

ZOMATO DATA ANALYSIS USING PYTHON AND POWER BI DASHBOARD

Renukha.M¹, Haritha.M², Kalaimugil.K³, Sathya.S⁴, Sharmila.A⁵

¹Assistant Professor, Artificial Intelligence and Data Science Bharathiyar Institute of Engineering for Women (Autonomous), Salem, Tamil Nadu, India.

^{2,3,4,5}UG Student., Artificial Intelligence and Data Science Bharathiyar Institute of Engineering for Women (Autonomous), Salem, Tamil Nadu, India.

Abstract – This research study focuses on analysing restaurant data collected from the Zomato platform using Python and Power BI tools. The main objective of the study is to transform raw restaurant datasets into meaningful insights that can support business decision making. The dataset contains information such as restaurant names, locations, cuisine types, cost for two people, delivery availability, and customer ratings.

Python libraries including Pandas and NumPy are used for data preprocessing, cleaning, and transformation. Exploratory Data Analysis (EDA) is performed to understand patterns in restaurant ratings, pricing distribution, and cuisine popularity. Visualization tools such as Matplotlib and Seaborn help represent patterns in graphical form.

Power BI is then used to build interactive dashboards that allow stakeholders to analyse restaurant performance through dynamic filters and KPI indicators. The analysis results identify high-demand cuisines, city-wise restaurant distribution, and pricing trends that influence customer ratings.

This integrated analytical framework demonstrates how data analytics and business intelligence tools can be applied in the food delivery industry to generate valuable insights for restaurants and online delivery platforms.

Key Words: Zomato, Data Analytics, Python, Power BI, Exploratory Data Analysis, Business Intelligence.

I. INTRODUCTION

Online food delivery platforms such as **Zomato** have significantly transformed the restaurant and hospitality industry by bridging the gap between customers and restaurants through digital technology. These platforms provide users with convenient access to restaurant information, online ordering services, and customer reviews, thereby influencing consumer dining decisions. With the rapid growth of smartphone usage and internet accessibility, digital food delivery platforms have become an integral part of modern urban lifestyles. As a result, they generate vast amounts of structured and semi-structured data related to restaurants, including customer ratings, textual reviews, menu pricing, cuisine types, restaurant locations, and service availability.

The availability of such large-scale datasets presents an opportunity for data-driven decision making in the food service industry. By analysing restaurant data, businesses can gain insights into market trends, customer preferences, and operational performance. For instance, restaurant ratings can indicate service quality and customer satisfaction, while pricing information can help identify competitive market positioning. Similarly, cuisine popularity data can reveal regional food preferences and emerging dining trends among customers.

However, raw datasets obtained from online platforms are often imperfect and require proper preprocessing before analysis. The data may contain missing values, duplicate entries, inconsistent formatting, and irrelevant attributes. If such issues are not addressed, they can negatively impact the accuracy and reliability of analytical results. Therefore, data preprocessing steps such as data cleaning, transformation, and normalization are essential to prepare the dataset for meaningful analysis. These processes ensure that the data is structured, consistent, and suitable for statistical evaluation and visualization.

Data analytics techniques play a crucial role in extracting valuable insights from restaurant datasets. Using analytical tools and programming languages, large volumes of data can be processed efficiently to uncover hidden patterns and correlations. For example, analysing rating distributions can help determine the overall quality trends among restaurants, while studying pricing patterns can reveal the relationship between cost and customer satisfaction. Furthermore, geographical analysis can identify areas with high restaurant density or popular dining zones within a city.

In this project, restaurant data obtained from **Zomato** is analysed using **Python** for data processing and **Microsoft Power BI** for data visualization and dashboard creation. Python is used for tasks such as data cleaning, transformation, and exploratory data analysis, while Power BI helps create interactive dashboards that present the analytical results in a visually understandable format. The project focuses on identifying patterns related to restaurant ratings, cost distribution, cuisine popularity, and location-based performance.

The results of this analysis provide valuable insights into customer behavior and restaurant performance. These insights can help restaurant owners and business analysts make informed decisions regarding menu planning, pricing strategies, and service improvements. Additionally, the interactive dashboards enable

stakeholders to monitor trends easily and support strategic planning within the restaurant industry.

II. PROPOSED SYSTEM

Several studies emphasize the importance of data analytics in the restaurant and food delivery industry. With the growth of digital platforms such as **Zomato**, large amounts of restaurant data are generated, including customer ratings, reviews, pricing information, and cuisine types. Researchers have analyzed this data to evaluate restaurant performance and understand customer preferences. Customer ratings and reviews are considered important indicators of service quality and overall customer satisfaction.

Many researchers have applied data mining techniques and machine learning algorithms to analyze restaurant datasets. Methods such as clustering, classification, and regression help identify patterns and relationships within the data. For example, clustering can group restaurants based on cuisine type or price range, while regression analysis can study the relationship between restaurant ratings and pricing. These analytical techniques help businesses identify market trends and improve their services.

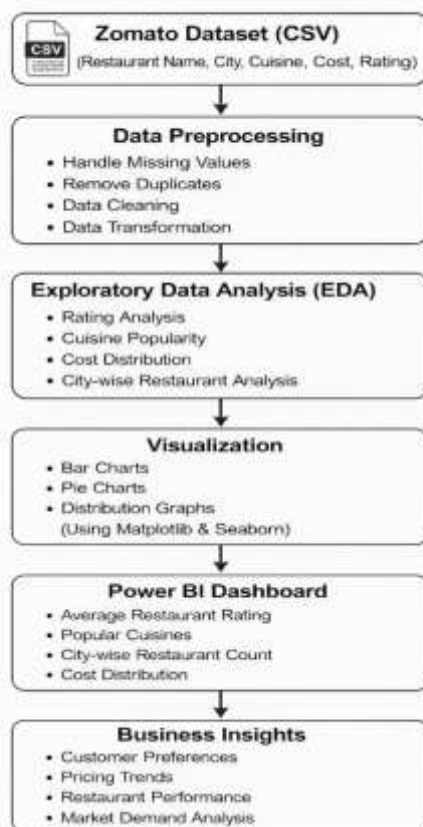
Python is widely used for data analysis because of its powerful libraries such as **Pandas**, **NumPy**, **Matplotlib**, and **Seaborn**. These libraries help in data cleaning, statistical analysis, and visualization of large datasets, making the analysis process more efficient and accurate.

Business intelligence tools such as **Microsoft Power BI** provide interactive dashboards that allow users to explore data visually. These dashboards help stakeholders easily understand key insights and monitor performance indicators. By combining Python-based data analysis with Power BI visualization, organizations can improve decision making and identify opportunities for growth in the restaurant industry.

III. METHODOLOGY

The proposed system follows a structured analytical framework consisting of several stages such as data collection, data preprocessing, exploratory data analysis, visualization, and dashboard development. These steps help transform raw restaurant data into meaningful insights that support decision-making in the food service industry.

Architecture of Zomato Data Analysis System
using Python and Power BI



Data Collection

The dataset used in this study is obtained from the online food delivery platform Zomato in CSV (Comma Separated Values) format. The dataset contains important information about restaurants such as restaurant name, city, cuisine type, average cost for two people, online delivery availability, and customer ratings. This data

provides a comprehensive overview of restaurant characteristics and customer feedback across different locations.

Data Preprocessing

Before analysis, the dataset is cleaned and prepared using Python. Data preprocessing involves handling missing values, removing duplicate records, and correcting inconsistent data formats. In addition, data types such as numerical and categorical variables are standardized to ensure accurate analysis. These preprocessing steps improve the quality of the dataset and make it suitable for further analytical processes.

Exploratory Data Analysis (EDA)

Exploratory Data Analysis is performed to understand the structure and patterns within the dataset. Statistical methods are used to analyze relationships between different variables such as restaurant ratings, pricing levels, and cuisine types. This analysis helps identify trends such as popular cuisines, highly rated restaurants, and variations in restaurant pricing across different cities.

Visualization

To better understand the dataset, graphical visualizations are created using Python libraries such as Matplotlib and Seaborn. Charts such as bar graphs, pie charts, and distribution plots are used to represent rating patterns, cuisine popularity, and cost distribution. These visualizations make it easier to interpret complex data and highlight key trends.

Dashboard Development

Finally, interactive dashboards are created using Microsoft Power BI. These dashboards present important insights including average restaurant ratings, top cuisine categories, city-wise restaurant distribution, and cost analysis. The interactive nature of Power BI dashboards allows users to filter and explore the data easily, enabling a better understanding of restaurant performance and customer preferences.

Table 1: Zomato Dataset Attributes

Attribute	Description	Example
Restaurant Name	Name of the restaurant listed on Zomato	Spice Garden
City	Location where the restaurant operates	Chennai
Cuisine Type	Type of food served by the restaurant	South Indian
Cost for Two	Average cost for two people dining	₹800
Rating	Customer rating provided on the platform	4.2
Online Delivery	Availability of food delivery service	Yes

IV. EXPERIMENT AND RESULT

The analysis of the restaurant dataset from Zomato provided several important insights into restaurant performance, customer preferences, and market trends. By applying data analytics techniques and visualization tools, meaningful patterns were identified from the dataset.

Cuisine Analysis

The analysis of cuisine categories revealed that certain food types dominate the online restaurant market. Cuisines such as North Indian, Chinese, and Fast Food were found to be the most commonly listed and ordered categories across multiple cities. This indicates that these cuisines have a consistently high demand among customers. Restaurants offering multiple cuisine options

often receive more customer engagement and better ratings, as they are able to cater to a wider audience.

Rating Analysis

Restaurant rating analysis shows a noticeable relationship between pricing and customer satisfaction. Restaurants with moderate pricing levels tend to receive higher customer ratings compared to very expensive restaurants. This suggests that customers often prefer restaurants that provide good quality food at reasonable prices. Additionally, highly rated restaurants generally maintain consistent service quality, good food taste, and positive customer experiences.

City-wise Distribution

The city-wise analysis indicates that metropolitan cities have the highest concentration of restaurants listed on the platform. Cities with larger populations and higher digital adoption rates show greater restaurant availability and online delivery services. This trend highlights the growing importance of online food delivery platforms in urban areas where customers prefer convenient ordering options.

Dashboard Insights

Interactive dashboards developed using Microsoft Power BI provide a clear visual representation of the analytical results. These dashboards allow users to apply filters based on city, cuisine type, and rating categories. As a result, stakeholders can easily identify high-performing restaurants, popular cuisine categories, and pricing trends across different regions. The visual insights help businesses understand customer behavior and improve their operational strategies.



Overall, the results demonstrate how data analytics can support better decision making in the restaurant industry by revealing hidden patterns within large datasets.

V. CONCLUSION

This research demonstrates how data analytics techniques can be applied to analyse restaurant datasets and extract meaningful business insights. By using **Python** for data preprocessing and analysis, and **Microsoft Power BI** for visualization, the study successfully transforms raw restaurant data into clear and interpretable insights. The combination of these tools provides an effective approach for handling large datasets and presenting analytical results in an understandable format.

The data preprocessing and exploratory analysis stages helped identify important patterns related to cuisine popularity, restaurant pricing trends, and customer rating distribution. The analysis revealed that certain cuisines dominate the market, while restaurants offering moderate pricing often receive higher customer satisfaction ratings. These insights highlight the importance of understanding customer preferences and market demand in the restaurant industry.

The interactive dashboards developed using Power BI provide a user-friendly platform for visualizing and exploring restaurant performance metrics. Through graphical representations and interactive filters,

stakeholders can easily analyse city-wise restaurant distribution, average ratings, and popular cuisine categories. Such visual insights can assist restaurant owners, analysts, and decision makers in developing better marketing strategies, pricing policies, and service improvements.

Overall, the proposed analytical framework demonstrates the potential of combining data analytics and business intelligence tools to support data-driven decision making in the restaurant industry. By analysing large-scale datasets from platforms such as **Zomato**, businesses can gain a deeper understanding of customer behaviour and market dynamics.

For future research, this study can be extended by integrating real-time datasets and applying advanced analytical techniques. Sentiment analysis can be performed on customer reviews to understand customer opinions more deeply, and machine learning algorithms can be used for predictive analytics such as forecasting restaurant ratings, demand patterns, and customer satisfaction levels. These improvements can further enhance the effectiveness of data analytics in the food delivery and restaurant sector.

• Technical Indicator

For this project, several technical indicators are used to evaluate restaurant performance and customer preferences based on the dataset obtained from **Zomato**. These indicators help transform raw data into meaningful insights that support data-driven decision making in the restaurant industry. Key indicators include the **average restaurant rating**, which measures overall customer satisfaction; **restaurant density by city**, which shows the concentration of restaurants in different urban areas; and the **cuisine popularity index**, which identifies the most demanded food categories among customers. In addition, **cost distribution analysis** is used to understand pricing

patterns and categorize restaurants into low, medium, and high price ranges. Another important indicator is the **online delivery availability rate**, which measures the percentage of restaurants providing delivery services through digital platforms. These technical indicators are calculated and analyzed using **Python**, and the results are visualized through interactive dashboards created in **Microsoft Power BI**. Together, these indicators provide a comprehensive view of restaurant performance, customer behavior, and market trends, enabling stakeholders to make informed business decisions.

VI. REFERENCES

1. W. McKinney, *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and Python*, 2nd ed., O'Reilly Media, 2017.
2. J. VanderPlas, *Python Data Science Handbook: Essential Tools for Working with Data*, O'Reilly Media, 2016.
3. F. Provost and T. Fawcett, *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking*, O'Reilly Media, 2013.
4. M. Waskom, "Seaborn: Statistical Data Visualization," *Journal of Open Source Software*, vol. 6, no. 60, 2021.
5. J. D. Hunter, "Matplotlib: A 2D Graphics Environment," *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90–95, 2007.
6. Microsoft Corporation, "Microsoft Power BI Documentation," Available: <https://learn.microsoft.com/power-bi/>
7. Kaggle, "Zomato Restaurant Dataset," Available: <https://www.kaggle.com/datasets/himanshupoddar/zomato-bangalore-restaurants>
8. R. S. Pressman, *Software Engineering: A Practitioner's Approach*, McGraw-Hill Education, 2014.
9. T. Davenport and J. Harris, *Competing on Analytics: The New Science of Winning*, Harvard Business Review Press, 2007.
10. Zomato, "Online Food Delivery Platform and Restaurant Data," Available: <https://www.zomato.com/>