

Project

Kyleigh Fertitta

4/3/2019

Reproducibility

```
parHIV <- read.delim("/Users/kyleigh/Desktop/math130/data/hiv.txt", header=TRUE, stringsAsFactors = FALSE)
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(ggplot2)
library(scales)
library(gridExtra)
```

```
##
## Attaching package: 'gridExtra'
## The following object is masked from 'package:dplyr':
##
##   combine
```

```
library(forcats)
```

Data Description

The data set parHIV is data collected as part of a clinical trial to evaluate behavioral interventions for families with a parent with HIV. The data include information on a subset of 252 adolescent children of parents with HIV. The data is owned by Dr. Mary Jane Rotheram-Borus, Professor of Psychology and Behavioral Sciences, Director of the Center for Community Health, Neuropsychiatric Institute, UCLA and used with permission in conjunction with the textbook Practical Multivariate Analysis by Afifi et.al. I will be exploring the JOBMO, HOWREL, and NGHB5 variables.

Univariate Description

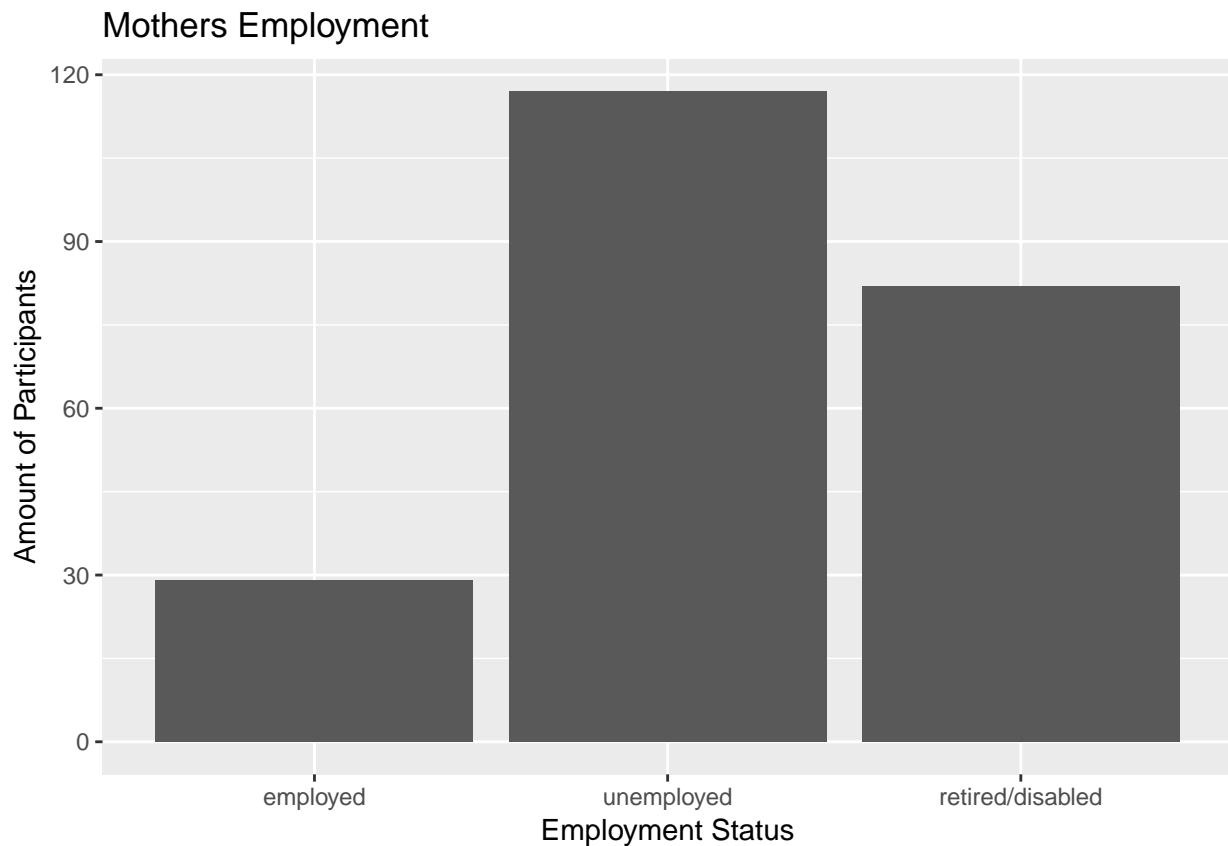
JOBMO

This variable represents the mother's job status where the variables are employed, unemployed, and retired/disabled. I will be constructing a table and a bar graph to show job status.

```
parHIV$job <- factor(parHIV$JOBMO, labels=c("employed", "unemployed", "retired/disabled"))
plot.job <- parHIV %>% select(job) %>% na.omit()
table(plot.job$job)
```

```
##
##      employed      unemployed  retired/disabled
##           29           117           82
```

```
ggplot(plot.job, aes(x=job)) + geom_bar() + ggtitle("Mothers Employment") + xlab("Employment Status") +
```



This bar graph shows how many of the mothers are employed, unemployed, or retired/disabled. Most the participants are unemployed. Least of the participants are employed.

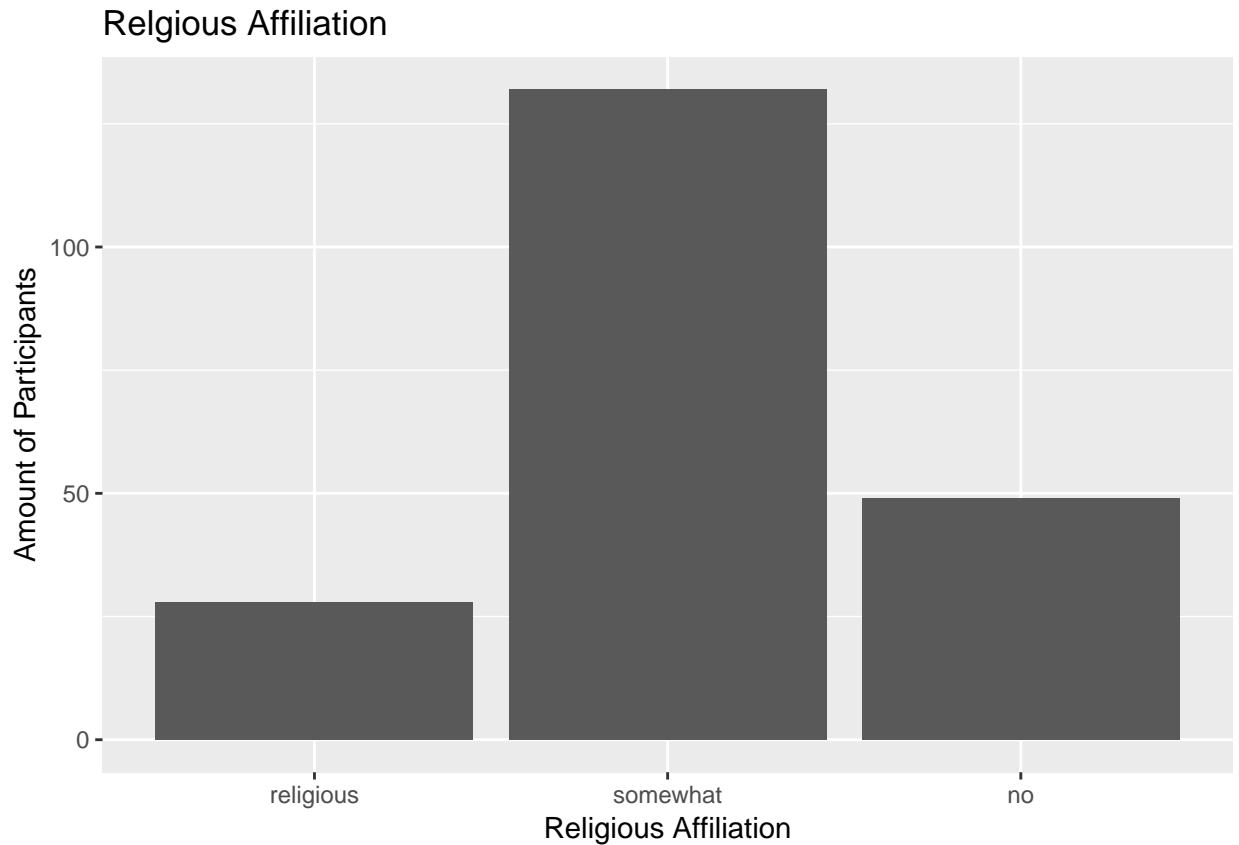
HOWREL

This variable represents how religious the HIV parent is where the variables are religious, somewhat, and not at all. I will be constructing a table and a bar graph to show religious affiliation.

```
parHIV$religion <- factor(parHIV$HOWREL, labels=c("religious", "somewhat", "no"))
plot.rel <- parHIV %>% select(religion) %>% na.omit()
table(plot.rel$religion)
```

```
##
## religious  somewhat      no
##          28          132      49
```

```
ggplot(plot.rel, aes(x=religion)) + geom_bar() + ggtitle("Religious Affiliation") + xlab("Religious Affi
```



This bar graph shows how religious the participants consider themselves to be based off the answers religious, somewhat, no, and NA. Most the participants claim to be somewhat religious. Least of the participants claim to be religious.

NGHB5

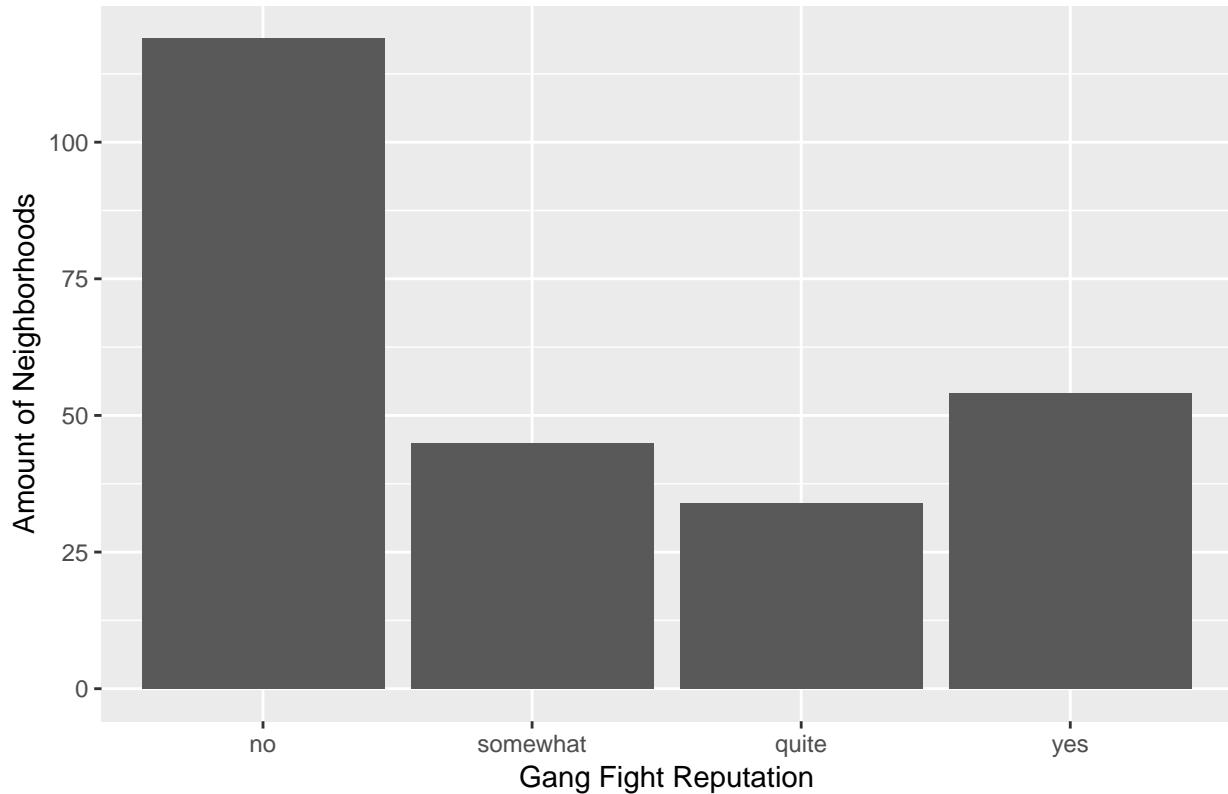
This variable represents youth gang fights in the neighborhood where the variables are not a problem, somewhat of a problem, quite a problem, and a very serious problem. I will be constructing a table and a bar graph to show youth gang fights the participants neighborhoods.

```
parHIV$gangs <- factor(parHIV$NGHB5, labels=c("no", "somewhat", "quite", "yes"))
table(parHIV$gangs)
```

```
##
##      no somewhat   quite    yes
##    119      45     34    54
```

```
ggplot(parHIV, aes(x=gangs)) + geom_bar() + ggtitle("Youth Gang Fights In Neighborhood") + xlab("Gang Fights In Neighborhood")
```

Youth Gang Fights In Neighborhood



This bar graph shows the amount of neighborhoods where youth gang fights are common based off the answers no, somewhat, quite, and yes. Most of the participants said they did not have gang fights. Least of the participants claim to have quite a gang fight problem.

Bivariate Comparison

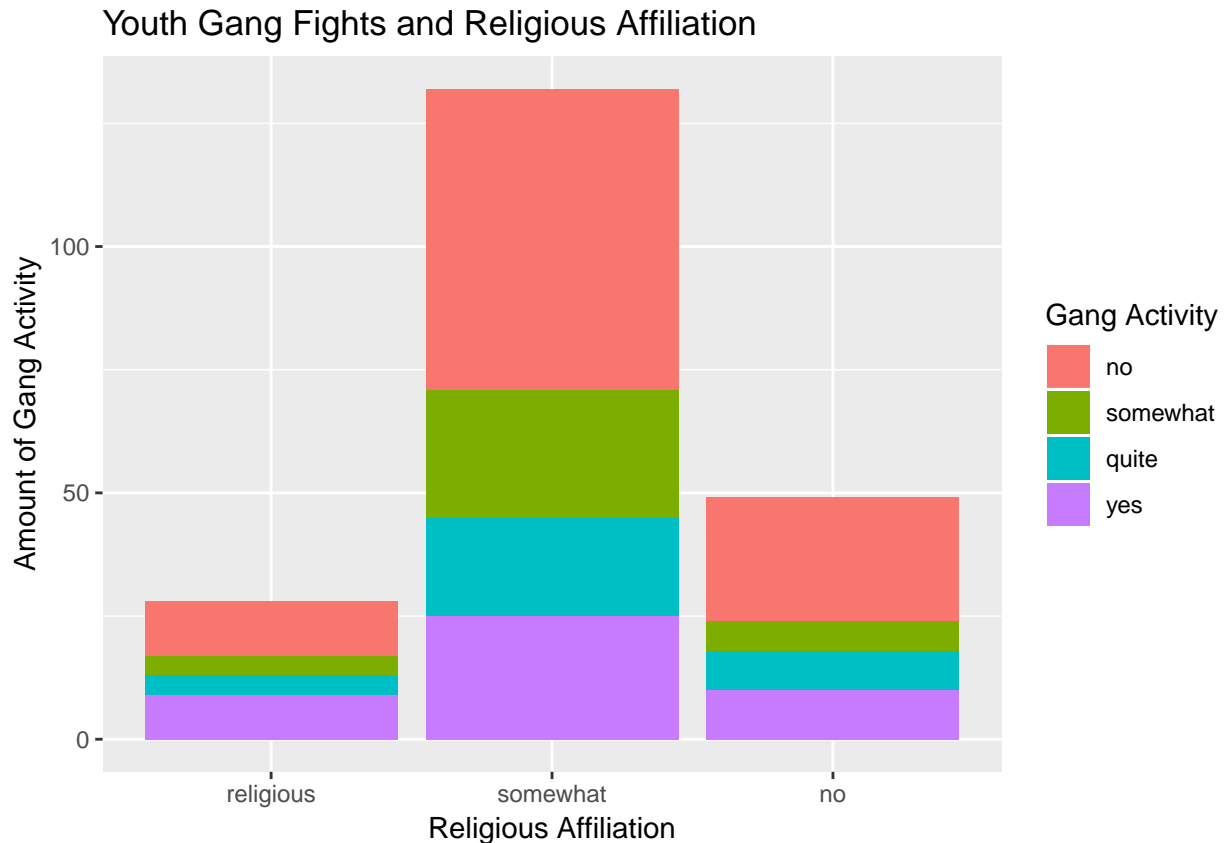
Youth Gang Fights and Religious Affiliation

Now we will see the frequency of gang fights by religious affiliation. I will be constructing a table and a bar graph to show gang fights and religious affiliation.

```
plot.relgan <- parHIV %>% select(gangs, religion) %>% na.omit()  
table(plot.relgan$religion, plot.relgan$gangs) %>% prop.table()
```

```
##  
##           no  somewhat  quite  yes  
## religious 0.05263158 0.01913876 0.01913876 0.04306220  
## somewhat 0.29186603 0.12440191 0.09569378 0.11961722  
## no        0.11961722 0.02870813 0.03827751 0.04784689
```

```
ggplot(plot.relgan, aes(x=religion, fill=gangs)) + geom_bar() + ggtitle("Youth Gang Fights and Religious Affiliation")
```



The table and graph show us that the most gang activity happens in the group where people claim to be somewhat religious, while gang activity is lower for people who claim to be or not be religious. This is surprising because you would believe nonreligious would be the most gang activity and religious would be the least.

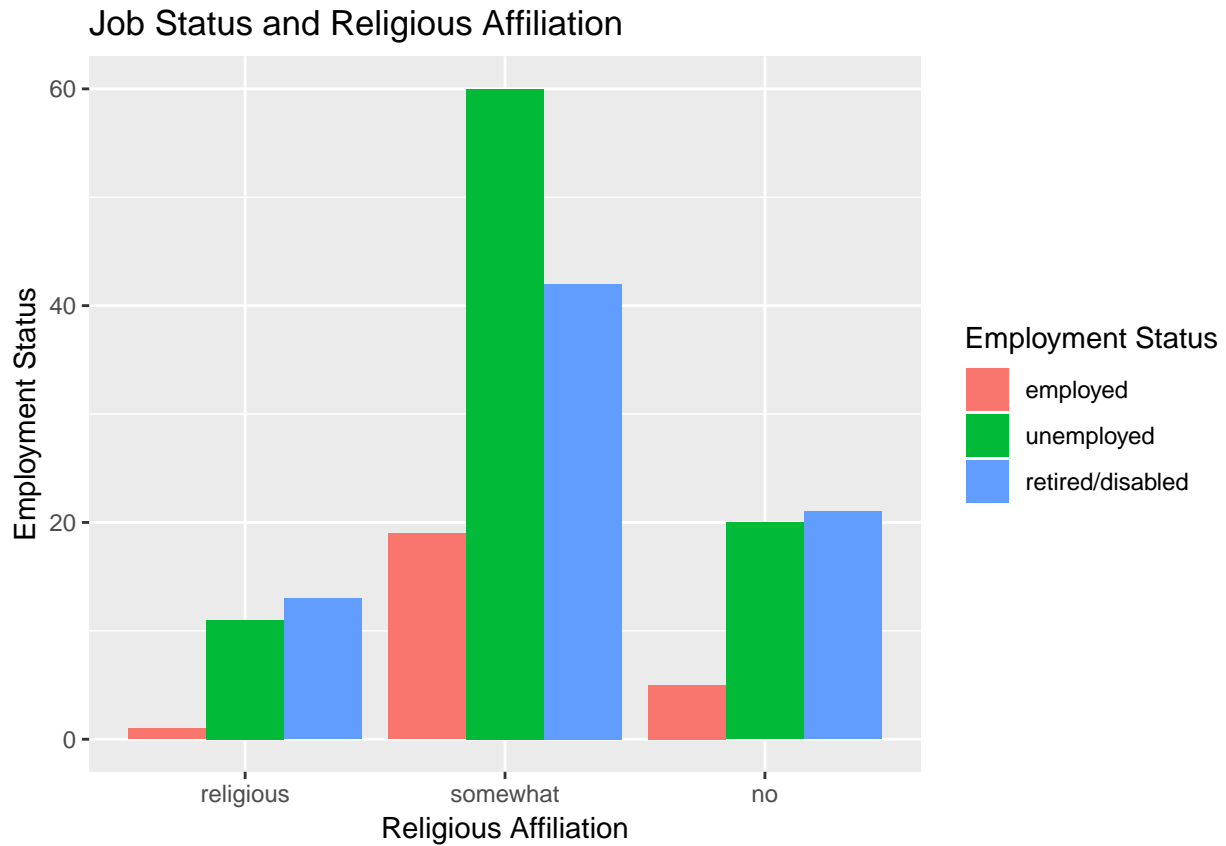
Job Status and Religious Affiliation

Now we will see the frequency of job status by religious affiliation. I will be constructing a table and a bar graph to show job status and religious affiliation.

```
plot.reljob <- parHIV %>% select(job, religion) %>% na.omit()
table(plot.reljob$religion, plot.reljob$job) %>% prop.table()
```

```
##
##           employed  unemployed  retired/disabled
## religious 0.005208333 0.057291667 0.067708333
## somewhat 0.098958333 0.312500000 0.218750000
## no       0.026041667 0.104166667 0.109375000
```

```
ggplot(plot.reljob, aes(x=religion, fill=job)) + geom_bar(position = "dodge") + ggtitle("Job Status and
```



The table and graph show us that participants who are religious are retired/disabled, somewhat religious are unemployed, and not religious are retired/disabled. This is interesting because you would think that the people who are religious or somewhat religious would be employed.