Jacob Dineen

IST 565

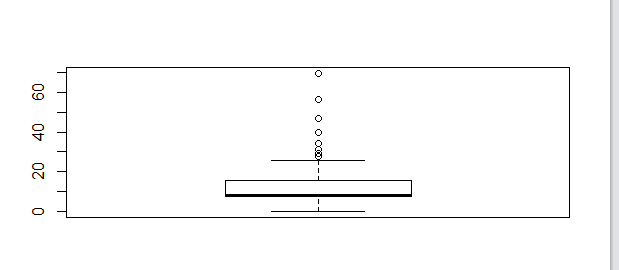
Show a box plot and histogram of "Fare" in third class (Pclass==3).

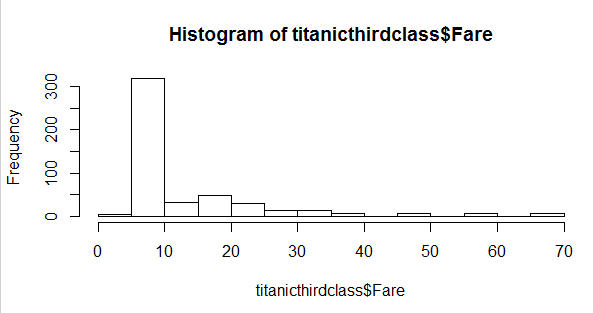
titanicthirdclass <- titanic[titanic$Pclass == 3,]

boxplot(titanicthirdclass$Fare)

hist(titanicthirdclass$Fare)

Looking at the boxplot of fares, filtering for only the third class, there aren’t any evident outliers, as all values fall between a range of $0-$70. Distortion in summary statistics is not likely.

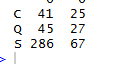




If we look at the relationship between embarked and survival, amongst the third class, we can see that people who embarked from S have a much higher rate of death than the other two ports.

t1 <- table(titanicthirdclass$Embarked, titanicthirdclass$Survived)

t1



If we look at the entire dataset (all three Pclasses), we can see that 372 out of the 549 recorded deaths come from people within the third P class.