

Coursera Battle of the neighbourhoods

Opening a new Jazz Club in Manhattan

Author: Mark Chinnoek

Date: June 2020

Content

1. Introduction
 - 1.1 Discussion of the "background situation" leading to the problem at hand:
 - 1.2 Problem to be resolved
 - 1.3 Audience for this project.
2. Data
 - 2.1 Data required to resolve the problem
 - 2.2 Data sources and data manipulation

1. Introduction

I am a successful Jazz musician wanting to move to New York City. I have limited knowledge of the area but I would like to open a new venue of my own in Manhattan. I want to locate my venue away from other clubs but in a similar neighborhood.

1.1 Background situation

Where is the best location to open a new Jazz Club in Manhattan? Determine similar neighborhoods to where existing Jazz venues are.

1.2 Business Understanding

The business understanding is to identify which neighborhood(s) are a best location, based on:

- Finding where existing Jazz Clubs are currently located
- Discovering what type of neighborhood category they are in
- Find a similar neighborhood that doesn't currently have a Jazz Venue

1.3 Audience

The target audience for this would be anyone considering moving to Manhattan without detailed knowledge of the area. This report is a specific study on Jazz Club categories but this approach and the search criteria could easily be amended to repeat the exercise for a different category of venue.

2. Data

To investigate this problem we will need the following datasets of information:

- List of neighborhoods in Manhattan
- Location of existing Jazz Clubs and other venues

Using the neighborhood data from the dataset previously provided on this course (https://cocl.us/new_york_dataset) we can obtain the geo coordinates of the Manhattan neighborhoods and store the following information:

Borough	Neighborhood	Latitude	Longitude
6 Manhattan	Marble Hill	40.876551	-73.910660
100 Manhattan	Chinatown	40.715618	-73.994279
101 Manhattan	Washington Heights	40.851903	-73.936900

This data provides:

1. Borough name - eg Manhattan
2. Neighborhood - eg marble hill, highlights the area within Manhattan
3. Latitude - latitude coordinate for mapping
4. Longitude - longitude coordinate for mapping

We can explore the foursquare database using the API call with Manhattan coordinates and retrieve all the existing venues in the area, including Jazz clubs

```
`address = 'Manhattan, NY'
```

```
geolocator = Nominatim(user_agent="foursquare_agent") location = geolocator.geocode(address)
latitude = location.latitude longitude = location.longitude print(latitude, longitude)`
```

```
`search_query = 'Jazz Club'
```

```
print(search_query + ' .... OK!')
```

```
url = 'https://api.foursquare.com/v2/venues/search?client_id={} & client_secret={} & ll={},
{} & v={} & query={} '.format( CLIENT_ID, CLIENT_SECRET, latitude, longitude, VERSION,
search_query, LIMIT )
```

```
results = requests.get(url).json()`
```

Here is an example of the data returned from foursquare search API

```
{'meta': {'code': 200, 'requestId': '5edf37fded78b8001b6bfcdd'}, 'response': {'venues': [{'id':
'54ed2952498ec36f92b69d0c', 'name': "Dizzy's Jazz Club", 'location': {'address': '10 Columbus Cir',
'lat': 40.768764, 'lng': -73.982944, 'labeledLatLngs': [{'label': 'display', 'lat': 40.768764, 'lng':
-73.982944}], 'distance': 3027, 'postalCode': '10019', 'cc': 'US', 'neighborhood': "Hell's Kitchen",
'city': 'New York', 'state': 'NY', 'country': 'United States', 'formattedAddress': ['10 Columbus Cir',
'New York, NY 10019', 'United States']}, 'categories': [{'id': '4bf58dd8d48988d1e7931735', 'name':
'Jazz Club', 'pluralName': 'Jazz Clubs', 'shortName': 'Jazz Club', 'icon': {'prefix':
'https://ss3.4sqi.net/img/categories_v2/arts_entertainment/musicvenue_jazzclub_', 'suffix': '.png'}},
{'primary': True}], 'referralId': 'v-1591687367', 'hasPerk': False}, {'id': '4aca81a1f964a52028c220e3',
'name': "Cecil's Jazz Club & Restaurant", 'location': {'address': '364 Valley Rd', 'lat':
40.77451572910885, 'lng': -74.23987875864671, 'labeledLatLngs': [{'label': 'display', 'lat':
40.77451572910885, 'lng': -74.23987875864671}, {'label': 'entrance', 'lat': 40.774557, 'lng':
-74.239832}], 'distance': 23659, 'postalCode': '07052', 'cc': 'US', 'city': 'West Orange', 'state': 'NJ',
'country': 'United States', 'formattedAddress': ['364 Valley Rd', 'West Orange, NJ 07052', 'United
States']}, 'categories': [{'id': '4bf58dd8d48988d1e7931735', 'name': 'Jazz Club', 'pluralName': 'Jazz
Clubs', 'shortName': 'Jazz Club', 'icon': {'prefix':
'https://ss3.4sqi.net/img/categories_v2/arts_entertainment/musicvenue_jazzclub_', 'suffix': '.png'}},
{'primary': True}], 'referralId': 'v-1591687367', 'hasPerk': False},
```

From this we can extract the following useful features:

1. venue name
2. venue location
3. venue address
4. distance from search coordinates ...

The neighborhood data will be compared using the venue data within to produce clusters of similar neighborhoods. The Jazz Club venues will then be placed in those clusters to see whether there is a pattern of cluster for a venue such as a jazz Club to help form a recommendation.