**Project 4 Group One Proposal**

**Analyzing the Selling price of new cars**

* The “Car Price Analysis and Prediction” project involves delving into a dataset encompassing various attributes of used cars, ranging from price and make to fuel type (electric, hybrid, gasoline), color, and horsepower. Through data analysis, we aim to uncover the key factors influencing car prices. Moreover, predictive modeling will enable us to estimate the price of cars based on their attributes, empowering sellers, such as Otis, to make informed pricing decisions. The data could then be used by an automobile seller as a guide on how to target their customer base and maximize sales

**Implementation Steps**

* **Install and Import Modules:** Begin by installing essential Python libraries like Pandas, NumPy, Matplotlib, Seaborn, and Scipy. Then, import these modules into the Python environment.
* **Data Loading:** Load the dataset, which may be in .csv or .data format, using Pandas.
* **Data Cleaning:** Identifies and handles missing or null values in the dataset correctly.
* **Data Exploration:** Analyze the variable distributions, summary statistics, and dataset structure.
* **Feature Engineering:**To improve model performance, add new or adjust existing features.
* **Data Visualization:** To graphically analyze data relationships, various types of plots and charts can be employed.
* **Model Building:** Train machine learning models, such as regression models, to predict car prices based on available features.
* **Model Evaluation:** Evaluate trained models’ performance using appropriate measures such as mean absolute error or root mean squared error.
* **Prediction:** Make price predictions for new instances using the trained models

**Process**

* Find and clean the data, using pandas, create a few visualizations to ensure the data is clean and usable
* Use a SQL database to create and join tables, the goal being to have all the pertinent data contained in one or two tables
* Create visualizations in Tableau
* Create machine learning model, probably in Colab Jupyter Notebook to handle the predictive analysis that is the ultimate goal of this study.