**Market Segment Analysis of Online Vehicle Booking**

*by*

*Khushi Pathak*

**Dataset used:**

<https://drive.google.com/file/d/1SWSL4qCW5ZG3r3alXOGgOTre9SXiS4JO/view>

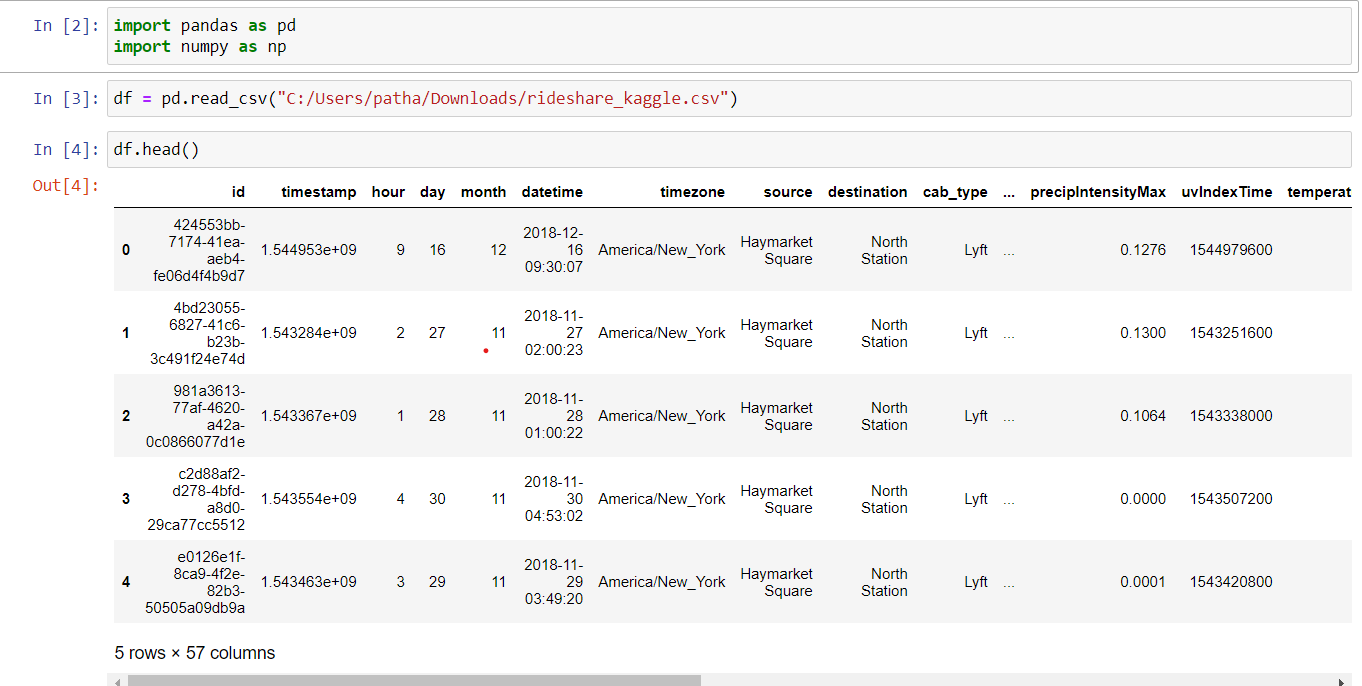
**Project Link:**

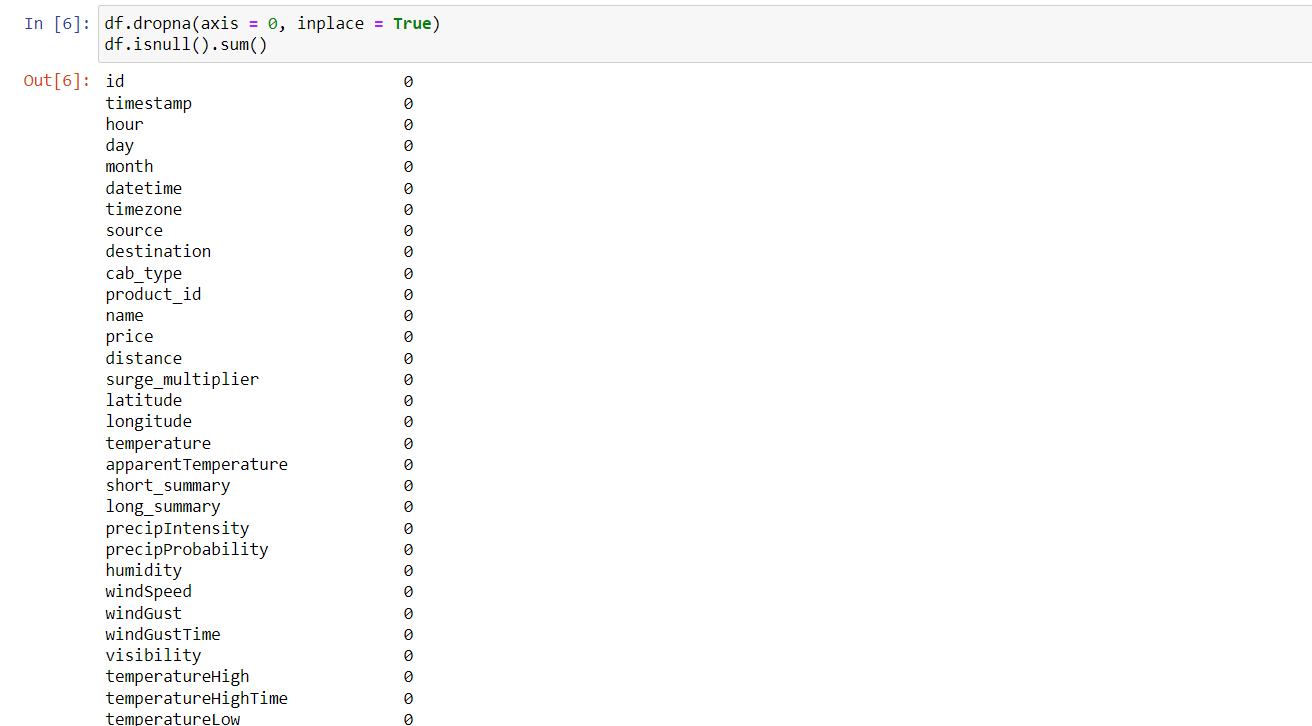
[**https://github.com/khushipathak3502/Feynn\_labs/tree/main**](https://github.com/khushipathak3502/Feynn_labs/tree/main)

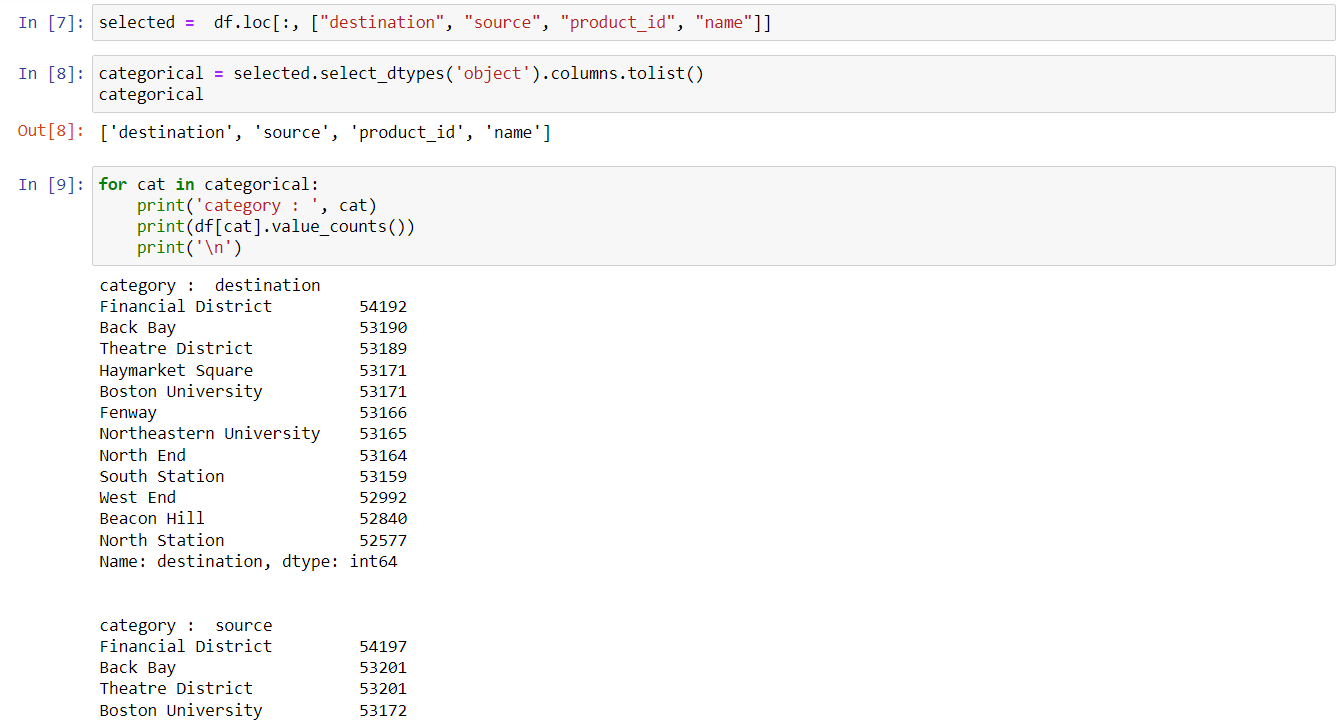
1. **Data Pre-Processing:**

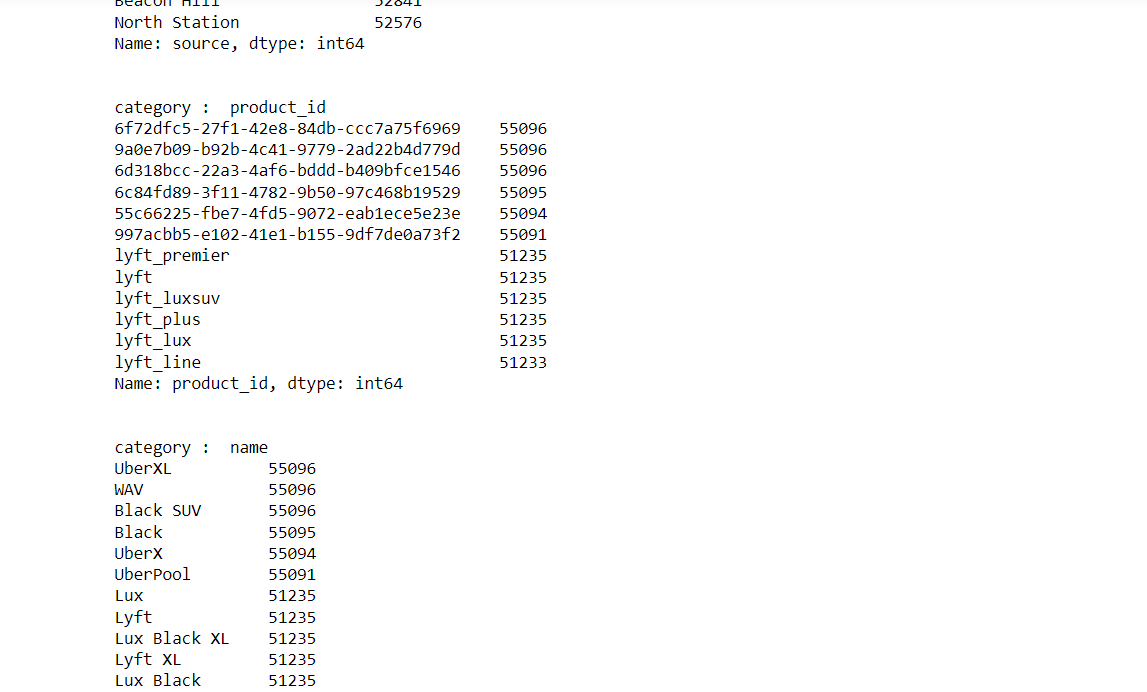
Data preprocessing is a crucial step in preparing raw data to make it suitable for machine learning models. The process involves cleaning the data, removing any errors or inconsistencies, and transforming it into a format that can be easily analyzed. It is essential to preprocess the data before performing any segmentation analysis.

To preprocess data, the first step is to import the raw data in a suitable format and create a data frame for further analysis. The next step is to identify any null values in the dataset and remove them to avoid any data inconsistencies.

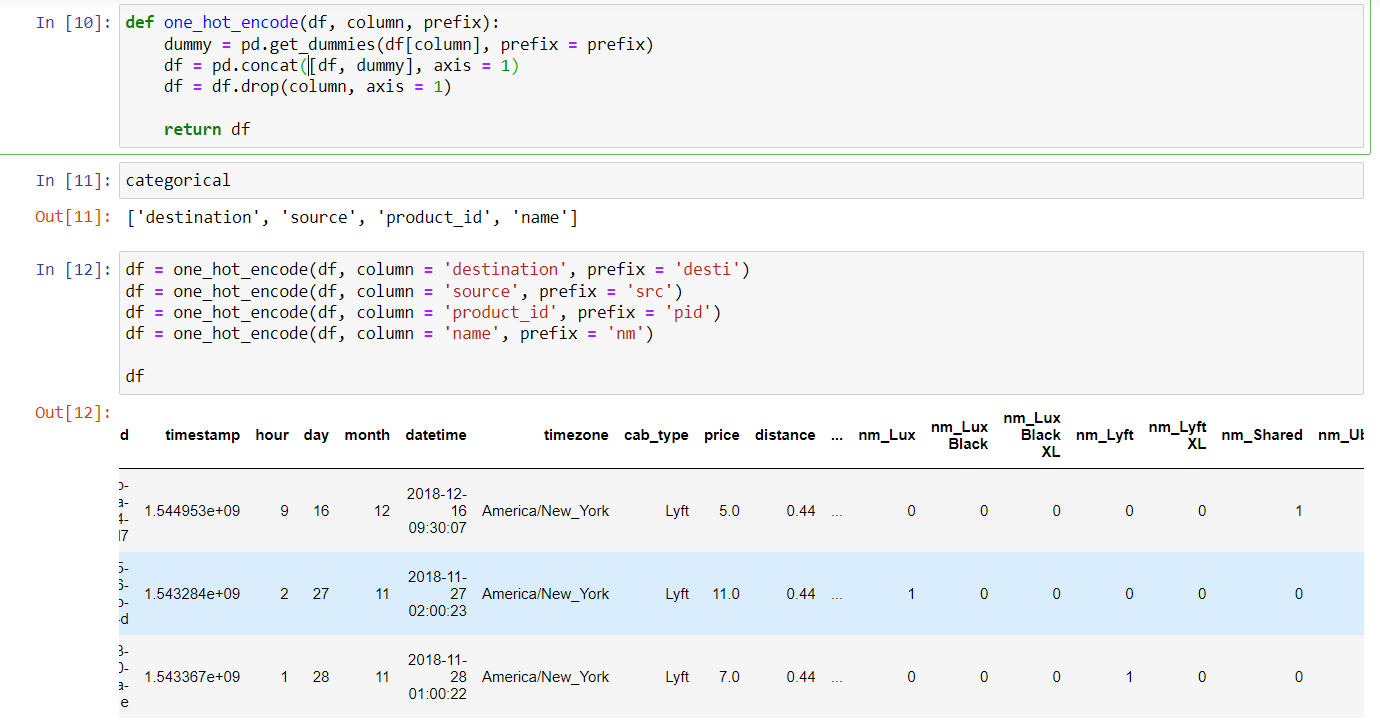






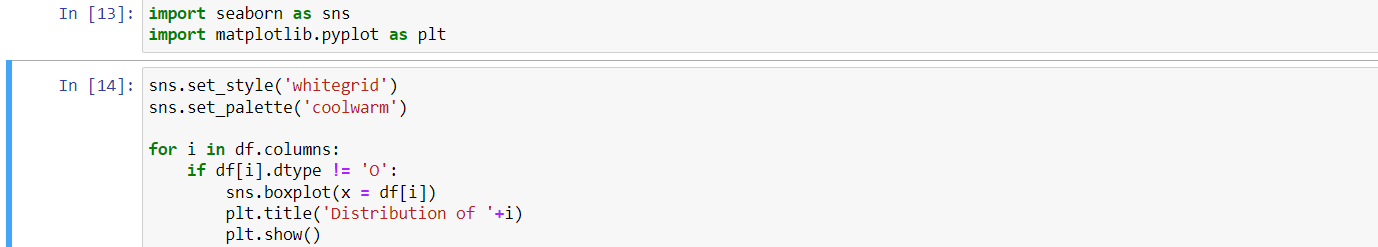


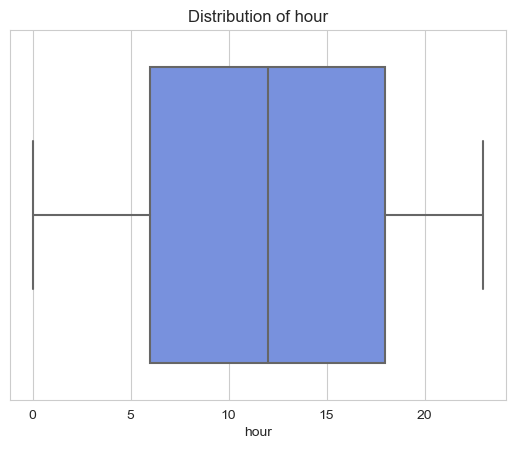
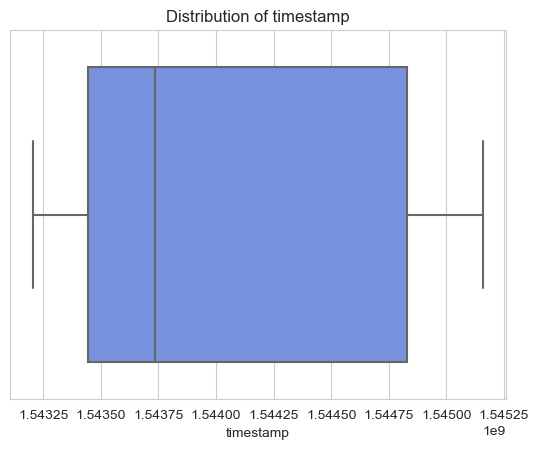
To make the attributes of data easier to understand we make changes to it known as One-hot encoding which is a technique used to represent categorical variables as numerical variables so that machine learning models can use them as inputs.

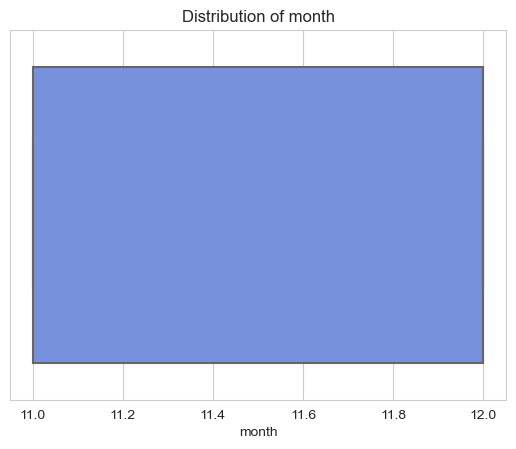
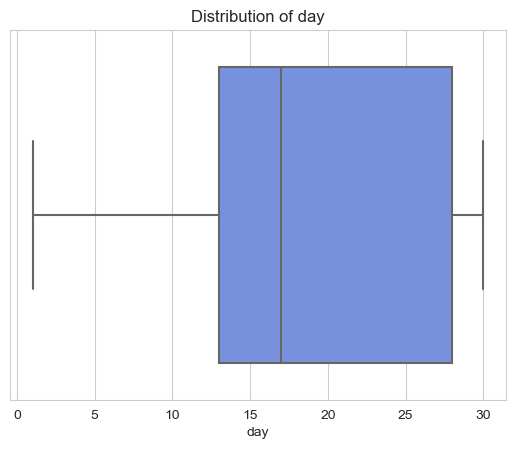


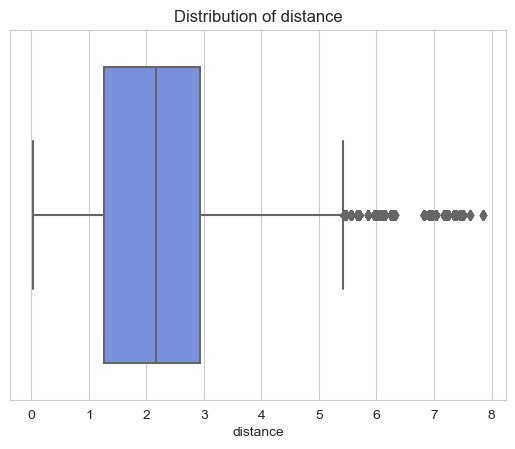
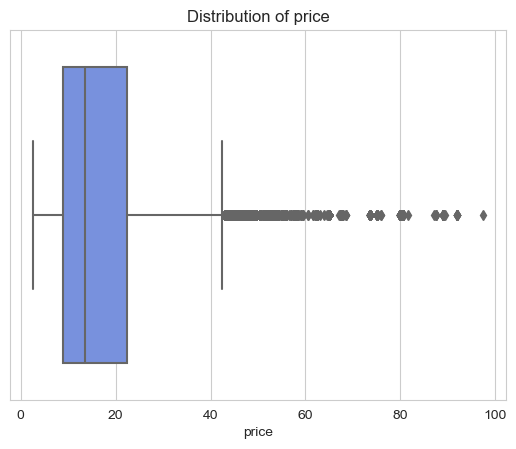
1. **Visualization**

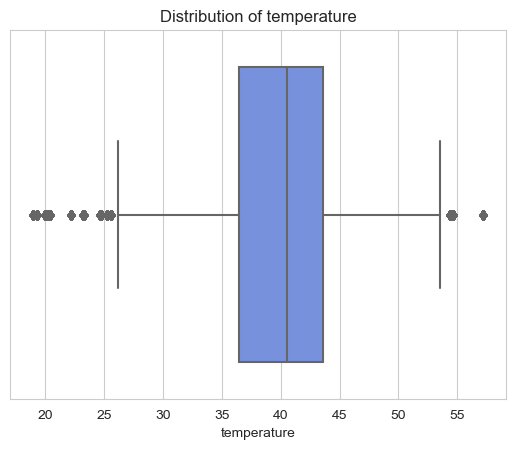
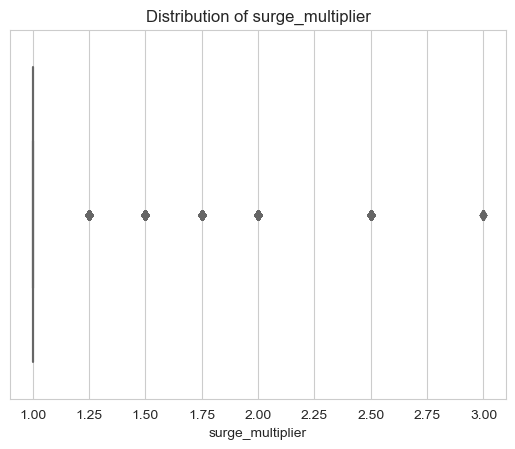
Data visualization is used to make complex data easier to understand, identify relationships and correlations, and communicate insights and findings to others. It also makes data more engaging, which can encourage people to explore it further. Finally, data visualization supports decision-making by providing a clear, visual representation of the data that can help identify trends and patterns that might be missed in other forms of analysis.











**Observations:**

The following are insights based on the given data:

* Data is almost equally distributed for source and destination:

The even distribution of data for source and destination suggests that the customers booking the cab are spread out across various locations, which can be advantageous for cab companies to optimize their services and coverage.

* Weather was cloudy when most of the customers booked cab and least of them booked on foggy day:

It appears that customers are more likely to book a cab on cloudy days and less likely on foggy days. This may be due to the fact that cloudy weather may not significantly impact transportation, while foggy weather can be challenging and even dangerous to drive in, causing people to avoid booking a cab altogether.

* Most of the cabs were booked in the midnight mostly after 10 P.M.:

The high number of bookings after 10 P.M. indicates that there is significant demand for transportation services during late hours. Cab companies can use this insight to optimize their services during these peak hours.

* Month end found to be the busiest days:

The observation that month-end is the busiest day for cab bookings could suggest that customers are more likely to book a cab during payday or while running errands to complete their month-end tasks.

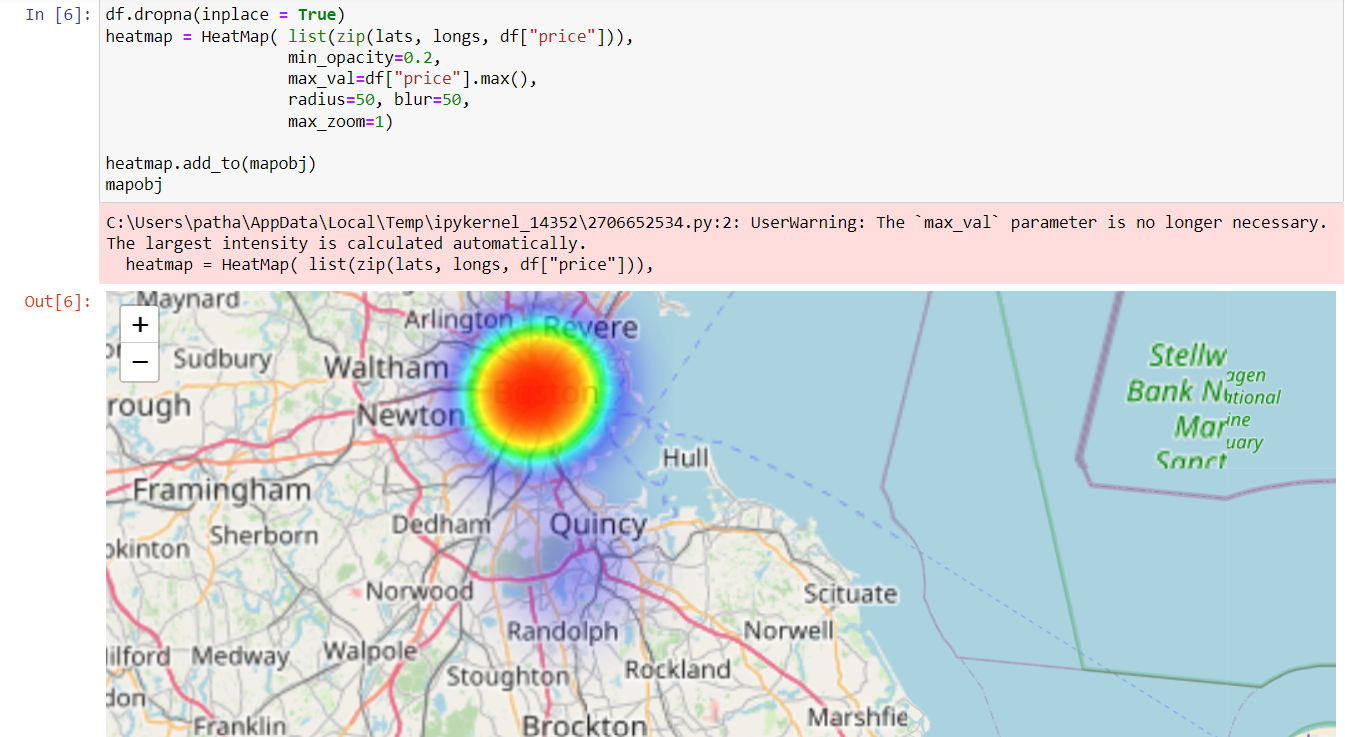
* People mostly booking the cabs which are budget-friendly:

Customers preferring budget-friendly cabs could indicate that price sensitivity is an essential factor for cab bookings. Cab companies can use this insight to offer more budget-friendly services or pricing strategies to attract more customers.

1. **Geometric Analysis**

Geometric analysis is used to study geometric objects and their properties such as shape, size, and position. It is used to provide a rigorous mathematical foundation for various areas such as physics, engineering, and computer science. Geometric analysis enables the development of powerful tools to solve complex problems in these fields.

****

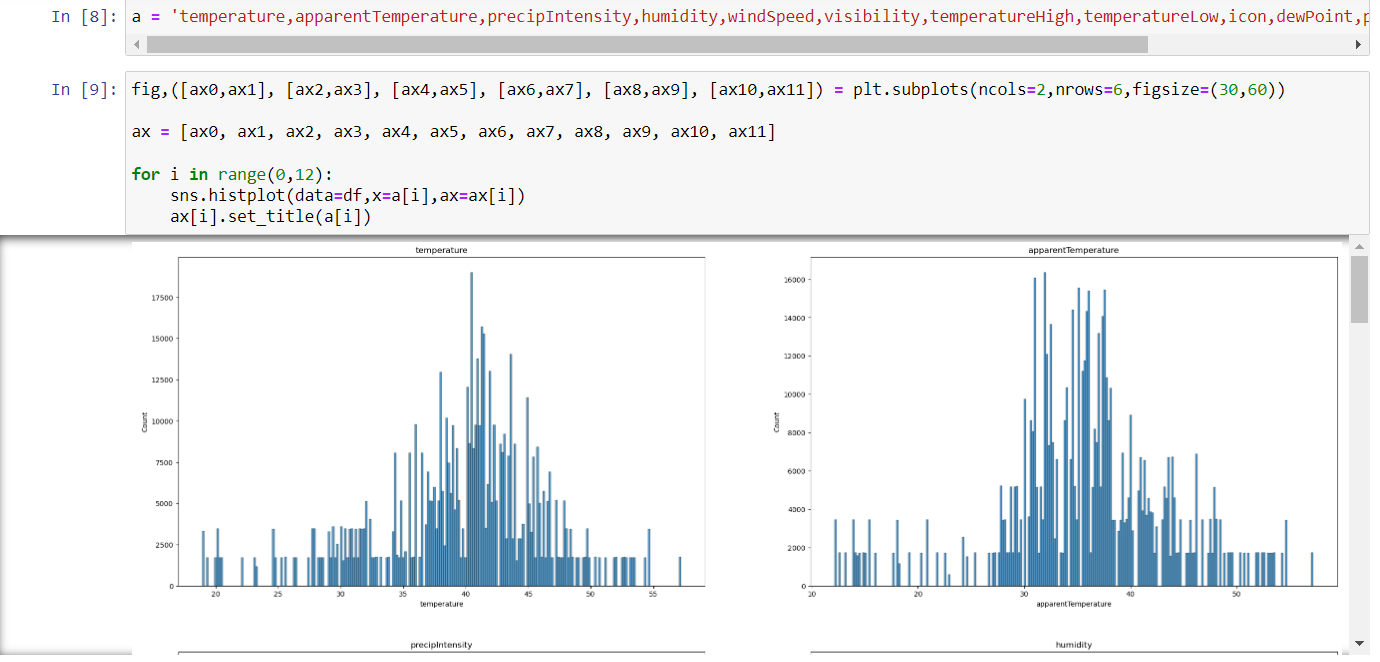


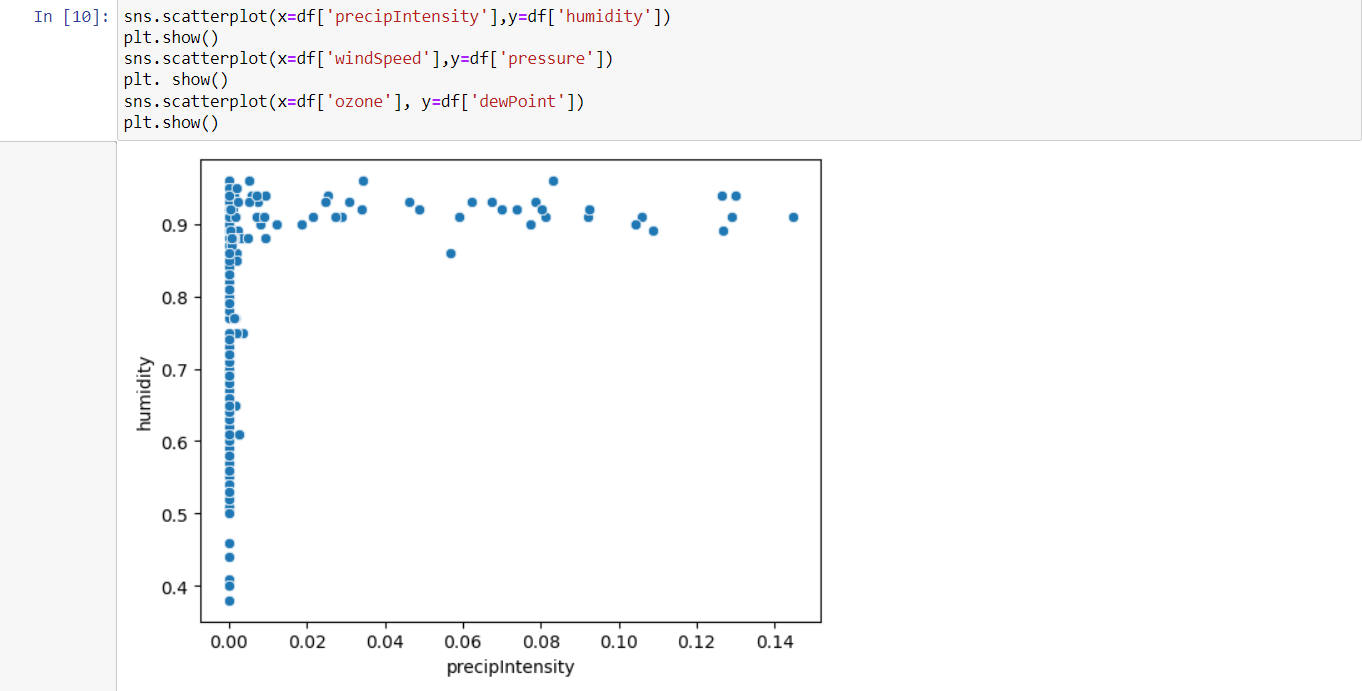
**Observations:**

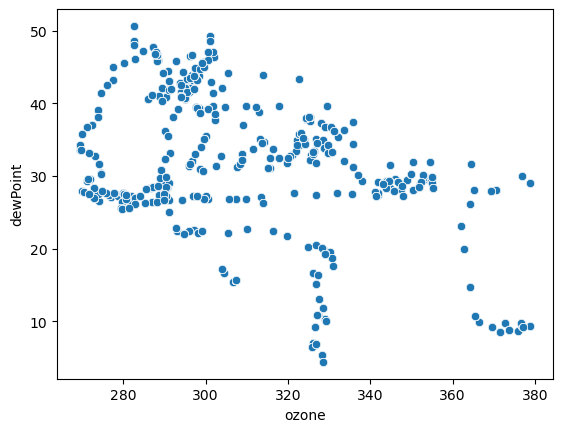
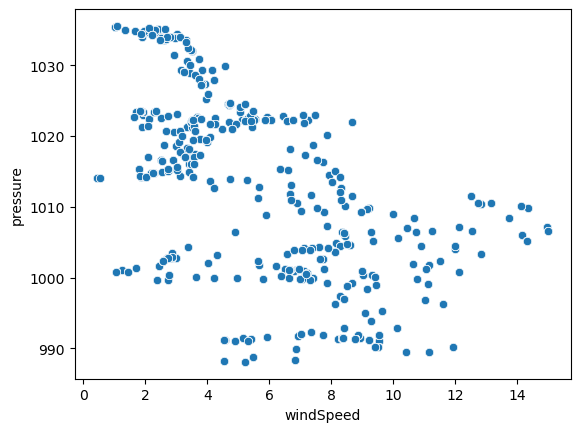
According to the available data, most of the cab bookings are from the Boston area. This could be due to various factors, such as a higher population density, greater business opportunities. Another possible factor could be the quality and reliability of cab services in the Boston area. Cab companies operating in this region may have a reputation for providing efficient and affordable transportation services.

1. **Psychographic Analysis**

Psychographics helps in understanding consumer behaviour by analyzing their personality, values, interests, and lifestyle. It provides insights into the motivations and attitudes of the target audience, which can help marketers create more effective marketing strategies. By understanding the psychographics of their target audience, businesses can tailor their products and services to better meet customer needs and preferences.





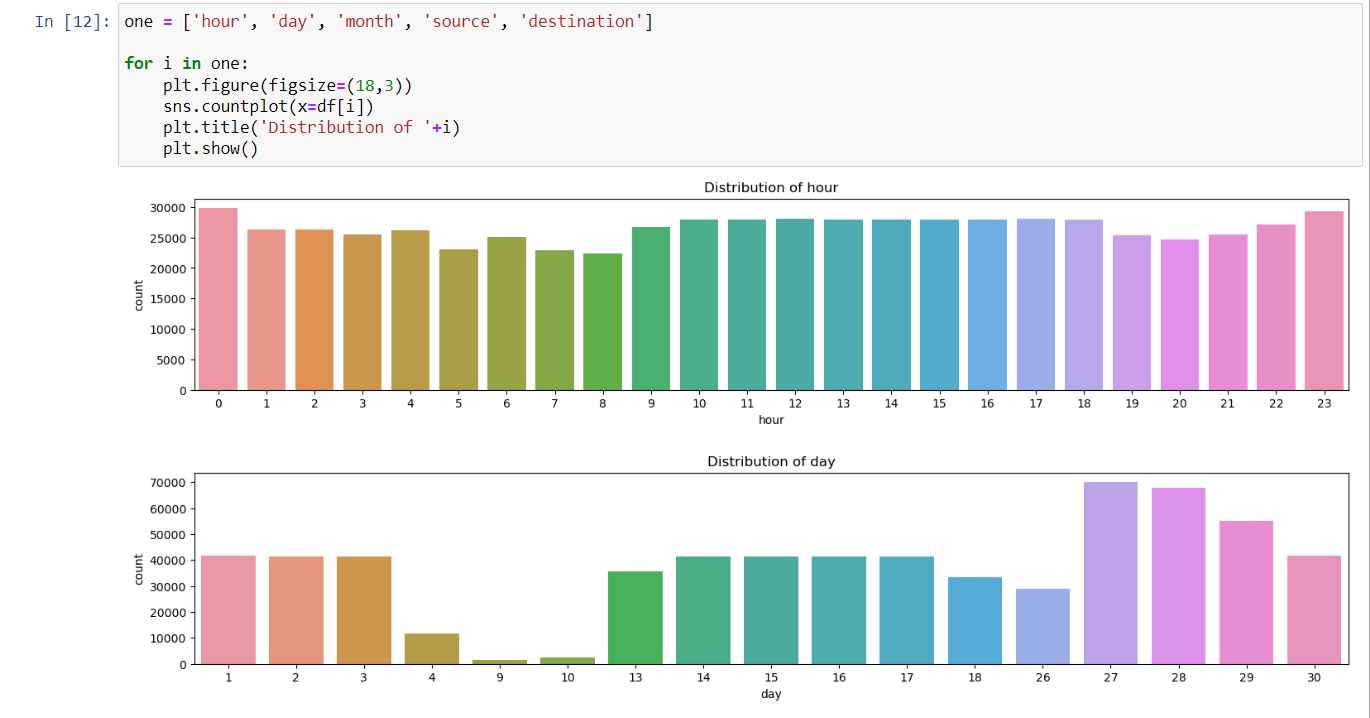


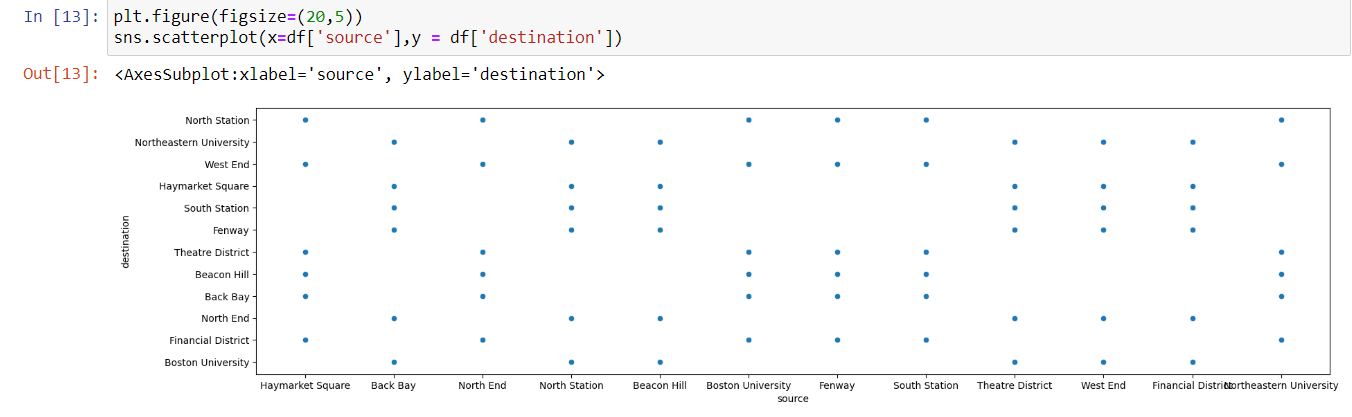
**Observations:**

The distribution of temperature is approximately normal, with a majority of values ranging from 35 to 45 degrees. The preparation intensity is centered around 0.00, and visibility is around 10. Interestingly, the busiest day for cab bookings was found to be cloudy, while the least busy day was surprisingly a foggy day. Additionally, customers are more likely to ride a cab when there is a precipitation intensity greater than 0.01 and humidity greater than 0.8. These findings provide valuable insights into the relationships between weather conditions and cab booking behaviour.

1. **Demographic Analysis:**

Demographic analysis helps in understanding the characteristics of a population, such as age, gender, income, and education. It provides insights into the preferences and behaviors of a particular group, which can help in developing effective marketing strategies. By understanding the demographic makeup of their target audience, businesses can tailor their products and services to better meet customer needs and preferences.



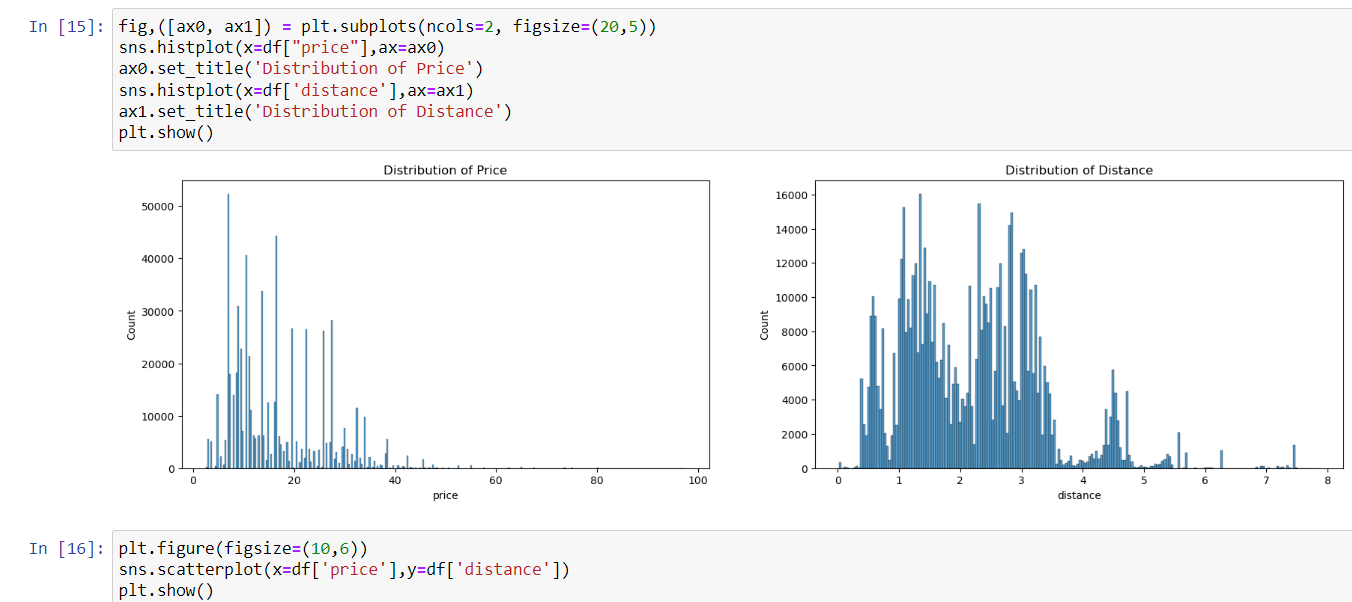


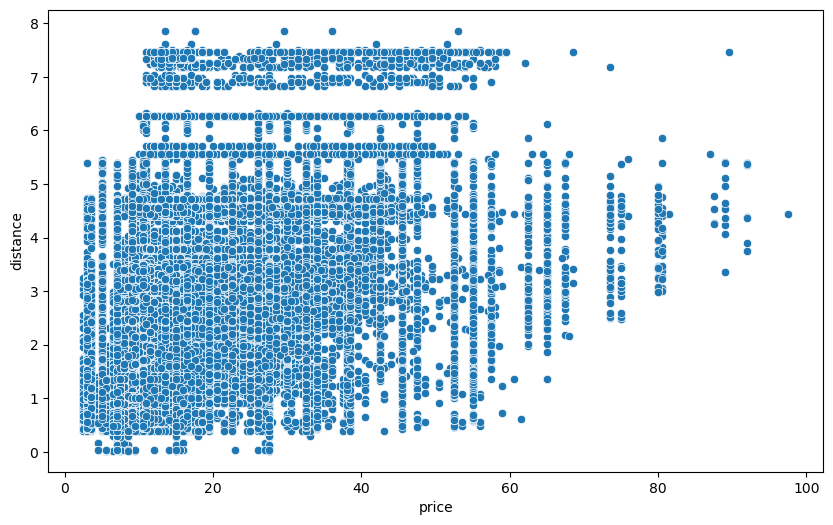
**Observations:**

According to the data from the last two months of 2018, the busiest times for cab bookings are between 10 A.M. to 6 P.M. and after 10 P.M. Similarly, the end of the month appears to be the busiest period for cab drivers, while the period from the 4th to the 13th of each month sees relatively fewer cab bookings. The data also indicates that there is a relatively even distribution of cab bookings across all source and destination points.

1. **Behaviour Analysis:**

Behaviour analysis helps in understanding the actions and choices made by individuals, providing insights into their preferences and motivations. It helps businesses identify the factors that influence consumer behaviour and develop effective marketing strategies. By understanding consumer behaviour, businesses can improve their products and services, enhance customer satisfaction, and increase profitability.





**Observations:**

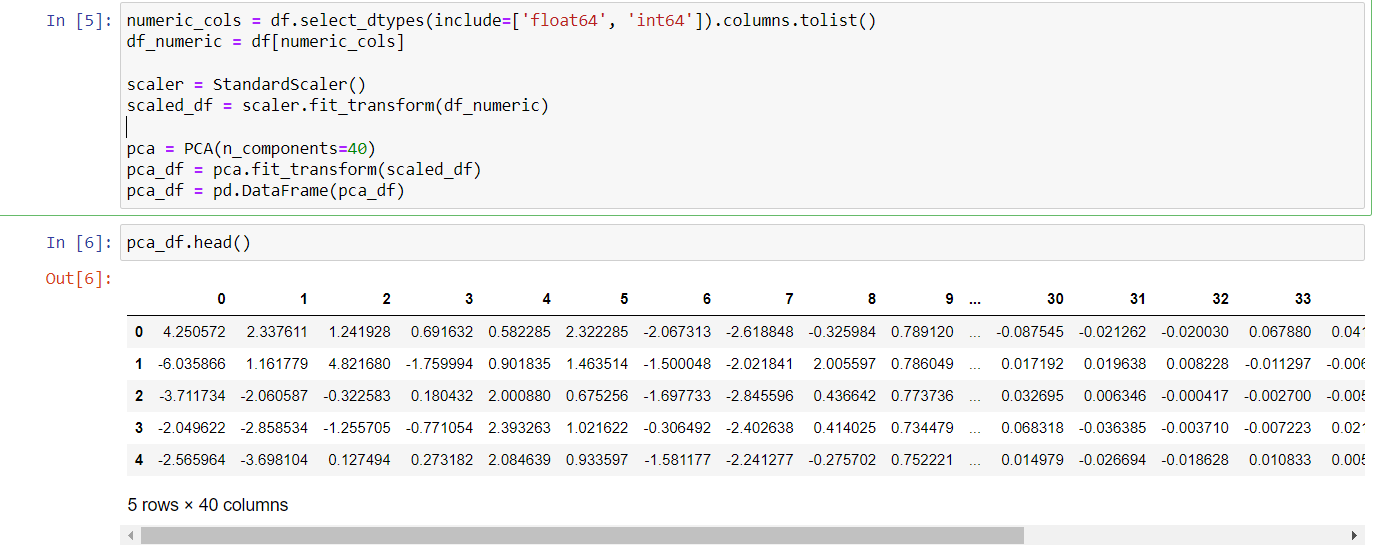
The majority of customers tend to book budget-friendly cabs with fares ranging from 5 to 25. Additionally, most customers prefer to book cabs for shorter distances, typically between 0.5 to 3.5 units. As the distance and price of the cab increases, the likelihood of customers booking decreases. These findings highlight the importance of price and distance in customer decision-making when it comes to booking a cab, and suggest that businesses should consider offering more affordable options for shorter trips to attract and retain customers.

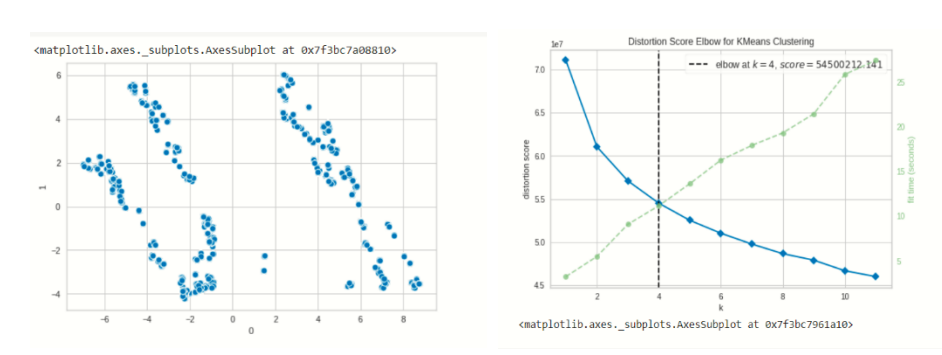
**Segment Extraction**

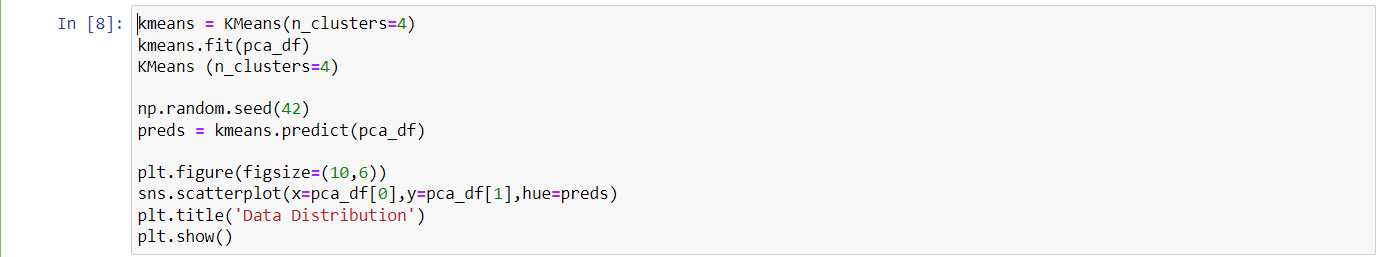
Segment extraction is a process of dividing a larger population or market into smaller subgroups or segments based on certain criteria or characteristics such as demographics, psychographics, behaviour, and geographic location. This process allows businesses to understand their target audience in more detail and develop targeted marketing strategies to better meet their needs and preferences. By identifying distinct customer segments, businesses can tailor their products, services, and messaging to effectively reach and engage each group, which can lead to increased customer satisfaction and loyalty, as well as improved profitability. Overall, segment extraction is a crucial step in the marketing process that helps businesses optimize their resources and drive growth.

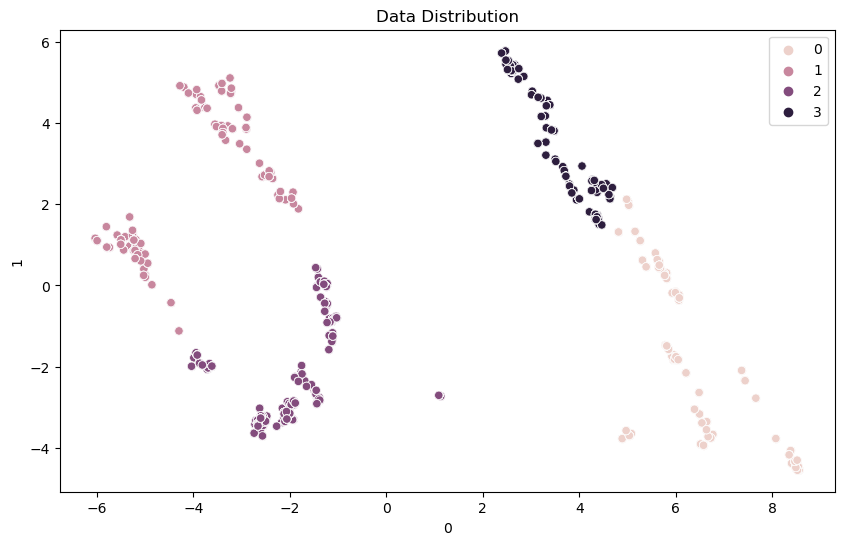
1. **Clustering:**

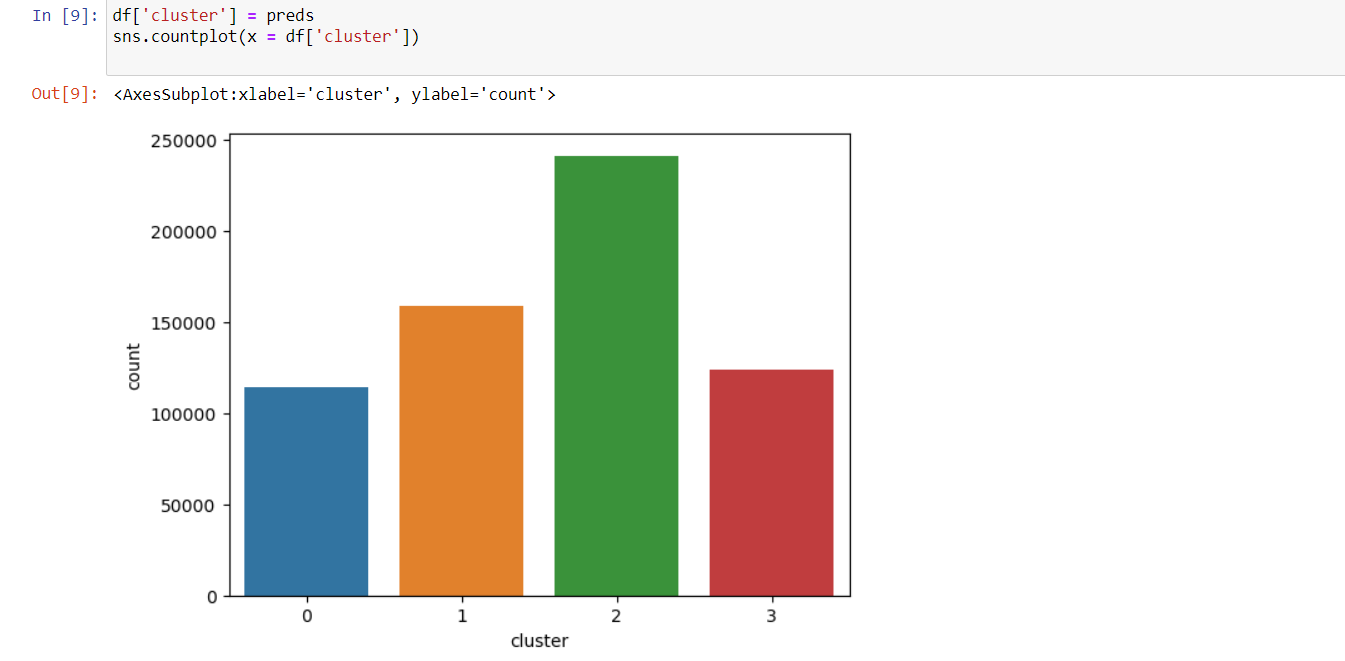
Clustering is important in identifying patterns and grouping similar data points together. It helps in understanding complex data sets, identifying market segments, and improving decision-making. Here by using K means clustering we have divided the set into 4 clusters using Elbow Curve method.

****

****

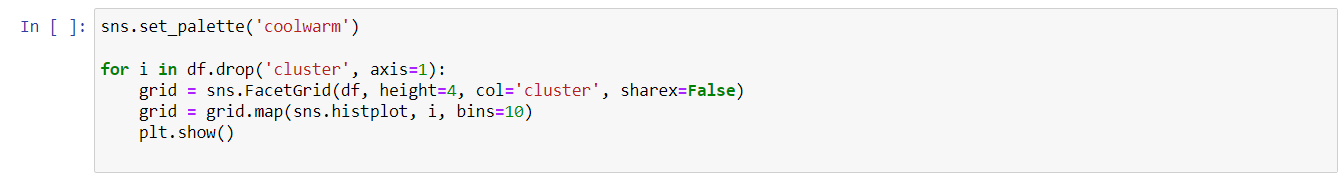
****

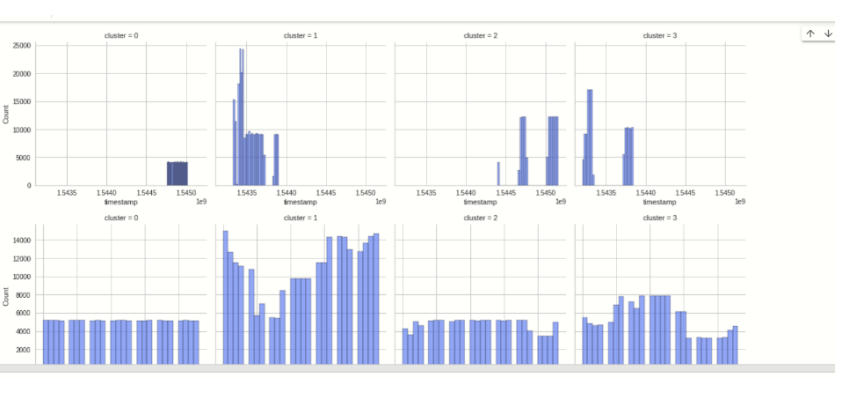
****

****

1. **Profiling Segments:**

Profiling segments involves analyzing the characteristics and behaviours of each segment identified through clustering. This analysis includes demographic and psychographic factors, as well as consumer behaviour patterns, such as purchase history and preferences. Profiling helps businesses understand the unique needs and preferences of each segment, allowing them to tailor their marketing strategies to effectively engage and meet the needs of each group.





**Final conclusions**

Based on the analysis of the cab booking data from the last two months of 2018, we can draw several observations and conclusions:

1. The majority of customers prefer to book budget-friendly cabs, with fares ranging from 5 to 25, for shorter distances ranging from 0.5 to 3.5 units.
2. Customers are more likely to book a cab during the day, with the busiest times being between 10 A.M. to 6 P.M. and after 10 P.M.
3. The end of the month is the busiest period for cab drivers, while the period from the 4th to the 13th of each month sees relatively fewer cab bookings.
4. The data shows an even distribution of cab bookings across all source and destination points.
5. The temperature distribution is approximately normal, with a majority of values ranging from 35 to 45 degrees.
6. The busiest day for cab bookings was found to be cloudy, while the least busy day was a foggy day.
7. Customers are more likely to ride a cab when there is a precipitation intensity greater than 0.01 and humidity greater than 0.8.
8. Most of the cab bookings are from the Boston area, indicating a high demand for efficient and affordable transportation services in the region.
9. Price sensitivity is an essential factor for cab bookings, with customers preferring budget-friendly options.
10. Cab companies can use these insights to optimize their services, pricing strategies, and coverage to attract and retain more customers.

To start a new cab service, various parameters should be considered, including customer demographics, booking behaviour, weather conditions, and time of day. For example, offering budget-friendly options for shorter distances could attract customers who prioritize affordability. It is also important to consider the busiest times for cab bookings, such as late evenings and month-end periods, and optimize services accordingly. Furthermore, monitoring weather conditions and understanding the relationship between weather and cab bookings can help in predicting demand and optimizing services accordingly. By considering these factors, a new cab service can tailor its services to meet customer needs and preferences and potentially gain a competitive advantage in the market.