Zomato Bangalore Data Analysis with Python

Group 4

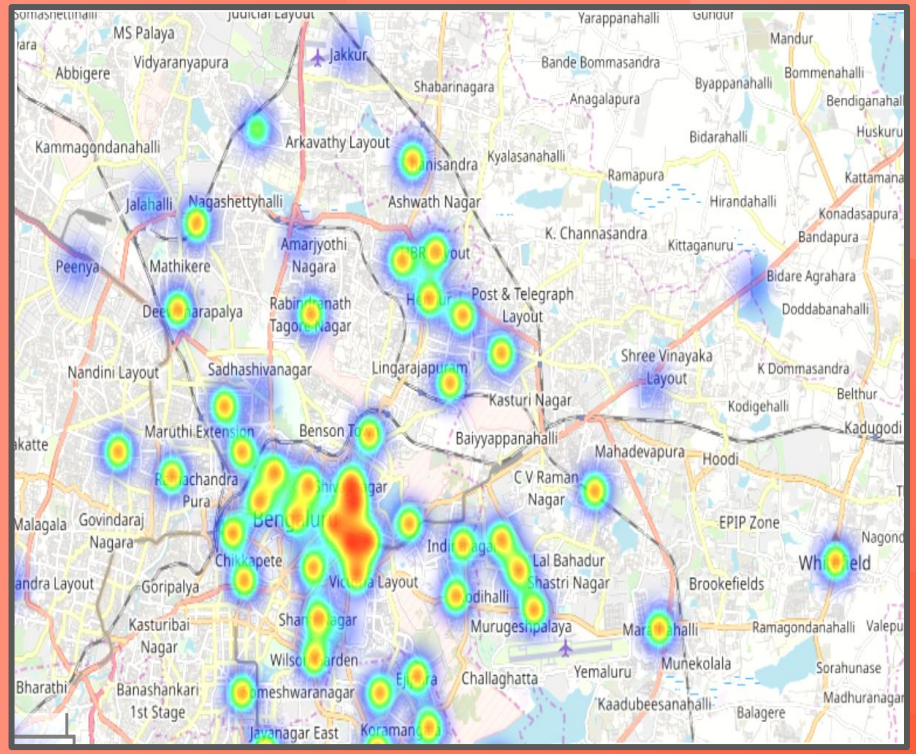
Adarsh Salapaka

Jaskaran Kohli

Sanjana Ramankandath

Vidhi Gandhi

Overview



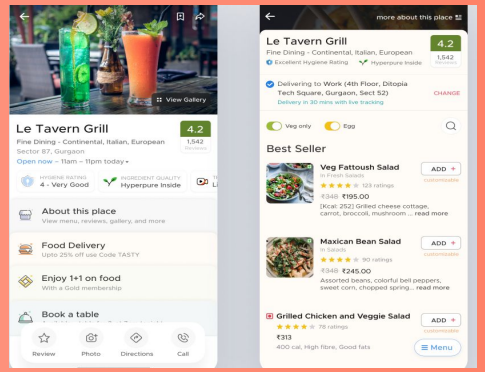
Data Source:
https://www.kaggle.com/himanshupoddar/zomato-bangalore-restaurants

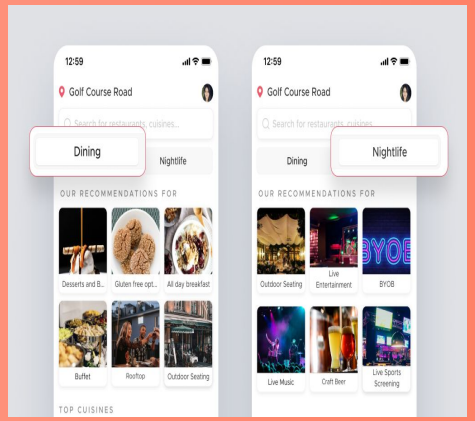
Bangalore is the Silicon Valley of India. Delivery, Dine-out, Pubs, Bars, Drinks, Buffet, Desserts you name it and Bengaluru has it.

Bengaluru is best place for foodies. The number of restaurant are increasing day by day. Currently which stands at approximately 12,000 restaurants.

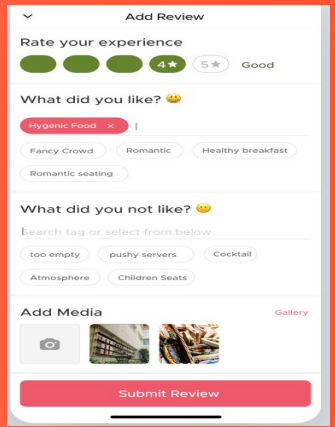
This Zomato data aims at analysing demography of the location. Most importantly it will help food entrepreneurs in deciding their theme, menus, cuisine, cost etc for a particular location.

Data Description





- **url** : contains the url of the restaurant in the zomato website
- **address :** 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_list** : 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



Data Cleaning

	url	address	name	online_order	book_table	rate	votes	phone	location	rest_type	dish_liked	cuisines	approx_cost(for two people)	reviews_list	menu_item	listed_in(type)	listed_in(city)
0	https://www.zomato.com/bangalore/jalsa-banasha...	942, 21st Main Road, 2nd Stage, Banashankari, ...	Jalsa	Yes	Yes	4.1/5	775	080 42297555\r\n+91 9743772233	Banashankari	Casual Dining	Pasta, Lunch Buffet, Masala Papad, Paneer Laja...	North Indian, Mughlai, Chinese	800	[('Rated 4.0', 'RATED\n A beautiful place to ...	[]	Buffet	Banashankari
1	https://www.zomato.com/bangalore/spice-elephan...	2nd Floor, 80 Feet Road, Near Big Bazaar, 6th ...	Spice Elephant	Yes	No	4.1/5	787	080 41714161	Banashankari	Casual Dining	Momos, Lunch Buffet, Chocolate Nirvana, Thai G...	Chinese, North Indian, Thai	800	[('Rated 4.0', 'RATED\n Had been here for din...	[]	Buffet	Banashankari

- `Approx_Cost_FOR_2` has a high cardinality: 71 distinct values
- `Contact` has 1208 (2.3%) missing values
- `Contact` has a high cardinality: 14927 distinct values
- `Cuisine` has a high cardinality: 2724 distinct values
- `Location` has a high cardinality: 94 distinct values
- `Menu` has a high cardinality: 9098 distinct values
- `Most_Liked_Dishes` has 28078 (54.3%) missing values
- `Most_Liked_Dishes` has a high cardinality: 5272 distinct values
- `Rating` has 7775 (15.0%) missing values
- `Rating` has a high cardinality: 65 distinct values
- `Restaurant_Address` has a high cardinality: 11495 distinct values
- `Restaurant_Name` has a high cardinality: 8792 distinct values

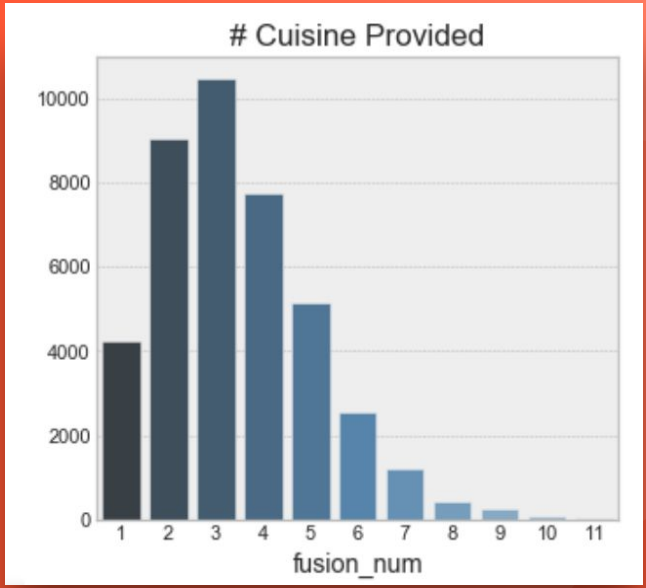
Data Cleaning Summary

- Renaming the columns.
- Deleting redundant columns.
- Replacing Gibberish Symbols ((©,Â)).
- Changing the Data Types for Analysis
- Dropping duplicates.
- Converting Strings to Bool .
- Cleaning individual columns.

- `Cuisine` has a high cardinality: 1769 distinct values
- `Location` has a high cardinality: 92 distinct values
- `Rating` has 2259 (5.2%) zeros
- `Restaurant_Name` has a high cardinality: 7112 distinct values
- `Reviews` has a high cardinality: 21575 distinct values
- `Type_of_Restaurant` has a high cardinality: 66 distinct values
- `Votes` has 2278 (5.2%) zeros

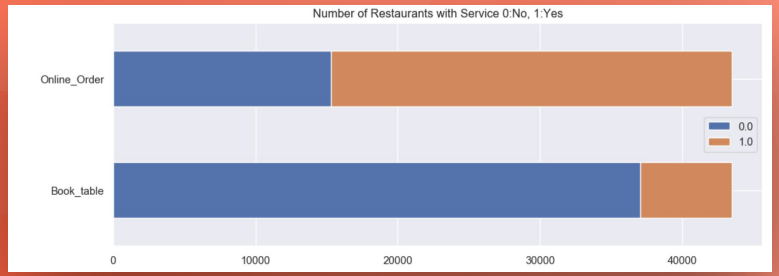
Restaurant_Name	Online_Order	Book_table	Rating	Votes	Location	Type_of_Restaurant	Cuisine	Approx_Cost_FOR_2	Reviews	Restaurant_Type_Main
Jalsa	1.0	1.0	4.1	775	Banashankari	Casual Dining	Chinese Indian Mughlai North	800	A beautiful place to dine in.The interiors tak...	Buffet
Spice Elephant	1.0	0.0	4.1	787	Banashankari	Casual Dining	Chinese Indian North Thai	800	Had been here for dinner with family. Turned o...	Buffet

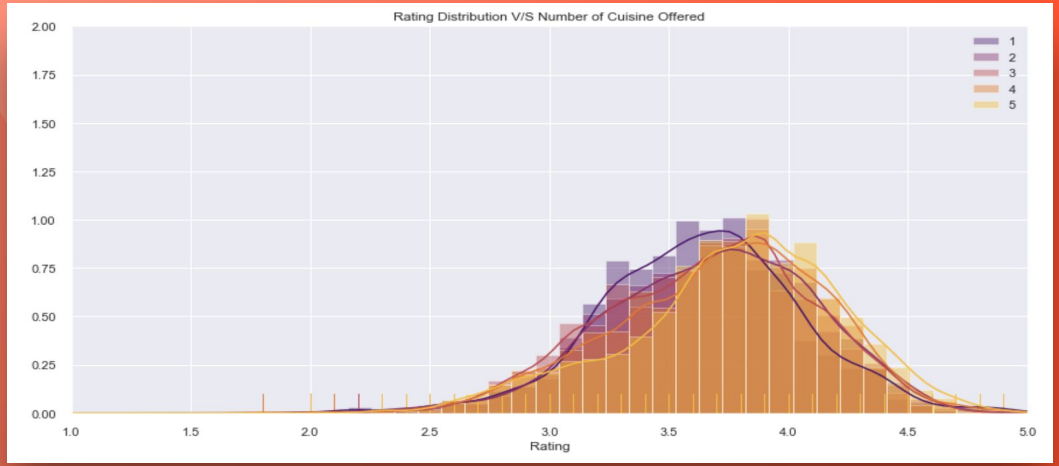
Higher the number of services/cuisines provided by a restaurant, more likely it is to receive higher ratings



Managerial Insight:

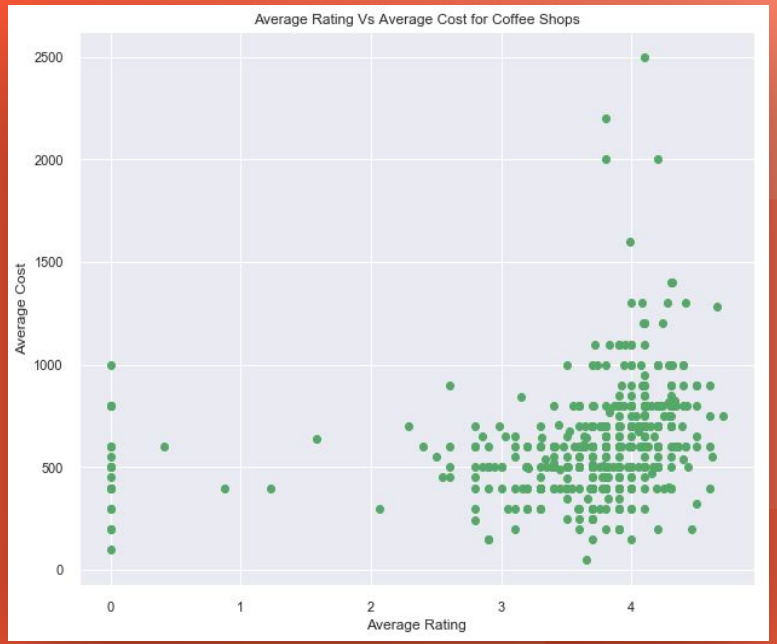
The finding is explained by the fact - Restaurants offering more number of services(online ordering/table booking facility) or multiple cuisines to customers, tend to have higher ratings irrespective of the location and price. This is probably because customers enjoy the additional ease, convenience and variety in their dining experience. We could use this finding as follows - it would be highly lucrative to serve more than a single cuisine and have both delivery and booking table options.




Rating Distribution V/S Number of Cuisine Offered
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Rating
1
2
3
4
5


Rating Distribution V/S Number of Services Offered (Book Table / Online Delivery)
1.5
1.0
0.5
0.0
0
1
2
3
4
5
Rating
Both N
Deli
Book
Both Y

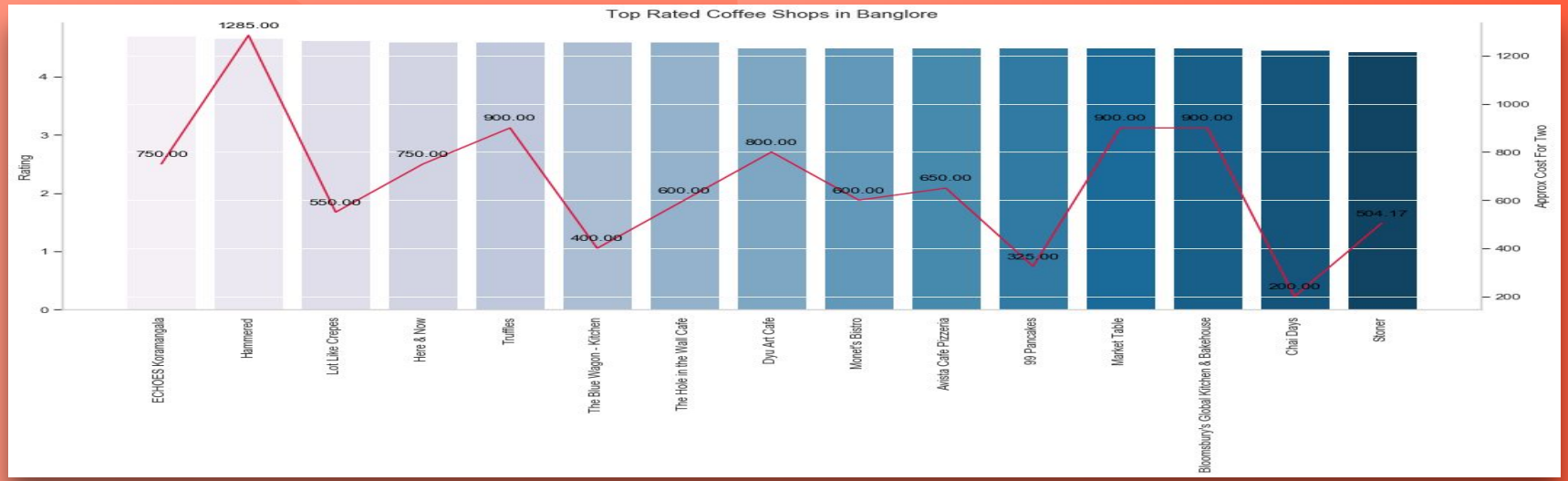
The higher the approximate Cost-for-2 for a cafe, the less likely it is to be popular

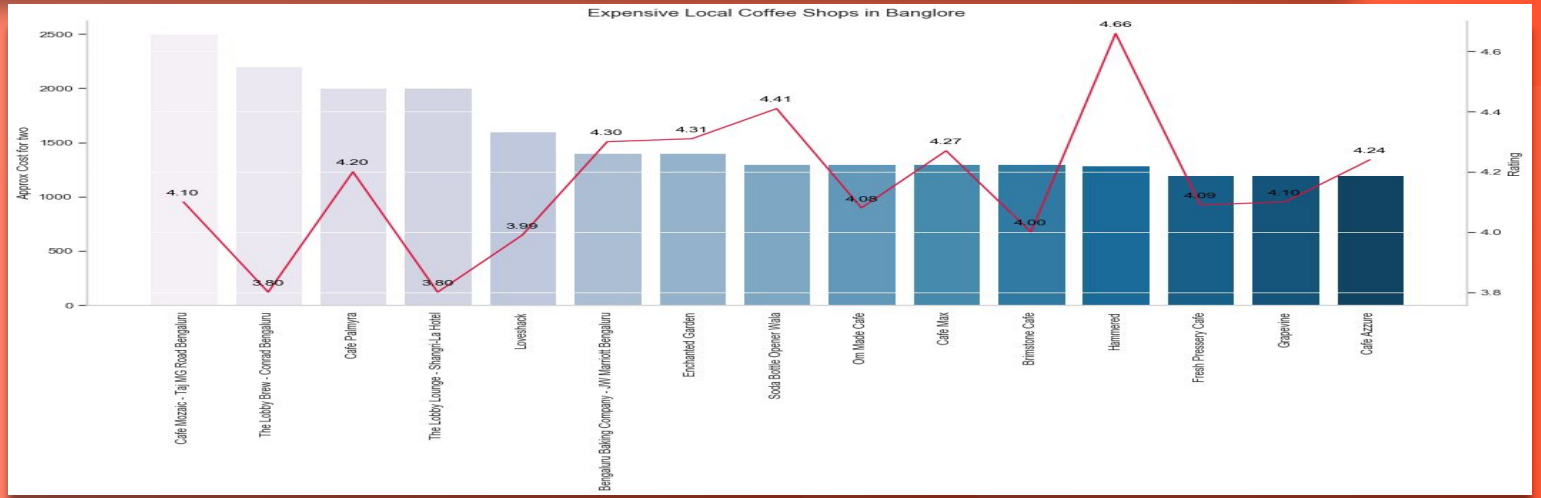


Managerial Insight:

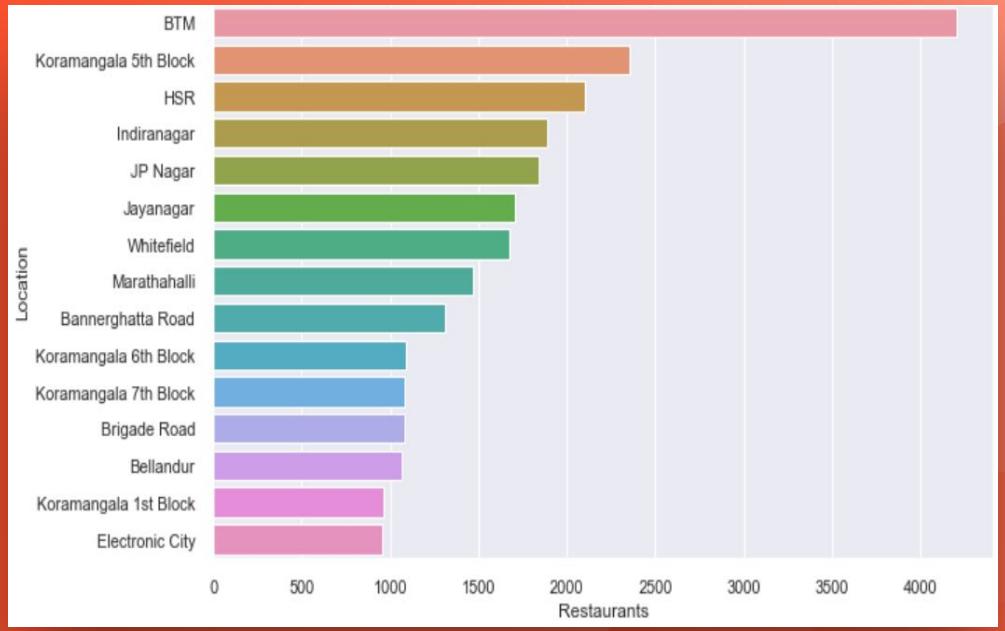
The finding is explained by the fact, the more expensive a cafe is, less popular it is amongst the customers. On the contrary, Cafes which are less expensive are more likely to be popular amongst customers.This is probably because cafes which are less expensive provide more wholesome customer experience than the expensive ones which are mainly concentrated in the Luxury hotels.

We could use this finding as follows - In order to open a profitable Cafe one should focus more on Customer experience and try to keep the pricing in the less expensive range.


Top Rated Coffee Shops in Banglore
Rating
Approx Cost For Two
750.00
1285.00
550.00
750.00
900.00
400.00
600.00
800.00
600.00
650.00
325.00
900.00
900.00
200.00
504.17
ECHOES Koramangala
Hammered
Lot Like Crepes
Here & Now
Truffles
The Blue Wagon - Kitchen
The Hole in the Wall Cafe
Dyu Art Cafe
Monet's Bistro
Arista Cafe Pizzeria
99 Pancakes
Market Table
Bloomsbury's Global Kitchen & Bakehouse
Chai Days
Stoner


Expensive Local Coffee Shops in Banglore
Approx Cost for two
Rating
4.10
3.80
4.20
3.80
3.99
4.30
4.31
4.41
4.08
4.27
4.00
4.66
4.09
4.10
4.24
Cafe Mocaic - Taj MG Road Bengaluru
The Lobby Brew - Conrad Bengaluru
Cafe Palmyra
The Lobby Lounge - Shangri-La Hotel
Loveshack
Bengaluru Baking Company - JW Marriott Bengaluru
Enchanted Garden
Soda Bottle Opener Wala
Om Made Cafe
Cafe Max
Brimstone Cafe
Hammered
Fresh Pressery Cafe
Grapevine
Cafe Azzure

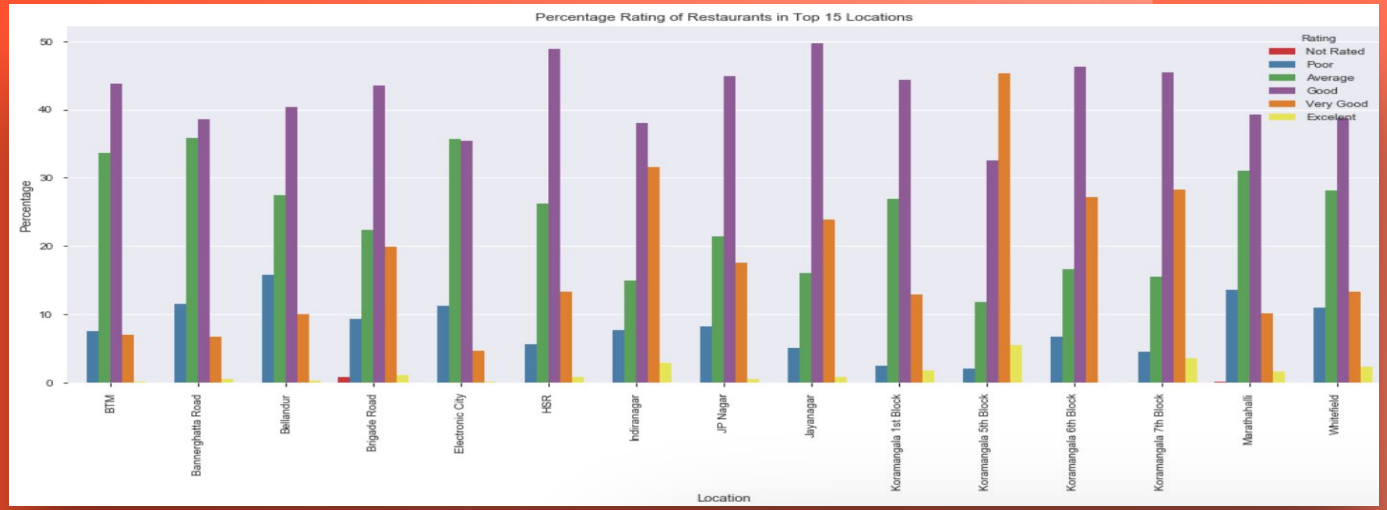
In high frequency restaurant locations,the category “ Drinks & Nightlife” has the most positive ratings

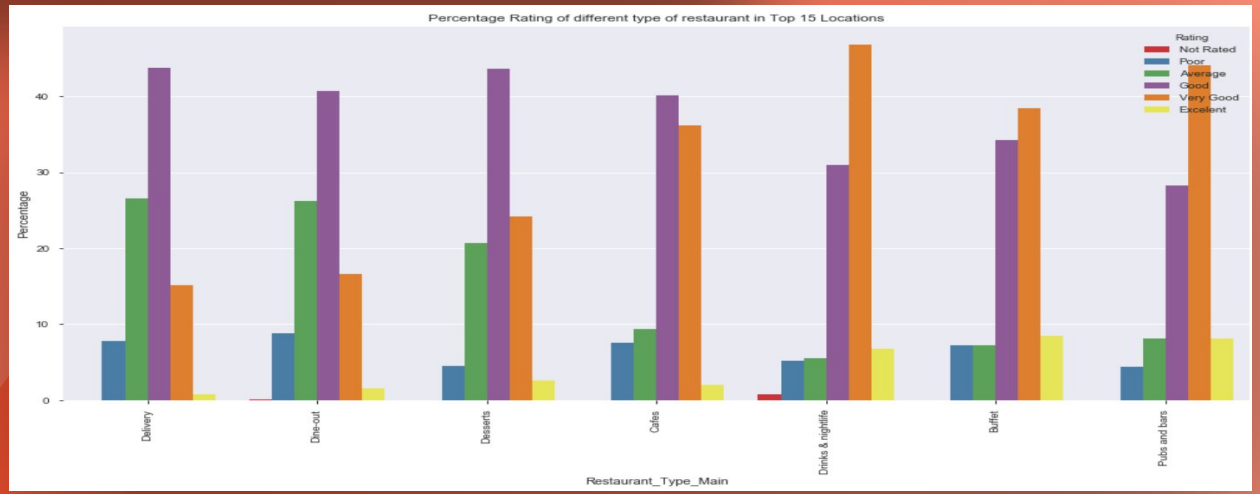


Managerial Insight:

The finding is explained by the fact that in the high frequency restaurant locations in Bangalore there was an overall positive rating, when compared to the low frequency restaurant locations. Further we noted that in the high frequency restaurant locations, **Drinks & Nightlife** had the greatest positive rating amongst all the restaurant types.

We could use this finding as follows - The location and the type of restaurant has an impact on the rating. In order to open a successful restaurant in the high frequency restaurant location, one could take this into consideration.


Percentage Rating of Restaurants in Top 15 Locations
Rating
Not Rated
Poor
Average
Good
Very Good
Excelent
Percentage
0
10
20
30
40
50
BTM
Bannerghatta Road
Bellandur
Brigade Road
Electronic City
HSR
Indiranagar
JP Nagar
Jayanagar
Koramangala 1st Block
Koramangala 5th Block
Koramangala 6th Block
Koramangala 7th Block
Marathahalli
Whitefield
Location


Percentage Rating of different type of restaurant in Top 15 Locations
Rating
Not Rated
Poor
Average
Good
Very Good
Excelent
Percentage
0
10
20
30
40
Delivery
Dine-out
Desserts
Cafes
Drinks & nightlife
Buffet
Pubs and bars
Restaurant_Type_Main

Additional Findings

- Using Random Forest we found the top 10 important variables impacting rating. Further, we were able to predict rating with an accuracy of 88%
- North Indian food is more popular in Bangalore: We observed through Location based analysis of North and South Indian cuisine that restaurants serving north indian cuisine are more popular than those serving south indian cuisine
- Finding best restaurant on the basis of approximate cost for two and ratings- We found the best restaurants both which were cheapest and highly rated vs expensive and highly rated
- Sentiment analysis on the basis of Restaurant review data