

ELECTRIC VEHICLE SALES BY STATE IN INDIA

By Himanshu Gautam



 [LinkedIn](#)

 [Github](#)

Market segmentation becomes a crucial tool for evolving transportation technology such as electric vehicles (EVs) in emerging markets to explore and implement for extensive adoption. EV adoption is expected to grow phenomenally shortly as low-emission and low-operating-cost cars, and thus, it drives a considerable amount of forthcoming academic research curiosity. The main aim of this study is to explore and identify distinct sets of potential buyer segments for EVs based on *psychographic, behavioral, and socio-economic* characterization by employing an integrated research framework of '*perceived benefits-attitude-intention*'.

Raw Data:

- Click [here](#) to access raw data
- This dataset is valuable for analysts, data scientists, and researchers aiming to understand electric vehicle (EV) adoption trends across India.
- It is versatile and ideal for geographic market segmentation, trend analysis, and predictive modeling.

Key Features:

- State: Names of Indian states with recorded EV sales data.
- Vehicle Type: Classifications of vehicles, such as two-wheelers and four-wheelers.
- Vehicle Category: Further classification into segments like commercial and passenger vehicles.
- Electric_Vehicle_Sales_Quantity: The number of EVs sold per state, essential for analyzing adoption trends.

Steps Involved:

- **Data Collection:** Load and inspect the dataset.
- **Data Preprocessing:** Handle missing values, convert date formats, and perform feature engineering.
- **Exploratory Data Analysis (EDA):** Visualize trends and relationships between variables.
- **Feature Engineering:** Create new features from the date column and encode categorical variables.
- **Modeling:** Build a regression model to predict EV sales.
- **Evaluation:** Evaluate the model performance and interpret the results.
- **Visualization:** Visualize the results and trends using graphs and charts.

Technologies Used:

Python, Numpy, Pandas, Matplotlib, Seaborn

```

import numpy as np
%pip install plotly==5.8.0
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
from tqdm import tqdm
import seaborn as sb
import statsmodels.api as sm
import plotly.express as px
from google.colab import files
pd.set_option("display.precision",3)
np.set_printoptions(precision=5, suppress=True)
pd.options.display.float_format = '{:.4f}'.format
import plotly.io as pio

pio.renderers.default = "svg"

```

```

df = pd.read_csv('/content/drive/MyDrive/Electric Vehicle Sales by State in India.csv')
df.head()

```

	Year	Month_Name	Date	State	Vehicle_Class	Vehicle_Category	Vehicle_Type	EV_Sales_Quantity
0	2014.0000	jan	1/1/2014	Andhra Pradesh	ADAPTED VEHICLE	Others	Others	0.0000
1	2014.0000	jan	1/1/2014	Andhra Pradesh	AGRICULTURAL TRACTOR	Others	Others	0.0000
2	2014.0000	jan	1/1/2014	Andhra Pradesh	AMBULANCE	Others	Others	0.0000
3	2014.0000	jan	1/1/2014	Andhra Pradesh	ARTICULATED VEHICLE	Others	Others	0.0000
4	2014.0000	jan	1/1/2014	Andhra Pradesh	BUS	Bus	Bus	0.0000

Information of the type of data in seach column

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 96845 entries, 0 to 96844
```

```
Data columns (total 8 columns):
```

#	Column	Non-Null Count	Dtype
0	Year	96845 non-null	float64
1	Month_Name	96845 non-null	object
2	Date	96845 non-null	object
3	State	96845 non-null	object
4	Vehicle_Class	96845 non-null	object
5	Vehicle_Category	96845 non-null	object
6	Vehicle_Type	96845 non-null	object
7	EV_Sales_Quantity	96845 non-null	float64

```
dtypes: float64(2), object(6)
```

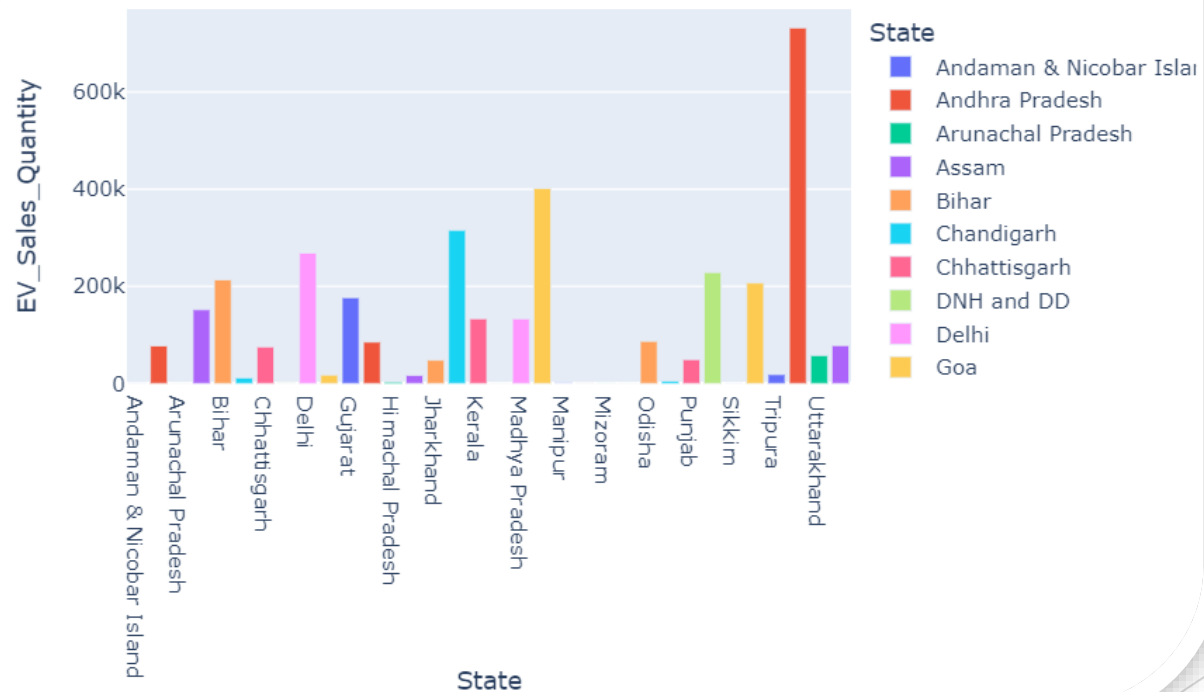
```
memory usage: 5.9+ MB
```

```
import kaleido
import plotly.io as pio
pio.renderers.default = "svg"
```

```
sales_by_state=df['EV_Sales_Quantity'].groupby(df['State']).sum()
sales_by_state=pd.DataFrame(sales_by_state)
sales_by_state=sales_by_state.reset_index()
```

```
fig = px.bar(sales_by_state,x='State',y = 'EV_Sales_Quantity',color = 'State',title = 'Sales In States',labels = {'x':'States','y':'Ev Sales'})
fig.show(renderer="svg")
```

Sales In States

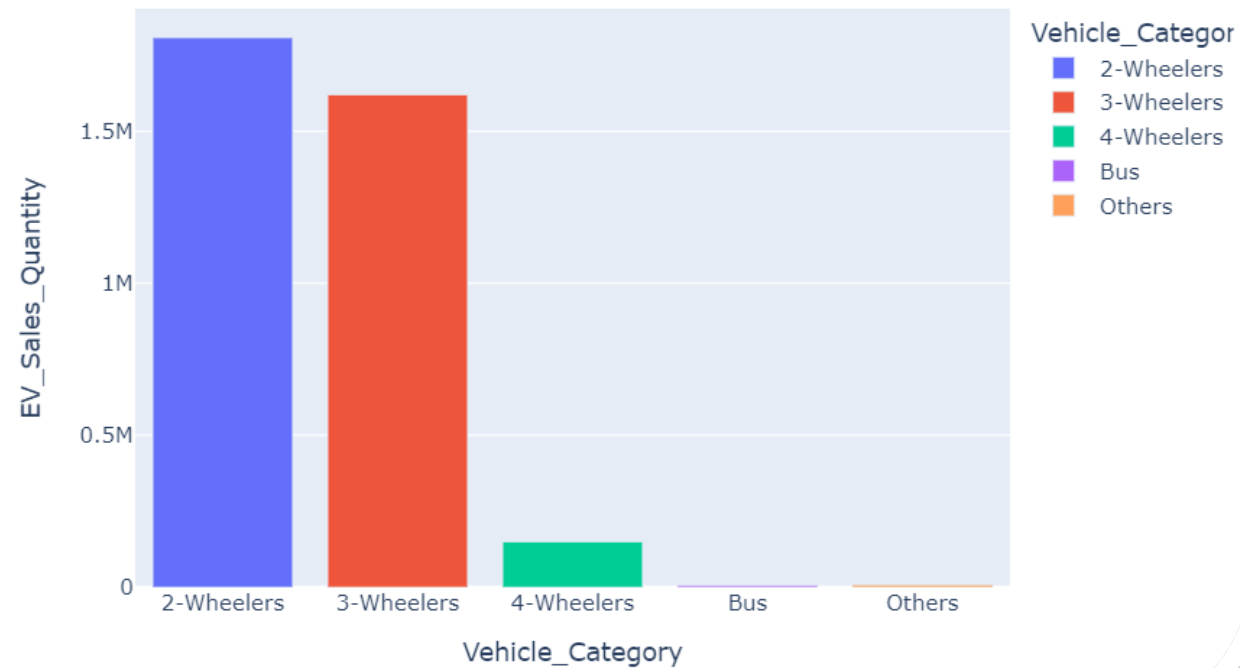


```
sales_by_vc = df['EV_Sales_Quantity'].groupby(df['Vehicle_Category']).sum()
sales_by_vc = pd.DataFrame(sales_by_vc)
```

Python

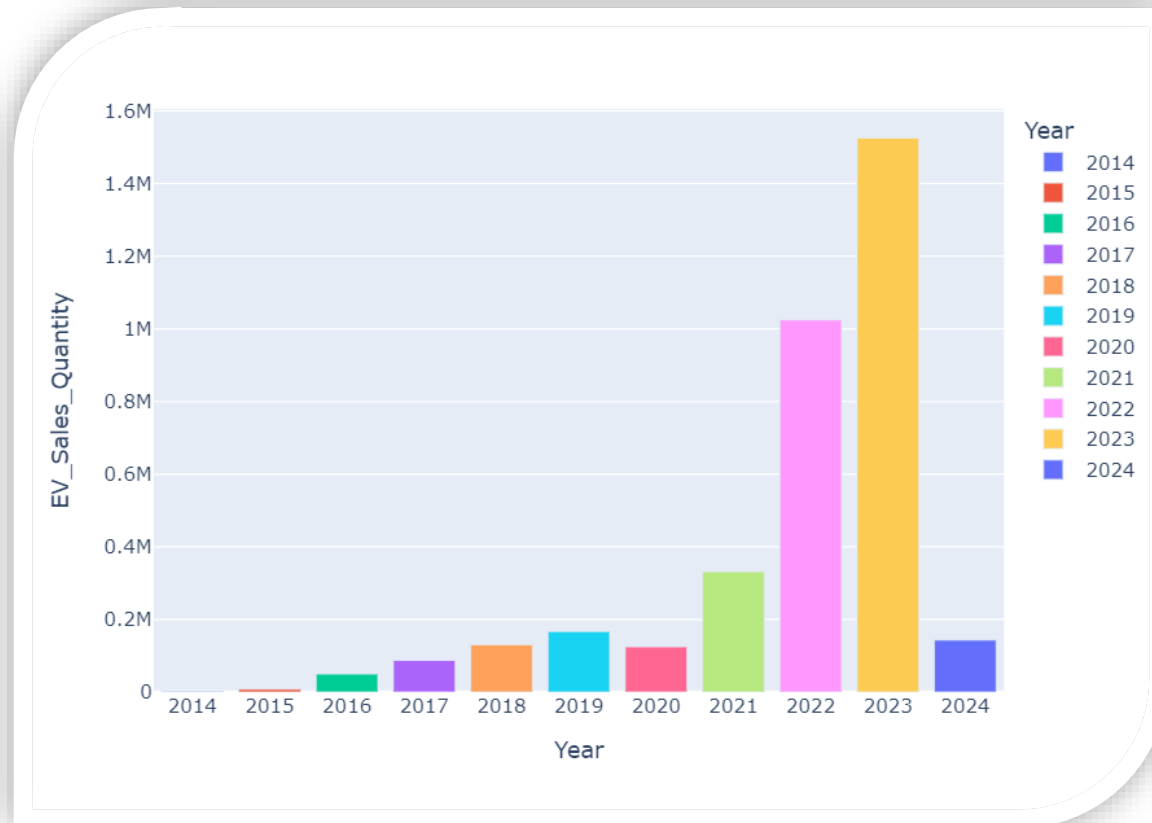
```
fig = px.bar(sales_by_vc,x='Vehicle_Category',y = 'EV_Sales_Quantity',color = 'Vehicle_Category',title = 'Which vehiclw class has more Sales?',labels = {'x':'vehicle class','y':'EV sales'})
pio.show(fig)
```

Which vehiclw class has more Sales?




```
sales_per_year = df['EV_Sales_Quantity'].groupby(df['Year']).sum()
sales_per_year = pd.DataFrame(sales_per_year)
sales_per_year = sales_per_year.reset_index()
sales_per_year['Year'] = sales_per_year['Year'].astype(str)
sales_per_year['Year'] = sales_per_year['Year'].str.split('.').str[0]
```

```
fig = px.bar(sales_per_year, x = 'Year', y = 'EV_Sales_Quantity', color = 'Year')
pio.show(fig)
```

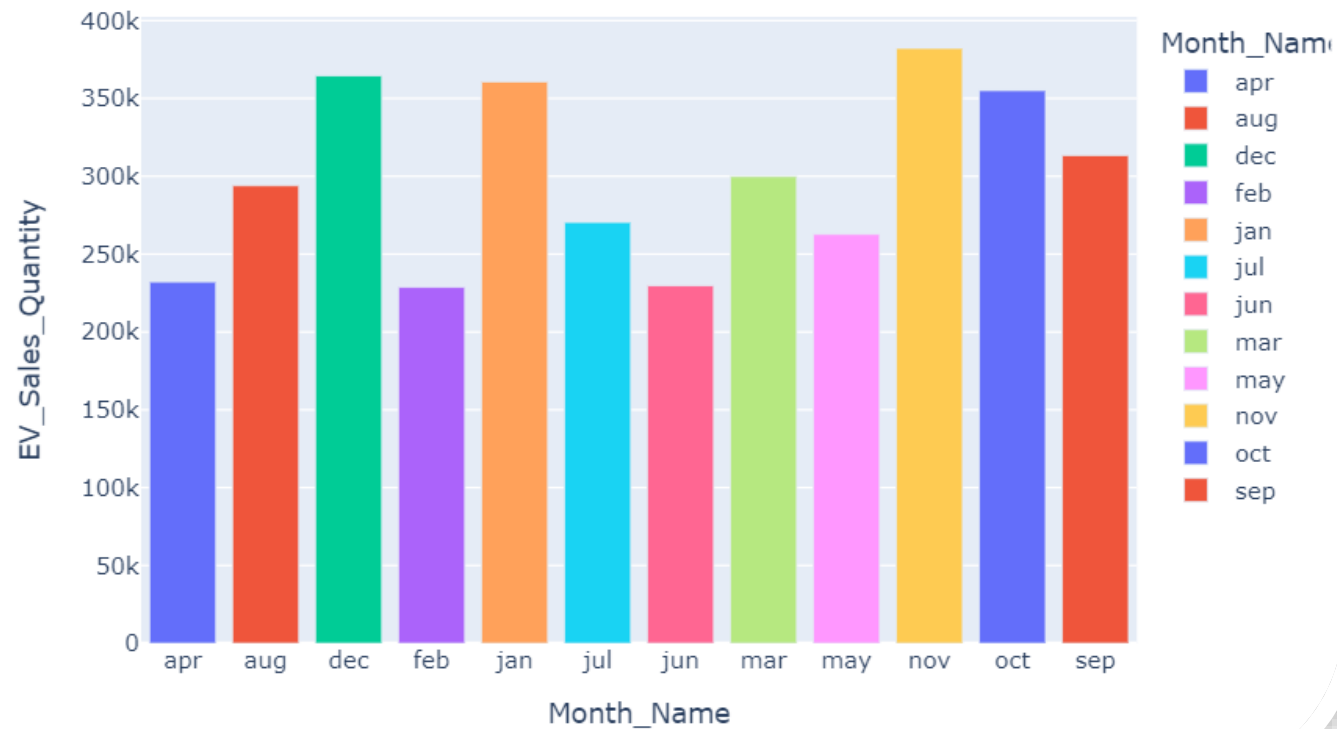


```
sales_by_month = df['EV_Sales_Quantity'].groupby(df['Month_Name']).sum()
sales_by_month = pd.DataFrame(sales_by_month)
sales_by_month = sales_by_month.reset_index()
```

Python

```
fig = px.bar(sales_by_month,x='Month_Name',y = 'EV_Sales_Quantity',color = 'Month_Name',title = 'Monthly sales of all time',labels : {'x':'Car Brands','y':'Top Speed Km/H'})
pio.show(fig)
```

Monthly sales of all time



Sample Insights:

- *Among all the years, 2023 recorded the highest number of sales of electric vehicle*
- *Among all the states, Utter Pradesh recorded the highest number of sales of electric vehicle*
- *The 2-Wheelers class has the highest number of sales*
- *Mostly electric vehicles were sold in the month of November*
- *Among the vehicle types, 56% of the market was captured by other vehicles right after 2W-Personal*