**ZOMATO DATA ANALYSIS**

**INTRODUCTION:**

Zomato is an Indian restaurant aggregator and food delivery start-up founded by Deepinder Goyal and Pankaj Chaddah in 2008. Zomato provides information, menus and user-reviews of restaurants, and also has food delivery options from partner restaurants in select cities.

The basic idea of analyzing the Zomato dataset is to get a fair idea about the factors affecting the aggregate rating of each restaurant, establishment of different types of restaurant at different places, Bengaluru being one such city has more than 12,000 restaurants with restaurants serving dishes from all over the world. With each day new restaurants opening the industry has'nt been saturated yet and the demand is increasing day by day. Inspite of increasing demand it however has become difficult for new restaurants to compete with established restaurants. Most of them serving the same food. Bengaluru being an IT capital of India. Most of the people here are dependent mainly on the restaurant food as they don't have time to cook for themselves. With such an overwhelming demand of restaurants it has therefore become important to study the demography of a location. What kind of a food is more popular in a locality. Do the entire locality loves vegetarian food. These kind of analysis can be done using the data, by studying different factors.

**GOAL:**

The basic idea is analyzing the Buisness Problem of Zomato to get a fair idea about the factors affecting the establishment of different types of restaurant at different places in Bengaluru, aggregate rating of each restaurant and many more.

**MOTIVATION:**

India is quite famous for its diverse multi cuisine available in a large number of restaurants and hotel resorts, which is reminiscent of unity in diversity. Restaurant business in India is always evolving. More Indians are warming up to the idea of eating restaurant food whether by dining outside or getting food delivered. The growing number of restaurants in every state of India has been a motivation to inspect the data to get some insights, interesting facts and figures about the Indian food industry in each city. So, this project focuses on analysing the Zomato restaurant data for each city in India.

The Project focuses on Customers and Company, we have analyzed the data and made some useful conclusion in the form of Visualizations. The data is vizualized as it becomes easy to analyse data at instant. The Analysis also solve some of the business cases that can directly help the customers finding the Best restaurant in their locality and for the company to grow up and work on the fields they are currently lagging in.

**CONCEPTUAL IDEA:**

* The Project focuses on Customers and Company, we have analyzed the data and made some useful conclusion in the form of Visualizations. The data is vizualized as it becomes easy to analyse data at instant. The Analysis also solve some of the business cases that can directly help the customers finding the Best restaurant in their locality and for the company to grow up and work on the fields they are currently lagging in.

**PRACTICAL IMPLICATIONS:**

Practiced this Data visualization on Zomato Dataset only for educational purpose for practising.

**DATA SOURCE:**

The data is taken from Kaggle named as Zomato Bangalore Dataset.

**INPUT DATA:**

The data in this dataset contains 51717 rows and 17 columns.

**Data Description**

url : contains the url of the restaurant in the zomato website

addres: contains the address of the restaurant in Bengaluru

name: contains the name of the restaurant

online\_order: whether online ordering is available in the restaurant or not

book\_table: table book option available or not

rate: contains the overall rating of the restaurant out of 5

votes: contains total number of rating for the restaurant as of the above mentioned date

phone: contains the phone number of the restaurant

location: contains the neighborhood in which the restaurant is located

rest\_type: restaurant type

dish\_liked: dishes people liked in the restaurant

cuisines: food styles, separated by comma

approx\_cost(for two people): contains the approximate cost for meal for two people

reviews\_lis: list of tuples containing reviews for the restaurant, each tuple

menu\_item contains: list of menus available in the restaurant

listed\_in(type): type of meal

listed\_in(city): contains the neighborhood in which the restaurant is listed

**METHODOLOGY**

In this section we are going to walk through the life cycle of a data analysis project.



**DATA CLEANING, & PRE-PROCESSING**

The next step is data cleaning process. This involves deletion of unnecessary and redundant features of the data.

Data Cleaning is the process of ensuring that your data is correct, consistent and useable by identifying any errors or corruptions in the data, correcting or deleting them, or manually processing them as needed to prevent the error from happening again. Data given to us was containing lot of errors like Null values, White Spaces, Hyphen, Latin Letters, Incorrect value. We have used Python Library Pandas for Data Cleaning purpose.

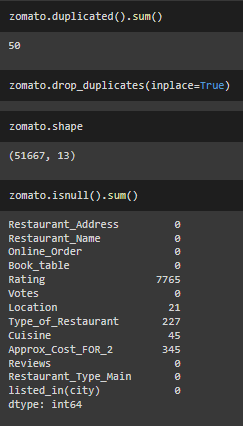
* Importing the data as Data Frame
* Explore the data and check for errors that make data Inefficient for further analysis
* Check the datatype and convert as per requirement
* Check for Null Values
* Replace or Drop Null values

For data analysis, we do not need the contact details of the restaurant so, the URL column is removed along with the Page Number column which is also of no use for the analysis.

Moreover, the timing column is not providing any additional information which may be used for the analysis since only the opening and closing time of the restaurant is specified so, this column is removed as well.

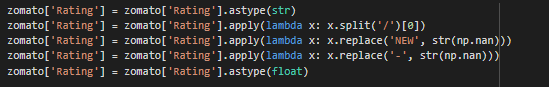


Similarly, we check the duplicate and the null values present in dataset or not. We can see there are 50 duplicate values and around 8000 null values so we drop it as our dataset is also big enough.



As you can see the rating column is string type with an extra /5 with all the ratings. This should be cleaned. We found this column has 'NEW' ,"-" values which should replaced by np.nan.

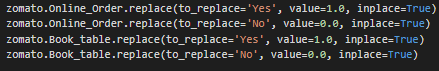
It is important to convert the string back to float because rating cannot be string value.



The Restaurant\_Name has some giberish symbols such as (©,Â) . This should also be cleaned.



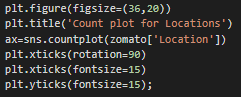
Online\_order & Book\_table are presented as strings instead of booleans (working with True/False make it easier to analyze)

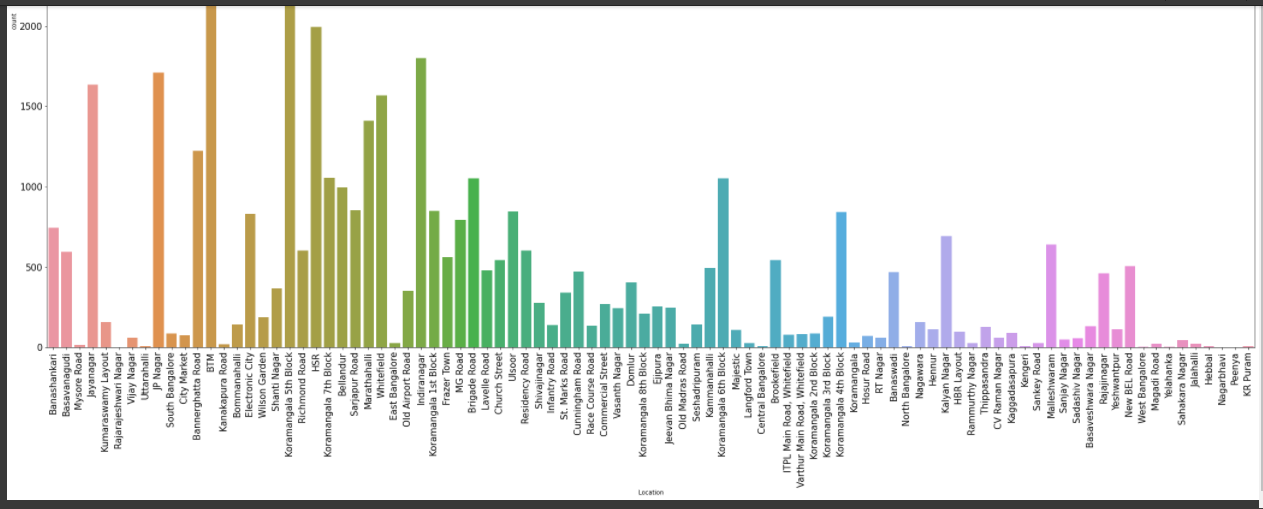


# **DATA VISUALIZATION**

Here in Data Visualization, we are going to analyze some queries or some questions which may arise and help to visualize business problems of Zomato through some graphs.

Q1. Count Plot for Various Locations

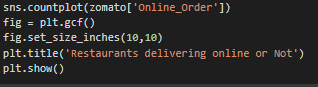


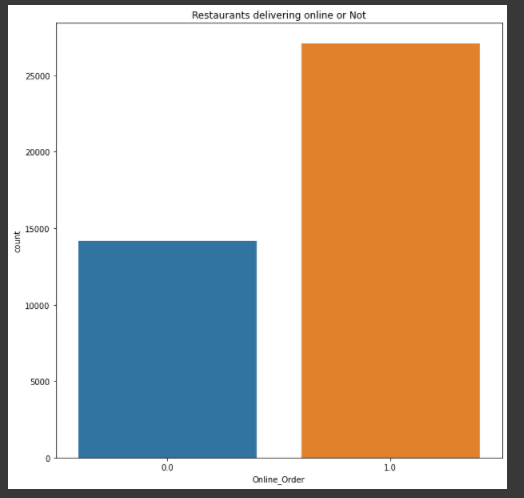


In this count plot, we can see maximum number of restaurants are present in BTM location followed by other locations.

So from this we come to point that if we want to open a restaurant in a location we will choose that location which is having less number of restaurants like Old Madras Road or Langford Town or Uttarahali,etc. So that there will be more profit as there are less number of restaurants.

Q2. Restaurants providing online delivery facility or not

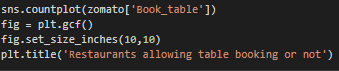


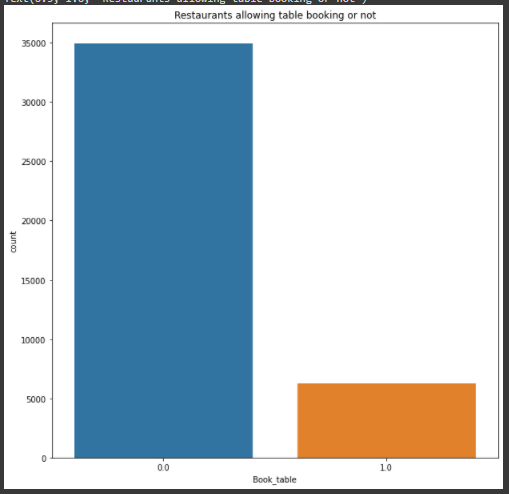


(0 is for No, 1 is for Yes)

We can see here that the number of restaurants providing online delivery facility are more than that restaurants which do not provide this facility.

Q3. Restaurants allowing table booking or not





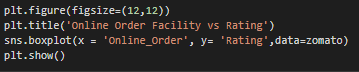
According to this graph, we can see there are less reataurants that provide online table booking.

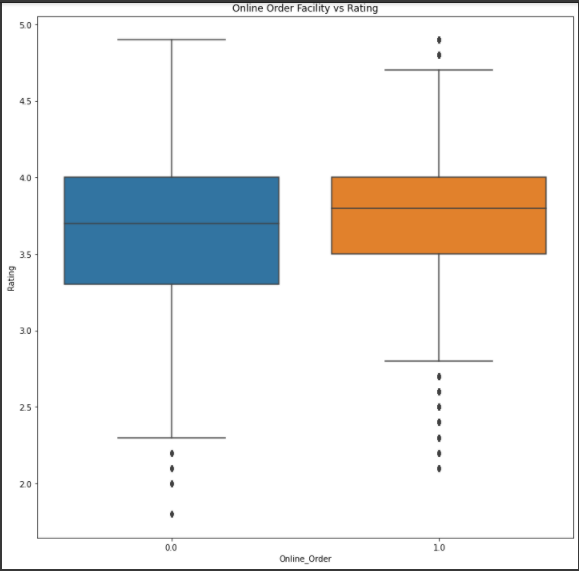
Q4. Visualizing Online Order Vs. Rating





Extracting the Rating and Online\_Order column from Dataframe for analysis.





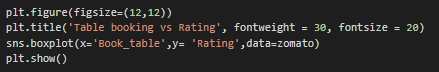
Here, we can analyze that the restaurants that are having online order facility ,their rating is lesser than the restaurants that are not having online order facility.

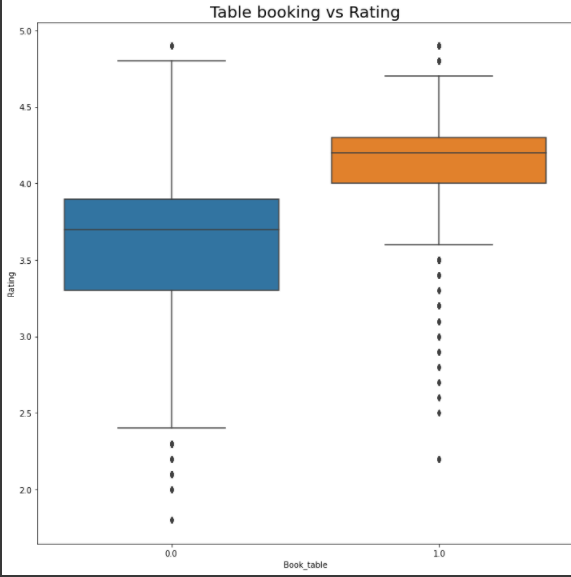
Maximum rating for restaurants with online order facility is 4.7 and that for not having facility is 4.9.

Q5. Visualizing Table booking vs Rating





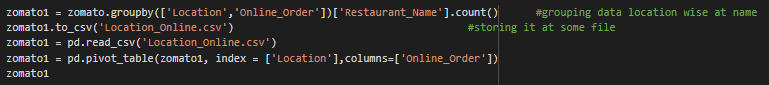


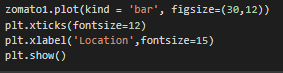


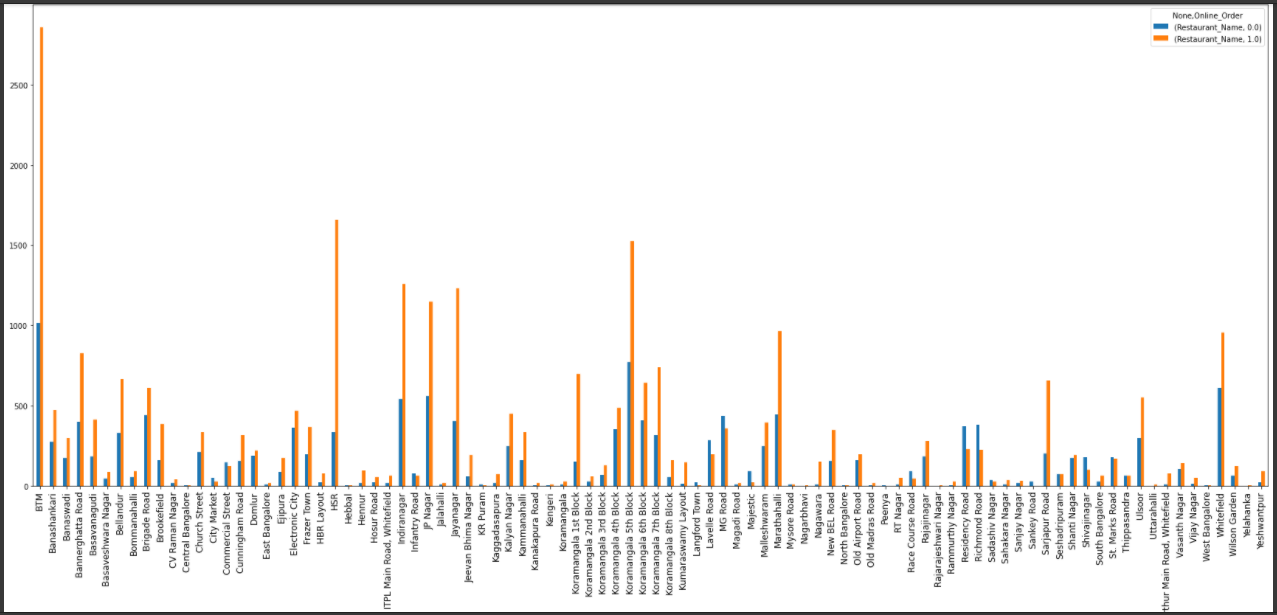
(1.0=Yes,0.0=No)

Restaurants that are having booking table facility, their average rating is lower than that who are not having the facility.

Q6. Visualizing Online Order Facility, Location Wise





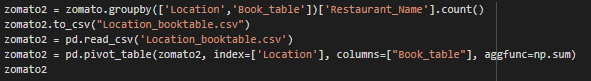


(orange=yes,blue=no)

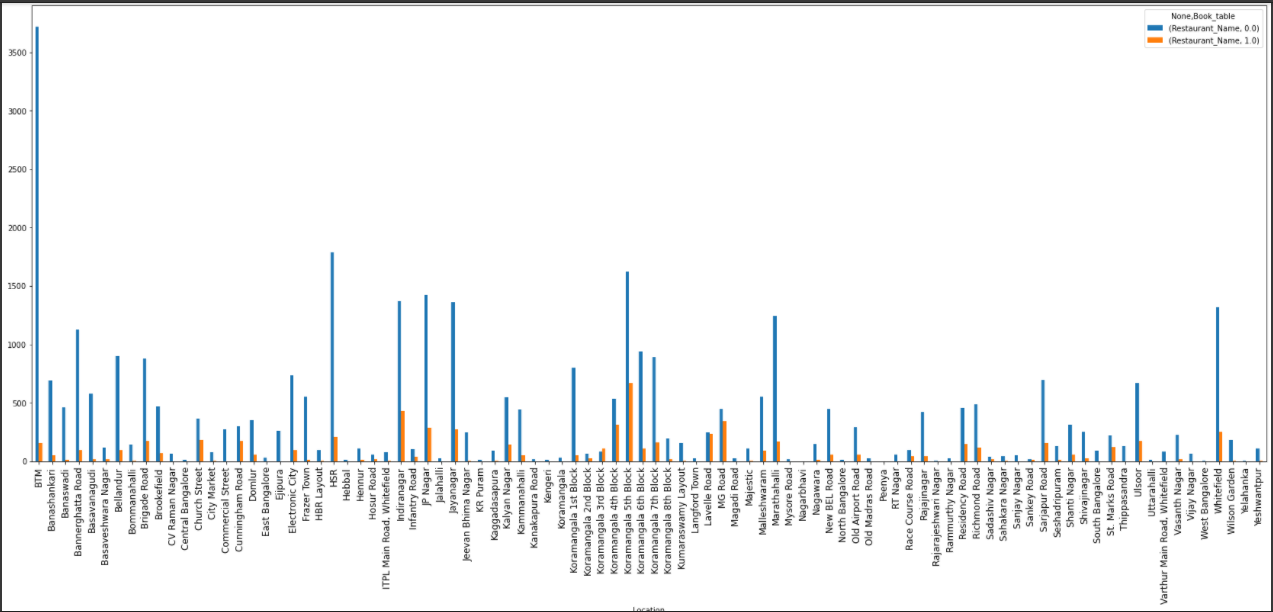
In BTM orange line is highest i.e. there are many restaurants that provide online facility.

But, in Location like Kanakapura Road or Mysore Road,etc. the blue line is smaller.

Q7. Visualizing Book Table Facility, Location Wise

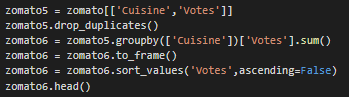






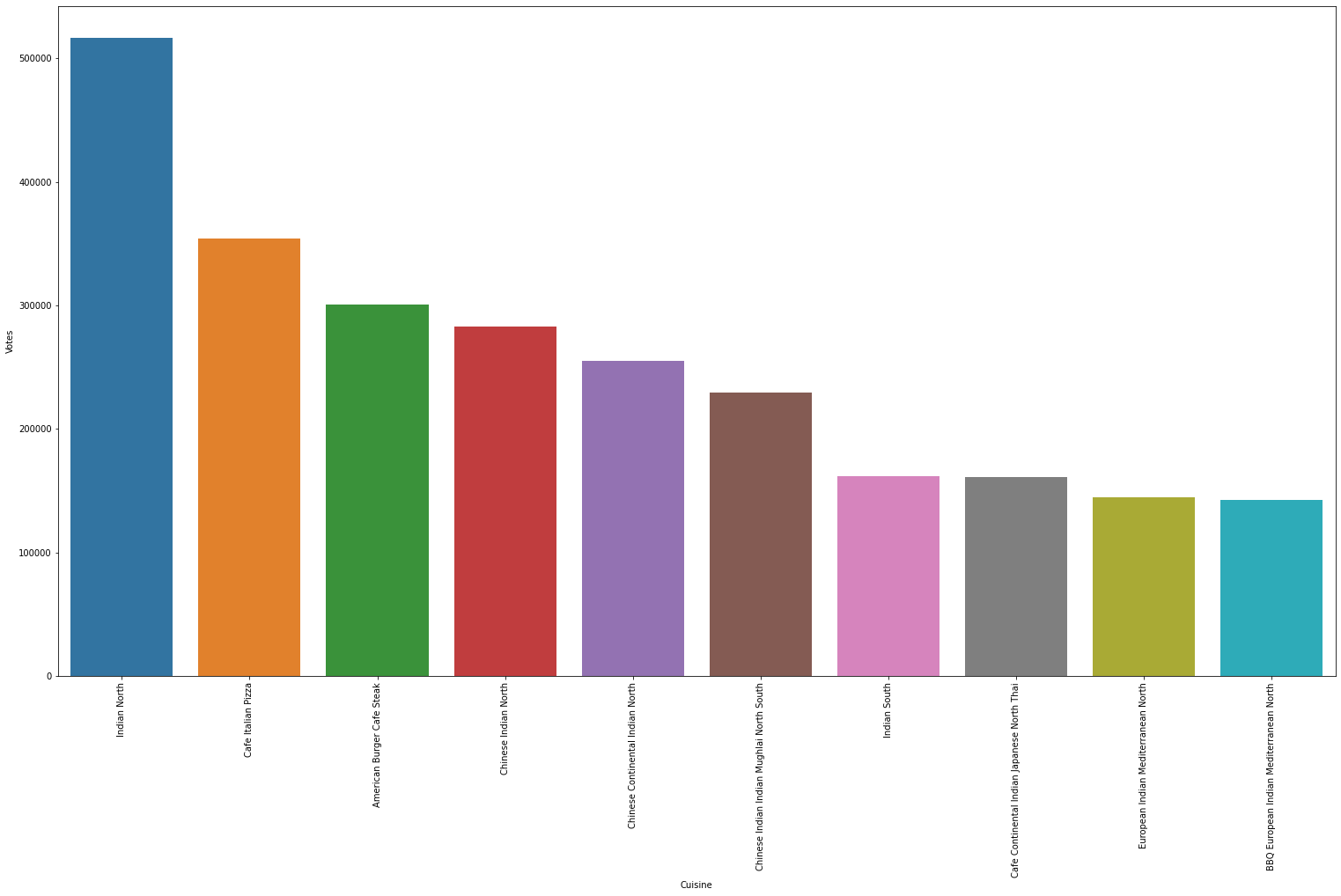
According to this barplot, we can see in BTM many restaurants are those who do not provide book table, so if someone want to open restaurant he can open in BTM with book table facility. But, BTM already have more restaurants so he can try opening in HSR where there are less restaurants so business may grow.

Q8. Visualizing Top Cuisines







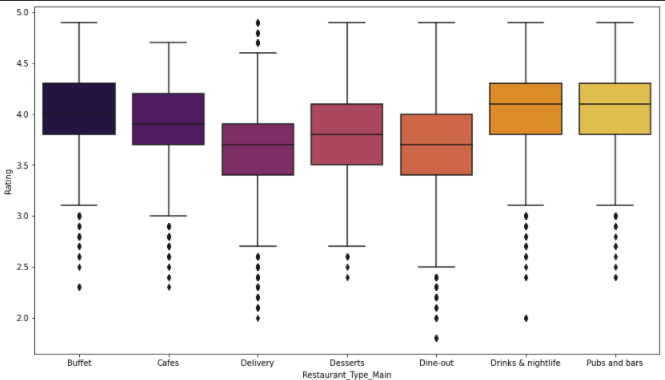


Here, we can see the top 10 cuisines that are top rated and are famous in Bangalore .

In this, North Indian Cuisine is rated highly followed by Cafe Italian Pizza and then American Burger Cafe Steak.

Q9. Type of Meal and their Rating





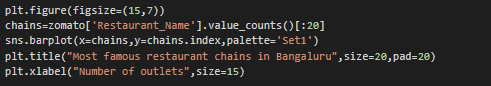
We can see that maximum average rating is given to Buffet type of meal.

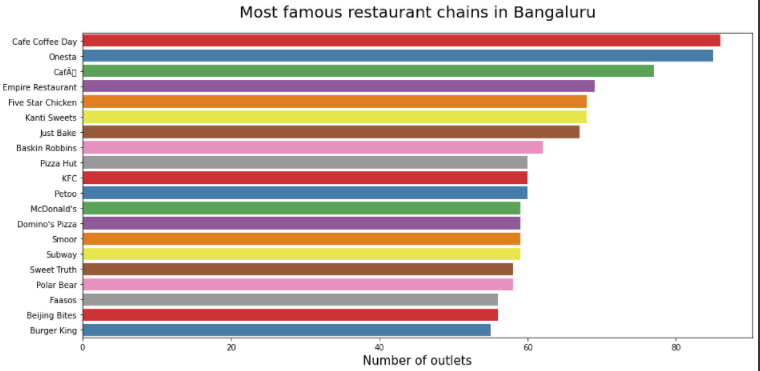
This type of restaurant are gievn maximum rating.

That means, peolpe are coming in this type of restaurant and have rated highest.

And also the Delivery type of restaurant is having lowest average rating.

Q10. Most famous Restaurant chains in Bengaluru

****

****

In Bangalore, most famous restaurant chain is CCD (Cafe Coffee Day).

**CONCLUSION**

After all the data processing and analysis, we can get this following conclusion: There are 51717 restaurants registered on Zomato Apps based in Bangalore. More than 3500 restaurants are located in BTM which is highest amongst the all in Bangalore. There are high number of restaurants that provide Online Order facility comparatively there are less number of restaurants that provide booking table facility. Also North Indian is the cuisine most famous in Bangalore. Restaurants which provide Buffet system are high rated . If we look for most famous restaurant chain in Bangalore, Cafe Coffee Day (CCD) stands at the top.

**REFERENCES**

[1] Blog on Zomato Data Analysis By Chandresh~ <https://chandresh.me/Zomato_data-Analysis/>

[2] From some Github repositories.

[3] Also from Kaggle notebooks.