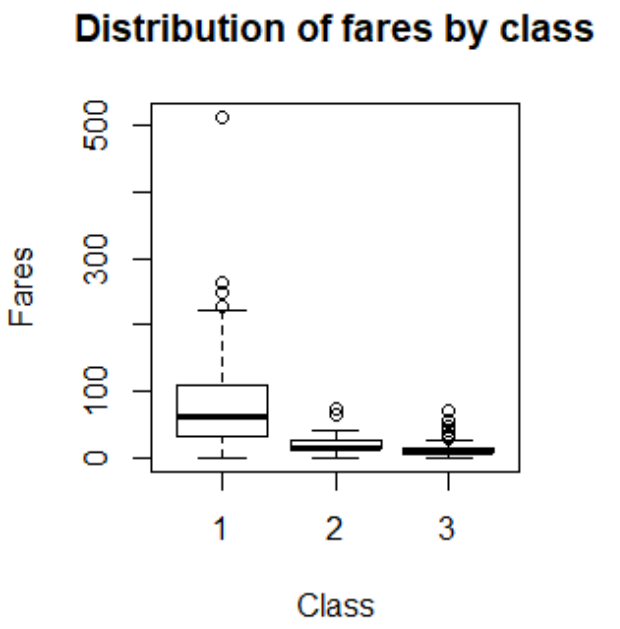
Task 1:

1. Import the Titanic Dataset from the link => Titanic Data Set.
   Perform the following:
   a) Is there any difference in fares by a different class of tickets?
      Note - Show a boxplot displaying the distribution of fares by class

   Ans:-

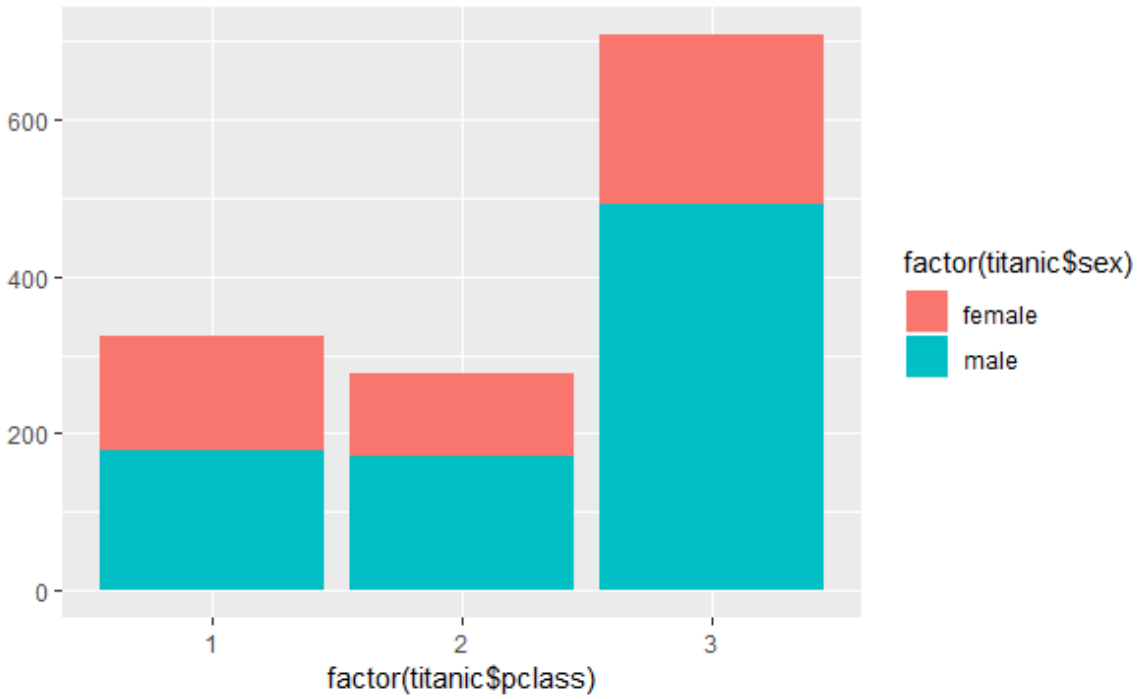
```
titanic = read.csv("titanic3.csv")
> head(titanic)
> View(titanic)
> df = data.frame(titanic$pclass,titanic$fare)
> View(df)
> boxplot(fare ~ pclass, data = titanic, xlab = "Class",
+         ylab = "Fares", main = "Distribution of fares by class")
```



   b) Is there any association with Passenger class and gender?
      Note – Show a stacked bar chart

   Ans:-

```
> titanic = read.csv("titanic3.csv")
> library(ggplot2)
> qplot(factor(titanic$pclass), data=titanic, geom="bar",
fill=factor(titanic$sex))
```

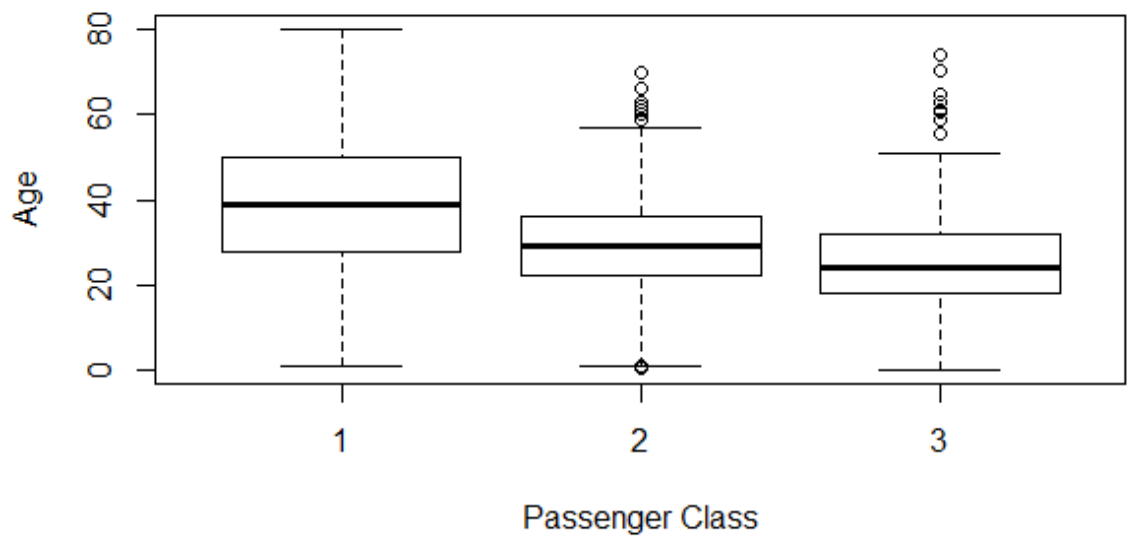


Task 2:

1. Create a box and whisker plot by class using mtcars dataset.

Ans:-

```
> titanic = read.csv("titanic3.csv")
> boxplot(age ~ pclass, data = titanic, xlab = "Passenger Class",
+ ylab = "Age")
```



Task 3:

1. A recent national study showed that approximately 44.7% of college students have used Wikipedia as a source in at least one of their term papers. Let X equal the number of students in a random sample of size n = 31 who have used Wikipedia as a source.

Perform the below functions:

a) Find the probability that X is equal to 17

Ans:-
```
> dbinom(x = 17, size = 31, prob = 0.447)
[1] 0.07532248
```

b) Find the probability that X is at most 13

Ans:-
```
> pbinom(q = 13, size = 31, prob = 0.447)
[1] 0.451357
```

c) Find the probability that X is bigger than 11.

Ans:-
```
> pbinom(q = 11, size = 31, prob = 0.447,
lower.tail=FALSE)
[1] 0.8020339
```

d) Find the probability that X is at least 15.

Ans:-
```
> pbinom(q = 15, size = 31, prob = 0.447,
lower.tail=FALSE)
[1] 0.2753716
```

e) Find the probability that X is between 16 and 19, inclusive

Ans:-
```
> diff(pbinom(c(19, 16), size = 31, prob = 0.447,
lower.tail=FALSE))
[1] 0.1488671
```