Data 607 Week 1 Assignment

Justin Hsi

2/1/2020

Overview

The article aggregates the approval/disapproval ratings of donald trump's presidency from various poll sources. The various polls are given weights and are also graded in terms of their historical accuracy. Finally, at the bottom of the article you can see similar plots for past presidents. Article at https://projects.fivethirtyeight.com/trump-approval-ratings/

```
approve_pollist = read.csv(file = 'approval_polllist.csv')
head(approve_pollist)
```

```
##      president  subgroup modeldate startdate   enddate        pollster grade
## 1 Donald Trump All polls  2/1/2020 1/20/2017 1/22/2017          Gallup     B
## 2 Donald Trump All polls  2/1/2020 1/20/2017 1/22/2017 Morning Consult   B/C
## 3 Donald Trump All polls  2/1/2020 1/20/2017 1/24/2017           Ipsos    B-
## 4 Donald Trump All polls  2/1/2020 1/21/2017 1/23/2017          Gallup     B
## 5 Donald Trump All polls  2/1/2020 1/21/2017 1/25/2017           Ipsos    B-
## 6 Donald Trump All polls  2/1/2020 1/22/2017 1/24/2017          Gallup     B
##   samplesize population    weight influence approve disapprove adjusted_approve
## 1       1500          a 0.2416820         0    45.0       45.0         45.82995
## 2       1992         rv 0.6969837         0    46.0       37.0         44.56065
## 3       1632          a 0.1507855         0    42.1       45.2         42.81977
## 4       1500          a 0.2232337         0    45.0       46.0         45.82995
## 5       1651          a 0.1389668         0    42.3       45.8         43.01977
## 6       1500          a 0.2086590         0    46.0       45.0         46.82995
##   adjusted_disapprove multiversions tracking
## 1            43.48533                  TRUE
## 2            38.42233                    NA
## 3            44.06066                  TRUE
## 4            44.48533                  TRUE
## 5            44.66066                  TRUE
## 6            43.48533                  TRUE
##                                                                                 url
## 1                         http://www.gallup.com/poll/201617/gallup-daily-trump-job-approval.aspx
## 2 http://static.politico.com/9b/13/82a3baf542ae9018e5b6e1008379/170103-topline-politico-v3-kd.pdf
## 3                                                    http://polling.reuters.com/#poll/CP3_2/
## 4                         http://www.gallup.com/poll/201617/gallup-daily-trump-job-approval.aspx
## 5                                                    http://polling.reuters.com/#poll/CP3_2/
## 6                         http://www.gallup.com/poll/201617/gallup-daily-trump-job-approval.aspx
##   poll_id question_id createddate            timestamp
## 1   49253       77265   1/23/2017 08:19:14  1 Feb 2020
```

```
## 2   49249       77261   1/23/2017 08:19:14  1 Feb 2020
## 3   49426       77599    3/1/2017 08:19:14  1 Feb 2020
## 4   49262       77274   1/24/2017 08:19:14  1 Feb 2020
## 5   49425       77598    3/1/2017 08:19:14  1 Feb 2020
## 6   49236       77248   1/25/2017 08:19:14  1 Feb 2020
```

```
approve_pollist$modeldate = as.Date(approve_pollist$modeldate, format='%m/%d/%Y')
approve_pollist$startdate = as.Date(approve_pollist$startdate, format='%m/%d/%Y')
approve_pollist$enddate = as.Date(approve_pollist$enddate, format='%m/%d/%Y')
recent_polls = subset(approve_pollist, startdate > as.Date("2019-1-1"), select=c(subgroup, modeldate, s
recent_polls = recent_polls %>%
  rename(
    polled_population = population
  )
recent_polls$polled_population = recode_factor(recent_polls$polled_population, a = "all adults", rv = '
head(recent_polls)
```

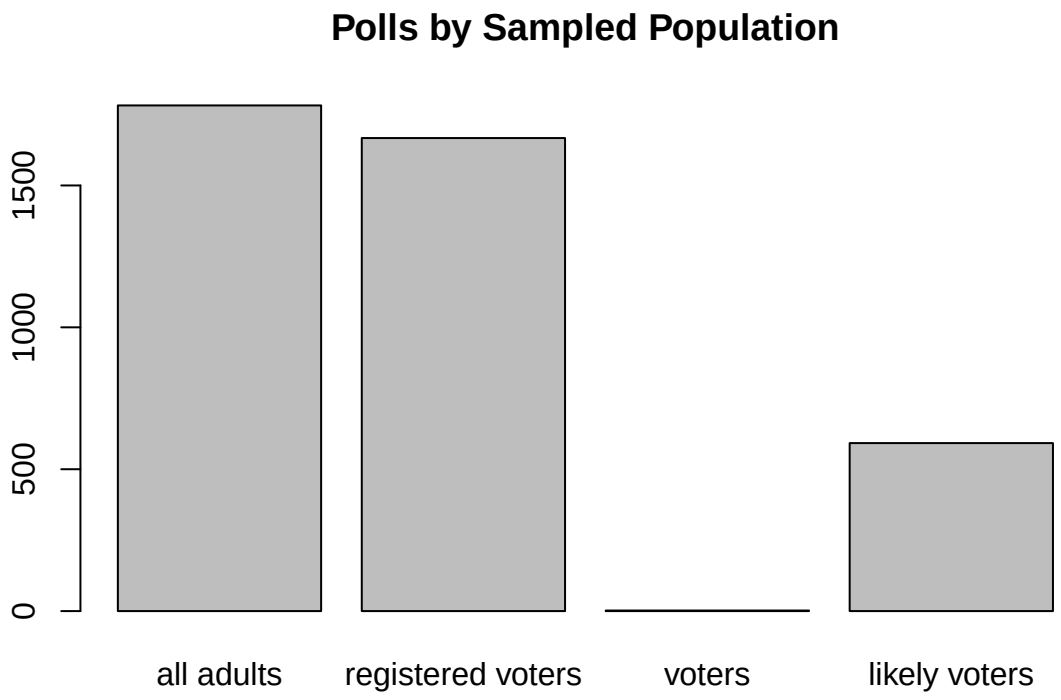
```
##       subgroup  modeldate  startdate    enddate
## 2820 All polls 2020-02-01 2019-01-02 2019-01-03
## 2821 All polls 2020-02-01 2019-01-02 2019-01-04
## 2822 All polls 2020-02-01 2019-01-03 2019-01-05
## 2823 All polls 2020-02-01 2019-01-03 2019-01-05
## 2824 All polls 2020-02-01 2019-01-03 2019-01-04
## 2826 All polls 2020-02-01 2019-01-02 2019-01-06
##                                     pollster grade samplesize
## 2820                                  YouGov    B-       1000
## 2821                                 HarrisX    C+       3003
## 2822                                  YouGov    B-       1000
## 2823                                 HarrisX    C+       3000
## 2824                   Public Policy Polling     B        658
## 2826 Rasmussen Reports/Pulse Opinion Research    C+       1500
##      polled_population     weight approve disapprove adjusted_approve
## 2820        all adults 0.11508766      41         52         41.75357
## 2821 registered voters 0.08019612      47         53         42.52428
## 2822        all adults 0.11564389      41         53         41.75357
## 2823 registered voters 0.07919949      48         52         43.52428
## 2824 registered voters 0.76291635      43         52         42.37991
## 2826     likely voters 0.08215774      46         53         40.10447
##      adjusted_disapprove
## 2820            54.40230
## 2821            52.36856
## 2822            55.40230
## 2823            51.36856
## 2824            52.33128
## 2826            54.62343
```

```
summary(recent_polls)
```

```
##       subgroup      modeldate            startdate
##  Adults   : 891   Min.   :2020-02-01   Min.   :2019-01-02
##  All polls:1698   1st Qu.:2020-02-01   1st Qu.:2019-03-19
##  Voters   :1454   Median :2020-02-01   Median :2019-06-22
##                   Mean   :2020-02-01   Mean   :2019-06-28
```

```
##                   3rd Qu.:2020-02-01   3rd Qu.:2019-09-28
##                   Max.   :2020-02-01   Max.   :2020-01-29
##
##     enddate                                                pollster
##  Min.   :2019-01-03   YouGov                                   :1358
##  1st Qu.:2019-03-22   HarrisX                                  : 702
##  Median :2019-06-25   Rasmussen Reports/Pulse Opinion Research: 542
##  Mean   :2019-07-01   Morning Consult                          : 415
##  3rd Qu.:2019-10-01   Ipsos                                    : 409
##  Max.   :2020-01-30   Gallup                                   :  44
##                       (Other)                                  : 573
##      grade        samplesize            polled_population     weight
##  B-     :1769   Min.   :  604   all adults       :1782    Min.   :0.0000
##  C+     :1278   1st Qu.: 1000   registered voters:1667    1st Qu.:0.1026
##  B/C    : 446   Median : 1500   voters           :   2    Median :0.1142
##  A/B    : 104   Mean   : 1801   likely voters    : 592    Mean   :0.3388
##  B      : 101   3rd Qu.: 2201                             3rd Qu.:0.2202
##         :  90   Max.   :19909                             Max.   :3.0935
##  (Other): 255
##     approve        disapprove    adjusted_approve adjusted_disapprove
##  Min.   :34.00   Min.   :45.00   Min.   :35.15    Min.   :45.78
##  1st Qu.:41.00   1st Qu.:52.00   1st Qu.:40.52    1st Qu.:52.40
##  Median :43.00   Median :53.00   Median :41.86    Median :53.64
##  Mean   :43.29   Mean   :53.19   Mean   :41.78    Mean   :53.62
##  3rd Qu.:45.00   3rd Qu.:55.00   3rd Qu.:43.10    3rd Qu.:54.69
##  Max.   :53.00   Max.   :65.00   Max.   :49.45    Max.   :60.77
##
```

```r
barplot(table(recent_polls$polled_population), main = "Polls by Sampled Population")
```



Findings and Recommendations

The analysis was on Trump's approval ratings across the various polling populations, but you could use this data to try and get a sense of the next presidential election results. If you assume that approval of Trump correlates to voting for Trump, then you might redo this analysis of only likely voters or registered voters and see how his approval rating stands. You could monitor this up until the next presidential election and it may give you a sense of sentiment on Trump being President for another term.