Exploring the BRFSS data

Load packages

```
library(ggplot2)
library(dplyr)
library(statsr)
```
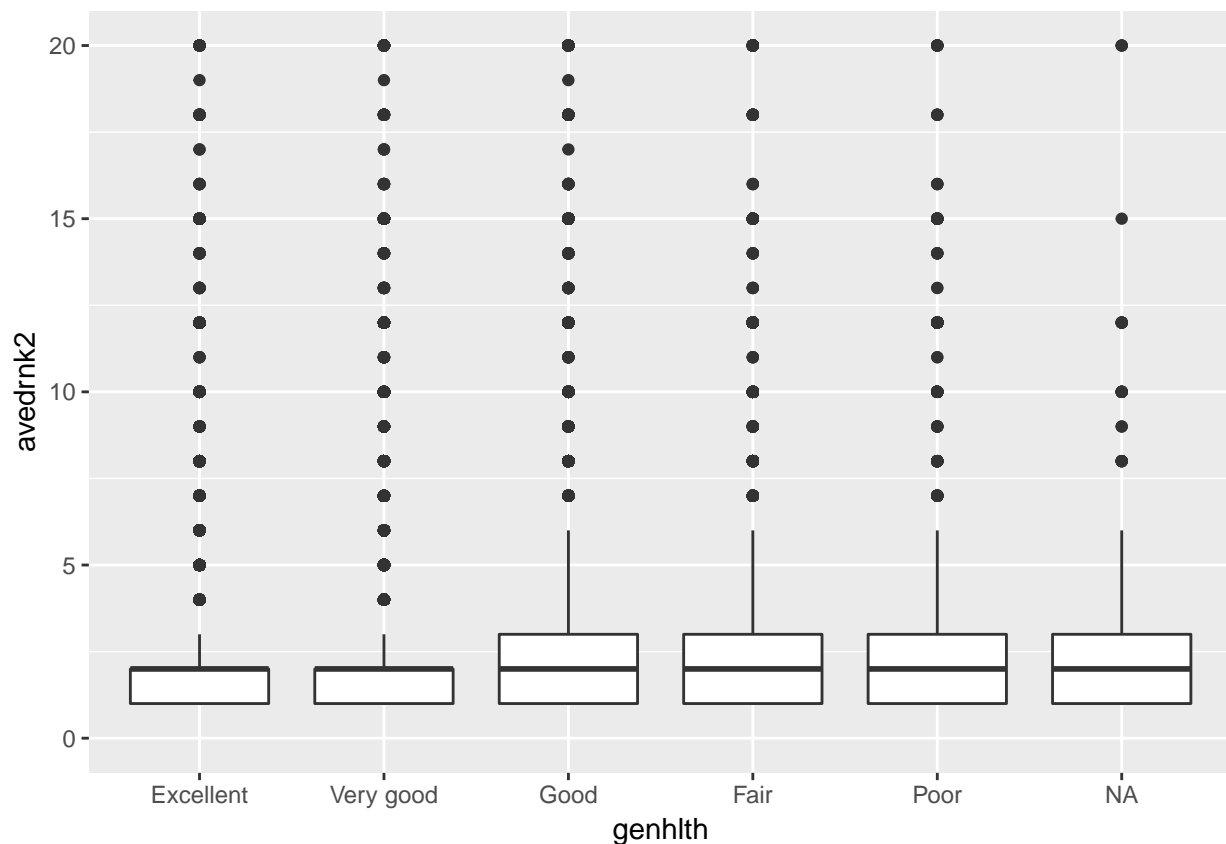
```
load("brfss2013.RData")
```

The data is collected through random sampling from a large sample size. This prevents biases from sneaking into the system. The conclusion can be to some extent be generalized toa larger population,but restraints should be followed as no causal relationship can be justified as the data is purely observational with no random assignment. "'

Research quesion 1:

Explore relation between drinking and the percieved level of general health in the soceity.This question is important as alcohol is considered one of the cause of multiple chronic lifestyle diseases
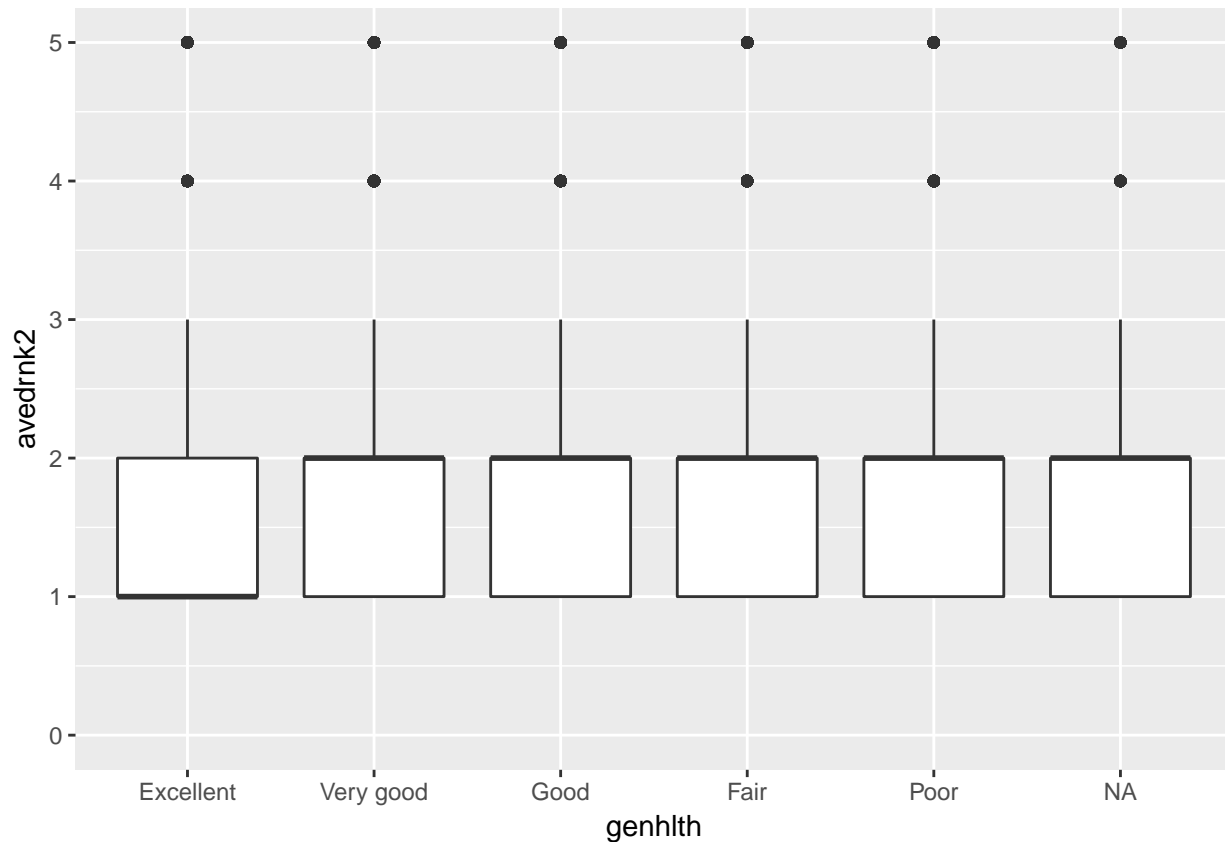
```
bp <- ggplot(brfss2013, aes(x=genhlth,y=avedrnk2)) +geom_boxplot()
bp +ylim(0,20)
```

```
## Warning: Removed 260978 rows containing non-finite values (stat_boxplot).
```

```
bp +ylim(0,5)
```

```
## Warning: Removed 273201 rows containing non-finite values (stat_boxplot).
```
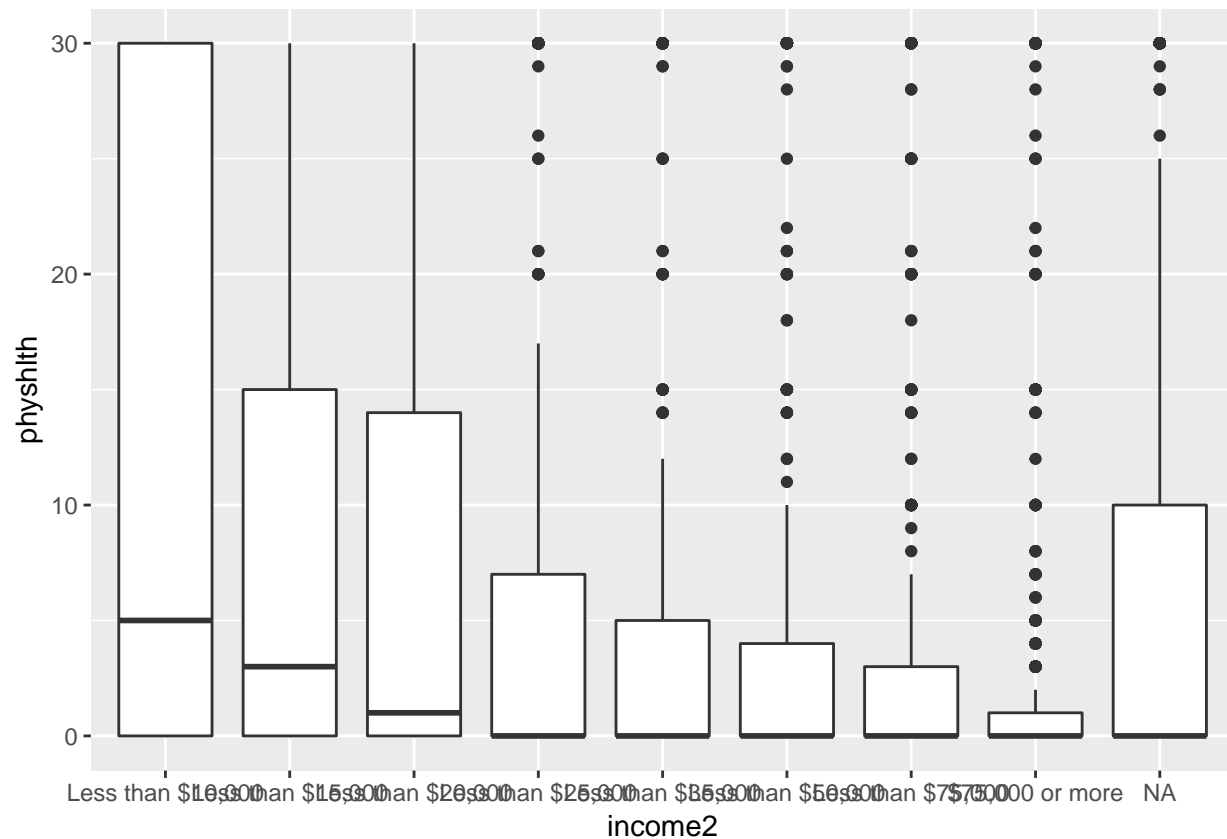

Contrary to the general perception, it was found out that people's opinion of their health is usually independent of the alcohol they consume. The median of average alcohol drink seems constant for different perception of helath. The only irregularity is that people who identify as being in excellent health tend to consume lesser value of alcohol.

Research quesion 2: To explore relationshp between income level and perception of physical health in Alabama. Income level can have great affect on your diet,exercise pattern etc. It is interesting to explore if data corresponds with this idea or not. "'

```
alabama<- brfss2013 %>% filter(X_state=="Alabama")
sp <- ggplot(alabama, aes(x=income2,y=physhlth)) +geom_boxplot()
sp +ylim(0,30)
```

```
## Warning: Removed 207 rows containing non-finite values (stat_boxplot).
```
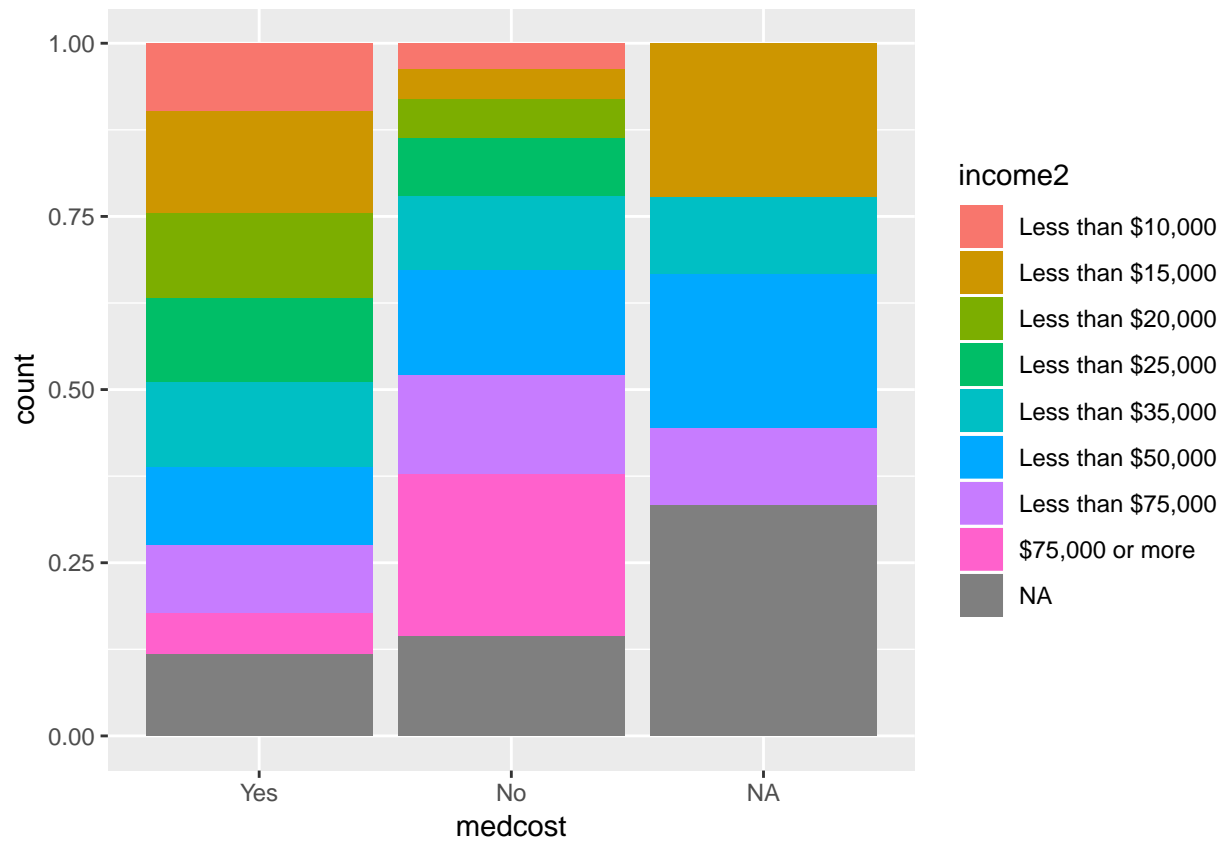
This question explores whether there is a possible correlation between income and health as rich could afford better facilites and diet. But the result show that people with lesser income tend to feel more healthy physically on the basis of higher median and higher IQR. It could be due to the fact their job are more physically active and thus help in maintaining their health

" '

Research quesion 3:

To explore relationship between the ability to afford medical care and the income level among the white population.This is critical with rising medical costs and the fear that a lot of americans are being left out.

```
White <- brfss2013 %>% filter( rrclass2 =="White")
 ggplot(data = White) +
 geom_bar(aes(x = medcost, fill = income2), position = "fill")
```

There is a weak correlation between the ability of a person and his income. Variation is seen significantly in two groups :- White with income more than $75000 and White with income less than $20000(0-20000). In the first group has only few people had the problem of affordability. But inthe other group, it was seen the number of who could afford were quite more than people who can. This seen the variation of the "fill" between the two bars. Nothing can be said about the NA group. This helps you to identify people to whom medical benefits like medicare should be extend at to.