#Assignment6.2\_Session6

#Problem 1

#1. Import the Titanic Dataset from the link Titanic Data Set.

#Perform the following:

# a. Is there any difference in fares by different class of tickets?

# Note - Show a boxplot displaying the distribution of fares by class

# b. Is there any association with Passenger class and gender?

# Note - Show a stacked bar chart

#Answer1

#a)

#use titanic dataset

boxplot(fare~pclass,data= titanic,

main="Fares Versus Pclass",xlab="Fares",ylab="Class",col=topo.colors(3))

#b)

#stacked bar chart

counts<-table(titanic$sex,titanic$pclass)

barplot(counts, main = "Distribution of Class by gender", xlab="Pclass", col=c("blue", "red"), legend = c("Female","Male"), names.arg = c("Pclass1st", "Pclass2nd","Pclass3rd"))

#or like this too

a1<-as.numeric(titanic$sex)

counts<-table(a1,titanic$pclass)

barplot(counts, main = "Distribution of Class by gender", xlab="Pclass", col=c("blue", "red"), legend = c("Female","Male"), names.arg = c("Pclass1st", "Pclass2nd","Pclass3rd"))

#we can do chisq test also for checking association

chisq.test(titanic$pclass ,titanic$sex)

#ho:there is no association

#since p value is 0.0002064<0.05

#we reject the null hypothesis and thus say there is association