Analysis the EV Sales

The electric vehicle (EV) market in India is growing fast, thanks to efforts to promote eco-friendly transportation and government support. This project looks at EV sales by state in India to see ho different regions are doing. Each state has its own way of encouraging EV use, which has led to levels of sales and popularity. By studying this data, we can understand which state policies are best, what people prefer, and how infrastructure is developing. This project aims to give a clear p EV sales in India, showing both the opportunities and challenges faced in each state.

In [3]:
```
# Import necessary libraries:-

import pandas as pd
import numpy as np
import seaborn as sns #visualisation
import matplotlib.pyplot as plt #visualisation
```

In [4]:
```
# Load the dataset:-

data=pd.read_csv("Electric Vehicle Sales by State in India.csv")
```

Data Overview

Let's take a look at the first few rows of the dataset to understand its structure.

In [7]:
```
data.head(10)
```

Out[7]:

	Year	Month_Name	Date	State	Vehicle_Class	Vehicle_Category	Vehicle_Ty
0	2014.0	jan	1/1/2014	Andhra Pradesh	ADAPTED VEHICLE	Others	Othe
1	2014.0	jan	1/1/2014	Andhra Pradesh	AGRICULTURAL TRACTOR	Others	Othe
2	2014.0	jan	1/1/2014	Andhra Pradesh	AMBULANCE	Others	Othe
3	2014.0	jan	1/1/2014	Andhra Pradesh	ARTICULATED VEHICLE	Others	Othe
4	2014.0	jan	1/1/2014	Andhra Pradesh	BUS	Bus	B
5	2014.0	jan	1/1/2014	Andhra Pradesh	CASH VAN	Others	Othe
6	2014.0	jan	1/1/2014	Andhra Pradesh	CRANE MOUNTED VEHICLE	Others	Othe
7	2014.0	jan	1/1/2014	Andhra Pradesh	EDUCATIONAL INSTITUTION BUS	Bus	Institut B
8	2014.0	jan	1/1/2014	Andhra Pradesh	EXCAVATOR (COMMERCIAL)	Others	Othe
9	2014.0	jan	1/1/2014	Andhra Pradesh	FORK LIFT	Others	Othe

In [9]:
```
data.dtypes
```

Out[9]:
```
Year                 float64
Month_Name            object
Date                  object
State                 object
Vehicle_Class         object
Vehicle_Category      object
Vehicle_Type          object
EV_Sales_Quantity    float64
dtype: object
```

In [11]:
```
data.columns
```

Out[11]:
```
Index(['Year', 'Month_Name', 'Date', 'State', 'Vehicle_Class',
       'Vehicle_Category', 'Vehicle_Type', 'EV_Sales_Quantity'],
      dtype='object')
```

In [13]:
```
data.shape
```

Out[13]:
```
(96845, 8)
```

In [15]:
```
data.info
```

Out[15]:
```
<bound method DataFrame.info of          Year Month_Name
Date                     State  \
0      2014.0        jan   1/1/2014            Andhra Pradesh
1      2014.0        jan   1/1/2014            Andhra Pradesh
2      2014.0        jan   1/1/2014            Andhra Pradesh
3      2014.0        jan   1/1/2014            Andhra Pradesh
4      2014.0        jan   1/1/2014            Andhra Pradesh
...       ...        ...        ...                       ...
96840  2023.0        dec  12/1/2023  Andaman & Nicobar Island
96841  2023.0        dec  12/1/2023  Andaman & Nicobar Island
96842  2023.0        dec  12/1/2023  Andaman & Nicobar Island
96843  2023.0        dec  12/1/2023  Andaman & Nicobar Island
96844  2023.0        dec  12/1/2023  Andaman & Nicobar Island

                           Vehicle_Class Vehicle_Category Vehicle_Type  \
0                       ADAPTED VEHICLE           Others       Others
1                    AGRICULTURAL TRACTOR           Others       Others
2                               AMBULANCE           Others       Others
3                     ARTICULATED VEHICLE           Others       Others
4                                     BUS              Bus          Bus
...                                   ...              ...          ...
96840                          MOTOR CAR       4-Wheelers  4W_Personal
96841  MOTOR CYCLE/SCOOTER-USED FOR HIRE       2-Wheelers    2W_Shared
96842                           OMNI BUS              Bus          Bus
96843               THREE WHEELER (GOODS)       3-Wheelers     3W_Goods
96844           THREE WHEELER (PASSENGER)       3-Wheelers    3W_Shared

       EV_Sales_Quantity
0                    0.0
1                    0.0
2                    0.0
3                    0.0
4                    0.0
...                  ...
96840                1.0
96841                5.0
96842                0.0
96843                0.0
96844                0.0

[96845 rows x 8 columns]>
```

Data Preprocessing

Before diving into analysis, we need to ensure our data is clean and ready for exploration. This ir parsing dates and checking for missing values.

In [17]:
```
# Convert 'Date' column to datetime format:-

data['Date'] = pd.to_datetime(data['Date'], format='%d/%m/%Y')
```

In [19]:
```
# Check for missing values:-

data.isnull().sum()
```

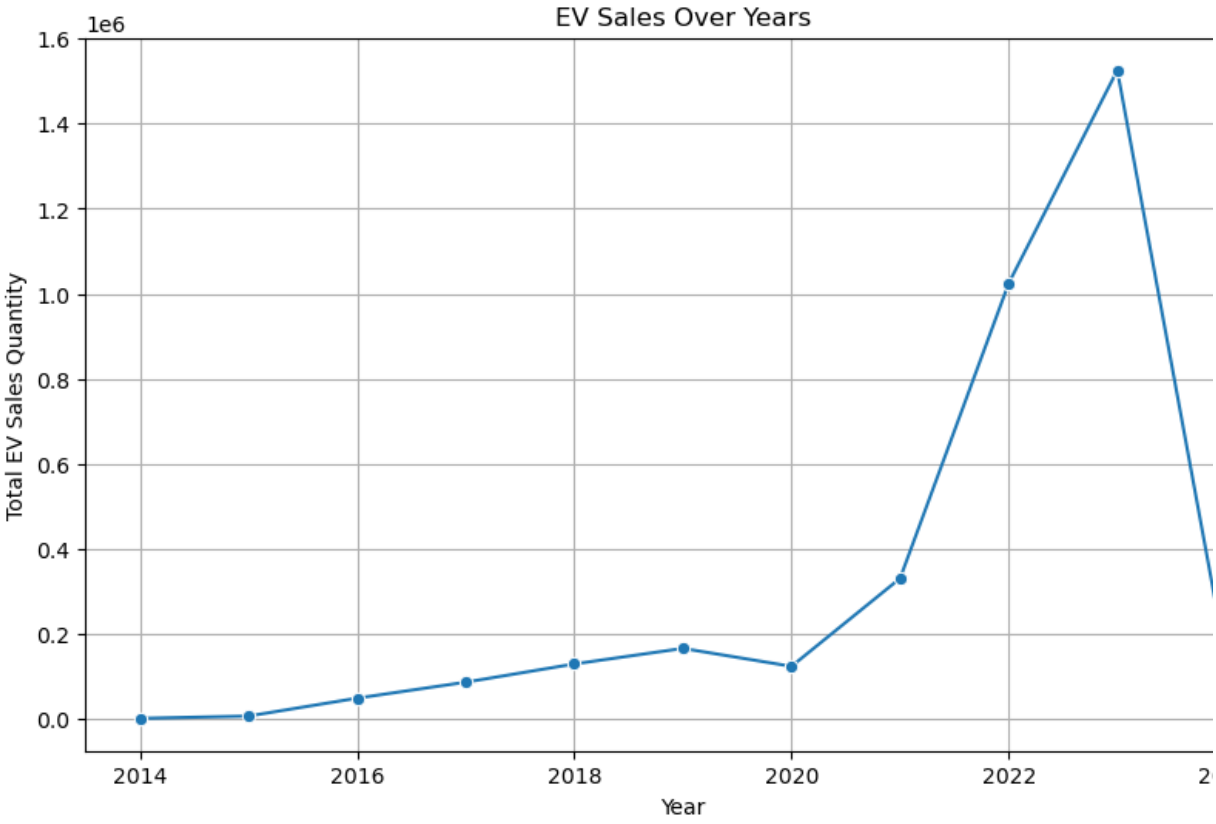
Out[19]:
```
Year                0
Month_Name          0
Date                0
State               0
Vehicle_Class       0
Vehicle_Category    0
Vehicle_Type        0
EV_Sales_Quantity   0
dtype: int64
```

Exploratory Data Analysis

Let's explore the data to uncover trends and patterns in EV sales across different states and veh categories.

In [21]:
```
# Plot EV sales over year:-

plt.figure(figsize=(10, 6))
yearly_sales = data.groupby('Year')['EV_Sales_Quantity'].sum()
sns.lineplot(x=yearly_sales.index, y=yearly_sales.values, marker='o')
plt.title("EV Sales Over Years")
plt.xlabel("Year")
plt.ylabel("Total EV Sales Quantity")
plt.grid()
plt.show()
```



Inference

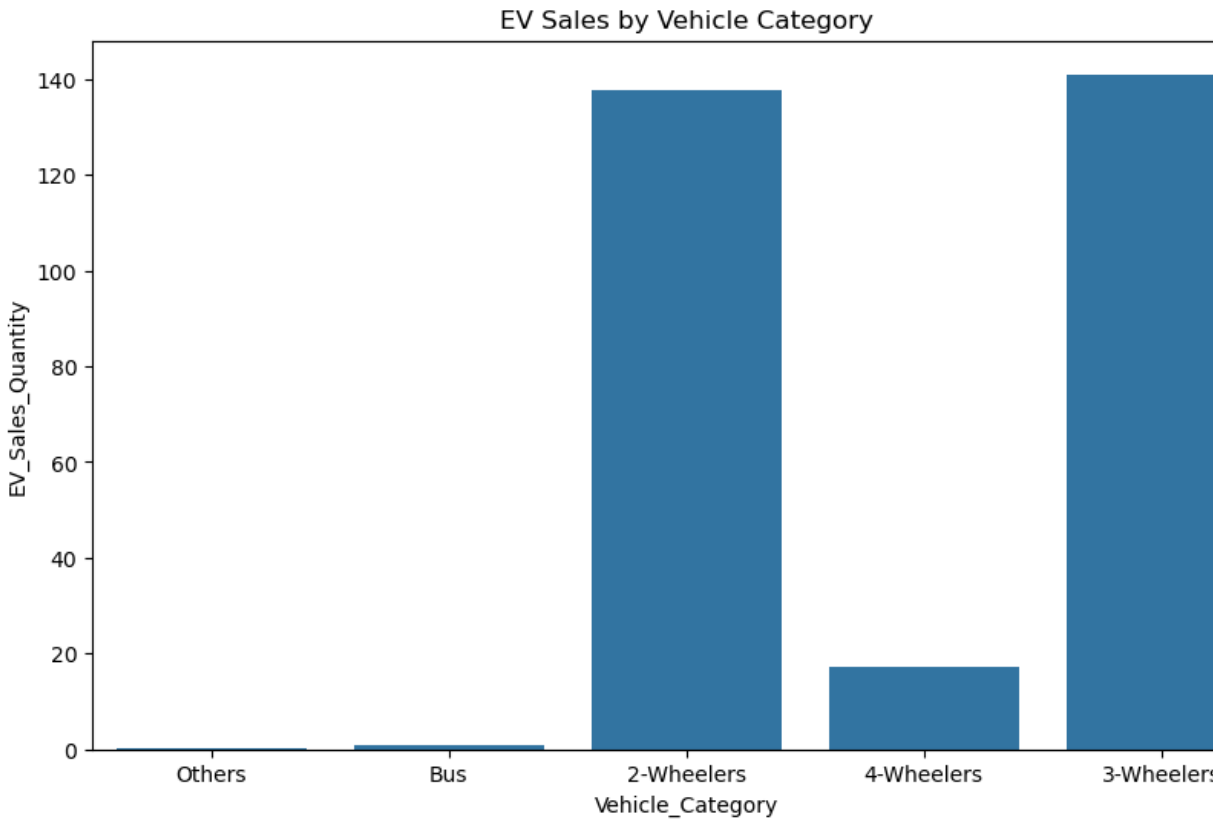
EV sales grew quickly in recent years, showing that more people are interested in them. The sud drop in 2024 suggests there might be new challenges that need attention to keep EV sales going

In [25]:
```
# Plot sales by vehicle category
plt.figure(figsize=(10, 6))
sns.barplot(x='Vehicle_Category', y='EV_Sales_Quantity',
data=data, ci=None)
plt.title('EV Sales by Vehicle Category')
plt.show()
```

```
C:\Users\maury\AppData\Local\Temp\ipykernel_3576\3105582523.py:3: FutureWarni

The `ci` parameter is deprecated. Use `errorbar=None` for the same effect.

  sns.barplot(x='Vehicle_Category', y='EV_Sales_Quantity',
```



Inference

The chart shows the total electric vehicle (EV) sales by category. It highlights that 2-wheelers hav highest sales, followed by 3-wheelers, while 4-wheelers, buses, and other vehicle categories hav significantly lower sales. This indicates that EV adoption is primarily driven by 2-wheelers and 3-wheelers.

Correlation Analysis

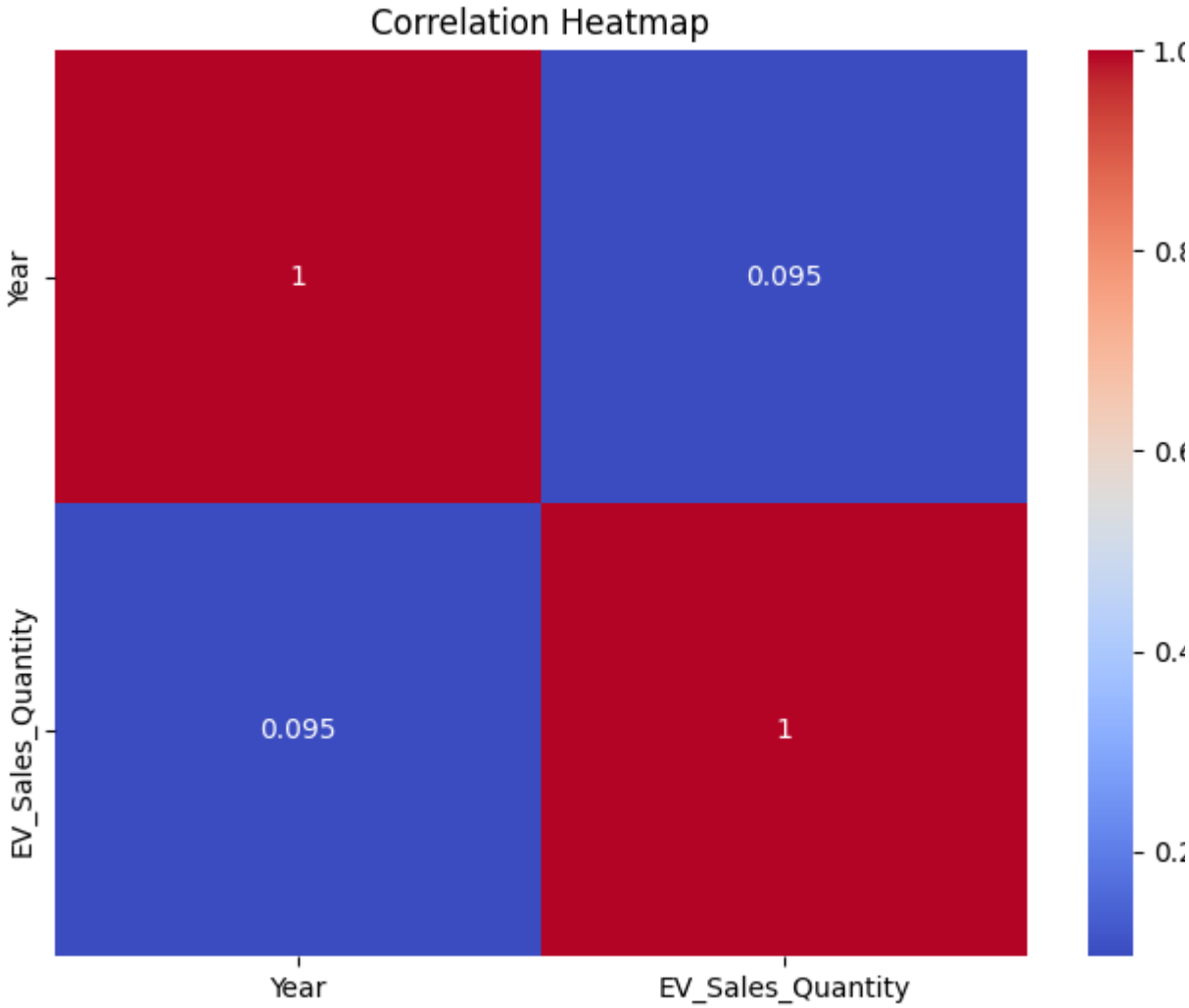
Let's examine the correlation between numeric variables to understand potential relationships.

In [12]:
```
# Select only numeric columns for correlation analysis:-

numeric_df = data.select_dtypes(include=[np.number])

# Plot the correlation heatmap:-

plt.figure(figsize=(8, 6))
sns.heatmap(numeric_df.corr(), annot=True, cmap='coolwarm')
plt.title('Correlation Heatmap')
plt.show()
```



Inference

The heatmap shows the correlation between the year and EV sales quantity. The correlation valu 0.095, which is very low. This means there is almost no linear relationship between the year and sales quantity in the given data.

Top 10 States by EV Sales

A horizontal bar chart highlighting the states with the highest EV sales.

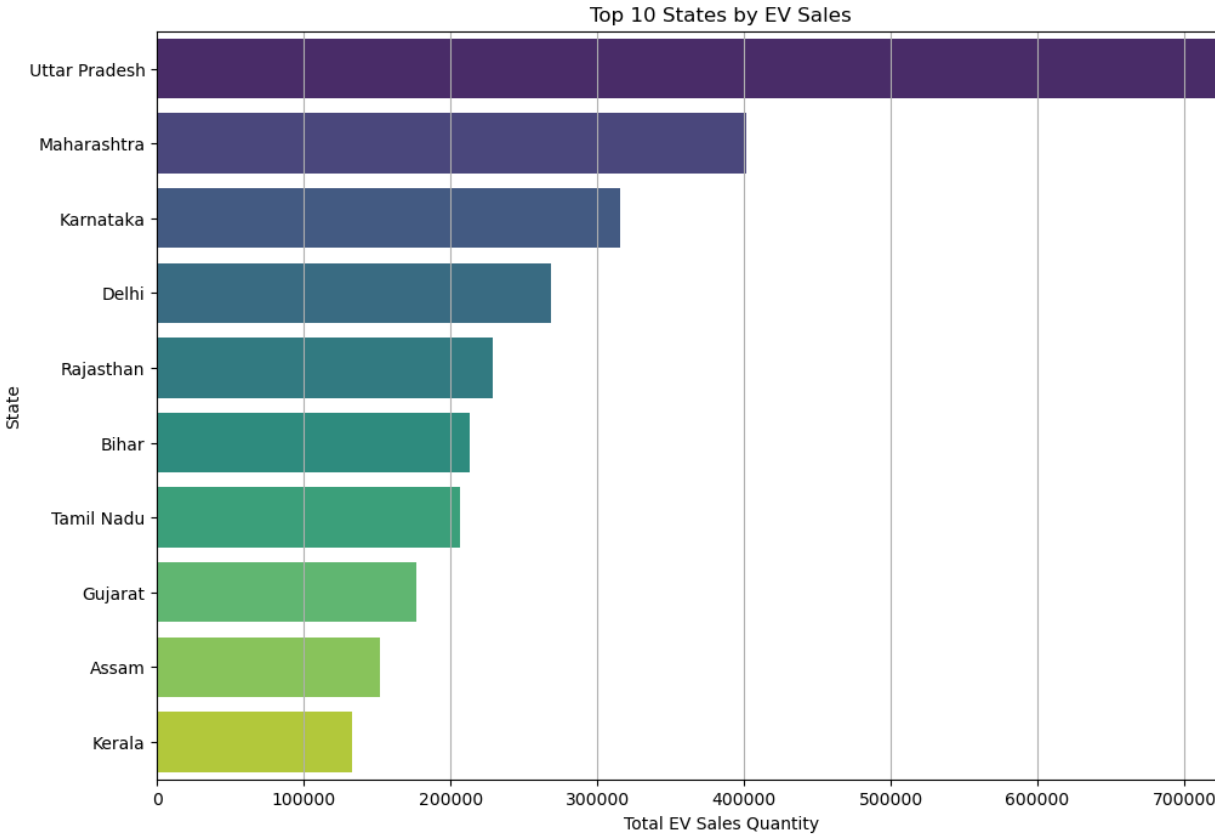
In [27]:
```
#EV sales by state (top 10 states):-

plt.figure(figsize=(12, 8))
state_sales = data.groupby('State')['EV_Sales_Quantity'].sum().sort_value
sns.barplot(x=state_sales.values, y=state_sales.index, palette='viridis')
plt.title("Top 10 States by EV Sales")
plt.xlabel("Total EV Sales Quantity")
plt.ylabel("State")
plt.grid(axis='x')
plt.show()
```

```
C:\Users\maury\AppData\Local\Temp\ipykernel_3576\3833774895.py:5: FutureWarni

Passing `palette` without assigning `hue` is deprecated and will be removed i
v0.14.0. Assign the `y` variable to `hue` and set `legend=False` for the same
effect.

  sns.barplot(x=state_sales.values, y=state_sales.index, palette='viridis')
```



Inference

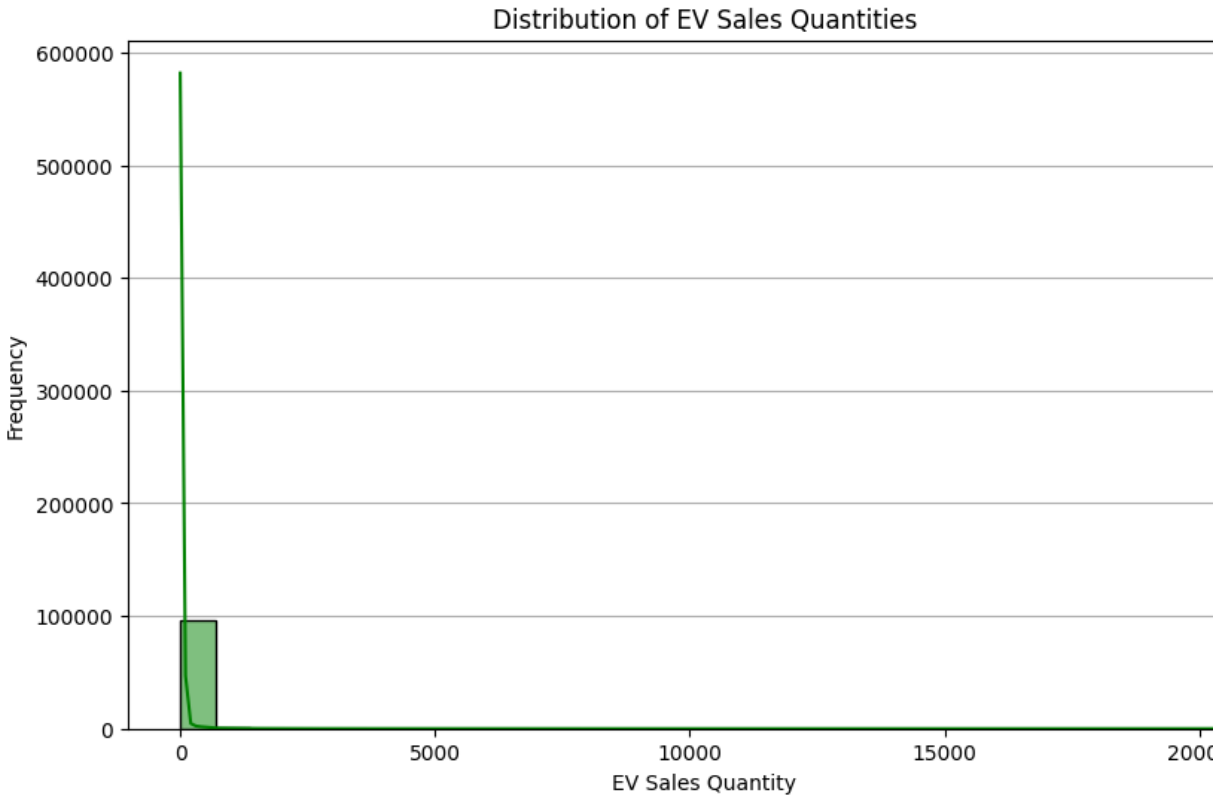
The chart shows the top 10 states by EV sales. Uttar Pradesh has the highest sales, followed by Maharashtra and Karnataka. States like Kerala, Assam, and Gujarat have lower sales compared top-performing states. This indicates that EV adoption varies significantly across states.

Distribution of EV Sales Quantities

Understand the distribution of sales quantities (e.g., skewness or typical sales quantities).

In [19]:
```
#Distribution of EV sales quantities:-

plt.figure(figsize=(10, 6))
sns.histplot(data['EV_Sales_Quantity'], bins=30, kde=True, color='green')
plt.title("Distribution of EV Sales Quantities")
plt.xlabel("EV Sales Quantity")
plt.ylabel("Frequency")
plt.grid(axis='y')
plt.show()
```



Inference

The graph shows the distribution of EV sales quantities, with a clear peak indicating a high frequ
sales in the range of around 100,000 units. The data suggests a skewed distribution, with the ma
sales falling within a relatively narrow range.

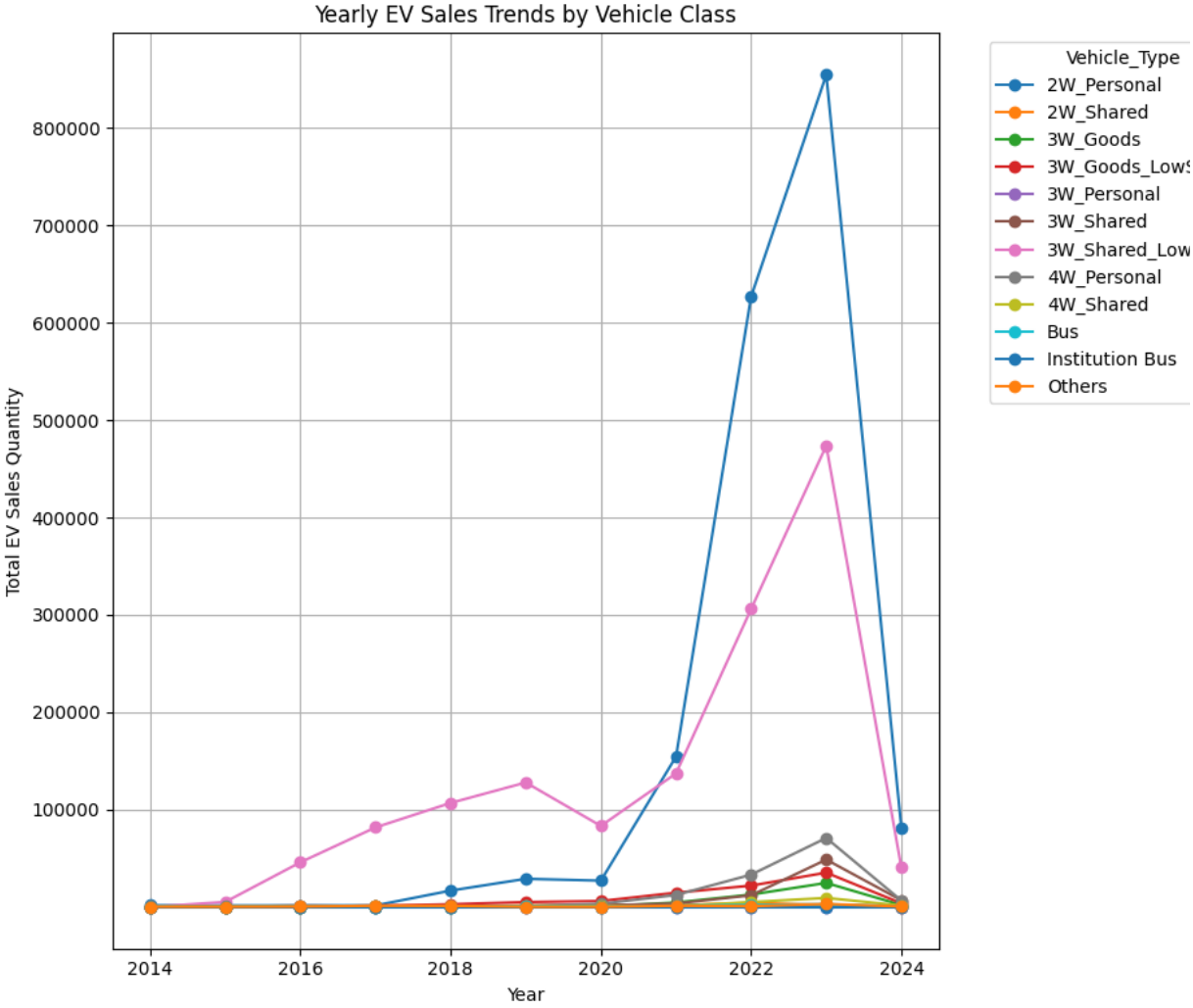
Yearly Sales Trends for Different Vehicle Classes

Compare trends of EV sales over years for different vehicle classes.

In [20]:
```
#Yearly sales trends for different vehicle classes:-

plt.figure(figsize=(10, 8))
class_trends = data.groupby(['Year', 'Vehicle_Type'])['EV_Sales_Quantity']
class_trends.plot(kind='line', marker='o', figsize=(10, 8))
plt.title("Yearly EV Sales Trends by Vehicle Class")
plt.xlabel("Year")
plt.ylabel("Total EV Sales Quantity")
plt.yticks([100000,200000,300000,400000,500000,600000,700000,800000])
plt.legend(title="Vehicle_Type", bbox_to_anchor=(1.05, 1), loc='upper lef
plt.grid()
plt.tight_layout()
plt.show()
```

```
<Figure size 1000x800 with 0 Axes>
```



Inference

The graph shows the yearly EV sales trends across various vehicle classes, with rapid growth pa in the 2-wheeler and 3-wheeler personal and shared vehicle segments. It highlights the significar increase in EV adoption across multiple vehicle types over the past decade.

EV Sales by Vehicle Category

Visualize the distribution of sales across different vehicle categories.

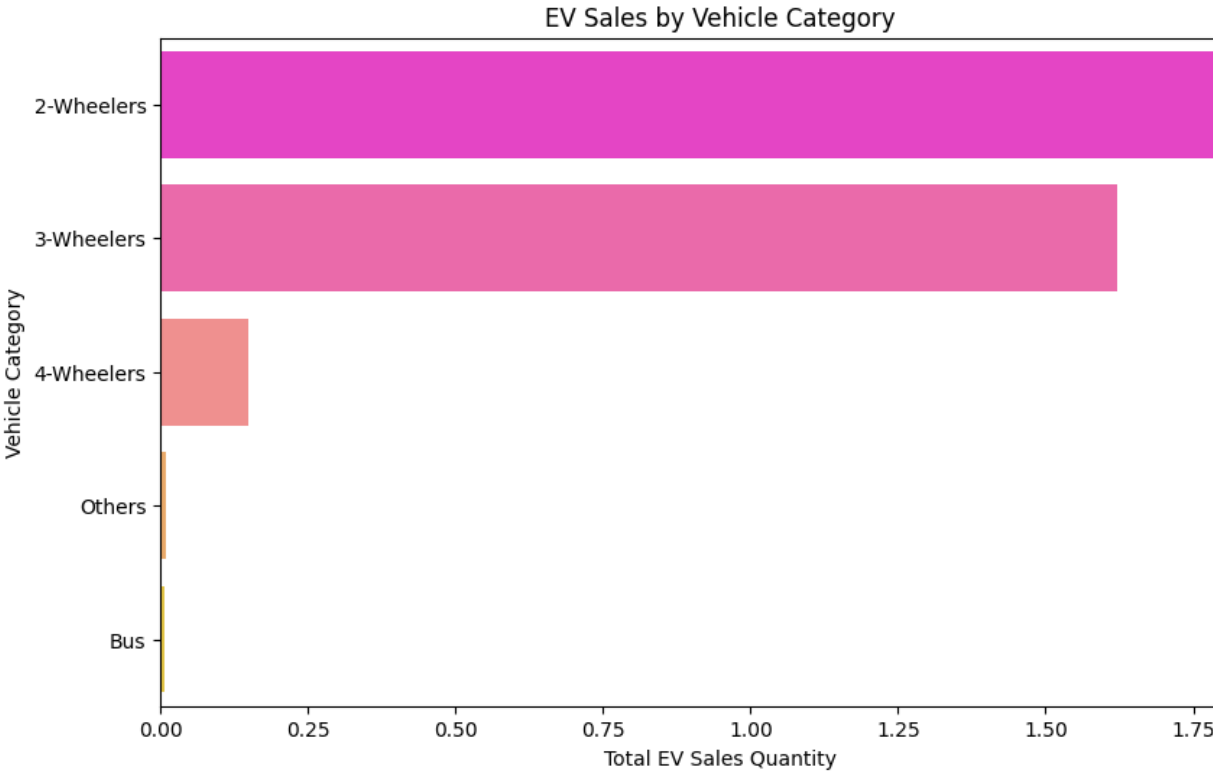
In [23]:
```
#EV sales by vehicle category:-

plt.figure(figsize=(10, 6))
vehicle_category_sales = data.groupby('Vehicle_Category')['EV_Sales_Quant
sns.barplot(x=vehicle_category_sales.values, y=vehicle_category_sales.inde
plt.title("EV Sales by Vehicle Category")
plt.xlabel("Total EV Sales Quantity")
plt.ylabel("Vehicle Category")
plt.show()
```

```
C:\Users\Abhijeet Kuanr\AppData\Local\Temp\ipykernel_15400\269902814.py:5:
FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed i
v0.14.0. Assign the `y` variable to `hue` and set `legend=False` for the same
effect.

  sns.barplot(x=vehicle_category_sales.values, y=vehicle_category_sales.index
palette='spring')
```



Inference

The graph shows EV sales by vehicle category. 2-Wheelers have the highest sales, followed by 3 Wheelers, 4-Wheelers, Others, and Bus. 2-Wheelers have significantly higher sales than the oth categories.

Conclusion

This analysis provided valuable insights into the evolving electric vehicle (EV) landscape across states. By visualizing sales trends over time and exploring vehicle categories, we identified patte adoption and usage. The correlation analysis revealed important relationships between numeric variables, offering clues about factors influencing EV sales.

Using a predictive model, we successfully estimated EV sales quantities with a mean absolute er units (replace with actual value), demonstrating the potential of data-driven approaches to anticip market behavior.

Future analyses could explore the effects of government policies, infrastructure development, an economic factors on EV adoption. Diving deeper into regional differences could also unveil uniqu opportunities and challenges across states.