

Exploring MPG and transmission type

K. Kurttila - November 19, 2015

Executive Summary

In this report, we will utilize the dataset (mtcars) from the 1974 Motor Trend magazine, to analyze the difference between automatic and manual transmissions with regards to miles per gallon (MPG).

Our focus:

1. Is an automatic or manual transmission better for MPG?
2. Quantify the MPG difference between automatic and manual transmissions.

After performing exploratory data analysis, and using hypothesis testing and regression models, our multivariate model, accounting for the major confounders of weight, quantity of cylinders, and horsepower, show that cars with manual transmissions average 1.81 MPG more than cars with an automatic transmission, however, our p-value is greater than .05, indicating these results are not statistically significant and could be due to chance.

Data Processing

We load the dataset and convert the necessary variables to factors for analysis. (For complete code, see the RMD file here: <https://github.com/kellisphere/RegressionModel/blob/master/Final-RegModel-ExecReport.Rmd>)

Our data is comprised of the following variables with regard to 32 cars from the 1973-74 model years:

Variables: **mpg** - Miles/(US) gallon, **cyl** - Number of cylinders, **disp** - Displacement (cu.in.), **hp** - Gross horsepower, **drat** - Rear axle ratio, **wt** - Weight (lb/1000), **qsec** - 1/4 mile time, **vs** - V/S, **am** - Transmission, **gear** - Number of forward gears, **car** - Number of carburetors

See **Appendix; Figure 1** for a sampling of the first 6 rows of the data (each row is the observation of one car model).

Exploratory analysis

The quickest way to conduct exploratory analysis is graphically. First, a simple boxplot (**Appendix; Figure 2**) for visual analysis of the data for transmission type (am) with regards to MPG (mpg) indicate a significant difference in mean values based on transmission type. A pairs plot for each factor level compared to all others (**Appendix; Figure 3**) provides some insight into which factors could be confounders. By utilizing conditioning plots (**Appendix; Figures 4 a-c**) focusing on MPG, given transmission, with regards to cyl, hp, and wt, respectively, as confounders we can easily see a multivariate linear model will be the most accurate. To test this theory, we analyze the variance (using anova) between the singular and multivariate models.

```
## Analysis of Variance Table
##
## Model 1: mpg ~ am
## Model 2: mpg ~ am + wt + hp + cyl
##   Res.Df    RSS Df Sum of Sq    F    Pr(>F)
## 1      30 720.90
## 2      26 151.03   4    569.87 24.527 1.688e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Reviewing the results: Based on anova, we can confirm that our second model, the multivariate will be the best fit due to the significant p-value, and we reject the null hypothesis that wt, hp, and cyl don't impact the outcome.

```
##           Estimate Std. Error  t value    Pr(>|t|)
## (Intercept) 33.70832390 2.60488618 12.940421 7.733392e-13
## amManual    1.80921138 1.39630450  1.295714 2.064597e-01
## wt         -2.49682942 0.88558779 -2.819404 9.081408e-03
## hp         -0.03210943 0.01369257 -2.345025 2.693461e-02
## cyl6       -3.03134449 1.40728351 -2.154040 4.068272e-02
## cyl8       -2.16367532 2.28425172 -0.947214 3.522509e-01
```

Results of the multivariate model: (Appendix; Figure 5) The intercept 33.71 is the average miles per gallon for cars with automatic transmissions, the slope 1.81 is the increased miles per gallon for manual transmissions. The R Squared value tells us our model explains 86.59% of the total variance. Our p-value is 2.60, ($>.05$) have to accept the null hypothesis - there are too many factors impacting the difference in MPG.

Conclusion

Based on the observations of the multivariate regression model we see the following results:

- Cars with manual transmissions get slightly higher miles per gallon compared to cars with automatic transmission. (1.81 miles per gallon adjusted by horsepower, cylinders, and weight).
- Miles per gallon will decrease by 2.5 (adjusted by horsepower, cylinders, and transmission) for every 1000 pounds of increase in weight.
- Miles per gallon decreases negligibly with increase of horsepower.
- If number of cylinders are increased from 4 to 6 and 8, miles per gallon will decrease by a factor of 3 and 2.2 respectively (adjusted by horsepower, weight, and transmission type).

Because of the impact of the confounders on the outcome, we cannot directly say if a manual transmission or automatic transmission is better as there are no vehicles to review that are equal in confounders so only miles per gallon can be assessed.

Assumptions

We would need to assume that all road tests were run in the same weather conditions, terrain, and with similar driving skills - thereby not impacting outcome.

Appendix

FIGURE 1

	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	Manual	4	4
Mazda RX4 Wag	21.0	6	160	110	3.90	2.875	17.02	0	Manual	4	4
Datsun 710	22.8	4	108	93	3.85	2.320	18.61	1	Manual	4	1
Hornet 4 Drive	21.4	6	258	110	3.08	3.215	19.44	1	Automatic	3	1
Hornet Sportabout	18.7	8	360	175	3.15	3.440	17.02	0	Automatic	3	2
Valiant	18.1	6	225	105	2.76	3.460	20.22	1	Automatic	3	1

FIGURE 2

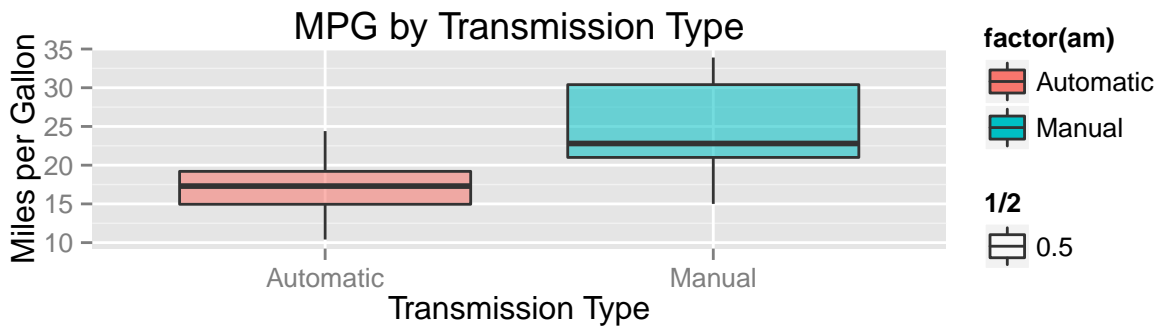


FIGURE 3

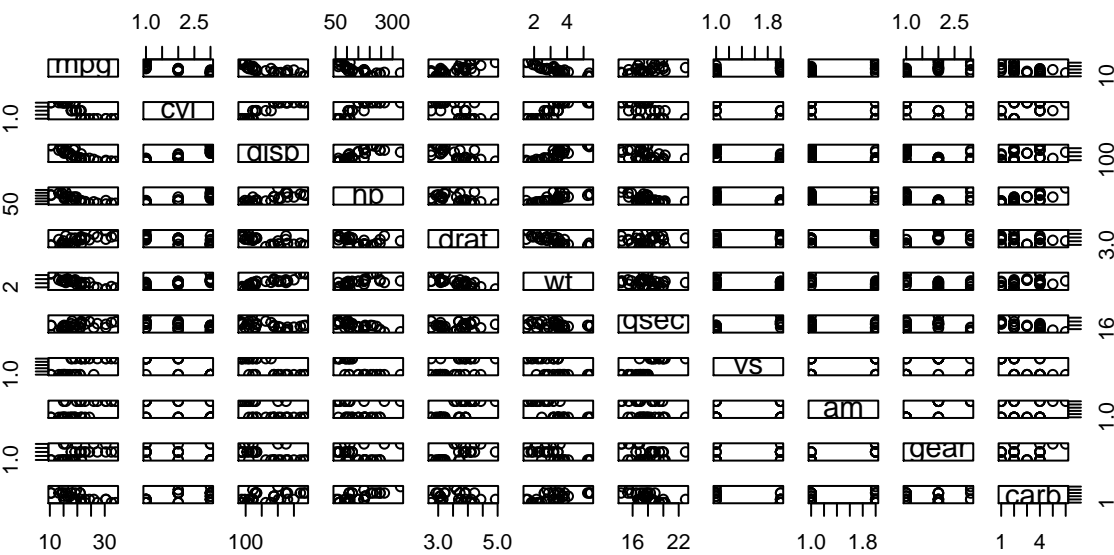


FIGURE 4a

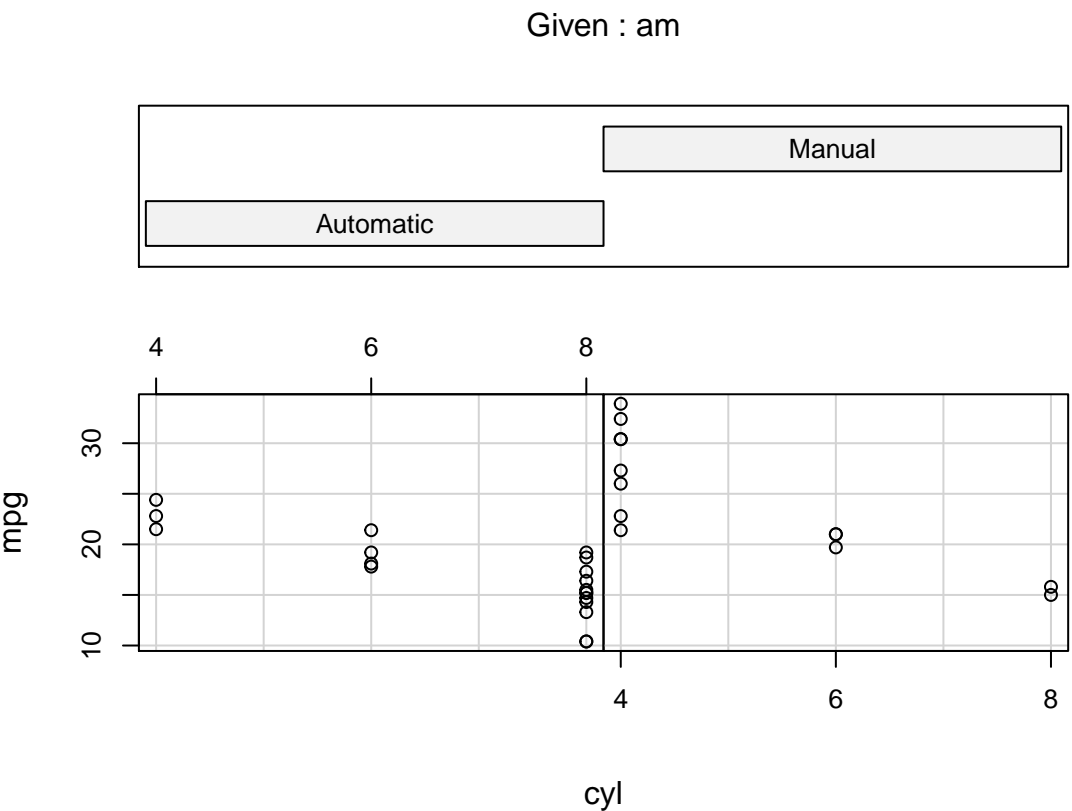


FIGURE 4b

Given : am

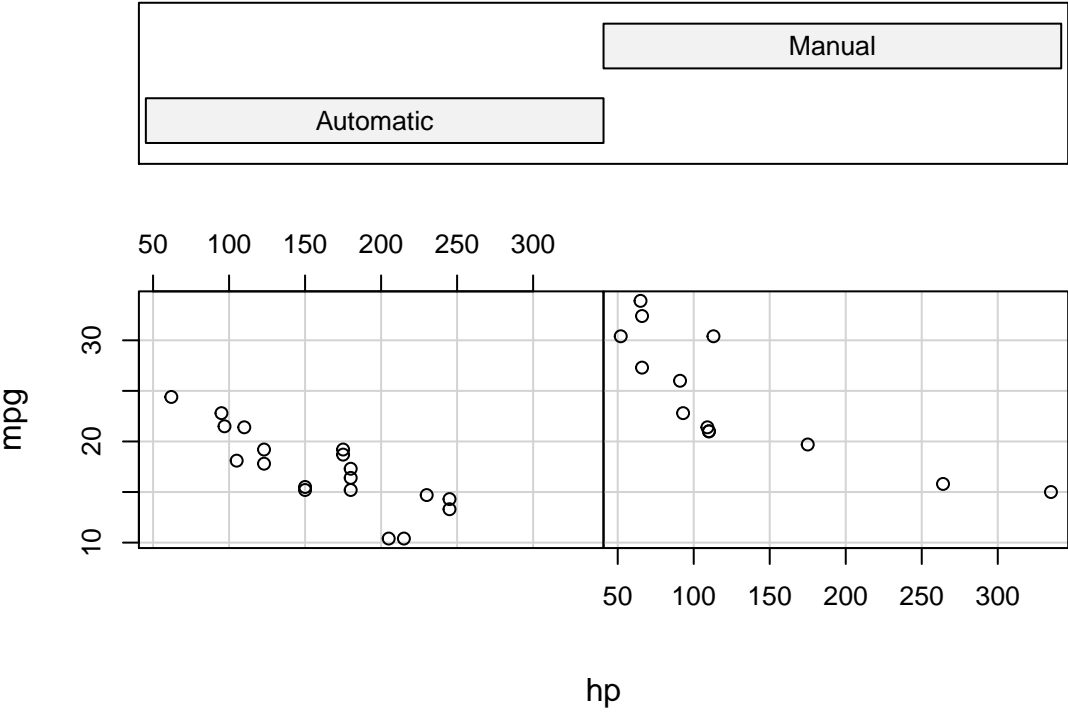


FIGURE 4c

Given : am

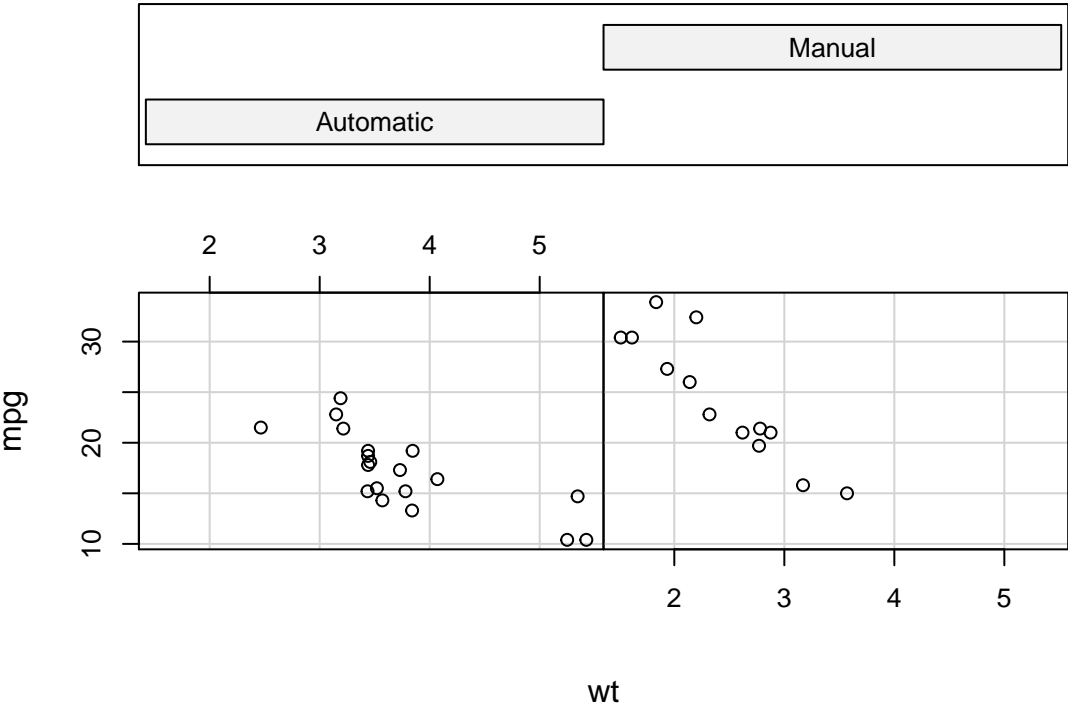
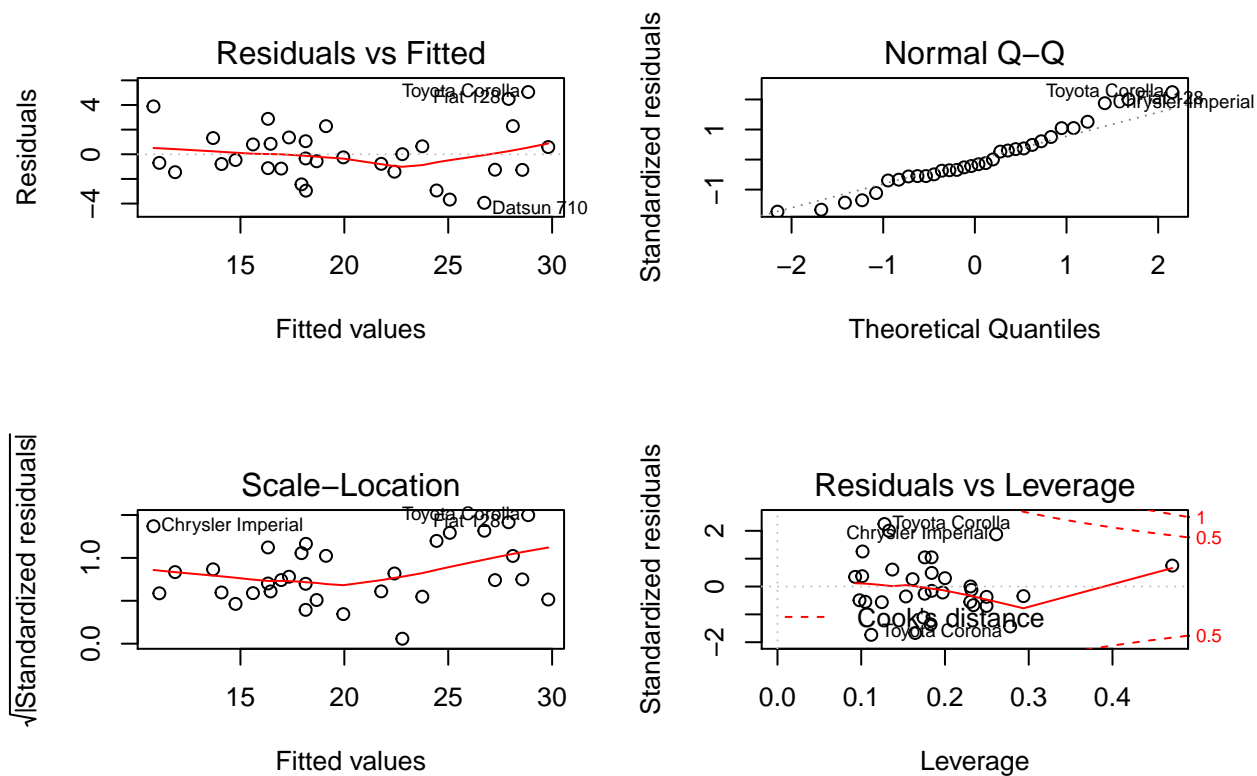


FIGURE 5**Figure 6**

```
sessionInfo() ## all libraries and tools used to create this report
```

```
## R version 3.2.0 (2015-04-16)
## Platform: x86_64-apple-darwin13.4.0 (64-bit)
## Running under: OS X 10.10.5 (Yosemite)
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats      graphics  grDevices utils      datasets  methods   base
##
## other attached packages:
## [1] broman_0.59-5 ggplot2_1.0.1 knitr_1.11
##
## loaded via a namespace (and not attached):
## [1] Rcpp_0.11.6      assertthat_0.1  digest_0.6.8    MASS_7.3-43
## [5] grid_3.2.0       plyr_1.8.3      gtable_0.1.2    formatR_1.2
## [9] magrittr_1.5     scales_0.2.5    evaluate_0.7.2   highr_0.5
## [13] stringi_0.5-5    reshape2_1.4.1  rmarkdown_0.8    labeling_0.3
## [17] proto_0.3-10     tools_3.2.0     stringr_1.0.0    munsell_0.4.2
## [21] yaml_2.1.13      colorspace_1.2-6 htmltools_0.2.6
```