James Velardi

Professor Nancy Binowski

Data Science Programming CMP-262

May 11th, 202515th, 2025

Final Project Summary

**Project Description**

This project investigates vehicle safety data from NHTSA in 2024. Data Science questions were asked and plotted in determining which cars in 2011-2024 are the safest for consumers. The tools used to do this were Pandas, Matplotlib’s Pyplot, Seaborn, Numpy, and regular expressions to filter and plot the data from the NHTSA CSV file.

**Questions Asked**

1. Which car manufacturers overall have the highest average safety ratings?
2. What is the distribution of vehicle types (Passenger Car, Multi-Passenger Car (SUV, Crossover), Vans, and Trucks) and their safety ratings?
3. Do cars with advanced safety features (Automatic Emergency Braking, Forward Collision Warning) help improve safety ratings?
4. How does the safety of vehicles compare overtime? Are newer vehicles getting much safer?
5. What do vehicles that rank the lowest in vehicle safety have the most in common? Could it be based on body type, frame type, lack of safety features and airbags, or does it all come down to brand?

**Recommendations**

Based on the findings in the data, I would encourage any potential client, like a rental car agency, to invest in Honda and Hyundai for economy cars and invest in Volvo and Lexus for their luxury lineup. As for which cars I would recommend avoiding altogether, Dodge Ram and Jeep, as they ranked a bit lower in this dataset.

**Future Work with This Dataset**

I would like to look at each individual model more to see which models rank the safest. I would also like to look at the rollover safety on some of the trucks and SUVs. Overall, there’s a lot of data in the CSV file and the various types of crash testing done, so taking a deep look at those numbers could also be interesting.