

# Data

Based on definition of the problem, the data is fetched from the following data sources:

1. A dataset that contains all the boroughs, main neighborhoods and a list of zip codes for each neighborhood of NYC:

<https://www.health.ny.gov/statistics/cancer/registry/appendix/neighborhoods.htm>

2. A list of the wards (analog of main neighborhoods of NYC) and major districts of Tokyo :

[https://en.wikipedia.org/wiki/Special\\_wards\\_of\\_Tokyo](https://en.wikipedia.org/wiki/Special_wards_of_Tokyo)

and postal codes of Tokyo prefecture:

<http://japanzipcodes.blogspot.com/search?q=Tokyo>

3. GeoPy library is used to get the latitude and longitude values of neighborhoods.
4. Foursquare API is used to explore venues of neighborhoods in New York City and Tokyo. Latitude and longitude values extracted before are used as parameters of API request.