

# Final Coursera Capstone Project – The Battle of Neighborhoods (Week 1)

## Chicago IL, USA



**Chicago** is the most populous city in the state Illinois, and the 3<sup>rd</sup> largest city in the U.S. with estimated population of 2.7M (2018). Chicago is famous as 'Windy City'. It's 58M domestic & international visitors in 2018 made it the second most visited city in the U.S. (as compared with NYC's 65M in 2018 – Wikipedia)

### 1. Description of the problem and a discussion of the background:

#### A. Background of the City of interest:

Whenever a visitor or likely investors search for new city, they are mostly interested in the best places – areas of concern city have to offer. They are eager to know about amount of best interesting venues concentrated, and which specific Community Areas (neighborhood) possess it.

Chicago as large populated city has expanded its boundaries over all the sides. As per collected dataset Chicago has 77 Community Areas (neighborhoods). We can explore venues of all Community Areas by API, but it can be somehow confusing. So, we have selected 10 neighborhoods which are most attractive and close to Downtown area of city. We will explore these areas for venues that can be more helpful to our audience.

#### B. Interested Audience:

To solve the problem of our target audience – mass of visitors of Chicago city and who are the proposed investors to start new businesses like Café, any Continental Restaurant Bistro or any kind of new venture of their interest could be interested in our data analysis. This data can be very helpful to reach at worth decision.

## 2. A description of the data and how it will be used to solve the problem:

The data we have used to create this project is retrieved from various sources:

Sources of Data:

- The City of Chicago – Data Portal
- Data of Census – Data Portal
- Boundaries of Neighborhoods
- FourSquare API
- Python Geocoders
- Python Folium

Method to extract and use data:

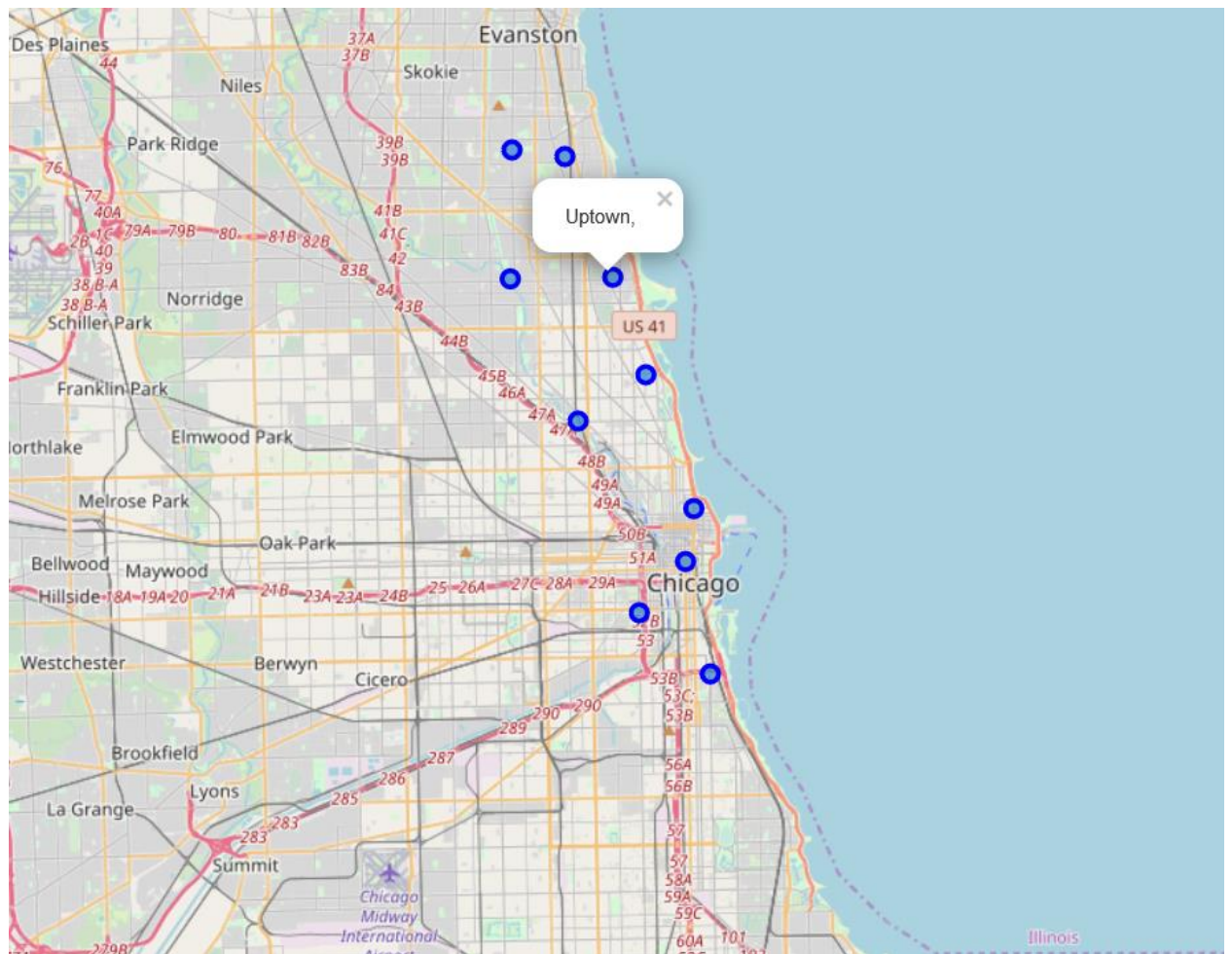
The City of Chicago – Data Portal is free and easily available source to extract names of City's Community Area (neighborhood) names. We can get the geographical coordinates of the all neighborhoods using Python Geocoders. It can give the Latitude & Longitude of the each neighborhood. Also we will get help of FourSquare API to have the venue data of neighborhoods. FourSquare have largest database of 150M+ places & widely used by Data Scientists and Developers in world. It provides various categories of the venue data.

In this project we will use different Data Science skills, Folium to create maps of city and neighborhoods and necessary Machine Learning Techniques.

	AREA NAME	Latitude	Longitude
0	Rogers Park	42.007012	-87.677996
1	West Ridge	42.008751	-87.699768
2	Uptown	41.969826	-87.657637
3	Lincoln Square	41.969389	-87.700489
4	Lake View	41.939539	-87.644384
5	Lincoln Park	41.925098	-87.672294
6	North Side	41.897983	-87.624096
7	West Side	41.865761	-87.646876
8	Loop	41.881598	-87.627758
9	South Side	41.846860	-87.617324

Map of Chicago using Folium package of Python: On next page

Map of Chicago using Folium package of Python:



Thanks.