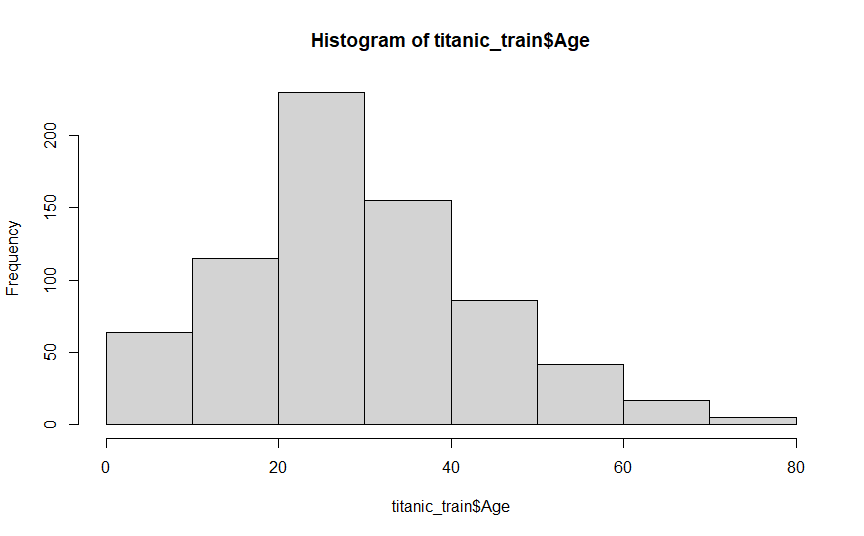
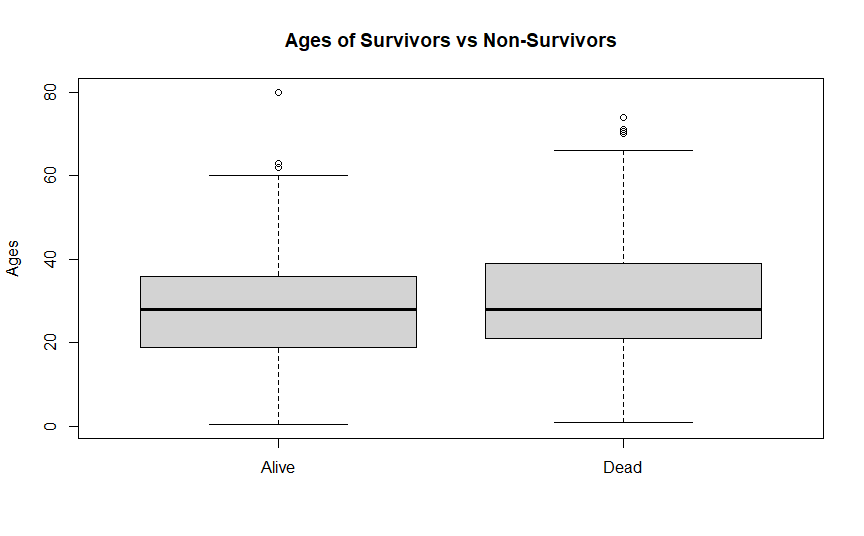
install.packages("titanic")  
library(titanic)  
  
  
###################################step 3##################################  
summary(titanic\_train)  
    
  
###################################step 4##################################  
?table  
titanic\_table <- table(titanic\_train$Survived, titanic\_train$Pclass)  
titanic\_table

      1   2   3  
  0  80  97 372  
  1 136  87 119

#1st class passengers were the only Pclass more likely to survive than not  
#there were less 2 Pclass passengers than either 1Pclass or 3Pclass  
  
  
###################################step 5##################################  
hist(titanic\_train$Age)

  
#younger passengers outlived older passengers  
#is that due to age or due to number of passengers within each age group?  
  
  
###################################step 6##################################  
?boxplot  
help(bx)  
#[https://www.programmingr.com/examples/how-to-make-a-side-by-side-boxplot-in-r/](about:blank)  
Alive = titanic\_train$Age[which(titanic\_train$Survived==1)]   
Dead = titanic\_train$Age[which(titanic\_train$Survived==0)]    
my.bp <- boxplot(Alive,Dead,  
          main = "Ages of Survivors vs Non-Survivors",  
          ylab = "Ages",  
          names = c("Alive", "Dead"))  
my.bp

  
#there appears to be no difference in the age distribution of who survived