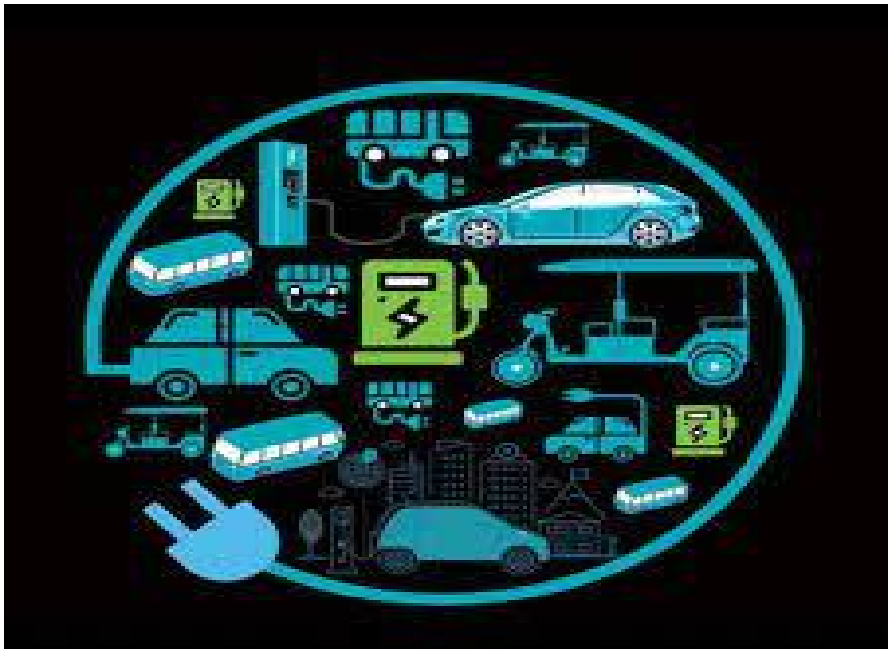
-Market Segmentation

Market segmentation for EV



Group Manikandan C

INTRODUCTION

An electric vehicle (EV) is a vehicle that uses one or more electric motors for propulsion.It can be powered by a collector system, with electricity from extravehicular sources, or it can be powered autonomously by a battery or other energy generating devices such as solar panels, dynamos.

EV vehicles were introduced as early as the 1900's along with steam powered vehicles, but were not fully developed and operational because there were shortcomings in the technology required to develop EVs

But after the advent of the 21st century and the discovery of Lithium Ion battery car manufacturers began speculating the use of Lithium Ion battery in the manufacturing of EV or electric vehicles and given the recent discovery of Lithium ores in India, it is predicted in the future that India will not only be a leading customer base for EVs but also a major hub for manufacturing them as well.

MARKET OVERVIEW

As stated above the world wants to shift from fossil fuel powered vehicles to electric vehicles and the world requires a proper supply chain for the production of electric vehicles and India is placed at a wonderful place as india can provide a continuous supply of lithium with the discovery of new lithium reserves in various states in india, it is predicted that in the next 10-15 years india will be a global hub for EV manufacturing and given the population could hold the majority market share for EVs.

The EV market in india is valued at over 114BN USD, many e-commerce companies such as Amazon, Swiggy etc.., have been pledging to use EV as their delivery system and India is also experimenting with the use of EV for public transport and other services.

Here is a Graph showing the rise in EV in India over a period of time.

Market Segmentation:

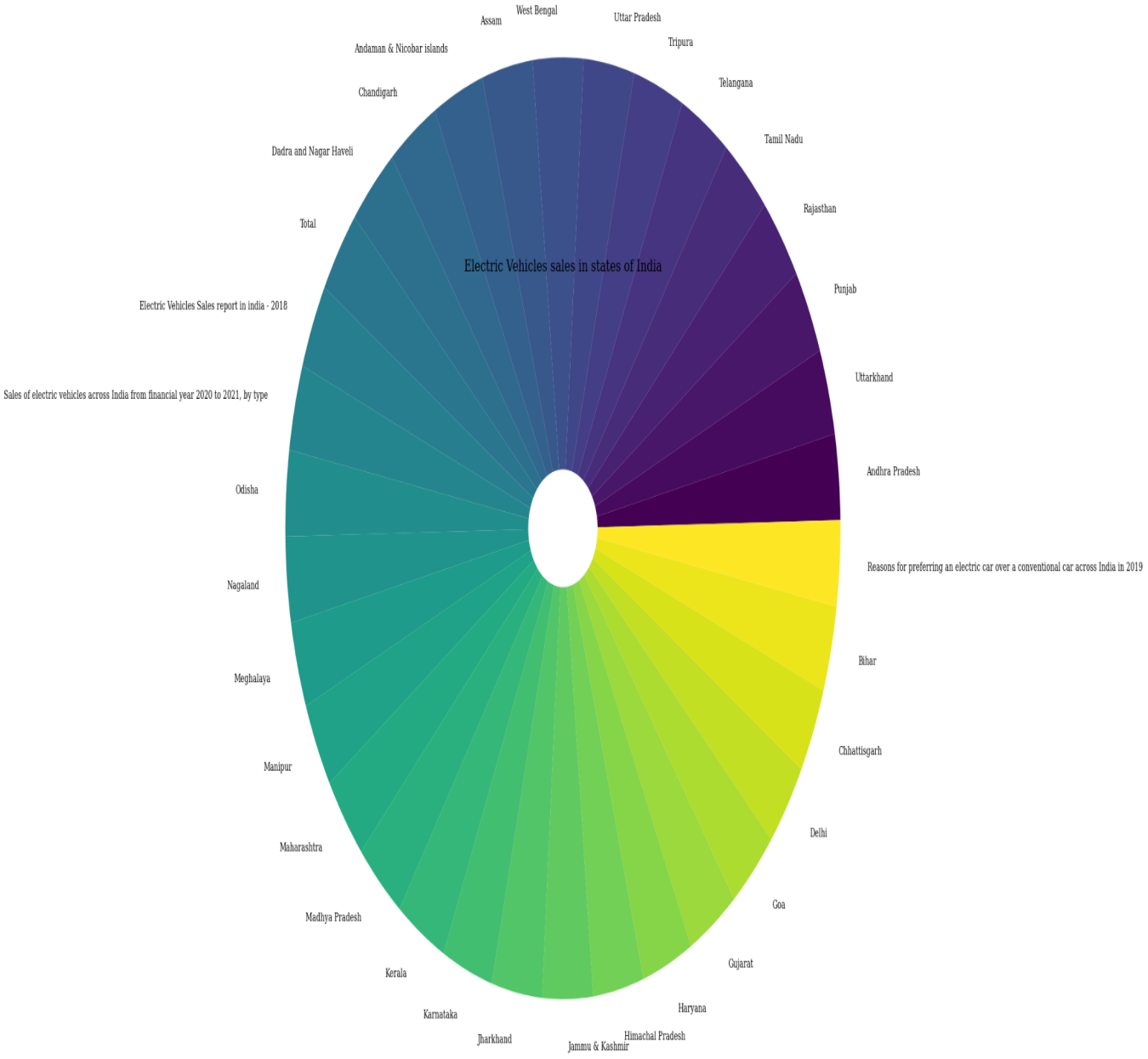
The indian EV sales in india can be segmented by Trustmark Dealer, Dealer, Individual. India is the second most populous country globally, following China, and is actively pursuing the electrification of buses to reduce greenhouse gas emissions. Similar to China, which boasts the largest electric bus fleet in the world, India is making significant efforts to adopt electric buses. Numerous state governments have already begun procuring electric buses from both Chinese and local electric bus manufacturers.

The Indian government recognizes the importance of controlling greenhouse gas emissions, particularly from vehicles, and is actively promoting the use of electric-powered vehicles throughout the country. This push has led to an increased demand for electric buses in India. Factors driving this demand include the growth of domestic manufacturing, rapid urbanization, and a rising environmental awareness among the population.

Overall, the Indian market for electric buses is experiencing substantial growth due to these factors and the country's commitment to reducing emissions and promoting sustainable transportation options.

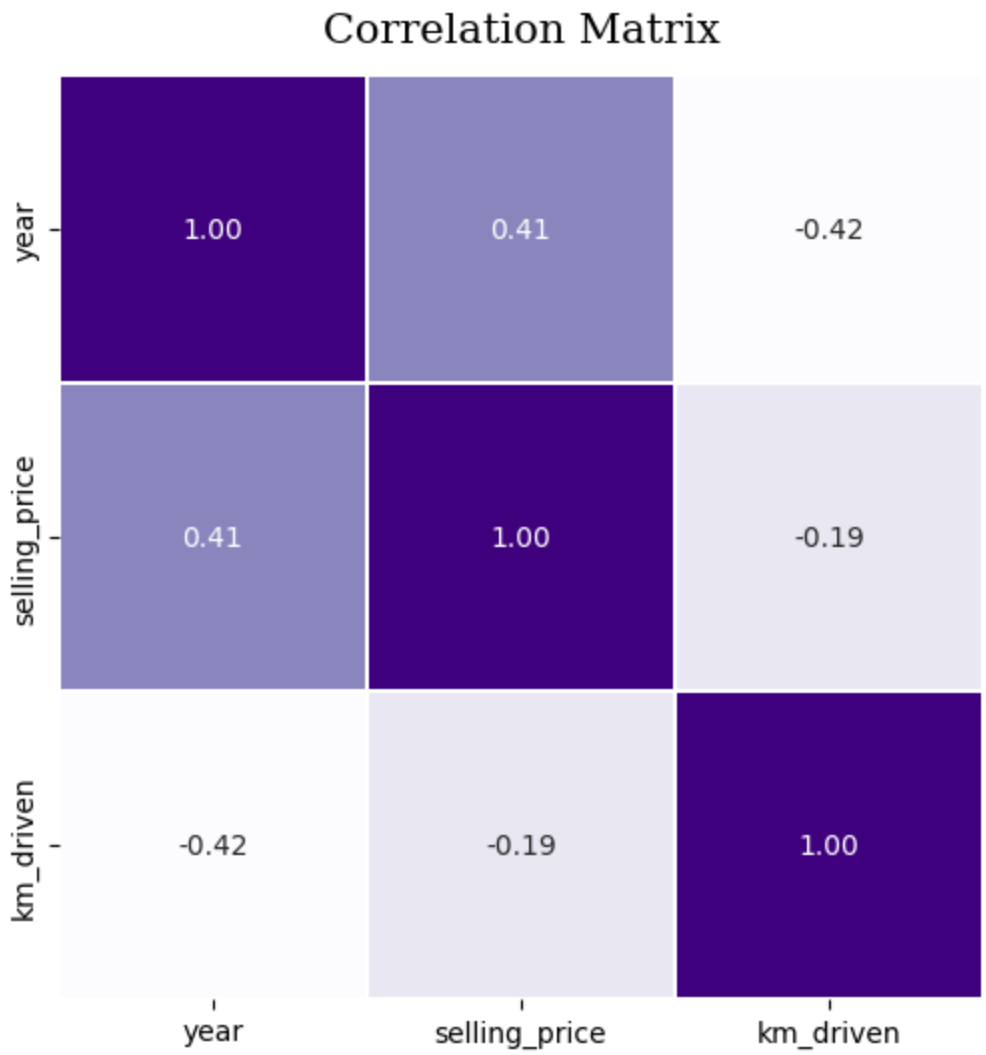
State wise distribution

The following Pie chart shows the distribution of electric vehicles in different states across India, it shows that EVs are equally distributed all across India specifically in the North eastern regions of India. It also has a large market base in Urban areas all across India, but in the future it could play a vital role in Public transport and Agriculture field etc..


Electric Vehicles sales in states of India
Electric Vehicles Sales report in india - 2018
Sales of electric vehicles across India from financial year 2020 to 2021, by type
Reasons for preferring an electric car over a conventional car across India in 2019
West Bengal
Assam
Uttar Pradesh
Tripura
Andaman & Nicobar islands
Telangana
Chandigarh
Tamil Nadu
Dadra and Nagar Haveli
Rajasthan
Total
Punjab
Uttarkhand
Andhra Pradesh
Odisha
Nagaland
Bihar
Meghalaya
Chhattisgarh
Manipur
Delhi
Maharashtra
Goa
Madhya Pradesh
Gujarat
Kerala
Haryana
Karnataka
Himachal Pradesh
Jharkhand
Jammu & Kashmir

Comparison Data

The below Matrix shows the correlation of the data variables which shows the different relationships between the increase in the selling_price as the years passed by along with the total journey of the electric vehicles.



K-Means Clustering:

Clustering is a fundamental technique in data science and machine learning that involves grouping similar data points together based on their characteristics or patterns. It is an unsupervised learning method, meaning it doesn't rely on predefined labels or target variables.

The goal of clustering is to identify inherent structures or natural groupings within the data, allowing for meaningful insights and analysis. It helps in understanding the underlying patterns, similarities, and differences among data points, which can be valuable for various applications such as customer segmentation, anomaly detection, image segmentation, document clustering, and more.

K Means Algorithm:

K-means is one of the most popular and widely used clustering algorithms in data science. It aims to partition a dataset into K clusters based on similarity or proximity of data points. The algorithm assumes that the number of clusters (K) is known in advance.

Here's a step-by-step explanation of the K-means algorithm:

1. **Initialization**: Randomly select K points from the dataset as initial centroids. These centroids will serve as the centers of the initial clusters.

2. **Assignment**: Assign each data point to the nearest centroid based on a distance metric, typically Euclidean distance. This step forms K clusters, with each data point belonging to the cluster associated with the nearest centroid.

3. **Update**: Recalculate the centroids of the K clusters based on the current assignments. The centroid is the mean of all the data points assigned to a particular cluster.

4. **Iteration**: Repeat steps 2 and 3 until convergence. Convergence occurs when either the centroids remain unchanged or the change falls below a predefined threshold. The algorithm iteratively refines the cluster assignments and updates the centroids to find the optimal cluster configuration.

5. **Final Result**: Once convergence is reached, the algorithm outputs the final cluster assignments and the corresponding centroids. Each data point belongs to the cluster associated with the nearest centroid.

The K-means algorithm aims to minimize the within-cluster sum of squares, also known as the inertia or distortion. It seeks to find centroids that minimize the distance between the data points within each cluster. This objective function is optimized during the iterative process of updating the assignments and centroids.

K-means has several advantages, including simplicity, scalability, and efficiency. However, it also has some limitations. It assumes that the clusters have a spherical shape and are of equal size, which might not always hold true for complex datasets. Moreover, the algorithm's performance can be sensitive to the initial centroid selection, and it may converge to a suboptimal solution.

To address some limitations, techniques like K-means++ initialization and the use of different distance metrics can be employed. Additionally, it's important to evaluate and interpret the clustering results carefully, considering domain knowledge and other validation metrics.

Overall, K-means is a powerful and widely used clustering algorithm for partitioning

data into groups based on similarity. It has applications in various fields, including customer segmentation, image compression, document clustering, and anomaly detection.

Applications of K=Means Clustering:

Customer Segmentation: K-means clustering can be used to segment customers based on their purchasing behavior, demographics, or other relevant features. This information can help businesses target specific customer groups with tailored marketing strategies.

Recommender Systems: K-means clustering can be used to create user segments or item clusters in recommender systems. This enables personalized recommendations by identifying similar users or items and suggesting items that are popular among the same cluster.

Packages/ Tools used:

1.Google Colab : google colab is a programming tool used for building the Machine Learning model.

2.Pandas : Pandas library is used for uploading and manipulating the data and performing Data Analysis.

3.Numpy : Numpy library is used for doing Numerical Operations.

4.Matplotlib, Seaborn : Matplotlib and Seaborn libraries are used for Visualizing the data.

5. Sklearn : Scikit learn library is used for the Machine Learning.

Data Pre-processing:

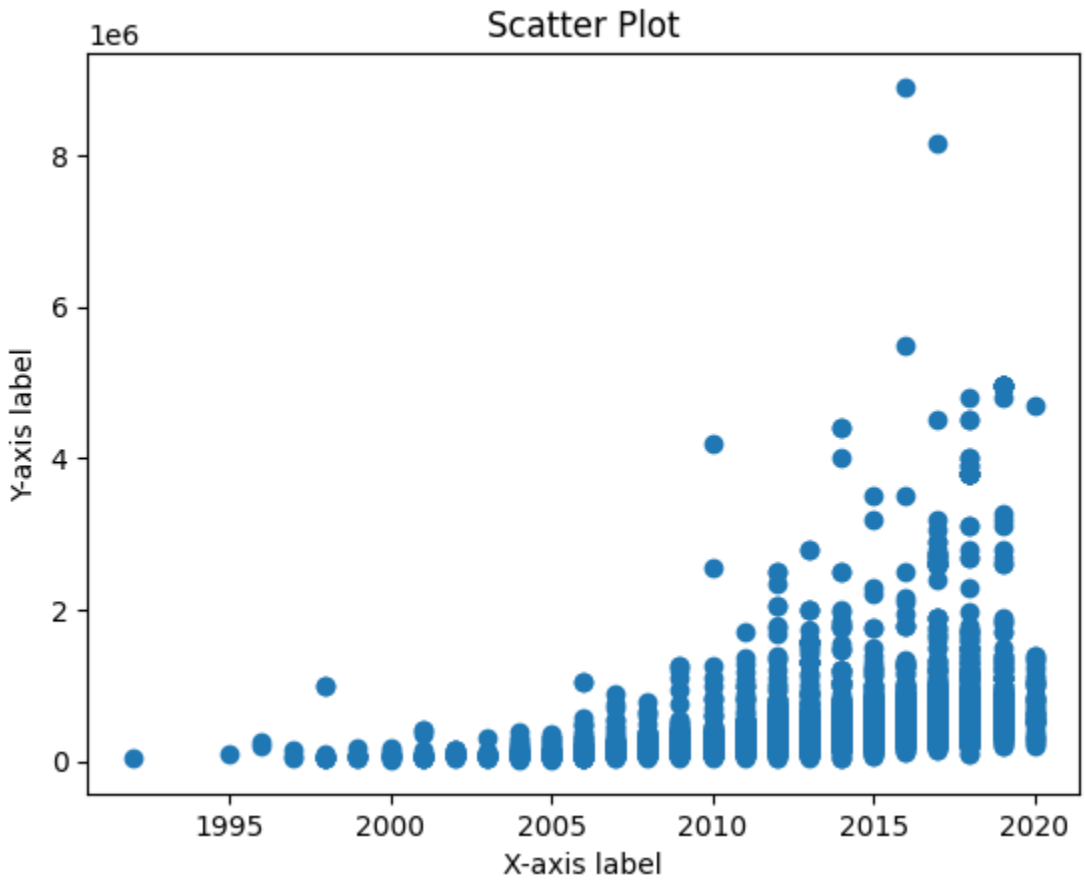
```
df = pd.read_csv('CAR DETAILS FROM CAR DEKHO.csv')

df.head()
```

```
df.info()
```

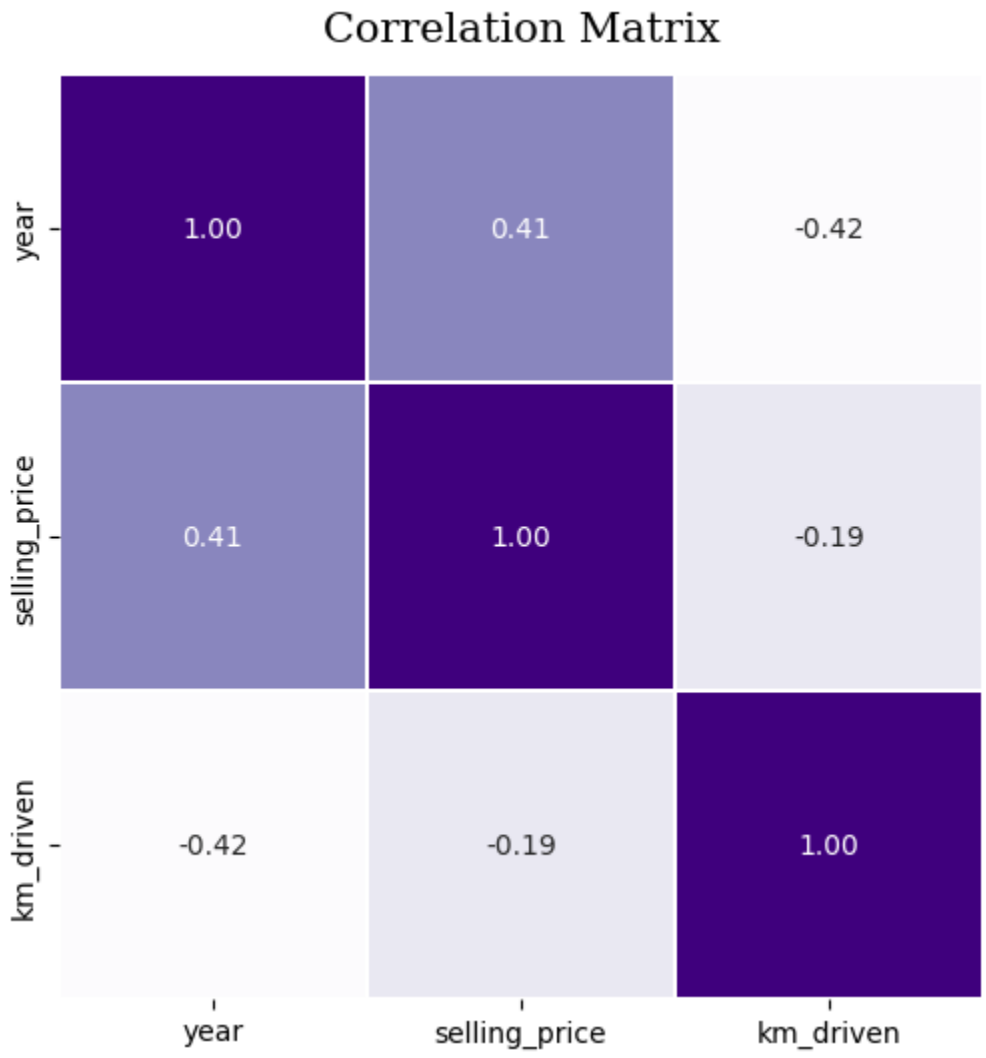
After the data is pre processed we can perform the EDA or where we can get the information of the data in the data set.

Then we visualize the data using Matplotlib.pyplot

```
plt.scatter(df['year'], df['selling_price'])

plt.xlabel('X-axis label')

plt.ylabel('Y-axis label')

plt.title('Scatter Plot')

plt.show()
```



```
plt.figure(figsize=(6,6))

sns.heatmap(data=df.corr(), annot=True, cmap='Purples', cbar=False,
square=True, fmt='.2f', linewidths=.3)

plt.title('Correlation Matrix', family='serif', size=15, pad=12);
```



The above code gives the correlation or the relationship between the three different parameters in the data set.

```
X = df[['selling_price']]
```

We select the target variable or parameter on which we will be doing market segmentation.

```
k = 3  # Number of clusters

n_init = 10  # Number of times the k-means algorithm will be run with
different centroid seeds

kmeans = KMeans(n_clusters=k, n_init=n_init)

kmeans.fit(X)

labels = kmeans.labels_

centroids = kmeans.cluster_centers_
```

The above code is derived using the Sklearn library.

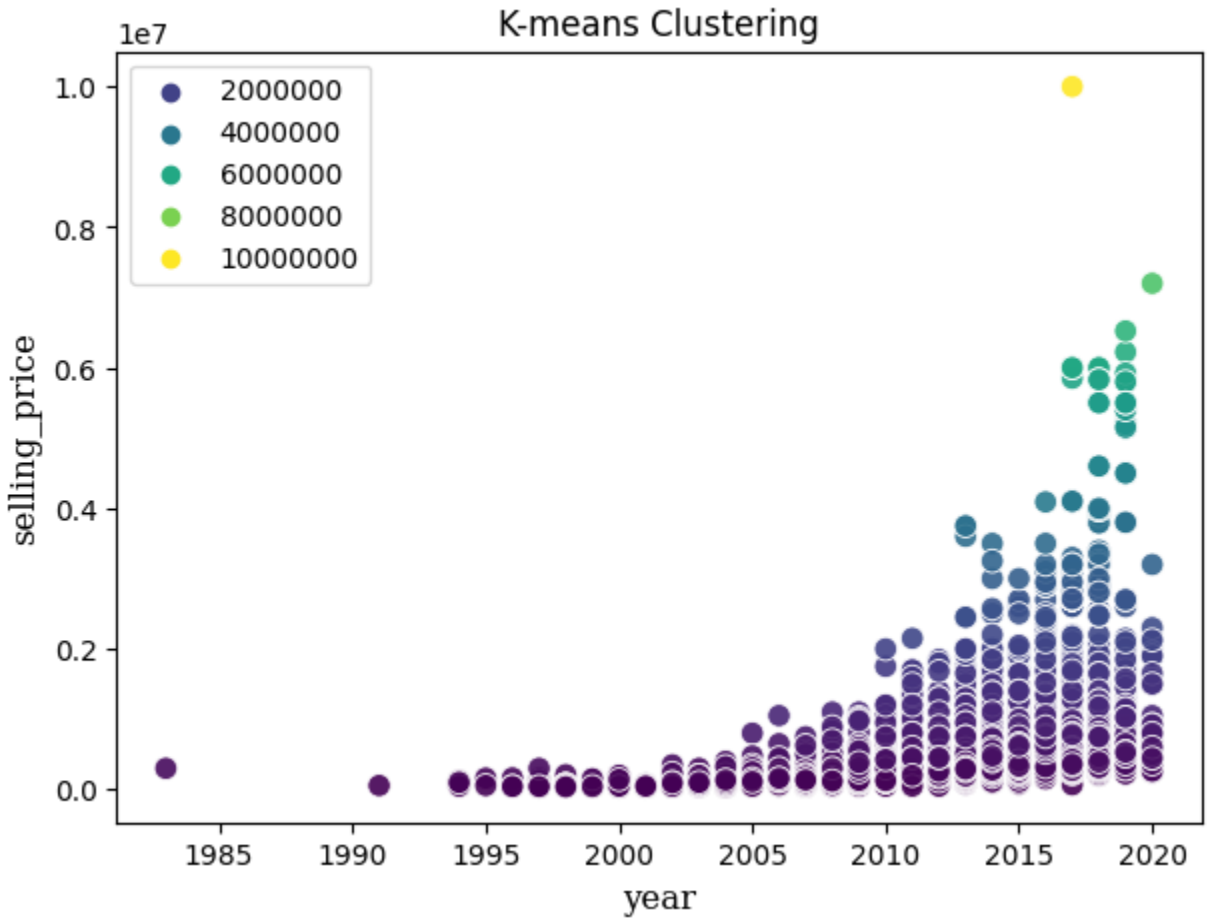
- First the number of clusters are selected based on the divisions.
- Then the centroids are selected and the K means clusters are initiated.
- The model is then fitted since this is an unsupervised ML model only one variable is fitted.
- The centroids are the selected using Kmeans.cluster_centers_

```
print(df['selling_price'].unique())
```

The code above provides the unique values which helps in grouping the variables based on the values.

```
plt.figure(figsize=(7, 5))

sns.scatterplot(data=df, x='year', y='selling price', hue='selling price',
s=70, palette='viridis', zorder=2, alpha=.9)

plt.xlabel('year', family='serif', size=12)

plt.ylabel('selling price', family='serif', size=12)

plt.title('K-means Clustering')

plt.legend()

plt.show()
```

The above code is used to visualize the clusters and then the cluster plot is then visualized and we can see how the demand has increased over the years and the selling_price clusters have increased and is clustered according to prices.



The above plot shows that the market share and price has increased post the 2010s and it has increased over the years post 2010s.

Target Segment:

The target audience for the EV business in India currently are public transport, pay for ride transport vehicles and young generation of riders and millennial and Gen z populations prefer EV, as a result with the increase in youth population EV will have a large market in India.

References:

- https://github.com/Rohit-Rannavre/Feynn-Labs-Internship-2022/tree/main/Project%202.0:%20Market%20Segmentation%20case%20study
- https://drive.google.com/drive/folders/137KIMhwpB1bx5zx0hTaa486bEKe3kXaB
- https://stackoverflow.com/questions/34113790/k-means-clustering-using-sklearn-cluster
- https://www.kaggle.com/code/prashant111/k-means-clustering-with-python
- https://www.kaggle.com/datasets/kkhandekar/electric-vehicles-india
- https://github.com/Chaganti-Reddy/EVMarket-India.git
- https://data.world/datasets/ev
- https://www.statista.com/statistics/1234761/india-electric-vehicle-sales-by-type/
- https://www.statista.com/statistics/1234761/india-electric-vehicle-sales-by-type/
- https://github.com/Mani512996/internship/blob/main/market_segmentation.ipynb
- https://github.com/DevanshG7/DevanshG7.git

Team Members:

- Manikandan C - Team Leader
- Devansh Gupta
- Aryan
- Deepika

Github:

- https://github.com/Mani512996/internship/blob/main/market_segmentation.ipynb
- https://github.com/DevanshG7/DevanshG7.git
- https://colab.research.google.com/drive/1XqUbOLHOnpGKz5Az6m8OtrR-EDDXlMW0?usp=sharing

Conclusion:

In conclusion we can note that with the depleting fossil fuel reserves in the world and the increasing environmental damage it causes there is a huge shift towards the EV market and the result can be seen as the selling price of the EV have increased with increasing demand.