

# Restaurant Project Description

## Description

Data Analysis is the process of creating a story using the data for easy and effective communication. It mostly utilizes visualization methods like plots, charts, and tables to convey what the data holds beyond the formal modeling or hypothesis testing task.

**Domain:** Marketing

Read the information given below and also refer to the data dictionary provided separately in an excel file to build your understanding.

**Problem Statement** A restaurant consolidator is looking to revamp its B-to-C portal using intelligent automation tech. It is in search of different matrix to identify and recommend restaurants. To make sure an effective model can be achieved it is important to understand the behaviour of the data in hand.

## Approach:

### 1. Data Preliminary analysis:

- Perform preliminary data inspection and report the findings as the structure of the data, missing values, duplicates cleaning variable names etc.
- Based on the findings from the previous questions identify duplicates and remove them.
- 

### 2. Prepare a preliminary report of the given data by answering following

questions. Expressing the results using graphs and plot will make it more appealing.

- Explore the geographical distribution of the restaurants, finding out the cities with maximum / minimum number of restaurants.
- Explore how ratings are distributed overall.

- Restaurant franchise is a thriving venture. So, it becomes very important to explore the franchise with most national presence.
- What is the ratio between restaurants that allow table booking vs that do not allow table booking?
- What is the percentage of restaurants providing online delivery?
- Is there a difference in no. of votes for the restaurants that deliver and the restaurant that don't?
- What are the top 10 cuisines served across cities?
- What is the maximum and minimum no. of cuisines that a restaurant serves? Also, what is the relationship between No. of cuisines served and Ratings
- Discuss the cost vs the other variables.
- Explain the factors in the data that may have an effect on ratings e.g. No. of cuisines, cost, delivery option etc.

All the information gathered here will lead to a better understanding of the data and allow for a better implementation of ML models.

### **Project Task1:**

#### **Importing, Understanding, and Inspecting Data :**

1. Perform preliminary data inspection and report the findings as the structure of the data, missing values, duplicates, etc.
2. Based on the findings from the previous questions, identify duplicates and remove them

#### **Performing EDA:**

1. Explore the geographical distribution of the restaurants and identify the cities with the maximum and minimum number of restaurants
2. Restaurant franchising is a thriving venture. So, it is very important to explore the franchise with most national presence
3. Find out the ratio between restaurants that allow table booking vs. those that do not allow table booking
4. Find out the percentage of restaurants providing online delivery
5. Calculate the difference in number of votes for the restaurants that deliver and the restaurants that do not deliver

## **Project Task 2:**

### **Performing EDA:**

1. What are the top 10 cuisines served across cities?
2. What is the maximum and minimum number of cuisines that a restaurant serves? Also, which is the most served cuisine across the restaurant for each city?
3. What is the distribution cost across the restaurants?
4. How ratings are distributed among the various factors?
5. Explain the factors in the data that may have an effect on ratings. For example, number of cuisines, cost, delivery option, etc.

### **Dashboarding:**

1. Visualize the variables using Tableau to help user explore the data and create a better understanding of the restaurants to identify the "star" restaurant
2. Demonstrate the variables associated with each other and factors to build a dashboard