**Zomato Bangalore Restaurants**

**Group-14**

**Deliverable-1**

**GitHub -** https://github.com/Vishnureddy466/bigData

* 1. **Group Members:**
     + Vishnu Vardhan Reddy Addulamale
     + Anvesh Chidura
     + Jaideep Gurrapu
     + Jagadeshwar Reddy Panta
     + Sushant Sagar Saka
  2. **Communication plan :**
* We are connecting weekly twice in google meetings, like discussing in depth over the issues we are facing in the project.
* Meeting in Atkins Library during the weekends.
* We are also communicating in slack and sharing the files which are related to the project.

**2.Data we would consider:**

We would like to select the ”Zomato Bangalore Restaurants" “https://[www.kaggle.com/datasets/himanshupoddar/zomato-bangalore-restaurants?datasetId](http://www.kaggle.com/datasets/himanshupoddar/zomato-bangalore-restaurants?datasetId)

=153420&sortBy=voteCount”.

**3.Business Problem or Opportunity for Zomato Bangalore Restaurants:**

The restaurant business is extremely competitive, and restaurants are constantly seeking methods to enhance their offerings and stay ahead of the competition. Understanding the customer preferences, trying to identify popular cuisines and locales, and analyzing pricing and user reviews can all provide useful insights that can assist restaurants in making data-driven decisions.

**Domain Knowledge:**

The following domains can be used to evaluate and translate the information in the Zomato Bangalore Restaurants dataset:

Restaurant sector: While evaluating the dataset, understanding the restaurant industry and the variables that affect a restaurant's success, such as location, pricing, menu, and customer service, can be very helpful.

Cuisines: Studying Bangalore's various cuisines and their appeal can assist identify popular cuisines and present potential for developing new and distinctive cuisines.

Data analysis and visualization: Understanding of the methods and tools for dealing with and displaying data, including pandas, numpy, matplotlib, and seaborn.

**Research Objectives:**

The objective of this research is to analyze the Zomato Bangalore Restaurants dataset and extract meaningful insights that can help restaurants improve their offerings and stay ahead of their competitors.

**Research Questions:**

The research will make an effort to respond to the following questions:

How does a restaurant's pricing affect its rankings and reviews from customers?

Do restaurants that deliver food online receive better ratings and more ratings than those that don't?

What are the most common complaints and issues raised in restaurant reviews in Bangalore, and how do these vary by cuisine type and price range?

Are there any notable trends or patterns in the data that suggest opportunities for new restaurant concepts or innovations in the Bangalore market?

Going forward we may add a few more questions and try to resolve those.

**Deliverable-2**

**Data Understanding:**

a) Exploratory Data Analysis:

Once the dataset is loaded into an S3 Bucket, by Leveraging the AWS Sagemaker service, we performed Exploratory Data Analysis and performed data preparation tasks on Zomato dataset. The ipynb file of this detailed analysis is located here.

b) Dashboard:

By Leveraging AWS Quicksight service we have constructed dashboards on different columns of the dataset and gathered insights that aid to understand the nuances and existing patterns in the dataset. The PDF file consisting of screenshots of these interactive visualizations is present in the file here.

**Data Preparation:**

Data Preparation involves steps like cleaning and transforming raw data and making it suitable for building Machine Learning models and further analysis. We have performed the necessary data preparation steps as mentioned below:

1. Dropping unnecessary columns in the data.

2. Removing Duplicate Rows.

3. Removing Null values.

4. Transforming "Rate", "estimated\_cost\_for\_2\_ppl", "rest\_type", "location", and "cuisines" columns.

5. Encoding textual data using the LabelEncoder function.