

# Predictive Policing in Chicago

## Data Description

The dataset contains a comprehensive list of the crimes committed in Chicago from 2001 to March xxx 2017. It was downloaded from [data.cityofchicago.org](https://data.cityofchicago.org/)<sup>1</sup>. The auxiliary metadata includes a wide variety of information such as the type of crime (burglary, homicide), the place (bars, mall), communities/districts, whether an arrest was made, geolocation data e.t.c

## Objective

The aim of this project is to analyse the impact of using hotspot policing as the foundational basis for modern predictive policing technology (PPT).

This project is split into three main components:

1. Pre-processing (data cleaning) and descriptive statistics to get a holistic view of different crimes over the dataset 16 year time period . We will use a heatmap to highlight the clear hotspots for criminal activity and plot the yearly changes in the total number of crime, (split into domestic and violent crimes), number of arrests, crime rate per neighbourhood. An example of this is shown in Figure1.
2. Develop and test a predictive policing algorithm that will predict the location of crimes before they are committed. This will be shown on a map with a bounding box to include an acceptable margin for error. The dataset from 2001 to 2012 will be used to train our model and we will test our model from 2013 to the present day by comparing the location (latitude and longitude) of a crime with the one predicted by our algorithm.
3. Check the effectiveness of crime reduction in areas that have received an increase in police presence based on predictive policing technology with similar crime dense areas that have not only been subjected to PPT but also has been supplemented with active community support, investment and rehabilitation from organisations such as the non-profit Local Initiatives Support Corporation (LISC)

## Background

Historically, police patrols have been conducted randomly, which meant that the criteria for a successful patrol was whether or not an incident occurred/an arrest

---

<sup>1</sup><https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-present/ijzp-q8t2/data>

was made. Predictive Policing aims to optimise human resources by reducing the patrol regions and guiding police officers to specific areas based on previous crime data by neither targeting individuals nor using demographic data. The rationale is that the presence of the police will be a deterrent for future crimes, and if insufficient, police can respond a lot quicker as they are already in the neighbourhood. Over time, crime will decrease as the area becomes too “hot for criminal activity”. This is known as the theory of change model<sup>2</sup>.

## **Drawback of Predictive Policing Technology**

There is a high risk that officers become over-reliant on this new technology. In addition, the mere fact that police are searching for a crime in targetted areas already leads them towards probable cause; therefore, if an officer is in “action mode” or is personally discriminatory, the chances of making an arrest are higher.

In addition, crime rate is defined by crimes that were caught and recorded, rather than crimes that actually occurred; therefore, as more police officers are allocated to low-income areas which tend to be “hotspots”, more affluent areas could become a haven for crime, yet remain undetected as there will be little police presence, hence, be incorrectly classified as having little/no crime rate.

## **Hypothesis**

PPT must be supplemented by different forms of community engagements and improvement schemes such as those by organisations like LISC in order to improve the local areas and ensure that, in the long run, the propensity for crime in these hotspot areas is reduced.

This hypothesis will be tested by considering hotspot areas where LISC are actively involved and where they have little/no presence.

---

<sup>2</sup><https://www.ncjrs.gov/pdffiles1/nij/grants/239207.pdf>