Electrical and Computer Engineering Department

Machine Learning and Data Science

ENCS5341

Assignment #1

Prepared By:

Student Name	Student Number
Rahaf Naser	1201319
Rania Rimawi	1201179

Instructor's Name: Ismail Khater

Section: 2

October 29, 2024

About Dataset

View the column names

```
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')
```

Get a sneak peek of our data

	VIN (1-10)	County	City	State	Postal Code	Model Year	Make	Model	Electric Vehicle Type	Clean Alternative Fuel Vehicle (CAFV) Eligibility	Electric Range	Base MSRP	Legislative District	DOL Vehicle ID	Vehicle Location	Electric Utility	2020 Census Tract
0	5UXTA6C0XM	Kitsap	Seabeck	WA	98380.0	2021	BMW	X5	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	30.0	0.0	35.0	267929112	POINT (-122.8728334 47.5798304)	PUGET SOUND ENERGY INC	5.303509e+10
1	5YJ3E1EB1J	Kitsap	Poulsbo	WA	98370.0	2018	TESLA	MODEL 3	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	215.0	0.0	23.0	475911439	POINT (-122.6368884 47.7469547)	PUGET SOUND ENERGY INC	5.303509e+10
2	WP0AD2A73G	Snohomish	Bothell	WA	98012.0	2016	PORSCHE	PANAMERA	Plug-in Hybrid Electric Vehicle (PHEV)	Not eligible due to low battery range	15.0	0.0	1.0	101971278	POINT (-122.206146 47.839957)	PUGET SOUND ENERGY INC	5.306105e+10
3	5YJ3E1EB5J	Kitsap	Bremerton	WA	98310.0	2018	TESLA	MODEL 3	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	215.0	0.0	23.0	474363746	POINT (-122.6231895 47.5930874)	PUGET SOUND ENERGY INC	5.303508e+10
4	1N4AZ1CP3K	King	Redmond	WA	98052.0	2019	NISSAN	LEAF	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	150.0	0.0	45.0	476346482	POINT (-122.13158 47.67858)	PUGET SOUND ENERGY INC||CITY OF TACOMA	5.303303e+10

View the shape of the dataset

```
(210165, 17)
```

That's mean data set contain of (210165 rows, 17 columns)

View Number of columns and the type of each one

```
RangeIndex: 210165 entries, 0 to 210164
Data columns (total 17 columns):
 #   Column                                             Non-Null Count   Dtype  
---  ------                                             --------------   -----  
 0   VIN (1-10)                                         210165 non-null  object 
 1   County                                             210161 non-null  object 
 2   City                                               210161 non-null  object 
 3   State                                              210165 non-null  object 
 4   Postal Code                                        210161 non-null  float64
 5   Model Year                                         210165 non-null  int64  
 6   Make                                               210165 non-null  object 
 7   Model                                              210165 non-null  object 
 8   Electric Vehicle Type                              210165 non-null  object 
 9   Clean Alternative Fuel Vehicle (CAFV) Eligibility  210165 non-null  object 
 10  Electric Range                                     210160 non-null  float64
 11  Base MSRP                                          210160 non-null  float64
 12  Legislative District                               209720 non-null  float64
 13  DOL Vehicle ID                                     210165 non-null  int64  
 14  Vehicle Location                                   210155 non-null  object 
 15  Electric Utility                                   210161 non-null  object 
 16  2020 Census Tract                                  210161 non-null  float64
dtypes: float64(5), int64(2), object(10)
memory usage: 27.3+ MB
```

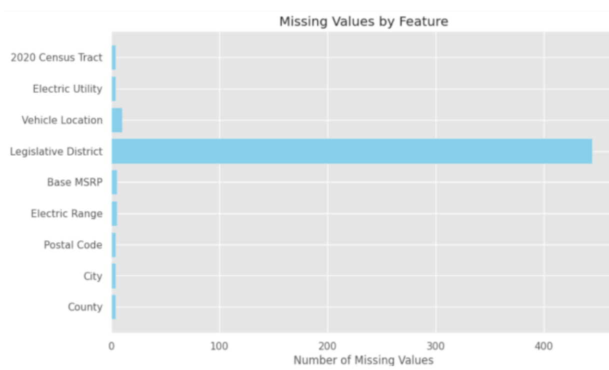
Unique values of all the columns in the data set

```
There are 12377 nos of unique values in VIN (1-10) column out of 210165
There are 203 nos of unique values in County column out of 210165
There are 758 nos of unique values in City column out of 210165
There are 47 nos of unique values in State column out of 210165
There are 931 nos of unique values in Postal Code column out of 210165
There are 21 nos of unique values in Model Year column out of 210165
There are 43 nos of unique values in Make column out of 210165
There are 153 nos of unique values in Model column out of 210165
There are 2 nos of unique values in Electric Vehicle Type column out of 210165
There are 3 nos of unique values in Clean Alternative Fuel Vehicle (CAFV) Eligibility column out of 210165
There are 105 nos of unique values in Electric Range column out of 210165
There are 31 nos of unique values in Base MSRP column out of 210165
There are 49 nos of unique values in Legislative District column out of 210165
There are 210165 nos of unique values in DOL Vehicle ID column out of 210165
There are 931 nos of unique values in Vehicle Location column out of 210165
There are 74 nos of unique values in Electric Utility column out of 210165
There are 2163 nos of unique values in 2020 Census Tract column out of 210165
```

1.Handling the missing values

A) number of missing values and percentage

	Missing Values	Percentage
County	4	0.001903
City	4	0.001903
Postal Code	4	0.001903
Electric Range	5	0.002379
Base MSRP	5	0.002379
Legislative District	445	0.211738
Vehicle Location	10	0.004758
Electric Utility	4	0.001903
2020 Census Tract	4	0.001903



We notice that Legislative District has the most number of missing Values.

B) records with missing values

	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	202
7440	WP1BM2AY4R	King	Kirkland	WA	98033.0	2024	PORSCHE	CAYENNE	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	NaN	NaN	45.0	270889363	POINT (-122.1925969 47.676241)	PUGET SOUND ENERGY INC||CITY OF TACOMA - (WA)	5.30
23186	WP1BM2AY1R	King	Seattle	WA	98108.0	2024	PORSCHE	CAYENNE	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	NaN	NaN	11.0	272220189	POINT (-122.3173531 47.5484673)	CITY OF SEATTLE - (WA)|CITY OF TACOMA - (WA)	5.30
40238	1GKB0RDCXR	Macomb	Warren	MI	48092.0	2024	GMC	HUMMER EV SUV	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0.0	0.0	NaN	273369309	POINT (-83.064593 42.512487)	NON WASHINGTON STATE ELECTRIC UTILITY	2.60
45209	1GKB0RDC0R	Macomb	Warren	MI	48092.0	2024	GMC	HUMMER EV SUV	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0.0	0.0	NaN	269827470	POINT (-83.064593 42.512487)	NON WASHINGTON STATE ELECTRIC UTILITY	2.60
53795	JTMAB3FV4N	Polk	Salem	OR	97304.0	2022	TOYOTA	RAV4 PRIME	Plug-in Hybrid Electric	Clean Alternative	42.0	0.0	NaN	218086949	POINT (-123.0771613	NON WASHINGTON STATE	4.10

2. Missing Value Strategies

We make handling to missing data in several methods:

A) Drop rows with missing values

```
After Dropping Rows with Missing Values:
Shape of original data: (210165, 17)
Shape after dropping rows: (209709, 17)
```

B) Handling missing data by filling(mean, median, mode).

Electric Utility

- Missing values are filled with the label 'Utility Not Available', which indicates that this information was not recorded.

Legislative District

- Missing values are replaced with 'Unknown'

. Vehicle Location

- Missing entries are filled with 'Unknown'.

Model

- fill missing values with the mode (most frequent value).

2020 Census Tract

- Missing values are filled with 'Unknown'.

Electric Range

- Missing values are filled with 'Unknown'

Base MSRP

- Missing values are filled with 'Unknown'

City

- fill missing values with the mode (most frequent value).

Postal Code

- Missing postal codes are filled with the **mean** value of Postal Code.

County

- Missing values are filled with county mode

Compare between the two methods :

Dropping missing value method will reduce the number of rows In data set , however the filling missing value method the number or rows in data set not changed.

Check number of missing values after handling: from table below we notice that after handling missing values the number of missing values will become zero.

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

3. Feature Encoding

A) One Hot Encoding

One-Hot Encoding represents each category as a separate column, making it suitable for non-ordinal (unordered) categorical data.

For each unique category in a categorical feature, One-Hot Encoding creates a new value (True or False) for each column. If a row belongs to a particular category, that column gets a value of Yes ; otherwise, it gets a No.

```
Shape of the dataset after One-Hot Encoding: (210165, 1287)
```

Make_SMART	Make_SUBARU	Make_TESLA	Make_TH!NK	Make_TOYOTA	Make_VINFAST	Make_VOLKSWAGEN	Model_LEAF	Model_LYRIQ	Model_MODEL 3	Model_MODEL S	Model_MODEL X
False	False	False	False	False	False	False	False	False	False	False	False
False	False	True	False	False	False	False	False	False	True	False	False
False	False	False	False	False	False	False	False	False	False	False	False
False	False	True	False	False	False	False	False	False	True	False	False
False	False	False	False	False	False	False	True	False	False	False	False

The above are the results of one hot encoding in Make and model columns.

We also do the encoding for 'Electric Vehicle Type', 'City', 'State', 'County', 'Clean Alternative Fuel Vehicle (CAFV) Eligibility', 'Electric Utility'

B)Label Encoding

Label encoding is a technique used to convert categorical data into numerical form. Each unique category in these columns is assigned a unique integer label, allowing algorithms to interpret these as numerical features.

```
Shape of the dataset after Label Encoding: (210165, 17)
```

Shape of the dataset after Label Encoding: (210165, 17)

	VIN (1-10)	County	City	State	Postal Code	Model Year	Make	Model	Electric Vehicle Type	Clean Alternative Fuel Vehicle (CAFV) Eligibility	Electric Range	Base MSRP	Legislative District	DOL Vehicle ID	Vehicle Location	Electric Utility	2020 Census Tract
0	5UXTA6C0XM	87	595	44	98380.0	2021	5	147	1	0	30.0	0.0	35.0	267929112	POINT (-122.8728334 47.5798304)	71	53035091301.0
1	5YJ3E1EB1J	87	524	44	98370.0	2018	36	88	0	0	215.0	0.0	23.0	475911439	POINT (-122.6368884 47.7469547)	71	53035091100.0
2	WP0AD2A73G	169	61	44	98012.0	2016	30	100	1	2	15.0	0.0	1.0	101971278	POINT (-122.206146 47.839957)	71	53061052009.0
3	5YJ3E1EB5J	87	64	44	98310.0	2018	36	88	0	0	215.0	0.0	23.0	474363746	POINT (-122.6231895 47.5930874)	71	53035080200.0
4	1N4AZ1CP3K	85	546	44	98052.0	2019	28	86	0	0	150.0	0.0	45.0	476346482	POINT (-122.13158 47.67858)	72	53033032323.0

We notice that in one hot encoding the number of columns will increase, but in label encoding not changed.

4. Normalization

A)We apply Min-Max scaling for numerical values.

Min-Max Scaling:

- Purpose: Compresses data to a specific range, typically [0, 1].
- Effect: Ensures all values fall within a set range .

Sample of normalized features:

	Base MSRP	Electric Range	Model Year	Postal Code	Legislative District
0	0.0	0.089021	0.846154	0.270328	0.708333
1	0.0	0.637982	0.730769	0.263195	0.458333
2	0.0	0.044510	0.653846	0.007846	0.000000
3	0.0	0.637982	0.730769	0.220399	0.458333
4	0.0	0.445104	0.769231	0.036377	0.916667

B)*Apply Standard Scaling* for numerical values.

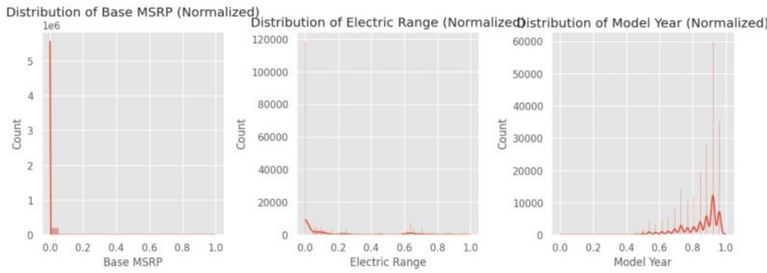
Standard Scaling:

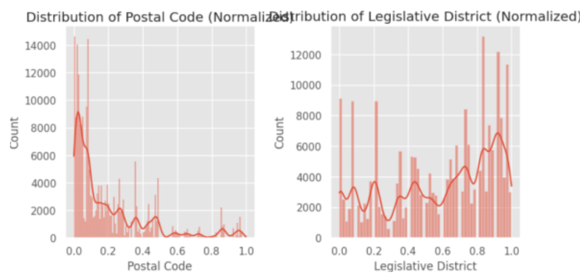
- Purpose: Standardizes features by removing the mean and scaling to unit variance.
- Effect: Rescales features to have a mean of 0 and standard deviation of 1, but does not bound them within a specific range.

Sample of standardized features:

	Base MSRP	Electric Range	Model Year	Postal Code	Legislative District
0	-0.117204	-0.236792	-0.016539	0.369959	0.407149
1	-0.117204	1.890633	-1.020128	0.337469	-0.397773
2	-0.117204	-0.409286	-1.689188	-0.825702	-1.873464
3	-0.117204	1.890633	-1.020128	0.142524	-0.397773
4	-0.117204	1.143160	-0.685598	-0.695738	1.077918

The figure bellow show the visualization shows the distributions of five normalized numerical features after applying Min-Max.





5. Descriptive Statistics

	Postal Code	Model Year	Electric Range	Legislative District	DOL Vehicle ID
count	210165.000000	210165.000000	210165.000000	210165.000000	210165.000000
unique	nan	nan	106.000000	50.000000	nan
top	nan	nan	0.000000	41.000000	nan
freq	nan	nan	118654.000000	13196.000000	nan
mean	98178.209406	2021.048657	nan	nan	229077434.800799
std	2445.406130	2.988941	nan	nan	71155185.129778
min	1731.000000	1999.000000	nan	nan	4469.000000
25%	98052.000000	2019.000000	nan	nan	194881648.000000
50%	98125.000000	2022.000000	nan	nan	240516391.000000
75%	98374.000000	2023.000000	nan	nan	262975800.000000
max	99577.000000	2025.000000	nan	nan	479254772.000000
median	98125.000000	2022.000000	nan	nan	240516391.000000

- Count: The total number of entries for each column is 210,165, except for Electric Range (106 unique values) and Legislative District (50 unique values).

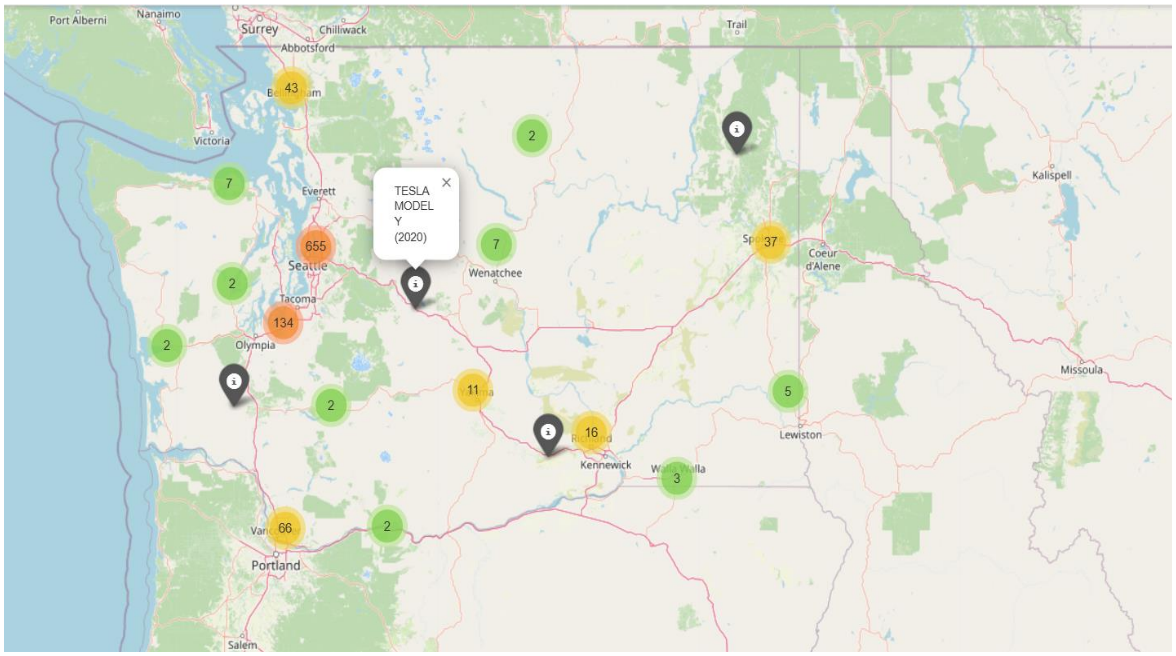
- Unique: This shows the count of unique entries in categorical columns. Since Model Year, Postal Code, and DOL Vehicle ID are not categorical, they show NaN, while Electric Range and Legislative District have 106 and 50 unique values, respectively.

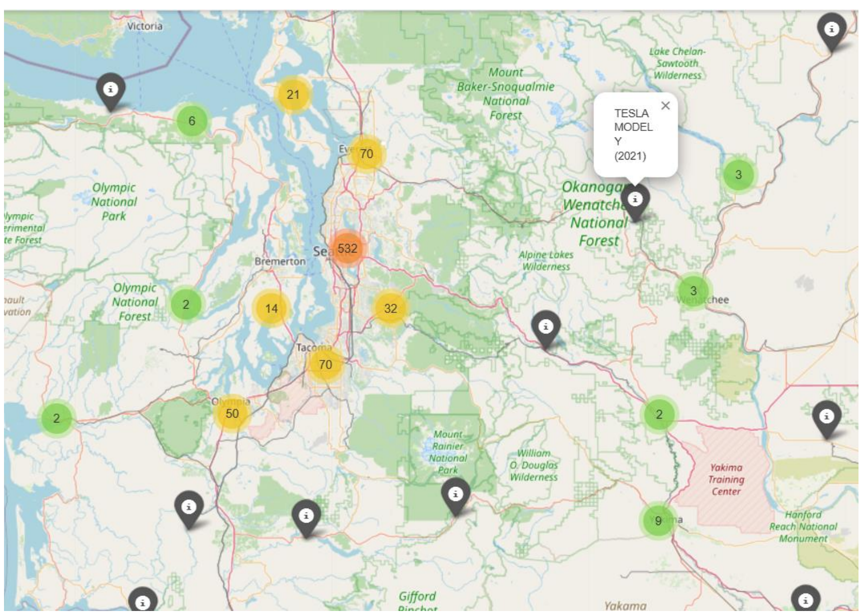
- Top and Freq: The most frequent value (top) and its frequency (freq) in categorical columns. For Electric Range, 0 is the most common value, appearing 118,654 times. For Legislative District, 41 is the top value, appearing 13,196 times.

- Mean, Std, Min, 25%, 50% (Median), 75%, Max: These statistical values are provided for numerical columns. For example, Model Year has an average of around 2021, with a minimum of 1999 and a maximum of 2022. The DOL Vehicle ID column has a mean of around 22,907,744, with a wide range, as shown by its minimum (4,469) and maximum (47,925,477).

6. Spatial Distribution

In this part we read a dataset containing vehicle location data, extracts the latitude and longitude coordinates from the "Vehicle Location" column. A sample of 1,000 entries is taken for efficient processing, and vehicles are color-coded based on their make (Tesla, Nissan, Chevrolet) with a default gray color for unspecified makes. A folium map is then initialized, centered on Washington state, and a marker cluster is added to organize and group the markers for better visualization. Finally, we places color-coded markers on the map for each sampled vehicle, with pop-ups showing the vehicle make, model, and model year.



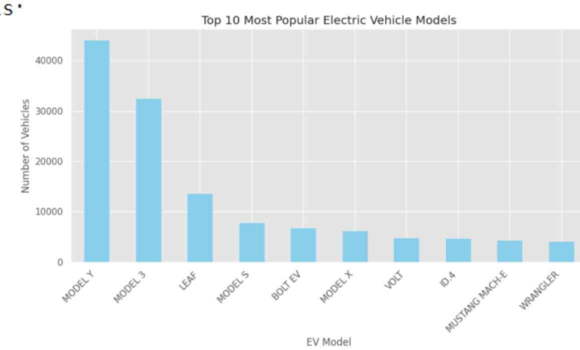


7. Model Popularity

A) We did work in the "Model" column to count the frequencies of each EV model, effectively measuring popularity by the number of vehicles.

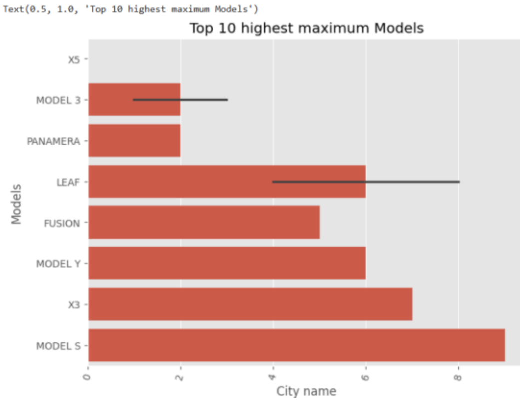
We have therefore presented the top 10 most popular electric car models by number, showing that the Model Y and Model 3 are the most popular by a large margin, with the Model Y leading with 44,038 vehicles, followed by the Model 3 with 32,520 vehicles. The popularity is gradually decreasing with models like Leaf, Model S and Bolt EV making the list.

```
Top 10 Most Popular EV Models:
 Model
MODEL Y           44038
MODEL 3           32520
LEAF              13606
MODEL S            7795
BOLT EV            6780
MODEL X            6239
VOLT               4815
ID.4               4716
MUSTANG MACH-E     4363
WRANGLER           4116
Name: count, dtype: int64
```



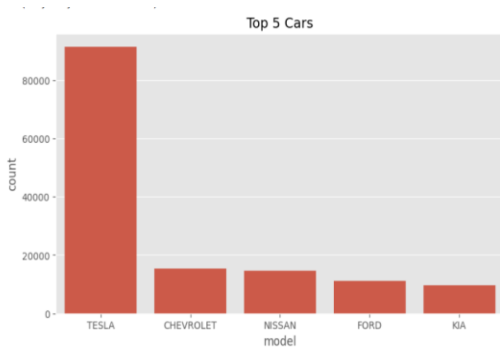
B) We Created a series of all electric car models and their numbers, ranked from most popular to least popular. We then grouped the data by city and retrieved the most recent model year for each city, then sorted the cities in descending order of model year (from newest to oldest).

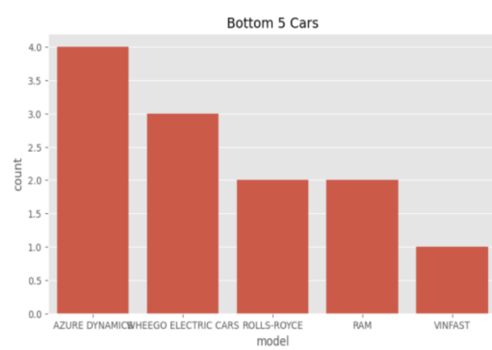
```
Text(0.5, 1.0, 'Top 10 highest maximum Models')
```



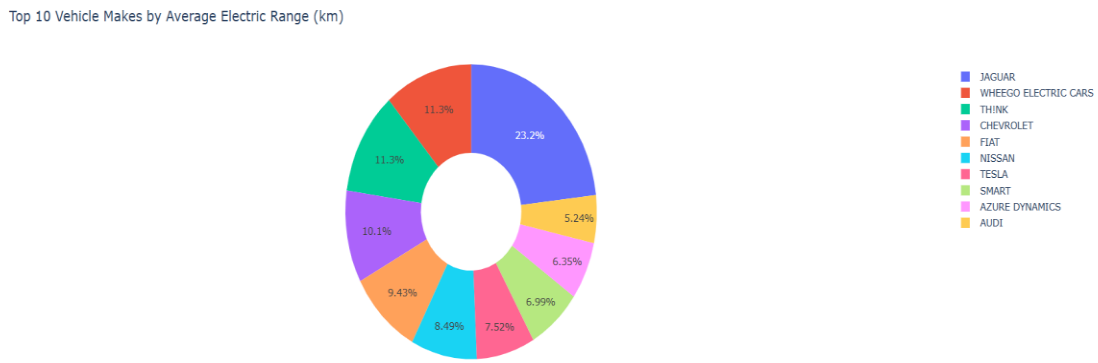
The title "Top 10 Highest Maximum Models" suggests an emphasis on cities with the latest model years.

C) We've compared insights into which manufacturers dominate the electric vehicle market (top 5) and which have a minimal presence (bottom 5).





D) We've illustrated using a pie chart the distribution of the top 10 manufacturers by their average electric range.



8. Investigate the relationship between every pair of numeric features

A) Using a correlation matrix to give insight into the relationships between digital features.

```
Correlation Matrix:
                Postal Code  Model Year  Electric Range  DOL Vehicle ID  \
Postal Code        1.000000   -0.001292       -0.000800        0.005861
Model Year        -0.001292    1.000000       -0.513534        0.215692
Electric Range    -0.000800   -0.513534        1.000000       -0.140689
DOL Vehicle ID     0.005861    0.215692       -0.140689        1.000000
Longitude         -0.751863   -0.002453        0.000990       -0.001487
Latitude           0.385764   -0.001386        0.003481       -0.009177

                Longitude  Latitude
Postal Code     -0.751863  0.385764
Model Year      -0.002453 -0.001386
Electric Range   0.000990  0.003481
DOL Vehicle ID  -0.001487 -0.009177
Longitude        1.000000 -0.375251
Latitude        -0.375251  1.000000
```

Postal Code:

- Longitude (-0.751863) and Latitude (0.385764) have moderate correlations with Postal Code, other correlations are very close to zero, meaning Postal Code has almost no linear relationship with features like Model Year, Electric Range, and DOL Vehicle ID.

Model Year:

- Electric Range (-0.513534): This is a moderate negative correlation, indicating that newer models tend to have shorter electric ranges. DOL Vehicle ID (0.215692) has a small positive correlation, suggesting a weak relationship. No substantial correlation with other features.

Electric Range:

- Model Year (-0.513534), as mentioned, has a moderate negative correlation with Electric Range.Other correlations are very weak, close to zero, showing that Electric Range is generally independent of most other features except Model Year.

DOL Vehicle ID:

- Very low correlation with all other features, indicating that DOL Vehicle ID is largely independent of the rest.

Longitude and Latitude:

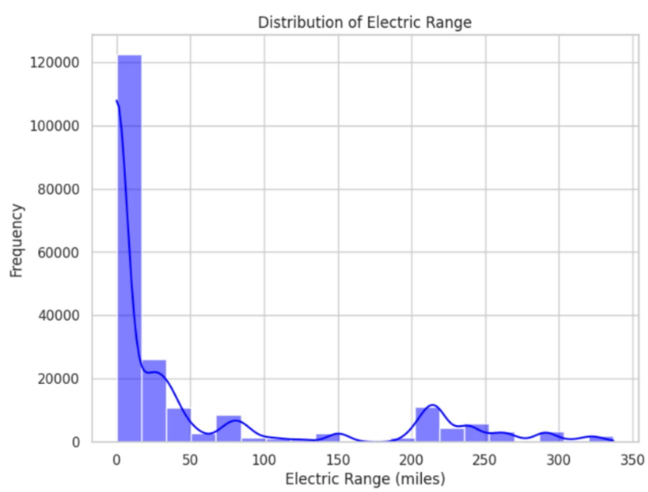
- Longitude and Latitude (-0.375251) have a moderate negative correlation with each other, which makes sense as these two values represent geographic locations. Postal Code shows moderate correlations with both Longitude and Latitude, which is consistent with how geographic data often correlates with postal codes.

B) visualizing the correlation matrix using a heatmap, Visualization provides a more intuitive look at the relationships between numerical features in a data set, making identifying patterns easier compared to a raw correlation matrix table.

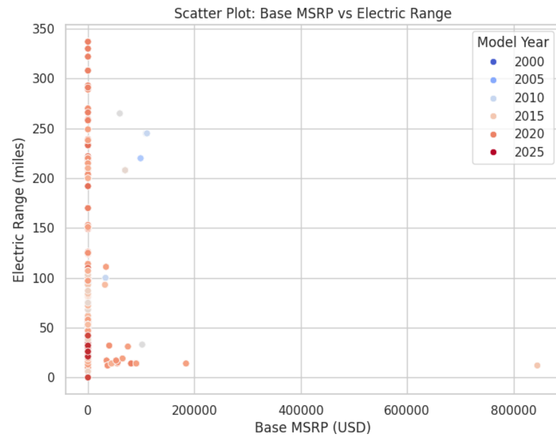
9. Data Exploration Visualizations

A)The graph is skewed to the right, with most vehicles having less electric range (big peak at 0-50 miles).

→Most vehicles have an electric range under 50 miles, likely plug-in hybrids. A smaller group, likely fully electric vehicles, reaches higher ranges (up to 200–250 miles).

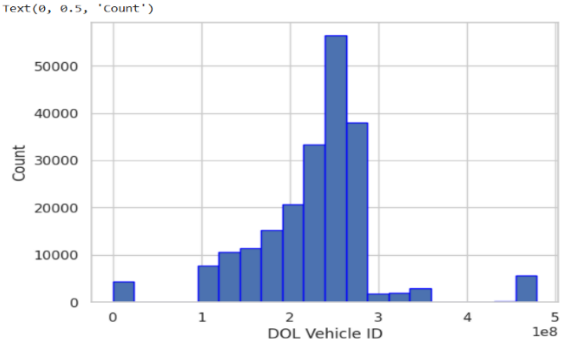


B) Here the scatter plot shows us the relationship between Base MSRP (vehicle price) and Electric Range (miles).

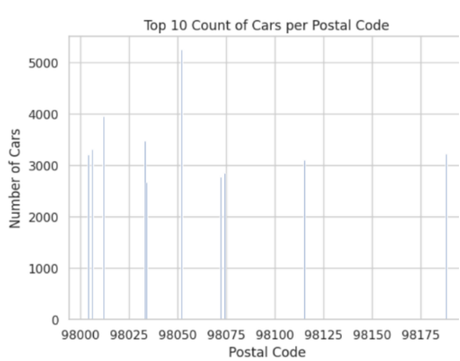


C) The histogram shows us the distribution of DOL Vehicle IDs. Most IDs cluster between 1.5e8 and 3e8, indicating a high frequency in this range. There's a sharp peak near 2.5e8, suggesting a concentration of IDs here, while fewer IDs fall outside this range, especially beyond 3e8.

```
Text(0, 0.5, 'Count')
```

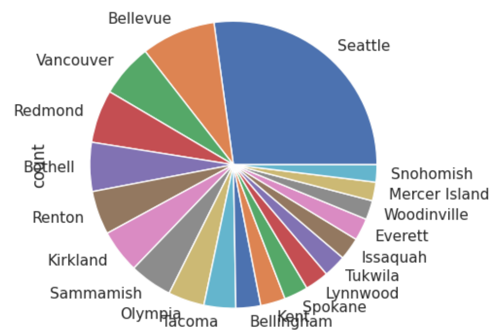


D)The bar chart shows us the number of cars per zip code, where each bar represents a zip code and the height shows the number of cars. The y-axis is labeled "Number of Cars", and the x-axis shows different zip codes. The spreadsheet below this plot ranks zip codes, with ZIP code 98052 having the highest number (5,256 cars) and 98034 having the least in the top 10 (2,675 cars). The table also has a gradient background, with darker blue shades representing higher numbers.

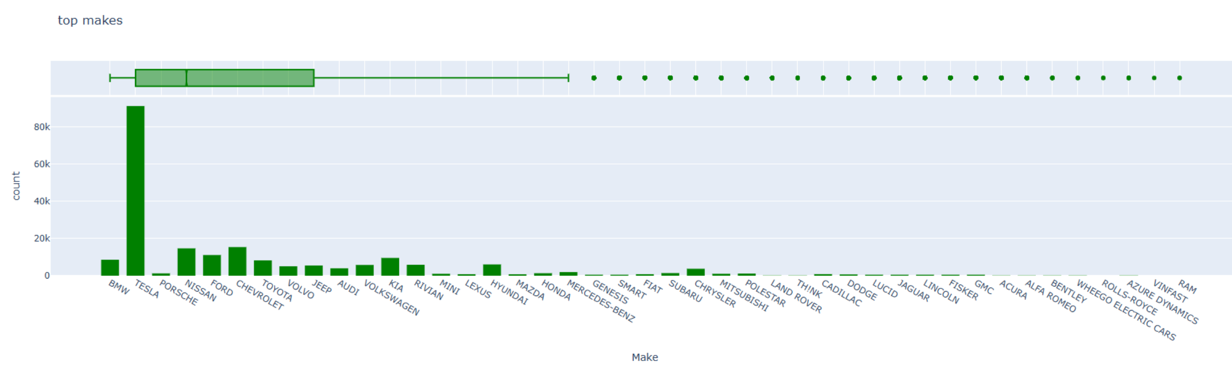


	count
Postal Code	
98052.000000	5256
98012.000000	3959
98033.000000	3485
98006.000000	3314
98188.000000	3229
98004.000000	3209
98115.000000	3103
98074.000000	2863
98072.000000	2779
98034.000000	2675

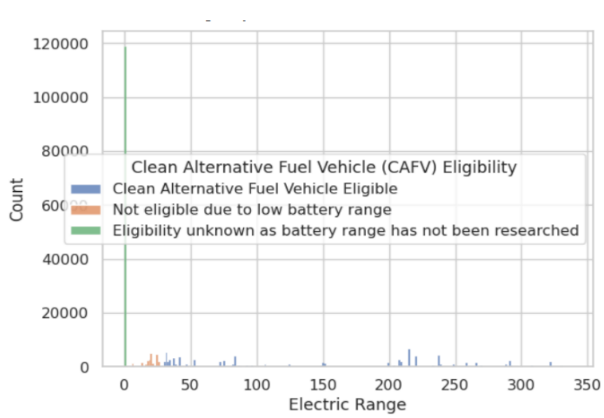
E) Our pie chart depicts the distribution of electric vehicle numbers across the top 20 cities. Each part represents a city, and the size of each part is proportional to the number of electric cars. Seattle has the largest segment, indicating it has the highest number, followed by Bellevue, Vancouver, and Redmond. Smaller cities like Tukwila, Lynwood, and others are represented by smaller segments, indicating lower car counts. The graph provides a clear comparison of the use of electric cars in these cities.



F) This histogram displays the counts of various car makes, with Tesla clearly leading by a substantial margin, showing over 80,000 units. Other notable brands, including BMW, Nissan, Ford, and Chevrolet, follow with much smaller counts.

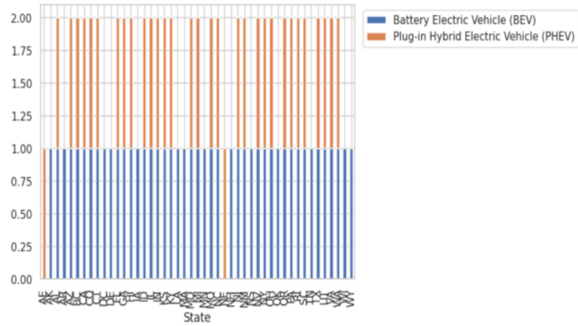


G) This histogram below visualizes the distribution of electric vehicle ranges based on their eligibility for the Clean Alternative Fuel Vehicle (CAFV) program. The x-axis represents the electric range (in miles), while the y-axis shows the count of vehicles. The data is categorized by CAFV eligibility: Blue: Vehicles eligible for the CAFV program. Orange: Vehicles not eligible due to low battery range. Green: Vehicles with unknown eligibility, as their battery range hasn't been researched.



H) The bar chart visualizes the distribution of electric vehicle types across various states, focusing on Battery Electric Vehicles (BEVs) and Plug-in Hybrid Electric Vehicles (PHEVs).

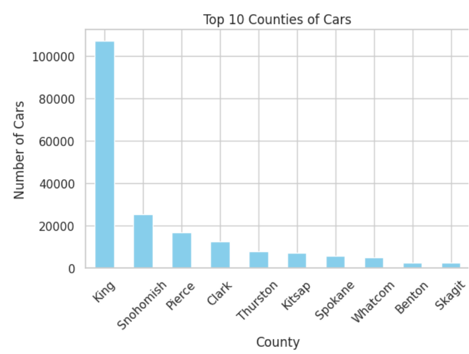
The table provides a sample of the data, showing the top counts by state. California (CA) has the highest count for BEVs with 84, while other states show lower counts for both BEVs and PHEVs.



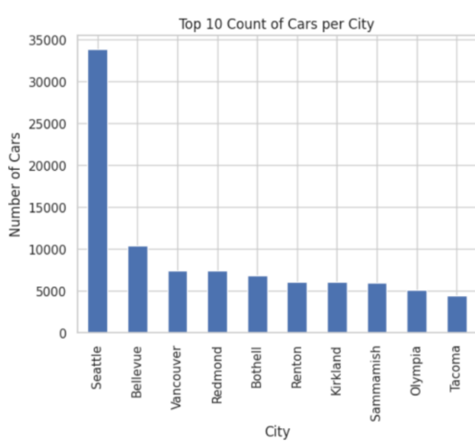
	State	Electric Vehicle Type	Count
9	CA	Battery Electric Vehicle (BEV)	84
2	AL	Battery Electric Vehicle (BEV)	7
5	AZ	Battery Electric Vehicle (BEV)	7
4	AR	Battery Electric Vehicle (BEV)	3
3	AL	Plug-in Hybrid Electric Vehicle (PHEV)	2
6	AZ	Plug-in Hybrid Electric Vehicle (PHEV)	2
0	AE	Plug-in Hybrid Electric Vehicle (PHEV)	1
1	AK	Battery Electric Vehicle (BEV)	1
7	BC	Battery Electric Vehicle (BEV)	1
8	BC	Plug-in Hybrid Electric Vehicle (PHEV)	1

10. Comparative Visualization

A) This bar chart shows the distribution of cars across the top 10 counties, with King County having the highest count by far.

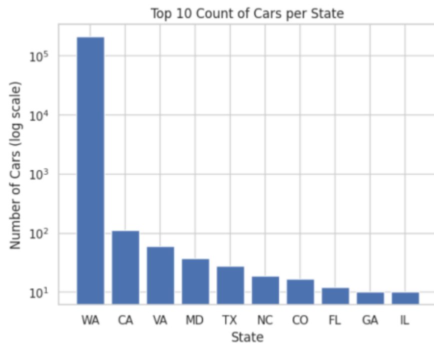


B) This bar chart shows the distribution of cars across the top 10 cities.



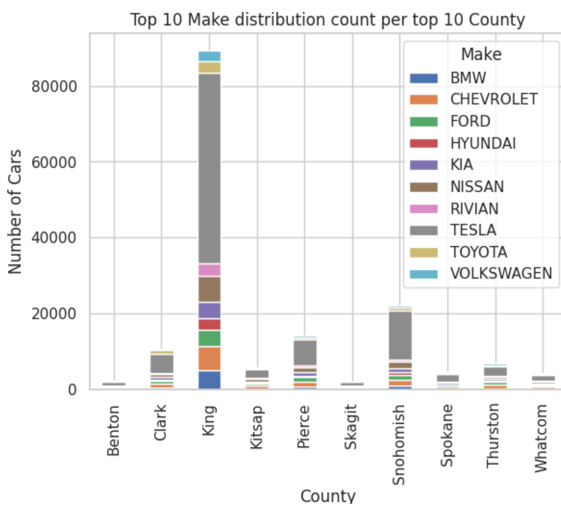
City	count
Seattle	33853
Bellevue	10414
Vancouver	7475
Redmond	7438
Bothell	6877
Renton	6131
Kirkland	6092
Sammamish	5995

C) This bar chart shows the distribution of cars across the top 10 cities.



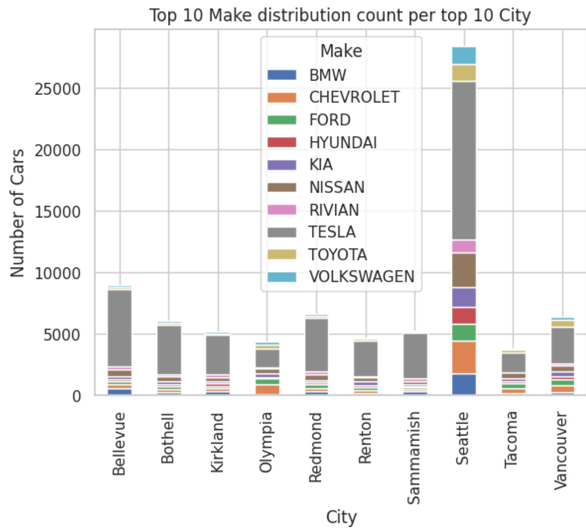
State	count
WA	209715
CA	113
VA	60
MD	37
TX	28
NC	19
CO	17
FL	12
GA	10
IL	10

D) In this chart, we see the distribution of car makes in the top 10 counties. King County has the highest number of cars across all makes, indicating a larger population or higher vehicle density. Popular makes include TESLA, NISSAN, and CHEVROLET, which have high counts across multiple counties, showing these brands' popularity.



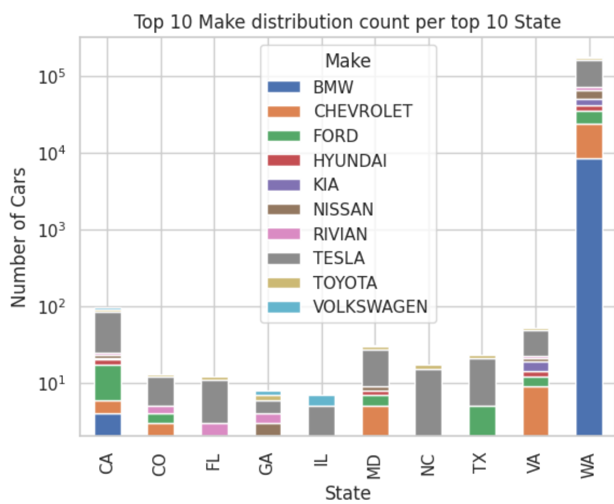
County	BMW	CHEVROLET	FORD	HYUNDAI	KIA	NISSAN	RIVIAN	TESLA	TOYOTA	VOLKSWAGEN
Benton	117	267	215	71	120	141	30	1068	172	48
Clark	467	959	706	394	604	851	245	5164	934	329
King	5101	6246	4203	3081	4287	7008	3333	50109	3144	2816
Kitsap	284	642	449	265	439	677	301	2239	303	233
Pierce	636	1298	1208	447	928	1152	554	6829	639	401
Skagit	57	288	174	109	130	176	121	798	125	82
Snohomish	844	1517	1350	734	1111	1798	544	12815	760	603
Spokane	241	481	419	137	316	315	99	2077	250	152
Thurston	182	1087	764	202	503	594	109	2467	406	351
Whatcom	130	527	295	154	316	655	117	1537	314	249

E) This chart displays the distribution of the top 10 car makes across the top 10 cities. Seattle stands out with the highest number of cars, particularly for popular makes like TESLA, TOYOTA, and NISSAN, suggesting a higher vehicle ownership or population density. Other cities like Bellevue and Redmond also show notable counts, but significantly fewer than Seattle.



City	BMW	CHEVROLET	FORD	HYUNDAI	KIA	NISSAN	RIVIAN	TESLA	TOYOTA	VOLKSWAGEN
Bellevue	577	315	244	182	250	563	264	6207	224	144
Bothell	274	245	236	159	247	408	178	4029	144	127
Kirkland	353	256	174	166	190	316	239	3205	134	124
Olympia	119	748	487	135	318	400	81	1559	280	272
Redmond	344	259	272	178	199	493	191	4358	146	158
Renton	180	266	253	155	331	299	99	2864	143	116
Sammamish	327	154	172	100	128	292	191	3731	91	103
Seattle	1810	2627	1423	1353	1606	2813	1068	12854	1418	1391
Tacoma	174	443	359	142	285	435	71	1589	203	128
Vancouver	251	615	425	248	382	550	117	2957	612	212

F) This chart and table show car counts for the top 10 brands across 10 states, using a logarithmic scale to make differences easier to see. Washington has notably high Tesla numbers compared to other states.



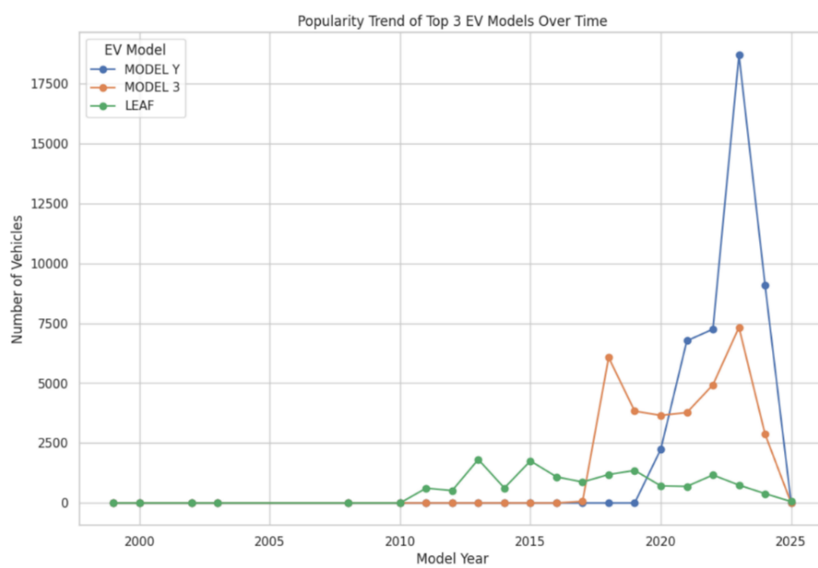
State	BMW	CHEVROLET	FORD	HYUNDAI	KIA	NISSAN	RIVIAN	TESLA	TOYOTA	VOLKSWAGEN
CA	4.0	2.0	11.0	3.0	1.0	2.0	2.0	60.0	6.0	5.0
CO	2.0	1.0	1.0	0.0	0.0	0.0	1.0	7.0	1.0	0.0
FL	0.0	1.0	0.0	0.0	1.0	0.0	1.0	8.0	1.0	0.0
GA	0.0	0.0	2.0	0.0	0.0	1.0	1.0	2.0	1.0	1.0
IL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.0	0.0	2.0
MD	2.0	3.0	2.0	1.0	0.0	1.0	0.0	18.0	3.0	1.0
NC	0.0	1.0	0.0	0.0	0.0	0.0	0.0	14.0	2.0	0.0
TX	0.0	2.0	3.0	0.0	0.0	0.0	0.0	16.0	2.0	0.0
VA	1.0	8.0	3.0	2.0	5.0	2.0	1.0	27.0	3.0	0.0
WA	8593.0	15392.0	11109.0	6060.0	9566.0	14711.0	5871.0	91158.0	8220.0	5771.0

11. Temporal Analysis

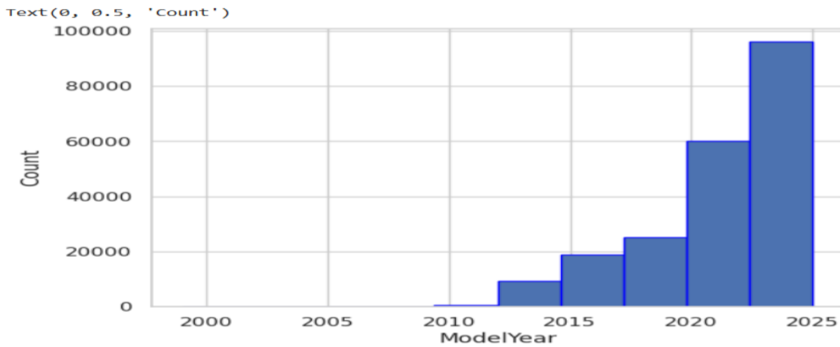
A) This table shows various car models and their production status across different years, represented by binary values ("0" indicates no production). The structure resembles a matrix, with models as columns and years as rows. Currently, only zeros are displayed, indicating no production for these models in listed years.

```
Model        330E  500  500E  530E  740E  745E  745LE  750E  918  A3  A7 E  \
Model Year
1999            0    0     0     0     0     0      0     0    0   0     0
2000            0    0     0     0     0     0      0     0    0   0     0
2002            0    0     0     0     0     0      0     0    0   0     0
2003            0    0     0     0     0     0      0     0    0   0     0
2008            0    0     0     0     0     0      0     0    0   0     0

Model        A8 E  ACCORD  AIR  ARIYA  AVIATOR  B-CLASS  BENTAYGA  BLAZER EV  \
Model Year
1999            0       0    0      0        0        0         0          0
2000            0       0    0      0        0        0         0          0
2002            0       0    0      0        0        0         0          0
2003            0       0    0      0        0        0         0          0
2008            0       0    0      0        0        0         0          0

Model        BOLT EUV  BOLT EV  BZ4X  C-CLASS  C-MAX  C40  CAYENNE  CITY  \
Model Year
1999                0        0     0        0      0    0        0     0
2000                0        0     0        0      0    0        0     0
2002                0        0     0        0      0    0        0     0
2003                0        0     0        0      0    0        0     0
2008                0        0     0        0      0    0        0     0

Model        CLARITY  CORSAIR  COUNTRYMAN  CR-V  CROSSTREK  CT6  CX-70  CX-90  \
Model Year
1999               0        0           0     0          0    0      0      0
2000               0        0           0     0          0    0      0      0
2002               0        0           0     0          0    0      0      0
2003               0        0           0     0          0    0      0      0
2008               0        0           0     0          0    0      0      0

Model        CYBERTRUCK  E-GOLF  E-TRON  E-TRON GT  E-TRON SPORTBACK  EDV  ELR  \
Model Year
1999                  0       0       0          0                 0    0    0
2000                  0       0       0          0                 0    0    0
2002                  0       0       0          0                 0    0    0
2003                  0       0       0          0                 0    0    0
2008                  0       0       0          0                 0    0    0

Model        EQ FORTWO  EQB-CLASS  EQE-CLASS SEDAN  EQE-CLASS SUV  \
Model Year
1999                 0          0                0              0
2000                 0          0                0              0
2002                 0          0                0              0
2003                 0          0                0              0
2008                 0          0                0              0

Model        EQS-CLASS SEDAN  EQS-CLASS SUV  EQUINOX EV  ESCAPE  ESPRINTER  \
```

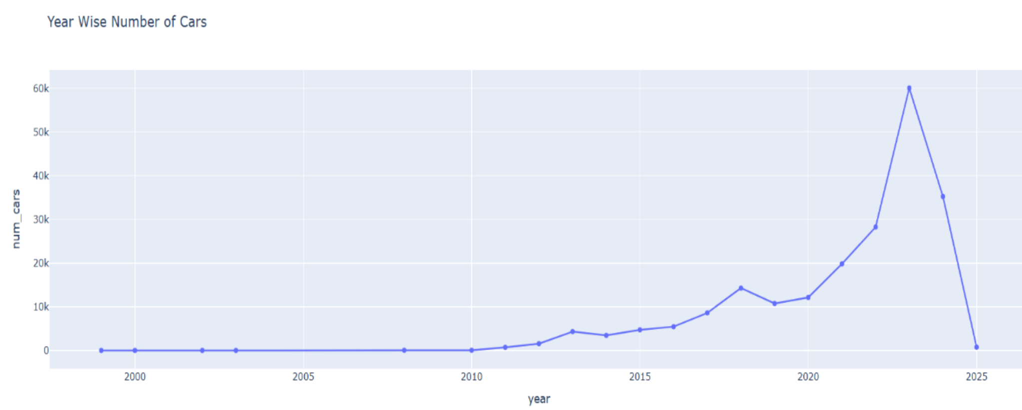
B) The below chart shows that Model Y's popularity surged around 2021, peaking higher than the Nissan Leaf and Model 3. The Leaf had steady growth from 2010, while Model 3 fluctuated, peaking around 2019. All models show a decline near 2025, possibly due to data limitations.



C) The data represented in the bar graph below shows an increasing trend in the count over the years from 2000 to 2025.



D) The bar graph below displays a steady increase in the number of cars from 2000 to 2025.



E) The bar graph depicts a gradual increase in car sales over the years from 2000 to 2025.