DATASET ANALYSIS

Asna E S

INTRODUCTION

- This report provides a straightforward analysis of a used car dataset.The data includes information about various car models,their selling prices,mileage,year,torque,power and other key features.The purpose of this analysis is to gain insights into the used cars listed for sale.The goal is to identify basic trends and insights that can guide decision-making for both sellers and buyers.

DATA CLEANING PROCESS

- Before performing the analysis,the dataset underwent a data cleaning process to correct any inaccuracies and prepare the data for analysis:

- The dataset was examined for outliers .Any values that seemed unrealistic were flagged for further investigation or it was removed from the dataset.

  eg:extremely high numbers like over 1.5 million km were removed.

- If an outlier was confirmed to be a data entry error,it was removed from the dataset.

  for instance ,the "km_driven" value was 2,360,457 km,it considered as an error and removed.

- Fields like "year","km_driven", "mileage","engine" etc…were standardized to ensure consistency in data types.

- Duplicates records:the dataset was checked for duplicate records and removed the duplicates

- Initially, the dataset contained 7927 rows,after data cleaning it was reduced to 6719.

DATA ANALYSIS

- The dataset includes cars from 31 different companies,with Maruti being the most common.

- The cars range from budget friendly option like Maruti 800 AC to luxury model like volvo XC90 T8 Excellence BSTV,BME etc...

➢ Year of manufacture:range from 1994-2023

   most of the cars were manufactured in 2017

➢ Selling price:

   Range:prices vary from 29999 for the Maruti 800 AC to 10000000 for the Volvo XC90 T8 Excellence BSTV

- most of the cars are priced between 29999 and 600000.

➢ Km driven: most cars have been driven above 100k km.
   Highest km driven car is Maruti wagen R LXI minor (577414 km)
   Average km driven 72700 km

➢ Fuel type: (Diesel,petrol,CNG,LPG,Electic)
    Diesel and petrol cars are the most common.
     there is only one Electric car – MG ZS EV Exclusive
       Diesel: 3657 cars
       Petrol: 2976 cars
       CNG: 50 cars
       LPG: 35 cars
       Electic: 1 car

➢ Seller type:most of the seller type were individual.
      Individual:6024
      Dealer: 668
      Trust mark dealer: 27

➢ Transmission:Manual Vs Automatic
     most of the cars are manuel
      count of Manual= 6142
      count of Automatic=577

➢ Mileage & Engine
   maximum mileage = 42 (Volvo XC90 TB Excellence BSIV)
   minimum mileage = 9
   average mileage= 19
   most of the cars offer average mileage.
   Engine capacities range from 624 cc to 3604 cc with power from 32 to400.

➢ Torque:
   maximum=789 (Maruti Zen D)
   minimum=4 (Tata Nano LX)
   average=160

➢ Seat:
   most of the cars have 5 seats.
   there is only one car with 14 setas.(Tata Winger Deluxe)

➢ Owner
   most of the owners were first owners,there is only 5 test drive car
   First owner: 4182 cars
   Second owner:1887 cars
   Third owner:492 cars
   four & above owner:153 cars
   Test drive car:5 cars

CONCLUSION

- The dataset suggests a diverse used car market with options for all types of buyers from economical to luxury preferences.

- Buyers should prioritize low – kilometer models,especially those from that latest manufacturing years to ensure value retention and minimize future maintenance cost.

- Power exhibits a positive correlation with the selling price.this suggest that higher powered cars are more desirable and thus,more expensive