Final Submission Write-Up

1. Project Design

**Sentiment Analysis**

***Project Design***

**Main Functions:**

* **cleanTweet () 🡪 Function to clean up Tweets**
* **percentage() 🡪 Function to calculate Percetange**
* **find\_reaction\_avg() 🡪 Function to Find the Average for Reaction**
* **process\_tweets() 🡪 Function to process the Tweets**
* **plotPie() 🡪 Function to plot Pie Chart**
* **plot\_cloud () 🡪 Function to plot a word cloud**
* **wordCloud() 🡪 Function that creates a word cloud object and plot it by calling the plot\_cloud() function**
* Connect to TweeterAPI
* Search Tweets about Chris Rock by calling the **.search\_tweets()** method
* Process the tweets by calling the **process\_tweets()** function
* Plot the Pie Chart by calling **plotPie()** function
* Same Process for searching Tweets for Jada Smith
* Same process for searching Tweets for Kim Kardashian
* Plot word cloud by calling **wordCloud()** function for Chris Rock, Jada Smith, and Kim Kardashian

2. Test Cases

Test Case 1: Kim Kardashian -- 50 tweets

People tweet's reactions to @KimKardashian by analyzing 50 tweets.

General Report: Polarity 0.1334506944444445

--Positive tweets--

Keep your head high You’ve won public opinion Your shows are rocking even more as you show compassio https://t.co/OeryTk2RGS

Good morning Chris Just wanted to compliment you on the 2009 Documentary “Good Hair. Thank you Be well ✌🏼🎵

...

25 people reacted neutral

32.00% people reacted positive

18.00% people reacted negative

50.00% people reacted neutral

Chart, pie chart

Description automatically generated

Test Case 2: Chris Rock – 39 tweets

People tweet's reactions to @chrisrock by analyzing 39 tweets.

General Report: Polarity 0.14562121011583903

--Positive tweets--

Keep your head high You’ve won public opinion Your shows are rocking even more as you show compassio https://t.co/OeryTk2RGS

Good morning Chris Just wanted to compliment you on the 2009 Documentary “Good Hair. Thank you Be well ✌🏼🎵

Happy Easter https://t.co/qe6rL72PEM

…

Detailed Report:

16 people reacted positive

2 people reacted negative

21 people reacted neutral

41.03% people reacted positive

5.13% people reacted negative

53.85% people reacted neutral

Chart

Description automatically generated

Test Case 3: Jada Smith – 49 Tweets

People tweet's reactions to @jadapsmith by analyzing 49 tweets.

General Report: Polarity 0.1419704700061843

--Positive tweets--

Keep your head high You’ve won public opinion Your shows are rocking even more as you show compassio https://t.co/OeryTk2RGS

Good morning Chris Just wanted to compliment you on the 2009 Documentary “Good Hair. Thank you Be well ✌🏼🎵

Happy Easter https://t.co/qe6rL72PEM

…

Detailed Report:

16 people reacted positive

1 people reacted negative

32 people reacted neutral

32.65% people reacted positive

2.04% people reacted negative

65.31% people reacted neutral

Chart, pie chart

Description automatically generated

Test Case 4: Data Visualization with Seaborn

Chart, bar chart

Description automatically generated

3. Learning Experience

The project 5 is a project that I will save for future research since I am interested in manipulating APIs from different platforms. The assignment was about to analyze tweets from real people and categorize them as positive, negative, and neutral comments. I learned how to connect the Twitter API with Python by using the TweePy library, and how to extract users’ information. To do so, I needed to create several functions to make the program work. However, one of the most important functions from this project is the **process\_tweets()** since it takes the tweets and convert them in positive, negative, and neutrals by using the polarity from **TextBlob** library, and then displays the results. We know that when polarity is 0 means neutral, greater than 0 is positive, and less than 0 is negative. That is how **process\_tweets()** works, then I needed to calculate the average of each category (positive, negative, neutral) in percentage to display this info in a chart.

Moreover, I also learned how to create charts in different libraries such as Matplotlib and Seaborn. For this project, I created a Pie Chart in Matplotlib to display the percentage of positive, negative, and neutral reactions from users towards a celebrity. One for each celebrity. Moreover, I also created a bar plot in two different libraries: Matplotlib and Seaborn. I found out that the Seaborn charts are more compelling and attractive than Matplotlib and easier to understand the syntax.

Furthermore, I created word clouds for each celebrity. For this part of the project, I review past assignments in order to program the word clouds. First, I processed the tweets as a whole string. Then, I converted the text as a list to cleaned up each word and converted again as a string text. Then, I re-used the functions from past assignments to plot a word cloud. Each word cloud represents most of the common words used in tweets. I created one word cloud per celebrity

At the end, the program worked successfully and had not complications with connecting to the Twitter API. I am interested to further investigate about APIs in programs and how to use them in different platforms such as Facebook, Google, Amazon, and others. Overall, it was a fun program to code, and I am happy with the outcome of this assignment, and I am ready for the next challenge coding program and projects.

4. Assumptions

* Assumption for Word Clouds:
* Quotes sentences may end a sentence in ( “ ) instead of a period ( . )
* Words may be compound or joined by a dash ( - ) or double dash ( -- )
* Words may be compound by (‘ll) or (‘s) or (‘t)
* Words may have special characters (()\"#$/@;:<>`+=~)
* All of these assumptions were addressed in the program
* All user comments will be classified into positive, negative, or neutral categories.
* There will be no comments without classification.
* Users chosen from Twitter will have comments to be analyzed