Electric Vehicles Market Segment Analysis

AUTHOR

Akhilesh Maurya

Problem Statement

The EV startup aims to determine which segments of the Indian market are most suitable for launching their electric vehicles. The startup needs to understand various customer and vehicle attributes to develop a feasible market entry strategy. The analysis focuses on identifying key market segments and understanding demographic, psychographic, behavioral, and other factors influencing the adoption of electric vehicles.

Overview

This report presents an in-depth analysis of the Electric Vehicle (EV) market in India. The goal is to identify potential customer and vehicle segments that an EV startup can target to establish a strong foothold in the market. The analysis leverages two datasets: the CarDekho dataset and an additional dataset on electric vehicles. The methodologies used include Principal Component Analysis (PCA) and KMeans clustering to derive actionable insights.

The Dataset two datasets Overview:

1. **Cardekho Dataset**: Contains data on various cars, including vehicle age, km driven, fuel type, transmission type, mileage, engine, max power, seats, and selling price.

	Unnamed: 0	car_name	brand	model	vehicle_age	km_driven	seller_type	fuel_type	transmission_type	mileage	engine	max_power	seats	selling_price
0	0	Maruti Alto	Maruti	Alto	9	120000	Individual	Petrol	Manual	19.70	796	46.30	5	120000
1	1	Hyundai Grand	Hyundai	Grand	5	20000	Individual	Petrol	Manual	18.90	1197	82.00	5	550000
2	2	Hyundai i20	Hyundai	i20	11	60000	Individual	Petrol	Manual	17.00	1197	80.00	5	215000
3	3	Maruti Alto	Maruti	Alto	9	37000	Individual	Petrol	Manual	20.92	998	67.10	5	226000
4	4	Ford Ecosport	Ford	Ecosport	6	30000	Dealer	Diesel	Manual	22.77	1498	98.59	5	570000

2. **Electric Car Data**: Contains data on electric cars, including car name, brand, acceleration, top speed, range, efficiency, fast charge speed, price in Germany, drive, power, number of seats, and type.

	Brand	Model	Accel	TopSpeed	Range	Efficiency	FastCharge	RapidCharge	PowerTrain	PlugType	BodyStyle	Segment	Seats	PriceEuro
0	Tesla	Model 3 Long Range Dual Motor	4.6 sec	233 km/h	450 km	161 Wh/km	940 km/h	Rapid charging possible	All Wheel Drive	Type 2 CCS	Sedan	D	5	55480
1	Volkswagen	ID.3 Pure	10.0 sec	160 km/h	270 km	167 Wh/km	250 km/h	Rapid charging possible	Rear Wheel Drive	Type 2 CCS	Hatchback	C	5	30000
2	Polestar	2	4.7 sec	210 km/h	400 km	181 Wh/km	620 km/h	Rapid charging possible	All Wheel Drive	Type 2 CCS	Liftback	D	5	56440
3	BMW	iX3	6.8 sec	180 km/h	360 km	206 Wh/km	560 km/h	Rapid charging possible	Rear Wheel Drive	Type 2 CCS	SUV	D	5	68040
4	Honda	e	9.5 sec	145 km/h	170 km	168 Wh/km	190 km/h	Rapid charging possible	Rear Wheel Drive	Type 2 CCS	Hatchback	B	4	32997

The goal is to use these datasets to identify market segments most likely to adopt EVs in India and provide strategic recommendations.

Tools and Libraries Used

- **Python**: Programming language used for data analysis.
- **Pandas**: Library for data manipulation and analysis.
- **NumPy**: Library for numerical operations.
- **Matplotlib**: Library for data visualization.
- **Seaborn**: Library for statistical data visualization.
- **Scikit-learn**: Machine learning library for clustering and PCA.

Analysis Steps

1. Data Collection and Preprocessing:

Load datasets

```
Df_ cardekho = pd.read_csv('/mnt/data/cardekho_dataset.csv')
```

```
Df_ electric = pd.read_csv('/mnt/data/ElectricCarData_Norm.csv')
```
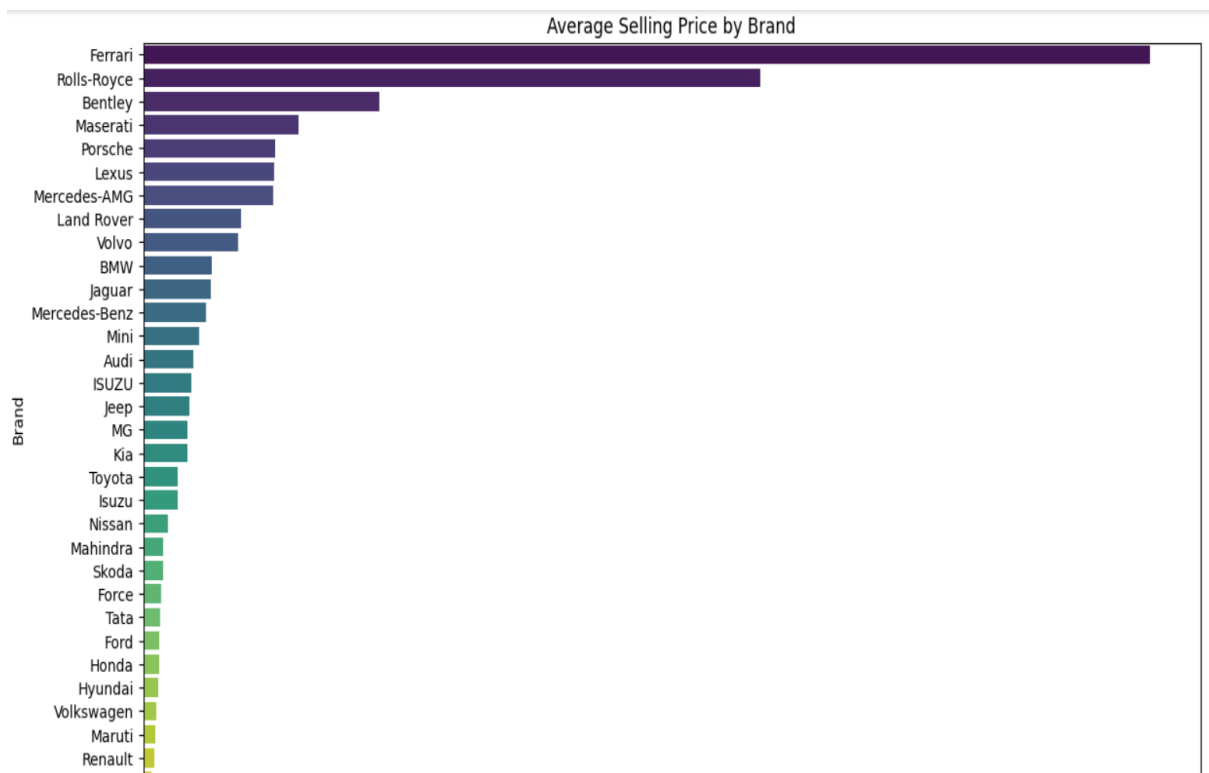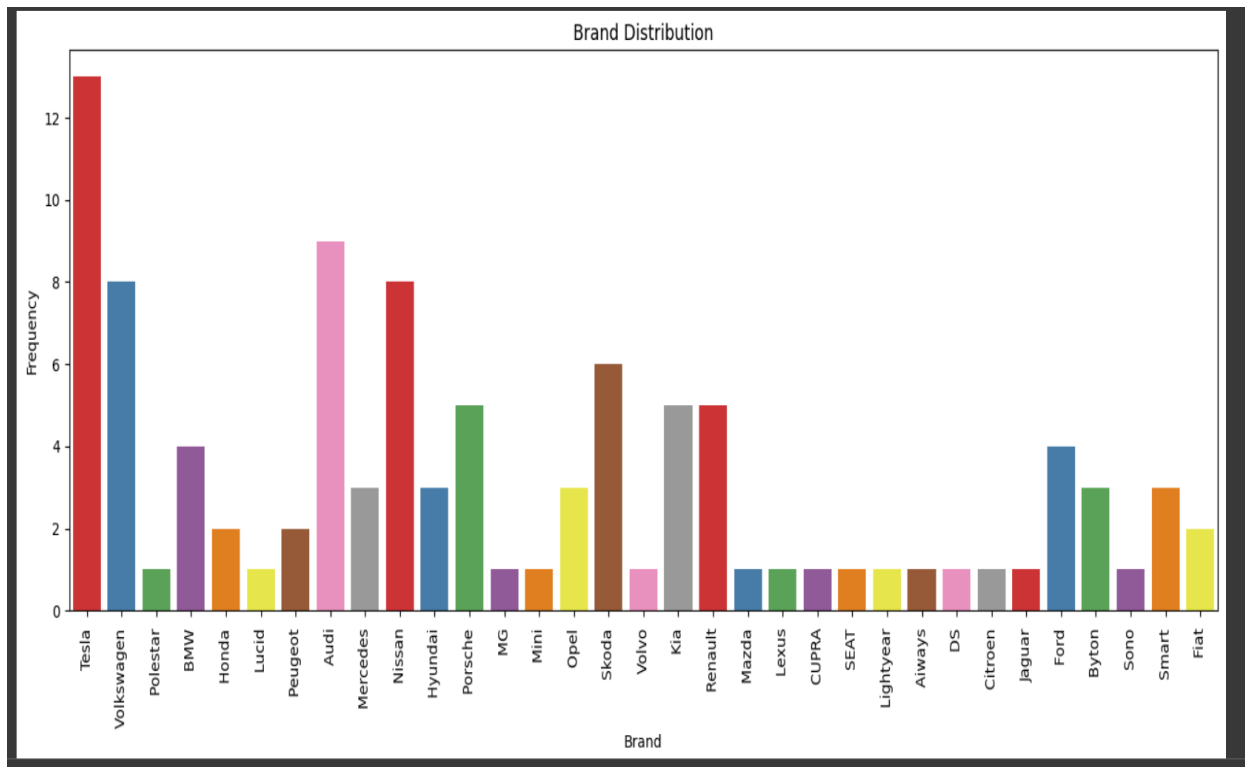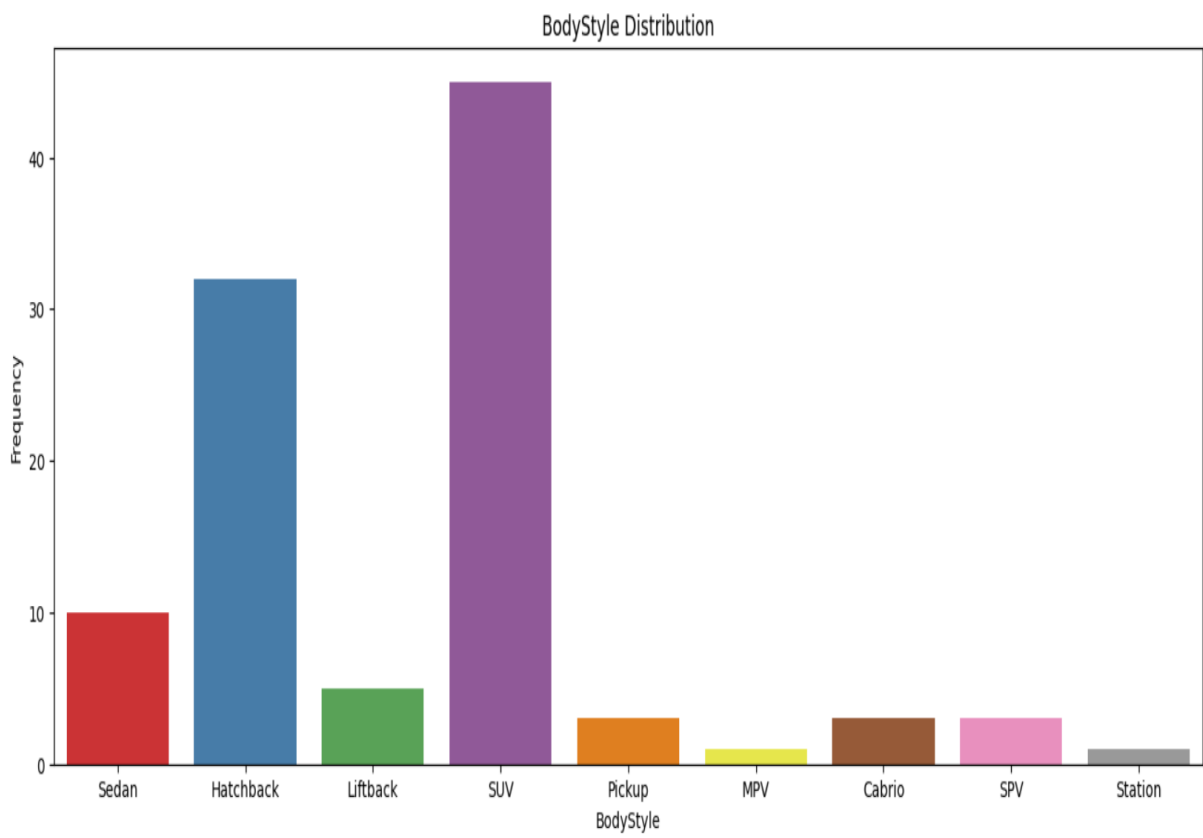
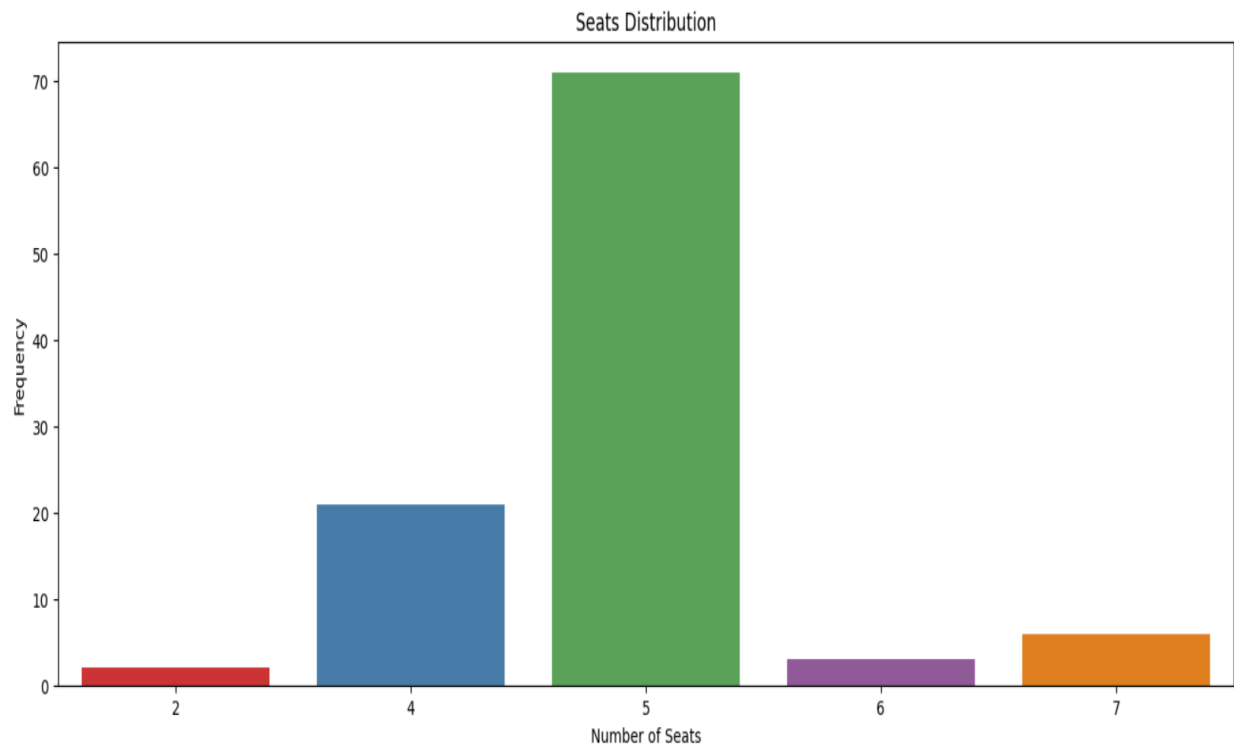# Display first few rows of each dataset

```
print(df_cardekho.head())
print(df_electric.head())
```
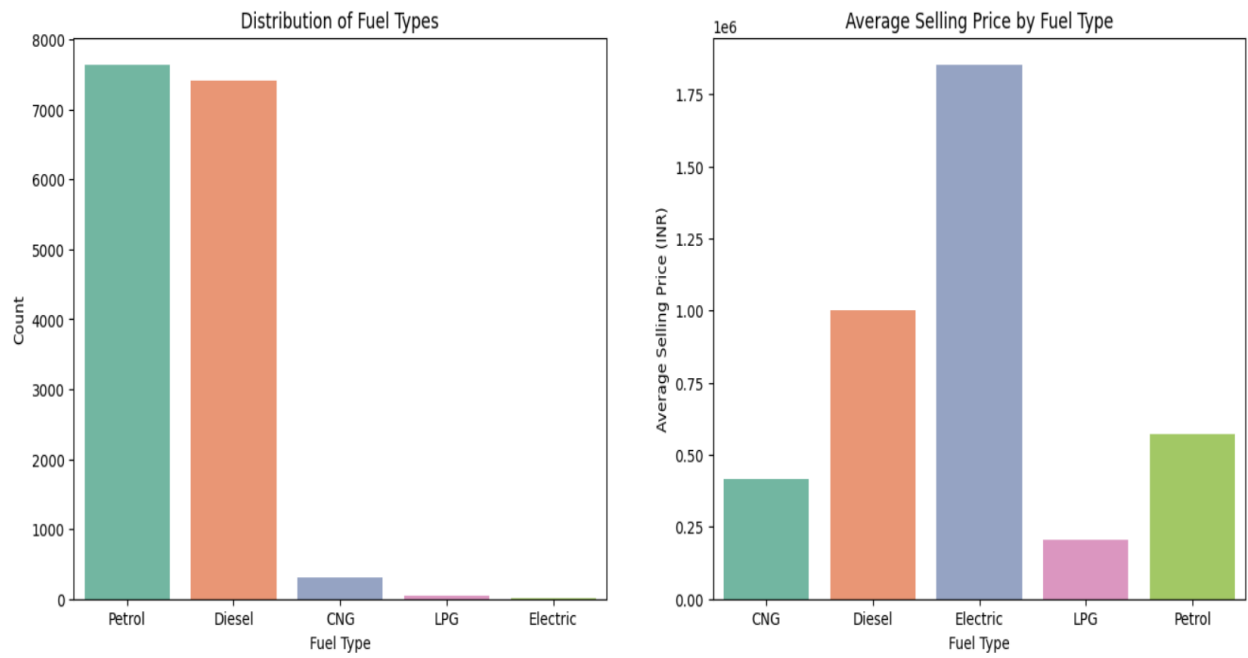
2. Data Cleaning and Handling Missing Values

```
df.drop(columns=['Unnamed: 0'], inplace=True)
df['FastCharge'] = df['FastCharge'].replace('-', np.nan)
df['FastCharge'] = df['FastCharge'].fillna(df['FastCharge'].mode())
df['Accel'] = df['Accel'].str.replace(' sec', '').astype(float)
df['TopSpeed'] = df['TopSpeed'].str.replace(' km/h', '').astype(float)
df['Range'] = df['Range'].str.replace(' km', '').astype(float)
df['Efficiency'] = df['Efficiency'].str.replace(' Wh/km', '').astype(float)
df['FastCharge'] = df['FastCharge'].str.replace(' km/h', '').astype(float)
df['Price'] = df['PriceEuro'] * 90
df.drop(['PriceEuro'], axis=1, inplace=True)
df['FastCharge'] = df['FastCharge'].fillna(df['FastCharge'].mean())
df[df['FastCharge'].str.contains('-')]
```
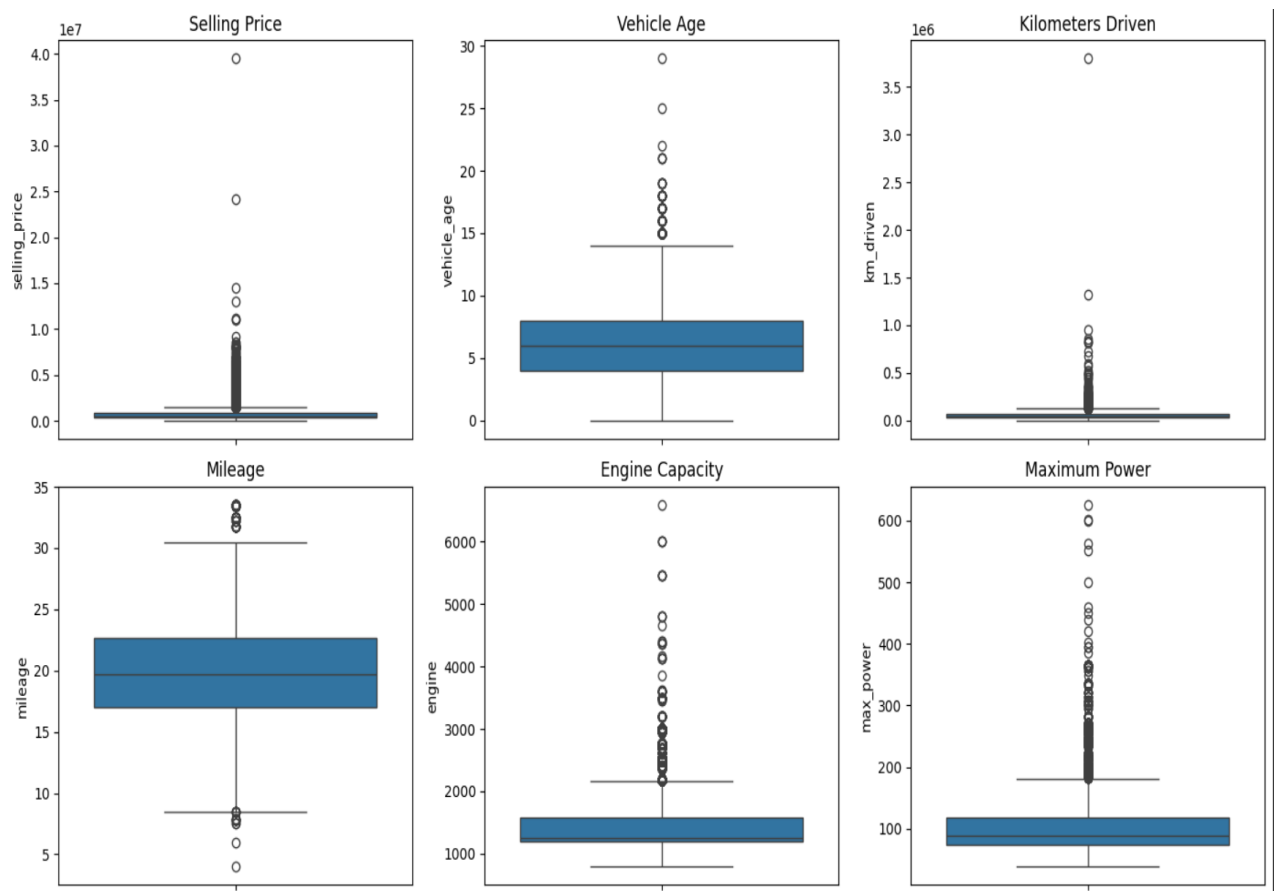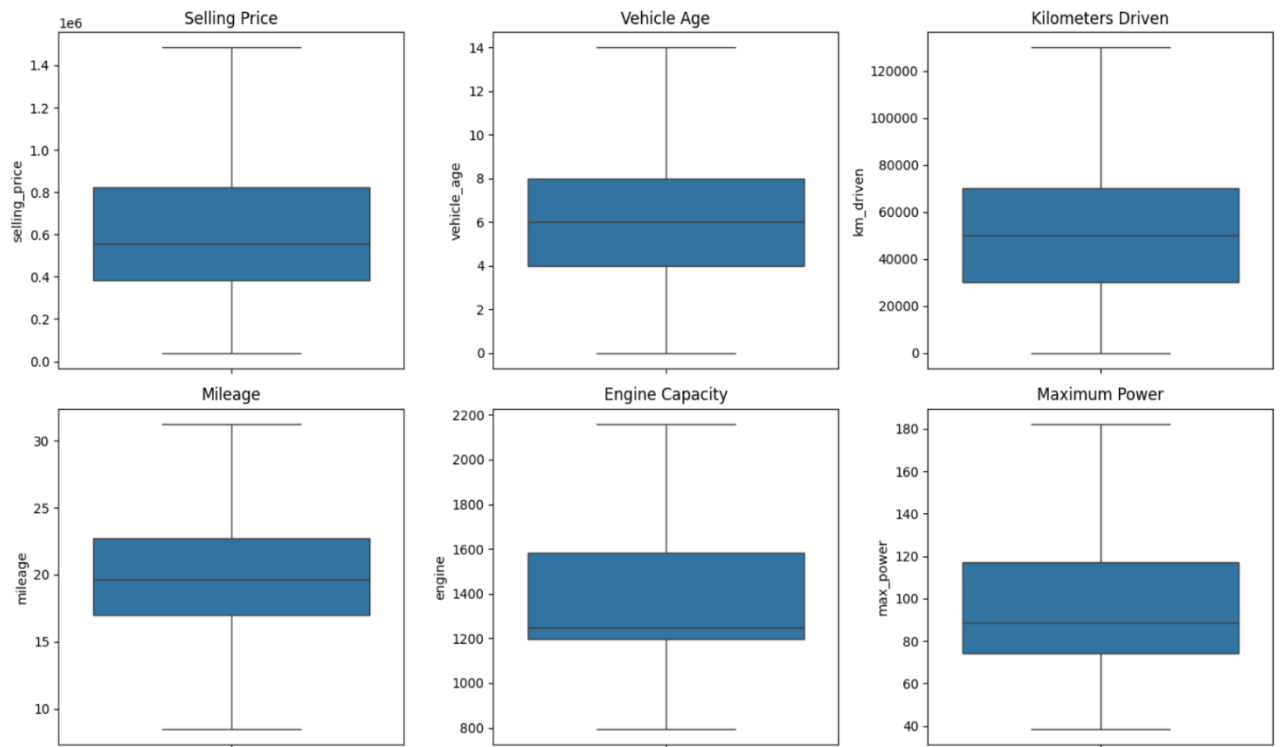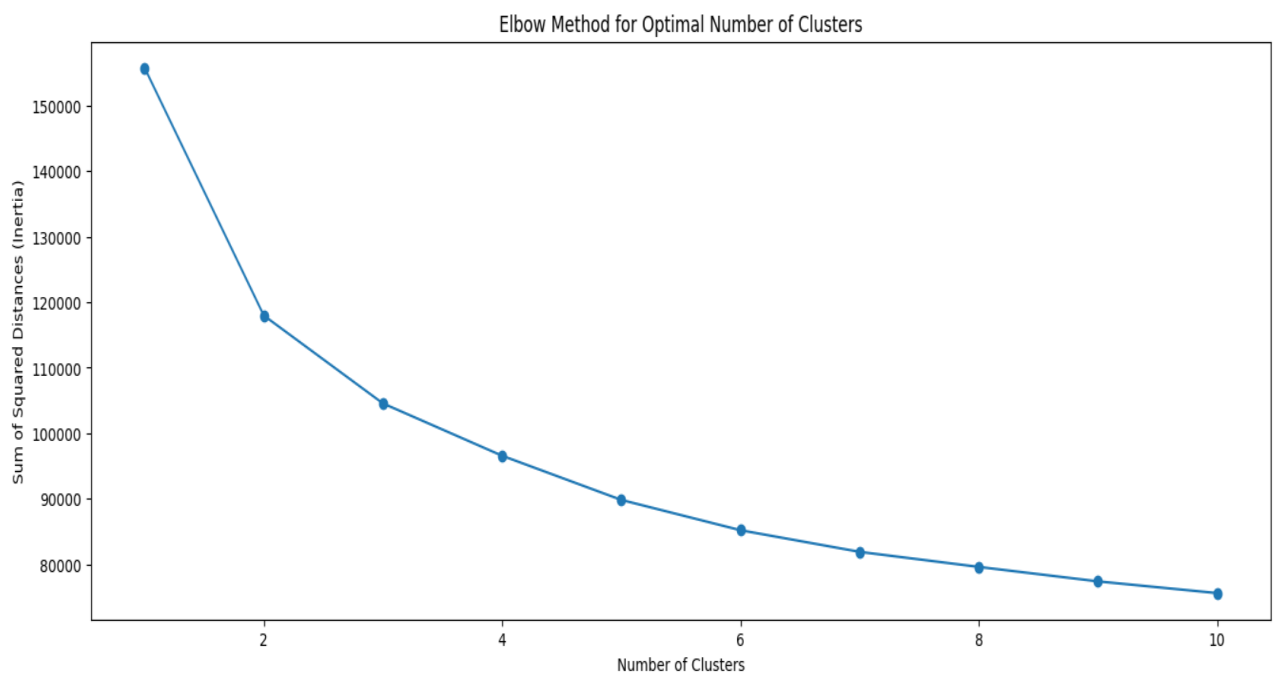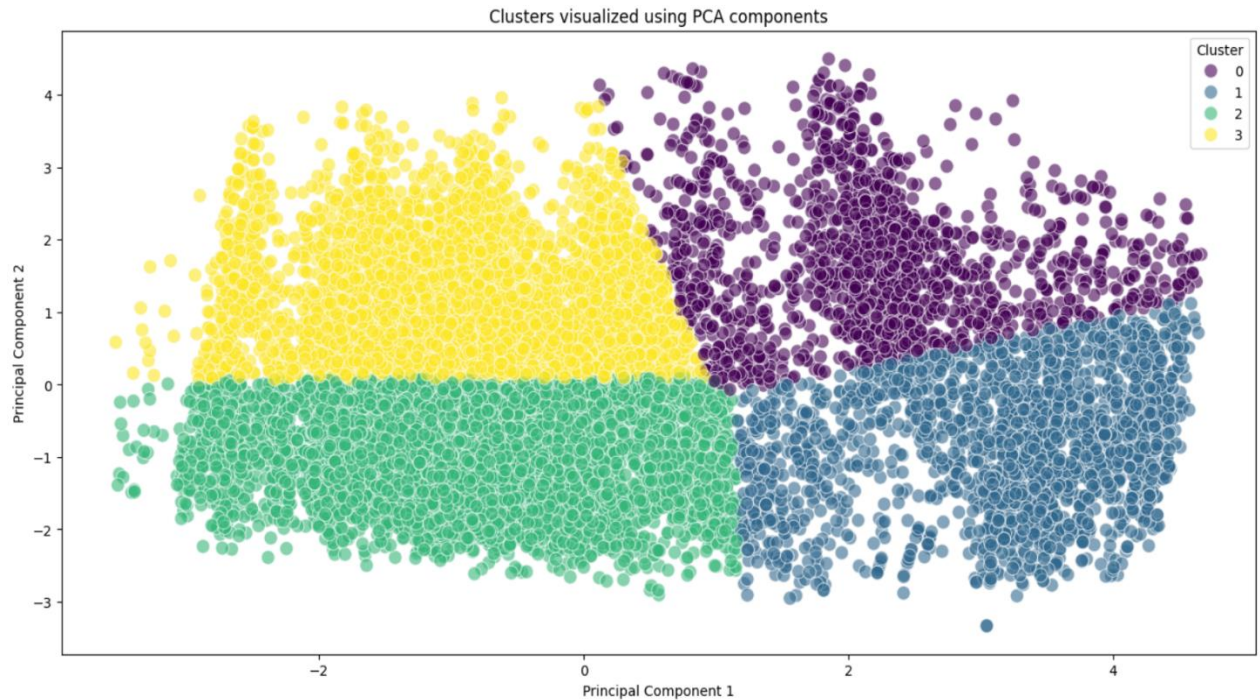
3. EDA (Exploratory Data Analysis)

Brand Distribution

Average Selling Price by Brand

Seats Distribution

BodyStyle Distribution

Distribution of Fuel Types — Average Selling Price by Fuel Type

4. Outliers

5. Remove Outliers

K-Means Clustering

Principal Component Analysis (PCA)

Clusters visualized using PCA components

K-Means Clustering Results Visualization

K-Means Clustering Results
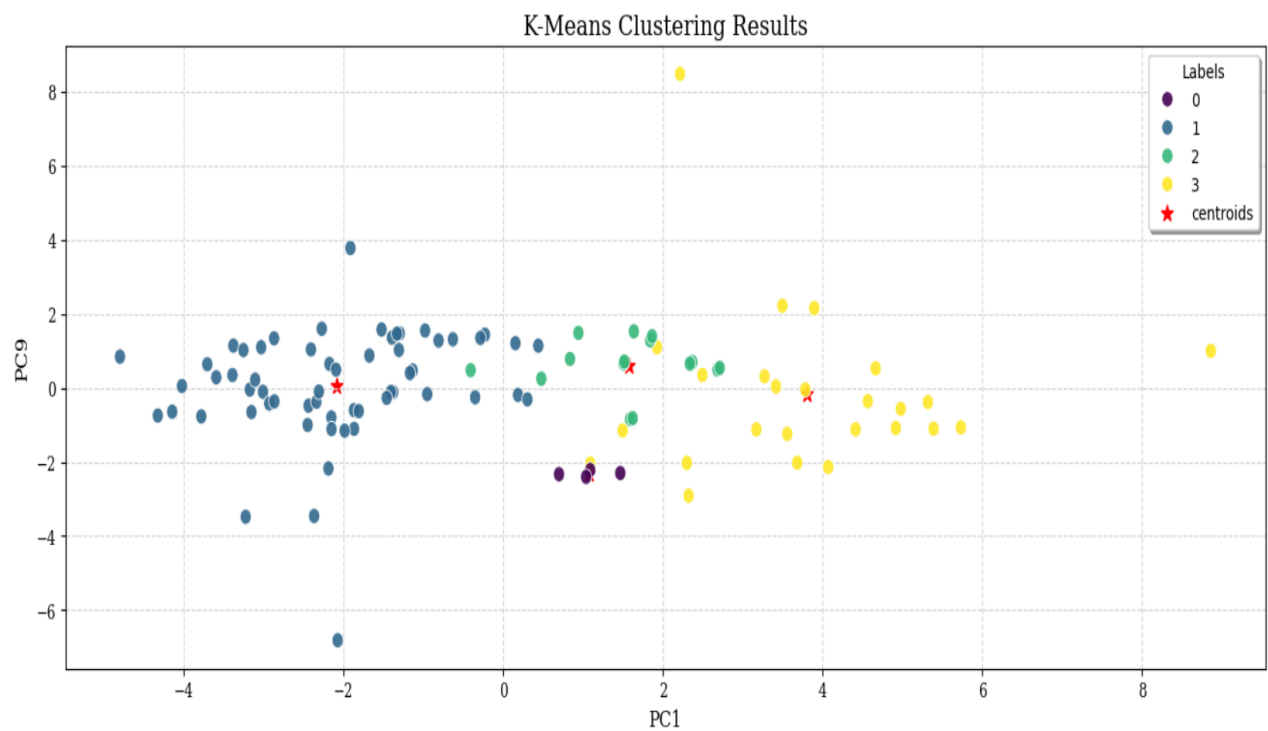
Conclusion

This project aimed to analyze the Electric Vehicle (EV) market in India using segmentation analysis to develop a feasible strategy for entering the market. By leveraging two datasets, one from Cardekho and another specific to electric cars, we performed a comprehensive analysis that included data preprocessing, outlier handling, dimensionality reduction using PCA, and clustering using K-Means.

Key Findings:

1. **Data Preprocessing and Outlier Handling**:
   - Successfully cleaned the data, handled missing values, and removed or capped outliers to ensure the dataset's quality for analysis.
   - Ensured that essential variables such as vehicle age, km driven, mileage, engine, max power, and selling price were prepared for further analysis.

2. **Dimensionality Reduction with PCA**:
   - Applied PCA to reduce the dimensionality of the dataset while retaining most of the variance. This step was crucial for visualizing high-dimensional data and simplifying the clustering process.
   - The elbow method was used to determine the optimal number of clusters, resulting in the selection of 4 clusters.

3. **Clustering with K-Means**:
   - Performed K-Means clustering to segment the data into distinct groups. Visualizations of the clusters using the first two principal components and the chosen principal components (PC1 and PC9) highlighted clear groupings in the data.
   - Each cluster represented a different market segment with unique characteristics, allowing for targeted strategic planning.

4. **Market Segmentation Insights**:
   - **Geographic Segmentation**: Identified potential early markets for EV adoption based on vehicle usage statistics and existing infrastructure such as charging stations.
   - **Demographic Segmentation**: Targeted specific demographics likely to adopt EVs early, such as environmentally conscious consumers and tech-savvy individuals.
   - **Psychographic Segmentation**: Focused on consumers with green values, a desire for the latest technology, and those seeking cost-efficient alternatives to traditional fuel vehicles.

- o **Behavioral Segmentation**: Considered the driving patterns and preferences of potential EV users to tailor product offerings and marketing strategies.

5. **Strategic Recommendations**:
   - o **Location**: Suggested focusing on metropolitan areas with higher rates of vehicle usage and better infrastructure for EVs.
   - o **Demographics**: Recommended targeting younger, environmentally conscious consumers with higher disposable incomes.
   - o **Product Pricing**: Proposed a strategic pricing range that considers the early market psychographics, balancing affordability with the perceived value of innovative EV technology.

PROJECTS LINK AND DATA SET :-

https://github.com/AkhileshMauryaa/Electrical_Vehical_MarketSegment_Analysis

Thanks You