



HỢP NHẤT TỔNG ĐIỀU TRA NÔNG NGHIỆP VỚI CHƯƠNG TRÌNH THỐNG KÊ KINH DOANH: CƠ SỞ CHÍNH CHO TỔNG ĐIỀU TRA TIẾP THEO Ở CANADA

Tóm tắt

Bài viết mô tả cách tích hợp Tổng điều tra nông nghiệp với Chương trình thống kê kinh doanh, sự hợp nhất này là một yếu tố quan trọng trong ngắn hạn để định vị chiến lược cho chương trình tổng điều tra trong tương lai.

1. Giới thiệu

Tổng điều tra nông nghiệp gần đây nhất ở Canada được thực hiện vào năm 2016, cách tiếp cận được Cơ quan Thống kê Canada (Statcan) sử dụng trong cuộc tổng điều tra này là liên hệ với toàn bộ nông dân 5 năm một lần để thu thập tất cả dữ liệu thông qua bảng hỏi. Các khái niệm, hệ thống, công cụ và phương pháp duy nhất đã được sử dụng trong suốt quy trình thống kê.

Trong bối cảnh phát triển nhanh chóng liên quan đến dữ liệu lĩnh vực nông nghiệp, đưa ra cả cơ hội và những thách thức. Người sử dụng dữ liệu, nhà hoạch định chính sách và nông dân muốn có được thông tin chất lượng cao, thời gian thực để đưa ra quyết định sáng suốt. Mục tiêu cuối cùng của Chương trình thống kê nông nghiệp (ASP) là tạo ra thông tin chi tiết sát với thời gian thực, với sự tiếp xúc tối thiểu với người trả lời bằng cách khai thác các nguồn thông tin thay thế chẳng hạn như hình ảnh vệ tinh, dữ liệu hành chính được thu thập bởi các tổ chức khác nhau hoặc trong các bộ dữ liệu khác và sử dụng các phương pháp tiên tiến mới nhất để xây dựng các mô hình hiệu quả. Để thích ứng với tình hình thay đổi nhanh chóng này,

người ta nhận thấy rằng các thực hành và phương pháp được sử dụng trong Tổng điều tra nông nghiệp và mức độ tích hợp của nó với phần còn lại của các hệ thống thống kê cần phải trải qua một sự chuyển đổi sâu sắc. Việc tích hợp Tổng điều tra nông nghiệp (CEAG) - xương sống của ASP - với Chương trình thống kê kinh doanh tích hợp (IBSP) là bước thiết yếu đầu tiên.

2. Bối cảnh thay đổi nhanh chóng

Từ năm 1956, CEAG đã sử dụng mô hình thu thập dựa trên bảng liệt kê đầy đủ các trang trại và dữ liệu thu được trực tiếp từ người trả lời. Cũng giống như nhiều lĩnh vực khác trong nền kinh tế, ngành nông nghiệp đang trải qua một sự chuyển đổi nhanh chóng và sâu sắc, khiến các nhà cung cấp số liệu thống kê hàng đầu phải thay đổi mô hình sản xuất truyền thống để đáp ứng yêu cầu của người dùng.

i) Việc hợp nhất và phức tạp hóa các cấu trúc

Kết quả từ các cuộc điều tra trước đây cho thấy một sự hợp nhất quy mô lớn của các trang trại nông nghiệp ở Canada. Số trang trại giảm 30% từ năm 1996 đến 2016,

nhưng quy mô đã phát triển lớn hơn, phức tạp hơn và tích hợp với nhau nhiều hơn. Trong năm 2016, 25,1% các hoạt động nông nghiệp được hợp nhất, so với 2,2% vào năm 1971. Khi các hoạt động nông nghiệp hợp nhất thì mức độ phức tạp liên quan đến hoạt động nông nghiệp cũng tăng lên. Việc tách một liên doanh nông nghiệp thành các đơn vị kinh doanh riêng biệt hợp pháp làm tăng thêm các thách thức trong việc thu thập và xử lý dữ liệu khi sử dụng phương pháp điều tra truyền thống. Trong hệ thống quy trình sản xuất của Statcan, trang trại là loại hình có thể được xử lý tốt hơn.

ii) Cần có thông tin thống nhất và kịp thời hơn

Nhu cầu của người dùng đang thay đổi nhanh chóng. Người dùng cần dữ liệu và bản phân tích chi tiết và dễ dàng truy cập. Các chính sách và vấn đề bao quát liên quan đến lĩnh vực nông nghiệp vượt xa phạm vi cơ bản. CEAG cần phải được điều chỉnh cho phù hợp.

Cho dù vấn đề là vận chuyển hiệu quả hàng nông sản từ trang trại đến thị trường bằng cách thu thập dữ liệu về năng lực sản xuất theo thời gian thực, hay là đánh giá tác động của biến đổi khí hậu đối với thực tiễn hoặc hỗ trợ phát triển bền vững, hay phát triển sản phẩm/thực phẩm mới hoặc xuất khẩu sang các thị trường quốc tế, thì để đo lường mức độ sẵn có của lương thực và an ninh lương thực và tác động của chúng đối với chất lượng cuộc sống, các bên liên quan hy vọng cơ quan thống kê sẽ tạo ra dữ liệu nông nghiệp có mức độ chi tiết và chất lượng cao một cách thường xuyên hơn, và dữ liệu được tích hợp đầy đủ với dữ liệu từ các ngành kinh tế khác để hỗ trợ ra quyết định dựa trên bằng chứng.

iii) Việc phổ biến dữ liệu và sự cải thiện các kỹ thuật xử lý các bộ dữ liệu lớn

Việc tăng cường khả năng tiếp cận vào dữ liệu hành chính và các giao dịch và truy cập vào hình ảnh vệ tinh chất lượng cao, được hỗ trợ bởi các kỹ thuật hiện đại để xử lý bộ dữ liệu lớn, cho phép việc xem xét các phương pháp mới để tạo số liệu thống kê phù hợp và kịp thời, giảm thiểu sự tiếp xúc với người trả lời.

Sự sẵn có của các nguồn dữ liệu thay thế đã tăng lên, đến nay, hơn 300 nguồn dữ liệu có sẵn sẽ được sử dụng ở các giai đoạn khác nhau của chu kỳ điều tra ASP. Các nguồn dữ liệu này bao gồm các nguồn từ Cơ quan Thuế Canada (ví dụ: dữ liệu thuế) và các lĩnh vực được quản lý cung ứng (bao gồm sữa, gà, trứng và gà tây), trong đó các bộ dữ liệu bao gồm hạn ngạch và giá trị sản xuất.

Dữ liệu bảo hiểm cây trồng, trong đó nêu chi tiết những cây trồng đã trồng và được bảo hiểm, cũng như năng suất của chúng, thể hiện một nguồn thông tin thực tế lớn có thể kết hợp với hình ảnh vệ tinh, dữ liệu khí hậu và mô hình kỹ thuật tiên tiến để sản xuất ra nông sản chất lượng cao thường xuyên và ước tính diện tích cây trồng ở phạm vi địa lý nhỏ mà không cần liên hệ với nông dân. Khả năng tiếp cận vào cơ sở dữ liệu truy xuất nguồn gốc động vật, theo dõi sự biến động và phương pháp tiếp cận điển hình trong kỹ thuật học máy có thể được sử dụng để tạo dữ liệu biến động hàng ngày của vật nuôi và kiểm kê thịt lợn theo thời gian thực từ thông tin biến động tổng hợp. Tận dụng các kỹ thuật mới và các nguồn thông tin thay thế này chắc chắn sẽ thay đổi cách thức thực hiện tổng điều tra theo truyền thống.

Việc chuyển dữ liệu CEAG sang mô hình hoạt động để sản xuất số liệu thống kê kinh tế là bước thiết yếu đầu tiên để điều chỉnh

➤➤➤ THÔNG KÊ QUỐC TẾ VÀ HỘI NHẬP

chương trình với bối cảnh đang thay đổi nhanh chóng.

3. Chương trình thống kê kinh doanh tích hợp

Năm 2014, Statcan đã đưa ra Chương trình thống kê kinh doanh tích hợp (IBSP) để có một mô hình sản xuất số liệu thống kê kinh tế hiệu quả hơn. IBSP cung cấp một khung chung và chuẩn từ quá trình thu thập đến phổ biến và một bộ công cụ và hệ thống phổ biến cho một số lượng lớn các cuộc điều tra kinh tế không đồng nhất do Statcan thực hiện. Đến năm 2021, gần 130 chương trình kinh tế (hàng năm, chu kỳ ngắn hạn, theo ngành, toàn nền kinh tế, dựa trên hoạt động, tài chính) sẽ được tích hợp vào khuôn khổ hài hòa này, bao gồm tất cả các cuộc điều tra thuộc APS và tổng điều tra nông nghiệp năm 2021.

4. Tổng điều tra nông nghiệp là một thành phần tích hợp của Chương trình thống kê kinh tế: Những thay đổi và cơ hội

Phần này nêu bật các yếu tố quan trọng nhất của việc chuyển CEAG sang IBSP và các cơ hội của chúng.

1) Sử dụng hệ thống đăng ký kinh doanh của Statcan làm một khung chung

Khi chương trình CEAG chuyển sang tích hợp các nguồn thông tin thay thế, việc sử dụng hệ thống đăng ký kinh doanh (BR) làm khung là một yếu tố thiết yếu. BR là công cụ chung cho tất cả các cuộc điều tra kinh tế. BR là một công cụ thiết yếu giúp duy trì và lập hồ sơ các doanh nghiệp nông nghiệp ngày càng lớn và phức tạp và hoạt động của chúng. Cơ sở dữ liệu này được cập nhật bằng cách kết hợp các nguồn hành chính (bao gồm dữ liệu thuế), phản hồi các cuộc điều tra và kết quả từ việc liên hệ trực tiếp với người

trả lời của doanh nghiệp.

Ngoài việc liên tục cập nhật, BR chứa nhiều biến được yêu cầu phải có liên kết hồ sơ hiệu quả từ các nguồn hành chính khác nhau (số doanh nghiệp, tên pháp lý, tên thương mại, địa chỉ, chủ trang trại, tọa độ địa lý, v.v...) và các chương trình thống kê. Các mối liên kết hiệu quả của nhiều bộ dữ liệu thay thế với BR là một yếu tố trung tâm để tiến tới việc sản xuất dữ liệu sát thời gian thực với hạn chế sự tiếp xúc với người trả lời. Việc sử dụng BR làm khung, cập nhật tính kịp thời và việc chia sẻ dữ liệu với cơ quan thuế đã dẫn đến việc tạo ra cuộc tổng điều tra các chủ trang trại và hộ gia đình có hoạt động nông nghiệp hàng năm để ước tính một loạt các biến tài chính mà không cần liên hệ với người trả lời. Các biến này sẽ không cần phải thu thập từ những người trả lời trong chương trình tổng điều tra 5 năm một lần.

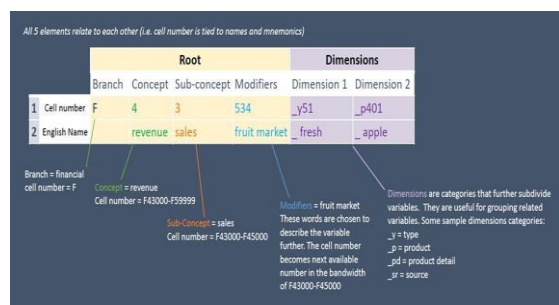
Trong cuộc tổng điều tra nông nghiệp năm 2016, khung gồm hai phiên bản: (1) các doanh nghiệp liên quan đến nông nghiệp được xác định trong BR, bao gồm tất cả các đơn vị thuộc tất cả các thành phần kinh tế liên quan đến sản xuất kinh tế ở Canada; và (2) các hộ gia đình được xác định thông qua bảng hỏi tổng điều tra. Để phù hợp hơn với khái niệm được sử dụng trong chương trình thống kê kinh tế và BR, khái niệm xác định số trang trại được sử dụng trong CEAG đang được xem xét. Thành phần thứ hai của khung xuất phát từ việc tự nhận dạng trong tổng điều tra dân số đã bị loại bỏ. Phần lớn các hoạt động được xác định thông qua tổng điều tra dân số đã có trong BR (86%). Những phần thiếu sót chủ yếu là các hoạt động không có tác động thực sự về kinh tế (0,4% giá trị bán hàng) hoặc tác động chính sách (ví dụ: thành lập trang trại vì sở thích). Việc xác định, thu thập và xử lý các đơn vị này rất tốn kém và có rất ít tác động đến

những con số tổng khác nhau. Với Tổng điều tra nông nghiệp năm 2021, chúng tôi đề xuất xác định một trang trại hoặc hộ nông nghiệp là một đơn vị sản xuất nông nghiệp và báo cáo thuế về doanh thu hoặc chi phí nông nghiệp. Thông tin từ cơ quan thuế được sử dụng để cập nhật BR. Sự hội tụ của các định nghĩa về cơ bản có ý nghĩa là số trang trại được sử dụng trong APS sẽ được cập nhật theo thời gian thực và sẽ không bị hạn chế giữa các cuộc điều tra.

ii) Nội dung và khái niệm được hài hòa hóa

Với việc chuyển sang IBSP, tất cả biến trong tổng điều tra nông nghiệp được đặt tên theo khung IBSP. Trong IBSP, khung đặt tên được áp dụng một cách nhất quán, mạch lạc và lô-gic có thể được nhân rộng cho tất cả các chương trình thống kê. Tất cả các biến được gán một mã số ô được tạo một cách hợp lý từ các tên biến và chứa ý nghĩa ngữ nghĩa. Với sự tích hợp của một cuộc điều tra vào IBSP, tất cả các biến được xem xét kỹ lưỡng, hài hòa hóa, phân tách. Hình 1 cho thấy một ví dụ về khung đặt tên.

Hình 1: Ví dụ về khung đặt tên trong IBSP



Nguồn: Cục thống kê Canada, Hướng dẫn tham khảo Quick Click dành cho Thông tin cơ bản

Ví dụ, câu hỏi về doanh thu đến từ việc bán táo tươi tại thị trường trái cây được phân tách và hài hòa hóa với khung đặt tên biến IBSP. Trong ví dụ này, khái niệm được sử dụng là 'doanh thu'. Trong IBSP, đối với tất

cả các chương trình thống kê, khái niệm doanh thu được xác định bằng các số ô từ F43000 đến F59999. Thành phần tiếp theo là khái niệm phụ. Đối với tất cả các chương trình IBSP, khái niệm phụ về doanh thu có tên 'doanh số' được xác định bằng số ô giữa F43000 và F45000. Phần thứ ba là mã sửa đổi cung cấp mô tả về các biến. Trong ví dụ ở đây, mã sửa đổi là thị trường trái cây. Thành phần quy mô phân loại các biến. Tổng điều tra nông nghiệp năm 2021 đã thông qua khung đặt tên doanh nghiệp được sử dụng trong IBSP. 630 biến khác nhau bao phủ 12 khái niệm, 21 khái niệm phụ, 86 mã sửa đổi và 15 quy mô biến và 331 quy mô nhỏ (biến duy nhất). CEAG đã chia sẻ các ô dữ liệu với 13 chương trình khác nhau.

Cấu trúc này cung cấp một khung mạch lạc, rõ ràng tạo điều kiện cho việc tích hợp dữ liệu của các chương trình khác trong chiến lược 'thu thập một lần sử dụng nhiều lần' hoặc sử dụng dữ liệu chương trình khác để xác thực hoặc phân tích theo chiều ngang. Với khung đặt tên hài hòa này, việc sản xuất, phân tích chuyên sâu và kịp thời theo yêu cầu của người dùng để đưa ra các quyết định dựa trên bằng chứng về các vấn đề ngày càng phức tạp ảnh hưởng đến các ngành liên quan đến nông nghiệp (ví dụ, vận chuyển chế biến thực phẩm, môi trường, nông thôn, thương mại quốc tế, và chỉ số giá nông nghiệp) sẽ dễ dàng hơn. Không cần xây dựng các phép biến đổi hoặc bảng thống kê phức tạp về các biến trong các chương trình hoặc nguồn dữ liệu khác nhau để tích hợp dữ liệu thay thế.

iii) Bắt buộc sử dụng các công cụ doanh nghiệp và hệ thống chung

Bằng cách áp dụng các công cụ và hệ thống doanh nghiệp được sử dụng bởi IBSP, CEAG sẽ được hưởng lợi, từ các phương pháp

➤ ➤ ➤ THÔNG KÊ QUỐC TẾ VÀ HỘI NHẬP

thống kê tiên tiến nhất được giới thiệu để cải thiện phương pháp, chất lượng và hiệu quả dữ liệu mà không phải đầu tư nhiều tại địa phương, lại có thể bù đắp sự lỗi thời của các hệ thống. Cho dù mục đích là để thiết kế phương tiện thu thập hay quản lý khung, liên kết các bộ dữ liệu hiệu quả, xử lý dữ liệu bị thiếu, lập bảng ước tính, trực quan hóa và phân tích dữ liệu, hay để quản lý các hoạt động hạn chế rủi ro tiết lộ thông tin, hay để tạo ra các chỉ số chất lượng và để công bố kết quả, CEAG sẽ sử dụng tất cả các hệ thống doanh nghiệp và công cụ có sẵn.

CEAG tiếp theo sẽ được hưởng lợi trực tiếp từ những tiến bộ này trong các hệ thống doanh nghiệp:

a. Để tiếp cận hơn trong việc tích hợp những sai lệch từ phương pháp gán giá trị vào hệ thống ước lượng chung. Việc chấp nhận IBSP làm đầu ra tiêu chuẩn có nghĩa là, các cuộc điều tra hiện có cách để định lượng sai số chọn mẫu luôn tồn tại nhưng không bao giờ được đo lường. Người dùng sẽ được thông báo tốt hơn về chất lượng tổng thể của các ước tính.

- Các ước tính tổng thể có chất lượng từ các nguồn dữ liệu không trực tiếp (dữ liệu hành chính, hình ảnh vệ tinh, mô hình, v.v...) với các phương pháp có thể chấp nhận.

- Giới thiệu các cách tiếp cận sáng tạo (ví dụ: điều chỉnh bằng ngẫu nhiên) để quản lý việc tránh tiết lộ thông tin mà không cần triệt tiêu dữ liệu để có thể công bố lượng dữ liệu lớn hơn, nhưng không cần phải phức tạp làm nhiều các bản ghi dữ liệu vi mô.

- Giới thiệu các phương pháp thu thập dữ liệu và ưu tiên trường hợp cải tiến: Việc tạo ra một bộ chỉ số chất lượng để hỗ trợ tích cực cho việc quản lý các hoạt động thu thập và theo dõi là một thủ tục mới thay thế các phương pháp chủ trọng quá khứ để xác định

hồ sơ nào được ưu tiên để theo dõi, khi bộ dữ liệu dừng ở một số lĩnh vực nhất định hoặc thậm chí là tổng thể.

- Cách tiếp cận tiêu chuẩn hơn để xử lý dữ liệu trước khi liên kết bản ghi, để đảm bảo rằng các nguồn dữ liệu cần liên kết được xử lý một cách nhất quán và có chất lượng cao, do đó cải thiện xác suất tạo ra các liên kết thành công. Nghiên cứu mới theo cách liên kết hồ sơ có thể được thực hiện mà không yêu cầu xác minh thủ công, nhưng vẫn kiểm soát mức độ sai khớp khi thực hiện.

- Các sáng kiến của doanh nghiệp có thể được sử dụng để phối hợp công việc, chia sẻ kiến thức trong học máy, trí tuệ nhân tạo giữa các bộ phận khác nhau. Điều này cũng đảm bảo một tập hợp các phương pháp được chấp nhận chung, tăng cường sự thống nhất của CEAG với các chương trình thống kê khác.

- Ngoài ra, việc sử dụng các công cụ và cơ sở dữ liệu phổ biến tạo điều kiện cho việc chia sẻ tài nguyên giữa các chương trình. Hiện tại có khoảng 750 người sử dụng các công cụ và hệ thống IBSP tại Statcan. Do tính chất chu kỳ của CEAG và nhu cầu cao trong một thời gian ngắn, đây là một yếu tố quan trọng để đảm bảo nhân viên được đào tạo tốt, luôn sẵn sàng làm việc khi cần.

5. Thảo luận và kết luận

APS phải áp dụng các phương pháp tiếp cận tiên tiến, hiện đại, hỗ trợ tích hợp nhiều nguồn dữ liệu thay thế trong mô hình sản xuất để tận dụng lợi thế các nguồn dữ liệu mới. Điều cần thiết là phải đáp ứng đầy đủ và kịp thời nhu cầu người dùng. Việc sử dụng phương pháp truyền thống trong việc quản lý, sử dụng bảng hỏi cho mọi trang trại để giải quyết nhu cầu thông tin ngày càng tăng là lỗi thời và phải được xem xét.

(Xem tiếp trang 23)

Tiếp theo trang 28

Sự hợp nhất và hài hòa APS với IBSP là bước đầu tiên thiết yếu để xây dựng Chiến lược tổng điều tra nông nghiệp cho tương lai. Việc sử dụng khung, khái niệm, phương pháp và công cụ chung sẽ cải thiện khả năng tích hợp các nguồn dữ liệu thay thế và hiệu quả để tạo ra những số liệu và bản phân tích vượt ra ngoài phạm vi cơ bản. Đồng thời, chương trình tổng điều tra sẽ được hưởng lợi trực tiếp từ các khoản đầu tư được thực hiện ở cấp độ doanh nghiệp trong việc phát triển và triển khai cơ sở hạ tầng, các phương pháp và công cụ, đây là một nguồn có hiệu quả đáng kể cho chương trình.

Khung quy trình hài hòa mới cùng IBSP được áp dụng với nhiều dịch vụ kinh doanh được tiêu chuẩn hóa. Số lượng lớn người tham gia vào nhiều quy trình CEAG theo IBSP làm cho việc phối hợp tất cả các hoạt động và quá trình ra quyết định phức tạp hơn một chút so với trường hợp tất cả các quy trình CEAG được quản lý đồng bộ. Một nền quản trị mạnh thực thi việc thúc đẩy các mục tiêu của doanh nghiệp đối với các ưu tiên địa phương, các thỏa thuận cấp độ dịch vụ được xác định rõ, hiệu quả và chất lượng cho các trung tâm dịch vụ khác nhau là chìa khóa thành công.

Thái Học (lược dịch)

Nguồn: The 62nd ISI World Statistics Congress, Special Topic Session (STS) - Volume 1, pp 81-88