KJPipho_PS_1

Krista Pipho

September 1, 2018

RStudio Fall 2018

Problem Set 1

1. HairEyeColor

How many people are included in this data set?

```
HEC <- HairEyeColor
sum(HEC)
```

```
## [1] 592
```

What are the male/female proportions of these 592 people?

```
prop.table(apply(HEC, c(3), sum))
```

```
##      Male    Female
## 0.4712838 0.5287162
```

What proportions are the hair and eye colors present in?

```
"Hair"
```

```
## [1] "Hair"
```

```
prop.table(apply(HEC, c(1), sum))
```

```
##     Black     Brown       Red     Blond
## 0.1824324 0.4831081 0.1199324 0.2145270
```

```
"Eye"
```

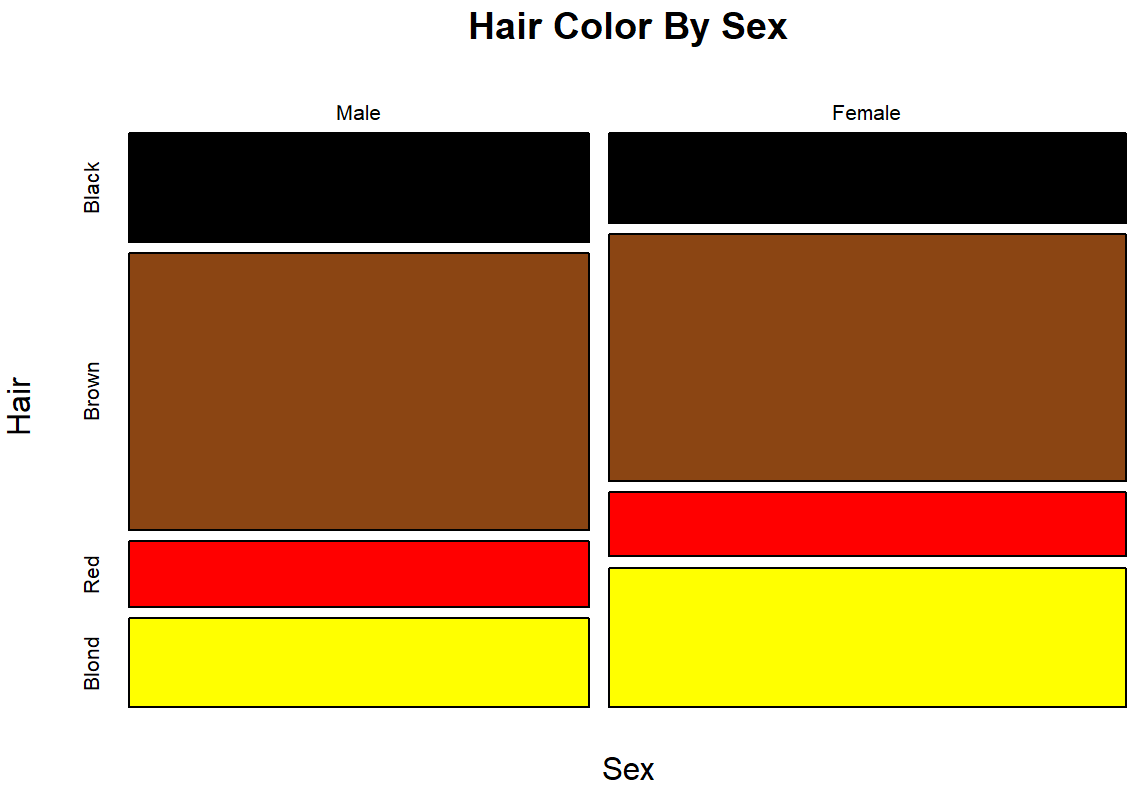
```
## [1] "Eye"
```

```
prop.table(apply(HEC, c(2), sum))
```

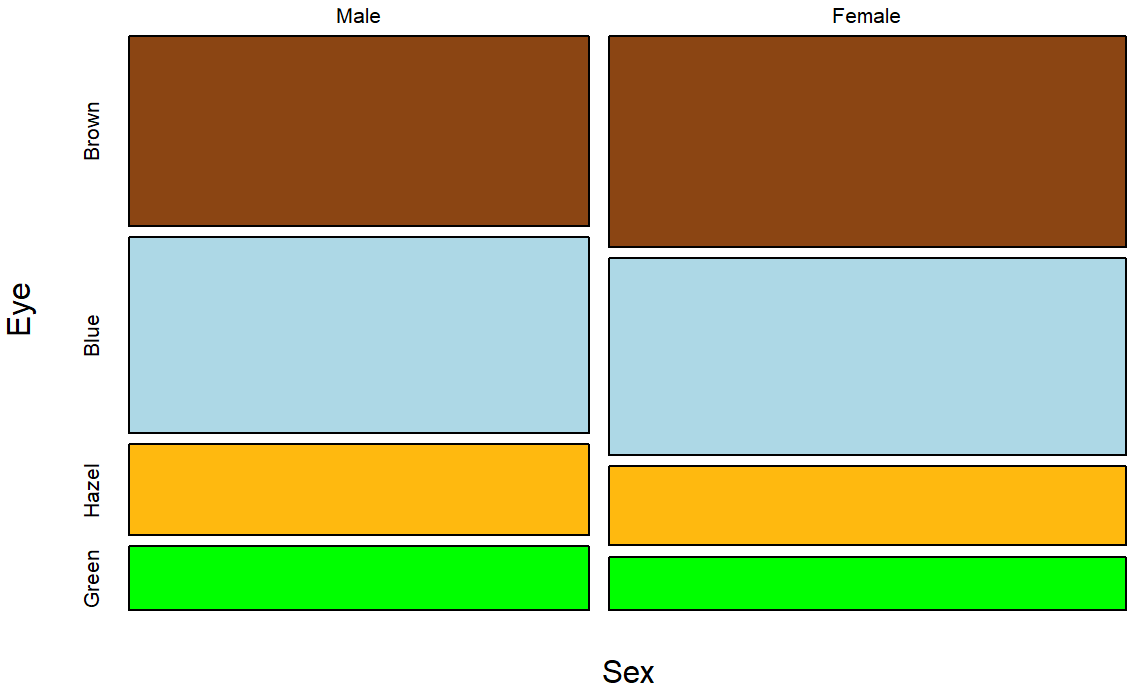
```
##     Brown      Blue     Hazel     Green
## 0.3716216 0.3631757 0.1570946 0.1081081
```

Are there large differences in hair and eye color frequencies between males and females?

```
mosaicplot(~Sex+Hair, data=HEC, col=c("Black","chocolate4", "red","Yellow"), main = "Hair Color By Sex")
```




```
mosaicplot(~Sex+Eye, data=HEC, col=c("chocolate4","light blue", "darkgoldenrod1","green"), main = "Eye Colo
r by Sex")
```



There do not appear to be large differences between the sexes.

Do hair and eye color sort independantly?

```
a <-prop.table(apply(HEC, c(1), sum))
b <-prop.table(apply(HEC, c(2), sum))
chisq.test(a,b)
```

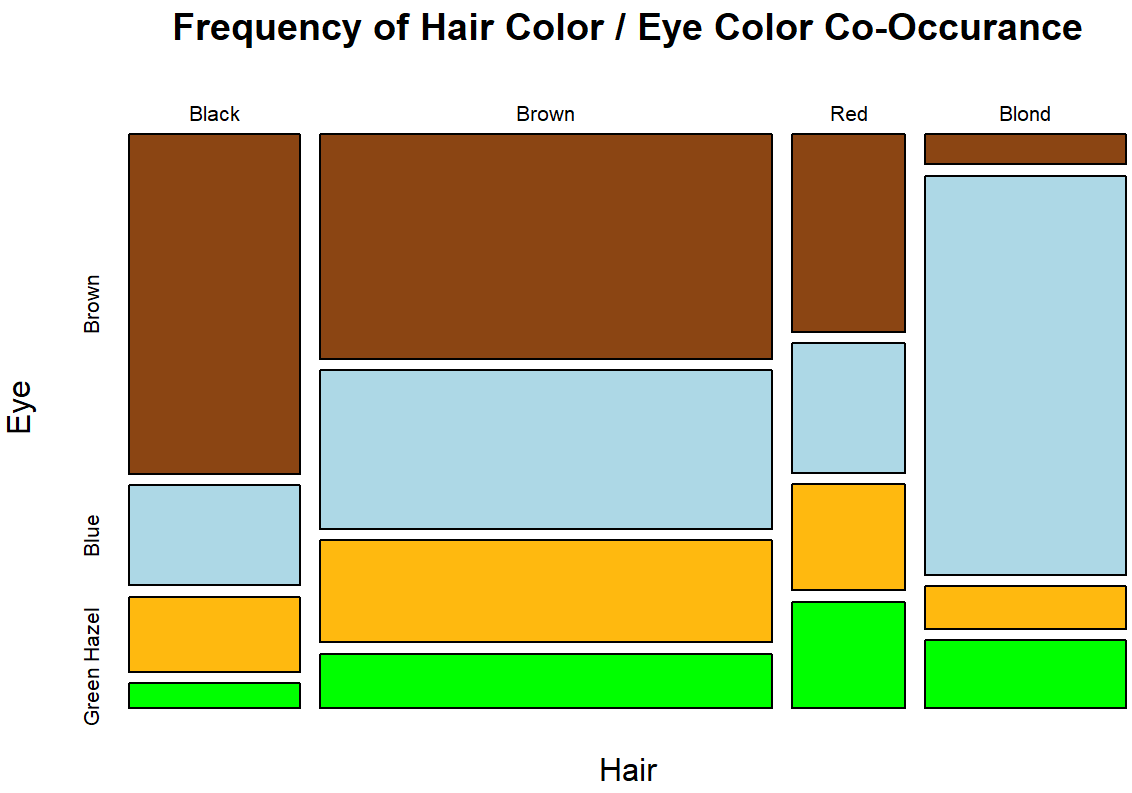
```
## Warning in chisq.test(a, b): Chi-squared approximation may be incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  a and b
## X-squared = 12, df = 9, p-value = 0.2133
```

The Chi Square value for Hair/Eye is .2133 This is less than .5, so an association seems to exist, but more than .05 so the certainty of an association is low.

Can we visualize which hair and eye colors are most likely to co-occur using a mosaic plot?

```
mosaicplot(~Hair+Eye, data=HEC, col=c("chocolate4","light blue", "darkgoldenrod1","green"), main= "Frequenc
y of Hair Color / Eye Color Co-Occurance")
```



It appears that (black hair / brown eyes) and (blond hair, blue eyes) are the most readily associated combinations.

2. Bat Tongues

```
Bat<-read.csv("C:\\Users\\xenon\\Desktop\\R Studio 2018\\bat tongues.csv")

plot(Bat$Palate.length.mm.,Bat$Tongue.length..mm., main="Palate Length vs. Tongue Length",
     xlab = "Palate", ylab = "Tongue",
     col.lab='dark green',
     col.main = 'purple',
     cex.lab = 1.5,
     cex.main=1.5,
     lwd = 2)
abline(lm(Bat$Tongue.length..mm. ~ Bat$Palate.length.mm.),
       col='red',
       lwd = 3,
       lty = "dotted")
```