

# Capstone Project – Moving to New York City

P.G.

21 May 2019

## 1. Introduction/Business Problem

Let us assume I live in Zurich at Helvetiaplatz, a multicultural area of Zurich, and found a new job in Manhattan, New York. However, I really like the spirit of the place I am currently living and therefore I am looking for a similar area in terms of culture, food, etc. to live in Manhattan. Which Neighborhoods in Manhattan should I consider to live in?

This report uses a k-means approach to find similar neighborhoods based on the 10 most common venues in each neighborhood. It can be used to find neighborhoods in a foreign city which are similar to a given area.

## 2. Data

For Manhattan I use the data set from [https://geo.nyu.edu/catalog/nyu\\_2451\\_34572](https://geo.nyu.edu/catalog/nyu_2451_34572) which contains all boroughs and neighborhoods of NYC as well as latitude and longitude of each neighborhood. As in the Lab, I use the data set provided on [https://cocl.us/new\\_york\\_dataset](https://cocl.us/new_york_dataset).

I use the geolocator to find the coordinates of Zurich Helvetiaplatz. The corresponding neighborhood is called Langstrasse.

To get data about the venues around Helvetiaplatz in Zurich as well as around all neighbourhoods in Manhattan, I use the Foursquare location data. Foursquare location data contains venues and their categories as well as coordinates.

## 3. Methodology

### 3.1 Idea

To check similarities between objects, clustering is a common procedure. The idea is to cluster all neighborhoods from Manhattan together with Langstrasse. Those neighborhoods which are in the same cluster as Langstrasse are considered similar to Langstrasse and are therefore eligible for living. I choose the most popular clustering algorithm, k-means. In case results would not be satisfactory, hierarchical clustering could be chosen to find out more about similarity among neighborhoods.

The features used to cluster neighborhoods are given by venue categories from Foursquare: For each venue category within a neighborhood, the portion of venues is calculated. All features have values between 0 and 1. Therefore, no further standardization has to be made.

### 3.2 Cluster Neighborhoods

In this section, I use the Foursquare API to explore the neighborhoods and segment them. Then, I choose an initial k equal to 5 and check in which cluster Langstrasse (or Helvetiaplatz) belongs. It actually turns out, that choosing a larger k results in a situation where Langstrasse is its own cluster, which is not in our favor. Hence, k = 5 is the best choice we can make using k-means.

## 4. Results

Langstrasse is in Cluster Label 0, together with 10 neighborhoods of Manhattan. So it shouldn't be a problem to find a place to live similar to the current home.

The ten neighborhoods which are most similar to Langstrasse in Zurich are depicted in Figure 1:

	Borough	Neighborhood	Latitude	Longitude
0	Manhattan	Marble Hill	40.876551	-73.910660
1	Manhattan	Manhattanville	40.816934	-73.957385
2	Manhattan	Yorkville	40.775930	-73.947118
3	Manhattan	Lenox Hill	40.768113	-73.958860
4	Manhattan	Roosevelt Island	40.762160	-73.949168
5	Manhattan	Upper West Side	40.787658	-73.977059
6	Manhattan	East Village	40.727847	-73.982226
7	Manhattan	Gramercy	40.737210	-73.981376
8	Manhattan	Financial District	40.707107	-74.010665
9	Manhattan	Carnegie Hill	40.782683	-73.953256

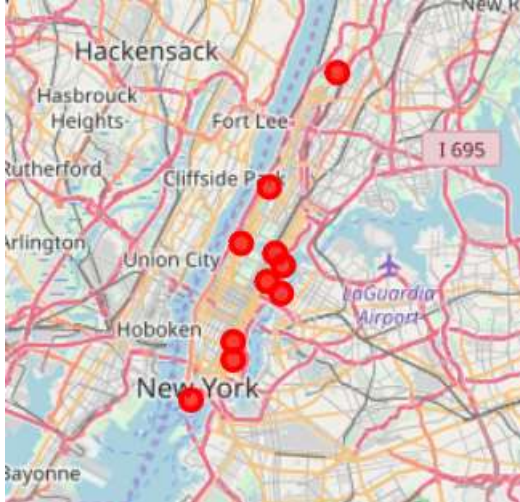


Figure 1: Neighborhoods in Manhattan which are similar to Langstrasse in Zurich.

## 5. Discussion

Out of the 40 neighborhoods of Manhattan, we found 10 neighborhoods similar to Langstrasse in Zurich. This is a good base to start searching for a flat based on the exact place of work as well as current availability and prices of flats.

I think this is a good result as it leaves enough places to look for a flat and can now be narrowed down by other criteria like time to commute, prices and availability of flats, noise, pollution, crime rate, etc.

Hierarchical clustering could have been used alternatively, especially if the amount of similar neighborhoods would have been too large.

## 6. Conclusion

Marble Hill, Manhattanville, Yorkville, Lenox Hill, Roosevelt Island, Upper West Side, East Village, Gramercy, Financial District and Carnegie Hill are suggested as neighborhoods to live in Manhattan which are most similar to Langstrasse (Helvetiaplatz) in Zurich in terms of nearby venues.