

# Depression In Men & Women

Katie Wilken

3/9/2022

```
depress <- read.delim("C:/Users/missk/Documents/Math130/Data/depress_081217.txt", header=TRUE, sep="\t")
dim(depress)
```

```
## [1] 294 37
```

## Introduction

The presented data set is on depression rates in men and women focused in the Los Angeles area. The information we will use in the data set is from 294 observations and 37 variables. The data consists of a series of interviews about the subjects in several categories such as their overall health, educational background, whether or not they drink regularly, and their income level.

## Univariate Exploration

To start, I want to see how many depressed men and women are participating in the data set compared to those that are not depressed.

```
depress$cases <- factor(depress$cases, labels=c("Depressed", "Normal"))
table(depress$cases)
```

```
##
## Depressed    Normal
##          244         50
```

Next, I want to see what kind of overall health status the participants have.

```
depress$health <- factor(depress$health, labels=c("Poor", "Fair", "Good", "Excellent"))
table(depress$health)
```

```
##
##      Poor      Fair      Good Excellent
##      130      115      35      14
```

Now, we will look at the educational level of the participants.

```
depress$educat <- factor(depress$educat, labels=c("Less than high school", "Some high school", "Finished high school"))
table(depress$educat)
```

```
##
##      Less than high school      Some high school
##              5                  43
##      Finished high school      Some college
##             114                  14
## Finished bachelor's degree  Finished master's degree
##              9                  48
##      Finished doctorate
##             61
```

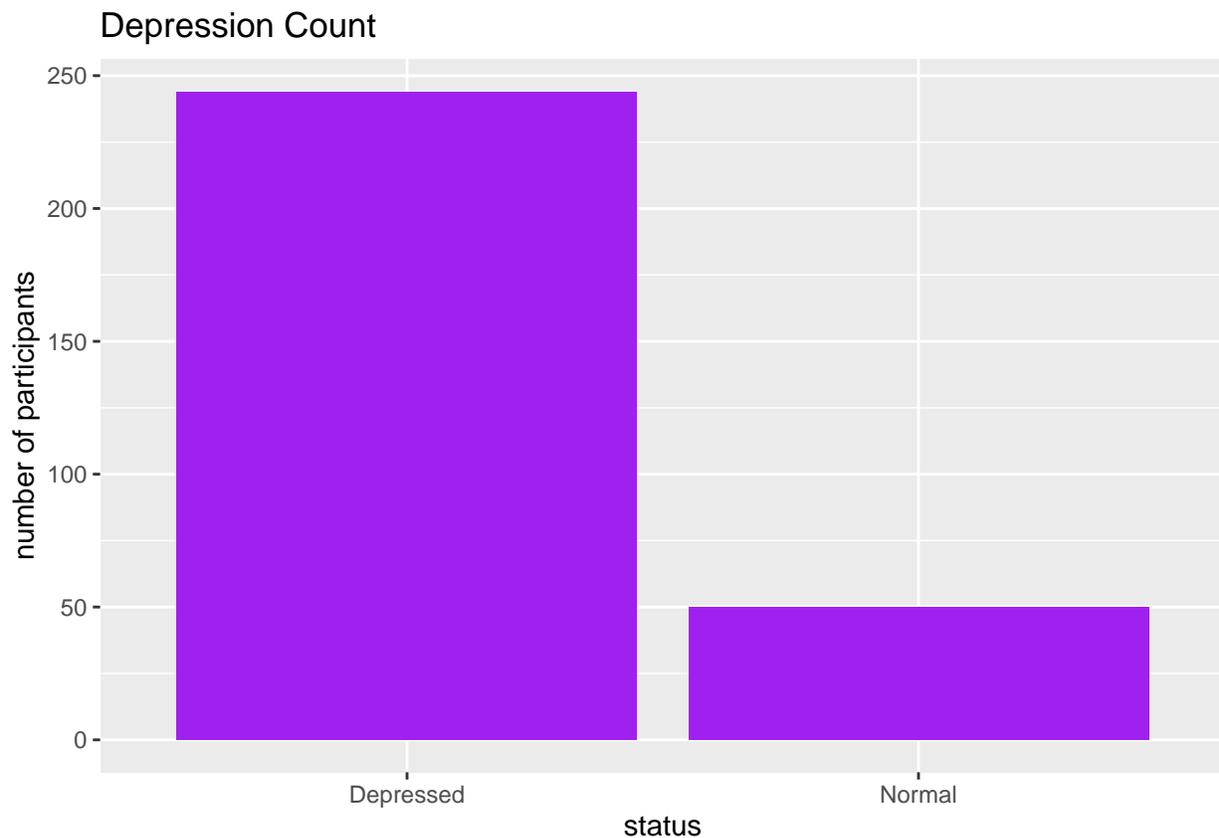
Now, I want to see what drinking habits are found within the sample population of the data set.

```
depress$drink <- factor(depress$drink, labels=c("No", "Yes"))
table(depress$drink)
```

```
##
## No Yes
## 60 234
```

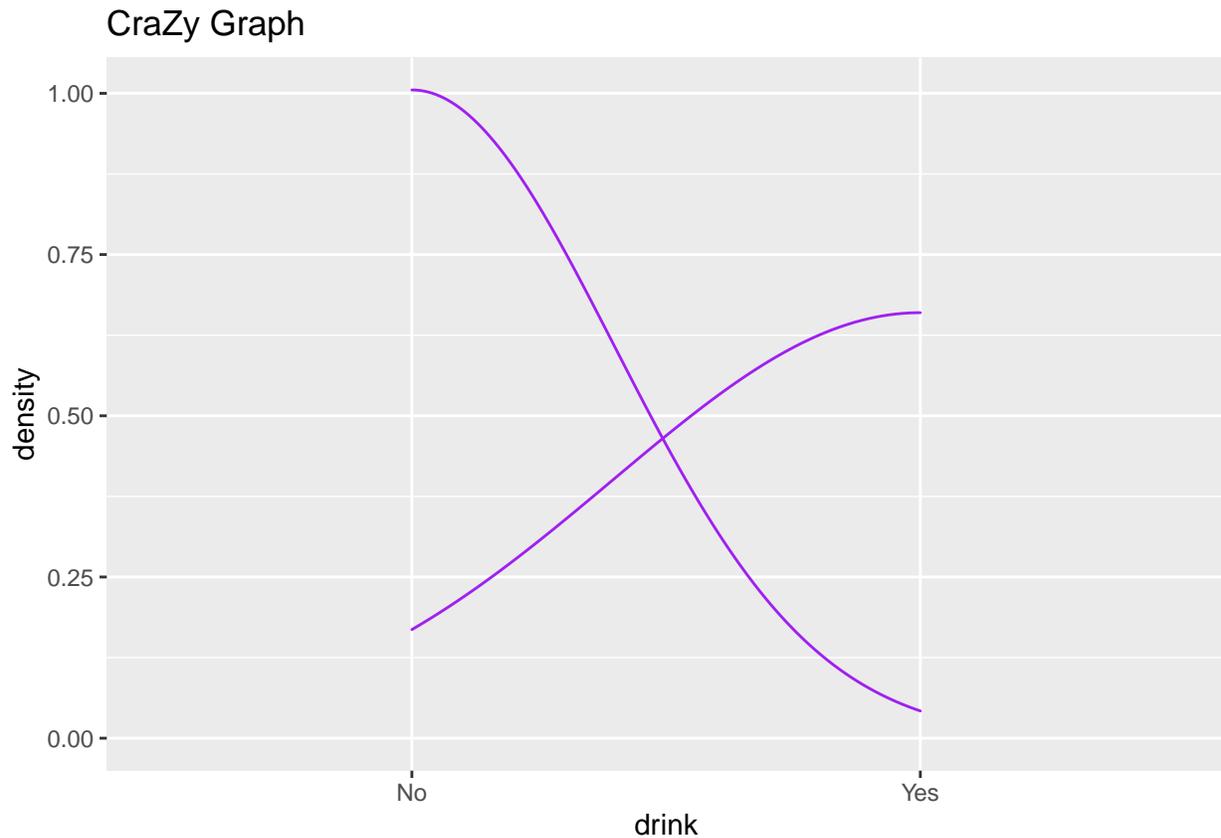
To visualize the data, we can use graphs instead of tables to show the amount of participants that are depressed compared to those that are not depressed.

```
ggplot(depress, aes(x=cases))+geom_bar(fill="purple") + ggtitle("Depression Count")+ylab("number of participants")
```



We can also view the density of our participants in one categorical space or another.

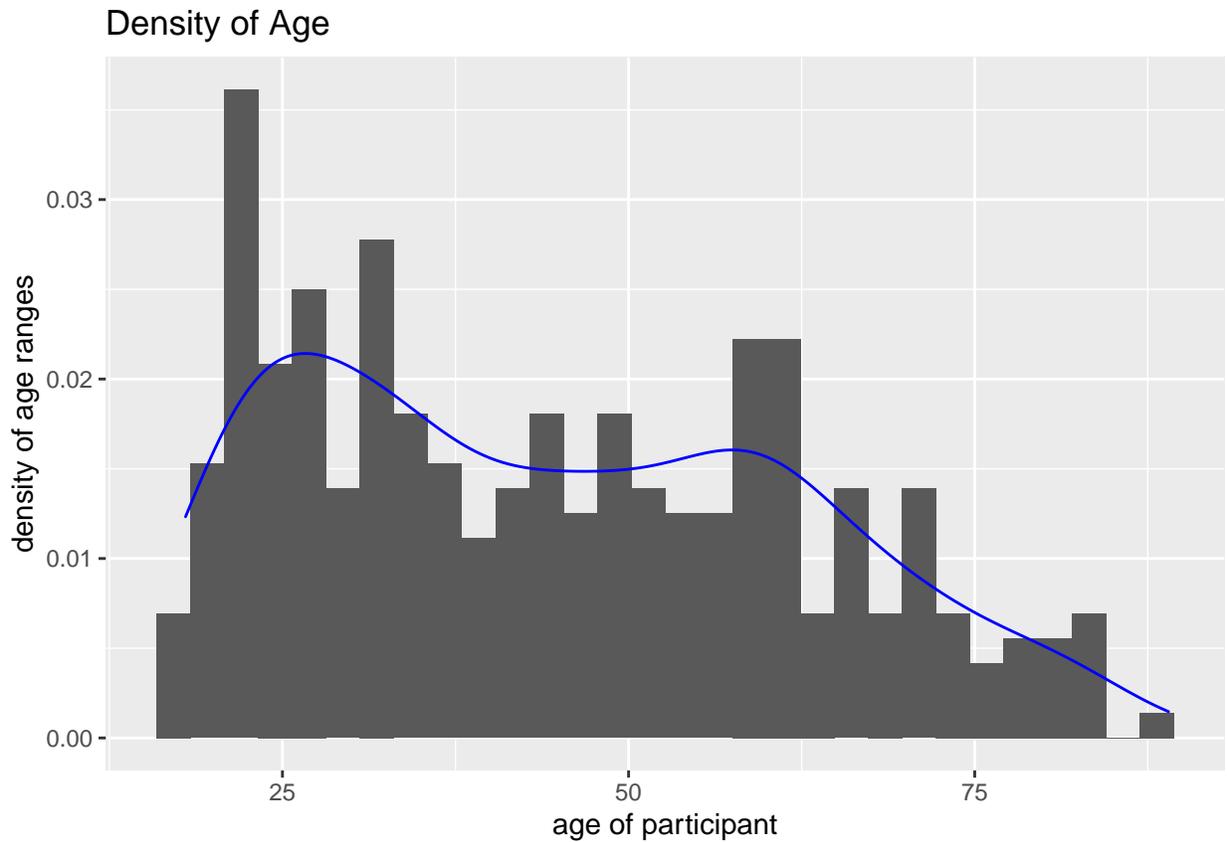
```
ggplot(depress, aes(x=drink)) + geom_density(col="purple") + ggtitle("CraZy Graph")
```



I don't even know what to make of this graph! It is super cool. I was trying to make a density curve and instead I made this, I guess x marks the spot! Now, I will try again to make a density graph but focus on the density of the ages of the participants, and I'm adding a histogram too.

```
ggplot(depress, aes(x=age))+ ggtitle("Density of Age") + geom_histogram(aes(y=..density..)) + geom_dens
```

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```



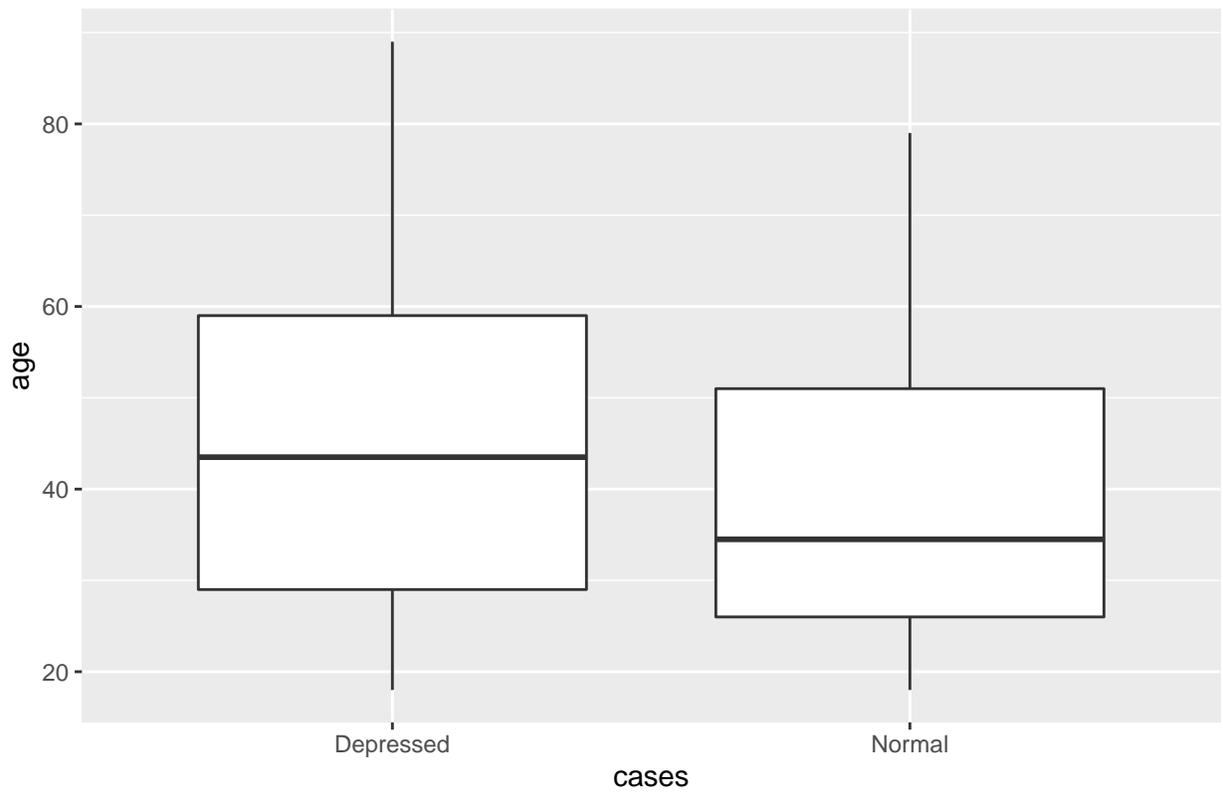
Based on the graph above, the majority of participants are in their mid-20s or their late 50s/early 60s.

### Bivariate Exploration

With a bivariate function, we can look at two variables like the participants ages and depression status.

```
ggplot(depress,aes(y=age, x=cases)) + geom_boxplot()+ggtitle("Relationship Between Age & Depression Sta
```

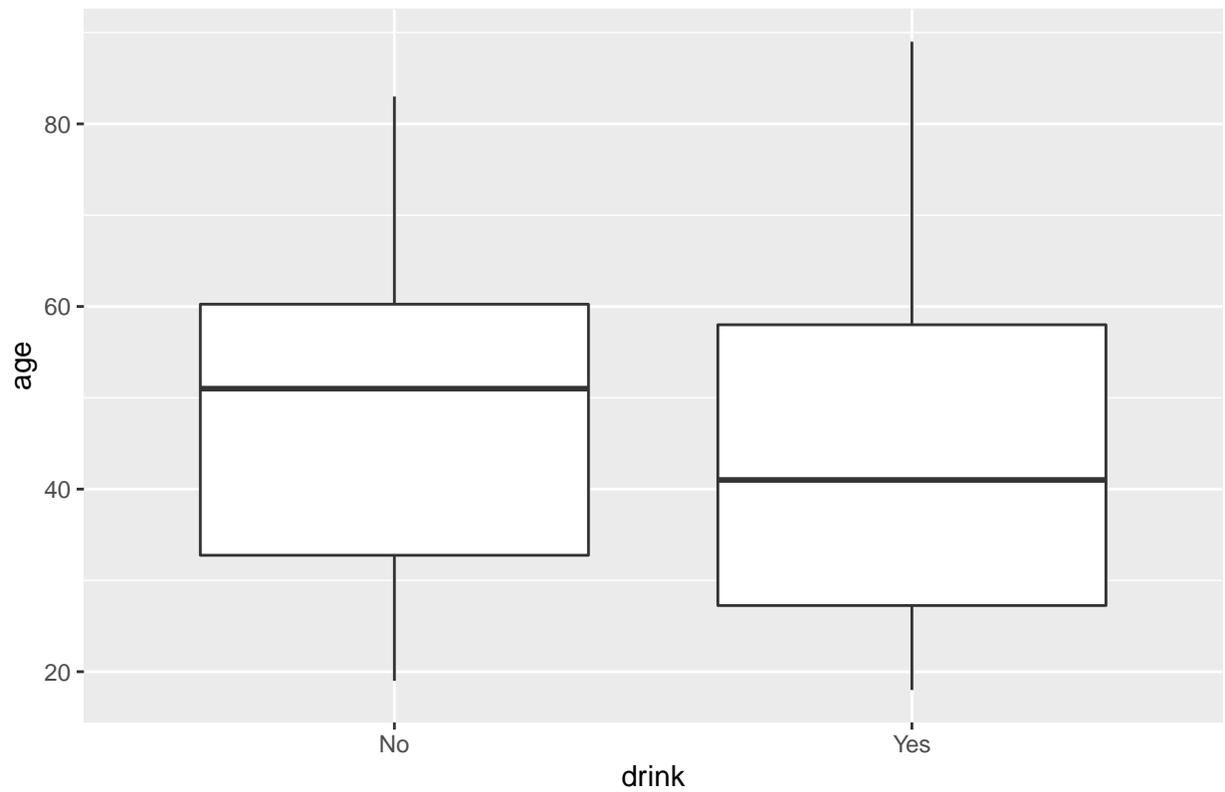
Relationship Between Age & Depression Status



Or, we could evaluate the drinking habits along with the age ranges.

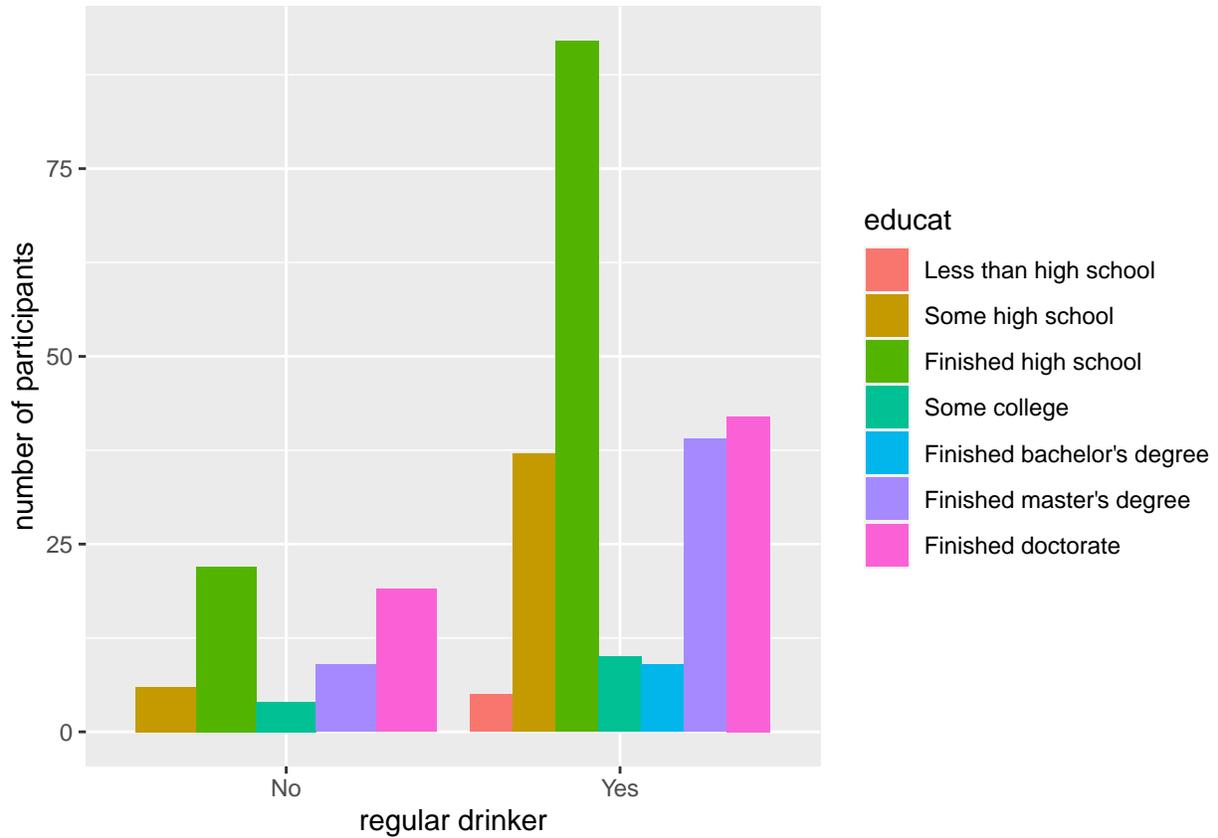
```
ggplot(depress, aes(y=age, x=drink))+ geom_boxplot() + ggtitle("Coorelations Between Age & Drinking Hab")
```

## Coorelations Between Age & Drinking Habits



Below, we can see how the education level is distributed among the participants based on their gender identities.

```
ggplot(depress, aes(x=drink, fill=educat)) + geom_bar(position = "dodge")+xlab("regular drinker")+ylab("education level")
```



From this data graphed, we can see there is a significant amount of subjects with higher educational levels that drink more regularly.

### Conclusion

The collected data shows that the sample population are in decent health overall. The population of depressed individuals is about 5 times the population of individuals that are not depressed. I found it interesting that higher education levels seemed to coorespond with higher drinking levels. The cases of depression appeared to be present most in middle age individuals, late 20s through 50s. It was interesting to me to see that the drinking levels and depression levels are similar, it must be the coping mechanism of choice for the participants in this data set.