

**Team Member's Name, Email and Contribution:**

- Vinayak Marathe: vinmarathe100@live.com
- Riya Patel: riyapatelrp8308@gmail.com

Contributor's Role:

**Vinayak Marathe:**

- Worked on Problem Statement with Business Objective
- Datasets Reading and Data Wrangling
- Worked on Missing/Null values
- Performed Outlier visualization (Anomaly Detection)
- Prepared different type Model and implemented it
- Performed Hypothesis Testing with 3 hypothetical statements
- Performed Model Explainability and Feature Importance
- Worked on Competitive Advantages and Business Goals of our Project
- Summary and Technical Documentation.

**Riya Patel:**

- Data Explanation and Data Pre-processing
- Data Visualization and Storytelling
- Performed Textual Data Preprocessing using NLP
- Feature Manipulation and Feature Selection
- Prepared different type Model and implemented it
- Evaluated Model Performance and Future Work
- Analyzed Overall Results of the Project and Conclusion
- Presentation Preparation

GitHub Repo link.

1. Vinayak Marathe: <https://github.com/v1git12>
2. Riya Patel: <https://github.com/riyapatelrp>

**Project Link:**

<https://colab.research.google.com/drive/1RFluzBmyBeT0oR05S7JMo2D6EtOGSG3e?usp=sharing#scrollTo=jDeoZs4rwuMr>

**Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)**

Zomato is an Indian restaurant aggregator and food delivery start-up founded by Deepinder Goyal and Pankaj Chaddah in 2008. Zomato provides information, menus and user-reviews of restaurants, and also has food delivery options from partner restaurants in select cities. India is quite famous for its diverse multi cuisine available in a large number of restaurants and hotel resorts, which is reminiscent of unity in diversity. Restaurant business in India is always evolving. More Indians are warming up to the idea of eating restaurant food whether by dining outside or getting food delivered. The growing number of restaurants in every state of India has been a motivation to inspect the data to get some insights, interesting facts and figures about the Indian food industry in each city. So, this project focuses on analyzing the Zomato restaurant data for each city in India.

The Project focuses on Customers and Company, you have to analyze the sentiments of the reviews given by the customer in the data and make some useful conclusions in the form of visualizations. Also, cluster the zomato restaurants into different segments: The data is visualized as it becomes easy to analyses data at instant. The Analysis also solves some of the business cases that can directly help the customers finding the best restaurant in their locality and for the company to grow up and work on the fields, they are currently lagging in. This could help in clustering the restaurants into segments. Also, the data has valuable information around cuisine and costing which can be used in cost vs. benefit analysis. Data could be used for sentiment analysis. Also, the metadata of reviewers can be used for identifying the critics in the industry.

The project begins by understanding the variables in the restaurant and review data frames. The restaurant data frame has 105 rows and 6 columns, with information about restaurant names, links, cost, collections, cuisines, and timings. The review data frame has 10000 rows and 7 columns, with information about the restaurant reviews, such as restaurant name, reviewer name, review, rating, metadata, time, and pictures. The data wrangling process is then performed to find meaningful insights from the data.

Data visualization is used to present the findings in the form of ten different charts, including univariate, bivariate, and multivariate visualizations. These visualizations help in understanding the relationships between different variables and the overall trends in the data. Hypothesis testing is performed to validate three different hypothetical statements. Feature engineering is also performed to prepare the data for modeling.

The project uses several clustering techniques, such as K-Means Clustering, PCA, Hierarchical Clustering, and DBSCAN Clustering, to cluster the restaurants into different segments. Clustering the restaurants helps to identify similarities and differences between different restaurants based on their features, such as cuisine, cost, collections, and timings. The project's conclusions provide useful insights for customers looking for the best restaurants in their locality and for the company to grow in the areas they are currently lagging in.

Finally, a Content-Based Recommendation System is implemented to recommend restaurants to customers based on their preferences. The recommendation system uses the data from the clustering process to identify restaurants that are similar to the customer's preferences. This helps in providing personalized recommendations to customers, improving their overall experience.

The conclusion of this Zomato restaurant clustering and metadata sentiment analysis project is that it is possible to use natural language processing and machine learning algorithms to build a model that can accurately cluster restaurants based on their reviews and sentiments. This project has helped identify customer preferences and restaurant ratings in order to better understand the impacts of customer feedback on the restaurant industry. This model can then be used to improve the decision-making process of a restaurant owner or manager in terms of advertising, pricing, customer acquisition, and other important business decisions. Furthermore, the insights provided by this project can be used for further business growth strategies.

