

Sean Hand

Bellarmino University

DS-160

03/25/25

Individual Project R-EDA

Conclusions

After diving into the mtcars dataset, a few interesting patterns stood out. For one, cars with 4 or 6 cylinders and manual transmissions tended to get better gas mileage, which kind of makes sense—they're lighter and generally less powerful. On the other hand, cars with 8 cylinders were heavier and had more horsepower, but that came at the cost of fuel efficiency. Another thing that caught my eye was how closely horsepower and weight were connected. Heavier cars need more power, but unfortunately, that usually means they burn more fuel. Manual cars, in general, performed a bit better when it came to MPG. Also, when I looked at the correlation heatmap, it was clear that MPG drops as weight, displacement, and horsepower go up. Cylinders also had a strong link with both horsepower and displacement, while gear and carburetor counts showed some moderate connections. It was really eye-opening to see how these variables play off each other.

Recommendations

Based on what I found, there are a few directions worth exploring further. First, it might be cool to build a basic prediction model—something like linear regression—to estimate MPG based on factors like horsepower, weight, and transmission type. Also, converting some of the variables, like transmission and engine type, into categories would make group comparisons easier down the line. Since this dataset is from the early '70s and only includes 32 cars, it would be awesome to compare it with something more recent and expansive. I also think trying out techniques like clustering or PCA could uncover some deeper trends or groups in the data. And if you're planning to do more of this kind of analysis, setting it up in R Markdown could save a lot of time by letting you generate full reports automatically.