Lab1Problem6.R

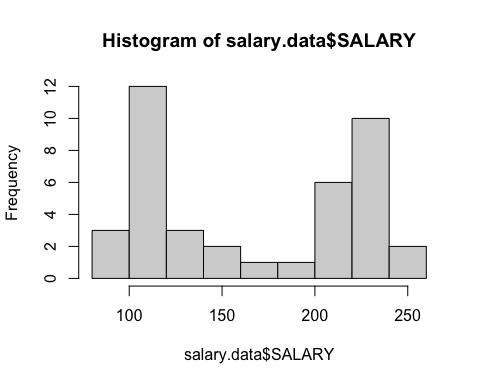
saint1729

2021-02-03

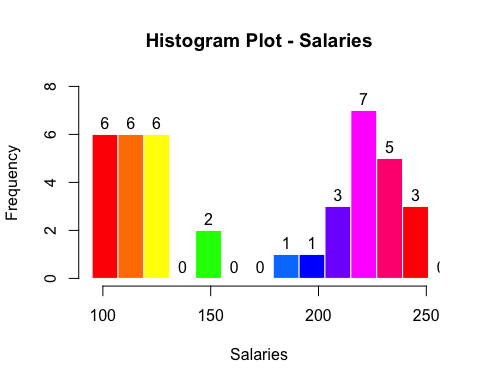
salary.data <- read.table(file="salary.txt", header=TRUE)  
salary.data

## GENDER SALARY  
## 1 F 100  
## 2 F 95  
## 3 F 105  
## 4 F 105  
## 5 F 110  
## 6 F 98  
## 7 F 105  
## 8 F 125  
## 9 F 130  
## 10 F 200  
## 11 F 120  
## 12 F 115  
## 13 F 110  
## 14 F 130  
## 15 F 120  
## 16 F 115  
## 17 F 110  
## 18 F 120  
## 19 F 115  
## 20 F 150  
## 21 M 150  
## 22 M 205  
## 23 M 210  
## 24 M 220  
## 25 M 205  
## 26 M 225  
## 27 M 230  
## 28 M 240  
## 29 M 220  
## 30 M 230  
## 31 M 235  
## 32 M 225  
## 33 M 230  
## 34 M 250  
## 35 M 245  
## 36 M 230  
## 37 M 225  
## 38 M 220  
## 39 M 180  
## 40 M 221

# a) Histogram of the salaries using R default setting  
hist(salary.data$SALARY)



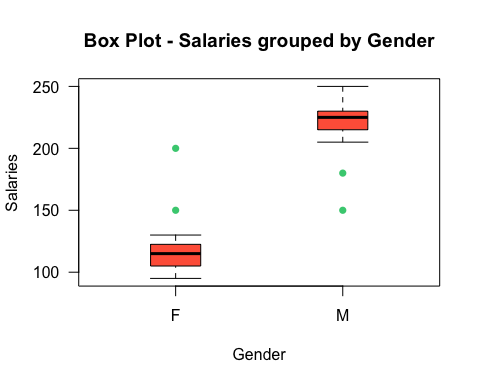
# b) Histogram with break points (at least 15 intervals)  
hist(salary.data$SALARY,  
 breaks=95+(0:14)\*12,  
 main = "Histogram Plot - Salaries",  
 xlab = "Salaries",  
 ylab = "Frequency",  
 border = FALSE,  
 labels = TRUE,  
 xlim = c(min(salary.data$SALARY), max(salary.data$SALARY)),  
 ylim = c(0, 8),  
 col = rainbow(12))



# c) Boxplot of the salaries  
boxplot(salary.data$SALARY,  
 main = "Box Plot - Salaries",  
 xlab = "All genders",  
 ylab = "Salaries",  
 labels = TRUE,  
 boxwex = 0.3,  
 outline = TRUE,  
 las = 1,  
 notch = FALSE,  
 staplewex = 1,  
 col = "purple")



# d) Boxplots of the salaries in two gender groups  
boxplot(SALARY~GENDER, data=salary.data,  
 main = "Box Plot - Salaries grouped by Gender",  
 xlab = "Gender",  
 ylab = "Salaries",  
 labels = TRUE,  
 boxwex = 0.3,  
 outline = TRUE,  
 outpch = 16,  
 outcol = "seagreen3",  
 las = 1,  
 notch = FALSE,  
 staplewex = 1,  
 col = "tomato")  
  
library(psych)



# e) summary statistics of salaries as one group and summary statistics within each gender  
describe(salary.data$SALARY)

## vars n mean sd median trimmed mad min max range skew kurtosis se  
## X1 1 40 169.35 55.83 165 168.94 81.54 95 250 155 0.02 -1.83 8.83

describeBy(salary.data$SALARY, salary.data$GENDER)

##   
## Descriptive statistics by group   
## group: F  
## vars n mean sd median trimmed mad min max range skew kurtosis se  
## X1 1 20 118.9 23.01 115 114.69 14.83 95 200 105 2.16 5.03 5.15  
## ------------------------------------------------------------   
## group: M  
## vars n mean sd median trimmed mad min max range skew kurtosis se  
## X1 1 20 219.8 22.57 225 223.19 7.41 150 250 100 -1.53 2.37 5.05