

Mtcars_DataAnalysis

Himanshi

2025-08-08

Project Overview

This project analyzes the `mtcars` dataset, which contains performance and design specifications for 32 car models from 1974 Motor Trend magazine.

The goal is to explore relationships between variables, compare groups, and build a simple regression model to understand what affects fuel efficiency (miles per gallon, `mpg`).

Loading the Data

```
data(mtcars)
```

```
# Basic overview
```

```
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
```

```
summary(mtcars)
```

```
##           mpg           cyl           disp           hp
##  Min.      :10.40  Min.      :4.000  Min.       : 71.1  Min.       : 52.0
## 1st Qu.:15.43  1st Qu.:4.000  1st Qu.:120.8  1st Qu.: 96.5
## Median :19.20  Median :6.000  Median :196.3  Median :123.0
## Mean   :20.09  Mean   :6.188  Mean   :230.7  Mean   :146.7
## 3rd Qu.:22.80  3rd Qu.:8.000  3rd Qu.:326.0  3rd Qu.:180.0
## Max.    :33.90  Max.    :8.000  Max.    :472.0  Max.    :335.0
##           drat           wt           qsec           vs
##  Min.       :2.760  Min.       :1.513  Min.       :14.50  Min.       :0.0000
## 1st Qu.:3.080  1st Qu.:2.581  1st Qu.:16.89  1st Qu.:0.0000
## Median :3.695  Median :3.325  Median :17.71  Median :0.0000
## Mean   :3.597  Mean   :3.217  Mean   :17.85  Mean   :0.4375
```

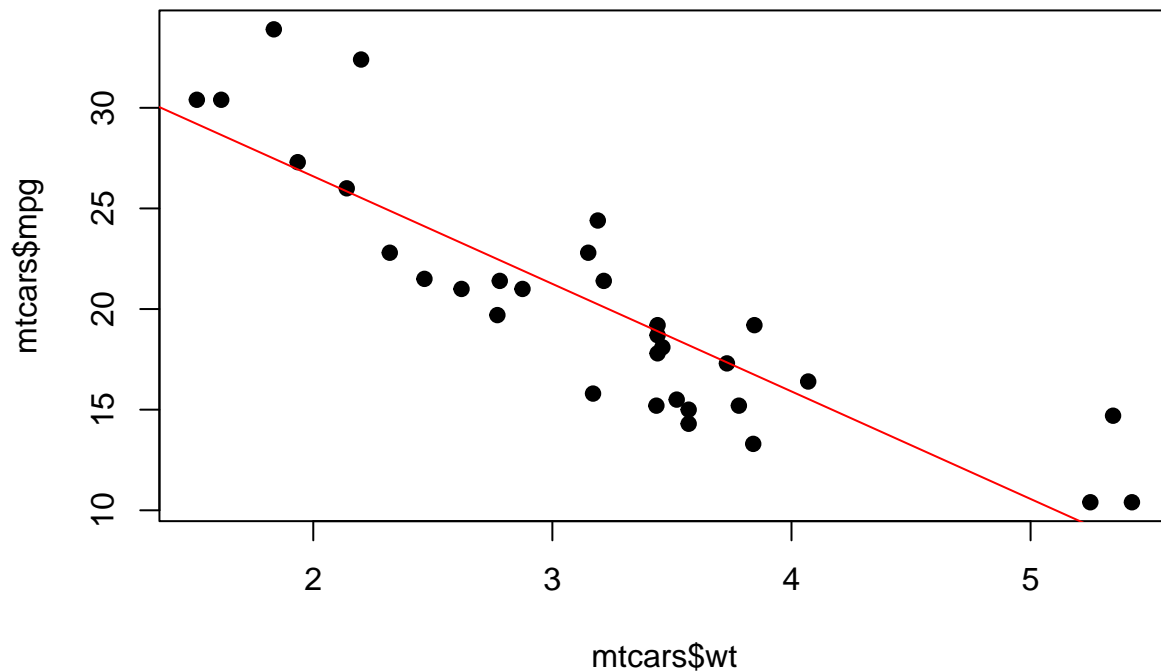
```
## 3rd Qu.:3.920 3rd Qu.:3.610 3rd Qu.:18.90 3rd Qu.:1.0000
## Max. :4.930 Max. :5.424 Max. :22.90 Max. :1.0000
##      am      gear      carb
## Min. :0.0000 Min. :3.000 Min. :1.000
## 1st Qu.:0.0000 1st Qu.:3.000 1st Qu.:2.000
## Median :0.0000 Median :4.000 Median :2.000
## Mean :0.4062 Mean :3.688 Mean :2.812
## 3rd Qu.:1.0000 3rd Qu.:4.000 3rd Qu.:4.000
## Max. :1.0000 Max. :5.000 Max. :8.000
```

Regression

```
model <- lm(mpg ~ wt + hp, data=mtcars)
summary(model)
```

```
##
## Call:
## lm(formula = mpg ~ wt + hp, data = mtcars)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.941 -1.600 -0.182  1.050  5.854
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 37.22727    1.59879   23.285 < 2e-16 ***
## wt          -3.87783    0.63273   -6.129 1.12e-06 ***
## hp           -0.03177    0.00903   -3.519 0.00145 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.593 on 29 degrees of freedom
## Multiple R-squared:  0.8268, Adjusted R-squared:  0.8148
## F-statistic: 69.21 on 2 and 29 DF, p-value: 9.109e-12
```

```
plot(mtcars$wt, mtcars$mpg, pch=19)
abline(lm(mpg ~ wt, data=mtcars), col="red")
```



Correlation Heatmap

```
install.packages("corrplot", repos = "https://cloud.r-project.org")
```

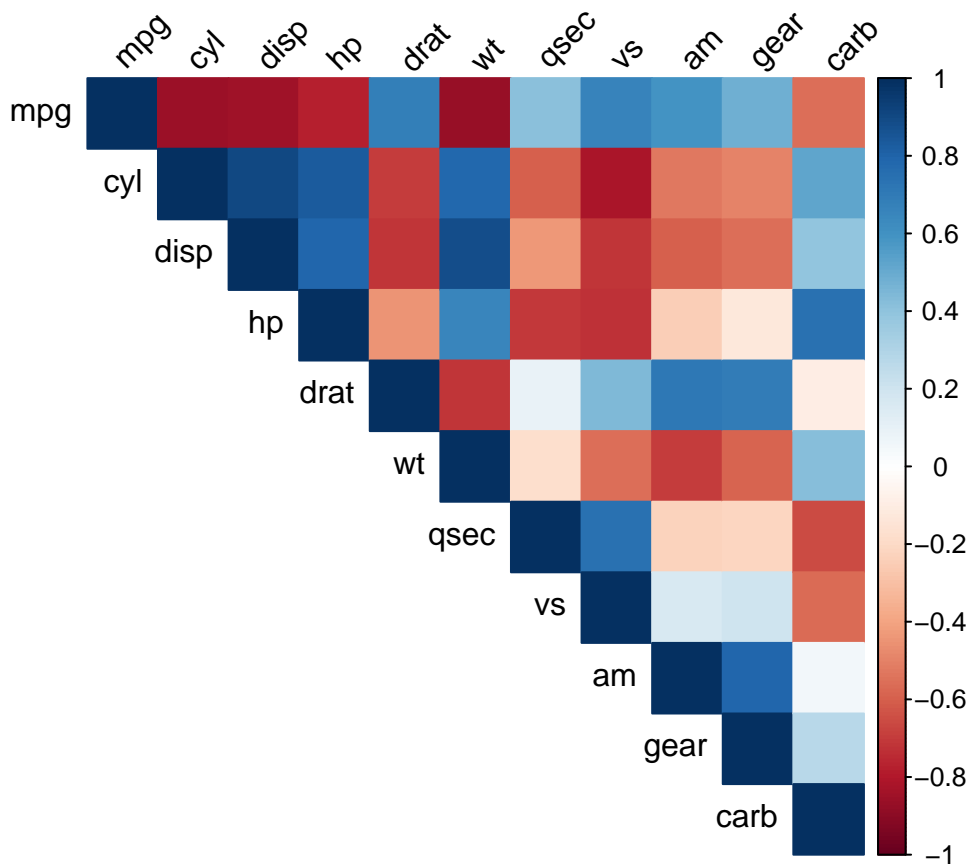
```
## Installing package into 'C:/Users/dkvh2/AppData/Local/R/win-library/4.5'  
## (as 'lib' is unspecified)
```

```
## package 'corrplot' successfully unpacked and MD5 sums checked  
##  
## The downloaded binary packages are in  
## C:\Users\dkvh2\AppData\Local\Temp\Rtmp8EVwoJ\downloaded_packages
```

```
library(corrplot)
```

```
## corrplot 0.95 loaded
```

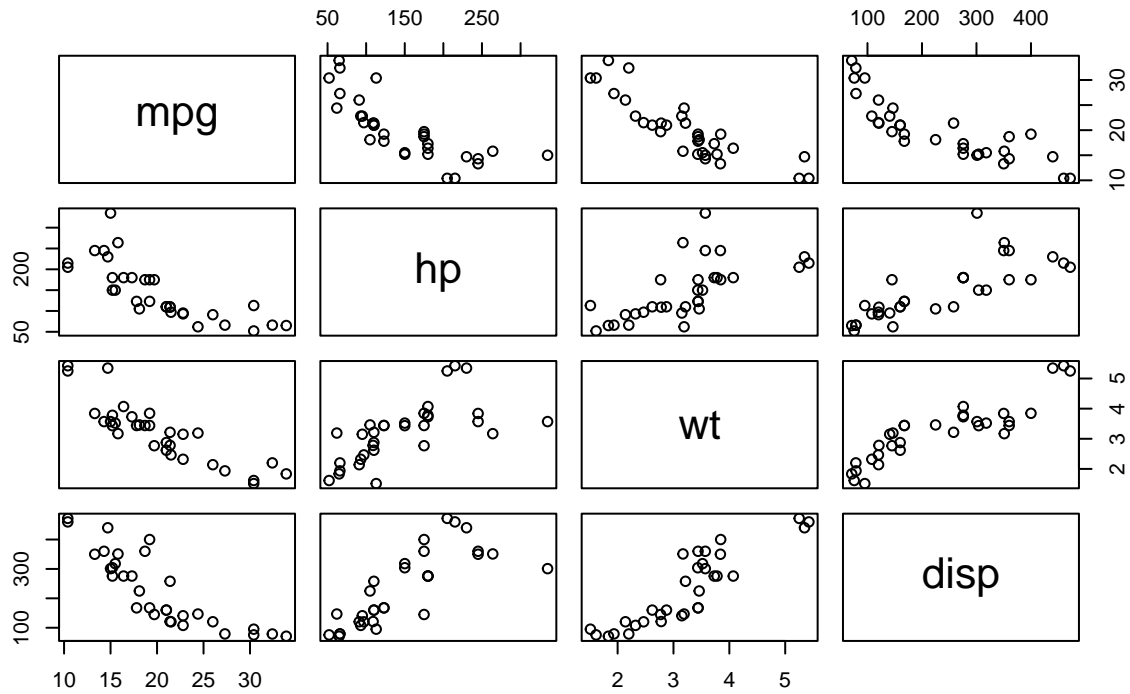
```
corrplot(cor(mtcars), method="color", type="upper", tl.col="black", tl.srt=45)
```



Pairwise Scatter Plots

```
pairs(~ mpg + hp + wt + disp, data=mtcars,  
      main="Pairwise Plots of Selected Variables")
```

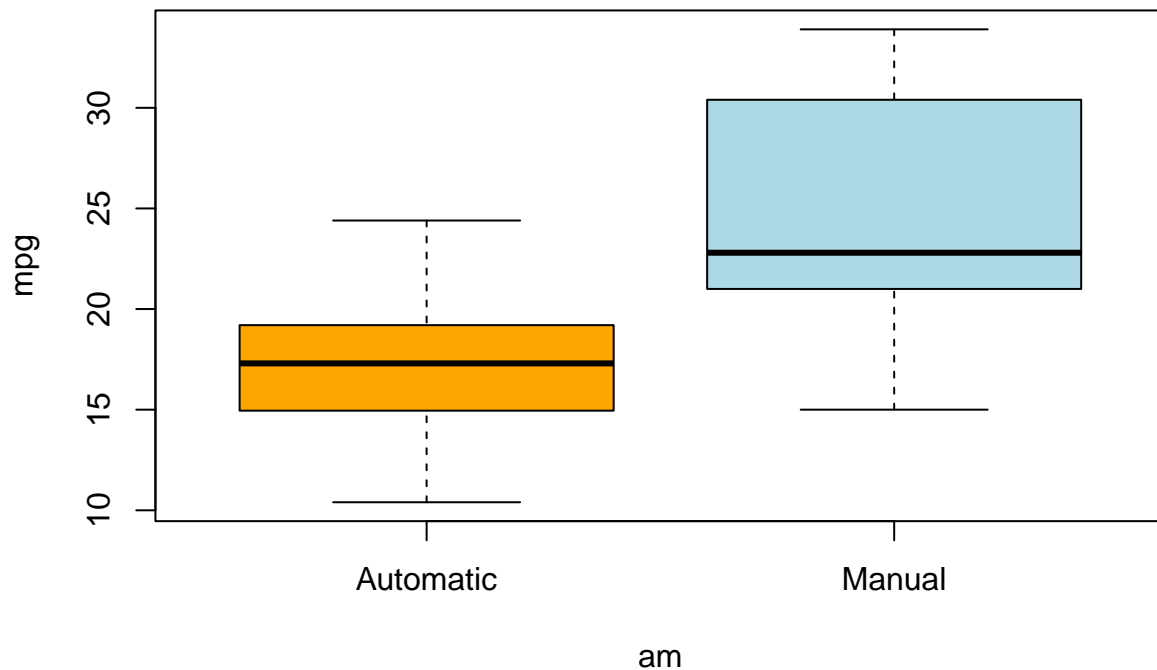
Pairwise Plots of Selected Variables



BoxPlots by Category

```
boxplot(mpg ~ am, data=mtcars,
        names=c("Automatic", "Manual"),
        col=c("orange", "lightblue"),
        main="MPG by Transmission Type")
```

MPG by Transmission Type



Multiple linear regression

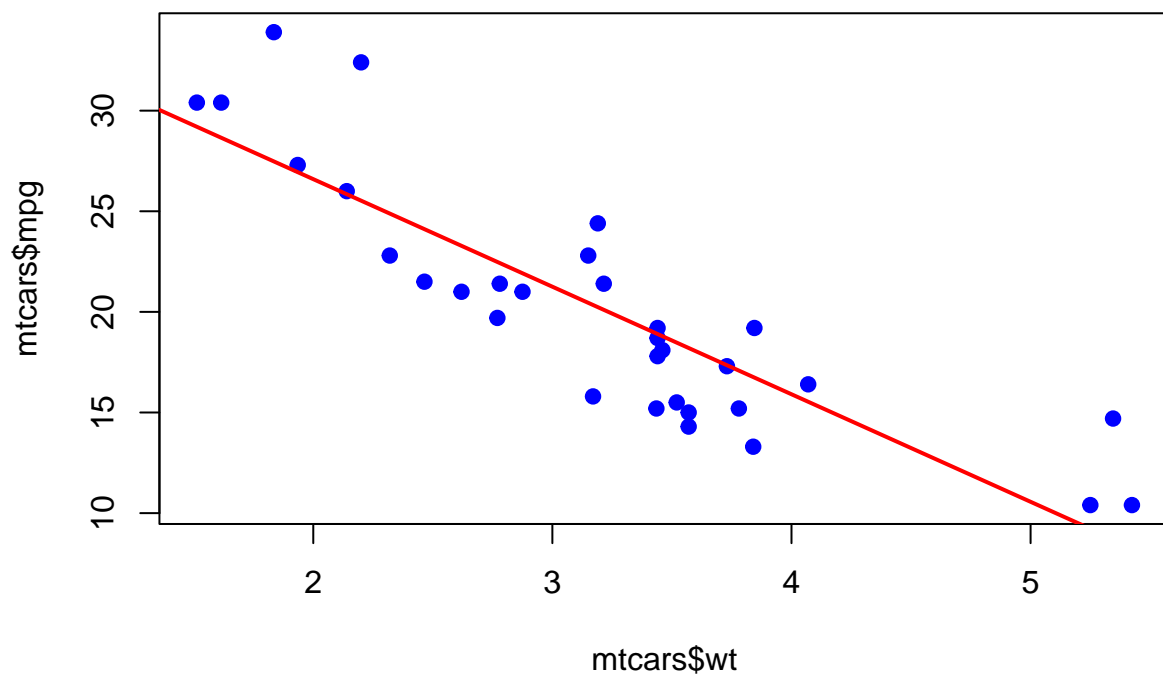
```
model <- lm(mpg ~ wt + hp, data=mtcars)
summary(model)
```

```
##
## Call:
## lm(formula = mpg ~ wt + hp, data = mtcars)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.941  -1.600  -0.182   1.050   5.854
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  37.22727    1.59879   23.285  < 2e-16 ***
## wt          -3.87783    0.63273   -6.129  1.12e-06 ***
## hp           -0.03177    0.00903   -3.519  0.00145 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.593 on 29 degrees of freedom
## Multiple R-squared:  0.8268, Adjusted R-squared:  0.8148
```

```
## F-statistic: 69.21 on 2 and 29 DF, p-value: 9.109e-12
```

Visualizing fit for weight

```
plot(mtcars$wt, mtcars$mpg, pch=19, col="blue")  
abline(lm(mpg ~ wt, data=mtcars), col="red", lwd=2)
```



Residual diagnostics

```
par(mfrow=c(2,2))  
plot(model)
```

