sta_100_hw_2

Tiffany Chan

2023-10-10

R Markdown

```
getwd()
```

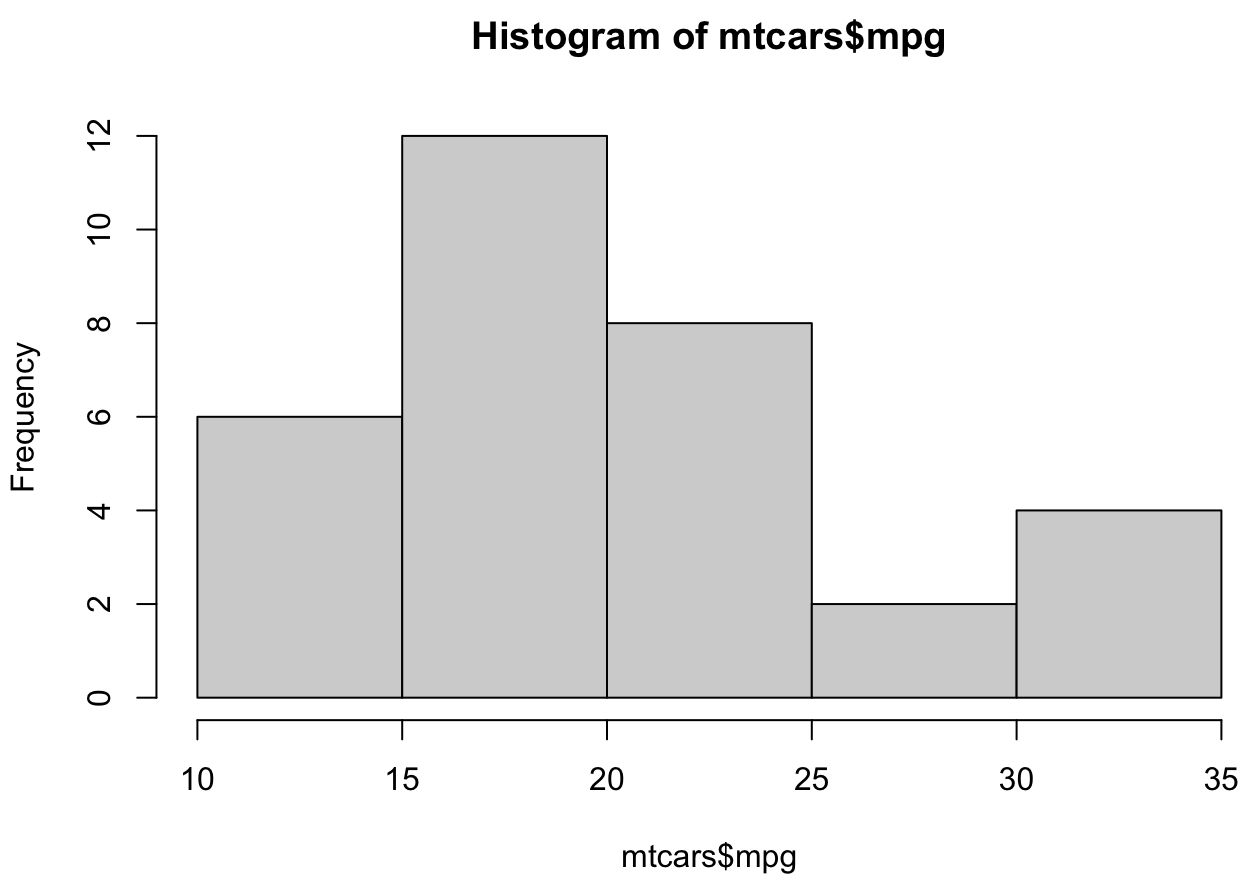
```
## [1] "/Users/tiffanysmacbookpro/Desktop/sta100F2023_tchan"
```

```
gender = rep(c("female","male"), c(1835, 2691))
admitted = rep(c("yes", "no", "yes", "no"), c(557, 1278, 1198, 1493))
dept = rep(c("A", "B", "C", "D", "E", "F", "A", "B", "C", "D", "E", "F"),
           c(89, 17, 202, 131, 94, 24, 19, 8, 391, 244, 299, 317))
dept2 = rep(c("A", "B", "C", "D", "E", "F", "A", "B", "C", "D", "E", "F"),
            c(512, 353, 120, 138, 53, 22, 313, 207, 205, 279, 138, 351))
department = c(dept, dept2)
ucb = data.frame(gender, admitted,department)
rm(gender, admitted, dept, dept2, department)
```

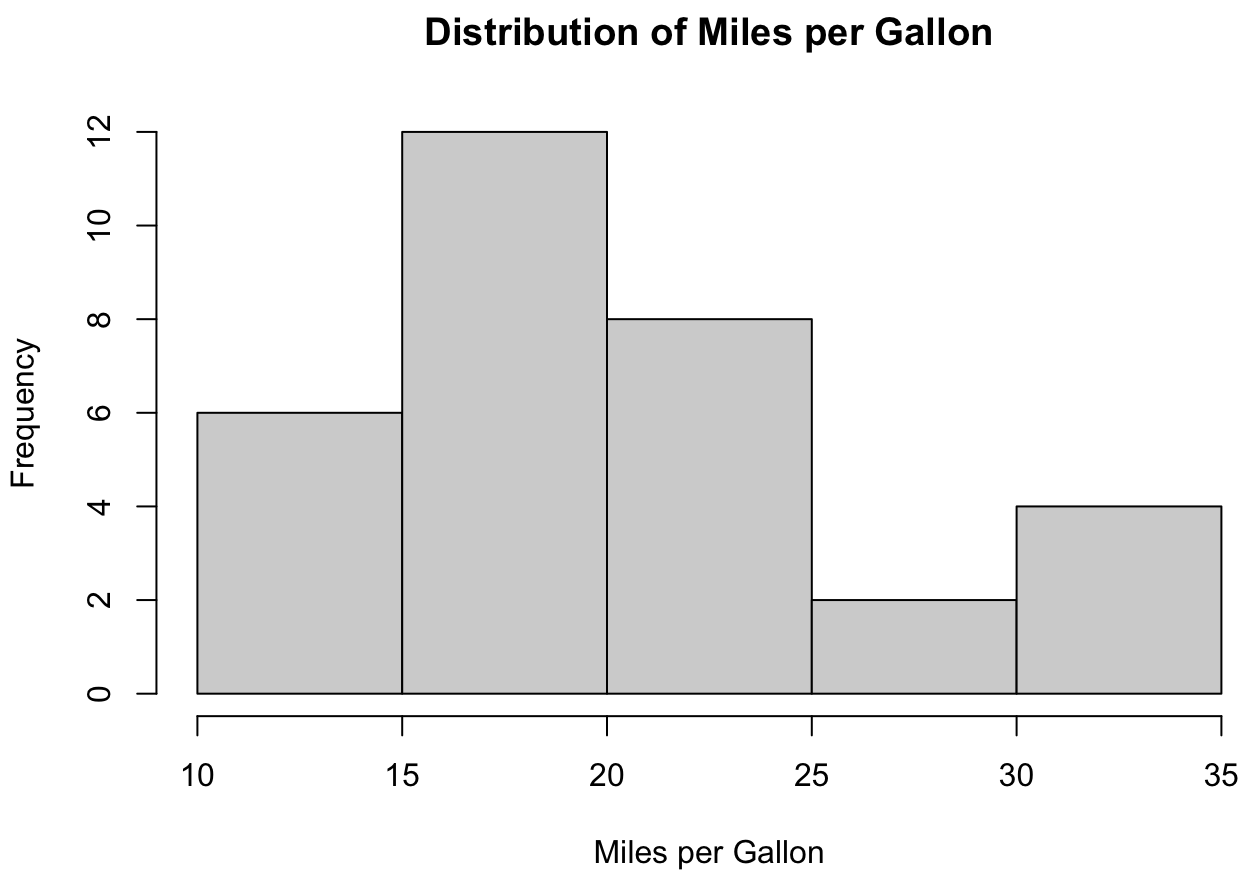
```
head(ucb)
```

```
##   gender admitted department
## 1 female      yes          A
## 2 female      yes          A
## 3 female      yes          A
## 4 female      yes          A
## 5 female      yes          A
## 6 female      yes          A
```

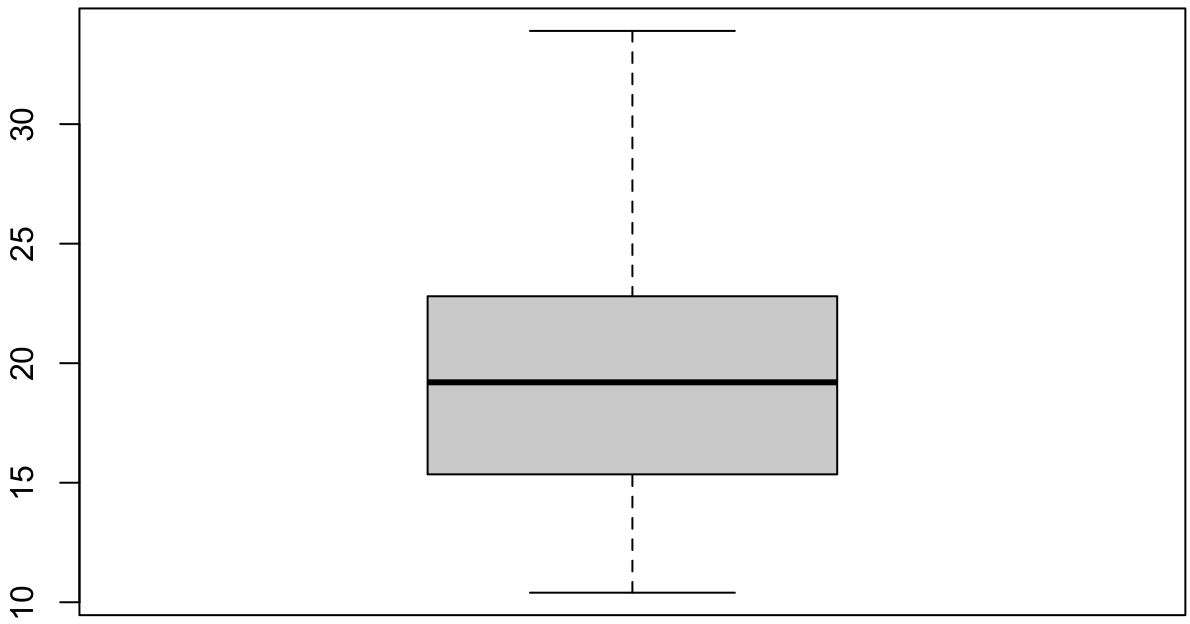
```
hist(mtcars$mpg)
```



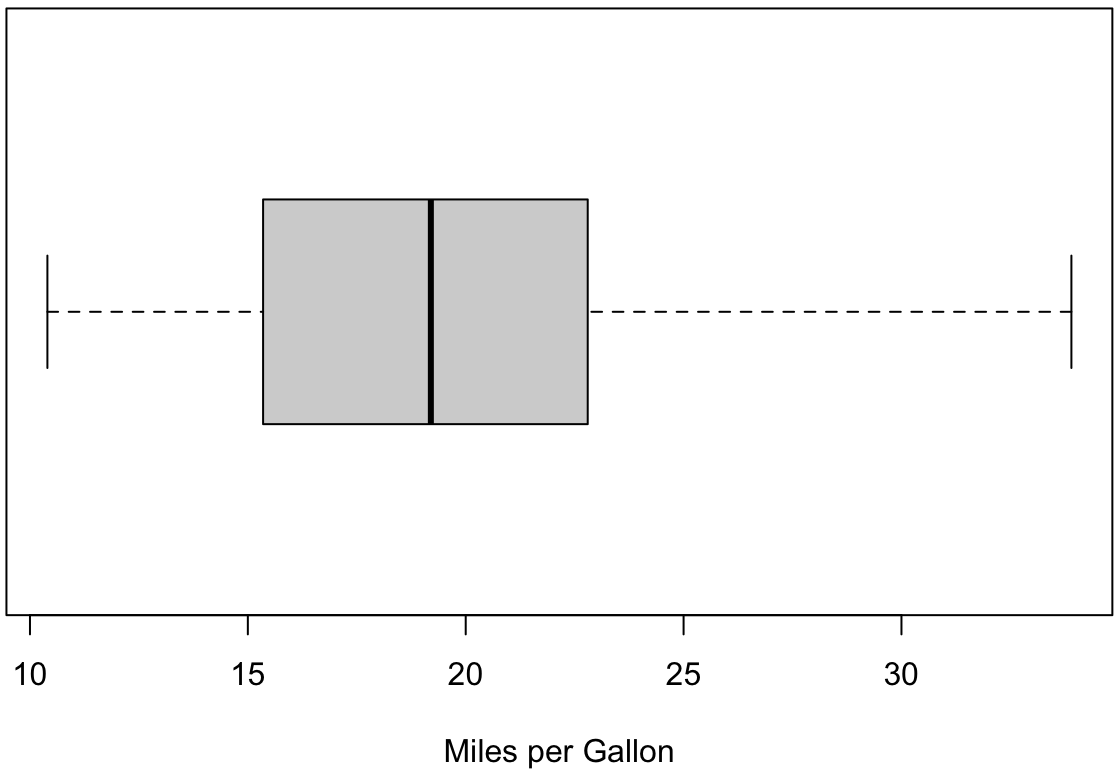

```
hist(mtcars$mpg, xlab = "Miles per Gallon", main = "Distribution of Miles per Gallon" )
```




```
boxplot(mtcars$mpg)
```




```
boxplot(mtcars$mpg, xlab="Miles per Gallon", main="Distribution of Miles per Gallon", ho
rizontal = TRUE)
```



You will be working with the dataset colors.csv, which has the following columns: Column 1: Eye: The eye color of the subject Column 2: Sex: The hair color of the subject Column 3: GPA: The college GPA of the subject

```
colors<-read.csv("/Users/tiffanysmacbookpro/Desktop/sta100F2023_tchan/data/colors.csv")
```

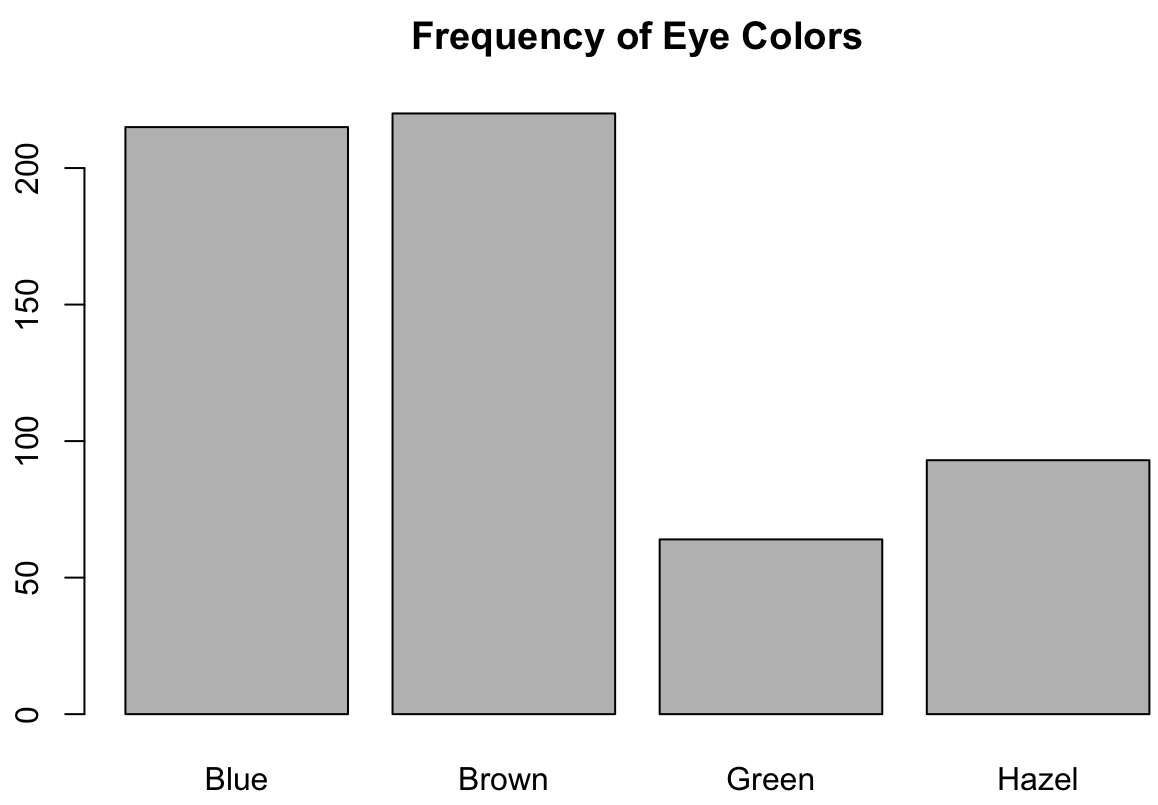
```
names(colors)
```

```
## [1] "Eye" "Sex" "GPA"
```

```
#(a) Plot a barplot of the eye color of the subjects. What
#color is the least common?
```

```
eye.table=table(colors$Eye)
```

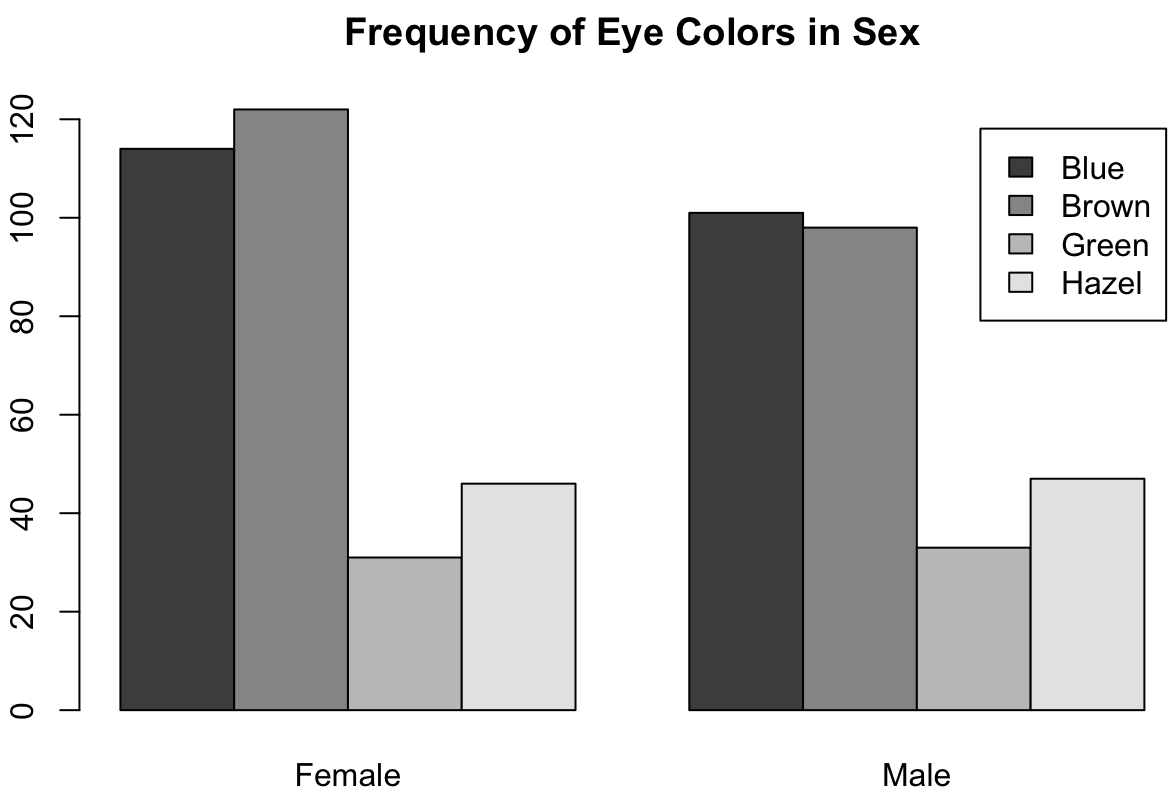
```
barplot(eye.table, main="Frequency of Eye Colors")
```




```
#Least common eye color is green eyes.
```

```
#(b) Plot a side-by-side barplot of the subjects, using sex
#and eye color. Comparing men and women, who
#has a higher probability of brown eyes? Be sure to
#choose the grouping of the bars that make it easier
#to read and interpret the result.
```

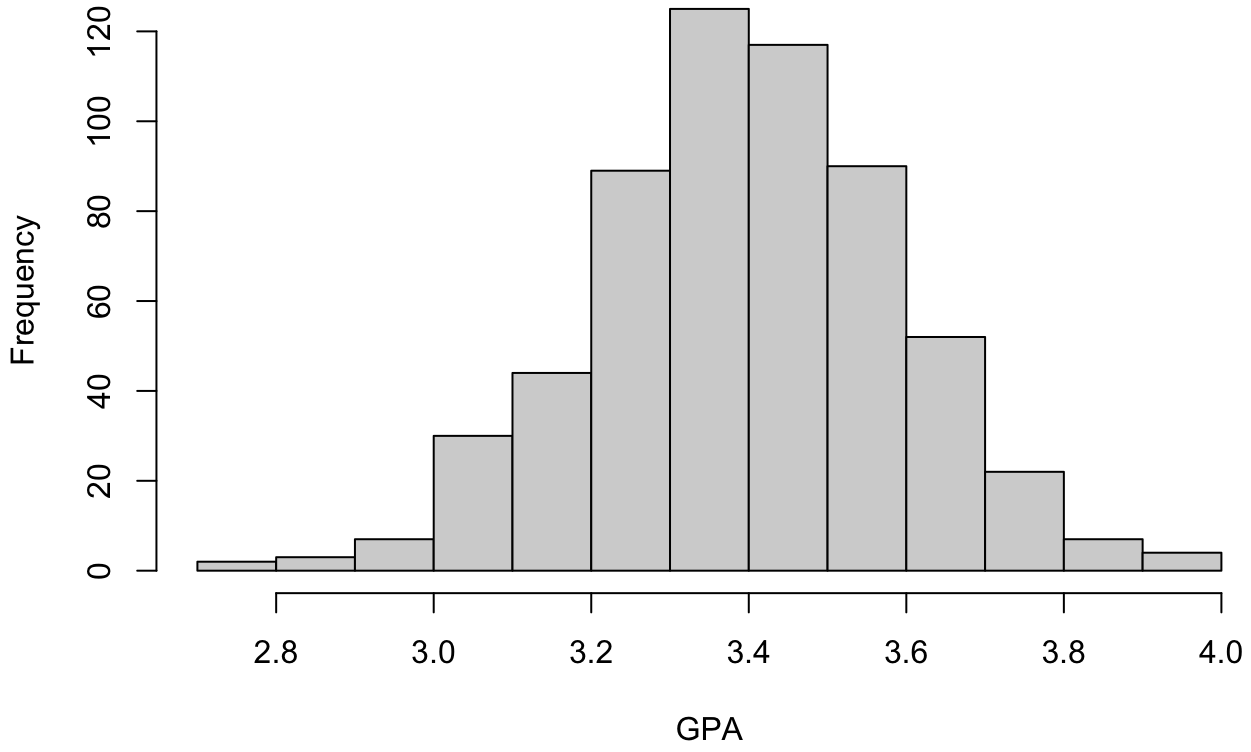
```
eye.sex.table=table(colors$Eye,colors$Sex)
barplot(eye.sex.table, main="Frequency of Eye Colors in Sex", beside = TRUE, legend=row.
names(eye.sex.table))
```




```
#Females had a higher probability of having brown eyes
```

```
#(c) Plot a histogram of GPA. What is the most common
#interval of GPA?
```

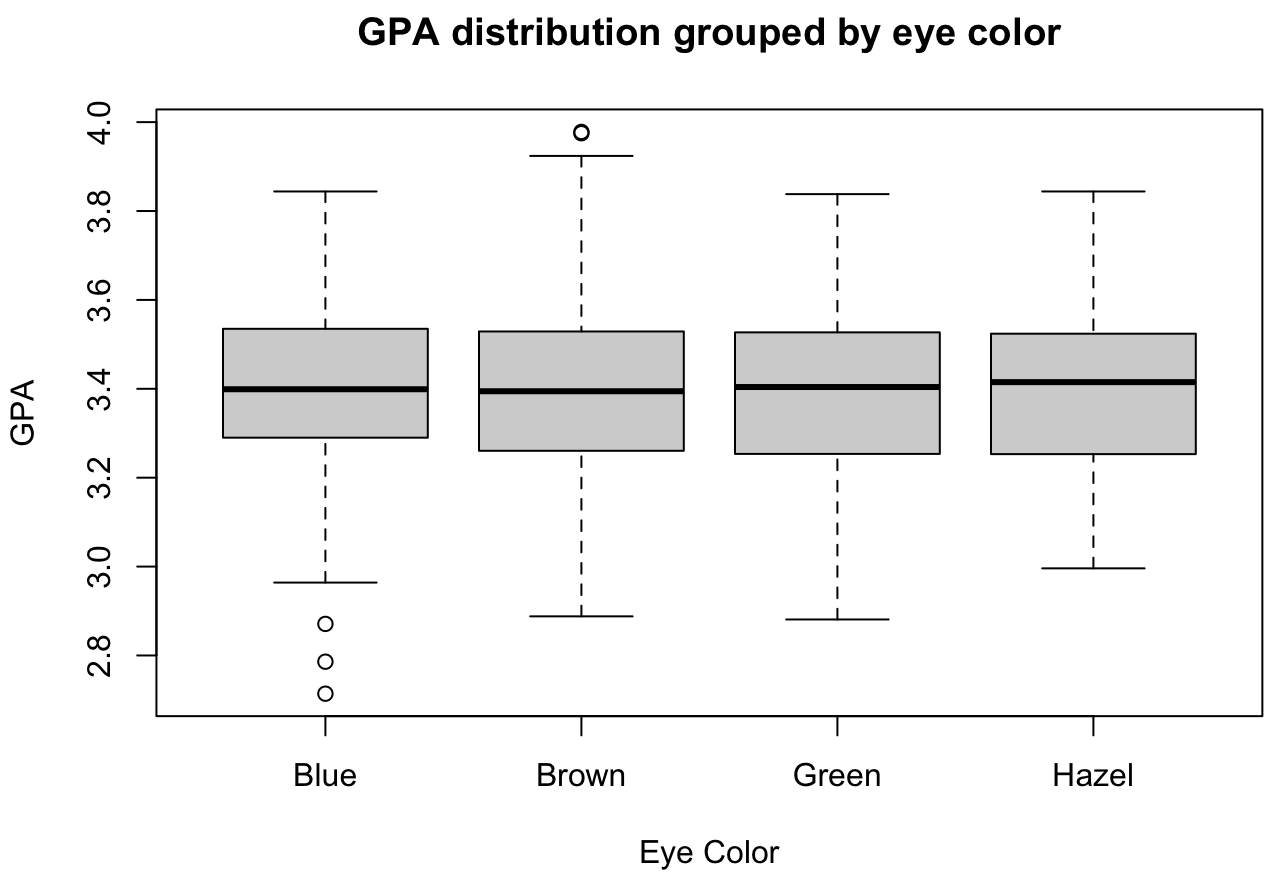
```
hist(colors$GPA, main="Distribution of GPA", xlab = "GPA")
```




```
#Most common interval ranged from 3.3 to 3.4
```

```
#(d) Plot a side-by-side boxplot of GPA by eye color.
#Which eye color has the highest minimum?
```

```
boxplot(GPA ~ Eye, data=colors, xlab = "Eye Color", main = "GPA distribution grouped by
eye color")
```




```
#Hazel eye color had the highest minimum.
```

```
#(e) Refer to the previous side-by-side boxplot of GPA
#by eye color. Which eye color has the highest 25th
#percentile?
```

```
#Blue eyes had the highest 25th percentile.
```

```
#(f) Refer to the previous side-by-side boxplot of GPA
#by eye color. Which eye color has the most outliers?
```

```
#Blue eyes had the most outliers.
```