Zomato Bangalore Restaurants

Source:

The Zomato Restaurants data was sourced from Kaggle. And its owner collected it directly from the Zomato website, a popular online restaurant discovery and food delivery platform. The dataset is open-sourced and available for educational and purposes.

Link: https://www.kaggle.com/datasets/himanshupoddar/zomato-bangalore-restaurants

Data Collection:

The data collection process occurred in two phases:

Phase I:

In this initial phase, the URLs, names, and addresses of restaurants were extracted from the Zomato website. These details were visible on the front page and were recorded to facilitate later individual data extraction for each restaurant.

Phase II:

During the second phase, detailed information for each restaurant and category was scraped individually. The dataset comprises 17 variables, including details such as online ordering availability, table booking options, ratings, votes, phone numbers, locations, restaurant types, cuisines, and more. Python, along with the Selenium API, was used for the scraping process.

Contents:

The dataset contains information on over 50,000 restaurants in Bengaluru, India. The variables cover diverse aspects, including restaurant details, customer reviews, online ordering, table booking, cuisines and approximate costs.

Limitations:

While the dataset provides valuable insights, it is crucial to acknowledge certain limitations:

- Temporal Limitation: The data is accurate only up to March 15, 2019, and may not reflect recent changes in the restaurant landscape.
- Sampling Bias: The dataset may not be fully representative of all restaurants in Bengaluru, as it relies on information available on the Zomato website.
- Data Accuracy: Although efforts were made to achieve 100 percent accuracy, errors may exist, and the creator disclaims responsibility for any inaccuracies.

Ethics:

Ethical considerations are paramount in any data analysis project. In this context:

- Terms of Use: The dataset is open-sourced, but users are expected to adhere to suitable citation practices when using the data.
- Copyright: All copyrights for the data belong to Zomato Media Pvt. Ltd. The data was collected solely for educational purposes.

Relevance:

The dataset is highly relevant for analysing the factors affecting restaurant ratings, exploring the types of restaurants in different locations, and understanding the culinary preferences and demographics of Bengaluru. This analysis can assist in decision-making for aspiring restaurant owners and provide valuable insights for customers.

In summary, the Zomato Restaurants dataset serves as a valuable tool for developing practical skills in business development, marketing, and strategic decision-making. The real-world applicability of the data ensures that the insights gained can directly contribute to your professional growth in these domains.

DATA PROFILE

Variable	Description	Time Variant/ Invariant	Structured/ Unstructured	Quantitative / Qualitative	Nominal/ Ordinal/ Discrete/ Continuous
url	Contains the url of the restaurant in the Zomato website	Invariant	Structured	Qualitative	Nominal
address	Contains the address of the restaurant in Bengaluru	Invariant	Structured	Qualitative	Nominal
name	Contains name of the restaurant	Invariant	Unstructured	Qualitative	Nominal
Online order	Whether online ordering is available in the restaurant or not	Invariant	Structured	Qualitative	Nominal
Book table	Table book option available or not	Invariant	Structured	Qualitative	Nominal
rate	Contains overall rating of the restaurant out of 5	Invariant	Structured	Quantitative	Ordinal
votes	contains total number of rating for the restaurant as of the above mentioned date	Time Variant	Structured	Quantitative	Discrete
phone	contains the phone number of the restaurant	Invariant	Structured	Qualitative	Nominal
location	Contains the neighbourhood in which the restaurant is located	Invariant	Structured	Qualitative	Nominal
rest type	Restaurant type	Invariant	Structured	Qualitative	Nominal
Dish liked	dishes people liked in the restaurant	Invariant	Structured	Qualitative	Nominal
cuisines	food styles, separated by comma	Invariant	Structured	Qualitative	Nominal
Approx. Cost(for two people)	contains the approximate cost for meal for two people	Invariant	Structured	Quantitative	Continuous
Reviews list	list of tuples containing reviews for the restaurant, each tuple consists of two values, rating and review by the customer	Time Variant	Structured	Quantitative and Qualitative	Ordinal (Rating and Textual) review

Menu item	contains list of menus available in the restaurant	Invariant	Structured	Qualitative	Nominal
Listed in(type)	Type of meal	Invariant	Structured	Qualitative	Nominal
Listed in(city)	contains the neighborhood in which the restaurant is listed	Invariant	Structured	Qualitative	Nominal