

TASK 1:Apply Exploratory Data Analysis(Univariate and Bivariate) using plotly.express library.

```
In [1]: import pandas as pd
import numpy as np

df=pd.read_csv(r"C:\Users\Surj\Downloads\dataset.csv")

In [2]: df.head()
```

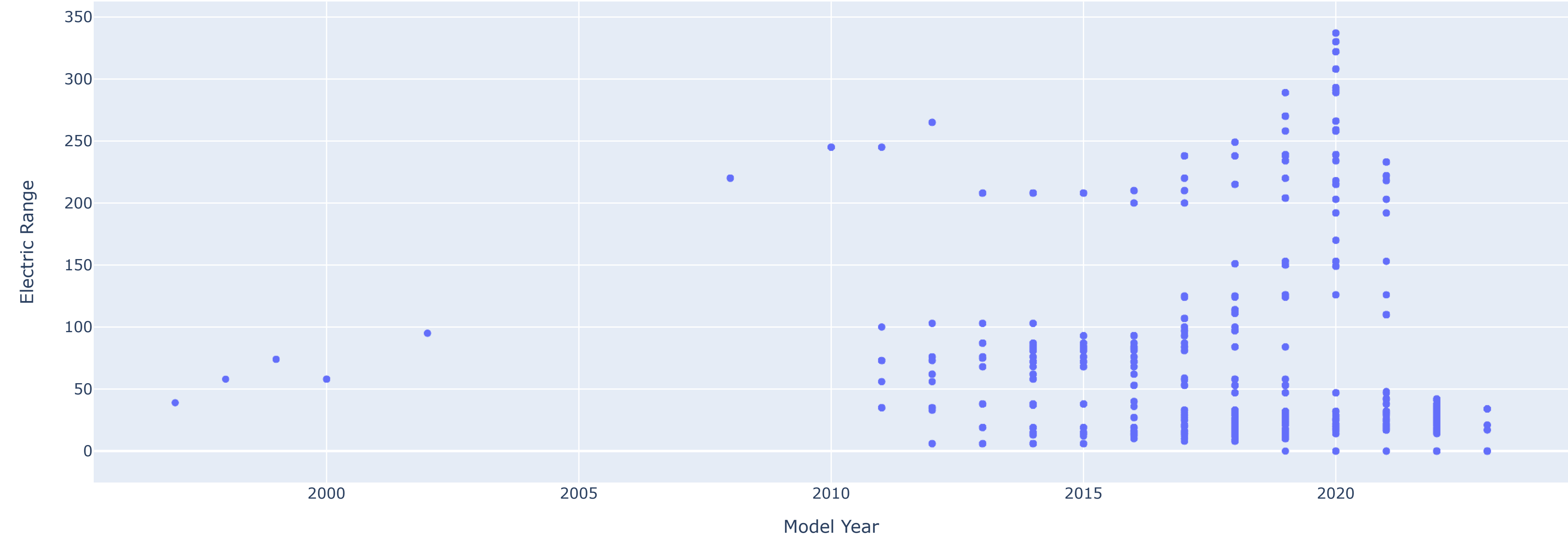
	VIN (1-10)	County	City	State	Postal Code	Model Year	Make	Model	Electric Vehicle Type	Clean Alternative Fuel Vehicle (CAFV) Eligibility	Electric Range	Base MSRP	Legislative District	DOL Vehicle ID	Vehicle Location	Electric Utility	2020 Census Tract
0	JTMEB3FV6N	Monroe	Key West	FL	33040	2022	TOYOTA	RAV4 PRIME	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	42	0	NaN	198968248	POINT (-81.80023 24.5545)	NaN	12087972100
1	1G1RD6E45D	Clark	Laughlin	NV	89029	2013	CHEVROLET	VOLT	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	38	0	NaN	5204412	POINT (-114.57245 35.16815)	NaN	32003005702
2	JN1AZ0CP8B	Yakima	Yakima	WA	98901	2011	NISSAN	LEAF	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	73	0	15.0	218972519	POINT (-120.50721 46.60448)	PACIFICORP	53077001602
3	1G1FW6S08H	Skagit	Concrete	WA	98237	2017	CHEVROLET	BOLT EV	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	238	0	39.0	186750406	POINT (-121.7515 48.53892)	PUGET SOUND ENERGY INC	53057951101
4	3FA6P0SU1K	Snohomish	Everett	WA	98201	2019	FORD	FUSION	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	26	0	38.0	2006714	POINT (-122.20596 47.97659)	PUGET SOUND ENERGY INC	53061041500

```
In [3]: import plotly.express as px

In [4]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 112634 entries, 0 to 112633
Data columns (total 17 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   VIN (1-10)                            112634 non-null object
1   County                                112634 non-null object
2   City                                  112634 non-null object
3   State                                 112634 non-null object
4   Postal Code                           112634 non-null int64
5   Model Year                            112634 non-null int64
6   Make                                  112634 non-null object
7   Model                                 112614 non-null object
8   Electric Vehicle Type                  112634 non-null object
9   Clean Alternative Fuel Vehicle (CAFV) Eligibility 112634 non-null object
10  Electric Range                         112634 non-null int64
11  Base MSRP                             112634 non-null int64
12  Legislative District                   112348 non-null float64
13  DOL Vehicle ID                        112634 non-null int64
14  Vehicle Location                       112610 non-null object
15  Electric Utility                       112191 non-null object
16  2020 Census Tract                     112634 non-null int64
dtypes: float64(1), int64(6), object(10)
memory usage: 14.6+ MB
```

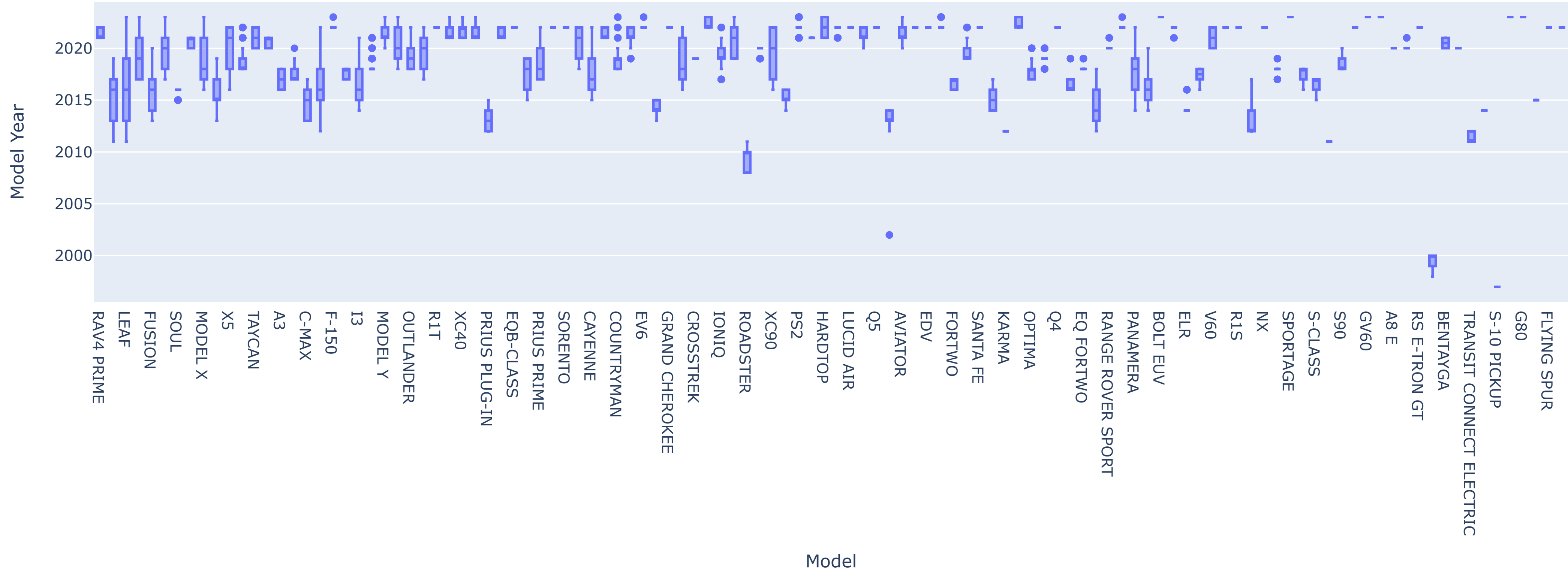
```
In [5]: #scatter plot
px.scatter(df,x="Model Year",y="Electric Range")
```



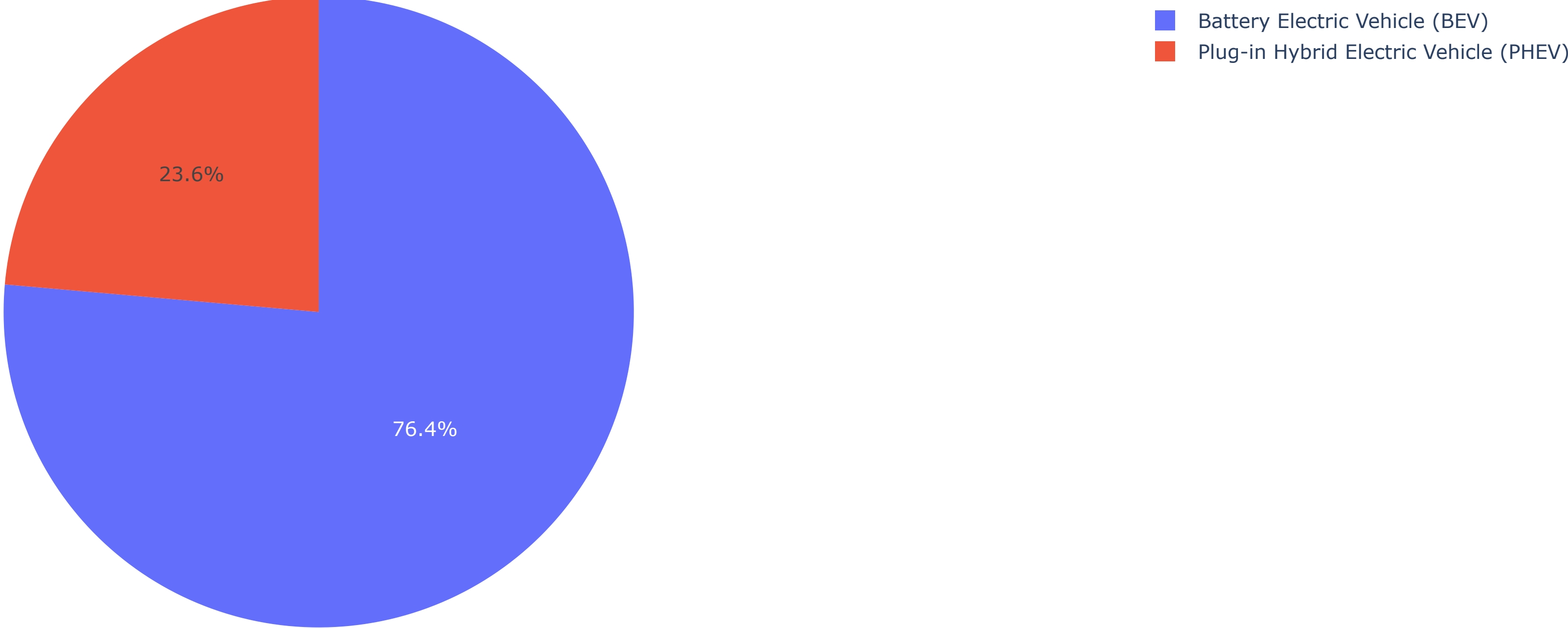
Observation : Electric range of vechiles got increased till year 2020 , high in year 2020 and got decreased in next 3 years

```
In [6]: # remvoing all the rows which contains null values for the Model
df=df[df["Model"].notnull()]

In [6]: #box plot
px.box(df,x="Model",y="Model Year")
```



```
In [7]: #pie chart
px.pie(df,names="Electric Vehicle Type")
```



Observation : In electric vechiles battery electric vechiles are of huge percentage(76.4%)

```
In [ ]:
```