

# SaimProjectFile

June 23, 2024

**0.0.1 Bootcamp DSAIML Project 2024 - Mohd. Saim(C4) System id : 2022006233**

## **1 List of Table**

## **2 List of Figure**

## **3 # Problem statement/Objective**

Market Size Analysis is the process of estimating the potential sales for a product or service within a particular market segment. In the context of electric vehicles (EVs), it involves assessing the total volume of EV registrations to understand the growth of the market, forecast future trends, and help stakeholders make informed decisions regarding production, infrastructure development, and policy-making.

The provided dataset contains the following columns, each representing different aspects of the electric vehicle (EV) population in the United States:

- VIN (1-10): Partial Vehicle Identification Number.
- County: The county in which the vehicle is registered.
- City: The city in which the vehicle is registered.
- State: The state in which the vehicle is registered. It appears that this dataset may be focused on Washington (WA) state.
- Postal Code: The postal code where the vehicle is registered.
- Model Year: The year of the vehicle model
- Make: The manufacturer of the vehicle.
- Model: The model of the vehicle.
- Electric Vehicle Type: The type of electric vehicle, e.g., Battery Electric Vehicle (BEV).
- Clean Alternative Fuel Vehicle (CAFV) Eligibility: Eligibility status for clean alternative fuel vehicle programs.
- Electric Range: The maximum range of the vehicle on a single charge (in miles).
- Base MSRP: The Manufacturer's Suggested Retail Price.
- Legislative District: The legislative district where the vehicle is registered.
- DOL Vehicle ID: Department of Licensing Vehicle Identification.
- Vehicle Location: Geographic coordinates of the vehicle location.
- Electric Utility: The electric utility service provider for the vehicle's location.
- 2020 Census Tract: The census tract for the vehicle's location.

### 3.1 Basic Steps

```
[1]: import os
os.getcwd()
```

```
[1]: 'C:\\Users\\anupi'
```

```
[11]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

```
[12]: dtype_specification = {
    10: str,
    12: str
}
data = pd.read_csv('6-Electric_Vehicle_Population_Data New.csv',
dtype=dtype_specification)
```

### 3.2 1) Display the top 5 rows

```
[13]: data.head()
```

```
[13]: VIN (1-10)      County      City State  Postal Code  Model Year  Make \
0  5YJYGDEE1L      King      Seattle  WA      98122.0      2020  TESLA
1  7SAYGDEE9P  Snohomish  Bothell  WA      98021.0      2023  TESLA
2  5YJSA1E4XK      King      Seattle  WA      98109.0      2019  TESLA
3  5YJSA1E27G      King      Issaquah  WA      98027.0      2016  TESLA
4  5YJYGDEE5M      Kitsap  Suquamish  WA      98392.0      2021   NaN
```

```
Model      Electric Vehicle Type \
0  MODEL Y  Battery Electric Vehicle (BEV)
1  MODEL Y  Battery Electric Vehicle (BEV)
2  MODEL S  Battery Electric Vehicle (BEV)
3  MODEL S  Battery Electric Vehicle (BEV)
4  MODEL Y  Battery Electric Vehicle (BEV)
```

```
Clean Alternative Fuel Vehicle (CAFV) Eligibility Electric Range \
0      Clean Alternative Fuel Vehicle Eligible      291
1  Eligibility unknown as battery range has not b...      0
2      NaN      270
3      Clean Alternative Fuel Vehicle Eligible      210
4  Eligibility unknown as battery range has not b...      0
```

```
Base MSRP Legislative District  DOL Vehicle ID \
0      0      37      125701579
1      0      1      244285107
2      0      36      156773144
```

3	0	5	165103011
4	0	23	205138552

```

Vehicle Location \
0 POINT (-122.30839 47.610365)
1 POINT (-122.179458 47.802589)
2 POINT (-122.34848 47.632405)
3 POINT (-122.03646 47.534065)
4 POINT (-122.55717 47.733415)

```

	Electric Utility	2020 Census Tract
0	CITY OF SEATTLE - (WA) CITY OF TACOMA - (WA)	5.303301e+10
1	PUGET SOUND ENERGY INC	5.306105e+10
2	CITY OF SEATTLE - (WA) CITY OF TACOMA - (WA)	5.303301e+10
3	PUGET SOUND ENERGY INC  CITY OF TACOMA - (WA)	5.303303e+10
4	PUGET SOUND ENERGY INC	5.303594e+10

### 3.3 2) Display the last 5 rows

```
[ ]: data.tail()
```

```
[ ]:
VIN (1-10) County City State Postal Code Model Year \
4350 KMHM34AC5P King Kirkland WA 98034 2023
4351 5YJSA1E2XK King Bellevue WA 98006 2019
4352 5YJ3E1EB9M Island Oak Harbor WA 98277 2021
4353 1G1FZ6S08N Thurston Rochester WA 98579 2022
4354 1N4BZ0CP4G King Covington WA 98042 2016
```

```

Make Model Electric Vehicle Type \
4350 HYUNDAI IONIQ 6 Battery Electric Vehicle (BEV)
4351 TESLA MODEL S Battery Electric Vehicle (BEV)
4352 TESLA MODEL 3 Battery Electric Vehicle (BEV)
4353 CHEVROLET BOLT EV Battery Electric Vehicle (BEV)
4354 NISSAN LEAF Battery Electric Vehicle (BEV)

```

```

Clean Alternative Fuel Vehicle (CAFV) Eligibility Electric Range \
4350 Eligibility unknown as battery range has not b... 0
4351 Clean Alternative Fuel Vehicle Eligible 270
4352 Eligibility unknown as battery range has not b... 0
4353 Eligibility unknown as battery range has not b... 0
4354 Clean Alternative Fuel Vehicle Eligible 84

```

```

Base MSRP Legislative District DOL Vehicle ID \
4350 0 45 260328554
4351 0 41 202143378
4352 0 10 181200813
4353 0 20 211177039

```

```
4354          0          47      176683636
```

```
          Vehicle Location \
4350    POINT (-122.209285 47.71124)
4351    POINT (-122.16937 47.571015)
4352    POINT (-122.6788673 48.2897314)
4353    POINT (-123.09575 46.82114)
4354    POINT (-122.111625 47.36078)
```

```
          Electric Utility  2020 Census Tract
4350  PUGET SOUND ENERGY INC||CITY OF TACOMA - (WA)      53033021903
4351  PUGET SOUND ENERGY INC||CITY OF TACOMA - (WA)      53033023902
4352          PUGET SOUND ENERGY INC      53029970402
4353          PUGET SOUND ENERGY INC      53067012730
4354  PUGET SOUND ENERGY INC||CITY OF TACOMA - (WA)      5303303200
```

### 3.4 3) Check the shape of dataset

```
[ ]: data.shape
```

```
[ ]: (13029, 17)
```

### 3.5 4) Check the datatypes of each feature.

```
[7]: data.dtypes
```

```
[7]: VIN (1-10)          object
County                object
City                 object
State                object
Postal Code          float64
Model Year           int64
Make                 object
Model                object
Electric Vehicle Type object
Clean Alternative Fuel Vehicle (CAFV) Eligibility object
Electric Range        object
Base MSRP             int64
Legislative District object
DOL Vehicle ID        int64
Vehicle Location      object
Electric Utility      object
2020 Census Tract     float64
dtype: object
```

### 3.6 5) Check the Statistical summary

```
[ ]: data.describe()
```

```
[ ]:      Postal Code      Model Year      Base MSRP  DOL Vehicle ID  \
count    13029.000000    13029.000000    13029.000000    1.302900e+04
mean     98179.398879     2020.496585     1119.917876    2.209412e+08
std        224.795648         2.960427     8595.903627    7.720531e+07
min       98001.000000     2000.000000         0.000000    1.861330e+05
25%       98033.000000     2019.000000         0.000000    1.823020e+08
50%       98103.000000     2022.000000         0.000000    2.276053e+08
75%       98178.000000     2023.000000         0.000000    2.549045e+08
max       99362.000000     2024.000000    184400.000000    4.789259e+08

      2020 Census Tract
count          1.302800e+04
mean           5.303352e+10
std            1.206012e+07
min            5.300796e+10
25%            5.303301e+10
50%            5.303302e+10
75%            5.303303e+10
max            5.307794e+10
```

### 3.7 6) Check the null values

```
[8]: data.isnull().sum()
```

```
[8]: VIN (1-10)                                0
County                                         5
City                                           5
State                                          0
Postal Code                                   5
Model Year                                    0
Make                                           7
Model                                          4
Electric Vehicle Type                         6
Clean Alternative Fuel Vehicle (CAFV) Eligibility 2
Electric Range                               3
Base MSRP                                     0
Legislative District                         389
DOL Vehicle ID                               0
Vehicle Location                             9
Electric Utility                             5
2020 Census Tract                           5
dtype: int64
```

### 3.8 7- Check the duplicate values

```
[ ]: data.duplicated().sum()
```

```
[ ]: 0
```

#### 3.8.1 8) Check the anomalies or wrong entries.

```
[14]: plt.figure(figsize=(12, 8))

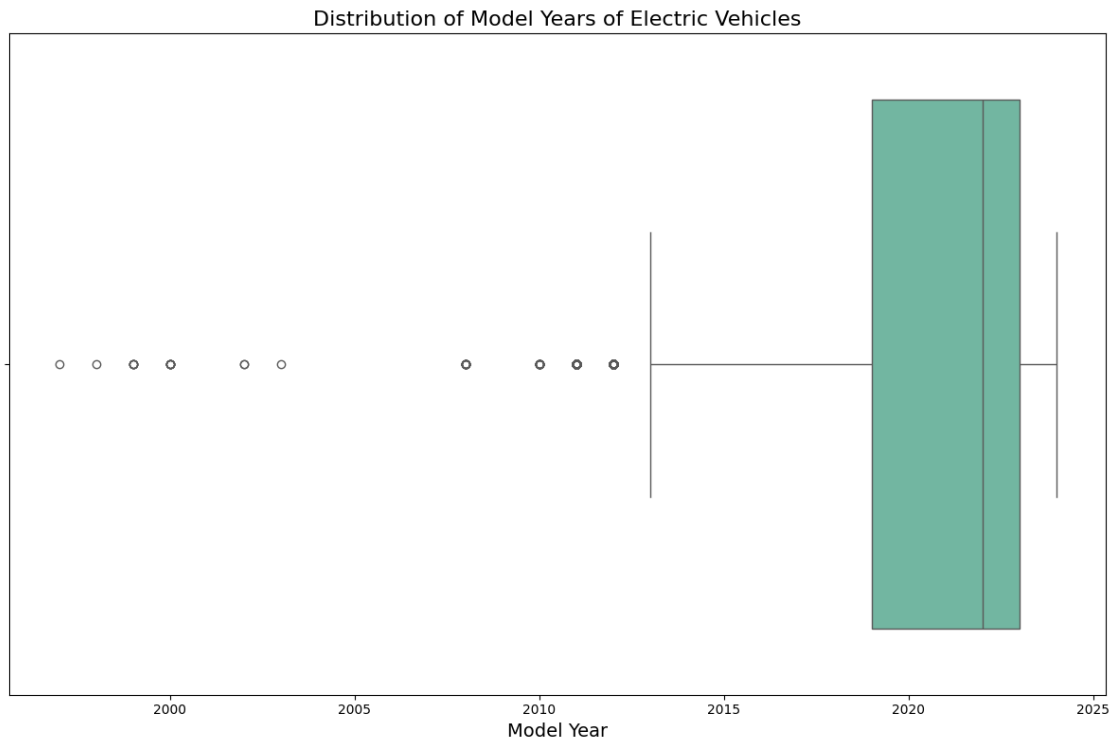
# Create a boxplot
sns.boxplot(x='Model Year', data=data, palette="Set2")
# Add a title and labels
plt.title('Distribution of Model Years of Electric Vehicles', fontsize=16)
plt.xlabel('Model Year', fontsize=14)
plt.ylabel('')

# Improve layout and show the plot
plt.tight_layout()
plt.show()
```

C:\Users\anupi\AppData\Local\Temp\ipykernel\_9408\53870542.py:4: FutureWarning:

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

```
sns.boxplot(x='Model Year', data=data, palette="Set2")
```



```
plt.figure(figsize=(10,10)) sns.boxplot(data=data[''],orient='h')
```

### 3.8.2 9- Check the outliers and their authenticity.

```
[15]: data[data['2020 Census Tract']=='?']
```

```
[15]: Empty DataFrame
Columns: [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]
Index: []
```

### 3.8.3 10- Do the necessary data cleaning steps like dropping duplicates, unnecessary columns, null value imputation, outliers treatment etc.

```
[ ]: data = data.drop_duplicates()
```

```
[ ]: data['2020 Census Tract']=data['2020 Census Tract'].replace('?',np.nan)
```

```
[ ]: data[data['2020 Census Tract']=='?']
```

[ ]: Empty DataFrame

Columns: [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]  
Index: []

[ ]: data.sample(5)

[ ]:	VIN (1-10)	County	City	State	Postal Code	Model Year	\
9278	1G1RA6S5XH	Kitsap	Port Orchard	WA	98367	2017	
11875	5YJ3E1EA7P	King	Kirkland	WA	98034	2023	
4205	WBY43AW05P	King	Bellevue	WA	98004	2023	
7850	1FTVW1EV4P	King	Kent	WA	98031	2023	
6808	1C4JJXP65N	Thurston	Yelm	WA	98597	2022	

	Make	Model	Electric Vehicle Type	\
9278	CHEVROLET	VOLT	Plug-in Hybrid Electric Vehicle (PHEV)	
11875	TESLA	MODEL 3	Battery Electric Vehicle (BEV)	
4205	BMW	I4	Battery Electric Vehicle (BEV)	
7850	FORD	F-150	Battery Electric Vehicle (BEV)	
6808	JEEP	WRANGLER	Plug-in Hybrid Electric Vehicle (PHEV)	

	Clean Alternative Fuel Vehicle (CAFV) Eligibility	Electric Range	\
9278	Clean Alternative Fuel Vehicle Eligible	53	
11875	Eligibility unknown as battery range has not b...	0	
4205	Eligibility unknown as battery range has not b...	0	
7850	Eligibility unknown as battery range has not b...	0	
6808	Not eligible due to low battery range	21	

	Base MSRP	Legislative District	DOL Vehicle ID	\
9278	0	26	198140960	
11875	0	1	240897989	
4205	0	41	244394363	
7850	0	11	258067407	
6808	0	2	207670669	

	Vehicle Location	\
9278	POINT (-122.6847073 47.50524)	
11875	POINT (-122.209285 47.71124)	
4205	POINT (-122.201905 47.61385)	
7850	POINT (-122.2012521 47.3931814)	
6808	POINT (-122.61023 46.94126)	

	Electric Utility	2020 Census Tract
9278	PUGET SOUND ENERGY INC	5.303509e+10
11875	PUGET SOUND ENERGY INC  CITY OF TACOMA - (WA)	5.303302e+10



4205	PUGET SOUND ENERGY INC  CITY OF TACOMA - (WA)	5.303302e+10
7850	PUGET SOUND ENERGY INC  CITY OF TACOMA - (WA)	5.303303e+10
6808	PUGET SOUND ENERGY INC	5.306701e+10

```
[ ]: data[data['Base MSRP']==data['Base MSRP'].min()]
```

	VIN (1-10)	County	City	State	Postal Code	Model Year	Make	\
0	5YJYGDEE1L	King	Seattle	WA	98122	2020	TESLA	
1	7SAYGDEE9P	Snohomish	Bothell	WA	98021	2023	TESLA	
2	5YJSA1E4XK	King	Seattle	WA	98109	2019	TESLA	
3	5YJSA1E27G	King	Issaquah	WA	98027	2016	TESLA	
4	5YJYGDEE5M	Kitsap	Suquamish	WA	98392	2021	NaN	
...	...	...	...	...	...	...	...	
13024	5YJ3E1EB0N	Clark	Vancouver	WA	98685	2022	TESLA	
13025	5YJYGDEE9M	King	Newcastle	WA	98056	2021	TESLA	
13026	5YJYGDEE1M	King	Bellevue	WA	98007	2021	TESLA	
13027	7SAYGDEE4N	Clark	Vancouver	WA	98661	2022	TESLA	
13028	WA1L2BFZXN	King	Seattle	WA	98177	2022	AUDI	

	Model	Electric Vehicle Type	\
0	MODEL Y	Battery Electric Vehicle (BEV)	
1	MODEL Y	Battery Electric Vehicle (BEV)	
2	MODEL S	Battery Electric Vehicle (BEV)	
3	MODEL S	Battery Electric Vehicle (BEV)	
4	MODEL Y	Battery Electric Vehicle (BEV)	
...	...	...	
13024	MODEL 3	Battery Electric Vehicle (BEV)	
13025	MODEL Y	Battery Electric Vehicle (BEV)	
13026	MODEL Y	Battery Electric Vehicle (BEV)	
13027	MODEL Y	Battery Electric Vehicle (BEV)	
13028	Q4	Battery Electric Vehicle (BEV)	

	Clean Alternative Fuel Vehicle (CAFV) Eligibility	Electric Range	\
0	Clean Alternative Fuel Vehicle Eligible	291	
1	Eligibility unknown as battery range has not b...	0	
2	NaN	270	
3	Clean Alternative Fuel Vehicle Eligible	210	
4	Eligibility unknown as battery range has not b...	0	
...	...	...	
13024	Eligibility unknown as battery range has not b...	0	
13025	Eligibility unknown as battery range has not b...	0	
13026	Eligibility unknown as battery range has not b...	0	
13027	Eligibility unknown as battery range has not b...	0	
13028	Eligibility unknown as battery range has not b...	0	

	Base MSRP	Legislative District	DOL Vehicle ID	\
0	0	37	125701579	

1	0	1	244285107
2	0	36	156773144
3	0	5	165103011
4	0	23	205138552
...	...	...	...
13024	0	18	220742007
13025	0	41	171497959
13026	0	48	138322005
13027	0	49	227325850
13028	0	32	219952008

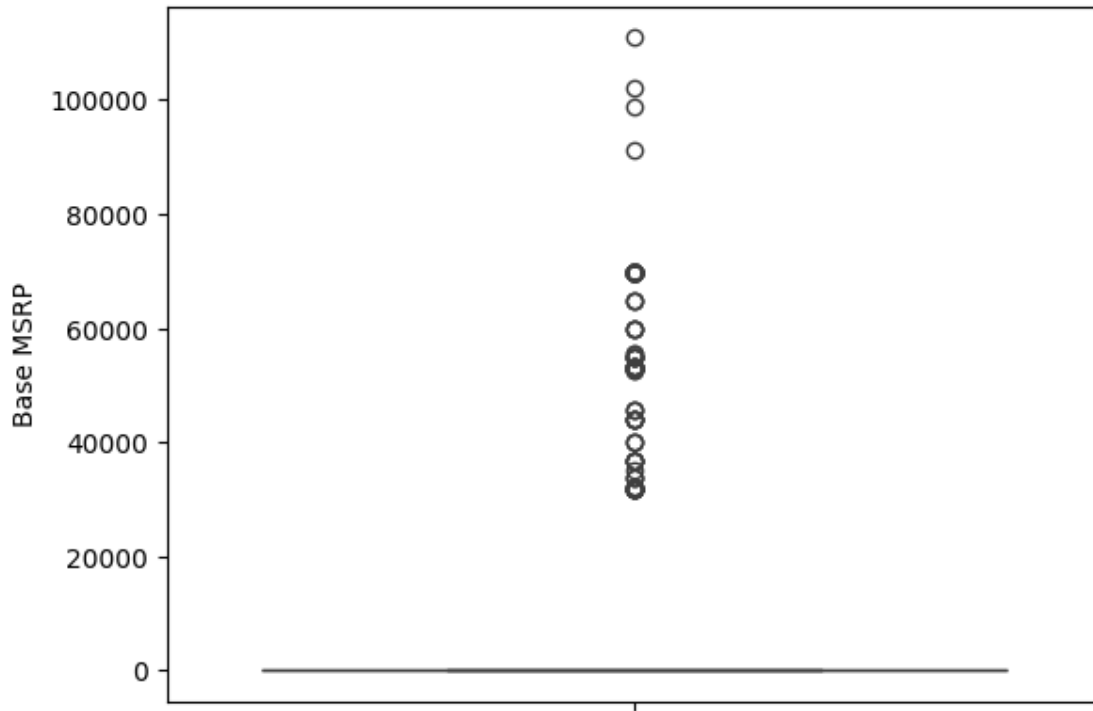
	Vehicle Location \
0	POINT (-122.30839 47.610365)
1	POINT (-122.179458 47.802589)
2	POINT (-122.34848 47.632405)
3	POINT (-122.03646 47.534065)
4	POINT (-122.55717 47.733415)
...	...
13024	POINT (-122.70302 45.703706)
13025	POINT (-122.180505 47.500055)
13026	POINT (-122.147385 47.599975)
13027	POINT (-122.641835 45.638545)
13028	POINT (-122.382425 47.77279)

	Electric Utility	2020 Census Tract
0	CITY OF SEATTLE - (WA) CITY OF TACOMA - (WA)	5.303301e+10
1	PUGET SOUND ENERGY INC	5.306105e+10
2	CITY OF SEATTLE - (WA) CITY OF TACOMA - (WA)	5.303301e+10
3	PUGET SOUND ENERGY INC  CITY OF TACOMA - (WA)	5.303303e+10
4	PUGET SOUND ENERGY INC	5.303594e+10
...	...	...
13024	BONNEVILLE POWER ADMINISTRATION  PUD NO 1 OF C...	5.301104e+10
13025	PUGET SOUND ENERGY INC  CITY OF TACOMA - (WA)	5.303302e+10
13026	PUGET SOUND ENERGY INC  CITY OF TACOMA - (WA)	5.303302e+10
13027	BONNEVILLE POWER ADMINISTRATION  PUD NO 1 OF C...	5.301104e+10
13028	CITY OF SEATTLE - (WA) CITY OF TACOMA	NaN

[12782 rows x 17 columns]

```
[ ]: sns.boxplot(data['Base MSRP'])
```

```
[ ]: <Axes: ylabel='Base MSRP'>
```



```
[ ]: data['Base MSRP'].min()
```

```
[ ]: 0
```

```
[ ]: data['Base MSRP'].max()
```

```
[ ]: 102000
```

```
[ ]: data[data['Base MSRP']==data['Base MSRP'].min()]
```

```
[ ]:
  VIN (1-10)  County      City State  Postal Code  Model Year  \
0  5YJYGDEE1L   King      Seattle  WA      98122      2020
1  7SAYGDEE9P  Snohomish  Bothell  WA      98021      2023
2  5YJSA1E4XK   King      Seattle  WA      98109      2019
3  5YJSA1E27G   King      Issaquah  WA      98027      2016
4  5YJYGDEE5M   Kitsap    Suquamish  WA      98392      2021
...         ...      ...      ...      ...      ...
4350  KMHM34AC5P   King      Kirkland  WA      98034      2023
4351  5YJSA1E2XK   King      Bellevue  WA      98006      2019
4352  5YJ3E1EB9M   Island  Oak Harbor  WA      98277      2021
4353  1G1FZ6S08N  Thurston  Rochester  WA      98579      2022
4354  1N4BZ0CP4G   King      Covington  WA      98042      2016
```

```
      Make      Model      Electric Vehicle Type  \
```

0	TESLA	MODEL Y	Battery Electric Vehicle (BEV)
1	TESLA	MODEL Y	Battery Electric Vehicle (BEV)
2	TESLA	MODEL S	Battery Electric Vehicle (BEV)
3	TESLA	MODEL S	Battery Electric Vehicle (BEV)
4	NaN	MODEL Y	Battery Electric Vehicle (BEV)
...	...	...	...
4350	HYUNDAI	IONIQ 6	Battery Electric Vehicle (BEV)
4351	TESLA	MODEL S	Battery Electric Vehicle (BEV)
4352	TESLA	MODEL 3	Battery Electric Vehicle (BEV)
4353	CHEVROLET	BOLT EV	Battery Electric Vehicle (BEV)
4354	NISSAN	LEAF	Battery Electric Vehicle (BEV)

	Clean Alternative Fuel Vehicle (CAFV) Eligibility	Electric Range \
0	Clean Alternative Fuel Vehicle Eligible	291
1	Eligibility unknown as battery range has not b...	0
2	NaN	270
3	Clean Alternative Fuel Vehicle Eligible	210
4	Eligibility unknown as battery range has not b...	0
...	...	...
4350	Eligibility unknown as battery range has not b...	0
4351	Clean Alternative Fuel Vehicle Eligible	270
4352	Eligibility unknown as battery range has not b...	0
4353	Eligibility unknown as battery range has not b...	0
4354	Clean Alternative Fuel Vehicle Eligible	84

	Base MSRP	Legislative District	DOL Vehicle ID \
0	0	37	125701579
1	0	1	244285107
2	0	36	156773144
3	0	5	165103011
4	0	23	205138552
...	...	...	...
4350	0	45	260328554
4351	0	41	202143378
4352	0	10	181200813
4353	0	20	211177039
4354	0	47	176683636

	Vehicle Location \
0	POINT (-122.30839 47.610365)
1	POINT (-122.179458 47.802589)
2	POINT (-122.34848 47.632405)
3	POINT (-122.03646 47.534065)
4	POINT (-122.55717 47.733415)
...	...
4350	POINT (-122.209285 47.71124)
4351	POINT (-122.16937 47.571015)

```

4352 POINT (-122.6788673 48.2897314)
4353     POINT (-123.09575 46.82114)
4354     POINT (-122.111625 47.36078)

```

	Electric Utility	2020 Census Tract
0	CITY OF SEATTLE - (WA) CITY OF TACOMA - (WA)	53033007800
1	PUGET SOUND ENERGY INC	53061051938
2	CITY OF SEATTLE - (WA) CITY OF TACOMA - (WA)	53033006800
3	PUGET SOUND ENERGY INC  CITY OF TACOMA - (WA)	53033032104
4	PUGET SOUND ENERGY INC	53035940100
...	...	...
4350	PUGET SOUND ENERGY INC  CITY OF TACOMA - (WA)	53033021903
4351	PUGET SOUND ENERGY INC  CITY OF TACOMA - (WA)	53033023902
4352	PUGET SOUND ENERGY INC	53029970402
4353	PUGET SOUND ENERGY INC	53067012730
4354	PUGET SOUND ENERGY INC  CITY OF TACOMA - (WA)	5303303200

```
[4280 rows x 17 columns]
```

### 3.9 1. Descriptive Statistics:

**3.9.1** What are the mean, median, and standard deviation of the base MSRP for the vehicles in the dataset?

#### 3.10 mean

```
[ ]: data['Base MSRP'].mean()
```

```
[ ]: 1119.917875508481
```

#### 3.10.1 Median

```
[ ]: data['Base MSRP'].median()
```

```
[ ]: 0.0
```

Standard Deviation

```
[ ]: data['Base MSRP'].std()
```

```
[ ]: 8595.903627066002
```

### 3.11 2. Data Distribution:

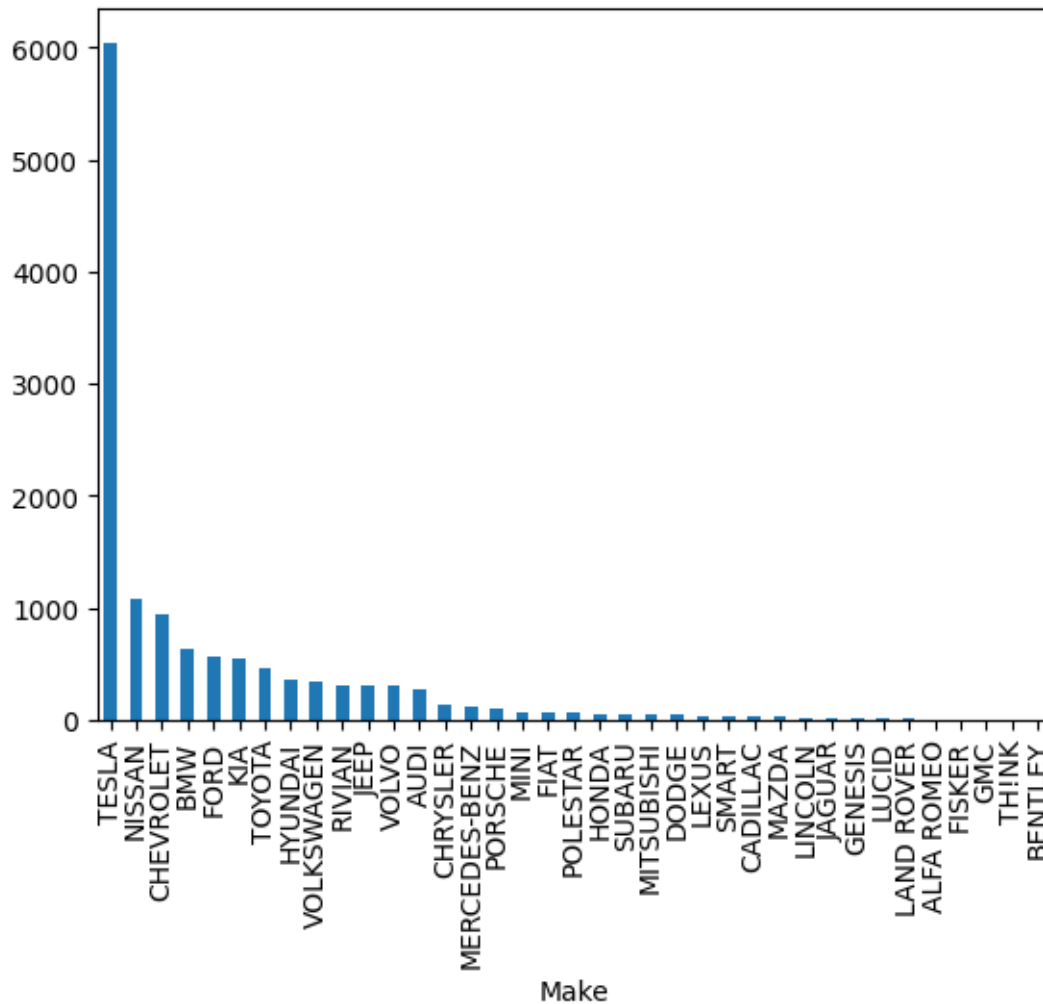
**3.12** What is the distribution of vehicle makes in the dataset? Represent it using a bar chart.

```
[ ]: data['Make'].value_counts()
```

```
[ ]: Make
      TESLA          6047
      NISSAN         1071
      CHEVROLET       935
      BMW            624
      FORD           556
      KIA            539
      TOYOTA         460
      HYUNDAI        359
      VOLKSWAGEN     346
      RIVIAN         307
      JEEP           306
      VOLVO          302
      AUDI           271
      CHRYSLER       137
      MERCEDES-BENZ  119
      PORSCHE        104
      MINI           69
      FIAT           63
      POLESTAR       56
      HONDA          50
      SUBARU         43
      MITSUBISHI     42
      DODGE          42
      LEXUS          29
      SMART          26
      CADILLAC       26
      MAZDA          25
      LINCOLN        15
      JAGUAR         13
      GENESIS        13
      LUCID          12
      LAND ROVER      5
      ALFA ROMEO      3
      FISKER          3
      GMC            2
      TH!NK          1
      BENTLEY         1
      Name: count, dtype: int64
```

```
[ ]: data['Make'].value_counts().plot(kind='bar')
```

```
[ ]: <Axes: xlabel='Make'>
```



### 3.13 3. Model Year Analysis:

3.13.1 What are the most common model years in the dataset? Provide a frequency table and histogram.

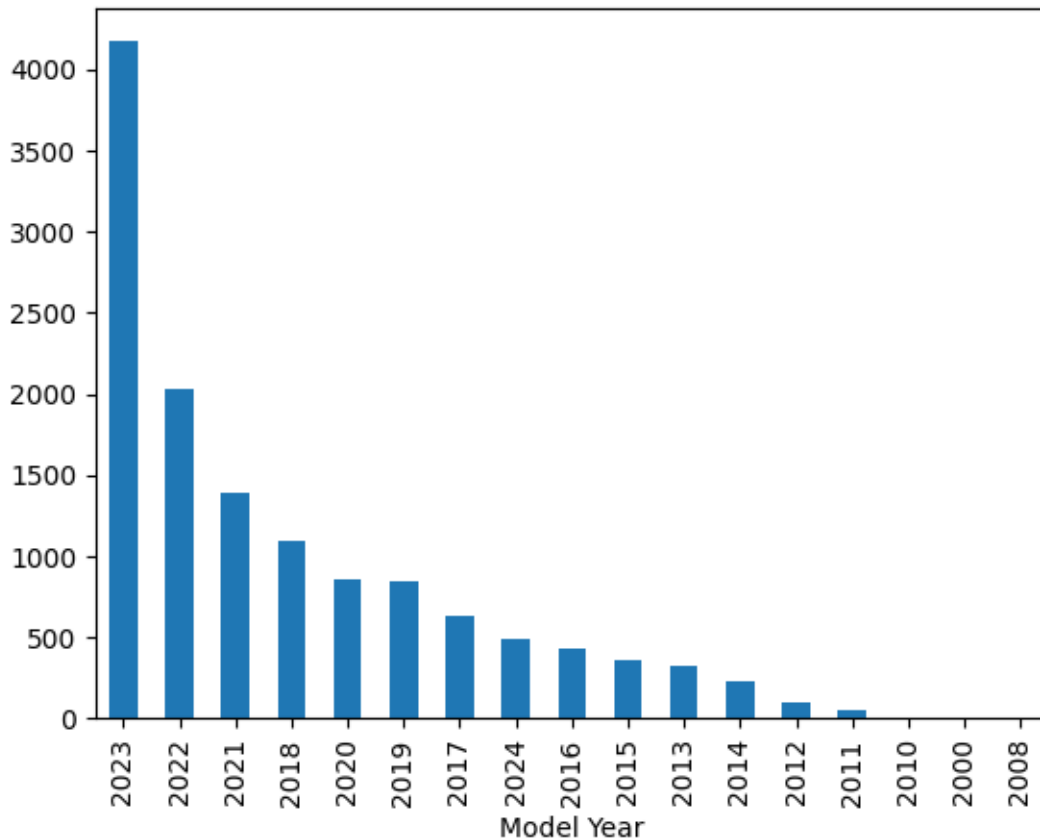
```
[ ]: data['Model Year'].value_counts()
```

```
[ ]: Model Year
2023    4169
2022    2026
2021    1388
2018    1093
2020     863
2019     852
2017     636
2024     496
```

```
2016    431
2015    362
2013    321
2014    236
2012    103
2011     48
2010     3
2000     1
2008     1
Name: count, dtype: int64
```

```
[ ]: data['Model Year'].value_counts().plot(kind='bar')
```

```
[ ]: <Axes: xlabel='Model Year'>
```





### 3.14 4. Electric Vehicle Type:

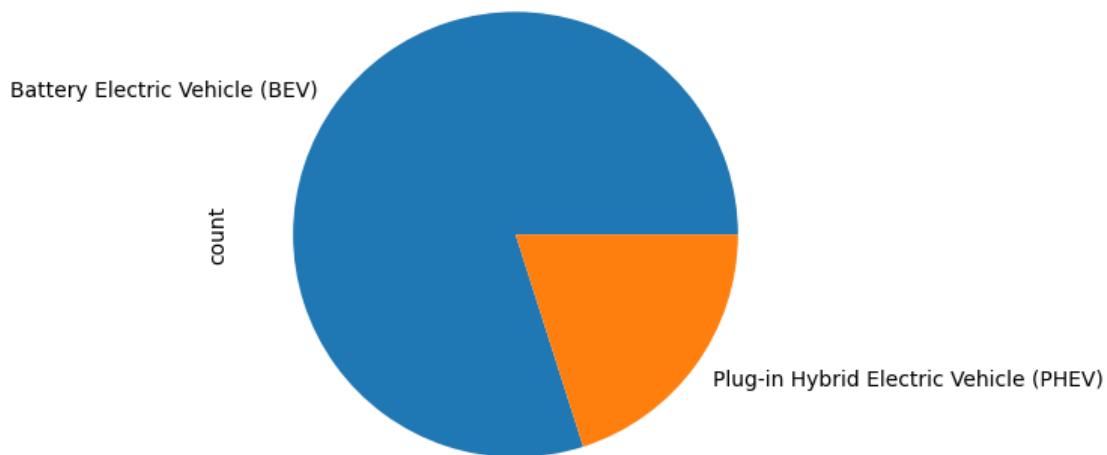
3.14.1 What is the proportion of Battery Electric Vehicles (BEV) versus other types of electric vehicles?

```
[ ]: data['Electric Vehicle Type'].value_counts()
```

```
[ ]: Electric Vehicle Type
      Battery Electric Vehicle (BEV)          10408
      Plug-in Hybrid Electric Vehicle (PHEV)    2615
      Name: count, dtype: int64
```

```
[ ]: data['Electric Vehicle Type'].value_counts().plot(kind='pie')
```

```
[ ]: <Axes: ylabel='count'>
```



### 3.15 6. County Distribution:

3.15.1 How are vehicles distributed across different counties in Washington state? Represent the distribution using a pie chart.

```
[ ]: data['County'].value_counts()
```

```
[ ]: County
      King          92740
      Snohomish     21001
      Pierce        13782
      Clark         10416
      Thurston       6428
```

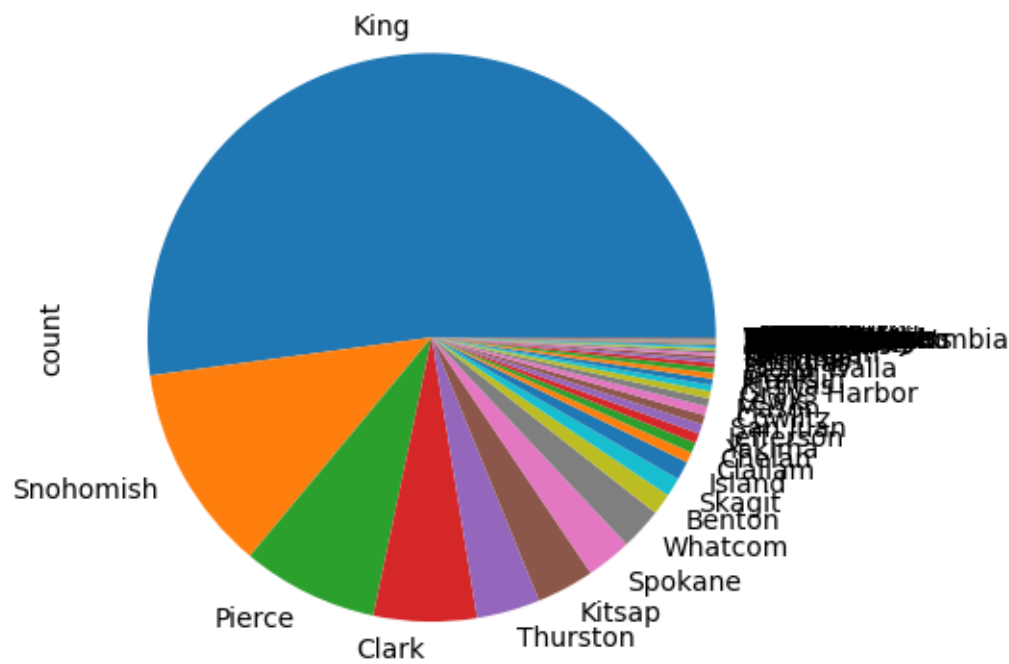
```

Sarpy      1
Suffolk    1
Beaufort   1
Meade      1
Hardin     1
Name: count, Length: 196, dtype: int64

```

```
[ ]: data['County'].value_counts().plot(kind='pie')
```

```
[ ]: <Axes: ylabel='count'>
```



### 3.16 7. Price Analysis:

#### 3.16.1 Compare the average base MSRP of vehicles eligible for the Clean Alternative Fuel Vehicle (CAFV) program versus those that are not.

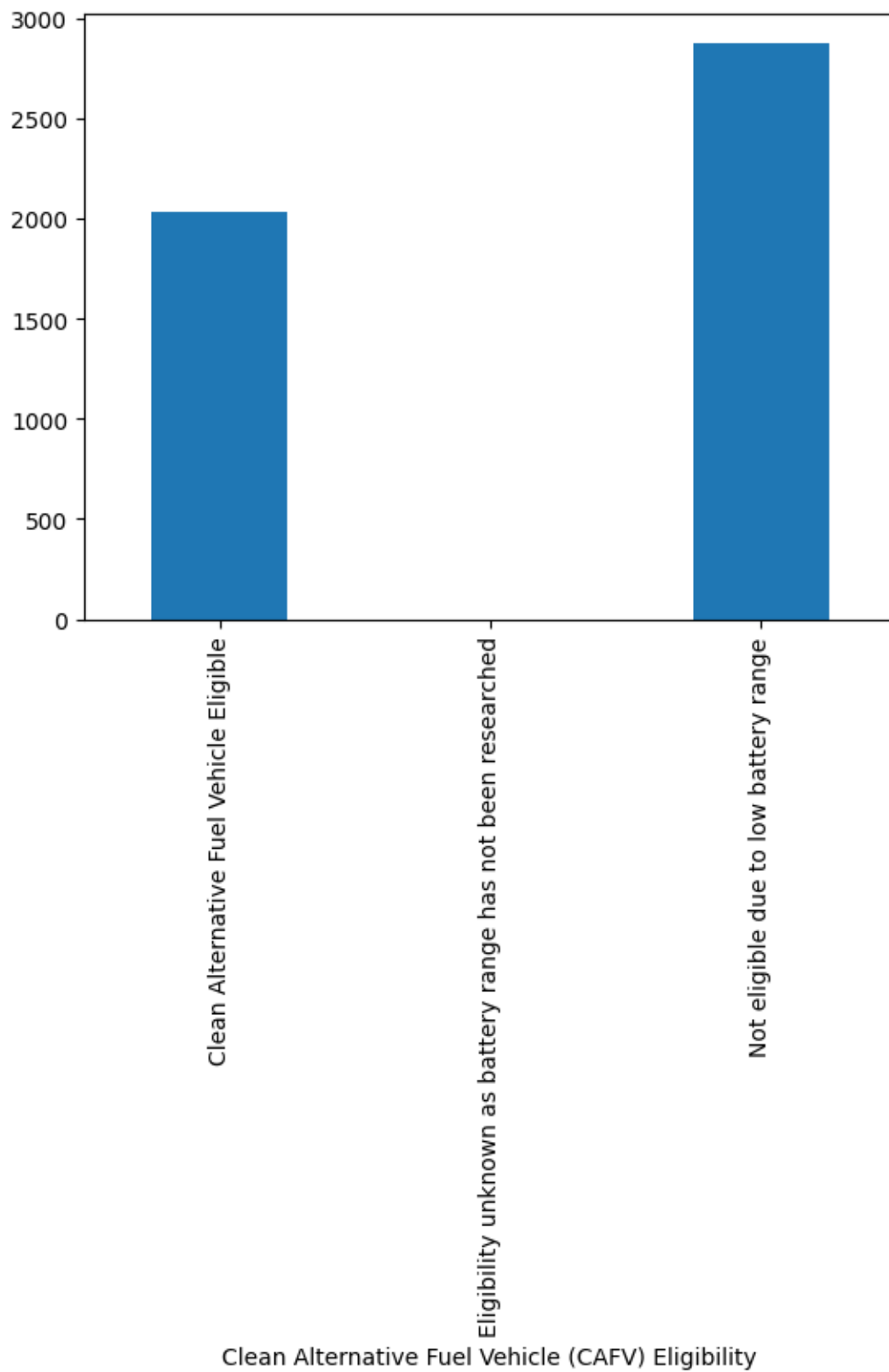
```
[ ]: data.groupby('Clean Alternative Fuel Vehicle (CAFV) Eligibility')['Base MSRP'].
      ↪mean()
```

```
[ ]: Clean Alternative Fuel Vehicle (CAFV) Eligibility
Clean Alternative Fuel Vehicle Eligible      2028.133575
Eligibility unknown as battery range has not been researched      0.000000
Not eligible due to low battery range      2877.020527
```

Name: Base MSRP, dtype: float64

```
[ ]: data.groupby('Clean Alternative Fuel Vehicle (CAFV) Eligibility')['Base MSRP'].  
      ↪mean().plot(kind='bar')
```

```
[ ]: <Axes: xlabel='Clean Alternative Fuel Vehicle (CAFV) Eligibility'>
```



### 3.17 8. Geographical Analysis:

#### 3.17.1 How does the base MSRP vary across different cities in Washington state

```
[ ]: data.groupby('City')['Base MSRP'].mean()
```

```
[ ]: City
      Algona          0.00000
      Anacortes       0.00000
      Ariel           0.00000
      Arlington       0.00000
      Auburn          0.00000
      ...
      Woodway         0.00000
      Yacolt          0.00000
      Yakima          2254.83871
      Yelm            0.00000
      Zillah          0.00000
      Name: Base MSRP, Length: 122, dtype: float64
```

```
[ ]: data.groupby('City')['Base MSRP'].mean().plot(kind='bar')
```

```
[ ]: <Axes: xlabel='City'>
```

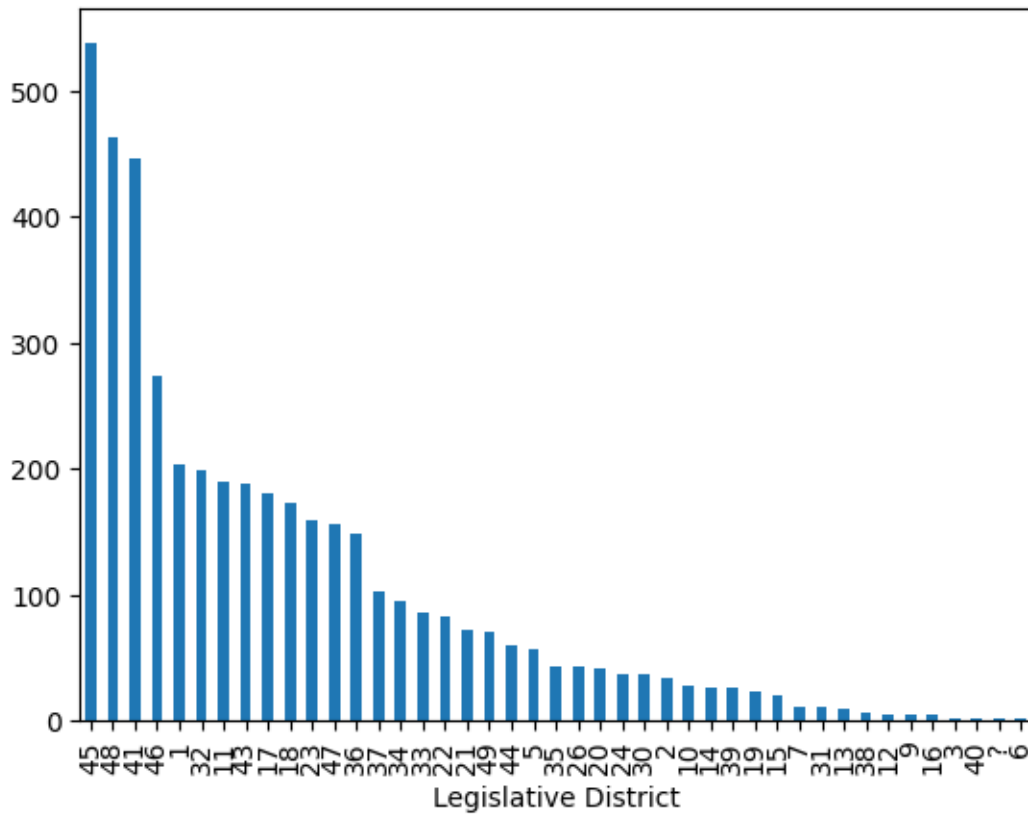


43	188
17	181
18	173
23	159
47	156
36	148
37	102
34	95
33	85
22	82
21	72
49	70
44	60
5	57
35	43
26	43
20	41
24	37
30	37
2	34
10	28
14	26
39	26
19	23
15	20
7	11
31	11
13	9
38	6
12	5
9	4
16	4
3	2
40	2
?	1
6	1

Name: count, dtype: int64

```
[ ]: data['Legislative District'].value_counts().plot(kind='bar')
```

```
[ ]: <Axes: xlabel='Legislative District'>
```



### 3.19 10. Electric Utility Providers:

#### 3.19.1 What is the distribution of electric utility service providers for the vehicles in the dataset?

```
[ ]: data['Electric Utility'].value_counts()
```

```
[ ]: Electric Utility
PUGET SOUND ENERGY INC|CITY OF TACOMA - (WA)
2129
CITY OF SEATTLE - (WA)|CITY OF TACOMA - (WA)
1002
PUGET SOUND ENERGY INC
619
BONNEVILLE POWER ADMINISTRATION|PUD NO 1 OF CLARK COUNTY - (WA)
418
BONNEVILLE POWER ADMINISTRATION|PUD NO 1 OF COWLITZ COUNTY
47
PACIFICORP
46
BONNEVILLE POWER ADMINISTRATION|PUGET SOUND ENERGY INC|PUD NO 1 OF JEFFERSON
```



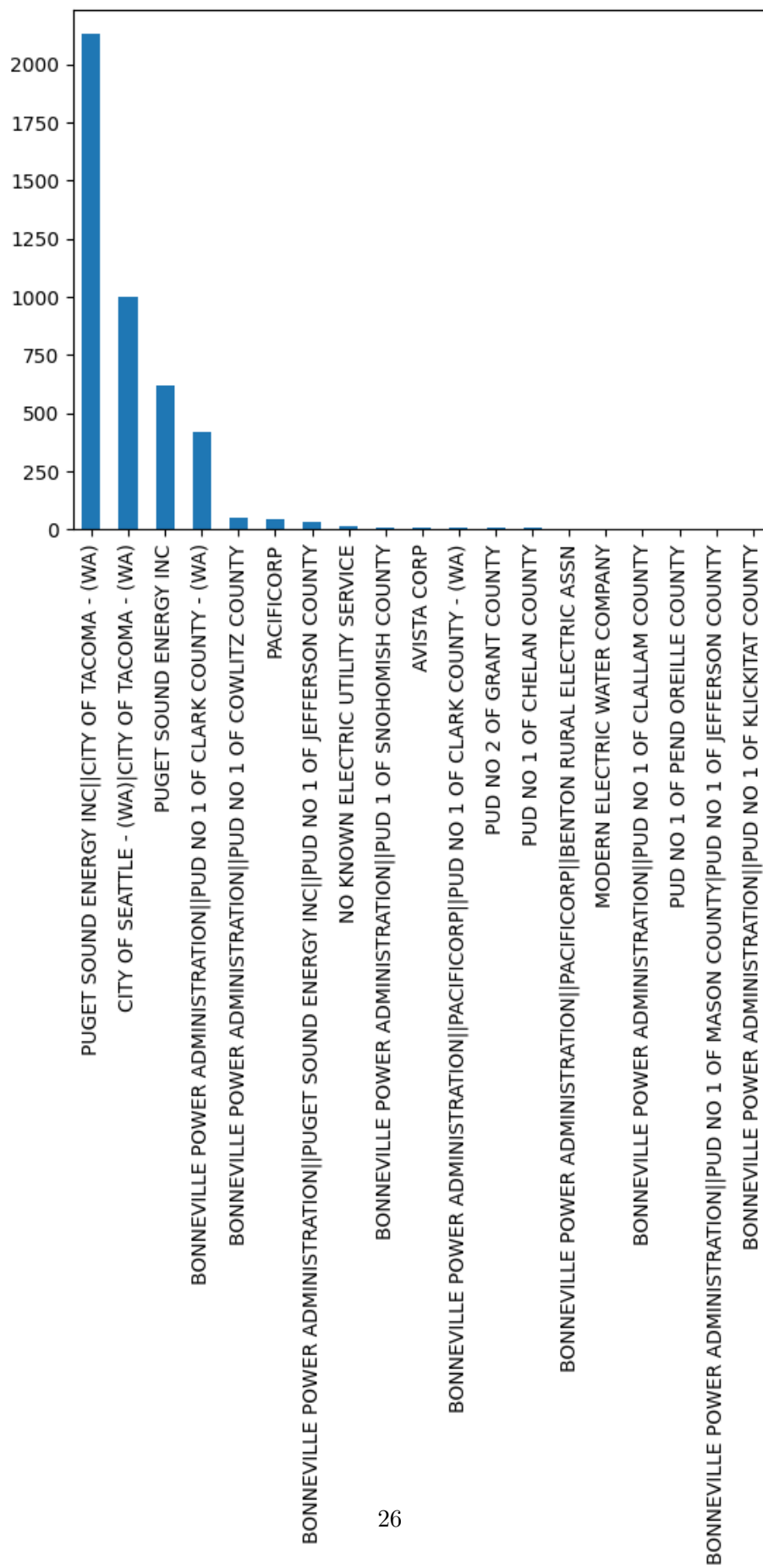
```

COUNTY          34
NO KNOWN ELECTRIC UTILITY SERVICE
11
BONNEVILLE POWER ADMINISTRATION||PUD 1 OF SNOHOMISH COUNTY
10
AVISTA CORP
8
BONNEVILLE POWER ADMINISTRATION||PACIFICORP||PUD NO 1 OF CLARK COUNTY - (WA)
8
PUD NO 2 OF GRANT COUNTY
6
PUD NO 1 OF CHELAN COUNTY
5
BONNEVILLE POWER ADMINISTRATION||PACIFICORP||BENTON RURAL ELECTRIC ASSN
4
MODERN ELECTRIC WATER COMPANY
3
BONNEVILLE POWER ADMINISTRATION||PUD NO 1 OF CLALLAM COUNTY
2
PUD NO 1 OF PEND OREILLE COUNTY
1
BONNEVILLE POWER ADMINISTRATION||PUD NO 1 OF MASON COUNTY|PUD NO 1 OF JEFFERSON
COUNTY          1
BONNEVILLE POWER ADMINISTRATION||PUD NO 1 OF KLINKITAT COUNTY
1
Name: count, dtype: int64

```

```
[ ]: data['Electric Utility'].value_counts().plot(kind='bar')
```

```
[ ]: <Axes: xlabel='Electric Utility'>
```



### 3.20 11. Census Tract Analysis:

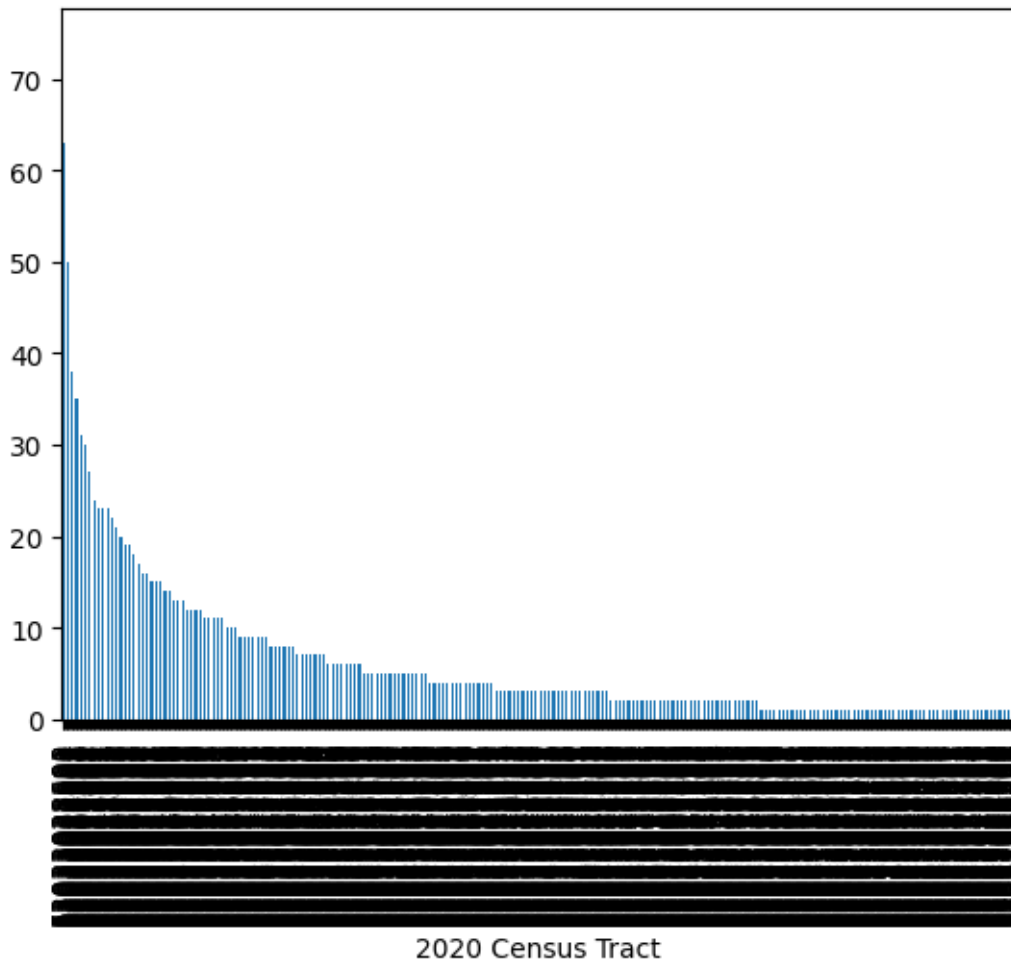
3.20.1 How are vehicles distributed across different 2020 Census Tracts? Provide insights based on vehicle counts per tract.

```
[ ]: data['2020 Census Tract'].value_counts()
```

```
[ ]: 2020 Census Tract
      53033028200      74
      53033022603      63
      53033032321      62
      53033032323      53
      53033021904      50
      ..
      53051970500       1
      53011040413       1
      53035091701       1
      53057940302       1
      5303303200        1
      Name: count, Length: 712, dtype: int64
```

```
[ ]: data['2020 Census Tract'].value_counts().plot(kind='bar')
```

```
[ ]: <Axes: xlabel='2020 Census Tract'>
```



### 3.21 12. Electric Range Correlation:

3.21.1 Is there a correlation between the electric range and the base MSRP of the vehicles? Provide the correlation coefficient and interpret the result.

[ ]:

[ ]:

### 3.22 14. Eligibility Status:

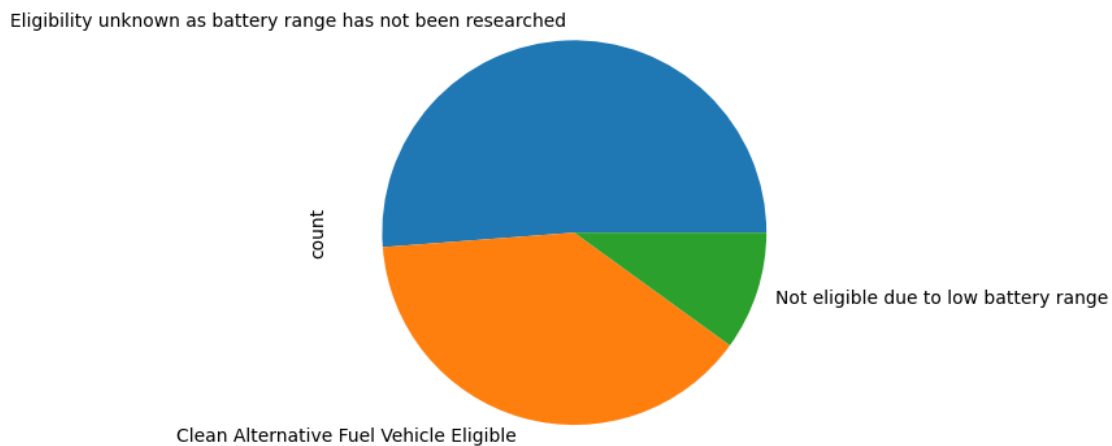
3.22.1 What percentage of vehicles are eligible for the Clean Alternative Fuel Vehicle (CAFV) program?

[ ]:

```
[ ]: Clean Alternative Fuel Vehicle (CAFV) Eligibility
Eligibility unknown as battery range has not been researched    2229
Clean Alternative Fuel Vehicle Eligible                        1690
Not eligible due to low battery range                          434
Name: count, dtype: int64
```

```
[ ]: data['Clean Alternative Fuel Vehicle (CAFV) Eligibility'].value_counts().
      plot(kind='pie')
```

```
[ ]: <Axes: ylabel='count'>
```



### 3.23 15. Model Popularity:

3.23.1 Which vehicle models are the most popular in the dataset? Provide a frequency table of the top 10 models.

```
[ ]: data['Model'].value_counts()
```

```
[ ]: Model
MODEL Y      893
MODEL 3      799
LEAF         378
MODEL S      190
MODEL X      158
...
CROSSTREK     1
I-MIEV        1
I5            1
TONALE        1
740E          1
```

Name: count, Length: 110, dtype: int64

```
[ ]: data['Model'].value_counts().head(10)
```

```
[ ]: Model
MODEL Y      893
MODEL 3      799
LEAF         378
MODEL S      190
MODEL X      158
BOLT EV      156
VOLT         98
X5           89
WRANGLER     89
ID.4         79
Name: count, dtype: int64
```

### 3.24 16. Postal Code Distribution:

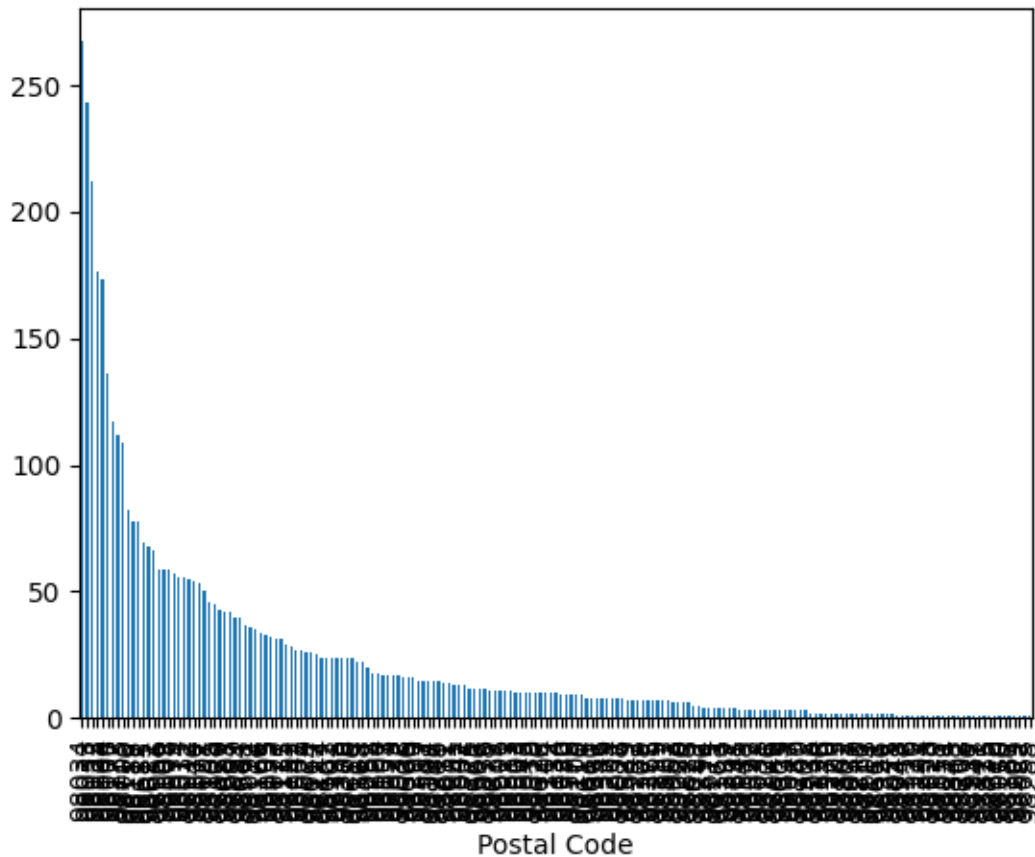
**3.24.1** How are vehicles distributed across different postal codes? Provide a heatmap or density plot.

```
[ ]: data['Postal Code'].value_counts()
```

```
[ ]: Postal Code
98034      267
98033      243
98052      212
98125      176
98004      173
...
98382       1
99109       1
98922       1
98053       1
98288       1
Name: count, Length: 187, dtype: int64
```

```
[ ]: data['Postal Code'].value_counts().plot(kind='bar')
```

```
[ ]: <Axes: xlabel='Postal Code'>
```



### 3.25 17. Vehicle Location Analysis:

3.25.1 Analyze the geographic coordinates to determine any clusters of electric vehicles in certain areas of Washington state.

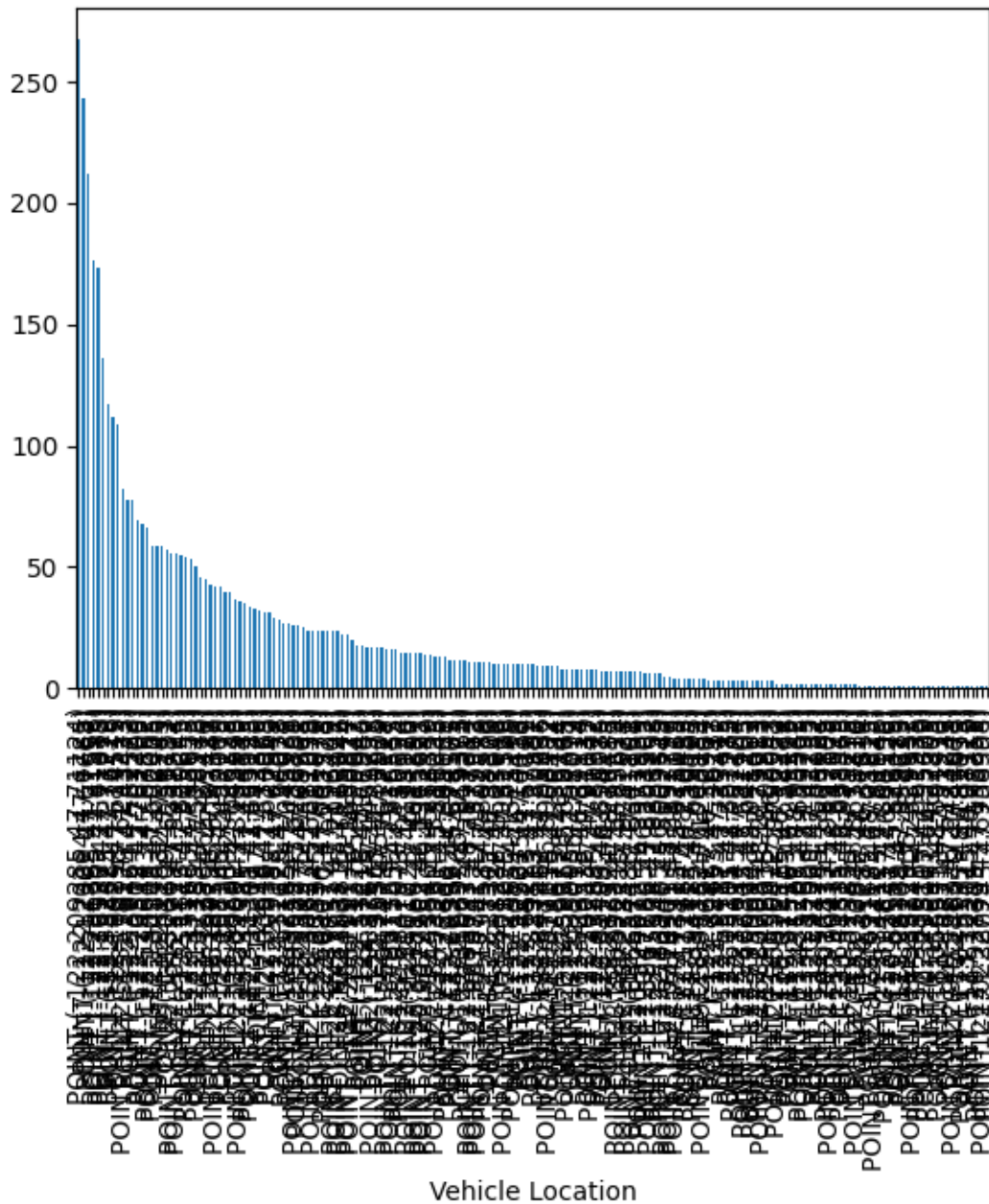
```
[ ]: data['Vehicle Location'].value_counts()
```

```
[ ]: Vehicle Location
POINT (-122.209285 47.71124)      267
POINT (-122.20264 47.6785)       243
POINT (-122.12302 47.67668)      212
POINT (-122.296385 47.71558)     176
POINT (-122.201905 47.61385)     173

...
POINT (-123.105015 48.08125)      1
POINT (-117.722145 48.27719)      1
POINT (-120.938305 47.195355)     1
POINT (-122.0222799 47.6958998)   1
POINT (-121.360745 47.709505)     1
Name: count, Length: 187, dtype: int64
```

```
[ ]: data['Vehicle Location'].value_counts().plot(kind='bar')
```

```
[ ]: <Axes: xlabel='Vehicle Location'>
```





### 3.26 18. Model Year Trend:

**3.26.1** Analyze the trend in the number of registered electric vehicles by model year. Provide a line chart to show any increase or decrease over the years.

```
[ ]: data['Model Year'].value_counts()
```

```
[ ]: Model Year
      2023      1389
      2022       662
      2021       469
      2018       390
      2019       308
      2020       285
      2017       193
      2024       156
      2016       144
      2015       125
      2013       114
      2014        71
      2012        34
      2011        15
      Name: count, dtype: int64
```

### 3.27 19. Range vs. Year:

**3.27.1** Is there a trend between the model year and the electric range of the vehicles? Provide a scatter plot and analyze the trend.

```
[ ]: data['Model Year'].value_counts()
```

```
[ ]: Model Year
      2023      1389
      2022       662
      2021       469
      2018       390
      2019       308
      2020       285
      2017       193
      2024       156
      2016       144
      2015       125
      2013       114
      2014        71
      2012        34
      2011        15
      Name: count, dtype: int64
```

```
[ ]: data['Electric Range'].value_counts()
```

```
[ ]: Electric Range
0      2228
215    193
220    110
84      99
21      88

...
110     1
62      1
?       1
68      1
203     1
Name: count, Length: 89, dtype: int64
```

### 3.28 20) Legislative District and MSRP:

#### 3.28.1 How does the average base MSRP vary across different legislative districts?

```
[ ]: data['Legislative District'].value_counts()
```

```
[ ]: Legislative District
45    538
48    463
41    447
46    274
1     204
32    198
11    189
43    188
17    181
18    173
23    159
47    156
36    148
37    102
34     95
33     85
22     82
21     72
49     70
44     60
5      57
35     43
26     43
20     41
24     37
30     37
2      34
```

10	28
14	26
39	26
19	23
15	20
7	11
31	11
13	9
38	6
12	5
9	4
16	4
3	2
40	2
?	1
6	1

Name: count, dtype: int64