Ncollin 607 HW1

Noah Collin

R Markdown

Noah Collin

NOAH.COLLIN35@login.cuny.edu

This is Noah's HW1 for 607. I'm using 538's Covid Polls data set available here: https://github.com/fivethirtyeight/covid-19-polls

The specific CSV I'm using describes the approval percentage of a President's performance in responding to Covid-19.

Below is a summary of the Polls CSV:

```
#setwd("")
polls <- read.csv("covid_approval_polls.csv")
summary(polls)
```

```
##   start_date          end_date           pollster           sponsor
##  Length:2809        Length:2809        Length:2809        Length:2809
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##   sample_size      population          party              subject
##  Min.   :    55   Length:2809        Length:2809        Length:2809
##  1st Qu.:   389   Class :character   Class :character   Class :character
##  Median :   640   Mode  :character   Mode  :character   Mode  :character
##  Mean   :  2379
##  3rd Qu.:  1226
##  Max.   :325970
##  NA's   :22
##   tracking           text              approve        disapprove
##  Mode :logical   Length:2809        Min.   : 1.00   Min.   : 1.00
##  FALSE:2559      Class :character   1st Qu.:30.00   1st Qu.:28.00
##  TRUE :242       Mode  :character   Median :42.00   Median :53.00
##  NA's :8                            Mean   :46.42   Mean   :48.48
##                                     3rd Qu.:66.00   3rd Qu.:63.00
##                                     Max.   :98.00   Max.   :98.00
##                                     NA's   :3       NA's   :15
##      url
##  Length:2809
##  Class :character
```

```
##  Mode  :character
##
##
##
##
```

Here are the top 5 rows of the uncleaned CSV:

```
head(polls)
```

```
##   start_date   end_date        pollster   sponsor sample_size population party
## 1 2020-02-02 2020-02-04          YouGov Economist        1500          a   all
## 2 2020-02-02 2020-02-04          YouGov Economist         376          a     R
## 3 2020-02-02 2020-02-04          YouGov Economist         523          a     D
## 4 2020-02-02 2020-02-04          YouGov Economist         599          a     I
## 5 2020-02-07 2020-02-09 Morning Consult                  2200          a   all
## 6 2020-02-07 2020-02-09 Morning Consult                   684          a     R
##   subject tracking
## 1   Trump    FALSE
## 2   Trump    FALSE
## 3   Trump    FALSE
## 4   Trump    FALSE
## 5   Trump    FALSE
## 6   Trump    FALSE
##
## 1                                                                    Do you approve or disapprove of
## 2                                                                    Do you approve or disapprove of
## 3                                                                    Do you approve or disapprove of
## 4                                                                    Do you approve or disapprove of
## 5 Do you approve or disapprove of the job each of the following is doing in handling the spread of co
## 6 Do you approve or disapprove of the job each of the following is doing in handling the spread of co
##   approve disapprove
## 1      42         29
## 2      75          6
## 3      21         51
## 4      39         25
## 5      57         22
## 6      88          4
##                                                                                                    url
## 1         https://d25d2506sfb94s.cloudfront.net/cumulus_uploads/document/73jqd6u5mv/econTabReport.pdf
## 2         https://d25d2506sfb94s.cloudfront.net/cumulus_uploads/document/73jqd6u5mv/econTabReport.pdf
## 3         https://d25d2506sfb94s.cloudfront.net/cumulus_uploads/document/73jqd6u5mv/econTabReport.pdf
## 4         https://d25d2506sfb94s.cloudfront.net/cumulus_uploads/document/73jqd6u5mv/econTabReport.pdf
## 5 https://morningconsult.com/wp-content/uploads/2020/02/200214_crosstabs_CORONAVIRUS_Adults_v4_JB.pdf
## 6 https://morningconsult.com/wp-content/uploads/2020/02/200214_crosstabs_CORONAVIRUS_Adults_v4_JB.pdf
```

##Subset of Data

```
library(dplyr)
```

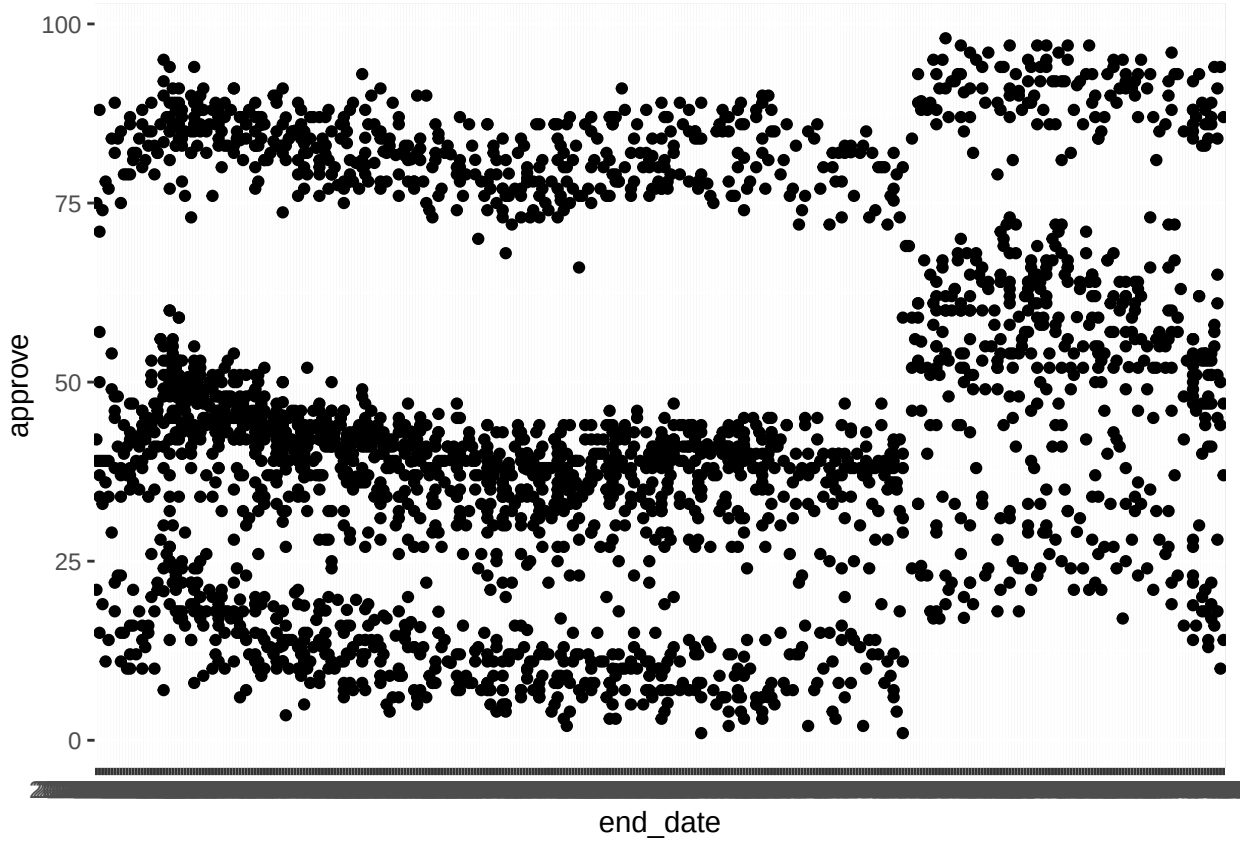
Here's just 4 columns of the data:

```
##     end_date approve disapprove subject
## 1 2020-02-04      42         29   Trump
## 2 2020-02-04      75          6   Trump
## 3 2020-02-04      21         51   Trump
## 4 2020-02-04      39         25   Trump
## 5 2020-02-09      57         22   Trump
## 6 2020-02-09      88          4   Trump
```

##Graph (I tried a few things that I couldn't get to work. I'd hoped to turn in something better but ran out of time.)

```
ggplot(data = temp) +
  geom_point(mapping = aes(x = end_date, y = approve))
```

```
## Warning: Removed 3 rows containing missing values (geom_point).
```




```
#+
 # facet_grid(subject == "Trump" ~ subject == "Biden")
```