Problem Statement

The electric vehicle market has witnessed significant growth, and with increasing interest in sustainable transportation, it is crucial to analyse and understand the performance and characteristics of electric cars. However, there is a lack of simple datasets on electric vehicles that can facilitate data science analysis and visualization. The problem at hand is to explore the provided Electric Vehicle dataset and address the following key questions:

1. Fastest 0-100 Acceleration: Determine which electric vehicle in the dataset has the fastest acceleration time from 0 to 100 km/h (0-60 mph). This analysis will help identify high-performance electric cars in terms of acceleration.

2. Highest Efficiency: Identify the electric vehicle with the highest efficiency in terms of energy consumption or range per charge. This insight is vital for understanding which electric car can travel the longest distance on a single charge or consume the least amount of energy.

3. Impact of Powertrain on Performance: Investigate whether the powertrain type (e.g., BEV, PHEV, or HEV) affects the vehicle's range, top speed, and efficiency. This

analysis will provide valuable information for consumers and manufacturers regarding the trade-offs between different powertrain options.

4. Manufacturer Analysis: Determine which manufacturer has the most number of electric vehicles in the dataset. This analysis will shed light on the dominant players in the electric vehicle market.

5. Price and Rapid Charging Relationship: Explore the relationship between the vehicle's price and its rapid charging capability. Understand if vehicles with higher prices offer better rapid charging options and if there is a correlation between price and rapid charging capabilities.

Objective:

The primary objective of this analysis is to gain insights into the performance, efficiency, and characteristics of electric vehicles using the provided dataset. By addressing the above questions, we aim to provide valuable information to various stakeholders:

- Consumers: Help potential electric car buyers make informed decisions based on performance metrics, efficiency, and pricing.

- Manufacturers: Understand how different powertrain types influence the performance of electric vehicles and identify areas for improvement in terms of efficiency and rapid charging capabilities.

- Policy Makers: Gain insights into the current state of the electric vehicle market and assess the impact of electric mobility on transportation sustainability.

Expected Outcome:

The analysis of the Electric Vehicle dataset will provide clear and concise answers to the questions posed, highlighting the top-performing electric cars, the most efficient models, and the influence of powertrain types on vehicle performance. Additionally, the analysis will identify the leading manufacturers in the electric vehicle market and shed light on the relationship between vehicle price and rapid charging capabilities. The outcome will serve as a valuable resource for data-driven decision-making, fostering the adoption and development of electric vehicles to achieve a more sustainable future in the automotive industry.

Data pre-processing

Required libraries

In order to perform EDA and clustering on the collected data, the following Python libraries are

used:

1. Pandas: for data handling/manipulation

2. Matplotlib and Seaborn: for data visualization

3. Numpy: to perform mathematical functions

4. Statmodels: for statical data exploration

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sb
import statsmodels.api as sm
```

Import the csv file

```
df= pd.read_csv('../input/evs-one-electric-vehicle-
dataset/ElectricCarData_Clean.csv')
```

From the following dataset we will find

1. Frequency of the Brand

2. Top speed of the car

3. Range of the car

4. Efficiency of the cars

5. Number of seats in a car

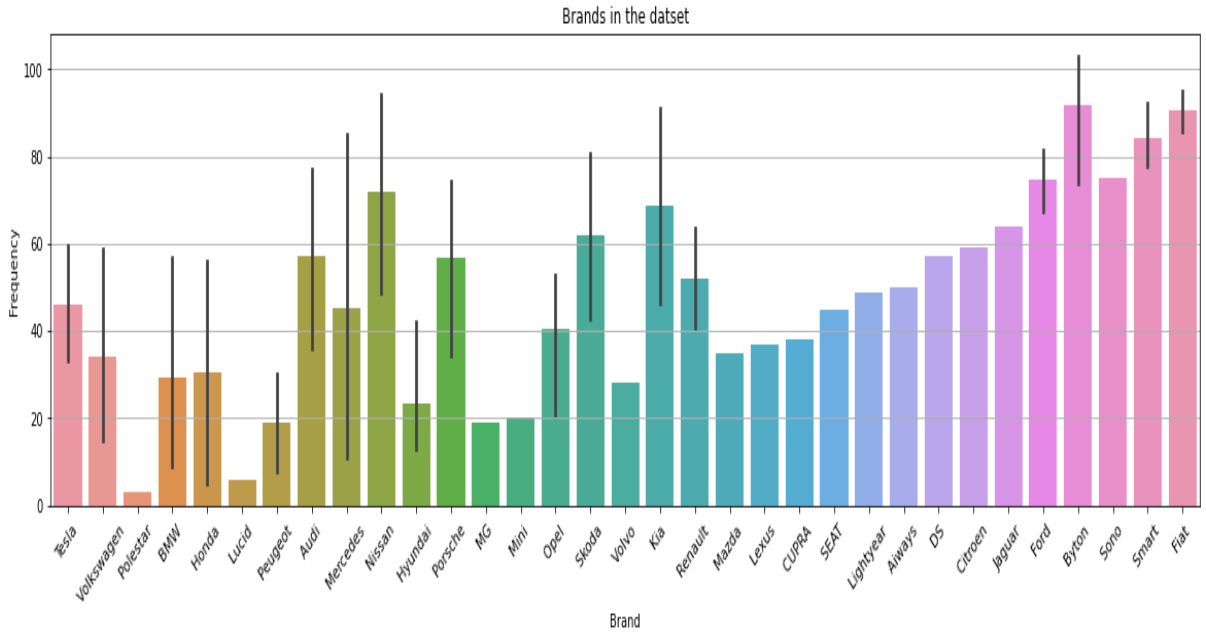
6. Price of the car (in euro , it can be converted into rupees as in the dataset it is given in euro)

7. Type of plug used for charging

8. Type of body of the car

Frequency of the Brand in the dataset

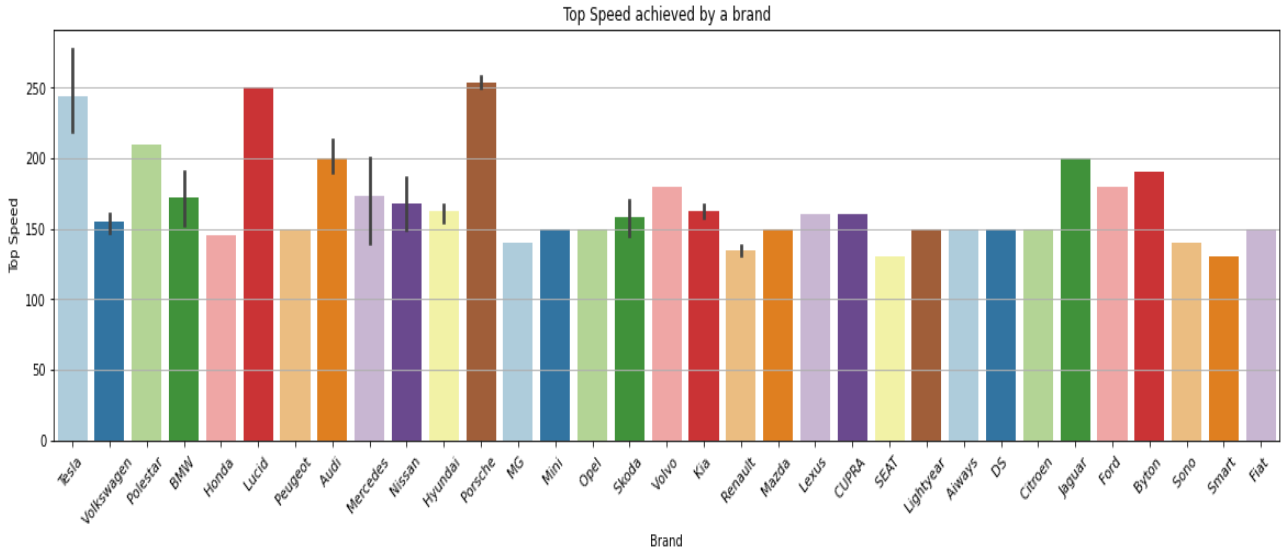
```
ax= plt.figure(figsize=(20,5))
sb.barplot(x='Brand',y=a,data=df)
plt.grid(axis='y')
plt.title('Brands in the datset')
plt.xlabel('Brand')
plt.ylabel('Frequency')
plt.xticks(rotation=45)
```



Byton , Fiat and smart are the prominent brands and Polestar being the least

Top speeds achieved by the cars of a brand

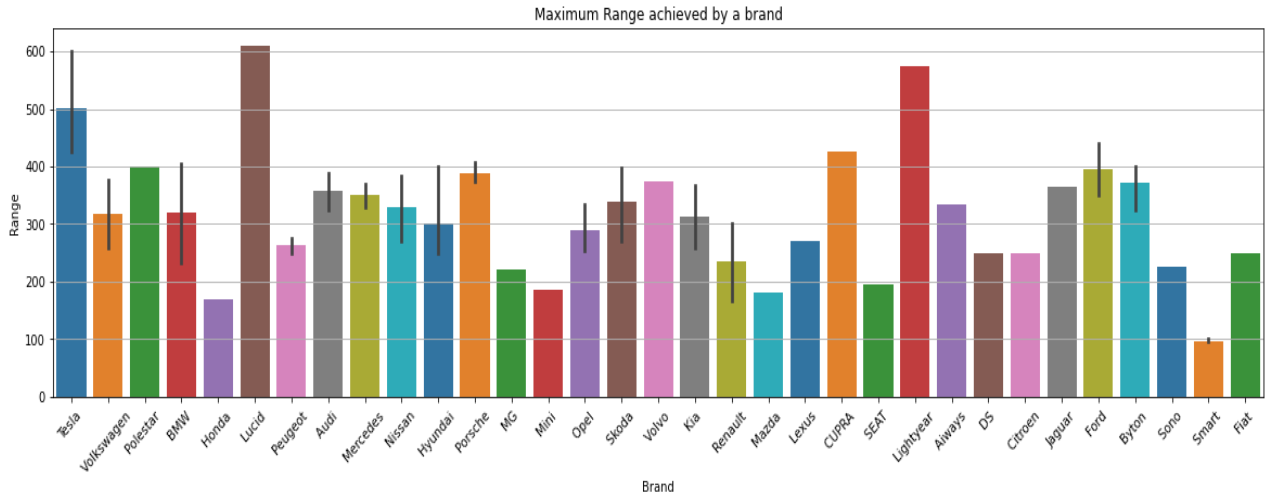
```
ax= plt.figure(figsize=(20,5))
sb.barplot(x='Brand',y='TopSpeed_KmH',data=df,palette='Paired')
plt.grid(axis='y')
plt.title('Top Speed achieved by a brand')
plt.xlabel('Brand')
plt.ylabel('Top Speed')
plt.xticks(rotation=45)
```



Porsche, Lucid and Tesla produce the fastest cars and Smart the lowest.

Range a car can achieve

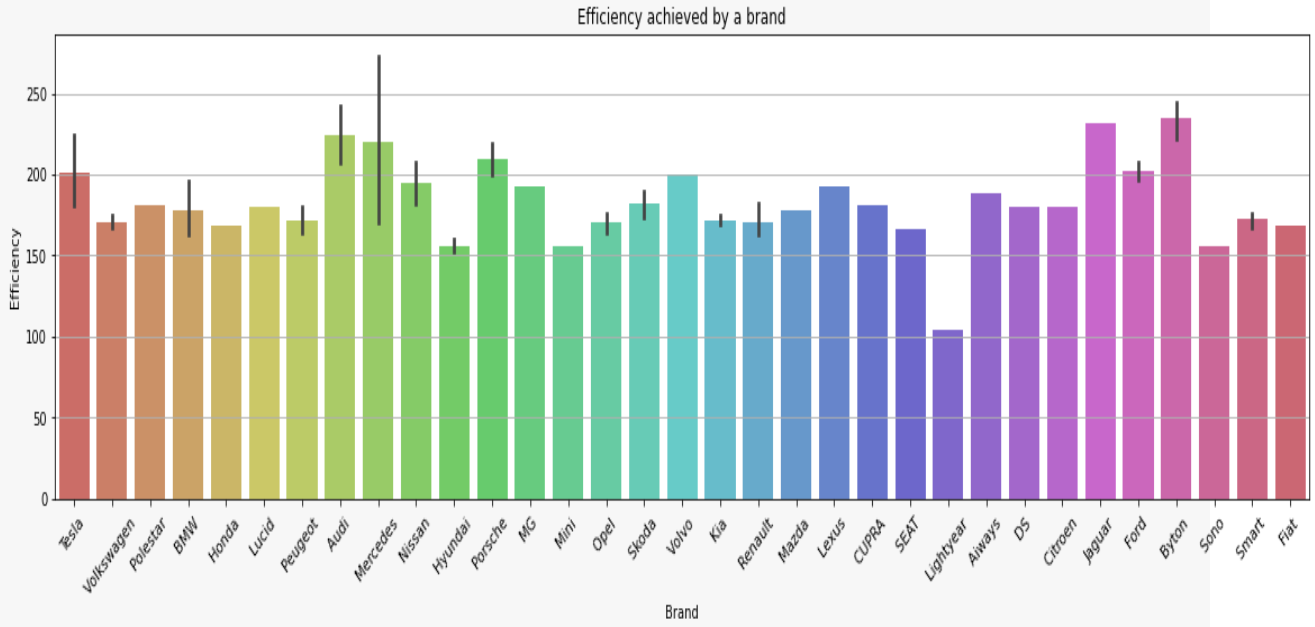
```
ax= plt.figure(figsize=(20,5))
sb.barplot(x='Brand',y='Range_Km',data=df,palette='tab10')
plt.grid(axis='y')
plt.title('Maximum Range achieved by a brand')
plt.xlabel('Brand')
plt.ylabel('Range')
plt.xticks(rotation=45)
```



Lucid, Lightyear and Tesla have the highest range and Smart the lowest.

Car efficiency

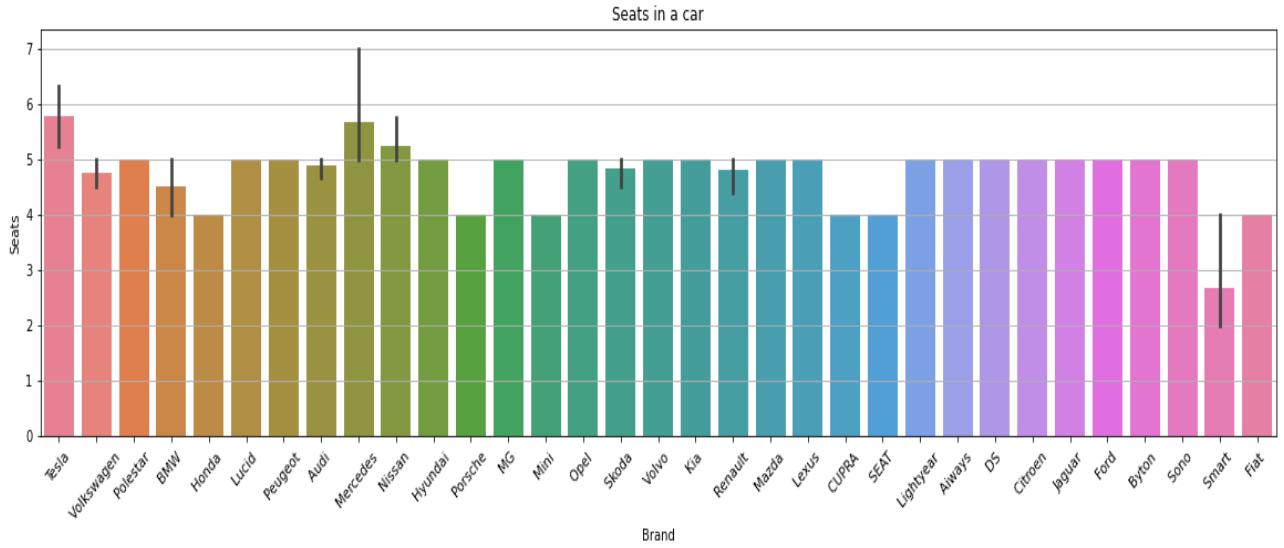
```
ax= plt.figure(figsize=(20,5))
sb.barplot(x='Brand',y='Efficiency_WhKm',data=df,palette='hls')
plt.grid(axis='y')
plt.title('Efficiency achieved by a brand')
plt.xlabel('Brand')
plt.ylabel('Efficiency')
plt.xticks(rotation=45)
```



Byton , Jaguar and Audi are the most efficient and Lightyear the least

Number of seats in each car

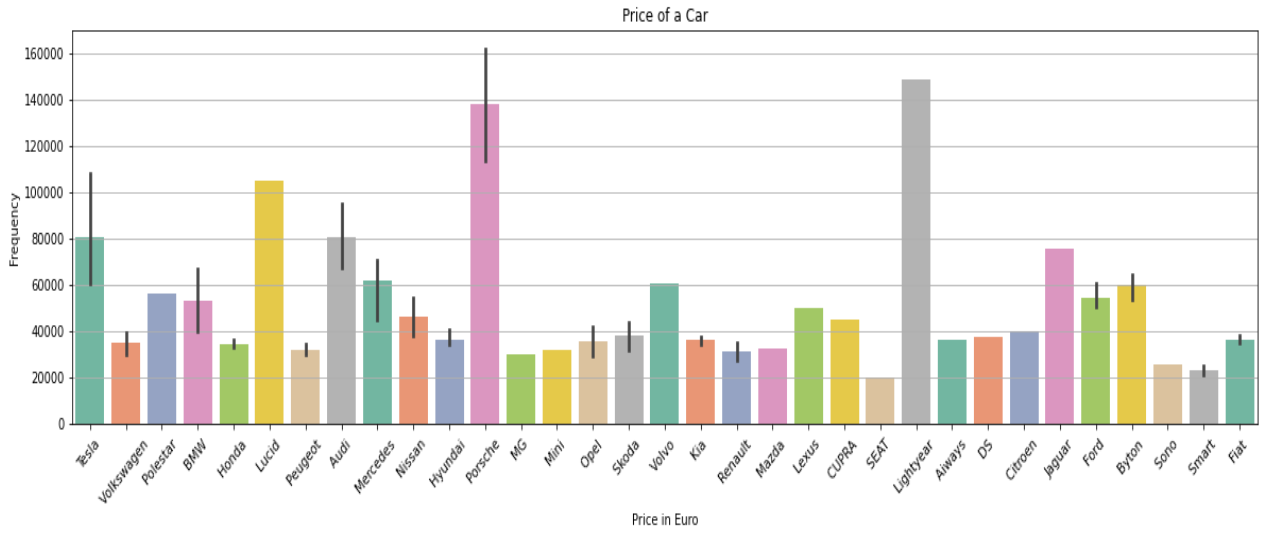
```
ax= plt.figure(figsize=(20,5))
sb.barplot(x='Brand',y='Seats',data=df,palette='husl')
plt.grid(axis='y')
plt.title('Seats in a car')
plt.xlabel('Brand')
plt.ylabel('Seats')
plt.xticks(rotation=45)
```



Mercedes, Tesla and Nissan have the highest number of seats and Smart the lowest

Price of cars (in Euro , it can converted into indian currency)

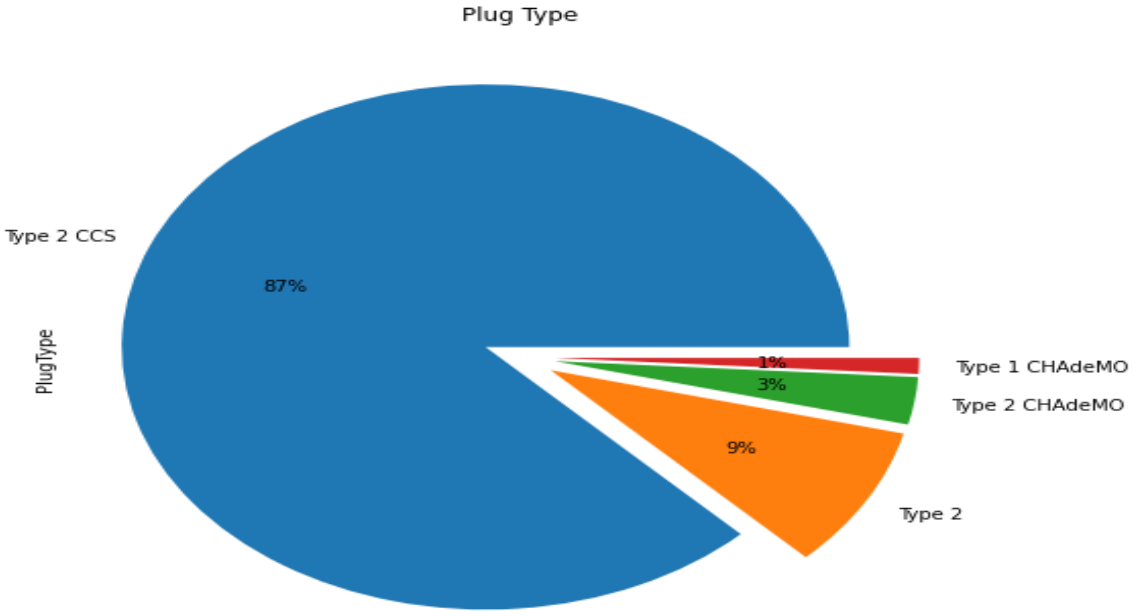
```
ax= plt.figure(figsize=(20,5))
sb.barplot(x='Brand',y='PriceEuro',data=df,palette='Set2')
plt.title('Price of a Car')
plt.xlabel('Price in Euro')
plt.grid(axis='y')
plt.ylabel('Frequency')
plt.xticks(rotation=45)
```



Lightyear, Porsche and Lucid are the most expensive and SEAT and Smart the least

Type of Plug used for charging

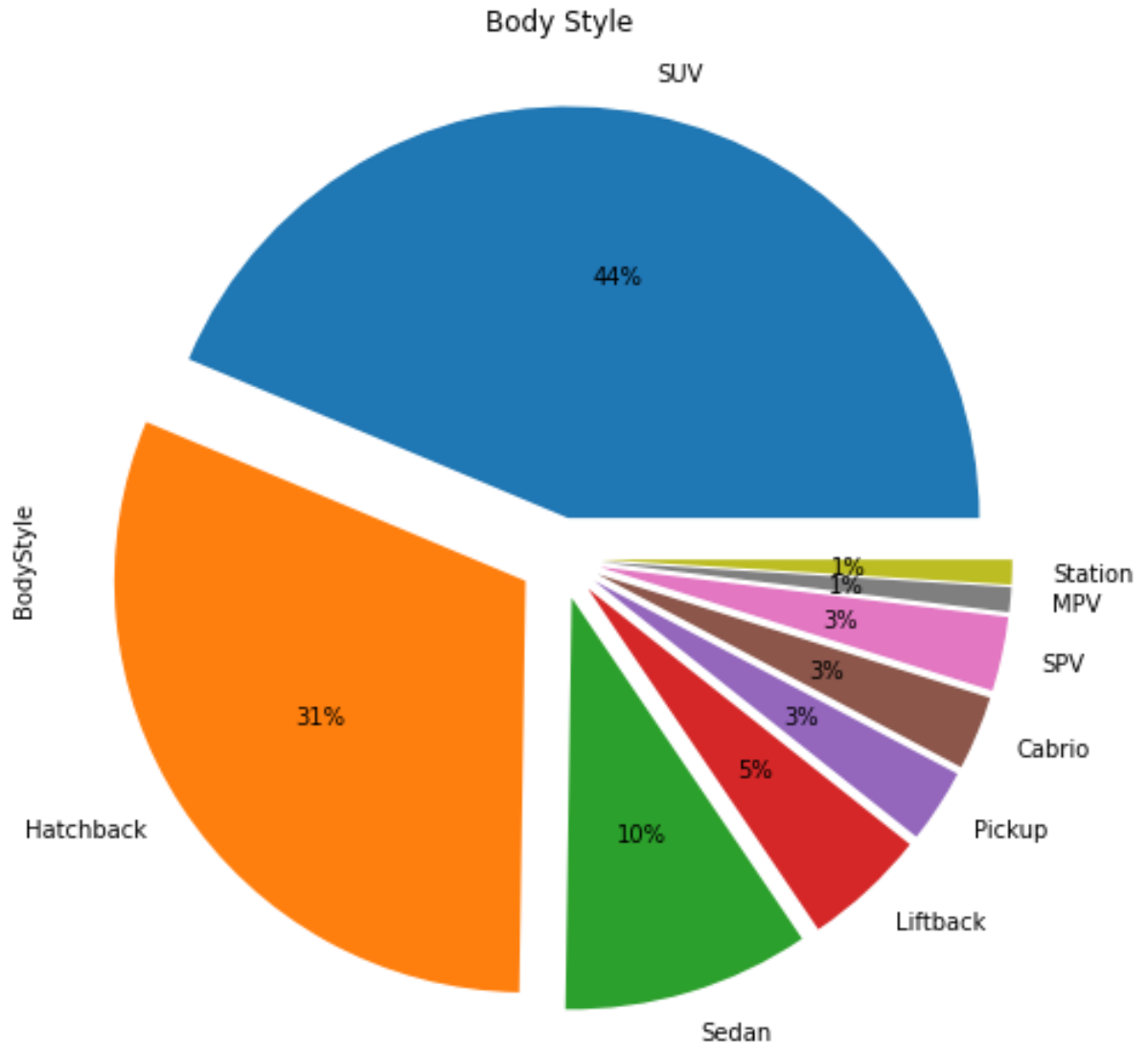
```
df['PlugType'].value_counts().plot.pie(figsize=(8,15),autopct='%.0f%%',
explode=(.1,.1,.1,.1))
plt.title('Plug Type')
```



Most companies use Type 2 CCS and Type 1 CHAdeMo the least

Cars and their body style

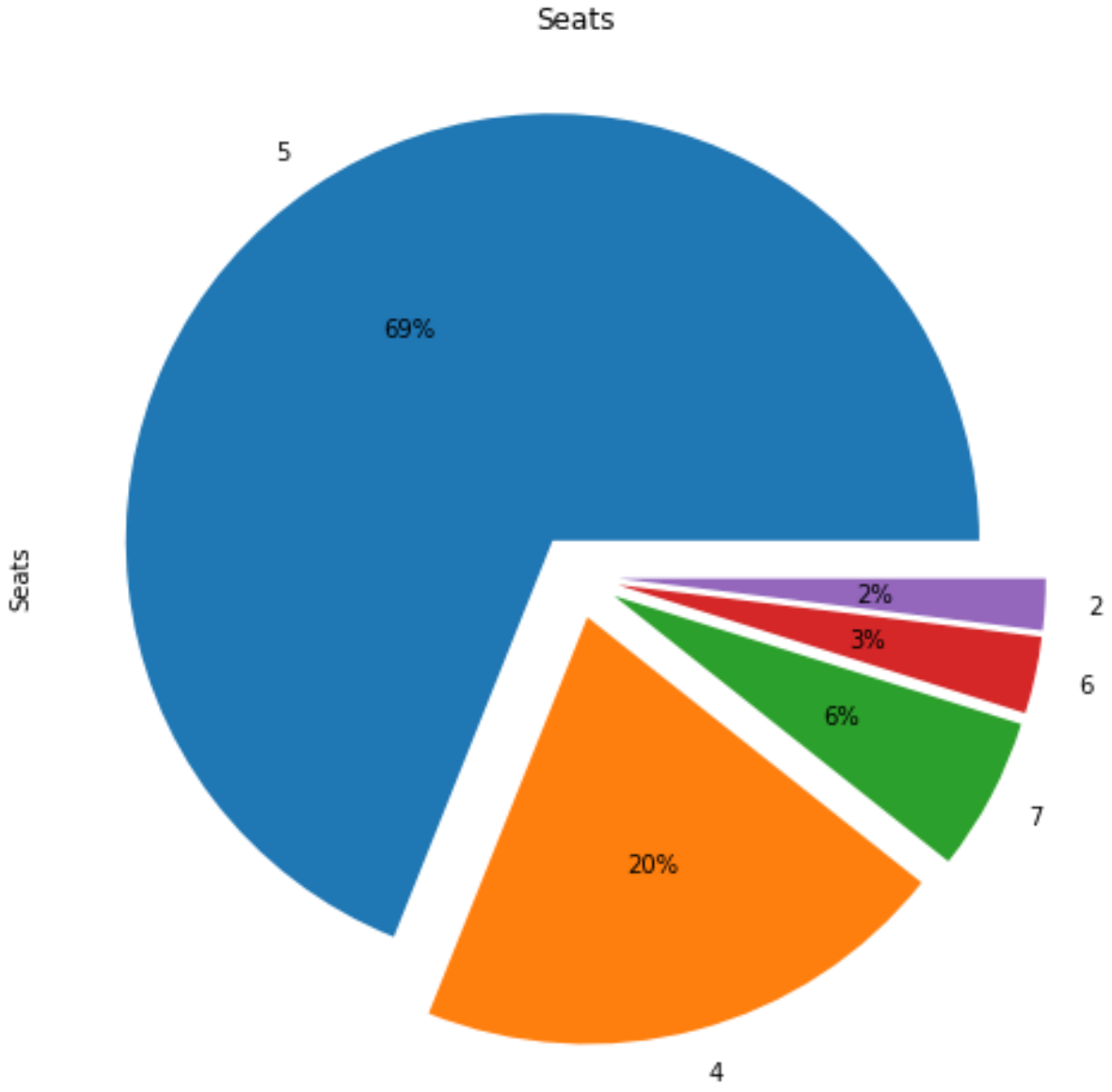
```
df['BodyStyle'].value_counts().plot.pie(figsize=(8,15),autopct='%.0f%%'
,explode=(0.1,0.1,0.1,0.1,0.1,0.1,0.1,0.1,0.1,0.1))
plt.title('Body Style')
```



Most cars are either SUV or Hatchback

Number of Seats

```
df['Seats'].value_counts().plot.pie(figsize=(8,15),autopct='%.0f%%',exp
lode=(0.1,0.1,0.1,0.1,0.1))
plt.title('Seats')
```



Majority of cars have 5 seats

References

Dataset - https://www.kaggle.com/datasets/geoffnel/evs-one-electric-vehicle-dataset