



## ĐO LƯỜNG CHẤT LƯỢNG

### DỮ LIỆU THỐNG KÊ NGẮN HẠN

*Rudi Sljak, Metka Zaletel, Cơ quan Thống kê Cộng hòa Slovenia*

*Từ khóa: Các chỉ tiêu chất lượng, chất lượng số liệu thống kê ngắn hạn, hệ thống kiểm soát chất lượng*

#### Tóm tắt

Có một sự gia tăng bền vững về nhu cầu của việc cung cấp dữ liệu nhanh chóng, đặc biệt là trường hợp các cuộc điều tra kinh doanh. Các cuộc điều tra hàng tháng hoặc hàng quý nên cung cấp các dữ liệu này thường xuyên trên cơ sở các mẫu và chỉ nên tập trung vào một vài biến (ví dụ, doanh thu, số lượng lao động). Mục tiêu chủ yếu là có được các kết quả đạt chất lượng một cách nhanh nhất có thể. Vì thời hạn công bố các kết quả này ngày càng ngắn nên xuất hiện tình trạng công bố các dữ liệu kém chất lượng nhằm mục đích kịp thời hạn. Bởi vậy, đo lường một cách chắc chắn các khía cạnh khác nhau của chất lượng dữ liệu là điều hết sức quan trọng để tránh việc công bố các kết quả không đáp ứng được các tiêu chuẩn chất lượng.

Trong bài viết này chúng tôi trình bày Hệ thống Quản lý chất lượng kết quả điều tra doanh nghiệp ở 3 lĩnh vực: Bán lẻ; Nhà hàng và khách sạn; Dịch vụ. Kết quả điều tra được công bố là các chỉ số cơ bản, được công bố vào các thời gian khác nhau. Chất lượng được kiểm soát bởi hệ thống chỉ tiêu đã được xác định trên cơ sở của Báo cáo tiêu chuẩn chất lượng và các phương pháp luận của Cơ quan thống kê châu Âu (Eurostat). Bởi hầu hết quy trình thống kê được tự động hóa nên mục đích là nhằm kết hợp việc xử lý cùng việc tính toán phần lớn các chỉ tiêu chất lượng xác định vào quá trình này, kết quả chúng tôi sẽ có các chỉ tiêu sắp xếp cùng một lúc với các chỉ số được ước tính. Chúng tôi mô tả những chỉ tiêu nào được tính toán tự động, những chỉ tiêu nào là vấn đề mà chúng tôi phải đối mặt và kế hoạch tương lai của chúng tôi trong lĩnh vực này.

#### 1. Giới thiệu

Chất lượng các cuộc điều tra và sự hiểu biết thực tế về chất lượng các sản phẩm thống kê ngày càng quan trọng không chỉ đối với người sử dụng dữ liệu mà còn đối với các nhà quản lý các cuộc điều tra và công tác quản lý tại các Viện Thống kê. Họ đều cần có những kết quả một cách nhanh chóng và dễ sử dụng, có thể so sánh được với kết quả của các cuộc điều tra thống kê khác (tương tự trong việc thiết kế, không gian hoặc thời gian).

Những người sử dụng các kết quả đo lường chất lượng tương tự như những người sử dụng kết quả của các cuộc điều tra thống kê: một số yêu cầu báo cáo đầy đủ với những mô tả chính xác của tất cả các ý kiến có thể, những người còn lại muốn có các kết quả thật vắn tắt chỉ một vài bảng, nhiều đồ thị minh họa và nhiều sự so sánh với các dữ liệu có thể áp dụng được. Hiểu được thực tế này, việc triển khai đo lường quản lý chất lượng các cuộc điều tra thống kê đã được thực hiện trước tiên là theo hướng

dẫn của Báo cáo tiêu chuẩn chất lượng, sau đó là tập hợp các chỉ tiêu chất lượng đã được phát triển. Cơ quan Thống kê Cộng hòa Slovenia (SORS) đã theo sát chiến lược này. Chúng tôi đang cố gắng kết hợp việc tính toán các chỉ tiêu chất lượng vào quá trình dữ liệu một cách hoàn toàn tự động. Mục tiêu là có được bộ chỉ tiêu chất lượng hàng tháng một cách nhanh chóng giống như những kết quả đầu tiên của cuộc điều tra và đưa ra Báo cáo tiêu chuẩn chất lượng hoàn toàn tự động.

Thử nghiệm đầu tiên của quá trình đã được tiến hành ở ba cuộc điều tra hàng tháng sẽ được mô tả ở các phần tiếp theo. Sau đó, chúng tôi sẽ tiếp tục việc mở rộng một hệ thống tương tự tới các cuộc điều tra khác, đặc biệt là các cuộc điều tra hàng tháng hoặc hàng quý.

## 2. Thông tin chung của hệ thống

Quá trình quản lý chất lượng là trọng tâm chính của bài viết đã được phát triển hàng đầu cho các cuộc điều tra kinh doanh ngắn hạn, đặc biệt là ba cuộc điều tra hàng tháng: Điều tra kinh doanh bán lẻ, điều tra nhà hàng khách sạn, và điều tra trong lĩnh vực dịch vụ. Đặc điểm phân biệt của các cuộc điều tra này là:

- Chúng đều là các cuộc điều tra hàng tháng với một danh sách ngắn các biến trong bảng hỏi;
- Các kết quả được công bố thường là các loại chỉ số khác nhau;
- Luôn có sự gia tăng nhu cầu về các kết quả nhanh chóng của cuộc điều tra.

Năm 2003 tất cả các phương pháp luận chính của ba cuộc điều tra này đã được chuẩn hóa. Các khía cạnh chính của phương pháp luận đã được chuẩn hóa là:

- Bảng hỏi chỉ bao gồm hai biến: doanh thu và số lượng nhân viên;

- Các cuộc điều tra dựa trên các mẫu mà tất cả các doanh nghiệp lớn được lựa chọn một cách chắc chắn, trong khi đó mẫu ngẫu nhiên được rút ra từ các doanh nghiệp vừa và nhỏ;

- Thiết kế kế mẫu luân phiên với 75% các đơn vị trùng lặp, nghĩa là các đơn vị cũ được giữ lại trong mẫu 12 tháng và sau đó một phần tư mẫu được thay thế bằng các đơn vị mới;

- Các giá trị khuyết (hay đơn vị và mục không có câu trả lời) được ước lượng bằng phương pháp thay thế các giá trị bị khuyết thường được gọi là việc gán xu thế mang tính lịch sử;

- Hệ thống gán trọng số đặc biệt, tính xác suất lựa chọn hàng năm và tỷ lệ ước tính giữa các đơn vị sinh mới và các đơn vị chết đã được phát triển;

- Kết quả đầu tiên của các cuộc điều tra được công bố sau khi kết thúc thời gian chu kỳ cơ sở khoảng 55 ngày.<sup>1</sup>

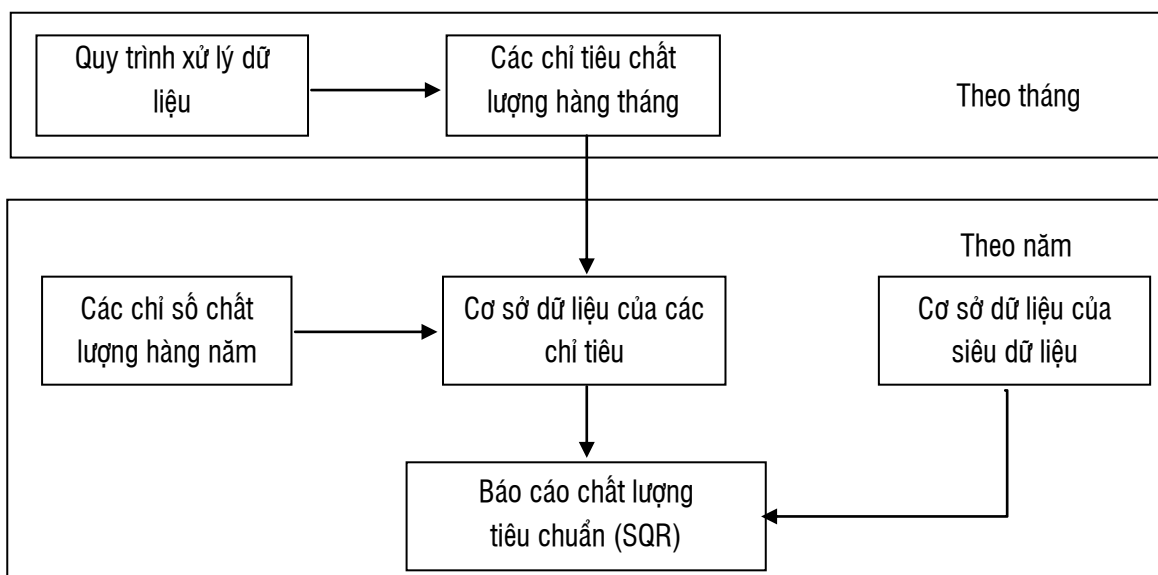
Dữ liệu của các cuộc điều tra được lưu trữ trong hai cơ sở dữ liệu. Cơ sở dữ liệu đầu tiên gồm có thông tin về các đơn vị được xác định tại thời điểm mẫu được rút ra (chẳng hạn như quy mô lớp học, mã các hoạt động thống kê (NACE), doanh thu năm vừa rồi, địa chỉ...). Cơ sở dữ liệu thứ hai gồm các thông tin chúng ta có hàng tháng giống như kết quả điều tra của chúng ta (dữ liệu thô). Dĩ nhiên dữ liệu trong cơ sở dữ liệu này được chốt hàng tháng. Việc hiệu chỉnh dữ liệu báo cáo được cho phép ở năm hiện tại và năm trước đó.

<sup>1</sup> Ngoại lệ với các kết quả của tháng 1 (bởi vì sự luân phiên mẫu) chúng tôi công bố các kết quả khoảng 70 ngày trước ngày kết thúc chu kỳ liên quan.

Việc xử lý dữ liệu gồm có thay thế giá trị khuyết, gán trọng số và việc lập bảng được hoàn toàn tự động thông qua thiết lập hoạt động SAS macro's nhờ việc sử dụng giao diện đồ họa MS-Access thích hợp. Chúng tôi tối thiểu hóa thời gian phần lập bảng này khoảng 15 phút và kết quả là phần lớn thời gian hữu ích còn lại chúng tôi có thể dùng cho các phần khác (ví dụ làm sạch dữ liệu, phân tích dữ liệu...) của việc xử lý.

qua một cách nhanh chóng dẫn đến nguy cơ công bố kết quả kém chất lượng. Để tránh điều đó và để đảm bảo cho việc kiểm soát chất lượng các kết quả đưa ra được nhanh chóng và hiệu quả chúng tôi đã thiết lập một hệ thống quản lý chất lượng, hệ thống này sẽ được trình bày ở các phần tiếp theo. Mục tiêu chính nhằm kết hợp một cách triệt để hệ thống này với quy trình xử lý dữ liệu tự động đã được giải thích ngắn gọn ở trên. Nội dung cơ bản được trình bày ở hình sau:

Như đã đề cập, do nhu cầu có được các kết



Hệ thống được trình bày ngắn gọn như sau:

- Việc kiểm soát các kết quả sản xuất hàng tháng được căn cứ vào các chỉ tiêu chất lượng là tập con của danh sách các chỉ tiêu chất lượng đầy đủ xác định mục tiêu của Báo cáo chất lượng tiêu chuẩn (SQR). Chúng tôi sẽ đề cập tới các chỉ tiêu này như là các chỉ tiêu hàng tháng. Việc tính toán các chỉ tiêu hàng tháng được phối hợp với hệ thống xử lý dữ liệu và được thực hiện hàng tháng.

- Kết thúc năm, chúng ta tính “các chỉ số chất lượng hàng năm” theo phương pháp luận quy định. Một vài chỉ tiêu được tính toán bằng việc sử dụng các chỉ tiêu tháng và một trong số chúng được tính

độc lập. Toàn bộ các chỉ tiêu chất lượng đã thiết lập được lưu trữ trong cơ sở dữ liệu bao gồm các chỉ tiêu cho các cuộc điều tra khác nhau và các năm liên quan khác nhau.

- Ngoài ra một vài thông tin dữ liệu nguyên bản (ví dụ thông tin về dàn mẫu, thiết kế mẫu, phương tiện truyền thông cho việc xuất bản...) cần có trong SQR cũng được lưu trữ trong cơ sở dữ liệu đặc biệt bao gồm các thông tin về các cuộc điều tra và các năm liên quan khác nhau.

Thông tin từ cả hai cơ sở dữ liệu sau đó được gộp thành một mẫu theo quy định chuẩn của SQR. Mẫu chuẩn của SQR được xác định theo

khuôn mẫu của thế giới liên kết trực tiếp với cả hai cơ sở dữ liệu cho phép hoàn thiện quy trình phiên bản SQR cuối cùng nhanh chóng và thuận tiện với người sử dụng. Mỗi mục trên sẽ được giải thích chi tiết ở các phần tiếp theo.

### 3. Danh sách đầy đủ các chỉ tiêu

Danh sách đầy đủ của các chỉ tiêu chất lượng cần có trong SQR được xác định trên cơ sở danh sách đề xuất bởi Nhóm chuyên gia châu Âu về các chỉ số chất lượng. Danh sách đề xuất được nghiên cứu và thảo luận trong nhóm làm việc đặc biệt của SORS. Kết quả các buổi thảo luận này là một danh sách gồm 18 chỉ tiêu được xác định theo yêu cầu tiên hành SQR đưa ra ở SORS. Phương pháp luận của việc tính toán các chỉ tiêu này được dựa theo các quy tắc dưới đây:

- Các giá trị của những chỉ tiêu thuộc khoảng [0,1]. Ngoài ra các giá trị {-2, -1} cũng được cho phép, -1 đại diện cho trường hợp khi giá trị của các

chỉ tiêu không khả dụng và -2 là cho trường hợp khi các chỉ số không thích hợp.

- Các chỉ tiêu nên được xác định theo cách giá trị của chỉ tiêu càng nhỏ nghĩa là chất lượng càng cao. Vì vậy giá trị 0 chính là mức chất lượng lý tưởng.

Trong một số trường hợp giới hạn các giá trị trong khoảng [0,1] theo sát việc xác định các chỉ tiêu, đồng thời bổ sung thêm một vài tính toán nhằm đảm bảo phạm vi quy định là cần thiết. Đối với các trường hợp gần đây được đề xuất nên xuất bản đồng thời, giá trị ban đầu (chúng ta nên gọi là giá trị danh nghĩa) cũng như giá trị được tính toán lại trong khoảng [0,1] (chúng ta nên gọi là giá trị được tiêu chuẩn hóa).

Đầu tiên, chúng tôi trình bày danh sách các chỉ số đầy đủ. Vì mục đích của việc so sánh với danh sách được đề xuất bởi Eurostat, chúng tôi đã sử dụng các kí hiệu tương tự cho các chỉ tiêu thông dụng.

Bộ phận chất lượng	Kí hiệu	Đề mục
Tính phù hợp	R1	Chỉ số hài lòng của người dùng
	R2	Tỷ lệ số liệu thống kê không khả dụng
Tính chính xác	A1	Hệ số biến thiên
	A2	Tỷ lệ đơn vị không có câu trả lời
	A3	Tỷ lệ các mục không có câu trả lời
	A4	Tỷ lệ hiệu chỉnh
	A5	Tỷ lệ thay thế giá trị khuyết
	A6	Tỷ lệ ngoài phạm vi
	A7	Phạm vi trung bình của các sửa đổi
	A8	Tỷ lệ chưa được phân loại
Tính kịp thời và đúng hạn	T1	Tính kịp thời của lần công bố đầu tiên
	T2	Độ trễ trung bình giữa thời điểm kết thúc của kì liên quan và ngày công bố đầu tiên
	T3	Độ trễ trung bình giữa thời điểm kết thúc của kì liên quan và ngày công bố các kết quả cuối cùng

Bộ phận chất lượng	Kí hiệu	Đề mục
Khả năng truy cập và tính rõ ràng	AC1	Loại phương tiện truyền thông được dùng cho việc phổ biến
	AC2	Mức khả năng được dùng cho việc phổ biến
Khả năng so sánh	C1	Độ dài của chuỗi thời gian so sánh
Tính nhất quán	CH2	Sự nhất quán giữa các kết quả đầu tiên và cuối cùng
	CH3	Sự nhất quán giữa dữ liệu so sánh với các nguồn dữ liệu khác

Một số các chỉ tiêu được liệt kê cũng như các giá trị quyền số của các chỉ tiêu nên được công bố. Những chỉ tiêu (nếu khả thi) nên được tính toán là: A2, A3, A4, A5. Quyền số sử dụng cho việc tính toán chỉ nên là quyền số chọn mẫu.

**Tính phù hợp**

Chỉ số hài lòng của người sử dụng được dựa trên cuộc khảo sát sự hài lòng của họ. Ở SORS một cuộc điều tra tương tự đã được lên kế hoạch cho năm này. Vì phương pháp luận của điều tra và tính toán các kết quả vẫn chưa được xác định đầy đủ, nên chúng ta chưa thể xác định chính xác được phương pháp luận về cách tính chỉ tiêu R1.

Chỉ tiêu R2 cho biết tỷ lệ số liệu thông kê không có giá trị trong tổng số các số liệu thông kê, các số liệu này cần được khả dụng nhờ việc điều chỉnh.

**Tính chính xác**

Hầu hết các chỉ tiêu đã được hiểu rõ và trình bày chi tiết ở các nguồn khác nhau của báo cáo. Trong trường hợp chỉ tiêu được xem như các biến đặc trưng, thì chỉ mới nên tính toán những giá trị của các biến quan trọng. Như đã đề cập trước đó, chúng tôi có một số lượng nhất định các biến quan trọng để mô tả các cuộc điều tra (tối đa 10 biến).

Việc tính toán chỉ tiêu A7 (phạm vi trung bình của các sửa đổi) được tính theo công thức sau:

$$A7 = \frac{1}{k-1} \sum_{i=1}^{k-1} |r(i)|$$

Trong đó:

$X_1 \dots X_{k-1}$ : các giá trị những kết quả tạm thời

$X_k$ : giá trị của kết quả cuối cùng

$r(i) = \frac{X_i - X_k}{X_k}$ : hệ số điều chỉnh tương đối

**Tính kịp thời và đúng hạn**

Tất cả các chỉ tiêu T1, T2 và T3 đều đo lường độ trễ giữa hai ngày. Chúng tôi quyết định sử dụng ngày như một đơn vị thời gian cho tất cả các cuộc điều tra khác nhau và các chu kì khác nhau. Chúng tôi cũng phải cân nhắc xem liệu tính toán lại các giá trị thu được trực tiếp trong khoảng xác định [0,1] như thế nào mới đáp ứng được một trong những quy tắc cơ bản. Chúng tôi đã quyết định sử dụng việc tính toán sau:

Với mỗi chu kì khác nhau chúng tôi xác định giới hạn trên cho phép tại đó giá trị của chỉ tiêu có thể bằng 1. Giới hạn trên cho phép được đưa ra là:

**T1** (Tính kịp thời của lần công bố đầu tiên):

Các cuộc điều tra hàng tháng: 10 ngày

Các cuộc điều tra hàng quý: 20 ngày

Các cuộc điều tra hàng năm (vài năm): 30 ngày

**T2** (Độ trễ trung bình giữa thời điểm kết thúc của kì liên quan và ngày công bố đầu tiên)

Các cuộc điều tra hàng tháng: 120 ngày

Các cuộc điều tra hàng quý: 210 ngày

Các cuộc điều tra hàng năm (vài năm): 730 ngày

**T3** (Độ trễ trung bình giữa thời điểm kết thúc của kì liên quan và ngày công bố các kết quả cuối cùng)

Các cuộc điều tra hàng tháng: 485 ngày

Các cuộc điều tra hàng quý: 575 ngày

Các cuộc điều tra hàng năm (và vài năm): 1095 ngày

Giá trị chuẩn hóa của chỉ tiêu sau đó sẽ được tính theo công thức sau:

$$T = \frac{T_m - T_{nom}}{T_m}$$

Trong đó:

T: Giá trị chuẩn hóa của chỉ tiêu

T<sub>nom</sub>: Giá trị ban đầu của chỉ tiêu

T<sub>m</sub>: Giới hạn trên cho phép

#### Khả năng truy cập và tính rõ ràng

Việc tính toán các chỉ tiêu được đưa ra (AC1, AC2) dựa trên hai danh sách: danh sách tất cả các phương tiện truyền thông có thể được sử dụng cho việc phổ biến, danh sách tất cả các khả năng có thể được dùng cho việc phổ biến. Danh sách đầu tiên hiện tại bao gồm 7 mục và danh sách thứ hai gồm 22 mục. Các danh sách có thể sẽ được chỉnh lý và hoàn chỉnh trong tương lai.

#### Khả năng so sánh

Giá trị danh nghĩa của chỉ tiêu C2 được xác định như là số thời điểm khác nhau trong chuỗi thời

gian của các giá trị riêng biệt. Giá trị chuẩn hóa của chỉ tiêu sẽ được tính bởi công thức sau:

$$C2 = 1 - [Y/4] \cdot 0,2$$

Y: Số năm trong chuỗi thời gian

[.]: Giá trị rút gọn của số

Nếu giá trị thu được từ công thức trên nhỏ hơn 0, thì giá trị của chỉ tiêu bằng 0.

#### Tính nhất quán

Chỉ tiêu CH2 so sánh các kết quả điều tra với các kết quả của một vài nguồn liên quan đã được xác định trước đó. Nguồn liên quan có thể là dữ liệu từ các cuộc điều tra tương tự, từ hệ thống tài khoản quốc gia hoặc là những dữ liệu hành chính.

#### 4. Các chỉ tiêu hàng tháng

Như chúng tôi đã đề cập trước đó trường hợp các mục điều tra ngắn hạn một vài chỉ tiêu chất lượng được tính toán tự động trong quá trình lập bảng dữ liệu. Các chỉ tiêu này là: A1, A2, A3, A5, A6, A8, T1, T2.

#### Chú ý:

Kết quả của các cuộc điều tra ngắn hạn là những chỉ số riêng biệt của các loại khác nhau. Rõ ràng các chỉ số đó không phải là ước lượng tuyến tính, vì vậy việc tính toán hệ số biến thiên (chỉ tiêu A1) là không đơn giản. Chúng tôi sẽ tính tới việc các chỉ số là các ước lượng thông thường từ dữ liệu của hai mẫu khác nhau ngay cả khi các đơn vị tương tự có các quyền số khác nhau. Bởi vì phần mềm ước lượng hệ số biến thiên cổ điển không thể sử dụng để giải quyết tình huống tương tự như thế này, nên chúng tôi đã phát triển lên phần mềm SAS macro's riêng của mình (SAS/SURVEYMENS) để ước lượng hệ số biến thiên của các chỉ số. Hệ số biến thiên



được ước lượng cho hai loại chỉ số khác nhau với tất cả các lĩnh vực khác nhau tại đó các kết quả được công bố.

Bởi vì tất cả các ngày khuyết (không có câu trả lời) đều được thay thế nên giá trị của chỉ tiêu A5 bằng với giá trị của chỉ tiêu A3.

Các giá trị của chỉ tiêu A6 và A8 phản ánh một cách lý tưởng tỷ lệ ngoài phạm vi và tỷ lệ chưa được phân loại với toàn bộ dân mẫu. Do chúng ta chỉ có thông tin về đơn vị bao gồm trong mẫu nên hai chỉ tiêu này chỉ có thể ước lượng.

Chỉ tiêu T1 và T2 được tính toán hoàn toàn tự động như những phần còn lại của các chỉ tiêu đã được liệt kê ở trên. Người phụ trách điều tra chỉ cần nhập ngày công bố đầu tiên vào ô xác định trong bảng tính Excel và các giá trị của những chỉ tiêu sẽ được tính toán một cách tự động.

Các chỉ tiêu hàng tháng được tính toán tự động bằng quy trình SAS macro's sau đó xuất ra

một vài tệp Excel tuy nhiên các nhà phương pháp luận có thể quan sát chúng chỉ trong một tệp Excel được liên kết với toàn bộ các tệp kết quả công việc. Việc tính toán các chỉ tiêu luôn được thực hiện cho các kết quả mỗi tháng của năm hiện tại. Điều này cho phép các nhà phương pháp luận không chỉ quan sát được các kết quả của các chỉ tiêu được công bố mà còn so sánh được các kết quả này với kết quả của các tháng khác. Các chỉ tiêu có thể quan sát được dưới hình thức bảng Excel và cũng có thể được trình bày dạng đồ thị bằng biểu đồ Excel.

Ví dụ chúng tôi trình bày các kết quả về sự biến thiên một trong các chỉ số đã được tính ở cuộc Điều tra kinh doanh bán lẻ 9 tháng năm 2003. Phần bảng dưới đây cho thấy các chỉ tiêu ở ba lĩnh vực. Như chúng tôi đã mô tả, cùng với các chỉ tiêu của tháng hiện tại, các chỉ tiêu của tất cả các tháng sau trong năm hiện tại cũng luôn luôn được trình bày trong bảng.

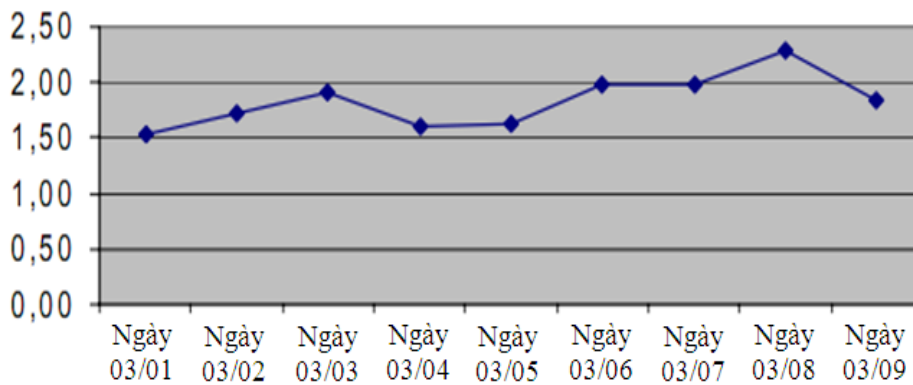
**Hệ số biến thiên (phần trăm)**

	Ngày 03/01	Ngày 03/02	Ngày 03/3	Ngày 03/4	Ngày 03/5	Ngày 03/6	Ngày 03/7	Ngày 03/8	Ngày 03/9
Tổng	1.53	1.73	1.91	1.60	1.62	1.98	1.98	2.28	1.84
Hàng tạp hóa	4.81	5.63	3.98	4.01	4.23	4.47	4.30	3.81	4.67
Kinh doanh xe cộ	2.22	2.91	2.84	3.58	2.93	3.43	3.22	2.92	2.72

Nếu các nhà nghiên cứu muốn quan sát các kết quả được trình bày bằng đồ thị chỉ cần click vào nút lệnh trên màn hình tại đây có thể chọn các lựa chọn như sau:

Các biểu đồ đường biểu thị các kết quả của các khu vực được chọn trong tất cả các tháng ở năm hiện tại. Ví dụ của chúng tôi cho lựa chọn khu vực tổng thì đồ thị sẽ là:

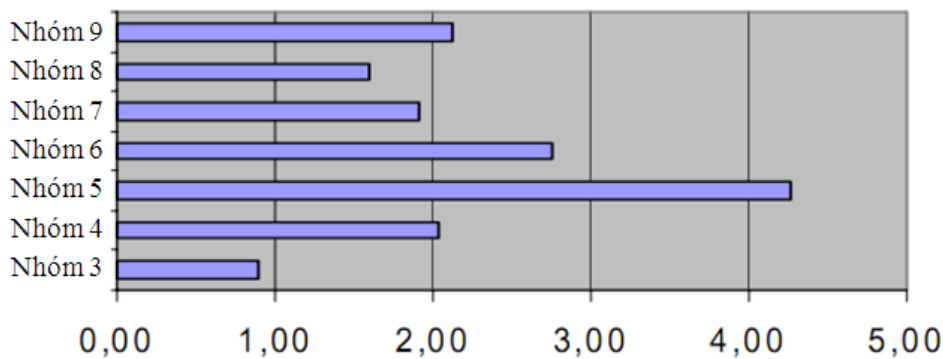
Tổng các hệ số biến thiên



Biểu đồ cột của hệ số biến thiên trung bình cho tất cả các nhóm công việc trong các năm hiện tại. Biểu đồ này cho phép những nhà nghiên cứu phát hiện những nhóm công việc mà hệ số biến thiên của các

chỉ số được tính toán cao và có thể coi như là “có vấn đề”. Trường hợp Điều tra kinh doanh bán lẻ chúng tôi có 15 nhóm công việc. Ở đây chúng tôi chỉ trình bày 7 nhóm.

Hệ số biến thiên trung bình



### 5. Báo cáo tiêu chuẩn chất lượng

Danh sách các chỉ tiêu hàng tháng đã được mô tả được tính cho mỗi tháng nhưng Báo cáo tiêu chuẩn chất lượng sẽ được chuẩn bị trong cả năm. Bởi vậy khi các kết quả của từng năm được công bố lần cuối cùng thì toàn bộ các chỉ tiêu hàng tháng cần được tổng hợp (thường tính trung bình các chỉ tiêu tháng). Đồng thời phần còn lại của các chỉ tiêu chỉ có thể được tính sau khi hoàn thành kết quả 12 tháng, chúng có thể được lồng ghép vào cơ sở dữ liệu các chỉ tiêu chất lượng. Như chúng tôi đã đề cập trước đó cơ sở

dữ liệu bao gồm các chỉ tiêu cho vài cuộc điều tra. Dữ liệu trong cơ sở dữ liệu được tổ chức theo cách tất cả các chỉ tiêu của các cuộc điều tra riêng biệt trong mỗi năm được gộp lại với nhau trong cùng một hồ sơ. Số xác định hồ sơ chính là mã cuộc điều tra cùng với năm liên quan. Một số chỉ tiêu được soạn thảo theo các mẫu khác nhau: không quyền số, quyền số, định danh, chuẩn hóa. Trường hợp chỉ tiêu liên quan tới các biến đặc trưng, các chỉ tiêu của tất cả các biến quan trọng đều được lưu trữ. Vì lý do này hồ sơ trong cơ sở dữ liệu có tới 136 biến.



Các chỉ tiêu được chèn vào cơ sở dữ liệu bằng giao diện MS-Access cũng bao gồm một số lệnh cơ bản kiểm soát tính hợp lý (ví dụ giá trị nhập vào chỉ có thể là các giá trị trong khoảng [0,1]). Giao diện cho phép người dùng khả năng tìm kiếm, biên tập và chèn thêm các chỉ tiêu mới.

Cơ sở dữ liệu thứ hai là cơ sở dữ liệu thông tin nguyên bản của cuộc điều tra nên điền sau khi hoàn thành từng năm liên quan. Cơ sở dữ liệu được tổ chức tương tự theo cách tổ chức của cơ sở dữ liệu các tiêu chí chất lượng nghĩa là tất cả thông tin của các năm riêng biệt được lưu trữ trong hồ sơ của cơ sở dữ liệu và sự xác định hồ sơ cũng là mã điều tra cùng với năm liên quan. Giao diện cũng sắp xếp tương tự như trường hợp cơ sở dữ liệu các chỉ tiêu chất lượng.

Có một lượng lớn thông tin nguyên bản nên cung cấp cho mỗi cuộc điều tra. Để minh họa đặc điểm của những thông tin này ở đây chúng tôi chỉ đưa ra một phần nhỏ của danh sách hoàn chỉnh các thông tin được yêu cầu:

- Sự mô tả cấu trúc của những người điều tra. Cấu trúc nên dựa vào việc phân loại tiêu chuẩn của người sử dụng.
- Sự mô tả mục tiêu tổng thể và quy trình xây dựng dàn mẫu.
- Sự mô tả việc thiết kế mẫu.
- Sự mô tả hệ thống quyền số và sử dụng phương pháp mô thay thế giá trị khuyết.
- Nguyên nhân khả năng tạm dừng lần công bố đầu tiên.
- Danh sách chi tiết của tất cả các khả năng dùng cho việc phổ biến.

- Sự mô tả quy trình dùng cho việc quản lý dữ liệu công khai.

- Người dùng có thể truy cập dữ liệu được công bố theo điều kiện nào và truy cập ở đâu.

- Nguyên nhân độ lệch lớn của các kết quả cuối cùng so với kết quả đầu tiên.

Bước cuối cùng của quy trình chuẩn bị báo cáo chất lượng tiêu chuẩn SQR với cuộc từng điều tra và từng năm, cần gộp những dữ liệu từ cả hai cơ sở dữ liệu đã được mô tả thành một mẫu tài liệu chuẩn theo quy định. Phần chuẩn hóa của tài liệu được chuẩn bị theo mẫu của MS-Word. Khuôn mẫu chính là các lĩnh vực đã xác định được liên quan trực tiếp tới các biến trong các cơ sở dữ liệu. Khi người chịu trách nhiệm điều tra muốn chuẩn bị bản SQR cuối cùng anh ta chỉ cần chọn đúng hồ sơ trong cơ sở dữ liệu (có thể tiến hành một vài khách hàng được lựa chọn) và lưu lại tài liệu như một bản SQR.

## **6. Kết luận**

Sự thử nghiệm về hệ thống tính toán các chỉ tiêu chất lượng vừa được mô tả bên trên. Cuối cùng, ta có thể nói rằng hệ thống xử lý dữ liệu với việc mô tả các cuộc điều tra đã được phát triển trong những điều kiện đặc biệt nguyên nhân làm cho hệ thống này khác biệt với các hệ thống “thông thường” ở cơ quan tương tự. Chúng tôi dự tính rằng khả năng việc thiết lập hệ thống tương tự cho công tác quản lý chất lượng các cuộc điều tra khác có thể sẽ có thêm nhiều thách thức và yêu cầu hơn. Mặt khác, kinh nghiệm thực tế rất lạc quan cho thấy nhiều nhà quản lý điều tra mơ ước có thể thúc đẩy sự phát triển mạnh mẽ của hệ thống các cuộc điều tra quan trọng khác theo hướng tương tự.

*(Xem tiếp trang 13)*

---

***(Tiếp theo trang 36)***

Tóm lại, ưu điểm của hệ thống là:

- Hệ thống cho phép các nhà nghiên cứu truy cập một cách nhanh chóng thông tin về chất lượng dữ liệu.

- Như chúng tôi đã thiết lập suốt khoảng thời gian thử nghiệm hệ thống, nó không chỉ có khả năng kiểm soát chất lượng mà còn cho phép phát hiện được các lỗi trong hệ thống xử lý hoặc ngay cả trong các dữ liệu được báo cáo.

- Mẫu được chuẩn hóa của SQR sẽ làm cho toàn bộ lượng thông tin trong SQR trở nên dễ đọc dễ so sánh hơn.

- Việc lưu trữ các chỉ tiêu chất lượng trong cơ sở dữ liệu tương tự cho phép kiểm soát một cách dễ dàng và hiệu quả mức độ chất lượng đạt được của từng cuộc điều tra theo thời gian.

Nhược điểm chính của hệ thống và thách thức trong tương lai:

- Một số chỉ tiêu được mô tả theo định nghĩa vẫn còn nhạy cảm trong khâu sửa đổi và hoàn thiện.

- Có một mối đe dọa từ việc giải thích sai tập hợp các chỉ số nếu ai đó khi tiến hành phân tích lại không tính đến việc giải thích sai lầm này cũng như các thông tin nguyên bản.

- Vì chúng tôi hy vọng sự phát triển hơn của lĩnh vực các chỉ tiêu chất lượng trong hệ thống Thống kê châu Âu nên khả năng chúng tôi sẽ phải luôn luôn điều chỉnh hệ thống của mình. Bởi vậy mục tiêu là có được một hệ thống đủ linh hoạt có thể phù hợp với sự phát triển trong tương lai.

***Minh Ánh (dịch)***

*Nguồn: <http://www.oecd.org/std/35238995.pdf>*