

Analysis of Zomato Bangalore Restaurant dataset

Zomato is an Indian restaurant aggregator and food delivery start-up founded in 2008. Zomato provides information, menus and user-reviews of restaurants as well as food delivery options from partner restaurants in selected cities.

Dataset

The dataset used for this project was found on Kaggle. The basic idea of analyzing the Zomato dataset is 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, 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 hasn't been saturated yet and the demand is increasing day by day. Inspire 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.

. These kind of analysis can be done using the data, by studying the factors such as

- Location of the restaurant
- Approx Price of food
- Which locality of that city serves that cuisines with maximum number of restaurants.
- The needs of people who are striving to get the best cuisine of the neighborhood.
- Is a particular neighborhood famous for its own kind of food.

Tools & Libraries

• Python • Jupyter Notebook • Pandas • Numpy • Seaborn • Matplotlib • Plotly & Cufflinks

Data Description

The dataset contains the following Columns:

- Name: Name of the restaurant
- rest_type: Type of restaurant
- online_order : Online order is available or not for restaurant
- approx_cost(for two people) : Contains the approximate cost of meal for two
- Cuisine Category: Consists of different food categories
- City: Name of the city the restaurant is located in.
- location: Contains the neighborhood in which the restaurant is located
- URL: This feature contains the URL of the restaurant on the Zomato website
- Cuisine: Cuisine available in Restaurant
- book_table: Table booking available or not
- address : Address of the restaurant
- rate: Contains the overall rating of the restaurant out of 5
- Votes: Contains total number of votes for the restaurant
- listed_in(type): Type in which restaurant is listed on Zomato
- phone: Contact number of restaurant

Data Cleaning

I made the following changes and created the following variables:

- Deleted the columns URL, address and phone as they were not important for analysis
- Removed Duplicate Rows
- Deleted dish_liked Column it was having too many null values
- Renamed the column 'approx_cost(for two people)'
:'cost','listed_in(type)':'type','listed_in(city)':'city'
- Cleaned the cost column as it was an object data type. Removed the comma and converted it into the numeric data type
- Cleaned the rate column and converted it into the numeric data type

EDA

I looked at the different-different trends of the data and Below is a few highlights of the analysis.

- Largest Food Chain in Bangalore
- Most Popular Restaurant types in Bangalore
- Expensive Restaurants in Bangalore
- Ratings by Type of Restaurants
- Restaurants per location
- Best economical Restaurants
- Best Average cost Restaurants
- Finding best Restaurants by Cost & Location