

Data Description:

New York city's neighborhood will be loaded, explored, analysed in this project. We will be using the below datasets, Wikipage and links for analysing Newyork city neighborhood.

Data A: For determining the geography of New York and its respective neighborhoods.

As per available information on web New York is famous for its Boroughs which has respective counties. For this exercise (i.e) to load the information and explore it we require a data set fortunately it was available. Also, this data contains the Latitude and Longitude coordinates of each neighborhood. For this analysis below mentioned link will be utilized.

LINK: https://geo.nyu.edu/catalog/nyu_2451_34572

Data B: To understand the demography of new York's boroughs and neighborhood.

In this step to obtain the information on demographics of New York there is availability of Wiki page. Unfortunately, wiki page cannot be used as it is in this exercise. So, the above-mentioned HTML file will be converted into CSV using Web scrapping Beautiful Soup.

Beautiful Soup is a Python package for parsing HTML and XML documents. It works with your favourite parser to provide idiomatic ways of navigating, searching, and modifying the parse tree. It commonly saves programmers hours or days of work. Please refer to the below mentioned link.

LINK: https://en.wikipedia.org/wiki/Demographic_history_of_New_York_City

Data C: Segmentation and clustering of neighborhoods, Farmer Markets, Cafes

For this exercise will use Foursquare API to explore neighborhoods in New York City. Which will help in obtaining coordinate information for each neighborhood of New York. Please refer to below mentioned snapshots for example of data obtained from Foursquare API.

| | name | categories | address | lat | lng | postalCode | state |
|---|--------------------------------------|------------|---------------|-----------|------------|------------|-------|
| 0 | Smyth Hotel | Hotel | 85 W Broadway | 40.715144 | -74.009183 | 10007 | NY |
| 1 | Four Seasons Hotel New York Downtown | Hotel | 27 Barclay St | 40.712612 | -74.009380 | 10007 | NY |
| 2 | The Beekman - A Thompson Hotel | Hotel | 123 Nassau St | 40.711173 | -74.006702 | 10038 | NY |
| 3 | Duane Street Hotel | Hotel | 130 Duane St | 40.715858 | -74.007446 | 10013 | NY |
| 4 | Arlezen Hotel | Hotel | 24 John St | 40.709639 | -74.008685 | 10038 | NY |
| 5 | Cosmopolitan Hotel - TriBeCa | Hotel | 95 W Broadway | 40.715622 | -74.009115 | 10007 | NY |

| | name | categories | address | lat | lng | neighborhood | postalCode | state |
|---|-------------------------------|-------------------------|------------------------|-----------|------------|--------------|------------|-------|
| 0 | American Express HQ Cafeteria | Corporate Cafeteria | 220 Vesey St | 40.713631 | -74.014871 | NaN | 10080 | NY |
| 1 | Spotify Cafeteria Hub - FL 60 | Corporate Cafeteria | 150 Greenwich St | 40.710189 | -74.012075 | NaN | 10007 | NY |
| 2 | Goldman Sachs Cafeteria | Corporate Cafeteria | 200 West St | 40.714716 | -74.014284 | NaN | 10282 | NY |
| 3 | Cafeteria | New American Restaurant | 119 7th Ave | 40.740648 | -73.998176 | NaN | 10011 | NY |
| 4 | Spotify Cafeteria Hub - FL 70 | Food Service | 150 Greenwich St Fl 70 | 40.710318 | -74.011946 | NaN | 10007 | NY |
| 5 | Citi Cafeteria | Café | 388 Greenwich St Fl 2 | 40.720315 | -74.010494 | NaN | 10013 | NY |

Data D: Extra Data collected for this project – New York Cuisine

https://en.wikipedia.org/wiki/Cuisine_of_New_York_City

https://en.wikipedia.org/wiki/List_of_Michelin_starred_restaurants_in_New_York_City

