SUMMER BOOTCAMP PROJECT ON - EV_DATA_ANALYSIS

Submitted by- kamaL AHMAD
SYSTEM ID – 2022603519

Data

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.

```
Start coding or generate with AI.
```

```
'/content'
```

importing required libraries-

```
Start coding or generate with AI.
```

Loading the dataset-

```
Start coding or generate with AI.
```

```
<ipython-input-9-9e9c147b00ad>:1: DtypeWarning: Columns (10,12) have mixed types. Specify dtype option on import o
  df = pd.read_csv('6-Electric_Vehicle_Population_Data New.csv')
```

1- Display the top 5 rows.

```
Start coding or generate with AI.
```

	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
0	5YJYGDEE1L	King	Seattle	WA	98122.0	2020	TESLA	MODEL Y	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	291	0
1	7SAYGDEE9P	Snohomish	Bothell	WA	98021.0	2023	TESLA	MODEL Y	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0
2	5YJSA1E4XK	King	Seattle	WA	98109.0	2019	TESLA	MODEL S	Battery Electric Vehicle (BEV)	NaN	270	0
3	5YJSA1E27G	King	Issaquah	WA	98027.0	2016	TESLA	MODEL S	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	210	0
4	5YJYGDEE5M	Kitsap	Suquamish	WA	98392.0	2021	NaN	MODEL Y	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	0

2- Display the last 5 rows

Start coding or generate with AI.

	VIN (1-10)	County	City	State	Postal Code	Model Year	Make	Model	Electric Vehicle Type	Clean Alternative Fuel Vehicle (CAFV) Eligibility	Ele
177861	7SAYGDEE3N	Pierce	Bonney Lake	WA	98391.0	2022	TESLA	MODEL Y	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	
177862	KM8K23AG1P	Mason	Shelton	WA	98584.0	2023	HYUNDAI	KONA ELECTRIC	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	
177863	5YJYGDEE6M	Grant	Quincy	WA	98848.0	2021	TESLA	MODEL Y	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	
177864	WVGKMPE27M	King	Black Diamond	WA	98010.0	2021	VOLKSWAGEN	ID.4	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	
177865	5YJ3E1EA8M	Pierce	Tacoma	WA	98422.0	2021	TESLA	MODEL 3	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	

3- Check the shape of dataset.

Start coding or generate with AI.

```
(177866, 17)
```

4- Check the datatypes of each feature.

Start coding or generate with AI.

```
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
```

5- Check the Statistical summary

Start coding or generate with AI.

	Postal Code	Model Year	Base MSRP	DOL Vehicle ID	2020 Census Tract
count	177861.000000	177866.000000	177866.000000	1.778660e+05	1.778610e+05
mean	98172.453506	2020.515512	1073.109363	2.202313e+08	5.297672e+10
std	2442.450668	2.989384	8358.624956	7.584987e+07	1.578047e+09
min	1545.000000	1997.000000	0.000000	4.385000e+03	1.001020e+09
25%	98052.000000	2019.000000	0.000000	1.814743e+08	5.303301e+10
50%	98122.000000	2022.000000	0.000000	2.282522e+08	5.303303e+10
75%	98370.000000	2023.000000	0.000000	2.548445e+08	5.305307e+10
max	99577.000000	2024.000000	845000.000000	4.792548e+08	5.603300e+10

6- Check the null values

Start coding or generate with AI.

```
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
```

7- Check the duplicate values

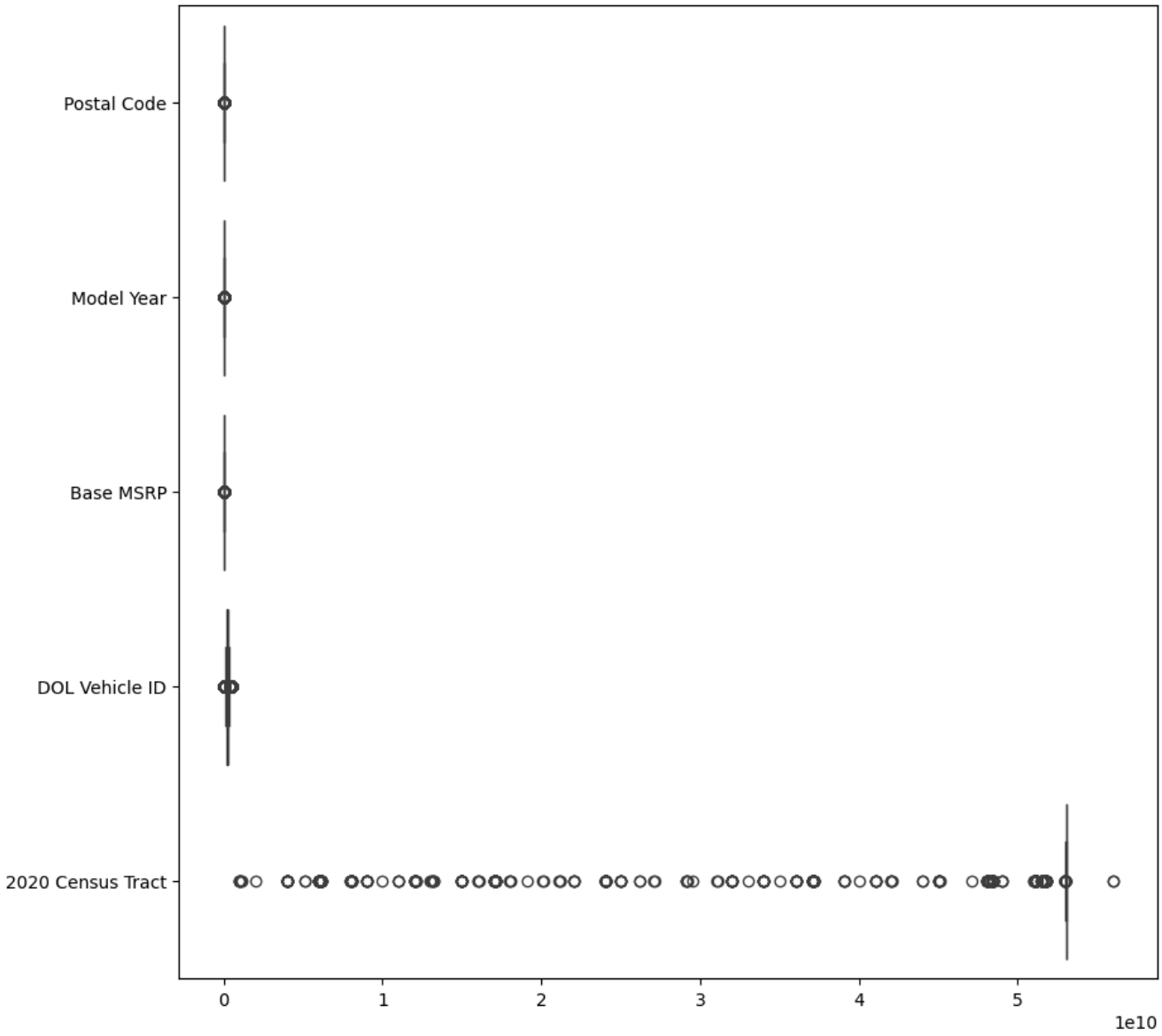
Start coding or generate with AI.

```
Number of duplicate rows: 0
```

8- Check the anomalies or wrong entries.

Start coding or generate with AI.

```
<Axes: >
```



9- Check the outliers and their authenticity.

Start coding or generate with AI.

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

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

Start coding or generate with AI.

Start coding or generate with AI.

Start coding or generate with AI.

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

Start coding or generate with AI.

	VIN (1-10)	County	City	State	Postal Code	Model Year	Make	Model	Electric Vehicle Type	Clean Alternative Fuel Vehicle (CAFV) Eligibility	Electric Range	E M
2746	5YJ3E1EA6M	Clark	Vancouver	WA	98682.0	2021	TESLA	MODEL 3	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	
25132	7SAYGDEE6P	King	Bellevue	WA	98006.0	2023	TESLA	MODEL Y	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	
175816	1N4AZ0CP3D	Grays Harbor	Montesano	WA	98563.0	2013	NISSAN	LEAF	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	75	
938	WA1VCBGE5N	King	Kirkland	WA	98033.0	2022	AUDI	E-TRON	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	
136390	7SAYGDEE7N	Benton	Kennewick	WA	99338.0	2022	TESLA	MODEL Y	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0	

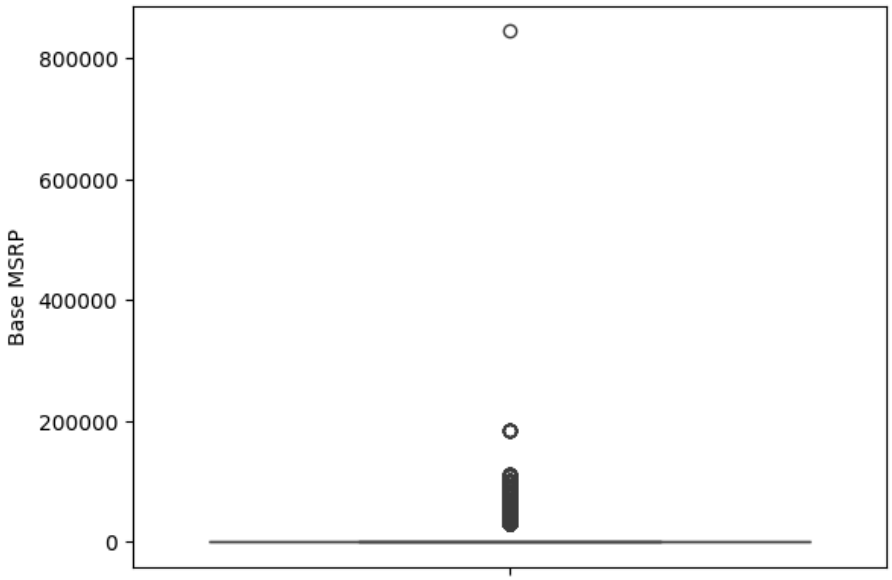
Start coding or generate with AI.

	VIN (1-10)	County	City	State	Postal Code	Model Year	Make	Model	Electric Vehicle Type	Clean Alternative Fuel Vehicle (CAFV) Eligibility
0	5YJYGDEE1L	King	Seattle	WA	98122.0	2020	TESLA	MODEL Y	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible
1	7SAYGDEE9P	Snohomish	Bothell	WA	98021.0	2023	TESLA	MODEL Y	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...
2	5YJSA1E4XK	King	Seattle	WA	98109.0	2019	TESLA	MODEL S	Battery Electric Vehicle (BEV)	NaN
3	5YJSA1E27G	King	Issaquah	WA	98027.0	2016	TESLA	MODEL S	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible
4	5YJYGDEE5M	Kitsap	Suquamish	WA	98392.0	2021	NaN	MODEL Y	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...
...	...	...	...	...	...	...	...	...	...	...
177861	7SAYGDEE3N	Pierce	Bonney Lake	WA	98391.0	2022	TESLA	MODEL Y	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...
177862	KM8K23AG1P	Mason	Shelton	WA	98584.0	2023	HYUNDAI	KONA ELECTRIC	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...
177863	5YJYGDEE6M	Grant	Quincy	WA	98848.0	2021	TESLA	MODEL Y	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...
177864	WVGKMPE27M	King	Black Diamond	WA	98010.0	2021	VOLKSWAGEN	ID.4	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...
177865	5YJ3E1EA8M	Pierce	Tacoma	WA	98422.0	2021	TESLA	MODEL 3	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...

174522 rows × 17 columns

Start coding or generate with AI.

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



Start coding or generate with AI.

```
0
```

Start coding or generate with AI.

```
845000
```

Start coding or generate with AI.

	VIN (1-10)	County	City	State	Postal Code	Model Year	Make	Model	Electric Vehicle Type	Clean Alternative Fuel Vehicle (CAFV) Eligibility
0	5YJYGDEE1L	King	Seattle	WA	98122.0	2020	TESLA	MODEL Y	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible
1	7SAYGDEE9P	Snohomish	Bothell	WA	98021.0	2023	TESLA	MODEL Y	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...
2	5YJSA1E4XK	King	Seattle	WA	98109.0	2019	TESLA	MODEL S	Battery Electric Vehicle (BEV)	NaN
3	5YJSA1E27G	King	Issaquah	WA	98027.0	2016	TESLA	MODEL S	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible
4	5YJYGDEE5M	Kitsap	Suquamish	WA	98392.0	2021	NaN	MODEL Y	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...
...	...	...	...	...	...	...	...	...	...	...
177861	7SAYGDEE3N	Pierce	Bonney Lake	WA	98391.0	2022	TESLA	MODEL Y	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...
177862	KM8K23AG1P	Mason	Shelton	WA	98584.0	2023	HYUNDAI	KONA ELECTRIC	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...
177863	5YJYGDEE6M	Grant	Quincy	WA	98848.0	2021	TESLA	MODEL Y	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...
177864	WVGKMPE27M	King	Black Diamond	WA	98010.0	2021	VOLKSWAGEN	ID.4	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...
177865	5YJ3E1EA8M	Pierce	Tacoma	WA	98422.0	2021	TESLA	MODEL 3	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...

174522 rows × 17 columns

Start coding or generate with AI.

	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
63678	WP0CA2A13F	King	Hunts Point	WA	98004.0	2015	PORSCHE	918	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	12	845000

1. Descriptive Statistics:

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

Mean :

Start coding or generate with AI.

```
1073.1093632284978
```

Median :

Start coding or generate with AI.

```
0.0
```

Standard deviation :

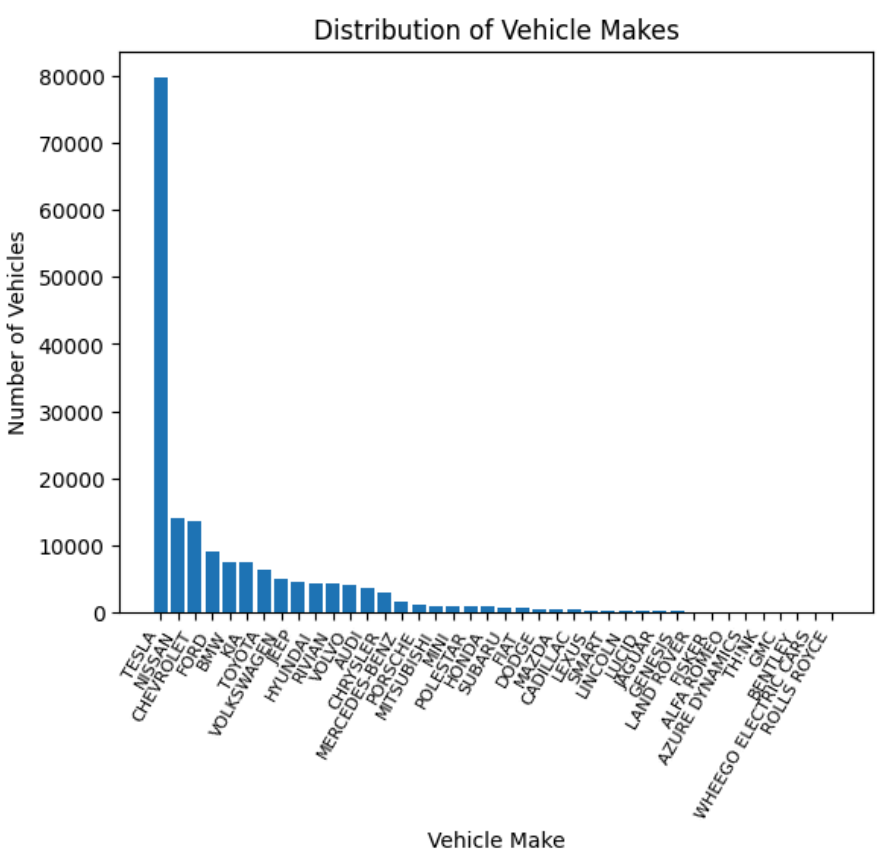
Start coding or generate with AI.

```
8358.624956091093
```

2. Data Distribution:

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

Start coding or generate with AI.

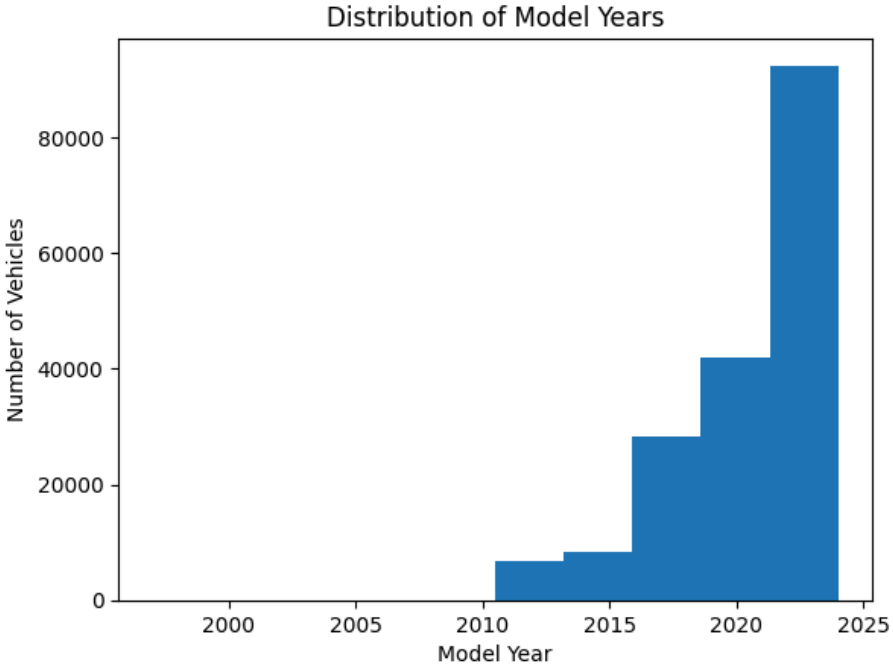


3. Model Year Analysis:

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

Start coding or generate with AI.

```
Model Year
2023    57587
2022    27776
2021    19132
2018    14323
2020    11768
2019    10940
2017     8562
2024     7080
2016     5483
2015     4844
2013     4409
2014     3509
2012     1618
2011      775
2010       23
2008       20
2000        7
1999        5
2002        2
1998        1
1997        1
2003        1
Name: count, dtype: int64
```



4. Electric Vehicle Type:

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

```
Start coding or generate with AI.
```

```
Proportion of BEVs: 0.0
Proportion of other EVs: 1.0
```

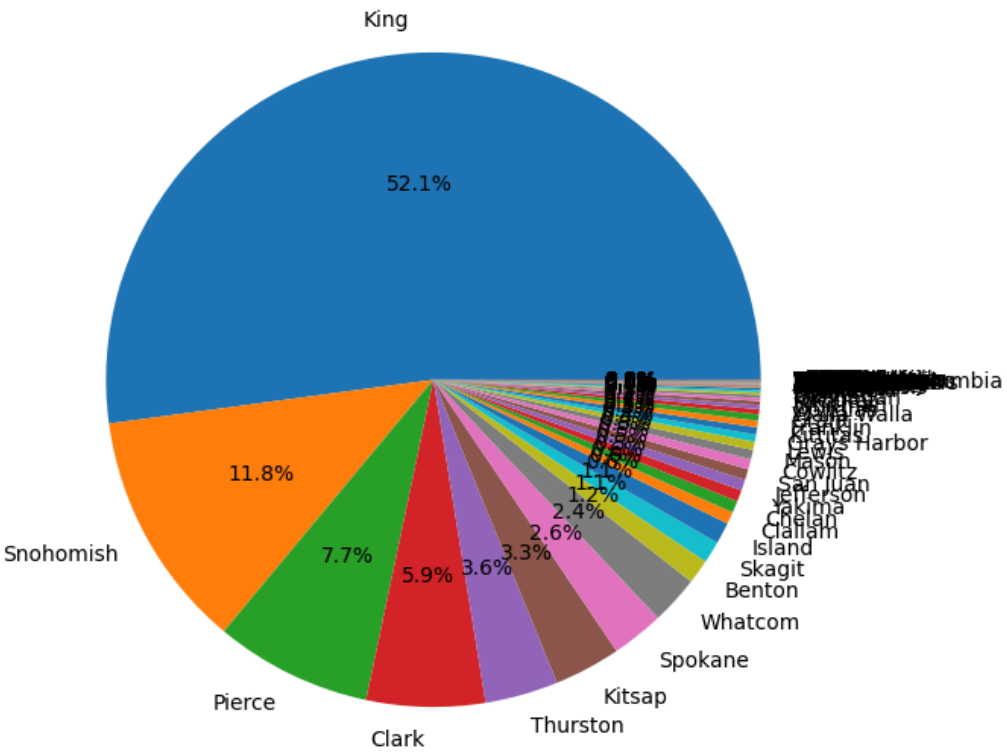
5. Electric Range Analysis:

• What is the average electric range for vehicles of different makes? Provide a summary table.

6. County Distribution:

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

Start coding or generate with AI.



7. Price Analysis:

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

Start coding or generate with AI.

```
Average Base MSRP by CAFV Eligibility:
Clean Alternative Fuel Vehicle (CAFV) Eligibility
Clean Alternative Fuel Vehicle Eligible                          2028.13
Eligibility unknown as battery range has not been researched        0.00
Not eligible due to low battery range                            2877.02
Name: Base MSRP, dtype: float64
```

8. Geographical Analysis:

• How does the base MSRP vary across different cities in Washington state?

Start coding or generate with AI.

```
Average Base MSRP by City:
City
Aberdeen                   2758.97
Aberdeen Proving Ground       0.00
Acme                          0.00
```

```
Adairsville                  0.00
Addy                         0.00
                            ...
Yakima                     532.97
Yarrow Point              1597.59
Yelm                      1523.02
Yorktown                     0.00
Zillah                    1997.14
Name: Base MSRP, Length: 723, dtype: float64
```

9. Legislative Districts:

• Which legislative districts have the highest number of registered electric vehicles? Provide a ranked list.

Start coding or generate with AI.

```
Legislative Districts with the Highest Number of Registered Electric Vehicles:
Legislative District
41.0    8441
45.0    7425
5.0     6810
48.0    6631
1.0     6265
        ...
16         4
3          2
40         2
?          1
6          1
Name: count, Length: 92, dtype: int64
```

10. Electric Utility Providers:

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

Start coding or generate with AI.

```
Distribution of Electric Utility Service Providers:
Electric Utility
PUGET SOUND ENERGY INC||CITY OF TACOMA - (WA)                                    65990
PUGET SOUND ENERGY INC                                                           35882
CITY OF SEATTLE - (WA)|CITY OF TACOMA - (WA)                                     31381
BONNEVILLE POWER ADMINISTRATION||PUD NO 1 OF CLARK COUNTY - (WA)                 10173
BONNEVILLE POWER ADMINISTRATION||CITY OF TACOMA - (WA)||PENINSULA LIGHT COMPANY   7828
                                                                                 ...
BONNEVILLE POWER ADMINISTRATION||PUD NO 1 OF ASOTIN COUNTY||INLAND POWER & LIGHT COMPANY      2
BONNEVILLE POWER ADMINISTRATION||PUD NO 1 OF CLALLAM COUNTY|PUD NO 1 OF JEFFERSON COUNTY      1
BONNEVILLE POWER ADMINISTRATION||PENINSULA LIGHT COMPANY                                      1
CITY OF SEATTLE - (WA)                                                                        1
BONNEVILLE POWER ADMINISTRATION||PUD NO 1 OF JEFFERSON COUNTY                                 1
Name: count, Length: 76, dtype: int64
```

11. Census Tract Analysis:

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

Start coding or generate with AI.

```
Top 10 Census Tracts with the Most Vehicles:
2020 Census Tract
5.303303e+10    2479
5.303303e+10     983
5.303303e+10     820
5.303303e+10     801
5.306701e+10     672
5.303301e+10     651
5.303303e+10     601
```

```
5.306105e+10     581
5.303303e+10     577
5.303302e+10     558
Name: count, dtype: int64
```

12. Electric Range Correlation:

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

13. VIN Analysis:

• Identify any patterns or commonalities in the VIN (1-10) for the vehicles. Are there any frequent prefixes or suffixes?

```
Start coding or generate with AI.
```

```
Most Frequent VIN Prefixes:
VIN (1-10)
7SAYGDEE6P    1239
7SAYGDEE7P    1235
7SAYGDEE8P    1197
7SAYGDEEXP    1191
7SAYGDEE5P    1177
7SAYGDEE2P    1163
7SAYGDEE9P    1161
7SAYGDEE0P    1148
7SAYGDEE3P    1147
7SAYGDEE1P    1119
Name: count, dtype: int64

Most Frequent VIN Suffixes:
VIN (1-10)
EE6P    1565
EE7P    1526
EE2P    1501
EE5P    1483
EE8P    1478
EEXP    1470
EE9P    1470
EE3P    1446
EE0P    1431
EE4P    1411
Name: count, dtype: int64
```

14. Eligibility Status:

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

```
Start coding or generate with AI.
```

```
0.00% of vehicles are eligible for the CAFV program.
```

15. Model Popularity:

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

```
Start coding or generate with AI.
```

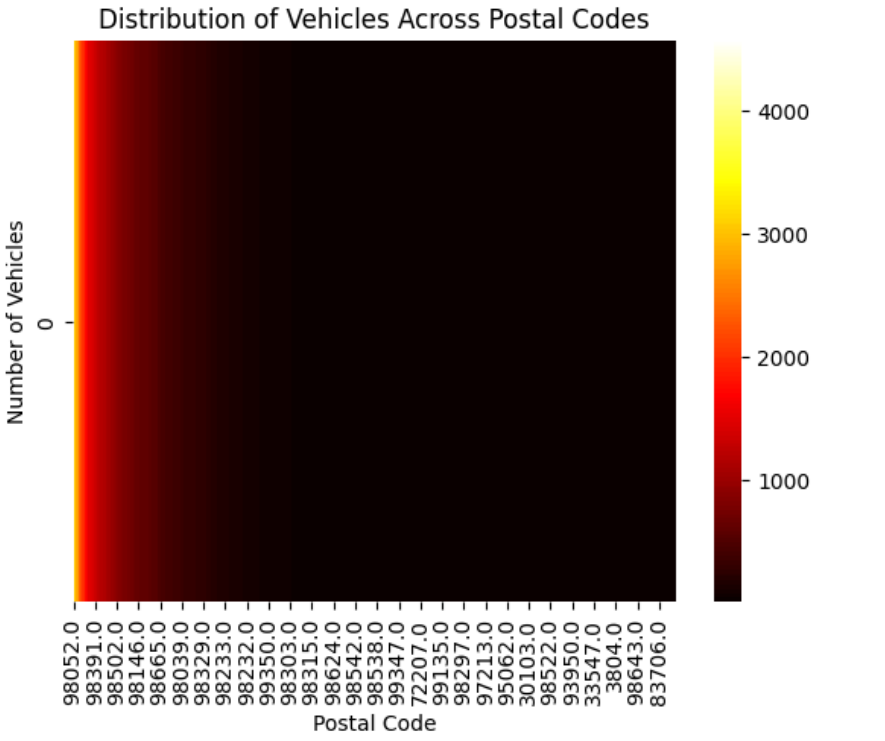
```
Top 10 Most Popular Vehicle Models:
Model
MODEL Y           35989
MODEL 3           30091
LEAF              13365
```

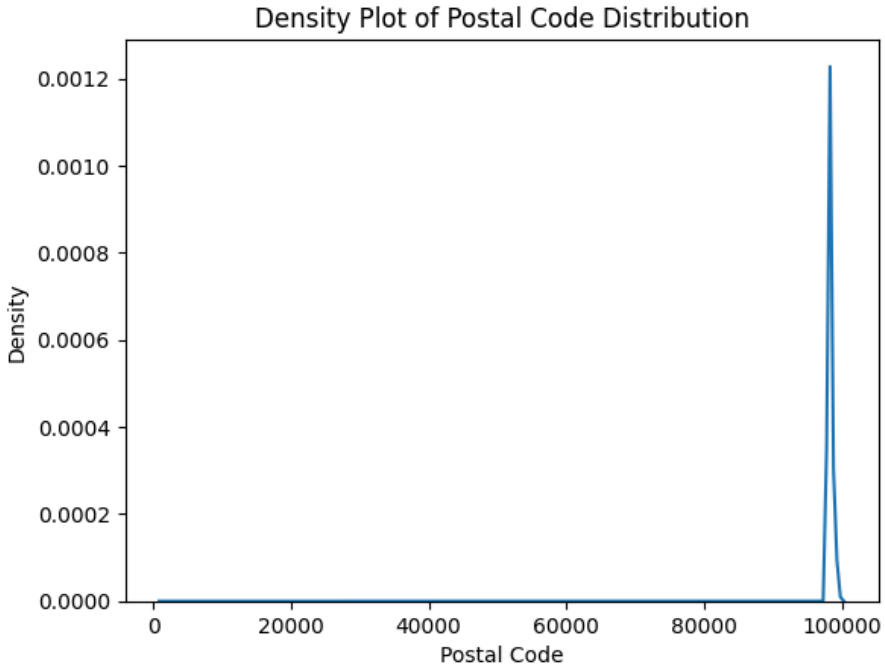
```
MODEL S            7734
BOLT EV            6821
MODEL X            5796
VOLT               4796
ID.4               3937
WRANGLER           3392
MUSTANG MACH-E     3322
Name: count, dtype: int64
```

16. Postal Code Distribution:

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

Start coding or generate with AI.





17. Vehicle Location Analysis:

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

18. Model Year Trend:

• 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.

Start coding or generate with AI.