

Course Project

Matthew Dunne

September 29, 2017

Executive Summary

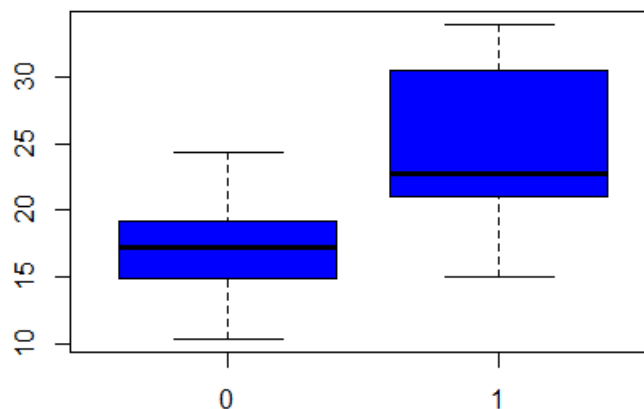
Is an automatic or manual transmission better for MPG? While a quick inspection of the mtcars data would suggest automatic transmissions is significantly better for mpg, we see essentially no difference when we control for weight.

Exploring the Data

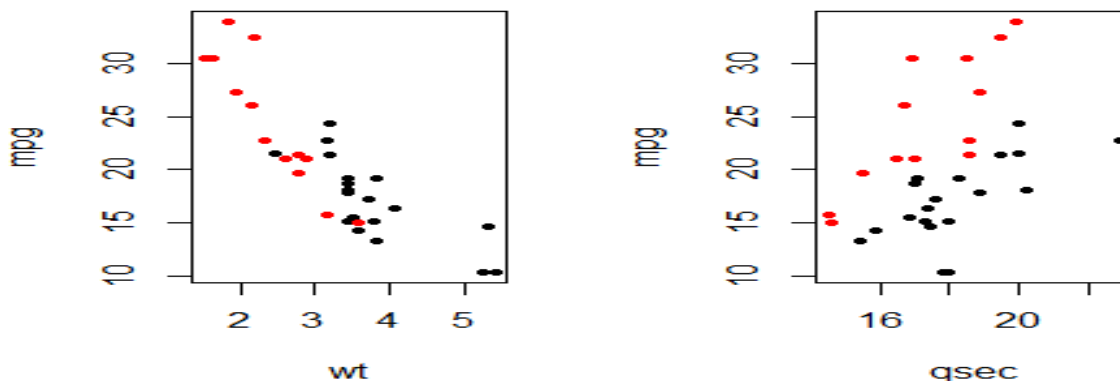
Looking at the data we see several variables that might affect mpg in addition to whether a car has an automatic or manual transmission. These include number of cylinders, weight, and speed (qsec=1/4 mile time).

##		mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
##	Mazda RX4	21.0	6	160	110	3.90	2.620	16.46	0	1	4	4
##	Mazda RX4 Wag	21.0	6	160	110	3.90	2.875	17.02	0	1	4	4
##	Datsun 710	22.8	4	108	93	3.85	2.320	18.61	1	1	4	1

Here we see that the median mpg (y-axis) for automatics (1) is significantly above the median mpg for manuals (0). Also the spread of data points suggest that cars with automatic transmission tend to have higher mpg than manuals.



Now let's see what happens when we separately account for weight and speed. In the plot below on the left we see that, even when distinguishing by transmission, there seems to be a linear relationship between mpg and weight. In the plot on the right, there is no apparent linear relationship between mpg and speed (qsec). Automatics get better mpg than manuals as previously seen.



Choice of Model, Interpretation, and Uncertainty

Because we see possible linear relationship between mpg and weight we will incorporate that into our model. Specifically, we will look whether automatic or manual transmissions have better mpg when accounting for weight. We will mean center the weight variable because no car has a weight of 0.

	Estimate	Std. Error	t value	Pr(> t)
## (Intercept)	20.10021868	0.8331837	24.12459551	9.639957e-21
## I(wt - mean(wt))	-5.35281145	0.7882438	-6.79080719	1.867415e-07
## am	-0.02361522	1.5456453	-0.01527855	9.879146e-01

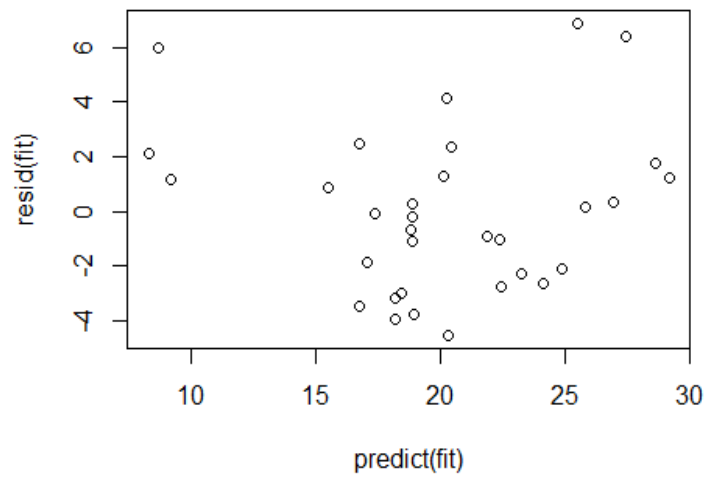
Under this model, automatics have a slightly lower mpg (20.10-0.02 or 20.08). This is somewhat surprising given our boxplot. However this requires some interpretation. First we will look at the confidence interval of the intercept for automatic and second its probability

```
## [1] -3.184815  3.137584
```

As we might have suspected, the confidence interval is almost evenly split between values above the intercept for manuals and values below. Also notice the probability of the t-value is about 0.99 which means we fail to reject the null hypothesis (that there is no difference in mpg for automatics and manuals when we hold weight constant).

Residuals and Diagnostics

To examine how well our model fits the data we will plot the interaction of our predicted values (x-axis) and residuals (y-axis)



There is no discernible pattern to the residuals, suggesting the model is a relatively good fit.