Topic: Car Price vs. different factors (How do factors impact the car pricing in the market?)
Dataset link: https://www.kaggle.com/datasets/asinow/car-price-dataset?resource=download

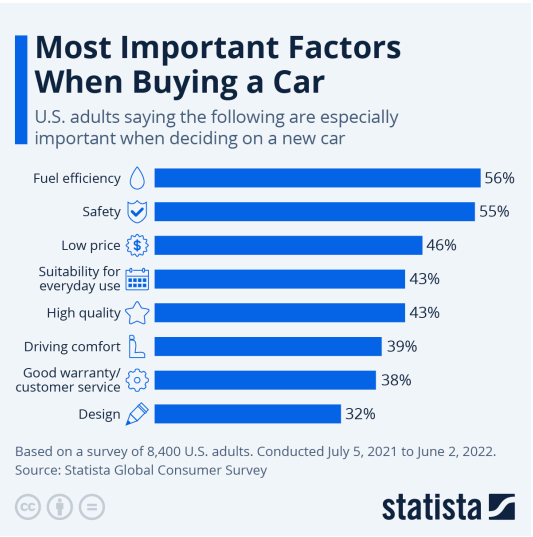
Title: Unveiling Car Price Determinants

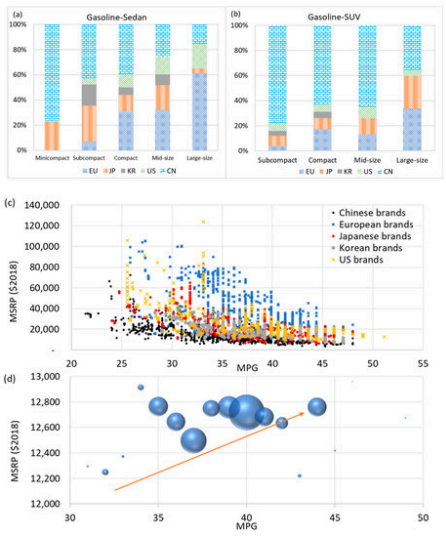
Introduction:
Presently, numerous people need cars to get around without an accessible transportation system nearby. Car pricing plays a vital role within the automotive industry and consumers worldwide. Numerous factors, including production costs, brand reputation, market demand, technological aspects, economic conditions, and profit margin, determine a car's price. Automakers must carefully balance affordability, profitability, and competitiveness when setting car prices. At the same time, consumers navigate pricing strategies such as discounts, financing options, and depreciation rates to make informed purchasing decisions. Car pricing is important because it affects accessibility and affordability to vehicles, influencing how individuals and families choose vehicles to meet their needs and budgets. Additionally, car pricing trends can provide insights into shifts in consumer preferences, innovation, environmental considerations, and global market fluctuations, which makes this dataset an intriguing subject for analysis and discussion.

Two reference paper

- https://www.emerald.com/insight/content/doi/10.1108/ijrdm-02-2012-0017/full/html
- https://www.sciencedirect.com/science/article/abs/pii/S0965856414002134

Two static images from online or other sources that may be related to your topic.





An introduction to your data, including the size and source.
We selected an existing dataset from Kaggle that contains 10,000 observations and 10 features. The dataset includes both categorical and numerical variables:

- Categorical Features: Brand, Model, Fuel_Type, Transmission
- Numerical Features: Year, Engine_Size, Mileage, Doors, Owner_Count, Price

This dataset can be utilized for studying car market price prediction. There are key numerical variables, such as Engine Size and Mileage, that can help estimate vehicle prices or cost. Additionally, some categorical features like Fuel Type and Transmission provide insights into customer preferences and market trends.

A plan for the later analysis, including any potential pre-processing to the data

Data Cleaning:

1. Remove duplicate entries (if any).
2. Handle missing values by either imputing or removing records.
3. Standardize units (e.g., mileage in kilometers vs. miles).
4. Convert categorical variables (e.g., Fuel_type, Transmission) into numerical representations for analysis.

Feature Engineering:

1. Create a depreciation rate feature: Calculate depreciation based on the difference between the original price and the listed resale price.
2. Generate a fuel efficiency cost metric: Estimate fuel costs based on Mileage and Fuel_type.
3. Add a brand reliability score: Integrate an external source or use existing brand reputation rankings.

Exploratory Data Analysis (EDA):

1. Examine distributions of key variables such as Price, Mileage, and Year.
2. Identify correlations between features (e.g., how does Mileage affect price?).
3. Detect and remove potential outliers that could skew analysis.

Data Visualization:

1. Static Visualization #1: Price Distribution by Car Brand to Show price variations across different brands.
2. Static Visualization #2: Mileage vs. Price Scatter Plot to Identify the impact of mileage on price
3. Interactive Visualization #3: Yearly Depreciation Rate by Brand
4. Interactive Visualization #4: Filterable Dashboard for Used Car Pricing Trends to Allow users to filter by brand, fuel type, year, and mileage to see trends

Describe each group member's duties.

Each member contributed equally by doing two parts of the proposal, and we will assign fairly for later stages of this project.