Project Proposal

Problem statement: How to decide the car price for next launch in 2022 by implementing popular features?

● Context: BMW, a German multinational corporation which produces luxury vehicles, plans to launch a brand new series in 2022. The market team and management team want to determine the MSRP by utilizing the right features to maximize the sales.

● Criteria for success: Determine the price for new car launch that give the best value by the end of Spring 2021

● Scope of solution space: Based on the market performance of existing BMW models and competitors' models to decide what features to include in the new launch and what would be a reasonable MSRP.

● Constraints: The data only includes models from 1990 to 2017. More recent data is not available.

The cost behind features is unknown. How much time and effort up-front affect the price as well.

● Stakeholders: Marketing Lead, Executive Director, Director of Development

● Data sources: CSV file scraped from Edmund and Twitter with features including make, model, year, engine and other properties of car.

The purpose of this data science project is to come up with a pricing model for BMW's next car model. BMW's management team tries to maximize the returns relative to its position in the market. The team does not have a strong sense of what car features matter most for the customers and bring the most profit. This project aims to build a predictive model for car price based on a number of car features for different car models in the market shares.

Exploring the data is the first step for the analysis. It is important to validate the dataset before using it for our prediction model. Then an exploratory data analysis shall be performed to identify what are the most correlated features with car price. Whether to isolate a portion of data or use the whole set of data for prediction? Grouping by market category or study only on BMW models might give some different idea on the car price. It is important to understand the brand value as well. When the data is ready, it is time to pick which modeling method works better. It would be a good idea to consider implementing Machine learning into the price prediction modeling. The final findings would be presented with a slide deck and a project report.