

**Sentiment Analysis on Elon Musk’s tweets using Python**

**Bhargava Ranga**

**2685185**

**Sentiment Analysis:**

Sentiment Analysis is the procedure of 'computationally' deciding if a bit of composing is certain, negative or impartial. It's otherwise called opinion mining, inferring the state of mind of a speaker.

Sentiment analysis is used in various categories like:

* **Business**: In advertising field organizations utilize it to build up their procedures, to comprehend clients' emotions towards items or brand, how individuals react to their item dispatches and why buyers don't get a few items.
* **Politics:** In political field, it is utilized to monitor political view, to identify consistency and irregularity amongst proclamations and activities at the administration level. It can be utilized to foresee election results also.
* **Public actions:** It is used to screen and break down social phenomena, for the spotting of conceivably perilous circumstances and deciding the general inclination of the blogosphere.

**Twitter sentiment analysis:**

Sentiment analysis is the examination of feelings, attitudes and sentiments that are helpful for settling on better business choices. A considerable measure of organizations are doing sentiment analysis of Twitter clients by getting conclusions from their tweets. The most critical reason is to check how your image is performing in the market and what your clients are saying in regard to it. Additionally, with the assistance of tweets occasion recognition, area identification and more can be performed flawlessly.

The steps to be followed are:

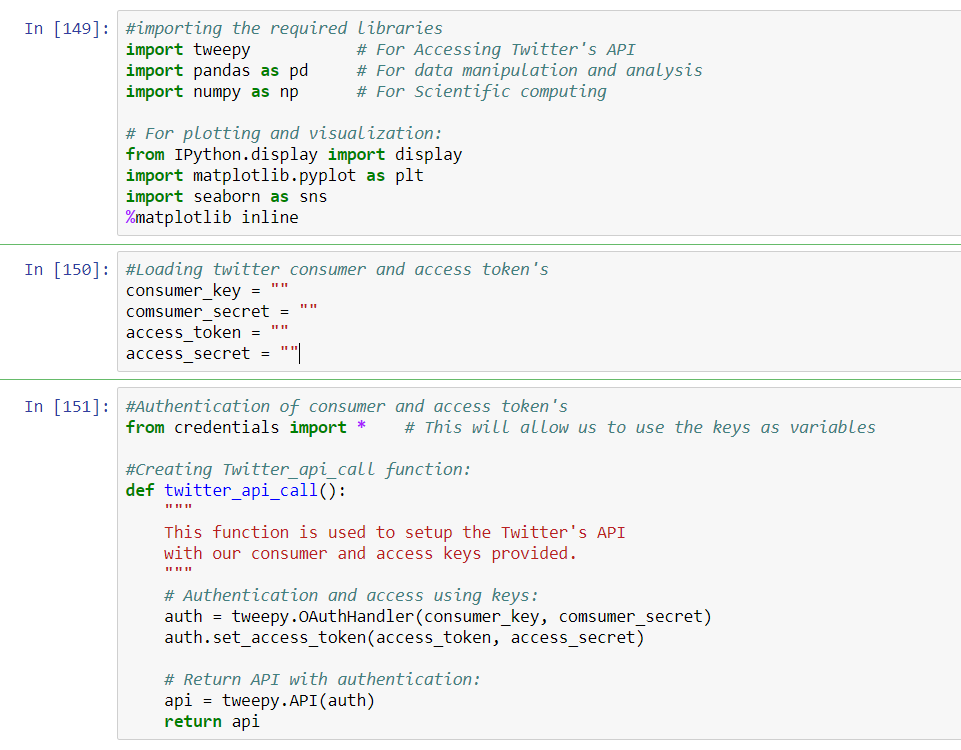
1. Extract twitter information utilizing tweepy and handle it with pandas.
2. Perform sentiment analysis on the tweets using textblob.
3. Display some essential insights and representation utilizing numpy and matlplotlib.

The requirements for performing twitter sentiment analysis are:

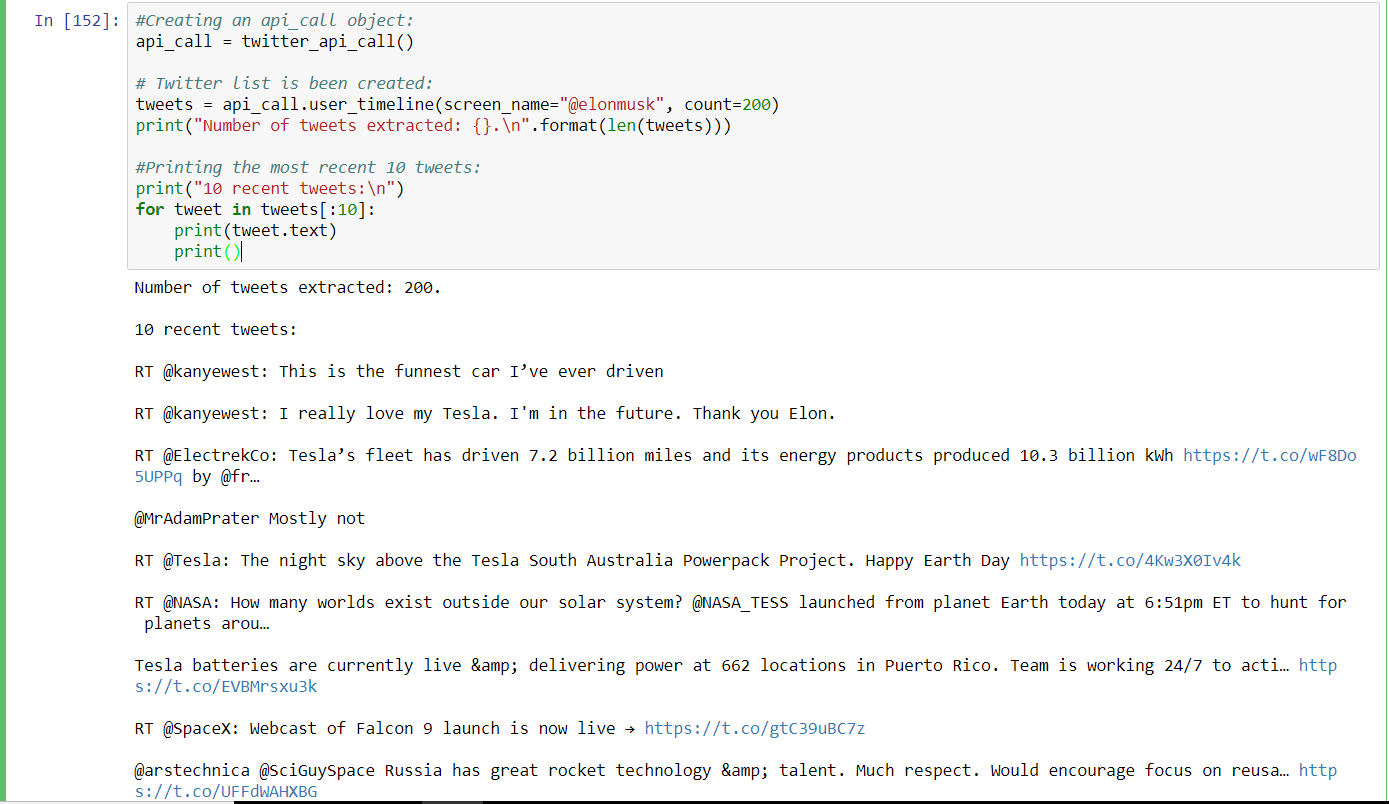
* Python must be installed
* NumPy: This is a python library which is used to perform scientific computation.
* Pandas: This library is used for data manipulation and analysis.
* Matplotlib: This library is used for 2D plotting and visualization
* Tweepy: This library is used for accessing twitter API.
* Textblob: This library is used for processing sentiment analysis using textual data.

**Extract twitter information utilizing tweepy and handle it with pandas:**

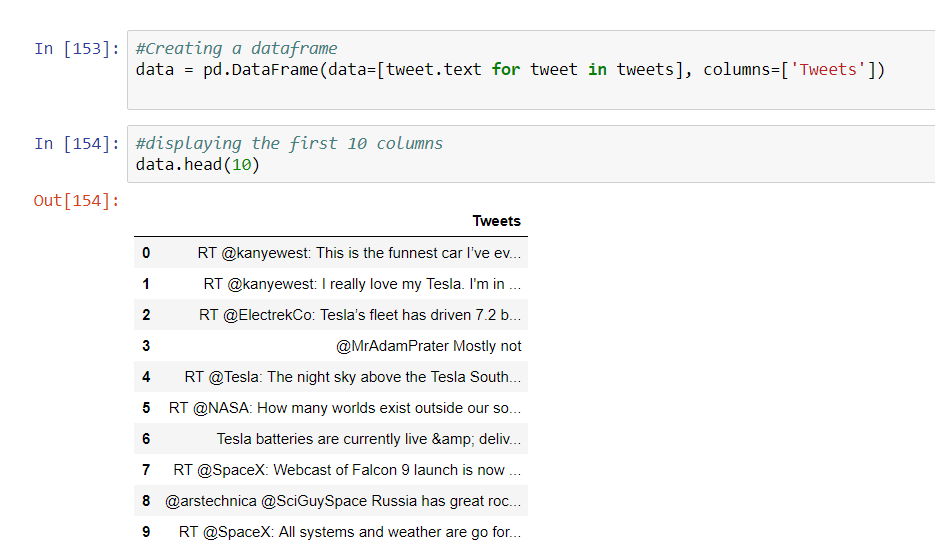
* First, we must import the required libraries for the twitter sentiment analysis and run it.
* To extract the tweets, we need to have a twitter account, so we must create a twitter account and then save all the credentials like consumer key, consumer secret, access token, access secret in a script called credentials.py.
* By using those credentials, we can access the twitter account. Every user has different credentials and they are confidential. So I have not included my credentials in the main code.
* We are currently ready to expend Twitter's API. To do this, we will create a twitter api call function to permit us our keys verification.



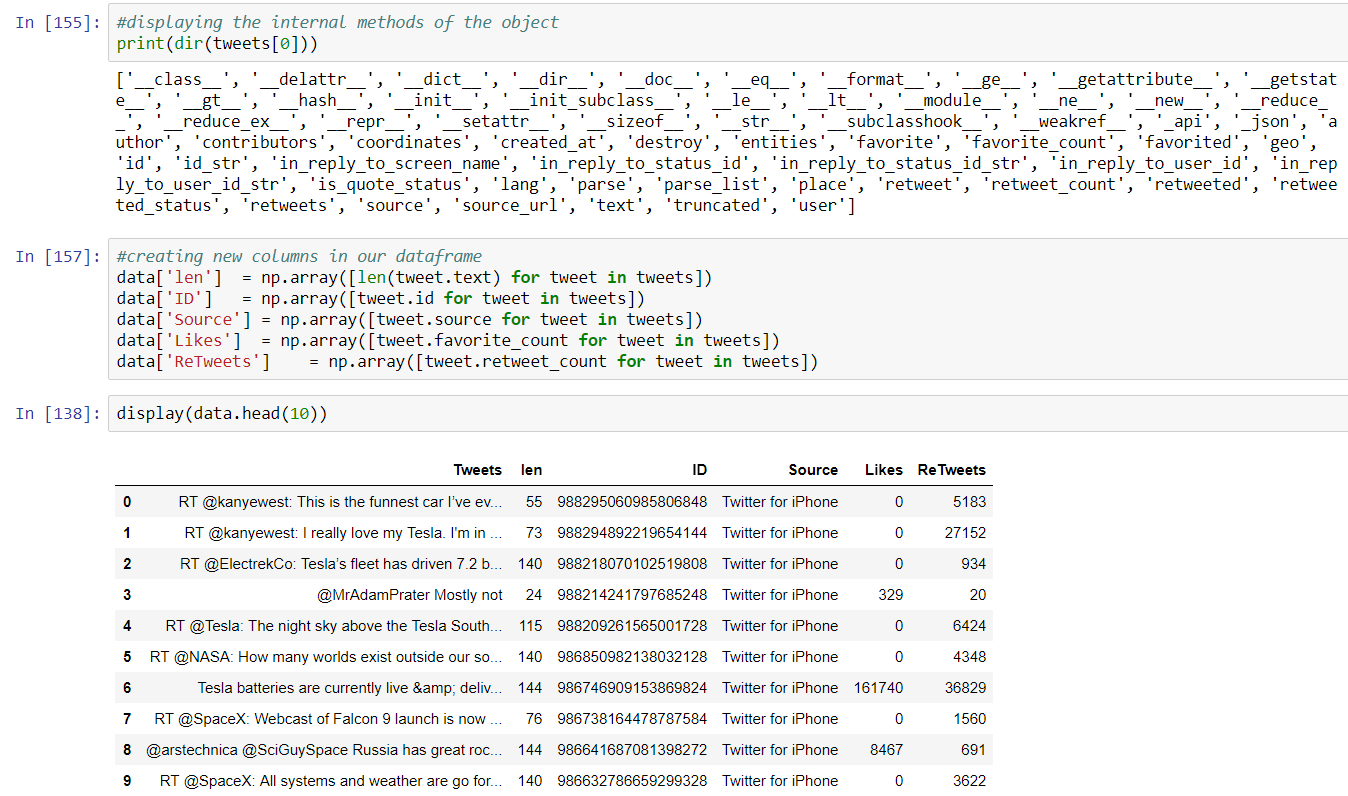
* We can utilize twitter api call function to make an "api call" object. After this, we can utilize Tweepy’ s function to extract from screen name's client the number of total tweets.
* As it is specified in the title, I've picked @elonmusk as the client to extract twitter data. Then we print the most recent 10 tweets.
* The output is shown below, and we can also compare our output with the twitter account to check if we are correct or not.



* We now have introductory data to build a pandas DataFrame, keeping in mind the end goal to control the information in a simple manner.
* The head method of a dataframe allows us to visualize the first n elements of the dataframe.
* By using python’s list comprehension, we create a dataframe and display the first 10 columns of the dataframe.
* The output is shown below.



* We can display the internal methods of the tweet structure using the tweepy
* As should be obvious, we can get a considerable measure of information from a single tweet. But not all the data is useful for our sentiment analysis. For our situation we well simply add a few information to our dataframe.
* We select some of the internal methods which we want to display in our dataframe, I have selected length of the tweet, ID, Source, likes Retweets
* We will utilize Pythons list comprehension and new column will be added to the dataframe. Then we display the first 10 columns of our dataframe to see our changes.
* The output is shown below.

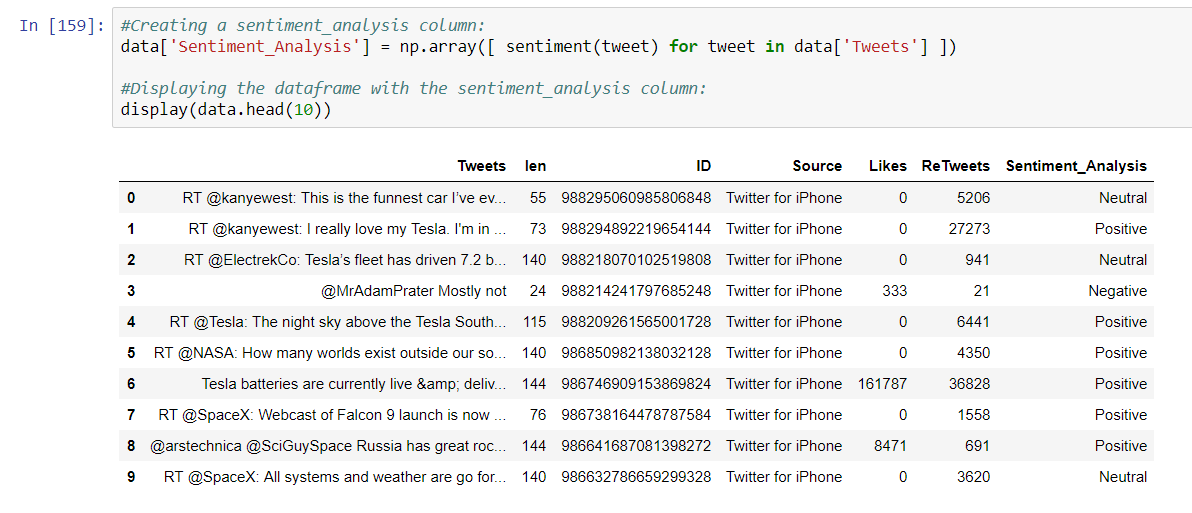


**Perform sentiment analysis on the tweets using textblob:**

* As we specified toward the start of this post, textblob will enable us to do sentiment analysis in an exceptionally basic manner. We will likewise utilize the re library from Python, which is utilized to work with regular expressions
* We create a text cleaning function which is used to clean the text by removing the links and special characters using regular expression library.
* After cleaning the tweets, we create a sentiment function which is used to classify the polarity of a tweet using textbolb.
* If the polarity is greater than 0, then the tweet is Positive.
* If the polarity is equal to 0, then the tweet is Neutral.
* If the polarity is less than 0, then the tweet is Negative.



* Finally, we create a sentiment analysis column which is our result and add it to our dataframe.
* We display the first 10 columns of the dataframe with our result.
* The output is shown below.



**Display some essential insights and representation utilizing numpy and matlplotlib:**

* We are trying to display the tweets which has the maximum number of likes and retweets using numpy max function.
* Since more than one could have a similar number of likes/retweets (maximum) we simply need to take the first found, and that is the reason we utilize index[0].
* Then we print the tweet which satisfies our condition.
* The output is shown below.



* We are using pandas value\_count method to display the unique values in the column sentiment analysis.
* Out of 200 tweets which are extracted for our user @elonmusk, the count of positive, negative and neutral tweets are shown below.
* Then we plot the graph of the unique values displayed using matplotlib.

