Motor Trend - Transmission Efficiency

Brett Taylor

August 3, 2015

Executive Summary

Our company Motor Trend Magazine did a study in 1974 that collected data on 32 cars that included information about fuel consumption, and also the design factors that may have an influence in this area. The industry has gone through a transformation where the majority of automobiles now use automatic transmissions instead of the traditional manual transmission. We have identified that there appears to be a significant correlation between the transmission type, and the miles per gallon. While this appears to be true, when performing a deeper analysis, we see that other automobile design factors do impact the miles per gallon. You can drive automatic or manual transmission based cars and recognize one of the most impactful aspect of your car is the mass of the vehicle. When evaluating the weight vs. the transmission type, we then see that the manual transmission is more efficient for light cars, and automatic is better for heavy cars. If you want the most efficient, then go with a car under 2,800 lbs. and with a standard transmission.

The Data

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). A data frame with 32 observations on 11 variables is used for this analysis.

Table 1: Description of data captured in the mtcars data set

fields	descriptions
mpg	Miles/(US) gallon
cyl	Number of cylinders
disp	Displacement (cu.in.)
hp	Gross horsepower
drat	Rear axle ratio
wt	Weight (lb/1000)
qsec	1/4 mile time
vs	V/S - Engine type - V Shape or Straight
am	Transmission (0 = automatic, 1 = manual)
gear	Number of forward gears
carb	Number of carburetors

Analysis

Which variables should we include in the analysis?

A good prediction of whether the trend towards the automatic transmission has had a positive impact on the gas mileage of cars requires that first we look to ensure that the data we utilize includes the variables that we know could cause gas mileage to go up or down. Next we need to evaulate if the variables that do look like they matter, have a relationship between each other. If this is true, we will remove variables that are too

tighlty correlated, so that we do not create an invalid model.

The outcome we are looking at is the Miles per US Gallon (MPG) and is known as the dependent variable. The independent variables may allow us to predict the outcome, start with the remaining variables in the data set.

Our first review of this information is the *Figure 1.* scatter plot found in the appendix. We reviewed each of the variables, and provided a linear regression which is displayed in each graph. We also looked at the correlation between miles per gallon and the independent variable to see which variables more than likely do not have any impact on the mpg. We identified in this review that there were 2 variables that had minimal impact on the miles per gallon of the automobiles. These variables are **qsec** *the time it takes to drive 1/4th mile*, and **gear** *the number of forward acting gears in the transmission.* Since both were under |.5| correlation, and they had too high of p-values, we rejected them from the analysis.

We also identified several variable with high correlation between each other and this made them suspect. The initial review was performed by reviewing *Figure 2.* in the appendix which shows a correlation plot between all variables with MPG as the dependent variable. One of the most important relationships was found between wt, cyl, and disp which made sense. To complete the analysis, we created linear models and used ANOVA to determine which independent variables should remain.

Table 2: ANOVA model evaluation results

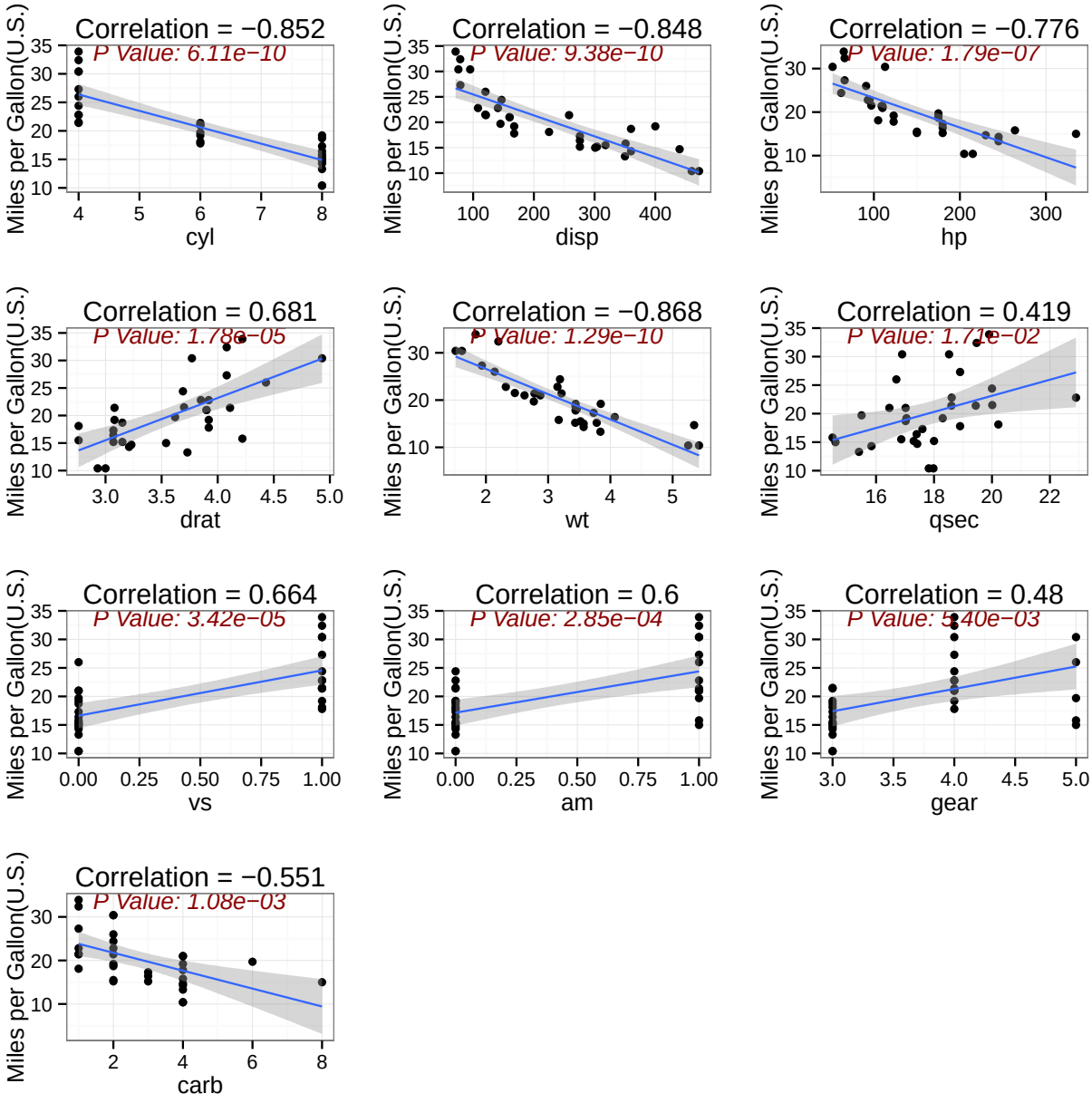
Analysis of Variance Table
Model 1: mpg ~ am
Model 2: mpg ~ am + wt
Model 3: mpg ~ am + wt + disp
Model 4: mpg ~ am + wt + disp + vs
Model 5: mpg ~ am + wt + disp + vs + carb + qsec
Model 6: mpg ~ am + wt + disp + vs + carb + qsec + hp + drat
Model 7: mpg ~ am + wt + disp + vs + carb + qsec + hp + drat + gear

model	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
Model 1	30	720.8966	NA	NA	NA	NA
Model 2	29	278.3197	1	442.5769	57.2876	0.0000
Model 3	28	246.5563	1	31.7634	4.1115	0.0586
Model 4	27	214.6765	1	31.8797	4.1265	0.0582
Model 5	21	156.6331	6	58.0434	1.2522	0.3296
Model 6	19	134.1485	2	22.4846	1.4552	0.2610
Model 7	17	131.3341	2	2.8145	0.1822	0.8351

The ANOVA analysis identified that there is only one coefficient under .5 % which is the *wt* variable which means the mass of the car is the variable worth analyzing. To determine wheter the manual or automatic transmission was more efficient, we first created a regression model that compared both auto to manual, and created a scatter plot (*Figure 3.*). This demonstrated that when a car is under the weight of approximately 2808 pounds, the mpg are more efficient for a manual automobile. Over this mass, then we see that the automatic transmission is more efficient.

Appendix

Independent vs. Dependent



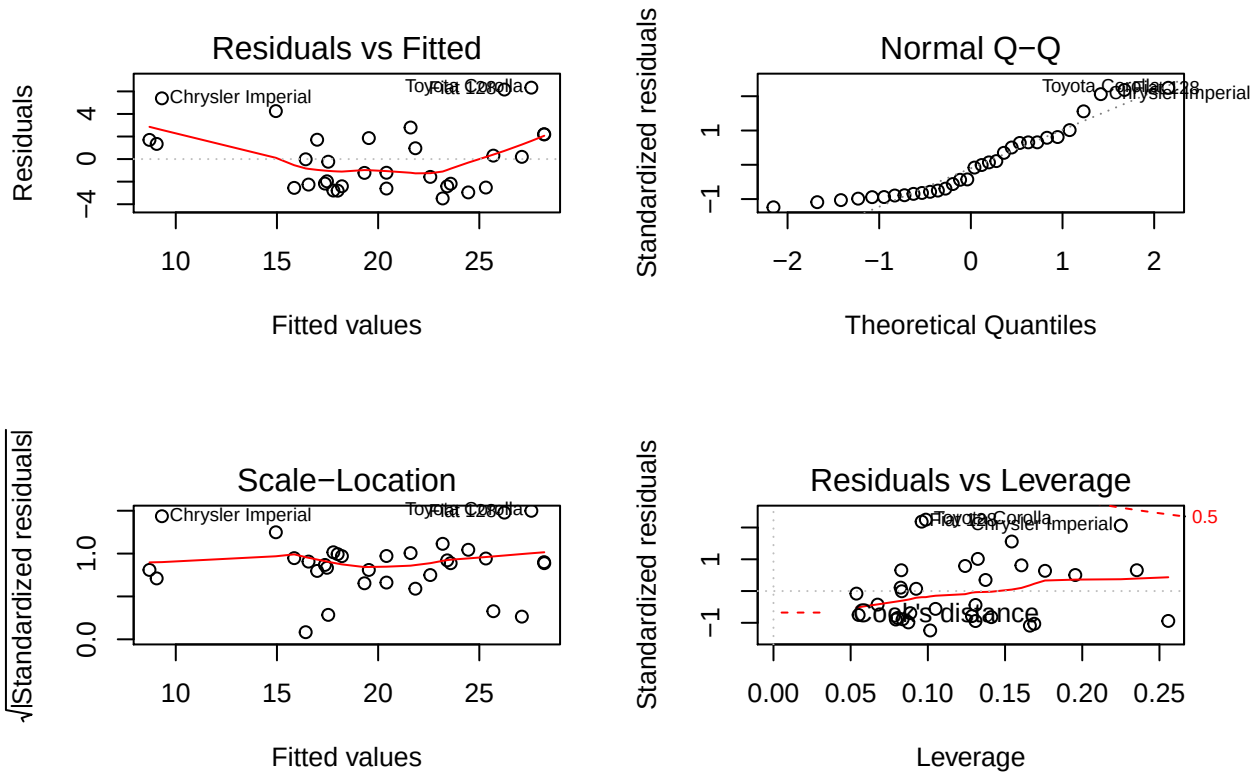


Figure 1: Regression fit plot

Regression Variance

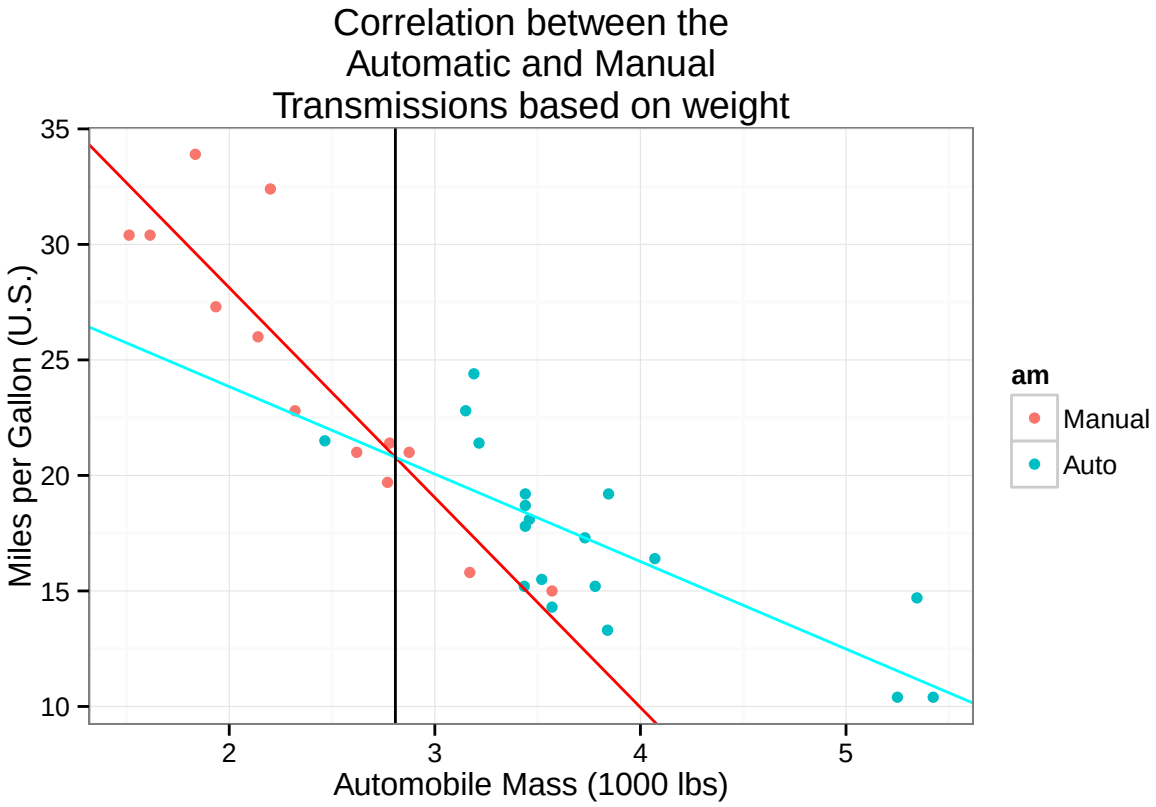


Figure 2:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	46.294478	3.010149	15.379464	0.0000000
wt	-9.084268	1.212400	-7.492798	0.0000000
amAuto	-14.878422	4.264042	-3.489276	0.0016210
wt:amAuto	5.298361	1.444699	3.667449	0.0010171

$$y_a = \beta_{0a} + \beta_{1a} * x$$
$$y_m = \beta_{0m} + \beta_{1m} * x$$

Let

$$y_a = y_m$$
$$\beta_{0a} + \beta_{1a} * x = \beta_{0m} + \beta_{1m} * x$$
$$\beta_{0a} - \beta_{0m} = (\beta_{1m} - \beta_{1a}) * x$$
$$x = \frac{\beta_{0a} - \beta_{0m}}{\beta_{1m} - \beta_{1a}}$$

The x intercept = 2.8 or 2808.12 lbs where an automatic transmission is more efficient.

BONUS - Correlation Analysis

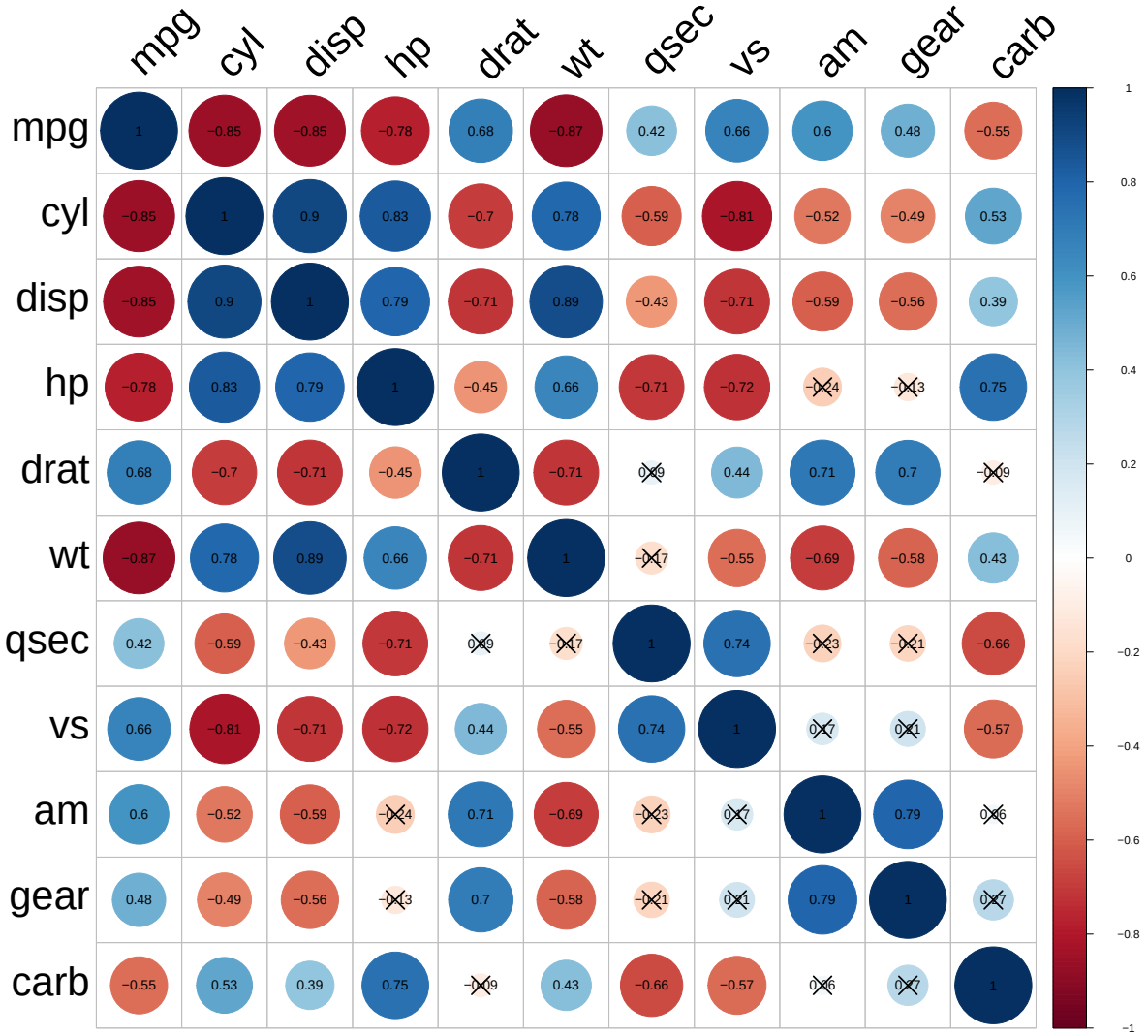


Figure 3: Significance Test with Correlation Analysis