Analyzing the effect of transmition type on MPG

Jamamel

Abstract

The following analysis identifies alternative scenarios in car MPG efficiency attributable to different types of transmition (automatic vs. manual). These are described through statistical inference and provide a quantifiable measure of benefit from one type compared to the other. All analyses use R dataset *mtcars*, and combine a series of exploratory and regression based techniques to quantify efficiency based on observed data.

Objective

We use dataset *mtcars* from R package *datasets* to determine the impact of several potential variables on fuel consumption efficiency. Specifically, two questions will be answered:

1. Is an automatic or manual transmission better for miles-per-galon (*MPG*)?
2. How much better (measured by change in *MPG*) is one type of transmition vs. the other?

We start by identifying possible correlation among variables. A series of exploratory analyses helped identify candidate relationships; specifically transmission type (automatic vs. manual; variable *am*) & mpg (variable *mpg*).

mtcars contains data for 32 unique car models, with the following measures:

Measure (Variable)	Definition
mpg	Miles per gallon
cyl	Number of cylinders
disp	Displacement (cubic inches)
hp	Gross horsepower
drat	Rear axle ration
wt	Weight (lb/1000)
qsec	Quarter mile time
vs	V vs. Straight engine
am	Transmission (automatic vs. manual)
gear	Number of forward gears
carb	Number of carburetors

Exploring & Modelling MPG

A combination of correlation, and plotting was performed. *mpg* has high correlation with a few variables relating to car dimensions. Also, a car's weight (*wt*), displacement (*disp*), and gross horse power (*hp*) are highly correlated. This suggests any linear-based regression modelling approach could confound simultaneous effects (see *Figure 1*). There also appear to be clear relationships between transmision type (*am*) and attributes like number of cylinders (*cyl*) or carburetors (*carb*), or engine type (*vs*). However, the distribution of the number of forward gears (*gear*) is less clear, perhaps due to the small number of observations in our dataset (see *Figure 2*). While some patterns are visible between transmission and other attributes, it is sometimes difficult to determine whether these patterns are logical and expected, rather than purely random observations in this particular data. While manual transmission cars tend to have more forward gears,

cylinders, and carburetors, there's no immediate evidence of these being linked more directly to engine type.

Next, a few of the relationships identified were modelled through linear regression to try and quantify the effect of different car specs on *MPG*, namely transmission type. The exploratory phase confirmed linear regression was a suitable approach given *mpg* fits the assumption of Normality. After a series of model iterations (described in comments of R Markdown file), the best model was that in which *mpg* was described by the interaction between type of transmision and weight, number of cylinders (as a factor referencing 4 cylinders), type of transmision on its own, and two dummy variables for Toyota Corolla and Fiat 128 cars. These 2 cars were modelled uniquely given their leverage on the model (see results below) and unique behaviour.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	23.908	1.359	17.590	0.000
ammanual	-1.725	0.974	-1.772	0.089
cwt	-2.478	0.637	-3.889	0.001
cyl	-2.180	0.591	-3.692	0.001
toy	6.024	1.908	3.156	0.004
fiat	6.603	1.852	3.565	0.001
ammanual:cwt	-3.218	1.045	-3.079	0.005

All estimated coefficients have acceptable p-values (lowest confidence is for type single effect at 0.91), and adjusted overall fit stands at a high **0.92 with 25 degrees of freedom** . Of more interest is the fact the residual standard error is low at **1.75**. Residual analysis also shows all residuals are within 1.5 standard deviations of the mean. Normality is fairly consistent apart from a couple of cases which don't have enough leverage on the analysis to make a meaningful difference (*See Figure 3*).

Results & Conclusions

At average car weight, manual transmision on its own appears to have a slight negative effect on MPG (-1.73 vs. the equivalent automatic transmision). This effect, however, is mostly compensated by other usual attributes automatic transmision cars which greatly hinder MPG. For every ton above the average, MPG decreases by -2.48, and for every extra cylinder, MPG decreases by a further -2.18. When one considers the average difference in weight and usually higher number of cylinders of automatic transmision cars, the counter effect becomes clear.

am	mpg	weight	cylinders	n
automatic	17	3.8	3	19
manual	24	2.4	1	13

The model does suggest that an increase of one ton for manual transmision vehicles has a much more profound effect on the efficiency of the car (decrease of -3.22 MPG).

While the model suggests a manual transmision on its own is not enough to make a difference, it sits at the core of a lot dimensions that ultimately define how efficient a car is. Automatic transmisions are difficult to place in smaller cars, rather these usually tend to be heavier vehicles with more cylinders. What is clear, is that designing a manual transmision vehicle that's small, powerful, and efficient enough is a challenge. At the same time manual transmision plays a fundamental role in allowing the car to be lighter and with a smaller engine, essential requirements for a more efficient car.

References

- MotorTrendMPG Github Repo

Appendix

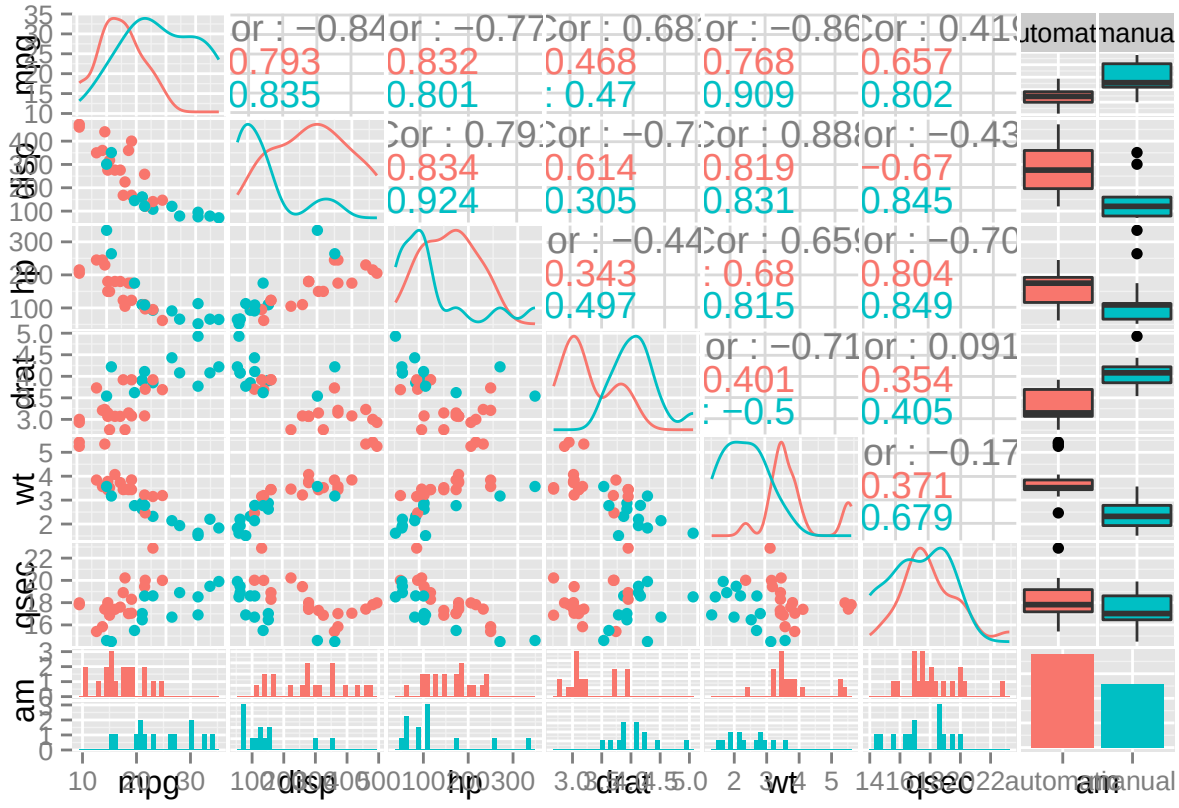

Figure 1: Correlation of car dimensions to mpg split by transmision type.

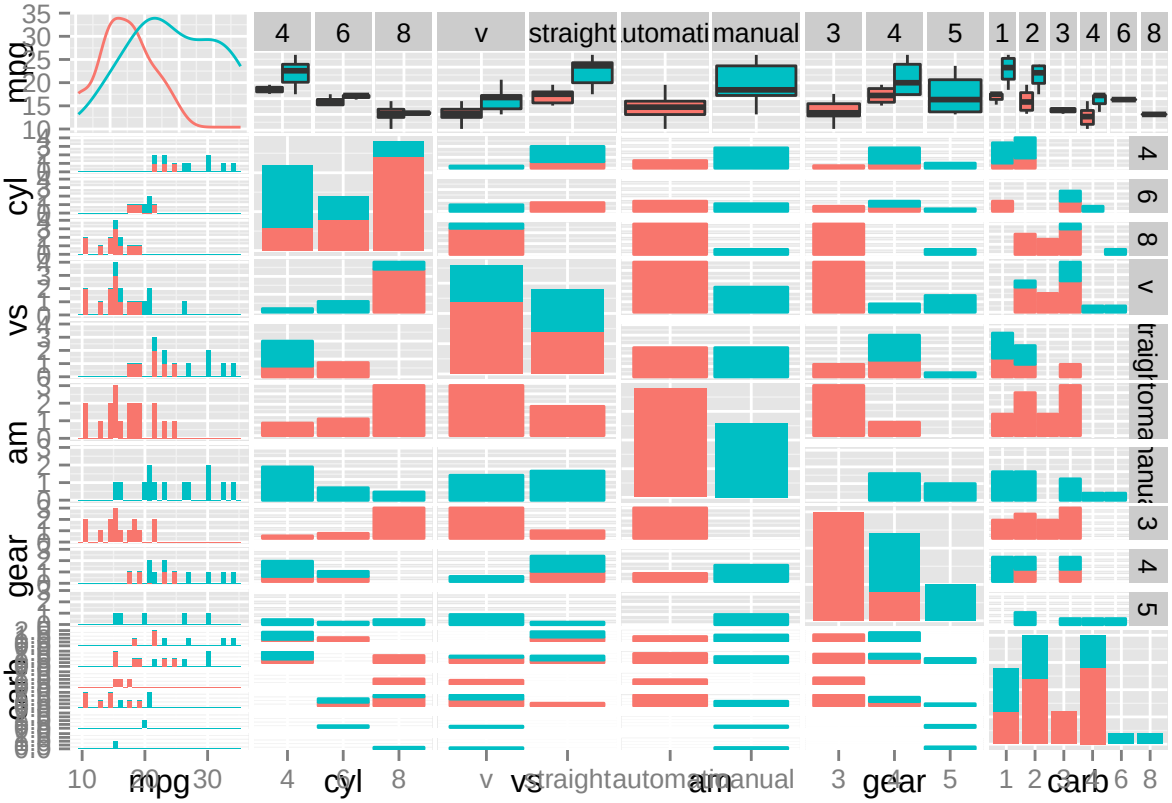

Figure 2: Distribution of car attributes to and mpg split by transmision type.

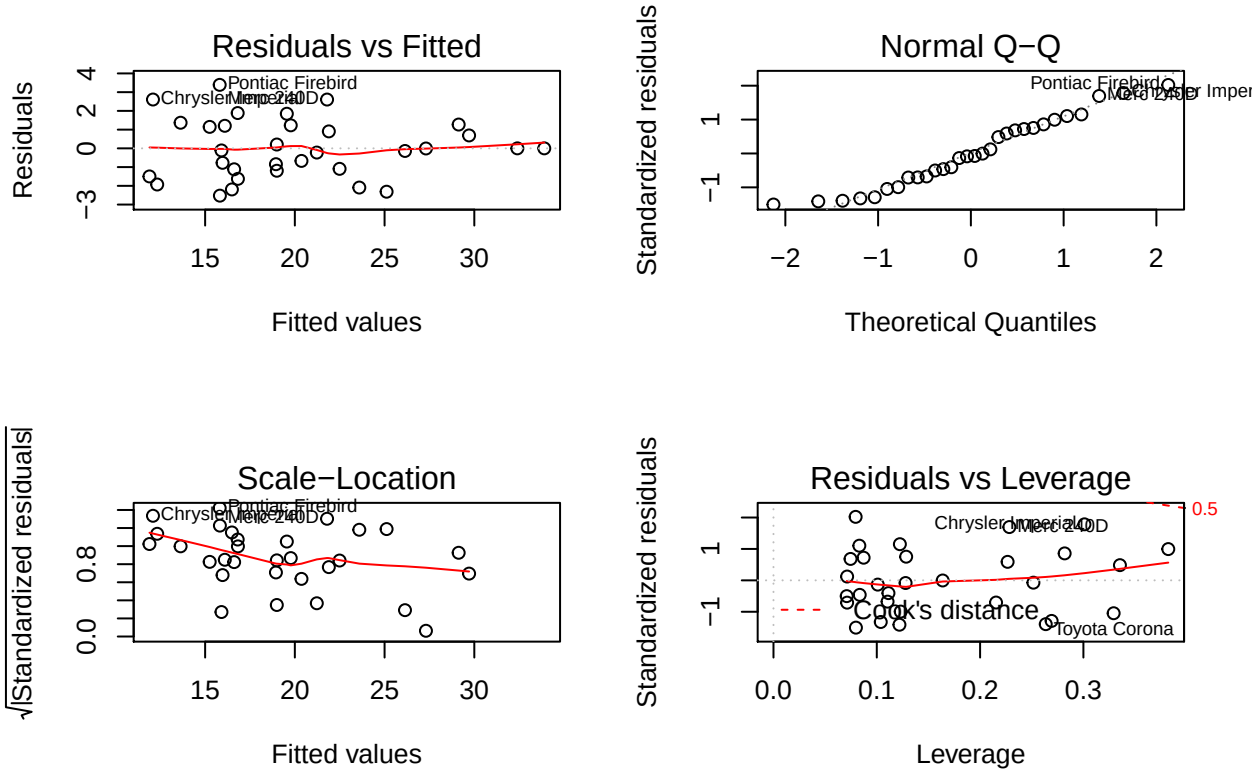

Figure 3: Final model Fit Diagnostics