# 130 Project

Diana Avila

2/24/2022

```
library(ggplot2)
```

```
## Warning in register(): Can't find generic `scale_type` in package ggplot2 to
## register S3 method.
```

```
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(sjPlot)
```

Introduction: The data we have chosen is depression, this is about 294 people being observed in LA county for a study of depression with about 19 variables.The variables we have chosen are employed, drink, and marital. We will test to see who is more likely to get depression. We would like to see the corolation of different variables that are common in depressive people.

```
depress <- read.delim("https://norcalbiostat.netlify.app/data/depress_081217.txt", header=TRUE,sep="\t")
dim(depress)
```

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

UNIVARIATE EXPLORATION

Variables being observed Employment Status:

```
table(depress$employ)
```

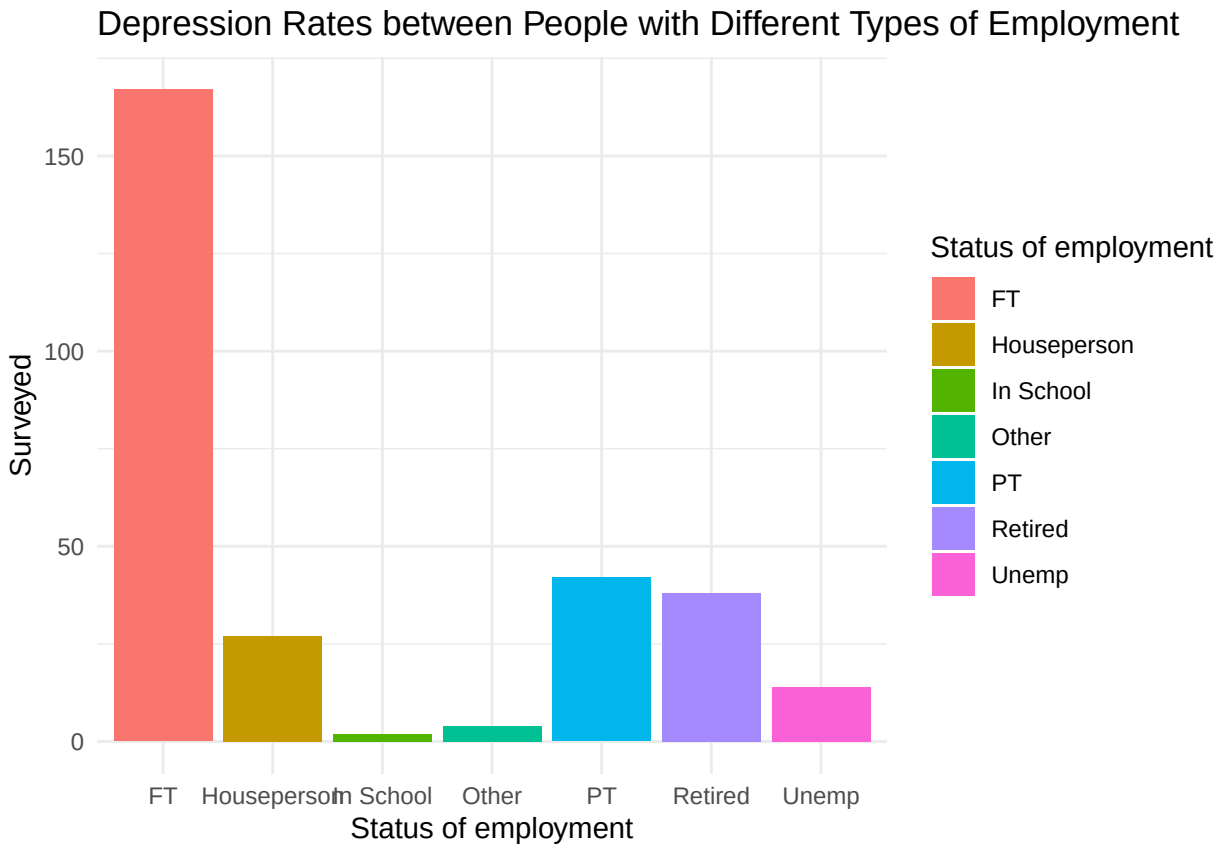
```
##
##          FT Houseperson    In School       Other          PT     Retired
##         167          27            2           4          42          38
##       Unemp
##          14
```

```
summary(depress$employrename)
```

```
## Length  Class   Mode
##      0   NULL   NULL
```

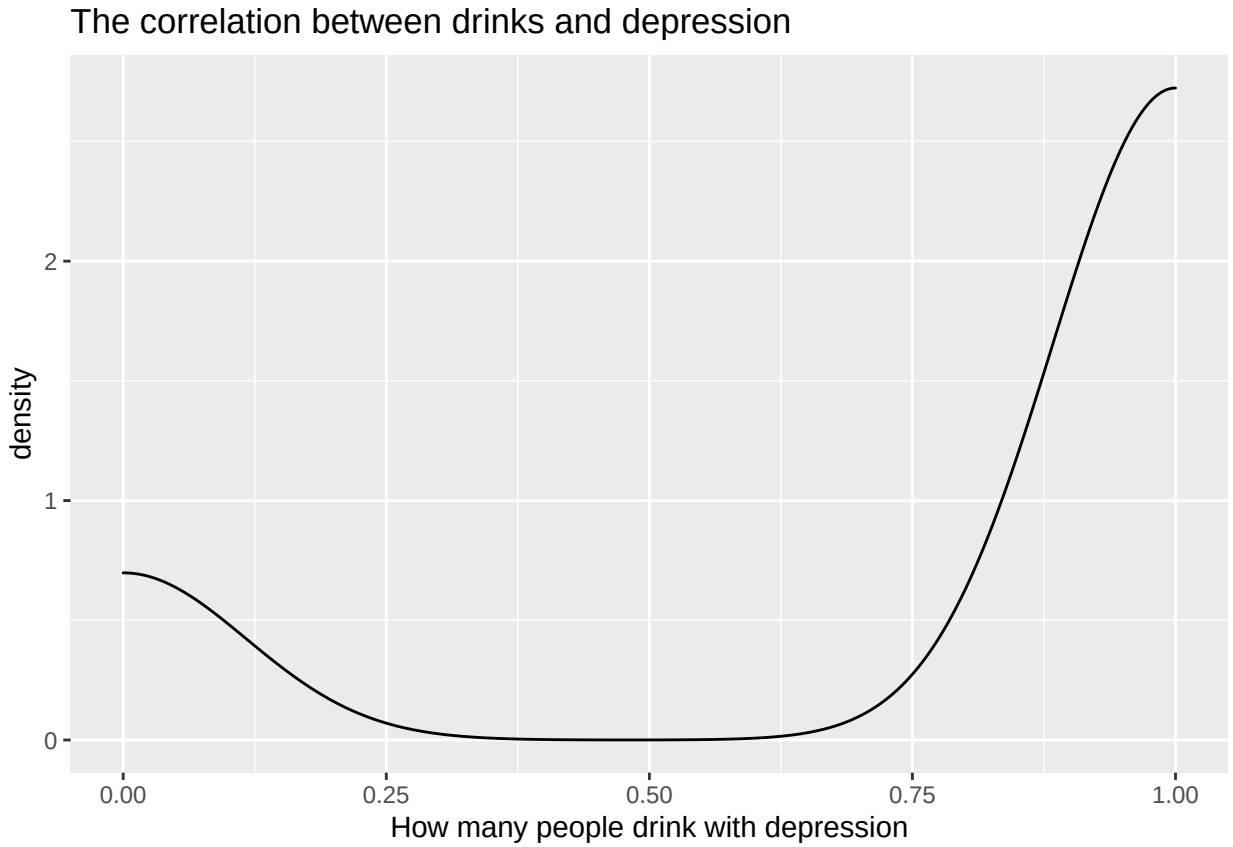
The table shown above represents people who suffer from depression and their employment status. People who don't have a job tend to be more sad.

```
ggplot(depress, aes(x=employ, fill=employ)) + geom_bar() + xlab("Status of employment") + ylab("Surveye
```




The chart shown above is also employment status but the bar graph helps see numbers easier. Drink:

```
ggplot(depress, aes(x=drink)) + geom_density() +xlab("How many people drink with depression") + ggtitle
```




The graph shown above shows the higher the depression, the higher the person is likely to drink more.

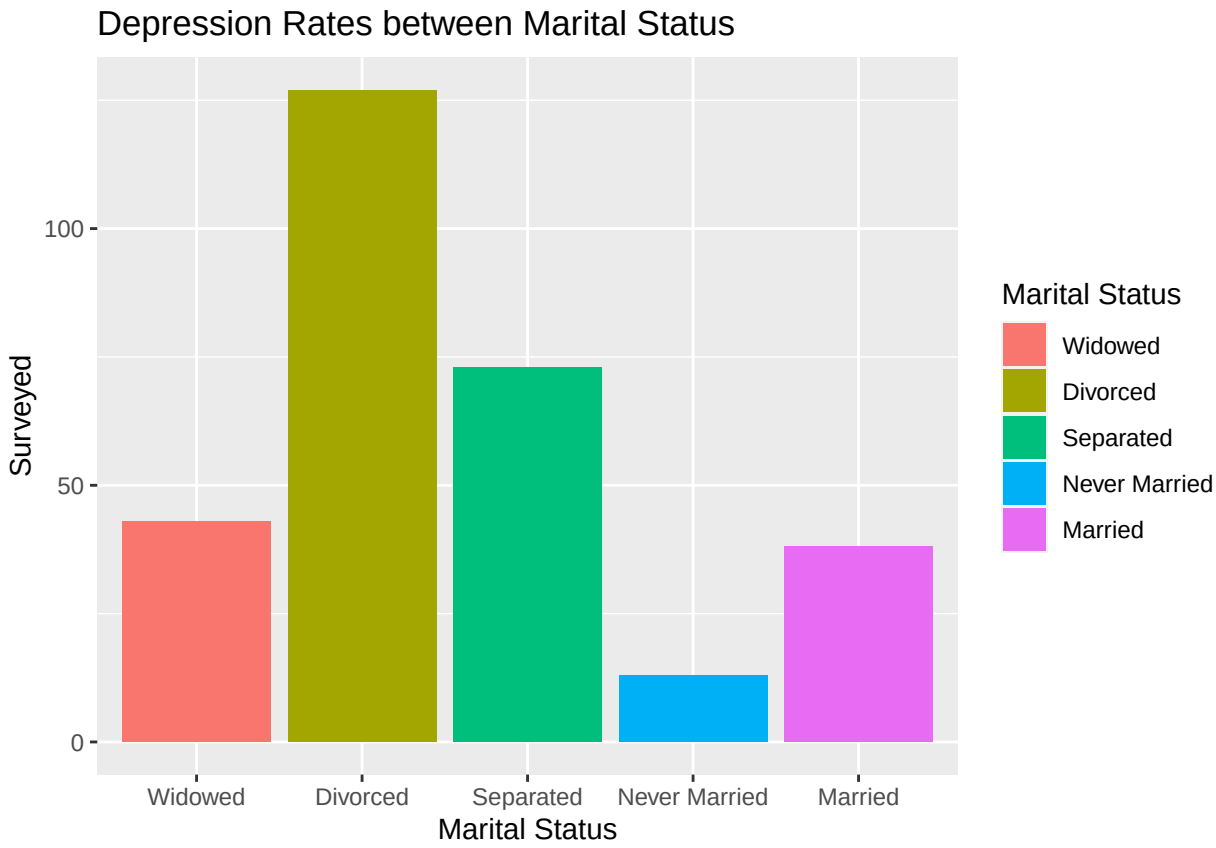
Marital

```
depress$maritalrename <- factor(depress$marital, labels=c("Widowed", "Divorced", "Separated", "Never Ma
 summary(depress$maritalrename)
```

```
##       Widowed      Divorced     Separated Never Married       Married
##            43           127            73            13            38
```

The table above shows depressive people's marital status.

```
 ggplot(depress, aes(x=maritalrename, fill=maritalrename)) + geom_bar() + xlab("Marital Status") + ylab
```




CESD Depression Scores

```
summary(depress$cesd)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   3.000   7.000   8.884  12.000  47.000
```

CESD depression scores are results from a scale of 1-60. You self reflect and score how you are feeling mentally.

BIVARIATE EXPLORATION

Employment and CESD

```
summary(depress$cesd)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   3.000   7.000   8.884  12.000  47.000
```

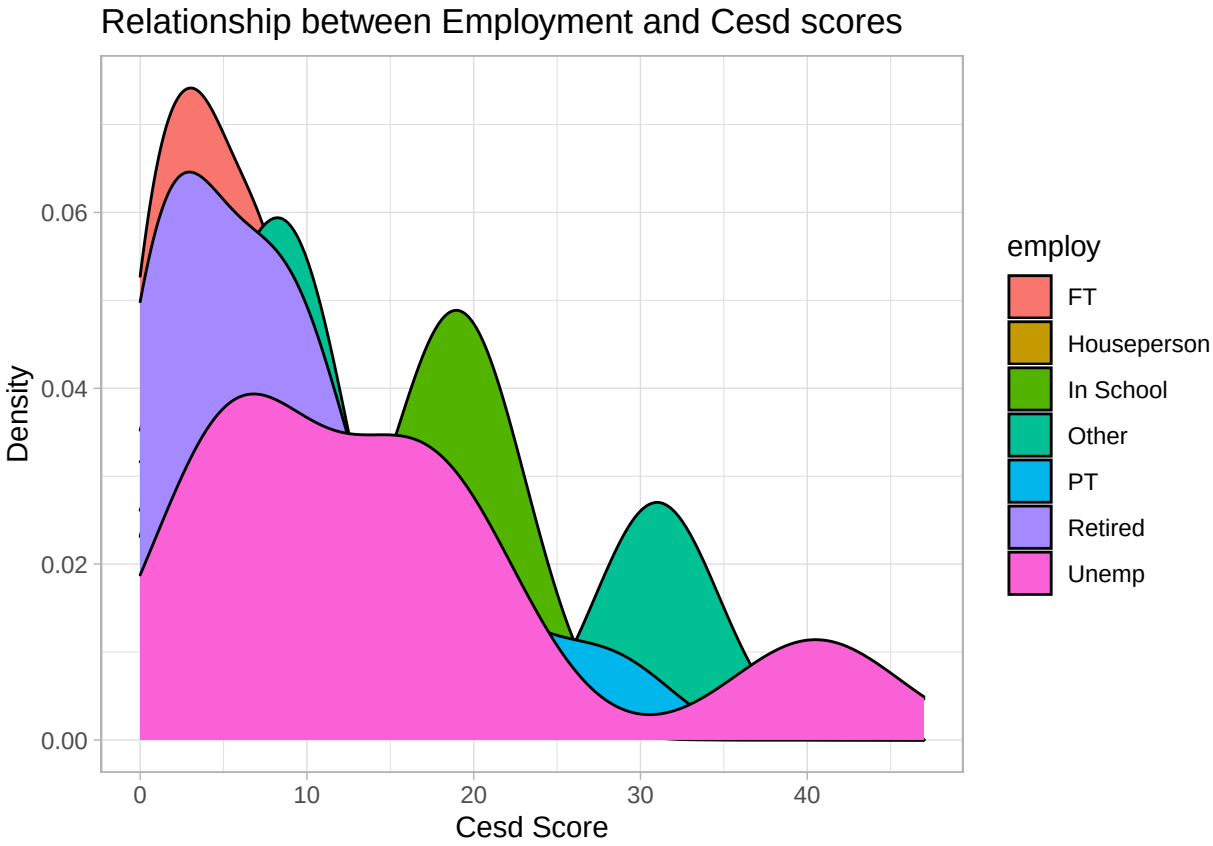
The table above is a summary of the average of employment scores with CESD scores.

```
summary(depress$employrename)
```

```
## Length  Class   Mode
##      0   NULL   NULL
```

The table above is the summary,supports our hypothesis.

```
ggplot(depress, aes(x=cesd, fill=employ)) + geom_density() + scale_fill_discrete(name="employ") + xlab
```




The above graph shows the different levels of average between each type of employment.

Drink and CESD Scores

```
summary(depress$drink)
```

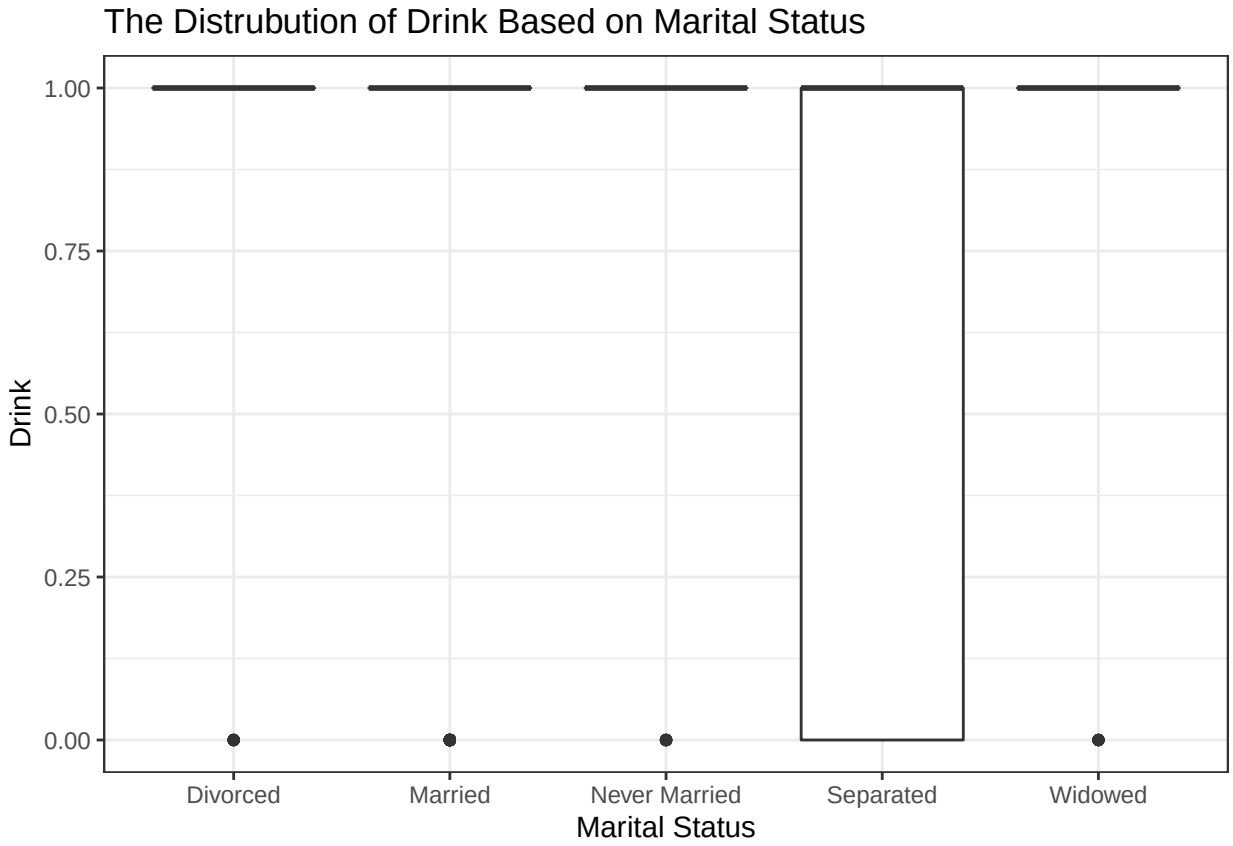
```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0000  1.0000  1.0000  0.7959  1.0000  1.0000
```

The above chart is the average of drink and CESD scores.

```
summary(depress$marital)
```

```
##    Length     Class      Mode
##       294 character character
```

```
ggplot(depress, aes(y=drink, x=marital)) + geom_boxplot() + theme_bw() + xlab("Marital Status") + ylab(
```




The boxplot above shows the distribution of drinks with different marital statuses. The box is more elongated than the other variables looking at the separated variable. The other variables barely made a dent in the graph.

Conclusion: Starting this project, I was curious about what common triggers affect depressed people. After looking at the study of depressed people in LA county and variables like employment, marital status, and if they drink. After identifying the correlation between CESD scores and these variables, unemployed people who only work in the house or don't have much in their lives are more likely to suffer from depression. The data backs up my claim. For example, looking at the bar graph of marital status, divorced people have depression.