Impact of Transmission Type on Automobile Fuel Consumption (Miles per Gallon)

Scott Robertson

January 17, 2019

Executive summary

The purpose of this paper is to answer two specific questions:

- Is an automatic or manual transmission better for Miles per Gallon(MPG)?; and
- Quantify the MPG difference between automatic and manual transmissions?

Using regression analysis we can show that a manual transmission is better for MPG. This provides an increase of 1.8 MPG compared to automatic transmission.

However, the number of cylinders, horsepower and weight have a more statistically significant impact on MPG.

Data Source

For the purpose of this analysis we will be using the mtcars datatset which is included in the datasets package.

This data set contains fuel consumption, design specifications and performance measures for 32 automobiles from 1974-74.

Exploratory Analysis

As a start point a piece of exploratory analysis was undertaken to determine if there is a relationship between transmission and MPG. The output of this analysis is in Figure 1 in Appendix A. This box plot shows a clear relationship between the two with manual transmission having a higher average MPG.

In order to understand and quantify this relationship we will use regression models. As part of the exploratory analysis a correlation matrix was produced to look at any existing relationships between variables in the data set. This matrix is included in Figure 2 in Appendix A.

Based on this analysis there are a number of highly correlated predictor variables that appear to have a greater impact

Regression Analysis

To begin with we create a liner model with mpg as the dependent variable and use all remaining variables as predictors. As we have already identified correlated variables this will require additional work to improve the model.

Observations	32
Dependent variable	mpg
Type	OLS linear regression

F(16,15)	7.83
R^2	0.89
Adj. R^2	0.78

	Est.	S.E.	t val.	p	
(Intercept)	23.88	20.07	1.19	0.25	
cyl6	-2.65	3.04	-0.87	0.40	
cyl8	-0.34	7.16	-0.05	0.96	
disp	0.04	0.03	1.11	0.28	
hp	-0.07	0.04	-1.79	0.09	.
drat	1.18	2.48	0.48	0.64	
wt	-4.53	2.54	-1.78	0.09	.
qsec	0.37	0.94	0.39	0.70	
vs1	1.93	2.87	0.67	0.51	
amManual	1.21	3.21	0.38	0.71	
gear4	1.11	3.80	0.29	0.77	
gear5	2.53	3.74	0.68	0.51	
carb2	-0.98	2.32	-0.42	0.68	
carb3	3.00	4.29	0.70	0.50	
carb4	1.09	4.45	0.25	0.81	
carb6	4.48	6.38	0.70	0.49	
carb8	7.25	8.36	0.87	0.40	

Standard errors: OLS

Since none of the coefficients have a p-value less than 0.05 we cannot conclude which variables are more statistically significant. In order to identify which variables are most significant in determining MPG we will use stepwise selection to identify the best combination of predictor variables to use in our model.

Observations	32
Dependent variable	mpg
Type	OLS linear regression

F(5,26)	33.57
R^2	0.87
Adj. R^2	0.84

	Est.	S.E.	t val.	p	
(Intercept)	33.71	2.60	12.94	0.00	***
cyl6	-3.03	1.41	-2.15	0.04	*
cyl8	-2.16	2.28	-0.95	0.35	
hp	-0.03	0.01	-2.35	0.03	*
wt	-2.50	0.89	-2.82	0.01	**
amManual	1.81	1.40	1.30	0.21	

Standard errors: OLS

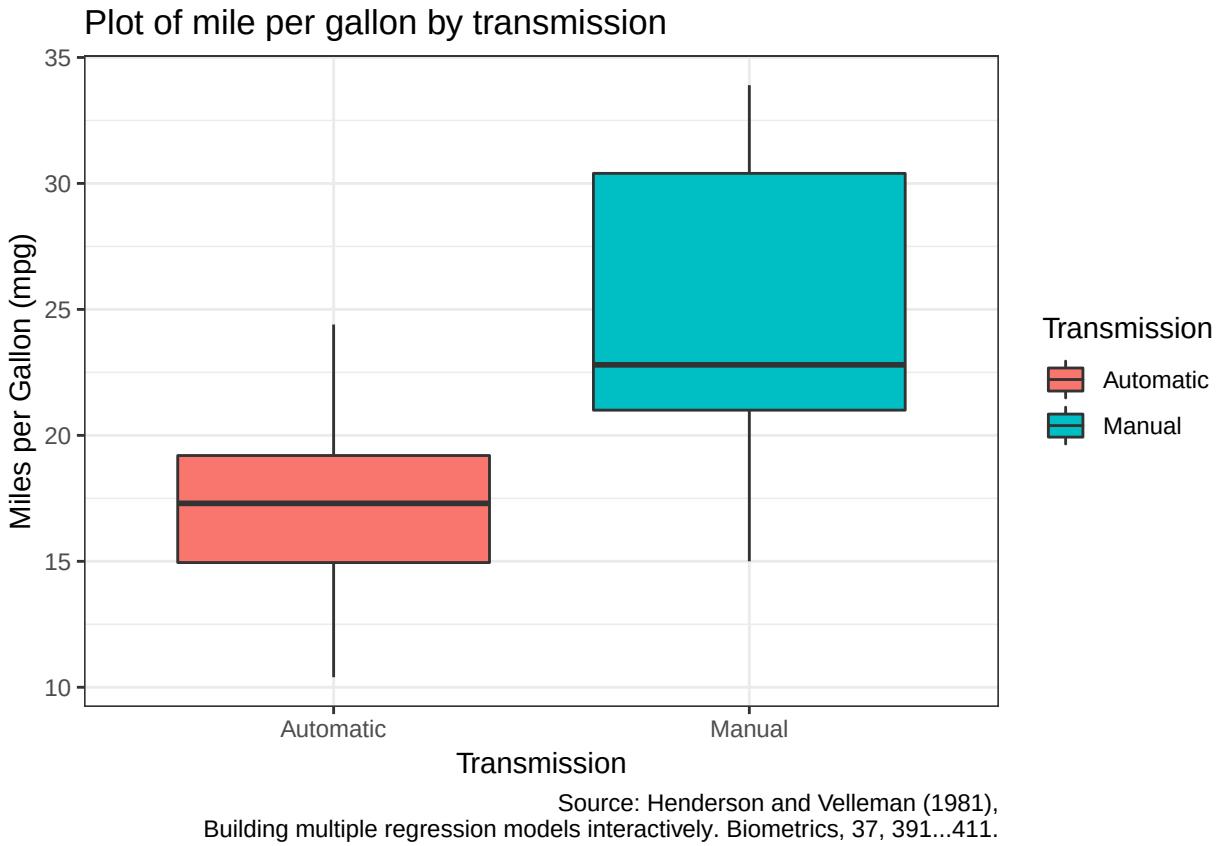
The new model has 4 variables; cylinders, horsepower, weight and transmission. The R-squared value of 0.8659 confirms that this model explains about 87% of the variance in MPG.

The p-values also are statistically significantly because they are less than 0.05. The coefficients conclude that:

- increasing the number of cylinders to 6 decreases MPG by 3.03
- increasing the number of cylinders to 8 decreases MPG by 2.16
- increasing horsepower by 1 decreases MPG by 0.03
- increasing weight by 1kg decreases MPG by 2.5
- manual transmission improves the MPG by 1.81 compared to automatic transmission

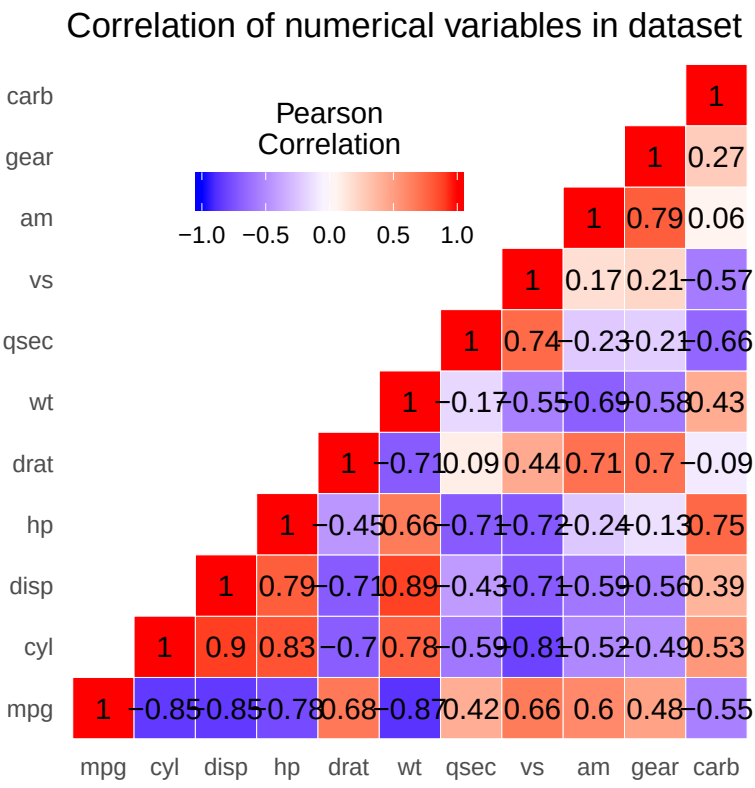
Appendix A - Figures

Figure 1 - MPG/Transmission Boxplot



Source: Henderson and Velleman (1981),
Building multiple regression models interactively. Biometrics, 37, 391...411.

Figure 2 - Variable correlation plot



Source: Henderson and Velleman (1981),
Building multiple regression models interactively. Biometrics, 37, 391...411.

Figure 3 - All Variable Model Diagnostics

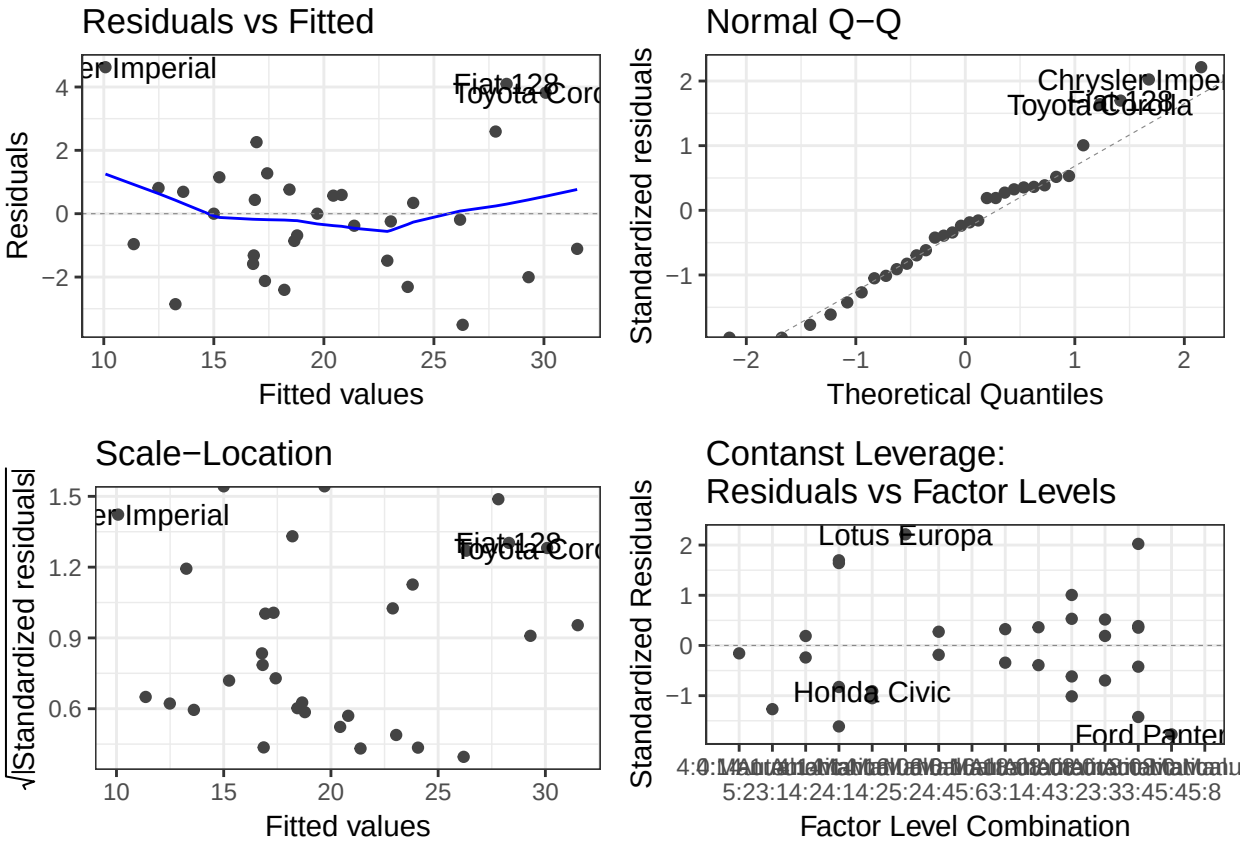


Figure 4 - Stepwise Algorithm Model Diagnostics