

Statistical Methodology Research and Development Program Achievements, 2025-2026

Release date: October 8, 2026



Statistics
Canada

Statistique
Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

Email at infostats@statcan.gc.ca

Telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

- | | |
|---|----------------|
| • Statistical Information Service | 1-800-263-1136 |
| • National telecommunications device for the hearing impaired | 1-800-363-7629 |
| • Fax line | 1-514-283-9350 |

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, the Agency has developed standards of service which its employees observe in serving its clients. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" > "[Standards of service to the public](#)."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Published by authority of the Minister responsible for Statistics Canada

© His Majesty the King in Right of Canada, as represented by the Minister of Industry, 2026

Use of this publication is governed by the Statistics Canada [Open Licence Agreement](#).

An [HTML version](#) is also available.

Cette publication est aussi disponible en français.

Statistical Methodology Research and Development Program

Achievements, 2025-2026

This report summarizes the 2025-2026 achievements of the Methodology Research and Development Program (MRDP) sponsored by the Modern Statistical Methods and Data Science Branch at Statistics Canada. This program covers research and development activities in statistical and data science methods with potentially broad application in the agency's statistical programs; these activities would otherwise be less likely to be carried out during the provision of regular methodology services to those programs. The MRDP also includes activities that provide client support in the application of past successful developments in order to promote the use of the results of research and development work. Contact names are provided for obtaining more information on any of the projects described. For more information on the MRDP as a whole, please contact:

Jean-François Beaumont

(Email : jean-francois.beaumont@statcan.gc.ca)

Statistical Methodology Research and Development Program

Achievements, 2025-2026

Table of Contents

1	Data integration	3
1.1	Integration of probability and non-probability samples	3
1.2	Record linkage	4
1.3	Small area estimation	6
2	Data science methods and applications	10
3	Estimation issues in surveys.....	17
4	Confidentiality and Access	24
5	Support (Resource Centres)	26
5.1	Time Series Research and Analysis Centre	26
5.2	Resource Centre for Economic Statistical Tools and Innovation.....	29
5.3	Record Linkage Resource Centre.....	31
5.4	Data Analysis Resource Centre	32
5.5	Centre for Confidentiality and Access	33
5.6	Questionnaire Design Resource Centre.....	34
5.7	Data Ethics Secretariat	34
5.8	Quality Secretariat.....	35
6	Other activities.....	37
6.1	Survey Methodology Journal.....	37
6.2	Knowledge Transfer – Statistical Training	38
6.3	Statistics Canada’s International Methodology Symposium.....	38
7	Research papers sponsored by the Methodology Research and Development Program.....	38

1 Data integration

1.1 Integration of probability and non-probability samples

PROJECT: Prediction inference for finite population totals or means with no or limited probability survey data

For a number of years, National Statistical Organizations (NSOs) in several countries have sought to reduce data collection costs of probability surveys and the burden on survey respondents by leveraging alternative data sources and modern prediction techniques. This has led some NSOs to investigate the use of non-probability samples for the production of official statistics. At Statistics Canada, an avenue that is currently being explored is to decrease the frequency of probability surveys and substitute missing survey data with model-based predictions. For instance, an annual probability survey could be conducted every other year, and missing survey data in non-survey years could be filled in with predictions provided that auxiliary data are available in the non-survey years along with training data (e.g., historical data or current non-probability sample data).

In this context, our research focused on investigating the following two questions:

- 1) Under what methods and assumptions can approximately unbiased predictors of finite population totals or means be obtained when the variables of interest are not collected through a probability survey?
- 2) How to evaluate the quality of these predictors? More specifically, how to estimate their prediction variance?

Progress:

In previous work, we developed theory, including estimators of the prediction variance, when missing survey data are filled in with predictions based on a linear working model or nearest-neighbour predictions. This year, we completed a simulation study that compares these prediction methods along with random forest predictions and wrote a paper (Beaumont, Bocci, Bosa, Dasylva and Maranda, 2026) that was presented at the 2025 Italian Conference on Survey Methodology in July 2025.

In our paper, we also briefly review the case where auxiliary data are available for the entire population along with survey data in a small probability sample. The availability of survey data allows us to make valid inferences without relying on model assumptions. A new approach, called prediction-powered inference, has recently been proposed to handle this case when observations are independent and identically distributed. It consists in eliminating the bias of machine learning predictions by leveraging the observed survey data. Prediction-powered inference is shown to be essentially equivalent to the well-known model-assisted approach.

For more information, please contact:

Jean-François Beaumont (jean-francois.beaumont@statcan.gc.ca).

Reference

Beaumont, J.-F., Bocci, C., Bosa, K., Dasylva, A. and Maranda, G. (2026). Prediction inference for finite population totals or means. Submitted for publication in a peer-reviewed journal.

1.2 Record linkage

PROJECT: Joint modeling of linkage errors and duplication for the Canadian census of population

In the Canadian Census of Population, duplication occurs when some individual is counted multiple times. While this duplication is a source of bias, it is challenging to detect without a unique identifier. So far, this challenge has been addressed by the Census Overcoverage Study (COS), which consists in linking census records among themselves and conducting costly clerical (i.e., manual) reviews to verify potential instances of duplication that are based on small groups of linked records (Statistics Canada, 2024). The COS may be improved (better quality) and expedited (i.e., less costs) if it is possible to evaluate the duplication and linkage errors with an adequate statistical model and without clerical reviews. In previous work, Dasylva and Goussanou (2022) proposed such a model, which was constrained by the assumption that each person generates at most two census records, i.e., allowing only single or double enumerations. However, the number of higher order enumerations is not negligible in the census.

Progress:

The methodology has been generalized and improved to allow any choice of the maximum number of enumerations per person and to provide point and interval estimates for the duplication and rates of linkage error (Dasylva and Goussanou, 2026). It provides encouraging results in simulations and tests on past census data.

For more information, please contact:

Abel Dasylva (abel.dasylva@statcan.gc.ca).

References

Dasylva, A. and Goussanou, A. (2022). [A new model for the automated identification of duplicate records](#). SSC Annual Meeting, Proceedings of the Survey Methods Section. Available at: https://ssc.ca/sites/default/files/imce/dasylva_ssc2022.pdf.

Dasylva, A. and Goussanou, A. (2026). A joint model of record duplication and linkage error for the Canadian Census of Population. Internal report, Statistics Canada, Ottawa.

Statistics Canada (2024). [Coverage Technical Report, Census of Population, 2021](#). Catalogue No. 98-303-X, Issue 2021001. Available at: <https://www12.statcan.gc.ca/census-recensement/2021/ref/98-303/98-303-x2021001-eng.pdf>.

PROJECT: A differentially private method for record linkage

Statistics Canada could produce many valuable statistics if it could link its microdata to that held by other organizations, without requiring that the latter fully disclose their microdata based on a data sharing

agreement, which may be hard to negotiate. This requires novel solutions for linking and performing calculations with microdata while it is being protected through perturbation or special forms of encryption such as homomorphic encryption. While many such solutions exist, to date none of them have been shown to fully satisfy the stringent requirements of differential privacy, which has become the de-facto gold standard for the control of statistical disclosure (Dwork and Roth, 2014). Indeed, each one of these solutions incorporates at least one step where confidentiality is entirely based on encryption, which is known to be insufficient for differential privacy (Escobar, Canard, Laguillaumie and Phan, 2024). Addressing this current limitation may ease the adoption of such technologies by national statistical organizations, who have been reluctant to send their microdata beyond their premises without a data sharing agreement, regardless of whether encryption is used in the process.

Progress:

A new fully differentially private procedure was developed (Dasylva, 2025) to estimate a finite population total by linking two data sources, which are separately held by two statistical organizations, and computing a total over the resulting links. In this procedure, which draws its inspiration from Inan, Kantarcioglu, Ghinita and Bertino (2010), each organisation separately hashes its records into a finite number of bins. Then, the first organization computes a total for each bin according to the records selected therein, and it shares a perturbed version of the result with its peer after adding some noise according to negotiated privacy loss parameters. The perturbed bin totals are provided to the peer organization in a dataset of microaggregates (one for each bin), which is similar to public use microdata as far as confidentiality is concerned. Based on this information, the second organization estimates the desired finite population total, and it publishes a perturbed version of this estimate after adding more noise also according to the negotiated privacy loss parameters. When compared to previous work (e.g., Dasylva, De Cubellis, De Fausti and Franssen, 2025; Dasylva, Santos, Franssen, De Cubellis, De Fausti, Pappagallo, Berrios and Fitzsimons, 2025), the proposed procedure may provide less utility and less flexibility regarding the statistics that may be derived from the microdata. However, a simulation study has shown that a high utility can be achieved with a moderate privacy loss (Dasylva and Goussanou, 2026). Also, the new procedure has many important advantages over existing solutions such as its low cost (no specialized hardware), scalability to datasets with tens of millions of records, and a stronger protection for the confidentiality, because all statistical outputs and exchanged microdata are differentially private. Furthermore, the procedure represents a practical example of multiparty computation since no trust is required between the parties, and no trusted third party is involved.

For more information, please contact:

Abel Dasylva (abel.dasylva@statcan.gc.ca).

References

- Dasylva, A. (2025). A differentially private record linkage for the production of official statistics. Paper presented at the Advisory Committee on Statistical Methods, October 2025, Statistics Canada.
- Dasylva, A., De Cubellis, M., De Fausti, F. and Franssen, L. (2025). Linking trade data from different national statistical offices through a private set intersection. *Journal of Official Statistics*, special issue, 1-29. Available at: <https://journals.sagepub.com/doi/10.1177/0282423X251329407>.
- Dasylva, A. and Goussanou, A. (2026). A joint model of record duplication and linkage error for the Canadian Census of Population. Internal report, Statistics Canada, Ottawa.

Dasilva, A., Santos, B., Franssen, L., De Cubellis, M., De Fausti, F., Pappagallo, A., Berrios, N. and Fitzsimons, J. (2025). Private linkage of international trade microdata in a cloud-based secure enclave. Internal report.

Dwork, C. and Roth, A. (2014). The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, 9, 211-407.

Escobar, F., Canard, S., Laguillaumie, F. and Phan, D.-H. (2024). [Computational differential privacy for encrypted databases supporting linear queries](https://petsymposium.org/popets/2024/popets-2024-0131.pdf). *Proceedings on Privacy Enhancing Technologies*, 2024(4), 583-604. Available at: <https://petsymposium.org/popets/2024/popets-2024-0131.pdf>.

Inan, A., Kantarcioglu, M., Ghinita, G. and Bertino, E. (2010). Private record matching using differential privacy. *Proceedings of the 13th International Conference on Extending Database Technology*, 123-134. Available at: <https://doi.org/10.1145/1739041.1739059>.

1.3 Small area estimation

PROJECT: Use of random forests in small area estimation

When domain sample sizes are small, design-consistent direct estimators of population parameters are usually unstable. To improve the accuracy of direct estimators, the area-level Fay-Herriot model is often used. It includes two components: a sampling model and a linking model. The latter specifies the relationship between the population parameters of interest and the auxiliary variables available at the area level. In its initial form, the Fay-Herriot model assumes a linear linking model with constant error variance. It also requires estimating the smoothed sampling variance of the direct estimators, that is, the expectation with respect to the model of the sampling variance of the direct estimators.

The underlying assumptions of the Fay-Herriot and smoothing models are not always met in practice, and it can be difficult and time-consuming to adequately correct the models. In this context, it may be desirable to have access to non-parametric methods, especially when the number of areas is large, because they rely less heavily on the validity of model assumptions and can speed up the production of estimates for small areas. We are particularly interested in random forests for three reasons: i) they can be easily applied when combining categorical and continuous auxiliary variables; ii) they do not require specifying interactions; and iii) they produce predictions that always remain within the range of observed values.

First, we developed and evaluated, through empirical studies, a non-parametric version of the empirical best predictor when random forests are used to replace fully parametric models. Subsequently, we used “out-of-bag” predictions as an auxiliary variable in a Fay-Herriot linear model. Our results show that random forests provide robustness against model misspecification, increase the efficiency of small-area estimates, and simplify (though without eliminating) the modeling effort.

In 2024-2025, this project was presented at the International Conference on Establishment Statistics in Glasgow in June 2024 and at the INCASS conference in Ottawa in July 2024. We also wrote a paper (Bosa, Beaumont, Bocci and Sombo, 2025) which was presented to the Advisory Committee on Statistical Methods of Statistics Canada in June 2025.

Progress:

We previously developed a bootstrap approach to estimate the mean squared error when predictions from random forests are used as an auxiliary variable. Over the past year, R functions have been developed to implement the proposed methodology, including the bootstrap method. We have also reprogrammed the basic Fay-Herriot methodology, which has reduced execution time compared to existing R and SAS functions. A beta version of the functions is available for Statistics Canada employees.

For more information, please contact:

Jean-François Beaumont (jean-francois.beaumont@statcan.gc.ca) or

Keven Bosa (keven.bosa@statcan.gc.ca).

Reference

Bosa, K., Beaumont, J.-F., Bocci, C. and Sombo, S. (2025). The use of a random forest algorithm in small area estimation. Paper presented at the Advisory Committee on Statistical Methods, June 2025, Statistics Canada.

PROJECT: Reverse-engineering a hypothetical raking to obtain the covariance matrix of direct estimates in the Fay-Herriot model

The basic area-level Fay-Herriot (FH) small area estimation (SAE) model assumes independence of the area direct estimates. It was observed in the Labour Force Survey (LFS) context for SAE of total employment for a mapping of the provinces that direct estimates are very negatively correlated. This was made evident because a raking process is applied to the small area estimates and the correlations greatly affect quality indicators of those estimates. This project focuses on the extension of the basic FH model to the case of a strong dependence. The aim is to obtain an improved estimate of the design variance matrix of the direct estimates, which is assumed to be known in the basic as well as the extended FH models. To achieve this, a hypothetical raking process is reverse engineered to obtain the covariance terms given the (smoothed) variance terms and a variance matrix of aggregated direct estimates (the provincial total estimates in the case of the LFS). In addition, both unraked and raked small area estimates are studied.

Progress:

In 2025-2026, the work done in 2024-2025 was presented at the 2025 conference of the Statistical Society of Canada and published in the conference's proceedings (Verret and Walker, 2025). This covered the development and implementation of the full variance matrix in the extended FH model, including the extension of many features of Statistics Canada's generalised system software, G-Est (backward variable selection, diagnostics, outlier detection). Furthermore, this methodology was used starting in July 2025 to publish the revised version of the LFS small area estimates rebased for the 2021 Census. The programs were converted from SAS to R later during the year. To further validate the method, an extensive Monte Carlo (MC) simulation study was started where the 2016 Census long-form data was used as the target population and LFS was simulated with a simplified design (simple random sampling) and no coverage nor non-response errors. The theoretical variance matrices of the Horvitz-Thompson and GREG estimators were derived and first compared to the corresponding MC matrices. They were then compared to the matrix obtained from the reverse engineering method and to a compromise matrix described in Isaki, Tsay and Fuller (2000). The simulations revealed potential areas of improvement for the covariance terms, by

better integrating the sampling dependence structure and the weight calibration to the population total of the largest areas. Future simulations are planned to address these needs and to test the impact of the various variance matrices on SAE with the extended FH model.

For more information, please contact:

François Verret (francois.verret@statcan.gc.ca) or

Braedan Walker (braedan.walker@statcan.gc.ca).

References

Isaki, C.T., Tsay, J.H. and Fuller, W.A. (2000). [Estimation of census adjustment factor](#). *Survey Methodology*, 26(1), 31-42. Available at: <https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2000001/article/5176-eng.pdf>.

Verret, F. and Walker, B. (2025). Application of the Fay-Herriot model with correlated design errors: The Labour Force Survey experience. Proceedings of the 2025 SSC Annual Meeting, Statistical Society of Canada, Saskatoon, Canada.

PROJECT: Small area estimation with benchmarking: evaluation and simulation study

In this project, we evaluate and study the Empirical Best Linear Unbiased Predictor (EBLUP) and benchmarked EBLUP estimators for small area estimation. We consider the self-benchmarked EBLUPs of Wang, Fuller and Qu (2008) and You, Rao and Hidirolou (2013) under the Fay-Herriot model with correct modeling and incorrect modeling. We apply the EBLUP estimators to the LFS unemployment rate data and compare the model-based estimates with the census values. We also conduct simulation studies to evaluate and compare the EBLUP and benchmarked EBLUPs under correct and incorrect modeling. The main point of this project is to evaluate the EBLUP and benchmarked EBLUP under the use of correct and incorrect modeling.

Progress:

We have conducted data analysis and simulation study. Our results show that both the WFQ (Wang, Fuller and Qu, 2008) and YRH (You, Rao and Hidirolou, 2013) benchmarked EBLUPs perform very similarly. However, the benchmarking procedure does not improve the EBLUP under incorrect modeling or model misspecification. A research paper, You (2026), has been written and submitted to a journal for publication.

For more information, please contact:

Yong You (yong.you@statcan.gc.ca).

References

Wang, J., Fuller, W.A. and Qu, Y. (2008). [Small area estimation under a restriction](#). *Survey Methodology*, 34(1), 29-36. Available at: <https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2008001/article/10619-eng.pdf>.

You, Y. (2026). Comparison and evaluation of small area EBLUP estimators under benchmarking with application and simulation study. Submitted to a peer-reviewed journal for possible publication.

You, Y., Rao, J.N.K. and Hidioglou, M.A. (2013). [On the performance of self benchmarked small area estimators under the Fay-Herriot area level model](https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2013001/article/11830-eng.pdf). *Survey Methodology*, 39(1), 217-229. Available at: <https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2013001/article/11830-eng.pdf>.

PROJECT: Empirical study, evaluation of borrowing effects and constructing informative priors in small area estimation

Hierarchical Bayes (HB) modeling is very popular in small area estimation, and prior specification is very important in the HB modeling approach. In this project, we study the performance of HB small area estimators using noninformative and informative priors for the regression parameters and variance components based on the Bayesian models of You and Chapman (2006) and You (2021) with application to the Canadian Labor Force Survey (LFS) data. We study the impact of correct and incorrect informative priors and non-informative priors for HB small area estimation using both a LFS application and a simulation study.

Progress:

We have published a paper, You and Bosa (2025), based on this project. Following the work of You and Bosa (2025), we propose an iterative prior specification approach for the use of previous month of data. Then we evaluate and compare the HB estimates using noninformative and informative priors through an empirical study. Our results indicate that informative priors can perform very well and the proposed iterative prior specification approach can lead to stable and better results in terms of smaller bias and coefficient of variation (CV). A draft research paper (You and Bosa, 2026) has been written.

For more information, please contact:

Yong You (yong.you@statcan.gc.ca).

References

You, Y. (2021). [Small area estimation using Fay-Herriot area level model with sampling variance smoothing and modeling](https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2021002/article/00007-eng.pdf). *Survey Methodology*, 47(2), 361-370. Available at: <https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2021002/article/00007-eng.pdf>.

You, Y. and Bosa, K. (2025). [Performance of hierarchical Bayes small area estimators using noninformative and informative priors with an application to the Canadian Labor Force Survey](https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2025002/article/00010-eng.pdf). *Survey Methodology*, 51(2), 425-448. Available at: <https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2025002/article/00010-eng.pdf>.

You, Y. and Bosa, K. (2026). An empirical study on the use of informative priors for the hierarchical Bayes small area estimation with application. Research paper in progress to be submitted for possible publication.

You, Y. and Chapman, B. (2006). [Small area estimation using area level models and estimated sampling variances](https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2006001/article/9263-eng.pdf). *Survey Methodology*, 32(1), 97-103. Available at: <https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2006001/article/9263-eng.pdf>.

PROJECT: Small area estimation for producer price indices

The goal of this project was to evaluate small area estimation (SAE) methods for producing more granular producer price index estimates in industry and commodity groups and to improve the quality of current estimates in the face of low response rates and budget constraints. We considered the cross-sectional Fay-Herriot model (FH), the Rao-Yu model (RY), and a multi-level model with a time series term.

Progress:

In the 2025-2026 year, we adapted established SAE theory towards estimating a Lowe or Young price index. We showed that the Empirical Best Linear Unbiased Predictor (EBLUP) under the FH model produces a transitive price index under certain conditions, but that the EBLUP of the RY model and Hierarchical Bayes estimates for the multi-level model do not have desired properties for price indices. We modelled the monthly Retail Services Price Index using a large number of covariate time series including broad market indicators, industry characteristics, and other price indices. Exploratory analysis and model selection unfortunately showed a lack of good predictors. Monthly SAE estimates were produced but had high mean square error and did not seem of sufficient quality to implement. This underscores the importance of good price collection as modelling cannot easily replace price data.

For more information, please contact:

Justin Francis (justin.francis@statcan.gc.ca).

2 Data science methods and applications

PROJECT: Using parameter efficient fine tuning on small language models to address statistical coding classification and factuality checking.

This project explores the potential of smaller, open-weight transformer language models for specialized statistical coding classification tasks. By using open-source models, we investigated whether parameter-efficient fine-tuning (PEFT) on curated datasets could enable a compact model to learn complex reasoning and decision patterns. The objective was to demonstrate that smaller, resource-efficient models can achieve comparable performance to larger proprietary models while facilitating cost-effective deployment.

Progress:

Two separate use cases were conducted: fine tuning pre-trained embedding and re-ranking models as a first step in the statistical coding of the North American Product Classification System codes; and fine-tuning smaller open-source Large Language Models on the reasoning of a GPT-5 teacher model, to perform National Occupational Classification rebasing from 2016 to 2021 for the Labour Force Survey.

The first use case proved inconclusive, with the fine-tuned models underperforming the base models. One challenge was the noisy data as well as the lack of proper graphics processing unit (GPU) resources to perform fine-tuning on sufficient data. Also, fine tuning embedding models is less documented and highly dependent on the formatting and quality of the data, making the task itself more difficult.

For the second use case, we managed to implement PEFT and QLoRA (Quantized Low-Rank Adapters) on a A100 GPU. We created training data leveraging the write-ins, codes definitions and the responses from the teacher model, enabling the learner model to mimic the reasoning and classification accuracy of the

teacher. Fine-tuning was performed on 10,000 and 30,000 examples, the later resulting in the open-source student surpassing the proprietary teacher model in average weighted F1 score (+1.4% - 79.93 vs 78.56). One possible explanation, which warrants future investigation, is that by training the student exclusively on the teacher's successful reasoning extracts, the student model may have learned the more effective assessment patterns, ultimately resulting in a slightly higher score. Beyond the modest score improvement, having a capable open-source model provides greater control over deployment and access to features such as the raw log probabilities, which were unavailable with the GPT-5 mini model. This access enables additional capabilities, such as uncertainty quantification based on these log probabilities.

For more information, please contact:

Alexandre Istrate (alexandre.istrate@statcan.gc.ca).

PROJECT: Enhancing text classification with conformal prediction

Large-scale text classification systems often require assigning text responses to one of many possible categories while also providing reliable measures of prediction uncertainty. This project develops uncertainty-aware classification methods using conformal prediction, moving from initial small-scale experiments (25 classes) to systems with a few thousand classes. This approach enables reliable uncertainty quantification for large language model-based coding of open-text survey responses. Benchmarked on North American Product Classification System (NAPCS) 3-digit and 7-digit datasets, this work aims to explicitly capture prediction uncertainty to support greater automation, reduce manual review, and improve trust in large-scale text classification systems.

Progress:

As part of this research project, a logit-based conformal prediction framework was developed and evaluated for large-class classification problems. An end-to-end experimental pipeline was implemented to extract model outputs, generate conformal prediction sets, and conduct large-scale experiments. Initial testing was performed using a Simple Math dataset (a 125-question synthetic dataset of basic arithmetic problems grouped into 25 answer-based classes) to validate the methodology in a controlled environment and to study the impact of model output constraints on prediction reliability. The approach was subsequently evaluated using datasets derived from the NAPCS, including both broad product-group (three-digit) classification and detailed seven-digit product-code classification tasks.

Results showed that the underlying models generally performed well when ranking candidate classes within smaller subsets but faced challenges when predictions were aggregated across very large class spaces. When combined with conformal prediction, the framework produced promising results for broader classification tasks; however, in fine-grained settings involving thousands of classes, it yielded prediction sets that were too large to be operationally useful. These findings indicate that the effectiveness of conformal prediction in large-class settings depends strongly on the performance and calibration of the underlying language model. Further research is planned to evaluate more capable models and to assess whether model logits can provide a reliable basis for conformal prediction in large-scale classification applications.

For more information, please contact:

Mihir Gajjar (mihir.gajjar@statcan.gc.ca).

PROJECT: Benchmarking LLM-Based coding for large-scale survey data

This project evaluates the use of large language models (LLM) to automate the coding of survey responses into standardized classification systems (such as occupation and industry codes) compared with existing machine learning models such as fastText and DistilBERT. The main objective is to support Statistics Canada's goal of modernizing and scaling coding processes for large volumes of unstructured data.

A standardized benchmarking framework was created, including a gold standard dataset, a unified evaluation pipeline, and consistent performance metrics (e.g., automation rate, F1 scores, and error rates). The project also developed an LLM-based coding pipeline with data preprocessing, candidate code retrieval and verification, and calibrated confidence scoring to enable fair comparisons with traditional trained models.

Progress:

The framework was applied to and tested with data from the Labour Force Survey and Job Vacancy and Wage Survey. Results show that traditional machine learning (ML) models (fastText and DistilBERT) outperform LLMs in both accuracy and automation rates, particularly at higher confidence thresholds. These findings highlight current limitations of pre-trained LLMs for coding tasks that rely on external information sources and subject matter expertise, as well as identifying challenges in generating calibrated confidence scores for LLM-based predictions.

Phase 2 research will focus on strengthening ML-based approaches and refining benchmarking strategies while assessing impacts on survey outputs in collaboration with subject matter experts.

For more information, please contact:

Soufiane Fadel (soufiane.fadel@statcan.gc.ca).

PROJECT: Prediction-powered estimation for official statistics

National statistical organizations increasingly face declining response rates, shrinking sample sizes, and budget constraints, while simultaneously gaining access to rich auxiliary data sources such as administrative records, images, and text. At the same time, advances in machine learning (ML), artificial intelligence (AI), and Large Language Models (LLMs) provide powerful predictive capabilities that can potentially improve statistical estimation. However, traditional model-assisted survey estimators typically require models to be fitted using survey sample data and therefore cannot fully leverage large pre-trained black-box models trained on external datasets.

This research revisits the use of pre-trained models for design-based survey estimation and introduces a new framework for integrating AI into official statistics while preserving design-based properties. The paper proposes the Prediction-Powered Estimator (PPE) together with the Prediction-Powered Difference (PPD) estimator and an adaptation of the Generalized Regression Estimator (GREG) using model predictions as auxiliary information. The proposed methods enable the use of pre-trained ML models and LLMs without requiring re-training on survey data, thereby expanding the survey methodologist's toolbox to include modern AI systems and novel data sources.

Progress:

This research culminated in a peer-reviewed publication (Denis and Haddou, 2026) demonstrating how pre-trained ML and AI models can be incorporated into design-based estimation while maintaining unbiased estimation and valid variance estimation. The study evaluated the proposed estimators using real-world survey data from Statistics Canada, complemented by simulation experiments and innovative LLM-based applications involving unstructured data such as text and images. Across multiple performance metrics—including bias, mean squared error, confidence interval coverage, and interval width—the proposed approaches consistently outperformed standard baseline estimators.

The research contributes to the emerging field of Inferential AI by providing principled methods for integrating frontier AI models into official statistics. The results suggest that pre-trained models can improve estimation efficiency, potentially reduce required survey sample sizes, and lower respondent burden while preserving statistical rigor. The work also provides a foundation for future applications of AI-enhanced inference within national statistical organizations.

For more information, please contact:

Mohammed Haddou (mohammed.haddou@statcan.gc.ca).

Reference

Denis, N. and Haddou, M. (2026). Prediction-powered estimation: Unbiased model-assisted estimation. *Journal of Official Statistics*. Available at: <https://doi.org/10.1177/0282423X261451320>.

PROJECT: UNECE handbook on uncertainty quantification for AI/ML-enhanced statistical inference

As national statistical Organizations (NSOs) increasingly adopt machine learning (ML) and artificial intelligence (AI) to improve the efficiency and scope of statistical production, ensuring the statistical validity and trustworthiness of AI-enabled outputs has become a critical challenge. While ML methods can substantially enhance prediction accuracy and operational efficiency, they often lack the uncertainty quantification and inferential guarantees traditionally required for official statistics. This creates an important methodological question: how can NSOs leverage modern AI while preserving the rigor and public trust that underpin official statistics?

To address this challenge, Statistics Canada led the development of the United Nations Economic Commission for Europe (UNECE) Handbook on Uncertainty Quantification for AI/ML-Enhanced Statistical Inference through the High-Level Group for the Modernisation of Official Statistics (HLG-MOS) ModernStats initiative, in collaboration with several national statistical offices. The handbook presents principled frameworks for integrating modern AI methods into official statistics while maintaining inferential validity. In addition to reviewing traditional uncertainty quantification approaches, it provides comprehensive guidance on two complementary frameworks: conformal prediction, which delivers finite-sample, distribution-free uncertainty guarantees for predictions, and prediction-powered inference, which enables the use of pretrained models while preserving design-based validity.

Progress:

The handbook was completed and published by the United Nations Economic Commission for Europe (UNECE, 2026). Statistics Canada led the international task team responsible for developing the handbook and coordinating contributions from multiple NSOs. The handbook includes methodological guidance, implementation considerations, decision frameworks, and operational case studies from several national statistical offices, demonstrating practical applications of uncertainty quantification methods in production environments. Case studies cover automated coding, economic classification, and prevalence estimation using modern AI techniques.

The handbook contributes to the emerging field of Inferential AI by demonstrating that modern AI methods can be integrated into existing survey methodology frameworks without compromising statistical rigor or trustworthiness. By extending classical model-assisted estimation to accommodate complex ML models and foundation models, the handbook provides practitioners with practical tools for deploying AI systems while preserving valid confidence intervals, design-based inference, and public trust in official statistics.

For more information, please contact:

Mohammed Haddou (mohammed.haddou@statcan.gc.ca).

Reference

United Nations Economic Commission for Europe (UNECE) (2026). UNECE handbook on Uncertainty quantification in ML-enhanced statistical inference - A practitioner handbook with applications in official statistics. United Nations Economic Commission for Europe (UNECE), HLG-MOS ModernStats Initiative. Available at: <https://unece.org/statistics/documents/2026/06/reports/uncertainty-quantification-ml-enhanced-statistical-inference>.

PROJECT: Validity and reliability of large language models as evaluators

The rapid adoption of large language models (LLMs) has led to their increasing use not only as generative systems but also as evaluators of machine learning outputs, commonly referred to as “LLMs-as-judges” (LLJs). These systems are now widely used in natural language generation (NLG) evaluation, benchmarking, and model comparison tasks due to their scalability and low cost. However, despite their growing popularity, fundamental questions remain regarding their statistical validity and reliability as measurement instruments. In particular, it is unclear whether LLJs provide evaluations that are consistent, unbiased, and aligned with human judgment in a way that is suitable for scientific inference and decision-making.

This research applies measurement theory from the social sciences to systematically examine the assumptions underlying the use of LLMs as evaluators. It critically investigates four core assumptions: (i) that LLJs can serve as reliable proxies for human judgment, (ii) that they function as valid evaluators of model outputs, (iii) that they scale effectively without degrading measurement quality, and (iv) that they provide meaningful cost advantages over traditional evaluation methods. The paper further analyzes three representative applications—text summarization, data annotation, and safety alignment—to illustrate how these assumptions may fail in practice and how such failures can affect downstream evaluation outcomes.

Progress:

This research was published in the NeurIPS 2026 proceedings (Chehbouni, Haddou, Cheung and Farnadi, 2026) following acceptance for presentation at the conference. The work provides a structured critique of the emerging paradigm of LLM-based evaluation and situates it within a formal measurement theory framework. The findings highlight that while LLJs are attractive due to scalability and efficiency, their widespread adoption may be premature in the absence of rigorous validation of their reliability and construct validity.

The paper contributes to the broader agenda of responsible AI evaluation by identifying key limitations in current LLM-based assessment practices and emphasizing the importance of measurement validity in machine learning evaluation pipelines. It supports the development of more robust evaluation standards for NLG systems and underscores the need for methodological rigor when deploying LLMs in roles beyond generation, particularly in contexts where evaluation results inform research priorities, benchmarking, and policy-relevant decisions.

For more information, please contact:

Mohammed Haddou (mohammed.haddou@statcan.gc.ca).

Reference

Chehbouni, K., Haddou, M., Cheung, J.C.K. and Farnadi, G. (2026). Neither valid nor reliable? Investigating the use of LLMs as judges. Proceedings of NeurIPS 2025. Available at: https://papers.nips.cc/paper_files/paper/2025/hash/829e8f32d76a6248815bf5b01633811d-Abstract-Position_Paper_Track.html.

PROJECT: Responsible AI framework for official statistics

The rapid advancement of Artificial Intelligence (AI) has led to its increasing adoption across National Statistical Organizations (NSOs), enabling new capabilities in data processing, analysis, and dissemination. However, alongside these opportunities, AI introduces significant risks related to bias, privacy, transparency, accountability, and public trust. For NSOs, which operate under strict mandates to produce trustworthy official statistics for the public good, these risks necessitate a structured and principled approach to AI governance. A key challenge is therefore to ensure that AI systems are deployed in a way that is ethical, transparent, and aligned with the Fundamental Principles of Official Statistics.

This research contributed to the development of a United Nations Economic Commission for Europe (UNECE) framework on Responsible AI for Official Statistics, produced under the High-Level Group for the Modernisation of Official Statistics (HLG-MOS) initiative. The framework synthesizes international guidance from statistical agencies and organizations such as Statistics Canada, the Organisation for Economic Co-operation and Development, and the European Commission, and establishes a set of guiding principles for the responsible design, deployment, and monitoring of AI systems in NSOs. These principles include ethical purpose and public good, accountability, transparency and explainability, fairness and inclusiveness, privacy and security, and validity and robustness. The framework provides both conceptual foundations and operational guidance for integrating AI into statistical production systems while maintaining institutional trust.

Progress:

This work was published as a UNECE report under the HLG-MOS ModernStats initiative (UNECE, 2025). The framework was developed collaboratively with international partners and builds on existing governance approaches while tailoring them specifically to the needs and constraints of official statistics. It provides practical guidance for NSOs on how to operationalize responsible AI principles throughout the statistical production lifecycle, from model development and validation to deployment and monitoring.

The framework strengthens the governance foundation for AI adoption in official statistics by ensuring that innovation is balanced with accountability and public trust. It emphasizes that AI systems must be not only technically robust but also socially and institutionally aligned with the mandate of NSOs. The work contributes to establishing shared international standards for responsible AI in official statistics and supports the safe and effective integration of AI technologies into production environments.

For more information, please contact:

Mohammed Haddou (mohammed.haddou@statcan.gc.ca).

Reference

United Nations Economic Commission for Europe (UNECE) (2025). [Framework on Responsible AI for Official Statistics](https://unece.org/sites/default/files/2025-10/Responsible%20AI%20for%20Official%20Statistics%20Framework_ADSaMM%20Group_HLG-MOS_0.pdf). Available at: https://unece.org/sites/default/files/2025-10/Responsible%20AI%20for%20Official%20Statistics%20Framework_ADSaMM%20Group_HLG-MOS_0.pdf.

PROJECT: Fairness and bias in machine learning for official statistics

As Artificial Intelligence (AI) and Machine Learning (ML) systems are increasingly integrated into statistical production and decision-support systems, concerns about fairness and bias have become central to their responsible deployment. Bias in AI systems can arise in multiple forms and at different stages of the machine learning lifecycle and may lead to systematically unequal outcomes across population groups. In the context of official statistics, where outputs inform public policy and resource allocation, understanding and mitigating bias is essential to ensure fairness, transparency, and public trust.

This work contributed to a United Nations Economic Commission for Europe (UNECE) companion paper to the Responsible AI Framework for Official Statistics, focusing specifically on the concept of fairness and the different forms of bias in ML systems. The paper clarifies how “bias” is defined across disciplines, distinguishing statistical notions of bias from legal and ethical interpretations related to discrimination and harm. It further develops a structured taxonomy of bias in ML systems, including historical bias, representation bias, algorithmic processing bias, deployment bias, and feedback loop bias. The work emphasizes that these forms of bias are often interrelated and can jointly influence outcomes in complex sociotechnical systems.

Progress:

This work was completed and published as a UNECE companion paper under the High-Level Group for the Modernisation of Official Statistics (HLG-MOS) ModernStats initiative (Haddou, Seimandi and Del Monaco, 2025). The paper complements the broader Responsible AI framework by providing a focused conceptual and practical treatment of fairness and bias in machine learning applications for official statistics. It synthesizes insights from statistics, computer science, law, and ethics to support a

shared understanding of bias across disciplines and to inform mitigation strategies in operational environments.

The contribution strengthens the conceptual foundations for responsible AI in official statistics by providing a clear and structured classification of bias types and their implications for statistical systems. It supports national statistical organizations in identifying potential sources of unfairness in AI/ML pipelines and reinforces the importance of considering sociotechnical context when deploying machine learning models in high-stakes environments. This work contributes to the development of harmonized international guidance on fairness in AI for official statistics.

For more information, please contact:

Mohammed Haddou (mohammed.haddou@statcan.gc.ca).

Reference

Haddou, M., Seimandi, T. and Del Monaco, A. (2025). [Fairness in Machine Learning: Companion paper to the Framework on Responsible AI for Official Statistics](https://unece.org/sites/default/files/2025-10/Companion%20Paper%20on%20Fairness%20in%20Machine%20Learning_Responsible%20AI%20Framework_ADSaMM%20group.pdf). Available at: https://unece.org/sites/default/files/2025-10/Companion%20Paper%20on%20Fairness%20in%20Machine%20Learning_Responsible%20AI%20Framework_ADSaMM%20group.pdf.

3 Estimation issues in surveys

PROJECT: Degrees of freedom related to long-form census questionnaire estimation

Statistics Canada now publishes confidence intervals to express the quality of the census estimates. The length of a confidence interval can testify to the quality of an estimate as long as the stated coverage is respected. A parameter that plays an important role in calculating confidence intervals is the number of degrees of freedom. In practice, this value is generally determined using a rule of thumb. In the case of small domains, this rule of thumb often overestimates the actual number of degrees of freedom, which leads to undercoverage of the confidence intervals.

In this project, the Satterthwaite approximation is used to derive a more precise estimate of the degrees of freedom in the context of estimates obtained using data from the long-form population census questionnaire. Variance estimation is based on an adaptation of the balanced half-sample method, as described by Devin and Verret (2016). A simulation study makes it possible to evaluate the gain in the coverage of confidence intervals for the estimation of totals of continuous and dichotomous variables. The results suggest that by using a more accurate number of degrees of freedom, the coverage is enhanced and often makes it possible to reach the nominal threshold in the problematic case of small domains.

The objective of the project for the year 2025-2026 was to make the requested corrections to the article by Toupin and Martin (2026) with a view to its publication in the *Survey Methodology* journal.

Progress:

The article by Toupin and Martin (2026) has been accepted for publication in the June 2026 issue of the *Survey Methodology* journal.

For more information, please contact:

Marie-Hélène Toupin (marie-helene.toupin@statcan.gc.ca).

References

Devin, N. and Verret, F. (2016). The development of a variance estimation methodology for large-scale dissemination of quality indicators for the 2016 Canadian census long form sample. *Proceedings of the Survey Research Methods Section*, American Statistical Association, Alexandria VA.

Toupin, M.-H. and Martin, V. (2026). [Improving the coverage of confidence intervals with respect to degrees of freedom: Application to the Canadian Census](https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2026001/article/00005-eng.pdf). *Survey Methodology*, 52(1), 187-211. Available at: <https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2026001/article/00005-eng.pdf>.

PROJECT: Dependent random group variance estimation using Fay's method

Replication-based methods are widely used in survey sampling to estimate the variance of population parameter estimates. This process involves repeatedly drawing subsamples from the original sample and measuring how the resulting estimates vary. However, replication methods can become unstable when estimating variances for ratios involving rare characteristics, since such characteristics may disappear from some subsamples. To address this issue, Robert Fay proposed a smoothing approach for replicate weights in the Balanced Repeated Replication (BRR) method, ensuring that all sampled units contribute to every subsample estimate. In this project, we extended Fay's smoothing approach (Dippo, Fay and Morganstein, 1984), called Fay's method, to the Dependent Random Group (DRG) replication method, which is used to estimate variances for the Canadian Census Individuals Public Use Microdata File (PUMF).

Progress:

We modified the dependent random group replicate weighting scheme using the ideas proposed by Robert Fay for BRR. To offset the variance underestimation introduced by the smoothed replicate weights, we derived an adjustment factor and incorporated it into the replicate variance formula. We then conducted a series of simulation studies to evaluate the bias and variability of the resulting DRG variance estimator for a range of statistics using different levels of replicate weight smoothing. The simulations showed that moderate smoothing was sufficient to stabilize the variance estimation of rare ratios while maintaining good performance for other types of estimators. We also incorporated a finite population correction factor into the smoothed replicate weights which produced unbiased variance estimates in the presence of non-negligible sampling fractions. The technical details and simulation results were compiled into an article which was submitted to the *Survey Methodology* journal (Imbrogno and Verret, 2025). The methodological developments were incorporated into the Census Individuals PUMF to improve the stability of variance estimates going forward.

For more information, please contact:

Alex Imbrogno (alexander.imbrogno@statcan.gc.ca) or

François Verret (francois.verret@statcan.gc.ca).

References

Dippo, C.S., Fay, R.E. and Morganstein, D.H. (1984). Computing variances from complex samples with replicate weights. *Proceedings of the Survey Research Methods Section*, American Statistical Association, Alexandria, VA.

Imbrogno, A. and Verret, F. (2025). Dependent random group variance estimation using Fay's method. Paper submitted for publication in the *Survey Methodology* journal.

PROJECT: Bootstrap variance estimation for calibration estimators

Survey estimates are often calibrated so that they match known population totals. The Generalized REGression estimator (GREG estimator) is a popular calibration estimator that assumes a linear relationship between the survey variable and the auxiliary variables. The theory underlying this estimator is well established, and there exist linearization variance estimators. Deville and Särndal (1992) showed that, for a family of calibration estimators, all members of the family are asymptotically equivalent to the GREG estimator. This suggests using the variance estimator of the GREG estimator, or an adaptation of it, for other calibration estimators. It is this type of variance estimator, similar to the GREG case, that is generally implemented in practice.

Some calibration methods involve constraints on the range of weights, such as ridge calibration. Linearization in the presence of such constraints is not always simple, and the usual variance estimator for the GREG estimator may not be appropriate. The objective of the present research was to find an alternative variance estimator in this situation.

This project was started in 2023-24, during which we developed a bootstrap variance estimator that properly accounts for constraints on weights and uses appropriate control totals. A simulation study comparing four bootstrap approaches for estimating the variance of a ridge-calibrated estimator was completed. In this study, we considered two different populations and four sample sizes. The results of this study demonstrated the importance of using constraints and appropriate control totals in order to accurately estimate the variance and mean squared error.

Progress:

We considered a new bootstrap variance estimator where the constraints on the weights require an additional adjustment when bootstrap weights are calibrated on the population totals at each bootstrap replicate. The results show that the approach developed in 2023-24, where the bootstrap weights are calibrated on the Horvitz-Thompson estimates, is slightly more stable than the new approach. We have also started to write an article explaining the theory and containing the main results of the simulation study. Our goal is to submit this article to a statistical journal in 2026-27.

For more information, please contact:

Keven Bosa (keven.bosa@statcan.gc.ca).

Reference

Deville, J.-C. and Särndal, C.-E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, 87, 376-382.

PROJECT: Automatic weighting for nonresponse

The score method is a popular method used to adjust survey weights to counteract the effects of nonresponse in sample surveys. A key component of the score method is the specification of a nonresponse model that models each unit's propensity to respond based on available auxiliary information. In some instances, particularly in situations where there may be limited information about the relationship between the response mechanism and available covariates, effectively specifying this response model can prove challenging and time-consuming. For this reason, we investigated the use of different machine learning algorithms to replace this modeling step in the score method.

This work builds off the work of Larbi, Tsang, Haziza and Dagdoug (2025), which investigated the use of machine learning methods for weight adjustment. This project analyzes the effect of a machine-learning score method weighting strategy on survey estimates' relative bias and efficiency. In addition, other dimensions were added to the analysis, including the effect of using cross-validation to select the "best" model, model aggregation using LASSO regression and neural networks, and various means of forming response homogeneity groups (RHGs).

Progress:

This project has been completed. We conducted a simulation study to investigate our research question. We created a synthetic dataset containing a survey variable, auxiliary variables, and response indicators generated through a defined response mechanism. We also created an automated workflow that selects auxiliary variables for a response model and fits response models using 1) a logistic regression; 2) machine learning methods selected from a pool of methods based on cross-validation metrics; and 3) an aggregate method that uses all machine learning methods considered. Multiple methods of grouping units into RHGs were tested. Monte Carlo relative bias and efficiency were computed for each estimate over 1,000 replicates to compare the effectiveness of the different modeling methods.

The results of this project demonstrated that an aggregated (stacked) model may provide a suitable alternative to a logistic regression model and could potentially provide some robustness to model misspecification. Certain machine learning models also demonstrated potential for modeling response, but there are some limitations in the generalisability of this study as few nonresponse mechanisms were experimented. This project also shed a light on the impact of the method used to create RHGs has on bias correction and efficiency.

For more information, please contact:

Gavi Arganosa (gavriel.arganosa@statcan.gc.ca), **Ilya Hekimi** (ilya.hekimi@statcan.gc.ca) or **Oleksii Volkov** (oleksii.volkov@statcan.gc.ca).

Reference

Larbi, K., Tsang, J., Haziza, D. and Dagdoug, M. (2025). [On the use of machine learning methods for the treatment of unit nonresponse in surveys](https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2025001/article/00003-eng.pdf). *Survey Methodology*, 51(1), 275-303. Available at: <https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2025001/article/00003-eng.pdf>.

PROJECT: Sampling plans for social surveys integrating response propensity

Several stages of the survey process are based on the assumption that the true response propensities of a survey, p_i , are exactly or approximately well estimated by a propensity model $\hat{p}_i$ for all units i in the sample. This is the starting point for a nonresponse adjustment and several adaptive collection strategies.

It is observed that the factors influencing response are relatively stable between social surveys. This observation suggests that indicators of response propensity could also be integrated into the sampling plan, based on models developed during previous cycles and the auxiliary information available from the survey database. Since this practice is not widespread, this project explored its benefits and limitations.

Progress:

Several theoretical explorations have been conducted, such as the derivation of an optimal allocation for a stratified without replacement simple random sampling plan, with strata combined with homogeneous response groups. This allocation has also been extended to high-entropy designs.

Analyses based on the loss of efficiency function introduced by Kish (1976), similar to the design effect, were carried out to quantify the loss according to the homogeneity of the response propensity in the population or due to an inaccurate approximation of the propensity model. It follows that, for example, optimality is reached when the estimated response propensity is proportional to the true response propensity. The optimal allocation adjusted by response propensities is more efficient when the strata are both intra-homogeneous and inter-heterogeneous with respect to the response propensities.

Simulations have been conducted to measure efficiency gains according to different model scenarios and homogeneity of strata with respect to the response propensity. Efficiency gains between 5% and 10% are expected, depending on the distribution of propensities and the quality of their modeling.

This project will continue. It will aim to integrate a realistic and practical cost indicator in the allocation and to validate efficiency gains through the simulation of a realistic social survey production context.

For more information, please contact:

Gautier Gissler (gautier.gissler@statcan.gc.ca).

Reference

Kish, L. (1976). Optima and proxima in linear sample designs. *Journal of the Royal Statistical Society*, 139, Part 1, 80-95.

PROJECT: Weight trimming methods for survey data

The typical weighting process for survey data involves three major stages: (1) assigning a base weight, (2) correcting the weights for unit nonresponse, and (3) modifying the weights to ensure consistency between survey estimates and known population totals, often called calibration. In some cases, the weights undergo a final modification through weight trimming to improve the efficiency of survey estimates by limiting the variability of the weights and thus ensuring greater stability of the estimates.

Trimming procedures require the specification of a tuning constant, where weights greater than the constant are trimmed to the value of the constant, and the excess weight is redistributed among remaining units. The procedures used in practice are often ad-hoc and a formal, and theoretical treatment of weight trimming is currently lacking. The goals of this project are (1) develop and explore weight trimming procedures to handle multiple survey variables, (2) conduct a vast simulation study to compare the performance of several weight trimming procedures (see Potter, 1988, 1990; Potter and Zheng, 2015; Haziza and Beaumont, 2017) in terms of bias and efficiency and to assess the performance of mean square error estimators in terms of bias, and finally, (3) apply and demonstrate the proposed methodology on an existing dataset.

Progress:

A simulation study was created and conducted in R, making use of Statistics Canada's generalized system for calibration, Yoho, by using a maximum weight constraint during the calibration process. The study compares ten weight trimming methods in terms of bias and variance when applied to categorical and continuous survey variables with varying correlation to the weights. Results demonstrate that when correlation is absent, there is little to lose by trimming weights, and more aggressive trimming tends to deliver the largest variance reductions without compromising bias. When correlation is introduced, the methods that trim the highest number of weights no longer yield desired results, and instead, those which trim a low number of weights perform better. This difference is even more pronounced as weights are more dispersed. The minimum Mean Squared Error (MSE) method considers specific important survey variables and selects a trimming threshold that yields the lowest combined MSE among them. This method shows adaptability and provides a good compromise in terms of MSE among the chosen survey variables. Performance of methods was similar when using continuous or categorical variables with the same correlation to the weights.

The methods were applied to a dataset provided by the General Social Survey team, and the analysis is in progress. Future steps would include the completion of the real data example, comparing additional methods such as weight smoothing, and measuring the effect on variance estimation using bootstrap weights.

For more information, please contact:

Jody Krahn (jody.krahn@statcan.gc.ca) or

David Haziza (dhaziza@uottawa.ca).

References

Haziza, D. and Beaumont, J.-F. (2017). Construction of weights in surveys: A review. *Statistical Science*, 32, 206-226.

Potter, F. (1988). Survey of procedures to control extreme sampling weights. *Proceedings of the Survey Research Methods Section*, American Statistical Association, 453-458.

Potter, F. (1990). A study of procedures to identify and trim extreme sampling weights. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 225-230.

Potter, R. and Zheng, Y. (2015). Methods and issues in trimming extreme weights in sample surveys. *Proceedings of the Section on Survey Research Methods*, American Statistical Association.

PROJECT: Variance estimation under advanced imputation methods

Machine learning is useful for survey imputation because it captures complex relations well. To estimate variance due to imputation (VDI), Statistics Canada uses the System for the Estimation of Variance due to Non-response and Imputation (SEVANI) (Beaumont, Bissonnette and Bocci, 2010; Beaumont and Bissonnette, 2011). Since variance is a key indicator of data quality, this system must keep up with these new methods. To bridge this gap, this project explores options to estimate VDI when tree-based models are used.

Progress:

We considered two approaches to estimate VDI when the imputation is done using machine learning methods. The first extends SEVANI to regression trees and random forests. We fitted these models with open-source R packages *rpart* and *ranger* and unpacked the model structures to identify how much each respondent contributed to each imputed value. The second way uses the Pseudo-Population Bootstrap (PPB) (Chen, Haziza, Léger and Mashreghi, 2019), which repeats the full workflow from sampling to estimation.

For SEVANI, we used cross-fitting to limit overfitting in error estimates of respondents. Then, we used the nearest-neighbour model to estimate errors of imputed values, as in Beaumont et al. (2010). In our simulation, this reduced bias in the variance estimates when imputation models are regression trees and random forests. For PPB, rather than unpacking regression trees and Cubist, we repeated the full process across all bootstrap replicates. This approach performed well for both imputation methods in the simulation. Overall, SEVANI remains strong when we can easily recover the contributions. PPB is practical when the unpacking is complex, as with Cubist, but the computation takes longer due to bootstrapping.

The simulation study points to two next steps. First, we can study how artificial intelligence (AI) hardware, such as graphical processing units (GPUs), can help with variance estimation. Parallel and distributed computing can speed up the process and make the two methods more practical at an enterprise scale. Therefore, benefits from AI tools could spill over and strengthen classical statistical methods. We could also study ways to reduce the impact of poor imputation models on inference. One option is model-agnostic inference.

For more information, please contact:

John Tsang (john.tsang@statcan.gc.ca) or

Darren Gray (darren.gray@statcan.gc.ca).

References

Beaumont, J.-F. and Bissonnette, J. (2011). [Variance estimation under composite imputation: The methodology behind SEVANI](https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2011002/article/11605-eng.pdf). *Survey Methodology*, 37(2), 171-179. Available at: <https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2011002/article/11605-eng.pdf>.

Beaumont, J.-F., Bissonnette, J. and Bocci, C. (2010). SEVANI, Methodology Guide. Statistics Canada.

Chen, S., Haziza, D., Léger, C. and Mashreghi, Z. (2019). Pseudo-population bootstrap methods for imputed survey data. *Biometrika*, 106(2), 369-384.

PROJECT: Estimation with nonprobability samples using multiple reference surveys

The goal of this project was to propose a method to calculate pseudo-weights for nonprobability surveys using two or more probability samples as reference surveys.

Progress:

The participation mechanism in nonprobability samples is unknown and estimated target parameters from nonprobability samples might suffer from selection bias. Current approaches developed to support statistical inference from nonprobability samples are unable to accommodate more than one reference sample. In this paper, we propose a general framework to address participation bias in nonprobability samples using multiple reference surveys. We focus primarily on the calibration estimators. These estimators have great flexibility in situations with limited access to survey microdata and are straightforward for practical implementation. We describe two methods for variance estimation that account for all sources of variability of the proposed estimators. We assess the performance of the various methods through an extensive simulation study, which demonstrated satisfactory performance of the raking ratio calibration estimator in situations with highly dispersed participation probabilities in nonprobability samples and markedly smaller variance estimates for continuous outcomes. Finally, we illustrate the application of these methods using data from a real-world study of working adults in Canada.

In the 2025-2026 year, we published the paper Landsman, Wang, Carrillo-Garcia, Mitani, Smith, Graubard, Bui and Carnide (2026) in *Statistics in Medicine*.

For more information, please contact:

Ivan Carrillo-Garcia (ivan.carrillo-garcia@statcan.gc.ca).

Reference

Landsman, V., Wang, L., Carrillo-Garcia, I., Mitani, A.A., Smith, P.M., Graubard, B.I., Bui, T. and Carnide, N. (2026). Correction for participation bias in nonprobability samples using multiple reference surveys. *Statistics in Medicine*, 45, 3-5, e70403. Available at: <https://doi.org/10.1002/sim.70403>.

4 Confidentiality and Access

Confidentiality research at Statistics Canada continues to focus on developing new methods and ideas that offer alternative forms of access while continuing to ensure that private individual and business information is not disclosed in any way. Progress was made on the projects described below. The team responsible for the Centre for Confidentiality and Access at Statistics Canada also continued to offer consultation services to internal and external partners to help develop capacity in disclosure risk identification and treatment (see [Section 5.5](#)).

PROJECT: Synthetic data

Work continued developing governance frameworks for synthetic data at Statistics Canada. Synthetic data can address confidentiality issues with personal data while retaining as much analytical value as possible. Synthetic data can be especially useful when looking for collaborative opportunities with external stakeholders that may not have access to the confidential microdata.

Progress:

An internal working group was developed to better understand the potential use of synthetic data and update guidance material to encourage its use and development. This includes updates to key governance pieces as well as a more pragmatic handbook (Gauvin, 2026).

For more information, please contact:

Héloïse Gauvin (heloise.gauvin@statcan.gc.ca) or

Steven Thomas (steven.thomas@statcan.gc.ca).

Reference

Gauvin, H. (2026). Handbook for the creation of public synthetic data files (DRAFT). Internal document, Statistics Canada.

PROJECT: Optimization strategies for complementary cell suppression

The Complementary Cell Suppression (CCS) methodology, used at Statistics Canada for many of its economic surveys, is a standard method for suppressing sensitive cells when releasing tabular magnitude variables. The methodology is well developed and supported with Statistics Canada's SAS based G-Confid software. The methods rely on linear programming optimization where optimal suppression solutions are obtained that ensure that suppression patterns are valid and minimize the amount of information being suppressed.

Progress:

Statistics Canada has continued its development of a Python-based solution for G-Confid. The software will be released internally in June 2026. The updated software includes an optimal additive rounding solution, a tool for the calculation of sensitivity measures, the creation of an optimal CCS pattern, and an audit program to ensure that patterns are valid. Open-source optimization solutions available through the PuLP Python package were studied as options for finding optimal solutions for the CCS pattern. With careful consideration of these options, the Python solution was successfully applied to the most complex cases at Statistics Canada. The group also investigated other options for ensuring that a valid suppression pattern can be found under the specified constraints for large scale problems. The solutions combine scaling of microdata, modifying constraints, utilizing available solver options and taking advantage of cloud infrastructure.

For more information, please contact:

Steven Thomas (steven.thomas@statcan.gc.ca) or

Haoluan Chen (Haoluan.chen@statcan.gc.ca).

PROJECT: A utility-disclosure risk-based impact assessment of the implementation of differential privacy in the context of tabular data dissemination

Disclosure control strategies, that support a Differential Privacy (DP) guarantee, continue to be studied at Statistics Canada in terms of their practical implications on the utility of the information being released as well as their risk guarantees. Our focus continues to be in the context of the Canadian Census Dissemination.

Progress:

A paper (Miller, 2026) was developed that includes a discussion about the various versions of DP, their advantages and disadvantages, and a formal analysis of the core properties of DP. A statistically intuitive interpretation of DP was also provided. The key findings of the paper are that DP provides strong and rigorous privacy guarantees, but this may come at the cost of a large loss of utility relative to what Statistics Canada typically tolerates. There comes a variety of practical challenges with implementing a DP framework and therefore its adoption is not a trivial task. We concluded that widescale adoption of DP is not presently operationally feasible in the context of Census dissemination. However, DP could at present be readily applied to administrative data when the statistics to be released are counts and the number of unique ways each contributor provides their information is low. It could also be feasible to generate synthetic datasets and release them in remote tabulation tools, such as Real-Time-Remote Access (RTRA). We do not rule out the possibility of the future discovery of DP compliant algorithms that provide more attractive privacy-utility trade-offs. If this occurs, the adoption of DP compliant algorithms could be revisited.

For more information, please contact:

Joshua Miller (Joshua.miller@statcan.gc.ca) or
Steven Thomas (steven.thomas@statcan.gc.ca).

Reference

Miller, J. (2026). *A Privacy-Utility Based Study of Differential Privacy*. Branch working paper, Modern Statistical Methods and Data Science Branch (SSMD-2026-001E).

5 Support (Resource Centres)

5.1 Time Series Research and Analysis Centre

The Time Series Research and Analysis Centre maintains high-level expertise in time series (seasonal adjustment, trend-cycle estimation, reconciliation, benchmarking, linking, calendarization, modelling, forecasting/nowcasting, etc.). The Centre provides specialised services to ensure that time series analysis methods used in Statistics Canada are appropriate and in line with recent advances; develops time series analytical and processing software; and offers services, consultation, and training in time series throughout the agency and to clients outside Statistics Canada.

The projects can be split into four sub-topics:

- Consultation and training in time series;
- Time series methods;
- Support and enhancement of the time series processing system and tools;
- Methodological support to consumer and producer price index programs.

Progress:

Consultation and training in time series

The Time Series Research and Analysis Centre is responsible for developing and delivering training on time series methods including seasonal adjustment, benchmarking, reconciliation, and time series modelling

to participants from Statistics Canada as well as those from other agencies. This year, the Centre offered courses on time series components, seasonal adjustment, benchmarking, and reconciliation to internal and external participants through the Statistics Canada Training Institute (Statistics Canada, 2026). The examples for the reconciliation course have been converted from SAS to R. The Centre also participated in outreach and ad hoc training to other groups in Statistics Canada on time series topics (special course offerings for groups of analysts, seasonal adjustment dashboard demonstrations).

The Centre also provides guidance and consultation on time series projects in general for programs throughout Statistics Canada. This year, the Centre has offered consultation to various internal partners: vital statistics (seasonal adjustment of weekly series), economic analysis (benchmarking), national economic accounts (seasonal adjustment), industry accounts (benchmarking), international trade in services (seasonal adjustment), labour force statistics (custom table of seasonally adjusted data), and public sector statistics (nowcasting/forecasting). Representatives from the Centre also periodically attend a weekly analyst forum to maintain a presence in the analyst community. The Centre regularly consults on back-casting to preserve or restore comparability across time. The Centre was also involved in consultations with federal and provincial public agencies, national statistical organizations, and other external agencies. The Centre consulted on calendarization of retail trade data, in particular.

Time series methods

The COVID-19 pandemic period heavily affected many economic time series during the 2020-2022 period. With data now available until 2025, the impact of the pandemic on seasonal patterns, the trend-cycle, and volatility was assessed. The Centre completed an investigation using representative time series from various statistical programs (labour, trade, manufacturing, building permits and investment in building construction, tourism, and consumer price index) to determine whether these series have returned to pre-pandemic patterns or have shifted to a new post-pandemic regime. The results of the study were presented at the 2025 OECD-INSEE Workshop on Time Series Analysis for Official Statistics (Cheng, An and Guest, 2025).

Due to changes in employment insurance eligibility rules at the beginning of 2017, two sets of seasonal adjustment options are currently maintained for the time series of the Employment Insurance Statistics (EIS) program. A study conducted by the Centre demonstrated that this is not necessary anymore and that seasonal adjustment could be performed with a single set of options while maintaining similar quality. This will be implemented when EIS data is next revised historically and will significantly improve the efficiency of the annual review of seasonal adjustment options for this program.

The Centre continued an exploration of machine learning to detect seasonality in time series – a complicated problem in seasonal adjustment. The algorithms tested performed well on simulated data, even on data unseen during training. Unfortunately, the performance of the algorithms deteriorated significantly on real data. The problem lies firstly in the limited quantity of labeled data available for training, and secondly in the difficulty of simulating series with characteristics similar to real series to enlarge the quantity of labeled data. The exploration is ongoing.

The Centre completed a new set of guidelines on time series methods at Statistics Canada, which will be published in 2026-2027. In addition to refreshing previous guidelines, sections on benchmarking and reconciliation techniques were expanded, and a section on temporal comparability was introduced.

Support and enhancement of the time series processing system and tools

The Time Series Research and Analysis Centre develops and maintains a number of important tools used to process and analyse time series data for the Statistics Canada programs producing seasonally adjusted data, in particular the generalized system G-Series, for benchmarking and reconciliation (raking and balancing) (Statistics Canada, 2016; Ferland, 2025), the Time Series Processing System (Ferland, 2022), and the Seasonal Adjustment Dashboard (Verret, 2021).

The remaining work on a new, open-source, version of G-Series in R was completed this year. Version 3 of G-Series was released officially as the R package [gseries](#) in June 2025 (Ferland, 2025).

The Time Series Processing System (TSPS) is a customizable SAS-based application to apply time series techniques including seasonal adjustment, benchmarking, and reconciliation, used extensively in the production of seasonally adjusted estimates for sub-annual programs within Statistics Canada (many of them being mission critical). The system is in a mature and stable state. However, it requires updating on an ongoing basis to broaden functionality and address new needs of programs in the agency. Just like G-Series, the TSPS is being redeveloped in R. The work started this year and is planned to end in December 2027. It will provide the flexibility to incorporate tools and new techniques available from open-source software.

A feature was added this year to the Seasonal Adjustment Dashboard to allow analysts to easily filter time series based on whether the current reference month was identified as an outlier by the seasonal adjustment process. The Centre also provided training to subject-matter analysts.

The Centre continued improving a system to seasonally adjust series in bulk that could process large numbers of time series without manual intervention. The resulting seasonal adjustment would not meet the quality criteria for official publication but could yield acceptable quality for experimental products.

Methodological support to consumer and producer price index programs

The Time Series Research and Analysis Centre also has a unit dedicated to providing methodological support to consumer and producer price index (CPI and PPI) programs.

The Centre continued work on the development of a set of quality indicators for the CPI. Measures of coverage were produced for the four dimensions of sampling: geographic coverage, market coverage of outlets, product coverage, and time coverage. For the product and time dimensions, CPI sampling bias was also estimated for the major retail outlets providing scanner data.

The Centre investigated modelling approaches that could provide a sound methodological framework for incorporating alternative data with survey data in some producer price index estimates (industrial products, raw materials). A Bayesian framework was proposed and presented to an internal scientific review committee. The approach was also discussed with experts in Bayesian statistics internally and will be presented to a committee of international experts (Statistics Canada's Advisory Committee on Statistical Methods) in June 2026 (Francis, Dewan and Carrillo-Garcia, 2026).

For more information, please contact:

Etienne Rassart (etienne.rassart@statcan.gc.ca).

References

Cheng, B., An, L. and Guest, J. (2025). Assessing the impact of the COVID-19 pandemic on Statistics Canada's economic time series. Internal document, Statistics Canada, Presented at the 2025 OECD-INSEE Workshop on Time Series Analysis for Official Statistics.

Ferland, M. (2022). Time series processing system – v3.08. Internal document, Statistics Canada.

Ferland, M. (2025). [gseries: Improve the Coherence of Your Time Series Data](https://StatCan.github.io/gensol-gseries/en/). R package, version 3.0.2. Available at: <https://StatCan.github.io/gensol-gseries/en/>. (Package [gseries](https://CRAN.R-project.org/package=gseries) also available on CRAN at: <https://CRAN.R-project.org/package=gseries>).

Francis, J., Dewan, A. and Carrillo-Garcia, I. (2026). *Bayesian Methods to Incorporate Auxiliary Price Data into Price Indices*. Presentation to the Statistics Canada Advisory Committee on Statistical Methods, June 2026, Statistics Canada.

Statistics Canada (2016). *G-Series 2.00.001 User Guides*. Internal document, Statistics Canada.

Statistics Canada (2026). [Training](https://www.statcan.gc.ca/eng/wtc/training). Available at: <https://www.statcan.gc.ca/eng/wtc/training>.

Verret, F. (2021). Statistics Canada's seasonal adjustment dashboard. *Proceedings: Symposium 2021, Adopting Data Science in Official Statistics to Meet Society's Emerging Needs*, Statistics Canada, Ottawa, Canada.

5.2 Resource Centre for Economic Statistical Tools and Innovation

The Generalized Systems Unit is part of the Integrated Business Statistics Program (IBSP) Infrastructure Section and is responsible for the support and development of three existing generalized systems - **G-Sam** (sampling), **Banff** (statistical data editing), and **G-Est** (estimation) – as well as two new systems, **Jasper** (the open-source replacement for G-Sam) and **Yoho** (the open-source replacement for G-Est).

Progress:

Projects in the unit can generally be categorized as support (for users of the system), research and development.

The unit handled a typical volume of support cases. These include support cases for both the current SAS-based systems, as well as the new open-source Python (Banff) and R (Jasper, Yoho) systems. A significant number of support cases involved providing guidance on the changes between system versions and advice on SAS to open-source migration. While some cases were resolved by a few email exchanges, others required more in-depth involvement and included recommendations on appropriate implementation of the system to achieve statistical goals. A summary of selected support cases is provided below.

The G-Sam/Jasper team:

- Provided assistance to various surveys as they converted their SAS-based sampling programs using G-Sam to R-based sampling programs using Jasper. Surveys included the Quarterly Survey of Financial Statements and Canada Emergency Business Accounts.

The Banff team:

- Provided guidance and assistance to Canada's Core Public Infrastructure Survey as they converted their SAS-based edit and imputation process to a metadata-driven process using the Python-based Banff Processor. The Banff team also worked on two custom modules, *ApplyRules* and *CubistR*, as part of the migration process.
- Worked directly with the T1 Family File Plus team to develop a Banff Processor process to handle deterministic editing rules, i.e., IF-THEN rules. This included a strategy to handle larger-than-memory files in the Banff Processor.

The G-Est/Yoho team:

- Supported request from partners in North-West Territories using G-Est.
- Provided assistance to the General Social Survey to migrate from G-Est to Yoho calibration.

Most of the research and development in the fiscal year was devoted to Statistics Canada's Shift to Open Source initiative. The team took an active role in updating users, both inside and outside the Branch, on the progress for the systems. This included presentations to the Generalized Systems Steering Committee, Branch Seminars, and Field Planning Board Meetings. A progress report for all generalized systems was provided to external users in Gray (2026).

The R package **Jasper** was released internally, via GitLab, on October 31, 2025, offering open-source allocation and selection functions to replace those found in the SAS-based generalized system G-Sam. These functions expand upon the equivalent macros in G-Sam, solving a wider variety of problems. Improvements include an expanded allocation problem, solved via optimization, and additional selection algorithms. The Jasper package also significantly reduced the codebase relative to G-Sam, by an approximate factor of one tenth, with only two external dependencies, reducing future maintenance and improvement costs, while increasing significantly the performance. The Jasper methodology was presented to the Branch's Scientific Review Committee (Stinner, 2026), and peer-review of the mathematics is underway. A migration guide was created on GitLab to aid existing G-Sam users in their migration to Jasper. An external release on GitHub and the CRAN is planned for the coming year.

G-Sam stratification methods were converted into R in a separate set of functions, outside of Jasper, and put in a GitLab repository. A beta version of this repository was made available to internal users with a planned release date of May 2026. Research into new stratification methods (Chase, Gray and Stinner, 2026) showed promises and will continue into the new year.

Development continues on the Python version of **Banff**, which was originally released in January 2025. As part of the development cycle, various production and beta versions of the Banff procedures and Banff Processor systems were released internally on GitLab and externally on GitHub and PyPi. Significant progress was made on new Banff Processor modules, to offer new, modern, imputation methods. The *ApplyRules* module applies determinist IF-THEN editing rules implemented using Pandas syntax, providing users a simple, metadata-driven approach to specify, execute, and track editing rules alongside more complex edit and imputation procedures. A second module, *CubistR*, offering imputation via the Cubist algorithm, was developed in collaboration with members of the IBSP imputation team, as part of a research project exploring the use of Cubist imputation on survey data (Mireuta, Bowditch, Gray, Kuromi, Ngabo Nsengiyaremye and Sivathayalan, 2026).

Finally, a beta version of **Yoho**, the R replacement for G-Est, was released in December 23, 2025. The beta release includes functions for calibration, point and variance estimation, and bootstrap replicates. Internal tests by the development team and beta users indicate consistency with G-Est outputs with a significant improvement in performance. The Yoho codebase is remarkably streamlined when compared to G-Est: approximately one percent of the size when measured in kilobytes. Yoho only requires two external dependencies, reducing expected maintenance costs. An official release version of Yoho, including the functions mentioned above, is planned for June 2026.

For more information, please contact:

Darren Gray (darren.gray@statcan.gc.ca) or

Pierre-Olivier Julien (pierre-olivier.julien@statcan.gc.ca).

References

Chase, P., Gray, D. and Stinner, M. (2026). Stratification via hierarchical clustering. Internal presentation, Statistics Canada.

Gray, D. (2026). Statistics Canada's generalized systems are migrating to R and Python. *The Survey Statistician*, 93, January 2026.

Mireuta, M., Bowditch, F., Gray, D., Kuromi, G., Ngabo Nsengiyaremye, B. and Sivathayalan, A. (2026). Modern Imputation methods applied via Open-Source and BANFF. Internal report, Statistics Canada.

Stinner, M. (2026). *The Jasper Package – An Open-Source-Style Sampling System*. Presented at the Modern Statistical Methods and Data Science Branch Scientific Review Committee, Meeting 59, Statistics Canada.

5.3 Record Linkage Resource Centre

The objectives of the Record Linkage Resource Center (RLRC) are to provide consultation services to internal and external users of record linkage methods, which includes making recommendations about the software and methods to be used, and collaborative work on record linkage applications. We also facilitate the dissemination of information on record linkage methods, software, and policy as well as the analysis of linked data to interested parties inside and outside Statistics Canada.

Progress:

During the year, most of the methodological work focused on maintenance and support for users of version 3.5 of G-Link on SAS servers in cloud computing. A typical volume of support cases for G-Link was processed by the project team. Most of these were resolved with suggestions on how to apply the system in practical terms; however, several required more involvement.

The team also undertook work on organizing and updating documentation on linkage strategy and synthetic data as part of the Methodological Acceleration program. The resulting products are intended to supplement resources for users of the forthcoming Python-based G-Link Web software.

The RLRC has also worked on a variety of other probabilistic linkages in the Social Data Linkage Environment (SDLE). These linkages helped us to analyze the performance of the software and the

solutions to be provided. Work was also undertaken on production and planning of redevelopment for Canadian Cancer Registry linkages and porting the regular and sampling versions of Longitudinal Social Data Development Program's (LSDDP's) Population Group Model to Python ahead of a cessation of LSDDP's activities so that it may be used in future programs including the Social Demographic Database.

Members of the team offered formal courses with Statistics Canada's Training Centre, as well as started planning for a workshop to be offered to Colombian official statistics personnel from CENSAS and updated reference and strategy documentation for users of probabilistic record linkage techniques.

For more information, please contact:

Eric Hortop (eric.hortop@statcan.gc.ca).

5.4 Data Analysis Resource Centre

The main goal of the Data Analysis Resource Centre (DARC) is to provide advice on the appropriate use of data analysis tools and methods, and to promote best practices in this area. DARC's services – which focus mainly on surveys, censuses, or administrative data – are available to the employees of the Agency and other departments, as well as to analysts and researchers from academia and Research Data Centers (RDCs).

Progress:

Consultations

Consultation services were provided as requested by internal and external clients. Between April 1, 2025 and March 31, 2026, DARC responded to approximately 40 requests. The questions varied in complexity and included topics such as age-standardized analysis, confidence interval overlap test, logistic regression on rare events with survey data, pooling data from several sources, model diagnostics, Fay adjustment in variance estimation, confidence intervals for small proportions, comparison of dependent subpopulations, multiple comparisons, predictive margins and variance of medians. DARC also helped clients with the implementation of statistical methods in R, SAS, SUDAAN and STATA software. In addition, DARC reviewed analytical papers for external and internal publications.

Provision of training

DARC presented at Statistics Canada's Data Interpretation Workshop on data analysis with complex survey data, in English and French.

DARC also presented the sessions on linear and logistic regression with complex survey data of the Statistical Modelling Course at Statistics Canada, in English.

DARC prepared and presented the following bilingual branch-level seminars:

- *Avoiding Confidence Interval Pitfalls: Tips and Guiding Principles*
- *AI-Assisted Methodology Problem-Solving: A Case Study*
- *Meeting Data Disaggregation Needs by Achieving a Better-than-Average Pooling Estimates Strategy*

Collaboration

DARC collaborated in developing measurement strategies for the Workplace Mental Health Performance Measurement Project with the Treasury Board Secretariat (TBS). This project used data from the 2019, 2020, 2022 and 2024 cycles of the Public Service Employee Survey (PSES) to measure latent variables like psychological risk factors, behaviors, etc. and to calculate factor scores for different levels of aggregation. The factor scores developed for this project were used to create the Federal Public Service Workplace Mental Health Dashboard: [Mental Health Dashboard- Canada.ca \(tbs-sct.gc.ca\)](https://tbs-sct.gc.ca/mental-health-dashboard). The measurement models were developed using factor analysis and structural equation modelling as discussed by Blais, Mach, Michaud and Simard (2020) and Blais, Michaud, Simard, Mach and Houle (2021). This year, improvements were made to the model, with some factors being added. All scores since 2019 were re-estimated using the new model.

For more information, please contact:

Ryan Chepita (ryan.chepita@statcan.gc.ca) or

Isabelle Michaud (isabelle.michaud@statcan.gc.ca).

References

Blais, A.-R., Mach, L., Michaud, I. and Simard, J.-F. (2020). Analysis of the Public Service Employee Survey Items as Measures of the Psychosocial Risk Factors. Presentation to the Workplace Mental Health Performance Measurement Steering Committee, October 7, 2020.

Blais, A.-R., Michaud, I., Simard, J.-S., Mach, L. and Houle, S. (2021). [Measuring workplace psychosocial factors in the federal government](https://www150.statcan.gc.ca/n1/en/pub/82-003-x/2021012/article/00001-eng.pdf). *Health Reports*, 32, 1-12. Available at: <https://www150.statcan.gc.ca/n1/en/pub/82-003-x/2021012/article/00001-eng.pdf>.

5.5 Centre for Confidentiality and Access

The methodology group responsible for confidentiality and access methods continued to offer consultation and support services to internal and external partners on the various access solutions and disclosure avoidance strategies.

Anonymized data

The Centre for Confidentiality and Access has continued to offer governance, guidance and strategic advice in producing open datasets. This year, over 12 new Public Use Microdata Files were reviewed and made available on the Statistics Canada website. Statistics Canada has also continued to offer its expertise to several groups internationally and within Canada.

Consultation

Domestically, we are continuing to work with Health Canada and the Public Health Agency of Canada in the development of strategies for releasing more information. We have also worked with the University of Sherbrooke and internal stakeholders in developing modern access options for federated analyses. Internationally, we have been consulting with the Australian Bureau of Statistics on improvements to our complementary cell suppression methodology that is supported through our open-source tool G-Confid.

Output vetting streamlining

The Centre has been working with our Research Data Centre colleagues on developing standardized vetting strategies to simplify and streamline the output vetting process. These standards will pave the way for more advanced vetting with automated processes being considered for the future.

For more information, please contact:

Steven Thomas (steven.thomas@statcan.gc.ca).

5.6 Questionnaire Design Resource Centre

The Questionnaire Design Resource Centre (QDRC) is a focal point of expertise at Statistics Canada for questionnaire design and evaluation. The QDRC provides consultation and support services and carries out projects and research related to the development, testing and evaluation of survey questionnaires.

The QDRC plays a very important role in quality management and responds to program requirements throughout Statistics Canada by consulting with clients, respondents, and data users, and by pre-testing survey questionnaires.

While much of the QDRC's work is carried out on a cost-recovery basis, the Centre is frequently approached on an ad hoc basis for expert reviews and consultation services on a wide variety of surveys. The group also offers courses on questionnaire design.

Progress:

The QDRC conducted numerous reviews of survey questionnaires. Most of these reviews involved Statistics Canada's questionnaires, while several were completed for surveys conducted by other government organizations, including the Administrative Tribunals Support Service of Canada, the Bank of Canada, the Canadian Human Rights Commission, the Canada Revenue Agency, Natural Resources Canada, and Shared Services Canada. The group also contributed to various corporate consultation initiatives.

For more information, please contact:

Jeremy Solomon (jeremy.solomon@statcan.gc.ca).

5.7 Data Ethics Secretariat

The Data Ethics Secretariat was created to support the internal Data Ethics Committee and to focus on identifying questions and issues of an ethical nature pertaining to collection initiatives at Statistics Canada. It also facilitates the operations by offering balanced recommendations to the Principal Data Ethics, Quality, and Scientific Integrity Officer (PDEQSIO) on issues pertaining to data ethics. These recommendations help determine whether a case must be presented to the internal Data Ethics Committee and the external Advisory Council on Ethics and Modernization of Microdata Access.

PROJECT: Ethics in artificial intelligence

This project aims to introduce a distinction between hard and easy problems in the field of Artificial Intelligence (AI) or Machine Learning (ML) ethics. It mirrors a well-known distinction in the literature on the philosophy of mind between the hard and easy problems of consciousness. That distinction is used to highlight the importance of existing ethical guidance and to show how we can improve the chances at

finding actionable solutions in the field of AI/ML ethics. This is especially relevant for organizations working under established governance instruments on topics such as scientific integrity or data ethics. Such organizations already have ethical expectations for their members. The essence of these expectations remains relevant even as they fully or partially automate tasks performed by individuals with the help of AI or ML algorithms. We can replace a person's involvement in a process by using algorithms, but we cannot throw away existing ethical guidance about that process.

Progress:

The project concluded with a peer-reviewed journal paper by Rochefort-Maranda (2026a).

For more information, please contact:

Guillaume Maranda (guillaume.maranda@statcan.gc.ca) or
Ryan Chepita (ryan.chepita@statcan.gc.ca).

Reference

Rochefort-Maranda, G. (2026a). The hard and easy problems of AI ethics: how to leverage existing guidance. *AI Ethics*, 6, 278. Available at: <https://doi.org/10.1007/s43681-025-00907-1>.

PROJECT: Ethical challenges for National Statistical Offices (NSOs)

This project aimed at understanding how ethical challenges for NSOs can appear by using literature on moral psychology, specifically literature on situationism. Situational variables can influence morally significant actions, and important lessons can be learned from this fact to improve ethical guidance. This study is attractive to the extent that the emergence of ethical challenges for NSOs does not seem to stem from a change in moral character of the employees nor from a lack of awareness of the relevant responsibilities or principles.

Progress:

The project concluded with an article by Rochefort-Maranda (2026b) published in a peer-reviewed journal.

For more information, please contact:

Guillaume Maranda (guillaume.maranda@statcan.gc.ca) or
Ryan Chepita (ryan.chepita@statcan.gc.ca).

Reference

Rochefort-Maranda, G. (2026b). Data ethics: Lessons from situationism for a National Statistical Organization. *Statistical Journal of the IAOS* (IOS Press/IAOS). Available at: <https://doi.org/10.1177/18747655261439346>.

5.8 Quality Secretariat

The Quality Secretariat ensures the production of relevant, high quality and ethically sourced data by supporting the development and implementation of policies and procedures that promote sound quality-management practices. A new mission for the Quality Secretariat has been drafted and will be reviewed by Statistics Canada's Strategic Management Committee (SMC).

PROJECT: Quality indicators for administrative data

While quality indicators (e.g. accuracy) for data from sample surveys are well-established, there are currently no consistent and established means of providing quality indicators for administrative data. While research has been conducted internationally, makeshift quality indicators at Statistics Canada for programs ingesting alternate data sources have been largely developed at the program level. Often, these indicators take the form of metadata, given the challenges of quantifying the quality of such a dataset or estimates. The purpose of this project is to identify candidates for quality indicators that are appropriate, robust, and easy to apply and understand.

Progress:

Drawing upon international research, including work accomplished by a sub-group of the United Nations Expert Group on National Quality Assurance Frameworks, a series of candidates of good quality indicators were identified. The relative importance of each indicator may be determined by the data user, which is consistent with the well-established concept that quality is akin to fitness for use. It was proposed that the weighted indicators could then be combined into an overall metric. This high-level proposal was presented to the Methodology Accelerator group at the March 2026 session. Based on feedback from the group, the next steps will be to identify a program to test the proposed measures, to implement the process and document any challenges, and to determine what to retain and modify for future applications.

For more information, please contact:

Ryan Chepita (ryan.chepita@statcan.gc.ca) or

Yuliya Romanyuk (yuliya.romanyuk@statcan.gc.ca).

PROJECT: Quality indicators for the Generic Statistical Business Process Model (GSBPM)

Within the ongoing United Nations Expert Group on National Quality Assurance Frameworks, a subgroup was formed to derive quality indicators to correspond to each sub-phase of the GSBPM. The GSBPM describes eight chronological phases that comprise the statistical production process (Specify Needs, Design, Build, Collect, Process, Analyze, Disseminate and Evaluate), and each phase has three to eight sub-processes. The need for modernized quality indicators is driven by expanding possibilities for some sub-processes (e.g. how data are ingested) and by the increasingly complex types of data available to and appropriate for National Statistical Offices.

Progress:

Throughout 2025-26, the Quality Secretariat (QS) played an active role in participating in the subgroup's monthly meetings. The QS collaborated with another member country to derive additional quality indicators and good practices for some of the GSBPM phases. It then contributed to the editing of other phases. At the end of the fiscal year, a draft document was being finalized by the group's co-leads, with an eye to making it available for consultation in May 2026.

For more information, please contact:

Ryan Chepita (ryan.chepita@statcan.gc.ca).

6 Other activities

6.1 Survey Methodology Journal

Survey Methodology is a free online peer-reviewed statistical journal published twice a year by Statistics Canada since 1975. The journal aims to publish innovative theoretical or applied research papers, and sometimes review papers, that provide new insights on statistical methods relevant to National Statistical Offices and other statistical organizations. Papers are published free of charge in both official languages and released at: www.statcan.gc.ca/surveymethodology. Its [editorial board](#) includes world-renowned leaders in survey methods from the government, academic and private sectors.

Progress:

The June 2025 and December 2025 issues (issues 51-1 and 51-2) were released. The [June 2025](#) issue is a special issue for the celebration of the 50th anniversary of *Survey Methodology*. It features two discussion papers. The first discussion paper, by Carl-Erik Särndal, is entitled “Progress in survey science: yesterday – today – tomorrow” whereas the second discussion paper, by J.N.K. Rao and Sharon Lohr, is entitled “Trends and directions in sample survey theory and methods”. Both papers are followed by discussions from eminent survey statisticians and a rejoinder. With this special issue, we launched a new section of the journal for interview papers. These papers are based on conversations with survey statisticians or methodologists that had a significant impact in the survey sampling field. The June 2025 special issue includes two interview papers, one interview with Ivan P. Fellegi, a former Chief Statistician of Canada, and the other with Geoffrey Hole, a former Director at Statistics Canada. It also features seven invited papers by renowned experts in survey statistics and methodology.

Thirteen papers, including a short note, were published in the [December 2025](#) issue, which featured the 2025 Waksberg invited paper by Mike Hidioglou entitled “Sampling for business surveys at Statistics Canada”.

In 2025, 77 papers were submitted to the journal. The average number of days from submission to initial decision was 46. All but two submitted papers were reviewed within 150 days, and 82% of them were reviewed within 90 days. Among those 77 papers, 49 were rejected, 20 were accepted, and 8 had not received a final decision (including papers that were not revised by the authors before the deadline) as of May 23, 2026. From April 2025 to March 2026, the *Survey Methodology* pages were viewed 51,204 times.

We are currently planning two forthcoming special issues in 2026 or 2027. The first of these special issues will include selected papers presented at the seventh International Conference on Establishment Statistics in Glasgow in June 2024. Paul Smith and Mojca Bavdaz are the guest editors for this special issue, which is expected for December 2026. The second special issue whose guest editors are Maria Rosaria Ferrante and Natalie Shlomo is expected for December 2027. It will be devoted to the theme “Shaping the future of survey statistics in the data-driven era”.

For more information, please contact:

Jean-François Beaumont (jean-francois.beaumont@statcan.gc.ca).

6.2 Knowledge Transfer – Statistical Training

The Statistical Talent Development Working Group has as its main mandate training in statistics within the Branch and the organization. Several courses were offered this year, including those related to time series, questionnaire design, sampling, record linkage, imputation, weighting, modeling and statistical analysis of survey data, as well as an introductory simulation workshop.

As for the new activities, the group continued to design and prioritize learning activities that can be developed in a timely manner and focused on active learning. This year, work to revise the second course on statistical analysis of survey data as well as the course on variance estimation continued.

Regarding the next year, we will continue to offer curriculum courses based on demand and instructor availability. In addition, we aim to offer the course on variance estimation in both official languages.

The Talent Development Working Group offers various types of training opportunities so that employees can enjoy some flexibility in their professional development. In addition to the activities mentioned above, there are many opportunities for self-training and self-learning as well as communities of practice.

For more information, please contact:

Keven Bosa (keven.bosa@statcan.gc.ca).

6.3 Statistics Canada's International Methodology Symposium

The *Proceedings* of Statistics Canada's 2024 International Methodology Symposium whose theme was "The future of official statistics" were published in September 2025. The *Proceedings* are available on our website at: <https://www150.statcan.gc.ca/n1/en/catalogue/11-522-X>.

Statistics Canada will organize the next International Methodology Symposium in a future year to be determined.

As an interim measure, a workshop internal to Statistics Canada employees has been proposed. Suggested topics include leadership, responsible artificial intelligence, the definition of official statistics, data integration and governance.

For more information, please contact:

Peter Wright (peter.wright@statcan.gc.ca).

7 Research papers sponsored by the Methodology Research and Development Program

Beaumont, J.-F., Bocci, C., Bosa, K., Dasylva, A. and Maranda, G. (2026). Prediction inference for finite population totals or means. Submitted for publication in a peer-reviewed journal.

Bosa, K., Beaumont, J.-F., Bocci, C. and Sombo, S. (2025). The use of a random forest algorithm in small area estimation. Paper presented at the Advisory Committee on Statistical Methods, June 2025, Statistics Canada.

Chase, P., Gray, D. and Stinner, M. (2026). Stratification via hierarchical clustering. Internal presentation, Statistics Canada.

Chehbouni, K., Haddou, M., Cheung, J.C.K. and Farnadi, G. (2026). Neither valid nor reliable? Investigating the use of LLMs as judges. Proceedings of NeurIPS 2025. Available at: https://papers.nips.cc/paper_files/paper/2025/hash/829e8f32d76a6248815bf5b01633811d-Abstract-Position_Paper_Track.html.

Cheng, B., An, L. and Guest, J. (2025). Assessing the impact of the COVID-19 pandemic on Statistics Canada's economic time series. Internal document, Statistics Canada, Presented at the 2025 OECD-INSEE Workshop on Time Series Analysis for Official Statistics.

Dasyilva, A. (2025). A differentially private record linkage for the production of official statistics. Paper presented at the Advisory Committee on Statistical Methods, October 2025, Statistics Canada.

Dasyilva, A. and Goussanou, A. (2026). A joint model of record duplication and linkage error for the Canadian Census of Population. Internal report, Statistics Canada, Ottawa.

Denis, N. and Haddou, M. (2026). Prediction-powered estimation: Unbiased model-assisted estimation. *Journal of Official Statistics*. Available at: <https://doi.org/10.1177/0282423X261451320>.

Ferland, M. (2025). [gseries: Improve the Coherence of Your Time Series Data](#). R package, version 3.0.2. Available at: <https://StatCan.github.io/gensol-gseries/en/>. (Package [gseries](#) also available on CRAN at: <https://CRAN.R-project.org/package=gseries>).

Francis, J., Dewan, A. and Carrillo-Garcia, I. (2026). *Bayesian Methods to Incorporate Auxiliary Price Data into Price Indices*. Presentation to the Statistics Canada Advisory Committee on Statistical Methods, June 2026, Statistics Canada.

Gauvin, H. (2026). Handbook for the creation of public synthetic data files (DRAFT). Internal document, Statistics Canada.

Gray, D. (2026). Statistics Canada's generalized systems are migrating to R and Python. *The Survey Statistician*, 93, January 2026.

Haddou, M., Seimandi, T. and Del Monaco, A. (2025). [Fairness in Machine Learning: Companion paper to the Framework on Responsible AI for Official Statistics](#). Available at: https://unece.org/sites/default/files/2025-10/Companion%20Paper%20on%20Fairness%20in%20Machine%20Learning_Responsible%20AI%20Framework_ADSaMM%20group.pdf.

Imbrogno, A. and Verret, F. (2025). Dependent random group variance estimation using Fay's method. Paper submitted for publication in the *Survey Methodology* journal.

Landsman, V., Wang, L., Carrillo-Garcia, I., Mitani, A.A., Smith, P.M., Graubard, B.I., Bui, T. and Carnide, N. (2026). Correction for participation bias in nonprobability samples using multiple reference surveys. *Statistics in Medicine*, 45, 3-5, e70403. Available at: <https://doi.org/10.1002/sim.70403>.

Miller, J. (2026). *A Privacy-Utility Based Study of Differential Privacy*. Branch working paper, Modern Statistical Methods and Data Science Branch (SSMD-2026-001E).

Mireuta, M., Bowditch, F., Gray, D., Kuromi, G., Ngabo Nsengiyaremye, B. and Sivathayalan, A. (2026). Modern Imputation methods applied via Open-Source and BANFF. Internal report, Statistics Canada.

Rocheffort-Maranda, G. (2026a). The hard and easy problems of AI ethics: how to leverage existing guidance. *AI Ethics*, 6, 278. Available at: <https://doi.org/10.1007/s43681-025-00907-1>.

Rocheffort-Maranda, G. (2026b). Data ethics: Lessons from situationism for a National Statistical Organization. *Statistical Journal of the IAOS* (IOS Press/IAOS). Available at: <https://doi.org/10.1177/18747655261439346>.

Statistics Canada (2026). [Training](https://www.statcan.gc.ca/eng/wtc/training). Available at: <https://www.statcan.gc.ca/eng/wtc/training>.

Stinner, M. (2026). *The Jasper Package – An Open-Source-Style Sampling System*. Presented at the Modern Statistical Methods and Data Science Branch Scientific Review Committee, Meeting 59, Statistics Canada.

Toupin, M.-H. and Martin, V. (2026). [Improving the coverage of confidence intervals with respect to degrees of freedom: Application to the Canadian Census](https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2026001/article/00005-eng.pdf). *Survey Methodology*, 52(1), 187-211. Available at: <https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2026001/article/00005-eng.pdf>.

United Nations Economic Commission for Europe (UNECE) (2025). [Framework on Responsible AI for Official Statistics](https://unece.org/sites/default/files/2025-10/Responsible%20AI%20for%20Official%20Statistics%20Framework_ADSaMM%20Group_HLG-MOS_0.pdf). Available at: https://unece.org/sites/default/files/2025-10/Responsible%20AI%20for%20Official%20Statistics%20Framework_ADSaMM%20Group_HLG-MOS_0.pdf.

United Nations Economic Commission for Europe (UNECE) (2026). UNECE handbook on Uncertainty quantification in ML-enhanced statistical inference - A practitioner handbook with applications in official statistics. United Nations Economic Commission for Europe (UNECE), HLG-MOS ModernStats Initiative. Available at: <https://unece.org/statistics/documents/2026/06/reports/uncertainty-quantification-ml-enhanced-statistical-inference>.

Verret, F. and Walker, B. (2025). Application of the Fay-Herriot model with correlated design errors: The Labour Force Survey experience. Proceedings of the 2025 SSC Annual Meeting, Statistical Society of Canada, Saskatoon, Canada.

You, Y. (2026). Comparison and evaluation of small area EBLUP estimators under benchmarking with application and simulation study. Submitted to a peer-reviewed journal for possible publication.

You, Y. and Bosa, K. (2026). An empirical study on the use of informative priors for the hierarchical Bayes small area estimation with application. Research paper in progress to be submitted for possible publication.