

Abstract:

We have scraped data from twitter based on the keyword “elections”, since we are only interested in 120 sentences, so we decided to limit our tweet scope to Atlanta. We have scraped more than 10000 tweets and filtered them based on word count of tweets, we have chosen only tweet lengths greater than 53 and then scanned every tweet and filtered them further to come with 120 tweets.

Instructions for Labeling:

We have a text column; we want you to label the text column based on whether the tweet is negative or not negative based on your understanding of the tweet. For e.g.,

```
@Truth_Tornado @CNN @jaketapper Again...you are believing a lie because it f
its a narrative which you already accept. There were no lines anywhere in
GA on Election Day in 2020. The law on water is no 3rd party can enter the
150 ft area near polling to do it. You are being lied to for political rea
sons.
```

In the above tweet, we can clearly identify that the tweet is negative. Similarly, we want you to label the data.

Some approach to make the task easier,

Steps:

```
Read the file in python using df = pandas.read.csv(filename)
Then pass a for loop
for i in range(len(df)):
    print("\n")
    print(i)
    print(df.iloc[i])
```

This code will give you all the tweets in a much readable format

Once you have all the sentences, analyze them, and create two lists to store index numbers of negative and not negative tweets. And once the list is finalized create two labels “negative” and “not negative” and update them in the label column.

You can use any other approaches that are convenient to you as well.

Thank you.