Project Overview

Data-Driven Car Price Prediction and Analysis

This project focuses on building an **end-to-end data analytics and machine learning solution** to analyze factors affecting car prices and predict the **market value of cars** using historical data.
The project integrates **data cleaning, exploratory data analysis (EDA), SQL querying, machine learning modeling, and interactive visualization** to deliver actionable insights for buyers, sellers, and dealerships.

🔹 **Objectives of the Project**

- Analyze key factors influencing car prices
- Perform structured analysis using **SQL**
- Build a **machine learning model** to predict car prices
- Create an **interactive Power BI dashboard** for business insights
- Enable **data-driven decision making** in the automotive domain

---

🔹 **End-to-End Workflow**

1️⃣ **Data Collection**

- Dataset contains historical car listings with attributes such as:
    - Brand & Model
    - Year of Manufacture
    - Mileage
    - Fuel Type
    - Transmission
    - Engine Capacity
    - Selling Price

2️⃣ **Data Cleaning & Preprocessing**

- Removed duplicate records
- Handled missing values
- Standardized categorical values
- Converted data types for analysis

- Feature engineering (car age, price per km, etc.)

3️⃣ Exploratory Data Analysis (EDA)

- Price distribution analysis
- Relationship between:
    - Price vs Year
    - Price vs Mileage
    - Price vs Fuel Type
- Brand-wise and model-wise comparisons
- Outlier detection

4️⃣ SQL Analysis

- Created structured tables
- Wrote queries to:
    - Count total cars
    - Find average, min, max prices
    - Identify most popular brands
    - Analyze fuel type and transmission trends
    - Rank cars by price and mileage

5️⃣ Machine Learning Model

- Algorithms used (example):
    - Linear Regression
    - Random Forest Regressor
- Train-test split
- Model evaluation using:
    - R^2 Score
    - Mean Absolute Error (MAE)
    - Mean Squared Error (MSE)
- Optimized model for better prediction accuracy

6️⃣ Power BI Dashboard

- Interactive dashboard with:
    - Total Cars Overview
    - Average Price KPI
    - Brand-wise price comparison
    - Fuel & transmission analysis
    - Price trends over years

🔹 Key Dataset Insights (Example Format)

- 🚙 **Total Cars:** 2501
- 🏷️ **Total Brands:** 7
- 🗺️ **Total Models:** 28
- 💰 **Average Car Price:** ₹ 52638.02
- 📉 **Minimum Price:** ₹ 5011
- 📈 **Maximum Price:** ₹ 99982.59