6.1

1. Import the Titanic Dataset from the link Titanic Data Set.

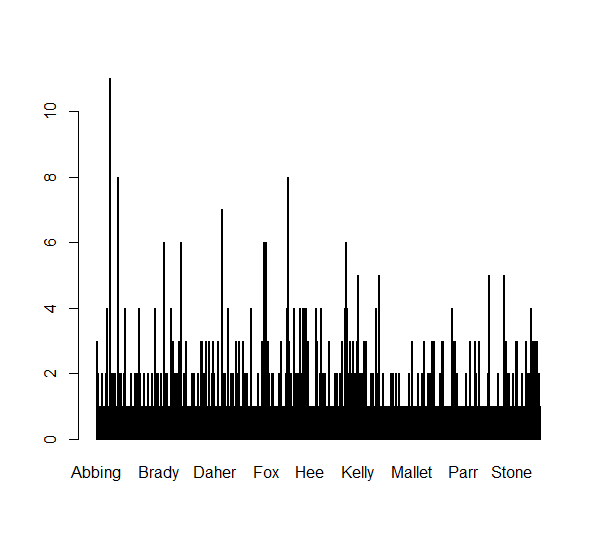
> setwd("C:/Users/satish/Desktop")

> titanic<-read.table(file = 'titanic.txt', sep = '\t', header = TRUE)

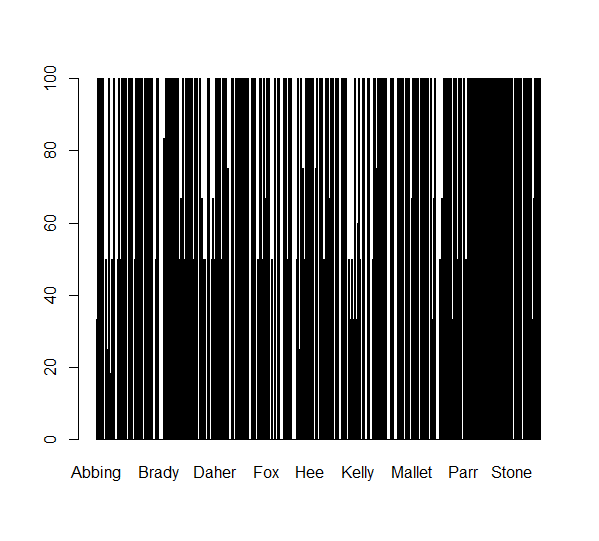
Perform the following:

a. Preprocess the passenger names to come up with a list of titles that represent families

and represent using appropriate visualization graph.



b. Represent the proportion of people survived from the family size using a graph.



c. Impute the missing values in Age variable using Mice Library, create two different

graph showing Age distribution and after imputation

sum(is.na(train$Age))

[1] 177

> mice\_imputes = mice(train, m=5, maxit = 40)

iter imp variable

1 1 Age

