

Effect of Transmission Type on Fuel Efficiency for “mtcars” Data Set

Executive Summary:

In terms of standardized fuel efficiency, the difference between an automatic transmission and a manual transmission is negligible. For this analysis, the “mtcars” data set was “subsetting” by weight in order that the analysis address only comparable vehicles. The vehicle weight range in the analysis is 2465# to 3570#. The average Miles Per Gallon (MPG) and standard deviation for the comparable automatic transmission vehicles are 18.98MPG and 3.27MPG respectively. The average MPG and standard deviation for the comparable manual transmission vehicles are 18.99MPG and 2.85MPG respectively.

Analysis:

Eleven parameters are addressed in the “mtcars” data set (see appendix figure “mtcars Parameters” for full list). MPG was linearly regressed against the other 10 parameters; $\text{lm}(\text{mpg} \sim . -1, \text{data}=\text{mtcars})$. The primary parameters influencing MPG are weight (slope = -3.83) and transmission type (slope = 2.83). The next most influential parameter is rear axle ratio (slope = 1.24, less than half of transmission). Also, in reviewing the plots of the parameter pairs (see appendix figure “Pairs Plots for mtcars Data Set”) “weight” and the effect of “weight” on the other parameters is the primary driver of standardized fuel efficiency the effect/influence of weight is apparent. Based on the above, weight is used as the influencing parameter in this analysis of the impact of transmission type on MPG.

A box-and-whisker plot of fuel efficiency versus transmission type shows a very large difference in MPG (see appendix figure “Fuel Efficiency for All 32 Cars in mtcars Data Set”). The average fuel efficiencies are 17.15MPG for the automatic transmission vehicles and 24.39MPG for the manual transmission vehicles. Noting that all of the “very low” MPG vehicles are also the “high weight” vehicles it is obvious that weight is the driving parameter (weight is also the major driver in required horsepower/displacement, rear end ratio, gearing, etc). To get an analysis set of vehicles that are comparable for fuel efficiency purposes a subset of “mtcars” is used. The subset is composed of manual and automatic transmission vehicles in the region in which the vehicle weights overlap; i.e., from the minimum weight automatic transmission vehicle (2465#) to the maximum weight manual transmission vehicle (3570#). The subset contains 6 manual transmission vehicles and 11 automatic transmission vehicles. The vehicles and parameters in the analysis subset are;

<u>Manual trans:</u>	mpg	cyl	displacement	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
Ford Pantera L	15.8	8	351	264	4.22	3.170	14.50	0	1	5	4
Ferrari Dino	19.7	6	145	175	3.62	2.770	15.50	0	1	5	6
Maserati Bora	15.0	8	301	335	3.54	3.570	14.60	0	1	5	8
Volvo 142E	21.4	4	121	109	4.11	2.780	18.60	1	1	4	2

<u>Automatic trans:</u>	mpg	cyl	displacement	hp	drat	wt	qsec	vs	am	gear	carb
Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215	19.44	1	0	3	1
Hornet Sportabout	18.7	8	360.0	175	3.15	3.440	17.02	0	0	3	2
Valiant	18.1	6	225.0	105	2.76	3.460	20.22	1	0	3	1
Duster 360	14.3	8	360.0	245	3.21	3.570	15.84	0	0	3	4
Merc 240D	24.4	4	146.7	62	3.69	3.190	20.00	1	0	4	2
Merc 230	22.8	4	140.8	95	3.92	3.150	22.90	1	0	4	2
Merc 280	19.2	6	167.6	123	3.92	3.440	18.30	1	0	4	4
Merc 280C	17.8	6	167.6	123	3.92	3.440	18.90	1	0	4	4
Toyota Corona	21.5	4	120.1	97	3.70	2.465	20.01	1	0	3	1

Dodge Challenger	15.5	8	318.0	150	2.76	3.520	16.87	0	0	3	2
AMC Javelin	15.2	8	304.0	150	3.15	3.435	17.30	0	0	3	2

For this subset, average Miles Per Gallon (MPG) and standard deviation for the automatic transmission vehicles are 18.98MPG and 3.27MPG respectively. The average MPG and standard deviation for the subset manual transmission vehicles are 18.99MPG and 2.85MPG respectively.

A “t test” for the null hypothesis that the average MPG's are equal has a P-value of over 0.99.

A linear regression fit (lm) was run for each transmission type. Plots of the residuals are presented in appendix figures “Residuals for Automatic Transmission Subset Model” and “Residuals for Manual Transmission Subset Model.” There are no obvious patterns in the residuals plots. In the “automatic residuals plot” the Toyota Corona (vehicle 9) appears to be an outlier. In the “manual residuals plot” the Ford Pantera L (vehicle 3) and the Maserati Bora (vehicle 5) may be outliers. Influence measures “tools,” “hatvalues” (relative leverage) and “dfbetas” (coefficient sensitivity) were run to identify anomalies. The “hatvalues” results showed that the Toyota Corona and the Maserati Bora are high leverage data points as expected. The leverage of the Ford Pantera L was not out of line with respect to the other vehicles. The “dfbetas” indicate potential problems with the Toyota Corona and the Maserati Bora. The “dfbetas” results for the Ford Pantera L were higher than the other four data points, but less than half of the Maserati Bora.

Discussion and conclusions:

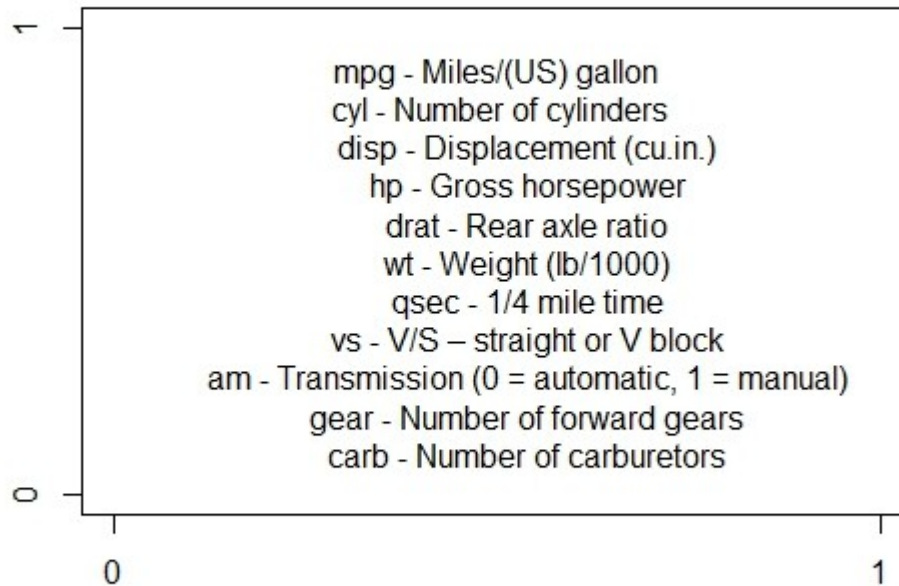
The analysis is summarized in appendix figure “Affect of Transmission Type on MPG: mtcars Data Set.” The plot presents fuel efficiency as a function of vehicle weight. All 32 “mtcars” data points are plotted. Data for the automatic transmission vehicles is in blue. Data for the manual transmission vehicles is in red. The vertical lines are the weight subset limits used for the analysis; all analyses address only the points within the subset limits. The right “red” line is at the maximum weight of the manual transmission vehicles. The left “blue” line is at the minimum weight of the automatic transmission vehicles. The solid slanted lines are the linear regression fits (manual transmission fit: $41.073 - 7.452 \cdot \text{wt}$ and automatic transmission fit: $40.711 - 6.577 \cdot \text{wt}$). The dashed lines are at the 95% confidence limits. The 3 points that were flagged by the “hatvalues” and “dfbetas” (Toyota Corona, Ford Pantera L, and Maserati Bora) are labeled.

The confidence intervals do not overlap to the extent expected based on the P-value. The Toyota Corona is significantly separated from the other “automatic transmission” vehicles and does appear to be an outlier. Eliminating the Corona would actually reduce the average fuel efficiency of the “automatic transmission” and would be counterproductive for the comparison. The Corona was not eliminated. The Maserati Bora essentially falls on the regression fit and therefore its influence is minimal. The Ford Pantera L is just inside the confidence limit. Neither the Maserati nor the Ford were eliminated from the analyses.

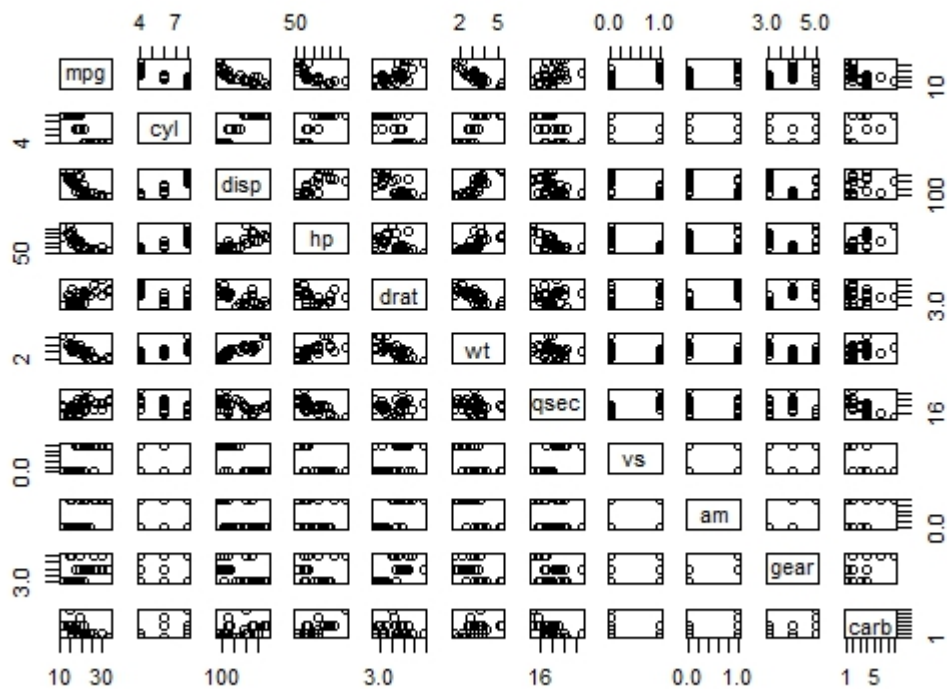
In terms of standardized fuel efficiency, the difference between an automatic transmission and a manual transmission is negligible. Contrary to popular belief, if there is a bias for better fuel efficiency it will favor the automatic transmission, but it will difficult to prove the bias based on only the “mtcars” data set.

Appendix:

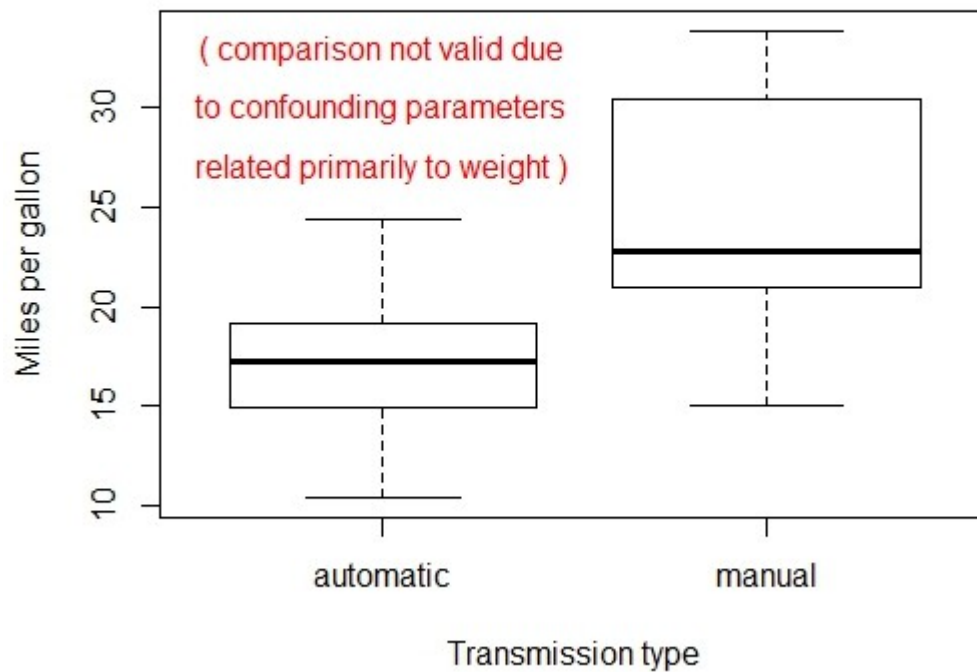
mtcars Parameters



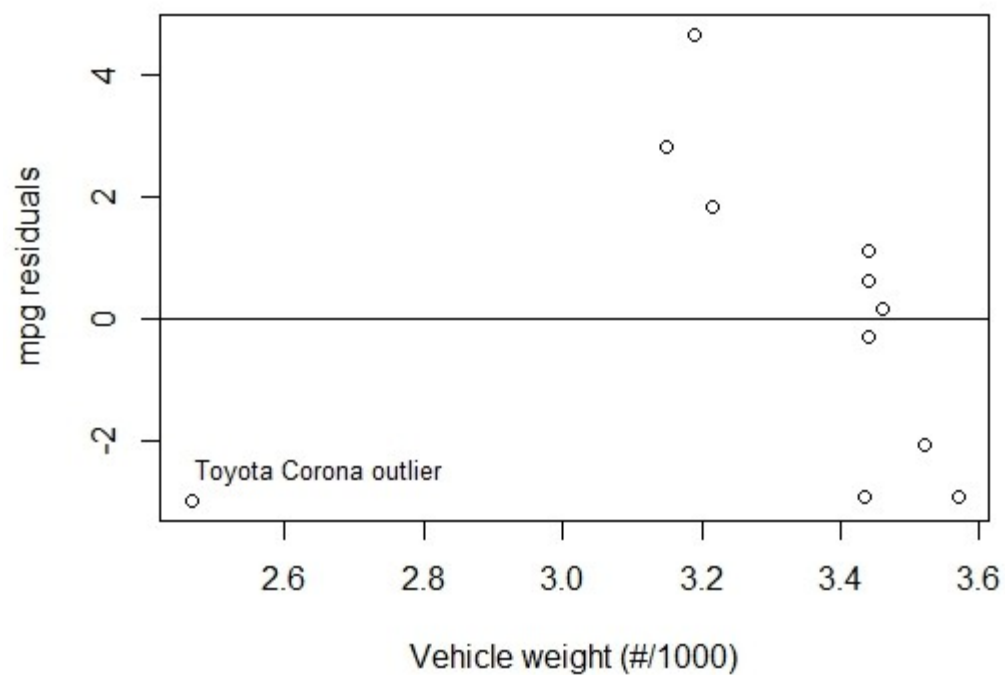
Pairs Plots for mtcars Data Set



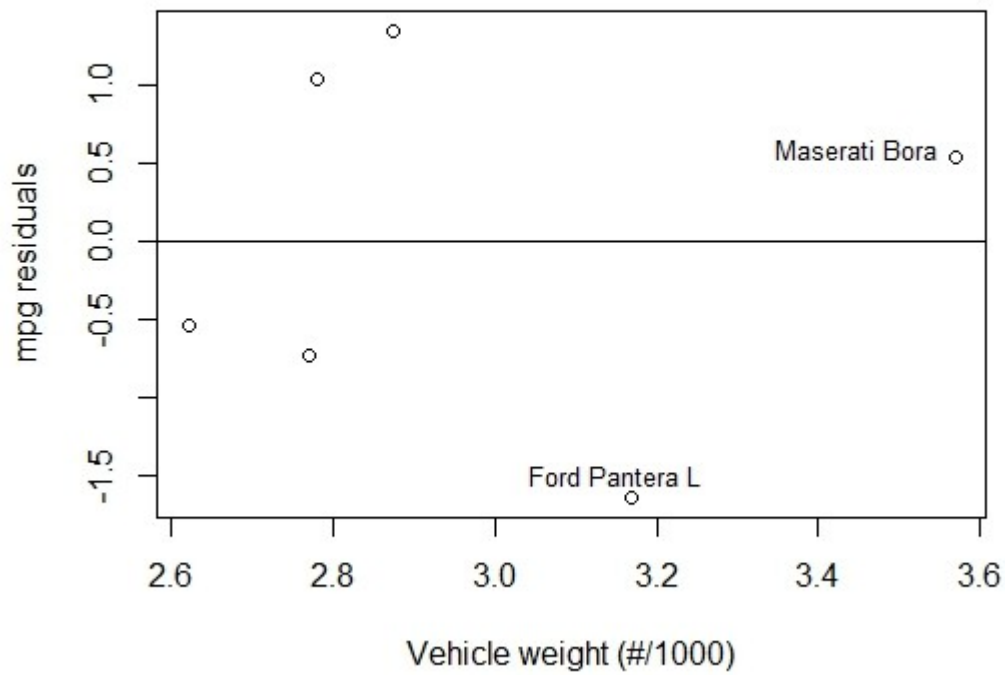
Fuel Efficiency for All 32 Cars in mtcars Data Set



Residuals for Automatic Transmission Subset Model



Residuals for Manual Transmission Subset Model



Affect of Transmission Type on MPG: mtcars Data Set

