1. **Problem Description:**

The problem statement states that given a dataset consisting restaurants, cuisines served, and their coordinates, we need to identify similar localities, based on which localities prefer a similar kind of cuisines. Similar cuisines add to the similarity of two localities, and using K Means clustering, we can identify clusters with similar localities. A cluster can have localities with different “1st popular cuisine”, but there can be other less popular cuisines that contribute to their similarity.

Restaurants can often be used to analyse localities. Our analysis can be extended to recommending localities to new and upcoming restaurants, and giving them alternative cuisines to try in these localities. The analysis can also be extended to analyse the average spending or the localities that have similar restaurants in terms of probable average spending.

Food being a very integral part of our cultures, should give us neighbourhoods/localities that are similar to each other, with an acceptable accuracy.

1. **Data Description:**

The data we use has the following columns. Every Restaurant contains the following variables:

• Restaurant Id: Unique id of every restaurant across various cities of the world  
• Restaurant Name: Name of the restaurant  
• Country Code: Country in which restaurant is located  
• City: City in which restaurant is located  
• Address: Address of the restaurant  
• Locality: Location in the city  
• Locality Verbose: Detailed description of the locality  
• Longitude: Longitude coordinate of the restaurant's location  
• Latitude: Latitude coordinate of the restaurant's location  
• Cuisines: Cuisines offered by the restaurant  
• Average Cost for two: Cost for two people in different currencies  
• Currency: Currency of the country  
• Has Table booking: yes/no  
• Has Online delivery: yes/ no  
• Is delivering: yes/ no  
• Switch to order menu: yes/no  
• Price range: range of price of food  
• Aggregate Rating: Average rating out of 5  
• Rating colour: depending upon the average rating colour  
• Rating text: text on the basis of rating of rating  
• Votes: Number of ratings casted by people

The collected data has been stored in the Comma Separated Value file Zomato.csv. Each restaurant in the dataset is uniquely identified by its Restaurant Id.

Zomato API Analysis is one of the most useful analysis for foodies who want to taste the best cuisines of every part of the world which lies in their budget. This analysis is also for those who want to find the value for money restaurants in various parts of the country for the cuisines.

Dataset credits: Shruti Mehta, Kaggle.