



Predicting suitable locations to
open new cuisine restaurants.

Data Acquisition

- ▶ The dataset `zomato.csv` is downloaded from Kaggle dataset.
- ▶ Restaurant, City, Locality, Type of Cuisine served alongwith Price Range and Rating are the major columns for consideration.
- ▶ Latitude & Longitude for each restaurant are also available in the dataset.
- ▶ In total, 9551 rows and 21 columns.

Data Cleaning

- ▶ For sake of simplicity only data for city='Gurgaon' is considered and rest of the data is dropped.
- ▶ Duplicate, highly similar or highly correlated features were dropped.
- ▶ For example- Rating, Rating Color and Rating text all conveys the same information therefore any two of these columns can be dropped without any loss of information.