Homework 3 – Due Monday September 28, 2015

Cheryl Calhoun

09/21/2015

Add your code blocks to this document and alter the Author name to yours. Your submission will consist of your own R Markdown file plus the compiled HTML version of the document.

I encourage you to submit your completed version as soon as possible. We will grade this quickly so that any opportunities for a regrade can be done quickly.

Assignment

Let's expand the GSS to include some new variables on guns and gun control. Specifically, the file `GSS-extract-3.RData` contains five new survery questions:

- `gunlaw`: do you favor or oppose permit requirements for gun ownership?
- `owngun`: does the respondent have a gun in their home?
- `hgunlaw`: should there be more restrictions on handguns?
- `hguncrim`: do restrictions on handguns decrease crime?
- `gunimp`: how important is the gun control issue to the respondent?

```
## Seting up the session.
## install.packages("ggplot2")                      ## which is obvious
## install.packages("MASS")                         ## which we need for some data.
## setwd("C:/Users/07001412/OneDrive/Education/UF/2015/Fall/EDF6938/Week 4")
load ("GSS-extract-3.RData")
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(ggplot2)
##library(MASS)

## Rename variable for same-sex marriage
subGSS.3 <- rename(subGSS.3, SSMar = marhomo)
head(subGSS.3, 2)
```

```
##   year grass       marital age          degree tvhours radiohrs racmar
## 1 1972  <NA> never married  23        bachelor      NA       NA     no
## 2 1972  <NA>       married  70 lt high school      NA       NA    yes
##          premarsx SSMar conrinc owngun hgunlaw hguncrim gunlaw gunimp
## 1 not wrong at all  <NA>      NA   <NA>    <NA>     <NA>  favor   <NA>
## 2     always wrong  <NA>      NA   <NA>    <NA>     <NA>  favor   <NA>
```

Results: The work environment is now setup and we are ready to begin analysing data.

1. The first two have very high response counts compared to the last three. What fraction of respondents provided no response to each question?

Results: The resulting calculations show the fraction of respondents that provided no response to each questions.

```
## Determining the fraction of responents who provided no response to each question.
Missing = c(
sum(is.na(subGSS.3$gunlaw))/summarize(subGSS.3, nagunlaw=n()),
sum(is.na(subGSS.3$owngun))/summarize(subGSS.3, naowngun=n()),
sum(is.na(subGSS.3$hgunlaw))/summarize(subGSS.3, nahgunlaw=n()),
sum(is.na(subGSS.3$hguncrim))/summarize(subGSS.3, nahguncrim=n()),
sum(is.na(subGSS.3$gunimp))/summarize(subGSS.3, nagunimp=n()))
Missing
```

```
## $nagunlaw
## [1] 0.3521032
##
## $naowngun
## [1] 0.3931106
##
## $nahgunlaw
## [1] 0.9780533
##
## $nahguncrim
## [1] 0.9782714
##
## $nagunimp
## [1] 0.9672142
```

2. For the latter three questions: during what survey years were these ones asked?

Results: The results show the years in which the questions were asked and number of present and missing responses for each year.

```
## Determining the years in which "should there be more restrictions on handguns?" is asked.
hgunlaw_asked = subGSS.3 %>% group_by (year) %>% summarize (present = sum(!is.na(hgunlaw)), missing=sum
hgunlaw_asked <- filter(hgunlaw_asked, present > 0)
hgunlaw_asked                                   ## Display the results.
```

```
## # A tibble: 1 × 3
##    year present missing
##   <int>   <int>   <int>
## 1  2004    1308    1504
```

```
## Determining the years in which "do restrictions on handguns decrease crime?" is asked.
hguncrim_asked = subGSS.3 %>% group_by (year) %>% summarize (present = sum(!is.na(hguncrim)), missing=su
hguncrim_asked <- filter(hguncrim_asked, present > 0)
hguncrim_asked                                  ## Display the results.
```

```
## # A tibble: 1 × 3
##    year present missing
##   <int>   <int>   <int>
## 1  2004    1295    1517
```

```
## Determining the years in which "how important is the gun control issue to the respondent?" is asked.
gunimp_asked = subGSS.3 %>% group_by (year) %>% summarize (present = sum(!is.na(gunimp)), missing=sum(is
```

```
gunimp_asked <- filter(gunimp_asked, present > 0)
gunimp_asked                                      ## Display the results.
```

```
## # A tibble: 2 × 3
##    year present missing
##   <int>   <int>   <int>
## 1  1976    1480      19
## 2  1984     474     999
```

Let's examine the trends over time for the first two with `ggplot2` and `dplyr`.

3. Create a new data frame, grouping by year, marital status and degree, that shows the share of people who favor permit requirements, the share of people who have a gun in their home, and the share who favor permits who also have a gun in their home. How many rows and columns does this data frame have?

Results: The resulting data set has 823 rows and 6 columns.

```
subGSS.33 <- subGSS.3 %>% select(year, marital, degree, gunlaw, owngun)          ## Selecting the d
subGSS.33 <- filter(subGSS.33, !is.na(marital))          ## Remove records missing marital data
subGSS.33 <- filter(subGSS.33, !is.na(degree))           ## Remove records missing degree data
subGSS.33 <- subGSS.33 %>%  arrange(year, marital, degree) %>%
  group_by(year, marital, degree) %>%
  summarize(gunlaw.share=sum(gunlaw=="favor", na.rm = TRUE)/n(),
            owngun.share=sum(owngun=="yes", na.rm = TRUE)/n(),
            favor.both=sum((gunlaw=="favor") && (owngun=="yes"), na.rm = TRUE)/n())

subGSS.33
```

```
## Source: local data frame [735 x 6]
## Groups: year, marital [?]
##
##      year  marital        degree gunlaw.share owngun.share favor.both
##     <int>   <fctr>        <fctr>        <dbl>        <dbl>      <dbl>
## 1    1972  married lt high school    0.6866667            0          0
## 2    1972  married    high school    0.6782450            0          0
## 3    1972  married junior college    0.8750000            0          0
## 4    1972  married       bachelor    0.6781609            0          0
## 5    1972  married       graduate    0.7380952            0          0
## 6    1972  widowed lt high school    0.7500000            0          0
## 7    1972  widowed    high school    0.8367347            0          0
## 8    1972  widowed       graduate    1.0000000            0          0
## 9    1972 divorced lt high school    0.7727273            0          0
## 10   1972 divorced    high school    0.8235294            0          0
## # ... with 725 more rows
```

4. Use `ggplot` to show the trend in these three responses over time. Do this first without any facet division (but coloring by degree); then by faceting by degree to show the overall trends within each group. What do you notice, if anything, about these trends?

Results: All of the charts seem to be working. It does seem that gun ownership and support for gun permits is dropping over time.

```
## Create a chart showing the share of people who favor permit requirements over time.

ggplot(subGSS.33, aes(x=year, y=gunlaw.share, fill=degree, order=desc(degree), na.rm = TRUE)) +
  geom_bar(stat="identity") +
```
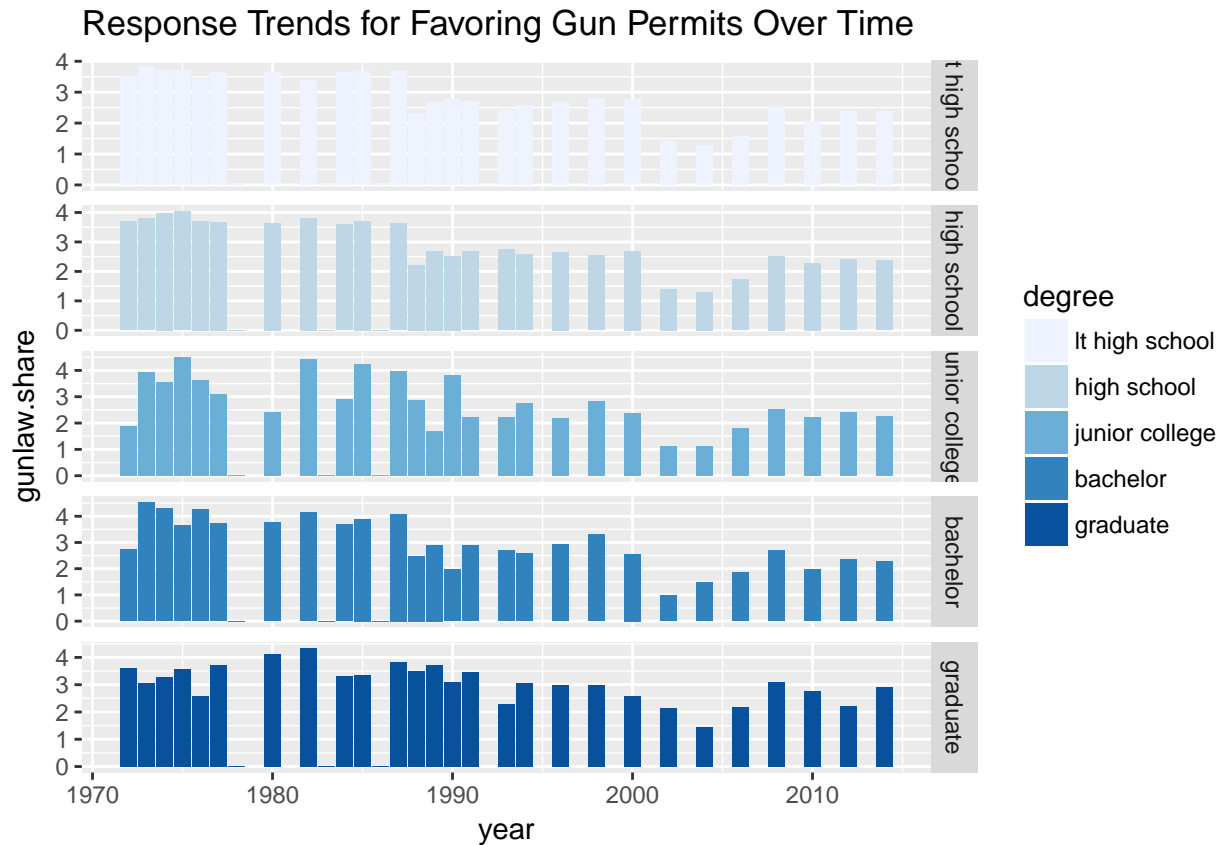
```
scale_fill_brewer(palette=1) +
ggtitle ("Response Trends for Favoring Gun Permits Over Time")
```

Response Trends for Favoring Gun Permits Over Time
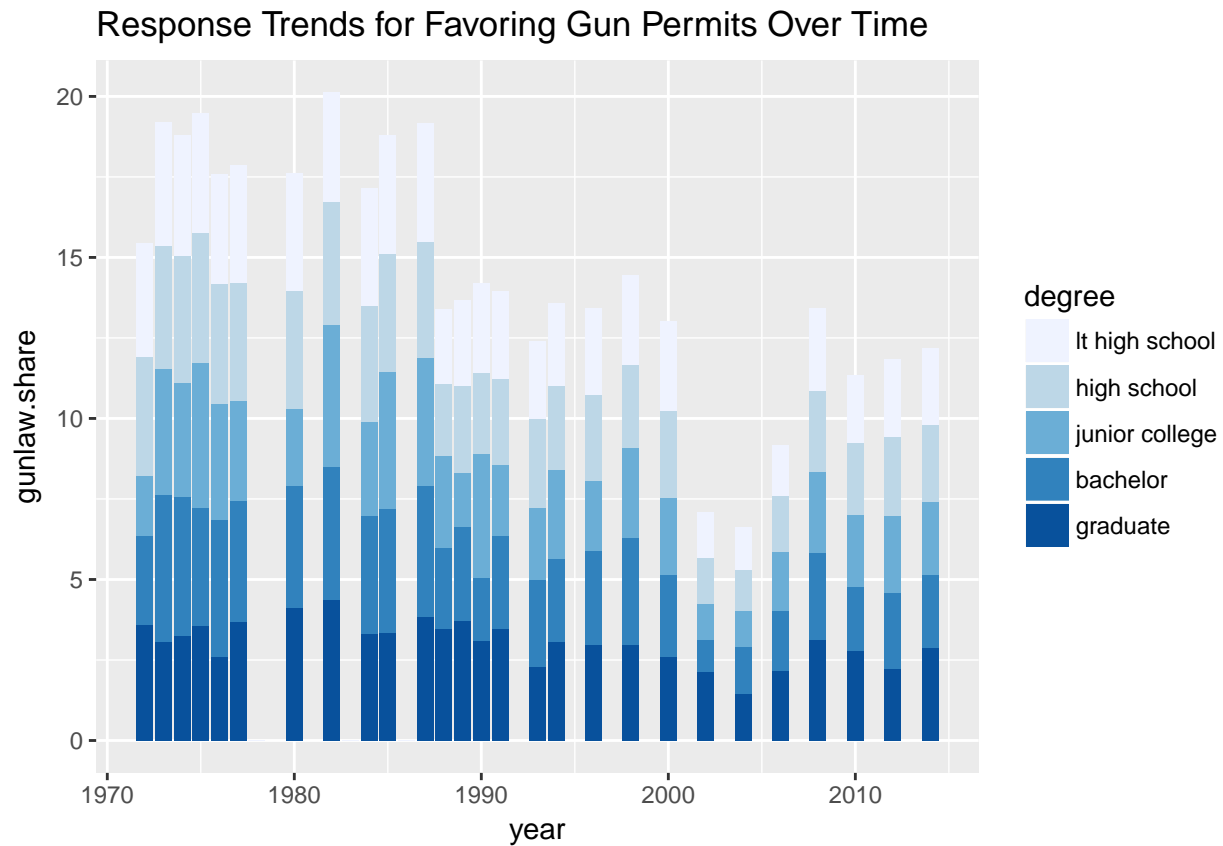
Results: From this chart, it does appear that responses favoring gun permits are dropping over time. There is an interesting drop in 2002 & 2004. I wonder if this has something to do with sentiments following the 9/11 attack.

```
## Create a chart showing the share of people who favor permit requirements over time. With facet divisi
```

```
ggplot(subGSS.33, aes(x=year, y=gunlaw.share, fill=degree, order=desc(degree))) +
  geom_bar(stat="identity") +
  scale_fill_brewer(palette=1) +
  facet_grid (degree ~ ., scales="free")  +
  ggtitle ("Response Trends for Favoring Gun Permits Over Time")
```
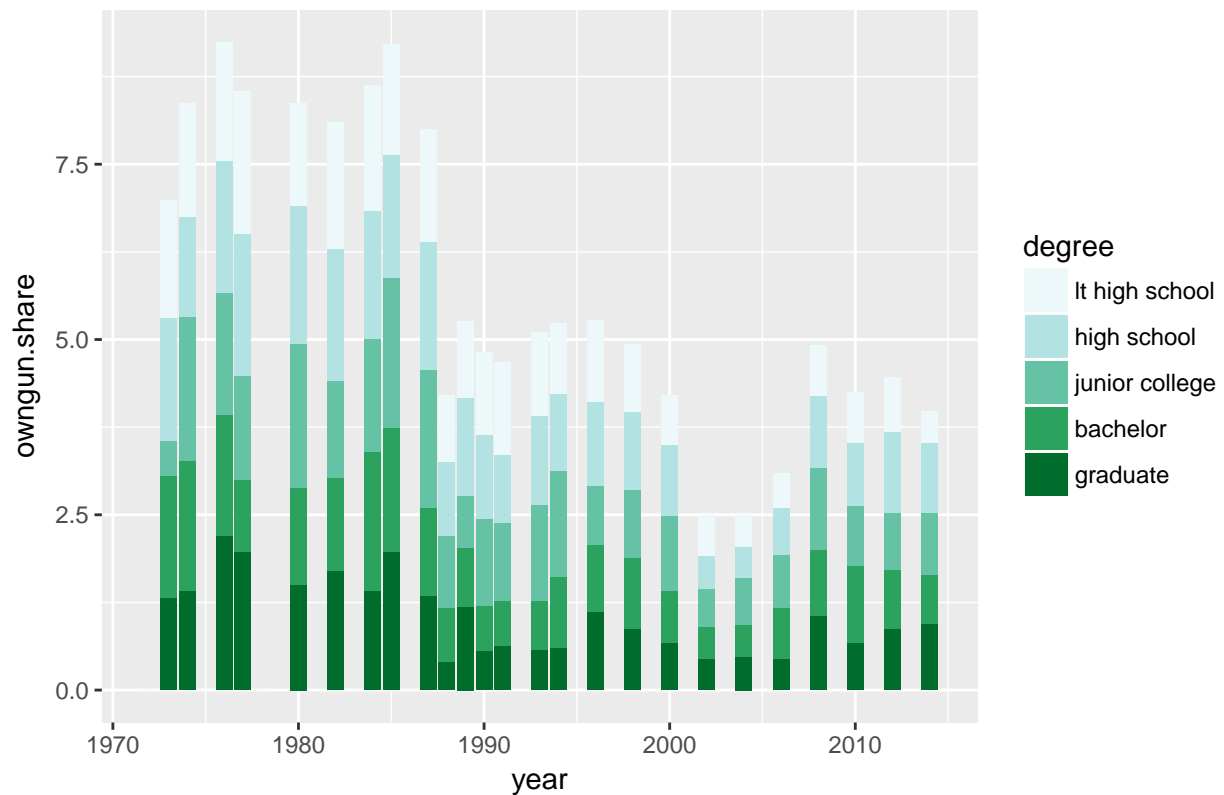
Response Trends for Favoring Gun Permits Over Time

Results: From this chart, you can see some differences in favoring gun permits due to degree, but it seems the overal patterns are very similar. Again, it would be interesting to look at outside influences.

```
## Create a chart showing the share of people who have a gun in their home over time.

ggplot(subGSS.33, aes(x=year, y=owngun.share, fill=degree, order=desc(degree))) +
  geom_bar(stat="identity") +
  scale_fill_brewer(palette=2)  +
  ggtitle ("Response Trends for Gun Ownership Over Time")
```

Response Trends for Gun Ownership Over Time

Results: From this chart, it does appear that responses for gun ownership are dropping over time. There is an interesting to see the same drop here in 2002 & 2004.

```
## Create a chart showing the share of people who have a gun in their home over time. With facet divisi

ggplot(subGSS.33, aes(x=year, y=owngun.share, fill=degree, order=desc(degree))) +
  geom_bar(stat="identity") +
  scale_fill_brewer(palette=2) +
  facet_grid (degree ~ ., scales="free", space="free")  +
  ggtitle ("Response Trends for Gun Ownership Over Time")
```

Response Trends for Gun Ownership Over Time
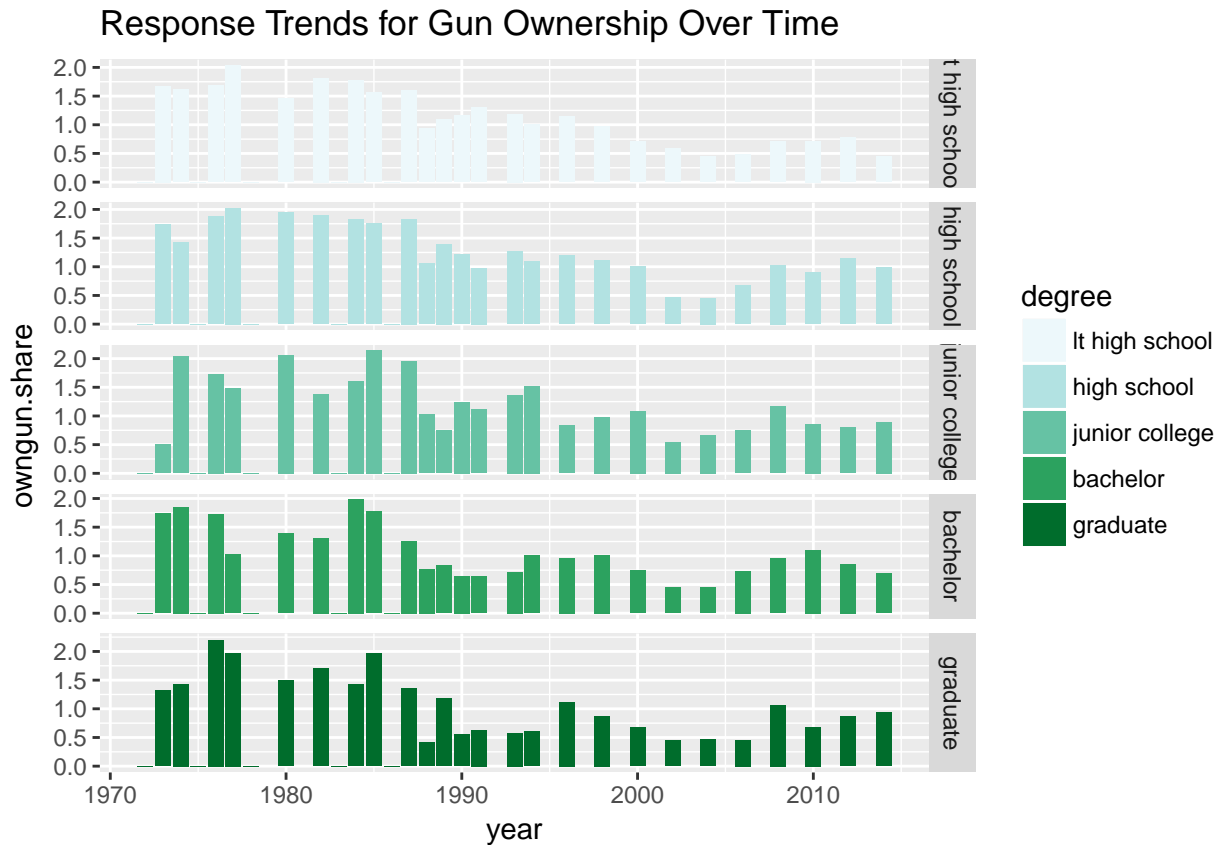
Results: Again, you can see some differences by degree, but it seems the overal patterns are very similar, suggesting outside influences have more influence than level of education.

```
## Create a chart showing the share who favor permits who also have a gun in their home over time.

ggplot(subGSS.33, aes(x=year, y=favor.both, fill=degree, order=desc(degree), na.rm = TRUE)) +
  geom_bar(stat="identity") +
  scale_fill_brewer(palette=3) +
    ggtitle ("Response Trends for Favoring Gun Permits by Gun Owners")
```

Response Trends for Favoring Gun Permits by Gun Owners
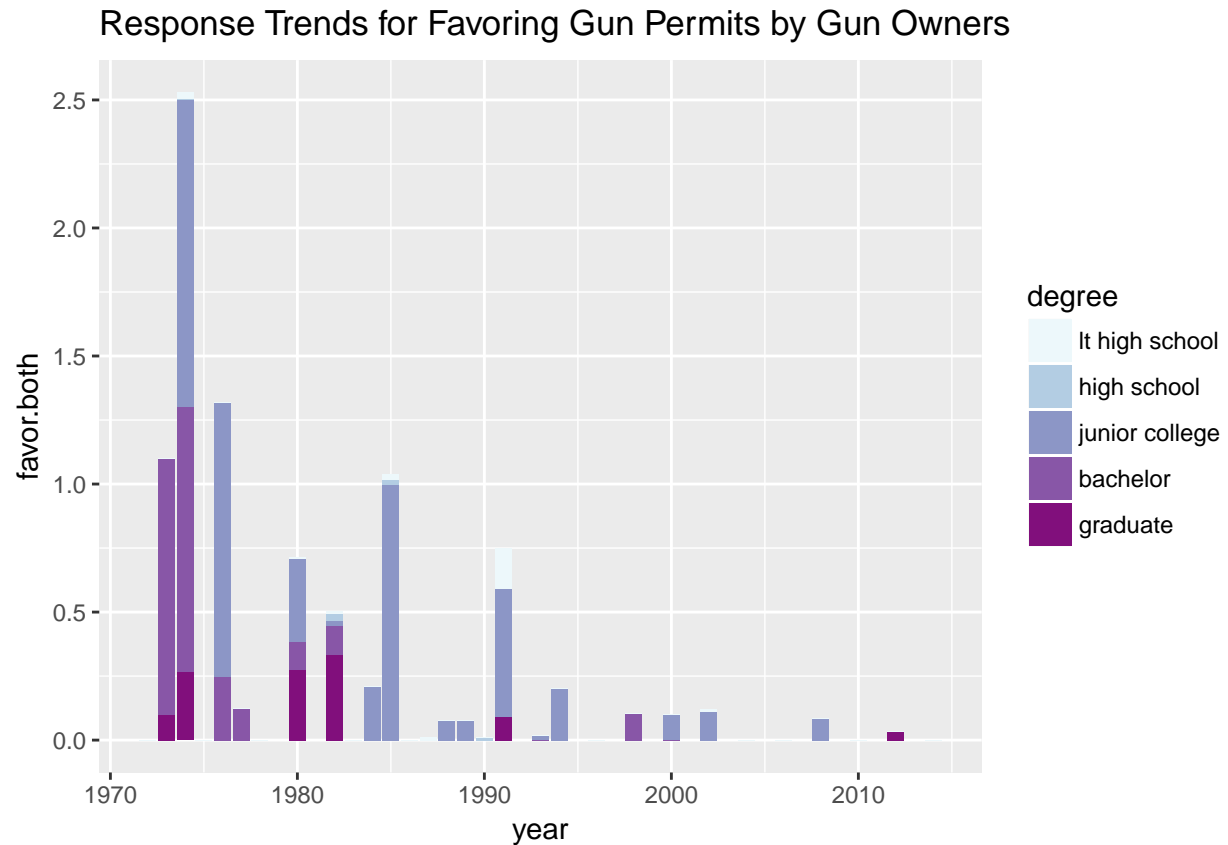
Results: The data set is much smaller her as there are fewer individuals who own guns and also favor gun permits. There is also a strong decrease from the 1970s to the 2000s.

```
## Create a chart showing the share who favor permits who also have a gun in their home over time. With

ggplot(subGSS.33, aes(x=year, y=favor.both, fill=degree, order=desc(degree), na.rm = TRUE)) +
  geom_bar(stat="identity") +
  scale_fill_brewer(palette=3)  +
  facet_grid (degree ~ ., scales="free") +
  ggtitle ("Response Trends for Favoring Gun Permits by Gun Owners")
```

Response Trends for Favoring Gun Permits by Gun Owners
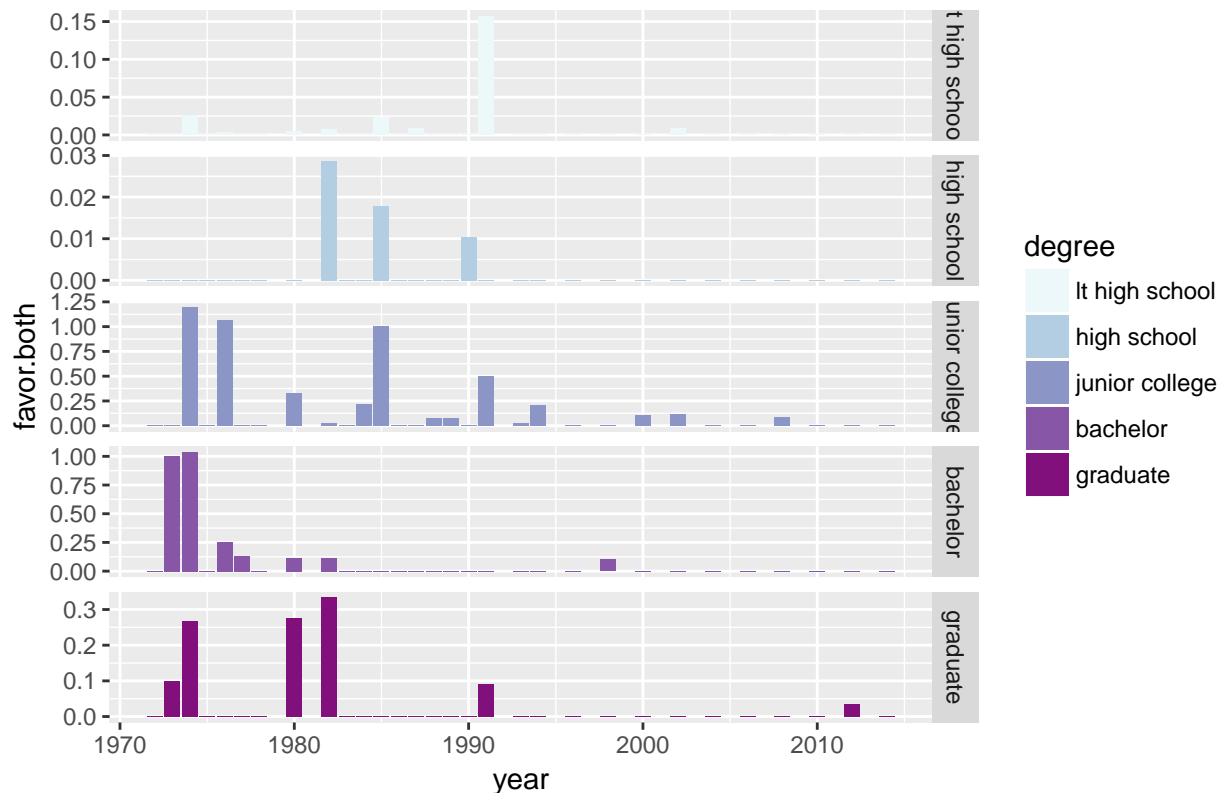
Results: Again, you can see some differences by degree. It is interesting, that there is very little overlab in favoring permits and gun ownership in the lt HS and HS degrees, much more consistent overlab in JC, and higher overlap in 1970s-1980s for B&G.

5. Take the instructions you used for the data frame in question **3**, and produce a new data frame with the same columns plus the other social variables that we had in previous cases: marijuana legalization, interracial marriage, premarital sex ("not wrong" vs the other categories). Add also the mean income within each group. How many rows and columns does this data frame have?

Results: The resulting data set has 823 rows and 9 columns.

```
subGSS.35 <- subGSS.3 %>%
  select(year, marital, degree, grass, racmar, premarsx, conrinc, gunlaw, owngun)
subGSS.35 <- filter(subGSS.35, !is.na(marital))           ## Remove records missing marital data
subGSS.35 <- filter(subGSS.35, !is.na(degree))            ## Remove records missing degree data %>%
subGSS.35 <- subGSS.35 %>%
  arrange(year, marital, degree) %>%
  group_by(year, marital, degree) %>%
  summarize(grass.share=sum(grass=="legal", na.rm = TRUE)/n(),
            racmar.share=sum(racmar=="yes", na.rm = TRUE)/n(),
            premarsx.share=sum(premarsx!="always wrong", na.rm = TRUE)/n(),
            gunlaw.share=sum(gunlaw=="favor", na.rm = TRUE)/n(),
            owngun.share=sum(owngun=="yes", na.rm = TRUE)/n(),
            favor.both=sum((gunlaw=="favor") && (owngun=="yes"), na.rm = TRUE)/n())
subGSS.35

## Source: local data frame [735 x 9]
## Groups: year, marital [?]
```

```
## 
##      year  marital         degree grass.share racmar.share premarsx.share
##     <int>   <fctr>         <fctr>        <dbl>        <dbl>          <dbl>
## 1   1972  married lt high school           0    0.4600000      0.5111111
## 2   1972  married    high school           0    0.2943327      0.6106033
## 3   1972  married junior college           0    0.3125000      0.6875000
## 4   1972  married        bachelor          0    0.1839080      0.7931034
## 5   1972  married        graduate          0    0.1190476      0.7857143
## 6   1972  widowed lt high school           0    0.4886364      0.3409091
## 7   1972  widowed    high school           0    0.4081633      0.3673469
## 8   1972  widowed        graduate          0    0.0000000      0.0000000
## 9   1972 divorced lt high school           0    0.3181818      0.6818182
## 10  1972 divorced    high school           0    0.1176471      0.8235294
## # ... with 725 more rows, and 3 more variables: gunlaw.share <dbl>,
## #   owngun.share <dbl>, favor.both <dbl>
```

Using this bigger data frame:

6. Plot the share of gun ownership against the share of favoring permits across all years (that is, do not have year as a value on the x or y axis) using `geom_point()`. Change the fill color to one of the other variables to see if you can spot a trend; do the same with point size, using the help menu for `geom_point()` to learn how to do this. You may pick any two that you choose from the social variables.

Results: The chart shows the share of gun ownership against the share of favoring gun permits across all years. Green shading is used to display share of favorable marijana legalization responses, and the size of the dots is used to indicate the share of favorable interracial marriage responses.

```
## Create a plot for the share of gun ownership against the share of favoring permits.
plot1 <- ggplot(subGSS.35, aes(x=owngun.share, y=gunlaw.share, size = racmar.share, colour = grass.share
  geom_point(stat="identity") +
  scale_colour_gradient(low = "green") +
  scale_size_area() +
  ggtitle ("Share of Gun Ownership vs. Favoring Permits") +
  xlab("Gun Ownership") +
  ylab("Favoring Gun Permits")
plot1
```

Share of Gun Ownership vs. Favoring Permits
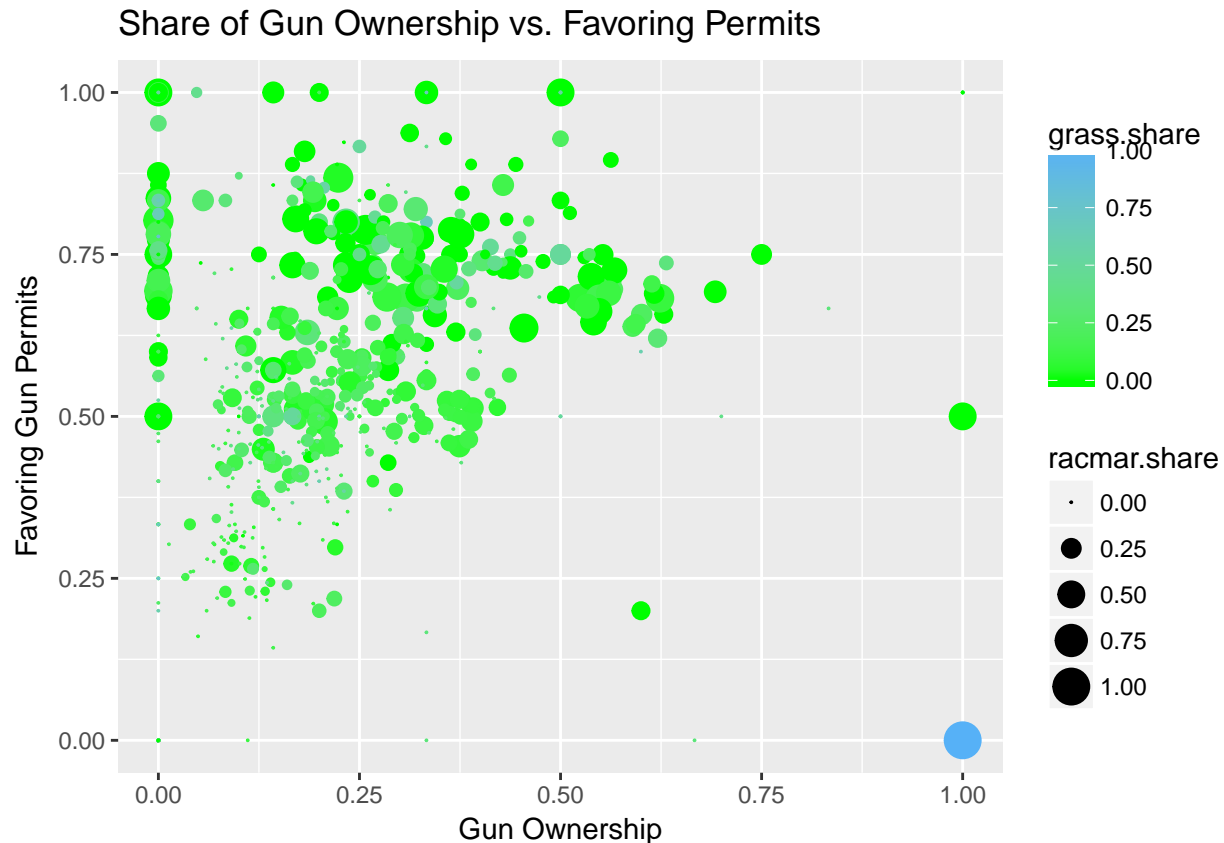
Results This plot seems to show that the more likely one is to favor gun permits, the less likely they are to own a gun. The big blue dot at X=0, y=1 seems to indicate a strong correlation with favoring gun permits and not owning guns. The dots are harder to interpret, but it does seem to indicate a stronger relationship between favoring gun permits and approving of interracial marriage.

7. Similarly, create a series of new plots for the share of gun ownership against each of the other social variables (as many plots as variables). In each case, pick two other social variables to affect the color and size of the remaining points.

Results: There are three additional plots showing share of favorable responses to Interracial Marriage, Marijuana Legalization, and Premarrital Sex. Each chart uses two other variables to affect color gradient and size of dot. These variables are labled in the chart legend.

```
## Create a plot for the share of gun ownership against the approval of interracial marriage.
plot2 <-
  ggplot(subGSS.35, aes(x=owngun.share, y=racmar.share, size = grass.share, colour = premarsx.share)) +
  geom_point(stat="identity") +
  scale_colour_gradient(low = "red") +
  scale_size_area() +
  ggtitle ("Share of Gun Ownership vs. Interracial Marriage") +
  xlab("Gun Ownership") +
  ylab("Favoring Interracial Marriage")
plot2
```

## Share of Gun Ownership vs. Interracial Marriage
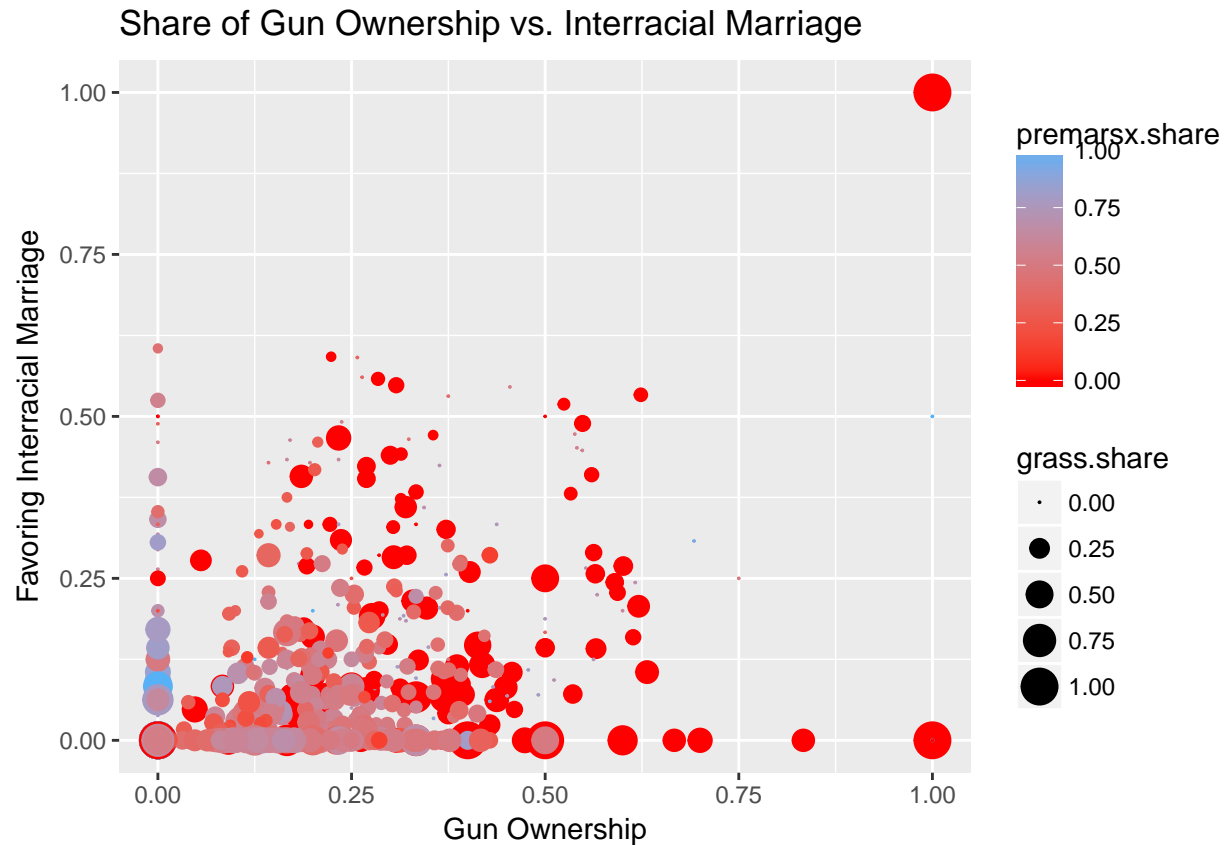
Results This plot seems to show that the less likely you are to own a gun, the less likely you are to favor interracial marriage.

```
## Create a plot for the share of gun ownership against the share of favoring interracial marriage
plot3 <-
  ggplot(subGSS.35, aes(x=owngun.share, y=grass.share, size = premarsx.share, colour = gunlaw.share)) +
  geom_point(stat="identity") +
  scale_colour_gradient(low = "purple") +
  scale_size_area() +
  ggtitle ("Share of Gun Ownership vs. Marijuana Legalization") +
  xlab("Gun Ownership") +
  ylab("Favoring Marijuana Legalization")
plot3
```

Share of Gun Ownership vs. Marijuana Legalization
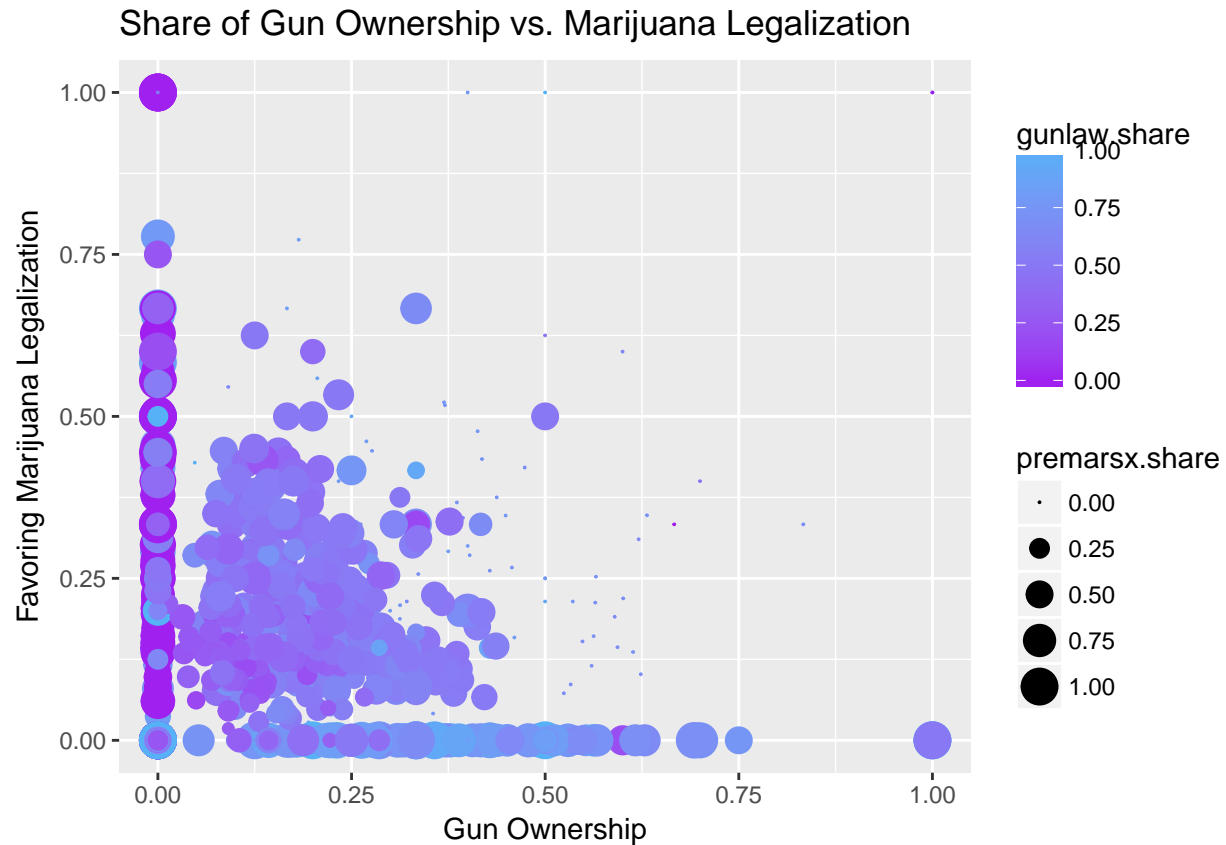
Results: This plot seems to show that those least likely to favor marijuana legalization are also least likely to own a gun. (This seems counter intuitive to me.) Again, I can't really say that the dot size, or coloring helps me much. Maybe there are just too many data points for this to be effective.

```
## Create a plot for the share of gun ownership against the share of individuals who approve of premari
plot4 <-
  ggplot(subGSS.35, aes(x=owngun.share, y=premarsx.share, size = gunlaw.share, colour = grass.share)) +
  geom_point(stat="identity") +
  scale_colour_gradient(low = "green") +
  scale_size_area() +
  ggtitle ("Share of Gun Ownership vs. Premarital Sex") +
  xlab("Gun Ownership") +
  ylab("Acceptance of  Premarital Sex")
plot4
```
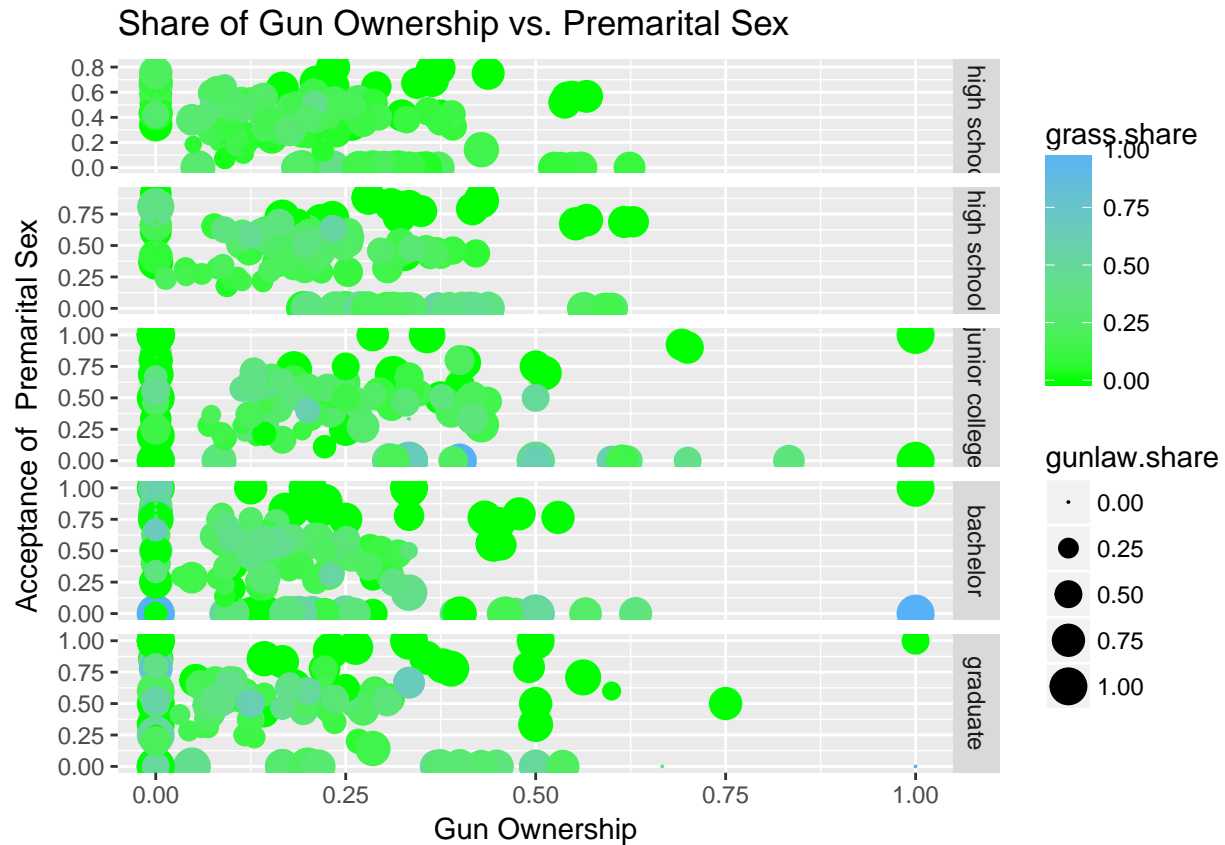
Share of Gun Ownership vs. Premarital Sex
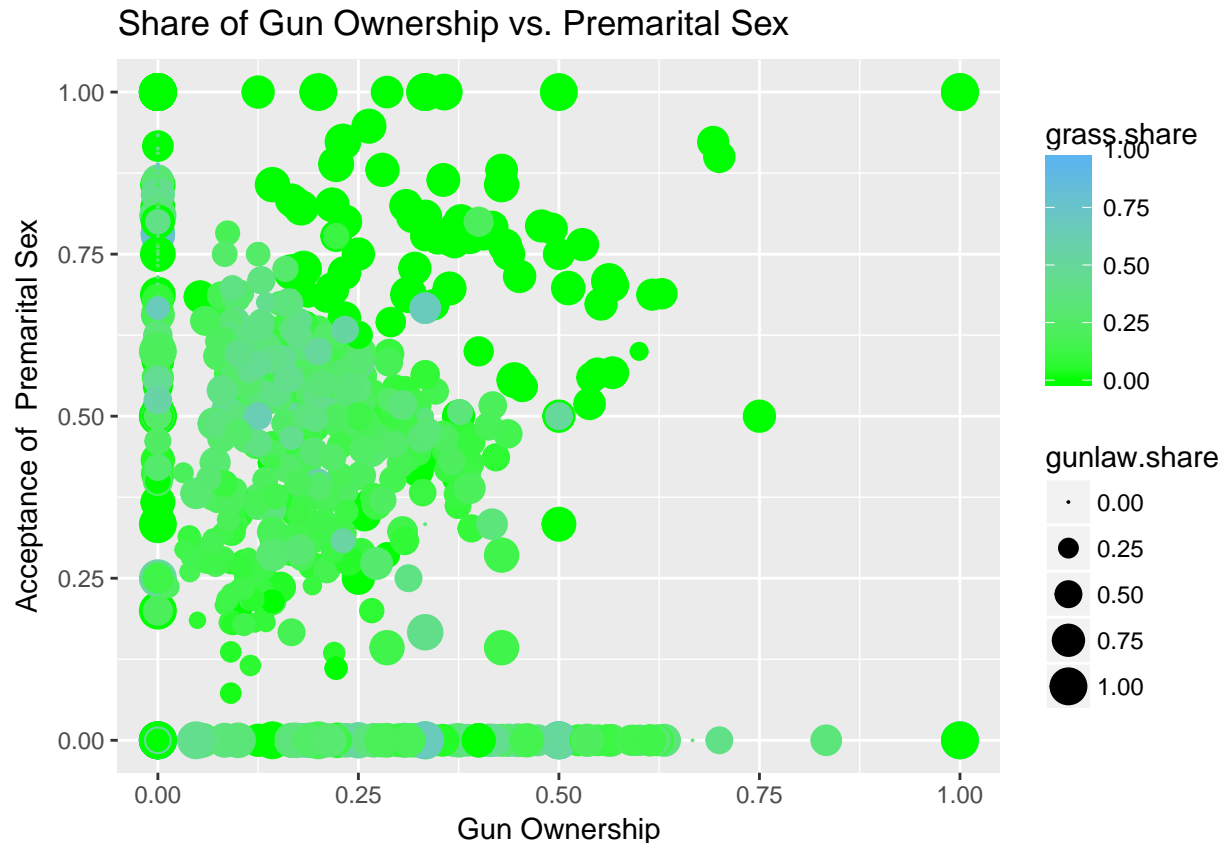
Results: From this plot it seems if you believe premarital sex is "not wrong", you are less likely to own a gun. Or if you don't own a gun, you are more likely to believe premarital sex is "not wrong". (Note: I suspect we could extract some religous, conservative right values from this plot)

8. From the previous question, choose one of these plots that you find the most interesting based on the size and color features you have observed. Use the facets feature to produce a new plot that explores this on the grouping variables. Describe what you see as the most interesting relationship across that facet decomposition.

Results: I choose the Share of Gun Ownership vs. Premarital Sex, partially because I feld the original chart was too jumbled and I wanted to better understand why so many dots were concentrated along the 0:1 horizontal and vertical axis.

```
## Create a plot for the share of gun ownership against the share of individuals who approve of premari
ggplot(subGSS.35, aes(x=owngun.share, y=premarsx.share, size = gunlaw.share, colour = grass.share)) +
  geom_point(stat="identity") +
  scale_colour_gradient(low = "green") +
  scale_size_area() +
  ggtitle ("Share of Gun Ownership vs. Premarital Sex") +
  xlab("Gun Ownership") +
  ylab("Acceptance of  Premarital Sex") +
  facet_grid(degree ~ ., scales="free", space="free")
```

Share of Gun Ownership vs. Premarital Sex

Results: By faciting out the plots by degree, and allowing the scales & spacing to adjust to the data, we get a better picture of the data for each category of degree.

```
ggplot(subGSS.35, aes(x=owngun.share, y=premarsx.share, size = gunlaw.share, colour = grass.share)) +
  geom_point(stat="identity") +
  scale_colour_gradient(low = "green") +
  scale_size_area() +
  ggtitle ("Share of Gun Ownership vs. Premarital Sex") +
  xlab("Gun Ownership") +
  ylab("Acceptance of Premarital Sex") +
  facet_grid(. ~ degree, scales="free", space="free")
```

Share of Gun Ownership vs. Premarital Sex
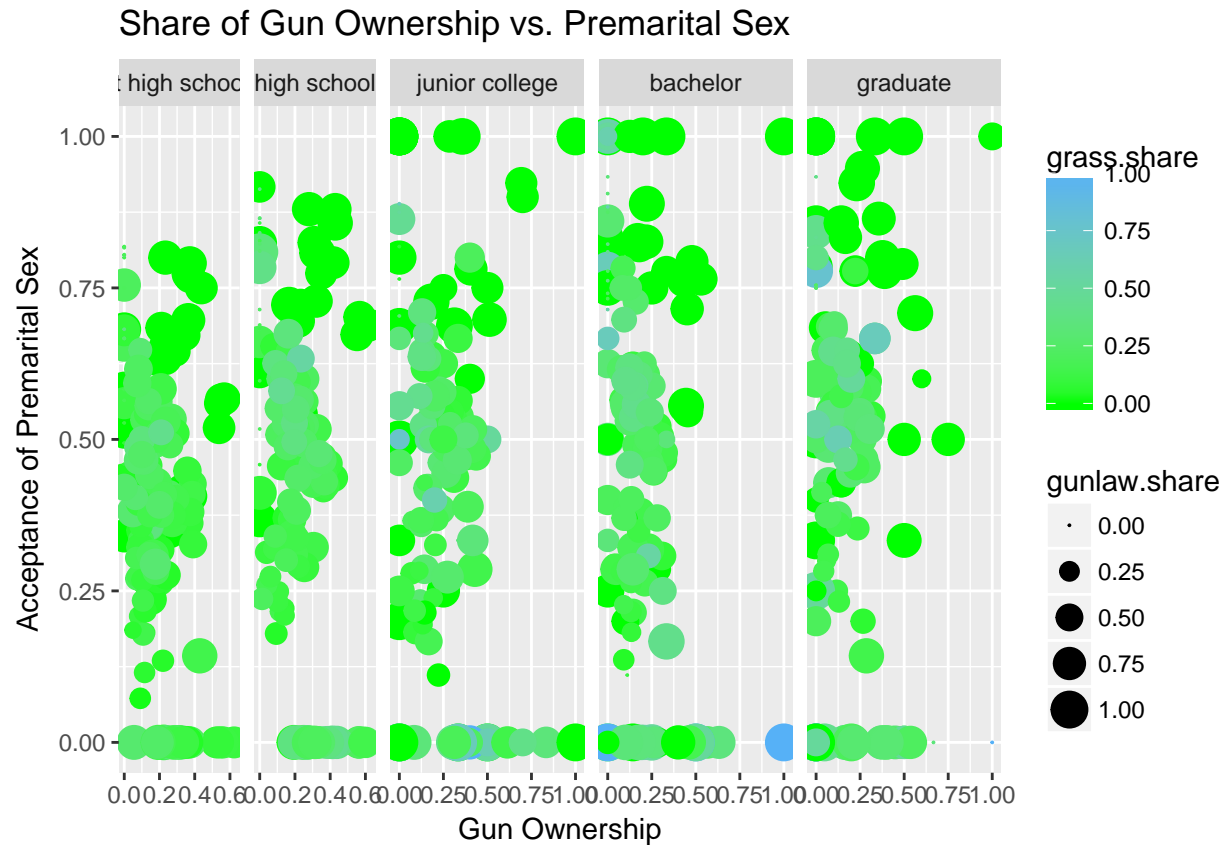
Results: By faciting out the plots by degree, and allowing the scales & spacing to adjust to the data, we get a better picture of the data for each category of degree. This chart shows the charts in a vertical orientation which allows us to better see how the data stretches out across the y axis (Acceptance of Premarital Sex)