

221003 Data Assinment 1

Jongrak

2022-10-03

```
library(tidyverse)

## — Attaching packages ————— tidyverse 1.3.
1 —

## ✓ ggplot2 3.3.6   ✓ purrr 0.3.4
## ✓ tibble 3.1.7   ✓ dplyr 1.0.9
## ✓ tidyr 1.2.0    ✓ stringr 1.4.0
## ✓ readr 2.1.2    ✓ forcats 0.5.1

## — Conflicts ————— tidyverse_conflicts
() —

## ✗ dplyr::filter() masks stats::filter()
## ✗ dplyr::lag()   masks stats::lag()

library(dplyr)
```

Part 1. Social Pressure and Voter Turnout

Default Setting

```
# set working directory
setwd("~/Library/Mobile Documents/com~apple~CloudDocs/Study/Univ/Lecture/2022-2/POLI223 POLITICAL METHODOLOGY/Assignment/Data Assinment 1")

# clear the environment and console
rm(list=ls())
```

Load the data

```
social <- read.csv("social.csv")
```

Data Overview

```
dim(social)

## [1] 305866    6

head(social)

##   sex yearofbirth primary2004 messages primary2006 hhsize
## 1  male      1941         0 Civic Duty         0      2
## 2 female      1947         0 Civic Duty         0      2
## 3  male      1951         0 Hawthorne          1      3
## 4 female      1950         0 Hawthorne          1      3
```

```
## 5 female      1982      0 Hawthorne      1      3
## 6  male      1981      0   Control      0      3

summary(social)

##      sex      yearofbirth  primary2004      messages
## Length:305866   Min.   :1900   Min.   :0.0000   Length:305866
## Class :character 1st Qu.:1947   1st Qu.:0.0000   Class :character
## Mode  :character Median :1956   Median :0.0000   Mode  :character
##           Mean  :1956   Mean  :0.4014
##           3rd Qu.:1965   3rd Qu.:1.0000
##           Max.  :1986   Max.  :1.0000
## primary2006      hhsizes
## Min.   :0.0000   Min.   :1.000
## 1st Qu.:0.0000   1st Qu.:2.000
## Median :0.0000   Median :2.000
## Mean   :0.3122   Mean   :2.184
## 3rd Qu.:1.0000   3rd Qu.:2.000
## Max.   :1.0000   Max.   :8.000
```

Question 1.

Calculate the average voter turnout rates in the 2006 Primary election separately for four different groups that received different messages (i.e. Civic duty, Neighbors, Hawthorne, Control).

```
# We can calculate the average voter turnout rates by using 'tapply' function.
tapply(social$primary2006, social$messages, mean, na.rm = TRUE)

## Civic Duty   Control Hawthorne Neighbors
## 0.3145377 0.2966383 0.3223746 0.3779482
```

Question 2.

How would you evaluate the causal effects of receiving different messages on voters' decision to go out to vote? Which type of message is the most effective in increasing voter turnout? Can we make a causal claim based on the analysis? Why? Or why not?

```
average_turnout_gap <- tapply(social$primary2006 - social$primary2004, social$messages, mean, na.rm = TRUE)

average_turnout_gap

## Civic Duty   Control Hawthorne Neighbors
## -0.08490764 -0.10370053 -0.08085541 -0.02871653
```

Q.2-1

How would you evaluate the causal effects of receiving different messages on voters' decision to go out to vote?

We can compare the differences of the average turnout rates between 'primary 2004' and 'primary2006' among the groups

Q.2-2

Which type of message is the most effective in increasing voter turnout?

Comparing only the average turnout, we can say the message of 'Neighbors' is most effective. Although the overall turnout rates has declined, the rates of all four groups have decreased less than the one of control group. Also, the turnout in group 'Neighbors' decreased the least among those four groups.

Q.2-3

Can we make a causal claim based on the analysis?

However, we cannot easily conclude which type of message is the most effective in 'increasing' voter turnout in that overall turnout rates has declined. We cannot analyze whether these messages affect on the increase or decrease of the rates. Probably, we can say all four messages affect on 'decreasing' rates, and the effect in group 'Neighbors' is most least.

In addition, this is just a comparison of only the average value, and the effects of other factors - such as gender, household size, and age - should be considered. We need to analyze it with other methods such as linear regression.

Part 2. Voting in the UN General Assembly

Default Setting

```
# set working directory
setwd("~/Library/Mobile Documents/com~apple~CloudDocs/Study/Univ/Lecture/2022-2/POLI223 POLITICAL METHODOLOGY/Assignment/Data Assinment 1")

# clear the environment and console
rm(list=ls())

# Load the package
library(foreign)
```

Load the data

```
un <- read.dta("UNVoting.dta")
```

Data Overview

```
names(un)

## [1] "countryname" "year" "idealpoint"
## [4] "pctagreeus" "pctagreerussia" "pctagreebrazil"
## [7] "pctagreechina" "pctagreeindia" "pctagreeisrael"
## [10] "region" "KOFindex" "foreign_aid_us"
## [13] "foreign_aid_notus" "us_trade" "gdp_pc"

dim(un)

## [1] 9505 15

head(un)

## countryname year idealpoint pctagreeus pctagreerussia
## 1 United States of America 1946 1.742829 1 0.2142857
## 2 United States of America 1947 1.750921 1 0.2631579
## 3 United States of America 1948 1.886371 1 0.1274510
## 4 United States of America 1949 1.798327 1 0.1111111
## 5 United States of America 1950 1.725135 1 0.1730769
## 6 United States of America 1951 1.765150 1 0.1200000
## pctagreebrazil pctagreechina pctagreeindia pctagreeisrael region
## 1 0.6428571 NA 0.4761905 NA Northern America
## 2 0.8421053 NA 0.2972973 NA Northern America
## 3 0.7766990 NA 0.3700000 0.1666667 Northern America
## 4 0.5396825 NA 0.3650794 0.5161290 Northern America
## 5 0.8113208 NA 0.5094340 0.6041667 Northern America
## 6 0.6400000 NA 0.3600000 0.6521739 Northern America
## KOFindex foreign_aid_us foreign_aid_notus us_trade gdp_pc
## 1 NA NA NA NA NA
## 2 NA NA NA NA NA
## 3 NA NA NA NA NA
## 4 NA NA NA NA NA
## 5 NA NA NA NA NA
## 6 NA NA NA NA NA

summary(un)
```

```

## countryname      year      idealpoint      pctagreeus
## Length:9505      Min.   :1946 Min.   :-2.7574 Min.   :0.0000
## Class :character  1st Qu.:1973 1st Qu.: -0.6330 1st Qu.:0.1419
## Mode  :character  Median :1988 Median :-0.1618 Median :0.2432
##           Mean  :1986 Mean  : 0.0000 Mean  :0.2973
##           3rd Qu.:2002 3rd Qu.: 0.7930 3rd Qu.:0.3913
##           Max.   :2014 Max.   : 3.0603 Max.   :1.0000
##           NA's   :1
## pctagreerussia  pctagreebrazil  pctagreechina  pctagreeindia
## Min.   :0.0000 Min.   :0.0000 Min.   :0.0000 Min.   :0.0000
## 1st Qu.:0.5128 1st Qu.:0.6154 1st Qu.:0.6295 1st Qu.:0.5270
## Median :0.6547 Median :0.8000 Median :0.8276 Median :0.7572
## Mean   :0.6213 Mean   :0.7324 Mean   :0.7547 Mean   :0.6884
## 3rd Qu.:0.7388 3rd Qu.:0.8784 3rd Qu.:0.8857 3rd Qu.:0.8378
## Max.   :1.0000 Max.   :1.0000 Max.   :1.0000 Max.   :1.0000
## NA's   :5      NA's   :1      NA's   :2089 NA's   :1
## pctagreeisrael  region      KOFindex  foreign_aid_us
## Min.   :0.0000 Length:9505      Min.   : 9.56 Min.   : 0.001
## 1st Qu.:0.1944 Class :character 1st Qu.:32.47 1st Qu.: 5.936
## Median :0.3254 Mode  :character Median :43.07 Median : 26.751
## Mean   :0.3530           Mean  :45.60 Mean  :116.931
## 3rd Qu.:0.4722           3rd Qu.:56.13 3rd Qu.: 80.391
## Max.   :1.0000           Max.   :92.77 Max.   :10751.479
## NA's   :113           NA's   :3482 NA's   :6061
## foreign_aid_notus  us_trade      gdp_pc
## Min.   : 0.0 Min.   : -18.0 Min.   : 62.24
## 1st Qu.: 62.7 1st Qu.: 39.9 1st Qu.: 488.86
## Median : 236.0 Median : 254.0 Median : 1734.70
## Mean   : 762.4 Mean   : 6043.9 Mean   : 5712.66
## 3rd Qu.: 713.5 3rd Qu.: 1693.0 3rd Qu.: 6364.26
## Max.   :58680.4 Max.   :575191.0 Max.   :116772.66
## NA's   :3666      NA's   :2168      NA's   :2852

```

Question 1

Let's explore the `idealpoint` variable that captures what international relations scholars have called countries' liberalism on issues such as political freedom, democratization, and financial liberalization. How has the distribution of state ideal points changed since the end of communism? Plot the distribution of ideal points separately for 1985, 2000, and 2014. Use the `hist()` function for plotting the distribution. In this exercise, pay attention to the x-axis and the y-axis in each plot. Add median to each plot as a vertical line. How do the two distributions differ? Describe the plots paying attention to the degree of polarization. Use the `quantile()` function to quantify the patterns you identified.

Q 1-1.

How has the distribution of state ideal points changed since the end of communism

Considering the histogram and quantile, it can be seen that the distribution of ideal points were polarized from 1985 to 2000. Also, we can see the anti-liberalization across countries in that the median has decreased from -0.2578553 to -0.3410604. There are particularly many countries with numbers between -1 and 0.5.

From 2000 to 2014, the polarization has eased in 2014. However, compared to 1985, the median has decreased from -0.2578553 to -0.4232631.

Median and Histogram of Ideal Points Separately for 1985, 2000, and 2014

```
un1985 <- subset(un, year == 1985)
un2000 <- subset(un, year == 2000)
un2014 <- subset(un, year == 2014)

median(un1985$idealpoint, na.rm = TRUE)
## [1] -0.2578553

median(un2000$idealpoint, na.rm = TRUE)
## [1] -0.3410604

median(un2014$idealpoint, na.rm = TRUE)
## [1] -0.4232631
```

Plot the histograms

```
par(mfrow=c(3,1))

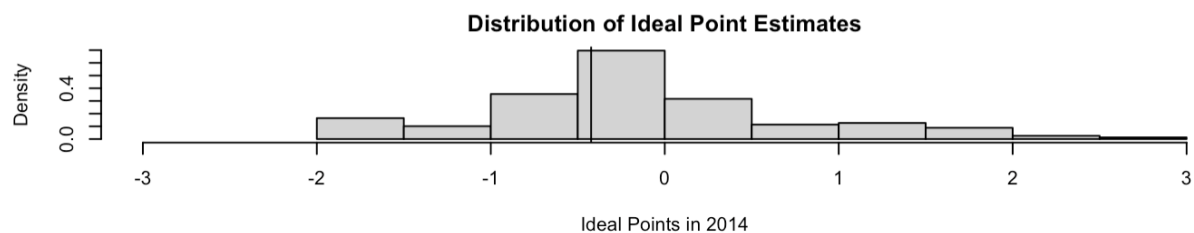
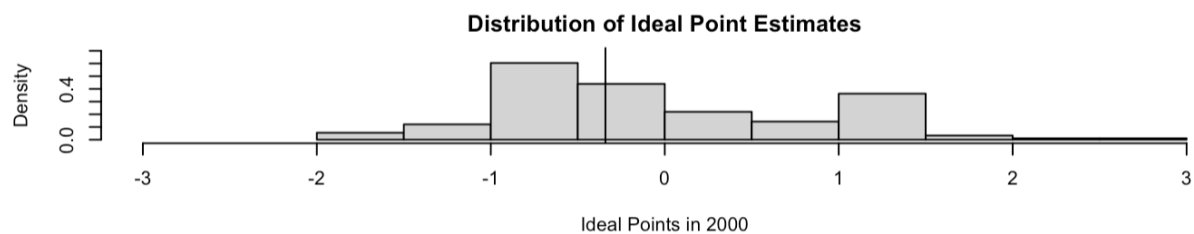
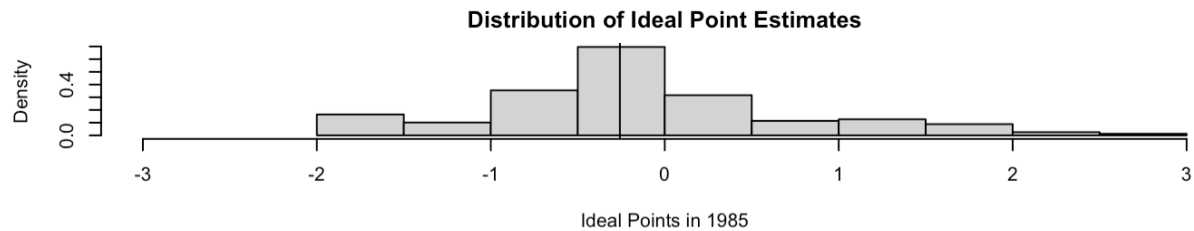
hist(un1985$idealpoint, freq = FALSE, main = "Distribution of Ideal Point Estimates",
     xlab = "Ideal Points in 1985", ylim = c(0,0.7), xlim = c(-3,3))
abline(v=median(un$idealpoint[un$year==1985]))

hist(un2000$idealpoint, freq = FALSE, main = "Distribution of Ideal Point Estimates",
     xlab = "Ideal Points in 2000", ylim = c(0,0.7), xlim = c(-3,3))
```

```
abline(v=median(un$Idealpoint[un$year==2000]))
```

```
hist(un1985$Idealpoint, freq = FALSE, main = "Distribution of Ideal Point Estimates",  
     xlab = "Ideal Points in 2014", ylim = c(0,0.7), xlim = c(-3,3))
```

```
abline(v=median(un$Idealpoint[un$year==2014]))
```



Quantile

```
quantile(un1985$Idealpoint, probs = seq(from = 0, to = 1, by = 0.25), na.rm = TRUE)
```

```
##      0%      25%      50%      75%     100%  
## -1.8290260 -0.5780610 -0.2578553  0.2370236  2.7688861
```

```
quantile(un2000$Idealpoint, probs = seq(from = 0, to = 1, by = 0.25), na.rm = TRUE)
```

```
##      0%      25%      50%      75%     100%  
## -1.6797190 -0.6910735 -0.3410604  0.7087527  2.7007489
```

```
quantile(un2014$Idealpoint, probs = seq(from = 0, to = 1, by = 0.25), na.rm = TRUE)
```

```
##      0%      25%      50%      75%     100%  
## -1.8487680 -0.7153760 -0.4232631  0.7876102  2.5304980
```

Question 2

Next, examine how the number of countries voting with the US has changed over time. Plot the average percent agreement with the US (pctagreeus) across all countries over time. In other words, you need to present the plot with year in your x-axis, and the average value of pctagreeus in your y-axis. Also, add the average percent agreement with Russia and China, respectively. Using the `tapply()` function may be helpful. What do you think of the pattern you see from the plot? Does the US appear to be becoming more or less isolated over time, as compared to Russia?

subset and calculate the average

```
us_year_mean <- tapply(un$pctagreeus, un$year, mean, na.rm = TRUE)
su_year_mean <- tapply(un$pctagreerussia, un$year, mean, na.rm = TRUE)
ch_year_mean <- tapply(un$pctagreechina, un$year, mean, na.rm = TRUE)

min(ch_year_mean, na.rm = TRUE)
## [1] 0.6225979

max(ch_year_mean, na.rm = TRUE)
## [1] 0.843394
```

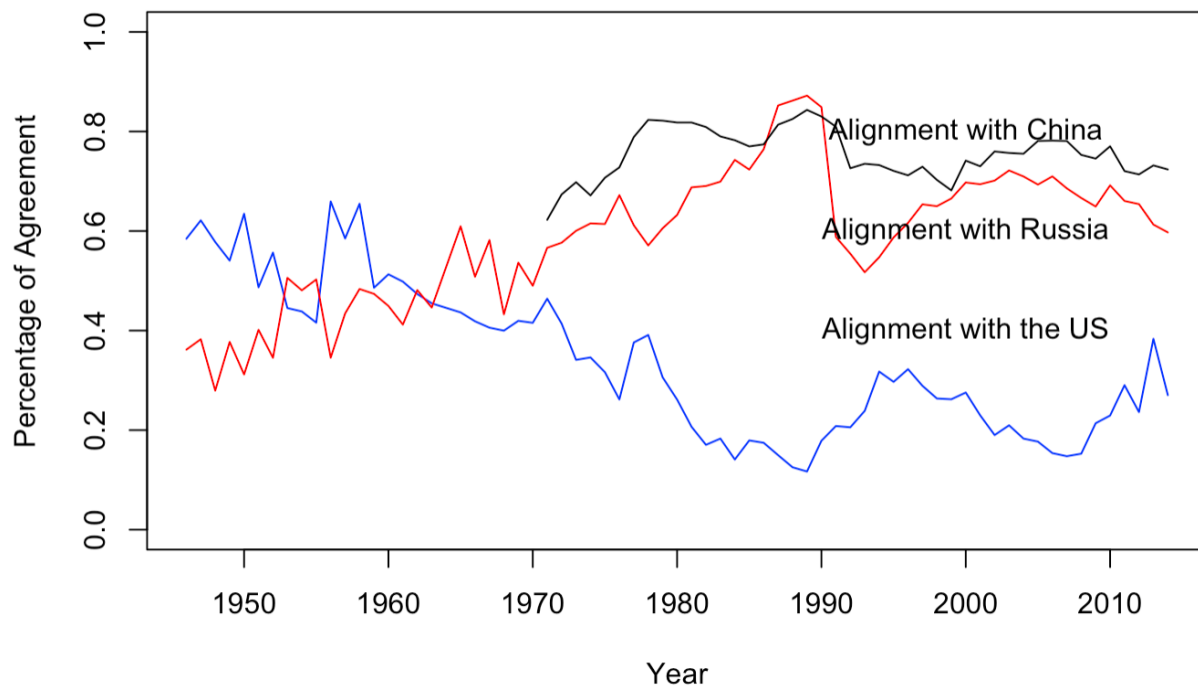
Q 2-1.

Plot the average percent agreement with the US (pctagreeus) across all countries over time. In other words, you need to present the plot with year in your x-axis, and the average value of pctagreeus in your y-axis. Also, add the average percent agreement with Russia and China, respectively.

plot the graph

```
par(mfrow=c(1,1))
plot(names(us_year_mean), us_year_mean, col = "blue", type = "l",
     xlab = "Year",
     ylab = "Percentage of Agreement", ylim=c(0,1))
lines(names(su_year_mean), su_year_mean, col = "red")
lines(names(ch_year_mean), ch_year_mean, col = "black")

text(2000, 0.4, "Alignment with the US")
text(2000, 0.6, "Alignment with Russia")
text(2000, 0.8, "Alignment with China")
```

Q 2-2.

What do you think of the pattern you see from the plot? Does the US appear to be becoming more or less isolated over time, as compared to Russia?

Considering the average of pro-US rates has decreased, the US has become more isolated from 1970s to 1990s compared to Russia. Although the average rates rose slightly in the 2010s, it is still consistently lower than Russia(or USSR) since the early 1960s. Therefore, we can say the US appear to becoming more isolated over time compared to Russia.

In 1990s, however, the rates of Russia also decreased. Although the overall average of the US has been low, both of the US and Russia can be seen as relatively isolated compared to the past, given that Russia is also lower than in the late 1980s. In other words, it can be interpreted that there can be an effect of the rise of China or the multi-polarization of the international system.

Question 3

Let's examine the voting patterns of African countries (Eastern Africa, Northern Africa, Southern Africa, Western Africa). What are some countries that are consistently pro-US? What are the most pro-China countries? Plot the average percent agreement with the US (pctagreeus) and with China (pctagreechina) across African countries over time. What explains the voting patterns of African countries? You can provide any perspectives on this.

subset the data

```
unAfrica <- subset(un, region == "Eastern Africa" | region == "Northern Africa" | region == "Southern Africa" | region == "Western Africa")
```

Q 3-1. What are some countries that are consistently pro-US

In order to analyze the 'consistency', the variance of pro-US rates in each country can be considered. Although all of these top 10 countries have low average rates, the countries with the lowest variance is 'South Africa'. Other countries cannot easily be considered as 'consistently pro-US' countries in that their average rates are low and their variances are also low. Rather, these countries can be considered as 'consistently anti-US' countries.

Therefore, we can say there are no consistently pro-US countries in Africa except 'South Africa'.

analyze the average pro-US rates of countries

```
us_country_mean_Africa <- tapply(unAfrica$pctagreeus, unAfrica$countryname, mean) %>%
  sort(decreasing = TRUE)

top_proUS_Africa <- us_country_mean_Africa[1:10] %>%
  as.data.frame.table()

us_country_var_Africa <- tapply(unAfrica$pctagreeus, unAfrica$countryname, var) %>%
  as.data.frame.table()

top_proUS_con_Africa <- inner_join(top_proUS_Africa, us_country_var_Africa, by = "Var1", na.rm = TRUE)

names(top_proUS_con_Africa) <- c("Countires", "pro US rates", "variance")

top_proUS_con_Africa
```

##	Countires	pro US rates	variance
## 1	South Africa	0.4513085	0.076188930
## 2	Liberia	0.3554425	0.044628974
## 3	South Sudan	0.2973300	0.005939531
## 4	Malawi	0.2792384	0.027736117
## 5	Ivory Coast	0.2509752	0.018309336
## 6	Ethiopia	0.2481347	0.024079210
## 7	Rwanda	0.2461881	0.021899050
## 8	Madagascar	0.2438762	0.021729757

```
## 9    Lesotho  0.2351554 0.026800438
## 10   Gambia  0.2333332 0.039832082
```

Q 3-2. What are the most pro-China countries

Zimbabwe is the most pro-China country.

Other countires: Djibouti, Comoros, Guinea-Bissau, Namibia, Seychelles, Mozambique, Cape Verde, Eritrea, South Sudan

Of the top 10 countries, 9 countries - except South Sudan - averaged about 90 percent.

analyze the average pro-China rates of countries

```
ch_country_mean_Africa <- tapply(unAfrica$pctagreechina, unAfrica$countryname, mean) %>%
  sort(decreasing = TRUE)

top_proChina_Africa <- ch_country_mean_Africa[1:10] %>%
  as.data.frame.table()

ch_country_var_Africa <- tapply(unAfrica$pctagreechina, unAfrica$countryname, var) %>%
  as.data.frame.table()

top_proChina_con_Africa <- inner_join(top_proChina_Africa, us_country_var_Africa, by = "Var1", na.rm = TRUE)

names(top_proChina_con_Africa) <- c("Countires", "pro China rates", "variance")

top_proChina_con_Africa
```

##	Countires	pro China rates	variance
## 1	Zimbabwe	0.9014954	0.002628914
## 2	Djibouti	0.8996057	0.005491896
## 3	Comoros	0.8958039	0.007871259
## 4	Guinea-Bissau	0.8813090	0.005939954
## 5	Namibia	0.8754067	0.004540565
## 6	Seychelles	0.8745607	0.011752176
## 7	Mozambique	0.8700153	0.006857537
## 8	Cape Verde	0.8645217	0.007143040
## 9	Eritrea	0.8362479	0.006347678
## 10	South Sudan	0.5329800	0.005939531

Q 3-3. Plot the average percent agreement with the US (pctagreeus) and with China (pctagreechina) across countries over time.

Q 3-4. What explains the voting patterns of African countries? You can provide any perspectives on this.

The average rates of African countries also showed a similar tendency to the one considering all countries around the world. The average of African countries has been consistently high since China joined the UN.

It can be explained that due to the rise of China and the multi-polarization of the international system, the average rates of the US has become relatively lower than the one of China.

Or, it could be because the aid of China for African countries has increased since it began 'reform and opening up' policy or aggressive international policies. In fact, some African countries are considered as China-dependent countries - such as Zimbabwe*.

** Samuel Ramani. (JAN 11, 2016). Zimbabwe: China's 'All-Weather' Friend in Africa While many worry about China's economy, Zimbabwe adopts the yuan as its international currency.". THE DIPLOMAT. <https://thediplomat.com/2016/01/zimbabwe-chinas-all-weather-friend-in-africa/>*

calculate the average pro-US rates of countries (year)

```
us_year_mean_Africa <- tapply(unAfrica$pctagreeus, unAfrica$year, mean, na.rm = TRUE)
min(us_year_mean_Africa, na.rm = TRUE)
```

```
## [1] 0.05965129
```

```
max(us_year_mean_Africa, na.rm = TRUE)
```

```
## [1] 0.6326253
```

calculate the average pro-China rates of countries (year)

```
ch_year_mean_Africa <- tapply(unAfrica$pctagreechina, unAfrica$year, mean, na.rm = TRUE)
min(ch_year_mean_Africa, na.rm = TRUE)
```

```
## [1] 0.6737618
```

```
max(ch_year_mean_Africa, na.rm = TRUE)
```

```
## [1] 0.9206248
```

plot the graph

```
par(mfrow=c(1,1))
```

```
plot(names(us_year_mean_Africa), us_year_mean_Africa, col = "blue", type = "l",  
     xlab = "Year",  
     ylab = "Percentage of Agreement among African Countries", ylim = c(0,1))  
lines(names(ch_year_mean_Africa), ch_year_mean_Africa, col = "red")
```

```
text(2000, 0.4, "Alignment with the US")
```

```
text(2000, 0.7, "Alignment with China")
```

