Regression Analysis Course Project

G.Stephens

June 21, 2015

Executive Summary

An analysis of the "Motor Trends Car Road Tests" data set was performed to better understand the relationship between the gas mileage (mpg) and the other ten variables measured. The question of whether there was a significant relationship between transmission type and mpg. My preliminary analysis suggested that cars with a manual transmission typically had a better mpg ratio; however, further analysis suggests that this interpretaion is an artifact. This artifact was due to the fact that manual transmission cars typically contained a higher number of forward gears, which provided a proved to underlye the differences between manual and automatic transmission cars in terms of mpg.

Given, the number of dependent variables, it is perhaps not surprising that no statistically significant differences were apparent when regressing mpg versus all the other data points other variables.

```
##                Estimate  Std. Error    t value   Pr(>|t|)
## (Intercept) 12.30337416 18.71788443  0.6573058 0.51812440
## cyl         -0.11144048  1.04502336 -0.1066392 0.91608738
## disp         0.01333524  0.01785750  0.7467585 0.46348865
## hp          -0.02148212  0.02176858 -0.9868407 0.33495531
## drat         0.78711097  1.63537307  0.4813036 0.63527790
## wt          -3.71530393  1.89441430 -1.9611887 0.06325215
## qsec         0.82104075  0.73084480  1.1234133 0.27394127
## vs           0.31776281  2.10450861  0.1509915 0.88142347
## am           2.52022689  2.05665055  1.2254035 0.23398971
## gear         0.65541302  1.49325996  0.4389142 0.66520643
## carb        -0.19941925  0.82875250 -0.2406258 0.81217871
```

However, a simple correlation analysis of the does suggest a positive correlation between mpg and several of the variables, including tranmission (am). This is further illustrated in figure 1 (Appendix). This is further illustrated in the table below.

```
##            [,1]
## mpg   1.0000000
## cyl  -0.8521620
## disp -0.8475514
## hp   -0.7761684
## drat  0.6811719
## wt   -0.8676594
## qsec  0.4186840
## vs    0.6640389
## am    0.5998324
## gear  0.4802848
## carb -0.5509251
```

This is further shown from the analysis of the single varible linear model coefficients, and the scatter plot shown in figure 2 (Appendix).

```
##              Estimate Std. Error   t value     Pr(>|t|)
## (Intercept) 17.147368   1.124603 15.247492 1.133983e-15
## am           7.244939   1.764422  4.106127 2.850207e-04
```

```
summary(lm(mpg~am+gear-1, mtcars))$coeff
```

However, as shown in figure 3 (Appendix), much of the data for manual transmission cars is confounded by the fact that cars with higher numbers of forward gears are over-represented as compared to automatic transmission cars. By factoring in the impact of the number of forward gears, we in fact find a model that provides a better fit to the data set.

```
##      Estimate Std. Error    t value     Pr(>|t|)
## am   1.684774  2.2562237  0.7467227 4.610434e-01
## gear 5.178911  0.3826349 13.5348630 2.604569e-14
```

Appendix

Figure 1

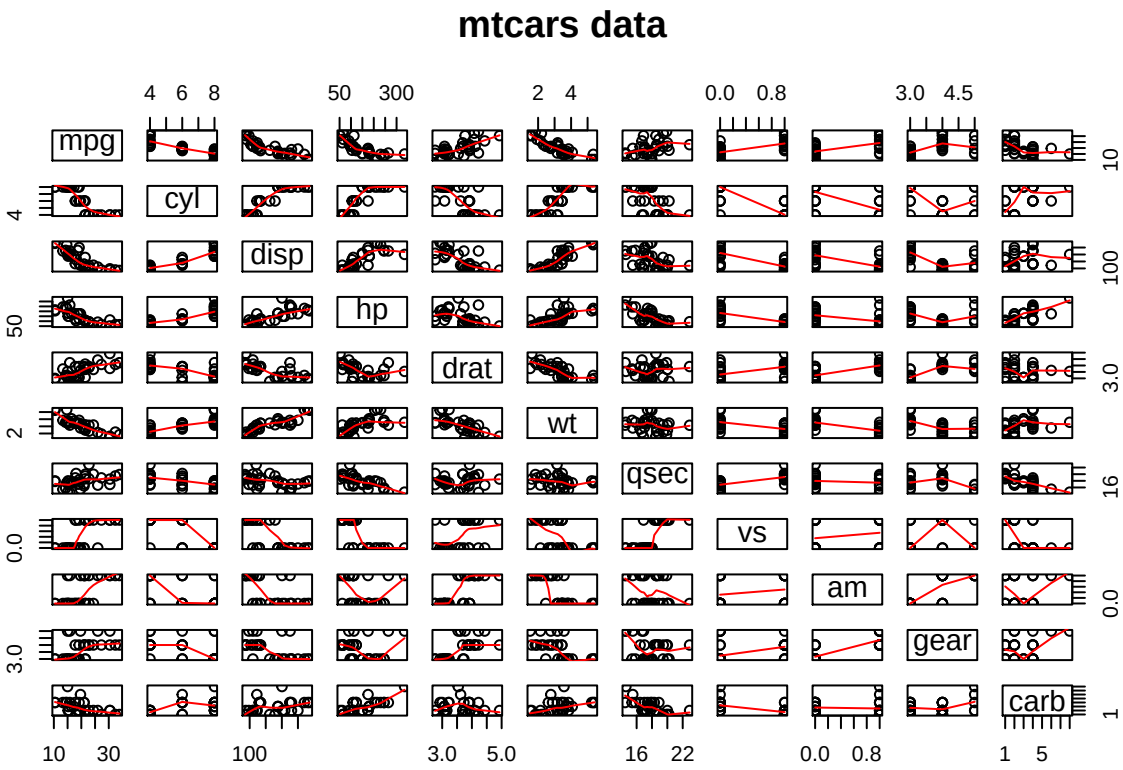

Figure 2

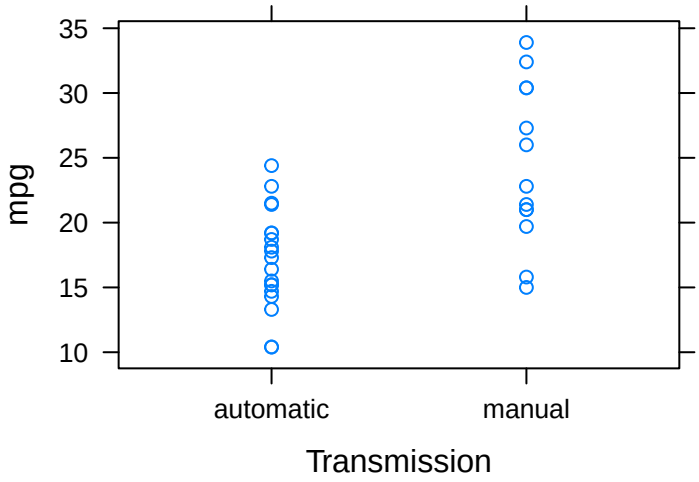

Figure 3

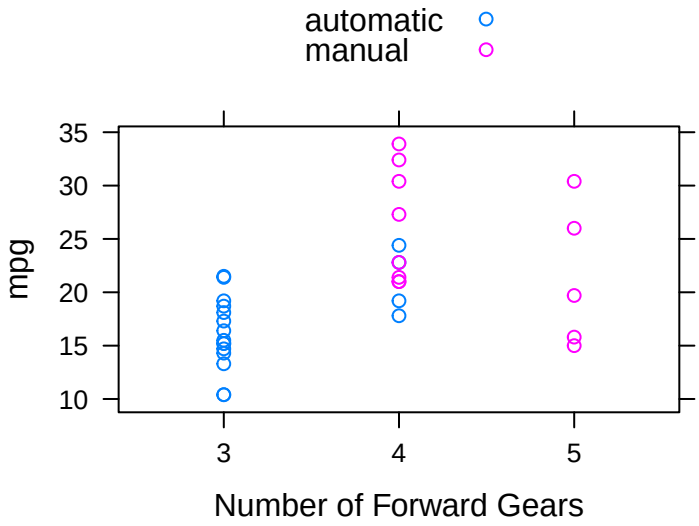

Figure 4

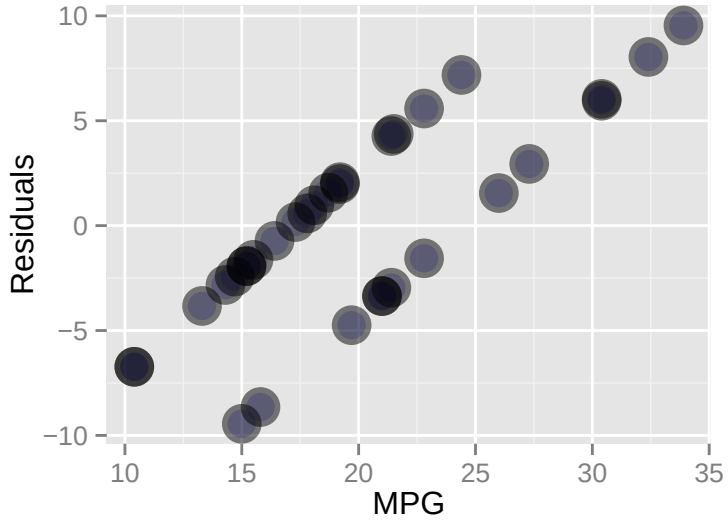

Figure 5