ASSIGNMENT – 4

**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****.***

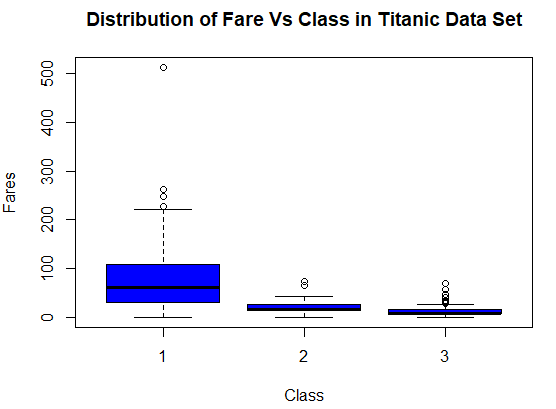
**Solution:**

*Titanic1 = read.csv("Titanic.csv")*

*boxplot(fare ~ pclass, data = Titanic1, xlab = "Class",*

*ylab = "Fares", main = "Distribution of Fare Vs Class in Titanic Data Set", col="Blue")*

***Output:***



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

***Note –*** *Show a stacked bar chart*

**Solution:**

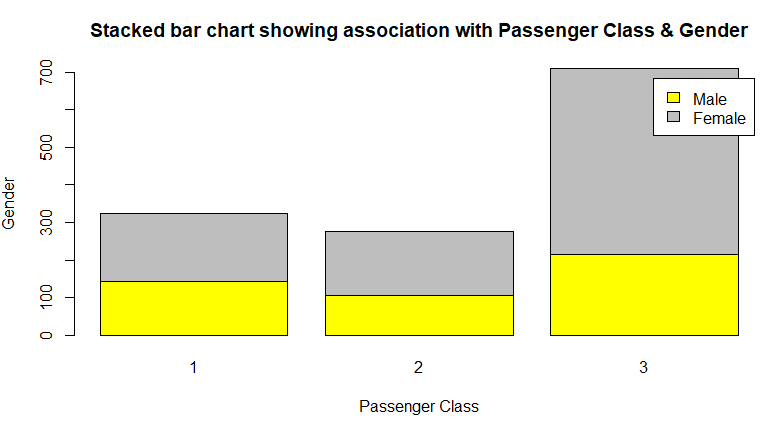
*Titanic1 = read.csv("Titanic.csv")*

*counts <- table(Titanic1$sex, Titanic1$pclass)*

*barplot(counts, main="Stacked bar chart showing association with Passenger Class & Gender",*

*xlab="Passenger Class",ylab = "Gender",col=c("grey","yellow"),legend.text = c("Female", "Male"))*

***Output:***



**2. Create a box and whisker plot by class and age using titanic dataset.**

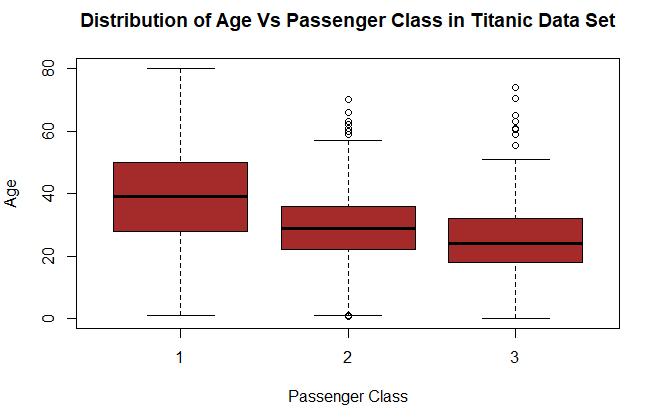
**Solution:**

*Titanic1 = read.csv(&quot;Titanic.csv&quot;)*

*boxplot(age ~ pclass, data = Titanic1, xlab = &quot;Passenger Class&quot;,*

*ylab = &quot;Age&quot;, main = &quot;Distribution of Age Vs Passenger Class in Titanic Data Set&quot;, col=&quot;Brown&quot;)*

***Output:***



**3. 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.***

**Solution:**

*dbinom(17, size = 31, prob = 0.447)*

***Output:***



***b. Find the probability that X is at most 13***

**Solution:**

*pbinom(13, size = 31, prob = 0.447)*

***Output:***



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

**Solution:**

*pbinom(11, size = 31, prob = 0.447, lower.tail = FALSE)*

***Output:***



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

**Solution:**

*pbinom(14, size = 31, prob = 0.447, lower.tail = FALSE)*

***Output:***



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

**Solution-1:**

*sum(dbinom(16:19, size = 31, prob = 0.447))*

***Output:***

