Requesting Access and Loading Data

for the Data & Analytics for Good journal

(3/29/2023)

**Sources**

Ron Freeze, Michael Gibbs, Odessa Elie, Tal Miranda, Samuel Kreinberg

Enterprise Systems, Sam M. Walton College of Business, University of Arkansas, Fayetteville

Teradata Viewpoint 16.50.01.00-b710

Copyright © 2020 *For educational uses only - adapted from sources with permission. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission from the author/presenter.*

# Use Case – Data & Analytics for Good journal

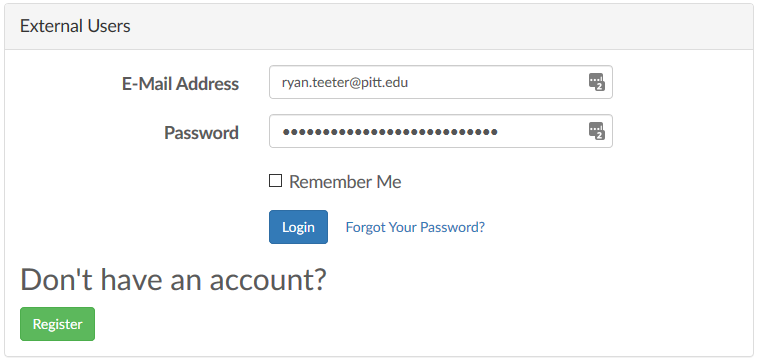
The purpose of this document is to demonstrate the processes of requesting accounts and loading data for articles accepted in the Data & Analytics for Good journal. These steps require access to the University of Arkansas applications: VMware and Teradata Studio Express (also referred to as Teradata Studio in this document).

Journal authors are allotted permanent personal space on the Teradata Intelliflex system to create their own databases, views, and more. Authors are also granted additional permissions to view prior journal article databases.

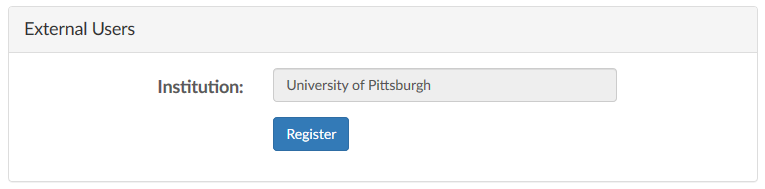
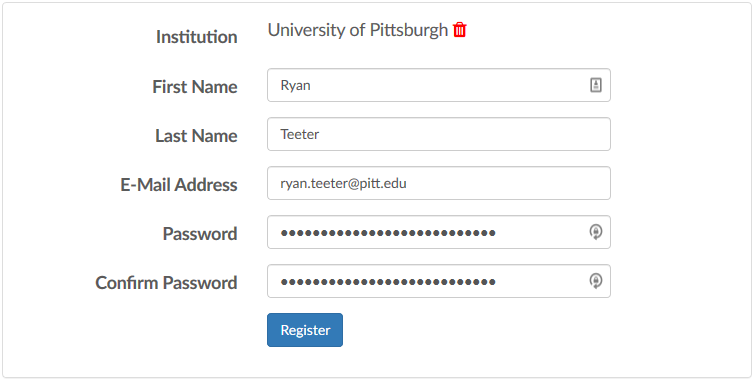
Authors who have additional questions concerning creating accounts, accessing enterprise data sets, and general questions about the Teradata Intelliflex system at the University of Arkansas should contact Ron Freeze or Michael Gibbs.

1. Ron Freeze at [rfreeze@walton.uark.edu](mailto:rfreeze@walton.uark.edu)
2. Michael Gibbs at [mgibbs@walton.uark.edu](mailto:mgibbs@walton.uark.edu)

## Step 1: Request Access to Teradata System

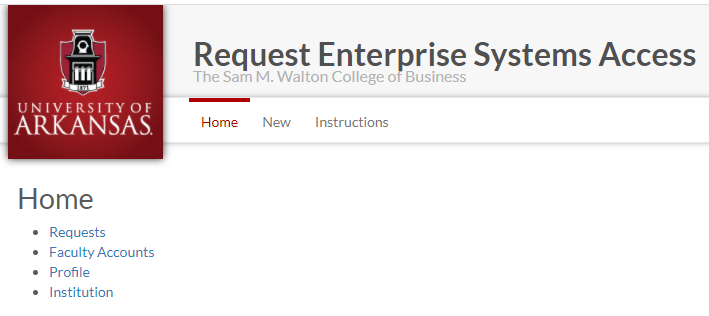
1. Go to <https://request.information-systems.uark.edu/login>.
2. Enter your e-mail address and password and click **Login.**

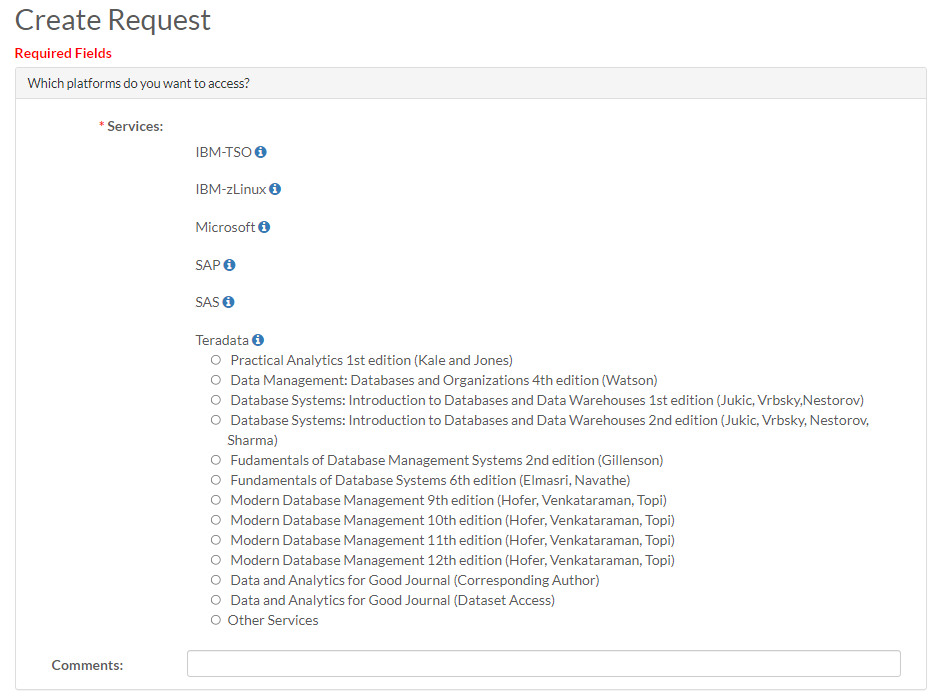
If you don’t already have an enterprise systems account, register for an account at the same place

1. Go to <https://request.information-systems.uark.edu/login>.
2. Click Register.
3. Type the name of your institution and click **Register.**
4. Type in your first name, last name, e-mail address, and password, and click **Register.**

You only need to create an author account once. If you plan on submitting multiple datasets, you can continue to use your prior account to load your data for the new acceptances.

## Step 2: Request Access to the Journal databases

Walton Enterprise Systems provides support for multiple authors and their datasets as well as access to many applications used in industry. This example outlines the process for requesting space and access for the data sets being shared with the Data & Analytics for Good journals.

1. Log in to <https://request.information-systems.uark.edu/login>.
2. Click “New” on the Home screen.
3. On the Create Request screen, under “Which platforms do you want to access?”, select the Teradata platform

There are two selections for the Data & Analytics for Good journal. The first is for the corresponding author (or co-authors). Selecting this will allow you to upload your data to the appropriate Issue and Article #. If you are the corresponding author, the following information will be requested.

* Corresponding Author
* Accepting Senior Editor
* Manuscript Titles
* Journal Issue
* Article Acceptance Number

The second selection is for those wishing read access to all datasets in the Journal.

1. Select appropriate Data & Analytics for Good Journal request.
2. For “How many new accounts would you like?”, there is no need to indicate anything for new accounts unless you are getting accounts for students where they will be using the DAG datasets.
3. In the “Anything else you would like us to know?” box, enter additional information that could help provide context behind your request.
4. Read the agreement, and check **I agree.**
5. Click **Submit** to finalize your request.

Note: you won’t receive confirmation of account creation until the accounts are created. This includes the Teradata databases that will be created for you.

## Step 3: Wait for the e-mail

A systems specialist at the University of Arkansas will manually review your request and create the account usernames and passwords requested. Account creation should be completed within 2 business days. NOTE: ATTENTION (Spring 2023): New requests will take approximately 10-15 business days to process and fulfill. We apologize for any inconvenience.

You will receive an e-mail with your login username and password, similar to the image on the right:

Once your author account is created, you may continue with the Step 4: Using VMware in order to log into the Walton Enterprise System.

## Step 4: Using VMware

In order to upload your dataset, you will need to use Teradata Studio which is accessible from the Walton Enterprise System virtual desktop – VMware. You can install VMware on your local desktop or use the browser version to access your virtual desktop. Either the downloaded VMware or browser version are acceptable and provide you with the same access.

1. Refer to the document “[VMware Guide via Windows Client](https://walton.uark.edu/enterprise/VMwareGuideWindowsClient_Updated_07.02.pdf)” to view the tutorial on how to download, install, and use VMware on your computer.
2. Refer to the document “[VMware Guide via Browser](https://walton.uark.edu/enterprise/VMwareGuideBrowser_Updated_07.02-converted.pdf)” to view the tutorial on how to access VMware through your browser.
3. Access the site directly, here: <https://waltonlab.uark.edu/>

## Step 5: Data Loading verification

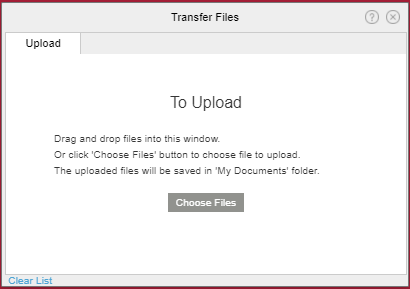
This step provides a short check list for the guidelines you will use in Step 6 & Step 7. Once you have uploaded all your tables, the following bullet points are the data loading verification that will be performed by the Data & Analytics for Good reviewers prior to final acceptance of the article.

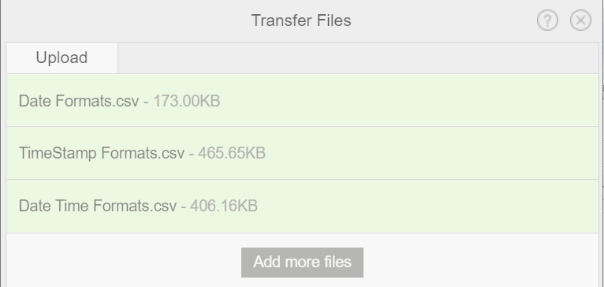
* Do the table names uploaded to Teradata match the table names in the:
  + Data Dictionary
  + Entity Relationship Diagram
  + Other locations in the article?
  + Column Headers in the CSV or Excel File
* Do the data types in the Teradata database match:
  + The variables in the Entity Relationship Diagram
  + The variables loaded in the Data Dictionary?

With concurrence of data load, final acceptance of the article will be issued.

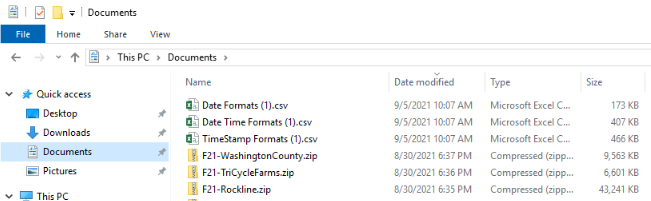
## Step 6: Transferring Files to VMWare

For the ease of loading your .csv data files to Teradata, it is recommended that you first transfer your files to either your VMWare desktop or your folder in the S: drive that is accessible in VMWare. This step shows how to transfer files to VMWare. For this example, you will need to make your file(s) available to the VMWare desktop.

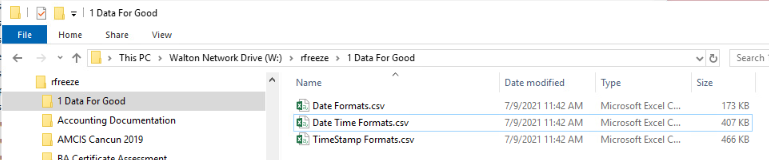
1. Expand the Navigation pane for your Desktop by selecting the tab located in the middle of your screen on the left side
2. Open the Transfer Files pop-up by selecting the icon
3. Select the “Choose Files” on the Transfer Files pop-up and navigate to the directory on your computer to select the files to be uploaded to the tables in your database.



1. Select each of the files you have created for uploading to Teradata. The screen shot shows three files (tables) moved to VMWare that are to be uploaded to Teradata. These files were for testing the Date, TimeStamp and Date Time formats. (See the section on Checking Date//Time formats.)



The files show in the Transfer Files pop-up will be in your VMWare Documents directory for selection by Teradata. You could also place them in your personal S: drive folder or on your VMWare Desktop. Screen shots for all three locations have been provided.



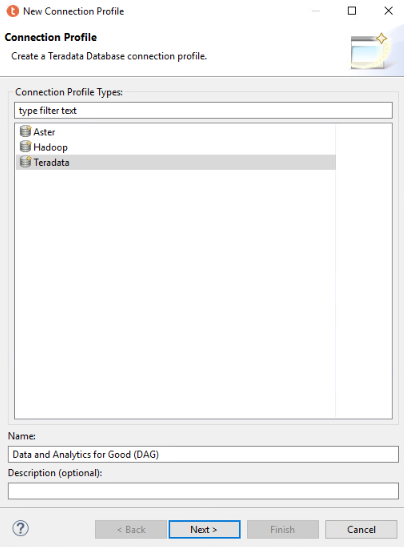
You have now loaded your data to VMWare and are ready to load your data to the Teradata tables.

## Step 7: Loading into Teradata Studio

Teradata Studio is a more robust application used to access the Teradata IntelliFlex system. Teradata Studio will be used to upload your dataset.

### 7.1 Accessing

1. On your desktop, you may see the Teradata Studio icon. If so, double-click the Teradata Studio icon. If not, click the Microsoft Start button in the lower left hand corner and select the icon there.

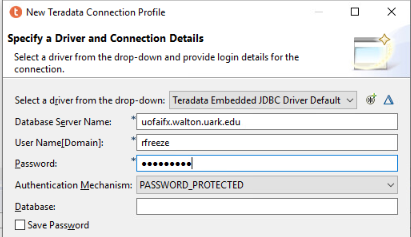
Teradata Studio will open. If you have logged in before, Teradata Studio may prompt you for your password to the University of Arkansas IntelliFlex system. Please enter your provided password. If you have not logged in before, you will be prompted with a Quick Tour which you may go through if you wish. After clicking out of the Quick Tour, you will be prompted to create a connection profile. NOTE: If you are not prompted to create a conncection profile, then select the Query Development icon in the upper-right hand side of Teradata Studio, and then you will need to right click on Database Connections and select ”new.”

1. Select “Teradata” from the Connection Profile window
2. Place the database name for your connection as “Data and Analytics for Good (DAG)”
3. Click “Next”

After clicking next, you will be prompted with a new window, “Specify a Driver and Connection Details”. Specify the drive and connection details by entering in the database server name, and your username and password given to you by the University of Arkansas.

**Database Server Name:** uofaifx.walton.uark.edu

**User Name (Domain):** (your username)

**Password:** (your password)

You may leave the rest of the dialog boxes blank.

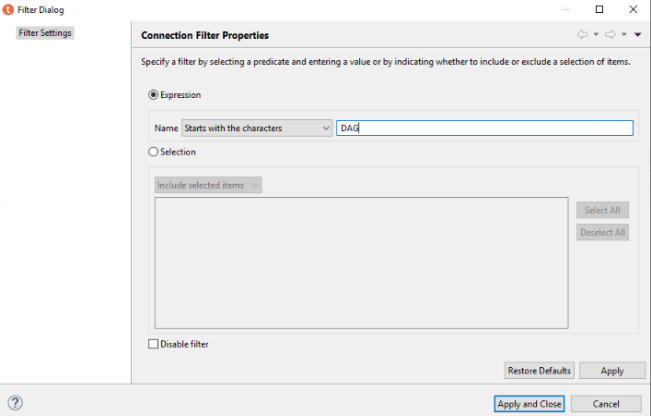
1. Select “Finish”
2. Please note that you may need to re-establish this connection with each new Teradata session.

### 7.2 Viewing Datasets

This section explains how to navigate Teradata Studio and view your dataset and all available datasets.

1. Access the Data Source Explorer by selecting the Query Development icon in the upper-right hand side of Teradata Studio:

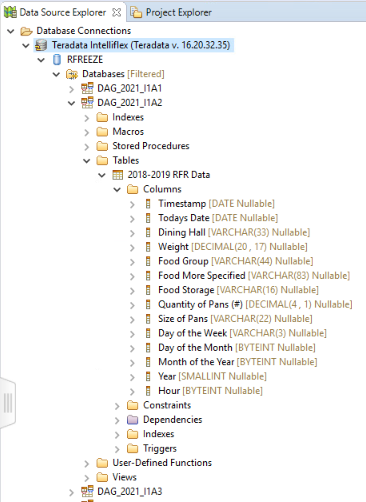
Data Source Explorer shows you the databases and Foreign Servers you can view in the Database Connections folder. This is where you can explore which databases are available to you and expand the Data and Analytics for Good (DAG) (Teradata v 17.10,3.07) icon to see the databases.

Due to the large number of databases, you will need to filter to only those databases for which you are interested. Since you will be uploading the dataset associated with your accepted article in the journal, you will need to search for only those databases associated with your article. These have already been set up for ease of uploading.

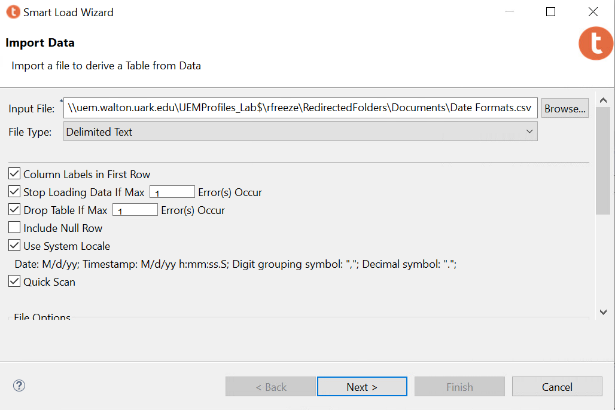
1. Click the arrow to the left of Data and Analytics for Good (DAG) icon to expand
2. Click the arrow to the left of your username icon to expand
3. Right click on the Databases folder
4. Select Filter

A Filter Dialog pop-up will appear.

1. Deselect “Disable Filter” at the bottom
2. Enter “DAG” in the Expression
3. Select “Apply and Close”

You can use the same process to select other databases for which you are able to view the data. The example shows the first issue of the Data & Analytics for Good journal. Additionally, Article 2’s (A2) table and variables have been displayed. The data types of the Teradata Intelliflex table should match the data dictionary provided in your article. As new articles are accepted, you will have read access to the new data.

### 7.3 Checking Date/Time formats

Data types are critical to communicating through analytics. One of the most frequent issues is with the uploading of date and time formats. The following highlights the uploading of date and time formats associated with the private database of rfreeze. There are two important aspects of uploading your data that will enable our community to replicate and expand on your initial article. The first aspect is ensuring correct Date, Date-Time and Timestamp formatting in your CSV file. In this example, the CSV file was created in Excel and saved appropriately for uploading to Teradata. The examples provided are from a CSV file named “Date Formats.CSV”, “TimeStamp Formats.CSV” and “Date Time Formats.CSV”. Please replicate this section with your own .CSV file in order to practice your data load from your article. The Date Formats section shows the process of uploading the Date Formats.CSV file. The next two sections provide the results of uploading the TimeStamp Formats.CSV and Date Time Formats.CSV files. For this example, the Date Formats.CSV file is being uploaded to the personal rfreeze database. An example of this formatting has been provided. All three examples will be uploaded to this location.

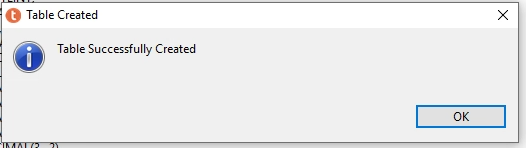
1. Select the Data Transfer icon in the upper-right hand side of Teradata Studio
2. Click the arrow to the left of the appropriate DAG database icon to expand
3. Right click the Tables folder in Teradata
4. Select Teradata
5. Select Data Transfer

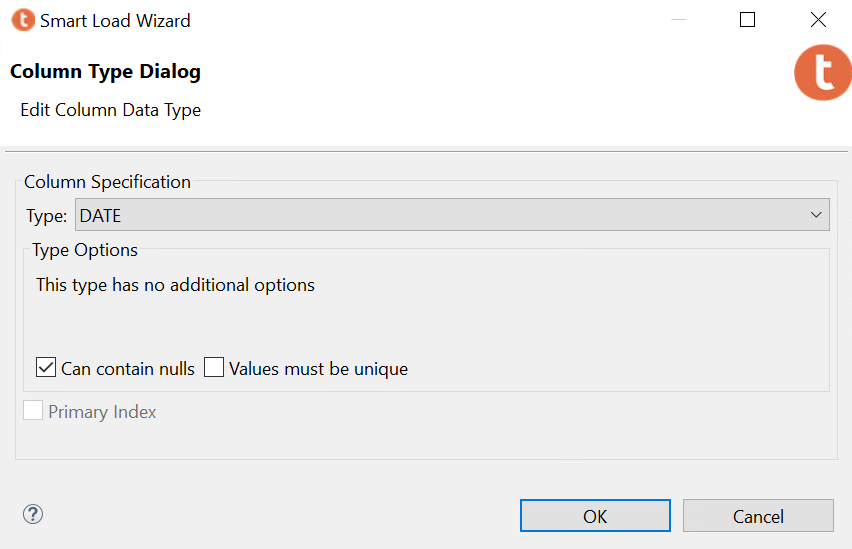
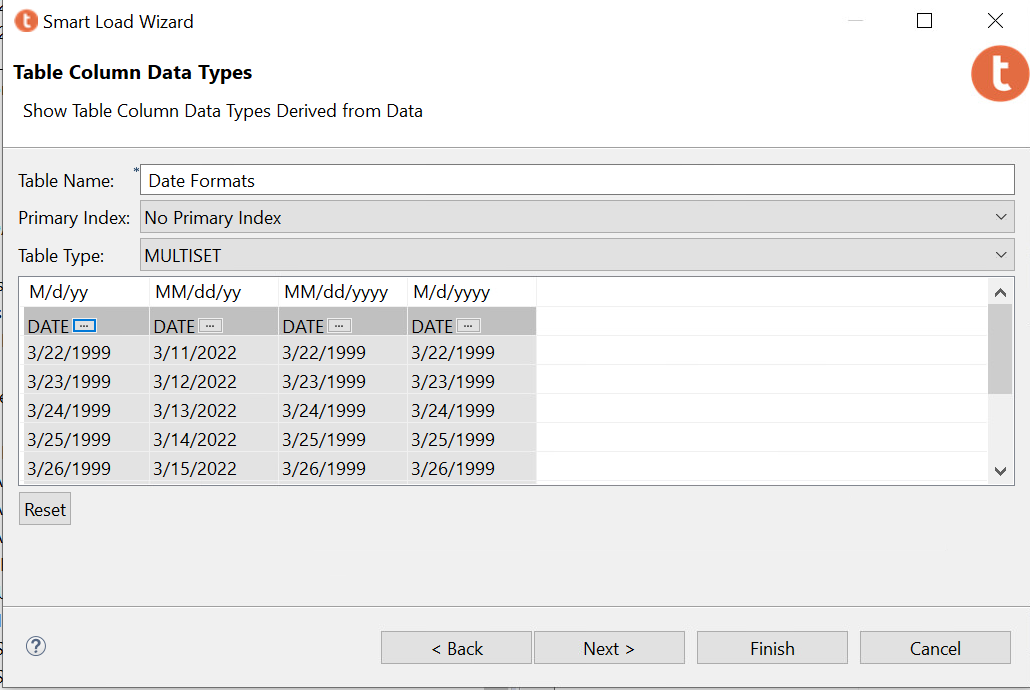
The Data Transfer Wizard pop-up will appear.

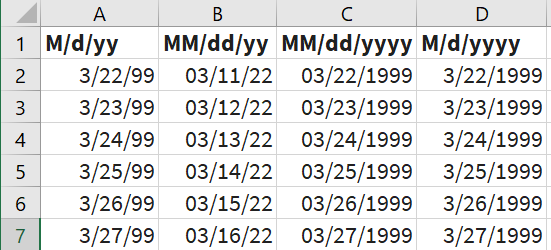
1. In the Data Transfer Wizard, select “External File (Smart Load) as the Source Type and then click Launch.

The File Type for all Data & Analytics for Good data sets should be a Comma Separated Text (.CSV file). Be sure to select “Delimited Text” as the File Type. The example uses the “Date Formats.csv” file.

1. Click on the “Browse” button to locate the file you would like to import. Your file should be either on your Desktop or in your S: drive folder as laid out in Step 6.
2. Select Next

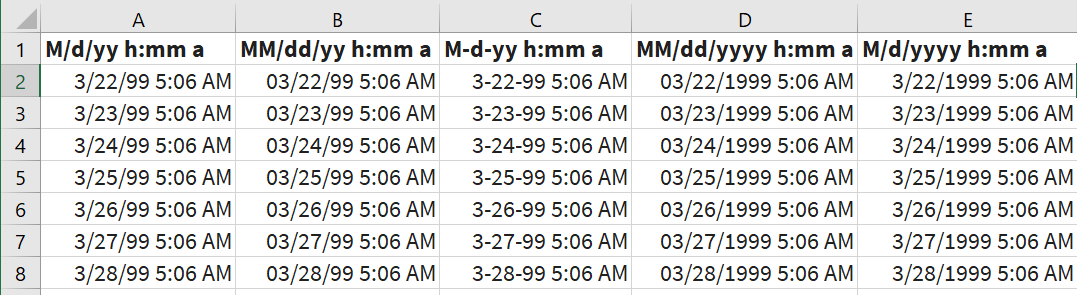
The next pop-up will allow you to verify and change any of the data types. In this case, the formatting has already been identified as a DATE and is the correct data type. To change the data type, you can select the three dots to the right of each variables data type. NOTE: This is also the pop-up where you can select the variables that are the primary index (primary keys) for your dataset

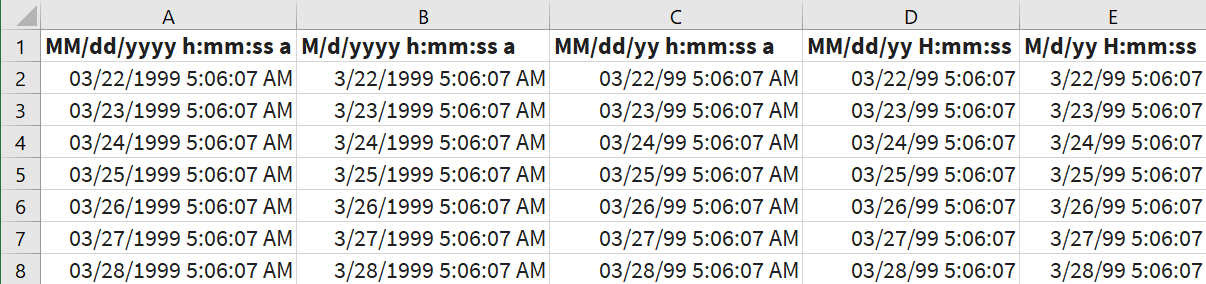


1. Select Finish once each of your varaiables have been validated.
2. You should receive a pop-up indicating that your Table was successfully created.

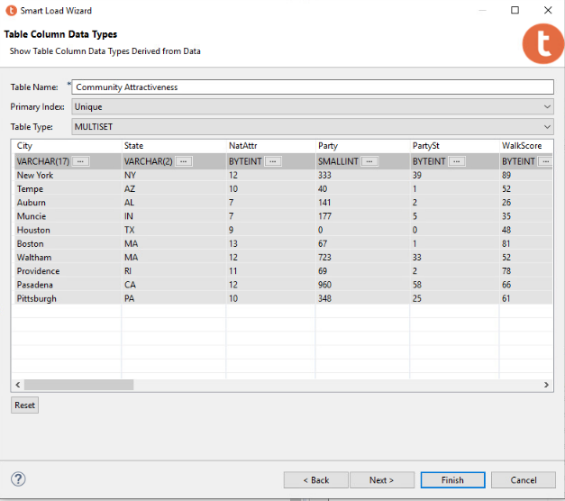
This process was duplicated for the indicated TimeStamp and Date Time Format files. The CSV formatting representations in Excel are shown below. NOTE: the formatting in the CSV file is critical for Teradata to correctly identify the variable data type for both dates and times.

**TimeStamp Formats**

******

**Date Time Formats**

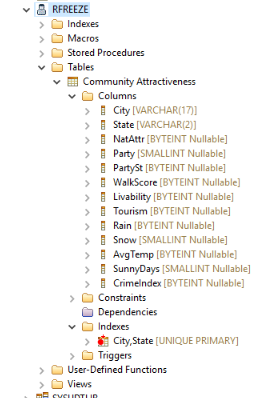
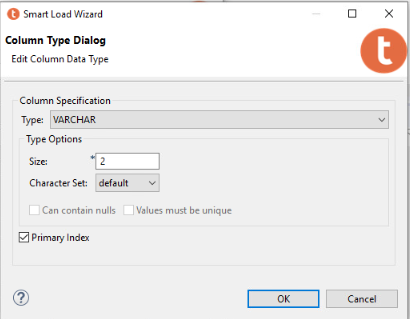
### 7.4 Setting Primary and Composite Keys



The second aspect is to ensure that your tables have been correctly identified with the appropriate keys and match the ERD that has been provided with your data set. For this example, the table “Community Attractiveness” from I1A3 article is used. The table “Community Attractiveness” has a composite key consisting of the two variables: “City” and “State”. When the table is loaded, there are two places to indicate the Primary Key (Composite Key) for the database. The first place to indicate that there is a Primary Index is in the second pop-up when you are using the Teradata Smart Load Wizard. At the top in the Primary Index: box, change the selection to “Unique”. This indicates that this table has a Unique primary key.

The second place is within the named variables (City and State) for that table. Select the three horizontal dots that are next to each variable in order to make appropriate selections. You will need to go to each of the variables to make these selections. Screenshots of these points have been provided.

Once all of your keys have been designated, you will need to hit NEXT to get a preview of the Generated SQL. Select the “Show in Editor” and select OK when the pop-up indicates you are switching to editor. **NOTE**: for this version of Teradata, you must go to the editor and run the SQL code from here in order to continue loading other tables.



Run the SQL that is in the editor. The table (Community Attractiveness) will load and you can see the City,State UNIQUE PRIMARY key when you open the Indexes. NOTE: this does not load the data. You will need to go to the section on Uploading Your Data to complete the process.

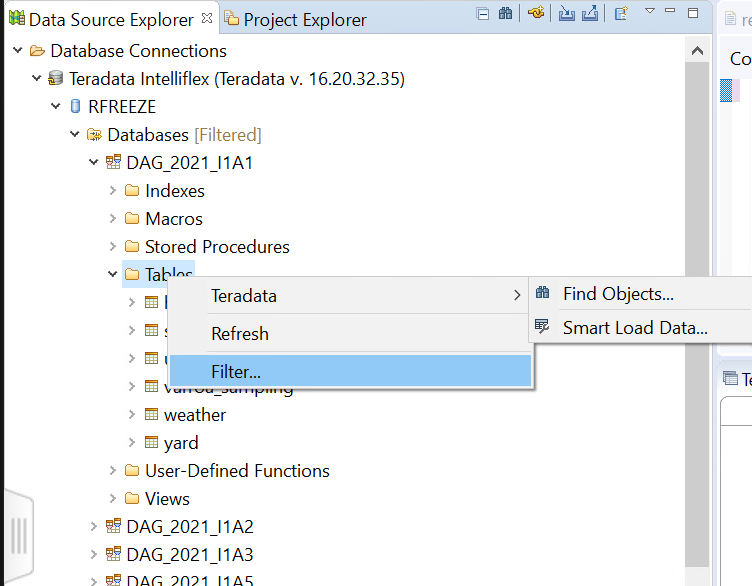
### 7.5 SQL

The authors of the Issue and Article acceptance will be responsible for uploading the data to the appropriate tables. During this process you may need to delete tables that were not properly uploaded so you can reload them correctly. If this is the case, the Teradata SQL statement needed to remove your tables are as follows. The first example is an actual statement with the database name and the table name. The second example provides you with the generic form of the SQL statement needed to remove your table. You should replace the “Database Name” with the appropriate Data & Analytics for Good Year-Issue-Article that you were given when the article was accepted. Only the authors are allowed to create and remove tables from their assigned database.

**Drop Table** “DAG\_2021\_I1A2”.”2018-2019 RFR Data2”;

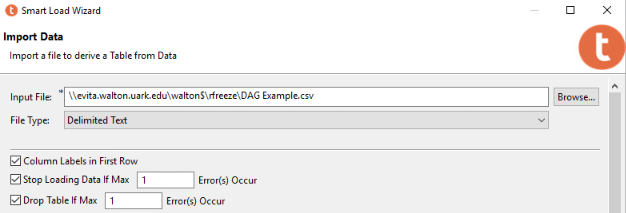
**Drop Table** “Database Name”.”Table Name”;

### 7.6 Uploading Your Data

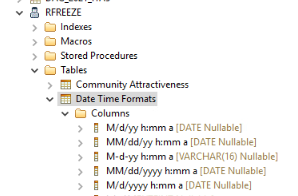
Each author requesting access has been provided with a private database area for testing the upload of the data associated with your article. This testing allows you to Read, Write and Create tables in your database. **NOTE:** you must have already been given the issue number and article number in which your article will appear prior to uploading your data. You can request access to Teradata to view datasets. You can practice in your private database area prior to article acceptance. The following steps are for uploading your article’s data.

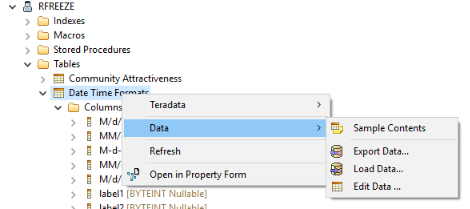
1. Select the Data Transfer icon in the upper-right hand side of Teradata Studio
2. Expand the appropriate location for your accepted article – DAG\_YYYY\_InAn
   1. YYYY – the year of your articles publication
   2. InAn – Issue number Article number
3. Click the arrow to the left of the Tables icon to expand
4. Right click the appropriate table that you loaded in Step 7.4Select DataSelect Load Data to open the Data Transfer WizardIn the Data Transfer Wizard, select “External File (FastLoad)” as the Source Type and then click Launch

The Load Data Wizard pop-up will appear.

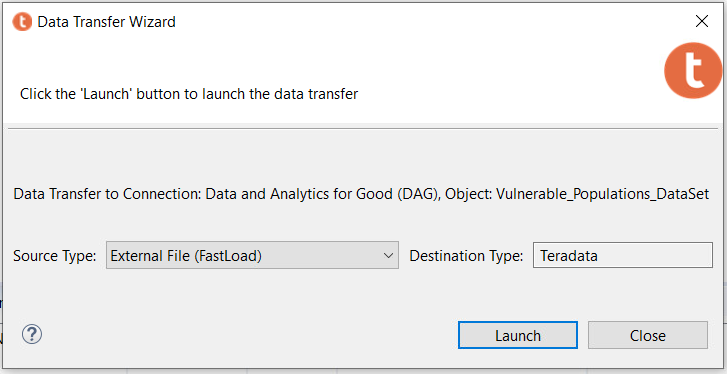
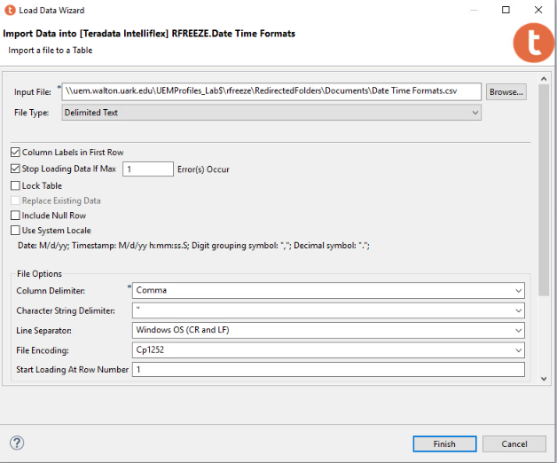
1. In the Data Load Wizard, click on the “Browse” button to locate the file you would like to import.
2. The File Type should be Delimited Text (a .csv file). All Data & Analytics for Good data sets should be imported from a .csv file. This is especially true if one of your tables will be exceeding the approximately 1 million records that excel files are limited to. The example uses either a file named “DAG Example.csv or the Date Time Formats.csv”. **NOTE**: Excel files may be needed for more appropriate data type specification.
3. Make sure your file has column labels in the first row and check the box for “Column Labels in First Row”.
4. Click Next

The next pop-up allows you to check the variable types along with your labels. This procedure was covered in the [Checking Date/Time formats](#_Checking_Date/Time_formats) and [Setting Primary Keys](#_Setting_Primary_Keys) sections above.

1. Click NEXT to get a preview of the Generated SQL.
2. Select “Show in Editor” and select OK when the pop-up indicates you are switching to editor. **NOTE**: for this version of Teradata, you must go to the editor and run the SQL code from there in order to continue loading other tables. **The script created will only create the data structure**.
3. Run the SQL that is in the editor. This will create the data structure to receive your data

If you successfully created the data structure, you should “Refresh” your database and open the Tables folder. To verify that the table was created, expand the “Tables” option inside of the database. In this case, we created the Table inside of the “rfreeze” user, so we expand the “Tables” folder and see that the Date Time Formats table was created. Expanding the “Columns” subfolder also indicates which columns were successfully imported from the file and whether all the columns were imported as the correct data type.

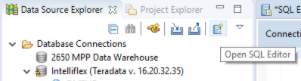
The next steps will be to **load the contents of the file into the data structure created**.

1. Right click on the table you just created and navigate to the Load Data selection
2. The Load Data Wizard pop-up will appear. Reselect the file you used to create the data structure using “External File (FastLoad)” as the Source Type. Use the setting as specficied earlier:
3. Select Finish

### 7.7 Verify Data Load

If you successfully imported the data, you should “Refresh” your database and open the Tables folder and Verify that the contents of the file were loaded.

1. Click the “Open SQL Editor” button at the top of the Data Source Explorer
2. When the editor opens on the right, type the query “select \* from rfreeze.<table name>;”.
3. Click the “Execute button” to run the query



You may have to select “No” when asked whether you wish to stop at the default 2000 records for viewing. The Teradata Result Set Viewer should show the result of the query at the bottom of the page, which should be the full contents of the table selected.