New York City Crime Analysis

Author : Mo. Helmi Boudhraa

1. Introduction

1.1 Background

New York is one of the biggest cities in the world having a large population density and relatively high crime rate making the job of law enforcement agents extremely difficult.
One way to help analyze and predict crimes in New York is by exploring various data about the city, a process made easier by the NYC Open Data Project that makes the wealth of public data generated by various New York City agencies and other City organizations available for public use.

1.2 Problem

Crime analysis based on traditional socio-demographic data is of limited value because it fails to capture the complexity and dynamicity of human activity in cities. With the rise of ubiquitous computing, there is the opportunity to improve crime analysis (and prediction) with crowdsourced data (such as Foursquare data) that make for better proxies of human activity.

1.3 Interest

- Law Enforcement agencies are really interested / encourages any crime analyzing / prediction made by the public which will be helpful in their war against crime in the city.

- We will be doing some Venues - Crimes analysis so people that want to visit “safe” venues (we will discuss the criteria of safety later on) will be interested.

2. Data acquisition and cleaning

2.1 Data sources

- **New York City official Incidents Data : NYPD Complaint Data Current (Year To Date)**
  This Data contains a lot of interesting information including the coordinates (longitude, latitude) of each incident which we can exploit using Foursquare API.

- **New York City Population Data : NYC Population by Borough**
  This Data will be helpful in calculating the Incidents / Population / Borough

- **New York City Census Data : NYC Census Data**
  This Data will allow us to explore The (Demographics - Crime) relation.
  We will explore this data and extend the results with data acquired from Foursquare API.

- **Foursquare API Data : Foursquare API**
  We will be using The **Explore** feature of Foursquare API to fetch Nearby Venues for each Incidents and analyze the resulting data.