Impact of Transmission Type on MPG of Cars in 1974

Pete Petersen III

6/23/2020

Abstract

In this paper we will analyze the relationships of the 10 variables present in the MTCARS data set in order to make a determination as to wether or not manual versus automatic transmission are more fuel efficient once the uncorrelated regressors are removed from the model. We conclude that the data alludes to the reasonable conclusion that manual transmision likely increases fuel efficiency overall. Unfortunately, we don't have enough evidence from the data at hand to say there is a significant true difference or to make stonger claims. It is likely that the data is confounded by a much higher concentration of displacements at the low end of the manual types.

Exploratory Data Analysis

The data was extracted from the 1974 Motor Trend US magazine, and comprises fuel consumption and 10 aspects of automobile design and performance for 32 automobiles (1973–74 models). We begin our analysis by exploring nature of the data in (MTCARS) in order to familiarize ourselves with the size and properties of the data. Upon inspection, we notice several numerical vectors of data that would probably serve best for modeling if we covert them to factors.

```
require(graphics)

## possibly more meaningful, e.g., for summary() or bivariate plots:
mtcars2 <- within(mtcars, {
   vs <- factor(vs, labels = c("V", "S"))
   am <- factor(am, labels = c("automatic", "manual"))
   cyl  <- ordered(cyl)
   gear <- ordered(gear)
   carb <- ordered(carb)
})
```

Data Visualization

The next step in our analysis included a pairwise plot of our regressors (see Apendix II) in order to determine which ones seem to be correlated with MPG. We note that cyl, hp, wt, and am each visiably demonstrate a correlation with our outcome variable MPG.

EDA Plots

The coplot below seems to support aur thesis that manual transmissions are more fuel efficient. Although it is interesting to note that manual transmisions are associated with smaller displacement engines. This is a fact we will need to be cautious of when we make inferences from our final model. The dot colors in the plot represent 4-cyl, 6-cyl, and 8-cyl as black, red, and green respectively.

```
coplot(mpg ~ disp | as.factor(am) , data = mtcars2,
       panel = panel.smooth, rows = 1, col = mtcars2$cyl)
legend(7,4.3,unique(mtcars2$cyl),col=1:length(mtcars2$cyl),pch=1)
```

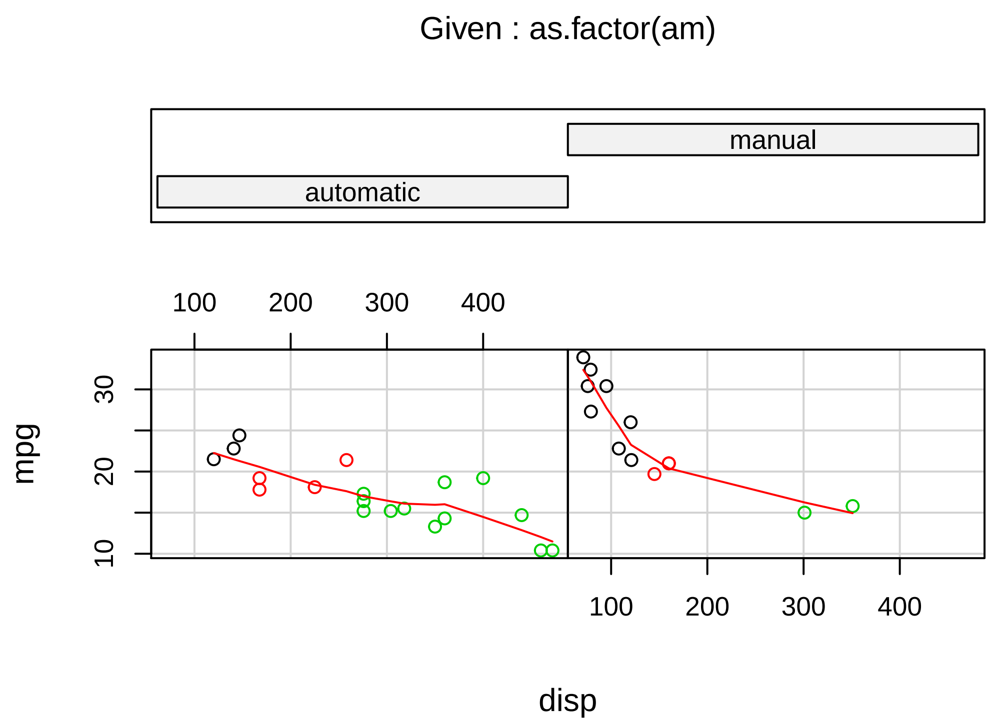


Figure 1: Figure 1

The Boxplots seem to also indicate that there is some correlation of MPG with manual transmission type.

```
boxplot(mpg ~ am, data =  mtcars2, col="beige")
```

Build Models

We first constructed a model that predicts MPG from all columns of data (Model.All). We performed a backword stepping analysis (see Apendix III for details) that resulted in Model.Best.Fit. This model was very much aligned with our expectations.

```
Model.All <- lm(mpg ~ ., mtcars2)
Model.BestFit <- lm( mpg ~  hp + wt + am + cyl , data = mtcars2)
summary(Model.BestFit)
```

```
##
## Call:
## lm(formula = mpg ~ hp + wt + am + cyl, data = mtcars2)
##
```

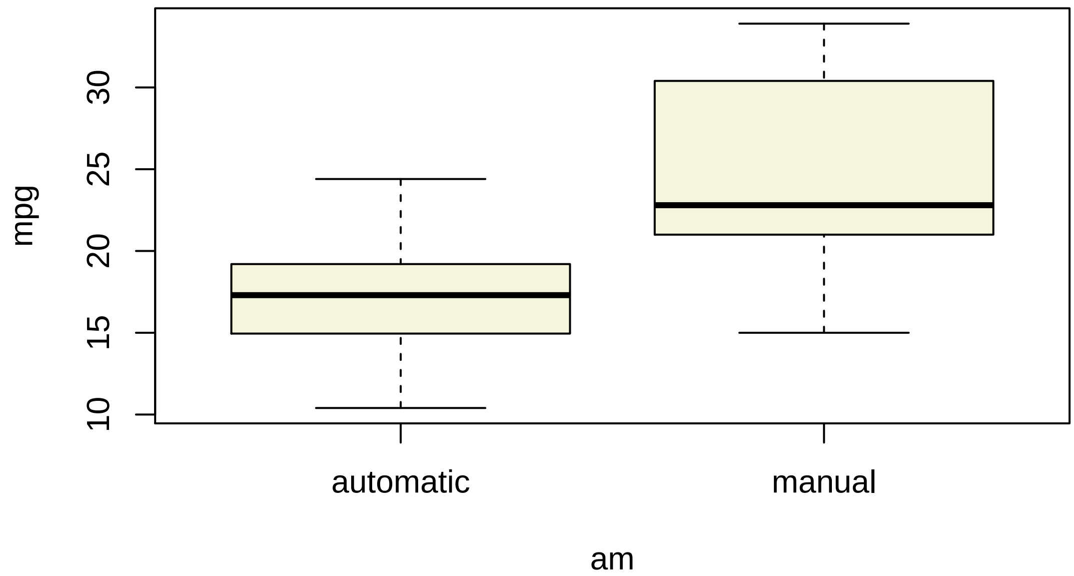


Figure 2: Figure 2

```
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.9387 -1.2560 -0.4013  1.1253  5.0513
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 31.97665    3.06337  10.438 8.61e-11 ***
## hp          -0.03211    0.01369  -2.345  0.02693 *
## wt          -2.49683    0.88559  -2.819  0.00908 **
## ammanual     1.80921    1.39630   1.296  0.20646
## cyl.L       -1.52995    1.61521  -0.947  0.35225
## cyl.Q        1.59177    0.88076   1.807  0.08231 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.41 on 26 degrees of freedom
## Multiple R-squared:  0.8659, Adjusted R-squared:  0.8401
## F-statistic: 33.57 on 5 and 26 DF,  p-value: 1.506e-10
```

```
anova(Model.All,Model.BestFit)
```

```
## Analysis of Variance Table
##
## Model 1: mpg ~ cyl + disp + hp + drat + wt + qsec + vs + am + gear + carb
## Model 2: mpg ~ hp + wt + am + cyl
##   Res.Df    RSS  Df Sum of Sq      F Pr(>F)
## 1     15 120.40
## 2     26 151.03 -11   -30.623 0.3468 0.9588
```

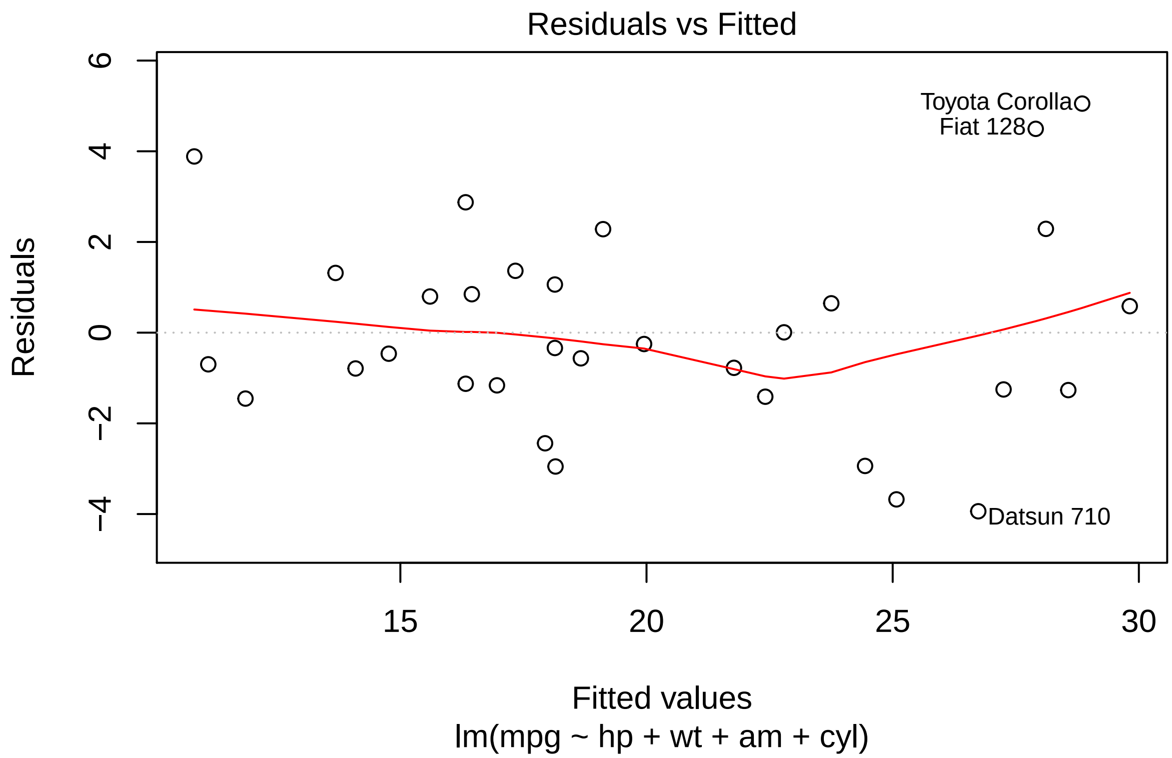
Model Interpretation

Based on the model, the manual transmision adds about 1.8 MPG over automatic transmission types although this is with a troubling 0.2 p-value. Therefore, under the terms of this model, we can not rule out the the Null hypothesis that there is no difference in MPG based on trnsmission type. Because this result is so counter to actually experience, we assume that there is either a sampling problem or we are missing an important predicter and biasing our coeficient.

Error Analysis

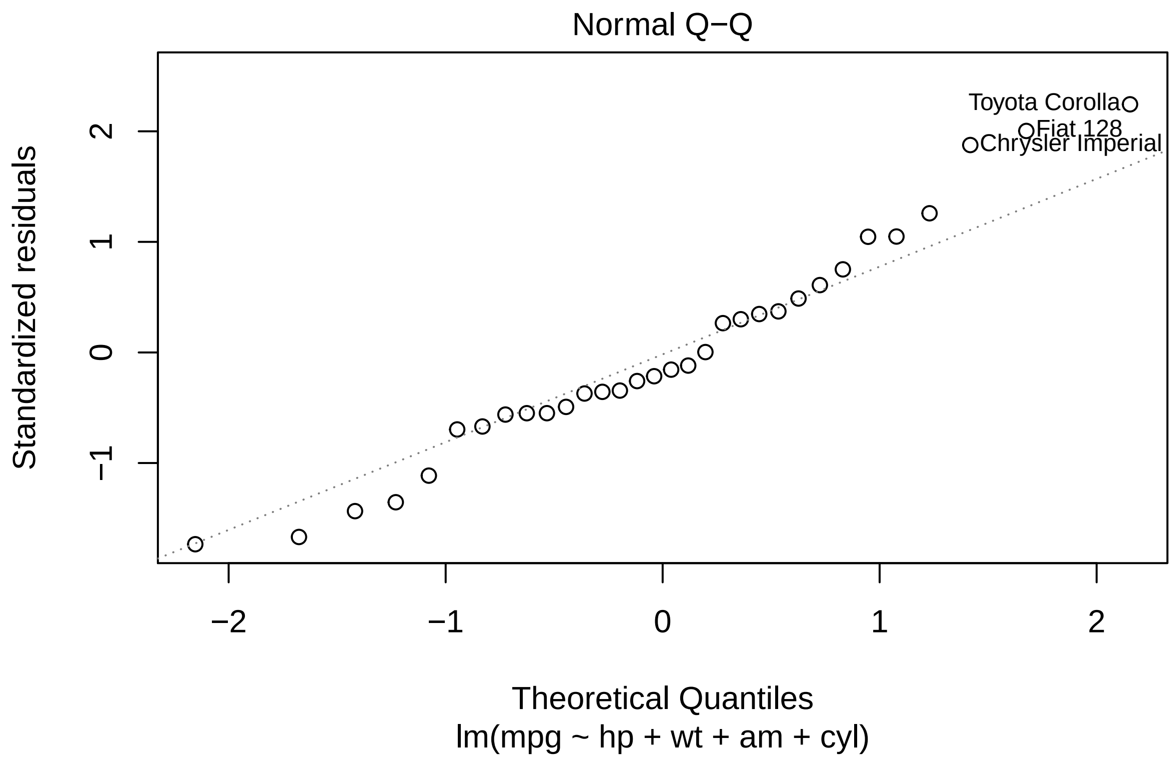
We see that residuals versus that fitted values are in line with no notable patterns. The plot highlights the relatively strong performance of the Toyota Corolla and Fiat 128 (+10%) and the weak performance (-10%) by the Datsun 710 versus their predicted MPG.

```
plot(Model.BestFit, which = 1)
```



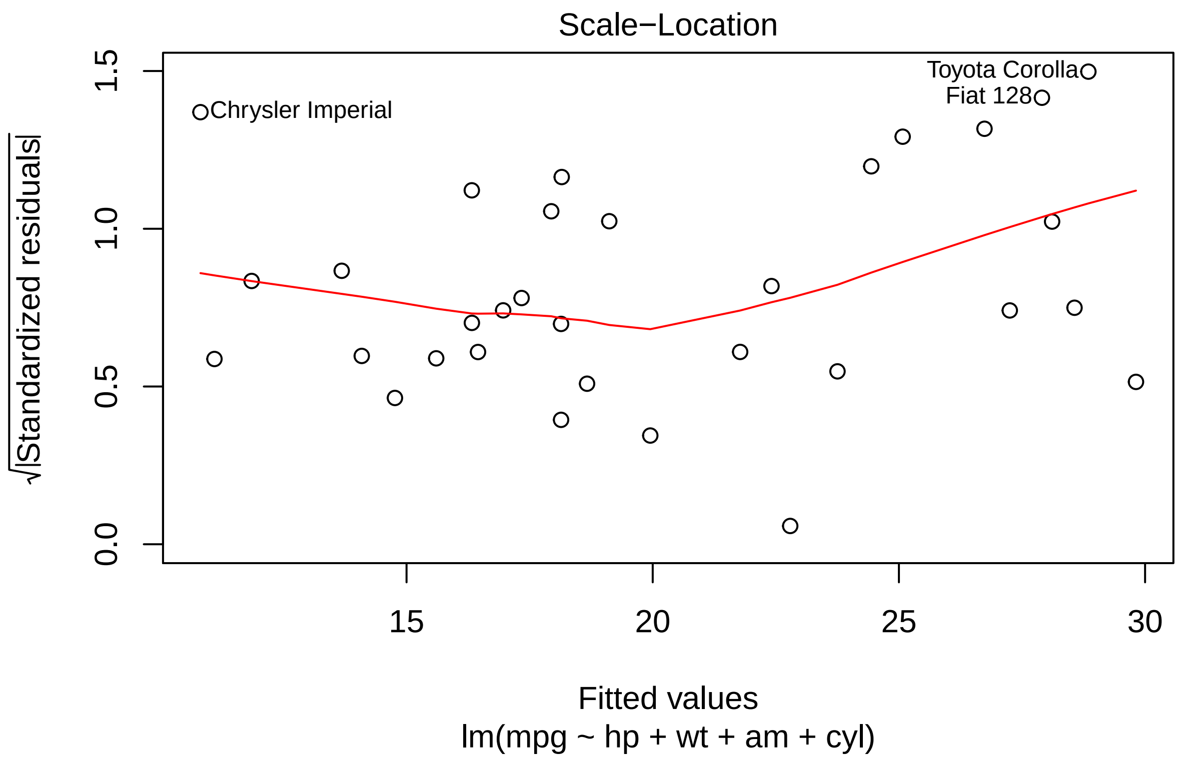
In the QQ plot below, we see that the standardized residuals are for the most part normally distributed with the exception of Toyota, Fiat, and Chrysler being string performers relative to their predicted values.

```r
plot(Model.BestFit, which = 2)
```



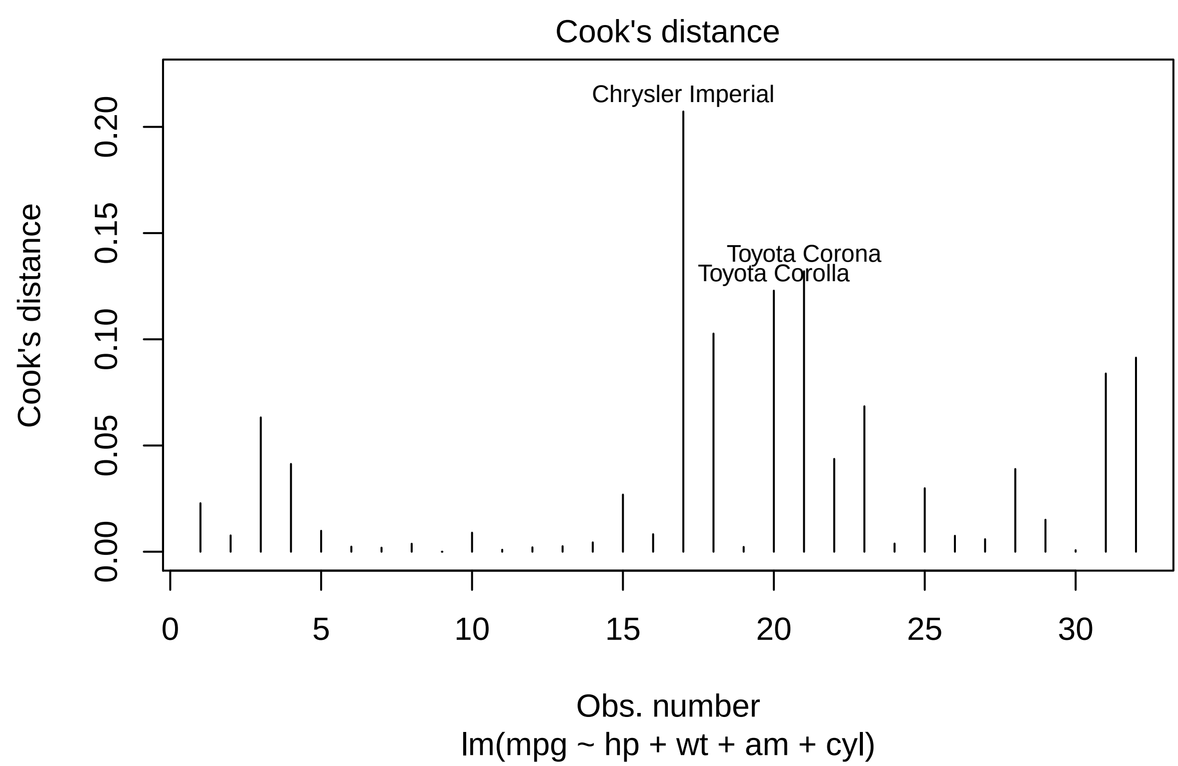
The Scale-Location Plot is indicating that as residuals spread wider from each other the red spread line goes up. In our case the data is homoscedastic and has uniform variance but later it becomes heteroscedastic. This is highlighting a bit of a problem in our linear model that the variances are not as homoscedastic as we would like. This is adding some uncertainty into the strength of the model.

```r
plot(Model.BestFit, which = 3)
```



The measurement is a combination of each observation's leverage and residual values; the higher the leverage and residuals, the higher the Cook's distance.

```
plot(Model.BestFit, which = 4)
```



Conclusion

In our tests, we've found that in some cars a manual transmission may improve gas mileage by a up to 5 mpg, compared with an automatic in the 95% confidence interval. This conclusion is supported by the fact that older automatic cars are fuel inefficient because the torque converter generally slips, wasting fuel as the car shifts. The uncertaintity in our conclusions stem from an unequal distribution of light, low displacement cars in manual transmissions. This is true most notably for the Toyota Corolla and the Fiat 128 in the manual category which have much higher MPG for the displacement due to their lower weight. On the Low MPG side of the automatic transmission, we see that the Lincoln and Maserati are exerting their influence as well. Additionally, because of the heteroscedaticity of the model their does appear to me a quadrative term to the model which we have not discovered that could lead to a better fit when known and included.

```
Lastly, we would like to make the following assumptions:

  1) The test were accurate and that all cars experienced the same road conditions, mix of highway and

  2) The capabilities of the drivers of the manual transmisions are consistent.
```

```r
confint(Model.BestFit)[4,]
```

```
##      2.5 %    97.5 %
## -1.060934  4.679356
```

Appendix

```r
summary(mtcars)
```

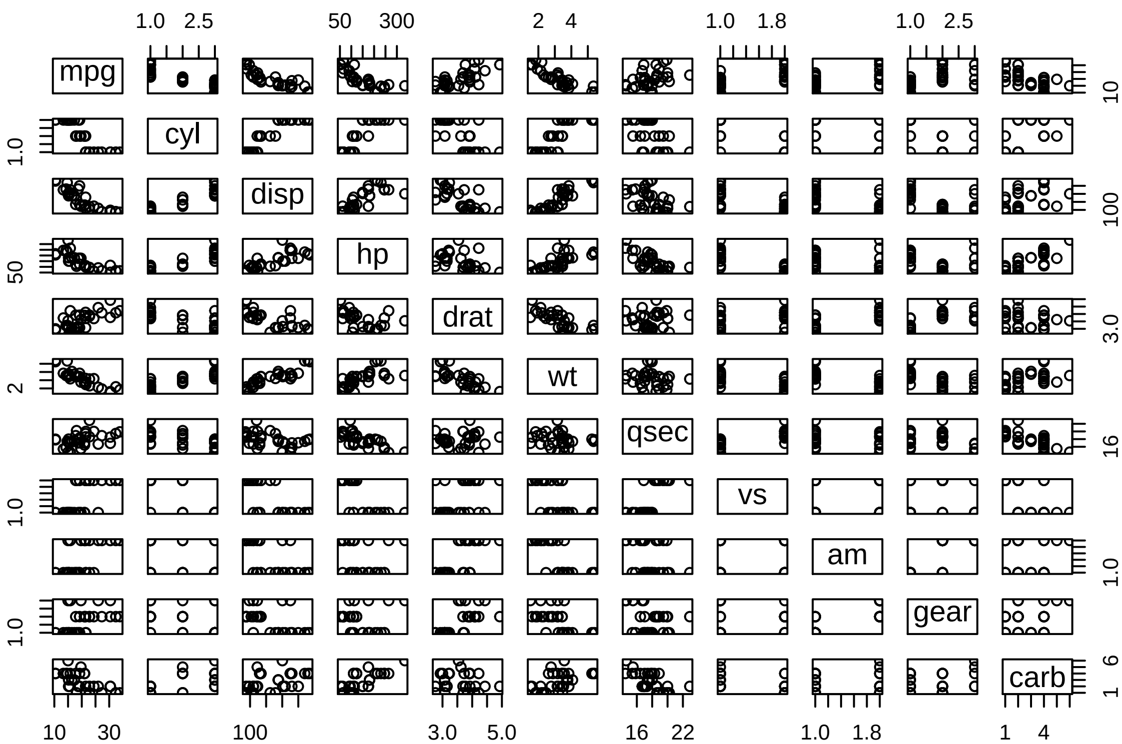
```
##       mpg             cyl             disp             hp       
##  Min.   :10.40   Min.   :4.000   Min.   : 71.1   Min.   : 52.0  
##  1st Qu.:15.43   1st Qu.:4.000   1st Qu.:120.8   1st Qu.: 96.5  
##  Median :19.20   Median :6.000   Median :196.3   Median :123.0  
##  Mean   :20.09   Mean   :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  
```

Apendix I - Variables in MTCARS

[, 1] mpg Miles/(US) gallon [, 2] cyl Number of cylinders [, 3] disp Displacement (cu.in.) [, 4] hp Gross horsepower [, 5] drat Rear axle ratio [, 6] wt Weight (1000 lbs) [, 7] qsec 1/4 mile time [, 8] vs Engine (0 = V-shaped, 1 = straight) [, 9] am Transmission (0 = automatic, 1 = manual) [,10] gear Number of forward gears [,11] carb Number of carburetors

Apendix II - Plot of Potential Regressors in MTCARS

```
plot(mtcars2)
```



Appendix III - Stepping all to Best Fit

```
Best.Fit <- step(Model.All)
```

```
## Start:  AIC=76.4
## mpg ~ cyl + disp + hp + drat + wt + qsec + vs + am + gear + carb
##
##        Df Sum of Sq    RSS    AIC
## - carb  5   13.5989 134.00 69.828
## - gear  2    3.9729 124.38 73.442
## - am    1    1.1420 121.55 74.705
## - qsec  1    1.2413 121.64 74.732
## - drat  1    1.8208 122.22 74.884
## - cyl   2   10.9314 131.33 75.184
```

```
## - vs    1    3.6299 124.03 75.354
## <none>              120.40 76.403
## - disp  1    9.9672 130.37 76.948
## - wt    1   25.5541 145.96 80.562
## - hp    1   25.6715 146.07 80.588
##
## Step:  AIC=69.83
## mpg ~ cyl + disp + hp + drat + wt + qsec + vs + am + gear
##
##        Df Sum of Sq    RSS    AIC
## - gear  2    5.0215 139.02 67.005
## - disp  1    0.9934 135.00 68.064
## - drat  1    1.1854 135.19 68.110
## - vs    1    3.6763 137.68 68.694
## - cyl   2   12.5642 146.57 68.696
## - qsec  1    5.2634 139.26 69.061
## <none>              134.00 69.828
## - am    1   11.9255 145.93 70.556
## - wt    1   19.7963 153.80 72.237
## - hp    1   22.7935 156.79 72.855
##
## Step:  AIC=67
## mpg ~ cyl + disp + hp + drat + wt + qsec + vs + am
##
##        Df Sum of Sq    RSS    AIC
## - drat  1    0.9672 139.99 65.227
## - cyl   2   10.4247 149.45 65.319
## - disp  1    1.5483 140.57 65.359
## - vs    1    2.1829 141.21 65.503
## - qsec  1    3.6324 142.66 65.830
## <none>              139.02 67.005
## - am    1   16.5665 155.59 68.608
## - hp    1   18.1768 157.20 68.937
## - wt    1   31.1896 170.21 71.482
##
## Step:  AIC=65.23
## mpg ~ cyl + disp + hp + wt + qsec + vs + am
##
##        Df Sum of Sq    RSS    AIC
## - disp  1    1.2474 141.24 63.511
## - vs    1    2.3403 142.33 63.757
## - cyl   2   12.3267 152.32 63.927
## - qsec  1    3.1000 143.09 63.928
## <none>              139.99 65.227
## - hp    1   17.7382 157.73 67.044
## - am    1   19.4660 159.46 67.393
## - wt    1   30.7151 170.71 69.574
##
## Step:  AIC=63.51
## mpg ~ cyl + hp + wt + qsec + vs + am
##
##        Df Sum of Sq    RSS    AIC
## - qsec  1     2.442 143.68 62.059
## - vs    1     2.744 143.98 62.126
```

```
## - cyl   2    18.580 159.82 63.466
## <none>              141.24 63.511
## - hp    1    18.184 159.42 65.386
## - am    1    18.885 160.12 65.527
## - wt    1    39.645 180.88 69.428
##
## Step:  AIC=62.06
## mpg ~ cyl + hp + wt + vs + am
##
##        Df Sum of Sq    RSS    AIC
## - vs    1     7.346 151.03 61.655
## <none>              143.68 62.059
## - cyl   2    25.284 168.96 63.246
## - am    1    16.443 160.12 63.527
## - hp    1    36.344 180.02 67.275
## - wt    1    41.088 184.77 68.108
##
## Step:  AIC=61.65
## mpg ~ cyl + hp + wt + am
##
##        Df Sum of Sq    RSS    AIC
## <none>              151.03 61.655
## - am    1     9.752 160.78 61.657
## - cyl   2    29.265 180.29 63.323
## - hp    1    31.943 182.97 65.794
## - wt    1    46.173 197.20 68.191
```

Appendix IV - Influence Measures

```
influence.measures(Model.BestFit)
```

```
## Influence measures of
##   lm(formula = mpg ~ hp + wt + am + cyl, data = mtcars2) :
##
##                        dfb.1_    dfb.hp    dfb.wt dfb.ammn dfb.cy.L  dfb.cy.Q
## Mazda RX4           -0.048802  0.136304 -0.027828 -0.19481 -0.15453  0.202846
## Mazda RX4 Wag        0.005592  0.088387 -0.055026 -0.13215 -0.07772  0.110344
## Datsun 710           0.129468  0.048243 -0.169544 -0.26038  0.16707 -0.203649
## Hornet 4 Drive       0.150492  0.051990 -0.104932 -0.22266 -0.04928 -0.382717
## Hornet Sportabout    0.174584 -0.056898 -0.125483 -0.06921  0.15062  0.071941
## Valiant             -0.017165  0.003358 -0.002441  0.03454  0.01141  0.088116
## Duster 360          -0.032393 -0.069419  0.063810  0.06590  0.01608  0.000161
## Merc 240D           -0.015001 -0.002700  0.038832 -0.05399 -0.07467  0.033949
## Merc 230            -0.000288  0.000689  0.000165 -0.00100 -0.00135  0.000179
## Merc 280             0.023218  0.049238 -0.016367 -0.09551 -0.06060 -0.183388
## Merc 280C           -0.007371 -0.015632  0.005196  0.03032  0.01924  0.058223
## Merc 450SE           0.025401 -0.047373  0.010320  0.00604  0.05825  0.042541
## Merc 450SL           0.067198 -0.035780 -0.037006 -0.02221  0.07281  0.042410
## Merc 450SLC         -0.081357  0.050414  0.039708  0.02386 -0.09459 -0.056903
## Cadillac Fleetwood   0.256129  0.100635 -0.332411 -0.16198  0.02663 -0.085419
## Lincoln Continental  0.156484  0.036186 -0.187001 -0.08584  0.03572 -0.038420
## Chrysler Imperial   -0.883284  0.010564  0.938908  0.35075 -0.33301  0.157983
```

```
## Fiat 128             0.010895 -0.377756  0.212130  0.42920  0.07900  0.338302
## Honda Civic          0.057244 -0.050161 -0.027681  0.02816  0.03927  0.044229
## Toyota Corolla       0.307931 -0.336379 -0.104583  0.28854  0.17610  0.367159
## Toyota Corona       -0.223238 -0.473446  0.364326  0.73054  0.54368 -0.046880
## Dodge Challenger    -0.342679  0.294580  0.147480  0.02138 -0.40669 -0.189939
## AMC Javelin         -0.463555  0.349562  0.230111  0.05486 -0.51338 -0.231778
## Camaro Z28          -0.021025 -0.100331  0.065376  0.08398  0.03402 -0.002327
## Pontiac Firebird     0.196280 -0.180886 -0.060383 -0.02429  0.26994  0.162946
## Fiat X1-9           -0.050682  0.079428  0.002903 -0.07643 -0.03267 -0.083419
## Porsche 914-2        0.005695  0.008997 -0.014763 -0.05839  0.03704 -0.060051
## Lotus Europa         0.175693  0.209744 -0.281548 -0.11405 -0.12358  0.061112
## Ford Pantera L       0.007339 -0.096003  0.065645 -0.09705 -0.05199 -0.021035
## Ferrari Dino         0.008902 -0.022826  0.003213 -0.01371  0.01029  0.049298
## Maserati Bora       -0.233083  0.525088 -0.088177  0.04138 -0.24785 -0.074335
## Volvo 142E           0.433531 -0.023931 -0.433473 -0.36084  0.35893 -0.173175
##                       dffit cov.r   cook.d    hat inf
## Mazda RX4           -0.3660 1.488 2.28e-02 0.2339
## Mazda RX4 Wag       -0.2107 1.633 7.65e-03 0.2496
## Datsun 710          -0.6422 0.681 6.32e-02 0.1120
## Hornet 4 Drive       0.4989 1.197 4.13e-02 0.1840
## Hornet Sportabout    0.2399 1.345 9.83e-03 0.1372
## Valiant             -0.1174 1.512 2.38e-03 0.1758
## Duster 360          -0.1046 1.560 1.89e-03 0.1973
## Merc 240D            0.1475 1.549 3.76e-03 0.2000
## Merc 230             0.0018 1.644 5.61e-07 0.2305
## Merc 280             0.2282 1.468 8.95e-03 0.1842
## Merc 280C           -0.0725 1.542 9.09e-04 0.1842
## Merc 450SE           0.1091 1.356 2.05e-03 0.0927
## Merc 450SL           0.1224 1.364 2.58e-03 0.1011
## Merc 450SLC         -0.1598 1.326 4.39e-03 0.0979
## Cadillac Fleetwood  -0.3975 1.506 2.69e-02 0.2496
## Lincoln Continental -0.2185 1.743 8.24e-03 0.2937   *
## Chrysler Imperial    1.1759 0.716 2.07e-01 0.2611
## Fiat 128             0.8370 0.534 1.03e-01 0.1331
## Honda Civic          0.1141 1.484 2.25e-03 0.1613
## Toyota Corolla       0.9378 0.399 1.23e-01 0.1277
## Toyota Corona       -0.9094 1.068 1.32e-01 0.2778
## Dodge Challenger    -0.5144 1.143 4.37e-02 0.1744
## AMC Javelin         -0.6518 0.998 6.84e-02 0.1828
## Camaro Z28          -0.1490 1.451 3.83e-03 0.1533
## Pontiac Firebird     0.4283 0.966 2.99e-02 0.1016
## Fiat X1-9           -0.2091 1.343 7.49e-03 0.1246
## Porsche 914-2       -0.1854 1.318 5.89e-03 0.1049
## Lotus Europa         0.4841 1.186 3.89e-02 0.1759
## Ford Pantera L      -0.2964 1.532 1.50e-02 0.2302
## Ferrari Dino        -0.0639 1.641 7.08e-04 0.2316
## Maserati Bora        0.7033 2.098 8.39e-02 0.4714   *
## Volvo 142E          -0.7683 0.767 9.14e-02 0.1645
```

Apendix V - Data Summary

```r
summary(mtcars2)
```

```
##       mpg        cyl        disp             hp             drat      
##  Min.   :10.40   4:11   Min.   : 71.1   Min.   : 52.0   Min.   :2.760  
##  1st Qu.:15.43   6: 7   1st Qu.:120.8   1st Qu.: 96.5   1st Qu.:3.080  
##  Median :19.20   8:14   Median :196.3   Median :123.0   Median :3.695  
##  Mean   :20.09          Mean   :230.7   Mean   :146.7   Mean   :3.597  
##  3rd Qu.:22.80          3rd Qu.:326.0   3rd Qu.:180.0   3rd Qu.:3.920  
##  Max.   :33.90          Max.   :472.0   Max.   :335.0   Max.   :4.930  
##        wt             qsec       vs             am     gear   carb  
##  Min.   :1.513   Min.   :14.50   V:18   automatic:19   3:15   1: 7  
##  1st Qu.:2.581   1st Qu.:16.89   S:14   manual   :13   4:12   2:10  
##  Median :3.325   Median :17.71                         5: 5   3: 3  
##  Mean   :3.217   Mean   :17.85                                4:10  
##  3rd Qu.:3.610   3rd Qu.:18.90                                6: 1  
##  Max.   :5.424   Max.   :22.90                                8: 1  
```