Twitter analysis

# **Adonis Shareef Fall Semester https://github.com/adonisShareef/DataScience**

# Which Domain?

What domain is this data going to come from? Please list 10 references (with a brief annotation) to use to make sense of what you’re doing with these data.

My domain is the world of social media and specifically I am looking into is from people with social significance. I will be able to see how their tweets are of three choices either positive, negative, or neutral. This is important because they we can investigate the distribution of those types of tweets and the way the world looks at these figures. Social media is a huge part of the world today so the biggest figures in a culture should in theory have the biggest impact on social media like Twitter.

1. <https://monkeylearn.com/blog/sentiment-analysis-of-twitter/#:~:text=Twitter%20sentiment%20analysis%20allows%20you,mentions%20before%20they%20they%20escalate>.
2. <https://www.analyticsvidhya.com/blog/2021/06/twitter-sentiment-analysis-a-nlp-use-case-for-beginners/>
3. <https://developer.twitter.com/en/docs/tutorials/how-to-analyze-the-sentiment-of-your-own-tweets>
4. <https://www.geeksforgeeks.org/twitter-sentiment-analysis-using-python/>
5. <https://towardsdatascience.com/step-by-step-twitter-sentiment-analysis-in-python-d6f650ade58d>
6. <https://www.coursera.org/projects/twitter-sentiment-analysis>
7. <https://www.kaggle.com/c/tweet-sentiment-extraction>
8. <https://brand24.com/blog/twitter-sentiment-analysis/>
9. <https://www.akkio.com/twitter-sentiment-analysis>
10. <https://betterprogramming.pub/twitter-sentiment-analysis-15d8892c0082>

# Which Data?

What is the dataset you’ll be examining? Please provide a codebook if there is one or a link to the dataset as well as a detailed description.

1. <https://www.kaggle.com/datacrux/barack-obama-twitterdata-from-20122019>

* Tweets by Barack Obama’s twitter account

1. <https://www.kaggle.com/vidyapb/elon-musk-tweets-2015-to-2020>

* Tweets by Elon Musk’s twitter account

# Research Questions? Benefits? Why analyze these data?

How are you proposing to analyze this dataset? This is about your approach. Here, you’ll be proposing your research questions as well as justifications for why you’d offer these data in this way.

What type of sentiment do popular figures tweet and how does that impact social media, specifically Twitter?

I want to analyze this data because the demographic for social media are younger people and kids which is going to be interesting if I find that negative tweets are coming from our biggest figures in the United States. So, I want to dig deeper and see what type of tweets have been put out into the world by these figures and now that the data is there I can analyze the results.

# What Method?

What methods will you be using? What will those methods provide in terms of analysis? How is this useful?

I will be using natural language processing NLP as well as tokenizing to determine the negatives and positives from the tweets. This will allow me to create new data frames with this information so I can then do more analysis on the tweets and come up with some answers on what impact these tweets are having.

# Potential Issues?

What challenges do you anticipate having? What could cause this project to go off schedule?

At first glance I do not think there will be any issues because I just want the text but maybe the account was hacked at certain points and some tweets were not accurate of that figure but that is a part of the security of their brand and platform so I will not throw that data out. Also, some figures like Elon Musk I now tweet some random things so choosing a threshold for that type of figure might be difficult.

# Concluding Remarks

Tie it all together. Think of this section as your final report’s abstract.

I believe from the data I have found and the methods I am choosing there should be no problems answering the question, most of the work will be to discover the trends within the results because the data is just text, which I have handled before. I just wanted to sentiment analysis on a topic that I was interested in. Maybe after this project I might have different views on the figures I am going to use as data for this project. There will be an answer that is solidified unlike my last project where there was an answer but I am still unsure if it was the correct answer.