**Market Segmentation**

**Based on Electric Vehicle**

**Data set : public-charging-stations-Hawaii**

**Name : Shashi Kumar Reddy Chintalapalli**

**OVERVIEW**

The electric vehicle (EV) industry is a rapidly growing market, with increasing demand for energy-efficient vehicles driven by environmental concerns and the need to reduce carbon emissions. As the EV market continues to expand, machine learning is being utilized to improve the performance and efficiency of these vehicles.

The EV market can be segmented into Battery Electric Vehicles (BEVs), Plug-in Hybrid Electric Vehicles (PHEVs), and Hybrid Electric Vehicles (HEVs). Machine learning algorithms can be applied to all of these types of vehicles to optimize their performance, battery life, and overall efficiency.

One of the key applications of machine learning in the EV industry is predictive maintenance. Machine learning algorithms can analyze data generated by the vehicle to predict when maintenance will be required, allowing for proactive maintenance and reducing the risk of breakdowns or failures. This can lead to cost savings for vehicle owners and improved reliability for the vehicles.

Battery management is another important application of machine learning in the EV industry. The performance of an EV's battery can be optimized using machine learning algorithms to analyze data on the battery's charge level, temperature, and other factors. This can help to extend the battery life and improve the vehicle's overall energy efficiency.

Machine learning can also be used for route optimization and energy management in EVs. By analyzing data on factors such as traffic patterns, weather conditions, and driving behavior, machine learning algorithms can optimize the vehicle's route to minimize energy consumption and reduce emissions.

The adoption of machine learning in the EV industry is being driven by several factors, including the increasing demand for energy-efficient vehicles, advancements in battery technology, and the availability of large amounts of data generated by EVs. Machine learning algorithms can analyze this data to optimize the performance of the vehicle and its components, resulting in better energy efficiency and longer battery life.

The global electric vehicle market is expected to continue growing at a rapid pace, with a projected market size of $802.81 billion by 2027, growing at a CAGR of 22.6% from 2020 to 2027. The integration of machine learning in the EV industry is expected to further drive this growth.

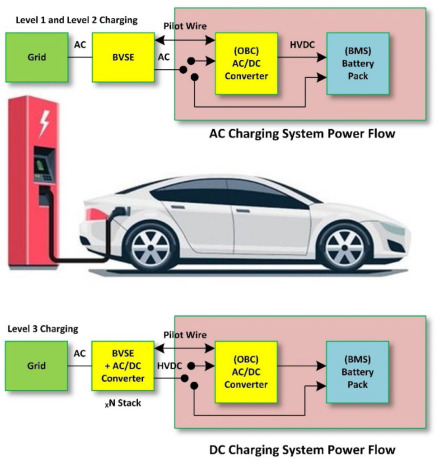
However, there are also challenges associated with the adoption of machine learning in the EV industry. The lack of standardization of data and models, the high cost of implementing machine learning solutions, and the need for skilled professionals are some of the challenges that need to be addressed.

Despite these challenges, the integration of machine learning in the EV industry has the potential to transform the way vehicles are designed, built, and operated. As the technology continues to evolve, it has the potential to make EVs more efficient, reliable, and cost-effective, driving further adoption in the global market.

Overall, machine learning is playing an increasingly important role in the EV industry, helping to optimize the performance and efficiency of these vehicles and contributing to the growth of the global electric vehicle market.

**Problem statement**

The problem statement revolves around the availability and accessibility of public charging stations for electric vehicles in the state of Hawaii, specifically on the island of Oahu. The dataset provided includes information about the location, number of chargers and ports, charger level, charge fee, and manufacturer of five charging stations on the island. However, it is evident from the limited number of charging stations that there may be a shortage of public charging infrastructure for electric vehicles in Hawaii. This shortage could act as a significant barrier to the adoption and usage of electric vehicles in the state.

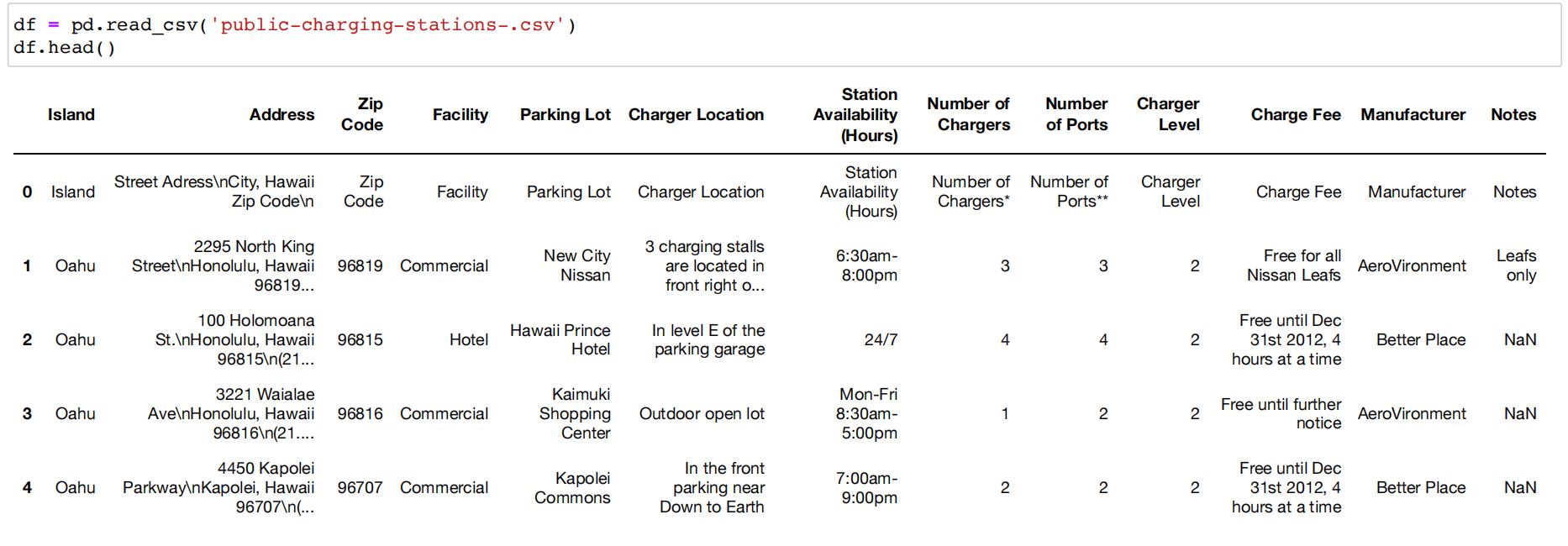


Therefore, the problem statement suggests that there is a need to address the availability and accessibility of public charging stations for electric vehicles in Hawaii to support the transition towards sustainable transportation. The state may need to develop policies and initiatives to promote the installation of more public charging stations, especially in areas with high EV traffic. Additionally, it may be necessary to ensure that the existing charging infrastructure is maintained and upgraded regularly to meet the growing demand for EV charging services. Solving the issue of the availability and accessibility of public charging stations in Hawaii can encourage the adoption and usage of electric vehicles, thereby contributing to the state's efforts to reduce greenhouse gas emissions and promote sustainable transportation.

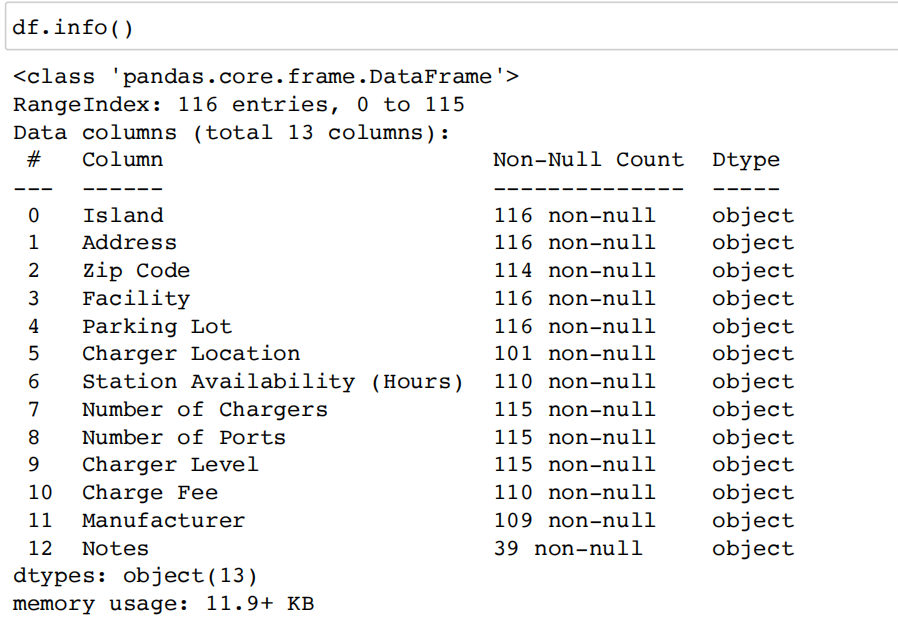


**Data Collection**

The website data.hawaii.gov provides a dataset that contains information on public charging stations for electric vehicles on the island of Oahu in Hawaii. The dataset includes details such as the location, zip code, facility type, parking lot, charger location, station availability, number of chargers and ports, charger level, charge fee, and manufacturer of each charging station.



This dataset can be a valuable resource for researchers, policymakers, and industry professionals who are interested in understanding the availability and accessibility of public charging stations for electric vehicles in Hawaii. The data can be used to identify areas with a high demand for charging infrastructure, inform decisions on the installation of new charging stations, and evaluate the effectiveness of existing policies and incentives for electric vehicle adoption.

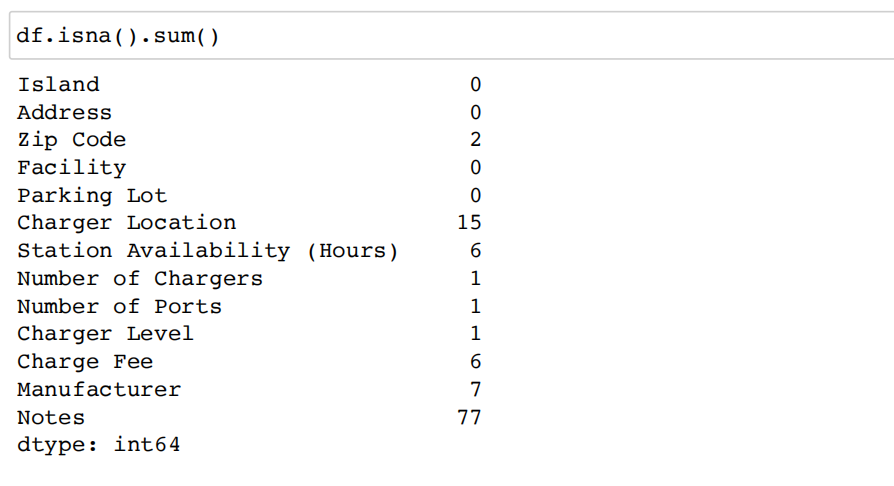


To utilize this dataset effectively, users may need to clean and preprocess the data to remove any inconsistencies or errors. Additionally, the dataset provides information on only a limited number of charging stations on Oahu, and users may need to gather additional data from other sources to obtain a more comprehensive understanding of the electric vehicle market in Hawaii.

Overall, the public charging stations dataset on data.hawaii.gov provides a valuable starting point for analyzing the electric vehicle market in Hawaii and can be used to inform decisions related to the development of electric vehicle infrastructure and policies in the state.

**Data Pre-Processing**

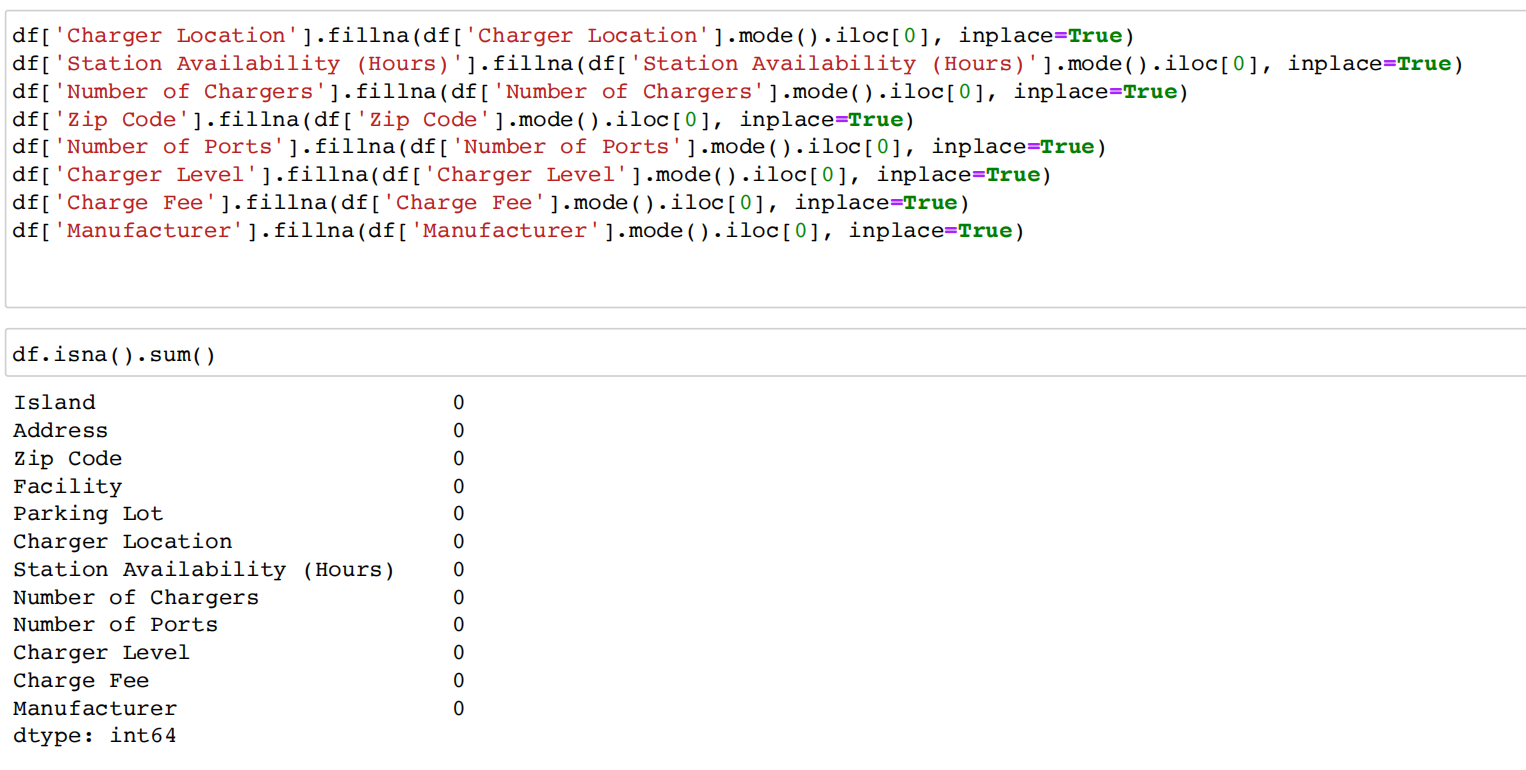
Data pre-processing is a crucial step in data analysis, which involves cleaning, transforming, and organizing raw data to make it suitable for further analysis. In this documentation, we will discuss the pre-processing steps and libraries used for the public charging stations dataset in Hawaii.



Firstly, we import the necessary libraries such as numpy, pandas, seaborn, and matplotlib. These libraries are widely used for data analysis and visualization in Python. We then read the CSV file containing the public charging stations data using the pandas library, and display the first few rows of the data using the head() method.

Next, we check for missing values in the data using the isna() method, which returns a boolean dataframe indicating the presence of missing values in each column. We can then sum up the number of missing values in each column using the sum() method, which gives us an overview of the amount of missing data in the dataset.

To handle missing values, we first drop the 'Notes' column using the drop() method since it contains a large number of missing values. We then fill the missing values in other columns with their respective mode (most frequent value) using the fillna() method. This helps us to replace the missing values with a representative value without introducing any bias in the data.

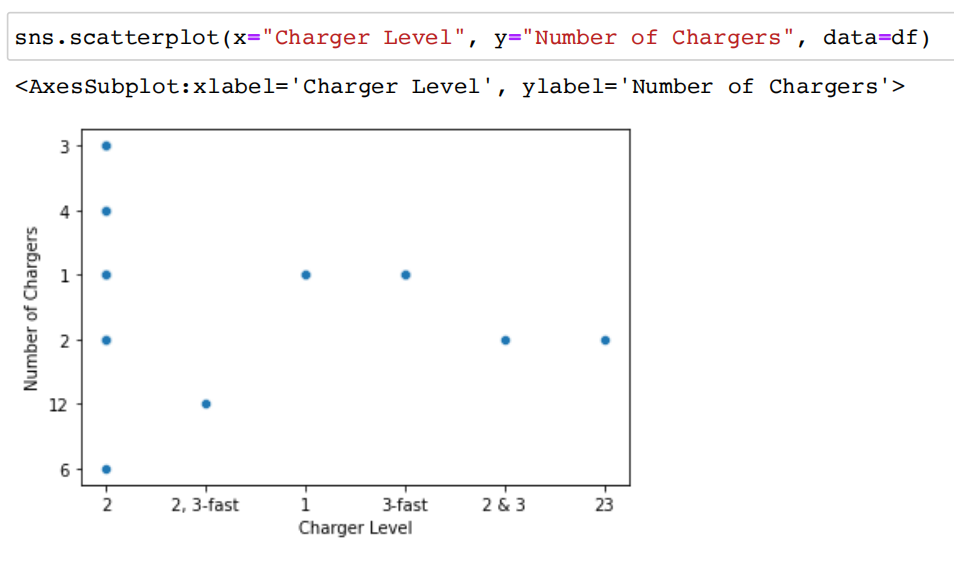


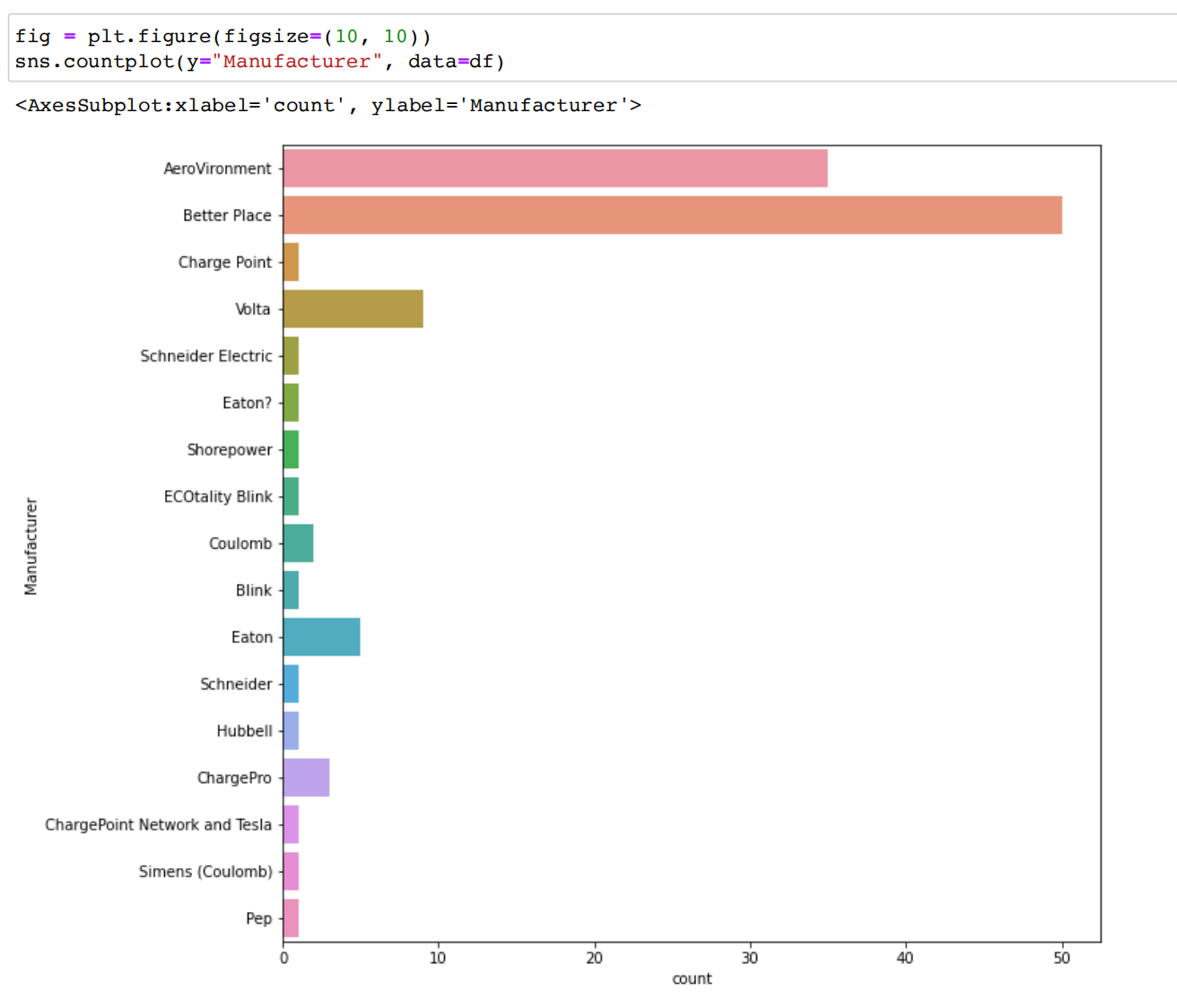
Finally, we can use visualization libraries such as seaborn and matplotlib to explore the data and gain insights. For example, we can plot a bar chart showing the number of charging stations by facility type using the countplot() method from seaborn.

In conclusion, data pre-processing is an essential step in data analysis, and libraries such as pandas, numpy, seaborn, and matplotlib provide useful tools to clean and transform the data. By following the steps outlined above, we can handle missing values and prepare the public charging stations data for further analysis and visualization.

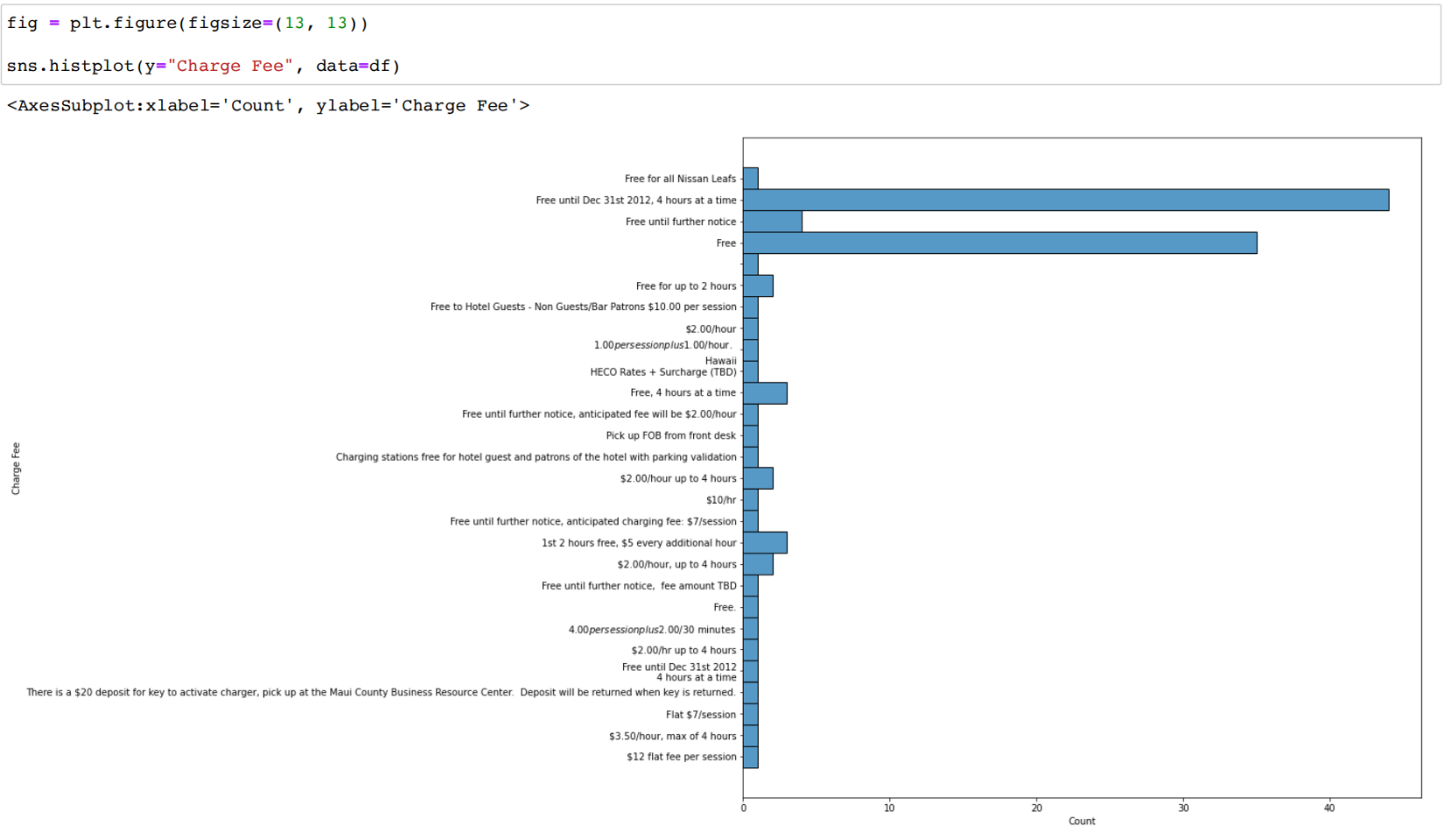
**Data visualization**

Data visualization is an essential step in the data analysis process as it helps to gain insights into the dataset and understand the patterns and trends. In the context of the provided dataset of public charging stations for electric vehicles in Hawaii, various visualizations can be created to understand the distribution of the charging stations, the types of chargers, the availability of stations, and other related features.

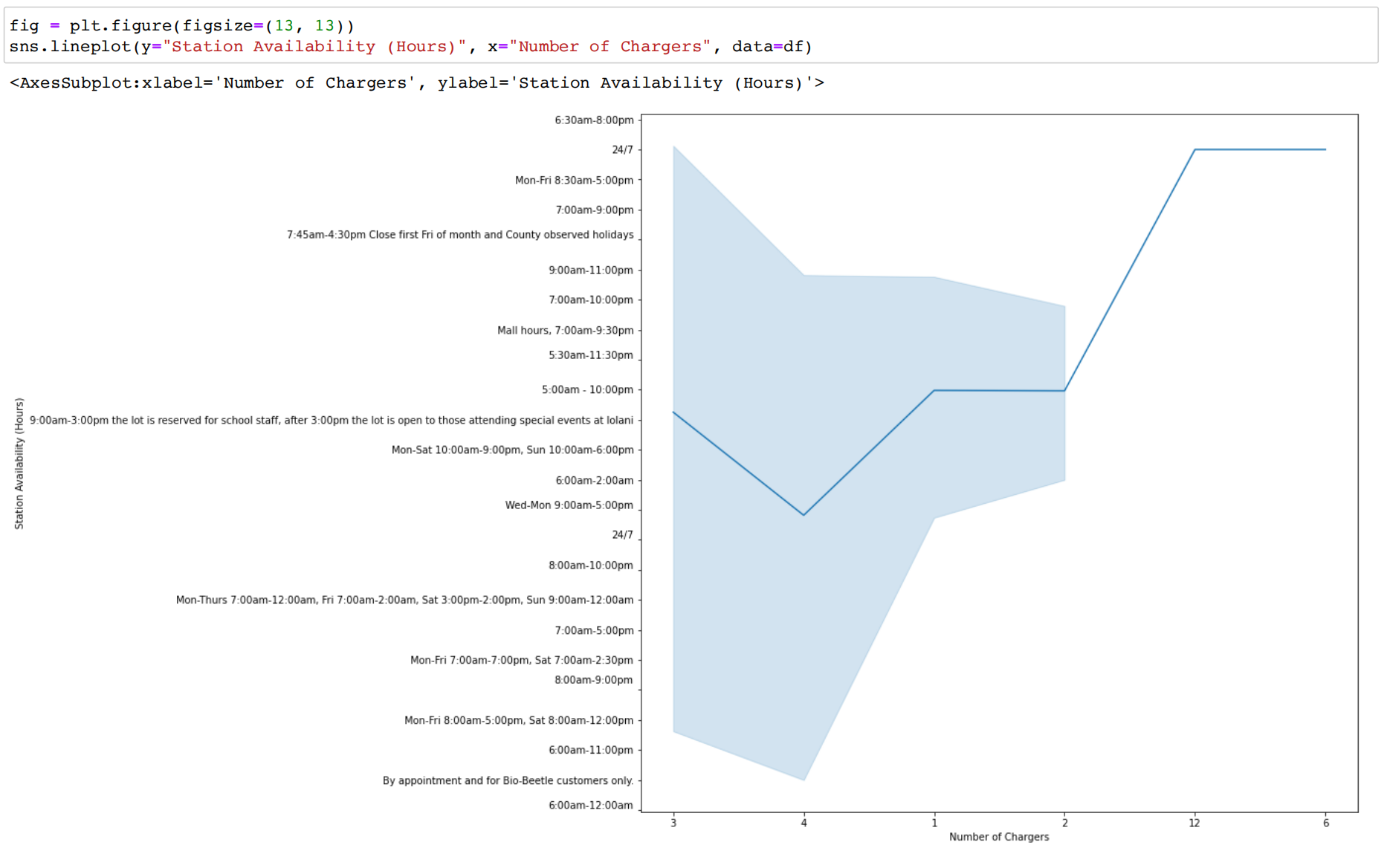




The seaborn and matplotlib libraries can be used to create different types of visualizations, such as scatter plots, count plots, histograms, and line plots. For instance, the scatter plot created using the sns.scatterplot() function displays the relationship between the charger level and the number of chargers available in the charging stations. The count plot created using the sns.countplot() function displays the distribution of charging station manufacturers in the dataset. The histogram created using the sns.histplot() function displays the distribution of charge fees for the charging stations. Lastly, the line plot created using the sns.lineplot() function shows the relationship between the station availability and the number of chargers available.



In addition to these libraries, Plotly can also be used to create interactive visualizations. The code snippet provided shows an example of creating an interactive bar chart using the go.Bar() function from the Plotly library. The bar chart displays the number of charging stations available on each island in Hawaii. The plot can be interactively explored using the Plotly interface and offers a more engaging way of visualizing the data.





Overall, data visualization is a crucial step in data analysis as it provides valuable insights into the data and helps to communicate the findings effectively. The seaborn, matplotlib, and Plotly libraries are powerful tools that can be used to create various types of visualizations and gain a better understanding of the data.