# Kolkata Zomato Dataset Analysis

The dataset contains 7388 rows and 7 columns. The details of the restaurant those provided are –

* 'name' [Name of the restaurant]
* 'voteCount' [Number of votes received from the user]
* 'rating' [rating given to the restaurant by the user]
* 'address' [Address of the restaurant]
* 'cusine' [Category of cuisine offered]
* 'cost' [Average price of the restaurant]
* 'timing' [Opening and closing time of the restaurant]

We will be performing EDA(exploratory data analysis) on the given Dataset.

## Libraries required –

* **Pandas** – Data exploration
* **Numpy** – Mathematical operation
* **Matplotlib** – Data Visualization
* **Seaborn** – Data Visualization

## WorkFlow –

Our workflow will primarily consist of two segments. Data cleaning and data visualization. The data cleaning process will consume more time than the visualization.

### Data Cleaning –

* Vote count , rating and cost , these three columns are type casted as object. We need to convert them in numerical category for our calculative functionality.
* We need to find the missing values and then handle them with mean, median or mode operation based on it’s relevance.
* We need to look for duplicate values in the dataset.
* Let’s create different data frames based on given conditions –
* high\_end\_restos – 1896 , AvgRating – 3.17 , AvgVotecount - 288
* cheap\_restos – 2193 , AvgRating – 2.11 , AvgVotecount - 32
* midnight\_restos – 670, AvgRating – 2.67 , AvgVotecount - 128
* genral\_timed\_resto – 6718 , AvgRating – 2.49 , AvgVotecount - 107
* Let’s find out the most popular cuisines in the market –
* North Indian
* Chinese
* Fast Food
* Café
* Biriyani
* Bengali

### Data Visualization –





