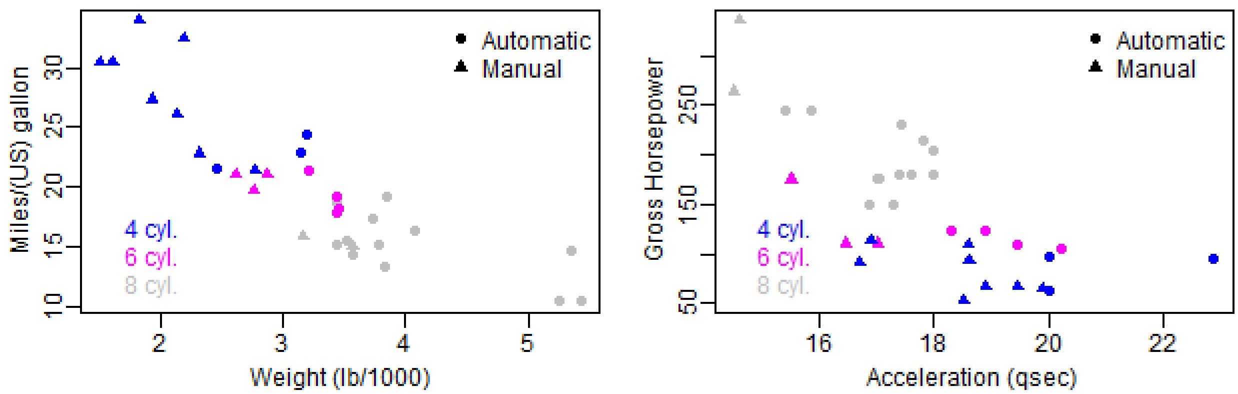
Transmission Effects On MPG In Cars

1. Executive Summary

This paper is based at 1974 Motor Trend Data that contains fuel consumption related with 10 aspects of automobile design and performance for 32 cars models 1973-74. The purpose is to estimate the effect of the type of transmission (automatic vs manual) on the miles per gallon (MPG). The results suggest that manual transmission have a higher MPG. However the weight and acceleration are attibutes that have a siginficant influence in the results.

2. Exploratory Data Analysis

Based on common sense we can assume that the mpg of a car depends on its weight and its power, also the cars with more cylinders have a great acceleration and more horse power. The following plot shows these theories.



The left plot offer following insights:

- The cars with automatic transmission are heavy.
- The cars with manual transmission are lightweight.
- A positive relationship between number of cylinders and weight.
- A negative relationship between MPG and weight.

The right plot offer following insights:

- A strong relationship between acceleration and horse power.
- Cars with more cylinders has great acceleration in lest time.
- Cars with more cylinders provides major horsepower.
- Cars with more horsepower have automatic transmission.

Also see the correlation matrix in the Appendix.

3. Models

The exploratory analysis suggested to use the wt variable with some power variables (cyl, hp) is a good start. An interesting insight is the acceleration that is strongly influenced by 8 cylinders cars therefore we will evaluate it. The models to analyze are:

```
## mpg ~ am:wt + am:cyl
```

```
## mpg ~ am:wt + am:hp
```

```
## mpg ~ am:wt + am:qsec
```

See the Appendix for results of model 1 and model 2. The stimation results reveal that model 3 is the best fit according following:

```
summary(model3)
```

```
##
## Call:
## lm(formula = mpg ~ am:wt + am:qsec, data = mtcarst)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -3.936 -1.402 -0.155  1.269  3.886
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)       13.969      5.776    2.42   0.0226 *
## amAutomatic:wt    -3.176      0.636   -4.99  3.1e-05 ***
## amManual:wt       -6.099      0.969   -6.30  9.7e-07 ***
## amAutomatic:qsec   0.834      0.260    3.20   0.0035 **
## amManual:qsec      1.446      0.269    5.37  1.1e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.1 on 27 degrees of freedom
## Multiple R-squared:  0.895,  Adjusted R-squared:  0.879
## F-statistic: 57.3 on 4 and 27 DF,  p-value: 8.42e-13
```

3.1 Model Validation

The model has a 88% Adjusted R-Squared, and contains three variables (transmission type, weight, and acceleration) along with the transmission interaction. The residuals are normal with 0 mean and constant variance. The “Residual vs Fitted” and “Scale Location” charts show that there is no trend to the residuals. The Q-Q plot shows that the errors are aproximately distributed. The p-values for am are considered statistically significant.

3.2 Model Interpretation

For constant quarter-mile-time: For each unit of increase in quarter-mile-time (1 unit = 1 sec), there is a 0.834 mpg increase for automatic cars while there is a 1.446 mpg increase for manual cars. This means manual cars have more efficient acceleration than automatic but in both cases require more fuel.

For constant weight (lb/1000): For each unit of increase in weight (1 unit = 1000 lb) there is 3.176 decrease for automatic cars while there is a 6.099 mpg decrease for manual cars. This means heavier cars in general require more fuel altought automatic cars consumes less fuel than manual.

Appendix

Data

The data frame consists of 32 observations on 11 variables.

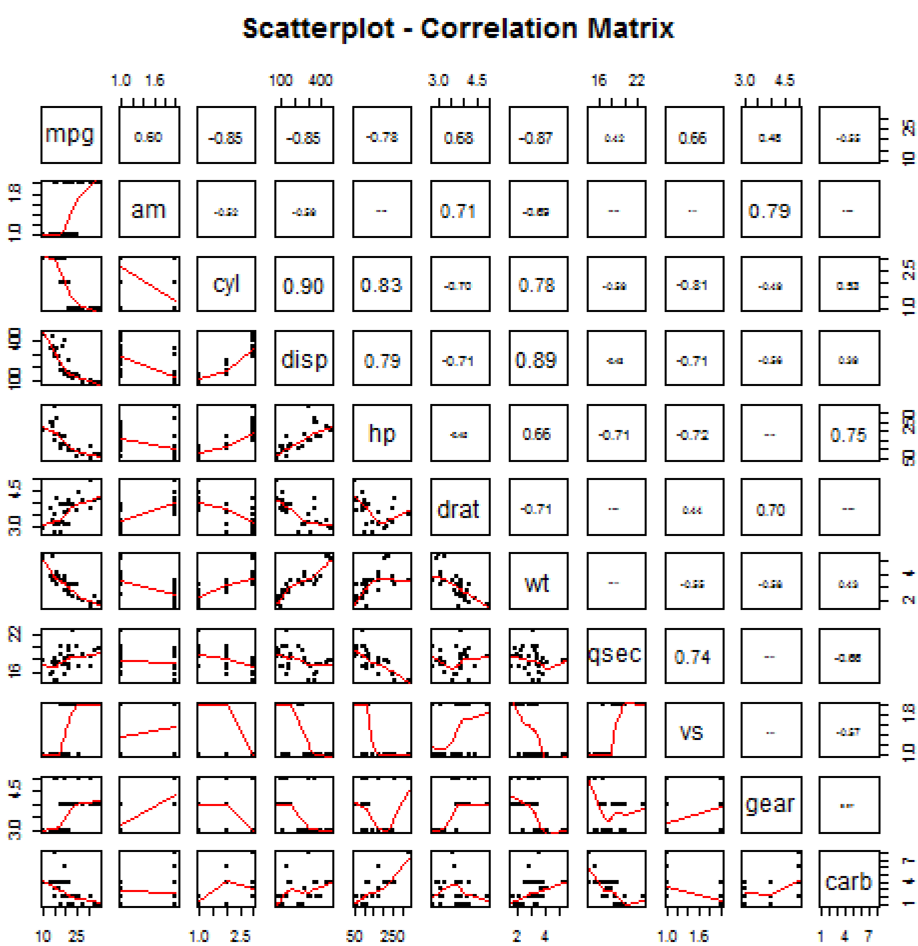
- mpg: Miles/(US) gallon
- cyl: Number of cylinders
- disp: Displacement (cu.in.)
- hp: Gross horsepower
- drat: Rear axle ratio
- wt: Weight (lb/1000)
- qsec: ¼ mile time
- vs: V/S
- am: Transmission (0 = automatic, 1 = manual)
- gear: Number of forward gears

```
##                      mpg cyl disp  hp drat    wt  qsec      vs        am
```

```
## Mazda RX4         21.0   6  160 110 3.90 2.620 16.46 Vengine    Manual
## Mazda RX4 Wag     21.0   6  160 110 3.90 2.875 17.02 Vengine    Manual
## Datsun 710        22.8   4  108  93 3.85 2.320 18.61 Iengine    Manual
## Hornet 4 Drive    21.4   6  258 110 3.08 3.215 19.44 Iengine Automatic
## Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02 Vengine Automatic
## Valiant           18.1   6  225 105 2.76 3.460 20.22 Iengine Automatic
##                   gear carb
## Mazda RX4            4    4
## Mazda RX4 Wag        4    4
## Datsun 710           4    1
## Hornet 4 Drive       3    1
## Hornet Sportabout    3    2
## Valiant              3    1
```

MPG against all variables

```
require(graphics)
panel.cor <- function(x, y, digits = 2, prefix = "", cex.cor, ...) {
    usr <- par("usr")
    on.exit(par(usr))
    par(usr = c(0, 1, 0, 1))
    r <- cor(x, y)
    txt <- format(c(r, 0.123456789), digits = digits)[1]
    txt <- paste(prefix, txt, sep = "")
    if (missing(cex.cor))
        cex.cor <- 0.8/strwidth(txt)
    text(0.5, 0.5, txt, cex = cex.cor * abs(r))
}
pairs(mpg ~ am:., data = mtcarst, lower.panel = panel.smooth, upper.panel = panel.cor,
    pch = 20, main = "Scatterplot - Correlation Matrix")
```



Models

Model 1

```
summary(model1)
```

```
##
## Call:
## lm(formula = mpg ~ am:wt + am:cyl, data = mtcarst)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -4.133 -1.223 -0.173  1.272  5.550
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)       33.820      2.731   12.38  3.7e-12 ***
## amAutomatic:wt    -3.409      0.867   -3.93  0.00059 ***
## amManual:wt       -2.981      1.363   -2.19  0.03831 *
## amAutomatic:cyl6  -3.143      2.052   -1.53  0.13822
## amManual:cyl6     -5.041      2.118   -2.38  0.02524 *
## amAutomatic:cyl8  -4.780      2.054   -2.33  0.02834 *
## amManual:cyl8     -8.374      2.882   -2.91  0.00757 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.63 on 25 degrees of freedom
## Multiple R-squared:  0.847,  Adjusted R-squared:  0.81
## F-statistic:   23 on 6 and 25 DF,  p-value: 4.82e-09
```

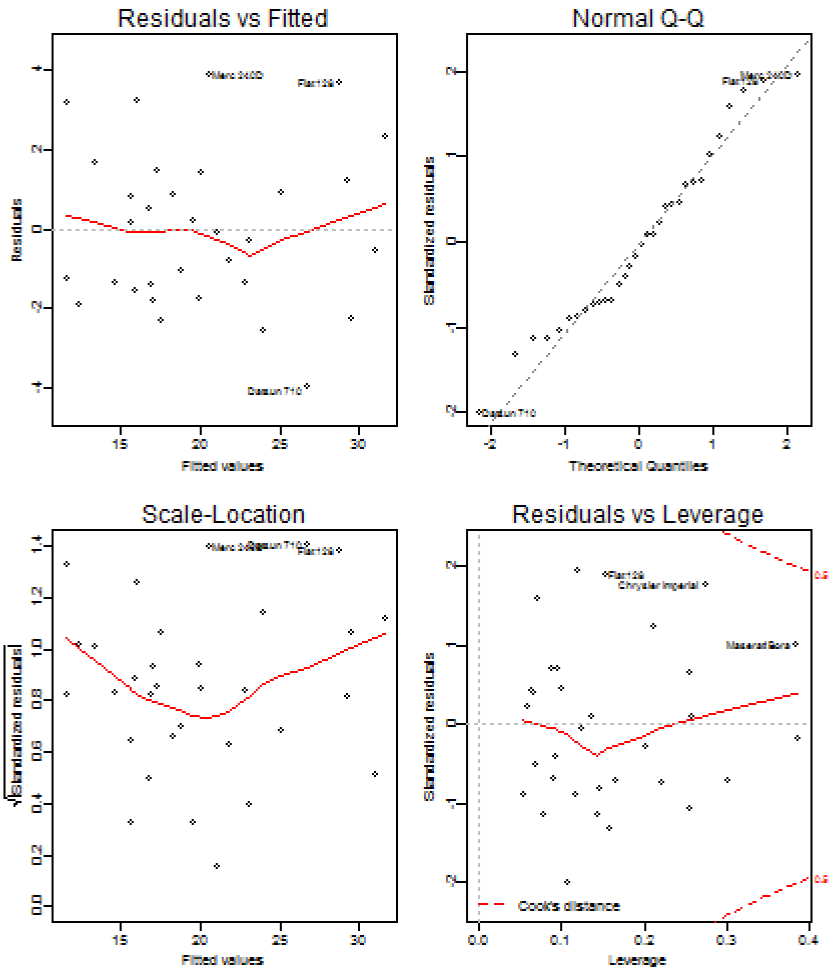
Model 2

```
summary(model2)
```

```
##
## Call:
## lm(formula = mpg ~ am:wt + am:hp, data = mtcarst)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -3.679 -2.066 -0.305  0.941  5.878
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)     36.2167     2.4096   15.03  1.2e-14 ***
## amAutomatic:wt  -3.3682     0.9579   -3.52   0.0016 **
## amManual:wt     -3.2964     1.4861   -2.22   0.0351 *
## amAutomatic:hp  -0.0385     0.0158   -2.43   0.0221 *
## amManual:hp     -0.0330     0.0141   -2.35   0.0265 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.66 on 27 degrees of freedom
## Multiple R-squared:  0.83,   Adjusted R-squared:  0.805
## F-statistic:   33 on 4 and 27 DF,  p-value: 4.91e-10
```

Selected Model Plot

```
par(mfrow = c(2, 2), cex.main = 0.1, cex = 0.6, mgp = c(1.7, 0.5, 0), mar = c(4,
    3.5, 2, 0.5), oma = c(0.5, 10, 0.5, 0.5))
plot(model3)
```



Residuals vs Fitted
Normal Q-Q
Scale-Location
Residuals vs Leverage
Fitted values
Residuals
Theoretical Quantiles
Standardized residuals
Leverage
Cook's distance
Maserati Bora
Chrysler Imperial