



Funded by the  
Government  
of Canada

**Canada**

Mai 2023

# Statistical Business Register Maintenance Guide - Excel

A technical guide to maintaining the Statistical Business Register in Excel

## Contents

<b>1. Introduction.....</b>	<b>3</b>
<b>2. Statistical Business Register Format .....</b>	<b>4</b>
<b>2.1 Protected Workbook .....</b>	<b>4</b>
2.1.1 Locked Cells.....	4
2.1.2 Risks of an unprotected Workbook .....	6
2.1.3 How to unprotect a worksheet: .....	7
2.1.4 How to protect a worksheet: .....	8
<b>2.2 Data Validations .....</b>	<b>9</b>
2.2.1 Data Validations on the Statistical Business Register .....	9
2.2.2 How to create data validations: .....	11
<b>2.3 Conditional Formatting .....</b>	<b>12</b>
2.3.1 Conditional Formatting on the Statistical Business Register .....	12
2.3.2 How to create conditional formatting rules: .....	13
<b>3. Statistical Business Register Maintenance.....</b>	<b>14</b>
<b>3.1 Organising Data .....</b>	<b>14</b>
3.1.1 Administrative Data Update Working Area .....	14
3.1.2 Backup Area .....	14
3.1.3 Snapshot Area .....	14
3.1.4 Work Area .....	15
<b>3.2 Creating Snapshots .....</b>	<b>16</b>
3.2.1 Benefits of snapshots.....	16
3.2.2 Standardised Procedures .....	17
3.2.3 How to create a Snapshot.....	18
<b>3.3 Updating the Statistical Business Register from Administrative Data.....</b>	<b>19</b>
3.3.1 Tools.....	19
3.3.2 Impact of Statistical Business Register values.....	20
3.3.3 Update Procedure.....	20
<b>4. Getting Information from the SBR.....</b>	<b>30</b>

<b>4.1 Extracting Data .....</b>	<b>30</b>
<b>4.2 Storing extracted data .....</b>	<b>30</b>
<b>4.3 Demographic Statistics (Pivot Tables) .....</b>	<b>31</b>
4.3.1 Statistical Business Register's Demographic Statistics .....	31
4.3.2 How to create a pivot table .....	32
<i>List of Acronyms .....</i>	<b>33</b>
<i>Definitions .....</i>	<b>34</b>

## 1. Introduction

Statistics Canada has created two versions of a Statistical Business Register (SBR) as part of the Project on the Regional Advancement of Statistics in the Caribbean (PRASC). One version of the SBR reside on an SQL platform and contains complex business rules, automated processes, and graphic user interfaces. This document refers to the Excel version, which offer just as many features, but provide a much simpler way of maintaining a Statistical Business Register. This document is a technical guide that will explain the format used to build the SBR in Excel, how to maintain the SBR with administrative data and how to use the SBR as a source of information. This document will start by highlighting the key features used to customise the format of the current SBR Excel file. Next, it will provide a step-by-step process for updating the SBR with administrative data. Finally, it will explain how to create snapshots, data extractions, and demographic counts for analysis and reporting purposes. Throughout the document, guidelines and tips will be provided to help users make full use of the SBR using various excel features.

## 2. Statistical Business Register Format

The first section of this document will detail the technical aspects of the current format of the SBR and provide an explanation of why certain decisions were made, as well as how the formatting rules was implemented.

### 2.1 Protected Workbook

In order to maintain the format of the Statistical Business Register and ensure that imbedded business rules are preserved, specific cells have been locked and each Worksheet of the Workbook has been protected. When a Worksheet is protected, locked cells cannot be modified. Locking certain cells prevents business rules from being overwritten and ensures that data validations and conditional formatting is kept intact. Other functions that may interfere with the business rules are also disabled when a worksheet is protected such as adding and deleting rows and sorting the data. Overall, protecting a worksheet prevents human error and contributes to ensuring high quality data on the SBR.

#### 2.1.1 Locked Cells

Some cells have been locked. The locked status of these cells should not be modified. These cells can be classified in two categories: cells that were locked to preserve the format of the SBR and cells that were locked to preserve SBR functionality.

Some of the cells that have been locked in order to preserve the format of the SBR include cells that are used as labels on the SBR such as the title and variable names along the top of the Worksheet. This also includes cells (or entire worksheets) that contain documentation that have been locked to ensure that they are available to users and cannot be inadvertently deleted. Other cells have been locked in order to preserve business rules, ensuring that the Statistical Business Register operates as expected. These cells contain formulas and validation rules that are dependent on other cells or worksheets. Column titles for these cells are highlighted in black. Figure 1 summarises some of the cells that have been locked on the Statistical Business Register and the reasoning for locking them.

Figure 1: Summary of locked cells

Locked Cell	Reason for Locking
<b><i>StatisticalID</i></b>	These cells contain the StatisticalID. The StatisticalID is a random number assigned to SBR Records in order to strengthen data confidentiality, facilitate record linkage and ensure consistency. The StatisticalID is tied to a record indefinitely. This number should never be altered or tampered with and is therefore locked. See Data dictionary for more details.
<b><i>ENT Code</i></b>	These cells contain a formula that automatically assigns values based on other SBR information. In this case, the formula refers to the administrative data IDs (REGISTRAR, TAX, VAT, SOCIAL SECURITY) and the Business Status of the same record. See Data dictionary for more details.
<b><i>EST Code</i></b>	These cells contain a formula that automatically assigns values based on other SBR information. In this case, the formula refers to the ParentID, Legal Code and Business Status of other records in the Statistical Business Register. See Data dictionary for more details.
<b><i>Legal Code</i></b>	These cells contain a formula that automatically assigns values based on other SBR information. In this case, the formula refers to the administrative data IDs (REGISTRAR, TAX, VAT, SOCIAL SECURITY) of the same record. See Data dictionary for more details.
<b><i>Prospect</i></b>	These cells contain a formula that automatically assigns values based on other SBR information. In this case, the formula refers to the administrative data IDs (REGISTRAR, TAX, VAT, SOCIAL SECURITY) and the Business Status of the same record. See Data dictionary for more details.
<b><i>ISIC Description</i></b>	These cells contain the descriptions of International Standard Industrial Classification (ISIC) codes. A validation rule has been implemented to link the correct ISIC description to its corresponding ISIC code, ensuring coherence between the variables.

### 2.1.2 Risks of an unprotected Workbook

While the Statistical Business Register workbook should be protected the vast majority of the time, there are specific circumstances when a representative from the Statistical Office (SO) might be required to unprotect the workbook.

It is therefore important to highlight the risk involved of manipulating the Worksheet while unprotected:

1. Because formulas are populated throughout multiple columns, inadvertently overwriting the content of a cell containing such a formula will prevent business rules from being applied properly within the entire row and potentially affect other variables such as business structure, survey eligibility, business status, etc.
2. Since the worksheet contains advanced formatting settings, adding or removing a row or column (or modifying the format of the sheet in any other way) might prevent certain functions such as sorting and filtering from working properly.
3. Since business rules have been implemented and might refer to other worksheets in the same workbook, any modifications to the worksheets might result in errors in cells where the business rule resides.
4. Since validation rules were implemented and refer to values in another sheet, tampering with either the source worksheet or the cells containing validations might prevent validation rules from being implemented and might cause some unexpected results.

Knowing these risks, the Statistical Office staff can proceed to unprotect the Statistical Business Register in order to make adjustments such as sorting data or updating the SBR through Administrative Data.

### 2.1.3 How to unprotect a worksheet:

The steps to unprotect the SBR are listed below. Figure 2 illustrates how to remove the protection from the SBR worksheet.

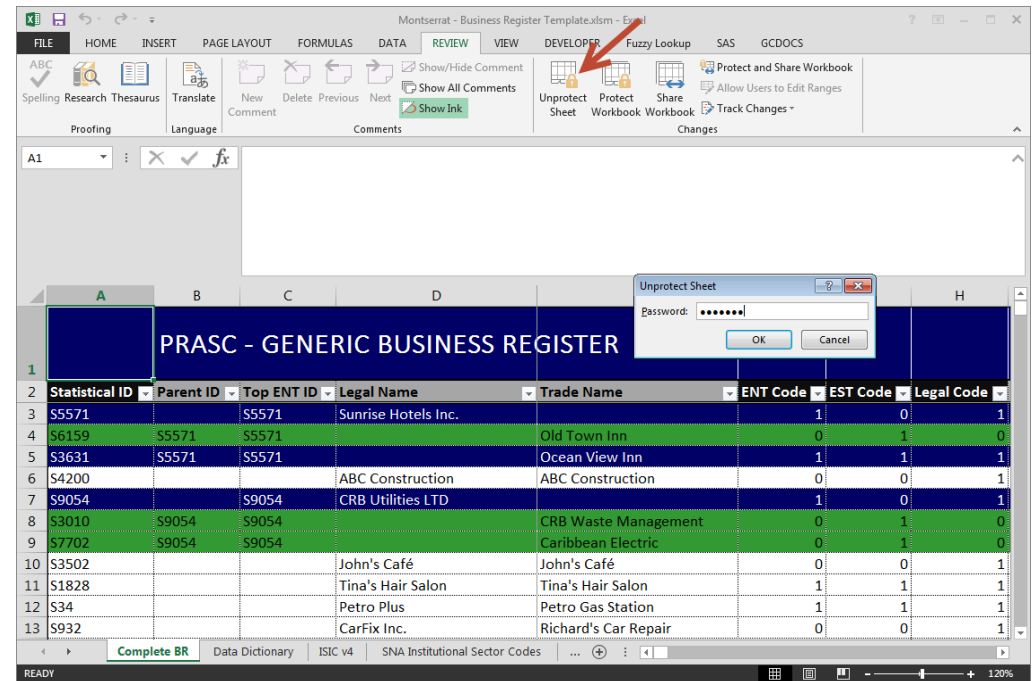
**Warning:** Working on the SBR while it is unprotected can have many adverse consequences as described above. For that reason, a password, known only by select NSO staff, should be assigned to unprotect the worksheets. Before removing the protection on a worksheet, the staff should always create a backup of the file following the procedures outlined in this document. The worksheets should be protected again as soon as possible. **The Complete SBR worksheet should only be unprotected when no one else is working on the file in order to prevent conflicts.**

**Figure 2: Removing the protection on the SBR**

#### Steps:

1. Create a backup of the file
2. Select the Review tab
3. Click on Unprotect Sheet
4. Enter Password (**statcan**)

\*Remember to protect the worksheet again as soon as possible to prevent the Statistical Business Register business rules and validations from being overwritten and potentially corrupting the data.



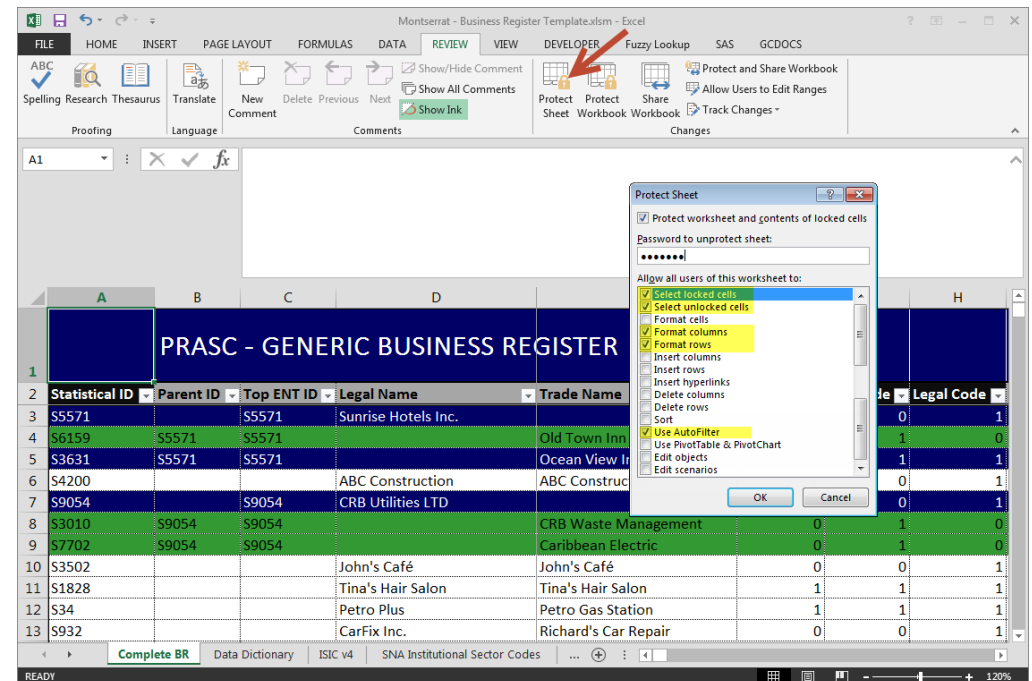
### 2.1.4 How to protect a worksheet:

The steps to protect the SBR are listed below. Figure 3 illustrates how to protect the SBR. The options that must be selected are highlighted in yellow.

**Figure 3: Protecting the SBR**

#### Steps:

1. Select the Review tab
2. Click on Protect Sheet
3. From the Protect Sheet pop up window
  - a. Select the following options under “Allow all users of this worksheet to”:
    - Select locked cells
    - Select unlocked cells
    - Format columns
    - Format rows
    - Use AutoFilters
  - b. Enter a password (**statcan**)
  - c. Re-enter password (**statcan**)



#### Note:

Since the worksheet contains a combination of locked and unlocked cells and the sort function conflicts with this format, the “sort” option should not be selected when protecting the Complete SBR worksheet. Allowing the sort function when the worksheet is locked causes Excel to stop responding and all unsaved work will be lost.

## 2.2 Data Validations

Data validation is a process that ensures the entry of clean and clear data to the programs, applications or services using it. It checks for the integrity and validity of data that is being inputted to different software and its components. For example, it can prevent typographical errors by rejecting invalid values, and ensure that data dependencies are respected.

### 2.2.1 Data Validations on the Statistical Business Register

Data validation rules have been imbedded in the Statistical Business Register Excel Workbook. Since some fields in the Statistical Business Register have defined code sets or a set of predetermined values, implementing data validation rules contributes in increasing the quality of the data being keyed in. Figure 4 summarises the fields on the SBR where data validations were implemented.

**Figure 4: Data Validations on the SBR.**

<b>Variables</b>	<b>Data Validation Rules</b>
<b><i>ParentID</i></b>	Data validation was implemented to ensure that a valid StatisticalID from the StatisticalID column was keyed in. Any other values will be rejected.
<b><i>Top ENT ID</i></b>	Data validation was implemented to ensure that a valid StatisticalID from the StatisticalID column was keyed in. Any other values will be rejected.
<b><i>Legal Code Overwrite</i></b>	Data validation was implemented to ensure that a valid Legal Code was keyed in. Any other values will be rejected. See Data dictionary for more details.
<b><i>Business Status Code</i></b>	Data validation was implemented to ensure that a valid Business Status Code based on a pre-defined code-set was keyed in. Any other values will be rejected. See Data dictionary for more details.
<b><i>Legal Type Code</i></b>	Data validation was implemented to ensure that a valid BR Legal Code based on a pre-defined code-set was keyed in. Any other values will be rejected. See Data dictionary in the Excel SBR for more details.
<b><i>ISIC</i></b>	Data validation was implemented to ensure that a valid ISIC Code (rev. 4) was keyed in. Any other values will be rejected. See Data dictionary for more details

*Note:*

Data Validation source references are contained within a hidden worksheet named “Drop Down List”. To view this worksheet, right click on a worksheet tab and select “Unhide”. It is recommended that the worksheet stay hidden because it adds unnecessary clutter to the look of the SBR.

## 2.2.2 How to create data validations:

The steps to create data validations are listed below. Figure 5 illustrates how data validations are created. The options that must be selected are highlighted in yellow.

### Steps:


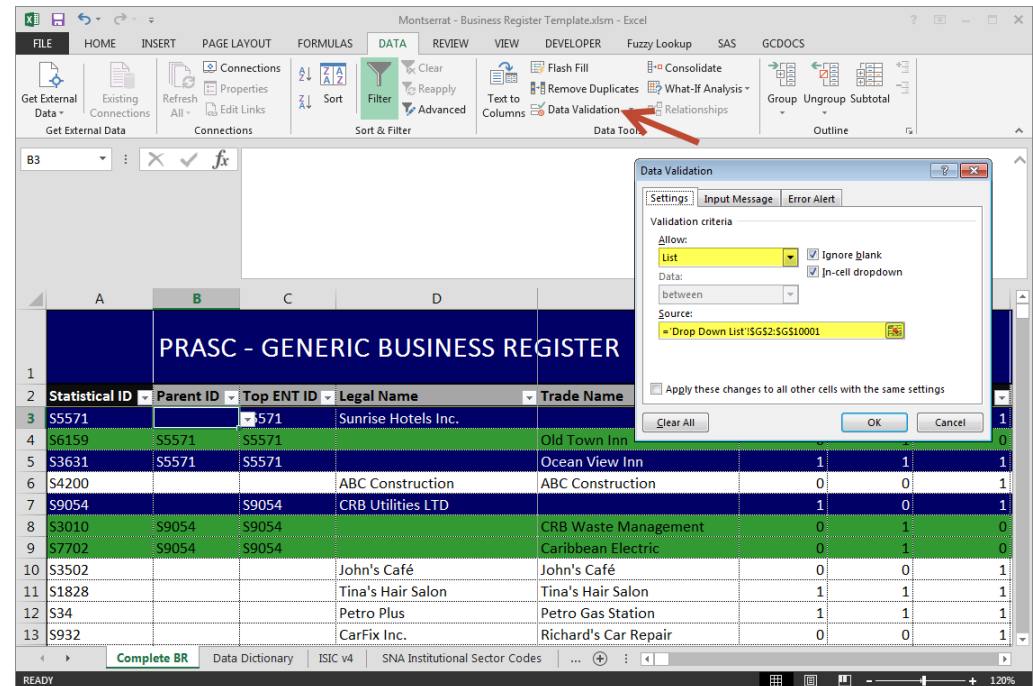
1. Create a list of valid values. In this case, the list of valid values are entered on a hidden worksheet named "Drop Down List"
2. Select all cells on which to apply the data validation on
3. From the Data tab, select Data Validation from the Data Tools section menu
4. Select Data Validation
5. Select the appropriate methodology to create rules. In this case, select "list" from the drill down menu under "Allow:"
6. Under "Source": click the icon  at the right of the text box.
7. Select the set of valid values and press "Enter". On the SBR, the lists of valid values for multiple data validation rules are all entered on a worksheet named "Drop Down List"
8. Click Ok to accept the creation of the new data validation rule.

Figure 5: Creating Data Validations



## 2.3 Conditional Formatting

Conditional formatting is a function that allows the formatting of specific cells, columns or rows when a specific condition is met.

### 2.3.1 Conditional Formatting on the Statistical Business Register

In order to facilitate the identification of certain events or states, conditional formatting has been imbedded into the SBR. On the SBR, conditional formatting was used to identify enterprises and establishments within complex structures. Figure 6 details these conditional formatting rules.

**Figure 6: Conditional Formatting rule**

<b>Variables</b>	<b>Conditional Formatting Rules</b>
<b><i>Enterprises</i></b>	<p><b>Enterprises in a complex structure</b> are identified when the ENT code =1 and that they have been identified as being part of a complex structure (Top ENT ID not empty). When this condition is met, the entire row is highlighted in blue and the text is set to white.</p> <p>Enterprise Formula: =AND(\$D3=1,\$C3&lt;&gt;0)</p>
<b><i>Establishments</i></b>	<p><b>Establishments in a complex structure</b> are identified when the EST code = 1 and that they have been identified as being part of a complex structure (Top ENT ID not empty). When this condition is met, the entire row is highlighted in green.</p> <p>Establishment Formula: =AND(\$D3=0,\$E3=1,\$B3&lt;&gt;0)</p>

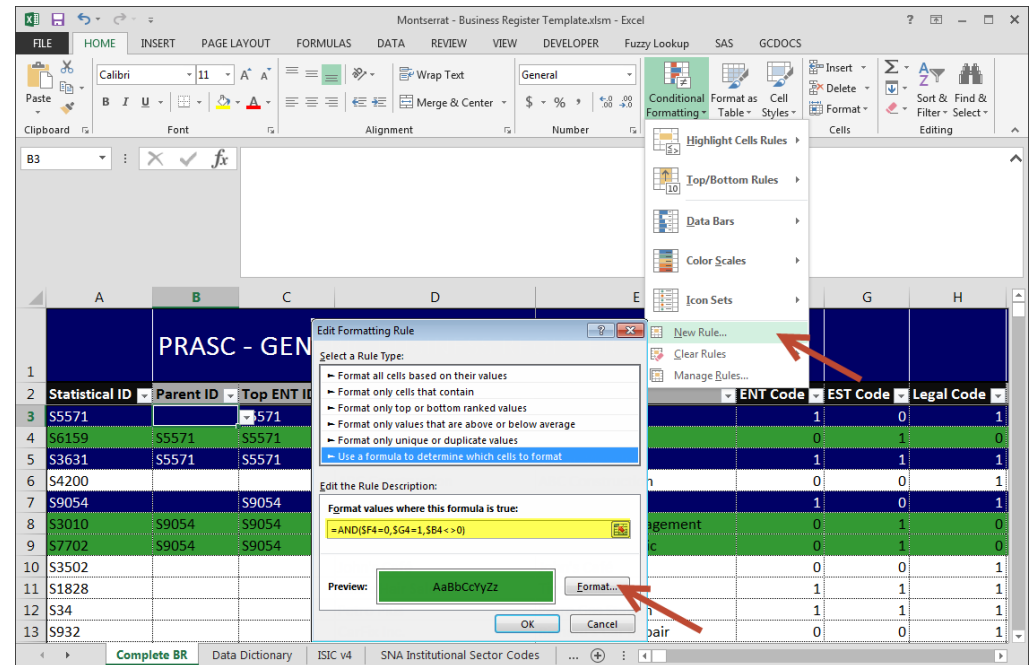
### 2.3.2 How to create conditional formatting rules:

The steps to create conditional formatting rules are listed below. Figure 7 is an image demonstrating how to access and how to complete the dialog box used to create conditional formatting.

#### Steps:

1. Select all cells on which to apply the conditional formatting on
2. From the Home tab, select Conditional Formatting from the Styles section menu
3. Select New Rules
4. Select the appropriate methodology to create rules. In this case, select "Use a formula to determine which cells to format."
5. Under "Format values where this formula is true:", enter desired formula
6. Click on "Format...", select the desired formatting from the "Format cells" options. In this case, select the "Fill" tab and select the appropriate color. To format the font of the text, select the "Font" tab and select a color (white for Enterprise condition).
7. Click Ok to close the "Format cells" pop up window.
8. Click Ok to close the "New formatting rule" pop up window.

Figure 7: Creating Conditional Formatting Rules



## 3. Statistical Business Register Maintenance

### 3.1 Organising Data

In order to have a proper file management system and work area to perform Statistical Business Register data manipulation procedures, it is recommended that the Statistical Office elaborate a standardised folder structure. A consistent and stable working environment will prevent data loss and will ensure that files can be accessed easily. It is therefore recommended that the files be grouped under the following file categories: Administrative Data Updates, Backups, Snapshots, and Working Area.

#### 3.1.1 Administrative Data Update Working Area

The administrative data update folder will contain files used to update the SBR with administrative data. The current process for updating the SBR with administrative data involves some manual steps that may require several days of work to complete. It is important to create a separate working environment for these files so that no other transactions are applied to the work in progress. The original administrative data files should also be kept in this folder (or a subfolder) and should be protected by setting their properties to Read-Only in order to prevent them from being overwritten. This can simply consist of creating designated folders, with names that are representative of their purpose.

#### 3.1.2 Backup Area

The Statistical Business Register contains complex formatting, business rules and data validation functions, it is therefore important to have an area to store Backup copies of the SBR in order to prevent data loss or corruption. This will be most important when updating the Statistical Business Register with Administrative Data, but having a regular backup plan is a recommended practice for any work activities involving data. Backup copies should be created at frequent intervals and before any major data manipulation process. This can simply consist of creating a designated folder to that effect and establishing a backup procedure and schedule. The folder name must be representative of its purpose. Out-of-date backups can be removed from the folder when they are no longer useful.

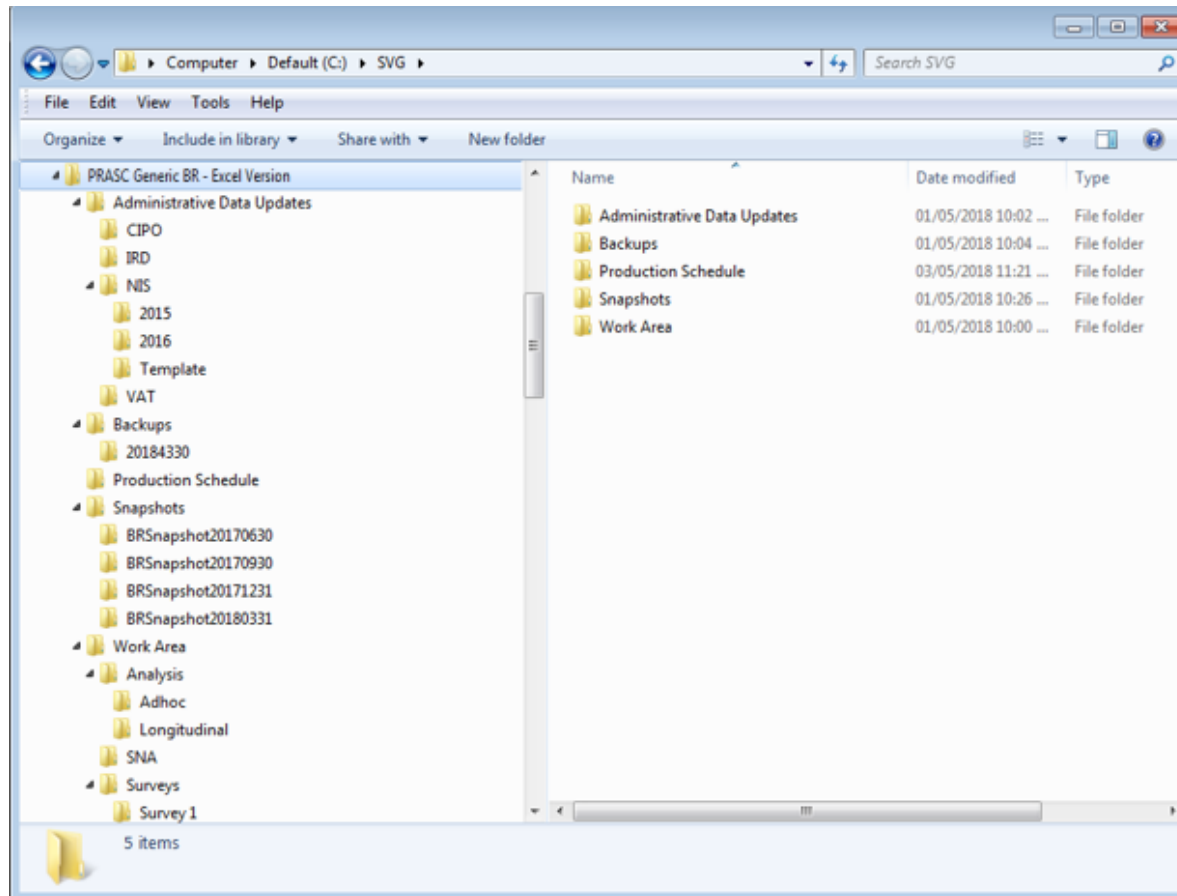
#### 3.1.3 Snapshot Area

It is also recommended that the Statistical Office create Snapshots of the Statistical Business Register on a regular basis and allocate a designated area to store these files. Snapshots should be taken using the same steps as backup files, but serve a different purpose. Snapshots are used as a historical log of what was on the SBR at a given time. The Snapshots should only be used as references and never be altered. For this reason, they should be stored in a different location than the current, working copy of the SBR. It is recommended that a designated Snapshot folder is created, and that a regular schedule for creating Snapshots is established. Snapshots should be stored indefinitely. For more details on the Snapshot creation, refer to the Creating Snapshots section of this document.

### 3.1.4 Work Area

Finally, since the Statistical Office staff needs an area to work on the SBR to research information, perform updates or data analysis or extract data, it is recommended that a work area is created to ensure that regular tasks performed on the current Statistical Business Register are not conflicting with the backup or administrative data load procedures and that snapshots are not accidentally modified. This can simply consist of creating a designated folder to that effect. The folder name must be representative of its purpose. Figure 8 illustrate an example of a recommended organised folder structure.

**Figure 8: Recommended Folder Structure**



### 3.2 Creating Snapshots

In order to document the continuity of the Statistical Business Register and be able to provide a picture of the economy at different times, it is recommended that Snapshots of the Statistical Business Register be created and kept in a secure location. The creation of Snapshots should simply involve taking a copy of the SBR at a specific date and time. These are copies which should never be altered thereafter. Snapshots should be completely locked and users should only be working from them in READ ONLY mode. Please refer to section 3.2.3 How to Create a Snapshot for information on how to set the Read-Only attribute on a file.

#### 3.2.1 Benefits of snapshots

Snapshots can be used for conducting longitudinal analysis, as well as performing quality assurance by comparing values to previous snapshots. They can also be used for survey sampling and act as an historical database from which reports can be generated. They are essential in ensuring data continuity and consistency and serve as a static reference point in time from which data can be analysed and published in a coherent manner. Figure 9 summarises the benefits of the creation of snapshots.

**Figure 9: Summary of Snapshot Benefits**

<b>Benefits</b>	<b>Description</b>
<b><i>Survey Sampling</i></b>	A SBR Snapshot provides the Statistical Office with a file that can be used for the design and development of survey processes. The SBR snapshot can be used to create survey sample files.
<b><i>Longitudinal Analysis</i></b>	SBR Snapshots can be compared from reference period to reference period to evaluate the economic trends in a country. By keeping consistent data frozen in time at particular reference periods, patterns can be monitored and economic trends can be identified. The information gathered can then help in the production of business demography statistics.
<b><i>Quality Assurance &amp; Reporting</i></b>	Users can query SBR snapshots, which is useful for writing reports. Since Snapshots are static and represent a point in time, snapshots of two different reference periods can be compared to identify data variation and discrepancies.
<b><i>Historical Data</i></b>	SBR snapshots can serve as references of particular points in time. These Snapshots can be archived and kept indefinitely and be used as reference data from which historical analysis or reports can be produced.

### 3.2.2 Standardised Procedures

For consistency purposes, it is important to establish a standardised pattern in the creation of Snapshots. Since snapshots will be used as references and represent the business population at a point in time, it is recommended that the Statistical Office establish procedures for its creation such as the implementation of a production plan and schedule to monitor data manipulation transactions and facilitate quality assurance. The adoption of a naming convention is also essential to ensure proper file management.

#### Production Schedule:

In order to establish a consistent methodology for maintaining and updating the Statistical Business Register, it is recommended that a production schedule be defined and documented to accompany the Statistical Business Register Maintenance Guide. Figure 10 displays an example of a Production Schedule.

The following are criterion that should be considered when establishing a Production schedule:

- 1. Person(s) responsible for the maintenance of the SBR
- 2. Frequency of Snapshots
- 3. Frequency of Backups
- 4. Frequency of administrative data updates
- 5. Server Maintenance
- 6. Dates of key deliverables

#### Naming Convention:

Since snapshots refer to specific point in time, it is important to differentiate each snapshot with a meaningful naming convention. A naming convention should therefore be composed of information that can clearly identify the source of the data and its reference period.

- Example: BusinessRegisterSnapshotyyyymmdd

**Figure 10: Production Schedule Example**

March 2018						
Black = Scheduled <i>Blue/Italic</i> = Executed    Red = Cancelled    Yellow = Tentative						
Server Maintenance						
Sun	Mon	Tue	Wed	Thu	Fri	Sat
				1	2	3
4	5	6	7 <i>BR backup</i> <i>Administrative Data Update (NIS)</i>	8 <i>Administrative Data Update (NIS)</i>	9	10
11	12	13	14 National Hero's Day	15	16 <i>BR Backup</i>	17
18	19 <i>Server Maintenance</i>	20	21	22	23	24
25	26	27 Administrative Data Update (IRD)	28	29	30 BR Snapshot	31

### 3.2.3 How to create a Snapshot

The steps to create Snapshots are listed below. Figure 11 illustrates how SBR Snapshots are created.

#### Steps:


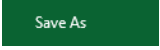
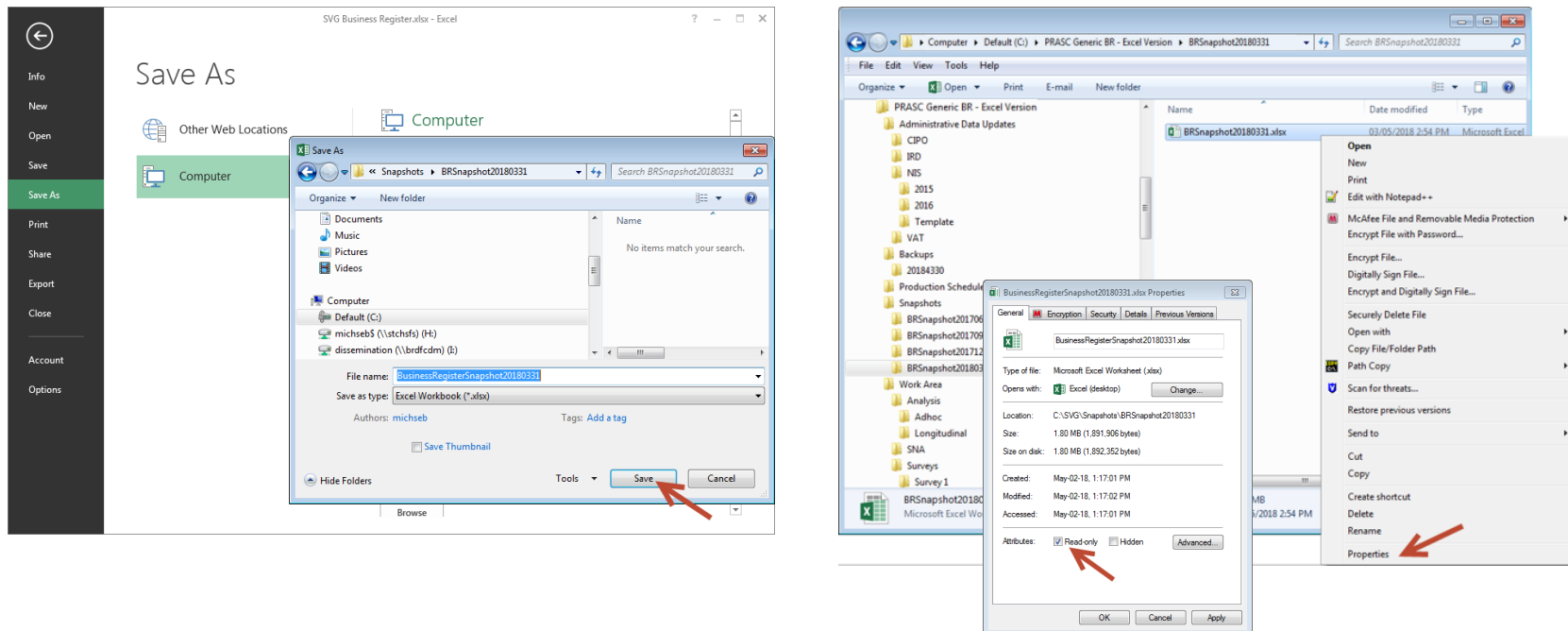
1. Open the Statistical Business Register Worksheet
2. In the top left corner on the Excel menu, click on File 
3. Select Save As 
4. Click on SBRrowse
5. Select the appropriate location for Snapshots
6. Enter the name of the Snapshot based on the defined naming convention, click Save
7. Using Windows explorer, locate file.
8. Select file, right-click and select Properties
9. Under the General tab of the **Properties** pop up window, put a checkmark beside **Read-only** in the Attributes section.

Figure 11: Creating Snapshots



### 3.3 Updating the Statistical Business Register from Administrative Data

Valuable information can be retrieved and processed from administrative data files to create, update and maintain a robust SBR.

Going forward, the Statistical Office will acquire files from their administrative partners at regular intervals. Upon receiving these files, the appointed production manager will be required to update the existing SBR with the new information found in the files. To facilitate this process, a record linkage process using an Excel add-on called Fuzzy lookup can be used. There is also a possibility of creating Excel macros to automate certain tasks of the record linkage process, but their need and implementation varies. For smaller population, a manual process could be simpler to implement, making it easier to maintain the Statistical Business Register, as long as the record linkage strategy and process is clearly defined and understood. If macros are required, the Statistical Office should contact their PRASC representative.

Since the Statistical Business Register is a moving target and that the intended purpose of administrative data does not necessarily align with that of the Statistical Office, manual review and intervention will always be required when updating the Statistical Business Register with administrative data.

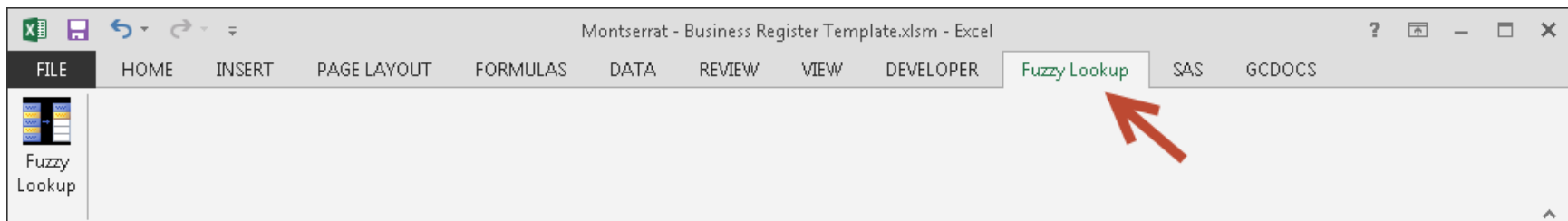
#### 3.3.1 Tools

##### 3.3.1.1 Backup

Before updating the SBR with administrative data, it is of utmost importance that a backup of the file be created. Since several automatic and manual interventions will be made to the file, a backup copy is necessary in order to undo certain changes or compare data for quality assurance analysis. Having a backup copy of the SBR before it is manipulated also ensures that the file can be restored in case of file corruption or data loss.

##### 3.3.1.2 Fuzzy Lookup

In order to link administrative data records to the Statistical Business Register business entities, record linkage must be applied. While multiple record linkage tools exist, since the SBR is in an Excel format, it is recommended to use the Fuzzy Lookup Excel add-on to do the record linkage. Before processing to the Administrative Data Update portion, ensure that the Fuzzy lookup Excel add-on has been installed. Please consult your Information Technology support on how to install the Fuzzy Lookup add-on.



### 3.3.2 Impact of Statistical Business Register values

There are different levels of impact when applying updates to the Statistical Business Register. While low impact updates can be applied on the SBR without extensive verifications, updates with high impact will need to be validated by the Statistical Office staff. For example, adding information to an empty field might constitute a low impact update while overwriting information might warrant validation from a Statistical Office staff member. Figure 12 illustrates the impact level of values to be updated.

**Figure 12: Impact of Record Update**

High Impact Values	Low Impact Values
<ul style="list-style-type: none"> <li>• Legal Name</li> <li>• Trade Name (if not empty)</li> <li>• Address (if not empty)</li> <li>• Social Security Status</li> <li>• Registration Date (if not empty)</li> <li>• Social Security Closure Date</li> <li>• ISIC</li> </ul>	<ul style="list-style-type: none"> <li>• Trade Name (if empty)</li> <li>• Telephone</li> <li>• Address (if empty)</li> <li>• Economic Sector</li> <li>• Number of Employees</li> <li>• Gross Wages</li> <li>• Social Security Registration Date (if empty)</li> </ul>

### 3.3.3 Update Procedure

As discussed, there are multiple Administrative Data sources that the Statistical Office will use to maintain their Statistical Business Register. Since Registrar, TAX and Social Security all maintain different files with slightly different information and formatting, all administrative data will need to be processed differently. A specific Administrative Data procedure should be established independently for each source. For that reason, the following procedures will present a generic way of updating the Statistical Business Register using administrative data.

On the current SBR, any known administrative data identifiers are recorded under the Registrar, TAX and Social Security variables, which provides a concordance of each administrative data identifiers and the Statistical ID. It is possible to match the new administrative data source with the Statistical Business Register using those identifiers. Once the data has been linked, it will be necessary to verify the suggested changes to the high impact values before applying them to the SBR. Low impact values can be automatically updated.

Once the updates have been applied, the Statistical Office can address the unmatched records. Not all records on the Statistical Business Register have a corresponding administrative data ID; some businesses on the SBR might not have been previously linked to an administrative record or might have been linked to a different administrative data source than the one we are loading. For this reason, the Statistical Office should use various record linkage techniques to link the unmatched administrative data records to the SBR. This record linkage process will use

alternate variables such as the Business Name, Employer Name, and any other identifying variables to determine a match. Once matches have been identified, they need to be manually verified by the Statistical Office staff to ensure that the matches are accurate and that complex business structures are identified and adjusted if need be. To perform manual updates, please refer to the Procedures for Updates section of the Statistical Business Register User Guide.

Once potential record linkage matches have been manually updated, the Statistical Office needs to analyse the remaining records; those that couldn't be matched on administrative data ID or through record linkage and are presumed to be new records. Before adding any record, it is extremely important that the Statistical Office staff search through the existing SBR records to ensure that the missing administrative data record doesn't already exist on the SBR. Adding a record that is already on the SBR would create duplication which has a negative impact on the quality of the SBR as a statistical frame, as well as its reliability for producing demographic statistics. When a duplicate is found, it will be necessary to adjust the record on the Statistical Business Register accordingly. If no duplicates are found, the updater can feel assured that a duplication will not be created by adding the new administrative data record and can therefore proceed to add it to the SBR. For more details on how to address duplicates and add new records, please refer to the Procedures for Update section of Statistical Business Register User Guide. The following section of this document provides the step-by-step procedures for updating the SBR with administrative data files using Fuzzy Lookup.

**Warning:** To prevent the introduction of duplicates on the SBR and potentially create over coverage, it is important to always search through the SBR before adding an unmatched record as a new business. Special care should be taken to that effect.

#### Pre-Processing:

The steps to prepare the file before proceeding with the administrative data update process are listed below.

##### *Steps:*

1. Create a backup of the current SBR file by following the steps for "Creating a Snapshot".
2. Move backup copy of the SBR into the designated Backup folder.
3. "Protect" current SBR so that no changes can be made by other updaters during the administrative data update.
4. Create a 2<sup>nd</sup> copy of the Statistical Business Register.
5. Move 2<sup>nd</sup> copy into a secure and designated folder (Administrative Data Updates) and rename to "Admin Data Load **source**yyyyymmdd" where **source** represent the source of the data being loaded and yyyyymmdd represent the reference date of the Administrative Data source.

Once the administrative data has been added to the “Admin Data Load *source*yyyyymmdd” file, the process can begin. Users can link the information with the assistance of Fuzzy Lookup and can perform the updates manually

#### Updating the Statistical Business Register Using Fuzzy Lookup

**Warning:** Before updating the Statistical Business Register with new administrative data, a backup copy of the SBR excel file should always be created and stored in a separate and secure location. Backup copies of the SBR serve as a safety measure for all unforeseen events while manipulating the data and can be used to retrieve data that has been corrupted or lost. Also, the Administrative Data load process should be done on a second copy of the SBR so that it doesn't hinder the work of other SO staff. Note that while the current SBR can still be accessed by the SO staff, it should not be modified until the Administrative Data update process is completed and a new version of the SBR is made available.

The steps to update the Statistical Business Register with administrative data using Fuzzy Lookup are listed below. Throughout the procedure, ***Admin*** (in bold and italic) refers to the Administrative Data Source acronym (Registrar, TAX, VAT, Social Security) being used.

*Steps:*

1. Creating tables in preparation for record linkage using Fuzzy Lookup
  - 1.1. From the Complete SBR tab, select all variables that will be required for record linkage and validations.
    - 1.1.1. Administrative ID
    - 1.1.2. Names
    - 1.1.3. Addresses
    - 1.1.4. Etc.
  - 1.2. Copy the data using the CTRL-C command.
  - 1.3. Select a new worksheet, rename it Table1\_SBR and select the top-left cell and paste the values using the CTRL-V command
  - 1.4. Select all values again and press CTRL-L to create a table and name it Table1\_SBR
  - 1.5. Repeat the process to create a table for the new Administrative Data, renaming the tab and table to Table2\_Admin
  - 1.6. Select or add a new Worksheet, rename it to ***Admin***IDMatch, and select the top-left cell.

Figure 13: Creating Tables

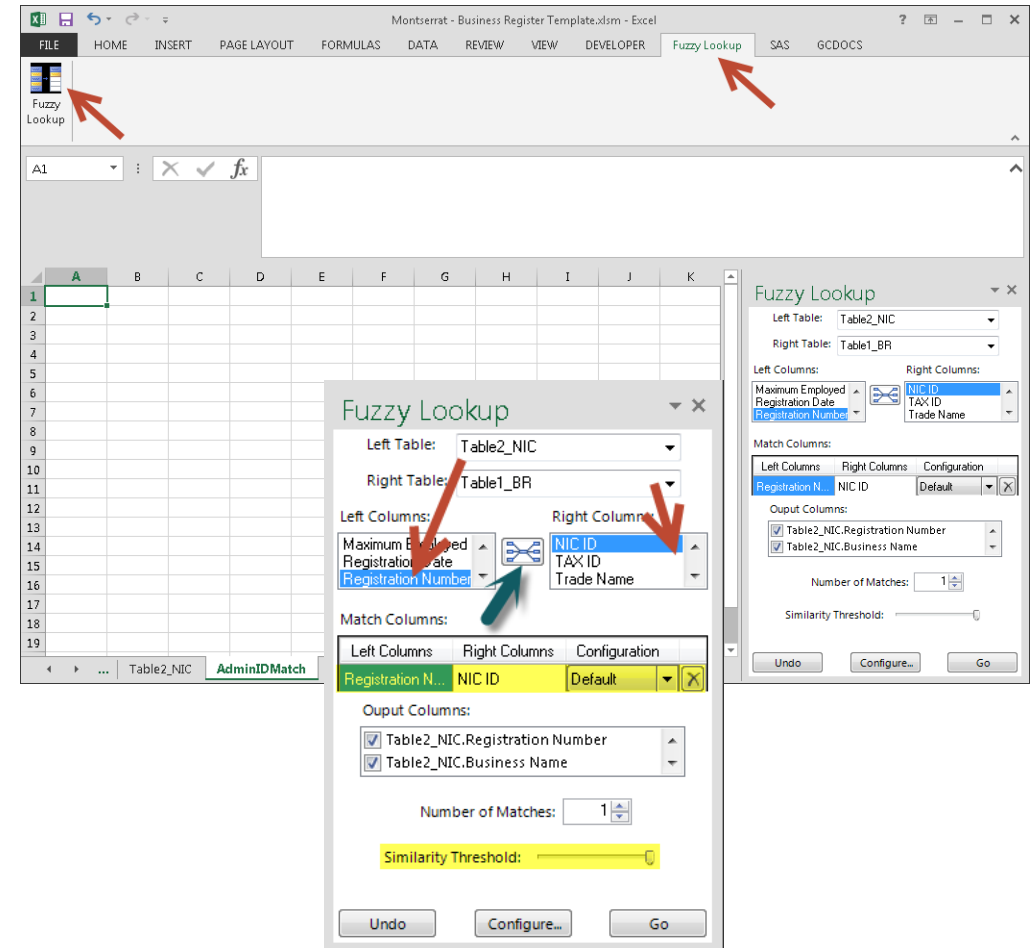
Legal Name	Trade Name	ENT Code	EST Code	Legal Code	Legal Code Overwrite	CIPO ID	TAX ID	VAT ID	NIC ID
Sunrise Hotels Inc.		0	0	1			1234		
	Old Town Inn	0	0	0					
	Ocean View Inn	0	0	1				987777	
ABC Constructi	ABC Constructi	0	0	1			5555		
CRB Utilities LTD		0	0	1		6789			
	CRB Waste Mar	0	0	0					
	Caribbean Elect	0	0	0					
John's Café	John's Café	0	0	1					121212
Tina's Hair Salo	Tina's Hair Salo	0	0	1					567899
Petro Plus	Petro Gas Statio	0	0	1			4322		545432
CarFix Inc.	Richard's Car Re	0	0	1			7876		
Crystal Mart In	Crystal Mart	0	0	1			9821		918273
123 Robert	Robert's Farm	0	0	0					
Comp-U-Tech	Comp-U-Tech	0	0	1		100012			
Lisa's Fashion Wear		0	0	0					
	Emerald Shoes	0	0	0					
Church of CRB	Church of CRB	0	0	1			222244		
	CRB Church	0	0	0					

2. Compare new Administrative Data information with Complete SBR on the Administrative Data ID
  - 2.1. From the Fuzzy Lookup tab, on the **AdminIDMatch** blank worksheet, Click on Fuzzy lookup. A Window will pop up on the bottom right.
  - 2.2. Identify variables on which you wish to perform record linkage on each table
    - 2.2.1. Left Columns: Registration Number from Table2\_Admin
    - 2.2.2. Right Column: **AdminID** from Table1\_SBR
  - 2.3. Click on the Join button (green arrow on Figure 14) to add the Fuzzy Join bonding.
  - 2.4. Adjust the Similarity Threshold to 1\*
  - 2.5. Press Go.

\*The similarity Threshold is a number between 0 and 1 that indicates the level of probability of a match. Since Administrative Data ID are unique and that we don't want to have partial matches, the threshold should be set to 1.

Updates and validations: As previously explained, some updates have a different level of impact. Low impact values should be applied on the SBR tab while high impact values should be verified and validated by the Statistical Office. Once validated, new values coming from administrative data source should manually be applied to the SBR.

**Figure 14: Record Linkage on Administrative ID using Fuzzy Lookup**



**NOTE:** Not all suggested updates to the Statistical Business Register from an administrative data source need to be applied. In some cases, a suggested update might not provide the most recent or accurate information. Statistical Office staff are therefore responsible for verifying and accepting or rejecting the suggested modifications.

### 3. Validating Administrative ID matches and updating values

- 3.1. Once the Fuzzy Lookup process is completed, the results will appear on the blank worksheet that was previously named **AdminIDMatch**.
- 3.2. On the far right, locate the Similarity column, select it and sort it in descending order (Sort Largest to Smallest) in order to isolate and regroup the matches from the Fuzzy Lookup.
- 3.3. Once, matches have been identified, the Statistical office can then manually verify the new values coming from the new Administrative Data source and apply the changes.
- 3.4. To update the Complete SBR from administrative data, follow these steps:
  - 3.4.1. Copy administrative ID of the matched record on the **AdminIDMatch** worksheet
  - 3.4.2. Select the Complete SBR and use the Find function (CTRL-F) and search the record with the correspond ID.
  - 3.4.3. Once the record has been found on the Complete SBR, manually copy and paste values from the **AdminIDMatch** worksheet to the Complete SBR worksheet.
  - 3.4.4. To facilitate this, the statistical office can use the New Window function located under the View menu and put the 2 windows side to side as shown in Figure 16.

\* The Fuzzy Lookup process returns all potential matches between the Complete SBR and the Administrative Data Source used. Since the similarity threshold was set to 1, only exact matches on the corresponding Administrative ID will be considered. These matches will have a Similarity score of 1, while non-matches will have a Similarity Score of 0.

Figure 15: Validating administrative ID matches and updating values

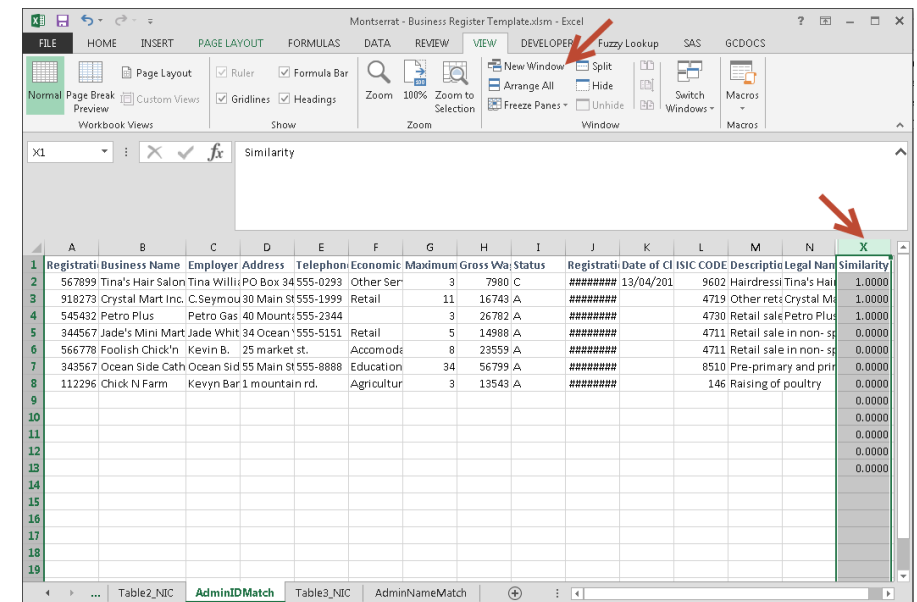
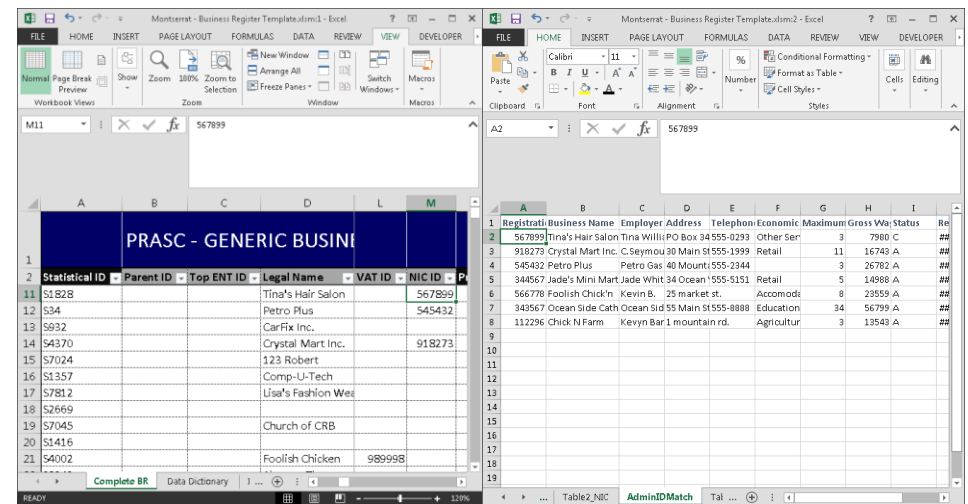


Figure 16: Using a New Window to perform manual updates



4. Compare unmatched records with the Complete SBR on Business Names (Legal, Trade, Employer, etc.)
  - 4.1. Once the Administrative Data matches have been validated and updated, the record linkage process can continue using other identifiable variables such as the business legal, trade or employer name.
  - 4.2. From the **AdminIDMatch** worksheet, select all unmatched records (records with a Similarity score of 0), copy (CTRL-C) them and paste (CTRL-V) them into a new worksheet. Rename the worksheet to **Table3\_Admin** as shown on Figure 17.
  - 4.3. Create/Add a new worksheet and name it **NameMatch** and select the top-left cell. **Name** refers to the desired name that will be used to perform the record linkage (Legal, Trade, Employer, etc.)
  - 4.4. From the Fuzzy Lookup tab, on the **NameMatch** blank worksheet, Click on Fuzzy lookup. A Window will pop up on the bottom right.
  - 4.5. Identify variables on which you wish to perform record linkage on each table
    - 4.5.1. Left Columns : **Name** from **Table3\_Admin**
    - 4.5.2. Right Column: **Name** from **Table1\_SBR**
  - 4.6. Click on the Join button (green arrow on Figure 18) to add the Fuzzy Join bonding.
  - 4.7. Adjust the Similarity Threshold to 0.50\*
  - 4.8. Press Go.

\*The similarity Threshold is a number between 0 and 1 that indicates the level of probability of a match. Since names can have multiple variations due to acronyms, accents, punctuations, etc., and that we want to include partial matches, the threshold should be set to anywhere between 0 and 1. It is up to the Statistical Office to decide what threshold is considered acceptable. It is recommended to set a low threshold and verify the potential matches discarding the matches with a scored deemed too low if necessary.

Figure 17: Creating table for subsequent record linkage

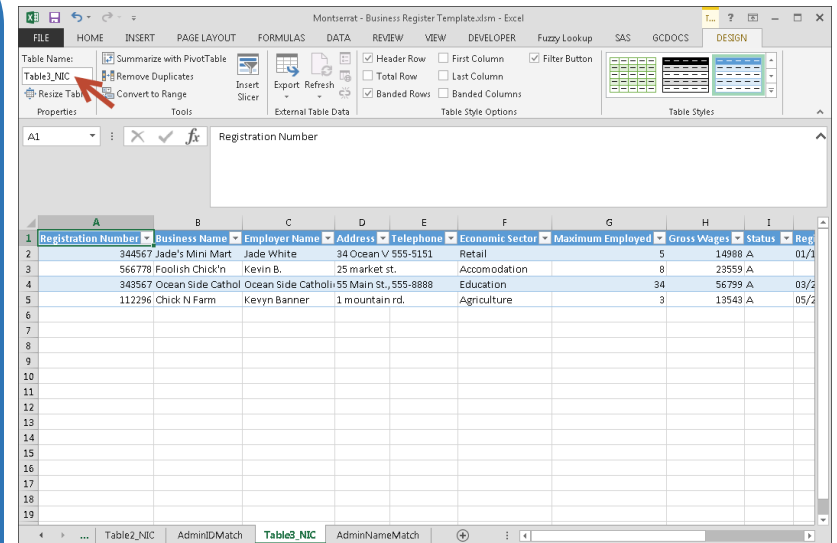
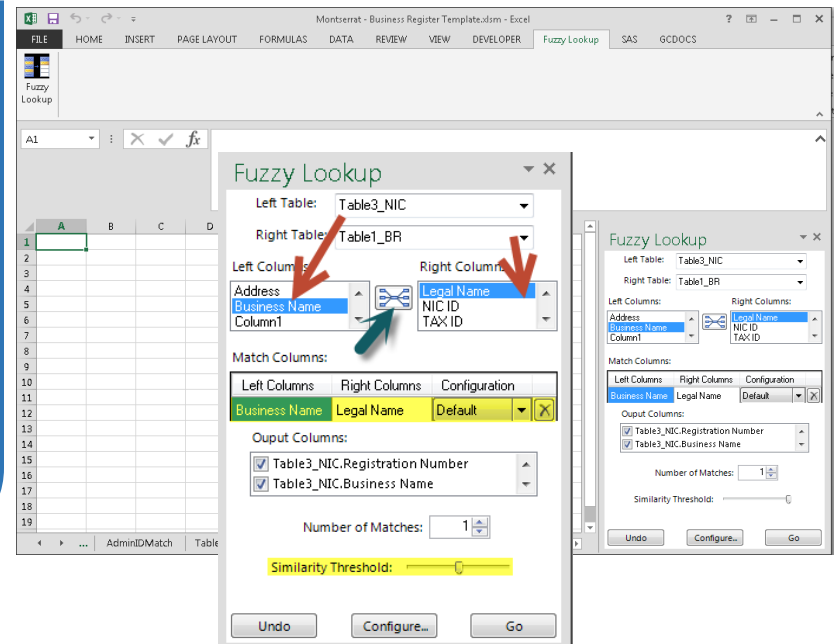


Figure 18: Record Linkage on Names using Fuzzy Lookup



## 5. Validating name matches and updating values

- 5.1. Once the Fuzzy Lookup process is completed, the results will appear on the blank worksheet that was previously named **NameMatch**.
- 5.2. On the far right, locate the Similarity column, select it and sort it in descending order (Sort Largest to Smallest) in order to isolate and regroup the matches from the Fuzzy Lookup. The Similarity score will range from the selected threshold (ex.: 0.50) to 1.
- 5.3. Verify all potential name matches and confirm all correct matches. Verification techniques will need to be established by the Statistical Office to identify all correct and incorrect matches.\*
- 5.4. Once correct matches have been identified, the Statistical office can then manually verify the new values coming from the new Administrative Data source and apply the changes to the SBR.
- 5.5. To update the Complete SBR from administrative data, follow these steps:
  - 5.5.1. Copy administrative ID of the matched record on the **NameMatch** worksheet
  - 5.5.2. Select the Complete SBR and use the Find function (CTRL-F) and search the record with the correspond ID.
  - 5.5.3. Once the record has been found on the Complete SBR, manually copy and paste values from the **NameMatch** worksheet to the Complete SBR worksheet.
  - 5.5.4. To facilitate this, the statistical office can use the New Window function located under the View menu and put the 2 windows side to side as shown in Figure 16 (From Step 3).
  - 5.5.5. Repeat step 4 to 5 for all subsequent matches on variables other than administrative identifiers.

\*Research might be required to determine if a match is correct. When there is not enough information to confirm a match, it is recommended to leave the record as is and add it as a prospect, later in the process.

Figure 18: Validating name matches and updating values

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	X
	Registrat Business Name	Employee Address	Telephone	Economic Activity	Maximur Gross Wt	Status	Registrat Date of	CTSI	CODI	Descripti	Legal Name	Similarity			
11	344567	Jade's Mini Mart	Jade Whi 34 Ocean	555-5151	Retail	5	14988	A	#####	90773	4711	Retail sal	Jade's Mini Mart	1.0000	
13	566778	Foolish Chicken	Kevin B. 25 market st.	Accomod	8	23559	A	#####	90773	4711	Retail sal	Foolish Chicken	0.8465		
14	343567	Ocean Side Catholi	Ocean Si 55 Main	555-8888	Educatio	34	56799	A	#####	8510	Pre-primary and primary edu	0.0000			
15	112296	Chick N Farm	Keyvyn Ba 1 mountain rd.	Agricultu	3	13543	A	#####		146	Raising of poultry	0.0000			

**WARNING:** Extra care needs to be taken when confirming a match. Creating incorrect matches will associate data of a business to another one, introducing errors. On the other hand, ignoring correct matches can introduce duplicates, creating overestimation of the business population.

When a record matched on names (and/or other variables) has two different and conflicting Administrative Data registration numbers, the record should be inspected thoroughly. This can occur if a record was previously incorrectly linked or if the structure of the business has changed. In any case, a complex profile of the business and its structure is required. Refer to the Statistical Business Register User's Guide for the steps to add new records, address duplicates and modify business structures.

6. Verify remaining records and confirm there are no records with similar information already existing on the SBR.

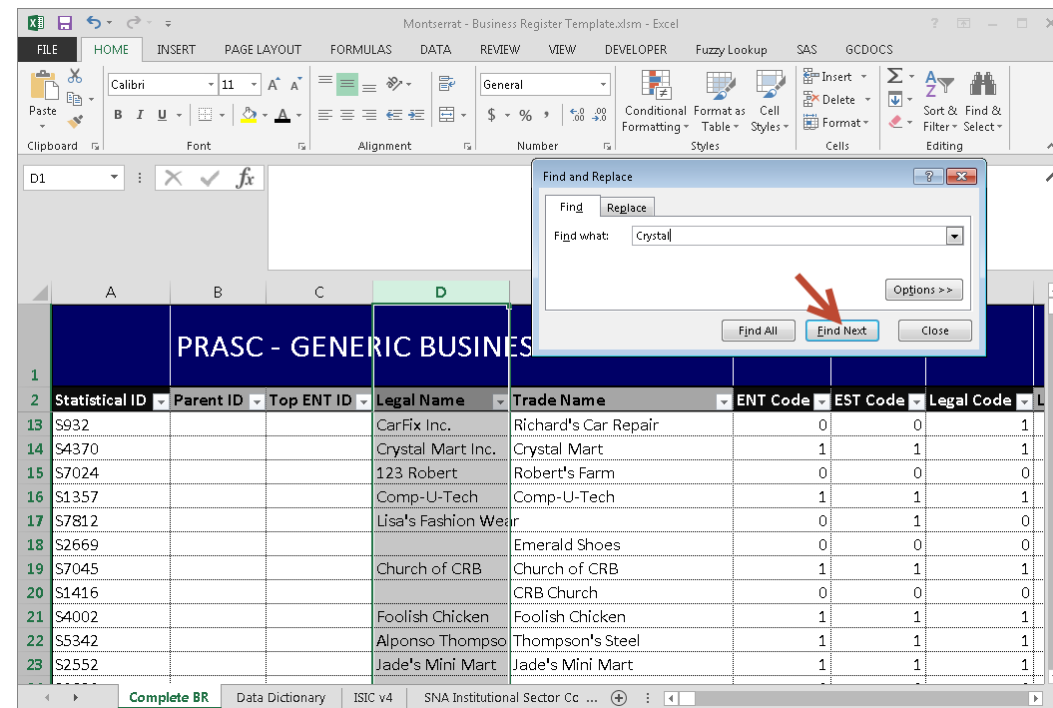
6.1. For each of the remaining unmatched record, search the Complete SBR worksheet to ensure that it is not a duplicate of a record already on the current SBR. Business names on either files might contain typos, or may be missing information causing problems with the registration number or the name match processes. The fact that an administrative data record couldn't be matched to a record on the SBR doesn't confirm that it should be added as a new record. Extra care should be taken to ensure that duplicates are not introduced on the SBR since it could lead to overestimation of economic activities in certain sectors and skew the estimates derived from the Statistical Business Register. Avoiding duplication is extremely important in order to ensure the quality of the Statistical Business Register.

6.1.1.1. From the Complete SBR sheet, perform string searches using the Find function. Refer to the Find section of this document for more details

6.1.1.2. Perform string searches on various fields such as:

- Legal Name (or partial legal name)
- Trade name (or partial trade name)
- Contact name
- Business Email
- Website
- Etc.

Figure 19: Searching Text String



- 6.1.1.3. When a match is found, adjust record on the Complete SBR with any new information from the administrative data file. Refer to the Statistical Business Register User's Guide for the steps for updating records and modifying business structures. Do not add them as new records; this will result in duplication.
- 6.1.1.4. Once a record has been identified as a new record, select the next available row on the Complete SBR worksheet and manually enter all relevant information from the administrative data file about the new business as displayed on Figure 20.

## Post Processing

### Steps:

9. Finalising administrative data update process:
  - 9.1. Save "Admin Data Load *sourceyyyyymmdd*" file and leave file in Administrative Data Updates folder.
  - 9.2. Create a backup copy and move under the appropriate backup folder.
  - 9.3. From the "Admin Data Load *sourceyyyyymmdd*" file, remove all worksheets used to update the SBR with administrative data.
  - 9.4. Save as "Statistical Business Register" to use as the current SBR and move into the corresponding Work Area folder.
  - 9.5. Create Snapshot according to pre-defined production schedule and file management strategy if required.

**Figure 20: Adding new records**

	Statistical ID	Parent ID	Top ENT ID	Legal Name	Trade Name	ENT Code	EST Code	Legal Code
19	S7045			Church of CRB	Church of CRB	1	1	1
20	S1416				CRB Church	0	0	0
21	S4002			Foolish Chicken	Foolish Chicken	1	1	1
22	S5342			Alponso Thompso	Thompson's Steel	1	1	1
23	S2552			Jade's Mini Mart	Jade's Mini Mart	1	1	1
24	S9639			Chick N F		0	0	0
25	S9342					0	0	0
26	S9372					0	0	0
27	S3804					0	0	0
28	S9440					0	0	0
29	S5122					0	0	0

## 4. Getting Information from the SBR

### 4.1 Extracting Data

There might be situations that requires users to work with specific variables, isolate records for analysis or simply extract a subset of the Statistical Business Register. Extractions should only be performed once the desired data has been identified.

The steps to extract data from the Statistical Business Register are listed below.

*Steps:*

1. Use Filters as desired
2. Select the desired data, including titles\*
  - a. Select cell in top left corner of the desired dataset.
  - b. Click and hold "Shift"
  - c. Select cell in bottom right corner of the desired dataset
3. Copy (right-click and select copy)
4. Open a new Excel Spreadsheet and paste the values (Right-Click and select Paste)
5. Save new Excel file in a secure location and ensure that its name clearly identifies it as a Statistical Business Register project or extraction.

\*It is important to not select the entire worksheet when selecting the data to extract once filters have been used. This operation will cause conflicts and might create unexpected results due to the multiple empty rows and the different formulas imbedded in the worksheet. Make sure that only the desired data is selected as described under step 2.

\*\*It is recommended to not copy extractions on worksheets of the Statistical Business Register file. The Statistical Business Register should remain a stable and clean product and should not carry over adhoc extractions and analysis projects. Create separate files to store these extractions.

### 4.2 Storing extracted data

As explained in the Organising Data section, it is important to designate areas for different work activities and file types. The Statistical Business Register should remain a stable and clean product and should not carry adhoc extractions and analysis projects. It is therefore recommended to save all extractions as separate files and store them in the appropriate folder. Refer to the Organising Data section for more details.

### 4.3 Demographic Statistics (Pivot Tables)

A pivot table is a data processing tool used to query, organise, and summarise data or information between spreadsheets, tables or databases. It is useful when analysing a large amount of data, as it allows users to apply specific criteria to reorganise data and create reports.

#### 4.3.1 Statistical Business Register's Demographic Statistics

The Demographic Statistics tab contains a pivot table that is linked to the Statistical Business Register data and will therefore reflect current counts as data is modified. The pivot table can be manipulated freely by users to obtain specific set of information for analysis or reporting purposes. For example, it can be used to retrieve business population counts, identify the total sales of a particular industry or the average number of employees in a specific sector. Counts can also be retrieved in 2 different snapshots and compared in order to identify the variations of businesses births and deaths in specific sectors or analyse the overall economic trend of the business population. The current Demographic Statistics table contains several dimensions: Business Status Code, ISIC (with description)\*, Enterprise Count, Establishment Count, Sum of Number of Employees, Sum of Wages and Sum of Sales.

#### Note

\*In order to create the ISIC with description dimension, two columns were combined on the Complete SBR tab. This field has been locked and hidden since its only purpose is to present the ISIC with its description as one variable within the pivot table.

**Figure 21: Demographic Statistics**

Business Status Code	Enterprise Count
<b>Demographic Statistics</b>	
<b>Business Status Code</b>	(All)
<b>ISIC - Description</b>	Enterprise Co
3510 - Electric power generation, transmission and distribution	
3811 - Collection of non-hazardous waste	
4100 - Construction of building	
4330 - Building completion and finishing	
4520 - Maintenance and repair of motor vehicles	
4711 - Retail sale in non- specialized stores with food, beverages or tobacco predominating	
4719 - Other retail sale in non-specialized stores	
4730 - Retail sale of automotive fuel in specialized stores	
4741 - Retail sale of computers, peripheral units, software and telecommunications equipment in specialized stores	
4771 - Retail sale of clothing, footwear and leather articles in specialized stores	
5510 - Short term accommodation activities	
5610 - Restaurants and mobile food service activities	
9491 - Activities of religious organization	

#### 4.3.2 How to create a pivot table

The steps to create pivot tables are listed below. Figure 25 displays how the fields were categorised for the Statistical Business Register's Demographic Statistics pivot table.

*Steps:*

1. Select a blank worksheet
2. From the Insert tab, select Pivot Table in the Tables section menu
3. Click on the icon (table with a red arrow) at the right of the Table/Range text box under the "Select a table or range" radio button
4. Select the table, including column titles, from which you want to create the pivot table for. In this case, select the entire Statistical Business Register (from the Complete SBR worksheet), including titles and blank records. Press Enter
5. From the PivotTable Fields pop up window, drag and drop the desired fields in the appropriate areas. Fields can be divided in 4 categories (Filters, Columns, Rows, Values).
6. Once a field is dragged under the Values area, select the item, and select Value Fields Settings
7. From the Value Fields Settings pop up window, adjust name if necessary and choose the type of calculation desired (Sum, Count, Average, etc.) and click OK

**Figure 22 Pivot Table fields by categories**

Categories	Fields
<b>Filters</b>	<ul style="list-style-type: none"> <li>○ Business Status Code</li> </ul>
<b>Columns</b>	<ul style="list-style-type: none"> <li>○ N/A (automatically populated when Values have been selected)</li> </ul>
<b>Rows</b>	<ul style="list-style-type: none"> <li>○ ISIC and Description Combined (hidden field on the Complete SBR worksheet)</li> </ul>
<b>Values</b>	<ul style="list-style-type: none"> <li>○ Enterprise Count (ENT Code – Renamed to Enterprise)</li> <li>○ Establishment Count (EST Code – Renamed to Establishment)</li> <li>○ Sum of Number of Employees</li> <li>○ Sum of Revenue</li> <li>○ Sum of Gross Wages</li> <li>○ Sum of Sales</li> </ul>

## List of Acronyms

SBR – Statistical Business Register

ENT – Enterprise

EST – Establishment

ISIC Rev. 4 – International Standard Industrial Classification of All Economic Activities, Revision 4

ID - Identifier

Social Security – National Insurance Service / National Insurance Corporation / Social Security

TAX – Tax filler data

VAT – Value Added Tax

## Definitions

*For additional definitions, please refer to the Data Dictionary.*

Active: Records with Active business statuses are engaged in an economic activity. This business status applies to both enterprises and establishments.

Administrative Data: Files of data collected by other government bodies such as Registrar, Tax, VAT, or Social Security for the purpose of administering their own respective government programs.

Administrative Identifier (ID): ID used by a particular administrative source to identify a company or business in its own records.

Administrative Source: Term referring to other government bodies that share administrative data (Registrar, TAX, VAT, or Social Security).

Amalgamation: The legal process of joining two or more companies to form a single new company. This process is overseen by the Registrar of Companies within Central Registrar Department. Companies must file “Articles of Amalgamation” with Social Security in order to undergo an amalgamation.

Backup: A Backup refers to the process of making copies of data or data files to use in the event the original data or data files are lost or destroyed.

Statistical Business Register (SBR): A comprehensive, structured list of businesses and companies engaged in economic activity. The SBR is used as a frame for economic surveys, and as a source of information for demography of the business population.

Business Status: Term used to reflect the actual legal or operating state of a business or company. There are 5 business statuses on the SBR: New Business, Active, Inactive, Closed or Amalgamated.

Business Rules: Specifications incorporated into the SBR excel file that dictate the value of certain fields based on the values of other variables within the Statistical Business Register.

Cell: In spreadsheet applications, a cell is a box in which you can enter a single piece of data. The data is usually text, a numeric value, or a formula. Individual cells are usually identified by a column letter and a row number.

Column: In spreadsheets, a column is a vertical group of cells. Spreadsheet columns are usually identified by letters.

Confidentiality: Within the context of the SBR, confidentiality refers to the concept of protecting sensitive statistical information that could result in disclosing information identifying or about a particular business or company.

Complex Structure: A group of related records on the SBRS make up a complex structure. There are two types of complex structures: complex operating and complex legal structures. Enterprises within complex structures are highlighted in blue and establishments within complex structures are highlighted in green.

Data: Data refers to distinct pieces of information, usually formatted in a special way. In this case, data is interpreted as computer information that is transmitted or stored.

Data Processing: Refers to a series of operations on data, especially the handling of information by a computer, to retrieve, transform, or classify information.

Duplication: Caused when two (or more) records are mistakenly created on the SBR to represent the same entity. For example, two enterprises with very similar legal names or two establishments at the same address with the same ISIC code.

Enterprise: Record on the SBR that represents a legal unit producing economic goods and services. Enterprises can be incorporated or unincorporated entities. Enterprise level information consists of legal information, as well as information about the economic activities of the entity as a whole. Highlighted in blue on the SBR when part of a complex structure.

Establishment: A record on the SBR representing a part of an enterprise that is characterised as having a single business activity and location. Highlighted in Green on the SBR.

Excel: Microsoft Excel is a spreadsheet program included in the Microsoft Office suite of applications. Spreadsheets present tables of values arranged in rows and columns that can be manipulated mathematically using both basic and complex arithmetic operations and functions.

Filter: Function in Excel that allows you to display only rows containing a specific string (of letters or numbers), while hiding rows that do not contain that particular string.

Invalid Structure: A complex structure on the SBR made up of one enterprise record with a single linked establishment. This is invalid and should be corrected.

ISIC Code: 4-digit code used to classify all economic activities.

Macro: A macro is an automated input sequence that imitates keystrokes or mouse actions. A macro is typically used to replace a repetitive series of keyboard and mouse actions and is common in spreadsheet and word processing applications.

Manual Update: An update made to the SBR by an updater and not through an automatic process of uploading administrative data.

Match: Refers to a record that has been successfully matched during a record linkage process.

New Business: Business Status assigned to a business or company on the SBR when it is known to exist in some form, but its legal registration and state of economic activity is not yet known.

Parent: Within a legal structure, the parent is the enterprise that owns the majority share of another enterprise.

Production Schedule: An operational schedule that displays the timetable for the use of resources and processes required by a business to produce products or provide services. A production schedule can help manage operational processes, work activities and systems maintenance.

Prospect: On the SBR, a prospect is a record from an original SBR that still had no legal identifiers after the record linkage process (Registrar ID, TAX ID, VAT ID, or Social Security ID). The business is presumed to exist, but it has not yet been confirmed to exist through administrative sources and is therefore ineligible for statistical sampling. All prospects have a Prospect Flag of 1, New Business as a Business Status and Legal Code of 0.

Read-only: Read-only is a designation for any object or construct which can no longer be altered after creation, it can only be read. It is a safety measure which protects files and data from accidental or intentional alteration or deletion and may be imposed only for select users or groups of users.

Record: The word record in reference to the SBR refers to one entity represented by one line on the current database. A record can contain enterprise level information only, establishment level information only, or both.

Record Linkage: Process of finding records that represent the same entity across different data sources.

Row: In spreadsheets, a row is a horizontal group of cells. Spreadsheet rows are usually identified by numbers.

Size Variables: Variables that indicate the economic impact of a business such as "Employees", "Social Security Number of Employees", "Revenue", "Gross Wages" and "Sales". See SBR Data Dictionary tab for more information.

Snapshot: Refers to a read-only, static view of a database at a point in time. In this case, we refer to the Statistical Business Register Snapshot as a copy of the Statistical Business Register Workbook at a point in time, which will be locked and never modified thereafter.

Spreadsheet: A spreadsheet is a table of values arranged in rows and columns. Each value can have a predefined relationship to the other values.

Statistical Identifier (ID): A unique number that is assigned to each record on the SBR to serve the purpose of identifying the unit.

Survey Eligibility: Concept describing whether or not a record can be surveyed, and what types of surveys can be sent to it. There are two levels of survey eligibility on the SBR: Enterprise and Establishment.

Survey Frame: A subset of units that match the specification of a survey's target population.

Table: Refers to data arranged in rows and columns. A spreadsheet, for example, is a table.

Template: A template refers to a document or file having a pre-set format, used as a starting point for a particular application so that the format does not have to be recreated each time it is used.

Top Enterprise: Enterprise at the top of a legal structure. It is not owned by any other company within the business population.

Unmatched: Refers to a record that couldn't be successfully matched during a record linkage process.

Workbook: A workbook is an Excel file that contains one or more worksheets. Each of the workbook's worksheets are in separate tabs on the bottom of the Excel window. You can switch between worksheets by clicking on the worksheet's tab on the bottom of the Excel window.

Worksheet: An Excel worksheet is a single spreadsheet that contains cells organised by rows and columns. A worksheet begins with row number one and column A. Each cell can contain a number, text or a formula. A cell can also reference another cell in the same worksheet, the same workbook or a different workbook.