

Settling Down in Chicago

Introduction

Suppose I'm a real estate broker in Chicago and I have a client who will be moving to Chicago but know little of this city. The client is of upper-middle class and has a family of four, he and his wife, together with their two young kids.

I am going to give him advice on **which community area to settle down in Chicago**. As far as I'm concerned, it has to meet 2 requirements,

- 1) **Safe**. Since we all know that Chicago is by no means a safe city, so it is the first thing that we would consider.
- 2) **Relaxed**. As we are looking for a community to live a life, this community should provide a calm and relax environment, and of course with sufficient venues to support daily life, such as dry cleaning, restaurants, etc.

Data

- **Regarding safety**

There are 77 communities in Chicago. In order to get the information of the safety level for each community, I looked for statistics in <https://data.cityofchicago.org/>. There is **a dataset reflects reported incidents of crime that occurred in Chicago** from 2001 to present, and for simplicity, I just downloaded the record for the year 2018.

But the above dataset only has community areas in numerical form. In order to get the names for each community, I have to scrape a website from Wikipedia to

match the numbers with the names. The website is https://en.wikipedia.org/wiki/Community_areas_in_Chicago .

With these statistics in hand, I could solve the problem of finding safe communities for my client. Let's just define that the communities which have crime incidents less than the average level is safe.

- **Regarding relaxation**

As for the second requirement, I'll turn to <https://foursquare.com/> to **separate the safe communities into 3 clusters**. Here, I'll use the **K-Means Clustering** method of machine learning to find out each cluster's characteristics and to decide which community/communities to recommend to my client.