**Final Project – Correlating Car Features To Predict MSRP**

**Team Members:**

Stas Beynarovich

Stacy Giauque

Step 1: Source data set with below 16 car features as the basis for predicting MSRP

1. Make
2. Model
3. Year
4. Engine Fuel Type
5. Engine Horse Power (HP)
6. Engine Cylinders
7. Transmission Type
8. Wheel Drive
9. Number of Doors
10. Market Category
11. Size Vehicle
12. Type of Vehicle
13. City MPG
14. Highway MGP
15. Popularity
16. MSRP

**Source:***Kaggle,*[https://www.kaggle.com/CooperUnion/cardataset](https://slack-redir.net/link?url=https%3A%2F%2Fwww.kaggle.com%2FCooperUnion%2Fcardataset)  
**Description:** Cars dataset with features including make, model, year, engine, and other properties of the car used to predict its price.  
**Format:** CSV file

**Number of Datapoints/Rows:** 11,915

Step 2: Using a combination of Python/Pandas, pgAdmin (SQL), Machine Learning (Scikit-Learn) we will create a model to predict vehicle MSRP based on make/model/specs.

* Will attempt to apply at least two different ML models in order to “train” and predict “best fit”

Step 3:  Design a website and visualizations using HTML/CSS/D3/Bootstrap, JavaScript and Tableau:

* Website: Multipage, interactive, user-driven summary of key car features and metrics
* Tableau: Create “gooey” visualizations of data used as input, feature correlation, and predictions by car make, model, year

Next steps for Thursday and Saturday classes:

* Familiarize and Review dataset
* Finalize ETL process
* Start ML fitting and training
* Strawman HTML/CSS code and Tableau on final visualizations