Syracuse University

Memo

|  |  |
| --- | --- |
| To: | Dr. Landowski |
| From: | David Madsen & Gayathri Sanjeev |
| Date: | May 10, 2019 |
| Re: | Project Proposal |

**Topic: Analyze the effects of weather on crime in Chicago**

**Data Description:**

The main data consists of Chicago crime incidents recorded from 2001 to 2017 (<https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-present/ijzp-q8t2>).

|  |  |  |
| --- | --- | --- |
| **Fields** | **Description** | **Example** |
| ID | Unique identifier for the record | 4677901 |
| Case Number | The police department case number (unique to the incident) | HM275615 |
| Date | Date incident was occurred (or best estimate) | 05/01/2003 01:00:00 AM |
| Block | Partially redacted address indicating the block where the crime occurred | 111XX S NORMAL AVE |
| IUCR | Illinois Uniform Crime Reporting Code (directly linked “Primary Type’ and ‘Description’ below | 0841 |
| Primary Type | The primary description of the IUCR code | THEFT |
| Description | The secondary description of the IUCR code | FINANCIAL ID THEFT: $300 &UNDER |
| Location Description | Description of the incident location | RESIDENCE |
| Arrest | Indicates if an arrest was made | False |
| Domestic | Indicates if the incident was domestic related (defined by the Illinois Domestic Violence Act) | False |
| Beat | The police beat where the incident occurred | 2233 |
| District | The police district where the incident occurred | 22.0 |
| Ward | The ward (or City Council district) where the incident occurred | 34.0 |
| Community Area | The community area (of 77 in Chicago) where the incident occurred | 49.0 |
| FBI Code | The NIBRS database classification of the crime | 06 |
| X Coordinate | A partially shifted x coordinate for the location | 1174948.0 |
| Y Coordinate | A partially shifted y coordinate for the location | 1831051.0 |
| Year | The year the incident occurred | 2003 |
| Updated On | Record last updated time | 04/15/2016 08:55:02 AM |
| Latitude | Partially shifted latitude from where the incident occurred | 41.691784636 |
| Longitude | Partially shifted longitude from where the incident occurred | -87.635115968 |
| Location | Partially shifted location formatted for mapping applications | (41.691784636, -87.635115968) |

This data can be combined with weather data from Chicago Midway Airport (2008 - 2017). This dataset gives hourly, daily and monthly aggregations of several weather statistics ranging from temperature to visibility and humidity. (<https://www.kaggle.com/yochanan/chicago-weather>)

We may also consider US Census data to compare crimes changing demographic data.

**Research Questions:**

* Do certain types of crime seem more frequent in higher temperatures?
* How do different types of precipitation affect crime?
* How does gentrification affect certain type of crime?

**Data Preparation Plan:**

1. Load the data into Python
2. Select data columns from each dataset for analysis
3. Select groups of data from each dataset which will be used for analysis
4. Determine common key columns for linking datasets
5. Combine data into a single data set
6. Examine resulting dataset for anomalies, missing values, etc.
7. Fix issues mentioned in step 6.
8. Consider features to extract from the data such as day of week if we choose to analyze periodicity of incidents.