Scraping and Regression Project Proposal

Overview:

Estimating the right price for a used car is a long and exhausting process, so to find a solution for the issue this project started. The goal of the project is to create a model that predicts the prices of used cars based on several known data. This model can be useful for cars dealerships, people who are interested in cars that hold their value, first-time car owners who know nothing about car prices.

Problem Statement:

A new dealership that sells luxury cars is opening in town, and they want a model to predict the price of used cars. Automating the pricing process will help in getting a closer estimate for the actual prices in the market, and will help in getting reducing many employees' errors and expenses.

Data Description:

- Cars.com scraped data from this website because it is the second-largest automotive classified site and easy to collect data from.
- We will focus on the three biggest luxury manufacturers in Germany BMW, Mercedes-Benz, and Audi.
- Our target is predicting used cars price through multiple features such as:
    - Car ID or Name
    - Model
    - Make
    - Year
    - Selling price
    - Mileage
    - Fuel consumption
    - Transmission
    - Drive train
    - Engine size
    - Number of cylinders
  - Number of observations: ~10K

Tools: Python, and Jupyter Notebook.

Libraries: BeautifulSoup, Requests, Pandas, Scikit-Learn, and Seaborn.