Statistical Business Register (SBR)
User's Guide

PRASC

Project for the Regional Advancement of Statistics in the Caribbean

Projet régional pour l'avancement de la statistique dans les Caraïbes

In partnership with Canada

June 2023

1. Introduction

1.1 The Project for the Regional Advancement of Statistics in the Caribbean Statistical Business Register

The Government of Canada is funding a statistical capacity building initiative for the Caribbean region, the Project for the Regional Advancement of Statistics in the Caribbean (PRASC). As part of the Business Survey Infrastructure component of PRASC, the Statistical Business Register (SBR) was developed in order to improve data quality and implement internationally recognized methods and tools across their economic survey programs. PRASC SBR is a relational database developed by Statistics Canada in response to this need. The software package and its accompanying concepts and procedures can be used by different countries as a basis for their own SBRs.

1.2 What is a Statistical Business Register?

A statistical business register (SBR) is a database that contains a complete list of businesses and related key attributes within a given country required for statistical purposes. There is also usually a user-friendly interface that accompanies the database to facilitate browsing and updating of records. SBRs can be built by combining different data sources as a base, including economic censuses or surveys of establishments. An administrative SBR combines data from different administrative sources to produce a complete, up-to-date list of businesses and business operations within a given country. The PRASC SBR is intended to be used as an administrative SBR.

1.3 Administrative data

Administrative data are files of data collected by organizations outside of the NSO. Administrative data commonly refers to data collected by other government bodies for the purpose of administering their respective government programs. For example, the files maintained in order to perform taxation of businesses within a given country can be referred to as administrative data. By entering into formal agreements with various administrative partners, NSOs can gain access to administrative data files on a predetermined basis. These files can be processed and used in statistical programs and consumed by the SBR.

Before SBR initialization, an NSO should determine what combination of administrative data would result in complete coverage of the business population in their country, as well as the frequency at which they can process and load the files. These decisions will depend on the resources available to the NSO and the selected **administrative partners** to prepare the data.

1.4 Uses of an Administrative Statistical Business Register:

- A cost-effective sample frame
- An important tool for stratification and estimation
- A central frame for integrating surveys and administrative data

- The database can be used for the direct production of demographic statistics
- Could be used as a tool for longitudinal analysis

1.5 Guide Overview

The following guide will explain all aspects of implementing, using and maintaining the PRASC SBR within a given country. It will begin by outlining SBR concepts and provide explanation of how to use the interface to browse the businesses information. It will move into performing manual updates and interpreting data extracted from the SBR. Finally, it will outline the processes behind setting up and maintaining the SBR from an administrator's perspective.

Although this guide is written specifically about the PRASC SBR, for simplicity the guide will refer generally to SBRs. This guide is written using standard terms outlined in the United Nations' Economic Commission for Europe (**UNECE**)'s guidelines on Statistical Business Registers published in 2015. It also follows international System of National Accounts (**SNA**) standards. For more detailed background on SBR concepts, please refer to documentation about these standards.

2. Concepts

2.1 SBR Coverage

The SBR was designed to meet the context of the individual country that uses it. It is recommended in the UNECE Guidelines that NSOs should aim to cover as many economic actors engaged in national production as possible in their SBRs. The SBR was designed to use administrative data as its basis. This means that at its most basic level, the units on an SBR are administrative records found on other administrative registers. These administrative units can be broken down into different components for the purpose of statistical analysis. The SBR stores business units throughout their lives as businesses, from registration to beyond their final closure.

2.2 SBR Components

The SBR uses three main building blocks to build a representative picture of businesses within a given country

2.3 Legal Component

The legal component can be thought of as a picture of business within a given country, "on paper". When a business owner registers their business name or corporation with the appropriate governing body, a **legal entity** is created. Each country will define its legal entities differently. Usually, a legal business must register with at least one administrative partner, will have an identification (ID) number assigned by that partner, a legal name and a legal type. Once registered, a legal entity has the power to negotiate contracts, assume financial obligations or pay off debts. The same legal entity can register for various government programs if they meet specific criteria, such as **National Insurance** programs if they have employees or **Tax Return** programs to file income tax or **Value-Added**

Tax programs, if they provide goods and services and earn revenues above the minimum threshold, as determined by individual governments.

Legal entities are the most reliable basis of the statistical business register. Many countries initialize their SBRs by starting with administrative data from the governing body responsible for business registration. Each legal entity registered with that governing body is created as a unit on the SBR and complemented with information from other administrative partners such as National Insurance, Tax Return etc.

2.4 Operating Component

Legal entities can conduct business in different ways. The operating component of the SBR can be thought of as the operational or physical picture of business within a given country. For example, the different restaurants, stores, offices and factories that can be found within a country.

A legal entity could operate at one or many locations and have one or many activities. To accurately reflect the diverse economic activities, a legal entity is broken down into smaller units, referred to as **operating entities**.

An operating entity is defined as a single location where a principal productive activity is conducted. On the SBR, information about a legal entity's respective operating entities is linked to the legal entity unit. When one legal entity has a single operating entity, one unit is created on the SBR with both legal and operating characteristics. This is called a **simple unit structure**. When one legal entity has multiple operating entities, separate units are created for the legal entity and each of the operating entities on the SBR. This is called a **complex unit structure**.

2.5 Statistical Component

2.5.1 Eligibility

For statistical and sampling purposes, analysts should only consider units that are part of the **eligible population**. **Statistical indicators** are assigned to all SBR units. Their primary use is to classify the business population by survey eligibility. **Only units that meet certain criteria are part of the eligible population and only eligible units have statistical indicators on the SBR.** Table 2.1 provides an example of eligibility criteria for surveys.

The **eligibility criteria** for a unit to gain statistical indicators are as follows:

- ✓ an industrial classification code (ISIC)
- ✓ an alive business status (business status=2)
- ✓ at least one recent (Less than 36 months from today's date) size variable with a value greater than zero (revenue, number of employees etc.)
- ✓ unduplicated status (prospect flag = 0)

Table 2.1: Survey Eligibility example

Record	ISIC	Alive Business Status	Any Size Variables	Prospect Flag = 0	Eligible
1	✓	✓	✓	✓	✓
2	×	✓	✓	✓	×
3	✓	×	✓	✓	×

2.5.2 Statistical Information Levels

Beyond eligibility, statistical indicators have a secondary classification function used to indicate what type of information can be collected from a unit. Business statistics are designed to provide high level information about legal entities, or lower level information about operating entities. Because of different structure types units can have on the SBR, some units will be able to provide both high-level and low-level information (simple unit structure), while units in complex structures can only provide one or the other. **Statistical indicators** are used on the SBR to signal the type of survey eligibility a unit has and the information that can be collected from it. Active legal entities are assigned **enterprise** level indicators and active operating entities are assigned **establishment** indicators.

An **enterprise** on the SBR is a statistical entity that aligns with a legally recognized entity that owns and has control over the operations of a business. Typically, enterprises are corporations, sole proprietorships or partnerships. As legal entities, enterprises are usually subject to financial reporting and can therefore make available a full set of financial statements, from which vital financial and activity information can be made available for SBR purposes.

An **establishment** is an operating unit that falls under the direct control of an enterprise. It can be a separate unit that is distinctly located from the enterprise's main unit, in which case it would be denoted by a separate record on the SBR that is linked to the parent enterprise. If not distinctly located, the establishment is integrated into the enterprise record. In either case, the establishment represents a unit of a business where economic production occurs. By definition, it does not require a full set of financial statements, but at a minimum, it must have accounting for its activities such that the data elements required to provide its revenues net of costs can be obtained. Establishments on the SBR form the basis of critical production statistics and national accounting indicators, such as the Gross Domestic Product (GDP).

Enterprise

An enterprise is a record on the SBR that enterprise level information can be obtained from.

Enterprise level information pertains to a legal entity and the whole business. Enterprise level information is information about the legal aspects of an entity such as its registered name and address, or information about the sum of its economic performance such as the enterprise level ISIC and size variables.

Establishment

An establishment is a record on the SBR that establishment level information can be obtained from.

Establishment level information is details about a specific operating entity. SBR information related to the day-to-day activities of a business include operating name, operating address, ISIC, etc.

2.5.3 Statistical Component Summary

Only businesses that meet the eligibility criteria have statistical indicators. When a legal entity is eligible for surveys, it becomes an enterprise. When an operating entity is eligible for surveys, it becomes an establishment. One unit can be both an enterprise and an establishment simultaneously and be flagged as both. Table 2.2 summarizes the differences between enterprise and establishment units.

Table 2.2: Enterprise and Establishment Statistical Flags (Survey Eligibility)

Record type	Description	Enterprise	Establishment
Legal entity	Unit represents the legal entity in a complex unit structure	✓	×
Operating entity	Unit represents an operating entity in a complex unit structure	×	✓
Legal with one operating entity	Unit has both legal and operating information (simple unit structure)	✓	✓

2.5.4 Statistical Numbers

The key identifier of the statistical component is the **statistical number**, which is a unique 9-digit number starting from 100000000 (to avoid leading-zero suppression). The statistical number is automatically assigned by the SBR to

each individual unit that is created. This number is crucial for the purpose of analysis and tabulation and is important for the purposes of secure statistical processing, communication, longitudinal analysis and record linkage.

The statistical number stays consistent despite changes to other identifying information such as the unit's name and address, as well as identifiers assigned by administrative partners. Throughout its life, an entity may open and close various accounts with other government agencies, change its name or move addresses. The statistical number ensures that a consistent, unique identifier (ID) is available for each unit throughout its life.

2.6 SBR Structure Types

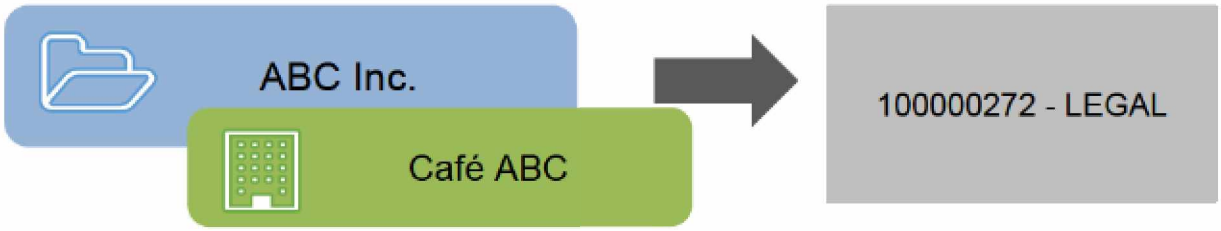
The SBR is used to represent business structures as they exist in the real world as accurately as possible. **Structure** refers to a group of related units on the SBR. Units can form structures on the SBR that are simple or complex. When simple units are eligible for both enterprise and establishment level surveys, they are referred to as **single-establishment enterprises**. More complex situations include one enterprise unit that has recorded relationships with multiple unique establishment units, which is referred to as a **multi-establishment enterprise**. It is also possible to have complex structures where multiple enterprise units that are related to one another through legal ownership, referred to as an **enterprise group**. These situations are covered more in depth through the following section.

2.6.1 Single-Establishment Enterprises

Most units on an SBR are **single-establishment enterprises** composed of one legal entity with only one operating entity. Both the enterprise and establishment information are reflected in a single unit on the SBR. An example of a single-establishment enterprise structure would be a business that operates a single store. This single unit would be eligible for both enterprise and establishment level surveys.

Figure 2.3 is an example of a single-establishment enterprise. Legal entity "ABC Inc.," operates one café in a single location (Café ABC). "ABC Inc." is therefore represented by one unit on the SBR that has both operating and legal characteristics. It is eligible for surveys seeking information at either the enterprise or establishment level.

Figure 2.3: Single-Establishment Enterprise

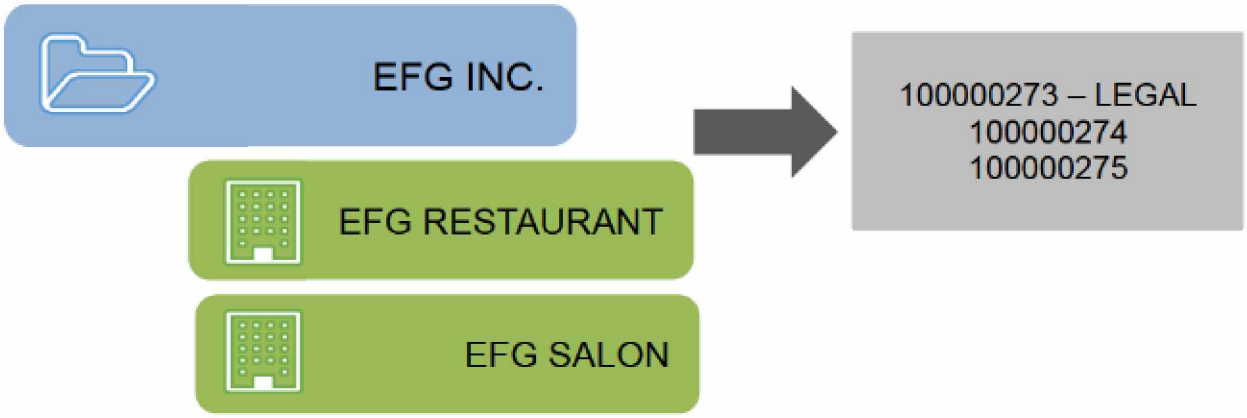


2.6.2 Multi-Establishment Enterprises

One enterprise can have two or more establishments. On the SBR these are represented by separate units. This is referred to as a **multi-establishment Enterprise.** In other words, the legal operating entity has more than one day-to-day operation or location. For example, a legal entity that operates three restaurants.

A multi-establishment enterprise is represented in Figure 2.4. "EFG Inc." is one legal entity with two unique business operations. "EFG Inc." represents the enterprise level information, and an establishment level unit is created to represent each of the business operations: one restaurant and one salon. On the SBR these would be three separate units.

Figure 2.4: Complex Operating Structure



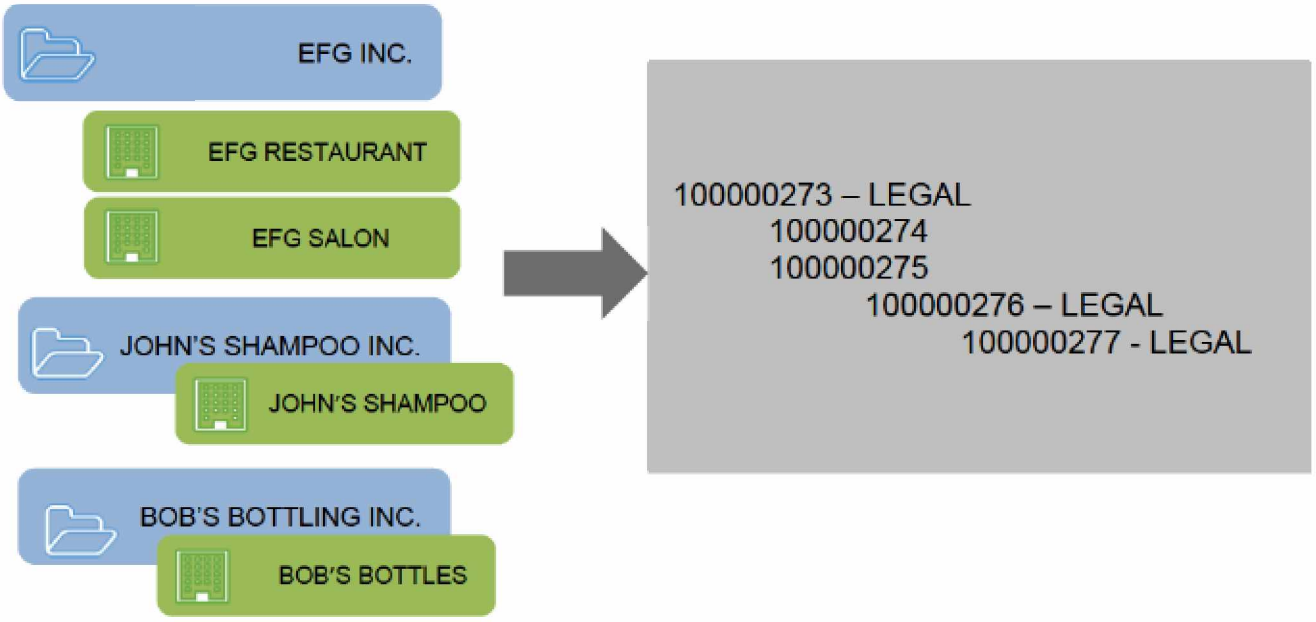
2.6.3 Enterprise Groups

An **enterprise group** is created when one enterprise is legally linked to other enterprises through majority ownership. Only incorporated legal entities can be part of a legal structure. If one enterprise owns the majority of the shares in another (more than 50% of the shares), a legal link can be created manually on the SBR. The enterprise that owns the other enterprise is referred to as the **parent**, and the enterprise that is owned is referred to as the **subsidiary** or **child**. The parent at the very top of a legal structure is referred to as the **top enterprise**.

2.6.4 Combining concepts

Figure 2.5 illustrates an enterprise group that combines multi-establishment enterprises and single-establishment enterprises. "EFG Inc." is a multi-establishment enterprise because it is one enterprise unit with two establishments. At the same time, it is a parent of "John's Shampoo Inc." "John's Shampoo Inc." is also a parent of "Bob's Bottling Inc." Both "John's Shampoo Inc." and "Bob's Bottling Inc." are single-establishment enterprises. "EFG Inc." is the top enterprise in the enterprise group.

Figure 2.5: Enterprise Group



3. BR Contents

The SBR is made up of over 100 variables, each with an intended purpose in producing high quality economic statistics. These variables can be grouped into 4 main categories: identification and contact variables, business status, stratification variables and relationship variables. The following section will provide an overview of some of the more salient variables on the SBR. For a full list of descriptions of the different variables, please refer to the SBR Data Dictionary (found within the Excel batch load template).

3.1 Identification and Contact Variables

Identifying and contact information is available for each SBR unit. Depending on the level of information the unit represents, slightly different information is given. At the enterprise level, the SBR stores legal names, legal addresses and contact information for the best enterprise level contact. At the establishment level, the operating name, operating address and a specific contact for that operation are stored.

3.2 Business Status

To facilitate longitudinal analysis, the SBR stores units throughout their life cycle and beyond. The **business status** is a variable used to classify the unit based on its legal or operational status. The business status effective date is also captured, referred to as the **business date**. This date should reflect the date the most recent status change occurred.

At the enterprise level there are eight different business statuses, as explained in Table 3.1. At this level, the business status relates to the status of a legal entity. Not all enterprises follow a linear path through these business statuses, but they do fall into one of these categories at any given time.

Table 3.1: Enterprise Level Business Statuses

Business Status	Description	Eligible for Surveys
New Business Business Status=1	Unit received from administrative data, but not yet alive, has no size variables	No
Alive Business Status=2	The enterprise is engaged in one (or more) economic activities and has current size variables	Eligible if has ISIC, a recent size variable and no duplicate on the frame (Prospect Flag=0)
Inactive Business Status=3	Enterprise has stopped all economic activities but still exists legally, has outdated size variables	No

Business Status	Description	Eligible for Surveys
Bankrupt Business Status=4	Formal legal process has granted the legal entity bankrupt status	No
Dissolved Business Status=5	Formal process has been undertaken to give up the legal status of the entity	No
Integrated Business Status=7	Status used by SBR updaters to indicate that this unit was a duplicate of another unit. The SBR updater has reconciled the two units, selected one to continue and one to cease. The unit with integrated status is the ceased unit.	No
Amalgamated Business Status=8	Formal legal process has granted the legal entity amalgamated status. This means that its operations have merged and are continuing under another legal entity	No

At the establishment level, there are two business statuses as summarized in Table 3.2. Operating entities are either active in economic production or not:

Table 3.2: Establishment Level Business Statuses

Business Status	Description	Eligible for Surveys
Alive Business Status=2	The establishment is actively engaged in producing goods and services	Eligible if has an ISIC code and a recent size variable
Ceased Operations Business Status=6	The establishment is no longer active or in operation	No

3.3 Stratification Variables

3.3.1 International Standard Industrial Classification (ISIC)

ISIC is used to classify the business activities of individual establishments, and the overall activities of a business at the enterprise level. The specific version of ISIC utilized may differ between countries, as well as between NSOs and administrative partners within the same country. ISIC is the basis for producing economic statistics, demographic statistics or selecting a sample from a population, which is why it is referred to as a stratification variable. It is a way to split up the population based on the individual characteristics of each unit. Currently the SBR is set up to accommodate ISIC Rev. 4. For more information

about ISIC Rev. 4, please refer to documentation prepared by the United Nations Statistics Division (UNSD). The administrator of the SBR with the collaboration of the database administrator can customize the version of ISIC, if needed.

At the enterprise level, an ISIC code is assigned that represents the dominant activity across all of the enterprise's operations. This is usually the primary activity or the ISIC that generates the most revenue. An ISIC code assigned at the establishment level is specific to that operating entity's activity.

A multi-establishment enterprise could have establishments with various ISIC codes, but the enterprise level unit would be assigned an ISIC that reflects the main activity of the all of its establishments taken as a whole. For example, if "CDE Inc." operates two restaurants and one hotel, but the most revenues are generated by the hotel and this is seen as the main activity, the enterprise level unit would be assigned the hotel ISIC code.

3.3.2 Size Variables

Another stratification variable used to segment the population is size variables. Size variables are also useful when using SBR data to produce economic estimates and demographic information. Depending on the administrative data available in each country, the SBR can accommodate various size variables such as revenues, sales, number of employees, gross wages, etc.

At the enterprise level, the size variables should reflect the performance of all the enterprise's establishments. The sum of size variables from the establishments should be consistent with the revenues on the enterprise level unit.

3.4 Record Linkage Variables

3.4.1 Administrative Identifiers (Administrative Keys)

Depending on the administrative sources associated with a given unit, the SBR can store the appropriate identifiers associated with these units from the various administrative databases. Maintaining the concordance between this information is crucial for avoiding duplication.

3.4.2 Relationship Variables

The SBR also records the relationships between different units within the SBR. This information is displayed in the structure viewer portion of the SBR interface. There is no relationship variable displayed on the SBR interface apart from the structure viewer, but this information can be obtained from the Survey Universe Files (SUF) – explained in section 6 – extracted from the SBR.

3.4.3 Prospect Flag

The Prospect Flag indicates that a legal entity was birthed as part of the non-primary administrative data source load. This functionality gives users a

mechanism to investigate the births to minimize the possibility of duplicates on the database.

For example, if your primary data source is the registrar file, consider a business that was loaded as a legal entity with RegistrarID 12345 and a legal name of 'Resort Beach' on the SBR, as indicated in Table 3.3 below.

Table 3.3: Example Registrar Data

RegistrarID	VAT_ID	LegalName
12345		Resort Beach

Then, a second file (non-primary file) is used to update the SBR from a VAT source with a VAT_ID 9999 and a legal name of 'Beach Resort' without a Registrar ID or after an unsuccessful matching process (see Table 3.4 below).

Table 3.4: Example VAT Data

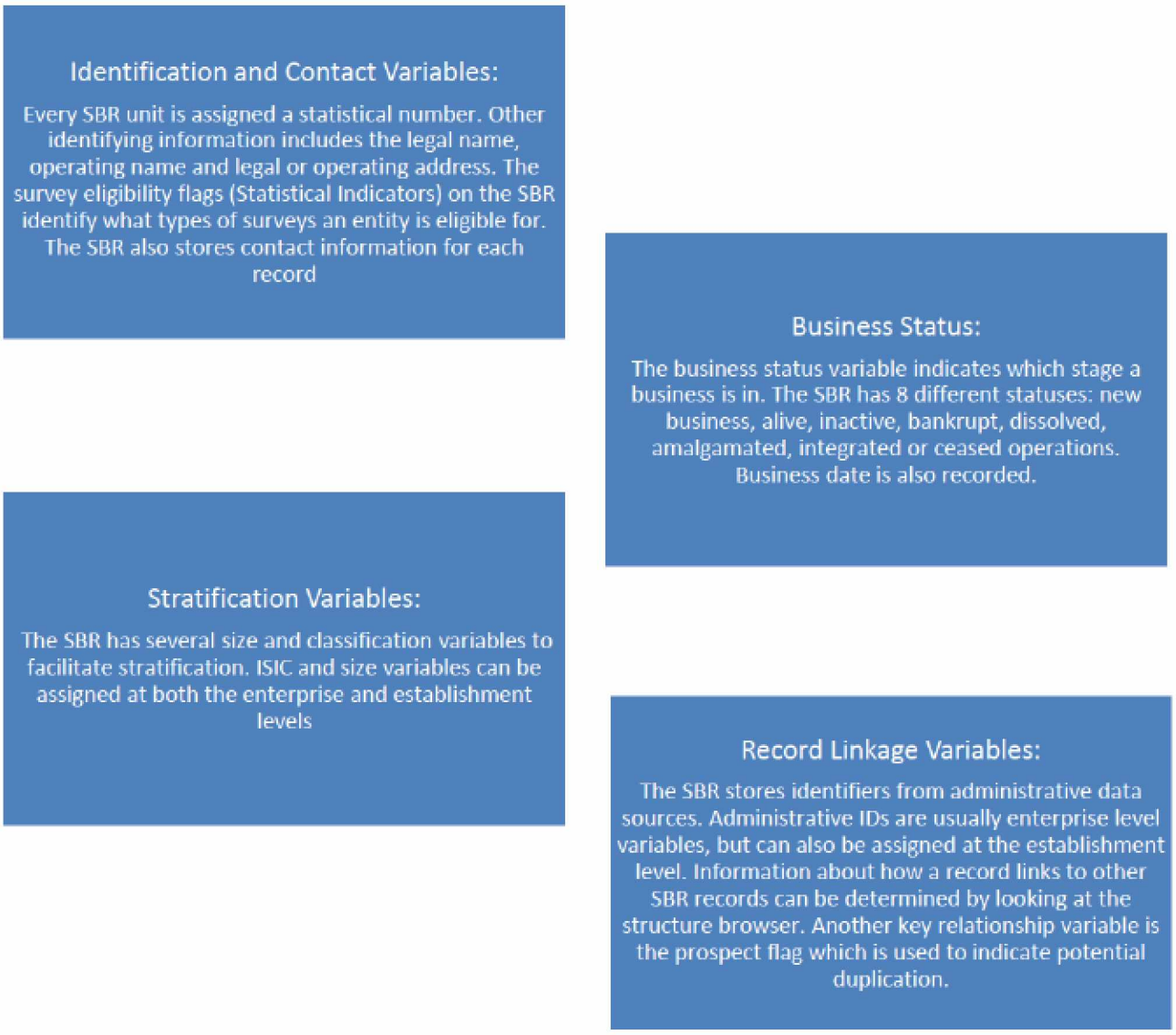
RegistrarID	VAT_ID	LegalName
	9999	Beach Resort

Each of these units in each of these files represents the same actual business. The second administrative source does not have a RegistrarID that corresponds to the first source and were unable to correctly match the legal name from both data sources. Without the RegistrarID on the second file to act as a link between the new file and the existing BR, the administrative process will birth a second unit to the SBR. The second unit would be a duplication of the first. To prevent such cases of duplication, the prospect flag must be manually updated (prospect flag=1) during the pre-processing step when birthing units from administrative sources to flag these types of records for review.

3.5 Content Summary

Figure 3.5 summarizes the type of content available from the 4 main information groups.

Figure 3.5: Summary of SBR content



3.6 Information available by unit type

Recall from section 2:

- Enterprise and establishment are two different statistical information levels
- Enterprise level units have information pertaining to the legal aspects of an entity and the sum of its operations (legal entity)
- Establishment level information pertains to information about the day-to-day operations of one location of an enterprise (operating entity)
- Single-establishment enterprises represent both enterprise and establishment in one unit
- Units can also be multi-establishment enterprises which means there is one unit that stores enterprise level information and multiple units that store information pertaining to each separate establishment

Table 3.6 summarizes which variables pertain to each level of information. Enterprise-only units store only enterprise level information, establishment-only units only have establishment level information and simple units store both levels of information.

Table 3.6: Information by Unit Type

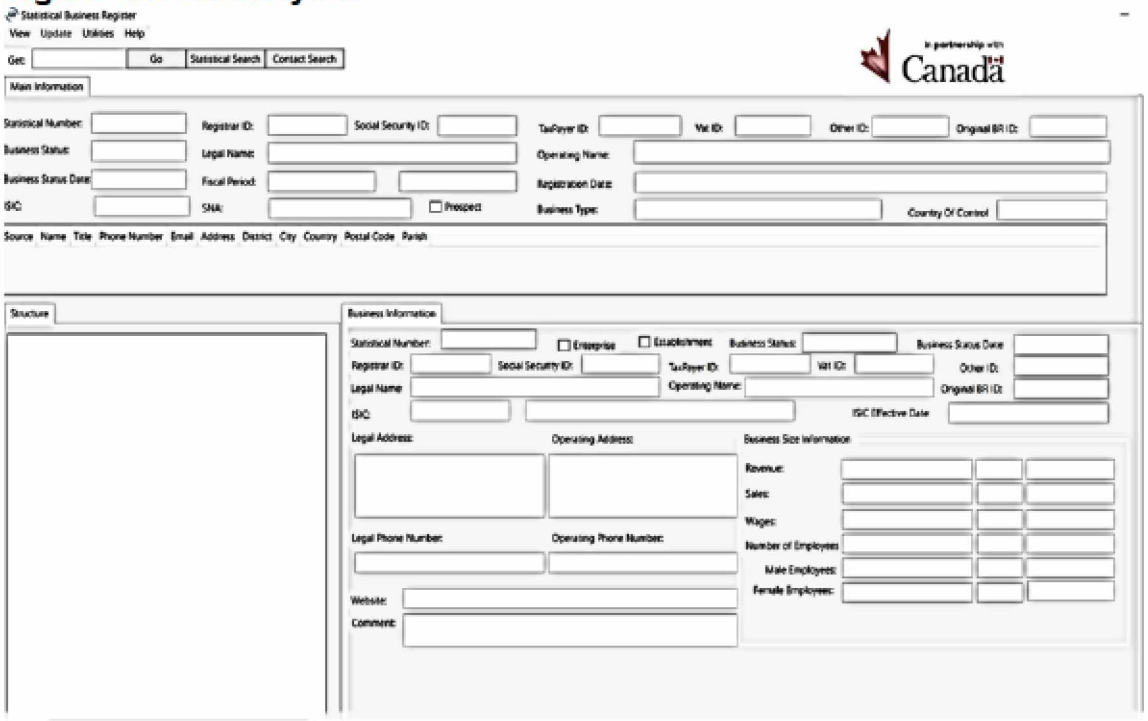
Variables	Enterprise (legal entity/business as a whole)	Establishment (specific to one operating entity)
Identification and Contact Information Variables		
Statistical Number	Y	Y
Legal Entity Information (name, address, telephone number, etc.)	Y	N
Operating Entity Information (name, address, telephone number, etc.)	N	Y
Contact Information	Y	Y
Survey Eligibility (Statistical Indicators)	Y	Y
Business Status		
Business Status	Y	Y
Stratification Variables		
ISIC and ISIC Description	Y	Y
Size Variables	Y	Y
Record Linkage Variables		
Administrative IDs	Y	Y
Relationship Variables	Y	Y
Prospect Flag	Y	N

4. Navigating the SBR

4.1 The SBR Layout

The SBR information is organized into a few main sections illustrated in Figure 4.1. The **menu** at the very top of the window allows a user to execute different functions depending on their permission level. The **search bar** allows a user to pull up a specific unit, or execute a statistical or contact search. The **structure browser** allows the user to look at the structure a unit falls within and toggle between units. The **main information** tab contains enterprise level information (legal information). The **business information** tab at the bottom right of the interface contains unit-specific information about the selected unit. It will display either the enterprise level or establishment level information depending on the type of unit selected in the structure browser.

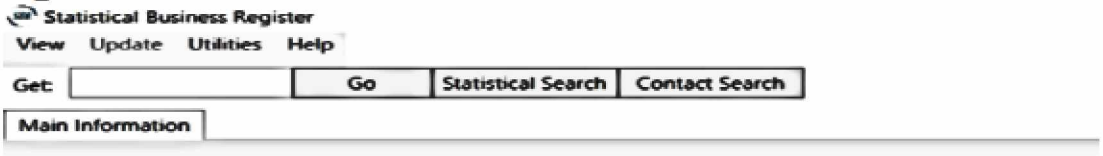
Figure 4.1: SBR Layout



4.2 Menu

At the top of the SBR interface, there is a toolbar that can be used to perform several functions. Under the “view” menu, all users can access the **survey universe tool, the logs** as well as the **demographic reports**. Under the “updater” menu, users with the necessary permissions can perform operating entity or legal entity updates. Under “utilities”, SBR administrators can manage user permissions, generate survey universe files (SUF), run inactivation, activation and update statistical processes. The “about” item in the “help” menu displays a pop-up window that shows the current version code of the SBR and its release date.

Figure 4.2: SBR Menu



4.3 Search Bar

The search bar (shown in Figure 4.3) allows a user to access a specific unit by typing in its 9-digit statistical number and clicking "go". The search bar also allows users to execute a statistical search or a contact search.

Figure 4.3: Search Bar

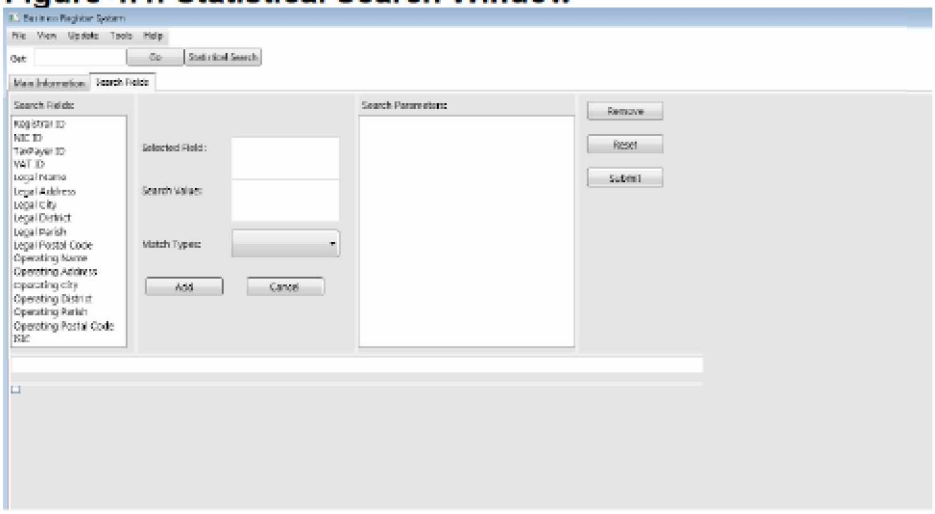
4.3.1 Statistical Search

The **statistical search** function is intended to assist with locating units that match a given search criteria. To execute a search, insert parameters by selecting a variable and adding a search value. There are four different **match types** that can be selected: starts with, contains, ends with and exact. Once the parameters have been entered, clicking submit will search the database for the appropriate matches for the search string and match type selected. The search function searches for exact matches to the search string, meaning that regardless of the match type selected any variations in spelling between the search term and the field on the database will not yield a match. If the search does not yield the intended results, re-attempt with a smaller search string.

For example, using the search term "ABC Ltd." and match type "exact" will not return the value "ABC Ltd," as the period after Ltd is not included in the legal entity's registered name. It is better to execute a search for the string "ABC" and match type "Starts with" to avoid excluding any potential matches.

Once a search has been executed, the search results will be displayed at the bottom of the window. If a user double clicks on a searched result unit, the entity will be loaded on the SBR browser tab.

Figure 4.4: Statistical Search Window



4.3.2 Contact Search

Contact search functions in the same manner as the statistical search, except that it searches for the business contact information that match the parameters entered. When working with names it is important to be mindful of inconsistent spelling. Entering only the first few letters of a name, rather than a full name, could yield more results. Searching for someone using a static identifier such as a telephone number, is another method to retrieve the desired contact person.

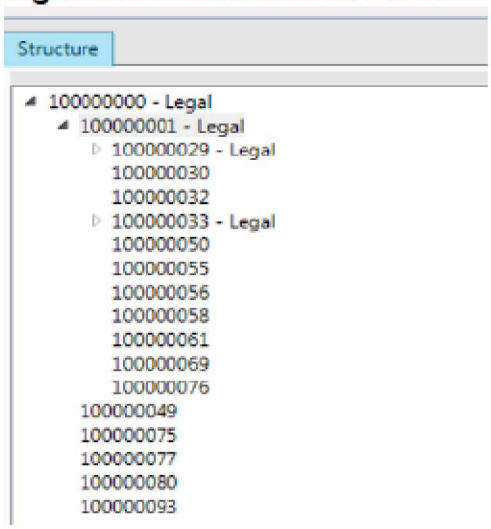
4.4 Structure Browser

On the left-hand side of the SBR there is a structure window. This window displays the full enterprise group (comprised of both legal and operating structures) of a given unit. This information is displayed in a hierarchical “tree” format using the statistical numbers of each unit to demonstrate parent-child links.

The structure browser allows users to browse through the different units within an enterprise group. Clicking on another unit within the structure will prompt the SBR to load the information for that unit. The entity selected is always highlighted (in light grey) in the structure browser.

Parents are listed above their children as demonstrated in Figure 4.5. The legal structure is made up of 100000000 as the parent of 100000001 which in turn is the parent of 100000029 and 100000033. These enterprise units are marked “legal”. Each of these enterprise units are also parents of several establishments which in turn are not marked as legal entities.

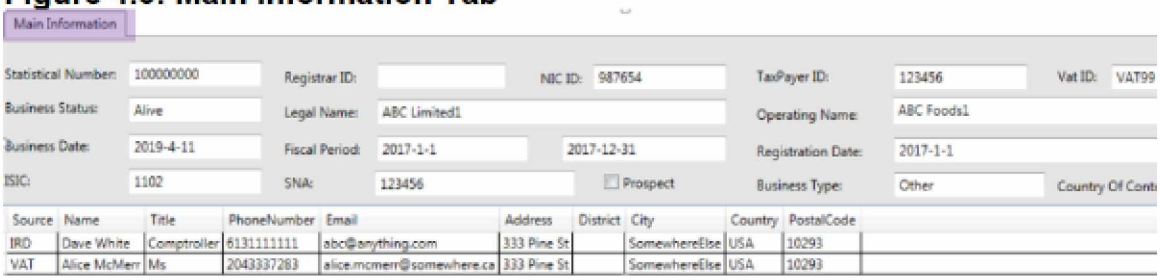
Figure 4.5: Structure Browser



4.5 Main Information

The main information tab towards the top of the interface shows enterprise level information about the unit itself or if the unit is an establishment-only unit, enterprise level information from its parent. Figure 4.6 is an example made with fake data that illustrates the type of information displayed in this part of the SBR interface. Keep in mind that if the unit selected is an establishment-only unit, the information displayed in the main information tab will pertain to the enterprise-level parent of that unit. The bottom of the main information tab shows the detailed contact information with the data source.

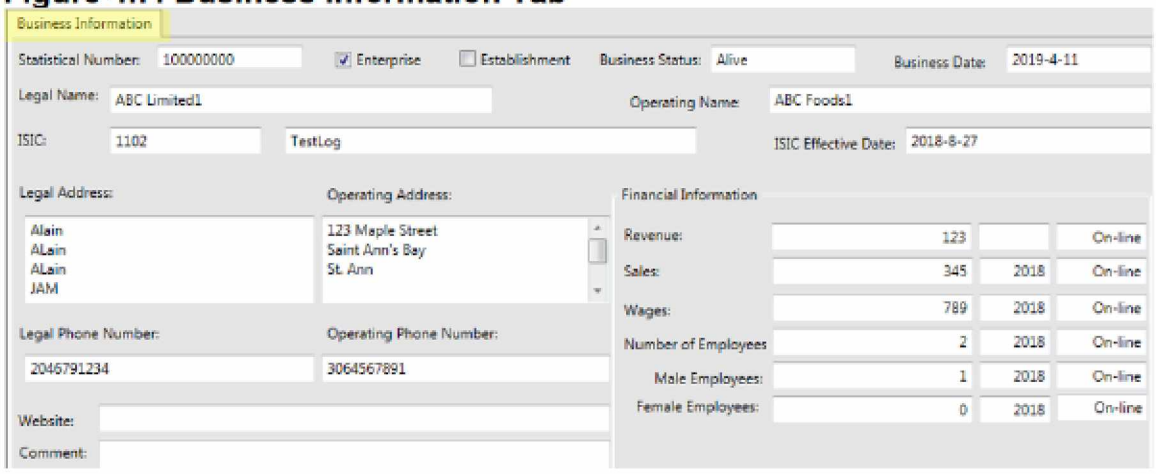
Figure 4.6: Main Information Tab



4.6 Business Information

The business information tab towards the bottom-right of the interface shows information specific to the unit that has been selected. It will indicate the statistical information level that unit represents, the specific ISIC code of that unit, its operating address and its financial information. On an enterprise level unit, this information will pertain to the business as a whole. On an establishment-only unit this information will pertain only to that specific activity.

Figure 4.7: Business Information Tab



5. Updating the SBR

This section of the document is intended to assist users with making updates to the SBR using the commands found under the "update" header on the menu. This section will also provide general best-practices for making SBR updates. The availability of administrative data files and the data requirements of individual countries will vary. As such, individual countries should define their own needs and provide additional guidelines to SBR updaters, as needed.

Under the update menu, users should see six different options: create entity, update, cessation, reactivation, inactivate legal and update relationship links as is displayed in Figure 5.1 and 5.2. The following section will provide some general update considerations, then it will explain each of these functionalities in order.

Figure 5.1: Update Menu

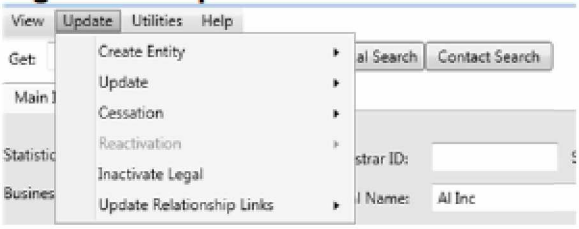
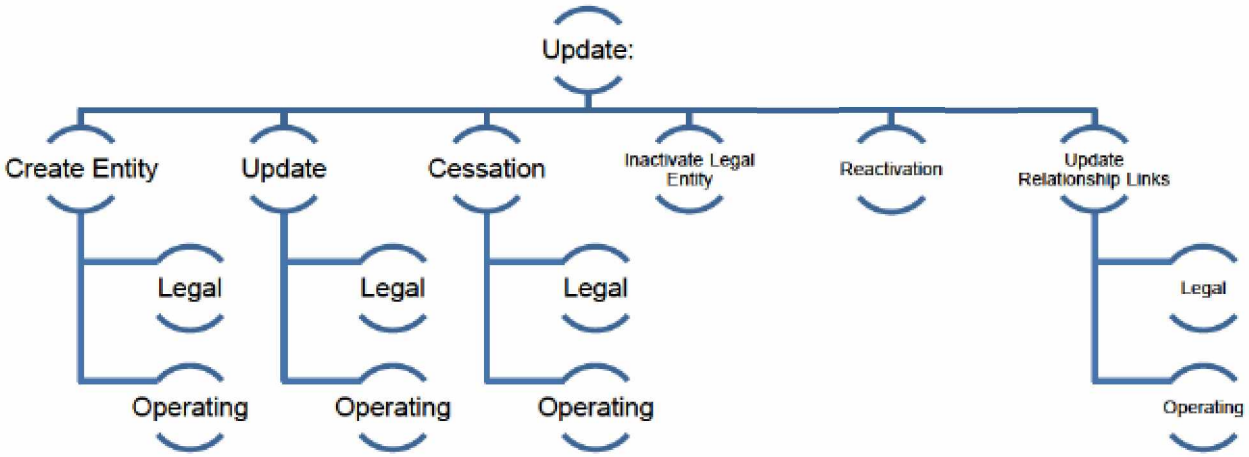


Figure 5.2: Manual Updates Chart Under Updater Menu



5.1 General Update Considerations

When making updates to an SBR, it is important to verify any changes using accurate, reliable, and up-to-date information. Good sources of information to update the SBR include information received directly from a business representative, administrative data partners, news articles or company websites.

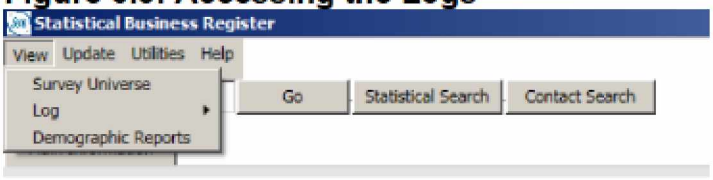
Information received directly from the respondent or from an administrative source are generally more reliable than news articles and websites.

It is also important to confirm that the information collected is as recent as possible. A good rule of thumb would be to only make changes where the information is confirmed to be newer than the last time a unit was updated on the SBR. The last update date can be confirmed by looking at the logs.

5.1.1 Logs

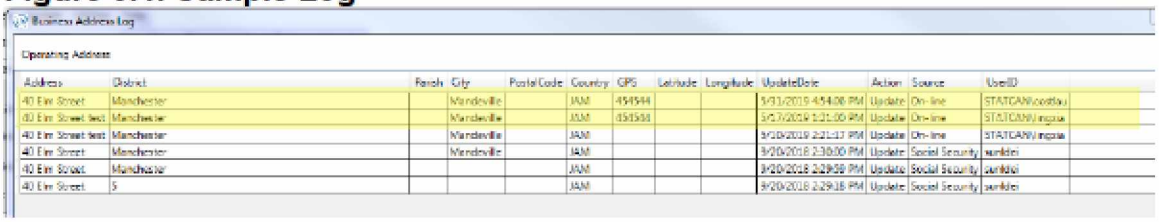
A log provides historical data on certain variables in the SBR. Browsing a unit's logs offers a summary of changes that have been made to that entity. Users can access the logs under the "view" menu as shown in Figure 5.3.

Figure 5.3: Accessing the Logs



The information available in a log includes: the changes made to a variable, the update date, the source of the change (was the change made via administrative data updates or by an SBR user) and the identity of the updater who made the change. Entries will be arranged in order of update date, with most recent changes appearing at the top of the log. To determine the change being made, compare the top two lines. The second line shows the value of the variable before the change was applied, and the top line shows the current value. Figure 5.4 provides an example of how the log information is displayed.

Figure 5.4: Sample Log



1.1.1.1 Which variables have logs

The variables for which logs are available are grouped into two categories: operating (unit information) and contact. The variables within each category are summarized in Table 5.5

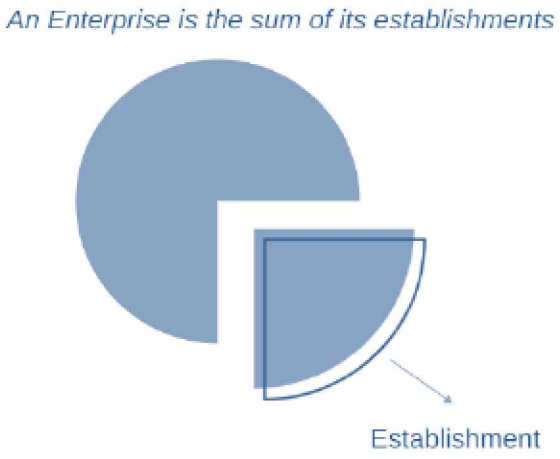
Table 5.5: Variables with Logs

Operating Information	Contact Information
Admin Keys Business Names Business Addresses Phone Numbers Size Information Business Statuses Business Type Comment Fiscal Dates ISIC Prospect Flag SNA Institutional Sector Statistical Flags Structure Websites *(includes both enterprise and establishment level information)*	Contact Name Contact Phone Number Contact Address Contact Email Contact Information Source

5.1.2 Coherency

Another general rule of SBR updates is coherency. Coherency refers to having consistency in values between enterprise and establishment units within a multi-establishment enterprise structure. The specific rules for coherence may vary from country to country. For example, the enterprise level ISIC should be the dominant activity of its operating structure. In terms of size variables, the size variables recorded on the enterprise unit should be the sum of the size variables recorded on its corresponding establishment units. As an example, the revenue on the enterprise unit should be the sum of revenues of each of the establishments. The establishment(s) are the pieces that make up a whole enterprise. This can be thought of as a pie; the establishments are slices and the enterprise is the whole pie.

Figure 5.6: Enterprise and establishments as a pie



5.2 Creating an Entity

5.2.1 Create Operating Entity

Most structures on an SBR are single-establishment enterprises. This means that one enterprise and one establishment are represented by the same unit. Multi-establishment enterprises are manually created on an SBR when an enterprise has more than one establishment fulfilling criteria outlined in Figure 5.7.

Figure 5.7: Criteria for Creating an Establishment

Dedicated capital	Dedicated process	Dedicated labour (employees)
• There must be some kind of investment in the location the establishment operates from such as equipment, buildings, inventory, etc.	• There needs to be a business activity carried out at the establishment. • There needs to be a combination of resources (like capital and labour) coming together to produce either a good or service.	• *Usually*, the establishment has one or more employees who report to it as their primary place of employment (but this is not a strict requirement for creating an establishment, as there can be exceptional cases of facilities that have no onsite workers for which economic data must be obtained).

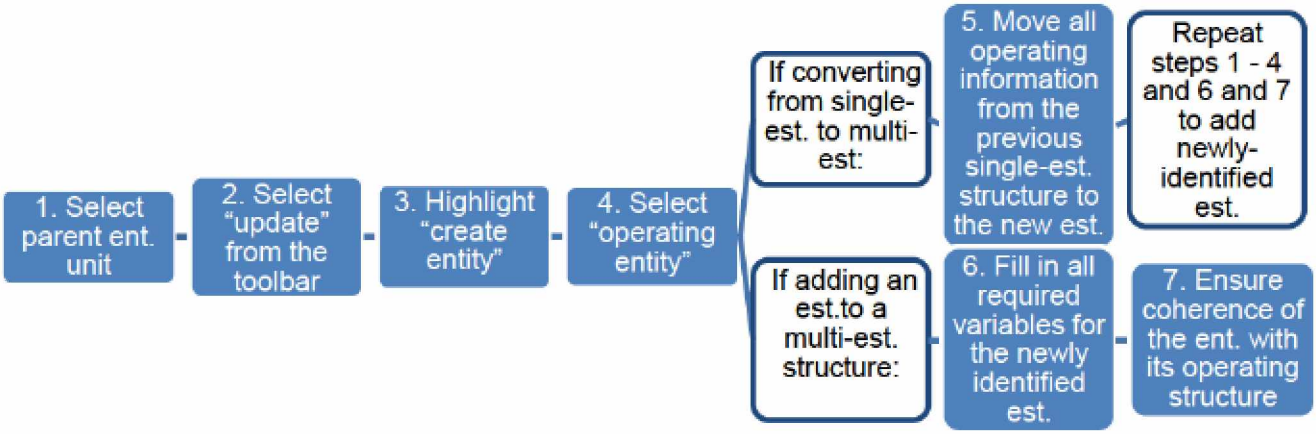
When **adding** an establishment to **an existing** single-establishment structure there are a few extra steps to be kept in mind (converting from single-establishment to multi-establishment)

- Create a new establishment under the enterprise to represent the existing establishment (operating activity of the legal unit) in the single-establishment structure

- Move any operating information from the enterprise record to the newly-created establishment.
- Add any additional establishment records required to accurately reflect the operating structure reporting to the enterprise record.
- When a new establishment is added to **an already** complex structure (multi-establishment structure to multi-establishment structure), it can simply be added to the existing operating structure without extra steps beforehand.
- After any adjustments to the operating structure, the updater should always review the information on the enterprise record to ensure coherence with its operating structure. Some data attributes associated with the enterprise unit such as enterprise revenues, employment or industry code may need to be adjusted to ensure that the operations of the newly added establishment are accurately reflected in the enterprise-level values. Coherency rules will vary from country to country.

The difference between adding an establishment to a single-establishment enterprise and adding an establishment to a multi-establishment enterprise is illustrated in Table 5.8. In both situations, you will need to enter an operating name, ISIC code, ISIC description, at least one address variable and at least one size variable in order to create an establishment. Otherwise, the update will be rejected by the system.

Figure 5.8: Adding Newly-Identified Establishments
Single-establishment enterprises versus multi-establishment enterprises



Summary

1. Select parent enterprise unit
2. Select "update" from the toolbar
3. Highlight "create entity"
4. Select "operating entity" (establishment)

If converting single-establishment to multi-establishment, proceed with step 5, if not skip directly to steps 6 and 7.

5. Move all operating information from the previous single-establishment structure to the new establishment, skip back to step 1 to add additional establishments.
6. Fill in all required variables for the newly-identified establishment.
7. Double-check coherence of the enterprise structure, ensuring to make any required adjustments to either enterprise-level or establishment-level information.

5.2.2 Create Legal Entity

5.2.2.1 Parent, Subsidiary or Stand Alone

Legal entities are usually created on the SBR automatically through the loading of administrative data. That said, legal entities can also be created manually on the SBR. The updater can choose to create a new legal entity as a parent or a subsidiary of an existing enterprise. An updater can also choose to create a **standalone entity** if the enterprise should not be part of an enterprise group.

Legal entity creation should be reserved for cases where it is known that a registered legal entity exists, but it has not yet been identified in the administrative data being received. Legal entity creation can also be used in cases where there is no registered legal entity, but a business is still operating (informal sector – see following section). Before creating anything on the SBR, the updater should ensure that the legal entity does not already exist by performing a statistical search. When creating an enterprise manually, the updater should include any possible administrative IDs so that the record can be updated through future administrative data loads.

Steps for updating

- Do a statistical search for an existing unit
- Select "Update"
- Select "Create Entity"
- Select "Legal Entity"
- Choose to create a, "Parent", "Subsidiary" or "Stand Alone" entity
- Enter all mandatory information to create legal entity, (Legal Name, Business Type and at least 1 legal address field). Do not forget to include any known administrative IDs
- Add any enterprise level information that is known (legal name, legal address, fiscal year, enterprise level ISIC, size variables)
- Add any establishment level information that is known (operating name, operating ISIC, operating address)
    - If the legal entity should be represented as a multi-establishment enterprise, this information can be left blank on the enterprise unit and separate establishments can be created following the creation of the enterprise unit

5.2.2.2 Informal sector

Some countries may want to keep units on the SBR of **informal** businesses that are not registered with any administrative partners, but still known to engage in economic production within a given country. These units can still be created on the SBR manually by an SBR updater as legal entities following the procedures outlined in the "create legal entity" section. They can be identified as informal units because no administrative IDs are entered on the unit because it is not registered with any administrative partners. These informal units will be coded as a business type = 10 for informal business on the SBR.

Please note that the informal businesses can also be created in batch mode on the SBR by using the Excel Batch Load template. As an example, a sequential number can be assigned to Other_ID as a unique business identification number. A reference period and a legal name will have to be supplied. For further information on the topic, please refer to Appendix F.

5.3 Update an Entity

5.3.1 Update Operating Entity

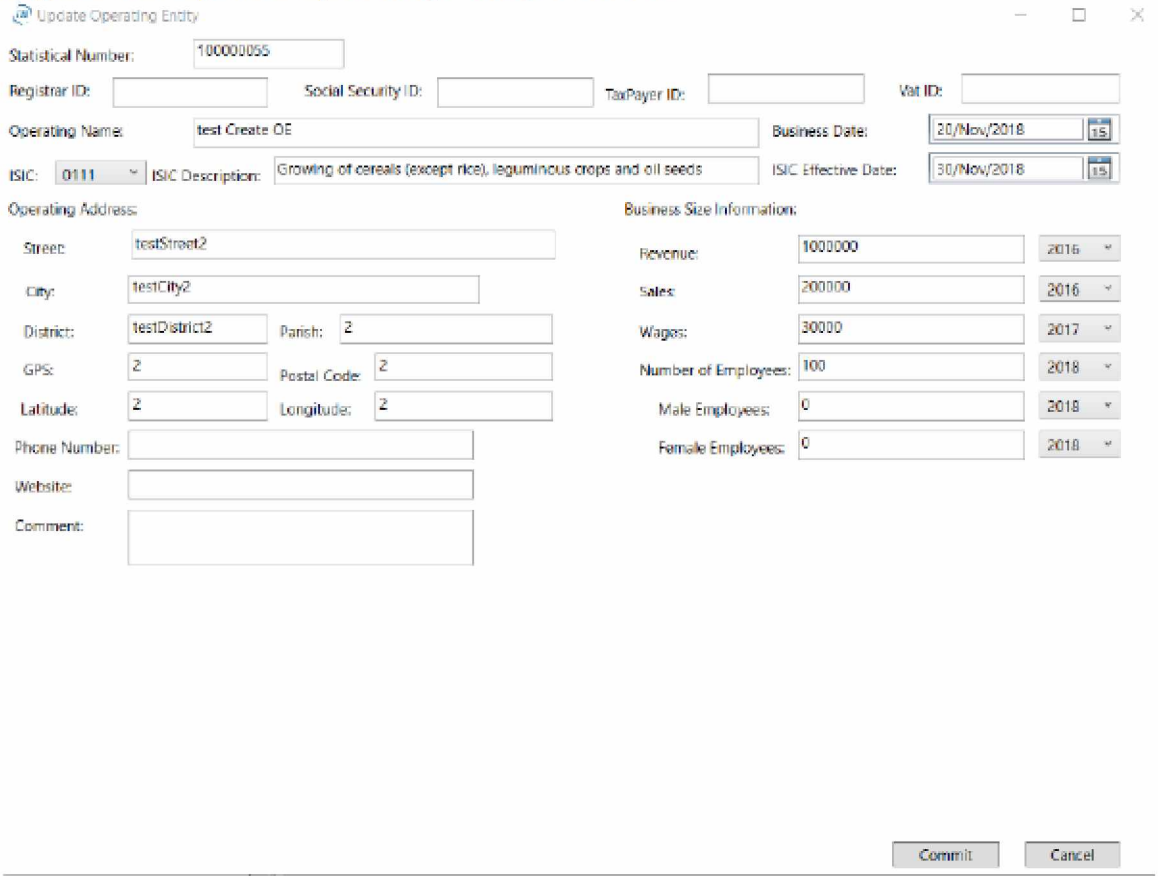
On the SBR, an operating entity refers to an establishment-only unit of a multi-establishment enterprise. The operating entity stores establishment level information such as the operating name, operating address and ISIC codes. If an enterprise has multiple establishments updaters need to ensure the information is coherent between the establishments and the enterprise unit.

Steps for updating

An establishment level unit must be selected to update operating information

- Select “update” from top menu
- Select “operating entity,” an update window will appear show in Figure 5.9
- Select field to update
- Select “Commit” at the bottom of the window to confirm changes

Figure 5.9: Update Operating Entity Window



5.3.2 Update Legal Entity

On the SBR, legal entities include both units that store enterprise level information only, as well as single-establishment enterprises which store both enterprise and establishment level information. Selecting “Legal Entity” under the

"Update" menu item will allow you to make modifications to both levels of information on the same unit. On single-establishment enterprise units, the information on the legal entity unit should reflect both the enterprise and its corresponding establishment.

On enterprise-only units found in multi-establishment enterprise structures, only enterprise level information should be stored on the enterprise unit. Enterprise level information includes enterprise level stratification variables like ISIC and legal information. Legal variables include the legal name of an entity, the legal address and the administrative identifiers assigned by other government entities at the enterprise level. The best source of legal information is the governing body responsible for granting legal status to business organizations. When modifying legal information on the SBR, future unit linkage and duplication are concerns. **To prevent problems, enterprise-level legal information should only be updated if changes are confirmed by the responsible party.**

Another important consideration when updating enterprise-level information on multi-establishment enterprises is coherency. The rules for coherence between the establishments compared to the variables on their enterprise counterpart can vary from country to country, but in general it is best practice to ensure the information is consistent throughout the structure. For example, the enterprise level ISIC should be the dominant activity of its operating structure. In terms of size variables, the size variables recorded on the enterprise unit should be the sum of the size variables recorded on its corresponding establishment units.

Steps for updating

To make updates to legal information, an enterprise level unit must be selected. An establishment level unit must be selected to update operating information

- Select "update" from top menu
- Select "legal entity"
- Select field to update
- For single-establishment enterprises both enterprise and establishment information can be modified
- For multi-establishment enterprises, the enterprise unit should be coherent with the rest of the structure, any establishment level information should be moved to its establishment units

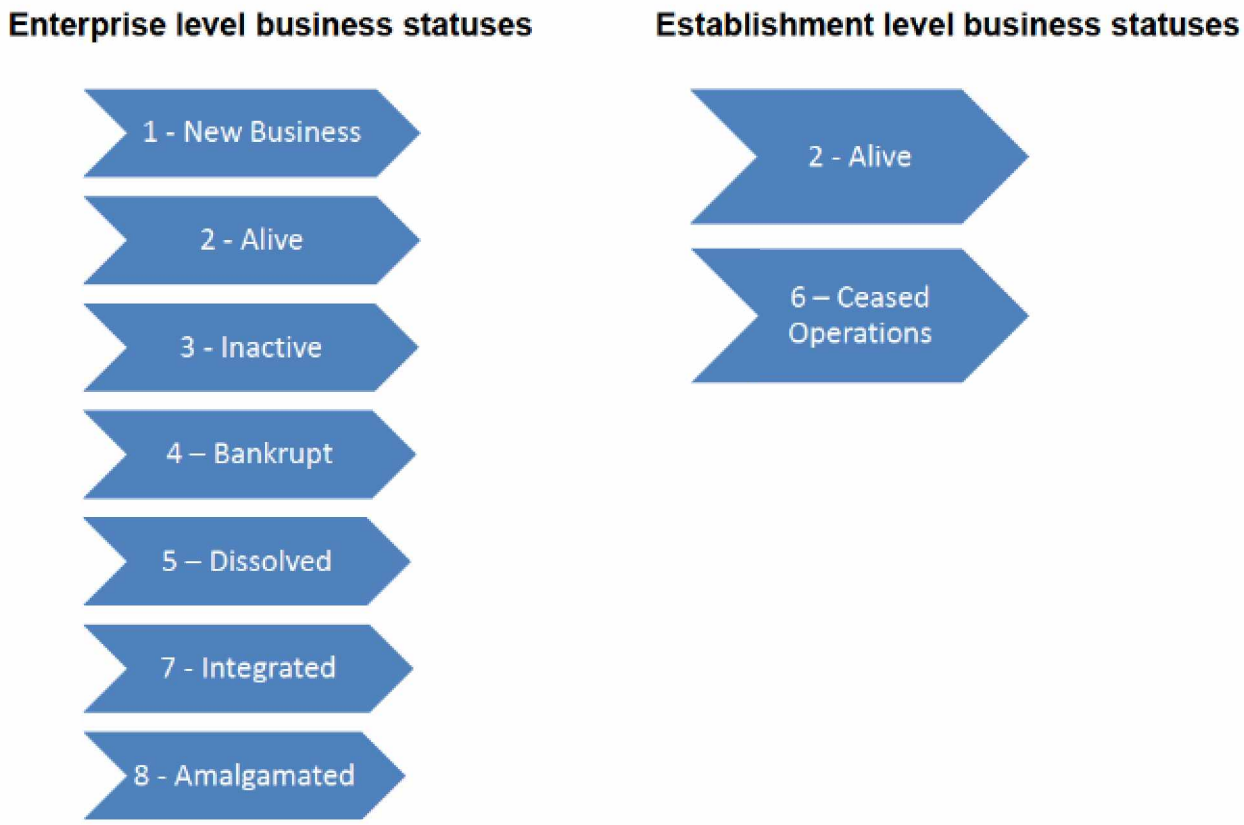
5.4 Business Status Updates

The business status of an SBR unit reflects the actual operating or legal status of an entity. The possible business statuses of an entity depend on the type of unit. Enterprise business statuses pertain to the legal entity or whole business. Establishment level business status refers to the status of an operating entity. Figure 5.10 lists the business statuses.

On the SBR, most business status updates are processed automatically based on administrative data. When needed, business statuses can also be processed manually. When a business status update is made, the "Business Date" should also be updated with the **effective date** of the change. The effective date should reflect the date that the change to the business status is presumed to have happened based on information from various sources.

Updaters should also be aware that when updating the business status of a unit, other variables may be affected or need updating.

Figure 5.10: Business Statuses on the SBR



5.4.1 Cessation

Cessation is a term that refers to a legal or operating entity that is definitively closed, this is the final "life stage" of a legal or operating entity. There are four possible enterprise level statuses that fall under cessation: bankrupt, dissolved, integrated and amalgamated. The establishment level status "ceased operations" is also found under cessation.

5.4.1.1 Cease Legal Entity

The business statuses of Bankrupt and Dissolved and Amalgamated are legal statuses that can only be applied to enterprise level units. These are all legal statuses granted to legal entities by the registration authority. This information should be confirmed by an administrative source before making changes on the SBR. With an amalgamation update, updaters should record the **continuing legal entity** in the comments and update the operating structure of the continuing entity if required to reflect any new establishments it may have inherited from the amalgamated legal entity or entities.

Integrated is a status used by SBR updaters to indicate the unit is a duplicate (this will be discussed further in the "Duplication" section).

Multi-establishment enterprises cannot be ceased. The updater will need to cease all establishments following the procedures outlined in the "cease operating entity" section before proceeding to cease the enterprise level unit.

Steps for updating

To make updates to legal information, an enterprise level unit must be selected. A multi-establishment enterprise will require updates to its operating structure before it can be ceased.

- Expand "update" from top menu
- Expand "cessation"
- Choose "legal entity"
- Select appropriate update
- Update effective date to reflect date change occurred
- If processing an amalgamation:
    - Update comment on "amalgamated" legal entity to specify the statistical number of the continuing legal entity
    - Update continuing legal entity accordingly
- If processing an integration:
    - Update comment on "integrated" legal entity to specify the correct unit

5.4.1.2 Cease Operating Entity

There are several important considerations when removing establishments from a multi-establishment enterprise. The main question an updater should have in mind when ceasing an establishment is whether just the one establishment has closed and the enterprise is continuing with different establishments or if the whole enterprise is now inactive or ceased.

General steps for updating operating entity to "ceased operations"

To make this update, an establishment level unit must be selected

- Expand “update” from top menu
- Expand “cessation”
- Choose “operating entity”
- Update effective date to reflect date change occurred

In cases where the entire business or company has stopped operating, the updater should:

- Update the business status of each of the establishments to ceased
- Update the business status of the enterprise to inactive, dissolved or bankrupt depending on the circumstances

In cases where one establishment has closed but there are still two or more establishments under the enterprise:

- Update the business status of the ceased establishment accordingly
- Make any changes required to the enterprise to ensure it is coherent with the remaining operating structure (double-check ISIC, ensure the financials and number of employees are in line with what is on the remaining establishments)

In cases where one establishment has closed and only one establishment remains:

- The updater will need to convert the structure back from a multi-establishment enterprise to a single-establishment enterprise
- Cease the closed establishment
- Move all the continuing establishment level information up to the enterprise unit (operating name, operating address, operating size values and operating ISIC)
- Cease the other establishment unit
- Double check the enterprise unit now has enterprise and establishment flags

5.4.2 Reactivation

A reactivation permits updaters to manually change the business status of a legal entity from dissolved, bankrupt or inactive back to inactive or alive. Some reasons an updater might want to make this type of change is if they have confirmed the legal entity still exists or is still operating.

This process cannot be done with operating entities, once they have been ceased they cannot be reactivated. If an establishment is ceased in error, it should be recreated manually.

Steps for updating

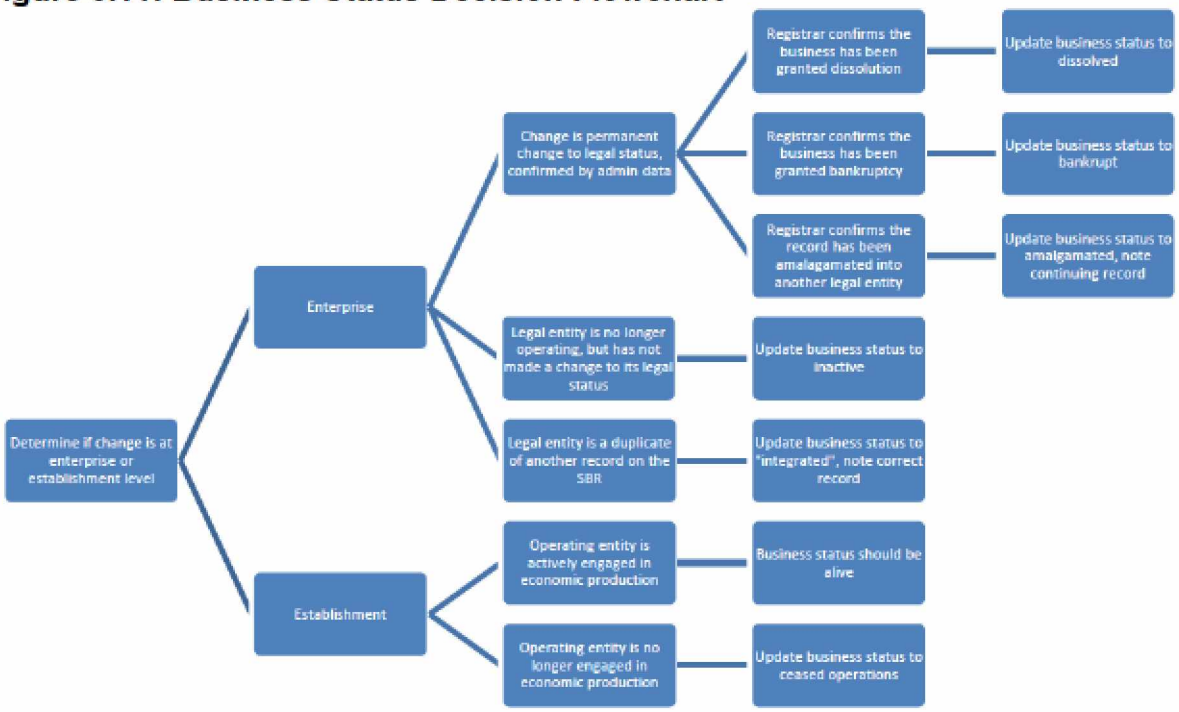
- Select an “inactive”, “dissolved” , “bankrupt”, “integrated” or “Amalgamated” legal entity unit
- Select “update” from the top menu
- Select “reactivation”

5.4.3 Inactivate Legal Entity

Inactive business status is used when a legal entity has no active operating entities, but has not undergone a legal process to unregister as a legal entity such as bankruptcy, dissolution or amalgamation. This status is normally assigned automatically to single-establishment enterprises that have outdated size variables. Multi-establishment enterprises cannot be updated to inactive status until their establishments have been ceased following the procedures outlined in the “cease operating entities” section.

5.4.4 Business Status Update Summary:

Figure 5.11: Business Status Decision Flowchart



5.5 Reorganization

5.5.1 Reorganize Legal Entity

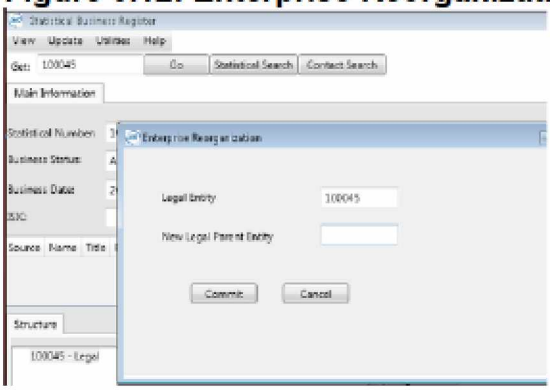
The reorganize legal entity function allows updaters to manage the position of legal entities within enterprise groups. An updater can change the legal parent of an enterprise using the update menu, as shown in Figure 5.12.

Steps for updating

- On an enterprise record, select “update” from the menu

- Select “update relationship links”
- Select “legal entity”
- Enter statistical identifier of new legal parent entity

Figure 5.12: Enterprise Reorganization



5.5.2 Reorganize Operating Entity

The reorganize operating entity function allows updaters to manage the position of establishment units within structures. An updater can update the relationship links of an establishment unit under the update menu, as shown in Figure 5.13.

Steps for updating

- On an establishment record, select “update” from the menu
- Select “update relationship links”
- Select “operating entity”
- Enter statistical identifier of new legal parent entity

Figure 5.13: Establishment Reorganization

5.6 Duplication

Because the SBR is built from administrative files that come from various administrative sources, the SBR administrator(s) will need to do **record linkage** before loading information to the SBR. This means they will undertake different processes to ensure that one unit with all the appropriate administrative identifiers is loaded to the BR as opposed to multiple units that represent the same entity. It is possible that some level of duplication will still occur. This is why all units from secondary data sources whose administrative records are not linked through record linkage to the primary data source should be created with a **prospect flag set to true (=1).**

Duplication can also occur if two units are created manually with the same information. It is important to confirm if a unit already exists before creating anything on the SBR.

Checking for duplication

Duplication research can be conducted at the enterprise or establishment level.

In cases with establishment duplication (establishment linked to more than one legal unit), confirm which enterprise unit the establishment truly belongs to and cease the duplicate establishment entity, no further reconciliation is required

Enterprise level duplication is more complex and takes place when you have two different enterprises on the database that seem to actually represent a single legal entity. In order to establish whether something has a duplicate, the best way is to use the statistical or contact search. Take key identifying information from the unit you want to search for duplicates of such as the names, addresses or phone numbers. Remember that information can be entered slightly differently which might be the reason the duplicate was created in the first place. For that reason, it is better to take the unique portion of a name and search for it rather than the entire name. For example, if the business name is J.D. Smith Incorporated, there might be various versions of the name recorded on various business datasets like JD Smith Inc. The best way to ensure you are not missing any matches is to create a unique search string from the name and use the "contains" function in the statistical search or contact search. In this example, a search contains "Smith" would bring back both results. It is possible that the search string is too generic and yields too many results. If that is the case, consider adding another filter on top of the search string such as a phone number area code or city.

Once a match is identified, double check that they are truly units representing the same registered legal entity. For the most part, the two units should have similar names, addresses, activities and business types. One of the units could likely have a prospect flag (True = 1) resulting from an administrative data load.

If the two activities do not match, consider the possibility that it is the same registered legal entity but should have more than one establishment. In this case, the complex structure should be created to represent the two establishments under one of the enterprises.

If there is a possibility that the two units are truly two different registered legal entities, leave them as-is and make sure that the prospect flag is set to false (=0) on both units. If the two units are truly duplicates of each other, proceed with the steps for reconciling the duplication.

Steps for resolving duplicates:

- Compare two enterprise units for matching information:
  - Variables should mostly match: should have same name, same address, same contacts
  - Confirm number of establishments
  - Choose one unit to continue, and one unit to cease
- Move all relevant, recent information to continuing unit
  - Do not forget to add any missing administrative IDs from the duplicate to the selected unit, and remove them from the duplicate entity
- Update business status of duplicate entity to "integrated" if enterprise and "ceased operations" if establishment
- Add a comment to the ceased unit to indicate the statistical number of the correct unit
- Remove prospect flag(s) (False = 0) from each unit to indicate that any potential duplication has been resolved

Steps for converting prospect flag to false (=0):

- Update **only available** at the enterprise level
- Select "update" from the top menu
- Select "update" again, then "legal entity"
- Clear the check box from the "prospect" field
- Hit "commit"
- ** Again, this update should only be done once an updater has validated there are no duplicates on the SBR, and preferably added any applicable administrative IDs that were missing from the record

**Please note that only legal entities with no children have the prospect flag updatable

5.6.1 Reviewing Prospects

In order to prevent duplication situations on the database, records with true **prospect flags (=1)** should be reviewed frequently. SBR administrators should regularly produce a list of prospect flag units for SBR updaters to go through and research for potential duplication.

The first step in reviewing a prospect should be to consider the original source (primary source) of the SBR and whether conceptually, the record should already be represented. For example, if NIS data is used to initialize the SBR as a primary data source, and VAT is reporting a prospect record (True = 1), one should consider whether such a record should be submitting NIS payments. If the prospect record should be submitting NIS payments, then it needs to be thoroughly researched (perhaps even contacted) prior to converting the prospect flag to FALSE (=0). Research should be done before converting the prospect flag to FALSE. A steady increase of duplicate records on the SBR can create a false upward trend in the economy or falsely mask a downward trend.

5.7 Manual Update Schedule

It is important to perform manual updates to the SBR on a regular basis to ensure it is as accurate and up-to-date as possible. In particular, it is important to perform supporting updates after an administrative data load is performed. SBR Updaters should be aware of the administrative update frequency followed by the system administrators.

After each administrative data upload, a spot check should be performed of differences between Survey Universe Files (SUFs) taken before and after the load, particularly large differences in size variables or unexpected business status changes. Other projects may be required based on the quality of data received. Updaters should inform SBR administrators of any concerns with data quality so that the administrative variable priority matrix can be updated accordingly.

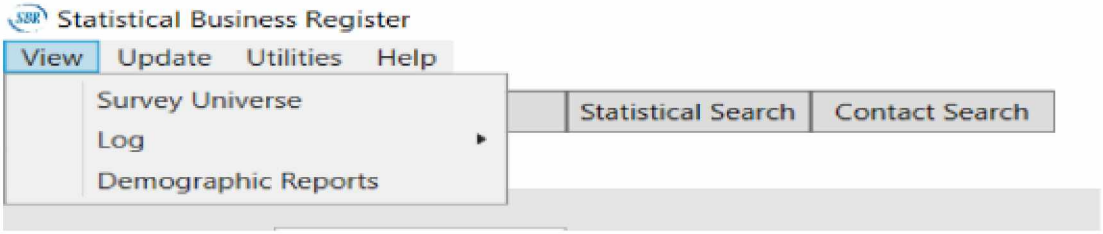
Updaters and SBR administrators should coordinate in developing a schedule for admin data and manual updates to be performed in harmony with each other.

6. Using and Analyzing Extraction Files

Extraction files (SUF) are snapshots of the SBR at a given time. The snapshot field includes all SBR variables of interest for all units. The SBR administrator creates these files at regular intervals and stores them according to their own specific procedures. When demographic statistics are required or the SBR is to be used as a survey frame, a user can access the most relevant snapshot and analyze it accordingly.

These files can be used to produce survey frames or to generate demographic statistics. Generating demographic statistics from these files can also be done in various ways depending on the preference of the analyst. The Demographic reports viewer allows users to easily produce demographic statistics about the business population directly through the GUI interface. The default SQL table can also be accessed and analyzed in various statistical programs like Excel, SQL, SAS, Power BI, and others.

Figure 6.1: View



6.1 SUF Viewer

The latest SUF information can be viewed directly through the SBR interface. In order to access the SUF viewer, select view from the top toolbar, then Survey Universe. A pop-up will appear allowing the user to select the most recent SUF extraction, and view the data therein at the record level.

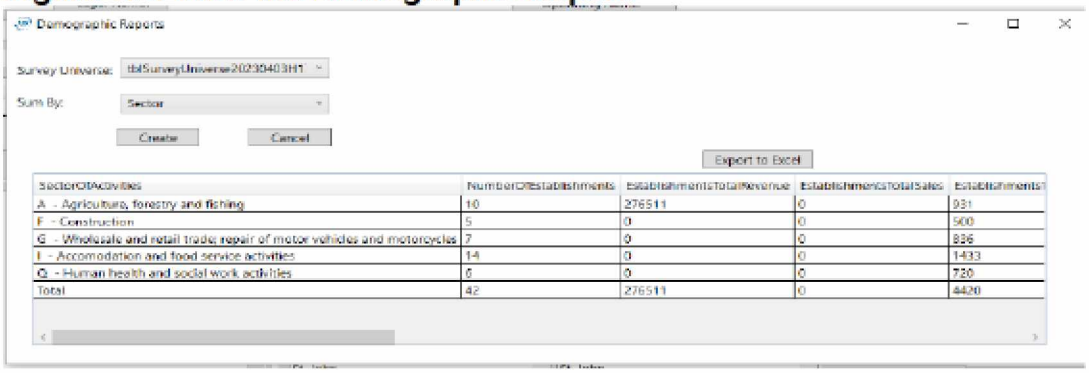
Figure 6.2: View the SUF

6.2 Demographic Reports

Beyond simply accessing the information, a user can also use the demographic reports tool to produce quick statistics on the SBR meeting certain criteria. The demographic reports can be accessed by selecting view on the top toolbar, then selecting demographic reports on the drop down menu. The user may select a SUF extraction to analyze, and whether they would like to produce a report by Parish or by Sector. This function returns a sum by Parish or Sector for the following variables:

- Establishments
- Establishment total revenues
- Establishment total sales
- Establishment total number of employees
- Establishment total number of male employees
- Establishment total number of female employees
- Enterprises
- Enterprise total revenue
- Enterprise total sales
- Enterprise total number of employees
- Enterprise total number of male employees
- Enterprise total number of female employees

Figure 6.3: View the Demographic Report



SectorOfActivities	NumberOfEstablishments	EstablishmentsTotalRevenue	EstablishmentsTotalSales	EstablishmentsT
A - Agriculture, forestry and fishing	10	276511	0	931
F - Construction	5	0	0	500
G - Wholesale and retail trade; repair of motor vehicles and motorcycles	7	0	0	826
I - Accomodation and food service activities	14	0	0	1433
Q - Human health and social work activities	6	0	0	730
Total	42	276511	0	4420

The demographic report can be directly exported to Excel by selecting the option Export to Excel.

6.3 Assessing Extractions Through Excel

For more complex analysis, a user may want to access the extractions in the analysis software of their choice. In order to access the SBR files in excel, they will need to be imported from the SQL server on which the SBR files are stored. Please contact the database or SBR administrator for the link. Figures 6.1 through 6.4 illustrate the process of opening an SQL Dataset in Excel. In order to import the SQL Data, you need to know the database name and enter your credentials (using windows authentication should be sufficient). After that, you

must select PRASCGenericBR to access extraction files. This will give you a list of tables from which you can select the desired extraction for analysis.

Figure 6.4: Open SQL Data in Excel

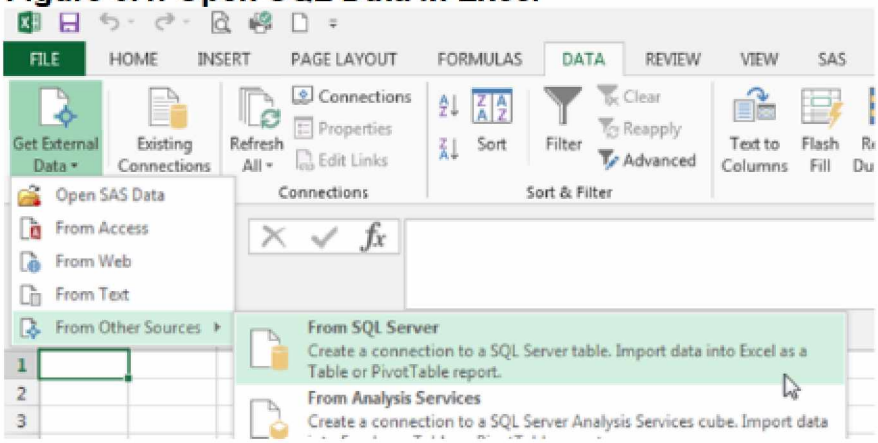


Figure 6.5: Enter Server Name and Authentication

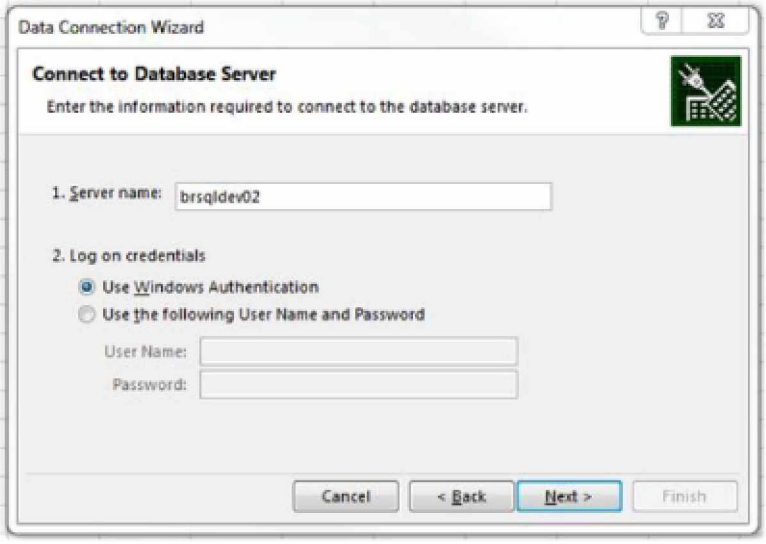


Figure 6.6: Select PRASCGenericBR to access extraction files

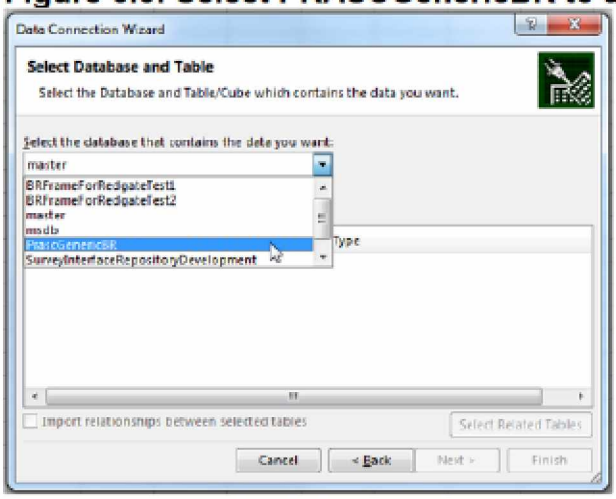
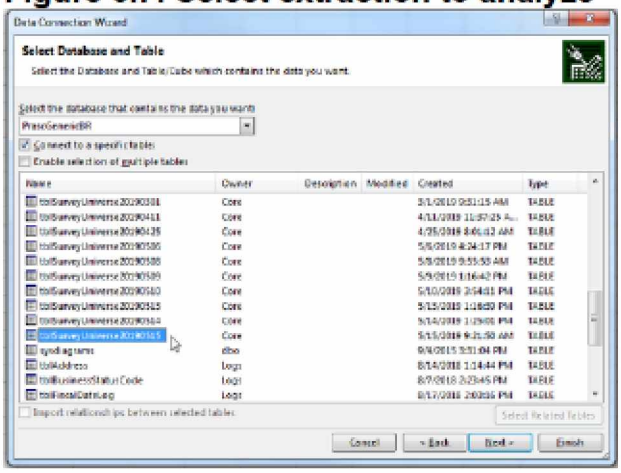


Figure 6.7: Select extraction to analyze



6.3.1 Permanent Random Numbers

The permanent random number (PRN) is a random number that is assigned by the SBR system for sampling purposes. The PRN ranges from 0 to 1 which allows it to be used in two different ways to select a simple random sample, both of which permit ease of long term sample maintenance.

The first way to use the PRN to select a sample is to determine the sampling rate, then select records that fall within a given interval. For example, if a 20% sample is to be selected, then all records having a PRN between 0.15 and 0.35 could be selected. It is noted that this will not produce the exact sample size that was desired, but it should be close.

The second way that the PRN can be used is to sort the sampling population by the PRN, select a starting point, then select the number of units within the sample. This will produce the exact sample size that is desired, but requires that the sampling population be sorted prior to selection.

To use the PRN under stratification, the selection process is performed within each stratum. Details about the use of the PRN can be found within the PRASC paper entitled “Using Permanent Random Numbers to Select and Maintain a Sample” (see Annex B), which covers topics such as how to handle new units entering the population (births), units exiting the population (deaths), units that change strata after being sampled (stratum jumpers), sample rotation, coordinating samples from multiple surveys and estimation.

7. SBR Administrator Roles

Users with the **SBR administrator** role are responsible for managing the overall function of the SBR. They can assign user permissions, load and prioritize administrative data and create extraction files.

Once the SBR setup has been completed, there are a few maintenance items that SBR administrators are responsible for:

- Updating user permissions
- Updating variable priority using Excel Template
- Pre-processing of administrative data files
- Loading of administrative and survey data using the Excel Template
- Customization of codesets (ISIC)
- Generating Survey Universe File (SUF)
- Creating and following a maintenance schedule

**Note: SBR administrator and database administrator roles are two different functions. The SBR administrator is to manage SBR users, loading files and extracting files from the SBR. The database administrator is a technical role to install and manage the database by using MS SQL Server Management Studio.

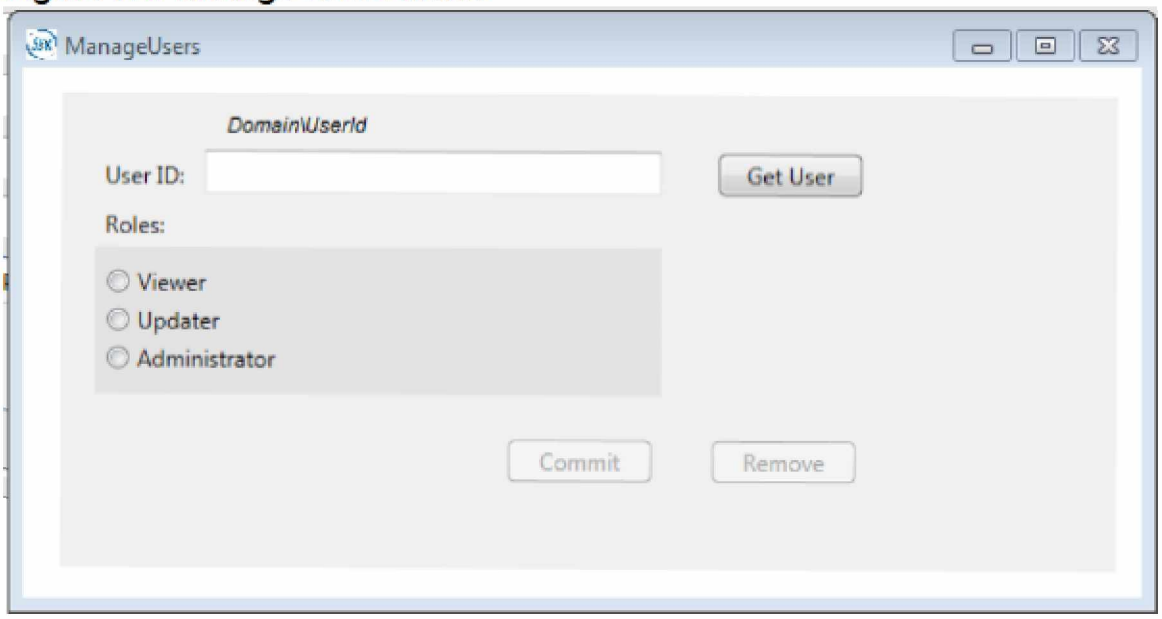
The SBR system has 3 user roles:

- SBR **administrator**: able to launch the application, browse operating structures, perform updates, execute utility jobs such as Manage User, SUF Generation, Batch Inactivation etc.
- SBR **updater**: able to launch the application, browse operating structures and perform updates. The Utilities menu is not visible.
- SBR **viewer**: able to launch the application and browse operating structures only. The Update menu and Utilities menu are not visible.

7.1 Updating User Permissions (Roles)

Under the “utilities” menu there is an item called “manage users”. Only those with SBR administrator permissions can access this menu item. The “utilities” menu is only viewable by someone with that permission. This module allows an SBR administrator to add, remove, and update user” role on the database. Selecting “Manage User” from the utilities menu will open the form “manage users”.

Figure 7.1: Manage Users Menu



The user ID text box is set to a maximum length of 30 characters.

7.1.1 Entering a Valid User (User Found on the Database)

The SBR administrator will enter a Domain\UserId and select “get user”. If the user is found in the database, the system will automatically populate the form and select their authorized role.

The SBR administrator will update information as needed. Note that viewer, updater and administrator will be radio buttons, and only one should be selected. Then SBR administrator will select “commit” to perform the action.

The system will do the following, depending on what has been changed:

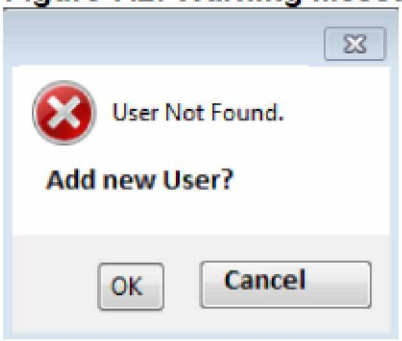
- If a role has been changed, the system will remove the existing role from the user and add the new role to the database.
- If a role has not been changed, the system will only show up a message to say the role is the same as before.

** Note, the SBR administrator cannot remove themselves from the SBR administrator role or update themselves to updater or viewer role. Only another SBR administrator can do that.

7.1.2 User Not Found on Database

A dialog box will open, indicating that the user was not found on the database, and whether the SBR administrator wants to add the user.

Figure 7.2: Warning Message – User not Found on Database



If SBR administrator selects "ok", the system will return to "manage users" form, leaving the user name in place and the viewer role will be checked as default. The SBR administrator will update information/change the role required if necessary and the SBR administrator will select "commit".
The system will add the user to the database along with their new role. If the steps above were successful, the system will display a dialog box indicating that the user was inserted. If any step fails, the system will display a dialog box indicating that there was an error. In a case of an error, the database administrator will probably have to become involved to adjust the security on the database.

If the SBR administrator selects "cancel" then the dialog box will close, control returns to the form, but the username will be blanked out, and no role checked.

7.1.3 Removing a User

SBR administrator has loaded an existing user into the form, as per step above and the SBR administrator clicks the "remove" button.

** Note that the SBR administrator cannot "remove" their own record.

The system removes the role from the user and removes the user from the database. The system refreshes the "manage users" to blank out userid and role checked. If not successful, the system would display a message indicating that there were problems removing the user from the database. The database administrator will probably have to become involved to adjust the security on the base.

7.1.4 Security within the SBR Browser

With the users and roles on the database, the SBR system can now action the security model. Upon launching the SBR application, the system will query the database for the roles for the user. If the user doesn't have permissions to access the database, the system will show a dialog indicating that they don't have permission to launch the application.

If the user has a role in the database, then the system would use that information in order to determine which menus should appear, and what functions the user is allowed to use:

- SBR viewer role: user will only be allowed to browse data and use the "view" menu. They will be able to use the two search features, access the SUF, Demographic Reports, logs, etc, but no updates will be allowed. With this role, the update menu and the utilities menu will not be visible.
- SBR updater role: user will have access to the update menu and all functions under that menu, as well as all SBR viewer permissions.
- SBR administrator role: utilities menu will be visible and available, and they will be able to execute any of the items under the utilities menu.

7.2 Updating the Variable Priority Matrix

The following section outlines the recommended procedures for using the Variable Priority Batch Update tool to set variable priorities for administrative data being loaded to the Statistical Business Register (SBR). The Generic SBR was developed with the intention that most updates to the database will be made by administrative data from various sources. In order to accommodate multiple sources with similar data, the Variable Priority Matrix allows database administrators to specify which variables should be updated by which sources and in which order.

7.2.1 Creating a Variable Priority Matrix

The matrix should be set up with all updatable variables along the left hand side and all administrative data sources across the top, as is illustrated in Figure 7.1. Each administrative data source is assigned a number for each variable corresponding to the source's priority. Administrative data sources that are blank for a given variable will never update that field. Variables with priority of 1 will overwrite all others for a given reference year. A priority of 1 should only be assigned to the most reliable source for a given variable. Assigning a priority of 2 means it is the second most reliable source, and so on. Manual SBR entries for a given reference year will never be overwritten with administrative data from the same reference year.

****It is noted that information from a newer reference year will overwrite higher priority information (administrative or manual) from previous years.**

Figure 7.1: An example of the Variable Priority Matrix

VariableName	TAXPriority	SocialSecurityPriority	VATPriority	RegPriority	OtherPriority	SurveyCollected	OriginalBR
StatisticalID	1	1	1	1	1	1	
SOCIAL SECURITY_ID		1					
BusinessType	2	3	4	1	5		
LegalCity	2	3	4			1	

For example, if a country is receiving Telephone Number information from the VAT and Social Security programs, the statistical office must decide which source provides a higher quality telephone number and assign priorities accordingly. If VAT is assigned a priority of 1 and Social Security a priority of 2 and both are received in 2019, the information from VAT will take precedence and be displayed on the BR. However, if National Security received in 2020 and VAT is not received, the information from National Security 2020 will overwrite the VAT 2019 values and the National Security 2020 will be displayed on the SBR.

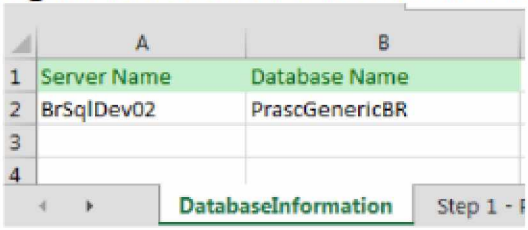
For this reason, it is very important to assign reference periods on input files accurately. Another important consideration is quality; for example, if the quality of National Security is not reliable, it might be better to assign a blank priority to it rather than a priority of 2 to prevent low quality information of a newer reference year overwriting information from previous reference years but from a more reliable source.

7.2.2 Variable Priority Excel Macro File

The following describes the recommended procedure; however, users may follow their own procedure, as desired.

- In the Excel macro file (Prasc_Variable_Priority_Batch_Update_Tool.xlsm), verify that the Server name and Database Name on the "DatabaseInformation" tab are correct. This information is used to populate the different dialog boxes. Adjust if necessary. Figure 7.2 is a screenshot of the Database Information Tab.

Figure 7.2: Database Information Tab

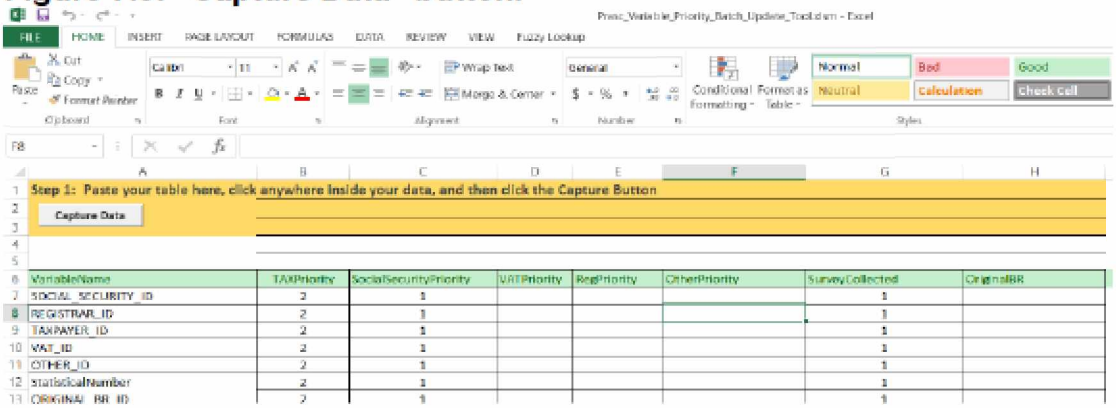


- Select the "Step 1 – Paste your Data" tab and populate your variable priority matrix. Ensure the format matches the formatting on the Data

Dictionary tab exactly. Validate the priorities for each data source and variable.

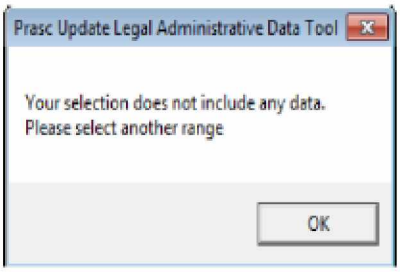
- Click on any data cell containing data, then click on “Capture Data”. Figure 7.3 is a screenshot of the “Step 1 – Paste your Data” tab and the Capture Data button.

Figure 7.3: “Capture Data” button.



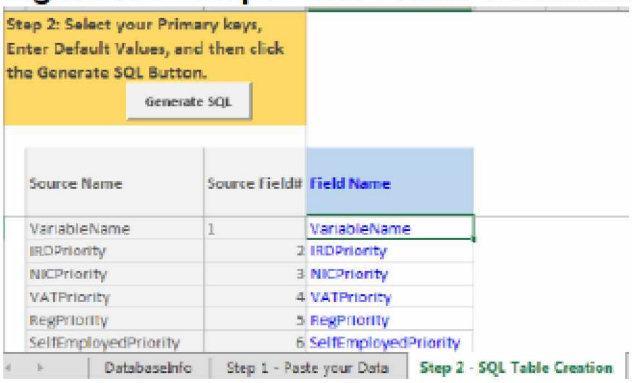
****Please Note:** If you don’t click on a cell with data before clicking on the “Capture Data” button, the dialog box in Figure 7.4 will appear. To fix the problem, simply close the dialog box, and repeat the process by clicking on a cell containing data and clicking on the “Capture Data” box.

Figure 7.4: Error caused when data is not selected



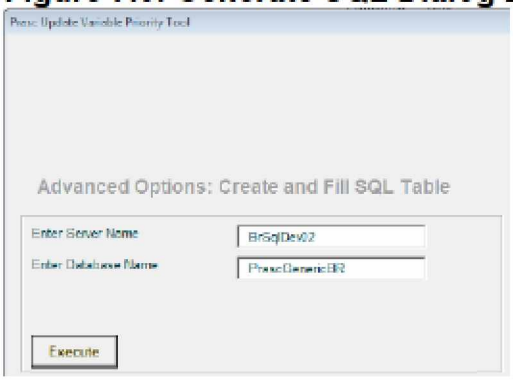
- The macro will automatically switch to the next sheet “Step 2 – SQL Table Creation”, as is seen in Figure 7.5.

Figure 7.5: "Step 2 - SQL Table Creation tab"



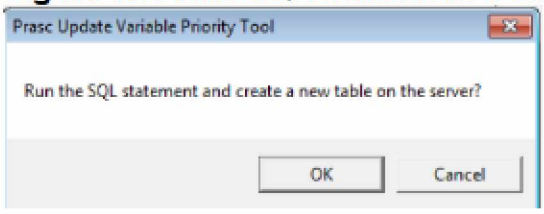
- Click the button "Generate SQL". The system will open a dialog box as seen in Figure 7.6.

Figure 7.6: Generate SQL Dialog Box



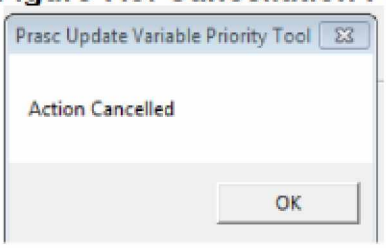
- Verify that the Server name and Database name are correct, and adjust if necessary. Note that this information comes from the Database Information tab. If the information is incorrect in the dialogue box, this indicates that it is also incorrect on the Database Information tab and should therefore be corrected there before proceeding.
- Click the button "Execute". The system will confirm that the user wants to proceed using a dialogue box seen in Figure 7.7.

Figure 7.7: Run SQL Statement Confirmation



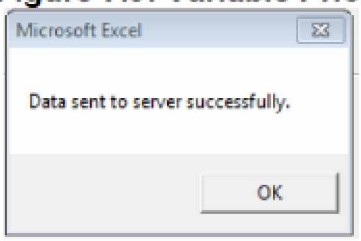
- If the user selects "Cancel", the system will display a pop-up notifying the user that the action was cancelled. Figure 7.8 is a screenshot of this pop-up. The system will then return to the previous dialog box.

Figure 7.8: Cancellation Pop-up



- If the user selects "OK" from the pop-up in Figure 7.7: Run SQL Statement Confirmation, the system will create the staging table on the database and load the variable priorities data. A pop-up will appear as illustrated in Figure 7.9.

Figure 7.9: Variable Priorities Loaded Successfully



- The system will also update the codeset table. Upon successful completion, the system will display a message to confirm this operation took place that states, "data insert into table successfully."

7.3 Preprocessing and Loading Administrative Data

The following section will review the tools and steps involved in pre-processing and loading administrative data files to the Generic SBR using the administrative data excel macro file. This step should only begin once other installation steps have been completed. The following procedures are recommended, although users may follow their own procedures, as desired.

7.3.1 Preprocessing

When receiving a file from administrative partners, there are several steps to accomplish before loading the file into the SBR. It is recommended to perform some quality assurance (QA) verifications. For example, check that the number of records received is as expected, a unique business identifier is present for all records, key requested variables (legal name, operating name, address, ISIC, size variable etc.) are populated and confirm the data vintage is up-to-date (recent reference period).

After QA is completed on the files, there might be a requirement to link information together from different administrative data sources (record linkage process). This process is accomplished using statistical software that can perform record linkage (SAS, Fuzzy Lookup Table in Excel etc.). Generally, the record linkage process requires some manual resolution to confirm matches, to

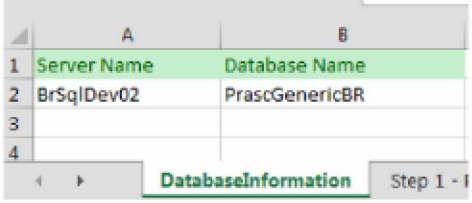
solve one-to-many matches and to properly review unmatched records. During the manual resolution of the unmatched records, the prospect flag variable has to be set as per the instructions in section 3.4.3 of this document. Finally, once the above steps are completed, the files will need to be prepared in the format that it will make-it loadable into the Excel ***Prasc.Admin_Data_Batch_Update_Tool.xlsm*** Excel Template.

Warning: When linking two units from two different administrative data sources and loading the two data sources on the SBR, it is very difficult to undo a linked of the unique identification numbers after the load of the two data sources. For more information on that subject, please see the FAQ below on: **Is it possible to modify two unique business identifiers linked together?**

7.3.2 Using the Administrative Data Excel Macro File

In the Excel macro file (***Prasc_Admin_Data_Batch_Update_Tool.xlsm***), verify that the Server name and Database Name on the DatabaseInformation tab is correct. This information is used to populate the different dialog boxes. Figure 7.10 shows the information displayed on this tab of the excel macro file.

Figure 7.10: Database Information Tab

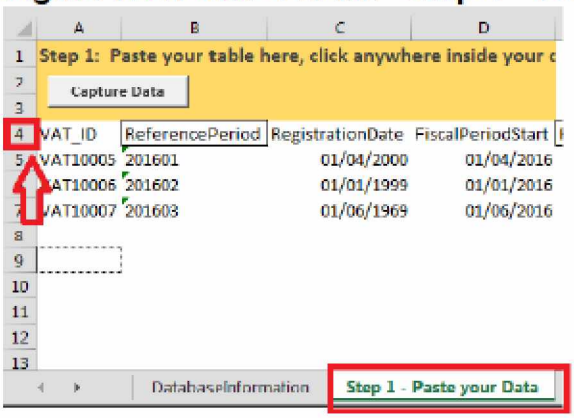


- In the incoming excel data file, ensure each column is named according to the standard names provided in the DataDictionary tab of the excel macro file (Prasc_Admin_Data_Batch_Update_Tool.xlsm). Be aware that the job is case sensitive, so "Isic" will not be updated, whereas "ISIC" will be. It is best to copy and paste the appropriate column names from the DataDictionary tab into the incoming excel data file, rather than type them in, to ensure that spelling and case are correct.

1.3.2.1 "Step 1 – Paste your Data" tab

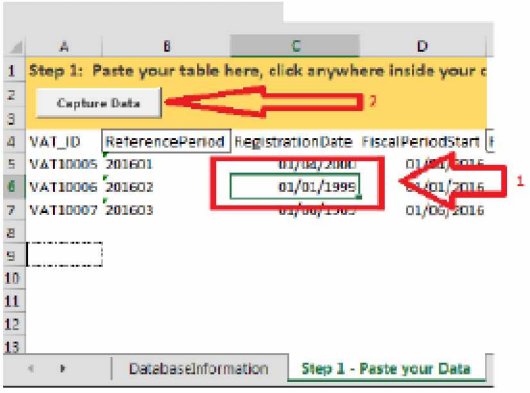
- Copy all data from the incoming excel file (including column names) and paste into the Excel macro file (Prasc_Admin_Data_Batch_Update_Tool.xlsm), in the tab "Step 1 – Paste your Data" starting at line 4. This is demonstrated in Figure 7.11.

Figure 7.11: Line 4 of the "Step 1 – Paste your Data" tab



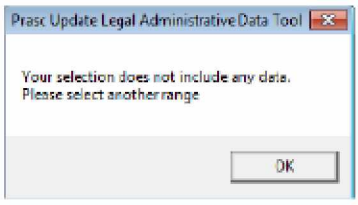
- Point the mouse and click on any cell containing pasted data (See Figure 7.12, arrow 1) and click the "Capture Data" button (see Figure 7.12, arrow 2).

Figure 7.12: Important areas of the "Step 1: Paste your Data" tab.



****Please Note:** If you don't click on a cell with data before clicking on the "Capture Data" button, the dialog box in Figure 7.13 will appear. To fix the problem, simply close the dialog box, and repeat the process by clicking on a cell containing data and clicking on the "Capture Data" box.

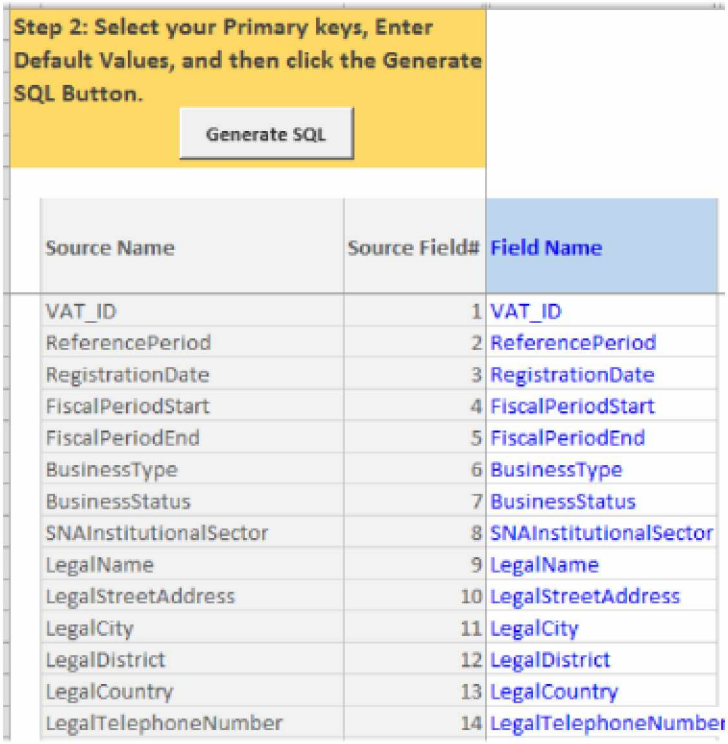
Figure 7.13: Error caused when data is not selected



1.3.2.2 "Step 2 – SQL Table Creation" tab

The Excel macro will automatically switch to the next sheet "Step 2 – SQL Table Creation". This will display a list of fields from which data will be loaded onto the database. An example of these fields is shown in Figure 7.14.

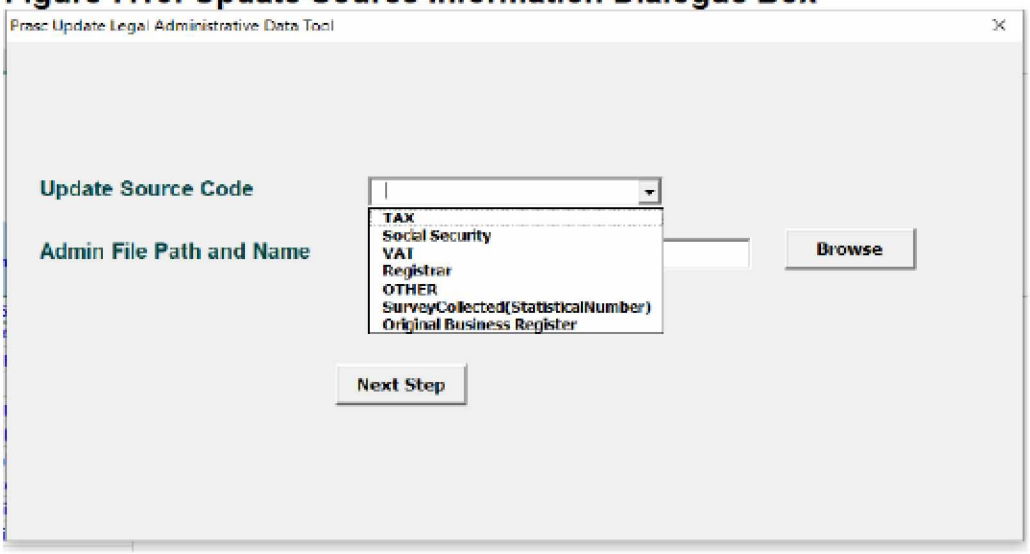
Figure 7.14: Fields from incoming file to be loaded on the database



Source Name	Source Field#	Field Name
VAT_ID	1	VAT_ID
ReferencePeriod	2	ReferencePeriod
RegistrationDate	3	RegistrationDate
FiscalPeriodStart	4	FiscalPeriodStart
FiscalPeriodEnd	5	FiscalPeriodEnd
BusinessType	6	BusinessType
BusinessStatus	7	BusinessStatus
SNAInstitutionalSector	8	SNAInstitutionalSector
LegalName	9	LegalName
LegalStreetAddress	10	LegalStreetAddress
LegalCity	11	LegalCity
LegalDistrict	12	LegalDistrict
LegalCountry	13	LegalCountry
LegalTelephoneNumber	14	LegalTelephoneNumber

- Click on the button "Generate SQL" at the top of the sheet.
- A dialog box will open, shown in Figure 7.15.
    - In the dialog box, select a value from the "Update Source Code" drop down list to identify the source of the Administrative data being loaded. There are 7 options available: TAX, SOCIAL SECURITY, VAT, Registrar, Other, Survey Collected and Original Business Register.
    - Next, click on the "Browse" button to select the file path name to store the information generated. This is strictly for documentation purposes. If an invalid path name is entered, it will have no impact on the process, but the batch job documentation will be incorrect.
    - Finally, click "Next Step".

Figure 7.15: Update Source Information Dialogue Box



**Please Note:

- If no "Update Source Code" is selected and the user clicks on "Next Step", a message will appear as shown in Figure 7.16. Simply select "ok" to close the dialog, then select an "Update Source Code".
- If no Admin File Path and Name is entered and the user clicks on "Next Step", a message will appear, as shown in Figure 7.17. Simply select "OK" to close the dialog, then enter a file path and name.
- Each incoming administrative data source has mandatory fields identified – such as the Unique ID and ReferencePeriod. If these fields are missing in the incoming data, the job cannot proceed. If these fields are missing, or if the Update Source Code is selected which is invalid for the incoming file (and therefore the mandatory fields are missing), a dialog box will appear as shown in Figure 7.18.

Figure 7.16: No Update Source code

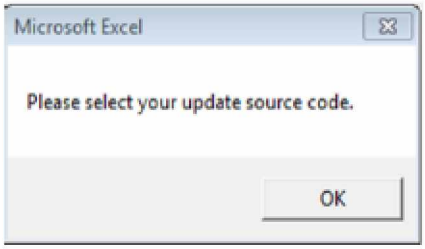


Figure 7.17: No Admin File Path and Name Error

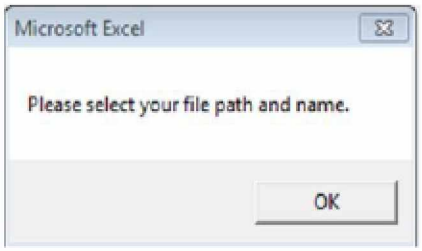
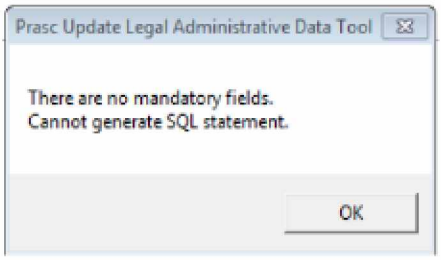


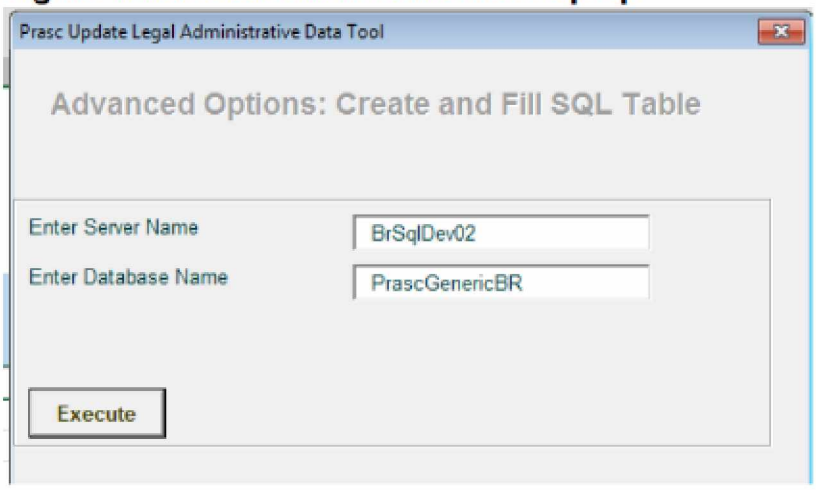
Figure 7.18: Mandatory Fields Missing Error



1.3.2.3 Final Steps

- Next, the system will pop up a window displaying the database information. This window is shown in figure 7.19. If the server name or database name is incorrect, it can be edited here. This information is retrieved from the "DatabaseInformation" tab in the Excel macro file.

Figure 7.19: Database Information Pop-up



- Click the "Execute" button on the Database Information pop-up.
- The system will prompt the user to verify that they want to proceed with a dialog box, shown in figure 7.20.
- If the user clicks "Cancel", then the system will display a dialog box as shown in Figure 7.21 and the information will not be uploaded.

- If the user clicks “OK”, the macro will create a table on the database and upload the data to the SBR by executing the stored procedure to update the database. Upon successful completion, the system will display a dialog box as shown in Figure 7.22.

Figure 7.20: Confirmation Dialog Box

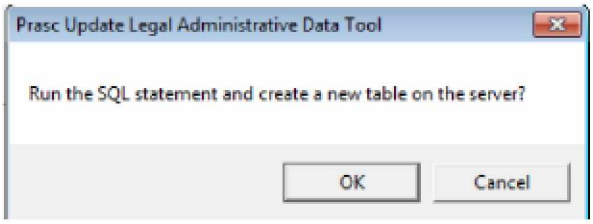


Figure 7.21: Cancellation Dialog Box

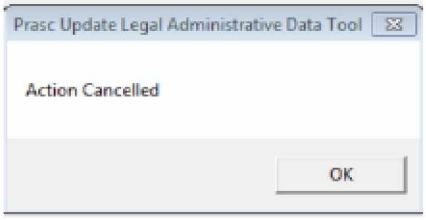
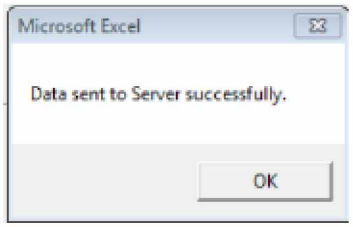
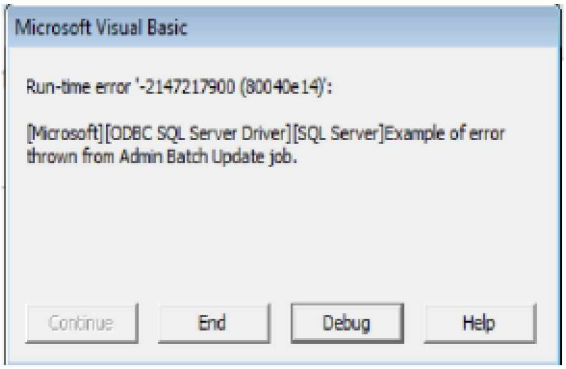


Figure 7.22: Successful Completion Dialog Box



****Please Note:** If there was an error, the system will display an error code returned from SQL.

Figure 7.23: Database Creation Error



7.4 Customization of Codesets (ISIC)

In certain situations, there may be a need to customize the ISIC codesets within the system. The system comes by default with the international ISIC rev. 4 loaded on the database. If a different version of the ISIC or country-specific codeset is needed, a collaborative effort between database administrators (IT Representative) and subject matter (SBR administrator) to update the ISIC codeset would be required in these cases. See the annex ?? for modifying ISIC code Table.

7.4.1 Necessary Preconditions

Changing version of ISIC require specific steps in order to obtain accurate results.

1. Generate a SUF from the Utilities Menu of the complete database and extract the file into Excel as specified in section 6.3 of this document. This will permit to extract the most recent version of ISIC code and description in Excel.
2. SBR Administrator will need to supply the complete new list of ISIC code and description to the Database Administrator (IT Representative) in order to update the codeset table.
3. The IT representative will update the codeset table with the new version of ISIC code and description.
4. Convert the ISIC code to the new version of ISIC. Sometimes it is simply a conversion of the code. In other occasions it could be a split of an ISIC code to two or three different ISIC. In this situation you would need to review the ISIC description provided by the respondent in order to determine the new ISIC code or contacting the respondent to get more details in order to determine the new ISIC code. You would need to repopulate your Excel template with the new ISIC codes and descriptions.
5. When you have converted all your ISIC code, you can simply reload your full or impacted business population by using the Excel Template with the new ISIC code with the description by using the Statistical Number as your main mandatory Unique keyId with the corresponding reference period.

**Please note while doing this exercise, it might not be possible to access your record on the database if your old ISIC code of your record is not part of the new ISIC list. Refer to the Annex ?? for more details.

7.4.2 Updating ISIC Codeset When the Database Does Not Contain Data

If the default ISIC codesets is not used by the country and there are no data loaded yet into the database, then the following steps will be required:

When the database is first being prepared, a list of ISIC codes will need to be provided to the database administrator with a list of ISIC codes, descriptions, and sector that ISIC represents. The database administrator (IT Representative) will then replace the Codes.tblISIC table in the database with the appropriate values. When the codeset table is loaded with the ISIC version, then the administrative data could be loaded into the database with the valid ISIC codes.

7.4.3 Updating ISIC Codeset When the Database Already Contains Data

In this case the SBR subject matter person (SBR administrator) will still need to supply a list of ISIC codes, descriptions, and sectors to the database administrator. However a careful mapping/analysis of whether the same ISIC codes are re-used between the old ISIC codes and new ISIC codes will need to take place. The database administrator will be able to add new codes, but for any ISIC code that needs to be removed, all businesses in that ISIC will need to be coded to the new ISIC code before the old code can be removed. This will take coordination between subject matter (SBR administrator) updating the SBR data and the database administrator adjusting the Codes.tblISIC table values.

Before starting the ISIC upgrade on the SBR, we recommend that the SBR administrator generate a SUF to be manipulated in Excel. That will make the job easier to convert your old ISIC code to the new version of ISIC. When conversion is completed for all businesses that required a new ISIC code, then you can simply reload your businesses impacted with the new ISIC code by using the excel template and the statistical identifier as your unique KeyId.

Careful attention should also be paid to ISIC codes that have changed meaning (description or sector) to confirm that the businesses that are coded to the old ISIC code are still accurately reflected with the new ISIC code and reference period. Refer to Annex ?? for more details.

7.5 Creating Extraction Files

Creating extraction files can only been done by users with the administrator role. There are several important preprocessing steps before the survey universe file (SUF) should be created to ensure that the SUF is as complete and accurate as possible. The creation of SUF should be done at regular, predetermined intervals so that updaters can ensure there are no in-progress updates when the SUF is created.

The SBR administrator can execute the following steps under the utilities menu (Figure 7.24) after loading administrative data from the Batch Load Excel Template:

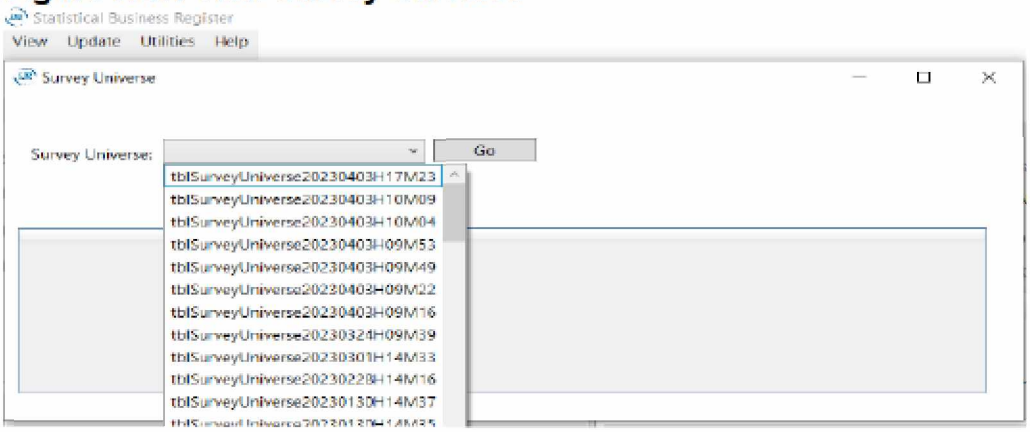
- Batch Inactivation
- Batch Activation
- Batch Reactivation
- Update Statistical Flags

Please note, when generating the SUF, all of the above functions will be executed automatically from the SUF Generation function.

Figure 7.24: Utilities Menu

After these steps have each been completed, the administrator can proceed with selecting "SUF Generation" under the Utilities menu. This will create a file accessible under the "view" menu. The file name will follow the naming convention tblSurveyUniverseYYYYMMDDHHMM (Figure 7.25) and the output will contain the full list of variables from the data dictionary. This file can also be accessed as an SQL table by accessing the server the database is stored on. See the section 2.4 Using and Analyzing Extraction files for more information about interpreting the SUF.

Figure 7.25: View Survey Universe



7.6 Scheduling maintenance and production

7.6.1 Developing a production schedule

It is suggested that the SBR administrator follow a specific schedule for SUF extraction to ensure that SBR data is consistent to facilitate appropriate comparisons (monthly, quarterly, annually, etc.).

8. Acronyms

- CARICOM – Caribbean Community
- ISIC – International Standard Industry Classification
- NSO – National Statistical Organization
- PRASC – Project for the Regional Advancement of Statistics in the Caribbean
- PRASC SBR - The Project for the Regional Advancement of Statistics in the Caribbean's Statistical Business Register
- SBR – Statistical Business Register
- SNA – System of National Accounts
- SUF – Survey Universe File
- UNECE – United Nations' Economic Commission for Europe
- UNSD – United Nations Statistical Division

9. Definitions

Term	Definition
Administrative Data	Files of data collected by other government bodies for the purpose of administering their own respective government programs.
Administrative Identifiers (Administrative Keys)	Identifiers assigned to units in administrative data files by the responsible organization.
Administrative Partners	Other government agencies that maintain administrative records, and have agreed to share this data with the NSO
Administrative Records	Administrative records are records found in administrative datasets.
Administrative Registers	Administrative registers are databases maintained by administrative partners for the purpose of administering their own respective programs.
SBR Administrators	SBR Administrator is the highest permission level on the SBR. Administrators can access all SBR content, make any type of update and make administrative changes to the database. Administrators are tasked with managing the initialization and maintenance of the SBR system.
Alive	Business status given to entities actively engaged in economic production.
Amalgamated	Business status of a unit that has undergone an amalgamation.
Amalgamation	Amalgamation is a legal process where two or more companies are joined to form a single new company. This process is usually overseen by the administrative partner responsible for the registration of businesses.
Bankrupt	Business status given to entities that have undergone the legal process for becoming bankrupt.
Business Date	The date on which the business status of an entity is presumed to have started.
Business Information	Section of the SBR interface that shows the key information about the selected unit. It displays the enterprise or establishment level information depending on the type of unit selected.
Business Status	The current state of a legal or operating business entity.
Cessation	Final state of a legal or operating entity when it has closed or given up its legal status as a business.

Term	Definition
Child (Or Subsidiary)	The enterprise that owns the other enterprise is referred to as the parent, and the enterprise that is owned is referred to as the subsidiary or child.
Ceased Operations	Business status assigned to establishment level units once they are no longer operating.
Coherency	Term that refers to the consistency in figures and variables between establishments and their corresponding enterprise.
Complex Unit Structure	Group of units that have recorded relationships between them
Contact Search	Function of the SBR that is used to search the repository of contacts on SBR units.
Continuing Legal Entity	In the case of amalgamation, one legal entity is selected to continue while one or more legal entities are ceased. The continuing legal entity is the entity selected for continuation
Demographic Reports	Reports generated through the SBR that provide statistical numbers on business population by Parish or Sector.
Dissolved	Dissolved is a form of cessation where the appropriate governing entity has a formal closure of a business with the state
Duplication	When two units on the SBR represent the same actual entity.
Eligibility Criteria	Characteristics of a unit that are required to be eligible for surveys (business status of alive, ISIC code and size variables)
Eligible Population	Population of units that have statistical indicators and are eligible for surveys.
Enterprise	Unit on the SBR that represents a legal entity and is eligible for enterprise level surveys
Enterprise Group	Group of enterprise level units (legal entities) that are linked together by legal ownership
Enterprise Level Information	Information about the legal entity or information about the business as a whole (sum of the enterprise and its establishments).
Enterprise-Only	A unit that only represents the enterprise and has no establishment-level information
Establishment	A unit on the SBR that represents an operating entity that has met the criteria for eligibility for surveys.

Term	Definition
Establishment Level Information	Information about a specific operating entity such as the operating name, operating address and ISIC.
Establishment-Only	A unit on the SBR that represents an establishment and has no enterprise-level information
Extraction	An extraction is the term used when the data from the SBR is downloaded into a file by the SBR administrator. The data file represents a snapshot of the SBR at a given time.
Government Units	These are units that exist on the SBR that represent a government entity.
Households	These are units on the SBR that represent a household.
Inactive	Business Status on the SBR that indicates a legal entity no longer has any active operating entities, but is still a registered legal entity.
Integrated	Business status assigned to an enterprise that has been part of an integration into another enterprise unit. This is classified as a type of cessation.
Integration	Business Status on the SBR that indicates a legal entity unit has been identified as a duplicate of another unit. The continuing unit should be identified in the comment section on the integrated unit.
ISIC	Classification system used to identify the main business activity of a unit.
Legal Entity	Unit on the SBR that represents a legal entity that has been granted its status as a business by the appropriate governing body.
Logs	Function of the SBR that displays the previous changes to a variable made on the SBR.
Match Types	This refers to the type of match function used in statistical or contact search, this defines how the system will use the search string in looking for a match. The different match types are: starts with, contains, ends with and exact.
Main Information	Section of the SBR GUI that shows information about the enterprise level unit selected or the enterprise level unit that is the parent of the selected establishment level unit.
Menu	Section of the SBR interface that allows a user to execute different SBR functions.

Term	Definition
Multi-Establishment Enterprise	An enterprise that operates multiple establishments concurrently.
New Business	Business status on the SBR that indicates a legal entity unit has been created on the SBR, but it has not yet reported size variables.
National Insurance	National Insurance refers to government agencies who provide National Insurance services to citizens.
Non-Profit Institutions	Type of unit on the SBR that represents a non-profit institution.
Operating entity	Unit on the SBR that represents the operating or establishment level of information.
One-Establishment Single-establishment Enterprise	Unit on the SBR that represents both a legal and operating entity simultaneously and is eligible for both enterprise and establishment level surveys.
Parent	A parent is a unit that is one hierarchical level higher than another. A parent can either be a unit that owns more than 50% of the common shares of another within an enterprise group, or the enterprise level unit in a complex operating structure.
Permanent Random Number	A random number that is assigned by the SBR system to each unit for sampling purposes.
Prospect Flag	Indicates a unit that the updater believes duplicates an existing unit on the SBR, though the record match has not yet been identified and further research is pending.
Reactivation	Process used to convert the business status of an enterprise unit from inactive or bankrupt or dissolved or integrated or amalgamated to alive.
Record Linkage	Process used in administrative data processing to link same units in the admin data file to the SBR units using a variety of techniques.
Relationship Variables	Refers to the relationships a unit has with other SBR units (Parent-Child, Legal-Operating, etc.).
SBR Linkage	Refers to variables on the SBR that refer to units on other databases. These variables are used to build and maintain concordance between different databases.
Search Bar	Bar at the top of the SBR interface that allows a user to type in a statistical number and retrieve a unit.

Term	Definition
Simple Unit Structure	One legal entity with a single operating entity represented by one unit on the SBR.
Single-establishment Enterprise	Unit on the SBR that represents both a legal and operating entity simultaneously and is eligible for both enterprise and establishment level surveys.
Size Variables	Variables on the SBR that indicate the size of the operating in terms of revenues, sales, number of employees, etc.
Stand Alone Entity	A unit on the SBR that has no recorded relationships with other legal or operating entity units.
Statistical Business Registers (SBR)	An SBR is a database that combines different administrative sources to produce a complete and up-to-date list of businesses and business operations within a given country. The SBR is also a frame used for sampling.
Statistical Indicators	Indicators used to denote the level of surveys a particular unit is eligible to receive (enterprise or establishment level).
Statistical Number	Number assigned randomly to each SBR unit for the purpose of identification, confidentiality, processing and analysis.
Statistical Search	Function of the SBR that allows a user to conduct a search of SBR units using specific parameters.
Stratification	Process where units are categorized based on different variables that permit for categorization and segmentation of the population (geographic location, size variables or ISIC codes).
Structure Browser	Section of the SBR interface that displays the relationships between units on the SBR in a hierarchical format.
Subsidiary or Child	A subsidiary is a unit that is one hierarchical level lower than another. A subsidiary or child can either be a unit that is owned by more than 50% of common shares by another legal entity. The term child can also refer to an operating entity unit that is below its legal entity "parent".
SUF Viewer	Tool available through the SBR interface that permits users to browse the most recent Survey Universe File (SUF).
Survey Eligibility	Survey eligibility refers to whether a unit has met the requirements to receive surveys. In order to be eligible to receive surveys a unit must be alive,

Term	Definition
	have at least one size variable, have an ISIC code and no prospect flag (=0). Eligible units have an enterprise, and/or an establishment flag in the business information section of the SBR interface.
Survey Universe Tool	The survey universe tool is a function of the SBR that permits a user to browse previous snapshots of SBR data created by the SBR administrator.
Toolbar	The toolbar is the bar at the top of the SBR interface that allows a user to execute various SBR functions.
Top Enterprise	The top enterprise is the enterprise that is hierarchically at the very top of an enterprise group. It is a parent of one or more subsidiaries, but it does not have a parent.
Updater	Updater is a permission level that is assigned to SBR users that are authorized to make updates to SBR data using the SBR interface. Administrators assign permission levels based on users' roles and training.
User Id	A user ID is the identity assigned to different SBR users. It is visible in the logs of updates so that changes can be traced back to the person responsible when needed.
Value Added Tax (VAT)	VAT refers to government programs structured around collecting Value-Added Tax from businesses added to the purchase of goods and services within a given jurisdiction.
Viewer	Viewer is a permission level on the SBR given to users who are authorized to look at SBR data using the SBR interface but do not have permission to make changes to the data or the system. Administrators assign permission levels to users based on their training and job requirements.

Appendix A: Business Rules

This annex contains summary information on the business rules for the Statistical Business Register (SBR). The business rules provide restrictions that exist in the business environment that must be represented in the system as system rules in order to preserve the data integrity.

1. Variable Priority Excel Template

Since the SBR will be used by different countries, the system needs to give the option to select the priority of the variable per Admin file. The template for the countries to set their priority of variables is located in the VariablePriority worksheet template.

The variable priority will be used to determine the source for *update* data where the reference periods are the same. There is a possibility that the reference periods for all the source are the same. The variable priority will be taken from the VariablePriority worksheet template.

As a general rule, the SBR will use the most recent data regardless of the source.

When *updating* data from secondary priorities, the system verifies that the reference period is more recent than what is on the base (all data must have a valid reference period). The SBR will only be updated from the secondary sources when the data received is more recent. If the reference periods are the same, then the variable priority will be followed.
If a priority is left blank for a given variable and admin source, the system is to disregard the update of that variable from that source. This is to give the possibility for a country to filter updates from certain sources on certain variables.
Data for updating the SBR can be supplied by in-house surveys conducted by the Statistical Office, these updates will be administered regardless of the reference period (it is the SBR Administrator that must determine the appropriateness of the data).

2. Batch Load Excel Template

This Excel template has been designed to automatically load the input file that will be used to initialize and maintain the SBR. It contains all the necessary variables required by the SBR. The SBR administrator needs to populate the data fields into the PasteYourData worksheet before loading the data.

The mandatory key columns need to be in the incoming file for each records otherwise the file will be rejected with an error.

- Each record must have data in at least one KeyID (e.g. SOCIAL_SECURITY_ID or REGISTRAR_ID or TAXPAYER_ID etc.)
- Each record must have a valid Reference Period populated in the ReferencePeriod variable.
- Each variable must have a standardized format as described in the data dictionary worksheet (e.g. Convert text file into excel format)
- Each record must have a maximum of 2 KeyIDs in the incoming file. The 1st would be from the source file, and the 2nd ID would be from the linked source. If two KeyIDs are provided, then the system will try to find a record with the 2 keys. (For example, the TAX file might contain the TaxPayer_ID and the SOCIAL SECURITY_ID as the linked KeyID.)
    - If it finds a StatisticalNumber for both keys then it will confirm that they are the same and it will proceed with the update.
    - If the system finds different StatisticalNumber for the two keys, the record will be rejected. The system will:
        - Create a record into BatchProcessingError sheet in Batch Load Excel Template to be manually researched and created
        - Write all the data fields that were used in the incoming file into the BatchProcessingError sheet.
        - Generate the following Error message in the BatchProcessingError sheet: '***Conflicting identifiers in the external keys table***'
    - If the system finds the record with only one 1 key, then the 2nd key will be inserted into the External Keys table, so it is linked to the same enterprise as the first key and it will proceed with the update.

Please remember that the External Keys Table contains the following variables:

- StatisticalNumber
- SOCIAL SECURITY_ID
- Registrar_ID
- TaxPayer_ID
- VAT_ID
- TEPA_ID

****Note:** All records from Administrative Files must have a unique key business identifier field to enable proper mapping and integrating of data. Expected Administrative files are generally Corporate and Individual Tax filers (TAX), Value Added Tax (VAT), Social Security (NIS) and Central Registrar (Registrar). Survey data must always have a StatisticalID in the input source file as a business unique identifier.

****Note:** If multiple data sources are used as an input to the SBR, it is assumed that these files have been pre-processed and linked together using external processes prior to loading them on the SBR. Finally, all updates will be applied regardless of the current Business Status of the Entity on the SBR.

3. Creating and Maintaining SBR Entities on the SBR

The SBR administrator is responsible to load the appropriate input data files using the Batch Load Excel Template in order to initialize and maintain the SBR. The Excel Template will upload the Excel input data file to the appropriate SQL tables. Excel macros exist for the Administrative files below:

- Inland Revenue Department (TAX) (Corporate and individual Tax Filers)
- Value Added Tax (VAT)
- National Insurance Security (Social Security)
- Registrar (Central Business Registrar)
- Previous Business Register File
- Other Batch Processes (Surveys and potentially other Administrative files)

3.1 Initialization of the SBR

Once the Input Data File is created, data needs to be loaded into the SBR. Each country will first identify a primary file (usually an Admin Data File) to initialize the SBR and this file will be complemented by other administrative files or survey files.

- For each record in the Input Data File:
  - The SBR system automatically assigns a unique ID – auto-generated from the system
    - Format – Integer starting at 1000000
  - Insert record in Business Core Table of the SBR
  - Insert Unique IDs and StatisticalNumber into SBR Concordance Table

***Note:** Business Status, Statistical Indicators (Enterprise and Establishment flags) and other key variables will be determined as per rules in subsequent sub-sections of section 10.3 which described the business rules for Automatic Update.

3.2 Creation of a Record after initialization

For variables that are null in the Variable Priority, when an entity is being created, they will be updated.

- If the record does not exists in the external keys table
  - Assign a unique ID (StatisticalNumber is auto-generated and the format of this ID is an integer starting at 1000000)
  - Insert record in SBR
    - If record has at least one size value then
      - Business Status is set to 2 – Alive
      - Business Status Date is set to processing date
      - Prospect Flag is set according to business rule in section 10.3.3
      - Update entity with all the variables that are in the input file.

- If record has no size values then
  - Business Status is set to 1 – New
  - Business Status Date is set to processing date
  - Prospect Flag is set according to business rule in section 10.3.3
  - Update entity with all the variables that are in the input file.

3.3 Setting of the Prospect Flag (ProspectiveLegalFlag Variable)

The Prospect Flag is updated under the following conditions **from an input data file.**

- If Legal Entity has Operating Entity children and Prospect flag = TRUE (=1) in the input data file then
  - Reject the record
  - Output the following message ***Record can't have a TRUE Prospect Flag because it has Operating children*** to the BatchProcessingError Sheet into the Batch Load Excel Template
- If Legal entity doesn't have an Operating Entity children and Prospect Flag is False (=0) on the SBR and Prospect Flag = True (=1) in the input file then
  - Set the Prospect Flag = True (1) on the SBR
  - Output the following error message ***Record had its Prospect Flag updated to TRUE*** to the BatchProcessingError Sheet into the Batch Load Excel Template
- If Legal entity doesn't have an Operating Entity children and Prospect Flag = False (=0) in the input file then
  - Set the Prospect Flag = False (=0)
- If Legal entity doesn't have an Operating Entity children and Prospect Flag = Null then
  - Prospect flag is not updated

Note: For all records that had their Prospect Flags updated to TRUE on the SBR then the following message is outputted to the BatchProcessingError Sheet into the Batch Load Excel Template: ***Record had its Prospect Flag updated to TRUE***

3.4 Setting the Business Type

The Business Type is updated under the following conditions

- If Business Type in Input Data File is Null then set Business Type to 'Other'

- If Business Type in Input Data File is not Null and valid then set Business Type to what is in the Input Data File
- If Business Type in Input Data File is invalid then set Business Type to 'Other'

3.5 Business Status

If at least one of the keys exist in the external keys table (i.e. record already exist on SBR) then update the Business Status under the following conditions

- If Business Status is invalid (i.e. not on codeset table) then
  - Reject the Business Status update (no update to Business Status)
  - Write the following message to BatchProcessingError Sheet into the Batch Load Excel Template : ***Invalid Business Status Code***
- If Business Status is invalid (i.e. not on codeset table) and no Legal Entity exist on the SBR for this record (i.e. Inserting a new business on the SBR) then
  - Set Business Status = 1 (New) if no Size variables or no ISIC Code or Prospect Flag = True (=1) In Input Data File
  - Set Business Status = 2 (New) if Size variable(s) exist and ISIC Code exist and Prospect Flag = False (=0) in Input Data File
  - Write the following message to BatchProcessingError Sheet into the Batch Load Excel Template : ***Invalid Business Status Code***
- If Business Status is greater than 2 **AND** Variable Priority for that source is not NULL and reference date of previous update (on the log table) of the variable is less recent than the reference date (on incoming data file)then
  - Update Business Status to what is in the Input Data File
- If Business Status is greater than 2 **AND** Variable Priority for that source is not NULL and reference date of previous update (on the log table) of the variable is the same as the reference date (on incoming data file) **AND** the priority for that variable is greater than the priority of the update then
  - Update Business Status to what is in the Input Data File
- If Business Status is greater than 2 **AND** Variable Priority for that source is not NULL **AND** reference date of previous update (on the log table) of the variable is more recent than the reference date (on incoming file) then
  - Don't update Business Status

3.6 Size Variable Updates (revenue, sales, wages and number of employees)

If at least one of the keys exist in the external keys table (i.e. record already exist on SBR) then update the Size Variables under the following conditions

- If Variable Priority of the Size Variable for that source is NULL then

    - Don't update that variable size from the Input Data File
- If Vintage Date (on the database) of the Size Variable is less recent than the reference date (on incoming Data file) then
    - Update the size variable from the Input Data File
- If Vintage Date (on the database) of the Size Variable is the same as the reference date (on incoming data file), and the priority for that variable from the Input Data File is greater than the priority of the updated size variable on the SBR then
    - Update the variable size from the Input Data File
- if Vintage Date (on the database) of the Size Variable is more recent than the reference date (on incoming data file) then
    - Don't update the variable size from the Input Data File

3.7 ISIC Update

The system validates incoming ISIC codes with the ISIC Table to ensure validity. The ISIC is updated under the following conditions from an input data file.

- If ISIC Code is invalid then
    - No ISIC code is updated
    - No ISIC description is updated
    - No effective date of the ISIC code is updated
    - Write the following error message in BatchProcessingError Sheet into the Batch Load Excel Template: ***Invalid ISIC code***
- If ISIC Code is valid **AND** if reference date of previous update (on the log table) is less recent than Reference date of incoming data record then
    - Update the ISIC Code
- If ISIC Effective Date is present in Input Data File then
    - Update the effective date of the ISIC code (The effective date is only updated when the ISIC changes)
- If ISIC Effective Date is Null in input data record then
    - Update ISIC Effective Date according to the Reference Period data variable of the input data record (Note that reference period is a mandatory field).
- If ISIC is Null then
    - No ISIC code is updated
    - No ISIC description is updated
    - No effective date of the ISIC code is updated

3.8 Address information

A codeset table exist with the complete list of country codes. The SBR has the Country Code of the host country stored into a system table. This variable stored in the system table is referred to the Host Country Code. The rules below is used

for all address updates that contain the country code variable (Legal, Operating and Contact addresses)

- If Country Code in input data file is Null **AND** Variable Priority is not null **AND** reference date on log table is less recent than reference date of incoming data record then
    - Set the Country Code to the Host Country Code
    - Update the other variables of the address
- If Country Code in input data file is invalid **AND** Variable Priority is not null **AND** reference date on log table is less recent than reference date of incoming data record then
    - Set the Country Code to the Host Country Code
    - Update the other variables of the address
- If Country Code in input data file is valid **AND** Variable Priority is not null **AND** reference date on log table is less recent than reference date of incoming data record then
    - Set the Country Code to the Country Code from the input data file
    - Update the other variables of the address

**Note that for Contact address, the reference date validation is unnecessary as each admin data source has their own contact, so it is assumed that each file will be the most recent file.

3.9 Assigning the Country of Control

When an entity is created or updated on the SBR, the Country of Control is assigned as follows:

- If no Legal Entity on the SBR (new Legal Entity) **AND** a legal address exist **AND** Country of Control is null in the input data file then
    - Set Country of Control to the Legal Entity Country Code from the legal address
- If no Legal Entity on the SBR (new Legal Entity) **AND** no legal address in the input data file then
    - Set Country of Control variable to Null.
- If Country of Control is not null in the input data file **AND** variable priority is not NULL **AND** reference date on log table is less recent than reference date of incoming record then
    - Set Country of Control variable with the value in the input data file

3.10 Enterprise and Establishment Statistical Code derivation

Below are the business rules to initialize the statistical indicators (Enterprise and Establishment) from the batch load initialization or update or online manual update

- If Prospect Flag = TRUE (=1) **OR** Business Status not = Alive (2) **OR** ISIC is Null then
    - Set Enterprise Flag to False
    - Set Establishment Flag to False
- If Prospect Flag = False (=0) **AND** Legal Entity (It is a legal entity) has a children operating entity (Establishment) **AND** Business Status = Alive (2) AND ISIC code is not Null then
    - Set Enterprise Flag to True
    - Set Establishment Flag to False
- If Prospect Flag = False (=0) **AND** Legal Entity(It is a legal entity) has no Children operating entity **AND** Business Status = Alive (2) **AND** ISIC code is not null then
    - Set Enterprise Flag to True
    - Set Establishment Flag to True
- If Prospect Flag = False (=0) **AND** not a Legal Entity (It is an operating entity) **AND** Business Status = Alive (2) **AND** ISIC code is not null then
    - Set Enterprise Flag to False
    - Set Establishment Flag to True

****Note** that the system will prevent users from creating operating entities children under operating entities. Operating entities can only be birth/created under Legal entities.

3.11 Business Status Derivation

Below are the business rules to initialize and update the business status

3.11.1 Activation Process

- If Legal Entity exist on the SBR **AND** Business Status = New (1) **AND** has a at least one size variable (Revenue, Sales, Wages or Number of Employees) then
    - Set Business Status to Alive (2)
    - Set Business Status Date to System Date
    - If ISIC is not Null **AND** Prospect Flag is False (=0) then
        - Set Enterprise Flag to 1 (True)
            - If Legal Entity has operating children (Establishments) then
            - Set Establishment Flag to False (=0)
            - Legal Entity has no operating children then
            - Set Establishment Flag to True (=1)
    - If ISIC is Null or Prospect Flag = True (=1)
        - Set Enterprise flag to False (=0)

- Set Establishment Flag to False (=0)

3.11.2 Inactivation Process

- If Legal Entity **AND** (Business Status = 'New' *OR* = 'Active') **AND** Size variable vintage date greater than (>) 24 months old (or the date in inactivation table parameter) then
  - Set Business Status = inactive (3)
  - Set Business Status Date = current processing date
  - Set Enterprise Flag = False (=0)
  - Set Establishment Flag = False (=0)
- If Legal Entity **AND** (Business Status = 'New' **OR** = 'Active') **AND** Size variable vintage date greater than (>) 24 months old (or the date in inactivation table parameter) **AND** have operating children with no recent size variable then
  - For the Legal Entity:
    - Set Business Status = inactive (3)
    - Set Business Status Date = current processing date
    - Set Enterprise Flag = False (=0)
    - Set Establishment flag = False (=0)
  - For each operating entity children:
    - Delete operational link
    - Set Business Status = ceased (6) for this operating entity
    - Set Business Status Date = current processing date
    - Set Establishment Flag = False (=0)

3.11.3 Dissolved

- If Legal Entity **AND** Business Status = Inactive (3) **AND** Business Status Date > 12 months old (or the date in inactivation table parameter) (i.e. it has been Inactive for the past 12 months) then
  - Set Business Status = dissolved (5)
  - Set Business Status Date = current processing date
  - Set Enterprise Flag = False (=0)
  - Set Establishment Flag = False (=0)
  - Remove ownership links

****Note:** There should be no operational links left as they as they were ceased as part of Inactive step. The SBR doesn't allow Operating Entities to be birthed under a Legal Entity with a business status different than 'New' or 'Active'.

3.11.4 Reactivation

- If Business Status is not = Alive (2) **AND** input data source indicates activity (recent Size Variable) then

- Set Business Status = Alive (2)
- Set Business Status Date = current processing date
- If ISIC is not Null **AND** Prospect Flag is False (=0) and size variable exist then
    - Set Enterprise Flag = True (=1)
    - Set Establishment Flag = True (=1)

3.11.5 All non-size Variable Updates

If at least one of the keys exist in the external keys table (i.e. record already exist on SBR) then update the remaining Variables under the following conditions

- If the Variable Priority for that source is NULL then
    - Don't update the variables from the Input Data File
- If reference date of previous update (on the log table) of the variable is less recent than the reference date (on incoming data file) then
    - Update variables from the Input Data File
- If reference date of previous update (on the log table) of the variable is the same as the reference date (on incoming data file) **AND** the priority for that variable input data source is greater than the priority of the update then
    - Update the variables from the Input Data File
- If reference date of previous update (on the log table) of the variable is more recent than the reference date (on incoming data file) then
    - Don't update the variables from the Input Data File

****Note:** Only variables that are in the incoming data file will be updated. For example, if there is no Operating Name in the incoming file, but an Operating Name on the database exist then the system will not update the Operating Name to null on the SBR. Same for all data variables including Size Variables.

Appendix B: Permanent Random Numbers

1. Introduction

It is important to keep a sampling plan as simple as possible, and yet allow it to serve its purpose in a robust fashion. For most business survey sampling needs within the Caribbean, the method outlined in this document should meet these requirements as it is easy to visualize, understand and apply. It also allows for some easy techniques for sample rotation, sample coordination between surveys and the management of births and deaths. Additionally, since it is a form of simple random sampling, the weighting theory is straightforward.

It needs to be noted that while stratification is not explicitly mentioned within most of the following sections, whenever sample selection, weighting or estimation are discussed, it is implied that the techniques are applied within each stratum. The exception would be the assignment of the Permanent Random Numbers (PRNs) which would be completely random across the entire common frame such as a Business Register (BR), and would not take into account any pre-existing information.

When referring to a survey frame, it is assumed that it has been selected from a common frame such as a BR. Typically, a BR contains all possible business units within a country. A survey frame is a subset of the BR which has been selected using criteria to identify the units of interest.

This document starts by outlining the basic idea of the method being suggested, then offers details about the various nuances that come into play over the lifespan of a survey. This is followed by sections describing how to handle new units being introduced to the frame (i.e., births), deaths and stratum jumpers. It is very important to handle units entering and exiting the population properly in order to ensure that survey estimates do not become biased over time, thus possibly introducing false trends in the estimates or masking true trends.

The next sections discuss ideas for sample rotation and coordinating samples between surveys which are followed by the weighting scheme for a simple random sample along with the non-response adjustment that may be required. Given the importance of weighting, an Excel file is attached which follows a stratum through a couple of survey cycles in order to provide a practical reference for determining survey weights. The section on stratum jumpers discusses domain estimation and provides a practical example; as this is an important aspect of calculating survey estimates, it is also summarized in a section of its own. The final two sections present considerations for redesigning the sample as well as some closing remarks and references.

2. Basic Concepts of Using Permanent Random Numbers to Select a Sample

The key to this method is that it selects a simple random sample through the use of a PRN which takes on a value between 0 and 1. As the name suggests, each frame unit is assigned a single PRN that is retained through time; it is important that these PRNs not change. For example, in the case of a business survey, if a unit goes out of business, then becomes active a few years later, it should be reactivated with its original PRN.

Additionally, to allow for sample coordination between surveys, the PRN needs to be carried in a common frame such as a BR so that each unit will carry the exact same PRN on all survey frames extracted from the common frame.

The PRNs are assigned within the interval of 0 and 1 for a few important reasons. Firstly, the interval is easy to remember and therefore helps to ensure that it remains constant through time. Any variation from the original PRN interval when introducing births will create a concentration (or lack) of births in certain parts of the PRN interval. This uneven distribution of births could lead to unpredictable sample sizes and will likely produce biased survey estimates.

A second reason for choosing the interval between 0 and 1 is to make it easy to identify; it is unlikely that another data item on any of the frames will take on such values.

However, the main reason for choosing an interval between 0 and 1 for the PRNs is that it allows sampling rates to be easily translated into sampling intervals. For example, if a 20% sample is desired within a stratum, the units with PRNs within the interval [0.3, 0.5)[1] could be selected into the sample; other intervals with different start and end points are possible, provided that their length is equal to the length of the sampling fraction. It is noted that while this sampling technique is the easiest to apply, it seldom selects the exact sample sizes that were set out at the beginning.

Another way to select a sample (within each stratum) is to sort the survey population by PRN, choose a starting point, then sequentially select the desired number of units. Unlike the previous method, this will select the exact sample size that was determined by the survey design. It is noted that there are studies indicating that this technique has a slight bias towards selecting birthed units. However, when compared to the non-sampling errors that can be introduced into a survey, this is likely of little concern.

---

[1] A sampling interval of [0.3, 0.5) is interpreted as: 0.3 <= PRN < 0.5. In other words, the square bracket "[" implies that the end value is included in the selection interval while the round bracket "(" implies that the end value of the interval is not included.

For both techniques, the direction of the selection can proceed either left or right of the starting point and is often varied when attempting to coordinate samples. This topic is discussed further in section 7 (Sample Rotation).

3. Nuances of assigning the PRN

There are a number of approaches for selecting random numbers. They range from paper lookup tables to using facilities within software that generate pseudo-random numbers using a starting seed. The use of random number lookup tables can be quickly eliminated in today's world due to it being labour intensive and supplying only a limited quantity of random numbers. Turning to software options, almost all software has a pseudo-random number generator which has been designed to generate series of numbers that possess the qualities of a random number but avoid duplicates. However, many of these facilities have pitfalls that the user needs to research and take into consideration.

For example, when working in Excel, generating pseudo-random numbers can be done quite easily using the RAND function. However, caution needs to be taken because if the cell is left as the RAND function, the value will change every time any change is made to the spreadsheet. To overcome this, the user can point to the formula bar and press F9, for each observation. Another option is to copy the formulae, and use *Paste Special / Values*, to either overwrite the existing formulae or to create a new column. In both situations, the PRNs should now appear as numbers.

In SAS, and likely in many other software, a seed is required. The benefit of using a seed is that the generation of the random numbers can be controlled: using the same seed will lead to the generation of the same sequence of pseudo-random numbers allowing reproducibility. To avoid the duplication of pseudo-random numbers when assigning the PRN to births, the starting seed should always be the last random number that was assigned on the previous occasion. If this is not done, there is a possibility of creating clusters of units with the same PRN, which will result in a number of issues such as the sample losing its randomness thereby creating unequal selection probabilities and complicating weighting.

Diagnostics should be run whenever PRNs are generated (either to initialize the common frame or to introduce birthed units), to ensure that anomalies in the PRNs were not created. The first and most important check is to test whether any duplicate random numbers have been created. In general, a good pseudo-random number generator should not create duplicate values, and in fact they should be very few and far between with almost all software. If one is able to produce basic statistical analysis, key percentiles would provide a quick and easy assessment of the PRN assignment. For example, the 90^{th} percentile should be close to 0.90. (Further information about pseudo-random number generators can

be found on Wikipedia:
https://en.wikipedia.org/wiki/Random_number_generation).

If any of the diagnostics indicate anomalies in the PRNs, then the assignment procedures need to be reviewed and fixed as soon as possible. The longer anomalies go undetected, the more they can wreak havoc on the sample selection and the estimates being produced from the samples using these procedures.

4. Introducing New Frame Units (Births)

As alluded to in the previous section, introducing new units to the common frame is as easy as assigning them a PRN. If the sample was selected using the interval method, then a birth would be sampled if it falls within the current sampling interval. On the other hand, if the sample was selected by counting out the desired number of records from a starting point, then new sample sizes will need to be calculated for each stratum such that they respect the initial sampling parameters, and the desired number of units would then be selected. Under both methods, the sampling weights within each stratum need to be updated[2] to be in line with the new population (N'_h) and sample (n'_h) sizes: $w'_{hi} = N'_h/n'_h$.

Just a reminder that the precautions that were previously mentioned need to be followed to ensure that the assignment of duplicated PRNs is a very rare event. To give this context, tests that have been run on well-designed pseudo-random number generators have seen a million random numbers being generated with no duplicates.

One aspect that should be noted is that if the population changes over time, the current effective sample sizes should be reviewed to determine whether the survey(s) should undergo a redesign as described in section 11 (Redesigning the Survey). That is, if the population grows over time and the initial sampling rates are maintained, the sample size will also grow. As such, it may become feasible to produce estimates of similar quality but with a smaller sample. On the other hand, if the population has experienced a high level of attrition, the predetermined targets of quality may no longer be getting met with the design sampling rates.

In general, introducing new records to the population is fairly straightforward, but the evolution of the sample needs to be monitored to ensure that it continues to meet the original objectives.

---

[2] Pure theorists would calculate a more complicated weight for the births, however, the non-sampling errors that are experienced within a survey will create larger issues than this simplification of the weighting scheme.

5. Dealing with Out-of-Scope Units (Deaths)

As a bit of background, out-of-scope units can occur for several reasons. For business surveys, it would include cases where businesses are no longer operating. Additionally, it could also include units that have changed characteristics which remove them from the scope of the survey. For example, a farm that switches its primary source of revenue from farming to Agrotourism would change ISIC industry and would become out-of-scope for agriculture surveys but in-scope for tourism surveys. However, for a survey that covers both industries, it would be a stratum jumper which is discussed in section 6 (Stratum Jumpers).

As previously mentioned, removing out-of-scope units from the population needs to be done with caution. Out-of-scope units should be kept on the frame and flagged as such. If they are selected into the sample, they are not sent for data collection: their data are simply imputed with 0's for all the data fields that were to be collected. This means that out-of-scope units are still part of the estimation process: so that their weighted values will represent the non-sampled out-of-scope units on the frame. Note that since the out-of-scope units have a 100% response rate, they will need to be omitted from the nonresponse adjustment, as discussed in the weighting section 9 (Weighting).

An extreme case of this can be visualized; a survey frame is used for only one survey, and the sample is never rotated. After a few years, a number of out-of-scope records will get detected within the sample, however, there is no mechanism to detect them outside of the sample. During estimation, the survey responses get weighted to represent similar units on the rest of the frame. However, if the out-of-scope units are removed from the frame after each collection, then there are no out-of-scope units in the sample to represent the rest of the population, which will lead to increasingly upward bias of the estimates over time. Similar situations can also be seen even when the sample is rotated, and the frame is updated from multiple sources.

As the number of out-of-scope units accumulate within the sample, it will erode the efficiency of the survey. As such, the number of out-of-scope units should occasionally be reviewed, and if the number is quite high, then it may be desirable to redesign the sample. Information about redesigning the sample can be found in section 11 (Redesigning the Sample).

Caution:
The only time that an out-of-scope unit should be considered for removal from a survey population is when it comes from an independent source such as administrative data or the survey is being completely redesigned, including the assignment of PRNs. Otherwise, it is certain that the estimates will become biased (Further information can be found in Lavallée, 1996).

6. Stratum Jumpers

As mentioned in the previous section (Dealing with Out-of-Scope Units (Deaths)), a stratum jumper is a unit that has a change in its stratifying characteristics. A stratum jumper will remain in its original stratum throughout the life of the survey and therefore be treated as a member of that stratum when calculating weights. However, it will be tabulated based on its updated characteristics. For example:

First Survey Occasion

Design			Post-Collection		
Stratum ID	Criteria	Weight	Stratum ID	Classification	Weight
STRAT11	Size: Medium ISIC=AA	100 / 20	STRAT11	Size: Medium ISIC=BB	100 / 16
STRAT11	Size: Medium ISIC=AA	100 / 20	STRAT11	Size: Large ISIC=AA	100 / 16

Second Survey Occasion

Design			Post-Collection		
Stratum ID	Criteria	Weight	Stratum ID	Classification	Weight
STRAT11	Size: Medium ISIC=AA	100 / 20	STRAT11	Size: Medium ISIC=BB	100 / 15
STRAT11	Size: Medium ISIC=AA	100 / 20	STRAT11	Size: Large ISIC=AA	100 / 15

In the above example, the first unit was assigned to stratum STRAT11, however, during collection it was discovered that it is in a different ISIC (coding error or a change in activity). The new ISIC code was recorded on the BR and the survey frame, however, the stratum on the survey frame remains as STRAT11. When assigning the nonresponse adjusted weight (4 units did not respond), the unit continues to be treated as part of STRAT11. The change in classification comes into play during estimation where it will contribute to the estimate for its new ISIC (or domain). Following this unit to the second survey occasion, it is still part of STRAT11 for sample selection and weight assignment, but again it will contribute its values to its new domain during estimation. When determining the weights for such units, they need to be computed using their original stratification in order to properly represent other non-sampled units that started in ISIC AA but are now in ISIC BB.

Similarly for the second unit in the example. It needs to retain its original stratum (STRAT11) for assigning the weight to reflect the non-sampled units that were either erroneously coded to the stratum of medium units or units that have grown through time.

It is recognized that occasionally, these types of units do not represent changes in the population and are simply outliers. For example, a small unit that grows rapidly could move from a stratum of small sized units to one of medium or large units. If the rapidly growing unit had a very high weight and now it also has a

high value, it could create serious issues with the survey estimates. If it is believed that this is a unique unit, it may be desirable to make adjustments to have it represent only itself by assigning the unit a weight of 1 and moving it to a take-all stratum of similar types of units. When this is done, the nonresponse adjusted weights for the remaining units in the stratum need to be adjusted such that the entire population continues to be represented (if the quickly growing unit has a weight of 20 and now has a weight of 1, if the weights are not adjusted there will be 19 units not accounted for in the estimate). The nonresponse adjustment technique that is presented in section 9 (Weighting) will automatically make the correct adjustments to the weights by effectively reducing both the population count and the number of responding units within the stratum by 1.

In the following example, the first unit has become a medium-sized unit (perhaps close to the limit between medium and large). However, it is so large that applying the proper weight is causing it to sway the estimates to the point that they are not representing the population. Since it was considered to be unique, it was assigned a weight of 1 and moved to the special stratum. The weight adjustment for STRAT01 reduced the population and sample sizes to reflect that this unique unit was moved to the special stratum (in addition to reducing the sample size from 20 to 16 to reflect the non-responses).

Design			Post-Collection		
Stratum ID	Criteria	Weight	Stratum ID	Classification	Weight
STRAT01	Size: Small ISIC=AA	500 / 20	Special_TakeAll	Size: n/a ISIC=AA	1
STRAT01	Size: Small ISIC=AA	500 / 20	STRAT01	Size: Small ISIC=AA	(500-1) / (16-1)

When placing these outlying units into the special take-all stratum, a very non-homogeneous stratum is being created. Because of this, adjusting weights to account for nonresponses will likely produce a very inaccurate estimate; therefore, it is critical that survey responses are available for each unit within this stratum. For this reason, it becomes very important to ensure that the use of this type of stratum be kept to a minimum by including only the true outliers that are having an adverse impact on the estimates.

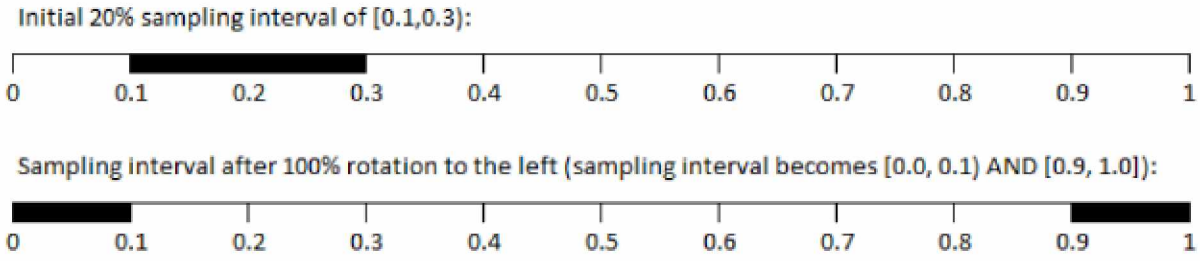
Note:
The examples in this section used non-descriptive stratum IDs in order help survey practitioners resist the urge to update them as units change strata. Similar to the out-of-scope units, updating the stratification for only the units within a sample will create a bias that grows over time because the sample will no longer represent the population units that have not been sampled.

7. Sample Rotation

Refreshing the entire sample or only a portion of the sample is simple: the sampling interval is shifted to the left or to the right. In the first section, there was an example of a 20% sampling rate using a sampling interval of [0.3, 0.5). A 100% sample rotation would result in the new sampling interval [0.5, 0.7) if the shift is to the right, or [0.1, 0.3) if the shift is to the left. Similarly, a 25% sample rotation would result in the new window becoming [0.35, 0.55) or [0.25, 0.45). It is noted that the sampling interval can be shifted either left or right, which may be desirable for sample coordination that is discussed in section 8 (Coordinated Sampling).

In the 100% rotation example, it is important to note that the sampling interval was just shifted to be adjacent to the previous one. This is for simplicity because non-adjacent sampling intervals would make it difficult to manage sample rotation over time. Looking at the example with 100% rotation, if the sampling interval went from [0.3, 0.5) to [0.6, 0.8), how would the records in [0.5, 0.6) get selected? If this was solved by moving the interval to [0.7, 0.9) then one needs to remember to go back to sample the ones within [0.5, 0.7) at a later date. Given that all units are in a random order, non-adjacent sampling intervals are unnecessary and will only make for a confusing sampling maintenance.

Taking the example a bit further, one might be wondering where the 100% rotation would go after moving to the lower PRNs of [0.1, 0.3), if the rotation took place by shifting the interval to the left. It would simply "wrap around" to start selecting the higher numbers such that the next sampling interval will be [0.0, 0.1) AND [0.9, 1.0]. Below is a diagram summarizing the intervals:



Someone thinking about setting up a sample rotation will need to decide upon the rotation rate. This decision would be based on considerations such as response burden / time in-sample, coverage of the entire population, estimate stability (continuity), loss of experienced respondents and increased costs that can be associated with new respondents. For example, a high rotation rate would reduce response burden and the entire population would be cycled through the sampling more quickly, thus speeding up frame updates. However, the stability of the estimates could be affected due to a completely different set of units frequently entering / exiting the sample. Additionally, since it often takes a few

survey occasions for a respondent to become used to responding to the survey and to learn how to properly answer the questions, the response rates and data quality could be negatively impacted. Therefore, to determine the rotation rate, one needs to assess items such as the goals of the survey, the quality of the frame, the difficulty of the questionnaire in terms of response burden and concepts, the quality of the responses and how the survey response rate could be impacted.

The frequency of rotation will be a trade-off of the same advantages and disadvantages that are considered when determining the amount of rotation. In general, a survey with more frequent collection such as monthly or quarterly could have a more frequent rotation than a survey with annual or biannual collection. However, it is suggested that samples in the Caribbean be rotated on an annual basis as this keeps the process fresh in everyone's minds and will allow all surveys to be on the same rotation schedule.

Note:
When considering the sampling rate and rotation rate of a survey, it may be necessary to make special considerations for strata with high sampling rates. Possible solutions include rotating these strata less frequently and assigning them a lower rotation rate. The key will be to determine the maximum length of time that a unit should be in the sample as well as the minimum time that it should be out of sample. It will also be key to keep the strategy simple and to document it to ensure that it is properly applied through time.

To ensure the ongoing cooperation of units with a high response burden, it may be desirable to send letters informing them of their importance to the survey and outlining the survey limits that have been developed (in-sample maximums and out-of-sample minimum). It may also be desirable to send letters when these types of units are rotated out of the sample, thanking them for their cooperation and indicating when they will be in sample again. Such communications help to build good will with the respondents while keeping them informed of the NSOs expectations.

8. Coordinated Sampling

Coordinated sampling would come into play when multiple surveys are selecting their frames (and implicitly their samples) from a common frame such as a BR. In such cases, it may be desirable to minimize the overlap of the samples and thus limit response burden of units selected in two or more survey frames at the same time. However, there could also be cases where it may be desirable to maximize sampling overlap in order to analyze relationships between data items collected by different surveys at the same time; although, in situations such as

this, one should first assess the possibility of merging the two surveys into one to reduce costs.

To coordinate samples from different surveys, a separate sampling interval would be required for each survey. To ensure that the goal of minimizing overlap is achieved through time, the performance of the proposed scheme should be assessed by simulating the sample selection and examining the overlap of the samples through the expected lifespan of the surveys. This is essential because each survey will have its own stratification, sampling rates and rotation rates which could eventually lead to a high overlap of the samples. Additionally, if the samples start to overlap, the situation could persist for a long time and lead to respondent fatigue, thereby compromising response rates and data quality. Note that if several surveys have high sampling rates or if there are many surveys using the same frame, it may be impossible to eliminate sample overlap. If this occurs, it could be desirable to minimize the number of times that survey units experience multiple survey contacts by simply not contacting in-sample units for some of the surveys or by collecting multiple questionnaires with one contact.

As mentioned in section 7 (Sample Rotation) the direction of the sample rotation can go to either the right or the left of the previous selection interval. For the example in the first graphic below, two well-spaced samples are selected and rotated to the right. In this scenario, when the samples eventually overlap, the situation will likely persist for a long time. However, as shown in the second graphic, if the starting points are the same (such as 0.5 or at observations N/2 and N/2+1), one sample could be selected to the left of 0.5 and the other to the right. In this scenario, the samples would overlap more often, but it would be for a much shorter time period. To visualize this, consider two cars on an oval racetrack. In order for a car to pass another going in the same direction, it will take more time than if they are travelling in opposing directions. However, the ones travelling in the opposing directions would pass one another much more often, but for a shorter time.

Another method for minimizing sample overlap across surveys that can be used when the survey population is very large, in comparison to the samples to be drawn, is to partition the frame. The usefulness of such a strategy in the Caribbean is not clear. For this reason, this concept has been summarized in Appendix A in the event that it might find a use.

It is noted that under the sampling method described in this document, sample coordination is only possible for surveys selecting the same level of units. That is,

when considering businesses with a complex operating structure, a survey that bases its sample selection on the *legal units* of a structure will not be able to be coordinated with a different survey that selects the *operating units* of the complex businesses. For the most part, complex businesses tend to be very large operations which would cause the legal units to fall into strata with fairly high sampling rates, where sample coordination would not be possible. Since being able to coordinate surveys that select units at different levels of an operating structure would be become quite complex and the advantages of it in smaller countries would not warrant this added complexity, it would be advisable to coordinate the sampling of only the simple units (i.e., those without legal and operating units / parent – child relationships).

9. Weighting

As previously mentioned, a sample based on the method described in this document can be treated as a simple random sample which lends itself to the simplest weighting procedure. However, one needs to account for out-of-scope units having a 100% response rate as they are not sent for collection and it is expected that all newly out-of-scope units would be detected. This means that only the responding units will be subjected to the nonresponse adjustment while all out-of-scope records will have their values imputed to zero and receive the sample design weight.

Hence, the weights for the out-of-scope records, within each stratum, would be the design weight[3]:

$$\frac{N_h}{n_h}$$

where:

N_h is the stratum population size, and
n_h is the stratum sample size (i.e., the number of units that were selected into the sample).

Additionally, all other in-sample records receive the weight of:

$$\frac{N_h}{n_h}\frac{n_{h\,(in-scope)}}{n_{hr(in-scope)}}$$

Where:

$\frac{N_h}{n_h}$ is the design weight for stratum h,

[3] As previously noted, when selecting a sample using an interval, the sample size can vary such that a 20% sample will not necessarily result in exactly 20% of the population being selected into the sample. Under pure sampling theory, the design weight should be 5 in order to reflect that each unit was given a 20% chance of being included in the sample. However, tracking the selection probabilities through time (and the addition of births) would become quite cumbersome, and therefore, this document is suggesting to simplify the design weight to N/n^*, where n^* *is* the number of units that fell within the sampling interval.

$n_{h\,(in-scope)}$ is the total number of sampled units in stratum *h* that are believed to be in-scope (active, in-business). It is calculated as:
(*number of sampled units minus number of out-of-scope units*), and

$n_{hr(in-scope)}$ is the total number responding units (complete and partial) in stratum h that are in-scope.

Caution:
If all non-responding sampling units are imputed, the survey could be considered to have a response rate of 100%, and the entire sample would be assigned the design weights. However, the estimated variance would need to reflect that some of the records have been imputed. Otherwise, the variance will be underestimated.

Whenever calculating weights, it is prudent to quickly check that the weights within each stratum sum to the total number of records within the stratum and that the sum of all weights across the entire sample sum to the size of the population. If this does not occur, an error has been made. The weights and the weighting scheme will need to be reviewed and revised until the weights sum correctly.

Examples

To help in applying the sequential sampling method, two examples can be found in the attached Excel files. The first example follows a simple example through initial sample selection, several iterations of rotation and estimation. To keep things simple, only 1 stratum is considered where the population size is 150 and the sample size 36, with a 25% rotation rate. This example demonstrates how Excel can be used to assign the PRN, assign the estimation weights and conduct rotation. It is noted that the formulae in the last sheet are slightly more automated for calculating the nonresponse adjusted weights. Additionally, it is recommended that for a survey application, the weights should be fully generated by the system in order to eliminate human error. Also, if Excel is used for processing a survey, procedures should be developed to check all formulae and results for accuracy due to calculations being hidden and easily perturbed through the various survey functions. In fact, due to these easily encountered issues with Excel, it should be a last choice for processing survey results.

For the second example, the coordination of two surveys is presented in order to demonstrate the importance of selecting starting points that minimize the occurrence of having the same unit selected in both samples. In this situation, two samples are followed through time. The first sheet sets up the sampling parameters that are simulated in all of the remaining sheets. With all results being driven by the sampling and rotation rates highlighted in orange. The 2nd and 3rd sheets show how the two different sampling schemes would perform over time, with the 3rd sheet showing how one needs to review the sampling overlap

every year because once the samples start to overlap, there is the potential for a high level of overlap for an extended period of time. The 4th sheet sets up an example where the take-some and take-few strata boundaries across two surveys differ. The 5th sheet shows how this situation essentially creates two subsamples cycling through the population, with there being a compromise of the survey overlap between the various combinations of the different strata. This last example was included to provide a visual understanding of the sampling dynamics that will be at play in a real-life situation where the strata boundaries differ between coordinated surveys.

Note that since the results in the file are driven by the parameters entered in the orange shaded area of the first sheet, someone developing a sampling procedure might be interested in this file for more than just understanding the dynamics of using sequential sampling to coordinate two surveys. Firstly, if reviewing the possible overlap outcomes of two strata overtime, the parameters in the first sheet can be updated with the potential results being displayed in the two following sheets for review. Secondly, if developing an automated procedure for conducting the sample selection and rotation, some of the formulae in these sheets could be helpful.


PRASC_Example_Seq uential Sampling - On


PRASC_Example_Seq uential Sampling - Co

10. Domain Estimation

Domain estimation is the process where estimates for subgroups of the population are calculated. As briefly described in section 6 (Stratum Jumpers), if a sampled unit was predominantly Manufacturing when the survey was designed (and therefore in the stratum with fellow manufacturers), then through an updating mechanism (administrative data, survey responses, etc.) it is detected to be predominantly Wholesaling, it will not change its stratum. When calculating its weight, it will also remain in its design stratum. The only time that the new information will be taken into account will be during the domain estimation phase. At this point, the unit will contribute its weighted values to their new group. This is necessary so that non-sampled units demonstrating the same change in characteristics will be properly accounted for.

Domain estimation is also used to calculate estimates for sub-groups of the population that were not included in the design phase. For example, a survey could have its strata based on 2-digit ISIC and groups of revenue, but it may be

desirable to calculate estimates for different subgroups[4] such as geographic regions or employer size groupings.

In a nutshell, for domain estimation, the weights are calculated based on the strata where the units were selected and the most recent information is used to determine the subgroup (domain) to which the unit will contribute. To calculate a domain estimate, the survey responses are multiplied by their associated survey weight (adjusted for nonresponse) and the values are summed within the domains (groups) of interest. If there is any confusion, please refer to the examples in section 6 (Stratum Jumpers).

11. Redesigning Samples

As the example through time showed, after a number of years, it is possible that the coordinated samples will eventually have a very high level of overlap. Also, as the frame ages a large number of deaths could be carried through the survey which can erode survey efficiency. As such, at some point in time, it will become necessary to redesign the sample.

To maximize the time before the samples start to have overlap issues, it is recommended that, prior to implementation, simulations of various sampling intervals be experimented with in order to select an optimal option. Similarly, prior to redesigning a survey, extra efforts should be invested in ensuring that the frame is as up to date as possible in order to optimize the long-term efficiency of the survey.

Once it is deemed that a group of coordinated samples have come to the end of their lifespan, the best strategy is to completely redesign the samples by assigning new PRNs to the common frame, recreating the strata and redrawing entirely new survey frames and samples. It is noted that certain troublesome strata can be redesigned without redoing the entire sample design; however, this could create more complications and issues than it is worth. This is because different surveys will define their strata differently, and therefore it will not be as simple as redesigning only a few strata in a survey (units in a stratum in Survey A will be in multiple strata in Survey B). The redesign will likely need to occur within a common level of stratification between the coordinated surveys, and in the case of business surveys, would likely involve redesigning all of the strata within certain industries. Trying to execute such an approach, maintain it long term and have it documented well enough for a new person to follow through time would likely create more long-term issues than the ones that were solved short term.

---

[4] It should be noted that important subgroups should be included within the stratification procedure in order to ensure that the survey is designed to ensure sufficient responses are received in order to produce reliable estimates and to profit from the estimating accuracies that come with sampling from homogeneous groups of units / strata. However, it is recognized that additional stratification criteria will generate additional strata and therefore a larger sample size will be required in order to have accurate estimate for each stratum. A balance needs to be struck.

Typically, a complete redesign of a sample could result in an undesirable instability of the estimates. Thankfully, business populations tend to be quite skewed which leads to the largest units having the highest sampling rates, and therefore some stability will naturally be achieved. The advantages and disadvantages of the redesign will need to be weighed prior to embarking on a redesign.

12. Closing remarks

As previously mentioned, using PRNs offers a fairly straightforward sampling strategy to implement. It allows for the coordination of samples across surveys and sample rotation. It also has the benefit that the sample selection can be viewed as a simple random sample which allows for a straightforward calculation of survey weights for estimation and for variance estimation. Note that although the method is simple, it is important that the statistical offices produce their own documentation about their sampling procedures. This is necessary to ensure that future team members will be able to maintain the sample through time, and in a consistent fashion. Such documentation will also offer guidance for the team when it comes time to redesign the sample.

The survey design should have two documents associated with it. The first should outline how the sample was initially designed, the decisions that were made and the reasons for the decisions. The second document should summarize how rotation is conducted and the reasons why the rotation rate and the rotation frequency were chosen. Both documents should also contain information about where all files associated with the processes are stored, along with clear and simple instructions of how the files are to be used and procedures executed. File names and locations will change, which will create the need to update the document each time the sample is rotated. At all times, all documentation should be stored in an area where it is easy to find.

For example, if sample maintenance is performed in an Excel spreadsheet, then there should be a tab containing all pertinent notes as well as notes within each sheet indicating the tasks and checks to be performed. Similarly, if the process is developed in an automated fashion using facilities such as SQL or SPSS, the program needs to contain comments and documentation for future programmers, and there needs to be a user manual for the personnel who will conduct the work. Again, the user manual should include step-by-step instructions for using the programs, checks to be performed in order to ensure that the sample maintenance was completed accurately as well as a summary of what each procedure does.

All pertinent information should be saved in a structured manner so that it is easy to find, and to determine which files belong to each survey year. The easiest

way to achieve this is to have a directory dedicated to each survey occasion, with subdirectories for each survey step (the Generic Statistical Business Process Model (GSBPM) is often used as a guide for setting up high-level directories). Within the sampling directory, there should be a copy of ALL documents related to the process. This will create duplicated information, but the convenience of having all related information for each year together will far outweigh this slight drawback.

Additionally, all directory and file names within the survey's directories should contain the reference period of the survey. When having two files open at the same time, it serves an easy way to distinguish between the years of information, while helping to prevent previous files accidentally being overwritten by the current year's work.

Hopefully, this summary will provide all of the details required by a Caribbean statistical office to design and maintain sample surveys. In the event that further information is required, some reference documents are provided.

The Hidiroglou and Lavallée (2009) paper in the list of references provides a nice overview of various survey steps. Some areas appear to present complex formulae; however, they are clearly explained and with a bit of time invested, most can be easily understood. At a minimum, it can offer ideas at various stages of the survey development that can be further researched in order to resolve issues that are encountered. In particular, related to this document, there is a description of using PRNs for sampling as well as adjusting sample sizes to account for expected non-responses and out of scope records.

The paper by Lavallée (1996) provides additional details about the bias that can be created when survey feedback about the current status of businesses is used to maintain the survey frame. The Ohlsson (1995) paper provides information about other variations of using Permanent Random Numbers to manage sample selections.

References

Hidiroglou, M. A. and Lavallée P. (2009). *"Sampling and Estimation in Business Surveys"*, The Handbook of Statistics 19 A.

Lavallée, P. (1996). *"Frame Update Problems with Panel Surveys"*, Proceedings of the Statistical Society of Slovenia, Radenci, 25-26 November 1996.

Ohlsson, E. (1995), "Coordination of Samples Using Permanent Random Numbers" In *Business Survey Methods*, John Wiley and Sons, New York, 1995.

Appendix C: Frequently Asked Questions (FAQ)

This annex provides a list of frequently asked questions with answers and procedures that arose during the PRASC technical assistance for the SQL SBR.

Q: If an establishment has been ceased on the SBR, can it be reactivated later or do we need to create a new establishment under the enterprise?

A: If an operating entity (establishment) has been ceased, it will not be possible to automatically reactivate the unit and the establishment will need to be recreated manually on the SBR system.

Q: How can I reactivate an enterprise with a ceased (6) business status?

A: If an enterprise (legal unit) got assigned a business status=6 (Ceased operating) after a long period of inactivity (more than 36 months) and you would like to reactivate the enterprise, because you have received an administrative record that indicates the unit is back in operation, then you would need to do the following two steps:

1. Batch load the record(s) by using the Excel Template with a recent reference period, with a business status = 3 and the data received from your administrative source. This batch load will reset the business status=6 (ceased operating) to a business status=3 (Inactive). This action will permit you to run the reactivation process.
2. Run the batch reactivation process from the utilities menu on the SBR **(Utilities>Batch Reactivation**). This process will reactivate all your enterprises from a business status=3 to a business status=2, if your reference period is recent enough (less than 36 months from today's date). If you have very few records (less than 10 or 20 records) in that situation, you can always manually run the reactivation process from the update menu. With the manual process, you will need to do it record by record.

Note: At step 1, the data could also be loaded with a business status=4 (Bankrupt) or 5 (Dissolved) similarly to a business status=3 (Inactive).

Q: If I ceased an enterprise or an establishment, is the record deleted from the SBR database?

A: No, records are never physically deleted from the SBR, only the business status is changed. Records are kept on the SBR for future analysis. When extracting a SUF, all records can be viewed.

Q: If I have linked two legal units by mistake, how can I remove the link between the two legal units?

A: There is no direct way to remove the link between two legal units. The following are the steps to follow to remove a link between two legal units, the parent and the subsidiary. In order to break the link between legal units and to keep the data integrity on the base, it is necessary to adopt a two-step approach. As an example, S1 is the parent of S2 where S2 is the subsidiary of S1.

1. First, from the tree view on the SBR browser, select the child legal unit (Subsidiary) S2. Select from the menu bar update>Cessation>Bankrupt or update>Cessation>Dissolved then commit in order to remove the link between S1 and S2. The two actions have the same result of removing the link.
2. To reactivate the Bankrupt or Dissolved legal unit S2, go to the S2 legal unit and select from the menu bar Update>Reactivation>Change status to alive then commit. The legal unit will be reactivated with a business status Alive.

Warning: If the subsidiary S2 has operating entities below as Childs, all the operating entities will be removed and lost. They will need to be manually recreated.

Q: Can I load Informal Businesses in Batch?

A: Yes, it is possible to load informal businesses in batch mode by using the Batch Load Template. You can simply use Other_ID as unique business identification number and assign a different sequential number to each business with a recent reference period. In addition, you can assign the owner's name as a legal business name if the informal business does not have a business name. For further information on the topic, please refer to Appendix F.

Q: Will the informal businesses be kept alive after 3 years without any updates on the SBR?

A: No, the informal businesses will be inactivated after 3 years with no recent update to the size variables, as it is the case with any other business types. In order to keep them alive after 3 years of inactivity on the SBR, another batch load of the informal businesses will need to be run with a recent reference period. As an example, you can use today's year and month as a reference period. By doing that, the informal businesses will be kept alive for three more years.

Q: Do I need to assign a size variable and an ISIC classification for the informal businesses?

A: We recommend assigning a size variable and an ISIC classification to each informal business in order to have them potentially available for survey sampling (Enterprise Flag = 1 and Establishment Flag=1). If you do not have any indication of the revenue or the number of employees, you can always assign a value of 1 for the number of employees.

Q: Do I need to assign a prospect flag (prospect Flag=1) to informal businesses?

A: No, you assign a prospect (Prospect Flag = 1) only if you are not sure if the business already exists on the SBR database.

Q: Is-it possible to modify two unique business identifiers linked together?

A: Yes, it is possible to manually modify a linked key by using the update function on the SBR. As an example, if you have linked Tax_ID=1 with SS_ID=12 during the load and you discovered that the link should have been Tax_ID=1 with SS_ID=11, then you can simply update the legal entity by using the Update function on the SBR (**Update>Update>Legal**) and changing SS_ID=12 to SS_ID=11. **This update can be done as long the SS_ID=11 is not linked to another Tax_ID and the SS_ID11 is not loaded yet on the SBR.** If the SS_ID11 is loaded on the SBR, then there is already a Statistical number assigned to SS_ID11 and it is not possible to link it to an existing TAX_ID.

In general, it is possible to manually change a link key if the secondary source is not loaded yet. It will become very limited to update linked keys when secondary sources are loaded, because records will have a statistical number assigned. In that situation, the only way to link two records with an assigned Statistical Number, it will be to integrate the two records (see section 5.6 above **– Steps for resolving duplicates** or the FAQ below - **Is there a specific procedure to follow when duplicates are found on the SBR?**)

Q: Does the SBR calculate the size variable for the legal entity based on the sum of the size variable for the operating entities?

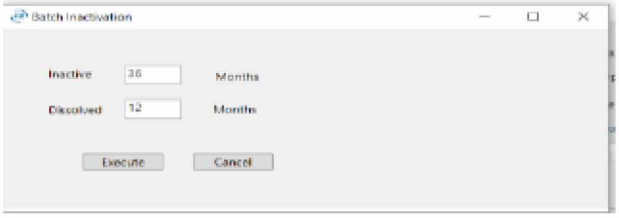
A: No, the size variables do not get recalculated at the enterprise level (legal entity) when updating the size variables at the establishment (operating) level.

Q: Is a legal entity with active operating entities would get inactivated if its size variables have not been updated over time.

A: No, the enterprise (legal entity) will not be inactivated when at least one active operating entity (at least one size variable more recent than 36 months) exists.

Q: How can I change the parameters of the batch inactivation function for the number of months before a unit got inactivated or dissolved?

A: By default, the SBR is set to 36 months before an enterprise got inactivated and 60 months (36 + 24 months) to get dissolved. In order to change the number of months, you will need to send a request to your IT expert to modify the table (TblInactivationCriteria) in the database. Your IT expert has received the necessary training on how to change the parameters in the batch inactivation table. The SQL script to modify the table content is in the last distribution kit (June 2023) of the SBR and the instruction on how to proceed can be found in the SBR Technical Guide under the section 2.2.3.



Q: If new data (size variable) is loaded for inactive units using the batch load template, would it reactivate the unit automatically on the SBR?

A: When updating the SBR with more recent size variables, the Batch Load Excel Template will not reactivate units automatically. It will be necessary to run the Batch Reactivation function from the Utilities menu in order to reactivate the units (Utilities>Batch Reactivation).

Q: Which business status can be reactivated with the Batch Reactivation Function?

A: Only legal units with a business status of inactive or bankrupt or dissolved will be reactivated.

Q: What is a prospect flag and what to do with it?

A: Please refer to section 3.4.3 and 5.6 for a definition and on how and when to set the prospect flag.

Q: Will the prospect flag be set to 0 if we update the data with a prospect flag equal to blank with the Batch Load Excel Template?

A: No, updating the data with a prospect flag = blank during a batch update will not alter the value of the prospect flag on the SBR. If you want to change a prospect flag from 1 to 0, then you need to set the prospect flag to 0 in your batch update load.

Q: Is there a specific procedure to follow when duplicates are found on the SBR?

A: Yes, there is a specific procedure before processing to the Integration of two units on the SBR.

1. Identify which entity will be the successor of the integration (ideally, pick the one from the base source).
2. Write down the following values from the unit to be integrated:
    - Size measure(s)
    - Administrative ID
    - Name and address, as required
3. Remove from the unit to be integrated by processing the following function Update>Update>Legal Entity:
    - Remove the prospect flag
    - Remove the administrative ID
    - Add a comment (the reason) and **commit changes**
4. Copy information from step 2 above to the successor unit by processing the following updates on the SBR Update>Update>Legal Entity:
    - Size measure(s) if applicable
    - Administrative ID
    - Name and address if applicable
5. Select the unit to be integrated by processing the following function: **Update > Cessation >Legal Entity> Integrated.**
    - Effective Date: Today's date and **commit changes**
    - The SBR will change the status of the entity to be integrated to Integrated (Business Status=7).

Q: Is it important to enter a priority number in the Priority Variables Excel Template?

A: Yes, it is important. If the priority is left blank for the data source you want to load, the SBR application will not update/overwrite the data even if the information is from a newer reference year.

Q: How can I establish my variable priorities?

A: When you have two or more input data sources, it is important to establish the priority of each data source. We usually assign priority 1 to the base source data file or the data file you are most confident with. As an example, if you receive the social security data and Tax data from your administrative data partners, then you will need to decide which information you trust the most and assign the priority accordingly. If you trust more the social security data because of the completeness of the information, the frequency of reception and the quality of the data, then you can assign a priority 1 to this source to all data variables and a priority 2 to the TAX data variables. However, if you believe that the ISIC from TAX is of a better quality or the ISIC version is newer than the one from Social Security, then you can always assign a priority 1 to the ISIC variable from Tax data and a priority 2 to ISIC variable from Social Security. Another example, the revenue from the Tax file could have a priority=1 while the number of employees from Social Security file could have a priority=1.

Q: Is it possible to have multiple priority '1' for various input data sources in the Priority Variables Excel template?

A: Yes, it is possible. However, it is recommended to assign different variable priority number to different input data sources in order to prioritize your updates. Remember, priority 1 is the highest priority and priority 9 is the lowest priority.

Q: I am trying to load records on the SBR and I am getting the following error: Run time error – Conversion failed when converting date and/or time from character string. What I should be doing?

A: You are getting this error because you have entered an invalid date format in one of the date fields in your Batch Load Excel Template. If you are not sure of the date format, you can verify in the All Columns worksheet for the fields format. In order to identify the record(s) with an invalid date format, you can always apply a filter to the worksheet Paste Your Data and try to locate a strange date format.

Q: I am trying to load records on the SBR and I am getting the following error: Either you are not connected or the server is down. What I should be doing?

A: You are getting this error either because your connection to the file server is down or you have not entered correctly your server name in your ***DatabaseInformation*** work sheet tab in your Batch Load Excel Template. If you are not sure of the server name, you can verify with your IT expert. If your server name is entered correctly, then verify with your IT expert if the file server is down.

Q: I received the following Error message in the 'BatchProcessingErrors' sheet in my Batch Load Excel Template: Warning: Already linked External Admin Key TAXPAYER_ID. What is the reason?

A: The error is caused by the fact that there is already another unique business identification (e.g., Social Security ID number) linked with that Taxpayer ID number on the SBR. As an example, the new record that you are trying to load has Tax_ID=12345 and SS_ID=101010. However, in the database, there is already the record SS_ID=202020 that has the Tax_ID=12345. One of the two records doesn't have the correct Tax_ID linked. You can use the SBR interface to look at the data to determine which record has the appropriate Tax_ID and make the necessary corrections.

Q: How can I map my administrative data to the Batch Load Excel Template?

A: To map administrative data to the Batch Load Excel Template:

1. Identify the data source received (e.g., Tax file or Social Security file, etc.). This information will indicate which Update Source Code (Tax or Social Security, etc.) to use when loading the data. It also indicates the unique identification number to be used to uniquely identify the legal entity (Taxpayer_ID or Social_Security_ID, etc.).
2. The vintage year of the administrative data source will help you to determine the ReferencePeriode (YYYYMM) to be used when loading the data.
3. The tombstone information will permit populating the legal name, the operating name and the addresses.
4. Map the classification variables such as ISIC and business type.
5. The size variables will indicate which size variable to populate. As an example, if you received the tax file then you might be able to populate the revenue variable.
6. The administrative file usually contains contact information that will permit you to populate the contact information on the Batch Load Excel Template.

Note: The Excel template has a specific format for each variable. The SBR application will only accept records that respect these formats. The spreadsheet 'All Columns' in the Batch Load Excel Template details the format of each variable.

Q: Why do I need to enter an 'Admin file path and name' while going through the Batch Load process?

A: This is strictly for documentation purposes. If an incorrect path and name are entered, it will have no impact on the process, but the batch job documentation will be simply incorrect.

Q: I obtained the following error 'Invalid ISIC' in the BatchProcessingError worksheet in the Batch Load Excel Template. What should I do?

A: This message is generated because you are trying to load an invalid ISIC. The system will accept only valid ISIC 4-digit code. There are two options to make the correction. First, you can simply reload the same record with a correct ISIC 4-digit code. As a second option, you can update or insert the ISIC code with the update option in the SBR. We suggest to simply reload the record with the valid ISIC code.

Q: I noticed that records loaded with a prospect flag of 1 are generating a warning message in the BatchProcessingError worksheet in the Batch Load Excel Template. Is this correct?

A: It is correct, it is a warning that is generated by the system to indicate that you have loaded a prospect flag = true (1). It is also a reminder to resolve that case in the future.

Q: Is there a need to supply mandatory variables for loading new records (initializing) on the SBR when using the Batch Load Excel Template?

A: Yes, there are mandatory set of variables that need to be supplied when creating new business on the SBR with the Batch Load Excel Template. You will need to supply a unique business identification (TAXPayer_ID or Social_Security_ID etc.) with a reference period and a legal name. The other variables are optional. However, we recommend supplying as many variables as possible.

Q: Is there a need to supply mandatory variables for updating existing records on the SBR when using the Batch Load Excel Template?

A: Yes, there are mandatory variables that need to be supplied when updating existing business on the SBR with the Batch Load Excel Template. You will need to supply a unique business identification (TAXPayer_ID or Social_Security_ID etc.) and a reference period. The other variables are optional; however, we recommend supplying as many variables as possible.

Q: I want to update the SBR from our existing business register for many years. Should I load year by year and in which order?

A: Yes, it would be preferable to load year by year for all years of data. We recommend that the files are loaded in ascending order of the year (i.e., start with the oldest year of data and finish with the most recent year of data).

Appendix D: Loading Administrative Data in Subsequent Cycles

This annex provides instructions for loading administrative data into the SBR on an ongoing basis, once the initial load has been performed.

1. When loading subsequent administrative data cycles, it is a good practice to generate a SUF before the subsequent load. This is done from the utility menu in the SBR GUI application. Once it's generated extract the SUF into an Excel spreadsheet and save it in a secure directory as a backup version. In the case of a complication during the subsequent loads, the SQL Survey Universe table that corresponded to the date and time prior to the subsequent loads could be used to recreate your SBR database.
2. Load your new base file (e.g., Social Security file) by using the Excel Batch Load Template (Admin Data Batch Update Tool).
3. After loading the new base file, run another SUF and extract it into an Excel Spreadsheet. This SUF will be your working copy to for the record linkage process.
4. In the working copy of the record linkage workbook:
   a. Copy full SUF into a new tab or worksheet named e.g., SUF Source.
   b. Remove establishment only entities (operating entities) from the SUF. These can be identified with Legal Flag = false.
   c. Convert the Prospect Flag variables from True and False to 1 and 0, respectively.
   d. Recommend reducing SUF variables to the ones that will be used to match to the complementary file plus the Prospect Flag e.g., Unique Business Identification variables (Registrar Id, Social Security Id, Taxpayer Id and Vat Id) and some common variables (Legal Name, Legal Address, Operating Name, etc.) and Prospect Flag.
   e. If not already converted into a table format, make sure to do so and give the table a user-friendly name to be used in Fuzzy Lookup. This can be accomplished under Insert and selecting table. Once the data is in a table format under design, you can assign a user-friendly table name e.g., SUFsource this will be the table to link the complementary file to.
5. Copy in the complementary file in a new tab, again insert table and apply a user-friendly table name e.g.Taxsource. For the complementary files, you will need to match your admin files (Unique Business Identification variables) with the SUF in order to determine whether or not your complementary file match with your base file (in this case, Social

Security). **Do not forget to transfer the prospect flag values of 0 or 1 from the SUF to the matched records on the complimentary file.**

6. In a new tab, start with a direct match by using the Business Unique Identification Number of your complementary file and the SUF file. As an example, if your complementary file is the Tax File then you will be matching by Taxpayer identification number with the SUF. This process will allow you to distinguish the already matched records from the unmatched and is necessary to preserve the integrity of the content of the previously loaded Prospect Flag variable.
7. To identify the matched records in the complementary file, apply the conditional formatting, new rule (or manage rule if already set up) to use a formula to determine which cells to format. The formula to use would be the COUNTIF function. As an example, if the common variable between the two files is the Tax_Payer_id in column 'A', then the formula will be defined as COUNTIF('Sheet Name'!$A:$A,A1).
8. Load the matched complementary file records from step 6, onto the SBR, using the Excel Batch Load Template (Admin Data Batch Update Tool).
9. For the unmatched records from step 6, rematch these records with your SUF file by using the common variables (Legal Name, Legal Address, Operating Name, etc.). We recommend this process in order to verify if we could increase the number of matching records between your base file and your complementary file.
   Note: In order to optimize the matching process, we recommend to only match the records from the SUF where there is no base file unique business identification (e.g., SUF records with no Taxpayer ID).
10. For the common variable matched records from step 8, add any Unique Business Identification variables, such as the Social security ID from the matched record on the SUF, Taxpayer ID and Prospect Flag value equal to zero (0) to the complementary file. Once file has all the necessary variables from the match, load it onto the SBR by using the Excel Batch Load Template (Admin Data Batch Update Tool).
11. For the unmatched records from step 8, add Prospect Flag value equal to one (1) to the complementary file and load them onto the SBR by using the Excel Batch Load Template (Admin Data Batch Update Tool).
12. Repeat steps 5 to 10 for each complementary file, if applicable.

Appendix E: Batch Load of multiple contact information

This section presents information on loading multiple contact information for the same entity onto the Statistical Business Register (SBR) in batch mode. It provides concrete examples and tips.

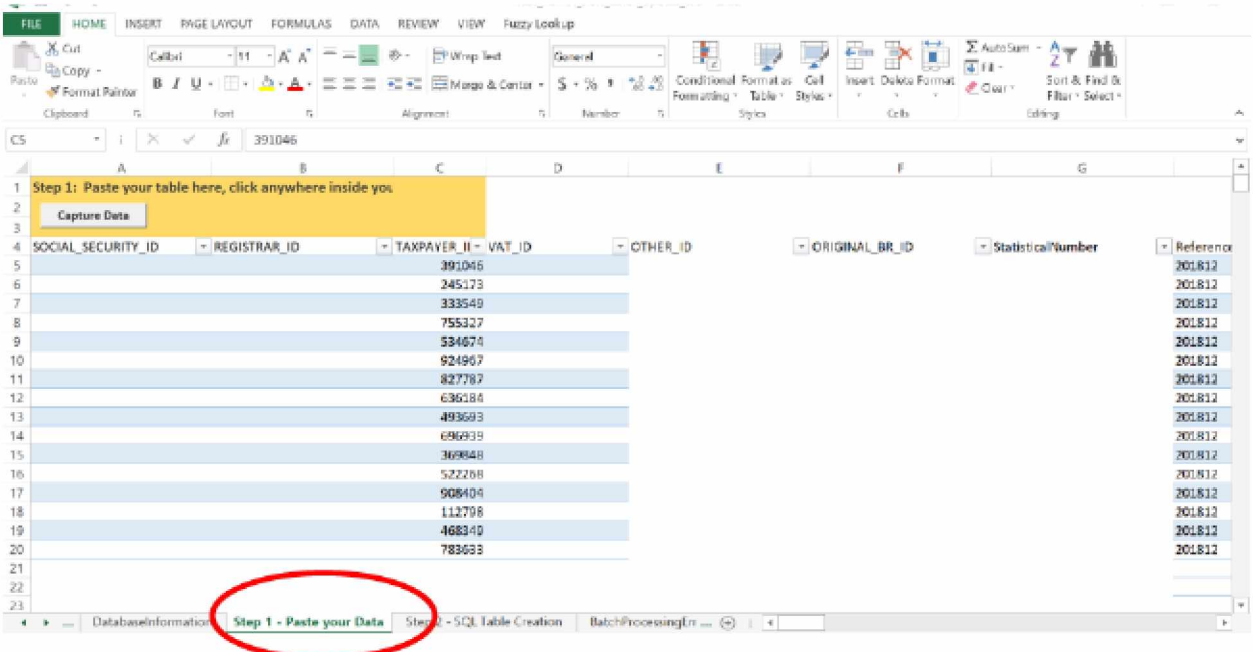
The example in table 1 provided below represent the multiple contact information that the administrative data source IRD contains.

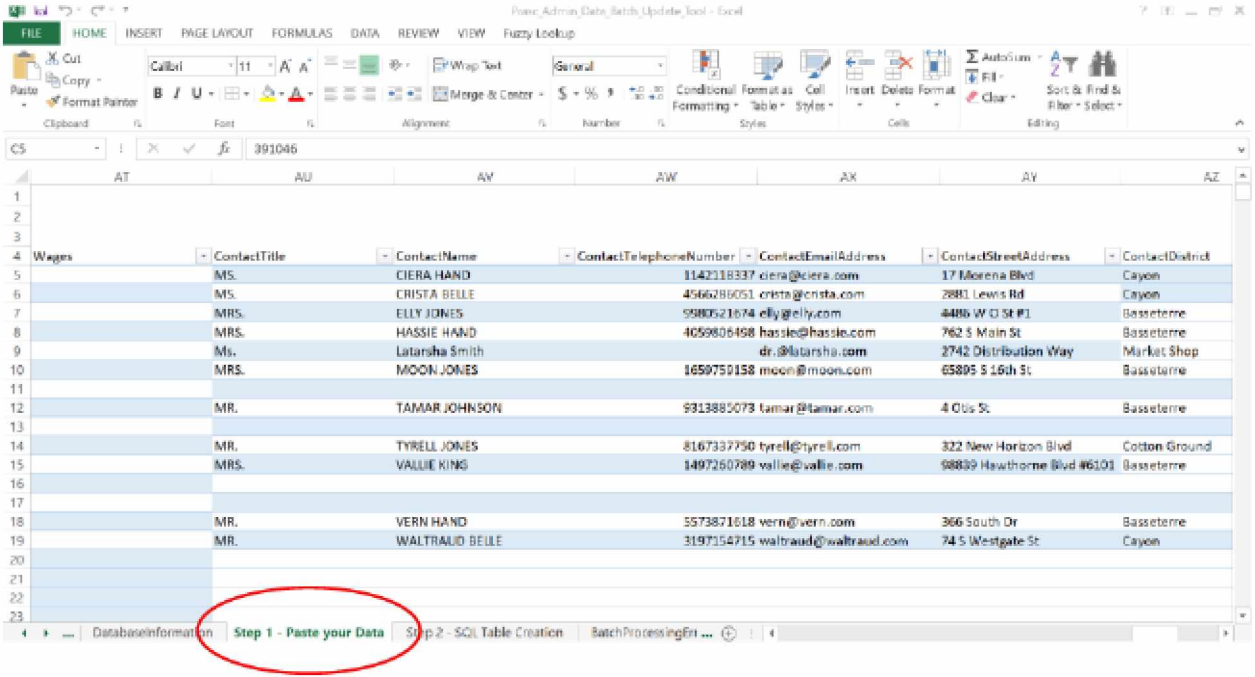
Table 1: Information collected from the collected from the Tax Department (IRD)

TAXPAYER ID	Reference Period	Contact Title	ContactName	Contact TelephoneNumber	ContactTitle2	ContactName2	Contact TelephoneNumber2
391046	201812	MS.	CIERA HAND	1142118337	MS.	CIERA HAND2	123456789
245173	201812	MS.	CRISTA BELLE	4566286051	MS.	CRISTA BELLE2	123456789
333549	201812	MRS.	ELLY JONES	9980521674	MRS.	ELLY JONES2	123456789
755327	201812	MRS.	HASSIE HAND	4059806498	MRS.	HASSIE HAND2	123456789
534674	201812	Ms.	Latarsha Smith		Ms.	Latarsha Smith2	123456789
924967	201812	MRS.	MOON JONES	1659759158	MRS.	MOON JONES2	123456789
827787	201812						
636184	201812	MR.	TAMAR JOHNSON	9313885073	MR.	TAMAR JOHNSON2	123456789
493693	201812						
696939	201812	MR.	TYRELL JONES	8167337750	MR.	TYRELL JONES2	123456789
369848	201812	MRS.	VALLIE KING	1497260789	MRS.	VALLIE KING2	123456789
522268	201812						
908404	201812						
112798	201812	MR.	VERN HAND	5573871618	MR.	VERN HAND2	123456789
468349	201812	MR.	WALTRAUD BELLE	3197154715	MR.	WALTRAUD BELLE2	123456789

The following steps describe how to batch load information onto the SBR in order to maintain more than one contact information.

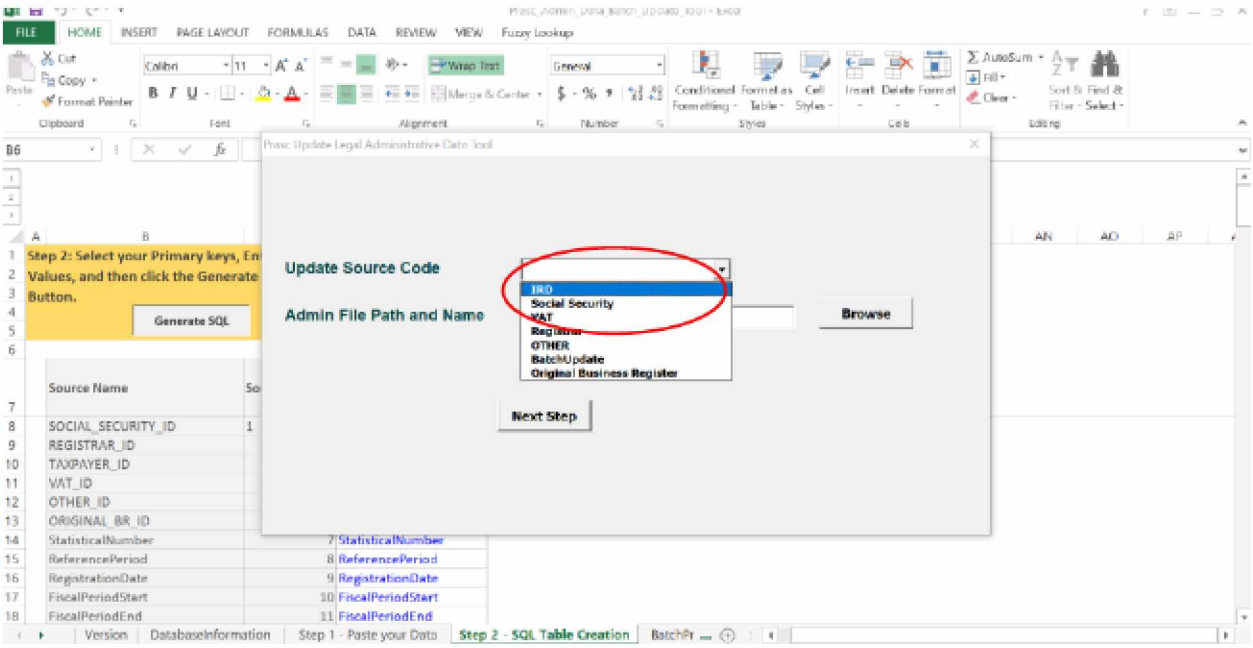
Step 1: Copy the data to '**Paste Your Data**' tab of the SBR Batch Load Excel Template keeping only the first contact. As an example, data from table 1 above have been copied to the Batch Load Excel Template. See the screenshot below:


FILE
HOME
INSERT
PAGE LAYOUT
FORMULAS
DATA
REVIEW
VIEW
Fuzzy Lookup
Step 1: Paste your table here, click anywhere inside you
Capture Data
SOCIAL_SECURITY_ID
REGISTRAR_ID
TAXPAYER_ID
VAT_ID
OTHER_ID
ORIGINAL_BR_ID
StatisticalNumber
Reference
391046
245173
333549
755327
534674
924967
827787
636184
493693
696939
369848
522268
908404
112798
468349
783693
201812
DatabaseInformation
Step 1 - Paste your Data
Step 2 - SQL Table Creation
BatchProcessingIn

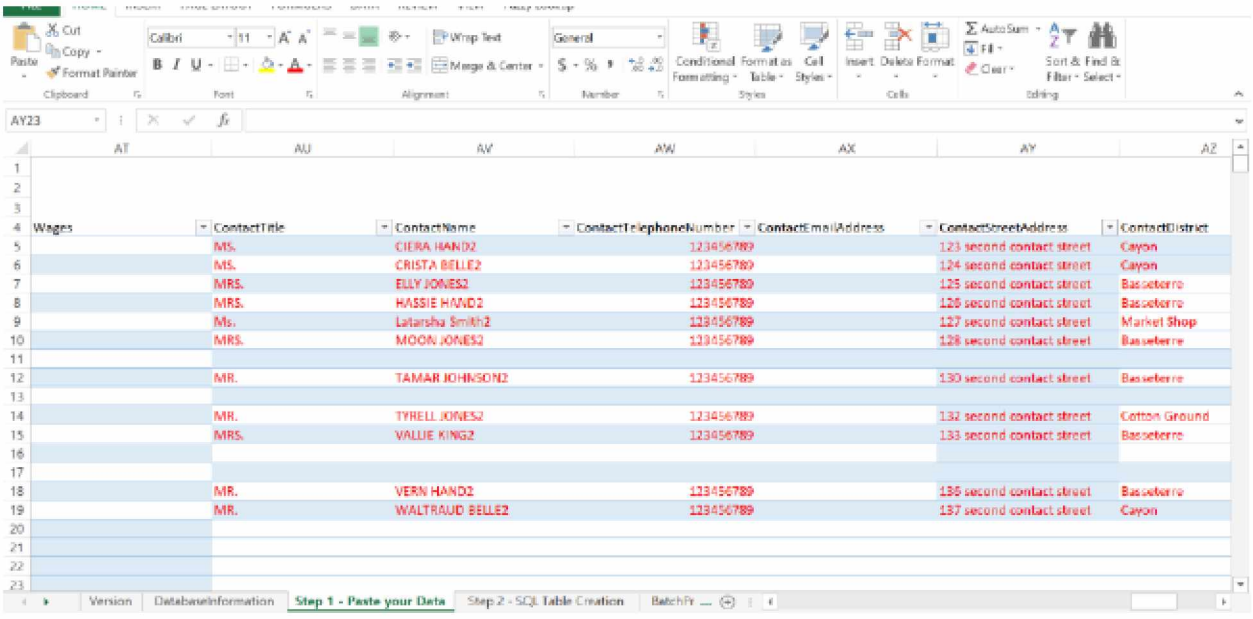


Wages	ContactTitle	ContactName	ContactTelephoneNumber	ContactEmailAddress	ContactStreetAddress	ContactDistrict
	MS.	CIERA HAND	1142118337	ciera@ciera.com	17 Morena Blvd	Cayon
	MS.	CRISTA BELLE	4566286051	crista@crista.com	2881 Lewis Rd	Cayon
	MRS.	ELLY JONES	5980521674	elly@elly.com	4486 W O St #1	Basseterre
	MRS.	HASSIE HAND	4059806458	hassie@hassie.com	762 S Main St	Basseterre
	Ms.	Latarsha Smith		dr.@latarsha.com	2742 Distribution Way	Market Shop
	MRS.	MOON JONES	1650750158	moon@moon.com	65895 S 16th St	Basseterre
	MR.	TAMAR JOHNSON	9313885073	tamar@tamar.com	4 Otis St	Basseterre
	MR.	TYRELL JONES	8167337750	tyrell@tyrell.com	322 New Horizon Blvd	Cotton Ground
	MRS.	VALLIE KING	1497260789	vallie@vallie.com	98839 Hawthorne Blvd #6101	Basseterre
	MR.	VERN HAND	5573871618	vern@vern.com	366 South Dr	Basseterre
	MR.	WALTRAUD BELLE	3197154715	waltraud@waltraud.com	74 S Westgate St	Cayon

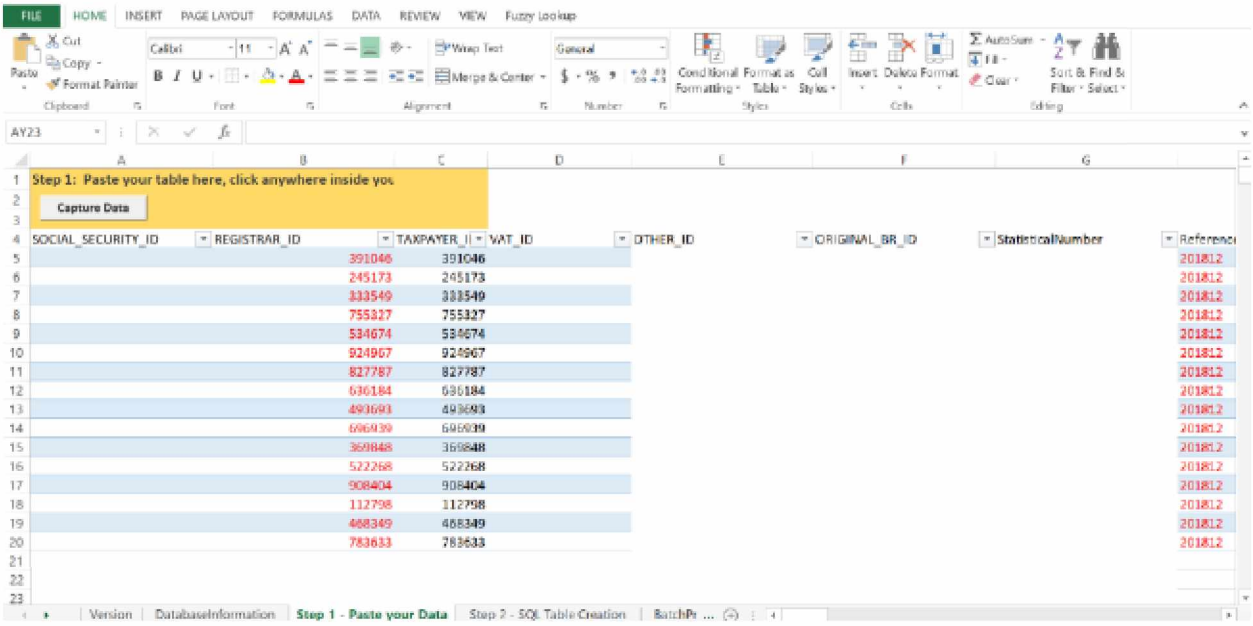
Step 2: Load the data onto the SBR, select capture data followed by Generate SQL with the Update Source Code **IRD**.



Step 3: Copy the information for the second contact information in the '**Paste Your Data**' tab of the SBR Batch Load Excel Template.



Step 4: To generate a unique ID for the second contact, paste the file unique identifier (TAXPAYER_ID in this case) in its appropriate column, but also in the column for an unique identifier of a source that is not likely to be used (REGISTRAR_ID in this case). Populate the column ReferencePeriod with the reference period of the file (201812 in this case). The data values in red are the necessary information to generate a second contact for a given legal entity.

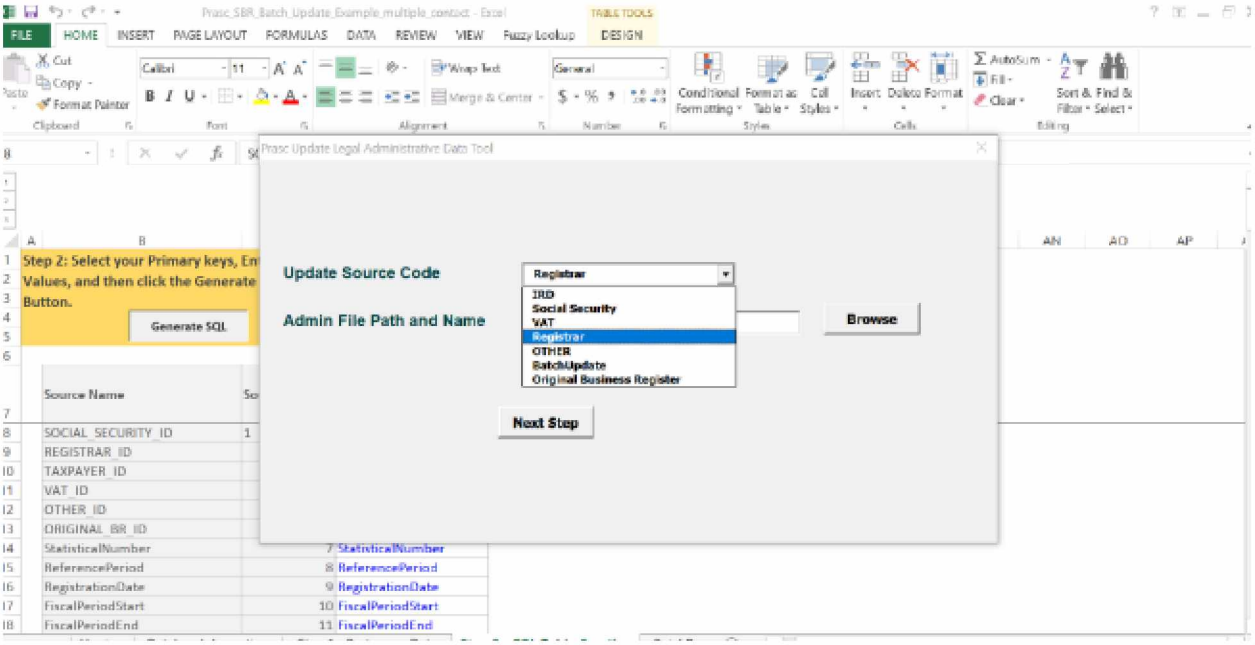


Step 1: Paste your table here, click anywhere inside you

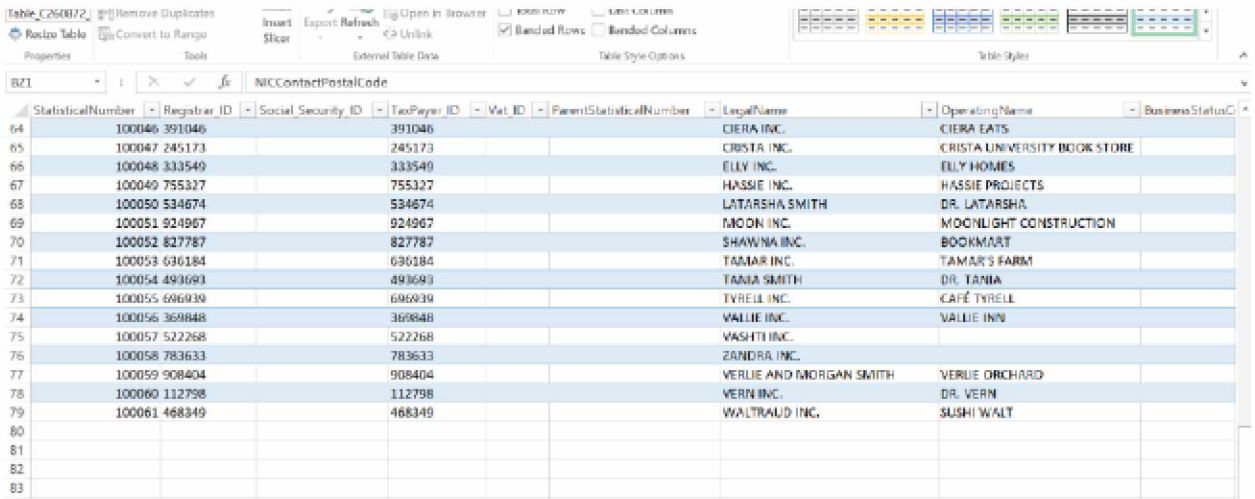
Capture Data

SOCIAL_SECURITY_ID	REGISTRAR_ID	TAXPAYER_I	VAT_ID	OTHER_ID	ORIGINAL_BR_ID	StatisticalNumber	Reference
	391046	391046					201812
	245173	245173					201812
	333549	333549					201812
	755327	755327					201812
	534674	534674					201812
	924967	924967					201812
	827787	827787					201812
	636184	636184					201812
	493693	493693					201812
	696939	696939					201812
	369848	369848					201812
	522268	522268					201812
	908404	908404					201812
	112798	112798					201812
	468349	468349					201812
	783633	783633					201812

Step 5: Load the data onto the SBR, select capture data followed by Generate SQL with the Update Source Code of the source that not likely to be used and for which the unique identifier was pasted into (**Registrar** in this case). This will store on the SBR a second contact information to the specified legal entities.



When generating the SUF, here's how the information has been stored on the SBR:



	StatisticalNumber	Registrar_ID	Social_Security_ID	TaxPayer_ID	Vat_ID	ParentStatisticalNumber	LegalName	OperatingName
64	100046	391046		391046			CIERA INC.	CIERA EATS
65	100047	245173		245173			CRISTA INC.	CRISTA UNIVERSITY BOOKSTORE
66	100048	333549		333549			ELLY INC.	ELLY HOMES
67	100049	755327		755327			HASSIE INC.	HASSIE PROJECTS
68	100050	534674		534674			LATARSHA SMITH	DR. LATARSHA
69	100051	924967		924967			MOON INC.	MOONLIGHT CONSTRUCTION
70	100052	827787		827787			SHAWNA INC.	BOOKMART
71	100053	636184		636184			TAMAR INC.	TAMAR'S FARM
72	100054	493693		493693			TANIA SMITH	DR. TANIA
73	100055	696939		696939			TYRELL INC.	CAFÉ TYRELL
74	100056	369848		369848			VALLIE INC.	VALLIE INN
75	100057	522268		522268			VASHTI INC.	
76	100058	783633		783633			ZANDRA INC.	
77	100059	908404		908404			VERLIE AND MORGAN SMITH	VERLIE ORCHARD
78	100060	112798		112798			VERN INC.	DR. VERN
79	100061	468349		468349			WALTRAUD INC.	SUSHI WALT

Appendix F: Batch Load of Informal Businesses

This section presents information on loading Informal Businesses onto the Statistical Business Register (SBR) in batch mode when collecting data from the census of population or other collection vehicle such as Labour Force Survey (LFS). It provides context, concrete examples and tips for loading large volume of data for this type of businesses.

The example in table 1 provided below assumed that the data are collected from a Visitation Record (VR) and a census questionnaire. From that process, the following information is collected for the Informal Businesses: The census record identification number (Electoral District etc.), the name of the business (often the business owner's name), the address, the business activity and the number of employees. In some countries, it is possible that some additional information such as the GPS coordinate are collected during the VR or on the census questionnaire.

Table 1: Information collected from the VR and the census questionnaire for Informal Businesses:

Census Record Identification Number	Person's / Business Name	Address	Parish	Business Activity	Number of employees	Other potential VR or census data
123456	John Smith	225 Main Street	St. Mary	Repair of Machinery	2	
123457	Mary Brown	122 Beach Road	St. Mary	Bakery Products	1	
123458	Peter Cooks	22 Jerrold Street	St. Mary	Tour Operator		
123459	Palmer King	553 Country Road	St. Mary	Raising Poultry		

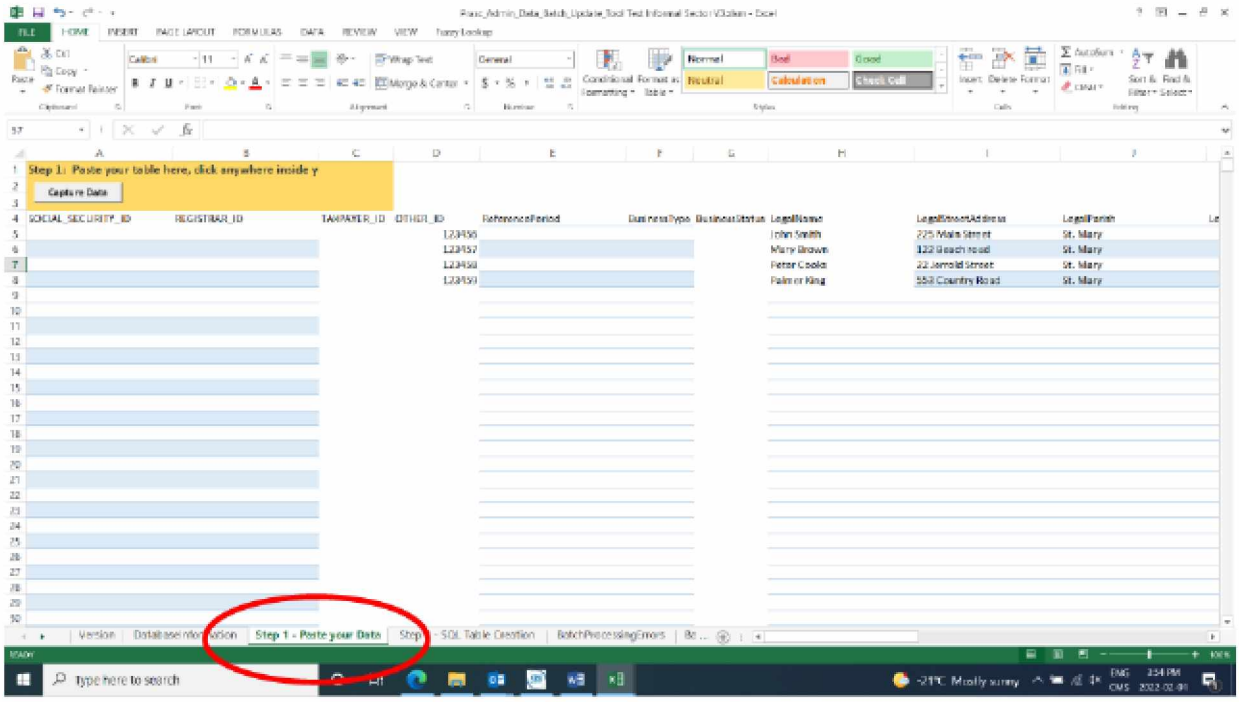
The following steps describe how to batch load information onto the SBR coming from the VR and the census of Population for informal businesses. It is noted that some data could be extracted directly from the VR and/or the census questionnaire but other data values would need to be derived.

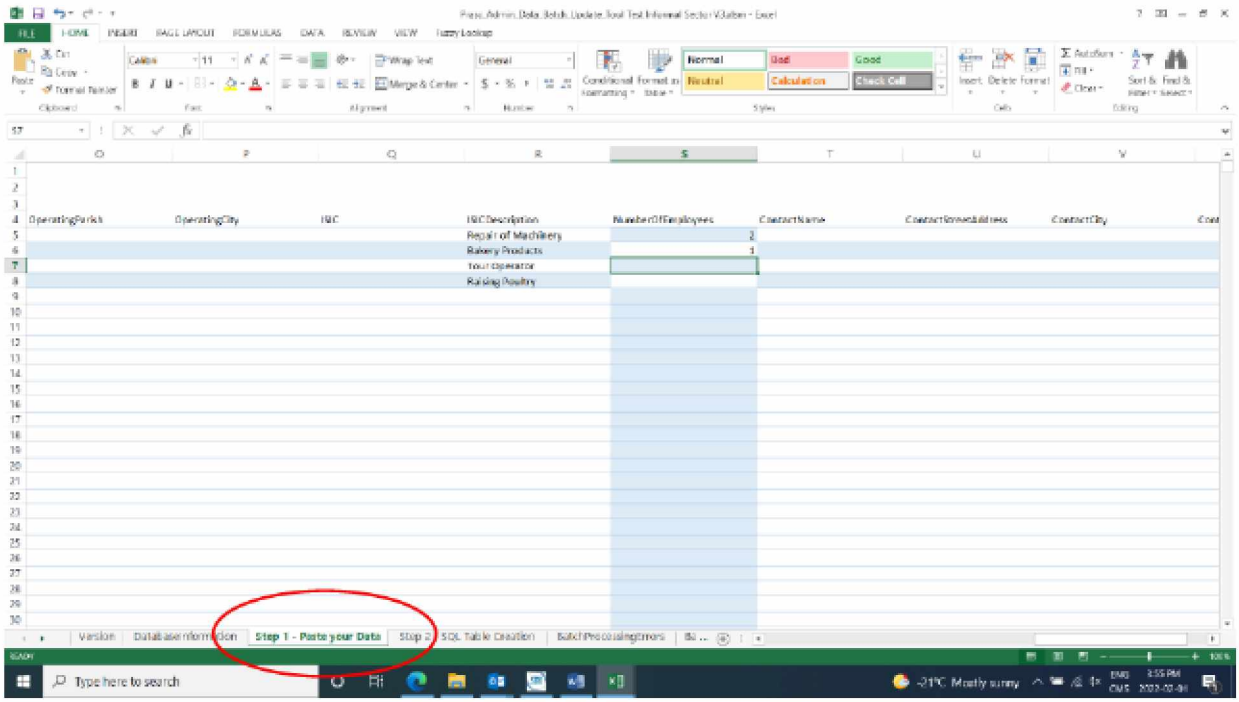
Step 1: Map the Census data variables to the data variables used on the Statistical Business Register (SBR):

VR/Census Data Variables	***SBR Data Variables***
Census Identification Number	OTHER_ID
Person's/Business Name	LegalName
Address	LegalStreetAddress
Parish	LegalParish
Business Activity	ISICDescription
Number of Employees	NumberOfEmployees
GPS Coordinate	LegalGps
Other variables if applicable	

Step2: Select the informal businesses only from the VR. Some countries collect information for registered and non-registered businesses during the VR. It will be important to only select the non-registered (informal) businesses in order to avoid loading duplicate businesses on the SBR. The registered businesses (formal) are usually loaded from your administrative data sources. Record linkage between the VR and the Survey Universe File (SUF) by business name (LegalName, OperatingName or owner's name, depending on the situation) and the address might be needed to identify the registered entities on the VR in order to exclude them from the batch load.

Step 3: Copy the data collected during the VR and/or the census of Population for the informal businesses only to '**Paste Your Data**' tab of the SBR Batch Load Excel Template. As an example, data from table 1 above have been copied to the Batch Load Excel Template. See the screenshot below:

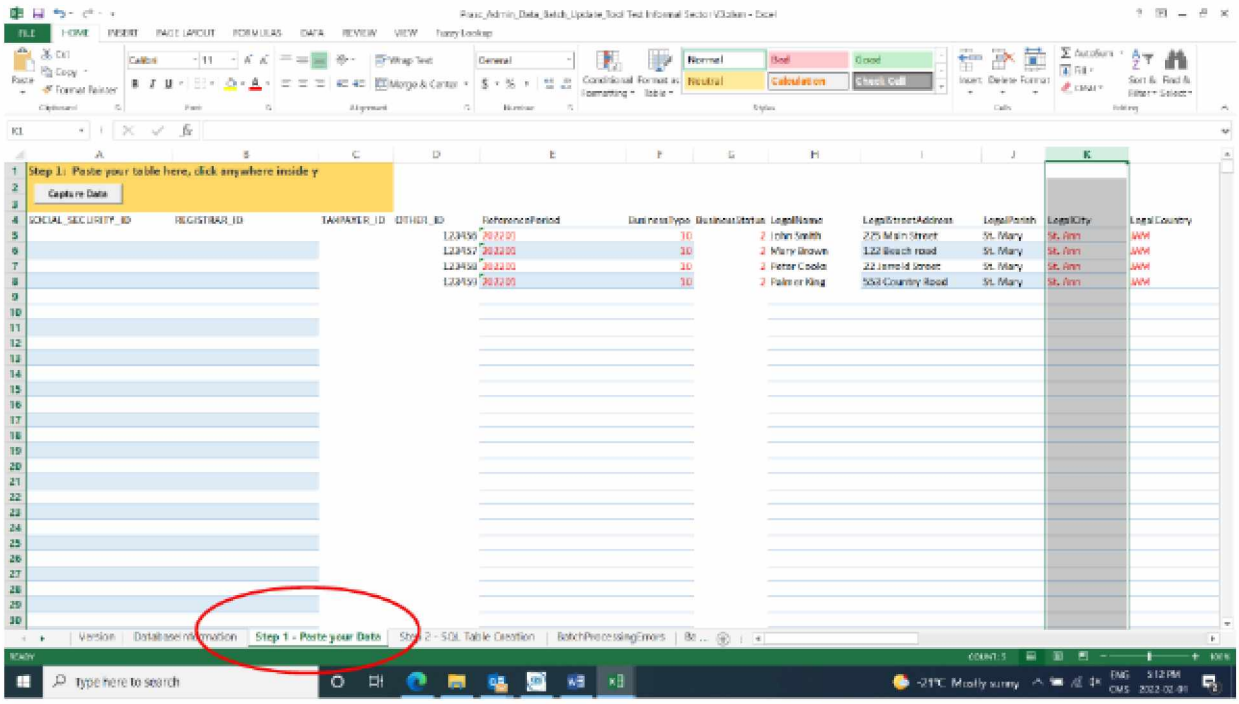


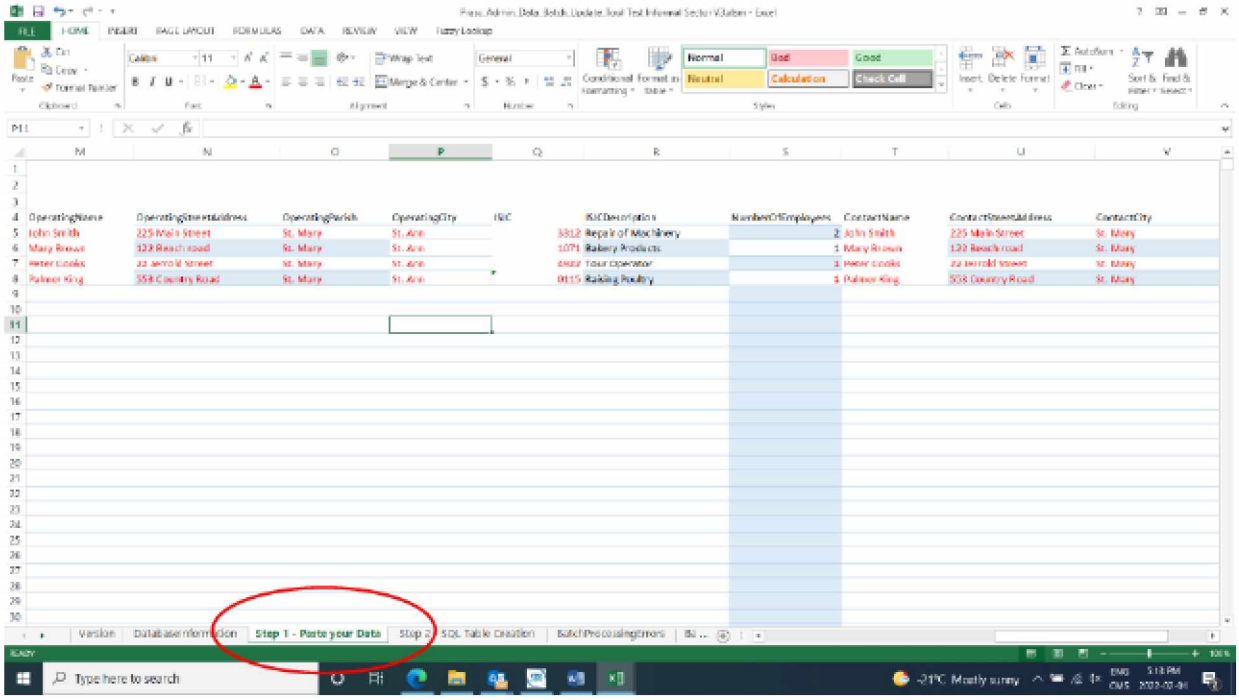


Step 4: From the information in table 1 above, the following SBR variables could be derived based on some basic assumptions:
ReferencePeriod, BusinessType, BusinessStatus, LegalCity, LegalCountry, OperatingName, OperatingAddress, OperatingParish, OperatingCity, ISIC, ConatcName, ContactAddress, ContactCity, ContactParish.

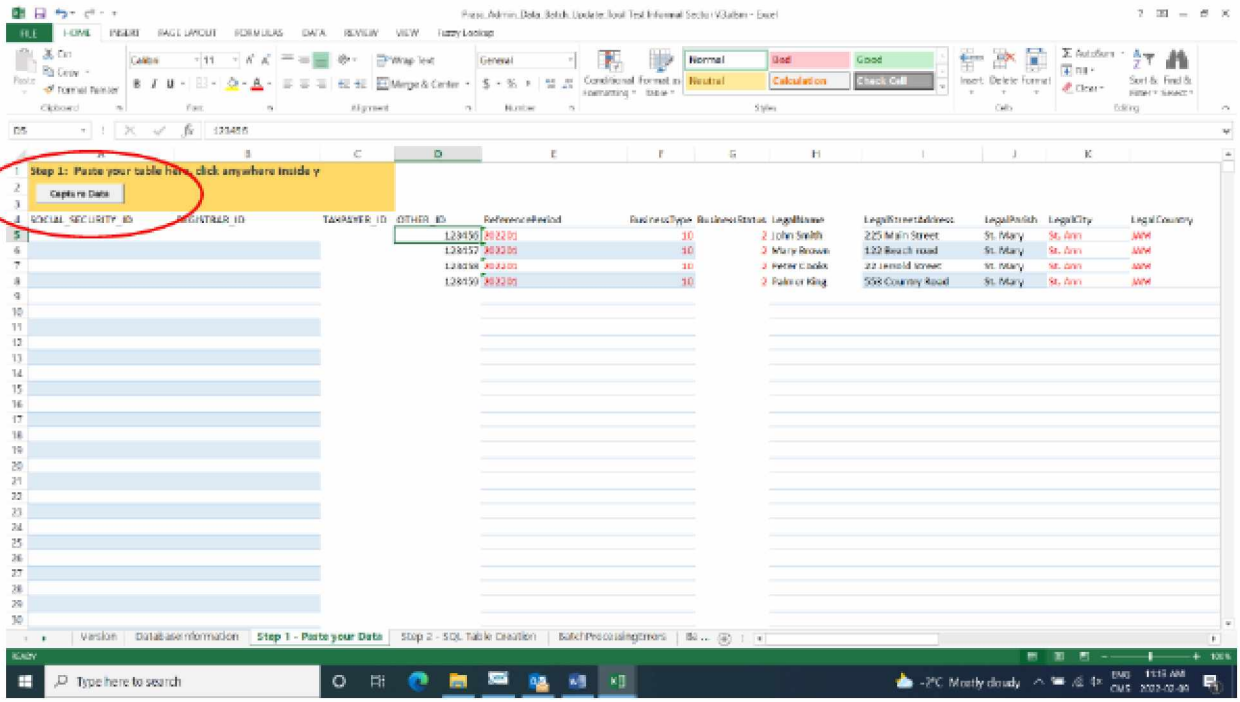
- The reference period is a mandatory field and it should be assigned as the collection period of the census of population. In this example, the **reference period** will be set to 202201, because the VR was done in January 2022.
- For the records with missing number of employees, it is possible to assume that the owner is the only employee. By assigning one (1) employee to the **number of employees**, this will contribute to ensure that the system will automatically set the EnterpriseFlag and EstablishmentFlag to true (1).
- In general, the VR contains a field to collect an activity description and not necessarily the corresponding **ISIC**. In order to load an ISIC on the SBR, it is recommended to code it from the business activity description reported on the VR. This is usually a manual process and it can be resource intensive. This will also contribute to ensure that the system will automatically set the EnterpriseFlag and EstablishmentFlag to true (1).
- Because only informal businesses got selected from the VR, the **business type** should be directly assigned to 10 for informal business.
- For **business status**, it is recommended to assign a business status 2 to indicate that an active business got loaded. This will also contribute to ensure that the system will automatically set the EnterpriseFlag and EstablishmentFlag to true (1).
- In general, for informal businesses, the **operating and contact information** are the same as the legal information of the business. In this scenario, it is assumed that the business information is the same as the business owner's information.

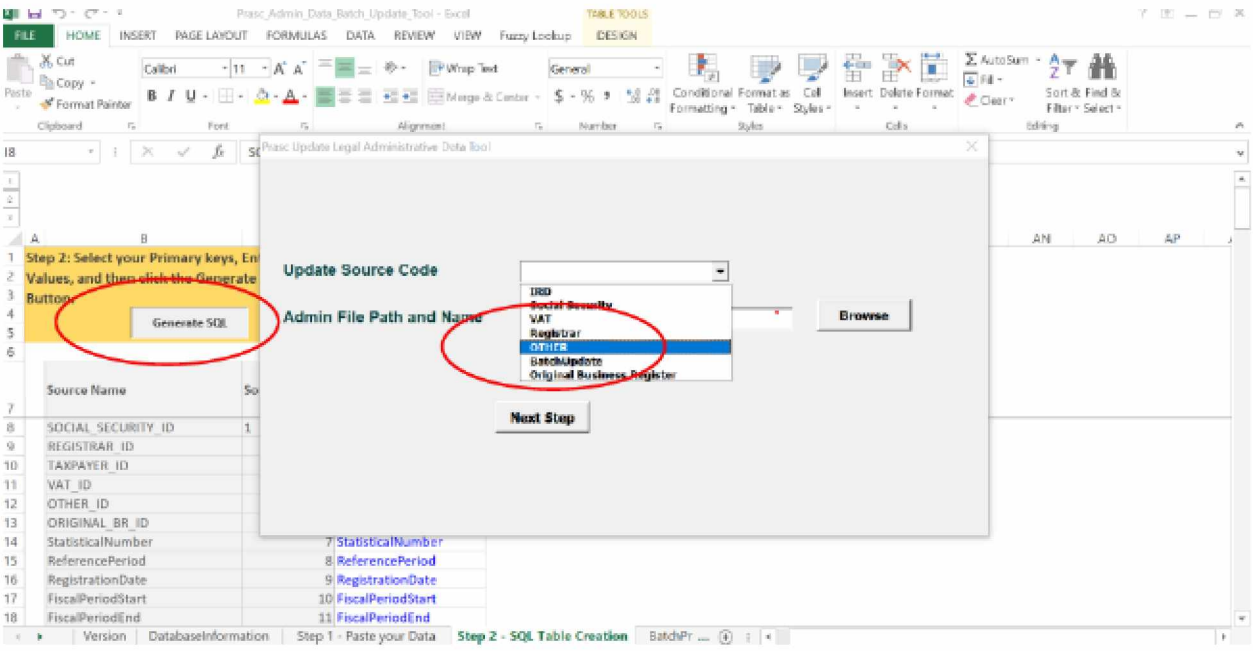
In the screenshot of the Batch Load Excel Template below, the derived field values are shown in red and inserted in the template under '**Paste Your Data**' tab.


Step 1: Paste your table here, click anywhere inside y
Capture Data
SOCIAL_SECURITY_ID REGISTRAR_ID TAXPAYER_ID OTHER_ID ReferencePeriod BusinessType BusinessStatus LegalName LegalStreetAddress LegalParish LegalCity LegalCountry
123456 202201 10 2 John Smith 225 Main Street St. Mary St. Ann JAM
123457 202201 10 2 Mary Brown 122 Beach road St. Mary St. Ann JAM
123458 202201 10 2 Peter Cooke 22 Jerrold Street St. Mary St. Ann JAM
123459 202201 10 2 Palmer King 558 Country Road St. Mary St. Ann JAM
Version DataBaseInformation Step 1 - Paste your Data Step 2 - SQL Table Creation BatchProcessingErrors
Type here to search


OperatingName OperatingStreetAddress OperatingParish OperatingCity ISIC ISICDescription NumberOfEmployees ContactName ContactStreetAddress ContactCity
John Smith 225 Main Street St. Mary St. Ann 3812 Repair of Machinery 2 John Smith 225 Main Street St. Mary
Mary Brown 122 Beach road St. Mary St. Ann 1071 Bakery Products 1 Mary Brown 122 Beach road St. Mary
Peter Cooke 22 Jerrold Street St. Mary St. Ann 4922 Tour Operator 3 Peter Cooke 22 Jerrold Street St. Mary
Palmer King 558 Country Road St. Mary St. Ann 0115 Raising Poultry 1 Palmer King 558 Country Road St. Mary
Version DataBaseInformation Step 1 - Paste your Data Step 2 - SQL Table Creation BatchProcessingErrors
Type here to search

Step 5: In order to load the data onto the SBR, select capture data followed by Generate SQL with the Update Source Code **OTHER**.





Appendix G: ISIC Table Modification Procedures

This section describes the procedures to follow in order to modify the codeset table for new ISIC classification. It is mandatory that Subject Matter and IT Experts collaborate in preparing and executing the transition plan toward the implementation of a new ISIC code on the SBR. This document is intended for subject-matter; IT should refer to the technical guide for the SBR for more technical details.

1. What Subject Matter need to do in preparation for modifying the ISIC code on the SBR.
   - Locate the documentation on new ISIC classification prepared by the United Nations.
   - Locate the concordance file from existing classification to new classification.
   - Analyze the new classification in order to evaluate the impact on the ISIC classification specific code used in your country.
   - Compare the new classification with the existing classification in order to identify the impacted classification.
   - Contact IT experts to bring them in the loop.
   - Based on your detail analysis, prepare an action plan in collaboration with the IT experts.

2. The high-level action plan may contain the following activities and dates
   - The SUF generation and extraction (SM)
   - The creation of the files for the ISIC conversion (SM)
     - New ISIC codes
     - ISIC codes to be removed or replaced
   - Adding the new ISIC code, description, sector description, description code and division code to the ISIC table (IT)
   - The identification of impacted businesses on the SBR data files (SM)
   - Preparation of the three files that will contain the businesses with ISIC to be converted (SM)
     - ***Identification of a detail strategy for each type of conversion (SM) (See Step D subsection V. below)***
   - Applying the ISIC update/conversion to the impacted businesses on the SBR data files (SM and IT). This task may require some time.
   - The load of the businesses with the new ISIC code on the SBR by using the Excel Batch Load template. This load will be done by using the statistical number (Select SurveyCollected(StatisticalNumber) from the drop-down menu as for the Update Source Code) as the unique business identification number (SM)

- The manual updates of the impacted establishments with the new ISIC code (SM)
- The creation of the final codeset table for the new ISIC code. Meaning, when the Subject-Matter is finished batch loading the impacted enterprises and manually updated the impacted establishments on the SBR with the new ISIC code, then IT can proceed to delete the ISIC codes that no longer exist in the new ISIC version. (IT)

3. Steps to follow by subject matter in preparation to transition to the new ISIC code.

   A. Run and Export in Excel a SUF in order to start analyzing your business population and the potential impacts of the new ISIC classification.

   B. In the new ISIC version, there are new codes, merged codes, codes that split and codes that are remaining the same, but their definitions have changed slightly, which might require the codes for some records to be modified. In your SUF file, locate businesses with the impacted ISIC (ISIC that is going to be changed) by filtering the SUF file by ISIC.

   C. Following your analysis of the SUF**, on your SBR data file, if no businesses are identified with an impacted ISIC, there will be no need to convert the SBR business data file.** Simply prepare the following files for the IT Experts to convert the ISIC codeset table.
      i. Prepare a first file containing the new ISIC code, the ISIC description, the sector description, the sector code and the division code.
      E.g.

ISIC Code	ISIC Description	Sector Description	Sector Code	Division Code
011	Growing of cereals (except rice), leguminous crops and oil seeds	Agriculture, forestry and fishing	A	01-03

      ii. Prepare a second file containing the existing ISIC code to be deleted due to the new version.
      iii. The IT expert would use the two files prepared above to update the ISIC codeset table. IT would also need a 3rd list – of any changes to ISIC descriptions (for codes that do not change, but the description or sector/division may change). The IT person will insert the new ISIC code to the codeset table from your first file and delete the ISIC code from the codeset table from your second file.

D. If the SUF contains businesses impacted by the ISIC for conversion, then there will be a need to go through the following detailed procedures for your SBR data file and ISIC codeset table conversion.

i. If modifications have been done to the ISIC variable due to a manual update on the SBR or through a batch load of new data file since your previous extraction in step 2.A, then you will need to run and Export a new SUF in order to prepare your detailed strategy.

ii. From your new extracted SUF, identify all **enterprises (LegalFlag=True on the SUF)** with an impacted ISIC to be converted and create a separate file (Enterprise File) containing these enterprises.

iii. From the same new extracted SUF, identify all **establishments (LegalFlag=False on the SUF)** with an impacted ISIC to be converted and create a separate file (Establishment File) containing these establishments. <u>These units have been manually created on the SBR and the establishments with an impacted ISIC will have to be manually updated on the SBR.</u> **There is no batch process to load establishment units on the SBR.**

iv. For the enterprise file created at step II, identify the three types of changes; **Changed, *Aggregate (merge) and Split sectors.***

v. For the enterprise file created at step IV, we suggest creating 3 additional separate files, one for each type of classification modification. By creating 3 separate files, it will make the modifications more explicit to that action to be taken. Each file should contain at least the unique statistical number generated by the SBR, the old ISIC code and the old activity description. In order to help your analysis, you may want to consider adding variables such as business name, operating name, address and reference period to each record in each file.

vi. Prepare a strategy for each modification.

1. For a **changed sector** such as ISIC xxxx changing to ISIC yyyy, then the action could simply be a rename all of xxxx to yyyy in your Excel file. This scenario could happen due to a need to change to the hierarchy or to the structure of the ISIC classification.
2. For ISIC sectors that will be aggregated such as ISIC code xxxx and ISIC code yyyy merging to ISIC code zzzz, then the action could simply be to rename xxxx and yyyy to zzzz ISIC code. This can happen when there is a reduction of importance in a given sector of activities that are similar.

3. For **an ISIC code split**, this will require more time and effort to resolve these cases. An ISIC split is when a single ISIC code split to 2 or more ISIC codes. As an example, the ISIC code XXXX split to ISIC code YYYY and ZZZZ. This type is a situation that happens when a given ISIC code containing multiple activities becomes so large that a split is recommended or when there is a need to have more detail activities estimates. Below are some strategies on how to split these ISIC codes.
    a. Could use the existing activity description on the SUF file to manually recode to the new ISIC. You could also have your IT expert to write a program to automatically scan the activity description and to recode the ISIC based on keywords found in the activity description. As a fictitious example, in the new classification version, if ISIC=0113 (Growing of vegetables and melons, roots and tubers) split to ISIC= 0117= (Growing of vegetables and melons) and ISIC=0118 (Growing of roots and tubers), then the program can search for businesses that report growing only vegetables and/or melons in their business activity for ISIC 0113 and recode it automatically to ISIC 0117. The same principle can apply to split the code 0113 to code 0118.
    b. If the activity description is not detailed enough, you could search the business on the Internet to identify the exact activity. If the business performs both activities, identify the major business activity.
    c. May need to contact the business to identify the exact business activity.
    d. For large volume of businesses, a NSO could potentially use a percentage distribution from the UN documentation, if available (Eg: 95% in YYYY and 5% in ZZZZ). Mean 95% of the records (enterprises) will receive ISIC code YYYY and 5% of the records (enterprises) will receive ISIC code ZZZZ.

vii. Initialize the variable priority SurveyCollected column for all data variables (especially the ISIC data variable) and to reload the Variable Priority Excel Template on the SBR. This action will permit the update of the ISIC data variable when reloading your updated ISIC code from the Batch Load Excel Template for the impacted enterprises (legal units).

viii. Manually update your impacted establishments. We also recommend creating 3 files for the changed, merged and split for the establishments. These files can guide you to target the establishments to be manually updated on the SBR.

Note to IT experts; it is not possible to just truncate the ISIC codeset table and insert new codes. Even if it were possible, this procedure would cause the data in the Core.tblBusiness to become invalid. The foreign keys are there to ensure data integrity, and so should not be removed. Insert the new codes into the corresponding ISIC codeset table. At this point, it would be impossible to delete the old codes if there is at least one record in Core.tblBusiness referencing that old ISIC code.

Please note that the above procedures can also be applied to the SNA codeset tables conversions.

Also note that the codeset table for Business Status should not be modified, because there are many business rules built-in within the system that drive the behaviour of the SBR application.