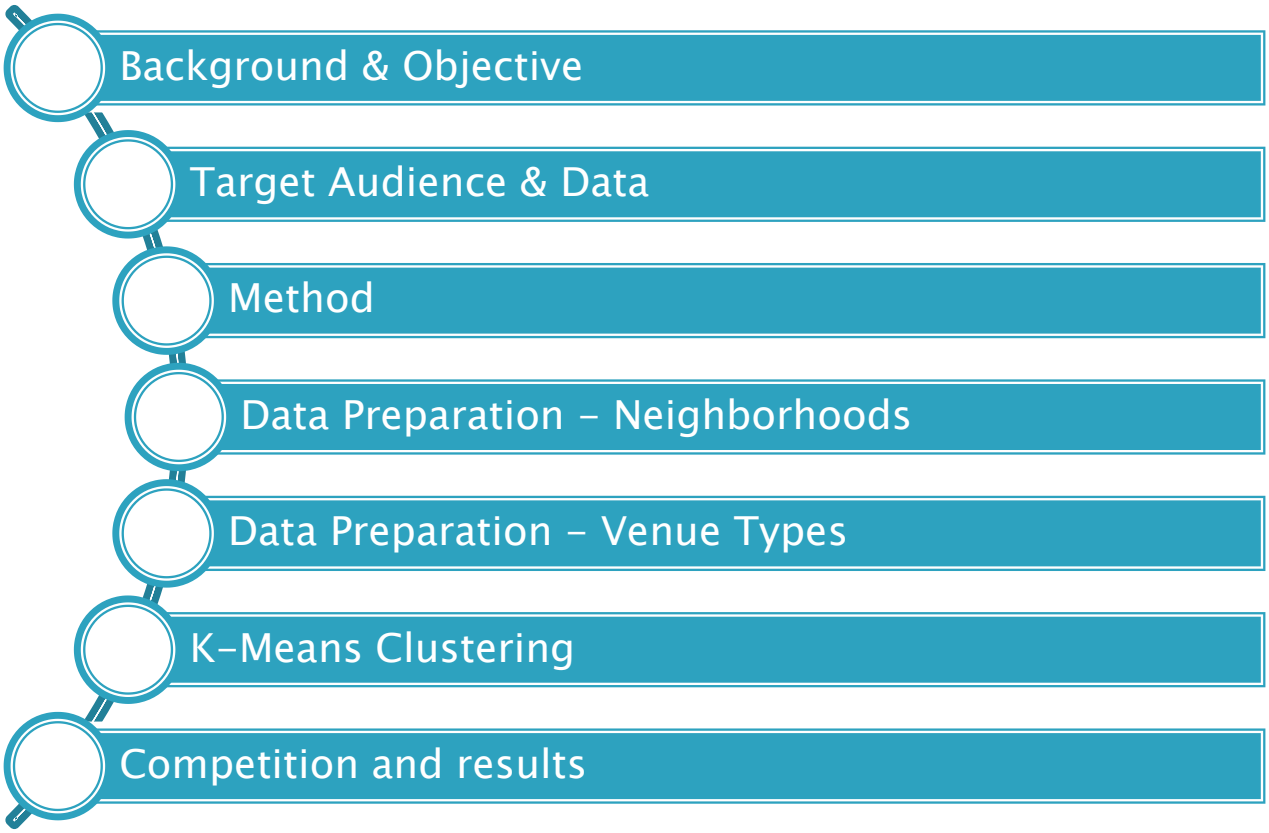
Pizzeria DiMari

Feb 2021 – Coursera Capstone

Presentation Agenda



Background & Objective

Background

Giuseppe DiMari is a renowned business man from little Italy, Manhattan. He has a few restaurants, mainly pizzerias, in that neighbourhood and he is thinking on expanding his operation into other neighbourhoods in other towns in America although he does not discard further growth in NY. Giuseppe has some friends in Chicago, Miami and Toronto. Thus he would like to select which neighbourhoods are the most similar to Little Italy based on the competition and nearby venues so that he can create a business in there. He is not just interested in one only neighbourhood, he wants a group of neighbourhoods he can chose from. If there are also neighbourhoods in NY that are similar to Little Italy, he would like to know those as well. Also, he is not particularly afraid of the competition, if anything, he prefers a neighborhood with competition since he know that this will help his Pizzeria to stay at its best. After all, it will be his inexperienced brother, Julio DiMari, who will be managing the restaurant. And he needs some healthy competition to become a successful business man.

Objective

Identify which neighbourhoods are the most similar to Little Italy(NY) in Miami, Chicago, Toronto and the broader neighbourhoods from NY. From these, identify which ones have enough competition.

Target Audience & Data

Target Audience

The target audience is just Giuseppe and possibly Donny. Perhaps the brother Julio will also be there but this will be just a project of business owenrs seeking clear guidance on similar neighboorhoods to Little Italy.

Data

First, it will be downloaded from Wikipedia and previous modules the neighborhoods from NY, Chicago, Toronto and Miami. For NY and Toronto we will use the data from other modules. For Miami and Chicago the data will be downloaded from wikipedia

After the names of the neighborhoods is obtained, we will obtain the latitude and longitude of each neighborhood. We will also double check that no error has occured (ie: neighborhoods located in LA).

After the neighborhoods and geoscpacial locations of these is obtained and these have been confirmed to belong to each city, we will download from Foursquare the 100 nearest venues to each geospacial address.

Method

1. We will obtain all the Neighbourhoods from Toronto, Miami, Chicago and NY. Neighbourhood data will be downloaded from Wikipedia or use existing data from other course modules. This data will be cleansed to ensure that these hoods are indeed in these cities. Based on my experience, Geolocator can give wrong locations if the address isn't sufficiently clear.

2. We will obtain all the venues per neighbourhood from Foursquare. We will characterise each neighbourhood using the most common types of venue. This will be done by using the one hot encoding.

3. We will classify the neighbourhoods into clusters based on the types of common venues using K-mean clusters. Number of clusters will also be optimised so that there are no more than 25 similar neighbourhoods in that group. The list of neighoorhoods has to be sufficiently small for Giuseppe,

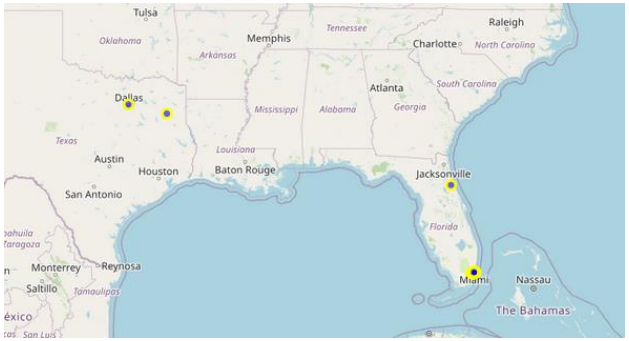
4. The cluster that contains Little Italy, NY, will be the target subgroup of neighbourhoods.

5. After clustering, we will find those neighborhoods that do have a pizzeria or an italian restaurant in the top 3 common venues.

Data Preparation- Neighbourhoods

1. We downloaded from Wikipedia the neighbourhoods for all relevant cities. Geolocation gave some incorrect addressess (latitude and longitude) in the case of Chicago and Miami. It can be seen that in the case of Chicago some hoods were incorrectly located outside the City. Similar with Miami.

Chicago neighbourhoods



Miami neighbourhoods

2. To fix this, a restriction was imposed on the addresses (Lat, Long) so that any neighourhood located outside these is excluded from the dataset.

3. Resulting neighbourhoods for all cities are shown below.

NY, Chicago, Toronto and Miami neighbourhoods

There are 406 neighbourhoods to compare

Data Preparation – Venue Types

1. We downloaded for each neighbourhood a maximum number of venues (100 per neighbourhood). This was downloaded through foursquare.

There are 13081 Venues in 406 neighbourhoods

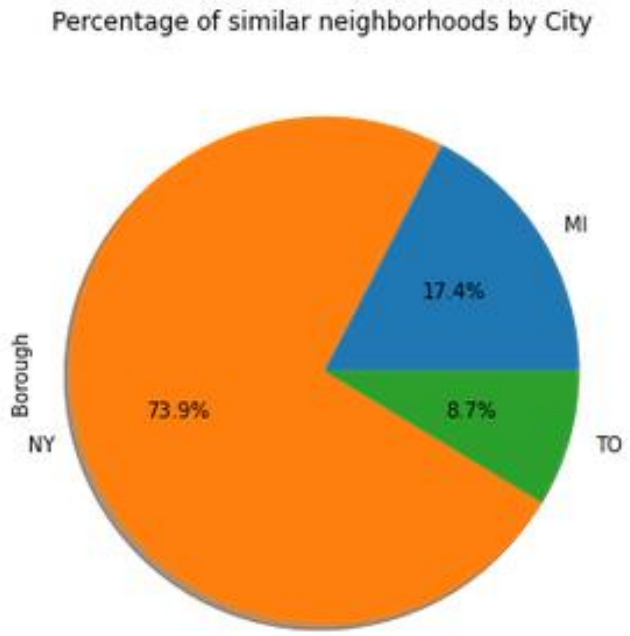
2. Once this data was downloaded, we performed – one hot encoding – for the venue types

There are 468 different venue types in 13081 venues

K-Means Clustering

1. We performend the K means clustering with groups of 5, then 50 and finally of 200 clusters. The idea was to reduce the number of neighborhoods to less than 25 in the cluster of Little Italy. With K=5, there were more than a 100 similar neighborhoods thus an interative proces was performed to get to the suitable group size.

2. Finally, with K=200, there were only 23 similar neighborhoods to Little Italy. With this, the results show that.
    1. Most of the more similar neighborhoods to Little Italy, are actually in NY
    2. Within NY, the neighborhoods more similar to Little italy are actually in Manhattan.
    3. There is no similar neighborhood to Little Italy in Chicago
    4. Between Miami and Toronto, there are more similar neighborhoods in Miami



Similar neighborhoods to Little Italy in NY.

Competition & Results

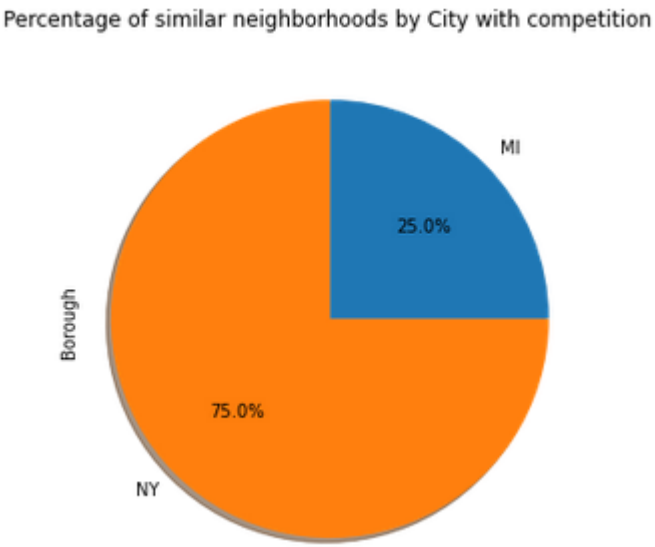
1. From the last cluster, we searched which neighborhoods have a pizzeria or an italian restaurant on the top 3 common venues. We included italian restaurants since these might serve pizza too.

2. With this, the results of this show that.
    1. Most of the competition is actually in Manhattan, NY.
    2. Threre are no neighborhoods in Toronto with this criteria (neither Chicago)
    3. In Miami there is also one neighborhood that shares the characterstics of Little italy and the competion desired.

Final Results

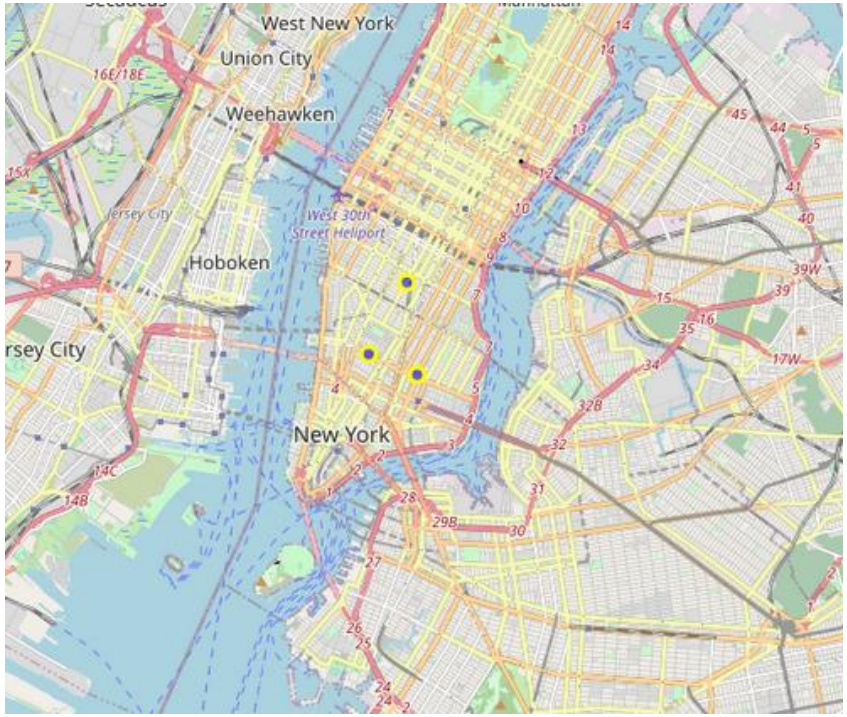
There are 23 neighboorhoods with similar venue profile as little Italy, Manhattan. Out of these, there are only four neighboorhoods with the desired level of competition. These neigbhoorhoods are:

	Borough	Neighborhood	Latitude	Longitude	City
209	Manhattan	Greenwich Village	40.7269	-73.9999	NY
340	Manhattan	Noho	40.7233	-73.9884	NY
368	Manhattan	Flatiron	40.7397	-73.9909	NY
399	Brickell	Brickell	25.7634	-80.1955	MI

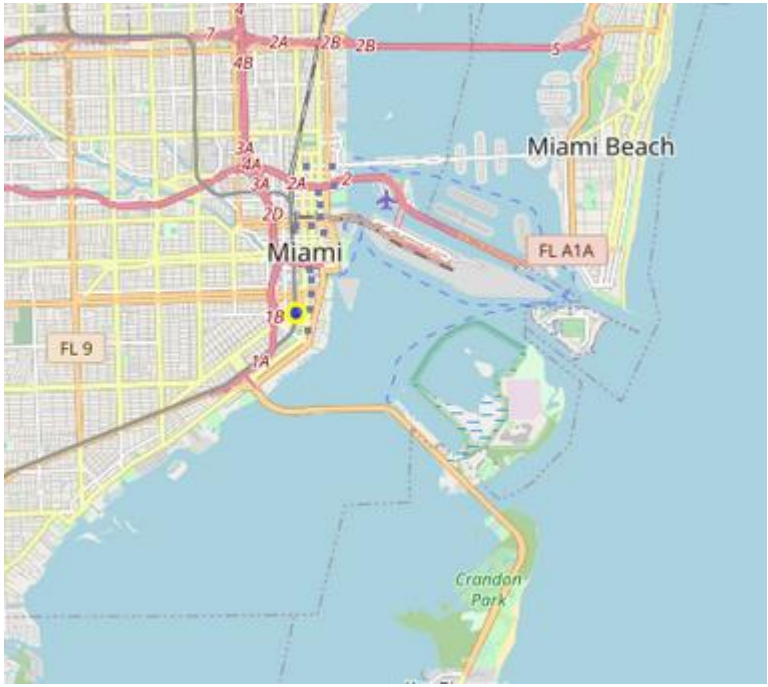


Location of Neighborhoods

NY, Manhattan



Miami



For completeness, the location of these neighborhoods are shown in the images above.

Further Questions?