# Dillards 2016 Data Extraction

## Teradata Connection

This tutorial assumes you have access to the University of Arkansas VMWare client or the Remote desktop client for participating universities (access can be requested from the University of Arkansas at https://walton.uark.edu/enterprise/. The intent of this exercise is to extract KPI data from the Dillard's 2016 dataset to an Excel file in order to load the information into an application of your choice. Questions can be directed to Ron Freeze at rfreeze@walton.uark.edu.

NOTE: the data used during this workshop should not be downloaded to your personal drives and should remain on the Remote Desktop S: drive provided by the University of Arkansas. This is due to our agreement with the data providers.
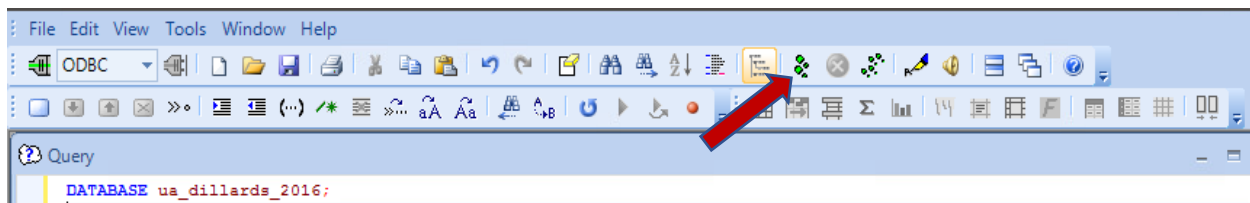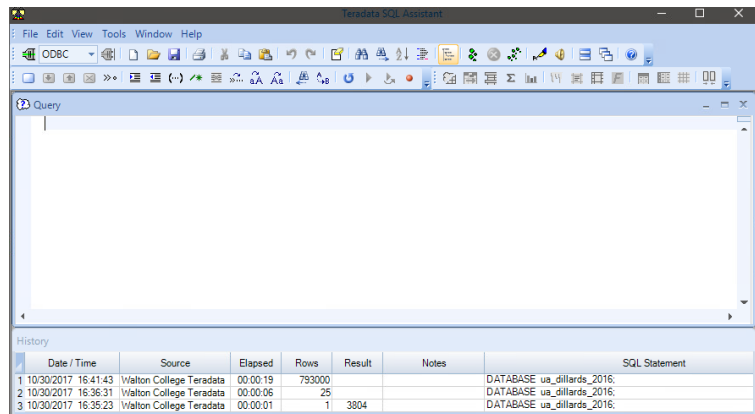
1. From the Desktop or Start Menu, search for and open the software Teradata SQL Assistant. You will see a Windows similar to the below screenshot.
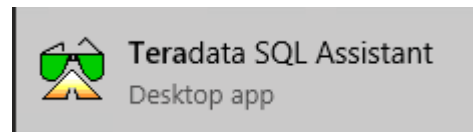
   

2. Copy and paste the query below onto the Query window to access the UA_Dillards_2016 and click on the "footsteps" execute button
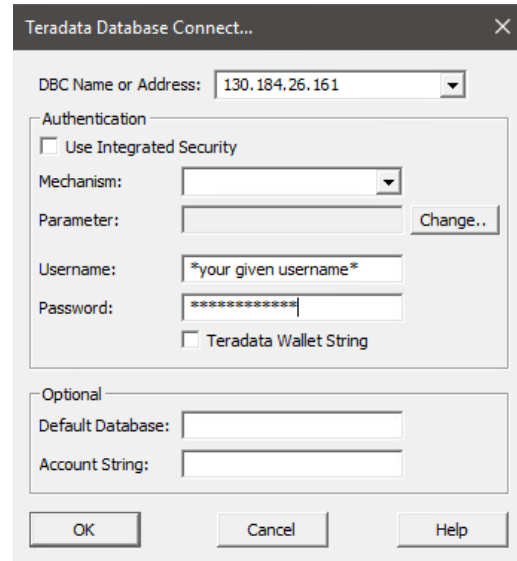
   

   **DATABASE** ua_dillards_2016;

   

   

3. A Select Data Source window will pop up to ask you the credentials to access the Dillard's database. Navigate to the **Machine Data Source** tab on top. Then double-click on **Walton College Teradata** (as shown to the right).

# Dillards 2016 Data Extraction

4.  Click **OK**

5.  Here you will see another pop-up window called **Teradata Database Connect…**

6.  Enter the **Username**: and **Password**: provided by your instructor. Click **OK**

7.  Congratulations! Now you are connected to the UA_Dillards_2016 database and ready to extract the data needed for your assignment.

## SQL Query to extract data

The Dillard's 2016 dataset has approximately 450 million rows. We need to narrow the data extraction down to only the KPI's we are interested in reviewing. To do this, we will run a query to extract transaction data from the TRANSACT table and associated identifying information (Store, City, State & Zip_Code) from the STORE table.

8.  Running the Query
    a.  Navigate to the query window and delete any current code
    b.  Copy and paste the SQL query below into the query window

```sql
SELECT      tran_date, Store.City, Store.Zip_code, Store.State,
            Store.Store, SUM(tran_amt) AS amount
FROM  Transact
INNER JOIN store
      ON transact.store = store.store
WHERE tran_type = 'p'
GROUP BY tran_date, city, State, Zip_Code, store.store
ORDER BY tran_date;
```

9.  Click the "footsteps" execute button           on the menu bar.

10. The final query should take **approximately 30 seconds to 1 minute to run** depending on the connection from your computer. You should have 298,516 rows of records in the final selection. Your result table should look like the picture below and the records can be verified on this screen.

# Dillards 2016 Data Extraction

## Export result table in a text file

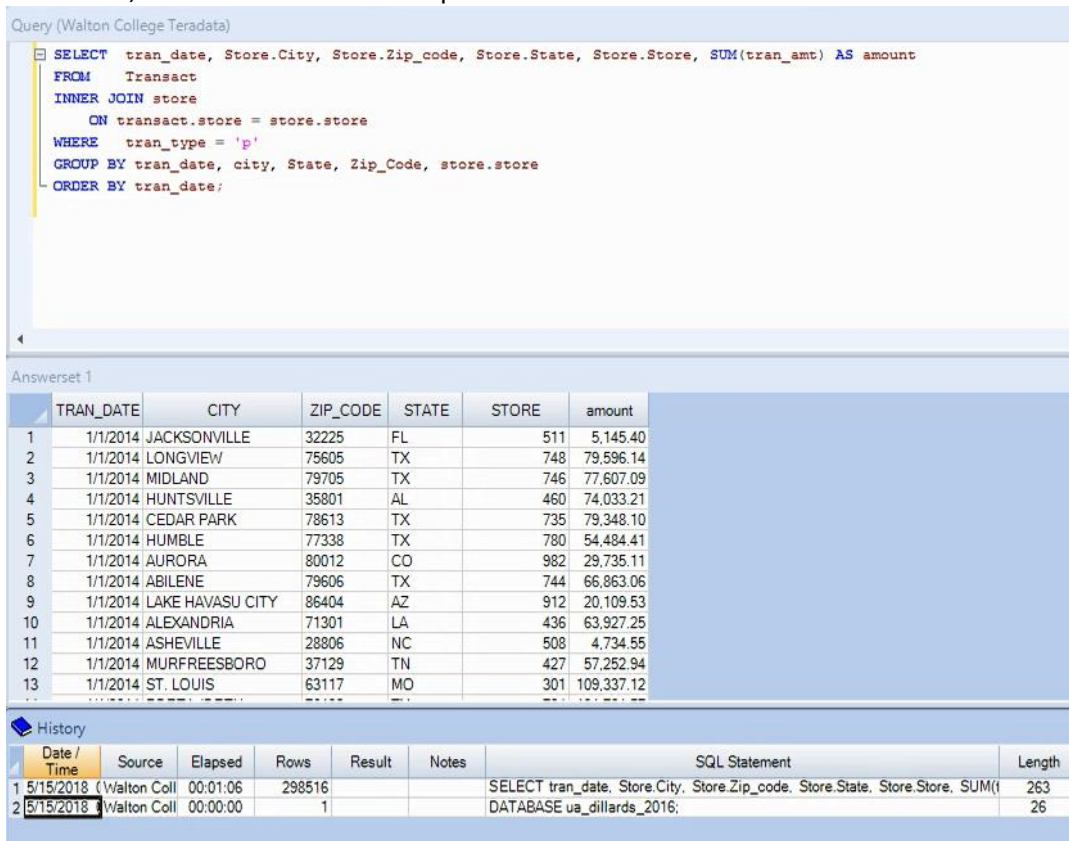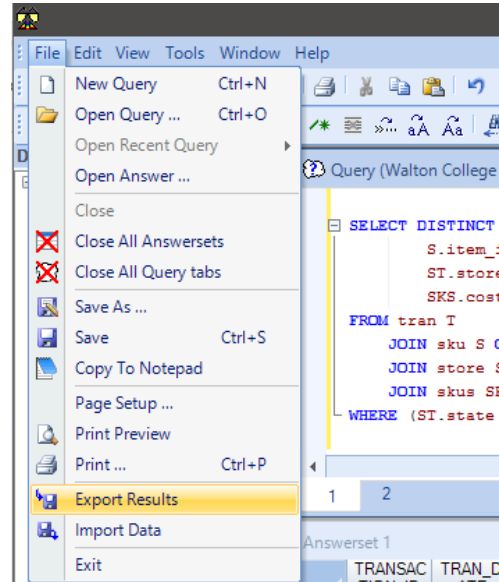Now we will export to a file in order to provide access to the application of our choice

11. Navigate to File on the top bar, and select Export Results…

12. Now click the "footsteps" [footsteps icon] execute button again to have Teradata run the query again and export the result directly.

13. You will now see a pop up window, prompting you for the location you want to save

14. Select your desired location, the File name: and make sure you Save as type: Delimited Text [ANSI] (*.txt).

15. **IMPORTANT**: after you click Save, Teradata will take a while to export everything. It will take about 3 minute or less depending on your computer, since there are 298,516 rows of data to be exported.

# Dillards 2016 Data Extraction

16. After Teradata finishes saving the result table as txt file in your folder, go to the location that you saved it to make sure it exports successfully. In this example, the file has been named DillardsTUN-Export-KPIs.txt

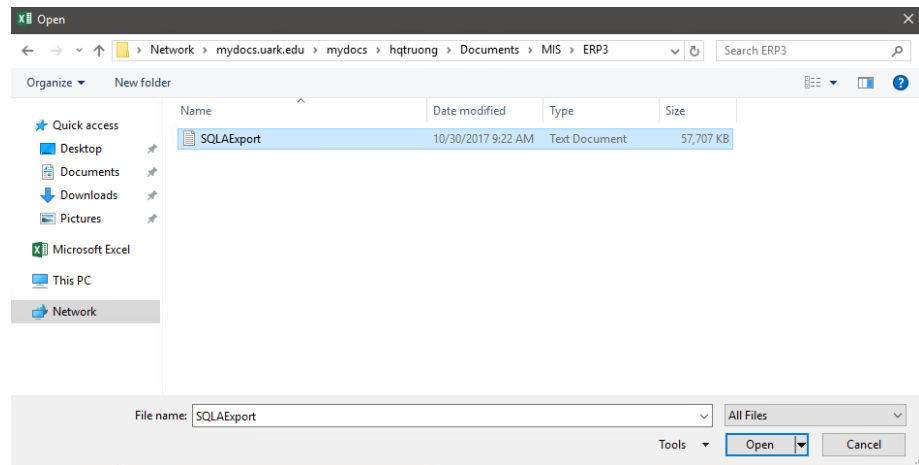| Name | Date modified | Type | Size |
|------|---------------|------|------|
| DillardsTUN-Export-KPIs | 5/15/2018 8:23 AM | Text Document | 12,267 KB |

## Import Text file into Excel

17. Open MS Excel.

18. Select **File** -> **Open** and Navigate to your .txt file folder

19. Choose the file format as **All Files** (as shown below)

20. Select your file

21. Click **Open**.

22. You will now see a window called **Text Import Wizard**, with different options of how you want to format your Excel file from a text file.

23. Accept the default settings and click **Finish** to import the file.

24. After couple seconds, you will now see the result is successfully imported into Excel. See next page for a screen shot.

25. **Save** the Excel file after you are done reviewing the data as a **.csv** comma delimited excel file
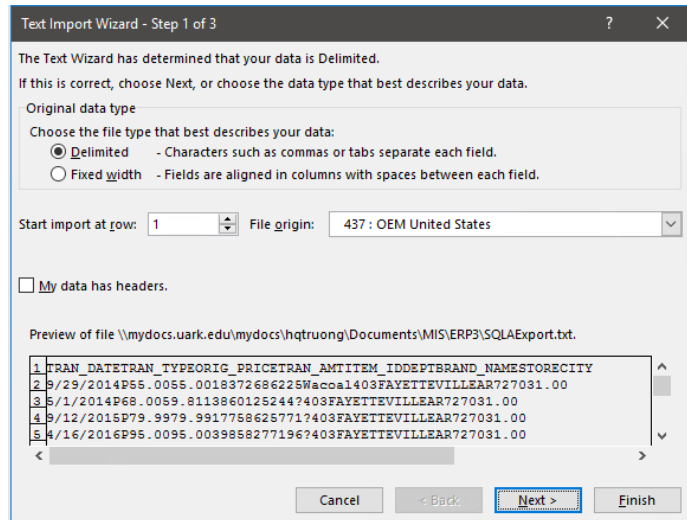
# Dillards 2016 Data Extraction

DillardsTUN-Export-KPIs - Excel

| | | File | Home | Insert | Page Layout | Formulas | Data | Review | View | ACROBAT | Tell me what you want to do... |

A1    TRAN_DATE

| | A | B | C | D | E | F | G | H | I | J | K | L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | TRAN_DATE | CITY | ZIP_CODE | STATE | STORE | amount | | | | | | |
| 2 | 1/1/2014 | JACKSONVILLE | 32225 | FL | 511 | 5145.4 | | | | | | |
| 3 | 1/1/2014 | LONGVIEW | 75605 | TX | 748 | 79596.14 | | | | | | |
| 4 | 1/1/2014 | HUNTSVILLE | 35801 | AL | 460 | 74033.21 | | | | | | |
| 5 | 1/1/2014 | MIDLAND | 79705 | TX | 746 | 77607.09 | | | | | | |
| 6 | 1/1/2014 | CEDAR PARK | 78613 | TX | 735 | 79348.1 | | | | | | |
| 7 | 1/1/2014 | HUMBLE | 77338 | TX | 780 | 54484.41 | | | | | | |
| 8 | 1/1/2014 | AURORA | 80012 | CO | 982 | 29735.11 | | | | | | |
| 9 | 1/1/2014 | ABILENE | 79606 | TX | 744 | 66863.06 | | | | | | |
| 10 | 1/1/2014 | LAKE HAVASU CITY | 86404 | AZ | 912 | 20109.53 | | | | | | |