

## ***2 Data preparation***

### **Data structure**

The analysis was based on the restaurant entities. The basic attributes includes the ID, name, longitude and latitude of a restaurant, and the relationship of the restaurant to other facilities and sites was summarized as the number of venues of different categories within a distance of 50 m, 200 m and 500 m from the restaurant. The 5 most popular categories were selected representing the combination of the restaurant and others facilities and sites.

### **Data collection and cleaning**

The data was collected from website Foursquare (<https://foursquare.com/>). Because of a maximum limit for a query, the map of Hong Kong was divided to build the search query. The ID, name, longitude and latitude of the restaurants in Hong Kong were collected by search queries and the ID, name, distance and category of the popular facilities near each restaurant were collected by explore queries. The table of the targeted structure was built by joining the converted data frame of the query feedback and screening the attributes.