Exploring the BRFSS data

Setup

Load packages

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

Load data

```
load("brfss2013.Rdata")
```

Part 1: Data

The Behavioral Risk Factor Surveillance System (BRFSS) is a system that measures behavioral risk factors by collecting data among all the states in USA. The system conducts health-related telephone surveys that collect state data about U.S. residents regarding their health-related risk behaviors, chronic health conditions, and use of preventive services. BRFSS tested by randomly selecting non-institutionalized adults, which qualifies the output a random sample that can be generalized to the adults of 18 years or older residing in US. BRFSS is an observational study and not an experimental study that establishes correlation between variables hence no causalities between the variables occured.

Part 2: Research questions

Research quesion 1:

It is said that the employment status and physical activities affect the amount of sleep. By observing the given data, how much do you agree that the given statement is true?

Research quesion 2:

People who maintain good health and weight are said to be less prone to certain diseases.Given the data, what do you conclude about the correlation of arthritis with both weight and health?

Research quesion 3:

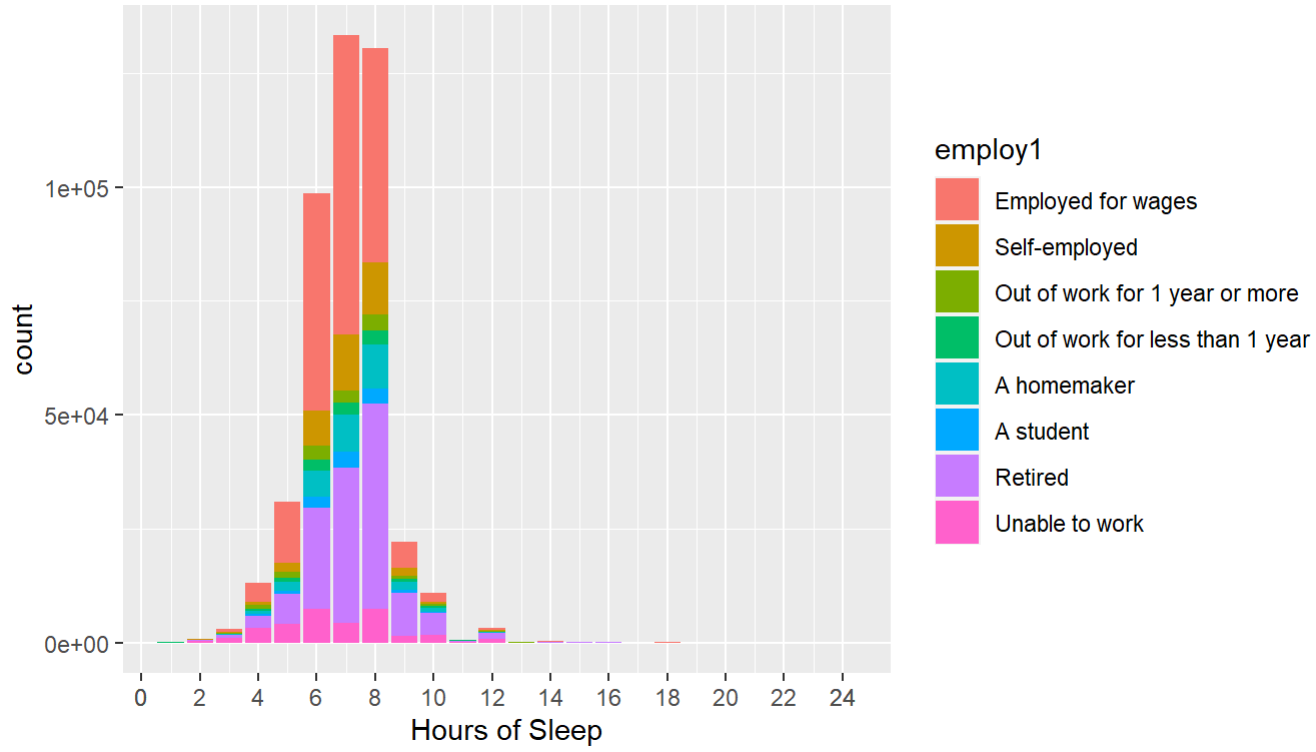
A person with more education is expected to be employed for high income. However, do you think that the gender of a person is correlated with these variables?

Part 3: Exploratory data analysis

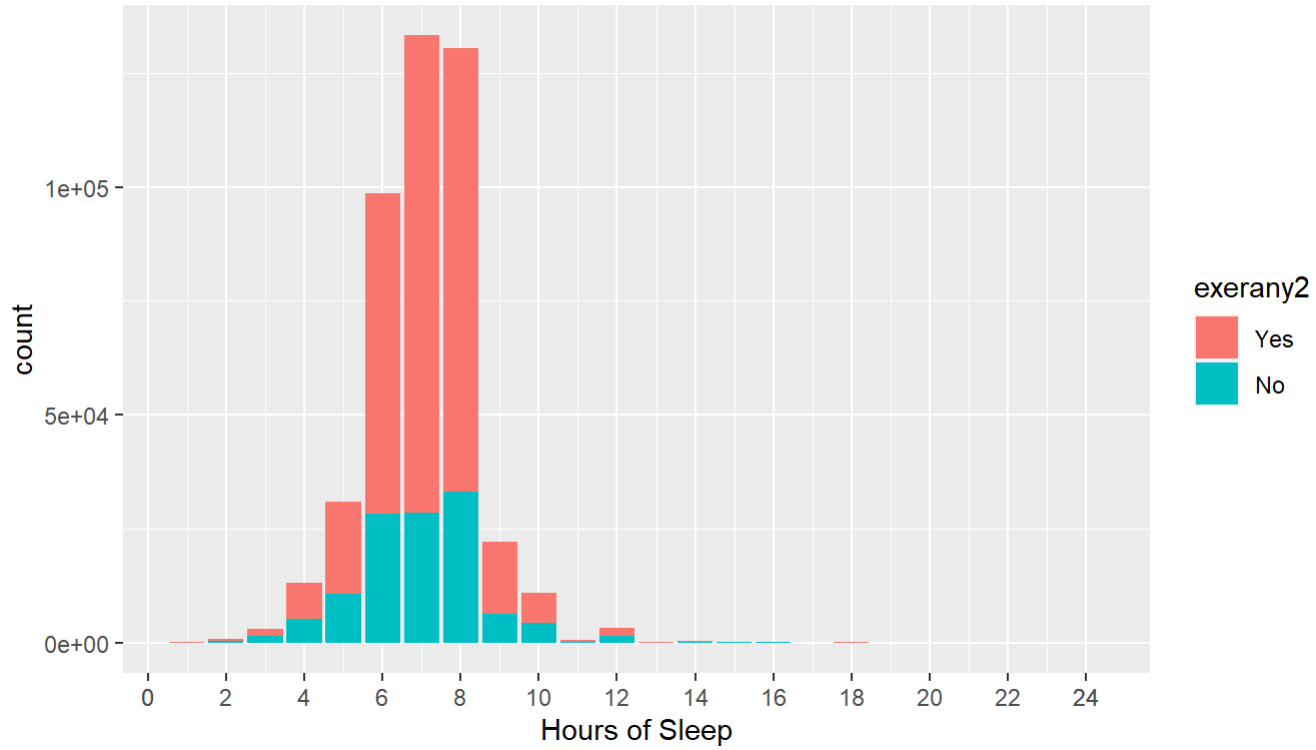
Research quesion 1:

```
Isleep<-brfss2013%>%
  filter(!is.na(sleptim1),!is.na(exerany2),!is.na(employ1))%>%
  select(sleptim1,exerany2,employ1)


ggplot(data = Isleep,aes(x=sleptim1,fill=employ1))+
  geom_bar()+scale_x_continuous(breaks =c(0,24, seq(0,24,2)))+
  xlab("Hours of Sleep")
```




```
ggplot(data = Isleep,aes(x=sleptim1,fill=exerany2))+
  geom_bar()+scale_x_continuous(breaks =c(0,24, seq(0,24,2)))+
  xlab("Hours of Sleep")
```




```
mean(Isleep$sleptim1)
```

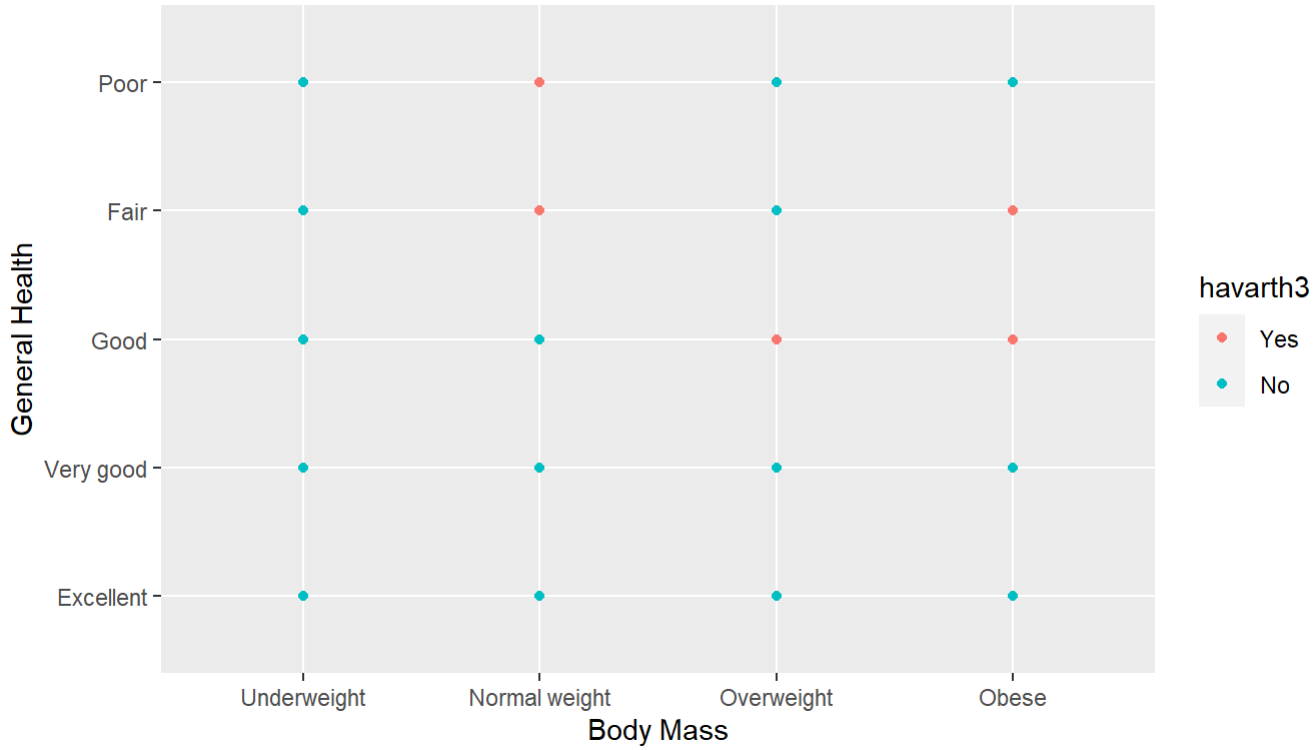
```
## [1] 7.049779
```

By looking at both the graphs, it seems that there is no significant difference in the amount of sleep a person gets regardless of involving in the physical activities and the employment status. according to the graphs, the average amount of sleep a person gets is 7-8 hours.

Research quesion 2:

```
Arthritis<-brfss2013%>%
  filter(!is.na(X_bmi5cat),!is.na(genhlth),!is.na(havarth3))%>%
  select(X_bmi5cat,genhlth,havarth3)

ggplot(data=Arthritis,aes(x=X_bmi5cat,y=genhlth,color=havarth3))+
  geom_point()+scale_fill_manual(values =c("red","blue"))+
xlab("Body Mass")+ylab("General Health")
```



According to the left half of the graph, we can see that, the quality of health and being underweight do not correlate with having an arthritis. However, a normal weight individual with fair and poor health quality correlate with arthritis.

From these observations, it seems that arthritis is an issue when the quality of health is compromised along with the increase in the body mass.

The other half of the graph shows that an overweight individual with good health and an obese individual with good and fair health, both are affected with arthritis. However, the graph shows that an obese individual with poor health is not affected by arthritis. This observation seems to contradict with the statement concluded for the first half of the graph where it was observed that compromise on health with an increase in body mass leads to arthritis. The reason for this contradiction could be due to the biasness in collecting the data.

Research quesion 3:

```
Income_edu<-brfss2013%>%
  filter(!is.na(educa),!is.na(sex),!is.na(income2),!is.na(employ1))%>%
  select(educa,sex,employ1,income2)

ggplot(data=Income_edu,aes(x=educa,y=income2,color=sex))+
  geom_point()+scale_fill_manual(values =c("red","blue"))+ theme(axis.text.x = element_text(size
  = 5,angle = 90, vjust = 0.5))+
  ylab("Income")+xlab("Education Level")
```

By looking at the trend in the graph, males are more likely to get higher income regardless of their education level. Whereas, females need to get more education in order to get a higher income.Therefore, according to the graph,gender does play a role in getting a better earning regardless of you eduction level.