Introduction

This report will analyse the geographic data in the Manhattan borough of New York City; specifically, the neighbourhoods surrounding Central Park. Central Park is the first public parks built in America and a notorious attraction that attracts 25 million guests a year. The frequent traffic should present ample business opportunity for people looking to open a restaurant business. The potential target audience that may take interest in this include:

* Entrepreneurs intending to start a restaurant in the vicinity of Central Park.
* Small business restaurant owners who could be considering relocating their business to a neighbourhood surrounding Central Park.
* Owners of a restaurant chain looking to expand their business into one of the most popular locations of Manhattan, New York.

By categorizing and clustering restaurants in the surrounding neighbourhoods of Central Park, the analysis will aim to deliver insights into:

* What type of restaurants are popular in the area surrounding Central Park?
* Which neighbourhoods favour specific types of restaurant categories?

This will in turn provide some insight into the restaurant market in local areas around the park and give guidance on the ideal type of restaurant to start and the neighbourhood to start it in.

Data

The analysis will be conducted using data downloaded from <https://geo.nyu.edu/catalog/nyu_2451_34572> which provides a geographic dataset of New York City that covers 5 boroughs and 306 neighbourhoods in total. The analysis will explore the data within the Manhattan borough and find and focus on the neighbourhoods in the surrounding area of Central Park. This dataset features attributes such as the geographic coordinates of the location, name of the neighbourhood it is in and also the New York City borough that it is located in.

An example of the data in json file format is:

{'type': 'Feature',

'id': 'nyu\_2451\_34572.1',

'geometry': {'type': 'Point',

'coordinates': [-73.84720052054902, 40.89470517661]},

'geometry\_name': 'geom',

'properties': {'name': 'Wakefield',

'stacked': 1,

'annoline1': 'Wakefield',

'annoline2': None,

'annoline3': None,

'annoangle': 0.0,

'borough': 'Bronx',

'bbox': [-73.84720052054902,

40.89470517661,

-73.84720052054902,

40.89470517661]}}

Then using location data available through the Foursquare API, the nearest 100 restaurants found within a 500m radius of neighbourhood areas surrounding Central Park will be processed within a Pandas dataframe. A sample of the Foursquare data will show that it has features attributes such as the name of the restaurant, its geographic location in the form of latitude and longitude, the borough and neighbourhood that it is located in, along with its category of restaurant which will usually reflect a type of cuisine.

Sample json:

{'meta': {'code': 200, 'requestId': '5ed0f0e771c428001bf4e546'},

'response': {'suggestedFilters': {'header': 'Tap to show:',

'filters': [{'name': 'Open now', 'key': 'openNow'},

{'name': '$-$$$$', 'key': 'price'}]},

'headerLocation': 'East Harlem',

'headerFullLocation': 'East Harlem, New York',

'headerLocationGranularity': 'neighborhood',

'query': 'restaurant',

'totalResults': 50,

'suggestedBounds': {'ne': {'lat': 40.79674947113033,

'lng': -73.93824946073823},

'sw': {'lat': 40.78774946213033, 'lng': -73.95011500223225}},

'groups': [{'type': 'Recommended Places',

'name': 'recommended',

'items': [{'reasons': {'count': 0,

'items': [{'summary': 'This spot is popular',

'type': 'general',

'reasonName': 'globalInteractionReason'}]},

'venue': {'id': '55a85a66498e68da40bc7212',

'name': 'Malii',

'location': {'address': '2028 2nd Ave',

'crossStreet': 'btwn 104th St & 105th St',

'lat': 40.789524977200365,

'lng': -73.94323226322045,

'labeledLatLngs': [{'label': 'display',

'lat': 40.789524977200365,

'lng': -73.94323226322045}],

'distance': 313,

'postalCode': '10029',

'cc': 'US',

'city': 'New York',

'state': 'NY',

'country': 'United States',

'formattedAddress': ['2028 2nd Ave (btwn 104th St & 105th St)',

'New York, NY 10029',

'United States']},

'categories': [{'id': '4bf58dd8d48988d149941735',

'name': 'Thai Restaurant',

'pluralName': 'Thai Restaurants',

'shortName': 'Thai',

'icon': {'prefix': 'https://ss3.4sqi.net/img/categories\_v2/food/thai\_',

'suffix': '.png'},

'primary': True}],

'delivery': {'id': '315186',

'url': 'https://www.seamless.com/menu/malii-thai-kitchen-2028-2nd-ave-new-york/315186?affiliate=1131&utm\_source=foursquare-affiliate-network&utm\_medium=affiliate&utm\_campaign=1131&utm\_content=315186',

'provider': {'name': 'seamless',

'icon': {'prefix': 'https://fastly.4sqi.net/img/general/cap/',

'sizes': [40, 50],

'name': '/delivery\_provider\_seamless\_20180129.png'}}},

'photos': {'count': 0, 'groups': []}},

'referralId': 'e-0-55a85a66498e68da40bc7212-0'},

{'reasons': {'count': 0,

'items': [{'summary': 'This spot is popular',

'type': 'general',

'reasonName': 'globalInteractionReason'}]},

By organising restaurants in the local areas according to categories and finding the frequency of each category within specific neighbourhoods, neighbourhoods with similar restaurant market can be clustered together according to k-means which can help determine what type and location is best to start around Central Park.