Problem Set 1

1. In a paragraph, describe the responses to the `racespeech` question. Your paragraph should include:

- The overall proportion of respondents who agree and disagree
- The proportion of respondents in each race/hisp category who agree and disagree
- The proportion of respondents in each degree category who agree and disagree

```
# Proportion table
pander(round(prop.table(table(ps1$racespeech)),3))
```

Allowed	Not allowed
0.617	0.383

```
# Proportion table with row proportions
pander(round(prop.table(table(ps1$racehisp, ps1$racespeech),1),3))
```

	Allowed	Not allowed
Black	0.538	0.462
Hispanic	0.444	0.556
Other	0.424	0.576
White	0.684	0.316

```
# Relevel degree categories
ps1$degree <- factor(ps1$degree,
                     levels=c("Less Than HS", "HS Diploma",
                              "Some College", "College Degree",
                              "Grad/Prof Degree"))

# Proportion table with row proportions
pander(round(prop.table(table(ps1$degree, ps1$racespeech),1),3))
```

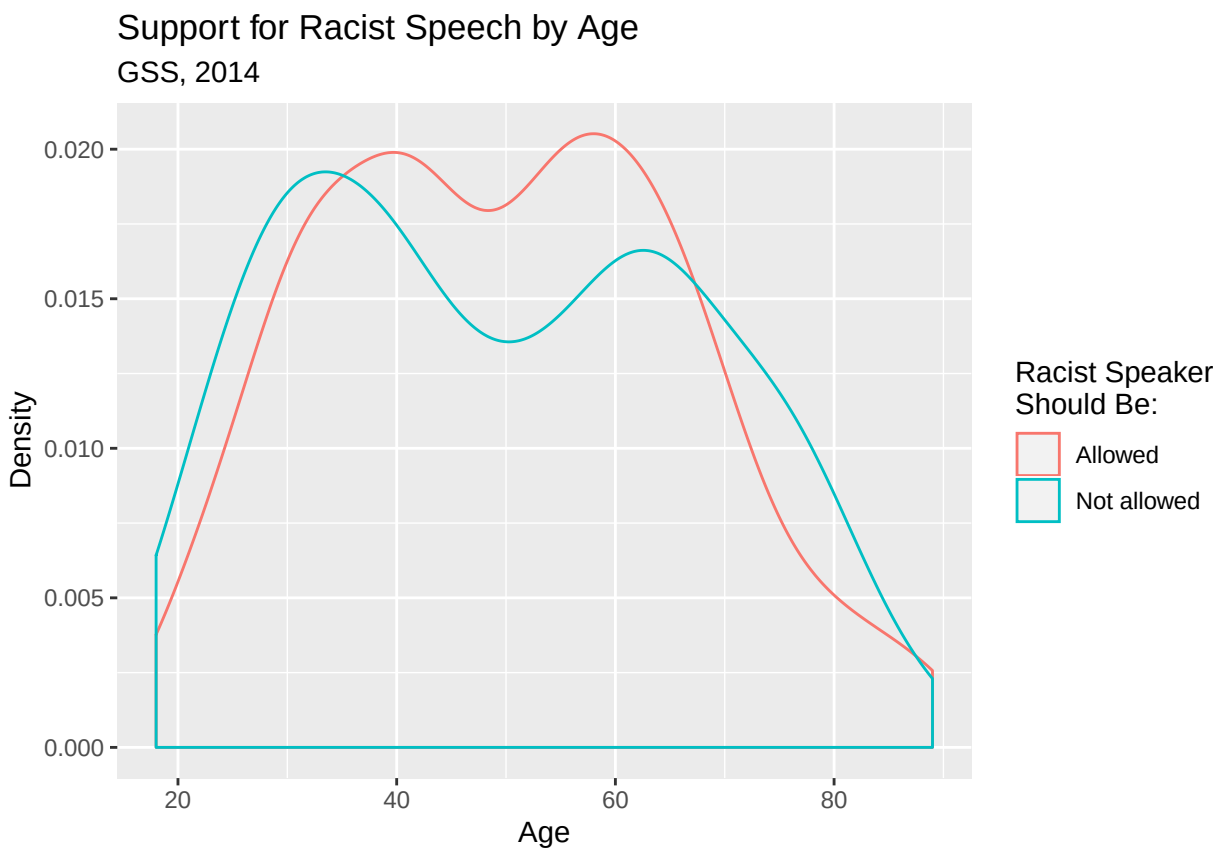
	Allowed	Not allowed
Less Than HS	0.437	0.563
HS Diploma	0.588	0.412
Some College	0.652	0.348
College Degree	0.708	0.292
Grad/Prof Degree	0.755	0.245

2. In a figure, describe how the age distribution differs for respondents who do and do not agree with the question of whether a racist should be allowed to speak. Interpret your figure in words, and offer a possible sociological explanation for the results.

```
# Density plot
racespeech_age <- ggplot(ps1, aes(x = age, color = racespeech))

racespeech_age + geom_density() +
     labs(x = "Age", y = "Density",
          title = "Support for Racist Speech by Age",
          subtitle = "GSS, 2014",
```

```
        color = "Racist Speaker\nShould Be:")
```



3. In a sentence or two, describe male-female differences in rally attendance. What is an additional variable you would want to have in order to explain the relationship between these variables?

```
# Reorder rally levels
ps1$attrally <- factor(ps1$attrally,
                       levels = c("In Past Year", "Before Last Year",
                       "No But I Might", "No And I Will Not"))

# Proportion table with row proportions
pander(round(prop.table(table(ps1$sex, ps1$attrally),1),3))
```

	In Past Year	Before Last Year	No But I Might	No And I Will Not
Female	0.084	0.199	0.373	0.344
Male	0.111	0.219	0.413	0.257

4. In a few sentences, describe class differences in demonstration participation. What is an additional variable you would want to have in order to explain the relationship between these variables?

```
# Reorder class levels
ps1$class <- factor(ps1$class,
                    levels = c("Lower", "Working", "Middle", "Upper"))

# Reorder demonstration levels
ps1$joindem <- factor(ps1$joindem,
                      levels = c("In Past Year", "Before Last Year",
```

```
                        "No But I Might", "No And I Will Not"))

# Proportion table with row proportions
pander(round(prop.table(table(ps1$class, ps1$joindem),1),3))
```

Table 5: Table continues below

	In Past Year	Before Last Year	No But I Might
Lower	0.121	0.155	0.293
Working	0.045	0.122	0.451
Middle	0.044	0.225	0.408
Upper	0.045	0.318	0.182

	No And I Will Not
Lower	0.431
Working	0.383
Middle	0.323
Upper	0.455

```
# If you do not want a long table broken into two...
panderOptions('table.split.table', Inf)
pander(round(prop.table(table(ps1$class, ps1$joindem),1),3))
```

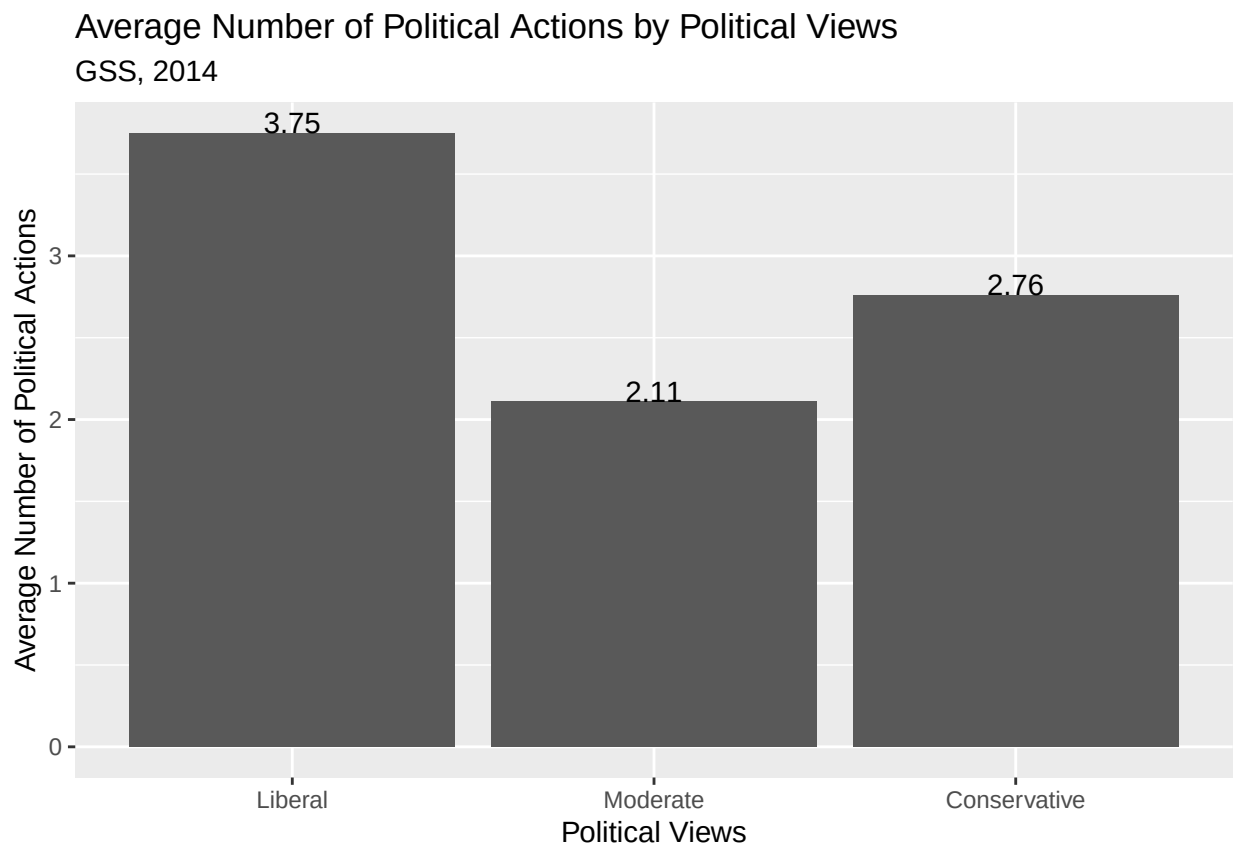
	In Past Year	Before Last Year	No But I Might	No And I Will Not
Lower	0.121	0.155	0.293	0.431
Working	0.045	0.122	0.451	0.383
Middle	0.044	0.225	0.408	0.323
Upper	0.045	0.318	0.182	0.455

5. Collapse the seven categories of political views into three categories: any liberal, moderate, any conservative. Create a figure showing how the average number of political actions varies across the three categories of political views. Interpret this figure in a few sentences.

```
# Create categories
ps1 <- ps1 %>%
     mutate(polcat = ifelse(polviews %in% 1:3, "Liberal",
                            ifelse(polviews == 4, "Moderate",
                                   "Conservative")),
            polcat = factor(polcat,
                            levels = c(
                                 "Liberal", "Moderate", "Conservative")))

# Find mean of actions by category
polcat_action <- ps1 %>%
               group_by(polcat) %>%
               summarize(avg = round(mean(num_action),2))
```

```
# Plot
action_by_polcat <- ggplot(polcat_action, aes(x = polcat, y = avg))
action_by_polcat + geom_col() +
     geom_text(aes(label = avg), vjust = "outward") +
     labs(x = "Political Views",
          y = "Average Number of Political Actions",
          title = "Average Number of Political Actions by Political Views",
          subtitle = "GSS, 2014")
```



6. In a sentence or two, describe the correlation between years of education and number of political actions. (No figure required here.)

```
cor(ps1$num_action, ps1$educ)
```

```
## [1] 0.4061416
```