Guidelines to run the scripts:

Software & Packages Required:

> MySQL
> Pandas

Files Attached:

Database Related files:

We have used MySQL database to store the 1.3 million tweets we have collected so far. We have exported the same into two formats such as:

A CSV file (TrumpDataset.csv). The separator used is comma (",")

An SQL dump (sys_tweetdataset77AffinCustom.sql).

Python Code:

The below mentioned python scripts access the Database and retrieves the data to perform computations accordingly.

ApprovalRatingCalculatorDaily.py – calculates the approval rating on a daily basis from 2nd April 2017 until 29th April 2017.

ApprovalRatingCalculatorWeekly.py –calculates the approval rating on a weekly basis for four weeks from 2nd April 2017 until 29th April 2017.

ApprovalRatingCalculatorStatewiseWeekly.py – calculates the approval rating on a weekly basis for particular state for four weeks from 2nd April 2017 until 29th April 2017.

When the above script is executed, it asks for the state combinations as input to provide the weekly approval ratings for that state. An instance of how the input should be is as follows:

Michigan,MI,Lansing
NewYork,NY,New York,NYC
Utah,UT,Salt Lake City,SaltLakeCity

If you are unable to connect to the Database or have any other issues. We have also written a sample script that reads data from sample CSV files and provides the approval ratings for the same.

ApprovalRatingCalculatorSample.py – calculates the approval ratings on a weekly basis for a given state by reading the data from a CSV file containing the appropriate state data.

➢ We have also included CSV files for three states such as New York, Utah and Michigan for four weeks named as follows:

Michigan_week1.csv
Michigan_week2.csv
Michigan_week3.csv
Michigan_week4.csv

NewYork_week1.csv
NewYork_week2.csv
NewYork_week3.csv
NewYork_week4.csv

Utah_Week1.csv
Utah_Week2.csv
Utah_Week3.csv
Utah_Week4.csv

When you run the script, it will ask for the file path to these CSV files. When you provide the same, it will output the approval rating for that state at that particular week after calculating the affin and custom lexicon score.