Midterm4063-Shantal-Cruz

```r
# A) Write a function that gets a number "n" bigger than 2 and returns (2*3*4*.....*n) If the number gi

factorial <- function(n) {
  if (n < 2) {
    return(-1)
  } else {
    answer <- 1
    for (i in 2:n) {
        answer <- answer * i
    }
    return(answer)
  }
}

# B) Write a script and call your function then display the result when 10 is passed to the function.
print(paste("Result when 10 is passed:", factorial(10)))
```

```
## [1] "Result when 10 is passed: 3628800"
```

```r
# Calculate the minimum and maximum incomes in the city assigned to you.
library(readr)
data <- readr::read_csv("4063Midterm.csv")
```

```
## Rows: 1000 Columns: 13
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (4): Fname, Lname, gender, City
## dbl (9): ID, FamilyIncome, EdYears, FamilySize, Grocery, Cosmatics, MF, Boug...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
# City Toronto only
myCity <- data[data$City == "Toronto",]

# minimum income
min(myCity$FamilyIncome)
```

```
## [1] 10945
```

```r
# maximum income
max(myCity$FamilyIncome)
```

```
## [1] 79186
```

```r
# Calculate the median dollar amount spent on cosmetics in the city assigned to you.
median(myCity$Cosmatics)
```

```
## [1] 454
```

```r
# Calculate the median dollar amount spent on Grocery in the city assigned to you.
median(myCity$Grocery)
```

```
## [1] 1099
```

```r
# Visualize the frequency of Male and Female customers in the city assigned to you using bar chart with
barplot(table(myCity$gender), main="Gender Frequency", col=c("#bd4b7d", "#44c7ce"))
```

Gender Frequency
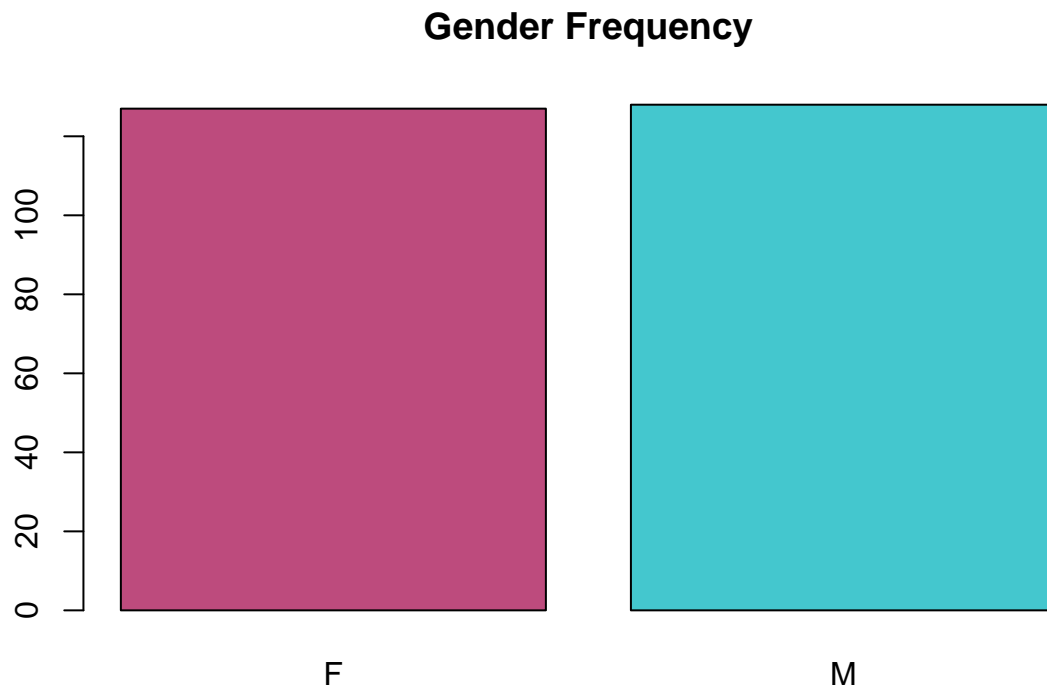

```r
# A) Visualize the distribution of family income in the city assigned to you by a histogram.
hist(myCity$FamilyIncome, main="Family Income Distribution", xlab="Family Income", labels=TRUE, col="#44
```

Family Income Distribution
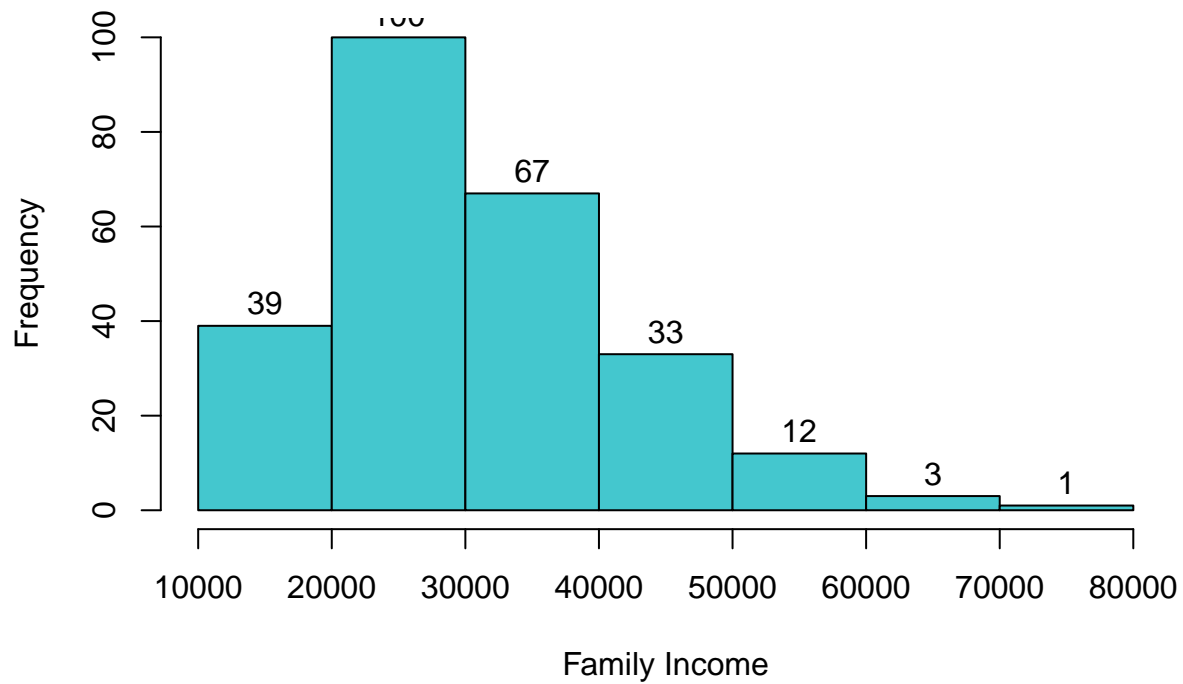

```r
# B) What is the mode for family income?
getmode <- function(v) {
  uniqv <- unique(v)
  uniqv[which.max(tabulate(match(v, uniqv)))]
}
getmode(myCity$FamilyIncome)
```

```
## [1] 22443
```

```r
# A) Visualize the distribution of Years of education in the city assigned to you by a histogram.
hist(myCity$EdYears, main="Years of Education Distribution", xlab="Years of Education", labels=TRUE, col
```

Years of Education Distribution
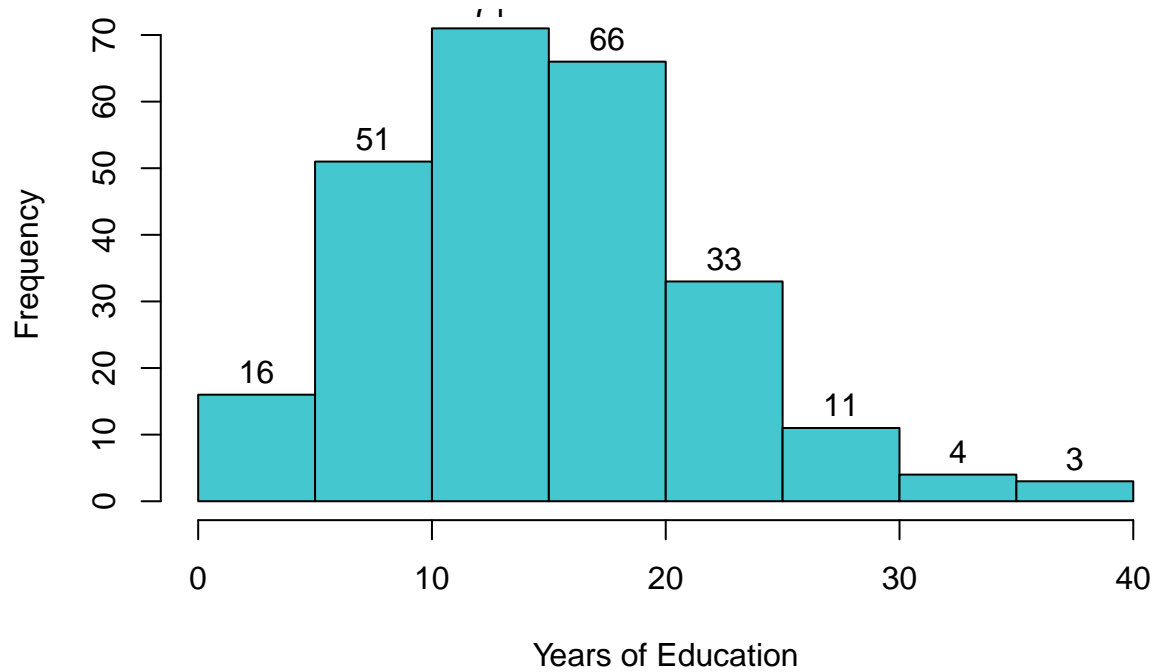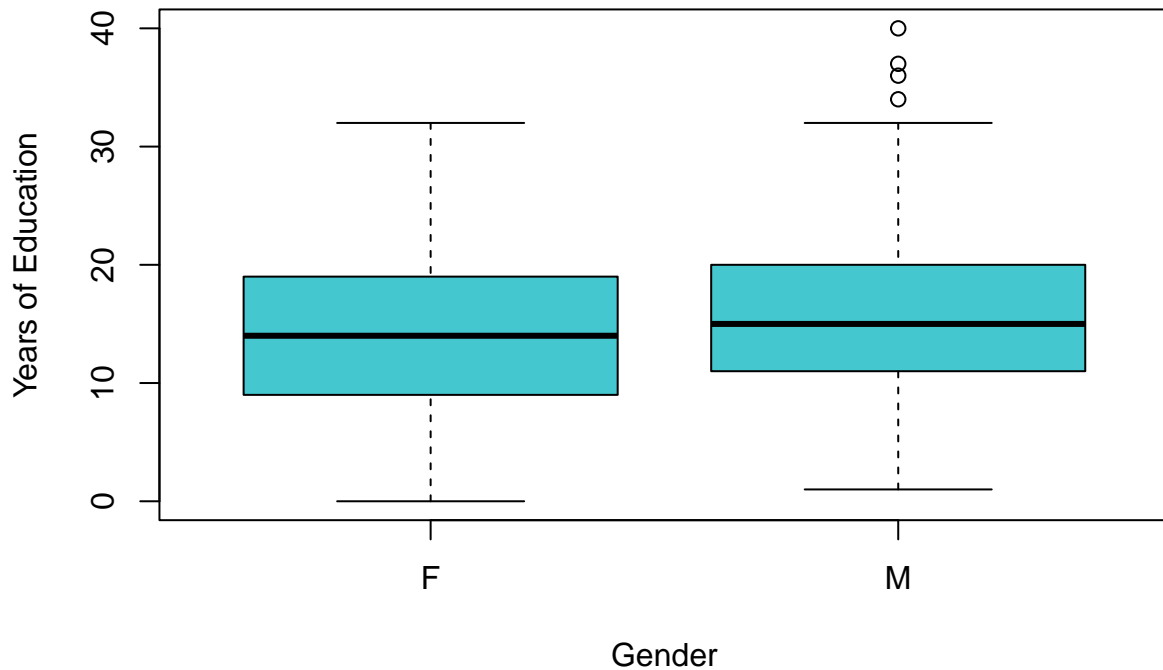


```r
# B) What is the mode for Years of education?
getmode(myCity$EdYears)
```

```
## [1] 12
```

```r
# Visualize the distribution of Years of education for male and female using a set of two boxplots.
boxplot(myCity$EdYears ~ myCity$gender, main="Education Years per Gender", xlab="Gender", ylab="Years o
```
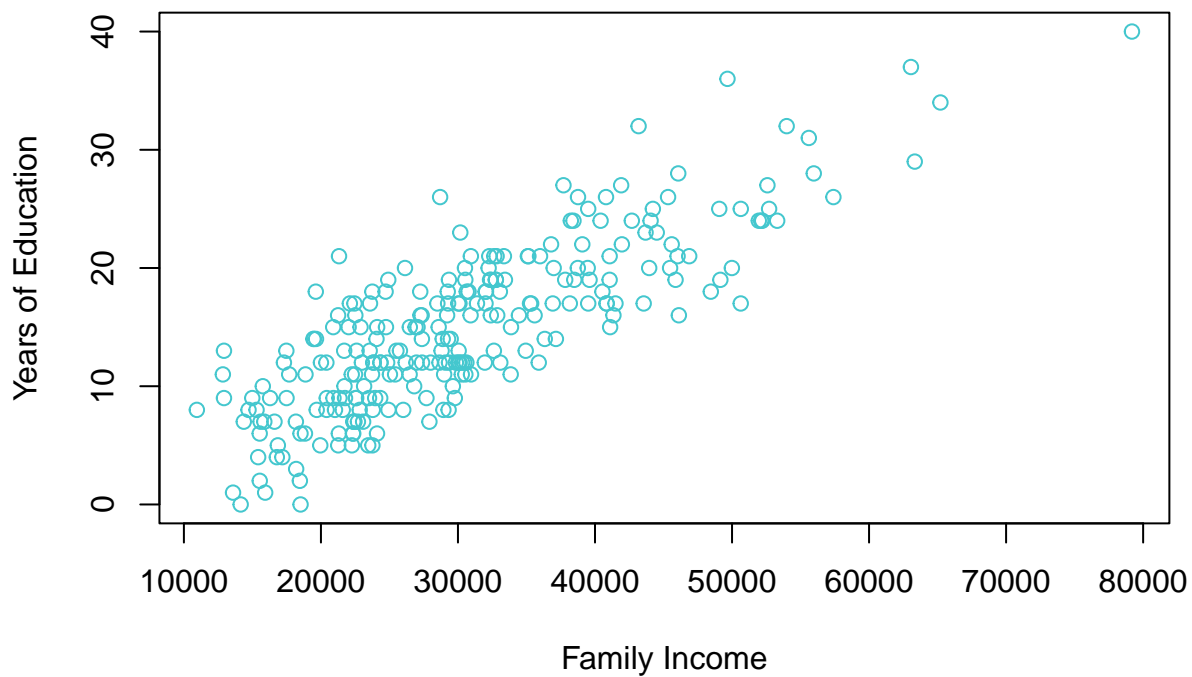
Education Years per Gender


```
# Use a scatter plot of the association between "family income" and "Years of education" to visualize t
plot(myCity$FamilyIncome, myCity$EdYears, main="Family Income vs. Years of Education", xlab="Family Inc
```

Family Income vs. Years of Education


```
# Create a 3x3 matrix of scatter plots of associations that shows the associations between 3 variables
histogramPanel <- function(x, ...) {
    par(new = TRUE)
```

```
    hist(x, ...)
}
pairs(myCity[,c("Cosmatics", "FamilyIncome", "EdYears")], diag.panel=histogramPanel, main="Scatterplot
```

Scatterplot Matrix for Cosmetics, Family Income, and Years of Education