Will My Gas Money Get Me Further If My Car Has A Manual Transmission ?

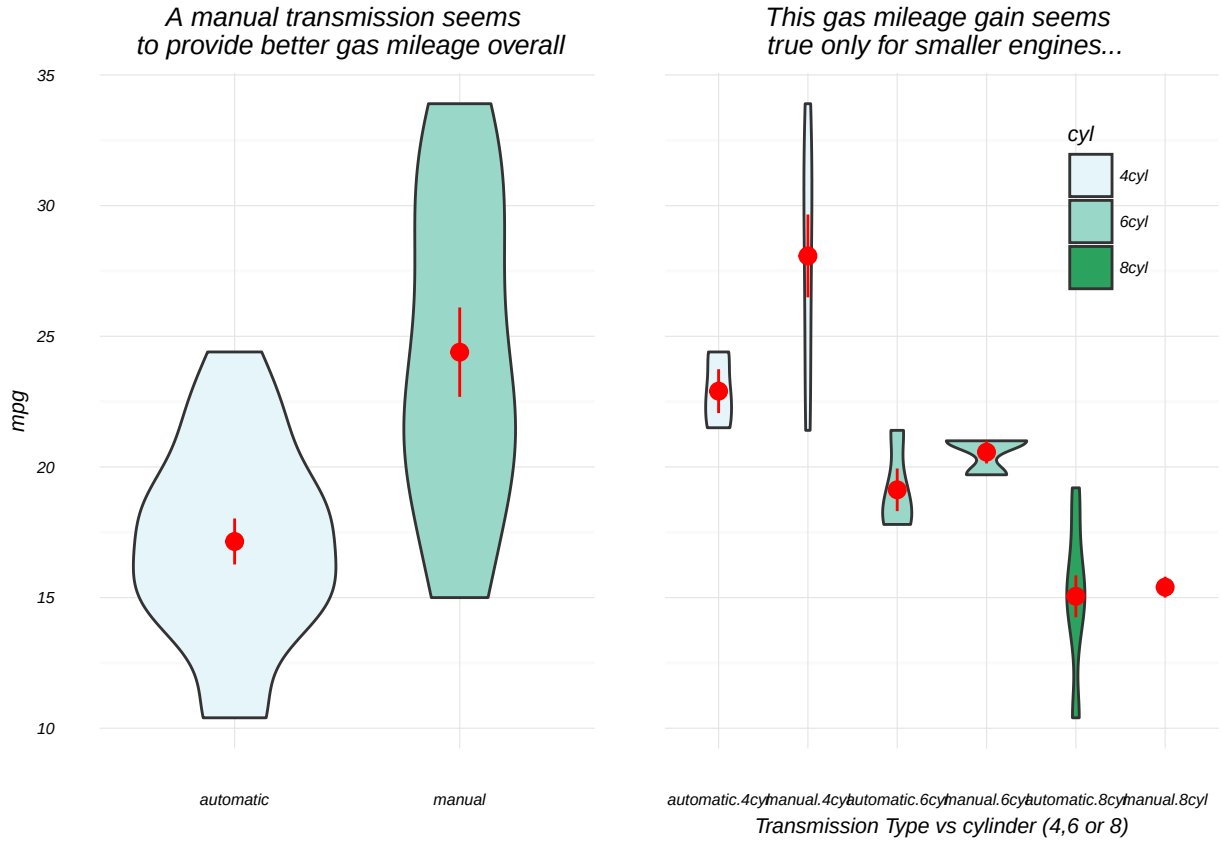
JRB

October 10, 2016

Executive Summary

In this short report we will evaluate the comparative gas efficiency of automatic vs. manual cars. We will use linear models to try explain fully gas efficiency. Our analysis reveals that on average manual cars offer better fuel efficiceny but this is only statistically significant for smaller engines (4 cylinder).

Exploratory Data Analysis



This simple data exploration helps us generate two hypothesis we need to verify...
1. *Hypothesis #1:* Manual transmissions provides better mpg
2. *Hypothesis #2:* Better mpg is only significant for smaller engines

Model selection

We will explore there multivariate regression models:

- Model 1: `mpg ~ factor(am)` mpg against transmission type

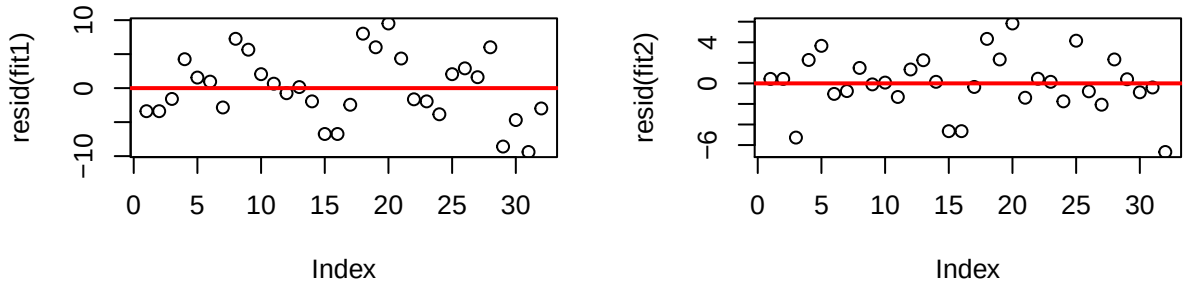


Figure 1: Residuals plots for model 1 and 2

- Model 2: `mpg ~ factor(am) + factor(am):factor(cyl)` mpg against transmission type as a regressor and an interaction term between number of cylinders and transmission
- Model 3: `mpg ~ factor(am) + factor(cyl) + factor(am):factor(cyl)` mpg against transmission type an dnumber of cylinders as regressors and an intercation factor bewteen transmission and number of cylinders.

Table 1: Analysis of Variance Table

Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
30	720.9	NA	NA	NA	NA
26	239.1	4	481.8	13.1	0
26	239.1	0	0	NA	NA

The anova test reveals that model 2 is the most approriate we could pick. Model 3 is not adding any additional information.
We plot (figure 1) the residuals of the two models we will be using and we see that there's no obvious pattern in the residual that would impact the models we chose.

Is there a mpg gain with manual cars ?

Model 1 is interesting to use to prove the assumption that on average (without distinction on cylinder count) manual transmission have better fuel economy.

Table 2: Manual vs Automatic mpg - Model 1

	Dependent variable:
	mpg
Manual Transimission (intercept)	24.39***
	(1.36)
Automatic Transmission	−7.24***
	(1.76)
Observations	32
R^2	0.36
Adjusted R^2	0.34
Residual Std. Error	4.90 (df = 30)
F Statistic	16.86*** (df = 1; 30)
Note:	*p<0.1; **p<0.05; ***p<0.01

We re-leveled the factors of `mtcars` to use manual as a reference point. The coeficient is interpreted by the difference of the mean between manual cars mpg and automatic cars mpg. A negative coefficient indicates a better mpg for manual cars. The significance level tells us that the coefficient is different that 0. This is confirmed by a t.test reproduce in appendix 1. Using this model we can quantify the average gain in mpg for manual cars using a confidence interval on the coefficient. The mpg gain for manual cars is between `10.85` and `3.64` So we can state that on average manual cars are more gas efficient than automatic cars.

Is this mpg gain true for all cylinders ?

Table 3:

	Dependent variable:
	mpg
Manual(4 cyl) - Intercept	28.075^{***} (1.072)
Automatic (4 cyl)	-5.175^{**} (2.053)
6 Cyl Manual	-7.508^{***} (2.053)
6 Cyl Automatic	-3.775 (2.316)
8 Cyl Manual	-12.675^{***} (2.397)
8 Cyl Automatic	-7.850^{***} (1.957)
Observations	32
R^2	0.788
Adjusted R^2	0.747
Residual Std. Error	3.032 (df = 26)
F Statistic	19.294^{***} (df = 5; 26)
Note:	*p<0.1; **p<0.05; ***p<0.01

Adding an interaction term (Model 2) between transmission and engine type, the model shows that only manual 4 cylinder cars have a better fuel ecomomy than automatic cars as shown in Table 3. All slope coefficients are negative from the point of reference (Manual 4cyl), all coefficients are either equal to zero (6 cylinder doesn't reject H_0) or different than 0. The t-test in appendix 1 reaches the same conclusion and further refine the assumption that the mpg gain is only true for smaller (4 cyl engine).We can quantify this gain for 4 cylinders engine for manual vs automatic using a 95% confidence interval on the first coefficient ("Automatic (4cyl)"). The mpg gain (on average) for manual 4 cylinder cars vs. automatic 4 cylinder car is `5.175` $\pm$ `4.22` The difference in means between automatic and manual cars is not statistically significant for 6 and 8 cylinder engines.

Conclusion

We were able to use linear regressions models to explore two hypothesis that our exploratory data analysis had suggested. We've concluded that while on average for the whole sample manual cars have superior gas mileage when we introduce a cylinder effect term to our model we established that the gas mileage superiority was only statistically significant for manual 4-cylinder cars versus other types of cars.

All models are bad but some models are useful....

Appendix 1

t-test of automatic vs. manual mpg

Table 4: Welch Two Sample t-test: `mycars$mpg` by `mycars$am` The t-test is of the mpg variable against the am variable as a factor with two levels and using manual as the reference level. so H_a : $\mu_{manual} > \mu_{automatic}$. We conclude that manual cars have a better mpg than automatic ones on average and without consideration for any other possible interactions.

Test statistic	df	P value	Alternative hypothesis
3.77	18.33	0 * * *	greater

t-test and confidence intervals of cyl vs am gas mileage

```
test4cyl <- t.test(mpg ~ am, subset(mycars, cyl == "4cyl"), alternative = "g")
test6cyl <- t.test(mpg ~ am, subset(mycars, cyl == "6cyl"))
test8cyl <- t.test(mpg ~ am, subset(mycars, cyl == "8cyl"))
```

Table 5: 4 cylinder manual cars shows greater fuel efficienty than automatic 4 cylinder cars

Test statistic	df	P value	Alternative hypothesis
2.89	9	0.01 * *	greater

Table 6: There's no statistical difference in fuel effiency between manual and automatic 6 cylinders cars

Test statistic	df	P value	Alternative hypothesis
1.56	4.41	0.19	two.sided

Table 7: There's no statistical difference in fuel effiency between manual and automatic 8 cylinders cars

Test statistic	df	P value	Alternative hypothesis
0.39	10.19	0.7	two.sided