**Project: Used Car Price Prediction using Vehicle Dataset**

### **📌 Problem Statement:**

In the thriving used car market, determining a fair resale price is essential. This assignment challenges you to predict the **selling price of used cars** using beloved features like manufacturing year, kilometers driven, fuel type, transmission, seller type, and ownership history. The goal is to build and evaluate regression models that provide reliable estimates—helping buyers and sellers navigate the market confidently.

### **📋 Guidelines:**

1. **Dataset Familiarization**
   * Review all columns: name, year, selling\_price, km\_driven, fuel, seller\_type, transmission, owner.
   * Consider how each feature logically affects the resale value (e.g., age, wear, fuel type).
2. **Data Cleaning & Preprocessing**
   * Address missing entries or duplicate rows.
   * Resolve inconsistencies (like unexpected category labels).
   * Prepare categorical features for modeling through encoding.
3. **Exploratory Data Analysis (EDA)**
   * Visualize price distribution to identify outliers.
   * Explore relationships such as selling\_price vs. year and selling\_price vs. km\_driven.
   * Analyze categorical impacts (e.g., fuel type, transmission) on price.
4. **Feature Engineering**
   * Introduce car\_age by subtracting year from the current year.
   * Encode categorical data (e.g., one-hot, ordinal).
   * Scale features if necessary to improve regression stability.
5. **Model Development**
   * Partition into training and test subsets.
   * Build and compare Regression Models
   * Optimize using cross-validation or hyperparameter tuning.
6. **Model Evaluation**
   * Compare model accuracy using:  
     + **R² Score**
     + **Mean Absolute Error (MAE)**
     + **Root Mean Squared Error (RMSE)**
   * Visualize predicted vs. actual values to assess fit.
7. **Insights & Recommendations**
   * Report the most accurate model and explain why it performs best.
   * Highlight which features (e.g., age, mileage, fuel type) have the most influence.
   * Offer insights and suggestions for market participants.

### **🪜 Step-by-Step Approach:**

1. Obtain and load the dataset.
2. Thoroughly inspect and clean your data.
3. Conduct EDA to uncover data patterns.
4. Engineer, encode, and scale features.
5. Train several regression models and evaluate them using robust metrics.
6. Summarize comparisons and derive practical insights.

### **🔗 Link to Dataset:**

📂 **Vehicle Dataset** [**https://www.kaggle.com/datasets/nehalbirla/vehicle-dataset-from-cardekho**](https://www.kaggle.com/datasets/nehalbirla/vehicle-dataset-from-cardekho)