|  |  |
| --- | --- |
|  | **Summary of Recommendations**  Prepared by Statistics Canada for the  Statistical Institute of Belize  **Belmopan, April 22, 2016** |

**Business Survey Infrastructure Component**

**Mission # 3**

**April 19-22, 2016**

**Belmopan, Belize**

The third mission of the Business Survey Infrastructure component of the Project for the Regional Advancement of Statistics in the Caribbean (PRASC) took place in Belmopan, Belize from April 19-22, 2016 between Statistics Canada (StatCan) and the Statistical Institute of Belize (SIB).

This working mission focused on developing and specifying the edit and imputation procedures to clean and complete the data received through the Part A questionnaire and the Part B series of questionnaires.

Following three days of productive meetings and discussions, the following set of recommendations were made by Statistics Canada on the fourth day:

1. It is recommended that SIB uses a standardized definition of response rates at collection. Most national statistical offices define their rates similarly, including StatCan (Ref.: StatCan [Survey Methods and Practices](http://www.statcan.gc.ca/pub/12-587-x/12-587-x2003001-eng.pdf) document, Chapter 9).
2. Should another BES take place in the future, it would be recommended that SIB review all the assumptions driving the collection, for instance, the tracing effort, the time spent per interview, the need for interview appointments, the response rates, the systematic use of an official letter from the SIB Director General introducing the survey, the target distinctive linguistic groups via appropriate media, the salary model.

1. SIB should adopt the decision table model to specify the editing and imputation processes.
2. SIB should use the Total Revenue as the only key variable defining a partial response and driving the edit and imputation process. Manual pre-processing of questionnaires would help in rescuing the Total Revenue when missing, as well as the Total Expenses itself in some cases.
3. Three databases should be defined: One for raw collected data, one for preliminary E&I data, one for post E&I data (final data). The collected data, unusable inactive data, and the post-E&I data should be archived.
4. A single record layout should be defined to cover variables from all form types. This will facilitate data management including versioning and back-ups.
5. The automated editing and imputation process should not start until field operation and data collection activities are terminated for the survey cycle (no batch editing).
6. An additional result code should be allowed in CSPro to track records manually imputed by subject matter experts (result code = 8) prior to automated editing/imputation.
7. A derived variable should be created for every percentage variable (Post E&I). The derived variable will contain the BZ$ value of the converted percentage. Estimates will be generated from the BZ$ value.
8. As an output of the preliminary E&I process, a distinct dataset should be created to list unusable active records that will be accounted via reweighting.
9. As an output of the preliminary E&I process, a distinct dataset should be created to list unusable inactive records for which variables will be assigned zero-values (Post E&I). These records will be re-integrated as input to estimation.
10. SIB should start using the StatCan Electronic File Transfer (EFT) service if data has to be shared.
11. It is recommended that an automated imputation process be developed and implemented.
12. The nearest neighbour imputation method should be used. This method will use a distance function based on the total revenue.
13. Imputation groups should be created. These groups should be derived using the ISIC classification (18 groupings) and the revenue (one or three groupings depending on the number of donors available).
14. SIB should review the decision tables developed for the Business Revenue Module and the Wholesale/Retail questionnaire and decide if new edits are required to ensure that BR2 (Sales of goods purchased for resale) and BE3 (Goods purchased for resale) are greater than 0.

1. SIB should implement the data processing solutions and high level process flow model defined jointly. The proposal was devised to ensure maximal automation while keeping the system complexity within the bounds of SIB expertise and IT resource allocation.

1. SIB should ensure that system solutions are kept simple and avoid defining complex alternate flows to deal with marginal cases and exceptions that can effectively be addressed with manual interventions.

1. As a good quality assurance practice, SIB should secure the availability of IT resources for testing functionalities and statistician resources for validating the output.