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 modelling or hypothesis testing task.

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

1. Prepare a preliminary report of the given data by answering following questions. Expressing the results using graphs and plot will make it more appealing.
    1. Explore the geographical distribution of the restaurants, finding out the cities with maximum / minimum number of restaurants.
    2. Explore how ratings are distributed overall.
    3. Restaurant franchise is a thriving venture. So, it becomes very important to explore the franchise with most national presence.
    4. What is the ratio between restaurants that allow table booking vs that do not allow table booking?
    5. What is the percentage of restaurants providing online delivery?
    6. Is there a difference in no. of votes for the restaurants that deliver and the restaurant that don't?
    7. What are the top 10 cuisines served across cities?
    8. 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
    9. Discuss the cost vs the other variables.
    10. Explain the factors in the data that may have an effect on ratings e.g. No. of cuisines, cost, delivery option etc.