# Data Analysis Project

Jacob Martinez

2022-09-19

```
library("ggplot2")
library("forcats")
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
```

```
library("RColorBrewer")
```
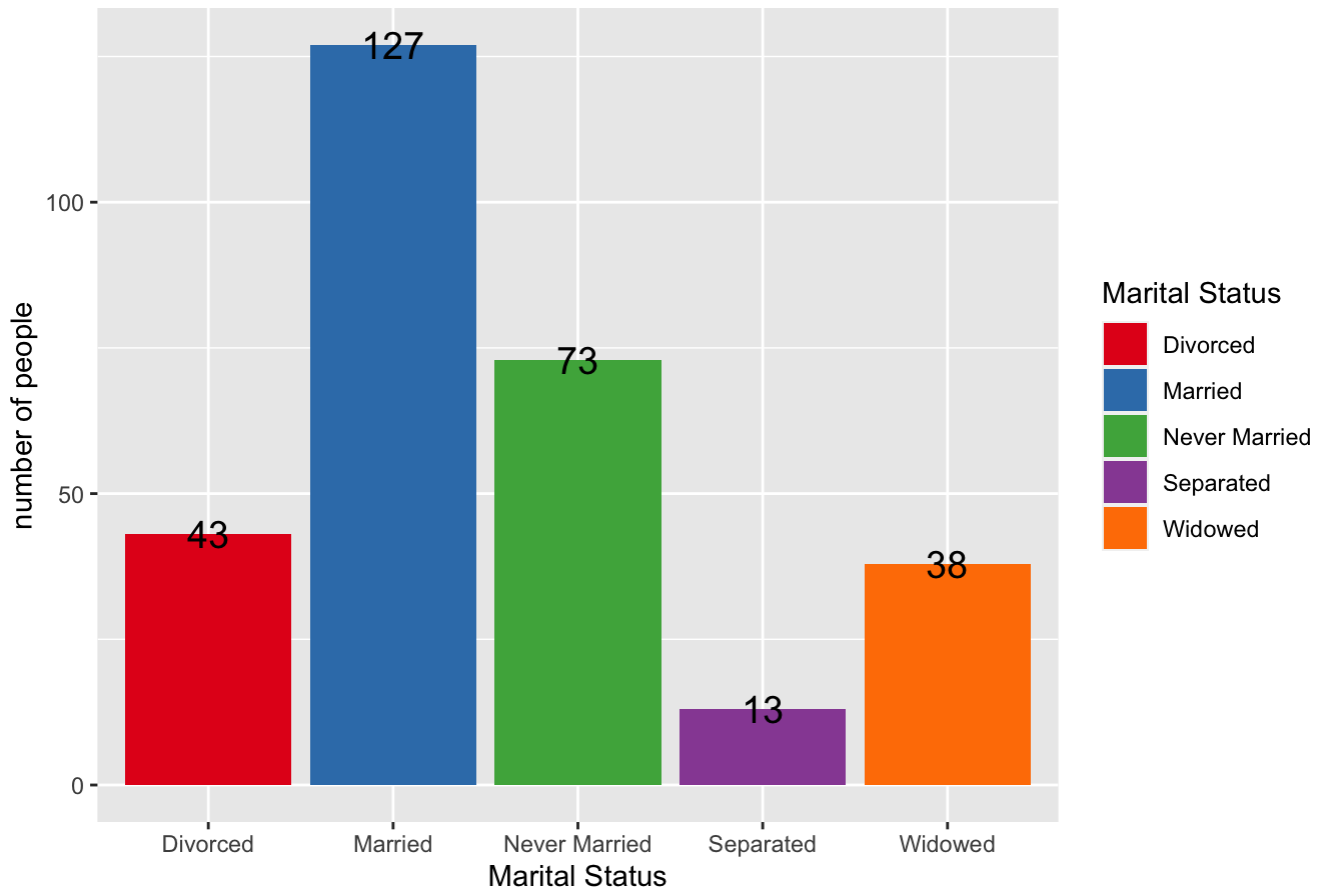
# Introduction

The data that I will be working with is depression. This data shows 294 people observed in LA county that are suffering from depression or not with around 19 variables. the variable that i have chosen are regular drinker, general health, and marital status. More specifically I will be looking at how your marital status affects you to be a regular drinker and how that affects your general health.

## Variable 1 - Marital Status

```
ggplot(depress, aes(x=marital,, fill=marital)) + geom_bar() + ylab("number of people") + xlab(
"Marital Status") + ggtitle("Number of Depressed poeple by Marital Status") + scale_fill_brewe
r(palette="Set1", name="Marital Status") + geom_text(aes(y=..count.. + 0, label=..count..), st
at='count', size = 5)
```

## Number of Depressed poeple by Marital Status



According to the graph we can see that people that are married, or never married are more depressed than people that are either divorced or widowed. And the people that are the least depressed are individuals that are separated.
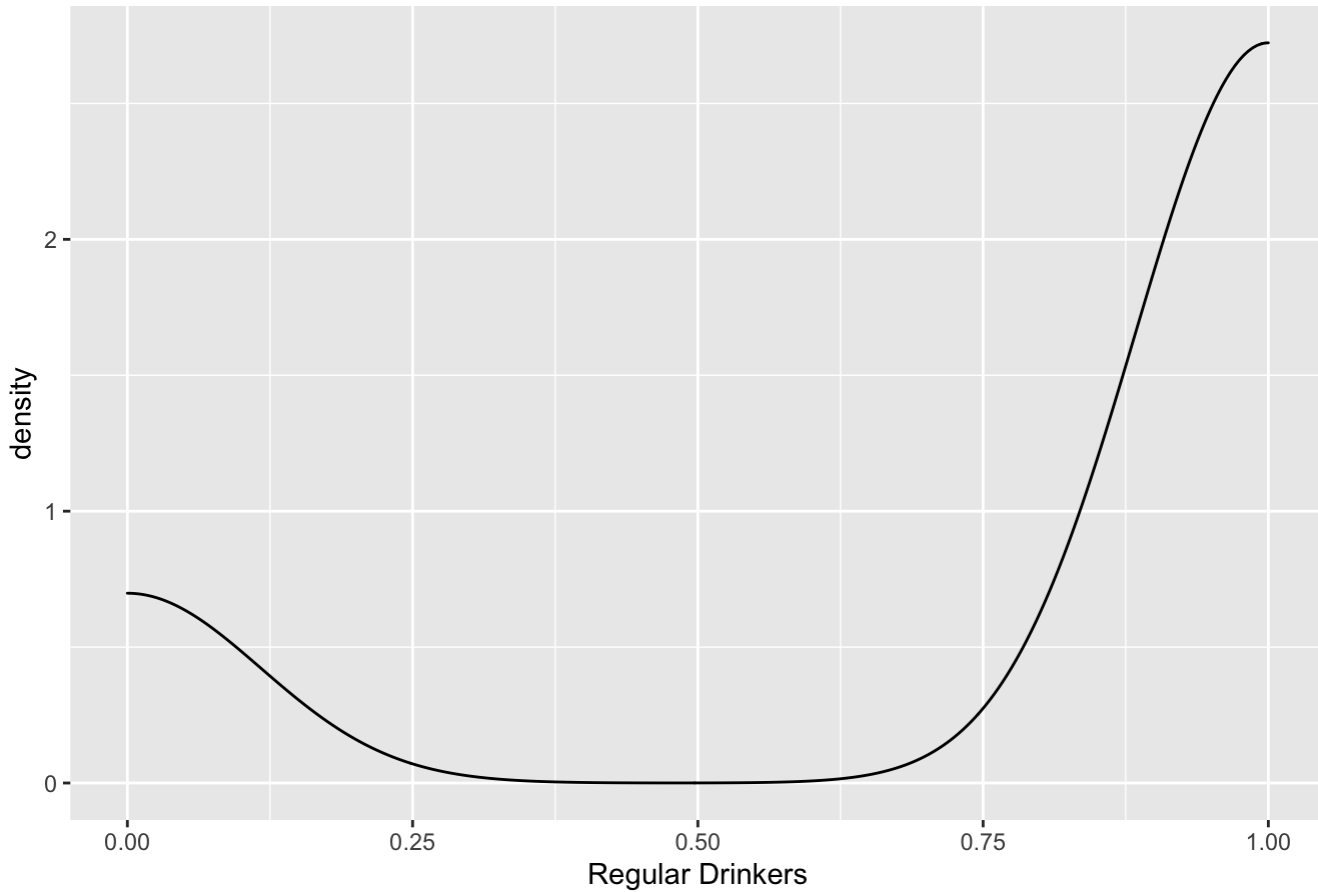
# Variable 2- Regular Drinker

```
table(depress$drink)
```

```
##
##   0   1
##  60 234
```

```
ggplot(depress, aes(x=drink)) + geom_density() + ylab("density") + xlab("Regular Drinkers") +
ggtitle("People that are Regular Drinkers Due to being Depressed")
```

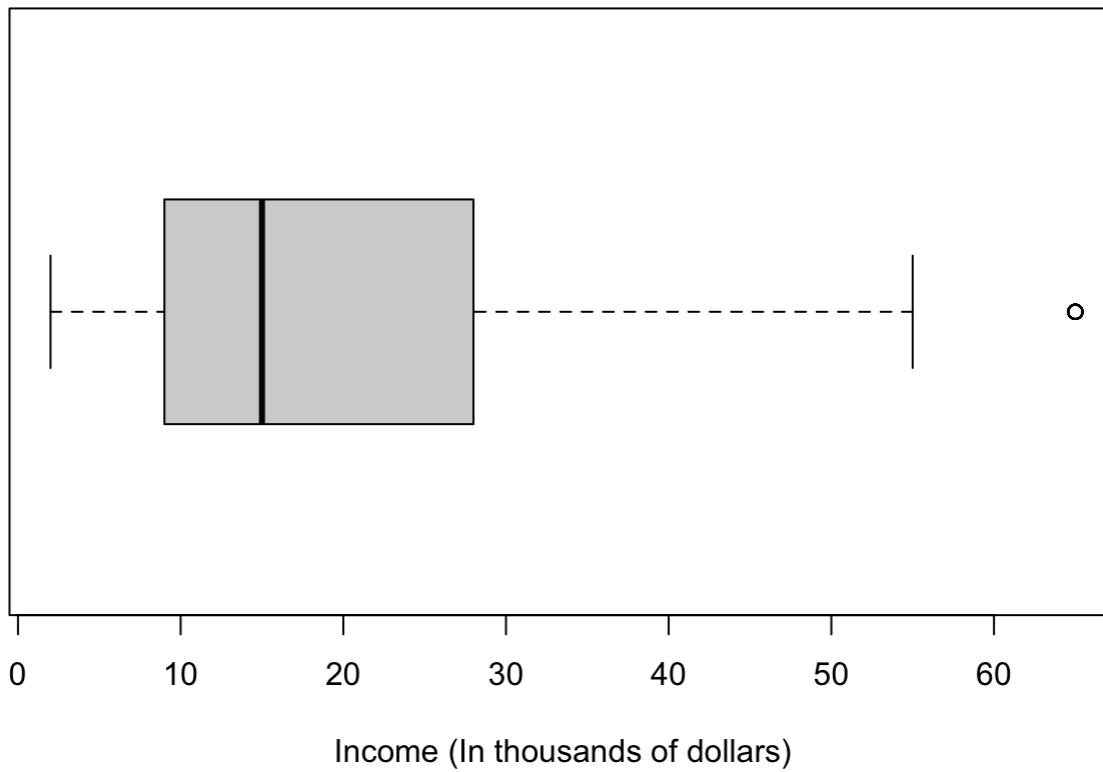## People that are Regular Drinkers Due to being Depressed



As the graph shows more people who suffer from depression drink regularly. And as the table shows there are 234 people that drink regularly and 60 people that donÕt.

## Variable 3- Income

```
boxplot(depress$income, horizontal= TRUE, main= "Distribution of Income Status",
xlab="Income (In thousands of dollars)")
```
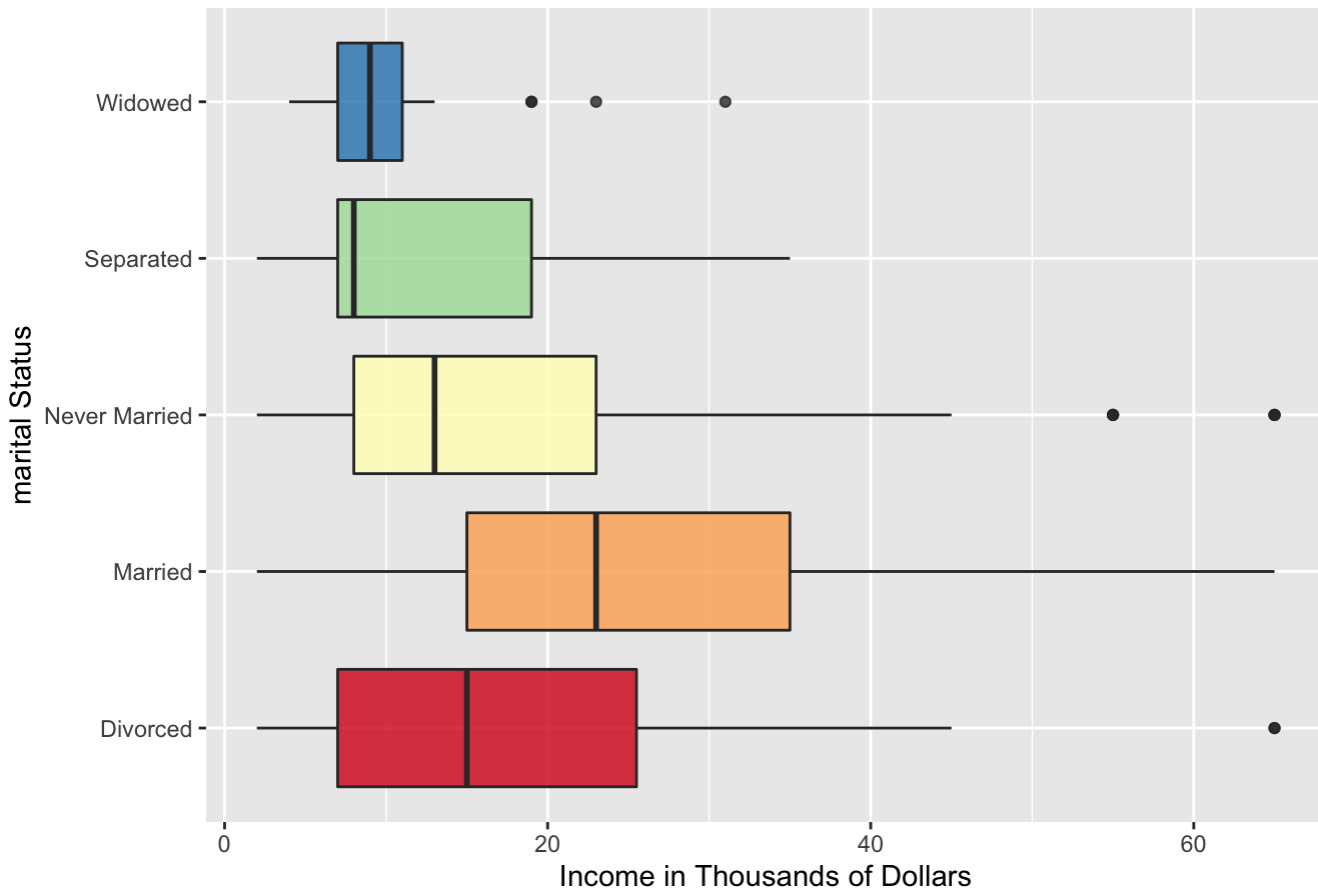
## Distribution of Income Status



Income (In thousands of dollars)

## Bivariate Exploration

```
ggplot(depress, aes(x=marital, y=income, fill=marital,)) + geom_boxplot(alpha=0.8) +
theme(legend.position="none") +
scale_fill_brewer(palette="Spectral") +ylab ("Income in Thousands of Dollars") + xlab ("marita
l Status") +ggtitle("Income vs Marital Status") +coord_flip()
```

## Income vs Marital Status



In the grouped box plot it shows the comparison between marital status and income. People that are married make more money than those that are widows. And what is surprising is that those who never married make a lot more money than those who are separated.

# Conclusion

Something that really surprised me was that there are a lot more people that are depressed that are married. I wouldÕve thought that people that are divorced or separated would have the most however that was not the case. My hypothesis was incorrect when comparing marital status to depression however it was correct when thinking about people drinking regularly if they are depressed. I think people who drink regularly when depressed think it is a way to get out of it and just forget about everything in the world at the moment however they should really go get help because there are numerous amounts of resources to sure depression.