Motor Trend Analysis

Cary Correia

November 7, 2014

Executive Summary: The mtcars database was extracted from the 1974 Motor Trend US magazine and comprises fuel consumption(mpg) and 10 possible regressors. This paper explores the relationship between these regressors and MPG. There was a focused attempt to see if an automatic or manual transmission was better for MPG and to quantify the MPG difference between automatic and manual transmissions.

A Boxplot and t-test showed that manual transmission's gave a median 5.5 mpg boost over automatic with a significant p-value of 0.0013736. However we could see that other factors also influenced MPG via covariance plot/tests so regression models were built to test those relationships.

Three regression models were created to explore mpg efficiency. The final regression equation consisted of the following terms auto/manual transmission, weight, cylinders, horsepower and hp interacting with cylinder and weight interacting with auto/manual transmissions. This analysis reveals the following with respect to fuel consumption:

1. A manual transmission significantly drives performance by 4.5329462 mpg
2. As weight increases by every 1 ton the mpg degrades by -2.3129274 mpg
3. As weight increases with a manual transmission there is a further degradation by -3.046846 mpg
4. There were other factors that influenced mpg (cylinder and horsepower) but their effects were minimal (see detailed analysis)
5. Overall the final model had a Rsq of 0.9036165 and Rsq Adj of 0.8700918 with a significant p-value for the overall model

Final Verdict: Although this analysis proved that manual transmissions can lead to better mileage it should be noted that some biasing could be present in the data. For example the data list and outliers showed that 32 makes were explored with only 1 measure/ vehicle and we don't know if there was 1 or many drivers involved or if other factors might have been present that could have introduced biasing.

Detailed Analysis- Data Exploration:

```
# Load and explore database mtcars
data(mtcars)
cor(mtcars$mpg, mtcars)
```

```
##      mpg       cyl       disp         hp      drat         wt     qsec
## [1,]   1 -0.852162 -0.8475514 -0.7761684 0.6811719 -0.8676594 0.418684
##             vs        am      gear       carb
## [1,] 0.6640389 0.5998324 0.4802848 -0.5509251
```

Data shows we have mpg along with the other 10 factors. Four factors (cyl, hp, wt and carb) showed negative correlations to mpg while the others were +ve.

Boxplot and Regression Results

The Boxplot shows a shift of 5.5 mpg. Manual transmissions have a significant effect on mpg. See boxplot figure in Appendix

Regression Summary The following table shows the output of the 3 regression models deployed (see Appendix for code and sample output: (note: the last two models we adjusted hp and wt to make the intercept easier to interpret)

	Model 1	Model 2	Model 3
(Intercept)	33.708***	27.781***	28.860***
	(2.605)	(5.635)	(2.050)
cyl: 6/4	-3.031*	-2.295	-4.101
	(1.407)	(7.781)	(3.178)
cyl: 8/4	-2.164	-5.035	-6.852*
	(2.284)	(6.068)	(2.781)
hp	-0.032*		
	(0.014)		
wt	-2.497**		
	(0.886)		
am: 1/0	1.809	5.760	4.533*
	(1.396)	(5.404)	(2.068)
wtFix		-1.218	-2.313**
		(3.659)	(0.812)
hpFix		-0.091*	-0.083*
		(0.038)	(0.034)
wtFix x am: 1/0		-5.176	-3.047
		(4.187)	(1.516)
cyl: 6/4 x hpFix		0.053	0.060
		(0.055)	(0.050)
cyl: 8/4 x hpFix		0.073	0.076*
		(0.038)	(0.036)
am: 1/0 x hpFix		0.019	
		(0.022)	
wtFix x cyl: 6/4		-0.926	
		(4.171)	
wtFix x cyl: 8/4		-0.782	
		(3.819)	
R-squared	0.866	0.907	0.904
adj. R-squared	0.840	0.856	0.870
p	0.000	0.000	0.000

The table above shows the output of the 3 regression models used. Model 1 utilized a step-wise regression, Model 2 utilized a step-wise with interactions and Model 3 was manually trimmed to maximize Rsq and Rsq adjust (by dropping non-significant terms)

```
anova(step.fit, step.fit.wI, best.fit.wI, step.fit)[6]
```

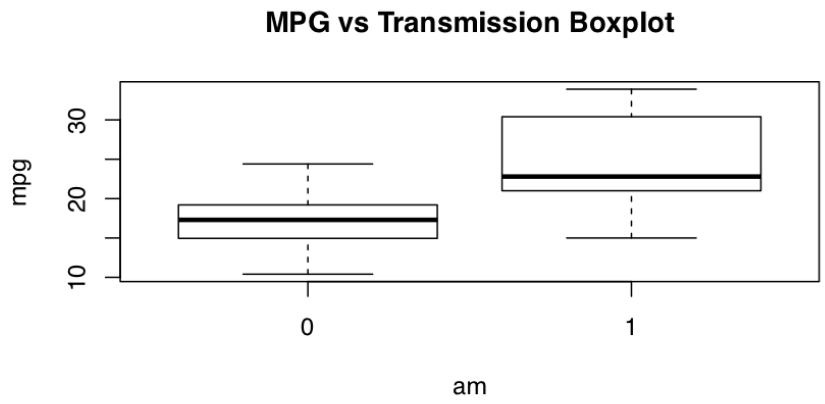
```
##    Pr(>F)
## 1
## 2 0.23107
## 3 0.84623
## 4 0.07199 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From the anova table we can see that model 2 is better than model 1 (line 2) but not better than model 3 (line 3). Also line 4 shows that Model 3 is better than model 1.

Conclusions: Manual transmissions do have a significant effect on mpg performance. (see detailed LM output in appendix) along with some of their key interaction variables. See detail in the executive summary.

Appendix:

Boxplot for MPG vs Transmission 0= Automatic 1= Manual Transmission



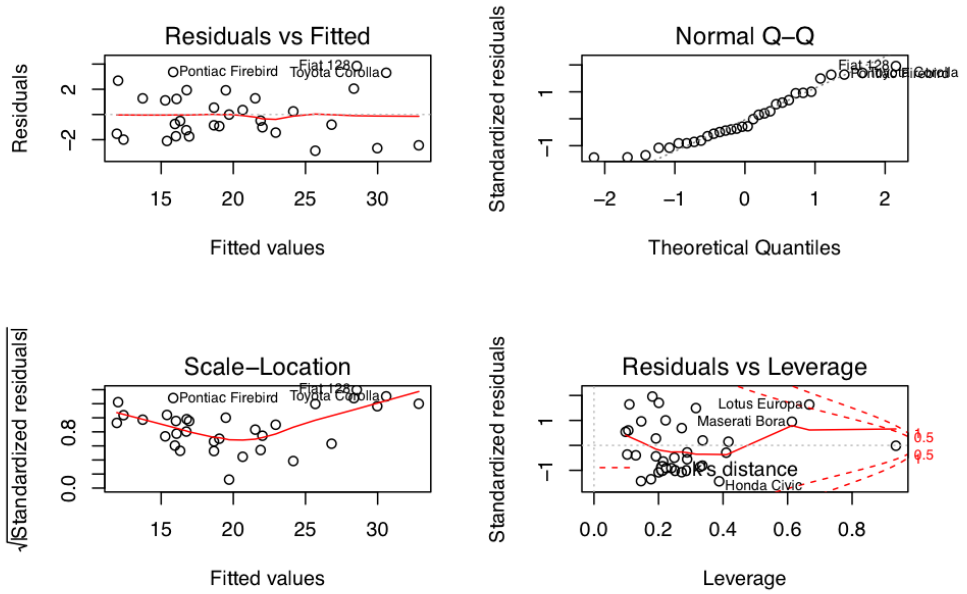
Full result of final regression (best.fit.wI)- note other regressions suppressed to conserve space

```
summary(best.fit.wI)
```

```
##
## Call:
## lm(formula = mpg ~ wtFix:am + cyl:hpFix + am + wtFix + cyl +
##     hpFix, data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.8777 -1.4603 -0.5024  1.2795  3.8468
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 28.85999    2.04992  14.079 8.57e-13 ***
## am1          4.53295    2.06844   2.191  0.03881 *
## wtFix       -2.31293    0.81181  -2.849  0.00908 **
## cyl6        -4.10116    3.17837  -1.290  0.20976
## cyl8        -6.85165    2.78119  -2.464  0.02166 *
## hpFix       -0.08268    0.03401  -2.431  0.02326 *
## wtFix:am1   -3.04685    1.51646  -2.009  0.05639 .
## cyl6:hpFix   0.05954    0.05035   1.182  0.24913
## cyl8:hpFix   0.07634    0.03565   2.142  0.04305 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.172 on 23 degrees of freedom
## Multiple R-squared:  0.9036, Adjusted R-squared:  0.8701
## F-statistic: 26.95 on 8 and 23 DF,  p-value: 6.391e-10
```

Diagnositc plots for best.fit.wI-

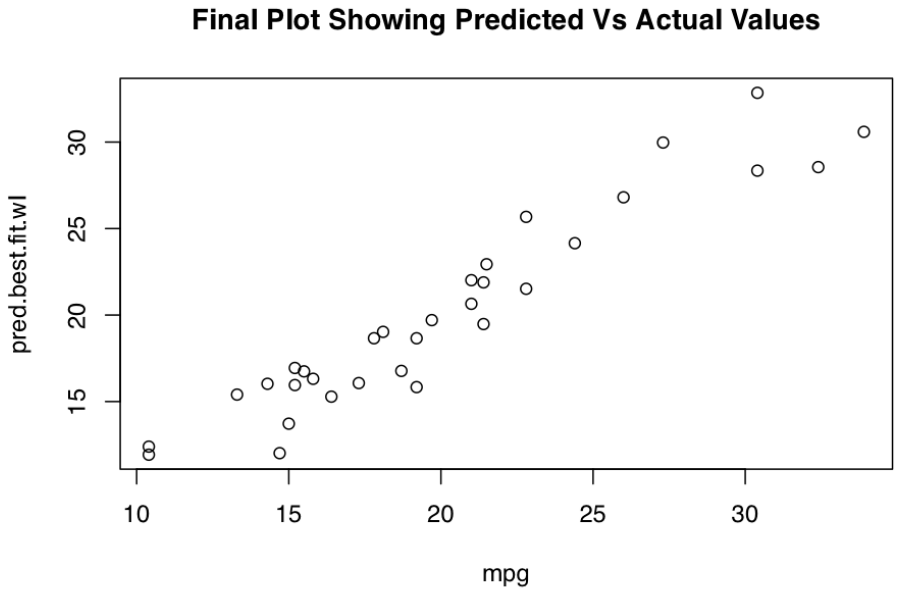
```
par(mfrow=c(2,2));
plot(best.fit.wI)
```



Analysis:

- Residuals vs Fit graph shows no defined pattern.
- The normality plot are normally distributed
- Overall there are a few outliers but these cannot be evaluated further within the scope of this project.

```
mtcars$pred.best.fit.wI<-predict(best.fit.wI)
par(mfrow=c(1,1))
plot(pred.best.fit.wI~ mpg, main="Final Plot Showing Predicted Vs Actual Values",  mtcars)
```



Analysis:

- This final plot of the Predicted values for mpg from the best.fit.wI model clearly aligns nicely to the actual MPG values that were given in the database
- This strong correlation does support the strength of the model
- Note there are a few outliers on the far right of the plot. Under normal circumstances it would be prudent to dig into these three points to look for other factors
- Suggestion for next study to include notes so these outliers could be diagnosed better