**Car Price Prediction**

**Problem statement:**

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.

They have contracted an automobile consulting company to understand the factors on which the pricing of cars depends. Specifically, they want to understand the factors affecting the pricing of cars in the American market, since those may be very different from the Chinese market. The company wants to know:

**Which variables are significant in predicting the price of a car?**

**How well those variables describe the price of a car?**

**Business Goal:**

You are required 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. Further, the model will be a good way for management to understand the pricing dynamics of a new market.

**About Dataset:**

The dataset contains **205** rows and **26** columns.

**1) Car\_ID:** Unique ID of each observation (Integer).

**2) Symboling:** Its assigned insurance risk rating, A value of +3 indicates that the auto is risky, -3 that it is probably pretty safe (Categorical).

**3) Car\_Company:** Name of Car company (Categorical). The following are the list of car companies available in the data set.

'alfa-romero', 'audi', 'bmw', 'chevrolet', 'dodge', 'honda', 'isuzu', 'jaguar', 'mazda', 'buick', 'mercury', 'mitsubishi', 'nissan', 'peugeot', 'plymouth', 'porsche', 'renault', 'saab', 'subaru', 'toyota', 'volkswagen', 'volvo'

**4) Fueltype:** Car fuel type i.e. gas or diesel (Categorical).

**5) Aspiration:** Aspiration used in a car i.e. Std or turbo (Categorical).Aspirated engine, is an internal combustion engine in which oxygen intake depends solely on atmospheric pressure and does not rely on forced induction through a turbocharger or a supercharger.

**6) DoorNumber:** Number of doors in a car i.e. 2 or 4 (Categorical).

**7) CarBody:** Body of a car sedan, hatchback, wagon, hardtop, convertible (Categorical).

**8) DriverWheel:** Type of driver wheel 4wd, fwd, rwd (Categorical).

**1)4wd:** Four-wheel drive, also called 4×4 or 4wd, refers to a two-axled vehicle drivetrain capable of providing torque to all of its wheels simultaneously.

**2)fwd:** Front wheel drive is called as fwd. A fwd can only supply power to the front wheels.

**3)rwd:** Rear wheel drive is called as rwd. A rwd can only supply power to the rear wheels.

**9) EngineLocation:** Location of car engine Front or Rear (Categorical). Out of 205 only 3 cars engine location is rear. All the 3 cars belong to Porsche company.

**10) WheelBase:** WheelBase is the distance between the front and rear wheels. Continuous from 86.6 to 120.9 (Float).

**11) CarLength:** Length of car continuous from 141.1 to 208.1 (Float).

**12) CarWidth:** Width of car continuous from 60.3 to 72.3 (Float).

**13) CarHeight:** Height of car continuous from 47.8 to 59.8 (Float).

**14) CurbWeight:** The weight of car continuous from 1488 to 4066 (Int).

**15) EngineType:** Type of engine dohc, ohcv, ohc, l, rotor, ohcf, dohcv (Categorical).

**16) CylinderNumber:** Number of cylinders continuous from 2 to 8 (Categorical). A cylinder is the power unit of an engine; it's the chamber where the gasoline is burned and turned into power.

**17) EngineSize:** The Size of engine continuous from 61 to 326 (Int).

**18) FuelSystem:** Fuel system is to inject a precise amount of automized and pressurized fuel in to each cylinder at a proper time. Types of fuel system are mpfi, 2bbl, mfi, 1bbl, spfi, 4bbl, idi, spdi (Categorical).

**19) BoreRatio:** Bore-Stroke Ratio is the ratio between the dimensions of the engine cylinder bore diameter to its piston stroke-length. Bore ratio continuous from 2.54 to 3.94 (Float).

**20) Stroke:** stroke or volume inside the engine continuous from 2.07 to 4.17 (Float).

**21) CompressionRatio:** Compression ratio, in an internal-combustion engine, degree to which the fuel mixture is compressed before ignition. Compression ratio continuous from 7.0 to 23.0 (Float).

**22) HorsePower:** Horsepower is a unit of power used to measure the forcefulness of a vehicle's engine. Horse power continuous from 48 to 288 (Int).

**23) PeakRPM:** RPM stands for revolutions per minute, and it's used as a measure of how fast any machine is operating at a given time. how many times each piston goes up and down in its cylinder. PeakRPM continuous from 4150 to 6600 (Int).

**24) CityMPG:** CityMPG refers to driving with occasional stopping and braking. Mileage in city continuous from 13 to 49 (Float).

**25) HighwayMPG**: It is based on continuous acceleration. Mileage on highway continuous from 16 to 54 (Float).

**26) Price**: Price of car continuous from 5118 to 45400 (Float).