

Group Exploration of the GSS Dataset

Rumi Albert, Lehan Li, Alice Tang

April 28, 2020

Contents

Introduction	1
Getting started	1
Initial Observations	3
Questions and Findings	4
Question #1: Change in political view over the years	4
Question #2: : A people more accepting of liberal ideas?	4
Question #3: Is there a bias in political view?	11
Question #4: Linear model predicting polviews	14
Conclusion	15

Introduction

We will be looking into the General Social Survey Cumulative Data also known as GSS. This data set gathers information from contemporary American Society in order to monitor and explain trends and changes in attitudes, behaviors, and attributes. This Data set has been tracking Americans since 1972 until 2018. The topics covered include civil liberties, crime & violence, intergroup tolerance, morality, national spending priorities, psychological well-being, social mobility, and stress and traumatic events. We will be exploring the whole dataset to find interesting patterns and obtain knowledge about the American population. First load all required packages and the data set.

Getting started

These are the packages utilized in this report.

Load packages

```
library(plotly)
library(tidyverse)
library(tidymodels)
library(dsbox)
library(patchwork)
library(drat)
library(gssr)
```

Load data and explore

Here we are setting up the actual data table called `gss_all`, we are also setting up an informative table about its variables called `gss_doc`.

```
data(gss_doc)
data(gss_all)
data_info <- gss_doc
data <- gss_all
data_info
```

```
## # A tibble: 6,144 x 5
##   id      description      properties      marginals      text
##   <chr>   <chr>           <list>         <list>         <chr>
## 1 caseid YEAR + Responde~ <tibble[,3] ~ <tibble[,3] ~ None
## 2 year   GSS year for th~ <tibble[,3] ~ <tibble[,5] ~ None
## 3 id     Respondent ID n~ <tibble[,3] ~ <tibble[,3] ~ None
## 4 age    Age of responde~ <tibble[,3] ~ <tibble[,3] ~ 13. Respondent's age
## 5 sex    Respondents sex  <tibble[,3] ~ <tibble[,5] ~ 23. Code respondent's sex
## 6 race   Race of respond~ <tibble[,3] ~ <tibble[,5] ~ 24. What race do you con~
## 7 racec~ What Is R's rac~ <tibble[,3] ~ <tibble[,5] ~ 1602. What is your race?~
## 8 racec~ What Is R's rac~ <tibble[,3] ~ <tibble[,5] ~ 1602. What is your race?~
## 9 racec~ What Is R's rac~ <tibble[,3] ~ <tibble[,5] ~ 1602. What is your race?~
## 10 hispa~ Hispanic specif~ <tibble[,3] ~ <tibble[,5] ~ 1601. IF R IS FEMALE, RE~
## # ... with 6,134 more rows
```

```
data
```

```
## # A tibble: 64,814 x 6,108
##   year   id wrkstat  hrs1  hrs2 evwork  occ prestige wrkslf wrkgovt commute
##   <dbl> <dbl>   <dbl> <dbl> <dbl> <dbl> <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
## 1 1972   1     1    NA    NA    NA   205     50     2     NA     NA
## 2 1972   2     5    NA    NA    1   441     45     2     NA     NA
## 3 1972   3     2    NA    NA    NA   270     44     2     NA     NA
## 4 1972   4     1    NA    NA    NA    1     57     2     NA     NA
## 5 1972   5     7    NA    NA    1   385     40     2     NA     NA
## 6 1972   6     1    NA    NA    NA   281     49     2     NA     NA
## 7 1972   7     1    NA    NA    NA   522     41     2     NA     NA
## 8 1972   8     1    NA    NA    NA   314     36     2     NA     NA
## 9 1972   9     2    NA    NA    NA   912     26     2     NA     NA
## 10 1972  10     1    NA    NA    NA   984     18     2     NA     NA
## # ... with 64,804 more rows, and 6,097 more variables: industry <dbl>,
## #   occ80 <dbl>, prestg80 <dbl>, indus80 <dbl>, indus07 <dbl>, occonet <dbl>,
## #   found <dbl>, occ10 <dbl>, occindv <dbl>, occstatus <dbl>, occtag <dbl>,
```

```
## #   prestg10 <dbl>, prestg105plus <dbl>, indus10 <dbl>, indstatus <dbl>,
## #   indtag <dbl>, marital <dbl>, martype <dbl>, agewed <dbl>, divorce <dbl>,
## #   widowed <dbl>, spwrksta <dbl>, sphrs1 <dbl>, sphrs2 <dbl>, spevwork <dbl>,
## #   cwrksta <dbl>, cwrkslf <dbl>, coevwork <dbl>, cohrrs1 <dbl>, cohrrs2 <dbl>,
## #   spocc <dbl>, sppres <dbl>, spwrkslf <dbl>, spind <dbl>, spocc80 <dbl>,
## #   sppres80 <dbl>, spind80 <dbl>, spocc10 <dbl>, spoccindv <dbl>,
## #   spoccstatus <dbl>, spocctag <dbl>, sppres10 <dbl>, sppres105plus <dbl>,
## #   spind10 <dbl>, spindstatus <dbl>, spindtag <dbl>, coocc10 <dbl>,
## #   coind10 <dbl>, paocc16 <dbl>, papres16 <dbl>, pawrkslf <dbl>,
## #   paid16 <dbl>, paocc80 <dbl>, papres80 <dbl>, paid80 <dbl>, paocc10 <dbl>,
## #   paoccindv <dbl>, paoccstatus <dbl>, paocctag <dbl>, papres10 <dbl>,
## #   papres105plus <dbl>, paid10 <dbl>, paidstatus <dbl>, paidtag <dbl>,
## #   maocc80 <dbl>, mapres80 <dbl>, mawrkslf <dbl>, maind80 <dbl>,
## #   maocc10 <dbl>, maoccindv <dbl>, maoccstatus <dbl>, maocctag <dbl>,
## #   mapres10 <dbl>, mapres105plus <dbl>, maind10 <dbl>, maindstatus <dbl>,
## #   maintag <dbl>, sibs <dbl>, childr <dbl>, age <dbl>, agekdrn <dbl>,
## #   educ <dbl>, paeduc <dbl>, maeduc <dbl>, speduc <dbl>, coeduc <dbl>,
## #   codeg <dbl>, degree <dbl>, padeg <dbl>, madeg <dbl>, spdeg <dbl>,
## #   major1 <dbl>, major2 <dbl>, dipged <dbl>, spdipged <dbl>, codipged <dbl>,
## #   cosector <dbl>, whenhs <dbl>, whencol <dbl>, sector <dbl>, ...
```

Initial Observations

The 6,108 variables:

These are the few variables the data table contains

- year: Gss year for this respondent
- id: Respondent id number
- wrkstat: Labor force status
- hrs1: Number of hours worked last week
- hrs2: Number of hours usually work a week
- evwork: Ever work as long as one year
- occ: census occupation code (1970)
- prestige: occupational prestige score (1970)
- wrkslf: self-emp or works for somebody
- wrkgovt: Govt or private employee
- commute: Travel time to work
- industry: industry code (1970)

The 2867 observations:

There are 64,814 observations in the GSS Dataset. This can be found in two ways to find this information, using tail or nrow function.

```
nrow(data)
```

```
## [1] 64814
```

Questions and Findings

Question #1: Change in political view over the years

Observing the visualization below, it is evident that the change in political view over the years has remained at a constant proportion. Looking at the proportional difference between political views in the year 1974 and 2018, one can see that the proportion remained constant, with a small rise in conservative views; though the count of each political view increased over the years, the proportion of the political views persisted staying roughly the same.

```
data_A1 <- filter(data, !is.na(polviews)) %>% select(year, polviews) %>% mutate(polviews = ifelse(polviews == "liberal", "liberal", "conservative"))
data_A1 <- tibble(summarise(group_by(data_A1, year, polviews), count = n()))
colors = c('red', 'blue', 'grey')
base <- data_A1 %>%
  plot_ly(x = ~polviews, y = ~count)

base %>%
  add_bars(marker = list(color = c('red', 'blue', 'silver')), type = 'bar', frame = ~year, ids = ~polviews)
  animation_opts(500, redraw = FALSE)
```

Question #2: : A people more accepting of liberal ideas?

Here we will analyze a variety of questions to see if the aforementioned is indeed the case. Utilize visualizations to observe any trends which suggest people are more accepting of liberal ideas. The following are the questions that the respondents were asked.

Is it wrong for same-sex adults to have sexual relations?

We visualize the percentage of respondents who chose the answer “Not wrong at all” and “Always wrong” over the years.

The following two visualizations reveal that people, generally speaking, are more accepting to same-sex adults having sexual relations. Viz, the collective of society, including conservatives and moderates, are accepting of homosexual behavior; something typically characterized of being of liberal nature.

```
dataA2 <- select(data, year, homosex) %>% filter(!is.na(homosex))
dataA2 <- mutate(dataA2, homosex = ifelse(homosex == 4, 1, 0))
table_Not_wrong_at_all <- tibble(summarise(group_by(dataA2, year), percentage = mean(homosex) ))

dataA2 <- select(data, year, homosex) %>% filter(!is.na(homosex))
dataA2 <- mutate(dataA2, homosex = ifelse(homosex == 1, 1, 0))
table_Always_wrong <- tibble(summarise(group_by(dataA2, year), percentage = mean(homosex) ))

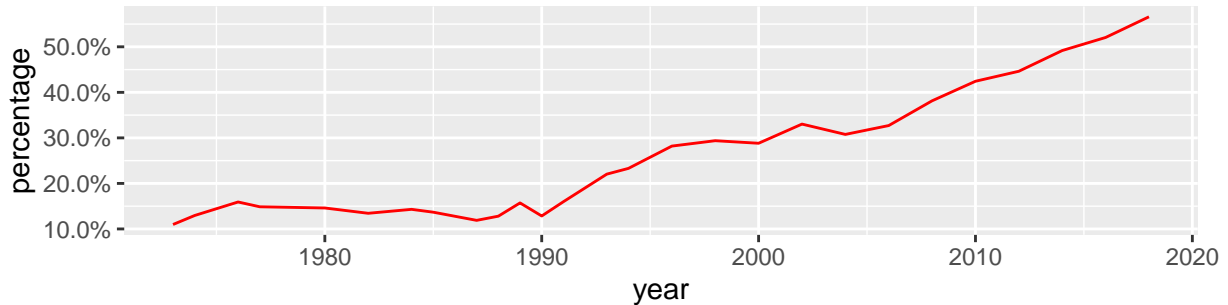
p_main <- ggplot(table_Not_wrong_at_all) +
  geom_line(mapping = aes(x = year, y = percentage), color = "red") +
  scale_y_continuous(labels = scales::percent) +
  labs(title = "Question: Is it wrong for same-sex adults to have sexual relations", subtitle = "Response: Not wrong at all")

p_int <- ggplot(table_Always_wrong) +
  geom_line(mapping = aes(x = year, y = percentage), color = "blue") +
  scale_y_continuous(labels = scales::percent) +
  labs(title = "Is it wrong for same-sex adults to have sexual relations", subtitle = "Response: Always wrong")
```

```
p_main /
p_int +
plot_layout(guides = "collect")
```

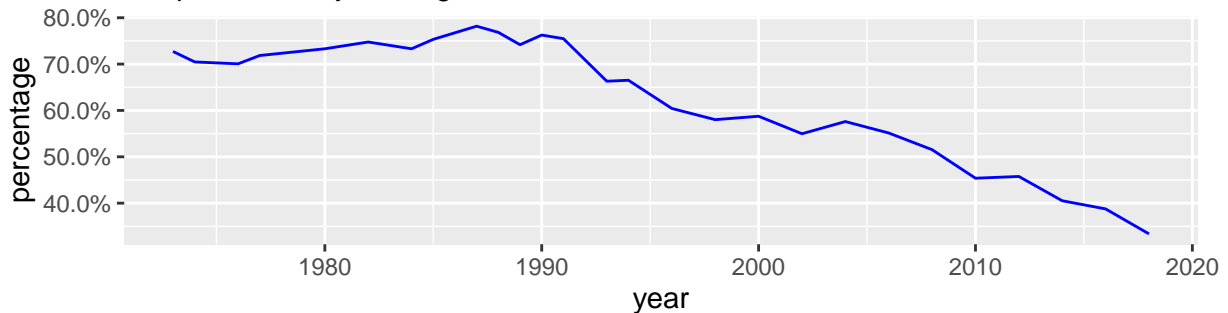
Question: Is it wrong for same-sex adults to have sexual relations

Response: Not wrong at all



Is it wrong for same-sex adults to have sexual relations

Response: Always wrong



What are your feelings about the bible?

The respondents were to answer either “Book of fables” or “Word of god”. We will see the percentage of respondents who responded with “Book of fables” and “Word of god” over the years.

The two distributions below illustrate that the number of people believing in the bible has been decreasing over the years. People are beginning to take the bible to be a “book of fables” and disregarding the “word of god”.

```
dataA4 <- select(data, year, bible) %>% filter(!is.na(bible))
dataA4 <- mutate(dataA4, bible = ifelse(bible == 1, 1, 0))
table_Word_of_god <- tibble(summarise(group_by(dataA4, year), percentage = mean(bible) ))

dataA4 <- select(data, year, bible) %>% filter(!is.na(bible))
dataA4 <- mutate(dataA4, bible = ifelse(bible == 3, 1, 0))
table_Book_of_fables <- tibble(summarise(group_by(dataA4, year), percentage = mean(bible) ))

p_main <- ggplot(table_Book_of_fables) +
  geom_line(mapping = aes(x = year, y = percentage), color = "red") +
  scale_y_continuous(labels = scales::percent) +
```

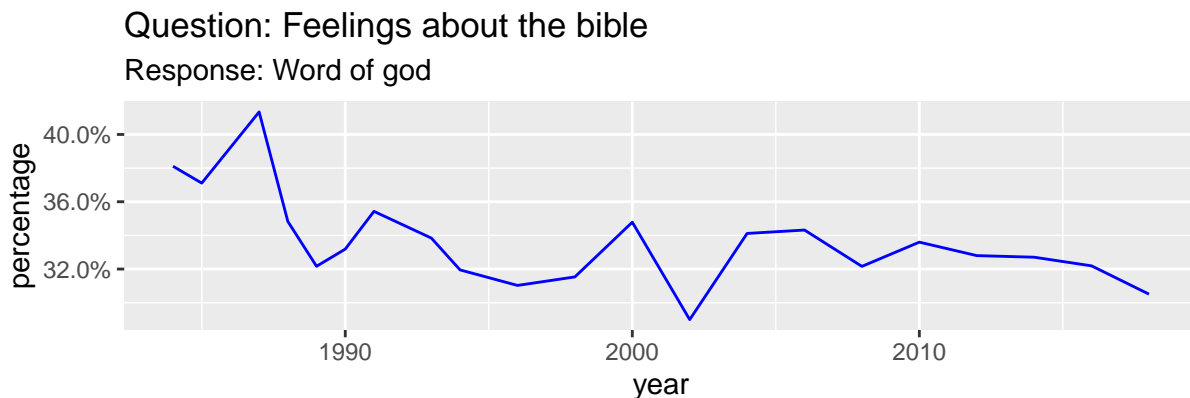
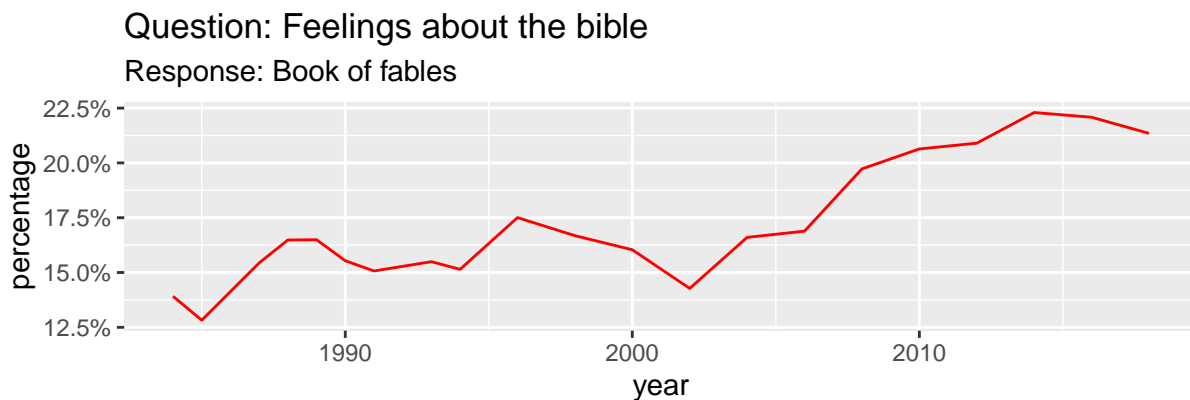
```

labs(title = "Question: Feelings about the bible", subtitle = "Response: Book of fables")

p_int <- ggplot(table_Word_of_god) +
  geom_line(mapping = aes(x = year, y = percentage), color = "blue") +
  scale_y_continuous(labels = scales::percent) +
  labs(title = "Question: Feelings about the bible", subtitle = "Response: Word of god")

p_main /
p_int +
plot_layout(guides = "collect")

```



Should marijuana be made legal?

We will analyze the percentage of respondents who responded with “Legal” and “Not Legal” over the years. The visualizations below inform us of an evident social opinion: that marijuana should be legalized. The percentage of people who believe marijuana should not be illegal, and should instead be legalized has increased dramatically since 1980.

```

dataA5 <- select(data, year, grass) %>% filter(!is.na(grass))
dataA5 <- mutate(dataA5, grass = ifelse(grass == 1, 1, 0))
table_Legal <- tibble(summarise(group_by(dataA5, year), percentage = mean(grass) ))

dataA5 <- select(data, year, grass) %>% filter(!is.na(grass))
dataA5 <- mutate(dataA5, grass = ifelse(grass == 2, 1, 0))

```

```

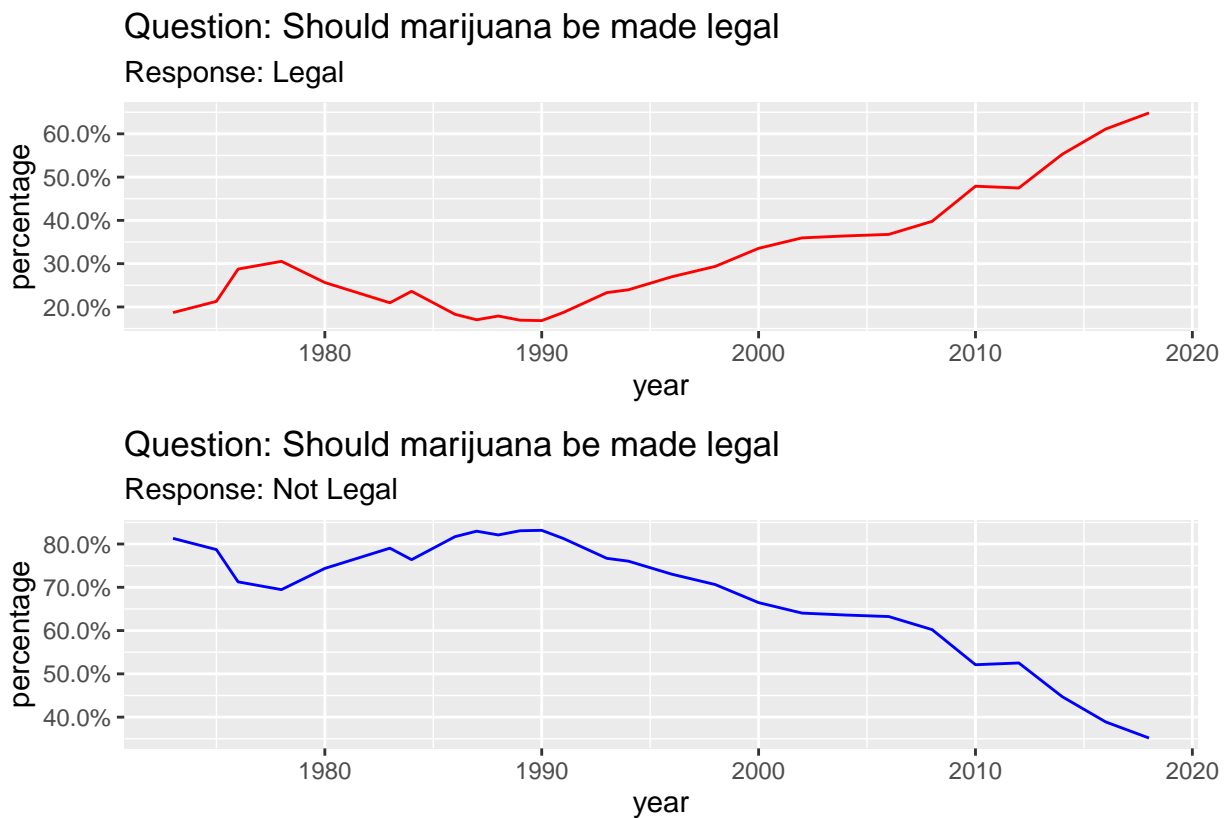
table_Not_Legal <- tibble(summarise(group_by(dataA5, year), percentage = mean(grass) ))

p_main <- ggplot(table_Legal) +
  geom_line(mapping = aes(x = year, y = percentage), color = "red") +
  scale_y_continuous(labels = scales::percent) +
  labs(title = "Question: Should marijuana be made legal", subtitle = "Response: Legal")

p_int <- ggplot(table_Not_Legal) +
  geom_line(mapping = aes(x = year, y = percentage), color = "blue") +
  scale_y_continuous(labels = scales::percent) +
  labs(title = "Question: Should marijuana be made legal", subtitle = "Response: Not Legal")

p_main /
p_int +
plot_layout(guides = "collect")

```



Should a pregnant woman obtain a legal abortion if the woman wants it for any reason?

We will analyze the percentage of respondents who responded with “should” and “should not” to this question over the years. The following figures inform us that as a whole, the percentage of respondents who believe that abortion should be possible for a pregnant woman regardless of the reason has increased remarkably.

```

dataA6 <- select(data, year, abany) %>% filter(!is.na(abany))
dataA6 <- mutate(dataA6, abany = ifelse(abany == 1, 1, 0))

```

```

table_should <- tibble(summarise(group_by(dataA6, year), percentage = mean(abany) ))

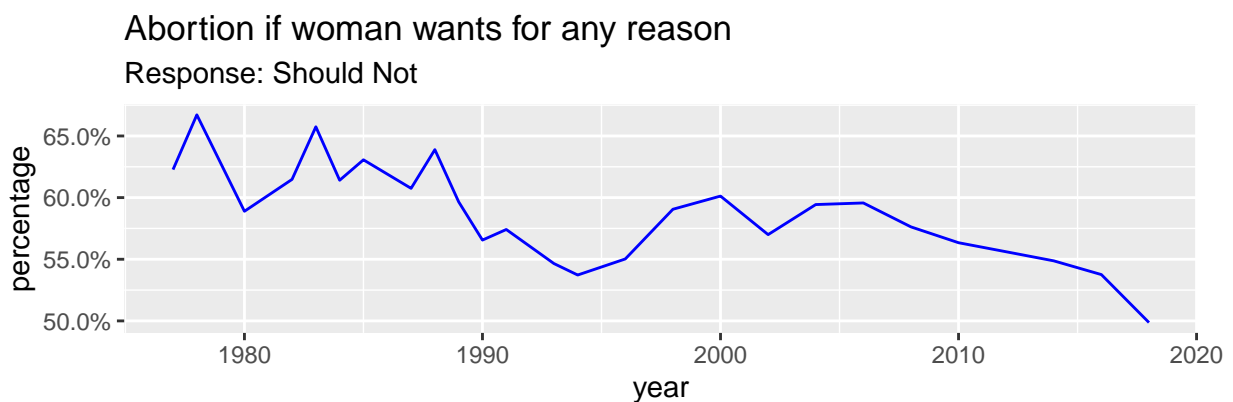
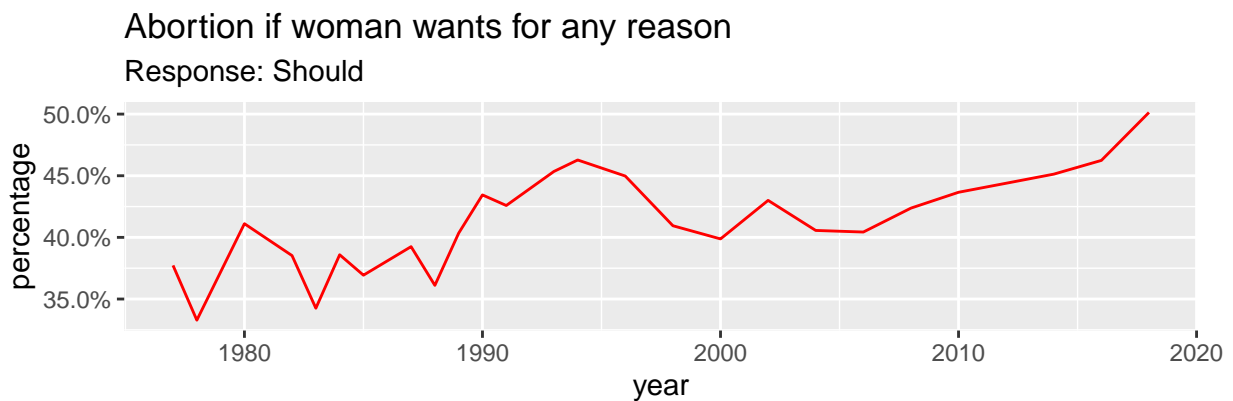
dataA6 <- select(data, year, abany) %>% filter(!is.na(abany))
dataA6 <- mutate(dataA6, abany = ifelse(abany == 2, 1, 0))
table_should_not <- tibble(summarise(group_by(dataA6, year), percentage = mean(abany) ))

p_main <- ggplot(table_should) +
  geom_line(mapping = aes(x = year, y = percentage), color = "red") +
  scale_y_continuous(labels = scales::percent) +
  labs(title = "Abortion if woman wants for any reason ", subtitle = "Response: Should")

p_int <- ggplot(table_should_not) +
  geom_line(mapping = aes(x = year, y = percentage), color = "blue") +
  scale_y_continuous(labels = scales::percent) +
  labs(title = "Abortion if woman wants for any reason ", subtitle = "Response: Should Not")

p_main /
p_int +
plot_layout(guides = "collect")

```



Should a college professor who claims to be communist be fired?

The following figures show that the respondents over the years believe that a communist teacher should not be fired if he or she is a college professor.


```

dataA7 <- select(data, year, colcom) %>% filter(!is.na(colcom))
dataA7 <- mutate(dataA7, colcom = ifelse(colcom == 4, 1, 0))
table_fired <- tibble(summarise(group_by(dataA7, year), percentage = mean(colcom) ))

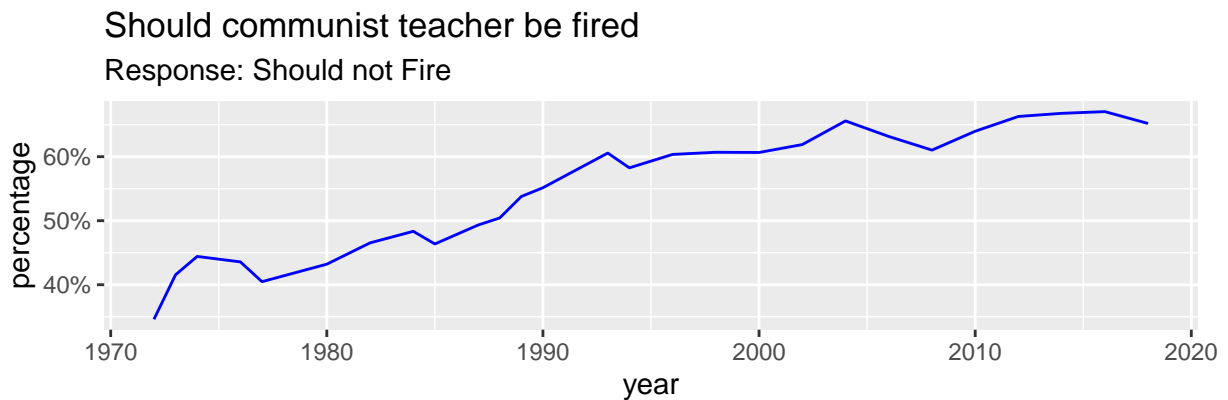
dataA7 <- select(data, year, colcom) %>% filter(!is.na(colcom))
dataA7 <- mutate(dataA7, colcom = ifelse(colcom == 5, 1, 0))
table_not_fired <- tibble(summarise(group_by(dataA7, year), percentage = mean(colcom) ))

p_main <- ggplot(table_fired) +
  geom_line(mapping = aes(x = year, y = percentage), color = "red") +
  scale_y_continuous(labels = scales::percent) +
  labs(title = "Should communist teacher be fired", subtitle = "Response: Should Fire")

p_int <- ggplot(table_not_fired) +
  geom_line(mapping = aes(x = year, y = percentage), color = "blue") +
  scale_y_continuous(labels = scales::percent) +
  labs(title = "Should communist teacher be fired ", subtitle = "Response: Should not Fire")

p_main /
p_int +
plot_layout(guides = "collect")

```



Is it wrong to have sex before marriage?

The figures below show that the percentage of people that believe that sex before marriage is acceptable has been increasing steadily throughout the years.

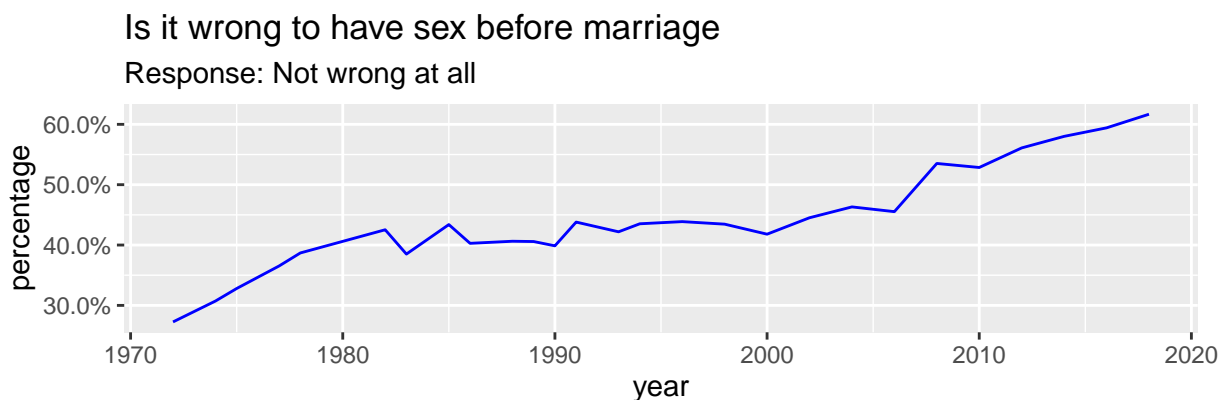
```
dataA8 <- select(data, year, premarsx) %>% filter(!is.na(premarsx))
dataA8 <- mutate(dataA8, premarsx = ifelse(premarsx == 1, 1, 0))
table_Always_wrong <- tibble(summarise(group_by(dataA8, year), percentage = mean(premarsx) ))

dataA8 <- select(data, year, premarsx) %>% filter(!is.na(premarsx))
dataA8 <- mutate(dataA8, premarsx = ifelse(premarsx == 4, 1, 0))
table_Not_wrong_at_all <- tibble(summarise(group_by(dataA8, year), percentage = mean(premarsx) ))

p_main <- ggplot(table_Always_wrong) +
  geom_line(mapping = aes(x = year, y = percentage), color = "red") +
  scale_y_continuous(labels = scales::percent) +
  labs(title = "Is it wrong to have sex before marriage", subtitle = "Response: Always wrong")

p_int <- ggplot(table_Not_wrong_at_all) +
  geom_line(mapping = aes(x = year, y = percentage), color = "blue") +
  scale_y_continuous(labels = scales::percent) +
  labs(title = "Is it wrong to have sex before marriage", subtitle = "Response: Not wrong at all")

p_main /
p_int +
plot_layout(guides = "collect")
```



Question #3: Is there a bias in political view?

In order to avoid bias, we should delve further into our findings to see whether it is only liberals who are answering the questions, or if moderates and conservatives also agree with these liberal ideas.

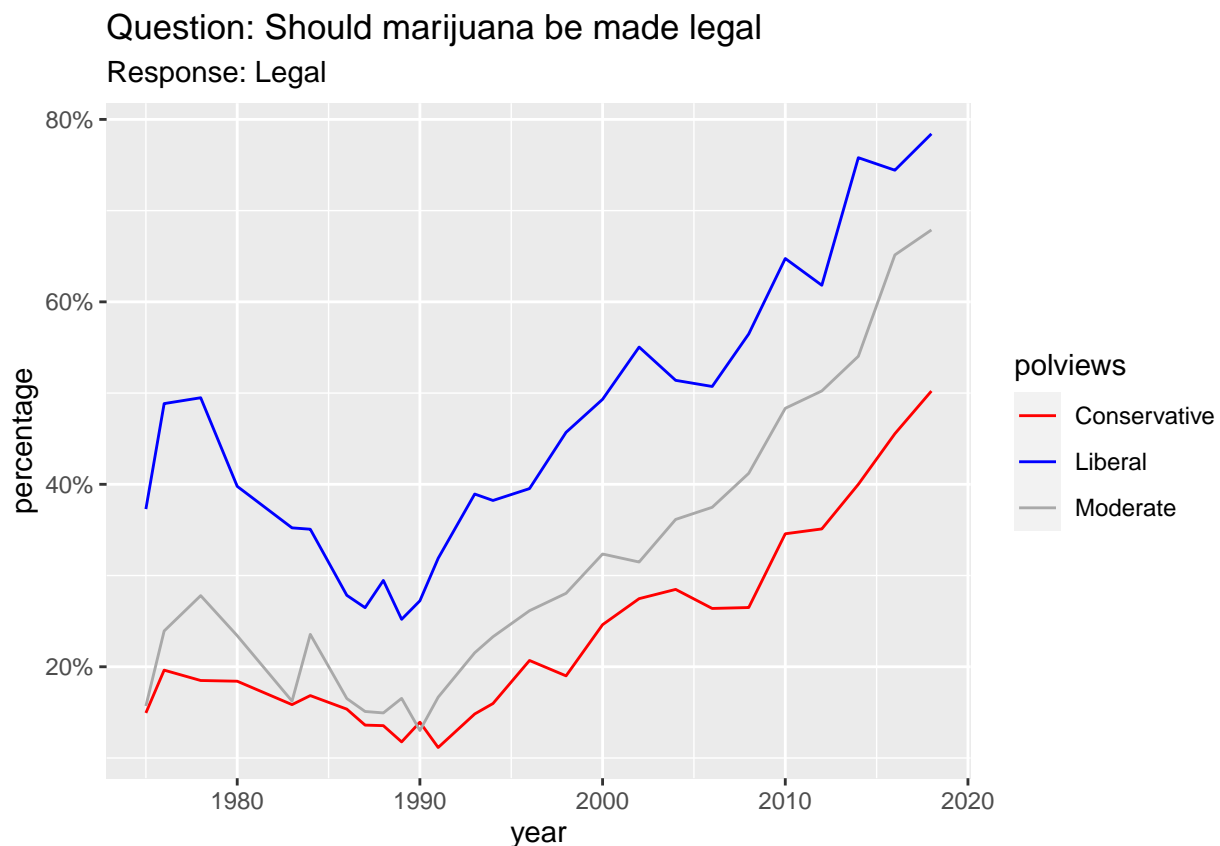
Should marijuana be made legal?

We will observe the percentage of respondents of each party over the years. The following figure illustrates that all three political views: conservative, liberal, and moderate all support the legalization of marijuana. The percentage of this agreement has grown remarkably. In regard to marijuana legalization, the percentage of liberal acceptance of such legalization is at the same level as conservative acceptance in 2007.

```
dataA5 <- select(data, year, grass, polviews) %>% filter(!is.na(grass), !is.na(polviews), polviews == 1)
dataA5 <- mutate(dataA5, grass = ifelse(grass == 1, 1, 0)) %>% mutate(polviews = ifelse(polviews == 1, 1, 0))

table_Legal <- tibble(summarise(group_by(dataA5, year, polviews), percentage = mean(grass)))

ggplot(table_Legal) +
  geom_line(mapping = aes(x = year, y = percentage, color = polviews)) +
  scale_y_continuous(labels = scales::percent) +
  labs(title = "Question: Should marijuana be made legal", subtitle = "Response: Legal") +
  scale_color_manual(values=c("red", "blue", "dark grey"))
```



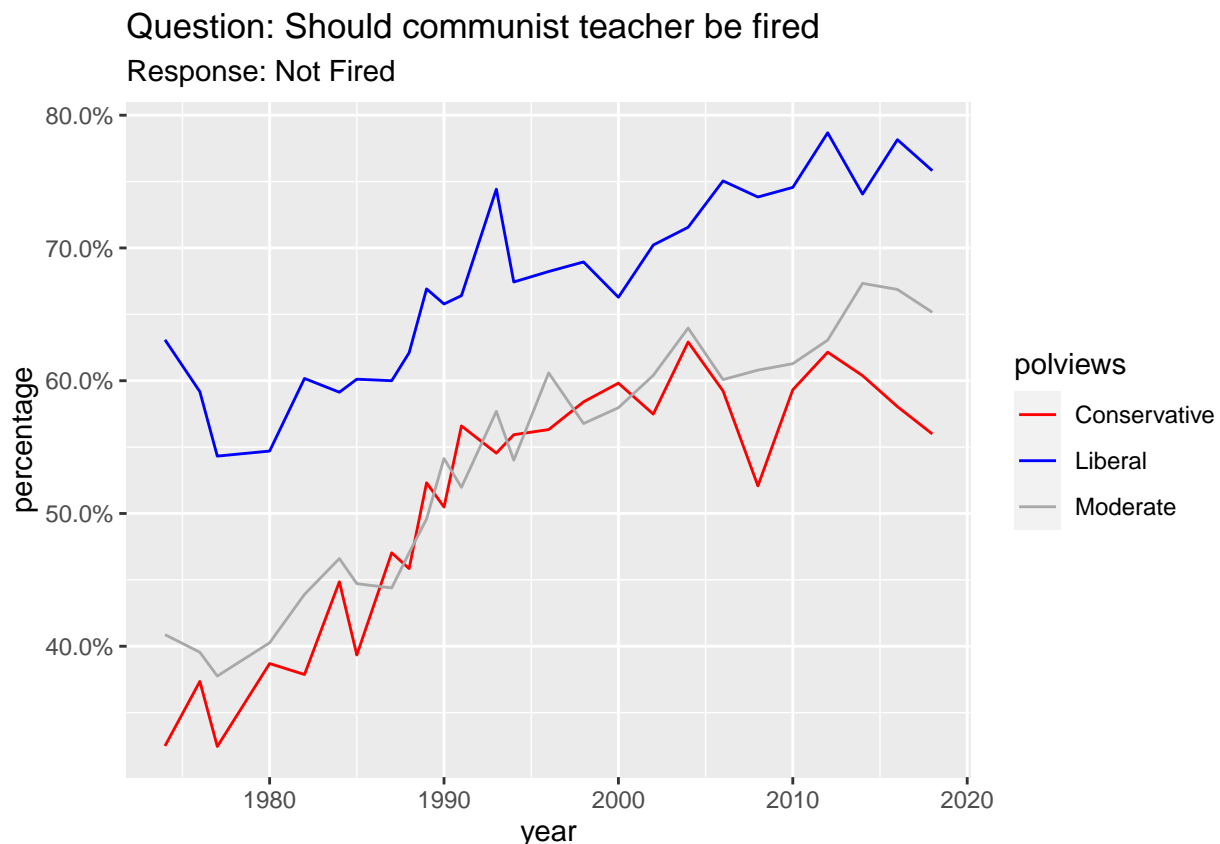
Should a college professor who claims to be communist be fired?

We will see the percentage of respondents of each political party over the years. The following figures illustrate that all three parties concur to the idea that a university professor who is a communist should not be fired. One can observe that the percentage of conservative acceptance of a communist professor not being fired in 1992 is equivalent to the liberal acceptance in 1980. One could purport the idea that the acceptance of a communist professor in university will have an immensely pernicious effect on society; by taking positions in education these professors will greatly impact the future of the country. Indoctrinating the youth allows one to change the course of the future.

```
dataA9 <- select(data, year, colcom, polviews) %>% filter(!is.na(colcom), !is.na(polviews), polviews == 1)
dataA9 <- mutate(dataA9, colcom = ifelse(colcom == 5, 1, 0)) %>% mutate(polviews = ifelse(polviews == 1, 1, 0))

table_Not_fired <- tibble(summarise(group_by(dataA9, year, polviews), percentage = mean(colcom)))

ggplot(table_Not_fired) +
  geom_line(mapping = aes(x = year, y = percentage, color = polviews)) +
  scale_y_continuous(labels = scales::percent) +
  labs(title = "Question: Should communist teacher be fired", subtitle = "Response: Not Fired") +
  scale_color_manual(values=c("red", "blue", "dark grey"))
```



Is it wrong for same-sex adults to have sexual relations?

We will see the percentage of respondents of each party over the years. Similar to our previous findings, in this figure we can observe that there is a increase of acceptance regarding liberal ideas; it is self-evident that

all political parties have over the years grown more leftist. Here, all three parties have grown significantly in percentage regarding the acceptance of same-sex relations.

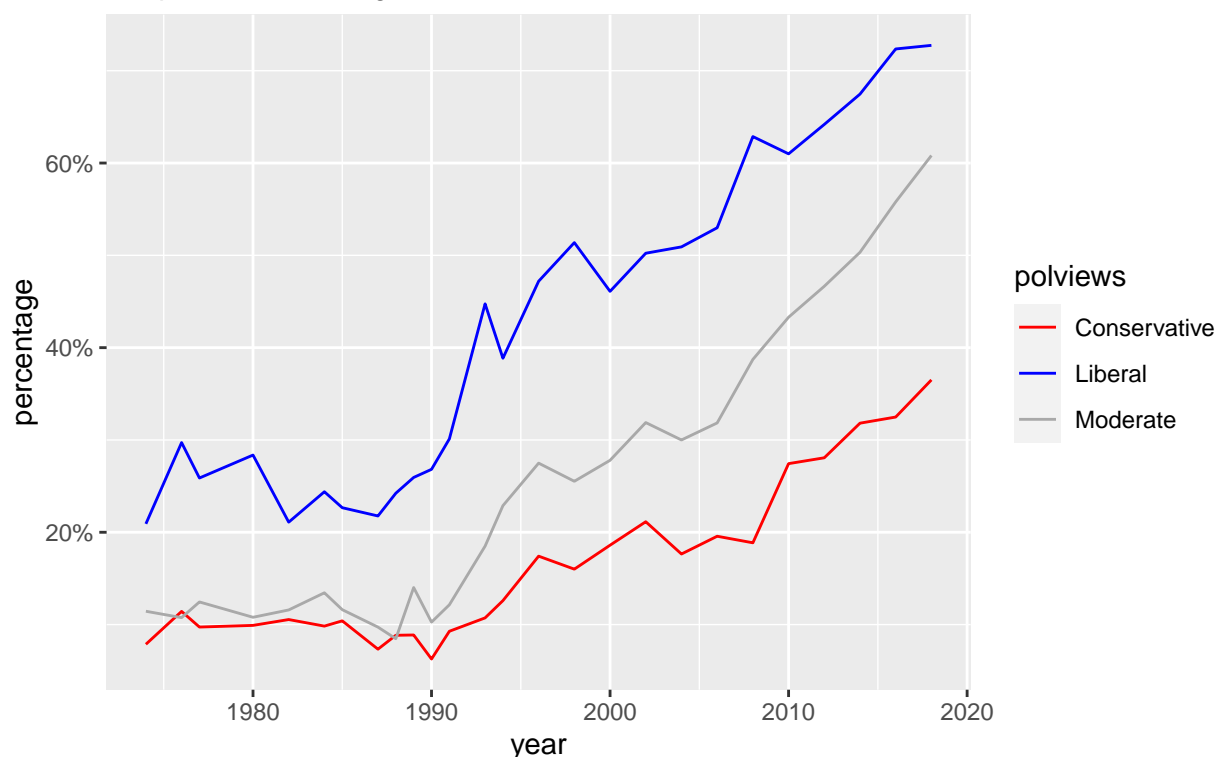
```
dataA10 <- select(data, year, homosex, polviews) %>% filter(!is.na(homosex), !is.na(polviews), polviews != 0)
dataA10 <- mutate(dataA10, homosex = ifelse(homosex == 4, 1, 0)) %>% mutate(polviews = ifelse(polviews == 4, 1, 0))

table_Not_wrong_at_all <- tibble(summarise(group_by(dataA10, year, polviews), percentage = mean(homosex)))

ggplot(table_Not_wrong_at_all) +
  geom_line(mapping = aes(x = year, y = percentage, color = polviews)) +
  scale_y_continuous(labels = scales::percent) +
  labs(title = "Question: Is it wrong for same-sex adults to have sexual relations", subtitle = "Response: Not wrong at all") +
  scale_color_manual(values=c("red", "blue", "dark grey"))
```

Question: Is it wrong for same-sex adults to have sexual relations

Response: Not wrong at all



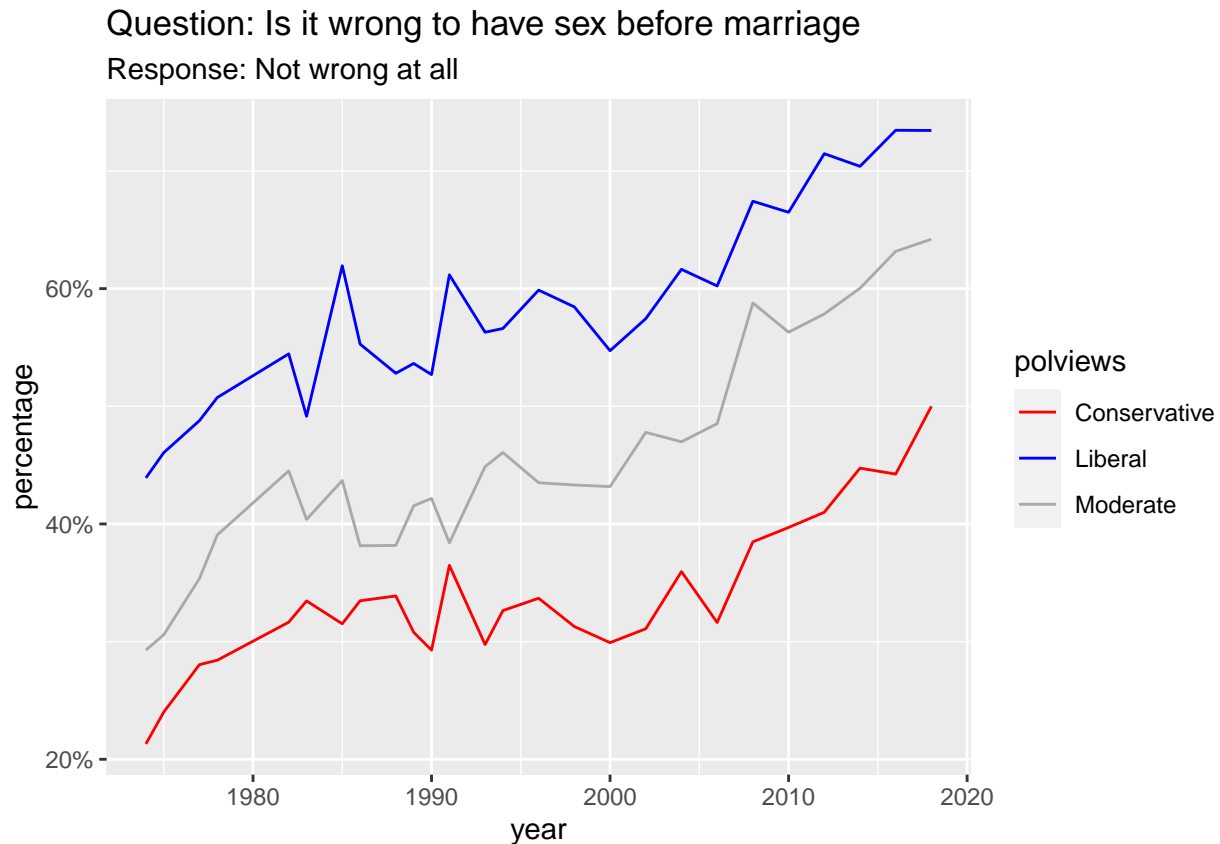
Is it wrong to have sex before marriage?

We will see the percentage of respondents of each party over the years. Again, not surprisingly, we find that all three parties have leaned towards a more liberal mindset over the years. Conservatives and moderates alike have grown to accept the liberal idea of sexual freedom and promiscuity. The percentage of conservative acceptance of sex before marriage in 2018 is equivalent to the acceptance percentage of liberals in the 1970s.

```
dataA11 <- select(data, year, premarsx, polviews) %>% filter(!is.na(premarsx), !is.na(polviews), polviews != 0)
dataA11 <- mutate(dataA11, premarsx = ifelse(premarsx == 4, 1, 0)) %>% mutate(polviews = ifelse(polviews == 4, 1, 0))

table_Not_wrong_at_all <- tibble(summarise(group_by(dataA11, year, polviews), percentage = mean(premarsx)))
```

```
ggplot(table_Not_wrong_at_all) +
  geom_line(mapping = aes(x = year, y = percentage, color = polviews)) +
  scale_y_continuous(labels = scales::percent) +
  labs(title = "Question: Is it wrong to have sex before marriage", subtitle = "Response: Not wrong at all") +
  scale_color_manual(values=c("red", "blue", "dark grey"))
```



Question #4: Linear model predicting polviews

Here we created a linear model to predict polviews according to the following parameters: bible, grass (marijuana), homosex, abany (abortion), and colcom. In this linear model we get an R-squared value of 0.144 which agrees with our previous findings.

```
data <- filter(data, polviews == 1 | polviews == 2 | polviews == 3 | polviews == 4 | polviews == 5 | polviews == 6)
polviews_fit <- linear_reg() %>% set_engine("lm") %>% fit(polviews ~ bible + grass + homosex + abany + colcom)
polviews_fit %>% tidy()
```

```
## # A tibble: 6 x 5
##   term          estimate std.error statistic    p.value
##   <chr>         <dbl>     <dbl>    <dbl>    <dbl>
## 1 (Intercept)   3.88      0.176    22.1 1.01e-104
## 2 bible        -0.129    0.0229   -5.62 1.92e- 8
## 3 grass         0.254    0.0335    7.59 3.50e- 14
## 4 homosex      -0.189    0.0129  -14.6 9.22e- 48
```

```
## 5 abany      0.469    0.0335    14.0  6.20e- 44
## 6 colcom    -0.0568   0.0321    -1.77 7.65e- 2
```

```
glance(polviews_fit)$adj.r.squared
```

```
## [1] 0.1427118
```

Conclusion

The GSS gathers information from contemporary American Society in order to monitor and explain trends and changes in attitudes, behaviors, and attributes. The data provided by the GSS inventively calls for a myriad of questions. In this exploratory data analysis (EDA) we utilized a variety of functions and packages to answer some questions which we deemed interesting and didactic.

Our first question aimed to describe and visualize the change in political parties over the years. First, we had to clean the data and methodically organize it in a more succinct and understandable manner. Then, utilizing persuasive graphing tools we were able to observe the change in proportion of the political parties (Liberal, Conservative, Moderate) over the years (1974-2018). It became evident that though the proportion did face some fluctuations over the years, it nevertheless maintained a constant rate. That is to say, the distribution of people with differing political views didn't undergo a great change over the span of 44 years.

Though the EDA conducted in question one suggested a sense of stability in the political field, we maintained some skepticism and wanted to further explore the change of political views over the years. Particularly, we aimed to probe more deeply into the data as to determine whether there were any trends towards the liberal political view. The aforementioned is precisely what we aimed to answer in question two. I.e. whether people, as a whole, are leaning towards liberal ideas.

In our EDA of the latter query we indeed found that people are collectively leaning towards liberal ideas. We identified and visualized the trend of liberal opinion in the following topics: homosexuality, religion (bible), marijuana legalization, abortion, communist educators, as well as fornication (sexual intercourse before marriage). In some of these trends we observed percentage changes as high as 40%, as in the case of marijuana legalization. While 20% of people in 1974 approved of marijuana legalization, the percentage escalated to over 60% in 2018.

Upon arriving to these findings, we decided to delve deeper into the data and explore these trends from another view point. Rather than asking whether people were leaning more towards liberal ideas, we aimed to see whether people were leaving conservative ideas and values. Comparable to question three, we approached the data similarly, and aimed to see the trends of change over the years given the same parameters as question three.

We found that from this point of analysis the same conclusion was still maintained: people are leaning towards liberal ideals. We identified that it is not only conservatives that are leaning more left, but even moderates, and liberals themselves have shifted left when compared to their 1974 counterparts. In some cases, such as when faced with the fornication topic, we found that conservatives in 2018 had the same percentage of acceptance of extramarital sex as did the liberals in the 1980s.

Given these findings, we decided to create a linear model to demonstrate the strength of correlation between the parameters investigated and that of political viewpoint. The R-square value was that of 0.144, a very insignificant amount. However, this value further verifies that our EDA was fruitful. Viz, the low R-squared value implies that these parameters assert a rather small amount of influence to that of political views. What this implies is that all three political views have skewed left towards liberal ideas in the past 44 years.

In conclusion, the findings in this EDA seem to suggest that in the past 44 years, from 1974 to 2018, public opinion and values have significantly inclined left; leaning towards a more liberal viewpoint.