



Annualizing Data for the Statistical Business Register (SBR)

October 2019

Background

A Statistical Business Register (SBR) is typically used to provide a common survey frame for all business surveys as it permits Statistical Offices to ensure that units are not unknowingly represented in multiple surveys and ensures similar coverage of the entire population of interest for all surveys.

A key aspect of any sampling frame is a size measure that permits efficient stratification of the population for sampling purposes. When using administrative data to populate these values, there are times when concepts may differ between sources, may differ from what is needed or it is desirable to make adjustments in order to create more stable values that are overall representative of the activity taking place. An issue that is often encountered with administrative data is that some data partners will require businesses to remit their payments by varying frequencies. The following sections discuss some issues that are often encountered and suggests methods for processing the data to meet the needs of a SBR.

GST / VAT Data

The GST data typically supply a Sales value for all businesses above a certain threshold. However, many countries have varying reporting frequencies for different types of businesses (monthly, quarterly or annually). This type of reporting not only makes the data difficult to compare with revenue from other sources, but it is also difficult to compare between the various records on the file.

The first step in annualizing the data should be to determine which time period the annualized figure should represent. Typically, a 12-month period is chosen such that most reportings should have been received, while keeping the reference period as current as possible.

To ensure that the annualized figure is made up of only 12 months' worth of data, it will be important to review the file to determine the criteria to be used to omit records outside of the time frame for value. The criteria should be documented and automated processes should be developed (a computer script or series of steps in Excel) in order to omit these records from the calculation in an efficient and dependable manner. The criteria can be applied to the file before moving to the next step or they can be applied at the time of the calculations. Care should be taken to ensure that any records which should apply to a future estimate are properly included in the calculation for the next period.

As with any administrative data, there can be revisions submitted therefore duplicated reporting need to be removed. Similar to the previous step, the files need to be manually reviewed to determine the criteria for identifying which records should be kept / removed from the calculation. These criteria should be included within the documentation for calculating the estimates and automated processes should be developed (a computer script or series of steps in Excel) in order to omit the duplicate records from the calculation in a quick and dependable manner. Again, the criteria can be applied to the file before moving to the next step or they can be applied at the time of the calculations.

Once the various criteria are determined and applied, annualizing is simply the task of summing the records by business entities. As VAT / GST data can be reported monthly, quarterly or annually, there may be a varying number of records to be summed based on the reporting frequency.

It should be noted that there can be occasions where some reportings are missing. This could be a valid situation where some monthly or quarterly reporters don't operate for all 12 months of the year. However, it is also possible that an entity is late in submitting their reports. For simplicity, this process does not deal with these types of situations. Instead, it is strongly suggested that part of the SBR QA processes include:

- A review of large changes in the size measures and
- Unusual ratios between the size measures

to understand (and manually adjust as required) these types of situations. The smaller Caribbean populations would not warrant the complexity required to account for these types of situations in an automated fashion.

To ensure that the data continue to be correctly processed, all procedures should be reviewed on a regular basis (minimum annually) to ensure that there have not been changes in the administrative data that could create anomalies within the annualized data. All changes and updates should be reflected in the documentation; ensuring to keep copies of previous procedures for future reference should it become necessary.

Social Security / National Insurance Data

Social Security data typically supply monthly values for employment. As these values often fluctuate throughout the year, it is important to populate the SBR with a fairly stable and representative value. Similar to the VAT / GST data, the following steps could be followed:

1. Determine which 12-month period to be used for the calculation.
2. Determine criteria to identify records that do not belong within the time period.
3. Determine criteria to identify duplicated reportings and to select the most appropriate one(s).
4. Apply the criteria to ensure that only the appropriate records enter the calculation.
5. Annualize by taking the average of all months in the time period.
6. Document all criteria and procedures.
7. Review significant changes from previous values and unusual ratios with other available data; making appropriate updates to the data.

8. Review procedures on a regular basis and update processing as well as documentation if required (minimum annually).

Inland Revenue / Income Tax Data

Income tax data, for both corporations as well as sole proprietorships, typically supply annual values for revenue. However, there can be situations where only partial years of data might be reported or amendments have been filed. Users of this data will need to take care to ensure that the appropriate values are reflected on the SBR.

1. Determine which 12-month period is being used.
2. Determine criteria to identify records that do not belong within the time period.
3. Determine criteria to identify duplicated reportings and to select the most appropriate one(s).
4. Develop strategies to ensure that all records are handled correctly and an appropriate value will be entered into the SBR.
5. Apply the criteria and strategies.
6. Document all criteria and procedures.
7. Review significant changes from previous values and unusual ratios with other available data; making appropriate updates to the data.
8. Review procedures on a regular basis and update processing as well as documentation if required (minimum annually).

Registrar / Intellectual Property Office

Typically, the Registrar will not be supplying size measures. However, if they are able to supply size measures, the quality of the values should be assessed to determine whether they would be appropriate to add to the SBR. For example, it is possible that these values are being captured from an administrative partner that is already supplying data for maintaining the SBR. In this case, it would be better to use the data that was received directly from the original source. It is also possible that the values were estimates that were supplied when the business was registered. This type of estimate could be quite outdated or inaccurate and therefore may not be appropriate for inclusion. However, it may be a useful input for estimating a size measure that is required for stratification.

Summary

When using any administrative data, especially size measures, it is important to understand the data that have been received from the partners and ensure that it is processed to the best of the NSOs ability. Otherwise, inaccuracies can be introduced which could significantly impact outputs from the SBR. For example, statistics derived directly from the SBR could be quite inaccurate and inaccurate survey frames could significantly impact the efficiency and accuracy of surveys.