Topic Proposal for Group 4

The data set is from the official open data web site of the Canadian government. It depicts numerous details of how CO2 emission by vehicles can differ based on certain characteristics. The data was collected over a seven year period, and there are 7,385 vehicles in this data set. There are nine variables in the data set. Five of these variables are categorical, while the other four are numeric. The research question for this analysis is the following: what are the best indicators of a vehicle with high CO2 emissions? We hypothesize that the variables fuel consumption and engine size are the best indictors for high CO2 emissions, thus this theory will be tested in our analysis. High levels of CO2 emissions can produce extremely detrimental environmental and health consequences. Dealing with this occurrence is an extremely tricky and pressing issue around the globe. The end goal of our analysis is to provide noteworthy statistical findings to an automotive manufacturer in order for them to produce vehicles that do not emit high levels of CO2 emissions.

*Link to Data:*

https://open.canada.ca/data/en/dataset/98f1a129-f628-4ce4-b24d-6f16bf24dd64#wb-auto-6

*Link to GitHub:*

https://github.com/mtrobi2/Group-4