Summary

To run the program, first we should set the Environment into our machine. We are using Windows to run the program so must install the following:

Python 2.7 – It is programming language

Anaconda – Software use for Data Science application

MongoDB – NoSQL database to store the streaming data

Tweepy - Install Tweepy library into the Anaconda terminal

In the TweetRead.py first we should import the libraries and connect with the applications

First, we must import everything which is require running the streaming tweets. Then need to paste the keys Consumer Key, Consumer Secret, Access Token, Access Token Secret to connect the program with Twitter.

Then create a class called **class StdOutListener(StreamListener):** which inherits from StreamListener.

Then need to connect with the MongoDB database.

If clients exceed a limited number of attempts to connect to the streaming API in a window of time, they will receive error 420. Tweepy’s Stream Listener usefully passes error messages to an on\_error stub.

Then we must manage the output with the authentication keys and filter. In our case we use the following tags.

stream.filter(track=['trump', 'modi', 'putin'])

Then run the program with python TwitterRead.py > twitter\_data.txt command into the anaconda terminal. This command will store the streaming data into the twiiter\_data.txt file which will use for analysis.

Next create another file as TweetAnalysis.py for analysis of the twitter\_data.txt

First as usual need to import everything with require making analysis as import pandas for data manipulation, matplotlib to draw charts.

Then define the method called word\_in\_text() to change them into lower case.

Into the main() method provide the tweet\_data file path. Then we set the compiler to read line by line and append them together.

Next, we should structure the tweets of pandas with text, language and country.

Now in the first query, we will compare the popular language i.e. number of tweets about the Trump, Modi and Putin in which language. So, from the result, we can analyse that as most of the tweets are in English language compare to Spanish, French and others.

In terms of ranking, Trump is most popular compare to Putin and Modi. Putin and Modi are not even close to trump. Similarly, in second attempt Trump is the most popular leader compare to Modi and Putin.