POL390 Week 2 Monday

Getting Started

In today's lab, we'll create some univariate graphs using base R. Univariate means one variable. Next time we'll extend our graphing skills by learning to use ggplot2. You can choose which variables you want to graph from the following variables:

- `V202557`: Was the coronavirus (covid-19) developed intentially in a lab? Recode this variable so that 0 = No and 1 = Yes.
- `V201393`: Were the limits placed on activities due to covid-19 too strict or not? Recode this variable so that 1 = Not nearly strict enough and 5 = Far too strict.
- `V202310`: How important should science be for decisions about covid? Recode this variable so that 1 = Not at all important and 5 = Extremely important.
- `V201624`: Has anyone in your household tested positive for covid? Recode this variable so that 0 = No and 1 = Yes. //
- `V202187`: Feelings toward the Center for Disease Control (CDC) on a scale from 0 (very cold feelings) to 100 (very warm feelings). This variable needs cleaning.
- `V202158`: Feelings toward Anthony Fauci on a scale from 0 (very cold feelings) to 100 (very warm feelings). This variable needs cleaning.

First, check/ set your working directory and load the data:

```
# Check/ set working directory
getwd()
```

```
## [1] "/Users/hyunmyungchoi/study_R/POL390"
```

```
setwd("/Users/hyunmyungchoi/study_R/POL390")
# Load data
dat <- read.csv("anes_timeseries_2020_csv_20220210.csv")
dim(dat)
```

```
## [1] 8280 1771
```

1. Creating a Bar Graph

Professor example of a bar graph

First, the prof will show you how to create a basic bar graph (or bar chart) in base R that shows the distribution of vote choice in the 2020 US election. First we'll do some recoding in base R. Bar graphs are useful when you want to show the distribution of a factor variable on the x-axis. We will recode the variable so that it is a dummy variable (two-category) where voted for Biden = 0 and voted for Trump = 1 and any other response is coded as missing. factor vari - for category, text & number <-> numeric vari - numb only

```
## str => structure
# Look at the data and recode
table(dat$V202073)
```

```
##
##   -9   -7   -6   -1    1    2    3    4    5    7    8   11   12
##   53   77  754 1497 3267 2462   69   23   56    1    3    2   16
```

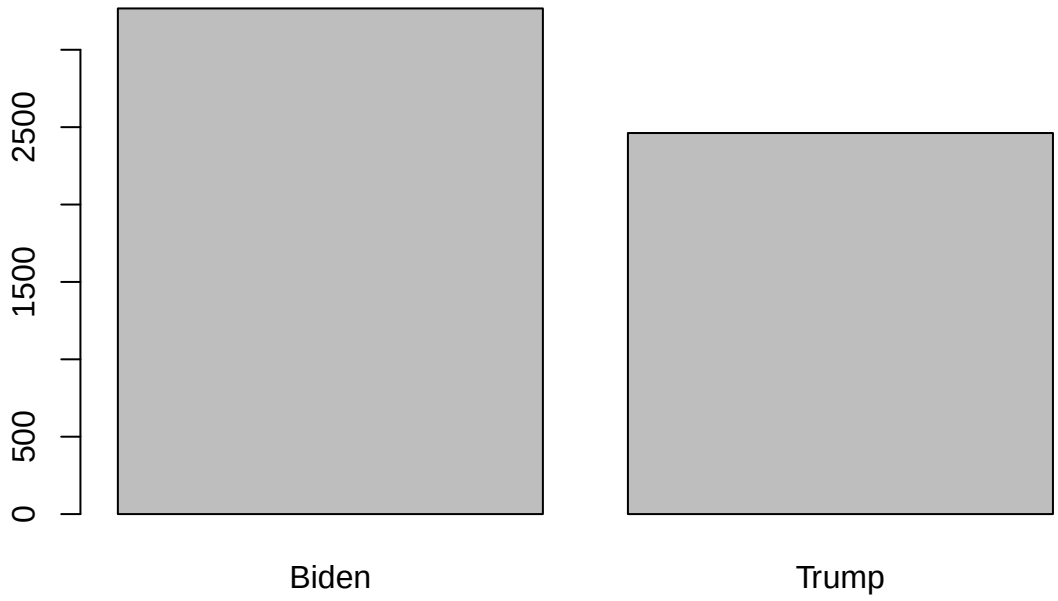
```
str(dat$V202073)
```

```
##  int [1:8280] -1 3 1 1 2 1 2 -1 -1 1 ...
```

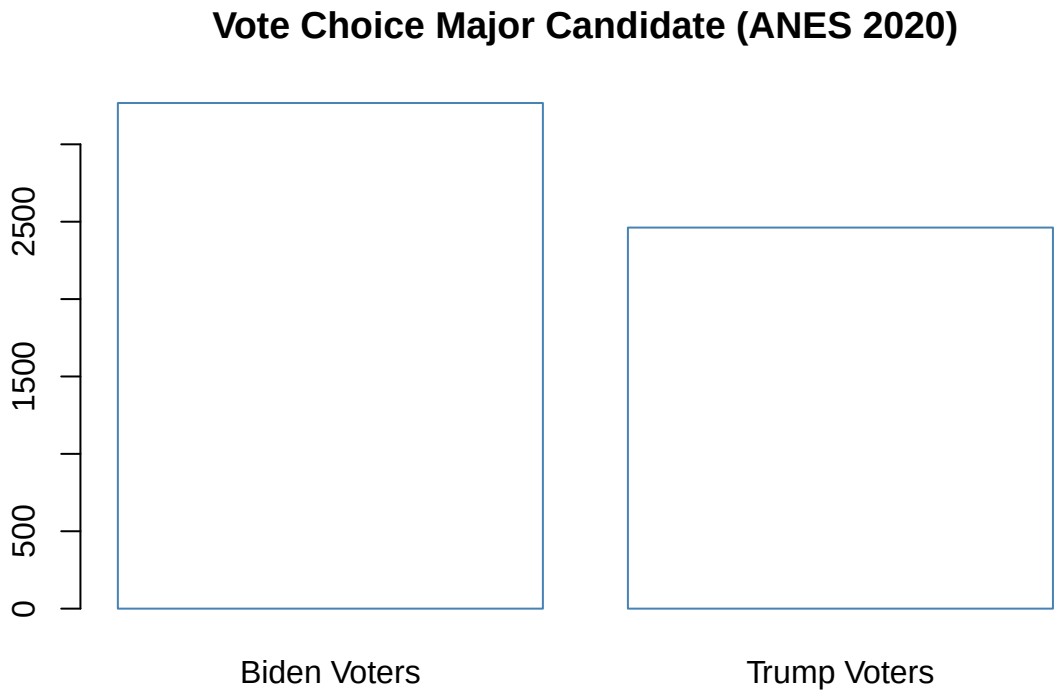
```
# Recode V202073 POST: FOR WHOM DID R VOTE FOR PRESIDENT
dat$vote_trump <- dat$V202073
dat$vote_trump[dat$V202073 < 1 | dat$V202073 > 2] <- NA
dat$vote_trump[dat$V202073 == 1] <- 0
dat$vote_trump[dat$V202073 == 2] <- 1
str(dat$vote_trump)
```

```
##  num [1:8280] NA NA 0 0 1 0 1 NA NA 0 ...
```

```
dat$vote_trump <- factor(dat$vote_trump, labels = c("Biden", "Trump"), levels = c(0, 1))
# Simple barchart
barplot(table(dat$vote_trump))
```



```r
# Adding some details:
### main -> title, names.arg -> x axis
barplot(table(dat$vote_trump), main ="Vote Choice Major Candidate (ANES 2020)",
        names.arg = c("Biden Voters", "Trump Voters"),
        col = "white", border ="steelblue")
```



Student practice creating a bar graph

Now it's your turn!

```r
# Code here
table(dat$V202557)
```

```
##
##   -9   -7   -6   -5    1    2
##   67   77  754  102 3374 3906
```

```r
str(dat$V202557)
```

```
##  int [1:8280] 1 1 2 1 1 2 1 2 2 2 ...
```

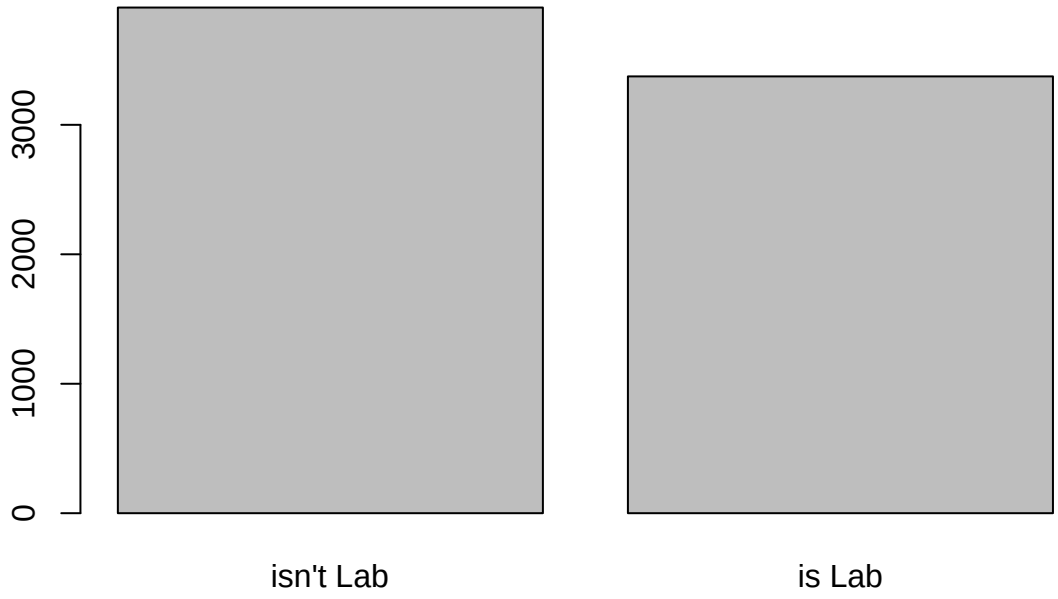
```
dat$isLab <- dat$V202557
dat$isLab[dat$V202557 < 1] <- NA
dat$isLab[dat$V202557 == 2] <- 0
dat$isLab[dat$V202557 == 1] <- 1
str(dat$isLab)
```

```
##  num [1:8280] 1 1 0 1 1 0 1 0 0 0 ...
```

```
summary(dat$isLab)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##  0.0000  0.0000  0.0000  0.4635  1.0000  1.0000    1000
```

```
dat$isLab <- factor(dat$isLab, labels = c("isn't Lab", "is Lab"), levels = c(0,1))
## 46% is proportion of "is lab"
barplot(table(dat$isLab))
```



2. Creating a Histogram

Professor example of a histogram

First, the prof will show you how to create a basic histogram in base R that shows the distribution of feelings toward the U.S. Immigration and Customs Enforcement (ICE) Agency. First we have to do some data cleaning.

```
# Look at the data and recode
# V202182 POST: FEELING THERMOMETER: IMMIGRATION AND CUSTOMS ENFORCEMENT (ICE) AGENCY
##higher value = more warm
str(dat$V202182)
```

```
##  int [1:8280] 100 50 0 85 95 40 70 0 60 0 ...
```

```
table(dat$V202182)
```

```
##
##   -9   -7   -6   -5   -4    0    1    2    4    5    7    8    9   10   12   13
##   82   77  754   16    1  876    8    6    1   30    1    2    2   84    1    1
##   15   17   18   20   22   23   25   27   30   35   40   43   45   48   49   50
##  453    3    3   73    1    1   47    1  512   27  666    1   39    1    6 1277
##   51   55   57   58   60   64   65   66   67   68   69   70   72   75   76   77
##    1   31    2    2  637    1   34    1    1    1    1  661    1   70    1    1
##   78   80   81   84   85   86   87   88   90   92   93   95   96   98   99  100
##    4   83    1    1  696    3    1    3   95    1    1   28    1    3    3  572
##  998  999
##    4  281
```

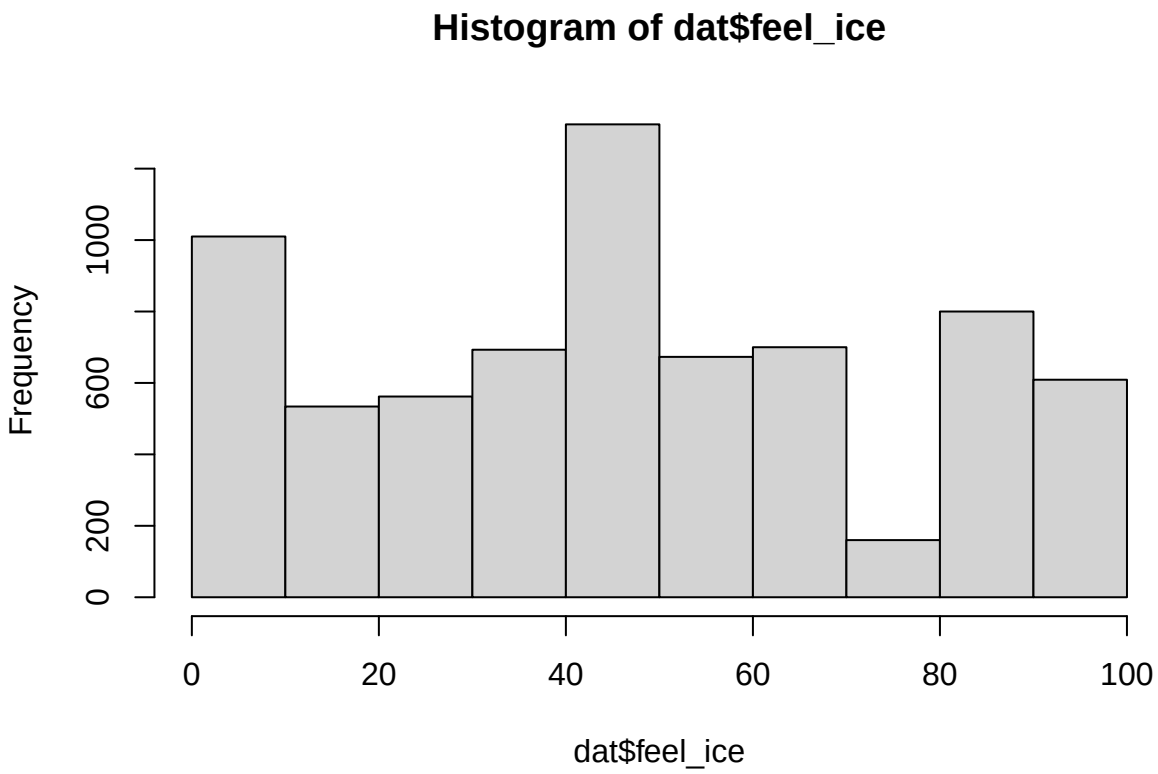
```
summary(dat$V202182)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   -9.00   15.00   50.00   75.98   70.00  999.00
```

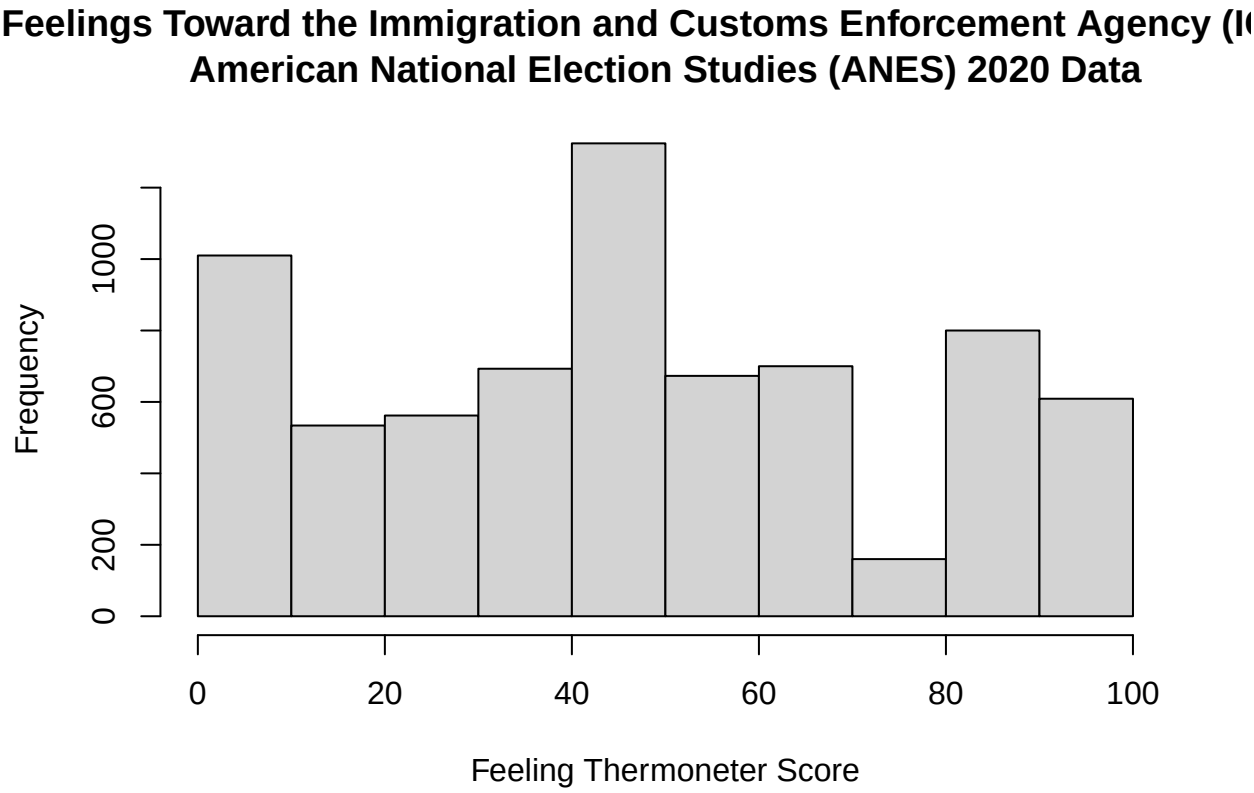
```
dat$feel_ice <- dat$V202182
dat$feel_ice[dat$V202182 < 0 | dat$V202182 > 100] <- NA
summary(dat$feel_ice)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##    0.00   30.00   50.00   49.58   70.00  100.00    1215
```

```
# Graph histogram
hist(dat$feel_ice)
```




```
hist(dat$feel_ice,
     main = "Feelings Toward the Immigration and Customs Enforcement Agency (ICE). \nAmerican National
     xlab = "Feeling Thermoneter Score")
```



Student practice creating a histogram

Now it's your turn! Pick a numeric variable to graph using a histogram in base R.

```
# Code here
str(dat$V202158)
```

```
##  int [1:8280] 50 50 100 100 0 85 0 -9 30 50 ...
```

```
table(dat$V202158)
```

```
##
##   -9   -7   -6   -5   -4    0    1    2    3    4    5    6    7    8    9   10
##  126   77  754    9    1  424    5    6    2    3   18    3    2    1    4   36
##   12   15   16   20   23   25   30   35   39   40   45   48   49   50   51   52
##    1  236    1   25    1   25  292   12    1  316   16    1    6 1095    1    1
##   53   55   59   60   65   69   70   72   75   77   78   79   80   83   84   85
##    1   14    1  488   32    1  629    1   85    1    1    1  109    1    1 1060
##   86   87   88   89   90   92   95   96   97   98   99  100  998  999
##    4    1    7    4  222    1  114    2    2   12   15 1949    5   15
```

```
summary(dat$V202158)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   -9.00   40.00   70.00   61.48   95.00  999.00
```

```
dat$feelsIce <- dat$V202158
dat$feelsIce[dat$V202158 < 0 | dat$V202158 > 100] <- NA
summary(dat$feelsIce)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##    0.00   50.00   70.00   67.92  100.00  100.00     987
```

```
hist(dat$feelsIce,
     main = "Feelings toward Anthony Fauci",
     xlab = "very cold toward 0, very warm toward 100")
```