Choosing a car to optimize miles per gallon

Marisa Casillas[1] & Natalie Dowling[1]

[1] University of Chicago

Author Note

Please email ndowling@uchicago.edu for questions about this important research.

Correspondence concerning this article should be addressed to Marisa Casillas, Rosenwald Hall, UChicago. E-mail: mcasillas@uchicago.edu

Abstract

Gas is expensive and burning it is bad for environmental health. How do I choose a car to optimize my gas mileage? We examine a few potential variables to help answer this question.

Keywords: Fuel efficiency, transmission, engine size

Choosing a car to optimize miles per gallon

1 Introduction

In the `mtcars` dataset (Henderson & Velleman, 1981) there are 32 cars documented, from 22 brands. The unique car types documented are: Mazda RX4, Mazda RX4 Wag, Datsun 710, Hornet 4 Drive, Hornet Sportabout, Valiant, Duster 360, Merc 240D, Merc 230, Merc 280, Merc 280C, Merc 450SE, Merc 450SL, Merc 450SLC, Cadillac Fleetwood, Lincoln Continental, Chrysler Imperial, Fiat 128, Honda Civic, Toyota Corolla, Toyota Corona, Dodge Challenger, AMC Javelin, Camaro Z28, Pontiac Firebird, Fiat X1-9, Porsche 914-2, Lotus Europa, Ford Pantera L, Ferrari Dino, Maserati Bora, and Volvo 142E. The mean mpg is 20 (median = 19.20; sd = 6.03; range = 10.40–33.90). The cars range in number of cylinders from 4 to 8, though there are no cars with odd numbers of cylinders.

We show the unique brands along with their average number of cylinders, average weight, and number of models represented in the dataset in Table 1.

2 Results

We modeled mileage with a linear mixed-effects regression, including fixed effects of number of cylinders (numeric), transmission (automatic or manual), and their interaction, as well as a random effect of brand.[1] Increases in number of cylinders was associated with significant decreases in mileage (B = -1.59, SE = 0.36, t = -4.41). Meanwhile, manual transmissions were associated with significant increases in mileage compared to automatic ones (B = 14.11, SE = 3.94, t = 3.59). That said, the decrease in mileage with more cylinders was significantly greater for manual transmission vehicles than automatic ones (B = -1.73, SE = 0.68, t = -2.55).

[1] `lmer(mileage ~ cylinders * transmission + (1|brand), data)`

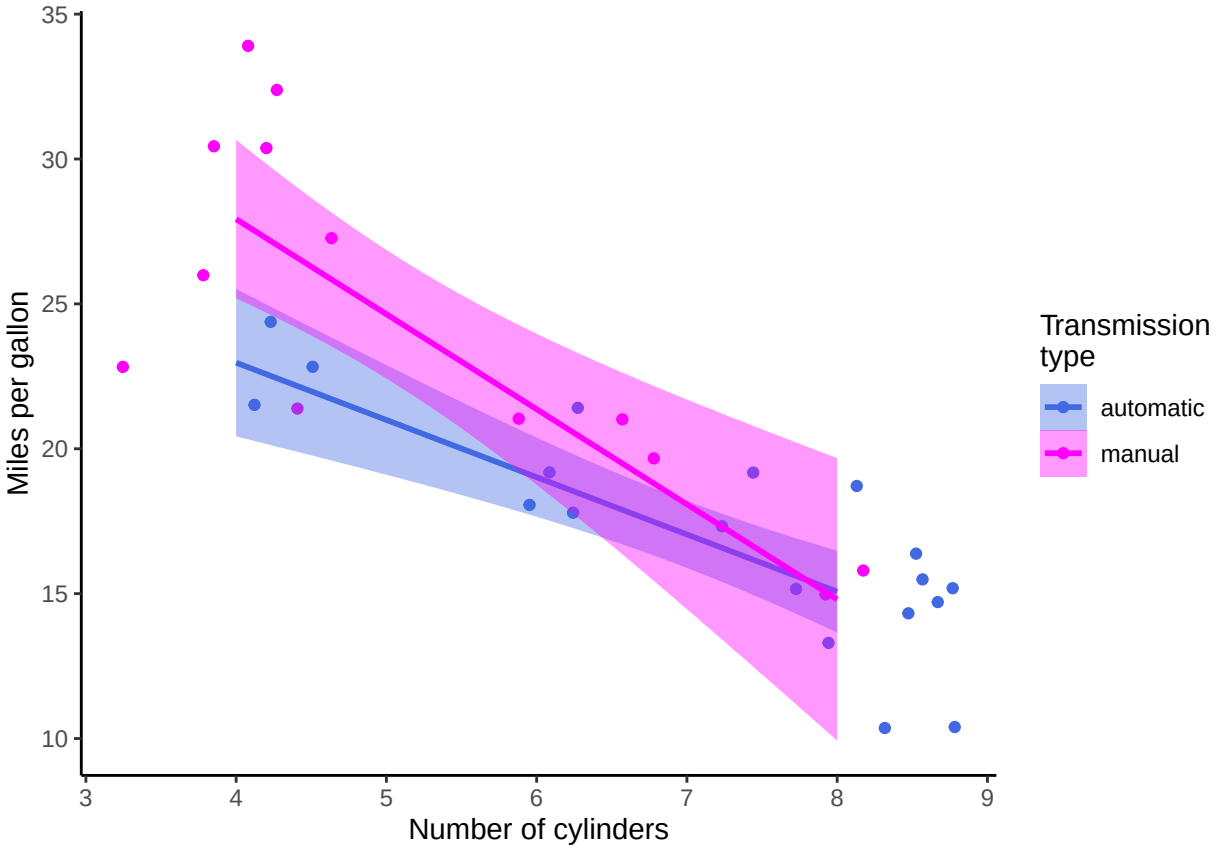

Figure 1. Miles per gallon as a function of nunber of cylinders and transmission type in the `mtcars` dataset.

These effects are visualized in Figure 1.

3 Discussion

In this study we investigated the relationship between miles per gallon, number of cylinders, and transmission type in the `mtcars` dataset. Unsurprisingly, cars with more cylinders and cars with automatic transmissions had lower mileage. The effect of transmission depended on the number of cylinders, such that the mileage benefit associated with manual transmissions decreased for cars with more cylinders. Based on these data, someone trying to maximize their mileage should choose a manual transmission, low-cylinder vehicle.

3.1 References we love

You may want to cite references in different formats depending on the surrounding sentential context, e.g.: Dowling, in her (2024) course, states that RMarkdown is awesome. RMarkdown is awesome (Dowling, 2021a, 2024; Xie, Allaire, & Grolemund, 2018). RMarkdown is awesome (Dowling, 2024; e.g., Xie et al., 2018). Same-year, same-author publications are automatically disambiguated (e.g., Dowling, 2021a, 2021b). We learned all about ggplot (see Wickham & Grolemund, 2016, ch. 1). We learned about ggplot with Wickham and Grolemund (2016, ch. 1).

This reproducible manuscript is written in RMarkdown (Xie et al., 2018), using the lme4 (Bates, Mächler, Bolker, & Walker, 2015) to run statistical models and ggplot2 (Wickham, 2016) within the tidyverse package (Wickham & Grolemund, 2016) to generate plots.

We summarize our findings in subsection 3 and note that subsection 3.1 is unrelated to this report.

4 References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Douglas, A., Roos, D., Mancini, F., Couto, A., & Lusseau, D. (2022). *An Introduction to R*. https://intro2r.com/.

Dowling, N. (2021a). *Reproducible science is great*. Various manuscripts.

Dowling, N. (2021b). *What a productive year*. Two publications!!

Dowling, N. (2024). *From Data to Manuscript in R*. UChicago coursework.

Henderson, H. V., & Velleman, P. F. (1981). Building multiple regression models interactively. *Biometrics*, 391–411.

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. Retrieved from https://ggplot2.tidyverse.org

Wickham, H., & Grolemund, G. (2016). *R for data science: Import, tidy, transform, visualize, and model data*. O'Reilly Media, Inc.

Xie, Y., Allaire, J. J., & Grolemund, G. (2018). *R markdown: The definitive guide*. Chapman; Hall/CRC.

Table 1

Brands included, along with average number of cylinders and average weight among the number of models included for each brand.

brand	# Cylinders	Weight	# models
Datsun	4.0	2.3	1
Fiat	4.0	2.1	2
Honda	4.0	1.6	1
Lotus	4.0	1.5	1
Porsche	4.0	2.1	1
Toyota	4.0	2.1	2
Volvo	4.0	2.8	1
Ferrari	6.0	2.8	1
Mazda	6.0	2.8	2
Valiant	6.0	3.5	1
Merc	6.3	3.5	7
Hornet	7.0	3.3	2
AMC	8.0	3.4	1
Cadillac	8.0	5.2	1
Camaro	8.0	3.8	1
Chrysler	8.0	5.3	1
Dodge	8.0	3.5	1
Duster	8.0	3.6	1
Ford	8.0	3.2	1
Lincoln	8.0	5.4	1
Maserati	8.0	3.6	1
Pontiac	8.0	3.9	1