CAR DEKHO

DATASET ANALYIS

INTRODUCTION

This report presents a comprehensive analysis of data sourced from the Car Dekho app. The primary objective of this analysis is to uncover key insights, trends and patterns within the dataset.

OBJECTIVE

1. Data collection

By acquiring a well-structured dataset. This dataset included details such as car name, year, selling price, km driven, fuel, seller type, transmission, owner, mileage, engine, max power, torque, seats. The data was stored in a relational database.

2. Data cleaning and preprocessing

Performed data cleaning tasks to handle missing values, duplicate records, and outliers, ensuring the dataset's integrity.

3. SQL queries

Designed and executed SQL queries to extract relevant information from the database. This involved a range of SQL operations, including SELECT, JOIN, GROUP BY, and aggregation functions.

DATA OVERVIEW

FIELDS	DATA TYPE
Company	Varchar 50
Year	Int
Selling price	Int
Km driven	Int
Fuel	Text
Seller type	Text
Transmission	Text
Owner	Text
Mileage	Int

Engine	Int
Max power	Int
Torque	Int
Seats	Int
Name	Varchar 200
Km driven set	Varchar 50
Mileage set	Varchar 50
Torque avg	Varchar 50
Current date	Date
Max power cat	Varchar 50

QUERIES AND CONCLUSIONS

- select company,count(*) from car_table group by company order by count(*)desc

  Mauti has the most cars for reselling with 2093 cars.
  Opel,Ashok and Lexus has the least cars with 1 each.

- select company,avg(max_power) from car_table group by company order by avg(max_power)desc;

  Isuzu has the highest average of max_power.
  Daewoo has the lowest average of max_power.

- select mileage_set,avg(selling_price) from car_table group by mileage_set;

medium	561172.9003
high	487194.3685
low	1822000.0000

  When the mileage is lower, the average price is higher.

- select max_power_cat,avg(selling_price) from car_table group by max_power_cat order by avg(selling_price);

low power	243716.9629
medium power	528018.7118
high power	877762.4840

  Higher the power, higher the average selling price.
  Lower the power, lower the selling price

- select max_power_cat,avg(mileage) from car_table group by max_power_cat order by avg(mileage);

high power	13.4890
medium power	19.8743
low power	20.6367

Higher the power, lower the mileage.
Lower the power, higher the mileage.

- select max_power_cat,avg(selling_price),avg(mileage) from car_table group by max_power_cat order by avg(selling_price);

  Vehicle with higher power will have higher price but lower mileage.
  Vehicle with lower power will have lower price but higher mileage.

- select km_driven_cat,avg(selling_price) from car_table group by km_driven_cat;

  Vehicle with higher km driven has lower price.
  Vehicle with lower km driven has higher price.

- select km_driven_cat,count(*) from car_table group by km_driven_cat;

10k to 100k	4704
low miles	219
above 100k	1799

Vehicle with higher km driven are for sale more than vehicles with lower km driven.

- select torque_avg,avg(max_power) from car_table group by torque_avg;

low	1107.6414
high	2276.9458
medium	1460.2058

Higher the torque, higher the average of max power.

- select torque_avg,avg(selling_price) from car_table group by torque_avg;

Higher the torque, higher the selling price.

- select seats,count(*) from car_table group by seats;

Most vehicles for sale is 5seater.

- select seats,count(*),avg(selling_price) from car_table group by seats order by avg(selling_price)desc limit 1;

  7 seater vehicles has the highest avg selling price.

- select transmission,avg(selling_price) from car_table group by transmission;
  automatic vehicles are more costlier than manual vehicles.