Analyzing Car CO2 Emissions

This project, aims to analyze vehicle CO2 emissions using multiple linear regression. Initially, a simple model was created using just city mileage, and later improved by incorporating the fuel type as a categorical variable.

Using pandas and statsmodels in Python, I preprocessed the dataset by encoding fuel types and adding a constant term.
The final model was trained using both predictors: City mileage and Fuel Type. The regression summary showed improved significance and fit.

The model was then used to predict emissions for four new cars from a test dataset. Dummy variables were aligned with the training set,
and predictions were successfully generated.

This project demonstrates key skills in data preprocessing, regression modeling, categorical encoding, and prediction on real-world datasets.