

1. Introduction

Background

The general tendency of Human life is change and people often change their place of stay. It is also very important for them that they select a good and a safe place to live. But most of the times this movement is because they select an inappropriate place to live or get wrong information. This leads to an unhappy and a discomforting life.

To avoid any of these, a proper research has to be done before they shift to a place. Consider the following factors while selecting a new place to live so that you don't end up wasting your valuable money and time making a move you will end up regretting there. Lots of people migrate to New York City like almost 40% of population include immigrants and its one of the growing metropolis in the world.

As many as 800 languages are spoken in New York, making it the most linguistically diverse city in the world. English remains the most widely spoken language, although there are areas in the outer boroughs in which up to 25% of people speak English as an alternate language, and/or have limited or no English language fluency. English is least spoken in neighbourhoods such as Flushing, Sunset Park, and Corona.

Problem

Here we are looking for :

1. Areas with Indian Cuisine Food.
2. Areas which have potential to have more Indian cuisine restaurants
3. Areas which are preferable to stay for Indian Cuisine Restaurants.

Data

The data which we are going to use is :

- **New York City data that contains list Boroughs, Neighborhoods along with their latitude and longitude.**(https://cocl.us/new_york_dataset)

This is dataset about various neighbour hoods of New York City.

- **Indian restaurant in each neighbourhood of New York City.(FOURSQUARE API)**

We use this API for getting map data

- **GeoSpace data**(<https://data.cityofnewyork.us/City-Government/Borough-Boundaries/tqmj-j8zm>)

This dataset is useful for getting all borough boundaries.

Method

We collect all data from New York City dataset and using foursquare API we collect all venues of all neighbourhoods, Then select all Indian Restaurants. We get all tips, ratings from foursquare API Then using scikit learn library get clusters of data using kmeans, using folium library in python we visualise the clusters.

