**W**estern **N**orth **C**arolina **D**ashboard

2018

JUNIOR DATA AND ECONOMIC ANALYST

OFFICIAL GUIDE BOOK

Table of Contents

**Getting Started2**

Tableau2

SQL Server and SSMS2

Erwin Data Modeler3

**SQL Server Basics4**

Loading Master Tables4

Loading a data file4

Loading a script5

Defining Organizations5

**ECONDEV6**

Connecting6

Importing Data (DDL)6

**Erwin Data Governance8**

Entities8

Inner Joins and Attributes9

Inner Joins and Attributes with Order10

**Updating11**

Data Series11

Detailed Updating Existing Data Series13

**Tableau19**

Login and Workbook Edits19

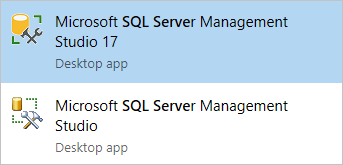
20

**Getting Started**: *This section will go over the necessary software installations on your personal machine*

1. **Tableau (Academic Installation) current version: 2018.1**
   1. Visit: <https://www.tableau.com/academic/students> and click “Get Tableau for free”.
   2. Fill out the form with your student email and await for a verification email from SheerID containing your license key.
   3. Visit: <https://www.tableau.com/products/desktop/download?signin=academic>

to download the academic version of Tableau.

* 1. Open Tableau -> Click the “Help” drop down tab -> Select “Manage Product Keys” -> Select “Activate” and enter in the license key that was emailed to you. You’re now activated!

1. **SQL Server 2017** 
   1. Visit: <https://www.microsoft.com/en-us/sql-server/sql-server-downloads> and click “download now” under developer edition.
   2. Select “Custom Install” -> Verify where it will install on your machine -> Click Install.
   3. Once SQL Server Installation Center opens up -> click under the left hand column “installation” -> choose the first option “New SQL Server Stand alone installation”.
   4. Once SQL Server 2017 Setup opens -> leave “specify a free edition ‘developer’ “ selected and click next.
   5. Accept license terms and click next (you may then be prompted to restart your machine) to return back after the restart use file explorer to browse out to “SQLServer2017Media -> Developer\_ENU and run “SETUP”
   6. Check “Use Microsoft Update to check for updates” and click next
   7. Click next after install rules runs \*the warning is normal next to Windows Firewall
   8. Feature Selection (Select all except PolyBase Query Service) -> click next
   9. Instance Configuration: Keep default and click next
   10. Server Configuration: Click next
   11. Database Engine Configuration: Leave windows authentication mode selected -> Under “specify SQL Server Administrator” click “Add Current User” -> Click Next
   12. Analysis Services Config: Leave default Tabular Mode and Add Current User below -> Click Next
   13. Leave defaults and click next until the “Distributed Replay Controller” -> click “Add Current User” then click next
   14. Click Next on the following window and “Accept” consent to install Microsoft R Open and then again for Python
   15. Preview the “Summary” and click “Install”
2. **SQL Server Management Studio**
   1. Open “SQL Server 2017 Installation Center” -> Click “Installation” under the left-hand side. **OR** go to: <https://docs.microsoft.com/en-us/sql/ssms/download-sql-server-management-studio-ssms?view=sql-server-2017> and click “download SQL Management Studio”
   2. Run the download and click “Install”
   3. To Connect to Server: Select the Server name drop down -> “browse for more….” -> Expand “database engine” and select your machines name -> click “OK” -> click connect
   4. You are now ready to use SQL Server!
3. **Erwin (Academic Edition)**
   1. Visit <http://go.erwin.com/erwin-academic-edition-free-trial> and fill out the form.
   2. Type in “Dan Clapper” for the professor name and “work study” for the course name.
   3. At the bottom, click “download academic edition” and install Erwin on your machine.
   4. Check your email and visit: <https://erwinhelp.zendesk.com/hc/en-us/requests/new?ticket_form_id=431568> to “submit a request” only enter the fields that the email tells you to enter below is an example:
      1. Choose erwin Data Modeling Licensing in the first drop down
      2. Enter your school email address
      3. For "Subject" enter Academic Edition
      4. For "Description," enter:

(First Name, Complete school name - no abbreviations,

Name of the course you are taking, Name of your professor)

* + 1. For "Severity," choose 4
    2. For "Product Name" drop-down, choose erwin Data Modeler
    3. For "Product Edition" drop-down, choose Academic
    4. Enter Telephone Number
  1. Once those fields are filled, you can then click Submit and await for an email containing your license code (this will take a few hours).
  2. Open Erwin and enter the license code at the first prompt, you have now completed the installation/activation of Erwin.

**SQL Server Basics:**

1. **Creating a Database**
   1. Right click “Databases” -> New Database -> set name such as “STG2” and leave owner as <default>
   2. Under “Database Files” scroll over to “Path” where you can select to store the filegroup to another drive to prevent filling up your machines primary drive. \*You may need to go in and create a few folders before hand.
   3. (STG2 and -> import tables (add tables / drop tables) -> import meta data. (View excel for descriptions and legends)
2. **Loading Master Tables in the DV2 Database (10 total)** 
   1. Right click the DV2 database in SSMS -> Tasks -> Import Flat File (If the data file being imported is a .TXT file)
   2. “Import Flat File” wizard will be ope: click next -> browse out to the desired file to be imported ex. *Metadata/Loaded/BEA\_GDP\_Component\_Type.txt* -> click next
   3. Click next once on the “preview data” menu
   4. Modify Columns to the desired datatype/size and specify null or not null -> click next
   5. Review summary and click finish, you should get a quick green checkbox to flash the screen showing the data file has been successfully loaded to that table.
   6. Continue this process for the remaining tables. Refer to **DV2 Master Tables Instructions** for specifics on datatypes/sizes and location of the .TXT datafile.
3. **Loading a Data File**
   1. If the file is in an excel spread sheet: open -> go to desired tab (spreadsheet) and save as a .txt (tab deliminated) file in the “loading zone” folder.
   2. Loading zone -> open in note pad to view and verify file.
   3. Open SSMS and right click on desired database -> Tasks -> Import Data which will launch the Import/export wizard
   4. Choose “Flat File Source” for Data Source -> Browse out and select your .txt file that you saved in the loading zone -> clear out the “<none> in “Text Qualifier” and replace it with a single quote: “
   5. Select the “advanced” tab on the left to view/edit the columns
   6. \*\*\*you will need to make sure and change the “OutputColumnWidth” on each column, this is where it’s helpful to expand the table and its corresponding columns in SSMS’s object explorer window put the import wizard beside to make it quick to compare the data type and output column width.

Ex: you may need to change Data Type from string to int and output column width from the default 50 to 20 or perhaps 200.

* 1. Click next and choose the Destination as “OLDB provider for SQL Server” (be sure all other locations of file are closed before proceeding) then click next and select the destination table you’d like that data to be imported to
  2. Go to mappings (verify create new table is selected or delete rows in destination table)
  3. Click Finish and then Finish again. Ensure all green arrows (success)

1. **Loading a Script**
   1. File -> open file -> Scripts -> posting facts -> dynamic insert
   2. Take table name from column source and paste into script where according
2. **Defining Organizations**

The process begins by preparing the data in the Data > Dimensions folder. In this folder one should find two Excel files:

* 1. **WDB\_NC\_County\_Xref.xlsx** containing data about Workforce Development Boards in North Carolina, and
  2. **GeoFIPS\_USStatesCounties\_NCPZWDB.xlsx** containing several worksheets.

The first workbook is used as input to the second workbook. Both workbooks are compiled by the NCDD support team from data gathered by visiting different pages on the North Carolina Department of Commerce website. The NCDOC website should be visited periodically and the two workbooks updated accordingly. Updates to these workbooks should be applied to the NCDD database. The following discussion outlines the process. In general, the process is the same for all dimensions:

1. Load a text file to the STG2 database in a format and layout conforming to the text file layout
2. Load a Work table specific to the dimension with data from the source-conforming table, the Work table having a layout specific to the dimension and designed to interact with the support team in an Excel-friendly manner. The name of the work table will contain the character strings “\_XLSX\_” and “\_WRK”
3. Optionally, the NCDD support team can query the XLSX Work table with a query in Excel, edit the contents, save the edited contents as a TXT file and load it back into the XLSX Work table.
4. The XLSX Work table is loaded into the STG Work table as Hash Key and Hash Difference values are created
5. The STG Work table is used as the primary driving input to the Dimension Populate and Update routines.

**ECONDEV**

**1) Connecting to Econdev/Econprod**

1. Click the Pulse Button on the task bar to (ensure Pulse is installed first):
   1. Establish connectivity through VPN ([VPN.WCU.EDU](mailto:VPN@WCU.EDU)), and
   2. Connect to Pulse client with username and password (Username and Password)
2. Use Remote Desktop (Right-click Windows Icon>Run>mstsc then Enter)
3. Connect to either Stein.wcu.edu (Development) or Greenspan.wcu.edu (Production) with wcu\lbergh as the username.
4. Once you are logged into the server, open SSMS. (Either pinned to the taskbar or left-click Windows icon and type ssms and click on the icon.)
5. On the login screen, make sure Windows Authentication is selected and hit connect.
6. To connect to OneDrive while connected to Stein or Greenspan:
   1. Open Web Browser
   2. Go to [http://email.wcu.edu](http://email.wcu.edu/)
   3. Login using your WCUid (everything before the @ symbol)
   4. Click the 3x3 icon in the upper left-hand corner and select “OneDrive”
7. When connecting to excel through the batch file(cmd): ensure all instances of excel on your local machine are closed out.

**2) Importing Data into ECONDEV Database DDL:**

1. Install SSMS on employee virtual machine
2. Refer to page 3 for instructions.

**3) Open SSMS**

1. Connect: “stein\econdev” without quotes for server name -> leave windows authentication checked. (*econprod is final product for future use)* -> Click “Connect”
2. Right Click on Databases -> New Database -> Name “STG” -> Click “OK”
   1. Follow that process for the remaining databases “DV1, STG2, ect.)

**Order for running scripts in DDL:**

1. **STG -> DV1-> STG2 -> DV2-> ACS -> MDV2**
2. Browse out to DDL -> STG (select first script and execute in SSMS) and follow this order with all the reaming scripts in that folder (be sure to work top to bottom)
3. Browse out to DDL -> ACS and run the scripts in chronological order (top to bottom)

**Troubleshooting an Error:**

* When the script executes with errors: Verify the views and potentially add or rename the view according to what the script is trying to execute it as.
* View the column name if the script has a red squiggly line under the name and correct potential typos in the script, this should allow you to execute.
* If script won’t run after troubleshooting, then close out -> and copy the script name in a word doc to keep track of which ones didn’t run.

**Import Data File to STG2:**

1. Right click “STG2” -> import data -> this will launch the wizard
2. Choose data source (file you wish to import ‘ZLLW’)
3. Import using the “create new table” feature.

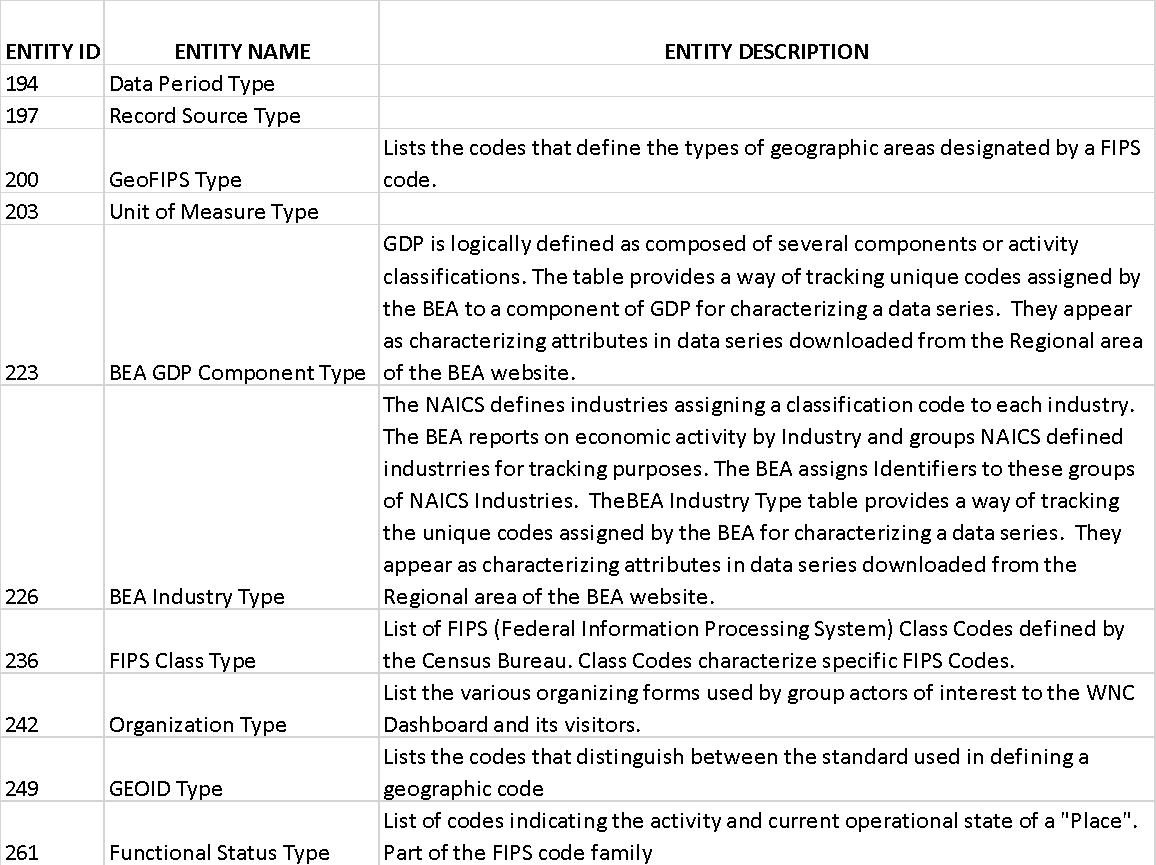
**Erwin Data Modeler**

*Overview: This section is critical for data governance, we must ensure the Erwin models line up with the database for quality control reasons. Later in this section you will learn how and why the order is important for things such as hash keys. Our goal is to produce the same table results as seen below but in SSMS when queried against our database.*

1. **Listing Entities**
   1. Open your desired Erwin data model such as: WNCD\_DV2\_Master\_Tables\_v01.0
   2. Click the “Tools” drop down and select “Query Tool”
      1. At the top of this window, select the “Connect to Erwin ODBC”
      2. Create a simple “select” “from” script selecting your desired fields such as entity id, name, description, ect.
      3. To list this entity data run this script in the query window:

**SELECT** ID@ AS 'Entity ID', NAME AS 'Entity Name', DEFINITION AS 'Entity Description'

**FROM** ENTITY

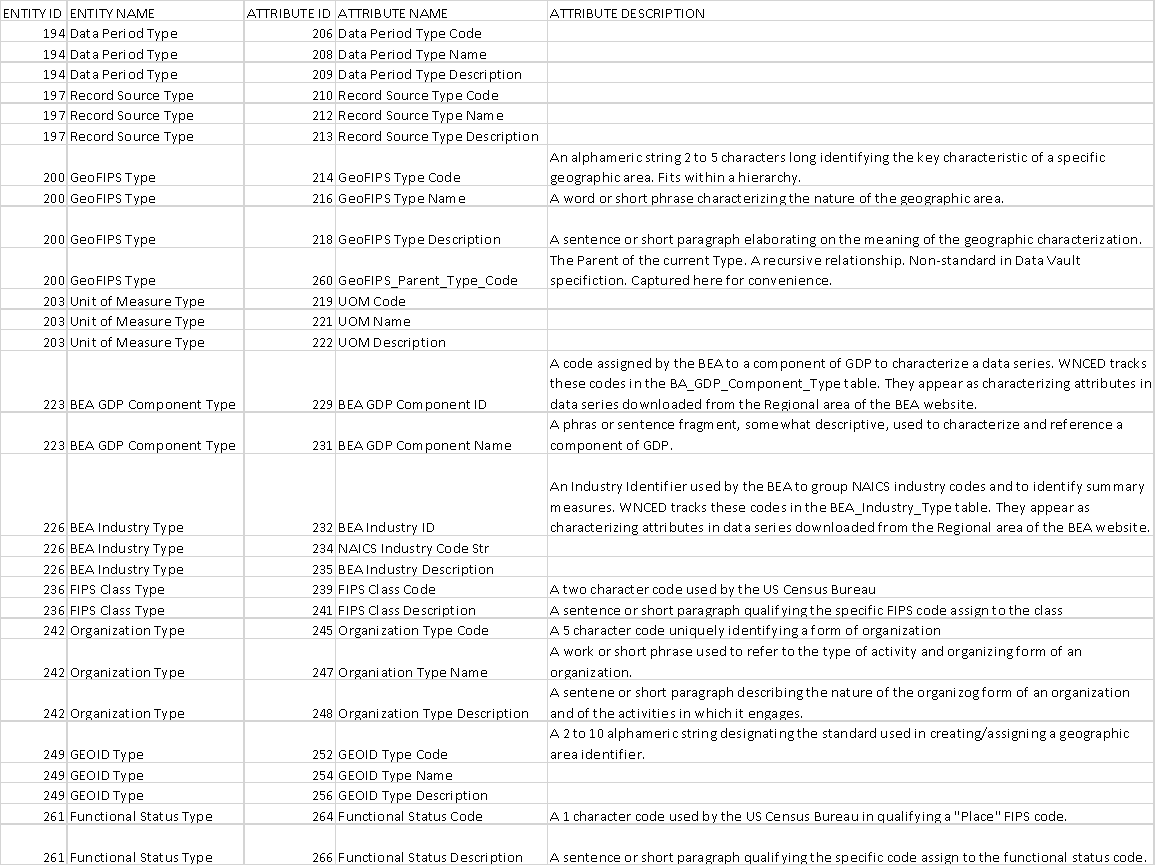
* 1. You can now click the save button in the middle right of the query tool window and save the results as a .CSV file to view and edit.
  2. Below is an example table of the DV2 model entities:

1. **Inner joins: Listing Entities an Attributes**
   1. Follow the same steps as part one to list entities
   2. Paste the following script in the query window and run:

**SELECT** E.ID@ AS 'ENTITY ID', E.Name AS 'ENTITY NAME' , A.ID@ AS 'ATTRIBUTE ID' , A.Name AS 'ATTRIBUTE NAME', A.Definition AS 'ATTRIBUTE DESCRIPTION'

**FROM** M0.ENTITY E INNER JOIN M0.ATTRIBUTE A

**ON** E.ID@ = A.OWNER@

* 1. Below is an example table of the results from running that query on the WNCD\_DV2\_Master\_Tables\_v01.0 Erwin Model:

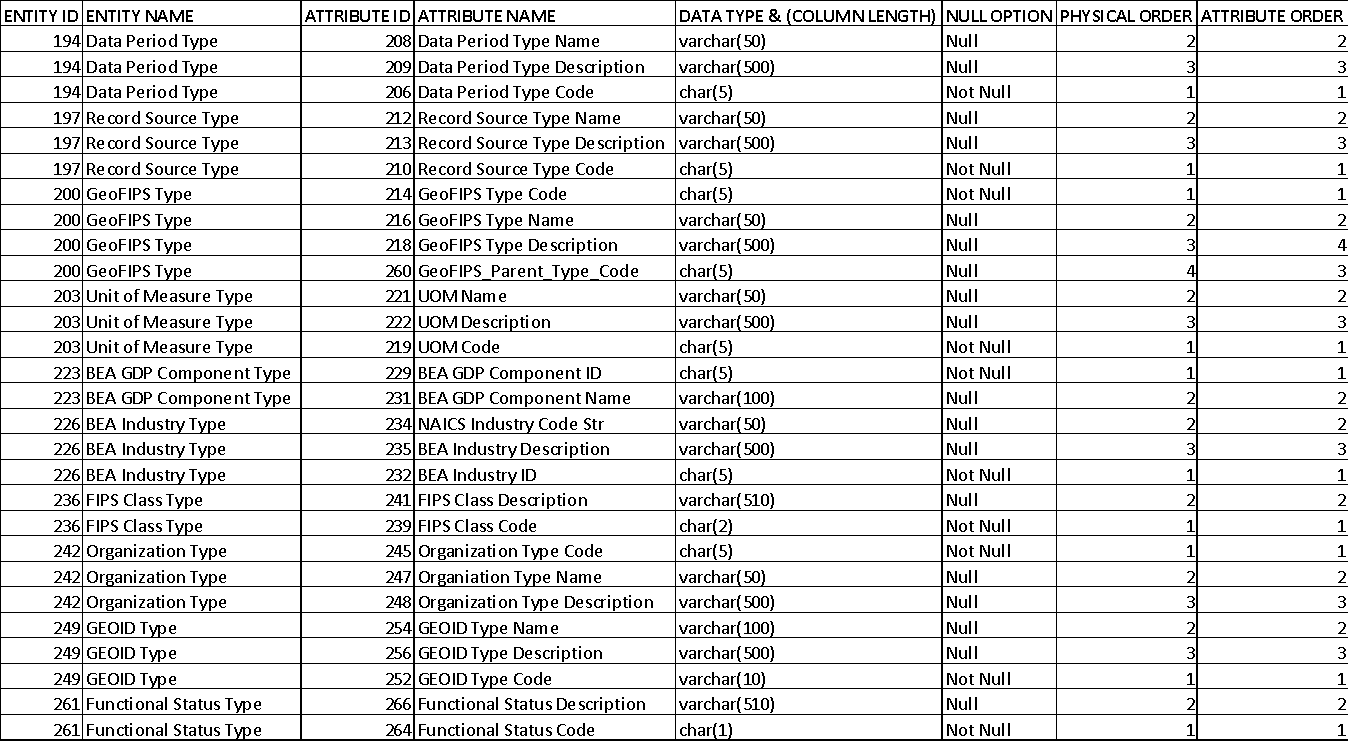
1. **Inner joins: Listing Entities an Attributes with Order**
   1. Follow the same steps as part two to list entities this time selecting to show more columns such as Null Option Type, Physical, and Attribute ‘Logical’ Order.
   2. Use an ORDER BY at the end of the script to sort the rows first by Entity ID then Attribute ID
   3. Paste the following script in the query window and run:

**SELECT** E.ID@ AS 'ENTITY ID', E.Name AS 'ENTITY NAME' , A.ID@ AS 'ATTRIBUTE ID' , A.Name AS 'ATTRIBUTE NAME', A.Physical\_Data\_Type AS 'DATA TYPE & (COLUMN LENGTH)', TRAN(Null\_Option\_Type) AS 'NULL OPTION', Physical\_Order AS 'PHYSICAL ORDER', Attribute\_Order AS 'ATTRIBUTE ORDER'

**FROM** M0.ENTITY E INNER JOIN M0.ATTRIBUTE A

**ON** E.ID@ = A.OWNER@

**ORDER BY** 1,2

* 1. Below is an example table of the results from running that query on the WNCD\_DV2\_Master\_Tables\_v01.0 Erwin Model (\*\*\*note in that this example, the attribute order differs from the physical order in only t

**Data series update process (more detailed with explanations starting on p13):**

Our data sources:

1. **BEA** (richest) 90%
   1. Annual data on the county level
   2. GDP at county level
2. **Federal reserve**
3. **Zillow**
4. **Census Bureau**

**Update Flow**

**Personal Machine** (Development Phase)

**Stein Server** (Testing Phase)

**Greenspan Server** (Production phase)

1. Monitor the Source and detect the release of new data (completely manual effort)

* Assign a dedicated person to setup and execute a new data checking process (browse out to zillow and download the latest CSV file and compare it to the older one to see if and data has been added)

1. Download new data from source and place it designated place on SharePoint

* Change the name of the columns so it matches the data series (Change 2018-06 to 2016M06)
* Once the CSV file has been confirmed to contain new data, upload the file or files to sharpoint (Documents > Data > Downloaded Files > Zillow)
* Change the name of the uploaded .CSV by adding ZLLW\_ at the beginning and save to LoadingZone directory on your computer (not sharepoint)

1. Create text file in Loading Zone for importing into STG2

* Open the downloaded .CSV file, Click insert tab on the excel toolbar > table, ensure the new table covers all the data
* Click File > Save As (browse out to your LoadingZone Folder)
* Name the file the same that you made it on SharePoint and set “save as type”: Text(Tab deliminated) and click save

1. Populate the appropriate STG\_XLSX\_???\_WRK tables in STG2 with the new imported data. The Measure Definition and Data Series tables will always need to be populated with new data. The Data Period and GeoArea Subject Areas should not require adjustment. Ignore these Dimensions for the time being.

* Right click the database > tasks > import data (this will launch the data import wizard)
* Select flat file source and then browse out to loading zone folder (change file type to .csv so it shows up)
* Set text qualifier to a single quote: “
* Destination naming “STG\_ZLLW\_...”
* Edit mapping and ensure “create destination table it checked”
* Next > run immediately
* Open “Populate STG\_XLSX” Script and edit the table name (this is marked with a green comment stating “needs to be updated manually”)
* Open the view “” in the ACS database and add the following statement right after the from line at the end: 
  + Where record sources = “ZLLW”
* Enter the appropriate measure business key in the populate script
* Pause here to troubleshoot some issues with editing the script -Brad will get back in touch in the coming days.88888888888888
* In the Populate\_STG\_XLSX\_DataSeries\_WRK\_DynamicInsert edit: Table Name, Record Source and use this query to find which one you need:
  + Select \*
  + From dob.STG\_XLSX\_MeasureDefn\_WRK
  + WHERE Record\_Source=’ZLLW’
* And edit measure business key, lastly \*comment out the truncate statement except for the first time

1. Edit the data in the WRK tables as needed. The Measure Definition will need to be updated with vintage data.
2. Execute the **PopulateDV2\_Measure\_Tables**script
3. Execute the **Post\_Observations** script
4. Refresh the data in the appropriate XLSX Tableau Data feed workbook(s)
5. Refresh the target Tableau workbooks using the XLSX Tableau Data feed workbook that was just refreshed.
6. Test/Validate the Refreshed Tableau workbook
7. Publish the refreshed Tableau workbook to the Public Tableau Server.

Updating Existing Data Series

# Overview

At a very high level, what needs to be accomplished to update a data series that currently exists in the database consists of three primary parts:

* Part 1 – Identify one or more data series to be updated from a single source file.
* Part 2 – Update the database with the new data
* Part 3 – Refresh the Tableau Visualizations with the new data.

Updating the database (Part 2 above) entails applying the new data to two groups of tables:

* The Measure Definition tables need to reflect:
  + The availability of observations for the most recent time periods
  + The likelihood of changes to previously posted observations
  + Changes that may have occurred in other metadata about a data series
* The Measure Observation tables need to include:
  + Observations for the more recent time periods not previously included.
  + Revised Observation values for those previously posted.

A more detail description of the process of applying new data to existing data series consists of the following steps:

1. Identify Likely Data Series to be Updated
2. Determine the ability to update one or more data series with newly available data
3. Apply new data to the database:
   1. Save the Incoming data file in the appropriate subfolder.
   2. Load the file to the Stage Database in a file-image table.
   3. Prepare the Stage tables
      1. Load the **STG\_XLSX\_MeasureDefn\_WRK** table from the file-image table with the Measure Definition update data based on an analysis of the downloaded data and metadata files.
      2. Load the **STG\_SLSX\_DataSeries\_WRK** table with Observations from file-image table.
   4. Update the Persistent Data Repository with the data in the Stage tables
      1. Execute the script “**Post\_Observations.sql”** to apply new observations to the existing data series.
      2. Execute the script “**PopulateDV2\_Measure\_Tables.sql”** to update the Measure Definition tables
4. Refresh the Tableau Workbooks with new data from the database:
   1. Identify the Tableau Workbooks that will be impacted by new data along with the Excel workbooks that are used to pass data from the database to Tableau.
   2. Refresh the Excel-to-Tableau workbooks and new data from the database,
   3. Save the Excel-to-Tableau workbooks in the appropriate Data subfolder
   4. Replace the appropriate Excel workbooks in the Tableau **Datasources** folder with a copy of the corresponding Excel-to-Tableau workbooks just refreshed.
   5. Refresh the impacted Tableau workbooks.
   6. Verify the integrity and completeness of the Tableau refresh.
   7. Retire the current version of the Tableau workbooks.
   8. Publish the new versions of the Tableau workbooks to the Public Tableau Server.
   9. Save the refreshed workbooks in the Tableau **Workbooks** folder

# Identify Likely Data Series to be Updated

See Appendix A for a comprehensive list of Measure Tables. Each series is identified by a unique line number within the Measure Table. Generally, new data is made available by the Measure Authority for all the data series appearing in a Measure Table at the same time.

Maintain a list of web addresses from which data is downloaded. Associate each Measure Table with a data source web address. Here is an example:

BEA Regional Data: <https://apps.bea.gov/regional/downloadzip.cfm>

CENSUS Poverty Rate: <https://www.census.gov/data-tools/demo/saipe/saipe.html?s_appName=saipe&map_yearSelector=2016&map_geoSelector=aa_c&menu=grid_proxy&s_year=2016,2015,2014,2013,2012,2011,2010,2009,2008,2007,2006,2005,2004,2003,2002,2001,2000,1999,1998,1997,1996,1995,1993,1989>

Maintain a list of Measure Tables by anticipated next release dates. Check the list of anticipated next release dates on a routine basis. Be aware that not all data sources publish anticipated release dates. Furthermore, not all sources publish data on a schedule. Often data is published as it becomes available. Zillow is a classic case in point. Such sources must be individually checked at reasonable intervals which vary depending on the number of days since the last data was published. For those sources with scheduled release dates always check on the scheduled date and daily after the scheduled date until the data is released.

# Determine the Ability to Update

1. To check for newly published data,
   1. Select a specific measure table for which one expects a high likelihood of change, navigate to the web address of the Source, and download the Data Series to the **Arrivals** folder.
   2. Examine the downloaded data in the **Arrivals** folder for new or changed data. Many sources will publish notes or metadata describing the most recent changes. In other cases, the analyst will need to look for the most recent data in the downloaded file and compare that with the data previously applied to the database.
   3. If no changes are detected:
      1. Remove the file from the **Arrivals** folder
      2. Check for data impacting other Measure Tables in the list
   4. If new data is present the ability to update a data series is assured. Proceed to the next step in the process

# Apply New Data to the NCDD Database:

## Save the Incoming data file

Copy or extract selected downloaded data file from **Arrivals** to appropriate **DataSeriesBySource** subfolder. For example after downloading a file from the BEA named ‘gdpstate\_naics\_all.zip’ and determining that it contained new data not currently in the NCDD database, we extracted the two components named ‘gdpstate\_naics\_all.csv’ and ‘GDPStateNAICSFootnotes.txt’ to the folder ‘**DataSeriesBySource/BEA/\_Regional/ALL\_GDPS\_NAICS**’.

After extracting/coping the useful files from the file downloaded to **Arrivals**, document the change, and save the documentation in the same **DataSeriesBySource** subfolder. Select a name for the file containing your documentation with a name that identifies it as describing the data series being processed.

Delete the downloaded file from the **Arrivals** folder

## Load the file to the Stage Database

Copy (do not MOVE) the downloaded data file (unzipped) from the **DataSeriesBySource** subfolder to the **LoadingZone** folder. The file must be in CSV or TXT format when placed in the **LoadingZone**. If the file was downloaded in Excel (xls, or xlsx) format, load the file using Excel and save the appropriate worksheet to the **LoadingZone** folder as a TXT file.

Use the SSMS Tasks>Import Data utility to create and load a new table with an image of the downloaded data file in the STG2 database. It is appropriate to load only the columns needed, as long as the data is not altered. It is not appropriate to modify the data at this point.

Use pattern: “**STG\_<DataSource>\_<Downloaded FileName>”** for the name of the table created in the STG2 database. For example: the file ‘gdpstate\_naics\_all.csv’ in our previous example was loaded into a table named **[STG2].[dbo].[STG\_BEA\_ gdpstate\_naics\_all].** If a table with the same name exists, left from a previous effort to load an earlier release from the same source, DROP the existing table before ruing the Input Data utility. The new table will most likely consist of different columns with different column names because the latest release probably contains observations for one or more new time periods and replacement observations for a different range of previous time periods.

Because our earlier analysis discovered that only the years 2010 through 2017 were subject to change and that only the first column could contain more characters than the default 50 characters, we created the stage table with only as many columns as needed for the 8 years of data, and we set the data length of the first column to 1000, taking the default of 50 for all the other columns.

All columns should be character format. The overall approach is to load all the relevant columns exactly as received. Invalid or extraneous data will be filtered out in a later step. Also, In a later step we may transform character data to appropriate numeric formats. But for now, we want to get the data into the Stage database where we can examine it more closely with SQL and the tools available.

Delete the TXT or CSV file from the **LoadingZone** folder after successfully loading the file-image table.

## Prepare the Stage tables

### Load the ”**STG\_XLSX\_MeasureDefn\_WRK**” table

Our objective in this step is to use the insight gained from examining the file-image table along with the data in the file-image table to populate the STG\_XLSX\_MeasureDefn\_WRK table.

The STG\_XLSX\_MeasureDefn\_WRK table is a standardized interface table enabling us to use one set of SQL scripts to load Measure Definitions to the NCDD persistent data repository database independent of where the measure definition data originated or the layout of the downloaded file. All we need to do is load the Measure definitions to the STG\_XLSX\_MeasureDefn\_WRK table and the existing routines will apply the data to the appropriate Measure Definition tables in the NCDD persistent data repository. All the logic necessary to map the data from the downloaded file to the NCDD database will be contained in no more than two routines.

Furthermore, the STG\_XLSX\_MeasureDefn\_WRK provides a standard data layout that facilitates the task of updating Measure definitions. Here are the steps:

1. Examine the file-image table and build a map from data in the file-image table to a Measure Business Key. The Measure Business Key uniquely identifies a single data series. The parts of a Measure Business Key are: 1) the Measure Authority, 2) the Measure Table ID, and 3) the Measure Table Line Number. The three parts are concatenated into a single value with a single underline character between the parts.
2. Fashion a script to extract Measure definitions from the NCDD database that match records in the File-Image Data table based on Measure Business Keys that can be constructed from data in the file-image table. This script may already exist. If so, it should be saved in the SQL Server Management Studio “**MeasureDimension**” subfolder. If such a SQL script does not exist, use the SQL script “**ListMeasuresToBeUpdated\_Tamplate.sql”** as a guide to building one. Once constructed, save the script in the SQL Server Management Studio “**MeasureDimension**” subfolder with a name that can be easily related to the file-image input data table. The name of the script used in our example of the STG\_BEA\_gdpstate\_naics\_all table is “**ListMeasuresToBeUpdated\_gdpstate.sql”**
3. Open a new Excel workbook. Use the new script to populate the worksheet with the current Measure definitions:
   1. Navigate to Data>Get Data>From Database>From SQL Server database
   2. Provide the name of the SQL Server Instance
   3. Specify ACS as the Database (it is NOT optional for our purpose)
   4. Click on Advanced options
   5. Paste the new script in the SQL statement box
   6. Click on OK
   7. Click the down arrow on the Load button and select “Load To”
   8. In the dialog box select “Existing worksheet” and click “OK”. The worksheet will return with Measure definitions for Measures that correspond to the file-image table entries.
4. Using the features of Excel, update the Measure definition entries based on your reading of the release metadata and your analysis of the file-image table. Pay particular attention to the columns

***,[Source\_Citation]***

***,[Accessed\_Date]***

***,[Vintage]***

***,[Revised\_Data\_Period]***

***,[New\_Data\_Period]***

***,*** and ***[WNCD\_Notes]***

1. Keep in mind that the column headings in the worksheet correspond one-for-one with the column names and position in the STG\_XLSX\_MesureDefn\_WRK table. DO NOT CHANGE DATA IN ***[Mesure\_Business\_Key], [Record\_Source], [Measure\_Authority], [TableID]***, or ***[Table\_Code]*** as doing so will lead to unpredictable and possibly catastrophic results.
2. Save the Excel workbook in the Data\Dimensions subfolder with a name that identifies the data and the date of the update process. In our example we used the file name: “**BEA\_gdpstate\_MeasureDefn\_WRK\_20180919.xlsx**”. Delete files that are over four release cycles old.
3. Save the worksheet a second time with the same name, but as a **TXT** file to the **Data\LoadingZone** folder.
4. Load the TXT file to the STG\_XLSX\_MeasureDefn\_WRK table using the script “**LOAD\_STG\_XLSX\_MeasureDefn\_from\_TXT.sql**”. Before executing the script you will need to provide the name of the input file in the script.
5. After successfully loading the STG\_XLSX\_MeasureDefn\_WRK table, move the TXT file from the **LoadingZone** folder to the **Data\DimensionsLoaded** folder. Remove any files that are over 18 months old from the **Data\DimensionsLoaded** folder.

### Load the “STG\_SLSX\_DataSeries\_WRK” table

Our objective in this step is similar to our objective in the previous step; we want to use our insight and knowledge of the incoming data and file layout to load data from the file-image table to the “STG\_SLSX\_DataSeries\_WRK” table

The “STG\_SLSX\_DataSeries\_WRK” table provides the same function for data series observations as the “STG\_XLSX\_MeasureDefn” table provides for Measure definitions; it enables use to use the same script to load observations into the persistent data repository independent of source file format or layout. All source-specific logic is located in the script that loads the “STG\_SLSX\_DataSeries\_WRK” from the file-image table.

The advantage we have when updating an existing series is the mapping between the observations in the file-image table and the “STG\_XLSX\_DataSeries\_WRK” table was probably recorded in a script when the Measure was initially added to the NCDD database. That script would have given a name that associates it with the data source and saved in the **PostingFacts** folder. Two such scripts are: **Populate STG\_XLSX\_DataSeries\_WRK\_DynamicInsertGDPS.sql** and **Populate STG\_XLSX\_DataSeries\_WRK\_SAIPES.sql**.

Find the script or build an appropriate script using one of these two scripts as a guide.

Make any adjustments that might be needed based on an analysis of the file-image table.

Execute the script.

**Tableau**

Login: Visit the public page (this will change once we have two workbooks -> one for public and one for us admins that will show all workbooks such as the testing ones: <https://public.tableau.com/profile/wnced#!/>

* + 1. Click the top right “sign in” button
       1. Username: [WNCED@email.wcu.edu](mailto:WNCED@email.wcu.edu)
       2. Password: wnced$%
    2. This login will allow you to make physical appearance edits and NOT changes in the data.

Updating a workbook:

1. Hover over one and click download (create a folder in your file directory to save the workbook to)
2. Double click and workbook opens in Tableau (if you’ve properly installed the software)
3. Make the needed changes (double click on text to open the edit window in Tableau)
4. When ready to publish after all changes are made:
   1. Click the save icon, click the overview tab in the bottom left, click server at the top toolbar -> mouse down to publish workbook.
   2. When publishing -> make sure the names all match.
   3. Make sure you publish only to the ones that say “WNCED”

Default Tableau Sheet Format:

1. Centered title
2. Brief instructions and year selector at top left corner (year selectors should not have the show all value checked and to verify; the year selector wouldn’t end but rather say “(all)”)
   1. To fix the above year selector: click the date selector panel -> click more options -> expand “customize” -> click show all value and view the options.
3. Map visual just under the year selector
4. Filter selectors should be in treemap format on the left side of each dashboard.
5. All graphs will be to the right side of the dashboard