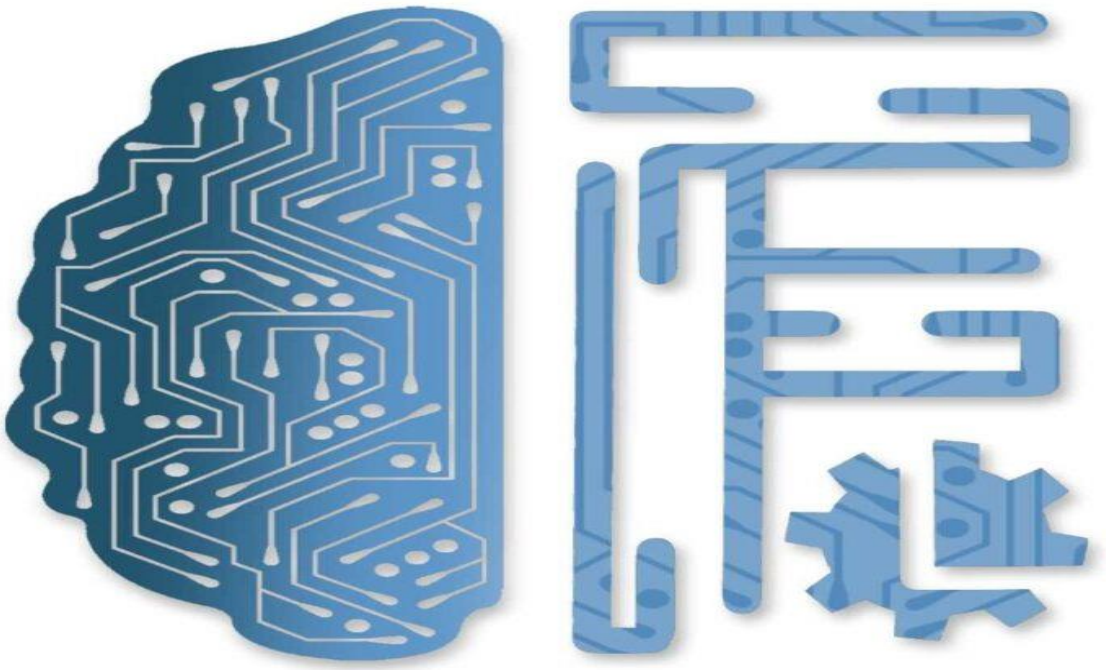


# ELECTRIC VEHICLE MARKET SEGMENTATION

Submitted by

**Kamatam Manoj**

On

**11-07-2023**

## Abstract:

The electric vehicle market is rapidly growing, and with the increasing number of electric vehicle models available, it has become essential for companies to identify and target specific customer segments. Market segmentation involves dividing a larger market into smaller groups of consumers with similar needs and characteristics. This approach helps companies to develop targeted marketing strategies and tailor their product offerings to meet the specific needs of each customer group. In the electric vehicle market, segmentation is typically based on factors such as customer demographics, driving habits, and charging infrastructure availability. By understanding the different segments within the electric vehicle market, companies can better position their products and services to meet the unique needs and preferences of each group, ultimately driving sales and market share.


## Problem Statement :

**(EV Market)**You are a team working under an Electric Vehicle Startup. The Startup is still deciding in which vehicle/customer space it will be develop its EVs. You have to analyse the Electric Vehicle market in India using Segmentation analysis and come up with a feasible strategy to enter the market, targeting the segments most likely to use Electric vehicles.

## Over view :

The India electric vehicle market size was valued at USD 220.1 million in 2020 and is expected to expand at a compound annual growth rate (CAGR) of 94.4% from 2021 to 2030. The attractive incentives being offered by the Indian government on the production and purchase of electric vehicles to encourage the adoption of electric vehicles are anticipated to drive the growth of the market over the forecast period. The outbreak of the COVID19 pandemic triggered a significant decline in the overall sales of passenger and commercial vehicles in 2020. However, the sales of electric vehicles in India remained unaffected. The post-lockdown sale of pure and hybrid electric vehicles is a prominent driving factor for the electric vehicle market in India. The stringent Greenhouse gas (GHG) emission norms drafted by the government, such as the Bharat Stage (BS) VI emission standards introduced by India's Ministry of Road Transport and Highways (MoRTH), are also expected to play a decisive role in driving the growth of the market. The increasing prices of conventional fuel are

expected to accentuate the development of vehicle electrification. The stringent emission norms being drafted by the government and the growing environmental awareness among Indian consumers are also expected to fuel the demand for electric vehicles. Furthermore, Indian automakers, such as Tata Motors, and Mahindra and Mahindra Ltd., have embarked upon aggressive efforts to add electrified vehicles to their product portfolio, which is expected to encourage Indian consumers to opt for electric vehicles. All these factors bode well for the growth of the electric vehicle market in India over the forecast period. The EV market in India comprised only two electric vehicle models in 2019. As a result, only 0.15% of the new passenger cars registered between April 2019 and March 2020 were BEVs. However, at the beginning of 2021, the India electric vehicle (EV) market consisted of around eight electric vehicle models, thereby offering more options for Indian consumers looking forward to buying electric vehicles. Moreover, the prices of electric vehicles are also expected to decline over the forecast period, thereby allowing EVs to provide a lower Total Cost of Ownership (TCO) as compared to conventional vehicles. This is expected to pave the way for the massmarket penetration of electric vehicles.

According to The Hindu News Report: India’s electric vehicle market grew 223% in 2022, adding around 48,000 EVs, according to market analyst firm Canalys. Tata leads in EVs with 86% of the market share from just two models: the Nexon EV and Tigor EV.

## Introduction :

Electric vehicles (EVs) have gained significant traction in recent years as a cleaner and more efficient alternative to traditional gasoline-powered cars. With advancements in battery technology, a growing network of charging infrastructure, and increasing consumer demand, EVs have become a viable option for many drivers around the world. Changing the way we think about driving, these vehicles are powered by electricity stored in a battery, rather than gasoline, and use electric motors to turn the wheels. India is the third largest automobile market globally in terms of sales, ahead of Germany and Japan. There is now a push for manufacturers and policymakers to collaborate to shift demand towards greener options. The automotive sector is a major contributor to India's economy, accounting for 7.1 percent of its GDP and providing significant employment. The Economic Survey 2023 predicts that

India's domestic electric vehicle market will see a 49 percent compound annual growth rate (CAGR) between 2022 and 2030, with 10 million annual sales by 2030. Additionally, the electric vehicle industry is projected to create around 50 million direct and indirect jobs by 2030. The Indian government has set a target to achieve 30 percent electrification of the country's vehicle fleet by 2030, and has introduced several incentives and policies to support the growth of the EV industry. The industry was given a major boost in the FY24 Union Budget for the production of electric vehicles, adoption of hydrogen fuel, and embracing changing technologies. In the 2023-24 Union Budget, Finance Minister Nirmala Sitharaman announced a budget allocation of INR 35,000 crore for crucial capital investments aimed at achieving energy transition and net-zero targets by 2070. Furthermore, she stated that the government will support Battery Energy Storage Systems with a capacity of 4,000 MWH through viability gap funding. For electric vehicle manufacturers, the government has already launched initiatives such as the Faster Adoption of Manufacturing of Electric Vehicles Scheme – II (FAME – II) and the Production Linked Incentive Scheme (PLI). The Budget has allocated INR 51.72 billion (approximately $ 631 million) towards its FAME-II scheme to subsidize and promote the adoption of clean energy vehicles. This represents an 80 percent increase in budget allocation from previous years. The reduced custom duty on Lithium-iOn batteries used in electric vehicles and excise duty exemptions on natural gas and biogas could result in more foreign electric vehicles being imported to India. The charging infrastructure is being expanded with investments from both government and private companies in setting up charging stations. The nation's first EV charging plaza was established by EESL in July 2020, and in just one year the number of charging stations has multiplied over five times. The increasing popularity of EVs in the national capital highlights the success of the Delhi EV Policy launched in 2020. In December 2022, EVs accounted for 16.8 percent of all vehicle sales in Delhi, marking a YoY growth of 86 percent. India's electric vehicle market is poised for significant growth in the coming years. With supportive government policies, increasing consumer awareness, and advancements in technology, the country is well positioned to transition towards a more sustainable and eco-friendly mode of transportation. As the demand for EVs increases, it presents a tremendous opportunity for both local and international companies to invest in and contribute to the growth of India's EV ecosystem.

**Customer Segmentation with Machine Learning :** Customer segmentation enables a company to customize its relationships with the customers, as we do in our daily lives. When you perform customer segmentation, you find similar characteristics in each customer's behaviour and needs. Then, those are generalized into groups to satisfy demands with various strategies. Moreover, those strategies can be an input of the

- Targeted marketing activities to specific groups
- Launch of features aligning with the customer demand
- Development of the product roadmap

There are different products/solutions available in the market from packaged software to CRM products. Today, I will apply an unsupervised machine learning algorithm with Python. I will apply K - Means clustering to the dataset with the following steps.

- Data Preparation
- Segmentation with K-means Clustering
- Hyperparameter Tuning
- Visualization and Interpretation of the Results

**Data Preprocessing :** Data preprocessing is an important step in the data mining process. It refers to the cleaning, transforming, and integrating of data in order to make it ready for analysis. The goal of data preprocessing is to improve the quality of the data and to make it more suitable for the specific data mining task. Data preprocessing transforms the data into a format that is more easily and effectively processed in data mining, machine learning and other data science tasks. The techniques are generally used at the earliest stages of the machine learning and AI development pipeline to ensure accurate results. There are several different tools and methods used for preprocessing data, including the following:

- sampling, which selects a representative subset from a large population of data;
- transformation, which manipulates raw data to produce a single input;
- denoising, which removes noise from data;
- imputation, which synthesizes statistically relevant data for missing values;

- normalization, which organizes data for more efficient access; and
- feature extraction, which pulls out a relevant feature subset that is significant in a particular context.

## Electric Vehicle Market Segmentation Analysis

In [105]:
```
# importing all the required in-built libraries

import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.preprocessing import LabelEncoder
```

In [106]:
```
# Loading dataset

df_Org = pd.read_csv(r"C:\Users\Manoj\Desktop\FynnL\Indian automoble buying behavour study 1.0.csv")

pd.set_option('display.float_format', lambda x: '%.3f' % x)

df = df_Org.copy()

df.head()

df.columns

df.shape

df.head()

df['Make'].unique()
```

Out[106]:
```
array(['i20', 'Ciaz', 'Duster', 'City', 'SUV', 'Baleno', 'Verna',
       'Luxuray', 'Creata'], dtype=object)
```

### Data Preprocessing

In [107]:
```
# Data preprocessing

df.isnull().sum()
```

### Data Preprocessing

In [107]:
```
# Data preprocessing

df.isnull().sum()

df.duplicated()
```

Out[107]:
```
0     False
1     False
2     False
3     False
4     False
      ...
94    False
95    False
96    False
97    False
98    False
Length: 99, dtype: bool
```

In [108]:
```
df['Age'].unique()
```

Out[108]:
```
array([27, 35, 45, 41, 31, 28, 33, 34, 29, 30, 49, 26, 37, 36, 43, 42, 32,
       44, 39, 46, 50, 51], dtype=int64)
```

In [109]:
```
df.describe()
```

Out[109]:

| | Age | No of Dependents | Salary | Wife Salary | Total Salary | Price |
|---|---|---|---|---|---|---|
| count | 99.000 | 99.000 | 99.000 | 99.000 | 99.000 | 99.000 |
| mean | 36.313 | 2.182 | 1736363.636 | 534343.434 | 2270707.071 | 1194040.404 |
| std | 6.246 | 1.335 | 673621.729 | 605444.956 | 1050777.411 | 437695.540 |
| min | 26.000 | 0.000 | 200000.000 | 0.000 | 200000.000 | 110000.000 |
| 25% | 31.000 | 2.000 | 1300000.000 | 0.000 | 1550000.000 | 800000.000 |
| 50% | 36.000 | 2.000 | 1600000.000 | 500000.000 | 2100000.000 | 1200000.000 |
| 75% | 41.000 | 3.000 | 2200000.000 | 900000.000 | 2700000.000 | 1500000.000 |

# Visualization :

## Profiling Describing Potential Segments

FeynnLabs

## Education (Graduate vs Post Graduate)

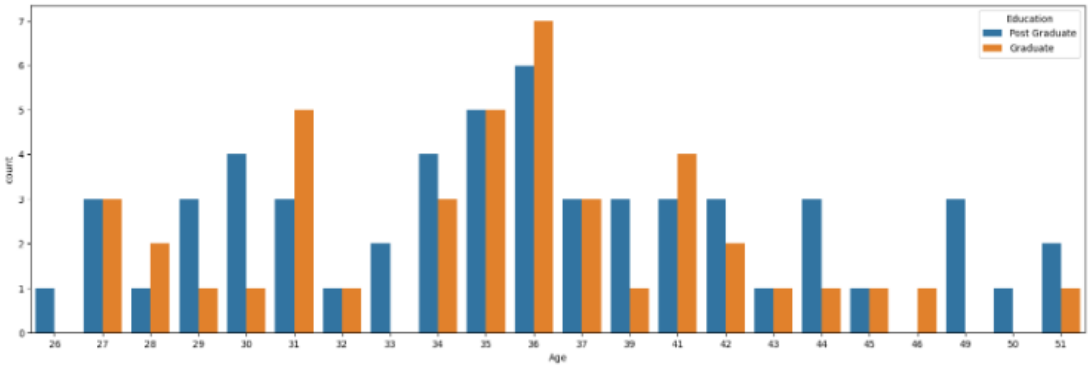
```
In [4]: plt.figure(figsize=(20,6))
        sns.countplot(x="Education", data=df, hue="Education")
Out[4]: <Axes: xlabel='Education', ylabel='count'>
```


## Marital status (Single vs Married)

```
In [5]: plt.figure(figsize=(20,6))
        sns.countplot(x="Marrital Status", data=df, hue="Marrital Status")
Out[5]: <Axes: xlabel='Marrital Status', ylabel='count'>
```


## Education with Age

```
In [73]: plt.figure(figsize=(20,6))
         sns.countplot(x="Age", data=df, hue="Education")
Out[73]: <Axes: xlabel='Age', ylabel='count'>
```




## Segmentation with K-means Clustering :

We are going to use K - means algorithm from scikit - learn. Let's first understand how the algorithm will form customer groups: Initialize k=n centroids=number-of-clusters randomly or smartly Assign each data point to the closest centroid based on Euclidian distance, thus forming the groups Move centres to the average of all points in the cluster While running the steps

through, the algorithm checks the sum of squared distances between clustered-point and centre for each cluster. Mathematically speaking, it tries to minimize — optimize the within-cluster sum-of-squared-distances or inertia of each cluster.

$$\sum_{i=0}^{n} \min_{\mu_j \in C} (||x_j - \mu_i||^2)$$

Mathematical expression of within-cluster sum-of-squared distances or inertia where X is the points in the cluster and μ is the current centroid When inertia value does not minimize further, algorithm converges. Thus, iteration stops. from sklearn.cluster import Kmeans kmeans_model = KMeans(init='k-means++', max_iter=500, random_state=42) init parameter with the k-means++ allows the algorithm to place initial centers smartly, rather than random. max_iter is the maximum number of iterations of the algorithm in a single run, default value is 300. random_state guarantees the reproducibility of the model results. This algorithm is easy to understand, fits well to large datasets in terms of computing times and guarantees convergence. However, when centroids are initialized randomly, algorithm may not assign the points to the groups in the most optimal way. One important consideration is the selection of k. In other words, how many groups should be formed? For example, Kmeans applied above uses k=8 as a default value

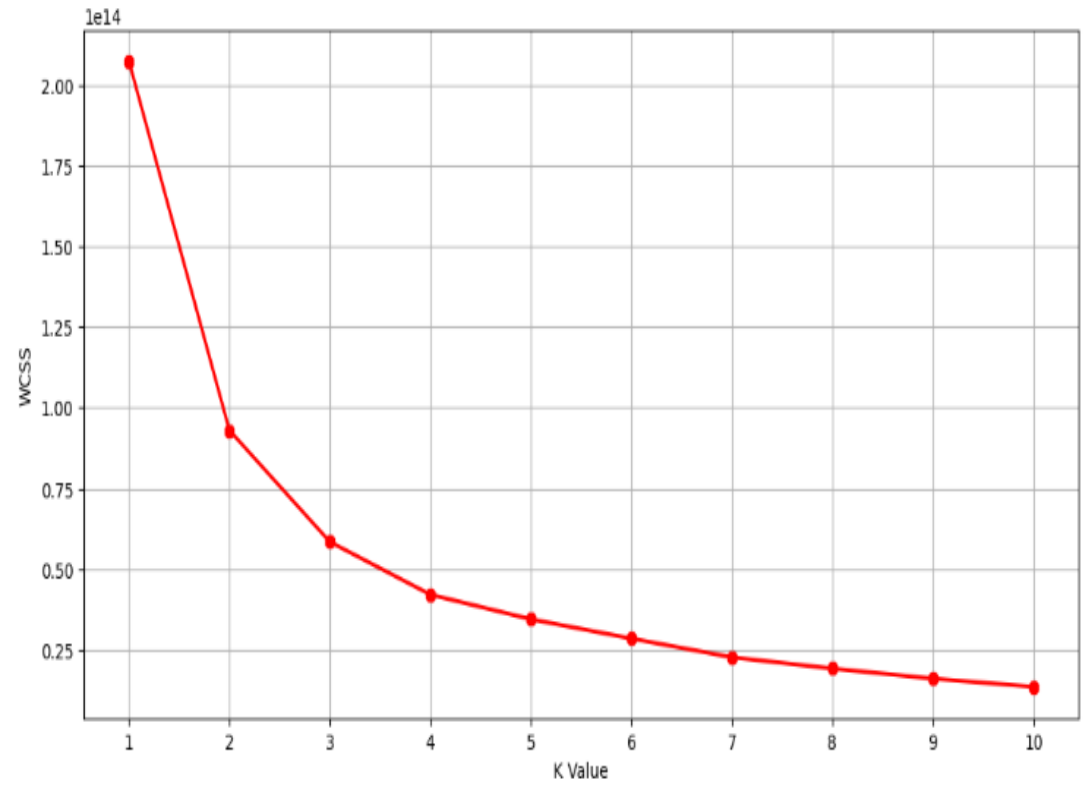
## Hyperparameter Tuning :

While selecting k, we are going to decide against the optimization criteria of the K-means, inertia, using elbow method. We are going to build different Kmeans models with k values 1 to 15, and save the corresponding inertia values. With the elbow method, we are going to select the k value where the decrease in the inertia stabilizes. When k=1 inertia is at the highest, meaning data is not grouped yet. Inertia decreases steeply until k=2. Between k=2 and 4, the curve continues to decrease fast. At k=4, the descent stabilizes and continues linearly afterwards, forming an elbow at k=4. This points out the optimal number of customer group is 4.

## K Means Model

In [79]:
```
# K-Means Model

from sklearn.cluster import KMeans
wcss = []
for k in range(1,11):
    kmeans = KMeans(n_clusters=k, init="k-means++",random_state=28)
    kmeans.fit(X)
    wcss.append(kmeans.inertia_)
plt.figure(figsize=(12,6))
plt.grid()
plt.plot(range(1,11),wcss, linewidth=2, color="red", marker ="8")
plt.xlabel("K Value")
plt.xticks(np.arange(1,11,1))
plt.ylabel("WCSS")
plt.show()
```




## Visualization and Interpretation of the Results :

Let’s plug in the k=4 to K-means and visualize how customer groups are created:

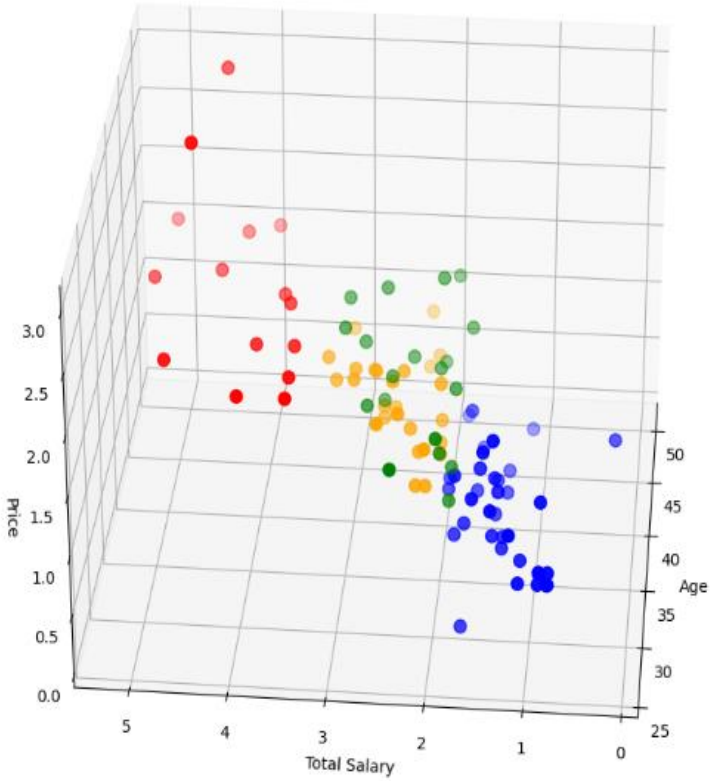
```
In [80]: km = KMeans(n_clusters=4, random_state=28)
clusters = km.fit_predict(df)
df["Cluster"] = clusters

df_Org["Cluster"] = clusters

from mpl_toolkits.mplot3d import Axes3D
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd

fig = plt.figure(figsize=(20,10))
ax = fig.add_subplot(111, projection='3d')
ax.scatter(df.Age[df.Cluster == 0], df["Total Salary"][df.Cluster == 0], df["Price"][df.Cluster == 0], c='blue', s=60)
ax.scatter(df.Age[df.Cluster == 1], df["Total Salary"][df.Cluster == 1], df["Price"][df.Cluster == 1], c='red', s=60)
ax.scatter(df.Age[df.Cluster == 2], df["Total Salary"][df.Cluster == 2], df["Price"][df.Cluster == 2], c='green', s=60)
ax.scatter(df.Age[df.Cluster == 3], df["Total Salary"][df.Cluster == 3], df["Price"][df.Cluster == 3], c='orange', s=60)

ax.view_init(30, 185)
plt.xlabel("Age")
plt.ylabel("Total Salary")
ax.set_zlabel('Price')
plt.show()
```




## Target Market :

A target market is a group of people that have been identified as the most likely potential customers for a product because of their shared characteristics such as age, income, and lifestyle. Identifying the target market is a key part of

the decision-making process when a company designs, packages, and advertises its product.

## Demographic:

These are the main characteristics that define your target market. Everyone can be identified as belonging to a specific age group, income level, gender, occupation, and education level.

```
In [91]: [Cluster_0['Education'].value_counts().head(),
          Cluster_1['Education'].value_counts().head(),
          Cluster_2['Education'].value_counts().head(),
          Cluster_3['Education'].value_counts().head()]

Out[91]: [Post Graduate    21
          Graduate         17
          Name: Education, dtype: int64,
          Post Graduate    10
          Graduate          5
          Name: Education, dtype: int64,
          Post Graduate    11
          Graduate          9
          Name: Education, dtype: int64,
          Post Graduate    14
          Graduate         12
          Name: Education, dtype: int64]
```

Graduates and post graduates both are interested in EV almost equally

```
In [86]: [Cluster_0['Profession'].value_counts().head(),
           Cluster_1['Profession'].value_counts().head(),
           Cluster_2['Profession'].value_counts().head(),
           Cluster_3['Profession'].value_counts().head()]

Out[86]: [Salaried    21
          Business    17
          Name: Profession, dtype: int64,
          Salaried    12
          Business     3
          Name: Profession, dtype: int64,
          Salaried    12
          Business     8
          Name: Profession, dtype: int64,
          Salaried    19
          Business     7
          Name: Profession, dtype: int64]
```

Salaried people are more interested on electric vehicles

```
In [85]: [Cluster_0['Age'].value_counts().head(3),
           Cluster_1['Age'].value_counts().head(3),
           Cluster_2['Age'].value_counts().head(3),
           Cluster_3['Age'].value_counts().head(3)]

Out[85]: [27    5
          35    5
          31    4
          Name: Age, dtype: int64,
          49    2
          34    2
          44    2
          Name: Age, dtype: int64,
          42    3
          34    2
          44    2
          Name: Age, dtype: int64,
          36    7
          35    5
          41    3
          Name: Age, dtype: int64]
```

Based on Cluster 0, 2 and 3 Age groups of 27-44 are targeted

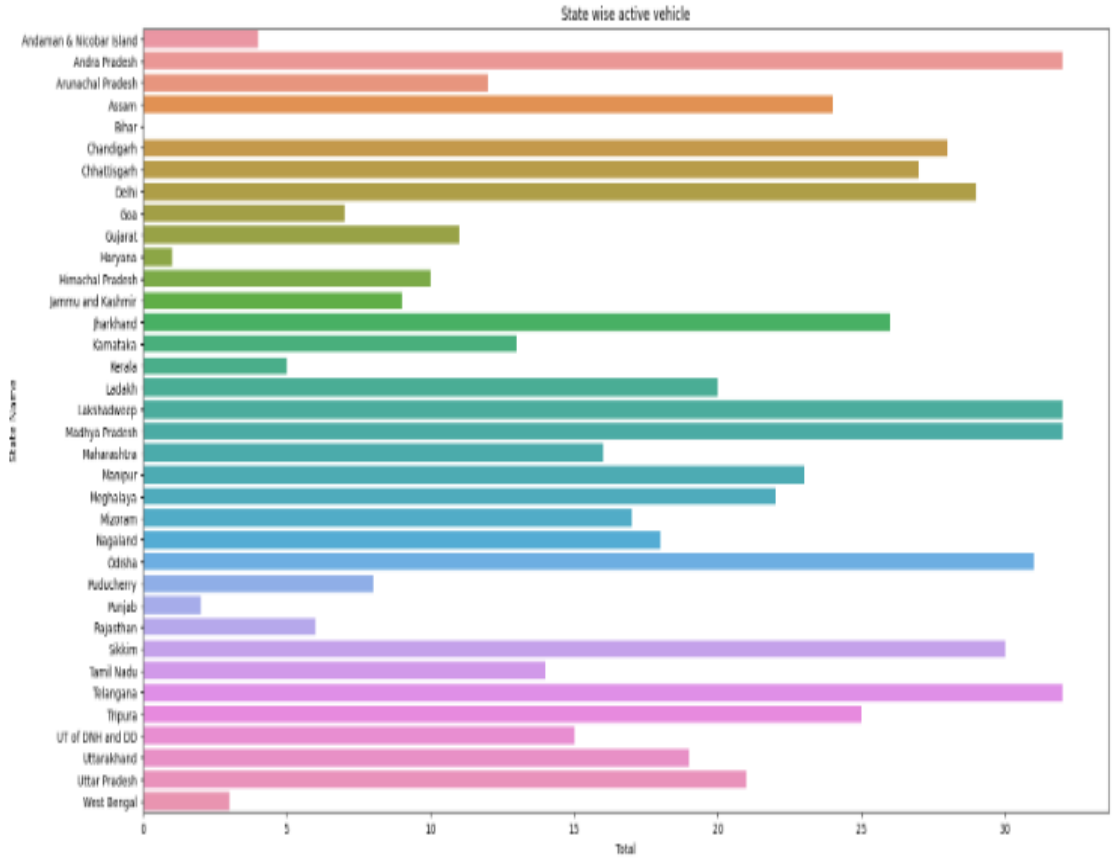
## Geographic:

This segment is increasingly relevant in the era of globalization. Regional preferences need to be taken into account.

## Visualization

In [99]:
```
plt.rcParams['figure.figsize'] = (20, 10)
f = sns.barplot(y=Active_EV['State Name'], x=Active_EV['Total'])
plt.title('State wise active vehicle')
plt.show()


plt.rcParams['figure.figsize'] = (20, 10)
f = sns.barplot(y=Active_EV['State Name'], x=Active_EV['Total Electric Vehicle'])
plt.title('State wise active vehicle')
plt.show()
```




**Psychographic:** This segment goes beyond the basics of demographics to consider lifestyle, attitudes, interests, and values.

```
In [94]: [Cluster_0['Price'].value_counts().head(),
          Cluster_1['Price'].value_counts().head(),
          Cluster_2['Price'].value_counts().head(),
          Cluster_3['Price'].value_counts().head()]

Out[94]: [700000     17
          800000     10
          1200000     5
          1100000     3
          110000      1
          Name: Price, dtype: int64,
          1500000    7
          1600000    6
          3000000    2
          Name: Price, dtype: int64,
          1100000    6
          1200000    3
          1500000    3
          800000     3
          1300000    2
          Name: Price, dtype: int64,
          1600000    9
          1200000    5
          1500000    5
          1100000    3
          1300000    3
          Name: Price, dtype: int64]
```

Cluster 0, 2, and 3 have more number of customers (91)

Based on cluster 0, 2, and 3

37% of people spend upto 10 lacks on EV

43% of people spend upto 15 lacks on EV

only 11% of people spend more than 15 lacks on EV

Based on this we can conclude that most of the customers ready to spend upto 15 lacks on EV

```
In [90]: [Cluster_0['No of Dependents'].value_counts().head(),
          Cluster_1['No of Dependents'].value_counts().head(),
          Cluster_2['No of Dependents'].value_counts().head(),
          Cluster_3['No of Dependents'].value_counts().head()]

Out[90]: [0    15
          3    11
          4     6
          2     6
          Name: No of Dependents, dtype: int64,
          2    8
          3    6
          4    1
          Name: No of Dependents, dtype: int64,
          3    7
          2    6
          4    5
          0    2
          Name: No of Dependents, dtype: int64,
          3    10
          2     9
          0     5
          4     2
          Name: No of Dependents, dtype: int64]
```

Family with 3 to 4 members are more interested in EV

```
In [89]: [Cluster_0['Marrital Status'].value_counts().head(),
          Cluster_1['Marrital Status'].value_counts().head(),
          Cluster_2['Marrital Status'].value_counts().head(),
          Cluster_3['Marrital Status'].value_counts().head()]

Out[89]: [Married    25
          Single     13
          Name: Marrital Status, dtype: int64,
          Married    15
          Name: Marrital Status, dtype: int64,
          Married    18
          Single      2
          Name: Marrital Status, dtype: int64,
          Married    26
          Name: Marrital Status, dtype: int64]
```

Married people are supposed to be targeted

## Behavorial:

This is the one segment that relies on research into the decisions of a company's current customers. New products may be introduced based on research into the proven appeal of past products.

## Result:

Charging and range remain the top priority for consumers entering the electric vehicle market. Whilst having their differences, the two go hand-in hand as consumers worry about how to charge, and how often they can charge their vehicle.

## Conclusion

### Demographic Segments :

Age target : 27 to 44 years

Income target : 1.3 lacks to 2.5 lacks

Education target : Post graduation

### Geographic Segments : ¶

Location : Andhra Pradesh, Assam, Bihar, Goa, Gujarat, Kerala, Lakshadweep, Madhya Pradesh, Manipur, Rajastan, TamilNadu Tripura, Telangana, West Bengal

*** Note : HighWay charging stations analysis also suggested

### Psychographic Segments :

Family - Status : Married and family members from 2 to 4

In [ ]:

GitHub : https://github.com/ManojKamatam/EV_Market_Segementation.git