Dillon Steiger Assignment 1

2025-09-04

rm(list=ls())  
url<- "https://raw.githubusercontent.com/mwaskom/seaborn-data/master/titanic.csv"  
titanic<-read.csv(url) #saving data into df called "titanic"  
head(titanic)

## survived pclass sex age sibsp parch fare embarked class who  
## 1 0 3 male 22 1 0 7.2500 S Third man  
## 2 1 1 female 38 1 0 71.2833 C First woman  
## 3 1 3 female 26 0 0 7.9250 S Third woman  
## 4 1 1 female 35 1 0 53.1000 S First woman  
## 5 0 3 male 35 0 0 8.0500 S Third man  
## 6 0 3 male NA 0 0 8.4583 Q Third man  
## adult\_male deck embark\_town alive alone  
## 1 True Southampton no False  
## 2 False C Cherbourg yes False  
## 3 False Southampton yes True  
## 4 False C Southampton yes False  
## 5 True Southampton no True  
## 6 True Queenstown no True

#3.  
#quantitative vars  
summary(titanic$age)

## Min. 1st Qu. Median Mean 3rd Qu. Max. NA's   
## 0.42 20.12 28.00 29.70 38.00 80.00 177

summary(titanic$fare)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.00 7.91 14.45 32.20 31.00 512.33

#categorical vars  
table(titanic$sex)

##   
## female male   
## 314 577

table(titanic$class)

##   
## First Second Third   
## 216 184 491

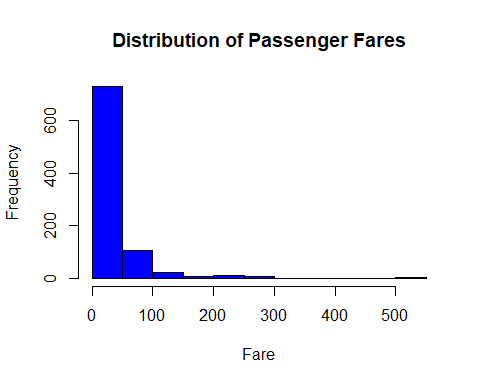
table(titanic$survived)

##   
## 0 1   
## 549 342

#4.  
titanic$is\_child<- ifelse(titanic$age < 18, 1, 0)  
#creates new var if the passenger is a child (if they are <18 yrs old)  
table(titanic$is\_child)

##   
## 0 1   
## 601 113

#5.  
#histogram of fare  
hist(titanic$fare, main="Distribution of Passenger Fares", xlab="Fare", col="blue")



#scatterplot of age vs fare  
plot(titanic$age, titanic$fare, main="Scatterplot of Age vs Fare", xlab="Age", ylab="Fare", col="blue", pch=19)

