

B. Data acquisition and cleaning:

New York City is one of the largest metropolises in the USA, it has various Neighborhoods and various kinds of restaurants within its neighborhood, but in our scenario we only need INDIAN restaurants, so we take help of FourSquare API to accomplish our task.

For this project we need following data :

❑ First we need all the neighborhoods of New York along with their latitude and longitude coordinates.

→ **Data Source** : https://en.wikipedia.org/wiki/Neighborhoods_in_New_York_City

→ **Description** : Wikipedia has a list of Neighborhood of New York, we scrap it using python libraries and get its location coordinates using geopy.geocoder library in python.

❑ Indian restaurants in each Neighborhood of New York

→ **Data Source** : **FourSquare API**

→ **Description** : By using this API we will get all the venues in each neighborhood. We can filter this venues to grab only INDIAN restaurants