Campaign Management

Execution with SQLR Scripts

For the purposes of a quick demo, we can to use a small dataset. To create a smaller dataset follow the steps in [**Data Setup**](Data%20Setup.docx)**.docx**.

Make sure you have set up your SQL Server and ODBC connection between SQL and PowerBI by following the instructions in [**START HERE.docx**](START%20HERE.docx). Then proceed with the steps below to run the solution template using the SQLR files.

Running these .sql scripts will walk through the operationalized steps of this solution – dataset creation, modeling, and scoring as described in the [data scientist page](https://github.com/Microsoft/SQL-Server-R-Services-Samples/tree/master/CampaignManagement/CampaignManagement/Resources/data-scientist.md) on GitHub.

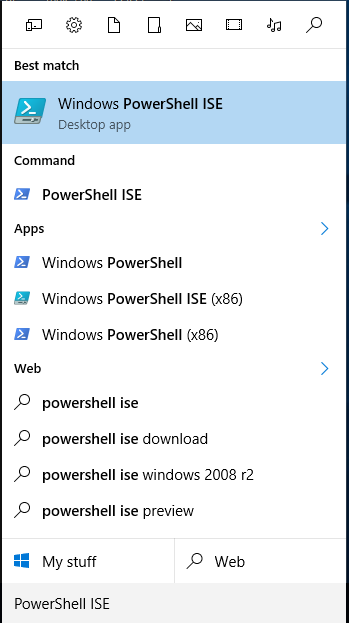
Solution Path: SQLR Scripts

Let’s start in SQL Server Management Studio where we have connected to the SQL 2016 server.

*Login to SSMS using the credential created in the Set up Instructions document.*

The first step we need to do is import the raw datasets into the SQL Server. In this solution we will use PowerShell to import the raw datasets into SQL Server.

1. *Click on the windows key on your keyboard. Type the words ‘PowerShell ISE’ and open the Windows PowerShell ISE app.*

**

*In the Powershell ISE command window, type the following command:*

Set-ExecutionPolicy Unrestricted -Scope Process

*Answer “Y” to the prompt to allow the following scripts to execute.*

1. *Click on File on the top left corner of the screen. Then click on Open and navigate to the folder location where you unzipped the CampaignManagement.zip file and open ‘Data Import.ps1’*
2. *Press F5 to run the PowerShell script*

You will get a warning here saying that you should only run scripts you trust

1. *Hit ‘Run Once’*

The command line window will prompt you to enter the Server Name

1. *Enter your Machine Name. If you do not know your Machine Name follow instructions from the Set up Instructions part of the document*

The command line window will prompt you to enter the Database Name

1. *Enter ‘CampaignManagement’*

The command line window will prompt you to enter the Schema Name

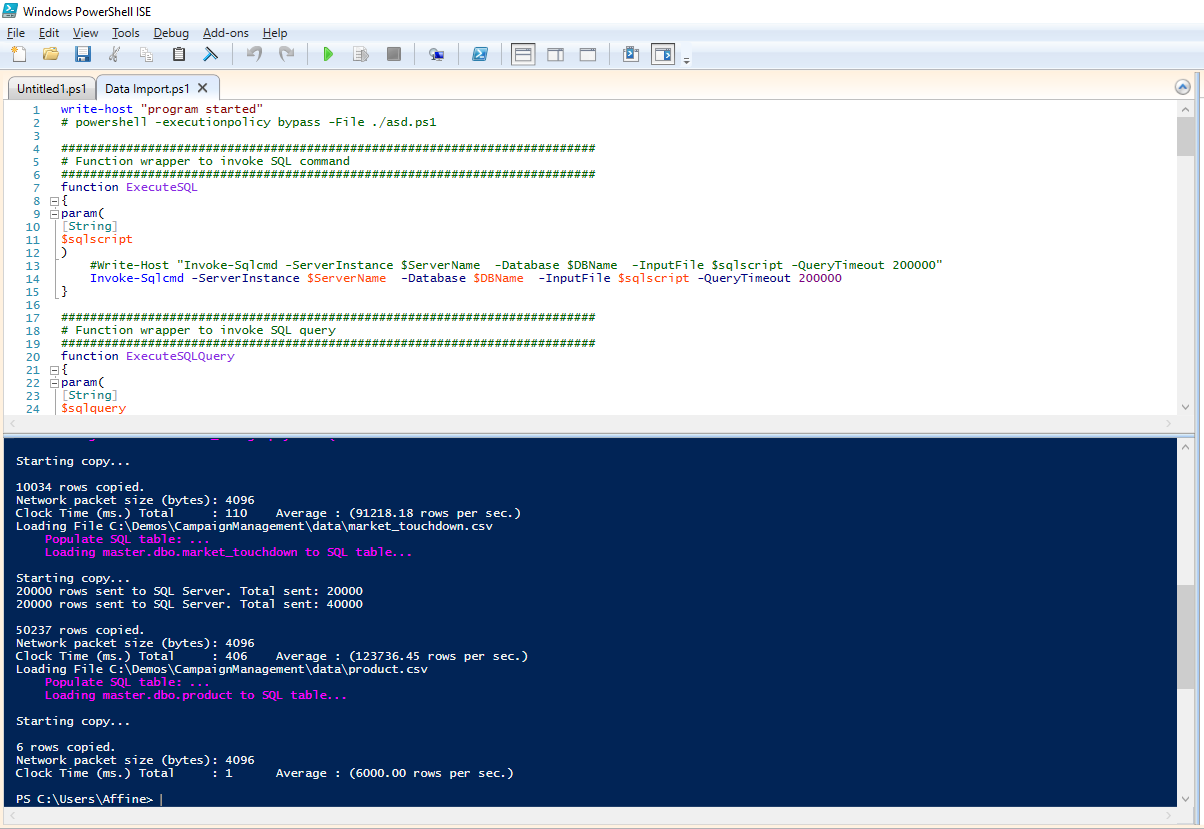
1. *Enter ‘dbo’*

The command line window will prompt you to enter your Username and Password. Refer the instructions from the Set up instructions document

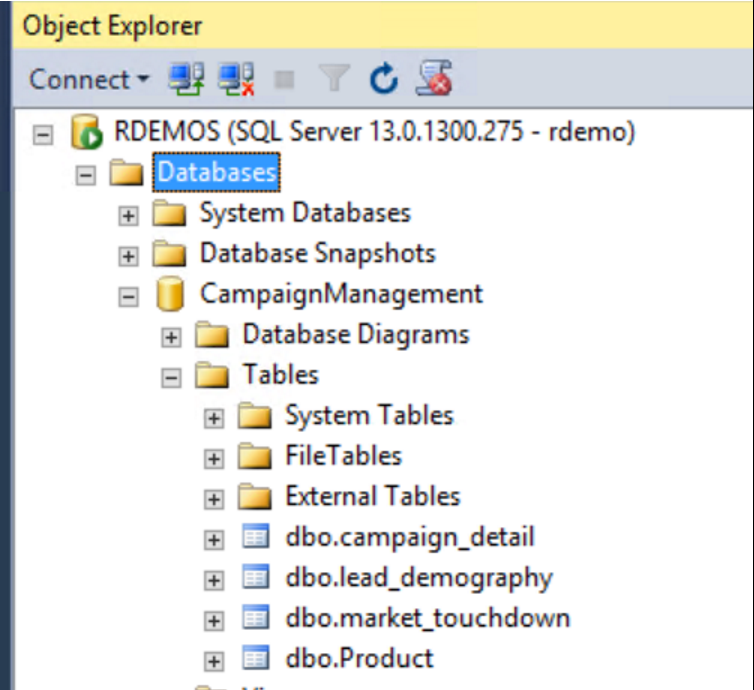
1. *Enter Your UserName & Password that you created earlier using the Set up Instructions document*

The command line window will prompt you to enter the Full Path of the Data Location

1. *Enter the full path of the folder you extracted the CampaignManagement.zip file (C:\Demos\CampaignManagement)*

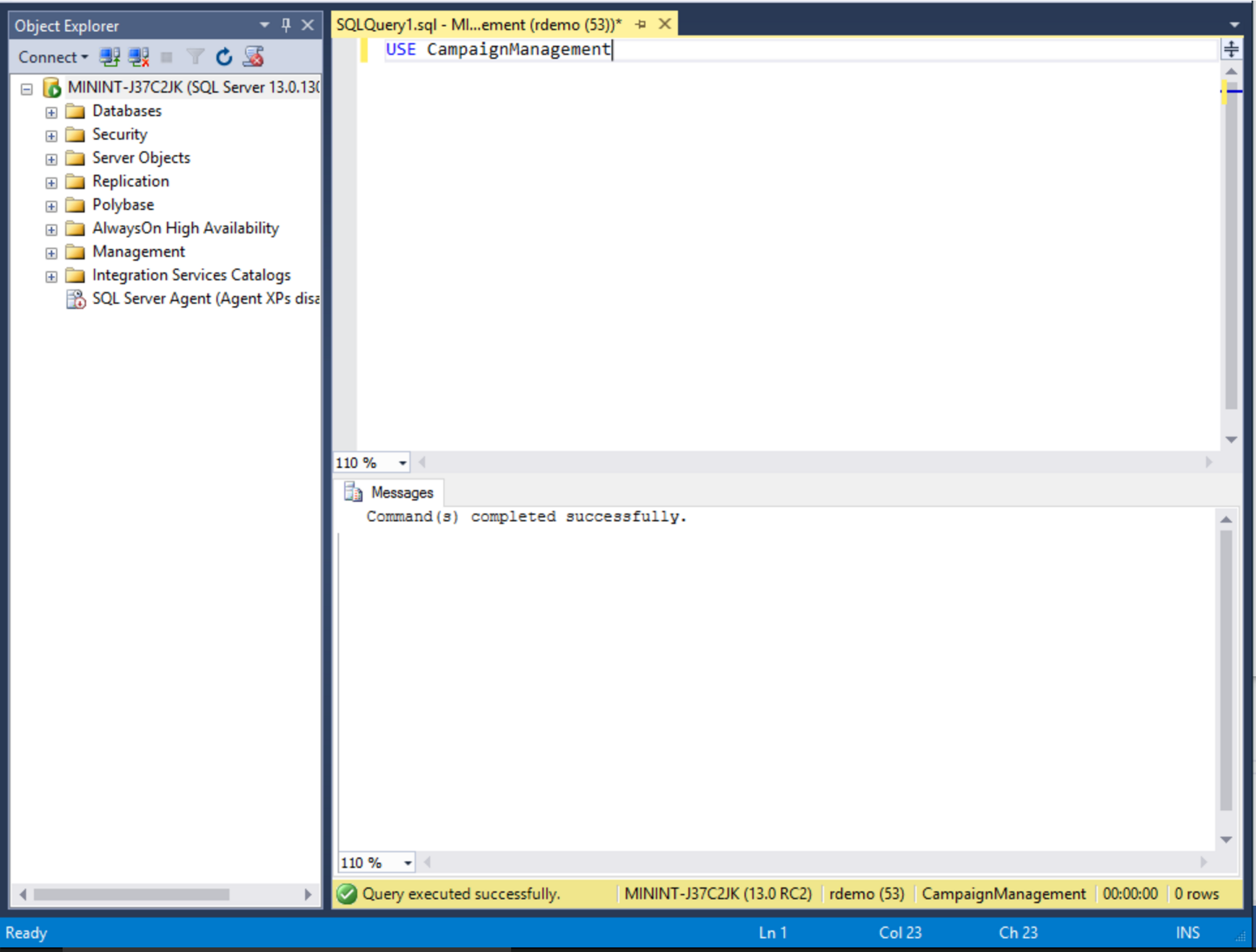
**

1. *Logon to SSMS using the credentials you created earlier. Follow the path shown in the image below and verify that the input datasets are created*



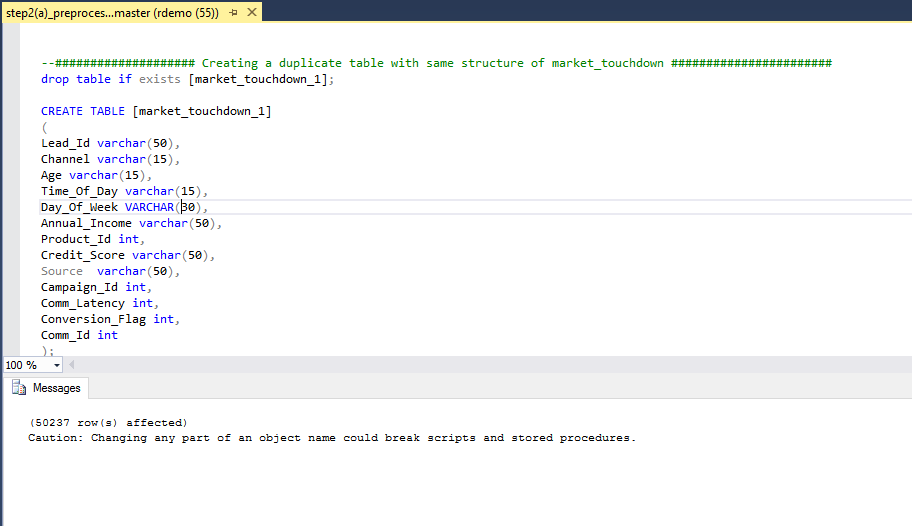
Now, all the required input data sets have been created. Now, we can move on to preprocessing the input datasets

1. *Open a new Query and type* USE CampaignManagement *into the query and Execute*



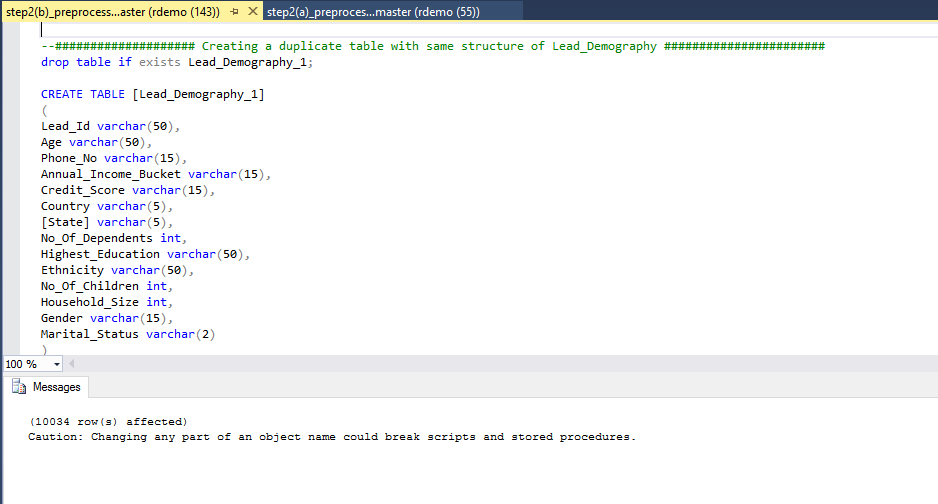
1. *Open step2(a)\_preprocessing\_market\_touchdown.sql and click on Execute*

This step creates the treats the market touchdown dataset for outliers



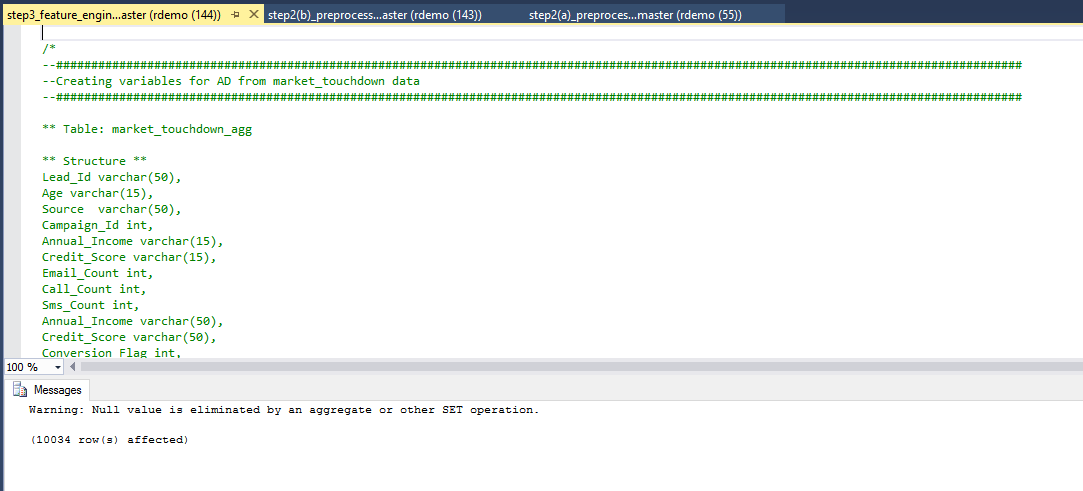
1. *Open 2. step2(b)\_preprocessing\_lead\_demography.sql and click on Execute*

This step creates the treats the target demography dataset for missing values



Once, the preprocessing is done. It’s time to perform feature engineering on the market touchdown dataset

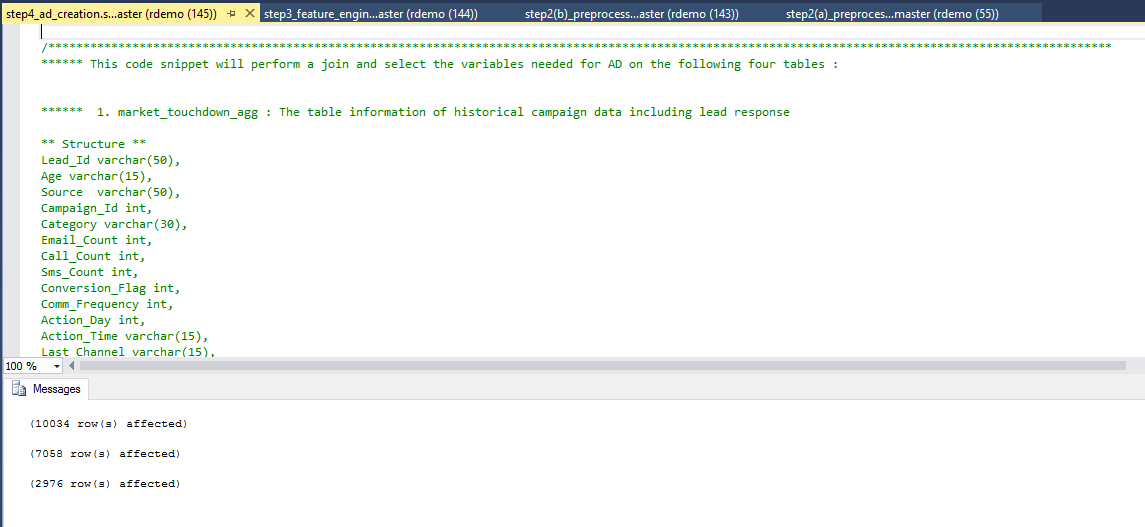
1. *Open step3\_feature\_engineering\_market\_touchdown.sql and click on Execute*

This step creates the new variables in the market touchdown dataset by aggregating the data in multiple levels. The table is aggregated at a lead level, so variables like channel which will have more than one value for each user are pivoted and aggregated to from variables like SMS count, Email count, Call Count, Last Communication Channel, Second Last Communication Channel etc.

Now that the feature engineering step is complete, we can now join all the input datasets to create the analytical dataset

1. *Open step4\_ad\_creation.sql and click on Execute*

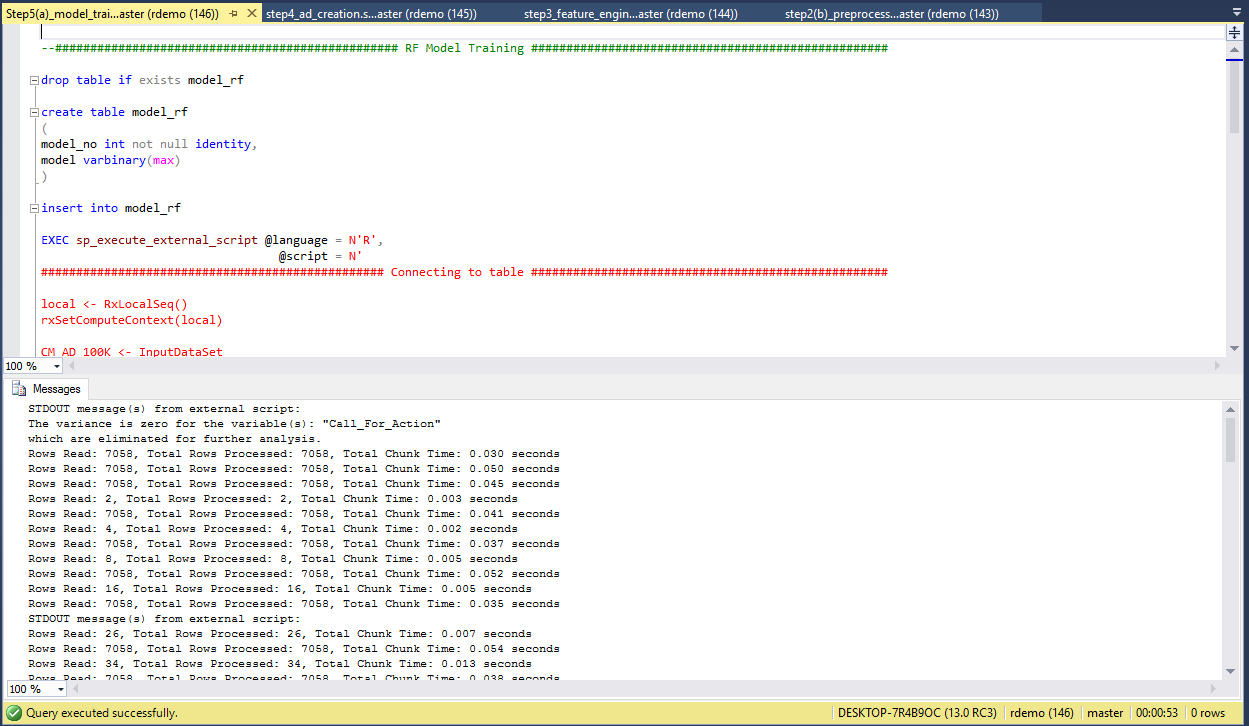
This step merges the processed input datasets to create the analytical dataset. This step also splits the analytical dataset into a 90-10 train-test random sample



The Analytical dataset is ready. We can start building the models now

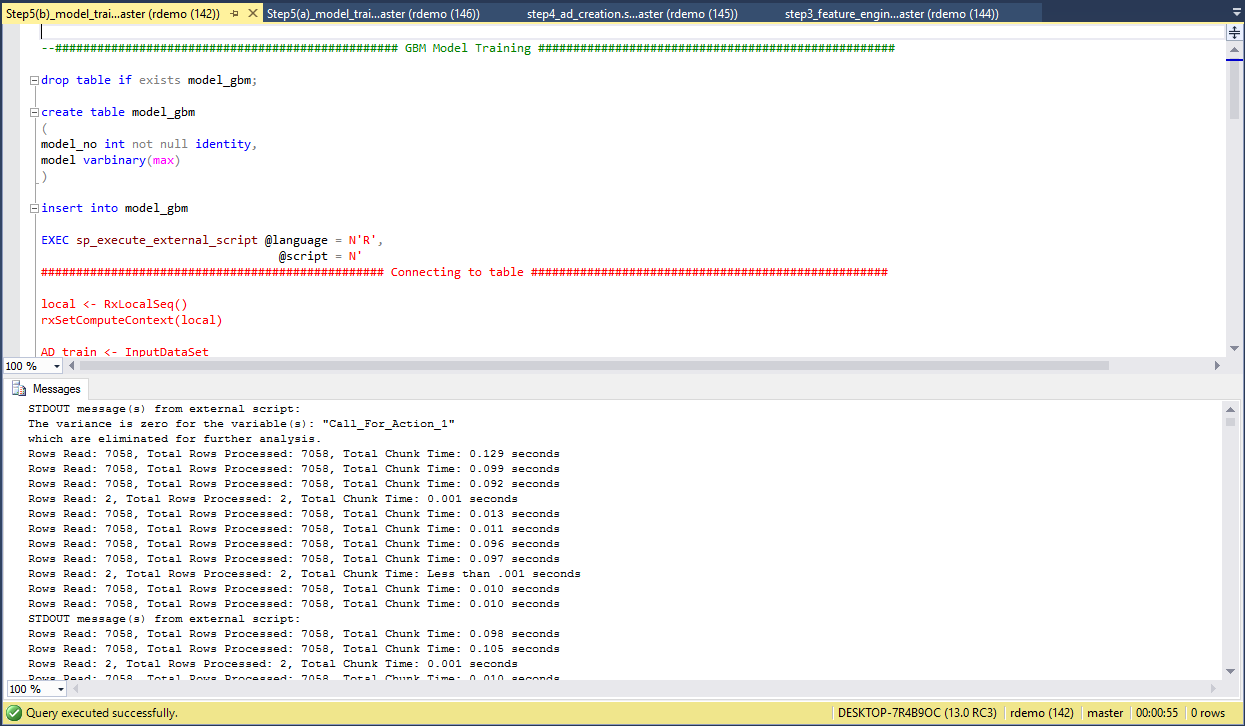
1. *Open Step5(a)\_model\_train\_rf.sql**and click on Execute*

In this step the random forest model is trained on the training dataset



1. *Open Step5(b)\_model\_train\_gbm.sql**and click on Execute*

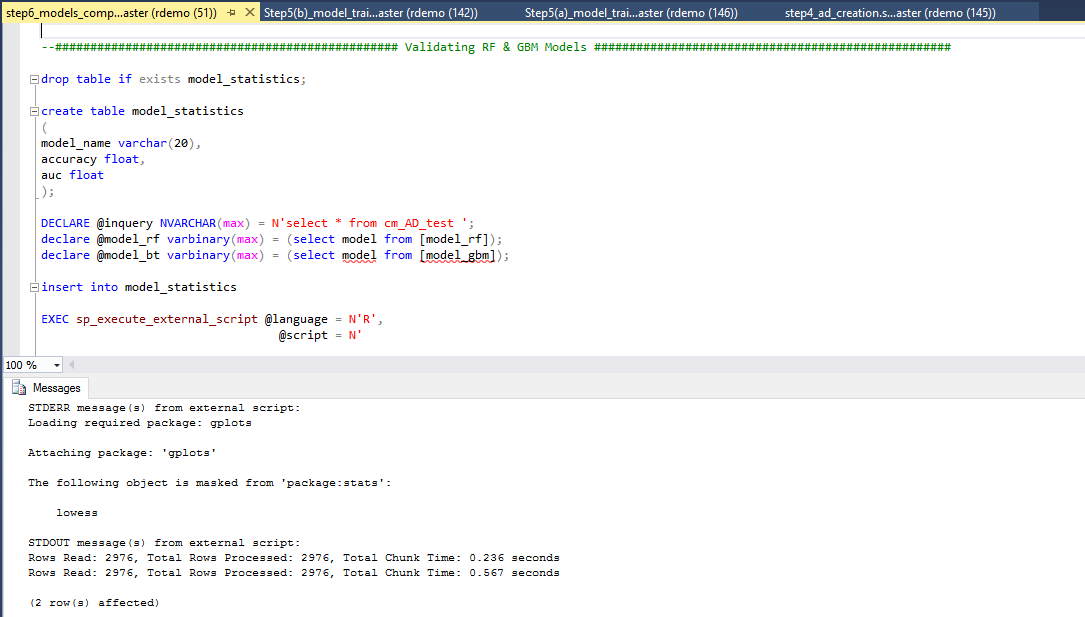
In this step the gradient boosting model is trained on the training dataset



The model algorithms have been created. Now, we need to compute the model statistics

1. *Open step6\_models\_comparision.sql and click on Execute*

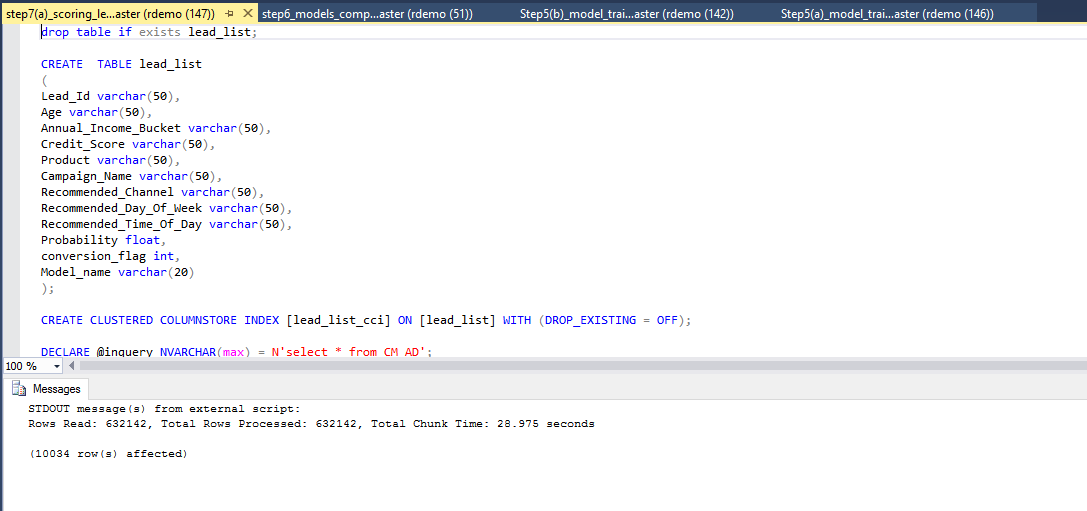
This step calculates model statistics like Accuracy and AUC on the test dataset for both the model algorithms



The next step is to select the champion model using the model statistics and score the analytical dataset using the champion model

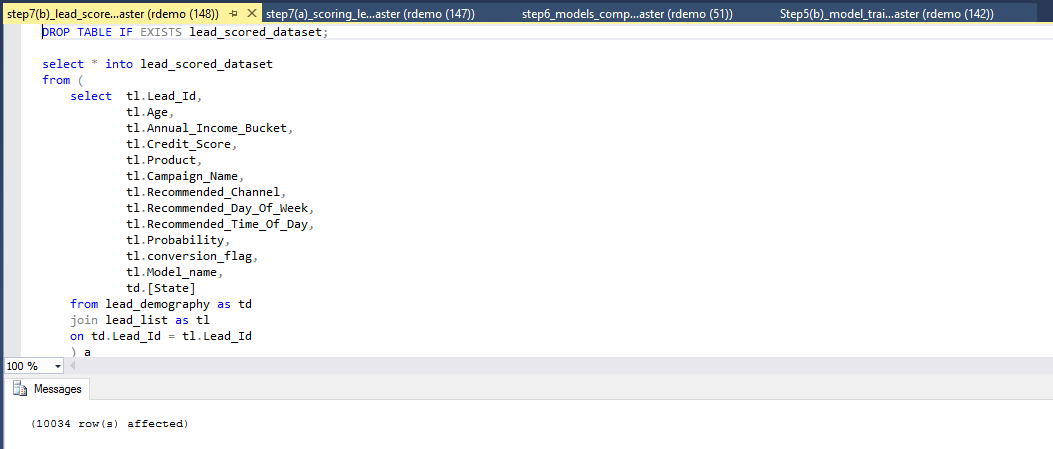
1. *Open step7(a)\_scoring\_leads.sql and click on Execute*

This step selects the champion model based on the AUC. It also scores the analytical dataset on the champion model

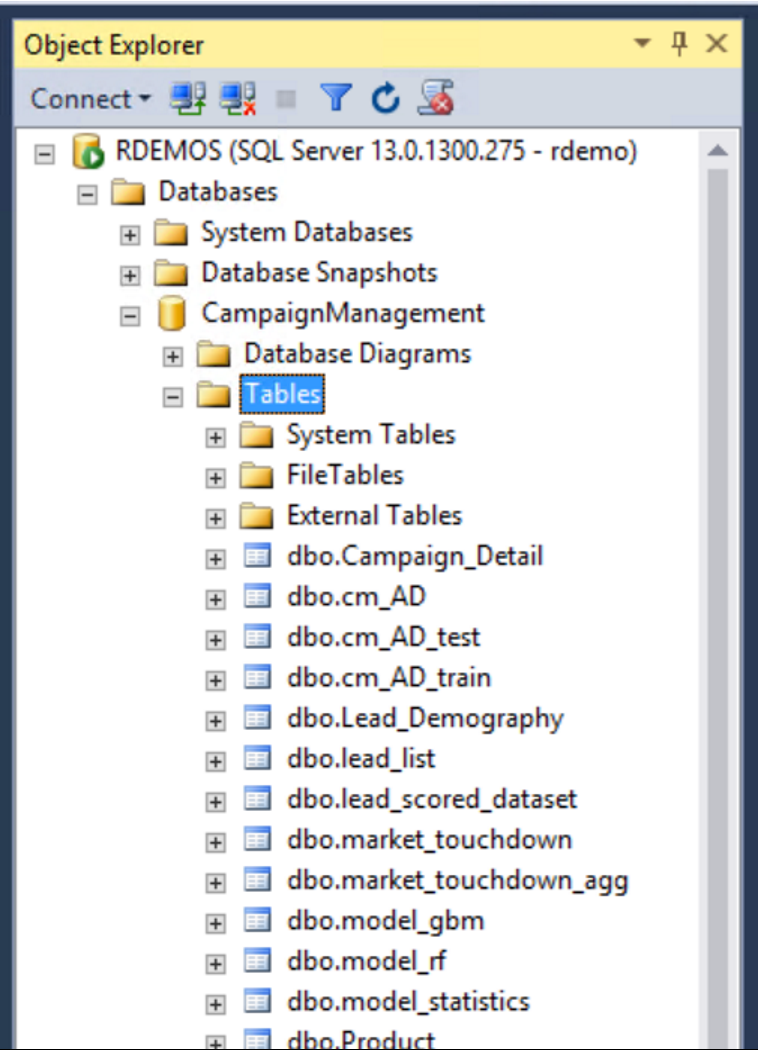


Now that the champion model has been scored, we need to create the final dataset which will be used as the input for the dashboard

1. *Open step7(b)\_lead\_scored\_dataset.sql and click on Execute*



1. *Follow the path shown in the image below and verify that the all the datasets are created. If you do not see them, right click the tables folder and click ‘refresh’*



1. *Now proceed to Visualizing Results with PowerBI by following the instructions in* [*Visualize Results*](Visualize%20Results.docx)*.docx.*