## 0.1 Question 1d

There are many ways we could choose to read tweets. Why might someone be interested in doing data analysis on tweets? Name a kind of person or institution which might be interested in this kind of analysis. Then, give two reasons why a data analysis of tweets might be interesting or useful for them. Answer in 2-3 sentences.

 ${\it Type\ your\ answer\ here,\ replacing\ this\ text.}$ 

## 0.1.1 Question 2e

What might we want to investigate further? Write a few sentences below.

## 0.1.2 Question 2f

We just looked at the top 5 most commonly used devices for each user. However, we used the number of tweets as a measure, when it might be better to compare these distributions by comparing *proportions* of tweets. Why might proportions of tweets be better measures than numbers of tweets?

## 0.1.3 Question 3b

Compare Cristiano's distribution with those of AOC and Elon Musk. In particular, compare the distributions before and after Hour 6. What differences did you notice? What might be a possible cause of that? Do the data plotted above seem reasonable?

 ${\it Type\ your\ answer\ here,\ replacing\ this\ text.}$ 

#### 0.1.4 Question 4a

Please score the sentiment of one of the following words, using your own personal interpretation. No code is required for this question!

- police
- order
- Democrat
- Republican
- gun
- dog
- $\bullet$  technology
- TikTok
- security
- $\bullet$  face-mask
- science
- climate change
- vaccine

What score did you give it and why? Can you think of a situation in which this word would carry the opposite sentiment to the one you've just assigned?

## 0.1.5 Question 4g

When grouping by mentions and aggregating the polarity of the tweets, what aggregation function should we use? What might be one drawback of using the mean?

# 0.1.6 Question 5a

Use this space to put your EDA code.

In [ ]: # perform your text analysis here

# 0.1.7 Question 5b

Use this space to put your EDA description.

Write your description here.