

# Depression in Los Angeles County Residents

Introduction:

For this exploratory data analysis project, I am using the Depression data set. This data set is from the first set of interviews of a study of depression in adult residents of Los Angeles County. It has 37 variables, but the two that I will be exploring are age and employment status.

```
depression <- read.table("/Users/harvinderathwal/Desktop/Math 130/depress_081217.txt", header=TRUE, sep = "\t")

library(ggplot2)
library(sjPlot)
```

```
## Registered S3 methods overwritten by 'parameters':
## method from
## as.double.parameters_kurtosis datawizard
## as.double.parameters_skewness datawizard
## as.double.parameters_smoothness datawizard
## as.numeric.parameters_kurtosis datawizard
## as.numeric.parameters_skewness datawizard
## as.numeric.parameters_smoothness datawizard
## print.parameters_distribution datawizard
## print.parameters_kurtosis datawizard
## print.parameters_skewness datawizard
## summary.parameters_kurtosis datawizard
## summary.parameters_skewness datawizard
```

```
## Install package "strengjeacke" from GitHub ('devtools::install_github("strengjeacke/strengjeacke")') to load all sj-packages at once!
```

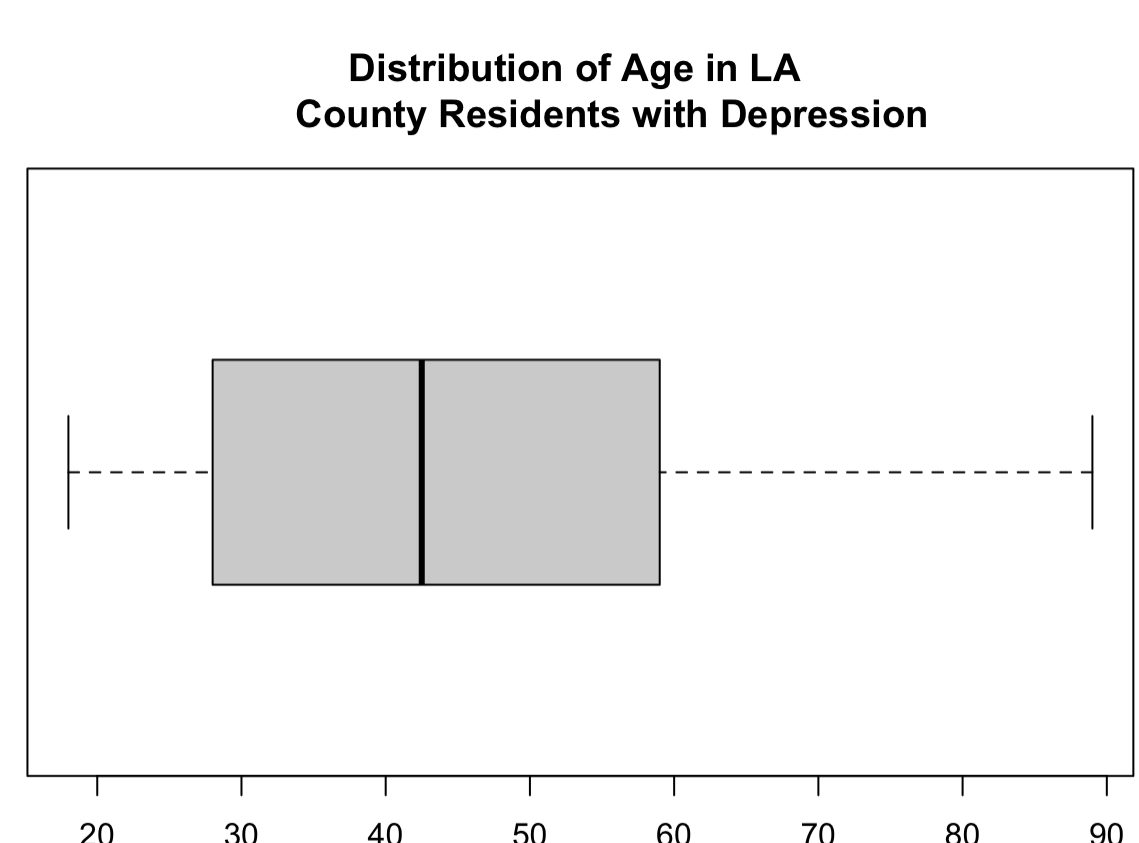
Univariate Description:

First, I want to explore the variation in age for depression. What's the youngest and oldest residents with depression?

```
summary(depression$age, digits = 3)
```

```
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 18.0 28.0 42.5 44.4 59.0 89.0
```

```
boxplot(depression$age, horizontal = TRUE, main= "Distribution of Age in LA County Residents with Depression")
```



The summary and help us clearly see

that the youngest age is 18 years old and oldest is 89. The average age of residents with depression is 44.

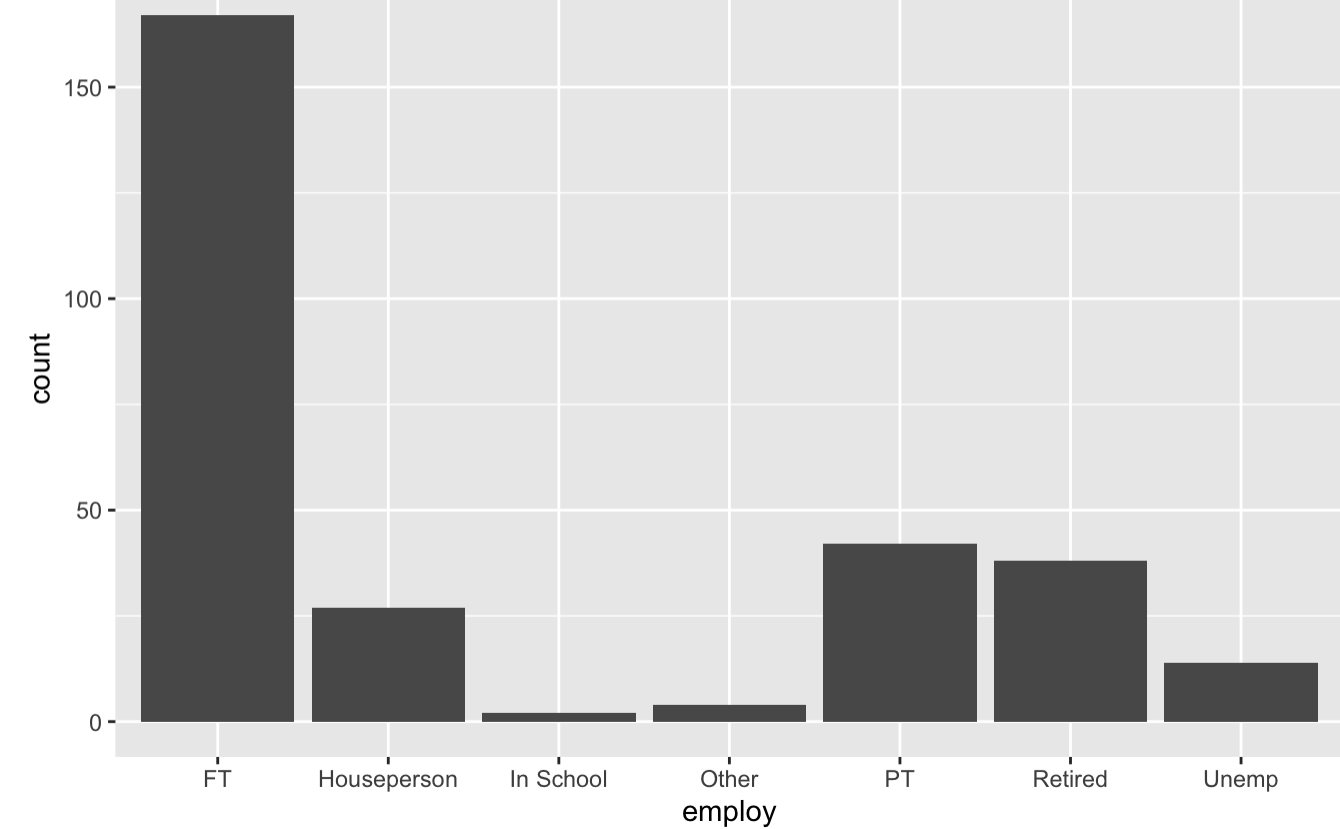
Next, I want to examine employment status. There are 7 categories in this variable; full person, houseperson, in school, part time, retired, unemployment, and other.

```
table(depression$employ)
```

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

```
ggplot(depression, aes(x=employ)) + geom_bar() + ggtitle("Employment Status of L.A. County Residents with Depression")
```

Employment Status of L.A. County Residents with Depression



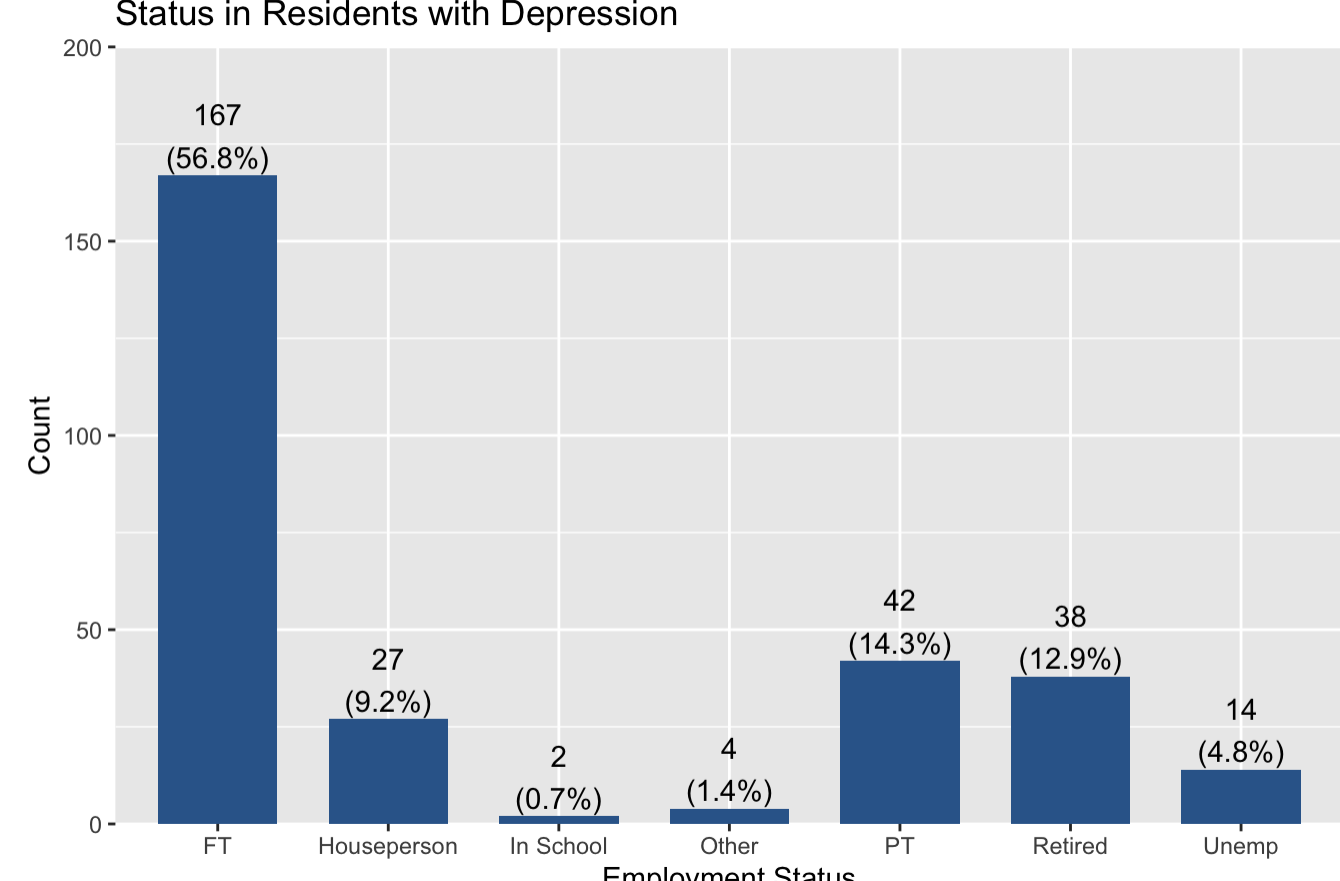
The data shows us that the majority

of residents (167 residents), who have depression are working full time. The least amount of residents (2 residents) who have depression are in school. I wanted to explore what percent of residents are working full time.

```
plot_frq(depression$employ) + xlab("Employment Status") + ylab("Count") + ggtitle("Percentage of Employment Status in Residents with Depression")
```

```
## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> = "none")` instead.
```

Percentage of Employment Status in Residents with Depression



This plot shows us that 56.8% of

residents that have depression work full time. Only 0.7% are in school. More than half of the residents with depression are working full time.

Bivariate Description:

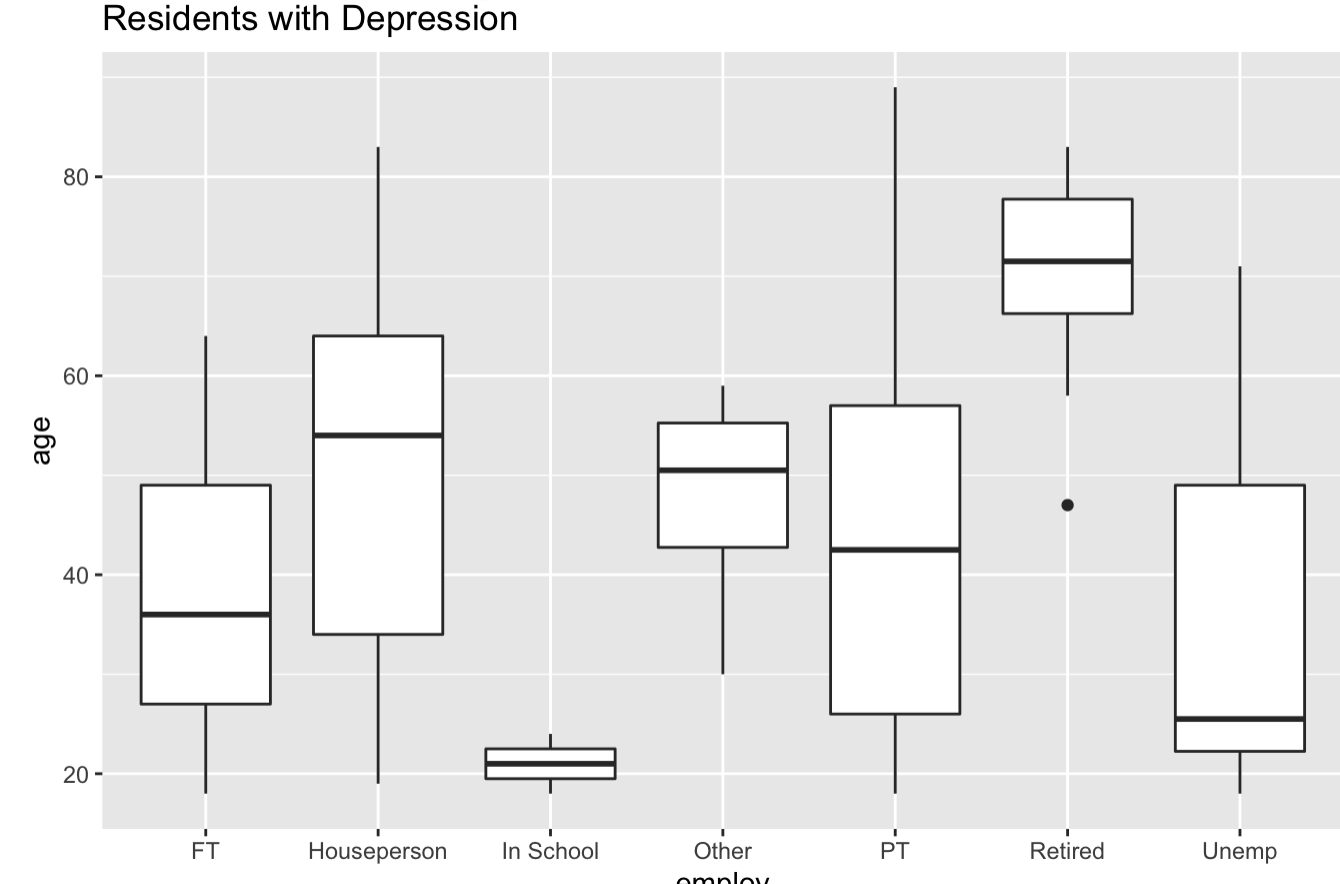
Next, I want to explore the relationship between age and employment status in residents with depression.

```
table(depression$employ, depression$age)
```

```
## FT 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39
## FT 1 2 4 5 6 7 6 4 6 4 3 3 4 5 7 4 8 2 6 4 1 1
## Houseperson 0 1 0 0 0 2 0 0 0 0 1 0 1 0 1 1 0 1 0 0 0 0
## In School 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## Other 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0
## PT 2 2 1 1 1 1 1 1 2 0 1 1 0 0 1 0 1 0 1 0 1 1
## Retired 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## Unemp 1 0 1 0 2 1 1 1 1 0 0 0 0 0 1 0 0 0 0 0 0 0
##
## FT 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61
## FT 5 1 7 5 2 3 2 3 3 4 3 2 2 2 1 1 2 2 5 5 5 4
## Houseperson 0 0 0 1 0 0 0 1 0 0 0 1 2 0 1 1 0 1 1 1 1 0
## In School 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## Other 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 1 0 0 0
## PT 0 0 2 1 0 1 0 0 1 0 1 3 0 0 0 0 1 3 0 0 1 1
## Retired 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1 1 0 1
## Unemp 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 2 0 0 0 0 0
##
## FT 62 63 64 65 66 67 68 69 70 71 72 73 74 75 77 78 79 80 81 82 83 89
## FT 4 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## Houseperson 0 1 0 1 0 1 0 0 1 1 0 0 0 1 0 0 0 0 0 0 2 0
## In School 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## Other 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## PT 0 0 1 2 1 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1
## Retired 0 1 1 2 2 1 3 1 3 1 2 2 3 0 2 2 2 1 2 1 2 0
## Unemp 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0
```

```
ggplot(depression, aes(y=age, x=employ)) + geom_boxplot() + ggtitle("Relationship Between Age and Employment Status of Residents with Depression")
```

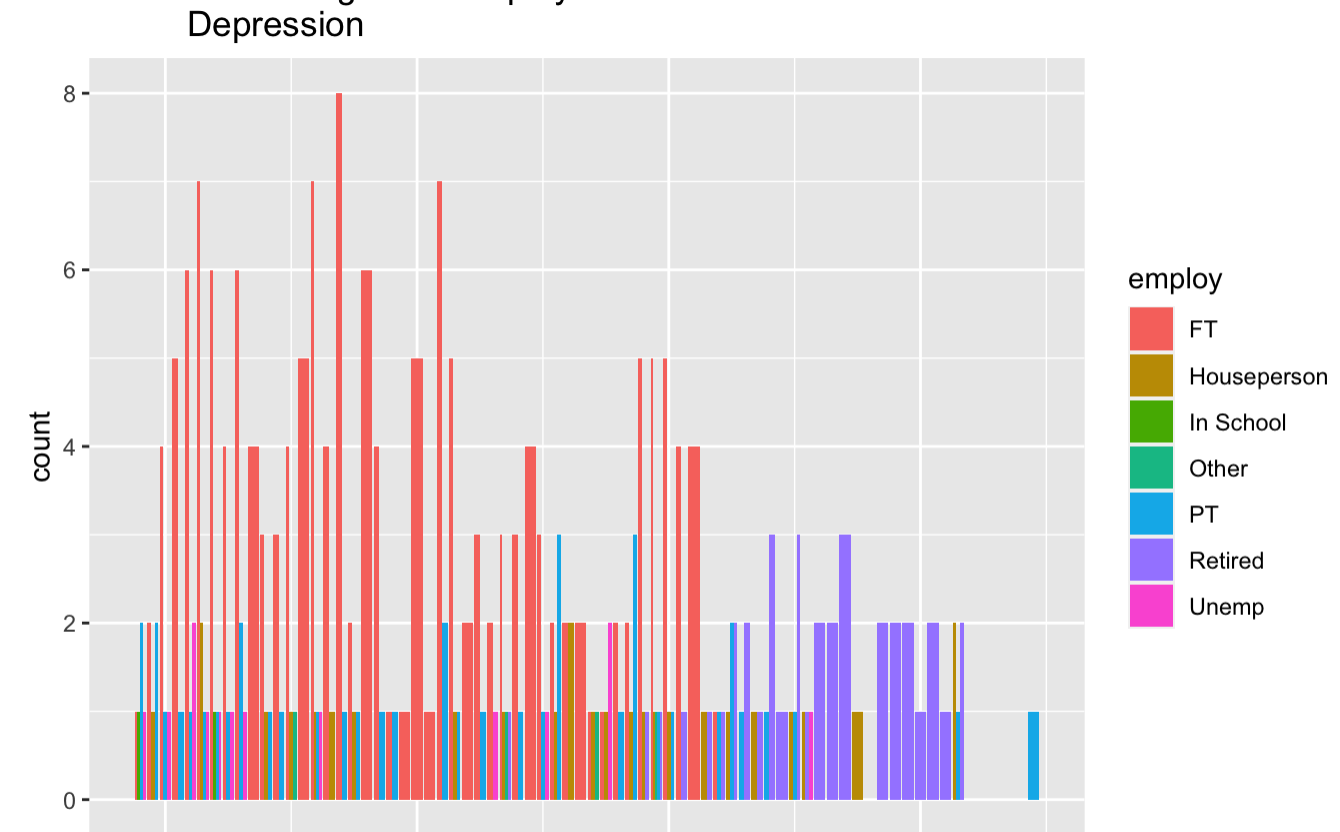
Relationship Between Age and Employment Status of Residents with Depression



This plot shows that residents who work full time are between the ages 18 and 63. The average age of full time residents is 35. Those in school are between the ages 18-24. Part time residents are between the minimum and maximum ages of 18 and 89.

```
ggplot(depression, aes(x=age, fill=employ)) + geom_bar(position="dodge") + ggtitle("Distribution of Age and Employment Status of Residents with Depression")
```

Distribution of Age and Employment Status of Residents with Depression



This plot shows us clearly shows us that majority of the residents who have depression are working full time between the ages of 18-63. Majority of residents aged above 60 that have depression are retired.

Conclusion:

Our main purpose of this analysis was to explore two variables of depression in Los Angeles County residents. These two variables were age and employment status. This analysis told us that the minimum age of these residents is 18 and maximum is 89. The average age is 44. 56.8% of these residents work full time.

Then, I did some analysis to see the relationship between age and employment status. I learned that the 56.8% of residents that work full time are the ages between 18 and 63, with the average age being 35. Therefore, the majority of residents with depression in Los Angeles County are working full time and between the ages 18 and 63.