

# Exploratory Data Analysis On Electric Vehicle





## What is Exploratory Data Analysis?

- Exploratory Data Analysis (EDA) is the process of analyzing a dataset in order to understand its main characteristics, patterns and identify anomalies. EDA is often the first step in the data analysis process.
  - It involves using different graphs and plots to help visualise the data and also uses statistical methods to draw inferences from the data.
  - The goal of EDA is not to arrive at a certain right answer or to confirm a pre-defined hypothesis. It is an exploratory process to draw inferences and get ideas on how the data can be further utilised to predict certain outcomes/develop ML models
- 
- An electric vehicle (EV) is a vehicle that uses one or more electric motors for propulsion. It can be powered by a collector system, with electricity from extravehicular sources, or it can be powered autonomously by a battery (sometimes charged by solar panels, or by converting fuel to electricity using fuel cells or a generator).
  - EVs include, but are not limited to, road and rail vehicles, surface and underwater vessels, electric aircraft, and electric spacecraft.
  - For road vehicles, together with other emerging automotive technologies such as autonomous driving, connected vehicles, and shared mobility, EVs form a future mobility vision called Connected, Autonomous, Shared, and Electric (CASE) Mobility.
  - EVs first came into existence in the late 19th century, when electricity was among the preferred methods for motor vehicle propulsion, providing a level of comfort and ease of operation that could not be achieved by the gasoline cars of the time.
  - Internal combustion engines were the dominant propulsion method for cars and trucks for about 100 years, but electric power remained commonplace in other vehicle types, such as trains and smaller vehicles of all types.
- 
- Data set link : <https://drive.google.com/file/d/1P742LU5OTXbfFG2F6drbABk1O8UGf4Cd/view?usp=sharing>  
(<https://drive.google.com/file/d/1P742LU5OTXbfFG2F6drbABk1O8UGf4Cd/view?usp=sharing>).

## About Dataset

This dataset shows the Battery Electric Vehicles (BEVs) and Plug-in Hybrid Electric Vehicles (PHEVs) that are currently registered through the Washington State Department of Licensing (DOL).

1.A Battery Electric Vehicle (BEV) is an all-electric vehicle using one or more batteries to store the electrical energy that powers the motor and is charged by plugging the vehicle into an electric power source.

2 Alternative Fuel Vehicle (CAFV) Eligibility is based on the fuel requirement and electric-only range requirement as outlined in RCW 82.08.809 and RCW 82.12.809 to be eligible for Alternative Fuel Vehicles retail sales and Washington State use tax exemptions.

3.Monthly count of vehicles for a county may change from this report and prior reports. Processes were implemented to more accurately assign county at the time of registration.

4.Electric Range is no longer maintained for Battery Electric Vehicles (BEV) because new BEVs have an electric range of 30 miles or more. Zero (0) will be entered where the electric range has not been researched.

5.Field 'Electric Utility' was added starting with the publication in March 2022.

6.Field '2020 Census Tract' was added starting with the publication in June 2022.

In [ ]:

## Importing Required Libraries

```
In [1]: import pandas as pd
import numpy as np
import plotly.express as px
import warnings
warnings.filterwarnings("ignore")
import matplotlib.pyplot as plt

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

df
```

Out[1]:

	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	Vehicle
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	1989€
1	1G1RD6E45D	Clark	Laughlin	NV	89029	2013	CHEVROLET	VOLT	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	38	0	NaN	52€
2	JN1AZ0CP8B	Yakima	Yakima	WA	98901	2011	NISSAN	LEAF	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	73	0	15.0	21897
3	1G1FW6S08H	Skagit	Concrete	WA	98237	2017	CHEVROLET	BOLT EV	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	238	0	39.0	18675
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	20€
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
112629	7SAYGDEF2N	King	Duvall	WA	98019	2022	TESLA	MODEL Y	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0	45.0	21795
112630	1N4BZ1CP7K	San Juan	Friday Harbor	WA	98250	2019	NISSAN	LEAF	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	150	0	40.0	1036€

	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	Vehic
112631	1FMCU0KZ4N	King	Vashon	WA	98070	2022	FORD	ESCAPE	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	38	0	34.0	19387
112632	KNDCD3LD4J	King	Covington	WA	98042	2018	KIA	NIRO	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	26	0	47.0	12503
112633	YV4BR0CL8N	King	Covington	WA	98042	2022	VOLVO	XC90	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	18	0	47.0	19467

112634 rows × 17 columns

In [2]: 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 [3]: df.duplicated().sum()

Out[3]: 0

```
In [4]: print(len(df.columns))  
df.columns
```

17

```
Out[4]: Index(['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'],  
             dtype='object')
```



In [5]: `df.head()`

Out[5]:

	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
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
1	1G1RD6E45D	Clark	Laughlin	NV	89029	2013	CHEVROLET	VOLT	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	38	0	NaN	5204412
2	JN1AZ0CP8B	Yakima	Yakima	WA	98901	2011	NISSAN	LEAF	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	73	0	15.0	218972519
3	1G1FW6S08H	Skagit	Concrete	WA	98237	2017	CHEVROLET	BOLT EV	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	238	0	39.0	186750406
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

In [6]: `df.shape`

Out[6]: (112634, 17)

## unique values

```
In [7]: cols = df.columns
def Unique_Values():
    for i in np.arange(0,len(cols)):
        print('{} column have {} number of unique values out of {}'.format( cols[i],df[cols[i]].nunique(), len(df)),end='\n')
Unique_Values()
```

VIN (1-10) column have 7548 number of unique values out of 112634

County column have 165 number of unique values out of 112634

City column have 629 number of unique values out of 112634

State column have 45 number of unique values out of 112634

Postal Code column have 773 number of unique values out of 112634

Model Year column have 20 number of unique values out of 112634

Make column have 34 number of unique values out of 112634

Model column have 114 number of unique values out of 112634

Electric Vehicle Type column have 2 number of unique values out of 112634

Clean Alternative Fuel Vehicle (CAFV) Eligibility column have 3 number of unique values out of 112634

Electric Range column have 101 number of unique values out of 112634

Base MSRP column have 30 number of unique values out of 112634

Legislative District column have 49 number of unique values out of 112634

DOL Vehicle ID column have 112634 number of unique values out of 112634

Vehicle Location column have 758 number of unique values out of 112634

Electric Utility column have 73 number of unique values out of 112634

2020 Census Tract column have 2026 number of unique values out of 112634

## Null values

```
In [8]: cols = df.columns
def Null_Values():
    for i in np.arange(0,len(cols)):
        print('{} column have {} number of Null values out of {}'.format( cols[i],df[cols[i]].isnull().sum(), len(df)
Null_Values()
```

VIN (1-10) column have 0 number of Null values out of 112634

County column have 0 number of Null values out of 112634

City column have 0 number of Null values out of 112634

State column have 0 number of Null values out of 112634

Postal Code column have 0 number of Null values out of 112634

Model Year column have 0 number of Null values out of 112634

Make column have 0 number of Null values out of 112634

Model column have 20 number of Null values out of 112634

Electric Vehicle Type column have 0 number of Null values out of 112634

Clean Alternative Fuel Vehicle (CAFV) Eligibility column have 0 number of Null values out of 112634

Electric Range column have 0 number of Null values out of 112634

Base MSRP column have 0 number of Null values out of 112634

Legislative District column have 286 number of Null values out of 112634

DOL Vehicle ID column have 0 number of Null values out of 112634

Vehicle Location column have 24 number of Null values out of 112634

Electric Utility column have 443 number of Null values out of 112634

2020 Census Tract column have 0 number of Null values out of 112634

```
In [9]: # to view the missing percentages
missing_percentages=df.isnull().sum()/len(df)
missing_percentages
```

```
Out[9]: VIN (1-10)          0.000000
County                    0.000000
City                      0.000000
State                     0.000000
Postal Code               0.000000
Model Year                0.000000
Make                      0.000000
Model                     0.000178
Electric Vehicle Type     0.000000
Clean Alternative Fuel Vehicle (CAFV) Eligibility 0.000000
Electric Range            0.000000
Base MSRP                 0.000000
Legislative District      0.002539
DOL Vehicle ID            0.000000
Vehicle Location          0.000213
Electric Utility           0.003933
2020 Census Tract         0.000000
dtype: float64
```

## Handling The Missing Values

- For handling the missing values we know the distributions of the variables by using statistics and vizualization techniques
- To fill the null values
- for numerical variables we use mean or median
- Mean is impact with outliers if ouliers present in the data we use median.
- if our data doesn't contain outliers then we use mean (to reduce the time complexity)
- for categorical(object)we use mode

**So in our data Model,Legislative District,Vehicle Location,Electric Utility columns having missing values.**

- numerical column - Legislative District

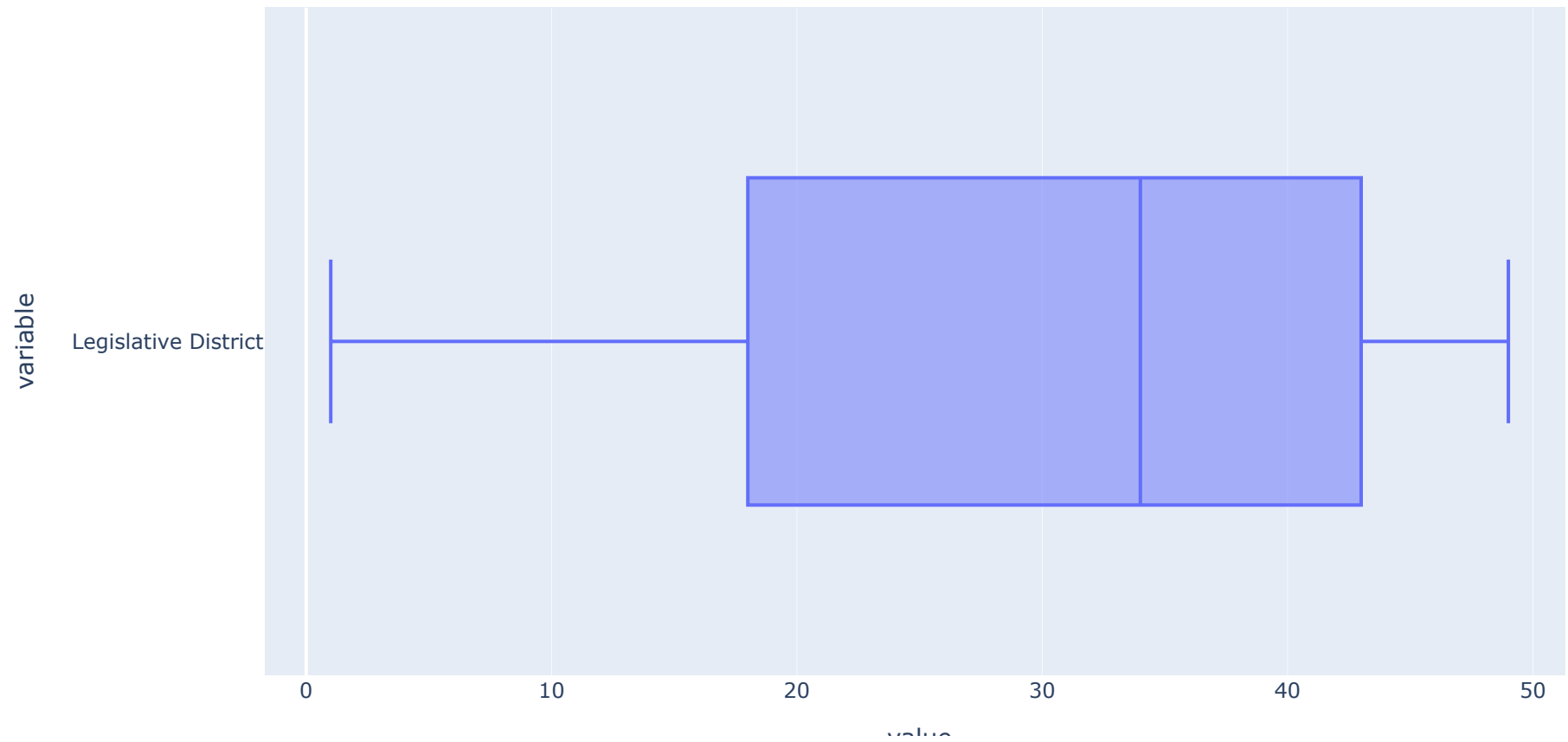
- categorical columns -Model, Vehicle Location,Electric Utility

In [10]: df.describe()

Out[10]:

	Postal Code	Model Year	Electric Range	Base MSRP	Legislative District	DOL Vehicle ID	2020 Census Tract
<b>count</b>	112634.000000	112634.000000	112634.000000	112634.000000	112348.000000	1.126340e+05	1.126340e+05
<b>mean</b>	98156.226850	2019.003365	87.812987	1793.439681	29.805604	1.994567e+08	5.296650e+10
<b>std</b>	2648.733064	2.892364	102.334216	10783.753486	14.700545	9.398427e+07	1.699104e+09
<b>min</b>	1730.000000	1997.000000	0.000000	0.000000	1.000000	4.777000e+03	1.101001e+09
<b>25%</b>	98052.000000	2017.000000	0.000000	0.000000	18.000000	1.484142e+08	5.303301e+10
<b>50%</b>	98119.000000	2020.000000	32.000000	0.000000	34.000000	1.923896e+08	5.303303e+10
<b>75%</b>	98370.000000	2022.000000	208.000000	0.000000	43.000000	2.191899e+08	5.305307e+10
<b>max</b>	99701.000000	2023.000000	337.000000	845000.000000	49.000000	4.792548e+08	5.603300e+10

```
In [11]: #for numerical columns we have to check distributions for this we find outliers  
px.box(df['Legislative District'],orientation='h')
```





## To check the outliers by using IQR method(statistical\_method)

```
In [12]: q1=df['Legislative District'].quantile(0.25)

q3=df['Legislative District'].quantile(0.75)

iqr=q3-q1

lb=q1-1.5*iqr
ub=q1+1.5*iqr

df[(df['Legislative District']<=lb) | (df['Legislative District']>=ub)]
```

Out[12]:

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
---------------	--------	------	-------	----------------	---------------	------	-------	-----------------------------	--	-------------------	--------------	-------------------------	----------------------	---------------------	---------------------	-------------------------

---

## Observation :

- Here also we can observe that there are no outliers in our data.

## Filling null values with mean.

```
In [13]: df['Legislative District']=df['Legislative District'].fillna(df['Legislative District'].mean())
```

```
In [14]: (df[df['Model'].isnull()])
```

Out[14]:

	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
13874	YV4ED3GM2P	King	Seattle	WA	98115	2023	VOLVO	NaN	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0	46.0	221526476
30517	YV4ED3UL3P	King	Seattle	WA	98115	2023	VOLVO	NaN	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0	43.0	223881556
31936	YV4ED3GM4P	Clallam	Sequim	WA	98382	2023	VOLVO	NaN	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0	24.0	219769000
37517	YV4ED3UW2P	Snohomish	Edmonds	WA	98026	2023	VOLVO	NaN	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0	32.0	218357779
58071	YV4ED3UM4P	King	Renton	WA	98058	2023	VOLVO	NaN	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0	11.0	224511766
61626	YV4ED3GM5P	Pierce	Tacoma	WA	98465	2023	VOLVO	NaN	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0	28.0	224496702
63240	YV4ED3GMXP	King	Redmond	WA	98052	2023	VOLVO	NaN	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0	48.0	221295224

	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
<b>63380</b>	YV4ED3GM7P	King	Seattle	WA	98122	2023	VOLVO	NaN	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0	37.0	224280472
<b>63462</b>	YV4ED3UW4P	King	Newcastle	WA	98059	2023	VOLVO	NaN	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0	41.0	218912410
<b>78472</b>	YV4ED3UM1P	King	Fall City	WA	98024	2023	VOLVO	NaN	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0	5.0	224631494
<b>81302</b>	YV4ED3UM5P	King	Redmond	WA	98052	2023	VOLVO	NaN	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0	48.0	220511791
<b>84142</b>	YV4ED3UM2P	King	North Bend	WA	98045	2023	VOLVO	NaN	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0	5.0	223998148
<b>86960</b>	YV4ED3UM9P	King	Sammamish	WA	98075	2023	VOLVO	NaN	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0	41.0	214714706
<b>88687</b>	YV4ED3GM5P	King	Maple Valley	WA	98038	2023	VOLVO	NaN	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0	5.0	224709726

	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
89882	YV4ED3UM5P	King	Bellevue	WA	98006	2023	VOLVO	NaN	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0	41.0	214731254
93197	YV4ED3GM8P	Snohomish	Bothell	WA	98021	2023	VOLVO	NaN	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0	1.0	220532063
103099	YV4ED3UW6P	Pierce	Milton	WA	98354	2023	VOLVO	NaN	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0	30.0	213335454
103394	YV4ED3GM5P	King	Seattle	WA	98133	2023	VOLVO	NaN	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0	46.0	220589967
108116	YV4ED3GL1P	King	Seattle	WA	98104	2023	VOLVO	NaN	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0	37.0	219268451
112622	YV4ED3GM0P	King	Covington	WA	98042	2023	VOLVO	NaN	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0	47.0	224307996

```
In [15]: crosstab=pd.crosstab(df['Make'],df['Model'])  
crosstab
```

Out[15]:

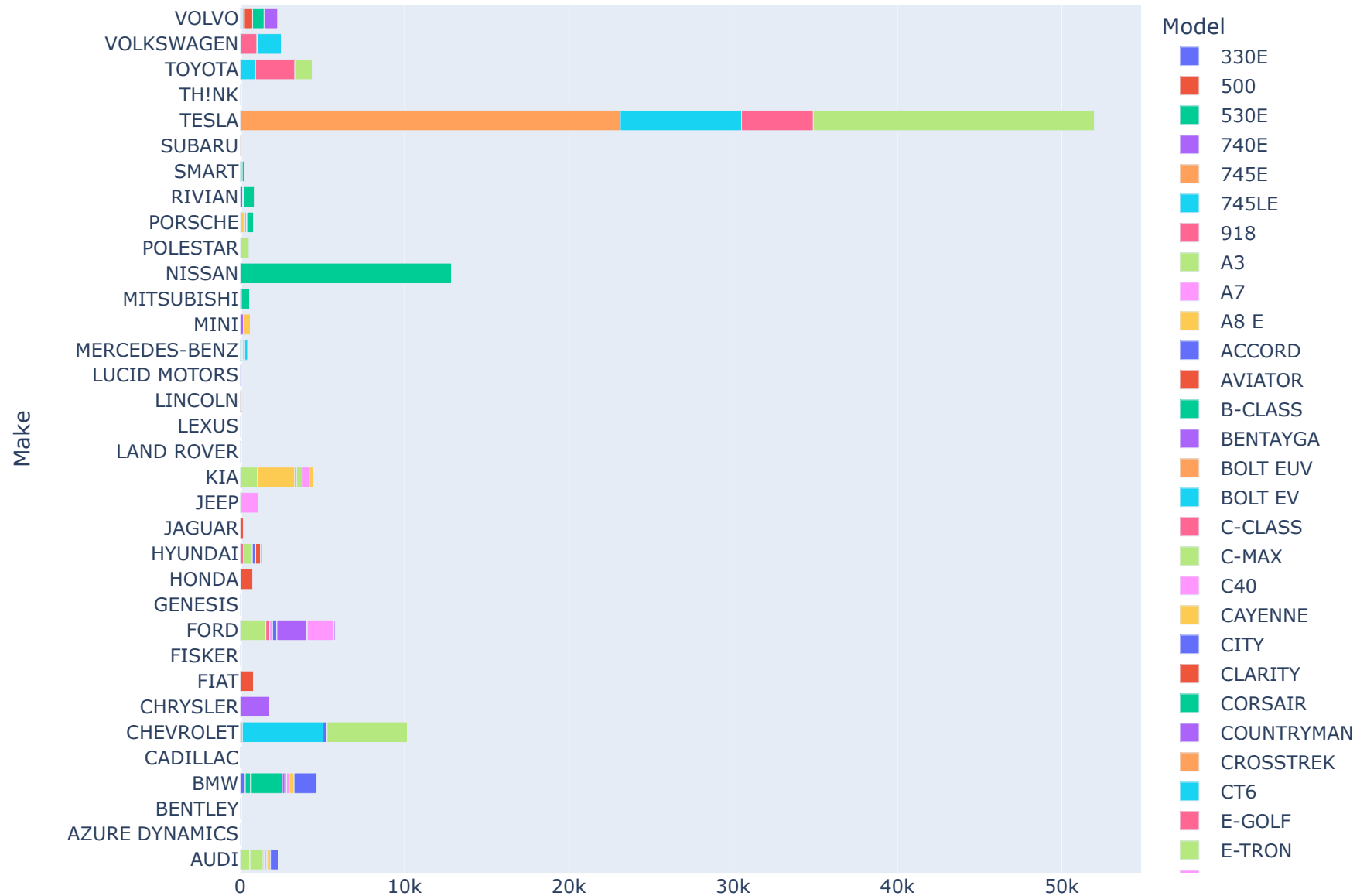
Model	330E	500	530E	740E	745E	745LE	918	A3	A7	A8 E	...	TRANSIT CONNECT ELECTRIC	TUCSON	V60	VOLT	WRANGLER	X3	X5	XC40	XC
Make																				
AUDI	0	0	0	0	0	0	0	575	11	3	...	0	0	0	0		0	0	0	0
AZURE DYNAMICS	0	0	0	0	0	0	0	0	0	0	...	7	0	0	0		0	0	0	0
BENTLEY	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
BMW	303	0	323	30	7	2	0	0	0	0	...	0	0	0	0		0	292	1407	0
CADILLAC	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
CHEVROLET	0	0	0	0	0	0	0	0	0	0	...	0	0	0	4896		0	0	0	0
CHRYSLER	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
FIAT	0	822	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
FISKER	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
FORD	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
GENESIS	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
HONDA	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
HYUNDAI	0	0	0	0	0	0	0	0	0	0	...	0	38	0	0		0	0	0	0
JAGUAR	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
JEEP	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		1104	0	0	0
KIA	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
LAND ROVER	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
LEXUS	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
LINCOLN	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
LUCID MOTORS	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
MERCEDES- BENZ	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
MINI	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0

Model	330E	500	530E	740E	745E	745LE	918	A3	A7	A8 E	...	TRANSIT CONNECT ELECTRIC	TUCSON	V60	VOLT	WRANGLER	X3	X5	XC40	XC
Make																				
MITSUBISHI	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
NISSAN	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
POLESTAR	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
PORSCHE	0	0	0	0	0	0	1	0	0	0	...	0	0	0	0		0	0	0	0
RIVIAN	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
SMART	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
SUBARU	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
TESLA	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
THINK	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
TOYOTA	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
VOLKSWAGEN	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0		0	0	0	0
VOLVO	0	0	0	0	0	0	0	0	0	0	...	0	0	12	0		0	0	0	495 7

34 rows × 114 columns



```
In [16]: px.bar(crosstab,orientation='h',height=700)
```



## In model we have missing values to fill these null values i choose condition based retrival

- only one Volvo brand having the null values so from volvo we find mode of model that is "XC90" now we use these value for null values.

```
In [17]: df['Model']=df['Model'].fillna("XC90")
```

```
In [18]: df['Model'].isnull().sum()
```

```
Out[18]: 0
```

```
In [19]: df[df['Vehicle Location'].isnull()]
```

```
Out[19]:
```

	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
16	1N4AZ0CP4D	Pierce	Kapowsin	WA	98344	2013	NISSAN	LEAF	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	75	0	2.0000
9196	3FA6P0SU9E	Hidalgo	Mcallen	TX	78501	2014	FORD	FUSION	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	19	0	29.8056
...	...	...	...	...	...	...	...	...	Battery Electric	Clean Alternative	...	...	...

```
In [20]: df['Electric Vehicle Type'].value_counts()
```

```
Out[20]: Battery Electric Vehicle (BEV)      86044
Plug-in Hybrid Electric Vehicle (PHEV)    26590
Name: Electric Vehicle Type, dtype: int64
```

```
In [21]: crosstab1=pd.crosstab(df['Vehicle Location'],df['Electric Vehicle Type'])
crosstab1['Battery Electric Vehicle (BEV)'].sort_values()
```

```
Out[21]: Vehicle Location
POINT (7.86484 51.32975)      0
POINT (-118.01268 33.83899)   0
POINT (-121.92442 36.55443)   0
POINT (-117.97378 47.30036)   0
POINT (-117.90629 47.20139)   0

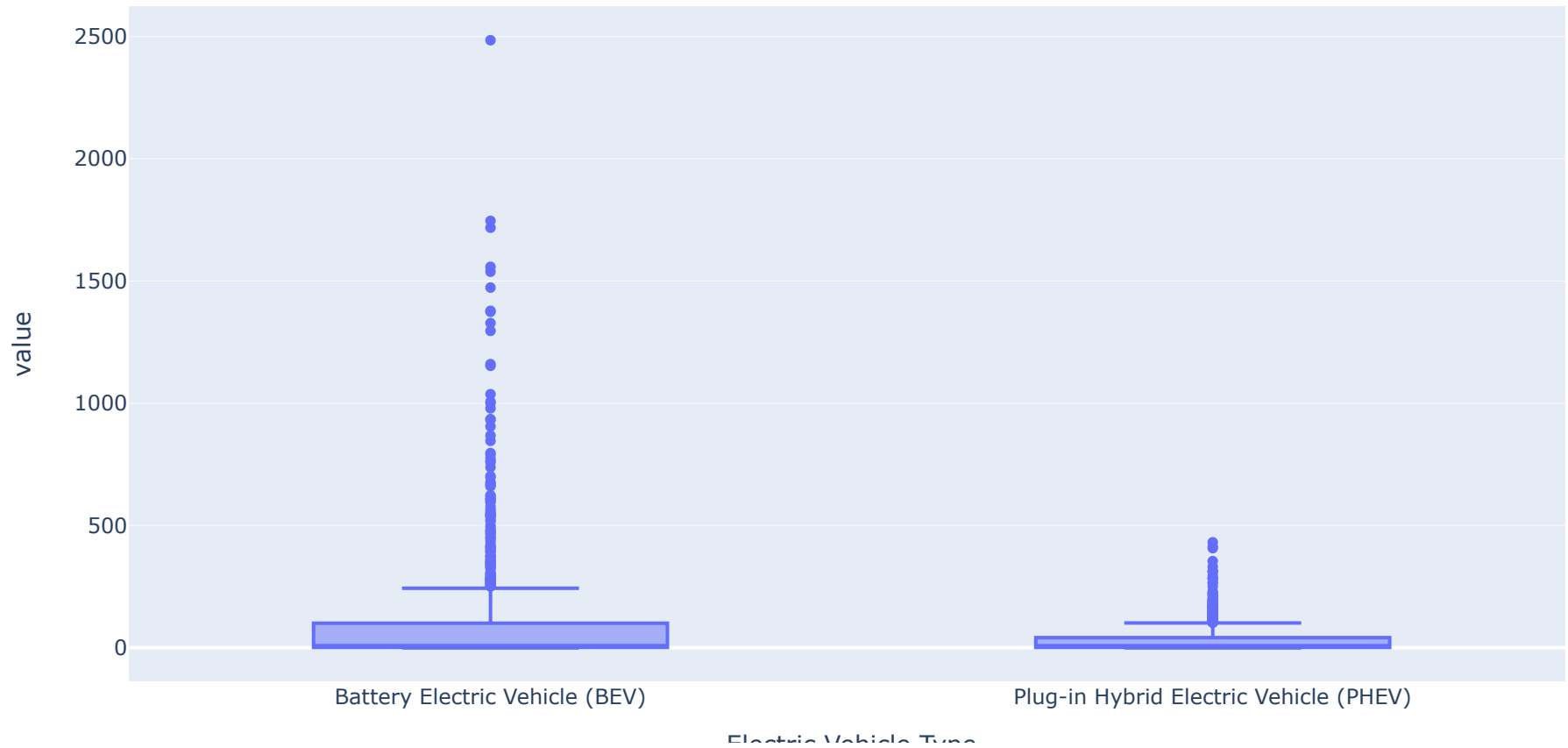
...
POINT (-122.21061 47.83448)   1538
POINT (-122.12096 47.55584)   1558
POINT (-122.1872 47.61001)    1718
POINT (-122.2066 47.67887)    1746
POINT (-122.13158 47.67858)    2485
Name: Battery Electric Vehicle (BEV), Length: 758, dtype: int64
```

```
In [22]: crosstab1['Plug-in Hybrid Electric Vehicle (PHEV)'].sort_values()
```

```
Out[22]: Vehicle Location
POINT (-102.69968 22.95716)    0
POINT (-76.8907 38.81605)      0
POINT (-118.50797 48.99237)    0
POINT (-118.59524 34.2271)     0
POINT (-76.73517 39.10852)     0

...
POINT (-122.521 47.62728)      331
POINT (-122.35436 47.67596)    354
POINT (-122.31765 47.70013)    407
POINT (-122.89166 47.03956)    413
POINT (-122.13158 47.67858)    431
Name: Plug-in Hybrid Electric Vehicle (PHEV), Length: 758, dtype: int64
```

```
In [23]: px.box(crosstab1)
```



```
In [24]: df['Vehicle Location']=df['Vehicle Location'].fillna(df['Vehicle Location'].mode()[0])
```

```
In [25]: df['Electric Utility']=df['Electric Utility'].fillna(df['Electric Utility'].mode()[0])
```

```
In [26]: df.isnull().sum()
```

```
Out[26]: VIN (1-10)          0
County          0
City            0
State           0
Postal Code     0
Model Year      0
Make            0
Model           0
Electric Vehicle Type  0
Clean Alternative Fuel Vehicle (CAFV) Eligibility  0
Electric Range  0
Base MSRP       0
Legislative District  0
DOL Vehicle ID  0
Vehicle Location  0
Electric Utility  0
2020 Census Tract  0
dtype: int64
```

## outliers

```
In [27]: df.columns
```

```
Out[27]: Index(['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'],
              dtype='object')
```

```
In [28]: num=df.select_dtypes(include='number')
```

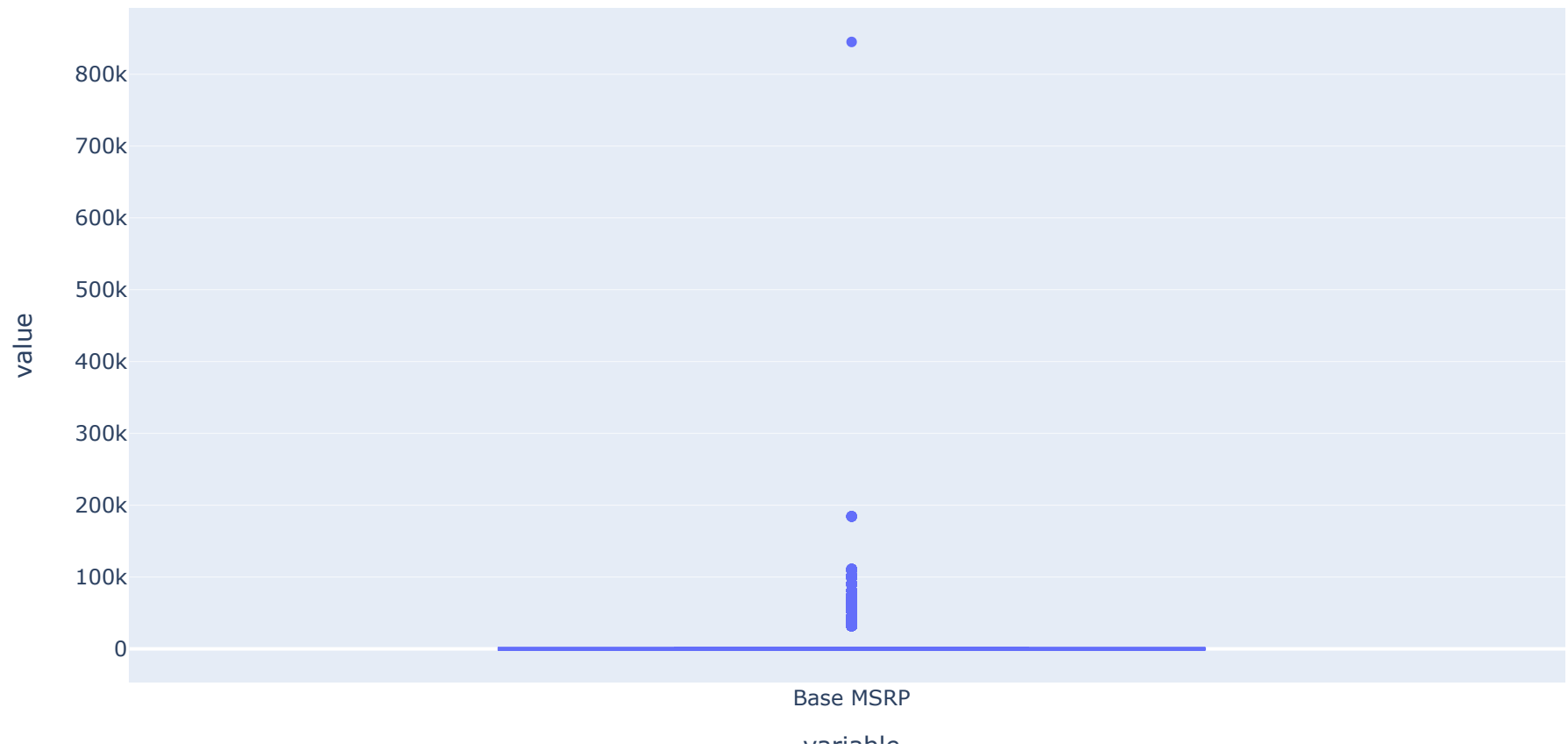
In [29]: num

Out[29]:

	Postal Code	Model Year	Electric Range	Base MSRP	Legislative District	DOL Vehicle ID	2020 Census Tract
0	33040	2022	42	0	29.805604	198968248	12087972100
1	89029	2013	38	0	29.805604	5204412	32003005702
2	98901	2011	73	0	15.000000	218972519	53077001602
3	98237	2017	238	0	39.000000	186750406	53057951101
4	98201	2019	26	0	38.000000	2006714	53061041500
...	...	...	...	...	...	...	...
112629	98019	2022	0	0	45.000000	217955265	53033032401
112630	98250	2019	150	0	40.000000	103663227	53055960301
112631	98070	2022	38	0	34.000000	193878387	53033027702
112632	98042	2018	26	0	47.000000	125039043	53033032007
112633	98042	2022	18	0	47.000000	194673692	53033032005

112634 rows × 7 columns

```
In [30]: px.box(df['Base MSRP'])
```





```
In [31]: q1=df['Base MSRP'].quantile(0.25)

q3=df['Base MSRP'].quantile(0.75)

iqr=q3-q1

lb=q1-1.5*iqr
ub=q1+1.5*iqr

df[(df['Base MSRP']>=lb) & (df['Base MSRP']>=200000)]
```

Out[31]:

	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	Ve Loc
62533	WP0CA2A13F	King	Hunts Point	WA	98004	2015	PORSCHE	918	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	12	845000	48.0	100479039	P (-122.47.6

```
In [32]: df = df.drop(index=62533)
```

```
In [33]: df.reset_index(drop='index',inplace=True)
```

In [34]: df

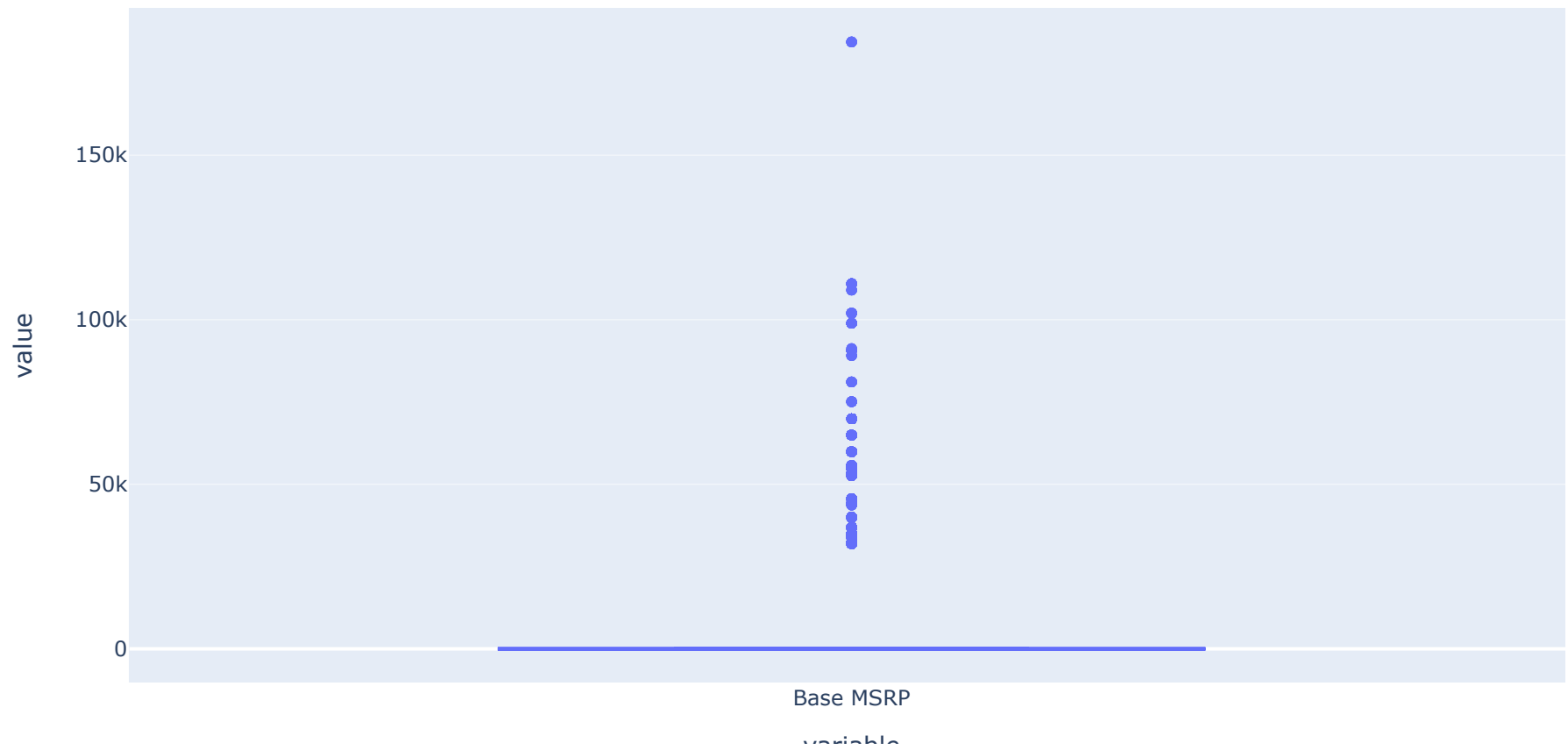
Out[34]:

	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	Vehicle
0	JTMEB3FV6N	Monroe	Key West	FL	33040	2022	TOYOTA	RAV4 PRIME	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	42	0	29.805604	1989€
1	1G1RD6E45D	Clark	Laughlin	NV	89029	2013	CHEVROLET	VOLT	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	38	0	29.805604	52€
2	JN1AZ0CP8B	Yakima	Yakima	WA	98901	2011	NISSAN	LEAF	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	73	0	15.000000	21897
3	1G1FW6S08H	Skagit	Concrete	WA	98237	2017	CHEVROLET	BOLT EV	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	238	0	39.000000	18675
4	3FA6P0SU1K	Snohomish	Everett	WA	98201	2019	FORD	FUSION	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	26	0	38.000000	20€
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
112628	7SAYGDEF2N	King	Duvall	WA	98019	2022	TESLA	MODEL Y	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0	45.000000	21795
112629	1N4BZ1CP7K	San Juan	Friday Harbor	WA	98250	2019	NISSAN	LEAF	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	150	0	40.000000	1036€

	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	Vehic
112630	1FMCU0KZ4N	King	Vashon	WA	98070	2022	FORD	ESCAPE	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	38	0	34.000000	19387
112631	KNDCD3LD4J	King	Covington	WA	98042	2018	KIA	NIRO	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	26	0	47.000000	12503
112632	YV4BR0CL8N	King	Covington	WA	98042	2022	VOLVO	XC90	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	18	0	47.000000	19467

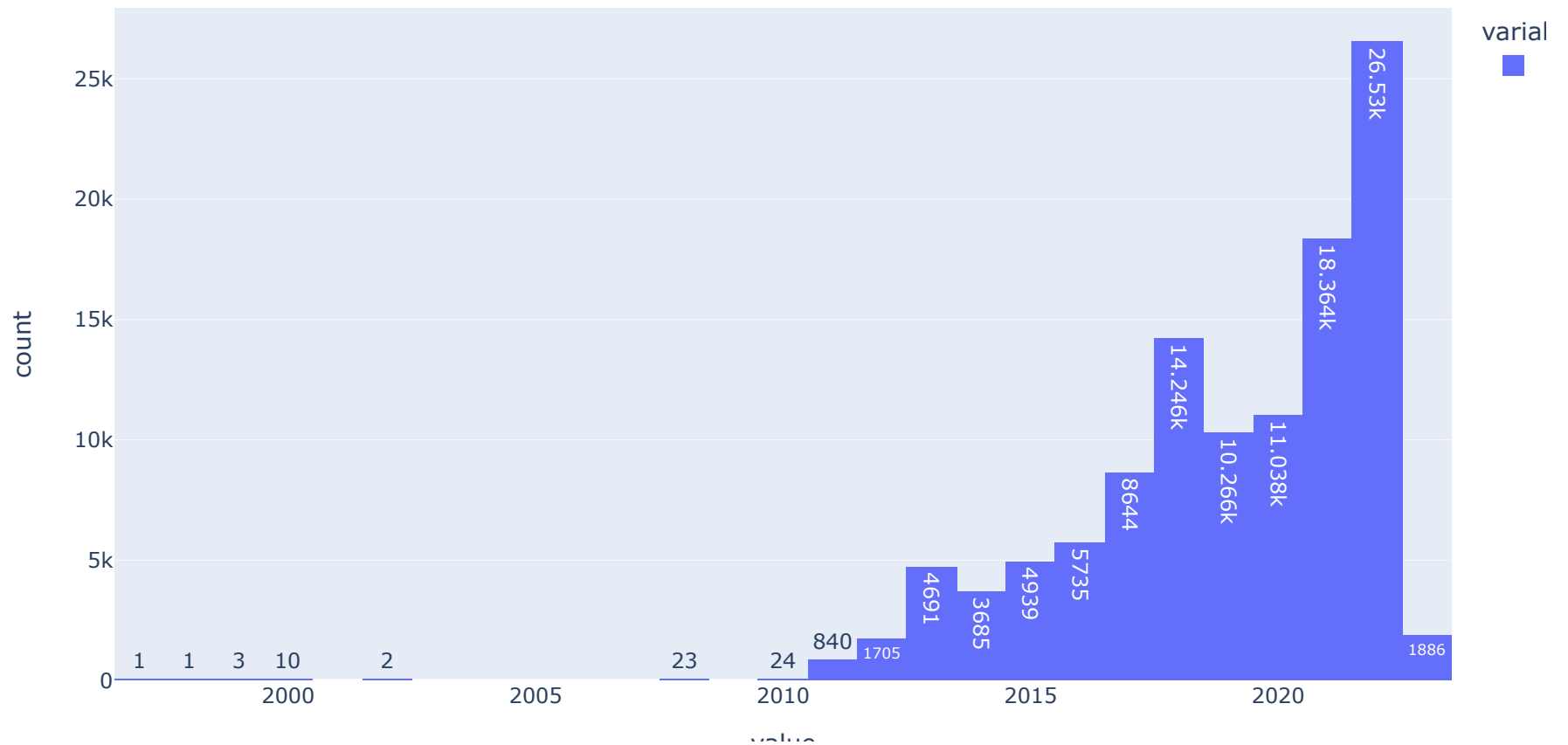
112633 rows × 17 columns

```
In [35]: px.box(df['Base MSRP'])
```



## Task1 (Description) - Apply Exploratory Data Analysis(Univariate and Bivariate) using plotly.express library.

```
In [36]: fig=px.histogram(df['Model Year'],orientation='v',text_auto=True)  
fig.show()
```



## Observation

- Every year the frequency will be increased

```
In [37]: cat=df.select_dtypes(exclude='number')
```

In [38]: cat



Out[38]:

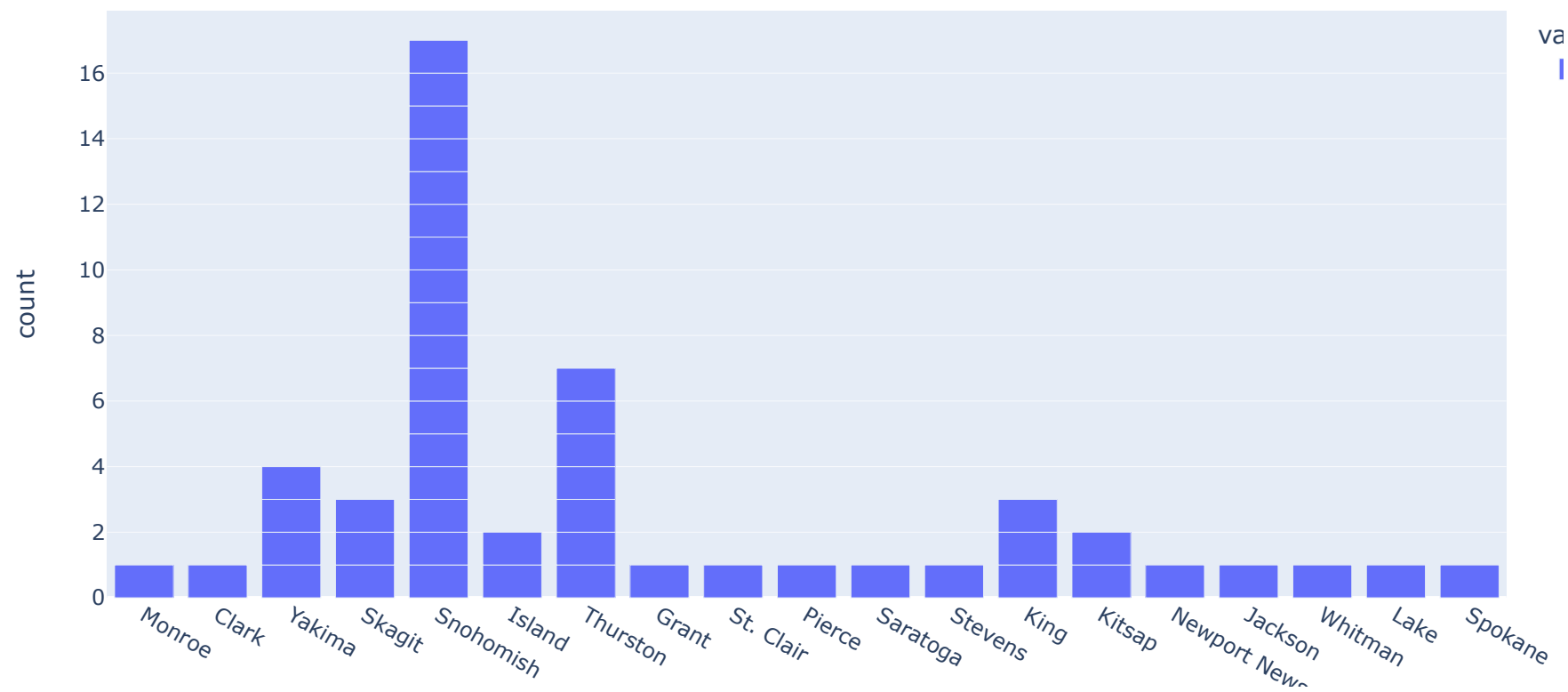
	VIN (1-10)	County	City	State	Make	Model	Electric Vehicle Type	Clean Alternative Fuel Vehicle (CAFV) Eligibility	Vehicle Location	Electric Utility
0	JTMEB3FV6N	Monroe	Key West	FL	TOYOTA	RAV4 PRIME	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	POINT (-81.80023 24.5545)	PUGET SOUND ENERGY INC  CITY OF TACOMA - (WA)
1	1G1RD6E45D	Clark	Laughlin	NV	CHEVROLET	VOLT	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	POINT (-114.57245 35.16815)	PUGET SOUND ENERGY INC  CITY OF TACOMA - (WA)
2	JN1AZ0CP8B	Yakima	Yakima	WA	NISSAN	LEAF	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	POINT (-120.50721 46.60448)	PACIFICORP
3	1G1FW6S08H	Skagit	Concrete	WA	CHEVROLET	BOLT EV	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	POINT (-121.7515 48.53892)	PUGET SOUND ENERGY INC
4	3FA6P0SU1K	Snohomish	Everett	WA	FORD	FUSION	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	POINT (-122.20596 47.97659)	PUGET SOUND ENERGY INC
...	...	...	...	...	...	...	...	...	...	...
112628	7SAYGDEF2N	King	Duvall	WA	TESLA	MODEL Y	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	POINT (-121.98609 47.74068)	PUGET SOUND ENERGY INC  CITY OF TACOMA - (WA)
112629	1N4BZ1CP7K	San Juan	Friday Harbor	WA	NISSAN	LEAF	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	POINT (-123.01648 48.53448)	BONNEVILLE POWER ADMINISTRATION  ORCAS POWER &...

	VIN (1-10)	County	City	State	Make	Model	Electric Vehicle Type	Clean Alternative Fuel Vehicle (CAFV) Eligibility	Vehicle Location	Electric Utility
112630	1FMCU0KZ4N	King	Vashon	WA	FORD	ESCAPE	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	POINT (-122.4573 47.44929)	PUGET SOUND ENERGY INC  CITY OF TACOMA - (WA)
112631	KND3CD3LD4J	King	Covington	WA	KIA	NIRO	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	POINT (-122.09124 47.33778)	PUGET SOUND ENERGY INC  CITY OF TACOMA - (WA)
112632	YV4BR0CL8N	King	Covington	WA	VOLVO	XC90	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	POINT (-122.09124 47.33778)	PUGET SOUND ENERGY INC  CITY OF TACOMA - (WA)

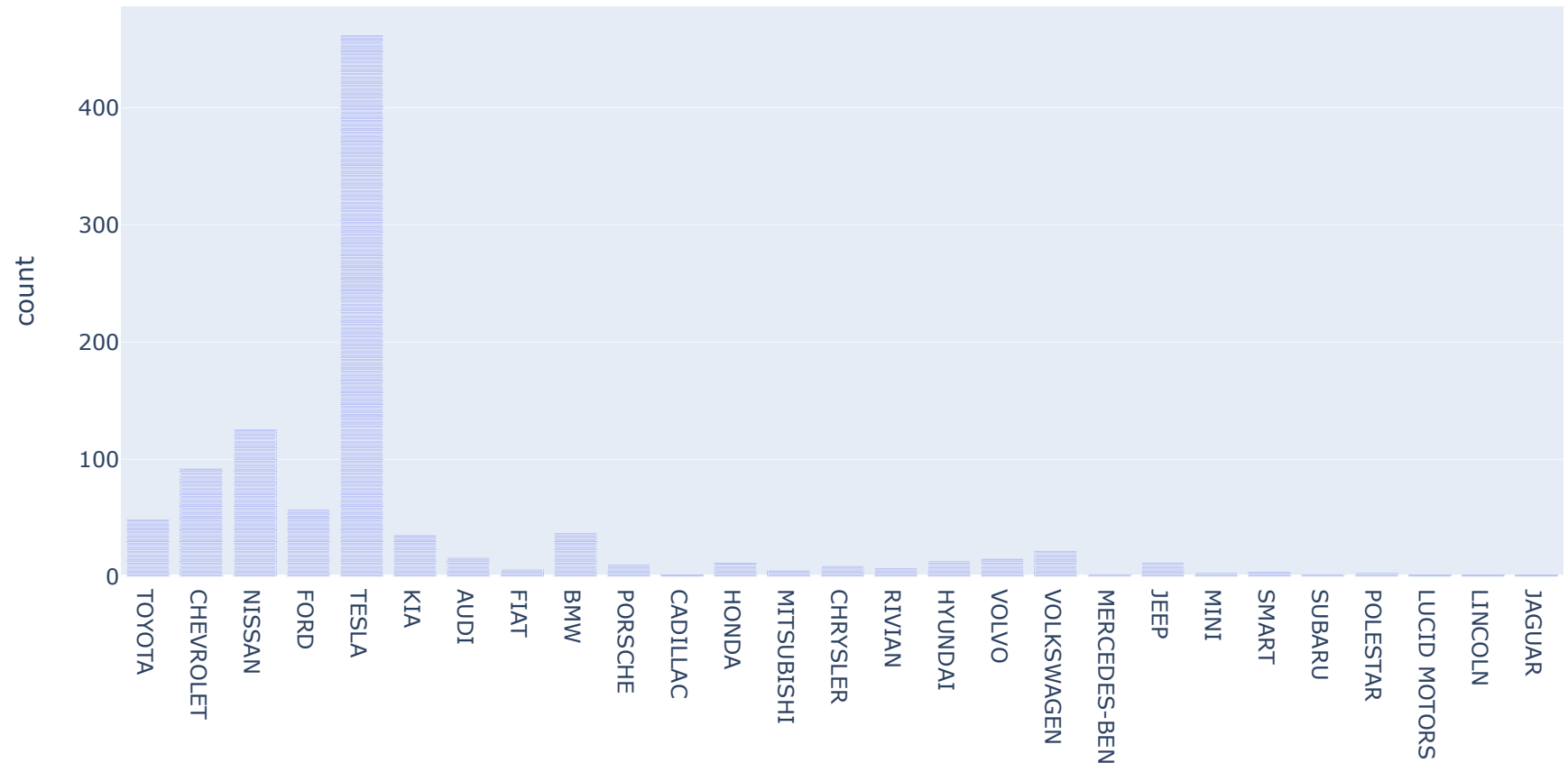
112633 rows × 10 columns

```
In [39]: px.bar(df['County'][0:50],title='Top 50 countries')
```

Top 50 countries



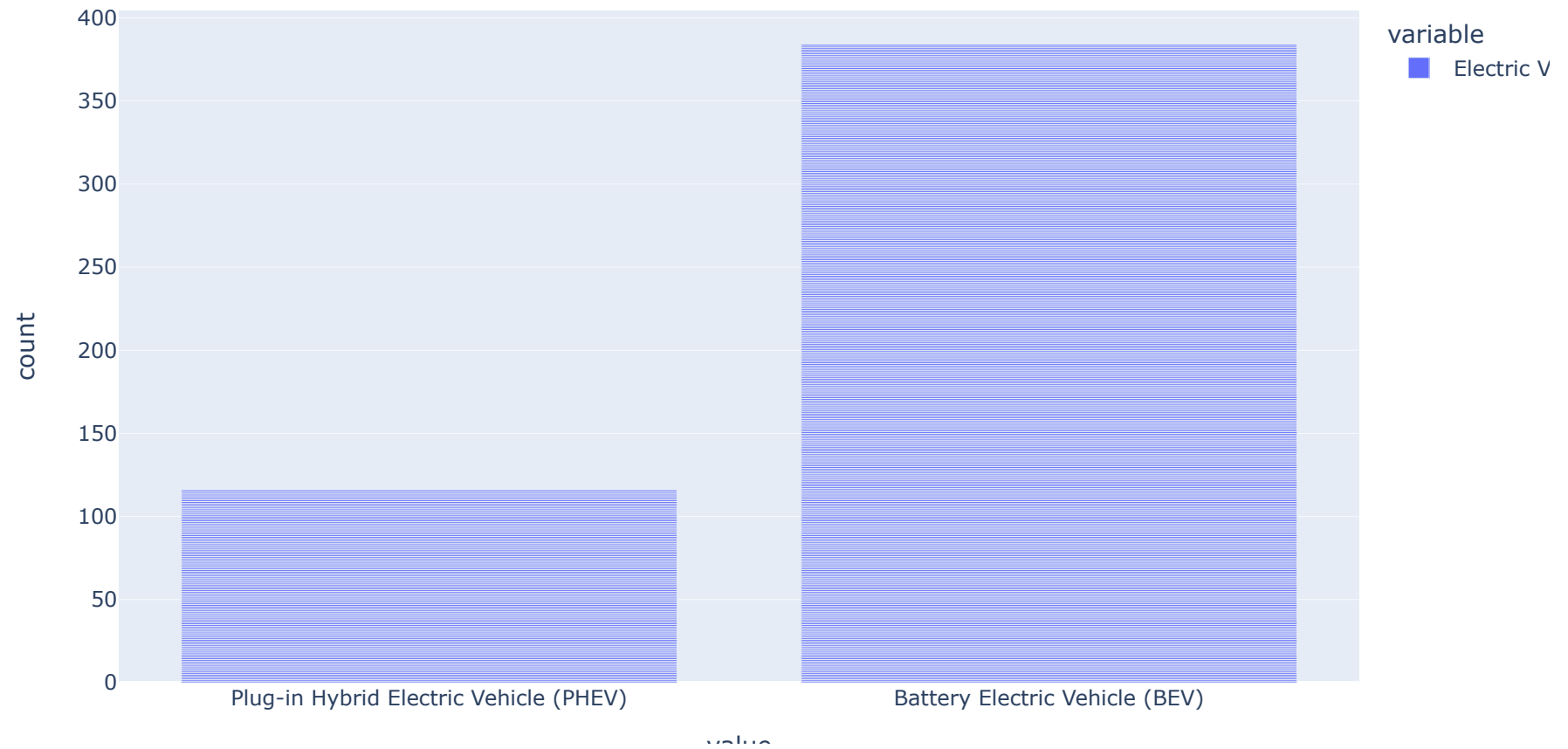
```
In [40]: px.bar(df['Make'][:1000])
```



## Observation:

- TESLA having more intrested to manufacturing the electronic vehicle

```
In [41]: px.bar(df['Electric Vehicle Type'][0:500])
```

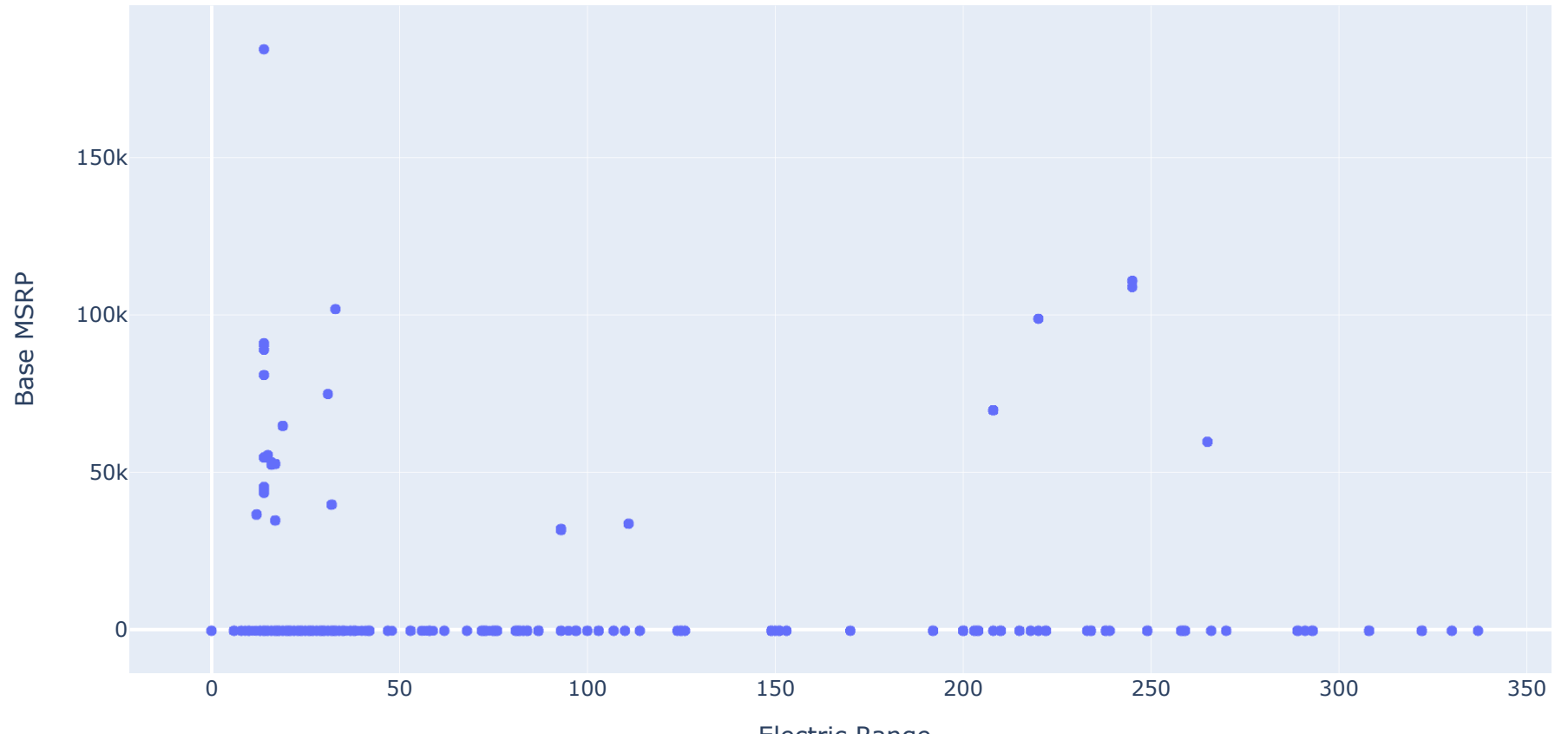


## Observation:

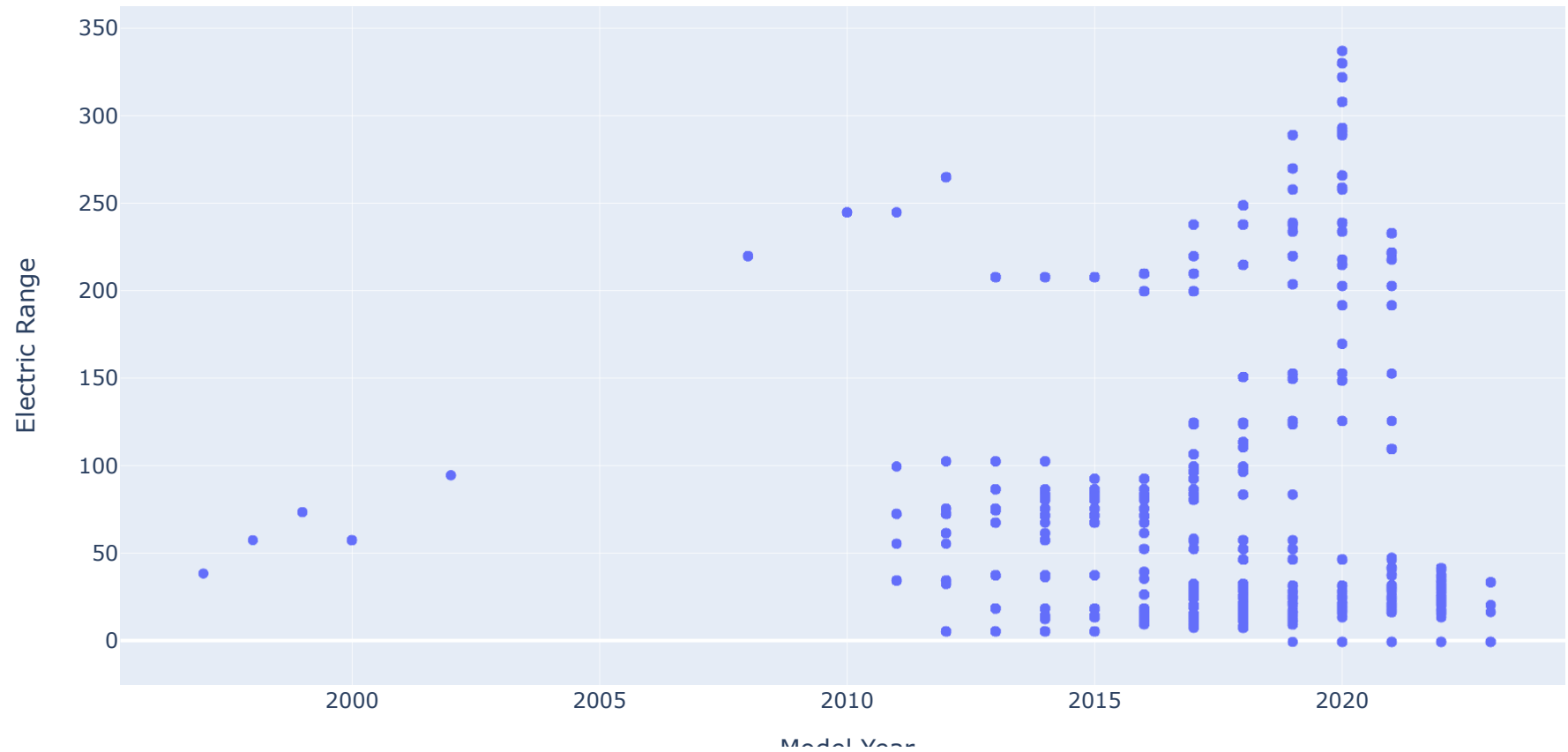
- Most of the companies are using Battery Electric Vehicles comparing with plug-in-Hybrid electric vehicle

# BI-VARIATE

```
In [42]: px.scatter(x=df['Electric Range'],y=df['Base MSRP'],data_frame=df)
```



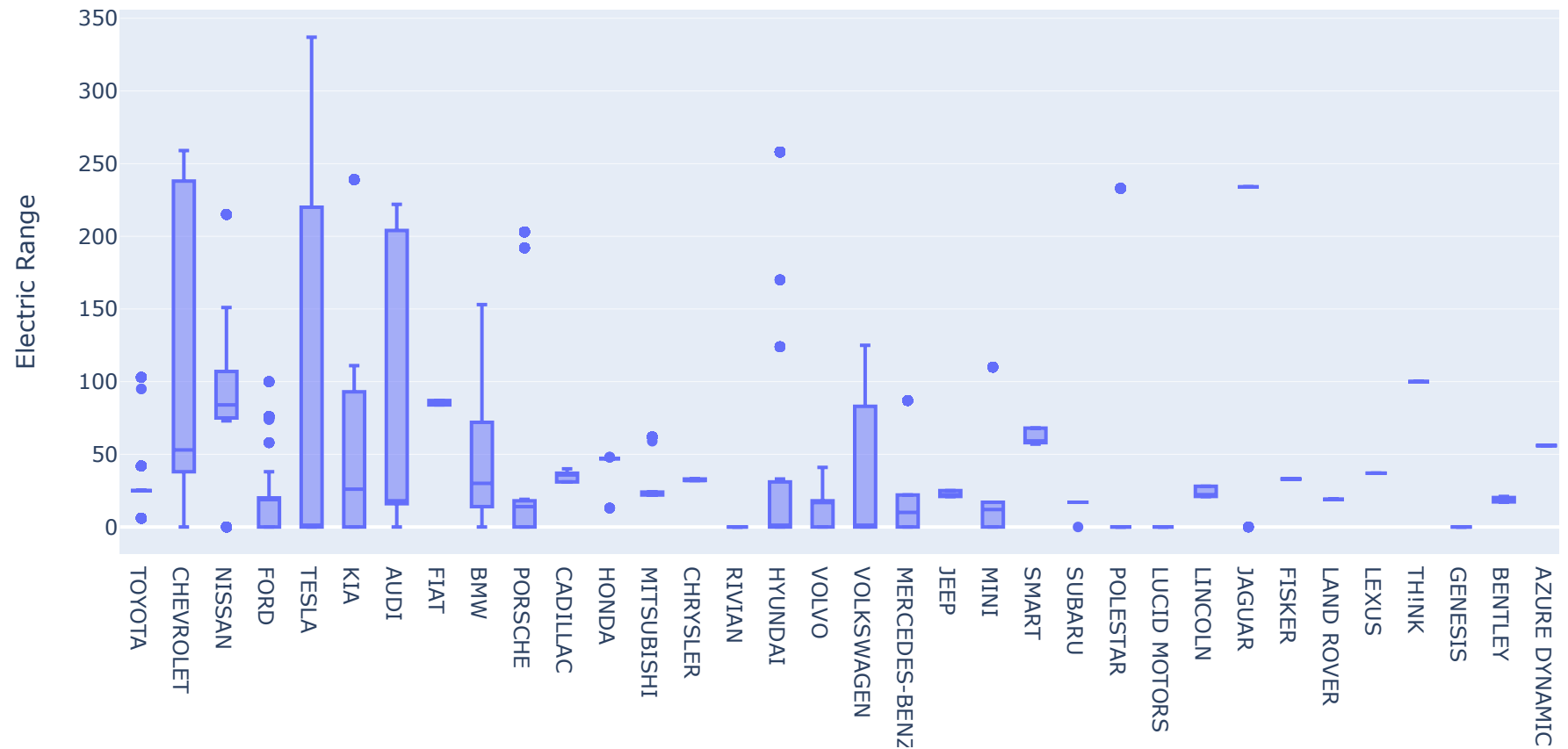
```
In [43]: px.scatter(x=df['Model Year'],y=df['Electric Range'],data_frame=df)
```



## Observation:

- In the Model Year 2020 having the high electric range that is 337 compare to the other model years

```
In [44]: px.box(x='Make',y='Electric Range',data_frame=df)
```

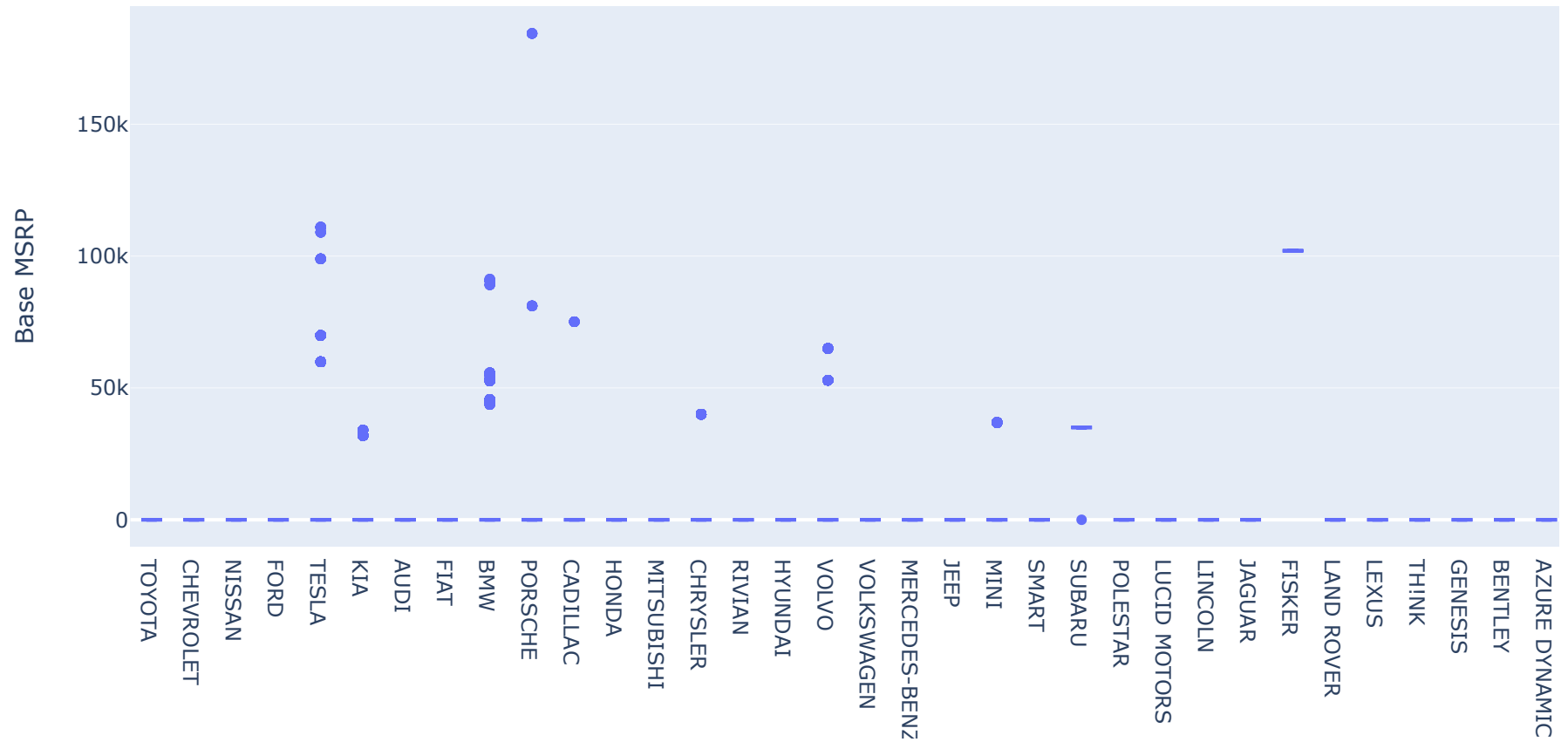


## Observation

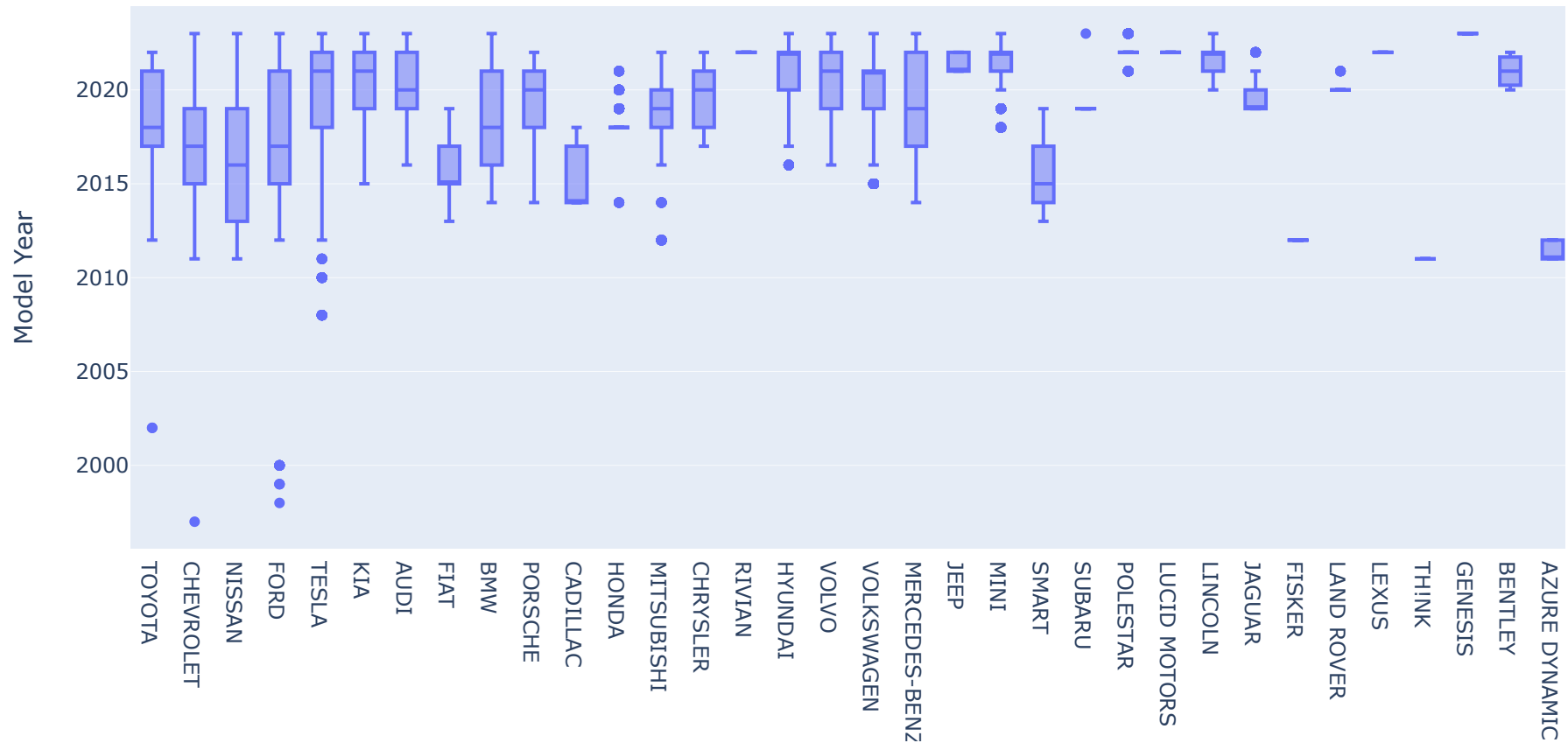
- Tesla having the maximum electric range that is 337.



```
In [45]: px.box(x=df['Make'],y=df['Base MSRP'],data_frame=df)
```



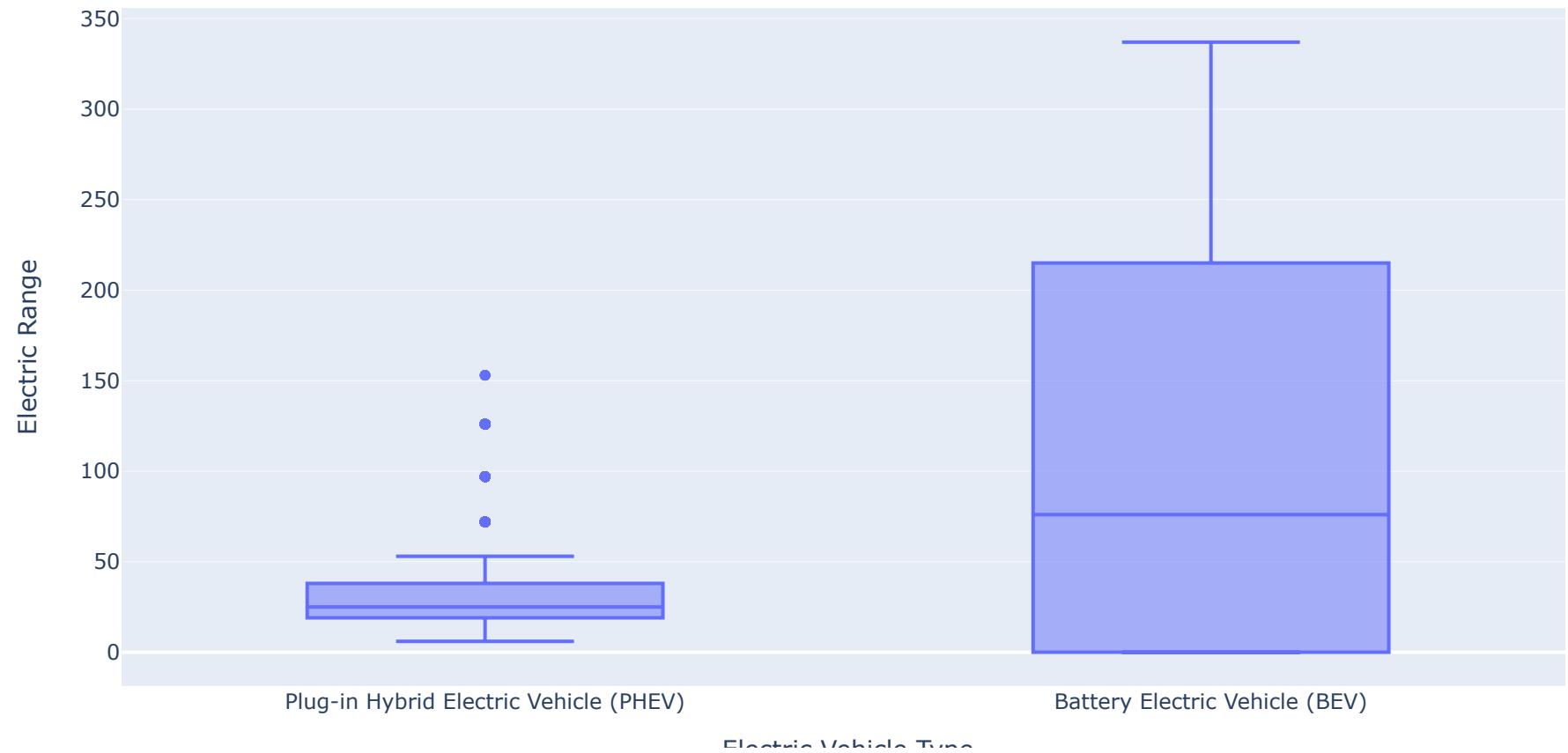
```
In [46]: px.box(x=df['Make'],y=df['Model Year'],data_frame=df)
```



## Observation:

- KIA And Tesla most of average electrical vehicles released in 2021 model-year because the median is 2021

```
In [47]: px.box(x=df['Electric Vehicle Type'],y=df['Electric Range'],data_frame=df)
```



## Observation:

- Battery Electric Vehicle have more electric range that is 337

In [48]: cat

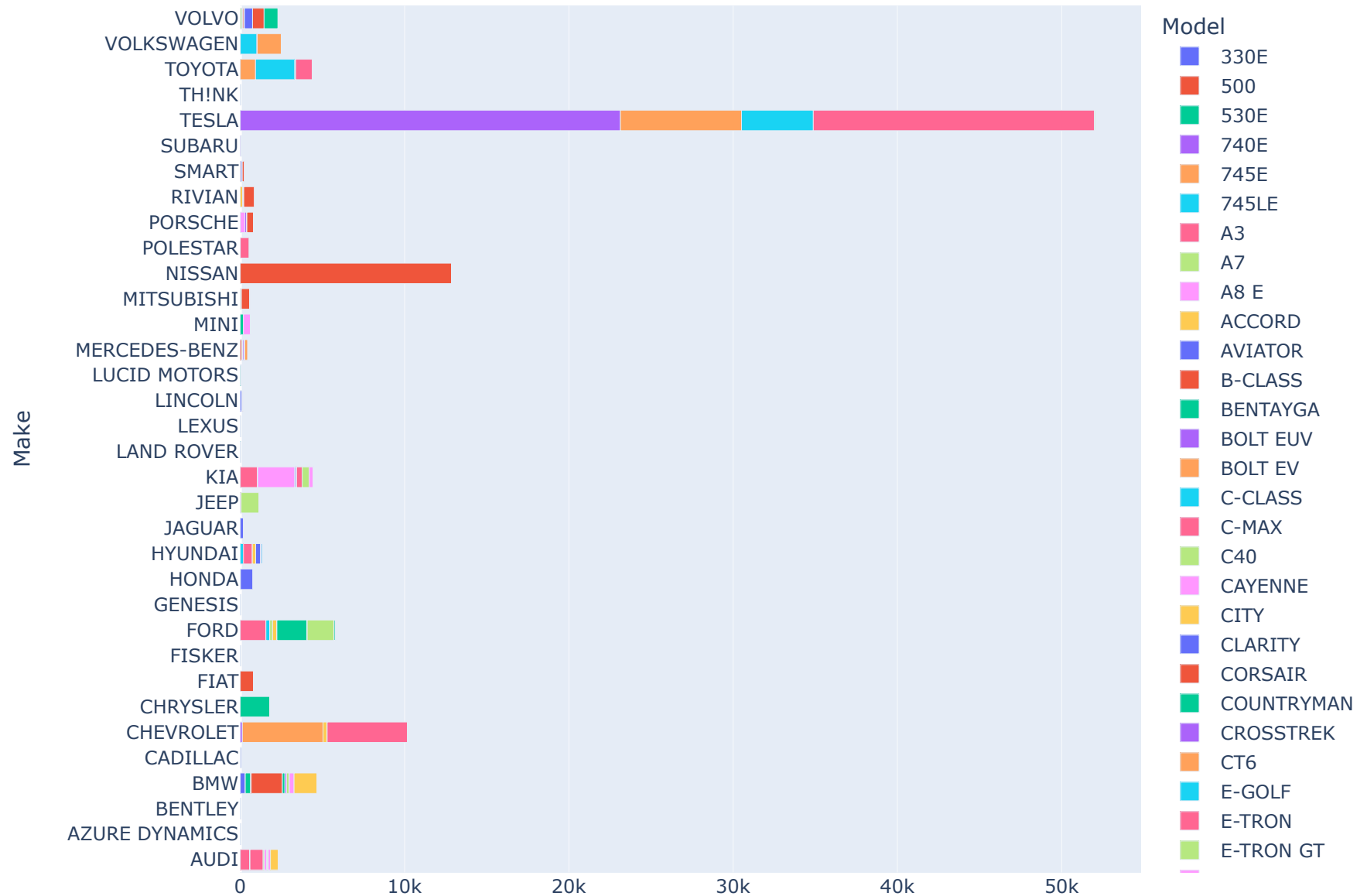
Out[48]:

	VIN (1-10)	County	City	State	Make	Model	Electric Vehicle Type	Clean Alternative Fuel Vehicle (CAFV) Eligibility	Vehicle Location	Electric Utility
0	JTMEB3FV6N	Monroe	Key West	FL	TOYOTA	RAV4 PRIME	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	POINT (-81.80023 24.5545)	PUGET SOUND ENERGY INC  CITY OF TACOMA - (WA)
1	1G1RD6E45D	Clark	Laughlin	NV	CHEVROLET	VOLT	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	POINT (-114.57245 35.16815)	PUGET SOUND ENERGY INC  CITY OF TACOMA - (WA)
2	JN1AZ0CP8B	Yakima	Yakima	WA	NISSAN	LEAF	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	POINT (-120.50721 46.60448)	PACIFICORP
3	1G1FW6S08H	Skagit	Concrete	WA	CHEVROLET	BOLT EV	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	POINT (-121.7515 48.53892)	PUGET SOUND ENERGY INC
4	3FA6P0SU1K	Snohomish	Everett	WA	FORD	FUSION	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	POINT (-122.20596 47.97659)	PUGET SOUND ENERGY INC
...	...	...	...	...	...	...	...	...	...	...
112628	7SAYGDEF2N	King	Duvall	WA	TESLA	MODEL Y	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	POINT (-121.98609 47.74068)	PUGET SOUND ENERGY INC  CITY OF TACOMA - (WA)
112629	1N4BZ1CP7K	San Juan	Friday Harbor	WA	NISSAN	LEAF	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	POINT (-123.01648 48.53448)	BONNEVILLE POWER ADMINISTRATION  ORCAS POWER &...

VIN (1-10)		County	City	State	Make	Model	Electric Vehicle Type	Clean Alternative Fuel Vehicle (CAFV) Eligibility	Vehicle Location	Electric Utility
112630	1FMCU0KZ4N	King	Vashon	WA	FORD	ESCAPE	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	POINT (-122.4573 47.44929)	PUGET SOUND ENERGY INC  CITY OF TACOMA - (WA)
112631	KND3CD3LD4J	King	Covington	WA	KIA	NIRO	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	POINT (-122.09124 47.33778)	PUGET SOUND ENERGY INC  CITY OF TACOMA - (WA)
112632	YV4BR0CL8N	King	Covington	WA	VOLVO	XC90	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	POINT (-122.09124 47.33778)	PUGET SOUND ENERGY INC  CITY OF TACOMA - (WA)

112633 rows × 10 columns

```
In [49]: crosstab_1=pd.crosstab(df['Make'],df['Model'])  
px.bar(crosstab_1,orientation='h',height=700)
```





## Observation:

- BMW making more model electric vehicle's like x5,x3,1x,l8,l4,l3,740E,530E,330E.

### statistical test

**Is there a relationship between Make and Country (i.e. Does the preference ofCountry depend on the Making company?)**

**h0:Make and Country has relationship**

**h1:Make and Country has no relationship**

```
In [50]: from scipy.stats import chi2_contingency  
from scipy.stats import chi2
```

```
In [51]: observed = pd.crosstab(df.Make,df.County)
```

```
In [52]: chi2_contingency(observed)
```

```
Out[52]: Chi2ContingencyResult(statistic=17206.387705438785, pvalue=0.0, dof=5412, expected_freq=array([[7.24654409e-01, 4.14
088233e-02, 1.65635293e-01, ...,
4.14088233e-02, 2.07044117e-02, 1.27746220e+01],
[2.17520620e-03, 1.24297497e-04, 4.97189989e-04, ...,
1.24297497e-04, 6.21487486e-05, 3.83457779e-02],
[9.32231229e-04, 5.32703559e-05, 2.13081424e-04, ...,
5.32703559e-05, 2.66351780e-05, 1.64339048e-02],
...,
[1.36882619e+00, 7.82186393e-02, 3.12874557e-01, ...,
7.82186393e-02, 3.91093196e-02, 2.41304502e+01],
[7.81209770e-01, 4.46405583e-02, 1.78562233e-01, ...,
4.46405583e-02, 2.23202791e-02, 1.37716122e+01],
[7.10981684e-01, 4.06275248e-02, 1.62510099e-01, ...,
4.06275248e-02, 2.03137624e-02, 1.25335914e+01]]))
```

```
In [53]: chi2_test_stat = chi2_contingency(observed)[0]
pval = chi2_contingency(observed)[1]
df = chi2_contingency(observed)[2]
```

```
In [54]: confidence_level = 0.90

alpha = 1 - confidence_level

chi2_critical = chi2.ppf(1 - alpha, df)

chi2_critical
```

```
Out[54]: 5545.751557653358
```

```
In [55]: # Plotting the chi2 distribution to visualise

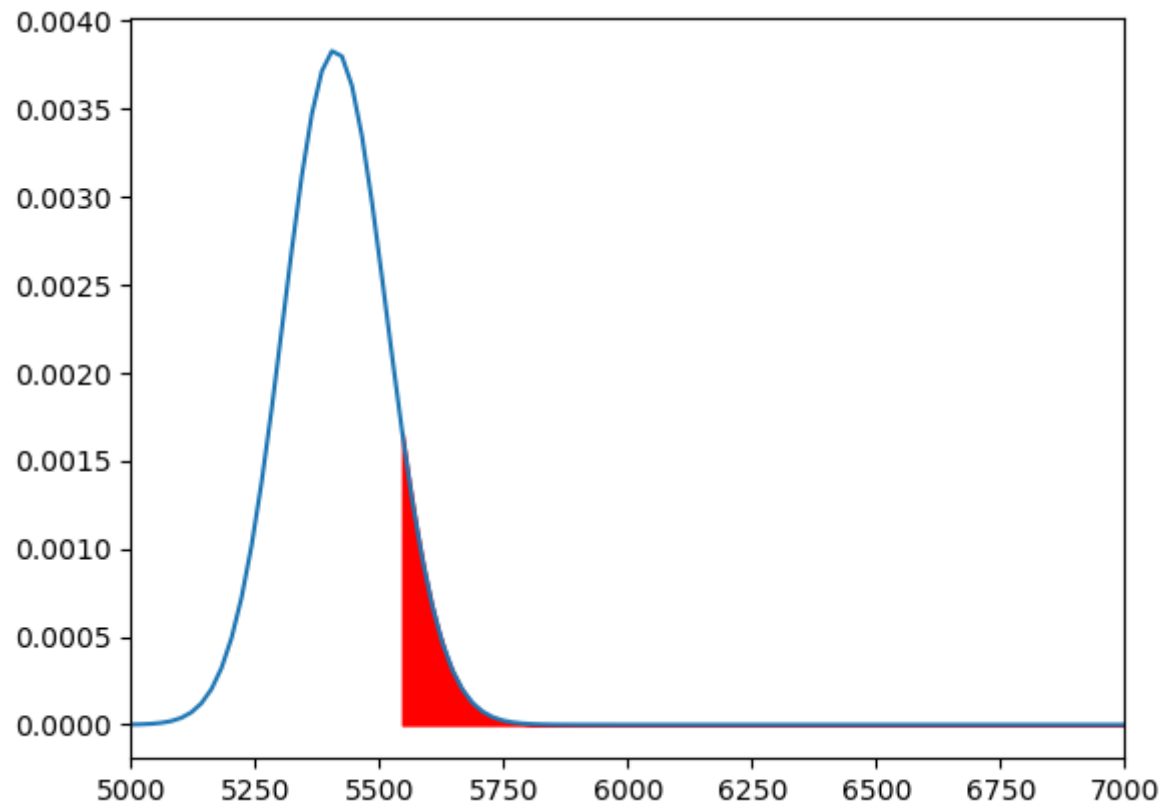
# Defining the x minimum and x maximum
#plt.figure(figsize=(15,6))
x_min = 5000
x_max = 7000

# Plotting the graph and setting the x limits
x = np.linspace(x_min, x_max, 100)
y = chi2.pdf(x, df)
plt.xlim(x_min, x_max)
plt.plot(x, y)

# Setting Chi2 Critical value
chi2_critical_right = chi2_critical

# Shading the right rejection region
x1 = np.linspace(chi2_critical_right, x_max, 100)
y1 = chi2.pdf(x1, df)
plt.fill_between(x1, y1, color='red')
```

```
Out[55]: <matplotlib.collections.PolyCollection at 0x1defabaa410>
```



```
In [56]: if(chi2_test_stat > chi2_critical):  
          print("Reject Null Hypothesis")  
        else:  
          print("Fail to Reject Null Hypothesis")
```

Reject Null Hypothesis

```
In [57]: if(pval < alpha):  
          print("Reject Null Hypothesis")  
        else:  
          print("Fail to Reject Null Hypothesis")
```

Reject Null Hypothesis

```
In [58]: df=pd.read_csv(r"C:\Users\Irfan\Downloads\dataset.csv")  
df
```

Out[58]:

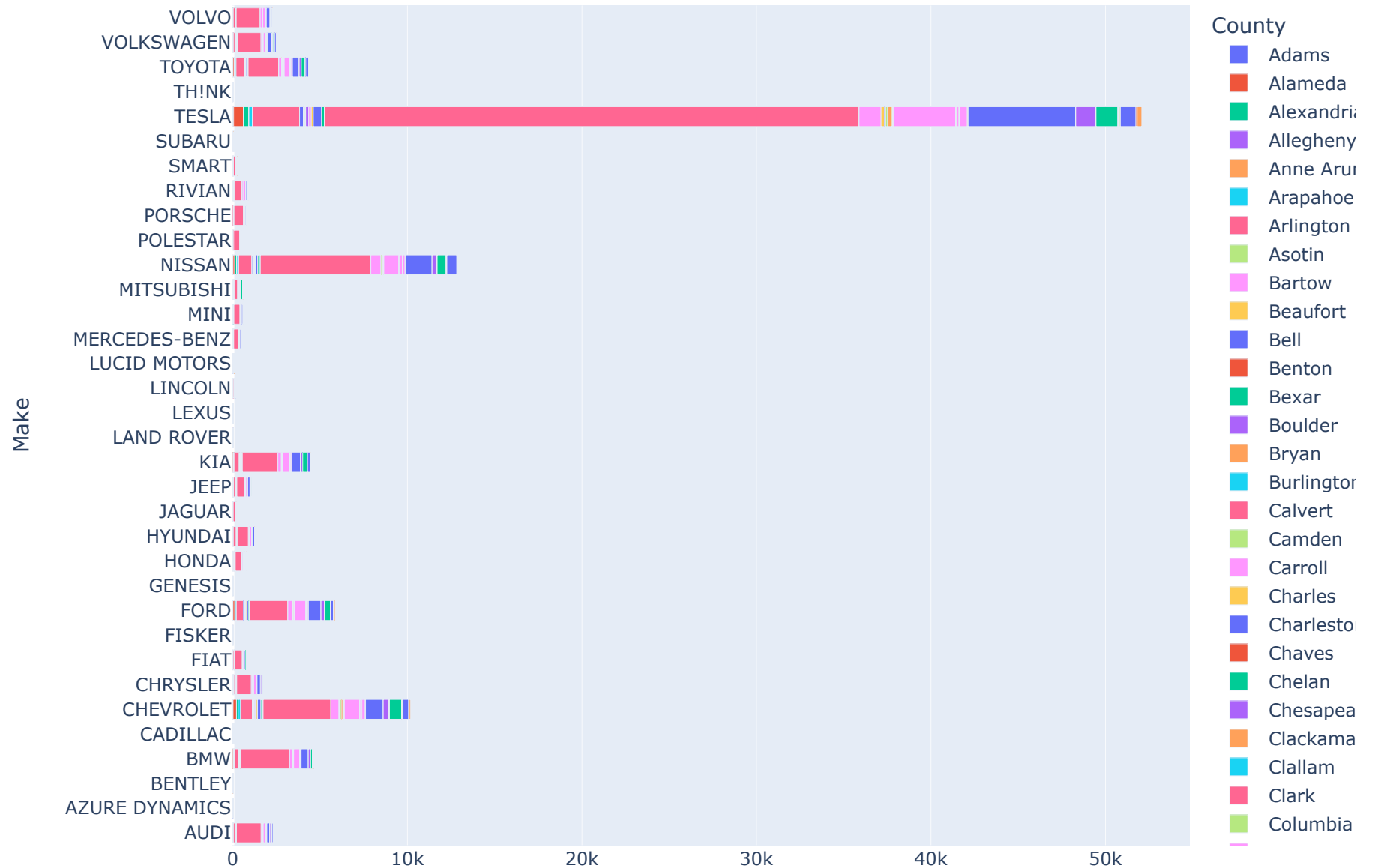
	VIN (1-10)	County	City	State	Postal Code	Model Year	Make	Model	Electric Vehicle Type	Clean Alternative Fuel Vehicle (CAFV) Eligibility	Electric Range	Base MSRP	Legislative District	Vehicle
0	JTMEB3FV6N	Monroe	Key West	FL	33040	2022	TOYOTA	RAV4 PRIME	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	42	0	NaN	1989€
1	1G1RD6E45D	Clark	Laughlin	NV	89029	2013	CHEVROLET	VOLT	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	38	0	NaN	52€
2	JN1AZ0CP8B	Yakima	Yakima	WA	98901	2011	NISSAN	LEAF	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	73	0	15.0	21897
3	1G1FW6S08H	Skagit	Concrete	WA	98237	2017	CHEVROLET	BOLT EV	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	238	0	39.0	18675
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	20€
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
112629	7SAYGDEF2N	King	Duvall	WA	98019	2022	TESLA	MODEL Y	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0	45.0	21795
112630	1N4BZ1CP7K	San Juan	Friday Harbor	WA	98250	2019	NISSAN	LEAF	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	150	0	40.0	1036€

	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	Vehic
112631	1FMCU0KZ4N	King	Vashon	WA	98070	2022	FORD	ESCAPE	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	38	0	34.0	19387
112632	KNDCD3LD4J	King	Covington	WA	98042	2018	KIA	NIRO	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	26	0	47.0	12503
112633	YV4BR0CL8N	King	Covington	WA	98042	2022	VOLVO	XC90	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	18	0	47.0	19467

112634 rows × 17 columns

```
In [59]: crosstab_2=pd.crosstab(df['Make'],df['County'])  
px.bar(crosstab_2,orientation='h',height=700)
```





## Observation:

- In king country having every type of company electrical vehicle so we can say that the electric vehicle buisness most popular in KING country

## Conclusion:¶

- Since the p-value is less than the significance level of 0.05, we can reject the null hypothesis. Therefore, we can conclude that there is a no relationship between Make and country.

## Task2 (Description) - Create a Choropleth to display the number of EV vehicles based on location.

```
In [60]: import plotly.graph_objects as go
def create_ev_choropleth_map(df):
    # Calculate the count of EV vehicles for each state
    ev_count_by_state = df['State'].value_counts().reset_index()
    ev_count_by_state.columns = ['State', 'EV Count']

    # Create the Choropleth map using plotly.graph_objects
    fig_choropleth = go.Figure(data=go.Choropleth(
        locations=ev_count_by_state['State'],
        z=ev_count_by_state['EV Count'],
        locationmode='USA-states',
        colorscale='Viridis',
        colorbar_title='Number of EV Vehicles',
    ))

    # Set the map title and layout
    fig_choropleth.update_layout(
        title_text='Choropleth Map of EV Vehicles by State',
        geo_scope='world',
    )

    return fig_choropleth

fig = create_ev_choropleth_map(df)
fig.show()
```

## Choropleth Map of EV Vehicles by State



**Task3 (Description) - Create a Racing Bar Plot to display the animation of EV Make and its count each year.**

```
In [61]: pip install bar_chart_race
```

```
Requirement already satisfied: bar_chart_race in c:\users\irfan\anaconda3\lib\site-packages (0.1.0)
Requirement already satisfied: matplotlib>=3.1 in c:\users\irfan\anaconda3\lib\site-packages (from bar_chart_race) (3.7.0)
Requirement already satisfied: pandas>=0.24 in c:\users\irfan\anaconda3\lib\site-packages (from bar_chart_race) (1.5.3)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\irfan\anaconda3\lib\site-packages (from matplotlib>=3.1->bar_chart_race) (1.4.4)
Requirement already satisfied: cycler>=0.10 in c:\users\irfan\anaconda3\lib\site-packages (from matplotlib>=3.1->bar_chart_race) (0.11.0)
Requirement already satisfied: python-dateutil>=2.7 in c:\users\irfan\anaconda3\lib\site-packages (from matplotlib>=3.1->bar_chart_race) (2.8.2)
Requirement already satisfied: fonttools>=4.22.0 in c:\users\irfan\anaconda3\lib\site-packages (from matplotlib>=3.1->bar_chart_race) (4.25.0)
Requirement already satisfied: contourpy>=1.0.1 in c:\users\irfan\anaconda3\lib\site-packages (from matplotlib>=3.1->bar_chart_race) (1.0.5)
Requirement already satisfied: pillow>=6.2.0 in c:\users\irfan\anaconda3\lib\site-packages (from matplotlib>=3.1->bar_chart_race) (9.4.0)
Requirement already satisfied: packaging>=20.0 in c:\users\irfan\anaconda3\lib\site-packages (from matplotlib>=3.1->bar_chart_race) (22.0)
Requirement already satisfied: pyparsing>=2.3.1 in c:\users\irfan\anaconda3\lib\site-packages (from matplotlib>=3.1->bar_chart_race) (3.0.9)
Requirement already satisfied: numpy>=1.20 in c:\users\irfan\anaconda3\lib\site-packages (from matplotlib>=3.1->bar_chart_race) (1.23.5)
Requirement already satisfied: pytz>=2020.1 in c:\users\irfan\anaconda3\lib\site-packages (from pandas>=0.24->bar_chart_race) (2022.7)
Requirement already satisfied: six>=1.5 in c:\users\irfan\anaconda3\lib\site-packages (from python-dateutil>=2.7->matplotlib>=3.1->bar_chart_race) (1.16.0)
Note: you may need to restart the kernel to use updated packages.
```

```
In [62]: import bar_chart_race as bcr
```

```

In [ ]: # Converting the 'Model Year' column to datetime type
df['Model Year'] = pd.to_datetime(df['Model Year'], format='%Y')

# Group by 'Model Year' and 'Make' to get the count of each make for each year
df_grouped = df.groupby(['Model Year', 'Make']).size().reset_index(name='Count')

import bar_chart_race as bcr

df_pivot = df_grouped.pivot(index='Model Year', columns='Make', values='Count')

# Fill missing values using forward fill (pad)
df_pivot = df_pivot.fillna(method='pad')

# Create the Racing Bar Plot
bcr.bar_chart_race(
    df=df_pivot,
    filename='ev_make_racing_bar_plot.mp4',
    orientation='h',
    sort='desc',
    nBars=10,
    fixed_order=False,
    title='EV Make Racing Bar Plot by Year',
    label_bars=True,
    period_label={'x': 0.99, 'y': 0.25, 'ha': 'right', 'va': 'center'},
)

```

## Conclusion

- Every year the frequency will be increased
- BMW making more model electric vehicle's like x5,x3,1x,l8,l4,l3,740E,530E,330E.
- Tesla having the maximum electric range that is 337.
- In the Model Year 2020 having the high electric range that is 337 compare to the other model years.
- Most of the companies are using Battery ELeetric Vehicles comparing witj plug-in-Hybrid electric vehicle.
- Seattle is the top city in top 10 with electric Cars.
- King County is the top in top 10 county with more electric Vehicles
- 98052 postal code contains the high electric cars.

- JAGUR have the more electric range comapre to other makes.
- Tesla is the most popular electric car make in Washington state, followed by Nissan, Chevrolet, and Toyota.
- Tesla is also the most popular make in Seattle, followed by Nissan, Chevrolet, and BMW.
- Washington state has the highest number of Audi, BMW, and Chevrolet electric cars registered among all states.

```
In [ ]: pip install --upgrade PyPDF2
```

```
In [ ]:
```