Capstone Project – The battle of Neighborhoods Report

Table of Contents

- Introduction
- Data Acquisition

Introduction

1.1 Background

Chicago is the most populous city in state of Illinois, and third most populous city in USA. It is quite evident that a lot of people migrate to Chicago from different parts of USA probably due to better opportunities and could be for many other reasons. As people move to Chicago for several reasons, this analysis on crime rate in and around Chicago would help people in making well-informed decisions.

1.2 Problem

In this report, we are looking at finding the safest neighborhoods In Chicago based on the least crime rate. We will also, look at the neighborhood of Chicago using Foursquare API, which would give us the most popular venues in and around the neighborhood of Chicago.

1.3 Interest

This report would help all the people who are considering migrating to neighborhoods of Chicago. Majority of the people have some preferences such as safest neighborhood, closer to restaurant or any food joint, or may be a Train/bus Station. This would help people in making a quick and well-informed decision making.

2. Data Acquisition

The dataset used here is obtained from Kaggle – Chicago Crime (https://www.kaggle.com/currie32/crimes-in-chicago/download). Below is the description of the dataset.

ID - Unique identifier for the record.
Case Number - The Chicago Police Department RD Number (Records Division Number), which is unique to the incident.
Date - Date when the incident occurred. this is sometimes a best estimate.
Block - The partially redacted address where the incident occurred, placing it on the same block as the actual address.
IUCR - The Illinois Unifrom Crime Reporting code. This is directly linked to the Primary Type and Description. See the list of IUCR codes at https://data.cityofchicago.org/d/c7ck-438e.
Primary Type - The primary description of the IUCR code.
Description - The secondary description of the IUCR code, a subcategory of the primary description.
Location Description - Description of the location where the incident occurred.
Arrest - Indicates whether an arrest was made.
Domestic - Indicates whether the incident was domestic-related as defined by the Illinois Domestic Violence Act.
Beat - Indicates the beat where the incident occurred. A beat is the smallest police geographic **area** – each beat has a dedicated police beat car. Three to five beats make up a police sector, and three sectors make up a police district. The Chicago Police Department has 22 police districts. See the beats at https://data.cityofchicago.org/d/aerh-rz74.

District - Indicates the police district where the incident occurred. See the districts at https://data.cityofchicago.org/d/fthy-xz3r.

Ward - The ward (City Council district) where the incident occurred. See the wards at https://data.cityofchicago.org/d/sp34-6z76.

Community Area - Indicates the community area where the incident occurred. Chicago has 77 community areas. See the community areas at https://data.cityofchicago.org/d/cauq-8yn6.

FBI Code - Indicates the crime classification as outlined in the FBI's National Incident-Based Reporting System (NIBRS). See the Chicago Police Department listing of these classifications at http://gis.chicagopolice.org/clearmap_crime_sums/crime_types.html.

X Coordinate - The x coordinate of the location where the incident occurred in State Plane Illinois East NAD 1983 projection. This location is shifted from the actual location for partial redaction but falls on the same block.

Y Coordinate - The y coordinate of the location where the incident occurred in State Plane Illinois East NAD 1983 projection. This location is shifted from the actual location for partial redaction but falls on the same block.

Year - Year the incident occurred.

Updated On - Date and time the record was last updated.

Latitude - The latitude of the location where the incident occurred. This location is shifted from the actual location for partial redaction but falls on the same block.

Longitude - The longitude of the location where the incident occurred. This location is shifted from the actual location for partial redaction but falls on the same block.

Location - The location where the incident occurred in a format that allows for creation of maps and other geographic operations on this data portal. This location is shifted from the actual location for partial redaction but falls on the same block.

- We will be looking at how the crime rate has changed over the years.
- We will study the most recent year to see how the city has evolved over the years.
- We will compare the most recent year with its neighborhood and cluster them using clustering technique.

From this dataset, we are going to use Year 2016 for finding the crime rate based on District (77), as we have the district numbers, we will obtain the Names of district from Wikipedia (https://en.wikipedia.org/wiki/Community_areas_in_Chicago). This would help us in getting the neighborhoods of districts.

- District Number
- Community Name

Using the Community Name, we will be able to use Foursquare API to obtain the neighborhoods of Chicago.