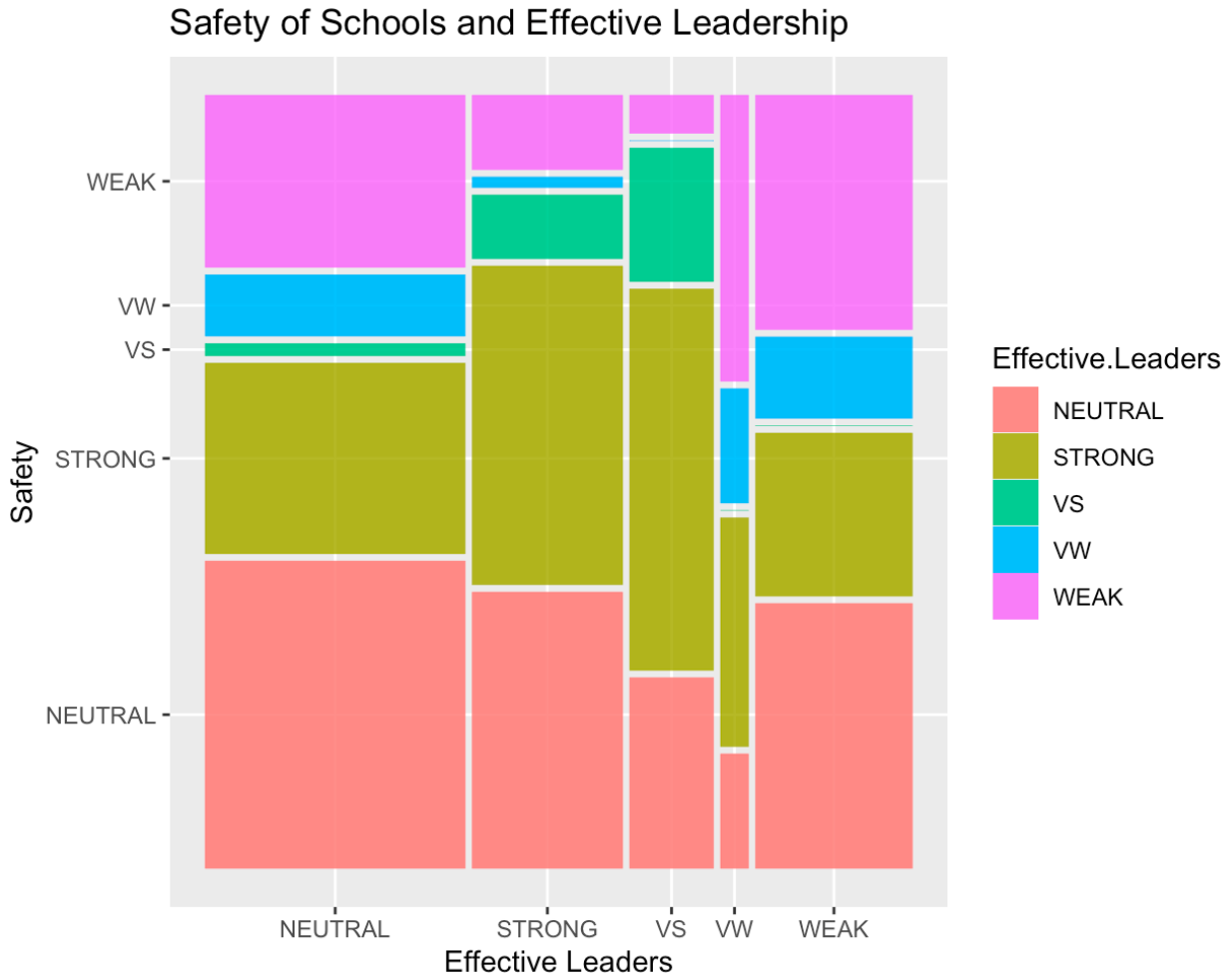
Homework 4

1. For my submission of a visualization for the second half of class I created a mosaic plot that compares the relationship between effective leadership and the safety ratings of schools. This plot demonstrates that schools with effective leaders have higher safety ratings or that schools with higher safety rating have more effective leaders. It is difficult to determine which variable is the cause and which is the effect, it is possible that neither are true and other outside factors cause both of the two variables.



R Code:
chi <- chicagopublicschoolsDataset

library(plyr)

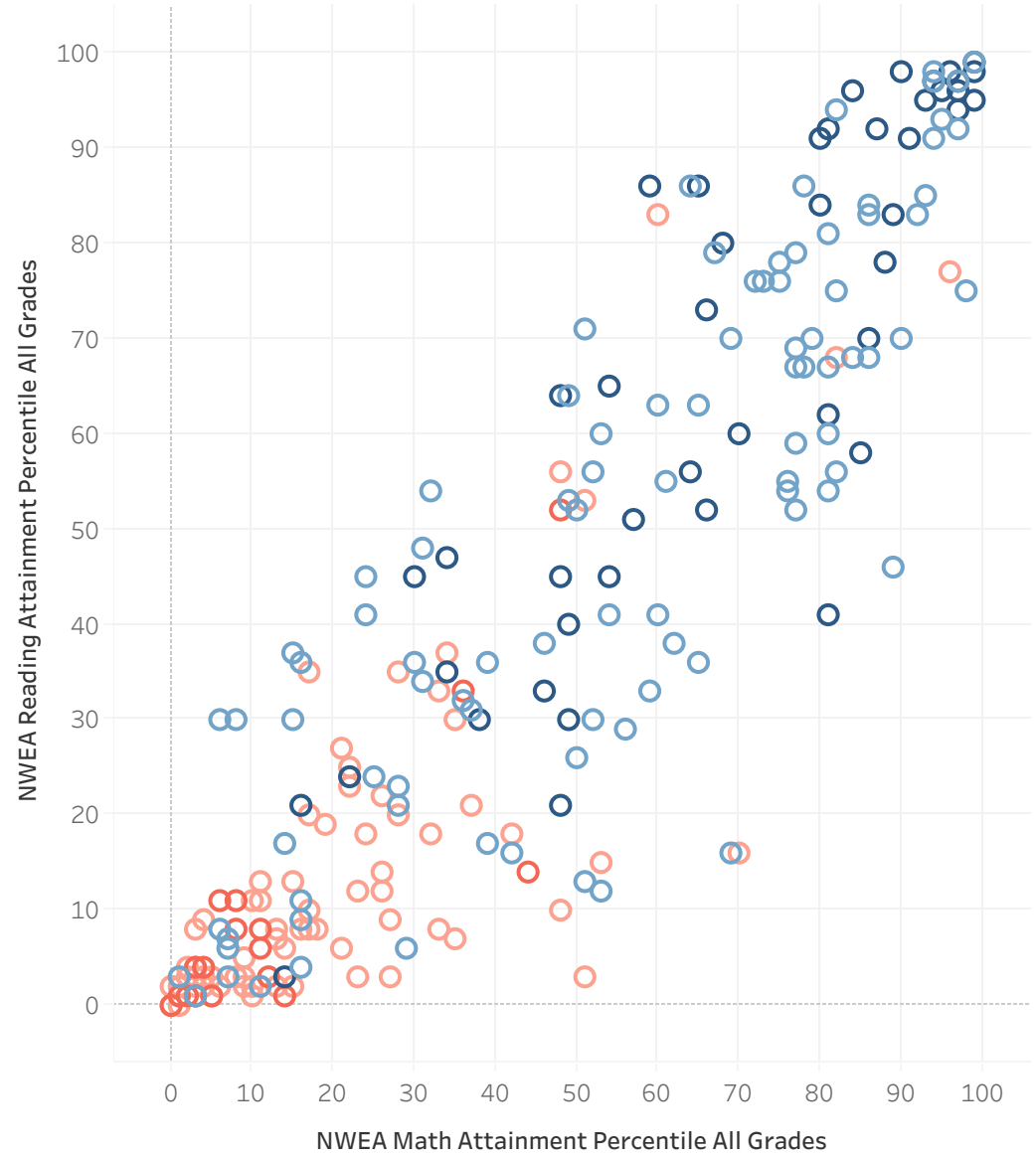
head(chi)

```
names(chi)

count(chi, 'ZIP.Code')

count(chi, 'Wards')

count(chi, 'CPS.Performance.Policy.Level')

count(chi, 'CPS.Performance.Policy.Status')

count(chi, 'Safe')

count(chi, 'Effective.Leaders')

chiTable <- matrix(chi=c())

mosaicplot(chi, main = 'Safe Schools Vs Policy Level',
        color = TRUE)
mosaicplot(~Safe + CPS.Performance.Policy.Level, data = chi,
        color = TRUE)

chi1 <- chi[!(chi$Safe=='NOT ENOUGH DATA'),]

count(chi1, 'Safe')

chi1 <- chi1[!(chi1$CPS.Performance.Policy.Level=='NOT ENOUGH DATA'),]

chi1 <- chi1[!(chi1$Effective.Leaders=='NOT ENOUGH DATA'),]

chi1 <- chi1[!(chi1$School.Community=='NOT ENOUGH DATA'),]


head(chi1)

count(chi1, 'CPS.Performance.Policy.Level')

chi1$Safe <- as.character(chi1$Safe)
chi1$Safe[chi1$Safe == 'VERY STRONG'] <- 'VS'

chi1$Safe <- as.character(chi1$Safe)
chi1$Safe[chi1$Safe == 'VERY WEAK'] <- 'VW'

chi1$Effective.Leaders <- as.character(chi1$Effective.Leaders)
```

```
chi1$Effective.Leaders[chi1$Effective.Leaders == 'VERY STRONG'] <- 'VS'

chi1$Effective.Leaders <- as.character(chi1$Effective.Leaders)
chi1$Effective.Leaders[chi1$Effective.Leaders == 'VERY WEAK'] <- 'VW'

count(chi1, 'Safe')
library(ggmosaic)
mosaic_examp <- ggplot(data = chi1) +
  geom_mosaic(aes(x = product(Safe), fill = Effective.Leaders)) +
  labs(y="Safety", x="Effective Leaders", title = "Safety of Schools and Effective Leadership")

mosaic_examp
```

2. For my visualization from the first half of class I compared reading and math test scores in a scatterplot in an attempt to see if there is a relationship between test scores and school safety, which is noted with color. Unsafe schools are marked with orange and red and safe schools are marked with blue. The first thing that I noticed is that there is a higher concentration of unsafe schools performing lower on tests and a high concentration of safe schools performing well on tests. Additionally, there appears to be a mostly linear relationship between reading and math scores, demonstrating that is a school performed well in one area, they are likely to perform well in other areas.

Math and Reading Attainment by Safety



NWEA Math Attainment Percentile All Grades vs. NWEA Reading Attainment Percentile All Grades. Color shows details about Safe. The view is filtered on Safe, which keeps STRONG, VERY STRONG, VERY WEAK and WEAK.

Safe

- STRONG
- VERY STRONG
- VERY WEAK
- WEAK