# Project Proposal.

# **Car Price Prediction**

#### **Project Title:**

# Car Price Prediction

#### **Purpose and Outcome:**

* **Purpose:** A Chinese automobile company Geely Auto aspires to enter the US market by setting up their manufacturing unit there and producing cars locally to give competition to their US and European counterparts.
* **Outcome:** To model the price of cars with the available independent variables. It will be used by the management to understand how exactly the prices vary with the independent variables. They can accordingly manipulate the design of the cars, the business strategy etc. to meet certain price levels.

#### **Dataset:**

* **Source:** Data set of different types of cars across the America market.
* **Description:** The dataset includes different types of cars, product information, and specifications.
* **Structure:**
  + **Car\_ID** Unique id of each observation (Interger)
  + **Symboling** Its assigned insurance risk rating, A value of +3 indicates that the auto is risky, -3 that it is probably pretty safe.(Categorical)
  + **carCompany** Name of car company (Categorical)
  + **fueltype** Car fuel type i.e gas or diesel (Categorical)
  + **aspiration** Aspiration used in a car (Categorical)
  + **doornumber** Number of doors in a car (Categorical)
  + **carbody** body of car (Categorical)
  + **drivewheel** type of drive wheel (Categorical)
  + **enginelocation** Location of car engine (Categorical)
  + **wheelbase** Weelbase of car (Numeric)
  + **carlength** Length of car (Numeric)
  + **carwidth** Width of car (Numeric)
  + **carheight** height of car (Numeric)
  + **curbweight** The weight of a car without occupants or baggage. (Numeric)
  + **enginetype** Type of engine. (Categorical)
  + **cylindernumber** cylinder placed in the car (Categorical)
  + **enginesize** Size of car (Numeric)
  + **fuelsystem** Fuel system of car (Categorical)
  + **boreratio**  Boreratio of car (Numeric)
  + **stroke**  Stroke or volume inside the engine (Numeric)
  + **compressionratio** compression ratio of car (Numeric)
  + **horsepower** Horsepower (Numeric)
  + **peakrpm** car peak rpm (Numeric)
  + **citympg**  Mileage in city (Numeric)
  + **highwaympg** Mileage on highway (Numeric)
  + **price(Dependent variable)** Price of car (Numeric)

#### **Initial Analysis Plan:**

* **Data Cleaning:** Handle missing values, correct data types, and remove duplicates.
* **EDA:** Generate summary statistics, visualize sales trends over time, and analyze seasonality and promotional impacts.
* **Analysis:** Build and evaluate forecasting models (e.g., ARIMA, Prophet), and compare model performances.
* **Visualization:** Create time series plots, seasonal decomposition plots, and forecast visualizations.
* **Data Storytelling:** Present the sales forecasts and provide actionable recommendations for inventory management and marketing strategies.