Problem Statement

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 char

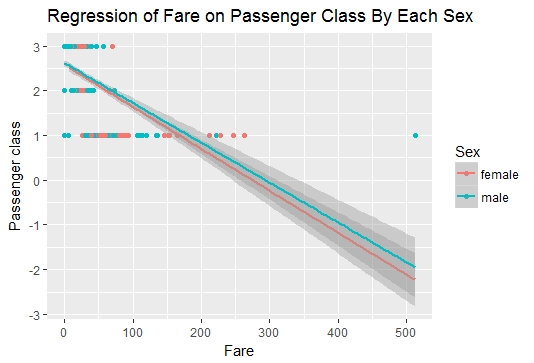
The graph is drawn after preprocessing the data and other appropriate libraries incorporated in session 6 assessment 1

tapply(training$Fare,training$Pclass,mean)

1 2 3

84.15469 20.66218 13.67555

qplot(Fare,Pclass,data = training, geom = c("point","smooth"), method = "lm", formula = y~x, col = Sex, main = "Regression of Fare on Passenger Class By Each Sex", ylab = "Passenger class", xlab = "Fare")



|  |
| --- |
| tapply(training$Fare,training$Deck,mean)  A B C D E F G T  39.62389 113.50576 100.15134 57.24458 46.02669 18.69679 13.58125 35.50000 |
|  |
| |  | | --- | | > | |

qplot(Deck,Fare, data = training, geom = c("boxplot"), fill = Sex, main = "Fare Per Deck",xlab = "", ylab = "Fare")

