**ACD Session 8 \_ Assignment – 1**

#Problem 1

#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?

#c Report the highest mean salary.

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

Solution :- a) We can generate a table listing the average salaries by gender with two methods.

The first uses tapply:

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

> x

gender

Female Male

16.46353 17.93035

The second method uses the by function:

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

gender: Female

[1] 16.46353

--------------------------------------------------------

gender: Male

[1] 17.93035

b) male

c) For these data, the highest mean salary is

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

Male

17.93035

d)

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

> y

gender

Female Male

2.122113 1.077183

For these data, the the largest standard deviation is approximately 2.12 which was attained

by the Female gender.