

Parental HIV

Raelene Ramirez

2022-09-21

Introduction

For this data set we will be looking at variables such as Gender, age started smoking, and Siblings. I will be question if there will be a difference between male and female children who's mother carries HIV and if they are more likely to have a working mother. I believe that there will be no big difference between the children's genders, because there is a 50-50 shot of a mother working. I think that Not very many of the mother will be low because most women still work until older age.

```
parHIV <- read.delim("/Users/raeleneramirez/Desktop/Math130/Data/PARHIV_081217.txt", header=TRUE, stringsAsFactors=FALSE)
dim(parHIV)
```

```
## [1] 252 123
```

```
library(ggplot2)
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
```

Univariate Description

```
kip<- select(parHIV, AGEALC, GENDER, JOBMO)
```

Including Plots

For this section we are going to be looking at our three different variables and creating two different forms of data. Before starting the data I needed to look t see if my data was going to be continuous or categorical.

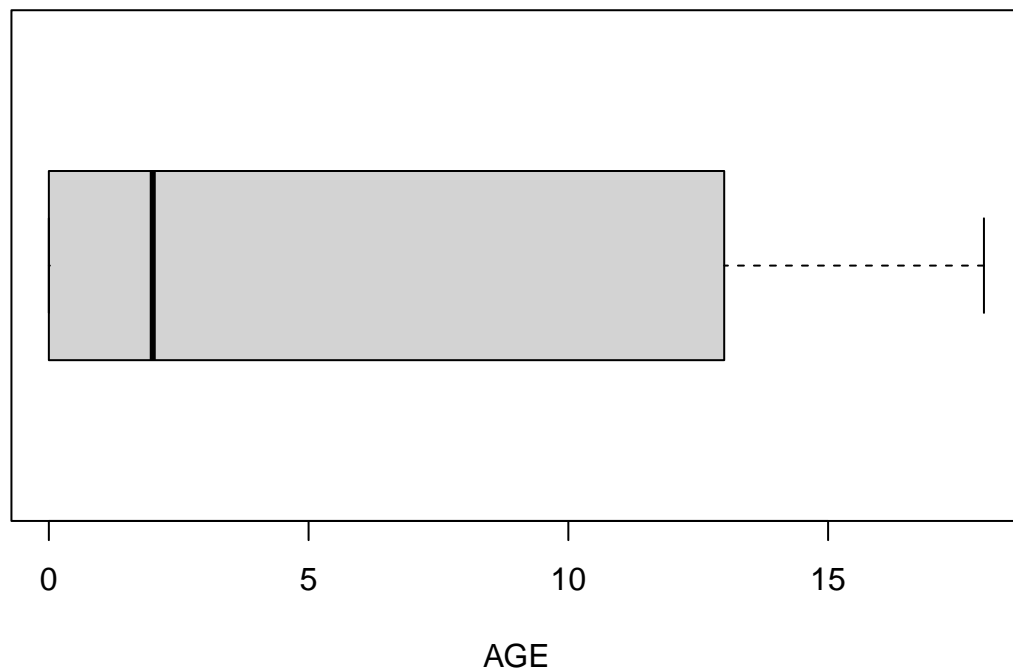
```
summary(kip$AGEALC)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   NA's  
##  0.000  0.000   2.000   6.386 13.000  18.000     1
```

Below you can find a box plot for the observed children. This allows us to see the range of the age of when they first started to consume alcohol.

```
boxplot(kip$AGEALC, horizontal = TRUE, main="Age of kids when they had first drink", xlab="AGE")
```

Age of kids when they had first drink



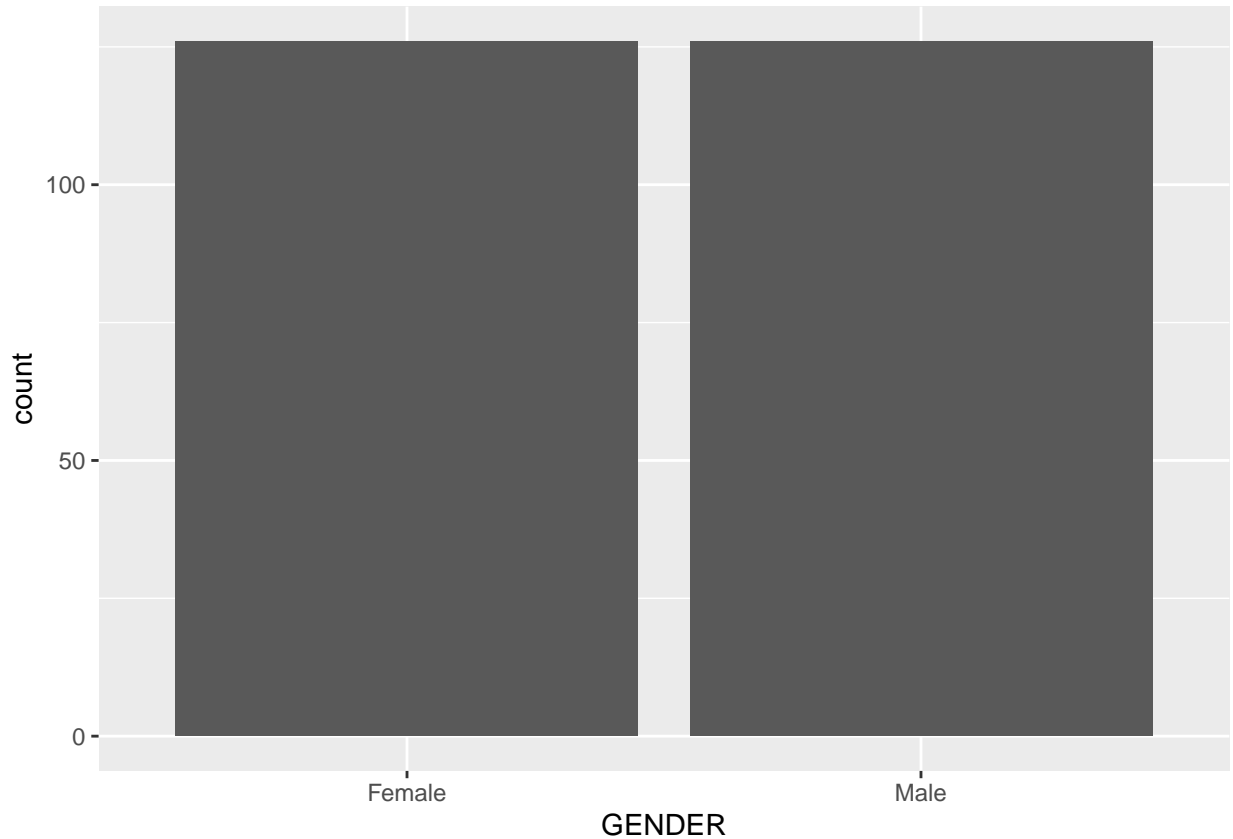
Next, we will be looking at gender using table, this will allow us to categorize the two separate genders.

```
table(kip$GENDER)
```

```
##  
## Female  Male  
##   126   126
```

With this data we will create a Bar graph that will help us compare the two gender for the children.

```
ggplot(kip, aes(x=GENDER)) + geom_bar()
```



Lastly we will repeat this step from before and create a table which will allow us to see how many of the childrens mother has a working status.

```
table(kip$AGEALC, kip$JOBMO)
```

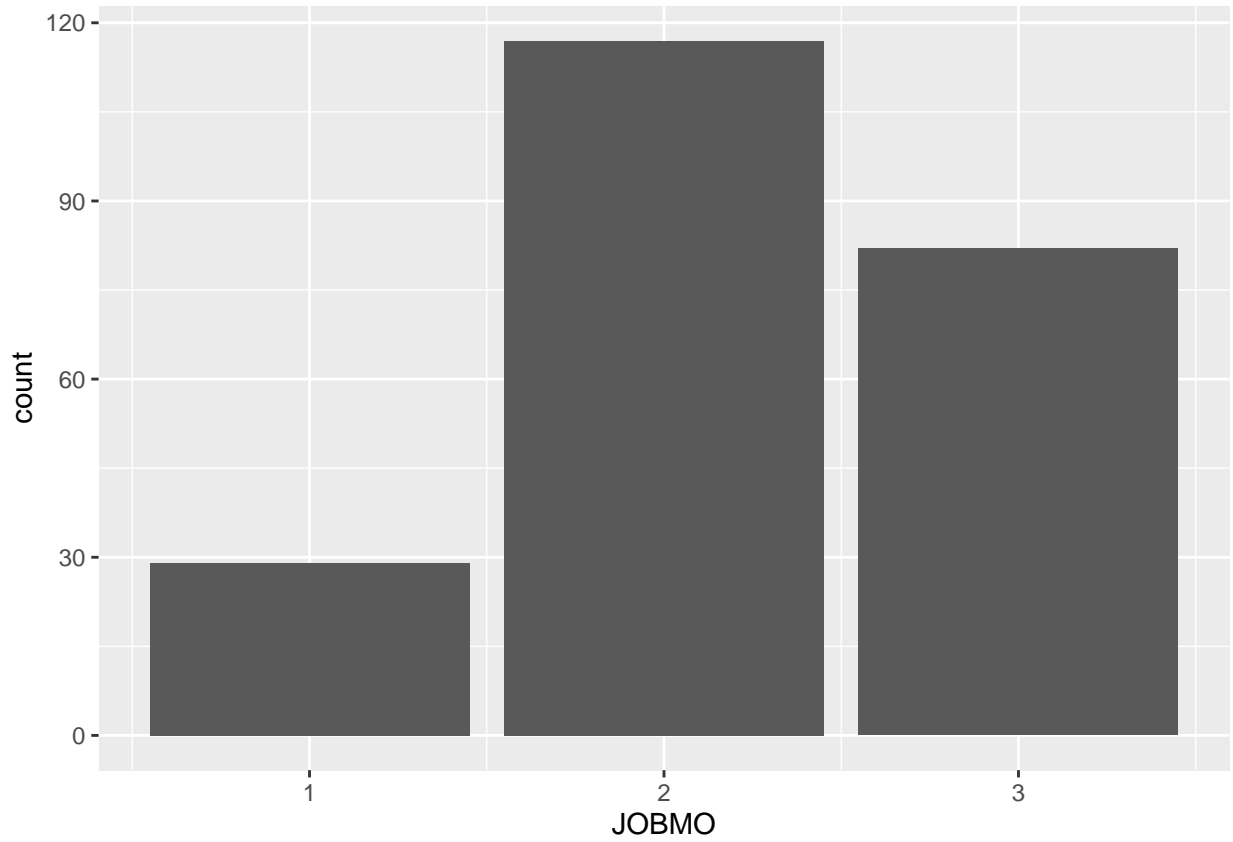
```
##
##      1  2  3
##  0 16 65 35
##  2   0  1  0
##  5   1  3  0
##  6   0  0  2
##  7   0  0  1
##  8   1  0  0
##  9   0  3  0
## 10   0  5  0
## 11   0  3  2
## 12   3  5  3
## 13   1 11 14
## 14   0 11 12
## 15   4  5  7
## 16   1  4  4
## 17   2  0  1
## 18   0  0  1
```

With this table we are able to see at what age these kids will be having their first drink. 0= that they have yet to strat drinking.

```
ggplot(kip, aes(x=JOBMO, fill=AGEALC)) + geom_bar(position = "dodge") + geom_bar()
```

```
## Warning: Removed 24 rows containing non-finite values (stat_count).
```

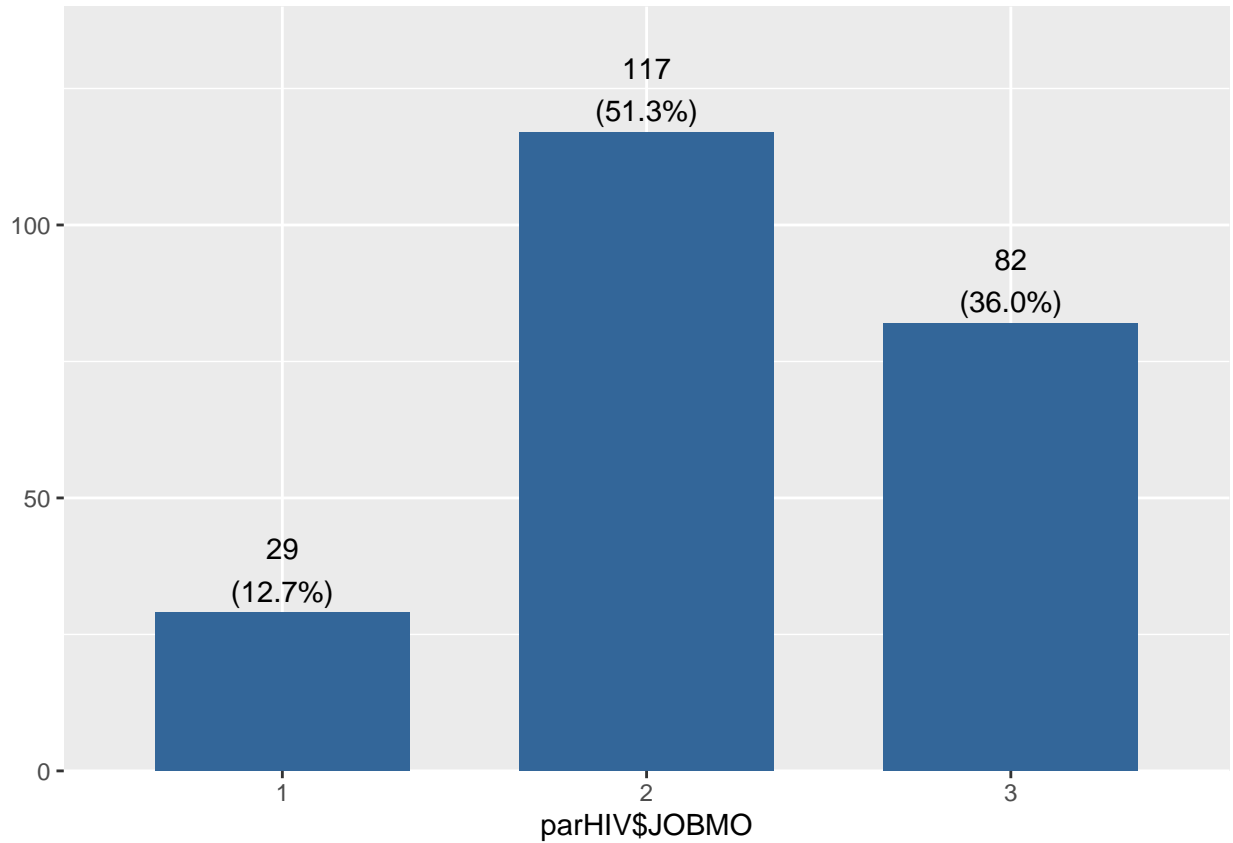
```
## Removed 24 rows containing non-finite values (stat_count).
```



```
library(sjPlot)
```

```
## Install package "strengjacke" from GitHub ('devtools::install_github("strengjacke/strengjacke")')
```

```
plot_frq(parHIV$JOBMO)
```



##Bivariate Comparison

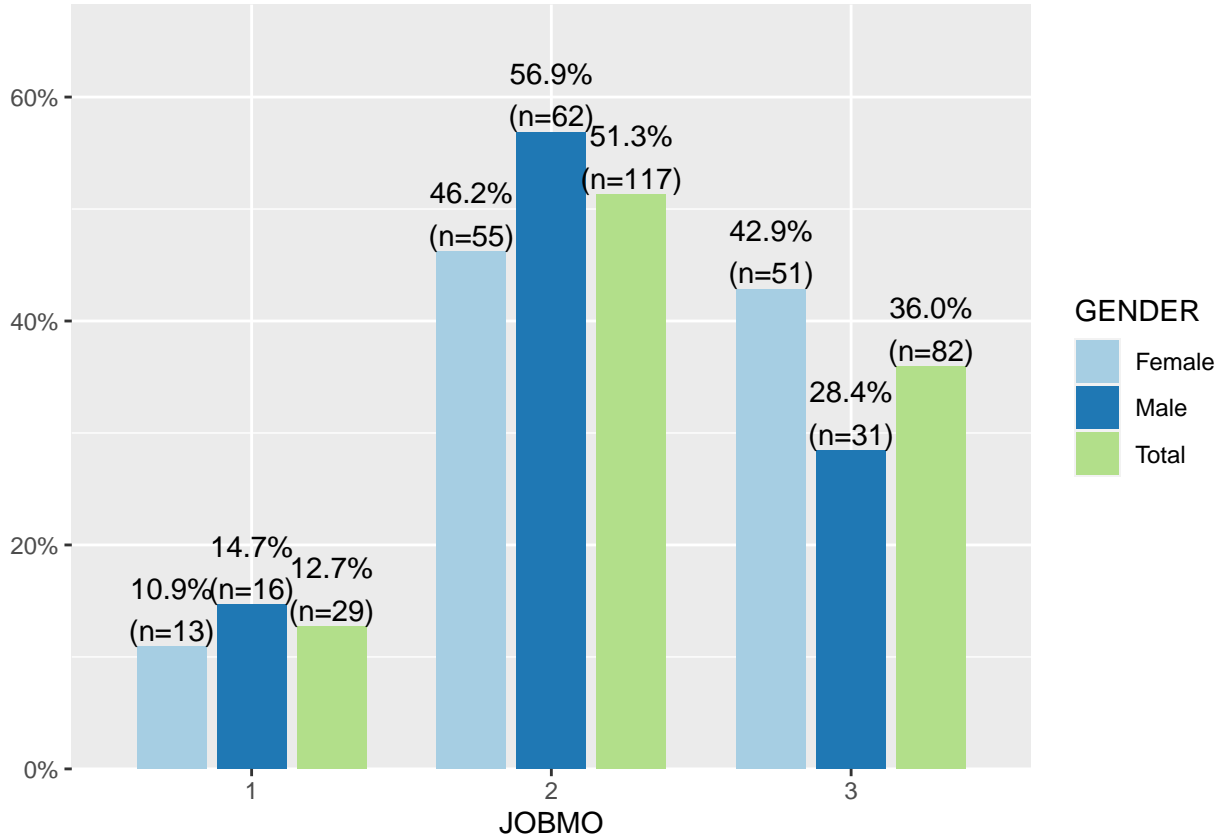
The two components I will be looking at is Gender and Mothers job status

```
table(parHIV$GENDER, parHIV$JOBMO)
```

```
##
##      1  2  3
## Female 13 55 51
## Male   16 62 31
```

We use table() to look at the two separate variables in a chart. Next we will use this same results in more of a visual by using a contingency table

```
plot_xtab(parHIV$JOBMO, parHIV$GENDER)
```



As you can see, in this grouped chart We find that group 2 has one of the highest percentages. We find that male is the leading category and we can then get the correct answer to our hypothesis.

##Conclusion

At the end, we can conclude that my hypothesis is that incorrect. Although I was to believe that there is 50-50 ratio, we find that the male gender has about a 10% difference from the female. We can also say that the most of the moms unemployment status is higher than ones who are employed. At the end of this study I would say that I am surprised by these results because I was not I was not expected retired to be a higher percentage than moms who are employed.