Assignments for session on "Exploratory Data Analytics"

install.packages("RcmdrPlugin.IPSUR")

library(RcmdrPlugin.IPSUR)

head(RcmdrTestDrive)

#a. Calculate the average salary by gender and smoking status.

#Gender-wise

tapply(RcmdrTestDrive$salary, RcmdrTestDrive$gender, mean)

#Smoking Status-wise

tapply(RcmdrTestDrive$salary, RcmdrTestDrive$smoking, mean)

#Which Gnder has the highest mean salary

#Answer is MALE as per the below output

Female   Male

698.0911 743.3915

Report the highest mean salary

which.max(RcmdrTestDrive$salary)

#Output is 152

View(RcmdrTestDrive)

#Male driver has highest salary 1156.16

#To Compare the Spread in Salary Gender wise we can plot graphs or compute mean

tapply(RcmdrTestDrive$salary, RcmdrTestDrive$gender, sd)

#This indicates SD is higher for Male

Female   Male

#130.7053 158.5423