Assignment 8.1

Name : Y Vasudev

Batch : DA with R , Excel and Tableau

1. Use the package RcmdrPlugin.IPSUR.

data(RcmdrTestDrive)

and perform the below operations:

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

b. Which gender has the highest mean salary?

Ans . Male

> x <- tapply(salary, list(gender = gender), mean)

> x

gender

Female Male

16.46353 17.93035

OR

> by(salary, gender, mean, na.rm = TRUE)

gender: Female

[1] 16.46353

gender: Male

[1] 17.93035

c. Report the highest mean salary.

mean(salary [gender == Male])

> x[which(x == max(x))]

Male

17.93035

d. Compare the spreads for the genders by calculating the standard deviation of salary by gender.

y <- tapply(salary, list(gender = gender), sd)

> y

gender

Female Male

2.122113 1.077183

Female gender has the highest SD = 2.12