**Project Title: Zomato Restaurants Data Analysis**

Team Member: Robin Sah (G48297336)

**Project Overview**

I have chosen to undertake a comprehensive analysis of the Zomato Restaurants dataset. This dataset offers a glimpse into the vast and varied world of restaurants listed on Zomato. The project aims to transform the raw JSON dataset into a structured format suitable for in-depth analysis using a SQL database. I intend to explore several facets of the dataset, such as restaurant ratings, popularity based on cuisine, geographical distribution, price ranges etc. and more as the project proceeds.

**Approach**

Technology: I plan to use PostgreSQL as SQL database system due to its support for JSON data types, and advanced analytical functions. For data transformation and preliminary analysis, Python, along with libraries such as Pandas for data manipulation and Matplotlib for visualization, will be employed. The Jupyter Notebook environment will serve as the primary tool for coding, data exploration, and documentation.

Data Sources: The primary source of the data is the Zomato Restaurants dataset from Kaggle (<https://www.kaggle.com/shrutimehta/zomato-restaurants-data>). This dataset contains information in JSON format, including restaurant names, locations, user ratings, price details, and cuisines.

Distribution of Work:

Data Exploration and Cleaning: Initially, I plan to explore the dataset to identify inconsistencies, missing values, and potential transformations needed. This phase aims to gain a thorough understanding of the dataset's structure and content.

Database Schema Design: Based on understanding of the dataset, I hope to design a normalized database schema that efficiently represents the data, paying attention to entities like restaurants, cuisines, and locations, and their relationships.

Data Transformation: I will use python scripts to transform the JSON dataset into a format that aligns with the database schema. This step includes parsing JSON, normalizing nested structures, and preparing CSV files for import into PostgreSQL.

Database Setup and Data Import: After setting up PostgreSQL database, the transformed data will be imported.

Analysis and Reporting: Once the data is in PostgreSQL, I plan to perform various analyses to uncover insights related to restaurant ratings, popular cuisines by region, price trends, and more using Python or Tableau or Power BI.