Summer Bootcamp Project Report 2024
Food Hub

Basharat Hassan

Index

List of Tables

Tables	
S No.	Table Description
1	The first 5 rows of the dataset.
2	The last 5 rows of the dataset.
3	Info of the dataset.
4	The statistical summary of the dataset.
5	Number fo null value per column.
6	Checking for unique values for devlivery time and rating column.
7	Checking for unique values for delivery time.
8	Checking value counts for deliver time column.
9	Showing the rows of delivery time is ?.
10	Checking for the unique values in delivery time column.
11	Checking info for the dataset.
12	Checking the rows which have atleat one entry null.
13	Counting the null value count per column.
14	Counting the null value count percentage per column.
15	Checking for the info of the cost of the order column.

List of Figures

Figures	
S No.	Figure Description
1	Boxplot figure to check outliers in the numerical features.
2	Boxplot figure food preparation time.
3	Boxplot figure food delivery time.
4	Countplot figure between Cuisine Type and value counts.
5	Boxplot figure to check outliers in cost of order feature.
6	Boxplot figure to verify the removal of outliers in cost of order column
7	Barplot figure between Days fo the week and Distribution.
8	Barplot figure between Restaurants and average food preparation.
9	Barplot figure between Restaurant and Average Delviery time.
11	Scatterplot figure between Cost of order and Rating.
12	Barplot figure between Cuisine type and Number of orders.
13	Barplot figure between Days fo the week and Number of orders.
14	Barplot figure between Cuisine type and Rating.
15	Barplot figure between Restaurant and average delviery time.
16	Scatterplot figure between Food preparation and Delivery time.
17	Scatterplot figure between Deliver time and Customer rating.
18	Pie figure between Multiple orders and Single Order.
19	Countplot figure between Rating and number of orders

Problem Statement / 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 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 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: 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

Importing the necessary Libraries

Loading the dataset

Basic Steps:

1- First Five 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	24.162447	28.0	15.0	24.0

Table No. 1

Observations

- There is some wrong entry in delivery_time column, will have to fix later
- There is a Not given entries in rating column, will have to check later

2 - Last Five 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

Table No. 2

Observations

- Here also there are some not given entries in column rating which we need to fix later

3 - Shape of dataset

Observations

- There are 1898 rows and 9 columns in the dataset

4 - Check datatypes of each feature

```
<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
```

Table No. 3

Observations

- There are 5 null entries in any column
- There are 2 float, 2 int and 5 object datatyes in the dataset
- We know that order_id and custormer_id are categories but they are stored as int which is alright for now
- Rating column should be stored as int but it is object here so there must be some string anamoly or not given entry, need to further explore it
- delivery time is appearing as object but it should be numerical column so we need to fix later

5 - Statistical summary

	order_id	customer_id	cost_of_the_order	food_preparation_time
count	1.898000e+03	1898.000000	1898.000000	1896.000000
mean	1.477496e+06	171168.478398	80.722007	27.371835
std	5.480497e+02	113698.139743	2798.141333	4.634211
min	1.476547e+06	1311.000000	0.000000	20.000000
25%	1.477021e+06	77787.750000	12.080000	23.000000
50%	1.477496e+06	128600.000000	14.160000	27.000000
75%	1.477970e+06	270525.000000	22.310000	31.000000
max	1.478444e+06	405334.000000	121920.000000	35.000000

Table No. 4

Observations

- minimum value of cost order can not be 0, so will have to check later
- maximum value fo cost of order is much higher, it is an outlier need to check later

6 - Check for 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
```

Table No. 5

Observations

- There are 5 null entries in the dataset, 3 in cuisine_type and 2 in food_preparation_time, will have to check later

7 - Duplicate values

Observations

- There are no duplicated rows in the dataset

8 - Check for anomalies or wrong entries

```
for deliver_time column: ['20' '?' '28' '15' '24' '21' '30' '26' '22' '17' '23' '25' '16' '29' '27'
 '18' '31' '32' '19' '33']
for rating column: ['Not given' '5' '3' '4']
```

Table No.6

Observations

- in deliver_time column we have a string "?" entry instead of a numeric so there is an anomaly here
- as for rating column there is an entry Not given in some rows so will have to check later

9 - Check for outliers and their authenticity

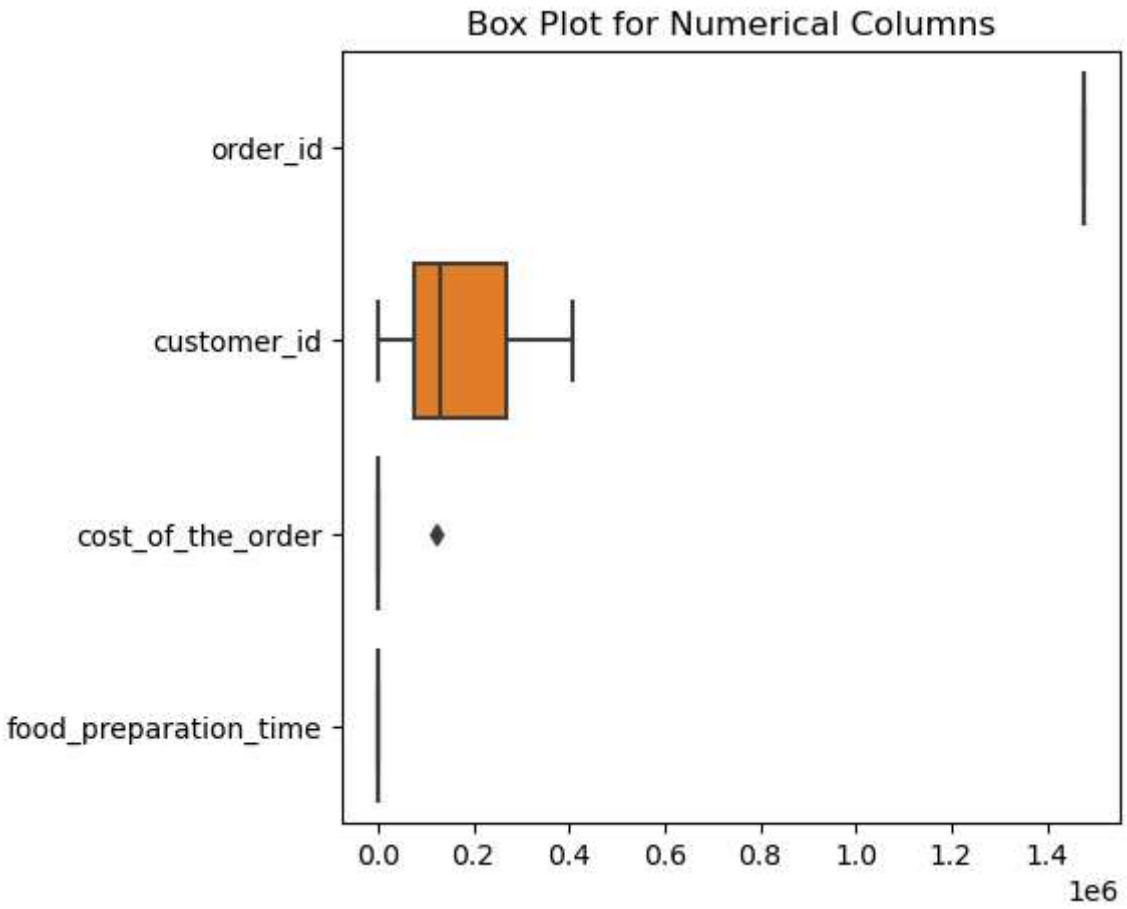


Fig No. 1

Observations

- after observing the graphs it is seen that only cost_of_the_order column has an outlier which we need to fix later

10 - Data Cleaning

1 - Correcting data types

Observations

- we will convert deliver_time from object to float later when we remove and replace the invalid entry

2 - Duplicate values

Observations

- as there are no duplicated rows so no need to drop rows

3 - Invalid values / Wrong Entries

```
array(['20', '?', '28', '15', '24', '21', '30', '26', '22', '17', '23',
       '25', '16', '29', '27', '18', '31', '32', '19', '33'], dtype=object)
```

Table No.7

- Getting count of each unique entry

```
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
```

Table No. 8

- getting rows with 'delivery time' as '?'

	order_id	customer_id	restaurant_name	cuisine_type	cost_of_the_order	day_of_the_week	rating	food_preparation_time	delivery_time
1	1477685	358141	Blue Ribbon Sushi Izakaya	Japanese	12.08	Weekend	Not given	25.0	?
180	1476808	84700	Pepe Giallo	Italian	14.60	Weekday	3	32.0	?

Table No. 9

- replacing the '?' with null value

```
array(['20', nan, '28', '15', '24', '21', '30', '26', '22', '17', '23',
       '25', '16', '29', '27', '18', '31', '32', '19', '33'], dtype=object)
```

Table No. 10

- Now we can change datatype of this column from object to numeric

```
<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          1896 non-null   float64
dtypes: float64(3), int64(2), object(4)
memory usage: 133.6+ KB
```

Table No. 11

4 - Missing values / Null values

	order_id	customer_id	restaurant_name	cuisine_type	cost_of_the_order	day_of_the_week	rating	food_preparation_time	delivery_time
1	1477685	358141	Blue Ribbon Sushi Izakaya	Japanese	12.08	Weekend	Not given	25.0	NaN
11	1478437	221206	Empanada Mama (closed)	NaN	8.10	Weekend	5	23.0	22.0
51	1477883	91817	Blue Ribbon Fried Chicken	NaN	29.39	Weekend	Not given	27.0	28.0
95	1477027	164016	Blue Ribbon Fried Chicken	NaN	16.39	Weekend	Not given	27.0	22.0
140	1477376	370372	Blue Ribbon Fried Chicken	American	11.59	Weekday	Not given	NaN	24.0
180	1476808	84700	Pepe Giallo	Italian	14.60	Weekday	3	32.0	NaN
188	1477872	300670	Shake Shack	American	13.39	Weekend	Not given	NaN	22.0

Table No. 12

- check for missing value in columns

```
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           2
dtype: int64
```

Table No. 13

- check for percentage wise missing values in columns

```
order_id                0.000000
customer_id             0.000000
restaurant_name         0.000000
cuisine_type            0.158061
cost_of_the_order       0.000000
day_of_the_week         0.000000
rating                  0.000000
food_preparation_time   0.105374
delivery_time           0.105374
dtype: float64
```

Table No. 14

- firstly checking for outliers in numeric columns where there are null values

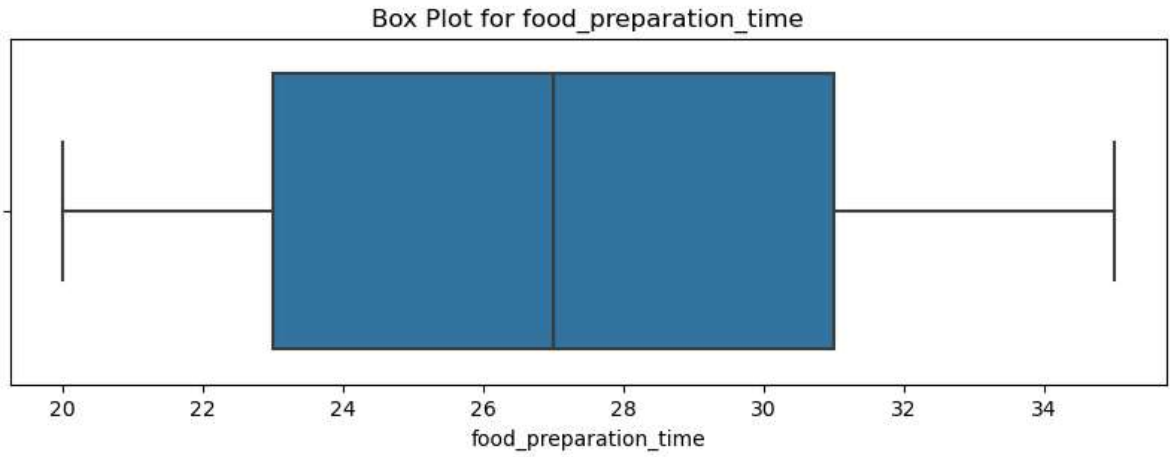


Fig No. 2

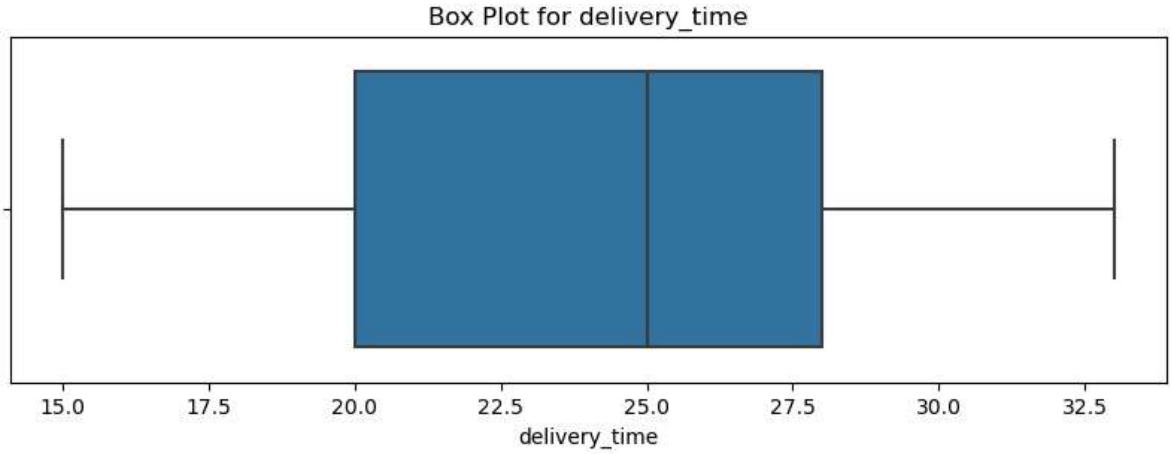


Fig No. 3

Observations

- as we can see there are no outliers in these 2 numeric columns where there are null entries

Now let us check for outliers in categorical column which has null entries

- Understanding the distribution

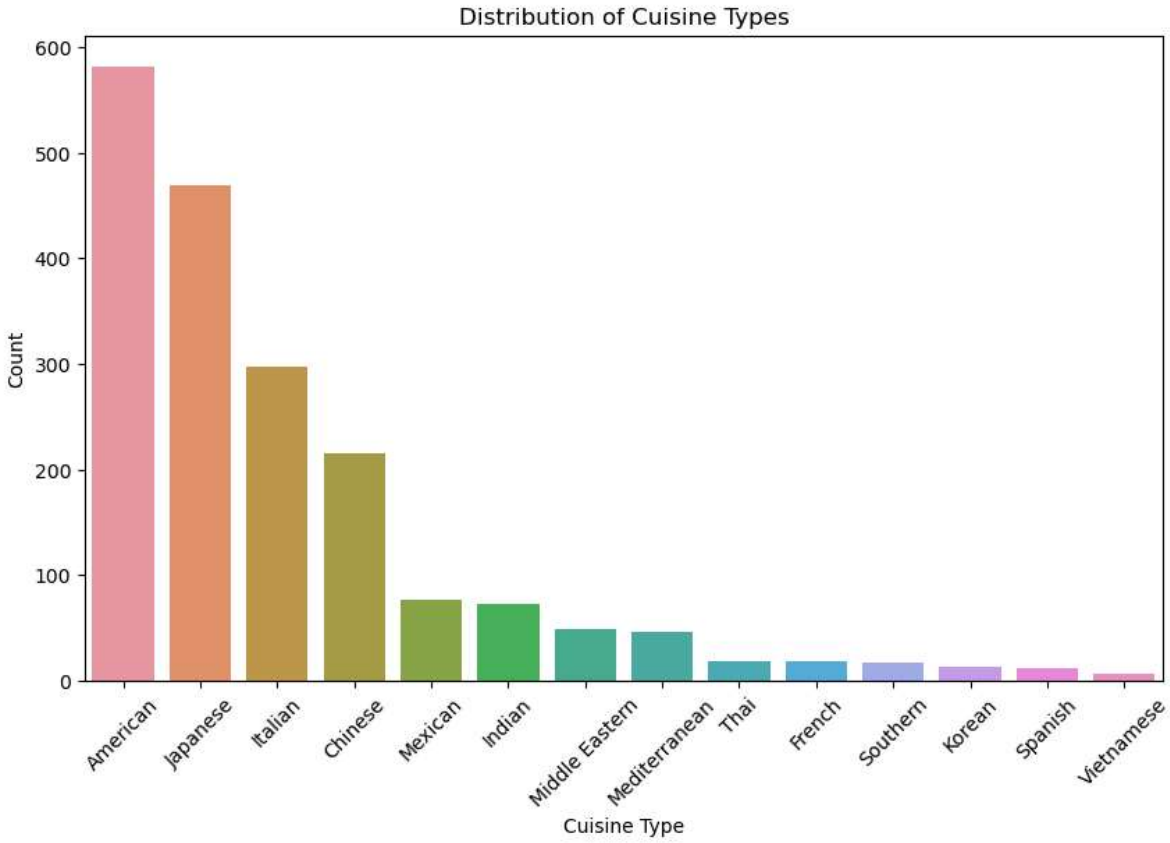


Fig No. 4

Observation

- here we can see american cuisine is the most ordered and vietnamese is least orderd

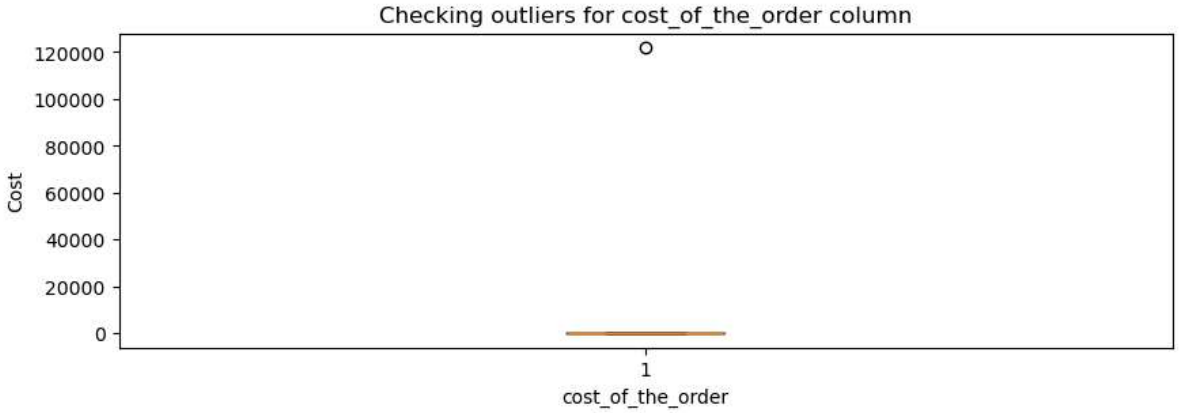


Fig No. 5

```
count    1898.000000
mean       16.505809
std         7.507834
min         0.000000
25%        12.080000
50%        14.160000
75%        22.310000
max        37.655000
Name: cost_of_the_order, dtype: float64
```

Table No.15

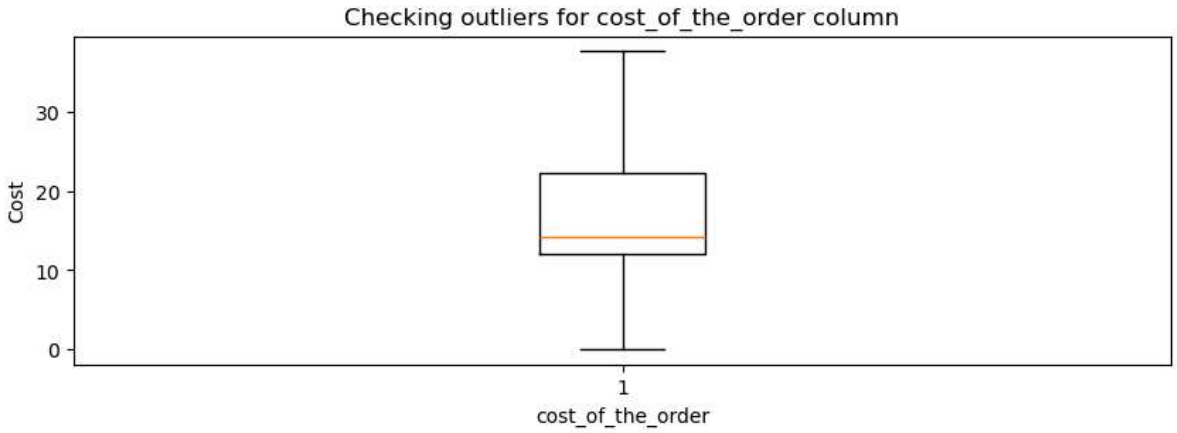


Fig No. 6

Observation

- checking for null values in my dataset and we have found that null percentage in are null entries in 3 columns

- Now verifying the changes

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

Table No. 16

Observation

- After completing my data cleaning let us check for null values we can see that all null entries were removed now we can perform further analysis

We have a rating column where Not given entry occurs we can not work with that so we fix it first

- Now well do for rating column

```
array(['Not given', '5', '3', '4'], dtype=object)
```

Table No. 17

```
rating
Not given    736
5            588
4            386
3            188
Name: count, dtype: int64
```

Table No. 18

- Here checking rows in datset where rating entry is not given

	order_id	customer_id	restaurant_name	cuisine_type	cost_of_the_order	day_of_the_week	rating	food_preparation_time	delivery_time
0	1477147	337525	Hangawi	Korean	30.75	Weekend	Not given	25.0	20.000000
1	1477685	358141	Blue Ribbon Sushi Izakaya	Japanese	12.08	Weekend	Not given	25.0	24.162447
6	1477894	157711	The Meatball Shop	Italian	6.07	Weekend	Not given	28.0	21.000000
10	1477895	143926	Big Wong Restaurant □_¤¾Ñ¼	Chinese	5.92	Weekday	Not given	34.0	28.000000
14	1478198	62667	Lucky's Famous Burgers	American	12.13	Weekday	Not given	23.0	30.000000
...	...	...	...	...	...	...	...	...	...
1887	1476873	237616	Shake Shack	American	5.82	Weekend	Not given	26.0	30.000000
1891	1476981	138586	Shake Shack	American	5.82	Weekend	Not given	22.0	28.000000
1892	1477473	97838	Han Dynasty	Chinese	29.15	Weekend	Not given	29.0	21.000000
1895	1477819	35309	Blue Ribbon Sushi	Japanese	25.22	Weekday	Not given	31.0	24.000000
1897	1478056	120353	Blue Ribbon Sushi	Japanese	19.45	Weekend	Not given	28.0	24.000000

736 rows × 9 columns

Table No. 19

Observation

- Here we have Not given entry in our rating column which occurs 736 times so we will have to extract a new dataframe from it where only rows which have a valid rating entries is given to perform analysis on questions where rating is required

- Now lets us print filtered_df info

```
<class 'pandas.core.frame.DataFrame'>
Index: 1162 entries, 2 to 1896
Data columns (total 9 columns):
 #   Column                 Non-Null Count  Dtype  
---  ------                 --------------  -----  
 0   order_id               1162 non-null   int64  
 1   customer_id            1162 non-null   int64  
 2   restaurant_name        1162 non-null   object 
 3   cuisine_type           1162 non-null   object 
 4   cost_of_the_order      1162 non-null   float64
 5   day_of_the_week        1162 non-null   object 
 6   rating                 1162 non-null   float64
 7   food_preparation_time  1162 non-null   float64
 8   delivery_time          1162 non-null   float64
dtypes: float64(4), int64(2), object(3)
memory usage: 90.8+ KB
```

Table No. 20

1. Order Analysis

• What is the total number of orders in the dataset?

Observation

- The total number of orders in the data set are: 1898

• What is the average cost of an order?

Observation

- The average cost of order is: 16.5058

• How many unique customers have placed orders?

Observation

- Number of unique customers are: 1200

• Which restaurant has received the highest number of orders?

Observation

- The restaurant with the highest number of orders is: Shake Shack

2. Customer Behavior

• What is the average rating given by customers?

Observation

- The average rating by customers is: 4.344234079173837

• How does the rating vary between weekdays and weekends?

Observation

- Average rating on Weekdays: 4.32
- Average rating on Weekends: 4.35

• Which cuisine type is ordered the most?

Obervation

- most ordered cuisine is: American

• What is the distribution of orders across different days of the week?

Fig No. 7

Observation

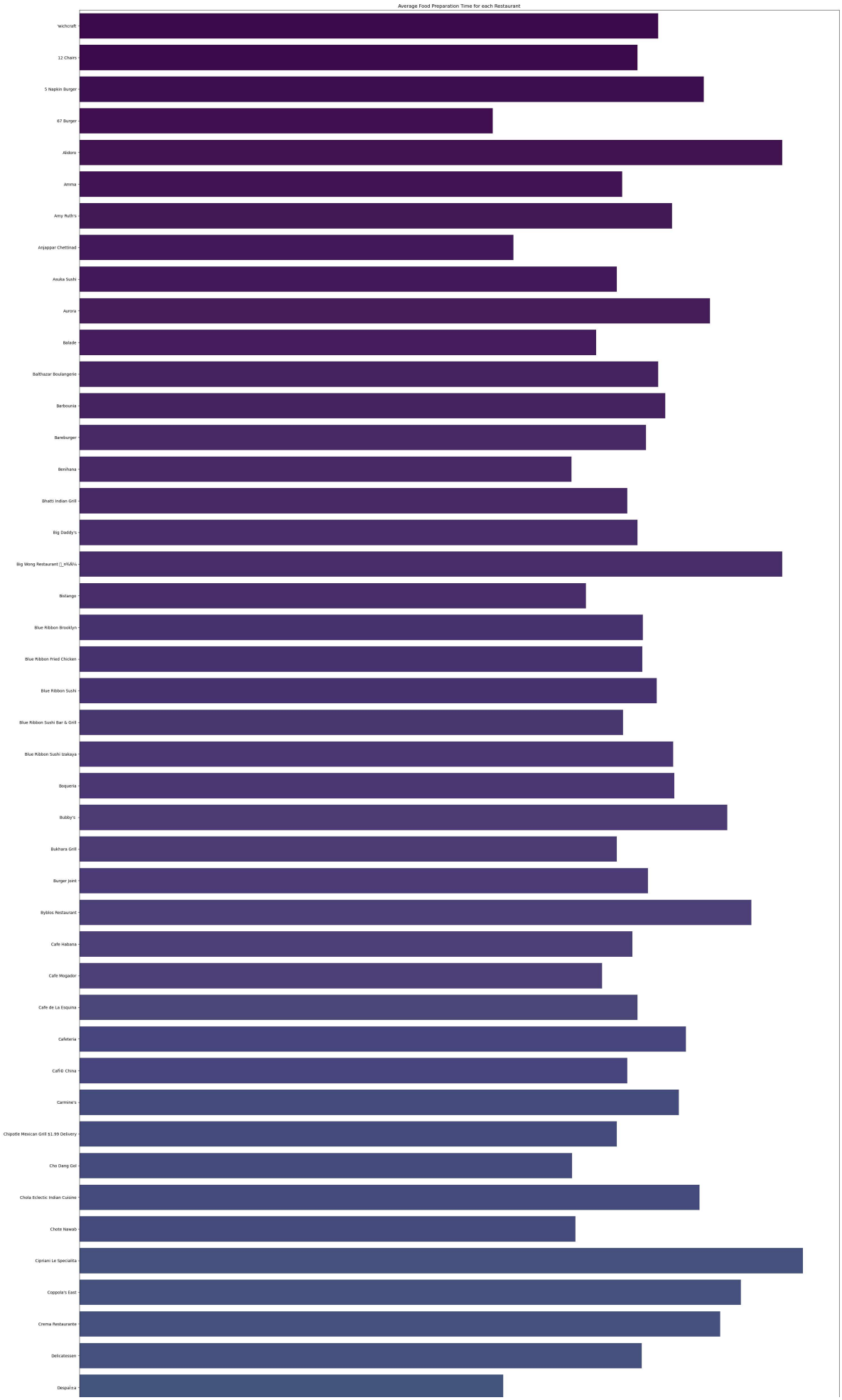
- day_of_the_week
- Weekend 1351
- Weekday 547

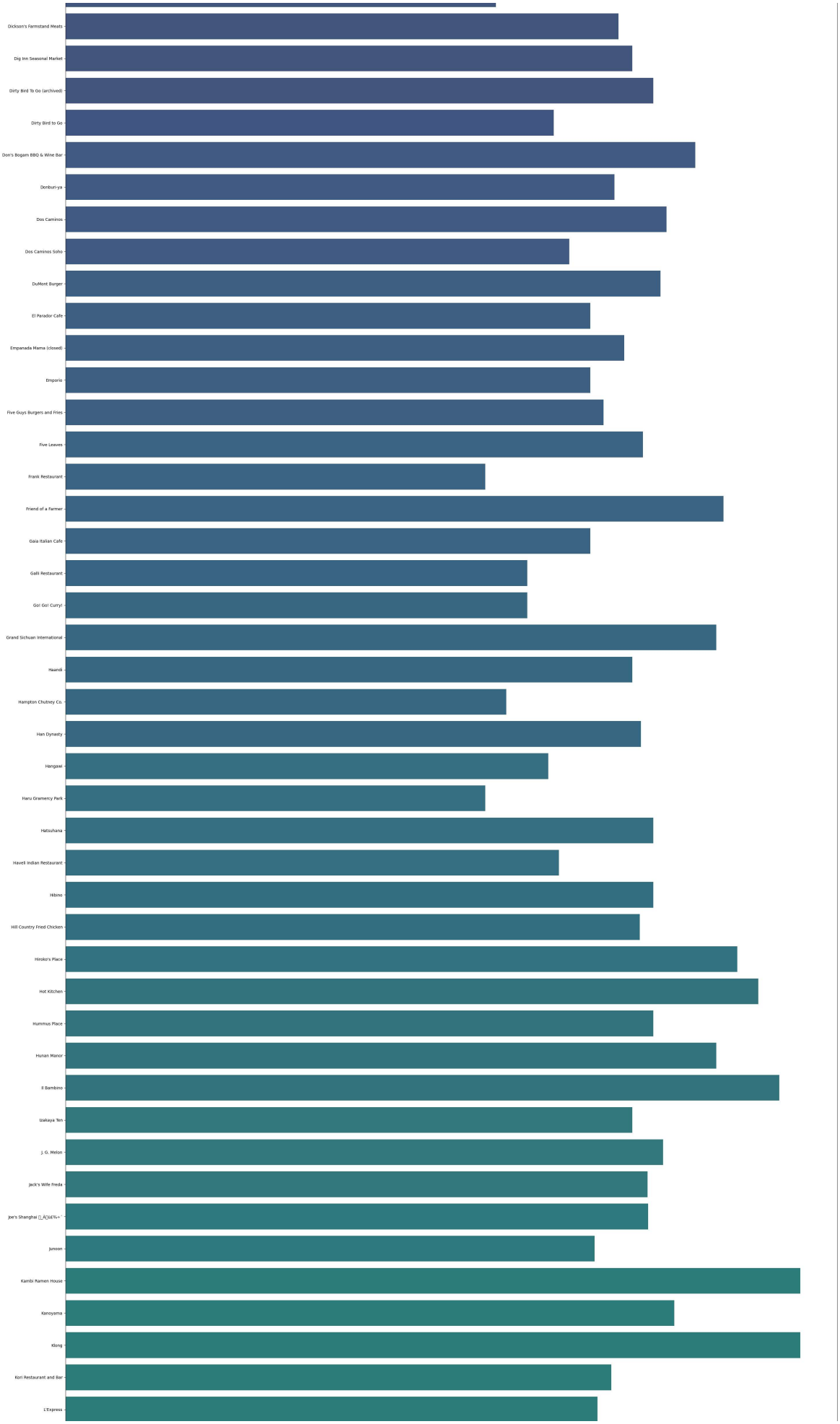
3. Restaurant Performance

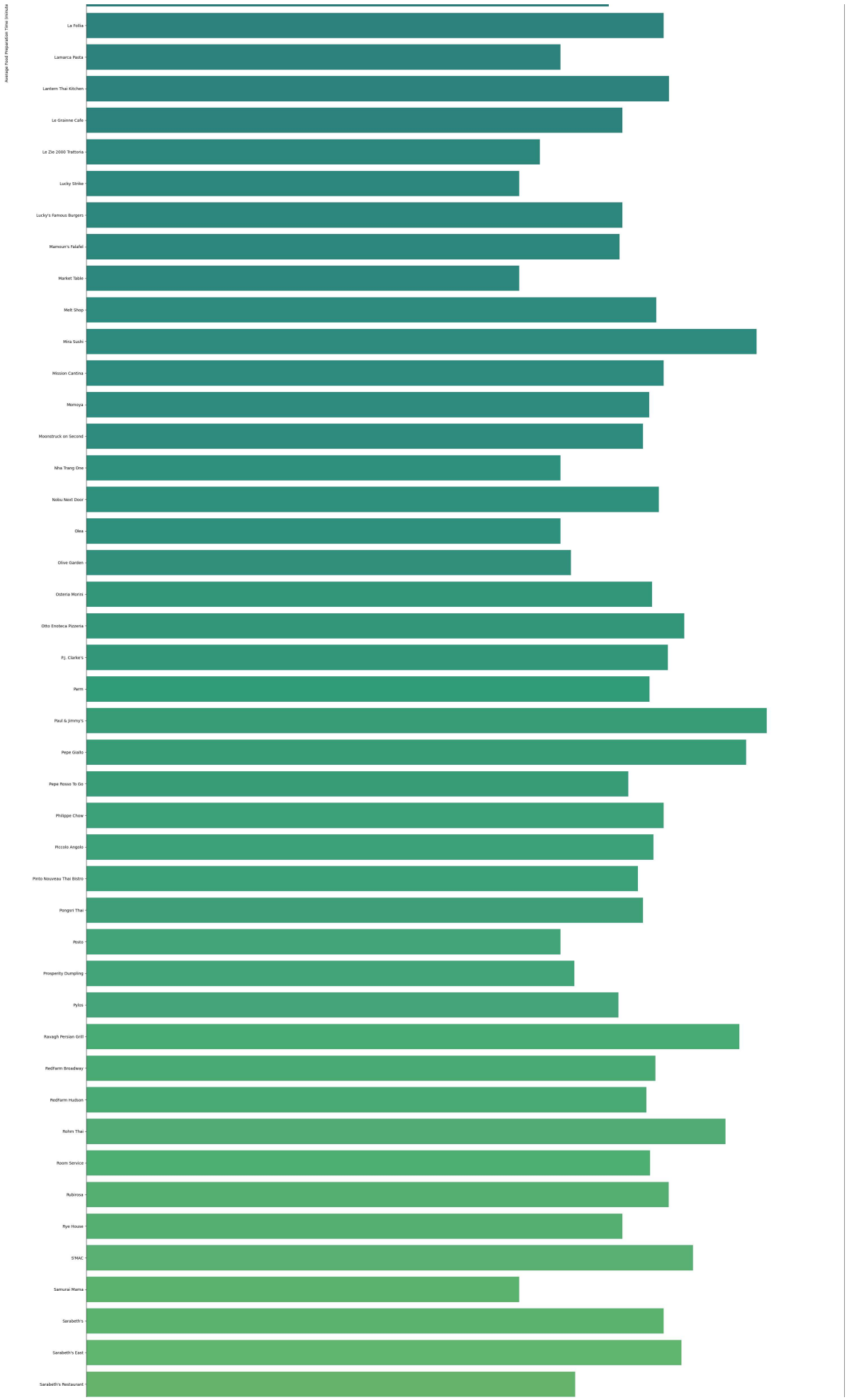
• What is the average food preparation time for each restaurant?

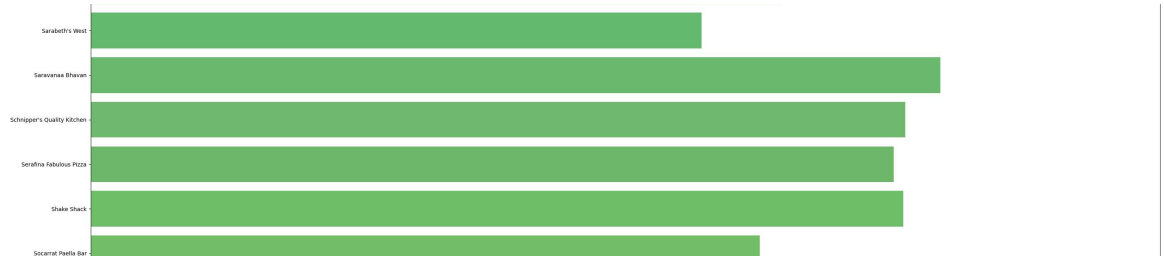
```
restaurant_name
'wichcraft          28.000000
12 Chairs           27.000000
5 Napkin Burger     30.200000
67 Burger           20.000000
Alidoro             34.000000
                        ...
Zero Otto Nove      30.000000
brgr                25.000000
da Umberto          24.333333
ilili Restaurant    26.388889
indikitch           30.750000
Name: food_preparation_time, Length: 178, dtype: float64
```

Table No.21


Average Food Preparation Time for each Restaurant
'wichcraft
12 Chairs
5 Napkin Burger
67 Burger
Alidoro
Amma
Amy Ruth's
Anjappar Chettinad
Asuka Sushi
Aurora
Balade
Balthazar Boulangerie
Barbounia
Bareburger
Benihana
Bhatti Indian Grill
Big Daddy's
Big Wong Restaurant □ ¤¼¾
Bistango
Blue Ribbon Brooklyn
Blue Ribbon Fried Chicken
Blue Ribbon Sushi
Blue Ribbon Sushi Bar & Grill
Blue Ribbon Sushi Izakaya
Boqueria
Bubby's
Bukhara Grill
Burger Joint
Byblos Restaurant
Cafe Habana
Cafe Mogador
Cafe de La Esquina
Cafeteria
Café © China
Carmine's
Chipotle Mexican Grill $1.99 Delivery
Cho Dang Gol
Chola Eclectic Indian Cuisine
Chote Nawab
Cipriani Le Specialita
Coppola's East
Crema Restaurante
Delicatessen
Despaña


Dickson's Farmstand Meats
Dig Inn Seasonal Market
Dirty Bird To Go (archived)
Dirty Bird to Go
Don's Bogam BBQ & Wine Bar
Donburi-ya
Dos Caminos
Dos Caminos Soho
DuMont Burger
El Parador Cafe
Empanada Mama (closed)
Emporio
Five Guys Burgers and Fries
Five Leaves
Frank Restaurant
Friend of a Farmer
Gaia Italian Cafe
Galli Restaurant
Go! Go! Curry!
Grand Sichuan International
Haandi
Hampton Chutney Co.
Han Dynasty
Hangawi
Haru Gramercy Park
Hatsuhana
Haveli Indian Restaurant
Hibino
Hill Country Fried Chicken
Hiroko's Place
Hot Kitchen
Hummus Place
Hunan Manor
Il Bambino
Izakaya Ten
J. G. Melon
Jack's Wife Freda
Joe's Shanghai □Ã□úÉ¥¬
Jonson
Kambi Ramen House
Kanoyama
Klong
Kori Restaurant and Bar
L'Express


Average Food Preparation Time (minute
La Follia
Lamarca Pasta
Lantern Thai Kitchen
Le Grainne Cafe
Le Zie 2000 Trattoria
Lucky Strike
Lucky's Famous Burgers
Mamoun's Falafel
Market Table
Melt Shop
Mira Sushi
Mission Cantina
Momoya
Moonstruck on Second
Nha Trang One
Nobu Next Door
Olea
Olive Garden
Osteria Morini
Otto Enoteca Pizzeria
P.J. Clarke's
Parm
Paul & Jimmy's
Pepe Giallo
Pepe Rosso To Go
Philippe Chow
Piccolo Angolo
Pinto Nouveau Thai Bistro
Pongsri Thai
Posto
Prosperity Dumpling
Pylos
Ravagh Persian Grill
Redfarm Broadway
Redfarm Hudson
Rohm Thai
Room Service
Rubirosa
Rye House
S'MAC
Samurai Mama
Sarabeth's
Sarabeth's East
Sarabeth's Restaurant



Observation

- Longest food preparation time restaurant is: Cipriani Le Specialita
- With avg preparation time: 35.0

• Which restaurant has the shortest average food preparation time?

Observation

- Shortest food preparation time restaurant is: 67 Burger
- With avg prepartion time: 20.0

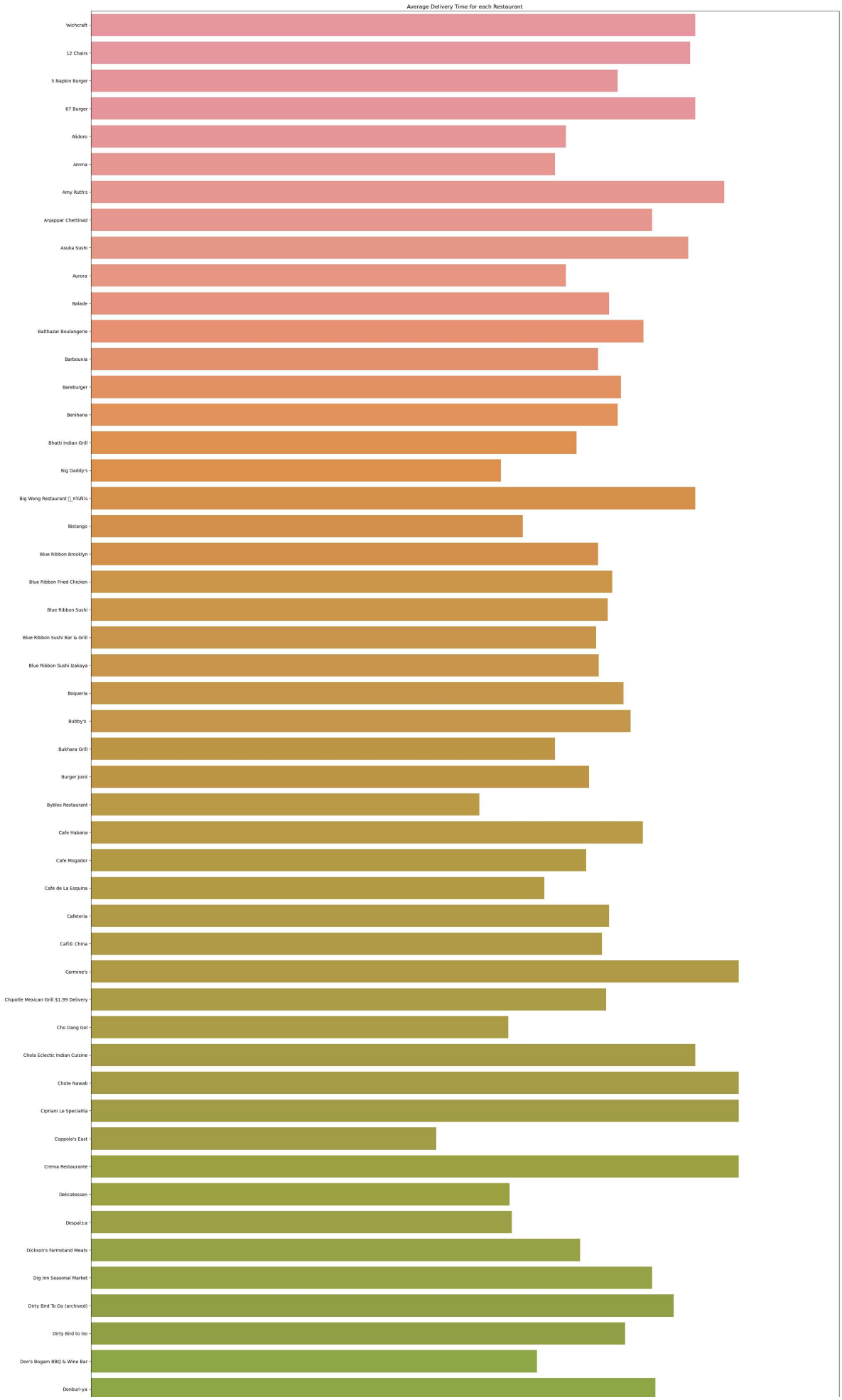
• How does the average delivery time compare across different restaurants?

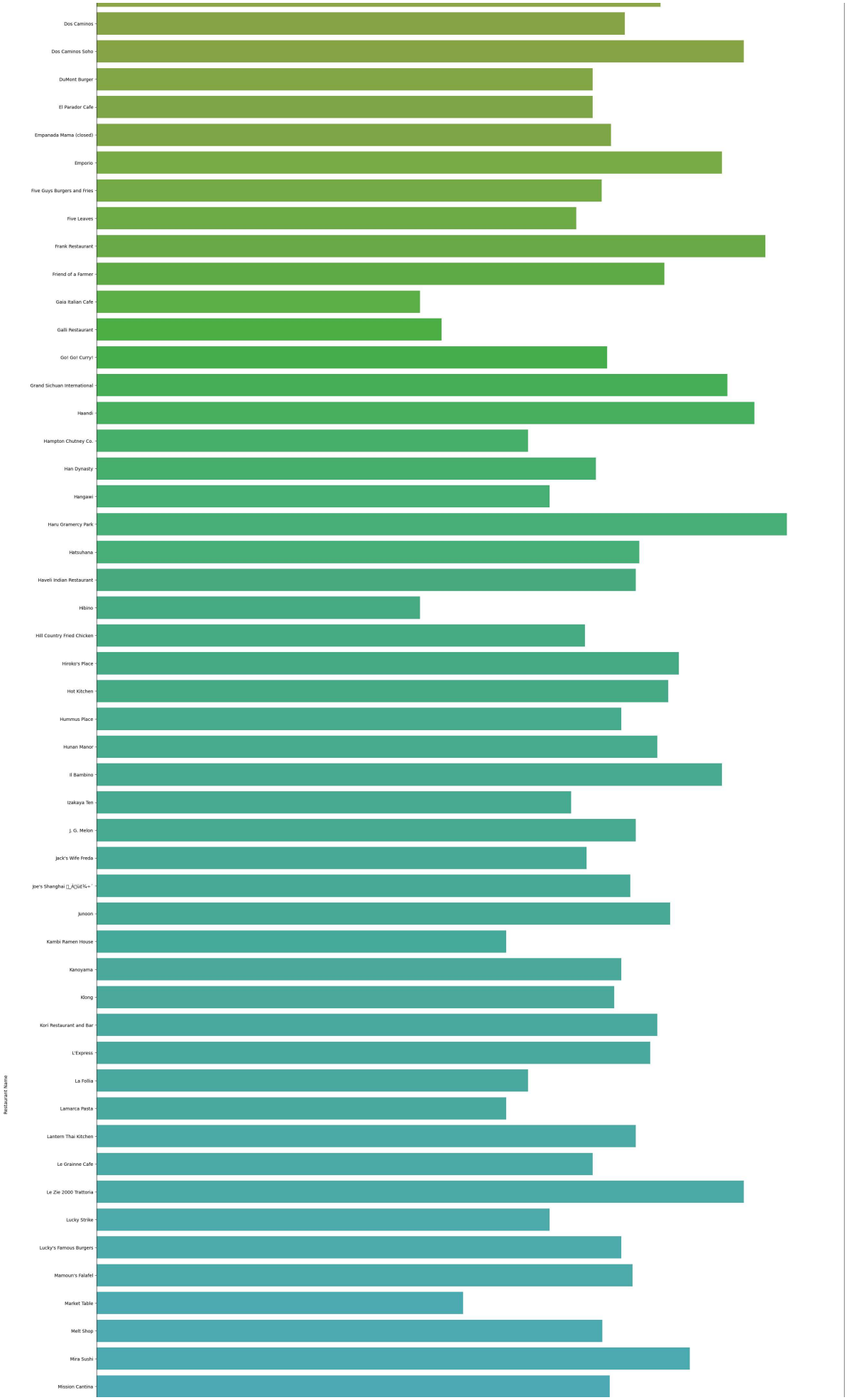
Average delivery time for each restaurant:

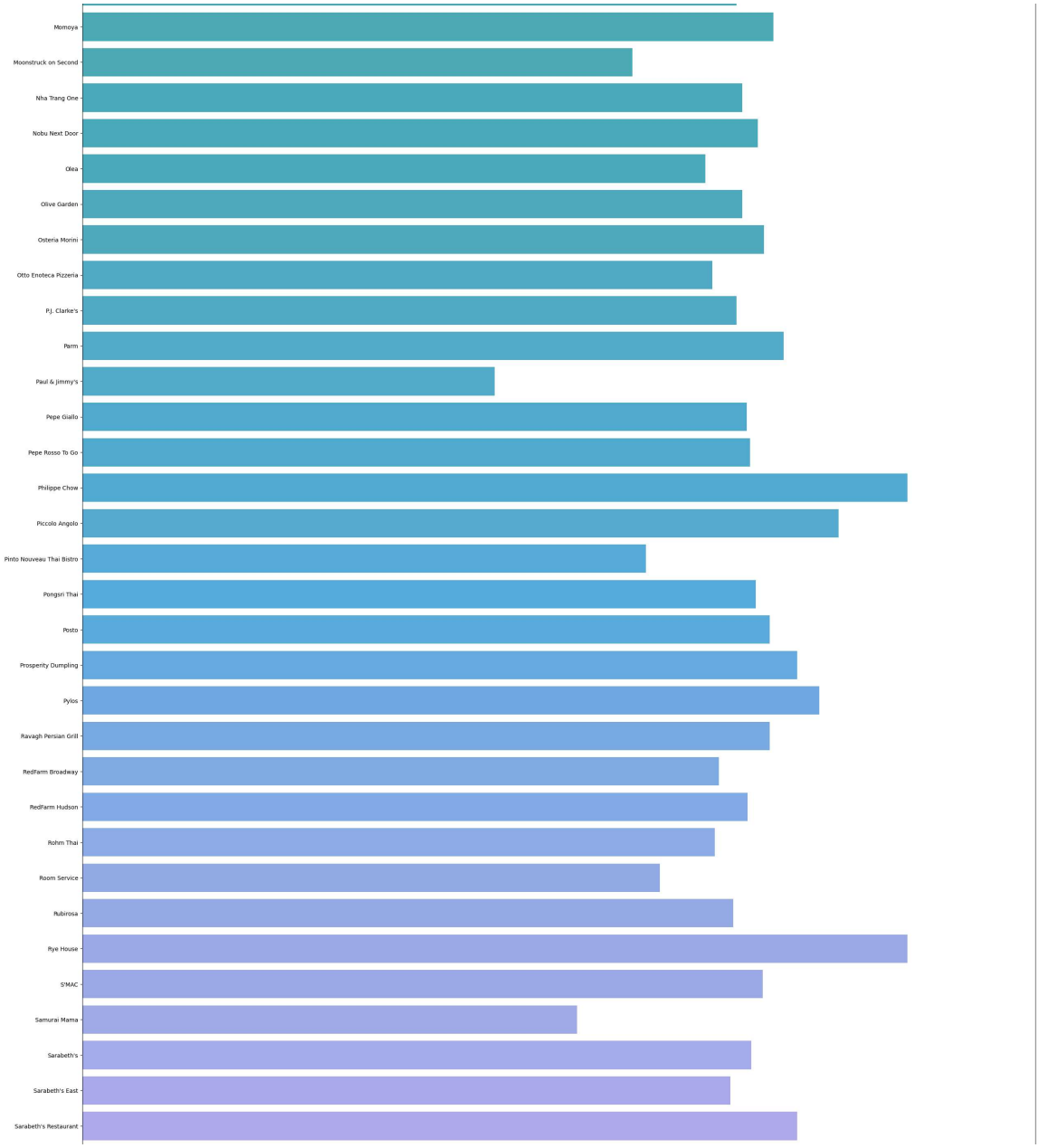
	restaurant_name	delivery_time
0	'wichcraft	28.000000
1	12 Chairs	27.750000
2	5 Napkin Burger	24.400000
3	67 Burger	28.000000
4	Alidoro	22.000000
..	...	...
173	Zero Otto Nove	21.500000
174	brgr	25.000000
175	da Umberto	28.000000
176	ilili Restaurant	24.888889
177	indikitch	25.500000

[178 rows x 2 columns]

Table No. 22


Average Delivery Time for each Restaurant
'wichcraft
12 Chairs
5 Napkin Burger
67 Burger
Alidoro
Amma
Amy Ruth's
Anjappar Chettinad
Asuka Sushi
Aurora
Balade
Balthazar Boulangerie
Barbounia
Bareburger
Benihana
Bhatti Indian Grill
Big Daddy's
Big Wong Restaurant □_¤¼Ψ¼
Bistango
Blue Ribbon Brooklyn
Blue Ribbon Fried Chicken
Blue Ribbon Sushi
Blue Ribbon Sushi Bar & Grill
Blue Ribbon Sushi Izakaya
Boqueria
Bubby's
Bukhara Grill
Burger Joint
Byblos Restaurant
Cafe Habana
Cafe Mogador
Cafe de La Esquina
Cafeteria
Cafi© China
Carmine's
Chipotle Mexican Grill $1.99 Delivery
Cho Dang Gol
Chola Eclectic Indian Cuisine
Chote Nawab
Cipriani Le Specialita
Coppola's East
Crema Restaurante
Delicatessen
Despaí±a
Dickson's Farmstand Meats
Dig Inn Seasonal Market
Dirty Bird To Go (archived)
Dirty Bird to Go
Don's Bogam BBQ & Wine Bar
Donburi-ya


Restaurant Name
Dos Caminos
Dos Caminos Soho
DuMont Burger
El Parador Cafe
Empanada Mama (closed)
Emporio
Five Guys Burgers and Fries
Five Leaves
Frank Restaurant
Friend of a Farmer
Gaia Italian Cafe
Galli Restaurant
Go! Go! Curry!
Grand Sichuan International
Haandi
Hampton Chutney Co.
Han Dynasty
Hangawi
Haru Gramercy Park
Hatsuhana
Haveli Indian Restaurant
Hibino
Hill Country Fried Chicken
Hiroko's Place
Hot Kitchen
Hummus Place
Hunan Manor
Il Bambino
Izakaya Ten
J. G. Melon
Jack's Wife Freda
Joe's Shanghai □_Ä□üE¾÷
Junoon
Kambi Ramen House
Kanoyama
Klong
Kori Restaurant and Bar
L'Express
La Follia
Lamarca Pasta
Lantern Thai Kitchen
Le Grainne Cafe
Le Zie 2000 Trattoria
Lucky Strike
Lucky's Famous Burgers
Mamoun's Falafel
Market Table
Melt Shop
Mira Sushi
Mission Cantina



Observation

- Restaurant with the highest average delivery time is: Sarabeth's West
- Highest average delivery time: 33.0

• Is there a correlation between the cost of the order and the rating given?

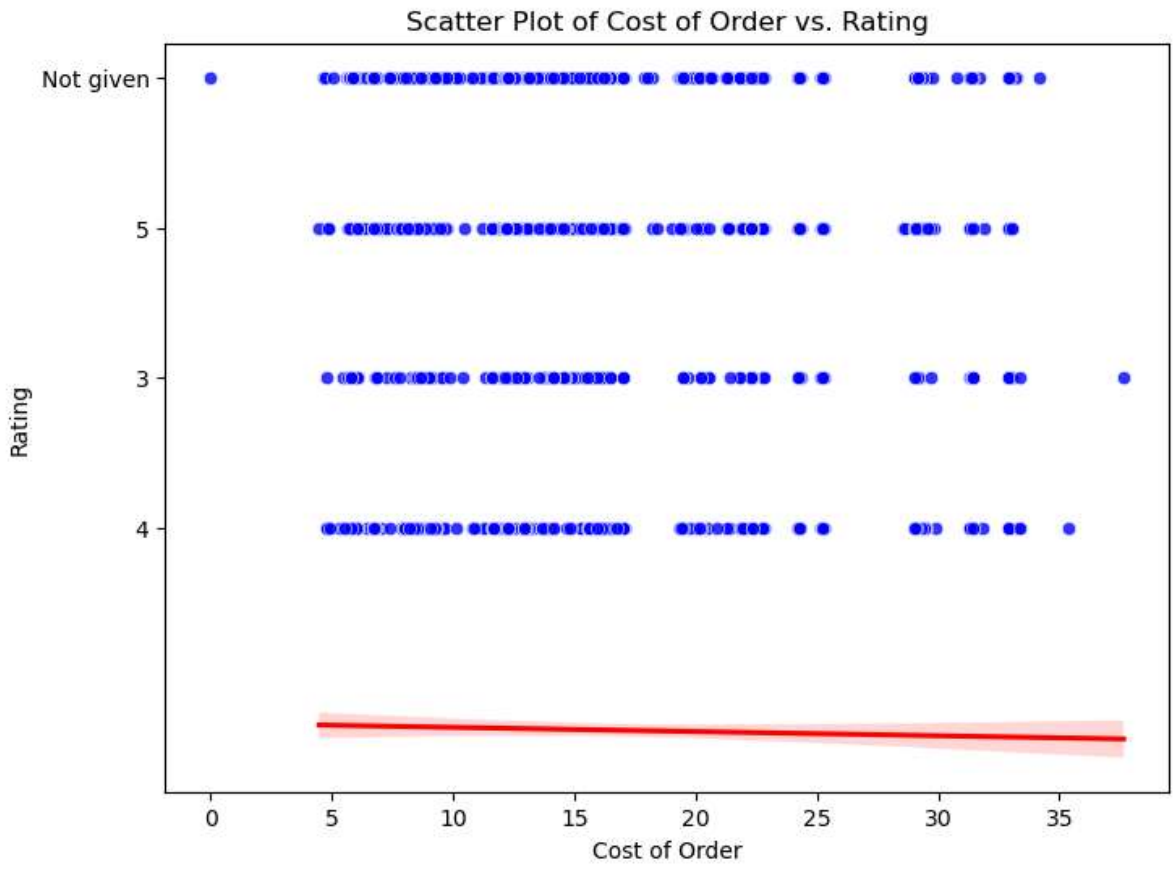


Fig No. 10

Observation

- Correlation between cost_of_the_order and rating: 0.03

4. Demand Patterns

• How does the demand for different cuisine types vary on weekdays versus weekends?

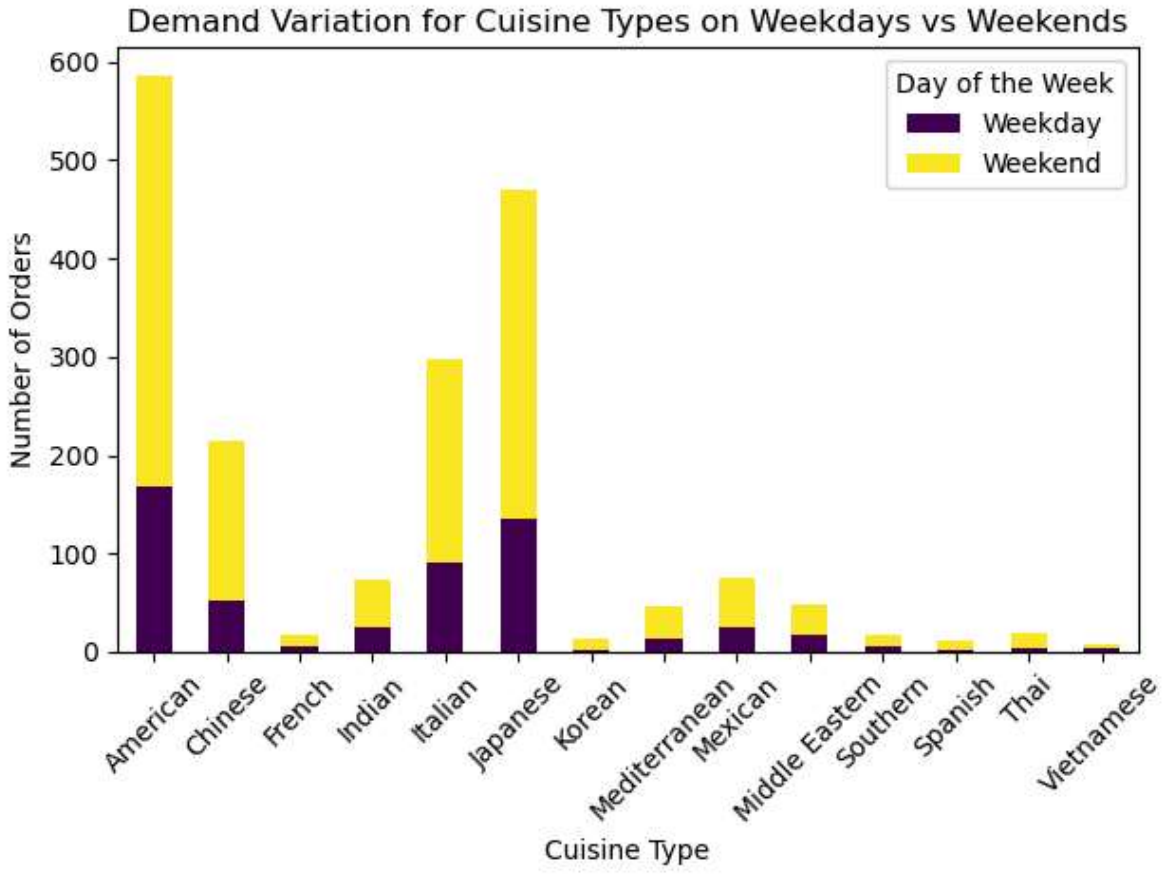


Fig No. 11

Observation

- The demand for American cuisine is highest in the Weekend
- The demand for American cuisine is also highest int he Weekday

• Which day of the week has the highest average order cost?

Observation

- Day with the highest average order cost: Weekend

• What is the most common day for orders to be placed?

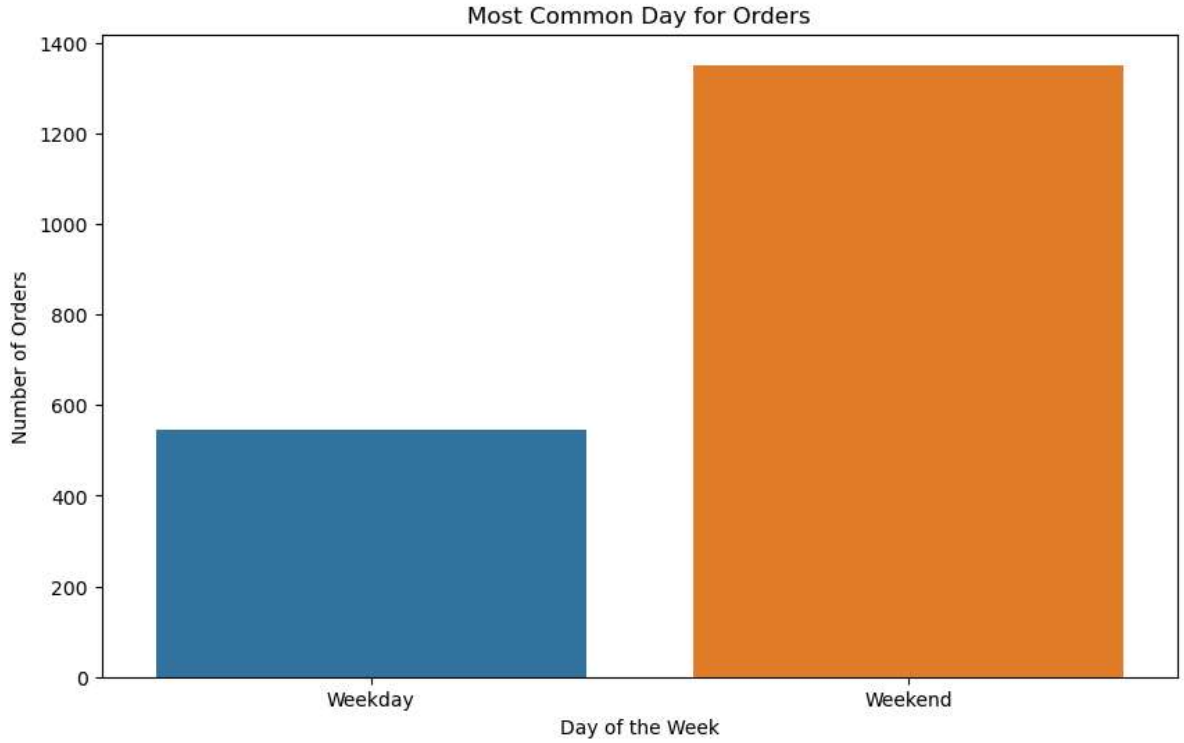


Fig No.12

Observation

- The most common day for orders to be placed is: Weekend

• How does the average rating vary by cuisine type?

```
cuisine_type
American          4.300813
Chinese           4.338346
French            4.300000
Indian            4.540000
Italian           4.360465
Japanese          4.373626
Korean            4.111111
Mediterranean     4.218750
Mexican           4.404255
Middle Eastern    4.235294
Southern          4.307692
Spanish           4.833333
Thai              4.666667
Vietnamese        4.000000
Name: rating, dtype: float64
```

Table No. 23

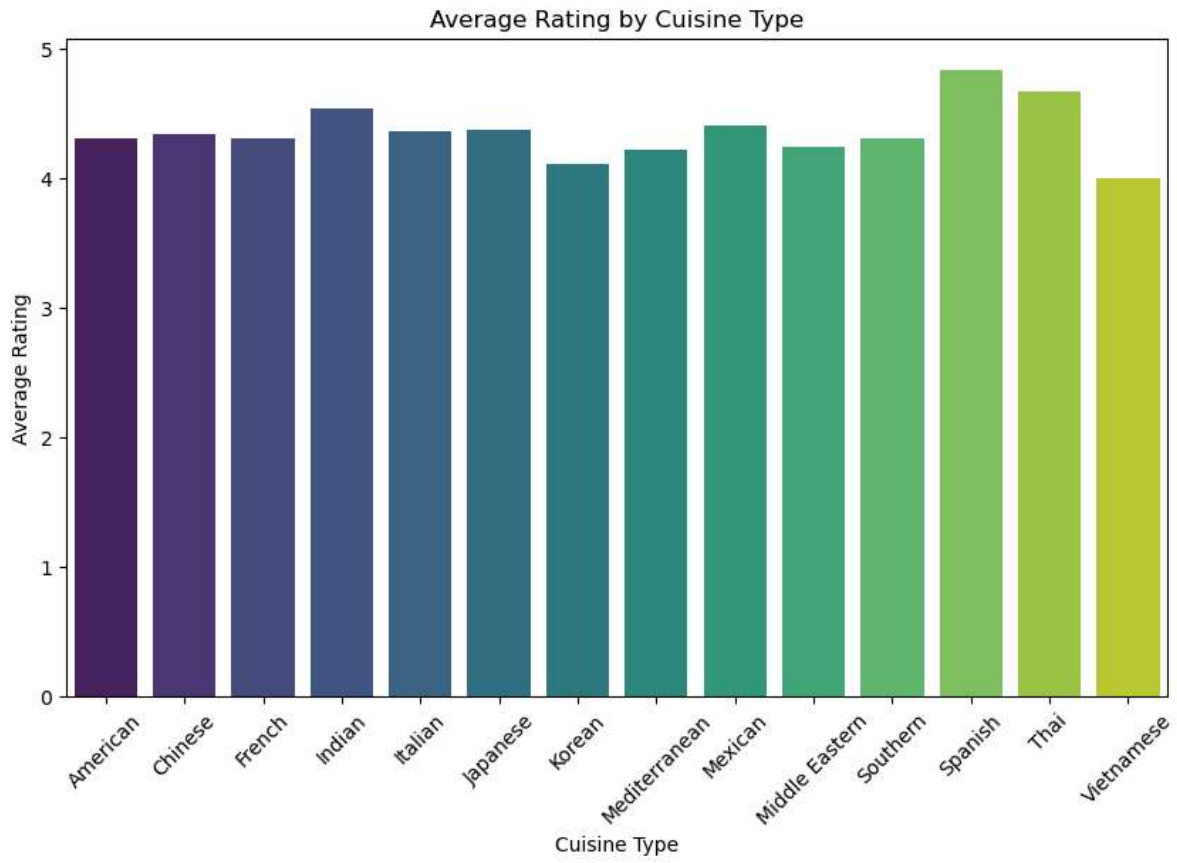


Fig No. 13

Observation

- Restaurant with highest average rating by cuisine is Spanish and the least rated is Vietnamese

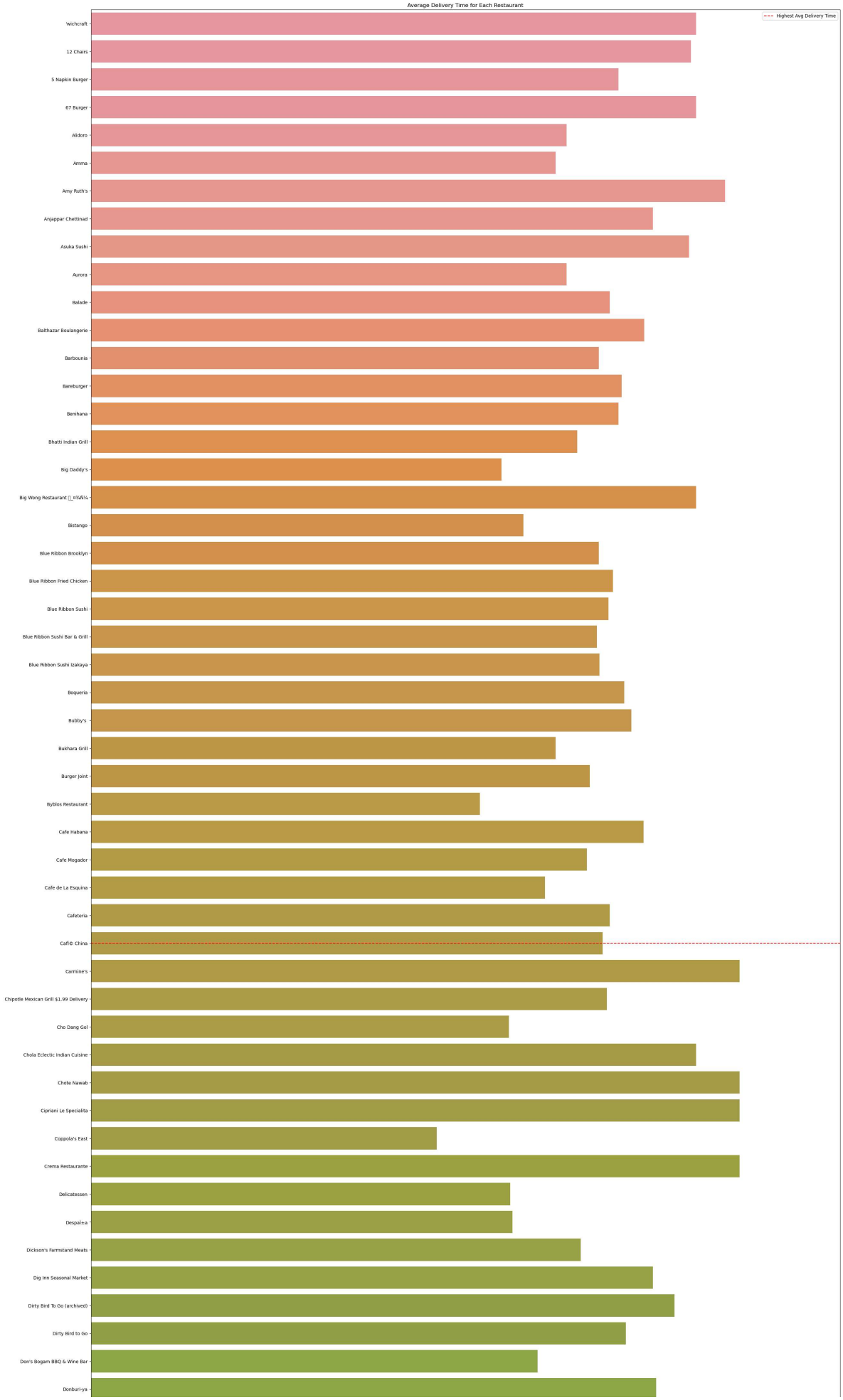
5. Operational Efficiency

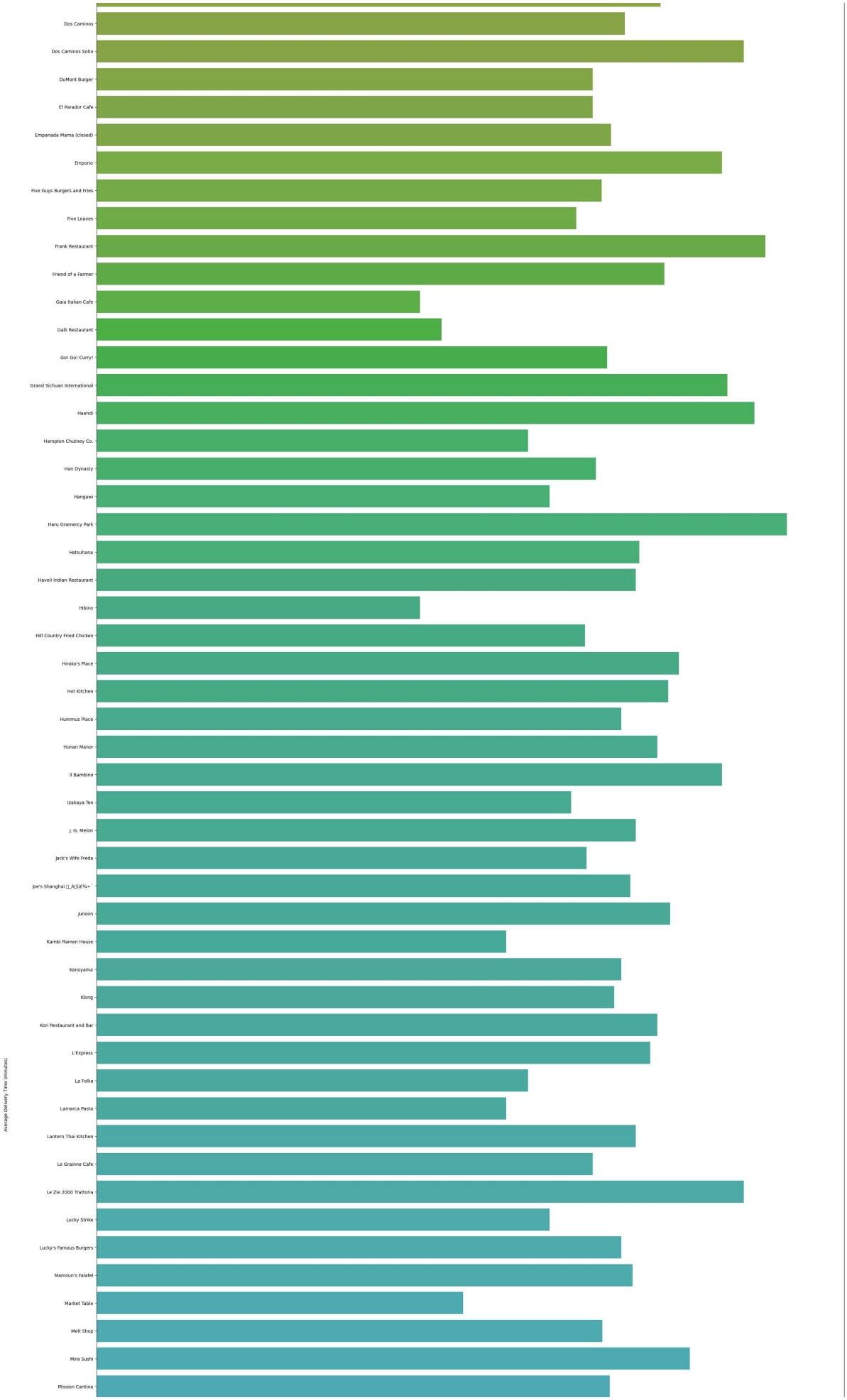
• What is the average delivery time for all orders?

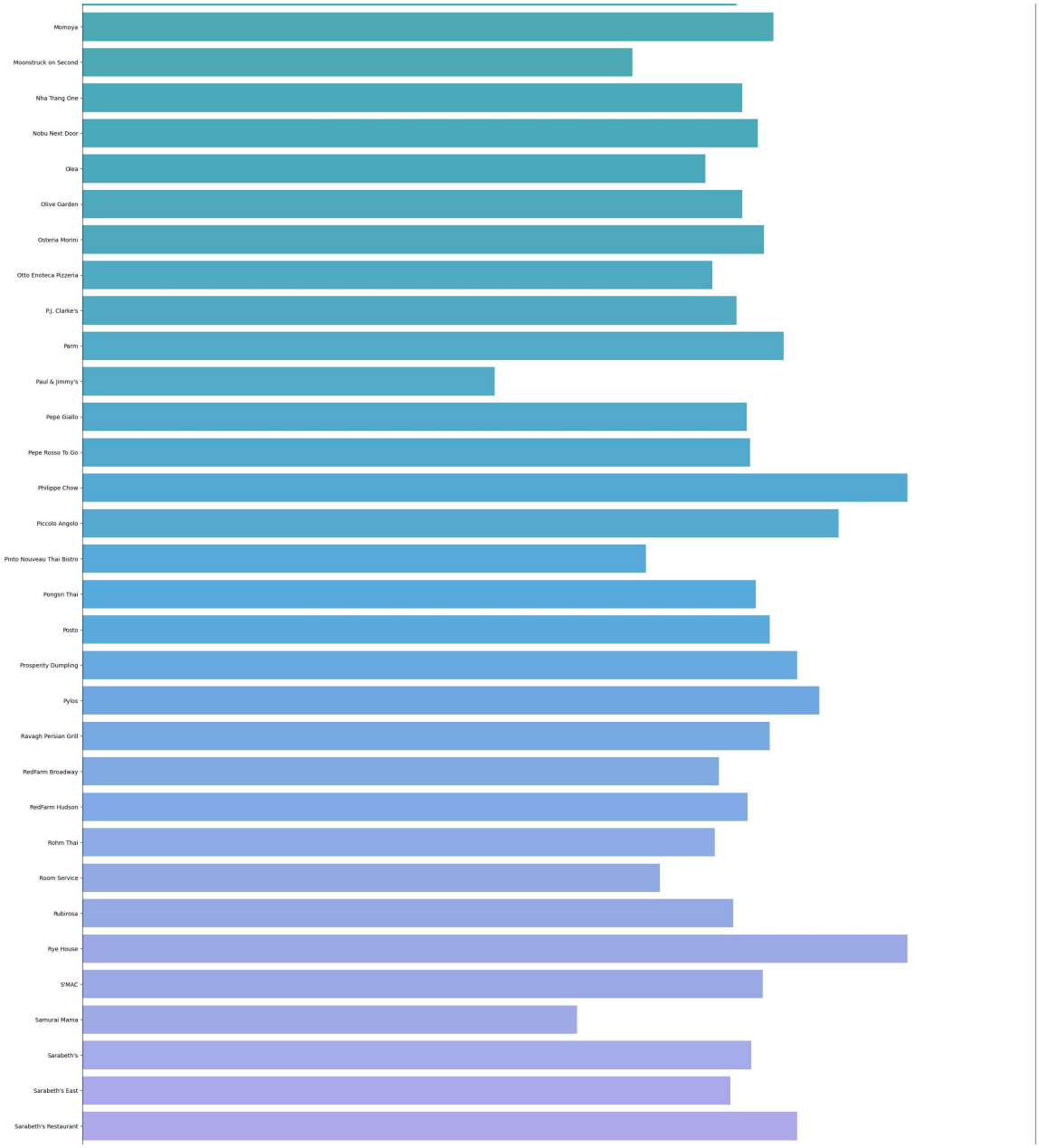
Observation

- Average delivery time for all orders is: 24.162447257383963

- **Which restaurant has the longest average delivery time?**


Average Delivery Time for Each Restaurant
Highest Avg Delivery Time
'wichcraft
12 Chairs
5 Napkin Burger
67 Burger
Alidoro
Amma
Amy Ruth's
Anjappar Chettinad
Asuka Sushi
Aurora
Balade
Balthazar Boulangerie
Barbounia
Bareburger
Benihana
Bhatti Indian Grill
Big Daddy's
Big Wong Restaurant □_¤¾Ñ¼
Bistango
Blue Ribbon Brooklyn
Blue Ribbon Fried Chicken
Blue Ribbon Sushi
Blue Ribbon Sushi Bar & Grill
Blue Ribbon Sushi Izakaya
Boqueria
Bubby's
Bukhara Grill
Burger Joint
Byblos Restaurant
Cafe Habana
Cafe Mogador
Cafe de La Esquina
Cafeteria
Cafi© China
Carmine's
Chipotle Mexican Grill $1.99 Delivery
Cho Dang Gol
Chola Eclectic Indian Cuisine
Chote Nawab
Cipriani Le Specialita
Coppola's East
Crema Restaurante
Delicatessen
Despaña
Dickson's Farmstand Meats
Dig Inn Seasonal Market
Dirty Bird To Go (archived)
Dirty Bird to Go
Don's Bogam BBQ & Wine Bar
Donburi-ya


Dos Caminos
Dos Caminos Soho
DuMont Burger
El Parador Cafe
Empanada Mama (closed)
Emporio
Five Guys Burgers and Fries
Five Leaves
Frank Restaurant
Friend of a Farmer
Gaia Italian Cafe
Galli Restaurant
Go! Go! Curry!
Grand Sichuan International
Haandi
Hampton Chutney Co.
Han Dynasty
Hangawi
Haru Gramercy Park
Hatsuhana
Haveli Indian Restaurant
Hibino
Hill Country Fried Chicken
Hiroko's Place
Hot Kitchen
Hummus Place
Hunan Manor
Il Bambino
Izakaya Ten
J. G. Melon
Jack's Wife Freda
Joe's Shanghai □_Ä□üE¾÷
Junoon
Kambi Ramen House
Kanoyama
Klong
Kori Restaurant and Bar
L'Express
La Follia
Lamarca Pasta
Lantern Thai Kitchen
Le Grainne Cafe
Le Zie 2000 Trattoria
Lucky Strike
Lucky's Famous Burgers
Mamoun's Falafel
Market Table
Melt Shop
Mira Sushi
Mission Cantina
Average Delivery Time (minutes)



Observation

- Restaurant with the highest average delivery time is: Sarabeth's West

• Is there a relationship between food preparation time and delivery time?

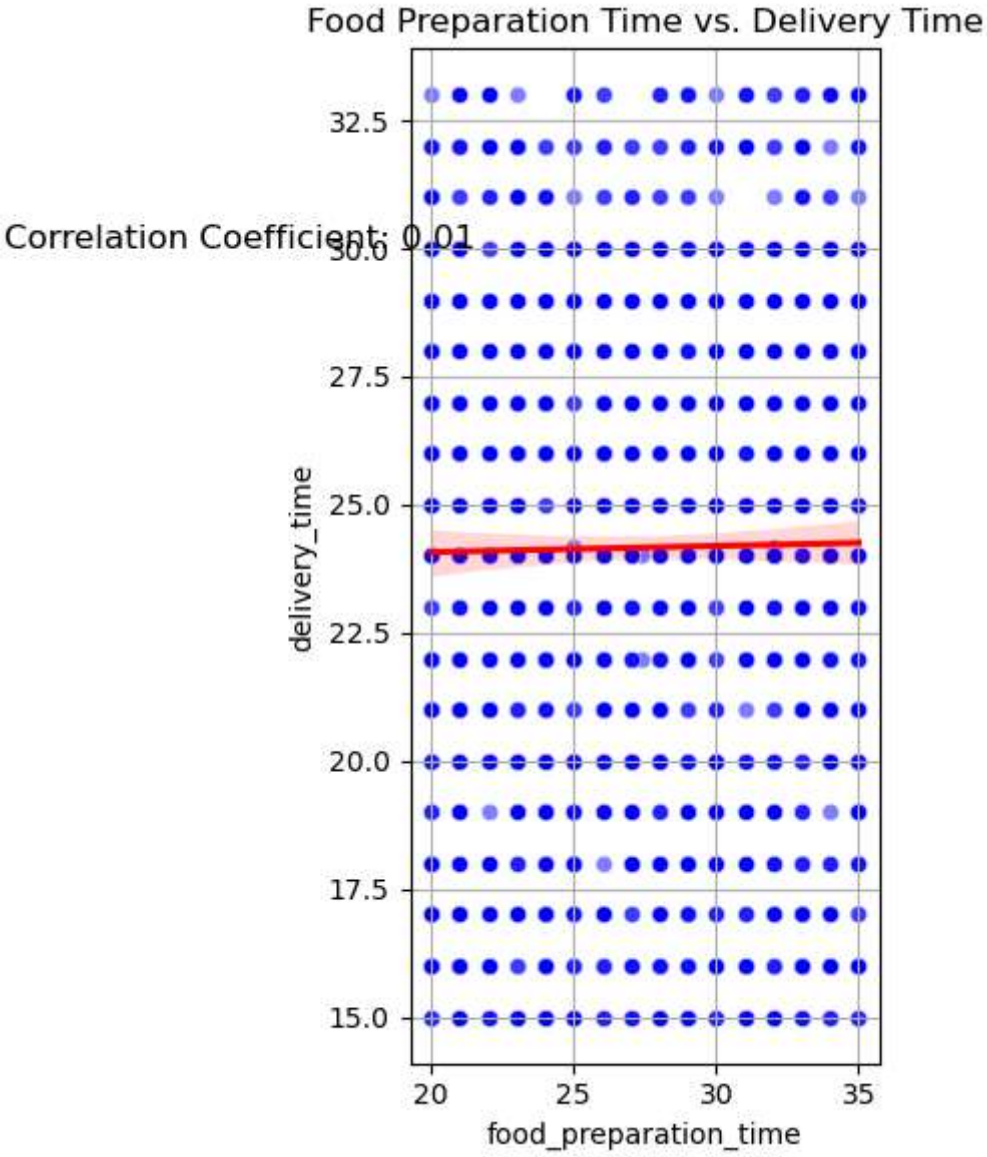


Fig No.15

Observation

- coefficient of correlation between food preparation and delviery time is 0.01

• How does the delivery time impact customer ratings?

```
<class 'pandas.core.frame.DataFrame'>
Index: 1162 entries, 2 to 1896
Data columns (total 9 columns):
 #   Column                 Non-Null Count  Dtype
---  ------                 --------------  -----
 0   order_id               1162 non-null   int64
 1   customer_id            1162 non-null   int64
 2   restaurant_name        1162 non-null   object
 3   cuisine_type           1162 non-null   object
 4   cost_of_the_order      1162 non-null   float64
 5   day_of_the_week        1162 non-null   object
 6   rating                 1162 non-null   float64
 7   food_preparation_time  1162 non-null   float64
 8   delivery_time          1162 non-null   float64
dtypes: float64(4), int64(2), object(3)
memory usage: 123.1+ KB
```

Table No. 24

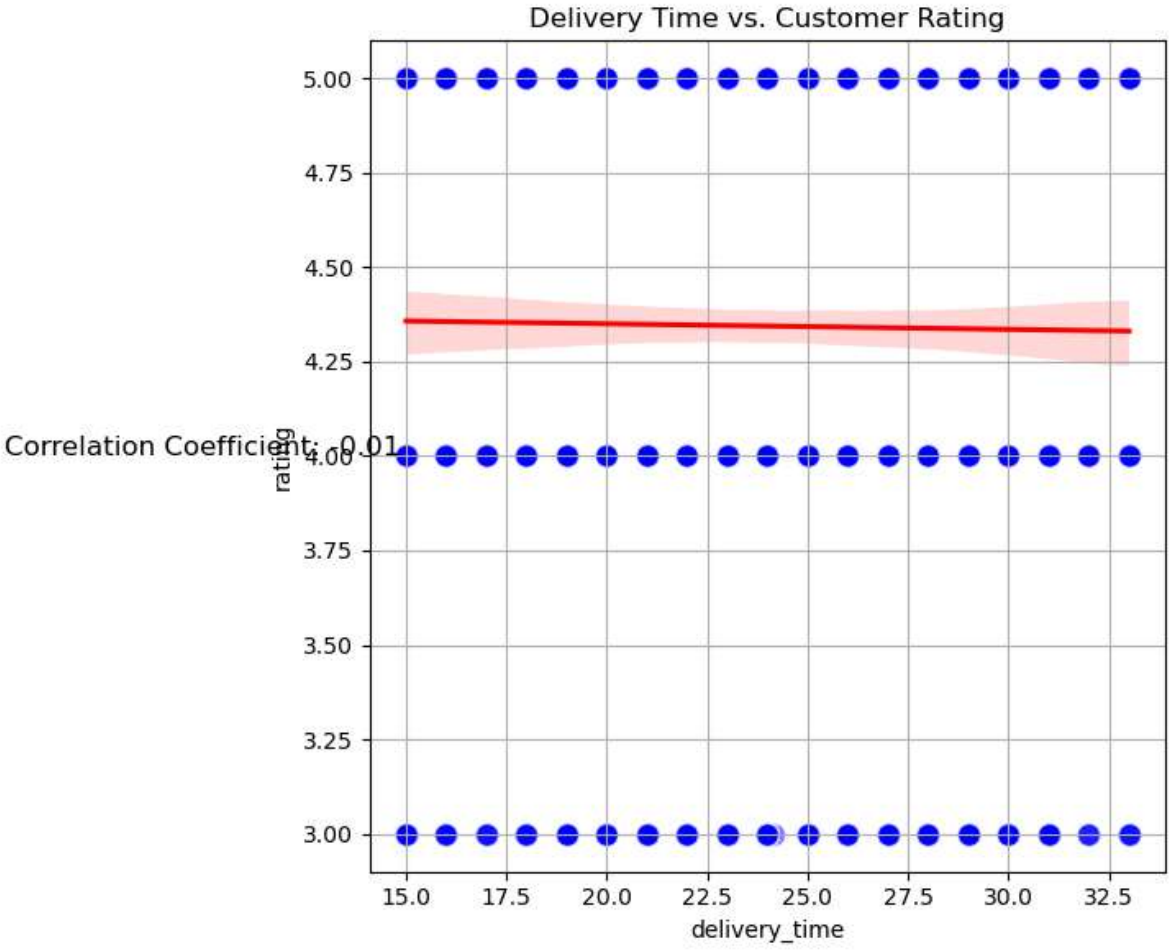


Fig No.16

Observation

- As we can see from the scatter plot between rating and delivery time the coefficient of correlation is 0.01

6. Customer Insights

• What is the repeat order rate (number of customers who have placed more than one order)?

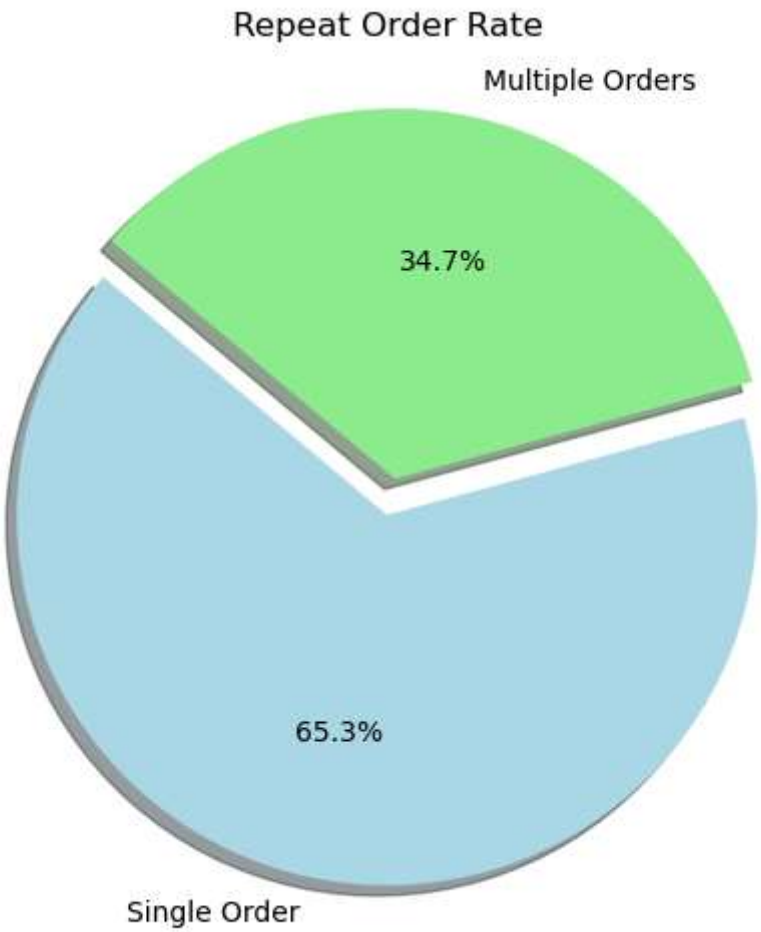


Fig No. 17

Observation

- Repeat Order Rate is 34.67%
- And we can see single order rate is 65.3%

• What percentage of orders receive a rating of 4 or higher?

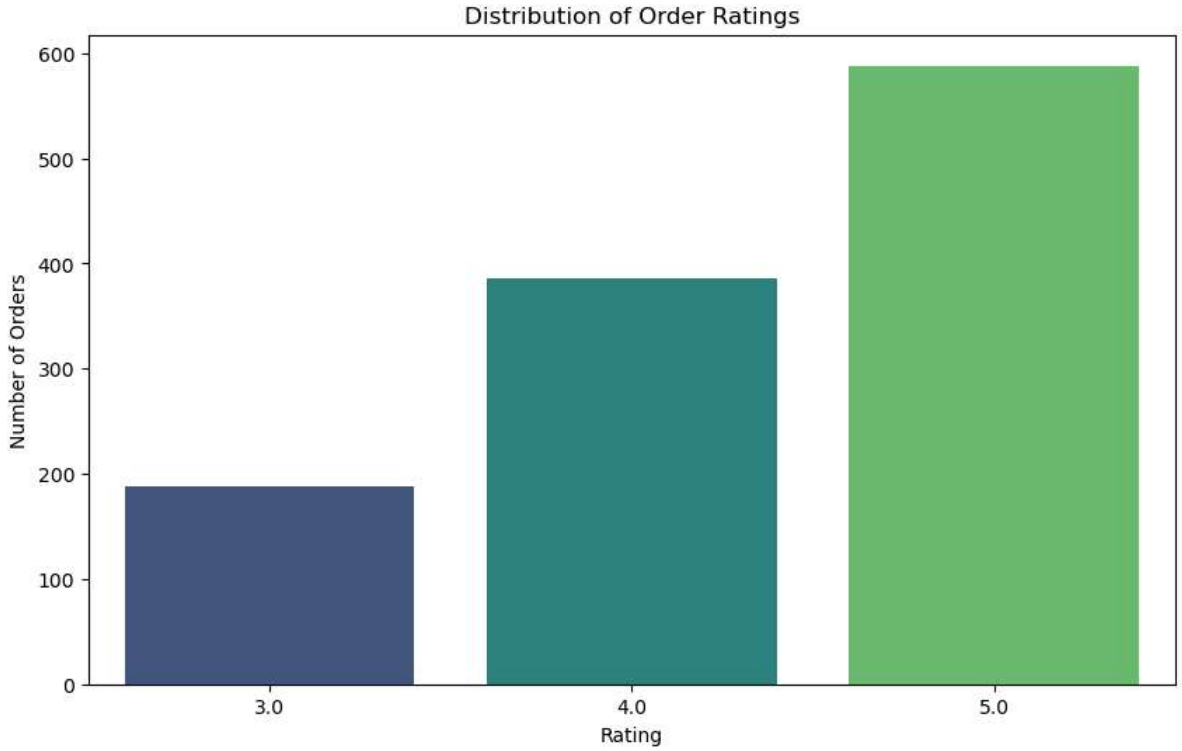


Fig No. 18

Observation

- Percentage of orders with rating greater than 4 is 50.60%
- Followed by 4 rating and 3 rating
- so we can say that more orders have a higher rating