64060 Assignment 1

Steven Kalinoff

2023-01-29

# 1/2 for the first assignment I choose to use a built in data set.  
data(mtcars)  
head(mtcars)

## mpg cyl disp hp drat wt qsec vs am gear carb  
## Mazda RX4 21.0 6 160 110 3.90 2.620 16.46 0 1 4 4  
## Mazda RX4 Wag 21.0 6 160 110 3.90 2.875 17.02 0 1 4 4  
## Datsun 710 22.8 4 108 93 3.85 2.320 18.61 1 1 4 1  
## Hornet 4 Drive 21.4 6 258 110 3.08 3.215 19.44 1 0 3 1  
## Hornet Sportabout 18.7 8 360 175 3.15 3.440 17.02 0 0 3 2  
## Valiant 18.1 6 225 105 2.76 3.460 20.22 1 0 3 1

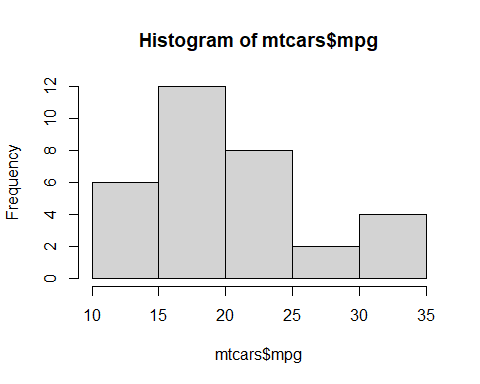
mtcars$model = rownames(mtcars)  
# 3 basic descriptive stats.   
  
summary(mtcars$mpg)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 10.40 15.43 19.20 20.09 22.80 33.90

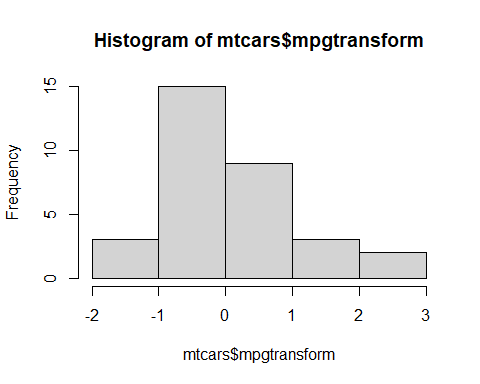
summary(mtcars$cyl)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 4.000 4.000 6.000 6.188 8.000 8.000

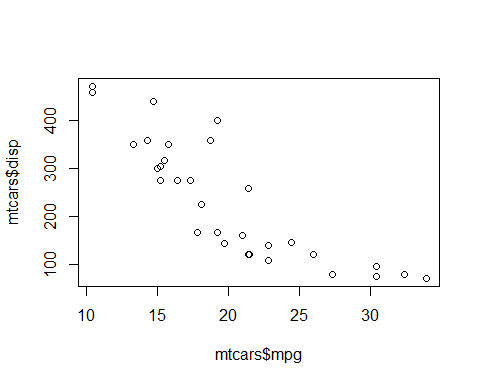
# normal transform  
mtcars$mpgtransformed = (mtcars$mpg - mean(mtcars$mpg))/sd(mtcars$mpg)  
  
# mtcars  
# 4 transforming mpg to look at values in an absolute sense.   
  
hist(mtcars$mpg, breaks = 5)



hist(mtcars$mpgtransform, breaks = 5)



# 5 we can see by the transform the data and its spread break out differently from the initial plot.  
  
plot(mtcars$mpg,mtcars$disp)



cor(mtcars$mpg,mtcars$disp)

## [1] -0.8475514

#from the graph and correlation between the 2 variables we can see that the mpg and disp (displacement) are pretty strongly inversely related.