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

Sol= TitanicData <- read.csv("C:/Users/Munmun/Downloads/Titanic/train.csv", header=FALSE)

str(TitanicData)

colnames(TitanicData) <- c("PassengerId","Survived","Pclass","Name",

"Sex","Age","SibSp","Parch","Ticket","Fare",

"Cabin","Embarked")

TitanicData <- TitanicData[,-13]

Titanic <- TitanicData %>% mutate(Pclass = as.factor(Pclass)) # Pasanger class as factor

str(Titanic)

View(Titanic)

boxplot(Fare~Pclass, data = Titanic, col = topo.colors(3),

xlab = "Class of Ticket", ylab = "Fares", main = "Fares by different Class of Tickets")

# Yes. Fares are different as per Class of Ticket.

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

Note – Show a stacked bar chart

Sol= A <- table(Titanic$Sex, Titanic$Pclass)

A

bp <- barplot(A, col= rainbow(length(A)), legend = rownames(A),

main = "Passenger class and gender",

xlab = "Class of Ticket", ylab = "No. of Passangers by Gender")

# Male paasengers are more than female in each class . Also the percentage of

# male passsengers over Female Passangers is more in class 3 as compared to class 1 & 2