Electric Vehicle Startup

by

Monalisa Burma

20th January, 2024

“Electric Vehicles are an essential strategy in the immediate term to reduce local emission and help improve local air quality”

-Kevin Emerson

Introduction

In the dynamic landscape of the Indian electric vehicle market, understanding consumer preferences and behaviors is paramount for industry stakeholders. This project aims to delve into the intricacies of electric vehicle adoption, employing a data-driven approach to uncover diverse market segments. Through the detailed analysis of two datasets encompassing electric car attributes and individual demographics, the project seeks to provide comprehensive insights, facilitate effective market segmentation, and deliver strategic recommendations for automakers, policymakers, and marketing teams. By bridging the gap between consumer behavior and market dynamics, this project endeavors to optimize strategies and foster the sustainable growth of the electric vehicle market in India.

• Background:

The automotive industry is undergoing a profound transformation driven by the imperative for sustainability, technological advancements, and a collective global commitment to mitigating climate change. This transformation is most prominently exemplified by the rise of Electric Vehicles (EVs), which represent a departure from the traditional internal combustion engine paradigm that has dominated the automotive landscape for over a century.

The push toward Electric Vehicles is propelled by the imperative to reduce carbon emissions, address environmental degradation, and transition toward cleaner energy sources. Unlike conventional vehicles reliant on fossil fuels, EVs operate on electric power stored in advanced batteries. These batteries, continuously evolving and becoming more efficient, are the linchpin of the electric mobility revolution.

1. Technological Advancements:

The evolution of battery technology, marked by enhanced energy density and prolonged battery life, has been instrumental in overcoming one of the primary barriers to EV adoption – range anxiety. Modern EVs boast impressive mileage on a single charge, making them viable options for both short commutes and long-distance travel.

Moreover, advancements in charging infrastructure have alleviated concerns about the availability and accessibility of charging stations. From home-based charging solutions to a burgeoning network of public charging stations, the infrastructure supporting EVs continues to expand, fostering increased consumer confidence in the feasibility of electric mobility.

2. Global Environmental Consciousness:

The imperative for sustainability, driven by a growing awareness of climate change and environmental degradation, has significantly influenced consumer preferences. Governments and regulatory bodies worldwide are incentivizing the adoption of electric vehicles through subsidies, tax credits, and stringent emission standards, signaling a collective commitment to sustainable transportation.

3. Emergence of New Market Players:

The surge in interest and investment in the EV sector has led to the emergence of new market players, ranging from established automotive giants to innovative startups. This diversification injects dynamism into the market, fostering competition, innovation, and a continual push toward technological excellence.

- ## Significance of Electric Vehicles:

Electric Vehicles (EVs) represent a pivotal and transformative force in the automotive landscape, offering a myriad of benefits that extend beyond individual transportation preferences. The significance of EVs is underscored by their impact on environmental sustainability, technological innovation, and the broader socio-economic landscape.

1. Environmental Sustainability:

At the forefront of the significance of EVs is their role in mitigating environmental challenges. Unlike traditional internal combustion engine vehicles that rely on fossil fuels, EVs operate on electricity, which can be sourced from renewable and cleaner energy options. The reduction in tailpipe emissions contributes to lower air pollution, diminished greenhouse gas emissions, and a decreased reliance on finite fossil fuel resources. As societies worldwide grapple with the consequences of climate change, the adoption of EVs emerges as a tangible and impactful step toward a more sustainable future.

2. Technological Innovation:

The widespread adoption of EVs propels the automotive industry into a new era of technological innovation. The development of high-performance batteries, advancements in electric drivetrain systems, and the integration of smart technologies within vehicles redefine the possibilities of modern transportation. EVs are not merely a mode of conveyance; they embody the convergence of automotive engineering, electronics, and software, fostering a landscape where innovation is a constant.

3. Energy Independence and Diversification:

The transition to electric mobility contributes to the diversification of energy sources. By enabling the integration of renewable energy into the grid to power EVs, nations reduce their dependency on conventional energy resources. This shift toward energy independence enhances energy security, reduces geopolitical vulnerabilities associated with fossil fuel dependence, and opens avenues for sustainable energy practices.

4. Economic Opportunities:

The rise of the EV sector creates a ripple effect across economies. It stimulates job creation in manufacturing, research and development, and the establishment of charging infrastructure. Governments and private entities investing in the EV ecosystem not only contribute to a greener planet but also foster economic growth through the development of a robust and forward-looking industry.

- ## Overview of the EV Market:

The EV market is dynamic, characterized by rapid technological innovation, evolving consumer preferences, and the emergence of new players. Understanding the nuances within this market is essential for stakeholders, including manufacturers, policymakers, and investors, to navigate and capitalize on opportunities in this evolving landscape.

Market Overview of Electric Vehicles in India:

1. Current Landscape:

India's electric vehicle market has witnessed significant developments in recent years, driven by a combination of environmental concerns, government initiatives, and a growing appetite for sustainable transportation. The market encompasses a diverse range of electric vehicles, including two-wheelers, three-wheelers, and four-wheelers.

2. Government Initiatives:

The Indian government has taken proactive steps to promote the adoption of electric vehicles. Initiatives include the Faster Adoption and Manufacturing of Hybrid and Electric Vehicles (FAME) scheme, which provides financial incentives, subsidies, and support for research and development in the electric mobility sector. These measures aim to accelerate the growth of EVs and reduce the dependence on traditional internal combustion engine vehicles.

3. Two-Wheeler Dominance:

Two-wheelers, particularly electric scooters and motorcycles, have emerged as dominant players in the Indian EV market. The affordability, ease of charging, and suitability for urban commuting make electric two-wheelers an attractive choice for a wide range of consumers. Several domestic and international manufacturers have entered this segment, contributing to its rapid expansion.

4. Three-Wheelers and Public Transport:

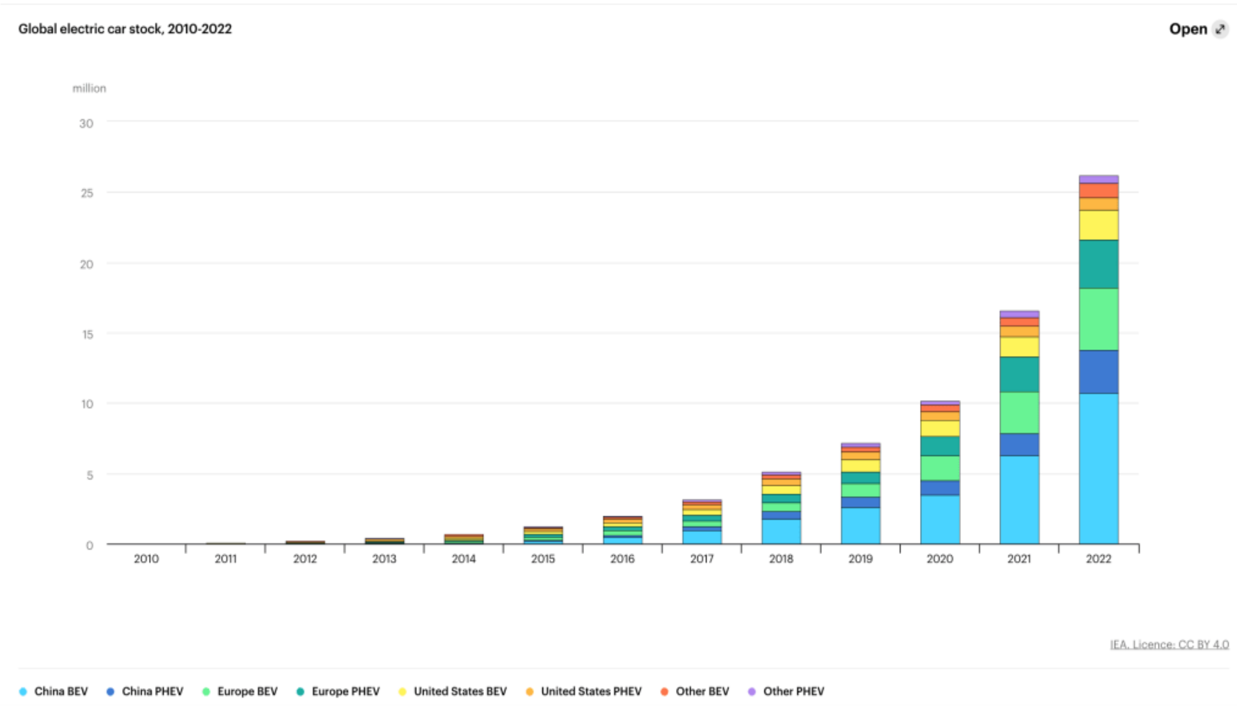
Electric three-wheelers, often used for last-mile connectivity and public transportation, have gained traction. E-rickshaws and electric three-wheeler autos provide eco-friendly alternatives in urban and peri-urban areas. Government support for electrifying public transport is a key driver in this segment.

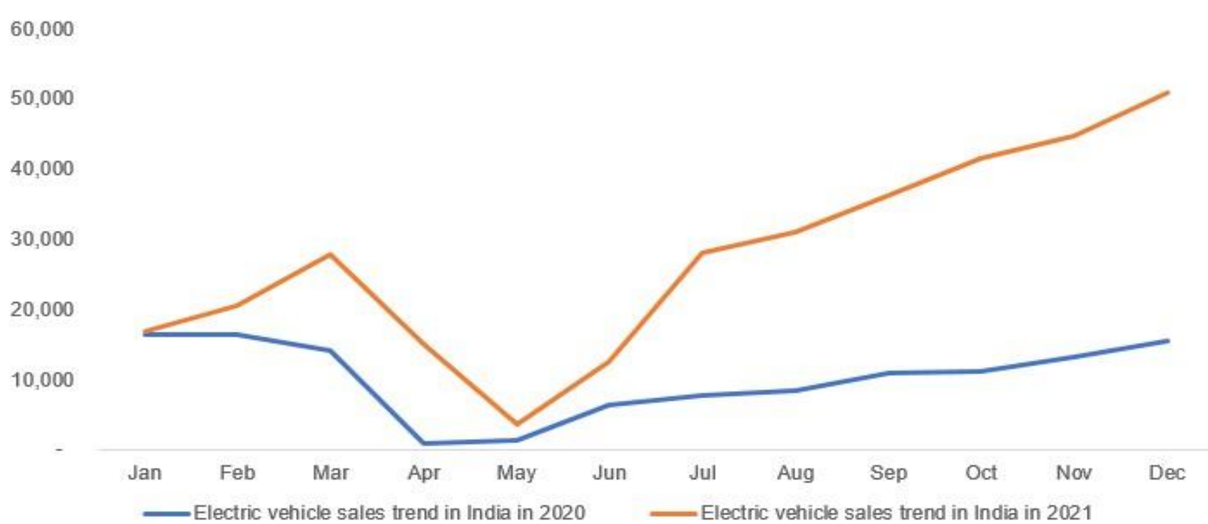
5. Four-Wheeler Segment:

While the penetration of electric four-wheelers is relatively modest compared to two-wheelers, there is a growing interest in electric cars. Several automakers are introducing electric models, and the government's push for electrification of the automotive fleet is expected to drive further adoption. The focus is not only on private electric cars but also on electric taxis and shared mobility solutions.

7. Industry Collaboration:

The electric vehicle market in India has seen collaborations between automotive manufacturers, technology companies, and energy providers. These collaborations are aimed at leveraging synergies, fostering innovation, and addressing challenges related to battery technology, charging infrastructure, and affordability.





Problem Statement

Despite the increasing popularity of electric vehicles, there exists a critical need to decipher the nuanced factors influencing consumer choices and market trends. The lack of a tailored understanding of diverse consumer segments poses a challenge for stakeholders in devising targeted strategies. This project addresses this gap by leveraging data analysis techniques to discern patterns, identify market segments, and offer strategic recommendations, ultimately contributing to the informed development and positioning of electric vehicles in the Indian market.

Key Goals of the Report

- To provide a comprehensive understanding of the electric vehicle market in India, including key trends, consumer preferences, and market dynamics.
- To offer actionable recommendations for industry stakeholders.
- To assist companies in effectively positioning their electric vehicles in the market by aligning with consumer preferences.
- To conduct a competitive analysis to identify strengths, weaknesses, opportunities, and threats in the Indian electric vehicle market.

Detailed Analysis

Dataset 1: Electric car data
link:
https://github.com/monalisaburma/EV_Market_Segment/blob/main/ElectricCarData_Norm.csv

Description:

The dataset consists of 103 entries with information on various electric car attributes. It includes features such as acceleration (Accel), top speed (TopSpeed), range, efficiency, and pricing. The dataset contains a mix of numerical and categorical variables, including details about powertrain, plug type, body style, and segment. Additionally, there are a few missing values in the 'FastCharge' column.

[2]:
```
df=pd.read_csv(r"C:\Users\ASUS\Downloads\ElectricCarData_Norm.csv")
df.head()
```

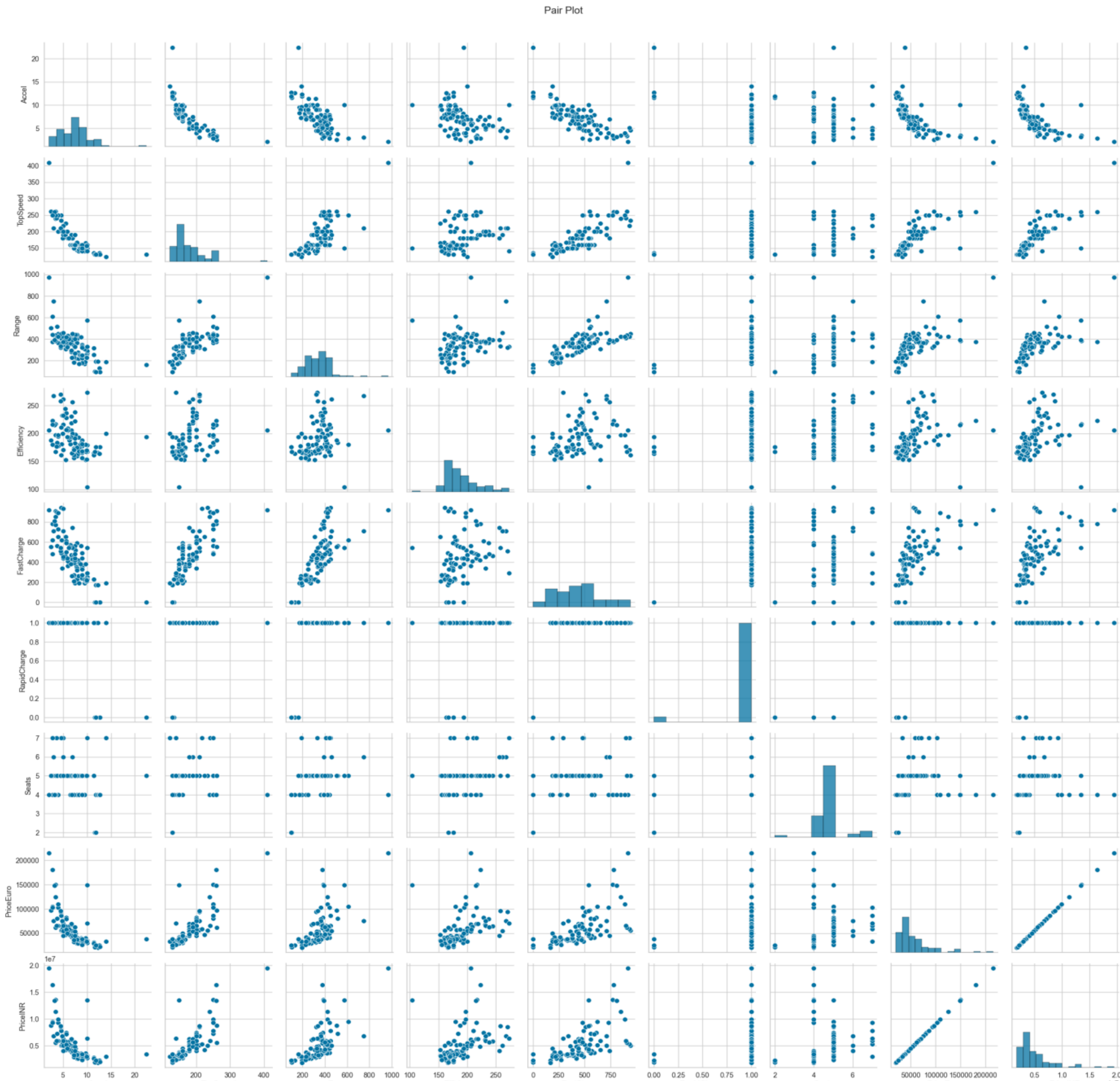
[2]:

	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

[5]:
```
df.info()
```

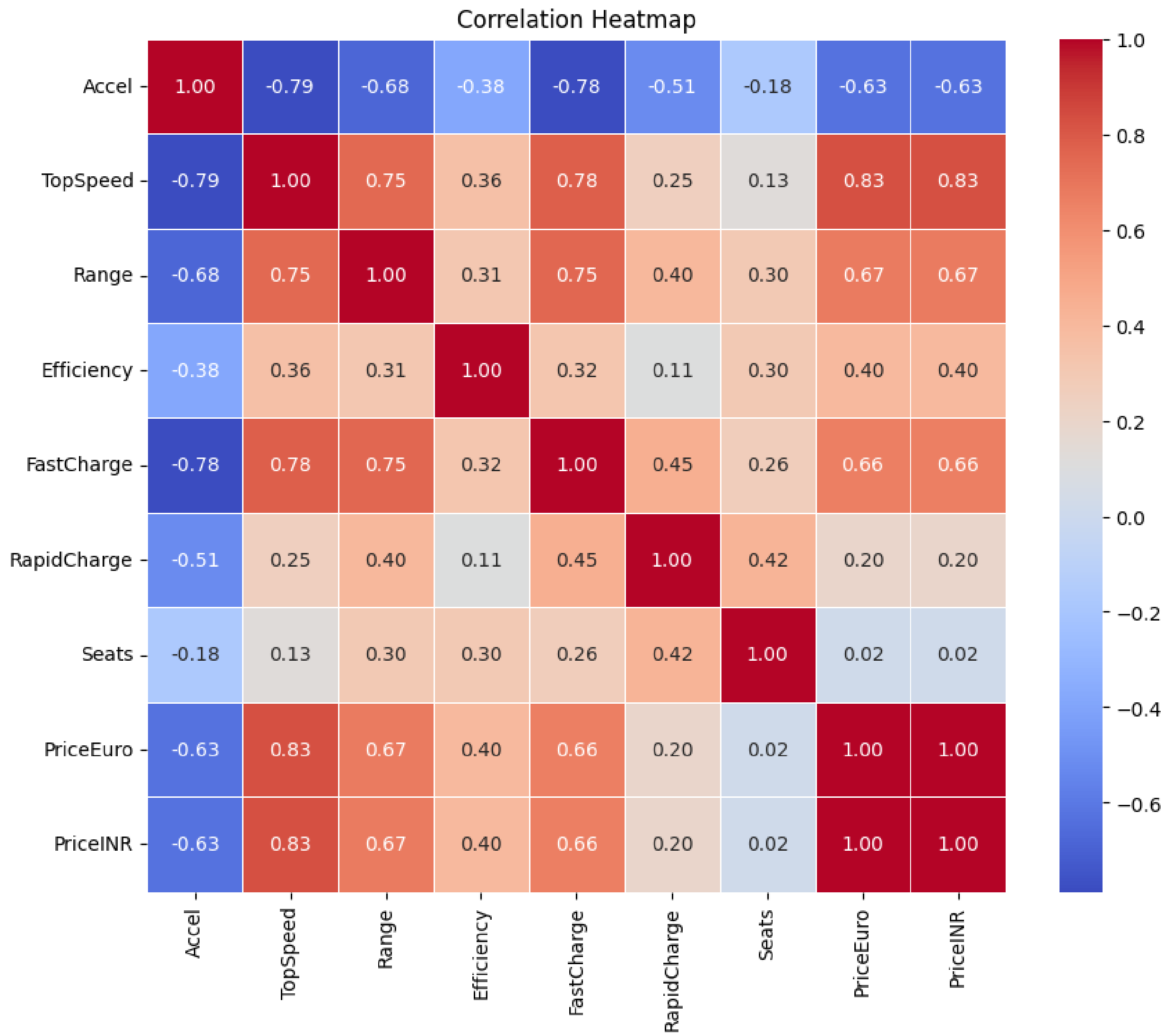
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 103 entries, 0 to 102
Data columns (total 15 columns):
 #   Column        Non-Null Count  Dtype  
---  ------        --------------  -----  
 0   Brand         103 non-null    object 
 1   Model         103 non-null    object 
 2   Accel         103 non-null    float64
 3   TopSpeed      103 non-null    int64  
 4   Range         103 non-null    int64  
 5   Efficiency    103 non-null    int64  
 6   FastCharge    98 non-null     float64
 7   RapidCharge   103 non-null    int64  
 8   PowerTrain    103 non-null    object 
 9   PlugType      103 non-null    object 
 10  BodyStyle     103 non-null    object 
 11  Segment       103 non-null    object 
 12  Seats         103 non-null    int64  
 13  PriceEuro     103 non-null    int64  
 14  PriceINR      103 non-null    float64
dtypes: float64(3), int64(6), object(6)
memory usage: 12.2+ KB
```

Pair Plot Insights:



From this pair plot, we can see that 'PriceEuro' and 'PriceINR' are highly positively correlated, as they represent the same price in different currencies. We can also see that 'Accel' and 'TopSpeed' are moderately negatively correlated, as they represent trade-offs between acceleration and top speed.

Correlation Matrix:



The correlation heatmap is providing insights into the relationships between different features of electric cars. For instance:

- ‘Accel’ (Acceleration) has a strong positive correlation with ‘TopSpeed’, meaning cars that accelerate faster tend to have higher top speeds.
- ‘PriceEuro’ and ‘PriceINR’ are highly correlated with ‘TopSpeed’, indicating that cars with higher top speeds are generally more expensive.
- There is a negative correlation between ‘FastCharge’ and both ‘Accel’ and ‘TopSpeed’, suggesting that faster charging might be associated with lower performance metrics.
- The ‘Efficiency’ is not strongly correlated with most of the other features, indicating it varies independently.

Principal Component Analysis (PCA):

```
[57]: from sklearn.decomposition import PCA
pca = PCA(n_components=8)
principal_components = pca.fit_transform(x_standardized)
data_pca = pd.DataFrame(principal_components, columns=[f'PC{i}' for i in range(1, 9)])
data_pca
```

[57]:

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
0	2.118016	0.244255	-1.430547	0.759397	-1.319033	0.071498	-0.713288	-0.326257
1	-1.528982	-0.587461	-0.450122	0.182961	0.294766	-0.139906	0.159079	-0.358488
2	1.262657	0.031005	-0.595544	0.131253	-0.660926	-0.030726	0.101642	-0.060218
3	0.758496	-0.128263	0.320195	-0.206427	0.003949	0.074760	-0.407353	0.137331
4	-2.211034	0.244843	-0.644147	-0.910628	0.342095	-0.302337	0.037561	-0.193960
...	...	...	...	...	...	...	...	...
98	-0.250036	-0.459912	-0.021948	-0.173428	-0.086265	0.186262	-0.124875	0.134913
99	1.989282	0.211116	1.778367	-1.128999	0.143838	-0.369675	-0.105644	0.100670
100	0.431311	-0.161516	-0.008665	-0.258134	-0.321145	-0.262346	0.333194	-0.241505
101	1.201606	-0.102469	1.026250	-0.714663	-0.083723	0.108249	0.352792	-0.084532
102	0.921885	-0.227871	1.266420	-0.480067	0.231364	0.524587	-0.150906	-0.343308

In this section, Principal Component Analysis (PCA) is applied to the standardized dataset (x_standardized). PCA is a dimensionality reduction technique used to transform the original features into a set of uncorrelated variables known as principal components. The goal is to capture the maximum variance in the data, allowing for a more concise representation. Here, n_components is set to 8, indicating that eight principal components will be generated.

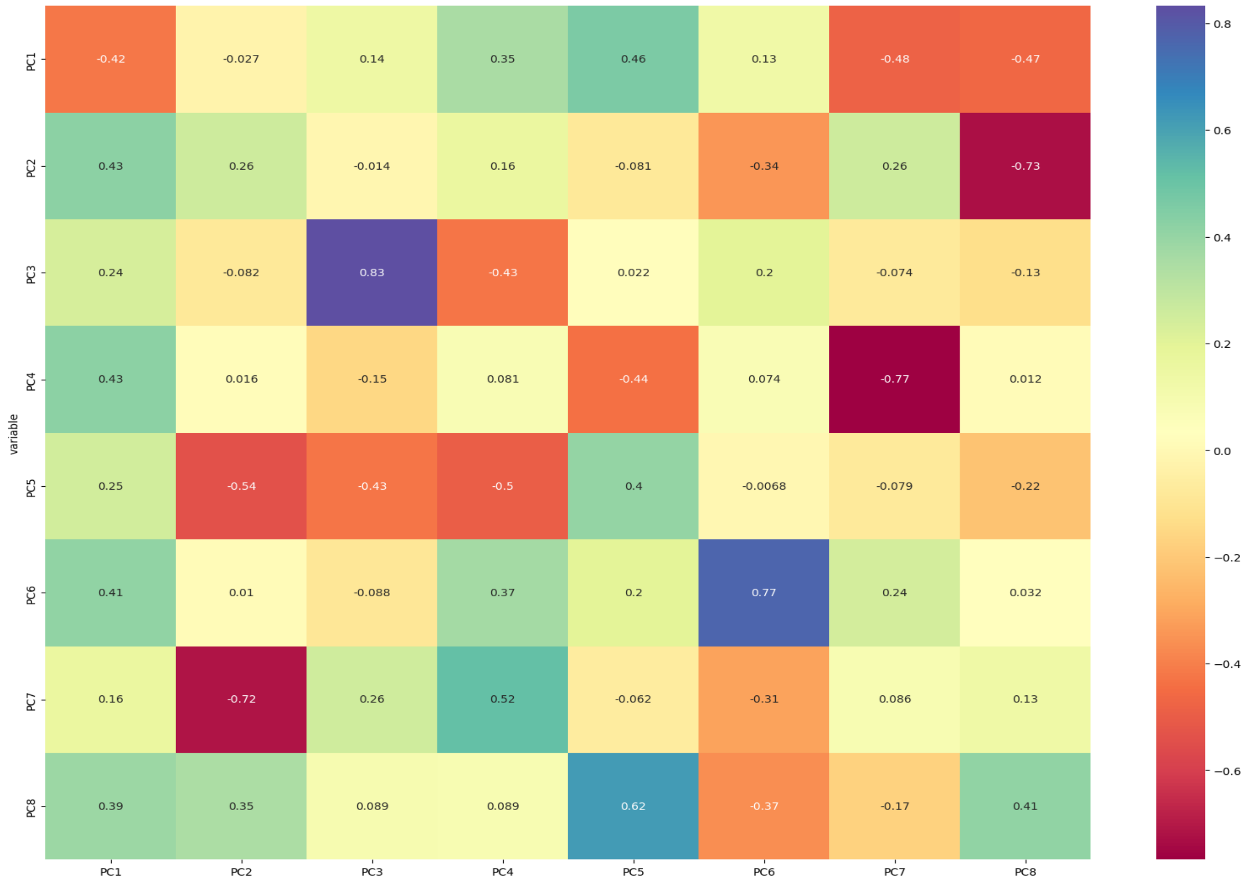
```
[73]: explained_variance_ratio = pca.explained_variance_ratio_
cumulative_explained_variance = np.cumsum(np.round(explained_variance_ratio, decimals=4) * 100)
```

```
[74]: print("Explained Variance Ratio:")
print(f"PC1: {explained_variance_ratio[0]*100:.2f}%")
print(f"PC2: {explained_variance_ratio[1]*100:.2f}%")

print("\nCumulative Explained Variance:")
print(f"PC1 and PC2: {cumulative_explained_variance[1]:.2f}%")
```

```
Explained Variance Ratio:
PC1: 55.19%
PC2: 16.01%

Cumulative Explained Variance:
PC1 and PC2: 71.20%
```

Here, PC1 is capturing 55.19% of the variance, and PC2 is contributing an additional 16.01%. Together, these two components explaining 71.20% of the cumulative variance in the standardized dataset. The high cumulative explained variance is suggesting that PC1 and PC2 effectively summarize the essential features of the data, providing valuable insights into its underlying structure.

Correlation Matrix:

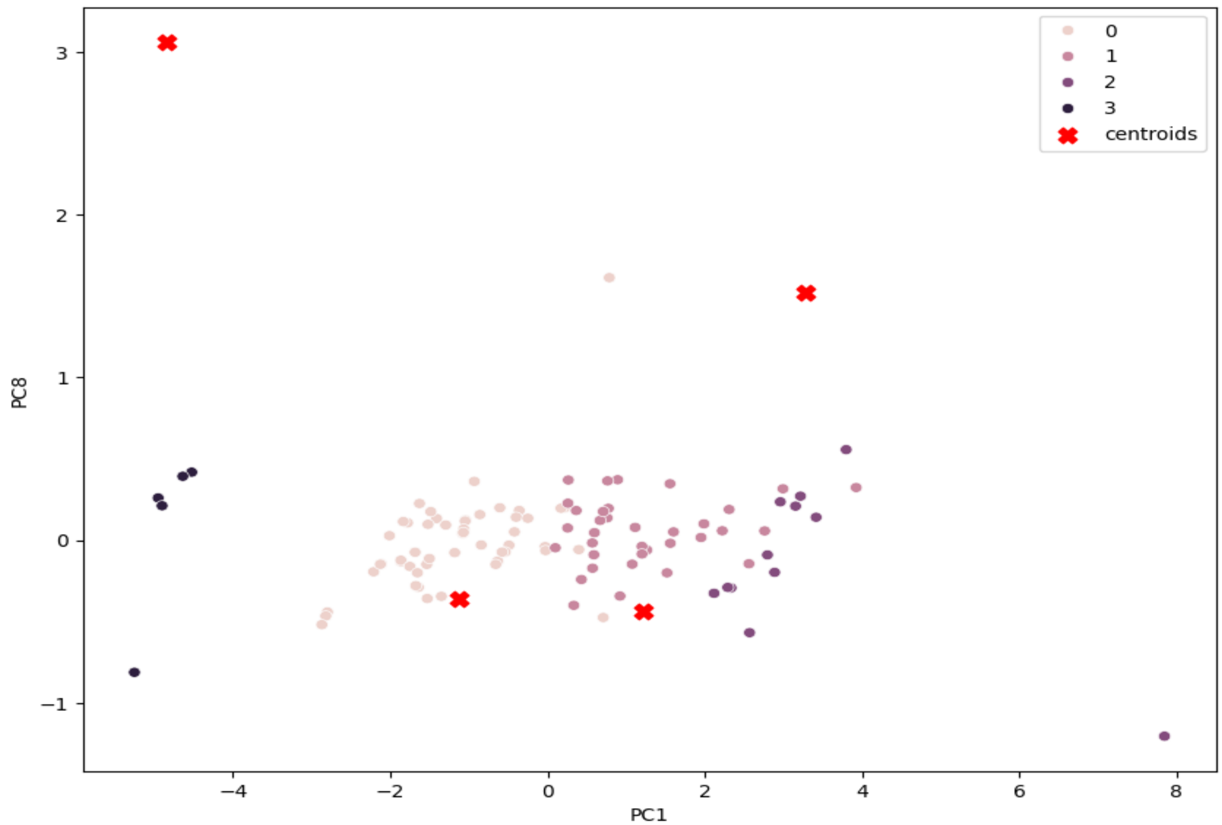


The correlation matrix plot is showing relationships between PCs and original variables:

- PC1 is strongly positively correlated with Range and PriceINR.
- PC2 is moderately negatively correlated with Accel, TopSpeed, and Range.
- PC3 is strongly positively correlated with TopSpeed and moderately negatively correlates with PriceINR.
- PC4 is strongly positively correlated with Accel and moderately negatively correlated with FastCharge.
- PC5 is strongly positively correlated with Efficiency.
- PC6 is moderately positively correlated with FastCharge.
- PC7 is strongly negatively correlated with RapidCharge.

KMeans Cluster Analysis:

```
[70]: #Visulazing clusters
      sn.scatterplot(data=data_pca, x="PC1", y="PC8", hue=kmeans.labels_)
      plt.scatter(kmeans.cluster_centers_[:,0], kmeans.cluster_centers_[:,1],
                  marker="X", c="r", s=80, label="centroids")
      plt.legend()
      plt.show()
```

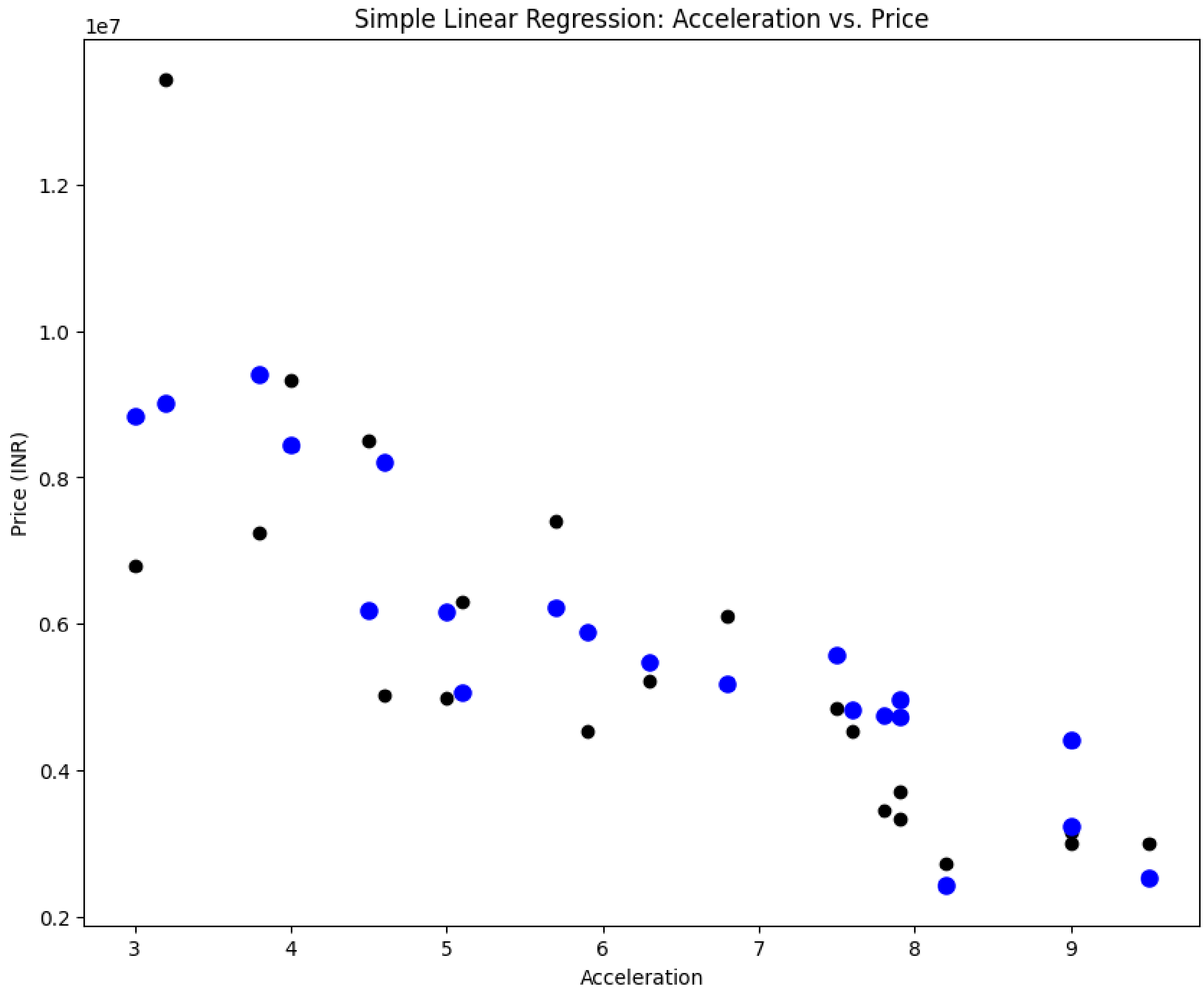


In the scatter plot, we can observe clear clusters and centroids resulting from the KMeans clustering algorithm. Each cluster is assigned a unique color, and centroids are marked with red X symbols. Here's a brief breakdown of the clusters:

- Cluster 0 (Purple): The data points are scattered around the plot's center, indicating moderate values for both PC1 and PC8.
- Cluster 1 (Light Pink): Spreading towards higher positive values of PC1 with moderate PC8 values.
- Cluster 2 (Dark Purple): Positioned at higher positive values for both PC1 and PC8, hinting at a correlation between these principal components.
- Cluster 3 (Blue): Positioned at negative PC1 values and moderate to high positive PC8 values.

The centroids act as central points, providing an average location for each cluster in terms of principal components. This KMeans clustering analysis, revealing four distinctive groups based on similarities in principal components.

Linear Regression for Price Prediction:



In the pursuit of forecasting car prices (in INR), I implemented a simple linear regression model based on acceleration, top speed, and range. The model underwent training on 80% of the dataset, subsequently being tested on the remaining 20%. The obtained R-squared value stands at approximately 56.16%, indicating the model's capability to elucidate around 56% of the variance in car prices. In the plot, black dots depict actual data points, and blue dots represent predicted values from the model. The absence of a clear linear trend suggests that a linear regression might not be the ideal fit for this data.

Dataset2 : Indian automoble buying behavour study 1.0.

Link:
https://github.com/monalisaburma/EV_Market_Segment/blob/main/Indian%20automoble%20buying%20behavour%20study%201.0.csv

```
[2]: df = pd.read_csv(r"C:\Users\ASUS\Downloads\Indian automoble buying behavour study 1.0.csv")
     df.head()
```

[2]:

	Age	Profession	Marrital Status	Education	No of Dependents	Personal loan	House Loan	Wife Working	Salary	Wife Salary	Total Salary	Make	Price
0	27	Salaried	Single	Post Graduate	0	Yes	No	No	800000	0	800000	i20	800000
1	35	Salaried	Married	Post Graduate	2	Yes	Yes	Yes	1400000	600000	2000000	Ciaz	1000000
2	45	Business	Married	Graduate	4	Yes	Yes	No	1800000	0	1800000	Duster	1200000
3	41	Business	Married	Post Graduate	3	No	No	Yes	1600000	600000	2200000	City	1200000
4	31	Salaried	Married	Post Graduate	2	Yes	No	Yes	1800000	800000	2600000	SUV	1600000

This dataset encompasses key information about individuals' demographic and financial aspects. Fields like 'Age', 'Profession', 'Marital Status', 'Education', and 'No of Dependents' shed light on personal attributes. Additionally, financial details such as 'Salary', 'Wife Salary', and 'Total Salary' are presented, alongside indicators of loan status ('Personal loan' and 'House Loan'). Insights into car-related preferences include 'Make' and 'Price', forming a comprehensive dataset for market segmentation analysis.

```
[3]: df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 99 entries, 0 to 98
Data columns (total 13 columns):
 #   Column           Non-Null Count  Dtype
---  ------           --------------  -----
 0   Age              99 non-null     int64
 1   Profession       99 non-null     object
 2   Marrital Status  99 non-null     object
 3   Education        99 non-null     object
 4   No of Dependents 99 non-null     int64
 5   Personal loan    99 non-null     object
 6   House Loan       99 non-null     object
 7   Wife Working     99 non-null     object
 8   Salary           99 non-null     int64
 9   Wife Salary      99 non-null     int64
 10  Total Salary     99 non-null     int64
 11  Make             99 non-null     object
 12  Price            99 non-null     int64
dtypes: int64(6), object(7)
memory usage: 10.2+ KB
```

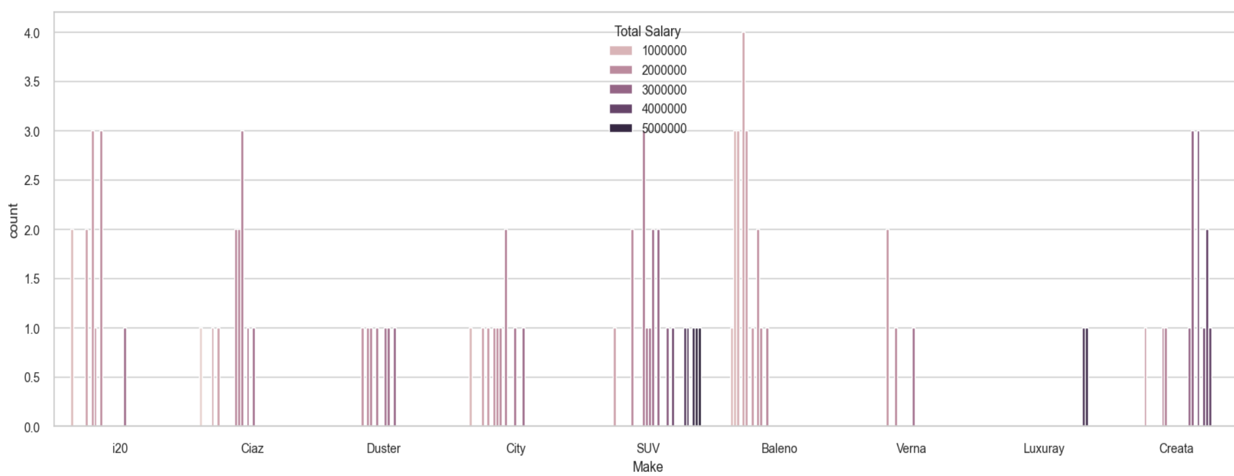
[4]:
```
df.describe()
```

[4]:

	Age	No of Dependents	Salary	Wife Salary	Total Salary	Price
count	99.000000	99.000000	9.900000e+01	9.900000e+01	9.900000e+01	9.900000e+01
mean	36.313131	2.181818	1.736364e+06	5.343434e+05	2.270707e+06	1.194040e+06
std	6.246054	1.335265	6.736217e+05	6.054450e+05	1.050777e+06	4.376955e+05
min	26.000000	0.000000	2.000000e+05	0.000000e+00	2.000000e+05	1.100000e+05
25%	31.000000	2.000000	1.300000e+06	0.000000e+00	1.550000e+06	8.000000e+05
50%	36.000000	2.000000	1.600000e+06	5.000000e+05	2.100000e+06	1.200000e+06
75%	41.000000	3.000000	2.200000e+06	9.000000e+05	2.700000e+06	1.500000e+06
max	51.000000	4.000000	3.800000e+06	2.100000e+06	5.200000e+06	3.000000e+06

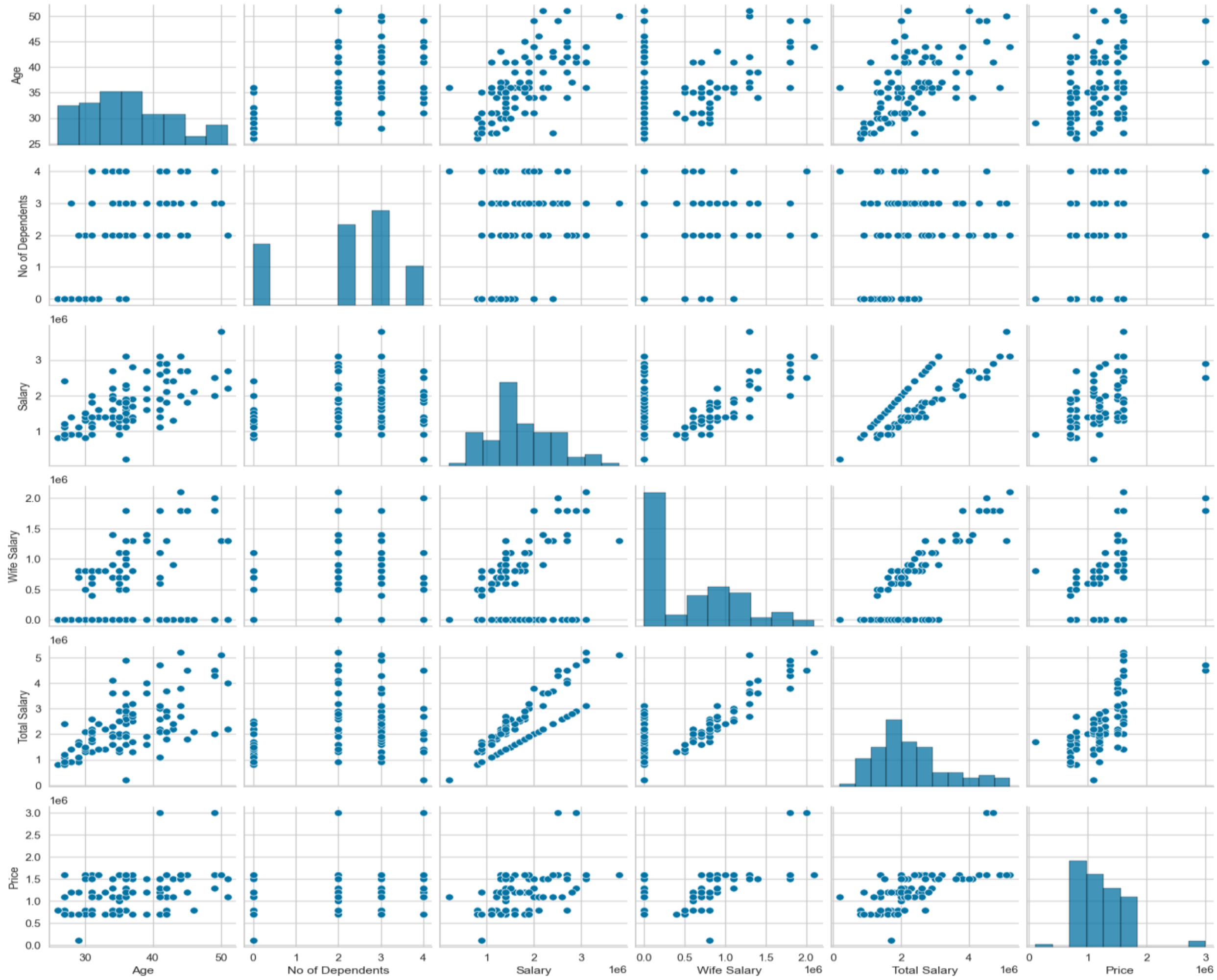
- The Indian behavior dataset provides valuable insights into individual demographics, including age, salary, and dependents, crucial factors influencing electric vehicle (EV) purchasing decisions.
- Understanding the behavior of potential EV buyers in India is essential for effective market segmentation, allowing for targeted strategies that align with customer preferences and financial capabilities.
- The dataset's information on marital status, education, and profession helps in identifying patterns and preferences that influence the adoption of electric vehicles among different demographic segments.
- Examining variables like 'Salary' and 'Wife Salary' aids in comprehending the financial aspects of potential EV customers, providing a foundation for pricing strategies and affordability considerations.
- With data on 'Make' and 'Price,' the dataset facilitates a comprehensive analysis of the factors influencing the choice of electric car brands and the price sensitivity of different customer segments in the Indian market.

In the graphs, we can observe the frequency distribution (histograms) and probability density function (KDE plots) for variables like 'Age,' 'No of Dependents,' 'Salary,' 'Wife Salary,' 'Total Salary,' and 'Price.' These distributions are revealing a right-skewed pattern, indicating a concentration of values at the lower end with some outliers at the higher end. Notably, 'Salary,' 'Wife Salary,' and 'Total Salary' sharing a similar shape and range due to common influencing factors like profession and education. The age distribution is peaking between 30 and 40 years, while the 'No of Dependents' is varying from 0 to 5.



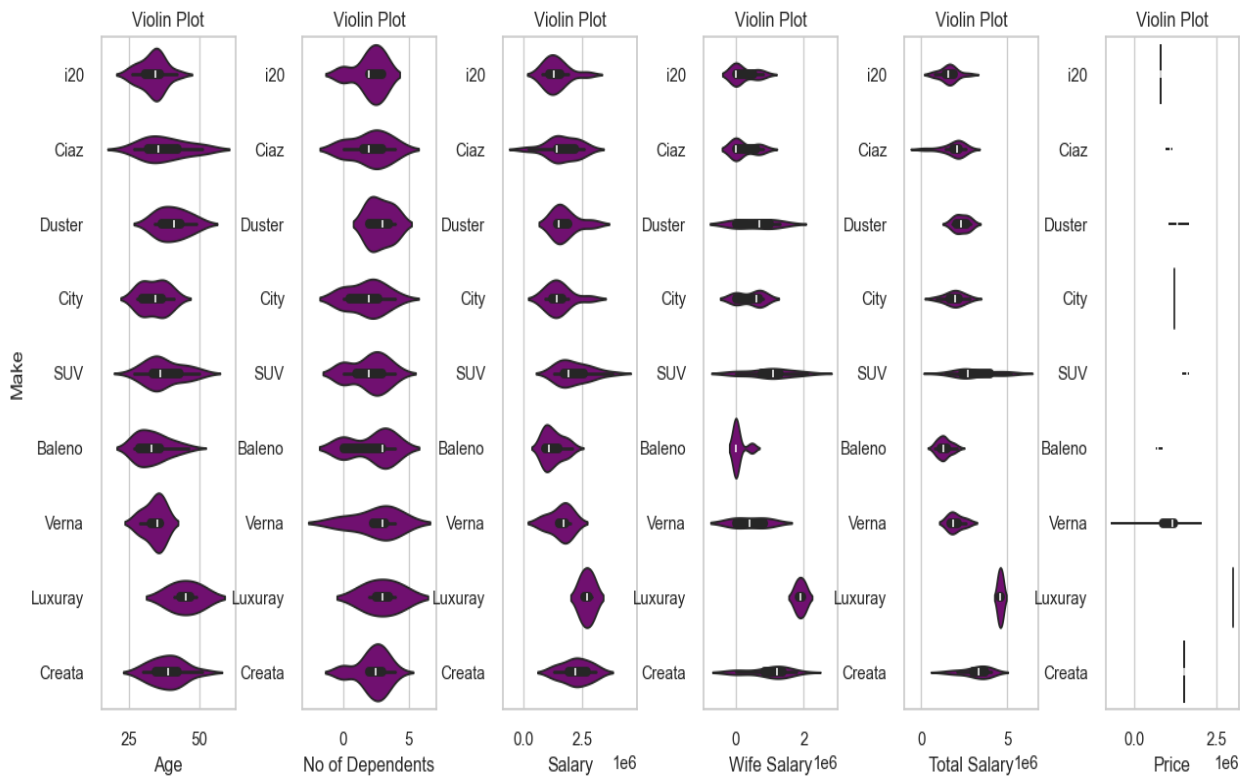
Here, the graph is showing the frequency of car makes (i20, Ciaz, Duster, etc.) with counts ranging from 0 to 4. It also illustrates the distribution of total salary within each car make, using colors for different salary ranges (1000000 to 5000000). Patterns emerge, indicating associations between car make and total salary; for instance, i20 buyers mostly have a salary of 1000000, while Luxury buyers tend to have a salary of 5000000. SUV and Creta exhibit diverse total salary distributions, appealing to varied income groups.

Pair plot:

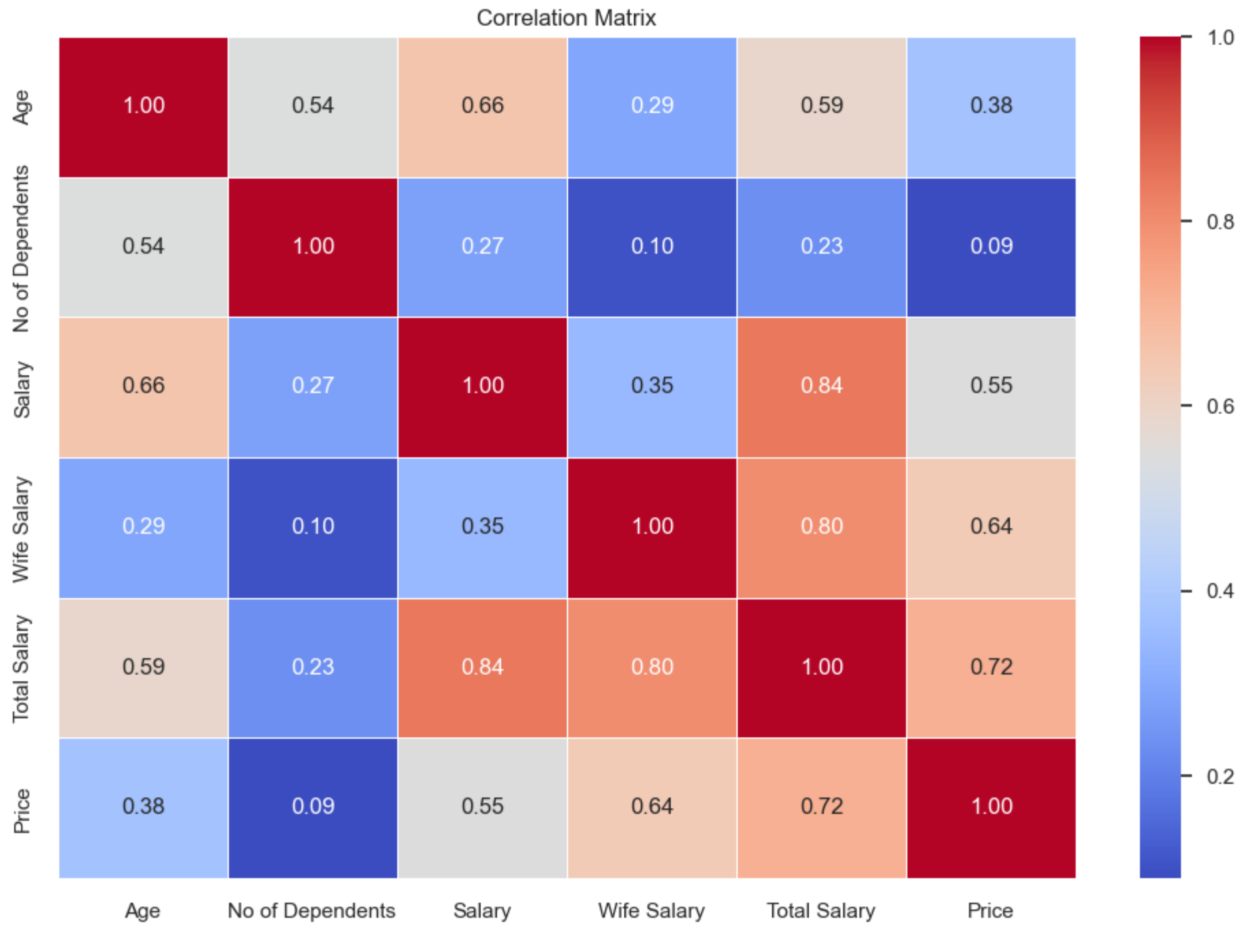


- In this graph, I have explored the frequency and correlation of variables like 'Age,' 'No of Dependents,' 'Salary,' 'Wife Salary,' 'Total Salary,' and 'Price.'
- Notably, these variables exhibit right-skewed distributions, indicating clustering at the lower end with some outliers at the higher end.
- 'Salary,' 'Wife Salary,' and 'Total Salary' are showing a strong positive correlation, aligning with the expected influence of similar factors.
- However, 'Price' displays a weak positive correlation with income variables, suggesting other factors influence the electric car's price, such as brand, model, features, or customer preferences.
- 'Age' and 'No of Dependents' show no clear correlation with other variables, indicating independence or a complex relationship with the dataset's other aspects.

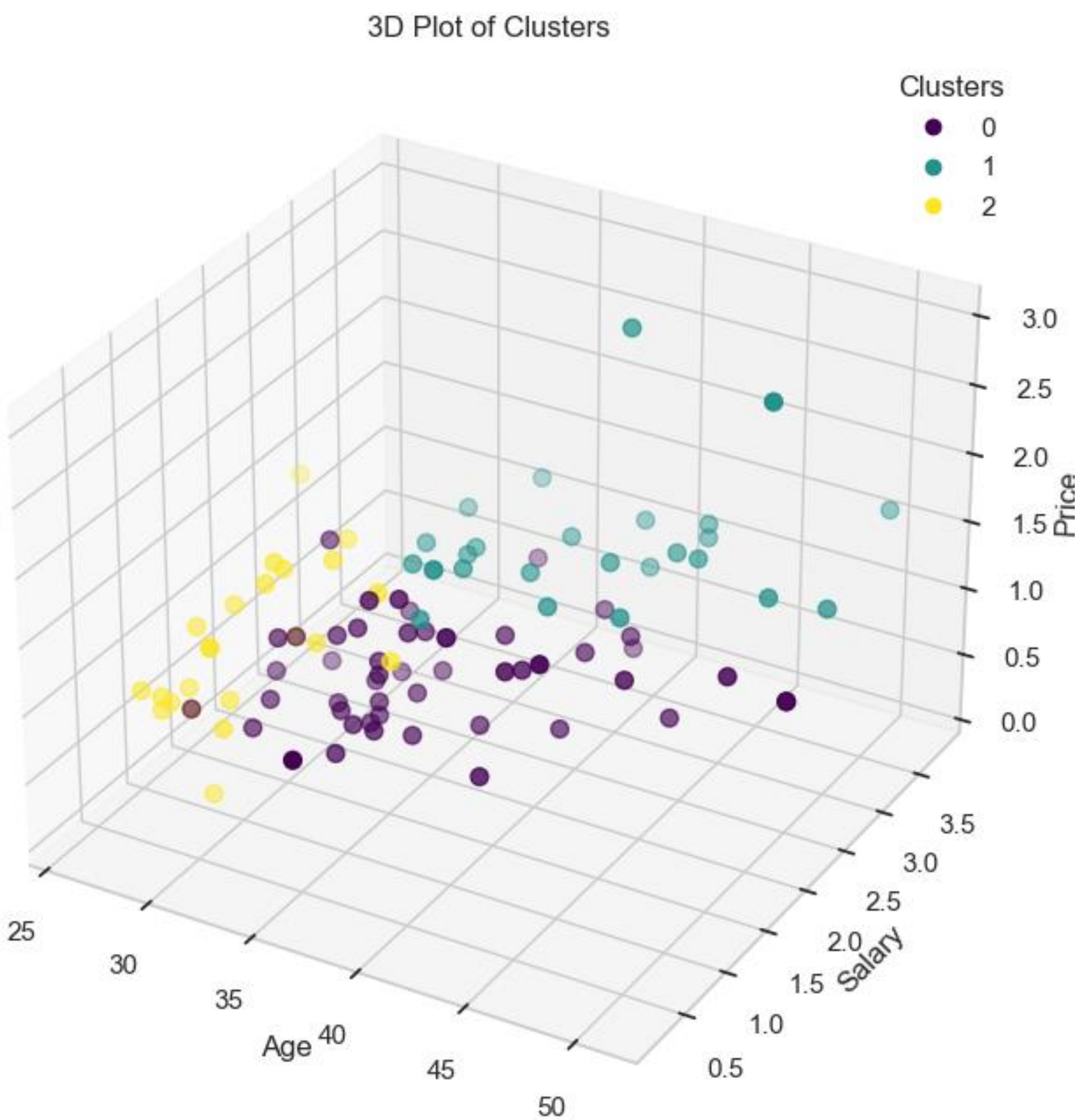
Violin plot:



- In this violin plot, thicker sections denote higher data density, while thinner parts indicate lower density. Median values are marked by white dots, and interquartile ranges by black bars.
- Notably, these variables exhibit diverse distributions across car makes. For instance, customers purchasing Ciaz are typically younger with fewer dependents, lower salaries, and prices compared to those opting for Luxury.
- SUV buyers tend to have more dependents, higher salaries and prices, and a broader age range than Baleno customers. This is also highlighting the outliers, like a Luxury customer aged over 100 and a Creta buyer with a salary exceeding 2 million, deviating significantly from median values.



From this heatmap we can see the linear relationships between variables, with coefficients ranging from -1 to 1. The heatmap color-codes these coefficients from blue to red, signifying magnitude and direction. Darker colors are denoting stronger correlations. Notably, 'Salary,' 'Wife Salary,' and 'Total Salary' are exhibiting a robust positive correlation, aligning with expectations. 'Price' showing a weaker positive correlation with income variables, suggesting other influencing factors like brand and features. 'Age' and 'No of Dependents' lack clear correlations, implying independence or complex relationships.



From this advanced 3d plot we can clearly identify the 3 cluster groups. While looking at the patterns we can observe that as age increases, the vehicle cost rises. Similarly, spending on cars tends to increase with the number of dependents and salary.

Observations

In this project, various exploratory data analysis (EDA) techniques were employed to gain insights into the market segmentation.

- **Principal Component Analysis (PCA):** PCA was applied to the first dataset to unearth underlying patterns in the standardized data. This dimensionality reduction technique transformed the original features into uncorrelated variables, known as principal components. PCA revealed that PC1 and PC2 effectively summarized essential features, explaining a substantial 71.20% of the cumulative variance. The correlation matrix plot further unveiled relationships between principal components and original variables, providing valuable insights into the dataset's structure.
- **KMeans Clustering:** This unsupervised machine learning algorithm was applied to both datasets to unveil distinctive market segments. For the first dataset, a scatter plot showcased clear clusters with centroids, aiding in the identification of groups based on principal components. The second dataset revealed three distinct clusters through an advanced 3D plot, offering insights into correlations between age, vehicle cost, dependents, and salary. These analyses lay the foundation for targeted strategies aligning with demographic and financial attributes in the electric vehicle market.
- **Linear Regression for Price Prediction:** A simple linear regression model was implemented to forecast car prices based on acceleration, top speed, and range. Although the model's fit was not optimal, it provided insights into the relationship between these attributes and pricing, aiding in pricing strategy considerations.

Identifying Optimal Market Segmentation Variables:
Based on the analysis, the top 4 variables/features for creating optimal market segments are,

- **Salary/Income:** Financial aspects play a crucial role in determining purchasing power and preferences.
- **Age:** Age can influence lifestyle choices, including the preference for electric vehicles.
- **Car Make/Model Preferences**: Analyzing the preferred electric car brands or models among consumers.
- **Education/Profession:** These demographic factors can contribute to understanding the lifestyle and preferences of potential electric vehicle buyers.

Conclusion

In conclusion, this project provides a comprehensive analysis of the Indian electric vehicle market, offering valuable insights into consumer behavior, preferences, and market dynamics. The outlined objectives aim to facilitate targeted marketing strategies, inform strategic decisions, and contribute to the sustainable growth of the electric vehicle sector. The meticulous data-driven approach employed in this project positions it as a valuable resource for industry stakeholders aiming to navigate and thrive in the evolving landscape of electric mobility in India.

- **Demographic Segmentation:**

  Understanding the demographic landscape revealed pivotal insights. Age emerged as a key determinant, with the peak adoption of electric vehicles occurring between 30 and 40 years. Marital status and education levels played a subtle yet influential role, hinting at the importance of lifestyle factors in consumer choices. By delving into these demographic intricacies, stakeholders can tailor marketing strategies to resonate with the preferences of specific age groups and lifestyles.

- **Behavioral Insights:**

  The behavioral analysis illuminated the interconnected web of factors influencing electric vehicle adoption. Financial aspects, such as salary and dependents, intertwined with preferences for specific car makes. SUV buyers, for instance, exhibited a broader age range, higher salaries, and prices compared to Baleno customers. These behavioral nuances unlock opportunities for targeted marketing campaigns, aligning product offerings with consumer preferences and financial capacities.

- **Target Segments:**

  The culmination of analyses, including KMeans clustering and linear regression, unveiled distinctive target segments within the market. Clusters based on demographic and behavioral attributes illustrated diverse consumer profiles. The linear regression model, while not yielding a clear linear trend, highlighted the complexity of pricing dynamics influenced by multiple factors. These target segments provide a roadmap for stakeholders, enabling them to tailor strategies for different consumer profiles, ensuring resonance and relevance.

As the electric vehicle market continues to evolve, this project provides a solid foundation for stakeholders to navigate the road ahead. The integration of demographic, behavioral, and target segment insights positions industry players to not only understand current market dynamics but also to proactively shape the future of electric mobility in India. The journey towards sustainable, consumer-centric electric transportation is now charted with clarity and precision.

Future Scope and Enhancements

- We can enhance our analysis by integrating smart grid data, gaining valuable insights into the impact on electric vehicle adoption. This integration allows us to understand the availability and efficiency of charging infrastructure, optimizing our strategies accordingly.

- By implementing predictive analytics, we can proactively forecast future consumer behavior. Leveraging machine learning models enables us to anticipate shifts in preferences and market dynamics, providing a strategic advantage in adapting our approach.

- We can conduct temporal trend analysis to identify changing patterns over time. By uncovering seasonality effects and understanding evolving consumer interests, we can adapt our strategies accordingly, staying ahead of market trends.

- Applying NLP techniques enables us to analyze online reviews, forums, and social media discussions. This allows us to extract qualitative insights and sentiments related to electric vehicles, providing a deeper understanding of consumer perceptions.

- We can implement real-time monitoring of charging infrastructure availability, reliability, and expansion. This approach allows us to align our marketing strategies with the dynamic charging infrastructure landscape, ensuring timely and informed decision-making.

- Integrating economic indicators into our analysis, such as interest rates and unemployment rates, provides insights into their influence on consumer purchasing power and electric vehicle adoption.

- Building interactive dashboards for stakeholders enables us to explore and visualize data trends in real-time. This enhances accessibility and usability, empowering decision-makers with timely and actionable insights.

- We can explore additional ML models like SVM and Decision Tree Classifier to uncover intricate patterns and enhance analysis accuracy.

Code links:

https://github.com/monalisaburma/EV_Market_Segment/blob/main/Electric_car.ipynb

https://github.com/monalisaburma/EV_Market_Segment/blob/main/Indian_behaviour.ipynb

Github link:

https://github.com/monalisaburma/EV_Market_Segment/tree/main