Final Project

Edgar Hernandez

2022-09-25

```
 dp <- read.table("/Users/edgarhernandez/Downloads/depress_081217.txt",
header=TRUE, sep = "\t")
library(ggplot2)
library(sjPlot)
```

```
## Learn more about sjPlot with 'browseVignettes("sjPlot")'.
```

```
library(knitr)
```

Introduction: My Exploratory Data Analysis project pertains to the results from a Depression study in Los Angeles County. The study involves several diverse variables; however, I am soley focusing on factors related to Health and Education. This will help me evaluate the influence that each factor has in the wellness of individuals by manner of determining whether or not the dataset can present a correlation between the two variables.

Univarate Exploration:

Health-

```
summary(dp$health)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.000   1.000   2.000   1.772   2.000   4.000
```

```
mean(dp$health)
```
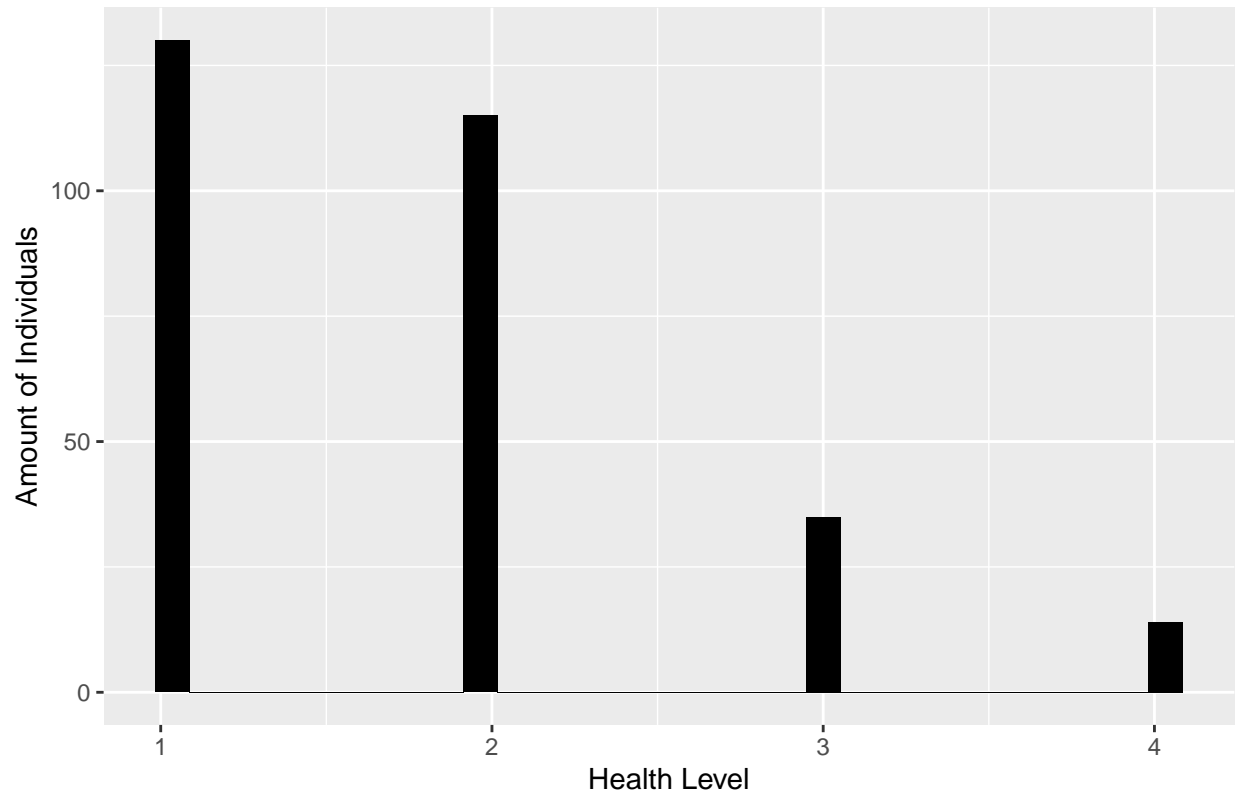
```
## [1] 1.772109
```

```
sd(dp$health)
```

```
## [1] 0.8379466
```

```
 health <- as.factor(dp$health)
levels=c("1", "2", "3", "4")
ggplot(dp, aes(x=health)) + geom_histogram(fill="black") + xlab("Health Level") + ggtitle("Health Status
```

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

Health Status of Surveyed Residents

This Graph illustrates the variety of health levels amongst the surveyed individuals. The repsonses were scored on a rangle of 1-4, where 1 is Excellent health and 4 is Poor health. In between is 2 and 3, which represents Good and Fair respectively.Evidenlty, the majority of individuals recorded a response of 1 (44%), whereas very few scored a 4 (5%)

Education

```
summary(dp$health)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.000   1.000   2.000   1.772   2.000   4.000
```
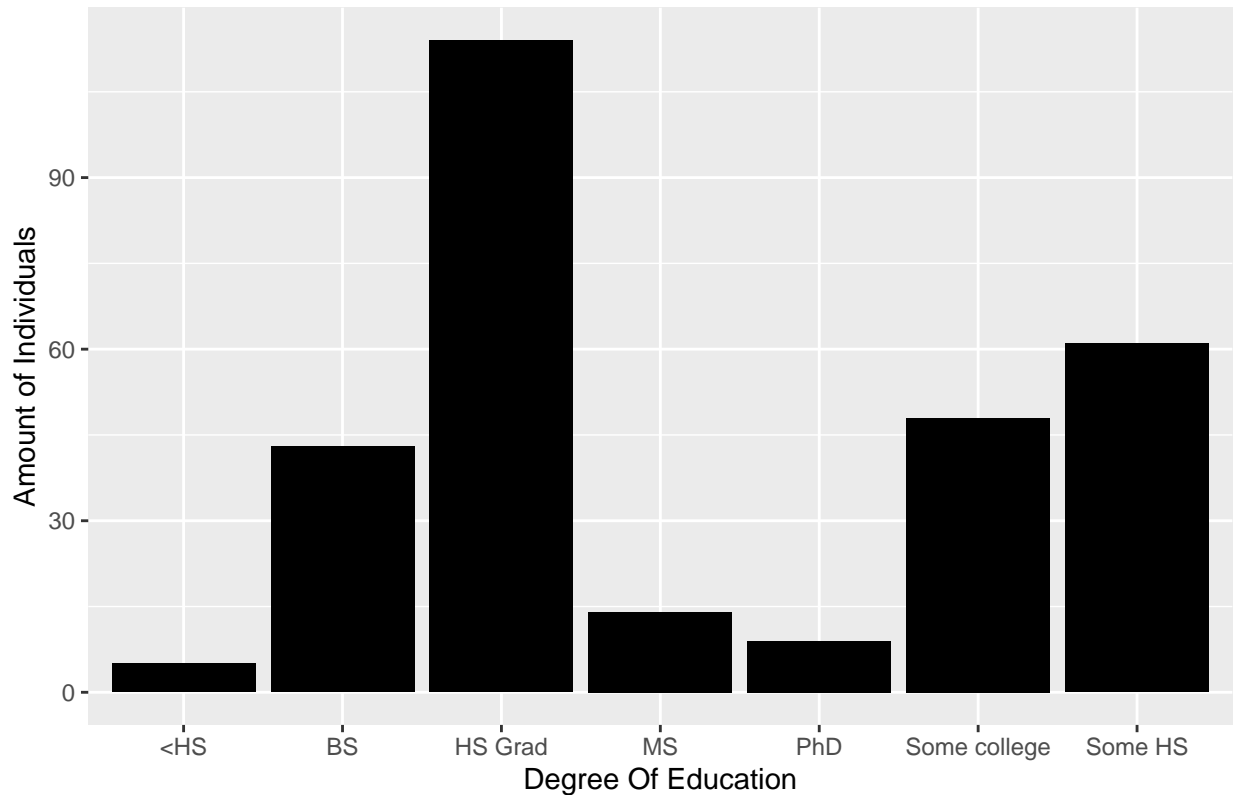
```
mean(dp$health)
```

```
## [1] 1.772109
```

```
sd(dp$health)
```

```
## [1] 0.8379466
```

```
ggplot(dp, aes(x=educat)) + geom_bar(fill="black") + xlab("Degree Of Education") + ggtitle("Educational
```

Educational Status of Surveyed Residents

This Graph displays the varying education levels of the individuals surveyed. As seen in the data, most individuals obtained a HS degree (39%), whereas the smallest recording was in fewer than HS (1.7%). Furthermore, about 37% of individuals replied with either some HS or College background, indicating that they either did not complete or are currently in progress of obtaining their degree.

Bivarate Exploration:

```
ggplot(dp, aes(x=health,fill=educat)) +
geom_bar()+ xlab("Health Level") + ylab("Amount of Indiviuals") + scale_fill_discrete(name="Education L
ggtitle("Health Vs Education Levels")
```

Health Vs Education Levels
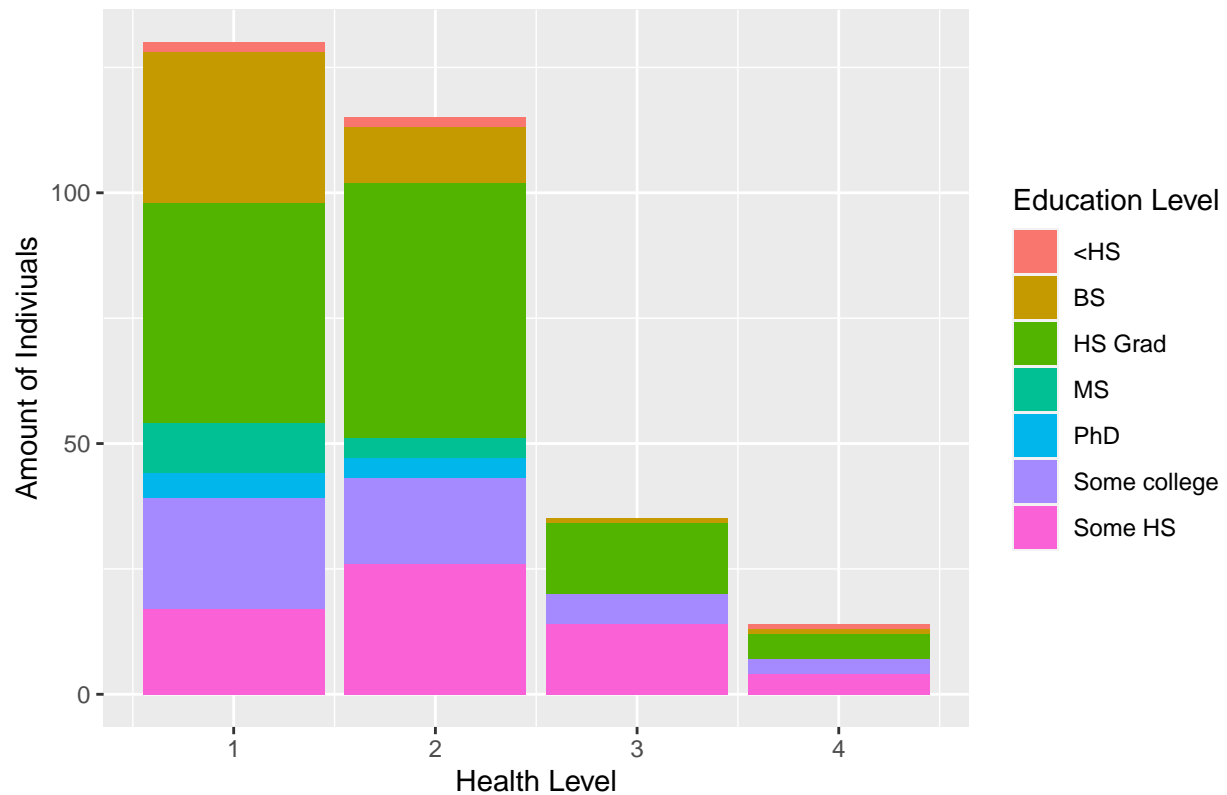

```
table(dp$health, dp$educat)
```

```
##
##     <HS BS HS Grad MS PhD Some college Some HS
##   1   2 30      44 10   5           22      17
##   2   2 11      51  4   4           17      26
##   3   0  1      14  0   0            6      14
##   4   1  1       5  0   0            3       4
```

This Graph compares two variables, in which we can see that the majority of surveyed invidiuals are HS Graduates who recorded a score of 2 (Good) for their health rating. Furthermore, the table allows us to determine that indi- viduals with a HS Graduate Degree or higher are evidently subject to being in either excellent or good health. Likewise, the select few indivudals who have either PhDs or MS did not record a fair or poor grade to their health.

Conclusion: Pertinent to the results of the data analysis, we can see that the education of a surveryed individual is a factor in evaluating their health. Surprisingly, I would predict that an increase in higher education would result in a decrease in health level; however, the opposite evidently occurred. Individuals who obtained a HS Degree or higher summed up the majority of Excellent grade responses. This helped me determine that wellness can be induced, in terms of health status, by means of pursuing a higher education.