Lim, Riana Mary Claire G.

CS 193.1

Tool Exercise 1: Data Story

Problems

I want to view the opinions of Twitter users on four popular songs used in the new live-action Aladdin movie. The four songs are A Whole New World, Arabian Nights, Friend Like Me, and Speechless.

I think this is an interesting topic to look into because other than changing the format of the movie from a 2D animation to a live-action movie with actors and CGI, they also made some changes to the songs, storyline, and even added new ones.

Datasets

I queried for a maximum of 300 tweets that mentioned Aladdin and also the song titles. Specifically, I looked for data using the following queries:

```
'aladdin whole new world'

'aladdin friend like me'

'aladdin arabian nights'

'aladdin speechless'
```

I also queried for 300 tweets that mentioned Aladdin in general.

After gathering the data, I processed them using the cleaning function already provided in class before.

Methods

For the exploratory data analysis, I did sentiment analysis on the tweets by checking the sentiment and the subjectivity of each. I did sentiment analysis because I

think it is the best way to get insights into the opinions of Twitter users on the four songs and the movie in general.

I also created heatmaps to see if there is a correlation between the sentiment of Twitter users for each song.

Results

Data Count

First I looked into how much data I was able to gather for each song.

	wnw_sent	flm_sent	an_sent	s_sent	wnw_subj	flm_subj	an_subj	s_subj
count	300.000000	202.000000	124.000000	300.000000	300.000000	202.000000	124.000000	300.000000

In the image above, I described the DataFrame which contains the sentiment and subjectivity of the four songs. For reference, *wnw* refers to A Whole New World, *flm* refers to Friend Like Me, *an* refers to Arabian Nights, and *s* refers to Speechless.

We can see that a lot of tweets were mined for the A Whole New World and also Speechless. From the count, we can deduce that these two songs are the most "tweeted" about.
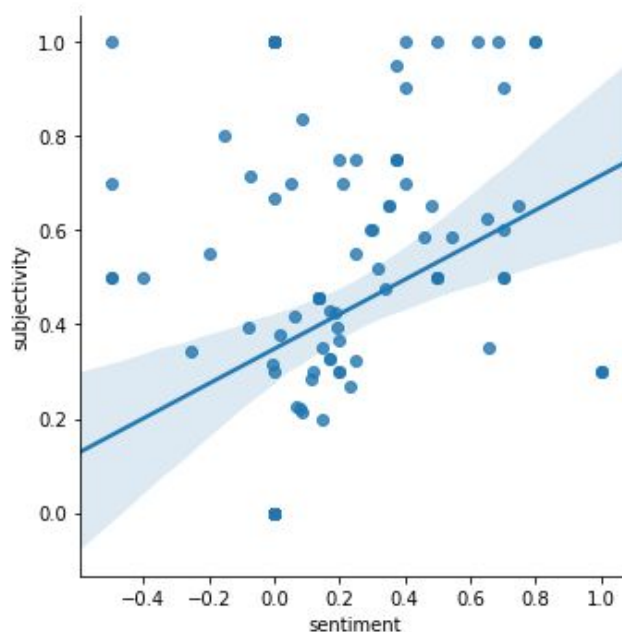
Sentiment and Subjectivity

Based on the graph A Whole New World, the general sentiment and subjectivity are neutral.
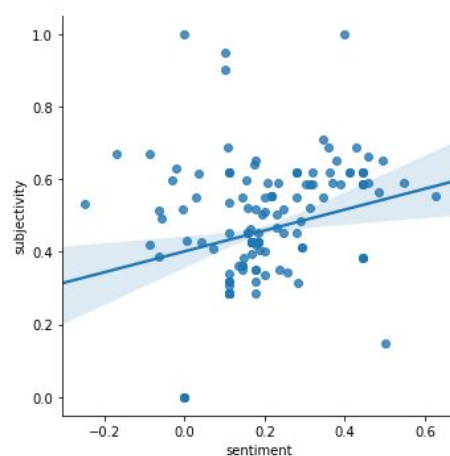
Based on the graph of Friend Like Me, the general sentiment and subjectivity is positive.
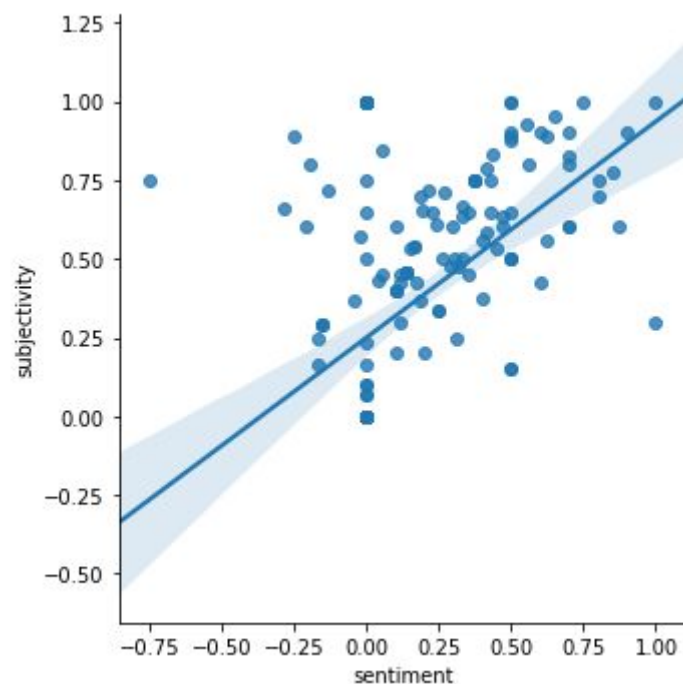
The sentiment and subjectivity of Arabian Nights is a lot more scattered. The sentiment of the tweets is generally negative, but they are a lot more subjective comapred to the others.
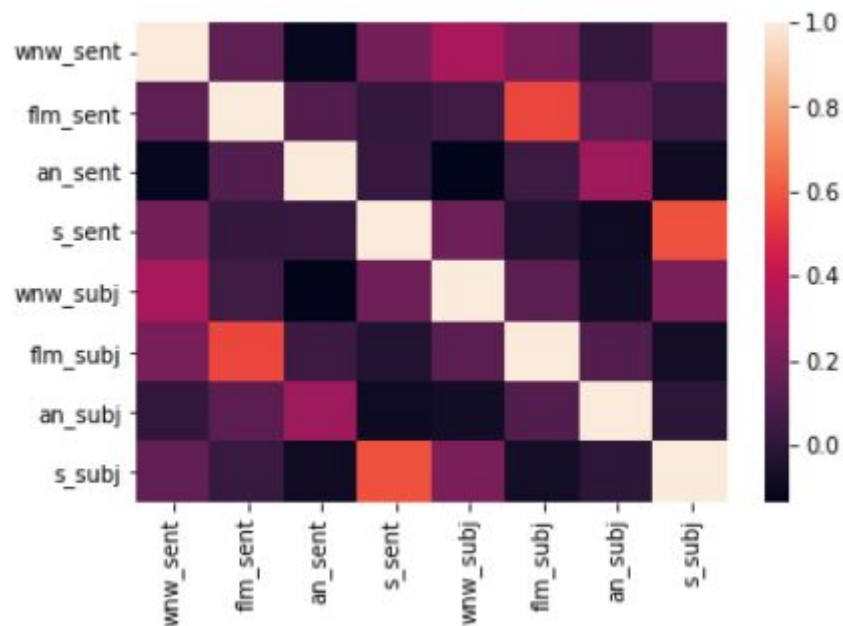
The sentiment and subjectivity of Speechless is also generally positive.
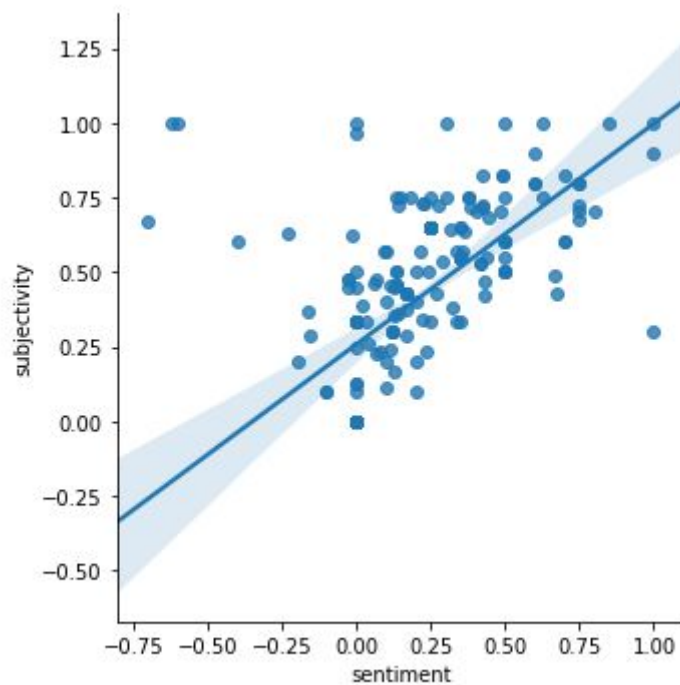
We can notice that the more positive the sentiment for the tweets, the more subjective they are.

I also wanted to see if there was a correlation between the sentiments and subjectivities of the different songs.

It's obvious that the subjectivity and sentiment of the same songs are the most correlated.

Although the different axes labels aren't very correlated in general, we can see that there is a slight correlation between the sentiment and subjectivity of the audience to A Whole New World with the sentiment of the audience to Speechless. This may also be due to those two being the most popular songs out of the four.

All in all, A Whole New World and Speechless gathered the most positive response from Twitter users. A Whole New World might be the song most people would associate with Aladdin because it is sung in the scene were Aladdin sneaks Jasmine out of the palace for a ride on the Magic Carpet, so it might be one of the most anticipated songs from the entire movie. Speechless is actually a new song which was not featured in the original animated movie. In the song, Princess Jasmine sings about how she feels like she doesn't have a voice, but nevertheless, won't stay silent. The popularity of the song may be due to the sudden influx of women empowerment initiatives which is the theme of the song. These may be factors that are causing the positive response of people to the two songs.