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.

1. Average Salary of Female for both smoking and non smoking calculated in R GUI

df<-RcmdrTestDrive %>%

+ filter(gender=="Female")%>%

+ select(smoking,gender,salary)

avg\_sal\_female<-mean(df$salary)

avg\_sal\_female

[1] 698.0911

Is mean salary of female with smoking status both

1. Average Salary of male for both smoking and non smoking calculated in R GUI

dfm<-RcmdrTestDrive%>%

+ filter(gender=="Male")%>%

+ select(smoking,gender,salary)

> dfm

mean(dfm$salary)

[1] 743.3915

SO Mean Salary Male=743.3915 with smoking status both

1. Male , NON Smoker

dfm<-RcmdrTestDrive%>%

+ filter(gender=="Male",smoking=="Nonsmoker")%>%

+ select(smoking,gender,salary)

>

avg\_male\_non\_smok<-mean(dfm$salary)

> avg\_male\_non\_smok

[1] 740.908

1. Male Smoker is 751.49

malesmoker<-RcmdrTestDrive%>%

+ filter(gender=="Male",smoking=="Smoker")%>%

+ select(smoking,gender,salary)

> mean(malesmoker$salary)

[1] 751.49

1. Female Non Smoker is 692.9093

femalenonsmoker<-RcmdrTestDrive%>%

+ filter(gender=="Female",smoking=="Nonsmoker")%>%

+ select(smoking,gender,salary)

mean(femalenonsmoker$salary)

[1] 692.9093

1. Female Smoker is 692.9093

femalesmoker<-RcmdrTestDrive%>%

+ filter(gender=="Female",smoking=="Nonsmoker")%>%

+ select(smoking,gender,salary)

mean(femalesmoker$salary)

[1] 692.9093

1. WHICH GENDER HAS THE HIGHEST MEAN SALARY?

Male Smoker is 751.49

1. Report the highest Mean Salary

MALE SMOKER-HISPANIC HAS THE HIGHEST MEAN SALARY of 829.39

ms<-RcmdrTestDrive%>%

+ filter(smoking=="Smoker",gender=="Male",race=="Hispanic")%>%

+ select(smoking,gender,race,salary)

> mean(ms$salary)

[1] 829.39

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

sdmale<-RcmdrTestDrive%>%

+ filter(gender=="Male")%>%

+ select(gender,salary)

sd(sdmale$salary)

[1] 158.5423

sdfemale<-RcmdrTestDrive%>%

+ filter(gender=="Female")%>%

+ select(gender,salary)

> sd(sdfemale$salary)

[1] 130.7053

Female salaries have lower sd compared to male salries