Math 170S HW4

Jun Ryu

2023-02-10

```
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
```

(a)

Run summaries for variables mpg, cyl, disp, hp, drat and wt:

```
df <- mtcars

summary(df$mpg)
```

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

```
summary(df$cyl)
```

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

```
summary(df$disp)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    71.1   120.8   196.3   230.7   326.0   472.0
```

```
summary(df$hp)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    52.0    96.5   123.0   146.7   180.0   335.0
```

```
summary(df$drat)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   2.760   3.080   3.695   3.597   3.920   4.930
```
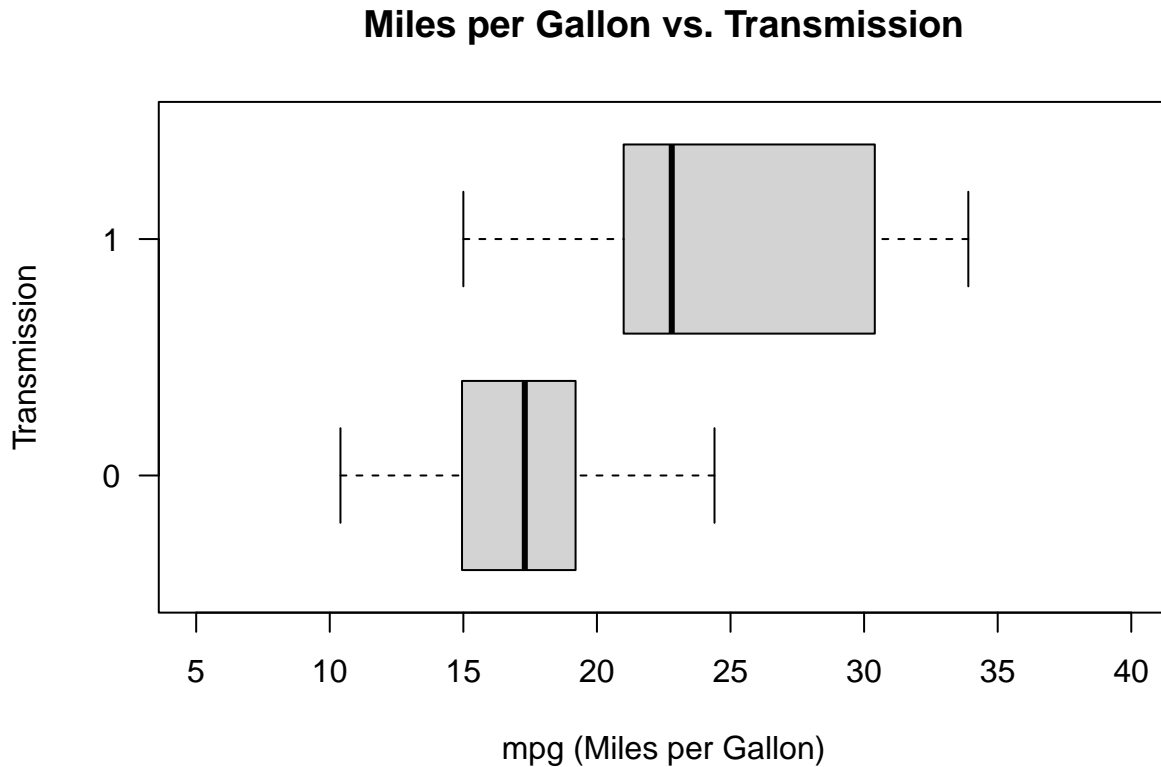
```
summary(df$wt)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.513   2.581   3.325   3.217   3.610   5.424
```

(b)

Generate and compare the boxplots for mpg against, am i.e. Transmission (0 = automatic, 1 = manual):

```r
boxplot(df$mpg ~ df$am, horizontal = T, xlab = "mpg (Miles per Gallon)",
        ylab = "Transmission", las = 1, ylim = c(5,40),
        main = "Miles per Gallon vs. Transmission")
```


The above boxplot shows that cars with manual transmission (1) generally had a better miles per gallon than cars with automatic transmission (0). All 5 summary statistics (min, Q1, median, Q3, and max) show a higher value for manual transmission.

(c)

Generate the scatter plot between variables mpg against cyl:

```r
plot(df$cyl, df$mpg, xlab = "Number of Cylinders", ylab = "mpg (Miles per Gallon)",
     main = "Miles per Gallon vs. Number of Cylinders", ylim = c(5,40))
```

Miles per Gallon vs. Number of Cylinders
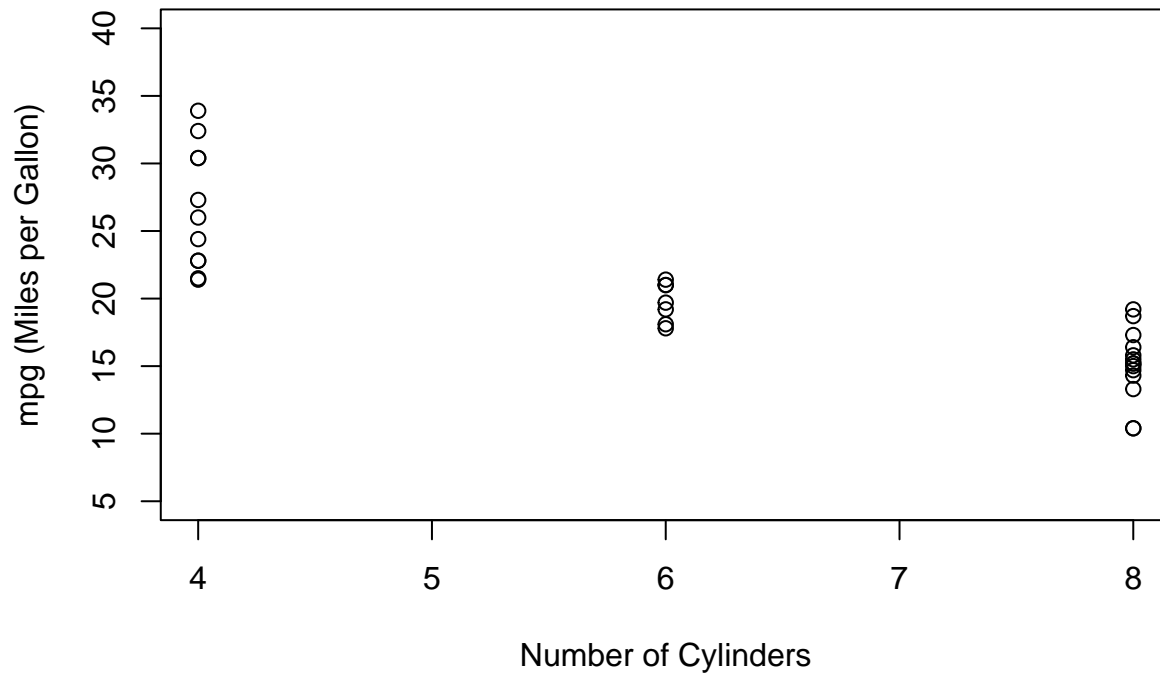
(d)

Compute the correlation coefficient of variables mpg and cyl:

```
cor(df$mpg, df$cyl)
```

```
## [1] -0.852162
```

(e)

Perform a linear regression using mpg as response variable and disp as exploratory variable:

```
SLR <- lm(df$mpg ~ df$disp)
SLR
```

```
##
## Call:
## lm(formula = df$mpg ~ df$disp)
##
## Coefficients:
## (Intercept)      df$disp
##    29.59985     -0.04122
```

```
plot(df$mpg ~ df$disp, xlim = c(0,600), ylim = c(0,40),
     xlab = "Displacement (cubic inch)", ylab = "mpg (Miles per Gallon)",
     xaxs = "i", yaxs = "i", main = "Miles per Gallon vs. Displacement")
abline(SLR)
mtext("Regression Line: mpg = -0.04122 disp + 29.59985", line = -2)
```

Miles per Gallon vs. Displacement

Regression Line: mpg = −0.04122 disp + 29.59985

mpg (Miles per Gallon)

Displacement (cubic inch)