```r
# Lab 14_Q2
getwd()  # Check current directory
```

```
## [1] "/Users/simransinha/Documents/Semester 3/SDA/Assignments/Week 7/Lab 14"
```

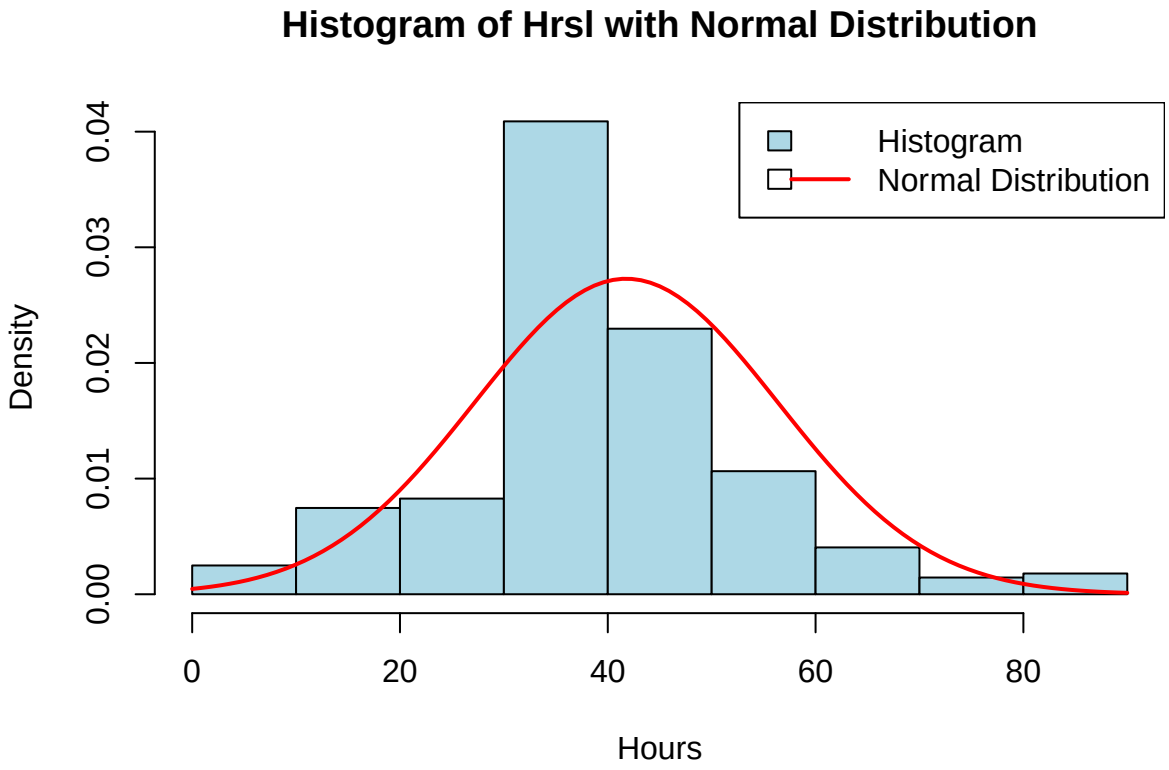
```r
# Load the dataset
load(file="descriptive_gss.Rdata")
```

```r
# Lab 14_Q1 A
head(descriptive_gss)
```

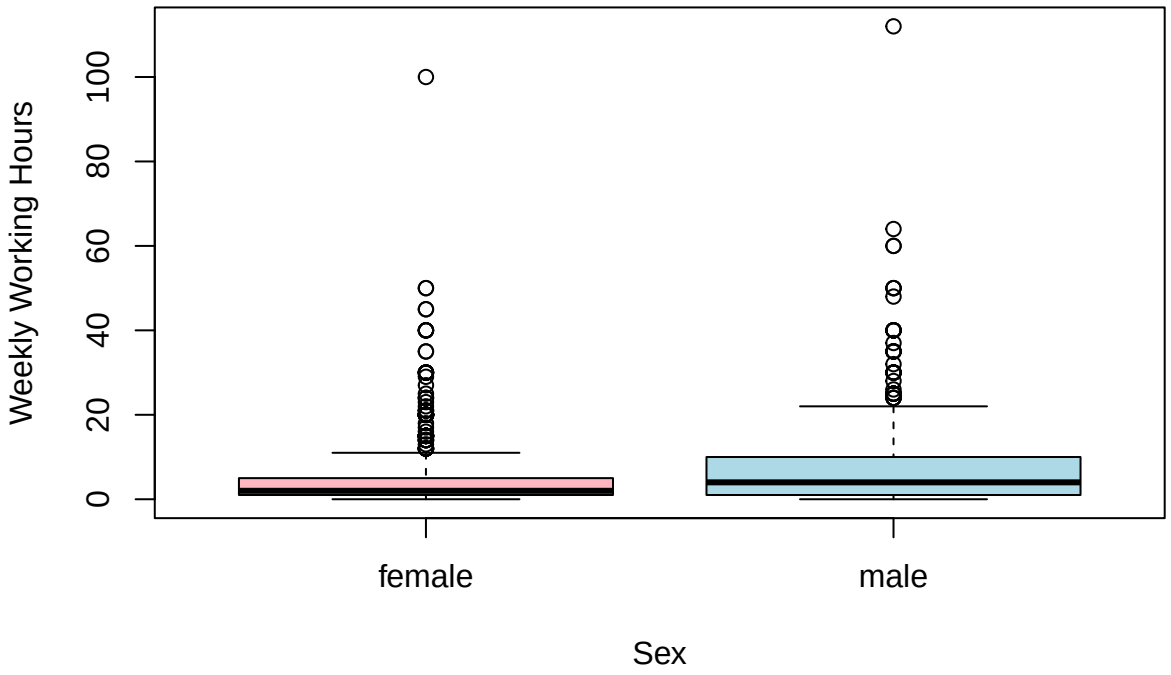
```
##     id hrs1  marital        childs age educ  sex             polviews wwwhr
## 1 2331   NA  married         three  71   18 male             moderate    NA
## 2 2003   NA divorced eight or more  69   11 male             moderate    NA
## 3 1221   NA  married         three  40   19 male                 <NA>     7
## 4 2051   NA  married           two  60   13 male slghtly conservative     1
## 5 2465   50  married          none  31   11 male             moderate     0
## 6  546   60  married          none  37   19 male              liberal     3
##         trustpeo wantbest                   advantge       goodlife
## 1       disagree    agree neither agree nor disagree       disagree
## 2 strongly agree disagree                      agree           <NA>
## 3           <NA>     <NA>                       <NA>           <NA>
## 4 strongly agree    agree             strongly agree           <NA>
## 5          agree    agree                      agree           <NA>
## 6          agree disagree                      agree strongly agree
##              deckids          strsswrk          satjob7
## 1               <NA>              <NA>             <NA>
## 2               <NA>              <NA>             <NA>
## 3 we decide together strongly disagree fairly satisfied
## 4               <NA>              <NA>             <NA>
## 5               <NA>              <NA>             <NA>
## 6               <NA>              <NA>             <NA>
```

```r
# 1.
hist(descriptive_gss$hrs1,
     main= "Basic Histogram of hrs1",
     xlab="Hours",
     ylab="Frequency",
     col="lightblue",
     border="black")
```

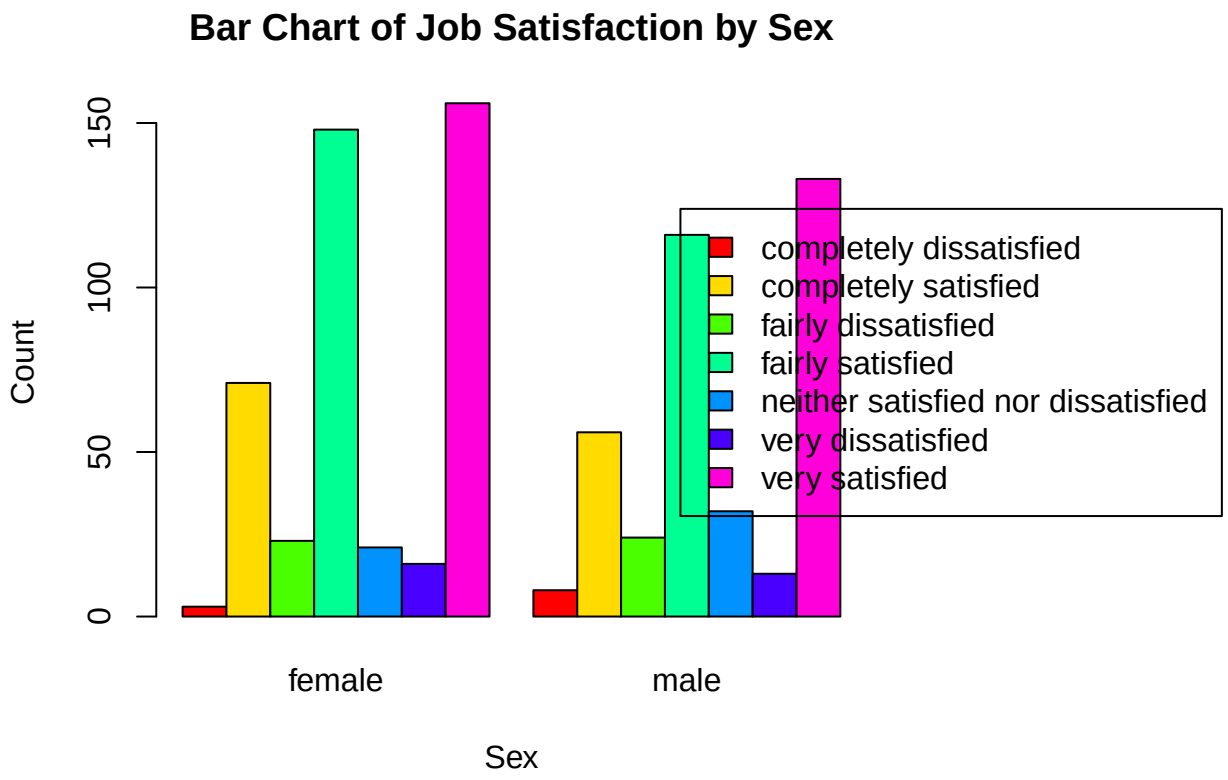

```r
# Lab 14_Q1 B
hrsl_mean <- mean(descriptive_gss$hrs1, na.rm=TRUE)
hrsl_sd <- sd(descriptive_gss$hrs1, na.rm=TRUE)

hist(descriptive_gss$hrs1,
     main="Histogram of Hrsl with Normal Distribution",
     xlab="Hours",
     ylab="Density",
     col="Lightblue",
     border="black",
     prob=TRUE)

curve(dnorm(x, mean=hrsl_mean, sd=hrsl_sd),
      add=TRUE, # add to existing plot
      col="red",
      lwd=2)

legend("topright",
       legend=c("Histogram", "Normal Distribution"),
       fill=c("lightblue", NA),
       border-c("black", NA),
       col=c(NA, "red"),
       lwd=c(NA, 2))
```




```
# Lab 14_Q1 C
boxplot(wwwhr ~ sex,
        data=descriptive_gss,
        main="Boxplot of Weekly Working Hours by Sex",
        xlab="Sex",
        ylab="Weekly Working Hours" ,
        col=c("lightpink", "lightblue"))
```

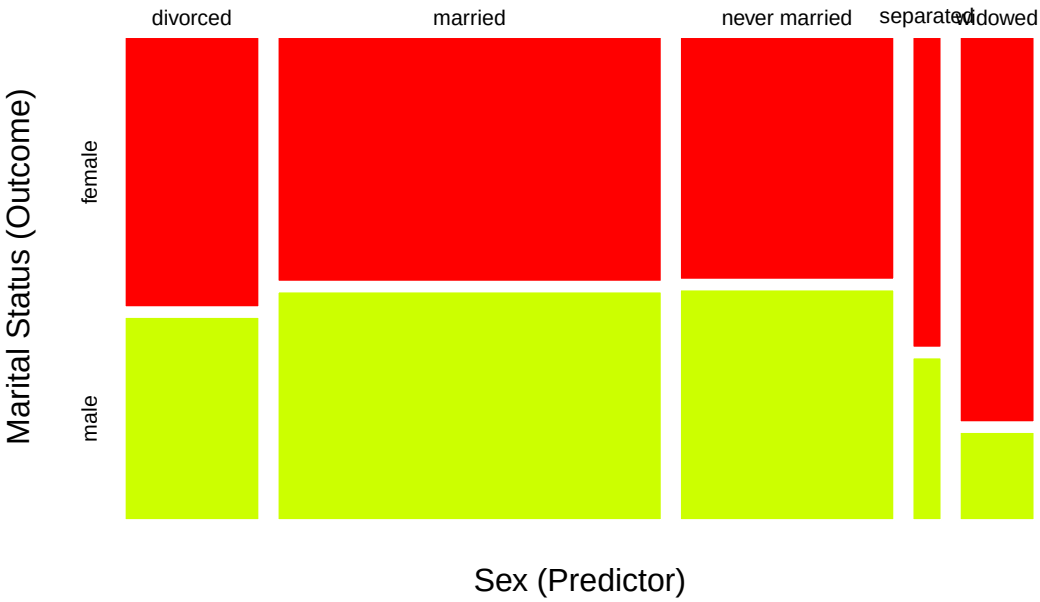



```r
# Lab 14_Q1 D
satjob_sex_table <- table(descriptive_gss$satjob7, descriptive_gss$sex)

par(mar = c(5, 4, 4, 10))
barplot(satjob_sex_table,
        main="Bar Chart of Job Satisfaction by Sex",
        xlab="Sex",
        ylab="Count",
        col=rainbow(nrow(satjob_sex_table)),
        beside=TRUE,
        args.legend = list(x = "right",
                           inset = c(-0.5, 0),
                           xpd = TRUE),
        legend.text = rownames (satjob_sex_table))
```




```
par (mar = c(5, 4, 4, 2) + 0.1)
```

```
# Lab 14_Q1 E
mosaicplot(table(descriptive_gss$marital, descriptive_gss$sex),
           main="Mosaic Plot of Marital Status by Sex",
           xlab="Sex (Predictor)",
           ylab="Marital Status (Outcome)",
           col=rainbow(length(unique(descriptive_gss$marital))),
           border="white")
```


Mosaic Plot of Marital Status by Sex
divorced
married
never married
separated
widowed
female
male
Marital Status (Outcome)
Sex (Predictor)