

SEGMENTATION REPORT ON EV MARKET

Dhanraj Kumar

ML Model:

Here, K-means clustering model is used, K-means clustering is an unsupervised machine learning algorithm that partitions a dataset into clusters. It assigns data points to the cluster with the nearest mean value, aiming to create K distinct clusters. It is popular because of its simplicity, scalability, and interpretability. However, there are limitations to consider. One drawback is the need to specify the number of clusters in advance, which may not always be known. The algorithm is also sensitive to the initial centroid selection, potentially leading to suboptimal results. Another limitation is its assumption that clusters are spherical and equally sized, which may not be true for all datasets. Researchers have proposed various extensions and modifications to overcome these limitations.

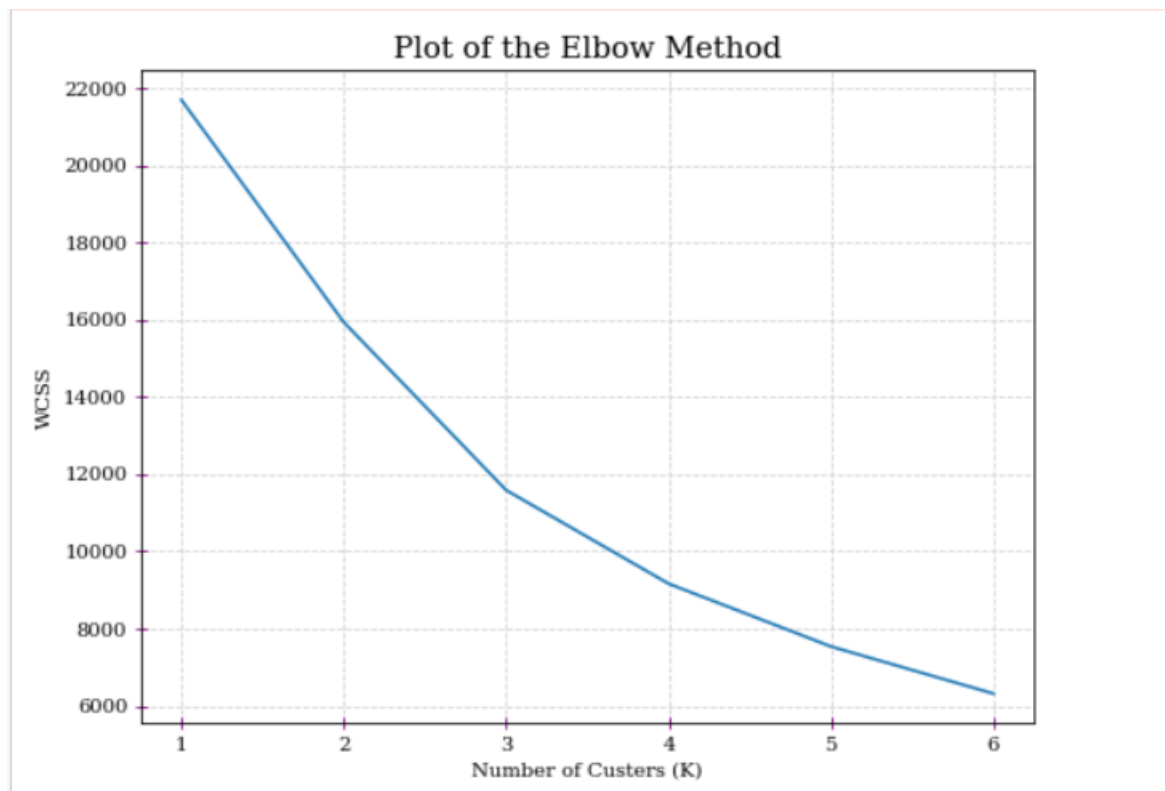
The elbow method is a strategy for determining the best number of clusters in K-means clustering. It entails plotting the sum of squared distances (WCSS) against various values of K and identifying the "elbow" point where the rate of WCSS decrease flattens out. This elbow point indicates a balance between the number of clusters and the quality of clustering, aiding in the selection of the optimal number of clusters. However, the elbow method may not always provide a clear answer, and supplementary evaluation measures or domain expertise may be required for a final decision.

```
df1 = pd.read_csv('CAR DETAILS FROM CAR DEKHO.csv')
df1.head()
```

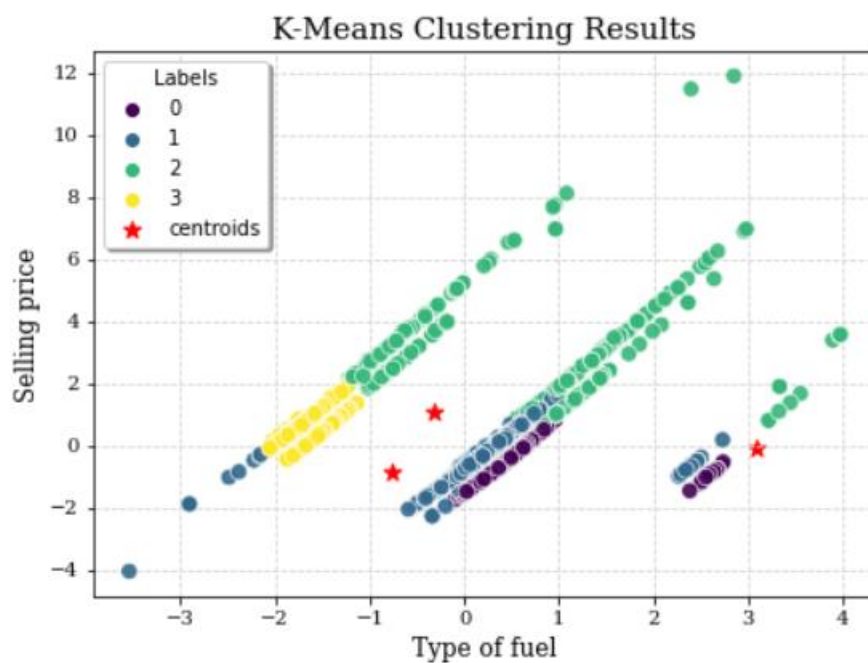
	name	year	selling_price	km_driven	fuel	seller_type	transmission	owner
0	Maruti 800 AC	2007	60000	70000	Petrol	Individual	Manual	First Owner
1	Maruti Wagon R LXI Minor	2007	135000	50000	Petrol	Individual	Manual	First Owner
2	Hyundai Verna 1.6 SX	2012	600000	100000	Diesel	Individual	Manual	First Owner
3	Datsun RediGO T Option	2017	250000	46000	Petrol	Individual	Manual	First Owner
4	Honda Amaze VX i-DTEC	2014	450000	141000	Diesel	Individual	Manual	Second Owner

```
<<< DATASET 2 -----
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4340 entries, 0 to 4339
Data columns (total 8 columns):
#   Column          Non-Null Count  Dtype  
---  -
0   name             4340 non-null   object  
1   year             4340 non-null   int64   
2   selling_price    4340 non-null   int64   
3   km_driven        4340 non-null   int64   
4   fuel             4340 non-null   object  
5   seller_type      4340 non-null   object  
6   transmission     4340 non-null   object  
7   owner            4340 non-null   object  
dtypes: int64(3), object(5)
memory usage: 271.4+ KB
None
```

Elbow method plotting:



Here, we get four clusters, $k = 4$.



In this plot, each point shows different group of customers. The x-axis represents types of fuel used in vehicles and y-axis represents total selling price of vehicle.

Final Conclusion And Insights:

On analysis the three datasets, we can say that number of vehicles sold each year is decrease after 2008 or it may be due to less data availability. It also shows that number of vehicles using diesel and petrol is highest but there is very less number of vehicles which used CNG, LPG, electricity. Vehicles consisting five seats have highest demand relative to other.

Analysis also shows how maximum power consumption of vehicles vary, how mileage of vehicles vary.

Scatter plot also shows groups of customer varying with types of fuel used in vehicle.

Improvements (Given Additional Time and Budget)-

- **Data Collection:** Enhancing data collection efforts by obtaining a broader range of information, including demographics, lifestyle preferences, and psychographic data, is a beneficial approach to enhance the accuracy of segmentation. Conducting surveys, interviews, and leveraging external data sources are effective methods to gather such information. Increasing the quantity of data points enables a more detailed and precise segmentation process, facilitating improved comprehension and targeted marketing towards specific customer segments.
- **Advanced Analytics:** By employing advanced analytics techniques like clustering algorithms, machine learning algorithms, and predictive modeling, it becomes possible to reveal concealed patterns and relationships within the data, thereby enhancing the process of segmentation. Clustering algorithms group similar data points together, while machine learning algorithms and predictive modeling identify crucial variables that set segments apart. Through ongoing refinement and validation of these models, businesses can create more meaningful and focused customer segments.
- **Real time data collection:** To enhance segmentation, it is advantageous to integrate real-time data sources that provide the latest information on customer behavior and preferences. This enables the segmentation process to be more dynamic and responsive, facilitating the development of personalized and targeted marketing strategies. By incorporating real-time data, businesses can ensure that their segmentation efforts are up-to-date and align with current customer insights, leading to more effective customer engagement.
- **Customer Feedback:** Maximize the benefits of customer feedback by gathering their opinions through surveys, interviews, or social media monitoring. This direct input provides valuable insights and perspectives that can refine and enhance segmentation efforts. By incorporating customer feedback into the segmentation process, businesses can gain a deeper understanding of customer preferences, needs, and behaviours, leading to more precise and targeted segmentation strategies.

Top 4 Variables/Features which can be used to create most optimal Market Segments for this Market Domain.

Following are the features which can be used to create optimal Market Segments for Electric Vehicle Market Domain:

- **Price Sensitivity and Affordability:**
Segmenting the EV market based on price sensitivity and affordability enables targeting customers with different budget considerations. Some segments are willing to pay more for advanced features, while others prioritize cost-effectiveness. By understanding the preferences of various segments, businesses can tailor their marketing strategies and offerings to meet the diverse needs and financial capabilities of their target audience.
- **Range and Battery Capacity:**
Segmenting the EV market based on range and battery capacity helps identify diverse customer needs. Some segments prioritize EVs with longer range and faster charging capabilities, while others focus on shorter commutes and lower battery capacity requirements. By understanding these preferences, businesses can tailor their marketing and product offerings to cater to specific segments, providing EV options that align with different customer usage patterns.
- **Charging behavior:**
By examining variables related to charging behavior, such as charging frequency, location preferences (home, workplace, public stations), and preferred charging times, businesses can identify segments with different charging needs and infrastructure requirements. This enables the customization of marketing strategies and infrastructure development to cater to the specific charging preferences of each segment in the EV market.
- **Psychographic Factors:**
By taking into account variables related to attitudes, values, and lifestyles, businesses can gain insights into consumer motivations and preferences, which is known as psychographic factors. For example, segments in the EV market that prioritize sustainability, innovation, or cost savings may have unique characteristics and preferences. Understanding these psychographic factors allows businesses to customize their marketing strategies and product offerings to align with the specific values and interests of each segment, leading to increased customer engagement and satisfaction.

[Segmentation Github Link](#)