Zomato Restaurants Data Bengaluru

By Ashritha.S

Terms of Use

The data is open sourced and anyone can use this data for education or commercial purpose. Suitable citation should be added wherever used.

Information about Copyrighted Content

Some instruments administered as part of this study may contain in whole or substantially in part contents from copyrighted instruments. However the data scraped was entirely for educational purposes only. Note that I don't claim any copyright for the data. All copyrights for the data is owned by Zomato Media Pvt. Ltd.. The data is correct to the best of my knowledge, to that available on the zomato website until 15 March 2019.

Contents

1 Introduction

Food. Everyone loves it, everyone has it. Everyone even talks about it. Food is something we can talk about for hours and hours. India is rightly called the Land of Spices. No country in the world produces as many varieties of spices as India. The restaurant industry in India has witnessed an unprecedented transformation with the entry of a variety of national and international players. This has, in turn, given birth to a huge demand for qualified professionals in the sector and all related industries. Thanks to the technological revolution, Indian restaurant setups have now gone online to gain more customers and serve them better.

But the demand-and-supply graph isn't quite the way it should be. With a noticeable shortage of skilled professionals, the restaurant industry presents a whole gamut of opportunities waiting to be grabbed. Enter culinary arts institutes. Traditional cooking schools and hotel management colleges have now expanded the range of education they offer in order to satisfy industry demands. Indian universities are investing time and money to train students to make them able and employable.

It's not surprising that the higher frequency of eating out has also evolved the market for the food services sector. The Indian food service market has come a long way from the early Nineties when it was dominated by unorganized players and few brands.

The revolution began in 1996 with McDonalds, Pizza Hut, Domino's Pizza, Subway and Yo!China, among others, setting up shop in the country. Since then, the food services market has been continuously growing.

The good news is that the food services industry is set to grow for many years to come, given the rising disposable incomes, a greater population of younger people, the growth of consumers in smaller towns and the widening exposure to new cultures and cuisines besides an increased propensity of eating outside the home. The analysis will mainly help new restaurants in examining the factors affecting their restaurant location.

2 Purpose of study

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 50,000 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. In spite 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 most people 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. If yes then is that locality populated by a particular sect of people for e.g. Jain, Marwaris, Guajarati's who are mostly vegetarian. This kind of analysis can be done using the data, by studying the factors such as

- Approx. Price of food
- Location of the restaurant
- Theme based restaurant or not
- 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?

- Just so that you have a good meal the next time you step out

3 Properties

3.1 File Format

The data is in csv format. For best results read the data using Python.

3.2 Content

The dataset contains 17 variables all of which were scraped from the Zomato website. The dataset contains details of more than 50,000 restaurants in Bengaluru in each of its neighborhood. The data is correct to the best of my knowledge, to that available on the Zomato website until 15 March 2019.

3.3 Size

The total size of the dataset is approximately 547MB. The dataset examined has the following dimensions:

Feature	Result
Number of Variables	7

4 Source

The data was scraped from the website of Zomato. All possible ways were tried to keep the data error free and I have tried to achieve 100 percent accuracy in the dataset. However I will not be responsible if any error is found in the dataset.

5 Variable names and description

Variable	Type	Description
online order	category	whether online ordering is available in the restaurant or not
book table	category	table book option available or not
rate	object	contains the overall rating of the restaurant out of 5
votes	int	contains total number of rating for the restaurant as of the above mentioned date
location	category	contains the neighborhood in which the restaurant is located
rest type	category	restaurant type
dish liked	object	dishes people liked in the restaurant
approx cost(for two people)	float	contains the approximate cost for meal for two people
reviews list	object	list of tuples containing reviews for the restaurant, each tuple consists of two values, rating and review by the customer
menu item	object	contains list of menus available in the restaurant
listed in(type)	category	type of meal
listed in(city)	category	contains the neighborhood in which the restaurant is listed

6 Procedure

The data was scraped from Zomato in two phase. After going through the structure of the website I found that for each neighborhood there are 6-7 category of restaurants viz. Buffet, Cafes, Delivery, Desserts, Dine-out, Drinks & nightlife, Pubs and bars.

7.1 Phase I

In Phase I of extraction only the URL, name and address of the restaurant were extracted The URL's for each of the restaurants on the Zomato were recorded in the csv file so that later the data can be extracted individually for each restaurant. This made the extraction process easier and reduced the extra load on my machine and dropped unnecessary columns.

7.2 Phase II

In Phase II the recorded data for each restaurant and each category was read and data for each restaurant was scraped individually. 5 variables were scraped in this phase. For each of the neighborhood and for each category their online_order, book_table, rate, votes, phone, location, rest_type, dish_liked, cuisines, approx_cost(for two people), reviews_list, menu_item was extracted.

7 Software used

The data was scraped using the Python Programming language with the help of Selenium API.

Software	Version
Python	3.7.2
Pandas	0.24.2
Numpy	1.16.2

References

[1] India's Restaurant Evolution
https://www.thehindubusinessline.com/news/variety/indias-restaurant-revolution/article23029875.ece

[2] Rise of Restaurant Industry
https://theaims.ac.in/resources/rise-of-the-restaurant-industry-in-india.html

[3] Restaurant Boom in Bengaluru
https://timesofindia.indiatimes.com/city/bengaluru/Restaurant-boom-in-Bangalore-thanks-to-unending-appetite-of-young-professionals/articleshow/34629897.cms