ML1

Mark Christian

9/13/2020

GOT <- read.csv("character-deaths.csv")

Thi is the summary descriptive statistics where I identified the classification: chatagorial variables have character and quantitative have statistical summary of mean, meadian and mode.

summary(GOT)

## Name Allegiances Death.Year Book.of.Death   
## Length:917 Length:917 Min. :297.0 Min. :1.000   
## Class :character Class :character 1st Qu.:299.0 1st Qu.:2.000   
## Mode :character Mode :character Median :299.0 Median :3.000   
## Mean :299.2 Mean :2.928   
## 3rd Qu.:300.0 3rd Qu.:4.000   
## Max. :300.0 Max. :5.000   
## NA's :612 NA's :610   
## Death.Chapter Book.Intro.Chapter Gender Nobility   
## Min. : 0.00 Min. : 0.00 Min. :0.0000 Min. :0.0000   
## 1st Qu.:25.50 1st Qu.:11.00 1st Qu.:1.0000 1st Qu.:0.0000   
## Median :39.00 Median :27.00 Median :1.0000 Median :0.0000   
## Mean :40.07 Mean :28.86 Mean :0.8288 Mean :0.4689   
## 3rd Qu.:57.00 3rd Qu.:43.00 3rd Qu.:1.0000 3rd Qu.:1.0000   
## Max. :80.00 Max. :80.00 Max. :1.0000 Max. :1.0000   
## NA's :618 NA's :12   
## GoT CoK SoS FfC   
## Min. :0.0000 Min. :0.0000 Min. :0.0000 Min. :0.0000   
## 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000   
## Median :0.0000 Median :0.0000 Median :0.0000 Median :0.0000   
## Mean :0.2726 Mean :0.3533 Mean :0.4242 Mean :0.2726   
## 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:1.0000   
## Max. :1.0000 Max. :1.0000 Max. :1.0000 Max. :1.0000   
##   
## DwD   
## Min. :0.0000   
## 1st Qu.:0.0000   
## Median :0.0000   
## Mean :0.2846   
## 3rd Qu.:1.0000   
## Max. :1.0000   
##

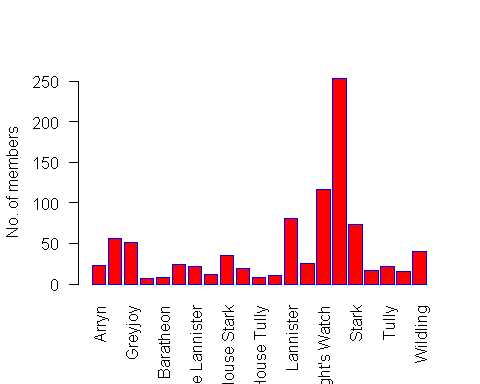
This Table shows the frequency of what characters of the show belonged to which royal family

FTable <- table(GOT$Allegiances)  
FTable

##   
## Arryn Baratheon Greyjoy House Arryn House Baratheon   
## 23 56 51 7 8   
## House Greyjoy House Lannister House Martell House Stark House Targaryen   
## 24 21 12 35 19   
## House Tully House Tyrell Lannister Martell Night's Watch   
## 8 11 81 25 116   
## None Stark Targaryen Tully Tyrell   
## 253 73 17 22 15   
## Wildling   
## 40

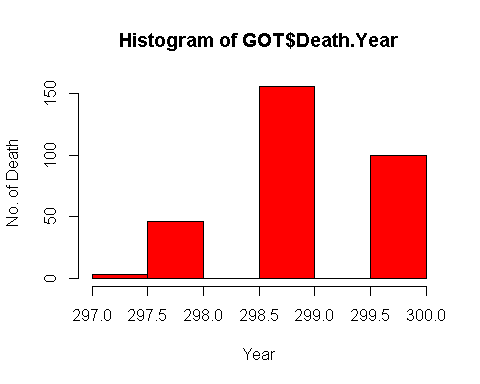
This is a bar graph for all the families and their member count

barplot(FTable, ylab="No. of members",col = "red", border = "blue", las = 2)



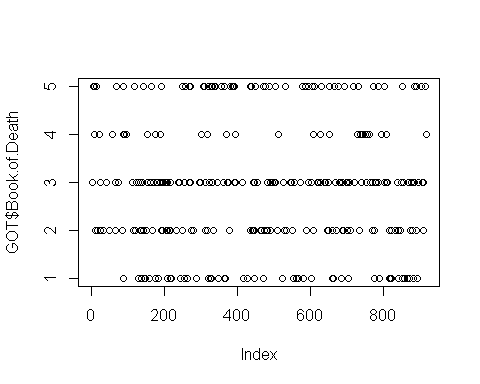
This is a quantitative stats. for all the characters and what year did they die in the show

hist( GOT$Death.Year,,xlab = "Year", ylab="No. of Death",col = "Red", border = "black" )



This is a scatter plot command to show which book of the show did the characters died.

plot(GOT$Book.of.Death)



Over here, I used the transform variable command and Since Game of Thrones started in year 297 as per the story timeline, the death year is 297 and onwards. I changed it to log10 which gave it the range from 1 to 10. Now the mean is 2.476 means on average, most deaths were in the first 25% of the show timeline.

summary(log10(GOT$Death.Year))

## Min. 1st Qu. Median Mean 3rd Qu. Max. NA's   
## 2.473 2.476 2.476 2.476 2.477 2.477 612