HW10

Tim McCormack

2/22/2021

```
##   gender race.ethnicity parental.level.of.education        lunch
## 1 female        group B           bachelor's degree     standard
## 2 female        group C                some college     standard
## 3 female        group B             master's degree     standard
## 4   male        group A          associate's degree free/reduced
## 5   male        group C                some college     standard
## 6 female        group B          associate's degree     standard
##   test.preparation.course math.score reading.score writing.score
## 1                    none         72            72            74
## 2               completed         69            90            88
## 3                    none         90            95            93
## 4                    none         47            57            44
## 5                    none         76            78            75
## 6                    none         71            83            78
```

Question 1

```
df2 <- df %>% group_by(parental.level.of.education) %>% summarise(count = n())
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
df3 <- arrange(df2, desc(count) )
```

```
e1 <- df$parental.level.of.education

table(e1)
```

```
## e1
## associate's degree  bachelor's degree        high school    master's degree
##                222                118                196                 59
##       some college   some high school
##                226                179
```

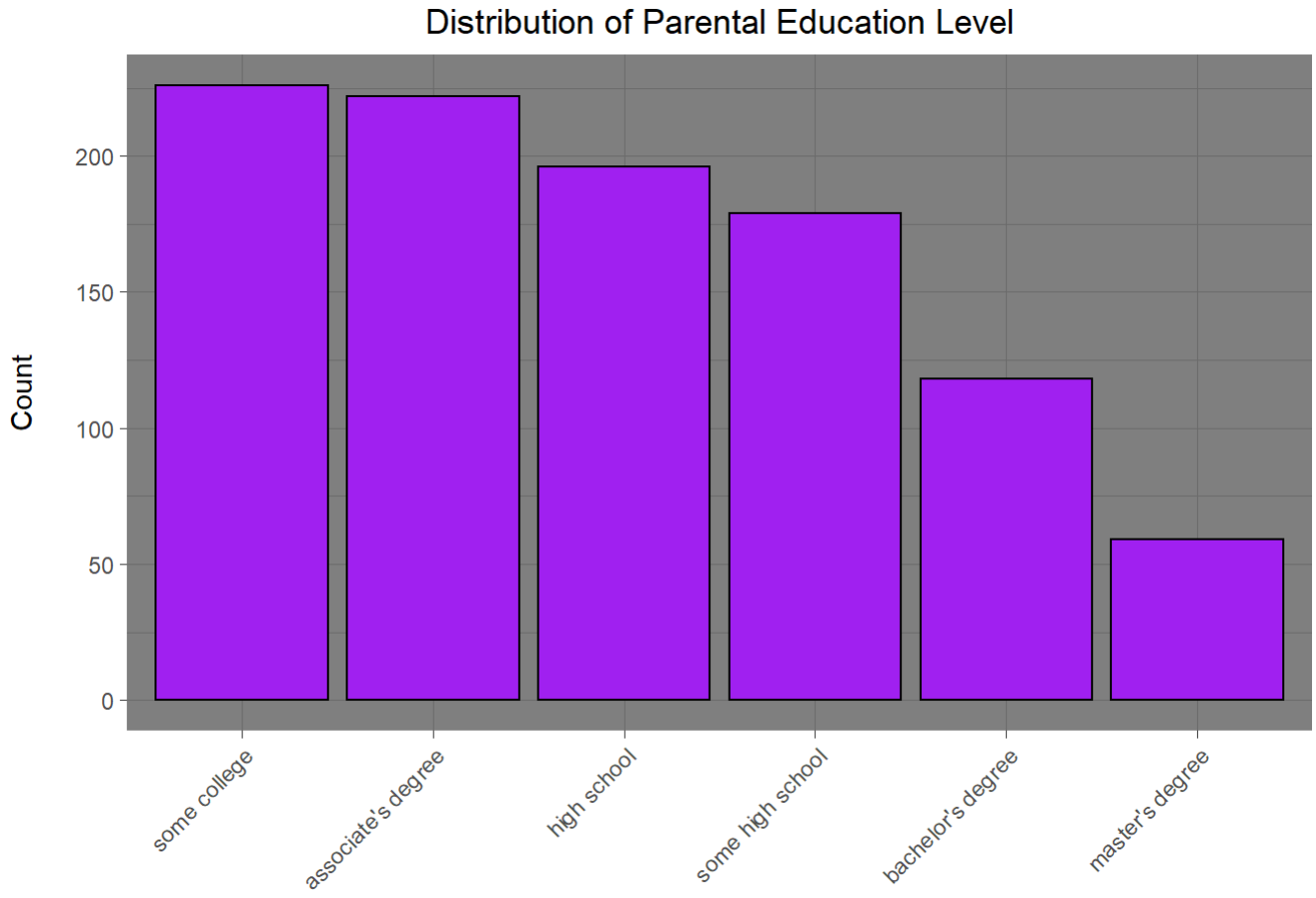
```
e2 <- fct_relevel(e1, c("some college", "associate's degree", "high school", "some high school",
"bachelor's degree", "master's degree"))

levels(e2)
```

```
## [1] "some college"       "associate's degree" "high school"
## [4] "some high school"   "bachelor's degree"  "master's degree"
```

```
df5 <- data.frame(e2)
df5 <- df5 %>% group_by(e2) %>% summarise(count = n())
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
df5 <- df5 %>% mutate(deg = e2) %>% select(-e2)

ggplot(data = df5, aes(x = deg, y = count)) + geom_bar(stat = "identity", fill = "purple", colou
r = "black") +  theme_dark()+theme(axis.text.x = element_text(angle = 45, vjust = 1, hjust=1)) +
xlab("Education Level") + ylab("Count\n") + ggtitle("Distribution of Parental Education Level")
 +  theme(plot.title = element_text(hjust = 0.5)) + xlab("")
```




```
df6 <- df %>% group_by(parental.level.of.education, lunch) %>% summarise(count = n())

unique(df$lunch)
```

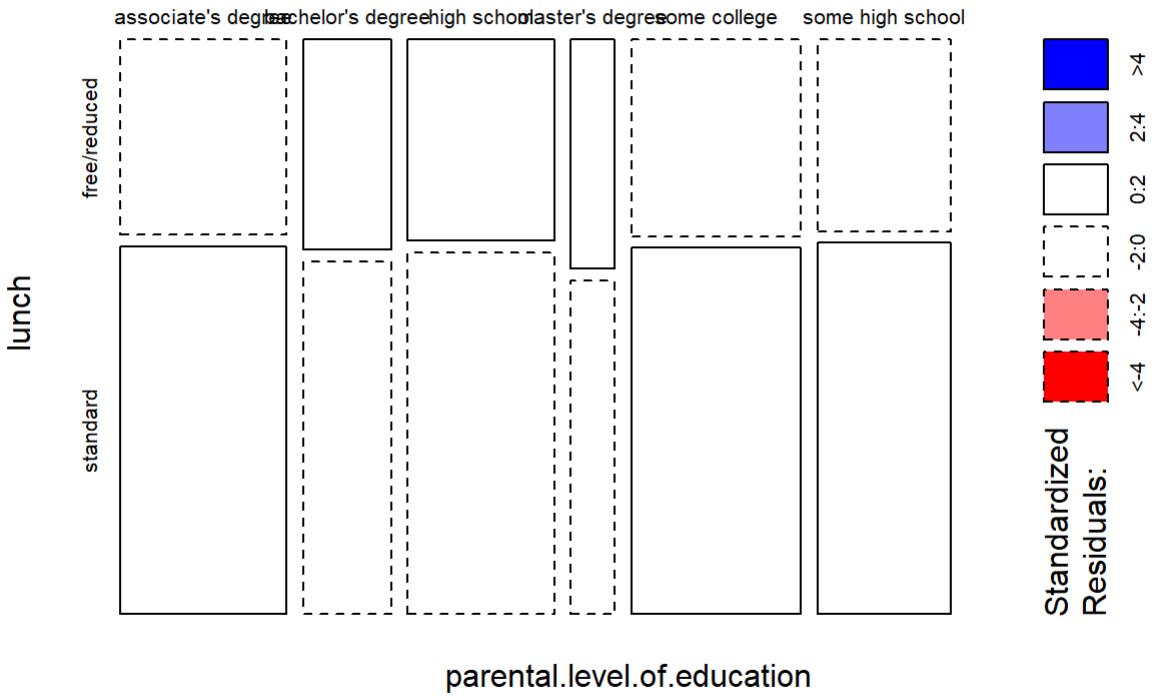
```
## [1] "standard"     "free/reduced"
```

```
degrees <- c(rep("associates", 222) , rep("bachelors", 118), rep("high school", 196), rep("maste
rs", 59), rep("some college", 226), rep("some high school", 179))

lunches <- c(rep("free", 77) , rep("standard", 145), rep("free", 44), rep("standard", 74), rep(
"free", 70), rep("standard", 126), rep("free", 24), rep("standard", 35), rep("free", 79), rep("s
tandard", 147), rep("free", 61), rep("standard", 118))

df8 <- data.frame(degrees, lunches)




chisq.test(table(df8) )
```

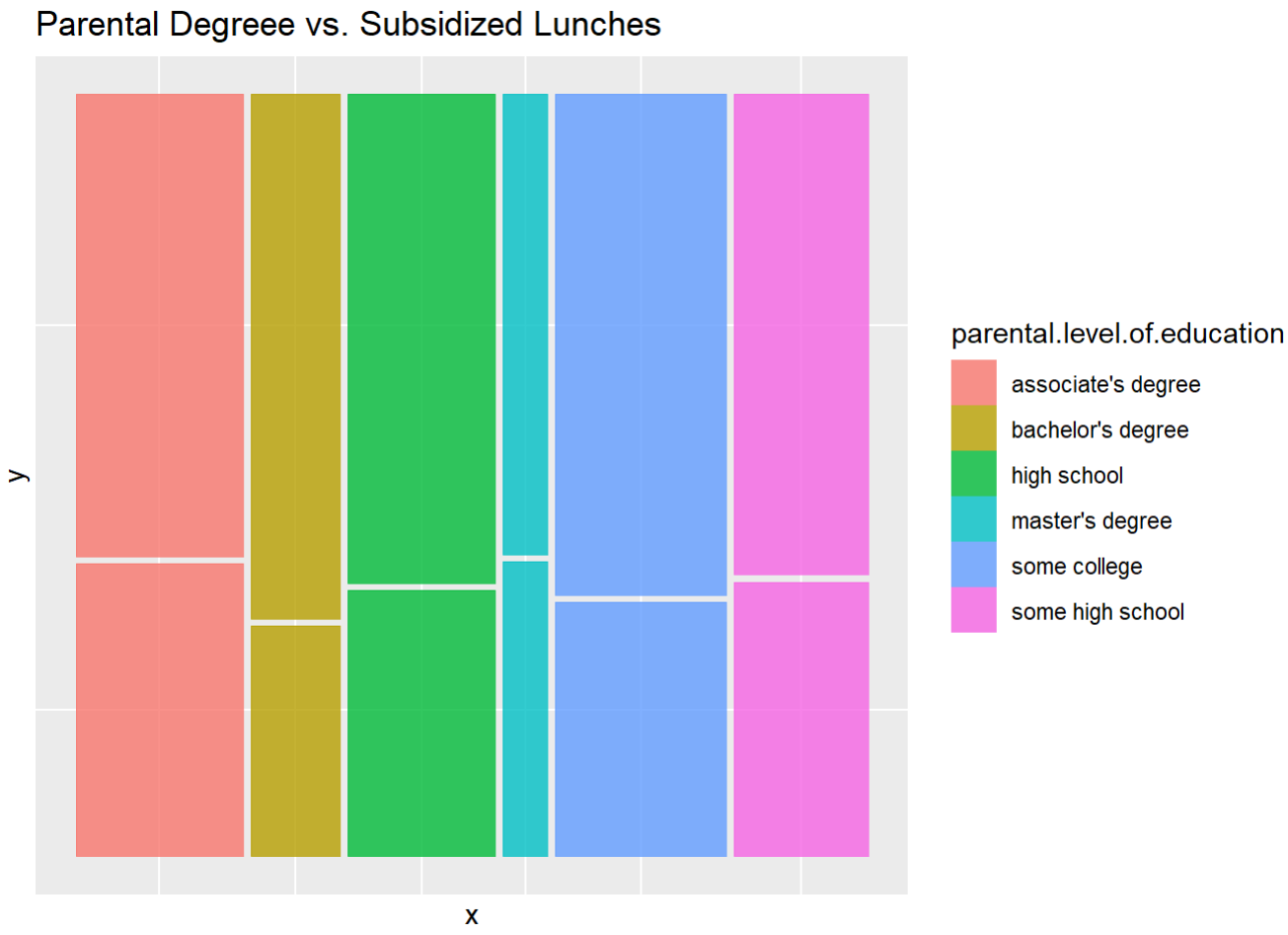
```
##
##  Pearson's Chi-squared test
##
## data:  table(df8)
## X-squared = 1.1113, df = 5, p-value = 0.9531
```

```
mosaicplot(~ parental.level.of.education + lunch, data = df, main = "Parental Degree vs. Subsidi
zed Lunch", shade = TRUE, direction = "v", rot_labels=c(90,90,0,0))
```

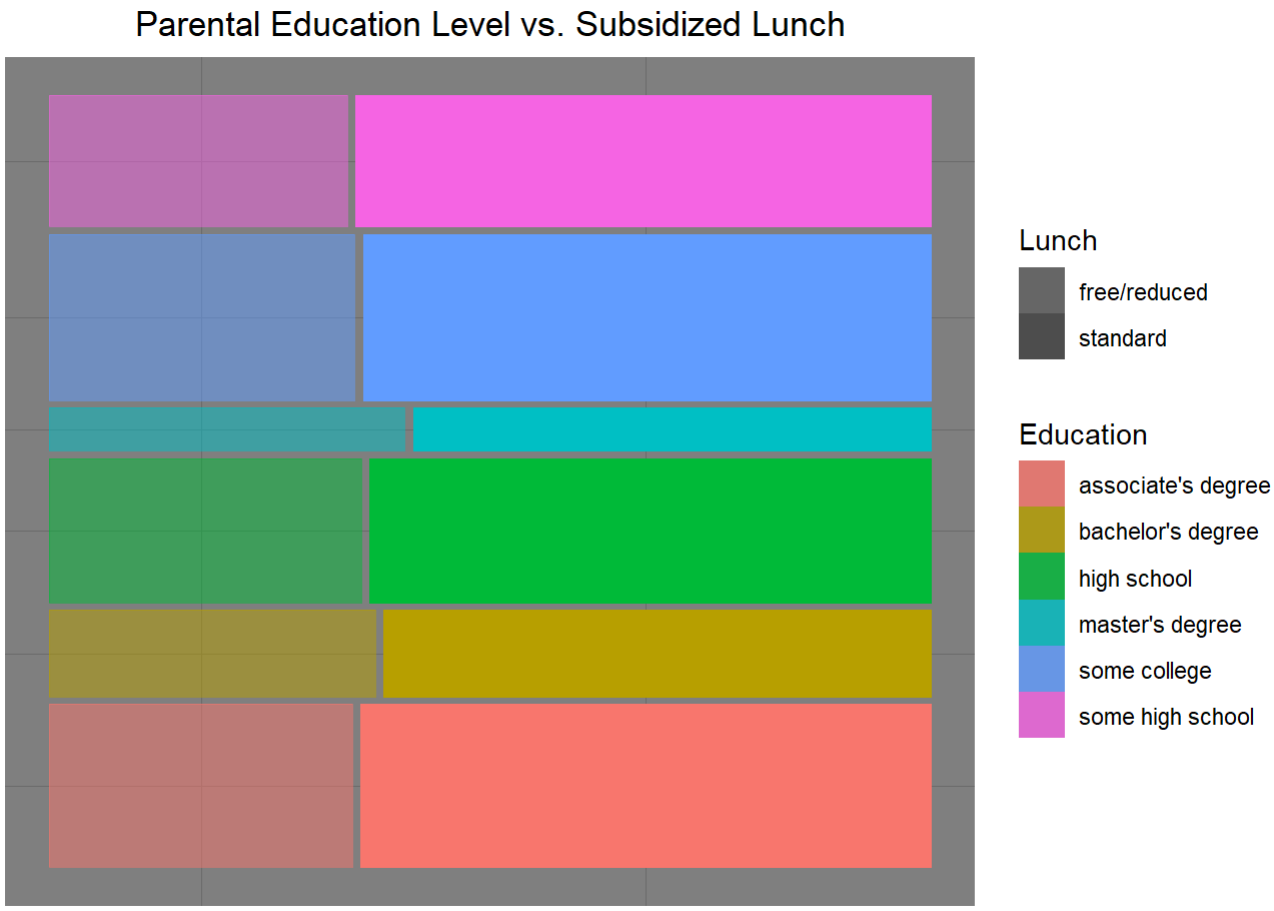
```
## Warning: In mosaicplot.default(table(mf), main = main, ...) :
##  extra arguments 'direction', 'rot_labels' will be disregarded
```



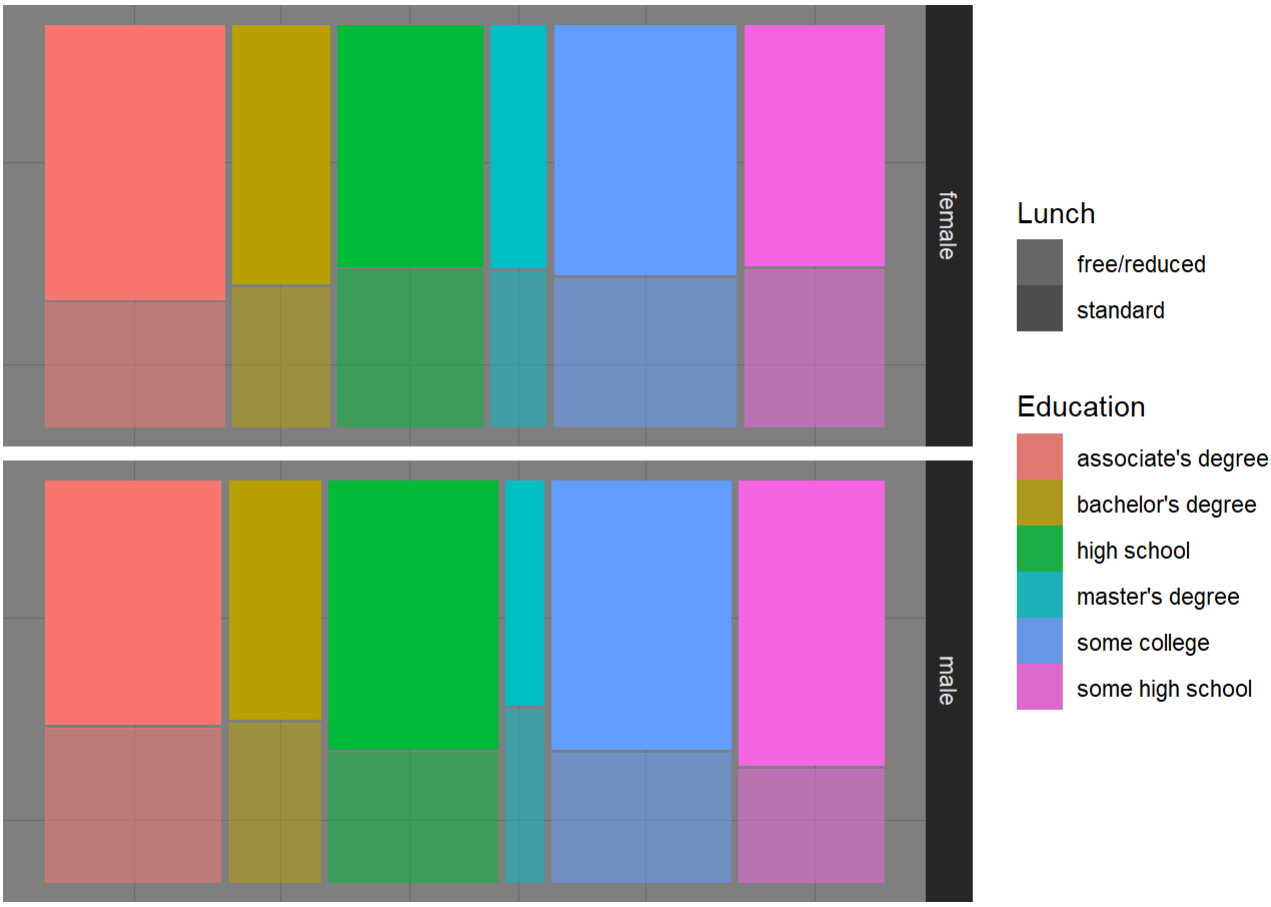

```
ggplot(data = df) + geom_mosaic(aes(x = product( lunches, parental.level.of.education), fill = p
arental.level.of.education), na.rm = TRUE) +
  labs( title='Parental Degreee vs. Subsidized Lunches')
```



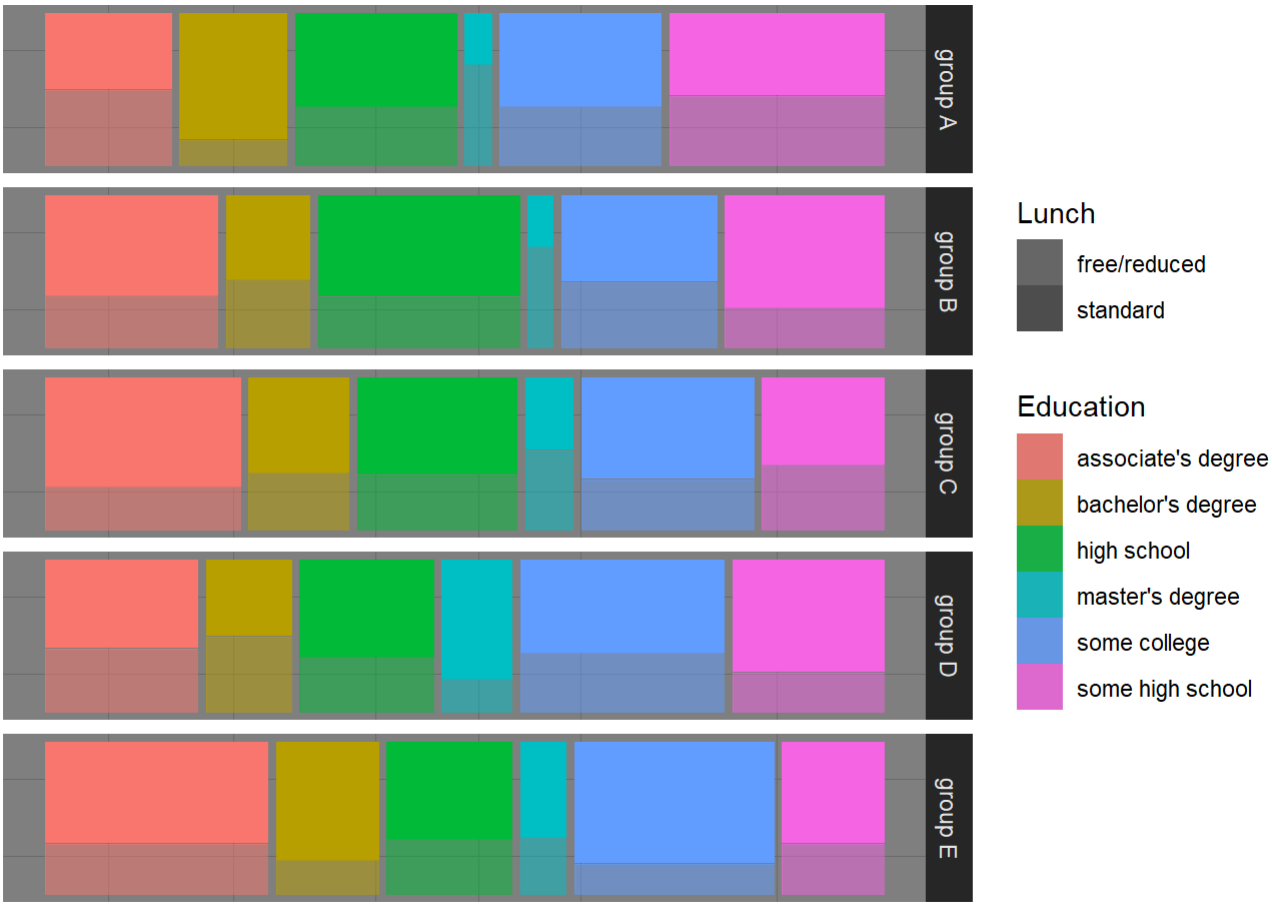

```
df <- df %>% mutate(Education = parental.level.of.education, Lunch = lunch) %>%select(Education, Lunch, gender, race.ethnicity)



ggplot(data = df) + geom_mosaic(aes(x = product(Lunch, Education),  fill = Education, alpha = Lunch)) + labs(x = "", y = "") + scale_alpha_manual(values = c(0.5,1)) + coord_flip() + theme_dark() + ggtitle("Parental Education Level vs. Subsidized Lunch") +  theme(plot.title = element_text(hjust = 0.5))
```




```
ggplot(data = df) + geom_mosaic(aes(x = product(Lunch, Education),  fill = Education, alpha = Lu
nch)) + labs(x = "", y = "") + facet_grid(gender~.) + theme_dark() + scale_alpha_manual(values =
c(0.5,1))
```




```
ggplot(data = df) + geom_mosaic(aes(x = product(Lunch, Education),  fill = Education, alpha = Lu
nch)) + labs(x = "", y = "") + facet_grid(race.ethnicity~.) + theme_dark() + scale_alpha_manual
(values = c(0.5,1))
```


group A
group B
group C
group D
group E
Lunch
free/reduced
standard
Education
associate's degree
bachelor's degree
high school
master's degree
some college
some high school