OVERVIEW :

Electric vehicles have come to light as a potential option to lessen carbon emissions and advance sustainability as people's awareness of the environmental impact of transportation grows. The adoption of electric vehicles is still constrained, though, due to a number of issues, such as their high cost, short range, and scarcity of charging stations. Understanding the current situation of the electric car market and identifying prospective areas for development are crucial for promoting the wide adoption of electric vehicles.

Agitate: In order to better understand the electric vehicle market, we analyse a dataset that contains data on electric vehicles in this study. We examine a number of facets of the market for electric cars, such as the distribution of electric vehicles by area, manufacturer, and model, as well as whether electric vehicles qualify for subsidies for clean alternative fuel vehicles. We also look at the distribution of electric cars by price range and the electric utility providers that supply the electricity for the electric vehicles.

Solution: By examining the electric car dataset, we can spot potential flaws and offer suggestions to organizations that support the use of electric vehicles. We discuss the prevalence of electric vehicles in cities and suggest that advocacy organizations concentrate their efforts on advancing the adoption of electric vehicles in metropolitan areas. We also advise promoting the use of renewable energy sources to power electric cars, working with electric utility companies to ensure that the infrastructure is in place to support the widespread adoption of electric vehicles, and pushing for the construction of more charging stations in public areas.

Benefits: This study provides valuable insights into the electric car market, including the distribution of electric cars by location, make, and model, as well as the eligibility of electric cars for clean alternative fuel vehicle incentives. The recommendations we make for electric car advocacy groups can help to promote the adoption of electric cars and support the transition to a more sustainable transportation system. By reducing carbon emissions and promoting sustainability, the widespread adoption of electric cars can have a significant positive impact on the environment and contribute to a more sustainable future.

DATA SOURCES:

- Global EV Outlook 2022

The Global EV Outlook is an annual publication that identifies and discusses recent developments in electric mobility across the globe. It is developed with the support of the members of the Electric Vehicles Initiative (EVI).

Combining historical analysis with projections to 2030, the report examines key areas of interest such as electric vehicle and charging infrastructure deployment, energy use, CO2 emissions, battery demand and related policy developments. The report includes policy recommendations that incorporate lessons learned from leading markets to inform policy makers and stakeholders with regard to policy frameworks and market systems for electric vehicle.

DATA PREPROCESSING:

Handling Missing Values by Dropping Rows

In our dataset, we have observed that only a small number of rows have missing values in certain columns. Given the overall size of the dataset, dropping these rows will likely have a minimal impact on the accuracy and representativeness of our analysis. Furthermore, dropping these rows with missing values is a simple and straightforward approach that allows us to work with a clean dataset without having to make assumptions or impute values that may introduce bias or errors into our analysis.

The following columns have missing values:

- Model (78 missing values)
- Electric Range (1 missing value)
- Base MSRP (1 missing value)
- Legislative District (148 missing values)
- DOL Vehicle ID (1 missing value)
- Vehicle Location (19 missing values)
- Electric Utility (227 missing values)
- 2020 Census Tract (1 missing value)

Considering the relatively small number of missing values in these columns, we can safely drop these rows without significantly affecting the overall dataset. This way, we can proceed with our analysis using a dataset free of missing values.

Feature Engineering: Extracting Latitude and Longitude

In our dataset, the Vehicle Location column contains both latitude and longitude coordinates as a string. To make these coordinates more accessible for further analysis, we will perform the following feature engineering steps:

- Convert the Vehicle Location column to string type.
- Create two new columns in the DataFrame, latitude and longitude.
- Extract latitude and longitude values from the Vehicle Location column and store them in the respective new columns.
- To achieve this, we define a function called extract_coordinates(), which takes two arguments: the input string (containing the coordinates) and the index (0 for latitude and

1 for longitude). The function uses regular expressions to find and return the floating-point numbers representing the coordinates. We then use the apply() function to apply this custom function to each element of the Vehicle Location column and populate the new latitude and longitude columns.

- Finally, we drop any rows with missing values in the latitude and longitude columns to ensure a clean dataset for further analysis.

Feature Engineering: Creating a 'Location' Column
In our dataset, we have three columns representing different levels of geographical divisions: County, City, and State. To create a more informative and combined representation of these geographical attributes, we will create a new column called Location.

- The Location column will be a concatenation of the County, City, and State columns, with each value separated by a comma. For example, if a row has the values "Yakima" for County, "Yakima" for City, and "WA" for State, the corresponding Location value will be "Yakima, Yakima, WA".

Feature Engineering: Creating a Price_Range_Category Column
In our dataset, we have observed an unusual distribution of values in the Base MSRP column, with a large number of vehicles having a value of 0. This could potentially indicate missing or unknown values in the dataset. To account for this uncertainty and still make use of the available data, we have decided to create a new column called Price_Range_Category based on the Base MSRP values.
We have defined four categories for the Price_Range_Category column:

- "Unknown": If the Base MSRP value is 0, we assign this category as it might indicate missing or unknown values.
- "Low": If the Base MSRP value is less than 40,000, we assign this category.
- "Medium": If the Base MSRP value is between 40,000 and 60,000, we assign this category.
- "High": If the Base MSRP value is greater than 60,000, we assign this category.

By creating this new column, we can better understand the distribution of electric vehicle prices in our dataset and account for the potential uncertainty introduced by the large number of 0 values in the Base MSRP column.

Feature Engineering: Creating an 'Electric_Range_Category' Column
In our dataset, we have observed an unusual distribution of values in the 'Electric Range' column, with a large number of vehicles having a value of 0. This could potentially indicate missing or unknown values in the dataset. To account for this uncertainty and still make use of the available data, we have decided to create a new column called 'Electric_Range_Category' based on the 'Electric Range' values.

We have defined four categories for the 'Electric_Range_Category' column:

- "Unknown": If the 'Electric Range' value is 0, we assign this category as it might indicate missing or unknown values.
- "Short": If the 'Electric Range' value is less than 150, we assign this category.
- "Medium": If the 'Electric Range' value is between 150 and 300, we assign this category.
- "Long": If the 'Electric Range' value is greater than 300, we assign this category.

By creating this new column, we can better understand the distribution of electric vehicle ranges in our dataset and account for the potential uncertainty introduced by the large number of 0 values in the 'Electric Range' column.

EXPLORATORY DATA ANALYSIS

The goal of EDA is to gain insights and understanding of the dataset, identify patterns, relationships, and anomalies. Through EDA, we can make informed decisions on how to preprocess and model the data, as well as generate hypotheses for further analysis.

For our electric car dataset, we will focus on the following columns:

- 'Location'
- 'Model Year'
- 'Make'
- 'Model'
- 'Electric Vehicle Type'
- 'Clean Alternative Fuel Vehicle (CAFV) Eligibility'
- 'Electric Range Category'
- 'Price Range Category'
- 'Electric Utility'

We will begin by examining the distribution of data points across various columns, such as the number of electric cars per city or county, the distribution of makes and models, and the distribution of electric vehicle types. We will also visualize the geographical distribution of electric cars using the 'County', 'City', and 'State' columns.

Next, we will explore the relationships between variables, such as the relationship between the electric range and the base MSRP of electric vehicles. We will also analyze the distribution of Clean Alternative Fuel Vehicle (CAFV) eligibility across different makes, models, and electric vehicle types.

Furthermore, we will investigate the trends in the adoption of electric cars over time, focusing on the 'Model Year' column. We will identify patterns and significant changes in the popularity of various makes and models of electric vehicles, as well as any trends related to electric vehicle types and CAFV eligibility.

Finally, we will examine any potential outliers or anomalies in the dataset that may warrant further investigation. Throughout the EDA process, we will visualize our findings using various plotting libraries such as Seaborn and Plotly, to help communicate the insights effectively.

By the end of the EDA, we will have a comprehensive understanding of the electric car dataset, enabling us to make informed decisions on how to proceed with preprocessing, feature engineering, and modeling, as well as guiding our hypotheses for further analysis.

Electric Car Dataset Analysis: Location

The 'Location' column in the Electric Car dataset provides information about the location of the electric cars. The following are the value counts of the 'Location' column:

Location	Count
King, Seattle, WA	21942
King, Bellevue, WA	6476
King, Redmond, WA	4641
Clark, Vancouver, WA	4462
King, Kirkland, WA	3920
King, Sammamish, WA	3650
King, Renton, WA	3112

Snohomish, Bothell, WA	3014
Thurston, Olympia, WA	3014
Pierce, Tacoma, WA	2651

Conclusion

From the above value counts, we can make the following inferences and gain insights:

- The top 10 electric cars in the dataset are primarily located in the state of Washington (WA), with no cars from other states.
- The highest number of electric cars are located in King county, with Seattle having the highest count of 21942 cars. This indicates that the majority of electric cars are concentrated in urban areas, particularly in large cities like Seattle and Bellevue.
- The next most common location is Clark county, with Vancouver having a count of 4462 cars.

Based on the analysis, the following recommendations can be made for electric car advocacies:

1. Focus on urban areas: Since the majority of electric cars are located in urban areas, electric car advocates should focus on promoting the use of electric cars in cities, particularly in large cities like Seattle and Bellevue.
2. Expand to other countries: While King and Clark counties have the highest number of electric cars, there is still potential to expand the use of electric cars in other counties in the state of Washington. Advocates should consider reaching out to these counties and promoting the benefits of electric cars.
3. Target specific cities: Within counties with a high number of electric cars, there may be specific cities or neighborhoods with lower adoption rates. Advocacies should target these areas and promote the use of electric cars through targeted marketing and outreach efforts.

Electric Car Dataset Analysis: CAFV Eligibility

The 'CAFV Eligibility' column in the Electric Car dataset provides information about whether a given electric car is eligible for Clean Alternative Fuel Vehicle (CAFV) incentives based on its battery range. The following are the value counts of the 'CAFV Eligibility' column:

CAFV Eligibility	Count
Clean Alternative Fuel Vehicle Eligible	59092
Eligibility unknown as battery range has not been researched	49346
Not eligible due to low battery range	15600

Electric vehicles (EVs) are gaining traction as a sustainable and environmentally friendly alternative to traditional internal combustion engine vehicles. As the transportation sector is a significant contributor to greenhouse gas emissions, the adoption of electric vehicles can play a critical role in reducing the overall carbon footprint. Analyzing the data available on electric vehicles can provide valuable insights and help inform future strategies to promote their adoption.

Based on the value counts provided for the electric car dataset, we can make a few inferences and provide the following recommendations:

1. **Promote clean alternative fuel vehicles:** A significant portion (about 47.5%) of the vehicles in the dataset are classified as "Clean Alternative Fuel Vehicle Eligible." This indicates that there is a growing interest in and adoption of electric vehicles. Efforts should be made to encourage the use of clean alternative fuel vehicles by offering incentives, improving infrastructure, and educating consumers about the environmental and financial benefits of owning such vehicles.
2. **Research and data collection for battery range:** There is a significant number of vehicles (about 39.8%) with "Eligibility unknown as battery range has not been researched." To improve the understanding of the electric vehicle market, it is essential to collect and analyze data on battery range for all electric vehicles. This will help in making better recommendations for consumers and also help in the development of policies and regulations surrounding electric vehicles.
3. **Improve battery technology:** About 12.7% of the vehicles are "Not eligible due to low battery range." This indicates that there is still room for improvement in battery technology. Investing in research and development to enhance battery performance will not only help increase the range of electric vehicles but also improve their overall efficiency and attractiveness to potential buyers.
4. **Expand charging infrastructure:** A significant barrier to the adoption of electric vehicles is the lack of charging infrastructure. Expanding the availability of charging stations, especially in urban and high-traffic areas, will make electric vehicles more accessible and convenient for users.
5. **Increase public awareness:** It is crucial to create awareness about the benefits of electric vehicles and clean alternative fuel options. Public campaigns, educational programs, and partnerships with local communities can help in promoting the adoption of electric vehicles and in reducing the overall carbon footprint.

Electric Car Dataset Analysis: Make

The 'Make' column in the Electric Car dataset provides information about the **top 10 makers** of electric cars. The following are the value counts of the 'Make' column:

Make	Count
TESLA	56906
NISSAN	12912

CHEVROLET	10797
FORD	6635
BMW	5556
KIA	4831
TOYOTA	4631
VOLKSWAGEN	3356
AUDI	2472
VOLVO	2324

Inference

From the above value counts, we can make the following inferences and gain insights:

- The majority of electric cars in the dataset are from the Tesla make, with 56,906 cars falling into this category.
- The next most common makes are Nissan and Chevrolet, with 12,912 and 10,797 cars respectively.
- The rest of the makes have relatively smaller counts, with Volvo having the lowest count of 2,324 cars.

Based on the analysis, the following recommendations can be made for electric car advocacies:

1. **Promote Tesla electric cars:** Since the majority of electric cars in the dataset are from the Tesla make, electric car advocates should focus on promoting Tesla electric cars to encourage more people to purchase them.
2. **Encourage diversity in make:** While Tesla is the dominant make in the dataset, there are other makes with a significant number of cars. Electric car advocates should encourage diversity in make and promote the benefits of each make to encourage more people to purchase electric cars.
3. **Advocate for increased availability of less common makes:** Electric car advocates should advocate for increased availability of less common makes, such as Volvo and Audi, to increase consumer choice and promote diversity in the electric car market.

We can also make the following inferences and gain insights:

- The most common model of electric car in the dataset is the Tesla Model 3, with 24,300 cars falling into this category.
- The next most common model is the Tesla Model Y, with 20,609 cars.
- The Nissan Leaf is the third most common model, with 12,890 cars.
- The remaining models have relatively smaller counts, with the Niro having the lowest count of 2,286 cars.

Based on the analysis, the following recommendations can be made for electric car advocacies:

1. **Promote popular models:** Since the most common models in the dataset are the Tesla Model 3 and Model Y, electric car advocates should focus on promoting these models to encourage more people to purchase them.
2. **Encourage diversity in model:** While the most common models are dominated by Tesla and Nissan, there are other models with a significant number of cars. Electric car advocates should encourage diversity in model and promote the benefits of each model to encourage more people to purchase electric cars.
3. **Advocate for increased availability of less common models:** Electric car advocates should advocate for increased availability of less common models, such as the VW ID.4 and Kia Niro, to increase consumer choice and promote diversity in the electric car market.

- Battery electric vehicles (BEVs) are the most common type of electric vehicle in the dataset, with 95,753 cars falling into this category.
- The next most common type is plug-in hybrid electric vehicles (PHEVs), with 28,285 cars.
- There are no other types of electric vehicles in the dataset, indicating that the majority of electric cars are either BEVs or PHEVs.

Recommendations for Advocacies:

Based on the analysis, the following recommendations can be made for electric car advocacies:

1. **Promote the benefits of BEVs and PHEVs:** Since BEVs and PHEVs are the most common types of electric vehicles in the dataset, electric car advocacies should focus on promoting the benefits of these types of vehicles to encourage more people to purchase them.
2. **Educate on the differences between BEVs and PHEVs:** While both types of electric vehicles are common, they have different characteristics and are suited for different use cases. Electric car advocacies should educate consumers on the differences between BEVs and PHEVs to help them make informed purchase decisions.
3. **Advocate for increased availability of different types of electric vehicles:** While BEVs and PHEVs are common, there are other types of electric vehicles, such as fuel cell electric vehicles (FCEVs), that are not represented in the dataset. Electric car advocacies should advocate for increased availability of different types of electric vehicles to increase consumer choice and promote diversity in the electric car market.

- The majority of electric cars in the dataset have an unknown price range category, with 120,656 cars falling into this category. This is likely due to a large number of vehicles having a value of 0 in the 'Base MSRP' column.
- Only a small percentage of electric cars have a defined price range category, with 1,653 cars falling into the 'High' category, 971 cars falling into the 'Low' category, and 758 cars falling into the 'Medium' category.
- The price range categories are defined based on the 'Base MSRP' values, with a value of 0 indicating an unknown price range.

Based on the analysis, the following recommendations can be made for electric car advocacies:

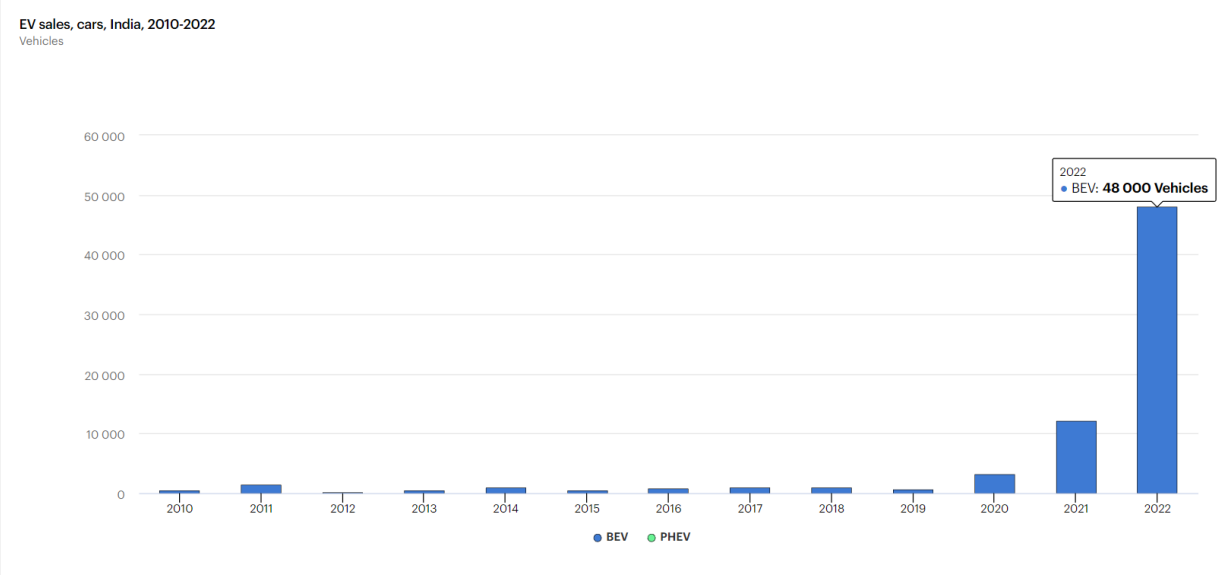
1. **Encourage more transparency in pricing:** Since the majority of electric cars in the dataset have an unknown price range category, electric car advocacies should encourage more transparency in pricing to help consumers make informed purchase decisions.
2. **Focus on affordable options:** Since only a small percentage of electric cars have a defined price range category and fall into the 'Low' and 'Medium' categories, electric car advocacies should focus on promoting affordable options to make electric cars more accessible to a wider range of consumers.
3. **Advocate for more incentives:** Given that electric cars tend to be more expensive than gasoline-powered cars, electric car advocacies should advocate for more incentives and subsidies to make electric cars more affordable and promote their adoption.

- Puget Sound Energy Inc. is the most common electric utility company in the dataset, with 44,926 electric cars being powered by this company.
- The next most common electric utility company is PUGET SOUND ENERGY INC (24,626 cars), followed by CITY OF SEATTLE (23,254 cars).
- There are several other electric utility companies in the dataset, but they have a much lower number of electric cars powered by them.

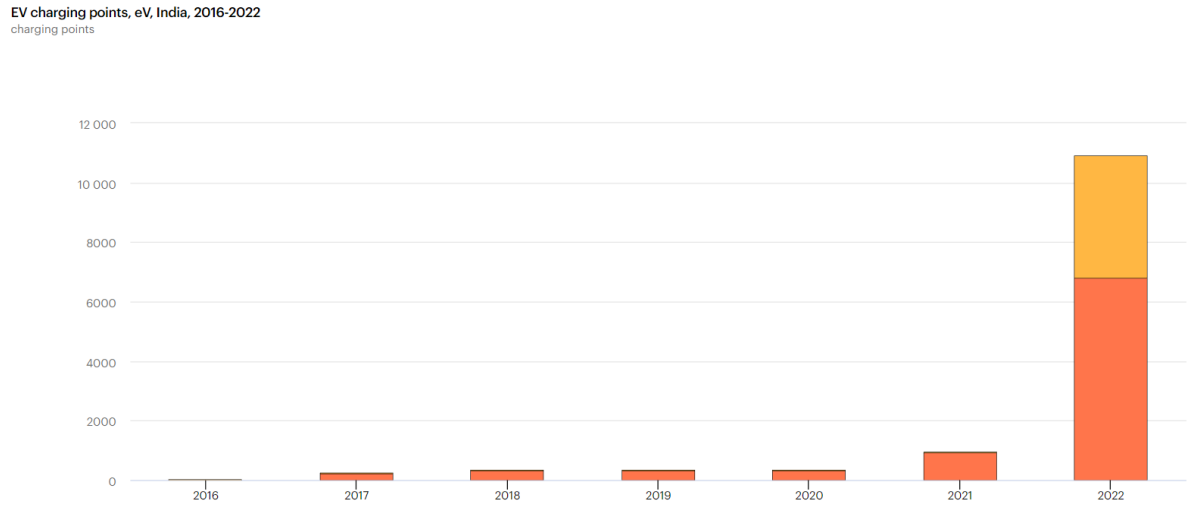
Based on the analysis, the following recommendations can be made for electric car advocacies:

1. **Collaborate with electric utility companies:** Since electric cars require a reliable and efficient electric grid to function, electric car advocacies should collaborate with electric utility companies to ensure that the necessary infrastructure is in place to support the widespread adoption of electric cars.
2. **Promote the use of renewable energy sources:** Many electric utility companies are transitioning to renewable energy sources, such as wind and solar power. Electric car advocacies should promote the use of these renewable energy sources to power electric cars, as they are more sustainable and environmentally friendly.
3. **Advocate for more charging stations:** To support the widespread adoption of electric cars, there need to be more charging stations available for electric car owners to use. Electric car advocacies should advocate for more charging stations to be built in public places, such as parking lots and rest areas, to make it more convenient for electric car owners to charge their cars.

FOR ELECTRIC CARS:
EV SALES:



EV CHARGING POINTS :



EV STOCK SHARE :