

Crime Analysis of Cities in the USA

Installation and Usage Document

Data Extraction and Pre-processing

Download the datasets (in .csv format) from the following links:

- Austin: <https://data.austintexas.gov/Public-Safety/Crime-Reports/fdj4-gpfu>
- Baltimore: <https://data.baltimorecity.gov/Public-Safety/BPD-Part-1-Victim-Based-Crime-Data/wsfq-mvij>
- Chicago: <https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-Present/ijzp-q8t2>
- LA City: <https://data.lacity.org/A-Safe-City/Crime-Data-from-2010-to-2019/63jg-8b9z>
- Rochester: <https://data-rpdny.opendata.arcgis.com/datasets/rpd-part-i-crime-2011-to-present>

Rename the datasets by their city name. There should be 5 CSV files like {Austin.csv, Baltimore.csv, Chicago.csv, LACity.csv, Rochester.csv}

How to Run the Program?

1. Open the *config* folder.
2. Set database name, hostname and port number in the *connection.json* file
3. Set the base path containing the csv files in *config/path.json*
4. Open the terminal and run the program as follows:

```
$ python3 preprocess.py
```