Analysis for Motor Trend

Michaela Schwaiger

16 May 2015

Executive Summary

The aim of this study is to explore the relationship between a set of variables measured for different cars and miles per gallon (MPG). In particular, we adress the following two questions:

- Is an automatic or manual transmission better for MPG?
- What is the MPG difference between automatic and manual transmissions?

Our results suggest that manual transmissions get better milage, with a difference to automatic transmission of about 2.9 mpg. However, we propose to conduct further experiments to fully answer these questions.

Main Analysis

To answer the questions above, we would ideally obtain a dataset where the same car has been tested for miles per gallon fuel consumption with automatic and with manual transmission. We would test different brands and types of cars, changing only the transmission but keeping all other parameters constant. However, the only available dataset is the mtcars dataset. Comparing fuel consumtion for the manual and automatic cars tested in this dataset, we find that manual cars can drive about 24 miles per gallon on average, while automatic cars only make it to 17 miles per gallon (Supplementary Figure 1), suggesting that cars with automatic transmission need more fuel. However, the manual and automatic cars that were analyzed were of different brands and differ in many other ways as well. Some of these differences, for example, the weight of the car, will surely also influence the fuel consumption. Indeed, we can see that heavier cars tend to have lower mpg, and unfortunately, that the cars with automatic and manual transmission are not equally distributed across car weights (Supplementary Figure 2). Therefore we will try to use a linear model to determine the relationship of fuel consumption to transmission, taking all possible confounding variables into account. To this end we first have to determine which other measurments we need to consider, and which ones we can leave out since they are not affecting milage or are dependent on variables we are already taking into account. Supplementary Figure 3 shows an overview of the mtcars dataset and the correlations among its variables. We can see that the biggest (negative) correlation with mpg is the weight of the cars, which is also not evenly distributed among manual and automatic cars. Therefore we definatly have to take weight into account. Sine weight correlates strongly with displacement,and it is not clear how this would affect the mileage of a car, we can omit this variable. The next group of variables contains the number of cylinders,horse power, V/S, carb and qsec. We will take horse power and 1/4 mile time (qsec) into account. Finally, we will take the gearing of the car into account by considering the rear axle ratio.

We first test if the inclusion of all these variables makes sense by fitting the linear model of predicting mpg by transmission type (am) alone and with adding more of the other variables. Then we perform an Anova on these models. The result below shows that we should definatly include weight, and also the horse power and 1/4 mile time (qsec), but the rear axle ratio does not seem to add anything in addition.

```
## Analysis of Variance Table
##
## Model 1: mpg ~ am
## Model 2: mpg ~ am + wt
## Model 3: mpg ~ am + wt + hp + qsec
## Model 4: mpg ~ am + wt + hp + qsec + drat
##   Res.Df RSS Df Sum of Sq     F  Pr(>F)
```

```
## 1     30 721
## 2     29 278  1       443 72.54 5.4e-09 ***
## 3     27 160  2       118  9.69 0.00072 ***
## 4     26 159  1         1  0.23 0.63261
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

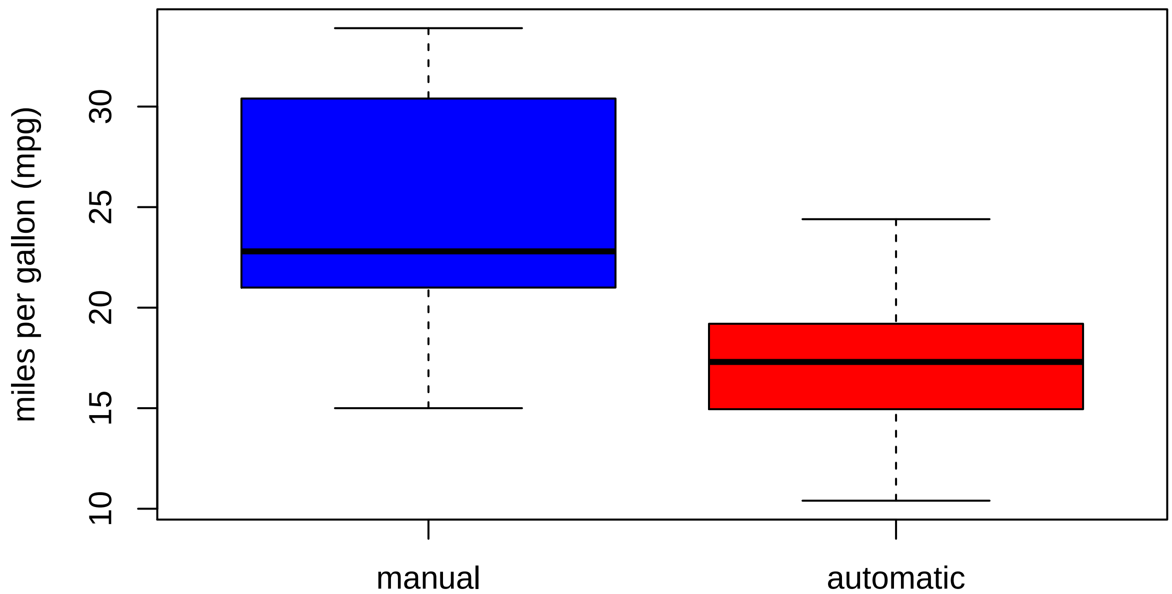
Therefore we consider the 3rd model below.Looking at the summary of this model we estimate a (barely) significant 2.9 mpg increase in fuel efficency for a change from automatic transmission (am=0) to manual transmission (am=1), while holding the other variables constant.

```
##
## Call:
## lm(formula = mpg ~ am + wt + hp + qsec, data = mtcars)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -3.497 -1.590 -0.112  1.180  4.540
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  17.4402     9.3189    1.87   0.0721 .
## am            2.9255     1.3971    2.09   0.0458 *
## wt           -3.2381     0.8899   -3.64   0.0011 **
## hp           -0.0176     0.0142   -1.25   0.2231
## qsec          0.8106     0.4389    1.85   0.0757 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.43 on 27 degrees of freedom
## Multiple R-squared:  0.858,  Adjusted R-squared:  0.837
## F-statistic: 40.7 on 4 and 27 DF,  p-value: 4.59e-11
```

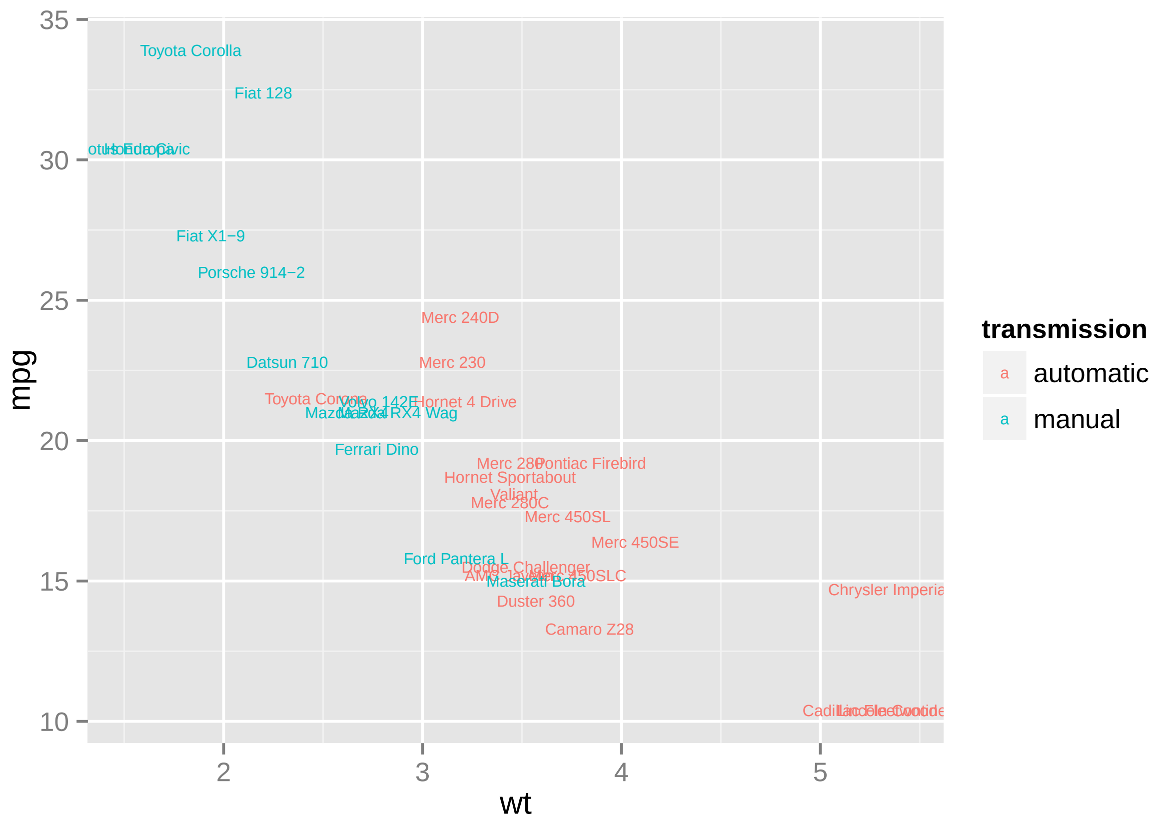
These results suggest that manual transmissions get better milage, with a difference to automatic transmission of about 2.9 mpg. Looking at the diagnostic plots (Supplementary Figure 4) we notice that the residuals are higher for higher predicted values, suggesting that the prediction of mpg by the model is worse in that range. We also see that two cars have higher influence on the model (and low residuals) than the other datapoints (Merc 230 and Maserati Bora).

Due to these caveats, and since this effect of transmission on mpg is not highly significant, and actually not significant at all when considering models 2 or 4, we would highly recommend to generate a new daset as described in the beginning to be able to answer the questions above with certainty.

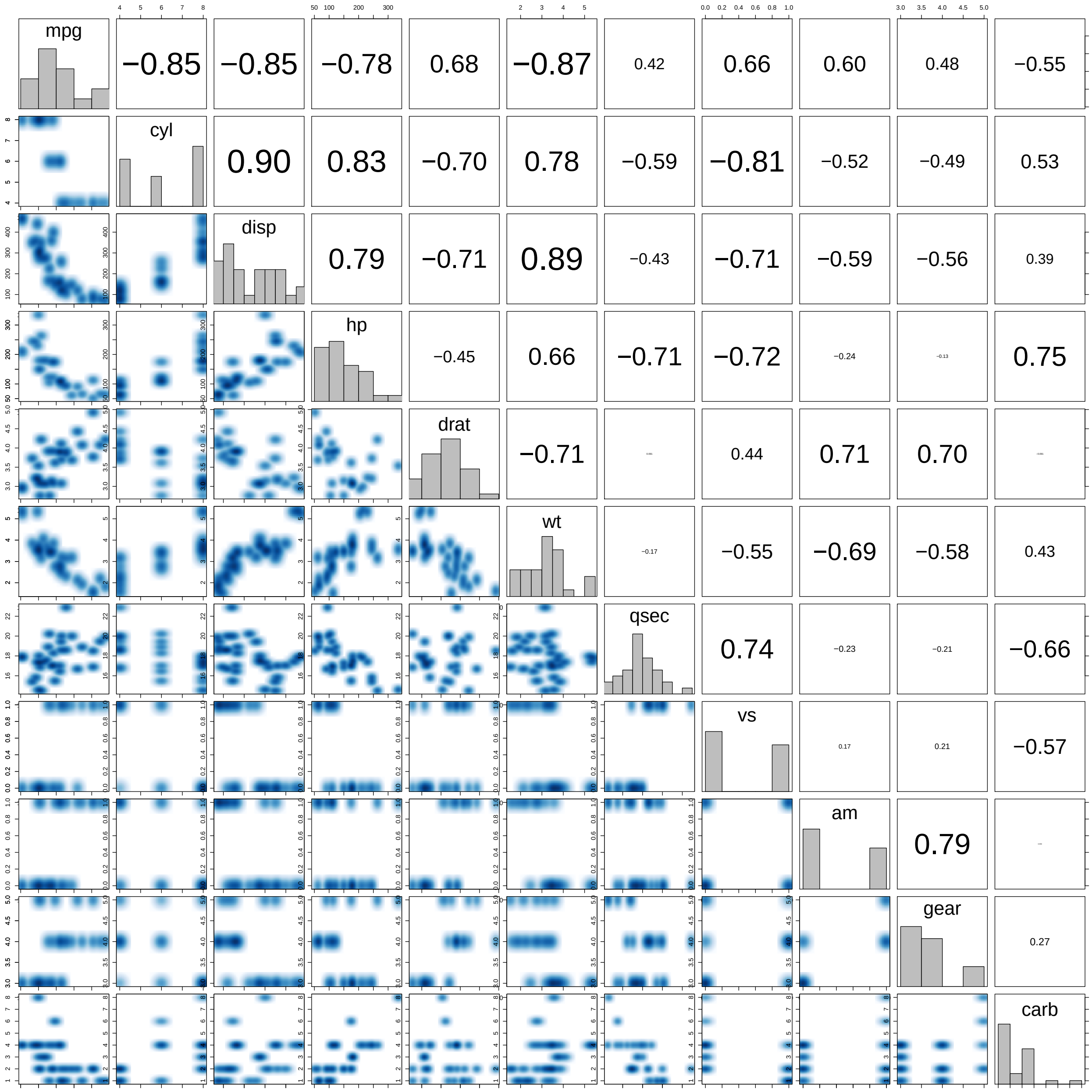
Supplementary Figures:



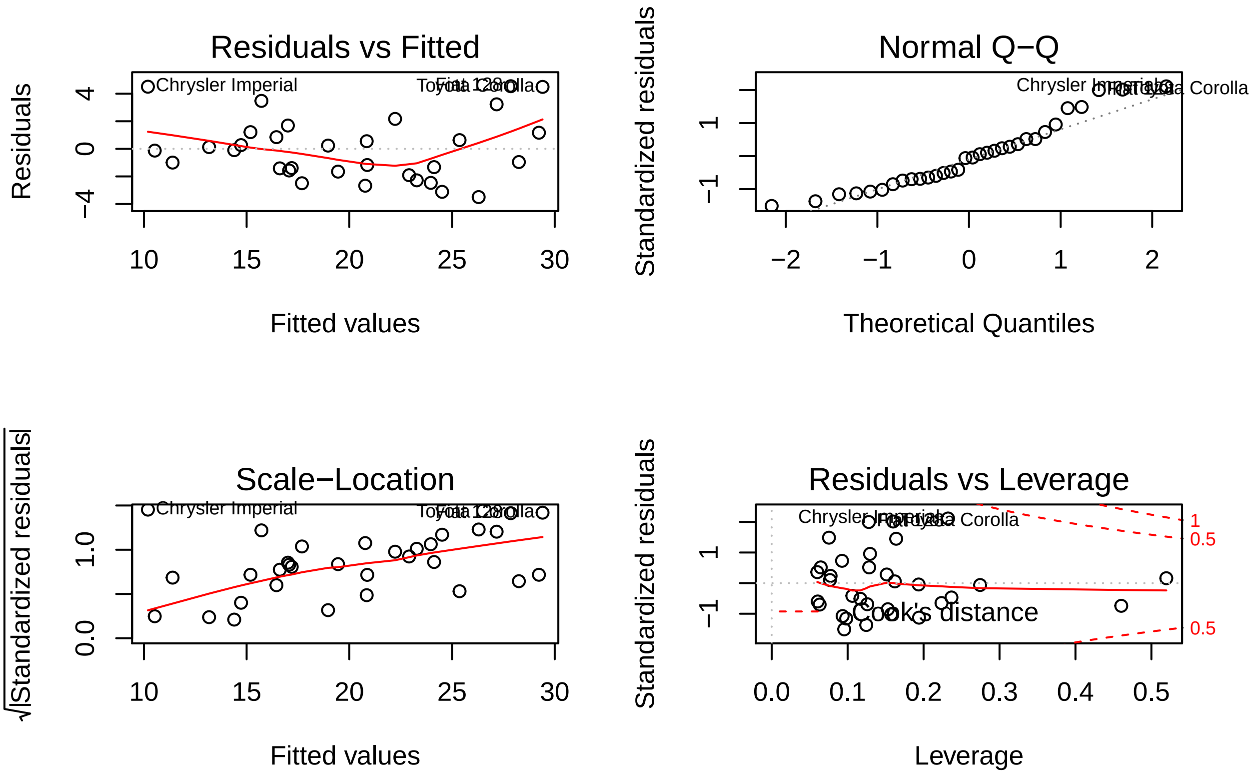
Supplementary Figure 1: Boxplot comparing fuel consumption (measured in mpg) in manual and automatic transmission cars.



Supplementary Figure 2: Comparison of fuel efficiency in mpg (y-axis) to car weight (x-axis). The datapoints are colored by transmission type.



Supplementary Figure 3: Pairwise correlations of all measurements in the mtcars dataset.



Supplementary Figure 4: Diagnostic plots of the linear model comparing mpg to transmission, taking weight, horse power, and 1/4 mile time (qsec) into account.