Problem Statement

1) Import the Titanic Dataset from the following link:
https://drive.google.com/file/d/1JTJCjdGuUxzKXYlwOavwovB01k6FWg3r/view?ts=5b42ea10
Perform the below operations:

a. Is there any difference in fares by different class of tickets?
Note- show a boxplot displaying the distribution of fares by class

Answer :

```
library(readr)
titanic <- read_csv("G:/DATA ANALYTICS/DATA/titanic3.csv")
str(titanic)

View(titanic)

boxplot(fare~pclass,data= titanic,
main="Fares Versus Pclass",xlab="Class",ylab="Fares",col=topo.colors(3))
```

b. Is there any association with Passenger class and gender?
Note- show a stacked bar chart

Answer :

```
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"))


#another way --> chisq test 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
```