Final Project

Kyle Colombi

February 19, 2019

The following packages were used for data analysis:

```
library("dplyr")
library("ggplot2")
library("knitr")
library("sjPlot")
```

Depress

The data set **depress** describes the results of a study conducted on depression in LA county adults. I'll be exploring the following variables: education (educat), health, and income. I want to study these variables in order to see the different ways health, income, and education affect a person's well-being and life.

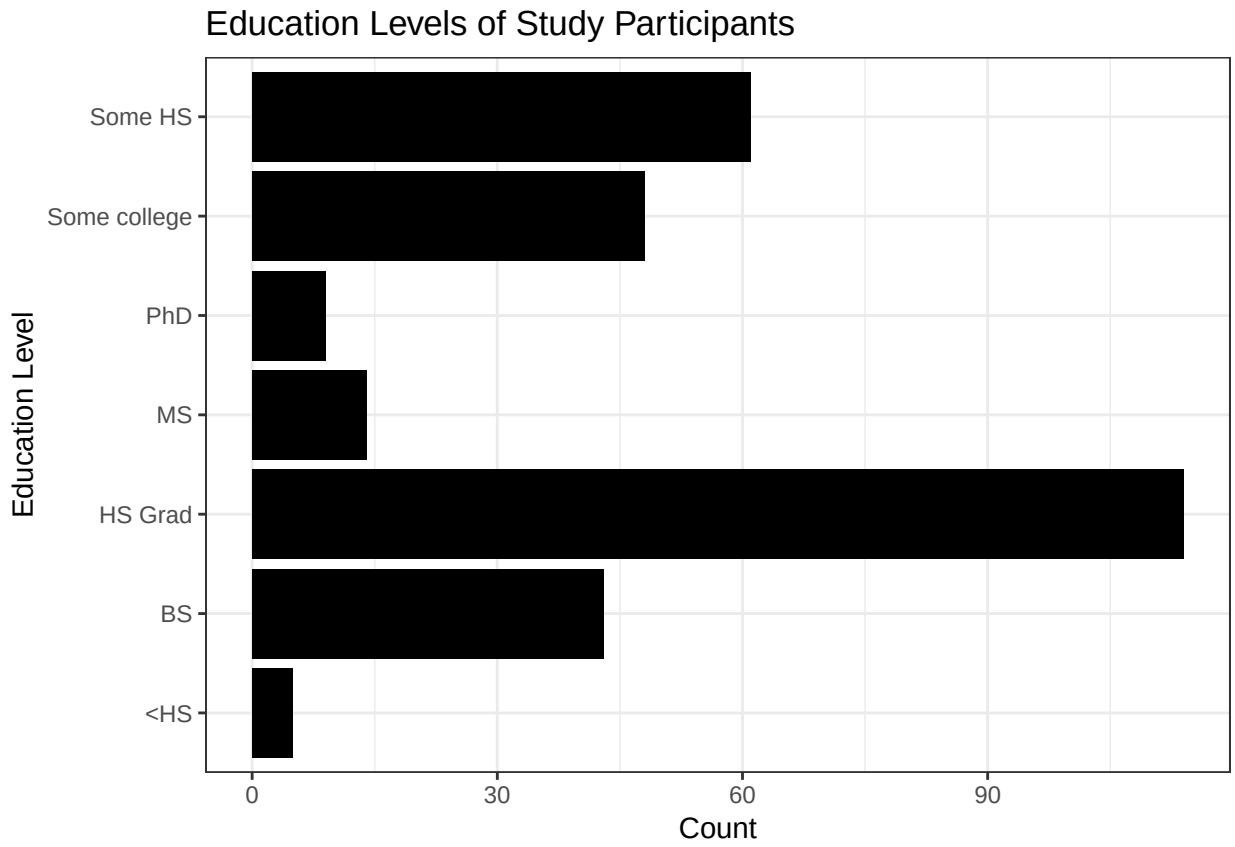
```
depress <- read.delim("C:/Users/kylec/OneDrive/Documents/spring2019/math130/depress_codebook.txt",
                      header = TRUE, sep="\t")
```

Education Summary and Univariate Graph

```
knitr::kable(t(c(summary(depress$educat))))
```

<HS	BS	HS Grad	MS	PhD	Some college	Some HS
5	43	114	14	9	48	61

```
ggplot(depress, aes(x=educat)) + geom_bar(fill="black") + coord_flip() +
  ylab("Count") + xlab("Education Level") +
  ggtitle("Education Levels of Study Participants") + theme_bw()
```



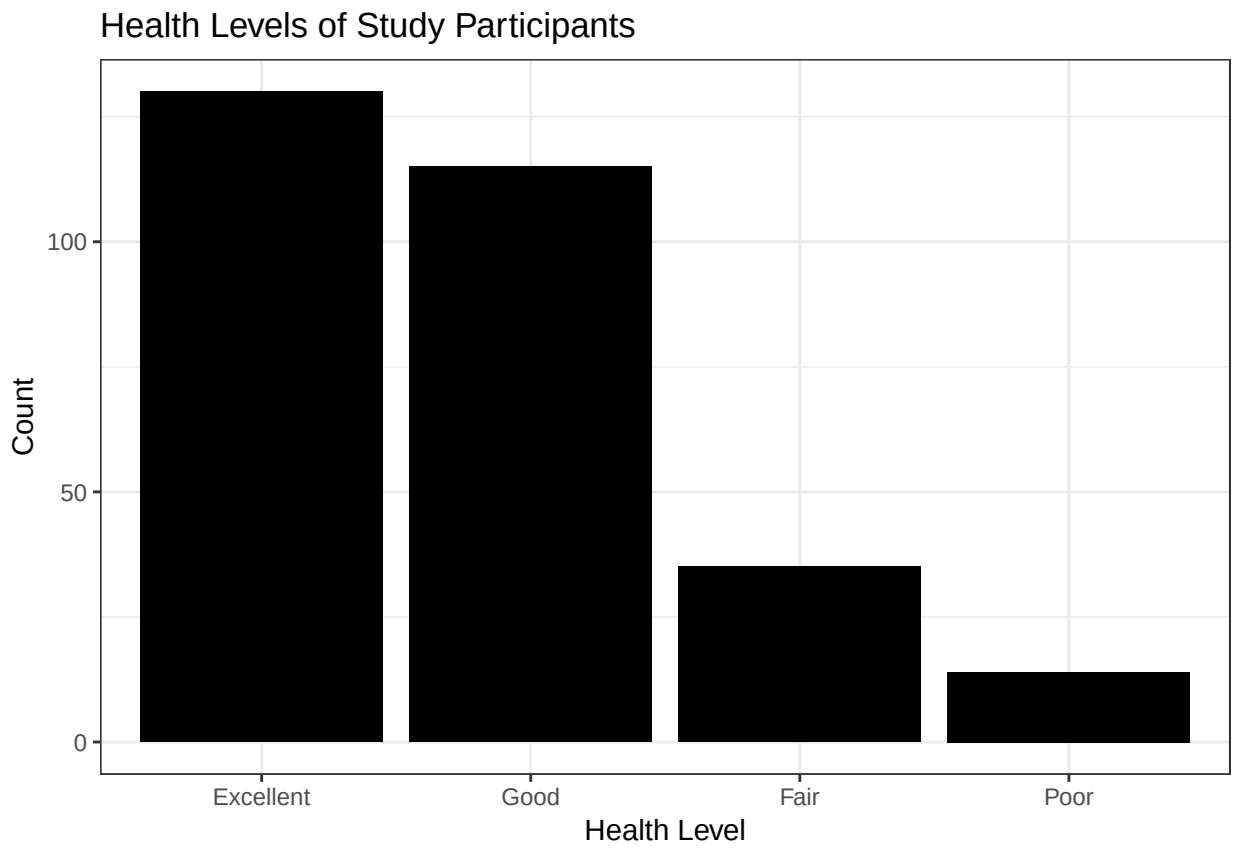
This graph displays the highest degree held by the study participants. About 39% of participants hold a high school degree (114, N=294). The smallest group present was those with less than a high school degree at 1.7% (5, N=294).

Health Summary and Univariate Graph

```
health_fac <- as.factor(depress$health)
levels=c("1", "2", "3", "4")
levels(health_fac) <- c("Excellent", "Good", "Fair", "Poor")
depress["Health_Factor"] <- health_fac
knitr::kable(t(c(summary(depress$Health_Factor))))
```

Excellent	Good	Fair	Poor
130	115	35	14

```
ggplot(depress, aes(x=Health_Factor)) + geom_bar(fill="black") +
  xlab("Health Level") + ggtitle("Health Levels of Study Participants") +
  ylab("Count") + theme_bw()
```



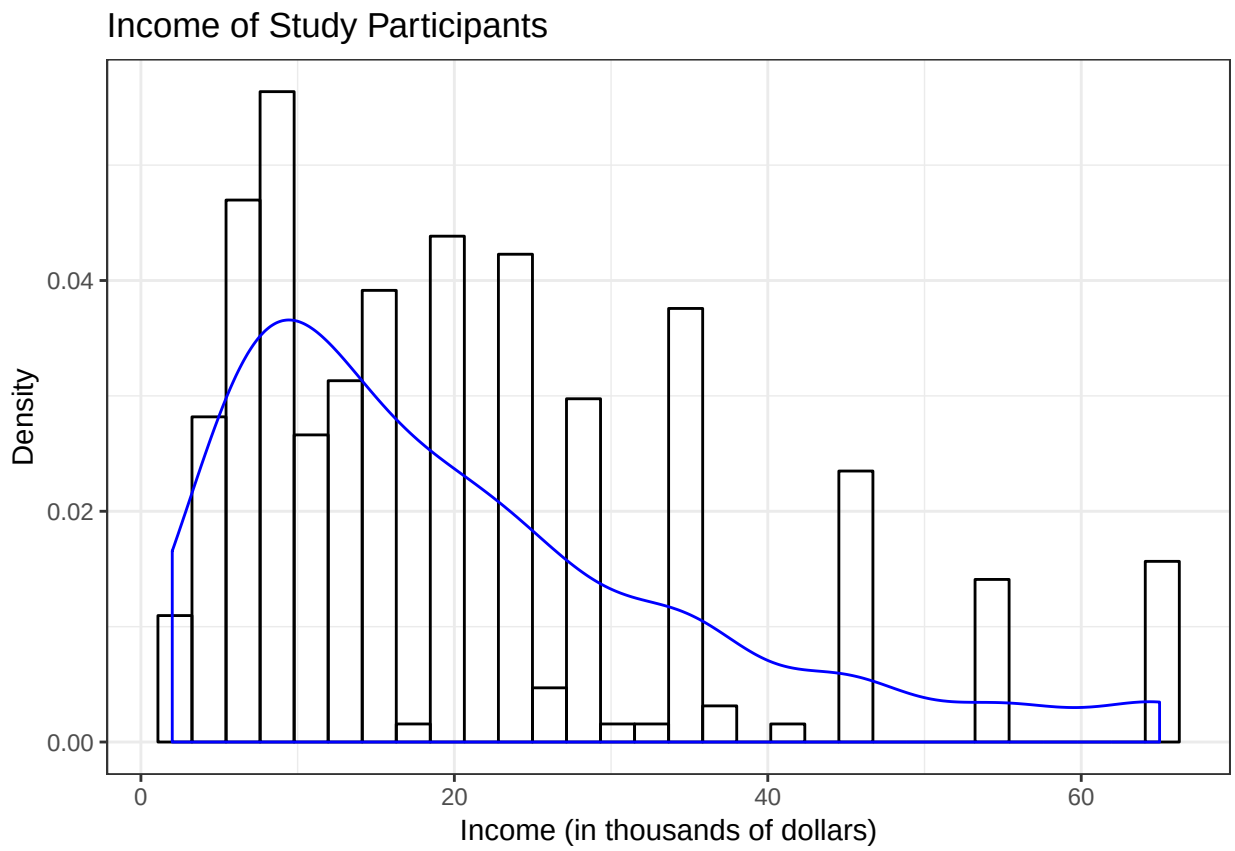
The health levels of study participants. 44% of participants had excellent health(130, N=294), 39% had good health (115, N=294), 12% had fair health (35, N=294), and 5% had poor health (14, N=294).

Income Summary and Univariate Graph

```
knitr::kable(t(c(summary(depress$income), Sd=sd(depress$income))), digits=1)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Sd
2	9	15	20.6	28	65	15.3

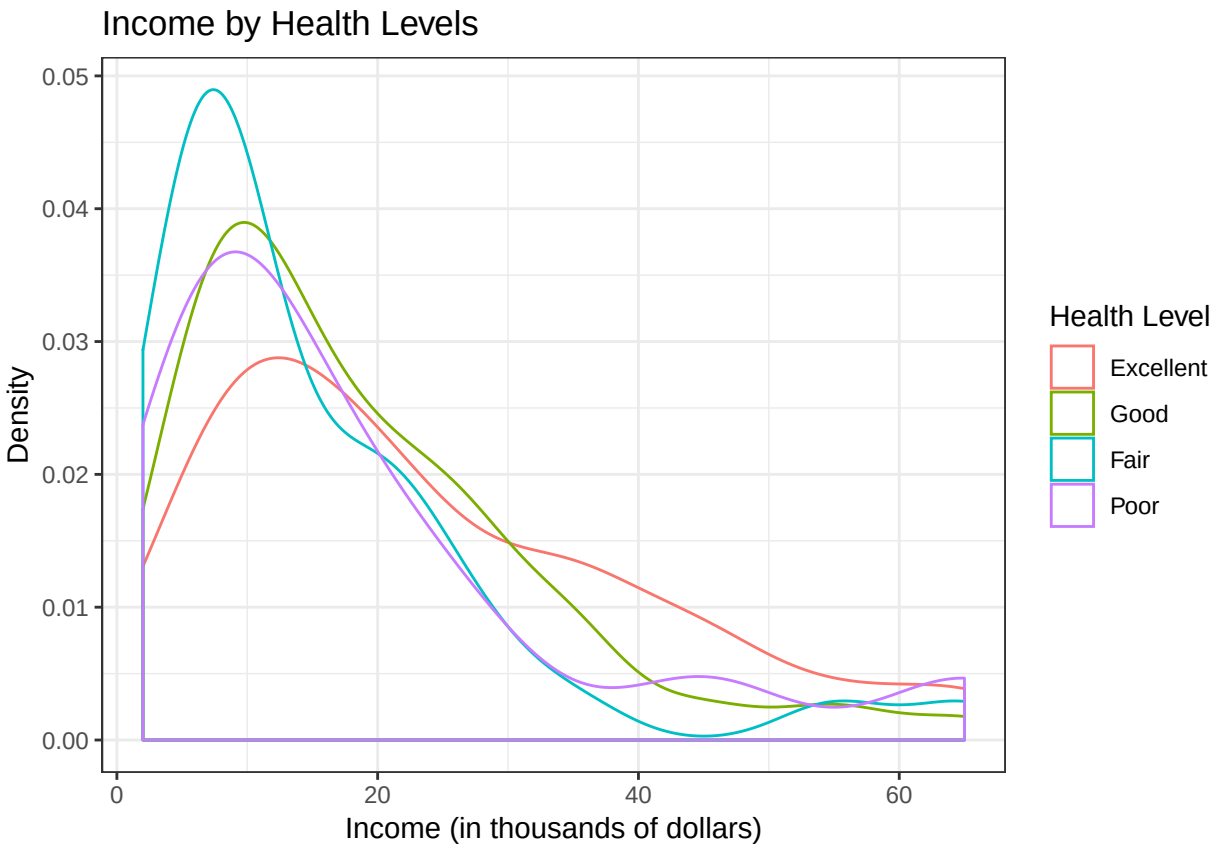
```
ggplot(depress, aes(x=income)) + geom_histogram(aes(y=..density..), color="black", fill=NA) +
  geom_density(col="blue") + xlab("Income (in thousands of dollars)") +
  ggtitle("Income of Study Participants") + ylab("Density") + theme_bw()
```



This graph shows the income level, in thousands, of study participants. The range was 2 to 65. The mean income was 20.6 and the median was 15. The standard deviation was 15.3. The graph is skewed to the right and shows the majority of participants earned less than 30 thousand.

Income vs Health Graph

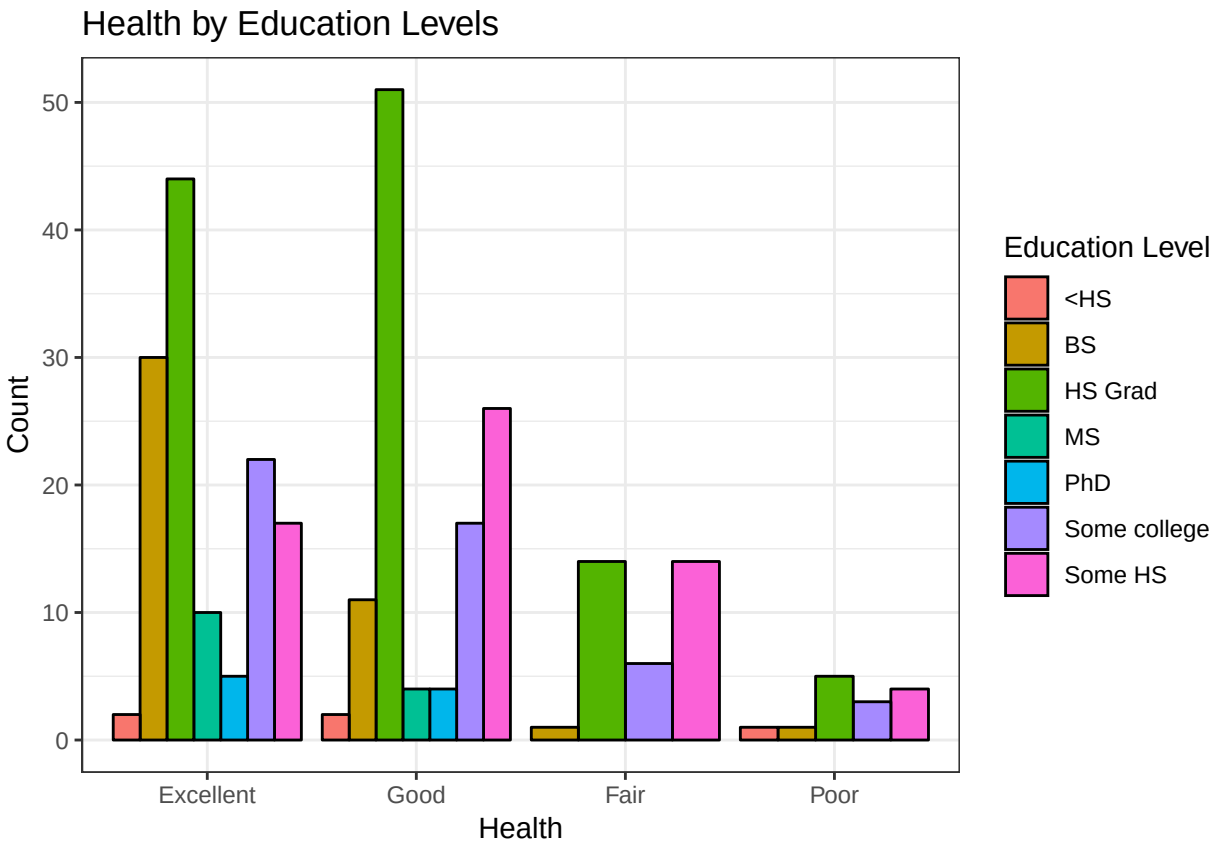
```
ggplot(depress, aes(x=income, color=Health_Factor)) + geom_density(alpha=.3) +
  ggtitle("Income by Health Levels") + scale_color_discrete(name="Health Level") +
  ylab("Density") + xlab("Income (in thousands of dollars)") + theme_bw()
```



The graph above displays income by health levels. Based on this data, the higher one's income is, the higher your overall health level will be. This can be seen from the 30-60 thousand region. The highest proportion of those with fair-poor health are those making less than 20 thousand a year.

Health vs Education Graph

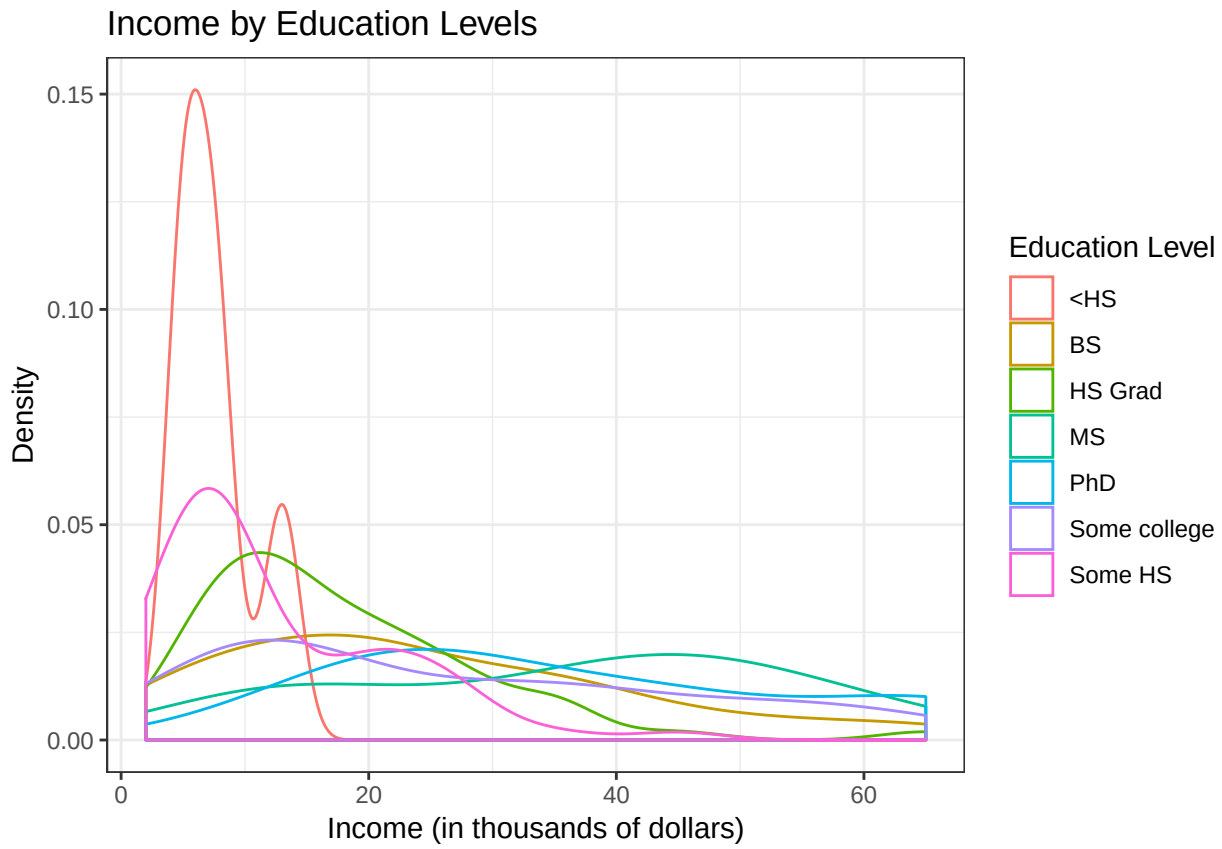
```
ggplot(depress, aes(x=Health_Factor, fill=educat)) +
  geom_bar(position = "dodge", color="black") + xlab("Health") + ylab("Count") +
  scale_fill_discrete(name="Education Level") +
  ggtitle("Health by Education Levels") + theme_bw()
```



This graph shows health levels by education levels. It shows that those with a BS degree or higher are more likely to be in excellent or good health when comparing proportions of each group. Interestingly, Those with MS and PhD degrees were only reported in the good-excellent health categories.

Income vs Education Graph

```
ggplot(depress, aes(x=income, color=educat)) + geom_density(alpha=0.3) +
  scale_color_discrete(name="Education Level") + ylab("Density") +
  xlab("Income (in thousands of dollars)") + ggtitle("Income by Education Levels") + theme_bw()
```



The final graph shows income by education levels. Those with less than a high school degree make far less than other groups, around 0-20 thousand, and by a significant proportion. Those with a PhD had a relatively consistent proportion across the income range, and those with a MS degree have the highest average income based on the proportions seen above.