Project 4

Nathen Byford

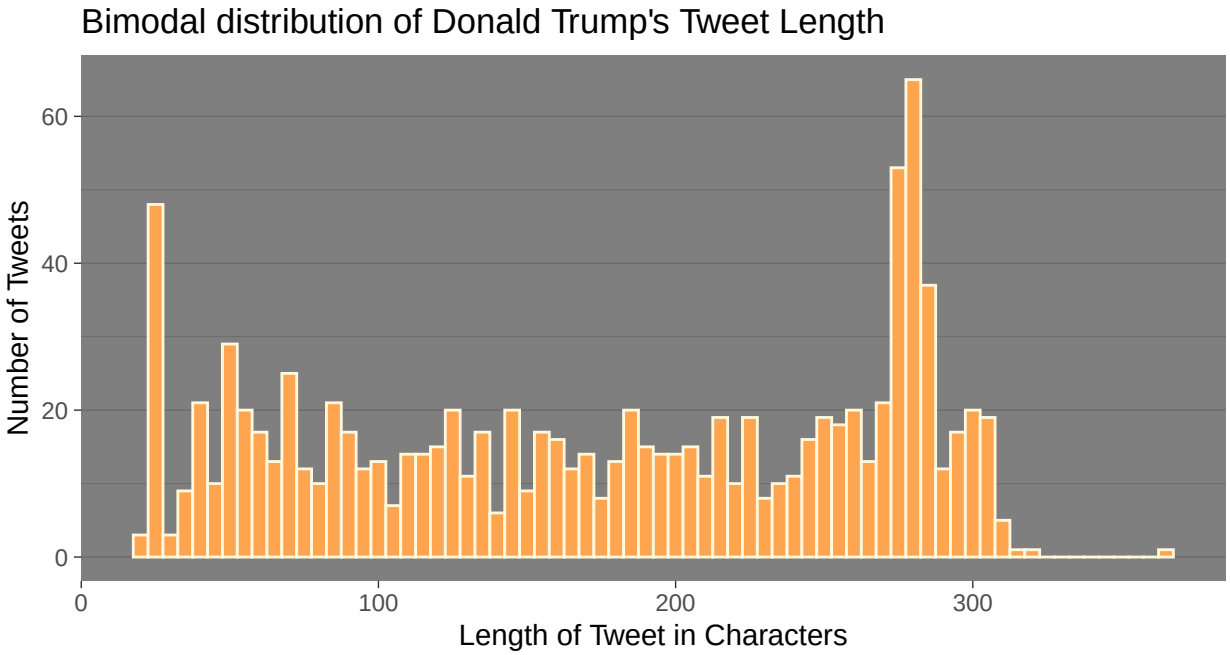
12/2/2021

Load data

```
dtweets <- read_rds(".//data//dtweets.rds")
```

Question 2

```
dtweets |>
  mutate(length = str_length(text)) |>
    ggplot() +
    geom_histogram(aes(x=length), fill = "tan1", color = "cornsilk1", binwidth =5) +
    #Color choices inspired by the color of Donald Trump's hair and face
    labs(title = "Bimodal distribution of Donald Trump's Tweet Length") +
    xlab("Length of Tweet in Characters")+
    ylab("Number of Tweets") +
    theme_dark() + #Theme chosen to emphasize color choice
    theme(panel.grid.major.x = element_blank(), panel.grid.minor.x = element_blank())
```



Question 3

```
good_grade <- "The correct thing here"
```

Question 4

```
mentions <- dtweets |>
  pull(text) |>
  str_extract_all(pattern = "\\@\\w{4,15}")

map(mentions, str_split, pattern = ", ") |>
  unlist() |>
  as_tibble() |>
  group_by(value) |>
  count() |>
  ungroup() |>
  arrange(desc(n))
## # A tibble: 198 x 2
##    value                n
##    <chr>            <int>
##  1 @realDonaldTrump   108
##  2 @FoxNews            18
##  3 @CDCgov             16
##  4 @POTUS              12
##  5 @foxandfriends       8
##  6 @WhiteHouse          8
##  7 @CMSGov              6
##  8 @nytimes             6
##  9 @SteveDaines         6
## 10 @Mike_Pence          5
## # ... with 188 more rows
```