

LỜI CẢM ƠN

Trước hết em xin bày tỏ tình cảm và lòng biết ơn đối với thầy Nguyễn Trịnh Đông – Khoa Công nghệ Thông tin – Trường Đại học Dân Lập Hải Phòng, người đã dành cho em rất nhiều thời gian quý báu, trực tiếp hướng dẫn tận tình giúp đỡ, chỉ bảo em trong suốt quá trình làm đồ án tốt nghiệp.

Em xin chân thành cảm ơn tất cả các thầy cô giáo trong khoa Công nghệ Thông tin - Trường ĐHDL Hải Phòng, chân thành cảm ơn các thầy giáo, cô giáo tham gia giảng dạy và truyền đạt những kiến thức quý báu trong suốt thời gian em học tập tại trường, đã đọc và phản biện đồ án của em giúp em hiểu rõ hơn các vấn đề mình nghiên cứu, để em có thể hoàn thành đồ án này.

Em xin cảm ơn GS.TS.NGUT Trần Hữu Nghị Hiệu trưởng Trường Đại học Dân lập Hải Phòng, Ban giám hiệu nhà trường, Bộ môn tin học, các Phòng ban nhà trường đã tạo điều kiện tốt nhất trong suốt thời gian học tập và làm tốt nghiệp.

Tuy có nhiều cố gắng trong quá trình học tập, trong thời gian thực tập cũng như trong quá trình làm đồ án nhưng không thể tránh khỏi những thiếu sót, em rất mong được sự góp ý quý báu của tất cả các thầy giáo, cô giáo cũng như tất cả các bạn để kết quả của em được hoàn thiện hơn.

Em xin chân thành cảm ơn!

Hải Phòng, ngày tháng năm 2013

Sinh viên

Nguyễn Trung Kiên

MỤC LỤC

LỜI CẢM ƠN	1
MỤC LỤC.....	2
DANH MỤC BẢNG VÀ HÌNH VẼ.....	4
LỜI MỞ ĐẦU	6
CHƯƠNG 1: BÀI TOÁN PHÂN TÍCH VÀ DỰ BÁO.....	8
1.1. Khái niệm.....	8
1.2. Ý nghĩa và vai trò của phân tích và dự báo trong quá trình ra quyết định kinh doanh.....	8
1.2.1. Ý nghĩa	8
1.2.2. Vai trò.....	9
1.3. Phân loại dự báo	9
1.3.1. Căn cứ vào độ dài thời gian dự báo:	9
1.3.2. Dựa vào các phương pháp dự báo:.....	10
1.3.3. Căn cứ vào nội dung (đối tượng dự báo)	11
1.4. Các phương pháp dự báo	12
1.4.1. Phương pháp dự báo định tính	12
1.4.2. Phương pháp dự báo định lượng.....	14
1.5. Quy trình dự báo	20
CHƯƠNG 2: HỆ THỐNG PHÂN TÍCH VÀ DỰ BÁO.....	23
2.1. Dự báo từ các mức độ bình quân.....	23
2.1.1. Dự báo từ số bình quân trượt (di động).....	23
2.1.2. Mô hình dự báo dựa vào lượng tăng (giảm) tuyệt đối bình quân	23
2.1.3. Mô hình dự báo dựa vào tốc độ phát triển bình quân	24
2.2. Mô hình dự báo theo phương trình hồi quy (dự báo dựa vào xu thế)	25

2.2.1. Mô hình hồi quy theo thời gian.....	25
2.2.2. Mô hình hồi quy giữa các tiêu thức.....	26
2.3. Dự báo dựa vào hàm xu thế và biến động thời vụ.....	27
2.3.1. Dự báo dựa vào mô hình cộng.....	27
2.3.2. Dự báo dựa vào mô hình nhân.....	28
2.4. Dự báo theo phương pháp san bằng mũ.....	29
2.4.1. Mô hình đơn giản (phương pháp san bằng mũ đơn giản).....	29
2.4.2. Mô hình xu thế tuyến tính và không có biến động thời vụ (Mô hình san mũ Holt – Winters).....	32
2.4.3. Mô hình xu thế tuyến tính và biến động thời vụ.....	33
CHƯƠNG 3: THỰC NGHIỆM CÁC BÀI TOÁN DỰ BÁO.....	36
3.1. Phần mềm IBM SPSS Modeler.....	36
3.1.1. Giới thiệu.....	36
3.1.2. Các chức năng trong SPSS Modeler.....	36
3.2. Áp dụng phần mềm IBM SPSS Modeler vào bài toán dự báo.....	39
3.2.1. Bài toán 1 (sử dụng phương pháp định tính).....	39
3.2.2. Bài toán 2 (sử dụng phương pháp định lượng).....	46
3.2.3. Bài toán 3 (sử dụng phương pháp định lượng).....	49
KẾT LUẬN.....	54
TÀI LIỆU THAM KHẢO.....	55

DANH MỤC BẢNG VÀ HÌNH VẼ

DANH MỤC BẢNG

Bảng 3.2. 1 Diễn giải.....	40
Bảng 3.2. 2 Giá bán sản phẩm.....	46
Bảng 3.2. 3 Sản lượng doanh nghiệp A	49

DANH MỤC HÌNH VẼ

Hình 3.1.2. 1 Nguồn dữ liệu.....	36
Hình 3.1.2. 2 Trích chọn dữ liệu	37
Hình 3.1.2. 3 Biến đổi dữ liệu	37
Hình 3.1.2. 4 Lọc các trường dữ liệu	37
Hình 3.1.2. 5 Biến đổi trường và thuộc tính dữ liệu	38
Hình 3.1.2. 6 Xử lý dữ liệu	38
Hình 3.1.2. 7 Đánh giá và biểu diễn kết quả	39
Hình 3.2.1. 1 Chọn nút Type.....	40
Hình 3.2.1. 2 Chọn nút Filter	41
Hình 3.2.1. 3 Chọn nút select.....	41
Hình 3.2.1. 4 Xử lý với tập Nguoi.sav	42
Hình 3.2.1. 5 Chọn nút Merge.....	42
Hình 3.2.1. 6 Chọn nút Statistic	43
Hình 3.2.1. 7 Kết quả số lượng	43
Hình 3.2.1. 8 Chọn nút Table.....	44
Hình 3.2.1. 10 Bảng kết quả.....	44
Hình 3.2.1. 11 Chọn nút Distribution.....	45
Hình 3.2.1. 12 Bảng kết quả.....	45
Hình 3.2.2. 1 Chọn nút Type.....	47

Hình 3.2.2. 2 Chọn nút Filter	47
Hình 3.2.2. 3 Chọn nút Derive	48
Hình 3.2.2. 4 Chọn nút Table	48
Hình 3.2.2. 5 Kết quả	48
Hình 3.2.3. 1 Chọn nút Filter	50
Hình 3.2.3. 2 Chọn nút Type	50
Hình 3.2.3. 3 Chọn nút Time Intervals	50
Hình 3.2.3. 4 Xử lý nút Time Intervals	51
Hình 3.2.3. 5 Chọn nút Time Series	51
Hình 3.2.3. 6 Kết quả xử lý Time Series	51
Hình 3.2.3. 7 Bảng kết quả dự báo	52
Hình 3.2.3. 8 Chọn nút Multiplot và Plot	52
Hình 3.2.3. 9 Biểu đồ kết quả qua Multiplot	53
Hình 3.2.3. 10 Biểu đồ kết quả qua Plot	53

LỜI MỞ ĐẦU

Dự báo luôn gắn liền với cuộc sống của con người, từ các dự báo đơn giản về thời tiết, môi trường sống,... đến các dự báo quan trọng trong lĩnh vực chính trị, quân sự, kinh doanh và trong các lĩnh vực khác. Với chiều dài phát triển như vậy dự báo vẫn chưa chính thức là một ngành khoa học độc lập cho đến những năm đầu của thập niên 60 của thế kỉ trước, khoa học dự báo với tư cách một ngành khoa học độc lập có đầy đủ hệ thống lí luận và phương pháp luận.

Mỗi lĩnh vực có một yêu cầu về dự báo riêng, đối với nhà quản trị khi lên kế hoạch, trong hiện tại họ xác định hướng tương lai cho các hoạt động mà họ sẽ thực hiện. Bước đầu tiên trong hoạch định là dự báo hay là ước lượng nhu cầu tương lai cho sản phẩm hoặc dịch vụ và các nguồn lực cần thiết để sản xuất sản phẩm hoặc dịch vụ đó.

Dự báo sẽ mở ra một cửa sổ để hướng tới tương lai. Nó là con đường dẫn tới việc lập kế hoạch cho sự phát triển của tương lai. Những tầm nhìn từ dự báo sẽ giúp mở ra nhiều lựa chọn hơn cho tương lai để ta có quyết định chọn hay không chọn.

Trong một thế giới thay đổi với nhịp độ nhanh như hiện nay, tương lai sẽ có xu hướng rất khác so với thực tế hiện tại với rất nhiều cách khác nhau. Hơn nữa, do sự phát triển của những tri thức mới và những tiến bộ trong khoa học (và tiếp theo là sự tiến bộ của công nghệ), xã hội học, chính trị, kinh tế và kinh doanh, xã hội toàn cầu của chúng ta có khả năng ngày càng tăng để hình thành (theo hướng tích cực hoặc tiêu cực) tương lai mà chúng ta sẽ phải đạt được.

Kết quả là xã hội và những tổ chức trong nó phải tìm những kiến thức dự báo về tương lai có thể xảy ra và những hậu quả đối với những hành động ngày hôm nay và những hành động cần thiết. Do đó việc ngày càng cần thiết là chúng ta có những công cụ dự báo tốt hơn và áp dụng chúng theo những cách ta có thể. Điều đó cho thấy dự báo là càng quan trọng nhiều hơn, ngành dự báo đã trở thành một công cụ cần thiết để mọi người sử dụng trong nỗ lực đưa ra định những quyết, kế hoạch, thiết kế, chỉ đạo, quản lý, thực hiện và kiểm soát thay đổi bằng cách xác định tương lai thích hợp hơn với những dự báo.

Xác định được tầm quan trọng của lĩnh vực dự báo như vậy nên em đã chọn đề tài **“Tìm hiểu và tích hợp các phương pháp dự báo.”**

Xuất phát từ yêu cầu của đề tài, đồ án được chia làm các phần như sau:

Lời mở đầu: Giới thiệu về dự báo và vai trò

Chương 1: BÀI TOÁN PHÂN TÍCH VÀ DỰ BÁO

Trong chương này trình bày các khái niệm và kiến thức cơ bản về dự báo cũng như các phương pháp dự báo.

Chương 2: HỆ THỐNG PHÂN TÍCH VÀ DỰ BÁO

Tìm hiểu các phương pháp cơ bản để xây dựng hệ thống phân tích và dự báo.

Chương 3: THỰC NGHIỆM CÁC BÀI TOÁN DỰ BÁO

Giới thiệu và áp dụng phần mềm SPSS Modeler của hãng IBM giải quyết các bài toán dự báo.

Kết luận

Tài liệu tham khảo

CHƯƠNG 1: BÀI TOÁN PHÂN TÍCH VÀ DỰ BÁO

1.1. Khái niệm

Dự báo là một khoa học và nghệ thuật tiên đoán những sự việc sẽ xảy ra trong tương lai, trên cơ sở phân tích khoa học về các dữ liệu đã thu thập được. Khi tiến hành dự báo ta căn cứ vào việc thu thập xử lý số liệu trong quá khứ và hiện tại để xác định xu hướng vận động của các hiện tượng trong tương lai nhờ vào một số mô hình toán học.

Dự báo có thể là một dự đoán chủ quan hoặc trực giác về tương lai. Nhưng để cho dự báo được chính xác hơn, người ta cố loại trừ những tính chủ quan của người dự báo.

Ngày nay, dự báo là một nhu cầu không thể thiếu được của mọi hoạt động kinh tế - xã hội, khoa học - kỹ thuật, được tất cả các ngành khoa học quan tâm nghiên cứu.

1.2. Ý nghĩa và vai trò của phân tích và dự báo trong quá trình ra quyết định kinh doanh

1.2.1. Ý nghĩa

- Dùng để dự báo các mức độ tương lai của hiện tượng, qua đó giúp các nhà quản trị doanh nghiệp chủ động trong việc đề ra các kế hoạch và các quyết định cần thiết phục vụ cho quá trình sản xuất kinh doanh, đầu tư, quảng bá, quy mô sản xuất, kênh phân phối sản phẩm, nguồn cung cấp tài chính,... và chuẩn bị đầy đủ điều kiện cơ sở vật chất, kỹ thuật cho sự phát triển trong thời gian tới (kế hoạch cung cấp các yếu tố đầu vào như: lao động, nguyên vật liệu, tư liệu lao động,... cũng như các yếu tố đầu ra dưới dạng sản phẩm vật chất và dịch vụ).

- Trong các doanh nghiệp nếu công tác dự báo được thực hiện một cách nghiêm túc còn tạo điều kiện nâng cao khả năng cạnh tranh trên thị trường.

- Dự báo chính xác sẽ giảm bớt mức độ rủi ro cho doanh nghiệp nói riêng và toàn bộ nền kinh tế nói chung.

- Dự báo chính xác là căn cứ để các nhà hoạch định các chính sách phát triển kinh tế văn hoá xã hội trong toàn bộ nền kinh tế quốc dân.

- Nhờ có dự báo các chính sách kinh tế, các kế hoạch và chương trình phát triển kinh tế được xây dựng có cơ sở khoa học và mang lại hiệu quả kinh tế cao.

- Nhờ có dự báo thường xuyên và kịp thời, các nhà quản trị doanh nghiệp có khả năng kịp thời đưa ra những biện pháp điều chỉnh các hoạt động kinh tế của doanh nghiệp nhằm thu được hiệu quả sản xuất kinh doanh cao nhất.

1.2.2. Vai trò

- Dự báo tạo ra lợi thế cạnh tranh.

- Công tác dự báo là một bộ phận không thể thiếu trong hoạt động của các doanh nghiệp, trong từng phòng ban như: phòng Kinh doanh hoặc Marketing, phòng Sản xuất hoặc phòng Nhân sự, phòng Kế toán – tài chính.

1.3. Phân loại dự báo

1.3.1. Căn cứ vào độ dài thời gian dự báo:

Dự báo dài hạn: Khoảng thời gian từ 3 năm trở lên. Dự báo dài hạn được ứng dụng cho lập kế hoạch sản xuất sản phẩm mới, kế hoạch nghiên cứu và ứng dụng công nghệ mới, định vị doanh nghiệp hay mở rộng doanh nghiệp.

Dự báo trung hạn: Khoảng thời gian dự báo thường là từ 3 tháng đến 3 năm. Nó cần cho việc lập kế hoạch sản xuất, kế hoạch bán hàng, dự thảo ngân sách, kế hoạch tiền mặt, huy động các nguồn lực và tổ chức hoạt động tác nghiệp.

Dự báo ngắn hạn: Khoảng thời gian dự báo có thể đến một năm, nhưng thường là ít hơn ba tháng. Loại dự báo này thường được dùng trong kế hoạch mua hàng, điều độ công việc, cân bằng nhân lực, phân chia công việc.

Dự báo trung hạn và dài hạn có ba đặc trưng khác với dự báo ngắn hạn:

Thứ nhất, dự báo trung hạn và dài hạn phải giải quyết nhiều vấn đề có tính toàn diện và yểm trợ cho các quyết định quản lý thuộc về hoạch định kế hoạch sản xuất sản phẩm và quá trình công nghệ.

Thứ hai, dự báo ngắn hạn thường dùng nhiều loại phương pháp luận hơn là dự báo dài hạn. Đối với các dự báo ngắn hạn người ta dùng phổ biến các kỹ thuật toán học như bình quân di động, san bằng mũ và hồi quy theo xu hướng. Nói cách khác thì các phương pháp ít định lượng được dùng để tiên đoán các vấn đề lớn toàn diện

như có cần đưa một sản phẩm mới nào đó vào danh sách các chủng loại mặt hàng của công ty không.

Thứ ba, dự báo ngắn hạn có khuynh hướng chính xác hơn dự báo dài hạn. Vì các yếu tố ảnh hưởng đến nhu cầu thay đổi hàng ngày, nếu kéo dài thời gian dự báo ra thì độ chính xác có khả năng giảm đi. Do vậy, cần phải thường xuyên cập nhật và hoàn thiện các phương pháp dự báo.

1.3.2. Dựa vào các phương pháp dự báo:

Dự báo có thể chia thành 3 nhóm:

- *Dự báo bằng phương pháp chuyên gia*: Loại dự báo này được tiến hành trên cơ sở tổng hợp, xử lý ý kiến của các chuyên gia thông thạo với hiện tượng được nghiên cứu, từ đó có phương pháp xử lý thích hợp đề ra các dự đoán, các dự đoán này được cân nhắc và đánh giá chủ quan từ các chuyên gia. Phương pháp này có ưu thế trong trường hợp dự đoán những hiện tượng hay quá trình bao quát rộng, phức tạp, chịu sự chi phối của khoa học - kỹ thuật, sự thay đổi của môi trường, thời tiết, chiến tranh trong khoảng thời gian dài. Một cải tiến của phương pháp Delphi – là phương pháp dự báo dựa trên cơ sở sử dụng một tập hợp những đánh giá của một nhóm chuyên gia. Mỗi chuyên gia được hỏi ý kiến và rồi dự báo của họ được trình bày dưới dạng thông kê tóm tắt. Việc trình bày những ý kiến này được thực hiện một cách gián tiếp (không có sự tiếp xúc trực tiếp) để tránh những sự tương tác trong nhóm nhỏ qua đó tạo nên những sai lệch nhất định trong kết quả dự báo. Sau đó người ta yêu cầu các chuyên gia duyệt xét lại những dự báo của họ trên cơ sở tóm tắt tất cả các dự báo có thể có những bổ sung thêm.

- *Dự báo theo phương trình hồi quy*: Theo phương pháp này, mức độ cần dự báo phải được xây dựng trên cơ sở xây dựng mô hình hồi quy, mô hình này được xây dựng phù hợp với đặc điểm và xu thế phát triển của hiện tượng nghiên cứu. Để xây dựng mô hình hồi quy, đòi hỏi phải có tài liệu về hiện tượng cần dự báo và các hiện tượng có liên quan. Loại dự báo này thường được sử dụng để dự báo trung hạn và dài hạn ở tầm vĩ mô.

- *Dự báo dựa vào dãy số thời gian*: Là dựa trên cơ sở dãy số thời gian phản ánh sự biến động của hiện tượng ở những thời gian đã qua để xác định mức độ của hiện tượng trong tương lai.

1.3.3. Căn cứ vào nội dung (đối tượng dự báo)

Có thể chia dự báo thành: Dự báo khoa học, dự báo kinh tế, dự báo xã hội, dự báo tự nhiên, thiên văn học...

- *Dự báo khoa học*: Là dự kiến, tiên đoán về những sự kiện, hiện tượng, trạng thái nào đó có thể hay nhất định sẽ xảy ra trong tương lai. Theo nghĩa hẹp hơn, đó là sự nghiên cứu khoa học về những triển vọng của một hiện tượng nào đó, chủ yếu là những đánh giá số lượng và chỉ ra khoảng thời gian mà trong đó hiện tượng có thể diễn ra những biến đổi.

- *Dự báo kinh tế*: Là khoa học dự báo các hiện tượng kinh tế trong tương lai. Dự báo kinh tế được coi là giai đoạn trước của công tác xây dựng chiến lược phát triển kinh tế - xã hội và dự án kế hoạch dài hạn; không đặt ra những nhiệm vụ cụ thể, nhưng chứa đựng những nội dung cần thiết làm căn cứ để xây dựng những nhiệm vụ đó. Dự báo kinh tế bao trùm sự phát triển kinh tế và xã hội của đất nước có tính đến sự phát triển của tình hình thế giới và các quan hệ quốc tế. Thường được thực hiện chủ yếu theo những hướng sau: dân số, nguồn lao động, việc sử dụng và tái sản xuất chúng, năng suất lao động; tái sản xuất xã hội trước hết là vốn sản xuất cố định: sự phát triển của cách mạng khoa học – kĩ thuật và công nghệ và khả năng ứng dụng vào kinh tế; mức sống của nhân dân, sự hình thành các nhu cầu phi sản xuất, động thái và cơ cấu tiêu dùng, thu nhập của nhân dân; động thái kinh tế quốc dân và sự chuyển dịch cơ cấu (nhịp độ, tỉ lệ, hiệu quả); sự phát triển các khu vực và ngành kinh tế (khối lượng động thái, cơ cấu, trình độ kĩ thuật, bộ máy, các mối liên hệ liên ngành); phân vùng sản xuất, khai thác tài nguyên thiên nhiên và phát triển các vùng kinh tế trong nước, các mối liên hệ liên vùng; dự báo sự phát triển kinh tế của thế giới kinh tế. Các kết quả dự báo kinh tế cho phép hiểu rõ đặc điểm của các điều kiện kinh tế - xã hội để đặt chiến lược phát triển kinh tế đúng đắn, xây dựng các chương trình, kế hoạch phát triển một cách chủ động, đạt hiệu quả cao và vững chắc.

- *Dự báo xã hội*: Dự báo xã hội là khoa học nghiên cứu những triển vọng cụ thể của một hiện tượng, một sự biến đổi, một quá trình xã hội, để đưa ra dự báo hay dự đoán về tình hình diễn biến, phát triển của một xã hội.

- *Dự báo tự nhiên, thiên văn học*, loại dự báo này thường bao gồm:

+ Dự báo thời tiết: Thông báo thời tiết dự kiến trong một thời gian nhất định trên một vùng nhất định. Trong dự báo thời tiết có dự báo chung, dự báo khu vực, dự báo địa phương, v.v. Về thời gian, có dự báo thời tiết ngắn (1-3 ngày) và dự báo thời tiết dài (tới một năm).

+ Dự báo thủy văn: Là loại dự báo nhằm tính để xác định trước sự phát triển các quá trình, hiện tượng thủy văn xảy ra ở các sông hồ, dựa trên các tài liệu liên quan tới khí tượng thủy văn. Dự báo thủy văn dựa trên sự hiểu biết những quy luật phát triển của các quá trình, khí tượng thủy văn, dự báo sự xuất hiện của hiện tượng hay yếu tố cần quan tâm. Căn cứ thời gian dự kiến, dự báo thủy văn được chia thành dự báo thủy văn hạn ngắn (thời gian không quá 2 ngày), hạn vừa (từ 2 đến 10 ngày); dự báo thủy văn mùa (thời gian dự báo vài tháng); cấp báo thủy văn: thông tin khẩn cấp về hiện tượng thủy văn gây nguy hiểm. Theo mục đích dự báo, có các loại: dự báo thủy văn phục vụ thi công, phục vụ vận tải, phục vụ phát điện, v.v. Theo yếu tố dự báo, có: dự báo lưu lượng lớn nhất, nhỏ nhất, dự báo lũ, v.v.

+ Dự báo địa lý: Là việc nghiên cứu về hướng phát triển của môi trường địa lí trong tương lai, nhằm đề ra trên cơ sở khoa học những giải pháp sử dụng hợp lí và bảo vệ môi trường.

+ Dự báo động đất: Là loại dự báo trước địa điểm và thời gian có khả năng xảy ra động đất. Động đất không đột nhiên xảy ra mà là một quá trình tích lũy lâu dài, có thể hiện ra trước bằng những biến đổi địa chất, những hiện tượng vật lí, những trạng thái sinh học bất thường ở động vật, v.v. Việc dự báo thực hiện trên cơ sở nghiên cứu bản đồ phân vùng động đất và những dấu hiệu báo trước. Cho đến nay, chưa thể dự báo chính xác về thời gian động đất sẽ xảy ra.

1.4. Các phương pháp dự báo

1.4.1. Phương pháp dự báo định tính

Phương pháp dự báo định tính dựa trên cơ sở doanh số của từng sản phẩm hay dịch vụ riêng biệt và dựa trên những ý kiến về các khả năng có liên hệ của những nhân tố nhân quả trong tương lai. Những phương pháp này có liên quan đến mức độ phức tạp của khảo sát ý kiến được tiến hành một cách khoa học để nhận biết về các sự kiện tương lai. Dưới đây là các dự báo định tính thường dùng:

1.4.1.1. Lấy ý kiến của ban điều hành

Phương pháp lấy ý kiến của ban điều hành sử dụng rộng rãi ở các tổ chức. Khi thực hiện công việc dự báo, họ tổng hợp ý kiến của các nhà quản trị cấp cao, trưởng các bộ phận của doanh nghiệp, và sử dụng các số liệu thống kê về những chỉ tiêu tổng hợp: doanh số, chi phí, lợi nhuận,...Ngoài ra cần lấy thêm ý kiến của các chuyên gia về marketing, tài chính, sản xuất, kỹ thuật.

Nhược điểm lớn nhất của phương pháp này là có tính chủ quan của các thành viên và ý kiến của người có chức vụ cao nhất thường chi phối ý kiến của những người khác.

1.4.1.2. Lấy ý kiến của người bán hàng

Những người bán hàng tiếp xúc thường xuyên với khách hàng, do đó họ hiểu rõ nhu cầu, thị hiếu của người tiêu dùng. Họ có thể dự đoán được lượng hàng tiêu thụ tại khu vực họ phụ trách.

Tập hợp ý kiến của nhiều người bán hàng tại nhiều khu vực khác nhau, ta có được lượng dự báo tổng hợp về nhu cầu đối với loại sản phẩm đang xét.

Nhược điểm của phương pháp này là phụ thuộc vào đánh giá chủ quan của người bán hàng. Một số có khuynh hướng lạc quan đánh giá cao lượng hàng bán ra của người bán. Ngược lại, một số khác lại muốn giảm xuống để dễ đạt định mức.

1.4.1.3. Phương pháp chuyên gia (Delphi)

Phương pháp này thu thập ý kiến của các chuyên gia trong hoặc ngoài doanh nghiệp theo những mẫu câu hỏi được in sẵn và được thực hiện như sau:

- Mỗi chuyên gia được phát một thư yêu cầu trả lời một số câu hỏi phục vụ cho việc dự báo.
- Nhân viên dự báo tập hợp các câu trả lời, sắp xếp chọn lọc và tóm tắt lại các ý kiến của các chuyên gia.
- Dựa vào bảng tóm tắt này nhân viên dự báo lại tiếp tục nêu ra các câu hỏi để các chuyên gia trả lời tiếp.
- Tập hợp các ý kiến mới của các chuyên gia. Nếu chưa thỏa mãn thì tiếp tục quá trình nêu trên cho đến khi đạt yêu cầu dự báo.

Ưu điểm của phương pháp này là tránh được các liên hệ cá nhân với nhau, không xảy ra va chạm giữa các chuyên gia và họ không bị ảnh hưởng bởi ý kiến của một người nào đó có ưu thế trong số người được hỏi ý kiến.

1.4.1.4. Phương pháp điều tra người tiêu dùng

Người ta áp dụng phương pháp để thu thập nguồn thông tin từ người tiêu dùng về nhu cầu hiện tại cũng như tương lai. Cuộc điều tra nhu cầu được thực hiện bởi những nhân viên bán hàng hoặc nhân viên nghiên cứu thị trường. Họ thu thập ý kiến khách hàng thông qua phiếu điều tra, phỏng vấn trực tiếp hay điện thoại,... Cách tiếp cận này không những giúp cho doanh nghiệp về dự báo nhu cầu mà cả trong việc cải tiến thiết kế sản phẩm. Phương pháp này mất nhiều thời gian, việc chuẩn bị phức tạp, khó khăn và tốn kém, có thể không chính xác trong các câu trả lời của người tiêu dùng.

1.4.2. Phương pháp dự báo định lượng

Dự báo định lượng dựa trên số liệu thống kê trong quá khứ, những số liệu này giả sử có liên quan đến tương lai. Tất cả các mô hình dự báo theo định lượng có thể sử dụng thông qua chuỗi thời gian và các giá trị này được quan sát đo lường các giai đoạn theo từng chuỗi.

- Tính chính xác của dự báo:

Tính chính xác của dự báo đề cập đến độ chênh lệch của dự báo với số liệu thực tế. Bởi vì dự báo được hình thành trước khi số liệu thực tế xảy ra, vì vậy tính chính xác của dự báo chỉ có thể đánh giá sau khi thời gian đã qua đi. Nếu dự báo càng gần với số liệu thực tế, ta nói dự báo có độ chính xác cao và lỗi trong dự báo càng thấp.

Người ta thường dùng độ sai lệch tuyệt đối bình quân (MAD) để tính toán:

$$\text{MAD} = \frac{\text{Tổng các sai số tuyệt đối của } n \text{ giai đoạn}}{n \text{ giai đoạn}}$$

$$\sum_{i=1}^n \text{ Nhu cầu thực tế - nhu cầu dự báo}$$

$$\text{MAD} = \frac{\quad}{N}$$

1.4.2.1. Dự báo ngắn hạn

Dự báo ngắn hạn ước lượng tương lai trong thời gian ngắn, có thể từ vài ngày đến vài tháng. Dự báo ngắn hạn cung cấp cho các nhà quản lý những thông tin để đưa ra quyết định về các vấn đề như:

- Cần dự trữ bao nhiêu đối với một loại sản phẩm cụ thể nào đó cho tháng tới ?
- Lên lịch sản xuất từng loại sản phẩm cho tháng tới như thế nào ?
- Số lượng nguyên vật liệu cần đặt hàng để nhận vào tuần tới là bao nhiêu ?

* *Dự báo sơ bộ:*

Mô hình dự báo sơ bộ là loại dự báo nhanh, không cần chi phí và dễ sử dụng.

Ví dụ như:

- Sử dụng số liệu hàng bán ngày hôm nay làm dự báo cho lượng hàng bán ở ngày mai.
- Sử dụng số liệu ngày này ở năm rồi như là dự báo lượng hàng bán cho ngày ấy ở năm nay.

Mô hình dự báo sơ bộ quá đơn giản cho nên thường hay gặp những sai sót trong dự báo.

* *Phương pháp bình quân di động có quyền số.*

Phương pháp bình quân di động xem vai trò của các số liệu trong quá khứ là như nhau. Trong một vài trường hợp, các số liệu này có ảnh hưởng khác nhau trên kết quả dự báo, vì thế, người ta thích sử dụng quyền số không đồng đều cho các số liệu quá khứ. Quyền số hay trọng số là các con số được gán cho các số liệu quá khứ để chỉ mức độ quan trọng của chúng ảnh hưởng đến kết quả dự báo. Quyền số lớn được gán cho

số liệu gần với kỳ dự báo nhất để ám chỉ ảnh hưởng của nó là lớn nhất. Việc chọn các quyền số phụ thuộc vào kinh nghiệm và sự nhạy cảm của người dự báo.

Công thức tính toán:

$$F_t = \frac{\sum_{i=1}^n A_{t-i} \cdot k_i}{\sum_{i=1}^n k_i}$$

Với: F_t - Dự báo thời kỳ thứ t

A_{t-i} - Số liệu thực tế thời kỳ trước ($i=1,2,\dots,n$)

k_i - Quyền số tương ứng ở thời kỳ i

Phương pháp bình quân di động có quyền số có ưu điểm là san bằng được các biến động ngẫu nhiên trong dãy số. Tuy vậy, phương pháp này có nhược điểm sau:

- Do việc san bằng các biến động ngẫu nhiên nên làm giảm độ nhạy cảm đối với những thay đổi thực đã được phản ánh trong dãy số.

- Số bình quân di động chưa cho chúng ta xu hướng phát triển của dãy số một cách tốt nhất. Nó chỉ thể hiện sự vận động trong quá khứ chứ chưa thể kéo dài sự vận động đó trong tương lai.

** Phương pháp điều hòa mũ.*

Điều hòa mũ đưa ra các dự báo cho giai đoạn trước và thêm vào đó một lượng điều chỉnh để có được lượng dự báo cho giai đoạn kế tiếp. Sự điều chỉnh này là một tỷ lệ nào đó của sai số dự báo ở giai đoạn trước và được tính bằng cách nhân số dự báo của giai đoạn trước với hệ số nằm giữa 0 và 1. Hệ số này gọi là hệ số điều hòa.

Công thức tính như sau: $F_t = F_{t-1} + \alpha (A_{t-1} - F_{t-1})$

Trong đó : F_t - Dự báo cho giai đoạn thứ t , giai đoạn kế tiếp.

F_{t-1} - Dự báo cho giai đoạn thứ $t-1$, giai đoạn trước.

A_{t-1} - Số liệu thực tế của giai đoạn thứ $t-1$

** Phương pháp điều hòa mũ theo xu hướng*

Đối với kế hoạch ngắn hạn thì mùa vụ và xu hướng là nhân tố không quan trọng. Khi chúng ta chuyển từ dự báo ngắn hạn sang dự báo trung hạn thì mùa vụ và xu hướng trở nên quan trọng hơn. Kết hợp nhân tố xu hướng vào dự báo điều hòa mũ được gọi là điều hòa mũ theo xu hướng hay điều hòa đôi.

Vì ước lượng cho số trung bình và ước lượng cho xu hướng cho số trung bình và hệ số điều hòa α được điều hòa cả hai. Hệ số điều hòa cho xu hướng, được sử dụng trong mô hình này β .

Công thức tính toán như sau:

$$FT_t = S_{t-1} + T_{t-1}(A_t - FT_t)\alpha$$

Với: $S_t = FT_t + (FT_t - FT_{t-1} - T_{t-1})\beta$ $T_t = T_{t-1}$

Trong đó FT_t - Dự báo theo xu hướng trong giai đoạn t

S_t - Dự báo đã được điều hòa trong giai đoạn t

T_t - Ước lượng xu hướng trong giai đoạn t

A_t - Số liệu thực tế trong giai đoạn t

t - Thời đoạn kế tiếp.

$t-1$ - Thời đoạn trước.

1.4.2.2. Dự báo dài hạn

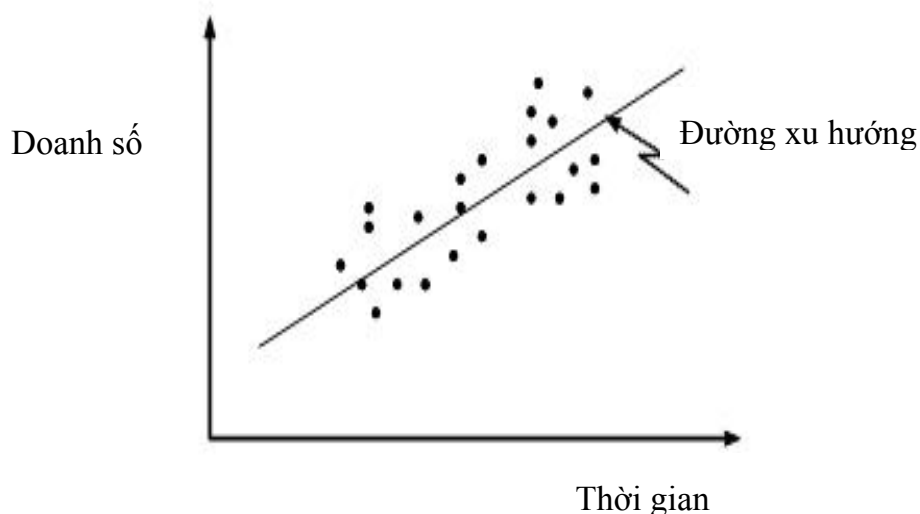
Dự báo dài hạn là ước lượng tương lai trong thời gian dài, thường hơn một năm. Dự báo dài hạn rất cần thiết trong quản trị sản xuất để trợ giúp các quyết định chiến lược về hoạch định sản phẩm, quy trình công nghệ và các phương tiện sản xuất. Ví dụ như:

- Thiết kế sản phẩm mới.

- Xác định năng lực sản xuất cần thiết là bao nhiêu? Máy móc, thiết bị nào cần sử dụng và chúng được đặt ở đâu ?

- Lên lịch trình cho những nhà cung ứng theo các – hợp đồng cung cấp nguyên vật liệu dài hạn.

Dự báo dài hạn có thể được xây dựng bằng cách vẽ một đường thẳng đi xuyên qua các số liệu quá khứ và kéo dài nó đến tương lai. Dự báo trong giai đoạn kế tiếp có thể được vẽ vượt ra khỏi đồ thị thông thường. Phương pháp tiếp cận theo kiểu đồ thị đối với dự báo dài hạn có thể dùng trong thực tế, nhưng điểm không thuận lợi của nó là vấn đề vẽ một đường tương ứng hợp lý nhất đi qua các số liệu quá khứ này.



Phân tích hồi qui sẽ cung cấp cho chúng ta một phương pháp làm việc chính xác để xây dựng đường dự báo theo xu hướng.

** Phương pháp hồi qui tuyến tính.*

Phân tích hồi qui tuyến tính là một mô hình dự báo thiết lập mối quan hệ giữa biến phụ thuộc với hai hay nhiều biến độc lập. Trong phần này, chúng ta chỉ xét đến một biến độc lập duy nhất. Nếu số liệu là một chuỗi theo thời gian thì biến độc lập là giai đoạn thời gian và biến phụ thuộc thông thường là doanh số bán ra hay bất kỳ chỉ tiêu nào khác mà ta muốn dự báo.

Mô hình này có công thức: $Y = ax + b$

$$a = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2} \quad b = \frac{\sum x^2 \sum y - \sum x \sum xy}{n \sum x^2 - (\sum x)^2}$$

Trong đó : y - Biến phụ thuộc cần dự báo.

x - Biến độc lập

a - Độ dốc của đường xu hướng

b - Tung độ gốc

n - Số lượng quan sát

Trong trường hợp biến độc lập x được trình bày thông qua từng giai đoạn theo thời gian và chúng phải cách đều nhau (như : $x = 0$. Vì vậy $\sum 2002, 2003, 2004, \dots$) thì ta có thể điều chỉnh lại để sao cho việc tính toán sẽ trở nên đơn giản và dễ dàng hơn nhiều.

Nếu có một số lẻ lượng mốc thời gian: chẳng hạn $x = 0$, \sum là 5, thì giá trị của x được ấn định như sau : -2, -1, 0, 1, 2 và như thế giá trị của x được sử dụng cho dự báo trong năm tới là +3.

Nếu có một số chẵn lượng mốc thời gian: chẳng hạn $x = 0$ và \sum là 6 thì giá trị của x được ấn định là : -5, -3, -1, 1, 3, 5. Như thế giá trị của x được dùng cho dự báo trong năm tới là +7.

Trường hợp biến độc lập không phải là biến thời gian, hồi qui tuyến tính là một nhóm các mô hình dự báo được gọi là mô hình nhân quả. Mô hình này đưa ra các dự báo sau khi thiết lập và đo lường các biến phụ thuộc với một hay nhiều biến độc lập.

** Tính chất mùa vụ trong dự báo chuỗi thời gian.*

Loại mùa vụ thông thường là sự lên xuống xảy ra trong vòng một năm và có xu hướng lặp lại hàng năm. Những vụ mùa này xảy ra có thể do điều kiện thời tiết, địa lý hoặc do tập quán của người tiêu dùng khác nhau...

Cách thức xây dựng dự báo với phân tích hồi qui tuyến tính khi vụ mùa hiện diện trong chuỗi số theo thời gian. Ta thực hiện các bước:

- Chọn lựa chuỗi số liệu quá khứ đại diện.
- Xây dựng chỉ số mùa vụ cho từng giai đoạn thời gian.

$$I_i = \frac{\bar{Y}_i}{\bar{Y}_0}$$

Với \bar{Y}_i - Số bình quân của các thời kỳ cùng tên.

\bar{Y}_0 - Số bình quân chung của tất cả các thời kỳ trong dãy số.

I_i - Chỉ số mùa vụ kỳ thứ i.

- Sử dụng các chỉ số mùa vụ để hóa giải tính chất mùa vụ của số liệu.
- Phân tích hồi qui tuyến tính dựa trên số liệu đã phi mùa vụ.
- Sử dụng phương trình hồi qui để dự báo cho tương lai.
- Sử dụng chỉ số mùa vụ để tái ứng dụng tính chất mùa vụ cho dự báo.

1.5. Quy trình dự báo

Quy trình dự báo được chia thành 9 bước:

Bước 1: Xác định mục tiêu:

- Các mục tiêu liên quan đến các quyết định cần đến dự báo phải được nói rõ. Nếu quyết định vẫn không thay đổi bất kể có dự báo hay không thì mọi nỗ lực thực hiện dự báo cũng vô ích.

- Nếu người sử dụng và người làm dự báo có cơ hội thảo luận các mục tiêu và kết quả dự báo sẽ được sử dụng như thế nào, thì kết quả dự báo sẽ có ý nghĩa quan trọng.

Bước 2: Xác định dự báo cái gì:

- Khi các mục tiêu tổng quát đã rõ ta phải xác định chính xác là dự báo cái gì (cần có sự trao đổi).

+ Ví dụ: Chỉ nói dự báo doanh số không thì chưa đủ, mà cần phải hỏi rõ hơn là: Dự báo doanh thu bán hàng (sales revenue) hay số đơn vị doanh số (unit sales). Dự báo theo năm, quý, tháng hay tuần.

+ Nên dự báo theo đơn vị để tránh những thay đổi của giá cả.

Bước 3: Xác định khía cạnh thời gian

Có 2 loại khía cạnh thời gian cần xem xét:

- Thứ nhất: Độ dài dự báo, cần lưu ý:

+ Đối với dự báo theo năm: từ 1 đến 5 năm.

+ Đối với dự báo quý: từ 1 hoặc 2 năm.

+ Đối với dự báo tháng: từ 12 đến 18 tháng.

- Thứ hai: Người sử dụng và người làm dự báo phải thống nhất tính cấp thiết của dự báo.

Bước 4: Xem xét dữ liệu:

- Dữ liệu cần để dự báo có thể từ 2 nguồn: bên trong và bên ngoài.
- Cần phải lưu ý dạng dữ liệu sẵn có (thời gian, đơn vị tính,...).
- Dữ liệu thường được tổng hợp theo cả biến và thời gian, nhưng tốt nhất là thu thập dữ liệu chưa được tổng hợp.
- Cần trao đổi giữa người sử dụng và người làm dự báo.

Bước 5: Lựa chọn mô hình:

- Làm sao để quyết định được phương pháp thích hợp nhất cho một tình huống nhất định?

- + Loại và lượng dữ liệu sẵn có.
- + Mô hình (bản chất) dữ liệu quá khứ.
- + Tính cấp thiết của dự báo.
- + Độ dài dự báo.
- + Kiến thức chuyên môn của người làm dự báo.

Bước 6: Đánh giá mô hình:

- Đối với các phương pháp định tính thì bước này ít phù hợp hơn so với phương pháp định lượng .

- Đối với các phương pháp định lượng, cần phải đánh giá mức độ phù hợp của mô hình (trong phạm vi mẫu dữ liệu).

- Đánh giá mức độ chính xác của dự báo (ngoài phạm vi mẫu dữ liệu).
- Nếu mô hình không phù hợp, quay lại bước 5.

Bước 7: Chuẩn bị dự báo:

- Nếu có thể nên sử dụng hơn một phương pháp dự báo, và nên là những loại phương pháp khác nhau (ví dụ mô hình hồi quy và san mũ Holt, thay vì cả 2 mô hình hồi quy khác nhau).

- Các phương pháp được chọn nên được sử dụng để chuẩn bị cho một số các dự báo (ví dụ trường hợp xấu nhất, tốt nhất và có thể nhất).

Bước 8: Trình bày kết quả dự báo:

- Kết quả dự báo phải được trình bày rõ ràng cho ban quản lý sao cho họ hiểu các con số được tính toán như thế nào và chỉ ra sự tin cậy trong kết quả dự báo.

- Người dự báo phải có khả năng trao đổi các kết quả dự báo theo ngôn ngữ mà các nhà quản lý hiểu được.

- Trình bày cả ở dạng viết và dạng nói.

- Bảng biểu phải ngắn gọn, rõ ràng.

- Chỉ cần trình bày các quan sát và dự báo gần đây thôi.

- Chuỗi dữ liệu dài có thể được trình bày dưới dạng đồ thị (cả giá trị thực và dự báo).

- Trình bày thuyết trình nên theo cùng hình thức và cùng mức độ với phần trình bày viết.

Bước 9: Theo dõi kết quả dự báo:

- Lệch giữa giá trị dự báo và giá trị thực phải được thảo luận một cách tích cực, khách quan và cởi mở.

- Mục tiêu của việc thảo luận là để hiểu tại sao có các sai số, để xác định độ lớn của sai số.

- Trao đổi và hợp tác giữa người sử dụng và người làm dự báo có vai trò rất quan trọng trong việc xây dựng và duy trì quy trình dự báo thành công.

CHƯƠNG 2: HỆ THỐNG PHÂN TÍCH VÀ DỰ BÁO

Các hệ thống phân tích và dự báo được xây dựng dựa trên yêu cầu thực tế của từng bài toán, qua đó lựa chọn các phương pháp dự báo phù hợp để xây dựng hệ thống là một yếu tố quan trọng của những chuyên gia làm trong lĩnh vực dự báo. Trong đề án này chúng tôi xây dựng hệ thống dự báo dựa trên các phương pháp phân tích và dự báo cơ bản để từ đó nắm được nguyên lý xây dựng hệ thống phân tích và dự báo.

2.1. Dự báo từ các mức độ bình quân

2.1.1. Dự báo từ số bình quân trượt (di động)

Phương pháp số bình quân di động là một trong những phương pháp biểu hiện xu hướng phát triển cơ bản của hiện tượng nghiên cứu, hay nói cách khác, mô hình hoá sự phát triển thực tế của hiện tượng nghiên cứu dưới dạng dãy các số bình quân di động.

Phương pháp bình quân di động còn được sử dụng trong dự báo thống kê. Trên cơ sở xây dựng một dãy số bình quân di động, người ta xây dựng mô hình dự báo.

Mô hình dự báo là: $\hat{y}_{n+1} = M_n$

Khoảng dự báo sẽ được xác định theo công thức sau:

$$\hat{y}_{n+L} \pm t_\alpha S \sqrt{1 + \frac{1}{k}}$$

Trong đó t_α là giá trị tra trong bảng tiêu chuẩn t- Student với $(k-1)$ bậc tự do và xác suất tin cậy $(1-\alpha)$. Độ lệch tiêu chuẩn mẫu điều chỉnh được tính theo công thức sau:

$$S = \sqrt{\frac{\sum (y_i - M_i)^2}{k-1}}$$

2.1.2. Mô hình dự báo dựa vào lượng tăng (giảm) tuyệt đối bình quân

- Phương pháp này được sử dụng trong trường hợp lượng tăng (giảm) tuyệt đối liên hoàn xấp xỉ nhau qua các năm (dãy số thời gian có dạng gần giống như cấp số cộng): $\Delta y = y_i - y_{i-1}$ xấp xỉ nhau ($i = z \div n$).

Mô hình dự báo theo phương trình:

$$\hat{Y}_{n+L} = y_n + \bar{\Delta}_y \cdot L$$

Trong đó:

\hat{Y}_{n+L} : Mức độ dự đoán ở thời gian (n+L)

y_n : Mức độ cuối cùng của dãy số thời gian

$\bar{\Delta}_y$: Lượng tăng, giảm tuyệt đối bình quân

L: Tầm xa của dự đoán (L=1,2,3,...năm)

Trong đó:

$$\begin{aligned} \bar{\Delta}_y &= \frac{\sum(y_i - y_{i-1})}{n-1} \quad (i = \overline{2, n}) \\ &= \frac{y_i - y_{i-1}}{n-1} \end{aligned}$$

2.1.3. Mô hình dự báo dựa vào tốc độ phát triển bình quân

Thường áp dụng trong trường hợp các mức độ của dãy số biến động theo thời gian có tốc độ phát triển (hoặc tốc độ tăng, giảm) từng kỳ gần nhau (dãy số thời gian có dạng gần như cấp số nhân).

Có hai mô hình dự đoán:

* *Dự đoán mức độ hàng năm*: (có thể dùng để dự báo trong dài hạn).

- Phương pháp này được áp dụng khi tốc độ phát triển hoàn toàn xấp xỉ nhau.

- Mô hình dự đoán:

$$\hat{Y}_{n+L} = y_n \cdot t^L$$

\hat{Y}_{n+L} : Mức độ dự đoán ở thời gian (n+L)

y_n : Mức độ được dùng làm kỳ gốc để ngoại suy

L: Tầm xa của dự đoán (L=1,2,3,...năm)

\bar{t} : Tốc độ phát triển bình quân hàng năm

$$\bar{t} = \sqrt[n-1]{\frac{y_n}{y_1}}$$

*Dự đoán mức độ của khoảng thời gian dưới 1 năm (quý, tháng- dự báo ngắn hạn)

$$\hat{Y}_{ij} = y_j \frac{\bar{t}^{i-1}}{S_r}$$

Trong đó:

\hat{Y}_{ij} : Là mức độ của hiện tượng ở thời gian j (j=1,m) của năm i

$y_j = \sum_{i=1}^n Y_{ij}$ - Tổng các mức độ của thời gian j của năm i (i=1...n)

$\bar{t} = \sqrt[n-1]{\frac{y_n}{y_1}}$: Tốc độ phát triển bình quân hàng

$$S_r = 1 + (\bar{t}) + (\bar{t})^2 + (\bar{t})^3 + \dots + (\bar{t})^{n-1}$$

n: có thể là số năm hoặc số lượng mức độ của từng năm.

2.2. Mô hình dự báo theo phương trình hồi quy (dự báo dựa vào xu thế)

Từ xu hướng phát triển của hiện tượng nghiên cứu ta xác định được phương trình hồi quy lý thuyết, đó là phương trình phù hợp với xu hướng và đặc điểm biến động của hiện tượng nghiên cứu, từ đó có thể ngoại suy hàm xu thế để xác định mức độ phát triển trong tương lai.

2.2.1. Mô hình hồi quy theo thời gian

* Ví dụ: Mô hình dự báo theo phương trình hồi quy đường thẳng:

$$\hat{Y} = a + bt$$

Trong đó: a,b là những tham số quy định vị trí của đường hồi quy

Từ phương trình này, bằng phương pháp bình phương nhỏ nhất hoặc thông qua việc đặt thứ tự thời gian (t) trong dãy số để tính các tham số a,b.

Nếu đặt thứ tự thời gian t sao cho $\sum t$ khác 0 ($\sum t \neq 0$), ta có các công thức tính tham số sau:

$$a = \frac{\overline{yt} - \bar{y} \cdot \bar{t}}{\overline{t^2} - \bar{t}^2}$$

$$b = \bar{y} - a \cdot \bar{t}$$

Nếu đặt thứ tự thời gian t sao cho $\sum t$ khác 0 ($\sum t = 0$), ta có các công thức tính tham số sau:

$$a = \frac{\sum y}{n} = \bar{y} \quad b = \frac{\sum y \cdot t}{\sum t^2}$$

2.2.2. Mô hình hồi quy giữa các tiêu thức

* Mô hình hồi quy tuyến tính giữa hai tiêu thức

Từ việc xây dựng phương trình hồi quy tuyến tính giữa các tiêu thức đã nêu ở phần trên, ta có thể dự đoán các giá trị của Y trong tương lai khi các biến trong hàm hồi quy thay đổi, cụ thể:

Đối với phương trình tuyến tính giản đơn: $Y_x = a + bx$

Trong đó: a, b là những tham số quy định vị trí của đường hồi quy. Hằng số a là điểm cắt trục tung (biểu hiện của tiêu thức kết quả) khi tiêu thức nguyên nhân x bằng 0. Độ dốc b chính là lượng tăng giảm của tiêu thức kết quả khi tiêu thức nguyên nhân thay đổi.

Từ phương trình này, ta sẽ dự đoán được giá trị của tiêu thức kết quả trong tương lai khi có sự thay đổi của tiêu thức nguyên nhân.

Tương tự như trong hồi quy giản đơn, trong hồi quy bội, giá trị dự đoán của Y có được tương ứng với các giá trị cho trước của k biến X bằng cách thay các giá trị của k biến X vào phương trình hồi quy bội.

Các giá trị cho trước của biến X lần lượt là $x_{1,n+1}, x_{2,n+1}, \dots, x_{k,n+1}$ thì giá trị dự đoán Y_{n+1} sẽ là:

$$Y_{n+1} = a + b_1 x_{1,n+1} + b_2 x_{2,n+1} + \dots + b_k x_{k,n+1}$$

2.3. Dự báo dựa vào hàm xu thế và biến động thời vụ

Phương pháp dự báo này áp dụng đối với hiện tượng nghiên cứu chịu tác động của nhiều nhân tố biến động. Như biến động thời vụ, biến động xu hướng và biến động bất thường.

- Mô hình dự báo sẽ có thể dựa vào hàm xu thế kết hợp với biến động thời vụ:

$$Y_t = \hat{Y} + tv + bt$$

- Hoặc dự báo dựa vào hàm xu thế kết hợp nhân tố với biến động thời vụ:

$$Y_t = \hat{Y}_x + tv + xbt$$

Trong đó:

\hat{Y} : Mức độ lý thuyết xác định từ hàm xu thế (hoặc các phương pháp nêu trên)

tv : Ảnh hưởng của nhân tố thời vụ

bt : Ảnh hưởng của nhân tố bất thường

Nhìn chung, hàm xu thế, chỉ số thời vụ được xác định từng mô hình còn những nhân tố biến động bất thường thường không dự báo được, do vậy mô hình chỉ còn lại hai nhân tố: biến động xu hướng và biến động thời vụ.

2.3.1. Dự báo dựa vào mô hình cộng

- * Trước tiên xác định hàm xu thế tuyến tính sản lượng doanh nghiệp có dạng là:

$$\hat{Y} = a + bt$$

Trong đó: a, b là các tham số quy định vị trí của hàm xu thế tuyến tính, được tính theo công thức sau:

$$b = \frac{12}{m.n(n^2 - 1)} \left(\frac{\sum t.y}{m} - \frac{n+1}{2m} \right) \sum y_j$$

$$a = \frac{\sum y_j}{m.n} - b \frac{m.n+1}{2}$$

Trong đó:

n: Số năm

m: Khoảng cách thời gian trong một năm (m= 4 đối với quý, m=12 đối với năm)

t: Thứ tự thời gian trong dãy số (năm)

* Tính các mức độ mang tính thời vụ theo công thức sau:

$$tv = \frac{\bar{y}_i}{\bar{y}_j} - b\left(i - \frac{m+1}{2}\right) \text{ với } i=1,2,3,4\dots$$

Sau khi xác định xong hàm xu thế và biến động thời vụ thì mô hình dự báo kết hợp cộng giữa xu thế biến động và tính thời vụ có dạng: $Y_t = \hat{Y} + tv$

2.3.2. Dự báo dựa vào mô hình nhân

Mô hình dự báo theo kết hợp nhân có dạng:

$$Y_t = \hat{Y} \times tv$$

Để dự báo theo mô hình này, trước hết phải tính được hàm xu thế, hàm xu thế trong trường hợp này phải được loại trừ biến động thời vụ bằng cách xây dựng dãy số bình quân trượt (\bar{y}_t) với số lượng mức độ bằng 4 với tài liệu quý và 12 với tài liệu tháng.

Từ đó ta tính được $\frac{y_t}{\bar{y}_t}$, từ đó xác định thành phần thời vụ (tv_t) bằng cách tính

các số bình quân \bar{tv}_t sau đó tính hệ số điều chỉnh H:

$$H = \frac{m}{\sum tv_t} \text{ (với } m=4 \text{ đối với tài liệu quý, } 12 \text{ đối với tài liệu tháng)}$$

Từ đó tính chỉ số thời vụ $I_{tv} = \bar{tv}_t \times H$

Sau khi xác định được tv_t thì xác định dãy số f_t là dãy số đã loại bỏ thành phần thời vụ như sau: $f_t = \frac{y_t}{tv_t}$

2.4. Dự báo theo phương pháp san bằng mũ

Phương pháp san bằng mũ (hay còn gọi là phương pháp dự đoán bình quân mũ) là một phương pháp dự đoán thống kê ngắn hạn hiện được sử dụng nhiều trong công tác dự đoán thực tế trên thế giới.

Phương pháp san bằng mũ coi giá trị thông tin của mỗi mức độ là tăng dần kể từ đầu dãy số cho đến cuối dãy số. Vì trên thực tế ở những thời gian khác nhau thì hiện tượng nghiên cứu chịu sự tác động của những nhân tố khác nhau và cường độ không giống nhau. Các mức độ ngày càng mới (ở cuối dãy số thời gian) càng cần phải được chú ý đến nhiều hơn so với các mức độ cũ (ở đầu dãy số). Hay nói cách khác, mức độ càng xa so với thời điểm hiện tại thì càng ít giá trị thông tin, do đó càng ít ảnh hưởng đến mức độ dự đoán.

Tuỳ thuộc vào đặc điểm dãy số thời gian (chuỗi thời gian) có biến động xu thế, biến động thời vụ hay không mà phương pháp san bằng mũ có thể sử dụng một trong các phương pháp cơ bản sau:

2.4.1. Mô hình đơn giản (phương pháp san bằng mũ đơn giản)

Điều kiện áp dụng: đối với dãy số thời gian không có xu thế và không có biến động thời vụ rõ rệt.

Trước hết, dãy số thời gian được san bằng nhờ có sự tham gia của các số bình quân mũ, tức là các số bình quân di động gia quyền theo quy luật hàm số mũ. Theo phương pháp này, ở thời gian t nào đó dựa vào các giá trị thực tế đã biết để ước lượng giá trị hiện tại (thời gian t) của hiện tượng và giá trị hiện tại này để dự toán giá trị tương lai (thời gian $t+1$). Mô hình san bằng mũ giản đơn được Brown xây dựng năm 1954 dựa trên 2 nguyên tắc:

- Trọng số của các quan sát trong dãy số thời gian càng giảm đi khi nó càng cách xa hiện tại.

- Sai số dự báo hiện tại (ký hiệu $e_t = y_t - \hat{Y}_t$) Phải được tính đến trong những dự báo kế tiếp.

Giả sử ở thời gian t , có mức độ thực tế là y_t , mức độ dự đoán là \hat{Y}_t .

Mức độ dự đoán của hiện tượng ở thời gian $(t+1)$ có thể viết:

$$\hat{Y}_{t+1} = \alpha y_t + (1 - \alpha) \hat{Y}_t$$

$$\text{Đặt } 1 - \alpha = \lambda, \text{ ta có: } \hat{Y}_{t+1} = \alpha y_t + \lambda \hat{Y}_t \quad (*)$$

α và λ được gọi là các tham số san bằng với $\alpha + \lambda = 1$ và $\alpha, \lambda \in [0; 1]$.

Như vậy mức độ dự đoán \hat{Y}_{t+1} là trung bình cộng gia quyền của y_t và \hat{Y}_t với quyền số tương ứng là α và λ .

- Mức độ dự đoán của hiện tượng ở thời gian t là:

$$\hat{Y}_t = \alpha \hat{Y}_{t-1} + \lambda \hat{Y}_{t-1} \text{ thay vào } (*) \text{ ta có:}$$

$$\hat{Y}_{t+1} = \alpha y_t + \lambda \alpha \hat{Y}_{t-1} + \lambda^2 \hat{Y}_{t-1}$$

- Mức độ dự đoán của hiện tượng ở thời gian $(t-1)$ là:

$$\hat{Y}_{t-1} = \alpha \hat{Y}_{t-2} + \lambda \hat{Y}_{t-2} \text{ thay vào } (*)$$

$$\text{Ta có: } \hat{Y}_{t+1} = \alpha y_t + \lambda \alpha \hat{Y}_{t-1} + \lambda^2 \alpha \hat{Y}_{t-2} + \lambda^3 \hat{Y}_{t-2}$$

- Mức độ dự đoán của hiện tượng ở thời gian $(t-2)$ là: $\hat{Y}_{t-2} = \alpha \hat{Y}_{t-3} + \lambda \hat{Y}_{t-3}$ thay vào ta có:

$$\hat{Y}_{t+1} = \alpha y_t + \lambda \alpha \hat{Y}_{t-1} + \lambda^2 \alpha \hat{Y}_{t-2} + \lambda^3 \hat{Y}_{t-3} + \lambda^4 \hat{Y}_{t-4}$$

Bằng cách tiếp tục tương tự thay vào các mức độ dự đoán $\hat{Y}_{t-3}, \hat{Y}_{t-4}$ ta sẽ có công thức tổng quát.

$$\hat{Y}_{t+1} = \alpha \sum_{i=1}^n \lambda^i y_{t-i} + \lambda^{i+1} \hat{Y}_{t-i}$$

Trong đó:

\hat{Y}_{t+1} : Số bình quân mũ tại thời điểm $t+1$

y_{t-i} : Các mức độ thực tế của của hiện tượng tại thời điểm $(t-i)$ ($i=0 \rightarrow n$)

\hat{Y}_{t-1} : Số bình quân mũ tại thời điểm $(t-i)$ ($i=0 \rightarrow n$)

α và λ được gọi là các tham số san bằng

(α và λ là hằng số với $\alpha + \lambda = 1$ và $\alpha, \lambda \in [0; 1]$.)

n : Số lượng các mức độ của dãy số thời gian

Vì $\lambda \in 0;1$ nên khi $i \rightarrow \infty$

$$\text{Thì } \begin{cases} \lambda^{i+1} \rightarrow 0 \Rightarrow \lambda^{i+1}Y_{t-i} \rightarrow 0 \\ \alpha \sum_{i=1}^{\infty} \lambda^i \rightarrow 1 \end{cases}$$

Khi đó công thức tổng quát trở thành:

$$\hat{Y}_{t+1} = \alpha \sum_{i=1}^n \lambda^i y_{t-i}$$

Như vậy: mức độ dự đoán \hat{Y}_{t+1} là trung bình cộng gia quyền của các mức độ của dãy số thời gian mà trong đó quyền số giảm dần theo dạng mũ (khi $i=0 \rightarrow n$) tùy thuộc vào mức độ cũ của dãy số. Vì thế, phương pháp này được gọi là phương pháp san bằng mũ.

Có 2 vấn đề quan trọng nhất trong phương pháp san bằng mũ:

- Thứ nhất: hệ số san bằng mũ α

α là hệ số san để điều chỉnh trọng số của các quan sát riêng biệt của dãy số thời gian. Vì vậy, khi lựa chọn α phải vừa đảm bảo kết quả dự báo sẽ gần với quan sát thực tế, vừa phải đảm bảo tính linh hoạt (nhạy với các thay đổi ở gần hiện tại).

Với $\alpha=1$ thì theo phương trình dự báo (1). Giá trị dự báo \hat{Y}_{t+1} bằng giá trị thực tế ở thời kỳ ngay liền trước (Y_{t+1}) và các mức độ trước đó không được tính đến.

Với $\alpha=0$ theo phương trình dự báo (1). Giá trị dự báo \hat{Y}_{t+1} bằng giá trị dự báo ở thời kỳ trước (\hat{Y}_t) và giá trị thực tế ở thời kỳ ngay liền trước không được tính đến.

Nếu α được chọn càng lớn thì các mức độ càng mới sẽ càng được chú ý, thích hợp với chuỗi thời gian không có tính ổn định cao.

Ngược lại, nếu α được chọn càng nhỏ thì các mức độ càng cũ sẽ càng được chú ý, thích hợp với chuỗi thời gian có tính ổn định cao.

Do đó, phải dựa vào đặc điểm biến động của hiện tượng qua thời gian và kinh nghiệm nghiên cứu để lựa chọn α cho phù hợp. Nói chung, giá trị α tốt nhất là giá trị làm cho tổng bình phương sai số dự đoán nhỏ nhất.

$$SSE = \sum (y_t - \hat{y}_t) \min$$

Đặt $e_t = y_t - \hat{Y}_t$ là các sai số dự đoán ở thời gian t hay còn gọi là phần dư ở thời gian t .

Theo kinh nghiệm của các nhà dự báo thì α thích hợp cho vận phương pháp san mũ có thể được chọn bằng.

$$\alpha = \frac{2}{n+1} n \quad : \text{độ dài chuỗi thời gian}$$

- *Thứ hai:* Xác định giá trị ban đầu (điều kiện ban đầu) ký hiệu y_0

Phương pháp san bằng mũ được thực hiện theo phép đệ quy, để tính \hat{Y}_{t+1} thì phải có \hat{Y}_t , để có \hat{Y}_t thì phải có \hat{Y}_{t-1} . Do đó để tính toán cần phải phải xác định giá trị ban đầu (y_0) dựa vào một số phương pháp.

+ Có thể lấy mức độ đầu tiên của dãy số.

+ Trung bình của một số các mức độ của dãy số

2.4.2. Mô hình xu thế tuyến tính và không có biến động thời vụ (Mô hình san mũ Holt – Winters)

Mô hình này thường áp dụng đối với sự biến động của hiện tượng qua thời gian có xu thế là tuyến tính và không có biến động thời vụ.

- Giả sử chúng ta có dãy số thời gian $y_1, y_2, y_3, \dots, y_n$ với biến động có tính xu thế.

Bước 1: Chọn các hệ số α, β ($0 < \alpha, \beta < 1$)

Nếu chọn hằng số san nhỏ tức là chúng ta coi các mức độ hiện thời của dãy số ít ảnh hưởng đến mức độ dự báo. Ngược lại nếu chọn hằng số san lớn tức là chúng ta muốn dãy số san số mũ phản ứng mạnh với những thay đổi hiện tại.

Bước 2: Tiến hành san mũ cho giá trị ước lượng và xu thế của dãy số:

Coi giá trị của dãy số thời gian là tổng của 2 thành phần: Thành phần trung bình có trọng số của các giá trị thực tế (ký hiệu là S_t – giá trị ước lượng của hiện tượng ở thời điểm t) và thành phần xu thế (ký hiệu là T_t). Ta có mô hình san số mũ:

$$\hat{y}_{t+1} = S_t + T_t$$

Trong đó:

$$S_t = \alpha y_t + (1 - \alpha) [S_{t-1} + T_{(t-1)}] = \alpha y_t + (1 - \alpha) S_t$$

$$T_t = \beta(S_t - S_{t-1}) + (1 - \beta) T_{(t-1)}$$

$$\text{Đặt } S_2 = Y_2 \quad T_2 = Y_2 - Y_1$$

Tiến hành san số mũ từ thời điểm thứ 3 trở đi, ta có:

$$S_3 = \alpha Y_3 + (1 - \alpha)(S_2 + T_2)$$

$$T_3 = \beta(S_3 - S_2) + (1 - \beta)T_2$$

$$S_4 = \alpha Y_4 + (1 - \alpha)(S_3 + T_3)$$

$$T_4 = \beta(S_4 - S_3) + (1 - \beta)T_3$$

...

Bước 3: Sử dụng mức và xu thế đã được san số mũ tại thời điểm để dự đoán cho các thời điểm trong tương lai để dự đoán giá trị của hiện tượng ở thời điểm tương lai $t + 1$:

$$\hat{y}_{t+1} = S_t + T_t$$

Ở thời điểm tương lai $(t + h)$ ($h=2, 3 \dots$)

$$\hat{y}_{t+h} = S_t + hT_t$$

2.4.3. Mô hình xu thế tuyến tính và biến động thời vụ

Mô hình này thường áp dụng đối với dự báo thời gian mà các mức độ của nó là tài liệu tháng hoặc quý của một số năm mà các mức độ trong dãy số được lập lại sau 1 khoảng thời gian h ($h = 4$ đối với quý, $h = 12$ đối với năm). Việc dự đoán có thể được thực hiện theo một trong hai mô hình sau:

$$+ \text{Mô hình cộng} \quad \hat{y}_{t+1} = S_t + T_t + V_{t+1}$$

Trong đó:

$$S_t = \alpha y_t - V(t-h) + (1 - \alpha) [S_{t-1} + T_{(t-1)}]$$

$$T_t = \beta(S_t - S_{t-1}) + (1 - \beta)T_{(t-1)}$$

$$V_t = \lambda(y_t - S_t) + (1 - \lambda)V_{(t-h)}$$

+ Mô hình nhân: $\hat{y}_{t+1} = (S_t + T_t).V_{t+1}$

Trong đó

$$S_t = \alpha \frac{y_t}{V(t-h)} + (1-\alpha)(S_{t-1} + T_{(t-1)})$$

$$T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1}$$

$$V_t = \lambda \frac{y_t}{S_t} + (1-\lambda).V_{(t-h)}$$

Với α, β, λ là các tham số san bằng nhận giá trị trong đoạn $[0;1]$.

α, β, λ nhận giá trị tốt nhất khi tổng bình phương sai số là nhỏ nhất.

$$SSE = \sum (y_t - \hat{y}_t)^2 \Rightarrow \min$$

Tham số α, β, λ không được xét một cách khách quan mà ít nhiều thông qua trực giác chủ quan, kết quả dự báo sẽ phụ thuộc vào sự lựa chọn các tham số này.

Với $a_0(0)$ có thể là mức độ đầu tiên trong dãy số.

$a_1(0)$ có thể là lượng tăng (giảm) tuyệt đối trung bình.

$S_j(0)$: Là các chỉ số thời vụ ban đầu ($j=1,2,3,\dots,k$); $k = 4$ đối với quý; $k = 12$ đối với tháng.

Nếu $t = 1, 2, 3, 4, 5, \dots, n$.

Là thứ tự thời gian hay tương ứng với thứ tự các mức độ theo thời ký trong chuỗi thời gian thì yếu tố thời vụ $V_j(0)$ của các mức độ trong chuỗi thời gian được tính sẽ tương ứng với các giá trị $t \leq k$.

$$V_j(0) = \hat{V}_j x H; \quad \hat{V}_j = \sum_{j=1}^k \frac{V_j}{k}$$

\bar{V}_j chỉ số bình quân thời vụ cho một quý hay một tháng của mỗi năm trong chuỗi thời gian.

$$V_j = \frac{y_t}{y_t}$$

Y_t mức độ trong chuỗi thời gian ở thời gian t .

V_j chỉ số thời vụ của từng quý hoặc tháng trong từng năm nay ở thời gian t :

\bar{y}_t Số bình quân trượt để loại trừ thành phần thời vụ và thành phần ngẫu nhiên với số lượng mức độ bằng 4 đối với tài liệu quý và bằng 12 đối với tài liệu tháng.

$$H = \frac{k}{\sum \bar{V}_j}$$

Nếu phải lựa chọn một trong hai mô hình để dự đoán thì tùy thuộc vào đặc điểm biến động của hiện tượng:

Đối với hiện tượng ít biến đổi qua thời gian thì dùng mô hình cộng.

Đối với hiện tượng biến đổi nhiều qua thời gian thì dùng mô hình nhân.

* *Ưu, nhược điểm của phương pháp san bằng mũ:*

- Ưu điểm:

Đơn giản và có kết quả tương đối chính xác phù hợp với dự đoán ngắn hạn cho các nhà kinh doanh cũng như lập kế hoạch ngắn hạn ở cấp vĩ mô.

Hệ thống dự báo có thể được điều chỉnh thông qua 1 tham số duy nhất (tham số san bằng mũ)

Dễ dàng chương trình hoá vì chỉ phải thực hiện một số phép toán sơ cấp để xác định giá trị dự báo.

- Hạn chế:

Phương pháp san bằng mũ chỉ bó hẹp trong phạm vi dự báo ngắn hạn vì không tính đến sự thay đổi cấu trúc của chuỗi thời gian mà phải tuân thủ tính ổn định theo thời gian của các quý trình kinh tế - xác hội.

CHƯƠNG 3: THỰC NGHIỆM CÁC BÀI TOÁN DỰ BÁO

3.1. Phần mềm IBM SPSS Modeler

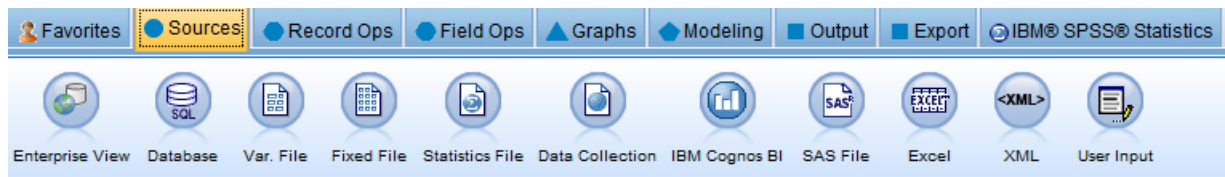
3.1.1. Giới thiệu

Hiện tại trên thế giới có rất nhiều phần mềm được áp dụng vào giải quyết các bài toán dự báo như: Foccat X, Eviews, SPSS... và SPSS Modeler của hãng IBM là 1 trong số đó. SPSS Modeler là một bộ công cụ khai thác dữ liệu cho phép người dùng nhanh chóng xây dựng các mô hình dự báo sử dụng chuyên môn kinh doanh và triển khai chúng vào hoạt động thương mại để cải thiện việc ra quyết định. SPSS Modeler hỗ trợ toàn bộ quá trình khai thác dữ liệu, từ dữ liệu tới những kết quả kinh doanh tốt hơn. SPSS Modeler cung cấp đa dạng các phương pháp mô hình hoá lấy từ học máy, trí tuệ nhân tạo và thống kê. Cho phép người dùng chuyển hóa dữ liệu thành những thông tin mới và phát triển các mô hình dự báo.

3.1.2. Các chức năng trong SPSS Modeler

1. Lấy nguồn dữ liệu:

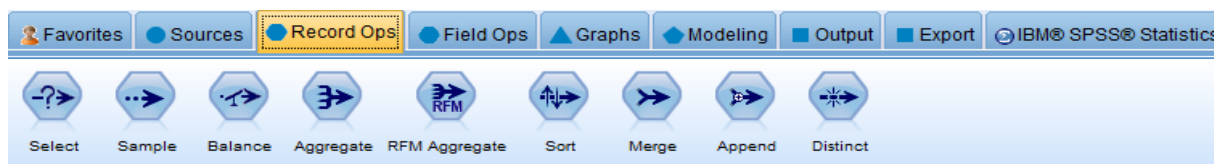
Nguồn dữ liệu bao gồm tập dữ liệu với rất nhiều các định dạng giúp người sử dụng dễ dàng đưa dữ liệu của mình vào để xử lý như là ; Exel, SPSS, SQL,...



Hình 3.1.2. 1 Nguồn dữ liệu

Nguồn dữ liệu hay còn gọi là dữ liệu thô nghĩa là dữ liệu chưa qua quá trình tinh chỉnh, là nguồn dữ liệu gốc, nguồn dữ liệu ban đầu.

2. Trích chọn dữ liệu:

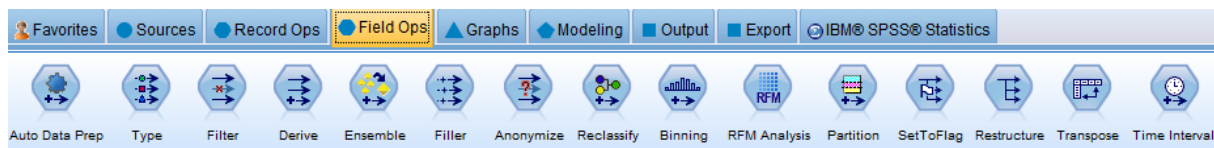


Hình 3.1.2. 2 Trích chọn dữ liệu

Chọn nguồn dữ liệu phù hợp nhất với yêu cầu bài toán đặt ra. Dữ liệu được chọn phải chứa những thông tin đầy đủ liên quan đến yêu cầu cần đặt ra, phải thỏa mãn các tiêu chí nhất định nào đó.

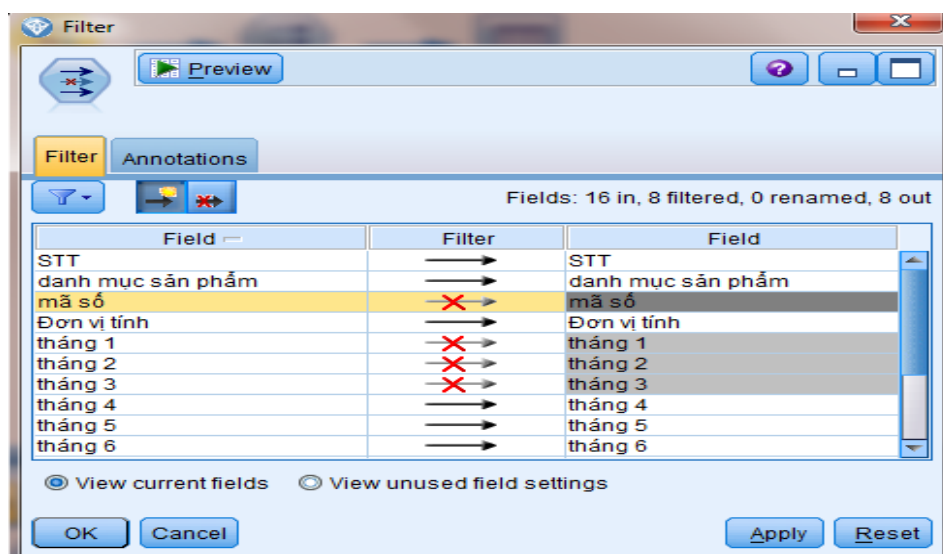
3. Biến đổi dữ liệu:

Biến đổi dữ liệu là quá trình chuẩn hóa và làm mịn dữ liệu để đưa dữ liệu về dạng ngắn gọn và đơn giản giúp giải quyết bài toán một cách nhanh nhất.



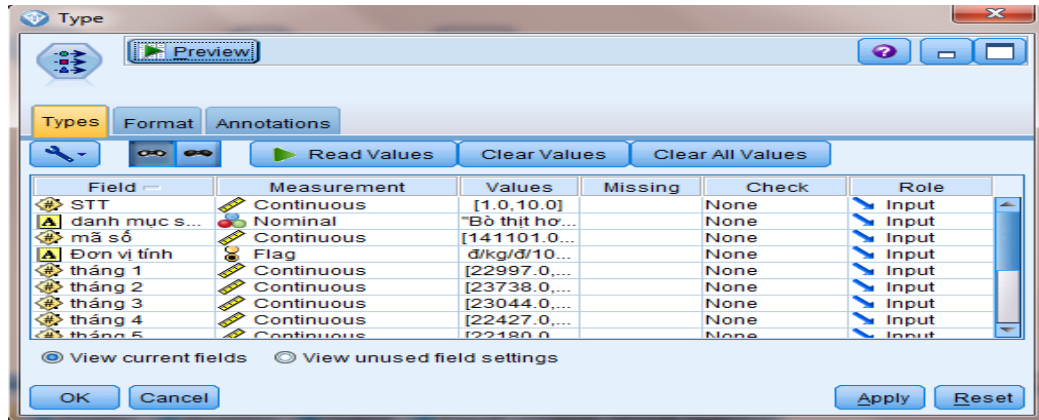
Hình 3.1.2. 3 Biến đổi dữ liệu

Biến đổi dữ liệu ban đầu thành các dữ liệu chuẩn nhất, có thể thêm các trường dữ liệu cần thiết hoặc bỏ đi các trường dữ liệu không cần thiết.



Hình 3.1.2. 4 Lọc các trường dữ liệu

Biến đổi dữ liệu thành các dữ liệu mới với những thuộc tính mới và các trường dữ liệu mới.

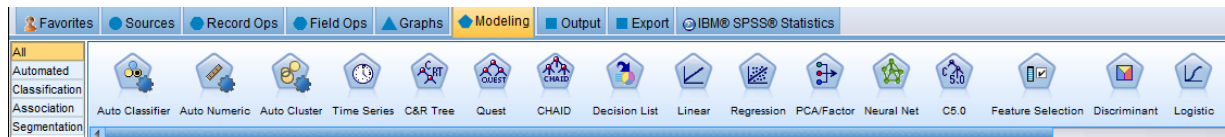


Hình 3.1.2. 5 Biến đổi trường và thuộc tính dữ liệu

Có thể biến đổi thành các loại dữ liệu sau: Range (khoảng cách, hàng), Default (mặc định), Flag (dạng cờ), Set (tập hợp), Ordered Set, Typeless, Discrete ...

4. Xử lý dữ liệu.

Đây được xem là bước quan trọng và tốn nhiều thời gian nhất của toàn quá trình.

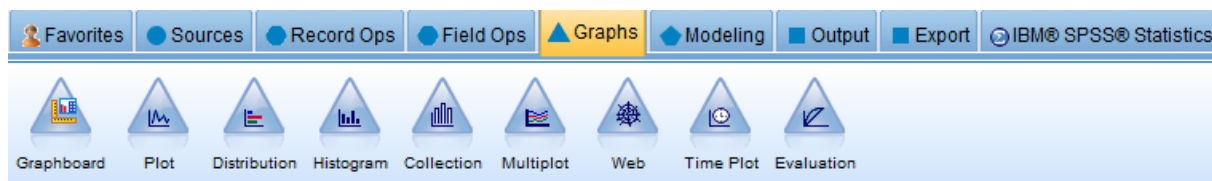


Hình 3.1.2. 6 Xử lý dữ liệu

Trong quá trình này sử dụng các thuật toán phân hoạch, các thuật toán phân cụm phân cấp để xử lý dữ liệu... như thuật toán K-means, PAM, CLARA, BIRCH,....

5. Đánh giá và biểu diễn kết quả.

Đây là kết quả của toàn bộ quá trình. Kết quả được thể hiện dưới các dạng khác nhau như bảng biểu (Excel, Table,...) hay dưới dạng đồ thị (Graphboard, Plot, Distribution, histogram, collection, multiplot, Web, Timelot, Evaluation,...) giúp đưa ra kết quả gần gũi với người sử dụng, có cái nhìn trực quan hơn đối với yêu cầu được đặt ra.



Hình 3.1.2. 7 Đánh giá và biểu diễn kết quả

Sau khi kết quả được đưa ra thì đánh giá kết quả đó có đúng yêu cầu của bài toán không, có thỏa mãn tiêu chí hay điều kiện nhất định nào đó hay không.

3.2. Áp dụng phần mềm IBM SPSS Modeler vào bài toán dự báo

3.2.1. Bài toán 1 (sử dụng phương pháp định tính)

3.2.1.1. Phát biểu bài toán:

Với số liệu có được từ Ban chỉ đạo tổng điều tra dân số và nhà ở trung ương, Tổng cục thống kê – Cục thống kê thành phố Hải Phòng, 11/2008. Hãy dự báo nhu cầu máy vi tính của thành phố Hải Phòng sau năm 2008.

3.2.1.2. Phương pháp thực hiện:

a. Lựa chọn phương pháp chuyên gia: Giả định để có nhu cầu mua máy vi tính các chuyên gia nhận định phải có đủ các tiêu chí sau:

- Độ tuổi từ 15-55 tuổi.
- Phải đang đi học hoặc đang đi làm.
- Trong gia đình phải có ti vi+ xe máy+ ĐT cố định+ chưa có máy vi tính.

Trong tập **Ho.sav** chứa rất nhiều thông tin trong đó có dữ liệu về độ tuổi, đi học, làm việc.

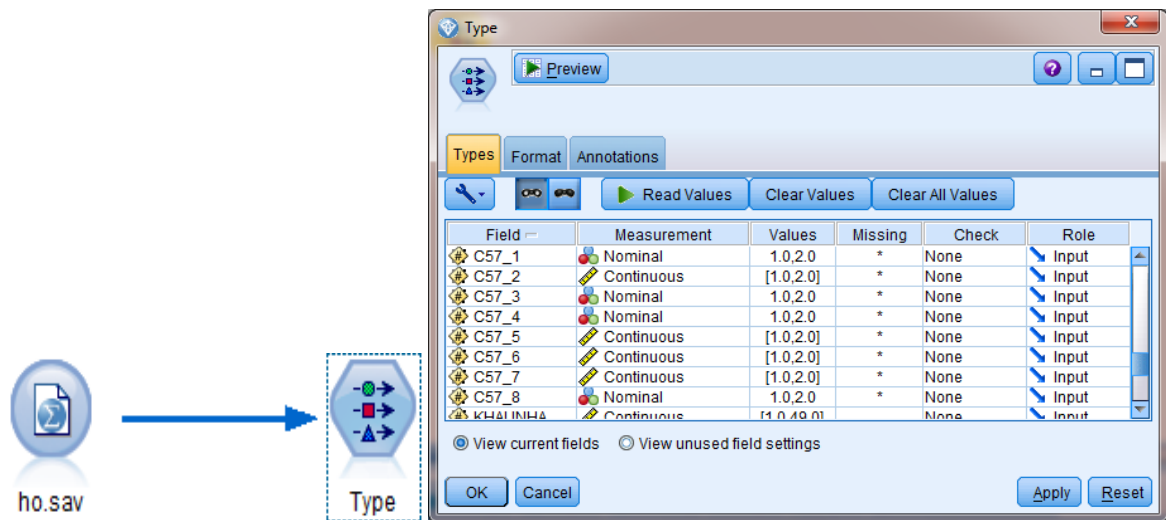
Trong tập **Ngnoi.sav** chứa rất nhiều thông tin trong đó có dữ liệu về tài sản của từng người như ti vi, xe máy, máy vi tính, điện thoại cố định.

Bảng 3.2. 1 Diễn giải

Ký hiệu	Diễn giải	Miền giá trị
C5	Độ tuổi	0-120
C12	Đi học	1.00 = “có” ; 2.00 = “không”
C20	Làm việc	1.00 = “có” ; 2.00 = “không”
C57_1	Ti vi	1.00 = “có” ; 2.00 = “không”
C57_3	Xe máy	1.00 = “có” ; 2.00 = “không”
C57_4	Máy vi tính	1.00 = “có” ; 2.00 = “không”
C57_8	ĐT cố định	1.00 = “có” ; 2.00 = “không”

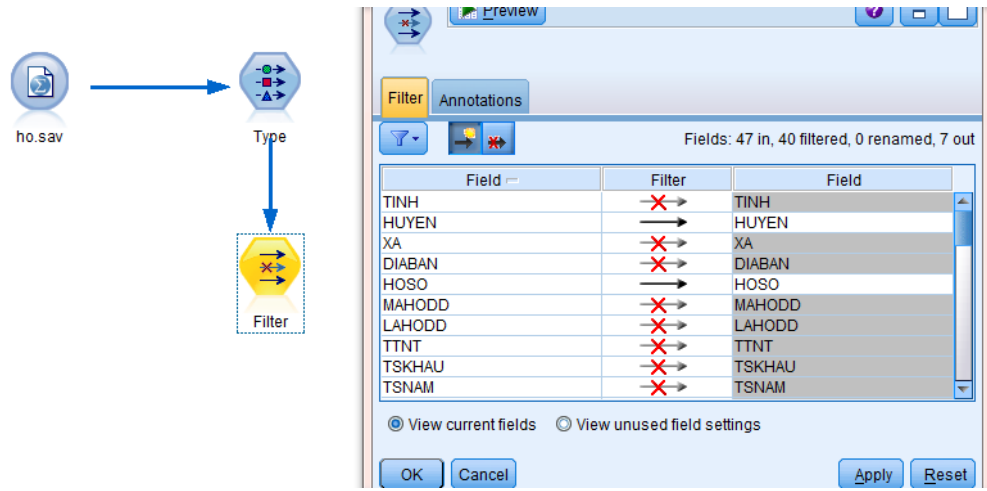
b. Sử dụng phần mềm IBM SPSS Modeler để dự báo và đưa ra kết quả dự báo:

Bước 1: Từ tập **Ho.sav**, chọn nút **Type**, thay đổi kiểu dữ liệu.



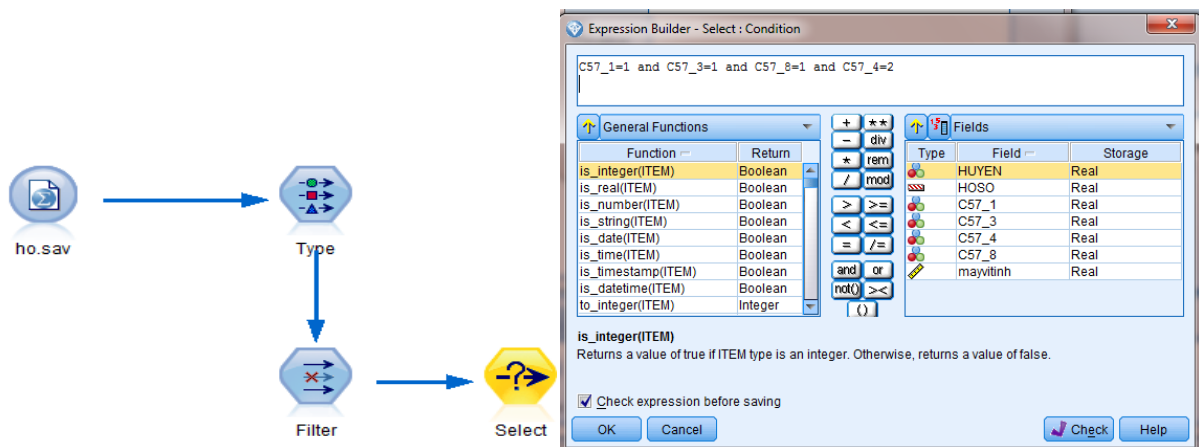
Hình 3.2.1. 1 Chọn nút Type

Bước 2: Chọn nút **Filter**, nhấn đúp vào nút **Filter**, loại bỏ một số trường dữ liệu không cần thiết.



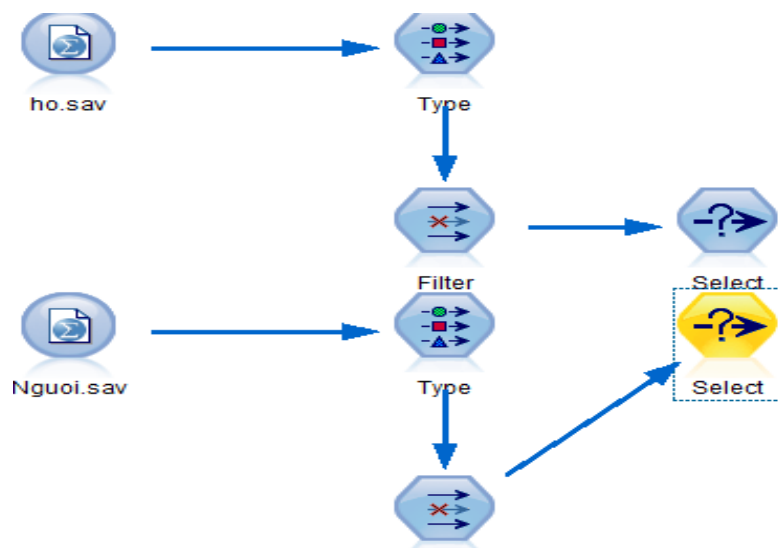
Hình 3.2.1. 2 Chọn nút Filter

Bước 3: Chọn nút **Select**, nhấn đúp chọn **Launch Expression Builder** (Tại đây chọn các trường và giá trị phù hợp theo yêu cầu bài toán)/ **Check/ OK/ OK**.



Hình 3.2.1. 3 Chọn nút select

Bước 4: Xử lý tương tự với tập *Nguoi.sav*



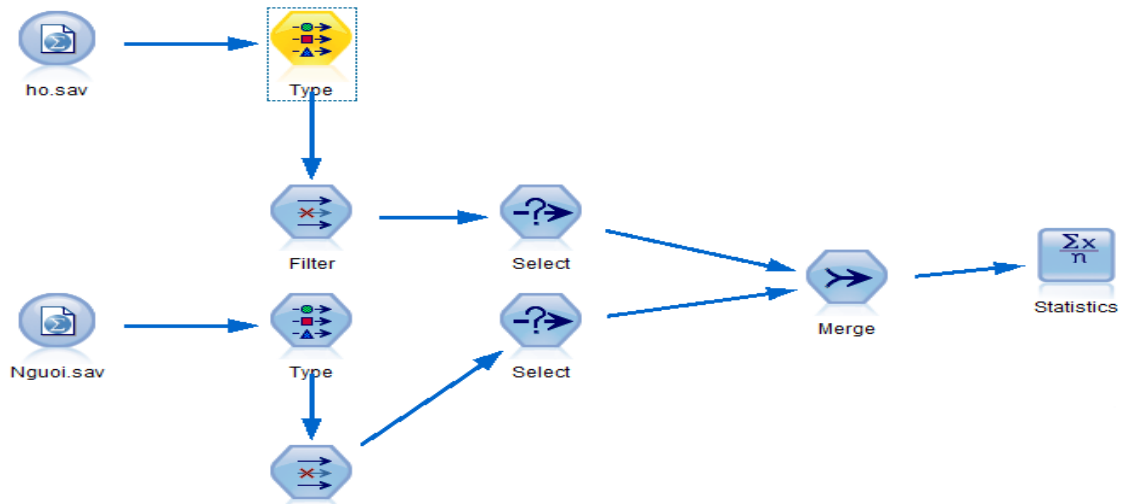
Hình 3.2.1. 4 Xử lý với tập *Nguoi.sav*

Bước 5: Chọn nút *Merge*, nhấn đúp chọn **Filter** (Tại đây tắt các trường trùng tên)/OK.

Field	Tag	Source ...	Connect...	Filter	Field
HUYEN	1	ho.sav	Select	→	HUYEN
HOSO	1	ho.sav	Select	→	HOSO
C57_1	1	ho.sav	Select	→	C57_1
C57_3	1	ho.sav	Select	→	C57_3
C57_4	1	ho.sav	Select	→	C57_4
C57_8	1	ho.sav	Select	→	C57_8
mayvtinh	1	ho.sav	Select	→	mayvtinh
HUYEN	2	Nguoi.sav	Select	→	HUYEN
HOSO	2	Nguoi.sav	Select	×	HOSO
C5	2	Nguoi.sav	Select	→	C5
C12	2	Nguoi.sav	Select	→	C12
C20	2	Nguoi.sav	Select	→	C20

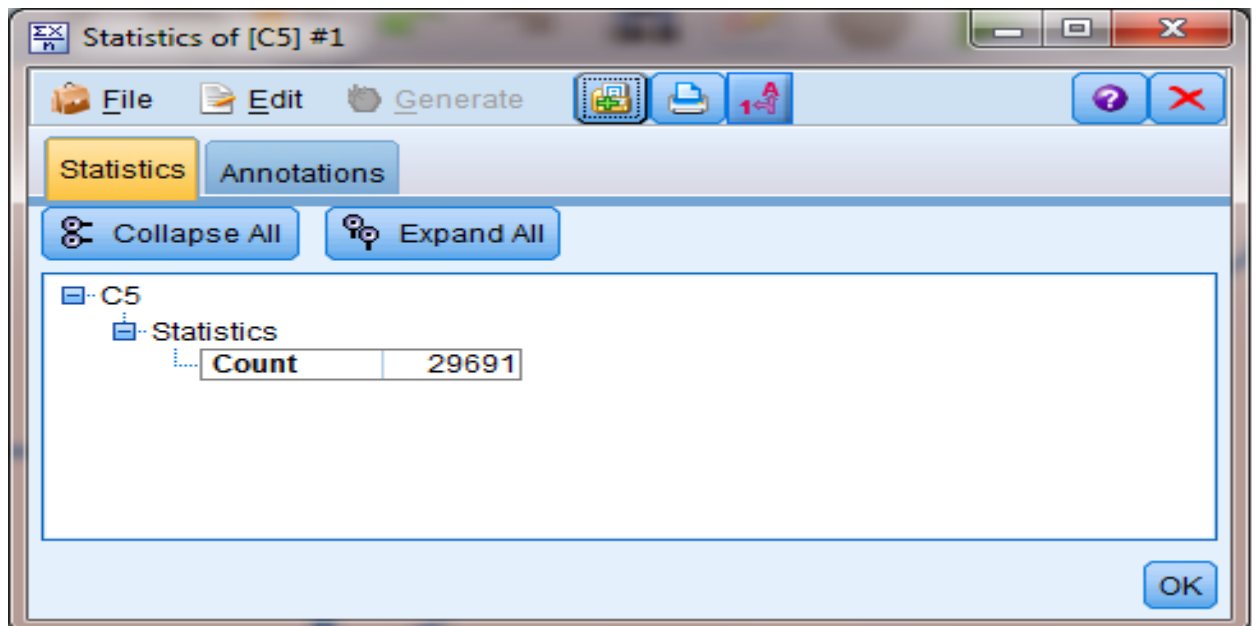
Hình 3.2.1. 5 Chọn nút *Merge*

Bước 6: Chọn nút **Statistics**, nhấn đúp chọn **Setting/Examine** (chọn trường C5) và **Statistics** (tích chọn ô **Count**)/ **Run**.



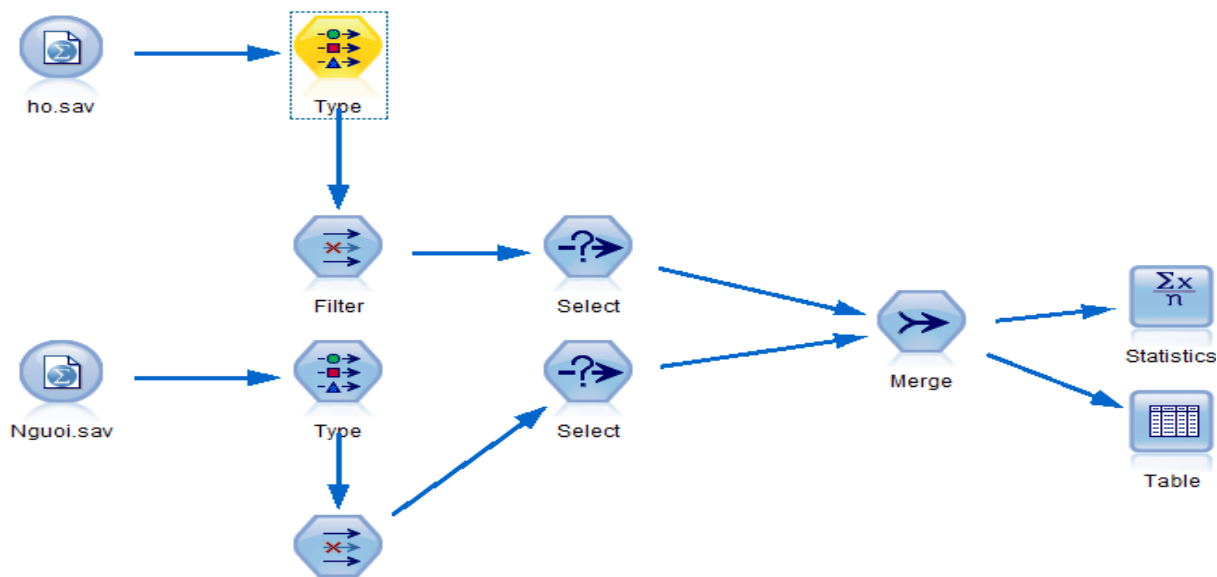
Hình 3.2.1. 6 Chọn nút Statistic

Bước 7: Tại dòng **Count** thể hiện nhu cầu số lượng về máy vi tính cần tìm.



Hình 3.2.1. 7 Kết quả số lượng

Bước 8: Chọn nút **Table**, nhấn đúp và chọn **Run**.



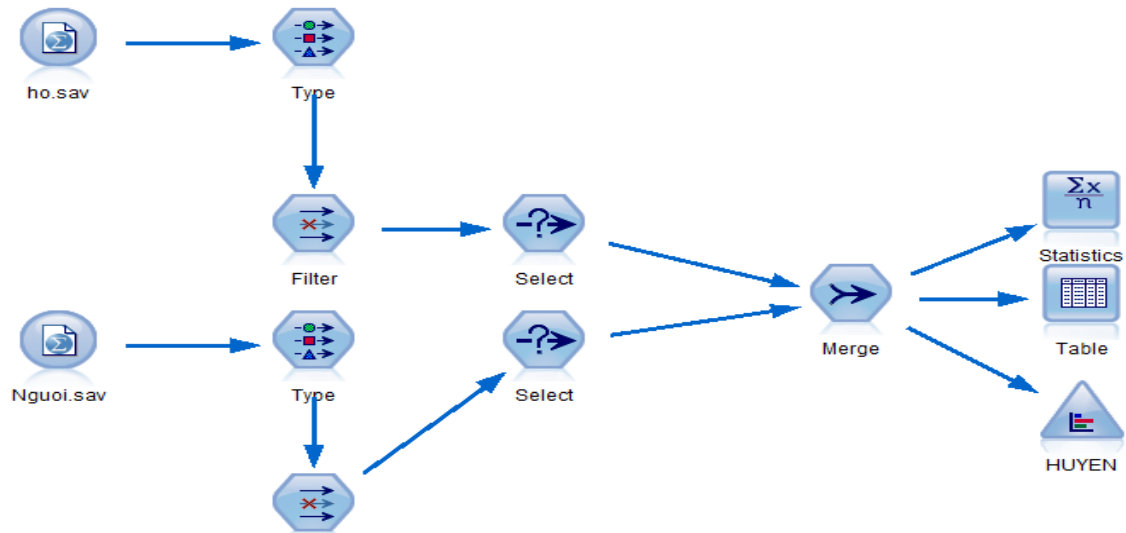
Hình 3.2.1. 8 Chọn nút Table

Qua bảng cho thấy kết quả đưa ra đúng với các tiêu chí lựa chọn của các chuyên gia.

	HUYEN	HOSO	C57_1	C57_3	C57_4	C57_8	C5	C12	C20
1	303	1	1	1	2	1	54	2	1
2	303	4	1	1	2	1	52	2	1
3	303	7	1	1	2	1	26	2	1
4	303	17	1	1	2	1	19	2	1
5	303	27	1	1	2	1	30	1	1
6	303	28	1	1	2	1	28	2	1
7	303	29	1	1	2	1	39	2	1
8	303	30	1	1	2	1	27	2	1
9	303	33	1	1	2	1	23	2	1
10	303	46	1	1	2	1	50	2	1
11	303	47	1	1	2	1	34	2	1
12	303	48	1	1	2	1	20	2	1
13	303	50	1	1	2	1	23	2	1
14	303	53	1	1	2	1	31	2	1
15	303	54	1	1	2	1	23	1	2
16	303	62	1	1	2	1	36	2	1
17	303	73	1	1	2	1	31	2	1
18	303	74	1	1	2	1	30	2	1
19	303	89	1	1	2	1	51	2	1
20	303	90	1	1	2	1	46	2	1

Hình 3.2.1. 9 Bảng kết quả

Bước 9: Chọn nút *Distribution*, nhấn đúp chọn **Plot** (chọn trường cho **Field** và **Color**)/ **Run**.



Hình 3.2.1. 10 Chọn nút *Distribution*

Tại tab **Table** thể hiện % và số lượng nhu cầu máy vi tính của từng huyện (quận).

Value	Proportion	%	Count
303		7.73	2294
304		7.41	2199
305		8.49	2520
306		7.66	2274
307		7.66	2275
308		6.33	1879
309		5.25	1560
311		8.78	2607
312		9.26	2749
313		7.43	2205
314		7.62	2261
315		5.99	1779
316		6.13	1821
317		4.27	1268

Hình 3.2.1. 11 Bảng kết quả

3.2.2. Bài toán 2 (sử dụng phương pháp định lượng)**3.2.2.1. Phát biểu bài toán:**

Cho bảng số liệu sau. Dự báo giá bán tháng 10 của các mặt hàng.

Bảng 3.2. 2 Giá bán sản phẩm

STT	Sản phẩm	Mã số	ĐVT	Tháng 1	Tháng 2	Tháng 3	Tháng 4	Tháng 5	Tháng 6	Tháng 7	Tháng 8	Tháng 9
1	Trâu thịt hơi	141101	đ/kg	64,749	64,965	65,000	65,000	64,119	64,245	64,121	64,624	65,737
2	Bò thịt hơi	141201	đ/kg	72,875	73,099	71,728	71,721	73,375	72,498	72,371	7,265	72,969
3	Lợn thịt hơi	145001	đ/kg	54,247	49,473	46,289	46,947	45,492	41,352	42,270	41,871	39,975
4	Lợn con giống	145002	đ/kg	44,994	54,542	51,488	54,496	51,710	48,150	46,524	46,622	47,133
5	Gà ta thịt hơi	146201	đ/kg	87,989	90,000	88,983	94,977	106,099	96,604	99,361	103,489	99,434
6	Gà CN thịt hơi	146202	đ/kg	50,743	53,747	50,925	46,706	45,252	46,483	41,170	42,249	50,915
7	Gà giống	146203	đ/kg	22,997	23,738	23,044	22,427	22,180	22,714	22,371	22,518	20,798
8	Trứng gà	146204	đ/10 quả	33,742	34,623	32,945	32,193	32,700	32,125	32,297	32,100	31,976
9	Vịt thịt hơi	146301	đ/kg	42,857	44,230	41,418	36,683	43,499	40,152	35,256	35,274	35,490
10	Ngan thịt hơi	146302	đ/kg	61,740	65,249	63,712	61,106	68,480	62,208	60,510	58,334	57,446

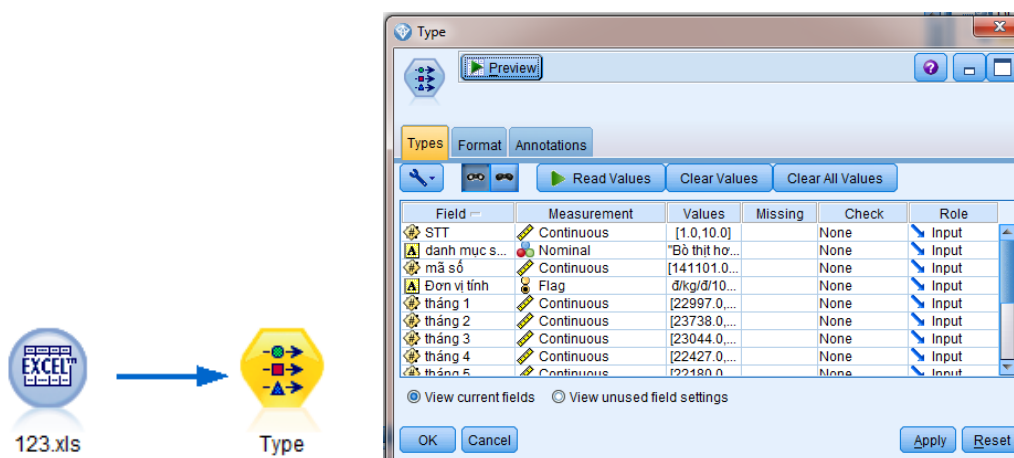
3.2.2.2. Phương pháp thực hiện:

a. Lựa chọn phương pháp dự báo: bình quân di động có quyền số.

Giả sử rằng ta có quyền số của tháng gần nhất là 3; cách 2 tháng trước là 2,5; cách 3 tháng trước là 2; 4 tháng trước là 1,5; 5 tháng trước là 1.

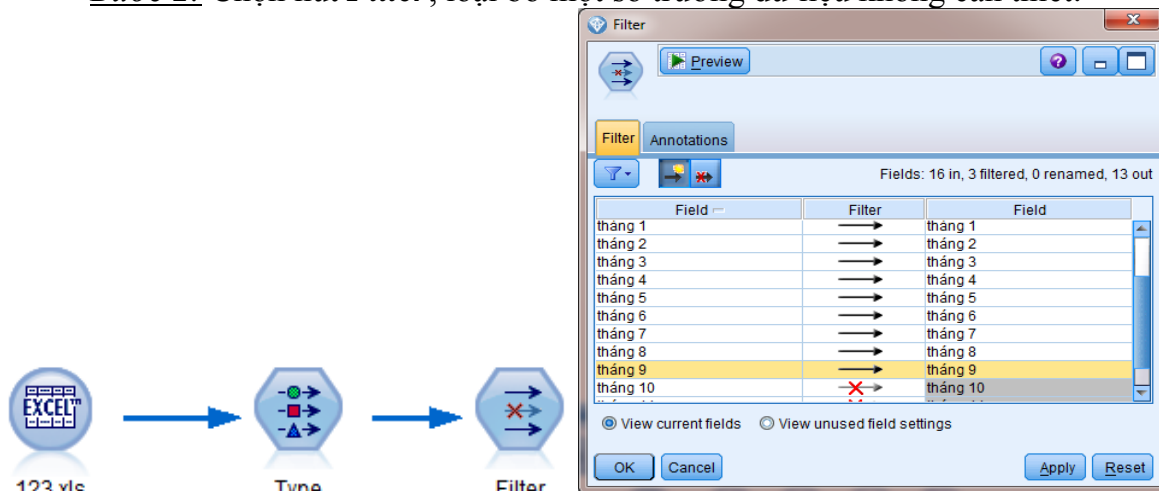
b. Sử dụng phần mềm IBM SPSS Modeler để dự báo và đưa ra kết quả dự báo:

Bước 1: Từ File excel **123.xls**, chọn nút **Type**, thay đổi kiểu dữ liệu.



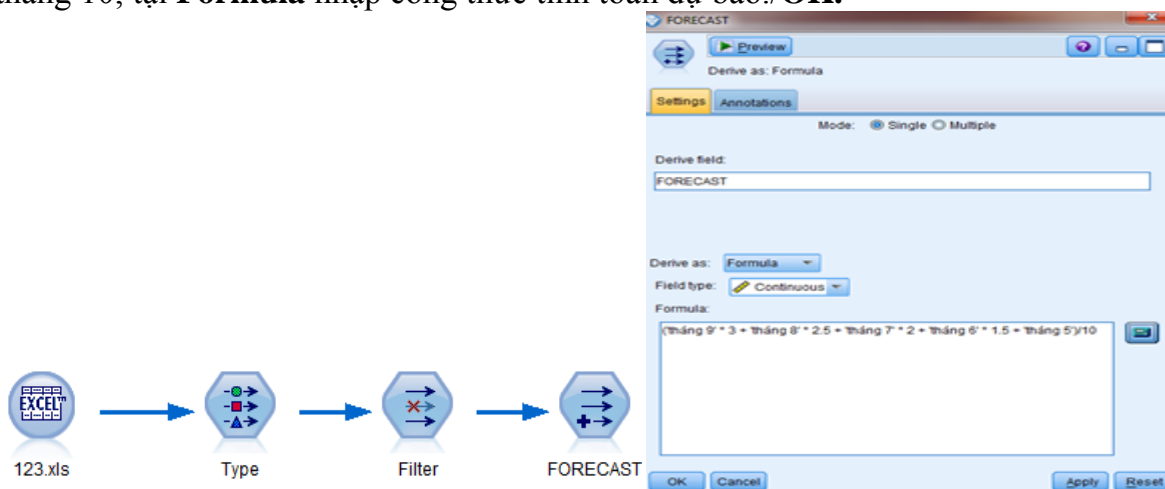
Hình 3.2.2. 1 Chọn nút Type

Bước 2: Chọn nút **Filter**, loại bỏ một số trường dữ liệu không cần thiết.



Hình 3.2.2. 2 Chọn nút Filter

Bước 3: Chọn nút *Derive*, nhấn đúp chọn tab **Setting**, tại **Derive Field** nhập tên tháng 10 sẽ dự báo; tại **Derive as** chọn *Formula*; tại **Field Type** chọn trường cho tháng 10; tại **Formula** nhập công thức tính toán dự báo./OK.



Hình 3.2.2. 3 Chọn nút *Derive*

Bước 4: Chọn nút *Table*, nhấn đúp và chọn **Run**.



Hình 3.2.2. 4 Chọn nút *Table*

Qua bảng cho thấy kết quả dự báo giá bán tháng 10 của các mặt hàng.

	STT	danh mục sản phẩm	mã số	Đơn vị tính	tháng 1	tháng 2	tháng 3	tháng 4	tháng 5	tháng 6	tháng 7	tháng 8	tháng 9	FORECA...
1	1	Trâu thịt hơi	141101	đ/kg	64749	64965	65000	65000	64119	64245	64121	64624	65737	64750
2	2	Bò thịt hơi	141201	đ/kg	72875	73099	71728	71721	73375	72498	72371	72265	72969	72643
3	3	Lợn thịt hơi	145001	đ/kg	54247	49473	46289	46947	45492	41352	42270	41871	39975	41666
4	4	Lợn con giống	145002	đ/kg	44994	54542	51488	54496	51710	48150	46524	46622	47133	47494
5	5	Gà ta thịt hơi	146201	đ/kg	87989	90000	88983	94977	106099	96604	99361	103489	99434	100675
6	6	Gà CN thịt hơi	146202	đ/kg	50743	53747	50925	46706	45252	46483	41170	42249	50915	45568
7	7	Gà giống	146203	đ/kg	22997	23738	23044	22427	22180	22714	22371	22518	20798	21968
8	8	Trứng gà	146204	đ/10quả	33742	34623	32945	32193	32700	32125	32297	32100	31976	32166
9	9	Vịt thịt hơi	146301	đ/kg	42857	44230	41418	36683	43499	40152	35256	35274	35490	36889
10	10	Ngan thịt hơi	146302	đ/kg	61740	65249	63712	61106	68480	62208	60510	58334	57446	60098

Hình 3.2.2. 5 Kết quả

3.2.3. Bài toán 3 (sử dụng phương pháp định lượng)**3.2.3.1. Phát biểu bài toán:**

Cho tài liệu về sản lượng của doanh nghiệp A. Dự báo sản lượng 4 quý tiếp theo của doanh nghiệp.

Bảng 3.2. 3 Sản lượng doanh nghiệp A

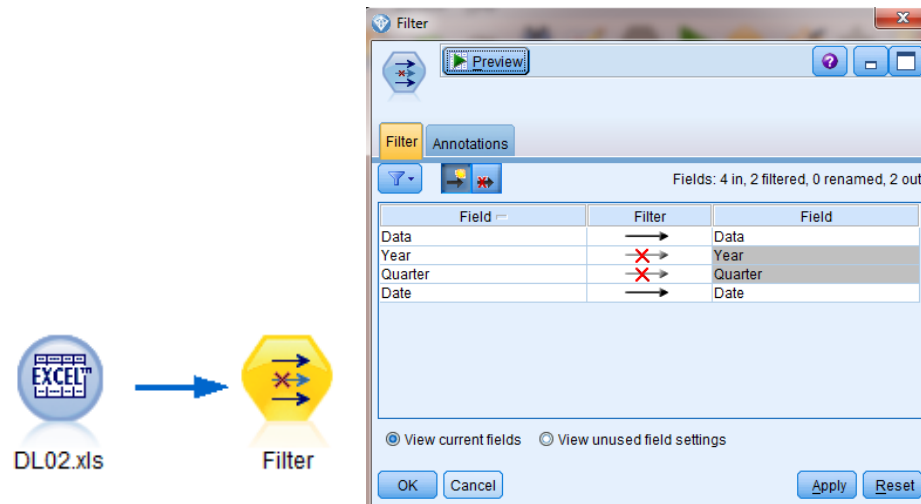
Năm (t) Quý	Sản lượng (nghìn tấn)					Cộng theo cùng quý $\sum y_j$	Mức độ bình quân từng quý \bar{y}_i	Chỉ số thời vụ $I_{tv} = \frac{\bar{y}_i}{\bar{y}}$
	2002	2003	2004	2005	2006			
I	20	25	27	31	29	132	26,4	0,678
II	25	32	30	37	36	160	32	0,82
II	38	38	45	44	47	212	42,4	1,14
IV	40	60	55	62	58	275	55	1,41
Cộng theo cùng năm $\sum y_j$	123	155	157	174	170	779	38,95	
$\sum t.y$	123	310	471	696	850			

3.2.3.2. Phương pháp thực hiện:

a. Lựa chọn phương pháp dự báo: dựa vào hàm xu thế và biến động thời vụ (dựa vào mô hình cộng).

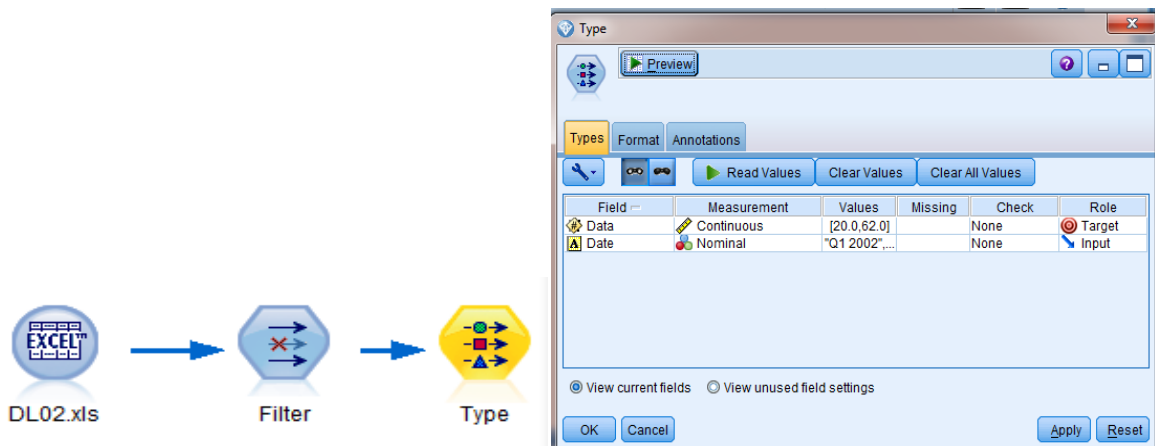
b. Sử dụng phần mềm IBM SPSS Modeler để dự báo và đưa ra kết quả dự báo:

Bước 1: Từ File **DL02.xls**, chọn nút **Filter**, loại bỏ một số trường dữ liệu không cần thiết.



Hình 3.2.3. 1 Chọn nút **Filter**

Bước 2: Chọn nút **Type**, nhấn đúp và thay đổi kiểu dữ liệu.



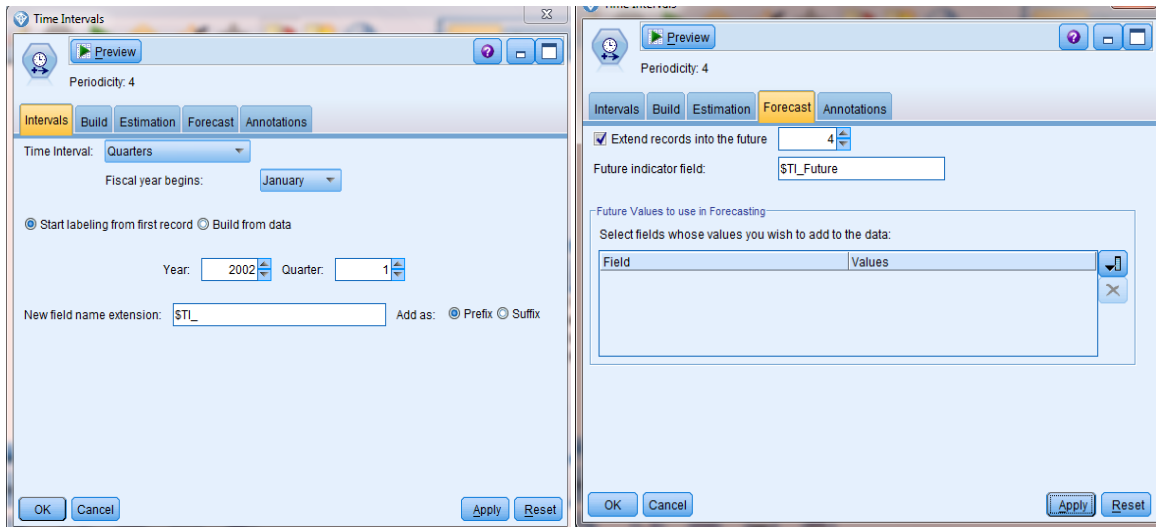
Hình 3.2.3. 2 Chọn nút **Type**

Bước 3: Chọn nút **Time Intervals**.



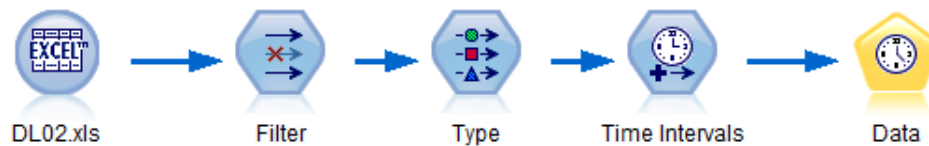
Hình 3.2.3. 3 Chọn nút **Time Intervals**

Chọn dữ liệu phù hợp cho 2 tab **Intervals** và **Forecast**.



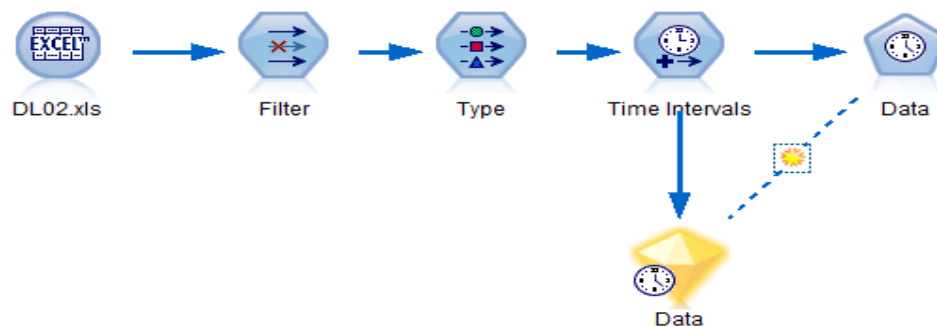
Hình 3.2.3. 4 Xử lý nút *Time Intervals*

Bước 4: Chọn nút **Time Series**, nhấn đúp chọn tab **Model**. Chọn *Exponential Smoothing* cho Method, trong *Criterial chọn Winters Additive* (mô hình cộng)/OK.



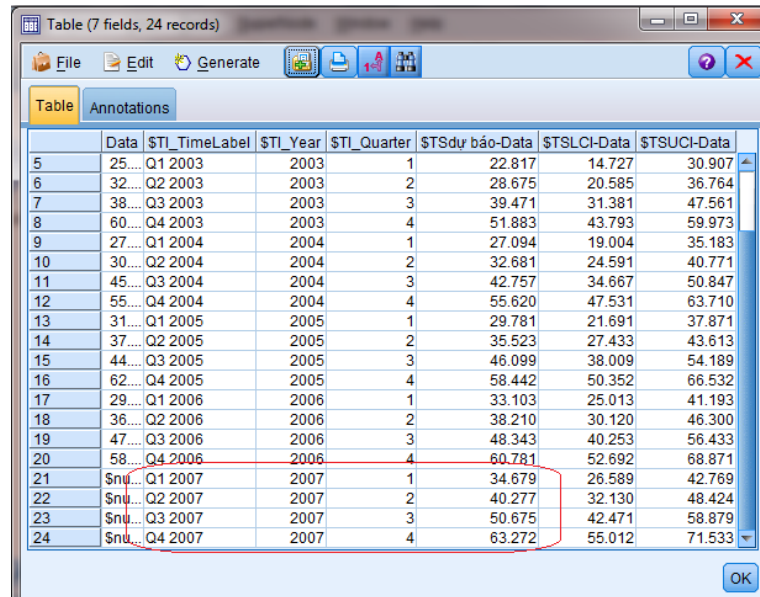
Hình 3.2.3. 5 Chọn nút *Time Series*

Nhấn **Run** được kết quả.



Hình 3.2.3. 6 Kết quả xử lý *Time Series*

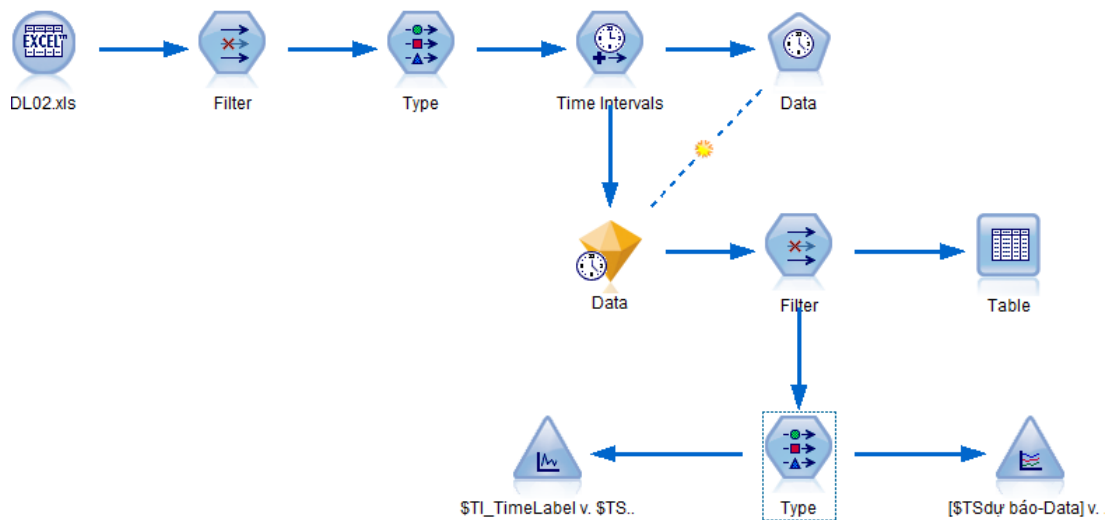
Bước 5: Chọn nút **Filter** để loại bỏ các trường không cần thiết và chọn nút **Table**. Sản lượng dự báo cho 4 quý năm 2009 được hiển thị trong bảng kết quả.



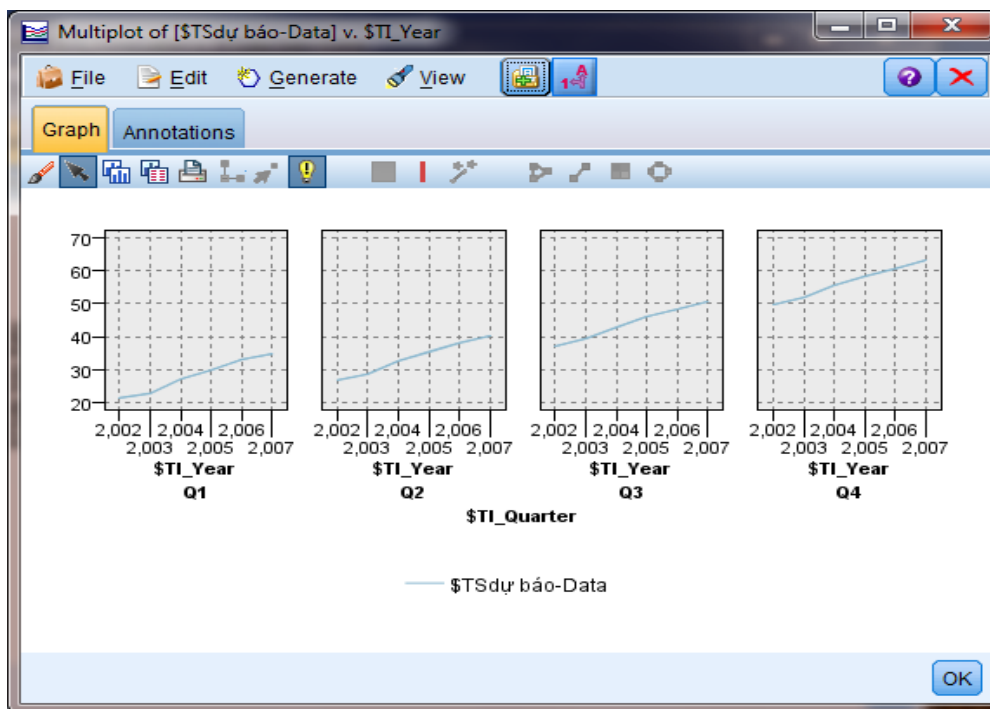
	Data	\$TI_TimeLabel	\$TI_Year	\$TI_Quarter	\$TSDự báo-Data	\$TSLCI-Data	\$TSMCI-Data
5	25...	Q1 2003	2003	1	22.817	14.727	30.907
6	32...	Q2 2003	2003	2	28.675	20.585	36.764
7	38...	Q3 2003	2003	3	39.471	31.381	47.561
8	60...	Q4 2003	2003	4	51.883	43.793	59.973
9	27...	Q1 2004	2004	1	27.094	19.004	35.183
10	30...	Q2 2004	2004	2	32.681	24.591	40.771
11	45...	Q3 2004	2004	3	42.757	34.667	50.847
12	55...	Q4 2004	2004	4	55.620	47.531	63.710
13	31...	Q1 2005	2005	1	29.781	21.691	37.871
14	37...	Q2 2005	2005	2	35.523	27.433	43.613
15	44...	Q3 2005	2005	3	46.099	38.009	54.189
16	62...	Q4 2005	2005	4	58.442	50.352	66.532
17	29...	Q1 2006	2006	1	33.103	25.013	41.193
18	36...	Q2 2006	2006	2	38.210	30.120	46.300
19	47...	Q3 2006	2006	3	48.343	40.253	56.433
20	58...	Q4 2006	2006	4	60.781	52.692	68.871
21	\$nu...	Q1 2007	2007	1	34.679	26.589	42.769
22	\$nu...	Q2 2007	2007	2	40.277	32.130	48.424
23	\$nu...	Q3 2007	2007	3	50.675	42.471	58.879
24	\$nu...	Q4 2007	2007	4	63.272	55.012	71.533

Hình 3.2.3. 7 Bảng kết quả dự báo

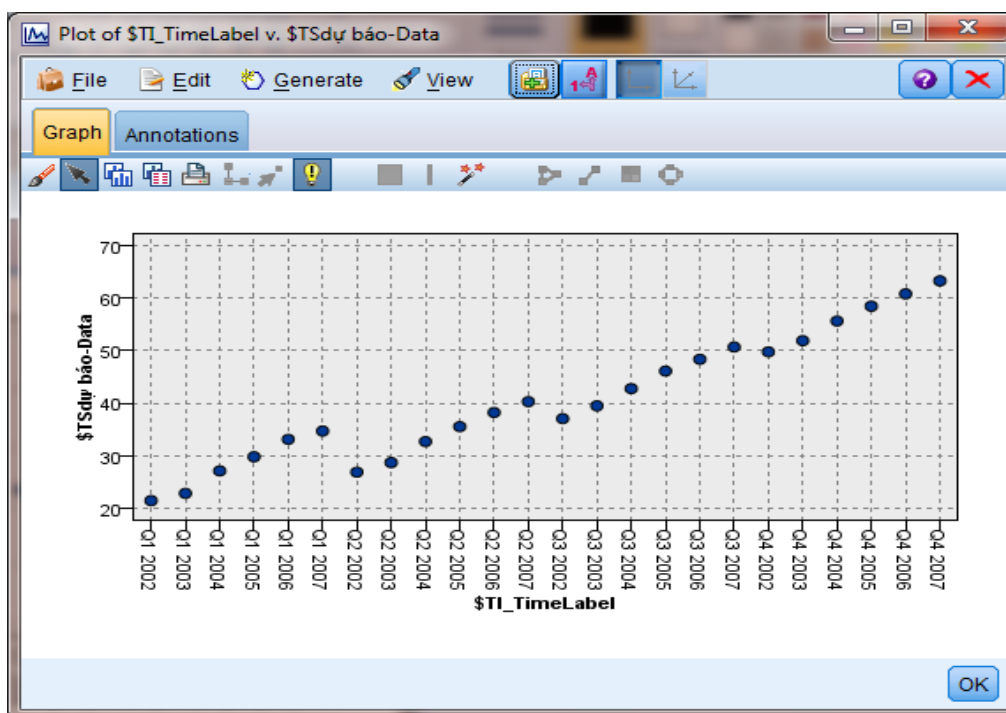
Bước 6: Hiển thị kết quả qua biểu đồ: chọn nút **Multiplot** và **Plot**.



Hình 3.2.3. 8 Chọn nút **Multiplot** và **Plot**



Hình 3.2.3. 9 Biểu đồ kết quả qua Multiplot



Hình 3.2.3. 10 Biểu đồ kết quả qua Plot

KẾT LUẬN

Dự báo nói chung là một bài toán khá thú vị và có ảnh hưởng rất lớn đến việc xác lập các kế hoạch hành động. Nó có một vai trò đặc biệt quan trọng nhất là đối với các đơn vị làm kinh tế trong giai đoạn kinh tế thị trường đầy tính cạnh tranh. Người nhận định đúng về tương lai sẽ là người chiến thắng. Bất cứ một nhận định nào cũng đều có cơ sở suy luận của nó. Cơ sở đó có thể là việc tổng kết kinh nghiệm đã diễn ra trong quá khứ, có thể là các quy luật đã được xác lập hoặc các suy luận logic nhưng nhìn chung mọi cơ sở đều xuất phát từ dữ liệu lịch sử. Song không có nghĩa là nếu ta có dữ liệu lịch sử tốt, có các phương pháp tiên hành dự báo phù hợp, giải thích được quá khứ một cách chính xác là ta có thể yên tâm với kết quả dự báo của mình.

Dự báo bao giờ cũng chỉ là dự báo. Các sự kiện, biến cố ngẫu nhiên làm ảnh hưởng đến kết quả dự báo thì trong một chừng mực nào đó ta không thể xác định được do không thể lập luận đơn giản từ các con số nhưng nó lại là một yếu tố chi phối rất lớn đến kết quả dự báo. Ảnh hưởng của các yếu tố ngẫu nhiên có tác động thực như thế nào đến dự báo thì hoàn toàn phụ thuộc vào nội dung, bản chất của từng biến cố đã hoặc sắp xảy ra mà ta nhận biết được. Việc cảm nhận về các ảnh hưởng đó là tùy vào nhận định của chuyên gia. Rõ ràng là yêu cầu về việc tích hợp hai phương pháp định lượng - định tính là một nhu cầu có thực và sẽ làm tăng độ chính xác cho các bài toán dự báo.

Qua đồ án chúng tôi đã giới thiệu 1 cách khái quát về khái niệm, các phương pháp cũng như nguyên lý xây dựng hệ thống phân tích và dự báo. Bên cạnh đó việc áp dụng phần mềm SPSS Modeler giải quyết các bài toán dự báo sẽ giúp người dùng chuyển hóa dữ liệu và giải quyết nhanh chóng cũng như phát triển các bài toán, mô hình dự báo, hỗ trợ việc ra quyết định.

TÀI LIỆU THAM KHẢO

TÀI LIỆU TIẾNG VIỆT

[1] Lê Văn Duyệt: *Mô hình dự báo ngắn hạn*, Thông tin Khoa học Thống kê-Viện Khoa học Thống kê, 2008.

[2] Lê Xuân Phương: *Ứng dụng phương pháp chuyên gia dự báo trong lĩnh vực Bưu chính – Viễn thông*, Tạp chí Công nghệ Thông tin và Truyền thông, 2004.

[3] Nguyễn Thị Thanh Huyền, Nguyễn Văn Huân, Vũ Xuân Nam: *Bài giảng phân tích và dự báo kinh tế*, Trường đại học Thái Nguyên, 2009.

[4] Nguyễn Trọng Hoài, Phùng Thanh Bình, Nguyễn Khánh Duy: *Dự báo và phân tích dữ liệu trong kinh tế và tài chính*, NXB ĐH Thống kê, 2009.

[5] Nguyễn Văn Cao, Trần Thái Ninh: *Giáo trình Lý thuyết xác suất và Thống kê toán*, NXB ĐH Kinh Tế Quốc Dân, 2008.

[6] Trần Ngọc Vũ: *Giáo trình Hướng dẫn sử dụng SPSS*, Hà Nội, 3/2005.

[7] Sổ tay điều tra viên địa bàn mẫu, Ban chỉ đạo tổng điều tra dân số và nhà ở trung ương, Tổng cục thống kê – Cục thống kê thành phố Hải Phòng, 11/2008.

[8]<http://isponre.gov.vn/home/dien-dan/463-tong-quan-ve-phuong-phap-du-bao-va-kha-nang-ap-dung-mot-so-mo-hinh-trong-du-bao-bien-dong-tai-nguyen-va-moi-truong-tai-viet-nam>

TÀI LIỆU TIẾNG ANH

[1] John E.Hanke, et al: *Business forecastin*, NXB Pearson 2010.

[2] Thomas L, Saaty, Luis G. Vargas: *Prediction - Projection and Forecasting*, Kluwer Academic Publishers, 1991.

[3] <http://www.forecastingmethods.net/>

[4]<ftp://ftp.software.ibm.com/software/analytics/spss/documentation/modeler/14.2/en/>