Data Analysis for FoodHub Co.

Sarah Choi

Date: 8/4/2023

Contents / Agenda

- Executive Summary
- Business Problem Overview and Solution Approach
- Data Overview
- EDA - Univariate Analysis
- EDA - Multivariate Analysis
- Appendix

Executive Summary - Conclusions

	Conclusion
Cuisine Type	• The most popular cuisine that Customers order is American cuisine. • More orders placed on the weekends than during the weekdays. • Among all the orders placed on the weekend, American cuisine has the highest count and Japanese cuisine comes second.
Cost of Orders	• The average cost of an order is around 14.14 dollars and most customers order cuisine that cost around 13 dollars. • There are more Southern cuisine orders that cost more than $20 dollars as compared to all other cuisines.
Feedback Ratings	• Different ratings (3, 4, 5) are provided by the customers on the orders; however, there is no rating provided for 736 orders which is around 61% of the total. For the orders that received the most counts are also rated the highest. • The orders with consistently higher cost receive higher rating, it seems that orders which cost more may be prepared with higher quality & good ingredients and therefore received better customer satisfaction and hence better rating. • Customers gave the lowest rating when the mean delivery time is longer than 24.5 mins.
Food preparation & delivery time	• Food preparation time for all cuisines are quite constant with an average of 27 mins to complete. There is no relationship or direct impact between customer rating and the time it took to prepare the food. • The number of orders that required total time of more than 60 mins to deliver is: 200 which made up 10.54% of the total. • The bigger impact to total time required to deliver the food is the food delivery time. • The higher cost of the food, the shorter is the delivery time and the higher cost of food is also associated with higher customer rating in the other metric. These two factors contribute to overall great customer satisfaction. • There are more food deliveries completed during weekends than weekdays because there is higher number of food order placed during the weekends. • It took longer to deliver the orders during weekdays compared to weekend, on an average of 28 mins during weekdays and 22 mins on weekends.

Executive Summary - Recommendations

	Recommendations
Cuisine Type	• Top 5 restaurants that sold the most orders are: Shake Shack, The Meatball Shop, Blue Ribbon Sushi, Blue Ribbon Fried Chicken, Parm. • Company (FoodHub) should develop promotion offers for these restaurants for growth opportunities.
Cost of Orders	• The number of total orders that cost above 20 dollars is 555 which made up of 29.24% of total. • There are more orders that cost more than 20 dollars for the Southern cuisine which shows a lot of potential. • Company should consider charging the restaurant 25% on the orders having cost above 20 dollars and 15% on the orders having cost greater than 5 dollars, this will generate a net revenue of around 6166.3 dollars.
Feedback Ratings	• Customers feedback rating is crucial to the company's business, we need to procure more data to help understand the reason why more than 60% of the customer orders did not receive any rating. • Company should share the positive customer rating with the restaurants and encourage theses restaurants to keep up with the quality of the food order to drive more business. • The restaurants which have a rating count of more than 50 and the average rating greater than 4 are: The Meatball Shop, Blue Ribbon Fried Chicken, Shake Shack, Blue Ribbon Sushi. Company should provide a promotional offer in the advertisement of the restaurants for growth opportunities.
Food preparation & delivery time	• Based on the low customer rating of the orders that took long delivery time, the company will need improvement of the food delivery service to enhance customer experience. • Assign the food delivery persons who are located close to the food delivery destination to minimize travel time in order to shorten the time to deliver the food. • Ensure more delivery persons are available during weekdays to spread out the delivery orders and in turn shorten the overall food delivery time for the customers.
Others	• There are quite a few categories in the dataset that have high variance and skewed data, e.g. the cost of the order and the delivery time. One of the drawbacks of data variance is that it results in a value in terms of units which can be difficult to interpret, and there could also be high error rates on test data when building machine learning model. We should try to curate the data as much as possible.

Business Problem Overview and Solution Approach

Problem Statement

Online food delivery service has become a very popular option for many students and busy professionals who rely on the restaurants due to their hectic lifestyles in the New York metropolitan area. A food aggregator company FoodHub offers access to multiple restaurants through a single smartphone app. 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. These services are the backbone for FoodHub to generate revenue by collecting a fixed margin of the delivery order from the restaurants. For this reason, it is important for FoodHub to understand the demand of different restaurants and to be able to detect patterns and to enhance the customer experience with using their services.

Solution Approach

Identity focus area	Perform Data Analysis	Summarize Recommendation
To extract actionable insights from the data that FoodHub has collected to identify key areas of focus.	Perform exploratory data analysis in the focus area : • Variables that influence the food orders • Factors that affect Customer Rating on the orders and the respective reasons • Type of cuisine and cost of order to formulate different pricing model to generate additional revenue	Summarize the business recommendations for growth and improvement

Data Overview

The data contains the different data related to a food order. The detailed data dictionary is given below.

Variable	Description
order_id	Unique ID of the order
customer_id	ID of the customer who ordered the food
restaurant_name	Name of the restaurant
cuisine_type	Cuisine ordered by the customer
cost_of_the_order	Cost of the order
day_of_the_week	Indicates whether the order is placed on a weekday or weekend
rating	Rating given by the customer out of 5
food_preparation_time	Time (in minutes) taken by the restaurant to prepare the food.
delivery_time	Time (in minutes) taken by the delivery person to deliver the food package

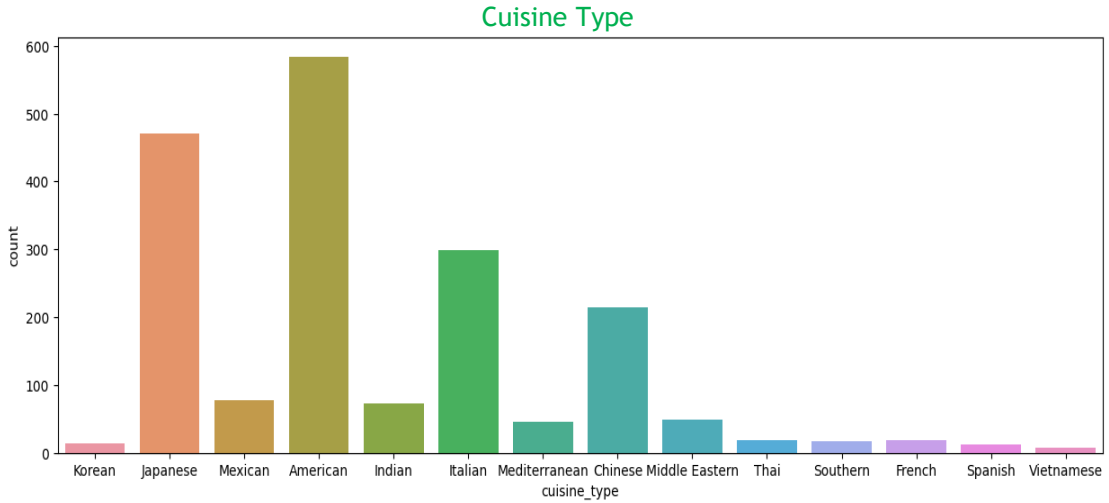
Observations	Variables
1898	9

Note:

- There is a total number of 1,898 records in the data frame with 9 columns for the data analysis
- There is no missing value in the data
- The food preparation time is calculated by taking the difference between the timestamps of the restaurant's order confirmation and the delivery person's pick-up confirmation.
- The delivery time is calculated by taking the difference between the timestamps of the delivery person's pick-up confirmation and drop-off information.

EDA - Univariate Data Analysis

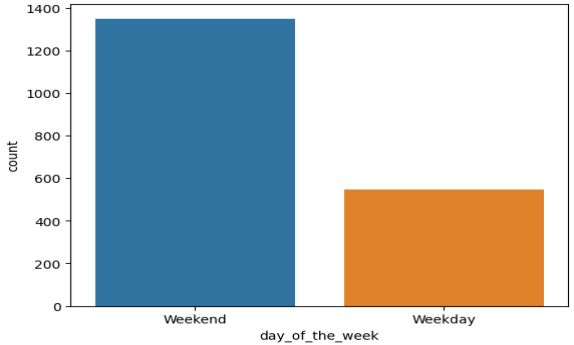
In this data, the exploratory data analysis is performed on 178 restaurants that offer 14 different type of cuisine, the data set consists of 1,200 customers who had placed 1,898 food orders. Let's explore the variables and find out how they are distributed.



- The most popular cuisine that Customers order is American cuisine
- Top 5 restaurants that sold the most orders are : Shake Shack, The Meatball Shop, Blue Ribbon Sushi, Blue Ribbon Fried Chicken, Parm

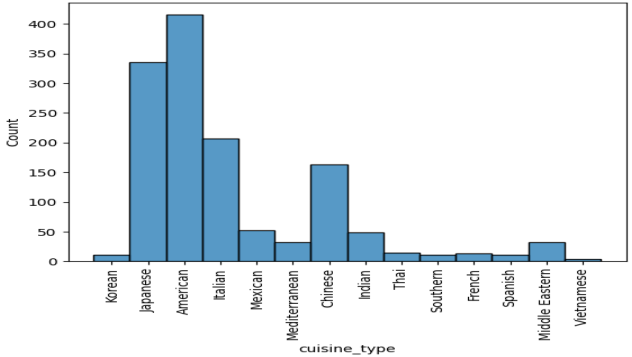
EDA - Univariate Data Analysis

Order placed on Weekday or Weekend



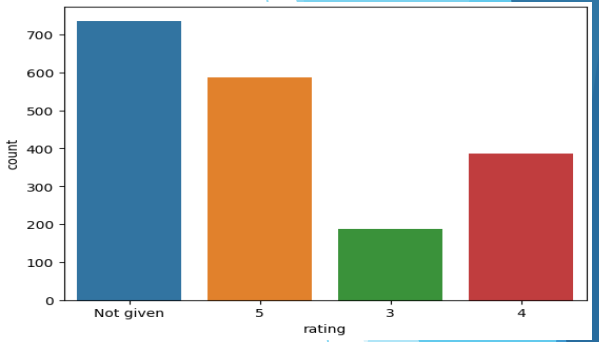
There are more orders placed on the weekends than during the weekdays

Order placed on Weekend by Cuisine



Among all the orders placed on the weekend, American cuisine has the highest count

Customer Ratings



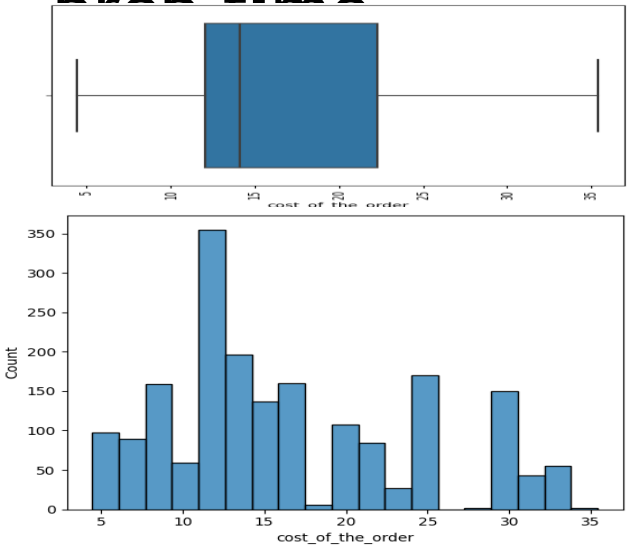
Different ratings are provided by customers on the orders; however, there is no rating provided for 736 out of the 1,898 orders

The company has decided to give 20% discount vouchers to the top 3 most frequent customers. The following is the customers' IDs and the number of orders they placed :

Customer ID	# of orders
52832	13
47440	10
83287	9

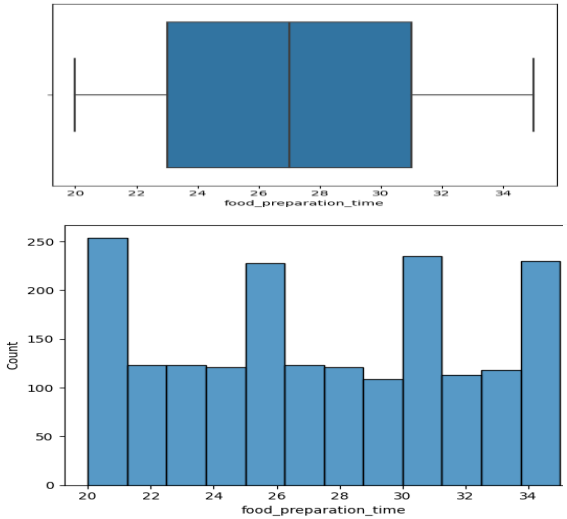
Univariate Analysis – Cost, Food delivery time, Food prep time

Cost of Orders



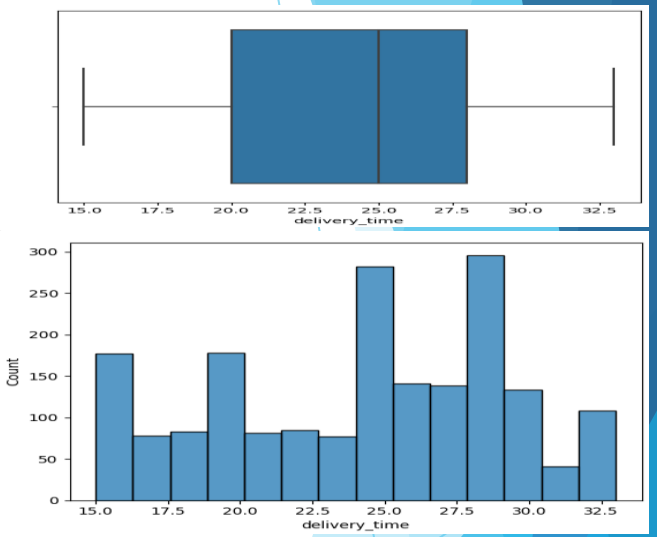
- The cost of order is positively skewed
- Average cost of the order is around $14.14
- Most customers order food that cost around $13
- The number of total orders that cost above $20 is 555
- Percentage of orders above $20 is 29.24 %

Food Preparation Time



- Food preparation time does not have any outliers
- 50% of the food preparation time required 27 mins
- Minimum preparation time is 20 mins and maximum being 35 mins
- The distribution of preparation time is quite even

Food Delivery Time

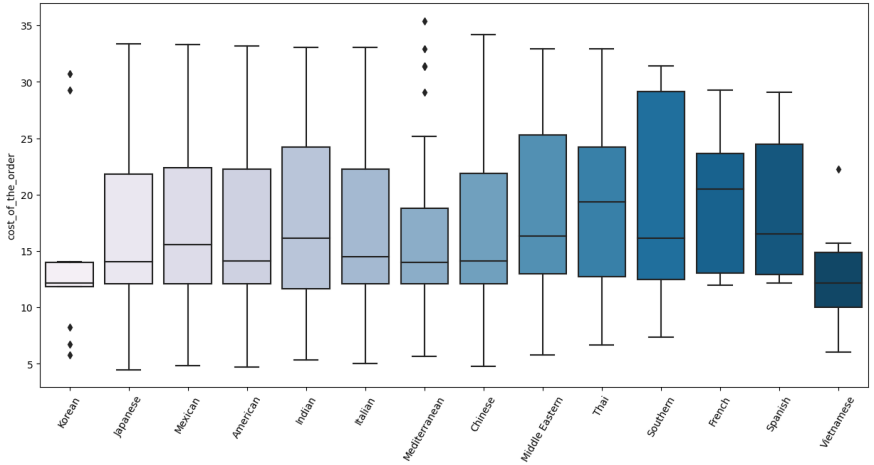


- The mean delivery time for this dataset is 25 minutes
- Minimum delivery time is 15 mins and maximum being 33 mins
- The delivery time is left skewed

Multivariate Analysis – Cusine vs Cost & Food prep time

Please ad

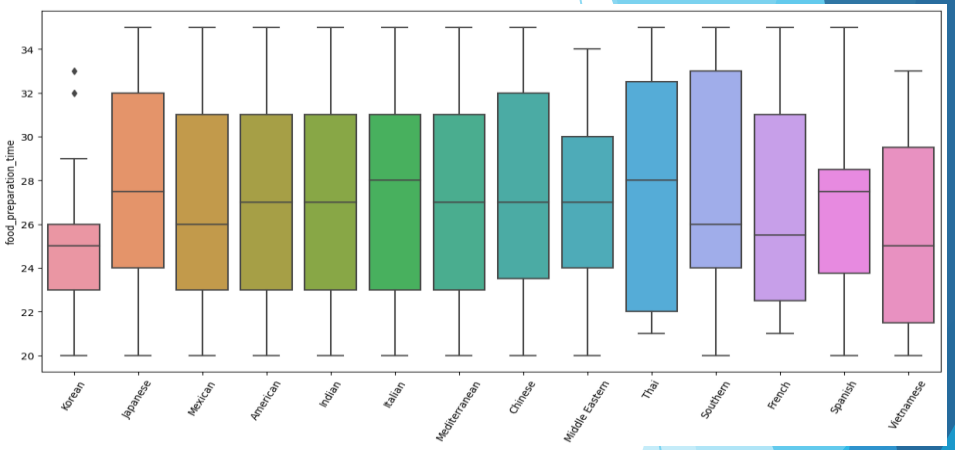
Cuisine vs Cost of Order



Observations:

- The Chinese cuisine has the highest variability of the cost of order, all the other cuisine have similar variance in the cost of order.
- There are more orders that cost more than $20 dollars for the Southern cuisine as compared to all other cuisines.
- The data for the Southern cuisine is very skewed, and the Korean cuisine seems to have the most outliners. The skewness and outliners from the data of these cuisines makes the data not normal and hard to predict.

Cuisine vs Food Preparation time

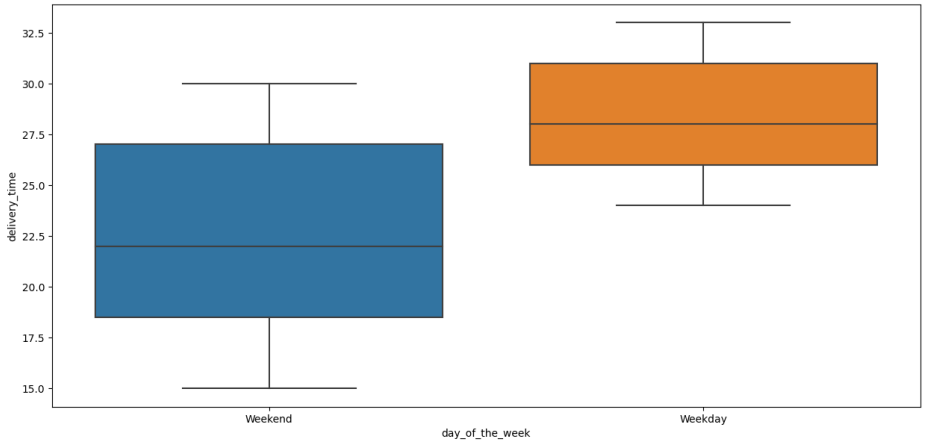


Observations:

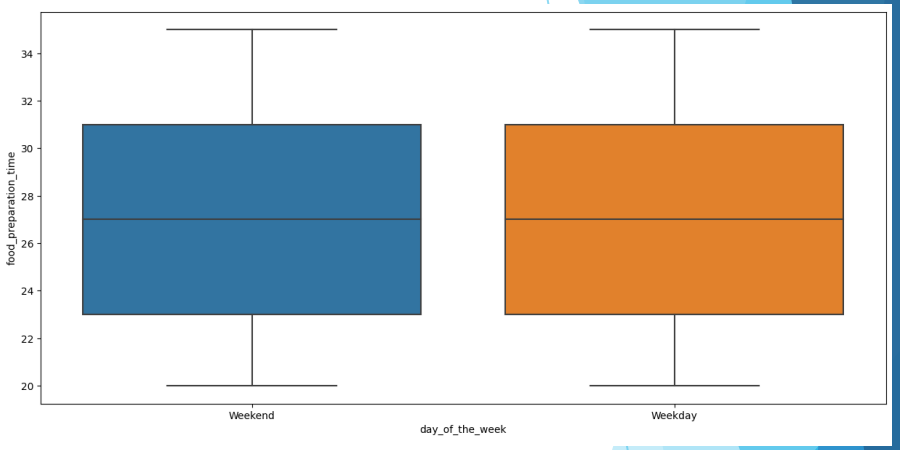
- Spanish cuisine has the highest variance in food preparation time
- Average food preparation time for all cuisine is 27 minutes
- Similar data skewness is seen in the Southern cuisine orders

Day of weeks vs Delivery time & Food prep time

Day of weeks vs Delivery Time



Day of weeks vs Food Preparation Time



Observations:

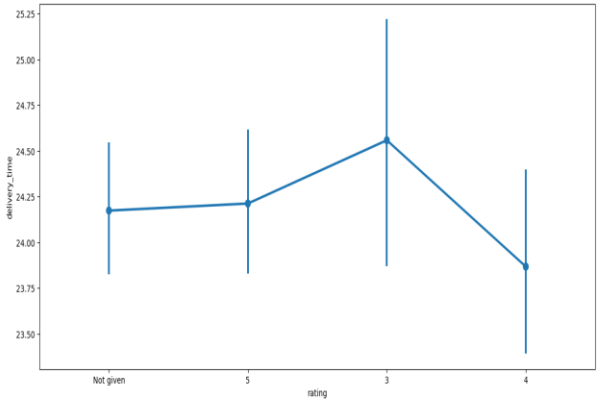
- It took longer to deliver the order during weekday compared to weekend.
- There is more food delivery completed during weekends than during weekdays because there is higher number of food orders placed during the weekends.

Observations:

- Food preparation time for the order to be completed is the same during weekends and weekdays.

Rating Multivariate Analysis

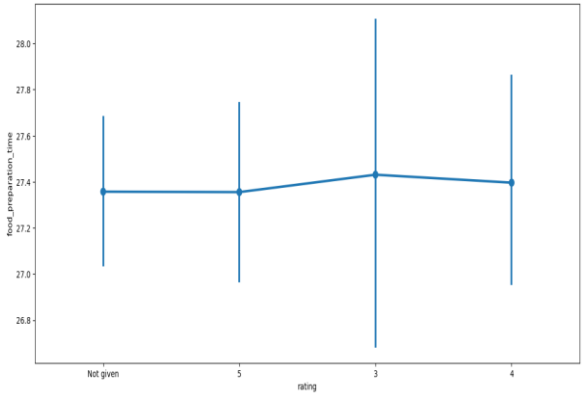
Rating vs Delivery Time



Observations:

- Customers gave the lowest rating when the average delivery time is longer than 24.5 mins

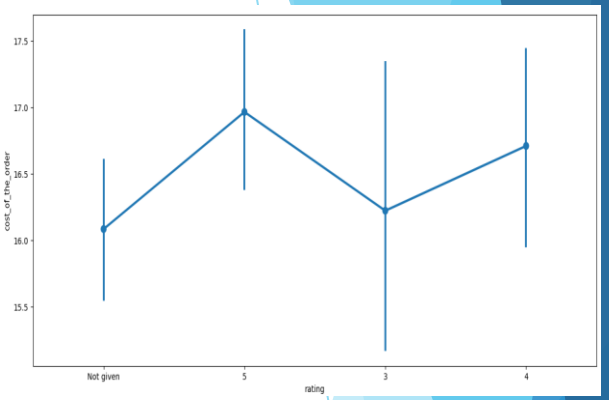
Rating vs Food Preparation Time



Observations:

- The mean values of the food preparation time for each rating are very much the same which indicates there is not much relationship or impact between customer rating and the time it took to prepare the food

Rating vs Cost of the order



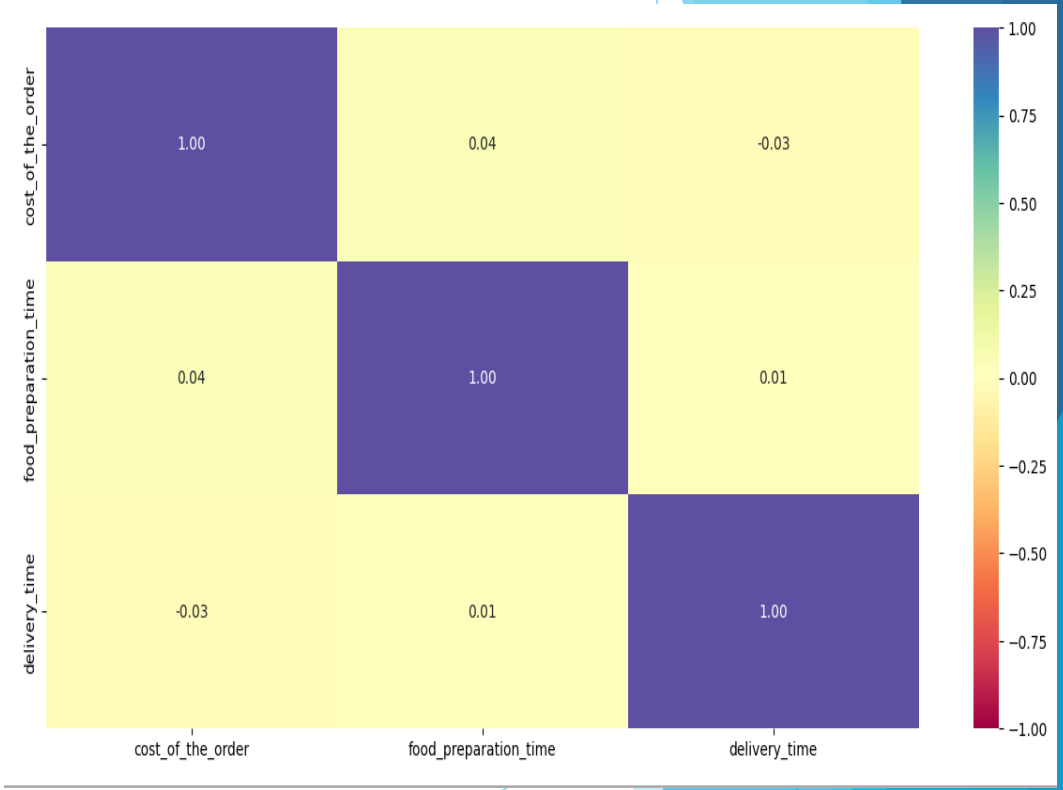
Observations:

- The order with consistently higher cost receive higher rating, it seems that order which cost more may be done with higher quality and get more customer preference and hence better rating

Correlation Matrix

Observations:

- Delivery time has positive correlation with food preparation time indicates as food preparation time goes up so will delivery time.
- Food preparation time has the highest positive correlation with cost of the order which makes sense, the longer to prepare the order means more labor hours for the restaurant hence it will cost more for the customer order
- Cost of the order has a negative correlation with the delivery time indicates as the cost of the order is higher, the delivery time is lower



APPENDIX

1. **The company charges the restaurant 25% on the orders having cost greater than 20 dollars and 15% on the orders having cost greater than 5 dollars. Find the net revenue generated by the company across all orders.**

	order_id	customer_id	restaurant_name	cuisine_type	cost_of_the_order	day_of_the_week	rating	food_preparation_time	delivery_time	Revenue
0	1477147	337525	Hangawi	Korean	30.75	Weekend	Not given	25	20	7.6875
1	1477685	358141	Blue Ribbon Sushi Izakaya	Japanese	12.08	Weekend	Not given	25	23	1.8120
2	1477070	66393	Cafe Habana	Mexican	12.23	Weekday	5	23	28	1.8345
3	1477334	106968	Blue Ribbon Fried Chicken	American	29.20	Weekend	3	25	15	7.3000
4	1478249	76942	Dirty Bird to Go	American	11.59	Weekday	4	25	24	1.7385

```
The net revenue is around 6166.3 dollars
```

2. **The company wants to analyze the delivery time of the orders on weekdays and weekends. How does the mean delivery time vary during weekdays and weekends?**

```
The mean delivery time on weekdays is around 28 minutes
The mean delivery time on Weekend is around 22 minutes
```

3. **The company wants to analyze the total time required to deliver the food. What percentage of orders take more than 60 minutes to get delivered from the time the order is placed? (The food has to be prepared and then delivered.)**

```
The number of total orders that have more than 60 minutes total delivery time is: 200
Percentage of orders that took more than 60 minutes total delivery time: 10.54
```

APPENDIX

4. The company wants to provide a promotional offer in the advertisement of the restaurants. The condition to get the offer is that the restaurants must have a rating count of more than 50 and the average rating should be greater than 4. Find the restaurants fulfilling the criteria to get the promotional offer

	restaurant_name	rating
0	The Meatball Shop	4.511905
1	Blue Ribbon Fried Chicken	4.328125
2	Shake Shack	4.278195
3	Blue Ribbon Sushi	4.219178