**Assignment 8.1**

**Question No.1:**

Calculate the average salary by gender and smoking status

> aggregate(RcmdrTestDrive$salary, list(RcmdrTestDrive$gender, RcmdrTestDrive$smoking), mean)

> colnames(gensmokemean)[c(1,2,3)] <- c(“Gender”, "SmokStat", "Mean Salary")

Gender SmokStat Mean Salary

1 Female Nonsmoker 692.9093

2 Male Nonsmoker 740.9080

3 Female Smoker 733.2122

4 Male Smoker 751.4900

> with(RcmdrTestDrive, tapply(salary, list(gender, smoking), mean,

Rcmdr+ na.rm=TRUE))

Nonsmoker Smoker

Female 692.9093 733.2122

Male 740.9080 751.4900

USING Rcmdr. PLUGIn

Rcmdr> numSummary(RcmdrTestDrive[,"salary", drop=FALSE],

Rcmdr+ groups=RcmdrTestDrive$gender, statistics=c("mean"), quantiles=c(0,.25,.5,

Rcmdr+ .75,1))

mean n

Female 698.0911 70

Male 743.3915 98

Rcmdr> numSummary(RcmdrTestDrive[,"salary", drop=FALSE],

Rcmdr+ groups=RcmdrTestDrive$smoking, statistics=c("mean"), quantiles=c(0,.25,.5,

Rcmdr+ .75,1))

mean n

Nonsmoker 719.3792 136

Smoker 746.3494 32

Rcmdr> numSummary(RcmdrTestDrive[,"salary", drop=FALSE],

Rcmdr+ groups=RcmdrTestDrive$gender, statistics=c("sd"), quantiles=c(0,.25,.5,.75,

Rcmdr+ 1))

sd n

Female 130.7053 70

Male 158.5423 98

**Question No.2:**

Which of the genders has the highest average salary: **The answer is Male**

> aggregate(RcmdrTestDrive$salary, list(RcmdrTestDrive$gender), mean)

Group.1 x

1 Female 698.0911

2 Male 743.3915

**Question No.3:**

Report the highest average salary

> gensalary <- aggregate(RcmdrTestDrive$salary, list(RcmdrTestDrive$gender), mean)

> gensalary <- as.data.frame(gensalary)

> colnames(gensalary)[1,2] <- c("Gender","Mean Salary")

> max(gensalary[,2])

[1] 743.3915

> max(gensalary$`Mean Salary`)

[1] 743.3915

**Question No.4:**

Compare the Spreads for the genders by calculating the standard deviations of salary by gender

> aggregate(RcmdrTestDrive$salary, list(RcmdrTestDrive$gender), sd)

Group.1 Std.Dev.

1 Female 130.7053

2 Male 158.5423

> gensum <- cbind(gensalary,gendisp$Std.Devn.)

> gensum

Gender Mean Salary gendisp$Std.Devn.

1 Female 698.0911 130.7053

2 Male 743.3915 158.5423

> class(gensum)

[1] "data.frame"

> colnames(gensum)[3] <- "Std.Devn."

> gensum[,4] <- gensum$`Mean Salary` + gensum$Std.Devn.

> gensum

Gender Mean Salary Std.Devn. V4

1 Female 698.0911 130.7053 828.7964

2 Male 743.3915 158.5423 901.9338

> colnames(gensum)[4] <- "Hi-En-Sal"

> gensum[,5] <- gensum$`Mean Salary` - gensum$Std.Devn.

> colnames(gensum)[5] <- "Lo-En-Sal"

> gensum

Gender Mean Salary Std.Devn. Hi-En-Sal Lo-En-Sal

1 Female 698.0911 130.7053 828.7964 567.3859

2 Male 743.3915 158.5423 901.9338 584.8492