Electric Vehicle Sales Analysis

About Dataset:

- The dataset contains the following columns:
- ● Year: The year of the sales.
- ● Month_Name: The month in which sales occurred.
- ● Date: The specific date of the sales.
- ● State: The state in India where the sales occurred.
- ● Vehicle_Class: The class of the vehicle (e.g., sedan, SUV, etc.).
- ● Vehicle_Category: The category of the vehicle (e.g., commercial, passenger).
- ● Vehicle_Type: The type of the vehicle (e.g., 2-wheeler, 4-wheeler).
- ● EV_Sales_Quantity: The quantity of EV sales.

Objective:

- This project aims to analyze and predict the sales of Electric Vehicles (EV) by state in India using machine learning.

Steps:

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

Data Loading:

```
df = pd.read_csv(r'C:\Users\HP\Downloads\Electric Vehicle Sales by State in India.csv')
df
```

	Year	Month_Name	Date	State	Vehicle_Class	Vehicle_Category	Vehicle_Type	EV_Sales_Quantity
0	2014.0	jan	1/1/2014	Andhra Pradesh	ADAPTED VEHICLE	Others	Others	0.0
1	2014.0	jan	1/1/2014	Andhra Pradesh	AGRICULTURAL TRACTOR	Others	Others	0.0
2	2014.0	jan	1/1/2014	Andhra Pradesh	AMBULANCE	Others	Others	0.0
3	2014.0	jan	1/1/2014	Andhra Pradesh	ARTICULATED VEHICLE	Others	Others	0.0
4	2014.0	jan	1/1/2014	Andhra Pradesh	BUS	Bus	Bus	0.0
...	...	...	...	...	...	...	...	...
96840	2023.0	dec	12/1/2023	Andaman & Nicobar Island	MOTOR CAR	4-Wheelers	4W_Personal	1.0
96841	2023.0	dec	12/1/2023	Andaman & Nicobar Island	MOTOR CYCLE/SCOOTER-USED FOR HIRE	2-Wheelers	2W_Shared	5.0
96842	2023.0	dec	12/1/2023	Andaman & Nicobar Island	OMNI BUS	Bus	Bus	0.0
96843	2023.0	dec	12/1/2023	Andaman & Nicobar Island	THREE WHEELER (GOODS)	3-Wheelers	3W_Goods	0.0
96844	2023.0	dec	12/1/2023	Andaman & Nicobar Island	THREE WHEELER (PASSENGER)	3-Wheelers	3W_Shared	0.0

Data Preparation and Cleaning:

```
df.shape
```

```
(96845, 8)
```

```
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
```

```
df.nunique()
```

```
Year                 11
Month_Name           12
Date                121
State                34
Vehicle_Class        73
Vehicle_Category      5
Vehicle_Type         12
EV_Sales_Quantity  1447
dtype: int64
```

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

```
Year               0
Month_Name         0
Date               0
State              0
Vehicle_Class      0
Vehicle_Category   0
Vehicle_Type       0
EV_Sales_Quantity  0
dtype: int64
```

```
df.duplicated().sum()
```

```
np.int64(0)
```

```
df['Year'] = df['Year'].astype(int)
```

```
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  int64  
 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(1), int64(1), object(6)
memory usage: 5.9+ MB
```

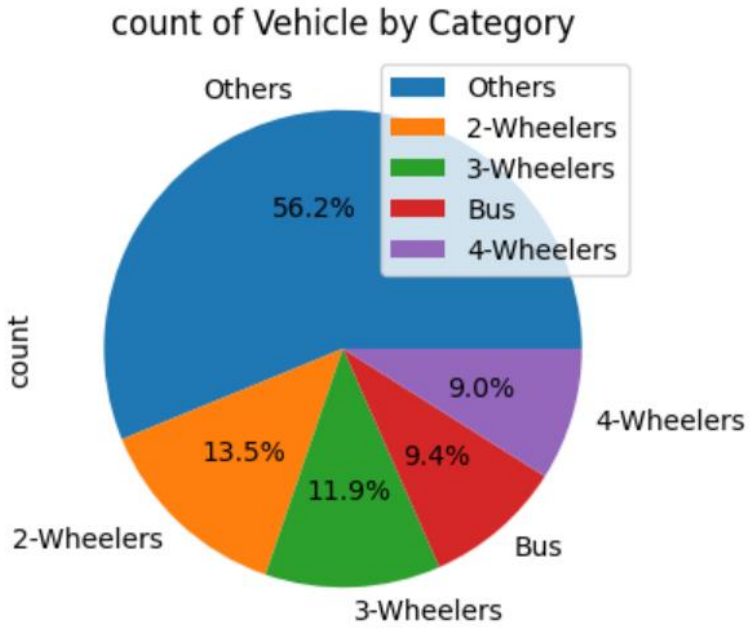
count of vehicles in different category:

```
vc = df['Vehicle_Category'].value_counts()
vc
```

```
Vehicle_Category
Others          54423
2-Wheelers      13121
3-Wheelers      11491
Bus              9119
4-Wheelers       8691
Name: count, dtype: int64
```

```
vc.plot(kind= "pie", figsize = (5,4), autopct = "%1.1f%%")

plt.title("count of Vehicle by Category")
plt.legend(labels = vc.index)
plt.show()
```

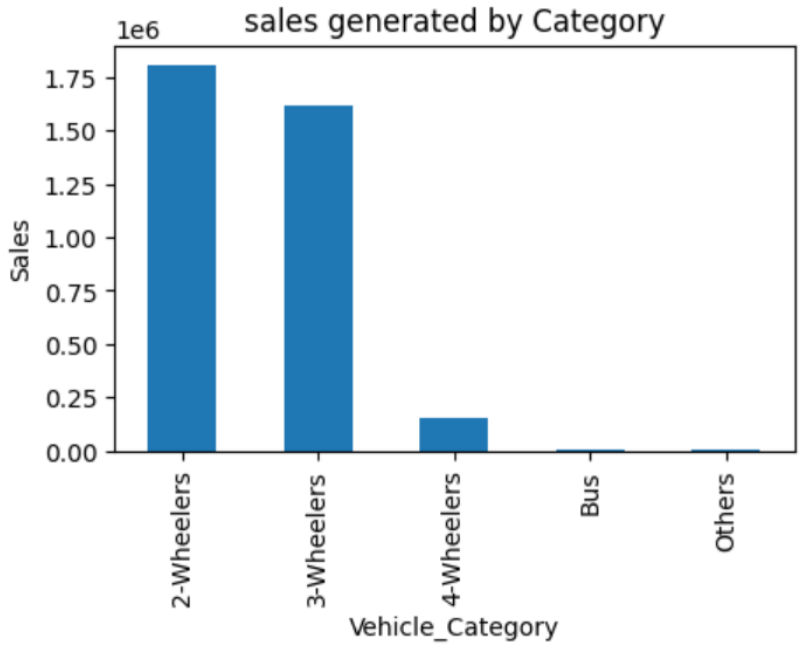


Sales Contribution by Category:

```
vc_sales = df.groupby('Vehicle_Category')['EV_Sales_Quantity'].sum()
vc_sales
```

```
Vehicle_Category
2-Wheelers    1808105.0
3-Wheelers    1620310.0
4-Wheelers     149775.0
Bus              7009.0
Others           8612.0
Name: EV_Sales_Quantity, dtype: float64
```

```
vc_sales.plot(kind = "bar", figsize = (5,3))
plt.title("sales generated by Category")
plt.ylabel("Sales")
plt.show()
```

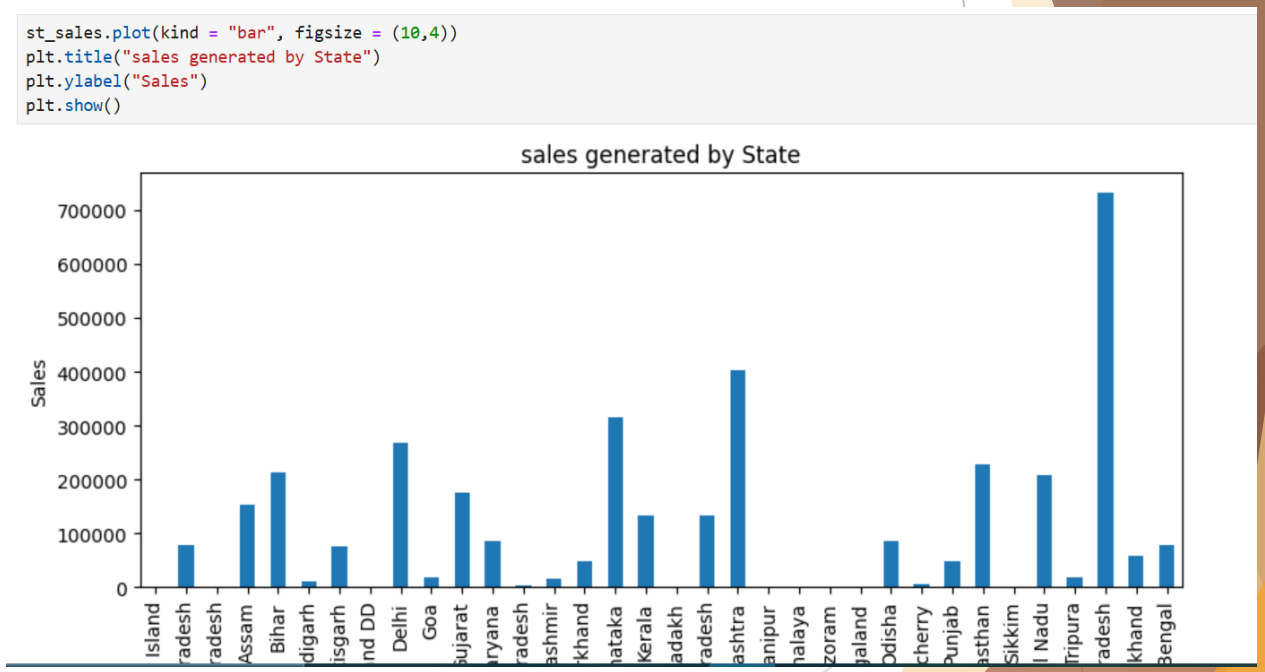


Sales Contribution by State:

```
st_sales = df.groupby('State')['EV_Sales_Quantity'].sum()
st_sales
```

```
State
Andaman & Nicobar Island       202.0
Andhra Pradesh               77356.0
Arunachal Pradesh               40.0
Assam                       151917.0
Bihar                       213465.0
Chandigarh                   11453.0
Chhattisgarh                 75275.0
DNH and DD                     431.0
Delhi                       268538.0
Goa                          17382.0
Gujarat                     176713.0
Haryana                      85250.0
Himachal Pradesh              2886.0
Jammu and Kashmir            16840.0
Jharkhand                    47871.0
Karnataka                   315498.0
Kerala                      133246.0
Ladakh                         88.0
```

```
st_sales.plot(kind = "bar", figsize = (10,4))
plt.title("sales generated by State")
plt.ylabel("Sales")
plt.show()
```

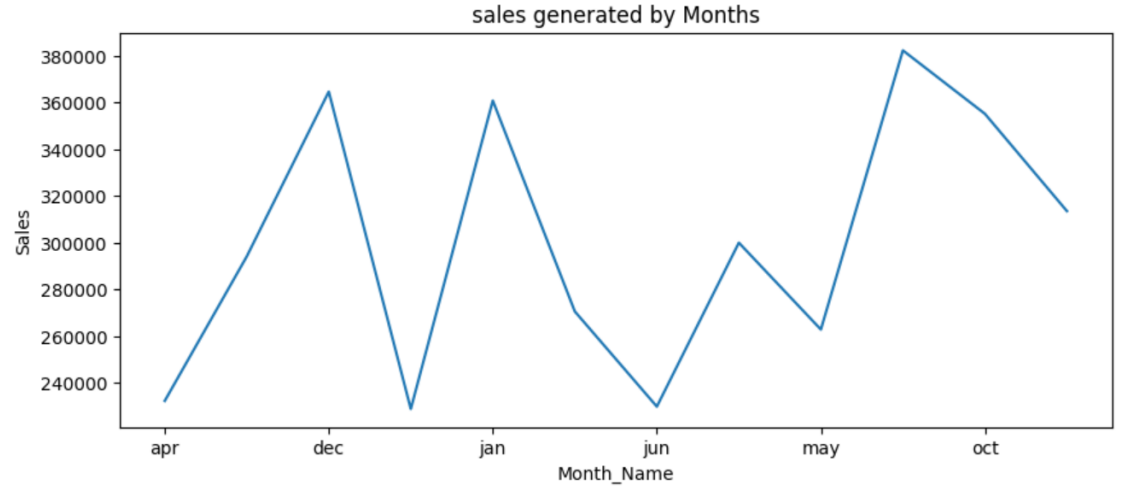


Sales by Month:

```
mn_sales = df.groupby('Month_Name')['EV_Sales_Quantity'].sum()
mn_sales
```

```
Month_Name
apr    232194.0
aug    294022.0
dec    364558.0
feb    228739.0
jan    360703.0
jul    270473.0
jun    229754.0
mar    299888.0
may    262747.0
nov    382217.0
oct    355083.0
sep    313433.0
Name: EV_Sales_Quantity, dtype: float64
```

```
mn_sales.plot(kind = "line", figsize = (10,4))
plt.title("sales generated by Months")
plt.ylabel("Sales")
plt.show()
```

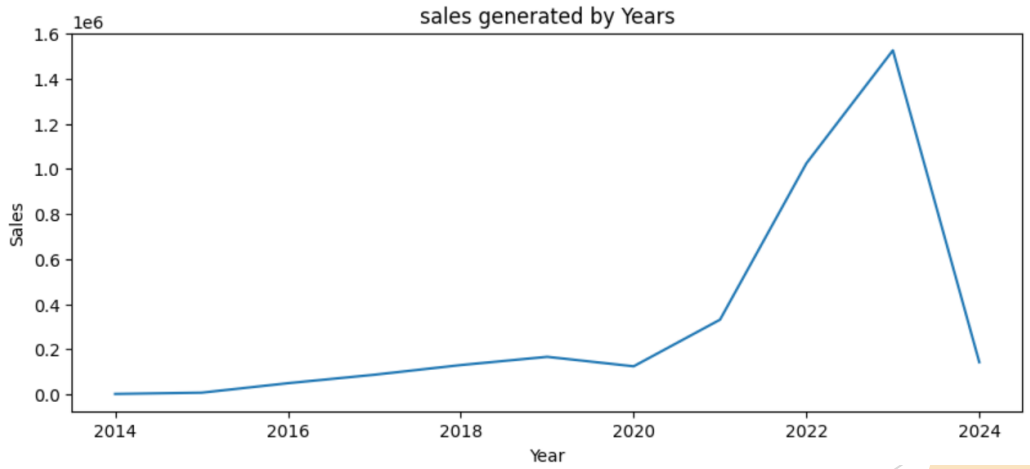


Sales generated by Year:

```
yr_sales = df.groupby('Year')['EV_Sales_Quantity'].sum()
yr_sales
```

```
Year
2014       2392.0
2015       7805.0
2016      49855.0
2017      87420.0
2018     130254.0
2019     166819.0
2020     124684.0
2021     331498.0
2022    1024723.0
2023    1525179.0
2024     143182.0
Name: EV_Sales_Quantity, dtype: float64
```

```
yr_sales.plot(kind = "line", figsize = (10,4))
plt.title("sales generated by Years")
plt.ylabel("Sales")
plt.show()
```

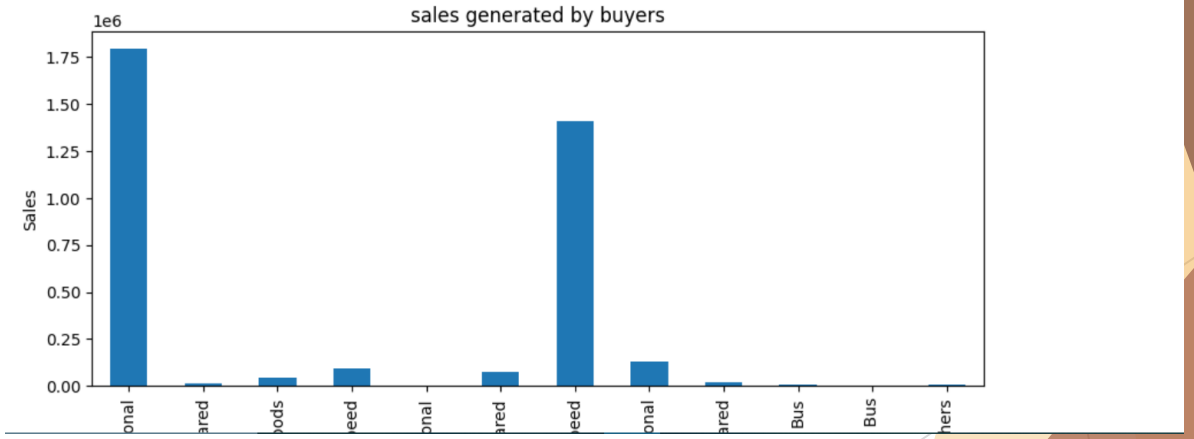


Sales Generated by Vehicle Type:

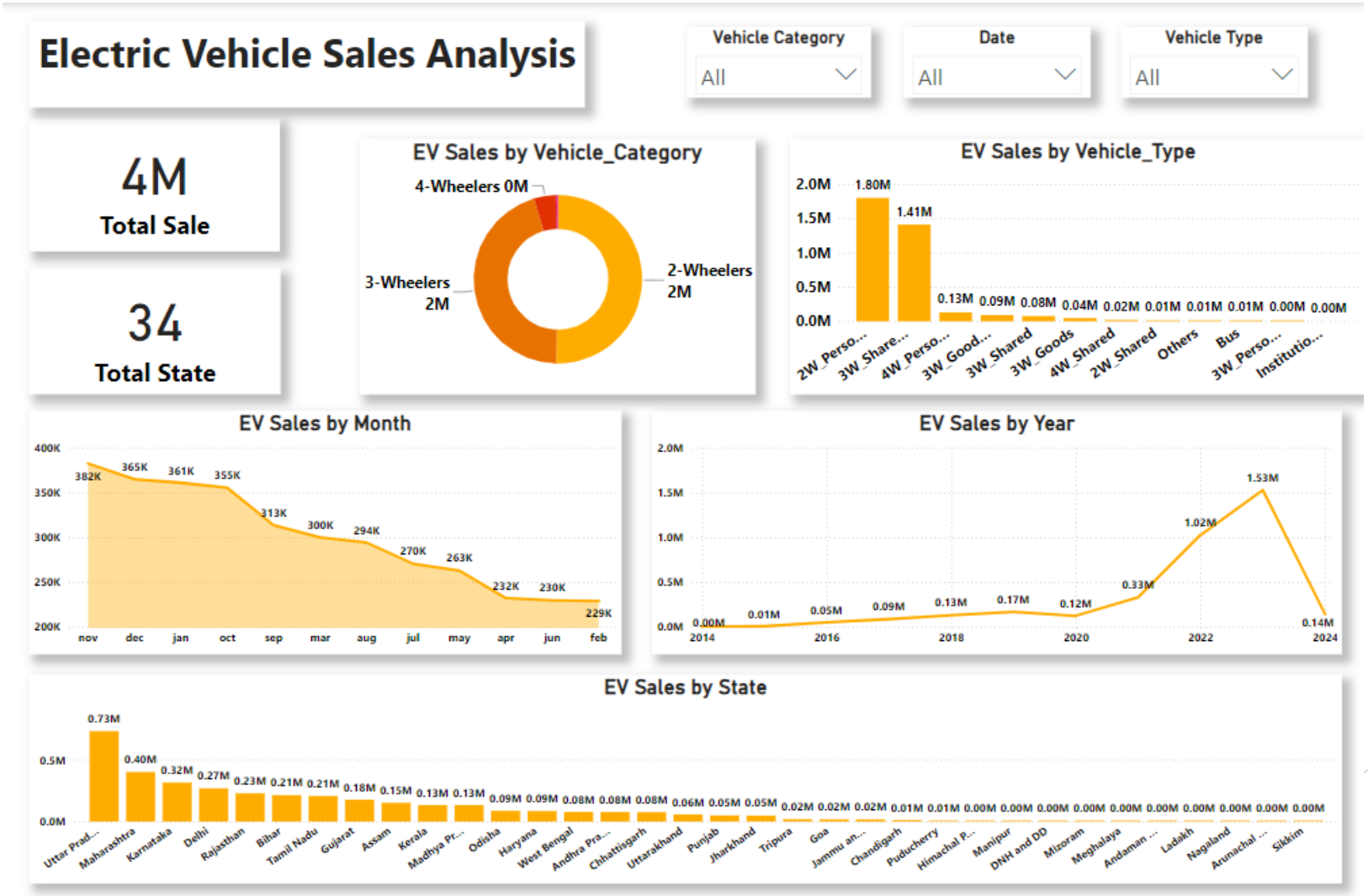
```
vt_sales = df.groupby('Vehicle_Type')['EV_Sales_Quantity'].sum()
vt_sales
```

```
Vehicle_Type
2W_Personal           1796340.0
2W_Shared               11765.0
3W_Goods                44974.0
3W_Goods_LowSpeed       90656.0
3W_Personal               421.0
3W_Shared               76132.0
3W_Shared_LowSpeed    1408127.0
4W_Personal            130676.0
4W_Shared               19099.0
Bus                      7009.0
Institution Bus             0.0
Others                   8612.0
Name: EV_Sales_Quantity, dtype: float64
```

```
vt_sales.plot(kind = "bar", figsize = (10,4))
plt.title("sales generated by buyers")
plt.ylabel("Sales")
plt.show()
```



Dashboard:


Electric Vehicle Sales Analysis
Vehicle Category
All
Date
All
Vehicle Type
All
4M
Total Sale
34
Total State
EV Sales by Vehicle_Category
4-Wheelers 0M
2-Wheelers 2M
3-Wheelers 2M
EV Sales by Vehicle_Type
2.0M
1.5M
1.0M
0.5M
0.0M
1.80M
1.41M
0.13M
0.09M
0.08M
0.04M
0.02M
0.01M
0.01M
0.01M
0.00M
0.00M
2W_Perso...
3W_Share...
4W_Perso...
3W_Good...
3W_Shared
3W_Goods
4W_Shared
2W_Shared
Others
Bus
3W_Perso...
Institutio...
EV Sales by Month
400K
350K
300K
250K
200K
382K
365K
361K
355K
313K
300K
294K
270K
263K
232K
230K
229K
nov
dec
jan
oct
sep
mar
aug
jul
may
apr
jun
feb
EV Sales by Year
2.0M
1.5M
1.0M
0.5M
0.0M
0.00M
0.01M
0.05M
0.09M
0.13M
0.17M
0.12M
0.33M
1.02M
1.53M
0.14M
2014
2016
2018
2020
2022
2024
EV Sales by State
0.5M
0.0M
0.73M
0.40M
0.32M
0.27M
0.23M
0.21M
0.21M
0.18M
0.15M
0.13M
0.13M
0.09M
0.09M
0.08M
0.08M
0.08M
0.06M
0.05M
0.05M
0.02M
0.02M
0.02M
0.01M
0.01M
0.00M
0.00M
0.00M
0.00M
0.00M
0.00M
0.00M
0.00M
0.00M
Uttar Prad...
Maharashtra
Karnataka
Delhi
Rajasthan
Bihar
Tamil Nadu
Gujarat
Assam
Kerala
Madhya Pr...
Odisha
Haryana
West Bengal
Andhra Pra...
Chhattisgarh
Uttarakhand
Punjab
Jharkhand
Tripura
Goa
Jammu an...
Chandigarh
Puducherry
Himachal P...
Manipur
DNH and DD
Mizoram
Meghalaya
Andaman ...
Ladakh
Nagaland
Arunachal ...
Sikkim

Insights:

- 1.A significant number of vehicles are categorized under the "Other" segment.
- 2.The 2-wheeler category accounts for the highest sales, underscoring its popularity as a preferred mode of transportation, likely due to affordability, ease of use, and adaptability for personal use.
- 3.Uttar Pradesh and Maharashtra lead in EV sales, showcasing strong demand in these states. This could be attributed to factors like population density, urbanization, and state-level EV incentives.
- 4.November records the highest sales volume, likely influenced by festive seasons, year-end discounts, or government policies promoting EV adoption during this time.
- 5.The year 2023 witnessed the most significant growth in sales, reflecting increased consumer awareness, improved EV infrastructure, and aggressive marketing campaigns by manufacturers.
- 6.The majority of sales come from 2-wheeler personal vehicles, reinforcing their utility as an efficient and economical transportation option for individual consumers.