Week 5 project

Chris Pinney

2/23/2022

```
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(sjPlot)
library(forcats)
knitr::opts_chunk$set(warning=FALSE, message=FALSE)
library(RColorBrewer)
parHIV <- read.delim("../data/PARHIV_081217.txt", header=TRUE, stringsAsFactors = FALSE, sep="\t")
dim(parHIV)
```

```
## [1] 252 123
```

Introduction:

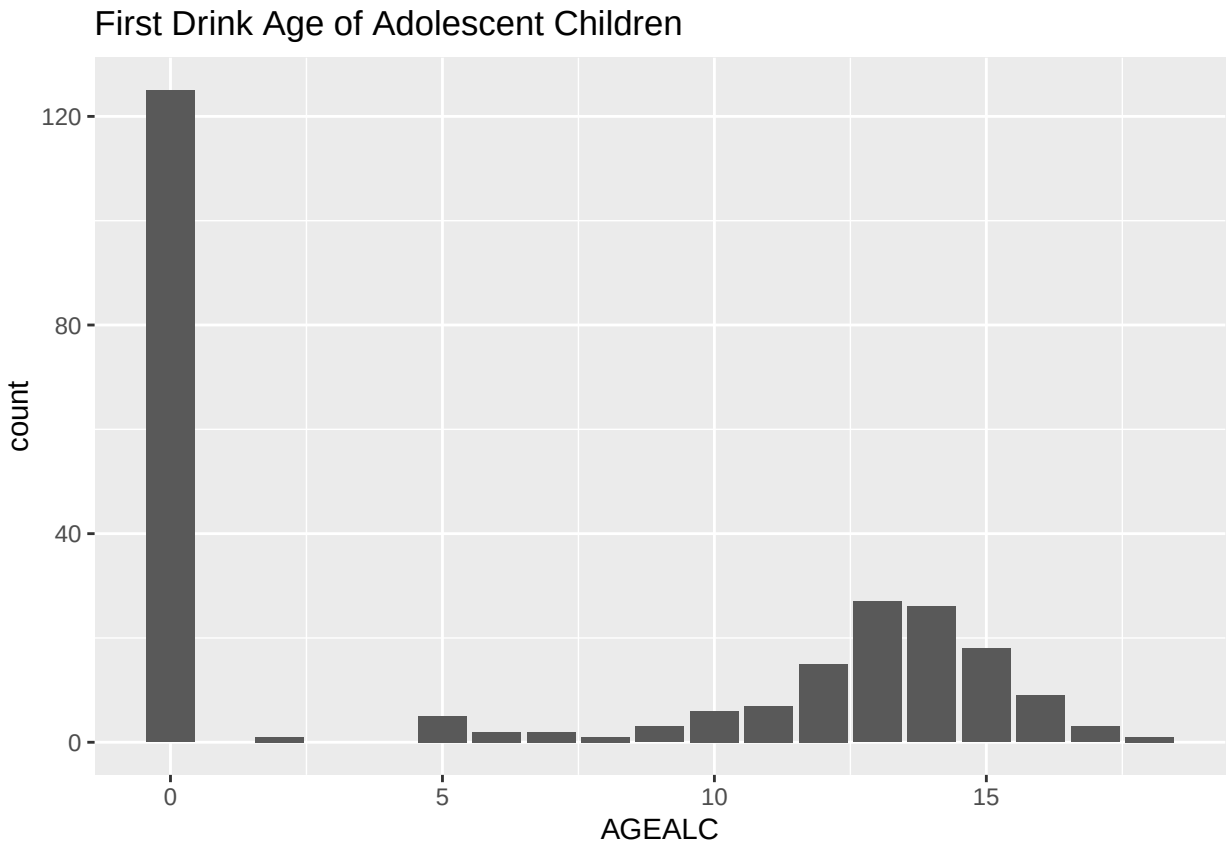
This data set focuses on the life experiences of children of adults that are HIV positive and was collected as part of a clinical trial involving 252 adolescent children of HIV+ parents. Through analyzing the following variables, AGEALC (age first had alcohol), SCHOOL (are they attending school), NGHB8 (drug dealing in neighborhood), and FINSIT (financial situation of the household), I hope to see in what way may children of HIV+ parents be impacted and how one's environment and socioeconomic status is associated with being HIV+ and I expect that the data will show that children with HIV+ parents will have tougher lives than children whose parents dont have HIV.

Age at First Drink:

```
table(parHIV$AGEALC)
```

```
##
##   0   2   5   6   7   8   9  10  11  12  13  14  15  16  17  18
## 125   1   5   2   2   1   3   6   7  15  27  26  18   9   3   1
```

```
ggplot(parHIV,aes(x=AGEALC, fill=AGEALC))+geom_bar()+ggtitle("First Drink Age of Adolescent Children")
```




```
125/252
```

```
## [1] 0.4960317
```

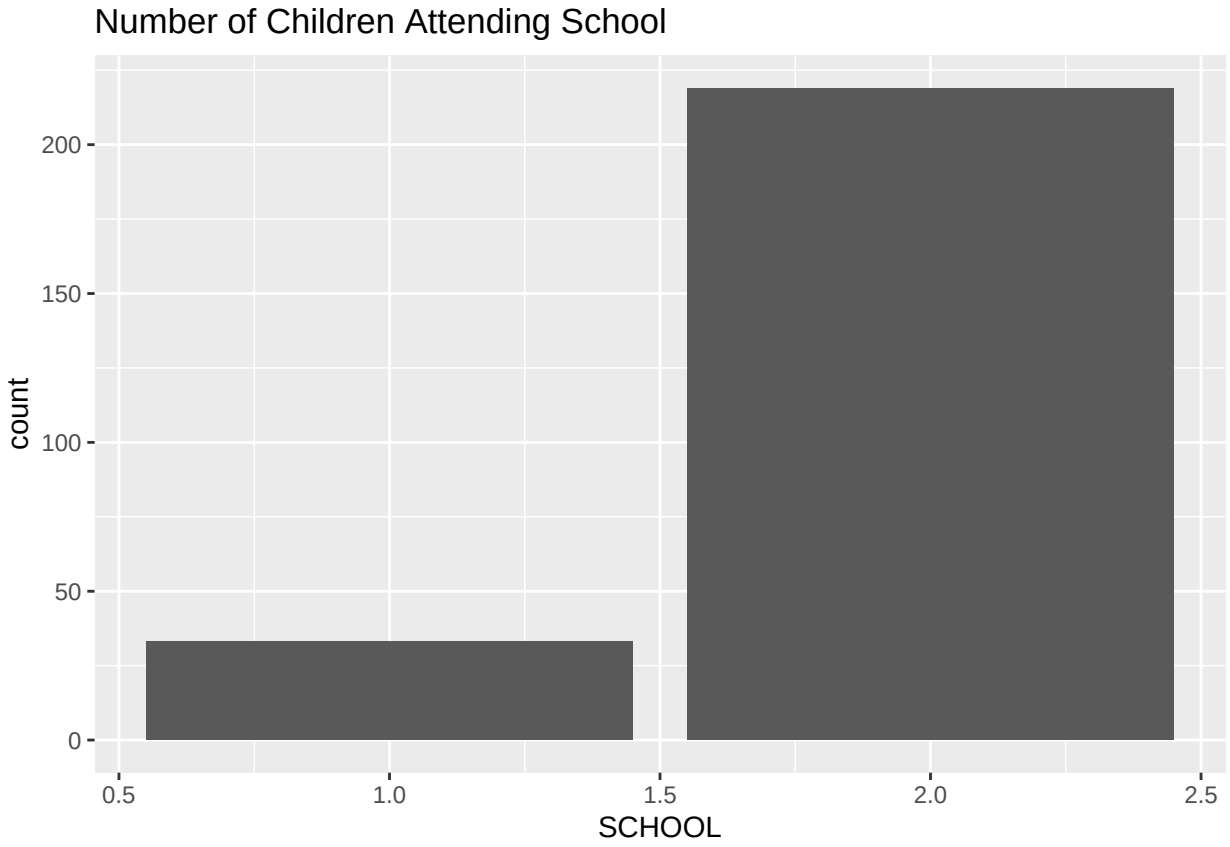
Only ~50% of adolescent children have not had an alcoholic drink yet.

School Attendance:

```
table(parHIV$SCHOOL)
```

```
##
##   1   2
##  33 219
```

```
ggplot(parHIV,aes(x=SCHOOL, fill=SCHOOL))+geom_bar()+ggtitle("Number of Children Attending School")
```




```
(219-33)/219
```

```
## [1] 0.8493151
```

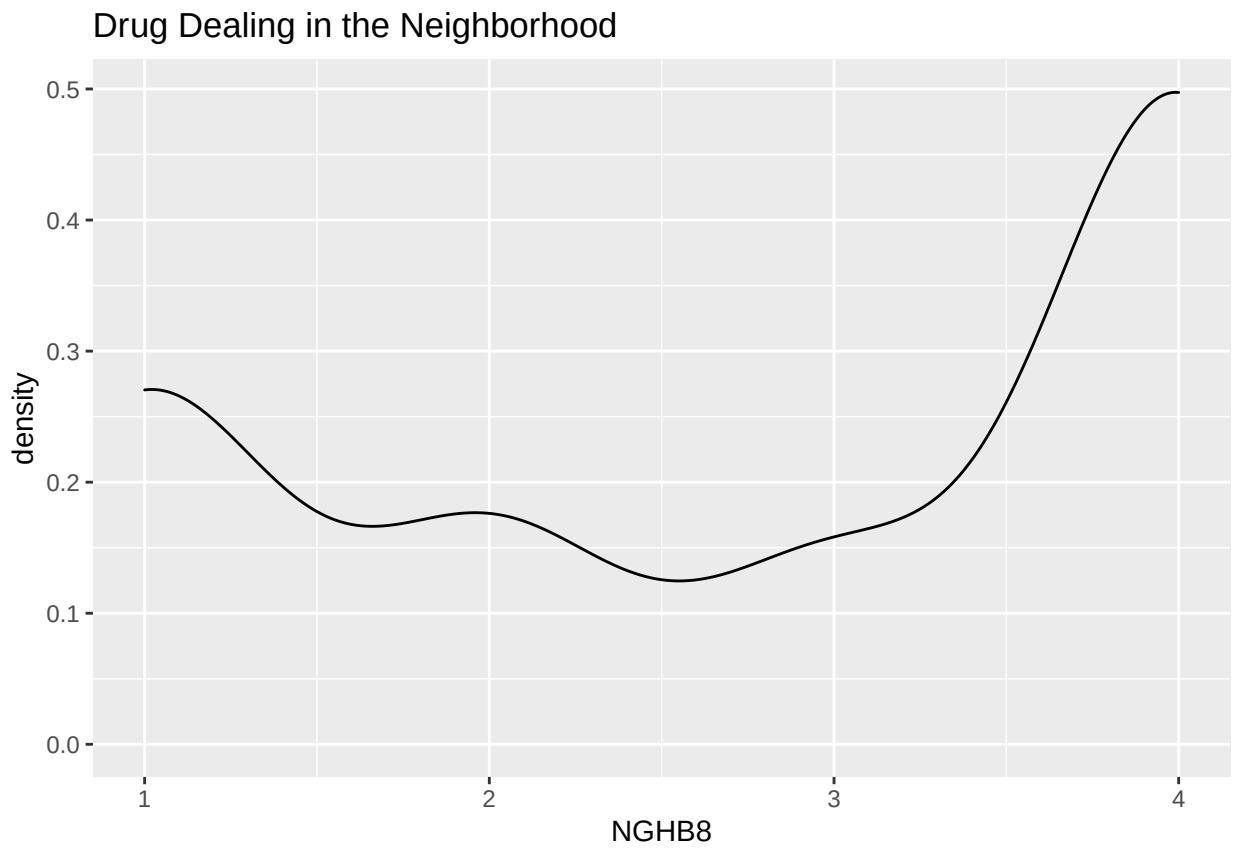
From the graph and table above, it can be seen that a majority of the children, 85%, are attending school. A "1" means that the child is not attending school, while a "2" means that the child is a student.

Drug Dealing in Neighborhood:

```
table(parHIV$NGHB8)
```

```
##
##   1   2   3   4
##  63  39  33 117
```

```
ggplot(parHIV,aes(x=NGHB8, fill=NGHB8))+geom_density()+ggtitle("Drug Dealing in the Neighborhood")
```




```
(39+33+117)/252
```

```
## [1] 0.75
```

```
117/(39+33+117)
```

```
## [1] 0.6190476
```

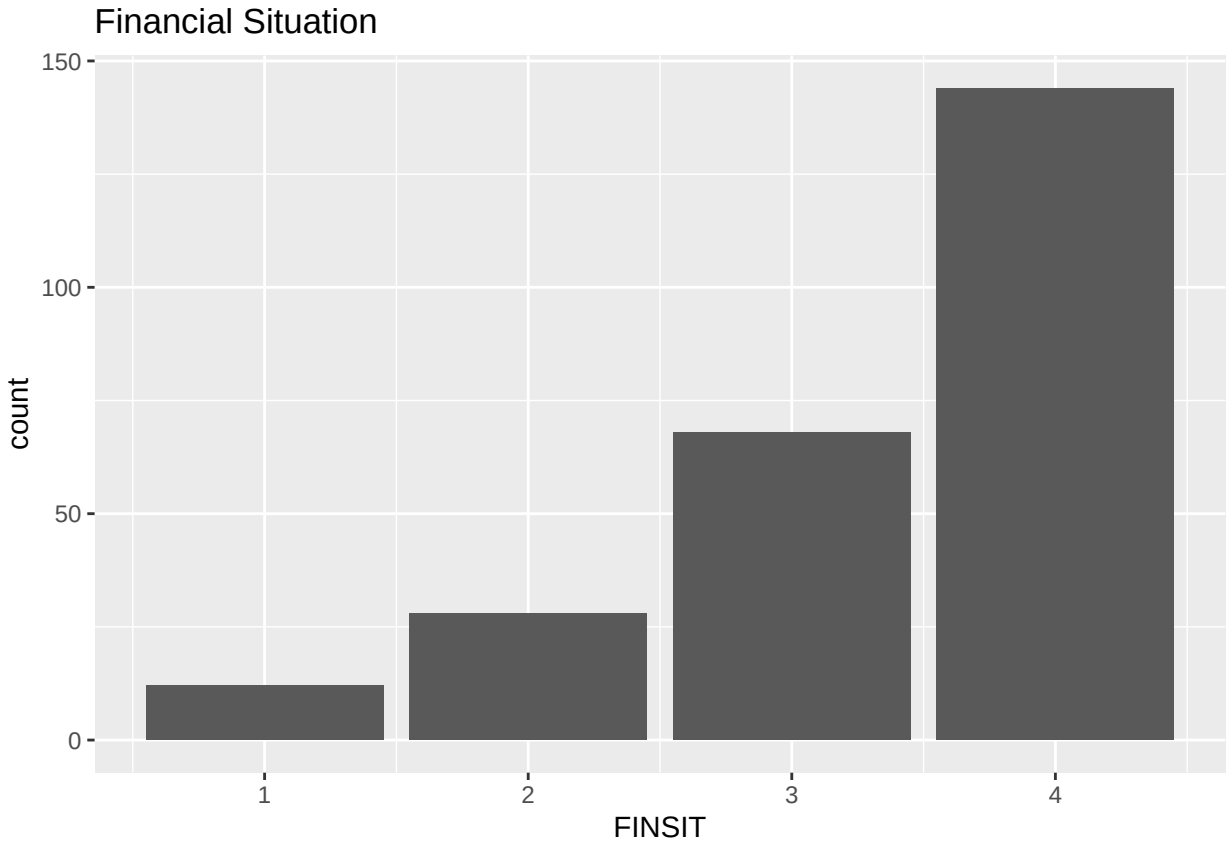
Drug dealing responses varied from 1, not an issue at all to 4, a very serious problem. An alarmingly high amount of neighborhoods, 75%, were described as having issues with drug dealing with 62% of these situations being described as severe.

Financial Situation of Household:

```
table(parHIV$FINSIT)
```

```
##
##   1   2   3   4
##  12  28  68 144
```

```
ggplot(parHIV,aes(x=FINSIT, fill=FINSIT))+geom_bar()+ggtitle("Financial Situation")
```




```
(12+28+68)/252
```

```
## [1] 0.4285714
```

```
(12+28)/68
```

```
## [1] 0.5882353
```

Financial situation responses varied from 1, very poor, struggling to survive, followed by 2, poor, barely paying the bills, to 3 having the necessities, and 4, comfortable living situation. Most children, 42%, lived in households that ranged from having the necessities to struggling to survive with 59% of that number being barely able to pay the bills and are very poor.

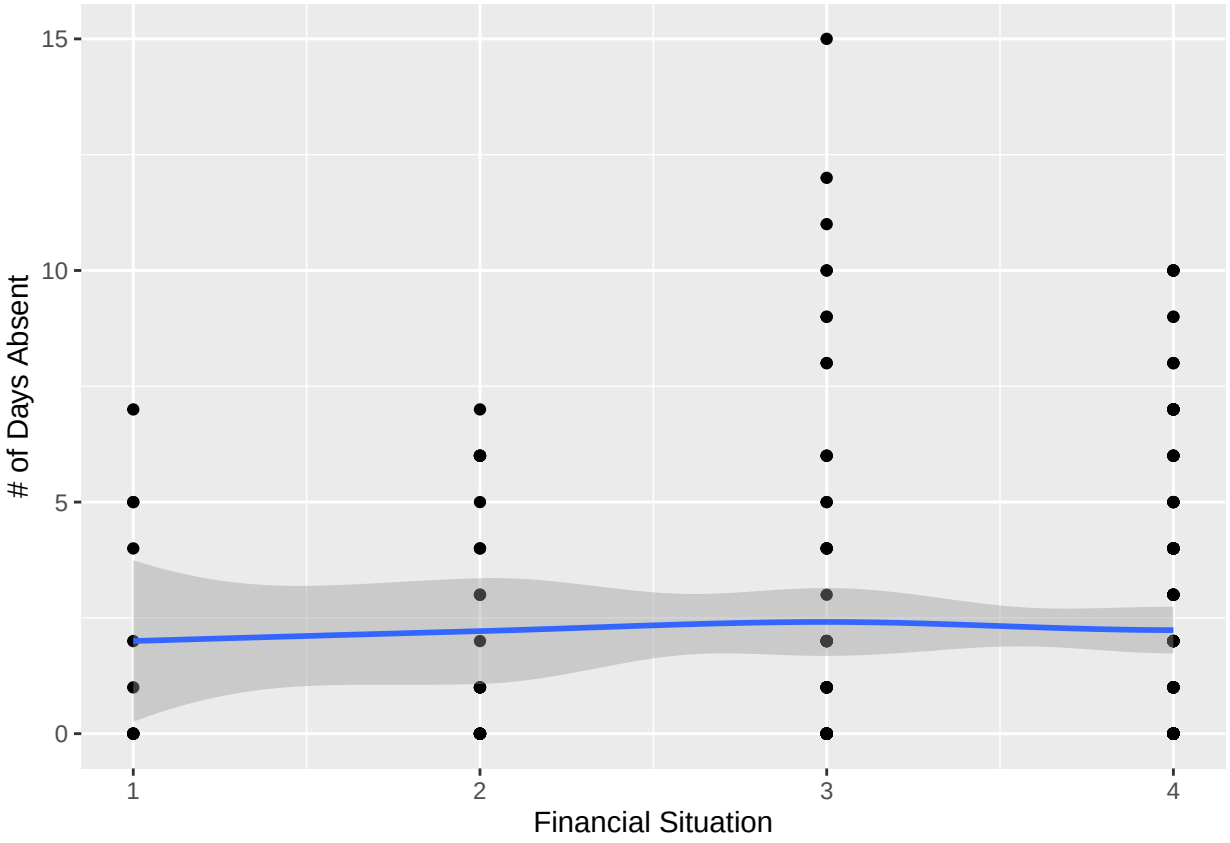
Bivariate Variable:

I wanted to see if there was a correlation between the financial situation of a child's household and their school attendance. To visualize this, a scatter plot was used.

```
table(parHIV$SCHOOL, parHIV$LIKESCH) %>% prop.table(1) %>% round(3)
```

```
##
##         1     2     3     4     5
##   1 0.212 0.242 0.121 0.061 0.364
##   2 0.237 0.315 0.292 0.059 0.096
```

```
ggplot(parHIV, aes(x=FINSIT,y=NHOOKEY)) +
geom_point() +
xlab("Financial Situation")+ylab("# of Days Absent")+geom_smooth(method="loess")
```



With a slope near 0, it is apparent there is little correlation between a child's financial situation and their school attendance.

Conclusion:

I cannot conclude anything based off this data set alone, but my hypothesis that children of HIV+ parents have, in general, tougher lives, is supported by the fact that around 50% of children have already had their first drink before the legal drinking age, 42% of children struggle with financial situations that range from struggling to survive to having the bare necessities, and 75% of the neighborhoods these children live in are described as having issues with drug dealing.

Additionally, while this limited data set can not directly show how one's environment and socioeconomic status is associated with having HIV and that there's little correlation between school attendance and financial situation, it does show us that a majority of HIV+ parents do live in less safe neighborhoods and many have struggles regarding their financial situation.

In essence, this data set has show us that HIV is correlated with having a tougher living situation for adults and the children of these adults alike.