Jim Ryan DS710Final Project Proposal

1. My question is based upon the #Brexit hashtag. The UK is supposed to leave the European Union on 01-31-2020. They also have an election coming up December 12, 2019 in an attempt by Prime Minister Boris Johnson to seek a parliamentary majority to enact his Brexit plan. The two largest parties are the Conservatives, which is the pro-Brexit party and Labour which is the anti-Brexit party. I am curious as to which one will appear more as the election approaches and if it provides any insight as to which of the two parties will win the election.
2. My audience for the answer to this question would be the politicians and political parties in Britain, Other audiences would be anyone interested in the outcome of the UK December Election. This would certainly include companies in the US and European Union that currently ship goods to the UK.
3. The information I will gather about each tweet will be the username and text of each tweet. I will also gather the sentiment and subjectivity of the tweet. I am curious to see if the tweets reflect a positive or negative sentiment of Brexit. I am also curious to determine the subjectivity of each tweet. For sentiment analysis the values range from -1 for negative to +1 for positive. The subjectivity values range from 0 for not subjective to +1 for most subjective.
4. I will start off using the Streaming API to collect data. I am confident that there will be enough tweets to gather what I need via the Streaming API. I will run the collected tweet text through Textblob to get a sentiment and subjectivity score.
5. I will count whether the tweet expresses positive sentiment of negative sentiment. I plan to use a Two-sample Test for a Difference of Proportions to compare the positive vs negative Tweets. I will also graph whether a tweets sentiment (positive or negative) is a function as to whether it is retweeted or not. So, of the tweets I gather that have been retweeted, does the sentiment make it more or less likely as to whether it is retweeted or not.