FoodHub Data Analysis

Python - Foundations

Contents / Agenda

- Executive Summary
- Business Problem Overview
- Solution Approach
- Data Overview
- EDA - Univariate Analysis
- Univariate Analysis - Key Questions
- EDA - Multivariate Analysis
- Multivariate Analysis - Key Questions

Executive Summary

- FoodHub should integrate with restaurants serving American, Japanese, Italian, and Chinese cuisines as these cuisines are very popular among FoodHub customers (accounting for ~80% of the orders).
- FoodHub should provide promotional offers to top-rated popular restaurants that serve most of the orders as the top 5 restaurants account for ~80% of the orders.
- Order volumes are higher (by ~60%) on the weekends compared to the weekdays. As such, more delivery executives should be employed during the weekends to ensure timely delivery of the orders. Weekend promotional offers can also be rolled out to the customers to increase the food orders during weekends.
- Delivery time over the weekends is less compared to the weekdays despite the higher number of orders. This could possibly be due to the dip in traffic volume over the weekends, but further analysis is needed to verify the same.
- The customer rating is a very important factor to gauge customer satisfaction and ~39% of the orders were not rated. The company should investigate the reason behind the low count of ratings. They can redesign the rating page in the app and make it more interactive to lure the customers to rate the order. Customer feedback comments should also be analyzed for further insights.
- Around 11% of the total orders have more than 60 minutes of total delivery time. FoodHub should try to minimize such instances in order to avoid customer dissatisfaction. They can provide some reward to the punctual delivery persons.

Business Problem Overview

- The number of restaurants in New York is increasing day by day. Lots of students and busy professionals rely on those restaurants due to their hectic lifestyles. Online food delivery service is a great option for them. It provides them with good food from their favorite restaurants. A food aggregator company FoodHub offers access to multiple restaurants through a single smartphone app.
- The app allows the restaurants to receive a direct online order from a customer. The app assigns a delivery person from the company to pick up the order after it is confirmed by the restaurant. The delivery person then uses the map to reach the restaurant and waits for the food package.
- Once the food package is handed over to the delivery person, he/she confirms the pick-up in the app and travels to the customer's location to deliver the food. The delivery person confirms the drop-off in the app after delivering the food package to the customer. The customer can rate the order in the app. The food aggregator earns money by collecting a fixed margin of the delivery order from the restaurants.

Solution Approach

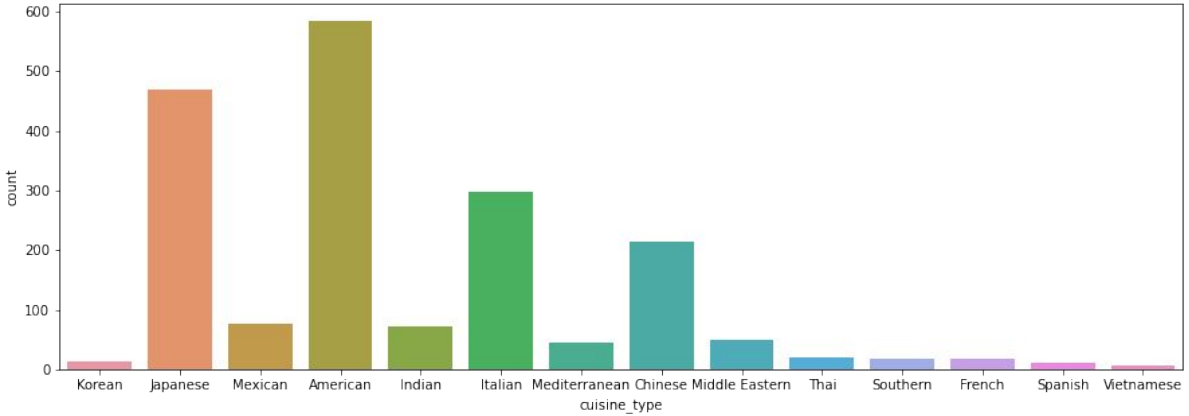
- Understand the demand of restaurants in the FoodHub portal
- Cuisine preference of the New York customers
- Get an idea about the cost of the ordered food
- Understand the volume of the orders over weekdays and weekends
- Estimate the revenue generated by the company
- Help the company to take decision on promotional offers
- Order rating analysis

Data Overview

- The data provided is of various customers who have ordered food in New York from FoodHub
- There are total of 1898 observations with 9 attributes that includes information about the various orders including order id, customer id, restaurant name, cuisine type, cost of the order, day of the week, rating, food preparation time and delivery time
- There are no missing values in the data
- There are no outliers in the data

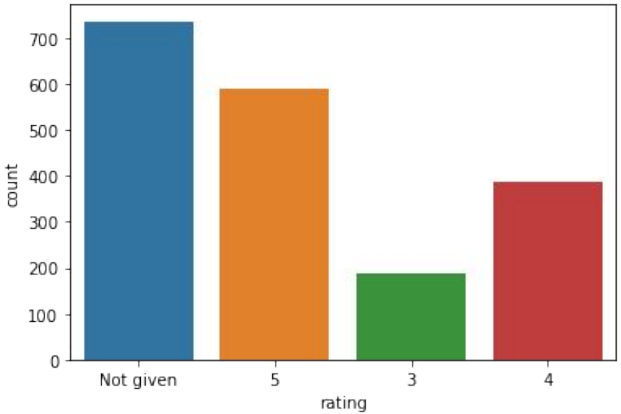
Link to Appendix slide on data background check

Univariate Analysis

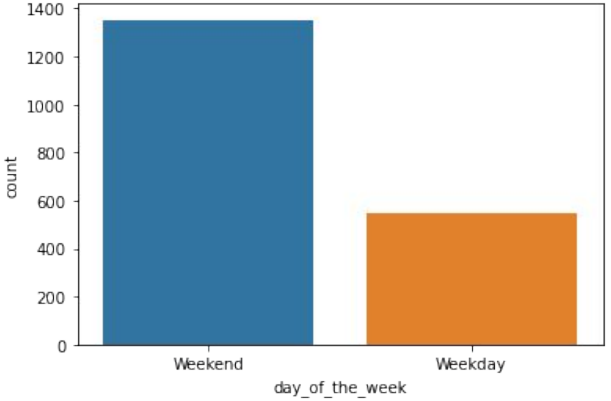


- There are 14 unique cuisines in the dataset.
- The distribution of cuisine types show that cuisine types are not equally distributed.
- The most frequent cuisine type is American followed by Japanese and Italian.
- Vietnamese appears to be the least popular of all the cuisines.

Univariate Analysis

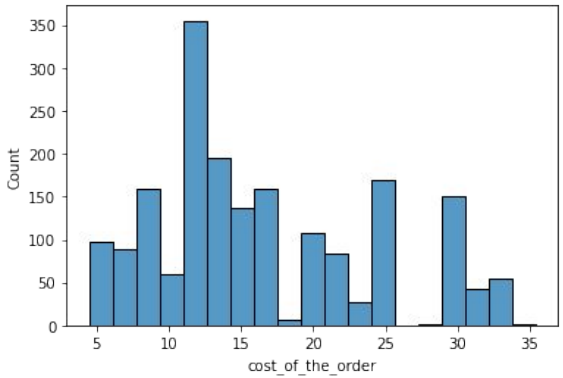


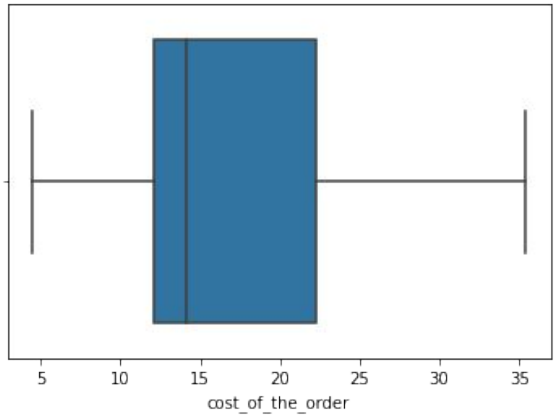
- The distribution of 'rating' shows that the most frequent rating category is 'not given', followed by a rating of 5.
- Only around 200 orders have been rated 3.



- The 'day_of_the_week' columns consists of 2 unique values - Weekday and Weekend
- The distribution shows that around 1300 orders are placed on weekends.

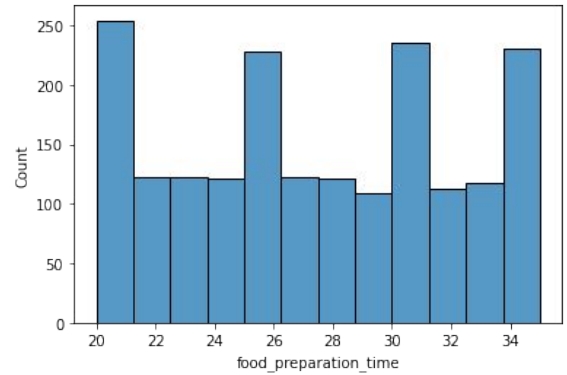
Univariate Analysis

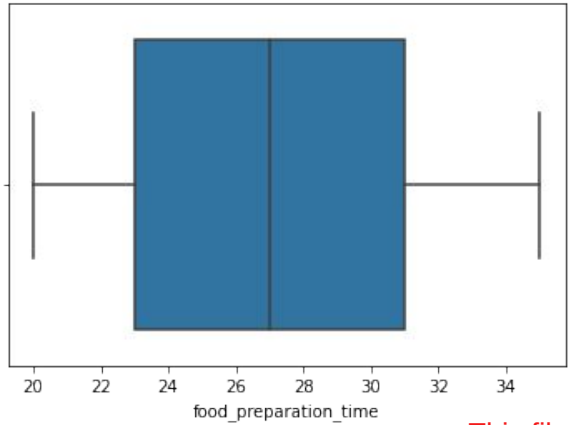




- The average cost of the order is greater than the median cost indicating that the distribution for the cost of the order is right-skewed.
- The mode of the distribution indicates that a large chunk of people prefer to order food that costs around 10-12 dollars.
- There are few orders that cost greater than 30 dollars. These orders might be for some expensive meals.

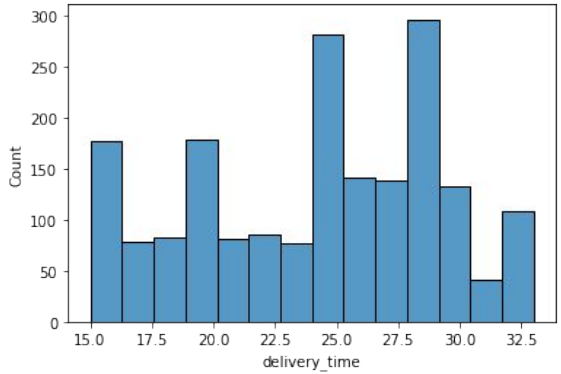
Univariate Analysis

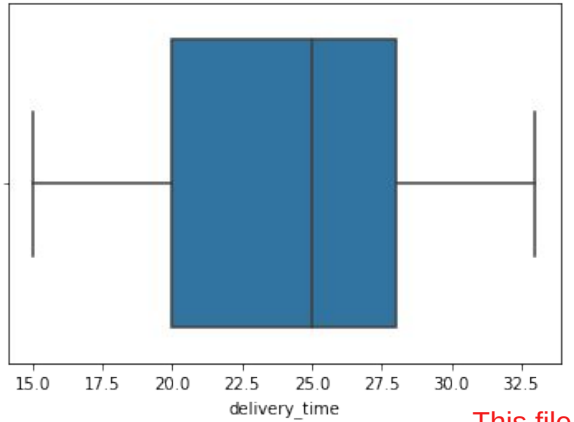




- The average food preparation time is almost equal to the median food preparation time indicating that the distribution is nearly symmetrical.
- The food preparation time is pretty evenly distributed between 20 and 35 minutes.
- There are no outliers in this column

Univariate Analysis





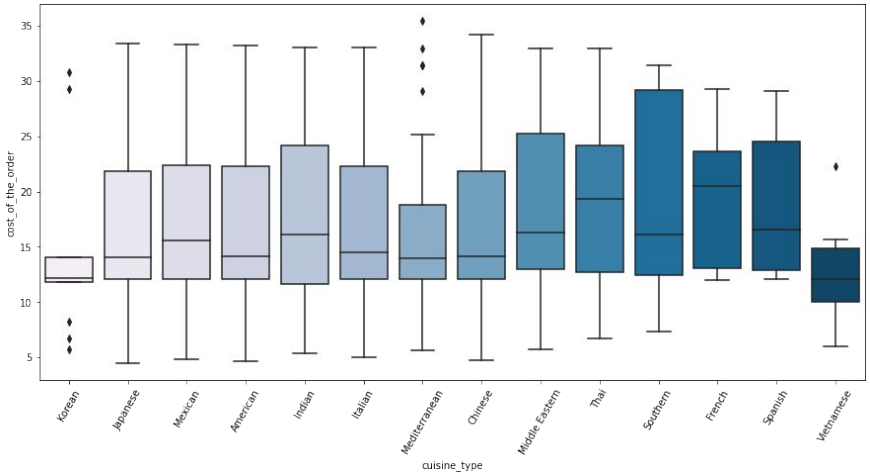
- The average delivery time is a bit smaller than the median delivery time indicating that the distribution is a bit left-skewed.
- Comparatively more number of orders have delivery time between 24 and 30 minutes.
- There are no outliers in this column.

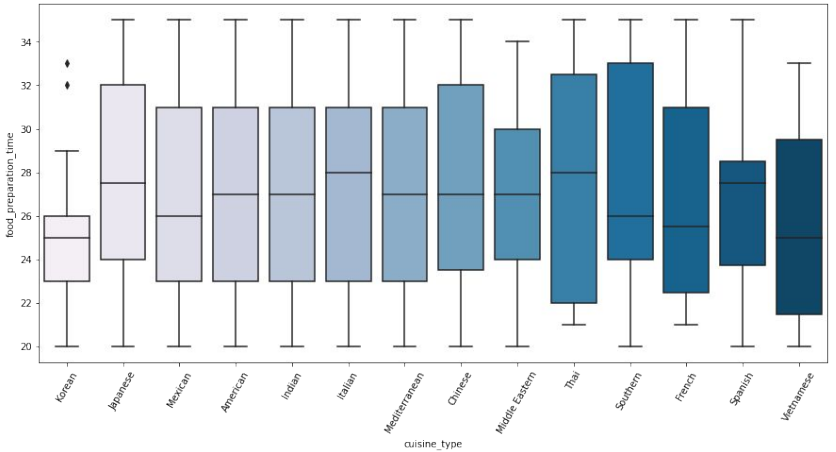
Univariate Analysis - Key Questions

- Delivery time ranges from 15 to 33 minutes, with an average of around 24 minutes and a standard deviation of 5 minutes. The spread is not too high for delivery time either
- There are 736 orders that are not rated
- Top 5 popular restaurants that have received the highest number of orders 'Shake Shack', 'The Meatball Shop', 'Blue Ribbon Sushi', 'Blue Ribbon Fried Chicken' and 'Parm'
- Almost 33% of the orders in the dataset are from these restaurants
- The most popular cuisine type on weekends is American
- There are a total of 555 orders that cost above 20 dollars. The percentage of such orders in the dataset is around 29.24%
- The mean delivery time is around 24.16 minutes
- Customer with ID 52832 has ordered 13 times. Other customers with Customer IDs 52832, 47440, 83287, 250494 and 259341 can avail discount offer of 20%

Link to Appendix slide on data background check

Multivariate Analysis

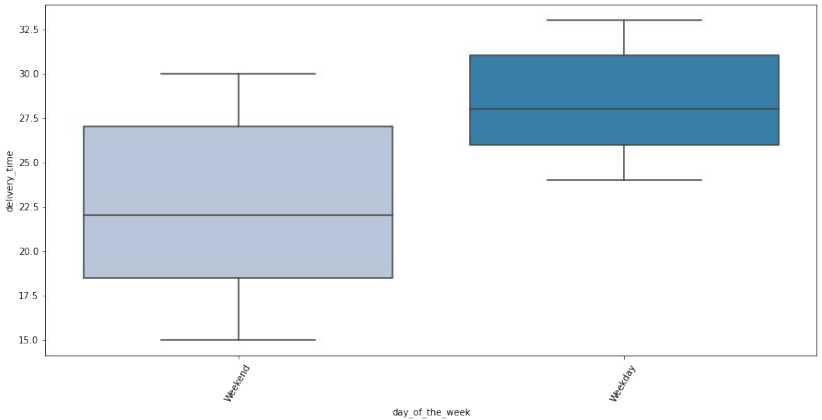


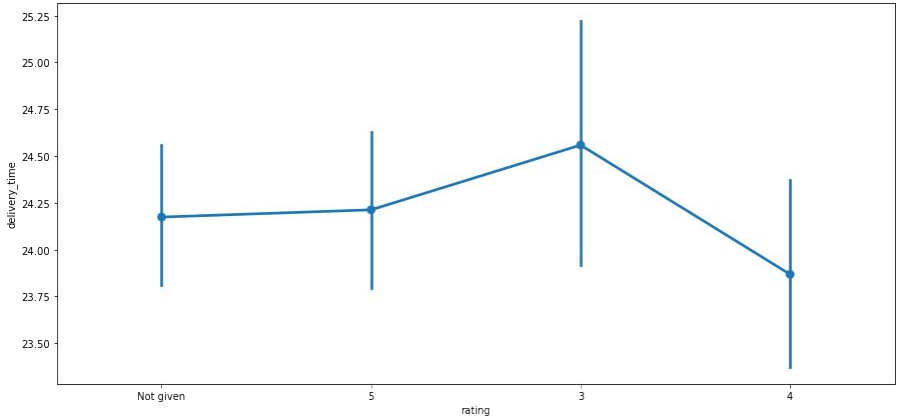


- Vietnamese and Korean cuisines cost less compared to other cuisines.
- The boxplots for Italian, American, Chinese, Japanese cuisines are quite similar. This indicates that the quartile costs for these cuisines are quite similar.
- French and Spanish cuisines are costlier compared to other cuisines.

- Food preparation time is very consistent for most of the cuisines.
- The median food preparation time lies between 24 and 30 minutes for all the cuisines.
- Outliers are present for the food preparation time of Korean cuisine.
- Korean cuisine takes less time compared to the other cuisines.

Multivariate Analysis

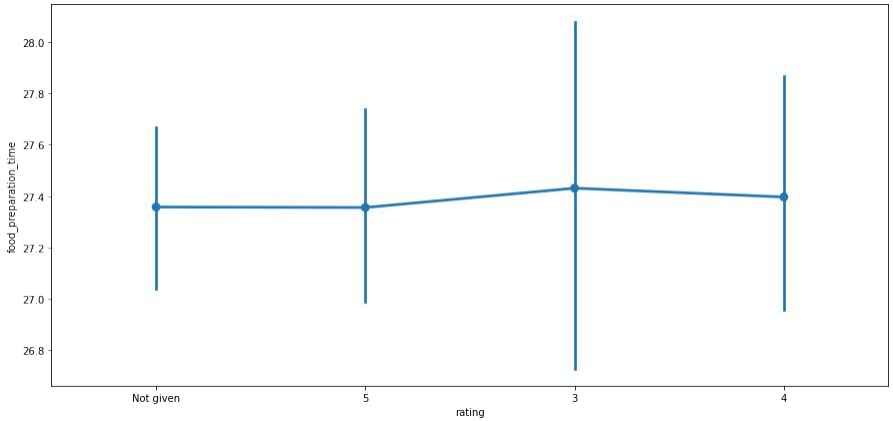




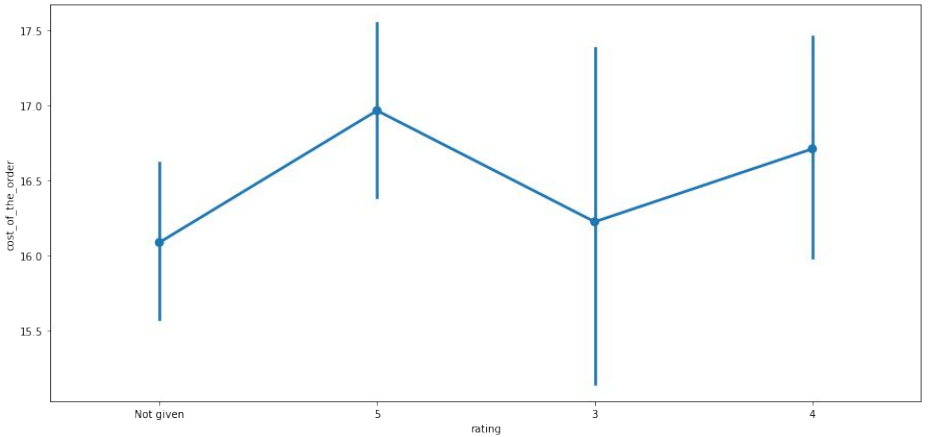
- The delivery time for all the orders over the weekends is less compared to weekdays. This could be due to the dip in traffic over the weekends

- It is possible that delivery time plays a role in the low-rating of the orders.

Multivariate Analysis



- It seems that food preparation time does not play a role in the low-rating of the orders.



- It seems that high-cost orders have been rated well and low-cost orders have not been rated.

Multivariate Analysis - Key Questions

- The restaurants fulfilling the criteria to get the promotional offer are: 'The Meatball Shop', 'Blue Ribbon Fried Chicken', 'Shake Shack' and 'Blue Ribbon Sushi'.
- The net revenue generated on all the orders given in the dataset is around 6166.3 dollars.
- Approximately 10.54 % of the total orders have more than 60 minutes of total delivery time.
- The mean delivery time on weekdays is around 28 minutes whereas the mean delivery time on weekends is around 22 minutes.
- This could be due to the dip of traffic volume in the weekends.

Great Learning

Happy Learning !