

TRƯỜNG ĐẠI HỌC TÀI CHÍNH – MARKETING
BỘ MÔN TOÁN THỐNG KÊ

Giáo Trình

PHƯƠNG PHÁP NGHIÊN
CỨU KHOA HỌC

(Dành cho chương trình chất lượng cao)

Mã số : GT – 05 – 19

Nhóm biên soạn:

Nguyễn Huy Hoàng (Chủ biên)

Nguyễn Trung Đông

Nguyễn Văn Phong

Dương Thị Phương Liên

THÀNH PHỐ HỒ CHÍ MINH - 2020

MỤC LỤC

	Trang
Lời mở đầu.....	7
Một số ký hiệu.....	9
Chương 1. Giới thiệu chung về nghiên cứu khoa học và phương pháp nghiên cứu kinh tế.....	10
1.1. Giới thiệu nghiên cứu là gì.....	10
1.1.1. Nghiên cứu.....	10
1.1.2. Nghiên cứu khoa học.....	10
1.1.3. Vai trò của nghiên cứu khoa học.....	11
1.1.4. Nghiên cứu kinh tế.....	11
1.1.5. Phương pháp nghiên cứu	11
1.2. Phân biệt loại hình nghiên cứu.....	12
1.2.1. Nghiên cứu cơ bản.....	12
1.2.2. Nghiên cứu ứng dụng.....	12
1.2.3. So sánh giữa nghiên cứu cơ bản và nghiên cứu ứng dụng.....	12
1.3. Phân biệt phương pháp nghiên cứu định tính, định lượng và hỗn hợp.....	13
1.3.1. Giới thiệu về phương pháp nghiên cứu định tính.....	13
1.3.2. Giới thiệu về phương pháp nghiên cứu định lượng.....	14
1.3.3. So sánh phương pháp nghiên cứu định tính, định lượng.....	14
1.3.4. Giới thiệu phương pháp nghiên cứu hỗn hợp.....	15
1.4. Quy trình nghiên cứu.....	16
1.4.1. Khái niệm quy trình nghiên cứu.....	16
1.4.2. Nội dung các bước của quy trình nghiên cứu.....	16
1.5. Các cấu phần cơ bản của một nghiên cứu.....	21
1.6. Đạo đức trong nghiên cứu khoa học.....	22
1.6.1. Thế nào là đạo đức nghiên cứu khoa học.....	22
1.6.2. Các nguyên tắc đạo đức trong nghiên cứu khoa học.....	22
1.7. Câu hỏi thảo luận.....	23
Thuật ngữ chính chương 1.....	24

Chương 2. Tổng quan tình hình nghiên cứu và câu hỏi nghiên cứu.....	25
2.1. Giới thiệu về tổng quan nghiên cứu.....	25
2.1.1. Định nghĩa tổng quan nghiên cứu.....	25
2.1.2. Vai trò của tổng quan nghiên cứu.....	25
2.1.3. Tổng quan nghiên cứu tốt.....	25
2.2. Nội dung và yêu cầu phần tổng quan.....	26
2.2.1. Các trường phái lý thuyết là cơ sở lý luận cho nghiên cứu.....	26
2.2.2. Bối cảnh nghiên cứu và các nhân tố chính.....	26
2.2.3. Các phương pháp nghiên cứu chính.....	27
2.2.4. Các kết quả nghiên cứu chính.....	27
2.2.5. Hạn chế của những nghiên cứu trước và khoảng trống tri thức.....	27
2.3. Một số kỹ năng tiến hành tổng quan.....	27
2.4. Giới thiệu câu hỏi nghiên cứu.....	28
2.4.1. Khái niệm.....	28
2.4.2. Các loại câu hỏi nghiên cứu.....	28
2.4.3. Làm thế nào để xác lập câu hỏi nghiên cứu tốt.....	29
2.5. Tiêu chuẩn và cách xây dựng câu hỏi nghiên cứu.....	30
2.5.1. Câu hỏi hướng tới vấn đề mang tính quy luật.....	30
2.5.2. Câu hỏi có cơ sở thực tiễn/hoặc lý thuyết.....	30
2.5.3. Các nhân tố, yếu tố trong câu hỏi có phạm vi, ý nghĩa rõ ràng.....	30
2.5.4. Câu hỏi có khả năng trả lời được.....	31
2.6. Câu hỏi thảo luận.....	31
Thuật ngữ chính chương 2.....	32
Chương 3. Phát triển khung lý thuyết, khung khái niệm và khung phân tích.....	33
3.1. Khung lý thuyết (theoretical framework).....	33
3.1.1. Giới thiệu về khung lý thuyết.....	33
3.1.2. Các cấu phần cơ bản của khung lý thuyết.....	33
3.1.3. Các bước xây dựng khung lý thuyết.....	34
3.2. Khung khái niệm (conceptual framework).....	35
3.3. Khung phân tích (analytic Framework).....	35
3.4. Câu hỏi thảo luận.....	36

Thuật ngữ chính chương 3.....	37
Chương 4. Nghiên cứu định lượng: phương pháp khảo sát	38
4.1. Giới thiệu khái niệm.....	38
4.1.1. Phương pháp khảo sát là gì?.....	38
4.1.2. Khi nào dùng phương pháp khảo sát?.....	38
4.2. Xác định mẫu khảo sát.....	38
4.2.1. Mẫu và tổng thể.....	38
4.2.2. Quy trình chọn mẫu.....	39
4.2.3. Các phương pháp chọn mẫu cơ bản.....	39
4.2.4. Tính đại diện của mẫu.....	41
4.2.5. Xác định cỡ mẫu.....	42
4.3. Thiết kế bảng khảo sát.....	43
4.3.1. Những bước chính khi thiết kế bảng khảo sát.....	43
4.3.2. Những chú ý khi thiết kế từng câu hỏi.....	44
4.3.3. Những chú ý khi thiết kế tổng thể bảng câu hỏi.....	45
4.4. Xây dựng và thực hiện quy trình khảo sát.....	46
4.5. Quy trình chuẩn bị số liệu.....	47
4.5.1. Nhập liệu.....	47
4.5.2. Kiểm định các thước đo.....	47
4.6. Câu hỏi thảo luận.....	48
Thuật ngữ chính chương 4.....	49
Chương 5. Nghiên cứu định lượng: phương pháp thử nghiệm.....	50
5.1. Giới thiệu.....	50
5.2. Yêu cầu cơ bản của phương pháp thử nghiệm.....	50
5.2.1. Đảm bảo phân nhóm phải ngẫu nhiên.....	50
5.2.2. Sử dụng nhóm đối chứng.....	51
5.2.3. Biến độc lập (yếu tố can thiệp) đủ mạnh.....	51

5.3. Thiết kế thử nghiệm có đối chứng.....	51
5.3.1. Chỉ đo lường sau thử nghiệm.....	51
5.3.2. Đo lường trước – sau thử nghiệm.....	52
5.4. Áp dụng nghiên cứu thử nghiệm trên thực địa.....	53
5.4.1. Đánh giá tác động của dự án hoặc chính sách.....	53
5.4.2. Đánh giá tác động của biến động trên thực địa	53
5.5. Câu hỏi thảo luận.....	54
Thuật ngữ chính chương 5.....	55
Chương 6. Thiết kế nghiên cứu tổng thể.....	56
6.1. Giới thiệu khái niệm.....	56
6.1.1. Khái niệm.....	56
6.1.2. Vai trò của thiết kế nghiên cứu.....	56
6.1.3. Phân biệt thiết kế nghiên cứu với đề cương báo cáo.....	56
6.2. Các yêu cầu chính trong thiết kế nghiên cứu.....	57
6.2.1. Tính chặt chẽ.....	57
6.2.2. Tính khái quát.....	57
6.2.3. Tính khả thi.....	58
6.3. Giới thiệu một số thiết kế nghiên cứu.....	58
6.3.1. Các bước thiết kế nghiên cứu	58
6.3.2. Ví dụ về một thiết kế nghiên cứu cụ thể.....	58
6.4. Câu hỏi thảo luận.....	60
Thuật ngữ chính chương 6.....	61
Chương 7. Xử lý và phân tích dữ liệu định lượng cho nghiên cứu.....	62
7.1. Tổng quan kiến thức về thống kê và sử dụng các kỹ thuật thống kê.....	62
7.1.1. Phân tích mô tả và khám phá.....	62
7.1.2. So sánh nhóm.....	64
7.2. Hồi quy tuyến tính cho phân tích dữ liệu định lượng.....	65
7.2.1. Mô hình hồi quy đơn.....	65

7.2.2. Mô hình hồi quy tuyến tính k biến.....	73
7.3. Câu hỏi thảo luận và bài tập.....	77
Thuật ngữ chính chương 7.....	80
Chương 8. Hướng dẫn viết đề cương nghiên cứu và báo cáo nghiên cứu.....	81
8.1. Tổng hợp kiến thức để viết đề cương.....	81
8.2. Hình thức và trình tự của đề cương nghiên cứu khoa học.....	81
8.3. Báo cáo kết quả nghiên cứu.....	82
8.3.1. Tổng quan.....	82
8.3.2. Các thành phần của một báo cáo.....	82
8.3.3. Các nguyên tắc khi viết báo cáo.....	84
8.3.4. Thuyết trình kết quả.....	85
8.4. Hướng dẫn phần tài liệu tham khảo.....	85
8.4.1. Ghi và sử dụng trích dẫn từ sách.....	85
8.4.2. Đối với tài liệu ngoài Việt Nam, nên ghi rõ tên thành phố, bang (nếu có) và tên nước.....	86
8.4.3. Trích dẫn bài báo đăng trên tạp chí khoa học.....	86
8.5. Ví dụ về đề cương chi tiết.....	86
8.6. Câu hỏi thảo luận.....	94
Thuật ngữ chính chương 8.....	95
Phụ lục 1. Bảng câu hỏi khảo sát.....	96
Phụ lục 2. Hướng dẫn sử dụng phần mềm SPSS cơ bản.....	105
Phụ lục 3. Hướng dẫn sử dụng phần mềm Eviews 8.0.....	142
Tài liệu tham khảo.....	191

LỜI MỞ ĐẦU

Phương pháp nghiên cứu khoa học là môn học bắt buộc, đối với sinh viên hệ đại học, chương trình chất lượng cao, của Trường đại học Tài chính – Marketing. Tuy nhiên sinh viên cũng như giảng viên còn gặp nhiều khó khăn trong việc lựa chọn tài liệu học tập phù hợp với đối tượng; chính vì vậy được sự đồng ý của Nhà trường chúng tôi mạnh dạn biên soạn “ **Giáo trình Phương pháp nghiên cứu khoa học**”. Đây là giáo trình dành cho đối tượng sinh viên khối ngành kinh tế và quản trị kinh doanh, thời lượng 3 tín chỉ (45 tiết giảng); Khó khăn là, sinh viên nắm nhất kiến thức về kinh tế cũng như các công cụ định lượng còn rất khiêm tốn; sinh viên mới chỉ học Toán cao cấp và Kinh tế vi mô, Kinh tế vĩ mô; do đó chúng tôi cố gắng lựa chọn các nội dung căn bản, trọng yếu và có thể áp dụng trong nghiên cứu kinh tế và quản trị kinh doanh; lấy ví dụ trực tiếp từ các nghiên cứu cụ thể; giáo trình được biên tập trên cơ sở tham khảo nhiều giáo trình quốc tế cũng như trong nước (xem phần tài liệu tham khảo), và kinh nghiệm giảng dạy nhiều năm của các tác giả; giáo trình dành cho sinh viên chương trình chất lượng cao, nên chúng tôi cũng quan tâm việc giới thiệu thuật ngữ Anh – Việt, giúp sinh viên có thể tự đọc, tự nghiên cứu các tài liệu tham khảo bằng tiếng Anh.

Nội dung giáo trình đã được thiết kế phù hợp với chương trình đào tạo và trình độ của sinh viên khối ngành kinh tế và quản trị kinh doanh. Giáo trình bao gồm 8 chương và một số phụ lục:

Chương 1. Giới thiệu chung về nghiên cứu khoa học và phương pháp nghiên cứu kinh tế. Cuối chương là một số thuật ngữ Anh – Việt.

Chương 2. Tổng quan tình hình nghiên cứu và câu hỏi nghiên cứu. Cuối chương là một số thuật ngữ Anh – Việt.

Chương 3. Phát triển khung lý thuyết, khung khái niệm và khung phân tích. Cuối chương là một số thuật ngữ Anh – Việt.

Chương 4. Nghiên cứu định lượng phương pháp khảo sát. Cuối chương là một số thuật ngữ Anh – Việt.

Chương 5. Nghiên cứu định lượng phương pháp thử nghiệm. Cuối chương là một số thuật ngữ Anh – Việt.

Chương 6. Thiết kế nghiên cứu tổng thể. Cuối chương là một số thuật ngữ Anh – Việt.

Chương 7. Xử lý và phân tích dữ liệu định lượng cho nghiên cứu. Cuối chương là một số thuật ngữ Anh – Việt.

Chương 8. Hướng dẫn viết đề cương nghiên cứu và báo cáo nghiên cứu. Cuối chương là một số thuật ngữ Anh – Việt.

Phần cuối, chúng tôi biên soạn một số số phụ lục cần thiết về sử dụng các phần mềm thông dụng như: SPSS, EVIEWS, và cách lập bảng hỏi, giúp sinh viên có thể tự tra cứu (ThS. Nguyễn Văn Phong và ThS. Dương Thị Phương Liên tham gia biên soạn phụ lục và một số ví dụ minh họa).

Giáo trình cũng là tài liệu tham khảo tốt cho các học viên cao học và nghiên cứu sinh!

Giáo trình do Giảng viên cao cấp, TS. Nguyễn Huy Hoàng và ThS. Nguyễn Trung Đông là các giảng viên của Bộ môn Toán – Thống kê, Khoa Kinh tế – Luật trường đại học Tài chính – Marketing, đã có nhiều năm kinh nghiệm giảng dạy cùng biên tập. Lần đầu biên soạn, nên giáo trình này không tránh khỏi còn thiếu sót. Rất mong nhận được sự góp ý của các độc giả để lần sau giáo trình được hoàn thiện hơn.

Mọi ý kiến đóng góp xin gửi về địa chỉ email:

hoangtoanb@ufm.edu.vn và nguyendong@ufm.edu.vn.

Xin trân trọng cảm ơn Trường đại học Tài chính – Marketing đã hỗ trợ kinh phí và tạo điều kiện cho giáo trình sớm đến với bạn đọc!

Tp. HCM, Tháng 05 năm 2020

Các tác giả

MỘT SỐ KÝ HIỆU

1. $E(X)$: Kỳ vọng của X .
2. \bar{X} : Trung bình (mẫu) của X .
3. $\text{Var}(X)$: Phương sai của X .
4. $\text{Se}(X)$: Độ lệch chuẩn của X .
5. S_X^2 : Phương sai (mẫu) của X .
6. $\text{Cov}(X, Y)$: Hiệp phương sai của hai biến X, Y .
7. $S_{X,Y}$: Hiệp phương sai (mẫu) của hai biến X, Y .
8. $\text{Cor}(X, Y)$: Hệ số tương quan của hai biến X, Y .
9. $r_{X,Y}$: Hệ số tương quan (mẫu) của hai biến X, Y .
10. ε : Sai số ngẫu nhiên tổng thể.
11. e : Phần dư (sai số ngẫu nhiên mẫu).
12. H_0 : Giả thuyết H_0 . H_1 : Đối thuyết H_1 .
13. $X \sim N(\mu, \sigma^2)$: X có phân phối chuẩn với kỳ vọng μ và phương sai σ^2 .
14. $T \sim \text{St}(n)$: T có phân phối Student với n bậc tự do.
15. $\chi^2(n)$: Phân phối chi bình phương với n bậc tự do.
16. $F \sim F(n, m)$: F có phân phối Fisher với bậc tự do của tử là n và bậc tự do của mẫu là m .
17. R^2 : Hệ số xác định mô hình.
18. R : Ma trận tương quan giữa các biến.
19. RSS : Tổng bình phương các sai lệch.
20. SRF : Hàm hồi quy mẫu.
21. PRF : Hàm hồi quy tổng thể.
22. OLS : Phương pháp bình phương bé nhất.
23. $\sum_{i=1}^n X_i = X_1 + X_2 + \dots + X_n$: Tổng n giá trị của X_i .
24. C : Giá trị tới hạn của các phân phối.
25. f : Tỷ lệ mẫu.

GIỚI THIỆU CHUNG VỀ NGHIÊN CỨU KHOA HỌC VÀ PHƯƠNG PHÁP NGHIÊN CỨU KINH TẾ

1.1. Giới thiệu nghiên cứu là gì

1.1.1. Nghiên cứu

Nghiên cứu là sự tìm kiếm kiến thức, hoặc là sự điều tra mang tính hệ thống, với suy nghĩ mở rộng để khám phá, giải thích và phát triển các phương pháp nhằm vào sự tiến bộ kiến thức của nhân loại.

Theo Babbie (1986) : Nghiên cứu là quá trình thu thập dữ liệu và phân tích thông tin một cách hệ thống nhằm tăng cường sự hiểu biết về một hiện tượng.

Theo Kothari (2004) : Nghiên cứu là quá trình thu thập, phân tích dữ liệu một cách có hệ thống nhằm khám phá các vấn đề liên quan.

Theo Kumar (2014) : Nghiên cứu là một trong những cách tìm ra các câu trả lời cho các câu hỏi.

Theo Shuttleworth (2008) : Nghiên cứu là bao gồm mọi cách thức thu thập dữ liệu, thông tin và sự kiện cho sự phát triển kiến thức.

Nghiên cứu là quá trình thu thập và phân tích thông tin một cách có hệ thống nhằm tăng cường sự hiểu biết về một hiện tượng hay một vấn đề nào đó.

1.1.2. Nghiên cứu khoa học

Theo Babbie (2011) : Nghiên cứu khoa học là cách thức : con người tìm hiểu các hiện tượng khoa học một cách có hệ thống và quá trình áp dụng các ý tưởng, nguyên lý để tìm ra các kiến thức mới nhằm giải thích các sự vật hiện tượng.

Theo Armstrong và Sperry (1994) : Nghiên cứu khoa học dựa vào việc ứng dụng các phương pháp khoa học để phát hiện ra những cái mới về bản chất sự vật, về thế giới tự nhiên và xã hội, và để sáng tạo phương pháp và phương tiện kỹ thuật mới cao hơn, giá trị hơn. Hình thức nghiên cứu này cung cấp thông tin và lý thuyết khoa học nhằm giải thích bản chất và tính chất của thế giới.

Nghiên cứu khoa học là hoạt động tìm kiếm, xem xét, điều tra, hoặc thử nghiệm. Dựa trên những số liệu, tài liệu, kiến thức,... đạt được từ các thí nghiệm nghiên cứu khoa học để phát hiện những cái mới về bản chất sự vật, về thế giới tự nhiên và xã hội, và để

sáng tạo phương pháp và phương tiện kỹ thuật mới cao hơn, giá trị hơn. Con người muốn nghiên cứu khoa học phải có kiến thức nhất định về lĩnh vực nghiên cứu và cái chính là phải rèn luyện cách làm việc tự lực, có phương pháp từ lúc ngồi trên ghế nhà trường.

1.1.3. Vai trò của nghiên cứu khoa học

Nghiên cứu là để truyền tải thông tin. Tuy nhiên một bài nghiên cứu hiệu quả phải:

- Làm thay đổi cách nhìn nhận vấn đề của người đọc.
- Thuyết phục người đọc tin vào một điều gì đó.
- Đưa người đọc đến quyết định và hành động.
- Dẫn dắt người đọc theo một quy trình nào đó.

1.1.4. Nghiên cứu kinh tế

Dựa trên khái niệm về nghiên cứu. Nghiên cứu kinh tế là nghiên cứu liên quan đến lĩnh vực kinh tế. Nói cách khác, nghiên cứu kinh tế là quá trình thu thập thông tin, dữ liệu, chứng cứ, vận dụng các công cụ kiến thức và công cụ phân tích xử lý thông tin dữ liệu nhằm đạt được sự hiểu biết về vai trò và hành vi của từng cá nhân, hộ gia đình, doanh nghiệp, ngành, thị trường, quốc gia hoặc toàn bộ nền kinh tế đối với việc đưa ra quyết định kinh tế.

1.1.5. Phương pháp nghiên cứu

Theo Yang (2001), phương pháp nghiên cứu cung cấp các chi tiết của quy trình và phương pháp cụ thể để thực hiện một vấn đề nghiên cứu. Phương pháp nghiên cứu cung cấp các quy trình cụ thể và chi tiết làm thế nào để bắt đầu, thực hiện và hoàn thành nhiệm vụ nghiên cứu và chủ yếu là tập trung vào làm thế nào để thực hiện được nghiên cứu.

Theo nghĩa hẹp, phương pháp nghiên cứu là phương pháp thu thập thông tin. Connaway và Powell (2003) cho rằng có rất nhiều cách để có được thông tin. Các phương pháp nghiên cứu phổ biến nhất là tìm kiếm tài liệu, hội thảo, hội thảo nhóm, phỏng vấn cá nhân, các cuộc điều tra qua điện thoại, các cuộc điều tra qua thư bưu điện và điều tra qua thư điện tử và mạng.

Phương pháp nghiên cứu đối với nhà kinh tế là tìm hiểu bản chất những vấn đề kinh tế đang cần giải quyết của hộ gia đình, doanh nghiệp, cộng đồng hay các nhà hoạch định chính sách ở phạm vi địa phương, quốc gia hay cộng đồng quốc tế nói chung. Các nhà kinh tế đã phát triển lý thuyết về cách thức hoạt động của thị trường, làm thế nào các hoạt động kinh tế được tiến hành trong các quốc gia hoặc trên phạm vi toàn cầu. Do đó, nghiên cứu kinh tế được chia thành hai nhóm chính: Nghiên cứu cơ bản và nghiên cứu ứng dụng.

1.2. Phân biệt loại hình nghiên cứu

Nghiên cứu trong lĩnh vực kinh tế, quản lý có thể được chia làm hai loại: Nghiên cứu cơ bản và nghiên cứu ứng dụng.

1.2.1. Nghiên cứu cơ bản

Nghiên cứu cơ bản có mục tiêu phát hiện và kiểm định quy luật mới. Đây là những nghiên cứu nặng về phát triển lý thuyết hơn là áp dụng vào thực tiễn. Đây là những nghiên cứu có tính hội nhập quốc tế cao, cần được các nhà khoa học quốc tế phản biện. Thông thường những nghiên cứu cơ bản có thể công bố trên các tạp chí quốc tế uy tín.

1.2.2. Nghiên cứu ứng dụng

Giữa nghiên cứu cơ bản và nghiên cứu ứng dụng trong thực tiễn có một khoảng cách rất lớn. Nhiều phát hiện mới, mang tính lý thuyết/hàn lâm, thường khó hiểu và không được thể hiện dưới dạng có thể ứng dụng ngay vào thực tiễn được. Để giải quyết vấn đề này cần có những nghiên cứu nhằm chuyển tải những phát minh khoa học mới vào thực tiễn. Trong lĩnh vực kinh tế, quản lý, quản trị kinh doanh, đó là các nghiên cứu chính sách, các nghiên cứu tư vấn. Các nghiên cứu này có thể dựa trên các quy luật đã nghiên cứu, thu thập dữ liệu để phân tích vấn đề thực tiễn (mối quan hệ giữa các nhân tố), từ đó đề xuất bài học hoặc giải pháp cho nhà hoạt động thực tiễn.

1.2.3. So sánh giữa nghiên cứu cơ bản và nghiên cứu ứng dụng

	Nghiên cứu cơ bản	Nghiên cứu ứng dụng
Mục tiêu	Phát triển lý thuyết trong ngành kinh tế - quản lý	Ứng dụng lý thuyết vào phân tích thực tiễn ở đơn vị, ngành, địa phương cụ thể
Kết quả nghiên cứu	Công trình nghiên cứu mang nặng tính lý thuyết với kết quả chính là luận điểm, mô hình, hoặc học thuyết mới	Công trình nghiên cứu vừa mang tính lý thuyết lại vừa có khả năng ứng dụng cao, trực tiếp vào những khung cảnh nghiên cứu cụ thể
Đặc điểm của các công trình	Coi trọng tính tổng quát hóa và trường tồn của kết quả nghiên cứu theo	Coi trọng tính phù hợp của kết quả nghiên cứu đối với một hoặc một số bối cảnh cụ

ngiên cứu	không gian và thời gian	thể
Người phản biện phù hợp	Các chuyên gia lý thuyết (quốc tế) là những người phù hợp để phản biện luận án hoặc công trình nghiên cứu	Các chuyên gia lý thuyết kết hợp với nhà hoạt động thực tiễn là những người phù hợp để phản biện luận án hoặc công trình nghiên cứu
Nơi công bố - xuất bản công trình	Công bố ở những tạp chí chuyên ngành lý thuyết (quốc tế)	Công bố ở những tạp chí dành cho các nhà hoạt động thực tiễn

Bảng 1.1. So sánh nghiên cứu cơ bản và nghiên cứu ứng dụng.

1.3. Phân biệt phương pháp nghiên cứu định tính, định lượng và hỗn hợp

1.3.1. Giới thiệu về phương pháp nghiên cứu định tính

Nghiên cứu định tính là một dạng nghiên cứu được tạo thành từ các phương pháp kỹ thuật chuyên môn được sử dụng để tìm hiểu sâu về các phản ứng từ trong suy nghĩ và tình cảm của con người

Theo Marshall và Rossman (1998): Nghiên cứu định tính là một phương pháp điều tra được sử dụng trong nhiều lĩnh vực khoa học kỹ thuật khác nhau, không những trong khoa học xã hội mà còn trong nghiên cứu kinh tế. Mục đích của nghiên cứu định tính nhằm tìm kiếm sự hiểu biết sâu sắc về hành vi của con người và lý do chi phối hành vi như vậy. Nghiên cứu định tính điều tra lý do tại sao và làm như thế nào của việc ra quyết định, không chỉ là những gì, mà còn ở đâu, khi nào. Nghiên cứu định tính đưa ra các kết luận tổng quát hơn là các kết luận cụ thể:

Nghiên cứu định tính là loại hình nghiên cứu nhằm mô tả sự vật hiện tượng mà không quan tâm đến sự biến thiên của đối tượng nghiên cứu và không nhằm lượng hóa sự biến thiên này.

Nghiên cứu định tính hướng đến ý nghĩa các khái niệm, định nghĩa, đặc điểm và sự mô tả đối tượng nghiên cứu.

Nghiên cứu định tính thường được áp dụng trong xã hội học, tâm lý học, kinh tế học, kinh tế chính trị, luật,...

Nghiên cứu định tính thường dùng các phương pháp phân tích, so sánh, tổng hợp, mô tả, logic,...

Nghiên cứu định tính thường được áp dụng giai đoạn thu thập dữ liệu và phân tích dữ liệu. Ở giai đoạn thu thập dữ liệu, các kỹ thuật nghiên cứu định tính thường được áp dụng gồm nghiên cứu lý thuyết nền, phỏng vấn nhóm, phỏng vấn chuyên gia, nghiên cứu tình huống, quan sát,...

Ở giai đoạn phân tích dữ liệu, nghiên cứu định tính được sử dụng các kỹ thuật phân tích nội dung với các dữ liệu thu thập, quan sát hành vi cũng như các chứng cứ, sự kiện thu thập được.

1.3.2. Giới thiệu về phương pháp nghiên cứu định lượng

Theo Ehrenberg (1994): Nghiên cứu định lượng thường được sử dụng để kiểm định lý thuyết dựa vào cách tiếp cận suy diễn.

Theo Daniel Muijs, (2004) : Nghiên cứu định lượng là phương pháp giải thích hiện tượng thông qua phân tích thống kê với dữ liệu định lượng thu thập được.

Nghiên cứu định lượng là loại hình nghiên cứu mà ta muốn lượng hóa sự biến thiên của đối tượng nghiên cứu và công cụ thống kê, mô hình hóa được sử dụng cho việc lượng hóa các thông tin của nghiên cứu định lượng.

Các phương pháp định lượng bao gồm các quy trình thu thập dữ liệu, phân tích dữ liệu, giải thích và viết kết quả nghiên cứu.

Các phương pháp này liên quan đến sự xác định mẫu, chiến lược điều tra, thu thập dữ liệu, phân tích dữ liệu, trình bày kết quả nghiên cứu, thảo luận kết quả và viết công trình nghiên cứu.

Nghiên cứu định lượng phù hợp với các nghiên cứu xác định các yếu tố ảnh hưởng một kết quả nào đó. Cách tiếp cận định lượng thực hiện khi cần kiểm định các giả thuyết khác nhau và một lý thuyết nào đó.

1.3.3. So sánh phương pháp nghiên cứu định tính, định lượng

Thứ tự	Nội dung	Định tính	Định lượng
1	Mục tiêu Nghiên cứu	Hiểu sâu sắc, xây dựng lý thuyết	Mô tả hoặc dự báo, xây dựng hoặc kiểm định lý thuyết
2	Thiết kế nghiên cứu	Có thể điều chỉnh trong quá trình thực hiện.	Được quyết định trước khi bắt đầu nghiên cứu.

	cứ	Thường phối hợp nhiều phương pháp	Sử dụng một hay phối hợp nhiều phương pháp.
3	Chọn mẫu, cỡ mẫu	Phi xác suất, có mục đích. Cỡ mẫu nhỏ	Xác suất, cỡ mẫu lớn
4	Phân tích dữ liệu	Phân tích bằng con người và thực hiện liên tục trong quá trình nghiên cứu	Phân tích bằng máy. Các phương pháp toán và thống kê làm chủ đạo. Phân tích có thể diễn ra suốt quá trình nghiên cứu.

Bảng 1.2. So sánh nghiên cứu định tính và nghiên cứu định lượng.

1.3.4. Giới thiệu phương pháp nghiên cứu hỗn hợp

Nghiên cứu phối hợp giữa định tính và định lượng được sử dụng khá phổ biến trong các ngành kinh tế, quản trị, tài chính,...

Trong quá trình nghiên cứu, chúng ta muốn hiểu rõ bản chất sự vật, nghiên cứu cơ sở lý thuyết, xây dựng mô hình hay khung phân tích thì phải dùng phương pháp nghiên cứu định tính với các công cụ tổng hợp, phân tích, so sánh đối chiếu, chuyên gia,...

Đồng thời, trong nghiên cứu chúng ta thường dựa trên một quan sát với cỡ mẫu đủ lớn để có kết quả tin cậy cần thiết. Chúng ta dùng dữ liệu, thông tin của mẫu để ước đoán số liệu, thông tin tổng thể nghiên cứu. Vì vậy phương pháp định lượng là hiển nhiên.

Khi so sánh nghiên cứu định lượng và định tính, Wilson (1982) lập luận rằng, việc sử dụng cân bằng cả hai nên được sử dụng trong nghiên cứu hiện đại. Thực hiện sự kết hợp này là sử dụng thế mạnh của cả hai nghiên cứu định tính và định lượng. Hơn nữa, phương pháp hỗn hợp có thể giải quyết sự phức tạp ngày càng tăng của thực tế xã hội. Hiểu biết có được sự kết hợp của cả hai phương pháp nghiên cứu định tính và định lượng cung cấp một sự hiểu biết tốt hơn và mở rộng của chủ đề nghiên cứu. Ta có bảng kết quả so sánh về quy trình các phương pháp như sau:

Các phương pháp nghiên cứu định tính	Các phương pháp nghiên cứu định lượng	Các phương pháp nghiên cứu hỗn hợp
Các phương pháp mới nổi. Các câu hỏi mở.	Các câu hỏi dựa vào một công cụ xác định trước. Dữ liệu về kết quả hoạt	Cả hai phương pháp mới nổi và xác định trước. Cả câu hỏi có mở và đóng.

Dữ liệu phỏng vấn, dữ liệu quan sát, dữ liệu văn bản, và dữ liệu nghe nhìn.	động, dữ liệu về thái độ, dữ liệu quan sát, và dữ liệu tổng điều tra thống kê. Phân tích thống kê.	Nhiều hình thức thu thập dữ liệu từ mọi khả năng. Phân tích thống kê và văn bản.
---	---	---

Bảng 1.3. So sánh quy trình phương pháp định tính, định lượng và hỗn hợp.

1.4. Quy trình nghiên cứu

1.4.1. Khái niệm quy trình nghiên cứu

Theo Kumar (2005), Quy trình nghiên cứu là một chuỗi các hành động diễn ra theo trình tự và gắn liền với nền tảng kiến thức cũng như các bước tư duy logic. Trong khái niệm này, quy trình nghiên cứu bao gồm một chuỗi các bước tư duy và vận dụng kiến thức về phương pháp nghiên cứu, kiến thức chuyên ngành, khởi đầu từ đặt vấn đề cho đến khi tìm ra câu trả lời. Các bước trong quy trình nghiên cứu phải theo một trình tự nhất định.

1.4.2. Nội dung các bước của quy trình nghiên cứu

1.4.2.1. Xác định và mô tả vấn đề nghiên cứu

a. Xác định vấn đề nghiên cứu

Đây là bước khá quan trọng trong quá trình nghiên cứu vì nếu không xác định đúng đề tài nghiên cứu sẽ khó thực hiện được. Để xác định vấn đề nghiên cứu ta cần làm rõ từng bước sau:

Bước 1. Xác định rõ lĩnh vực nghiên cứu.

Bước 2. Xác định loại vấn đề nghiên cứu.

Bước 3. Xác định sự cần thiết của nghiên cứu.

Bước 4. Đánh giá tính khả thi của nghiên cứu.

Bước 5. Trao đổi với giáo viên hướng dẫn, bạn bè, chuyên gia trong lĩnh vực này.

Bước 6. Thỏa mãn sự ưu thích, đam mê và sở trường của người nghiên cứu.

Ví dụ 1. Xác định vấn đề nghiên cứu. Đề tài: “ Nghiên cứu các yếu tố ảnh hưởng đến thu hút đầu tư trực tiếp nước ngoài vào khu công nghiệp ở TPHCM”.

Để có thể thu hút được các doanh nghiệp đầu tư vào các khu công nghiệp ở TPHCM ta cần tìm hiểu yếu tố môi trường đầu tư nào tác động đến khả năng thu hút đầu tư vào các khu công nghiệp, từ đó tìm giải pháp thu hút đầu tư trực tiếp từ nước ngoài vào khu công nghiệp TPHCM.

b. Mô tả vấn đề nghiên cứu

Vấn đề nghiên cứu cần được mô tả làm rõ ràng các bước sau:

Bước 1. Mục tiêu nghiên cứu

Bước 2. Câu hỏi nghiên cứu

Bước 3. Đối tượng nghiên cứu

Bước 4. Phạm vi nghiên cứu

Bước 5. Ý nghĩa và giới hạn của nghiên cứu.

Ví dụ 2. Xác định mục tiêu nghiên cứu. Đề tài: “ Nghiên cứu các yếu tố ảnh hưởng đến thu hút đầu tư trực tiếp nước ngoài vào khu công nghiệp ở TPHCM”.

Mục tiêu chung: Đánh giá thực trạng các khu công nghiệp về thu hút đầu tư trực tiếp nước ngoài (FDI) và xác định các yếu tố ảnh hưởng thu hút FDI trên cơ sở đó đề xuất các giải pháp nhằm thu hút FDI vào khu công nghiệp.

Mục tiêu cụ thể:

- a. Đánh giá thực trạng các khu công nghiệp về thu hút FDI
- b. Xác định các yếu tố ảnh hưởng thu hút FDI vào khu công nghiệp.
- c. Đề xuất các giải pháp nhằm thu hút FDI vào các khu công nghiệp.

Ví dụ 3. Xác định câu hỏi nghiên cứu. Đề tài: “ Nghiên cứu các yếu tố ảnh hưởng đến thu hút đầu tư trực tiếp nước ngoài vào khu công nghiệp ở TPHCM”.

Câu hỏi nghiên cứu:

- a. Các yếu tố môi trường đầu tư của các KCN là gì?
- b. Những yếu tố nào làm thỏa mãn nhà đầu tư đang đầu tư và sẽ đầu tư tại đây?
- c. Mức độ tác động của các yếu tố môi trường đầu tư đến sự thỏa mãn của nhà đầu tư như thế nào?

Ví dụ 4. Xác định đối tượng và phạm vi nghiên cứu. Đề tài: “ Nghiên cứu các yếu tố ảnh hưởng đến thu hút đầu tư trực tiếp nước ngoài vào khu công nghiệp ở TPHCM”.

Đối tượng nghiên cứu: Lý thuyết thực trạng của môi trường đầu tư, các yếu tố ảnh hưởng đến thu hút đầu tư trực tiếp nước ngoài vào các khu công nghiệp TPHCM.

Phạm vi nghiên cứu:

Nội dung: Đánh giá các nguồn lực cho phát triển - điều kiện tự nhiên, kinh tế, xã hội, môi trường và các KCN; môi trường đầu tư và thực trạng đầu tư vào các KCN; mức độ hài lòng và các yếu tố ảnh hưởng đến độ hài lòng của nhà đầu tư về môi trường đầu tư ở các KCN và đề xuất giải pháp thu hút đầu tư vào các KCN.

Không gian: Các khu công nghiệp trên địa bàn TPHCM.

Thời gian: Dữ liệu thu thập của 5 năm gần đây (2010 -2014) và dữ liệu do tác giả tiến hành thu thập trong năm 2014.

1.4.2.2. Tìm hiểu cơ sở lý thuyết

Cơ sở lý thuyết của nghiên cứu bao hàm các khái niệm, lý thuyết kinh tế học, các công trình nghiên cứu thực nghiệm trong và ngoài nước có liên quan:

a. Các khái niệm

Trong giai đoạn tìm hiểu về cơ sở lý thuyết, cần nêu rõ một số khái niệm trực tiếp liên quan đến vấn đề nghiên cứu. Tất cả các khái niệm được đề cập phải có nguồn gốc, trích dẫn rõ ràng để có giá trị tham khảo và phải có trong mục “tài liệu tham khảo” của đề cương nghiên cứu.

b. Các lý thuyết kinh tế học và các công trình nghiên cứu thực nghiệm

Các lý thuyết kinh tế học là những kết quả nghiên cứu đã được công nhận trong giới học thuật trên phạm vi toàn cầu, trong khi các công trình nghiên cứu thực nghiệm chỉ cho kết quả nghiên cứu riêng lẻ trong từng quốc gia, hay từng vùng, từng địa phương trong một quốc gia. Tất cả các lý thuyết được đề cập phải có nguồn gốc, trích dẫn rõ ràng để có giá trị tham khảo và phải có trong mục “tài liệu tham khảo” của đề cương nghiên cứu.

1.4.2.3. Xây dựng khung phân tích và giả thuyết nghiên cứu

a. Khung phân tích của nghiên cứu

Dựa trên lý thuyết, nghiên cứu thực nghiệm, thực tiễn địa phương, tác giả hình thành khung phân tích cho đề tài nghiên cứu của mình. Khung phân tích này cho thấy tác giả đã am tường các lý thuyết liên quan, kế thừa kết quả từ các nghiên cứu liên quan trước đây và cũng thể hiện được điểm mới trong nghiên cứu của mình.

b. Giả thuyết nghiên cứu

Giả thuyết nghiên cứu là phát biểu về mối liên hệ giữa các biến, nhà nghiên cứu sẽ kiểm định giả thuyết này trong quá trình nghiên cứu. Dựa trên các lý thuyết kinh tế học và kế thừa kết quả từ các công trình nghiên cứu thực nghiệm, tác giả hình thành quan hệ giữa biến độc lập và biến phụ thuộc, tương quan cùng chiều hay ngược chiều.

1.4.2.4. Viết đề cương nghiên cứu

a. Khái niệm

Đề cương nghiên cứu là một báo cáo trình bày toàn bộ các bước nghiên cứu từ vấn đề đặt ra đến tiến độ thực nghiệm nghiên cứu.

b. Nội dung của đề cương nghiên cứu

Hiện nay trên thế giới cũng như ở Việt Nam chưa có một quy định thống nhất nào về nội dung đề cương nghiên cứu. Tùy thuộc vào từng trường đại học, từng quốc gia nước khác nhau có một vài điểm giống nhau như sau:

1. Giới thiệu (mở đầu)
1.1. Xác định vấn đề nghiên cứu
1.2. Tính cấp thiết của đề tài
1.3. Mục tiêu nghiên cứu
1.4. Câu hỏi nghiên cứu
1.5. Phạm vi nghiên cứu
1.6. Ý nghĩa và giới hạn của nghiên cứu
2. Tổng quan tài liệu và cơ sở lý thuyết
2.1. Khái niệm
2.2. Lý thuyết liên quan
2.3. Các nghiên cứu trong và ngoài nước liên quan đề tài nghiên cứu
3. Khung phân tích và giả thuyết nghiên cứu
3.1. Khung phân tích của nghiên cứu
3.2. Các giả thuyết nghiên cứu
4. Phương pháp nghiên cứu (thiết kế nghiên cứu)
4.1. Phương pháp chọn mẫu
4.2. Dữ liệu thu thập
4.3. Công cụ phân tích dữ liệu
5. Kết cấu của đề tài
6. Tiến độ thực hiện
7. Tài liệu tham khảo

Bảng 1.4. Nội dung cơ bản của đề cương nghiên cứu.

Để minh họa cho đề cương nghiên cứu, chúng tôi giới thiệu chi tiết một đề cương nghiên cứu luận văn thạc sĩ của Cô Phùng Vũ Bảo Ngọc, Khoa Du lịch của trường Đại học Tài chính - Marketing (xem chương 8).

1.4.2.5. Thiết kế nghiên cứu

a. Phương pháp chọn mẫu

Phương pháp chọn mẫu bao gồm các nội dung như sau:

+) Khái niệm mẫu

- +) Lý do chọn mẫu
- +) Một số định nghĩa về mẫu
- +) Xác định quy mô mẫu
- +) Phương pháp và hình thức chọn mẫu

b. Phương pháp thu thập dữ liệu

Có hai loại dữ liệu chính để thu thập

+) Dữ liệu thứ cấp: là dữ liệu do người khác thu thập, không phải do người nghiên cứu trực tiếp thu thập.

+) Dữ liệu sơ cấp: là dữ liệu do người nghiên cứu trực tiếp thu thập.

1.4.2.6. Phân tích dữ liệu và kiểm định giả thuyết

a. Phân tích dữ liệu: Sau khi thu thập dữ liệu, ta cần tiến hành phân tích dữ liệu theo các mục tiêu và các giả thuyết nghiên cứu đưa ra.

+) Phân tích mô tả: Sử dụng thống kê mô tả để làm rõ các thuộc tính của đối tượng khảo sát.

+) Kiểm định sự khác biệt giữa hai tham số trung bình: Sử dụng các kiểm định như kiểm định t đối với mẫu độc lập, kiểm định Chi – bình phương, kiểm định phương sai một yếu tố để xác định sự khác biệt của các tham số trung bình có ý nghĩa.

+) Kiểm định chất lượng thang đo: Sử dụng kiểm định Cronbach’s Alpha để xác định chất lượng thang đo xây dựng.

+) Phân tích nhân tố khám phá: Sử dụng kiểm định KMO, Bartlett, và phương sai trích để xác định hệ thống thang đo đại diện.

+) Phân tích hồi quy tuyến tính: Sử dụng các kiểm định các hệ số hồi quy, mức độ phù hợp của mô hình, tự tương quan và phương sai phần dư để xác định các yếu tố và mức độ ảnh hưởng của các yếu tố.

b. Kiểm định giả thuyết: Sau khi có kết quả của việc phân tích dữ liệu, ta cần tiến hành kiểm định các giả thuyết đưa ra ban đầu. Có thể kết quả phân tích cho biết dữ liệu là phù hợp nhưng có thể khác với giả thuyết ban đầu. Trong trường hợp khác với giả thuyết ban đầu, ta cần điều chỉnh lại giả thuyết.

1.4.2.7. Giải thích kết quả và viết báo cáo

- +) Rút ra được những phát hiện nào, kết luận nào từ kết quả?
- +) Kết quả phân tích có phù hợp với lý thuyết và thực tiễn hay không?
- +) Kết quả có gì mới hay không?

+) Có thể đề xuất gì hay không?

1.5. Các cấu phần cơ bản của một nghiên cứu

Cấu phần	Nội dung	Cở sở
Ý tưởng	<p>Nghiên cứu nhằm phát hiện điều gì và vì sao nó có ý nghĩa?</p> <p>+) Phát hiện khoảng trống quan trọng trong nghiên cứu trước.</p> <p>+) Câu hỏi nghiên cứu.</p> <p>+) Hình thành ý tưởng điểm mới và đóng góp của nghiên cứu</p>	<p>+) Hiểu biết của nhà nghiên cứu về lĩnh vực chuyên ngành.</p> <p>+) Tổng quan tài liệu.</p> <p>+) Trải nghiệm với bối cảnh cụ thể.</p>
Thiết kế	<p>Cần những bằng chứng gì để trả lời câu hỏi nghiên cứu và thu thập/phân tích chúng bằng cách nào?</p> <p>+) Cách tiếp cận và thiết kế tổng thể.</p> <p>+) Quy trình, phương pháp, công cụ cụ thể.</p> <p>+) Kế hoạch tiến hành và yêu cầu về nguồn lực, dữ liệu.</p> <p>+) Mức độ đáp ứng các yêu cầu của một nghiên cứu khoa học.</p> <p>+) Rủi ro và hạn chế của nghiên cứu.</p>	<p>+) Nắm rõ mục tiêu và khung nghiên cứu.</p> <p>+) Xác định các giả thuyết “cạnh tranh” để có kế hoạch thu thập bằng chứng, kiểm định và bảo vệ giả thuyết chính.</p> <p>+) Nắm rõ các phương pháp và công cụ có thể sử dụng.</p>
Phương pháp và công cụ	<p>Thực hiện việc thu thập và phân tích dữ liệu bằng cách nào?</p> <p>+) Các phương pháp thu thập dữ liệu.</p> <p>+) Các phương pháp và công cụ phân tích dữ liệu.</p>	<p>+) Nắm rõ câu hỏi và khung nghiên cứu.</p> <p>+) Nắm rõ quy trình và kỹ thuật của từng phương pháp.</p>

Bảng 1.5. Các cấu phần cơ bản của một nghiên cứu.

1.6. Đạo đức trong nghiên cứu khoa học

1.6.1. Thế nào là đạo đức nghiên cứu khoa học?

Một định nghĩa về đạo đức nghiên cứu trong các nghiên cứu xã hội được John Barnes đưa ra năm 1979. Ông cho rằng “ Đạo đức trong nghiên cứu khoa học là những điều đặt ra khi chúng ta quyết định giữa việc cần thực hiện một hành động này với những điều khác không chỉ xét trên tính thích hợp hay hiệu quả mà còn bằng việc tham khảo các tiêu chuẩn đúng hay sai về mặt đạo đức” (J.A. Barnes, 1979; trang 16).

Barnes đã đưa ra sự phân định và đặt cơ sở cho quan điểm các quyết định đạo đức cần dựa trên các nguyên tắc chứ không dựa vào sự thích hợp. Đây là một điển ghi nhớ quan trọng. Các quyết định đạo đức không chỉ được xác định trên cơ sở sự thuận tiện cho nhà nghiên cứu hay đề tài nghiên cứu mà nhà nghiên cứu đó đang tham gia. Người ta cần quan tâm đến điều gì là đúng, không chỉ đối với đề tài nghiên cứu, nhà tài trợ nghiên cứu hay nhà nghiên cứu mà còn đối với những người tham gia trong nghiên cứu. Các quyết định đạo đức sẽ phải dựa trên các giá trị của nhà nghiên cứu và cộng đồng nghiên cứu, nhà tài trợ, những người tham gia vào nhóm nghiên cứu và sẽ dựa trên những thương lượng giữa nhóm nghiên cứu nói trên và kể cả những người đóng vai trò kiểm soát, đánh giá các thông tin mà nhà nghiên cứu thu được. Việc thực hiện các giám sát trong quá trình nghiên cứu sẽ ảnh hưởng đến chính các quyết định đạo đức của các nhà nghiên cứu khoa học.

1.6.2. Các nguyên tắc đạo đức trong nghiên cứu khoa học

1.6.2.1. Nguyên tắc tôn trọng con người

Nguyên tắc tôn trọng con người kết hợp ít nhất hai vấn đề đạo đức nền tảng là:

a) Tôn trọng quyền tự quyết: Các cá nhân có quyền quyết định tham gia hoặc không tham gia vào một nhóm nghiên cứu khoa học. Họ phải được đối xử một cách tôn trọng. Nhà nghiên cứu sẽ vi phạm đạo đức nghiên cứu nếu ép buộc bất kỳ người nào đó tham gia vào nhóm nghiên cứu của mình dù là người đó có khả năng hoặc không có khả năng tự cân nhắc để đưa ra các quyết định cá nhân.

b) Bảo vệ những người bị hạn chế hoặc giảm sút quyền tự quyết: Nhà nghiên cứu khi tiến hành một nghiên cứu khoa học cần đảm bảo an toàn, chống lại mọi xâm hại, lạm dụng đối với người bị phụ thuộc vào hoàn cảnh hoặc bị tổn thương, ví dụ như: phụ nữ, trẻ em, người nghèo...

1.6.2.2. Nguyên tắc hướng thiện

Nguyên tắc này được đề ra nhằm hướng tới mục tiêu đảm bảo nghĩa vụ của mọi nghiên cứu khoa học là phải tối đa hóa lợi ích và tối thiểu hóa tác hại. Đây là nguyên tắc nâng dần đến các chuẩn mực đòi hỏi rằng rủi ro nghiên cứu phải ở mức hợp lý so với lợi ích mong đợi; thiết kế nghiên cứu phải hợp lý và người thực hiện nghiên cứu phải có đủ năng lực chuyên môn để bảo vệ lợi ích của đối tượng nghiên cứu. Nguyên tắc này còn ngăn cấm mọi sự gây hại có chủ tâm đến con người, xã hội nói chung khi thực hiện các nghiên cứu khoa học.

1.6.2.3. Nguyên tắc công bằng

Nguyên tắc công bằng đề cập đến nghĩa vụ đạo đức là phải đối xử với mọi người nghiên cứu một cách đúng đắn và phù hợp về mặt đạo đức, phải đảm bảo mỗi cá nhân tham gia vào nghiên cứu phải nhận được tất cả những gì mà họ có quyền được hưởng. Trong đạo đức nghiên cứu liên quan đến con người, nguyên tắc là chủ yếu đề cập tới sự phân chia công bằng, trong đó công bằng cả thiệt thòi lẫn lợi ích của việc tham gia nghiên cứu cho mọi đối tượng. Sự khác nhau về phân chia thiệt thòi và lợi ích chỉ được chấp thuận khi sự khác biệt về đạo đức giữa người này với người kia; một trong những khác biệt đó chính là “tính dễ bị tổn thương”. “Tính dễ bị tổn thương” đề cập đến một người không có đủ khả năng để bảo vệ lợi ích của mình hay thiếu năng lực để đưa ra quyết định đồng ý, ví dụ người nghèo, thành viên cấp dưới, thành viên phụ thuộc của nhóm người có cấu trúc thứ bậc.

1.7. Câu hỏi thảo luận

Câu hỏi 1. Nghiên cứu khoa học là gì? Nêu các chuẩn mực quan trọng nhất của nghiên cứu khoa học?

Câu hỏi 2. Nghiên cứu kinh tế là gì? Phân biệt nghiên cứu kinh tế và nghiên cứu khoa học?

Câu hỏi 3. Nêu các cấu phần quan trọng nhất của một đề tài nghiên cứu khoa học. Các cấu phần đó tương ứng với những hoạt động nào trong nghiên cứu?

Câu hỏi 4. Tại sao phải quan tâm đến đạo đức trong nghiên cứu khoa học?

Câu hỏi 5. Phân biệt giả thuyết và giả thiết trong nghiên cứu khoa học.

Câu hỏi 6. Đề cương nghiên cứu khoa học là gì? Có nhất thiết phải xây dựng đề cương nghiên cứu hay không?

Thuật ngữ chính chương 1

Tiếng Anh	Tiếng Việt
Applied research	Nghiên cứu ứng dụng
Brief notes	Bảng ghi chú ngắn gọn
Cross – sectional data	Dữ liệu chéo
Deductive research approach	Cách tiếp cận suy diễn
Hypothesis	Giả thuyết
Introduction	Giới thiệu
Inductive research approach	Cách tiếp cận quy nạp
Model	Mô hình
Mixed – methods approach	Phương pháp hỗn hợp
Rationale for the research	Tính cấp thiết
Reductionism	Quy giản luận
References	Tài liệu tham khảo
Relevant concepts	Khái niệm
Relevant theories	Lý thuyết liên quan
Research	Nghiên cứu
Research methods	Phương pháp nghiên cứu
Primary data	Dữ liệu sơ cấp
Sample size	Quy mô mẫu (cỡ mẫu)
Secondary data	Dữ liệu thứ cấp
Scientific	Khoa học
Scientific research	Nghiên cứu khoa học
Scientific method	Phương pháp khoa học
Statement of the problem	Xác định vấn đề nghiên cứu
Scope of study	Phạm vi nghiên cứu
Outline of draft chapters	Kết cấu luận văn
Quantitative research	Nghiên cứu định lượng
Qualitative research	Nghiên cứu định tính
Tentative hypothesis	Giả thuyết dự kiến

TỔNG QUAN TÌNH HÌNH NGHIÊN CỨU VÀ CÂU HỎI NGHIÊN CỨU

2.1. Giới thiệu về tổng quan nghiên cứu

2.1.1. Định nghĩa tổng quan nghiên cứu

Tổng quan nghiên cứu là một bản miêu tả chi tiết để chỉ ra lý thuyết nào sẽ được nói đến và sử dụng trong đề tài nghiên cứu của mình.

Mục đích của tổng quan nghiên cứu là tóm lược các kiến thức và sự hiểu biết của cộng đồng khoa học trong và ngoài nước đã công bố liên quan đến vấn đề nghiên cứu của mình.

2.1.2. Vai trò của tổng quan nghiên cứu

- Cải thiện hiểu biết của người nghiên cứu về vấn đề nghiên cứu.
- Chọn lọc những lý thuyết và các nghiên cứu thực nghiệm liên quan hữu ích để áp dụng cho nghiên cứu của mình.
- Cung cấp nền tảng lý thuyết cho việc nghiên cứu.
- Định lượng cho nghiên cứu, giúp người nghiên cứu xác định có nên theo đuổi nghiên cứu này hay không.
- Kết quả cụ thể của tổng quan tài liệu giúp người nghiên cứu có đủ thông tin cần thiết để xây dựng khung khái niệm, khung phân tích cho các vấn đề nghiên cứu và là sơ đồ liên kết các khía cạnh nghiên cứu như mục tiêu, phương pháp nghiên cứu, điểm mới,...

Lưu ý khi viết tổng quan nghiên cứu

Viết tổng quan tình hình nghiên cứu không phải liệt kê hay miêu tả các nghiên cứu trước đây.

Phải là một bảng tổng hợp khoa học theo vấn đề nghiên cứu và đánh giá có mục đích.

Chất lượng tổng quan nghiên cứu phụ thuộc

- Khả năng tìm kiếm thông tin, dữ liệu.
- Khả năng tổng hợp và đánh giá về vấn đề nghiên cứu.

2.1.3. Tổng quan nghiên cứu tốt

- i) Được viết theo một trình tự hợp lý

- Khái niệm, định nghĩa.
- Mô hình lý thuyết.
- Các mô hình nghiên cứu thực nghiệm.
- Kết quả đạt được của các nghiên cứu.
- Các bài học kinh nghiệm tự rút ra.

ii) Chỉ ra được các thông tin, dữ liệu quan trọng cần phải thu thập để giải quyết vấn đề nghiên cứu.

iii) Chỉ ra được phương thức thu thập dữ liệu, phương thức xử lý và phân tích dữ liệu.

iv) Có đủ thông tin nền tảng giúp phát họa được phiếu điều tra cho nghiên cứu.

v) Tìm ra khoảng trống nghiên cứu và hướng đi mới

2.2. Nội dung và yêu cầu phần tổng quan

Tổng quan các công trình nghiên cứu trước có thể làm nền móng và định hướng tốt cho các nghiên cứu mới, phần tổng quan thường có những nội dung sau:

2.2.1. Các trường phái lý thuyết là cơ sở lý luận cho nghiên cứu

Phần tổng quan cần nêu các nghiên cứu trước đã áp dụng những trường phái lý thuyết nào khi nghiên cứu chủ đề này. Các tác giả cần tóm tắt luận điểm chính của các trường phái và một số công trình tiêu biểu đã áp dụng từng trường phái. Phần tổng quan về các trường phái lý thuyết có thể tóm tắt dưới dạng sau:

- Cách tiếp cận hiệu quả
- Cách tiếp cận dựa vào năng lực
- Cách tiếp cận thể chế

2.2.2. Bối cảnh nghiên cứu và các nhân tố chính

Các nghiên cứu trước đây đã thực hiện trong bối cảnh nào? Bối cảnh có thể là vùng, ngành, quốc gia, nhóm đối tượng nghiên cứu: Bối cảnh là một yếu tố quan trọng khi viết tổng quan vì bối cảnh khác nhau có thể đưa ra các kết quả rất khác nhau.

Tương tự với từng bối cảnh, phần tổng quan cũng cần chỉ rõ những nhân tố mục tiêu và nhân tố tác động nào đã được nghiên cứu. Những nhân tố nào được nghiên cứu nhiều nhất? Những nhân tố nào ít được chú ý?

Tóm lại, mục này cần thể hiện rõ bối cảnh và những nhân tố (mô hình) đã được các công trình trước nghiên cứu đề cập đến. Đó có thể là nhân tố mục tiêu, nhân tố tác động, nhân tố kết quả, nhân tố điều tiết hay nhân tố trung gian. Định nghĩa chi tiết về các

loại nhân tố này sẽ được trình bày trong phần khung lý thuyết và mô hình nghiên cứu.

2.2.3. Các phương pháp nghiên cứu chính

Các nghiên cứu trước đây đã sử dụng những phương pháp nghiên cứu nào? Nghiên cứu hiện tại cần điểm lại các phương pháp nghiên cứu tương ứng với bối cảnh và mô hình mà các nghiên cứu trước áp dụng. Điều này sẽ rất hữu ích cho phần bình luận về hạn chế của nghiên cứu hiện tại cũng như thiết kế của nghiên cứu của nó.

2.2.4. Các kết quả nghiên cứu chính

Các kết quả nghiên cứu chính thể hiện chủ yếu bằng mối quan hệ giữa các nhân tố. Khi thực hiện tổng quan về kết quả nghiên cứu đã tiến hành trước đây cần chú ý nhóm chúng theo các nhóm sau:

- Các kết quả có nhất quán cao nhất giữa các nghiên cứu.
- Các kết quả còn nhiều mâu thuẫn giữa các nghiên cứu.
- Sự nhất quán hay mâu thuẫn của kết quả có liên quan tới bối cảnh hay phương pháp nghiên cứu khác nhau hay không?

2.2.5. Hạn chế của những nghiên cứu trước và khoảng trống tri thức

Trong phần này đòi hỏi tác giả phải đánh giá được những đóng góp cũng như những hạn chế của các nghiên cứu trước. Nếu làm tốt các nội dung trên thì phần này sẽ dễ dàng hơn.

Trên cơ sở hạn chế của các nghiên cứu trước, các tác giả có thể đề xuất hướng nghiên cứu mới. Các hướng nghiên cứu này có thể cần nhiều hơn một đề tài để thực hiện. Các hướng nghiên cứu mới có thể đề xuất dưới dạng sau:

- Chủ đề nghiên cứu mới.
- Câu hỏi nghiên cứu mới.
- Bối cảnh nghiên cứu mới.
- Mô hình nghiên cứu mới.
- Phương pháp nghiên cứu mới.

2.3. Một số kỹ năng tiến hành tổng quan

Các bước thực hiện tổng quan tài liệu

Bước 1. Thu thập tài liệu lý thuyết, các đề tài và bài báo liên quan đến vấn đề nghiên cứu

- Thu thập từ các nguồn có thể.
- Đánh giá các nguồn.

- Đọc các nguồn quan trọng, có chất lượng.

Bước 2. Quản lý tài liệu

- Phát triển một cách thức ghi nhận tài liệu: tên tác giả, năm, tên bài báo, sách,...
- Lập danh sách các tài liệu liên quan.
- Ghi chú, đánh dấu lại các nội dung quan trọng khi đọc

Bước 3. Đọc các lý thuyết, bài báo khoa học về chủ đề

- Đọc, phát hiện, phân tích và tổng hợp các tranh luận khoa học.
- Phân tích các tranh luận khoa học khi đọc và đánh giá các chỉ trích một cách cẩn thận và có suy nghĩ.
- Viết lại các chỉ trích đó.

Bước 4. Tổng quan

- Viết tổng quan như một văn bản đánh giá, phê bình chứ không đơn giản là liệt kê hay tóm lược.
- Nên tổng quan các bài báo đăng các tạp chí có uy tín.
- Tổng quan các vấn đề liên quan có tính đánh giá, phê phán, suy nghĩ, so sánh.
- Có thể tóm lược các thông tin.
- Tìm ra một khoảng trống nghiên cứu và hướng đi mới của đề tài.

2.4. Giới thiệu câu hỏi nghiên cứu

2.4.1. Khái niệm

Nghiên cứu một vấn đề là tìm ra câu trả lời cho vấn đề đó. Đặt câu hỏi nghiên cứu là cách tốt nhất để xác định vấn đề nghiên cứu.

Ngược lại, khi ta đã xác định được vấn đề nghiên cứu thì ta đặt câu hỏi để trả lời vấn đề nghiên cứu đó.

Bản chất câu hỏi nghiên cứu liên quan đến các hành động: khám phá, mô tả, kiểm định, so sánh, đánh giá tác động, đánh giá quan hệ nhân quả.

2.4.2. Các loại câu hỏi nghiên cứu

- Câu hỏi nhằm mô tả sự vật, hiện tượng nghiên cứu.
- Câu hỏi nhằm so sánh các sự vật, hiện tượng nghiên cứu.
- Câu hỏi nhằm tìm hiểu mối quan hệ nhân quả giữa các đặc tính (biến) của sự vật hiện tượng.
- Câu hỏi về các giải pháp, hàm ý chính sách khả thi.

2.4.3. Làm thế nào để xác lập câu hỏi nghiên cứu tốt

Câu hỏi nghiên cứu được rút ra trực tiếp từ vấn đề nghiên cứu. Có thể có một câu hỏi duy nhất hay một vài câu hỏi cho vấn đề nghiên cứu.

Theo Nguyễn Văn Tuấn (2011), câu hỏi nghiên cứu tốt phải đáp ứng ít nhất 3 trong 5 tiêu chuẩn FINER.

- F là viết tắt của feasibility (khả thi): Một câu hỏi nghiên cứu tốt phải khả thi, có nghĩa là phải có khả năng trả lời được câu hỏi đó.

- I là viết tắt của interesting (thú vị): Một câu hỏi nghiên cứu tốt phải thú vị đối với nhà khoa học, xứng đáng để theo đuổi.

- N là viết tắt của novelty (có cái mới): Làm nghiên cứu là một việc làm tạo ra thông tin mới, phương pháp mới, ý tưởng mới hay phát hiện mới. Một nghiên cứu chỉ lặp lại những gì người khác đã làm thì không có cái gì mới, giá trị nghiên cứu thấp.

- E là viết tắt của ethics (đạo đức): Một nghiên cứu kinh tế phải tôn trọng quyền con người, doanh nghiệp, luật pháp quốc gia, không làm tổn thương người, doanh nghiệp, lợi ích quốc gia và phải bảo mật tuyệt đối (không được tiết lộ thông tin cá nhân ra ngoài).

- R là viết tắt của relevant (liên đới): Thật ra, chữ “liên đới” ở đây có nghĩa là có ảnh hưởng. Một câu hỏi nghiên cứu mà nếu tìm được câu trả lời và có thể làm thay đổi một chính sách là một câu hỏi quan trọng.

Ví dụ 1. Xác định câu hỏi nghiên cứu. Đề tài: “Đổi mới kiểm tra đánh giá trong dạy học hiện nay”.

Câu hỏi nghiên cứu:

Câu hỏi 1. Các hình thức đánh giá đang được áp dụng là gì?

Câu hỏi 2. Có sự khác biệt ở giảng viên nam hay giảng viên nữ không?

Câu hỏi 3. Có sự tương quan giữa mức độ ứng dụng và sự hỗ trợ của nhà trường không?

Ví dụ 2. Xác định câu hỏi nghiên cứu. Đề tài: “Tăng trưởng kinh tế và vốn đầu tư trực tiếp nước ngoài ở TPHCM”.

Câu hỏi nghiên cứu:

Câu hỏi 1. Yếu tố FDI có ảnh hưởng đến tăng trưởng không?

Câu hỏi 2. Những yếu tố nào ảnh hưởng đến thu hút vốn FDI?

Câu hỏi 3. Mức độ tác động của các yếu tố ảnh hưởng đến thu hút vốn FDI là như thế nào?

2.5. Tiêu chuẩn và cách xây dựng câu hỏi nghiên cứu

2.5.1. Câu hỏi hướng tới vấn đề mang tính quy luật

Câu hỏi hướng tới mối quan hệ (bản chất, lặp đi lặp lại) giữa các nhân tố. Những hiểu biết về mối quan hệ giữa các nhân tố thường tồn tại theo thời gian. Câu hỏi dạng này vì vậy khác với câu hỏi mang tính mô tả hoặc câu hỏi hướng vào giải pháp.

- Nếu câu hỏi nghiên cứu mang tính mô tả, dạng như “thực trạng của chất lượng nguồn nhân lực” thì câu trả lời sẽ chỉ có ý nghĩa vào đúng thời điểm nghiên cứu này. Ngay sau khi công bố kết quả “thực trạng” đã thay đổi.

- Nếu câu hỏi hướng vào giải pháp thì cần nhớ rằng không có giải pháp vạn năng cho mọi tổ chức, ngành, địa phương. Như vậy giải pháp đề xuất, nếu may mắn là đúng, sẽ chỉ có ý nghĩa cho “đơn vị” được nghiên cứu mà không có ý nghĩa rộng rãi.

2.5.2. Câu hỏi có cơ sở thực tiễn/hoặc lý thuyết

Câu hỏi nghiên cứu không thể được đề xuất một cách tùy tiện theo cảm tính và ý thích của nhà nghiên cứu. Về cơ bản, câu hỏi phải có cơ sở thực tiễn và cơ sở khoa học.

- Cơ sở thực tiễn thể hiện ở chỗ câu hỏi nghiên cứu gắn với vấn đề thực tiễn quan tâm.

- Cơ sở khoa học thể hiện ở việc câu hỏi nghiên cứu hướng vào khoảng trống tri thức mà các nhà nghiên cứu để lại.

2.5.3. Các nhân tố, yếu tố trong câu hỏi có phạm vi, ý nghĩa rõ ràng

Sự rõ ràng của câu hỏi phụ thuộc vào sự rõ ràng trong ý nghĩa và phạm vi của nhân tố đề cập tới. Nếu các nhân tố đề cập đã được định nghĩa, đo lường hoặc có phạm vi rõ ràng trong các nghiên cứu trước thì sẽ dễ dàng định hướng nghiên cứu. Ngược lại, nếu đây là những nhân tố trừu tượng, nhân tố có phạm vi rộng rãi hoặc chứa đựng nhiều cách hiểu khác nhau thì câu hỏi nghiên cứu cũng sẽ không rõ ràng.

Ví dụ câu hỏi nghiên cứu: “Hội nhập quốc tế ảnh hưởng như thế nào tới tái cấu trúc doanh nghiệp ở ngành A” là một câu hỏi có nhân tố không rõ ràng. Thứ nhất “hội nhập quốc tế” là một thuật ngữ lớn, không nói rõ hội nhập của ai (nền kinh tế, ngành hay địa phương) và về những gì. Thứ hai, “tái cấu trúc doanh nghiệp” là một thuật ngữ có nhiều nghĩa. “Tái cấu trúc doanh nghiệp” có thể nói tới việc thay đổi cơ cấu tổ chức, cơ cấu sản phẩm, thị trường, thậm chí cả việc sắp xếp lại lao động,..Câu hỏi liên quan tới các nhân tố trừu tượng và không rõ nghĩa như vậy sẽ không thể thực hiện tốt vai trò định hướng và xác lập giá trị khoa học của đề tài.

2.5.4. Câu hỏi có khả năng trả lời được

Câu hỏi nghiên cứu phải có tính khả thi trong việc tìm bằng chứng để trả lời. Nếu câu hỏi quan trọng, rất thú vị nhưng không có khả thi thì nên loại bỏ khỏi đề tài nghiên cứu. Ví dụ mặc dù việc nghiên cứu tác động của một số đặc điểm trong chương trình đào tạo đại học về quản trị kinh doanh tới sự thành công của các doanh nhân là một chủ đề thú vị, song nếu nghiên cứu sự thành công thì có thể phải cần tới 10 năm và đó là khoảng thời gian quá dài để một nghiên cứu sinh có thể thực hiện được. Đây là câu hỏi không khả thi cho một nghiên cứu sinh.

2.6. Câu hỏi thảo luận

Câu hỏi 1. Vai trò của tổng quan nghiên cứu là gì

Câu hỏi 2. Các yêu cầu đặt ra với tổng quan nghiên cứu là gì?

Câu hỏi 3. Hãy nêu các công việc bạn cần làm khi viết tổng quan nghiên cứu ?

Câu hỏi 4. Thế nào là một tổng quan nghiên cứu tốt?

Câu hỏi 5. Trình bày vai trò của câu hỏi nghiên cứu?

Câu hỏi 6. Các tiêu chuẩn của một câu hỏi nghiên cứu là gì?

Câu hỏi 7. Hãy đặt một số câu hỏi nghiên cứu cho đề tài sau: “Phân tích các yếu tố tác động với kết quả học tập của sinh viên”.

Câu hỏi 8. Hãy đọc một công trình nghiên cứu khoa học bất kỳ. Xác định câu hỏi nghiên cứu của công trình này. Câu hỏi nghiên cứu trong công trình này có đáp ứng tốt các tiêu chuẩn mà chương này đề cập hay không?

Thuật ngữ chính chương 2

Tiếng Anh	Tiếng Việt
Advocacy	Quan điểm ủng hộ
Correlational research	Nghiên cứu tương quan
Descriptive research	Nghiên cứu mô tả
Discussion guideline	Dàn bài thảo luận
Dialectical	Biện chứng
Ethics	Đạo đức
Exploratory research	Nghiên cứu khám phá
Empirical research	Nghiên cứu thực nghiệm
Explanatory research	Nghiên cứu giải thích
Feasibility	Khả thi
Hypothesis	Giả thuyết
Interesting	Thú vị
Knowledge claims	Các nhận định tri thức
Law	Quy luật
Leading question	Câu hỏi tắt định, câu hỏi dẫn dắt
Literature review	Tổng quan lý thuyết
Novelty	Có cái mới
Research objectives	Mục tiêu nghiên cứu
Research question	Câu hỏi nghiên cứu
Research hypothesis	Giả thuyết nghiên cứu
Relevant	Liên đới
Recursive	Tính đệ quy
Overall objectives	Mục tiêu cụ thể
Open-ended	Câu hỏi mở
Pre-test	Tiền kiểm định
Strategies of inquiries	Các chiến lược tìm hiểu
Theory	Lý thuyết
Theoretical research	Nghiên cứu lý thuyết
Way in research	Phương thức nghiên cứu

PHÁT TRIỂN KHUNG LÝ THUYẾT, KHUNG KHÁI NIỆM VÀ KHUNG PHÂN TÍCH

3.1. Khung lý thuyết

3.1.1. Giới thiệu về khung lý thuyết

3.1.1.1. Lý thuyết

Lý thuyết là một hệ thống khái niệm về các nhân tố và mối quan hệ giữa chúng, thể hiện cách nhìn nhận về quy luật thế giới. Linh hồn của lý thuyết là các luận điểm về mối quan hệ bản chất, lặp lại giữa các nhân tố và biến số.

Ví dụ 1. Lý thuyết về cung cầu là luận điểm về mối quan hệ giữa khối lượng với giá cả.

3.1.1.2. Khung lý thuyết

Khung lý thuyết là sự thể hiện có logic các nhân tố, biến số và mối quan hệ liên quan trong công trình nghiên cứu. Khung lý thuyết xác định rõ điều cần đo lường, mô tả, khám phá, hoặc kiểm định.

Khung lý thuyết là sự cụ thể hóa của lý thuyết cơ sở thành phần nhân tố, biến số và mối quan hệ cần phát hiện, kiểm định. Mỗi khung lý thuyết thường là sự áp dụng của lý thuyết hoặc sự kết hợp của một vài lý thuyết cơ sở. Vì vậy không có khung lý thuyết đúng hoặc sai. Các tác giả cần luận giải liệu có khung lý thuyết phù hợp với chủ đề và khung cảnh nghiên cứu này hay không mà thôi.

3.1.2. Các cấu phần cơ bản của khung lý thuyết

3.1.2.1. Nhân tố mục tiêu (biến phụ thuộc)

Nhân tố mục tiêu chính là nhân tố trọng tâm của đề tài nghiên cứu.

Ví dụ 2. Một đề tài nghiên cứu về thu hút vốn đầu tư nước ngoài (FDI) thì nhân tố trọng tâm có thể là lượng vốn, số dự án, loại dự án FDI được thu hút.

Trong nghiên cứu định tính, nhân tố trọng tâm thường được nghiên cứu, mô tả và phân tích dưới dạng:

- Các hình thái khác nhau của nhân tố
- Các cấu phần khác nhau của nhân tố
- Sự thay đổi của nhân tố qua thời gian

Trong nghiên cứu định lượng, nhân tố trọng tâm thường được thể hiện là biến

phụ thuộc (đôi khi là biến trung gian) trong mô hình.

3.1.2.2. Nhân tố tác động (biến độc lập) và các nhân tố khác

Các nhân tố có quan hệ tương quan trực tiếp với nhân tố mục tiêu gọi là nhân tố tác động. Trong mô hình kinh tế lượng, nhân tố tác động thường được gọi là biến độc lập. Ngoài ra, một khung lý thuyết (mô hình) còn có thể có các nhân tố khác, như nhân tố điều kiện, nhân tố trung gian, v.v...

3.1.2.3. Mối quan hệ giữa các nhân tố

- Mối quan hệ tương quan: Đây là mối quan hệ giữa cặp hai nhân tố. Mối quan hệ này có thể là đồng biến hoặc là nghịch biến.

- Mối quan hệ nhân quả: Đây là trường hợp đặc biệt trong quan hệ tương quan. Sự thay đổi của A tác động hoặc gây nên sự thay đổi của B.

- Mối quan hệ điều kiện: Đây là mối quan hệ “tay ba”, trong đó quan hệ giữa hai nhân tố phụ thuộc vào nhân tố thứ ba. Sự thay đổi của A chỉ dẫn tới sự thay đổi của B nếu có C.

- Mối quan hệ trung gian: Đây là mối quan hệ “tay ba”, nhưng nhân tố thứ ba lại là trung gian cho hai nhân tố ban đầu.

3.1.3. Các bước xây dựng khung lý thuyết

Bước 1. Lựa chọn cơ sở lý thuyết (trường phái) cơ bản cho nghiên cứu

Một vấn đề nghiên cứu từ các góc nhìn khác nhau. Mỗi trường phái lý thuyết là một góc nhìn và nhà nghiên cứu thường phải lựa chọn cơ sở lý thuyết phù hợp cho nghiên cứu của mình. Các tác giả phải hiểu được các trường phái lý thuyết để có thể áp dụng giải thích cho vấn đề nghiên cứu của mình quan tâm.

Bước 2. Trọng tâm hóa câu hỏi nghiên cứu dựa trên cơ sở lý thuyết

Các nhà nghiên cứu có thể lựa chọn một số câu hỏi trọng tâm phù hợp với trường phái lý thuyết chính. Đây là quá trình tương tác hai chiều: câu hỏi nghiên cứu ban đầu định hướng lựa chọn trường phái lý thuyết. Ngược lại, việc lựa chọn trường phái lý thuyết lại giúp cụ thể và trọng tâm hóa câu hỏi nghiên cứu.

Bước 3. Định nghĩa rõ các nhân tố

Để xây dựng được khung lý thuyết, yêu cầu đầu tiên là phải định nghĩa rõ nhân tố trọng tâm. Các nhân tố trọng tâm có các đặc điểm sau:

- Nhân tố có nội dung, phạm vi rõ ràng, cụ thể
- Nhân tố có sự khác biệt giữa các đơn vị.

- Sự khác biệt giữa các đơn vị đối với từng nhân tố là có thể đo lường hoặc kiểm soát được.

Bước 4. Xác định mối quan hệ giả thuyết của các nhân tố

Dựa trên cơ sở lý thuyết các nhà nghiên cứu có thể đặt giả thuyết về mối quan hệ giữa các nhân tố. Đặc biệt là nhân tố tác động/điều tiết đến nhân tố mục tiêu.

3.2. Khung khái niệm

Khung khái niệm được sinh ra trực tiếp từ khung lý thuyết và chỉ tập trung vào một phần của khung lý thuyết mà phần này là nền tảng của nghiên cứu.

Khung khái niệm là cốt lõi của vấn đề nghiên cứu

Khung khái niệm là một dạng lý thuyết liên quan do người nghiên cứu xây dựng và có khả năng kết nối tất cả các khía cạnh nghiên cứu như: xác định vấn đề, mục tiêu, tổng quan, phương pháp thu thập và phân tích dữ liệu.

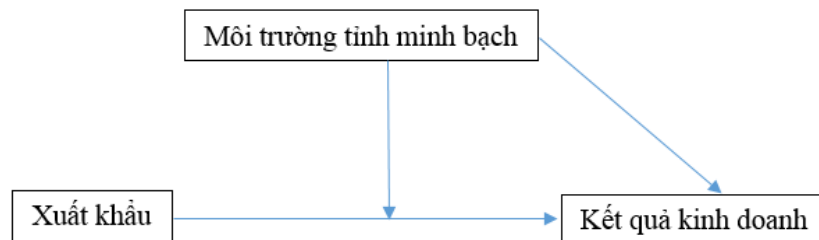
3.3. Khung phân tích

Khung phân tích là hình thức sơ đồ hóa tất cả các quan hệ tương quan, nhân quả giữa các biến số, các chỉ tiêu theo bản chất và trình tự của chúng. Từ đó, ta có thể mô tả trực quan cách thức mà ta phải phân tích vấn đề nghiên cứu.

Khung phân tích giúp ta hình dung được bản chất của dữ liệu, nguồn dữ liệu, tiến trình thu thập, phương thức xử lý để trả lời các câu hỏi nghiên cứu.

Ví dụ trong mô hình nghiên cứu về môi trường thể chế cấp tỉnh, chiến lược xuất khẩu và hiệu quả hoạt động của doanh nghiệp, nhóm tác giả Nguyễn Văn Thắng, Lê Thị Bích Ngọc và Scott Bryant (2013) sử dụng mô hình như hình bên dưới. Trong mô hình này có 3 nhân tố được nghiên cứu:

- Biến độc lập: Chiến lược xuất khẩu của doanh nghiệp.
- Biến phụ thuộc: Hiệu quả kinh doanh của doanh nghiệp.
- Biến điều tiết: Môi trường thể chế cấp tỉnh.



Mô hình nghiên cứu trong công trình của Nguyễn và công sự

3.4. Câu hỏi thảo luận

Câu hỏi 1. Lý thuyết là gì? Khung lý thuyết là gì? Khung lý thuyết khác gì với quy trình nghiên cứu?

Câu hỏi 2. Hãy nêu những bước cơ bản khi xây dựng khung lý thuyết?

Câu hỏi 3. Nêu những yếu tố cơ bản của khung lý thuyết? Yếu tố nào cần được chú ý nhất trong khung lý thuyết?

Câu hỏi 4. Vì sao cần khung lý thuyết trong nghiên cứu khoa học? Các nhà nghiên cứu có thể tìm khung lý thuyết ở đâu?

Câu hỏi 5. Thế nào là khung khái niệm? khung phân tích?

Câu hỏi 6. Hãy lấy một công trình khoa học bất kỳ và chỉ rõ khung lý thuyết của công trình này, nếu có. Khung lý thuyết này được xây dựng trên cơ sở nào? Được thể hiện như thế nào? Khung lý thuyết này đã định hướng cho nghiên cứu trong công trình như thế nào?

Thuật ngữ chính chương 3

Tiếng Anh	Tiếng Việt
Analytic	Phân tích
Analytic Framework	Khung phân tích
Cause and Effect	Nhân quả
Concept	Khái niệm
Conceptual framework	Khung khái niệm
Correlation relationship	Mối quan hệ tương quan
Factor	Nhân tố
Frameworks	Khung mẫu
Four framework approach	Tiếp cận bốn khung mẫu
Intermediate relationship	Mối quan hệ trung gian
Key concept	Khái niệm chính
Relationship conditions	Mối quan hệ điều kiện
Research hypotheis	Giả thuyết nghiên cứu
Research objectives	Mục tiêu nghiên cứu
Research methods	Phương pháp nghiên cứu
Research questions	Câu hỏi nghiên cứu
Methodological framework	Khung phương pháp
Scientific research	Nghiên cứu khoa học
Scientific method	Phương pháp khoa học
Statement of the problem	Xác định vấn đề nghiên cứu
Scope of study	Phạm vi nghiên cứu
Opinion	Quan điểm
Theory	Lý thuyết
Theoretical framework	Khung lý thuyết
Theoretical basis	Cơ sở lý thuyết
The hypothesis of factor	Giả thuyết của nhân tố
Variable	Biến số

NGHIÊN CỨU ĐỊNH LƯỢNG: PHƯƠNG PHÁP KHẢO SÁT

4.1. Giới thiệu khái niệm

4.1.1. Phương pháp khảo sát là gì?

Khảo sát là phương pháp sử dụng bảng hỏi (phiếu câu hỏi) để thu thập dữ liệu phục vụ cho mục tiêu nghiên cứu. Về mặt lý thuyết, phương pháp khảo sát có thể sử dụng cả trong nghiên cứu định tính và định lượng. Trên thực tế, phương pháp khảo sát thường được sử dụng nhằm thu thập dữ liệu trên diện rộng phục vụ các phương pháp nghiên cứu định lượng. Đây là phương pháp thu thập dữ liệu sơ cấp khá thông dụng trong nghiên cứu quản lý.

4.1.2. Khi nào dùng phương pháp khảo sát?

Phương pháp khảo sát phù hợp nhất khi dữ liệu cần thu thập có những đặc điểm như sau:

- Dữ liệu cần thu thập nằm rải rác ở từng đối tượng.
- Dữ liệu có sự khác biệt giữa các đối tượng.
- Dữ liệu thu thập từ các đối tượng là đáng tin cậy.
- Dữ liệu thu thập trên diện rộng.

Ngoài các vấn đề chung, trong nghiên cứu khảo sát, bốn vấn đề cơ bản cần được chú ý là:

- Xác định mẫu khảo sát (hỏi ai).
- Xây dựng phiếu khảo sát (hỏi cái gì).
- Quy trình khảo sát (hỏi như thế nào).
- Quy trình chuẩn bị số liệu (chuẩn bị số liệu như thế nào).

4.2. Xác định mẫu khảo sát

4.2.1. Mẫu và tổng thể

Tổng thể hay còn gọi là đám đông nghiên cứu là tập hợp tất cả phần tử được định nghĩa là thuộc phạm vi nghiên cứu.

Tổng thể nghiên cứu là tập hợp các phần tử mà thực tế có thể nhận dạng và lấy mẫu.

Mẫu nghiên cứu là một phần của tổng thể được lựa chọn để thu thập dữ liệu.

Đơn vị lấy mẫu là một hay một nhóm các phần tử để từ đó thực hiện việc lấy mẫu trong mỗi giai đoạn của quá trình chọn mẫu.

Phần tử là đơn vị mà nhà nghiên cứu cần quan sát và thu thập dữ liệu (cá nhân, hộ gia đình, tổ chức,...).

Khung mẫu là danh sách các đơn vị lấy mẫu có sẵn để phục vụ cho việc lấy mẫu.

4.2.2. Quy trình chọn mẫu

Bước 1. Định nghĩa tổng thể và phần tử.

Bước 2. Xác định khung lấy mẫu.

Bước 3. Xác định kích thước mẫu.

Bước 4. Xác định phương pháp chọn mẫu.

Bước 5. Tiến hành lấy mẫu theo phương pháp đã chọn.

4.2.3. Các phương pháp chọn mẫu cơ bản

4.2.3.1. Phương pháp chọn mẫu ngẫu nhiên đơn giản

Đây là phương pháp mà mỗi đối tượng trong tổng thể được gán một con số, sau đó các con số được lựa chọn một cách ngẫu nhiên. Thông thường các nhà nghiên cứu có thể dùng bảng ngẫu nhiên để chọn phần tử cho mẫu.

- **Ưu điểm:** Đơn giản nếu có một khung mẫu đầy đủ.

- **Nhược điểm:** Khó khả thi khi tổng thể lớn.

Ví dụ 1. Chọn 100 sinh viên trong 1000 sinh viên.

4.2.3.2. Phương pháp chọn mẫu ngẫu nhiên có hệ thống

Theo phương pháp này, toàn thể đối tượng trong tổng thể được liệt kê theo thứ tự định trước. Sau đó tùy vào quy mô mẫu và tổng thể mà quyết định khoảng cách các đối tượng được lựa chọn. Đây là phương pháp sử dụng phổ biến hơn phương pháp ngẫu nhiên đơn giản.

- **Ưu điểm:** không cần khung mẫu hoàn chỉnh.

- **Nhược điểm:** Mẫu sẽ bị lệch khi khung mẫu xếp theo chu kỳ và tần số bằng với bước nhảy.

Ví dụ 2. Dựa vào danh sách bầu cử tại một thành phố, ta có danh sách theo thứ tự vần của tên chủ hộ, bao gồm 240000 hộ. Ta muốn chọn ra một mẫu có 2000 hộ. Vậy khoảng cách chọn là: $k = 240000/2000 = 120$, có nghĩa là cứ cách 120 hộ thì ta chọn một hộ vào mẫu.

4.2.3.3. Phương pháp chọn mẫu phân tầng

Khi mẫu tương đối nhỏ, việc lựa chọn ngẫu nhiên theo hai phương pháp trên có thể dẫn tới một số đối tượng có tỷ lệ quá cao hoặc quá thấp trong mẫu. Phương pháp chọn mẫu

phân tầng giúp giải quyết vấn đề này. Theo phương pháp này, các đối tượng được chia theo nhóm. Sau đó đối tượng được chọn ngẫu nhiên trong từng nhóm theo tỷ lệ tương ứng với tổng thể.

- **Phân tầng ngẫu nhiên theo tỷ lệ:** Số phần tử trong mỗi tầng tỷ lệ với quy mô của mỗi tầng trong tổng thể.

- **Phân tầng ngẫu nhiên không theo tỷ lệ:** Sử dụng khi độ phân tán các phần tử trong mỗi tầng khác nhau đáng kể. Số phần tử trong mỗi tầng được chọn phụ thuộc vào độ phân tán của biến quan sát trong các tầng.

Ví dụ 3. Một tòa soạn báo muốn tiến hành nghiên cứu trên một mẫu 1000 doanh nghiệp trên cả nước về sự quan tâm của họ đối với tờ báo nhằm tiếp thị việc đưa thông tin quảng cáo trên báo. Tòa soạn có thể căn cứ vào các tiêu thức: vùng địa lý (miền Bắc, miền Trung, miền Nam); hình thức sở hữu (quốc doanh, ngoài quốc doanh, công ty 100% vốn nước ngoài,...) để quyết định cơ cấu của mẫu nghiên cứu.

4.2.3.4. Phương pháp chọn mẫu theo khu vực

Trong trường hợp mà các nhóm nghiên cứu không có khả năng di chuyển quá nhiều để phỏng vấn đối tượng, họ có thể áp dụng phương pháp chọn mẫu theo khu vực. Phương pháp này không lựa chọn các đối tượng mà lựa chọn một cách ngẫu nhiên khu vực, sau đó phỏng vấn toàn bộ đối tượng trong khu vực.

Ví dụ 4. Tổng thể chung là sinh viên của một trường đại học. Khi đó ta sẽ lập danh sách các lớp chứ không lập danh sách sinh viên, sau đó chọn ra các lớp để điều tra.

4.2.3.5. Phương pháp chọn mẫu thuận tiện

Nhà nghiên cứu có thể chọn những phần tử nào mà họ có thể tiếp cận được. Nói cách khác, hình thức chọn mẫu này dựa trên sự thuận tiện hay dựa trên tính dễ tiếp cận của đối tượng khảo sát.

Ví dụ 5. Chọn mẫu $n = 100$ doanh nghiệp bất động sản ở Thành Phố Hồ Chí Minh. Bất kỳ doanh nghiệp nào ở Thành Phố Hồ Chí Minh thuộc bất động sản và đồng ý tham gia vào mẫu đều có thể được chọn.

4.2.3.6. Phương pháp chọn mẫu phán đoán

Phương pháp chọn mẫu theo phán đoán là phương pháp mà người phỏng vấn là người tự đưa ra phán đoán về đối tượng cần chọn vào mẫu. Như vậy tính đại diện của mẫu phụ thuộc nhiều vào kinh nghiệm và sự hiểu biết của người tổ chức việc điều tra và cả người đi thu thập dữ liệu.

Ví dụ 6. Nhân viên phỏng vấn được yêu cầu đến các trung tâm thương mại chọn các phụ nữ ăn mặc sang trọng để phỏng vấn. Như vậy không có tiêu chuẩn cụ thể “thế nào là sang trọng” mà hoàn toàn dựa vào phán đoán để chọn ra người cần phỏng vấn.

4.2.3.7. Phương pháp chọn mẫu theo lớp

- Dựa vào một số thuộc tính kiểm soát xác định một số phần tử sao cho chúng đảm bảo tỷ lệ của tổng thể và các đặc trưng kiểm soát.

- Sử dụng phổ biến nhất trong thực tiễn nghiên cứu.

- Có thể dùng một hoặc nhiều thuộc tính kiểm soát như tuổi, giới tính, thu nhập, loại hình doanh nghiệp,...

Ví dụ 7. Chọn 200 hộ trong 1000 hộ ở các huyện như sau

Huyện	Số hộ trong mỗi lớp	Cỡ mẫu
A	250	50
B	150	30
C	400	80
D	200	40
Tổng	1000	200

Bảng 4.1. Thống kê lấy mẫu theo lớp.

4.2.4. Tính đại diện của mẫu

4.2.4.1. Quy mô mẫu

Quy mô mẫu càng lớn thì tính đại diện càng cao, nếu các điều kiện khác không đổi. Quy mô mẫu quá nhỏ thì không thể đại diện cho tổng thể. Trong nghiên cứu định lượng quy mô mẫu tối thiểu là 30 quan sát mới có thể áp dụng các công cụ thống kê suy diễn hay kiểm định. Quy mô mẫu cho các kiểm định thống kê hoặc hàm thống kê có nhiều biến số thường phải lớn hơn 100. Tuy nhiên, quy mô mẫu là điều kiện cần, không phải là điều kiện quyết định nhất tới tính đại diện của mẫu. Quy mô lấy mẫu nó phụ thuộc vào số lượng biến, nó phụ thuộc vào phương pháp phân tích,...

4.2.4.2. Quy trình và phương pháp chọn mẫu

Khi quy mô mẫu đã đảm bảo tương đối phù hợp với các phân tích thống kê (hơn 100 quan sát). Quy trình chọn mẫu trở thành yếu tố có tính chất quyết định tới đại diện cho mẫu. Trong điều kiện có thể, nhóm nghiên cứu nên áp dụng các phương pháp và quy trình chọn mẫu chuẩn mực ở trên.

4.2.5. Xác định cỡ mẫu

- Một câu hỏi luôn đặt ra với nhà nghiên cứu là cần phải điều tra bao nhiêu đơn vị mẫu để nó đại diện và có thể suy rộng cho tổng thể, để phân tích có ý nghĩa và kết quả nghiên cứu có giá trị về mặt khoa học?

- Một cách đơn giản và dễ nhất là dựa vào các nghiên cứu có cùng nội dung đã được thực hiện trước đó để lấy mẫu.

- Có thể hỏi ý kiến các chuyên gia, những người có kinh nghiệm thực hiện các dự án điều tra khảo sát.

- Có thể tính toán theo công thức tính mẫu.

- Với trường hợp cỡ mẫu lớn và không biết tổng thể.

$$n = \frac{C^2 f(1-f)}{\varepsilon^2}$$

Trong đó:

n : là cỡ mẫu (quy mô mẫu),

C : Giá trị tới hạn tương ứng với độ tin cậy lựa chọn,

f : Tỷ lệ mẫu (là ước tính tỷ lệ % của tổng thể), thường là tỷ lệ ước tính là 50%,

ε : Sai số cho phép ($\pm 3\%$, $\pm 4\%$, $\pm 5\%$...).

Ví dụ 8. Tính cỡ mẫu của một cuộc bầu cử với độ tin cậy là 95%, sai số cho phép là 5%, tỷ lệ ước lượng là 0,5

Với độ tin cậy 95% thì giá trị tới hạn là $C = 1,96$. Cỡ mẫu là

$$n = \frac{1,96^2(0,5*0,5)}{0,05^2} = 385$$

Nếu biết số lượng đơn vị trong tổng thể là N, tỷ lệ tổng thể là p, sai số cho phép là K thì cỡ mẫu được tính theo công thức

$$n = \left[\frac{1}{N} + \frac{N-1}{N} \frac{1}{p(1-p)} \left(\frac{K}{C} \right)^2 \right]^{-1}$$

Nếu tổng thể nhỏ và biết được tổng thể thì dùng công thức sau:

Với n là cỡ mẫu, N là số lượng tổng thể, e là sai số tiêu chuẩn

$$n = \frac{N}{1 + N(e)^2}$$

Ví dụ 9. Tính cỡ mẫu của một cuộc điều tra với tổng thể là $N = 2000$, độ chính xác là 95%, sai số tiêu chuẩn là 5%. Cỡ mẫu sẽ được tính là:

$$n = \frac{N}{1 + N(e)^2} = \frac{2000}{1 + 2000(0,05)^2} = 333.$$

4.3. Thiết kế bảng khảo sát

4.3.1. Những bước chính khi thiết kế bảng khảo sát

Bảng khảo sát là tập hợp các câu hỏi được trình bày theo một trình tự nhất định để người được hỏi trả lời dễ dàng và chính xác. Khi tiến hành thiết kế bảng khảo sát, người nghiên cứu cần phải trải qua 7 bước chính sau:

Bước 1. Xác định thông tin cần thu thập

Làm thế nào để xác định đầy đủ và chi tiết các thông tin cần thu thập

Khi thiết kế bảng câu hỏi phải dựa vào:

- Dựa vào vấn đề nghiên cứu.
- Dựa vào nhu cầu thông tin.
- Dựa vào khung lý thuyết.

Bước 2. Xác định phương pháp phỏng vấn

Có ba phương pháp phỏng vấn chính: phỏng vấn trực diện, phỏng vấn qua điện thoại, và phỏng vấn bằng cách gửi thư/email/câu hỏi điện tử. Đối với mỗi phương pháp khác nhau người nghiên cứu sẽ xây dựng cấu trúc bảng câu hỏi khác nhau.

Bước 3. Xác định nội dung câu hỏi

Nội dung các câu hỏi thường xoay quanh việc thu thập thông tin về:

- Các sự kiện thực tế.
- Kiến thức của đối tượng được hỏi.
- Ý kiến thái độ của người đó.
- Một số dữ liệu căn bản về cá nhân đối tượng nghiên cứu để phân loại, thông tin

liên lạc, và tìm kiếm các biến số liên quan.

Bước 4. Xác định hình thức câu trả lời

Trả lời cho các câu hỏi đóng, gồm các dạng:

- Chọn một trong nhiều lựa chọn.
- Chọn nhiều lựa chọn.
- Xếp theo thứ tự.

Trả lời cho các câu hỏi mở

- Câu hỏi trả lời tự do.
- Câu hỏi có tính chất thăm dò.

Bước 5. Xác định cách sử dụng từ ngữ

- Nên dùng từ ngữ quen thuộc, tránh dùng tiếng lóng hoặc từ chuyên môn.
- Nên dùng từ ngữ dễ hiểu, để mọi người ở bất cứ trình độ nào cũng có thể hiểu được.
- Tránh đưa ra câu hỏi dài quá.
- Tránh đặt câu hỏi mơ hồ, không rõ ràng.

Bước 6. Xác định trình tự và hình thức bảng câu hỏi

Nguyên tắc để có bảng câu hỏi đẹp:

- Mỗi phần nên được trình bày phân biệt (dùng màu giấy khác nhau).
- Đánh số các câu hỏi theo thứ tự.
- Mã hóa các phương án trả lời.
- Sử dụng dạng chữ rộng, rõ ràng.
- Dùng để các câu hỏi bị ngắt khi sang trang mới.
- Đưa ra các hướng dẫn cụ thể cho bảng hỏi nếu cần, tiếp đó là câu hỏi.
- Phân biệt giữa hướng dẫn với câu hỏi.

Bước 7. Phỏng vấn thử và hoàn thiện bảng câu hỏi

Câu hỏi đánh giá được nội dung theo mục đích đưa ra cho nó.

- Tất cả đều hiểu được câu hỏi và hiểu theo nghĩa giống nhau.
- Các hướng dẫn dễ hiểu hoặc dễ theo dõi.
- Liệu đã đưa ra hết các câu trả lời cho vấn đề chưa?
- Có thiếu câu hỏi nào một cách hệ thống/thường xuyên không?
- Kiểm tra các lỗi kỹ thuật cơ bản.

4.3.2. Những chú ý khi thiết kế từng câu hỏi

4.3.2.1. Cơ sở quan trọng khi xây dựng câu hỏi

Hai yếu tố cực kỳ quan trọng cần nắm rõ trước khi nghiên cứu:

- Thứ nhất là đặc điểm của đối tượng, ví dụ trình độ học vấn, văn hóa, điều kiện kinh tế, độ tuổi, ... Câu hỏi cần phù hợp với đặc điểm của đối tượng để đối tượng có thể và muốn trả lời.
- Thứ hai là thông tin cần thu thập theo khung nghiên cứu. Thông tin cần thu thập là gốc để đặt câu hỏi. Tuy nhiên, câu hỏi không nhất thiết hỏi thẳng vào thông tin

cần mà phải hỏi những thông tin mà đối tượng có thể trả lời.

4.3.2.2. Các loại câu hỏi

a. Phân theo hình thức, có câu hỏi đóng, câu hỏi có nhiều lựa chọn và câu hỏi mở

- Câu hỏi đóng đơn giản: là dạng câu hỏi chỉ có hai thái cực trả lời như “Có”/ “Không”, “Đúng”/ “Sai”,...

- Câu hỏi có lựa chọn định sẵn và đối tượng có thể chọn nhiều phương án phù hợp: Đây là một dạng khác của câu hỏi đóng đơn giản khi bản thân mỗi phương án là một câu hỏi đóng.

- Câu hỏi có lựa chọn định sẵn nhưng chỉ chọn một phương án.

- Câu hỏi mở: Dạng câu hỏi này không có các phương án để lựa chọn mà đối tượng có thể điền câu trả lời theo ý của mình. Câu hỏi mở được sử dụng hạn chế trong khảo sát định lượng vì sẽ mất công mã hóa.

b. Phân theo nội dung, câu hỏi có thể chia làm ba loại

- Câu hỏi về thông tin khách quan.

- Câu hỏi về hành vi hoặc trải nghiệm cụ thể.

- Câu hỏi về cảm nhận, thái độ và đánh giá của đối tượng.

4.3.3. Những chú ý khi thiết kế tổng thể bảng câu hỏi

Thiết kế tổng thể bảng câu hỏi cũng là một công đoạn quan trọng để đảm bảo đối tượng muốn trả lời bảng câu hỏi. Có một số kinh nghiệm khi thiết kế bảng câu hỏi như sau

4.3.3.1. Hình thức

Bảng câu hỏi cần được trình bày cẩn thận, dễ nhìn và nhất quán. Việc thiết kế cũng đảm bảo thuận lợi cho đối tượng lựa chọn và điền câu trả lời.

4.3.3.2. Giới thiệu

Bảng câu hỏi nên có phần giới thiệu hoặc thư giới thiệu đính kèm. Phần giới thiệu cần nêu mục đích cuộc khảo sát (không nhất thiết phải quá cụ thể - nên dừng ở mức mà đối tượng quan tâm). Phần này cũng nên khẳng định việc bảo mật danh tính người trả lời và cung cấp địa chỉ liên hệ của nhóm nghiên cứu.

4.3.3.3. Các câu hỏi cơ bản

Có thể phân chia câu hỏi theo các phần để đối tượng dễ trả lời. Nên bắt đầu bằng những phần dễ trả lời, ít nhạy cảm.

Trong một số trường hợp đối tượng trả lời có thể bỏ qua một số câu hỏi. Khi đó,

việc hướng dẫn chuyển câu hỏi cần được ghi rõ ràng (ví dụ: Nếu trả lời “Không”, chuyển sang câu 10).

Khi có các câu hỏi nhạy cảm, nên đan xen với những câu hỏi ít nhạy cảm hơn.

Ưu tiên các câu hỏi về thông tin khách hàng, sau đó đến câu hỏi về trải nghiệm và hành vi. Các câu hỏi về cảm nhận và đánh giá có ưu tiên thấp hơn, trừ khi chính cảm nhận và đánh giá của đối tượng là mục tiêu cần nghiên cứu.

4.3.3.4. Các câu hỏi theo nhóm

Các câu hỏi phân nhóm thường là đặc điểm của đối tượng trả lời (Cá nhân, hộ gia đình, doanh nghiệp,...). Các thông tin này dùng để phân nhóm, so sánh nhóm và để kiểm soát khi sử dụng các mô hình kiểm định thống kê.

4.3.3.5. Độ dài bảng câu hỏi

Độ dài bảng câu hỏi phụ thuộc vào hai yếu tố cơ bản: thông tin cần thu thập và nguồn lực của đề tài. Một bảng câu hỏi quá dài thường khó thuyết phục các đối tượng trả lời. Ngược lại, một bảng câu hỏi quá ngắn có thể không thu thập đủ thông tin cần thiết. Khi không có lợi ích đi kèm (ví dụ : quà tặng), một đối tượng có thể chỉ sẵn sàng dành 20 – 25 phút để trả lời bảng câu hỏi.

4.4. Xây dựng và thực hiện quy trình khảo sát

Sau khi xác định mẫu khảo sát và xây dựng bảng khảo sát, bước tiếp theo là xác định quy trình khảo sát để đảm bảo thu thập được thông tin với độ tin cậy cao.

Đối với khảo sát qua thư, quy trình như sau:

- Kiểm tra lại địa chỉ liên hệ của đối tượng.
- Tiến hành gửi thư tới các đối tượng.
- Gọi điện thông báo trước.
- Gửi thư cảm ơn.

Đối với khảo sát bằng phương pháp phỏng vấn, quy trình như sau:

- Tập huấn cho các bộ phỏng vấn: Phỏng vấn trực tiếp đòi hỏi có nhiều người cùng tham gia cùng phỏng vấn. Các cán bộ phỏng vấn cần thực hiện đúng quy trình phỏng vấn một cách nhất quán.

- Gọi điện liên hệ và hẹn thời gian phỏng vấn.
- Tiến hành phỏng vấn: Nhóm nghiên cứu cần đảm bảo địa điểm phỏng vấn thuận lợi cho việc trả lời một cách khách quan.
- Giám sát và đảm bảo chất lượng.

4.5. Quy trình chuẩn bị số liệu

4.5.1. Nhập liệu

Công đoạn này cần thiết đối với phương pháp thu thập qua thư hoặc phỏng vấn trực tiếp bằng phiếu giấy. Cần lưu ý:

- Mỗi dòng được dành cho một quan sát (thường là một phiếu).
- Mỗi cột là một trường dữ liệu.
- Mỗi phiếu câu hỏi gán một mã.
- Nhập dữ liệu theo trình tự câu hỏi và trung thành với giá trị trong bảng câu hỏi.

Không tiến hành điều chỉnh khi nhập số liệu trừ khi nhận rõ sai sót khi nhập số liệu.

- Nhập phiếu hai lần độc lập.

- File dữ liệu có thể được kiểm tra bằng các lệnh tần suất đơn giản. Nếu có các giá trị nằm ngoài khoảng cho phép hoặc đáng ngờ thì nhóm nghiên cứu có thể đối chiếu lại với phiếu câu hỏi.

4.5.2. Kiểm định các thước đo

Các biến số về thái độ, hành vi, hay cảm nhận thường được đo lường bằng một số câu hỏi hoặc mệnh đề. Kể cả khi những thước đo được kiểm định cẩn thận ở những nghiên cứu trước đó, đối với mỗi cuộc khảo sát, những thước đo này vẫn cần được kiểm tra về độ tin cậy.

- Phân tích nhân tố (factor analysis) : Phân tích nhân tố chính là việc kiểm tra xem các mệnh đề/câu hỏi có thực sự nhóm lại với nhau thành thước đo như trong lý thuyết hay không. Với các khảo sát khác nhau, có thể một số mệnh đề không vào cùng nhóm với các mệnh đề khác. Khi đó nhóm nghiên cứu cần tiếp tục kiểm tra độ tin cậy để ra quyết định.

- Phân tích độ tin cậy (Reliability analysis) : Phân tích độ tin cậy là xem các mệnh đề có thực sự “thống nhất” với nhau để cùng đo lường biến số cần đo hay không. Chỉ số đo lường sự thống nhất này là Cronbach’s alpha. Chỉ số này tốt là từ 0,7 trở lên và tối thiểu cần đạt là 0,63 (D’Vellis, 1990).

Các phần mềm thống kê có thể giúp thực hiện hai phép phân tích này khá nhanh chóng và dễ dàng. Bạn đọc có thể tìm hiểu kỹ thuật phân tích nhân tố ở các sách vở viết về thống kê toán.

4.6. Câu hỏi thảo luận

Câu hỏi 1. Nghiên cứu khảo sát là gì? Khi nào ta cần áp dụng nghiên cứu khảo sát? Nêu một ví dụ áp dụng phương pháp khảo sát?

Câu hỏi 2. Khi thiết kế khảo sát cần lưu ý những vấn đề gì? Nêu những hạn chế của phương pháp khảo sát?

Câu hỏi 3. Mẫu khảo sát cần đảm bảo yêu cầu gì? Yếu tố nào ảnh hưởng tới tính đại diện của mẫu.

Câu hỏi 4. Hãy nêu các phương pháp chọn mẫu? Phương pháp khảo sát nên chọn phương pháp chọn mẫu nào?

Câu hỏi 5. Trình bày phương pháp lấy mẫu theo xác suất? cho ví dụ minh họa.

Câu hỏi 6. Những loại nghiên cứu nào thì sử dụng phương pháp khảo sát?

Thuật ngữ chính chương 4

Tiếng Anh	Tiếng Việt
Area sampling	Chọn mẫu theo khu vực
Cluster	Khu vực, cụm
Convenience sampling	Chọn mẫu thuận tiện
Cross – sectional survey	Các nghiên cứu khảo sát ở một thời điểm
Element	Phần tử
Factor analysis	Phân tích nhân tố
Frequency	Tần suất
Model fit	Mô hình thích hợp
Method	Phương pháp
Method of stratified sampling	Phương pháp chọn mẫu phân tầng
Methods for selecting samples in clusters	Phương pháp chọn mẫu theo cụm
Non – Probability sampling methods	Chọn mẫu phi xác suất
Observe	Quan sát
Study population	Tổng thể (đám đông trong nghiên cứu)
Sampling	Mẫu
Sampling error	Sai số mẫu
Sampling frame	Khung lấy mẫu
Sampling method	Phương pháp chọn mẫu
Simple sampling method	Phương pháp chọn mẫu đơn giản
Sampling unit	Đơn vị mẫu
Sample size	Cỡ mẫu (quy mô mẫu)
Snowball	Quả bóng tuyết
Reliability analysis	Phân tích độ tin cậy
Probability sampling methods	Phương pháp chọn mẫu theo xác suất
Population	Tổng thể
Quota sampling	Chọn mẫu hạn ngạch
Starting point	Điểm xuất phát
Sampling interval	Bước nhảy
Systematic sampling method	Phương pháp chọn mẫu theo hệ thống

NGHIÊN CỨU ĐỊNH LƯỢNG: PHƯƠNG PHÁP THỬ NGHIỆM

5.1. Giới thiệu

Thử nghiệm là phương pháp mà nhà nghiên cứu chủ động thay đổi giá trị một biến số (biến độc lập) và quan sát xem sự thay đổi đó có ảnh hưởng tới biến số khác (biến phụ thuộc) hay không. Ưu điểm của phương pháp này là khả năng kiểm soát các biến khác, đảm bảo tính ngẫu nhiên trong việc chọn và phân nhóm đối tượng nghiên cứu và chủ động điều chỉnh giá trị các biến độc lập để kiểm định giả thuyết.

Phương pháp thử nghiệm là phương pháp tốt nhất để kiểm định mối quan hệ nhân quả. Tuy nhiên, hạn chế của phương pháp này là tính tổng quát hóa. Các phương pháp thử nghiệm thường được thực hiện trong điều kiện được kiểm soát chặt chẽ (ví dụ: phòng thí nghiệm), với một số đối tượng nhất định. Vì vậy khả năng áp dụng kết quả nghiên cứu trong điều kiện thực tiễn hoặc với đối tượng khác luôn là một câu hỏi đáng chú ý.

Phương pháp này được sử dụng thông dụng ở các ngành khoa học kỹ thuật như nông học, sinh học, y học, ... Trong lĩnh vực kinh tế - quản lý, phương pháp này cũng được nhiều nhà nghiên cứu tâm lý, marketing, hành vi tổ chức trên thế giới áp dụng. Một dạng nghiên cứu thử nghiệm (cận thử nghiệm) cũng được áp dụng rộng rãi trong lĩnh vực quản trị kinh doanh và kinh tế học. Tuy nhiên, phương pháp này chưa thực sự thông dụng trong các nghiên cứu kinh tế và quản lý ở Việt Nam.

5.2. Yêu cầu cơ bản của phương pháp thử nghiệm

5.2.1. Đảm bảo phân nhóm ngẫu nhiên

Để đảm bảo loại bỏ tác động của biến ngoại lai, tính ngẫu nhiên trong lựa chọn và phân nhóm đối tượng là hết sức quan trọng. Kỹ thuật để phân nhóm ngẫu nhiên đối tượng vào nhóm đối chứng (không nhân sự can thiệp) và nhóm thử nghiệm (nhận sự can thiệp) cũng có thể áp dụng như phần chọn mẫu ngẫu nhiên. Ví dụ một công trình nghiên cứu của một tác giả là muốn nghiên cứu tác động của phương pháp giảng dạy tới học tập của sinh viên. Các sinh viên đăng ký trước được phân vào nhóm “đối chứng” (giảng dạy truyền thống), trong khi các sinh viên đăng ký sau được phân vào nhóm “thực nghiệm” (giảng dạy theo phương pháp mới). Việc phân nhóm như vậy không đảm bảo tính ngẫu nhiên mà tác giả phải tập hợp danh sách tất cả sinh viên đăng ký

sau đó mới phân bổ ngẫu nhiên vào hai nhóm. Hiện giờ có rất nhiều phần mềm có thể sử dụng để phân nhóm ngẫu nhiên.

5.2.2. Sử dụng nhóm đối chứng

Sử dụng nhóm đối chứng là một yêu cầu hết sức quan trọng trong nghiên cứu thử nghiệm. Nhóm đối chứng có vai trò chính là cơ sở để so sánh về kết quả của “can thiệp thử nghiệm” và là cơ sở để kiểm định các giả thuyết khác (ngoài giả thuyết của nghiên cứu). Cần lưu ý đối tượng tham gia nhóm đối chứng cần tương đồng với đối tượng tham gia nhóm thử nghiệm.

5.2.3. Biến độc lập (yếu tố can thiệp) đủ mạnh

Trong nghiên cứu thử nghiệm, biến độc lập được các nhà nghiên cứu chủ động điều chỉnh “giá trị”. Nhà nghiên cứu chủ động “can thiệp” vào biến độc lập và quan sát sự thay đổi của biến phụ thuộc. Sự can thiệp được chủ động tạo ra này cần đủ mạnh để đối tượng tham gia nhóm thử nghiệm phải “cảm thấy” được sự khác biệt, so với nhóm đối chứng.

5.3. Thiết kế thử nghiệm có đối chứng

5.3.1. Chỉ đo lường sau thử nghiệm

Thiết kế này chỉ đo lường biến phụ thuộc sau khi đã tiến hành thử nghiệm. Thiết kế này được sử dụng khá nhiều trong lĩnh vực Marketing.

Ví dụ 1. Một nhóm nghiên cứu muốn tìm hiểu tác động của chương trình quảng cáo sử dụng sản phẩm mẫu. Họ có thể áp dụng thiết kế nghiên cứu này như sau:

Bước 1. Phân nhóm ngẫu nhiên khách hàng tiềm năng vào hai nhóm: Nhóm thử nghiệm và nhóm đối chứng. Các tác giả có thể kiểm tra mức độ tương đồng của hai nhóm về các chỉ số cơ bản như tuổi, giới tính, thu nhập, trình độ học vấn,...

Bước 2. Tiến hành thử nghiệm: Cho các thành viên trong nhóm thử nghiệm được sử dụng sản phẩm mẫu, trong khi đó nhóm đối chứng không được sử dụng sản phẩm mẫu.

Bước 3. Các thành viên tham gia nghiên cứu trong hai nhóm đều được gửi phiếu giảm giá khi mua sản phẩm ở siêu thị. Phiếu giảm giá được mã hóa để phân biệt, nhận dạng được cả hai nhóm.

Bước 4. Sau một khoảng thời gian (ví dụ: 1 tháng), số phiếu giảm giá của mỗi nhóm thu lại từ siêu thị sẽ được đếm. So sánh kết quả hai nhóm để đánh giá sự tác động của hoạt động quảng cáo sản phẩm mẫu.

5.3.2. Đo lường trước – sau thử nghiệm

Thiết kế trước – sau thử nghiệm khác với thiết kế “chỉ đo lường sau khi thử nghiệm” ở chỗ các biến phụ thuộc được đo lường trước và sau khi tiến hành thử nghiệm. Thiết kế này giúp kiểm soát tốt các tác động ngoại lai và rất phù hợp với việc đánh giá tác động ngắn hạn của thí nghiệm.

Ví dụ 2. Giả sử một nhóm tác giả muốn nghiên cứu tác động của chuyến thăm và nói chuyện về bóng đá của danh thủ Messi tới niềm đam mê bóng đá của trẻ em Việt Nam. Nhóm nghiên cứu có thể áp dụng thiết kế nghiên cứu thử nghiệm “trước – sau” có đối chứng như sau:

Bước 1. Chọn một mẫu trẻ em. Phân ngẫu nhiên họ thành hai nhóm: Nhóm đối chứng và nhóm thử nghiệm. Kiểm tra sự tương đồng của hai nhóm về các chỉ số cơ bản như giới tính, tuổi, sở thích bóng đá,...

Bước 2. Đo lường trước: cả hai nhóm đều được đo lường về sự đam mê bóng đá và hiểu biết về bóng đá.

Bước 3. Khi danh thủ Messi đến thăm và chia sẻ về bóng đá, chỉ có nhóm thử nghiệm được tham gia dự buổi nói chuyện và chia sẻ của danh thủ Messi. Nhóm đối chứng không được tiếp xúc hoặc nghe bất kỳ bài nói chuyện nào của danh thủ Messi.

Bước 4. Sau buổi nói chuyện của danh thủ Messi, nhóm nghiên cứu có thể đo lường lại niềm đam mê bóng đá của hai nhóm (khoảng sau 1 tuần).

Bước 5. Nhóm nghiên cứu so sánh sự thay đổi về suy nghĩ, thái độ đối với môn bóng đá của hai nhóm và sự khác biệt này là do được tham dự nghe buổi nói chuyện của danh thủ Messi.

Khác với thử nghiệm đo lường sau, thử nghiệm trước – sau không chỉ cho phép so sánh các kết quả cuối cùng mà so sánh sự khác biệt giữa kết quả đo lường sau và trước. Thiết kế thử nghiệm trước - sau vì vậy cho phép đo lường trực tiếp tác động của thử nghiệm và so sánh tác động của thử nghiệm so với không có thử nghiệm.

	Đo lường trước	Thực hành	Đo lường sau	
Thử nghiệm	Y_0	X	Y_1	$Y_0 - Y_1$
Đối chứng	Y_0	Y_1		$Y_0 - Y_1$

Bảng 5.1. Mô hình thiết kế thử nghiệm “trước – sau”.

5.4. Áp dụng nghiên cứu thử nghiệm trên thực địa

Thiết kế thử nghiệm như đã trình bày ở trên là thiết kế đầy đủ với điều kiện nhà nghiên cứu có thể kiểm soát toàn bộ quá trình, kể từ việc lựa chọn đối tượng, loại bỏ ảnh hưởng ngoại lai, tới việc điều tiết các mức độ/giá trị của biến độc lập. Điều này thường được đảm bảo với các thiết kế thử nghiệm ở phòng thí nghiệm.

Nghiên cứu thử nghiệm ngoài thực địa thường khó có thể đảm bảo điều kiện trên. Vì vậy, thiết kế nghiên cứu cận thử nghiệm được gọi là quasi experiment thường được áp dụng. Với dạng thiết kế thử nghiệm này, các nhà nghiên cứu coi những biến động trên thực địa (chính sách, thị trường, chính trị,...) là “sự can thiệp” giống như biến độc lập được điều chỉnh trong thử nghiệm và tìm cách đánh giá tác động của những can thiệp đó.

Một số dạng áp dụng thông dụng bao gồm:

5.4.1. Đánh giá tác động của dự án hoặc chính sách

Mỗi một dự án hoặc chính sách mới có thể coi là một sự can thiệp, tương đồng với điều tiết biến độc lập trong nghiên cứu thử nghiệm. Vì vậy đánh giá tác động của của dự án hoặc chính sách, các nhà nghiên cứu có thể áp dụng phương pháp thử nghiệm không đầy đủ. Trong đánh giá các nhà nghiên cứu có thể chọn nhóm đối chứng (những cá thể tương đồng song không thuộc nhóm điều chỉnh chính sách hoặc dự án) và nhóm thuộc diện chính /dự án. Chỉ số tác động (biến phụ thuộc), ví dụ như chất lượng cuộc sống hay nhận thức về một vấn đề gì đó được đo lường trước và sau dự án/chính sách. Các nhà nghiên cứu có thể so sánh sự khác biệt trong thay đổi giữa trước – sau của hai nhóm để đánh giá sự tác động của chính sách/dự án. Khó khăn thường là khó tìm nhóm đối chứng tương đồng, đặc biệt khi đánh giá tác động của chính sách có tầm bao phủ toàn quốc và cho mọi đối tượng. Một số kỹ thuật thống kê có thể giúp xác định nhóm đối chứng tương đồng nhất trên thực địa, song không thể có độ tương đồng cao như trong thiết kế thử nghiệm đầy đủ (ở phòng thí nghiệm).

5.4.2. Đánh giá tác động của biến động trên thực địa (chính trị, thị trường, hoặc tự nhiên)

Các nhà nghiên cứu cũng có thể đánh giá tác động của biến động chính trị, biến động của thị trường hay tự nhiên tới hành vi của doanh nghiệp, người dân,... Chỉ có điều khác là biến động này không có tính “chủ động” như chính sách dự án. Vì vậy thường khó

có những khảo sát cơ sở theo đúng mục tiêu và các nhà nghiên cứu phải sáng tạo trong việc sử dụng các dữ liệu sẵn có trước biến động để làm cơ sở so sánh.

5.5. Câu hỏi thảo luận

Câu hỏi 1. Phương pháp thử nghiệm là gì? Vì sao nói đây là phương pháp tốt nhất để kiểm chứng mối quan hệ nhân quả?

Câu hỏi 2. Các yêu cầu chính của phương pháp thử nghiệm là gì?

Câu hỏi 3. Trình bày một phương pháp thử nghiệm mà bạn biết? Nếu điểm mạnh, điểm yếu của phương pháp này?

Câu hỏi 4. Lĩnh vực nào thì ta nên dùng phương pháp thử nghiệm? Tại sao?

Câu hỏi 5. Hãy tìm một ví dụ sử dụng phương pháp thử nghiệm không đầy đủ trong lĩnh vực mà bạn quan tâm. Mô tả thiết kế nghiên cứu này?

Câu hỏi 6. Nghiên cứu thử nghiệm không đầy đủ là gì? So sánh phương pháp này với nghiên cứu thử nghiệm?

Thuật ngữ chính chương 5

Tiếng Anh	Tiếng Việt
Affect	Có ảnh hưởng
Baseline survey	Khảo sát cơ sở theo đúng mục tiêu
Control	Kiểm soát
Control variables	Biến số kiểm soát
Dependent variable	Biến phụ thuộc
Difference in differences	Khác biệt trong sự khác biệt
Designing	Thiết kế
Experiment	Thử nghiệm
Exotic variable	Biến ngoại lai
Field test	Thử nghiệm trên thực địa
Group test	Nhóm thử nghiệm
History	Lịch sử
Independent variable	Biến số độc lập
Intervening variable	Biến số can thiệp
Intervening factor	Yếu tố can thiệp
Method	Phương pháp
Measure	Thước đo
Random	Ngẫu nhiên
Random group	Nhóm ngẫu nhiên
Research subjects	Đối tượng nghiên cứu
Quasi – experiment	Tựa thử nghiệm (cận thử nghiệm)
Similarities group	Nhóm tương đồng
Simulation	Mô phỏng

THIẾT KẾ NGHIÊN CỨU TỔNG THỂ

6.1. Giới thiệu khái niệm

6.1.1. Khái niệm

Thiết kế nghiên cứu là tầm nhìn của nhà nghiên cứu về kết quả nghiên cứu với các chuẩn mực về độ chặt chẽ và tính phổ quát, cùng quy trình và nguồn lực tương ứng nhằm đạt các chuẩn mực đó. Thiết kế nghiên cứu phải gắn chặt với câu hỏi nghiên cứu và khung lý thuyết. Thiết kế nghiên cứu là trung tâm kết nối các hoạt động cơ bản của dự án nghiên cứu: xác định mẫu, thước đo, quy trình thu thập dữ liệu, phương pháp phân tích dữ liệu để trả lời cho câu hỏi nghiên cứu.

Thiết kế nghiên cứu có thể dưới dạng đơn lẻ hoặc kết hợp. Thiết kế nghiên cứu đơn lẻ là việc sử dụng một phương pháp nghiên cứu cụ thể (ví dụ: khảo sát hoặc thực nghiệm). Thiết kế kết hợp là việc sử dụng nhiều hơn một phương pháp.

6.1.2. Vai trò của thiết kế nghiên cứu

Vai trò đảm bảo chuẩn mực nghiên cứu: Thiết kế nghiên cứu giúp đảm bảo các dữ liệu thu thập cho phép trả lời câu hỏi một cách rõ ràng và chặt chẽ nhất có thể.

Nói khác đi, thiết kế nghiên cứu chính là việc trả lời câu hỏi: Cần dữ liệu gì để trả lời câu hỏi một cách thuyết phục nhất?

Vai trò kế hoạch: Thiết kế nghiên cứu có vai trò giúp nhà nghiên cứu chuẩn bị nguồn lực và lên kế hoạch cho các hoạt động một cách phù hợp nhất. Khi lựa chọn một thiết kế, các nhà nghiên cứu đã tính toán trước hạn chế của nghiên cứu và xác định liệu hạn chế này chấp nhận được không.

6.1.3. Phân biệt thiết kế nghiên cứu với đề cương báo cáo

Thiết kế nghiên cứu là việc hoạch định quy trình, phương pháp và nguồn lực nhằm trả lời câu hỏi nghiên cứu một cách chặt chẽ, thuyết phục. Trong khi đó đề cương báo cáo chỉ đơn thuần là việc dự kiến các mục trình bày trong báo cáo. Thiết kế quan tâm tới quá trình thực hiện nghiên cứu, còn đề cương báo cáo quan tâm tới trình bày kết quả nghiên cứu. Thiết kế là vấn đề tư duy vì nó phải thể hiện sự gắn kết giữa mục tiêu nghiên cứu với quy trình, phương pháp, nguồn lực. Báo cáo nghiên cứu thiên về vấn đề trình bày: Cùng một nghiên cứu nhưng báo cáo có thể khác nhau tùy theo đối tượng. Trước khi tiến hành nghiên cứu, các nhà nghiên cứu nên xây dựng và bảo vệ thiết kế

nghiên cứu. Đề cương báo cáo chỉ nên dừng ở mức độ tham khảo.

6.2. Các yêu cầu chính trong thiết kế nghiên cứu

6.2.1. Tính chặt chẽ

Một nghiên cứu không chỉ đơn giản là đi tìm dữ liệu và bằng chứng phù hợp với giả thuyết hay luận điểm định trước. Tính chặt chẽ đòi hỏi nghiên cứu phải tìm đủ bằng chứng/dữ liệu để bác bỏ hoặc kiểm soát các giả thuyết “cạnh tranh” khác.

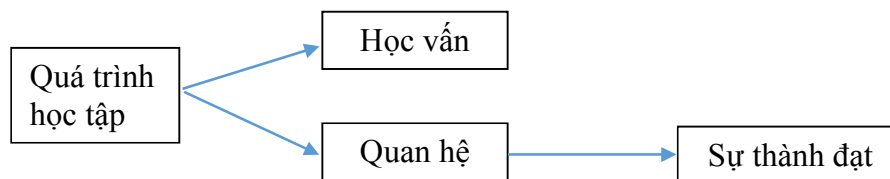
Ví dụ dưới đây minh họa điều này:

Một người nghiên cứu muốn kiểm định vai trò của vốn con người tới sự thành đạt của các cá nhân. Tác giả thu thập dữ liệu và kết quả minh chứng rằng học vấn càng cao (vốn con người cao) thì sự thăng tiến trong công việc càng cao (sự thành đạt). Nghiên cứu này được mô phỏng như sau:



Sơ đồ 6.1. Mô phỏng học vấn càng cao thì càng thành đạt

Nếu nghiên cứu dừng ở đây sẽ chưa đảm bảo sự chặt chẽ vì chưa tính tới các giả thuyết “cạnh tranh”. Một trong những giả thuyết cạnh tranh có thể là *Học vấn* không hề ảnh hưởng tới *Sự thành đạt*. Chẳng qua người có học vấn cao là người có quan hệ nhiều hơn với những người thành đạt hơn. Như vậy không phải vốn con người (học vấn) mà là *vốn xã hội (quan hệ) mang lại sự thành đạt*. Giả thuyết này có thể được mô phỏng như sau:



Sơ đồ 6.2. Mô phỏng quá trình học tập dẫn tới sự thành đạt.

6.2.2. Tính khái quát

Một trong những chuẩn mực của nghiên cứu khoa học là tính phổ biến của kết quả nghiên cứu. Tính khái quát hóa của nghiên cứu đòi hỏi kết quả nghiên cứu phải có khả năng suy rộng. Có 3 loại tổng quát hóa cơ bản sau:

- Khái quát cho tổng thể đối tượng nghiên cứu: Kết quả từ một mẫu nghiên cứu liệu có thể suy rộng cho tổng thể nghiên cứu hay không?

Ví dụ 1. Kết quả nghiên cứu từ một mẫu gồm 200 sinh viên đại học liệu có thể suy rộng cho tổng thể là sinh viên đại học được hay không? hoặc rộng hơn nữa, liệu có

thể suy rộng cho trí thức trẻ được không? Điều này phụ thuộc rất nhiều tính đại diện của mẫu nghiên cứu, trong đó quy trình chọn mẫu và quy mô mẫu có ý nghĩa quyết định.

- Khái quát cho các bối cảnh nghiên cứu khác nhau: Kết quả nghiên cứu có thể suy rộng cho các bối cảnh khác nhau được hay không?

Ví dụ 2. Một nghiên cứu cảm nhận và trải nghiệm của người dân về tham nhũng ở 10 tỉnh, thành phố có thể suy rộng cho các tỉnh thành phố trong cả nước hay không? Điều này phụ thuộc vào việc 10 tỉnh, thành phố được nghiên cứu có thể đại diện cho 63 tỉnh thành phố về cảnh kinh tế, xã hội, văn hóa hay không?

- Khái quát cho các thời điểm khác nhau: Liệu kết quả nghiên cứu có trường tồn theo thời gian không? Điều này phụ thuộc rất nhiều liệu thời gian cho làm thay đổi khung cảnh và làm thay đổi kết quả nghiên cứu hay không. Những nghiên cứu mang tính mô tả (ví dụ: mô tả thực trạng nền kinh tế hay giá trị văn hóa) không có tính trường tồn cao. Những nghiên cứu hướng vào mối quan hệ có tính quy luật có tính trường tồn cao hơn. Tuy nhiên, khi bối cảnh thay đổi lớn thì kết quả cũng có thể thay đổi.

6.2.3. Tính khả thi

Không có nghiên cứu nào có nguồn lực vô hạn. Ngoài ra, kết quả nghiên cứu còn phụ thuộc vào dữ liệu sẵn có. Vì vậy, nếu thiết kế nghiên cứu vượt ra ngoài khả năng về nguồn lực và tiếp cận dữ liệu thì cũng không có ý nghĩa thực thi. Vì vậy, các nhà nghiên cứu cần cân đối giữa hai yêu cầu trên (tính chặt chẽ và tính khái quát hóa) với nguồn lực và khả năng tiếp cận dữ liệu trong thiết kế của mình.

6.3. Giới thiệu một số thiết kế nghiên cứu

6.3.1. Các bước thiết kế nghiên cứu

Bước 1. Chọn đề tài nghiên cứu

Bước 2. Xác định câu hỏi nghiên cứu

Bước 3. Mô tả thiết kế nghiên cứu để thực hiện

- Cách thu thập số liệu.
- Những thông tin cần thu thập để trả lời câu hỏi nghiên cứu.
- Các phương pháp đề tài sử dụng để thu thập dữ liệu.
- Ưu và nhược điểm của thiết kế nghiên cứu này.

6.3.2. Ví dụ về một thiết kế nghiên cứu cụ thể

Xét một thiết kế cụ thể được cho dưới bảng sau

Mục tiêu nghiên cứu	Giả thuyết nghiên cứu	Dữ liệu cần thiết	Phương pháp thu thập	Hạn chế
Mục tiêu 1: Nghiên cứu mối quan hệ học vấn và thu nhập	Giả thuyết nghiên cứu: Học vấn càng cao càng có thu nhập cao Giả thuyết cạnh tranh 1: Càng nhiều mối quan hệ càng có thu nhập cao Giả thuyết cạnh tranh 2: Sự trợ giúp của bố mẹ càng nhiều càng có thu nhập cao.	Dữ liệu ở cấp độ từng cá nhân: Thu nhập Học vấn Mối quan hệ xã hội (số lượng, nhóm,..) Sự trợ giúp của bố mẹ (tài chính, định hướng nghề nghiệp) Tuổi, giới tính, dân tộc,..	Khảo sát ngẫu nhiên một mẫu người trưởng thành.	Dữ liệu khảo sát không cho biết mối quan hệ nhân quả: Vì học vấn cao, có nhiều mối quan hệ nên thu nhập cao hay ngược lại.
Mục tiêu 2: Nghiên cứu mối quan hệ giữa học vấn và sự thăng tiến trong công việc

Bảng 6.1. Ví dụ về một thiết kế.

6.4. Câu hỏi thảo luận

Câu hỏi 1. Thiết kế nghiên cứu tổng thể là gì? Nêu những yêu cầu chính khi thiết kế nghiên cứu tổng thể.

Câu hỏi 2. Giả thuyết cạnh tranh là gì? Tại sao khi thiết kế nghiên cứu chúng ta phải quan tâm tới giả thuyết cạnh tranh? Làm thế nào để đảm bảo giả thuyết cạnh tranh trọng được kiểm soát trong thiết kế nghiên cứu?

Câu hỏi 3. Thiết kế nghiên cứu tổng thể kết hợp nhiều phương pháp có ưu/nhược điểm gì? Hãy nêu một ví dụ cụ thể về một thiết kế nghiên cứu kết hợp và bình luận điểm mạnh, điểm yếu của thiết kế này.

Câu hỏi 4. Giả thuyết cạnh tranh là gì? Tại sao khi thiết kế nghiên cứu chúng ta phải quan tâm tới giả thuyết cạnh tranh?

Câu hỏi 5. Một người cho rằng “ Hướng đổi mới phương pháp dạy học hiện nay là các nghiên cứu phải áp dụng nghiên cứu định tính”. Anh/chị có đồng tình với nhận định đó hay không? Tại sao?

Thuật ngữ chính chương 6

Tiếng Anh	Tiếng Việt
Assumptions	Giả thiết
Data collection process	Quy trình thu thập dữ liệu
Data analysis	Phân tích dữ liệu
Data over time	Dữ liệu theo thời gian
Data collection	Thu thập dữ liệu
External validity	Tính khái quát hóa
Factor scale	Thang đo nhân tố
For example	Ví dụ
General	Khái quát
Internal validity	Tính chặt chẽ
Main design	Thiết kế chính
Models	Mô hình
New points of research	Điểm mới của nghiên cứu
Research outline	Đề cương nghiên cứu
Report	Báo cáo
Research report	Báo cáo nghiên cứu
Research restrictions	Hạn chế nghiên cứu
Research summary	Tóm tắt nghiên cứu
Research models	Mô hình nghiên cứu
Scale	Thang đo
Scheme	Kế hoạch

XỬ LÝ VÀ PHÂN TÍCH DỮ LIỆU ĐỊNH LƯỢNG CHO NGHIÊN CỨU

7.1. Tổng quan kiến thức về thống kê và sử dụng các kỹ thuật thống kê

7.1.1. Phân tích mô tả và khám phá

7.1.1.1. Thống kê mô tả

Thống kê mô tả cung cấp các chỉ số cơ bản của biến số với dữ liệu của mẫu nghiên cứu. Hầu hết các nghiên cứu định lượng đều cần cung cấp các chỉ số thống kê mô tả để giúp người đọc hiểu về dữ liệu sử dụng. Các chỉ số và cách trình bày có thể khác nhau với biến định lượng và biến định danh.

Đối với các biến có giá trị liên tục (biến định lượng). Các nhà nghiên cứu thường cung cấp các chỉ số như giá trị trung bình, giá trị lớn nhất, giá trị nhỏ nhất và độ lệch chuẩn của biến. Trong một số trường hợp thì cần thêm giá trị trung vị.

Đối với các biến định danh. Các nhà nghiên cứu thường cung cấp các chỉ số như tần suất, tỷ lệ phần trăm trong tổng số, giá trị trung vị, giá trị yếu vị.

Ví dụ 1. Cho bảng thống kê mô tả của các biến định lượng

	Y	X2	X3
	Y	X2	X3
Mean	141.3333	20.41667	12.16667
Median	141.5000	20.50000	12.00000
Maximum	180.0000	26.00000	17.00000
Minimum	102.0000	15.00000	6.000000
Std. Dev.	23.20789	4.187825	3.352972
Skewness	-0.218444	-0.034924	-0.453123
Kurtosis	2.301560	1.401529	2.378514
Jarque-Bera	0.339345	1.279994	0.603763
Probability	0.843941	0.527294	0.739426
Sum	1696.000	245.0000	146.0000
Sum Sq. Dev.	5924.667	192.9167	123.6667
Observations	12	12	12

Bảng 7.1. Thống kê mô tả.

Trong đó

- Mean: Trung bình của lượng biến.

- Median: Trung vị của lượng biến.
- Maximun: Giá trị lớn nhất của lượng biến.
- Minimun: Giá trị nhỏ nhất của lượng biến.
- Std. Dev: Độ lệch chuẩn của lượng biến.
- Skewness: Hệ số bất đối xứng.
- Kurtosis: Hệ số nhọn.
- Jarque – Bera: Giá trị thống kê Jarque – Bera dùng để kiểm định phân phối chuẩn.
- Probability: Giá trị xác suất của thống kê Jarque – Bera dùng để kiểm định phân phối chuẩn.
- Sum: Tổng các giá trị của lượng biến.
- Sum Sq. Dev: Độ lệch chuẩn của tổng.
- Observations: Tổng số quan sát.

7.1.1.2. Ma trận hệ số tương quan

Các biến số có thể có tương quan với nhau, ma trận hệ số tương quan là một công cụ ban đầu để giúp các tác giả và người đọc quan sát về mối tương quan của từng cặp biến. Công đoạn này cũng giúp các tác giả nhận biết các hiện tượng bất thường hoặc đề phòng trường hợp đa cộng tuyến khi các biến độc lập có tương quan lớn.

Ví dụ 2. Cho ma trận tương quan các biến định lượng như sau:

Correlation			
	Y	X2	X3
Y	1.000000	0.782281	0.904627
X2	0.782281	1.000000	0.480173
X3	0.904627	0.480173	1.000000

Bảng 7.2. Ma trận tương quan giữa các biến.

Ý nghĩa. Ma trận tương quan cho biết mối tương quan giữa các biến trong mô hình. Ví dụ hệ số tương quan của X_2 và X_3 là 0,480173; hệ số tương quan của Y và X_2 là 0,782281; hệ số tương quan của Y và X_3 là 0,904627.

7.1.1.3. Phân tích nhân tố khám phá

Phân tích nhân tố khám phá là một phương pháp phân tích thống kê dùng để rút gọn một tập gồm nhiều biến quan sát có liên hệ với nhau thành một tập biến (gọi là các nhân tố) ít hơn để chúng có ý nghĩa hơn nhưng vẫn chứa đựng hầu hết nội dung thông

tin của tập biên bản đầu (Hair và cộng sự, 1998).

Phân tích nhân tố khám phá thường được sử dụng nhiều nhất trong nghiên cứu khảo sát khi mà tác giả phải dùng nhiều câu hỏi để thu thập thông tin một vấn đề trừu tượng hơn, đặc biệt là những thông tin về tâm lý, thái độ, thậm chí hành vi. Kể cả khi chúng ta sử dụng thước đo đã được các tác giả trước phát triển và kiểm định thì cũng vẫn nên thực hiện kỹ thuật này xem liệu các mệnh đề/câu hỏi có “nhóm” đúng theo thước đo ban đầu hay không. Kết quả phân tích nhân tố là cơ sở để tạo biên số cho các phân tích tiếp theo.

7.1.1.4. Phân tích độ tin cậy của thước đo

Phân tích độ tin cậy cho phép chúng ta xác định thuộc tính của thước đo mà chủ yếu là liệu các mệnh đề/câu hỏi của thước đo có “thống nhất” với nhau hay không. Thông thường các tác giả sử dụng chỉ số Cronbach’ alpha từ 0,7 trở lên, song giá trị tối thiểu để thước đo có thể sử dụng được là 0,63 (DeVellis, 1990).

Phân tích độ tin cậy của thước đo thường được sử dụng cùng với phân tích nhân tố khám phá để quyết định các mệnh đề/câu hỏi cho từng thước đo. Lý tưởng nhất thước đo đủ cả ba điều kiện:

- Các câu hỏi/mệnh đề của thước đo được phát triển dựa trên lý thuyết hoặc đã được các tác giả trước xây dựng và kiểm định.
- Các câu hỏi/mệnh đề của thước đo “nhóm” cùng với nhau khi thực hiện phân tích nhân tố khám phá.
- Các câu hỏi/mệnh đề có chỉ số Cronbach’s alpha từ 0,7 trở lên, hoặc ít nhất cũng là 0,63.

7.1.2. So sánh nhóm

Một dạng nghiên cứu định lượng khá thông dụng là so sánh sự khác biệt giữa các nhóm về một hoặc một số chỉ số nào đó. Dưới đây là liệt kê các công cụ chính:

- T – test (kiểm định t) : được sử dụng để so sánh hai giá trị trung bình.
- ANOVA và ANCOVA (Analysis of Covariance) : Khi có nhiều hơn hai nhóm cần so sánh thì sử dụng ANOVA. Khi so sánh các nhóm, đồng thời kiểm soát tác động của một biến liên tục khác thì sử dụng ANCOVA.
- MANOVA (Multivariate Analysis of Variance): Tương tự như ANOVA nhưng được sử dụng khi có nhiều hơn một biến phụ thuộc và các biến phụ thuộc lại tương quan chặt với nhau.

7.2. Hồi quy tuyến tính cho phân tích dữ liệu định lượng

7.2.1. Mô hình hồi quy đơn

7.2.1.1. Hàm hồi quy tổng thể

Giả sử ta có các bộ số liệu $(X_i, Y_{i,j})$ cho tổng thể, với $i = 1, 2, \dots, n$; $j = 1, 2, \dots, m(i)$. Ứng với mỗi giá trị của X , $X = X_i$, với $i = 1, 2, \dots, n$, ta có thể có nhiều giá trị của Y tương ứng nên quan hệ của Y theo X không là quan hệ “hàm số”. Tuy nhiên, ứng với mỗi giá trị của X , $X = X_i$, ta có duy nhất giá trị trung bình $E(Y | X = X_i)$, nên quan hệ này trở thành quan hệ hàm số

$$E(Y | X = X_i) \equiv f(X_i)$$

và hàm số này được gọi là *hàm hồi quy tổng thể*, PRF (Population Regression Functions) mà trong trường hợp này, ta còn gọi là *hàm hồi quy đơn (hồi quy hai biến)*, do nó chỉ có một biến độc lập. Trường hợp có nhiều hơn một biến độc lập, ta gọi là *hàm hồi quy bội*.

Trước hết, giả sử PRF là hàm tuyến tính

$$E(Y | X = X_i) = \beta_1 + \beta_2 X_i$$

mà ta còn viết là

$$E(Y | X) = \beta_1 + \beta_2 X,$$

trong đó β_1 và β_2 là các tham số chưa biết nhưng cố định, được gọi là các *hệ số hồi quy*; β_1 gọi là *hệ số tự do* hay *hệ số chặn*, β_2 gọi là *hệ số góc* (nó cho biết tỷ lệ thay đổi của Y đối với X).

Dạng ngẫu nhiên

$$Y = \beta_1 + \beta_2 X + \varepsilon,$$

trong đó ε là một đại lượng ngẫu nhiên.

7.2.1.2. Hàm hồi quy mẫu

Hàm hồi quy quy tuyến tính mẫu có dạng

$$\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X,$$

trong đó \hat{Y} là ước lượng điểm của $E(Y | X)$, $\hat{\beta}_1$ là ước lượng điểm của β_1 và $\hat{\beta}_2$ là ước lượng điểm β_2 .

7.2.1.3. Phương pháp bình phương bé nhất

Phương pháp bình phương nhỏ nhất, OLS (Ordinary Least Square), do nhà toán học Đức Carl Fredrich Gauss đưa ra. Với phương pháp này, kèm theo một vài giả thiết, các ước lượng thu được có một số tính chất đặc biệt mà nhờ đó nó trở thành phương pháp hồi quy mạnh và phổ biến nhất.

Nội dung phương pháp OLS

Giả sử $Y_i = \beta_1 + \beta_2 X_i$ là PRF cần tìm. Ta tìm cách ước lượng nó bằng cách xây dựng SRF dạng

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$$

từ một mẫu gồm n quan sát (X_i, Y_i) , với $i = 1, 2, \dots, n$.

Khi đó, ứng với mỗi i, sai biệt giữa giá trị chính xác, Y_i , và giá trị ước lượng, $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$, là $e_i = Y_i - \hat{Y}_i = Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i$, mà ta gọi là các *phần dư*.

Phương pháp OLS nhằm xác định các tham số $(\hat{\beta}_1, \hat{\beta}_2)$ sao cho tổng bình phương các phần dư, $RSS = e_1^2 + e_2^2 + \dots + e_n^2 = \sum_{i=1}^n e_i^2$, là đạt nhỏ nhất.

Chú ý rằng tổng bình phương các phần dư này là hàm theo hai biến $(\hat{\beta}_1, \hat{\beta}_2)$,

$$RSS = \sum_{i=1}^n (Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)^2$$

với đạo hàm riêng theo các biến

$$\frac{\partial RSS}{\partial \hat{\beta}_1}(\hat{\beta}_1, \hat{\beta}_2) = -\sum_{i=1}^n 2(Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)$$

$$\frac{\partial RSS}{\partial \hat{\beta}_2}(\hat{\beta}_1, \hat{\beta}_2) = -\sum_{i=1}^n 2X_i(Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_i)$$

Giá trị nhỏ nhất của RSS, nếu có, phải đạt tại điểm dừng của nó, nghĩa là

$$\begin{cases} \frac{\partial RSS}{\partial \hat{\beta}_1}(\hat{\beta}_1, \hat{\beta}_2) = 0 \\ \frac{\partial RSS}{\partial \hat{\beta}_2}(\hat{\beta}_1, \hat{\beta}_2) = 0 \end{cases}$$

Do đó, ta nhận được hệ phương trình

$$\begin{cases} n\hat{\beta}_1 + \hat{\beta}_2 \sum_{i=1}^n X_i = \sum_{i=1}^n Y_i \\ \hat{\beta}_1 \sum_{i=1}^n X_i + \hat{\beta}_2 \sum_{i=1}^n X_i^2 = \sum_{i=1}^n X_i Y_i \end{cases}$$

Giải hệ phương trình trên, ta được

$$\hat{\beta}_2 = \frac{n \sum_{i=1}^n X_i Y_i - \sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

và

$$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{X},$$

trong đó \bar{X} , \bar{Y} là các trung bình của mẫu X, Y.

Các giá trị $\hat{\beta}_1$ và $\hat{\beta}_2$ nhận được bằng các công thức trên được gọi là các *ước lượng bình phương nhỏ nhất* của β_1 và β_2 .

Ví dụ 3. Bảng sau cho số liệu về lãi suất ngân hàng (Y) và tỷ lệ lạm phát (X) trong năm 1988 ở 9 nước

X	7,2	4,0	3,1	1,6	4,8	51,0	2,0	6,6	4,4
Y	11,9	9,4	7,5	4,0	11,3	66,3	2,2	10,3	7,6

Giả sử ra rằng X và Y có quan hệ tuyến tính. Hãy ước lượng hàm hồi quy mẫu.

Giải. Lập bảng ta tính được một số đại lượng sau :

X_i	Y_i	$X_i Y_i$	X_i^2
7,2	11,9	85,68	51,84
4	9,4	37,6	16
3,1	7,5	23,25	9,61
1,6	4	6,4	2,56
4,8	11,3	54,24	23,04
51	66,3	3381,3	2601
2	2,2	4,4	4
6,6	10,3	67,98	43,56
4,4	7,6	33,44	19,36
$\sum X_i = 84,7$	$\sum Y_i = 130,5$	$\sum X_i Y_i = 3694,29$	$\sum X_i^2 = 2770,97$

Bảng 7.3. Kết quả tính toán các tổng.

Thay vào hệ phương trình trên, ta có

$$\begin{cases} 9\hat{\beta}_1 + 84,7\hat{\beta}_2 = 130,5 \\ 84,7\hat{\beta}_1 + 2770,97\hat{\beta}_2 = 3694,29 \end{cases} \Leftrightarrow \begin{cases} \hat{\beta}_1 = 2,74169 \\ \hat{\beta}_2 = 1,24941 \end{cases}$$

và từ đó suy ra hệ số hồi quy

Ta nhận được SRF: $\hat{Y} = 2,74169 + 1,24941X$.

7.2.1.4. Các giả thiết của mô hình

Để có thể dùng các công cụ của thống kê toán nhằm đánh giá chất lượng của mô hình hồi quy tuyến tính, ta cần các giả thiết sau trên biến số X và đại lượng sai số ngẫu nhiên ε

Giả thiết 1. Biến giải thích X là biến phi ngẫu nhiên, nghĩa là các giá trị của nó được hoàn toàn xác định. Giả thiết này đương nhiên được thỏa trong mô hình hồi quy tuyến tính.

Theo giả thiết này thì các giá trị của biến độc lập X là những giá trị có sẵn. Ví dụ khi tìm mối quan hệ giữa thu nhập và chi tiêu thì các mức thu nhập được xác định từ trước để xem chi tiêu thay đổi như thế nào tương ứng với mỗi mức thu nhập sẵn có này.

Giả thiết 2. Kỳ vọng của sai số ngẫu nhiên ε bằng 0, nghĩa là $E(\varepsilon_i) = E(\varepsilon_i | X = X_i) = 0$. Giả thiết này có nghĩa là các yếu tố không xuất hiện trong mô hình, được đại diện bởi đại lượng ngẫu nhiên ε , không ảnh hưởng một cách có hệ thống đến giá trị trung bình của Y.

Với mỗi một giá trị cho trước của biến độc lập X_i có một tập hợp giá trị của biến phụ thuộc Y xoay quanh giá trị trung bình, trong đó có một số giá trị nằm trên đường trung bình và các giá trị khác nằm dưới đường trung bình tạo ra các sai số ε_i . Giả thiết 2 giả định rằng giá trị kỳ vọng của các sai số này bằng 0.

Nói cách khác, tác động dương của các sai số triệt tiêu tác động âm của những sai số khác làm cho tổng ảnh hưởng của sai số đến biến phụ thuộc là bằng 0.

Giả thiết 3. Phương sai của sai số không đổi (phương sai thuần nhất), nghĩa là $\text{Var}(\varepsilon_i) = \text{Var}(\varepsilon_j) = \sigma^2$, với mọi i, j. Giả thiết này có nghĩa là các giá trị cụ thể của Y tương ứng với cùng một giá trị của X đều được phân bố xung quanh giá trị trung bình của nó với cùng một mức độ phân tán như nhau.

Giả thiết 4. Không có tương quan giữa các sai số, nghĩa là $\text{Cov}(\varepsilon_i, \varepsilon_j) = 0$, với mọi i, j. Giả thiết này có nghĩa là sai số ở số liệu quan sát này không ảnh hưởng gì tới sai số ở số liệu quan sát khác.

Giả thiết 5. Sai số và biến giải thích không có tương quan, nghĩa là $\text{Cov}(\varepsilon, X) = 0$. Giả thiết này là cần thiết vì nó cho phép tách bạch ảnh hưởng của X và của các yếu tố không

xuất hiện trong mô hình đến các giá trị của Y . Giả thiết này đương nhiên được thỏa do X là phi ngẫu nhiên.

Giả thiết 6: ε_i tuân theo phân phối chuẩn với trị bình quân của ε_i bằng 0 và phương sai σ^2 hay $\varepsilon_i \sim N(0, \sigma^2)$.

7.2.1.5. Tính chất cho hàm hồi quy mẫu

Hàm hồi quy mẫu $\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X$ có các tính chất

i) SRF đi qua trung bình mẫu (\bar{X}, \bar{Y}) , nghĩa là $\bar{Y} = \hat{\beta}_1 + \hat{\beta}_2 \bar{X}$.

ii) Giá trị trung bình của \hat{Y} bằng với giá trị trung bình quan sát Y , nghĩa là $\overline{\hat{Y}} = \bar{Y}$.

iii) Giá trị trung bình của phần dư e bằng 0, nghĩa là $\bar{e} = \frac{1}{n} \sum_{i=1}^n e_i = 0$.

iv) Phần dư e và \hat{Y} không tương quan, nghĩa là $\sum_{i=1}^n e_i \hat{Y}_i = 0$.

iv) Phần dư e và X không tương quan, nghĩa là $\sum_{i=1}^n e_i X_i = 0$.

7.2.1.6. Hệ số xác định mô hình

Gọi TSS (Total Sum of Squares) là tổng bình phương các sai số giữa giá trị quan sát Y_i với giá trị trung bình của chúng,

$$TSS = \sum_{i=1}^n (Y_i - \bar{Y})^2.$$

ESS (Explained Sum of Squares) là tổng bình phương các sai lệch giữa giá trị của Y tính theo hàm hồi quy mẫu với giá trị trung bình,

$$ESS = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 = \hat{\beta}_2^2 \sum_{i=1}^n (X_i - \bar{X})^2.$$

RSS (Residual Sum of Squares) là tổng bình phương các sai lệch giữa giá trị quan sát của Y và các giá trị nhận được từ hàm hồi quy mẫu,

$$RSS = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2.$$

Khi đó,

TSS đo độ chính xác của số liệu thống kê,

ESS đo độ chính xác của hàm hồi quy so với trung bình, và

RSS đo độ chính xác của hàm hồi quy mẫu so với hàm hồi quy tổng thể.

Nhận xét rằng

$$TSS = ESS + RSS.$$

Nếu các số liệu quan sát của Y đều nằm trên SRF thì $RSS = 0$ và $ESS = TSS$.

Nếu hàm hồi quy mẫu phù hợp tốt với các số liệu quan sát thì ESS sẽ lớn so với RSS.

Ngược lại, nếu hàm hồi quy mẫu kém phù hợp với các số liệu quan sát thì RSS sẽ lớn so với ESS.

Với các nhận xét trên, ta dùng *hệ số xác định* (coefficient of determination)

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

để đo mức độ phù hợp của hàm hồi quy.

Ta có $0 \leq R^2 \leq 1$ và $R^2 = (r_{X,Y})^2$, với $r_{X,Y}$ là hệ số tương quan giữa X và Y. Do đó,

Khi $R^2 = 1$, ta được đường hồi quy “hoàn hảo”, mọi sai lệch của Y (so với trung bình) đều giải thích được bởi mô hình hồi quy.

Khi $R^2 = 0$, X và Y không có quan hệ tuyến tính.

7.2.1.7. Khoảng tin cậy và kiểm định giả thuyết

Khoảng tin cậy cho β_j .

Để ước lượng hệ số hồi quy tổng thể β_j , ta dùng thống kê sau

$$T = \frac{\hat{\beta}_j - \beta_j}{\text{se}(\hat{\beta}_j)} \sim \text{St}(n-2), \quad j=1,2$$

Do $\hat{\beta}_j$ và $\text{se}(\hat{\beta}_j)$ tính được từ mẫu nên với độ tin cậy γ cho trước, ta suy ra được mức ý nghĩa $\alpha=1-\gamma$, tra trong bảng phân phối Student ta có giá trị tới hạn $C=t_{\alpha/2}^{n-2}$. Từ khoảng tin cậy của T, ta suy ra được khoảng tin cậy cho β_j là

$$\beta_j \in \left[\hat{\beta}_j - C \text{se}(\hat{\beta}_j); \hat{\beta}_j + C \text{se}(\hat{\beta}_j) \right], \quad j=1,2.$$

So sánh β_2 với β_2^0 cho trước.

Ta có bài toán kiểm định

$$\begin{cases} H_0 : \beta_2 = \beta_2^0 \\ H_1 : \beta_2 \neq \beta_2^0 \end{cases}$$

Nếu giả thuyết H_0 đúng, ta có thống kê sau

$$T = \frac{\hat{\beta}_2 - \beta_2^0}{\text{se}(\hat{\beta}_2)} \sim \text{St}(n-2)$$

Với nguy cơ sai lầm α cho trước, tra trong bảng phân phối Student, ta có giá trị tới hạn

$$C = t_{\frac{\alpha}{2}}^{n-2}.$$

Nếu $|T| > C$, bác bỏ giả thuyết H_0 .

Nếu $|T| \leq C$, chưa đủ cơ sở bác bỏ H_0 nên ta chấp nhận H_0 .

Đặc biệt, với $\beta_2^0 = 0$, giả thiết H_0 có nghĩa là “*biến độc lập X không ảnh hưởng gì tới biến phụ thuộc Y*”. Khi đó, thống kê tương ứng là

$$T = \frac{\hat{\beta}_2}{\text{se}(\hat{\beta}_2)} \sim \text{St}(n-2).$$

Chẳng hạn, với số liệu của ví dụ 3, ta có $T = \frac{\hat{\beta}_2}{\text{se}(\hat{\beta}_2)} = \frac{1,249407}{0,0388} \approx 32,2$.

Với nguy cơ sai lầm $\alpha = 0,05$, tra bảng phân phối Student với $n-2 = 9-2 = 7$ độ tự do, ta được giá trị $C = 2,365$. Vì $|T| > C$ nên ta bác bỏ H_0 , chấp nhận H_1 và ta kết luận rằng *tỷ lệ lạm phát có ảnh hưởng đến lãi suất ngân hàng*, với nguy cơ sai lầm $\alpha = 0,05$.

7.2.1.8. Kiểm định sự phù hợp của mô hình

Trong phần này, ta khảo sát việc phân tích hồi quy theo quan điểm của phân tích phương sai. Việc phân tích này cung cấp cho ta một phương pháp hữu ích khác trong việc giải quyết vấn đề phán đoán thống kê.

Xuất phát từ tính chất

$$\text{TSS} = \text{ESS} + \text{RSS} \text{ và } R^2 = \frac{\text{ESS}}{\text{TSS}},$$

Từ thống kê này, ta cũng có thể kiểm định giả thiết

$$\begin{cases} H_0 : \beta_2 = 0 \\ H_1 : \beta_2 \neq 0 \end{cases}$$

nghĩa là kiểm định giả thiết cho rằng biến độc lập X không ảnh hưởng gì đến biến phụ thuộc Y. Chú ý rằng khi đó, nghĩa là khi (H_0) đúng, giá trị F còn có thể được tính bằng công thức

$$F = \frac{\hat{\beta}_2^2 n S_X^2}{\hat{\sigma}^2} = \frac{\frac{ESS}{1}}{\frac{RSS}{n-2}} = \frac{\frac{R^2 TSS}{1}}{\frac{(1-R^2)TSS}{n-2}} = \frac{(n-2)R^2}{1-R^2}.$$

Do đó, quá trình phân tích phương sai cho phép ta đưa ra các phán đoán thống kê về độ thích hợp của hàm hồi quy.

Chẳng hạn, với số liệu của ví dụ 3, ta có

$$F = \frac{(n-2)R^2}{1-R^2} = \frac{(9-2) \times 0,9933}{1-0,9933} \approx 1037,7761.$$

Với nguy cơ sai lầm $\alpha = 0,05$, tra bảng phân phối Fisher với độ tự do (1,7), ta được $C = f_{0,05}(1,7) = 5,59$. Vì $F > C$ nên ta bác bỏ H_0 , chấp nhận H_1 , nghĩa là biến độc lập X (tỷ lệ lạm phát) có ảnh hưởng đến biến phụ thuộc Y (lãi suất ngân hàng).

7.2.1.9. Ứng dụng phân tích hồi quy vào dự báo

Ta có thể dùng hàm hồi quy mẫu để dự báo giá trị của biến phụ thuộc Y tương ứng với một giá trị của biến độc lập X . Có hai loại dự báo

Dự báo giá trị trung bình của Y ứng với $X = X_0$.

Dự báo giá trị cá biệt của Y ứng với $X = X_0$.

Dự báo trung bình.

Với $X = X_0$, ta muốn dự báo giá trị trung bình của Y khi biết X lấy giá trị X_0 ,

$$E(Y | X = X_0) = \beta_1 + \beta_2 X_0.$$

Hàm hồi quy mẫu cho ta ước lượng điểm là $\hat{Y}_0 = \hat{\beta}_1 + \hat{\beta}_2 X_0$, trong đó \hat{Y}_0 là ước lượng điểm, không chệch và có phương sai nhỏ nhất của $E(Y | X = X_0)$. Tuy nhiên, \hat{Y}_0 vẫn sai khác so với giá trị thực của nó. Để có khoảng tin cậy cho $E(Y | X = X_0)$, ta chú ý rằng \hat{Y}_0 có phân phối chuẩn với trung bình $\beta_1 + \beta_2 X_0$ và phương sai

$$\text{Var}(\hat{Y}_0) = \sigma^2 \left[\frac{1}{n} + \frac{(X_0 - \bar{X})^2}{n S_X^2} \right].$$

Với mức ý nghĩa α , ta tìm được giá trị $C = t_{\alpha/2}^{n-2}$ trong bảng phân phối Student, suy ra ước lượng khoảng cho giá trị trung bình $E(Y | X = X_0)$

$$E(Y | X = X_0) \in \left[\hat{Y}_0 - C \text{se}(\hat{Y}_0); \hat{Y}_0 + C \text{se}(\hat{Y}_0) \right].$$

Dự báo giá trị riêng biệt Y_0 .

Nếu muốn dự báo giá trị riêng biệt Y_0 khi $X = X_0$, ta chú ý rằng $Y_0 - \hat{Y}_0$ là đại lượng ngẫu nhiên có phân phối chuẩn với trung bình là 0 và phương sai

$$\text{Var}(Y_0 - \hat{Y}_0) = \sigma^2 \left[1 + \frac{1}{n} + \frac{(X_0 - \bar{X})^2}{nS_X^2} \right].$$

Với mức ý nghĩa α , ta tìm được giá trị $C = t_{\frac{\alpha}{2}}^{n-2}$ trong bảng phân phối Student, suy ra ước lượng khoảng cho giá trị riêng biệt Y_0 .

$$Y_0 \in \left[\hat{Y}_0 - C \text{se}(Y_0 - \hat{Y}_0); \hat{Y}_0 + C \text{se}(Y_0 - \hat{Y}_0) \right].$$

7.2.2. Mô hình hồi quy tuyến tính k biến

7.2.2.1. Hàm hồi quy tổng thể

Xét hàm hồi quy tổng thể k biến dạng

$$E(Y | X_2, X_3, \dots, X_k) = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_k X_k,$$

Hay dạng ngẫu nhiên

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_k X_k + \varepsilon$$

7.2.2.2. Ước lượng tham số

Hàm hồi quy mẫu SRF có dạng

$$\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3 + \dots + \hat{\beta}_k X_k,$$

nghĩa là $Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2,i} + \hat{\beta}_3 X_{3,i} + \dots + \hat{\beta}_k X_{k,i} + e_i$,

Với phương pháp bình phương nhỏ nhất, các hệ số $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$ được chọn sao cho tổng bình phương của các phần dư (RSS) nhỏ nhất, nghĩa là

$$\text{RSS} \equiv \sum_{i=1}^n e_i^2 = \sum_{i=1}^n \left(Y_i - \hat{\beta}_1 - \hat{\beta}_2 X_{2,i} - \dots - \hat{\beta}_k X_{k,i} \right)^2$$

đạt giá trị nhỏ nhất.

Ví dụ 4. Số liệu quan sát của một mẫu cho ở bảng sau. Trong đó

Y : Lượng hàng bán được của một loại hàng hóa (tấn / tháng)

X_2 : Thu nhập của người tiêu dùng (triệu / năm)

X_3 : Giá bán của loại hàng này (ngàn đồng / kg)

Y	X_2	X_3	Y	X_2	X_3
---	-------	-------	---	-------	-------

20	8	2	17	6	5
18	7	3	16	5	6
19	8	4	15	5	7
18	8	4	13	4	8
17	6	5	12	3	8

Vậy hàm hồi quy cần tìm là

$$\hat{Y}_i = 14,99215 + 0,76178X_{2i} - 0,58901X_{3i}$$

7.2.2.3. Hệ số xác định hồi quy bội

Hệ số xác định hồi quy bội R^2 có thể tính bằng một trong hai công thức sau

$$R^2 = 1 - \frac{RSS}{TSS} = \frac{ESS}{TSS}$$

Trong đó

$$TSS = Y^T \cdot Y - n(\bar{Y})^2; \quad ESS = \hat{\beta}^T \cdot X^T \cdot Y - n(\bar{Y})^2;$$

$$RSS = TSS - ESS.$$

7.2.2.4. Ma trận tương quan

Giả sử $\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X_2 + \dots + \hat{\beta}_k X_k$. Ký hiệu $r_{t,j}$ chỉ hệ số tương quan giữa biến thứ t và biến thứ j . Với $t=1$ thì $r_{1,j}$ là hệ số tương quan giữa biến Y và biến X_j .

$$r_{1,j} = \frac{\sum_{i=1}^n y_i x_{i,j}}{\sqrt{\sum_{i=1}^n y_i^2 \sum_{i=1}^n x_{j,i}^2}}, \quad r_{t,j} = \frac{\sum_{i=1}^n x_{t,i} x_{j,i}}{\sqrt{\sum_{i=1}^n x_{t,i}^2 \sum_{i=1}^n x_{j,i}^2}},$$

trong đó: $x_{j,i} = X_{j,i} - \bar{X}_j$.

Dễ thấy rằng $r_{t,j} = r_{j,t}$ và $r_{j,j} = 1$ và ma trận hệ số tương quan có dạng

$$R = \begin{pmatrix} 1 & r_{1,2} & \dots & r_{1,k} \\ r_{2,1} & 1 & \dots & r_{2,k} \\ \dots & \dots & \dots & \dots \\ r_{k,1} & r_{k,2} & \dots & 1 \end{pmatrix}.$$

7.2.2.5. Ma trận hiệp phương sai

Để kiểm định giả thiết, tìm khoảng tin cậy cũng như thực hiện các suy đoán thống kê khác, ta cần phải tìm $\text{var}(\hat{\beta}_j)$, với $j=1,2,\dots,k$ và $\text{cov}(\hat{\beta}_i, \hat{\beta}_j)$. Phương pháp ma trận giúp ta có thể thực hiện điều này.

Ma trận hiệp phương sai của $\hat{\beta}$ có dạng tổng quát

$$\text{cov}(\hat{\beta}) = \begin{pmatrix} \text{cov}(\hat{\beta}_1, \hat{\beta}_1) & \text{cov}(\hat{\beta}_1, \hat{\beta}_2) & \dots & \text{cov}(\hat{\beta}_1, \hat{\beta}_k) \\ \text{cov}(\hat{\beta}_2, \hat{\beta}_1) & \text{cov}(\hat{\beta}_2, \hat{\beta}_2) & \dots & \text{cov}(\hat{\beta}_2, \hat{\beta}_k) \\ \dots & \dots & \dots & \dots \\ \text{cov}(\hat{\beta}_k, \hat{\beta}_1) & \text{cov}(\hat{\beta}_k, \hat{\beta}_2) & \dots & \text{cov}(\hat{\beta}_k, \hat{\beta}_k) \end{pmatrix}$$

7.2.2.6. Khoảng tin cậy và kiểm định giả thiết

Khoảng tin cậy cho β_j .

Để ước lượng hệ số hồi quy tổng thể β_j , ta dùng thống kê sau

$$T = \frac{\hat{\beta}_j - \beta_j}{\text{se}(\hat{\beta}_j)} \sim \text{St}(n-2), \quad j=1,2,\dots,k$$

Do $\hat{\beta}_j$ và $\text{se}(\hat{\beta}_j)$ tính được từ mẫu nên với độ tin cậy γ cho trước, ta suy ra được mức ý nghĩa $\alpha=1-\gamma$, tra trong bảng phân phối Student ta có giá trị tới hạn $C=t_{\alpha/2}^{n-k}$. Từ khoảng tin cậy của T, ta suy ra được khoảng tin cậy cho β_j là

$$\beta_j \in \left[\hat{\beta}_j - C \text{se}(\hat{\beta}_j); \hat{\beta}_j + C \text{se}(\hat{\beta}_j) \right], \quad j=1,2,\dots,k.$$

Kiểm định giả thiết về các hệ số hồi quy.

Xuất phát từ thống kê

$$T = \frac{\hat{\beta}_j - \beta_j}{\text{se}(\hat{\beta}_j)} \sim \text{St}(n-k).$$

ta có thể kiểm định giả thiết

$$\begin{cases} H_0 : \beta_j = \beta_j^0 \\ H_1 : \beta_j \neq \beta_j^0 \end{cases}$$

với β_j^0 cho trước.

Từ đó, ta dễ dàng kiểm định giả thiết về hệ số hồi quy riêng.

Đặc biệt, ta có thể kiểm định giả thiết

$$H_0 : \beta_2 = \beta_3 = \dots = \beta_k = 0.$$

Khi đó, nghịch thuyết là H_1 : tồn tại ít nhất một hệ số hồi quy riêng khác 0. Dùng đại lượng ngẫu nhiên

$$F = \frac{\frac{ESS}{k-1}}{\frac{RSS}{n-k}} \sim F(k-1; n-k).$$

hay

$$F = \frac{(n-k)R^2}{(k-1)(1-R^2)} \sim F(k-1; n-k).$$

7.2.2.7. Dự báo

Cho X^0 , với $(X^0)^T = (1 \quad X_2^0 \quad \dots \quad X_k^0)$, ta cần dự báo giá trị trung bình của biến phụ thuộc Y , nghĩa là dự báo $E(Y|X = X^0) = \beta_1 + \beta_2 X_2^0 + \dots + \beta_k X_k^0$.

Dự báo điểm (ước lượng điểm) của $E(Y|X = X^0)$ chính là \hat{Y}_0 ,

$$\hat{Y}_0 = \hat{\beta}_1 + \hat{\beta}_2 X_{2,i}^0 + \hat{\beta}_3 X_{3,i}^0 + \dots + \hat{\beta}_k X_{k,i}^0.$$

Ta tìm dự báo khoảng (ước lượng khoảng) của $E(Y|X = X^0)$ khi X_2 nhận giá trị X_2^0 , ..., X_k nhận giá trị X_k^0 .

Với $X = X^0$, ta có $\hat{Y}_0 = (X_0)^T \hat{\beta}$. Do đó

$$\text{var}(\hat{Y}_0) = \hat{\sigma}^2 (X_0)^T (X^T X)^{-1} X_0$$

Vì

$$\text{cov}(\hat{\beta}) = \sigma^2 (X^T X)^{-1}.$$

Do ta chưa biết σ^2 nên ta thay bằng ước lượng không chệch của nó là $\hat{\sigma}^2$, nghĩa là

$$\text{var}(\hat{Y}_0) = \hat{\sigma}^2 (X_0)^T (X^T X)^{-1} X_0 \text{ và } \text{se}(\hat{Y}_0) = \sqrt{\text{var}(\hat{Y}_0)}.$$

Từ đó, ta suy ra

$$T = \frac{E(Y|X = X^0) - \hat{Y}_0}{se(\hat{Y}_0)} \sim St(n - k)$$

và với độ tin cậy γ cho trước, ta suy ra khoảng tin cậy cần tìm.

Để tìm dự báo cá biệt của biến phụ thuộc Y khi $X = X^0$, nghĩa là tìm khoảng tin cậy cho Y_0 , ta dùng thống kê

$$T = \frac{Y_0 - \hat{Y}_0}{se(Y_0 - \hat{Y}_0)} \sim St(n - k),$$

trong đó

$$\text{var}(Y_0 - \hat{Y}_0) = \text{var}(\hat{Y}_0) + \sigma^2 \quad \text{và} \quad se(Y_0 - \hat{Y}_0) = \sqrt{\text{var}(Y_0 - \hat{Y}_0)}.$$

Với mức ý nghĩa α , ta tìm được giá trị $C = t_{\alpha/2}^{n-k}$ trong bảng phân phối Student, suy ra ước lượng khoảng cho giá trị trung bình $E(Y | X = X^0)$

$$E(Y | X = X^0) \in [\hat{Y}_0 - Cse(\hat{Y}_0); \hat{Y}_0 + Cse(\hat{Y}_0)].$$

và suy ra ước lượng khoảng cho giá trị riêng biệt Y_0 .

$$Y_0 \in [\hat{Y}_0 - Cse(Y_0 - \hat{Y}_0); \hat{Y}_0 + Cse(Y_0 - \hat{Y}_0)].$$

7.3. Câu hỏi thảo luận và bài tập

Câu hỏi thảo luận

Câu hỏi 1. Tại sao phải xây dựng mô hình hồi quy

Câu hỏi 2. Phân biệt mô hình hồi quy tổng thể; mô hình hồi quy tổng thể ngẫu nhiên; mô hình hồi quy mẫu và mô hình hồi quy mẫu ngẫu nhiên?

Câu hỏi 3. Hãy nêu những giả thiết của phương pháp ước lượng OLS?

Câu hỏi 4. Nếu sai số ngẫu nhiên không tuân theo luật phân phối chuẩn thì có thể ước lượng được các tham số trong mô hình hay không?

Câu hỏi 5. Giải thích tổng bình phương các biến động của biến phụ thuộc (TSS) là gì? ESS là gì? RSS là gì? Hệ số xác định mô hình?

Câu hỏi 6. Hãy nêu ý nghĩa của sai số ngẫu nhiên tổng thể.

Bài tập rèn luyện

Bài số 1. Một công ty nhà đất chuyên kinh doanh bất động sản. Trong nhiều năm nghiên cứu về giá nhà đất cho thấy giá nhà (Y : triệu đồng), phụ thuộc vào hai yếu tố là diện tích căn nhà (X_2 : m^2), vị trí của căn nhà (X_3 được cho điểm 1, 2, 3, 4, 5). Số liệu thu thập được cho trong bảng sau :

Y	X_2	X_3
1200	150	4
1100	175	3
380	180	1
620	58	4
410	82	2
390	158	1
1550	92	5
395	102	2
905	84	4
485	62	3

1. Tìm hàm hồi quy. Hãy giải thích ý nghĩa kinh tế của các hệ số nhận được.
2. Giải thích ý nghĩa các hệ số hồi quy.
3. Hãy tìm khoảng tin cậy cho các hệ số hồi quy với độ tin cậy 95%.
4. Kiểm định giả thuyết $H: R^2 = 0$, với mức ý nghĩa 1%.
5. Hãy dự báo giá bán trung bình cho một căn nhà có diện tích là $200m^2$ và vị trí là 3, với độ tin cậy 95%.
6. Với mức ý nghĩa 5%, hay cho biết vị trí (diện tích) có ảnh hưởng đến giá nhà hay không ?

Bài số 2. Bảng số liệu sau cho biết số lượng khách đi xe buýt (Y : ngàn người), giá vé xe buýt (X_2 : đồng/vé), và giá xăng hiện hành (X_3 : đồng/lít).

Y	X_2	X_3
22	2500	3400
43	1750	3550
26	2200	3500
79	1400	3600
121	1200	3850

227	1000	4000
16	2500	3000
25	2000	3450
35	2000	3400
63	2450	3550

1. Tìm hàm hồi quy. Hãy giải thích ý nghĩa kinh tế của các hệ số nhận được.
2. Giải thích ý nghĩa các hệ số hồi quy.
3. Hãy tìm khoảng tin cậy cho các hệ số hồi quy với độ tin cậy 95%.
4. Kiểm định giả thuyết $H: R^2 = 0$, với mức ý nghĩa 1%.
5. Hãy dự báo cho biết lượng khách đi xe buýt trung bình và cá biệt khi giá xe buýt là 3000 và giá xăng là 4500, với độ tin cậy 95%.
6. Với mức ý nghĩa 5%, hãy cho giá vé xe buýt (giá xăng hiện hành) có ảnh hưởng đến số lượng khách đi xe buýt hay không ?

Bài số 3. Số liệu về sản lượng Y (tạ), phân hóa học X_2 (tạ), thuốc trừ sâu X_3 (lít), tính trên một đơn vị diện tích ha, cho trong bảng sau

Y	40	44	46	48	52	58	60	69	74	80
X_2	6	10	12	14	16	18	22	24	26	32
X_3	4	4	5	7	9	12	14	20	21	24

Ước lượng mô hình hồi quy tuyến tính theo tham số của Y phụ thuộc vào X_2 và X_3 .

1. Kết quả ước lượng có phù hợp với thực tế không ? Hãy giải thích ý nghĩa kinh tế của các hệ số nhận được.
2. Với mức ý nghĩa 5%, hãy cho biết phân bón (thuốc trừ sâu) có ảnh hưởng đến năng suất của loại cây trồng trên hay không ?
3. Hãy tìm khoảng tin cậy cho các hệ số hồi quy với độ tin cậy 95%.
4. Hãy giải thích ý nghĩa của hệ số xác định mô hình.
5. Cả phân bón lẫn thuốc trừ sâu đều không ảnh đến năng suất ? với mức ý nghĩa 5%.
6. Với độ tin cậy 95%. Hãy dự báo giá trị trung bình và giá trị cá biệt của sản lượng khi mức phân bón là 20 tạ và thuốc trừ sâu là 16 lít.

Thuật ngữ chính chương 7

Tiếng Anh	Tiếng Việt
Column vector	Vecto cột
Corollary	Hệ quả
Correlation matrix	Ma trận tương quan
Covariance matrix	Ma trận hiệp phương sai
Coefficient of determination	Hệ số xác định
Descriptive analysis	Thống kê mô tả
Definition	Định nghĩa
Determinant	Định thức
Determinant of square matrix	Định thức của ma trận vuông
Diagonal matrix	Ma trận chéo
Exercise	Bài tập
Example	Ví dụ
Element of matrix	Phần tử của ma trận
Exploratory factor analysis	Phân tích nhân tố khám phá
Identity matrix	Ma trận đơn vị
Inverse matrix	Ma trận nghịch đảo
Linear algebra	Đại số tuyến tính
Matrix	Ma trận
Maximum	Giá trị lớn nhất
Minimum	Giá trị nhỏ nhất
Mean	Giá trị trung bình
Standard deviation	Độ lệch chuẩn
Median	Giá trị trung vị
Mode	Giá trị yếu vị (Giá trị tin chắc)
Multi – collinerity	Đa cộng tuyến
Reliability analysis	Phân tích độ tin cậy thước đo
Observations	Số quan sát
Ordinary least square	Bình phương bé nhất

HƯỚNG DẪN VIẾT ĐỀ CƯƠNG NGHIÊN CỨU VÀ BÁO CÁO NGHIÊN CỨU

8.1. Tổng hợp kiến thức để viết đề cương

Đề cương nghiên cứu khoa học là tác phẩm đầu tay của nhà khoa học trên con đường nghiên cứu, nó trình bày nội dung vấn đề nghiên cứu và kế hoạch tổ chức thực hiện. Đề cương nghiên cứu là việc làm đầu tiên có tầm quan trọng nhất, cụ thể là:

- Giúp cho người nghiên cứu xây dựng được phương hướng, chương trình, kế hoạch hoạt động nhằm đạt kết quả trong quá trình nghiên cứu đề tài đã được xác định.

- Giúp cho người nghiên cứu tập dượt và rèn luyện tác phong, rèn luyện khoa học, nâng cao tinh thần trách nhiệm trong công việc nhằm hoàn thành công trình nghiên cứu đúng tiến độ, thời gian, tránh lối làm ăn tùy tiện.

- Giúp người hướng dẫn khoa học có cơ sở để dẫn dắt, điều chỉnh hoạt động của người nghiên cứu, đánh giá kết quả nghiên cứu của người nghiên cứu.

8.2. Hình thức và trình tự của đề cương nghiên cứu khoa học

Tên đề tài: Tên đề tài cần ngắn gọn, thể hiện được mục tiêu và nội dung chính của nghiên cứu dự định thực hiện.

1. Giới thiệu (mở đầu)
 - 1.1. Xác định vấn đề nghiên cứu
 - 1.2. Tính cấp thiết của đề tài
 - 1.3. Mục tiêu nghiên cứu
 - 1.3.1. Mục tiêu tổng quát
 - 1.3.2. Mục tiêu cụ thể
 - 1.4. Câu hỏi nghiên cứu
 - 1.5. Đối tượng và phạm vi nghiên cứu
 - 1.5.1. Đối tượng nghiên cứu
 - 1.5.2. Đối tượng khảo sát
 - 1.5.3. Phạm vi nghiên cứu
 - 1.6. Ý nghĩa và giới hạn của nghiên cứu

1.6.1. Ý nghĩa của nghiên cứu
1.6.2. Giới hạn của nghiên cứu
2. Tổng quan tài liệu và cơ sở lý thuyết
2.1. Khái niệm
2.2. Lý thuyết liên quan
2.3. Các nghiên cứu trong và ngoài nước liên quan đề tài nghiên cứu
3. Khung phân tích và giả thuyết nghiên cứu
3.1. Khung phân tích của nghiên cứu
3.2. Các giả thuyết nghiên cứu
4. Phương pháp nghiên cứu (thiết kế nghiên cứu)
4.1. Phương pháp chọn mẫu
4.2. Phương pháp thu thập dữ liệu
4.3. Công cụ phân tích dữ liệu
5. Kết cấu của đề tài
6. Tiến độ thực hiện
7. Tài liệu tham khảo

Bảng 8.1. Nội dung cơ bản của đề cương nghiên cứu khoa học.

8.3. Báo cáo kết quả nghiên cứu

8.3.1. Khái niệm: Báo cáo kết quả nghiên cứu là giai đoạn cuối của một quá trình nghiên cứu. Đây là giai đoạn quan trọng giúp cho người đọc hiểu đúng giá trị và sử dụng đúng kết quả nghiên cứu.

8.3.2. Các thành phần của một báo cáo

8.3.2.1. Phần dẫn nhập gồm: Trang bìa, authorization, tóm tắt, mục lục.

a. Trang bìa

- Tựa đề thường gồm 3 thành tố: các biến/yếu tố nghiên cứu, loại quan hệ giữa chúng và tổng thể nghiên cứu.

- Tên người/tổ chức thực hiện.

- Tên người/tổ chức được báo cáo.

- Ngày (tháng, năm).

b. Authorization: Trình bày việc giao nhiệm vụ nghiên cứu, tài trợ nghiên cứu, v.v.

c. Tóm tắt

- Bao gồm tất cả nội dung của báo cáo nhưng thật ngắn gọn, hoặc
- Chỉ tóm tắt kết quả, kết luận và kiến nghị.

d. Mục lục

- Nếu báo cáo dài trên 10 trang nên có mục lục.
- Trường hợp có nhiều hình, bảng, v.v. nên có mục lục hình/bảng riêng.

8.3.2.2. Phần giới thiệu

- Lý do chọn đề tài
- Vấn đề nghiên cứu
- Mục tiêu nghiên cứu
- Phạm vi nghiên cứu

8.3.2.3. Phần cơ sở lý thuyết mô hình nghiên cứu

- Chủ yếu dành cho báo cáo khoa học hoặc nghiên cứu cơ bản.
- Trong nghiên cứu ứng dụng đơn giản có thể bỏ qua.

8.3.2.4. Phần phương pháp nghiên cứu

- Đối với báo cáo quản lý: nên viết thành một mục trong phần giới thiệu, sau “phạm vi nghiên cứu”. Chi tiết hơn nên đưa vào phụ lục.

- Đối với báo cáo kỹ thuật: quan trọng, cần nêu rõ:

Mô tả bản chất của thiết kế nghiên cứu.

Cách lấy mẫu và cỡ mẫu.

Cách đo và thu thập dữ liệu

Cách xử lý và phân tích dữ liệu.

8.3.2.5. Phần kết quả nghiên cứu

- Đây là phần dài nhất của báo cáo.
- Nên sắp xếp kết quả theo mục tiêu nghiên cứu và nhu cầu thông tin.
- Các thông tin phải tổ chức, trình bày theo trình tự (flow) chặt chẽ & logic.
- Phân biệt giữa các sự kiện và diễn dịch, cần có các minh họa cho diễn dịch nhưng không lạm dụng.

8.3.2.6. Phần kết luận và kiến nghị

- Tóm tắt kết quả (sự kiện) và kết luận (diễn dịch)
- Liên hệ những kết quả tìm được với những nhu cầu thông tin, mục tiêu nghiên cứu.

- Có hai quan điểm về kiến nghị: Dựa theo kết quả thông tin, kinh nghiệm để đưa ra kiến nghị; Không nên đưa ra những kiến nghị chủ quan để làm lệch lạc cho người nhận thông tin.

- Các hạn chế: nêu rõ các hạn chế có thể gây ảnh hưởng đến giá trị của kết quả.

8.3.2.7. Phần phụ lục

- Dùng để trình bày *chi tiết hơn* các thông số thống kê, bảng biểu nhưng báo cáo chính không thật sự cần.

- Dùng *cung cấp thêm thông tin* khi người đọc cần tìm hiểu sâu vấn đề.

8.3.2.8. Phần tài liệu tham khảo: Liệt kê các tài liệu tham khảo theo các tiêu chuẩn đã quy định trước, theo các hình thức được sử dụng rộng rãi.

8.3.3. Các nguyên tắc khi viết báo cáo

8.3.3.1. Trước khi viết

a. Cần xác định rõ: Mục đích của báo cáo là gì?; Ai là người đọc?

Có những yêu cầu gì về nội dung/ hình thức?

b. Thiết kế dàn ý chi tiết : Dàn ý các đề mục, dàn ý các nội dung.

c. Chuẩn bị tài liệu tham khảo/ hỗ trợ

8.3.3.2. Trong khi viết

a. Trình bày rõ mục tiêu và nhu cầu thông tin

- Báo cáo có tác dụng truyền đạt thông tin đến người ra quyết định.

- Trình bày rõ ràng các mục tiêu, nhu cầu thông tin, kết quả, các vấn đề quản lý, đề xuất.

b. Tính khách quan

- Báo cáo phải trung thực với các kết quả đã tìm được.

- Các trường hợp liên quan đến nhận định, phán đoán chủ quan của người trả lời thì cần nêu rõ.

c. Văn phong

- Câu ngắn gọn, từ thông dụng, khách quan

- Chặt chẽ, logic, nhất quán về cấu trúc

- Tránh viết tắt

- Dùng thì hiện tại đối với nội dung, quá khứ đối với cách thực hiện

- Tận dụng bảng, hình, đồ thị để minh họa

- Thống nhất các ghi chú, tài liệu tham khảo.

8.3.3.3. Sau khi viết

- Hiệu đính về nội dung.
- Chú ý về hình thức trình bày.
- Đọc kỹ nhiều lần (bởi nhiều người) để kiểm tra sai sót về nội dung và hình thức.

In ấn, đóng bìa, tạo soft – copy,...

8.3.4. Thuyết trình kết quả

8.3.4.1. Cần xác định trước: Thời gian trình bày, mục đích của buổi thuyết trình, đối tượng người nghe.

8.3.4.2. Thiết kế dàn ý và nội dung: Phần mở đầu, kết quả và kết luận, kiến nghị.

8.3.4.3. Phương tiện hỗ trợ: Bảng, Bút, Flip charts, Handouts, Slides, Powerpoints, Minh họa, v.v...

8.4. Hướng dẫn phần tài liệu tham khảo

- Tôn trọng và đề cao giá trị sản phẩm trí tuệ của học giả hoặc nguồn được nêu trong trích dẫn.

- Chứng minh tác giả đã tham khảo và xem xét vấn đề nghiên cứu một cách nghiêm túc.

- Cho phép người đọc xác nhận tính đúng đắn của thông tin được trích dẫn.

- Tác giả chịu trách nhiệm về tính chính xác của các trích dẫn trong luận văn.

- Việc trích dẫn đầy đủ tài liệu tham khảo giúp tác giả tránh được lỗi đạo văn.

8.4.1. Ghi và sử dụng trích dẫn từ sách:

- Quy chuẩn chung danh mục tài liệu tham khảo

Cách viết: Họ tên tác giả (năm xuất bản), *tên sách*, nhà xuất bản, nơi xuất bản

- Ví dụ 1. Danh mục tài liệu tham khảo

➤ Đào Hữu Hồ (2001), *Xác suất thống kê*, NXB Đại học quốc gia Hà Nội.

➤ Trần Ngọc Phát, Trần Thị Kim Thu (2006), *Giáo trình Lý thuyết thống kê*, NXB Thống kê.

- Ví dụ 2: Trích dẫn tương ứng trong bài viết

➤ Ví dụ trích dẫn gián tiếp: Định nghĩa biến cố (Đào Hữu Hồ, 2001)

➤ Ví dụ trích dẫn trực tiếp: Đào Hữu Hồ (2001, 8) nêu rõ định nghĩa biến cố

8.4.2. Đối với tài liệu ngoài Việt Nam, nên ghi rõ tên thành phố, bang (nếu có) và tên nước

- Ví dụ 1. Danh mục tài liệu tham khảo
 - Krugman, P. (1995), *Development, Geography, and Economic Theory*, MIT Press, Cambridge, Massachusetts, USA.
- Ví dụ 2. Trích dẫn tương ứng trong bài viết
 - Krugman (2009, 19) nêu rõ yếu tố X có ảnh hưởng tới sản lượng quốc dân

8.4.3. Trích dẫn bài báo đăng trên tạp chí khoa học

- Quy chuẩn chung danh mục tài liệu tham khảo
 - Họ tên tác giả (năm xuất bản), ‘tên bài báo’, tên tạp chí, số..., tập phát hành, trang chứa nội dung bài báo trên tạp chí.
- Ví dụ 1. Danh mục tài liệu tham khảo
 - Lê Xuân X (2009), ‘Tổng quan kinh tế Việt Nam năm 2010’, *Tạp chí Y*, 15 (4), 7– 13.
- Ví dụ 2: Trích dẫn tương ứng trong bài viết
 - Lê Xuân X (2009) tin rằng...

8.5. Ví dụ về đề cương chi tiết

Ví dụ dưới đây trích từ đề cương chi tiết luận văn thạc sĩ của Cô Phùng Vũ Bảo Ngọc, Khoa Du Lịch trường Đại học Tài chính Marketing Thành phố Hồ Chí Minh.

Đề tài: “Ảnh hưởng của hình ảnh điểm đến tới lòng trung thành của khách du lịch: Nghiên cứu trường hợp điểm đến du lịch TP. HCM”

1. Lý do chọn đề tài nghiên cứu

Du lịch là ngành công nghiệp không khói mang lại những lợi ích vô cùng to lớn cho mỗi quốc gia. Du lịch đóng góp phần lớn vào doanh thu của đất nước, mang lại hàng triệu việc làm cho người dân. Đặc biệt, du lịch còn là một phương tiện quảng bá hình ảnh đất nước mạnh mẽ nhất.

Từ năm 2001, Du lịch Việt Nam đã được chính phủ quy hoạch, định hướng đầu tư và phát triển để trở thành ngành kinh tế mũi nhọn. Với tiềm năng sẵn có và sự hỗ trợ tích cực của chính phủ trong hơn 10 năm qua, Việt Nam ngày nay đã trở thành tâm điểm thu hút nhiều khách du lịch trong và ngoài nước đến tham quan và khám phá. Theo hội

đồng lý hành và du lịch thế giới (WTTC), trong 181 quốc gia, vùng lãnh thổ thì du lịch Việt Nam đứng thứ 47 trên thế giới về phát triển tổng thể, đứng thứ 54 vì những đóng góp cho nền kinh tế quốc gia và đứng thứ 12 về sự tăng trưởng dài hạn trong vòng 10 năm tới (CafeF, 2010). Riêng trong khu vực Đông Nam Á, du lịch Việt Nam đứng thứ 2 về tốc độ tăng trưởng, đứng thứ 4 về tỷ trọng đóng góp vào nền kinh tế quốc dân và thứ 5 về kết quả tuyệt đối. Điều này khẳng định sức hấp dẫn, vị thế của du lịch Việt Nam trên bản đồ du lịch khu vực và thế giới (Nguyễn Hằng, 2011).

Do đó, ngành du lịch Việt Nam đang đứng trước sự cạnh tranh quyết liệt về thu hút du khách hơn bao giờ hết. Không chỉ cạnh tranh trong nội bộ ngành hay giữa các địa phương du lịch, sự cạnh tranh còn diễn ra gay gắt giữa các công ty du lịch lý hành quốc tế và trong nước. Vì vậy, việc quan trọng nhất của các công ty du lịch trong nước là làm thế nào để thu hút khách đến và quan trọng hơn là quay lại điểm đến, đồng nghĩa với việc nâng cao lòng trung thành của khách du lịch đối với điểm đến du lịch trong giai đoạn hiện nay.

Đặc biệt, điểm đến Thành phố Hồ Chí Minh (TP. HCM) đang ngày càng phát triển, thu hút ngày càng nhiều lượt khách du lịch trong và ngoài nước. Theo báo cáo của UBND TP. HCM trong 6 tháng đầu năm 2014, lượng khách quốc tế đến Thành phố ước đạt 2,1 triệu lượt, tăng 10% so cùng kỳ năm trước, đạt 48% kế hoạch năm 2014 (4,4 triệu lượt khách). Tổng doanh thu du lịch (khách sạn, nhà hàng và du lịch lý hành) ước đạt 44.299 tỷ đồng, tăng 8,5% so cùng kỳ năm 2013, đạt 45% kế hoạch năm 2014 (99.000 tỷ đồng).

Tuy nhiên so với tiềm năng vốn có thì du lịch TP. HCM vẫn chưa phát triển đúng mức và hình ảnh TP. HCM trong mắt du khách đang ngày càng xấu đi. Nguyên nhân một phần là do những người làm du lịch vẫn chỉ quan tâm đến lợi ích ngắn hạn và chưa thật sự hiểu cảm nhận của khách làm ảnh hưởng không tốt đến đánh giá của khách đối với điểm đến TP. HCM cũng như du lịch Việt Nam. Tình trạng chật chém vẫn xảy ra. Một số hoạt động du lịch còn do tự phát, chưa chú trọng đến chất lượng dịch vụ, chưa định hướng phát triển bền vững, nên khách thường có xu hướng “một đi không trở lại”.

Bên cạnh đó, một số nghiên cứu nước ngoài cũng đã có đề cập rằng hình ảnh điểm đến có liên quan đến lòng trung thành của khách du lịch. Trong khi đó, TP Hồ Chí Minh

cùng với thủ đô Hà Nội, phố cổ Hội An và vịnh Hạ Long được đánh giá là những địa chỉ du lịch hấp dẫn nhất châu Á năm 2013 do du khách bình chọn trên website Trip Advisor. Nhưng với thực trạng hình ảnh TP.HCM không mấy đẹp đẽ trong mắt khách du lịch như hiện nay. Vậy cơ sở nào để nâng cao hình ảnh điểm đến TP. HCM? Mức độ ảnh hưởng của hình ảnh điểm đến tới lòng trung thành khách du lịch như thế nào? Nguồn lực chung là hạn chế, phải ưu tiên giải quyết thành phần nào của hình ảnh điểm đến nhằm gia tăng lòng trung thành du khách?

Từ những lý do trên, tác giả đã chọn đề tài:

“Ảnh hưởng của hình ảnh điểm đến tới lòng trung thành của khách du lịch: Nghiên cứu trường hợp điểm đến du lịch TP. HCM”.

Bài nghiên cứu này sử dụng mô hình định lượng để khám phá và khẳng định những yếu tố quan trọng của hình ảnh điểm đến tác động tới lòng trung thành của khách du lịch, từ đó đưa ra các giải pháp thu hút khách du lịch và nâng cao lòng trung thành của du khách đối với điểm đến Thành phố Hồ Chí Minh.

Đồng thời, tác giả hy vọng những thông tin về các yếu tố hình ảnh điểm đến tác động đến lòng trung thành của khách du lịch sẽ góp phần là tài liệu tham khảo cho các nhà lãnh đạo, các tổ chức du lịch và người làm du lịch tại TP.HCM.

Hơn thế nữa, người nghiên cứu hy vọng có thể tổng hợp về phương diện lý luận trong mảng du lịch TP.HCM từ đó làm cơ sở tham khảo cho các doanh nghiệp mong muốn gia nhập lĩnh vực du lịch hoặc muốn khai thác tiềm năng du lịch tại đây.

2. Tình hình nghiên cứu liên quan đến đề tài

Trên thế giới, hình ảnh điểm đến được nghiên cứu bắt đầu từ những năm 1970 và sau đó trở thành chủ đề được quan tâm phổ biến trong lĩnh vực du lịch. Những bài viết nghiên cứu về hình ảnh điểm đến ở nước ngoài khá nhiều nhưng nghiên cứu về mối quan hệ giữa hình ảnh điểm đến và lòng trung thành của khách du lịch hầu như rất hiếm.

Các nghiên cứu nước ngoài có đề cập đến mối quan hệ giữa hình ảnh điểm đến và lòng trung thành của khách du lịch có thể kể đến như:

+) Nghiên cứu của Girish Prayag và Chris Ryan (2011): Tiền đề lòng trung thành du khách đối với đảo Mauritius – Vai trò và ảnh hưởng của hình ảnh điểm đến, gắn kết địa điểm, sự tham gia cá nhân và sự hài lòng của khách du lịch.

+) Và nghiên cứu gần đây: Nghiên cứu của R. Rajesh (2013) về tác động của nhận thức du khách, hình ảnh điểm đến và sự hài lòng khách du lịch đối với lòng trung thành điểm đến.

Trong nước: Nghiên cứu của Hồ Huy Tựu và Nguyễn Xuân Thọ: Ảnh hưởng của hình ảnh điểm đến và cảm nhận rủi ro đến ý định quay lại và truyền miệng tích cực của du khách đối với khu du lịch biển Cửa Lò, tỉnh Nghệ An.

Nhìn chung, các công trình thường đề cập đến khái niệm hình ảnh điểm đến, khái niệm lòng trung thành mà chưa thể hiện rõ nét mối quan hệ giữa hình ảnh điểm đến và lòng trung thành của khách du lịch. Cụ thể là hiện tại ở Việt Nam chưa có công trình nào nghiên cứu trực tiếp đề tài này.

Chính vì thế mà đề tài này không trùng lặp và mang tính cấp thiết. Người nghiên cứu cho rằng cần thiết phải nghiên cứu về vấn đề này.

3. Mục tiêu nghiên cứu và câu hỏi nghiên cứu

3.1. Mục tiêu nghiên cứu:

+) Xác định các thành phần của hình ảnh điểm đến tác động tới lòng trung thành của khách du lịch và phát triển thang đo những yếu tố này.

+) Xây dựng và kiểm định mô hình lý thuyết về mối quan hệ giữa thành phần của hình ảnh điểm đến với lòng trung thành của khách du lịch – áp dụng cho trường hợp điểm đến TP.HCM, từ đó xác định cường độ tác động của các yếu tố này.

+) Đề xuất một số hàm ý rút ra từ kết quả nghiên cứu để xây dựng hình ảnh điểm đến TP. HCM tốt đẹp trong mắt du khách, trên cơ sở đó gia tăng lòng trung thành của khách du lịch đến với TP.HCM.

3.2. Câu hỏi nghiên cứu: Từ mục tiêu nghiên cứu trên, người nghiên cứu đặt ra câu hỏi nghiên cứu sau:

+) Hình ảnh điểm đến du lịch là gì? Gồm những thành phần nào? Áp dụng cho trường hợp TP.HCM thì những thành phần đó là gì?

+) Quan hệ giữa các thành phần của hình ảnh điểm đến với lòng trung thành của khách du lịch tại TP. HCM như thế nào?

+) Giải pháp nào giúp xây dựng và cải thiện hình ảnh điểm đến TP.HCM nhằm nâng cao lòng trung thành du khách đối với điểm đến này? (Cần làm gì để nâng cao tính hấp dẫn của hình ảnh điểm đến, từ đó gia tăng lòng trung thành của khách du lịch?)

4. Đối tượng và phạm vi nghiên cứu

4.1. Đối tượng nghiên cứu

Hình ảnh điểm đến, các thành phần của hình ảnh điểm đến, lòng trung thành của khách du lịch và mối quan hệ giữa chúng: lý thuyết và thực tiễn áp dụng tại TP.HCM.

4.2. Đối tượng khảo sát

Khách du lịch nội địa và khách du lịch quốc tế đã và đang du lịch TP.HCM.

4.3. Phạm vi nghiên cứu

+) Về không gian: Nghiên cứu được thực hiện tập trung tại TP.HCM. Cụ thể khảo sát tại các điểm du lịch, địa điểm ăn uống thường được khách du lịch lui tới.

+) Về thời gian: Các tài liệu được sử dụng trong bài là các tài liệu trong và ngoài nước trong khoảng thời gian 1900 – 2014. Thời gian thực hiện bài nghiên cứu từ tháng 01/2014 – 09/2014.

+) Về nội dung: Từ dữ liệu thứ cấp, bài nghiên cứu tổng hợp lý thuyết Hình ảnh điểm đến và Lòng trung thành của khách du lịch, tổng hợp tình hình ngành du lịch từ những báo cáo của Tổng cục du lịch, báo cáo của Ủy ban nhân dân TP.HCM. Dữ liệu sơ cấp được thu thập từ bảng khảo sát thực tế 300 khách du lịch tại TP.HCM.

5. Phương pháp nghiên cứu

5.1. Nghiên cứu định tính:

Nghiên cứu định tính được sử dụng trong nghiên cứu này là phương pháp thảo luận và phỏng vấn sâu. Phỏng vấn hai đối tượng, đối tượng thứ nhất là các chuyên gia du lịch, doanh nghiệp đang hoạt động trong lĩnh vực du lịch. Đối tượng thứ hai là một số khách du lịch đã và đang tham quan và sử dụng dịch vụ du lịch tại TP.HCM.

5.2. Nghiên cứu định lượng:

Thu thập dữ liệu nghiên cứu bằng bảng câu hỏi phỏng vấn trực tiếp và bảng câu hỏi qua thư điện tử đến các đối tượng mục tiêu.

Kiểm định độ tin cậy và giá trị của thang đo bằng hệ số tin cậy Cronbach alpha. Sử dụng phân tích nhân tố khám phá (EFA) thông qua phần mềm SPSS 20.0 để xây

dựng mô hình và các giả thuyết về các yếu tố hình ảnh điểm đến tác động đến lòng trung thành của khách du lịch đối với TP.HCM.

Dùng phương pháp phân tích hồi quy với các quan hệ tuyến tính để kiểm định các nhân tố hình ảnh điểm đến có ảnh hưởng quan trọng lòng trung thành của khách du lịch đối với TP.HCM, dựa vào đó tính mức độ quan trọng của từng yếu tố.

6. Ý nghĩa và những đóng góp mới của nghiên cứu

6.1. Ý nghĩa của nghiên cứu

Về phương diện học thuật:

+) Hệ thống hóa lý thuyết về hình ảnh điểm đến và lòng trung thành của khách du lịch đối với điểm đến.

+) Kiểm nghiệm mô hình của nghiên cứu trước, hoàn thiện mô hình các yếu tố hình ảnh điểm đến tác động đến lòng trung thành của khách du lịch đối với điểm đến TP.HCM.

Về phương diện thực tiễn:

+) Trong bối cảnh cạnh tranh ngày càng gay gắt giữa các doanh nghiệp du lịch, hình ảnh điểm đến đối với khách du lịch nói riêng được coi là nhân tố đóng vai trò quyết định trong việc thu hút và giữ chân du khách. Do vậy, kết quả nghiên cứu này là cơ sở khoa học cho việc hoạch định các chiến lược nâng cao hình ảnh điểm đến và các giải pháp nhằm thu hút và giữ chân khách du lịch tại TP.HCM.

+) Tổng hợp kết quả nghiên cứu nhằm cung cấp những thông tin giá trị cho việc ra quyết định của các nhà lãnh đạo góp phần giúp phát triển ngành du lịch thành phố mang tên Bác.

+) Làm cơ sở tham khảo cho các doanh nghiệp mong muốn gia nhập lĩnh vực du lịch.

Đặc biệt là các doanh nghiệp muốn phát triển dịch vụ du lịch đối với TP.HCM.

6.2. Đóng góp mới của nghiên cứu

+) Nghiên cứu lòng trung thành của nhân viên, của khách hàng là đề tài rất phổ biến nhưng trong lĩnh vực du lịch, nghiên cứu về ảnh hưởng của hình ảnh điểm đến tới lòng trung thành của khách du lịch còn rất ít và chưa cụ thể; đặc biệt ở Việt Nam, qua trải nghiệm của người nghiên cứu tìm tòi, hiện tại chưa có đề tài nào nghiên cứu về vấn đề này.

+) Đề tài tổng hợp một số lý thuyết về hình ảnh điểm đến và lòng trung thành của khách du lịch, đồng thời xây dựng mô hình các yếu tố hình ảnh điểm đến tác động đến lòng trung thành của khách du lịch từ đó phát triển mô hình tại điểm đến TP.HCM.

7. Kết cấu của báo cáo nghiên cứu

Để thực hiện mục tiêu nghiên cứu như trên, nội dung đề tài được thiết kế thành Phần mở đầu và 4 chương với nội dung chính của các chương được mô tả dưới đây:
Phần mở đầu: Trình bày lý do chọn đề tài, đối tượng và phạm vi nghiên cứu, phương pháp nghiên cứu, ý nghĩa nghiên cứu và bố cục của đề tài nghiên cứu.

Chương 1: Cơ sở lý thuyết và mô hình nghiên cứu: Trình bày cơ sở lý thuyết về khách du lịch, hình ảnh điểm đến và lòng trung thành của khách du lịch. Đặc điểm của điểm đến du lịch TP. HCM. Khám phá và giải thích các yếu tố hình ảnh điểm đến tác động đến lòng trung thành của khách du lịch từ những bài báo, nghiên cứu trước. Đề xuất thêm những yếu tố hình ảnh điểm đến tác động đến lòng trung thành của khách du lịch. Từ đó đặt ra giả thuyết nghiên cứu và đề xuất mô hình nghiên cứu.

Chương 2: Phương pháp nghiên cứu: Trình bày qui trình nghiên cứu, phương pháp nghiên cứu định tính và định lượng, đưa ra mô hình nghiên cứu điều chỉnh và xây dựng thang đo.

Chương 3: Phân tích dữ liệu và kết quả nghiên cứu: Trình bày kết quả nghiên cứu gồm kiểm định độ tin cậy của các thang đo các biến, phân tích nhân tố EFA, mô hình hồi qui đa biến và kiểm định các giả thuyết nghiên cứu.

Chương 4: Kết luận và hàm ý: Trình bày kết luận và giới hạn nghiên cứu. Đưa ra các hàm ý cho các nhà quản trị du lịch, doanh nghiệp, tổ chức du lịch và đề xuất hướng nghiên cứu sắp tới.

8. Tài liệu tham khảo

Tiếng Việt

[1] Cục thống kê TP.HCM, 2014. Báo cáo Tình hình kinh tế - xã hội tháng 6 và 6 tháng năm 2014. [online] có sẵn tại: <<http://www.pso.hochiminhcity.gov.vn/web/guest/nam-2014>> [Truy cập ngày 28.7.2014].

[2] Hoàng Trọng và Chu Nguyễn Mộng Ngọc, 2008. Phân tích dữ liệu nghiên cứu với SPSS (tập 1, 2). HCM: NXB Hồng Đức.

[3] Hồ Huy Tựu và Nguyễn Xuân Thọ, 2012. Ảnh hưởng của hình ảnh điểm đến và cảm nhận rủi ro đến ý định quay lại và truyền miệng tích cực của du khách đối với khu du lịch biển cửa Lò, tỉnh Nghệ An. Trường Đại học Nha Trang.

[4] Luật du lịch Việt Nam, 2005. Nhà xuất bản Chính trị Quốc gia, Hà Nội.

[5] Nguyễn Đình Thọ và Nguyễn Thị Mai Trang, 2011. Nghiên cứu thị trường. HCM: NXB Lao động.

[6] Nguyễn Văn Đỉnh và Trần Thị Minh Hòa, 2008. Giáo trình kinh tế du lịch. NXB Đại học Kinh tế quốc dân.

Tiếng Anh

[7] Assael, H., 1987. Consumer behaviour and marketing action. Boston: PWS Kent.

[8] Baloglu, S., 2001. Image variations of Turkey by familiarity index: informational and experiential dimensions. *Tourism Management*, Vol. 22 No. 2, pp. 127-133.

[9] Baloglu, S. and McCleary, K.W., 1999. A model of destination image formation. *Annals of Tourism Research*, Vol. 26 No. 4, pp. 868-897.

[10] Beerli, A. and Martin, J.D., 2004. Tourists' characteristics and the perceived image of tourist destinations: a quantitative analysis - a case study of Lanzarote, Spain. *Tourism Management*, Vol. 25, pp. 623-636.

[11] Bhatia. A. K., 2005. *Tourism development – Principles and Practices*. Sterling Publishers Pvt. Ltd. New Delhi.

[12] Boulding, K. E., 1956. *The Image: Knowledge in Life and Society*. The University of Michigan Press.

[13] Bowen, John T. & Chen, Shiang-Lih, 2001. The relationship between customer loyalty and customer satisfaction. *International Journal of Contemporary Hospitality Management*, Vol. 13 No. 5, pp. 213 – 217.

[14] Burkart, A. J. and Medlik, S., 1974. *Tourism: past, present and future*. Business and Economics.

[15] Crompton, J.L., 1977. A systems model of the tourist's destination selection decision process with particular reference to the role of image and perceived constraints. College Station: Texas A & M University. Unpublished Ph.D. Dissertation.

Bảng 8.2. Đề cương chi tiết luận văn thạc sĩ.

8.6. Câu hỏi thảo luận

Câu hỏi 1. Vì sao nói “ báo cáo không phải là báo cáo diễn giải quy trình và kết quả nghiên cứu”?

Câu hỏi 2. Báo cáo kết quả nghiên cứu là gì? Nêu các nguyên tắc khi viết báo cáo kết quả nghiên cứu?

Câu hỏi 3. Những lỗi mắc phải khi trình bày báo cáo nghiên cứu khoa học.

Câu hỏi 4. Lấy một đề tài nghiên cứu khoa học bất kỳ mà bạn biết hãy thử viết báo cáo cho đề tài này.

Câu hỏi 5. Trình bày trích dẫn theo tiêu chuẩn APA là gì?

Câu hỏi 6. Nêu một số tiêu chuẩn trích dẫn mà bạn biết.

Câu hỏi 7. Đề cương nghiên cứu là gì? Nêu một số đề cương nghiên cứu mà bạn biết.

Câu hỏi 8. Nếu cho rằng “Đề cương nghiên cứu là xương sống của nghiên cứu” theo bạn nói như vậy đúng hay sai? Nếu quan điểm của bạn về đề cương nghiên cứu?

Thuật ngữ chính chương 8

Tiếng Anh	Tiếng Việt
A timetable for completing the study	Tiến độ thực hiện
Appendix	Phụ lục
Abbreviations	Viết tắt
Conclusions	Kết luận
Cover pages	Trang bìa
Domestic and foreign research related	Các nghiên cứu trong và ngoài nước
Highlights of research findings	Những điểm nổi bật trong nghiên cứu
Introduction	Mở đầu/giới thiệu
List of table	Danh mục bảng
List of figures	Danh mục hình
Literature review	Tổng quan tài liệu và cơ sở lý thuyết
Rationale for the research	Tính cấp thiết
Relevant concepts	Khái niệm
Relevant theories	Lý thuyết liên quan
References	Tài liệu tham khảo
Research results	Kết quả nghiên cứu
Recommendations/Implications	Gợi ý chính sách/kiến nghị
Outline of draft chapters	Kết cấu luận văn
Scope of study	Phạm vi nghiên cứu
Statement of the problem	Xác định vấn đề nghiên cứu
Theoretical research	Nghiên cứu lý thuyết
Table of contents	Mục lục
Topic outline	Dàn ý các đề mục

Phụ lục 1. BẢNG CÂU HỎI KHẢO SÁT

1. Bảng câu hỏi khảo sát “200 hộ gia đình ở vùng nông thôn ĐBSCL năm 2016”

Xin kính chào quý ông (bà)

Nghiên cứu “Các nhân tố tác động đến thu nhập hộ gia đình” nhằm đánh giá thực trạng thu nhập của hộ gia đình nông thôn, từ đó đề ra một số giải pháp để có thể nâng cao thu nhập hộ gia đình. Rất mong ông (bà) dành chút thời gian để trả lời phỏng vấn hoặc điền câu trả lời vào bảng khảo sát dưới đây. Những thông tin này sẽ được hoàn toàn giữ kín.

Phần A. Thông tin tổng quát

Q1. Họ và tên chủ hộ:.....

Q2. Địa chỉ của hộ:.....

Q3. Trình độ văn hóa của chủ hộ:.....

Q4. Thành phần dân tộc của chủ hộ:.....

Q5. Số nhân khẩu của hộ:.....

Q6. Số người lao động chính của hộ trong 12 tháng qua:.....

Q7. Xin ông (bà) cho biết thông tin của từng thành viên trong hộ:

Số thứ tự	Họ và tên	Quan hệ với chủ hộ	Giới tính	Tuổi	Trình độ học vấn	Nghề nghiệp chính	Kinh nghiệm	Hình thức nghề nghiệp chính	Nghề nghiệp phụ
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1		1							
2									
3									
4									
5									
6									
7									
8									
9									
...									

Hộ hàng, bạn bè

Q10. Giá trị khoản vay là bao nhiêu?.....

Q11. Hộ gia đình ông (bà) vay các khoản tiền này để làm gì?

Đầu tư cho sản xuất

Xây, sửa chữa nhà

Tiêu dùng, sinh hoạt

Trả nợ

Chữa bệnh

Đi học

Khác (Xin vui lòng ghi rõ).....

Phần C. Thông tin diện tích đất canh tác

Q12. Gia đình ông (bà) có đất nông nghiệp, lâm nghiệp để trồng trọt, sản xuất không?

(Nếu có thì làm tiếp câu Q13, Q14; nếu không thì chuyển sang câu Q16).

Có.

Không.

Q13. Xin ông (bà) cho biết diện tích thửa đất là bao nhiêu? Thuộc loại đất nào?

(1) Diện tích đất trồng cây hàng năm:.....

(2) Diện tích đất trồng cây lâu năm:.....

(3) Diện tích đất lâm nghiệp:.....

(4) Diện tích đất dùng để chăn nuôi:.....

(5) Diện tích đất nuôi trồng thủy sản:.....

(6) Tổng diện tích: (1)+(2)+(3)+(4)+(5):.....

Q14. Gia đình ông (bà) có canh tác, sản xuất trên diện tích đất này không?

(Nếu có thì làm tiếp câu Q15; nếu không thì chuyển sang câu Q16).

Có.

Không.

Q15. Nếu không sản xuất thì ông (bà) cho thuê hay bỏ hoang?

Cho thuê.

Bỏ hoang.

Phần D. Thu nhập của gia đình

Q16. Tổng thu và tổng chi cho hoạt động sản xuất kinh doanh của hộ trong 12 tháng qua

(tính từ thời điểm điều tra về trước)

Nguồn thu	Tổng thu	Tổng chi
1. Trồng trọt		
Cây hàng năm		

Cây lâu năm		
Sản phẩm phụ trồng trọt (thân, lá, ngọn cây, rơm rạ,...)		
Sản phẩm trồng trọt khác (cây giống, cây cảnh,...)		
2. Chăn nuôi		
Gia súc		
Gia cầm		
Sản phẩm khác (trứng, con giống,..)		
Sản phẩm phụ chăn nuôi (phân, lông, da,...)		
3. Hoạt động dịch vụ nông nghiệp		
4. Lâm nghiệp và dịch vụ lâm nghiệp		
5. Thủy sản và dịch vụ thủy sản		
6. Các hoạt động sản xuất kinh doanh dịch vụ phi nông nghiệp		
7. Tiền lương, tiền công		
8. Các khoản khác (thu từ hái lượm, quà tặng, lãi tiết kiệm, tiền cho thuê nhà, lương hưu, trợ cấp,..)		
Tổng cộng		

Q17. Theo ông (bà), Nhà nước cần hỗ trợ nào để giúp hộ gia đình phát triển kinh tế, nâng cao thu nhập?

- Phát triển cơ sở hạ tầng
 Vốn
 Kỹ thuật
 Đất sản xuất
 Tiêu thụ sản phẩm, giá ổn định.

Xin chân thành cảm ơn quý ông (bà) đã tham gia trả lời câu hỏi. Chúc gia đình ông (bà) thật nhiều sức khỏe và hạnh phúc!

2. Bảng câu hỏi khảo sát “Ảnh hưởng của hình ảnh điểm đến tới lòng trung thành của khách du lịch: Nghiên cứu trường hợp điểm đến du lịch TP. HCM”

Chào Anh (Chị), Tôi là học viên cao học khóa 22 trường Đại học Kinh tế TP.HCM đang thực hiện đề tài luận văn cao học: “Ảnh hưởng của hình ảnh điểm đến tới lòng trung thành của khách du lịch: Nghiên cứu trường hợp điểm đến du lịch TP. HCM”

Trước tiên, xin cảm ơn Anh (Chị) đã nhận lời tham gia cuộc phỏng vấn của chúng tôi. Cũng xin lưu ý mọi thông tin trung thực do Anh (Chị) cung cấp không có quan điểm nào là đúng hay sai và tất cả đều rất có giá trị cho nghiên cứu của chúng tôi, vì thế rất mong nhận được sự hợp tác của Anh (Chị).

Tiếp theo, xin Anh (Chị) vui lòng cho biết mức độ đồng ý của Anh (Chị) về các phát biểu dưới đây với quy ước:

1	2	3	4	5
Rất không đồng ý	Không đồng ý	Trung lập	Đồng ý	Rất đồng ý

(Lưu ý: Mỗi phát biểu chỉ chọn 1 mức độ, Anh (Chị) đồng ý ở mức độ nào thì khoanh tròn vào mức độ tương ứng, trường hợp chọn nhầm xin Anh (Chị) gạch chéo X để hủy chọn và chọn lại).

Số TT	Các tiêu thức	Mức độ đồng ý				
		1	2	3	4	5
I. Môi trường cảnh quan						
1	TP. HCM có khí hậu, thời tiết dễ chịu quanh năm	1	2	3	4	5
2	TP. HCM có nhiều cảnh quan để tham quan, khám phá (sông, hồ, thác, biển, núi, vv)	1	2	3	4	5
3	Môi trường sống tại TP. HCM ít bị ô nhiễm khói bụi, tiếng ồn	1	2	3	4	5
4	TP. HCM có nhiều tòa nhà, công trình xây dựng hiện đại	1	2	3	4	5
5	TP. HCM có nhiều đường phố, đô thị sạch đẹp	1	2	3	4	5
II. Môi trường kinh tế, chính trị						
6	TP. HCM có tình hình chính trị ổn định	1	2	3	4	5
7	TP. HCM có các quy tắc, luật lệ rõ ràng	1	2	3	4	5
8	Tại TP. HCM ít xảy ra trộm cướp và tệ nạn xã hội	1	2	3	4	5
9	TP. HCM là trung tâm kinh tế phát triển, năng động	1	2	3	4	5
10	Anh/Chị cảm thấy TP. HCM an ninh và an toàn	1	2	3	4	5
III. Khả năng tiếp cận điểm đến						
11	Anh/Chị dễ dàng có được thông tin về điểm đến TP. HCM	1	2	3	4	5
12	Anh/Chị dễ dàng đến TP. HCM bằng các phương tiện vận chuyển khác nhau (ô tô, tàu thủy, tàu lửa, máy bay)	1	2	3	4	5
13	Dịch vụ đưa đón khách du lịch (xe ôm, taxi, xích lô) tại TP. HCM thuận tiện	1	2	3	4	5
14	Chuyến đi của anh/chị đến TP. HCM thuận lợi	1	2	3	4	5

15	Chi phí đi lại, vận chuyển hợp lý	1	2	3	4	5
IV. Lịch sử, văn hóa						
16	TP. HCM có các di tích lịch sử, di sản văn hóa đặc sắc	1	2	3	4	5
17	TP. HCM có các sự kiện, lễ hội văn hóa đa dạng và hấp dẫn	1	2	3	4	5
18	TP. HCM có nhiều bảo tàng lịch sử, văn hóa, nghệ thuật	1	2	3	4	5
19	Hoạt động của khách du lịch không bị hạn chế bởi rào cản ngôn ngữ	1	2	3	4	5
20	Người dân TP. HCM thân thiện và hiếu khách	1	2	3	4	5
V. Ẩm thực và mua sắm						
21	TP. HCM có nền ẩm thực phong phú, đa dạng	1	2	3	4	5
22	TP. HCM có nhiều món ăn độc đáo	1	2	3	4	5
23	Hoạt động mua sắm tại TP. HCM đa dạng, văn minh	1	2	3	4	5
24	Phong cách phục vụ tại TP. HCM chuyên nghiệp	1	2	3	4	5
25	Giá cả ăn uống, mua sắm phải chăng	1	2	3	4	5
VI. Cơ sở hạ tầng						
26	TP. HCM có hệ thống giao thông phát triển	1	2	3	4	5
27	Dịch vụ y tế, chăm sóc sức khỏe tại TP. HCM phát triển	1	2	3	4	5
28	Mạng lưới thông tin, viễn thông tại TP. HCM phát triển	1	2	3	4	5
29	Hệ thống khách sạn, nhà hàng TP. HCM đa dạng và có chất lượng	1	2	3	4	5
30	Hệ thống trung tâm lễ hành, công tu du lịch tại TP. HCM phát triển	1	2	3	4	5
VII. Giải trí						
31	TP. HCM có nhiều địa danh để thăm thú, tìm cảm giác lạ	1	2	3	4	5
32	TP. HCM có cuộc sống, hoạt động về đêm đa dạng (uống café, đi bar, coi kịch, nghe ca nhạc, xem phim, vv)	1	2	3	4	5
33	TP. HCM có các hoạt động ngoài trời, thể thao phong phú	1	2	3	4	5
34	TP. HCM có hoạt động vui chơi, tiêu khiển đặc sắc	1	2	3	4	5
VIII. Lòng trung thành của khách du lịch						

35	Anh/Chị ưu tiên lựa chọn điểm đến TP. HCM khi có cơ hội đi du lịch	1	2	3	4	5
36	Anh/Chị sẽ trở lại TP. HCM khi có cơ hội	1	2	3	4	5
37	Anh/Chị sẽ nỗ lực trở lại TP. HCM trong một ngày không xa	1	2	3	4	5
38	Anh/Chị sẽ giới thiệu người thân/ bạn bè đến du lịch tại TP. HCM	1	2	3	4	5

IX. THÔNG TIN CHUYẾN ĐI

39) Anh/Chị đã du lịch đến TP. HCM bao nhiêu lần (bao gồm cả chuyến đi này)?

- 1 lần
- 2-3 lần
- 4-5 lần
- Trên 5 lần

40) Lý do anh/chị đi du lịch TP. HCM? (có thể chọn nhiều câu trả lời)

- Tham quan
- Nghỉ dưỡng
- Vui chơi, giải trí
- Kinh doanh/ đi công tác
- Thăm người thân/ bạn bè
- Dự hội nghị/ triển lãm
- Đám cưới/ tuần trăng mật
- Mua sắm
- Tham dự sự kiện (thể thao, văn hóa, nghệ thuật, vv)
- Trên đường đến nơi khác
- Khác: (ghi rõ)

41) Anh/Chị thường ở lại TP. HCM bao lâu?

- Dưới 1 ngày
- 1-2 ngày
- 3-5 ngày
- 1 tuần
- Trên 1 tuần

42) Anh/Chị du lịch đến TP. HCM bằng phương tiện nào? (có thể chọn nhiều câu trả lời)

- Xe máy
- Xe hơi
- Xe du lịch
- Taxi
- Tàu lửa
- Máy bay
- Khác: (ghi rõ)

43) Anh/Chị đi du lịch TP. HCM với ai?

- Một mình
- Vợ/ chồng
- Gia đình và con cái
- Bạn bè/ họ hàng
- Đối tác kinh doanh
- Đoàn du lịch

44) Anh/Chị biết đến TP. HCM thông qua kênh thông tin nào? (có thể chọn nhiều câu trả lời)

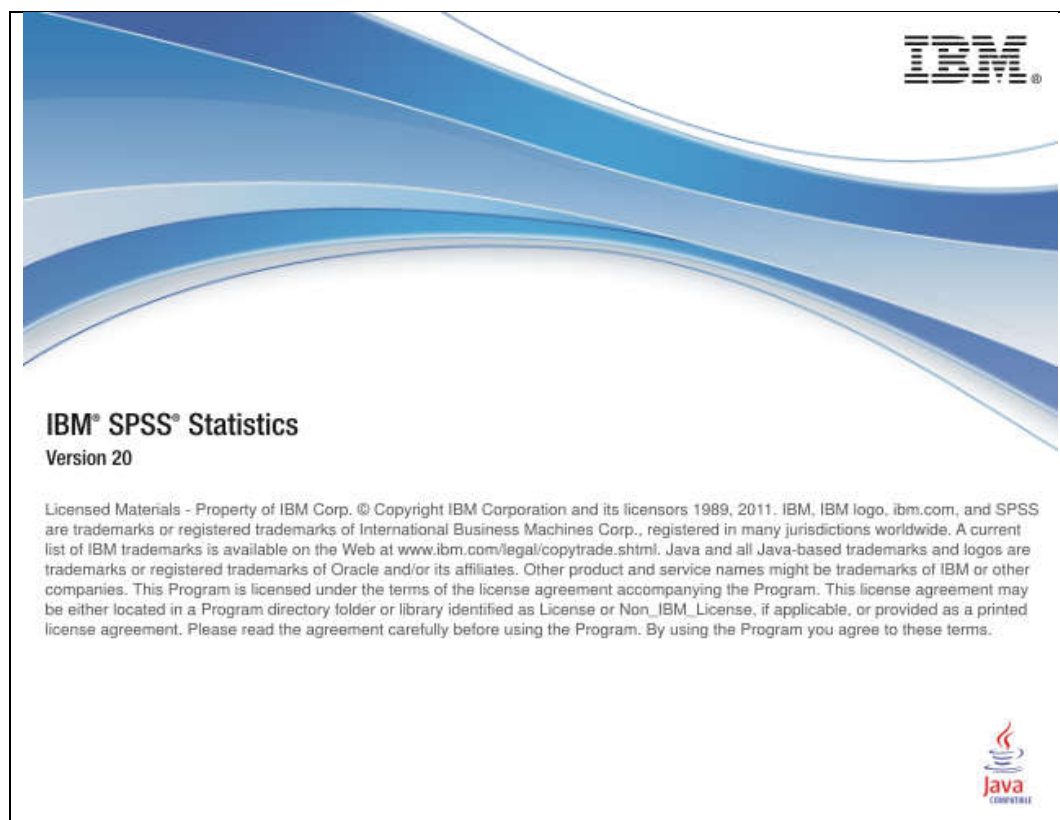
- Internet
- Đại lý du lịch
- Ấn phẩm quảng cáo (brochure) / sách hướng dẫn du lịch (travel guidebook)/ tạp chí
- Bạn bè/ người thân
- Quảng cáo truyền hình
- Khác: (ghi rõ)

45) Điểm du lịch nào tại TP. HCM mà anh/chị đã hoặc dự định đến?

(có thể chọn nhiều câu trả lời)

- | | |
|--|---|
| <input type="checkbox"/> Nhà hát lớn thành phố | <input type="checkbox"/> Khu du lịch Văn Thánh |
| <input type="checkbox"/> Thảo Cầm Viên | <input type="checkbox"/> Khu du lịch Bình Quới Thanh Đa |
| <input type="checkbox"/> Chợ Bến Thành | <input type="checkbox"/> Công viên văn hóa Đầm Sen |
| <input type="checkbox"/> Dinh Độc Lập | <input type="checkbox"/> Khu du lịch Suối Tiên |
| <input type="checkbox"/> Trụ sở UBND TP | <input type="checkbox"/> Khu giải trí BCR quận 9 |
| <input type="checkbox"/> Nhà thờ Đức Bà | <input type="checkbox"/> Vườn Cò Thủ Đức |
| <input type="checkbox"/> Trung tâm mua sắm | <input type="checkbox"/> Cần Giờ |

Phụ lục 2. HƯỚNG DẪN SỬ DỤNG PHẦN MỀM SPSS CƠ BẢN



1. Giới thiệu và cài đặt phần mềm SPSS

1.1. Giới thiệu phần mềm SPSS

SPSS (viết tắt của Statistical Package for the Social Sciences) là một chương trình máy tính phục vụ công tác thống kê. SPSS được sử dụng rộng rãi trong công tác thống kê xã hội.

Thế hệ đầu tiên của SPSS được đưa ra từ năm 1968. Thế hệ mới nhất là thế hệ 18 được giới thiệu từ tháng 8 năm 2008, có cả phiên bản cho các hệ điều hành Microsoft Windows, Mac, và Linux / UNIX.

SPSS là một hệ thống phần mềm thống kê toàn diện được thiết kế để thực hiện tất cả các bước trong các phân tích thống kê từ những tính toán trong thống kê mô tả (liệt kê dữ liệu, lập bảng tần số, lập biểu đồ, tính các đại lượng thống kê mô tả,...) đến thực hiện những bài toán trong thống kê suy diễn (kiểm định, tương quan, hồi quy, ...)

1.2. Cài đặt SPSS

Cho đến nay, SPSS có rất nhiều phiên bản khác nhau, phiên bản mới nhất là phiên bản 25.0. Với những tính toán thống kê phổ biến ta có thể cài SPSS với một trong các

phiên bản 14.0, 16.0, 18.0 hay 20.0, ... Tất cả các phân tích trong tài liệu này được thực hiện trên phiên bản SPSS 20.0.

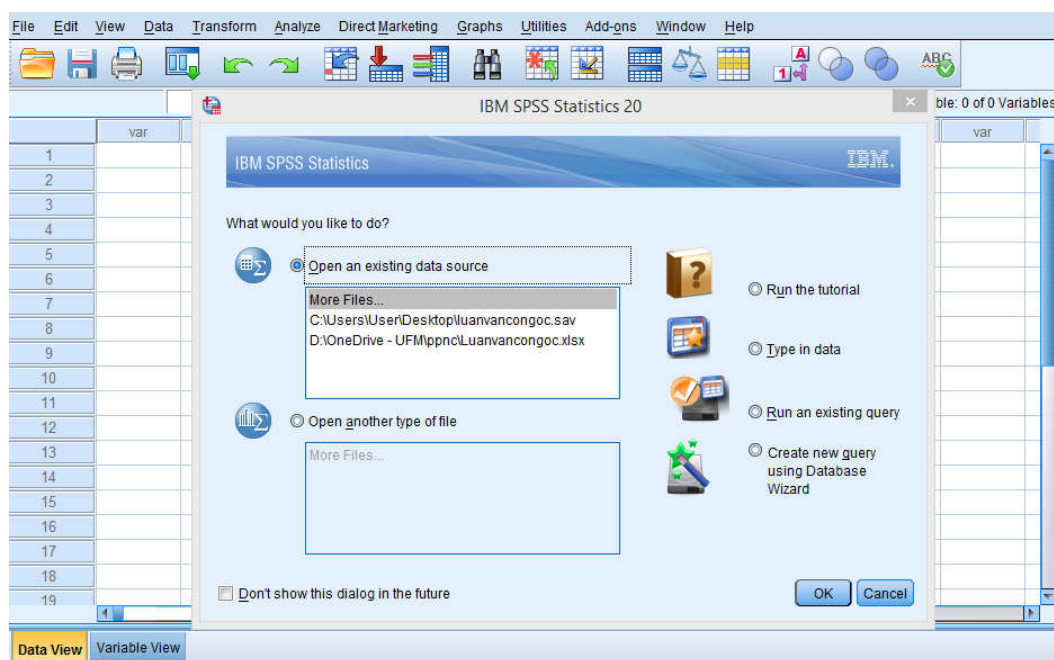
1.3. Khởi động SPSS

Để khởi động SPSS ta có thể thực hiện theo những cách sau:

+) Kích đúp vào biểu tượng SPSS trên màn hình;

+) Vào **Start** → **Program** → **SPSS for Windows** → **SPSS 20.0 for Windows**

Sau khi khởi động, ta được giao diện sau:



Hình 1

2. Nhập dữ liệu trong SPSS

2.1. Nhập trực tiếp vào SPSS

Để nhập trực tiếp một tập dữ liệu vào SPSS, ta thực hiện như sau:

Bước 1. Tại cửa sổ **IBM SPSS Statistics Data Editor**, ta ấn vào nút **Variable View** để khai báo thông tin về các thuộc tính của từng biến trong file dữ liệu.

Trong một bảng Variable View

+) Các hàng là tên các biến;

+) Các cột là tên các thuộc tính của biến.

Các thuộc tính của một biến bao gồm:

+) Tên biến (**Name**); Loại dữ liệu (**Type**)

+) Số lượng con số hoặc chữ (**Width**)

+) Số lượng chữ số thập phân (**Decimals**)

+) Mô tả biến/nhãn biến (**Label**) và nhãn trị số biến (**Values**)

+) Các giá trị khuyết thiếu do người sử dụng thiết lập (**Missing**)

+) Độ rộng của cột (**Columns**)

+) Căn lề (**Align**)

+) Thang đo của biến (**Measure**)

Khi khai báo hoặc chỉnh sửa các thuộc tính của biến trong cửa sổ **Variable View** cần chú ý một số điểm sau:

+) Tên biến phải bắt đầu bằng một chữ cái và không được kết thúc bằng một dấu chấm. Tên của biến là duy nhất, không được đặt trùng tên biến và tên biến không phân biệt chữ hoa, chữ thường.

+) Loại biến: Variable Type xác định loại dữ liệu đối với từng biến. Theo mặc định, mọi biến mới được giả sử là dạng số. Phụ thuộc vào loại dữ liệu được thu thập, ta có thể khai báo dưới những kiểu sau: dữ liệu là dạng số (**numeric**), dấu phẩy (**comma**), dấu chấm (**dot**), ghi chú khoa học (**Scientific notation**), ngày tháng (**Date**), đô-la (**Dollar**), đơn vị tiền riêng (**custom currency**) và chuỗi (**string**).

+) Nhãn của biến dùng để mô tả rõ hơn về tên của biến do tên của biến chỉ có độ dài tối đa là 8. Nhãn của biến có thể có độ dài đến 256 ký tự.

+) Ta có thể gán nhãn cho từng giá trị của biến. Tính năng này đặc biệt tiện lợi khi ta dùng các số để mã hóa các biến định tính. Ví dụ... Biến được mã hóa như vậy có thể dùng cho nhiều phân tích khác nhau. Hơn nữa, với nhãn của các giá trị, kết quả tính ra sẽ được trình bày rõ ràng hơn.

+) Các giá trị khuyết thiếu do người sử dụng thiết lập (**Missing**). Những giá trị không thích ứng với kiểu khai báo của biến sẽ được coi là giá trị khuyết thiếu. Đối với biến kiểu số, các ô trống được hiểu là giá trị khuyết và được đánh dấu bằng dấu phân cách thập phân. Nhiều thủ tục trong SPSS sẽ loại các giá trị khuyết ra khỏi các bước tính toán và các kết quả phân tích chỉ dựa trên phần số liệu không khuyết.

Ta có thể nhập đến 3 trị số khuyết riêng biệt, một phạm vi khoảng cách trị số khuyết hoặc một phạm vi cộng với một trị số khuyết riêng biệt;

Các phạm vi có thể được chỉ định cho các biến dạng số;

Các trị số khuyết cho các biến dạng chuỗi phải có độ dài không vượt quá 8 ký tự.

+) Số đo của biến có thể ở thang đo định danh (**Nominal**), thang đo thứ bậc (**Ordinal**) hoặc thang đo khoảng, tỉ lệ (**gọi chung là Scale**):

Thang đo định danh: Thang đo định danh dùng cho các biến định tính. Số đo của các biến này là các mã số để phân loại đối tượng. Giữa các mã số ở đây không có quan hệ

hơn kém, chỉ dùng để đếm tần số xuất hiện của các biểu hiện. Một số ví dụ về thang đo này là: biến giới tính với các số đo là: Nam hoặc Nữ; biến màu sắc với số đo là: xanh, đỏ, tím, vàng,...; biến khu vực sống với các số đo: Thành phố, Thị xã, Nông thôn, Miền núi,...

Thang đo thứ bậc: Thang đo thứ bậc thường dùng cho các biến định tính, đôi khi dùng cho cả biến định lượng. Trong thang đo này giữa các số đo của các biến có quan hệ thứ bậc hơn kém. Tuy nhiên, sự chênh lệch giữa các số đo không nhất thiết bằng nhau. Ví dụ biến đánh giá thái độ đối với chất lượng dịch vụ mạng Internet tại nhà có số đo là: Không hài lòng, hài lòng, rất hài lòng.

Thang đo khoảng: Thang đo thứ bậc thường dùng cho các biến định lượng. Thang đo khoảng là thang đo thứ bậc có các khoảng cách đều nhau. Các phép tính cộng trừ đều có nghĩa nhưng không có giá trị không xác định một cách chính xác và không thể lấy tỉ lệ giữa các số đo. Ví dụ số đo nhiệt độ, số đo chỉ số IQ, chỉ số EQ,...

Thang đo tỉ lệ: Thang đo tỉ lệ dùng cho các biến định lượng. Thang đo tỉ lệ là thang đo khoảng, hơn nữa thang đo này có giá trị không xác định một cách chính xác và có thể lấy tỉ lệ giữa các số đo. Ví dụ về thang đo này đơn vị đo tiền tệ (VND, dollar, pound, yen,...); đơn vị đo chiều dài (cm, m, km,...); đơn vị đo khối lượng (kg, tấn, tạ, yến,...).

Bước 2. Tại cửa sổ **Data Editor**, nhấn vào nút **Data View** để nhập từng dữ liệu trong mỗi ô. Trong bảng Data View

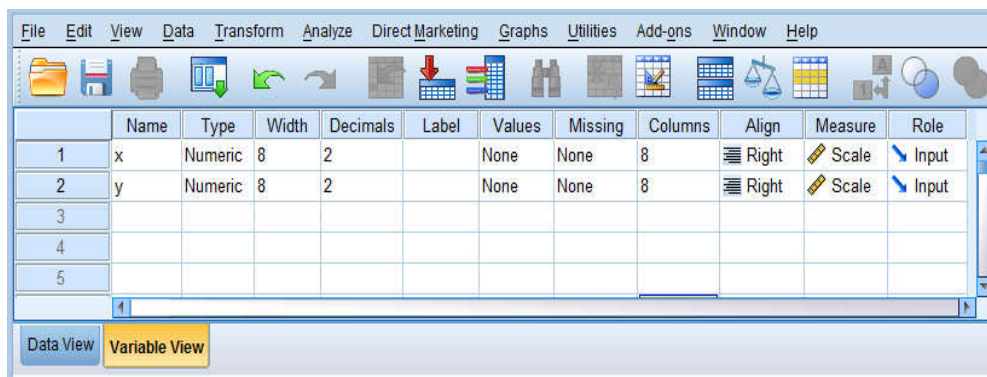
- +) Mỗi cột là mỗi biến **Variable**;
- +) Mỗi hàng là một đối tượng **cases**.

Ví dụ 1. Cho dữ liệu

X	6	10	12	14	16	18	22	24	26	32
Y	40	44	46	48	52	58	60	68	74	80

Bảng 1

Sau khi đã khai báo các thuộc tính của các biến trong cửa sổ Variable View



Hình 2

Ta vào cửa sổ Data View nhập giá trị cho từng biến như đã được điều tra.

	x	y	var	var	var	var
1	6.00	40.00				
2	10.00	44.00				
3	12.00	46.00				
4	14.00	48.00				
5	16.00	52.00				
6	18.00	58.00				
7	22.00	60.00				
8	24.00	68.00				
9	26.00	74.00				
10	32.00	80.00				

Hình 3

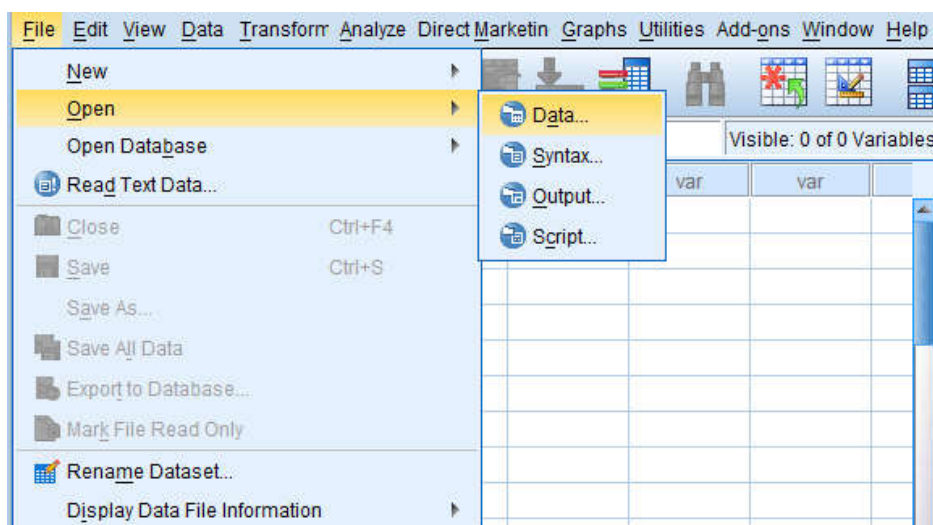
Để lưu dữ liệu vừa được tạo ra dưới đuôi của **SPSS.sav**, ta vào **File** → **Save as** để đánh tên file cần lưu và thư mục để lưu trong máy tính.

2.2. Đọc dữ liệu từ những file có sẵn

SPSS có thể đọc dữ liệu từ rất nhiều kiểu file khác nhau. Ngoài dữ liệu dưới dạng file.sav của SPSS, SPSS còn cho đọc file dữ liệu dưới một số dạng thông dụng như **file.xls** (Excel File), **file.txt** (Text File) và file dữ liệu tạo ra từ một số phần mềm thống kê khác như **file.dta** (Stata File), **file.wfl** (Eviews Workfile), ...

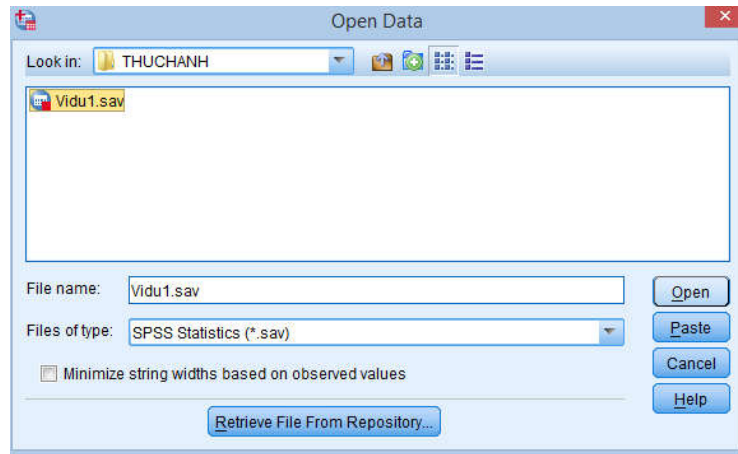
2.2.1. Đọc dữ liệu từ file.sav

Để đọc dữ liệu từ một file.sav, chẳng hạn file dữ liệu **Vidu1.sav**, ta vào **File** → **Open** → **Data** và chọn đến thư mục để file dữ liệu và mở file:



Hình 4

Kích đúp Data...



Hình 5

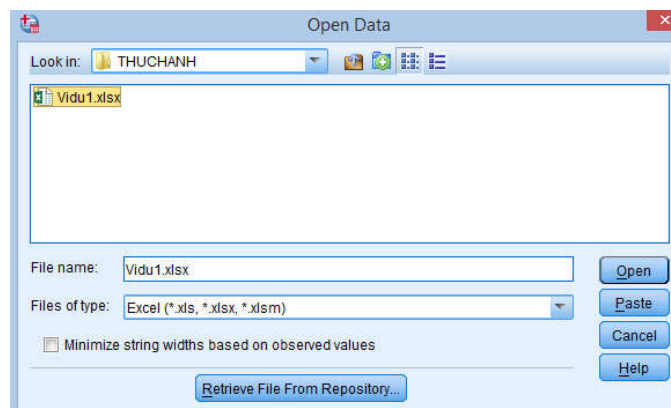
Chọn **Vidu1** rồi kích Open. Ta được

	x	y	var	var	var	var	var
1	6.00	40.00					
2	10.00	44.00					
3	12.00	46.00					
4	14.00	48.00					
5	16.00	52.00					
6	18.00	58.00					
7	22.00	60.00					
8	24.00	68.00					
9	26.00	74.00					
10	32.00	80.00					
11							

Hình 6

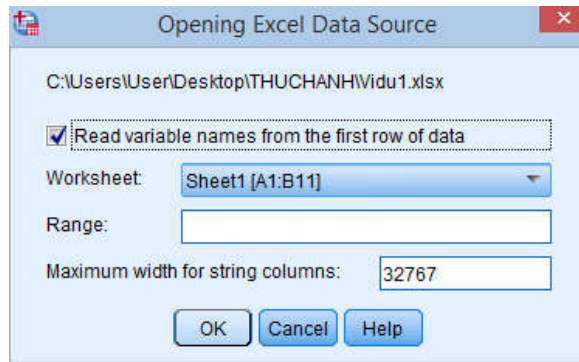
2.2.2. Đọc dữ liệu từ file .xls

Để đọc dữ liệu từ một file.xls, chẳng hạn file dữ liệu **Vidu1.xlsx**, ta vào **File** → **Open** → **Data** và chọn đến thư mục để file dữ liệu và mở file:



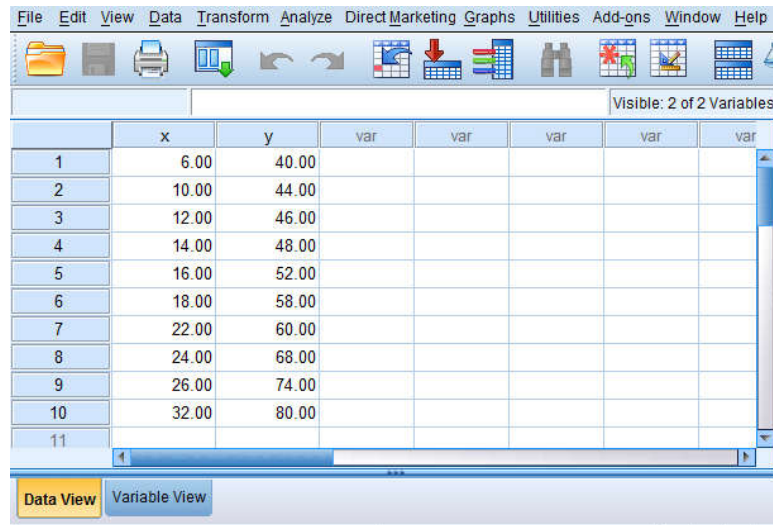
Hình 7

Chọn **Vidu1** rồi kích Open màn hình xuất hiện cửa sổ sau



Hình 8

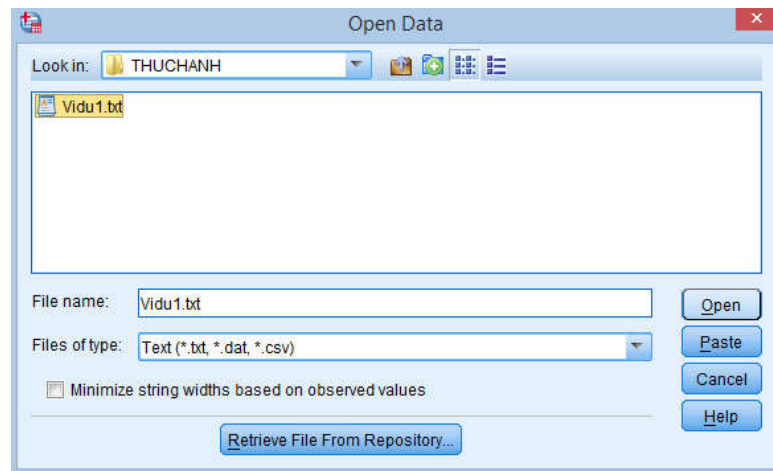
Kích **OK**, ta được



Hình 9

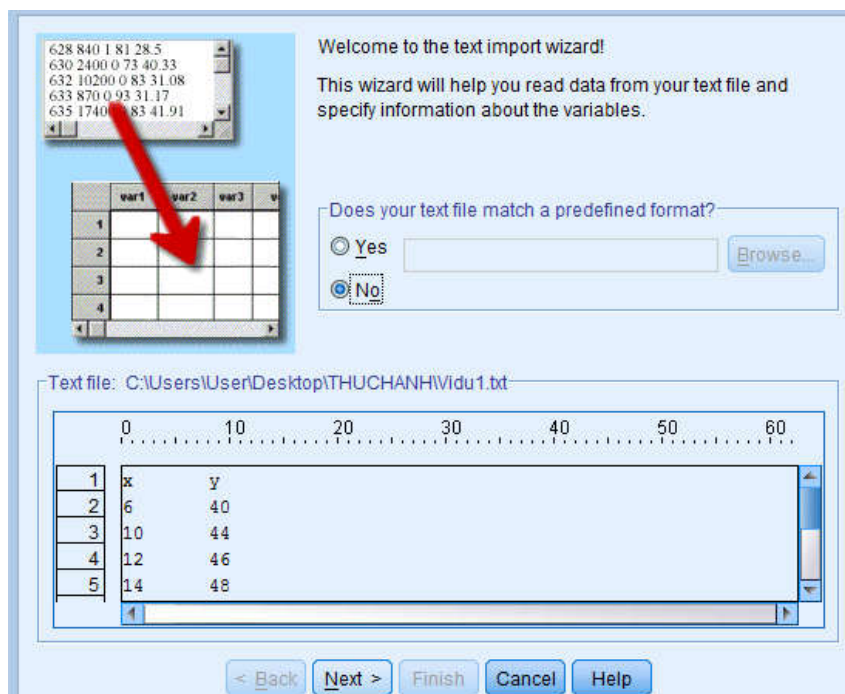
2.2.3. Đọc dữ liệu từ file.txt

Để đọc dữ liệu từ một file.txt, chẳng hạn file dữ liệu **Vidu1.txt**, ta vào **File** → **Open** → **Data** và chọn đến thư mục để file dữ liệu và mở file:



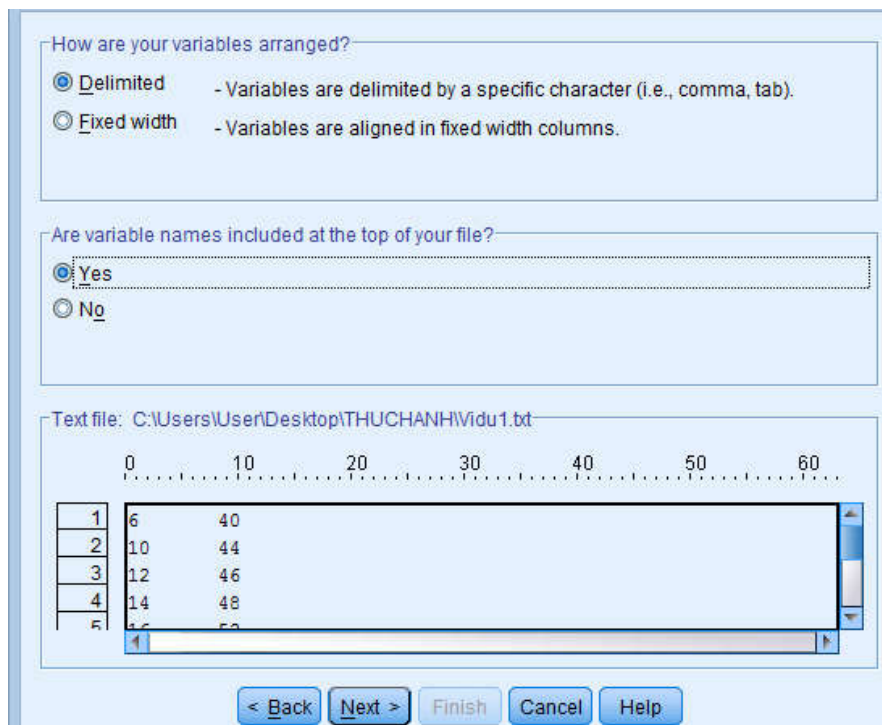
Hình 10

Khi màn hình trên hiện ra hộp thoại sau ta ấn vào **Open**



Hình 11

Ta ấn **Next**, cửa sổ sau xuất hiện



Hình 12

Khi trên màn hình hiện ra hộp thoại sau ta chọn **Yes** cho câu hỏi **Are variable names include at the top of your file?** và sau đó ấn vào **Next**. Tiếp tục nhấn **Next** → **Next** → **Finish**, ta được:

	x	y	var	var	var	var	var
1	6.00	40.00					
2	10.00	44.00					
3	12.00	46.00					
4	14.00	48.00					
5	16.00	52.00					
6	18.00	58.00					
7	22.00	60.00					
8	24.00	68.00					
9	26.00	74.00					
10	32.00	80.00					
11							

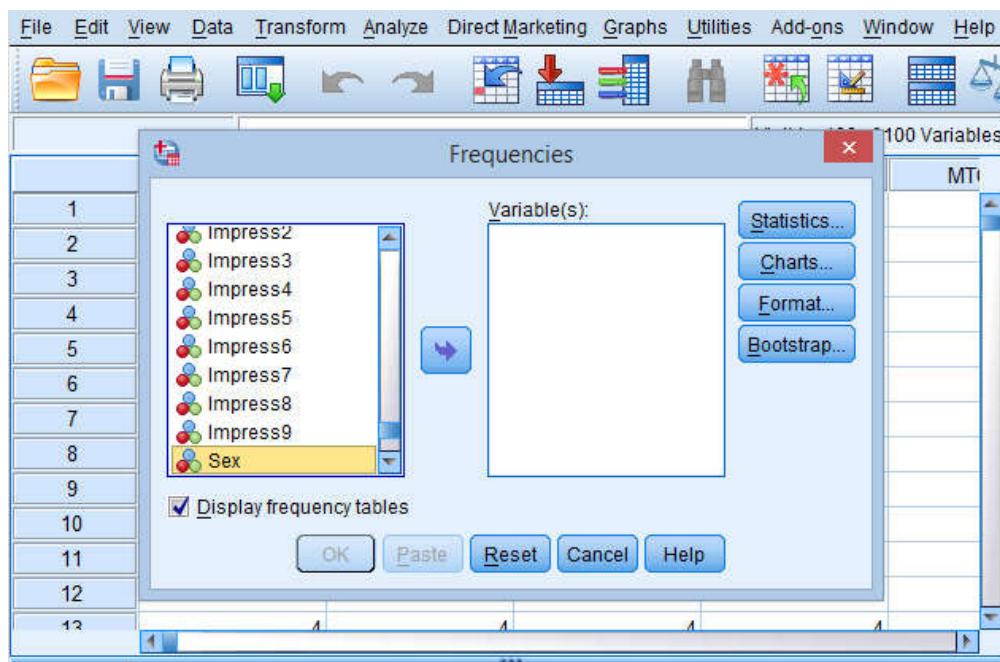
Hình 13

3. Lập bảng tần số và tính các đại lượng thống kê mô tả

3.1. Lập bảng tần số

Để lập bảng tần số cho các biến định tính hoặc dữ liệu định lượng có ít biểu hiện trong tập dữ liệu, chẳng hạn Ví dụ 1.

Bước 1: Vào **Analyze** → **Descriptive Statistics** → **Frequencies...** Màn hình hiện ra hộp thoại:



Hình 14

Bước 2: Chọn các biến cần lập bảng tần số bên trái đưa vào khung **Variable** bên phải và ấn **OK**.

Ta có thể tiến hành lập bảng tần số cho một biến hoặc một số biến cùng một lúc, chẳng hạn khi chọn biến Sex trong cửa sổ **Output** nhận được kết quả sau:

Sex

	Frequency	Percent	Valid Percent	Cumulative Percent
1	172	47.8	47.8	47.8
Valid 2	188	52.2	52.2	100.0
Total	360	100.0	100.0	

Bảng 2

3.2. Tính các đại lượng thống kê mô tả

Trong một tập dữ liệu ta có thể tính các đại lượng thống kê mô tả của tập dữ liệu như

- +) Các đại lượng đo độ tập trung: trung bình cộng, trung vị, mode.
- +) Các đại lượng đo độ phân bố: tứ phân vị và phân vị thứ p.
- +) Các đại lượng đo độ phân tán: khoảng biến thiên, độ trải giữa, phương sai, độ lệch chuẩn.
- +) Các đại lượng mô tả hình dáng của tập dữ liệu: hệ số bất đối xứng Skewness, hệ số đo độ nhọn Kurtosis.

Trước hết ta nhắc lại định nghĩa và công thức tính toán những đại lượng này.

+) Trung bình cộng đơn giản được tính bằng cách cộng tất cả các giá trị quan sát của tập dữ liệu rồi chia cho số quan sát của tập dữ liệu đó.

+) Trung vị là giá trị đứng giữa của tập dữ liệu đã được sắp thứ tự. Như vậy, không kể trung vị sẽ có 50% số quan sát của tập dữ liệu có giá trị lớn hơn trung vị và 50% số quan sát của tập dữ liệu có giá trị nhỏ hơn trung vị.

+) Mode của một tập dữ liệu là giá trị xuất hiện nhiều nhất trong tập dữ liệu.

+) Tứ phân vị chia tập dữ liệu đã sắp xếp theo trật tự tăng dần thành bốn phần có số quan sát bằng nhau. Tứ phân vị bao gồm ba tứ phân vị: tứ phân vị thứ nhất Q_1 , tứ phân vị thứ hai Q_2 , (chính là trung vị) và tứ phân vị thứ ba Q_3 . Trong tập dữ liệu có 25% số quan sát của tập dữ liệu nhỏ hơn hoặc bằng Q_1 , 50% số quan sát của tập dữ liệu nhỏ hơn hoặc bằng Q_2 và có 75% số quan sát của tập dữ liệu nhỏ hơn hoặc bằng Q_3 .

+) Phân vị thứ p của một tập dữ liệu đã được sắp thứ tự là giá trị chia tập dữ liệu thành hai phần, một phần gồm p% số quan sát có giá trị nhỏ hơn hoặc bằng phân vị thứ p, phần còn lại có (100 - p)% số quan sát lớn hơn hoặc bằng phân vị thứ p.

+) Khoảng biến thiên của một tập dữ liệu là hiệu giữa giá trị lớn nhất và giá trị nhỏ nhất của tập dữ liệu.

+) Độ trải giữa của một tập dữ liệu là hiệu độ chênh lệch giữa tứ phân vị thứ ba và tứ phân vị thứ nhất của tập dữ liệu.

+) Phương sai của một tập dữ liệu tổng thể, ký hiệu là σ^2 , được xác định bởi công thức:

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2,$$

trong đó μ là trung bình của tổng thể và N là số quan sát trong tổng thể.

+) Phương sai của một tập dữ liệu mẫu, ký hiệu là s^2 , được xác định bởi công thức:

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2,$$

trong đó \bar{x} là trung bình của mẫu và n là số quan sát trong mẫu.

+) Độ lệch chuẩn của một tập dữ liệu tổng thể, ký hiệu là σ , là căn bậc hai của phương sai của tổng thể:

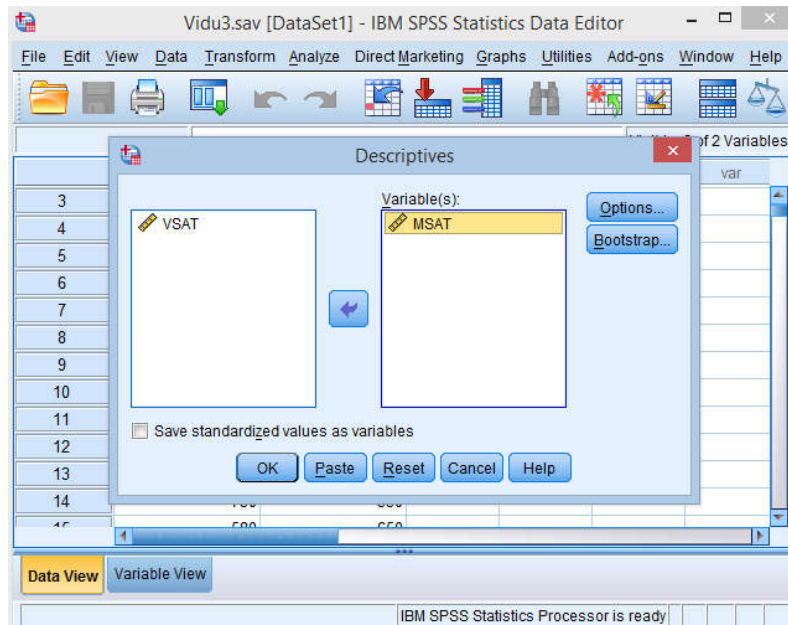
$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2}.$$

+) Độ lệch chuẩn của một tập dữ liệu mẫu, ký hiệu là s, là căn bậc hai của phương sai mẫu:

$$s = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}.$$

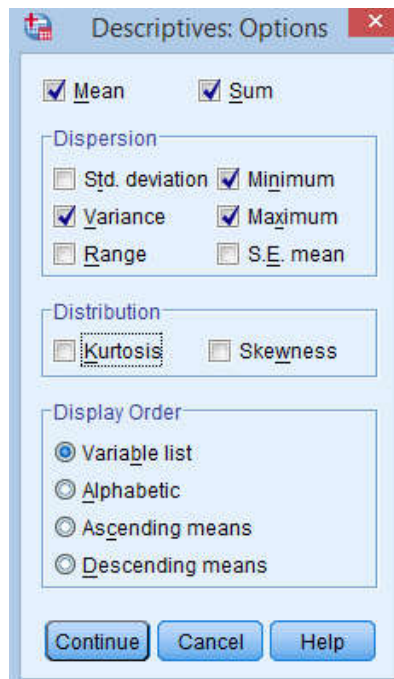
Chẳng hạn, để tính những đại lượng thống kê mô tả như: trung bình, trung vị, phương sai, độ lệch chuẩn của các biến từ file Vidu3.sav.

Bước 1. Vào **Analyze** → **Descriptive Statistics** → **Descriptive...** và khi màn hình sẽ hiện ra hộp thoại sau thì ta chọn biến MSAT để đưa vào khung **Variable(s)** ở bên phải như sau:



Hình 15

Bước 2. Ấn tiếp vào nút **Options** để chọn tính những đại lượng thống kê mô tả cho biến MSAT và màn hình hiện ra hộp thoại:



Hình 16

Trong đó

Mean	Trung bình cộng
Sum	Tổng các giá trị của biến
Std.Deviation	Độ lệch chuẩn
Minimum	Giá trị nhỏ nhất

Maximum	Giá trị lớn nhất
Variance	Phương sai
Range	Khoảng biến thiên
SE mean	Độ lệch chuẩn của trung bình mẫu
Kurtosis	Hệ số đo độ nhọn
Skewness	Hệ số bất đối xứng

Bảng 3

Ta muốn tính đại lượng nào thì chọn vào đại lượng đó để hiện kết quả ra cửa sổ Output sau đó ấn **Continue** và khi màn hình hiện ra hội thoại cũ thì ấn **OK**. Chẳng hạn, ta chọn như hình cửa sổ Output sẽ cho ta kết quả sau:

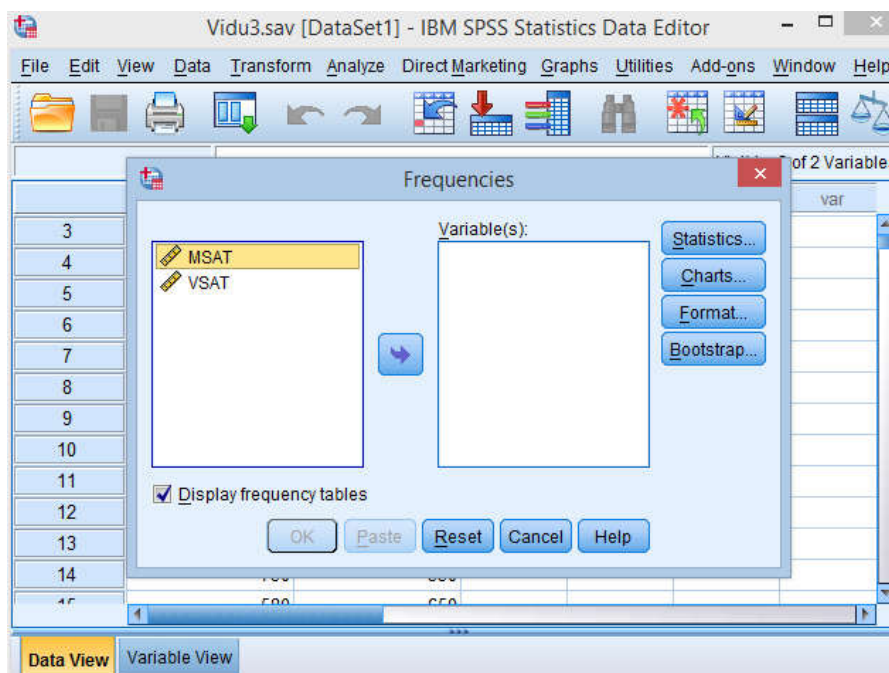
Descriptive Statistics

	N	Minimum	Maximum	Sum	Mean	Variance
MSAT	427	330	770	241820	566.32	8671.191
Valid N (listwise)	427					

Bảng 4

Tính tứ phân vị hay phân vị thứ p tùy ý của tập dữ liệu, chẳng hạn cho biến MSAT:

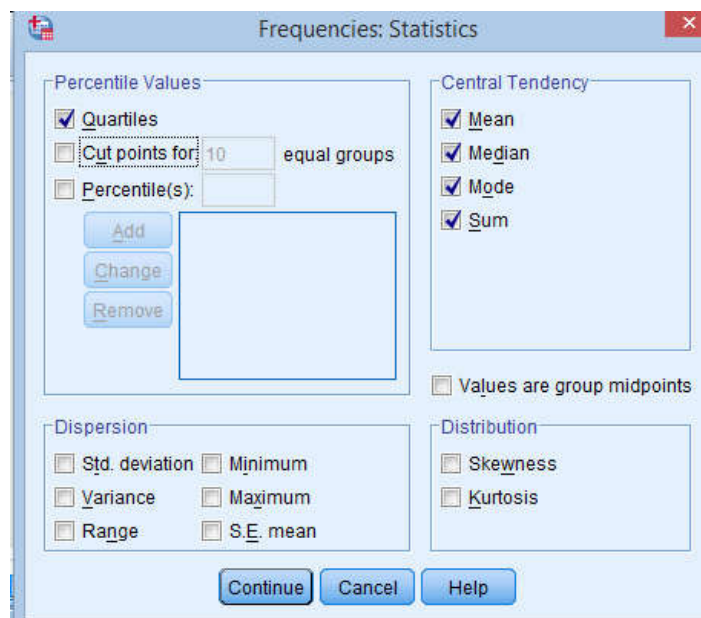
Bước 1. Vào **Analyze** → **Descriptive Statistics** → **Frequencies...** và màn hình hiện ra hội thoại:



Hình 17

Bước 2. Chọn các biến định lượng cần tính bên trái đưa vào khung **Variable** bên phải, ấn **Statistics** và nếu ta không muốn đưa bảng tần số thì không chọn ô **Display frequency tables**.

Bước 3. Khi màn hình hiện ra hộp thoại:



Hình 18

ta chọn những đại lượng cần tính, ấn **Continue** trở về hộp thoại trước ấn **OK**. Ta được

Statistics

MSAT

N	Valid	427
	Missing	0
Mean		566.32
Median		570.00
Mode		530
Sum		241820
Percentiles	25	510.00
	50	570.00
	75	640.00

Bảng 5

Kết quả tính toán này trên SPSS cho ta thấy:

+) Trung vị của MSAT là 570, mode của tuổi là 530;

+) Ba tứ phân vị là: $Q_1 = 510$; $Q_2 = 570$; $Q_3 = 640$;

4. Biểu đồ

4.1. Một số loại biểu đồ

+) Biểu đồ thanh, biểu đồ tròn: áp dụng cho dữ liệu định tính

+) Biểu đồ hộp và râu, biểu đồ tần số, biểu đồ tán xạ, biểu đồ gấp khúc,...: áp dụng cho dữ liệu định lượng.

4.2. Biểu đồ thanh

+) Biểu đồ thanh được dùng cho dữ liệu định danh hay thứ bậc.

+) Chiều cao của mỗi thanh biểu diễn tần số hay tần suất của trường hợp biểu diễn bởi thanh đó.

+) Có thể áp dụng để biểu diễn cho một hay nhiều tập dữ liệu trên cùng 1 biểu đồ.

Cách vẽ biểu đồ thanh

Bước 1. Vào menu **Graph** → **Bar** để mở cửa sổ **Bar Chart**

Bước 2. Tại cửa sổ **Bar Chart**

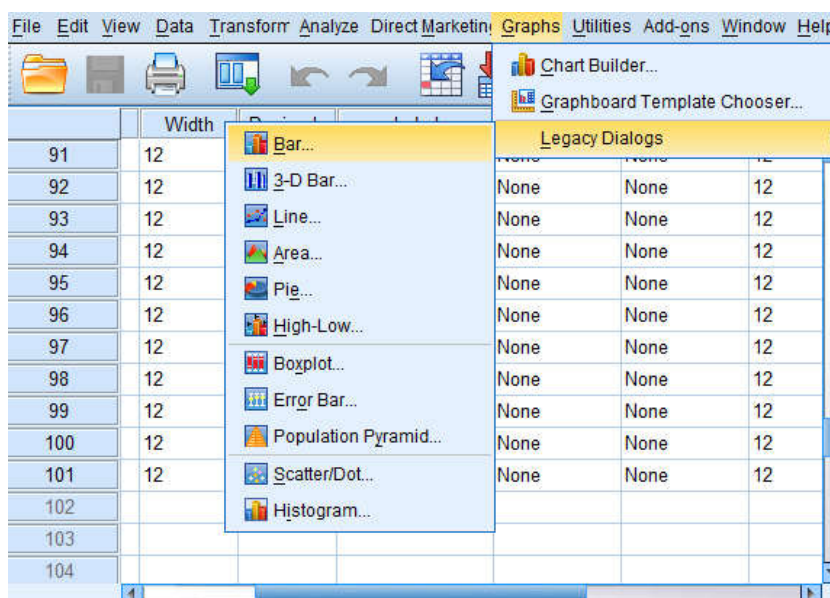
+) Chọn **Simple**: nếu biểu diễn tập dữ liệu một biến

+) Chọn **Clustered**: nếu muốn biểu diễn tập dữ liệu của một biến được chia nhóm bởi một biến khác, các thanh đứng kề nhau

+) Chọn **Stacked**: nếu muốn biểu diễn tập dữ liệu của một biến được chia nhóm bởi một biến khác, các thanh chồng lên nhau

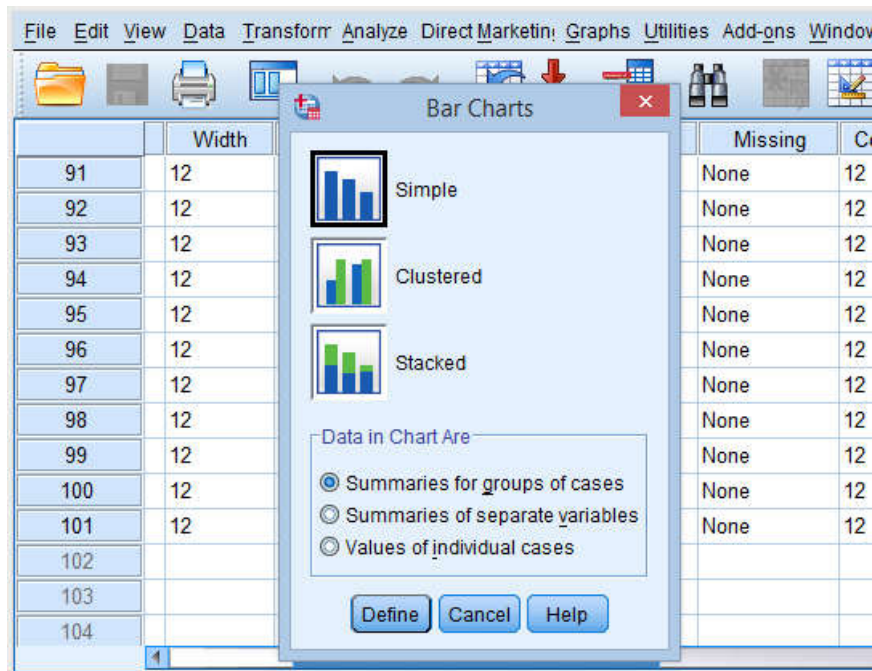
Vẽ biểu đồ thanh biểu diễn một biến: Để vẽ biểu đồ thanh biểu diễn biến nghề nghiệp (Career) trong file **Vidu2.sav** ta tiến hành như sau

Bước 1. Từ cửa sổ **Data View** → **Graphs** → **Legacy Dialogs** → **Bar**



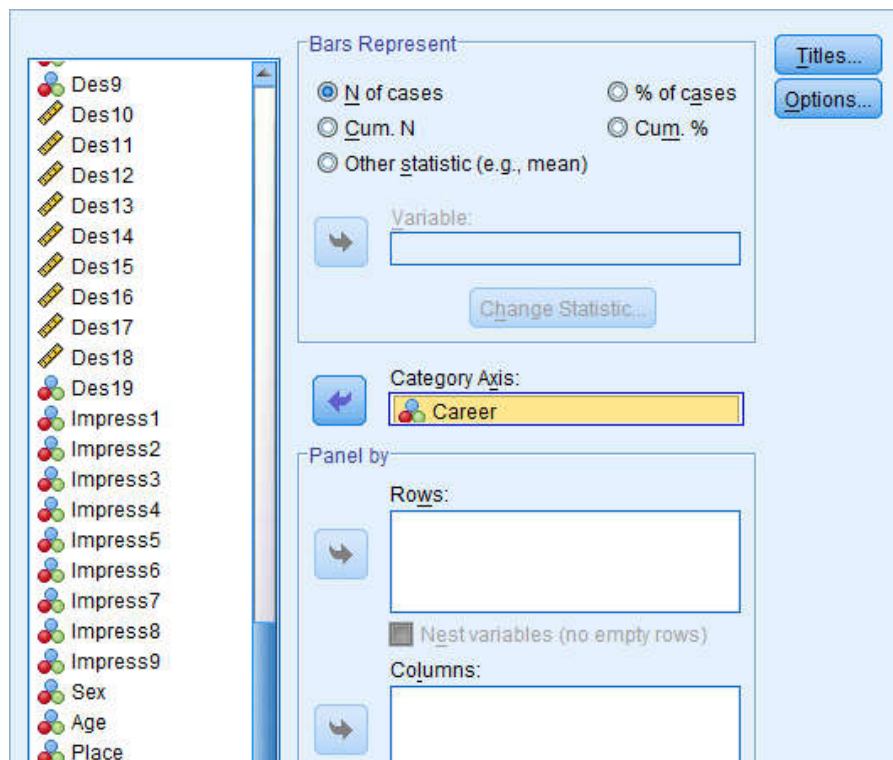
Hình 19

Kích đúp Bar, ta được



Hình 20


Bước 2. Tại hộp thoại **Bar Charts**, chọn **Simple**. Tại khu vực **Data in Chart Area** chọn **Summary for group of cases**. Ấn **Define**. Xuất hiện cửa sổ



Hình 21

Bước 3. Đưa biến **Career** vào khung **Category Axis**. Lựa chọn **N of cases** để trục tung biểu diễn tần số, **% of cases** để trục tung biểu diễn tần suất.

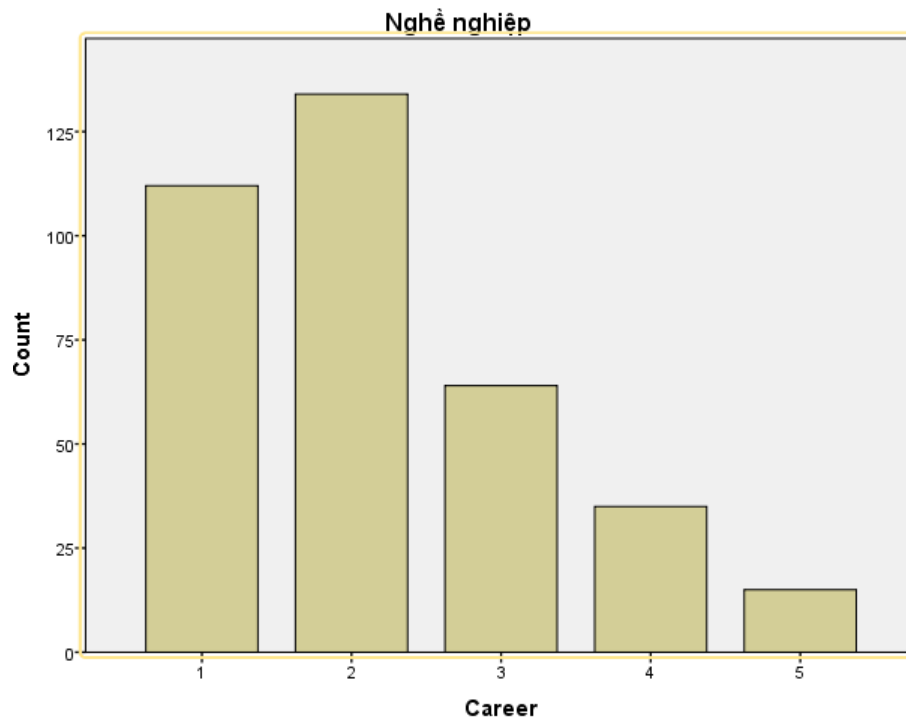
Bước 4. Ấn nút **Title** để đặt tiêu đề cho hình rồi ấn **Continue**.



The screenshot shows a dialog box titled "Titles" with a close button (X) in the top right corner. It contains three main sections: "Title", "Subtitle", and "Footnote". The "Title" section has two input fields: "Line 1" with the text "Nghề nghiệp" and "Line 2" which is empty. The "Subtitle" section has one empty input field. The "Footnote" section has two empty input fields: "Line 1" and "Line 2". At the bottom of the dialog are three buttons: "Continue", "Cancel", and "Help".

Hình 22

Bước 5. Ấn **OK**. Ta được

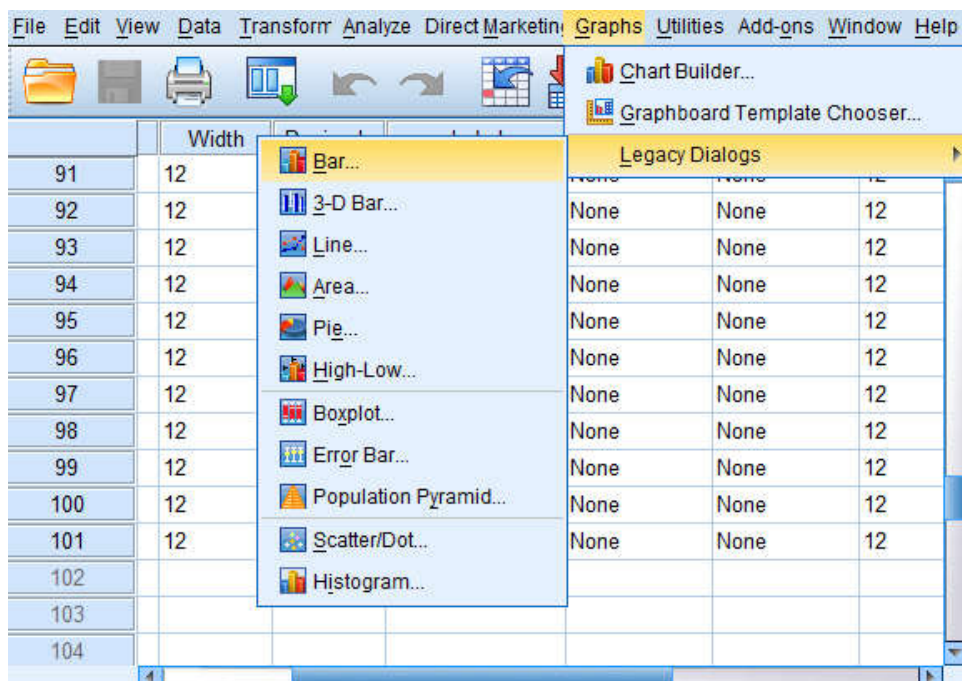


Hình 23

Ví dụ: vẽ biểu đồ thanh biểu diễn một biến được phân tách bởi một biến khác

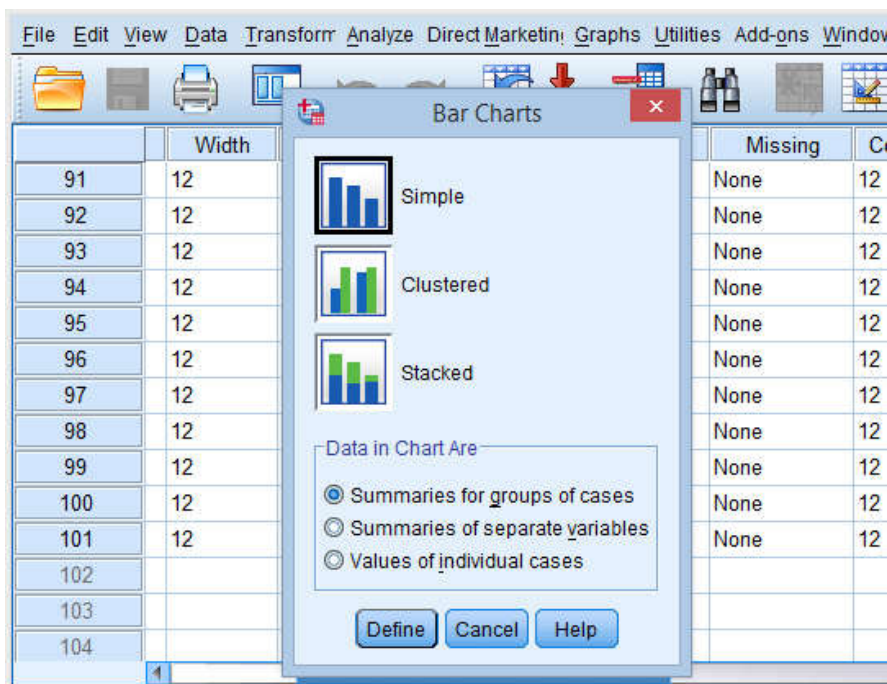
Để vẽ biểu đồ thanh biểu diễn biến nghề nghiệp (Career) theo từng nhóm nam, nữ (biến nghề nghiệp được phân tách bởi biến giới tính) trong file **Vidu2.sav** ta tiến hành như sau:

Bước 1. Từ cửa sổ **Data View** → **Graphs** → **Legacy Dialogs** → **Bar**



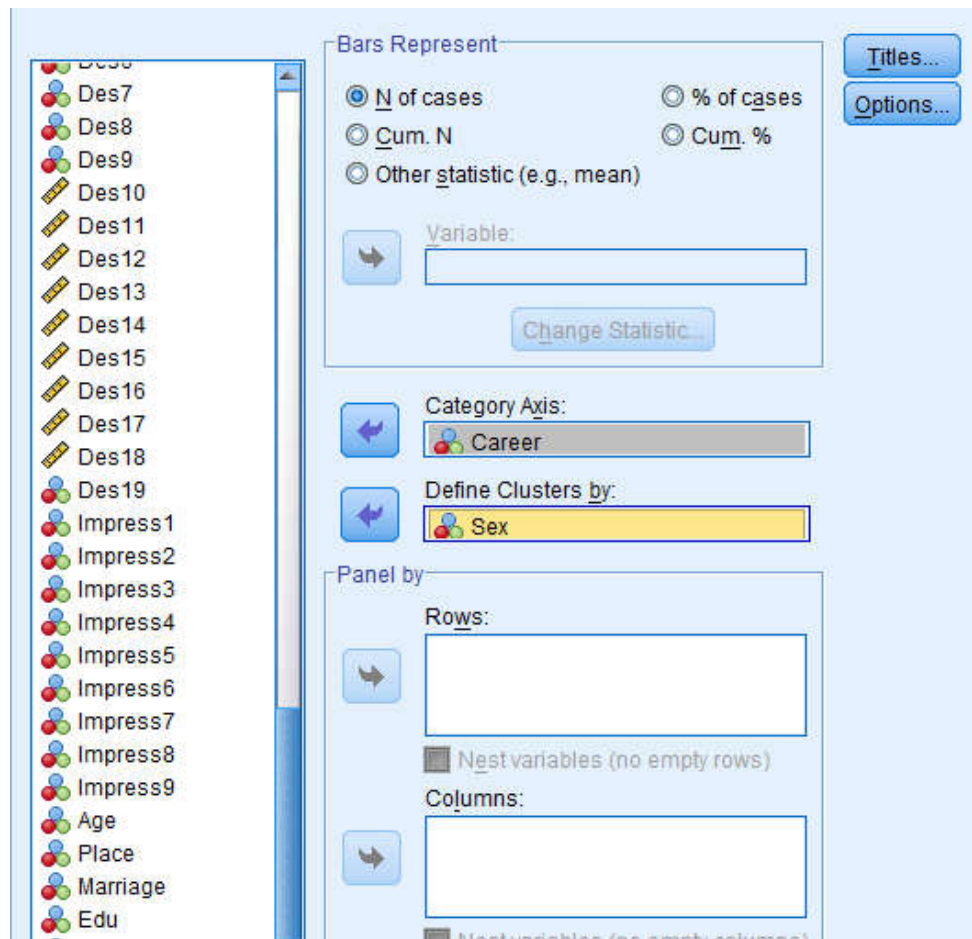
Hình 24

Kích đúp Bar, ta được



Hình 25

Bước 2. Tại hộp thoại **Bar Charts**, chọn **Clustered**. Tại khu vực **Data in Chart Area** chọn **Summary for group of cases**. Ấn **Define**. Xuất hiện cửa sổ



Hình 26

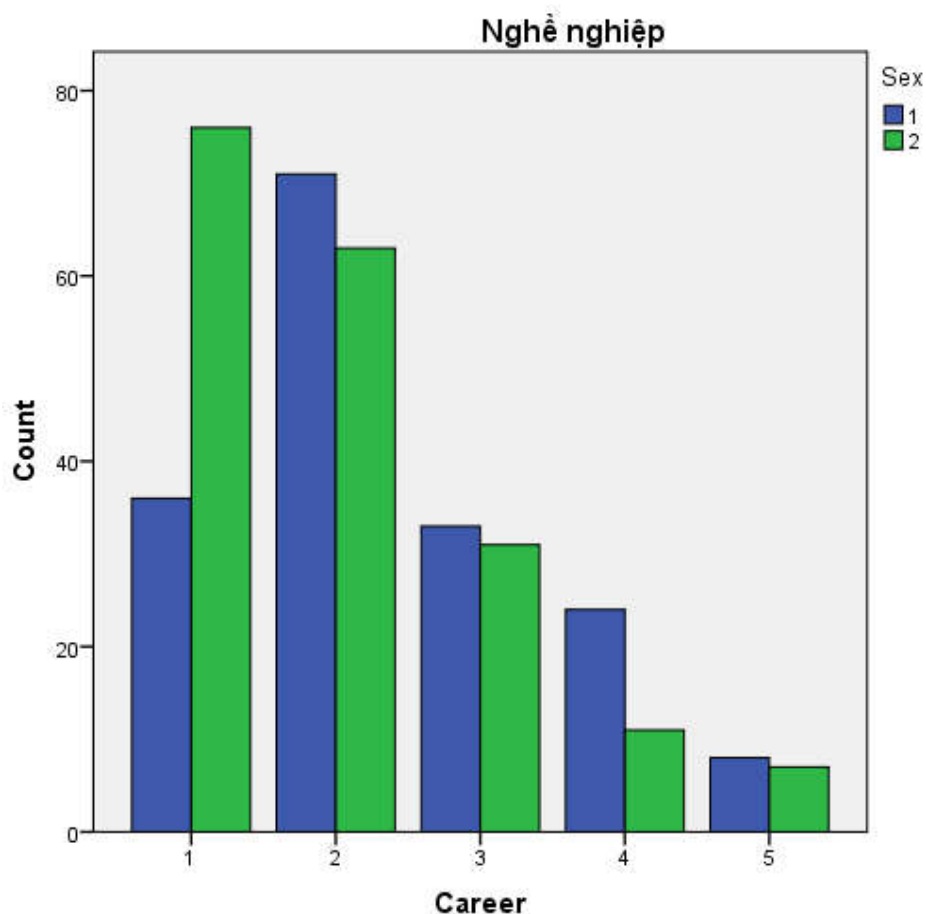
Bước 3. Đưa biến Career vào khung **Category Axis**. Lựa chọn **N of cases** để trực tung biểu diễn tần số, **% of cases** để trực tung biểu diễn tần suất.

Bước 4. Ấn nút **Title** để đặt tiêu đề cho hình rồi ấn **Continue**.



Hình 27

Bước 5. Ấn OK. Ta được



Hình 28

4.3. Biểu đồ hình tròn

- +) Biểu đồ tròn được dùng cho dữ liệu định danh hay thứ bậc
- +) Mỗi hình quạt biểu diễn tỷ lệ đóng góp của từng biểu hiện.

Cách vẽ biểu đồ tròn

Bước 1. Vào menu **Graph** → **Pie** để mở cửa sổ **Pie Chart**

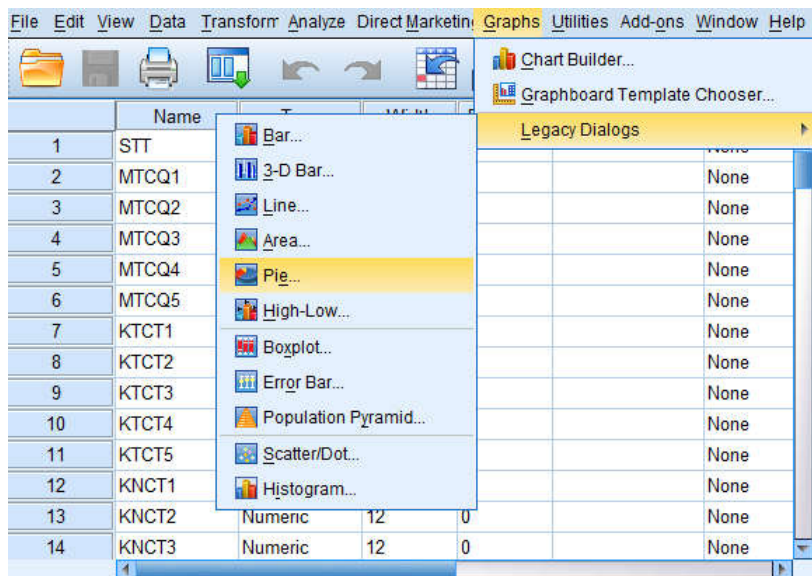
Bước 2. Tại cửa sổ **Pie Chart** chọn **Summary for group of cases**. Ấn **Define**. Xuất hiện cửa sổ sau

Bước 3. Đưa biến cần biểu diễn vào khung **Define Slice by**. Ấn **OK**

Ví dụ: vẽ biểu đồ hình tròn biểu diễn một biến

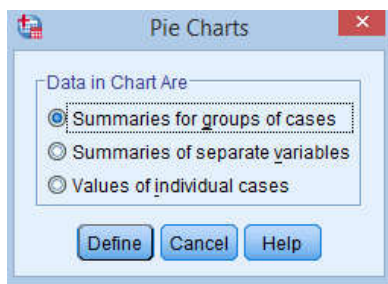
Để vẽ biểu đồ thanh biểu diễn biến nghề nghiệp (Career) trong file **Vidu2.sav** ta tiến hành như sau

Bước 1. Từ cửa sổ **Data View** → **Graphs** → **Legacy Dialogs** → **Pie**



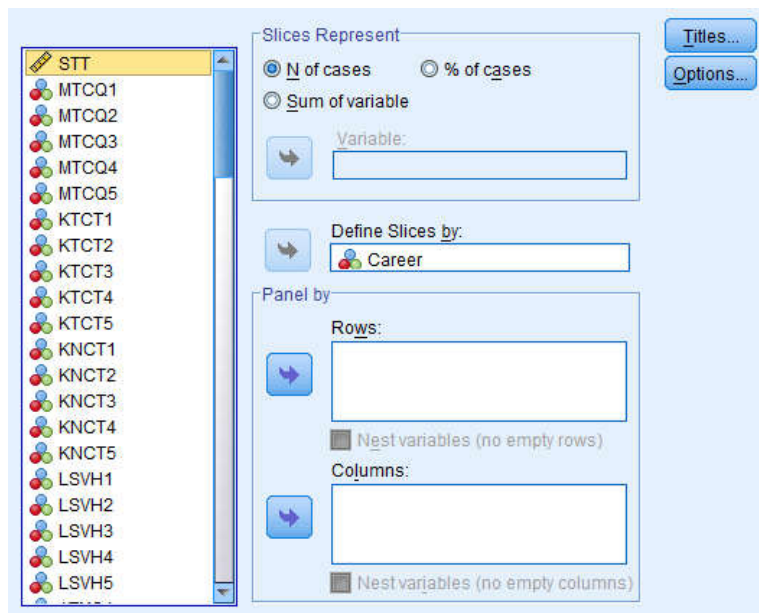
Hình 29

Kích đúp **Pie**, ta được



Hình 30


Bước 2. Tại cửa sổ **Pie Chart** chọn **Summary for group of cases**. Ấn **Define**. Xuất hiện cửa sổ sau



Hình 31

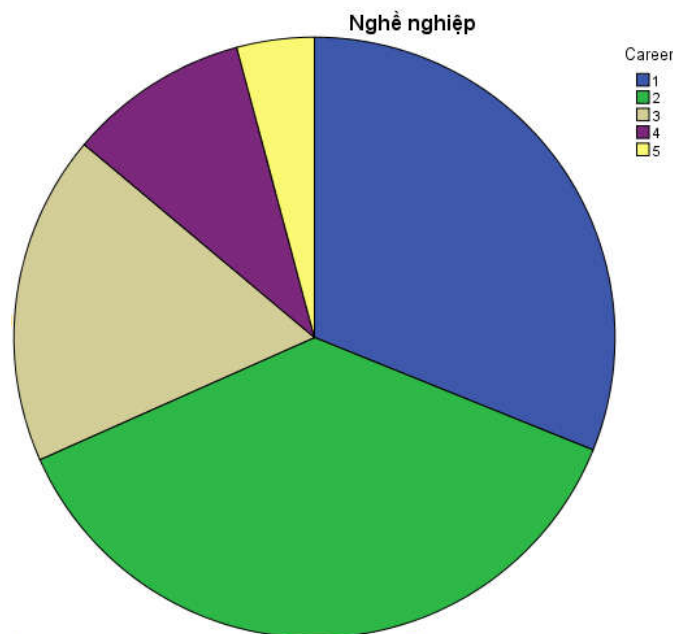
Bước 3. Đưa biến **Career** vào khung **Define Slice by**.

Bước 4. Ấn nút **Title** để đặt tiêu đề cho hình rồi ấn **Continue**.



Hình 32

Bước 5. Ấn **OK**. Ta được



Hình 33

4.4. Biểu đồ tần số (Histogram)

+) Biểu đồ tần số: dùng cho biến định lượng liên tục nhằm biểu diễn phân phối của tập dữ liệu.

+) Chiều cao mỗi thanh hình chữ nhật trong biểu đồ biểu diễn tần số của lớp tương ứng.

Cách vẽ biểu đồ tần số (Histogram)

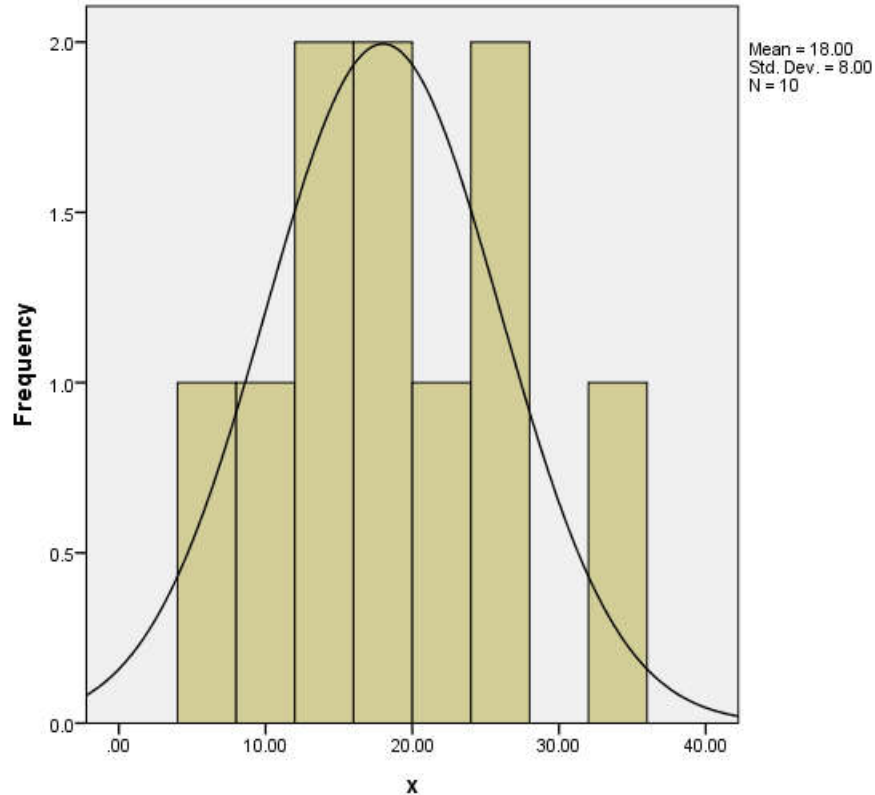
Bước 1. Chọn **Graph** → **Histogram**

Bước 2. Chọn một biến định lượng vào khung **Variable**

Bước 3. Nếu muốn hiển thị đường phân phối chuẩn thì chọn **Display normal curve**.

Ấn **OK**.

Ví dụ ta cần vẽ biểu đồ tần số cho biến tuổi trong file **Vidu1.sav**. Trong bước hai ta chọn biến định lượng là **X**. Hình vẽ như sau



Hình 34

4.5. Biểu đồ phân tán (Scatter plot)

Biểu đồ tán xạ dùng để biểu diễn mối quan hệ giữa hai biến định lượng.

Cách vẽ biểu đồ phân tán (Scatter plot)

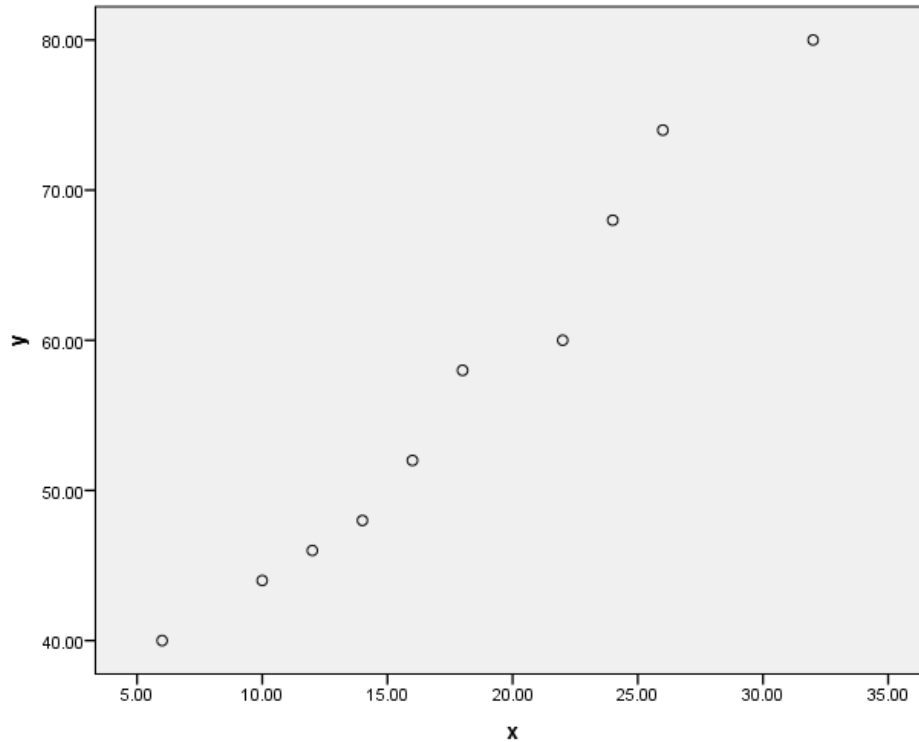
Bước 1: Chọn **Graph** → **Scatter/Dot**. Được cửa sổ sau

Bước 2: Tại cửa sổ **Scatter/Dot**, chọn **Simple** nếu vẽ biểu đồ cho một cặp biến.

(**Ovelay** biểu diễn nhiều đám mây đơn cùng biểu đồ,....)

Bước 3: Chọn biến đưa vào khung **Y-axis** (giá trị là tung độ của điểm), biến đưa vào khung **X-axis** (giá trị là hoành độ của điểm)

Ví dụ ta cần vẽ biểu đồ phân tán trong file **Vidu1.sav**. Hình vẽ như sau

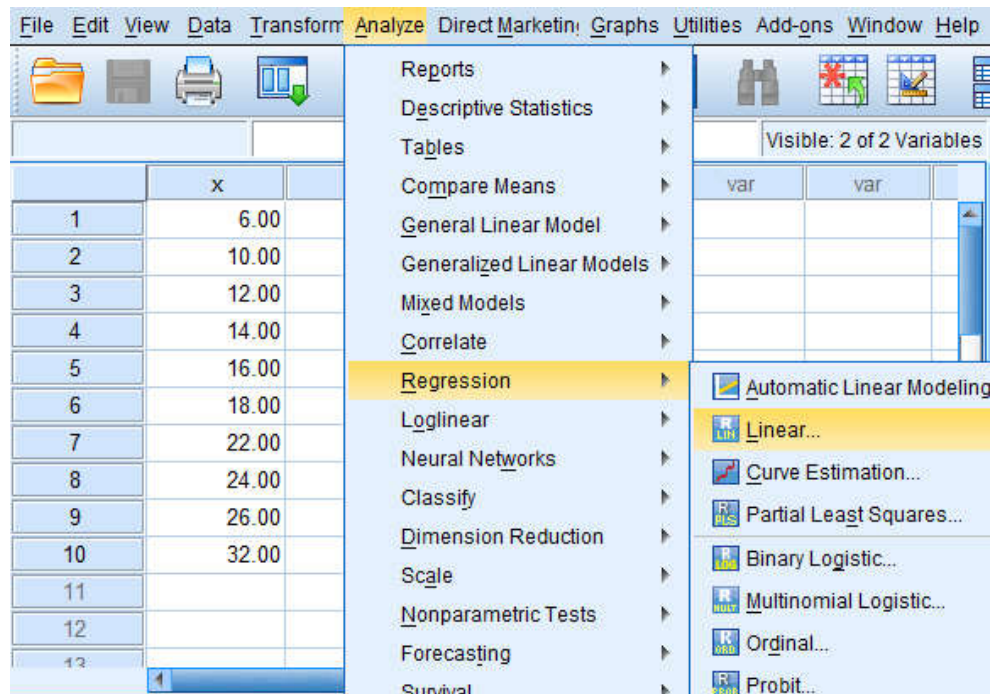


Hình 35

5. Tìm hàm hồi quy tuyến tính, kiểm định mô hình, ước lượng hệ số hồi quy

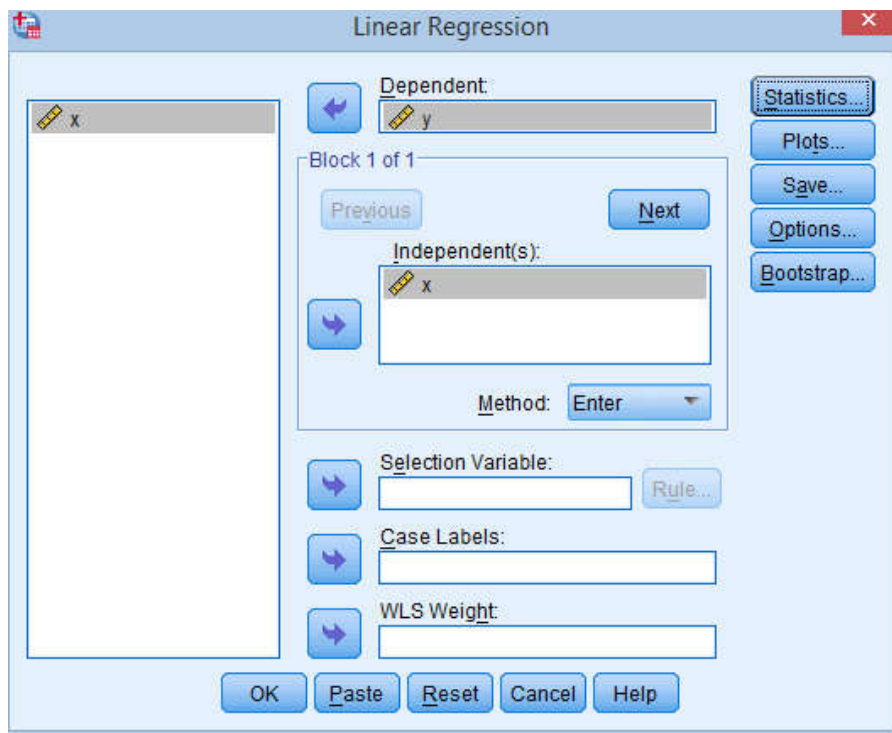
Ta cần tìm hàm hồi quy tuyến tính, kiểm định mô hình, ước lượng hệ số hồi quy trong file **Vidu1.sav**. Ta thực hiện như sau

Bước 1. Từ cửa sổ Data View chọn Analyze → Regression → Linear



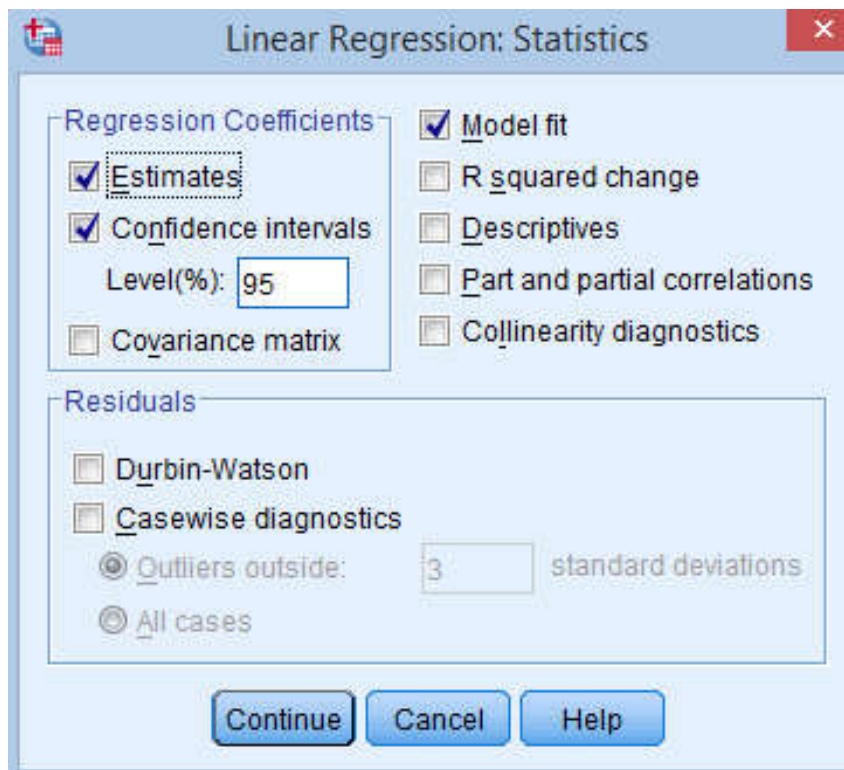
Hình 36

Kích đúp **Linear**, ta được



Hình 37

Bước 2. Đưa biến Y sau **Dependent** và biến X vào **Independent(s)**. Chọn **Statistics** ta được bảng sau



Hình 38

Tích vào **Confidence intervals**. Ấn **Continue**

Kích **OK**, ta được các bảng sau

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.985 ^a	.971	.967	2.43171

a. Predictors: (Constant), x

Bảng 6. Tóm tắt mô hình

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1586.694	1	1586.694	268.331	.000 ^b
	Residual	47.306	8	5.913		
	Total	1634.000	9			

a. Dependent Variable: y

b. Predictors: (Constant), x

Bảng 7. Phân tích phương sai

Coefficients^a

Model		Unstandardized		Standardized	t	Sig.	95.0% Confidence	
		Coefficients					Interval for B	
		B	Std. Error	Beta				Lower Bound
1	(Constant)	27.125	1.979		13.705	.000	22.561	31.689
	x	1.660	.101	.985	16.381	.000	1.426	1.893

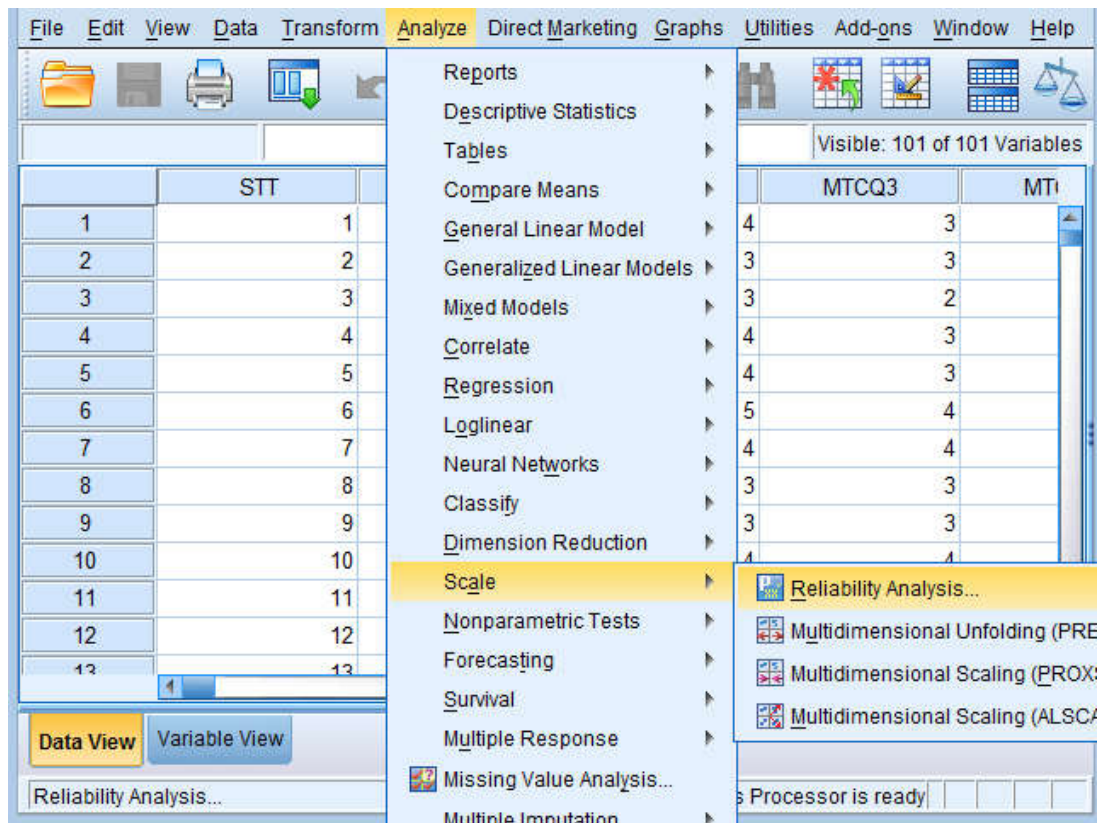
a. Dependent Variable: y

Bảng 8. Ước lượng hệ số hồi quy

6. Đánh giá độ tin cậy các thang đo bằng Cronbach 's alpha

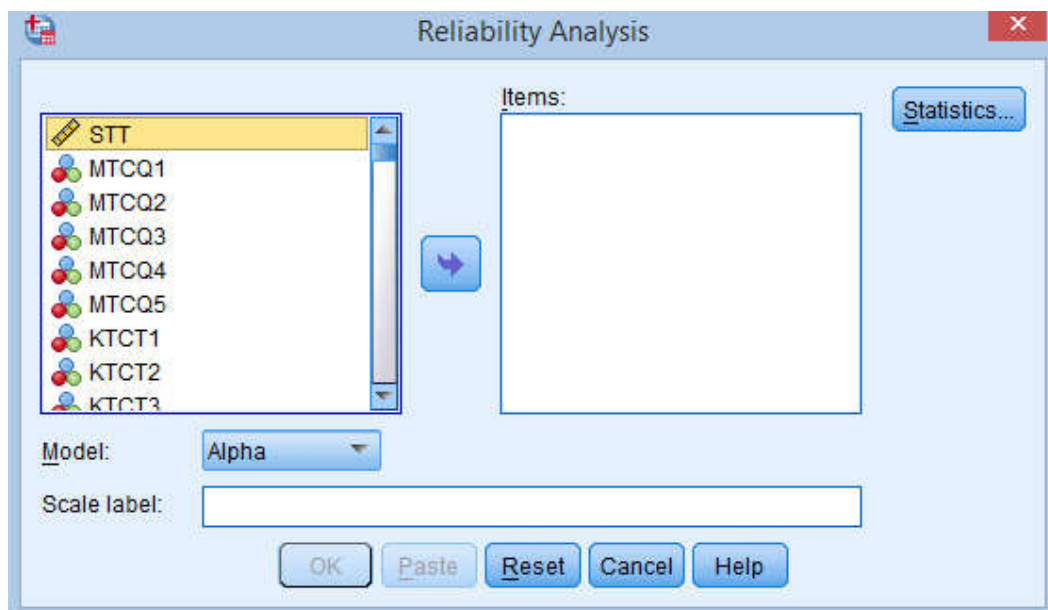
Trong phần này áp dụng cụ thể cho file **Vidu2.sav**. Ta kiểm định thang đo của môi trường cảnh quan (MTCQ1,..., MTCQ5) như sau

Từ cửa sổ **Data View**: Chọn **Analyze** → **Scale** → **Reliability Analysis...** như sau:



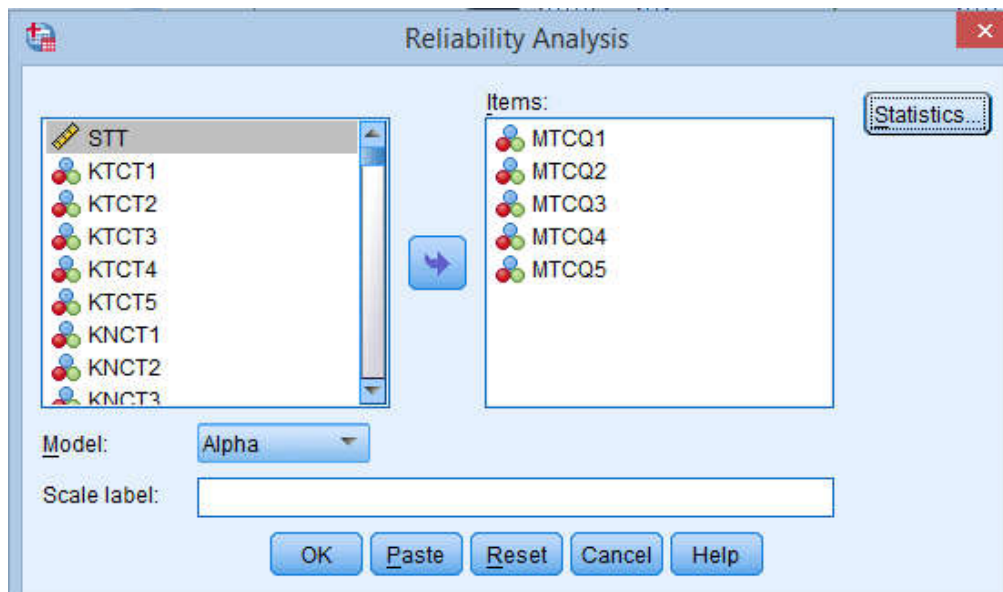
Hình 39

Kích đúp **Reliability Analysis**, ta được



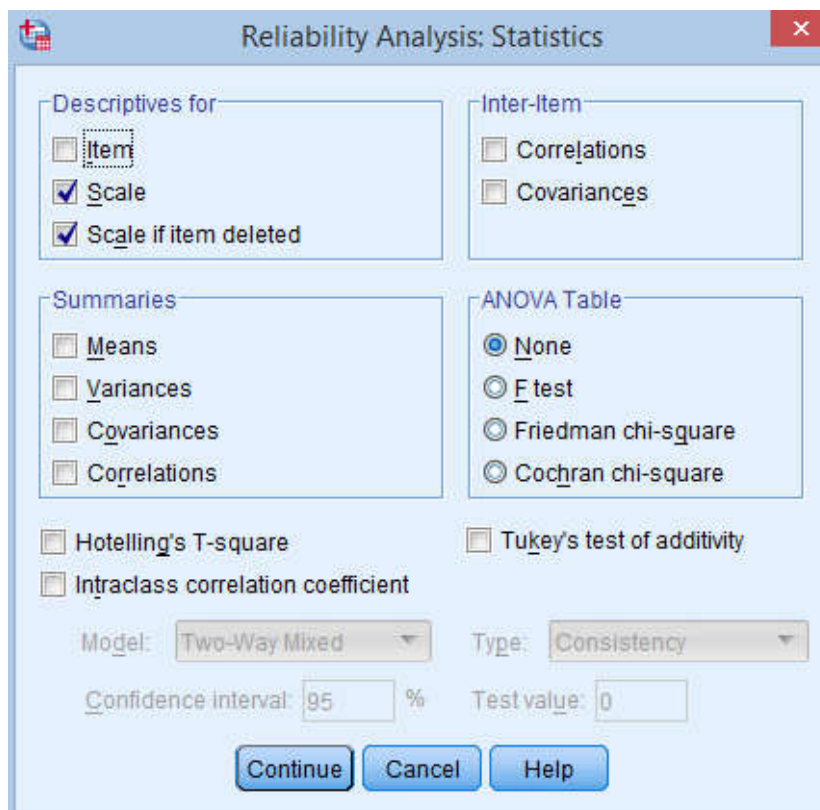
Hình 40

Trong hình 40, chọn nhóm biến môi trường cảnh quan (**MTCQ1, ..., MTCQ5**) đưa vào hộp **Item**, xuất hiện như sau:



Hình 41

Trong hình 41. Chọn **Statistics**, màn hình xuất hiện như sau:



Hình 42

Trong hình 42, chọn **Scale** và **Scale if Scale item deleted** của hộp **Descriptives for**, sau đó chọn **Continue** màn hình xuất hiện như hình 41. Chọn **OK**, ta được

Reliability Statistics

Cronbach's Alpha	N of Items
.742	5

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
MTCQ1	14.11	6.195	.421	.730
MTCQ2	13.96	5.823	.598	.663
MTCQ3	14.02	5.668	.553	.678
MTCQ4	13.99	6.237	.485	.705
MTCQ5	13.99	6.248	.479	.707

Bảng 9

Sau khi chạy Cronbach 's alpha từng yếu tố, biến KNTC1 bị loại vì hệ số tương quan biến - tổng (Corrected Item – total Correlation) là -0,138 (<0,3).

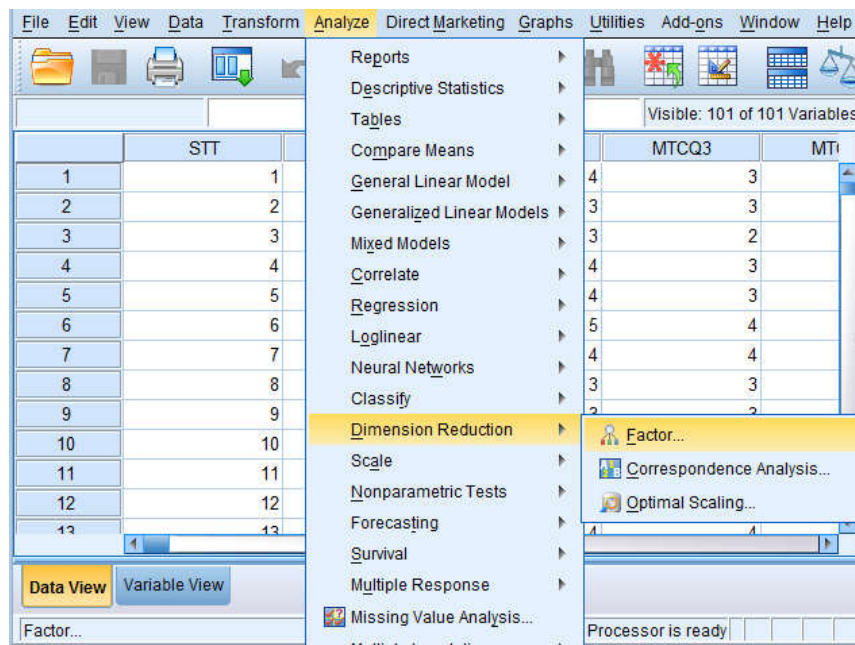
Kết quả đánh giá Cronbach 's alpha cho thấy các thang đo đều có hệ số alpha đạt yêu cầu (> 0,6), thấp nhất là thang đo khả năng tiếp cận điểm đến (KNTC) có alpha = 0,638 và cao nhất là thang đo cơ sở hạ tầng (CSHT) có alpha = 0,836.

7. Phân tích nhân tố khám phá (EFA)

Trong phần này áp dụng cụ thể cho file **Vidu2.sav**. Ta tiến hành như sau

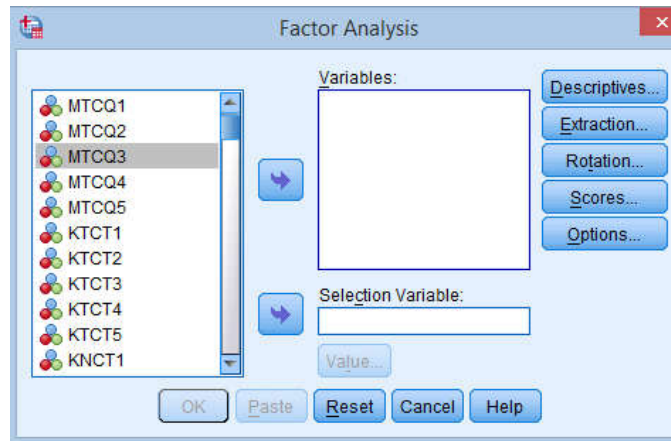
7.1. Sử dụng SPSS thực hiện phân tích EFA

Từ cửa sổ **Data View**: Chọn **Analyze** → **Dimension Reduction** → **Factor...** như sau



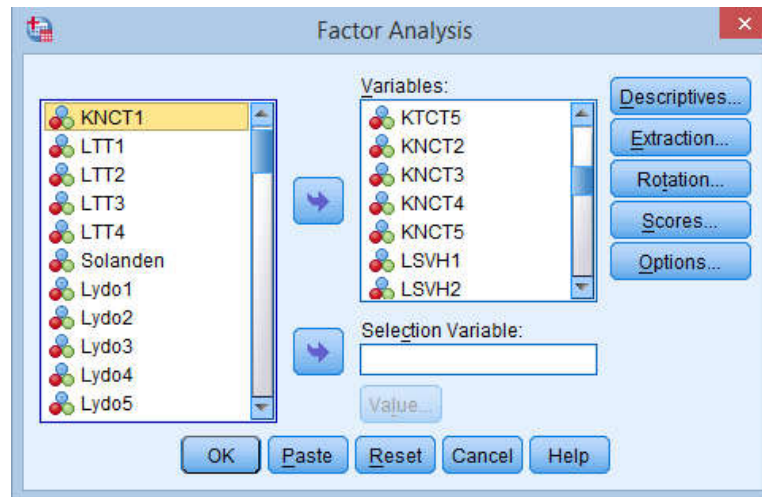
Hình 43

Kích đúp **Factor**, ta được



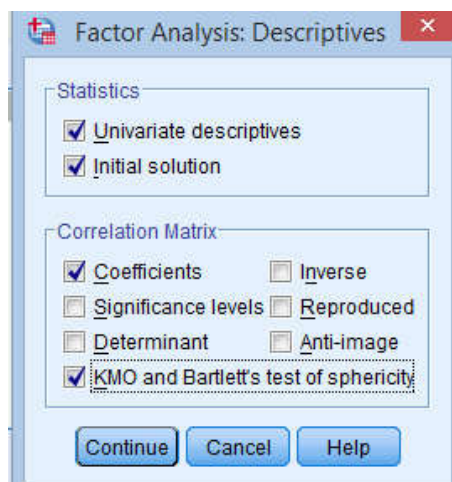
Hình 44

Trong hình 44, đưa tất cả các biến đặc trưng vào hộp **Variables** (trừ biến bị loại ở kiểm định Cronbach's Alpha và các biến của thước đo LLT, xuất hiện như sau:



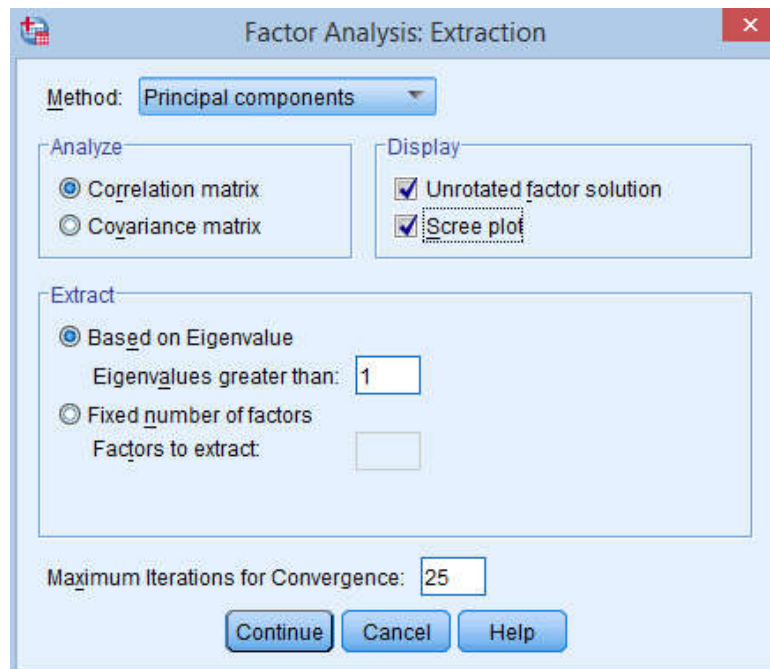
Hình 45

Chọn **Descriptives**, sẽ xuất hiện



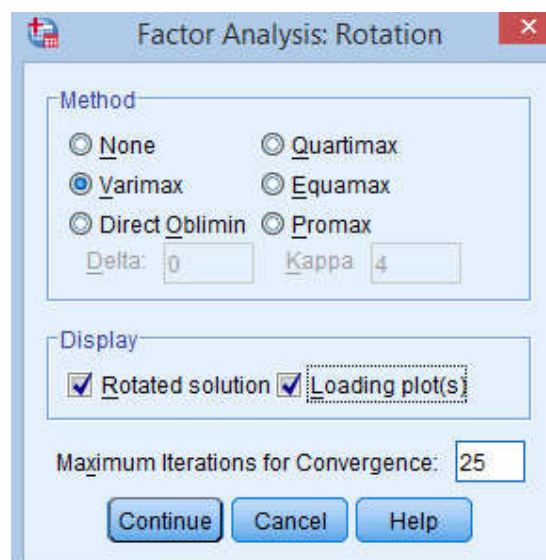
Hình 46

Trong hình 46, ta chọn **Univariate descriptives** và **Initial solution** trong hộp **Statistics**. Trong hộp **Correlation Matrix**, chọn **Coefficients** và **KMO**. Nhấn **Continue**, chọn **Extraction** sẽ xuất hiện



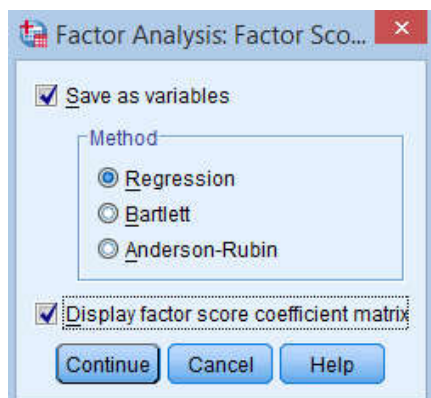
Hình 47

Trong hình 47, ta chọn **Principle components** trong hộp **Method**. Trong hộp **Analyze**, chọn **correlation matrix**. Trong hộp **Display**, chọn **Unrotated factor solution** và **Scree plot**. Trong hộp **Extract**, chọn **Based on Eigenvalue** và **Eigenvalues greater than 1**. Trong hộp **Maximum Iterations for Convergence 25**. Nhấn **Continue**, chọn **Rotation** sẽ xuất hiện



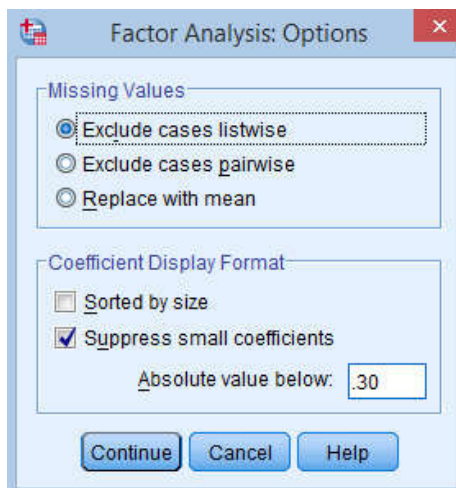
Hình 48

Trong hình 48, ta chọn **Varimax** trong hộp **Method**. Trong hộp **Display**, chọn **Rotated solution** và **Loading plot(s)**. Trong hộp **Maximum Iterations for Convergence** **25**. Nhấn **Continue**, chọn **Scores** sẽ xuất hiện



Hình 49

Trong hình 49, bên dưới **Factor Analysis**, ta chọn **Save as variables** và **Display factor score coefficient matrix**. Trong hộp **Method**, chọn **Regression**. Nhấn **Continue**, chọn **Option** sẽ xuất hiện



Hình 50

Trong hình 50, chọn **Exclude cases listwise**. Trong hộp **Coefficient Display Format**, chọn **Suppress small coefficient**. Đối với **Absolute value Below**, chọn giá trị bao nhiêu tùy thuộc vào cỡ mẫu nghiên cứu. Lưu ý

Cỡ mẫu	Giá trị chọn (Factor loadings)
Lớn hơn 350	0,3
Từ 100 đến 350	0,55
Nhỏ hơn 100	0,75

Bảng 10

Trong ví dụ trên cỡ mẫu 360 ta chọn 0,3.

Nhấn **Continue**, chọn **OK** ta có toàn bộ kết quả

Factor Analysis

[DataSet1] C:\Users\User\Desktop\luanvanconggoc.sav

Descriptive Statistics

	Mean	Std. Deviation	Analysis N
MTCQ1	3.41	.891	360
MTCQ2	3.56	.819	360
MTCQ3	3.49	.902	360
MTCQ4	3.53	.810	360
MTCQ5	3.53	.814	360
KTCT1	3.72	.783	360
KTCT2	3.98	.763	360
KTCT3	3.52	.824	360
KTCT4	3.72	.733	360
KTCT5	3.82	.722	360
KNCT2	3.75	1.058	360
KNCT3	3.79	.939	360
KNCT4	3.66	1.043	360

Bảng 11

Rotated Component Matrix^a

	Component						
	1	2	3	4	5	6	7
MTCQ1					.606		
MTCQ2					.777		
MTCQ3					.754		
MTCQ4					.691		
MTCQ5					.667		
KTCT1		.699					
KTCT2		.763					
KTCT3		.770					
KTCT4		.754					
KTCT5		.678					
KNCT2							.758
KNCT3							.713
KNCT4							.743
LSVH1			.640				
LSVH2			.711				
LSVH3			.768				
LSVH4			.682				
LSVH5			.704				

Bảng 12. Kết quả phân tích nhân tố khám phá

Ta chọn các kết quả cần thiết phục vụ cho phân tích.

7.2. Hệ thống kiểm định cho EFA

7.2.1. Kiểm định tính thích hợp của EFA

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	.759
Approx. Chi-Square	3239.022
Bartlett's Test of Sphericity	df
	496
	Sig.
	.000

Bảng 13 . Kiểm định KMO và Bartlett

Dựa vào kết quả ta có hệ số KMO là 0,759 ($>0,5$) và sig= 0,00 $<0,05$, nghĩa là các biến quan sát có tương quan tuyến tính với nhân tố đại diện.

7.2.2. Kiểm định mức độ giải thích của các biến quan sát với nhân tố

Total Variance Explained									
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.868	12.087	12.087	3.868	12.087	12.087	3.090	9.655	9.655
2	3.337	10.429	22.515	3.337	10.429	22.515	2.906	9.082	18.738
3	2.725	8.516	31.032	2.725	8.516	31.032	2.672	8.350	27.087
4	2.316	7.237	38.268	2.316	7.237	38.268	2.564	8.013	35.100
5	2.151	6.723	44.991	2.151	6.723	44.991	2.512	7.851	42.951
6	1.821	5.690	50.681	1.821	5.690	50.681	2.252	7.038	49.989
7	1.524	4.762	55.443	1.524	4.762	55.443	1.745	5.454	55.443
8	.973	3.039	58.482						
9	.961	3.002	61.484						
10	.808	2.526	64.010						

Bảng 14. Tổng phương sai được giải thích.

Phân tích EFA mang lại 7 nhóm nhân tố tác động đến LLT. Hệ số tải và hệ số tin cậy Cronbach Alpha đạt tại Eigenvalue = 1,524 và phương sai trích đạt 55,443 % ($> 0,50$). Vì thế kết quả EFA là đáng tin cậy và có thể sử dụng cho phân tích hồi qui ở bước theo.

7.2.3. Kết quả mô hình EFA

Rotated Component Matrix^a

	Component						
	1	2	3	4	5	6	7
MTCQ1					.606		

MTCQ2				.777	
MTCQ3				.754	
MTCQ4				.691	
MTCQ5				.667	
KTCT1	.699				
KTCT2	.763				
KTCT3	.770				
KTCT4	.754				
KTCT5	.678				
KNCT2					.758
KNCT3					.713
KNCT4					.743
LSVH1		.640			
LSVH2		.711			
LSVH3		.768			
LSVH4		.682			
LSVH5		.704			
ATMS1			.732		
ATMS2			.758		
ATMS3			.718		
ATMS4			.670		
ATMS5			.622		
CSHT1	.762				
CSHT2	.826				
CSHT3	.809				
CSHT4	.749				
CSHT5	.717				
GT1					.788
GT2					.728
GT3					.760
GT4					.695

Bảng 15. Ma trận nhân tố xoay.

Trong bảng 15. Cho biết các đặc trưng đều có hệ số tải nhân tố lớn hơn 0,5. Có 7 nhân tố

+) **Nhân tố 1.** Bao gồm các biến CSHT2, CSHT3, CSHT1, CSHT4, CSHT5. Ta đặt tên cho nhân tố này là **CSHT (F1)**.

+) **Nhân tố 2.** Bao gồm các biến KTCT3, KTCT2, KTCT4, KTCT1, KTCT5. Ta đặt tên cho nhân tố này là **KTCT (F2)**.

+) **Nhân tố 3.** Bao gồm các biến LSVH3, LSVH2, LSVH5, LSVH4, LSVH1. Ta đặt tên cho nhân tố này là **LSVH (F3)**.

+) **Nhân tố 4.** Bao gồm các biến ATMS2, ATMS1, ATMS3, ATMS4, ATMS5. Ta đặt tên cho nhân tố này là **ATMS (F4)**.

+) **Nhân tố 5.** Bao gồm các biến MTCQ2, MTCQ3, MTCQ4, MTCQ5, MTCQ1. Ta đặt tên cho nhân tố này là **MTCQ (F5)**.

+) **Nhân tố 6.** Bao gồm các biến GT1, GT3, GT2, GT4. Ta đặt tên cho nhân tố này là **GT (F6)**.

+) **Nhân tố 7.** Bao gồm các biến KNCT2, KNCT4, KNCT3. Ta đặt tên cho nhân tố này là **KNCT (F7)**.

7.2.4. Thực hiện phân tích hồi quy đa biến

Để nhận diện các yếu tố ảnh hưởng đến hình ảnh điểm đến, mô hình tương quan tổng thể có dạng:

$$LTT = f(F_1, F_2, F_3, F_4, F_5, F_6, F_7)$$

Trong đó

+) LTT là biến phụ thuộc;

+) F1, F2, F3, F4, F5, F6, F7 là biến độc lập.

Vì thế, chúng ta có thể dự đoán mô hình hồi quy bội với 7 nhân tố như đã xác định trong các chương trước có dạng như sau:

$$LTT = \beta_0 + \beta_1 F_1 + \beta_2 F_2 + \beta_3 F_3 + \beta_4 F_4 + \beta_5 F_5 + \beta_6 F_6 + \beta_7 F_7$$

Trong đó các biến đưa vào phân tích hồi quy được xác định bằng cách tính điểm của các nhân tố (Factor score).

Nhân tố thứ i, được xác định

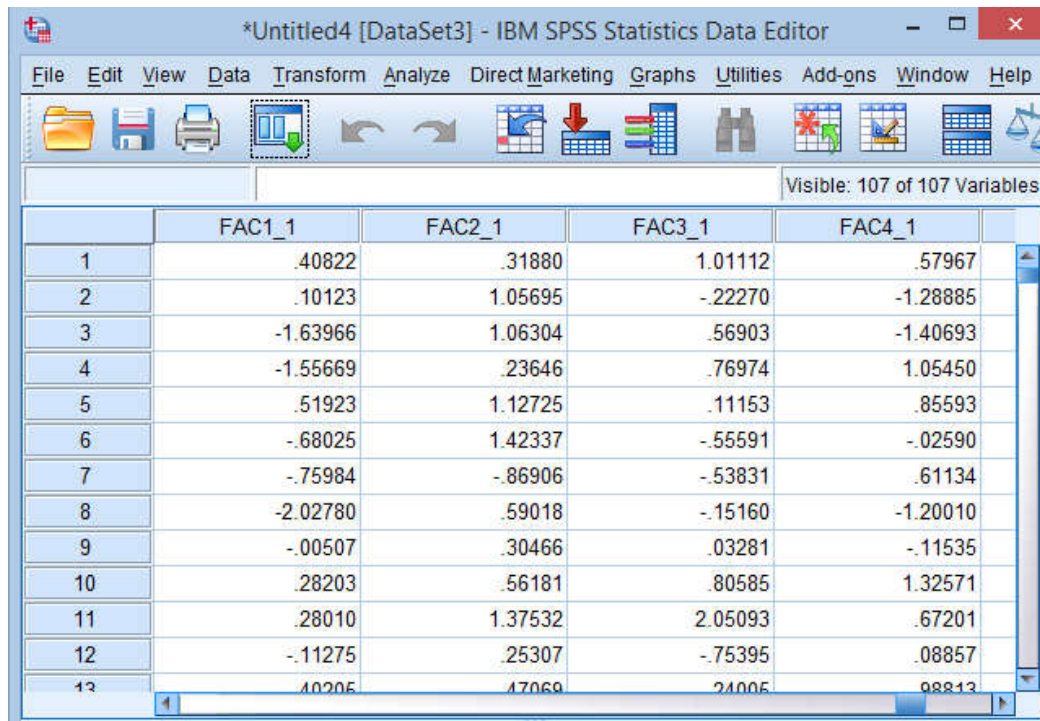
$$F_i = w_{i1} X_1 + w_{i2} X_2 + \dots + w_{ik} X_k$$

Với

+) w_{ik} : Hệ số nhân tố được trình bày trong ma trận hệ số nhân tố (Component Score Coefficient).

+) X_i : Biến quan sát trong nhận tổ i .

Đối với các biến độc lập, SPSS tính sẵn trong Data view khi phân tích nhân tố khám phá như hình 51

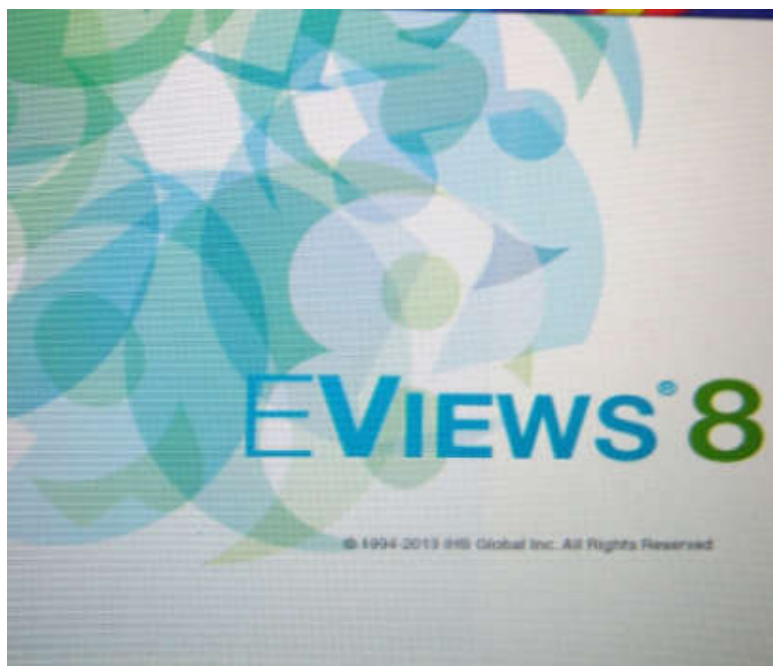


	FAC1_1	FAC2_1	FAC3_1	FAC4_1
1	.40822	.31880	1.01112	.57967
2	.10123	1.05695	-.22270	-1.28885
3	-1.63966	1.06304	.56903	-1.40693
4	-1.55669	.23646	.76974	1.05450
5	.51923	1.12725	.11153	.85593
6	-.68025	1.42337	-.55591	-.02590
7	-.75984	-.86906	-.53831	.61134
8	-2.02780	.59018	-.15160	-1.20010
9	-.00507	.30466	.03281	-.11535
10	.28203	.56181	.80585	1.32571
11	.28010	1.37532	2.05093	.67201
12	-.11275	.25307	-.75395	.08857
13	.40205	.47069	.24005	.98813

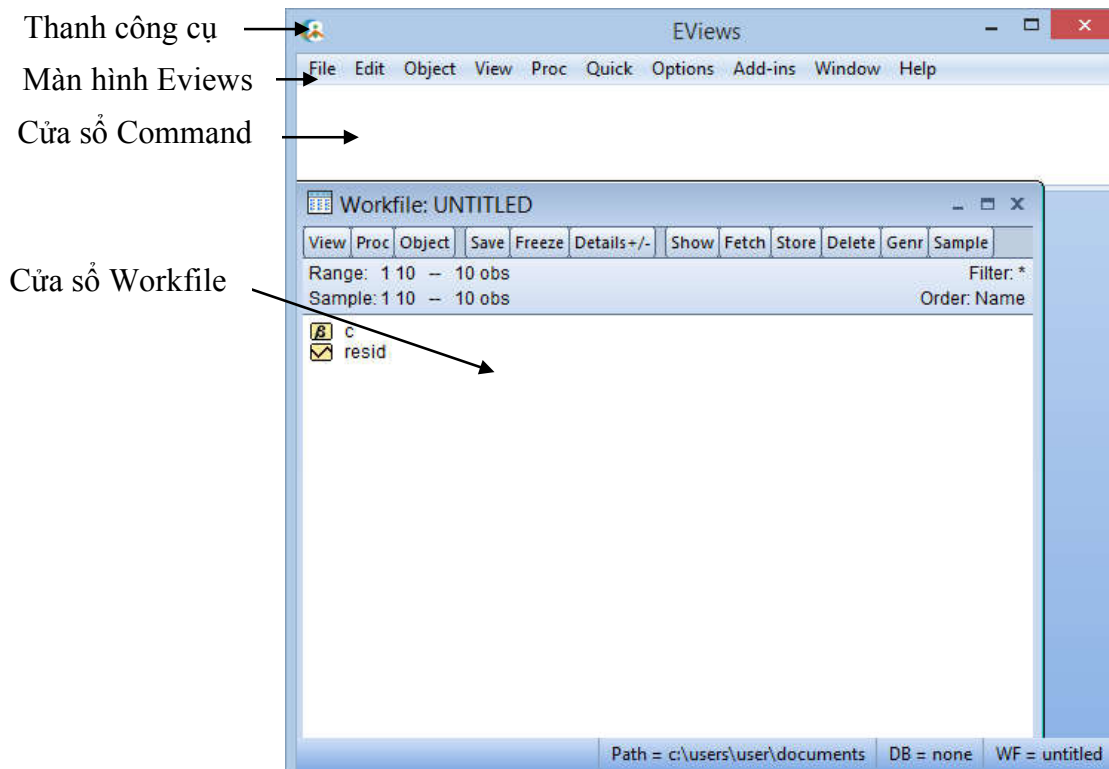
Hình 51

Trong hình 51, FAC1_1 là F1, FAC2_1 là F2, FAC3_1 là F3, FAC4_1 là F4, được xác định theo cách tính điểm của nhân tố, tương tự cho các nhân tố còn lại.

Phụ lục 3. HƯỚNG DẪN SỬ DỤNG PHẦN MỀM EVIEWS 8.0



1. Màn hình Eviews



Hình 1

2. Các kiểu dữ liệu thường dùng

2.1. Số liệu theo thời gian: là các số liệu thu thập tại nhiều thời điểm khác nhau trên cùng một đối tượng. Chẳng hạn như số liệu về GDP bình quân của Việt Nam từ 1998 – 2006 được cho trong bảng sau:

Năm	1998	1999	2000	2001	2002	2003	2004	2005	2006
GDP	360	374	401	413	440	489	553	618	655

Bảng 1

2.2. Số liệu chéo: là số liệu thu thập tại một thời điểm ở nhiều nơi, địa phương, đơn vị, khác nhau. Chẳng hạn như số liệu về GDP bình quân trong năm 2006 của các nước Brunei, Campuchia, Indonesia, Lào, Malaysia, Myanmar, Philippines, Singapore, Thái Lan, Việt Nam được cho như sau:

Nước	GDP	Nước	GDP
Brunei	30376	Myanmar	230
Campuchia	459	Philippines	1361
Indonesia	1581	Singapore	30162
Lào	570	Thái Lan	2959
Malaysia	5570	Việt Nam	655

Bảng 2

2.3. Số liệu hỗn hợp: là số liệu tổng hợp của hai loại trên, nghĩa là các số liệu thu thập tại nhiều thời điểm khác nhau ở nhiều địa phương, đơn vị khác nhau. Chẳng hạn như số liệu về GDP bình quân của các nước từ 1998 – 2006.

Nước	1998	1999	2000	2001	2002	2003	2004	2005	2006
Brunei	13065	14511	18465	16820	17135	18788	21989	25759	30376
Campuchia	255	281	285	302	317	333	373	430	459
Indonesia	516	746	807	773	928	1100	1176	1283	1581
Lào	255	286	329	326	329	378	432	485	570
Malaysia	3254	3485	3844	3665	3884	4161	4652	5042	5570
Myanmar	134	173	178	129	130	197	199	219	230
Philippines	910	1019	994	914	966	982	1049	1168	1361
Singapore	21009	20909	23075	20724	21210	22157	25345	26839	30162
Thái Lan	1829	1985	1967	1836	1999	2233	2484	2659	2959
Việt Nam	360	374	401	413	440	489	553	618	655

Bảng 3

3. Nhập dữ liệu

3.1. Nhập trực tiếp vào Eview

Để minh họa cho phần này, ta xét các ví dụ sau:

Ví dụ 1. Bảng 4 dưới đây cho biết số liệu về GDP bình quân đầu người của Việt Nam trong các năm 1998 – 2006.

Năm	1998	1999	2000	2001	2002	2003	2004	2005	2006
GDP	360	374	401	413	440	489	553	618	655

Bảng 4

Ví dụ 2. Bảng 5 dưới đây cho biết số liệu về doanh số của một công ty.

Năm	Quý			
	I	II	III	IV
2001	5280	4138	3959	7810
2002	6250	4565	4770	8712
2003	5883	5286	6142	9280
2004	7523	5758	5714	10144
2005	6783	6268	7618	11567

Bảng 5

Ví dụ 3. Bảng 6 dưới đây cho biết số liệu về năng suất (Y, đơn vị tạ/ha) và mức phân bón (X, đơn vị tạ/ha) cho một loại cây trồng tính trên một ha trong 10 năm từ 1988 đến 1997.

Năm	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
X	6	10	12	14	16	18	22	24	26	32
Y	40	44	46	48	52	58	60	68	74	80

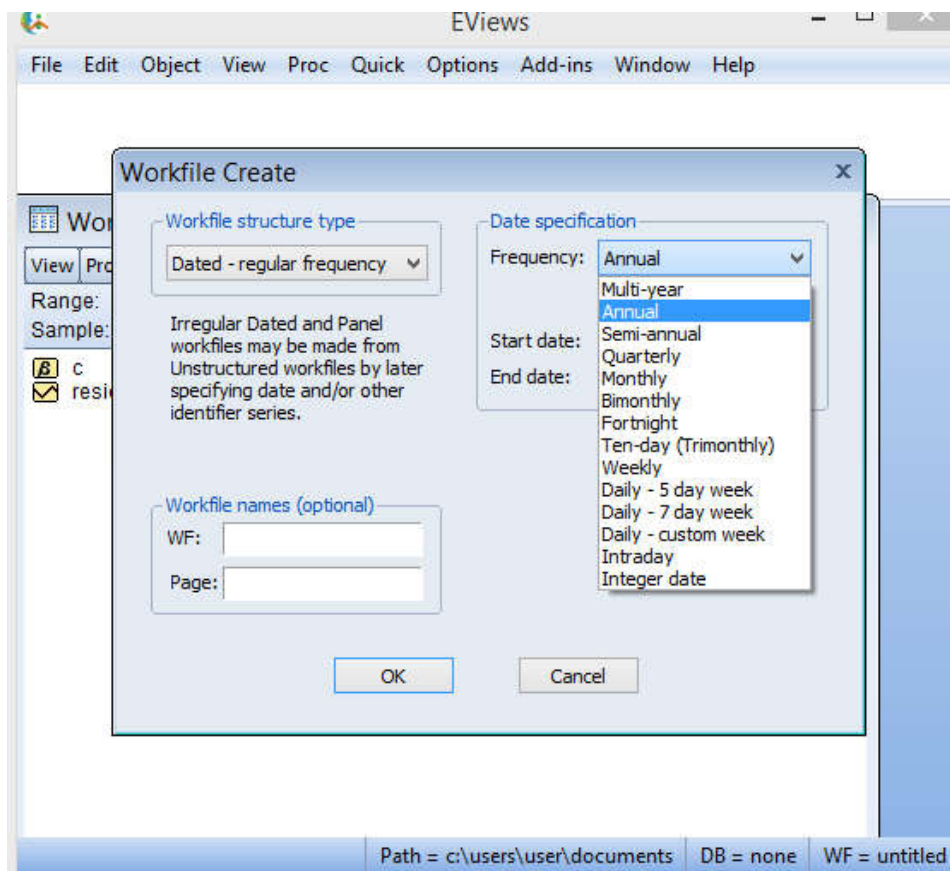
Bảng 6

Ví dụ 4. Bảng 7 dưới đây cho biết số liệu về doanh thu (Y), chi phí cho quảng cáo (X₂), tiền lương của nhân viên tiếp thị (X₃) của 12 công nhân (đơn vị triệu đồng).

STT	1	2	3	4	5	6	7	8	9	10	11	12
Y _i	127	149	106	163	102	180	161	128	139	144	159	138
X _{2i}	18	25	19	24	15	26	25	16	17	23	22	15
X _{3i}	10	11	6	16	7	17	14	12	12	12	14	15

Bảng 7

Mở Eview, để nhập dữ liệu: Chọn **File** → **New** → **Workfile**, ta có màn hình như sau:



Hình 2

Tùy vào kiểu dữ liệu cần khảo sát, ta có thể chọn được các kiểu sau :

Dated – regular frequency

Multi – year : Số liệu nhiều năm

Annual : Số liệu năm

Semi – Annual : Số liệu nửa năm

Quarterly : Số liệu theo quý

Monthly : Số liệu theo từng tháng

Bimonthly : Mỗi tháng 2 lần/2 tháng 1 lần

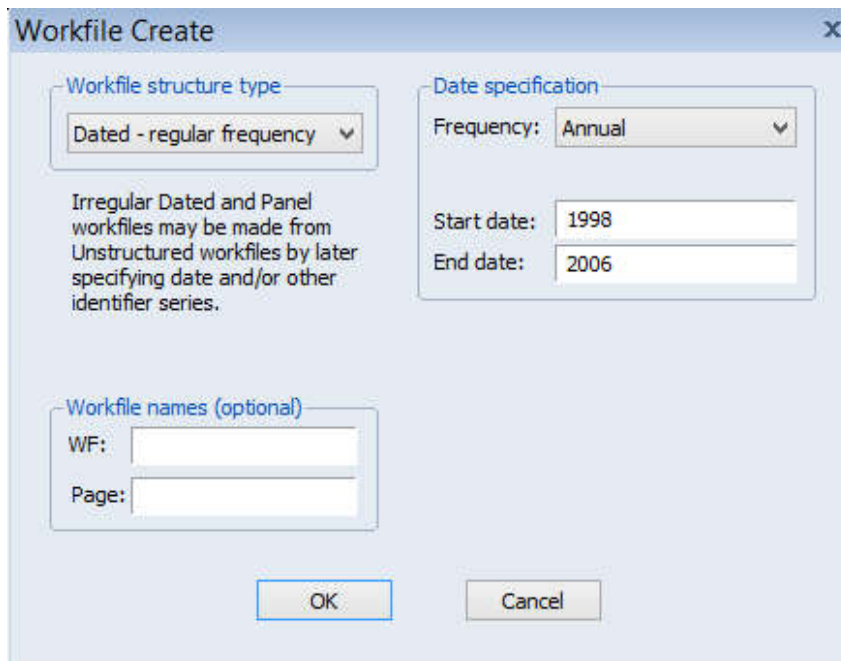
Fortnight : Hai tuần lễ/15 ngày

Ten – day (Trimonthly) :

Weekly : Số liệu theo từng tuần

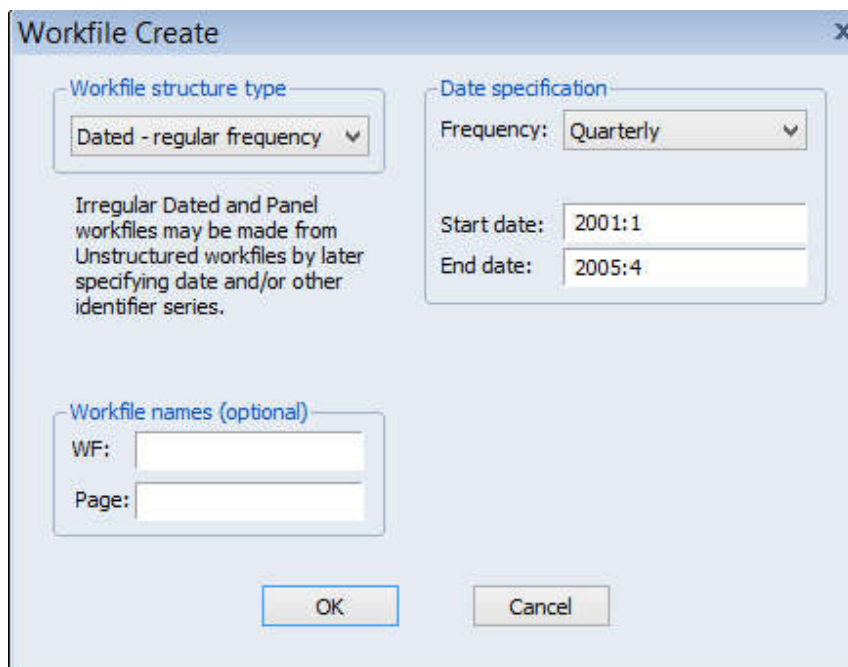
Unstructure / Undate : Số liệu chéo

Để nhập dữ liệu ở ví dụ 1, ta chọn các khai báo như trong hình 3 như sau:



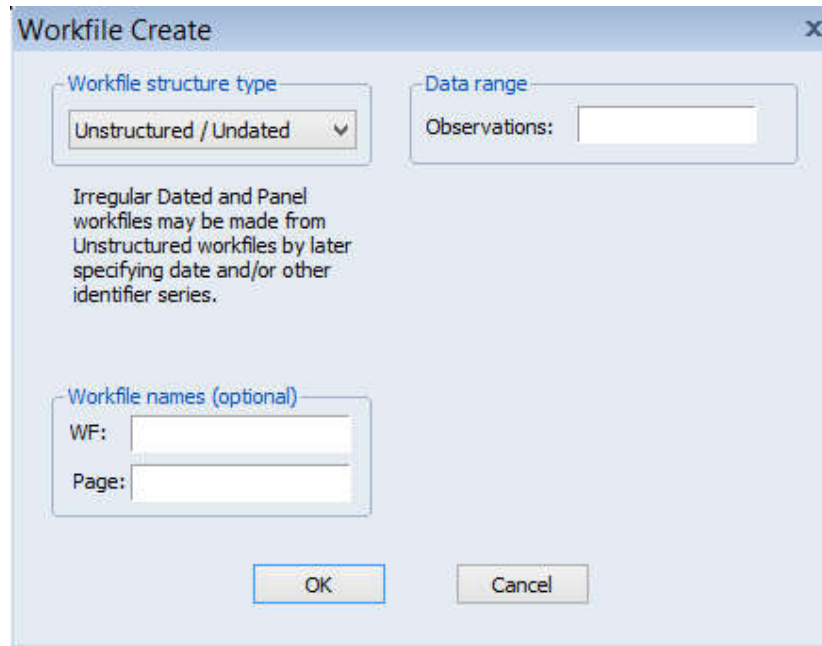
Hình 3

Để nhập dữ liệu ở ví dụ 2, ta chọn các khai báo như trong hình 4



Hình 4

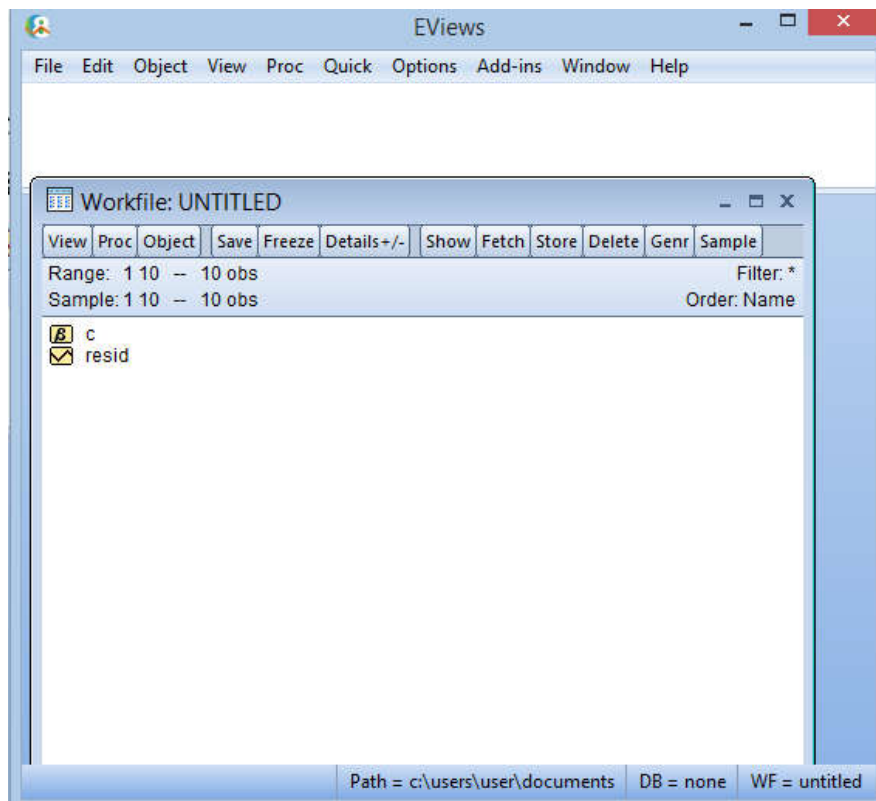
Để nhập dữ liệu cho ví dụ 3, ví dụ 4, ta có thể khai báo như trong hình 5.



Hình 5

Trong ô **Observations** ta nhập cỡ mẫu (số các quan sát)

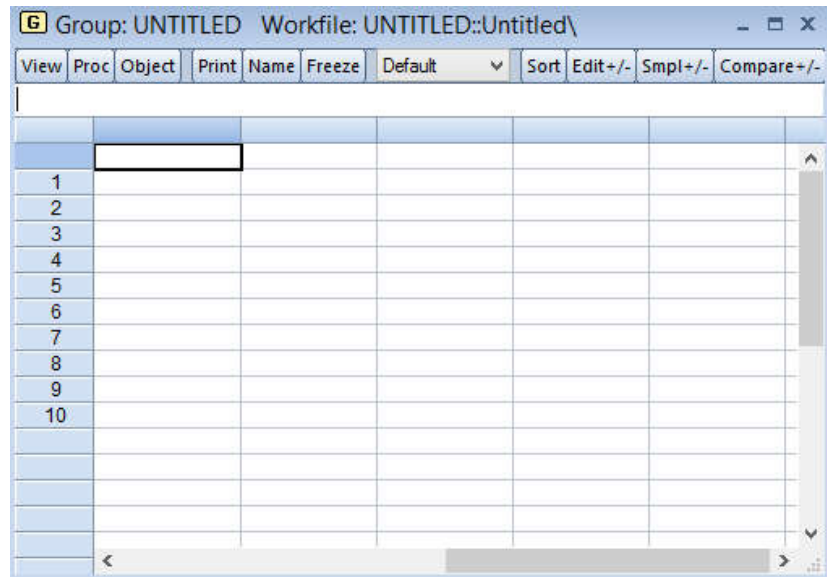
Chẳng hạn như trong ví dụ 3, ta nhập 10 rồi **nhấn OK** ta được hình 6



Hình 6

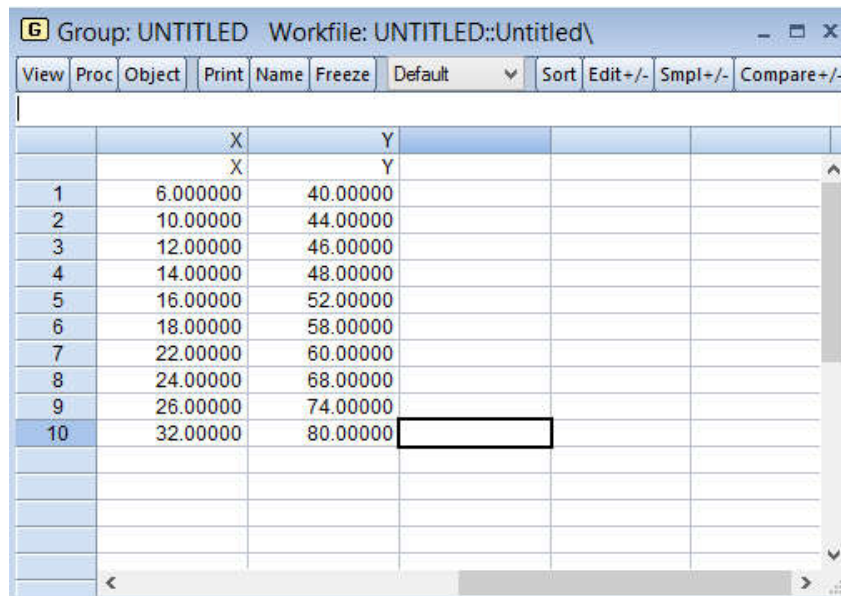
Để nhập số liệu ta chọn : **Quick** → **Empty Group (Edit Series)**, màn hình xuất hiện một cửa sổ như hình 7. Trong đó

- Cột **obs** ghi thứ tự quan sát.
- Các cột kế tiếp để khai báo các biến và nhập số liệu.



Hình 7

Ví dụ nhập số liệu cho biến Y vào cột số 2, ta nhấp chuột vào đầu cột này và gõ tên biến Y sau đó nhấp Enter và lần lượt gõ các giá trị vào các ô bên dưới có ghi chữ NA. Chẳng hạn như trong ví dụ 3 và ví dụ 4, ta khai báo và nhập số liệu tuần tự như trong các hình sau :



Hình 8

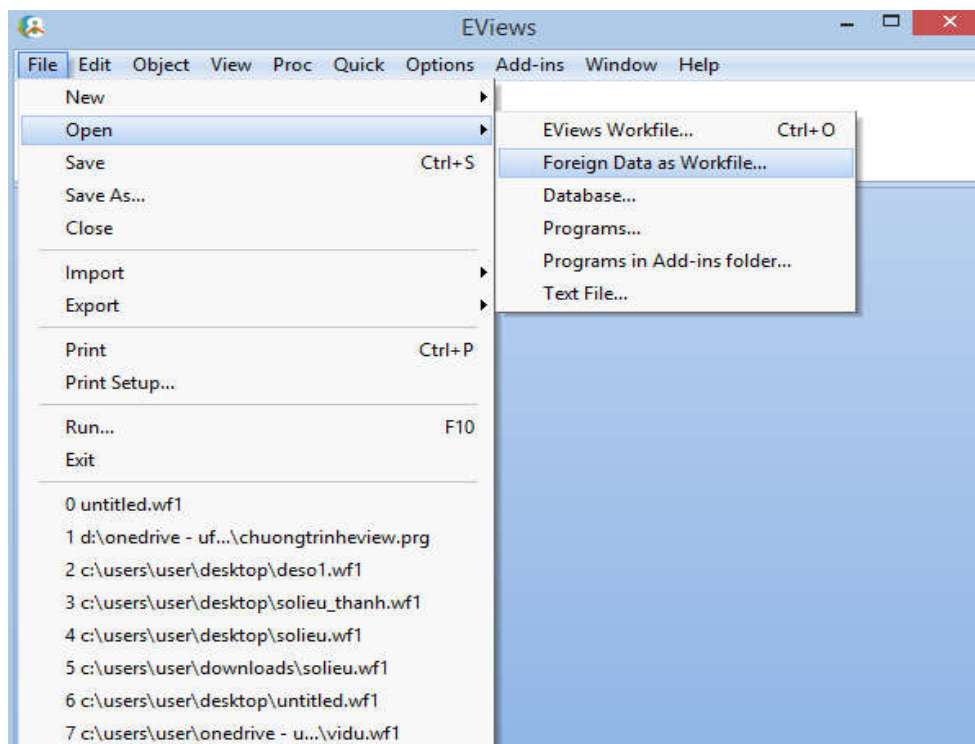
	Y	X2	X3
1	127.0000	18.00000	10.00000
2	149.0000	25.00000	11.00000
3	106.0000	19.00000	6.000000
4	163.0000	24.00000	16.00000
5	102.0000	15.00000	7.000000
6	180.0000	26.00000	17.00000
7	161.0000	25.00000	14.00000
8	128.0000	16.00000	12.00000
9	139.0000	17.00000	12.00000
10	144.0000	23.00000	12.00000
11	159.0000	22.00000	14.00000
12	138.0000	15.00000	15.00000

Hình 9

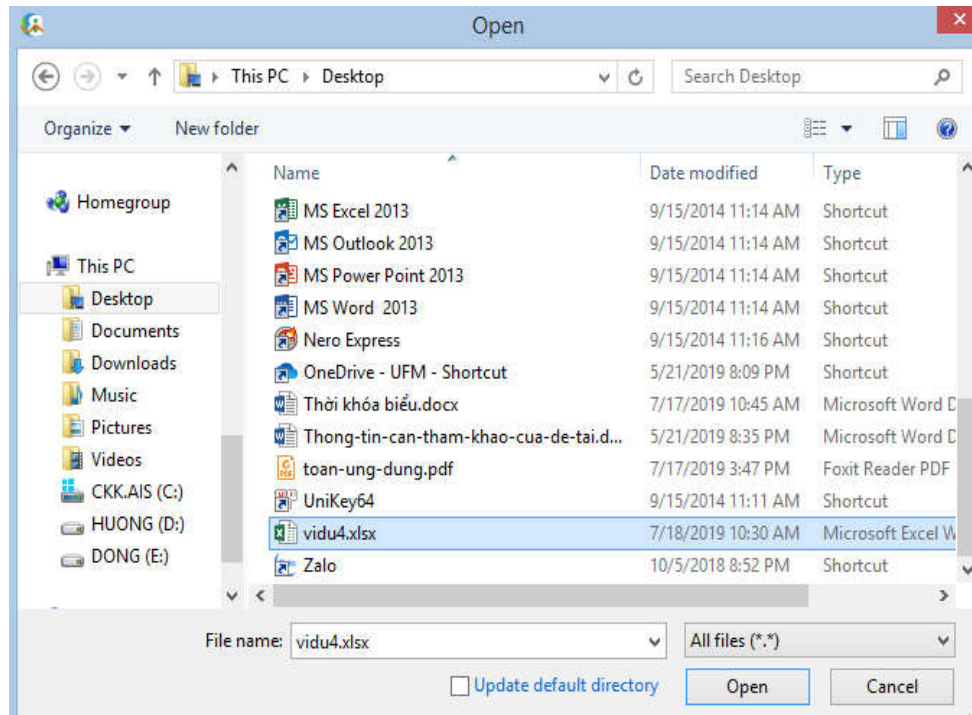
3.2. Nhập từ Excel và Word có sẵn

Giả sử ta có sẵn File Excel **vidu 4.xls** chứa số liệu của ví dụ 4. Khi đó ta thực hiện các bước Import sau:

Mở chương trình **Eviews** chọn **File** → **Open** → **Foreign Data as Workfile...** như sau

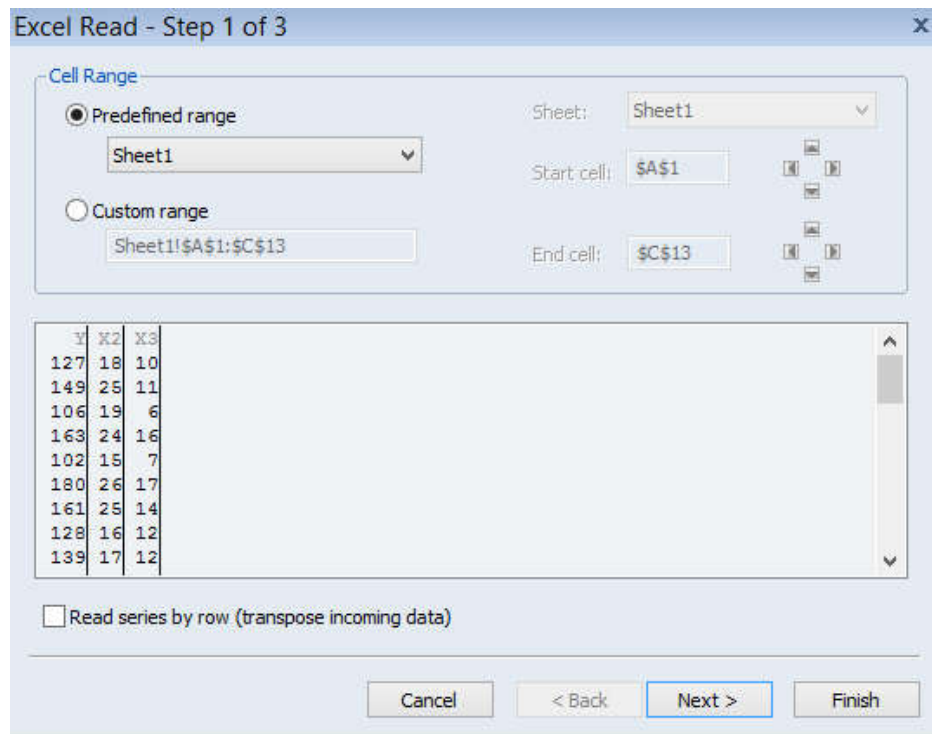


Hình 10



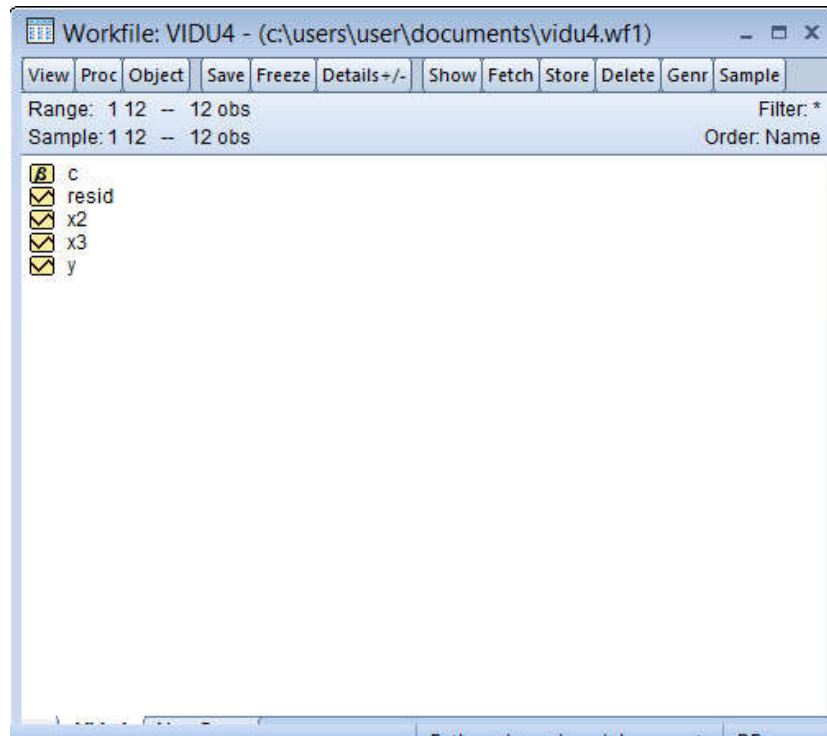
Hình 11

Chọn **Open** ta được kết quả như trong hình 12. Trong cửa sổ này chúng ta thấy có hai cột số liệu của X và Y tương ứng trong Sheet1 của File vidu 3.xls



Hình 12

Cuối cùng chọn **Finish** ta được kết quả như trong hình 13

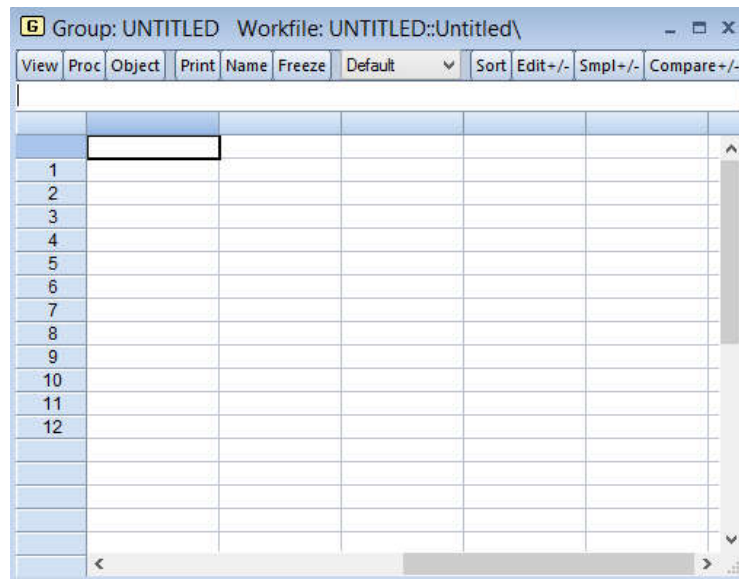


Hình 13

Lưu ý. Các bước trên được gọi là trích lọc dữ liệu từ một file dữ liệu có sẵn.

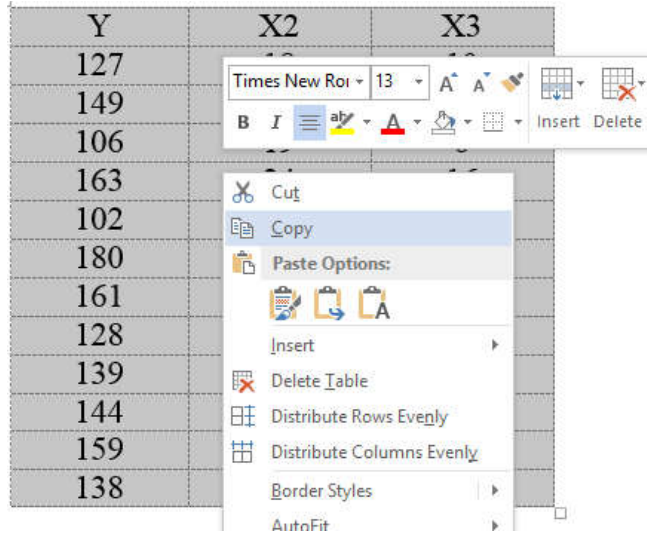
Ta có thể thực hiện copy trực tiếp từ một file Word hoặc Excel

Mở cửa sổ Group của Eview



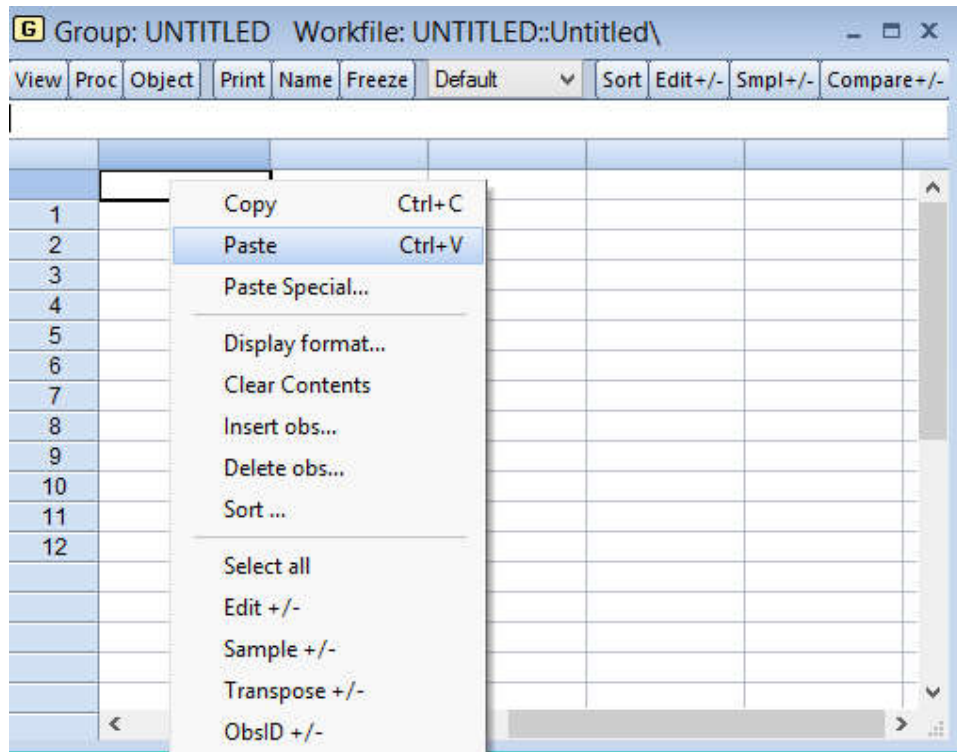
Hình 14

Từ file excel hoặc file word bôi đen rồi copy và paste vào file trên. Chẳng hạn ta có file word ta thực hiện như sau:



Hình 15

Ta paste vào của sổ Group như sau



Hình 16

Và được kết quả như sau:

	Y	X2	X3
	Y	X2	X3
1	127.0000	18.00000	10.00000
2	149.0000	25.00000	11.00000
3	106.0000	19.00000	6.000000
4	163.0000	24.00000	16.00000
5	102.0000	15.00000	7.000000
6	180.0000	26.00000	17.00000
7	161.0000	25.00000	14.00000
8	128.0000	16.00000	12.00000
9	139.0000	17.00000	12.00000
10	144.0000	23.00000	12.00000
11	159.0000	22.00000	14.00000
12	138.0000	15.00000	15.00000

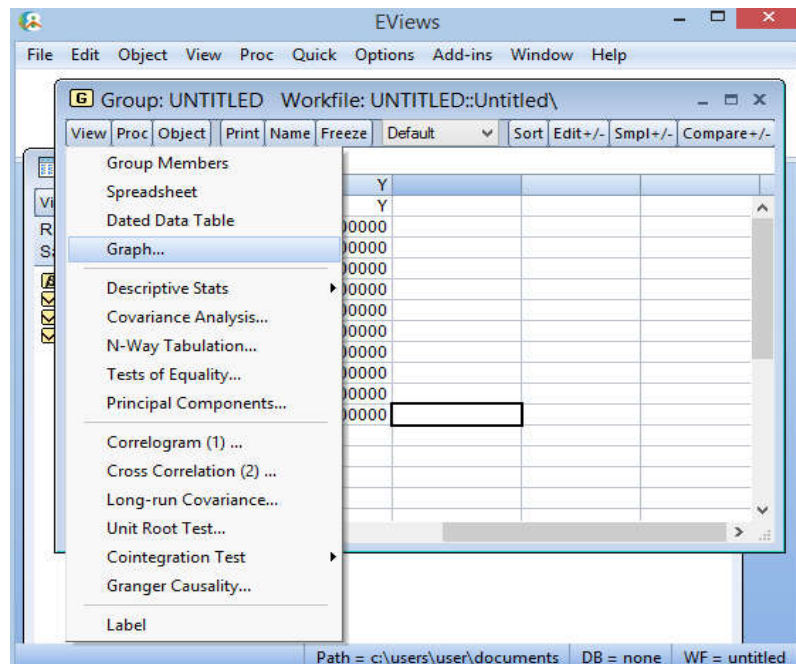
Hình 17

4. Vẽ đồ thị

4.1. Vẽ biểu đồ phân tán số liệu.

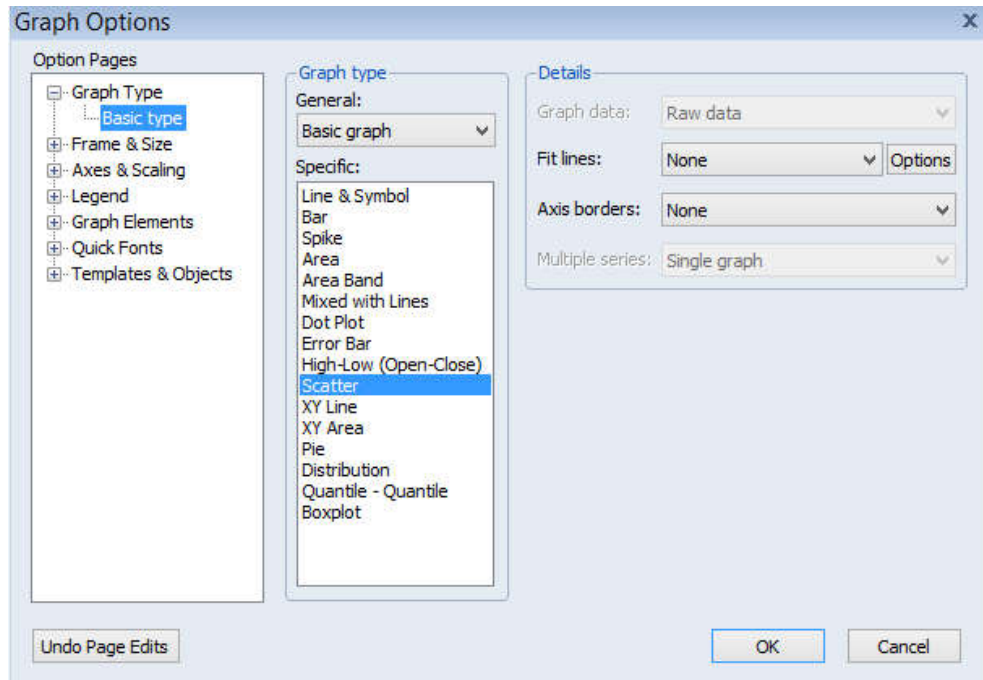
Mục đích của việc vẽ đồ thị này cho phép ta đánh giá sơ bộ về mối quan hệ cũng như hình dung được dạng hàm (mô hình) giữa hai biến với nhau. Để vẽ đồ thị phân tán của hai biến, chẳng hạn như trong ví dụ 3 ta vẽ đồ thị phân tán của Y và X.

Từ cửa sổ **Eviews** chọn **Quick**→**Graph**



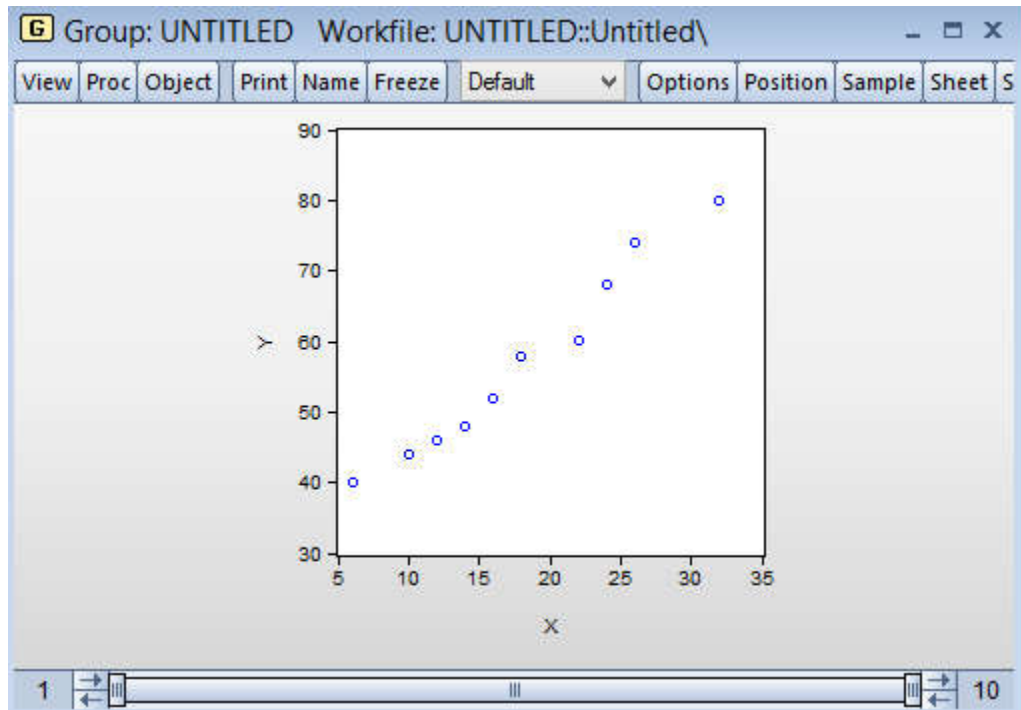
Hình 18

Ta được



Hình 19

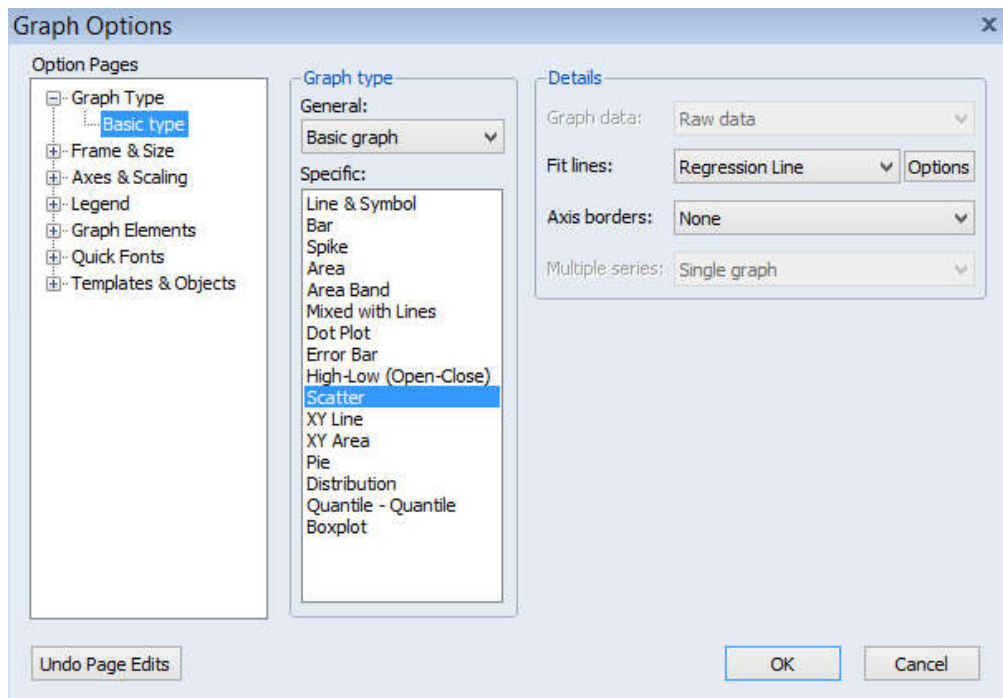
Ta chọn **Scatter** rồi nhấn **Ok**, ta được đồ thị phân tán dữ liệu như sau



Hình 20

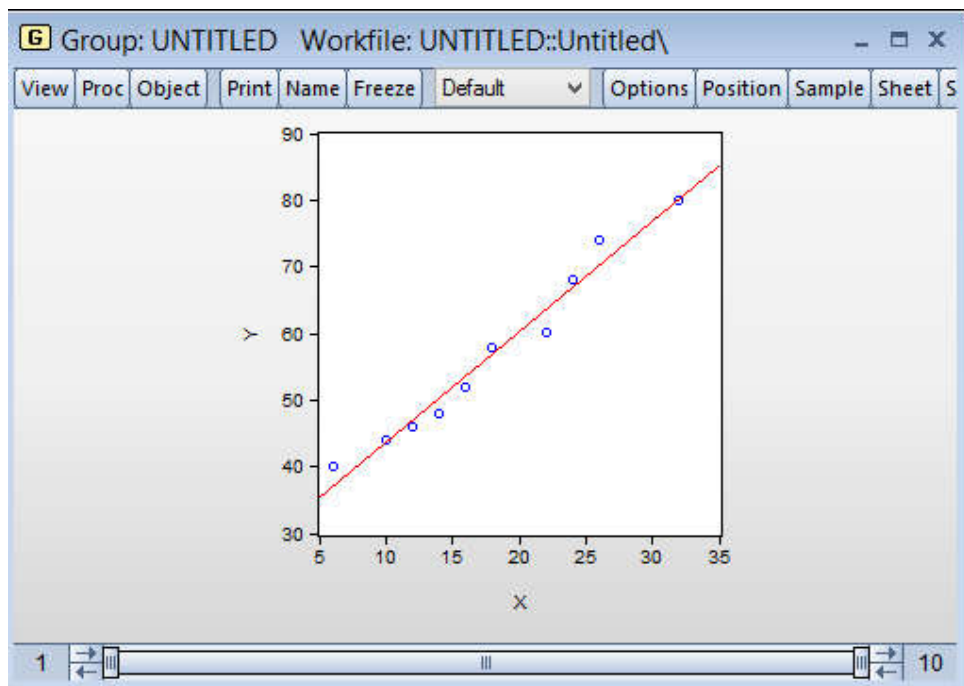
Làm tương tự như các bước trên ta có thể vẽ các loại đồ thị khác.

4.2. Vẽ đường hồi quy tuyến tính.



Hình 21

Thực hiện các bước tương tự như trên. Ta chọn **Scatter**→**Regression line** rồi nhấn **Ok**, ta được đồ thị đường hồi quy như sau:



Hình 22

5. Tìm hàm hồi quy tuyến tính mẫu (SRF)

Muốn tìm hàm hồi quy tuyến tính mẫu của Y theo X chẳng hạn như trong ví dụ 3 có nhiều cách làm sau đây tôi chỉ giới thiệu một cách đơn giản nhất.

Từ cửa sổ **Command** ta gõ dòng lệnh **ls y c x** và nhấn **Enter**. Ta có bảng hồi quy sau mà ta gọi là bảng **Equation**

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	27.12500	1.979265	13.70458	0.0000
X	1.659722	0.101321	16.38082	0.0000

R-squared	0.971049	Mean dependent var	57.00000
Adjusted R-squared	0.967430	S.D. dependent var	13.47426
S.E. of regression	2.431706	Akaike info criterion	4.791920
Sum squared resid	47.30556	Schwarz criterion	4.852437
Log likelihood	-21.95960	Hannan-Quinn criter.	4.725533
F-statistic	268.3312	Durbin-Watson stat	1.783613
Prob(F-statistic)	0.000000		

Hình 23

Các kết quả ở bảng trong **hình 23** lần lượt là

- **Dependent Variable** : Tên biến phụ thuộc
- **Method: Least Squares** : Phương pháp bình phương tối thiểu (nhỏ nhất).
- **Date – Time** : Ngày giờ thực hiện
- **Sample** : Số liệu mẫu 1 – 10
- **Included observations** : Cỡ mẫu là 10 (số các quan sát)
- **Cột Variable** : Các biến giải thích có trong mô hình (trong đó C là hệ số bị chặn)
- **Cột Coefficient** : Giá trị các hệ số hồi quy $\hat{\beta}_1; \hat{\beta}_2$.
- **Cột Std. Error** : Sai số chuẩn của các hệ số hồi quy.

$$se(\hat{\beta}_1) = \sqrt{\text{var}(\hat{\beta}_1)}; se(\hat{\beta}_2) = \sqrt{\text{var}(\hat{\beta}_2)}$$

- **Cột t – Statistic** : Giá trị thống kê t tương ứng: $t_1 = \frac{\hat{\beta}_1}{se(\hat{\beta}_1)}; t_2 = \frac{\hat{\beta}_2}{se(\hat{\beta}_2)}$

(Trong đó t là đại lượng ngẫu nhiên có phân phối Student với bậc tự do (n – 2)).

- **Cột Prob.** : Giá trị xác suất (p – value) của thống kê t tương ứng

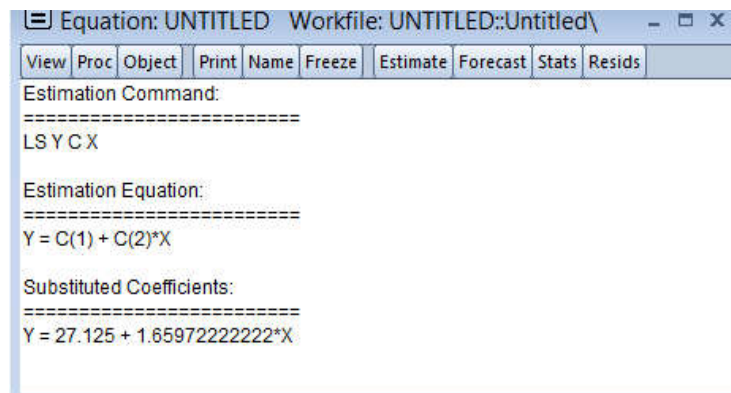
$$p_value_1 = P(t > t_1); p_value_2 = P(t > t_2)$$

- **R – Squared** : Hệ số xác định mô hình (R^2)
- **Adjusted R – Squared** : Hệ số xác định có hiệu chỉnh (\bar{R}^2)
- **S.E. of regression** : Giá trị ước lượng cho σ : $\hat{\sigma}$ (sai số chuẩn của hồi quy)
- **Sum squared resid** : Tổng bình phương các sai lệch (phần dư) (RSS)
- **Log likelihood** : Tiêu chuẩn ước lượng hợp lý (Logarit của hàm hợp lý)
- **Durbin – Watson stat** : Thống kê Durbin – Watson
- **Mean dependent var** : Giá trị trung bình mẫu của biến phụ thuộc
- **S.D. dependent var** : Độ lệch chuẩn mẫu của biến phụ thuộc
- **Akaike info criterion** : Tiêu chuẩn Akaike
- **Schwarz info criterion** : Tiêu chuẩn Schwarz
- **F – Statistic** : Giá trị của thống kê F
- **Prob (F – Statistic)** : Giá trị xác suất (p-value) của thống kê F tương ứng

$$p_value = P(F > F_statistic)$$

Với F là biến ngẫu nhiên có phân phối Fisher có bậc tự do $(k - 1, n - k)$.

Muốn thể hiện đường hồi quy. Từ bảng **Equation** → **View** → **Representations**, ta có kết quả sau:



Hình 24

6. Một số hàm trong Views

LOG(X) : $\ln(X)$

EXP(X) : e^X

ABS(X) : giá trị tuyệt đối của X

SQR(X) : căn bậc 2 của X

@SUM(X) : tổng của các X

@MEAN(X) : giá trị trung bình của X

@VAR(X) : phương sai của X

@COV(X,Y) : hiệp phương sai của X, Y

@COR(X,Y) : hệ số tương quan của X, Y

7. Cách tìm một số dạng hàm hồi quy

Giả sử ta có số liệu của các biến Y và X tại thời điểm t. Nếu tìm hàm hồi quy của Y_t theo X và Y_{t-1} (biến trễ thì câu lệnh sẽ là **y c x y(-1)**).

Giả sử ta có số liệu của các biến Y và X. Nếu tìm hàm hồi quy của $\ln(Y)$ theo $\ln(X)$ thì câu lệnh sẽ là **log(y) c log(x)**.

Giả sử ta có số liệu của các biến Y và X. Nếu tìm hàm hồi quy của Y theo \sqrt{X} thì câu lệnh sẽ là **y c sqr(x)**.

Giả sử ta có số liệu của các biến Y và X. Nếu tìm hàm hồi quy của Y theo e^x thì câu lệnh sẽ là **y c exp(x)**.

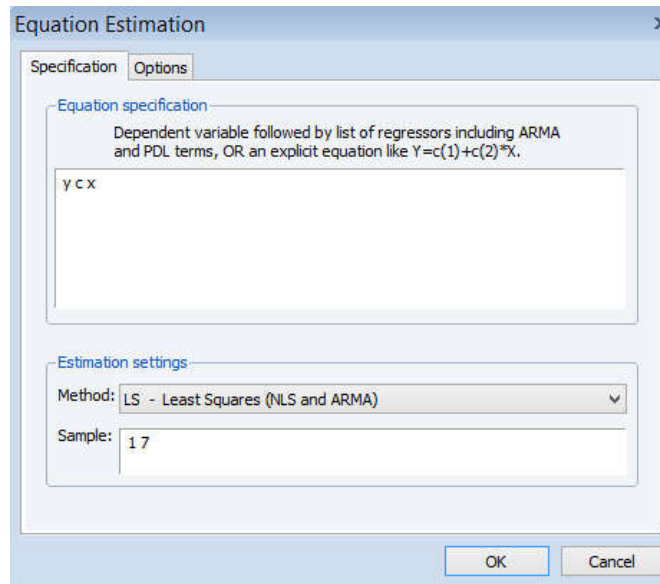
Giả sử ta có số liệu của các biến Y và X. Nếu tìm hàm hồi quy của Y theo X và X^2 thì câu lệnh sẽ là **y c x x^2**.

Giả sử ta có số liệu của các biến Y và X. Nếu tìm phương trình sai phân cấp 1 của Y theo X thì câu lệnh sẽ là **d(y) c d(x)**.

Giả sử ta có số liệu của các biến Y và X. Nếu tìm phương trình sai phân cấp k của Y theo X thì câu lệnh sẽ là **d(y,k) c d(x,k)**.

Nếu cần tìm hàm hồi quy nhưng không sử dụng hết các quan sát của mẫu, chẳng hạn ta tìm hàm hồi quy của Y theo X trong ví dụ 3 nhưng ta chỉ sử dụng 7 cặp quan sát đầu tiên. Khi đó ta thực hiện các thao tác như sau:

Từ bảng **Equation chọn Estimate**, ta có màn hình sau. Ta chỉnh 10 thành 7



Hình 25

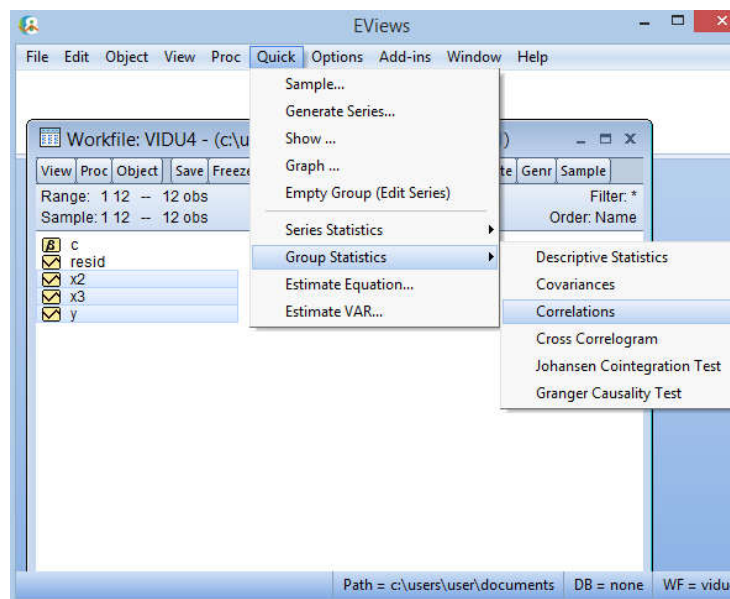
8. Tìm ma trận tương quan và ma trận hiệp phương sai của các hệ số hồi quy

8.1. Ma trận tương quan giữa các biến

Giả sử ta có mẫu gồm các biến Y, X2, X3 cho trong ví dụ 4. Để tìm ma trận tương quan của các biến này ta thực hiện như sau:

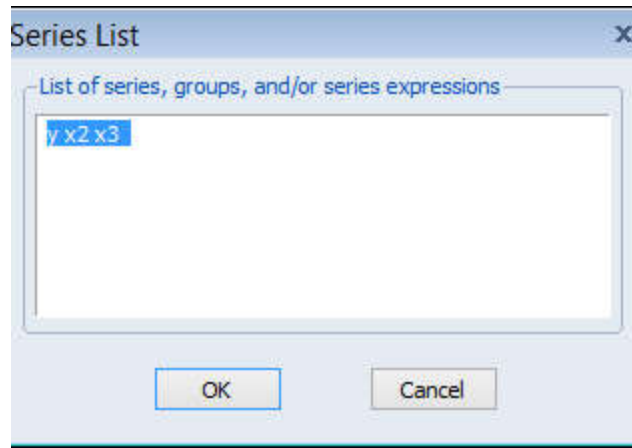
Từ cửa sổ EViews chọn **Quick** → **Group Statistics** → **Correlations**.

Khi đó màn hình xuất hiện như sau:



Hình 26

Nhấp chuột sẽ xuất hiện cửa sổ sau



Hình 27

Sau đó **nhấn OK**, ta được ma trận tương quan như sau

	Y	X2	X3
Y	1.000000	0.782281	0.904627
X2	0.782281	1.000000	0.480173
X3	0.904627	0.480173	1.000000

Hình 28

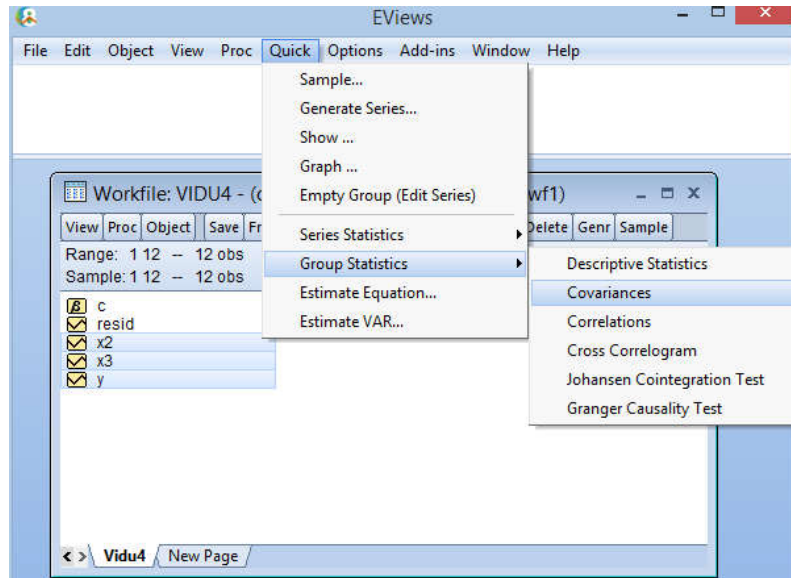
Ý nghĩa: Ma trận tương quan (**Correlation**) cho biết xu thế và mức độ tương quan tuyến tính giữa hai biến trong mô hình. Nhìn vào bảng ma trận tương quan ở trên ta thấy hệ số tương quan của X2 và X3 là 0.480173 khá nhỏ điều đó có nghĩa là X2 và X3 có tương quan tuyến tính ở mức độ yếu và tương quan thuận.

8.2. Ma trận hiệp phương sai giữa các biến

Giả sử ta có mẫu gồm các biến Y, X2, X3 cho trong ví dụ 4. Để tìm ma trận tương quan của các biến này ta thực hiện như sau:

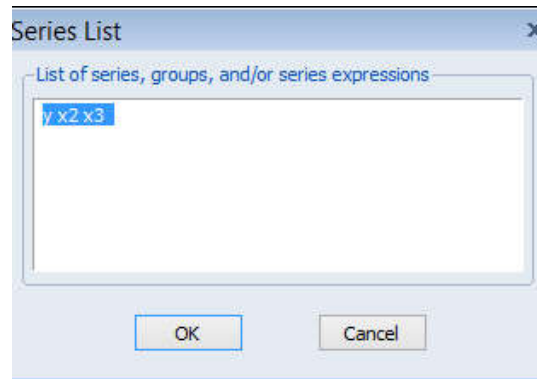
Từ cửa sổ **Eviews** chọn **Quick** → **Group Statistics** → **Covariances**

Khi đó màn hình xuất hiện như sau:



Hình 29

Nhấp chuột sẽ xuất hiện cửa sổ sau



Hình 30

Sau đó **nhấn OK**, ta được ma trận tương quan như sau

View	Proc	Object	Print	Name	Freeze	Sample	Sheet	Stats	Spec
Covariance									
				Y	X2	X3			
		Y		493.7222	69.69444	64.52778			
		X2		69.69444	16.07639	6.180556			
		X3		64.52778	6.180556	10.30556			

Hình 31

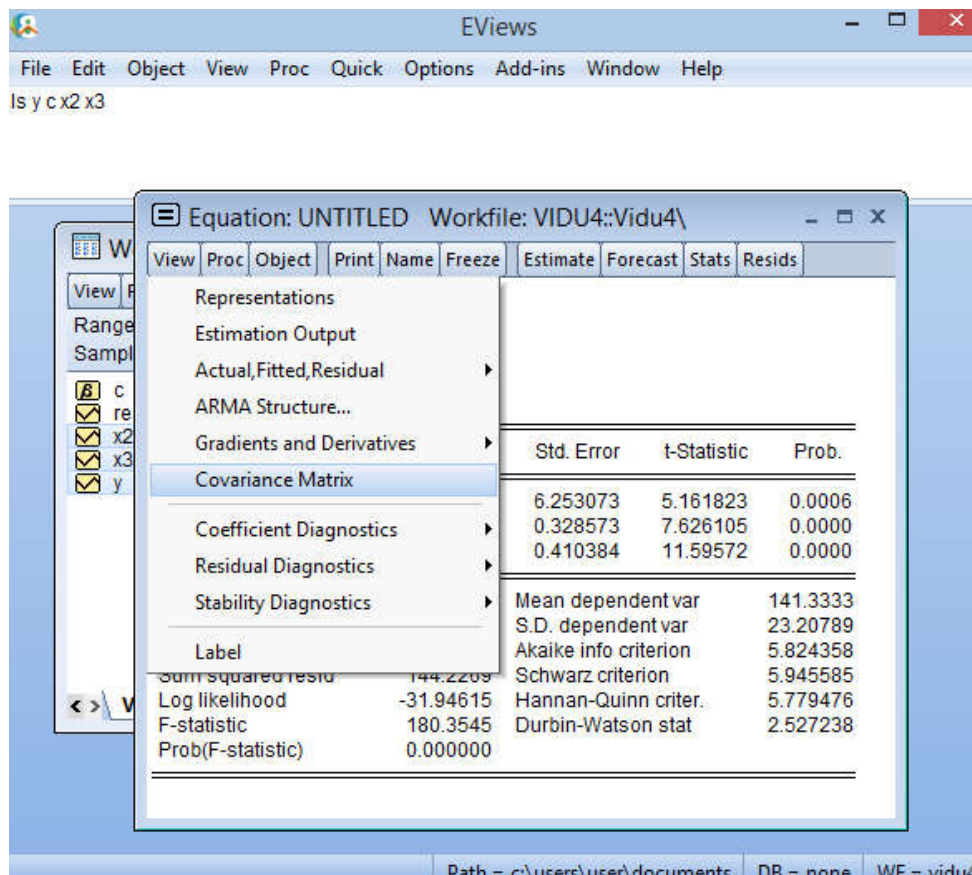
Ý nghĩa: Ma trận hiệp phương sai (Covariances) cho biết phương sai các biến nằm trên đường chéo chính.

8.3. Ma trận hiệp phương sai giữa các hệ số hồi quy

Giả sử ta có mẫu gồm các biến Y, X2, X3 cho trong ví dụ 4. Để tìm ma trận hiệp phương sai giữa các hệ số hồi quy, ta thực hiện như sau:

Từ cửa sổ **Equation** chọn **View** → **Covariance Matrix**.

Khi đó màn hình xuất hiện như sau:



Hình 32

Nhấp chuột, ta được ma trận hiệp phương sai giữa các hệ số hồi quy như sau

The screenshot shows the 'Coefficient Covariance Matrix' window. The matrix displays the following data:

	C	X2	X3
C	39.10093	-1.416429	-0.727129
X2	-1.416429	0.107960	-0.064747
X3	-0.727129	-0.064747	0.168415

Hình 33

Ý nghĩa: Ma trận hiệp phương sai của các hệ số hồi quy (**Coefficient Covariance matrix**) cho biết phương sai các hệ số hồi quy nằm trên đường chéo chính, các thành phần còn lại là hiệp phương sai của những hệ số trong mô hình.

Chẳng hạn, ví dụ 4 bên trên. Nhìn vào ma trận hiệp phương sai bên trên ta có phương sai của các hệ số hồi quy là:

$$\text{var}(\hat{\beta}_1) = 39.10093; \text{var}(\hat{\beta}_2) = 0.107960; \text{var}(\hat{\beta}_3) = 0.168415.$$

9. Bài toán tìm khoảng tin cậy cho các hệ số hồi quy (Khoảng tin cậy đối xứng)

Khoảng ước lượng các hệ số hồi quy tổng thể

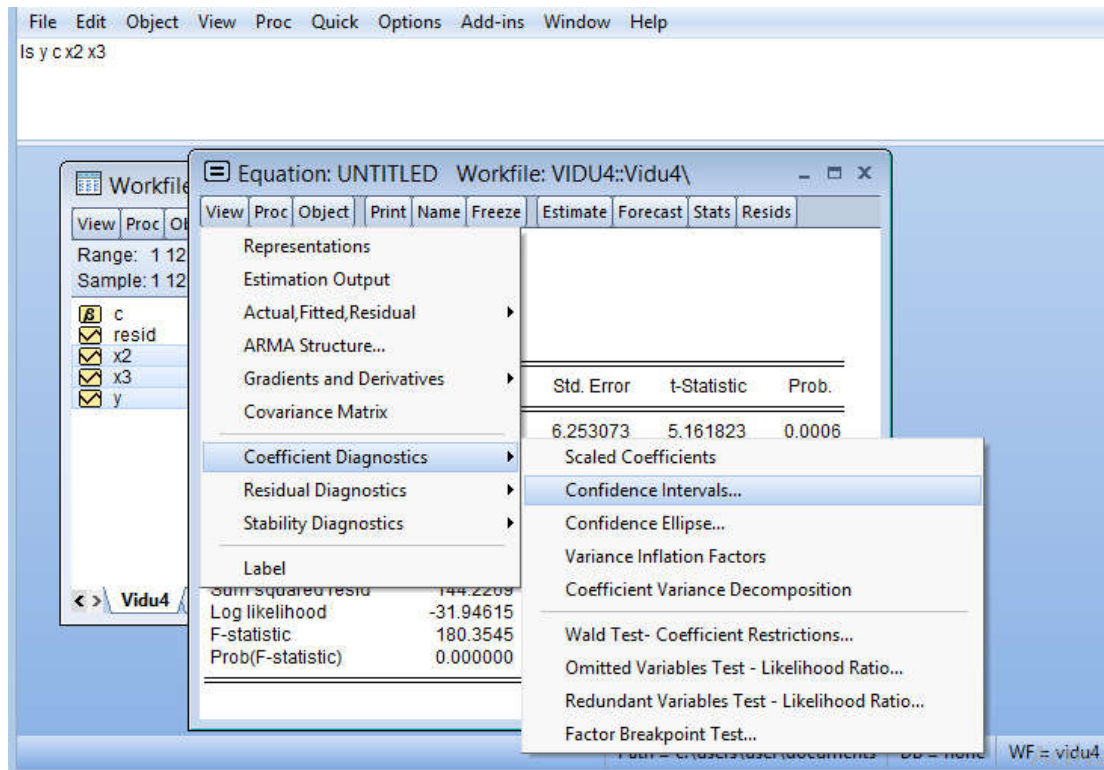
$$\beta_j \in \left[\hat{\beta}_j - Cse(\hat{\beta}_j); \hat{\beta}_j + Cse(\hat{\beta}_j) \right]; j = 1, 2, \dots, k; \quad C = t_{\alpha/2}^{n-k}$$

Trong đó C là giá trị được dò trong bảng phân phối Student với bậc tự do là (n-k).

Giả sử ta có mẫu gồm các biến Y, X2, X3 cho trong ví dụ 4. Để tìm khoảng tin cậy cho các hệ số hồi quy tổng thể, ta thực hiện như sau:

Từ cửa sổ **Equation** chọn **View** → **Coefficient Diagnostics** → **confidence Intervals...**

Khi đó màn hình xuất hiện như sau:



Hình 34

Nhấp chuột, ta được kết quả sau

Equation: UNTITLED Workfile: VIDU4::Vidu4\

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Coefficient Confidence Intervals
Date: 07/18/19 Time: 11:31
Sample: 1 12
Included observations: 12

Variable	Coefficient	90% CI		95% CI		99% CI	
		Low	High	Low	High	Low	High
C	32.27726	20.81467	43.73985	18.13183	46.42270	11.95580	52.59872
X2	2.505729	1.903418	3.108040	1.762446	3.249012	1.437922	3.573536
X3	4.758693	4.006414	5.510973	3.830341	5.687045	3.425015	6.092372

Hình 35

Bảng trên là kết quả ước lượng khoảng tin cậy của các hệ số hồi quy tổng thể ứng với độ tin cậy 90%, 95% và 99%.

10. Bài toán dự báo

Khoảng dự báo giá trị trung bình

$$E(Y | X = X^0) \in [\hat{Y}_0 - Cse(\hat{Y}_0); \hat{Y}_0 + Cse(\hat{Y}_0)]$$

Khoảng dự báo giá trị cá biệt

$$Y_0 \in [\hat{Y}_0 - Cse(Y_0 - \hat{Y}_0); \hat{Y}_0 + Cse(Y_0 - \hat{Y}_0)]$$

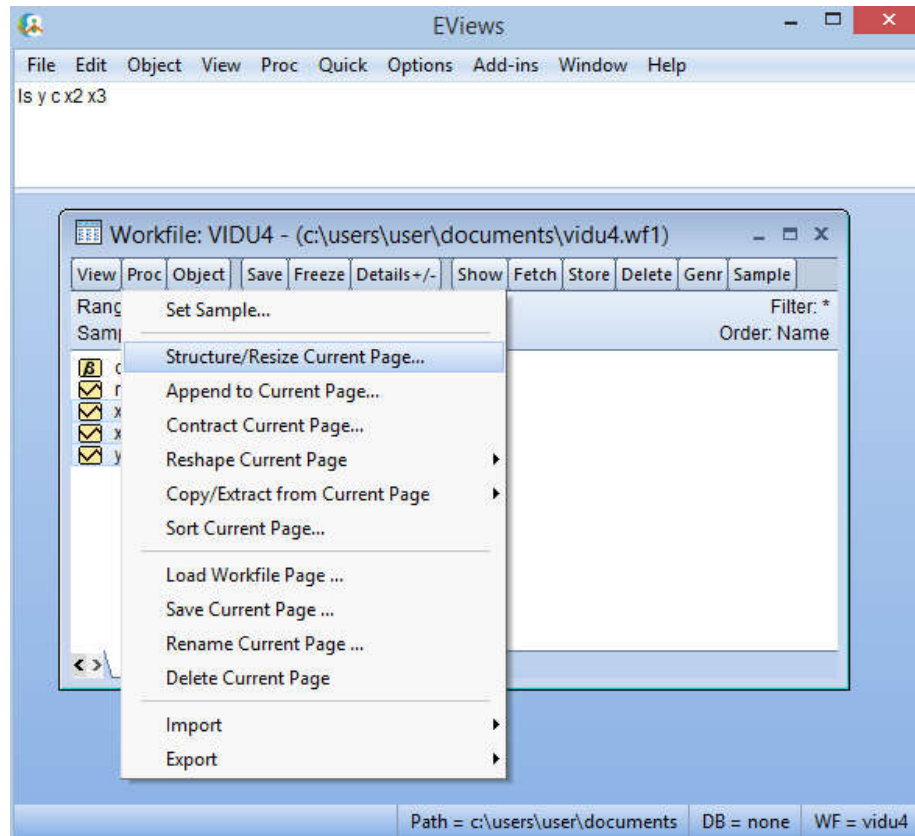
$$\text{Đặt } Y_{DB} = \hat{Y}_0; Se1 = se(Y_0 - \hat{Y}_0); Se2 = se(\hat{Y}_0); C = @qtdist\left(1 - \frac{\alpha}{2}, n - k\right)$$

$$MH = Equation$$

Xét ví dụ 4, để tìm khoảng dự báo giá trị trung bình và giá trị cá biệt của Y khi $X_2 = 20, X_3 = 16$, với độ tin cậy 95%, ta thực hiện như sau:

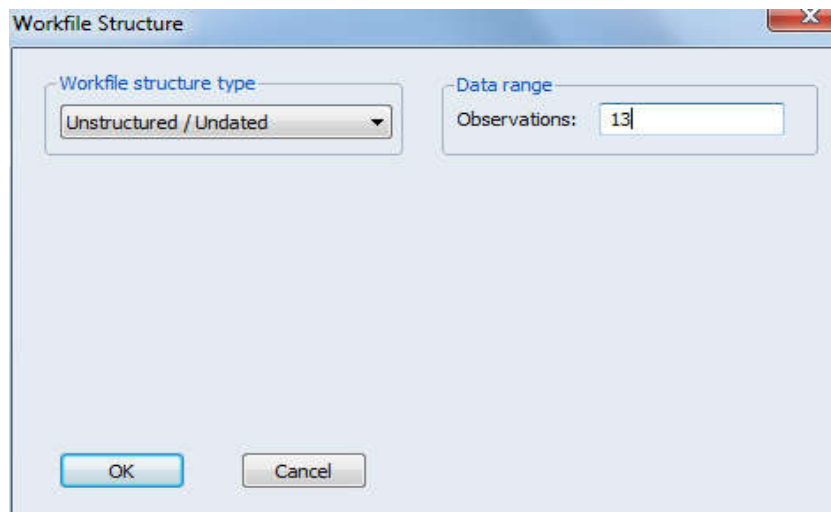
Bước 1. Nhập thêm dữ liệu vào bảng Group để dự báo

Từ bảng Workfile, chọn Proc → Structure/Resize Current Page... Màn hình sau



Hình 36

Nhấp chuột, màn hình sau xuất hiện. Ở ô quan sát (**Observations**) ta điều chỉnh 12 thành 13) như sau:



Hình 37

Nhấp OK. Từ bảng **Group**. Ta chọn **Edit+/-**, sau đó nhập $X_2 = 20, X_3 = 16$ vào hàng số 13 có chữ NA như sau:

	X2	X3	Y
1	18.00000	10.00000	127.0000
2	25.00000	11.00000	149.0000
3	19.00000	6.000000	106.0000
4	24.00000	16.00000	163.0000
5	15.00000	7.000000	102.0000
6	26.00000	17.00000	180.0000
7	25.00000	14.00000	161.0000
8	16.00000	12.00000	128.0000
9	17.00000	12.00000	139.0000
10	23.00000	12.00000	144.0000
11	22.00000	14.00000	159.0000
12	15.00000	15.00000	138.0000
13	20.00000	16.00000	NA

Hình 38

Tắt cửa sổ **Group**.

Bước 2. Tính giá trị $\hat{Y}_0 = Y_{DB}; se(Y_0 - \hat{Y}_0) = se1; se(\hat{Y}_0) = se2$.

Từ bảng **Equation**. Chọn **forecast** màn hình xuất hiện như sau

Equation: UNTITLED Workfile: VIDU4::Untitled\

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: Y
Method: Least Squares
Date: 06/07/13 Time: 21:31
Sample (adjusted): 1 12

Forecast

Forecast of
Equation: UNTITLED Series: Y

Series names
Forecast name: YDB
S.E. (optional): Se1
GARCH(optional):

Method
Static forecast (no dynamics in equation)
 Structural (ignore ARMA)
 Coef uncertainty in S.E. calc

Forecast sample
1 13

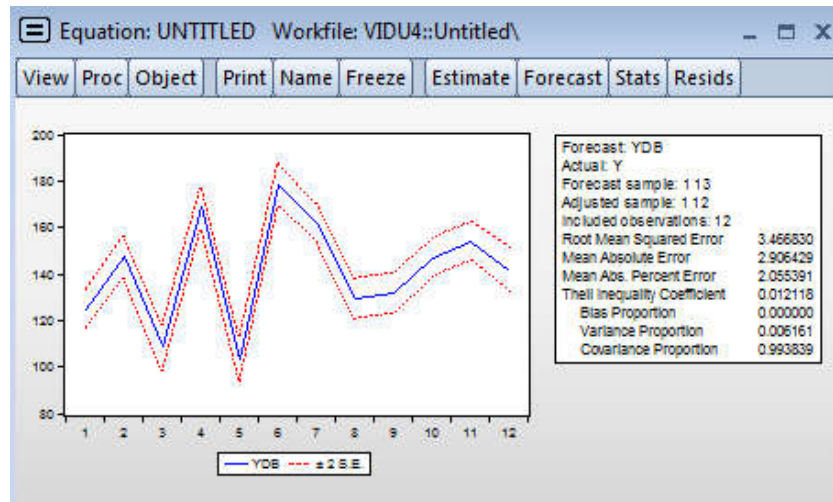
Output
 Forecast graph
 Forecast evaluation

Insert actuals for out-of-sample observations

OK Cancel

Hình 39

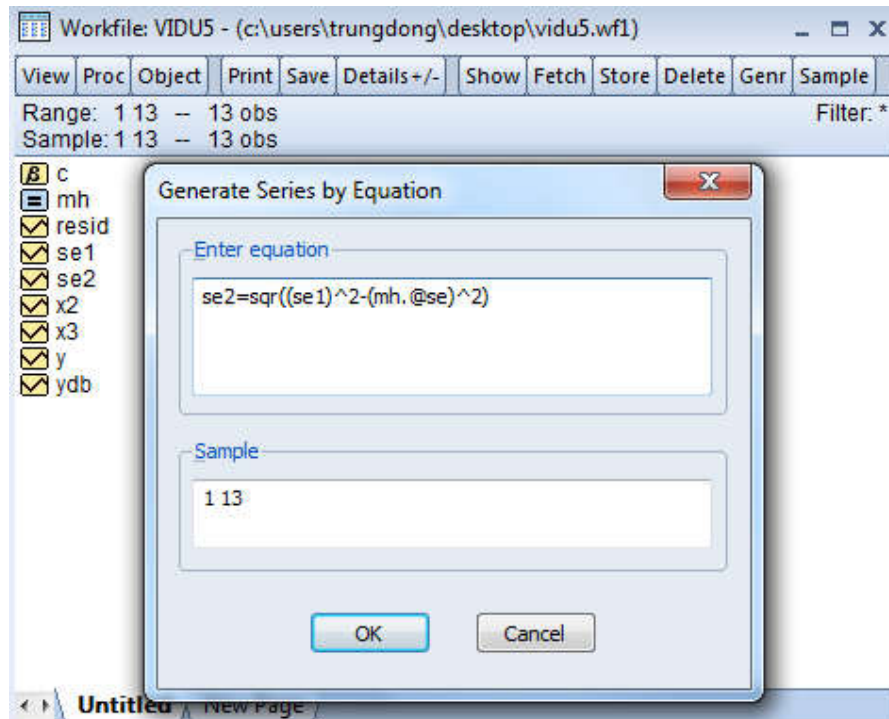
Ô **Forecast name** ta đổi Yf thành Y_{DB} , ô **S.E. (optional)** ta gõ Se1. **Nhấn OK**.



Hình 40

Tắt đồ thị dự báo

Từ bảng **Workfile**. Chọn **Genr** và gõ lệnh như sau rồi nhấn **Ok**.

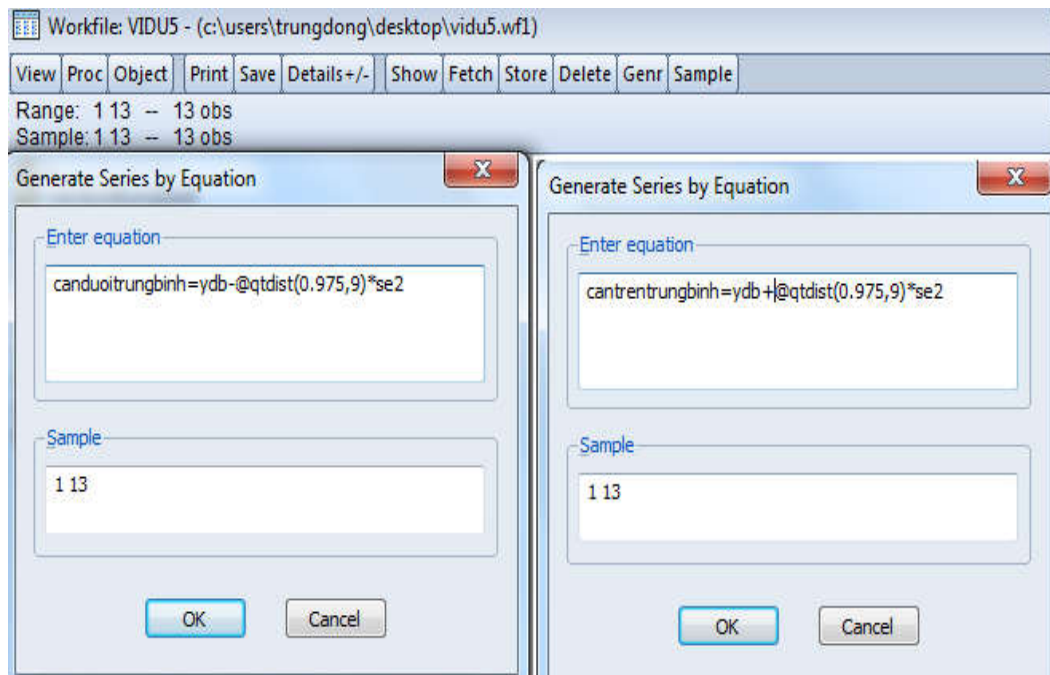


Hình 41

Bước 3. Tìm khoảng dự báo

- Dự báo giá trị trung bình

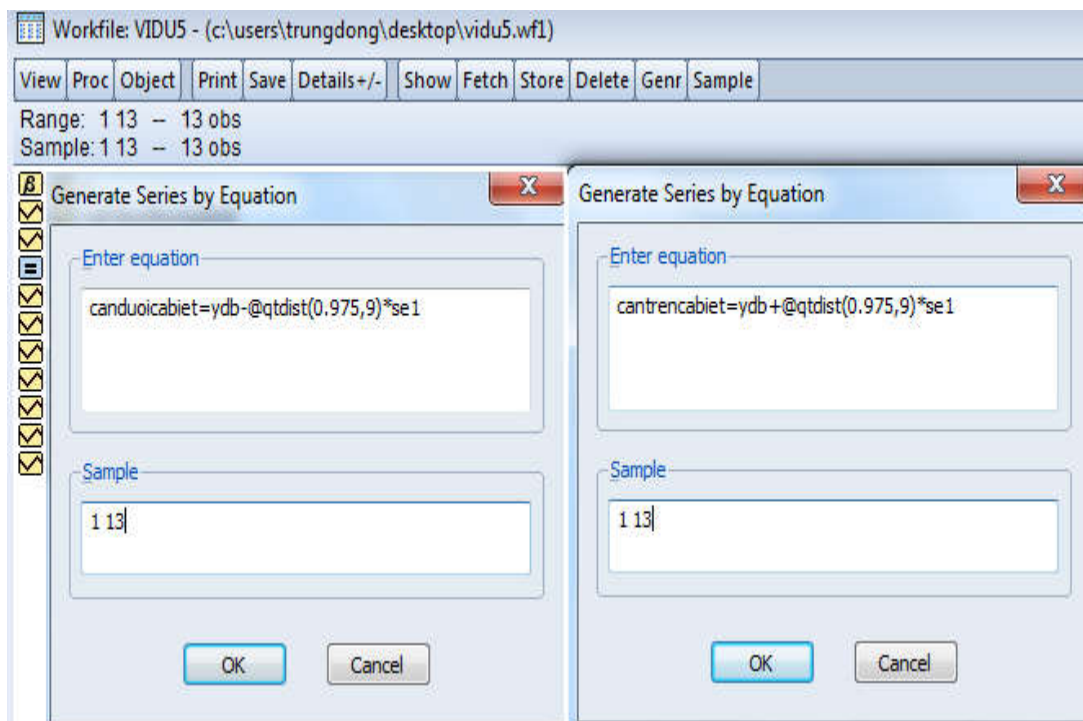
Từ bảng **Workfile**. Chọn **Genr** và gõ lệnh như sau rồi nhấn **Ok**.



Hình 42

Dự báo giá trị cá biệt.

Từ bảng **Workfile**. Chọn **Genr** và gõ lệnh như sau rồi nhấn **Ok**.



Hình 43

Để mở các kết quả trên cùng một bảng ta thực hiện như sau:

Từ cửa sổ **Workfile**, nhấn phím **Ctrl** rồi chọn **canduoicabiet**, **cantrencabiet**, **canduoitrungbinh**, **cantrentrungbinh** sau đó nhấn **Enter**, ta được kết quả sau (lưu ý nhìn vào hàng thứ 13)

View	Proc	Object	Print	Name	Freeze	Default	Sort	Transpose	Edit+/-	Smpl+/-	Title	Sample
obs		CANDUOITRUNGBINH		CANTRENTRUNGBINH		CANDUOICABIET		CANTRENCABIET				
obs		CANDUOITRUNGBINH		CANTRENTRUNGBINH		CANDUOICABIET		CANTRENCABIET				
1			121.7060		128.2287		115.3422					134.5925
2			142.4540		152.0782		137.0112					157.5210
3			102.5283		114.3482		97.62467					119.2519
4			164.3718		172.7359		158.5791					178.5286
5			97.92959		108.4185		92.70930					113.6388
6			173.1950		183.4530		167.9166					188.7314
7			157.5713		165.5131		151.6541					171.4303
8			125.3324		133.6141		119.5157					139.4308
9			128.3832		135.5748		122.2354					141.7225
10			143.7224		150.3043		137.3782					156.6485
11			150.9932		157.0569		144.4752					163.5748
12			134.9086		147.5786		130.1920					152.2952
13			153.9864		163.0754		148.3989					168.6630

Hình 44

Vậy khoảng dự báo giá trị trung bình và giá trị cá biệt của Y là

CANDUOITB	CANTRENTB	CANDUOICB	CANTRENCB
153.9864	163.0754	148.3989	168.6630

11. Định mẫu

Trước hết ta xét ví dụ sau

Ví dụ 5. Bảng số liệu sau cho biết số liệu về lượng hàng bán được (Y tấn/tháng), giá bán (X ngàn đồng/kg) ở 20 khu vực bán và được khảo sát tại hai nơi là Thành phố và Nông thôn.

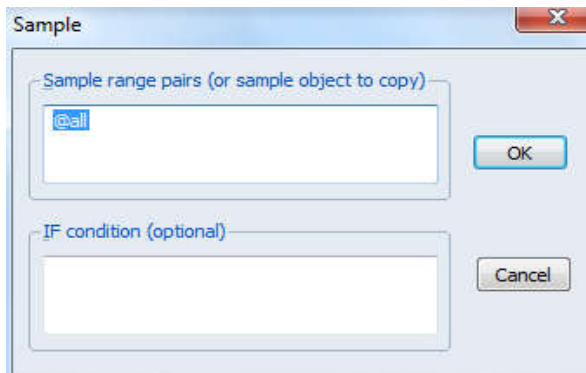
STT	Y	X	Z	STT	Y	X	Z
1	20	2	1	11	14	5	0
2	19	3	0	12	14	6	1
3	18	3	1	13	13	6	0
4	18	4	0	14	12	7	1
5	17	4	1	15	12	7	0
6	17	3	1	16	15	5	1
7	16	4	0	17	16	4	0
8	16	4	1	18	12	7	1
9	15	5	1	19	10	8	0
10	15	5	1	20	11	8	1

Bảng 8

Trong đó Z là biến giả với $Z = 0$: khảo sát ở nông thôn; $Z = 1$: khảo sát ở thành thị

Có nhiều trường hợp ta không sử dụng hết các số liệu của mẫu ban đầu, hay chỉ cần khảo sát sự phụ thuộc khi biến giả nhận một giá trị nào đó.

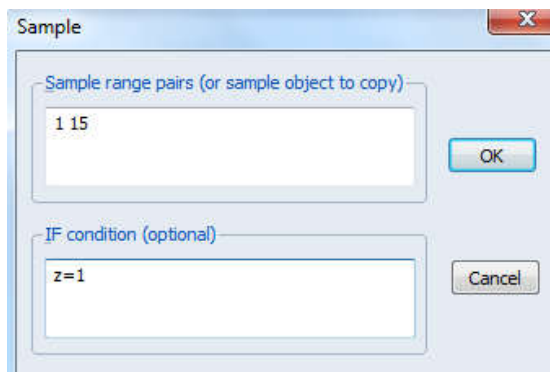
Để định mẫu lại, từ cửa sổ **Workfile** chọn **Sample**, màn hình xuất hiện như



Hình 45

Chẳng hạn ta chỉ khảo sát 15 mẫu đầu tiên và ở khu vực Thành phố ứng với $Z = 1$.

Ta khai báo vào ô Sample range pairs và IF condition như trong hình sau



Hình 46

Nhấn OK, ta thấy có sự thay đổi trong cửa sổ Workfile như sau

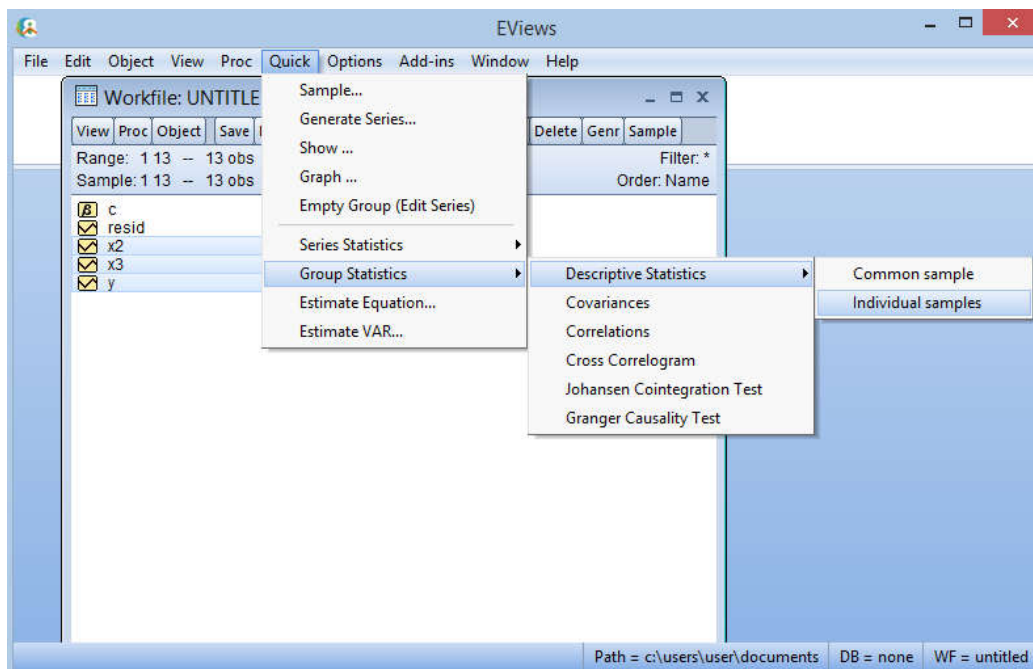


Hình 47

12. Tính các giá trị thống kê

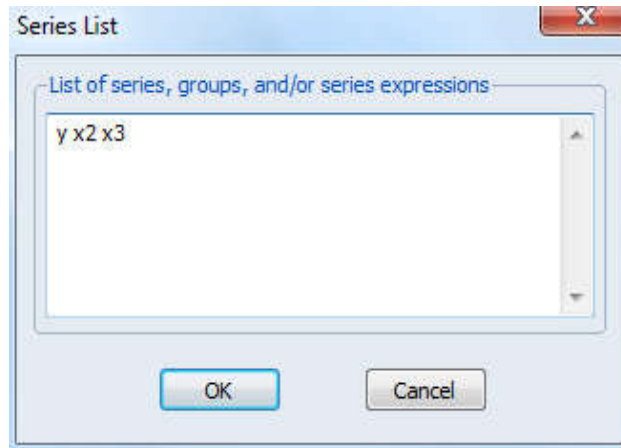
Để tính các giá trị thống kê như Trung bình, trung vị, độ lệch chuẩn, ... của các biến có trong mô hình chẳng hạn với số liệu cho trong ví dụ 4 ta làm như sau:

Từ cửa sổ **EViews** chọn **Quick** → **Group Statistics** → **Descriptive statistics** → **Common sample**, như hình sau



Hình 48

Nhấp chuột và nhập tên các biến vào cửa sổ **Series List** như hình sau



Hình 49

Nhấp OK, ta được bảng các giá trị thống kê sau:

View	Proc	Object	Print	Name	Freeze	Sample	Sheet	Stats	Spec
				Y		X2		X3	
Mean				141.3333		20.41667		12.16667	
Median				141.5000		20.50000		12.00000	
Maximum				180.0000		26.00000		17.00000	
Minimum				102.0000		15.00000		6.000000	
Std. Dev.				23.20789		4.187825		3.352972	
Skewness				-0.218444		-0.034924		-0.453123	
Kurtosis				2.301560		1.401529		2.378514	
Jarque-Bera				0.339345		1.279994		0.603763	
Probability				0.843941		0.527294		0.739426	
Sum				1696.000		245.0000		146.0000	
Sum Sq. Dev.				5924.667		192.9167		123.6667	
Observations				12		12		12	

Hình 50

Giải thích :

- Mean : trung bình.
- Median : trung vị
- Maximum : Giá trị lớn nhất
- Minimum : Giá trị nhỏ nhất
- Std. Dev : Độ lệch chuẩn
- Skewness : Hệ số bất đối xứng
- Kurtosis : Hệ số nhọn
- Jarque – Bera : Kiểm định phân phối chuẩn
- Sum : Tổng các quan sát

- Sum sq. Dev : Độ lệch chuẩn của tổng bình phương
- Observations : Số quan sát (cỡ mẫu)

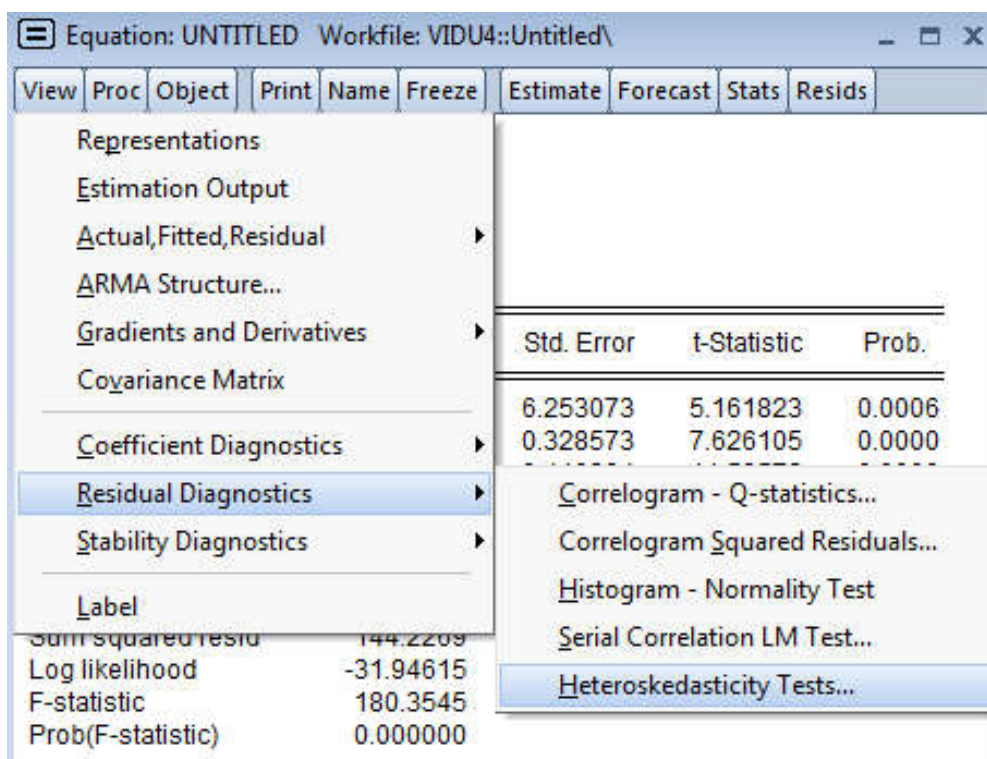
13. Các bài toán kiểm định giả thiết mô hình

13.1. Kiểm định phương sai thay đổi

13.1.1. Kiểm định White

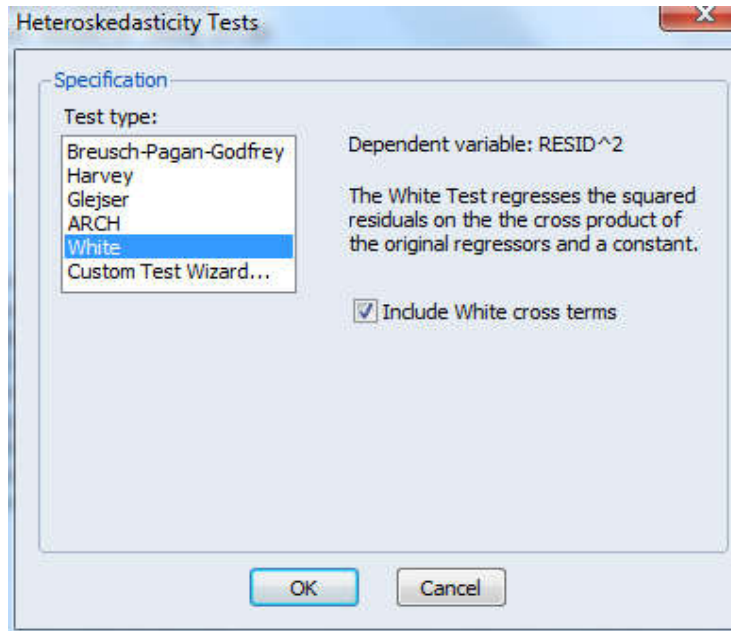
Chẳng hạn như trong ví dụ 4.

Để thực hiện việc kiểm định **White** bằng **Eview**, sau khi ước lượng mô hình hồi quy mẫu, từ cửa sổ **Equation** chọn **View**→**Residual Diagnostics** → **Heteroskedasticity tests...** Khi đó màn hình sẽ như sau:



Hình 51

Nhấp chuột, màn hình như sau



Hình 52

Ta chọn **White**, rồi **nhấn Ok**. Ta có kết quả như sau:

View	Proc	Object	Print	Name	Freeze	Estimate	Forecast	Stats	Resids
Equation: UNTITLED Workfile: VIDU4::Untitled\									
Heteroskedasticity Test: White									
F-statistic	0.843820	Prob. F(5,6)	0.5645						
Obs*R-squared	4.954369	Prob. Chi-Square(5)	0.4215						
Scaled explained SS	2.033496	Prob. Chi-Square(5)	0.8445						
Test Equation:									
Dependent Variable: RESID^2									
Method: Least Squares									
Date: 06/08/13 Time: 21:40									
Sample: 1 12									
Included observations: 12									
Variable	Coefficient	Std. Error	t-Statistic	Prob.					
C	-302.2919	201.0729	-1.503394	0.1834					
X2	31.71005	20.36090	1.557400	0.1704					
X2^2	-0.879882	0.520465	-1.690568	0.1419					
X2*X3	0.206625	0.536312	0.385270	0.7133					
X3	0.587451	10.18193	0.057695	0.9559					
X3^2	-0.075395	0.534322	-0.141104	0.8924					
R-squared	0.412864	Mean dependent var	12.01891						
Adjusted R-squared	-0.076416	S.D. dependent var	15.16494						
S.E. of regression	15.73369	Akaike info criterion	8.656339						
Sum squared resid	1485.295	Schwarz criterion	8.898792						
Log likelihood	-45.93803	Hannan-Quinn criter.	8.566574						
F-statistic	0.843820	Durbin-Watson stat	2.579780						
Prob(F-statistic)	0.564521								

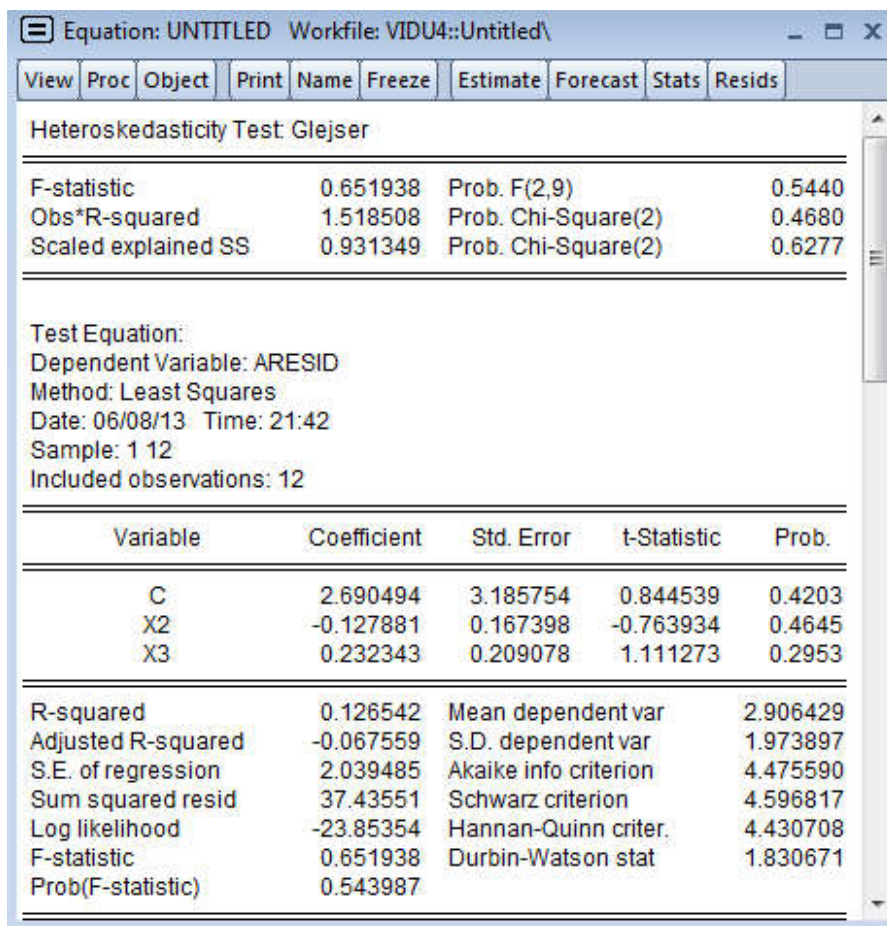
Hình 53

Ta đặt bài toán kiểm định như sau: H_0 : Mô hình không xảy ra hiện tượng phương sai thay đổi; H_1 : Mô hình xảy ra hiện tượng phương sai thay đổi.

Từ bảng kiểm định White ở trên, ta có $P_value = 0.4215 > \alpha$ cho trước nên chấp nhận H_0 . Vậy mô hình không xảy ra hiện tượng phương sai thay đổi.

13.1.2. Kiểm định Glejser.

Ta thực hiện các bước như trong kiểm định **White** nhưng ta chọn **Glejser**, rồi **nhấn Ok**. Ta có kết quả như sau:



Heteroskedasticity Test: Glejser				
F-statistic	0.651938	Prob. F(2,9)	0.5440	
Obs*R-squared	1.518508	Prob. Chi-Square(2)	0.4680	
Scaled explained SS	0.931349	Prob. Chi-Square(2)	0.6277	
Test Equation:				
Dependent Variable: ARESID				
Method: Least Squares				
Date: 06/08/13 Time: 21:42				
Sample: 1 12				
Included observations: 12				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.690494	3.185754	0.844539	0.4203
X2	-0.127881	0.167398	-0.763934	0.4645
X3	0.232343	0.209078	1.111273	0.2953
R-squared	0.126542	Mean dependent var	2.906429	
Adjusted R-squared	-0.067559	S.D. dependent var	1.973897	
S.E. of regression	2.039485	Akaike info criterion	4.475590	
Sum squared resid	37.43551	Schwarz criterion	4.596817	
Log likelihood	-23.85354	Hannan-Quinn criter.	4.430708	
F-statistic	0.651938	Durbin-Watson stat	1.830671	
Prob(F-statistic)	0.543987			

Hình 54

Ta đặt bài toán kiểm định như sau: H_0 : Mô hình không xảy ra hiện tượng phương sai thay đổi; H_1 : Mô hình xảy ra hiện tượng phương sai thay đổi.

Từ bảng kiểm định **Glejser** ở trên, ta có $P_value = 0.4680 > \alpha$ cho trước nên chấp nhận H_0 . Vậy mô hình không xảy ra hiện tượng phương sai thay đổi.

13.1.3. Kiểm định Breusch-Pagan-Godfrey

Ta thực hiện các bước tương tự như kiểm định **White** nhưng ta chọn **Breusch-Pagan-Godfrey**, rồi **nhấn Ok**. Ta có kết quả như sau:

Equation: UNTITLED Workfile: VIDU4::Untitled\

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Heteroskedasticity Test: Breusch-Pagan-Godfrey

F-statistic	0.596913	Prob. F(2,9)	0.5709
Obs*R-squared	1.405352	Prob. Chi-Square(2)	0.4953
Scaled explained SS	0.576820	Prob. Chi-Square(2)	0.7495

Test Equation:
 Dependent Variable: RESID^2
 Method: Least Squares
 Date: 06/08/13 Time: 21:49
 Sample: 1 12
 Included observations: 12

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	10.63993	24.60707	0.432393	0.6756
X2	-0.953765	1.292998	-0.737639	0.4795
X3	1.713837	1.614940	1.061239	0.3162

R-squared	0.117113	Mean dependent var	12.01891
Adjusted R-squared	-0.079085	S.D. dependent var	15.16494
S.E. of regression	15.75319	Akaike info criterion	8.564280
Sum squared resid	2233.466	Schwarz criterion	8.685507
Log likelihood	-48.38568	Hannan-Quinn criter.	8.519398
F-statistic	0.596913	Durbin-Watson stat	2.193097
Prob(F-statistic)	0.570918		

Hình 55

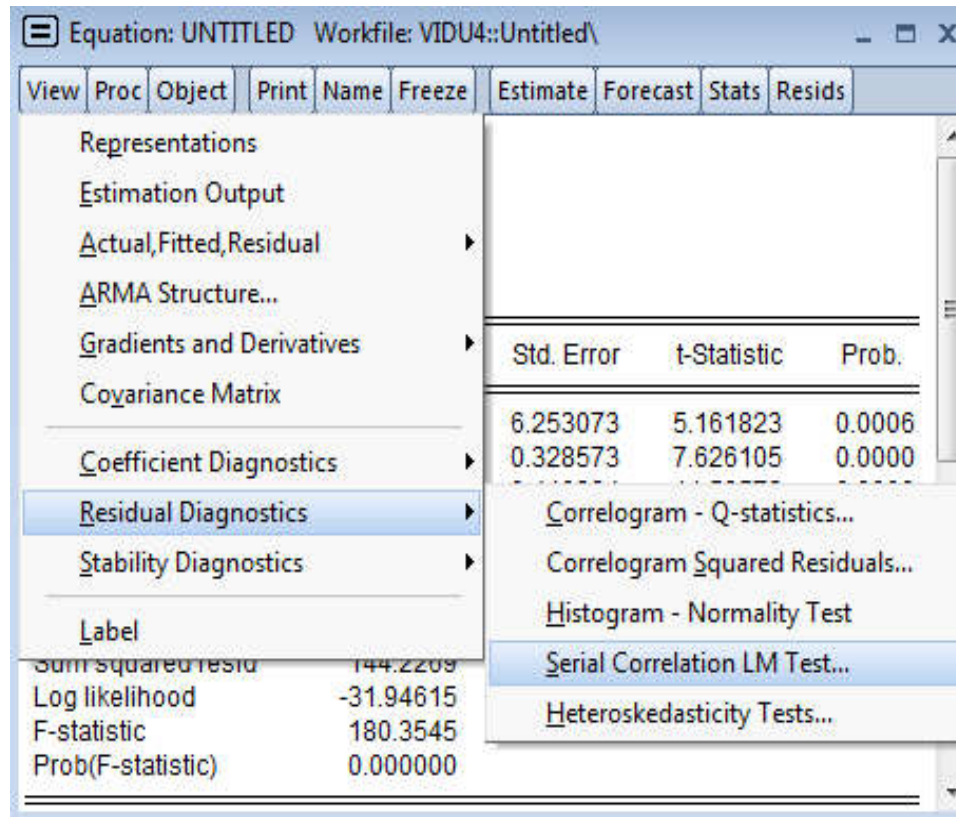
Ta đặt bài toán kiểm định như sau: H_0 : Mô hình không xảy ra hiện tượng phương sai thay đổi; H_1 : Mô hình xảy ra hiện tượng phương sai thay đổi.

Từ bảng kiểm định **Breusch – Pagan - Godfrey** ở trên, ta có $P_value = 0.4953 > \alpha$ cho trước nên chấp nhận H_0 . Vậy mô hình không xảy ra hiện tượng phương sai thay đổi.

13.2. Kiểm định tự tương quan (Kiểm định BG)

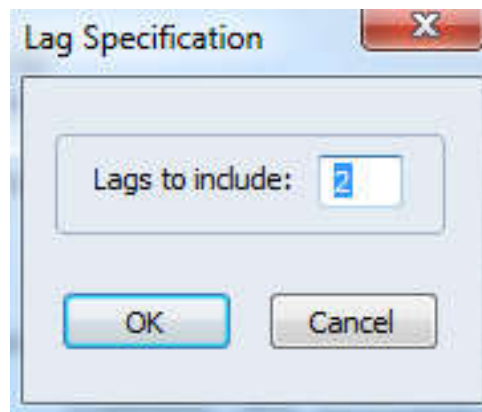
Chẳng hạn như trong ví dụ 4.

Để thực hiện việc kiểm định **BG** bằng Eview, sau khi ước lượng mô hình hồi quy mẫu, từ cửa sổ **Equation** chọn **View**→**Residual Diagnostics** → **Serial Correlation LM test...** Khi đó màn hình sẽ xuất hiện như sau:



Hình 56

Nhấp chuột, cửa sổ sau xuất hiện như sau:



Hình 57

Ô **Lags to include** ta gõ bậc tự tương quan vào (ví dụ như tự tương quan là bậc 2)
Nhấn Ok. Ta có kết quả như sau:

Equation: UNTITLED Workfile: VIDU4::Untitled\

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.481309	Prob. F(2,7)	0.6370
Obs*R-squared	1.450706	Prob. Chi-Square(2)	0.4842

Test Equation:
 Dependent Variable: RESID
 Method: Least Squares
 Date: 06/08/13 Time: 22:00
 Sample: 1 12
 Included observations: 12
 Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.843019	6.704450	-0.125740	0.9035
X2	0.027662	0.353556	0.078238	0.9398
X3	0.031521	0.442767	0.071191	0.9452
RESID(-1)	-0.321367	0.384510	-0.835784	0.4309
RESID(-2)	0.127392	0.426641	0.298592	0.7739

R-squared	0.120892	Mean dependent var	2.04E-14
Adjusted R-squared	-0.381455	S.D. dependent var	3.620986
S.E. of regression	4.255938	Akaike info criterion	6.028844
Sum squared resid	126.7910	Schwarz criterion	6.230888
Log likelihood	-31.17306	Hannan-Quinn criter.	5.954040
F-statistic	0.240655	Durbin-Watson stat	1.891410
Prob(F-statistic)	0.906546		

Hình 58

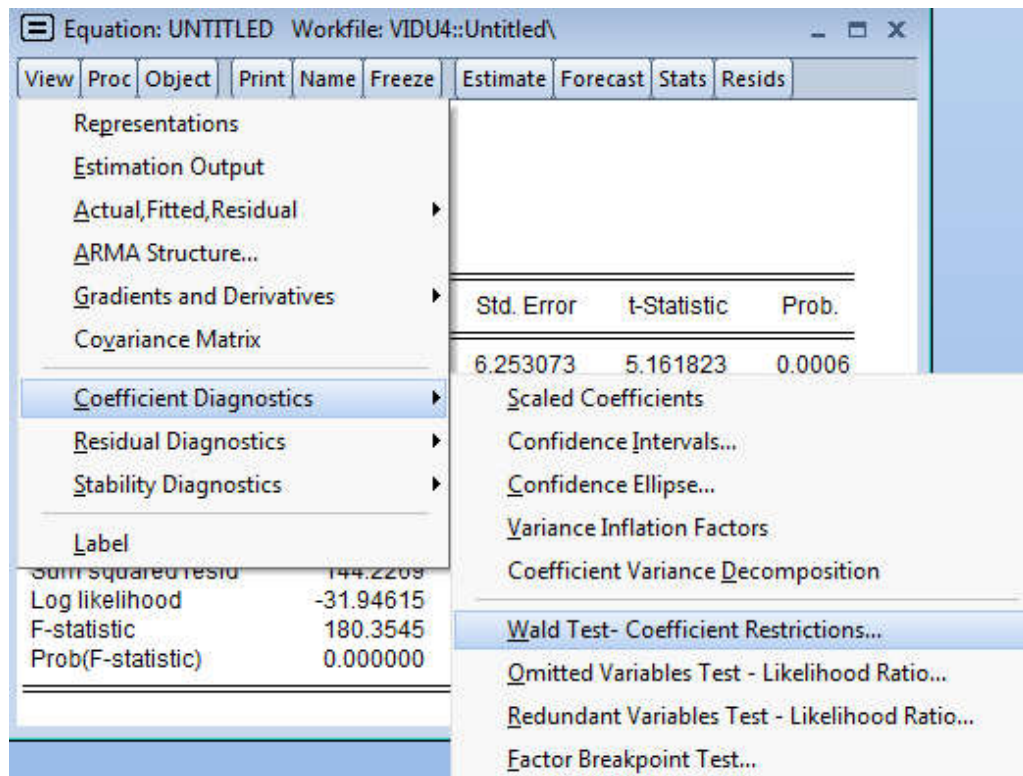
Ta đặt bài toán kiểm định như sau: H_0 : Mô hình không xảy ra hiện tượng tự tương quan bậc 2; H_1 : Mô hình xảy ra hiện tượng tự tương quan bậc 2.

Từ bảng kiểm định BG ở trên, ta có $P_value = 0.4842 > \alpha$ cho trước nên chấp nhận H_0 . Vậy mô hình không xảy ra hiện tượng tự tương quan bậc 2.

13.3. Kiểm định biến có cần thiết trong mô hình hay không (Kiểm định Wald)

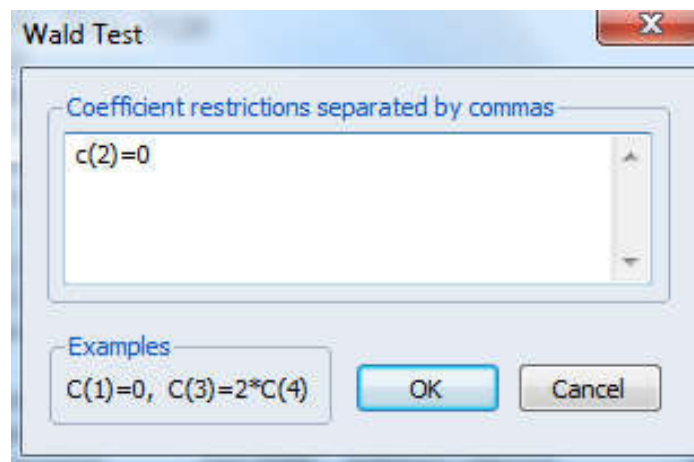
Chẳng hạn như trong ví dụ 4.

Để thực hiện việc kiểm định **Wald** bằng **Eview**, sau khi ước lượng mô hình hồi quy mẫu, từ cửa sổ **Equation** chọn **View**→**Coefficient Diagnostics** → **Wald test – Coefficient Restrictions...** Khi đó màn hình sẽ như sau:



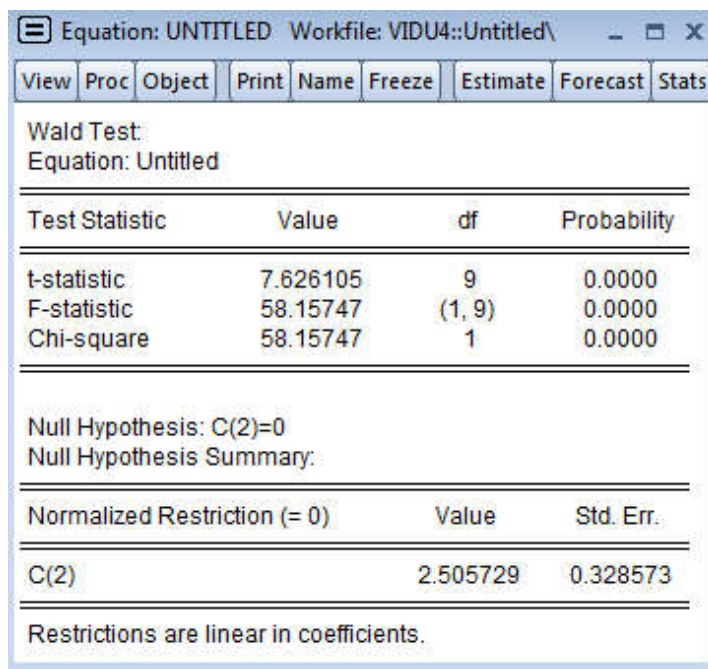
Hình 59

Nhấp chuột ta có cửa sổ sau xuất hiện: Gõ $c(2)=0$ vào



Hình 60

Nhấp Ok. Ta được kết quả như sau:



Test Statistic	Value	df	Probability
t-statistic	7.626105	9	0.0000
F-statistic	58.15747	(1, 9)	0.0000
Chi-square	58.15747	1	0.0000

Normalized Restriction (= 0)	Value	Std. Err.
C(2)	2.505729	0.328573

Hình 61

Ta đặt bài toán kiểm định như sau: H_0 : Biến X2 không cần thiết trong mô hình;
 H_1 : Biến X2 cần thiết trong mô hình.

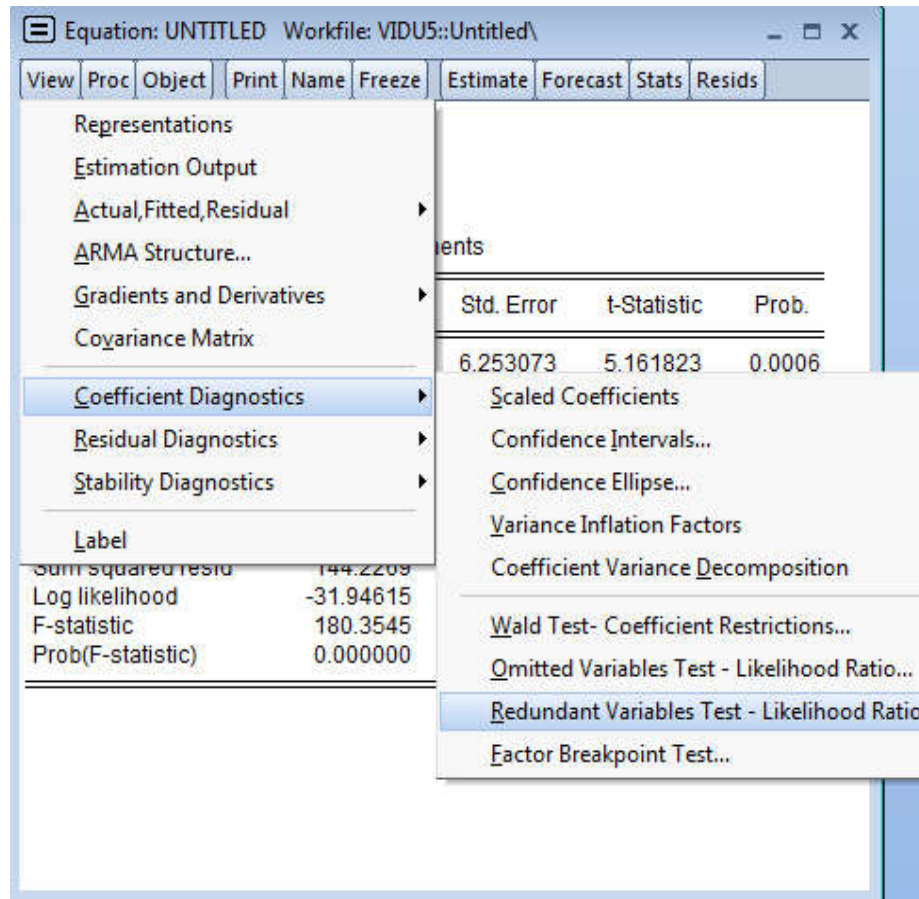
Từ bảng kiểm định **Wald** ở trên, ta có $P_value = 0.0000 < \alpha$ cho trước nên bác bỏ H_0 . Vậy X2 cần thiết trong mô hình.

Lưu ý: Trong trường hợp này ta chỉ khảo sát X2 nên ta có thể dùng giá trị xác suất của thống kê t hoặc giá trị xác suất của thống kê F đều được. Trong trường hợp ta khảo sát nhiều hơn hai biến thì ta chỉ dùng thống kê F.

13.4. Kiểm định thừa biến trong mô hình (biến không cần thiết)

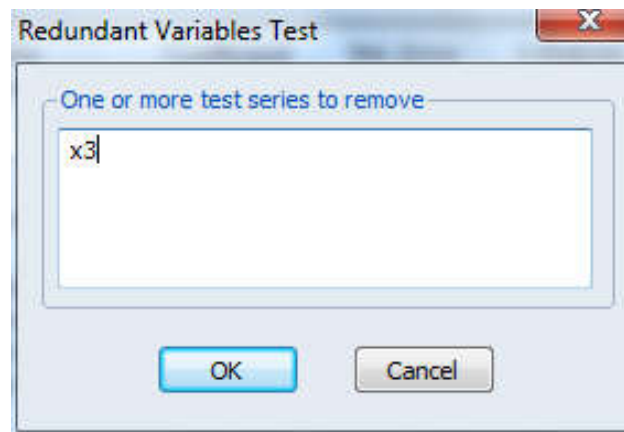
Giả sử xét ví dụ 4 bên trên, ta tiến hành như sau:

- Tìm hàm hồi quy của Y theo X2 và X3. Từ cửa sổ Equation, ta chọn **View**→**Coefficient Diagnostics** → **Redundant Variables Test – Likelihood ratio...** Khi đó màn hình sẽ như sau:



Hình 62

Nhấp chuột ta có cửa sổ **One or more test series to remove** xuất hiện, rồi gõ biến X3 vào



Hình 63

Nhấp **Ok**, ta có kết quả sau:

	Value	df	Probability
t-statistic	11.59572	9	0.0000
F-statistic	134.4608	(1, 9)	0.0000
Likelihood ratio	33.22605	1	0.0000

Hình 64

Ta đặt bài toán kiểm định như sau: $H_0 : \beta_3 = 0$: Biến X3 không cần thiết trong mô hình; $H_1 : \beta_3 \neq 0$: Biến X3 cần thiết trong mô hình.

Từ bảng kiểm định ở trên, ta có $P_value = 0.0000 < \alpha$ cho trước nên bác bỏ H_0 .

Vậy X3 cần thiết trong mô hình.

13.5. Kiểm định biến bị bỏ sót trong mô hình

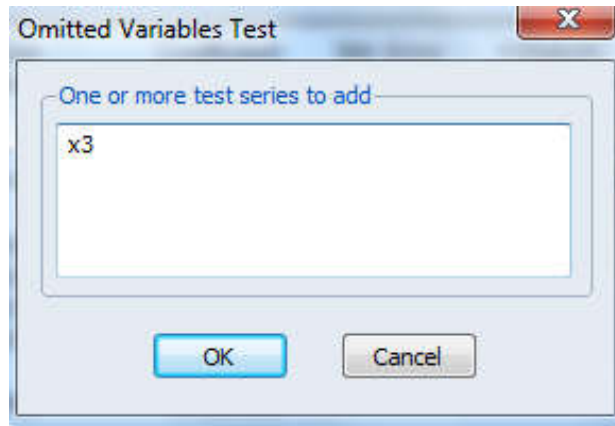
Giả sử xét ví dụ 4 bên trên, ta tiến hành như sau.

- Tìm hàm hồi quy mẫu của Y theo X2. Từ cửa sổ **Equation**, ta chọn **View**→**Coefficient Diagnostics** → **Omitted Variables Test – Likelihood ratio...** Khi đó màn hình sẽ như sau:

	Std. Error	t-Statistic	Prob.
	22.71359	2.325607	0.0424

Hình 65

Nhấp chuột ta có cửa sổ **One or more test series to add** xuất hiện. Ta gõ biến X3 vào



Hình 66

Nhấp Ok, ta được kết quả sau:

	Value	df	Probability
t-statistic	11.59572	9	0.0000
F-statistic	134.4608	(1, 9)	0.0000
Likelihood ratio	33.22605	1	0.0000

Hình 67

Ta đặt bài toán kiểm định như sau: $H_0 : \beta_3 = 0$: Biến X3 ảnh hưởng tới Y (X3 không bị bỏ sót); $H_1 : \beta_3 \neq 0$: Biến X3 bị bỏ sót trong mô hình.

Từ bảng kiểm định ở trên, ta có $P_value = 0.0000 < \alpha$ cho trước nên bác bỏ H_0 .

Vậy X3 bị bỏ sót trong mô hình.

13.6. Kiểm định Chow trong mô hình hồi quy với biến giả

Ví dụ 7. Giả sử số liệu về tiết kiệm và thu nhập cá nhân ở nước Anh từ năm 1946 đến 1963 (đơn vị pound) cho ở bảng sau:

Thời kỳ I	Y	X	Thời kỳ II	Y	X
1946	0.36	8.80	1955	0.59	15.5
1947	0.21	9.40	1956	0.90	16.7
1948	0.08	10.0	1957	0.95	17.7
1949	0.20	10.6	1958	0.82	18.6
1950	0.10	11.0	1959	1.04	19.7
1951	0.12	11.9	1960	1.53	21.1
1952	0.41	12.7	1961	1.94	22.8
1953	0.50	13.5	1962	1.75	23.9
1954	0.43	14.3	1963	1.99	25.2

Bảng 10

Trong đó, Y : Tiết kiệm ; X : Thu nhập.

Để kiểm định rằng có sự thay đổi về tiết kiệm giữa hai thời kỳ hay không, ta thực hiện các bước kiểm định Chow như sau:

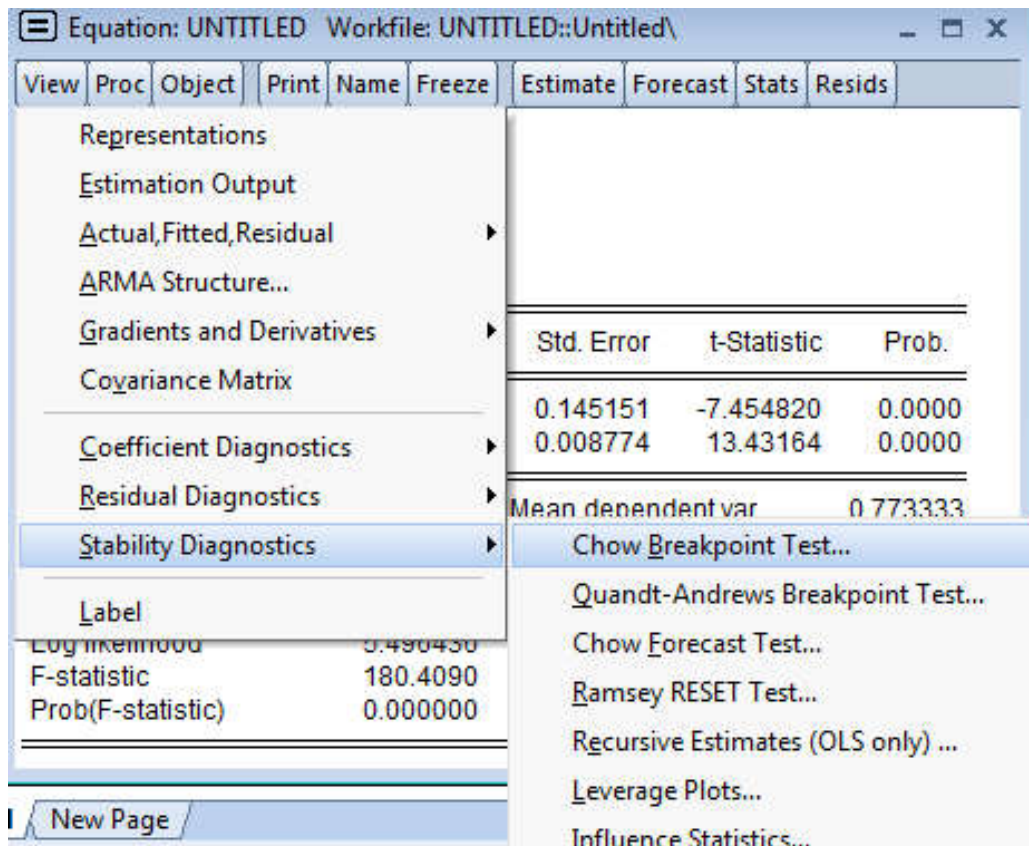
Hồi quy Y theo X, ta được kết quả

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.082071	0.145151	-7.454820	0.0000
X	0.117845	0.008774	13.43164	0.0000

R-squared	0.918537	Mean dependent var	0.773333
Adjusted R-squared	0.913446	S.D. dependent var	0.642806
S.E. of regression	0.189114	Akaike info criterion	-0.388493
Sum squared resid	0.572226	Schwarz criterion	-0.289563
Log likelihood	5.496436	Hannan-Quinn criter.	-0.374852
F-statistic	180.4090	Durbin-Watson stat	0.947783
Prob(F-statistic)	0.000000		

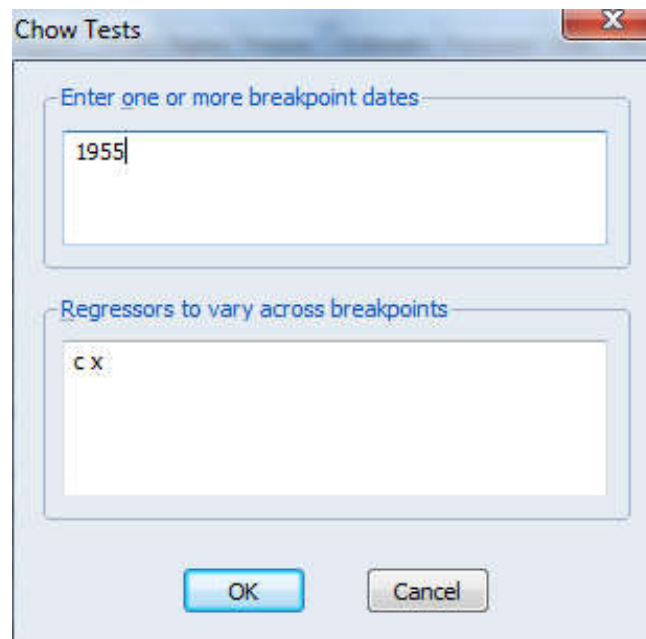
Hình 68

Từ cửa sổ **Equation**, chọn **View** → **Stability Diagnostics** → **Chow Breakpoint Test**... như hình sau:



Hình 69

Sau khi nhấp chuột, một cửa sổ xuất hiện như sau:



Hình 70

Ta gõ vào cửa sổ **Chow Test** giá trị **Breakpoint** là **1955** như hình trên, **nhấp OK**. Khi đó ta được kết quả sau:

Chow Breakpoint Test: 1955			
Null Hypothesis: No breaks at specified breakpoints			
Varying regressors: All equation variables			
Equation Sample: 1946 1963			
F-statistic	5.037060	Prob. F(2,14)	0.0225
Log likelihood ratio	9.757441	Prob. Chi-Square(2)	0.0076
Wald Statistic	10.07412	Prob. Chi-Square(2)	0.0065

Hình 71

và dựa vào bảng kết quả trên ta cũng có giá trị $F = 5.037$. Với giá trị xác suất là 0.022493, nên ta chấp nhận giả thuyết là hai mô hình hồi quy khác nhau.

14. Định dạng mô hình (Kiểm định Ramsey RESET)

Xét mô hình gốc:

$$Y_i = \beta_1 + \beta_2 X_i + \varepsilon_i \quad (1)$$

Kiểm định Ramsey RESET

$$Y_i = \beta_1 + \beta_2 X_i + \alpha_1 \hat{Y}_i^2 + \alpha_2 \hat{Y}_i^3 + \dots + \alpha_m \hat{Y}_i^{m+1} + \varepsilon_i \quad (2)$$

Bài toán kiểm định

$$\begin{cases} H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_m = 0 \\ H_1 : \exists \alpha_j \neq 0, j = \overline{1, m} \end{cases}$$

H_0 : Mô hình gốc không thiếu biến, dạng hàm đúng

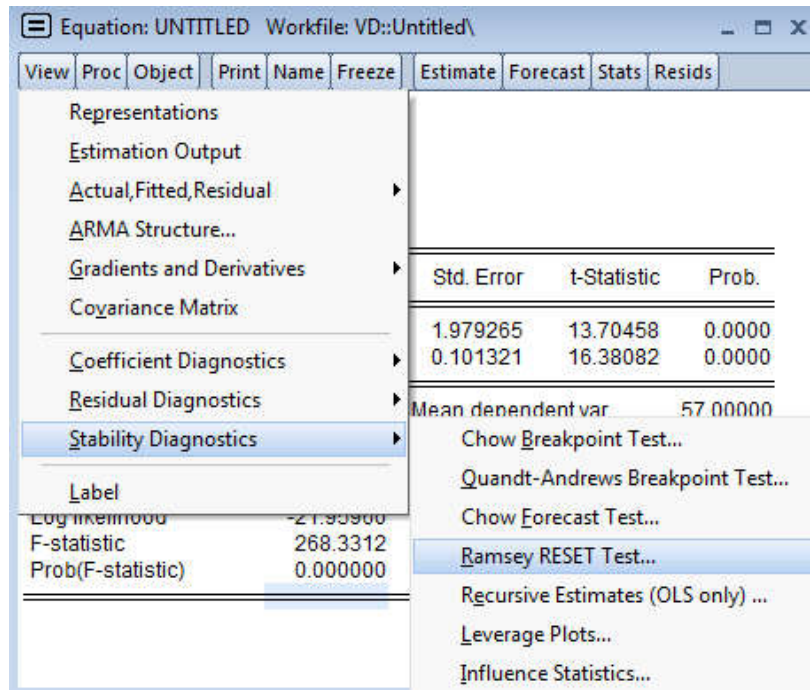
H_1 : Mô hình gốc thiếu biến, dạng hàm sai

$$F = \frac{R_2^2 - R_1^2}{1 - R_2^2} \cdot \frac{n - k_2}{m} \sim F(m, n - k_2)$$

Giả sử xét ví dụ 3 bên trên, ta tiến hành như sau:

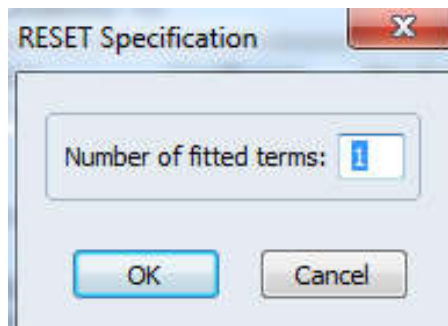
Tìm hàm hồi quy tuyến tính mẫu của Y theo X. Từ của số Equation.

Chọn **View** → **Stability Diagnostics** → **Ramsey RESET Test...** như hình sau:



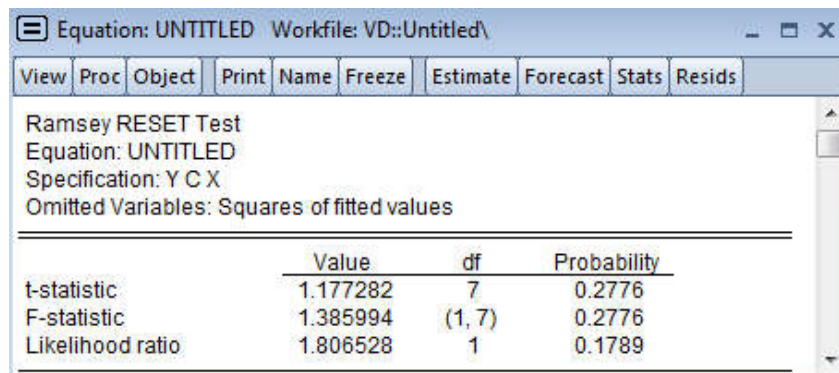
Hình 72

Nhấp chuột ta có cửa sổ **Number of fitted terms** xuất hiện. Ta gõ tham số $m=1$ vào



Hình 73

Nhấp **Ok**, ta được kết quả sau:



Hình 74

Ta đặt bài toán kiểm định như sau: $H_0 : \alpha_1 = 0$: Mô hình trên không thiếu biến, dạng hàm đúng; $H_1 : \alpha_1 \neq 0$: Mô hình trên thiếu biến dạng hàm sai.

Từ bảng kiểm định ở trên, ta có $P_value(F_statistic) = 0.2776 > \alpha$ cho trước nên chấp nhận H_0 . Vậy mô hình trên không thiếu biến, dạng hàm đúng.

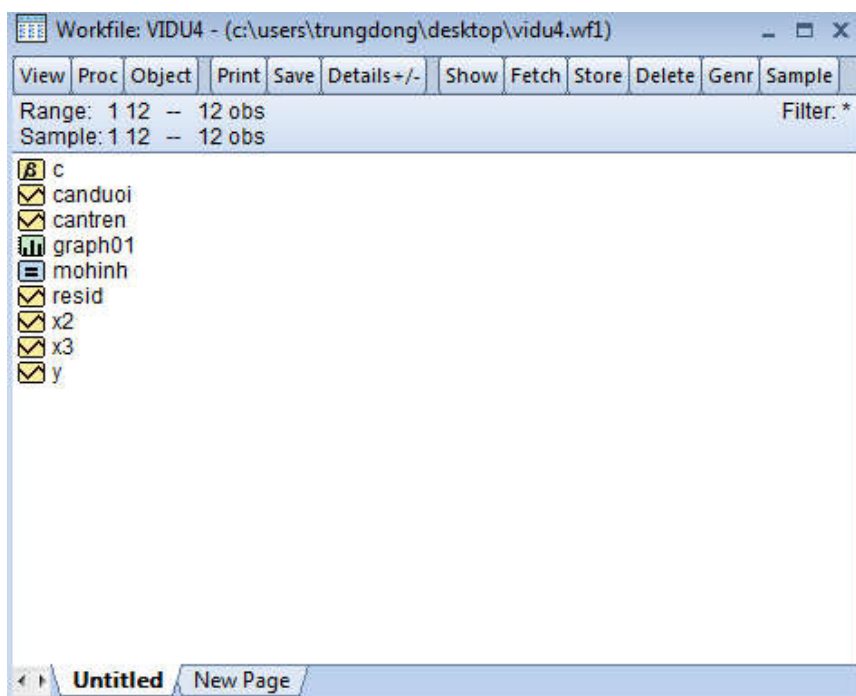
15. Lưu kết quả trong Eviews

15.1. Lưu file dữ liệu

Các thao tác được thực hiện như sau:

Sau khi làm xong các thao tác. Từ cửa sổ **Eviews** chọn **File** → **Save**

Lưu ý: Khi đó trên cửa sổ Workfile thì không có đối tượng nào được chọn (Nếu không ta chỉ lưu được một file dạng rác).



Hình 75

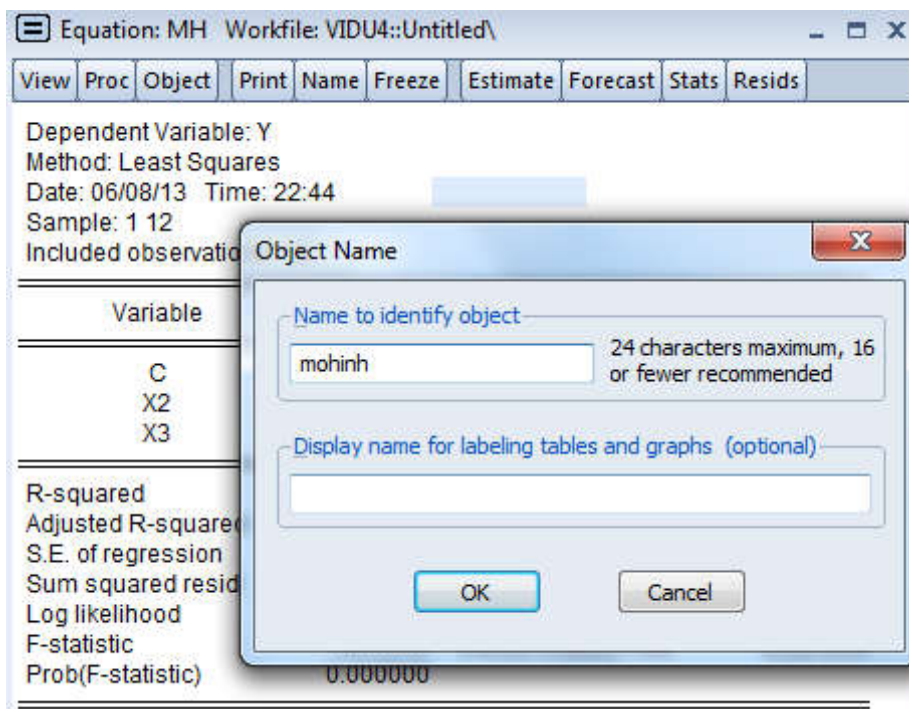
15.2. Lưu các bảng kết quả

Trên các cửa sổ như **Equation**, **Graph**, **Group**, ...Đều có thanh công cụ chứa hai nút

là : **Name** và **Freeze** dùng để lưu trữ các đối tượng hoặc các kết quả được tạo ra trong quá trình thao tác. Đối với chức năng **Name** cho phép ta lưu trữ các kết quả mà ta có thể dùng tiếp cho các thao tác sau. Mặt khác chức năng **Freeze** chỉ lưu các kết quả dưới dạng một Table (Kết quả đó được đóng băng).


Chẳng hạn với số liệu trong ví dụ 3 sau khi tìm được mô hình hồi quy xong và ta thực hiện lưu trữ như sau:

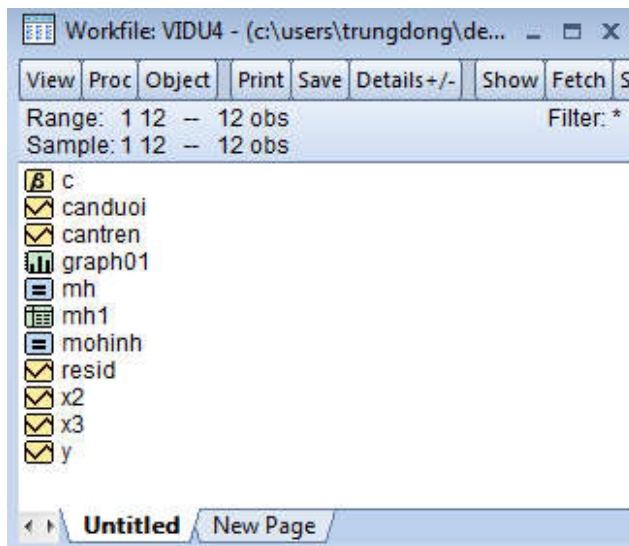
Từ cửa sổ **Equation**. Nếu ta chọn chức năng **Name** như hình 63



Hình 76

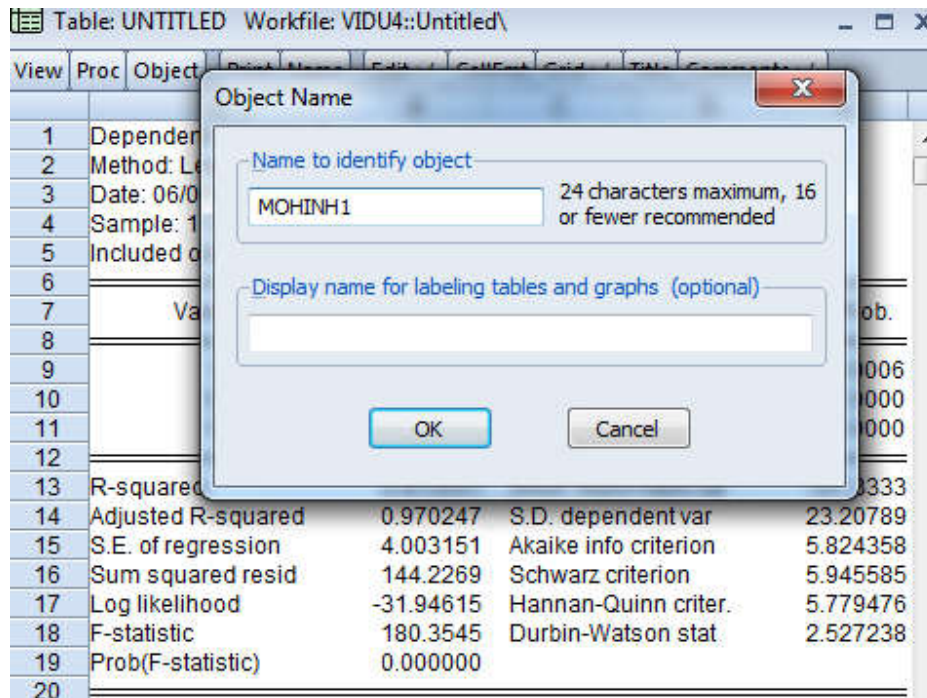
Chọn **OK** ta được kết quả có biểu tượng là

 mohinh



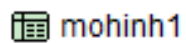
Hình 77

Từ cửa sổ **Equation**. Nếu ta chọn chức năng **Freeze** thì ta thấy một table mới xuất hiện như sau:



Hình 78

Chọn OK ta được kết quả có biểu tượng là



TÀI LIỆU THAM KHẢO

Tiếng Việt

- [1] Nguyễn Quang Dong, Nguyễn Thị Minh, (2013). *Giáo trình Kinh tế lượng*, Nhà xuất bản Đại học Kinh tế Quốc Dân.
- [2] Trần Tiến Khai, (2012). *Phương pháp nghiên cứu kinh tế - Kiến thức cơ bản*. Khoa Kinh tế Phát triển. Đại học Kinh Tế TP. Hồ Chí Minh. Nhà Xuất Bản Lao Động Xã Hội.
- [3] Nguyễn Thị Tuyết Mai, Nguyễn Vũ Hùng, (2015). *Phương pháp điều tra khảo sát: nguyên lý và thực tiễn*, Nhà xuất bản ĐHKQTĐ.
- [4] Hoàng Trọng – Chu Nguyễn Mộng Ngọc, (2008). *Phân tích dữ liệu nghiên cứu với SPSS*. TP.HCM. NXB Thống Kê.
- [5] Nguyễn Văn Hiến, (2016), *Nghiên cứu marketing thực hành*, Nhà xuất bản tài chính.
- [6] Lê Công Hoa, Nguyễn Thành Hiếu, (2014), *Nghiên cứu kinh doanh*, Nhà xuất bản Đại học Kinh tế Quốc Dân.
- [7] Đinh Phi Hổ, (2016), *Phương pháp nghiên cứu kinh tế & viết luận văn thạc sĩ*, Nhà xuất bản Đại học Kinh tế TP. Hồ Chí Minh.
- [8] Hồ Đăng Phúc, *Sử Dụng Phần Mềm SPSS Trong Phân Tích Số Liệu*, Nhà xuất bản Khoa học và Kỹ thuật.
- [9] Nguyễn Văn Thắng, (2014), *Thực hành nghiên cứu trong Kinh tế và Quản trị kinh doanh*, Nhà xuất bản Đại học Kinh tế Quốc Dân.
- [10] Nguyễn Văn Tuấn, (2011), *Cách viết đề cương nghiên cứu khoa học*. Bài giảng cho sinh viên Đại học Quốc gia TPHCM, Sydney: ĐH New South Wales.
- [11] Một số bài báo quốc tế, trong nước, đề cương, bài nghiên cứu khoa học, luận án làm mẫu.

Tiếng Anh

- [12] Babbie, E.R., (2011). *The Practice of Social Research*. Belmont CA: Wadsworth.
- [13] Daniel Muijs, (2004). *Doing Quantitative Research in Education with SPSS*. Sage Publications.
- [14] Ehrenberg, A.S.C., (1994). Theory or Well-Based Results: Which Comes First. *In Research Traditions in Marketing* (Laurent, G and Lilien, G.L.) Boston: Kluwer Academic.

- [15] Kumar, R. (2014). *Research Methodology. A Step-by-Step Guide for Beginners*. Fourth edition. SAGE Publications.
- [16] Kothari, C.R., (2004). *Research Methodology: Methods and Techniques*. New Age International (p) Ltd.
- [17] Marshall, C., & Rossman, G. B. (2006). *Designing Qualitative Research* (4 th ed.). Thousand Oaks, CA: Sage.
- [18] John A. Sharp, John Peters and Keith Howard. (2006). *The management of a Student Research Project*. Third edition. Gower Publishing Company.
- [19] Joseph F. Hair Jr. William C. Black Barry J. Babin Rolp E. Anderson, (2014). *Multivariate Data Analysis*. Pearson New International Seventh Edition.
- [20] Robert A. Day. (1998). *How to write and publish a scientific paper*. Fifth edition. Oryx Press.
- [21] Shuttleworth, M., (2008). Definition of Research. <<https://explorable.com/definition-of-research>>.
- [22] Uma Sekaran and Roger Bougie (2009). *Research methods for business: A skill building approach*.