Assignment Week3

Adam Peetz

MSDS660 Week 3 Assignment

Regis University

Dr. Siripun Sanguansintukul

November 6th 2022

Pew Survey on Foreign Students

The United States has a world class higher education system. Many international students travel to the US
for school and return to their native countries, bringing the skills they have learned with them. This transfer
of knowledge is concerning for some who fear the United States is training students who will help competing
countries in a technological arms race. The Pew Research Center conducted a survey in 2021 that included
questions about this topic. Answers to those questions will be analyzed in the code below to estimate how
American's feel about accepting international students.

Set working directory, load packages, and dataset

```
#set working directory
setwd("C:\\Users\\adamg\\Documents\\MSDS_660\\Week_3")
```

```
#load required libraries
library(tidyverse)
library(data.table)
#install.packages("foreign")
library(foreign)
#install.packages("survey")
library(survey)
#install.packages("knitr")
library(knitr)
#install.packages("viridis")
library("viridis")
```

```
#load dataset
atp<-read.spss("ATP W82.sav", to.data.frame = TRUE)
names(atp) <- tolower(names(atp))
```

Responses to question 46

```
#show answers to question 46
print("Do you think it's good or bad for U.S. colleges and universities to accept international students

## [1] "Do you think it's good or bad for U.S. colleges and universities to accept international student
```

```
table(atp$gap21q46_w82)
```

```
##
##    Good     Bad Refused
##    2106     457      33
```

Responses to question 47

```
#show answers to question 47
print("When it comes to whether or not to limit Chinese students studying in the U.S., do you...")
```

```
## [1] "When it comes to whether or not to limit Chinese students studying in the U.S., do you..."
```

```
table(atp$gap21q47_w82)
```

```
##
## Strongly support limiting Chinese students
##                                        562
## Somewhat support limiting Chinese students
##                                        902
##  Somewhat oppose limiting Chinese students
##                                        658
##  Strongly oppose limiting Chinese students
##                                        434
##                                    Refused
##                                         40
```

Combining Questions 46 and 47

Many respondents who said it was good to accept international students also answered that they wanted to limit Chinese enrollment. These answers are somewhat contradictory. To add clarity to poll responses, answers to question 46 and 47 will be combined into a new variable that reflects objection to the enrollment of Chinese students specifically.

```
#create new column based on contents of other columns
mutate_df<-atp %>%
  mutate(chi_stu = case_when(gap21q46_w82 == "Good" & gap21q47_w82 == "Strongly oppose limiting Chinese
                             gap21q46_w82 == "Good" & gap21q47_w82 == "Somewhat oppose limiting Chi
                             gap21q46_w82 == "Good" & gap21q47_w82 == "Strongly support limiting Ch
                             gap21q46_w82 == "Good" & gap21q47_w82 == "Somewhat support limiting Chi
                             gap21q46_w82 == "Good" & gap21q47_w82 == "Refused" ~ "Good",
                             gap21q46_w82 == "Bad" ~ "Bad",
                             gap21q46_w82 == "Refused" & gap21q47_w82 == "Strongly oppose limiting C
                             gap21q46_w82 == "Refused" & gap21q47_w82 == "Strongly support limiting
                             gap21q46_w82 == "Refused" & gap21q47_w82 == "Somewhat oppose limiting C
                             gap21q46_w82 == "Refused" & gap21q47_w82 == "Somewhat support limiting
                             gap21q46_w82 == "Refused" & gap21q47_w82 == "Refused" ~ "Refused",
                             ) %>%
         fct_relevel("Good",
                     "Good except for Chinese students",
                     "Bad",
                     "Refused")
      )
```

```r
#create grouped analysis dataframe for sum/total
analysis_df_1 <- mutate_df %>%
  group_by(chi_stu) %>%
  summarize(total = n_distinct(interview_start_w82))

#draw horizontal barplot
p1 <-ggplot(analysis_df_1, aes(x=fct_reorder(chi_stu,total), y=total,fill=total))+
        geom_bar(stat='identity')+
        coord_flip() +
        scale_x_discrete(position = "top")+
        scale_fill_viridis(discrete = FALSE, option="B") +
        ggtitle("Are international students good or bad?") +
        theme_bw() +
        theme(plot.title = element_text(size=12),
            panel.border = element_blank(),
            axis.line.y = element_blank(),
            axis.title.y=element_blank(),
            axis.title.x=element_blank(),
            panel.grid.major = element_blank(),
            panel.grid.minor = element_blank(),
            axis.line = element_line(colour = "black"),
            legend.position="none")

#show plot
p1
```

Are international students good or bad?
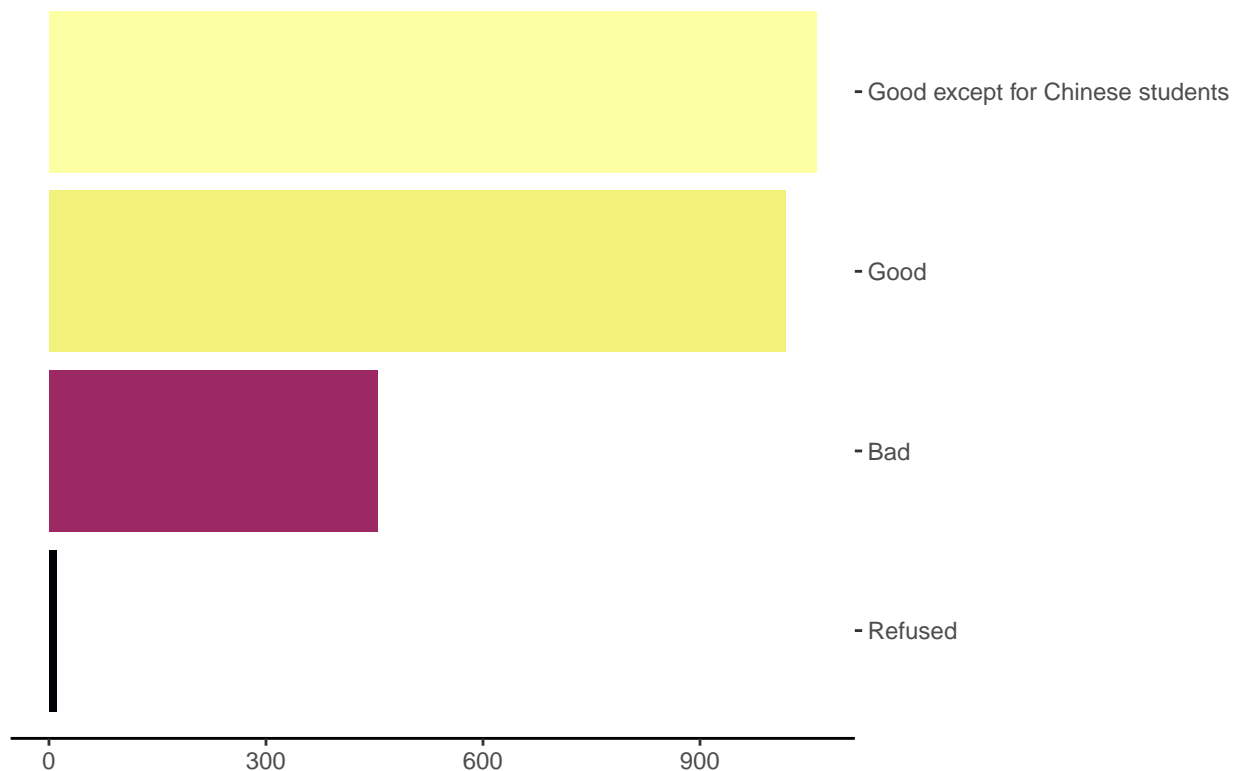
Investigating answers by demographics: Race, Education, and Political Party

There is a close to even split between responses for "Good" and "Good except for Chinese students" with prevailing opinion being a limit to Chinese enrollment in colleges and universities. The least given response was "Bad". These opinions on enrollment may show distinct trends when divided by various demographic groups participating in the survey. Responses to this question will be broken out by Education, Political Party, and Race to show how opinions differ between these groups.

```r
#create new DF with used features only
subset_df <- mutate_df %>%
  select(id = qkey,
         weight_w82,
         chi_stu,
         f_racecmb,
         f_educcat,
         f_party_final)

#change feature names
setnames(subset_df, old = c('f_racecmb', 'f_educcat', 'f_party_final'), new = c('Race', 'Education', 'Po

#convert subset to long DF
subset_df_long <- subset_df %>%
  gather(key = subgroup_variable, value = subgroup,
         Race, Education, Political_Party)

#create weighted estimates of long subset DF
subset_df_long_estimates <- subset_df_long %>%
  group_by(subgroup_variable, subgroup, chi_stu) %>%
  summarise(weighted_n = sum(weight_w82)) %>%
  group_by(subgroup) %>%
  mutate(weighted_group_size = sum(weighted_n),
         weighted_estimate = weighted_n/weighted_group_size)

#drop unused features
subset_df_long_estimates <- subset_df_long_estimates %>%
  select(-weighted_n, -weighted_group_size)

#orient data for charting.
subset_df_long_estimates %>%
  filter(chi_stu != "Refused") %>%
  filter(!(subgroup %in%
             c("Refused"))) %>%
  ggplot(aes(x = weighted_estimate,y = subgroup)) +
  geom_point() +
  labs(x = "Percent of Respondents")+
  scale_x_continuous(limits = c(0, .8),
                     breaks = seq(0, .6, by = .2),
                     labels = scales::percent(seq(0, .6, by = .2),
                     accuracy = 1)) +
  facet_grid(cols = vars(chi_stu),
             rows = vars(subgroup_variable),
             scales = "free_y",
             space = "free") +
```
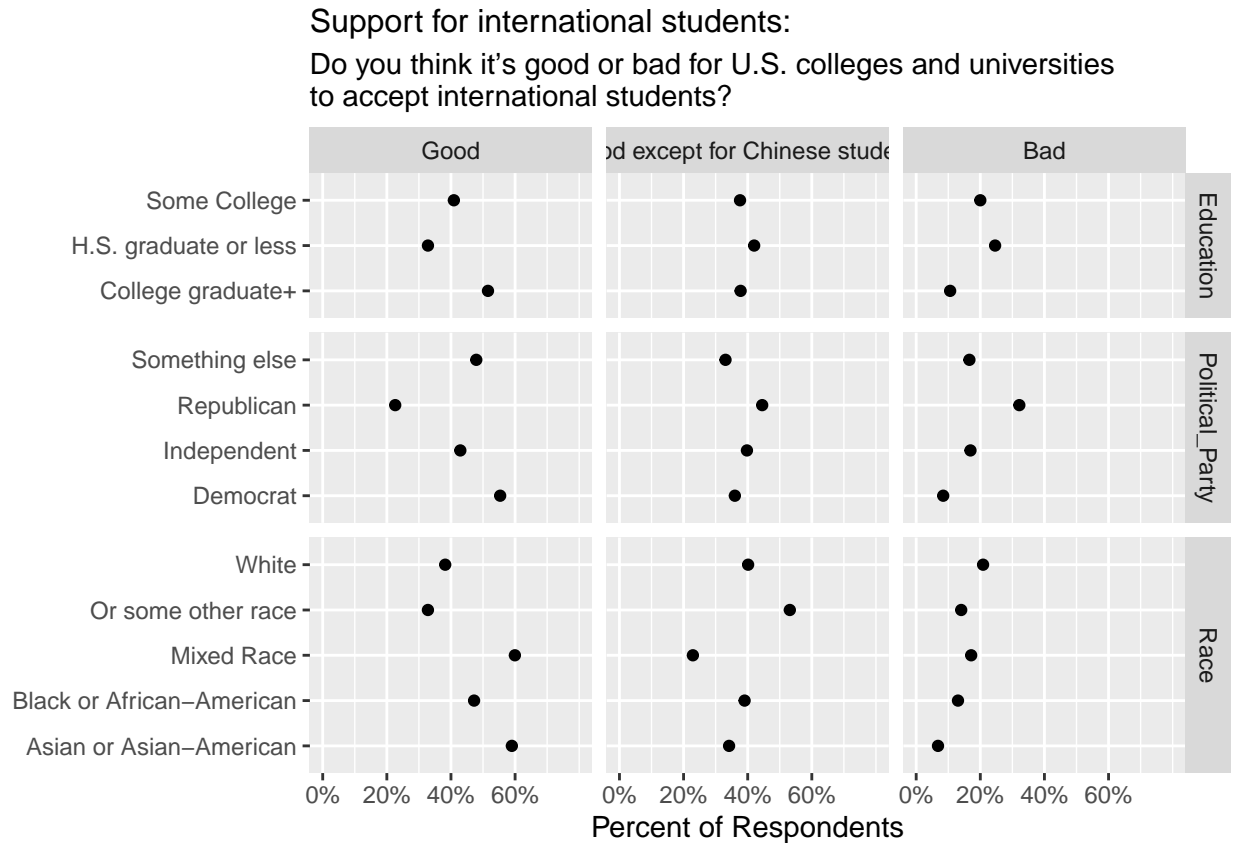
```
    theme(plot.title = element_text(size=12)) +
    theme(axis.title.y = element_blank())+
    labs(title = 'Support for international students:',
         subtitle = "Do you think it's good or bad for U.S. colleges and universities\nto accept internat:
```


Support for international students:

Do you think it's good or bad for U.S. colleges and universities
to accept international students?

Demographic Analysis

Education

Opinions on enrollment show support for international students increasing as education level go up.

Party

There is a split of opinion based on party affiliation. Republicans are less likely to support international student enrollment than democrats. All parties show the same resistance to Chinese student enrollment specifically.

Race

Race divides opinions. "Mixed race", "Black or African American", and "Asian or Asian-American", show more support for intentional students than other races participating in the survey. "Or some other race" shows the strongest resistance to the enrollment of Chinese students. "White" has the largest percent of respondents stating all international students are bad.

Conclusion

Opinions on the enrollment of international students vary based on which demographic group is asked. Each group's opinion can be evaluated using weighted responses to the Pew Survey. This allows researchers to evaluate support and resistance for each group participating in the survey. This type of information can be valuable to groups such as political parties who want to make decisions that align with their base.

References

MSDS660. (2022). Statistical Methods and Experimental Design. Taught by Dr. Siripun Sanguansintukul.

Pew Research Center. (2021). Global Attitudes Project U.S. survey. retrieved from https://www.pewresearch.org/american-trends-panel-datasets/