**CONTEXT:**

I was always fascinated by the food culture of Bengaluru. Restaurants from all over the world can be found here in Bengaluru. From United States to Japan, Russia to Antarctica, you get all type of cuisines here. Delivery, Dine-out, Pubs, Bars, Drinks,Buffet, Desserts you name it and Bengaluru has it. Bengaluru is best place for foodies. The number of restaurant are increasing day by day. Currently which stands at approximately 12,000 restaurants. With such an high number of restaurants. This industry hasn't been saturated yet. And new restaurants are opening every day. However it has become difficult for them to compete with already established restaurants. The key issues that continue to pose a challenge to them include high real estate costs, rising food costs, shortage of quality manpower, fragmented supply chain and over-licensing. This Zomato data aims at analyzing demography of the location. Most importantly it will help new restaurants in deciding their theme, menus, cuisine, cost etc for a particular location. It also aims at finding similarity between neighborhoods of Bengaluru on the basis of food. The dataset also contains reviews for each of the restaurant which will help in finding overall rating for the place.

# **Exploratory Data Analysis(EDA):**

The basic idea of analyzing the Zomato dataset is to get a fair idea about the factors affecting the establishment of different types of restaurant at different places in Bengaluru, aggregate rating of each restaurant, Bengaluru being one such city has more than 12,000 restaurants with restaurants serving dishes from all over the world. With each day new restaurants opening the industry hasn’t been saturated yet and the demand is increasing day by day. In spite of increasing demand it however has become difficult for new restaurants to compete with established restaurants. Most of them serving the same food. Bengaluru being an IT capital of India. Most of the people here are dependent mainly on the restaurant food as they don’t have time to cook for themselves. With such an overwhelming demand of restaurants it has therefore become important to study the demography of a location. What kind of a food is more popular in a locality. Do the entire locality loves vegetarian food. If yes then is that locality populated by a particular sect of people for eg. Jain, Marwaris, Gujaratis who are mostly vegetarian. These kind of analysis can be done using the data, by studying the factors such as • Location of the restaurant • Approx Price of food • Theme based restaurant or not • Which locality of that city serves that cuisines with maximum number of restaurants • The needs of people who are striving to get the best cuisine of the neighborhood • Is a particular neighborhood famous for its own kind of food.

# **Feature Engineering:**

After the Exploratory data analysis(EDA) the next important step is ****Feature Engineering****. ****Feature engineering**** is the process of using domain knowledge to extract features from raw data via data mining techniques. These features can be used to improve the performance of machine learning algorithms. Feature engineering can be considered as applied machine learning itself. By applying various feature engineering techniques we made our data ready for the next step which is ****Feature Selection****.

# **Feature Selection:**

In machine learning and statistics, ****feature selection****, also known as variable selection, attribute selection or variable subset selection, is the process of selecting a subset of relevant features (variables, predictors) for use in model construction. We used Extra tree regressor to extract top 10 most important feature which helped our model to make all the correct predictions.

# **Model Building:**

After all the pre-processing of the data and feature selection our data is ready for model building. The machine learning algorithms applied for model building are Linear Regression, Decision Tree Regressor, Random Forest Regressor.

# **Model Deployment:**

After building the model our model is ready for deployment. But before deployment we need to convert our model into pickle file(.pkl). After converting it to a pickle file we will be able to deploy our model over web successfully.

## **Motivation:**

I was always astonished by how each of the restaurants are able to keep up the pace in spite of that cutting edge competition. And what factors should be kept in mind if someone wants to open new restaurant. Does the demography of an area matters? Does location of a particular type of restaurant also depends on the people living in that area? Does the theme of the restaurant matters? Is a food chain category restaurant likely to have more customers than its counter part? Are any neighborhood similar ? If two neighborhood are similar does that mean these are related or particular group of people live in the neighborhood or these are the places to it? What kind of a food is more popular in a locality. Do the entire locality loves vegetarian food. If yes then is that locality populated by a particular sect of people for eg. Jain, Marwaris, Gujaratis who are mostly vegetarian. There are infacts dozens of question in my mind. lets try to find out the answer with this dataset.