

Applied Data Science Capstone Project

Metis XU

Data Section

New York City's demographics show that it is a large and ethnically diverse metropolis. With its diverse culture, comes diverse food items. There are many restaurants in New York City, each belonging to different categories like Chinese, Indian, and French etc.

For this project we need the following data:

- ❖ New York City data that contains list Boroughs, Neighborhoods along with their latitude and longitude.
 1. Data Source: https://cocl.us/new_york_dataset
 2. Description: This data set is the most important dataset that contains the information of the neighborhoods and the latitude & longitude as well.
- ❖ List of Chinese restaurants in each neighborhood of New York City.
 1. Data Source: Foursquare API
 2. Description: This data provides all the venues in each neighborhood. Filter the venues to get only Chinese restaurants.
- ❖ GeoSpace Dataset:
 1. Data Source: <https://data.cityofnewyork.us/CityGovernment/Borough-Boundaries/tqmj-j8zm>
 2. Description: Get the NY Borough boundaries that will help us to visualize the choropleth map.

Problems

Using data science methodology and machine learning techniques like clustering, this project aims to provide solutions to answer the main business question: What is best location in New York City for a Chinese Cuisine?

Along with some sub-questions like:

1. Which areas have potential Chinese Restaurants?
2. Which areas lack Chinese Restaurants?
3. Which is the best place to stay if you prefer Chinese Cuisine?