SMDM BUSINESS REPORT

FoodHub

11/5/2023

Context

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 smart phone 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.

Objective

The food aggregator company has stored the data of the different orders made by the registered customers in their online portal. They want to analyze the data to get a fair idea about the demand of different restaurants which will help them in enhancing their customer experience. Suppose you are a Data Scientist at FoodHub and the Data Science team has shared some of the key questions that need to be answered. Perform the data analysis to find answers to these questions that will help the company to improve the business.

Data Description

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

Data Dictionary

- 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 (The weekday is from Monday to Friday and the weekend is Saturday and Sunday)
- rating: Rating given by the customer out of 5
- food_preparation_time: Time (in minutes) taken by the restaurant to prepare the food. This is calculated by taking the difference between the timestamps of the restaurant's order confirmation and the delivery person's pick-up confirmation.
- delivery_time: Time (in minutes) taken by the delivery person to deliver the food package. This is calculated by taking the difference between the timestamps of the delivery person's pick-up confirmation and drop-off information.

Contents

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

3.5 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?

3.6 What are your conclusions from the analysis? What recommendations would you like to share to help improve the business? (You can use cuisine type and feedback ratings to drive your business recommendations.)

Data Analysis & A detailed solution around the problems:

Question 1: How many rows and columns are present in the data?
Answer: The data consists of 1898 rows and 9 columns.

Question 2: What are the data types of the different columns in the dataset?
Answer: The data is a combination of float, int., and object. Below is the chart for more clarity.

Observation:

1. Order Id, customer Id, food preparation time, delivery time are integers.
2. Restaurant name, cuisine type, day of the week, rating are objects.
3. Cost of the order is type float.
4. There are no missing/null values in the data.

Question 3: Are there any missing values in the data? If yes, treat them using an appropriate method.
Answer: It's very crucial to identify the missing values, and treat them before moving ahead with the data analysis, and as checked there are no missing values in the data, and we good to proceed. Below is the chart for the same.

Question 4: Check the statistical summary of the data. What is the minimum, average, and maximum time it takes for food to be prepared once an order is placed?
Answer: Statistical summary is nothing but the technique to identify variables such as min, max, mean, standard deviation, count, 25^{th}, 50^{th}, & 75^{th} percentile of the data.

Observation:

1. Minimum and maximum cost of order is 4.47 & 35.41 respectively.
2. Food preparation minimum time is 20 mins and maximum is 35mins.
3. Delivery time minimum has been 15 mins and maximum is 33 mins.

Question 5: How many orders are not rated?
Answer: The FoodHub has a mechanism of rating (out of 5), which helps them know which restaurant is the most rated, which is moderately rated, and which was not rated at all. Below is the figure for the same:

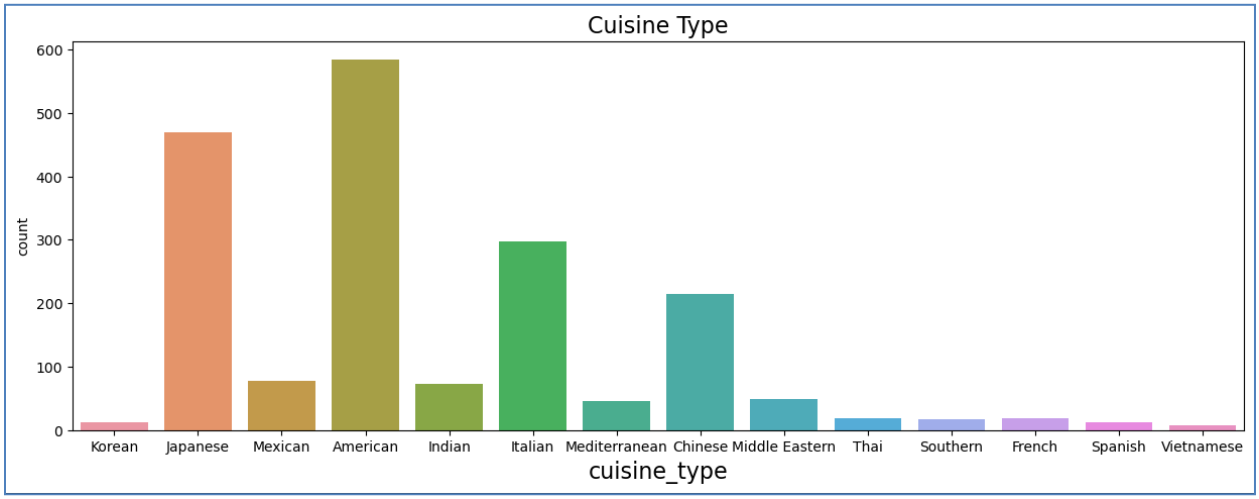
Observation: 736 orders are not rated out of all 1898 orders.

Univariate Analysis

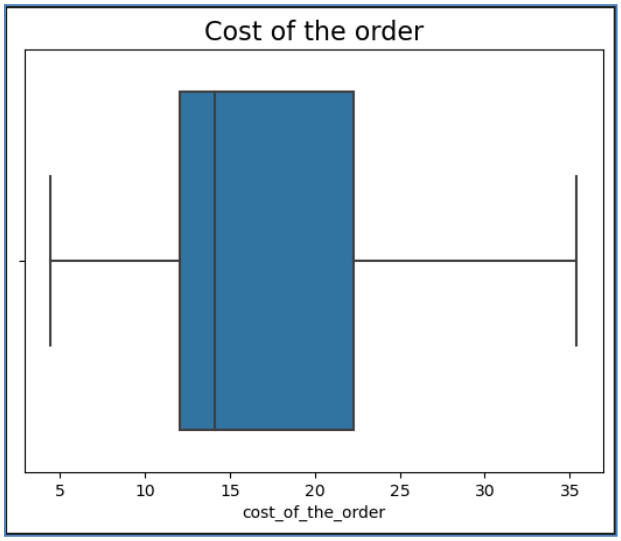
Question 6: Explore all the variables and provide observations on their distributions. (Generally, histograms, boxplots, countplots, etc. are used for univariate exploration.)

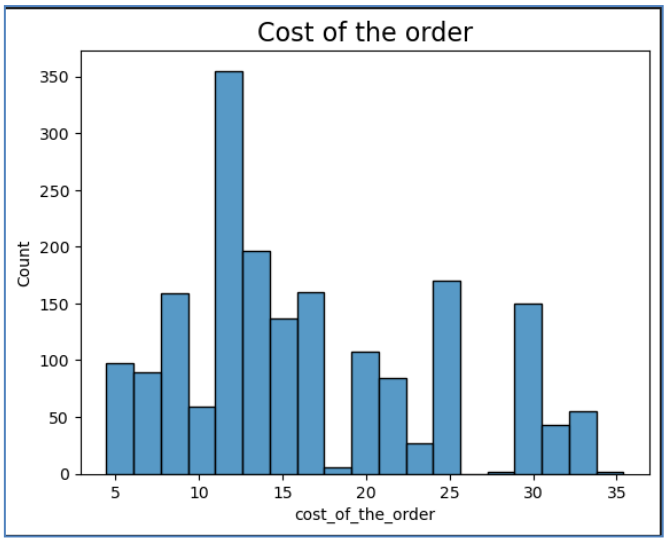
Unique Customer Id	1200
Total Unique Restaurants	178
Type of Cuisines	14
Days	Week & Weekend

6.1 **Cuisine Type:** The American food is the most popular, followed by Japanese and Italian. Vietnamese food is the least ordered.



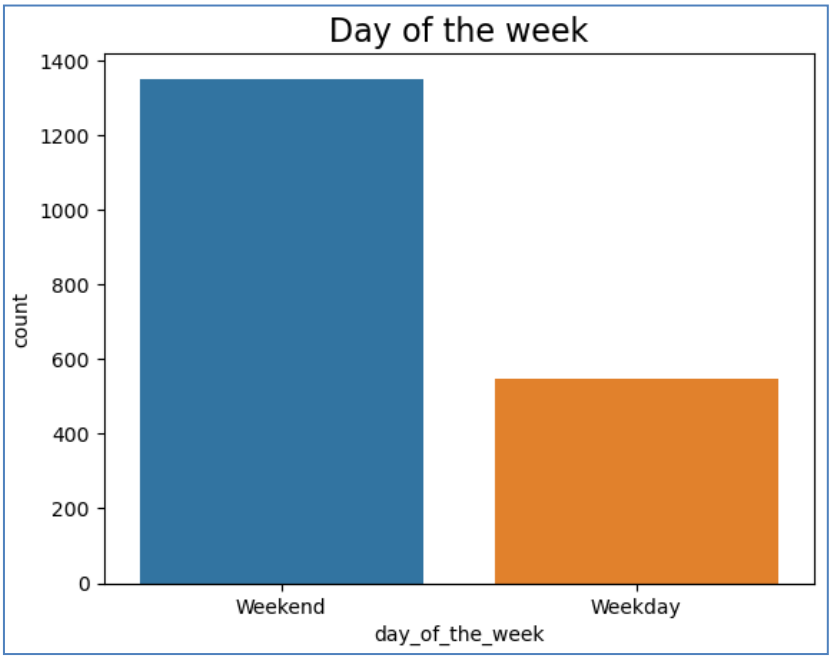
6.2 **Cost of the order:** As per the inference below most of the orders are placed around 12 dollars.



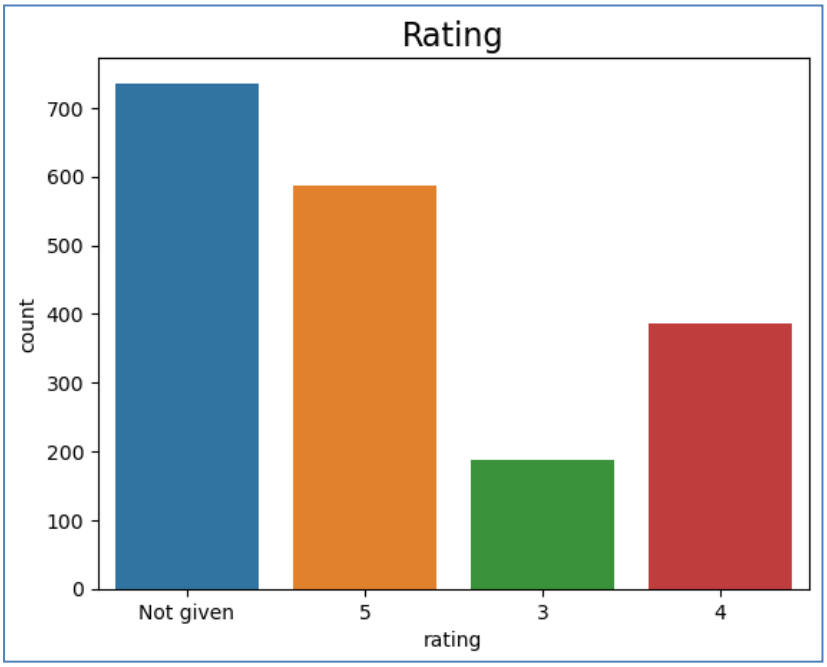


6.3 **Day of the week:** Maximum orders are placed on the weekend, with 1351 orders on weekend and 547 orders on weekdays.

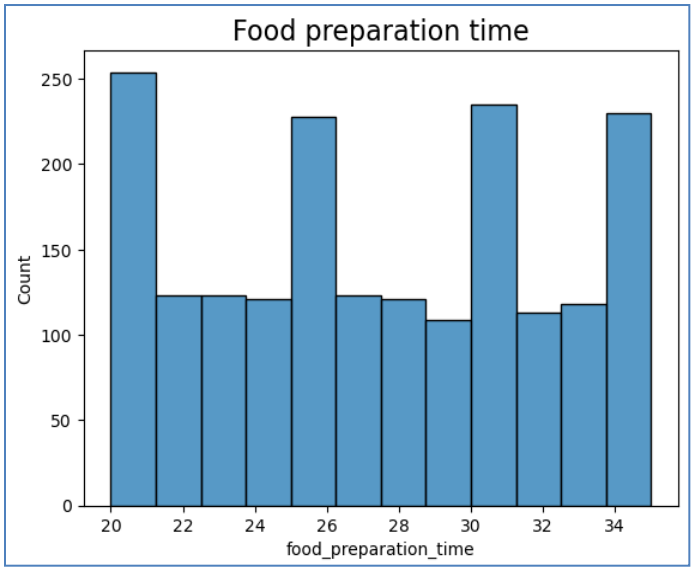
Day	Count of order_id
Weekday	547
Weekend	1351
Grand Total	**1898**

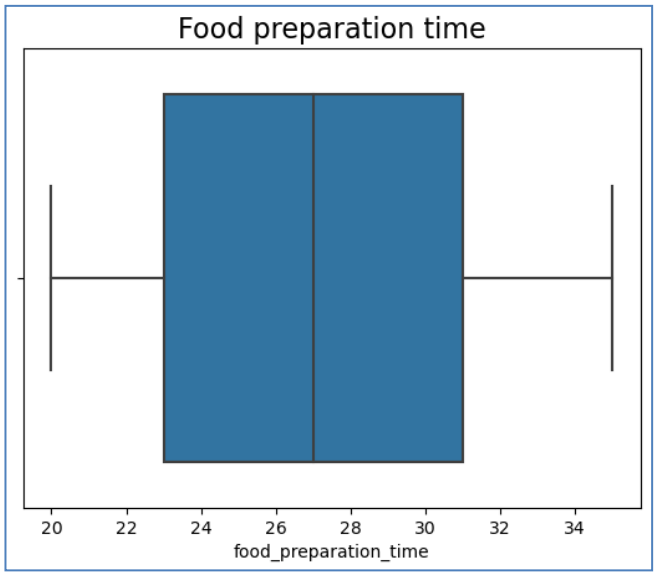


6.4 **Rating:** More than 700 orders did not receive any rating. Approximately 200 orders have rating of 3, 400 orders have rating of 4 and close to 600 orders have rating of 5.

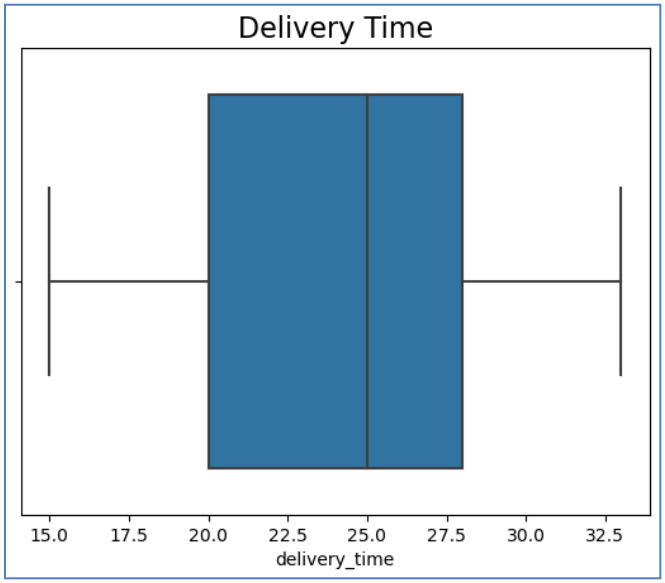


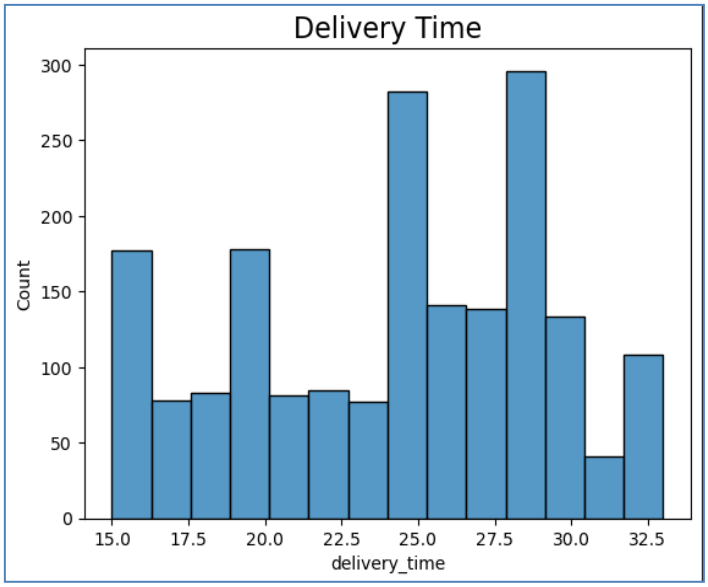
6.5 **Food preparation time:** A normally distributed data here, with no outliers, no skewness. Average food delivery time of 27 minutes observed.





6.6 **Delivery time:** Average delivery time of 24 minutes observed, with minimum of 15 minutes and maximum of 33 minutes.





Question 7: Which are the top 5 restaurants in terms of the number of orders received?
Answer: Shake shack, The Meatball Shop, Blue Ribbon Sushi, Blue Ribbon Fried Chicken & Parm are the top 5 restaurants to receive highest number of orders.

	index	restaurant_name
0	Shake Shack	219
1	The Meatball Shop	132
2	Blue Ribbon Sushi	119
3	Blue Ribbon Fried Chicken	96
4	Parm	68

Question 8: Which is the most popular cuisine on weekends?
Answer: American & Japanese cuisines are the most popular on weekends as compared to other cuisines.

	index	cuisine_type
0	American	415
1	Japanese	335
2	Italian	207
3	Chinese	163
4	Mexican	53
5	Indian	49
6	Mediterranean	32
7	Middle Eastern	32
8	Thai	15
9	French	13
10	Korean	11
11	Southern	11
12	Spanish	11
13	Vietnamese	4

Question 9: What percentage of the orders cost more than 20 dollars?
Answer: The number of total orders that cost above 20 dollars is 555. The percentage of the same is 29.24%.

Question 10: What is the mean order delivery time?
Answer: The average/mean delivery time of the order is 24.16 minutes.

Question 11: The Company has decided to give 20% discount vouchers to the top 3 most frequent customers. Find the IDs of these customers and the number of orders they placed.
Answer: Customer Id 52832, 47440 and 83287 are the top 3 customers who order frequently, hence will get 20% discount vouchers.

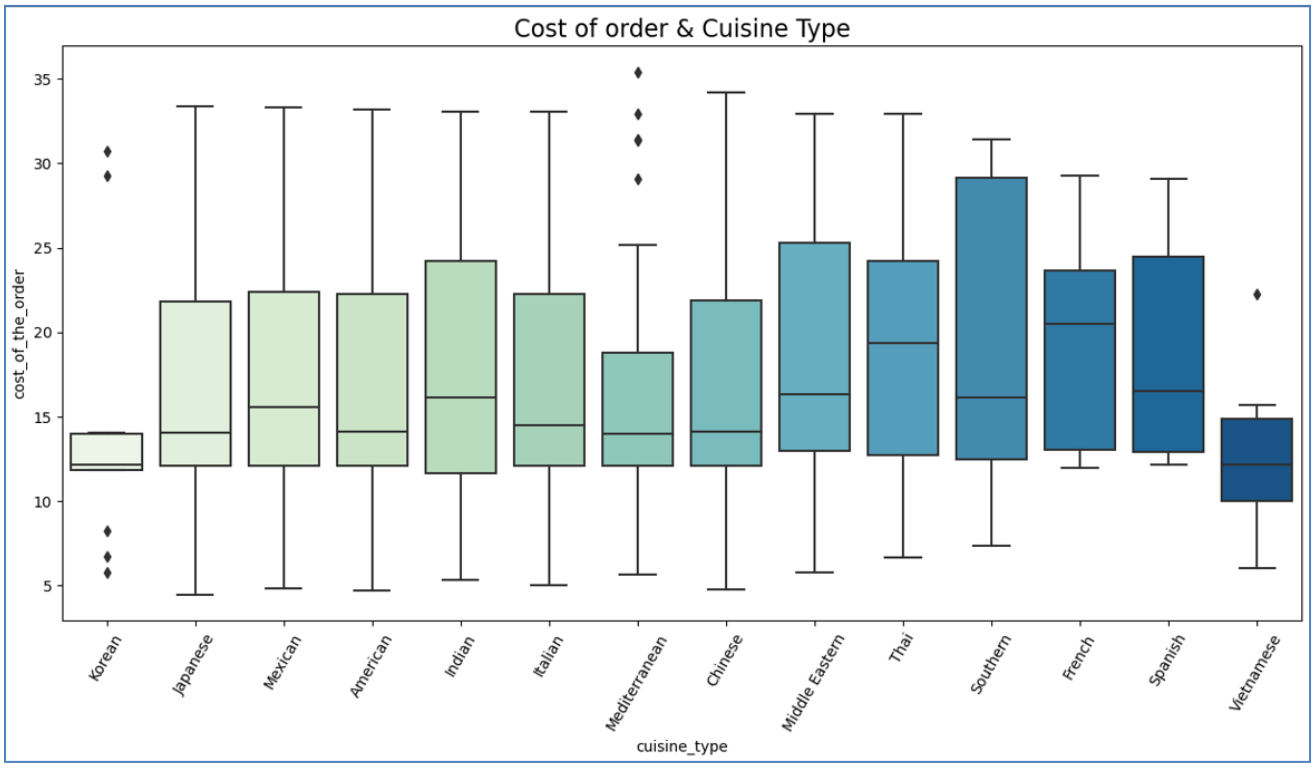
Multivariate Analysis

Question 12: Perform a multivariate analysis to explore relationships between the important variables in the dataset. (It is a good idea to explore relations between numerical variables as well as relations between numerical and categorical variables)
Answer:

12.1 **Cuisine v/s Cost of the order:**

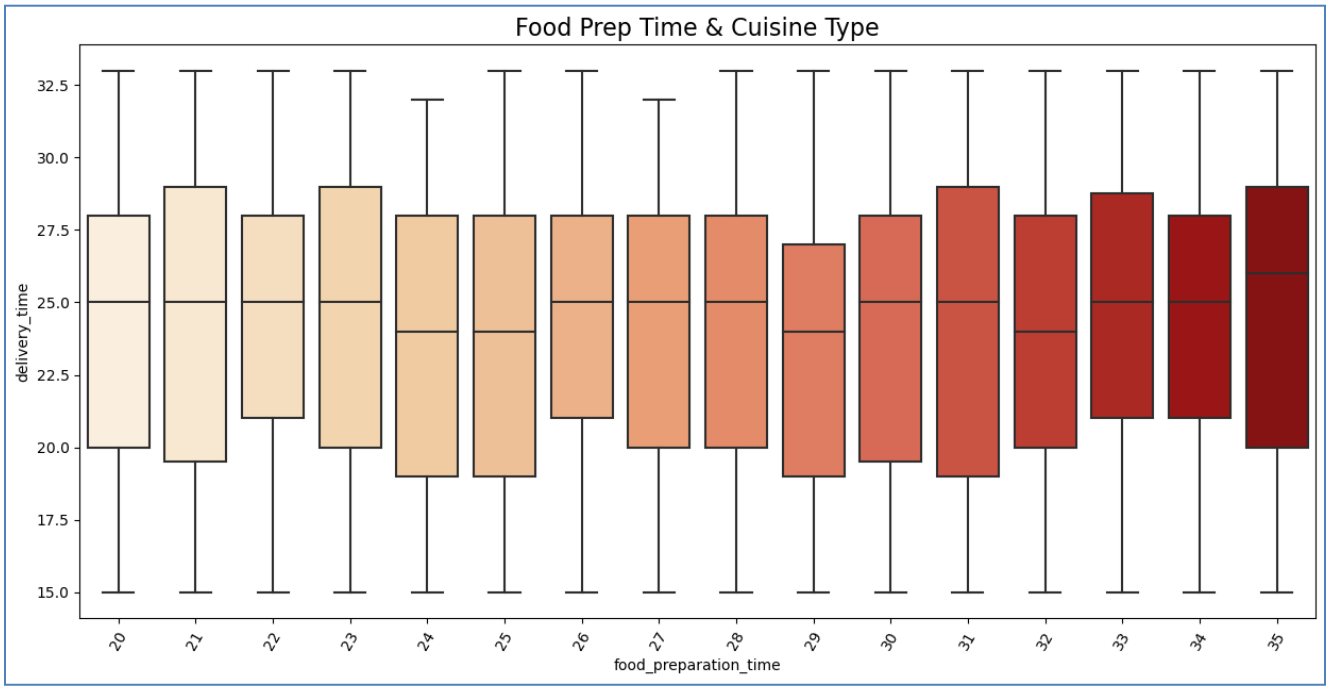
Observations:

1. American and Chinese cuisines have an average cost.
2. Mediterranean cuisine has outliers indicating towards expensive food but with no repetitions.
3. Vietnamese and Korean are the cheapest.



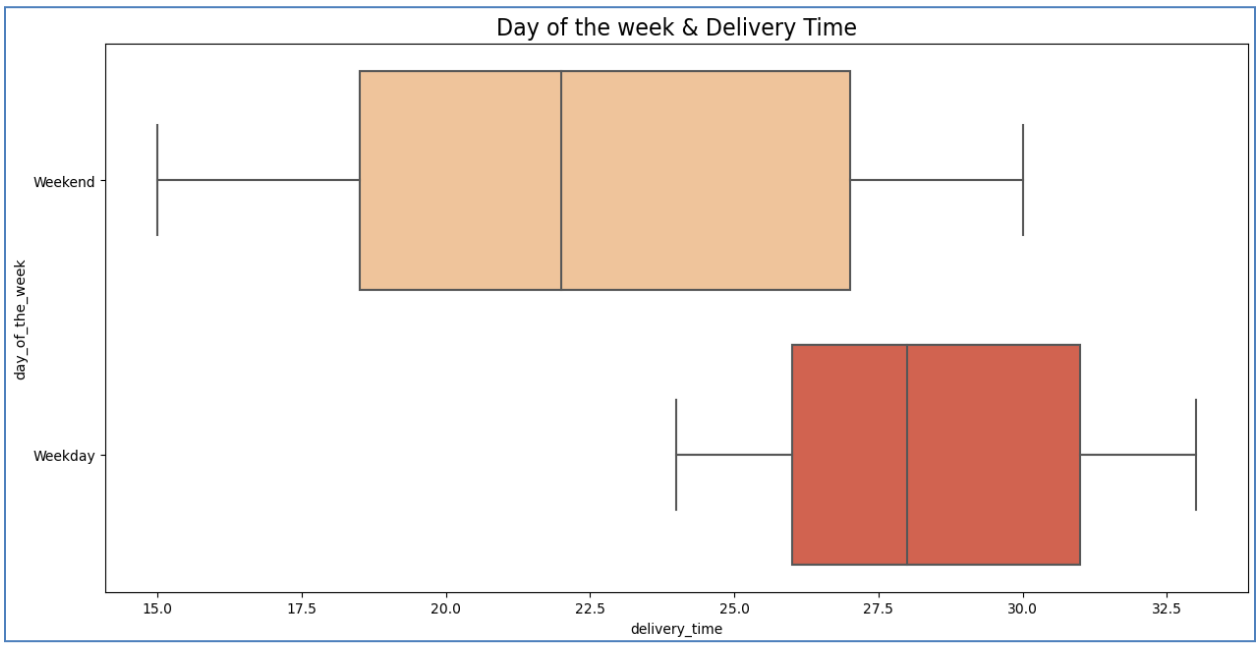
12.2 **Cuisine v/s Food preparation time:**

Observations: Most of the cuisine is prepared between 25 to 28 minutes.



12.3 **Day of the week v/s Delivery Time:**

Observation: Delivery on weekdays takes longer time as compared to on weekends.

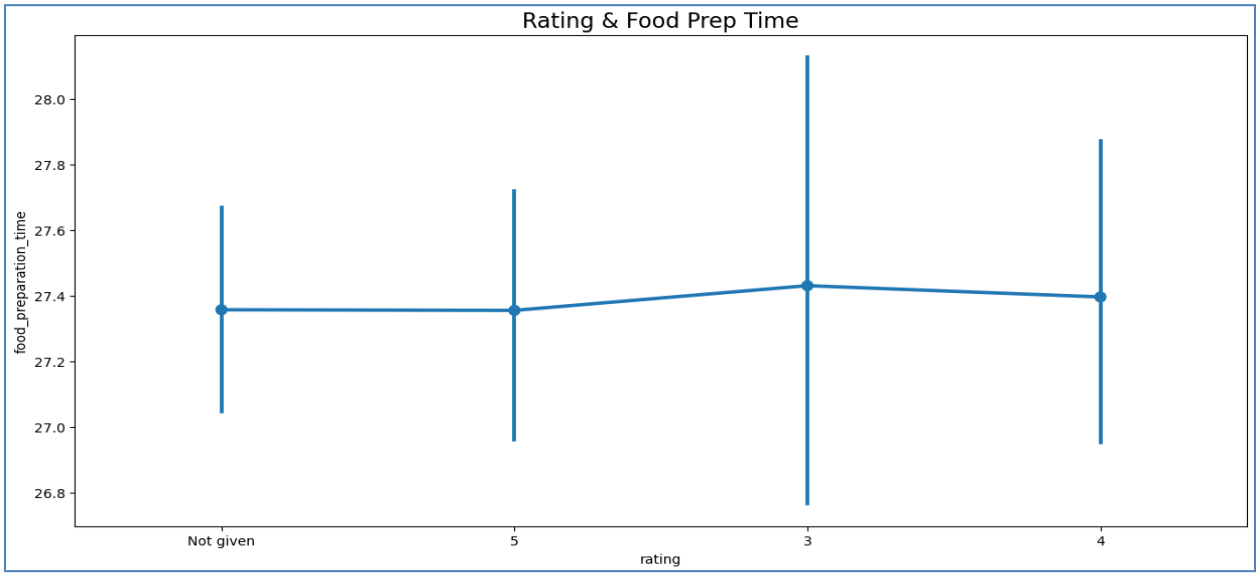


12.4 **Rating v/s Delivery Time:**

Observation: Evident decrease in rating, when delivery time is more.

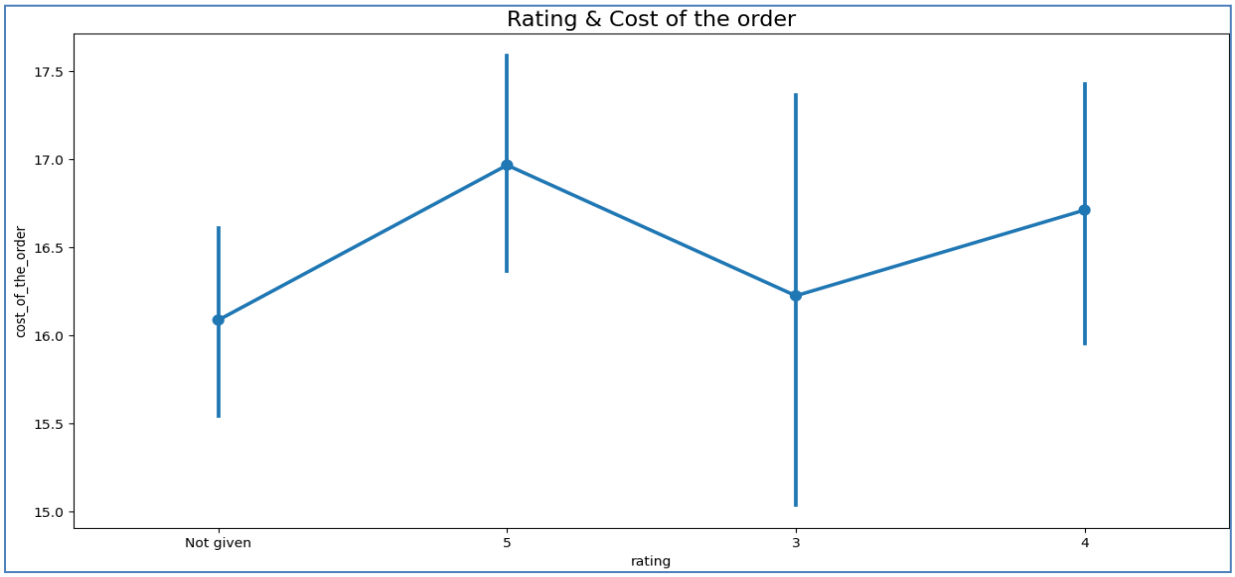
12.5 Rating v/s Food preparation time:

Observation: No evident movement or difference in rating in comparison with food preparation time.



12.6 Rating v/s Cost of the order:

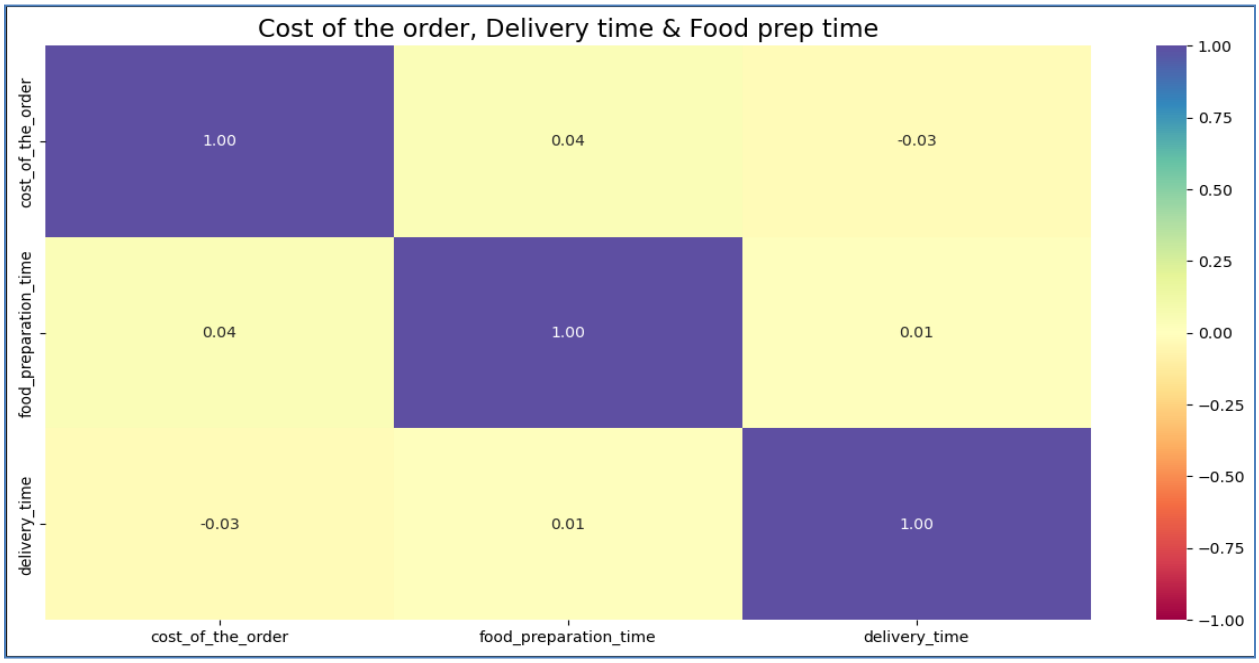
Observation: Highly rated order is also high in cost.



12.7 Correlation among variables:

Observations:

1. Cost of order and Delivery time has very less negative correlation.
2. Food preparation time and delivery time has very less positive correlation.
3. Food preparation time and cost of the order also have very less positive correlation.



Question 13: 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.

Answer: These below listed restaurants have rating count of 50 and their average rating is 4.

	index	restaurant_name	rating
0	136	Shake Shack	219.0
1	153	The Meatball Shop	132.0
2	21	Blue Ribbon Sushi	119.0
3	20	Blue Ribbon Fried Chicken	96.0
4	109	Parm	68.0
5	121	RedFarm Broadway	59.0
6	122	RedFarm Hudson	55.0

Question 14: 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.

Answer: The net revenue is around 6166.3 dollars.

	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

Question 15: The Company wants to analyze the total time required to deliver the food. What percentage of orders takes more than 60 minutes to get delivered from the time the order is placed? (The food has to be prepared and then delivered.)

Answer:

Observations:

1. The total no of orders that have more than 60 mins of total delivery time is 200.
2. Total no of orders is 1898
3. Percentage of orders takes more than 60 minutes to get delivered from the time the order is 10. 54%.

Question 16: 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?
Answer: The mean delivery time on weekdays is around 28 minutes.

Day of the week
Weekday 28.0
Weekend 22.0

Question 17: What are your conclusions from the analysis? What recommendations would you like to share to help improve the business? (You can use cuisine type and feedback ratings to drive your busi ness recommendations.)

Unique Customer Id	1200
Total Unique Restaurants	178
Type of Cuisines	14
Day	Week & Weekend
Rating	Out of 5 or Not Given
Food prep time	20-35mins
Delivery time	15-33mins
Cost	$4.47-$35.41

Answer:

1. Data comprises of 1898 rows and 9 columns.
2. Data consists of information about highly rated restaurants, famous cuisines, no. of orders received on weekdays or weekends, along with food preparation time & delivery time.

Given below is the following conclusion after detailed analysis on the data:

1. There are total 178 restaurants with Shake shack being the most popular.
2. Highly popular cuisines are American, Japanese, followed by Italian and Chinese.
3. Minimum cost of order is 4.47 dollars, and maximum is 35.41 dollars.
4. Food preparation time minimum is 20 minutes and maximum is 35 minutes.

5. Average of food preparation time is 27 mins.
6. Minimum delivery time is 15 minutes and maximum is 33 minutes.
7. Average delivery time is 24 mins.
8. Most of the orders were placed on weekends.
9. Total 736 orders out of 1898 were not rated.
10. Net revenue generated by company is 6166.3 dollars.

Recommendations:

1. Rating needs more marketing strategy that motivates consumers to rate, as 736 orders were not rated, more ratings will help improve better.
2. Weekday orders were less; this also needs attention and offerings in terms of discounts and vouchers.
3. Attractive initiatives not only around festive but also around regular days can be implemented.