**Data Analysis Project Report**

**Project Title:** Cars-24 Data Analysis of Used Car Dataset  
**Author:** Anshuman Dutta  
**Date:** 23 October 2025

**1. Introduction**

This project focuses on analyzing a dataset of **used cars** to extract insights on sales patterns, car preferences, fuel types, locations, and pricing trends. The main objective is to perform **data cleaning, visualization, and analysis** to support data-driven decisions.

**Dataset Source:** Kaggle / Provided CSV file

**Key Features in Dataset:**

* Car Name
* Year
* Fuel Type
* Location
* Distance
* Price
* Type

**2. Objective**

The goals of this project are:

1. Understand the dataset and its structure.
2. Identify and handle missing or inconsistent data.
3. Explore key patterns in car sales, pricing, and locations.
4. Visualize trends for better insights.
5. Highlight actionable insights for stakeholders.

**3. Data Cleaning**

* **Handling Missing Values:**
  + Missing Location values filled with "Punjab".
  + Missing Fuel or Car Name values handled appropriately.
* **Data Formatting:**
  + Standardized column names and sorting them.
  + Converted numerical columns (e.g., Price, Distance) to proper types.
* **Removing Duplicates:**
  + Duplicate rows were identified and removed to maintain data integrity.

**4. Exploratory Data Analysis (EDA)**

**4.1 Distribution of Cars by Fuel Type**

Fuel\_type = df["Fuel"].value\_counts()

Fuel\_type.plot(kind='pie', title='Cars by Fuel Type')

**Insight:** Petrol cars are the most common, followed by diesel and electric vehicles.

**4.2 Top Locations with Maximum Sales**

* Aggregated sales by Location.
* Visualized using bar charts or pie charts for clarity.

**Insight:** Certain cities show higher car sales — key for targeted marketing or inventory planning.

**4.3 Car Price vs Year**

* Scatter plots or line charts used to visualize price trends over years.

**Insight:** Older cars generally have lower prices; newer models command a premium.

**4.4 Top Car Models**

* Identified the **most sold car models** using frequency counts.

**Insight:** Popular models indicate customer preferences and market trends.

**5. Tools & Libraries Used**

* **Python**: Pandas, Matplotlib, Seaborn
* **Jupyter Notebook** for analysis
* **Git & GitHub** for version control

**6. Key Insights**

1. Petrol cars dominate the used car market in terms of quantity.
2. Some locations have significantly higher sales than others, indicating market hotspots.
3. Car prices decrease with age and distance covered, as expected.
4. Certain car models are preferred by users, which can guide inventory decisions.

**7. Conclusion**

This analysis provides a **comprehensive understanding** of the used car market available in Cars-24. Insights from the data can help **dealers, sellers, and marketers** make informed decisions regarding stock, pricing, and promotions.

**Next Steps:**

* Predictive modeling for car prices.
* Dashboard creation for real-time insights.
* Further analysis of customer preferences by location.