HW3

Nikhil Gopal

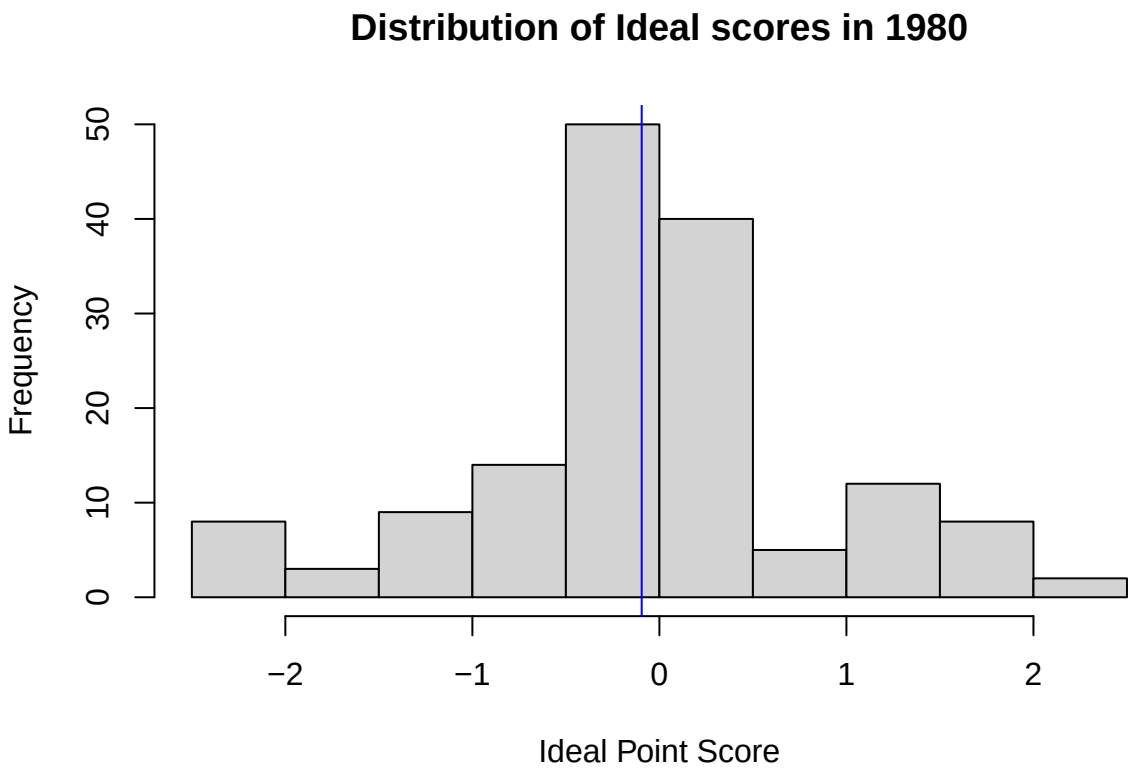
10/17/2020

Question 1

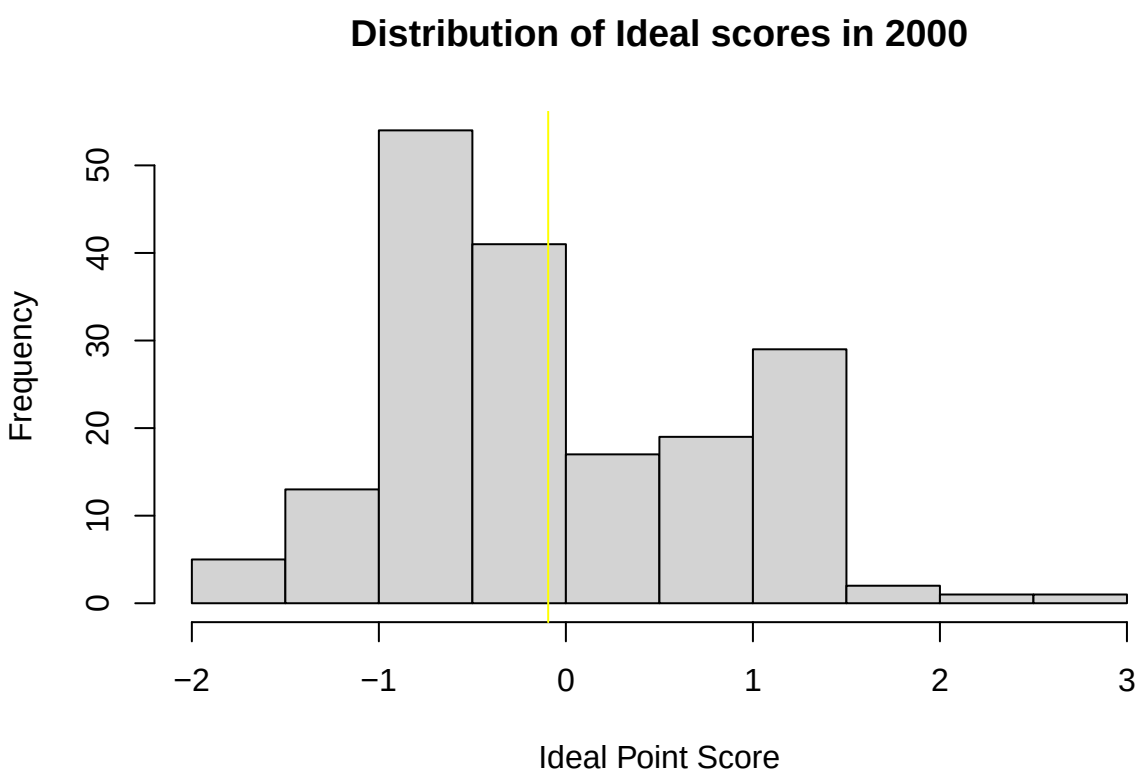
```
setwd("C:/Users/d/Google Drive/Notability/Research and Design_ Data Analysis/hw 3")
data <- read.csv("unvoting.csv")

#subset data
data1980 <- subset(data, Year==1980)

data2000 <- subset(data, Year==2000)

#1980 histogram and quantiles
hist(data1980$idealpoint, xlab = "Ideal Point Score", main = "Distribution of Ideal scores in 1980")
quantile(data1980$idealpoint)
```

```
##          0%         25%         50%         75%        100%
## -2.35749800 -0.34559800 -0.09465057  0.34365995  2.31760700
```

```
 #add median to the plot
abline(v = median(data1980$idealpoint), col = "blue")
```




```
#2000 histogram and quantiles
hist(data2000$idealpoint, xlab = "Ideal Point Score", main = "Distribution of Ideal scores in 2000")
abline(v = median(data1980$idealpoint), col = "yellow")
```




```
quantile(data2000$idealpoint)
```

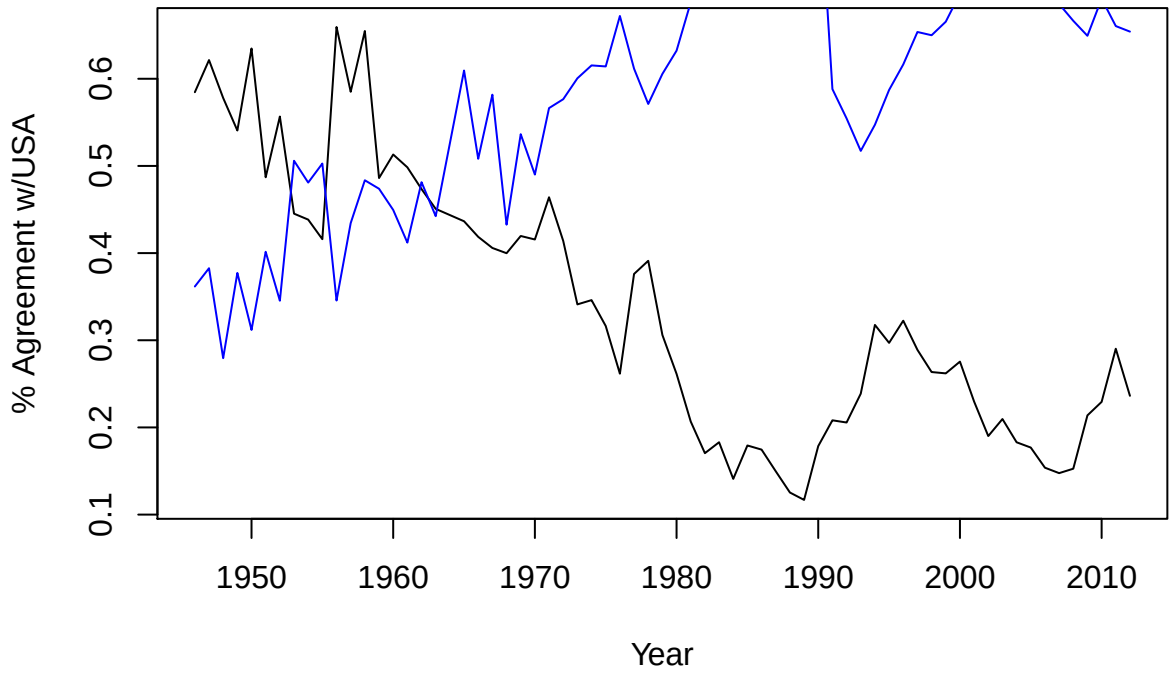
```
##         0%        25%        50%        75%       100%
## -1.7018510 -0.7021018 -0.3514302  0.6200154  2.6105410
```

The 1980 scores appear to be relatively normally distributed, while the 2000 scores appear to be skewed towards the negative side. The median for 1980 was about -0.1 while for 2000 it was -0.35. Looking at the quantiles, we can tell that the data became more spread out from 1980 - 2000 (including to the liberal side), but the median decreased and the data became skewed negative, indicating that countries became less liberal as time passed after the dissolution of the USSR.

Question 2

```
#Question 2

#get the avg agreement w/USA for each year
pctagreeUSA <- tapply(data$PctAgreeUS, data$Year, mean, na.rm = TRUE)
#get the unique years to, store it for graphing later
unique_years <- unique(data$Year)

plot(unique_years, pctagreeUSA, xlab = "Year", ylab = "% Agreement w/USA", type = 'l')

#agreement for russia
pctagreeRussia <- tapply(data$PctAgreeRUSSIA, data$Year, mean, na.rm = TRUE)
```

```
#add Russia to the graph
lines(unique_years, pctagreeRussia, type = "l", col = "blue")
```




```
#avg agreement w/US + Russia by country
agreement_USA_by_country <- tapply(data$PctAgreeUS, data$CountryName, mean, na.rm = TRUE)
agreement_Russia_by_country <- tapply(data$PctAgreeRUSSIA, data$CountryName, mean, na.rm = TRUE)

#print out consistently pro US/Russian countries in the markdown file
print(sort(agreement_Russia_by_country[1:4]))
```

```
##     Andorra Afghanistan     Algeria     Albania
##   0.6523980   0.6729391   0.7223443   0.7846117
```

```
print(tail(sort(agreement_USA_by_country), n = 6))
```

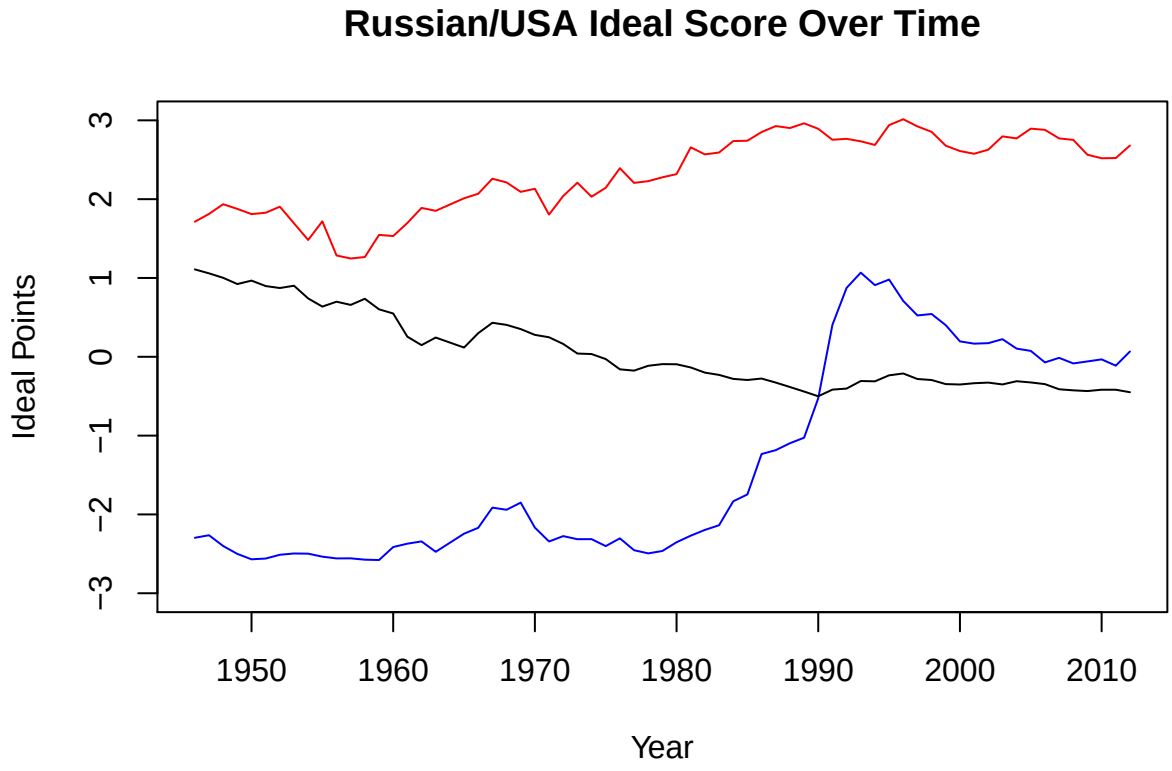
```
## Federated States of Micronesia                          Israel
##                      0.5937617                       0.6399077
##                         Taiwan                  United Kingdom
##                      0.6430106                       0.6521001
##                          Palau         United States of America
##                      0.7356335                       1.0000000
```

The United States appears to get more isolated over time. The plot shows that the Russia % agreement is basically a reflection of the American % agreement, and that the Russian one has been rising over time

while the American one has been falling. The blue line is the Russian line and the black represents the US. The top 5 countries that vote most consistently with Russia are Andorra, Afghanistan, Algeria and Albania. The top 5 countries that most consistently agree with the US are The Federated States of Micronesia, Israel, Taiwan, United Kingdom and Palau.

Question 3

```
#Question 3

plot(unique_years, data$idealpoint[data$CountryAbb == "USA"], ylim = c(-3,3), xlab = "Year", ylab = "Id
#add Russia to the graph
lines(unique_years, data$idealpoint[data$CountryAbb == "RUS"], col = "blue")

#add median of all countries ideal points to the graph
lines(unique_years, tapply(data$idealpoint, data$Year, median, ))
```

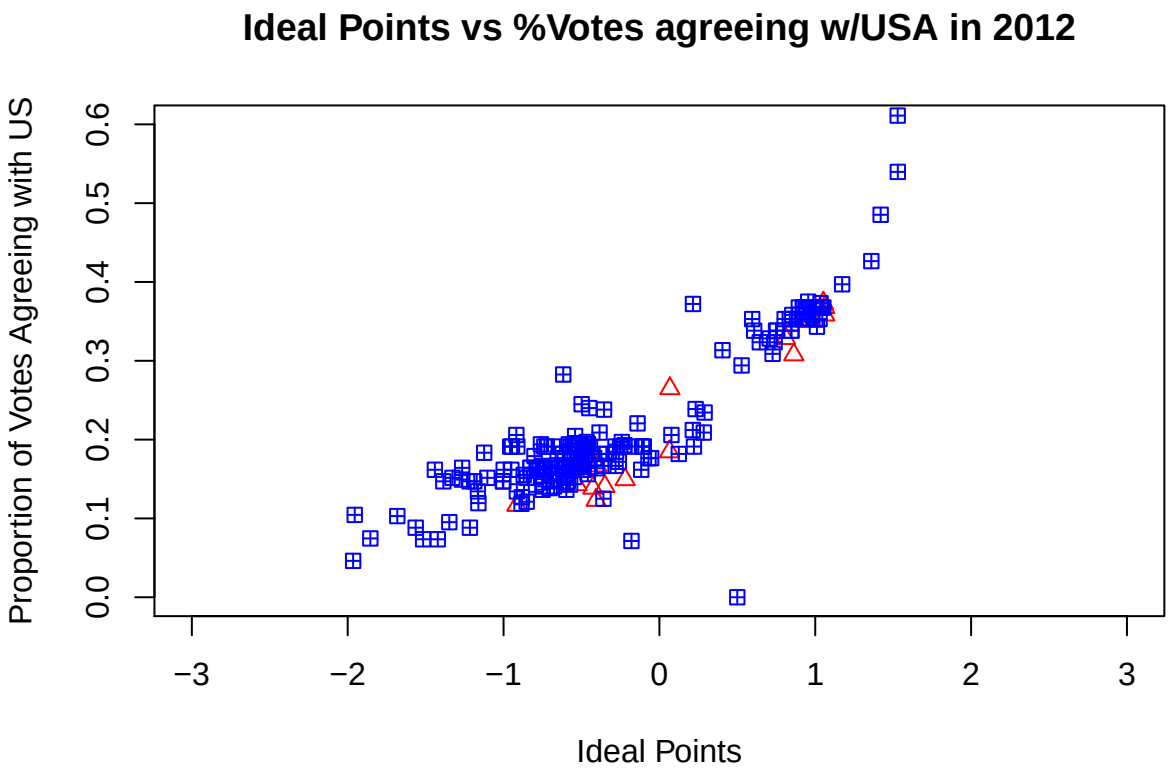


In this graph, the red line represents American ideal point score over time, blue the Russian and the black is the median ideal point score of all countries. The American score has increased slightly over time from about 2 to 2.5, the Russian has increased dramatically from about -2.5 to 0 and the median of all countries has decreased slightly from 1 to about 0. This does change our analysis of the previous question. Countries are becoming ideologically closer to Russia because Russian ideology is becoming closer to American ideology than it has been in the past.

Question 4

```
#Question 4
```

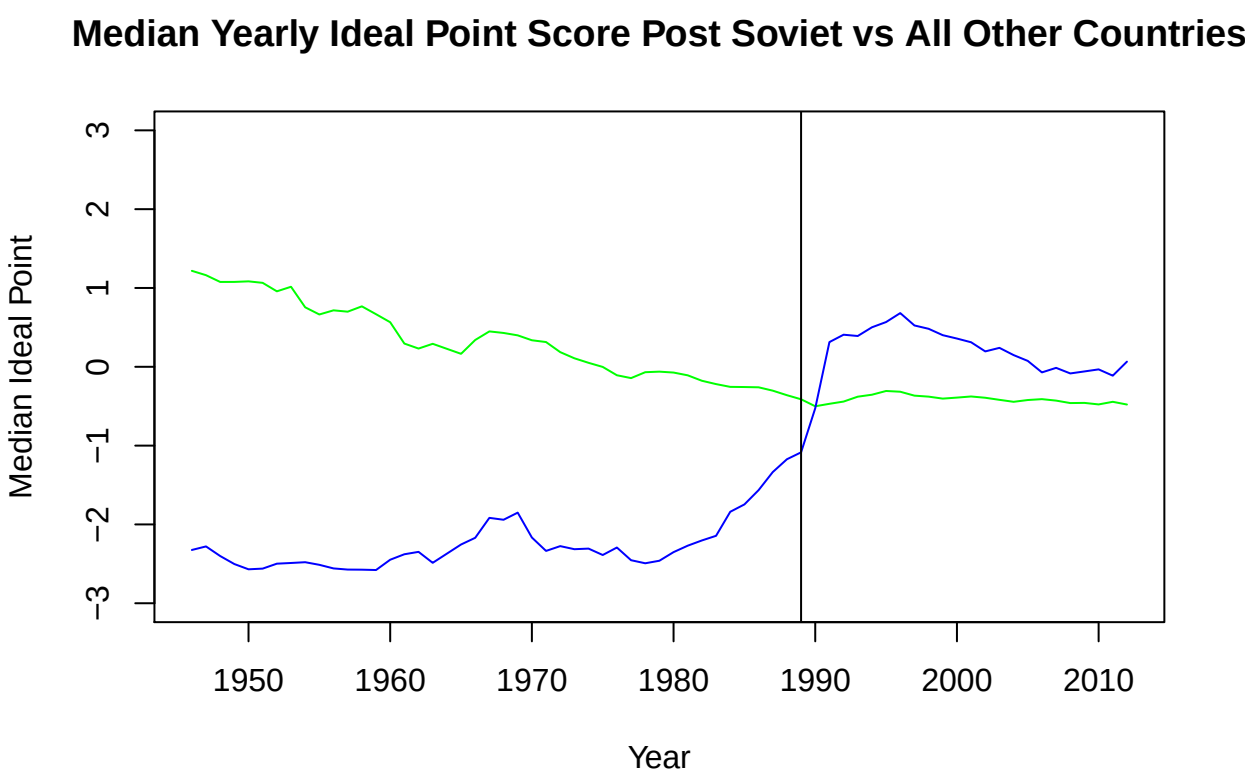
```r
#vector containing each country
postUSSRcountries <- c("Estonia", "Latvia", "Lithuania", "Belarus", "Moldova", "Ukraine", "Armenia", "A


#subset to only data in year 2012
data2012 <- subset(data, Year == 2012)
#subset again for only post soviet countries
data2012postUSSR <- subset(data2012, CountryName %in% postUSSRcountries)
plot(data2012$idealpoint[data2012$CountryName %in% postUSSRcountries], data2012postUSSR$PctAgreeUS, pch
     xlim = c(-3, 3), ylim = c(0, 0.6), col = "red",
     xlab = "Ideal Points", ylab = "Proportion of Votes Agreeing with US",
     main = "Ideal Points vs %Votes agreeing w/USA in 2012")

#vector with all country names
all_country_names <- unique(data2012$CountryName)
#empty vector to hold all country names minus the USSR country names
all_countries_not_post_USSR <- c()

#for loop thru vector, add the countries that aren't post USSR countries to new vector
for (country in all_country_names) {
  if(country %in% postUSSRcountries){
  }else{
    all_countries_not_post_USSR <- c(all_countries_not_post_USSR, country)
  }
}

#add other countries to the graph, need to do the above steps so that points aren't plotted twice
points(data2012$idealpoint[data2012$CountryName %in% all_countries_not_post_USSR], data2012$PctAgreeUS[
```



The red triangles represent the post soviet states and the blue squares represent all other countries. There is a positive correlation between ideal points score and agreement with the USA. The post soviet countries tend to fall in the middle of the graph (between -1/1 ideal score), where as other countries are more widely distributed in ideal score and %agreement with the US. Addtionally, the post soviet countries proportion of agreement with the US ranges from 0.1 to 0.4, but there is nonetheless a strong positive correlation for both post soviet and other countries between ideal points and proportion of votes agreeing with the US.

Question 5

```
#Question 5

#plot median ideal score by year for all countries except post soviet ones
plot(unique_years,
     tapply(data$idealpoint[data$CountryName %in% all_countries_not_post_USSR],
            data$Year[data$CountryName %in% all_countries_not_post_USSR],
            median, na.rm = TRUE),
     type = "l", ylim = c(-3, 3), xlab = "Year",
     ylab = "Median Ideal Point", col = "green",
     main = "Median Yearly Ideal Point Score Post Soviet vs All Other Countries")

#add post soviet countries to the graph
lines(unique_years, tapply(data$idealpoint[data$CountryName %in% postUSSRcountries], data$Year[data$Cou

#add line at 1989 to indicate fall of Berlin wall
abline(v=1989)
```



Here, the green line represents the median ideal point score for all countries (except the post soviet ones), while the blue line represents the post soviet countries. After the 1989 collapse of the Berlin wall and the decline of Communism in Eastern Europe, we see that the post Soviet Countries immediately increase their ideal point scores, while the rest of the world slowly continues on a downward trend. By about 1991, the post soviet countries actually surpass the world average. We can safely conclude that the post soviet countries have become more liberal since the fall of Communism.