

Exploring the BRFSS data

RafaCastle

R Markdown

Setup

Load packages

```
library(ggplot2)
library(dplyr)
```

Load data

Make sure your data and R Markdown files are in the same directory. When loaded your data file will be called `brfss2013`. Delete this note when before you submit your work.

```
load("brfss2013.RData")
```

Part 1: Data

Introduction

The Behavioral Risk Factor Surveillance System (BRFSS) is a United States health survey that looks at behavioral risk factors [Wikipedia]. BRFSS completes more than 400,000 adult interviews each year, making it the largest continuously conducted health survey system in the world [BRFSS Official Page].

Data of interest for this work

In this work, I expect to analyse some behavioral differences between men and women, taking into consideration some factors like their money income, marital status and healthy habits. The data im going to use in order to achieve that is the one presented above.

sex: man or woman maxdrnks: Most Drinks On Single Occasion Past 30 Days educa: Education Level marital: Marital Status menthlth: Number Of Days Mental Health Not Good genhlth: General Health exerany2: Exercise In Past 30 Days

Part 2: Research questions

Research question 1:

How much do Americans drink depending on their gender and education level?

Research question 2:

Is there a correlation between mental health and the marital status of Americans? And if it is, does this correlation depend on gender?

Research question 3:

How healthy do Americans perceive themselves depending on the exercise they do? Does this perception depend on gender?

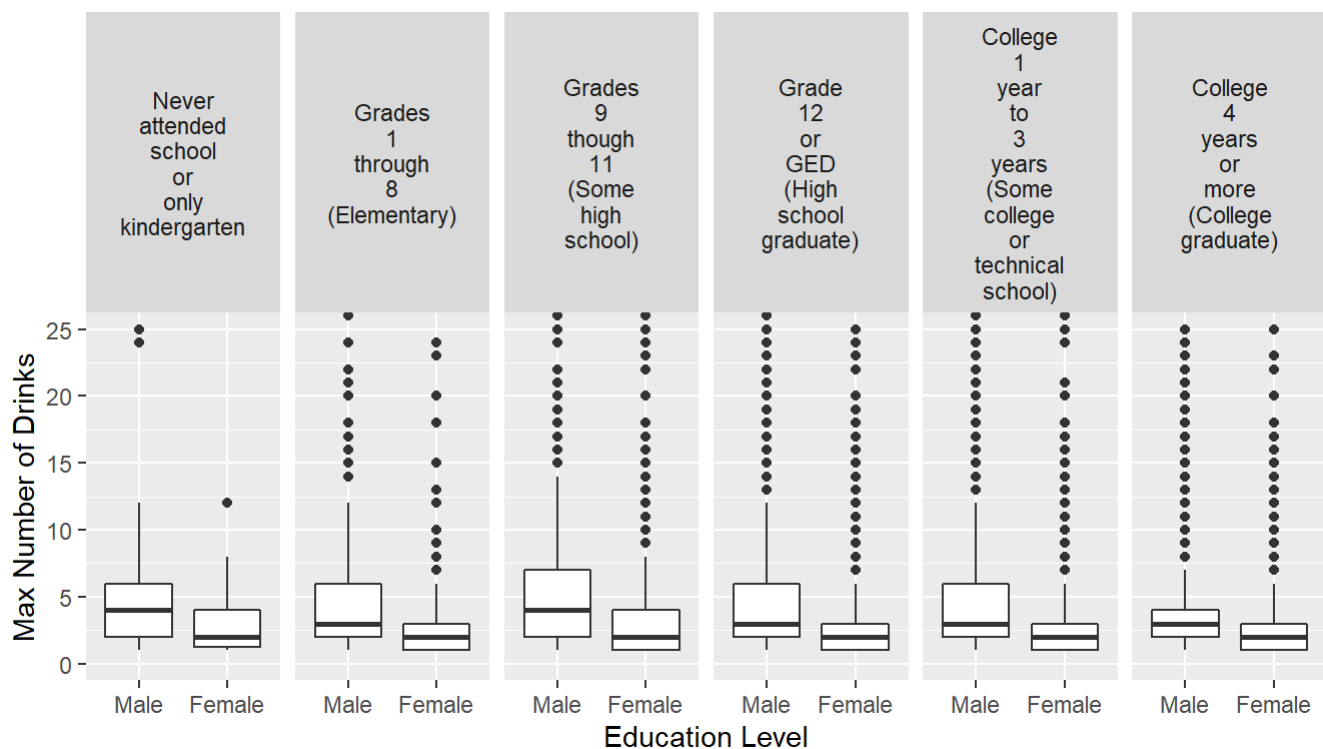
Part 3: Exploratory data analysis

Research question 1:

```
answer_1 <- brfss2013 %>% select(sex, maxdrnks, educa) %>%
  filter(!is.na(sex), !is.na(maxdrnks), !is.na(educa))

levels(answer_1$educa) <- gsub(" ", "\n", levels(answer_1$educa))

ggplot(answer_1, aes(x=sex, y=maxdrnks)) +
  geom_boxplot() +
  coord_cartesian(ylim = c(0, 25)) +
  labs(x = "Education Level", y = "Max Number of Drinks")+
  facet_grid(. ~ educa)
```



The data shows us that men drink more than women at every education level, and there is no clear trend that makes us relate the amount of alcohol with the education level of Americans.

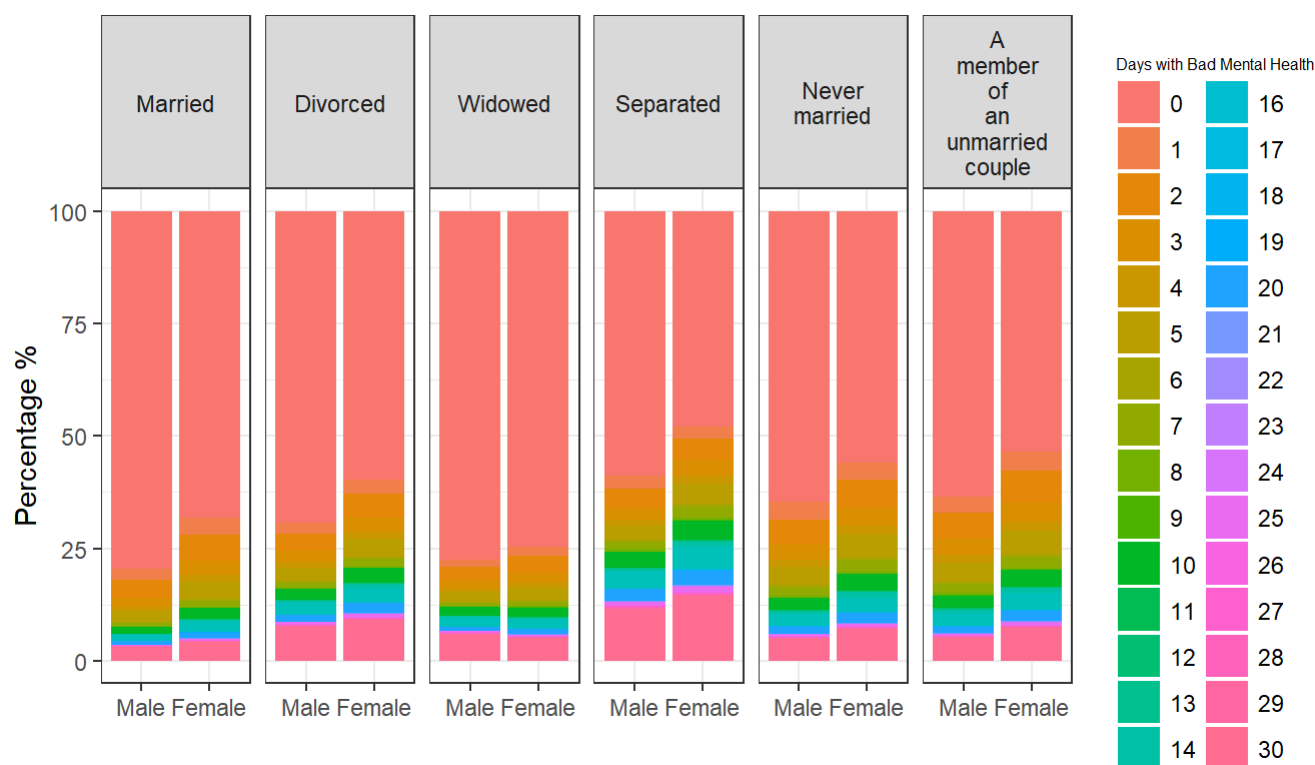
Research question 2:

```
answer_2 <- brfss2013 %>% select(marital , sex, menthlth) %>%
  filter(!is.na(marital), !is.na(sex),!is.na(menthlth)) %>%
  group_by(marital,sex,menthlth) %>%
  summarise(count=n()) %>%
  mutate(perc=count/sum(count))
```

`summarise()` has grouped output by 'marital', 'sex'. You can override using the `.groups` argument.

```
levels(answer_2$marital) <- gsub(" ", "\n", levels(answer_2$marital))
```

```
ggplot(answer_2, aes(x = factor(sex), y = perc*100, fill = factor(menthlth))) +
  geom_bar(stat="identity", width = 0.9) +
  labs(x = " ", y = "Percentage %", fill = "Days with Bad Mental Health") +
  theme_minimal(base_size = 8) +
  theme_bw() +
  facet_grid(. ~ marital) +
  theme(legend.title = element_text(size = 6))
```



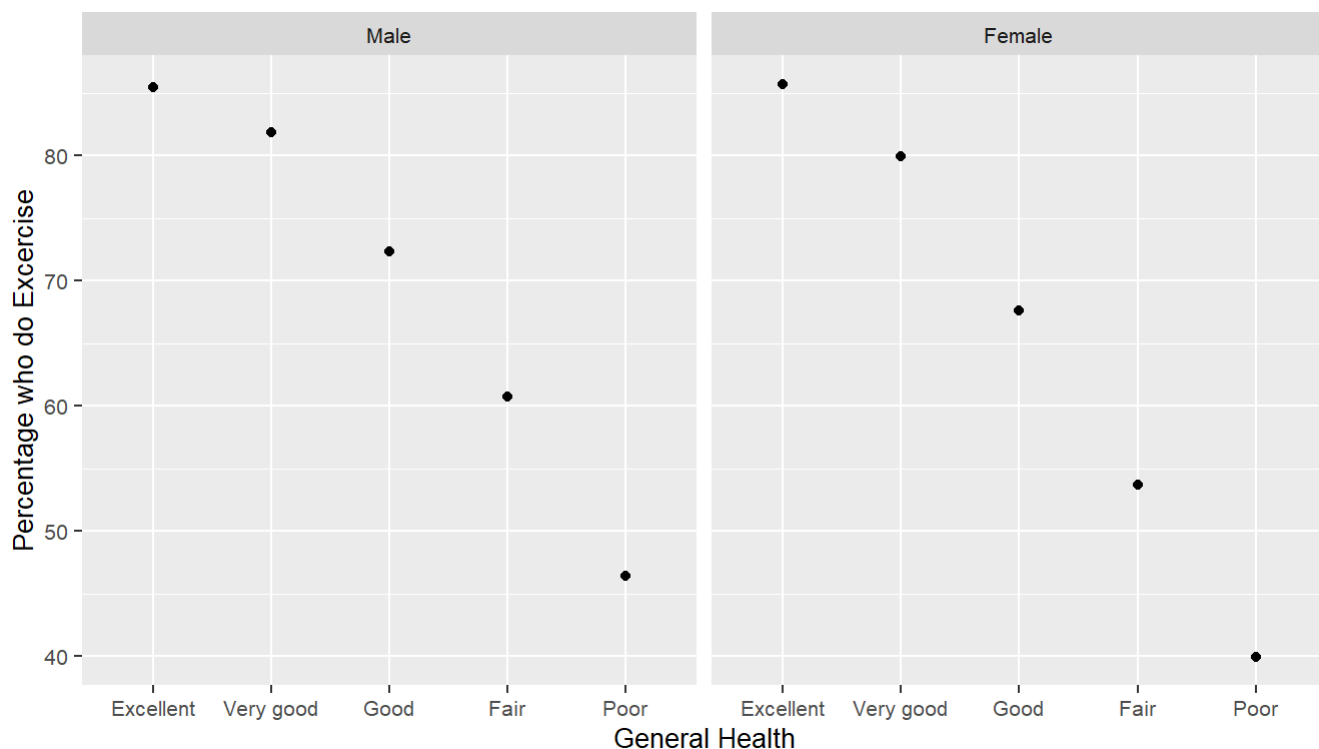
Data shows us, that in general, American men perceive themselves to be more mentally healthy than women, also, the data suggest a relationship between the marital status of Americans and their mental health. People separated tends to be more unhappy.

Research question 3:

```
answer_3 <- brfss2013 %>% select(sex, genhlth, exerany2) %>%
  filter(!is.na(genhlth), !is.na(exerany2), !is.na(sex)) %>%
  group_by(sex, genhlth) %>%
  summarise(amount_exer = 100*sum(exerany2 == "Yes") / n())
```

`summarise()` has grouped output by 'sex'. You can override using the `.groups` argument.

```
ggplot(answer_3, aes(genhlth, amount_exer)) +
  geom_point(aes(genhlth, amount_exer)) +
  labs(x="General Health", y="Percentage who do Exercise")+
  theme(text = element_text(size=10))+
  facet_wrap(. ~ sex)
```



We see that the previous pattern is repeated, American men perceive themselves to be healthier than women. Besides, it is clear that people who exercise perceive themselves healthier than people who don't. It would be interesting to test if men are indeed healthier than women, that may be my next project in this specialization.