên khuỷ...	{1, Hoàn toà...	None	8	Right	Scale	Input
28	ĐT3	Numeric	8	0	Nhân viên đượ...	{1, Hoàn toà...	None	8	Right	Scale	Input
29	ĐT4	Numeric	8	0	Công ty hoạt đ...	{1, Hoàn toà...	None	8	Right	Scale	Input
30	ĐT5	Numeric	8	0	Nhân viên đượ...	{1, Hoàn toà...	None	8	Right	Scale	Input
31	ĐT6	Numeric	8	0	Lãnh đạo có tá...	{1, Hoàn toà...	None	8	Right	Scale	Input
32	ĐT7	Numeric	8	0	Nhân viên đượ...	{1, Hoàn toà...	None	8	Right	Scale	Input
33	LTT1	Numeric	8	0	Anh/chỉ có ý đi...	{1, Hoàn toà...	None	8	Right	Scale	Input
34	LTT2	Numeric	8	0	Nếu có nơi khá...	{1, Hoàn toà...	None	8	Right	Scale	Input
35	LTT3	Numeric	8	0	Về nhiều phươ...	{1, Hoàn toà...	None	8	Right	Scale	Input
36	BP	Numeric	8	0	Bộ phận công tác	{1, Kinh doa...	None	6	Right	Nominal	Input
37	VITRI	Numeric	8	0	Vị trí công tác	{1, Nhân việ...	None	4	Right	Nominal	Input
38	GIOTINH	Numeric	8	0	Giới tính	{1, Nam}...	None	6	Right	Nominal	Input
39	CHMON	Numeric	8	0	Trình độ chuyê...	{1, Trung cầ...	None	6	Right	Ordinal	Input

Như vậy, sau khi việc định nghĩa biến được hoàn tất, việc nhập liệu sẽ được bắt đầu. Như đã đề cập ở phần trên, trước khi tiến hành nhập liệu, cần phải sắp xếp, đánh số thứ tự toàn bộ các bản câu hỏi thu thập được và sử dụng cho quá trình phân tích, cần nhập liệu thứ tự của bản câu hỏi trùng với số thứ tự của bản ghi (cột bên trái màn hình) để thuận tiện cho việc đối chiếu sau này.

Để tiến hành nhập liệu, từ cửa sổ **Data View**, kích vào **Variable View** để mở cửa sổ nhập liệu. Theo bản câu hỏi được giới thiệu ở phần thiết kế công cụ thu thập dữ liệu, tiến hành điều tra tại hiện trường với 200 mẫu, mỗi bản câu hỏi được nhập liệu cho mỗi bản ghi trong SPSS có thứ tự từ 1 đến n như phía bên trái trang màn hình.

Sau khi định nghĩa biến và nhập liệu, kết quả được thể hiện như sau:

THUC HANH - SPSS - CAP TOC.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

1: LTT1

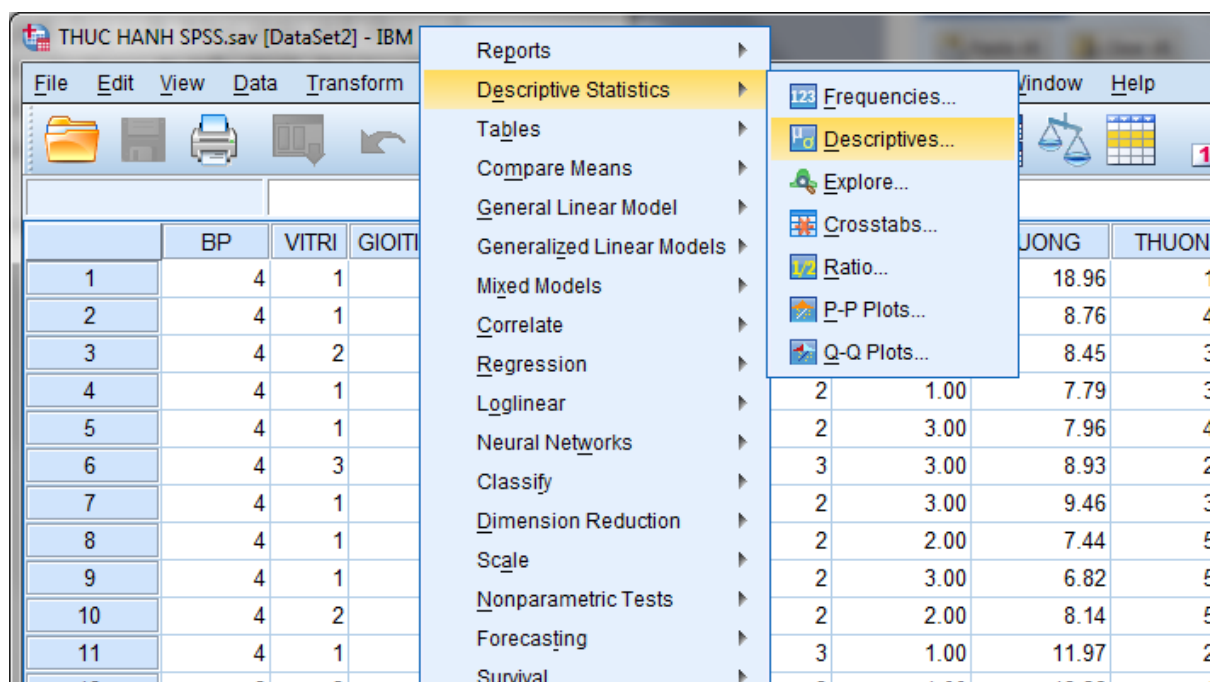
Visible: 67 of 67 Variables

	LTT1	LTT2	LTT3	BP	VITRI	GIOITINH	CHMON	TGLV	TUOI	DIADIEM	KECAN	DANHIEU	LUONG	THUONK
1	4	4	4	4	1	1	3	1	25	1.00	3	1.00	18.96	1
2	4	4	2	4	1	2	3	1	23	1.00	2	1.00	8.76	4
3	4	4	3	4	2	2	3	1	24	2.00	2	1.00	8.45	3
4	4	4	5	4	1	1	3	1	24	1.00	2	1.00	7.79	3
5	4	4	4	4	1	1	3	1	31	3.00	2	1.00	7.96	4
6	4	4	7	4	3	1	3	1	26	3.00	3	1.00	8.93	2
7	2	2	2	4	1	2	3	1	24	3.00	2	1.00	9.46	3
8	4	4	5	4	1	1	3	1	28	2.00	2	1.00	7.44	5
9	4	2	3	4	1	2	3	1	23	3.00	2	1.00	6.82	5
10	7	5	6	4	2	2	3	2	28	2.00	2	1.00	8.14	5
11	3	2	2	4	1	2	3	1	25	1.00	3	1.00	11.97	2
12	2	7	4	2	3	1	3	2	31	1.00	3	1.00	18.26	1
13	5	6	6	1	1	2	3	1	28	2.00	2	1.00	20.50	1
14	4	5	4	1	1	2	3	1	27	3.00	1	1.00	10.43	1
15	1	1	2	2	1	2	3	1	24	2.00	1	1.00	11.97	2
16	3	4	3	5	1	1	3	1	27	1.00	3	1.00	17.20	1
17	6	5	5	2	2	1	3	2	32	2.00	2	1.00	12.32	1

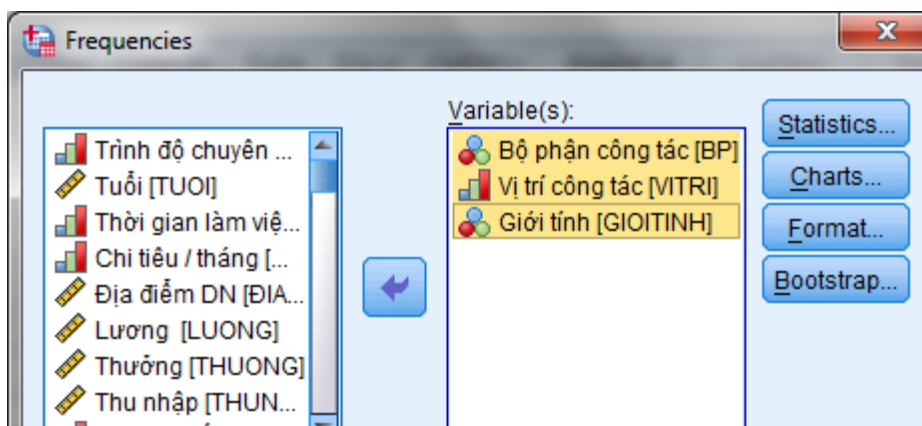
PHÂN TÍCH MÔ TẢ

Khi các dữ liệu đã được nhập, các nhà nghiên cứu có thể bắt đầu việc phân tích dữ liệu. Giả sử các nhà nghiên cứu quan tâm đến việc có được những thông số mô tả cho tất cả các biến được nhập vào dữ liệu thông qua bảng phân phối tần suất. Bảng phân phối tần suất được thể hiện với tất cả các biến định tính (rời rạc) với các cấp độ thang đo biểu danh, thứ tự và các biến định lượng (liên tục) với thang đo khoảng cách hoặc tỉ lệ.

- Vào **File / Open / Data** chọn thư mục chứa file cần phân tích để tiến hành mở dữ liệu đã nhập, giả sử ở đây là file **THUC HANH SPSS.sav**.
- Vào **Analyze / Descriptive Statistics / Frequencies**



- Cửa sổ **Frequencies** sẽ mở ra, chọn các biến cần phân tích vào ô **Variables**



- Kích vào **Statistics** để lựa chọn các chỉ tiêu cần mô tả.
- Kích vào **Charts** để vẽ các hình vẽ theo các cấu trúc cho sẵn **Pie charts / Bar charts / Histograms...**
- Kết quả và diễn giải:

Bộ phận công tác

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Kinh doanh	32	16.0	16.0	16.0
	Kỹ thuật	49	24.5	24.5	40.5
	Văn phòng	39	19.5	19.5	60.0
	Kế toán	51	25.5	25.5	85.5
	Khác	29	14.5	14.5	100.0
	Total	200	100.0	100.0	

Statistics

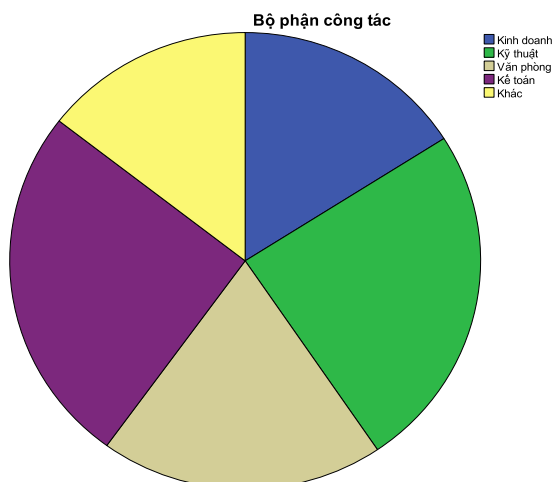
Tuổi

N	Valid	200
	Missing	0
Mean		26.47
Std. Deviation		3.161
Minimum		21
Maximum		42

Kết quả phân tích tần suất Frequencies khi thực hiện nghiên cứu trên 200 mẫu cho thấy có 32 người làm việc ở bộ phận kinh doanh (chiếm 16%), 49 người làm việc tại bộ phận kỹ thuật (chiếm 24.5%)...

Hai bộ phận có số lượng cán bộ cao nhất là bộ phận kế toán (51 người) và bộ phận kỹ thuật (49 người).

Độ tuổi trung bình là 26.47 tuổi với người có tuổi lớn nhất là 42 và nhỏ nhất là 21 tuổi...

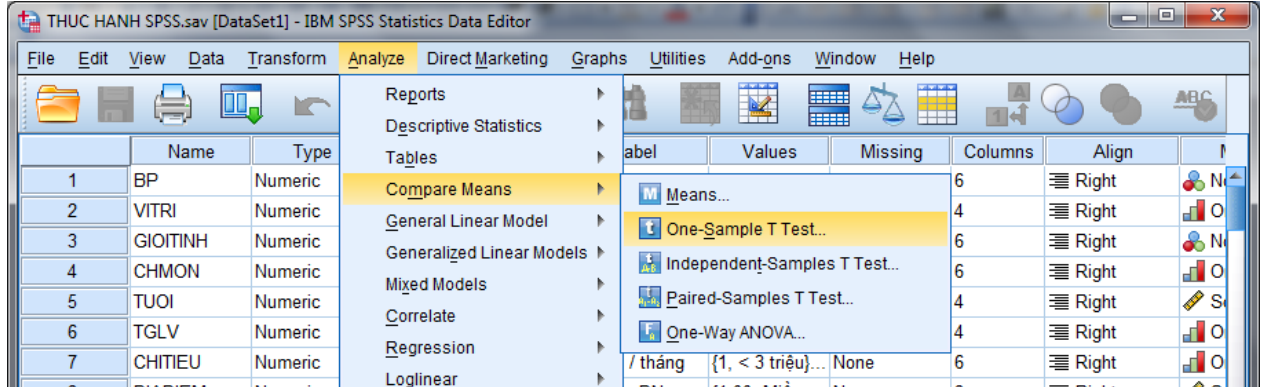


Phân tích một biến định lượng

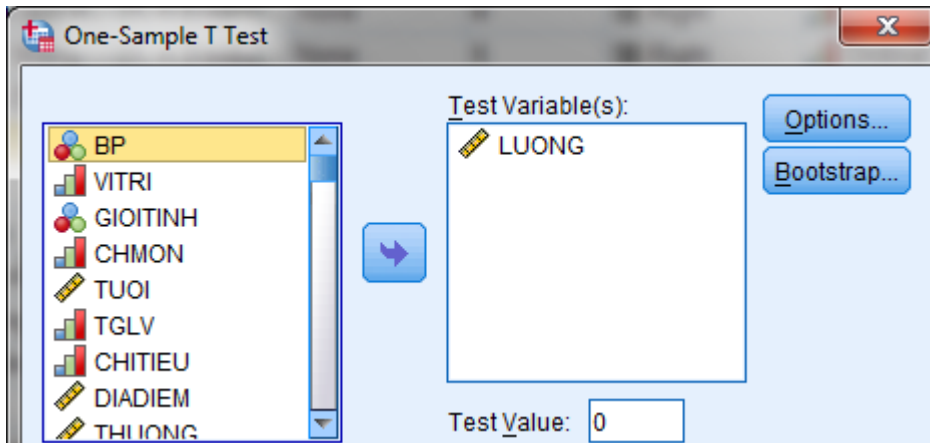
Ước lượng tham số trung bình

Giả sử người nghiên cứu muốn biết lương trung bình của cán bộ công nhân viên trong doanh nghiệp, quy trình thực hiện như sau:

- Trong thanh menu, vào **Analyze / Compare Mean / One sample T Test**.



- Khi đó, hộp hội thoại **One-Sample T Test** mở ra:



- Chọn biến lương vào ô **Test Variable** và **Test Value** mặc định bằng 0. Bấm **OK**, kết quả thu được như sau:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
LUONG	200	12.558	3.56823	.25231

Mẫu nghiên cứu gồm 200 cán bộ công nhân viên:

- + Tham số lương trung bình (Mean) = 12,5585 triệu đồng.
- + Độ lệch chuẩn (Std. Deviation) = 3.56823
- + Gọi μ là Trung bình tổng thể, với mức ý nghĩa $\alpha = 0.05$

One-Sample Test

	Test Value = 0					
	t	Df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
LUONG	49.774	199	.000	12.55848	12.0609	13.0560

Qua bảng One – Sample Test, ước lượng tham số lương trung bình của cán bộ công nhân

viên trong doanh nghiệp là $12.061 \leq \mu \leq 13.056$ triệu.

Kiểm định T đối với tham số trung bình của tổng thể

Thông thường, một bài toán kiểm định giả thiết bao gồm 5 bước, trong phần này sẽ giới thiệu tổng lược về 5 bước của bài toán kiểm định như sau:

B1: Phát biểu giả thiết và đối thiết trong trường hợp kiểm định đối xứng (2 đuôi):

Giả thuyết: $H_0: \mu = \mu_0$

Đối thuyết: $H_A: \mu \neq \mu_0$

B2: Xác định mức ý nghĩa α

B3: Xác định phương pháp kiểm định: Phương pháp kiểm định T-student đối với tham số trung bình tổng thể.

B 4: Tính tiêu chuẩn kiểm định: T-student.

B5: Kết luận:

- Nếu xác suất $P(T, df) \leq \alpha$ thì kết luận bác bỏ H_0 , chấp nhận H_1 .
- Nếu xác suất $P(T, df) > \alpha$ thì kết luận chưa có cơ sở để bác bỏ H_0 tức chưa có đủ cơ sở để chấp nhận H_A .

Ví dụ:

Như chúng ta đã biết, lương trung bình của các đối tượng phỏng vấn là 12,558 triệu/tháng, có nhận định cho rằng lương trung bình của cán bộ công nhân viên của doanh nghiệp khác 11 triệu/tháng, cần kết luận nhận định đó có đúng không.

Khi đó, giả thuyết của bài toán là:

$H_0: \mu = \mu_0 = 11$ (triệu) và $H_A: \mu \neq \mu_0 = 11$ (triệu)

- Trên thanh menu, kích **Analyze / Compare Means / One sample T Test**.
- Chọn biến cần phân tích vào ô **Test Variable(s)**, đặt giá trị μ_0 vào ô **Test Value**, khi đó **Test Value** = 11.
- Bấm **Continue** và bấm **OK**, kết quả thu được như sau:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
LUONG	200	12.5585	3.56823	.25231

One-Sample Test

	Test Value = 11					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
LUONG	6.177	199	.000	1.55848	1.0609	2.0560

Tại các biểu trên, lương trung bình của cán bộ công nhân viên là 12,558 triệu với độ lệch chuẩn là 3,558.

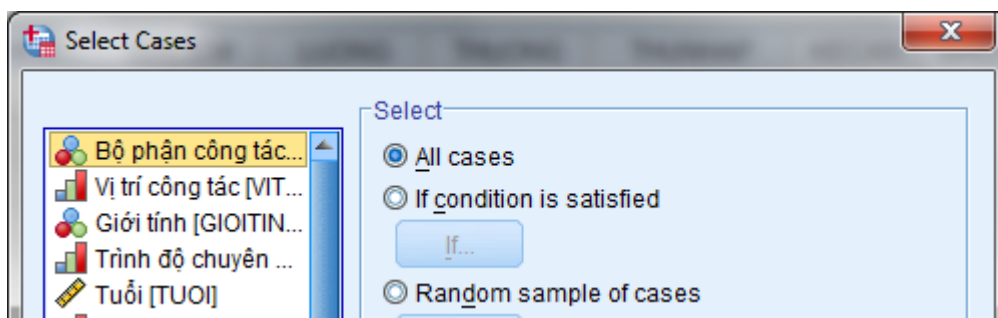
Vì giá trị t-student = 6.177 và df = 199, khi đó $P(t, df) = 0.000 < 0,05$ nên có thể bác bỏ H_0 và chấp nhận H_A nghĩa là lương trung bình của cán bộ công nhân viên của doanh nghiệp khác 11 triệu/tháng.

MỘT SỐ XỬ LÝ THƯỜNG GẶP SAU KHI NHẬP LIỆU

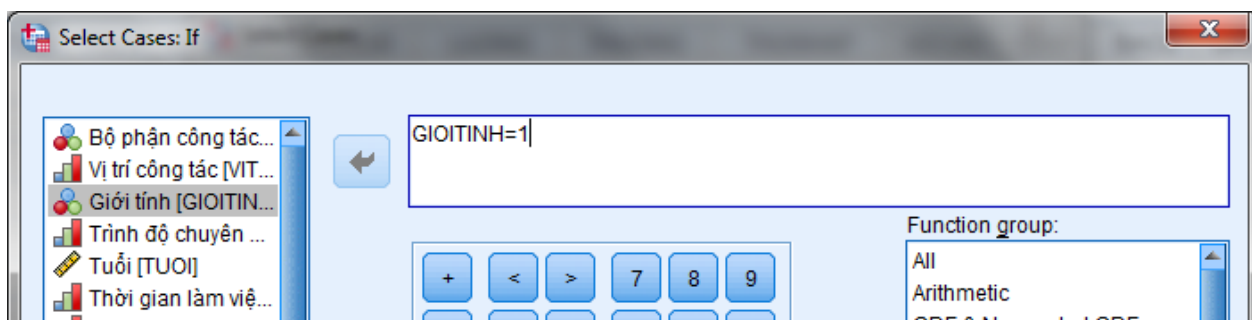
Chọn dữ liệu theo điều kiện

Sau khi nhập liệu và tiến hành phân tích, tùy theo mục tiêu mà nghiên cứu có thể thực hiện trên toàn bộ dữ liệu hoặc lựa chọn một số dữ liệu thỏa mãn những điều kiện nhất định. Ví dụ, nhà nghiên cứu muốn xem xét thông tin những người nam, khi đó người ta chỉ lựa chọn những người có giới tính là nam (GIOITINH=1).

- Kích vào **Data / Select Cases**, cửa sổ **Select Cases** sẽ xuất hiện:



- Chọn **If condition is satisfied** và nhấn vào **If**, trong cửa sổ **Select Cases: If** gán biến GIOITINH=1 (giới tính Nam).



- Kích **Continue / OK**.
- Tại cửa sổ **Data View** sẽ xuất hiện một số bản ghi có gạch chéo, những bản ghi này là những bản ghi của những người có giới tính là Nam.

	BP	VITRI	GIOITINH	CHMON	TUOI	TGLV	MUCCHITI EU	DIADIEM	LUONG	THUONG	THUNHAP	KECAM
1	4	1	1	3	25	1	3	1.00	18.96	1.48	20.44	
2	4	1	2	3	23	1	2	1.00	8.76	4.26	13.02	
3	4	2	2	3	24	1	2	2.00	8.45	3.79	12.23	
4	4	1	1	3	24	1	2	1.00	7.79	3.79	11.57	
5	4	1	1	3	31	1	2	3.00	7.96	4.95	12.91	
6	4	3	1	3	26	1	3	3.00	8.93	2.15	11.08	
7	4	1	2	3	24	1	2	3.00	9.46	3.79	13.25	
8	1	1	1	3	26	1	1	3.00	10.43	1.15	11.58	
9	1	1	1	3	28	1	2	3.00	11.97	2.21	14.18	
10	1	1	2	2	32	1	1	3.00	17.20	1.29	18.50	

- Chú ý rằng, những bản ghi không thỏa mãn điều kiện (có dấu gạch chéo) sẽ không tham gia vào các phân tích. Giả sử thực hiện lại phân tích mô tả với các biến Bộ phận (BP) và Vị trí (VITRI) như trên, kết quả mô tả của những người có giới tính là nam như sau:

		BP			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Kinh doanh	22	21.2	21.2	21.2
	Kỹ thuật	36	34.6	34.6	55.8
	Văn phòng	14	13.5	13.5	69.2
	Kế toán	20	19.2	19.2	88.5
	Khác	12	11.5	11.5	100.0
	Total	104	100.0	100.0	

Khi đó, số lượng nam nhân là 104 người, trong đó nam nhân viên kinh doanh là 22 người, nam nhân viên kỹ thuật là 36 người...

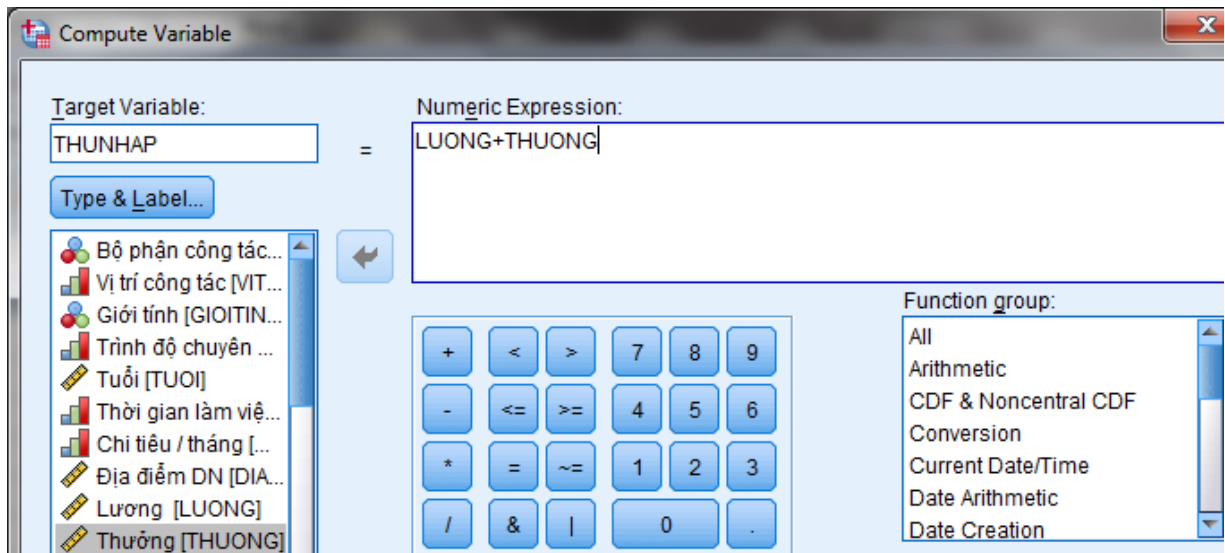
		VITRI			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Nhân viên ở các bộ phận	78	75.0	75.0	75.0
	Cán bộ phụ trách ở các bộ phận	14	13.5	13.5	88.5
	Trưởng phó phòng ban hoặc tương đương	12	11.5	11.5	100.0
	Total	104	100.0	100.0	

- Sau khi tiến hành phân tích theo điều kiện đã lựa chọn, để trả về giá trị cũ, vào **Data / Select Cases**.
- Tại cửa sổ **Select Cases**, chọn **All Cases / OK**.

Tạo biến mới không hoặc có ràng buộc một điều kiện

Trong quá trình nhập liệu, để có thể rút ngắn thời gian nhập liệu hoặc để phục vụ mục đích phân tích, chúng ta còn có thể tạo ra biến mới từ các dữ kiện và cấu trúc của biến đã nhập hay tạo ra một biến mới từ một hoặc nhiều biến đã có.

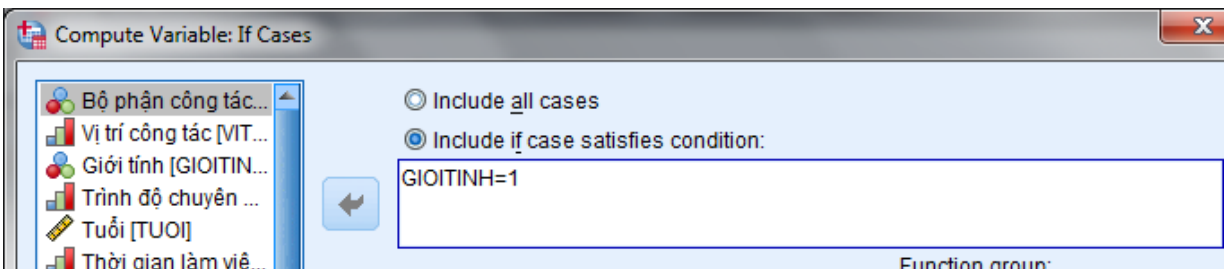
- Tạo biến mới không điều kiện: Giả sử theo số liệu nghiên cứu như trên, có hai biến lương và thưởng, giả sử thu nhập của một người sẽ bằng lương + thưởng. Như vậy, có thể tạo một biến mới (THUNHAP) sẽ bằng tổng của lương (LUONG) và thưởng (THUONG)
 - + Nhấn **Transform/Compute Variable**
 - + Trong ô **Target Variable** nhập biến mới (THUNHAP).
 - + Nên định nghĩa **Type & Label** để tiện cho việc quản lí và so sánh các giá trị sau này.
 - + Trong ô **Numeric Expression** nhập giá trị cần gán cho biến mới từ biến đích cho trước (LUONG+THUONG).



+ Kích **OK**, tại file dữ liệu sẽ xuất hiện một biến mới THUNHAP có giá trị bằng tổng hai biến lương và thưởng.

Chú ý: Khi gặp các biến thuộc kiểu chuỗi, ngày tháng... chúng ta cần phải tìm một hàm tương ứng để quy các giá trị này về giá trị tương đồng mà chúng ta có thể so sánh được (sử dụng hàm Function)

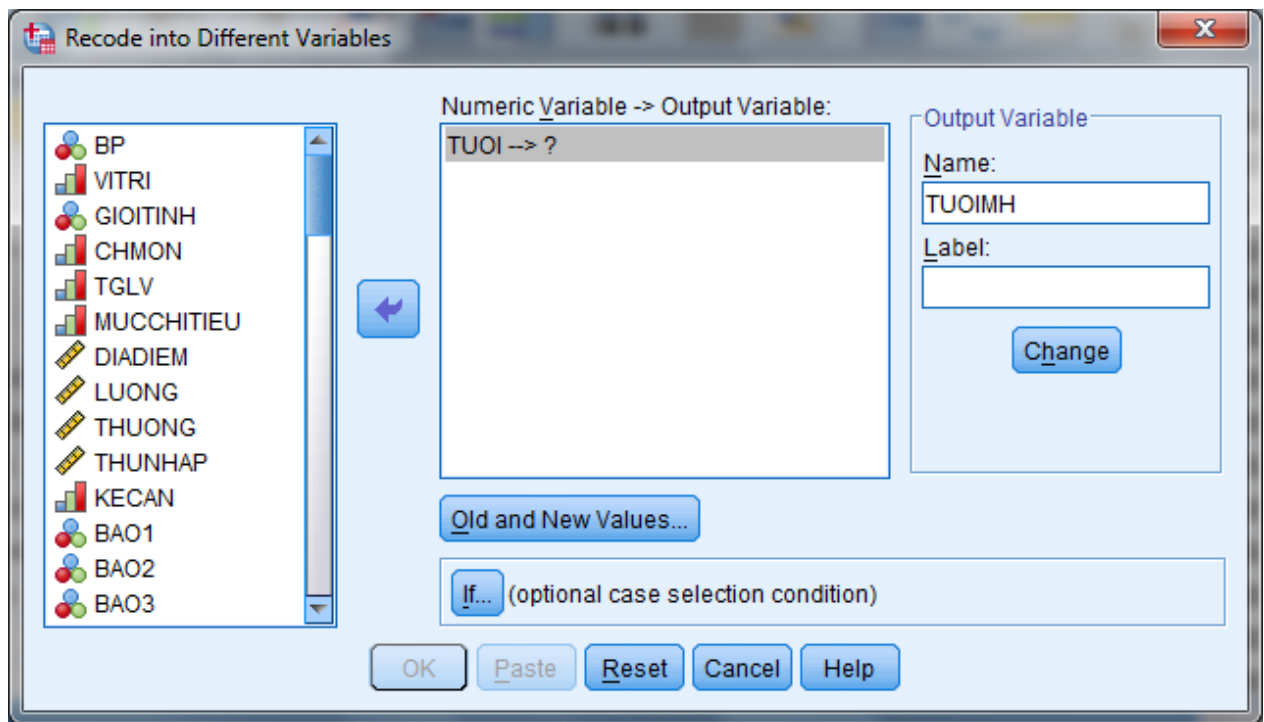
- Tạo biến mới có điều kiện: Nội dung cũng như ví dụ trên nhưng nhà nghiên cứu cần phân chia dữ liệu ra thành nam và nữ. Sau khi thiết đặt các giá trị như trên xong, nhấn **If** (ở phía dưới bên phải hộp hội thoại), tiếp theo nhấn **Include if case satisfies condition** trong hộp hội thoại để thiết đặt điều kiện (áp dụng cho những người có giới tính là nam thì điều kiện thiết đặt là GIOITINH=1 như trong hộp hội thoại).
- Sau khi thiết đặt xong, kích **Continue / OK** để tạo biến mới thỏa mãn điều kiện như thiết đặt.



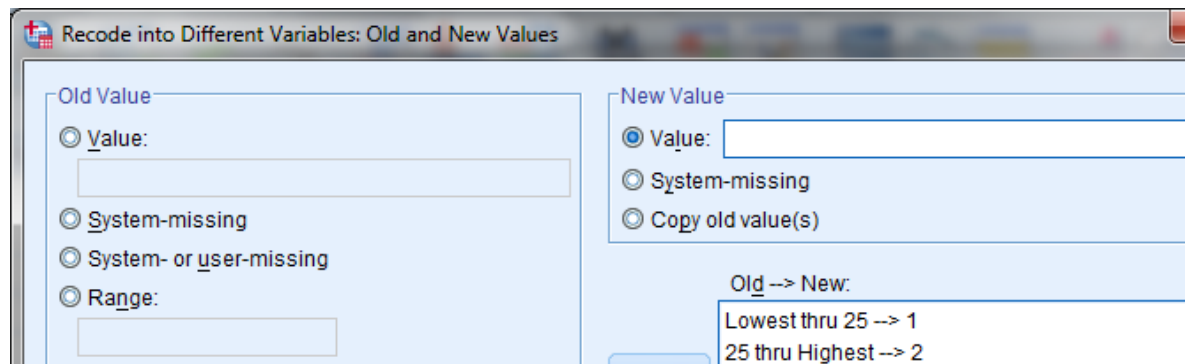
Mã hoá lại biến

Trong một số trường hợp, do nhu cầu của quá trình phân tích, chúng ta cần phải mã hóa lại các biến. Có hai hình thức mã hoá là mã hóa dùng lại tên biến cũ (**Recode Into Same Variables**) và mã hóa lưu trên một biến mới (**Recode Into Different Variables**). Để đảm bảo cho việc không bị mất dữ liệu gốc nhằm phục vụ cho mục đích khác, nhà nghiên cứu nên mã hóa lại và lưu trên một biến mới.

- Nhấn **Transform / Recode Into Different Variables**,
- Tên biến cũ đặt ở ô **Input Variable**, tên biến mới được đặt ở ô **Output Variable** với các thông số thỏa mãn một biến bình thường.



- Nhấn cửa biến được thiết đặt tại ô **Label**.
- Nhấn **If** để thiết đặt các điều kiện (nếu có).
- Nhấn **Old and New Values** để thay đổi bộ mã hoá.
 - Trong ô **Old Value** là giá trị cũ và **New Value** là giá trị mới cần nhập (sau mỗi lần định nghĩa giá trị cũ thành giá trị mới cần phải kích Add để nhận giá trị).
 - Nếu nhập giá trị mới ở thang điểm biểu danh, khoảng cách, tỷ lệ thì nhập tại ô **Value**.
 - Nếu mã hoá giá trị với thang điểm khoảng cách - Nhấn **Range** để định nghĩa các giá trị.

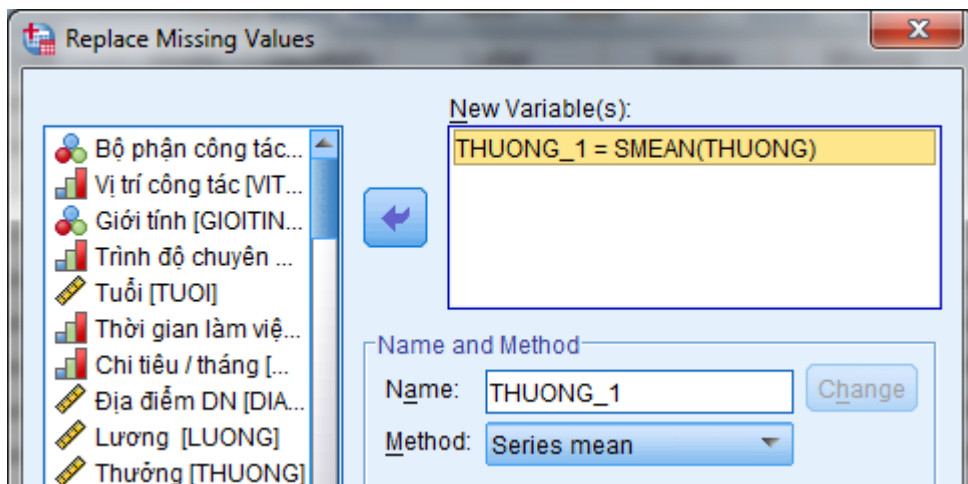


- Sau đó nhấn **Change** để thực hiện việc mã hóa lại biến.

Thay thế giá trị lỗi (missing)

Như đã đề cập ở phần trên, trong quá trình thu thập dữ liệu, do nhiều nguyên nhân khác nhau nên vẫn còn tồn tại một số giá trị lỗi (missing), để đảm bảo điều kiện của một số phân tích, các giá trị lỗi của biến định lượng (thang đo khoảng cách hoặc tỷ lệ) có thể thay thế bằng giá trị trung bình của tất cả ý kiến trả lời, giá trị trung bình của bản ghi gần với bản ghi chứa giá trị lỗi...

Vào **Transform / Replace Missing Values**, hộp thoại Replace Missing Values xuất hiện, chọn biến có chứa giá trị lỗi cần thay thế (ví dụ biến THUONG) vào ô **New Variable(s)**, chọn phương pháp thay thế (giả sử như **series mean**) trong ô **Method**, kích **OK**. Khi đó, trong cơ sở dữ liệu sẽ xuất hiện iến mới có ký hiệu mở rộng từ biến cũ (THUONG_1) ở cuối cùng của dãy biến.



ĐỊNH NGHĨA BIẾN CHO CÂU HỎI NHIỀU LỰA CHỌN (MULTIPLE RESPONSE)

Ví dụ về câu hỏi nhiều lựa chọn

Câu hỏi cho nhiều lựa chọn là loại câu hỏi cho phép người trả lời chọn được nhiều phương án trong các phương án được liệt kê. Ví dụ, một nhà nghiên cứu mong muốn điều tra tên của tất cả các tờ báo đọc trong tuần qua, hoặc vòng tròn tất cả các tờ báo đọc trong tuần trước từ một danh sách các tờ báo.

Ví dụ:

Hàng tuần, Anh (Chị) đọc loại báo nào dưới đây (có thể chọn nhiều loại báo):

- ☐ Tiền phong
- ☐ Thanh niên
- ☐ Tuổi trẻ
- ☐ Hà Nội mới
- ☐ Lao động

Có hai cách định nghĩa cho loại biến gồm phương pháp **Multiple-Dichotomy** và phương pháp **Categories**, cả hai phương pháp này đều cho ra một kết quả như nhau. Quyển sách này sẽ trình bày phương pháp **Multiple - Dichotomy**, còn phương pháp categories được thực hiện tương tự, tuy nhiên, phần nhập liệu chỉ nhập tương ứng số tại mỗi biến, các biến còn lại không chọn sẽ không nhập dữ liệu.

Phương pháp Multiple-Dichotomy (Multiple-Dichotomy method)

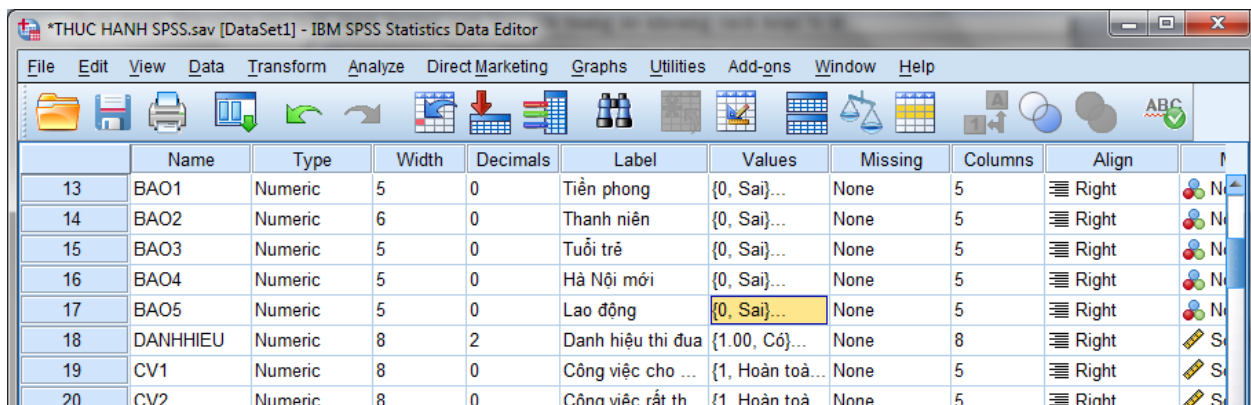
Phương pháp này dựa trên cơ sở giả định với mỗi phương án trả lời là một câu hỏi phân đôi với câu trả lời là có / không, đúng / sai, đồng ý / không đồng ý... cho mỗi phương án. Như vậy, câu hỏi trên bao gồm nhiều câu hỏi có hai sự lựa chọn (gọi là câu hỏi dạng dichotomies. Câu hỏi trên có thể được hiểu:

Trong tuần vừa qua, Anh (Chị) đọc loại báo nào dưới đây:

Tiền phong
Thanh niên
Tuổi trẻ
Hà Nội mới
Lao động

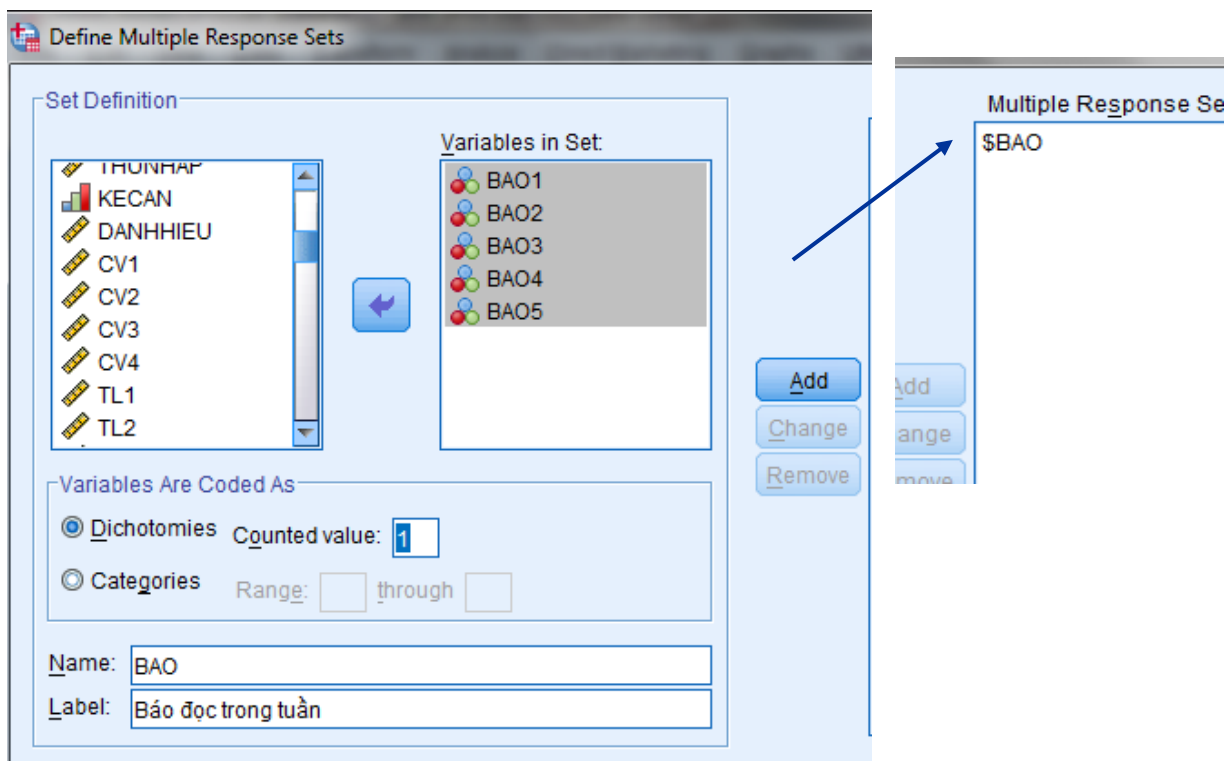
Đúng	Sai
☐	☐
☐	☐
☐	☐
☐	☐
☐	☐

- Trước tiên, cần định nghĩa 5 biến BAO1, 2, 3, 4, 5 tương ứng với 5 loại báo nhận giá trị với 1 là đúng và 0 là sai và nhập liệu các biến.



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	
13	BAO1	Numeric	5	0	Tiền phong	{0, Sai}...	None	5	Right	
14	BAO2	Numeric	6	0	Thanh niên	{0, Sai}...	None	5	Right	
15	BAO3	Numeric	5	0	Tuổi trẻ	{0, Sai}...	None	5	Right	
16	BAO4	Numeric	5	0	Hà Nội mới	{0, Sai}...	None	5	Right	
17	BAO5	Numeric	5	0	Lao động	{0, Sai}...	None	5	Right	
18	DANHIEU	Numeric	8	2	Danh hiệu thi đua	{1.00, Có}...	None	8	Right	
19	CV1	Numeric	8	0	Công việc cho ...	{1, Hoàn toả...	None	5	Right	
20	CV2	Numeric	8	0	Công việc rất th...	{1, Hoàn toả...	None	5	Right	

- Định nghĩa biến nhiều lựa chọn (Biến BAO) trên cơ sở các biến BAO1, 2, 3, 4, 5: Vào **Analyze / Multiple Response / Define Variables Sets** (hoặc có thể là **Analyze / Table / Multiple Response Sets**).



Define Multiple Response Sets

Set Definition

Variables in Set:

- BAO1
- BAO2
- BAO3
- BAO4
- BAO5

Variables Are Coded As:

☒ Dichotomies Counted value: 1

☐ Categories Range: through

Name: BAO

Label: Báo đọc trong tuần

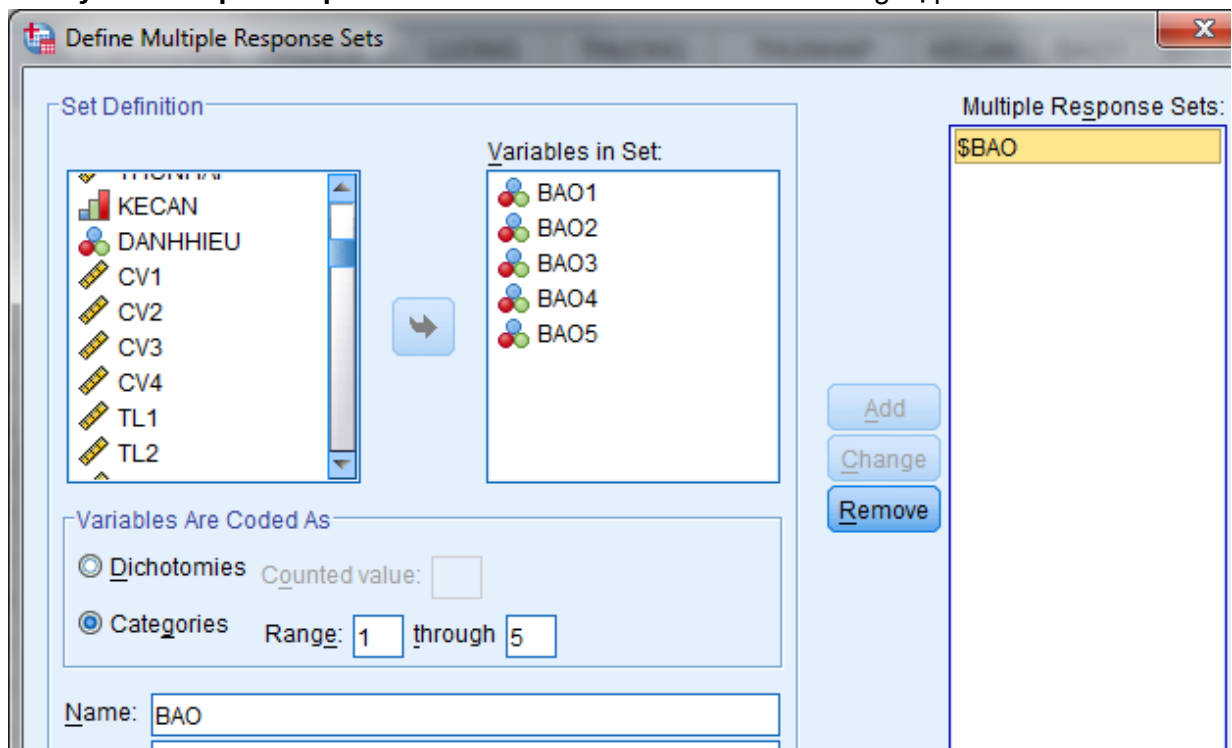
Buttons: Add, Change, Remove

- Tại cửa sổ **Define Multiple Response Sets**, chọn các biến BAO1, 2, 3, 4, 5 sang ô **Variables in Set**, đặt biến BAO tại ô **Name** với **Label** là "Báo đọc trong tuần".

- Trong ô **Variable Are Coded As**, chọn **Dichotomies** với **Counted Value = 1** (chú ý giá trị này tương ứng với giá trị "đúng" trong định nghĩa biến BAO1, 2, 3, 4, 5).
- Nhấn **Add**, khi đó tại ô **Multiple Response Sets** sẽ xuất hiện biến \$BAO, biến này sẽ chứa đựng các thông tin của năm biến BAO1, 2, 3, 4, 5.
- Có thể kích **Change** để thay đổi các biến hoặc **Remove** để xóa các biến đã định nghĩa.
- Nhấn **Close** để đóng cửa sổ định nghĩa biến.

Phương pháp Categories

- Thực hiện định nghĩa tương tự như phần **Multiple-Dichotomy (Multiple-Dichotomy method)**, tuy nhiên, ở phần nhập liệu các biến BAO1, 2, 3, 4, 5 đều định nghĩa các giá trị 1, 2, 3, 4, 5 tương ứng với sự lựa chọn như sau:
Hàng tuần, Anh (Chị) đọc loại báo nào dưới đây (có thể chọn nhiều loại báo):
 - 1 Tiền phong
 - 2 Thanh niên
 - 3 Tuổi trẻ
 - 4 Hà Nội mới
 - 5 Lao động
- Chú ý: Khi nhập liệu, tại vị trí BAO(i), nếu đáp viên lựa chọn loại báo i để đọc thì phải nhập liệu giá trị i tương ứng. Ví dụ, một người lựa chọn loại báo Thanh niên và Lao động thì tại BAO2 nhập số 2 và tại BAO5 nhập số 5.
- Định nghĩa biến nhiều lựa chọn (Biến BAO) trên cơ sở các biến BAO1, 2, 3, 4, 5: Vào **Analyze / Multiple Response / Define Variables Sets** như trường hợp trên.



- Kích vào **Categories** và đặt giá trị **Range** từ 1 đến 5, thực hiện tương tự như trường hợp trên sẽ cho kết quả tương tự.

Tính tần suất đối với câu hỏi cho nhiều lựa chọn

- Vào **Analyze / Multiple Response / Frequencies**
- Chọn biến \$BAO từ ô **Multiple Response Sets** sang ô **Table(s) for**, kích **OK**.

Case Summary						
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
\$BAO ^a	200	100.0%	0	.0%	200	100.0%

a. Dichotomy group tabulated at value 1.

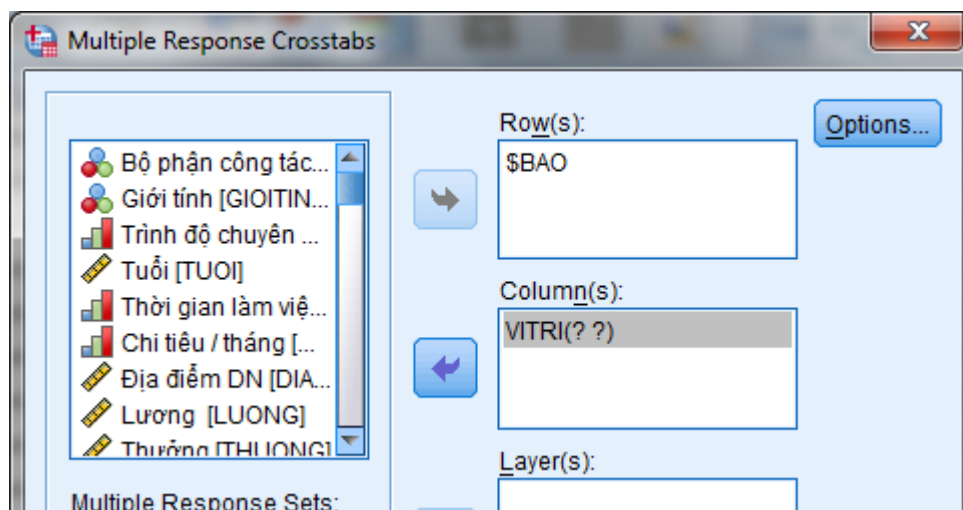
\$BAO Frequencies				
		Responses		Percent of Cases
		N	Percent	
\$BAO ^a	Tiền phong	130	30.2%	65.0%
	Thanh niên	100	23.3%	50.0%
	Tuổi trẻ	70	16.3%	35.0%
	Hà Nội mới	70	16.3%	35.0%
	Lao động	60	14.0%	30.0%
Total		430	100.0%	215.0%

a. Dichotomy group tabulated at value 1.

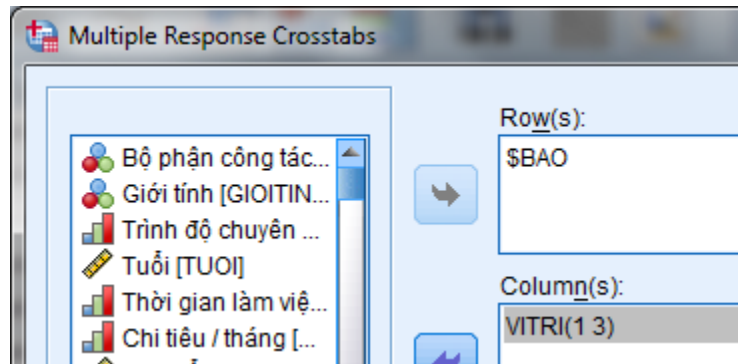
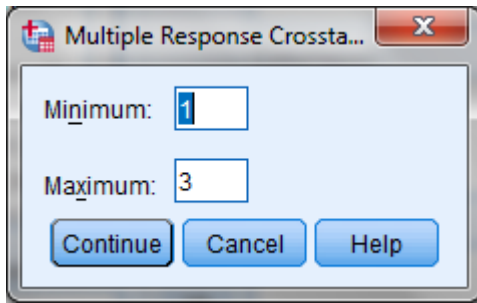
- Kết quả và diễn giải dữ liệu: Với mẫu nghiên cứu n = 200, có tổng số 430 lượt báo được đọc, trong đó có 130 lượt đọc báo Tiền phong (tỷ lệ 30.2%), 100 lượt đọc báo Thanh niên (23.3%)...

Lập bảng chéo (Crosstabs) với các biến nhiều sự lựa chọn

- Vào **Analyze / Multiple Response / Crosstabs**
- Chọn biến \$BAO từ ô **Multiple Response Sets** sang ô Dòng (Rows) hoặc Cột (Column).
- Chọn biến cần mô tả phối hợp với biến \$BAO (ví dụ biến VITRI) vào ô Dòng, Cột còn lại.



- Kích vào **Define Ranges** để định nghĩa giá trị Minimum và Maximum của biến VITRI (minimum = 1 và maximum = 3) và **Continue**.



- Kích **OK** để hiển thị kết quả.
- Kết quả và bình luận: Đa phần nhân viên ở các bộ phận là những người thường đọc báo, trong đó có 101 lượt người đọc báo Tiền phong, 80 người đọc báo Thanh niên; tương tự như vậy là các Trưởng phó phòng ban hoặc tương đương...

S\$BAO*VITRI Crosstabulation						
			Vị trí công tác			Total
			Nhân viên ở các bộ phận	Cán bộ quản lí ở các bộ phận	Trưởng phó phòng ban hoặc tương đương	
\$BAO ^a	Tiền phong	Count	101	17	12	130
	Thanh niên	Count	80	9	11	100
	Tuổi trẻ	Count	54	11	5	70
	Hà Nội mới	Count	57	10	3	70
	Lao động	Count	44	11	5	60
Total		Count	159	24	17	200

Percentages and totals are based on respondents.

a. Dichotomy group tabulated at value 1.

SƠ ĐỒ LỰA CHỌN KỸ THUẬT PHÂN TÍCH

Trong quá trình phân tích, người thực hiện phải cân nhắc việc lựa chọn kỹ thuật phân tích nào trong từng trường hợp, việc lựa chọn kỹ thuật phân tích phụ thuộc vào mục tiêu của thiết kế nghiên cứu, loại cấp độ thang đo được sử dụng để thiết kế các câu hỏi trong công cụ đo lường, độ lớn của biến, mối quan hệ độc lập hay phụ thuộc giữa các biến, phân tích tham số hay phi tham số.

Người ta thường chia phân tích mối quan hệ giữa các biến số thành ba loại là (1) phân tích đơn biến (Univariate Analysis), (2) phân tích hai biến (Bivariate Analysis) và (3) phân tích đa biến (Multivariate Analysis)

Phân tích đơn biến (Univariate Analysis)

Phân tích đơn biến thường rất đơn giản, thông thường đó là phân tích giá trị lớn nhất, nhỏ nhất; đo lường khuynh hướng hội tụ như giá trị trung bình - mean, median, mode, tổng các giá trị...; đo lường độ phân tán như độ lệch chuẩn (std. deviation), phương sai hoặc xem xét phân phối chuẩn thông qua các hệ số Skewness và Kurtosis, ước lượng và kiểm định tham số trung bình mẫu.

Phân tích hai biến (Bivariate Analysis)

Phân tích hai biến được xem xét trên cơ sở phân tích tham số hay phi tham số gồm kiểm định:

- Kiểm định tham số:
 - *Mối quan hệ giữa một biến định tính (thang đo biểu danh – nominal hoặc thứ tự ordinal) và một biến định lượng (thang đo khoảng cách – interval hoặc tỷ lệ - ratio):*
 - + Kiểm định sự khác biệt tham số trung bình của hai biến độc lập (T-Tets for Independent Groups).
 - + Kiểm định sự khác biệt tham số trung bình của hai biến phụ thuộc (Paired-Samples T-Tetst).
 - + Phân tích phương sai một chiều (One-Way ANOVA).
 - *Mối quan hệ giữa hai biến định lượng (thang đo khoảng cách – interval hoặc tỷ lệ - ratio):*
 - + Phân tích tích quan (Correlation)
 - + Hồi quy tuyến tính đơn (Linear Regression)
- Kiểm định phi tham số:
 - *Kiểm định Chi bình phương (Chi-Square) sự độc lập giữa hai biến,*
 - *Kiểm định Mann – Whitney U cho hai mẫu độc lập,*
 - *Kiểm định dấu và hạng Wilcoxon cho hai mẫu phụ thuộc.*

Phân tích đa biến (Multivariate Data Analysis)

- Phân tích các thành phần chính và phân tích nhân tố khám phá (Principal components and explore factor analysis)
- Tương quan bội và hồi quy tuyến tính bội (Multiple correlation and Multiple regression)
- Phân tích tách biệt và hồi quy logistic (Multiple discriminant analysis and logistic regression)
- Phân tích tương quan chuẩn tắc (Canonical correlation analysis)
- Phân tích phương sai đa biến (Multivariate analysis of variance)
- Phân tích liên đồng (Conjoint analysis)
- Phân tích cụm (Cluster analysis)
- Biểu đồ nhận thức (Perceptual mapping)
- Phân tích tương hợp (Correspondence analysis)
- Phương trình mô hình cấu trúc và phân tích nhân tố khẳng định

Việc lựa chọn một số phân tích trên cơ sở xem xét các biến định tính và định lượng như sau:

Phân tích phương sai đa biến (Multiple Analysis of Variance)

$$\begin{array}{ccc} Y_1 + Y_2 + Y_3 + \dots + Y_n & = & X_1 + X_2 + X_3 + \dots + X_n \\ \text{(Định lượng)} & & \text{(Định tính)} \end{array}$$

Phân tích sự tách biệt (Discriminant Analysis)

$$\begin{array}{ccc} Y & = & X_1 + X_2 + X_3 + \dots + X_n \\ \text{(Định tính)} & & \text{(Định lượng)} \end{array}$$

Hồi quy tương quan bội (Multiple Regression Analysis)

$$\begin{array}{ccc} Y & = & X_1 + X_2 + X_3 + \dots + X_n \\ \text{(Định lượng)} & & \text{(Định lượng, định tính)} \end{array}$$

Phân tích liên đồng

$$\begin{array}{ccc} Y & = & X_1 + X_2 + X_3 + \dots + X_n \\ \text{(Định tính, định lượng)} & & \text{(Định tính)} \end{array}$$

Phân tích tương quan chuẩn tắc (Canonical Correlation)

$$\begin{array}{ccc} Y_1 + Y_2 + Y_3 + \dots + Y_n & = & X_1 + X_2 + X_3 + \dots + X_n \\ \text{(Định tính, định lượng)} & & \text{(Định tính, định lượng)} \end{array}$$

Phương trình cấu trúc tuyến tính (Structural Equation Modeling)

$$\begin{array}{ccc} Y_1 & = & X_{11} + X_{12} + X_{13} + \dots + X_{1n} \\ Y_2 & = & X_{21} + X_{22} + X_{23} + \dots + X_{2n} \\ \dots & & \dots \dots \dots \\ Y_k & = & X_{k1} + X_{k2} + X_{k3} + \dots + X_{kn} \\ \text{(Định lượng)} & & \text{(Định lượng, định tính)} \end{array}$$

CHƯƠNG

8.1

KIỂM ĐỊNH KHÁC BIỆT GIỮA HAI THAM SỐ TRUNG BÌNH: KIỂM ĐỊNH T (T-TEST) VÀ PHÂN TÍCH PHƯƠNG SAI MỘT CHIỀU (ONE-WAY ANALYSIS OF VARIANCE)

Trong chương trước chúng ta đã giới thiệu sơ lược về phần mềm SPSS và ứng dụng trong phân tích mô tả mẫu nghiên cứu, trong chương này sẽ thảo luận chi tiết những kỹ thuật liên quan đến việc kiểm định T (T-test) và phân tích phương sai một chiều (One-way Analysis of Variance).

SƠ LƯỢC VỀ QUY TRÌNH KIỂM ĐỊNH

Một bài toán kiểm định thường có 5 bước gồm:

- Bước 1. Phát biểu giả thuyết (Null hypothesis – H_0) và đối thuyết (Alternative hypothesis – H_A): Tùy theo mục tiêu và dạng quy luật phân phối, người ta sẽ xác định giả thuyết và đối thuyết tương ứng. Giả thuyết nghiên cứu là một giả thuyết có liên quan đến một trong 3 vấn đề sau: (1) Tính độc lập hay phụ thuộc của đại lượng ngẫu nhiên cần nghiên cứu; (2) Dạng của qui luật phân phối xác suất của đại lượng ngẫu nhiên; (3) Giá trị của tham số của đại lượng ngẫu nhiên.
- Bước 2. Xác định mức ý nghĩa (level of significance): Đối với một nghiên cứu ứng dụng, mức ý nghĩa thường được xác định $\alpha=0.05$.



- Bước 3. Lựa chọn phương pháp kiểm định: Như đã đề cập ở bước 1, việc lựa chọn phương pháp kiểm định sẽ phụ thuộc vào dữ liệu nghiên cứu (thang đo dùng để thu thập dữ liệu), quy luật phân phối của dữ liệu, mối quan hệ giữa các biến cần kiểm định (nếu có).
- Bước 4. Xác định giá trị kiểm định và mức ý nghĩa tương ứng: Với mỗi phương pháp kiểm định được lựa chọn sẽ tồn tại tương ứng một giá trị kiểm định. Giá trị này được xác định

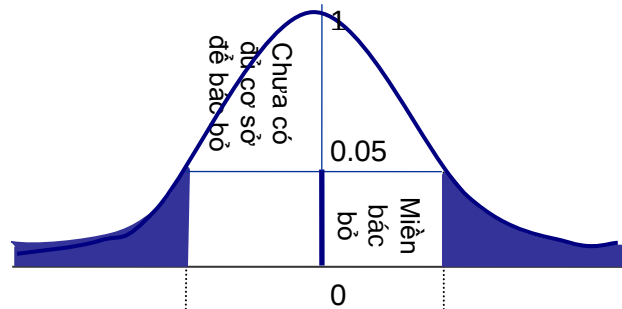
thông qua hàm ngược của hàm phân phối. Ví dụ với hàm phân phối chuẩn giá trị kiểm định $Z = 1.645$ (kiểm định một đuôi – one-tailed test) thì mức ý nghĩa tương ứng là $p=\alpha=0.05$.

- Bước 5. Kết luận: Với mỗi kiểm định luôn tồn tại một mức ý nghĩa tương ứng (với giá trị kiểm định) p :

+ nếu $p \leq 0.05$ (tức $0 \leq p \leq 0.05$): Kết luận bác bỏ H_0 , chấp nhận H_A .

+ nếu $p > 0.05$ (tức $0.05 < p \leq 1$): Kết luận *chưa có đủ cơ sở để bác bỏ H_0* , tức *chưa có đủ cơ sở để chấp nhận H_A* .

Chúng ta có thể giải thích tương tự với trường hợp kiểm định 2 đuôi (two-tailed tests).



KIỂM ĐỊNH SỰ KHÁC BIỆT GIỮA HAI THAM SỐ TRUNG BÌNH: T-TEST

Kiểm định tham số trung bình hai mẫu độc lập: T-test for Independent Groups

Vấn đề nghiên cứu

Trong nghiên cứu khoa học, chúng ta thường gặp việc so sánh giữa hai giá trị trung bình dựa trên hai mẫu độc lập. Ví dụ: Nhà nghiên cứu muốn trả lời câu hỏi (1) tuổi bình quân của nam và nữ có khác nhau hay không?; (2) thu nhập trung bình của những người cán bộ quản lý và những người không phải là cán bộ quản lý có khác nhau hay không?;... người ta thường dùng phương pháp kiểm định tham số trung bình T-test Independent Groups.

Sơ lược lý thuyết

B1. Giả thuyết và đối thuyết:

Giả thuyết: $H_0: \mu_x - \mu_y = D_0$

Đối thuyết: $H_A: \mu_x - \mu_y \neq D_0$

B2. Chọn mức ý nghĩa $\alpha = 0.05$

B3. Xác định phương pháp kiểm định: Phương pháp kiểm định sự khác biệt tham số trung bình hai mẫu (độc lập).

B4. Xác định giá trị kiểm định: Giá trị t (T-Student) với
$$t = \frac{\bar{x} - \bar{y} - D_0}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}}$$

B5. Kết luận: Nếu $p(t) \leq \alpha$, kết luận đủ cơ sở để bác bỏ giả thuyết và chấp nhận đối thuyết. Nếu $p(t) > \alpha$, chưa có đủ cơ sở để bác bỏ giả thuyết tức là chưa có đủ cơ sở để chấp nhận đối thuyết.

Yêu cầu

- Chỉ có một biến độc lập (independent variable) - ví dụ biến giới tính (GIOITINH)
- Chỉ có hai mức (hai sự lựa chọn) trong biến độc lập – ví dụ chỉ có hai giá trị là 1 là nam và 2 là nữ.
- Chỉ có một biến phụ thuộc – ví dụ biến độ tuổi (TUOI). Biến phụ thuộc luôn là biến định lượng (đo lường bởi cấp độ thang đo khoảng cách – interval scale hoặc tỷ lệ - ratio scale)

Giả định

- Phân phối mẫu của các khác biệt trung bình là phân phối chuẩn.
- Phương sai đồng nhất (homogeneity of variance).

Ví dụ

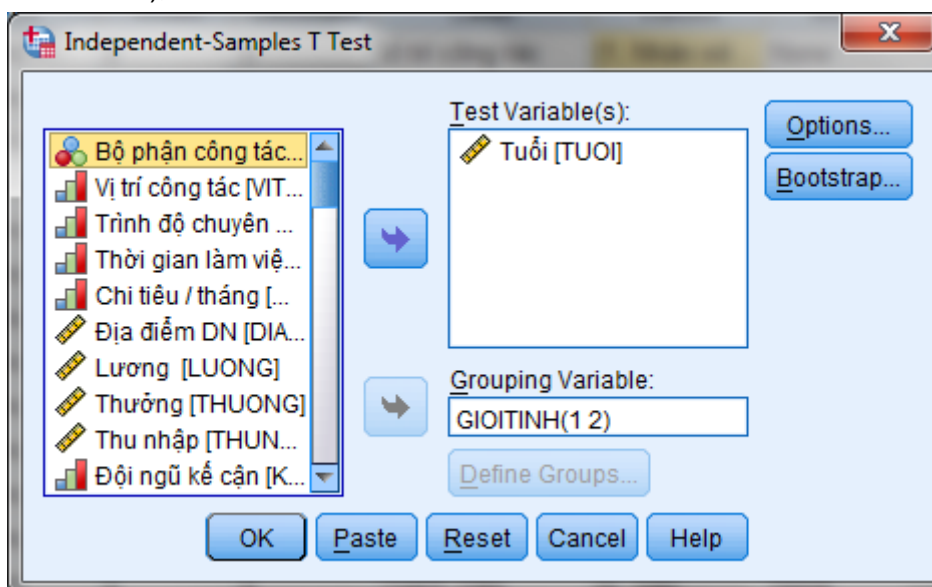
Người nghiên cứu muốn biết sự khác biệt trung bình về độ tuổi của nhân viên theo giới tính (nam và nữ) trong nghiên cứu 200 cán bộ công nhân viên, kết quả thu được nhập liệu tại file THUC HANH SPSS.sav.

Giả thuyết

- H_0 : Không có sự khác biệt giữa độ tuổi trung bình của nam và nữ trên tổng thể.
- H_A : Có sự khác biệt giữa độ tuổi trung bình của nam và nữ trên tổng thể.

Thực hành trên SPSS

- Từ thanh menu, kích **Analyze / Compare means / Independent-samples T- test**.
- Chọn biến TUOI vào ô **Test Variable(s)** và biến GIOITINH vào ô **Grouping Variable**.
- Kích vào **Define Groups** để định nghĩa hai giá trị của biến GIOITINH (chú ý giới tính nam = 1 và nữ =2), kích **Continue**.



- Kích **OK** để hoàn tất việc thiết đặt và phân tích kết quả.
- Kết quả và diễn giải:

Kiểm định Levene cho sự bằng nhau của của phương sai kiểm định cho phép khẳng định

rằng hai phương sai là bằng nhau, trong ví dụ này, Levene với $F = 0.345$ với mức ý nghĩa tương ứng là $p = 0.558 > \alpha = 0.05$. Do đó, có thể khẳng định giả định phương sai đồng nhất được chấp nhận và các giá trị thống kê t (t-test) tham chiếu theo dòng **Equal variances assumed**. Ngược lại, nếu mức ý nghĩa thống kê Levene nhỏ ($p < 0.05$) thì khẳng định phương sai tổng thể không bằng nhau (không đồng nhất), các thống kê t (t-test) được tham chiếu trên dòng **Equal variances not assumed**.

Vì $p(t, df) = p(2.783, 198) = 0.06 < \alpha = 0.05$ nên khẳng định có sự khác biệt giữa độ tuổi trung bình của nam và nữ trên tổng thể (tuổi trung bình của nam là 27.06 và nữ là 25.83).

Group Statistics					
	Giới tính	N	Mean	Std. Deviation	Std. Error Mean
Tuổi	Nam	104	27.06	3.353	.329
	Nu	96	25.83	2.820	.288

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Tuổi	Equal variances assumed	.345	.558	2.783	198	.006	1.224	.440	.357	2.092
	Equal variances not assumed			2.802	196.3	.006	1.224	.437	.363	2.086

Kiểm định tham số trung bình hai mẫu phụ thuộc hay mẫu phối hợp từng cặp: Paired – Samples T Test

Paired - Samples T Test được sử dụng trong đo lường lặp lại hoặc 2 nhóm tổng thể có mối quan hệ với nhau, trong đó mỗi chủ thể được kiểm định hai lần trên cùng một biến, kiểm định này thường được sử dụng trong thiết kế thử nghiệm trước và sau (pre-test và post-test). Ví dụ nhà nghiên cứu muốn thử nghiệm một quảng cáo, người ta tiến hành đo lường nhận thức của khách hàng trước quảng cáo, sau khi cho xem quảng cáo, người ta tiến hành đo lường lặp lại để xem xét có hay không sự khác biệt trong nhận thức của khách hàng trước và sau quảng cáo, khi đó, phương pháp kiểm định tham số trung bình hai mẫu phụ thuộc hay hai mẫu phối hợp từng cặp được sử dụng.

Sơ lược lý thuyết

B1. Giả thuyết và đối thuyết:

Giả thuyết: $H_0: \mu_x - \mu_y = D_0$

Đối thuyết: $H_A: \mu_x - \mu_y \neq D_0$

B2. Chọn mức ý nghĩa $\alpha = 0.05$

B3. Xác định phương pháp kiểm định: Phương pháp kiểm định sự khác biệt tham số trung bình giữa hai mẫu (phụ thuộc).

B4. Xác định giá trị kiểm định: Giá trị t (T-Student) với $t = \frac{(\bar{x} - D_0)\sqrt{n}}{S'_d}$

B5. Kết luận: Nếu $p(t) \leq \alpha$, kết luận đủ cơ sở để bác bỏ giả thuyết và chấp nhận đối thuyết. Nếu $p(t) > \alpha$, chưa có đủ cơ sở để bác bỏ giả thuyết tức là chưa có đủ cơ sở để chấp nhận đối thuyết.

Yêu cầu

- Trong bất kỳ một phân tích cần phải có 2 bộ dữ liệu;
- Hai bộ dữ liệu phải được thu thập từ (1) cùng một đối tượng hoặc (2) từ hai nhóm phối hợp từng cặp.

Giả định

- Độ lớn của hai mẫu phải tương đương nhau;
- Phân phối của sự khác biệt của các tham số trung bình phải là phân phối chuẩn.

Ví dụ

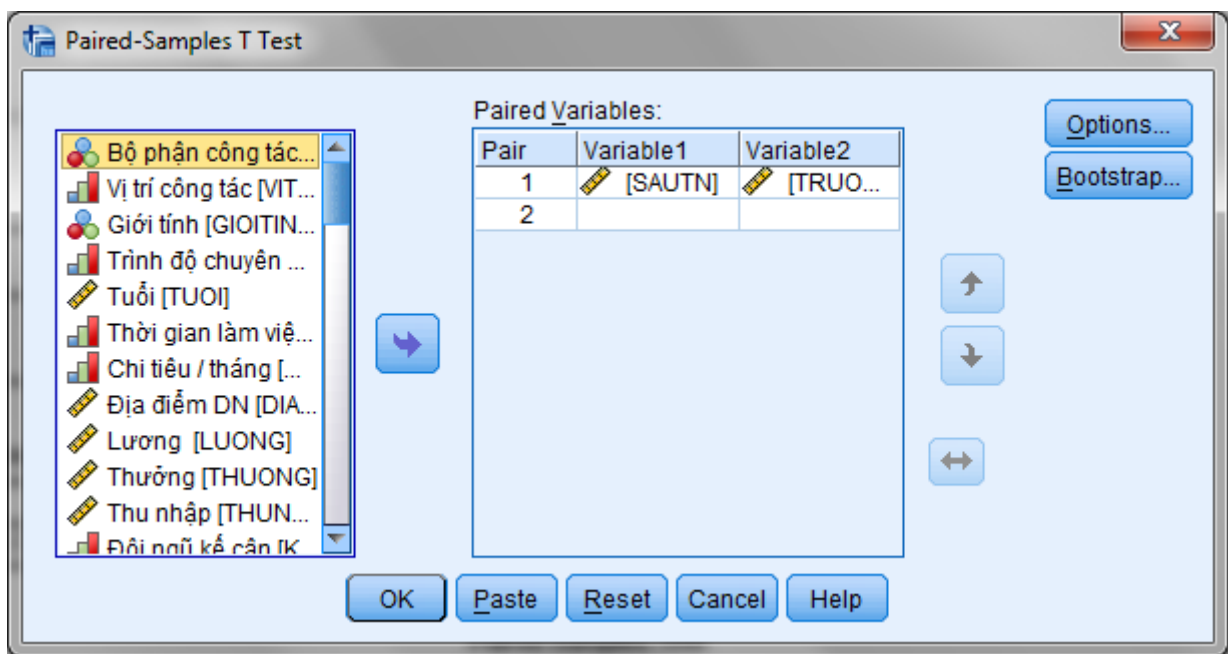
Để có cơ sở cho việc thực hiện thay đổi phương thức lãnh đạo, nhà quản trị quyết định thực hiện thử nghiệm phương thức lãnh đạo mới trên mẫu (có độ lớn $n=200$). Bằng cách đo lường sự nhận thức của cán bộ công nhân viên về mức độ tự chủ trước và sau thử nghiệm, việc nhận thức của cán bộ công nhân viên về mức độ tự chủ được đo lường bởi thang đo khoảng (interval scale) 7 lựa chọn với 1 là hoàn toàn không đồng ý và 7 là hoàn toàn đồng ý, kết quả nghiên cứu được nhập liệu tại hai biến TRUOCTN, SAUTN tại file THUC HANH SPSS.sav.

Giả thuyết

- H_0 : Không có sự khác biệt về nhận thức của cán bộ công nhân viên về mức độ tự chủ trước và sau khi thử nghiệm phương thức lãnh đạo mới trên tổng thể.
- H_A : Có sự khác biệt về nhận thức của cán bộ công nhân viên về mức độ tự chủ trước và sau khi thử nghiệm phương thức lãnh đạo mới trên tổng thể.

Thực hành trên SPSS

- Từ thanh menu, kích **Analyze / Compare means / Paired-Samples T Test**, khi đó sẽ xuất hiện hộp hội thoại **Paired-Samples T Test**
- Trong hộp hội thoại **Paired-Samples T Test** chọn biến TRUOCTN vào phần **Variable 1** và SAUTN vào phần **Variable 2**.
- Trong nghiên cứu này mẫu nghiên cứu là 200, nếu muốn phóng đại lên một cỡ mẫu lớn hơn (ví dụ 1000), kích vào **Bootstrap** để thiết đặt.
- Kích vào **OK** để thực hiện.



- Kết quả và diễn giải:

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	SAUTN - TRUOCTN	1.31500	2.18164	.15427	1.01080	1.61920	8.524	199	.000

Kết quả phân tích cho thấy trung bình chênh lệch trong nhận thức giữa sau và trước khi thực hiện thử nghiệm là 1.315 cấp độ với $t = 8.524$, $df = 199$ và $p(t, df) = p(8.524, 199) = 0.000 < 0.05$, điều đó cho phép khẳng định bác bỏ H_0 và chấp nhận H_A nghĩa là có một sự khác biệt đáng kể trong nhận thức của cán bộ công nhân viên trong doanh nghiệp trước và sau khi thực hiện thử nghiệm.

Giá trị trung bình chênh lệch $mean = 1.315 > 0$ có nghĩa là sau khi thử nghiệm tốt hơn trước khi thử nghiệm và nhà quản trị nên áp dụng phương pháp lãnh đạo mới trong doanh nghiệp.

KIỂM ĐỊNH MỐI QUAN HỆ GIỮA MỘT BIẾN ĐỊNH TÍNH VÀ MỘT BIẾN ĐỊNH LƯỢNG: PHÂN TÍCH PHƯƠNG SAI MỘT CHIỀU (ONE – WAY ANALYSIS OF VARIANCE)

Trong phần trước đã giới thiệu so sánh tham số trung bình hai mẫu độc lập, như vậy, biến định tính phân loại chỉ có 2 nhóm (2 lựa chọn, ví dụ như đúng / sai; đồng ý / không đồng ý; nam / nữ...). Trong trường hợp có hơn 3 lựa chọn chúng ta cũng có thể sử dụng kiểm định T để thực hiện việc phân tích, tuy nhiên phải thực hiện so sánh theo từng cặp, khi đó, sai số có thể tăng cao hơn theo số lần thực hiện kiểm định.

Để thuận tiện cho quá trình phân tích đối với biến định tính có hơn 2 sự loại chọn, phân tích phương sai một chiều (One – way ANOVA) được sử dụng, ANOVA là sự mở rộng của kiểm định T, phương pháp này cho phép so sánh giá trị trung bình của 3 nhóm trở lên.

Sơ lược lý thuyết

B1. Giả thiết và đối thiết trong phân tích phương sai một chiều được phát biểu như sau:

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k$$

H_1 : Tồn tại ít nhất một giá trị trung bình của nhóm thứ i (μ_i) khác với ít nhất giá trị trung bình của một nhóm còn lại.

B2. Lựa chọn mức ý nghĩa α .

B3. Xác định phương pháp kiểm định: Bài toán phân tích phương sai một chiều (One-way ANOVA).

B4. Xác định giá trị kiểm định:

Để tính tiêu chuẩn kiểm định trong phân tích phương sai (ANOVA), chúng ta cần tiến hành tính các chỉ tiêu sau:

+ Tổng độ lệch bình phương giữa các nhóm (Sum of squares between groups): phản ánh biến thiên của biến định lượng đánh nghiên cứu do tác động của biến phân loại đang xem xét

$$SSG = \sum_{i=1}^k (\bar{x} - \bar{x}_i)^2$$

+ Tổng độ lệch bình phương trong nội bộ nhóm (Sum of squares within groups) phản ánh biến thiên ngẫu nhiên do ảnh hưởng của các yếu tố khác không xem xét ở mẫu.

$$SSW = \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$$

+ Tổng các độ lệch bình phương toàn bộ (Total sum of squares): phản ánh toàn bộ biến thiên của biến định lượng đang nghiên cứu.

$$SST = \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2$$

hay $SST = SSW + SSG$.

+ Phương sai giữa các nhóm (Mean squares between groups):

$$MSG = \frac{SSG}{k-1}$$

+ Phương sai trong nội bộ các nhóm (Mean squares within groups):

$$MSW = \frac{SSW}{n-k}$$

+ Lúc đó tiêu chuẩn kiểm định F (Fisher) được tính bằng:

$$F = \frac{MSG}{MSW}$$

Chúng ta có thể tóm gọn cách tính thông qua bảng sau:

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	$SSG = \sum_{i=1}^k (\bar{x} - \bar{x}_i)^2$	k-1	$MCG = SSG/(k-1)$	MCG/MSW	p(F)
Within Groups	$SSW = \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$	n-k	$MSW = SSW/(n-k)$		
Total	$SST = \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2$ ($SST=SSG+SSW$)	n-1			

B5. Kết luận: Nếu $p(t) \leq \alpha$, kết luận đủ cơ sở để bác bỏ giả thuyết và chấp nhận đối thuyết. Nếu $p(t) > \alpha$, chưa có đủ cơ sở để bác bỏ giả thuyết tức là chưa có đủ cơ sở để chấp nhận đối thuyết.

Yêu cầu

- Phải có một biến độc lập (biến định tính) có nhiều hơn hai cấp độ (hai sự lựa chọn);
- Phải có một biến phụ thuộc (là một biến định lượng).

Giả định

- Các nhóm phải có phân phối chuẩn;
- Phương sai của các nhóm phải đồng nhất;
- Các nhóm phải độc lập và được chọn một cách ngẫu nhiên.

Ví dụ

Người nghiên cứu quan tâm độ tuổi trung bình của cán bộ công nhân viên ở các bộ phận có bằng nhau hay không? Mã hóa các biến trong dữ liệu SPSS như sau:

Độ tuổi TUOI

Bộ phận công tác BP (1: Kinh doanh, 2: Kỹ thuật, 3: Văn phòng; 4: Kế toán; 5: Khác)

Khi đó chúng ta sẽ có 5 nhóm gồm nhóm 1: Kinh doanh, 2: Kỹ thuật, 3: Văn phòng; 4: Kế toán; 5: Khác và sẽ so sánh sự khác nhau giữa giá trị trung bình về độ tuổi của các nhóm này.

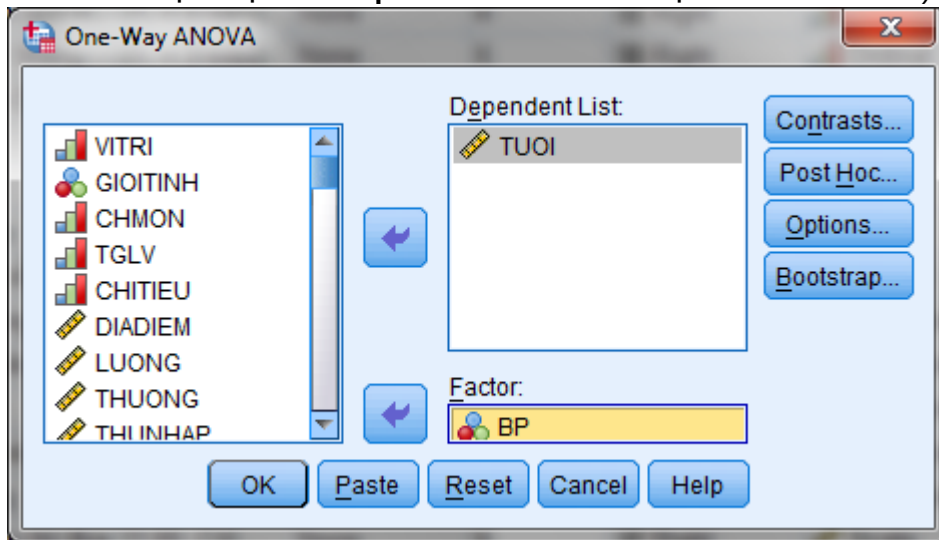
Giả thuyết

- H_0 : Không có sự khác biệt về độ tuổi trung bình giữa các bộ phận (hay tuổi trung bình giữa các bộ phận bằng nhau) trên tổng thể.
- H_A : Có sự khác biệt về độ tuổi trung bình giữa các bộ phận trên tổng thể (nghĩa là tồn tại độ tuổi trung bình của một bộ phận khác với trung bình của ít nhất một bộ phận còn lại).

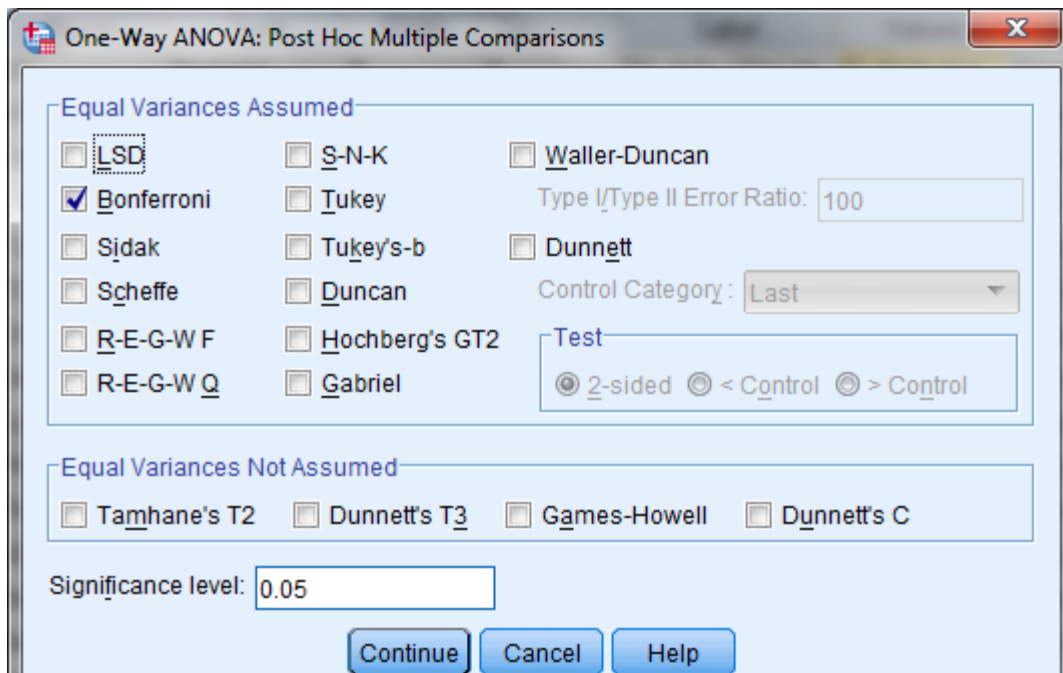
Cần phải chú ý rằng giả thuyết H_0 không đồng nghĩa với giá trị trung bình của tất cả các nhóm phải bằng nhau.

Thực hành trên SPSS

- Từ thanh menu, vào **Analyze / Compare Means / One-Way ANOVA**.
- Chọn biến TUOI vào ô **Dependent List** và biến BP vào ô **Factor** (chú ý: Biến định lượng luôn luôn được chọn ở ô **Dependent List** và biến định tính ở ô **Factor**).



- Để biết được độ tuổi trung bình của nhóm nào khác (hay bằng) với nhóm nào, kích vào **Post Hoc**, cửa sổ **One-Way ANOVA: Post Hoc Multiple Comparisons** xuất hiện, kích vào **Bonferroni** để so sánh trung bình từng cặp.



- Kích **Continue** để đóng cửa sổ **One-Way ANOVA: Post Hoc Multiple Comparisons**.
- Kích **OK** để hoàn tất thiết đặt và tính kết quả.
- Kết quả và diễn giải:

Qua phân tích ANOVA, kết quả cho được ở biểu sau:

ANOVA

Tuổi

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	127.202	4	31.800	3.333	.011
Within Groups	1860.618	195	9.542		
Total	1987.820	199			

Kết quả phân tích ANOVA được diễn giải (tính toán) theo các công thức được liệt kê dưới đây:

Vì $F = 3.333$ và $p(F) = 0.011 < 0.05$ nên đủ cơ sở để bác bỏ giả thuyết, chấp nhận đối thuyết hay có sự khác biệt về độ tuổi trung bình giữa các bộ phận trên tổng thể (nghĩa là tồn tại độ tuổi trung bình của một bộ phận khác với tuổi trung bình của ít nhất một bộ phận còn lại).

Multiple Comparisons

Tuổi

Bonferroni

(I) Bộ phận công tác	(J) Bộ phận công tác	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Kinh doanh	Kỹ thuật	-.090	.702	1.000	-2.08	1.90
	Văn phòng	.632	.737	1.000	-1.46	2.72
	Kế toán	1.917	.697	.065	-.06	3.90
	Khác	.232	.792	1.000	-2.02	2.48
Kỹ thuật	Kinh doanh	.090	.702	1.000	-1.90	2.08
	Văn phòng	.722	.663	1.000	-1.16	2.60
	Kế toán	2.007*	.618	.014	.25	3.76
	Khác	.322	.724	1.000	-1.73	2.38
Văn phòng	Kinh doanh	-.632	.737	1.000	-2.72	1.46
	Kỹ thuật	-.722	.663	1.000	-2.60	1.16
	Kế toán	1.285	.657	.519	-.58	3.15
	Khác	-.401	.757	1.000	-2.55	1.75
Kế toán	Kinh doanh	-1.917	.697	.065	-3.90	.06
	Kỹ thuật	-2.007*	.618	.014	-3.76	-.25
	Văn phòng	-1.285	.657	.519	-3.15	.58
	Khác	-1.686	.718	.200	-3.73	.35
Khác	Kinh doanh	-.232	.792	1.000	-2.48	2.02
	Kỹ thuật	-.322	.724	1.000	-2.38	1.73
	Văn phòng	.401	.757	1.000	-1.75	2.55
	Kế toán	1.686	.718	.200	-.35	3.73

*. The mean difference is significant at the 0.05 level.

Kết quả phân tích nhân tố ANOVA có $p = 0.011$ chỉ ra rằng tồn tại độ tuổi trung bình của một bộ phận khác với ít nhất một bộ phận còn lại. Khi đó, câu hỏi đặt ra là tuổi trung bình của bộ phận nào khác với bộ phận nào, so sánh **Post Hoc** (Post Hoc Comparisons) sẽ giúp nhà phân tích trả lời câu hỏi này.

Trong bảng **Multiple Comparision**, cột Mean Difference (I-J) đi kèm với dấu (*) tương ứng với Sig. < 0.05 cho phép khẳng định sự chênh lệch có ý nghĩa giữa hai tham số trung bình, nghĩa là giá trị trung bình của đại lượng I khác với trung bình đại lượng J. Trong ví dụ trên, tương ứng với dấu (*) ở cột Mean Difference cho phép khẳng định độ tuổi trung bình của bộ phận Kỹ thuật khác với độ tuổi bình của bộ phận Kế toán vì có Sig. = 0.014 < 0.05 (tuổi trung bình bộ phận Kỹ thuật lớn hơn bộ phận Kế toán là 2.007 tuổi).

CHƯƠNG

8.2

TƯƠNG QUAN VÀ HỒI QUY TUYẾN TÍNH (CORRELATION VÀ LINEAR REGRESSION)

Các phần trước đã đề cập mối quan hệ giữa hai biến định tính, một biến định tính và một biến định lượng, trong phần này sẽ giới thiệu mối quan hệ giữa hai biến định lượng trên cơ sở tương quan (correlation) và hồi quy đơn tuyến tính (linear regression).

TƯƠNG QUAN (CORRELATION)

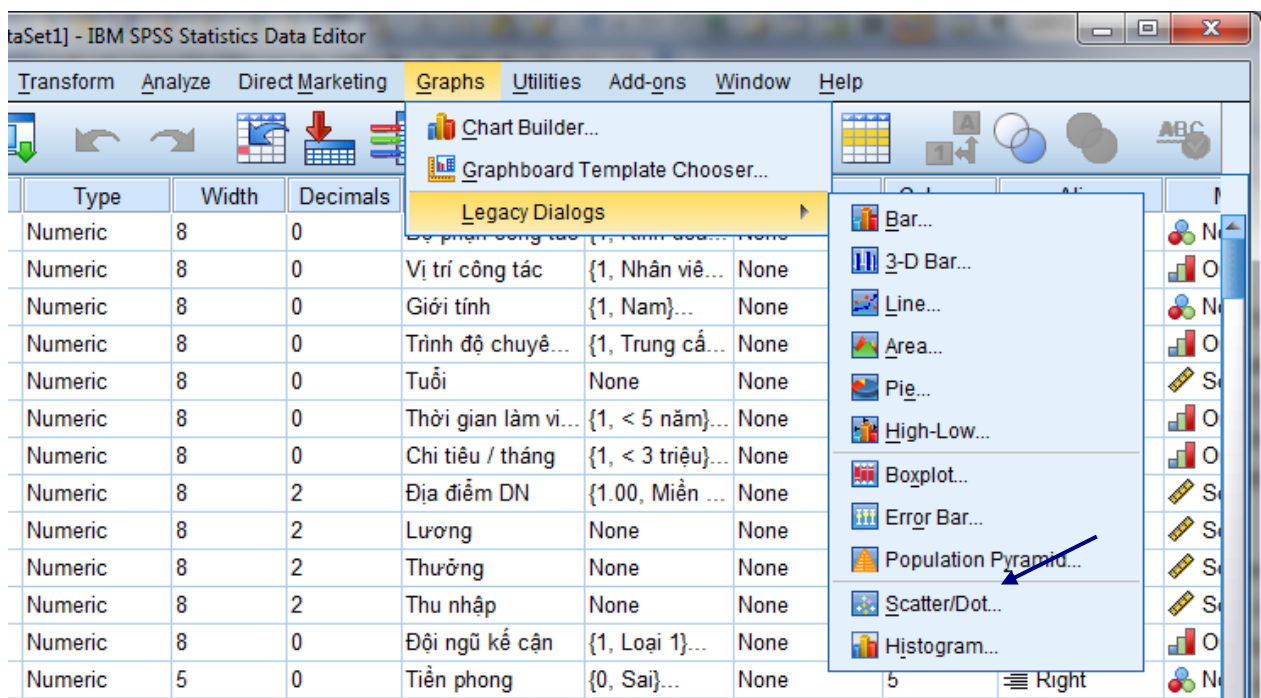
Sự tương quan tập trung chủ yếu đến việc tìm xem có hay không mối quan hệ tồn tại giữa hai biến với việc xác định mức độ và chiều hướng của nó. Ví dụ khi hai biến cùng thay đổi, cùng chiều hoặc ngược chiều, chúng được xem như có tương quan. Do đó, nghiên cứu sự tương quan là cố gắng tìm ra mức độ mà hai hay nhiều biến có liên quan nhau. Trong nghiên cứu sự tương quan, người nghiên cứu đo lường các sự kiện, hành vi hay đặc điểm tính cách xuất hiện tự nhiên và sau đó định rõ liệu các kết quả được đo lường được có hiệp biến hay không.

Nghiên cứu tương quan đơn giản nhất phải có được cặp quan sát hoặc phải đo lường trên hai biến khác nhau chứ không phải là một số cá thể riêng biệt. Các đo lường theo cặp được phân tích thống kê để xác định có bất kì mối quan hệ nào tồn tại giữa chúng hay không. Ví dụ, thực tế tại Việt Nam, có mối quan hệ giữa thời gian công tác (x) và lương của người đó được hưởng (y) hoặc có mối quan hệ giữa phong cách lãnh đạo (x) với lòng trung thành của nhân viên đối với tổ chức (y)...

Để biểu diễn định lượng mức độ mà hai biến tương quan với nhau thì phải cần tính hệ số tương quan (correlation coefficient). Có nhiều loại hệ số tương quan và việc sử dụng loại nào cho một tập tin dữ liệu cụ thể phụ thuộc vào các yếu tố sau:

- Mức độ phạm vi mà các biến được đo lường;
- Loại phân phối cơ bản (liên tục hay rời rạc);
- Bản chất của hàm phân phối (tuyến tính hay không tuyến tính).

Một cách trực quan, chúng ta có thể sử dụng đồ thị phân tán để cho thấy nhiều loại liên hệ giữa hai biến khảo sát, biểu đồ **Scatter** của SPSS để biểu diễn mối quan hệ của hai biến số nhằm xem xét có khả năng chúng có (1) không có liên hệ (các cặp giá trị phân tán ngẫu nhiên), (2) liên hệ tuyến tính thuận (mối quan hệ gần như tuyến tính và khi giá trị này tăng thì giá trị kia cũng sẽ tăng), (3) liên hệ tuyến tính nghịch (giống như trường hợp tuyến tính thuận nhưng chiều hướng ngược lại) và (4) liên hệ phi tuyến (đường đồ thị không phân bố theo một chiều hướng duy nhất) bằng cách kích vào **Graphs / Legacy Dialogs / Scatter/Dot** để vẽ các hình vẽ thể hiện mối liên hệ.



Trong chương này sẽ giới thiệu hai hệ số tương quan: hệ số tương quan Pearson (r), được sử dụng cho các biến có thang đo khoảng cách hoặc tỷ lệ (các biến hệ metric) và hệ số tương quan hạng Spearman (r_{rho}) sử dụng cho dữ liệu sắp xếp theo thứ tự. Cần lưu ý rằng, bất kỳ kỹ thuật tương quan nào được sử dụng thì chúng đều phải có các đặc điểm như sau:

- Hai tập hợp đo lường được lấy trên những cá thể giống nhau hoặc trên các cặp cá thể được kết hợp trên một vài nền tảng cơ bản.
- Các giá trị của các hệ số tương quan biến thiên trong khoảng từ +1.00 và -1.00, cả hai giá trị này đều tượng trưng cho mối quan hệ hoàn hảo giữa các biến số, còn 0.00 tượng trưng cho việc các biến không có quan hệ với nhau.
- Một mối quan hệ cùng chiều có nghĩa là các cá thể đạt được giá trị lớn trên một biến thì hướng đến việc đạt được những giá trị cao ở biến thứ hai. Điều ngược lại cũng đúng, các cá thể đạt được giá trị thấp trên một biến thì cũng sẽ như vậy ở biến thứ hai.
- Mối quan hệ ngược chiều có nghĩa là các cá thể đạt được giá trị thấp trên một biến thì sẽ có khuynh hướng đạt được giá trị cao ở biến thứ hai. Ngược lại, các cá thể đạt được giá trị cao trên một biến thì biến thứ hai sẽ có khuynh hướng đạt được giá trị thấp.

Giả định cho bài toán tương quan:

- Với mỗi đối tượng trong nghiên cứu, chắc chắn có các cặp liên quan về giá trị, nếu một đối tượng có một giá trị trên biến X, thì đối tượng tương tự chắc chắn cũng có một giá trị trên biến Y.
- Mối quan hệ giữa hai biến phải là tuyến tính, mối quan hệ này có thể được trình bày chính xác bằng một đường thẳng.
- Các biến số nên được đo lường tối thiểu ở cấp độ thứ tự.
- Tính biến thiên về độ lớn của biến Y nên được giữ nguyên không đổi với tất cả giá trị của

biến X, giả định này được gọi là phương sai không đổi.

Hệ số tương quan Pearson (r)

Sơ lược lý thuyết

Hệ số tương quan r (Pearson Correlation Coefficient) dùng để biểu hiện mức độ chặt chẽ của hai biến định lượng (đo bởi cấp độ thang đo khoảng cách hoặc tỷ lệ). Hệ số r được tính như sau:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y} \quad (-1 \leq r \leq +1)$$

Với n là số quan sát, s_x và s_y là độ lệch chuẩn của hai biến x và y.

Trong kiểm định giả thuyết về hệ số tương quan của tổng thể (kí hiệu là p), giả thuyết này cần một số giả định về phân phối chung của cả hai biến x và y là các mẫu ngẫu nhiên độc lập lấy từ một tổng thể trong đó cả hai biến đều phân phối chuẩn.

B1. Giả thuyết và đối thuyết:

Giả thuyết: H_0 : Không có mối quan hệ giữa hai biến.

Đối thuyết: H_A : Có mối quan hệ giữa hai biến.

B2. Chọn mức ý nghĩa $\alpha = 0.05$

B3. Xác định phương pháp kiểm định: Phương pháp kiểm định hệ số tương quan Pearson r với phân phối t (T-student).

B4. Xác định giá trị kiểm định: Giá trị t (T-Student) với $t = r \sqrt{\frac{n-2}{1-r^2}}$

B5. Kết luận: Nếu $p(t) \leq \alpha$, kết luận đủ cơ sở để bác bỏ giả thuyết và chấp nhận đối thuyết. Nếu $p(t) > \alpha$, chưa có đủ cơ sở để bác bỏ giả thuyết tức là chưa có đủ cơ sở để chấp nhận đối thuyết.

Ví dụ

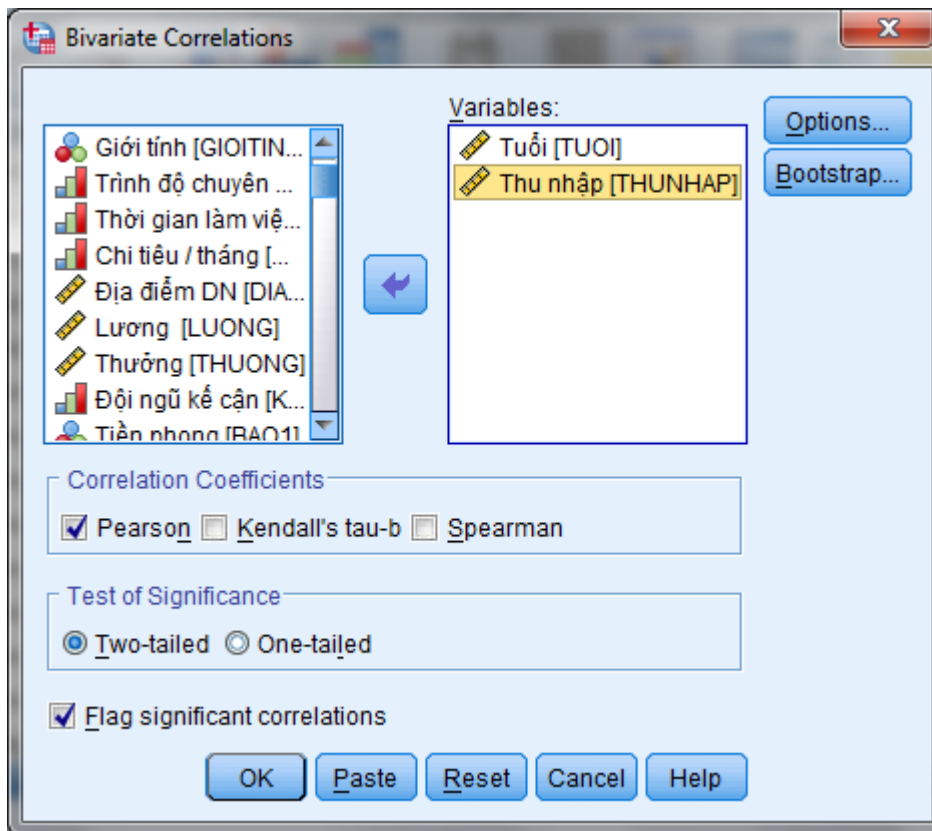
Trong doanh nghiệp, có nhiều ý kiến cho rằng, cách thức trả lương của nhà quản lý theo độ tuổi, nhà nghiên cứu đã ghi lại các cặp giá trị được nhập liệu và lưu tại file THUC HANH SPSS.sav.

- Câu hỏi nghiên cứu: Độ tuổi và thu nhập có mối quan hệ với nhau hay không ?
- Hai biến được chọn để thực hiện nghiên cứu là: Độ tuổi (TUOI) và thu nhập (THUNHAP), đây là 2 biến định lượng được đo lường bởi thang đo tỉ lệ.
- Giả thuyết H_0 : Độ tuổi và thu nhập không có liên hệ với nhau trên tổng thể.

Chú ý: Để đơn giản trong quá trình trình bày, trong nghiên cứu khoa học người ta chỉ phát biểu giả thuyết không (null hypothesis - H_0) và không phát biểu giả thuyết đối (Alternative hypothesis) nên người đọc phải ngầm hiểu đối thuyết này.

Thực hành trên SPSS

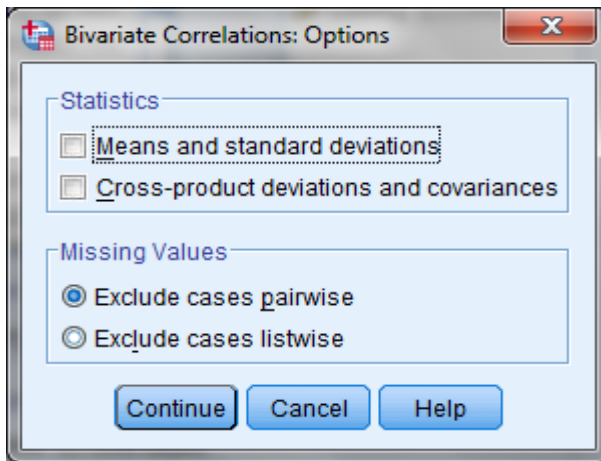
- Từ thanh công cụ, chọn **Analyze / Correlate / Bivariate...**, cửa sổ **Bivariate Correlations** sẽ xuất hiện.
- Chọn biến TUOI và LUONG vào ô **Variables**.
- Giá trị **Pearson** tại ô **Correlation Coefficients** được mặc định (chúng ta có thể sử dụng hệ số tương quan **Tau-b Kendall**).
- Kiểm định hai phía (Two-tailed) hoặc một phía (One-tailed) tại ô **Tests of Significance**, chú ý nếu chưa biết về chiều hướng của mối quan hệ (thuận hay nghịch) thì nên sử dụng kiểm định hai phía.



- Nhấp chuột vào **Options** để mở cửa sổ **Bivariate Correlation: Options**.

Theo mặc định, SPSS sử dụng phương pháp **Exclude cases pairwise** để sử dụng dữ liệu khuyết (được đánh dấu trong SPSS). Phương pháp này xem kết quả của mỗi cặp biến là riêng biệt, sử dụng tất cả các đối tượng với dữ liệu cặp đầy đủ. Với phương pháp này, chỉ các đối tượng có dữ liệu khuyết ở cặp biến cụ thể sẽ được đưa ra khỏi kết quả. Như vậy, sự tương quan giữa các cặp biến khác nhau dựa vào số lượng khác nhau của các đối tượng.

Như một tùy chọn, phương pháp **Exclude cases listwise** có thể chỉ được sử dụng cho các đối tượng có dữ liệu đầy đủ. Với phương pháp này, bất cứ đối tượng với dữ liệu khuyết hay cặp đều bị loại ra khỏi kết quả. Như vậy, kích thước mẫu cho sự tương quan giữa bất kỳ cặp biến nào đều có thể bị rút gọn ở phương pháp listwise.



- Nhấn **Continue / OK** để hoàn tất việc thiết đặt và cho kết quả nghiên cứu.
- Kết quả và diễn giải:

Correlations			
		Tuổi	Thu nhập
Tuổi	Pearson Correlation	1	.075
	Sig. (2-tailed)		.291
	N	200	200
Thu nhập	Pearson Correlation	.075	1
	Sig. (2-tailed)	.291	
	N	200	200

Như vậy, không có sự tương quan giữa kết quả tuổi và thu nhập của các bộ công nhân viên trên tổng thể vì $r = 0.075$ và Sig (2-tailed) hay $p = 0.291 > 0.05$.

Hệ số tương quan hạng (r)

Như đã đề cập ở phần trên, hệ số tương quan r (Pearson Correlation Coefficient) dùng để biểu hiện mức độ chặt chẽ của hai biến định lượng (đo bởi thang đo khoảng cách hoặc tỷ lệ) và giả định cần để kiểm định giả thuyết về hệ số tương quan tuyến tính là tổng thể có phân phối chuẩn. Trong trường hợp dữ liệu không thỏa mãn, cần sử dụng một hệ số tương quan khác là hệ số tương quan hạng **Spearman** (Spearman-rank-order – rho)

Hệ số tương quan hạng được tính dựa vào các hạng của dữ liệu, trước khi thực hiện trước hết cần sắp xếp dữ liệu theo hạng và sau đó tính toán hệ số Spearman (hệ số này có giá trị từ -1 đến +1 với -1 và +1 cho thấy mối quan hệ tuyến tính tuyệt đối giữa hạng của hai biến) và cần chú ý rằng đây là liên hệ giữa các hạng chứ không phải liên hệ giữa các giá trị. Việc thực hiện trên SPSS cũng tương tự như tương quan Pearson nhưng chọn giá trị **Spearman** tại ô **Correlation Coefficients**.

HỒI QUY TUYẾN TÍNH (LINEAR REGRESSION)

Hồi quy (correlation) và tương quan (regression) có quan hệ chặt chẽ với nhau, cả hai kỹ thuật bao gồm mối quan hệ giữa hai biến và chúng đều sử dụng tập các cặp kết quả giống nhau được lấy từ cùng đối tượng. Tuy nhiên, trong khi sự tương quan (correlation) quan tâm đến

mức độ và chiều hướng của mối quan hệ, thì hồi quy (regression) tập trung vào việc sử dụng mối quan hệ để dự đoán (prediction). Về mặt dự đoán, nếu hai biến tương quan hoàn toàn với nhau thì khi biết giá trị của một kết quả của biến thứ nhất sẽ cho phép dự đoán chính xác kết quả của biến thứ hai. Thông thường, khi hai biến hoàn toàn tương quan với nhau thì người nghiên cứu có thể sử dụng kết quả của một biến để dự báo kết quả biến thứ hai.

Để tiện cho việc theo dõi, phần này sẽ trình bày hồi quy đơn tuyến tính (linear regression) trên cơ sở mối quan hệ giữa một biến phụ thuộc (dependent variable – DV) định lượng và một biến độc lập (independent variable – IV) định lượng. Quan hệ tương quan và hồi quy tuyến tính bội (Multi – Regression) giữa một biến phụ thuộc (DV) định lượng và nhiều biến độc lập (IV) sẽ được trình bày ở phần sau.

Sơ lược về lý thuyết

Mô hình mối quan hệ tuyến tính giữa một biến phụ thuộc (dependent variable – DV) định lượng và một biến độc lập (independent variable – IV) định lượng trên mẫu được biểu hiện thông qua phương trình: $y_i = b_0 + b_1x_i$ và phương trình hồi quy trên tổng thể có dạng: $\hat{Y}_i = B_0 + B_1X_i + \varepsilon_i$ với X_i là giá trị quan sát thứ i của biến độc lập; $\hat{Y}_i$ là giá trị dự đoán thứ i của biến phụ thuộc; B_0 và B_1 là hệ số hồi quy (xác định thông qua phương pháp bình phương bé nhất thông thường (Ordinary Least Square – OLS) và ε_i là sai số của dự đoán. Có nhiều lý do tại sao những người nghiên cứu muốn dự đoán một biến từ biến kia, chẳng hạn, nếu biết cảm nhận về chính sách đào tạo, môi trường làm việc... của cán bộ công nhân viên trong một tổ chức, chúng ta có thể dự đoán gì về lòng trung thành của các cán bộ công nhân viên đối với doanh nghiệp...? Từ đó, các nhà quản lý doanh nghiệp có thể nhận biết rằng, muốn gia tăng lòng trung thành đối với doanh nghiệp, cần phải thực hiện (thay đổi) như thế nào chính sách về đào tạo, môi trường làm việc?. Trong phân tích hồi quy tuyến tính nói chung và hồi quy đơn nói riêng, một số vấn đề mà người phân tích dữ liệu cần phải được xem xét gồm:

Giả định

- Với mỗi đối tượng trong nghiên cứu, chắc hẳn sẽ có các cặp điểm tương quan, nếu một đối tượng có điểm (giá trị) ở biến X , thì đối tượng này cũng phải có điểm (giá trị) ở biến Y ,
- Mối quan hệ giữa hai biến phải là tuyến tính, tức là, mối quan hệ có thể được biểu diễn gần như là một đường thẳng,
- Các biến nên được đo lường bởi biến định lượng (cấp thang đo khoảng hoặc tỷ lệ),
- Sự thay đổi của kết quả ở biến Y luôn theo các giá trị của biến X , giả định này được gọi là hiệp phương sai đồng nhất.

Độ chính xác của ước lượng các tham số tổng thể từ các hệ số hồi quy mẫu

Trong phân tích hồi quy tuyến tính, giá trị thống kê tính từ mẫu giúp cho chúng ta ước lượng điểm về tham số của tổng thể và tham số B_0 và B_1 , chúng ta sẽ ước lượng các tham số này của tổng thể.

- Sai số chuẩn của B_0 là:
$$\sigma_{B_0} = \sigma \sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{(n-1)s_x^2}\right)}$$

- Sai số chuẩn của B_1 là: $\sigma_{B_1} = \frac{\sigma}{\sqrt{(n-1)s_x^2}}$
- Khi đó, sai số của ước lượng là: $S = \sqrt{\frac{\sum_{i=1}^n (Y_i - B_0 - B_1 x)^2}{n-2}}$

Chú ý: Các ước lượng sai số chuẩn của hệ số góc B_1 và hằng số B_0 được thể hiện trong cột Std. Error (Standard Error) của biểu tính hệ số Coefficients.

Đánh giá độ phù hợp của mô hình

Trên thực tế, các quan sát hầu như không phân bố theo một đường thẳng (mà gần như đường thẳng), vì vậy luôn có một sự sai lệch giữa các giá trị dự báo cho ra bởi đường thẳng và các giá trị thực tế thông qua phân dư.

Để biết mô hình hồi quy tuyến tính xây dựng trên mẫu có phù hợp với dữ liệu, một thước đo về sự phù hợp đó là hệ số xác định R^2 (coefficient of determination). Hệ số này được tính toán trên cơ sở công thức sau với Y_i là giá trị quan sát, $\bar{Y}$ là giá trị trung bình của Y và $\hat{Y}_i$ là giá trị ước lượng của Y .

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad 0 \leq R^2 \leq 1$$

- $R^2 = 1$ thể hiện mô hình hồi quy tuyến tính xây dựng được phù hợp 100% với dữ liệu mẫu, đây là một tình huống gần như không có trên thực tế vì mô hình có tốt đến đâu cũng không thể đạt được giá trị $R^2 = 1$.
- $R^2 = 0$ nghĩa là không có liên hệ tuyến tính giữa hai biến.
- $R^2 = a$ ($0 < a < 1$) nghĩa là mô hình hồi quy tuyến tính đã xây dựng phù hợp với dữ liệu là $(a \cdot 100)\%$ nghĩa là biến độc lập giải thích $(a \cdot 100)\%$ biến phụ thuộc (Ví dụ: $R^2 = 0.763$ nghĩa là biến độc lập giải thích 76.3% biến phụ thuộc).

Chú ý:	Nếu	$R^2 < 0.3$	Mối quan hệ yếu
	Nếu	$0.3 \leq R^2 < 0.5$	Mối quan hệ trung bình (chấp nhận)
	Nếu	$0.5 \leq R^2 < 0.7$	Mối quan hệ khá chặt chẽ
	Nếu	$1 \leq R^2$	Mối quan hệ rất chặt chẽ

Ví dụ

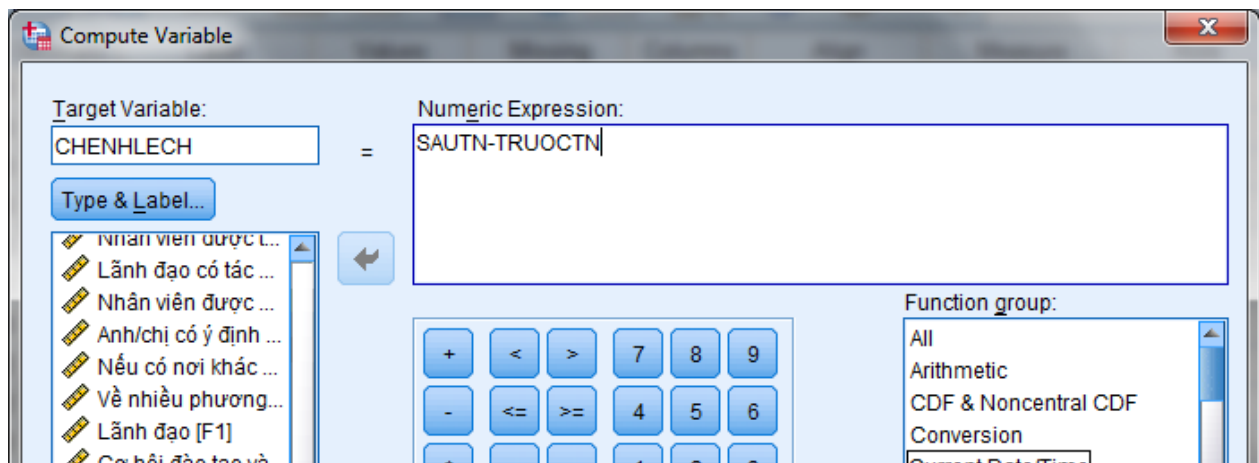
Trong phần trước, nhà quản lý thực hiện một cuộc thử nghiệm trước và sau để xem xét có sự chênh lệch về cảm nhận của cán bộ công nhân viên trước và sau khi thử nghiệm một phong cách quản lý mới. Đồng thời, để nghiên cứu đánh giá triển vọng về sự thành công của phong cách lãnh đạo mới này trong doanh nghiệp, người nghiên cứu thu thập biến số triển vọng (TRIENVONG) được đo lường bởi thang đo khoảng cách với (1) là hoàn toàn không triển vọng và (7) là hoàn toàn triển vọng. Kết quả được thu thập và nhập liệu tại file THUC HANH SPSS.sav.

- Câu hỏi nghiên cứu: Triển vọng về sự thành công của phong cách lãnh đạo mới có phụ thuộc vào sự cảm nhận của cán bộ công nhân viên sau và trước thử nghiệm?
- Hai biến được chọn: Biến độc lập là biến chênh lệch sau và trước thử nghiệm được tạo ra (CHENHLECH), biến phụ thuộc là biến TRIENVONG.
- Giả thuyết: H_0 : Không có mối quan hệ tuyến tính giữa chênh lệch cảm nhận sau và trước thử nghiệm (CHENHLECH) và triển vọng thành công của việc áp dụng phong cách lãnh đạo mới (TRIENVONG).

Thực hành trên SPSS

Cần phải hiểu rằng, sự thay đổi trong nhận thức của cán bộ công nhân viên trong doanh nghiệp được tạo thành từ chênh lệch trong nhận thức sau và trước thử nghiệm. Như vậy, để phân tích được, cần phải tạo ra một biến mới (CHENHLECH) bằng chênh lệch giữa nhận thức (giá trị) sau khi thử nghiệm – nhận thức trước khi thử nghiệm (thang đo nhận thức trước và sau thử nghiệm được đo bằng thang đo khoảng cách 7 sự lựa chọn).

Trong phần trước đã giới thiệu cách thức để tạo một biến mới, việc tạo ra biến mới chênh lệch (CHENHLECH) bằng cách vào **Transform / Compute** và thiết đặt thông tin trong hộp hội thoại **Compute Variable** như sau:

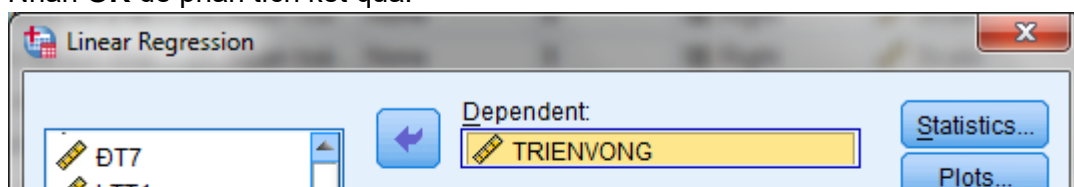


Như vậy trong cơ sở dữ liệu đã hình thành nên biến CHENHLECH bằng hiệu hai giá trị của biến SAUTN và TRUOCTN. Từ đó, chúng ta có thể thực hiện phân tích hồi quy giữa biến CHENHLECH và biến TRIENVONG.

- Vào **Analyze / Regression / Linear**.
- Trong hộp hội thoại **Linear Regression**, chọn biến độc lập (Independent Variable – IV) là CHENHLECH vào ô **Independent(s)**. Chọn biến phụ thuộc (Dependent Variable – DV) vào ô **Dependent**.

Chú ý: Trong phân tích hồi quy tuyến tính, cần phải xác định một cách rõ ràng biến nào là biến độc lập (Independent Variable – IV) và biến nào là biến phụ thuộc (Dependent Variable – DV).

- Nhấn **OK** để phân tích kết quả.



- Kết quả và diễn giải: Trong kết quả phân tích hồi quy tương quan cần giải thích, diễn giải các thông tin ít nhất từ 3 biểu sau:

+ Bảng ANOVA: Vì F (Fisher) = 12.726 và p (F) = 0.000 < 0.05 nên có thể khẳng định tồn tại mối quan hệ giữa biến chênh lệch (CHENHLECH) và triển vọng (TRIENVONG) trên tổng thể.

ANOVA ^b						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	33.491	1	33.491	12.726	.000 ^a
	Residual	521.064	198	2.632		
	Total	554.555	199			

a. Predictors: (Constant), CHENHLECH

b. Dependent Variable: TRIENVONG

+ Phương trình thể hiện mối quan hệ tuyến tính theo đường thẳng như sau:

Theo phương trình hồi quy lý thuyết $TRIENVONG_i = B_0 + B_1CHENHLECH_i + \varepsilon_i$.

Coefficients ^a						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4.582	.134		34.190	.000
	CHENHLECH	-.188	.053	-.246	-3.567	.000

a. Dependent Variable: TRIENVONG

Thông tin liên quan đến việc xây dựng phương trình hồi quy (dự báo) bình phương bé nhất được biểu diễn trong bảng Coefficients, với $t = -3.567$ và p (t) = 0.000 < 0.05 nên khẳng định có mối hệ giữa biến độc lập (CHENHLECH) và biến phụ thuộc TRIENVONG. Sử dụng các giá trị (giá trị Constant (B_0) và B_1) được đưa ra trong cột Unstandardized Coefficients, phương trình dự báo sẽ là:

$$TRIENVONG_i = 4.582 - 0.188CHENHLECH_i + \varepsilon_i.$$

Như vậy, với sự thay đổi mỗi đơn vị (1 cấp độ trong 7 cấp độ của thang đo biến chênh lệch) sẽ làm giảm triển vọng thành công áp dụng phong cách lãnh đạo mới là 0.188 đơn vị (trong 7 cấp độ của thang đo triển vọng). Dựa vào kết quả này có thể đưa ra những gợi ý cho nhà

quản lý trong việc có tiếp tục áp dụng phương thức quản lý mới cho doanh nghiệp trong tương hay không...

+ Hệ số tương quan R^2 :

Hệ số tương quan R^2 được trích từ biểu Model Summary, vì $R^2 = 0.060$ nên có thể khẳng định mối quan hệ giữa hai biến CHENHLECH và TRIENVONG là rất yếu và biến độc lập CHENHLECH chỉ giải thích 0.6% biến phụ thuộc TRIENVONG.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.246 ^a	.060	.056	1.622

a. Predictors: (Constant), CHENHLECH

CHƯƠNG

9.1

PHÂN TÍCH NHÂN TỐ KHÁM PHÁ (EXPLORE FACTOR ANALYSIS)

Trong các phần trước, chúng ta đã giới thiệu mối quan hệ của một vài biến, tuy nhiên, trong một số trường hợp, số lượng các biến được xem xét trong kĩ thuật thống kê tăng lên. Trên thực tế, hầu hết các biến này có liên hệ với nhau, vì vậy, cần phải giảm bớt xuống đến một số lượng nhân tố đại diện mà chúng ta có thể sử dụng để phân tích sau này.

Phân tích nhân tố khám phá (Explore Factor Analysis gọi tắt là EFA và sau đây gọi là phân tích nhân tố) được dựa trên giả định rằng tất cả các biến tương quan ở một mức độ nào đó. Chính vì vậy một vài biến quan sát tương tự nhau có tương quan cao, những biến đo lường không tương tự nhau sẽ có tương quan thấp.

Hay nói một cách khác, phân tích nhân tố khám phá bằng các thành phần chính (principal components) cho phép rút gọn nhiều biến số (variables, indicators hoặc items) ít nhiều có tương quan lẫn nhau thành những đại lượng được thể hiện dưới dạng mối tương quan theo đường thẳng được gọi là những nhân tố (factors).

Một điều cần phải chú ý ở đây là trong các nghiên cứu liên quan đến thái độ, người ta thường không nghiên cứu trực tiếp các biến số (ví dụ những vấn đề liên quan đến hành vi, động cơ của một con người) và các biến không đo lường trực tiếp được gọi là (latent variables) mà thông thường phát triển thông qua các biến quan sát nguyên thủy hay các chỉ báo (observed variables hay items). Do vậy, trong quá trình phân tích cần phải rút trích từ tập các chỉ báo đo lường trực tiếp được (items) thành tập các nhân tố (factors) để hình thành các biến tiềm ẩn (latent variables) và thể hiện mối quan hệ giữa nhân tố (latent variables) và biến quan sát nguyên thủy (observed variables, items).

Phân tích nhân tố cung cấp cho người nghiên cứu 2 khả năng: tổng hợp dữ liệu và cắt giảm dữ liệu. Ví dụ, chúng ta có những biến riêng lẻ (items) là “tôi mua hàng đặc biệt”, “tôi thường quan tâm đến giá thấp nhất”, “tôi mua hàng thương lượng”, “nhãn hiệu quốc gia đáng giá hơn nhãn hiệu của cửa hàng”, tập hợp lại, những biến này có thể được sử dụng để xác định khách hàng “biết rõ giá” hoặc là “chuyên gia thương lượng mua hàng” hay còn gọi đây là biến số tiềm ẩn (latent variable).

Trong phân tích phương sai, hồi quy... cần phải xem xét và phân biệt biến độc lập và biến phụ thuộc, trong phân tích nhân tố không có sự phân biệt này nên gọi là một kĩ thuật phụ thuộc lẫn nhau.

MỘT SỐ THÔNG SỐ TRONG PHÂN TÍCH NHÂN TỐ KHÁM PHÁ

- Kiểm định Barlett's test of sphericity: Kiểm định Barlett's là kiểm định thống kê nhằm xem xét giả thuyết các biến không có tương quan trong tổng thể. Điều kiện cần để áp dụng phân tích

nhân tố là các biến phải có tương quan với nhau, do vậy nếu kiểm định cho thấy không có ý nghĩa thống kê thì không nên áp dụng phân tích nhân tố cho các biến đang xem xét.

- Kaiser – Meyer – Olkin (KMO) measure of sampling adequacy: là chỉ số dùng để xem xét sự thích hợp của phân tích nhân tố. Giá trị KMO hay MSA nằm trong khoảng từ 0 đến 1, khi $KMO \geq 0.5$ thì có thể khẳng định dữ liệu là thích hợp để phân tích nhân tố. Đồng thời, giá trị trên đường chéo của **Anti-image Correlation** trong ma trận Anti-image (**Anti-image Matrices**) phải lớn hơn 0.5. Khi $KMO < 0.5$ hay tồn tại giá trị trên đường chéo nhỏ hơn 0.5 thì cần phải *lần lượt bỏ items tương ứng với giá trị nhỏ nhất trên đường chéo và lần lượt thực hiện* lại phân tích nhân tố khi nào giá trị này thỏa mãn.
- Correlation matrix: ma trận thể hiện hệ số tương quan giữa các cặp biến trong phân tích.
- Communality: là lượng biến thiên của một biến được giải thích chung với các biến khác được xem xét trong phân tích và giá trị này phải lớn hơn hoặc bằng 0.5 (Hair, Black, Babin, & Anderson, 2009).
- Total Variance Explained: Tổng phương sai giải thích được (thông thường, tổng phương sai được phải lớn hơn 50%).
- Eigenvalue: phần biến thiên được giải thích bởi mỗi nhân tố.
- Factor loading (hệ số tải nhân tố) là hệ số tương quan đơn giữa các biến (items) và các nhân tố (factors). Hệ số này được xác định tương ứng với kích thước mẫu (được giới thiệu cụ thể ở phần sau).
- Cross-loading: là hệ số tương quan của một chỉ báo (items) với nhân tố thứ 2 trở lên, hệ số cross-loading phải nhỏ hơn hệ số tải nhân tố tương ứng với mẫu nghiên cứu (được giới thiệu ở phần sau). Ví dụ: nếu mẫu là 200 thì cross-loading phải nhỏ hơn 0.4, nếu mẫu là 350 thì cross-loading phải nhỏ hơn 0.3...
- Component matrix (ma trận thành phần (nhân tố)): chứa các hệ số tải nhân tố của tất cả các biến với các nhân tố được rút trích.
- Rotated component matrix (ma trận thành phần (nhân tố) sau khi xoay): chứa các hệ số tải nhân tố của tất cả các biến với các nhân tố được rút trích sau khi thực hiện phép xoay.

TIẾN TRÌNH PHÂN TÍCH NHÂN TỐ

Có 7 bước cơ bản trong phân tích nhân tố khám phá: (1) Mục tiêu của phân tích nhân tố, (2) Thiết kế phân tích nhân tố, (3) Giả định phân tích nhân tố, (4) Rút trích và xác định số lượng nhân tố, (5) Giải thích các nhân tố, (6) Giá trị hóa của phân tích nhân tố, (7) Sử dụng kết quả phân tích nhân tố. Cụ thể, trong phân tích tương quan cần chú ý các vấn đề sau đây:

Tính ma trận tương quan

Phân tích nhân tố dựa trên sự tương quan giữa các biến đo được, một ma trận tương quan bao gồm hệ số tương quan của các biến phải được tính toán. Ma trận được thể hiện thông qua mối quan hệ giữa các biến tham gia vào quá trình phân tích và thể hiện cơ sở để thực hiện phân tích nhân tố.

Giả sử có các biến CV1, CV2, CV3, CV4 và TL1, TL2, TL3, TL4, ma trận tương quan được thiết lập như sau:

	CV1	CV2	CV3	CV4	TL1	TL2	TL3	TL4
CV1	1.000							
CV2	.659	1.000						
CV3	.682	.671	1.000					
CV4	.676	.408	.422	1.000				
TL1	.486	.275	.313	.332	1.000			
TL2	.351	.380	.360	.316	.539	1.000		
TL3	.213	.202	.284	.071	.400	.628	1.000	
TL4	.248	.275	.227	.279	.589	.673	.394	1.000

Trong ma trận tương quan, chúng ta dễ dàng nhận ra biến CV1 có mối quan hệ chặt với biến CV2, CV3, CV4 và TL2 có mối quan hệ chặt với TL3, TL4 nhưng TL1 cũng có mối quan hệ chặt với TL1, TL4. Như vậy, làm thế nào để giải tích các mối quan hệ này? Phép xoay nhân tố sẽ giúp cho chúng ta xác định một cách dễ dàng hơn, nó hướng đến sự đơn giản, ý nghĩa hơn cho các nhân tố.

Rút trích các nhân tố ban đầu (initial factors)

Trong bước này, cần rút trích các nhân tố từ các biến số ban đầu, để làm việc này các nhà nghiên cứu phải quyết định dùng (1) phương pháp rút trích và (2) số lượng nhân tố được rút trích (được chọn để trình bày dưới dạng cấu trúc của dữ liệu).

Phương pháp rút trích

Có 2 phương pháp cơ bản vẫn thường được sử dụng cho phép phân tích nhân tố, đó là phân tích **Principal Components** (thành phần chính) và là **Common Factor Analysis** (phân tích nhân tố chung), lưu ý rằng SPSS cung cấp 6 phương pháp để rút trích đối với mô hình phân tích nhân tố phổ biến; đó là principal-axis factoring, unweighted least-squares, generalized least-squares, maximum-likelihood, alpha factoring và image factoring. Nếu mục đích không ngoài việc “giảm dữ liệu” để thu được số nhân tố là bé nhất thì cần phải trình bày dữ liệu nguyên gốc của nó, khi đó sẽ thực hiện bởi phân tích thành phần chính (principal components). Ngược lại, khi mà các mục tiêu đầu tiên được xác định một cách lý thuyết về kích thước cơ bản, phương pháp phân tích nhân tố chung (common factor analysis) là mô hình thích hợp. Đa phần khi thực hiện phân tích nhân tố với mục tiêu giảm dữ liệu, do vậy phương pháp thành phần chính (principal components) thường được sử dụng rộng rãi hơn.

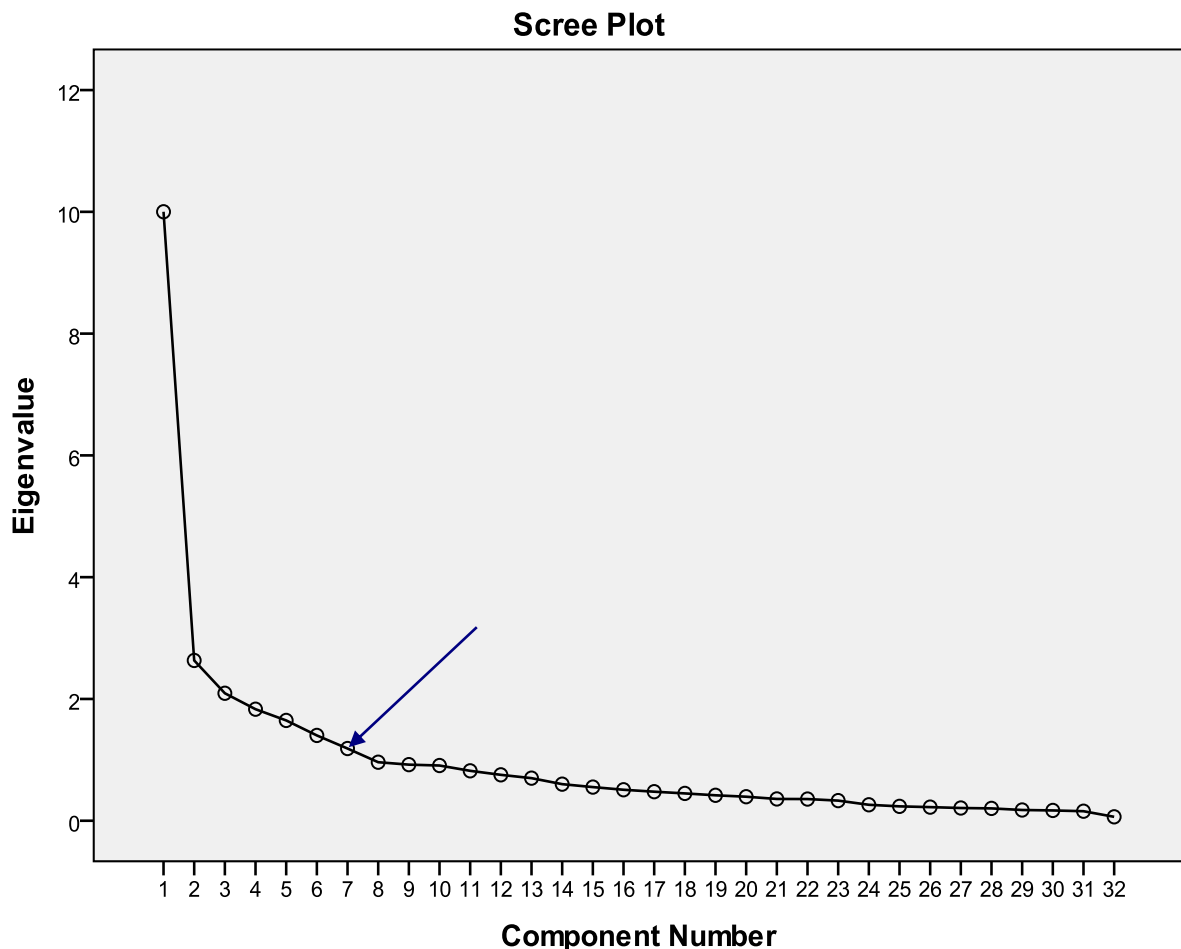
Xác định số lượng nhân tố rút trích

Có một số quy ước cho việc định nghĩa về số lượng các nhân tố được rút trích như phương pháp xác định từ trước (priori determination), dựa vào giá trị riêng eigenvalues (determination based on eigenvalue), phương pháp scree test và phần trăm phương sai biến thiên giải thích được (percentage of variance).

- Phương pháp xác định từ trước: Trong quá trình triển khai nghiên cứu, thông qua hợp tuyển lý thuyết hoặc từ các nghiên cứu trước, chúng ta đã xác định được k nhân tố được rút trích

(hay k biến số tiềm ẩn được hình thành) thì cần chỉ định số lượng nhân tố trước để phân mềm hình thành số lượng nhân tố như đã xác định.

- Phương pháp dựa vào giá trị riêng eigenvalue: Chỉ có những nhân tố có giá trị riêng (eigenvalue) bằng hoặc lớn hơn 1 mới được xem là có ý nghĩa, chúng ta không thực sự cần quan tâm đến các nhân tố có giá trị riêng bé hơn. Giá trị riêng (eigenvalue) là tỷ lệ giữa phương sai thông thường và phương sai riêng được giải thích bằng một nhân tố cụ thể được rút trích. Cơ sở cho việc sử dụng phương pháp giá trị riêng là số lượng các phương sai thông thường được giải thích bởi nhân tố được rút ra ít nhất phải tương đương với phương sai được giải thích bởi một biến đơn (phương sai duy nhất) nếu nhân tố được giữ lại cho việc giải thích. Giá trị riêng lớn hơn 1 cho biết phương sai thông thường hơn là phương sai duy nhất được giải thích bằng nhân tố đó.
- Phương pháp Scree test: Scree test được lấy từ biểu đồ của eigenvalue (trên trục Y) ngược lại số lượng các nhân tố được rút ra từ trục X. Các nhân tố đầu tiên được rút ra là các nhân tố lớn (với eigenvalue cao), tiếp theo là các nhân tố bé hơn. Bằng đồ thị, biểu đồ sẽ chỉ ra có một dốc lớn giữa các nhân tố lớn và một vệt dài gần như nằm ngang. Tại điểm mà đường cong bắt đầu thẳng được xem như là dấu hiệu để nhận biết số lớn nhất các nhân tố được rút trích. Như vậy, số thứ tự gắn liền với điểm trên điểm uốn cong này là số lượng nhân tố (chính) được rút trích thông qua phân tích nhân tố.



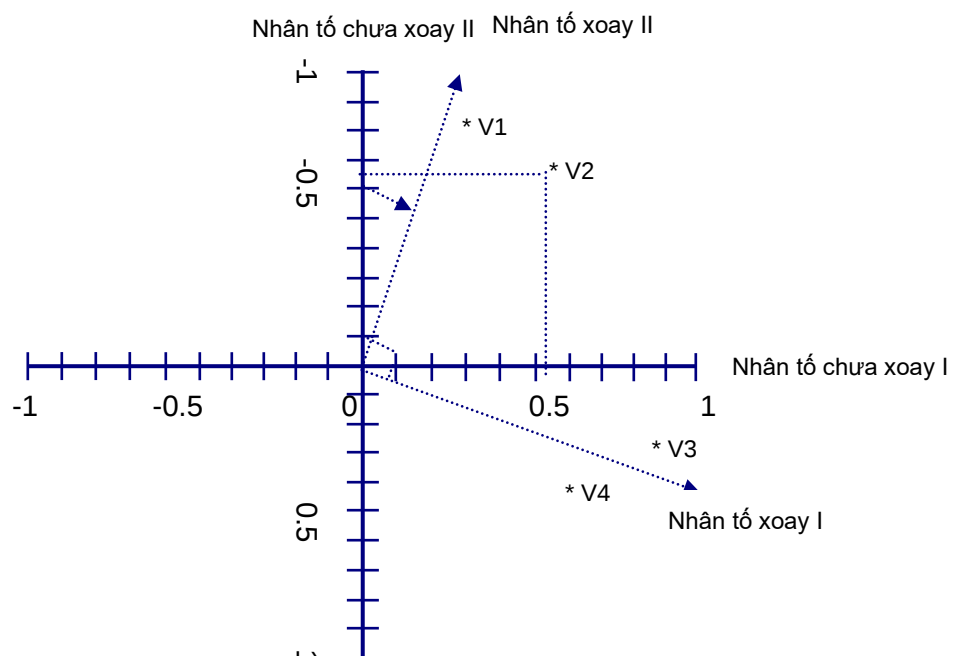
- Phần trăm phương sai biến thiên: Mục đích của phương pháp này nhằm đảm bảo ý nghĩa nghiên cứu của nhân tố gốc bởi chúng giải thích lượng xác định tối thiểu sự khác biệt. Không có một ngưỡng tuyệt đối trong tất cả ứng dụng, tuy nhiên, trong khoa học tự nhiên quy trình nhân tố thường không nên dừng lại cho đến khi nhân tố trích giải thích cho ít nhất 95% sự khác biệt còn các nhân tố còn lại giải thích cho chỉ 1 phần nhỏ (nhỏ hơn 5%). Ngược lại, trong khoa học xã hội khi thông tin ít rõ ràng, phần trăm phương sai giải thích thường không nhỏ hơn 0.6 (hay 60%).

Xoay nhân tố

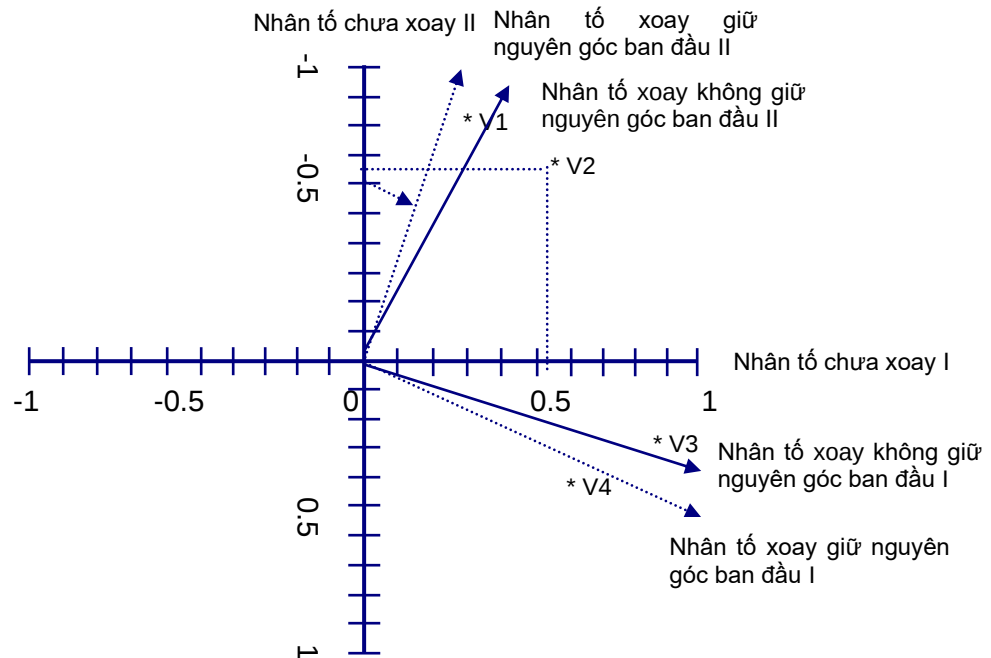
Phương pháp phân tích nhân tố sẽ hình thành ma trận nhân tố (Component Matrix), ma trận này chứa các hệ số (factor loading) biểu diễn mối quan hệ giữa các nhân tố các các biến và hệ số này càng lớn thì mối quan hệ này càng chặt chẽ.

Có 2 cách sắp xếp chính của phương pháp quay tạt gọi là phương pháp trực giao (**orthogonal**) và xiên (**oblique**). Phép quay **Orthogonal** thực hiện trên cơ sở giả định các nhân tố độc lập và thực hiện phép quay 90 độ. **Oblique** cho phép các nhân tố tương quan thay vì duy trì sự độc lập giữa các nhân tố xoay, quy trình của **Oblique** không đòi hỏi trục phải duy trì ở 90 độ.

Trong việc chọn lựa giữa **Orthogonal** và **Oblique rotation**, không có lý do nào thuyết phục trong việc thiên về dùng phương pháp này hơn là phương pháp kia. Thực vậy, không có quy luật để hướng dẫn các nhà nghiên cứu lựa chọn cụ thể phương pháp **Orthogonal** hay **Oblique rotation**. Tuy nhiên, quy ước được đưa ra là theo các nguyên tắc có thể giúp cho việc chọn lọc. Nếu mục đích của nghiên cứu không nhiều hơn ngoài việc giảm dữ, không chú ý đến kết quả các nhân tố có thể có ý nghĩa như thế nào và nếu đó là lý do để thừa nhận rằng các nhân tố không tương quan, khi đó **Orthogonal rotation** nên được dùng. Ngược lại, nếu mục đích của việc nghiên cứu là tìm ra các nhân tố có nghĩa một cách lý thuyết và nếu đó là lý do để thừa nhận rằng nhân tố đó tương quan, khi đó dùng **Oblique rotation** thích hợp.



Phép xoay Orthogonal: xoay giữ nguyên góc ban đầu



Phép xoay Oblique: xoay không giữ nguyên góc ban đầu

Đôi khi các nhà nghiên cứu có thể không biết có hay không việc các nhân tố rút ra có thể tương quan không. Trong trường hợp này các nhà nghiên cứu nên dùng giải pháp **oblique** trước. Đề nghị này dựa trên giả định rằng có rất ít các biến trong một dự án nghiên cứu cụ thể sẽ không tương quan. Nếu sự tương quan giữa các nhân tố là rất thấp (<0.20) nhà nghiên cứu có thể thực hiện lại phân tích với phương pháp **Orthogonal rotation**

Có 3 phương pháp chính của **Orthogonal rotation** là:

- Phép quay **varimax**: xoay nguyên góc các nhân tố để tối thiểu số lượng các biến có hệ số lớn tại cùng một nhân tố, điều này giúp cho việc dễ dàng giải thích các biến số quan sát có quan hệ chặt chẽ với một nhân tố.
- Phép quay **quartimax**: xoay nguyên góc các nhân tố để tối thiểu số nhân tố có hệ số lớn tại cùng một biến.
- Phép quay **equimax**: xoay các nhân tố để đơn giản việc giải tích cả biến lẫn nhân tố.

Trong 3 phương pháp này, **varimax** được sử dụng rộng rãi nhất vì nó chỉ ra sự chia cắt rõ nhất của các nhân tố. Phương pháp này thực hiện bằng cách đơn giản hóa số cột lớn nhất có thể bên trong các nhân tố ma trận trong khi đó cả **quartimax** và **equimax** đều không được sử dụng rộng rãi.

Số lần thực hiện nhân tố

Nên chú ý rằng phân tích nhân tố dùng cho nghiên cứu (kể cả cho mục đích giảm bớt dữ liệu hoặc nhận dạng một cách lý thuyết các phương diện (dimensions) có ý nghĩa), có ít nhất 2 cách thực hiện. Trong lần đầu, nhà nghiên cứu cho phép phân tích nhân tố rút trích các nhân tố với việc xoay nhân tố. Tất cả các nhân tố với eigenvalues lớn hơn hay bằng 1 thông qua phép quay varimax được thực hiện trong SPSS. Tuy nhiên, trong một số trường hợp, ngay cả sau khi xoay, không phải tất cả nhân tố rút trích đều có nghĩa. Ví dụ: một vài nhân tố chỉ có đại diện bởi ít

quá items hoặc nó vẫn có thể có một vài chỉ báo (items) có hệ số tải lớn trên nhiều hơn một nhân tố (cross-loading). Trong giai đoạn này, nhà nghiên cứu phải quyết định nhân tố nào có ý nghĩa (dựa trên cả lý thuyết hoặc bằng trực giác) và có giữ lại những nhân tố này cho lần thực hiện tiếp theo. Cũng không phải là hiếm thấy dữ liệu nghiên cứu phải thực hiện một chuỗi phân tích nhân tố và phép xoay trước khi thu được các nhân tố có thể được xem như là “sạch” và có thể diễn giải được.

Xác định mức ý nghĩa của hệ tải nhân tố dựa trên độ lớn của mẫu

Như chúng ta đã biết, hệ số tải nhân tố (factor loading) tượng trưng cho sự tương quan giữa một biến ban đầu và yếu tố của nó. Khi xác định một mức độ ý nghĩa cho việc giải thích của hệ số tải nhân tố, một cách tiếp cận tương tự để xác định ý nghĩa thống kê của hệ số tương quan có thể được sử dụng. Ở phần trước đã giới thiệu một nghiên tắc chung giản đơn cho việc xác định mẫu nghiên cứu là theo quy tắc 5 lần số biến quan sát (items). Để xác định hệ số tải nhân tố, Hair và cộng sự (2009) đã cho rằng hệ số tải nhân tố được xác định theo kích cỡ mẫu và nếu hệ số tải càng nhỏ thì kích thước mẫu nghiên cứu càng lớn. Ví dụ, trong một mẫu của 100 người trả lời, hệ số tải nhân tố là 0,55, tuy nhiên, trong một mẫu số 50, một yếu tố tải 0,75 là cần thiết. Thường người ta lựa chọn và sử dụng hệ số tải nhân tố lớn hơn từ ± 0.3 đến ± 0.4 là chấp nhận được, lớn hơn ± 0.5 là có ý nghĩa thực tế nhưng theo đề nghị Hair và cộng sự (2009), hệ số tải nhân tố cần được xác định theo kích thước mẫu cho ở biểu sau:

Hệ số tải nhân tố	Kích thước mẫu
0.30	350
0.35	250
0.40	200
0.45	150
0.50	120
0.55	100
0.60	85
0.65	70
0.70	60
0.75	50

Nguồn: Hair và cộng sự (2009)

Đặt tên các nhân tố

Như chúng ta đã đề cập ở trên, phân tích nhân tố cung cấp cho người nghiên cứu với 2 khả năng: tổng hợp dữ liệu và cắt giảm dữ liệu. Để dễ dàng cho việc thực hiện những công việc liên quan tiếp theo, chúng ta phải đặt tên của các nhân tố, tên của các nhân tố phải đại diện cho nội dung của các chỉ báo đo lường (đại diện) cho chúng. Như ví dụ đã đề cập ở phần trên, chúng ta có những biến riêng lẻ (items) là “tôi mua hàng đặc biệt”, “tôi thường quan tâm đến giá thấp

nhất”, “tôi mua hàng thương lượng”, “nhãn hiệu quốc gia đáng giá hơn nhãn hiệu của cửa hàng”, tập hợp lại, những biến này có thể được sử dụng để xác định khách hàng “biết rõ giá” hoặc là “chuyên gia thương lượng mua hàng” hay còn gọi đây là biến số tiềm ẩn (latent variable).

Nhân số (factor score)

Giả sử có n biến số quan sát tham gia vào quá trình phân tích nhân tố, bằng cách sử dụng phương pháp phân tích nhân tố, giả sử phân tích nhân tố rút ra được i nhân tố (factors), ta có:

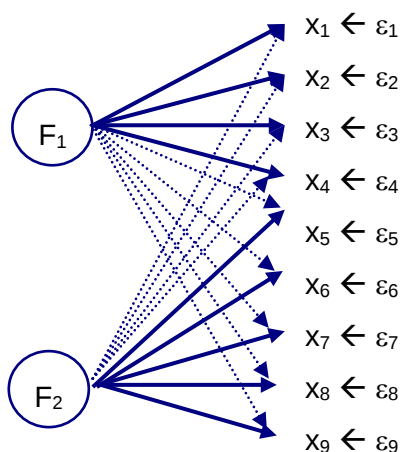
$$F_i = W_{i1}X_1 + W_{i2}X_2 + W_{i3}X_3 + \dots + W_{in}X_n$$

Với: F_i là ước lượng trị số của nhân tố (factor) thứ i, W_{ik} là quyền số hay trọng số nhân tố (weight or factor score coefficient) của biến số thứ k đến nhân tố i và n là số biến số (items).

Diễn giải các nhân tố

Trong quá trình phân tích nhân tố, kết quả phân tích được sử dụng để diễn giải các nhân tố, cần chú ý một số nguyên tắc sau:

- Mỗi biến số quan sát (items) chỉ có hệ số tải nhân tố lớn với duy nhất một nhân tố.



Trong hình trên có hai nhân tố F_1 và F_2 được rút trích từ 9 biến số quan sát (items), các biến số quan sát x_1, x_2, x_3, x_4 có mối quan hệ chặt chẽ với nhân tố F_1 và các biến số x_5, x_6, x_7, x_8, x_9 có mối quan hệ chặt chẽ với nhân tố F_2 .

- Nếu xảy ra cross-loading (một items có hệ số tải nhân tố lớn hơn hoặc bằng 2 nhân tố) sẽ bị xóa.
- Thông thường, các biến phải có hệ số communalities lớn hơn hoặc bằng 0.5.

THỰC HIỆN VÀ DIỄN GIẢI KẾT QUẢ PHÂN TÍCH NHÂN TỐ KHÁM PHÁ TRONG SPSS

Giả sử cần nghiên cứu các nhân tố tác động đến lòng trung thành của nhân viên đối với tổ chức, thực hiện nghiên cứu lý thuyết 7 biến số độc lập và một biến số phụ thuộc với các chỉ báo cho ở bảng sau:

Công việc	CV1	Công việc cho phép anh/chị sử dụng tốt các năng lực cá nhân
	CV2	Công việc rất thú vị
	CV3	Công việc có nhiều thách thức
	CV4	Khi công việc hoàn thành tốt, sẽ được công ty rất hoan nghênh

Cảm nhận về tiền lương, thưởng	TL1	Anh/ chị được trả lương cao
	TL2	Anh/ chị thường được tăng lương
	TL3	Anh/ chị có thể sống hoàn toàn dựa vào thu nhập từ công ty
	TL4	Tiền lương tương xứng với kết quả làm việc
	TL5	Tiền lương, thu nhập được trả công bằng
Phúc lợi	PL1	Công ty có chế độ phúc lợi tốt
	PL2	Công ty thực hiện chế độ bảo hiểm xã hội tốt
	PL3	Công ty thực hiện chế độ bảo hiểm y tế tốt
	PL4	Anh chị hài lòng với chế độ tiền thưởng của công ty
Môi trường làm việc	MT1	Công việc không bị áp lực cao
	MT2	Công việc không đòi hỏi thường xuyên phải làm ngoài giờ
	MT3	Trang thiết bị nơi làm việc an toàn, sạch sẽ
	MT4	Công việc ổn định, không phải lo lắng về mất việc làm
Đồng nghiệp	DN1	Đồng nghiệp của anh/chị thoải mái và dễ chịu
	DN2	Anh/ chị và các đồng nghiệp phối hợp làm việc tốt
	DN3	Những người mà anh/ chị làm việc với thường giúp đỡ lẫn nhau
	DN4	Anh/ chị cảm thấy có nhiều động lực trau dồi chuyên môn khi được làm việc với các đồng nghiệp của mình
Phong cách lãnh đạo	LD1	Anh/ chị được biết những điều kiện để được thăng tiến
	LD2	Công ty tạo cho anh chị nhiều cơ hội thăng tiến
	LD3	Anh/ chị được cung cấp kiến thức/ kỹ năng cần thiết cho công việc
	LD4	Công ty tạo cho anh/chị nhiều cơ hội phát triển cá nhân
Chính sách đào tạo	DT1	Cấp trên hỏi ý kiến anh/chị khi có vấn đề liên quan đến công việc anh/chị
	DT2	Cấp trên khuyến khích cấp dưới tham gia vào những quyết định quan trọng
	DT3	Nhân viên được sự hỗ trợ của cấp trên
	DT4	Công ty hoạt động có hiệu quả
	DT5	Nhân viên được tôn trọng và tin cậy trong công việc
	DT6	Lãnh đạo có tác phong lịch sự, hòa nhã
	DT7	Nhân viên được đối xử công bằng, không phân biệt
Lòng trung thành (biến phụ thuộc)	LTT1	Anh/chị có ý định ở lại lâu dài cùng công ty
	LTT2	Nếu có nơi khác có lời đề nghị lương bổng tương đối hấp dẫn hơn, Anh/Chị vẫn sẽ ở lại cùng công ty
	LTT3	Về nhiều phương diện, anh chị coi công ty là mái nhà thứ hai của mình

Các biến số được đo lường bởi thang đo Likert (7 sự lựa chọn) với (1) là rất không đồng ý và (7) là rất đồng ý, thực hiện khảo sát trên 200 mẫu là cán bộ công nhân viên của doanh nghiệp và kết quả được sử dụng để phân tích nhân tố.

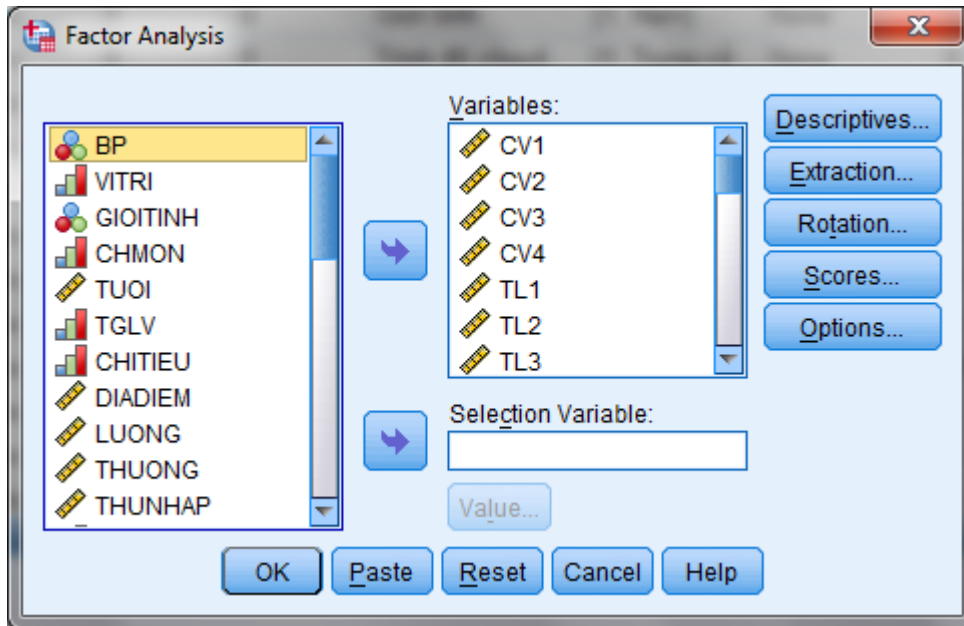
Mục tiêu của phân tích nhân tố

- Mục tiêu của phân tích nhân tố nhằm giảm số lượng lớn biến quan sát (32 items) từ biến CV1 đến biến DT7 thành một tập biến nhỏ hơn (nhân tố).

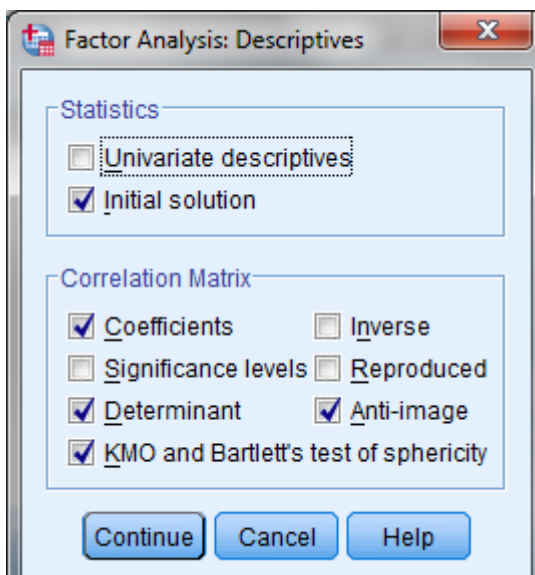
- Kích thước mẫu là 200, đảm bảo yêu cầu theo tỉ lệ 5:1.

Thực hiện phân tích nhân tố trên SPSS

- Từ thanh thực đơn, kích vào **Analyze / Dimension Reduction / Factors**, của số **Factor Analysis** mở ra.
- Chọn tất cả các biến muốn rút gọn dữ liệu vào ô **Variables**.



- Kích vào **Descriptives**, chọn các mục chọn như hình vẽ và kích **Continue**.

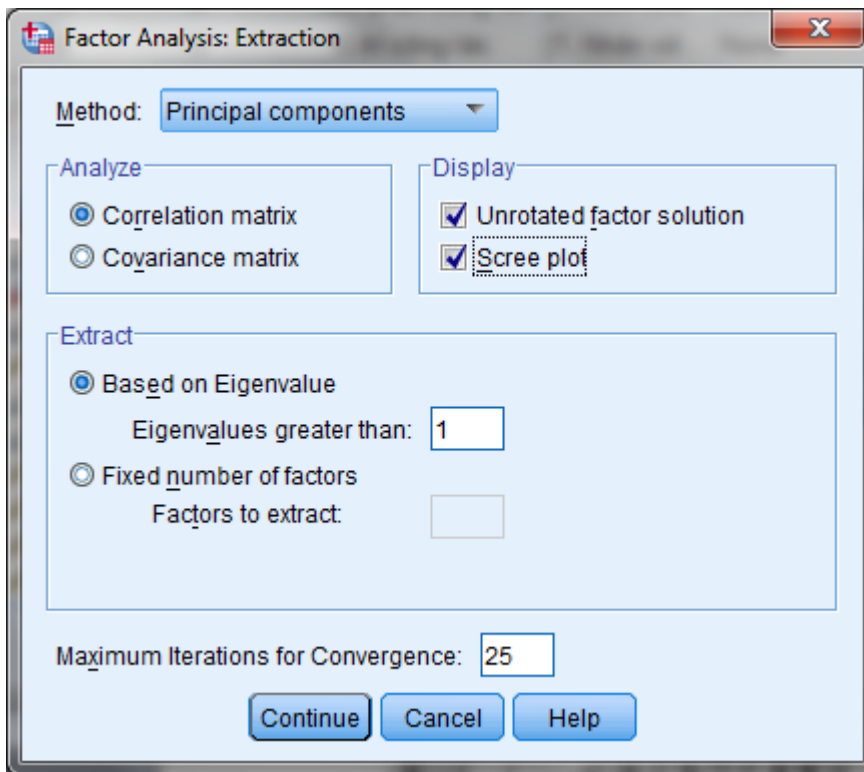


- Kích vào **Extraction**, mục tiêu phân tích nhân tố ở đây là rút trích dữ liệu nên ta lựa chọn phương pháp **Principal components** và đây là phương pháp trích được SPSS mặc định.

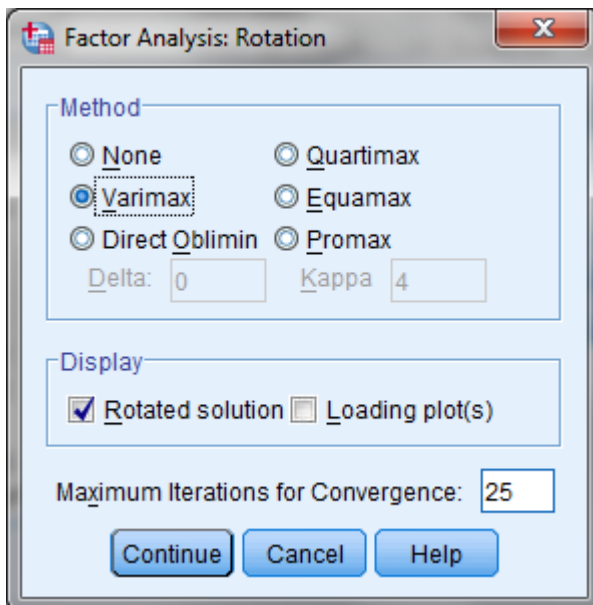
Chú ý:

+ Phần mềm sẽ sử dụng rút trích (extract) dựa trên giá trị riêng Eigenvalue ≥ 1 nên mặc định phương pháp **Based on Eigenvalue** và **Eigenvalues greater than 1**.

+ Nếu lựa chọn phương pháp dựa trên số lượng nhân tố đã xác định trước thì kích vào **Fixed number of factors** và đánh số nhân tố đã xác định trước thông quan nghiên cứu lý thuyết vào ô **Factors to extract**.

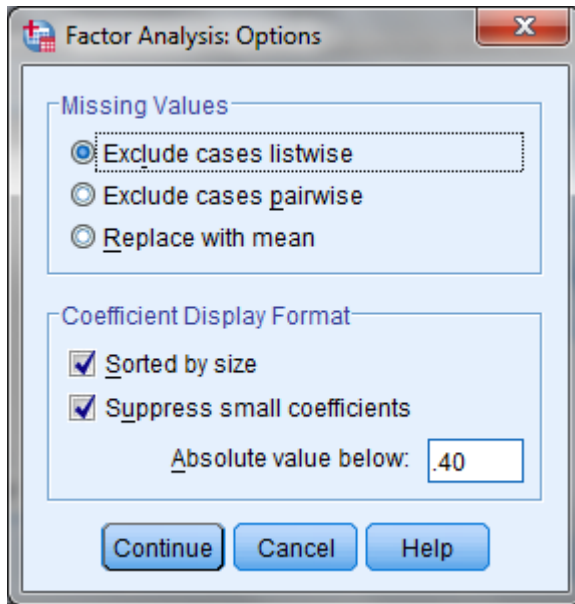


- Kích vào **Rotation** và lựa chọn phương pháp quay là **Varimax**:



- Vì kích thước mẫu $n = 200$ nên lựa chọn hệ số tải nhân tố của các items phải lớn hơn hoặc bằng 0 (Hair và cộng sự, 2009).
- Kích vào **Option**, lựa chọn **Sorted by size** để sắp xếp từ lớn đến nhỏ hệ số tải nhân tố và bỏ những hệ số tải có giá trị nhỏ hơn 0.4 vì mẫu $n=200$ (theo Hair và cộng sự, 2009), khi đó,

hộp hội thoại **Factor Analysis: Options** được mở ra.



- Kích **OK** để hoàn tất việc thiết đặt thông số phân tích và phân tích kết quả.

Phân tích và diễn giải kết quả

- Kiểm định Bartlett và đo lường sự thích hợp của dữ liệu (MSA):

Kết quả phân tích cho thấy, giá trị KMO = 0.863 > 0.5 và kiểm định Bartlett có Chi-square = 3614.823, df = 496 nên p (chi-square, df) = 0.000 < 0.05 và đo lường sự tương thích của dữ liệu (MSA) với tất cả giá trị trên đường chéo đều lớn hơn 0.5 nên khẳng định dữ liệu là thích hợp để phân tích nhân tố.

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.863
Bartlett's Test of Sphericity	Approx. Chi-Square	3614.823
	df	496
	Sig.	.000

Vấn đề đặt ra: Nếu KMO < 0.5 hoặc giá trị MSA **trên đường chéo** của **Anti-image Correlation** trong ma trận Anti-image (**Anti-image Matrices**) nhỏ hơn 0.5 nên có thể khẳng định dữ liệu chưa thích hợp để phân tích nhân tố.

Vì vậy cần phải cải thiện dữ liệu bằng cách thực hiện lại phân tích nhân tố nhưng *lần lượt* bỏ items có giá trị **trên đường chéo** của **Anti-image Correlation** trong ma trận Anti-image (**Anti-image Matrices**) có giá trị nhỏ hơn 0.5, cần nhớ rằng chúng ta phải thực hiện bỏ lần lượt với giá trị trên đường chéo đã nhắc ở trên càng nhỏ hơn 0.5 cần phải bỏ trước. Chúng ta thực hiện đến khi nào KMO > 0.5 và MSA có giá trị lớn hơn 0.5 nghĩa là giá trị thích hợp để phân tích nhân tố.

- Số lượng nhân tố rút trích:

Trong biểu thể hiện tổng phương sai trích cho thấy có thể rút trích từ 32 items thành 7 nhân tố có giá trị riêng eigenvalue lớn hơn hoặc bằng 1 (ở đây nghiên cứu sử dụng nguyên tắc rút trích với giá trị riêng lớn hơn hoặc bằng 1), với phương sai trích tích lũy bằng 64.979% thỏa

mãn điều kiện (vì phương sai tích lũy >50%).

Component	Total Variance Explained		
	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	10.001	31.253	31.253
2	2.633	8.227	39.480
3	2.093	6.542	46.022
4	1.832	5.724	51.747
5	1.647	5.148	56.894
6	1.401	4.378	61.273
7	1.186	3.706	64.979
8	.961	3.004	67.982
9	.921	2.879	70.861
10	.907	2.834	73.695
11	.820	2.562	76.257
12	.754	2.356	78.613
13	.699	2.185	80.798
14	.601	1.877	82.675
15	.553	1.727	84.402
16	.509	1.590	85.992
17	.477	1.491	87.483
18	.449	1.402	88.885
19	.418	1.306	90.191
20	.395	1.233	91.424
21	.358	1.118	92.542
22	.356	1.113	93.655
23	.331	1.033	94.688
24	.262	.819	95.507
25	.236	.738	96.244
26	.224	.699	96.943
27	.209	.654	97.597
28	.203	.633	98.230
29	.177	.552	98.782
30	.169	.528	99.310
31	.156	.489	99.799
32	.064	.201	100.000

Extraction Method: Principal Component Analysis.

- Ứng dụng phép quay Orthogonal (Varimax):

Với phép quay varimax, kích thước mẫu $n = 200$, với việc thiết đặt các thông số như phân hướng dẫn SPSS cho phân tích nhân tố, kết quả đã ẩn các hệ số tải nhân tố nhỏ hơn 0.4 và sắp xếp theo thứ tự giảm dần của hệ số tải. Cần phải kiểm tra và hệ số communalities (nhỏ hơn 0.5), các giá trị hệ số tải nhân tố (nhỏ hơn 0.4), các cross-loading (nhỏ hơn 0.4). Lần

lượt thực hiện việc loại bỏ các items thỏa mãn điều kiện, kết quả cho thấy, so với kết quả ở biểu **Total Variance Explained** trước có 7 nhân tố được rút trích thì ở đây có 6 nhân tố được rút trích. Lúc này 6 nhân tố giải thích được 69.97% tổng phương sai của 20 items sau khi loại các biến không thỏa điều kiện.

Total Variance Explained

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	6.224	31.122	31.122
2	1.968	9.839	40.961
3	1.847	9.236	50.196
4	1.560	7.799	57.996
5	1.247	6.237	64.233
6	1.147	5.737	69.970
7	.844	4.222	74.192
8	.766	3.830	78.022
...	...	...	...
20	.071	.354	100.000

Extraction Method: Principal Component Analysis.

Kết quả của ma trận thành phần sau khi xoay với tất cả các items thỏa mãn điều kiện hệ số tải nhân tố lớn hơn 0.4 và không xảy ra hiện tượng cross-loading.

Rotated Component Matrix^a

	Component					
	1	2	3	4	5	6
DT6	.801					
DT5	.788					
DT3	.761					
DT7	.675					
DT4	.651					
TL4		.802				
TL5		.731				
TL2		.712				
TL1		.686				
TL3		.664				
LD1			.821			
LD2			.769			
LD3			.707			
LD4			.701			
CV2			.446			
DN1				.872		
DN2				.830		
DN3				.813		
PL3					.952	

PL2					.940	
MT1						.794
MT2						.773

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 6 iterations.

- Đặt tên các biến:

- Nhân tố F1: Chính sách đào tạo (được đo lường bởi các items (DT3, DT4, DT5, DT6, DT7)
- Nhân tố F2: Cảm nhận về tiền lương, thưởng (được đo lường bởi các items TL1, TL2, TL3, TL4, TL5)
- Nhân tố F3: Phong cách lãnh đạo (được đo lường bởi các items LD1, LD2, LD3, LD4, CV2)
- Nhân tố F4: Đồng nghiệp (được đo lường bởi các items DN1, DN2, DN3)
- Nhân tố F5: Phúc lợi (được đo lường bởi các items PL2, PL3)
- Nhân tố F6: Môi trường (được đo lường bởi các items MT1, MT2)

Thực hiện tương tự với các chỉ báo (LTT1, LTT2, LTT3) của Lòng trung thành đối với tổ chức, kết quả như sau:

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.701
Bartlett's Test of Sphericity	Approx. Chi-Square	188.787
	Df	3
	Sig.	.000

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.146	71.523	71.523	2.146	71.523	71.523
2	.487	16.230	87.753			
3	.367	12.247	100.000			

Extraction Method: Principal Component Analysis.

Dễ dàng nhận thấy rằng, kết quả phân tích nhân tố khám phá với $KMO = 0.701 > 0.5$ và kiểm định Bartlett với $p(\text{Chi-square}, df) = 0.000 < 0.05$ nên khẳng định dữ liệu phù hợp để phân tích nhân tố.

Phân tích cũng đã rút trích từ 3 chỉ báo thành một nhân tố chính có Eigenvalue = 2.146 và tổng phương sai trích tích lũy là $71.523 > 0.05$ với tất cả hệ số tải nhân tố đều lớn hơn 0.4, ma trận thành phần được giới thiệu tại biểu component matrix.

Component Matrix^a

	Component
--	-----------

	1
LTT1	.874
LTT3	.834
LTT2	.829

Extraction Method: Principal Component Analysis.

a. 1 components extracted.

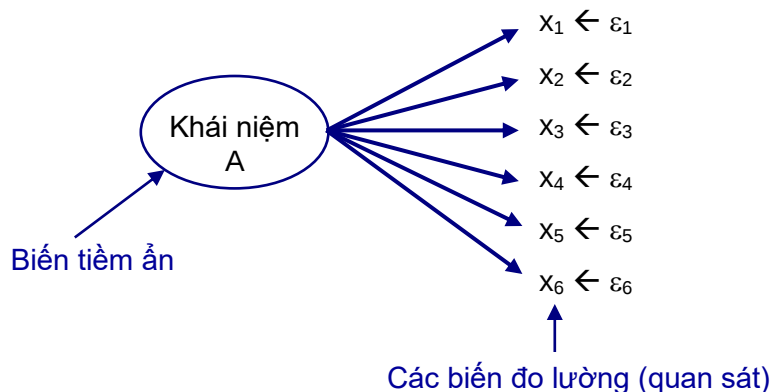
CHƯƠNG

9.2

ĐÁNH GIÁ ĐỘ TIN CẬY CỦA THANG ĐO (RELIABILITY)

Như đã đề cập ở các chương trước, khi thực hiện nghiên cứu định lượng cần phải lượng hóa các khái niệm cần phải đo lường, có những khái niệm rất dễ đo lường như độ tuổi, số năm học, cân nặng, giới tính... Tuy nhiên, có nhiều khái niệm rất khó để lượng hóa như những khái niệm liên quan đến hành vi, thái độ... của một con người. Những khái niệm này không thể đo lường đơn giản bởi một chỉ báo (items) mà đo lường thông qua nhiều chỉ báo mới đo lường chính xác nội dung của khái niệm – những biến đo lường khái niệm như trên được gọi biến tiềm ẩn (latent variable).

Một biến tiềm ẩn cần được đo lường bằng nhiều biến quan sát (gọi là thang đo), các biến quan sát này cùng đo lường một biến tiềm ẩn – vì vậy chúng phải có tương quan với nhau.



Độ tin cậy (reliability) của công cụ đo lường được định nghĩa như là khả năng đo lường sự phù hợp các hiện tượng mà nó được thiết kế để đo lường. Do đó, độ tin cậy liên quan đến việc kiểm định tính nhất quán (test consistency) của thang đo để khẳng định thang đo có thể đo lường đúng khái niệm cần đo lường.

Tầm quan trọng của độ tin cậy ở việc nó là điều kiện tiên quyết về giá trị của kiểm định. Đơn giản như, để giá trị của một công cụ đo lường được xác nhận thì nó phải đáng tin cậy, bất kỳ công cụ đo lường nào tồn tại một vài thuộc tính phù hợp thì sẽ khó được xem là một thuộc tính đo lường có giá trị.

Có một vài phương pháp để xác định độ tin cậy của một công cụ đo lường, những phương pháp này có thể được chia thành hai loại:

- Tính nhất quán bên ngoài phù hợp bên ngoài (external consistency)
- Tính nhất quán bên trong (internal consistency).

TÍNH NHẤT QUÁN BÊN NGOÀI (EXTERNAL CONSISTENCY)

Quy trình phù hợp bên ngoài lấy các kết quả của kiểm định tích lũy cho so sánh, tương phản với chính chúng để kiểm tra độ tin cậy của đo lường. Hai phương pháp chính để xác định độ tin cậy của một kiểm định bằng mức phù hợp bên ngoài là:

- Kiểm định - tái kiểm định (Test – retest): Kết quả của kiểm định cho một nhóm người được so sánh trong hai thời điểm khác nhau.
- Kiểm định song song: So sánh hai tập kết quả tương đương nhưng từ các kiểm định khác nhau.

TÍNH NHẤT QUÁN BÊN TRONG (INTERNAL CONSISTENCY)

Tính nhất quán bên trong liên quan đến nội dung mà các biến quan sát (items) trong một thang đo lường đo lường một cấu trúc hay khái niệm nghiên cứu (construct). Các chỉ báo cùng đo lường một hiện tượng nên phải liên kết với nhau một cách logic. Kiểm tra tính nhất quán bên trong của kiểm định cho phép người nghiên cứu xác định những biến nào không nhất quán (phù hợp) với kiểm định trong việc đo lường hiện tượng được điều tra. Đối tượng sẽ loại bỏ những biến không nhất quán và gia tăng sự nhất quán bên trong. Kiểm định tính nhất quán bên trong sẽ gia tăng xác suất đáng tin cậy của kiểm định.

Ba phương pháp xác định tính nhất quán bên trong là:

- Kỹ thuật phân đôi (**split-half technique**): Phương pháp này cho một nửa các biến kiểm định tương quan với nửa kia. Sự tương quan càng cao thì đo lường bên trong càng phù hợp. Phương pháp Spearman-Brown được sử dụng rộng rãi trong việc xác định độ tin cậy bằng phương pháp phân đôi.
- Phân tích chỉ báo (**item analysis**): Quy trình này trình bày sự “tinh lọc” của kiểm định độ tin cậy bằng cách nhận dạng vấn đề của các chỉ báo trong kiểm định, tức là, những biến này mang lại sự tương quan yếu so với toàn bộ các kết quả của những biến còn lại. Loại bỏ những biến này, những biến không phù hợp với các biến còn lại (nghĩa là giữ lại những biến có mức tương quan ít nhất từ trung bình trở lên với các biến khác), sẽ gia tăng mức độ phù hợp bên trong của các công cụ đo lường. Phân tích chỉ báo được thực hiện thông qua quy trình mối quan hệ tương quan biến tổng (Item-Total correlation).
- Hệ số **Cronbach's alpha**: Đây là một hệ số tương quan đơn, dùng để ước lượng mức trung bình của tất cả các hệ số tương quan của các biến trong kiểm định. Theo quy ước, một tập hợp các mục hỏi dùng để đo lường được đánh giá là tốt phải có hệ số Cronbach's Alpha nằm lân cận 0.8. Nếu giá trị Alpha thấp thì có ít nhất một biến không đáng tin cậy và biến đó phải được nhận dạng thông qua quy trình phân tích.

Hệ số Cronbach's Alpha được tính theo công thức:

$$\alpha = \frac{k}{k-1} \left[1 - \frac{\sum \sigma^2(x_i)}{\sigma_x^2} \right]$$

Trong đó: α = Cronbach's Alpha
 k = Số lượng biến trong thang đo
 $\sum \sigma^2(x_i)$ = Tổng phương sai các biến
 σ_x^2 = Phương sai tổng thang đo

Tùy theo bối cảnh nghiên cứu là hoàn toàn mới hoặc mới trong bối cảnh nghiên cứu mà nhà nghiên cứu quyết định hệ số Alpha phải lớn hơn 0.6; 0.7 hoặc 0.8.

Theo Nunnally (1978), Peterson (1994) và Slate (1995), hệ số Alpha được xem xét trong các trường hợp sau:

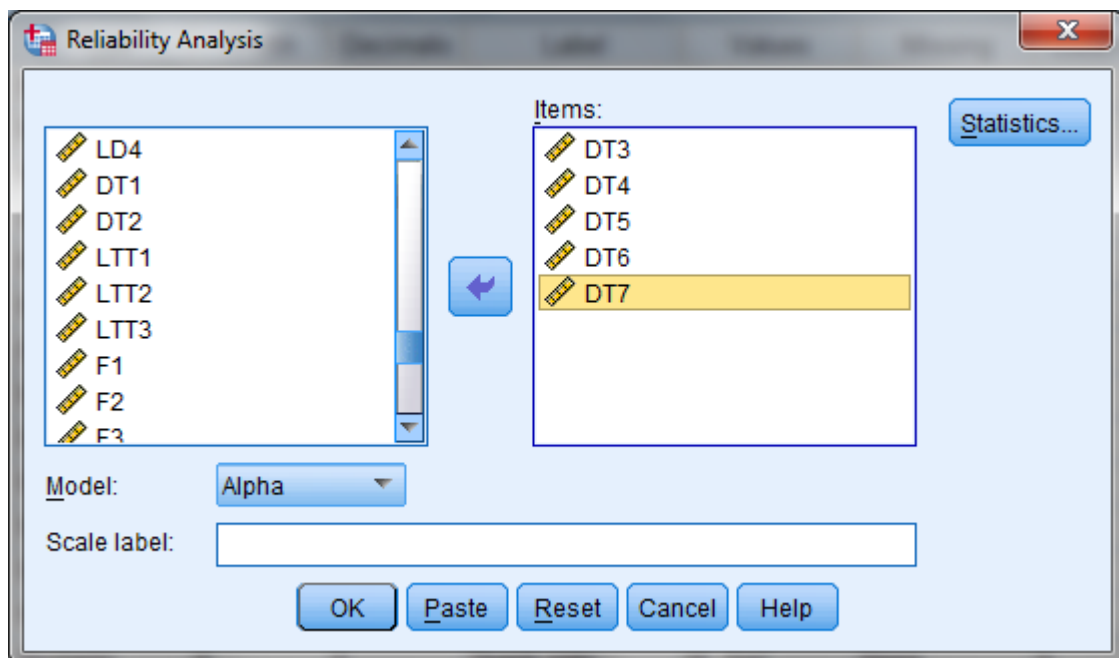
- $\alpha \geq [0.60 - 0.70]$: chấp nhận được trong trường hợp nghiên cứu hoàn toàn mới hoặc mới trong bối cảnh nghiên cứu.
- $\alpha \geq [0.70 - 0.80]$: chấp nhận được.
- $\alpha \in [0.70 - 0.90]$: tốt
- $\alpha > [0.90 - 1.0]$: chấp nhận được – không tốt

TÍNH HỆ SỐ CRONBACH'S ALPHA TRÊN SPSS

Để tính tính nhất quán bên trong (internal consistency) bằng hệ số Cronbach's Alpha cần thực hiện trên các chỉ báo của từng biến. Qua kết quả nghiên cứu lý thuyết, biến Chính sách đào tạo được đo lường bởi 7 chỉ báo (từ DT1 đến DT7), tuy nhiên, kết quả của phân tích nhân tố ở trong phần trước đã khẳng định biến Chính sách đào tạo đo lường bởi các items (DT3, DT4, DT5, DT6, DT7), các chỉ báo này được sử dụng để kiểm tra tính nhất quán nội tại của thang đo.

Ví dụ 1: Phân tích tính nhất quán nội tại đối với biến Chính sách đào tạo

- Từ thanh công cụ, chọn **Analyze / Scale / Reliability Analysis**, cửa sổ **Reliability Analysis** sẽ mở ra.
- Chọn các chỉ báo DT3, DT4, DT5, DT6, DT7 vào ô **Items**.



- Chọn vào các ô **Item**, **Scale**, **Scale if item deleted** và kích **Continue** để trở về cửa sổ **Reliability Analysis**.
- Sau đó bấm **OK** để hoàn tất thiết đặt và tiến hành phân tích.

- Kết quả phân tích và diễn giải dữ liệu:

Case Processing Summary			
		N	%
Cases	Valid	200	100.0
	Excluded ^a	0	.0
	Total	200	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics	
Cronbach's Alpha	N of Items
.860	5

Trong toàn bộ mẫu bao gồm 397 đối tượng đều tham gia vào quá trình phân tích (không có đối tượng nào bị loại vì thiếu dữ liệu).

Trong biểu Reliability Statistics cho thấy hệ số Cronbach's Alpha = 0.860, điều đó cho phép khẳng định rằng mức độ nhất quán bên trong giữa năm chỉ báo của biến Chính sách đào tạo là cao.

Item Statistics			
	Mean	Std. Deviation	N

DT3	4.76	1.407	200
DT4	5.38	1.274	200
DT5	5.01	1.402	200
DT6	5.28	1.453	200
DT7	4.64	1.598	200

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
DT3	20.31	21.391	.728	.819
DT4	19.70	24.163	.563	.858
DT5	20.06	21.343	.736	.817
DT6	19.80	21.340	.699	.826
DT7	20.44	20.529	.674	.835

Hệ số tương quan biến tổng Corrected Item-Total Correlation cho thấy sự tương quan (phù hợp) giữa mỗi mục với toàn bộ các mục còn lại. Điều kiện để một chỉ báo được giữ lại nếu hệ số tương quan biến tổng Item – Total Correlation của chỉ báo đó phải lớn hơn 0.3. Dựa vào tiêu chuẩn này, thông qua cột Corrected Item-Total Correlation tất cả các chỉ báo đều lớn hơn 0.3 nên đều được giữ lại. Thật vậy, loại bỏ bất kỳ chỉ báo nào trong 3 mục này đều làm giảm đi độ tin cậy của thang đo, điều này được biểu thị ở cột Cronbach's Alpha if Item Deleted.

Ví dụ 2: Phân tích tính nhất quán nội tại đối với biến Phong cách lãnh đạo

Về mặt lý thuyết, biến Phong cách lãnh đạo được đo lường bởi 4 chỉ báo (từ LD1 đến LD4), tuy nhiên, dựa vào kết quả phân tích nhân tố khám phá EFA có thể thấy rằng biến Phong cách lãnh đạo được đo lường bởi các items LD1, LD2, LD3, LD4, CV2.

Các items này được sử dụng để phân tích nhân tố khám phá, thực hiện với SPSS theo quy trình như trên, kết quả cho được như sau:

Reliability Statistics

Cronbach's Alpha	N of Items
.830	5

Có thể nhận thấy rằng Cronbach's Alpha = 0.830 nằm trong khoảng từ 0.7 đến 0.9 nên ta có thể khẳng định tính nhất quán nội tại của thang đo Phong cách lãnh đạo là tốt.

Tuy nhiên, có thể xem xét thêm hệ số tương quan biến tổng của chỉ báo CV2 = 0.48 (không thực sự lớn) và hệ số Cronbach's Alpha sẽ tăng (từ 0.830 lên 0.835) nếu bỏ đi chỉ báo CV2 (xem trong cột Cronbach's Alpha if Item Deleted) nên có thể bỏ chỉ báo CV2 trong biến Phong cách lãnh đạo và chắc chắn khi bỏ chỉ báo CV2, độ tin cậy nội tại (Cronbach's Alpha) sẽ là 0.835.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
LD1	18.19	22.547	.678	.781
LD2	18.34	23.101	.736	.765
LD3	17.73	25.103	.568	.813
LD4	17.91	23.550	.687	.779
CV2	17.90	26.629	.480	.835

Chú ý:

- Khi thực hiện phân tích tính nhất quán nội tại bằng Cronbach's Alpha, nếu hệ số Alpha nhỏ hơn 0.6 thì ta sẽ *lần lượt* bỏ chỉ báo (item) có hệ số tương quan biến tổng nhỏ hơn 0.3 (hệ số càng nhỏ càng ưu tiên bỏ trước) hoặc *lần lượt* bỏ chỉ báo có giá trị ở cột Cronbach's Alpha if Item Deleted cao nhất (hệ số ở cột này càng cao thì ưu tiên bỏ trước vì đây là giá trị Alpha khi bỏ item tương ứng) và sẽ thực hiện đến khi nào hệ số Cronbach's Alpha thỏa mãn điều kiện.
- Như đã đề cập ở phần trước, việc loại hoặc giữ lại chỉ báo trong thang đo lường cần phải xem xét giá trị về mặt nội dung (content validity) của thang đo, trong một số trường hợp nhiều chỉ báo không thỏa mãn điều kiện (ví dụ tương quan biến tổng < 0.3 nhưng vẫn giữ lại vì chỉ báo đó đóng vai trò quan trọng trong nội dung của thang đo nhằm duy trì giá trị về mặt nội dung – content validity).

KẾT HỢP ĐO LƯỜNG GIÁ TRỊ (VALIDITY) VÀ ĐỘ TIN CẬY (RELIABILITY) CỦA THANG ĐO TRONG TRƯỜNG HỢP ỨNG DỤNG THANG ĐO LẬP LẠI

Trong quá trình xây dựng thang đo, các nhà nghiên cứu thường áp dụng quy trình xây dựng thang đo của Churchill (1979), tuy nhiên, khi ứng dụng thang đo (thang đo đã có sẵn được sử dụng từ các nghiên cứu trước), các nhà nghiên cứu chỉ cần thực hiện kiểm tra tính nhất quán nội tại của thang đo hoặc kết hợp phân tích nhân tố khám phá (EFA) và kiểm tra tính nhất quán nội tại (bằng Cronbach's Alpha) để kiểm tra giá trị và độ tin cậy của thang đo.

Không có một quy định nào cho việc thực hiện phân tích nhân tố (EFA) hay kiểm tra tính nhất quán nội tại trước. Theo kinh nghiệm của tác giả, đối với những nghiên cứu cơ bản, thang đo hoàn toàn mới, người ta thường xem xét Cronbach Alpha trước cho từng khái niệm nhằm loại bỏ những biến rác trong mô hình, sau đó thực hiện phân tích nhân tố và kết quả các chỉ báo cho từng nhân tố trong phân tích nhân tố khám phá (EFA) sẽ sử dụng để tính hệ số Cronbach Alpha cuối cùng. Đối với những nghiên cứu các tác giả sử dụng lại thang đo để kiểm tra trong một môi trường được xem là mới thì sử dụng phân tích nhân tố khám phá, sau đó sử dụng kết quả thực tế các chỉ báo trong phân tích này để thực hiện kiểm tra hệ số Cronbach Alpha.

Ví dụ 1: Các nhân tố ảnh hưởng đến lòng trung thành của nhân viên đối với tổ chức

Như trong chương phân tích nhân tố khám phá (EFA), chúng ta đã kết luận có 6 nhân tố (6 biến số tiềm ẩn) được rút trích từ 32 chỉ báo (vì có 6 nhân tố có eigenvalue lớn hơn 1 và tổng phương sai trích của 6 nhân tố chính là 69.970% > 50%). Với KMO = 0.863 và sig. = 0.000

<0.05 nên ta có thể khẳng định dữ liệu phù hợp để thực hiện phân tích nhân tố.

Ở phần trên, chúng ta đã thực hiện phân tích Alpha Cronbach cho biến số Chính sách đào tạo (DT), tương tự như vậy tính toán cho các biến số TL, LD, DN, PL, MT. Thông thường trong trình bày ở các công trình khoa học để đảm bảo ngắn gọn trong trình bày nhưng đầy đủ về nội dung, người trình bày thường tổng hợp kết quả phân tích Alpha Cronbach và kết hợp với kết quả phân tích nhân tố để trình bày trong cùng một biểu. Như vậy, có thể khẳng định rằng tồn tại 6 nhân tố trong mô hình được biểu thị ở biểu sau:

	Component					
	1	2	3	4	5	6
DT6	.801					
DT5	.788					
DT3	.761					
DT7	.675					
DT4	.651					
TL4		.802				
TL5		.731				
TL2		.712				
TL1		.686				
TL3		.664				
LD1			.821			
LD2			.769			
LD3			.707			
LD4			.701			
CV2			.446			
DN1				.872		
DN2				.830		
DN3				.813		
PL3					.952	
PL2					.940	
MT1						.794
MT2						.773
Eigenvalue	6.224	1.968	1.847	1.560	1.247	1.147
% of Variance	31.122	9.839	9.236	7.799	6.237	5.737
Cumulative % of Variance	31.122	40.961	50.196	57.996	64.233	69.970
Cronbach Alpha	0.860	0.813	0.830	0.857	0.957	0.726

Các nhân tố này có thể đặt tên biểu hiện các biến số tiềm ẩn trong mô hình như sau:

- Nhân tố F1: Chính sách đào tạo (được đo lường bởi các items (DT3, DT4, DT5, DT6, DT7))
- Nhân tố F2: Cảm nhận về tiền lương, thưởng (được đo lường bởi các items TL1, TL2, TL3, TL4, TL5)
- Nhân tố F3: Phong cách lãnh đạo (được đo lường bởi các items LD1, LD2, LD3, LD4, CV2)

- Nhân tố F4: Đồng nghiệp (được đo lường bởi các items DN1, DN2, DN3)
- Nhân tố F5: Phúc lợi (được đo lường bởi các items PL2, PL3)
- Nhân tố F6: Môi trường (được đo lường bởi các items MT1, MT2)

Sau khi đã biết được các items đo lường cho mỗi biến tiềm ẩn, sử dụng các items này để tính hệ số Cronbach Alpha của từng biến số nhằm đo lường tính nhất quán nội tại.

Sau khi thực hiện phân tích, các biến số này đều có giá trị Alpha > 0.7, tương quan biến tổng khi phân tích Cronbach Alpha đều lớn hơn 0.3 nên có thể kết luận các biến số này đảm bảo tính nhất quán nội tại.

Ví dụ 2: Lòng trung thành của nhân viên đối với tổ chức

Thực hiện tương tự như vậy đối với biến số lòng trung thành của nhân viên đối với tổ chức, kết quả phân tích nhân tố khám phá đối với biến số lòng trung thành của nhân viên đối với tổ chức cho phép khẳng định rút trích thành một nhân tố (biến số lòng trung thành) từ 3 chỉ báo. Hệ số Alpha = 0.797 > 0.7 nên khẳng định biến số đảm bảo tính nhất quán nội tại.

Component Matrix ^a	
	Component
	1
LTT1	.874
LTT3	.834
LTT2	.829
Eigenvalue	2.146
% of Variance	71.523
Cumulative % of Variance	71.523
Cronbach Alpha	.797

CHƯƠNG

9.3

HỒI QUY BỘI (MULTIPLE REGRESSION)

Trong phần hồi quy đơn đã giới thiệu phương pháp thống kê để phân tích mối quan hệ giữa một biến phụ thuộc định lượng với một biến độc lập định lượng. Mô hình hồi quy bội (Multiple Regression) mở rộng mô hình hồi quy đơn bằng cách thêm vào một số biến độc lập để giải thích biến phụ thuộc hay nói một cách khác hơn, hồi quy bội là kỹ thuật thống kê phân tích mối quan hệ giữa một biến phụ thuộc (Dependent Variable - DV) và một tập các biến độc lập (Independent Variable – IV).

Hồi quy bội là một công cụ thống kê thường được sử dụng nhằm đạt được 3 mục tiêu:

- Tìm ra phương trình dự báo tốt nhất cho tập các biến, ví dụ giữa các biến $X_1, X_2, X_3, \dots$ và Y ?
- Kiểm soát các nhân tố gây “nhiều” để đánh giá sự đóng góp của một biến đặc biệt, tập các biến hay xác định các mối quan hệ không phụ thuộc.
- Để xác định mối liên hệ cấu trúc và đưa ra lời giải thích cho mỗi quan hệ đa biến phức hợp, như được tiến hành trong phân tích đường dẫn (path analysis).

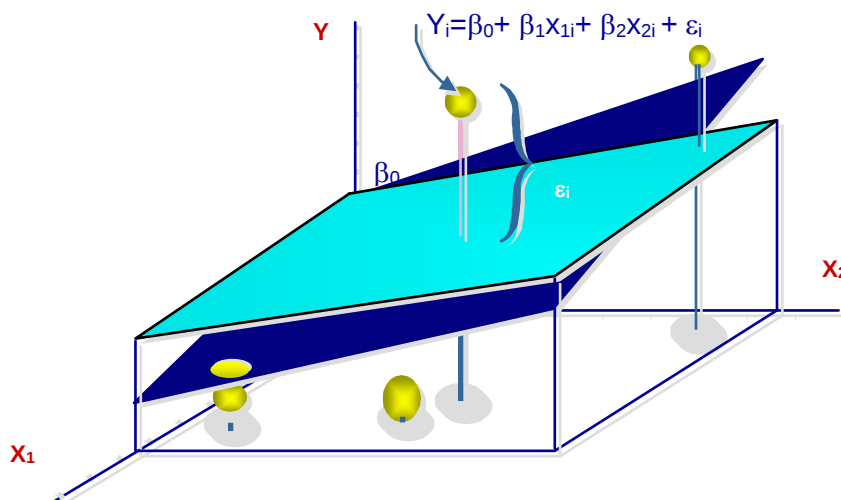
Mô hình hồi quy bội có dạng:

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \varepsilon_i$$

Hệ số chặn (hằng số) Hệ số hồi quy của tổng thể Sai số ngẫu nhiên

Biến phụ thuộc Các biến độc lập

Mô hình hồi quy bội được thể hiện trên hình vẽ như sau:



CÁC YÊU CẦU ĐỐI VỚI PHÂN TÍCH HỒI QUY BỘI

- Kích thước mẫu ảnh hưởng trực tiếp đến quyền năng trong hồi quy bội. Vì vậy, một nguyên tắc đơn giản theo Cohen và Cohend (1983), độ lớn của mẫu (số đối tượng) ít nhất nên gấp 20 lần so với số biến độc lập. Ví dụ, nếu một nghiên cứu thực hiện phân tích mối quan hệ chặt chẽ giữa 5 biến độc lập với biến phụ thuộc thì nên có ít nhất là 100 mẫu (đối tượng).
- Đối với biến độc lập, các biến đo lường có thể là liên tục (định lượng) hoặc nhị phân. Khi các biến độc lập định tính nhiều hơn 2 lựa chọn, chúng sẽ được chuyển đổi thành tập các biến nhị phân bằng cách mã hoá biến thành biến giả (dummy variable coding). Biến giả là biến nhị phân và đại diện cho một loại biến độc lập định tính. Bất kỳ biến định tính với k lựa chọn sẽ được mã hóa thành k-1 biến giả. Ví dụ khi khảo sát các bộ công nhân viên ở ba miền Bắc, Trung, Nam, để có thể thực hiện hồi quy đối với biến này thì cần mã hóa lại biến địa điểm thành 2 biến giả là Miền Bắc (nhận các giá trị 1: Miền Bắc, 0: các miền còn lại) và Miền Trung (nhận giá trị 1: Miền Trung, 0: các miền còn lại).

ĐIỀU KIỆN ỨNG DỤNG PHÂN TÍCH HỒI QUY BỘI

- Mối quan hệ giữa biến độc lập và biến phụ thuộc là tuyến tính: kiểm tra bằng phân tích đồ thị phân tán (scatter plots),
- Phương sai của các sai số là hằng số: SPSS cung cấp kiểm định Levene cho sự biến đổi đồng nhất, kiểm định này đo lường sự bằng nhau của các thay đổi cho từng cặp biến.
- Các sai số phải độc lập lẫn nhau: SPSS sẽ cung cấp thống kê Durbin-Watson (d) giúp kiểm định sự tương quan giữa các sai số liên kề, nếu giá trị này có ý nghĩa thì đó là dấu hiệu của sự không độc lập của sai số.
- Các sai số phải tuân theo quy luật phân phối chuẩn: Người ta giả định rằng các sai số của các dự báo (hiệu giữa kết quả của biến phụ thuộc có được và kết quả biến phụ thuộc được dự báo) có phân phối chuẩn.

Việc vi phạm giả định này có thể được nhận biết thông qua việc xây dựng biểu đồ phần dư, kiểm tra trực quan để thấy phân phối có xấp xỉ phân phối chuẩn hay không hoặc sử dụng kiểm định Skewness và Kurtosis.

CÁC KỸ THUẬT HỒI QUY BỘI

Có 3 kỹ thuật hồi quy bội chủ yếu là **standard multiple regression**, **hierarchical regression**, và **statistical (stepwise) regression**.

Hồi quy bội chuẩn (Standard multiple regression)

Trong mô hình hồi quy này, tất cả các biến độc lập của nghiên cứu đều được nhập vào phương trình một lần, sau đó, mỗi biến độc lập được ước lượng trong giới hạn phương sai độc lập (riêng lẻ) mà nó giải thích. Nhược điểm của mô hình hồi quy chuẩn là một biến độc lập có thể quan hệ chặt chẽ với biến phụ thuộc, nhưng được xem là biến không quan trọng trong dự báo nếu sự đóng góp độc lập (riêng lẻ) của nó trong việc giải thích biến phụ thuộc là nhỏ.

Hồi quy thứ bậc (Hierarchical regression)

Đây là mô hình hồi quy linh hoạt nhất vì nó cho phép người nghiên cứu quyết định trật tự các biến độc lập đưa vào phương trình hồi quy. Mỗi biến độc lập được đánh giá vào thời điểm riêng của nó đưa vào trong thời hạn cho phép để giải thích sự đóng góp của biến vào phương trình. Thứ tự nhập vào thường theo sự xem xét một cách logic hoặc lý thuyết. Chẳng hạn như, dựa theo lý thuyết, người nghiên cứu có thể quyết định 2 biến độc lập cụ thể nào (từ tập các biến độc lập) sẽ dự báo tốt nhất biến phụ thuộc. Vì vậy, 2 biến độc lập này sẽ được ưu tiên đưa vào trước và tổng khả năng giải thích của các biến này (nghĩa là tổng số sự thay đổi được giải thích) được ước lượng. Các biến độc lập ít quan trọng hơn sẽ được nhập vào và được đánh giá chúng đóng góp thêm vào việc giải thích được tạo ra bởi 2 biến độc lập đầu tiên như thế nào.

Hồi quy thống kê (Statistical (Stepwise) regression)

Đối với mô hình hồi quy thống kê này, thứ tự đưa vào các biến giải thích chỉ dựa vào tiêu chuẩn thống kê. Các biến tương quan mạnh với biến phụ thuộc sẽ được ưu tiên nhập vào trước mà không liên quan đến việc xem xét một cách lý thuyết. Nhược điểm của mô hình này là tiêu chuẩn thống kê được sử dụng để xác định biến ưu tiên đưa vào trước có thể đặc biệt cho mẫu thuận tiện. Còn với mẫu khác thì tiêu chuẩn thống kê được tính có thể sẽ khác, dẫn đến việc với cùng một biến nhưng thứ tự nhập vào sẽ khác nhau. Mô hình hồi quy thống kê được sử dụng chủ yếu trong nghiên cứu thăm dò, khi người nghiên cứu không chắc chắn về mối quan hệ dự báo của các biến độc lập.

Hồi quy thống kê (statistical stepwise regression) có thể được thực hiện thông qua một trong 3 phương pháp: **forward selection**, **backward deletion** và **stepwise regression**.

- Trong phương pháp **forward selection**, các biến được ước lượng ngược lại với tiêu chuẩn thống kê và nếu chúng đáp ứng những tiêu chuẩn này thì được ưu tiên nhập vào dựa trên mối quan hệ tương quan của chúng với biến phụ thuộc. Biến nào tương quan với biến phụ thuộc mạnh nhất sẽ được nhập vào phương trình đầu tiên và khi xuất hiện trong phương trình thì nó luôn được giữ lại.
- Với phương pháp **backward deletion** thì phương trình bắt đầu tiến hành với tất cả các biến độc lập được nhập vào. Mỗi biến sau đó sẽ được ước lượng một lần, theo đóng góp của nó vào phương trình hồi quy, các biến nào không có ý nghĩa thống kê sẽ bị loại bỏ.
- Đối với phương pháp **stepwise regression** thì các biến được ước lượng để nhập vào phương trình theo cả 2 tiêu chuẩn lựa chọn chuyển tiếp (forward selection) và loại bỏ dần (backward deletion). Nghĩa là các biến được nhập vào một lần nếu đáp ứng tiêu chuẩn thống kê, nhưng chúng cũng có thể bị loại bỏ ở bất kỳ bước nào nếu nó không có ý nghĩa thống kê bền vững đối với mô hình hồi quy.

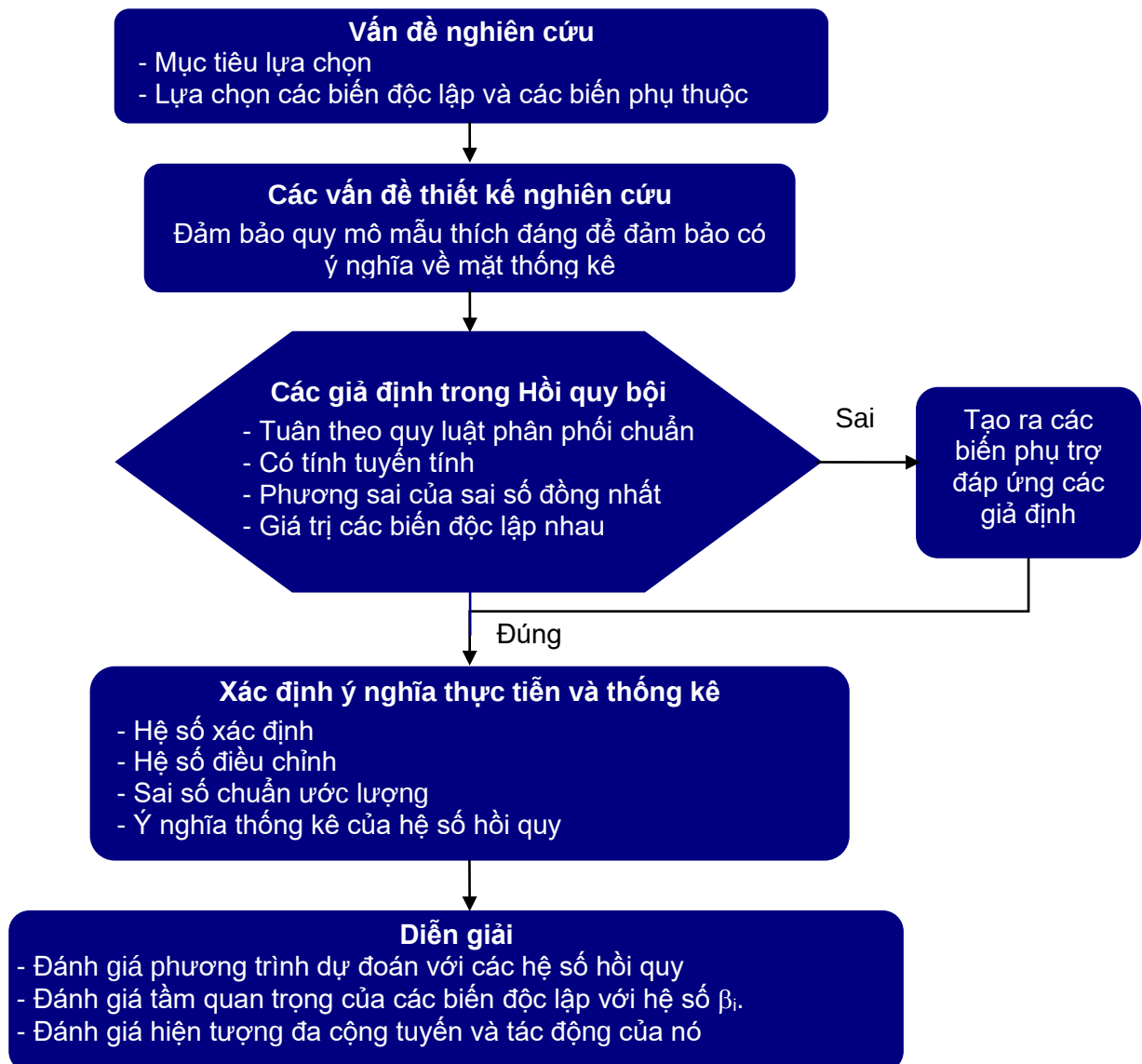
HIỆN TƯỢNG ĐA CỘNG TUYẾN (MULTICOLLINEARITY)

Đa cộng tuyến là hiện tượng tồn tại mối quan hệ ở mức độ cao giữa các biến độc lập (biến giải thích). Khi các biến độc lập xuất hiện đa cộng tuyến, thì có sự chồng chéo hay phân chia khả năng dự báo, điều này có thể dẫn đến kết quả nghịch lý, tuy mô hình hồi quy thích hợp với dữ liệu, nhưng không có biến dự báo nào có sự ảnh hưởng đáng kể trong việc dự báo biến phụ

thuộc. Điều này xảy ra là bởi khi các biến dự báo có mối tương quan cao, chúng sẽ phân chia những thông tin chung. Vì vậy, chúng có thể cùng giải thích nhiều về biến phụ thuộc, nhưng không có đóng góp riêng lẻ đáng kể vào mô hình. Như vậy, đa cộng tuyến làm giảm khả năng dự báo riêng biệt của biến độc lập ở bất kỳ phạm vi nào bởi nó được kết hợp với các biến độc lập khác. Nghĩa là, không có biến dự báo nào đóng góp riêng biệt và đáng kể vào mô hình dự đoán khi chúng được kết hợp với nhau.

Các công cụ dò tìm sự tồn tại của đa cộng tuyến gồm:

- Xem xét hệ số tương quan tuyến tính giữa các biến độc lập trong ma trận hệ số tương quan, nếu một biến có tương quan lớn với các biến còn lại thì dễ xảy ra hiện tượng đa cộng tuyến.



- Hệ số Tolerance: Hệ số Tolerance = $1 - R^2_k$ và trong đó R^2_k là hệ số tương quan bội khi biến độc lập X_k được dự đoán từ các biến độc lập khác (hay phần trăm sự biến đổi của biến X_k từ các biến độc lập khác). Khi Tolerance càng nhỏ thì càng có dấu hiệu đa cộng tuyến giữa các biến độc lập.
- Hệ số phóng đại phương sai VIF (Variance Inflation Factor):

Hệ số VIF = $1/\text{Tolerance} = 1/(1 - R^2_k)$, khi Tolerance càng nhỏ thì VIF càng lớn và thông thường VIF > 10 là xảy ra cộng tuyến.

MÔ HÌNH HỒI QUY TUYẾN TÍNH BỘI

Ma trận hệ số tương quan (correlation matrix)

Khi tiến hành phân tích hồi quy tuyến tính bội, chúng ta cần xem xét mối quan hệ giữa các biến độc lập (IV) cũng như giữa các biến độc lập (IV) và các biến phụ thuộc (DV) trên cơ sở ma trận hệ số tương quan (correlation matrix). Đặc biệt, cần chú ý đến mối quan hệ chặt chẽ giữa các một biến độc lập với các biến độc lập còn lại vì có thể sẽ xảy ra hiện tượng đa cộng tuyến.

Đánh giá mức độ phù hợp của mô hình hồi quy tuyến tính bội

Các hệ số R^2 (R-square) và R^2_a điều chỉnh (R^2_a viết tắt của Adjusted R – square) là các hệ số đánh giá mức độ phù hợp của mô hình hồi quy tuyến tính.

Hệ số R^2 được xem là hệ số có mối quan hệ đồng biến với số lượng các biến độc lập đưa vào mô hình và dĩ nhiên, khi thêm càng nhiều biến độc lập vào mô hình thì R^2 càng tăng.

Hệ số R^2 điều chỉnh là hệ số được điều chỉnh từ R^2 và được sử dụng để phản ánh sát hơn mức độ phù hợp của mô hình hồi quy tuyến tính đa biến và R^2 điều chỉnh không (nhất thiết) tăng khi nhiều biến độc lập được vào phương trình. R^2_a được điều chỉnh từ R^2 theo công thức sau:

$$R^2_a = R^2 - \frac{p(1-R^2)}{n-p-1}$$

Trong đó n là độ lớn của mẫu và p là số biến độc lập được đưa vào phân tích.

Kiểm định độ phù hợp của mô hình

Kiểm định giả thuyết về sự phù hợp của mô hình hồi quy tuyến tính tổng thể thông qua hệ số Fisher với giả thuyết không là không có mối quan hệ giữa các biến độc lập và biến phụ thuộc, nghĩa là $\beta_1 = \beta_2 = \dots = \beta_k = 0$. Nếu bác bỏ giả thuyết không nghĩa là các biến độc lập có mối quan hệ với biến phụ thuộc hay mô hình tuyến tính phù hợp với dữ liệu.

Giải thích ý nghĩa của các hệ số hồi quy β_k

Hệ số β_k đo lường sự thay đổi của giá trị Y khi X_k thay đổi một đơn vị và các biến số khác không đổi.

Hệ số Beta_k

Dễ dàng nhận thấy rằng, hệ số β_k luôn đi kèm với biến số X_k , các biến số này đôi khi có đơn vị đo khác nhau nên không thể so sánh được với nhau, vì vậy, hệ số Beta sẽ cho phép đo lường được trên cơ sở hệ số này được biểu diễn bằng đơn vị đo lường là độ lệch chuẩn. Hệ số Beta_k được tính như sau:

$$\text{Beta}_k = \beta_k \frac{S_k}{S_Y}$$

Trong đó S_k là độ lệch chuẩn của biến độc lập thứ k.

Đại lượng thống kê Durbin - Watson

Đại lượng Durbin – Watson (d) dùng để kiểm định sự tương quan của các sai số liên tiếp, theo phần điều kiện ứng dụng mô hình ta thấy các hệ số tương quan tổng thể của các phần dư bằng 0, giá trị d được tính như sau:

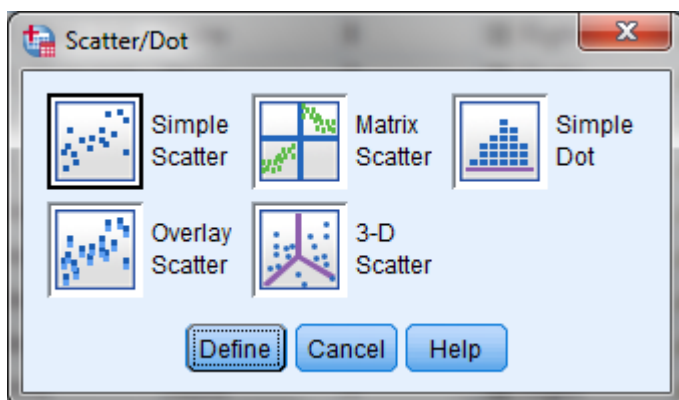
$$d = \frac{\sum_{i=1}^n (e_i - e_{i-1})^2}{\sum_{i=2}^n e_i^2}$$

Khi đó, $0 \leq d \leq 4$ và khi phần dư không có tương quan chuỗi bậc nhất thì giá trị d sẽ gần bằng 2. Miền tương quan và miền không tương quan có thể tham chiếu cụ thể như sau:

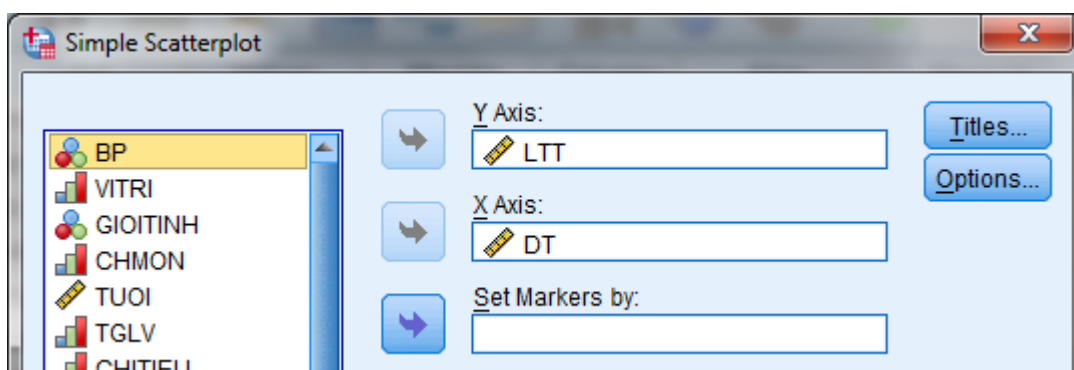
Có tương quan thuận chiều (+)	Không có kết luận	Không có tương quan chuỗi bậc nhất	Miền không có kết luận	Có tự tương quan ngược chiều (-)		
0	d_L	d_U	2	$4 - d_U$	$4 - d_L$	4

DÒ TÌM SỰ VI PHẠM CÁC GIẢI THUYẾT TRONG HÒI QUY TƯƠNG QUAN BỘI

- Mỗi quan hệ giữa biến độc lập và biến phụ thuộc là tuyến tính: kiểm tra bằng phân tích đồ thị phân tán (scatter plots):
 - + Vẽ biểu đồ Scatter bằng cách từ thanh trình đơn, vào **Graphs / Legacy Dialogs / Scatter/Dot**, kích vào **Simple Scatter** trên **Scatter/Dot**, chọn **Define**.

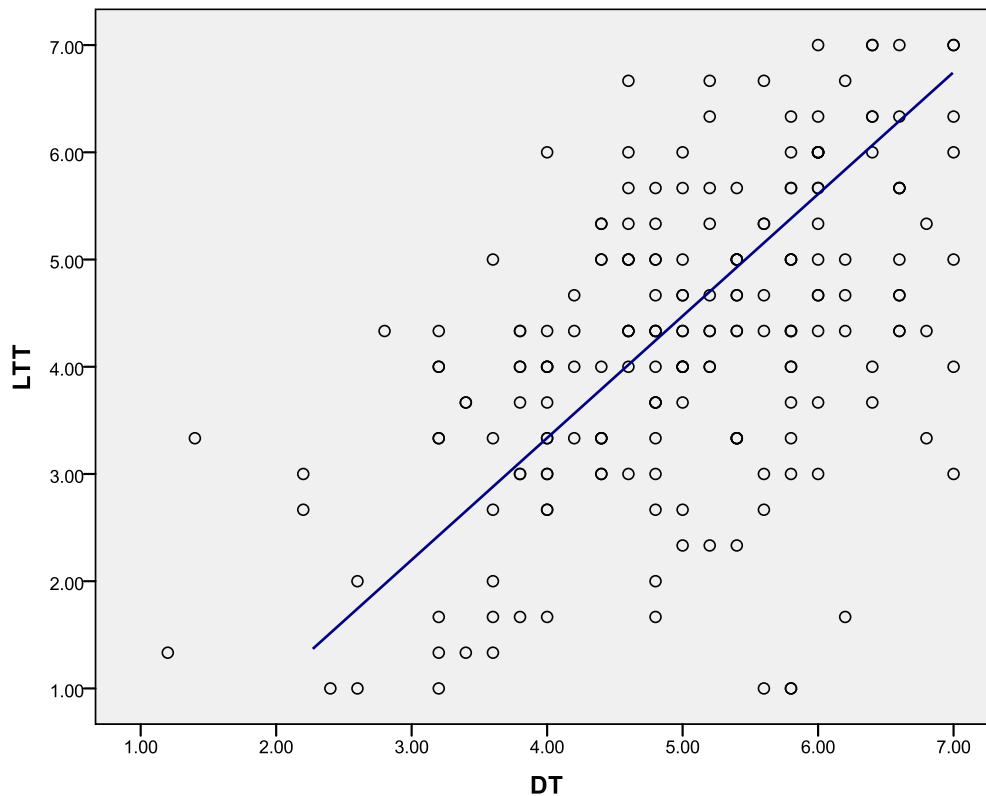


- + Chọn các giá trị cần xem xét vào các trục X và Y như hình vẽ (trong đó trục X chứa biến độc lập và trục Y chứa biến phụ thuộc), chọn **OK** để vẽ hình.



Biểu đồ phân tán của nhân viên được biểu diễn theo đường thẳng (điều đó có thể giải thích rằng khi cán bộ

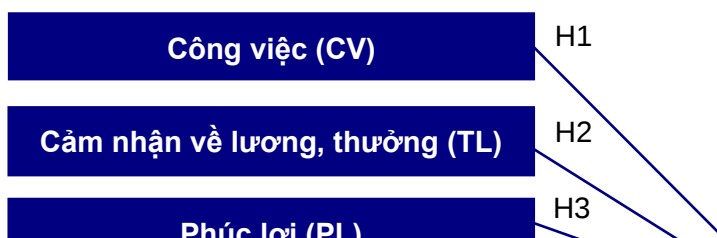
công nhân viên càng hài lòng với chính sách đào tạo của doanh nghiệp thì lòng trung thành của nhân viên đối với tổ chức càng cao).



- Phương sai của các sai số là hằng số: Trong khi thực hiện hồi quy (sẽ trình bày ở phần sau), kích vào **Save** để tiến hành sao lưu phần dư (sai số). Tạo ra một biến mới (bằng cách thực hiện **Compute** trong phần **Transform**) là giá trị tuyệt đối của phần dư. Thực hiện kiểm định Levene cho sự biến đổi đồng nhất hoặc kiểm định tương quan hạn Spearman.
- Các sai số phải độc lập lẫn nhau: Khi thực hiện hồi quy, kích vào **Durbin-Watson** (d) giúp kiểm định sự tương quan giữa các sai số liên kề (xuất hiện trong biểu **Model Summary**).
- Các sai số phải tuân theo quy luật phân phối chuẩn: Chúng ta có thể sử dụng kiểm định **Skewness** và **Kurtosis** để kiểm định phân phối chuẩn hoặc sử dụng kiểm định **Kolgomogrov** một mẫu.

THỰC HIỆN HỒI QUY BỘI TRÊN SPSS

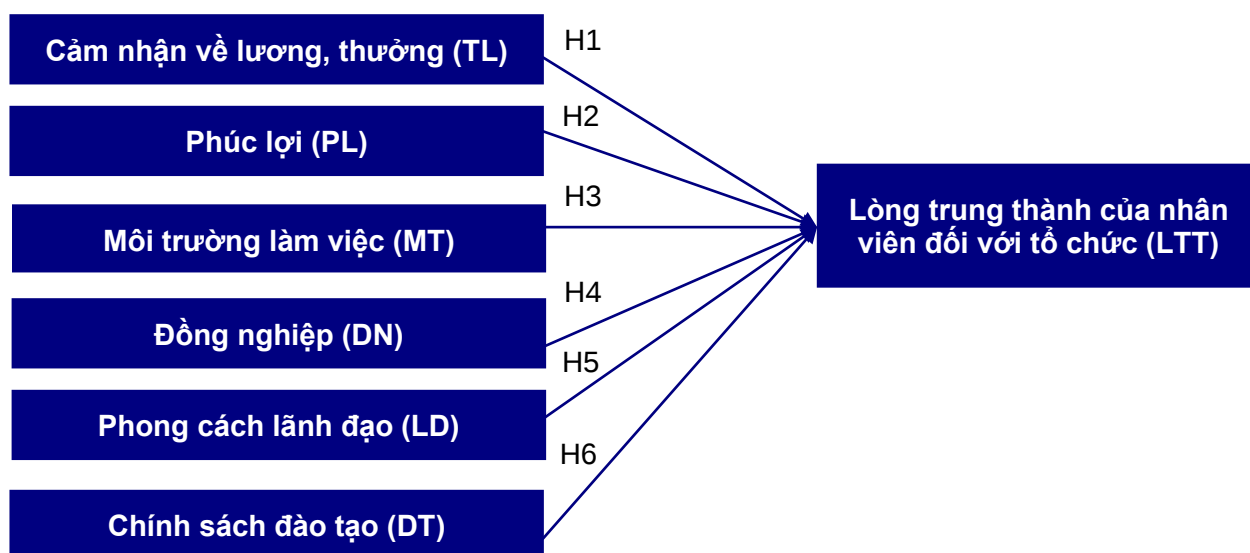
Như đã giới thiệu ở các phần trước, mô hình lý thuyết đề nghị nhằm nghiên cứu các nhân tố tác động đến lòng trung thành của nhân viên đối với tổ chức bao gồm 7 khái niệm là (1) công việc; (2) cảm nhận về lương, thưởng; (3) phúc lợi; (4) môi trường làm việc; (5) đồng nghiệp; (6) phong cách lãnh đạo; (7) chính sách đào tạo), các biến độc lập và các biến phụ thuộc đều là những biến định lượng (Likert 7 sự lựa chọn). Mối quan hệ giữa các khái niệm thể hiện trên hình vẽ:



Qua phân tích nhân tố khám phá đã rút gọn từ 32 chỉ báo đo lường 7 khái niệm trong mô hình lý thuyết xuống còn 6 khái niệm đo lường bởi 20 chỉ báo và kiểm định Cronbach Alpha cho thấy có tồn tại độ thống nhất nội tại giữa các items trong từng nhân tố.

Do có sự thay đổi số lượng nhân tố (từ 7 nhân tố xuống còn 6 nhân tố) nên cần điều chỉnh mô hình và phát biểu lại các thuyết. Lúc này mô hình gồm mối quan hệ giữa 6 khái niệm (biến độc lập) gồm: (1) cảm nhận về lương và thưởng; (2) phúc lợi; (3) môi trường làm việc; (4) đồng nghiệp; (5) phong cách lãnh đạo, (6) chính sách đào tạo và một khái niệm (biến phụ thuộc) là lòng trung thành của nhân viên đối với tổ chức.

Như vậy, tương ứng với 6 khái niệm (6 biến số trên hình vẽ), 6 giả thuyết tương ứng (phát biểu lại) về mối quan hệ giữa biến độc lập thứ *i* với biến phụ thuộc hay nói rõ hơn, tồn tại mối quan hệ giữa biến độc lập thứ *i* với lòng trung thành của nhân viên đối với tổ chức.



Tạo các biến tiềm ẩn

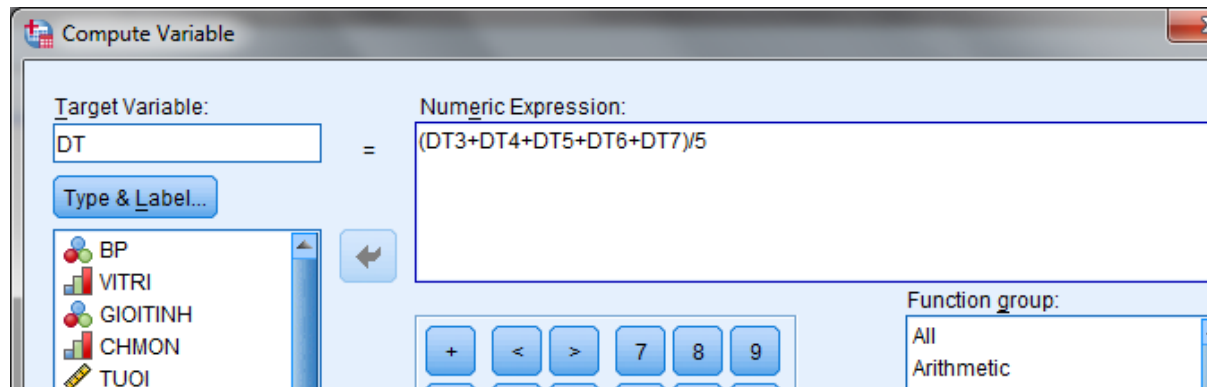
Đầu tiên, cần tạo ra các biến mới TL, PL, MT, DN, LD, DT từ các chỉ báo đo lường các biến số

trên, với kết luận ở chương trước, chính sách đào tạo được đo lường bởi các items (DT3, DT4, DT5, DT6, DT7). Như vậy biến DT được tạo lập như sau:

- Vào **Transform / Compute Variable**, cửa sổ **Compute Variable** mở ra,
- Tạo biến DT ở ô **Target Variable** và biến này được tính bởi trung bình cộng của giá trị các chỉ báo từ DT3 đến DT7, cụ thể là: $(DT3+DT4+DT5+DT6+DT7)/5$.

Chú ý: Các biến (biến tiềm ẩn) được ta tạo ra bởi giá trị của các chỉ báo nào cần dựa vào kết quả kết luận ở cuối chương đánh giá độ tin cậy của thang đo (Cronbach Alpha).

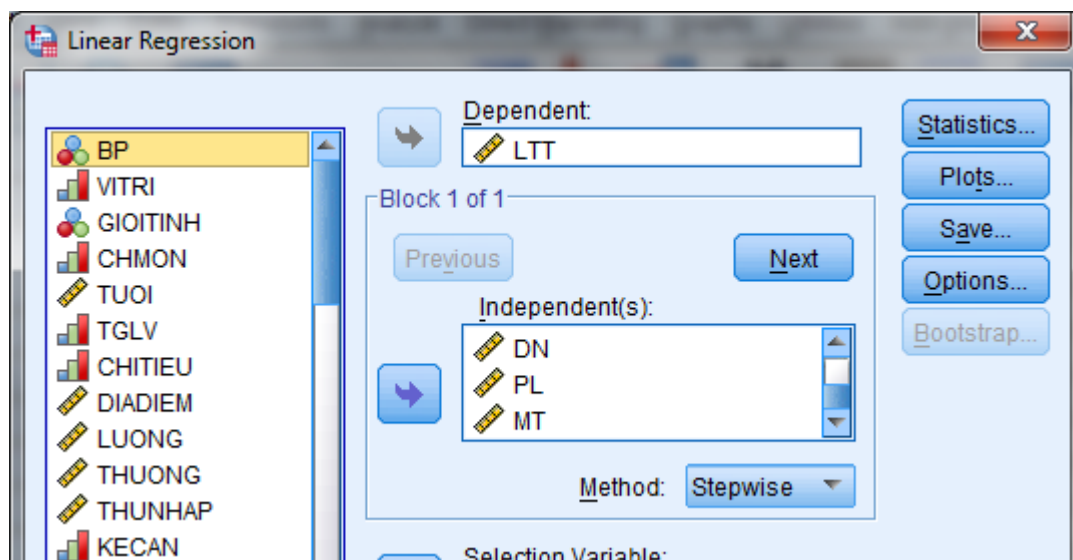
- Bấm **OK** để tạo biến.



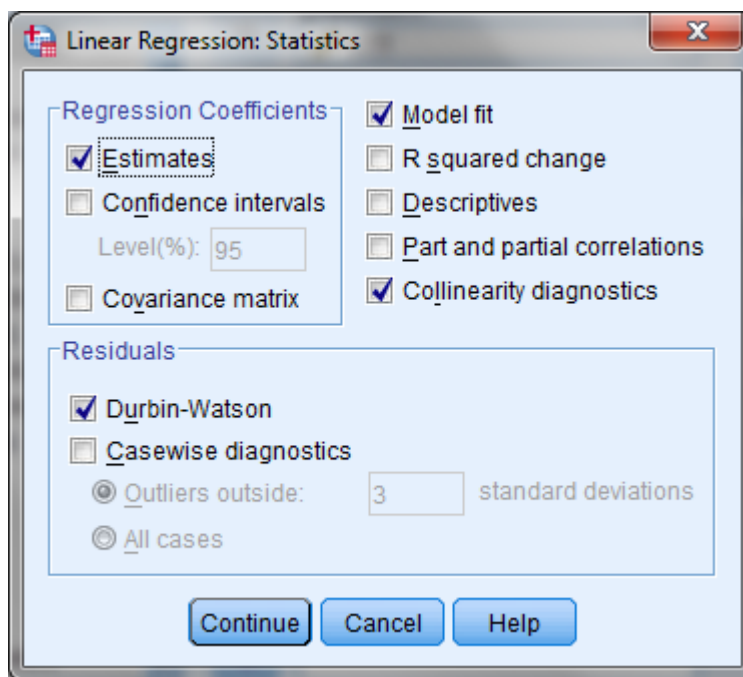
- Thực hiện tương tự với các biến TL, PL, MT, DN, LD, LTT.

Hồi quy tuyến tính bội với phương pháp stepwise

- Vào **Analyze / Regression / Linear**, cửa sổ **Linear Regression** mở ra,
- Trong phân tích hồi quy, cần xác định biến độc lập và biến phụ thuộc, trong ô **Dependent** (biến phụ thuộc) chọn biến lòng trung thành của nhân viên đối với tổ chức (LTT), trong ô **Independent(s)** chọn các biến độc lập DT, TL, LD, DN, PL và MT.



- Kích vào **Statistics**, ngoài các yếu tố mặc định, chọn **Collinearity diagnostics** và **Durbin-Watson**, sau đó bấm **Continue** để trở về cửa sổ ban đầu.



- Sau đó bấm **OK** để hoàn tất việc thiết đặt và phân tích kết quả.
- Kết quả phân tích và diễn giải dữ liệu:

Bảng ANOVA

Đối với hồi quy tuyến tính bội sử dụng phương pháp stepwise, mô hình sử dụng để diễn giải kết quả luôn là mô hình cuối cùng. Trong trường hợp này là mô hình thứ 3.

ANOVA ^d						
Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	136.272	1	136.272	101.575	.000 ^a
	Residual	265.635	198	1.342		
	Total	401.906	199			
2	Regression	180.282	2	90.141	80.126	.000 ^b
	Residual	221.624	197	1.125		
	Total	401.906	199			
3	Regression	189.284	3	63.095	58.162	.000 ^c
	Residual	212.622	196	1.085		
	Total	401.906	199			

a. Predictors: (Constant), TL

b. Predictors: (Constant), TL, LD

c. Predictors: (Constant), TL, LD, DT

d. Dependent Variable: LTT

Qua kết quả phân tích, vì $F = 58.162$ và $p(F) = 0.000 < 0.05$ nên có thể khẳng định tồn tại mối quan hệ giữa các biến TL, LD, DT với LTT.

Bảng Coefficients

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	1.282	.304		4.224	.000		
	TL	.705	.070	.582	10.078	.000	1.000	1.000
2	(Constant)	.171	.330		.517	.605		
	TL	.497	.072	.410	6.875	.000	.787	1.271
	LD	.440	.070	.373	6.255	.000	.787	1.271
3	(Constant)	-.309	.364		-.849	.397		
	TL	.446	.073	.368	6.100	.000	.741	1.349
	LD	.331	.079	.280	4.191	.000	.604	1.656
	DT	.236	.082	.191	2.881	.004	.616	1.624

a. Dependent Variable: LTT

Qua phân tích lý thuyết, có thể thiết lập mô hình lý thuyết thể hiện mối quan hệ giữa các biến độc lập và biến phụ thuộc trên tổng thể như sau:

$$LTT_i = B_0 + B_1TL_i + B_2PL_i + B_3MT_i + B_4DN_i + B_5LD_i + B_6DT_i + \varepsilon_i$$

Qua kết quả phân tích, tất cả các giá trị Sig. = p(t) tương ứng với các biến TL, LD và DT lần lượt là 0.000; 0.000 và 0.004 đều nhỏ hơn 0.05. Do vậy, có thể khẳng định các biến số này có ý nghĩa trong mô hình và mô hình thực tế được thiết lập như sau:

$$LTT_i = -0.309 + 0.446TL_i + 0.331LD_i + 0.236DT_i + \varepsilon_i$$

Qua phương trình trên, có thể nhận thấy rằng với một cấp độ thay đổi cảm nhận về tiền lương, thường thì lòng trung thành của nhân viên sẽ tăng 0.446 cấp độ. Giải thích tương tự như trên đối với phong cách lãnh đạo và chính sách đào tạo. *Chú ý:* Các biến độc lập và biến phụ thuộc được đo lường bởi 7 sự lựa chọn (cấp độ) với (1) là hoàn toàn không đồng ý và (7) là hoàn toàn đồng ý.

Bảng Model Summary

Model Summary ^d					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.582 ^a	.339	.336	1.15827	
2	.670 ^b	.449	.443	1.06066	
3	.686 ^c	.471	.463	1.04154	1.595

a. Predictors: (Constant), TL

b. Predictors: (Constant), TL, LD

c. Predictors: (Constant), TL, LD, DT

d. Dependent Variable: LTT

Kết quả phân tích cho thấy $R^2 = 0.471$ nên mối quan hệ giữa các biến TL, LD, DT và LTT là trung bình ($0 < R^2 < 0.5$) hay nói khác hơn ba biến TL, LD, DT giải thích được 47.1% lòng trung thành của nhân viên đối với tổ chức. Như vậy, $1 - R^2 = 0.529$ được giải thích bởi những nhân tố không được kiểm soát trong mô hình và đây cũng được xem là một trong những hạn chế của nghiên cứu.

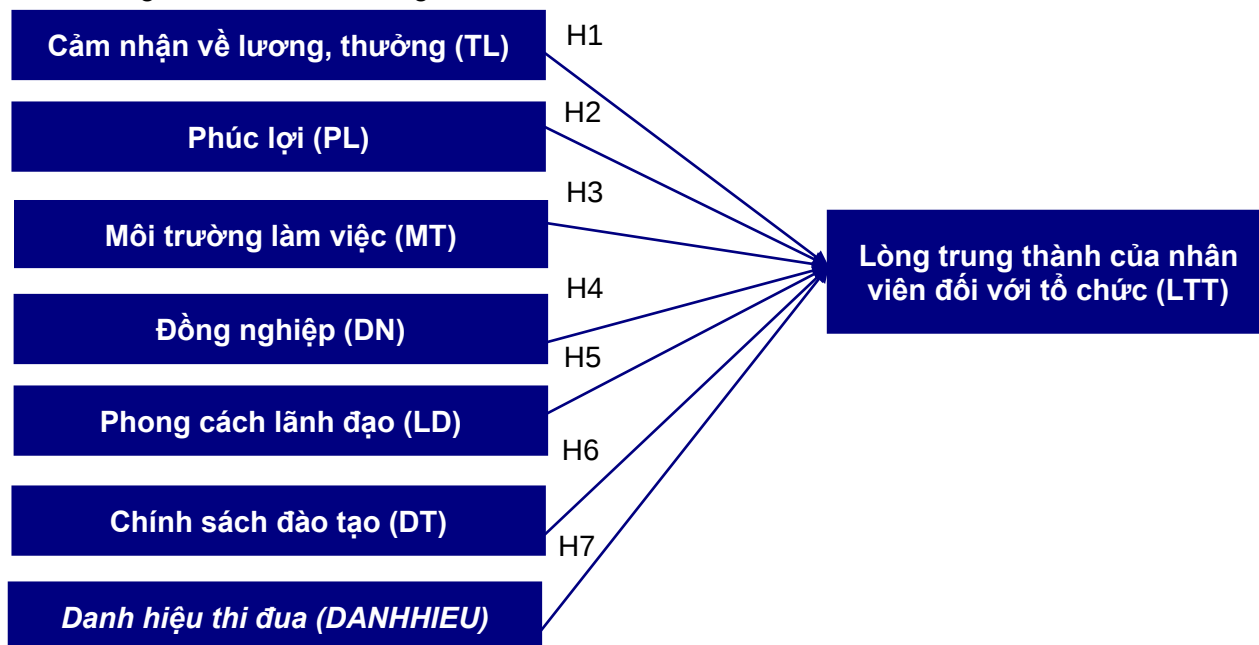
THỰC HIỆN HỒI QUY BỘI VỚI BIẾN GIẢ (DUMMY)

Ở phần trên giới thiệu hồi quy tuyến tính bội giữa biến phụ thuộc định lượng và các biến độc lập định lượng. Tuy nhiên, trên thực tế, nhà nghiên cứu muốn phân tích mối quan hệ giữa một biến phụ thuộc định lượng với các biến độc lập định lượng và định tính. Ví dụ, nhà nghiên cứu muốn tìm hiểu người có nhận được danh hiệu thi đua trong thời gian qua có trung thành hơn so với người không nhận danh hiệu hay không? Người nữ có lòng trung thành với tổ chức cao hơn người nam hay không? Vấn đề đặt ra là việc vận dụng mô hình hồi quy tuyến tính để giải quyết vấn đề này như thế nào?

Trong trường hợp như vậy, các biến định tính phải mã hóa lại thành các biến giả (biến dummy) nhận các giá trị (0, 1), một biến định tính có k lựa chọn thì số biến giả được tạo thành sẽ là k-1. Trờ lại ví dụ ở chương trước, các bộ công nhân viên khảo sát ở ba miền Bắc, Trung, Nam, để có thể thực hiện hồi quy với biến định tính (địa điểm), cần mã hóa lại biến địa điểm thành 2 biến giả là Miền Bắc (nhận các giá trị 1: Miền Bắc, 0: các miền còn lại) và Miền Trung (nhận giá trị 1: Miền Trung, 0: các miền còn lại) hoặc nếu muốn nghiên cứu người có nhận danh hiệu trong thời gian qua có lòng trung thành khác với người không nhận danh hiệu không, lúc đó, ngoài 6 biến độc lập như trên còn có biến danh hiệu (DANHHIEU). Vì biến danh hiệu thi đua DANHHIEU nhận các giá trị 1: có danh hiệu và 2: chưa có danh hiệu nên cần mã hóa lại biến DANHHIEU thành biến DANHHIEUMH với các giá trị 1: Có danh hiệu và 0: Chưa có danh hiệu.

Chú ý: Việc mã hóa lại biến được thực hiện bằng phương pháp **Recode into Different Variable** trong phần **Transform** trên thanh công cụ.

Mô hình nghiên cứu được đề nghị như sau:

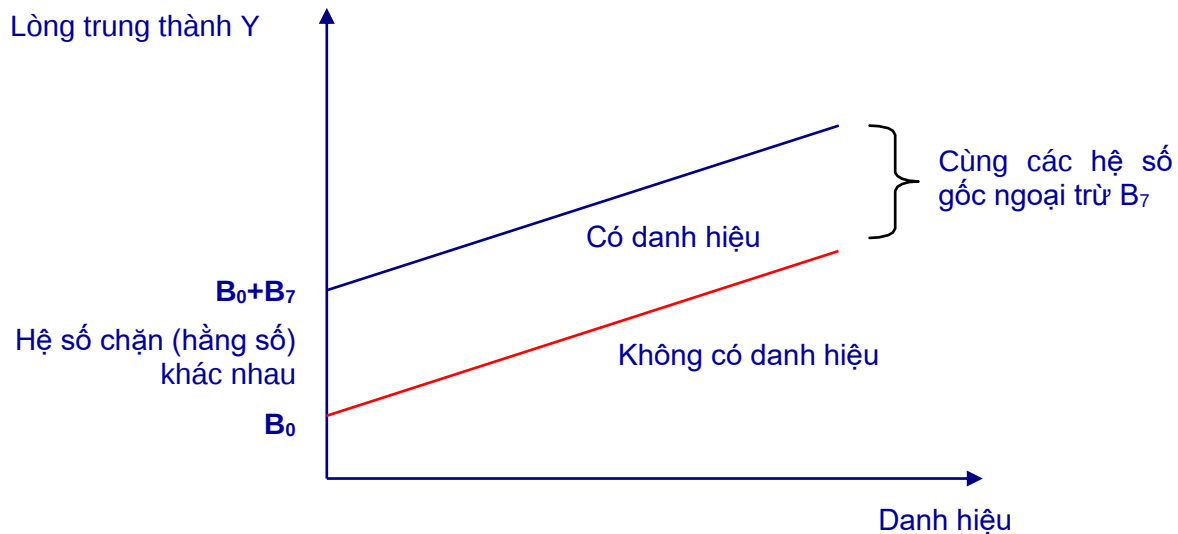


Phương trình lý thuyết mối quan hệ giữa các 7 biến độc lập và biến phụ thuộc trên tổng thể là:

$$LTT_i = B_0 + B_1TL_i + B_2PL_i + B_3MT_i + B_4DN_i + B_5LD_i + B_6DT_i + B_7DANHHIEUMH + \varepsilon_i$$

Khi đó, chênh lệch về lòng trung thành của người có danh hiệu với người không có danh hiệu bằng giá trị của hệ số góc B_7 .

$$LTT_i = (B_0 + B_7) + B_1TL_i + B_2PL_i + B_3MT_i + B_4DN_i + B_5LD_i + B_6DT_i + B_7DANHHIEUMH_i \varepsilon_i$$



Kết quả phân tích hồi quy giữa 7 biến độc lập và biến phụ thuộc (LTT) cho kết quả như sau:

Bảng ANOVA:

ANOVA ^e						
Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	136.272	1	136.272	101.575	.000 ^a
	Residual	265.635	198	1.342		
	Total	401.906	199			
2	Regression	180.282	2	90.141	80.126	.000 ^b
	Residual	221.624	197	1.125		
	Total	401.906	199			
3	Regression	189.284	3	63.095	58.162	.000 ^c
	Residual	212.622	196	1.085		
	Total	401.906	199			
4	Regression	195.106	4	48.777	45.993	.000 ^d
	Residual	206.800	195	1.061		
	Total	401.906	199			

a. Predictors: (Constant), TL

b. Predictors: (Constant), TL, LD

c. Predictors: (Constant), TL, LD, DT

d. Predictors: (Constant), TL, LD, DT, DANHHIEUMH

e. Dependent Variable: LTT

Như đã đề cập ở trên, trong hồi quy với stepwise, mô hình xem xét luôn là mô hình cuối cùng nên với $F = 45.993$ và $p(F) = 0.000 < 0.05$, có thể khẳng định tồn tại mối quan hệ giữa các biến TL, LD, DT, DANHHIEUMH với LTT.

Bảng Coefficients:

Qua kết quả phân tích, tất cả các giá trị Sig. = p(t) tương ứng với các biến TL, LD, DT và DANHHIEUMH đều nhỏ hơn 0.05. Do vậy, có thể khẳng định các biến số này có ý nghĩa trong

mô hình và mô hình thực tế được thiết lập như sau:

$$LTT_i = -0.309 + 0.446TL_i + 0.331LD_i + 0.236DT_i + 0.350DANHIEUMH_i + \varepsilon_i$$

Vì $B_7 = 0.350 > 0$ khẳng định rằng người có nhận danh hiệu có lòng trung thành đối với tổ chức cao hơn người không nhận danh hiệu là 0.35 (cấp độ)

Coefficients ^a								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	1.282	.304		4.224	.000		
	TL	.705	.070	.582	10.078	.000	1.000	1.000
2	(Constant)	.171	.330		.517	.605		
	TL	.497	.072	.410	6.875	.000	.787	1.271
	LD	.440	.070	.373	6.255	.000	.787	1.271
3	(Constant)	-.309	.364		-.849	.397		
	TL	.446	.073	.368	6.100	.000	.741	1.349
	LD	.331	.079	.280	4.191	.000	.604	1.656
	DT	.236	.082	.191	2.881	.004	.616	1.624
4	(Constant)	-.329	.360		-.913	.362		
	TL	.416	.073	.344	5.673	.000	.719	1.391
	LD	.324	.078	.274	4.145	.000	.603	1.658
	DT	.237	.081	.191	2.914	.004	.616	1.624
	DANHIEUMH	.350	.149	.124	2.343	.020	.950	1.053

a. Dependent Variable: LTT

Bảng Model Summary:

Model Summary ^e					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.582 ^a	.339	.336	1.15827	
2	.670 ^b	.449	.443	1.06066	
3	.686 ^c	.471	.463	1.04154	
4	.697 ^d	.485	.475	1.02981	1.624

a. Predictors: (Constant), TL

b. Predictors: (Constant), TL, LD

c. Predictors: (Constant), TL, LD, DT

d. Predictors: (Constant), TL, LD, DT, DANHIEUMH

e. Dependent Variable: LTT

Cũng như ở trên, kết quả phân tích cho thấy $R^2 = 0.475$ nên mối quan hệ giữa các biến TL, LD, DT, DANHIEUMH và LTT là trung bình ($0 < R^2 < 0.5$) hay nói khác hơn bốn biến TL, LD, DT, DANHIEUMH giải thích được 47.5% lòng trung thành của nhân viên đối với tổ chức. Như vậy, $1 - R^2 = 0.525$ được giải thích bởi những nhân tố không được kiểm soát trong mô hình và đây cũng có thể xem là một trong những hạn chế của nghiên cứu.

Model Summary^e

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.582 ^a	.339	.336	1.15827	
2	.670 ^b	.449	.443	1.06066	
3	.686 ^c	.471	.463	1.04154	
4	.697 ^d	.485	.475	1.02981	1.624

a. Predictors: (Constant), TL

b. Predictors: (Constant), TL, LD

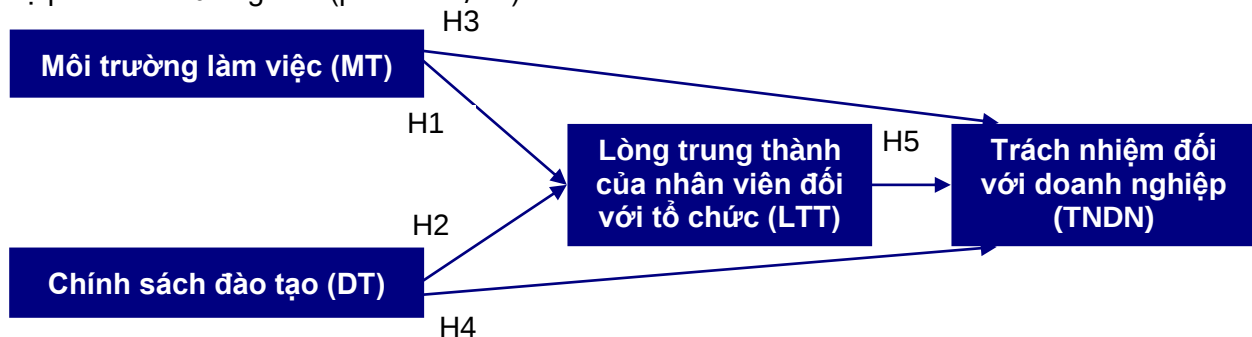
c. Predictors: (Constant), TL, LD, DT

d. Predictors: (Constant), TL, LD, DT, DANHHIEUMH

e. Dependent Variable: LTT

PHÂN TÍCH ĐƯỜNG DẪN (PATH ANALYSIS)

Ở ví dụ trên, các biến TL, PL, MT, DN, LD là biến độc lập và biến LTT, tuy nhiên, trong một số trường hợp, một biến có thể đóng vai trò là biến độc lập trong khía cạnh này nhưng cũng có thể đóng vai trò là biến độc lập trong khía cạnh khác. Ví dụ trong mô hình sau, biến lòng trung thành của nhân viên đối với tổ chức là biến phụ thuộc trong mối quan hệ với các biến môi trường làm việc và chính sách đào tạo nhưng đóng vai trò là biến độc lập trong mối quan hệ với biến trách nhiệm đối với doanh nghiệp. Trong những trường hợp như vậy, chúng ta cần xem xét dưới góc độ phân tích đường dẫn (path analysis).



Trong phân tích đường dẫn, hồi quy bội thường được sử dụng cùng với lý thuyết nhân quả, với mục đích mô tả toàn bộ cấu trúc của mối liên kết giữa biến độc lập và biến phụ thuộc được thiết kế từ nghiên cứu lý thuyết. Trong nghiên cứu này, chúng ta phải thực hiện lần lượt hai phân tích hồi quy tuyến tính:

- Mối quan hệ giữa môi trường làm việc, chính sách đào tạo với lòng trung thành của nhân viên với tổ chức.
- Mối quan hệ giữa môi trường làm việc, chính sách đào tạo, lòng trung thành của nhân viên đối với tổ chức với trách nhiệm đối với doanh nghiệp.

CHƯƠNG

9.4

PHÂN TÍCH PHƯƠNG SAI NHIỀU CHIỀU (FACTORIAL ANALYSIS OF VARIANCE)

Phân tích phương sai hai hoặc nhiều chiều (phân tích thừa số - Factorial Analysis of Variance) là một sự mở rộng của phân tích phương sai một chiều (One-way ANOVA) liên quan đến việc phân tích của hai hay nhiều biến độc lập. Phương pháp này được sử dụng trong các thiết kế nghiên cứu thực nghiệm nơi mà mỗi cấp độ của mỗi nhân tố kết hợp theo từng cặp với mỗi cấp độ của nhân tố khác, nó cho phép những người nghiên cứu ước lượng tác động của mỗi hệ số độc lập một cách tách biệt, cũng giống như tác động chung hay sự tương tác của các biến số. Có thể minh họa một cách cụ thể một mẫu thừa số với hai biến số độc lập (Ví dụ, giới tính và danh hiệu) và hai cấp độ cho mỗi biến số độc lập (giới tính: nam/ nữ; danh hiệu: có/ không) được gọi là một thừa số hai chiều hay phân tích thừa số 2×2 .

YÊU CẦU

- Trong bất kỳ một phân tích nào cũng phải có hai hoặc nhiều biến số độc lập.
- Có thể có hai (hay nhiều cấp độ) cho mỗi biến số độc lập.
- Chỉ có thể có 1 biến số phụ thuộc.

GIẢ ĐỊNH

- Tổng thể mà từ đó mẫu nghiên cứu được lựa chọn phải có phân phối chuẩn,
- Phương sai đồng nhất.
- Các quan sát phải độc lập.

PHÂN TÍCH THỪA SỐ (TWO-WAY ANOVA)

Vấn đề nghiên cứu

Một nhà nghiên cứu quan tâm đến ảnh hưởng của việc sử dụng ba chiến lược quản lý và mức độ linh hoạt của cán bộ công nhân viên:

- Định hướng vào công việc,
- Định hướng vào nhân viên,
- Định hướng vào quy trình,
- Mức độ linh hoạt cao,
- Mức độ linh hoạt trung bình
- Mức độ linh hoạt thấp

đến sự hài lòng của của cán bộ công nhân viên.

Giải sử biến số sự hài lòng (HAILONG) được tính điểm trung bình của 3 items đo lường bởi thang đo Likert (7 sự lựa chọn). Kết quả khảo sát và tính toán được lưu trữ tại file THUC HANH SPSS.sav

Giả thuyết

Bài toán trên có 3 giả thuyết như sau:

- Việc sử dụng ba chiến lược quản lý A – Định hướng vào công việc, B – Định hướng vào nhân viên và C – Định hướng vào quy trình tạo ra sự hài lòng của cán bộ công nhân viên là như nhau (bằng nhau).
- Việc sử dụng ba mức độ linh hoạt (mức độ linh hoạt cao, trung bình và thấp) tạo ra sự hài lòng của cán bộ công nhân viên là như nhau (bằng nhau).
- Ảnh hưởng của chiến lược quản lý (định hướng vào công việc, định hướng vào nhân viên, định hướng vào quy trình) đến sự hài lòng của cán bộ công nhân viên là như nhau đối với ba mức độ linh hoạt. Ảnh hưởng của việc áp dụng ba mức độ linh hoạt (cao, trung bình và thấp) đến sự hài lòng của cán bộ công nhân viên là như nhau đối với ba chiến lược quản lý. Hay nói một cách đơn giản là không có sự tương tác trong việc áp dụng chiến lược và mức độ linh hoạt đối với cán bộ công nhân viên.

Chiến lược (CHIENLUOC)	1	Định hướng vào công việc
	2	Định hướng vào nhân viên
	3	Định hướng vào quy trình
Mức độ linh hoạt (LINHHOAT)	1	Mức độ linh hoạt nhiều
	2	Mức độ linh hoạt trung bình
	3	Mức độ linh hoạt thấp

Khi đó có sự kiến hợp giữa chiến lược và mức độ linh hoạt gồm có:

- Định hướng vào công việc – mức độ linh hoạt cao,
- Định hướng vào công việc – mức độ linh hoạt trung bình,
- Định hướng vào công việc – mức độ linh hoạt thấp,
- Định hướng vào nhân viên – mức độ linh hoạt cao,
- Định hướng vào nhân viên – mức độ linh hoạt trung bình,
- Định hướng vào nhân viên – mức độ linh hoạt ít,
- Định hướng vào quy trình – mức độ linh hoạt cao,
- Định hướng vào quy trình – mức độ linh hoạt trung bình,
- Định hướng vào quy trình – mức độ linh hoạt ít.

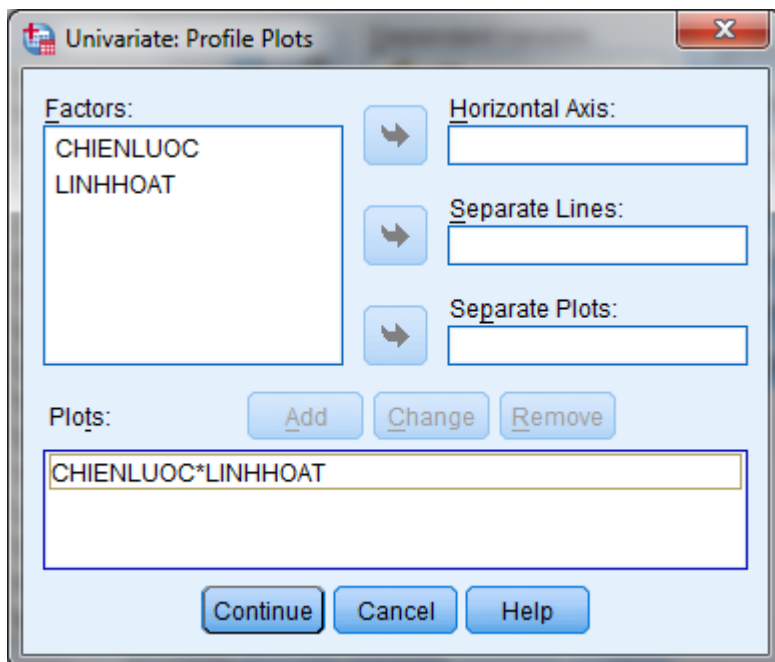
Thực hiện trên SPSS

- Trên thanh menu, vào **Analyze / General Linear Model / Univariate**.
- Tại cửa sổ **Univariate**, chọn biến độc lập (biến định lượng) là biến HAILONG vào ô



Dependent Variable, và các biến định tính phân loại gồm biến chiến lược quản lý (CHIENLUOC) và biến linh hoạt (LINHHOAT) vào ô **Fixed Factor(s)**

- Nhấp vào **Plots** để vẽ 1 biểu đồ cho tương tác CHIENLUOC*LINHHOAT bằng cách chọn LINHHOAT vào ô **Separate Lines** và CHIENLUOC vào ô **Horizontal Axis**. Sau đó bấm **Add**, biểu hiện tương tác CHIENLUOC*LINHHOAT sẽ hiện lên như hình vẽ. Kích vào **Continue** để tiếp tục.



- Bấm vào **Option** để thiết đặt hiện thị tham số trung bình và các thống kê khác

- Chọn các biến CHIENLUOC, LINHHOAT, tương tác CHIENLUOC*LINHHOAT từ ô **Factor(s) and Factor Interactions** vào ô **Display Means for:**
- Chọn các ô **Compare main effects, Descriptive statistics, Estimates of effect size, Homogeneity test** như hình vẽ, kích **Continue**

Univariate: Options

Estimated Marginal Means

Factor(s) and Factor Interactions:
 (OVERALL)
 CHIENLUOC
 LINHHOAT
 CHIENLUOC*LINHHOAT

Display Means for:
 CHIENLUOC
 LINHHOAT
 CHIENLUOC*LINHHOAT

☒ Compare main effects

Confidence interval adjustment:
 LSD(none)

Display

☒ Descriptive statistics
☒ Estimates of effect size
☐ Observed power
☐ Parameter estimates
☐ Contrast coefficient matrix

☒ Homogeneity tests
☐ Spread vs. level plot
☐ Residual plot
☐ Lack of fit
☐ General estimable function

Significance level: .05 Confidence intervals are 95.0 %

Continue Cancel Help

- Kết quả và diễn giải:

Between-Subjects Factors			
		Value Label	N
CHIENLUOC	1.00	A- Định hướng vào công việc	63
	2.00	B - Định hướng vào nhân viên	66
	3.00	C- Định hướng vào quy trình	71
LINHHOAT	1.00	X- Linh hoạt cao	63
	2.00	Y - Linh hoạt trung bình	63
	3.00	Z- Linh hoạt thấp	74

Biểu Between-Subjects Factors biểu hiện số người cán bộ công nhân viên tác động bởi ba chiến lược và ba mức độ linh hoạt trong thực hiện công việc.

Descriptive Statistics

Dependent Variable:HAILONG

CHIENLUOC	LINHHOAT	Mean	Std. Deviation	N
1.00	1.00	4.6143	.36508	21
	2.00	5.0429	.25014	21
	3.00	5.1286	.27775	21
	Total	4.9286	.37349	63
2.00	1.00	4.7571	.34143	21
	2.00	5.4000	.27928	21
	3.00	5.9208	.23770	24
	Total	5.3848	.55974	66
3.00	1.00	4.7286	.48902	21
	2.00	5.7571	.35436	21
	3.00	6.3517	.33447	29
	Total	5.6958	.78037	71
Total	1.00	4.7000	.40201	63
	2.00	5.4000	.41504	63
	3.00	5.8649	.57706	74
	Total	5.3515	.67771	200

Biểu Descriptive Statistics biểu hiện tham số trung bình về lòng trung thành khi thực hiện sự phối hợp giữa chiến lược và mức độ linh hoạt như chiến lược định hướng công việc – linh hoạt cao (4.6143), chiến lược định hướng công việc – linh hoạt trung bình (5.1286),...

Levene's Test of Equality of Error Variances^a

Dependent Variable:LTT

F	df1	df2	Sig.
.085	3	196	.968

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + CHIENLUOC + LINHHOAT + CHIENLUOC * LINHHOAT

Kết quả kiểm định Leneve cho phép khẳng định giả định về phương sai bằng nhau không bị vi phạm.

Tests of Between-Subjects Effects

Dependent Variable:HAILONG

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	70.322 ^a	8	8.790	79.654	.000	.769
Intercept	5556.761	1	5556.761	50353.725	.000	.996
CHIENLUOC	15.673	2	7.837	71.013	.000	.426
LINHHOAT	41.396	2	20.698	187.561	.000	.663
CHIENLUOC * LINHHOAT	7.081	4	1.770	16.042	.000	.251
Error	21.078	191	.110			

Total	5819.110	200				
Corrected Total	91.400	199				

a. R Squared = .769 (Adjusted R Squared = .760)

Đồng thời, kết quả kiểm định trong biểu Tests of Between-Subjects Effects tất cả các biến chiến lược, mức độ linh hoạt, tương tác giữa chiến lược và mức độ linh hoạt đều có mối quan hệ với giá trị Fisher (F) lần lượt là 71.013; 187.561 và 16.042 với giá trị Sig. tương ứng lần lượt đều bằng $0.000 < 0.05$. Do vậy, có thể bác bỏ giả thuyết và khẳng định:

(1) Việc sử dụng ba chiến lược quản lý A – Định hướng vào công việc, B – Định hướng vào nhân viên và C – Định hướng vào quy trình tạo ra sự hài lòng của cán bộ công nhân viên là khác nhau.

(2) Việc sử dụng ba mức độ linh hoạt (mức độ linh hoạt cao, trung bình và thấp) tạo ra sự hài lòng của cán bộ công nhân viên là khác nhau.

(3) Ảnh hưởng của chiến lược quản lý (định hướng vào công việc, định hướng vào nhân viên, định hướng vào quy trình) đến sự hài lòng của cán bộ công nhân viên là khác nhau đối với ba mức độ linh hoạt. Ảnh hưởng của việc áp dụng ba mức độ linh hoạt (cao, trung bình và thấp) đến sự hài lòng của cán bộ công nhân viên là khác nhau đối với ba chiến lược quản lý. Hay có thể khẳng định có sự tương tác trong việc áp dụng chiến lược và mức độ linh hoạt đối với cán bộ công nhân viên.

So sánh trung bình tổng thể về các chiến lược quản lý:

Estimates

Dependent Variable:HAILONG

Chiến lược	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1.00	4.929	.042	4.846	5.011
2.00	5.359	.041	5.279	5.440
3.00	5.612	.040	5.534	5.691

Pairwise Comparisons

Dependent Variable:HAILONG

(I) Chiến lược	(J) Chiến lược	Mean Difference (I-J)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^a	
					Lower Bound	Upper Bound
1.00	2.00	-.431*	.059	.000	-.546	-.315
	3.00	-.684*	.058	.000	-.798	-.570
2.00	1.00	.431*	.059	.000	.315	.546
	3.00	-.253*	.057	.000	-.366	-.140
3.00	1.00	.684*	.058	.000	.570	.798
	2.00	.253*	.057	.000	.140	.366

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Dễ dàng nhận thấy tất cả các hệ số Sig. đều nhỏ hơn 0.05 nên có thể kết luận có sự khác nhau về sự hài lòng của các chiến lược quản lý (định hướng vào công việc, nhân viên hay quy trình). Sự sai biệt cụ thể có thể rút ra từ cột Mean Difference (I-J). Có thể giải thích tương tự cho các trường hợp mức độ linh hoạt.

So sánh trung bình tổng thể về mức độ linh hoạt:

Estimates

Dependent Variable:HAILONG

Mức độ linh hoạt	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1.00	4.700	.042	4.617	4.783
2.00	5.400	.042	5.317	5.483
3.00	5.800	.039	5.724	5.877

Pairwise Comparisons

Dependent Variable:HAILONG

(I) Mức độ linh hoạt	(J) Mức độ linh hoạt	Mean Difference (I-J)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^a	
					Lower Bound	Upper Bound
1.00	2.00	-.700*	.059	.000	-.817	-.583
	3.00	-1.100*	.057	.000	-1.213	-.988
2.00	1.00	.700*	.059	.000	.583	.817
	3.00	-.400*	.057	.000	-.513	-.288
3.00	1.00	1.100*	.057	.000	.988	1.213
	2.00	.400*	.057	.000	.288	.513

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Tất cả các hệ số Sig. đều nhỏ hơn 0.05 nên có thể kết luận có sự khác nhau về sự hài lòng của các mức độ linh hoạt và dĩ nhiên, nhận biết sự sai biệt có thể rút ra từ cột Mean Difference (I-J).

Tương tác giữa chiến lược và mức độ linh hoạt:

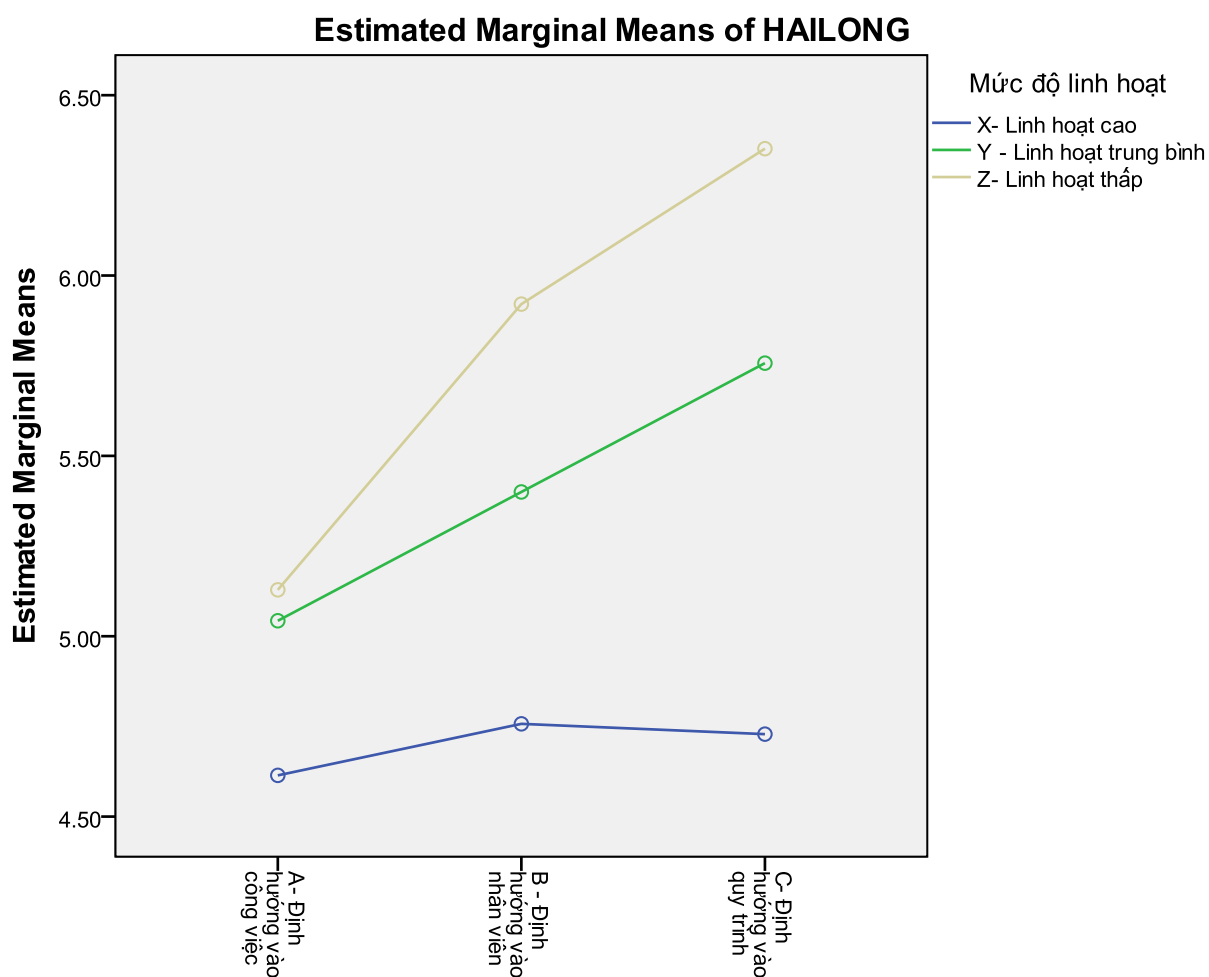
Điểm trung bình trong tương tác giữa chiến lược và mức độ linh hoạt được biểu hiện ở biểu sau:

4. Chiến lược * Mức độ linh hoạt

Dependent Variable:HAILONG

Chiến lược	Mức độ linh hoạt	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound

1.00	1.00	4.614	.072	4.471	4.757
	2.00	5.043	.072	4.900	5.186
	3.00	5.129	.072	4.986	5.272
2.00	1.00	4.757	.072	4.614	4.900
	2.00	5.400	.072	5.257	5.543
	3.00	5.921	.068	5.787	6.055
3.00	1.00	4.729	.072	4.586	4.872
	2.00	5.757	.072	5.614	5.900
	3.00	6.352	.062	6.230	6.473



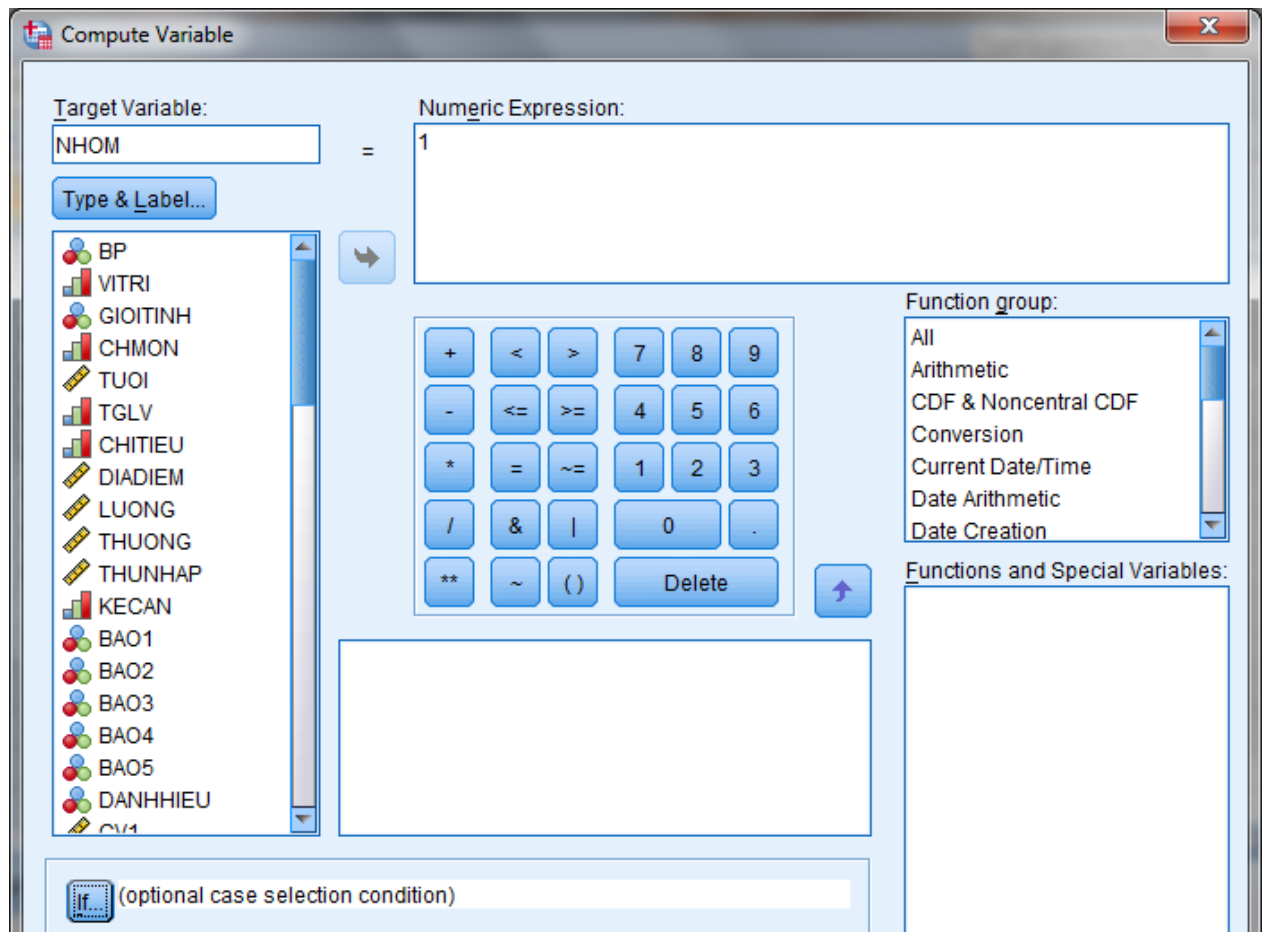
Như đã đề cập ở phần trên, chúng ta có 9 cặp phối hợp giữa các chiến lược và mức độ linh hoạt, để tiến hành nghiên cứu sự phối hợp này, chúng ta sẽ tạo ra biến nhóm (NHOM) bao gồm 9 nhóm phối hợp sau:

- Định hướng vào công việc – mức độ linh hoạt cao (nhóm = 1),
- Định hướng vào công việc – mức độ linh hoạt trung bình (nhóm = 2),
- Định hướng vào công việc – mức độ linh hoạt thấp (nhóm = 3),
- Định hướng vào nhân viên – mức độ linh hoạt cao (nhóm = 4),

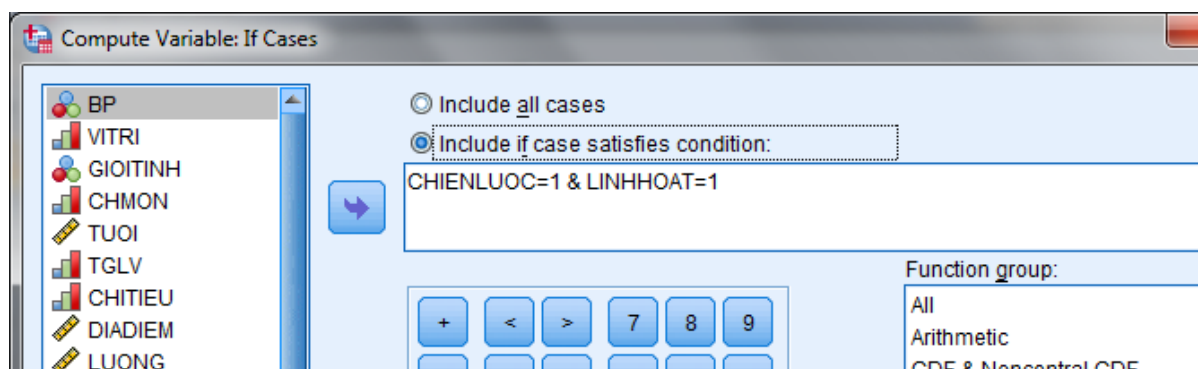
- Định hướng vào nhân viên – mức độ linh hoạt trung bình (nhóm = 5),
- Định hướng vào nhân viên – mức độ linh hoạt ít (nhóm = 6),
- Định hướng vào quy trình – mức độ linh hoạt cao (nhóm = 7)
- Định hướng vào quy trình – mức độ linh hoạt trung bình (nhóm = 8)
- Định hướng vào quy trình – mức độ linh hoạt ít (nhóm = 9)

ạo lập biến nhóm gồm các phối hợp trên SPSS như sau:

- Trên thanh menu, vào **Transform / Compute Variable**, khi đó hộp **Compute Variable** được mở ra. Tạo biến NHOM ở **Target Variable** với giá trị 1 ở **Numeric Expression**.



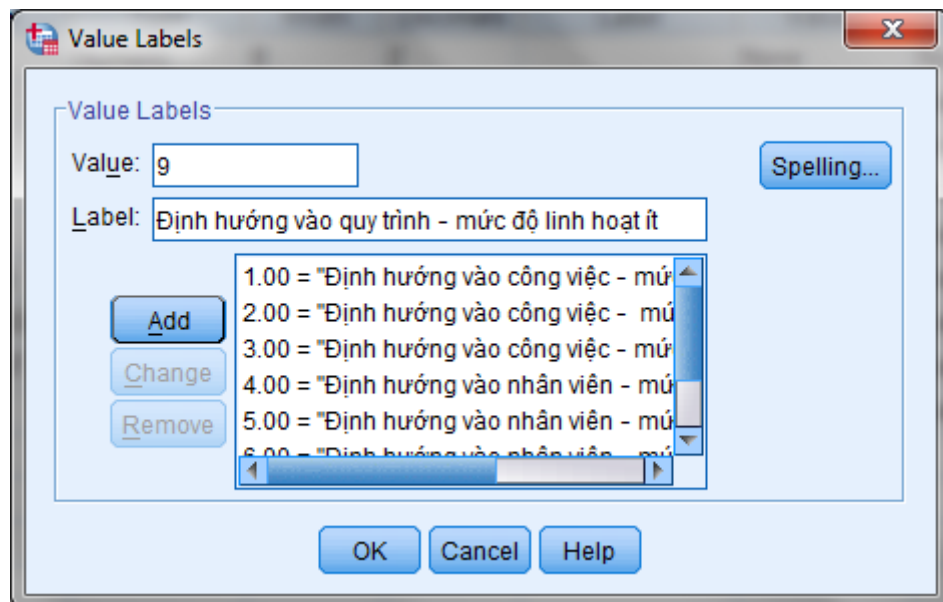
- Kích vào **If** để thiết đặt điều kiện,
- Kích vào **Include if case satisfies condition** và thiết đặt điều kiện với $CHIENTLUOC = 1$ & $LINHHOAT = 1$ (nghĩa là chiến lược định hướng vào công việc – mức độ linh hoạt cao).



- Kích vào Continue để hoàn thành và trở về cửa sổ **Compute Variable**.
- Bấm OK để hoàn thành việc thiết đặt nhóm có giá trị = 1 (chiến lược định hướng vào công việc – mức độ linh hoạt cao).
- Thực hiện tương tự như vậy với nhóm có giá trị từ 2 đến 9, vì vậy trong **Include if case satisfies condition** và thiết đặt điều kiện với $CHIENTLUOC = i$ & $LINHHOAT = j$ (với i và j lần lượt là các giá trị 1, 2, 3 của biến CHIENTLUOC và LINHHOAT).

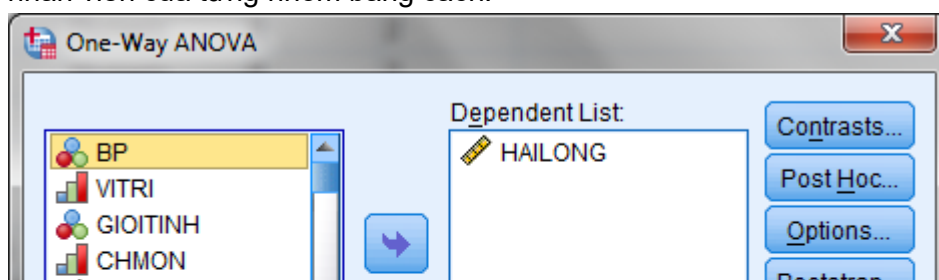
Chú ý: Với mỗi lần tạo nhóm mới, sau khi bấm **OK**, sẽ có xuất hiện hộp hội thoại **Change Exit Variable**, chúng ta bấm **OK** để thêm giá trị trong biến NHOM.

- Sau đó đặt tên và gán các giá trị của các biến NHOM (có 9 giá trị tương ứng với các cặp phối hợp).



So sánh Post Hoc (Post Hoc Comparisons)

Thực hiện One- Way ANOVA để so sánh tham số trung bình về hài lòng của cán bộ công nhân viên của từng nhóm bằng cách:



Kết quả như sau:

ANOVA

HAILONG

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	70.322	8	8.790	79.654	.000
Within Groups	21.078	191	.110		
Total	91.400	199			

Vì $F = 79.654$ và $p(F) = 0.000 < 0.05$ nên khẳng định tồn tại sự khác biệt trung bình về sự hài lòng của các nhóm. So sánh từng cặp được thực hiện trong biểu **Multile Comparisons**, giải thích sự so sánh theo cặp như đã trình bày trong phần One – Way ANOVA.

Multiple Comparisons

HAILONG

Bonferroni

(I) NHOM	(J) NHOM	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-.42857*	.10252	.002	-.7612	-.0959
	3.00	-.51429*	.10252	.000	-.8469	-.1817
	4.00	-.14286	.10252	1.000	-.4755	.1898
	5.00	-.78571*	.10252	.000	-1.1183	-.4531
	6.00	-1.30655*	.09926	.000	-1.6286	-.9845
	7.00	-.11429	.10252	1.000	-.4469	.2183
	8.00	-1.14286*	.10252	.000	-1.4755	-.8102
	9.00	-1.73744*	.09519	.000	-2.0463	-1.4286
2.00	1.00	.42857*	.10252	.002	.0959	.7612
	3.00	-.08571	.10252	1.000	-.4183	.2469
	4.00	.28571	.10252	.211	-.0469	.6183
	5.00	-.35714*	.10252	.022	-.6898	-.0245
	6.00	-.87798*	.09926	.000	-1.2000	-.5559
	7.00	.31429	.10252	.090	-.0183	.6469
	8.00	-.71429*	.10252	.000	-1.0469	-.3817
	9.00	-1.30887*	.09519	.000	-1.6177	-1.0000
3.00	1.00	.51429*	.10252	.000	.1817	.8469

	2.00	.08571	.10252	1.000	-.2469	.4183
	4.00	.37143*	.10252	.013	.0388	.7041
	5.00	-.27143	.10252	.316	-.6041	.0612
	6.00	-.79226*	.09926	.000	-1.1143	-.4702
	7.00	.40000*	.10252	.005	.0674	.7326
	8.00	-.62857*	.10252	.000	-.9612	-.2959
	9.00	-1.22315*	.09519	.000	-1.5320	-.9143
4.00	1.00	.14286	.10252	1.000	-.1898	.4755
	2.00	-.28571	.10252	.211	-.6183	.0469
	3.00	-.37143*	.10252	.013	-.7041	-.0388
	5.00	-.64286*	.10252	.000	-.9755	-.3102
	6.00	-1.16369*	.09926	.000	-1.4858	-.8416
	7.00	.02857	.10252	1.000	-.3041	.3612
	8.00	-1.00000*	.10252	.000	-1.3326	-.6674
	9.00	-1.59458*	.09519	.000	-1.9034	-1.2857
5.00	1.00	.78571*	.10252	.000	.4531	1.1183
	2.00	.35714*	.10252	.022	.0245	.6898
	3.00	.27143	.10252	.316	-.0612	.6041
	4.00	.64286*	.10252	.000	.3102	.9755
	6.00	-.52083*	.09926	.000	-.8429	-.1988
	7.00	.67143*	.10252	.000	.3388	1.0041
	8.00	-.35714*	.10252	.022	-.6898	-.0245
	9.00	-.95172*	.09519	.000	-1.2606	-.6429
6.00	1.00	1.30655*	.09926	.000	.9845	1.6286
	2.00	.87798*	.09926	.000	.5559	1.2000
	3.00	.79226*	.09926	.000	.4702	1.1143
	4.00	1.16369*	.09926	.000	.8416	1.4858
	5.00	.52083*	.09926	.000	.1988	.8429
	7.00	1.19226*	.09926	.000	.8702	1.5143
	8.00	.16369	.09926	1.000	-.1584	.4858
	9.00	-.43089*	.09167	.000	-.7283	-.1335
7.00	1.00	.11429	.10252	1.000	-.2183	.4469
	2.00	-.31429	.10252	.090	-.6469	.0183
	3.00	-.40000*	.10252	.005	-.7326	-.0674
	4.00	-.02857	.10252	1.000	-.3612	.3041
	5.00	-.67143*	.10252	.000	-1.0041	-.3388
	6.00	-1.19226*	.09926	.000	-1.5143	-.8702
	8.00	-1.02857*	.10252	.000	-1.3612	-.6959
	9.00	-1.62315*	.09519	.000	-1.9320	-1.3143
8.00	1.00	1.14286*	.10252	.000	.8102	1.4755
	2.00	.71429*	.10252	.000	.3817	1.0469

	3.00	.62857*	.10252	.000	.2959	.9612
	4.00	1.00000*	.10252	.000	.6674	1.3326
	5.00	.35714*	.10252	.022	.0245	.6898
	6.00	-.16369	.09926	1.000	-.4858	.1584
	7.00	1.02857*	.10252	.000	.6959	1.3612
	9.00	-.59458*	.09519	.000	-.9034	-.2857
9.00	1.00	1.73744*	.09519	.000	1.4286	2.0463
	2.00	1.30887*	.09519	.000	1.0000	1.6177
	3.00	1.22315*	.09519	.000	.9143	1.5320
	4.00	1.59458*	.09519	.000	1.2857	1.9034
	5.00	.95172*	.09519	.000	.6429	1.2606
	6.00	.43089*	.09167	.000	.1335	.7283
	7.00	1.62315*	.09519	.000	1.3143	1.9320
	8.00	.59458*	.09519	.000	.2857	.9034

*. The mean difference is significant at the 0.05 level.

CHƯƠNG

PHÂN TÍCH TÁCH BIỆT VÀ HỒI QUY BINARY LOGISTIC

9.5

(DISCRIMINANT ANALYSIS VÀ BINARY LOGISTIC REGRESSION)

PHÂN TÍCH TÁCH BIỆT (DISCRIMINANT ANALYSIS)

Trong phần hồi quy đơn hồi quy bội đã giới thiệu phương pháp thống kê để phân tích mối quan hệ giữa một biến phụ thuộc định lượng với tập hợp độc lập định lượng. Phân tích tách biệt sẽ giới thiệu mối quan hệ giữa một biến phụ thuộc định tính (có hai hoặc nhiều lựa chọn) và một hoặc nhiều biến độc lập định lượng, phương trình mối quan hệ giúp cho việc lựa chọn phương pháp phân tích như sau:

Hồi quy tương quan bội (Multiple Regression Analysis)

$$\begin{array}{ccc} Y & = & X_1 + X_2 + X_3 + \dots + X_n \\ \text{(Định lượng)} & & \text{(Định lượng, định tính)} \end{array}$$

Phân tích sự tách biệt (Discriminant Analysis)

$$\begin{array}{ccc} Y & = & X_1 + X_2 + X_3 + \dots + X_n \\ \text{(Định tính)} & & \text{(Định lượng)} \end{array}$$

Phân tích tách biệt chỉ ra sự khác nhau của hai hay nhiều nhóm trên cơ sở một tập hợp các biến số, ví dụ, biến phụ thuộc có thể là giới tính, bộ phận công tác và biến độc lập có thể là các biến tiền lương, chính sách đào tạo, môi trường làm việc, công việc... ví dụ, người nghiên cứu muốn phân biệt 2 nhóm người mua hàng và không mua dựa trên một số đặc điểm của sản phẩm như giá cả, độ bền, mẫu mã; những yếu tố giúp phân biệt người có nhạy cảm với giá hoặc không...

Loại phân tích tách biệt

Có hai trường hợp phân tích tách biệt là:

- Phân tích tách biệt hai nhóm (two-group discriminant analysis)
- Phân tích tách biệt bội (nhiều hơn hai nhóm) (multiple discriminant analysis).

Một số tham số trong phân tích tách biệt:

- Centroid: là trung bình các giá trị tách biệt hay gọi là "trọng tâm" của nhóm.
- Discriminant function coefficients: Hệ số hay trọng số của biến độc lập i.
- Classification matrix: ma trận phân loại chứa các quan sát phân loại đúng và số quan sát phân loại sai.

- Discriminant scores: giá trị tách biệt (tính bằng cách nhân các hệ số không chuẩn hóa với giá trị của các biến) theo hàm tách biệt.
- Standardized discriminant function coefficients: Hệ số hàm phân biệt chuẩn hóa.
- Group means and group standard deviations: Trung bình nhóm và độ lệch chuẩn nhóm được tính cho mỗi biến dự đoán cho mỗi nhóm.
- Discriminant loadings: Hệ số tải tách biệt.
- Total correlation matrix: ma trận tương quan toàn bộ là ma trận tương quan khi các quan sát được coi như xuất phát từ một mẫu duy nhất.
- Wilks' Lamda: Hệ số Wilks' Lamda được tính bằng tỉ số giữa tổng các độ lệch bình phương trong nội bộ các nhóm và tổng các độ lệch bình phương toàn bộ nhóm. Giá trị $0 \leq \text{Wilks' Lamda} \leq 1$, nếu Wilks' Lamda càng tiến tới 1 cho biết các trung bình nhóm gần như khác nhau và ngược lại.

Các bước phân tích tách biệt



Xác định vấn đề nghiên cứu

Phân tích tách biệt nhằm mục tiêu:

- Phát triển các hàm phân biệt kết hợp tuyến tính những nhân tố dự báo (các biến độc lập), các hàm này có sự phân biệt tốt nhất giữa các tiêu chuẩn nhóm đã phân loại (biến phụ thuộc).
- Xác định xem có sự khác biệt có ý nghĩa tồn tại giữa các nhóm về nội dung của các biến độc lập.
- Xác định biến độc lập nào gây ra sự khác biệt giữa các nhóm.
- Phân loại nhóm này so với nhóm khác dựa vào các giá trị của các biến độc lập.
- Đánh giá tính chính xác của việc phân loại.

Thiết kế nghiên cứu

- Lựa chọn các biến độc lập: là các biến định lượng và tuân theo quy luật phân phối chuẩn;
- Quy mô mẫu: Với mỗi sự lựa chọn trong biết định tính sẽ làm cơ sở cho việc phân tích thành các nhóm (ví dụ nhóm người nam và nhóm người nữ, nhóm mua hàng và nhóm không mua hàng). Theo Hair và cộng sự (2009), để thực hiện phân tích tách biệt, cần phải có ít nhất 20 quan sát thuộc mỗi nhóm.
- Sau đây là so sánh về sự giống nhau và khác nhau giữa ANOVA, hồi quy và phân tích tách biệt:

	Hồi quy	ANOVA	Tách biệt
<i>Điểm giống nhau</i>			
- Số lượng biến phụ thuộc	Một	Một	Một
- Số lượng biến độc lập	Nhiều biến	Nhiều biến	Nhiều biến
<i>Điểm khác nhau</i>			
- Bản chất của biến phụ thuộc	Định lượng	Định lượng	Phân loại
- Bản chất của biến độc lập	Định lượng	Phân loại	Định lượng

Các giả định

- Các biến độc lập phải phân phối chuẩn;
- Tồn tại mối quan hệ tuyến tính;
- Không xảy ra hiện tượng tự tương quan giữa các biến độc lập;
- Các ma trận phân tán như nhau.

Hàm tách biệt

Các hàm tách biệt có dạng tuyến tính như sau:

$$Z_{jk} = a + W_1X_{1k} + W_2X_{2k} + W_3X_{3k} + \dots + W_nX_{nk}$$

Trong đó: Z_{jk} Giá trị tách biệt (discriminant Z score) của hàm tách biệt j của đối tượng k

a Hệ số chặn

W_i Hệ số hay trọng số của biến độc lập i

X_{ik} Biến độc lập i của đối tượng k.

Trong hàm tách biệt, hệ số W_i sẽ được tính sao cho các nhóm có giá trị tách biệt khác nhau càng nhiều càng tốt. Nếu kết quả có k+1 phân tích thì sẽ có k hàm tách biệt tương ứng.

Giải thích kết quả

Giải thích kết quả phụ thuộc vào từng kết quả phân tích trong đó cần chú ý đến giá trị phân biệt (cutting score).

- Điểm phân biệt của hai nhóm có quy mô khác nhau:

$$Z_{CS} = \frac{N_A Z_B + N_B Z_A}{N_A + N_B}$$

- Điểm phân biệt của hai nhóm có quy mô bằng nhau:

$$Z_{CS} = \frac{Z_A + Z_B}{2}$$

Trong đó: Z_{CS} Giá trị (điểm) phân biệt.

N_A Quy mô (số quan sát) thuộc nhóm A

N_B Quy mô (số quan sát) thuộc nhóm B

Z_A Centroid nhóm A

Z_B Centroid nhóm B.

Đánh giá tính đúng đắn của hàm tách biệt

Phân tích tách biệt sẽ chia mẫu thành hai phần, một phần để ước lượng hàm phân biệt và một phần dùng để kiểm tra tính đúng đắn của hàm phân biệt được gọi là nhóm predicted group và original. Dựa vào biểu **Classification Result** để tính tỷ lệ phần trăm số quan sát được phân loại đúng (tỷ lệ này được tính bằng tổng các công số trên đường chéo của biểu này chia cho tổng số quan sát).

Thực hành trên SPSS với phân tích tách biệt hai nhóm (two-group discriminant analysis)

Xác định vấn đề và thiết kế nghiên cứu

Vấn đề được hình thành bằng cách nhận dạng mục tiêu phân tích, biến độc lập và biến phụ thuộc. Biến phụ thuộc phải là biến phân loại (biểu danh) bao gồm hai hay nhiều nhóm được

chọn. Nếu biến phụ thuộc sử dụng thang đo khoảng hay thang đo tỷ lệ, trước tiên phải mã hóa chúng về biến phân loại. Chẳng hạn như ý kiến của cán bộ công nhân viên về việc áp dụng phương thức quản lý mới được đo lường bằng 7 cấp độ (từ 1 đến 7), cần chuyển về thang đo phân loại gồm từ 1 đến 3 là không hài lòng và 4 đến 7 là hài lòng... để đảm bảo cho việc phân tích.

Ví dụ

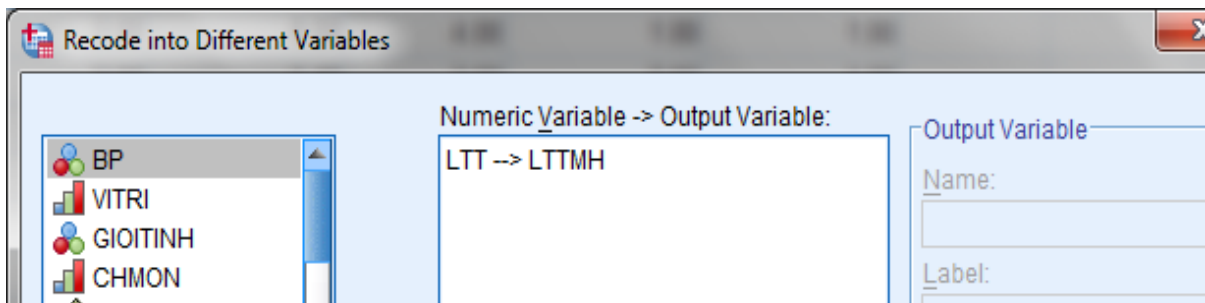
Nhà nghiên cứu muốn phân biệt cán bộ công nhân viên trung thành hay không dựa trên các biến số (1) cảm nhận về lương và thưởng; (2) phúc lợi; (3) môi trường làm việc; (4) đồng nghiệp; (5) phong cách lãnh đạo, (6) chính sách đào tạo của doanh nghiệp.

Khi đó biến lòng trung thành của cán bộ công nhân (LTT) được đo lường bởi thang điểm Likert (7 lựa chọn) sẽ được mã hóa thành biến (LTTMH) là thang đo định tính với các giá trị 1 – không trung thành (nhận giá trị từ 0 đến cận 4) và 2 – trung thành nhận giá trị từ 4 đến 7.

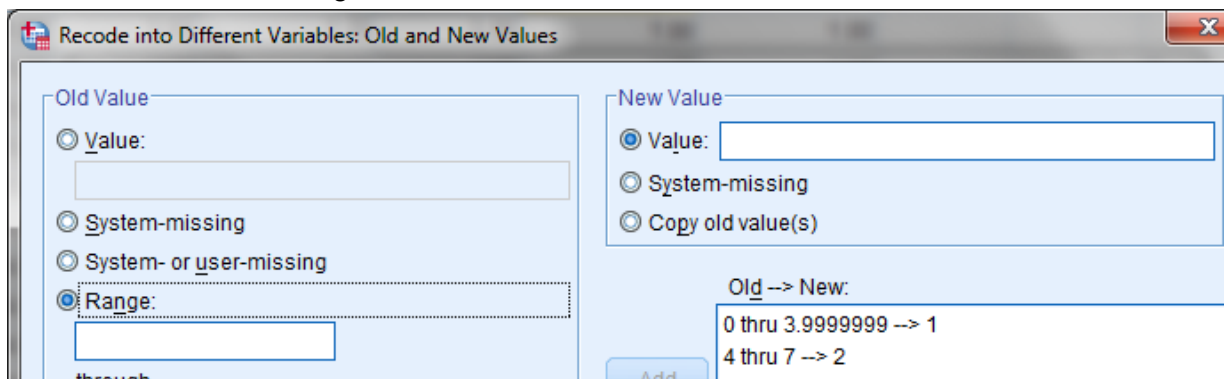
Các biến (1) cảm nhận về lương và thưởng; (2) phúc lợi; (3) môi trường làm việc; (4) đồng nghiệp; (5) phong cách lãnh đạo, (6) chính sách đào tạo của doanh nghiệp là biến định lượng (thang đo tỷ lệ - Likert 7 sự lựa chọn) được giả định phân phối chuẩn.

Thực hành trên SPSS

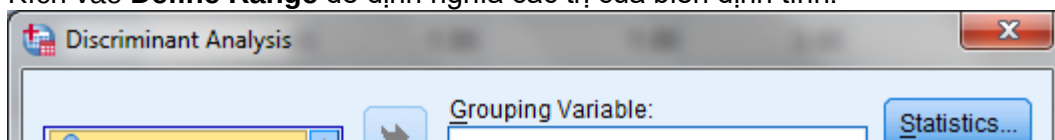
- Mã hóa thang đo lòng trung thành: Vào **Analyze / Recode into Different Variable**,
- + Mã hóa biến LTT thành LTTMH:



- + Biến LTTMH nhận các giá trị như sau:

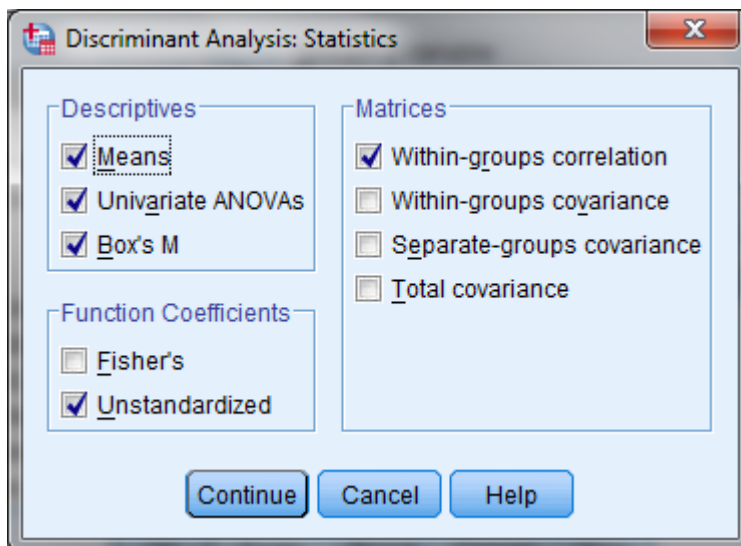


- Phân tích tách biệt bằng cách **Analyze / Classify / Discriminant** chọn các biến **DT, TL, LD, DN, PL, MT** vào ô **Independents** và biến LTTMH vào ô **Grouping Variable**.
- Kích vào **Define Range** để định nghĩa các trị của biến định tính.

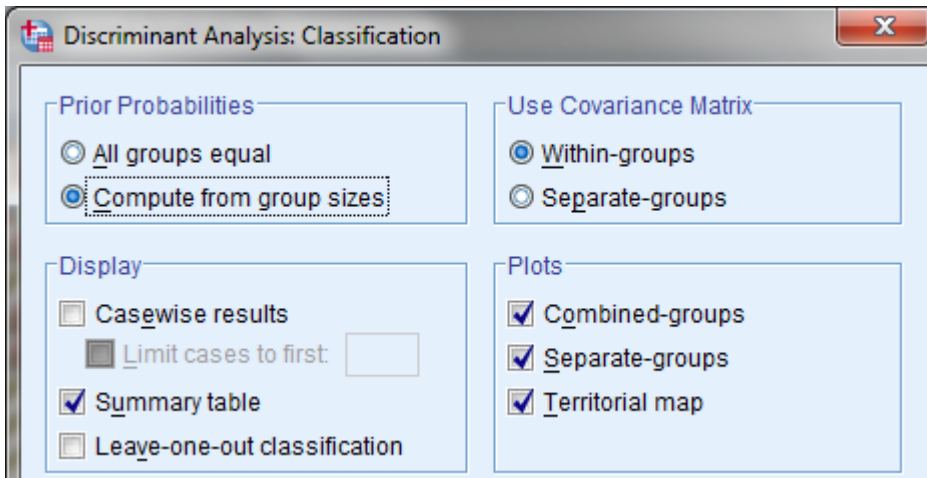


\

- Kích vào **Statistic** và thiết đặt các mục chọn như hình vẽ:



- Tương tự thiết đặt ở **Classify** như hình vẽ và kích **Continue**.



- Kích **OK** để hoàn tất việc thiết đặt.

- Kết quả và diễn giải:

Group Statistics					
Lồng trung thành (mã hóa)		Mean	Std. Deviation	Valid N (listwise)	
				Unweighted	Weighted
1.00	DT	4.3800	1.21924	70	70.000
	TL	3.3000	.93793	70	70.000
	LD	3.7629	1.29225	70	70.000
	DN	4.9143	1.29247	70	70.000
	PL	5.0500	1.83396	70	70.000
	MT	3.8071	1.54003	70	70.000
2.00	DT	5.3569	.94598	130	130.000
	TL	4.6477	1.00566	130	130.000
	LD	4.9031	.94080	130	130.000
	DN	5.4897	1.10452	130	130.000
	PL	5.4000	1.55693	130	130.000
	MT	4.3692	1.46080	130	130.000
Total	DT	5.0150	1.14619	200	200.000
	TL	4.1760	1.17299	200	200.000
	LD	4.5040	1.20417	200	200.000
	DN	5.2883	1.20240	200	200.000
	PL	5.2775	1.66300	200	200.000
	MT	4.1725	1.50927	200	200.000

Biểu **Group Statistic** sẽ thống kê giá trị trung bình của các biến số độc lập trong mô hình phân tích phân chia theo các nhóm, ví dụ nhóm không trung thành và nhóm trung thành có giá trị trung bình của biến DT lần lượt là 4.380 và 5.015; biến TL có giá trị trung bình lần lượt là 3.330 và 4.176...

Pooled Within-Groups Matrices							
		DT	TL	LD	DN	PL	MT
Correlation	DT	1.000	.289	.496	.360	.245	-.081
	TL	.289	1.000	.286	.188	.247	-.019
	LD	.496	.286	1.000	.364	.176	-.083
	DN	.360	.188	.364	1.000	.208	.047
	PL	.245	.247	.176	.208	1.000	.026
	MT	-.081	-.019	-.083	.047	.026	1.000

Qua ma trận tương quan trong nội bộ các nhóm (Pooled Within-Groups Matrices) cho thấy, các quan hệ giữa các biến độc lập thấp nên có thể kết luận không xảy ra hiện tượng đa cộng tuyến.

Tests of Equality of Group Means					
	Wilks' Lambda	F	df1	df2	Sig.

DT	.834	39.438	1	198	.000
TL	.698	85.596	1	198	.000
LD	.795	51.057	1	198	.000
DN	.948	10.942	1	198	.001
PL	.990	2.026	1	198	.156
MT	.968	6.485	1	198	.012

Vì Sig. của các biến DT, TL, LD, DN, MT có giá trị nhỏ hơn 0.05 nên có thể khẳng định có sự phân biệt (có ý nghĩa) giữa những người trung thành và không trung thành theo những biến này, như vậy, chỉ tồn tại duy nhất biến PL là không có ý nghĩa vì $F = 2.026$ và $\text{Sig.} = 0.156 > 0.05$).

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	.632 ^a	100.0	100.0	.622

a. First 1 canonical discriminant functions were used in the analysis.

Giá trị Eigenvalue = 0.632 và chiếm 100% phương sai giải thích nguyên nhân. Hệ số Canonical Correlation = 0.622 nên $(\text{Canonical Correlation})^2 = 0.3868 = 38.68\%$ nghĩa là 38.68% phương sai biến phụ thuộc (LTTMH) được giải thích bởi mô hình.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.613	95.493	6	.000

Vì Wilks' Lambda = 0.613, Chi-square = 95.493 với 6 bậc tự do (df) nên $p(\text{Chi-square}, \text{df}) = 0.000 < 0.05$ nên có thể khẳng định có sự phân biệt giữa hai nhóm trung thành và không trung thành bởi các biến độc lập.

Canonical Discriminant Function Coefficients

	Function
	1
DT	.238
TL	.714
LD	.359
DN	-.038
PL	-.104
MT	.201
(Constant)	-5.878

Unstandardized coefficients

Vì có hai nhóm (trung thành và không trung thành) nên chỉ tồn tại một hàm tách biệt, phương trình lý thuyết của hàm tách biệt là:

$$Z = a + W_1DT + W_2TL + W_3LD + W_4DN + W_5PL + W_6MT$$

Khi đó, phương trình thực tiễn được biểu diễn:

$$Z = -5.878 + 0.238DT + 0.714TL + 0.359LD - 0.038DN + 0.201MT.$$

Functions at Group Centroids

Lồng trung thành (mã hóa)	Function
	1
1.00	-1.078
2.00	.580

Unstandardized canonical discriminant functions evaluated at group means

Thông qua biểu Function at Group Centroids, giá trị Centroid của nhóm 1 là -1.078 và nhóm 2 à 0.580 nên điểm phân biệt giữa hai nhóm là:

$$Z_{CS} = \frac{N_A Z_B + N_B Z_A}{N_A + N_B} = \frac{70 * 0.580 + 130 * (-1.078)}{70 + 130} = -0.4977$$

Classification Results^{b,c}

		Lồng trung thành (mã hóa)	Predicted Group Membership		Total
			1.00	2.00	
Original	Count	1.00	49	21	70
		2.00	17	113	130
	%	1.00	70.0	30.0	100.0
		2.00	13.1	86.9	100.0
Cross-validated ^a	Count	1.00	48	22	70
		2.00	18	112	130
	%	1.00	68.6	31.4	100.0
		2.00	13.8	86.2	100.0

a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

b. 81.0% of original grouped cases correctly classified.

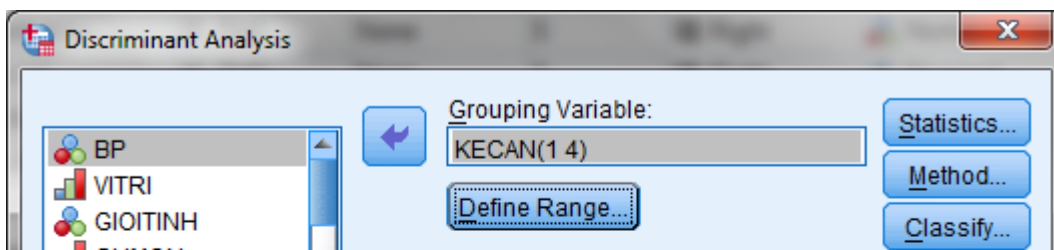
c. 80.0% of cross-validated grouped cases correctly classified.

Như vậy, tỷ lệ phân biệt đúng là $(49+113)/200 = 81\%$, thực hiện phân tích biệt số trên mẫu kiểm tra tỷ lệ phân biệt là $(48+112)/200 = 80\%$ hơi thấp so với 81%.

Thực hành trên SPSS với phân tích tách biệt bội (multiple discriminant analysis) với phương pháp stepwise

Nhà nghiên cứu muốn phân biệt cán bộ công nhân viên ở 4 loại đội ngũ kế cận dựa trên các biến số (1) cảm nhận về lương và thưởng; (2) phúc lợi; (3) môi trường làm việc; (4) đồng nghiệp; (5) phong cách lãnh đạo, (6) chính sách đào tạo của doanh nghiệp.

Thực hiện phân tích trên SPSS và giải thích giống như trên, tuy nhiên trong hộp hội thoại **Discriminant Analysis** cần chọn vào **Use stepwise method** để thực hiện bằng phương pháp stepwise.



HỒI QUY BINARY LOGISTIC (BINARY LOGISTIC REGRESSION)

Phương pháp phân tích biệt số và hồi quy Binary logistic là hai phương pháp có điểm tương đồng liên quan đến biến phụ thuộc là biến định tính. Hồi quy Binary logistic sử dụng biến phụ thuộc là một biến nhị phân nhằm ước lượng xác suất một sự kiện sẽ xảy ra trên cơ sở sự biến đổi của các biến độc lập. Đặc biệt, các nhà nghiên cứu sử dụng hồi quy Binary Logistic vì một số nguyên nhân sau:

- Yêu cầu ít giả định hơn phân tích tách biệt (đặc biệt là yêu cầu về phân phối của các biến độc lập).
- Phương thức thực hiện hồi quy thân thuộc (như các phương pháp hồi quy tuyến tính khác)
- Áp dụng được cho dữ liệu dùng để phân tích tách biệt.

Tuy nhiên, do yêu cầu ít giả định hơn nên độ tin cậy không cao, bên cạnh đó, biến định tính trong Binary Logistic chỉ có 2 lựa chọn nên trong trường hợp biến định tính có nhiều lựa chọn sẽ không sử dụng được phương pháp hồi quy này.

Mô hình hồi quy Binary Logistic

Đối với biến định tính có 2 biến lựa chọn, để sử dụng Binary Logistic cần phải chuyển về biến nhị phân có giá trị 0 là không xảy ra sự kiện và 1 là xảy ra sự kiện. Khi đó hàm hồi quy Binary Logistic như sau:

$$P_i = E(Y = 1/X) = \frac{e^{(B_0 + B_1 X)}}{1 + e^{(B_0 + B_1 X)}} \quad \text{hay} \quad \log_e \left[\frac{P_i}{1 - P_i} \right] = B_0 + B_1 X$$

Kiểm định ý nghĩa của các hệ số

Wald Chi-square là đại lượng dùng để kiểm định ý nghĩa thống kê của hệ số hồi quy tổng thể, khi B_0 và B_1 bằng 0 thì tỷ lệ chênh lệch giữa các xác suất sẽ bằng không hay xác suất để sự kiện xảy ra ($Y=1$) hay không xảy ra ($Y=0$) như nhau, khi đó hàm dự đoán không còn ý nghĩa.

Độ phù hợp của mô hình

Việc đo lường độ phù hợp của mô hình Binary Logistic dựa trên chỉ tiêu -2LL (-2 log likelihood) và khi -2LL càng nhỏ thì mô hình càng phù hợp và khi -2LL có nghĩa là mô hình hoàn hảo. *Chú ý:* Độ phù hợp của mô hình Binary Logistic (-2LL) ngược với độ phù hợp của mô hình hồi quy tuyến tính R^2 . Ngoài ra, thông qua bảng Classification Table cũng có thể xác định mô hình dự đoán tốt mức độ nào dựa trên việc so sánh trị số thực và trị số dự đoán cho từng biểu hiện và tỷ lệ dự đoán đúng sự kiện.

Kiểm định độ phù hợp tổng quát

Hồi quy Binary Logistic sẽ dùng kiểm định Chi-square để kiểm định các hệ số của mô hình khi giả thuyết rằng $B_0 = B_1 = \dots = B_k = 0$.

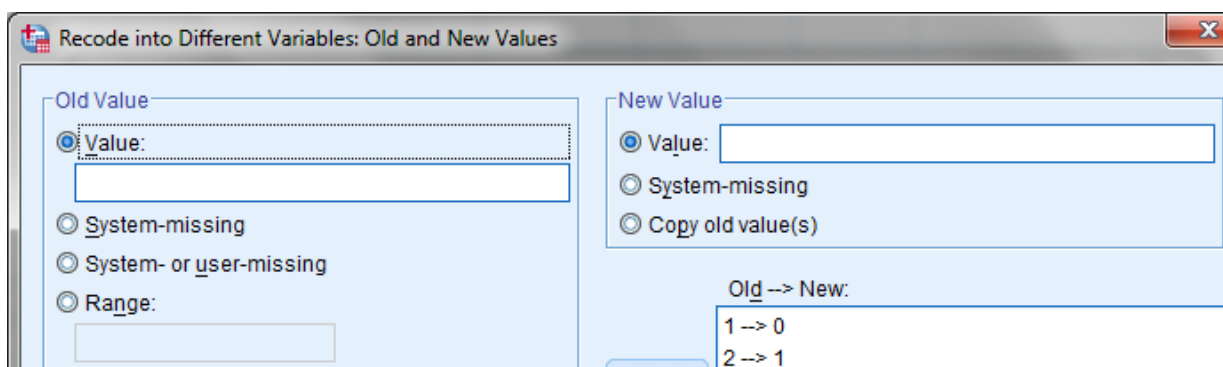
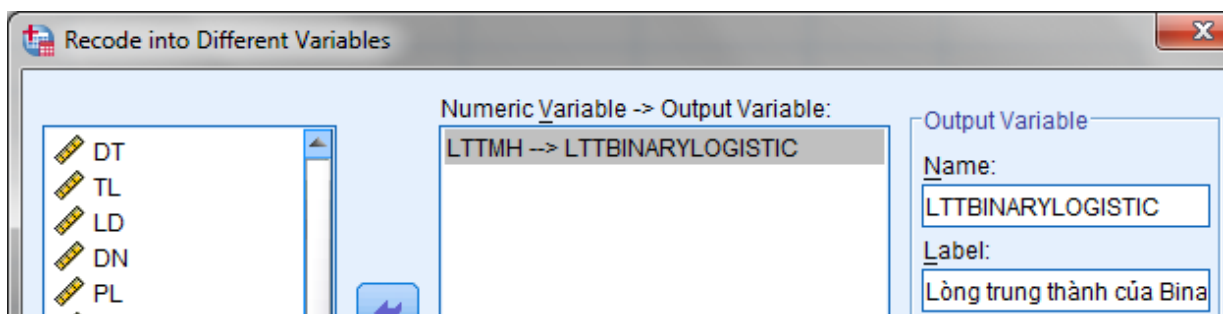
Thực hành trên SPSS với hồi quy Binary Logistic

Ví dụ

Như ví dụ ở phần phân tích tách biệt (discriminant analysis), nhà nghiên cứu muốn phân biệt cán bộ công nhân viên trung thành hay không dựa trên các biến số (1) cảm nhận về lương và thưởng; (2) phúc lợi; (3) môi trường làm việc; (4) đồng nghiệp; (5) phong cách lãnh đạo, (6) chính sách đào tạo của doanh nghiệp.

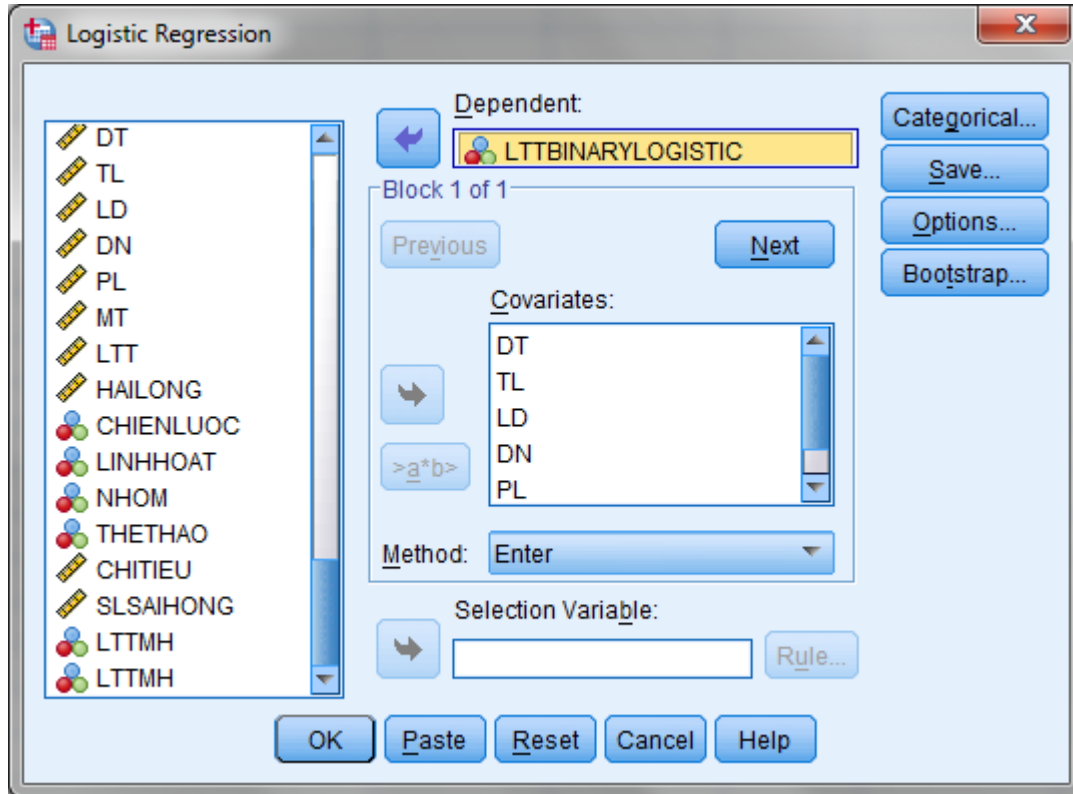
Thực hành trên SPSS

- Trước tiên, cần chú ý rằng biến lòng trung thành sau khi mã hóa để phân tích discriminant analysis có các giá trị 1 và 2, để hồi quy Binary logistic, cần mã hóa lại biến từ LTTMH thành LTTBINARYLOGISTIC với $1 \rightarrow 0$ (không trung thành) và $2 \rightarrow 1$ (trung thành)



- Để thực hiện hồi quy **Binary Logistic**, trên thanh trình đơn, vào **Analyze / Regression / Binary Logistic**.
- Ở cửa sổ **Logistic Regression**, chọn biến LTTBINARYLOGISTIC vào ô **Dependent** và các

biến **DT, TL, LD, DN, PL, MT** vào ô **Covariate**.



- Chọn phương pháp là **Enter**.
- Kích vào **OK** để hoàn tất việc thiết đặt và phân tích kết quả.
- Kết quả và diễn giải:

Omnibus Tests of Model Coefficients				
		Chi-square	df	Sig.
Step 1	Step	94.038	6	.000
	Block	94.038	6	.000
	Model	94.038	6	.000

Vì Chi-square = 94.038, df = 6 nên $p(\text{chi-square}, df) = 0.000 < 0.05$ nên có thể khẳng định rằng các hệ số góc $B_{DT}, B_{LT} \dots$ khác 0 (giả thuyết $B_{DT} = B_{LT} = \dots = 0$).

Model Summary			
Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	164.941 ^a	.375	.517

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than .001.

Theo biểu Model Sumary, độ phù hợp của mô hình là $-2LL = 164.941$.

Classification Table ^a			
		Predicted	
		Lòng trung thành (Binary Logistic)	Percentage Correct
Observed	.00		
	1.00		

Step 1	Lòng trung thành (Binary Logistic)	.00	49	21	70.0
		1.00	16	114	87.7
	Overall Percentage				81.5

a. The cut value is .500

Dễ dàng thấy rằng:

- Trong 65 (49+16) trường hợp được dự đoán là không trung thành có 49 trường hợp dự đoán đúng (trong tổng số 70 trường hợp thực tế không trung thành) chiếm tỷ lệ dự đoán đúng là 70%.
- Trong 130 (16 + 114) trường hợp thực tế là trung thành có 16 trường hợp dự đoán sai và 114 trường hợp dự đoán đúng chiếm tỷ lệ là 87.7%.
- Tỷ lệ dự đoán đúng của toàn bộ mô hình 81.5%.

Variables in the Equation							
		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	DT	.397	.216	3.379	1	.066	1.488
	TL	1.183	.234	25.576	1	.000	3.265
	LD	.536	.214	6.244	1	.012	1.709
	DN	-.024	.192	.016	1	.900	.976
	PL	-.219	.130	2.831	1	.092	.803
	MT	.332	.135	6.047	1	.014	1.393
	Constant	-8.404	1.469	32.721	1	.000	.000

a. Variable(s) entered on step 1: DT, TL, LD, DN, PL, MT.

Như vậy, các biến số trong mô hình gồm TL, LD, MT (vì có Sig. < 0.05), hàm hồi quy Binary Logistic sẽ là:

$$\log_e \left[\frac{P_i}{1-P_i} \right] = -8.404 + 1.183TL + 0.536LD + 0.332MT$$

Thay các giá trị TL, LD, MT vào mô hình dự đoán sẽ tính được xác suất có cán bộ công nhân viên trung thành theo phương trình:

$$P_i = E(Y = 1/X) = \frac{-8.404 + 1.183TL + 0.536LD + 0.332MT}{1 + e^{-8.404 + 1.183TL + 0.536LD + 0.332MT}}$$

CHƯƠNG

9.6

PHÂN TÍCH CỤM (CLUSTER ANALYSIS)

Phân tích cụm (cluster analysis) là kỹ thuật đa biến sử dụng nhằm phân loại và sắp xếp các quan sát vào các nhóm dựa vào các đặc tính của quan sát đó. Kỹ thuật này nhận diện và phân loại đối tượng sao cho thuộc tính của các quan sát trong cùng một cụm (nhóm) tương đồng cao và trong các cụm (nhóm) khác nhau có sự khác biệt lớn.

Chú ý rằng, trong phân tích tách biệt (discriminant analysis) cũng có thể cho phép sắp xếp các quan sát vào các nhóm, tuy nhiên, để sử dụng phân tích tách biệt cần phải biết số lượng nhóm trước khi thực hiện phân tích trong khi đó phân tích nhóm có thể không cần biết trước số nhóm mà sẽ do bản chất dữ liệu quyết định.

Phân tích cụm được sử dụng rất nhiều trong nghiên cứu kinh tế - xã hội như phân đoạn thị trường sản phẩm – dịch vụ, thử nghiệm các chiến lược của doanh nghiệp hay ứng nghiên cứu nghiên cứu địa lý, y học...

TIẾN TRÌNH PHÂN TÍCH CỤM

Phân tích cụm, giống như các kỹ thuật đa biến khác đã thảo luận trước đó, theo Hair và cộng sự (2009), phân tích cụm nhằm mục đích:

- Phân vùng tập hợp dữ liệu để hình thành cụm và lựa chọn giải pháp cụm;
- Giải thích các cụm để hiểu được đặc điểm của từng cụm (nhóm) và phát triển một tên hoặc nhãn phù hợp bản chất xác định của nó;
- Xác thực các kết quả của giải pháp cụm cuối cùng (tức là, xác định sự ổn định và tính tổng quát của nó).

Mục tiêu của phân tích cụm

Mục tiêu chính của phân tích cụm là phân vùng một tập hợp các nhóm thành hai hoặc nhiều nhóm dựa trên sự giống nhau của các đối tượng của một tập hợp các đặc điểm quy định (cụm biến ngẫu nhiên). Hoàn thành mục tiêu cơ bản này, các nhà nghiên cứu phải giải quyết hai vấn đề chính: các câu hỏi nghiên cứu được đề cập trong phân tích (mô tả phân loại, rút gọn dữ liệu hay xác định mối quan hệ) và các biến được sử dụng để mô tả các đối tượng trong quá trình phân cụm.

Thiết kế nghiên cứu trong phân tích cụm

Với các mục tiêu đã xác định và các biến đã chọn, trong phân tích cụm cần phải xác định kích thước mẫu có phù hợp hay không. Không có một quy định cụ thể mang tính chất định lượng về

độ lớn của mẫu trong phân tích cụm, tuy nhiên, cỡ mẫu phải đủ lớn để cung cấp đầy đủ đại diện của các nhóm nhỏ trong tập hợp và đại diện các cấu trúc cơ bản của các quan sát trong nhóm và tất nhiên, mẫu cần đủ lớn để khi phân nhóm sẽ làm tăng cơ hội của các nhóm nhỏ mang tính đại diện.

Đo lường trong phân tích cụm

Theo Trọng và Ngọc (2008), mục tiêu của phân tích cụm là nhóm các đối tượng giống nhau lại nên cần phải có một thức đo nào đó để đánh giá mức độ giống nhau hay khác nhau của đối tượng và phương pháp tưởng chừng nhất là đo lường mức độ giống nhau bằng khoảng cách giữa hai đối tượng trong một cặp đối tượng.

Có các cách để tính khoảng cách giữa hai đối tượng:

- Khoảng cách Euclid là căn bậc hai của tổng các độ lệch bình phương của các giá trị trên từng biến.
- Khoảng cách Manhattan giữa hai đối tượng là tổng các độ lệch tuyệt đối của các giá trị trên từng biến.
- Khoảng cách Chebyshev giữa hai đối tượng là chênh lệch tuyệt đối lớn nhất của các giá trị trên từng biến.

Thủ tục phân cụm

Phân cụm thứ bậc (Hierarchical clustering)

Phân cụm thứ bậc là phương pháp xây dựng theo một cấu trúc thứ bậc hay dạng hình cây, phương pháp này được tiến hành theo cách tích tụ lại hoặc phân chia ra. Phân cụm tích tụ thường được sử dụng trong nghiên cứu với các phương pháp (1) khoảng cách liên kết, (2) tổng các độ lệch bình phương hay phương sai và (3) khoảng cách trung tâm (Trọng & Ngọc, 2008).

Phân cụm không thứ bậc (Non-hierarchical clustering)

Phân cụm không thứ bậc là loại phân cụm mà nhà nghiên cứu đã biết trước số cụm thông qua nghiên cứu lý thuyết hoặc do ý kiến chủ quan của nhà quản lý (ví dụ nhà quản lý muốn phân chia thị trường thành k phân đoạn) và giả sử rằng, đặc điểm (thuộc tính) của các quan sát trong các nhóm (cluster) sẽ tương đồng nhưng giữa các nhóm khác nhau thì khác nhau. Theo Trọng và Ngọc (2008), phân cụm không thứ bậc gồm các phương pháp: phương pháp bắt đầu tuần tự, bắt đầu song song và phân chia tối ưu.

TIẾN TRÌNH PHÂN TÍCH CỤM THỨ BẬC

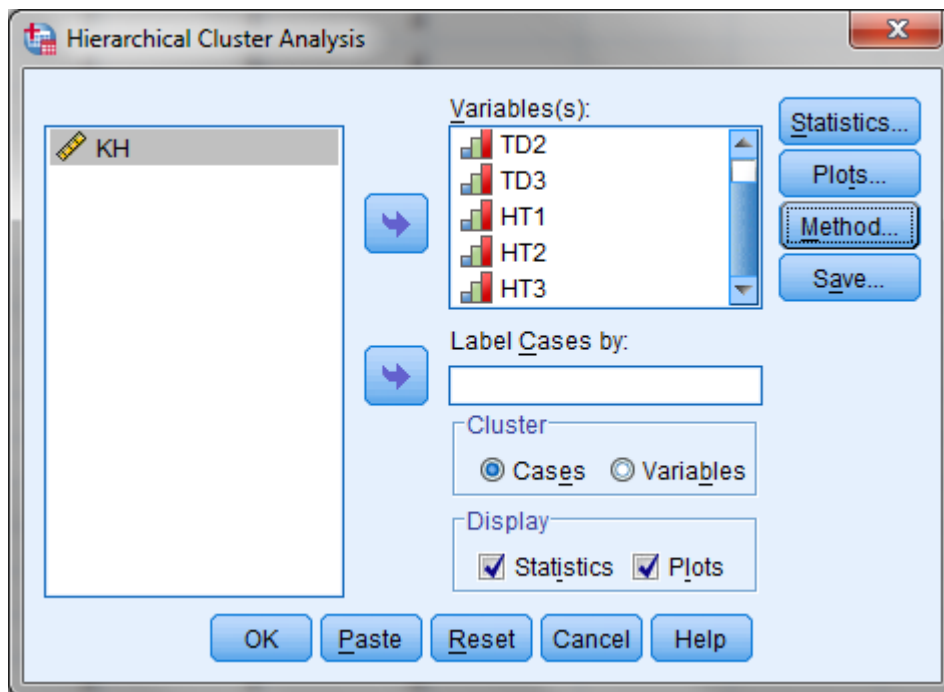
Xác định vấn đề

Một nhà nghiên cứu mục tiêu đi siêu thị của khách hàng bằng cách thực hiện phỏng vấn 50 khách hàng thông qua 6 chỉ báo về mục tiêu với (1) Hoàn toàn không đồng ý và (7) Hoàn toàn đồng ý. Thông qua phân tích cụm, với kết quả trả lời của đáp viên nhà nghiên cứu muốn phân loại các nhóm khách hàng và thuộc tính của các nhóm để có được chính sách marketing hợp lý cho từng nhóm khách hàng. Dữ liệu thu thập được được lưu trữ tại file MUA SAM GIAI TRI.sav

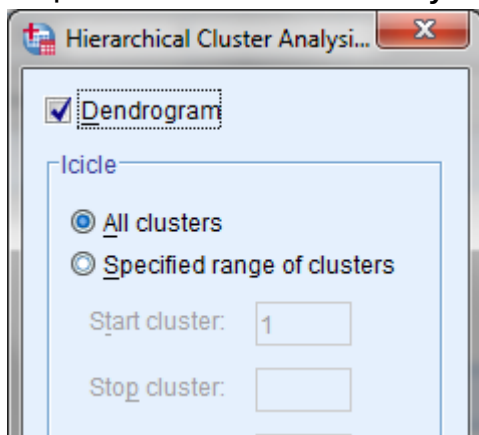
	1	2	3	4	5	6	7
Có thể tự do lựa chọn, mua sắm sản phẩm	☐	☐	☐	☐	☐	☐	☐
Có thể tìm thấy mặt hàng mà khách hàng muốn mua	☐	☐	☐	☐	☐	☐	☐
Dễ dàng giao tiếp trong quá trình mua hàng	☐	☐	☐	☐	☐	☐	☐
Có thể tìm hiểu tính năng / đặc điểm sản phẩm	☐	☐	☐	☐	☐	☐	☐
Đạt được những mục tiêu mua sắm dự định	☐	☐	☐	☐	☐	☐	☐
Khả năng mua được hàng hóa cao	☐	☐	☐	☐	☐	☐	☐

Thực hiện trên SPSS bằng phân cụm thứ bậc (hierarchical clustering)

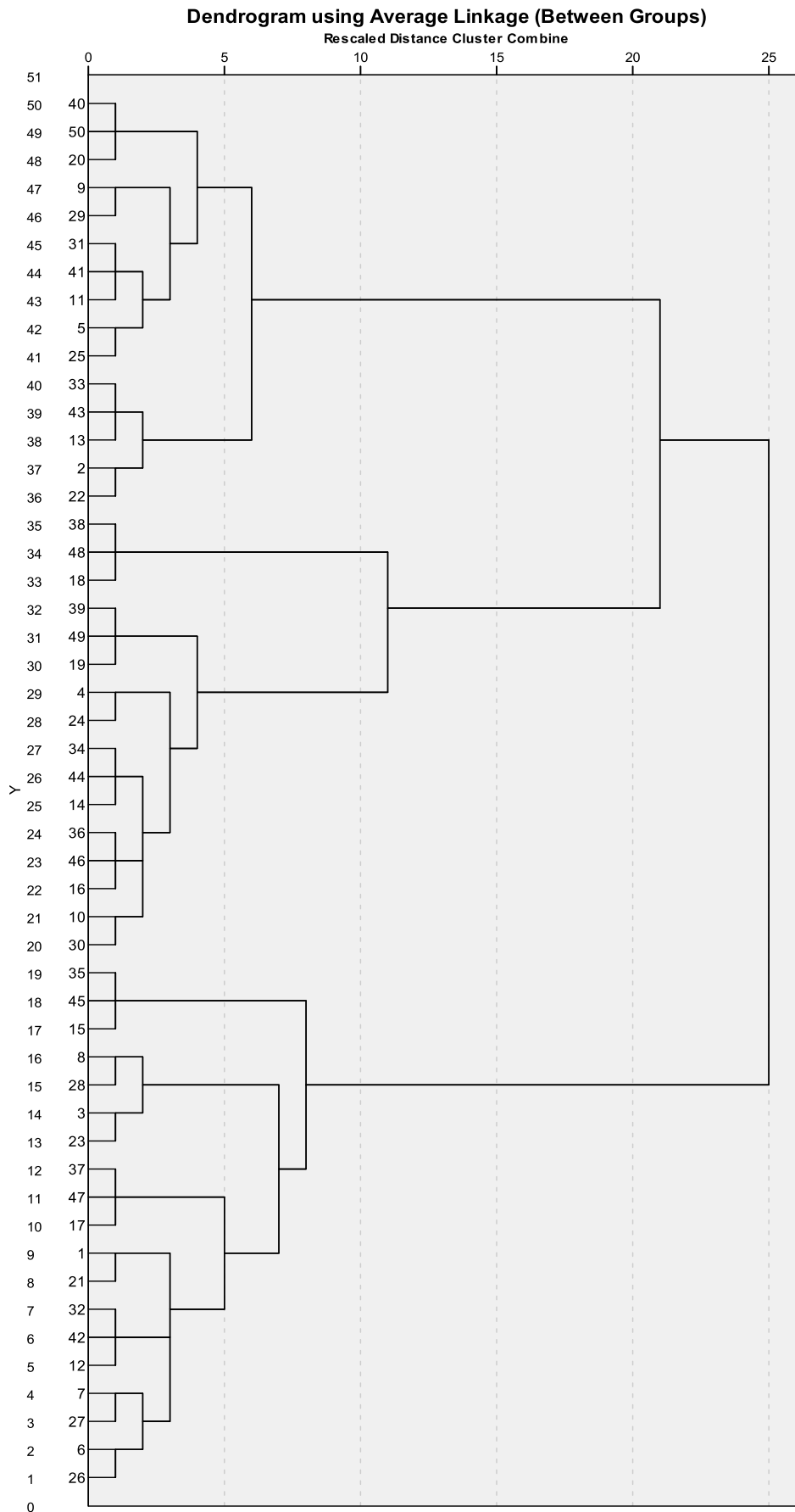
- Mở file MUA SAM GIAI TRI.sav, trên thanh trình đơn chọn **Analyze / Classify / Hierarchical Cluster** và chọn các tùy chọn như hình vẽ.



- Kích vào **Plots** và chọn **Dendrogram** để vẽ sơ đồ hình cây, kích **Continue** để trở về hộp hội thoại **Hierarchical Cluster Analysis**.



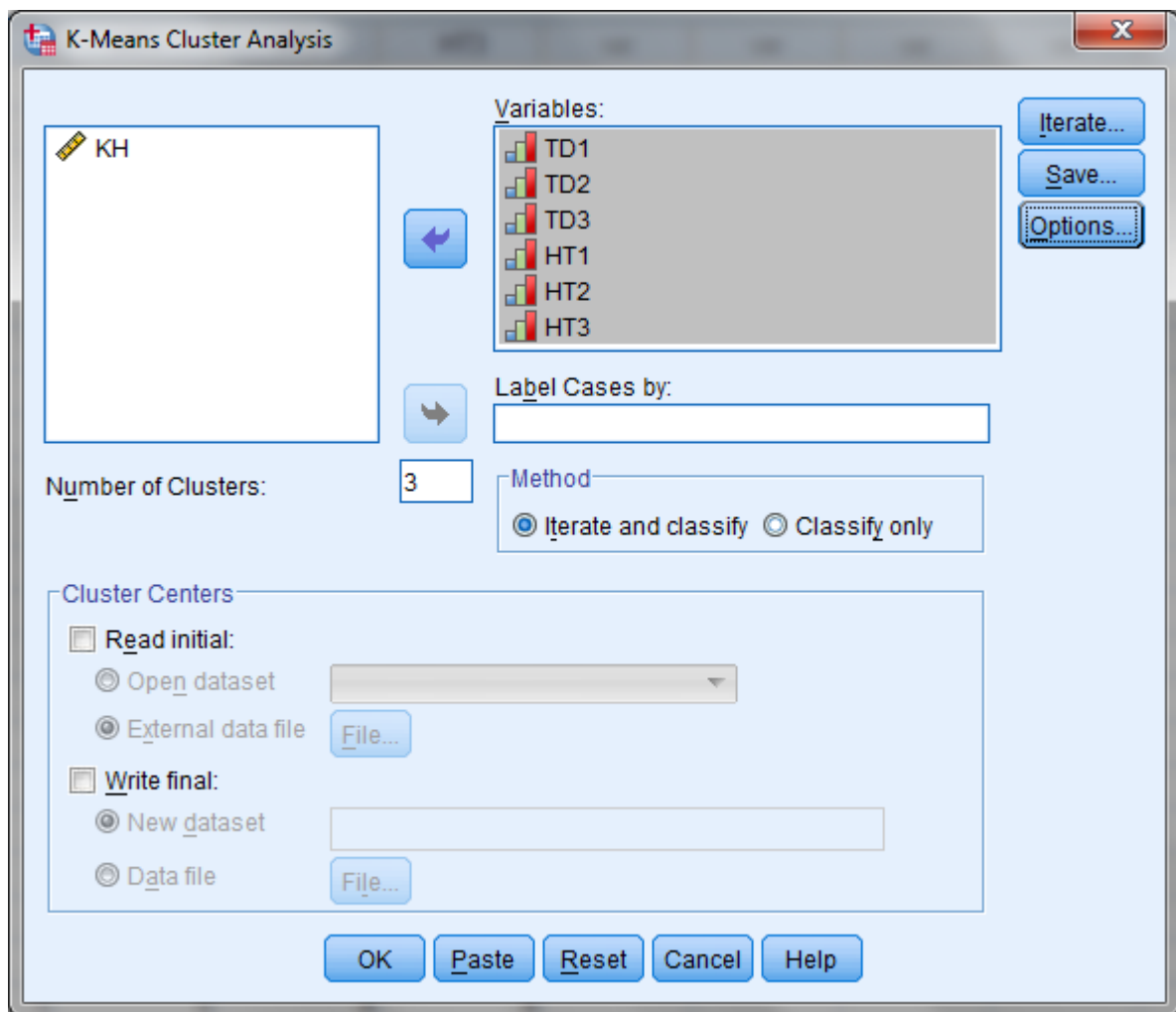
- Kích **OK** để hoàn tất thiết đặt và tính kết quả.



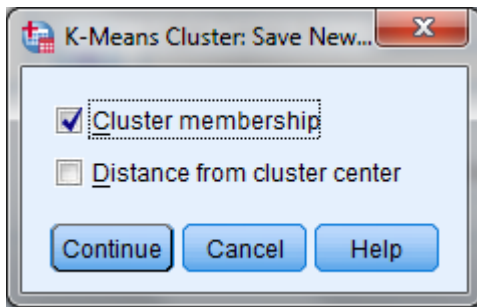
- Xác định số cụm: Cần nhấn mạnh rằng cho đến nay, chưa có một quy tắc rõ ràng cho việc xác định số cụm, khi xác định có thể dựa trên (1) các lý thuyết đi trước đã kết luận hoặc đã đề xuất có bao nhiêu nhóm (ví dụ, 2-3... nhóm), kết hợp với hình vẽ sẽ thuận tiện cho việc kết luận số nhóm; (2) quan sát sự thay đổi (đột ngột) trong hình vẽ hoặc trong hệ số (coefficient) tại biểu Agglomeration Schedule. Trên hình vẽ trên, có thể thấy có 3 sự thay đổi đột ngột nên có thể kết luận tồn tại 3 nhóm khách hàng.
- Cần chú ý, trong trường hợp này phương pháp phân tích đã mặc định phương pháp phân cụm **Between-groups linkage** với bình phương khoảng cách Euclid (Squared Euclidean distance).

Thực hiện trên SPSS bằng phân cụm không thứ bậc (non - hierarchical clustering)

- Với ví dụ trên thực hiện bằng phân cụm không thứ bậc bằng cách mở file MUA SAM GIAI TRI.sav, trên thanh trình đơn chọn **Analyze / Classify / K-means Cluster** và đưa các giá trị cần phân tích vào ô **Variables**, giả sử số nhóm được tạo thành được biết trước (theo phân tích lý thuyết) là 3 và thiết đặt số nhóm bằng 3 ở ô **Number of Clusters**.



- Nếu cần biết khách hàng thuộc nhóm nào, nhấn vào **Save** (để lưu lại trên biến có tên là **QCL_1** ở cuối file dữ liệu), bấm **Continue** để trở lại cửa sổ **K-means Cluster Analysis**.



- Bấm **OK** để hoàn tất thiết đặt và tiến hành phân tích.
- Kết quả phân tích và diễn giải:

Final Cluster Centers			
	Cluster		
	1	2	3
TD1	6	3	4
TD2	4	2	6
TD3	3	2	6
HT1	4	5	2
HT2	6	4	3
HT3	6	3	4

Biểu Final Cluster Centers sẽ cho biết thông tin về đặc điểm của các nhóm trên cơ sở tiêu chí sử dụng để phân nhóm (từ TD1 đến HT3). Ví dụ: Ở nhóm 1 có điểm trung bình của TD1 = 6, TD2 = 4, TD3 = 3...; nhóm 2 có TD1 = 3, TD2 = 2, TD3 = 2...

Number of Cases in each Cluster		
Cluster	1	16.000
	2	15.000
	3	19.000
Valid		50.000
Missing		.000

Biểu Number of Cases in each Cluster cho biết có 16 khách hàng thuộc nhóm 1, 15 khách hàng thuộc nhóm 2 và 19 khách hàng thuộc nhóm 3.

Nếu dữ liệu có bao gồm những thông khác, trên cơ sở thông tin về nhóm tại biến QCL_1, người phân tích có thể sử dụng biến này để phân tích sâu hơn về đặc điểm của nhóm như đặc điểm giới thích, thu nhập, trình độ học vấn của các thuộc các nhóm...

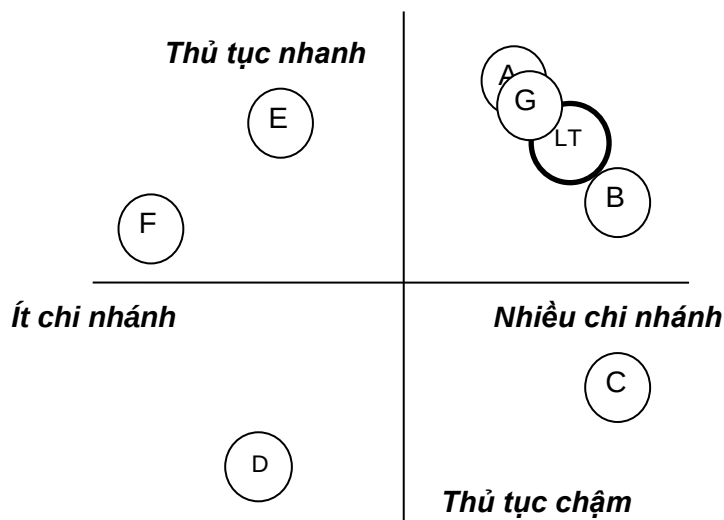
CHƯƠNG

9.7

LẬP BẢN ĐỒ NHẬN THỨC (PERCEPTUAL MAP)

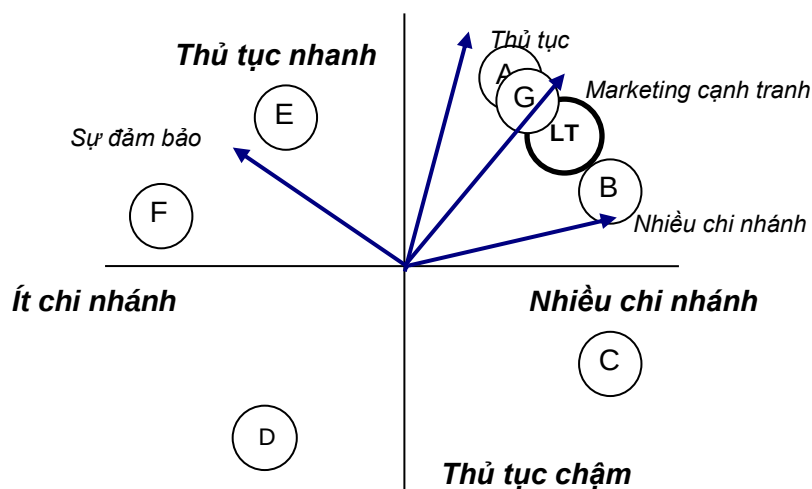
SƠ LƯỢC VỀ BIỂU ĐỒ NHẬN THỨC (Perceptual map)

Trong thực tế kinh doanh, các doanh nghiệp luôn mong muốn tìm biết được vị trí (position) sản phẩm, dịch vụ hoặc thương hiệu trong tâm trí của người tiêu dùng, thông thường, cần phải tìm hiểu sự cảm nhận, sự nhận thức của người tiêu dùng đối với sản phẩm, dịch vụ hay thương hiệu. Để biết được điều này, thông thường, người ta sử dụng bản đồ nhận thức (perceptual map). Đo lường đa hướng (Multidimensional scaling - MDS), còn được gọi là lập bản đồ nhận thức, là một thủ tục cho phép một nhà nghiên cứu để xác định hình ảnh liên hệ nhận thức của một tập các đối tượng (các công ty, sản phẩm, ý tưởng hoặc các chỉ báo...). Mục đích của MDS là để chuyển đổi sự đánh giá của người tiêu dùng về sự tương đồng (similarity) hoặc sự ưu tiên (preference) đối với các sản phẩm dịch vụ hoặc thương hiệu... vào khoảng cách đại diện trong không gian đa chiều. Để thực hiện phân tích MDS, nhà nghiên cứu thực hiện ba bước cơ bản: (1) Tập hợp các biện pháp tương tự hoặc ưu tiên trên toàn bộ các đối tượng được phân tích, (2) Sử dụng kỹ thuật MDS để ước lượng vị trí tương đối của mỗi đối tượng trong không gian đa chiều và (3) xác định các giải thích các trục của không gian về nhận thức v / hoặc các thuộc tính khách quan. Giả sử người nghiên cứu muốn xem xét hai thuộc tính (nhiều – ít chi nhánh, thủ tục nhanh – chậm) của ngân hàng A, B, C, D, E, F... trong nhận thức của khách hàng, kỹ thuật MDS sẽ giúp cho việc xác định vị trí các đối tượng A và B sao cho khoảng cách giữa chúng trong không gian đa chiều nhỏ hơn khoảng cách giữa các cặp khác trong số các đối tượng (ngân hàng). Bản đồ nhận thức cho thấy vị trí tương đối của tất cả các ngân hàng như sau:



Trên hình vẽ có một điểm lý tưởng (LT), đây là điểm lý tưởng mà các ngân hàng mong muốn hướng điểm nhằm thỏa mãn một cách tốt nhất mong muốn của khách hàng.

Trong quá trình kinh doanh, các ngân hàng sẽ xây dựng thương hiệu thông qua các thuộc tính (attributs), từ đó có chiến lược phát triển thương hiệu nhằm giúp cho khách hàng nhận biết thương hiệu. Giả sử có 4 thuộc tính liên quan đến xem xét thương hiệu ngân hàng gồm số lượng chi nhánh, thủ tục, chiến lược marketing cạnh tranh, công nghệ, sự đảm bảo. Các thuộc tính này sẽ thể hiện trên biểu đồ nhận thức như sau:



Thông qua biểu đồ nhận thức, dễ dàng nhận thấy ngân hàng A được khách hàng nhận thức là một ngân hàng có chiến lược marketing cạnh tranh, thủ tục nhanh chóng trong khi đó ngân hàng B được nhận thức là một ngân hàng có nhiều chi nhánh... Việc xem xét khoảng cách xa hay gần được xác định trên cơ sở hình chiếu của vị trí của ngân hàng lên các trục.

Các ngân hàng càng gần điểm lý tưởng LT càng có nhiều thuận lợi trong quá trình xây dựng và phát triển thương hiệu, các ngân hàng có vị trí càng xa điểm lý tưởng càng khó khăn và tốn nhiều chi phí trong việc định vị và phát triển thương hiệu. Đồng thời, các ngân hàng càng gần nhau càng cạnh tranh gay gắt nhằm tạo ra sự khác biệt cho mỗi thương hiệu.

LẬP BẢN ĐỒ NHẬN THỨC

Thông thường, có hai kỹ thuật để lập bản đồ nhận thức:

- Đo lường đa hướng (Multidimensional scaling - MDS): Thể hiện các đối tượng trong không gian đa chiều.
- Phân tích sự tương hợp (Correspondence analysis - CA): Thể hiện sự tương hợp của các đối tượng với các thuộc tính.

Để làm rõ hai kỹ thuật này, chúng ta sẽ xem xét hai ví dụ về việc sử dụng các kỹ thuật để lập biểu đồ nhận thức, cần chú ý các thức thu thập và thiết đặt dữ liệu trong SPSS để thực hiện hai kỹ thuật này.

Ví dụ 1

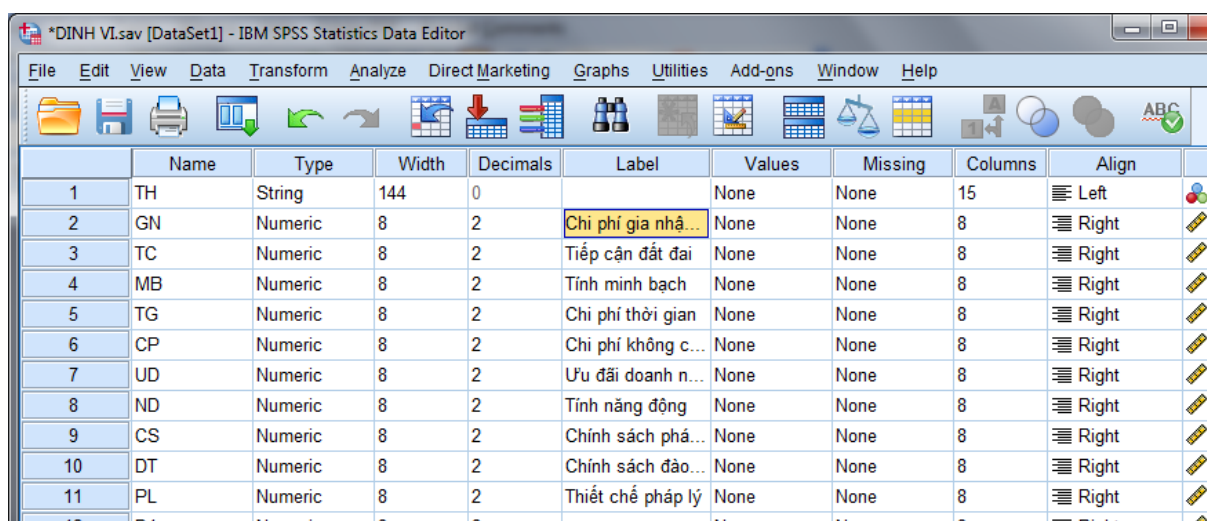
Giả sử có số liệu của 10 chỉ số năng lực cạnh tranh cấp tỉnh (PCI) xây dựng dựa trên việc đo

lường và đánh giá thực tiễn điều hành kinh tế của chính quyền 64 tỉnh/thành phố trên cơ sở khảo sát khu vực kinh tế tư nhân gồm:

- Chi phí gia nhập thị trường;
- Tiếp cận đất đai;
- Tính minh bạch;
- Chi phí thời gian;
- Chi phí không chính thức;
- Ưu đãi doanh nghiệp nhà nước;
- Tính năng động;
- Chính sách phát triển kinh tế tư nhân;
- Chính sách đào tạo;
- Thiết chế pháp lý;

Trong cơ sở dữ liệu đã thu thập, nhà nghiên cứu muốn định vị thương hiệu các Tỉnh, thành phố gồm Bình Định, Quảng Ngãi, Quảng Nam, Quảng Ngãi và Huế trên cơ sở 10 biến số thành phần của chỉ số PCI nhằm nhận biết trong nhận thức của các doanh nghiệp, mỗi Tỉnh, Thành phố được “nhận thức” như thế nào theo những thuộc tính đã xác định.

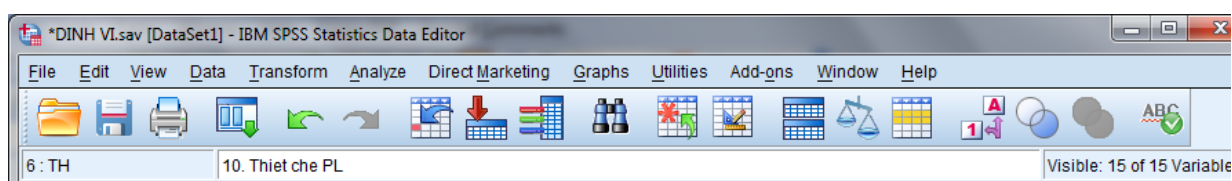
- Định nghĩa các biến: Trước khi tiến hành nhập liệu cần định nghĩa các biến gồm TH (biến Thương hiệu), các biến GN, TC, MB... lần lượt là các biến chi phí gia nhập thị trường, tiếp cận đất đai, tính minh bạch....



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align
1	TH	String	144	0		None	None	15	Left
2	GN	Numeric	8	2	Chi phí gia nhâ...	None	None	8	Right
3	TC	Numeric	8	2	Tiếp cận đất đai	None	None	8	Right
4	MB	Numeric	8	2	Tính minh bạch	None	None	8	Right
5	TG	Numeric	8	2	Chi phí thời gian	None	None	8	Right
6	CP	Numeric	8	2	Chi phí không c...	None	None	8	Right
7	UD	Numeric	8	2	Ưu đãi doanh n...	None	None	8	Right
8	ND	Numeric	8	2	Tính năng động	None	None	8	Right
9	CS	Numeric	8	2	Chính sách phá...	None	None	8	Right
10	DT	Numeric	8	2	Chính sách đào...	None	None	8	Right
11	PL	Numeric	8	2	Thiết chế pháp lý	None	None	8	Right

- Nhập liệu các chỉ số thành phần và thương hiệu các Tỉnh, Thành phố với giá trị chỉ số thành phần tương ứng với các Tỉnh, Thành phố.

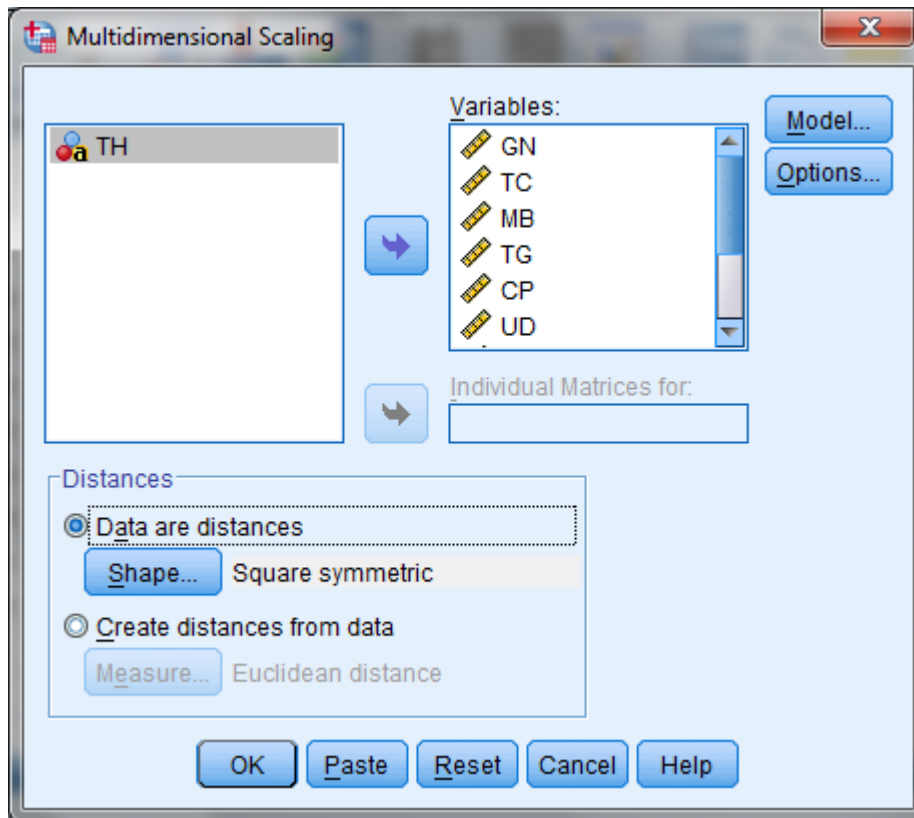
Chú ý: Trong trường hợp này đã có số liệu trung bình đánh giá các thuộc tính của các Tỉnh, Thành phố, nếu chưa có giá trị trung bình, người nghiên cứu phải tính giá trị trung bình trước khi thực hiện.



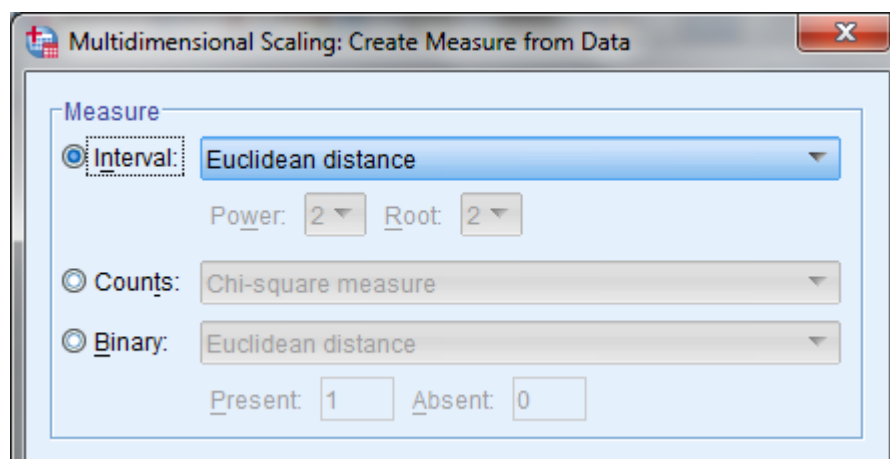
	6 : TH	10. Thiet che PL
1		
2		

Visible: 15 of 15 Variable

- Trên thanh trình đơn, vào **Analyze / Scale / Multidimensional Scaling (ALSCAL)**
- Đưa các biến thành phần vào ô **Variables**

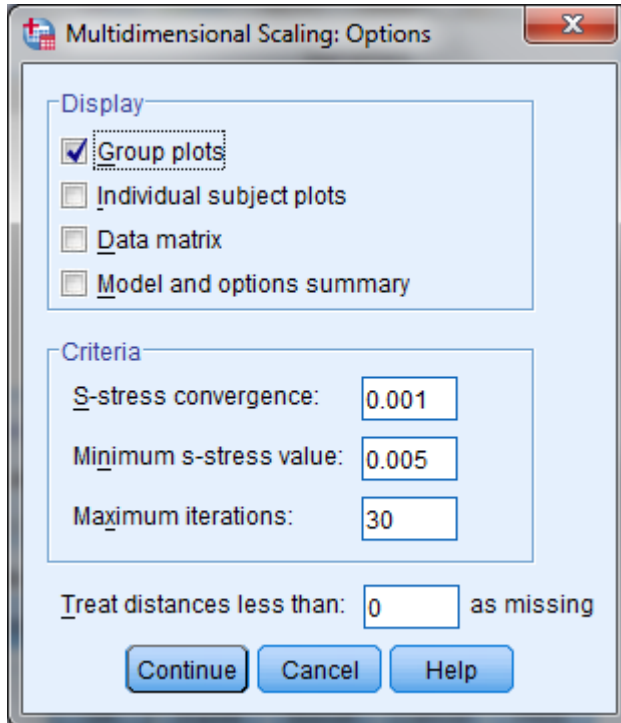


- Đánh dấu vào **Create distance from data** và chọn **Measure**, chọn **Between cases**



- Kích **Model**, chọn **Ratio** (thang đo cấp tỉ lệ) và kích **OK**,

- Kích **Option** chọn **Group plots**,
- Kích **Continue** để tiếp tục,



- Kích **OK** để hoàn tất việc thiết đặt và tính kết quả.
- Trong file Output, tìm dòng **Stimulus Coordinates** và có thông tin như sau:

Stimulus Coordinates				
Dimension				
Stimulus	Stimulus	1	2	
Number	Name			
1	VAR1	.5546	.7738	
2	VAR2	-1.9765	-.2930	
3	VAR3	-.2147	.4839	
4	VAR4	2.0408	-.5996	
5	VAR5	-.4042	-.3651	

- Chép (copy) giá trị của các biến VAR1 đến VAR5 vào trong phần nhập liệu của SPSS

	CS	DT	PL	VAR00002	VAR00003	var	var	var
1	39	6.87	5.85	3.94	.5458	.8913		
2	70	3.85	4.39	3.40	-1.8729	-.4632		
3	38	5.23	5.19	5.61	-.2293	.6857		
4	38	8.10	8.78	6.14	1.9649	-.6422		
5	51	5.19	6.24	4.15	-.4086	-.4716		

- Đổi tên các biến thành D1 và D2.

*DINH VI.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

Visible: 13 of 13 Variables

	CS	DT	PL	D1	D2	var	var	var
1	6.87	5.85	3.94	.5458	.8913			
2	3.85	4.39	3.40	-1.8729	-.4632			
3	5.23	5.19	5.61	-.2293	.6857			
4	8.10	8.78	6.14	1.9649	-.6422			
5	5.19	6.24	4.15	-.4086	-.4716			
6	.	.	.	2.7425	-.2300			
7	.	.	.	-.3294	1.0642			
8	.	.	.	.5536	-.0791			
9	.	.	.	-.6248	.4513			
10	.	.	.	.2324	.3604			
11	.	.	.	1.3060	.3669			
12	.	.	.	-.5302	-.0722			
13	.	.	.	-.6408	-.8508			

Data View Variable View

IBM SPSS Statistics Processor is ready

- Thực hiện tương tự như trên nhưng trong phần **Multidimensional Scaling: Create distance from data** chọn **Between variable**

Multidimensional Scaling: Create Measure from Data

Measure

☒ **Interval:** Euclidean distance

Power: 2 Root: 2

☐ **Counts:** Chi-square measure

☐ **Binary:** Euclidean distance

Present: 1 Absent: 0

Transform Values

Standardize: None

☐ By variable

☐ By case

Create Distance Matrix

☐ Between variables

☒ **Between cases**

Continue Cancel Help

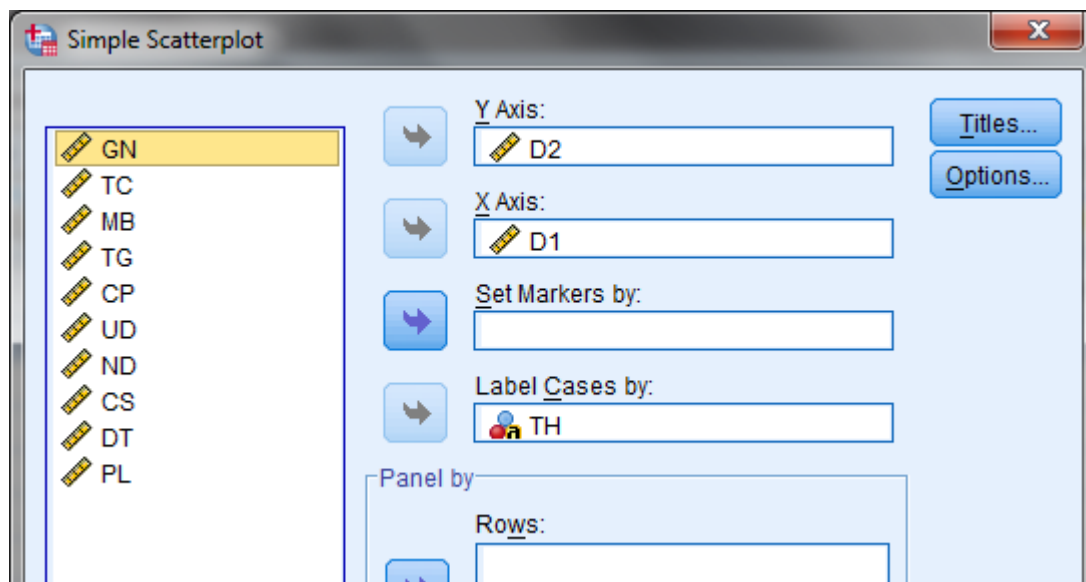
- Trong file Output, tìm dòng **Stimulus Coordinates** và có thông tin như sau:
Stimulus Coordinates

		Dimension	
Stimulus Number	Stimulus Name	1	2
1	GN	2.7425	-.2300
2	TC	-.3294	1.0642
3	MB	.5536	-.0791
4	TG	-.6248	.4513
5	CP	.2324	.3604
6	UD	1.3060	.3669
7	ND	-.5302	-.0722
8	CS	-.6408	-.8508
9	DT	-.3778	-1.1153
10	PL	-2.3314	.1046

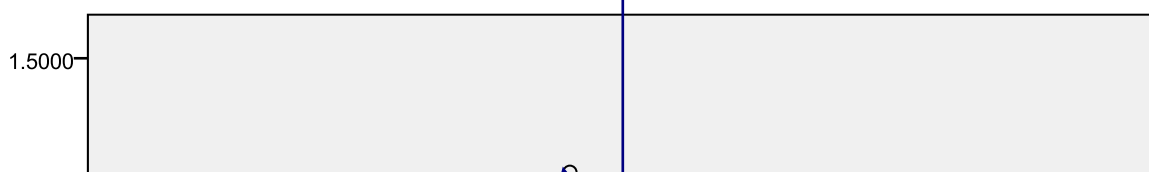
- Chép (copy) giá trị của các biến từ GN đến PL vào trong phần nhập liệu của SPSS phía dưới dữ liệu của các thương hiệu (VAR)

	CS	DT	PL	D1	D2	var	var	var
4	8.10	8.78	6.14	1.9649	-.6422			
5	5.19	6.24	4.15	-.4086	-.4716			
6	-	-	-	2.7425	-.2300			
7	-	-	-	-.3294	1.0642			
8	-	-	-	.5536	-.0791			
9	-	-	-	-.6248	.4513			
10	-	-	-	.2324	.3604			
11	-	-	-	1.3060	.3669			

- Vào **Graphs / Legacy Dialogs / Scatter/Dot** để vẽ sơ đồ:



- Khi đó, bản đồ nhận thức được thể hiện như sau:



Chú ý: Để cho hình vẽ bản đồ định vị dễ nhìn và đẹp hơn, có thể sao chép dữ liệu vào Excel để vẽ bản đồ.

Ví dụ 2

Một nhà nghiên cứu muốn xem xét việc định vị thương hiệu thẻ của các ngân hàng thông qua việc nhận thức của khách hàng nhằm xây dựng chiến lược định vị và truyền thông phù hợp với điều kiện cạnh tranh trong tương lai, người trả lời sẽ lựa chọn trả lời 4 trong tổng số 8 ngân hàng được đưa vào nghiên cứu gồm ngân hàng A, B, C, D, E, F, G, H.

Để thu thập dữ liệu phục vụ phân tích, một bản câu hỏi được xây dựng trên cơ sở những thuộc tính của dịch vụ thẻ được xem xét gồm (1) uy tín ngân hàng, (2) mạng lưới chi nhánh, (3) công nghệ sử dụng, (4) chiến lược marketing cạnh tranh, (5) mạng lưới cung cấp dịch vụ thẻ, (6) chất lượng dịch vụ thẻ (tính năng thẻ, độ an toàn trong thanh toán), (7) khả năng kết nối liên ngân hàng. Các thuộc tính được đo lường bởi thang đo khoảng cách (thang Likert) từ 1 đến 9 với (1) là rất đồng ý và (7) là rất không đồng ý. Để người trả lời cung cấp được thông tin cho nghiên cứu, cần lựa chọn người tham gia trả lời phải biết và hiểu các thuộc tính của các ngân hàng trên. Mẫu câu hỏi được thiết kế như sau:

Bạn lựa chọn thẻ của ngân hàng XYZ vì	1	2	3	4	5	6	7	8	9
1. Uy tín của ngân hàng	☐	☐	☐	☐	☐	☐	☐	☐	☐
2. Mạng lưới chi nhánh rộng khắp	☐	☐	☐	☐	☐	☐	☐	☐	☐

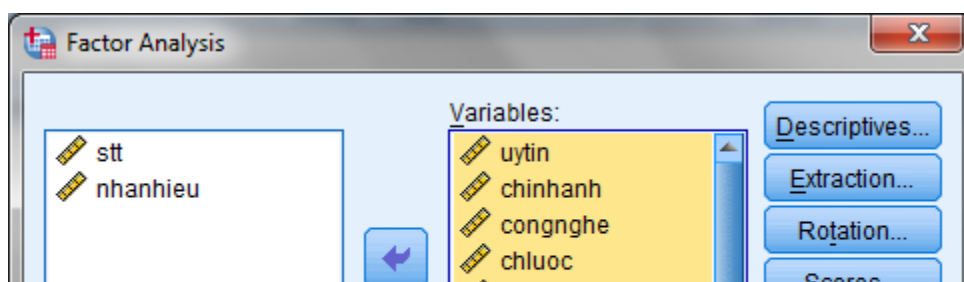
3. Công nghệ hiện đại	☐	☐	☐	☐	☐	☐	☐	☐	☐
4. Chiến lược marketing cạnh tranh	☐	☐	☐	☐	☐	☐	☐	☐	☐
5. Mạng lưới cung cấp dịch vụ thẻ rộng khắp	☐	☐	☐	☐	☐	☐	☐	☐	☐
6. Dịch vụ thẻ có chất lượng tốt	☐	☐	☐	☐	☐	☐	☐	☐	☐
7. Có khả năng kết nối liên ngân hàng	☐	☐	☐	☐	☐	☐	☐	☐	☐

Sau khi trả lời, kết quả được nhập liệu vào phần mềm SPSS, khi nhập liệu, cần chú ý là ở cột **stt** biểu thị số thứ tự của người tham gia trả lời (chúng ta thấy có 4 số 1 nghĩa là người thứ 1 trả lời cho 4 nhãn hiệu được lựa chọn trong 8 nhãn hiệu (gồm 1, 3, 5, 8 tương ứng với A, C, E, H) mà họ lựa chọn, nhãn hiệu là tên của ngân hàng được lựa chọn để trả lời các thuộc tính tương ứng.

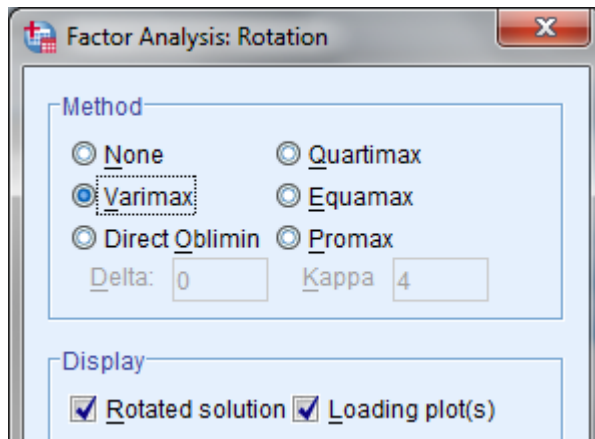
	stt	nhanhieu	uytin	chinhanh	congnghe	chluoc	mangluoi	chluong	ketnoi
1	1	1	7	4	4	4	5	1	1
2	1	3	7	7	8	3	3	3	3
3	1	5	7	7	6	4	8	3	3
4	1	8	7	7	6	2	2	8	3
5	2	2	6	6	4	5	5	2	2
6	2	3	6	6	2	1	1	2	2
7	2	7	7	7	5	5	5	3	3
8	2	8	4	4	3	2	2	8	4
9	3	1	7	7	3	8	8	4	4
10	3	2	7	7	5	6	6	3	3
11	3	6	7	7	6	5	5	4	6
12	3	4	6	6	5	6	6	8	6

Trình tự thực hiện đo lường đa hướng (Multidimensional scaling - MDS) như sau:

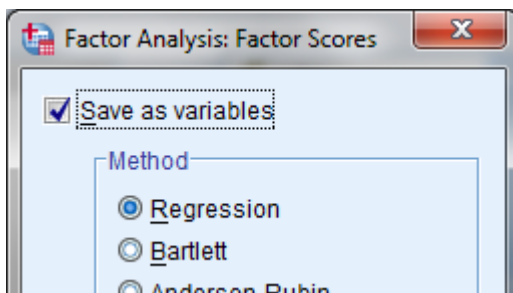
- Xác định nhân tố chính (principal dimension): Như đã giới thiệu trong phần phân tích nhân tố khám phá (EFA), phương pháp phân tích nhân tố khám phá (explore factor analysis) cho phép xác định số nhân tố chính (Hair và cộng sự, 2009). Theo đó, nhà nghiên cứu sẽ xác định bằng cách các thuộc tính của các thương hiệu sẽ được phân nhóm và thể hiện mối quan hệ chặt chẽ với mỗi nhân tố chính vừa xác định. Số nhân tố chính được sẽ được lựa chọn theo nguyên tắc giá trị riêng (eigenvalues > 1) hoặc lựa chọn theo số nhân tố cố định (Hair và cộng sự, 2009).
- Thực hiện phân tích nhân tố với phép quay Varimax, chọn hệ số tải nhân tố lớn hơn hoặc 0.5 cho 7 thuộc tính.
- Khi đó, hộp hội thoại Factor Analysis được mở ra, chọn 7 thuộc tính vào ô Variables như sau:



- Kích vào **Varimax**, **Rotated solution** và **Loading plot(s)** trong **Rotation**.



- Kích vào **Save as Variables** trong **Scores** để lưu lại giá trị của các nhân tố.



- Khi đó, kết quả phân tích có ma trận thành phần sau khi xoay **Rotated Component Matrix** như sau:

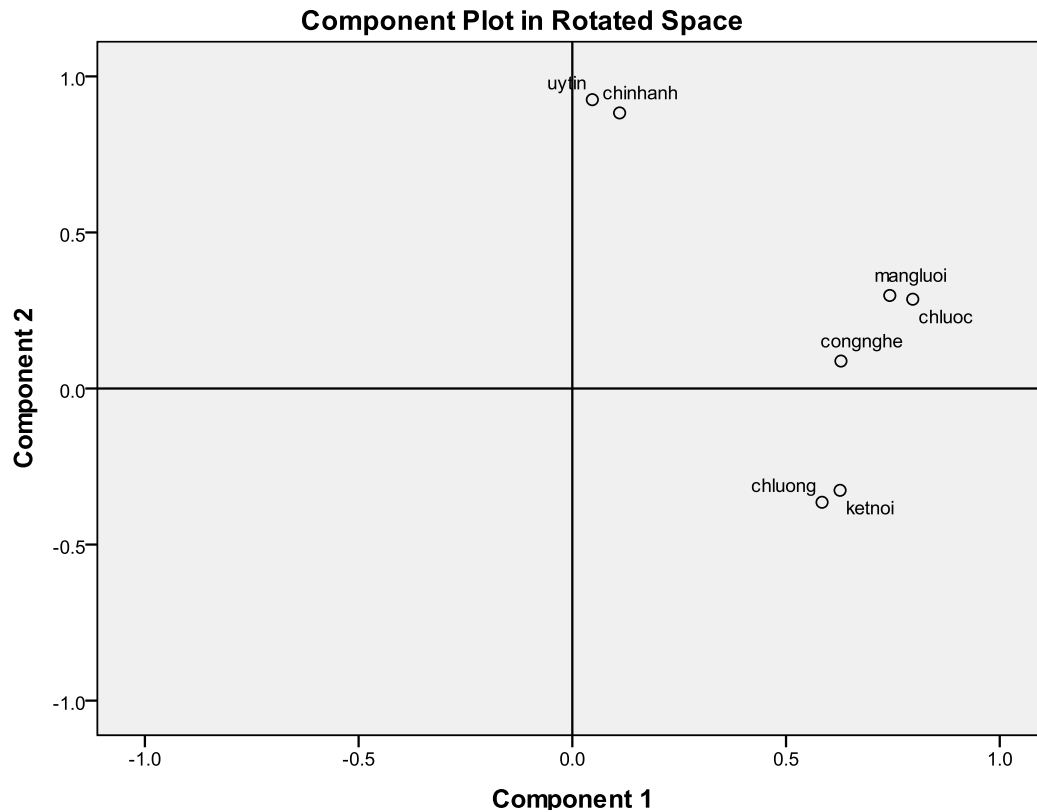
	Component	
	1	2
Chiến lược marketing cạnh tranh	.796	.286
Mạng lưới cung cấp	.743	.298
Công nghệ ngân hàng	.629	.088
Khả năng kết nối	.626	-.326
Chất lượng dịch vụ thẻ	.584	-.365
Uy tín ngân hàng	.047	.925
Mạng lưới chi nhánh	.111	.883

Extraction Method: Principal Component Analysis.

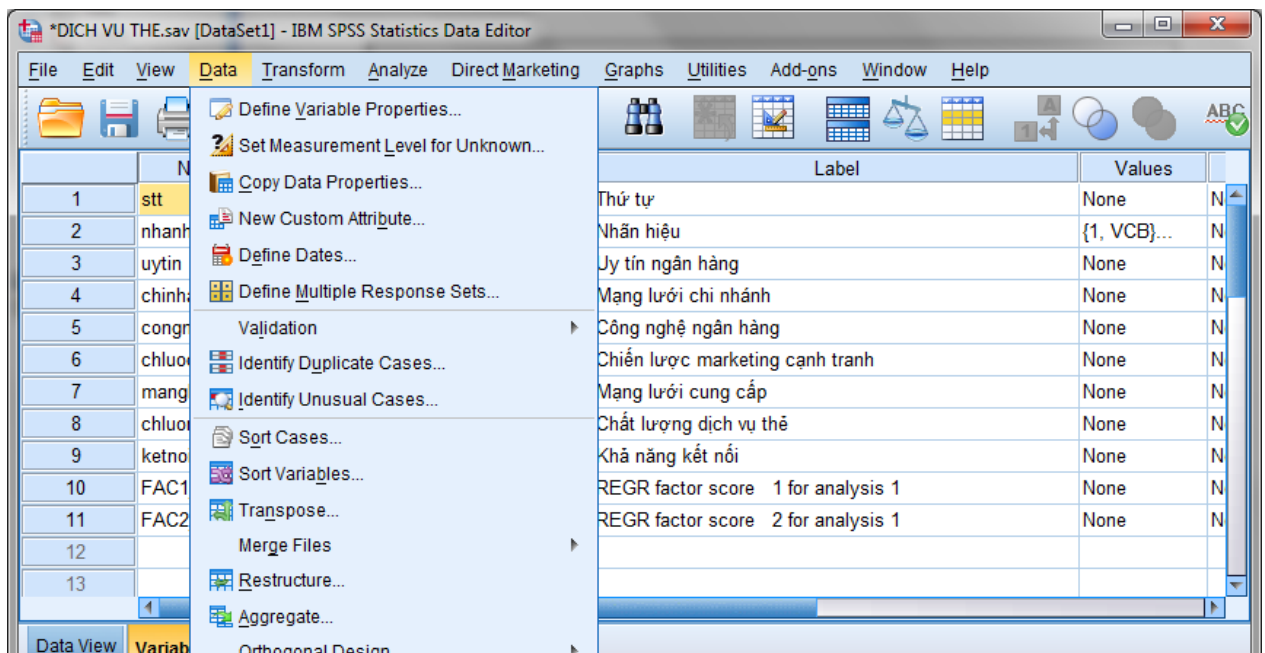
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 3 iterations.

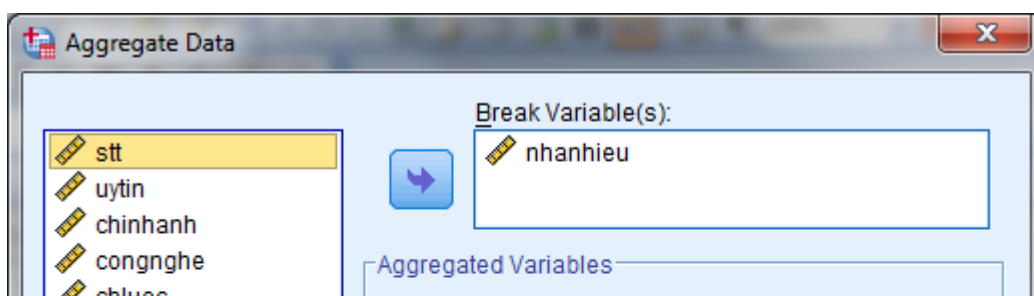
- Kết quả của **Rotated Component Matrix** được thể hiện thông qua bản đồ như sau:



- Gộp dữ liệu bằng **Aggregate**: Trên thanh trình đơn, vào **Data / Aggregate**



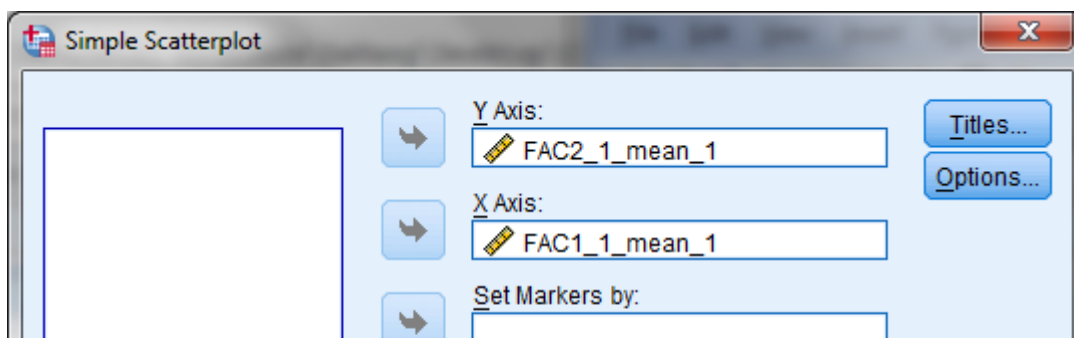
- Trên cửa sổ **Aggregate Data**:



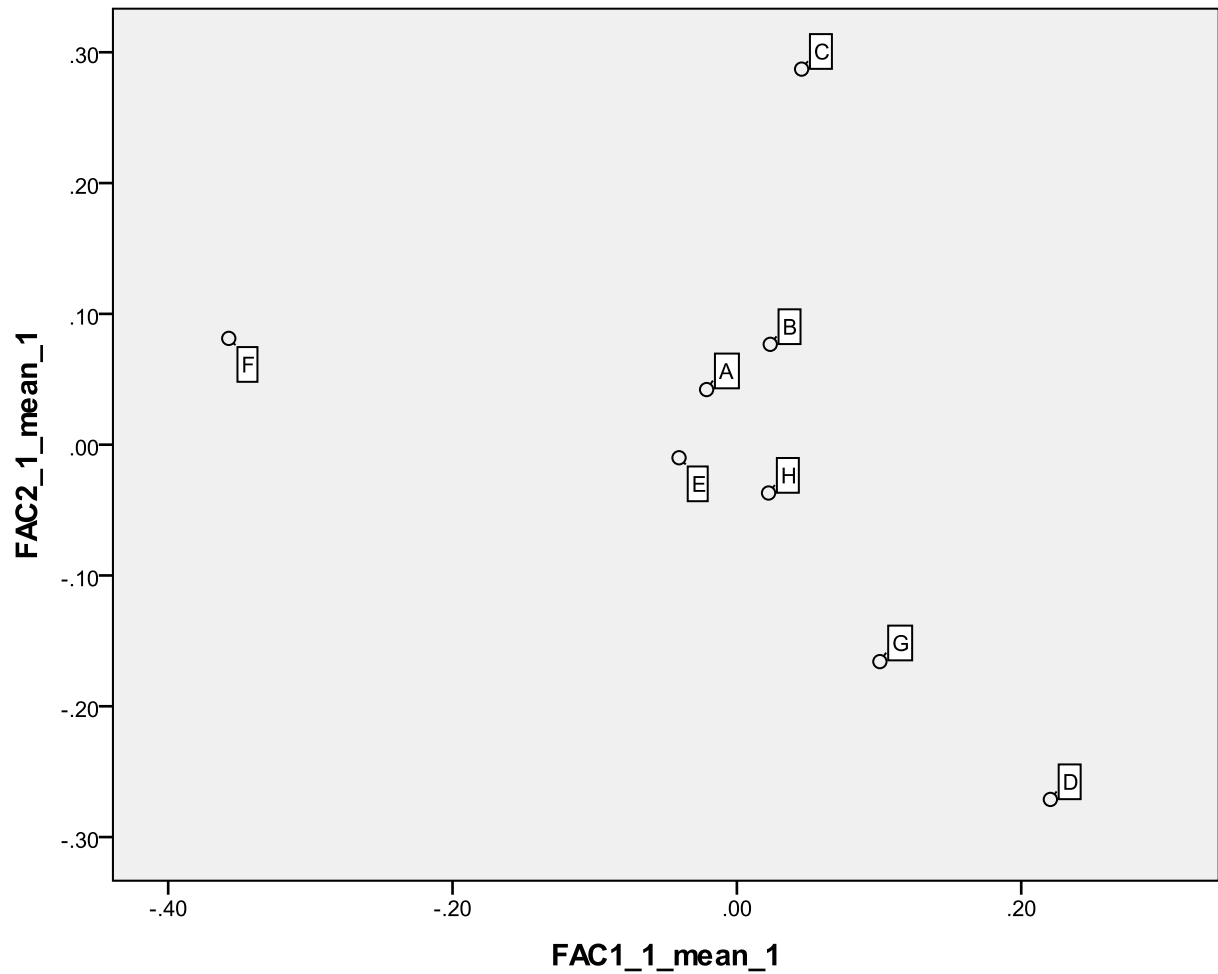
Đưa biến nhanhieu vào ô **Break Variable(s)** và 2 biến giá trị của các factor (được tạo ra khi chúng ta thực hiện việc **Save as Variables** trong **Scores** khi phân tích nhân tố.

Đồng thời, kích vào **Write a new data file containing only the aggregated variables** để lưu lại file **aggr.sav**.

- Vào thư mục có chứa file SPSS dùng để phân tích dữ liệu, trong đó tìm và mở file **aggr.sav** mới được tạo ra.
- Vào **Graphs / Legacy Dialogs / Scatter/Dot...**
- Chọn **Simple Scatter** và kích vào **Define** và thiết lập các biến vào các ô như hình vẽ, trong đó ô **Label Cases by**, chọn biến nhanhieu.



- Kích vào **OK** để vẽ sơ đồ:



- Sử dụng kết quả của hai bản đồ (một bản liên quan đến thuộc tính, một bản liên quan đến thương hiệu) sẽ hình thành biểu đồ nhận thức.
- Tương tự như hình trên, có thể sao chép dữ liệu vào SPSS để vẽ biểu đồ sao cho dễ nhìn và đẹp hơn.

CHƯƠNG

10

KIỂM ĐỊNH PHI THAM SỐ (NONPARAMETRIC ANALYSIS)

Trong các chương trước đã giới thiệu các phương pháp kiểm định tham số như kiểm định T (T-test), phân tích phương sai (ANOVA), hồi quy tuyến tính..., phần lớn những loại kiểm định này trong sách này đều có yêu cầu về phân phối của mẫu hoặc phương sai đồng nhất. Tuy nhiên, trong thực tiễn những dữ liệu thu thập được có thể vi phạm một hoặc tất cả các giả định mà các phân tích ra. Vì vậy, khi những dữ liệu đã thu thập về vi phạm của những giả định, thì nhà nghiên cứu cần phải lựa chọn và đánh giá kiểm định phi tham số (thay vì kiểm định tham số). Kết luận của kiểm định phi tham số có những yêu cầu hoặc giả định ít hơn về đặc trưng của mẫu hoặc tổng thể.

Trong khuôn khổ chương này sẽ giới thiệu các phương pháp kiểm định phi tham số sau:

- Kiểm định chi bình phương (chi-square) cho một biến đơn.
- Kiểm định chi bình phương (chi-square) về tính độc lập hay phụ thuộc giữa 2 biến.
- Kiểm định Mann-Whitney cho 2 mẫu độc lập.
- Kiểm định Kruskal-Wallis cho một nhiều mẫu độc lập.
- Kiểm định hạng có dấu Wilcoxon cho 2 mẫu phụ thuộc.
- Kiểm định Friedman cho nhiều mẫu phụ thuộc.

KIỂM ĐỊNH CHI BÌNH PHƯƠNG (CHI-SQUARE) CHO MỘT BIẾN ĐƠN

Kiểm định chi-bình phương là kiểm định được sử dụng thường xuyên nhất với dữ liệu định tính (thang đo biểu danh hoặc thang đo thứ tự), nơi tập hợp các dữ kiện thông tin quan sát được nhóm lại với nhau thành các nhóm riêng biệt, loại trừ lẫn nhau và nơi mà tần số xuất hiện trong mỗi loại. Kiểm định Chi-bình phương phương sai đơn so sánh tần số đã được quan sát với tần số sẽ được kỳ vọng nếu giả thuyết H_0 là đúng.

Thống kê chi-bình phương được tính bằng cách so sánh các giá trị đã được quan sát so với giá trị kỳ vọng cho mỗi loại, kiểm định sự khác biệt giữa chúng.

Giả định

- Các dữ liệu được giả định là một mẫu ngẫu nhiên.
- Độc lập giữa các quan sát ghi lại trong mỗi đối tượng (category).
- Tần suất mong đợi cho mỗi đối tượng cần có phải ít nhất là 5.

Ví dụ

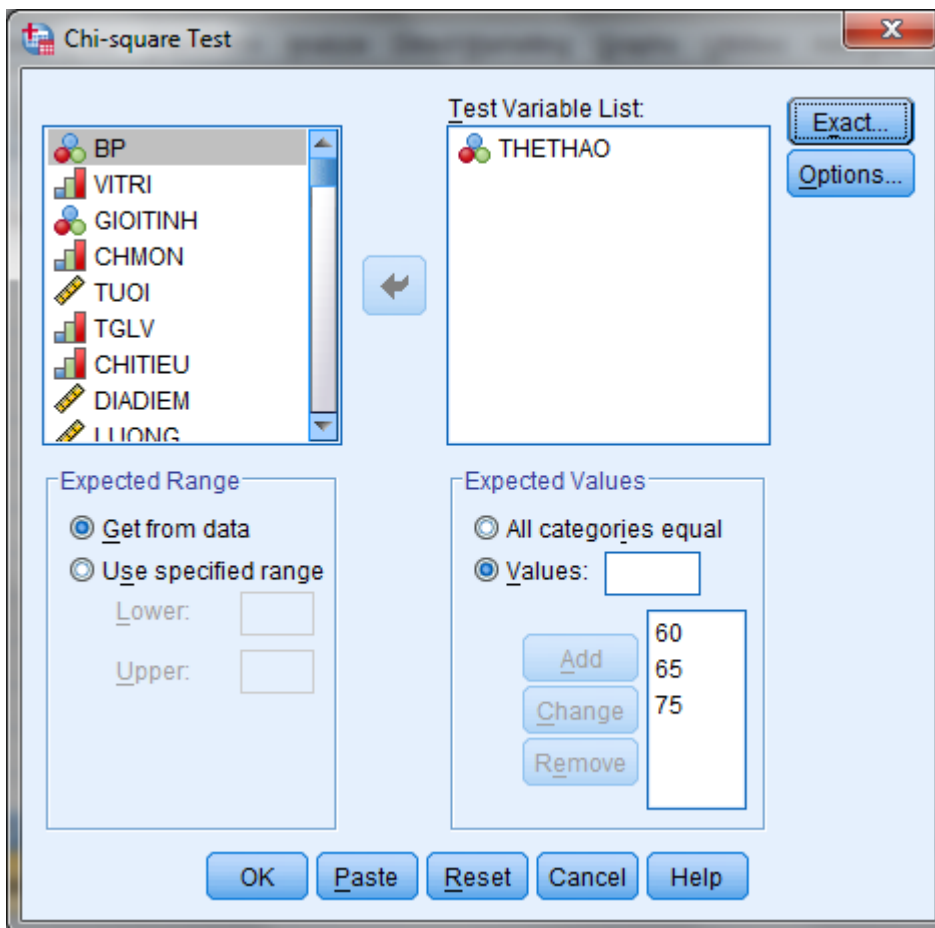
Giả sử nhà nghiên cứu quan tâm đến việc xác định liệu có một sự khác biệt về sở thích giữa những cán bộ công nhân viên đối với các loại hình thể thao được triển khai trong doanh nghiệp gồm bóng chuyền, cầu lông và tennis. Có hai hình thức kiểm định với tần suất bằng nhau và tần suất không bằng nhau:

- Giả định tần suất bằng nhau cho tất cả các nhóm;
- Tần suất không bằng nhau khi người nghiên cứu hi vọng rằng phân bố tần suất theo một số lượng nhất định;

Trong khuôn khổ nghiên cứu này sẽ trình bày phần tần suất không bằng nhau, nhà nghiên cứu hy vọng rằng sự phân bố tần số của 200 cán bộ công nhân viên đối với các hình thức thể thao là không giống nhau, cụ thể nhà nghiên cứu hy vọng sẽ là 60 người thích chơi bóng chuyền, 65 người thích chơi cầu lông và 75 người thích chơi tennis. Như vậy, câu hỏi nghiên cứu sẽ là liệu có một sự khác biệt giữa kết quả nghiên cứu thực tế và những gì nhà nghiên cứu hy vọng.

Thực hành trên SPSS

- Từ thanh trình đơn, kích **Analyze / Nonparametric Test / Legacy Dialogs / Chi-Square** và cửa sổ **Chi-square** sẽ mở ra.



- Chọn biến THETHAO vào ô **Test Variable List**, tại ô **Expected Values**, nếu để mặc định **All categories equal** thì sử dụng hình thức tần suất bằng nhau, nếu chọn **Values** và thiết đặt

giá trị thì theo mong muốn của người nghiên cứu. Trong trường hợp này chúng ta chọn **Values** thiết đặt 60 (mong đợi số người chơi bóng chuyền), 65 (mong đợi số người chơi cầu lông) và 75 (mong đợi số người chơi tennis).

- Kích **OK** để hoàn tất.

Loại thể thao triển khai			
	Observed N	Expected N	Residual
Bóng chuyền	57	60.0	-3.0
Cầu lông	60	65.0	-5.0
Tennis	83	75.0	8.0
Total	200		

Test Statistics	
	Loại thể thao triển khai
Chi-Square	1.388 ^a
Df	2
Asymp. Sig.	.500

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 60.0.

Từ bảng Test Statistics, ta có thể nhìn thấy rằng giá trị chi-square = 1.388, $p(\chi^2, df) = 0.50 > 0.05$ nên không có cơ sở để khẳng định rằng có sự khác biệt trong việc ưa thích ba loại thể thao triển khai tại doanh nghiệp.

KIỂM ĐỊNH CHI – BÌNH PHƯƠNG (χ^2) VỀ TÍNH ĐỘC LẬP HAY PHỤ THUỘC GIỮA HAI BIẾN

Đầu tiên sử dụng kiểm định Chi bình phương về tính độc lập hay phụ thuộc giữa hai biến, dĩ nhiên, đây là hai biến định tính (đo bằng thang đo biểu danh hoặc thứ tự). Một số chú ý về các thông số như sau:

- χ^2 (chi – bình phương) được thiết lập để xác định có hay không một mối liên hệ giữa hai biến, nhưng nó không chỉ ra được cường độ của mối liên hệ đó. Trong trường hợp này, cần sử dụng các đo lường kết hợp.
- χ^2 cho phép tìm ra những mối liên hệ phi tuyến tính giữa hai biến.
- Với kiểm định chi - bình phương, ta thành lập được các bảng chéo. Hệ số V Cramer được áp dụng cho tất cả các loại bảng chéo với k là chiều bé nhất của bảng chéo. Cường độ của nó biến động từ 0 đến 1.

$$V = \sqrt{\frac{\chi^2}{n(k-1)}}$$

Ví dụ

Nhà nghiên cứu muốn xem xét có tồn tại mối quan hệ giữa hai biến thích thể thao và giới tính. Giả thuyết được phát biểu như sau:

H_0 : Hai biến giới tính và thích thể thao độc lập với nhau trên tổng thể

H_A Hai biến giới tính và thích thể thao phụ thuộc lẫn nhau trên tổng thể.

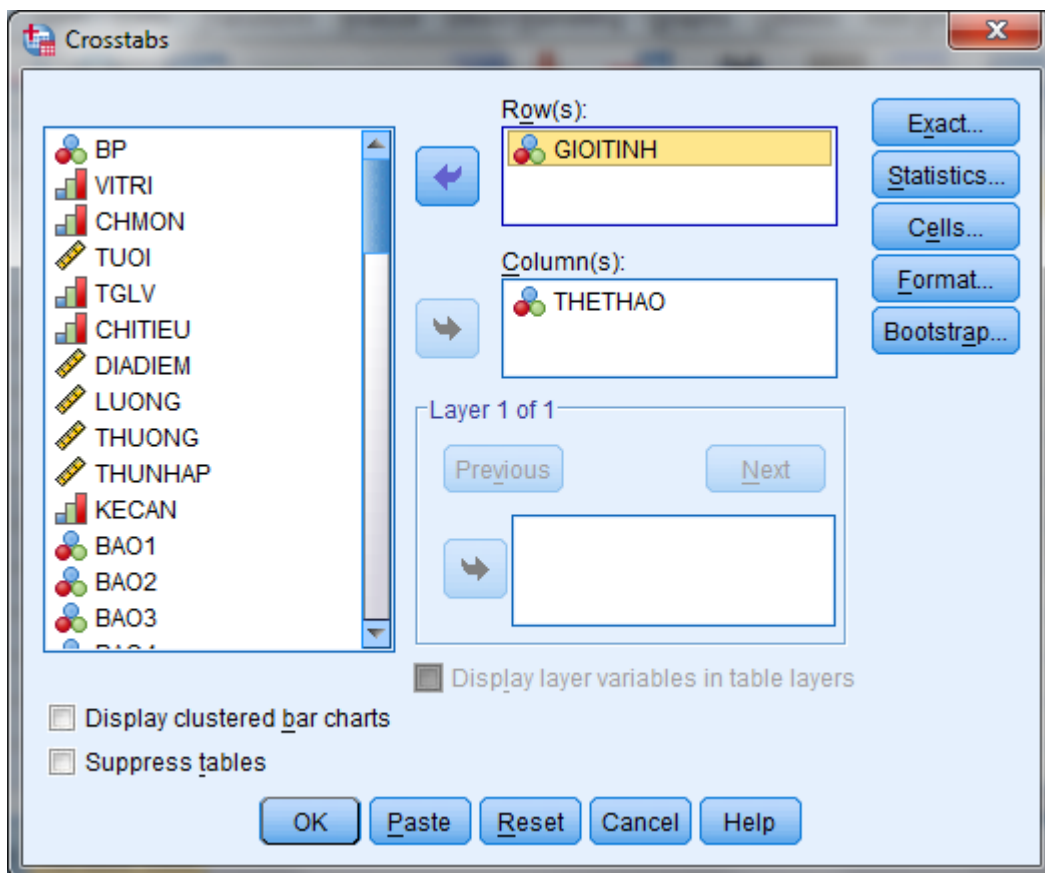
Giả định

Cũng như trường hợp chung, các giả định cho phương pháp này là:

- Các dữ liệu được giả định là một mẫu ngẫu nhiên.
- Độc lập giữa các quan sát ghi lại trong mỗi đối tượng (category).
- Tần suất mong đợi cho mỗi đối tượng cần có phải ít nhất là 5.

Thực hành trên SPSS

- Từ thanh thực đơn, chọn **Analyze / Descriptive Statistics / Crosstabs...** và cửa sổ **Crosstabs** sẽ mở ra, chọn GIOITINH vào ô Row(s) và THETHEO vào ô Column(s).



- Chọn **Cell** để mở cửa sổ **Crosstabs: Cell Display**. Ở phần **Counts**, kích vào các ô **Observed** và **Expected**. Ở phần **Percentages**, kích vào các ô **Row**, **Column**, **Total**. Chọn **Continue** để trở lại cửa sổ **Crosstabs**.
- Chọn **Statistics** để mở cửa sổ **Crosstabs: Statistics**, kích vào ô **Chi-square** và **Phi and Cramer's V** và sau đó chọn **Continue** để trở lại cửa sổ **Crosstabs**.
- Kích vào **OK** để thực hiện phân tích.
- Kết quả và diễn giải:

Giới tính * Loại thể thao triển khai Crosstabulation

			Loại thể thao triển khai			Total
			Bóng chuyền	Cầu lông	Tennis	
Giới tính	Nam	Count	28	34	42	104
		Expected Count	29.6	31.2	43.2	104.0
		% within Giới tính	26.9%	32.7%	40.4%	100.0%
		% within Loại thể thao triển khai	49.1%	56.7%	50.6%	52.0%
		% of Total	14.0%	17.0%	21.0%	52.0%
	Nu	Count	29	26	41	96
		Expected Count	27.4	28.8	39.8	96.0
		% within Giới tính	30.2%	27.1%	42.7%	100.0%
		% within Loại thể thao triển khai	50.9%	43.3%	49.4%	48.0%
		% of Total	14.5%	13.0%	20.5%	48.0%
Total		Count	57	60	83	200
		Expected Count	57.0	60.0	83.0	200.0
		% within Giới tính	28.5%	30.0%	41.5%	100.0%
		% within Loại thể thao triển khai	100.0%	100.0%	100.0%	100.0%
		% of Total	28.5%	30.0%	41.5%	100.0%

Trong bảng chéo Giới tính * Loại thể thao triển khai (**Giới tính * Loại thể thao triển khai Crosstabulation**) cho phép biết được số người nam, nữ thích triển khai các loại hình thể thao bóng chuyền, cầu lông hoặc tennis. Số lượng này có thể tính phần trăm theo dòng, cột và tổng số.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	.778 ^a	2	.678
Likelihood Ratio	.779	2	.677
Linear-by-Linear Association	.007	1	.935
N of Valid Cases	200		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 27.36.

Với kiểm định chi – bình phương (Chi-square Test) ta thấy, Chi-square = 0.778, df = 2 nên p (chi-square, df) = 0.678 > 0.05 nên chưa có đủ cơ sở để bác bỏ giả thuyết tức là chưa có đủ cơ sở để chấp nhận đối thuyết hay chưa có đủ cơ sở để khẳng định hai biến loại thể thao mà cán bộ công nhân viên tích với giới tính có mối quan hệ với nhau trên tổng thể.

KIỂM ĐỊNH MANN-WHITNEY CHO HAI MẪU ĐỘC LẬP

Kiểm định Mann-whitney là kiểm định phi tham số, nó được dùng để kiểm định giả thuyết đối với một biến định tính có 2 sự lựa chọn và một biến sử dụng ít nhất là thang đo thứ tự (hoặc khoảng cách, tỷ lệ). Nó thường dùng ở trong kiểm định t cho những nhóm độc lập khi có một sự vi phạm của giả thuyết thông thường hoặc khi dữ liệu là không thích hợp trong kiểm định t.

Ví dụ

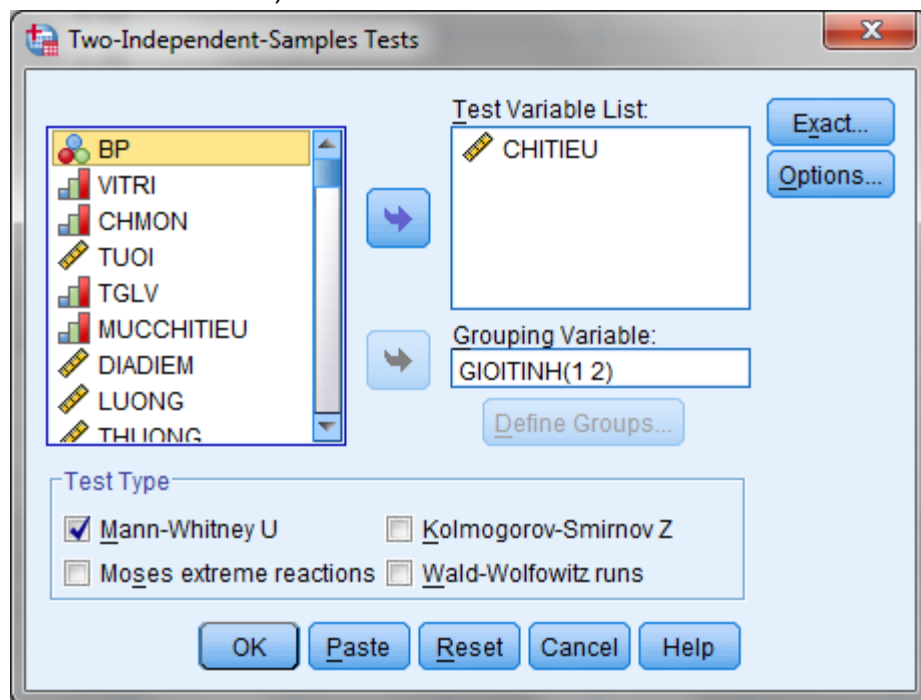
Nhà nghiên cứu muốn biết chi tiêu (CHITIEU) của cán bộ nam và nữ có khác nhau hay không nếu biết rằng biến số chi tiêu không thỏa mãn các điều kiện để phân tích tham số (t hoặc ANOVA). Kết quả khảo sát được lưu ở biến CHITIEU trong file THUC HANH SPSS.sav.

Giả định

- Dữ liệu phải độc lập, mẫu ngẫu nhiên.
- Dữ liệu phải đo lường tối thiểu ở cấp độ thứ bậc (ở đây dữ liệu biến chi tiêu được đo lường bằng thang tỷ lệ).

Thực hành trên SPSS

- Từ thanh thực đơn, kích **Analyze / Nonparametric Tests / 2 Independent samples**, cửa sổ **Two- Independent Samples tests** sẽ mở ra.
- Khi cửa sổ **Two Independent Samples tests** phần mềm SPSS đã mặc định phương pháp Mann-whitney U.
- Vì biến CHITIEU là biến phụ thuộc nên chọn vào ô **Test Variable List** và giới tính là biến độc lập nên chọn vào ô Grouping Variable, kích vào Define Groups để thiết đặt giá trị 1 và 2 (với 1 là Nam và 2 là Nữ).



- Kích **OK** để thực hiện phân tích.
- Kết quả và diễn giải:

Ranks				
	Giới tính	N	Mean Rank	Sum of Ranks
Mức chi tiêu	Nam	104	89.70	9329.00

	Nu	96	112.20	10771.00
	Total	200		

Test Statistics ^a	
	Mức chi tiêu
Mann-Whitney U	3869.000
Wilcoxon W	9329.000
Z	-2.752
Asymp. Sig. (2-tailed)	.006

a. Grouping Variable: Giới tính

Kiểm định giả thuyết bởi bản phân tích Mann-whitney là giá trị medians của 2 nhóm bằng nhau. Giá trị Mann - Whitney U bằng 3869 và Z-score bằng -2.752 với mức ý nghĩa là $0.006 < 0.05$ nên có thể rằng có sự khác biệt về chi tiêu giữa người nam và người nữ.

KIỂM ĐỊNH KRUSKAL – WALLIS CHO NHIỀU MẪU ĐỘC LẬP

Kiểm định Kruskal-Wallis là kiểm định phi tham số, trong đó các mẫu phải độc lập và có nhiều hơn 2 nhóm. Nó là gần giống như kiểm định ANOVA một chiều đã được trình bày ở phần trước nhưng vi phạm các yêu cầu của kiểm định tham số.

Ví dụ

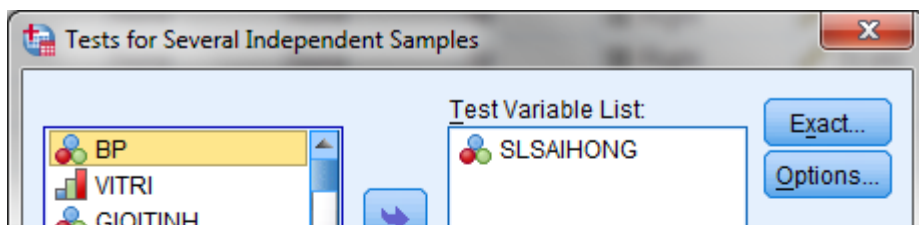
Với phương pháp quản lý cho phép sự linh hoạt của cán bộ công nhân viên với 3 mức là linh hoạt cao, trung bình và thấp. Nhà nghiên cứu muốn xem xét số lượng sai hỏng của 3 phương pháp quản lý. Kết quả sai hỏng được đo lường và nhập liệu tại file SLSAIHONG trong file THUC HANH SPSS.sav.

Giả thuyết

- Dữ liệu phải độc lập, mẫu ngẫu nhiên.
- Dữ kiện phải đo lường tối thiểu ở cấp bậc thứ bậc (ở đây được đo lường bởi thang đo tỷ lệ).
- Phải có ít nhất 5 số liệu cho mỗi nhóm trong mẫu.

Phương pháp cửa sổ

- Từ thanh thực đơn, kích **Analyze / Nonparametric Tests / Legacy Dialogs / K-Independent Sample**, sau đó, cửa sổ **Test for Several Independent Samples** sẽ được mở.
- Khi cửa sổ **Tests for Several Independent Samples** mở ra đã mặc định phương pháp Kruskal-Wallis H.
- Chọn biến SLSAIHONG vào ô **Test Variable List** và biến LINHHOAT vào ô **Grouping Variable**, sau đó định nghĩa các giá trị của biến linh hoạt (1, 3) trong **Define Range** như hình vẽ.



- Kích vào OK để hoàn tất việc thiết đặt và phân tích kết quả.
- Kết quả và diễn giải:

Ranks			
	Mức độ linh hoạt	N	Mean Rank
Số lượng sai hỏng / tháng	X- Linh hoạt cao	63	99.54
	Y - Linh hoạt trung bình	63	103.21
	Z- Linh hoạt thấp	74	99.01
	Total	200	

Test Statistics ^{a,b}	
	Số lượng sai hỏng / tháng
Chi-Square	.206
df	2
Asymp. Sig.	.902

a. Kruskal Wallis Test

b. Grouping Variable: Mức độ linh hoạt

Giả thuyết H_0 trong phương pháp kiểm định Kruskal-Willis là 3 phương pháp quản lý theo mức độ linh hoạt (cao, trung bình, thấp) có số lượng sai hỏng như nhau.

Vì Chi-square = 0.206, df = 2 nên p (Chi-square, df) = 0.902 > 0.05 nên chưa có cơ sở để bác bỏ giả thuyết H_0 tức là chưa có cơ sở chấp nhận H_1 hay nói rõ hơn là chưa có cơ sở để khẳng định với phương pháp quản lý theo mức độ linh hoạt có số sai hỏng khác nhau.

KIỂM ĐỊNH DẤU VÀ HẠNG WILCOXON CHO HAI MẪU CÓ MỐI QUAN HỆ VỚI NHAU

Kiểm định dấu và hạng là kiểm định phi tham số được sử dụng trong trường hợp hai mẫu quan hệ (liên hệ) nhằm so sánh sự khác nhau của giá trị trung bình của hai tổng thể nhưng không

cần quan tâm đến giả thuyết về phân phối của hai mẫu (hoặc hai mẫu không đảm bảo giả thuyết về phân phối). Kiểm định dấu và hạng cho phép người nghiên cứu biết được dấu (-/+) giữa các cặp quan sát và thông tin về độ lớn của chênh lệch giữa các cặp quan sát.

Trong phần kiểm định tham số trung bình hai mẫu phụ thuộc hay mẫu phối hợp từng cặp (Paired Samples T Test) đã giới thiệu một nghiên cứu liên quan đến thử nghiệm phương thức lãnh đạo mới về kiểm định có hay không sự khác biệt về nhận thức của cán bộ công nhân viên trước và sau khi thử nghiệm phương thức lãnh đạo mới. Để thực hiện Paired Sample T Test, giả định rằng:

- Độ lớn của hai mẫu phải tương đương nhau,
- Phân phối của sự khác biệt của các tham số trung bình phải là phân phối chuẩn.

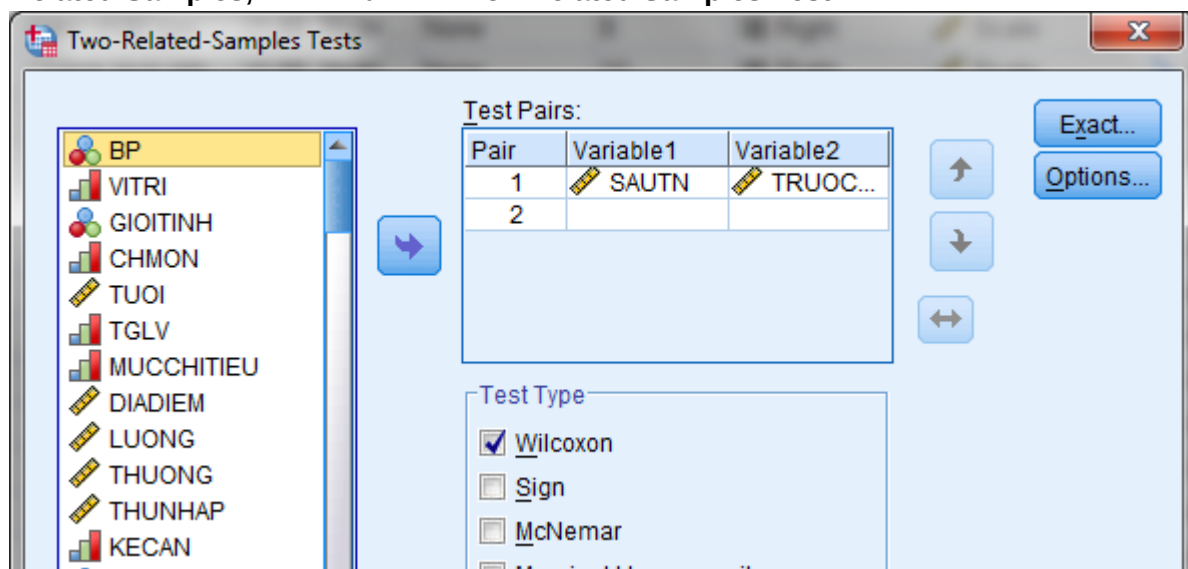
Tuy nhiên, trong trường hợp này, giả sử dữ liệu thu thập không thỏa mãn yêu cầu và các giả định đưa ra, vì vậy, kiểm định dấu và hạng Wilcoxon được dùng để trả lời câu hỏi trên.

Giả định

- Thang đo lường cho mỗi nhóm ít nhất là thang đo thứ tự (ở đây sử dụng thang đo tỷ lệ).
- Sự khác nhau (sự chênh lệch) điểm số cũng phải ít nhất là thang đo thứ tự.

Thực hành trên SPSS

- Từ thanh trình đơn, nhấp vào **Analyze / Nonparametric Tests / Legacy Dialogs / 2 Related Samples**, khi đó cửa sổ **Two - Related Samples Test** sẽ mở ra.



- Chọn SAUTN vào biến 1 (Variable 1) cặp 1 (pair 1) và TRUOCTN vào biến 2 (Variable 2) như hình vẽ.
- Mặc định kiểm định **Wilcoxon**.
- Kích vào **OK** để hoàn tất việc thiết đặt và tính kết quả.
- Kết quả và diễn giải:

		N	Mean Rank	Sum of Ranks
Cảm nhận trước khi thử nghiệm - Cảm nhận sau khi thử nghiệm	Negative Ranks	131 ^a	86.97	11392.50
	Positive Ranks	34 ^b	67.72	2302.50
	Ties	35 ^c		
	Total	200		

a. Cảm nhận trước khi thử nghiệm < Cảm nhận sau khi thử nghiệm

b. Cảm nhận trước khi thử nghiệm > Cảm nhận sau khi thử nghiệm

c. Cảm nhận trước khi thử nghiệm = Cảm nhận sau khi thử nghiệm

Hạng trung bình của 131 chênh lệch âm là 86.97 và hạng trung bình của 34 chênh lệch dương là 67.72, có 35 cặp quan sát có giá trị trước và sau khi thử nghiệm bằng nhau.

Test Statistics^b

	Cảm nhận trước khi thử nghiệm - Cảm nhận sau khi thử nghiệm
Z	-7.462 ^a
Asymp. Sig. (2-tailed)	.000

a. Based on positive ranks.

b. Wilcoxon Signed Ranks Test

Với giá trị $Z = -7.462$ và $p(Z) = 0.000 < 0.05$ nên bác bỏ giả thuyết H_0 nghĩa là cán bộ công nhân viên thích áp dụng chiến lược quản lý mới hơn (vì cảm nhận trước khi thử nghiệm < cảm nhận trước khi thử nghiệm).

KIỂM ĐỊNH FRIEDMAN NHIỀU HƠN HAI MẪU CÓ QUAN HỆ

Đối với kiểm định Wilcoxon được sử dụng để phân tích (so sánh) giá trị trung bình của 2 mẫu, kiểm định Friedman sử dụng để kiểm định giá trị của hơn 2 mẫu nhưng không thỏa mãn những yêu cầu và giả định của phân tích tham số (phân tích ANOVA). Giống như kiểm định Mann-Whitney và Kruskal-Wallis, sự tính toán của kiểm định Friedman dựa vào các hạng trong mỗi trường hợp.

Cách thức phân tích và diễn giải kết quả giống với kiểm định Wilcoxon, kiểm định này thực hiện bằng cách nhấp vào **Analyze / Nonparametric Tests / Legacy Dialogs / 2 Related Samples**, khi đó cửa sổ **Two - Related Samples Test** và thiết đặt các thông số vào ô Two – Related Samples Test.

TÀI LIỆU THAM KHẢO

- Al-Qirim, N. (2005). An Empirical Investigation of An E-Commerce Adoption-Capability Model in Small Business in New-Zealand. *Electronic Markets*, 15(4), 418-437.
- Auger, P., BarNir, A., & Gallaughier, J. M. (2003). Strategic Orientation, Competition, and Internet– Based Electronic Commerce. *Information Technology and Management*, 4(2), 139-146.
- Babbie, E. R. (1986). *The practice of social research* (4 ed.). Belmont, Calif.: Wadsworth Pub. Co. .
- Chieochan, O., Lindley, D., & Dunn, T. (2002). Factors Affecting the Use Information Technology in Thai Agricultural Cooperatives: A Work in Progress. *The Electronic Journal on Information Systems in Developing Countries, University of Hong Kong*, 2(1), 1-15.
- Chuang, T. T., Nakatani, K., Chen, J. C. H., & Huang, I. L. (2007). Examining the Impact of Organisational and Owner's Characteristics on the Extent of E-Commerce Adoption in SMEs. *Int. J. Business and Systems Research*, 1(1), 61-80.
- Chwelos, P., Benbasat, I., & Dexter, A. (2001). Research Report: Empirical Test of an EDI Adoption Model. *Information Systems Research*, 12(3), 304-321.
- Cooper, D. R., & Schindler, P. S. (2011). *Business research methods* (11 ed.). New York: McGraw-Hill Higher Education.
- Dholakia, R. R., & Kshetri, N. (2004). Factors Impacting the Adoption of the Internet Among SMEs. *Small Business Economics*, 23, 311-322.
- Gill, J., & Johnson, P. (2002). *Research Methods for Managers* (3 ed.). London: Sage Publishing. .
- Grover, V., & Goslar, M. D. (1993). The Initiation, Adoption, and Implementation of Telecommunications Technologies in U. S. Organizations. *Journal of Management Information Systems*, 10(1), 141-164.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate Data Analysis* (7 ed.): Prentice Hall.
- Hair, J. F., Bush, R. P., & Ortineau, D. J. (2002). *Marketing Research: The McGraw-Hill Companies*.
- Hong, W., & Zhu, K. (2006). Migrating to Internet-based E-commerce: Factors Affecting E-commerce Adoption and Migration at the Form Level. *Information & Management*, 41, 204-221.
- Huy, L. V., & Filiatrault, P. (2006). *The Adoption of e-Commerce in SMEs in Vietnam: A Study of Users and Prospectors*. Paper presented at the 10th Pacific Asia Conference on Information Systems (PACIS), Kuala Lumpur, Malaisie.
- Jeon, B. N., Han, K. S., & Lee, M. J. (2006). Determining Factors for the Adoption of e-Business: the Case of SMEs in Korea. *Applied Economics*, 38, 1905-1916.
- Kendall, J. D., Tung, L. L., Chua, K. H., Ng, C. H. D., & Tan, S. M. (2001). Receptivity of Singapore's SMEs to Electronic Commerce Adoption. *Journal of Strategic Information Systems*, 10, 223-242.
- Kerlinger, F. N. (1986). *Foundations of behavioral research* (3 ed.). New York: Holt, Rinehart and Winston
- Kothari, C. R. (2004). *Research Methodology: Methods and Techniques* (2 ed.): New Age International (P) Limited.

- Lertwongsatien, C., & Wongpinunwatana, N. (2003). E-commerce Adoption in Thailand: An Empirical Study of Small and Medium Enterprises (SMEs). *Journal of Global Technology Management*, 6(3), 67-83.
- Limthongchai, P., & Speece, M. W. (2003). E-commerce Adoption in Thailand: An Empirical Study of Small and Medium Enterprises (SMEs). *Journal of Global Technology Management*, 6(3), 67-83.
- Ling, C. Y. (2001). *Model of Factors Influences on Electronic Commerce Adoption and Diffusion in Small & Medium Sized Enterprises*. Paper presented at the ECIS Doctoral Consortium, 24-26 June, AIS region 2 (Europe, Africa, Middle-East).
- Looi, H. C. (2005). E-commerce Adoption in Brunei Darussalam: A Quantitative Analysis of Factors Influencing Its Adoption. *Communication of the Association for Information Systems*, 15, 61-81.
- Molla, A., & Licker, P. S. (2005). eCommerce Adoption in Developing Countries: A Model and Instrument. *Information & Management*, 41, 877-899.
- Pierre, J. L., & Delbecq, A. L. (1977). Organizational Structure, Individual Attitudes and Innovation. *Academy of Management Review*, 2(1), 27-37.
- Premkumar, G., & Ramamurthy, K. (1995). The Role if Inter-organizational and Organizational Factors on the Decision Mode for Adoption of Inter-organizational Systems. *Decision Science*, 26(3), 303-336.
- Premkumar, G., Ramamurthy, K., & Nilakanta, S. (1994). Implementation of Electronic Data Interchange: An Innovation Diffusion Perspective. *Journal of Management Information Systems*, 11(2), 157-186.
- Saunders, M., Lewis, P., & Thornhill, A. (2007). *Research method for Business Students*. England: Person Education Limited, Edinburgh Gate, Harlow, Essex CM202 JE.
- Scupola, A. (2003). The Adoption of Internet Commerce by SMEs in the South of Italy: An Environmental, Technological and Organisational Perspective. *Journal of Global Information Technology Management*, 6(1), 52-71.
- Seyal, A. H., Awais, M. M., Shamail, S., & Abbas, A. (2004). Determinants of Electronic Commerce in Pakistan: Preliminary Evidence from Small and Medium Enterprises. *Electronic Markets*, 14(4), 372-387.
- Seyal, A. H., & Rahman, M. N. N. A. (2003). A Preliminary Investigation of e-Commerce Adoption in Small & Medium Enterprises in Brunei. *Journal of Global Information Technology Management*, 6(2), 6-26.
- Tan, M., & Teo, T. S. H. (2000). Factors Influencing the Adoption of Internet Banking. *Journal of the Association for Information Systems*, 1(5), 15.
- Teo, T. S. H., & Tan, M. (1998). An Empirical Study of Adopters and Non-Adopters of the Internet in Singapore. *Information & Management*, 34, 339-345.
- Thọ, N. Đ. (2011). *Phương pháp nghiên cứu khoa học trong kinh doanh: Thiết kế và thực hiện*: Nhà xuất bản Lao động - Xã hội.
- Thong, J. Y. L. (1999). An Integrated Model of Information Systems Adoption in Small Businesses. *Journal of Management Information Systems*, 15(4), 187-214.
- Thong, J. Y. L., & Yap, C. S. (1995). An Integrated Model of Information Systems Adoption in Small Businesses. *Journal of Management Information System*, 15(4), 275-290.
- To, M. L., & Ngai, E. W. T. (2007). The Role of Managerial Attitudes in the adoption of Technological Innovations: An Application to B2C E-commerce. *International Journal of Enterprise Information Systems*, 3(2), 23-33.

$$MSW = \frac{SSW}{n-k}$$

Trọng, H., & Ngọc, C. N. M. (2008). *Phân tích dữ liệu với SPSS*: Nhà xuất bản Hồng Đức.

Truong, D., & Rao, S. S. (2002). *Development of A Contingency Model for Adoption of Electronic Commerce*. Paper presented at the Decision Sciences Institute, Annual meeting proceedings.

Valaseca-Requena, J., Torrent-Sellens, J., Messeguer-Artola, A., & Rodriguez-Ardura, I. (2007). An Integrated Model of the Adoption and Extent of E-Commerce in Firms. *International Advances in Economic Research*, 13, 222-241.