

final project

Kayla Saechao

2023-09-22

Introduction: In this EDA, we will explore the “High School and Beyond” dataset, focusing on three variables: Math score, Gender, and race. We are trying to find if there is a relationship between students GPA, their gender and participation in extracurricular activities.

UNIVARITE EXPLORATION

```
hsb2 <- read.delim("/Users/kayla/Downloads/hs and beyond.txt", sep="\t")
dim(hsb2)
```

```
## [1] 200 11
```

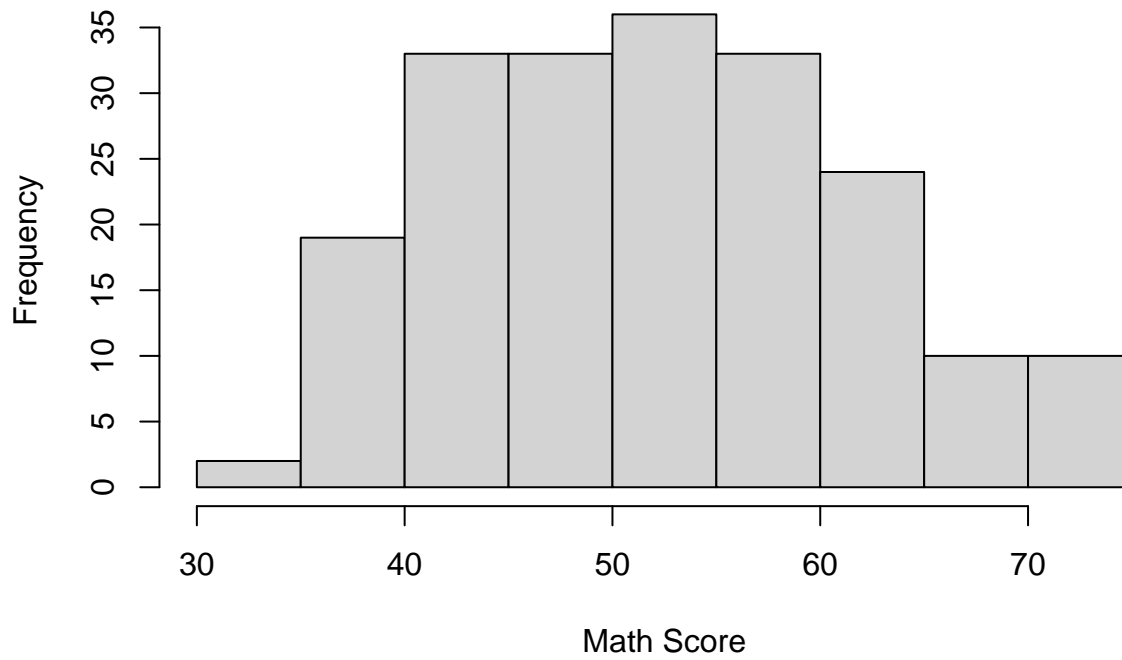
Now we will be finding the Math scores:

```
summary(hsb2$math)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  33.00  45.00   52.00   52.65  59.00   75.00
```

```
hist(hsb2$math, main="Math Score Distribution", xlab="Math Score")
```

Math Score Distribution



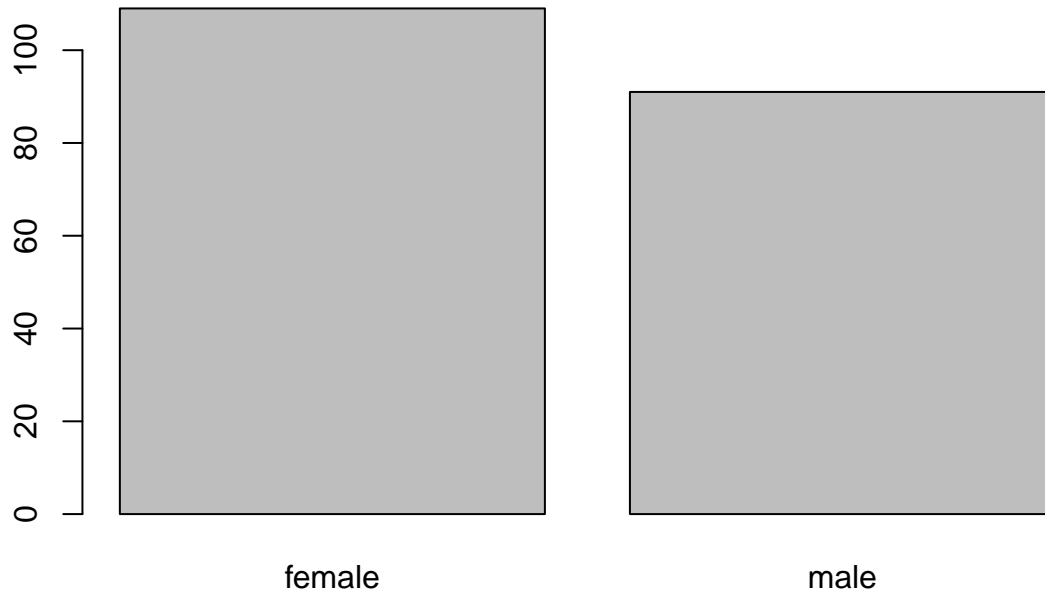
Now the Gender distribution:

```
table(hsb2$gender)
```

```
##  
## female  male  
##   109    91
```

```
barplot(table(hsb2$gender), main="Gender Distribution")
```

Gender Distribution



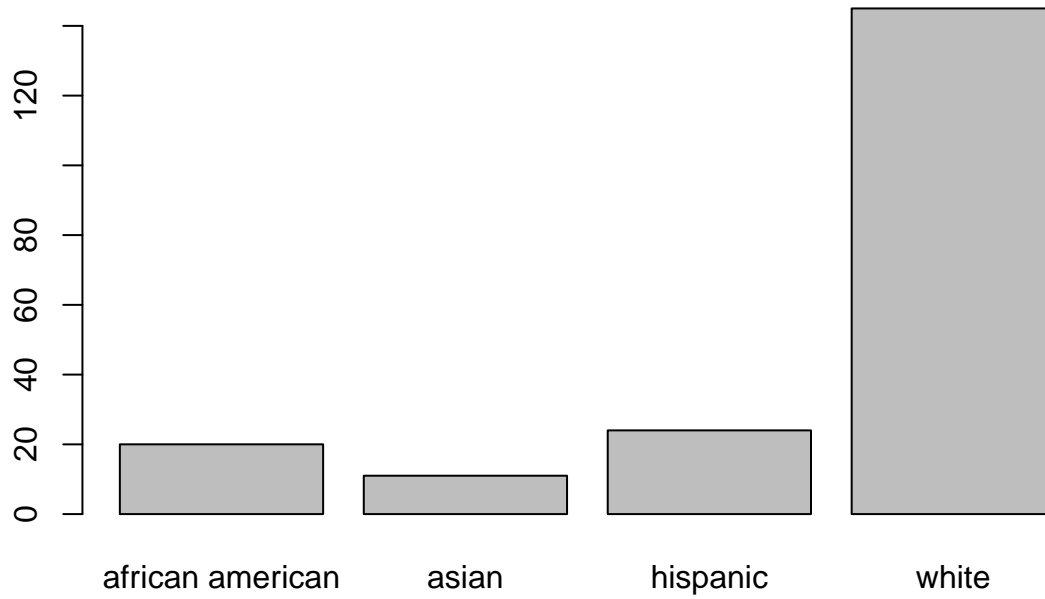
And lastly we do the race:

```
table(hsb2$race)
```

```
##  
## african american      asian      hispanic      white  
##           20           11           24           145
```

```
barplot(table(hsb2$race), main="Race Distribution")
```

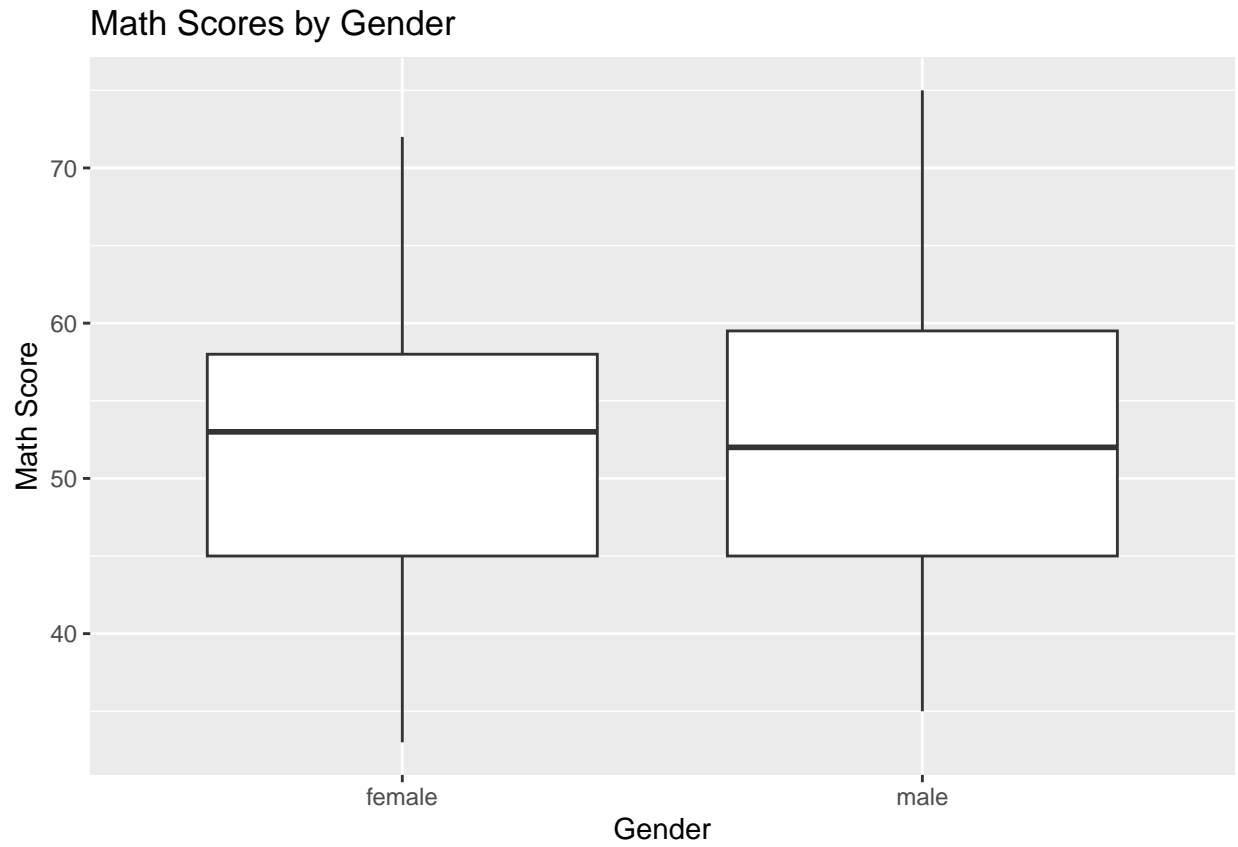
Race Distribution



Bivariate Exploration:

1. In this we will compare the math scores by gender using a boxplot:

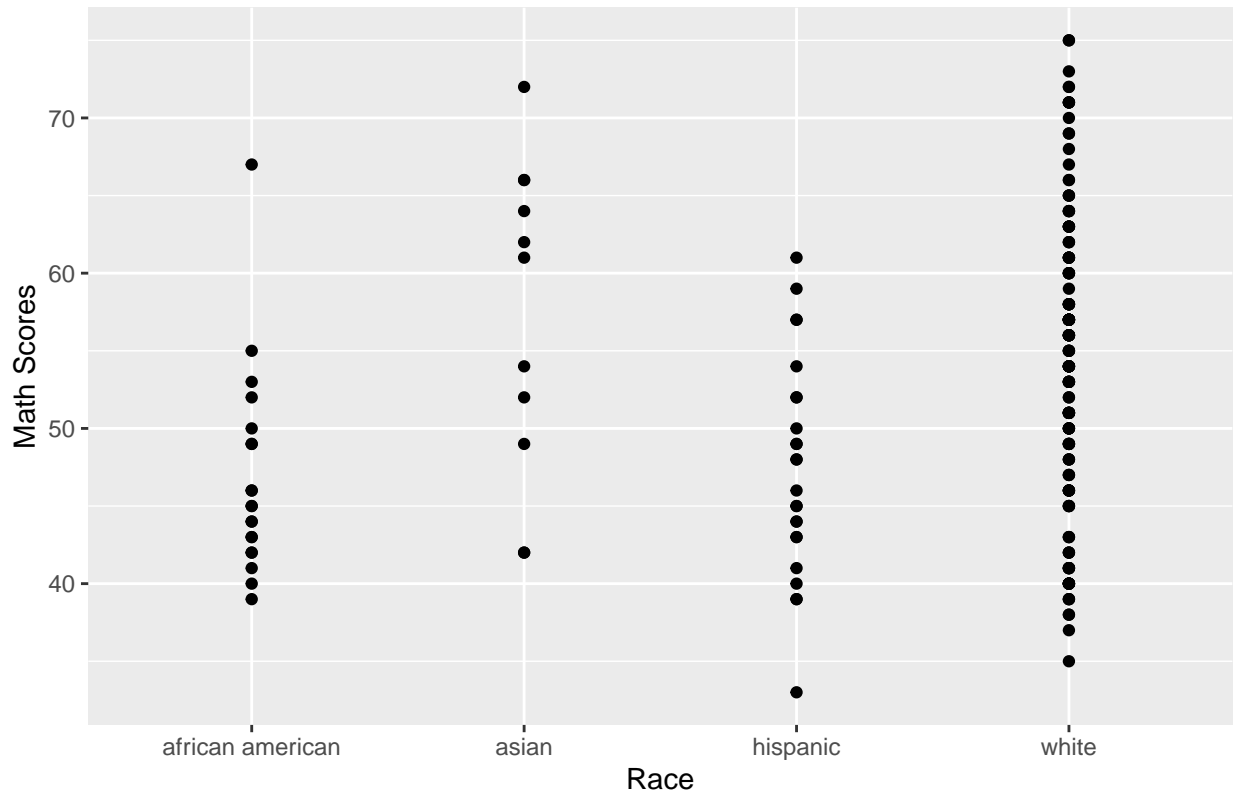
```
ggplot(hsb2, aes(x=gender, y=math))+ geom_boxplot()+ labs(title="Math Scores by Gender", x="Gender", y="Math Score")
```



2. Now we will also compare the math scores by race using a scatter plot:

```
ggplot(hsb2, aes(x=race, y=math))+geom_point()+labs(title="Math Scores by Race", x="Race", y="Math Scores")
```

Math Scores by Race



Conclusion: In conclusion, we saw that white people had better math scores but there were also a good amount of white people than any other race. We also saw that there are more females than males but males had a little bit more of a better score than females. So all in all, white people had more of higher scores because there were more of them than any other race, and the the males had higher scores than females.