INDEX	
S.No	Topic
1	Cover Page
2	Index
3	List of Tables
4	List of Figures
5	Problem Statement
6	Basic Steps
7	Order Analysis
8	Customer Behaviour
9	Restaurant Performance
10	Demand Patterns
11	Operation Efficiency
12	Customer Insights
13	Conclusion

List of Tables

List of figures

Problem Statement

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 hired as a Data Scientist in this company 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 improve its business

Data 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: 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

Basic steps

1.Display the top 5 rows

Table 1: Showing the top 5 rows

	0	1	2	3	4
order_id	1477147	1477685	1477070	1477334	1478249
customer_id	337525	358141	66393	106968	76942
restaurant_name	Hangawi	Blue Ribbon Sushi Izakaya	Cafe Habana	Blue Ribbon Fried Chicken	Dirty Bird to Go
cuisine_type	Korean	Japanese	Mexican	American	American
cost_of_the_order	30.75	12.08	12.23	29.2	11.59
day_of_the_week	Weekend	Weekend	Weekday	Weekend	Weekday
rating	Not given	Not given	5	3	4
food_preparation_time	25.0	25.0	23.0	25.0	25.0
delivery_time	20.0	25.0	28.0	15.0	24.0

2.Display the last 5 rows

Table 2: Showing the bottom 5 rows

	1893	1894	1895	1896	1897
order_id	1476701	1477421	1477819	1477513	1478056
customer_id	292602	397537	35309	64151	120353
restaurant_name	Chipotle Mexican Grill $1.99 Delivery	The Smile	Blue Ribbon Sushi	Jack's Wife Freda	Blue Ribbon Sushi
cuisine_type	Mexican	American	Japanese	Mediterranean	Japanese
cost_of_the_order	22.31	12.18	25.22	12.18	19.45
day_of_the_week	Weekend	Weekend	Weekday	Weekday	Weekend
rating	5	5	Not given	5	Not given
food_preparation_time	31.0	31.0	31.0	23.0	28.0
delivery_time	17.0	19.0	24.0	31.0	24.0

3. Check the shape of dataset

```
(1898, 9)
```

4.Check the datatypes of each features

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1898 entries, 0 to 1897
Data columns (total 9 columns):
 #   Column                 Non-Null Count  Dtype  
---  ------                 --------------  -----  
 0   order_id               1898 non-null   int64  
 1   customer_id            1898 non-null   int64  
 2   restaurant_name        1898 non-null   object 
 3   cuisine_type           1895 non-null   object 
 4   cost_of_the_order      1898 non-null   float64
 5   day_of_the_week        1898 non-null   object 
 6   rating                 1898 non-null   object 
 7   food_preparation_time  1896 non-null   float64
 8   delivery_time          1898 non-null   object 
dtypes: float64(2), int64(2), object(5)
memory usage: 133.6+ KB
```

Observations

- Datatype of delivery_time should be numerical but it is object instead. Hence, it needs to be checked

5.Check the statstical summary

Table 3: Showing the statistical summary

	order_id	customer_id	cost_of_the_order	food_preparation_time	delivery_time
count	1.898000e+03	1898.000000	1898.000000	1898.000000	1898.000000
mean	1.477496e+06	171168.478398	17.439145	27.371444	24.163330
std	5.480497e+02	113698.139743	41.823984	4.631783	4.972638
min	1.476547e+06	1311.000000	0.000000	20.000000	15.000000
25%	1.477021e+06	77787.750000	12.080000	23.000000	20.000000
50%	1.477496e+06	128600.000000	14.160000	27.000000	25.000000
75%	1.477970e+06	270525.000000	22.310000	31.000000	28.000000
max	1.478444e+06	405334.000000	1809.126912	35.000000	33.000000

6.Check the null values

```
order_id                  0
customer_id               0
restaurant_name           0
cuisine_type              3
cost_of_the_order         0
day_of_the_week           0
rating                    0
food_preparation_time     2
delivery_time             0
dtype: int64
```

7.Check the duplicate values

```
0
```

Observations

- There are no duplicate values

8.Check the anomolies or wrong entries

```
restaurant_name
Shake Shack                  219
The Meatball Shop            132
Blue Ribbon Sushi            119
Blue Ribbon Fried Chicken     96
Parm                          68
                             ...
Sushi Choshi                   1
Dos Caminos Soho               1
La Follia                      1
Philippe Chow                  1
'wichcraft                     1
Name: count, Length: 178, dtype: int64
```

```
cuisine_type
American          582
Japanese          470
Italian           298
Chinese           215
Mexican            76
Indian             73
Middle Eastern     49
Mediterranean      46
Thai               19
French             18
Southern           17
Korean             13
Spanish            12
Vietnamese          7
Name: count, dtype: int64
```

```
cost_of_the_order
12.18    86
12.13    81
12.23    47
24.20    42
29.10    37
         ..
6.26      1
9.61      1
4.47      1
15.04     1
29.59     1
Name: count, Length: 314, dtype: int64
```

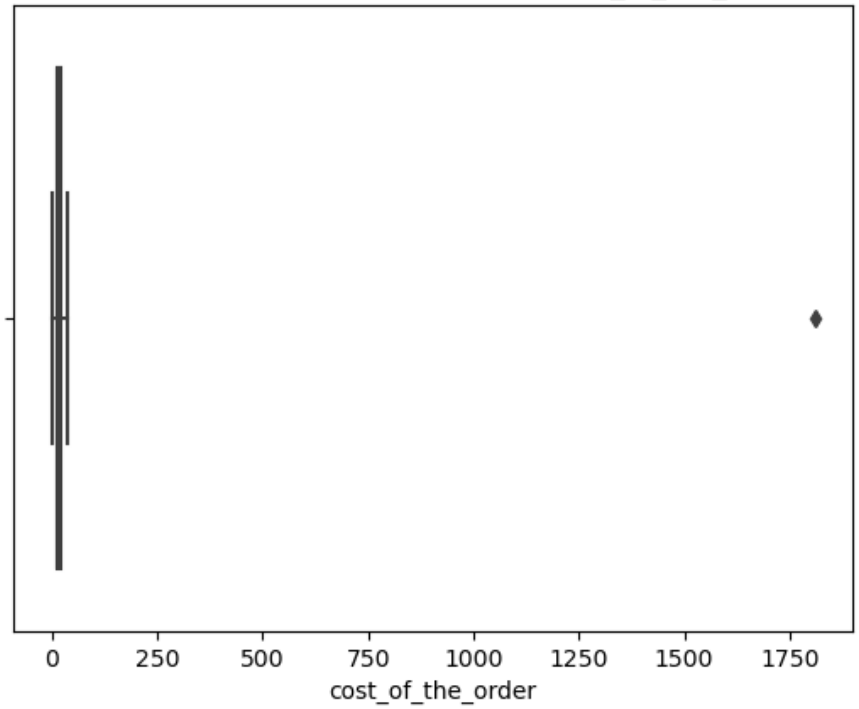
```
day_of_the_week
Weekend    1351
Weekday     547
Name: count, dtype: int64
```

```
food_preparation_time
21.0    135
23.0    123
27.0    123
22.0    123
28.0    121
24.0    121
20.0    119
33.0    118
30.0    118
35.0    117
31.0    116
26.0    115
34.0    113
32.0    113
25.0    112
29.0    109
Name: count, dtype: int64
```

```
delivery_time
24    161
28    148
29    148
26    141
27    138
30    133
25    120
19     90
16     90
20     88
15     87
22     85
18     83
21     81
17     78
23     76
32     59
33     49
31     41
?       2
Name: count, dtype: int64
```

9.Check the outliers and their authenticity

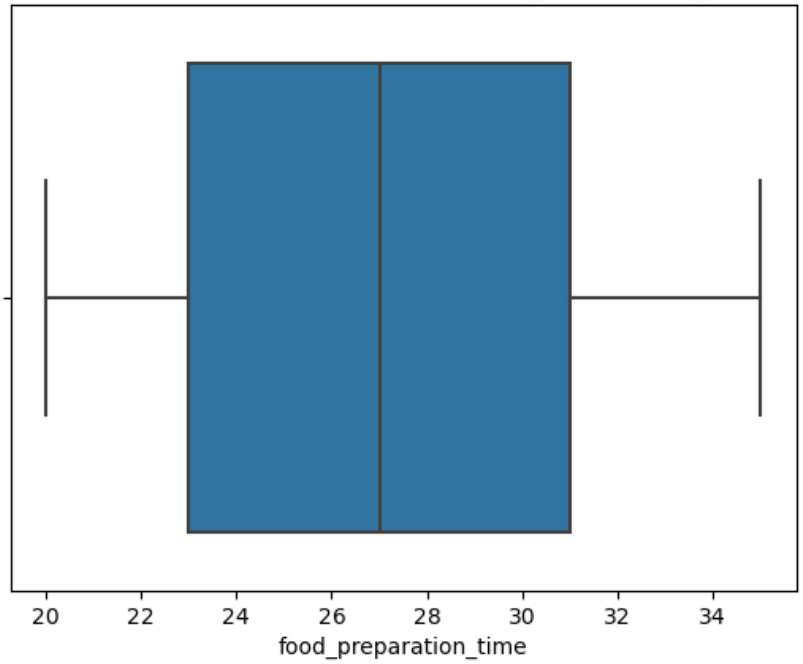
1.1 Boxplot showing distrbution of data in cost_of_the_order column



Observations

- There is one outlier in the 'cost_of_the_order' column which is order_id 214 with the value of 12192.0

1.2 Boxplot showing data distribution in food_preparation_time column



Observations

- There is no outlier in 'food_preparation_time' column

10.Do the necessary data cleaning steps like dropping duplicates, unnecessary columns, null value imputation, outliers treatment etc.

All values in column delivery time which was '?' replaced with null value

Note:- The outlier in cost_of_the_order column has been replaced with mean of the cost_of_the_order column where the restaurant name is same as the outlier's one

1.Order Analysis

- What is the total number of orders in the dataset?
- What is the average cost of an order?
- How many unique customers have placed orders?
- Which restaurant has received the highest number of orders?

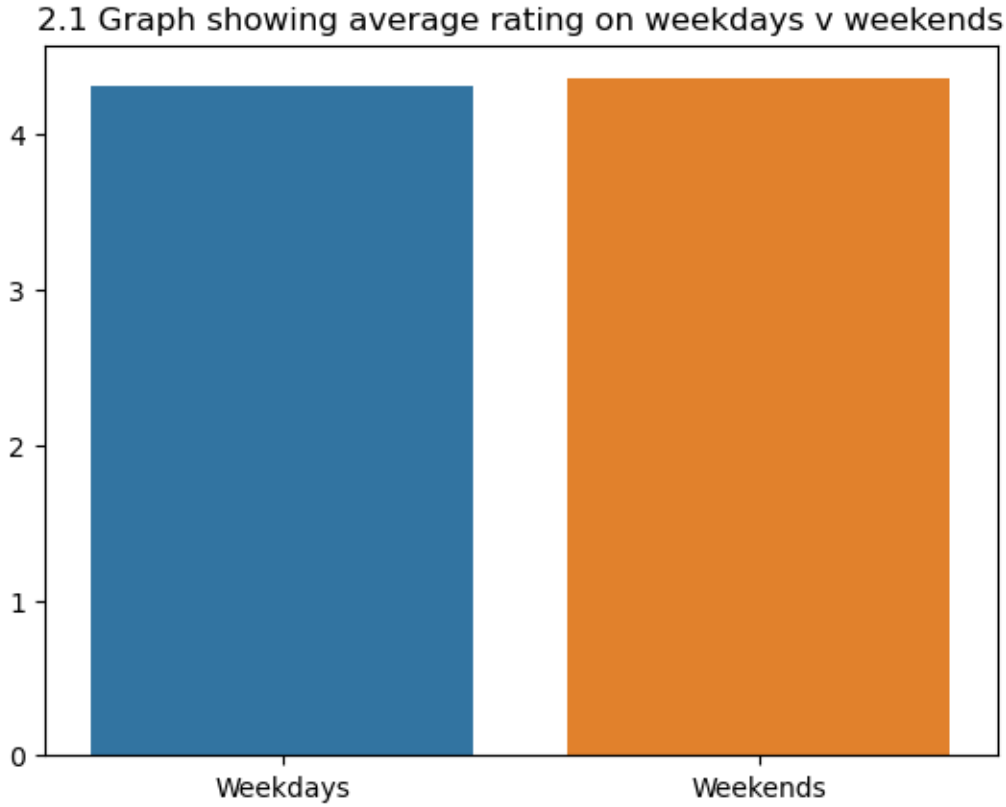
Answers of Order Analysis

- There are 1898 orders in the dataset
- The average cost of an order is 17.439144842868654
- 1200 unique customers have placed orders
- Restuarant Shake Shack has the highest number of orders with 219

Customer Behaviour

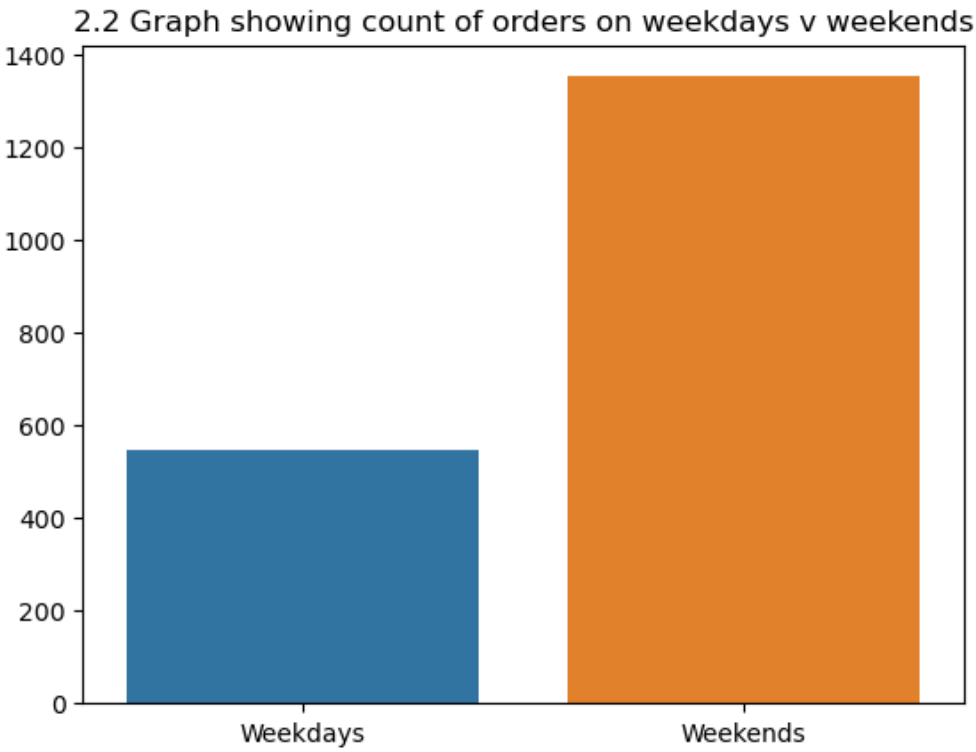
- What is the average rating given by customers?

- How does the rating vary between weekdays and weekends?
- Which cuisine type is ordered the most?
- What is the distribution of orders across different days of the week?



Observations

- The average rating on weekday is 4.3088235294117645 while the average rating on weekend is 4.358880778588808



Observations

- The distribution of orders on weekdays is to weekend is 547 to 1351

Answers of Customer behaviour

- The average rating given by consumers is 4.344234079173838
- The average rating on weekday is 4.3088235294117645 while the average rating on weekend is 4.358880778588808
- The cuisine type ordered the most is American with 582 Orders
- The distribution of orders on weekdays is to weekend is 547 to 1351

3.Restaurant performance

- What is the average food preparation time for each restaurant?
- Which restaurant has the shortest average food preparation time?
- How does the average delivery time compare across different restaurants?
- Is there a correlation between the cost of the order and the rating given?

Answers of Restaurant performance

```
Table 4:Showing the average food preparation time for each restaurant
Top 5
                  food_preparation_time
restaurant_name
'wichcraft                         28.0
12 Chairs                          27.0
5 Napkin Burger                    30.2
67 Burger                          20.0
Alidoro                            34.0
Bottom 5
                   food_preparation_time
restaurant_name
Zero Otto Nove                 30.000000
brgr                           25.000000
da Umberto                     24.333333
ilili Restaurant               26.388889
indikitch                      30.750000
```

- The restaurant with the shortest average food preparation time id 67 Burger with average preparation time of 20.0

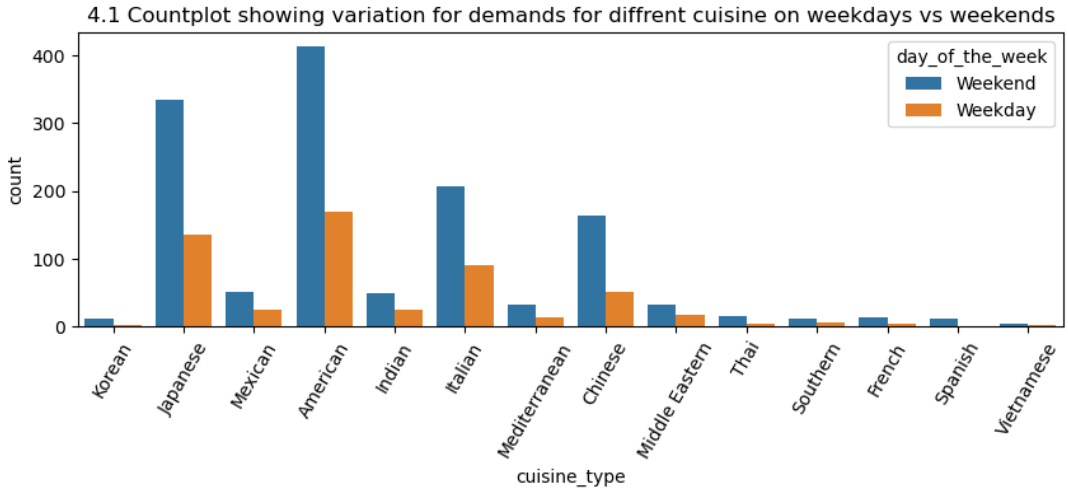
```
table 5:The average delivery time across diffrent restaurant is
Top 5
                    delivery_time
restaurant_name
Zero Otto Nove          21.500000
brgr                    25.000000
da Umberto              28.000000
ilili Restaurant        24.888889
indikitch               25.500000
Bottom 5
                    delivery_time
restaurant_name
Zero Otto Nove          21.500000
brgr                    25.000000
da Umberto              28.000000
ilili Restaurant        24.888889
indikitch               25.500000
```

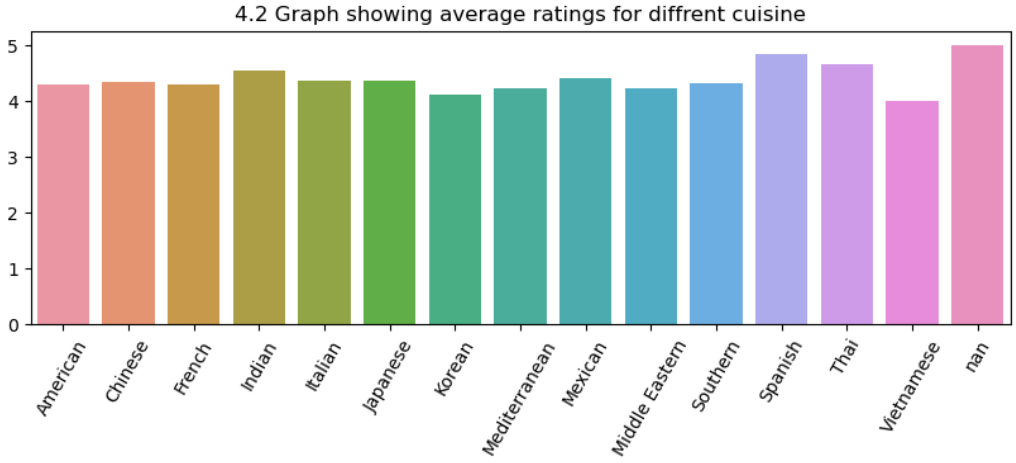
- The correlation coefficient of -0.04797741071835822 represents a very slight negative association (closed to negligible) between cost_of_the_order and rating

4.Demand Patterns

- How does the demand for different cuisine types vary on weekdays versus weekends?
- How does the average rating vary by cuisine type?

Answers of Demand Patterns





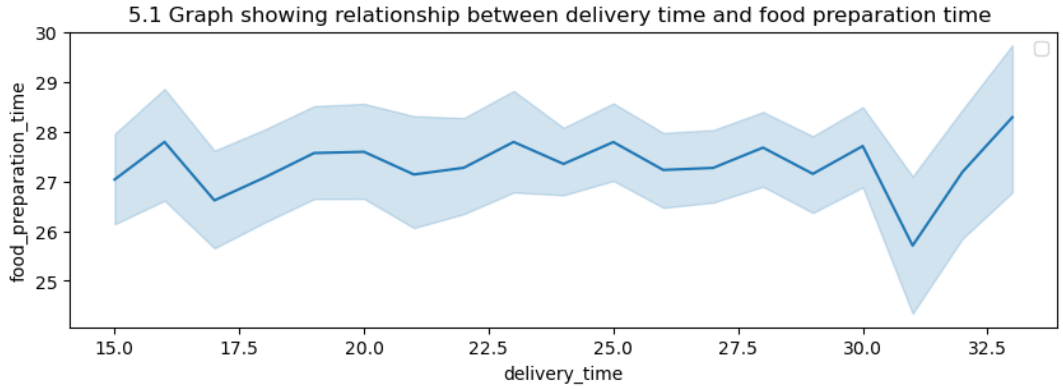
5.Operational efficiency

- What is the average delivery time for all orders?
- Which restaurant has the longest average delivery time?
- Is there a relationship between food preparation time and delivery time?
- How does the delivery time impact customer ratings?

Answers of Operational efficiency

- The average delivery time for all orders is 24.163329820864067

- The restaurant with longest average delivery time of 33.0 is Sarabeth's West



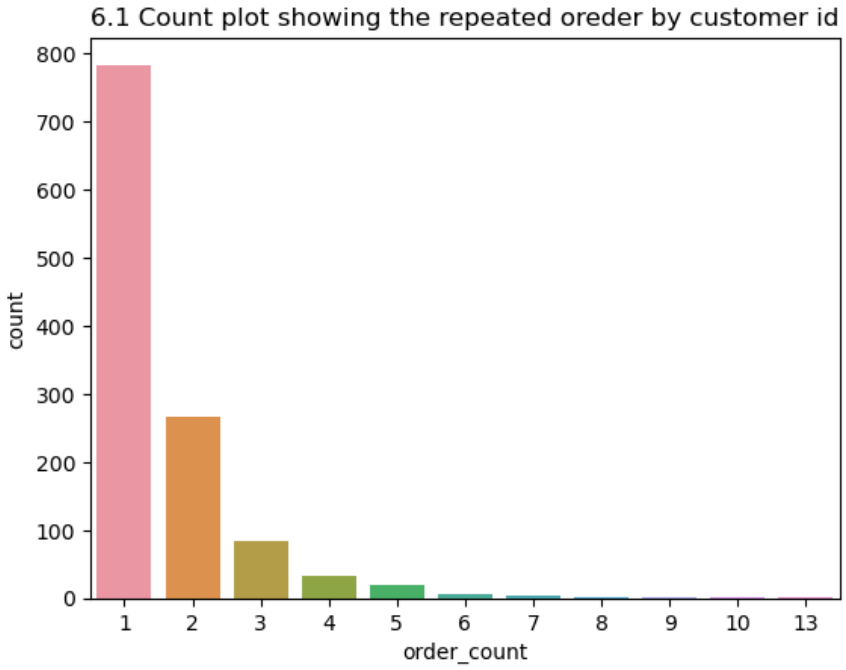
- The correlation coefficient is -0.010120801797414725

Observations

- There is very low (close to negligible) negative relatiosnhip between delivery time and rating

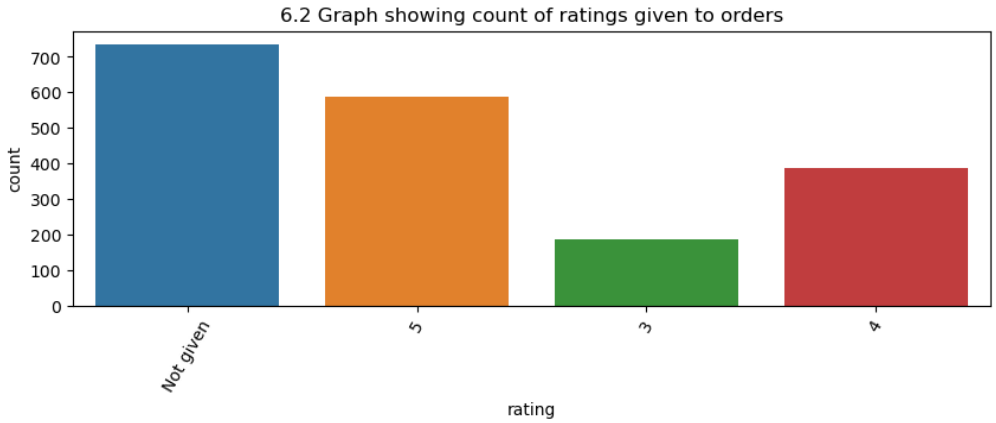
6.Customer Insights

- What is the repeat order rate (number of customers who have placed more than one order)?
- What percentage of orders receive a rating of 4 or higher?



Observations

- From the above graph the number of customers with repeated orders is 416 while total customers is 1200.Hence,repeated order rate is 416/1200*100 which is 34.67%



Observations

- The number of orders with rating of 4 or higher is 974 which is 974/1898*100 i.around 51.7%

Answers of Customer Heights

- Customers with repeated order rate is 34.67%
- The rate of orders with rating of 4 or higher is 51.7%

Conclusion

The data analysis on the food aggregator company's orders provides valuable insights that can help improve customer experience and optimize business operations. Here are the key findings and recommendations based on the analysis:

- Identifying the most popular restaurants and cuisine types can help the company prioritize partnerships and promotional activities. The restaurants with the highest number of orders indicate customer preference and demand.
- Monitoring the cost trends over time, particularly across weekdays and weekends, can aid in dynamic pricing strategies to maximize revenue
- Weekends typically show higher order volumes, suggesting the need for increased operational capacity during these times.
- Examining the food preparation and delivery times helps identify bottlenecks in the order fulfillment process. Restaurants with long preparation times might need operational improvements.

By leveraging these insights and implementing the recommended strategies, the food aggregator company can enhance customer experience, improve operational efficiency, and ultimately drive business growth. Continuous monitoring and analysis of order data will be crucial to adapt to changing customer preferences and market conditions.