Data Analysis of Zomato

Vaishali Kadam

Problem Statement

- Zomato is an Indian multinational restaurant aggregator and food delivery company .
- Zomato is one of the most comprehensive and user-friendly apps for finding nearby restaurants and cafés to dine in or to order food online. It also gives menus, reviews, and ratings to acquire factual information on eateries.
- Zomato Limited is an online restaurant guide and food ordering platform. It offered comprehensive details on more than 1.4 million establishments in 23 countries. There were restaurant names, menu items, pricing, reviews, and other information.
- It has evolved into an internet platform for meal delivery over the years. Zomato's tagline – "Never have a bad meal". It serves as a comprehensive encyclopedia of restaurants, replete with ratings, average pricing, menus, and reviews.
- Lets take the data of Zomato Worldwide and analyze it and give proper information to the customer so that the customer enjoy their meals without any interruption.
- In the data provide we need to identify the low or under rated restaurant and need to update the restaurant.
- Find out the cheapest restaurant belong to the city and its country similarly the most expensive restaurant belong to the city and country.
- The distribution of restaurant over the various countries such as India , USA, UAE and few other countries.
- Here we get the data and with respect to that data we analyze it and further which is in major part of data …we going on that path..

Proposed Solution

- User who try to find local restaurants of various cuisines. Restaurant who want their name to reached targeted audience.
- Customer Preferring Home Delivery.
- Database Companies
- Market Research Companies
- One stop shop for dines and offers restaurants a way to differentiate them
- Bridge the gap between Customers and Restaurants by efficient technology by reduce delivery time .
- Rating based price model for foods
- Provide local restaurants and hotels search services.
- Collect data for food menus , contacts and relevant information to users
- Large dataset present across cities.
- Presence across 15 countries.

Descriptive Analysis and loading of data

First we have to load all the Library which we are needed and then load the Data through csv file…

In [1]:
```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import matplotlib.colors as mcolors
```

In [2]:
```
df=pd.read_csv(r'C:\Users\LENOVO\OneDrive\Desktop\New folder (2)\zomato.csv',encoding='latin')
df
```

Out[2]:

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	Latitude	Cuisines	...	Currency	Has Table booking	H Onli delive
0	6317637	Le Petit Souffle	162	Makati City	Third Floor, Century City Mall, Kalayaan	Century City Mall, Poblacion, Makati City	Century City Mall, Poblacion, Makati City,	121.027535	14.565443	French, Japanese, Desserts	...	Botswana Pula(P)	Yes	

After Loading the data we have to find more information about the columns which we are dealing with…..

Data Cleaning

Checking if dataset contains any null values and have to clear it because it affect our data while analyzing.....

Cuisines seems to contain null values. Hence any further analysis involving Cuisines the NaN values has to be considered.

There is an other file which is also available along with this dataset .

We have to load that data as well

```
In [3]: df1 = pd.read_excel(r'C:\Users\LENOVO\OneDrive\Desktop\New folder (2)\Country-Code.xlsx')
        df1
```

Out[3]:

	Country Code	Country
0	1	India
1	14	Australia
2	30	Brazil
3	37	Canada
4	94	Indonesia

```
41]: nan_values = df.isna()
     nan_columns = nan_values.any()

     columns_with_nan = df.columns[nan_columns].tolist()
     print(columns_with_nan)
```

```
['Cuisines']
```

df.isna() nan_values.any() print(columns_with_nan)

```
42]: df.isna()
     nan_values.any()
```

```
42]: Restaurant ID          False
     Restaurant Name        False
     Country Code           False
     City                   False
     Address                False
     Locality               False
     Locality Verbose       False
     Longitude              False
     Latitude               False
     Cuisines                True
     Average Cost for two   False
     Currency               False
     Has Table booking      False
     Has Online delivery    False
     Is delivering now      False
     Switch to order menu   False
     Price range            False
     Aggregate rating       False
     Rating color           False
     Rating text            False
```

Merging of the Data

Let merge the both dataset this will help us to understand the data set country wise.

```
df2=pd.merge(df,df1,on='Country Code',how='left')
df2.head()
```

	Restaurant ID	Restaurant Name	Country Code	City	Address	Locality	Locality Verbose	Longitude	Latitude	Cuisines	...	Has Table booking	Has Online delivery	Is delivering now	Swi or me
0	6317637	Le Petit Souffle	162	Makati City	Third Floor, Century City Mall, Kalayaan Avenu...	Century City Mall, Poblacion, Makati City	Century City Mall, Poblacion, Makati City, Mak...	121.027535	14.565443	French, Japanese, Desserts	...	Yes	No	No	
1	6304287	Izakaya Kikufuji	162	Makati City	Little Tokyo, 2277 Chino Roces Avenue,	Little Tokyo, Legaspi Village, Makati City	Little Tokyo, Legaspi Village, Makati City, Ma...	121.014101	14.553708	Japanese	...	Yes	No	No	

We can't see columns properly so take the info of the whole data...and get knowledge about all columns...

This are all 21 columns which is further use in the data visualization.....

```
: df2.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 9551 entries, 0 to 9550
Data columns (total 22 columns):
 #   Column                Non-Null Count  Dtype  
---  ------                --------------  -----  
 0   Restaurant ID         9551 non-null   int64  
 1   Restaurant Name       9551 non-null   object 
 2   Country Code          9551 non-null   int64  
 3   City                  9551 non-null   object 
 4   Address               9551 non-null   object 
 5   Locality              9551 non-null   object 
 6   Locality Verbose      9551 non-null   object 
 7   Longitude             9551 non-null   float64
 8   Latitude              9551 non-null   float64
 9   Cuisines              9542 non-null   object 
 10  Average Cost for two  9551 non-null   int64  
 11  Currency              9551 non-null   object 
 12  Has Table booking     9551 non-null   object 
 13  Has Online delivery   9551 non-null   object 
 14  Is delivering now     9551 non-null   object 
 15  Switch to order menu  9551 non-null   object 
 16  Price range           9551 non-null   int64  
 17  Aggregate rating      9551 non-null   float64
 18  Rating color          9551 non-null   object 
 19  Rating text           9551 non-null   object 
 20  Votes                 9551 non-null   int64  
 21  Country               9551 non-null   object 
dtypes: float64(3), int64(5), object(14)
memory usage: 1.7+ MB
```

Brief Analysis and Data Visualization of Zomato Data

First up all, we have to get the whole restaurants data which is geographical spread , which help us to understand the rating,

Currency , Online Delivery , City Coverage ....and many more...

List of countries the survey is spread across

```
print('List of counteris the survey is spread accross - ')
for x in pd.unique(df2.Country): print(x)
print('Total number to country', len(pd.unique(df2.Country)))
```

```
List of counteris the survey is spread accross -
Phillipines
Brazil
United States
Australia
Canada
Singapore
UAE
India
Indonesia
New Zealand
United Kingdom
Qatar
South Africa
Sri Lanka
Turkey
Total number to country 15
```

The survey seems to have spread across 15 countries all over world.

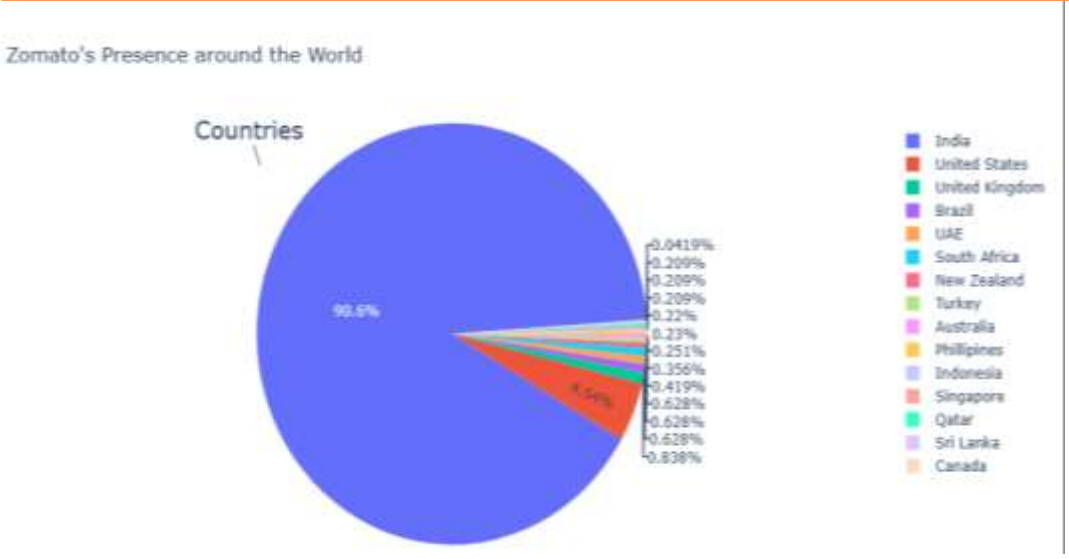
This shows that Zomato is a multinational company having actives business in all those countries.

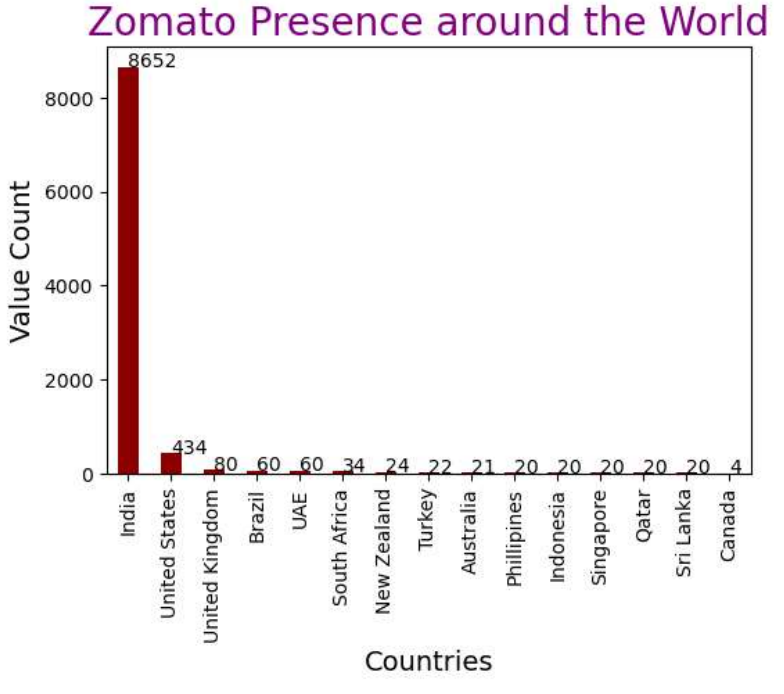
This is the value count of the Country wise data....

```
df2['Country'].value_counts()
```

```
India             8652
United States      434
United Kingdom      80
Brazil              60
UAE                 60
South Africa        60
New Zealand         40
Turkey              34
Australia           24
Phillipines         22
Indonesia           21
Singapore           20
Qatar               20
Sri Lanka           20
Canada               4
Name: Country, dtype: int64
```

Descriptive Analysis





As Zomato is a startup from India hence it makes sense that it has maximum business spread across restaurants in India

This Data show same information...India is in majority....

Descriptive Analysis

Understanding the Rating aggregate, color and text:

```
In [52]: df3 = df2.groupby(['Aggregate rating','Rating color', 'Rating text']).size().reset_index().rename(columns={0:'Rating Count'})
         df3
```

Out[52]:

	Aggregate rating	Rating color	Rating text	Rating Count
0	0.0	White	Not rated	2148
1	1.8	Red	Poor	1
2	1.9	Red	Poor	2
3	2.0	Red	Poor	7
4	2.1	Red	Poor	15

The above information helps us to understand the relation between Aggregate rating, color and text. We conclude the following color assigned to the ratings:

Rating 0 — White — Not rated

Rating 1.8 to 2.4 — Red — Poor

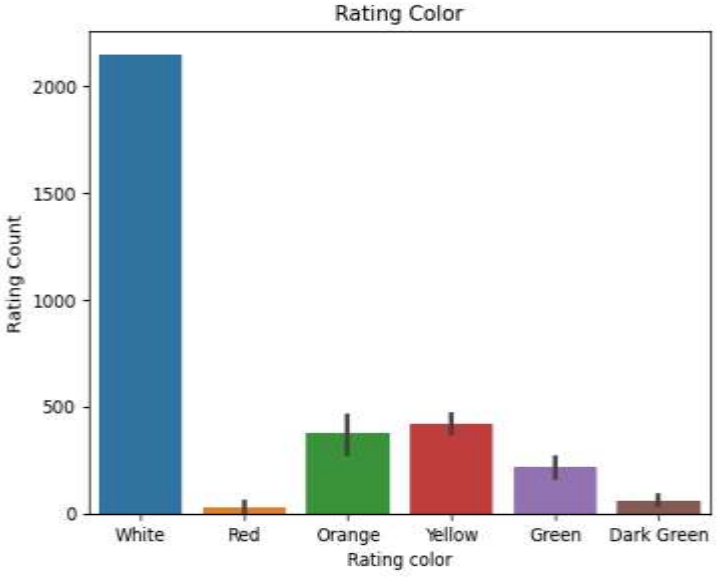
Rating 2.5 to 3.4 — Orange — Average

Rating 3.5 to 3.9 — Yellow — Good

Rating 4.0 to 4.4 — Green — Very Good

Rating 4.5 to 4.9 — Dark Green — Excellent

Let us try to understand the spread of rating across restaurants



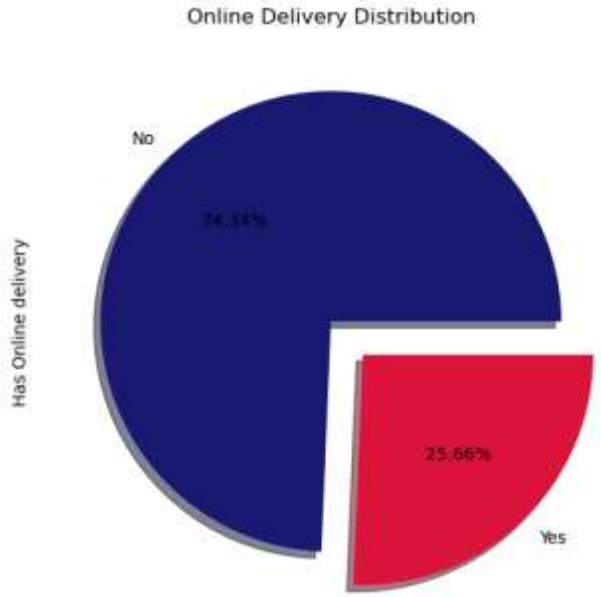
Interesting, Maximum restaurants seems to have gone No ratings.

Descriptive Analysis

Let us check if these restaurants belong to some specific country

```
No_rating = df2[df2['Rating color']=='White'].groupby('Country').
No_rating
```

	Country	Rating Count
0	Brazil	5
1	India	2139
2	United Kingdom	1
3	United States	3



India seems to have maximum unrated restaurants.

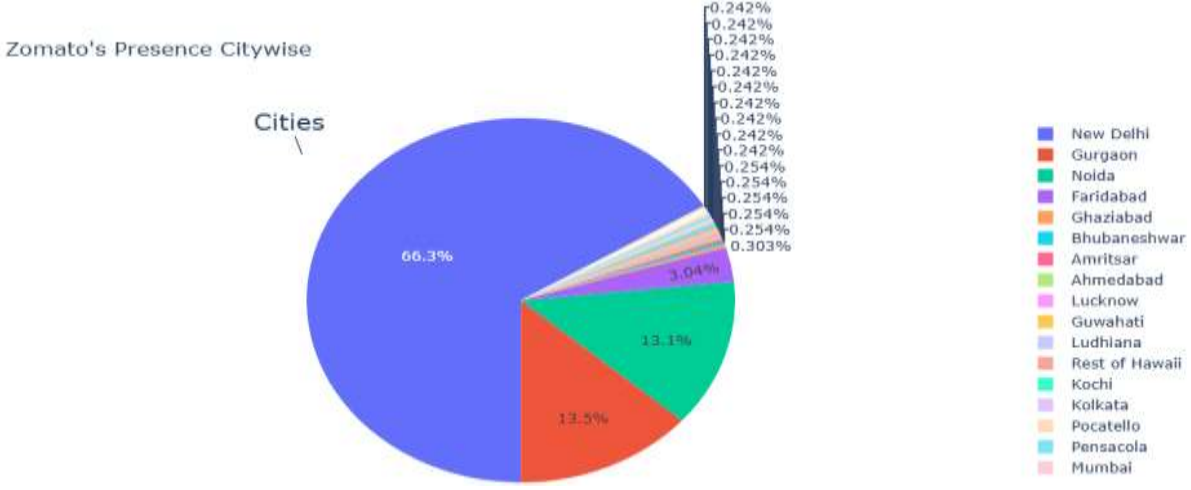
In India the culture of ordering online food is still gaining momentum hence most

of the restaurants are still unrated on Zomato as people might be preferring to visiting the restaurant for a meal..

Only 25% of restaurants accepts online delivery. This data might be biased as we have maximum restaurants listed here are from India. Maybe analysis over city wise would be more helpful.

Descriptive Analysis

Let us try to understand the coverage of city



The data seems to be skewed towards New Delhi, Gurgaon and Noida. I see minimal data for other cities.

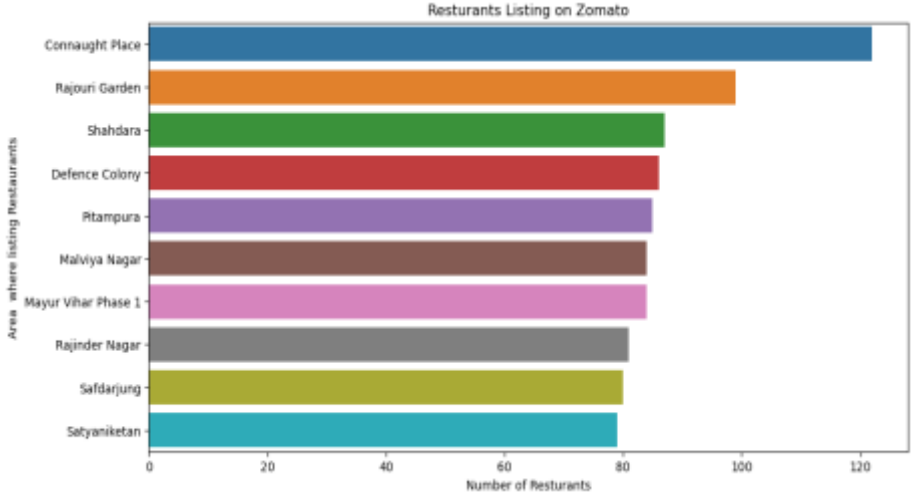
Hence I would do my analysis predominantly on New Delhi.

We've already gained several insights about the restaurants present in the survey.

Let us try some visualization and operation using this data....to get more information...about the restaurants...

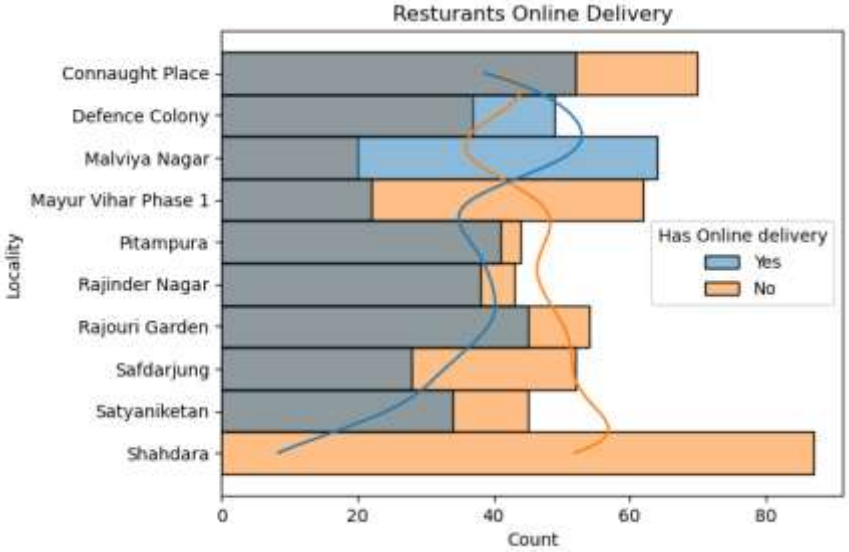
Descriptive Analysis

Lets us try some experiment how the Zomato is being spread throughout the different section.....
Locality having maximum hotels are listed in Zomato



Connaught place seems to have high no of restaurants registered with Zomato,

In Delhi high rated restaurant except online delivery or not.....

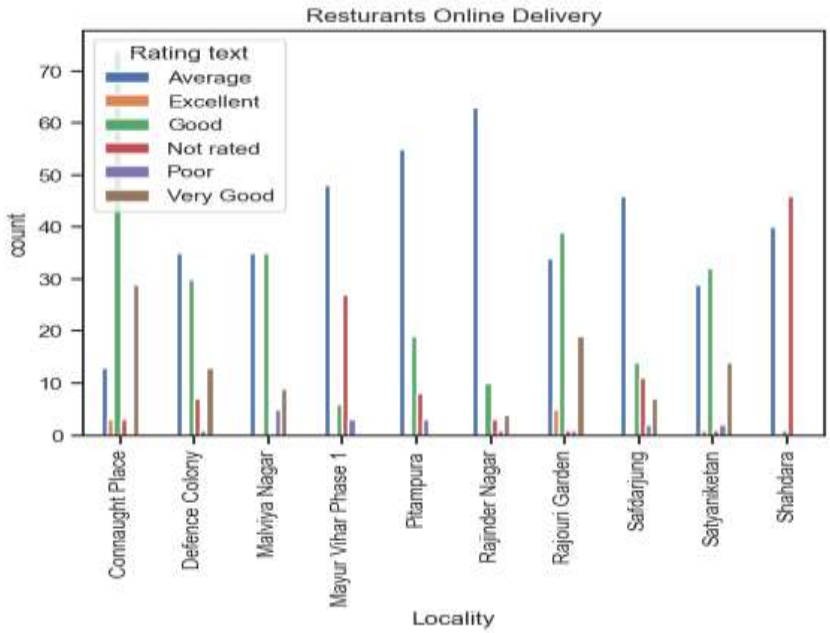


Online Delivery is high in Defence Colony and Malviya nagar whereas Shahdara don't except the online delivery

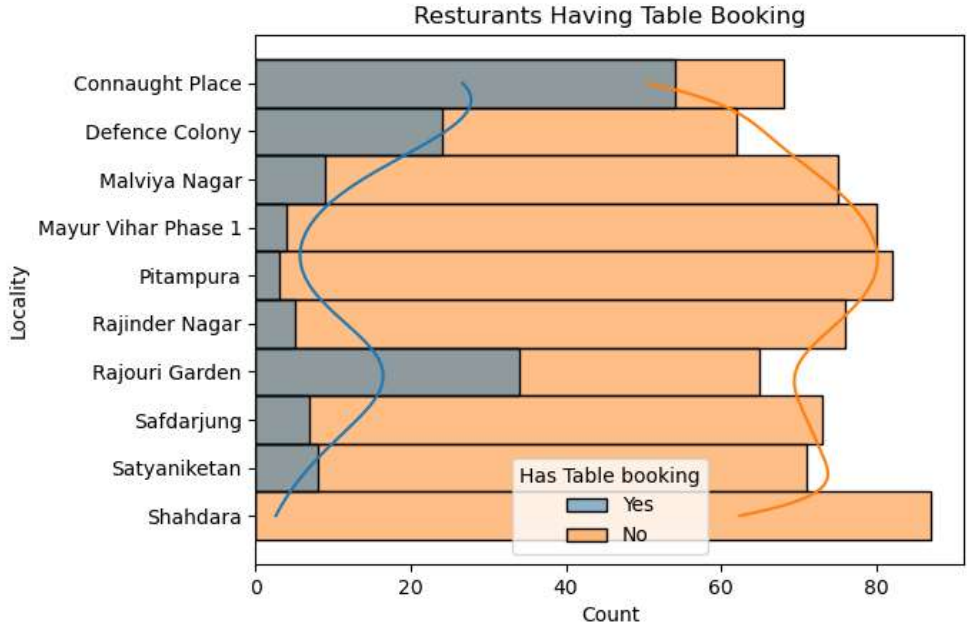
Descriptive Analysis

Lets us understand the Restaurants Rating localities..

I would now like to understand the rating of these restaurants that are providing online delivery .

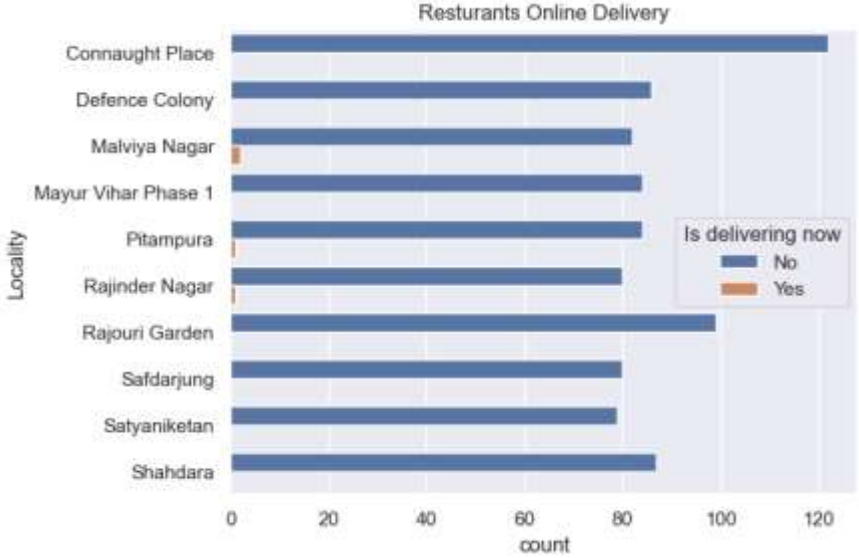


Restaurant Having Table Booking or Not



Descriptive Analysis

Restaurant Is Delivering Now

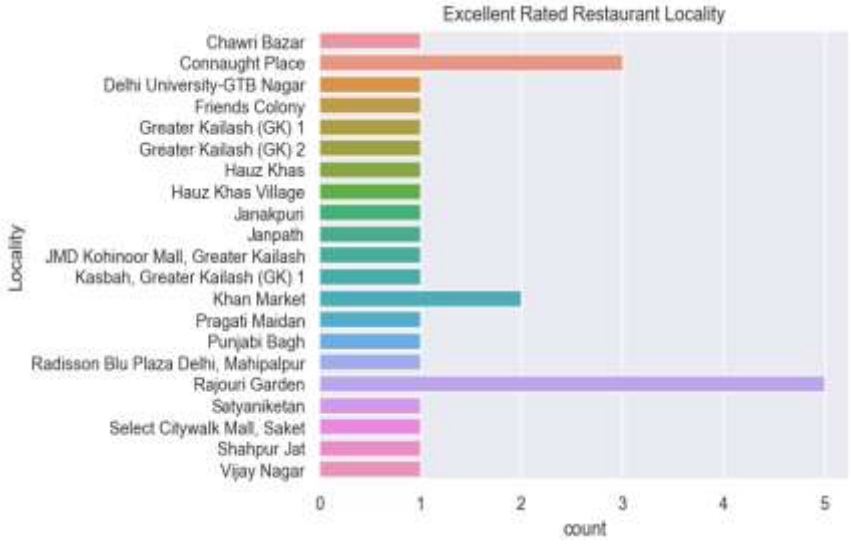


Let us understand the cost of Dinning vs their Aggregate Rating

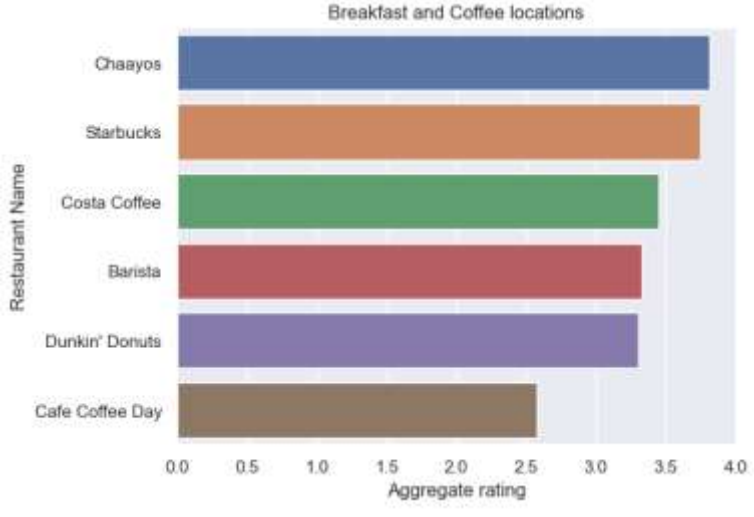
I observe there is no linear relation between price and rating. For instance, Restaurants with good rating (like 4–5) have restaurants with all the price range and spread across the entire X axis

Descriptive Analysis

Location of Highly rated restaurants across New Delhi



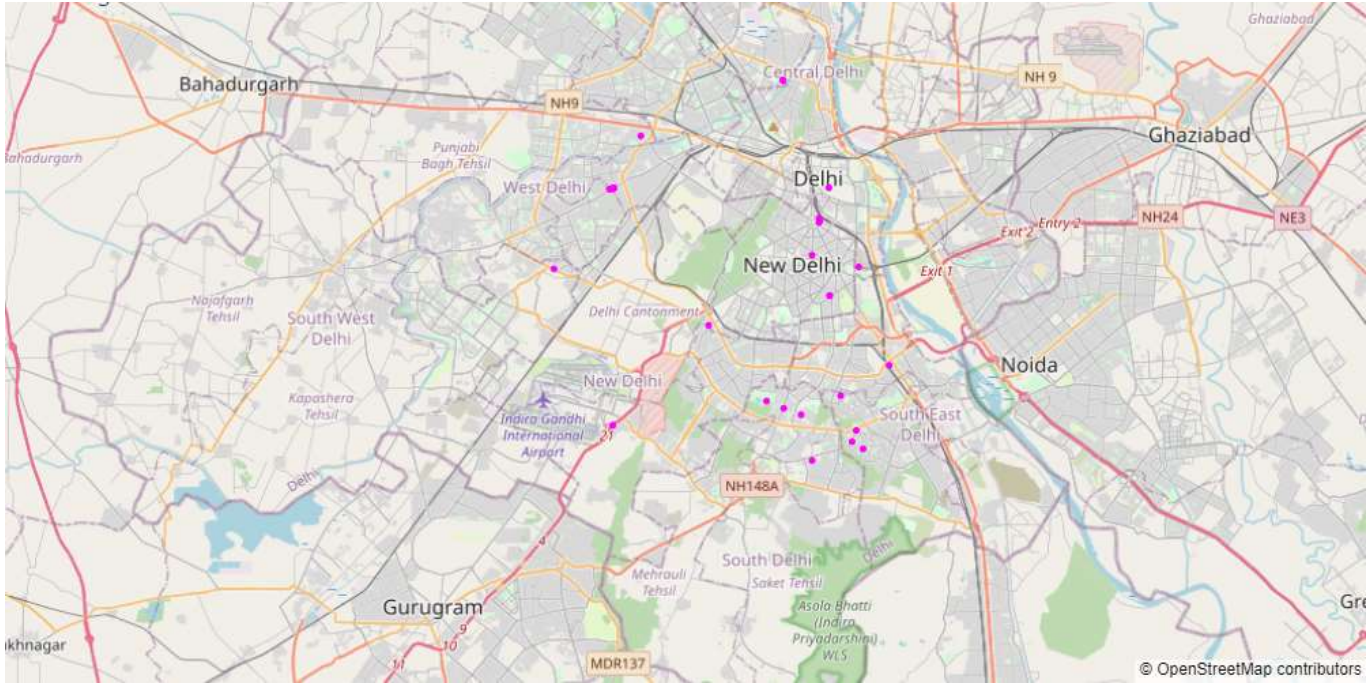
Common Eateries
Breakfast and Coffee locations



The aforementioned four cities represent nearly 65% of the total data available in the dataset. Apart from the higly rated local restaurants, it'd be intersting to know where the known-eateries that are commonplace. The verticles across which these can be located are -
Breakfast
American Fast Food
Ice Creams, Shakes & Desserts

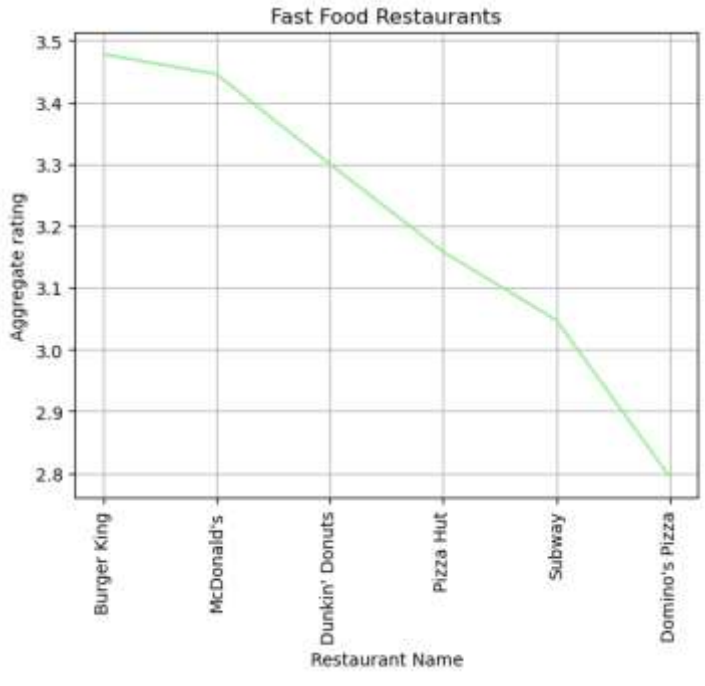
Descriptive Analysis

Location of Highly rated restaurants across New Delhi …We can see visually by using map....

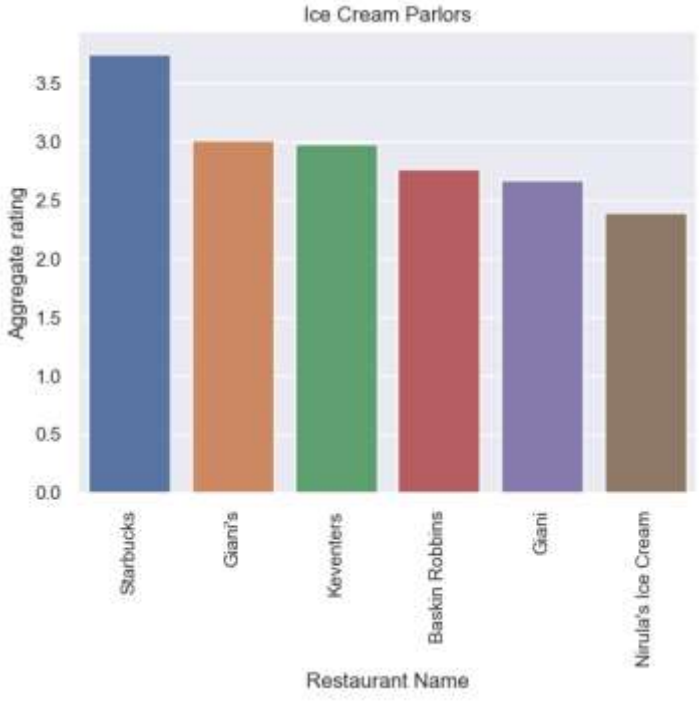


Descriptive Analysis

Fast Food Restaurants



Ice Cream Parlour



Conclusion

We've drawn many inferences from the survey. Here's a summary of a few of them:

The dataset is skewed towards India and doesn't represent the complete data of restaurants worldwide.
Restaurants rating is categorized in categories
Not Rated
Average
Good
Very Good
Excellent

Connaught Palace have maximum restaurants listed on Zomato but in terms of online delivery acceptance Defence colony and Malviya nagar seems to be doing better.

The top rated restaurants seems to be getting better rating on the following cuisine

North Indian
Chinese
American
Italian

There is no relation between cost and rating. Some of the best rated restaurants are low on cost and vice versa.
On common Eateries, For Breakfast and Coffee location Indian restaurants seems to be better rated but for Fast food chain and Ice cream parlors American restaurants seems to be doing better.

Future Scope

Future of Zomato in India.....

- As we all know Zomato is the Indian Company…And this data give Large information about India …So as Zomato is growing company its future is quite bright
- It helps people to get food on doorsteps in this busy scheduled life
- Here's why Zomato works well in these countries and it's hard for someone like Yelp to beat it: Cheap manual labor: The Zomato model needs a lot of paid manual labor. Countries with cheaper labor costs like India are the ones where Zomato has succeeded till now.

"Food has no religion" basic moto of Zomato”

- First mover advantage – One of the best competitive advantages of Zomato is that it is the first mover in many of the nations where it is establishing itself. Directories and other forms of restaurant ratings might exist.
- Convenience to use
- Healthier Delivery Options

Coins has two sides.. we have to look both sides pros and cron …

- Customers often back out from ordering their food after looking at the additional taxation and delivery charges. This practice becomes a disadvantage for this food delivery app
- Killing the Vibe.
- More Expensive
- Isolated Disconnect.

Zomato has traveled a long way to become one of the success stories among Indian startups in 2021, with almost 50% share in the food delivery market.

Your Food is Ready