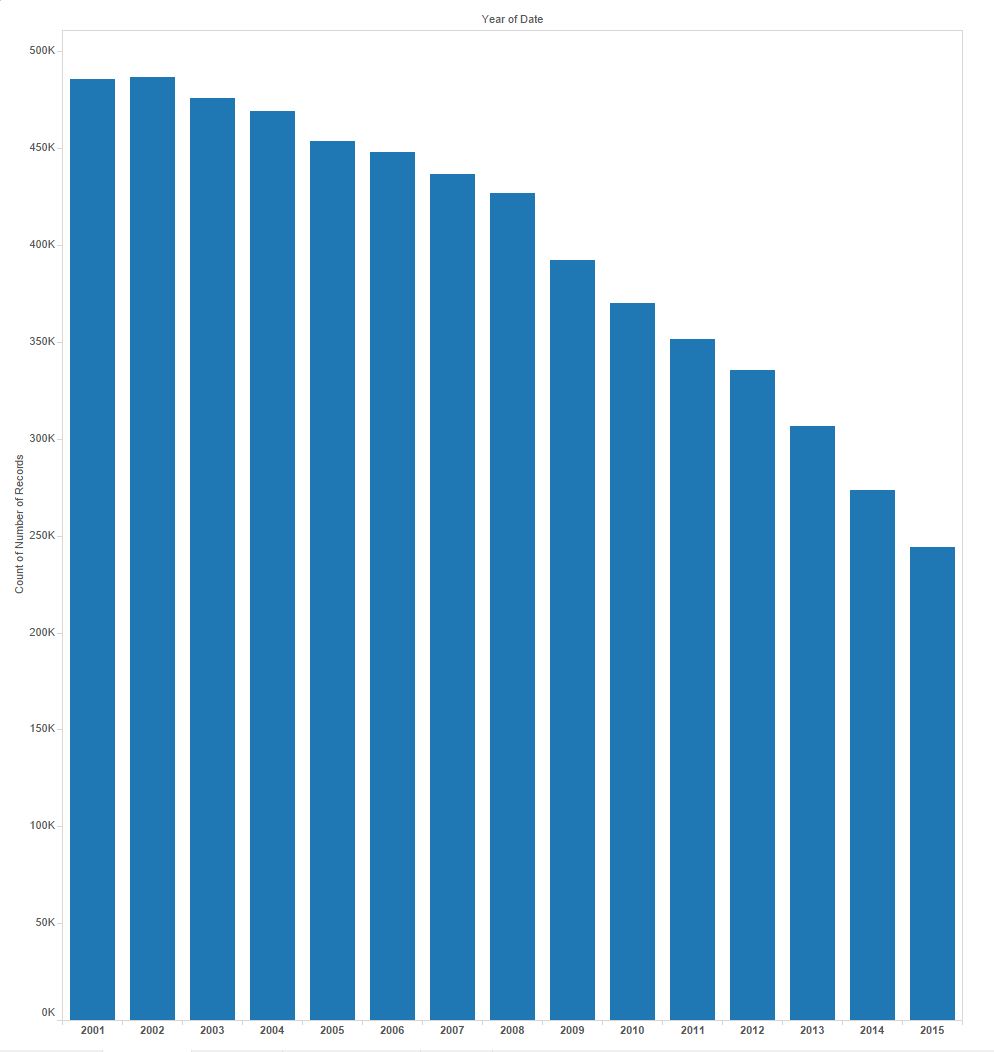
Chicago Crime

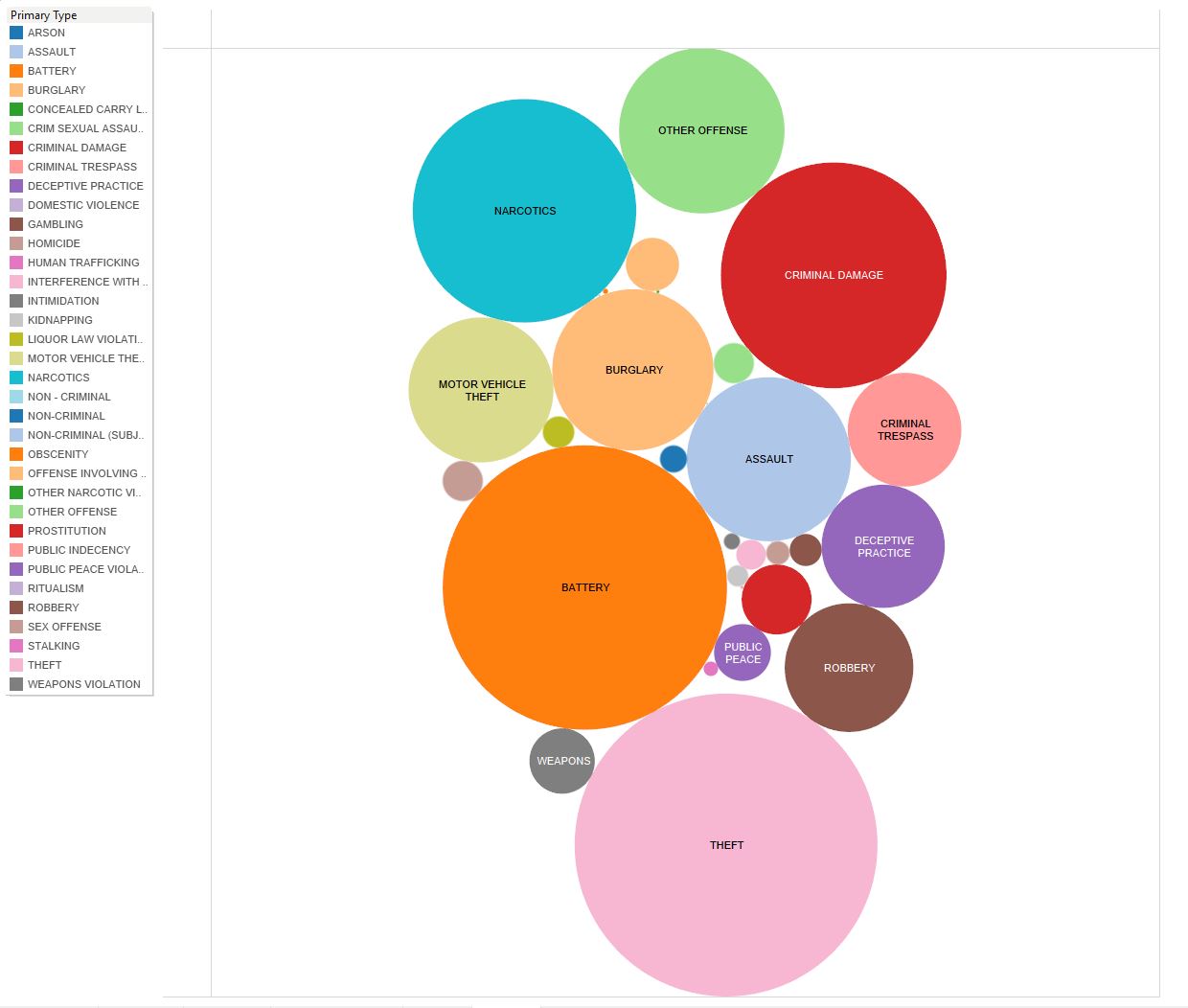
Crime in Chicago has been tracked by the Chicago Police Departments Bureau of Records since the beginning of the 20th century. Starting in the 1960’s, Chicago saw a major rise in the violent crime. The crimes peaked in 1974 with 970 murders when the city’s’ population was over 3 million, resulting in a murder rate of around 29 per 100,000, and also again in 1992 with a murder rate of 34 murders per 100,000 citizens. Using a dataset publicly available by the city of Chicago, I wanted to see how the trend of crime has progressed since 2001. Chicago police department classified crime based on 30 distinct categories. Based on the description of the crime, I classified the crime as violent or nonviolent.

Initially I used Python a little bit for filtering the data and cleaning it up, but for the most of the project I used Tableau. Tableau is a fairly new software that can translate data into pictures of optimized database inquiries. This makes it easier to see patterns, identify trends, and discover visual insights easier than before. Tableau has a very powerful data visualization tools that will let even beginners make sense of GB’s of code and present it in a way that is easier to understand.

**Crime By Year**

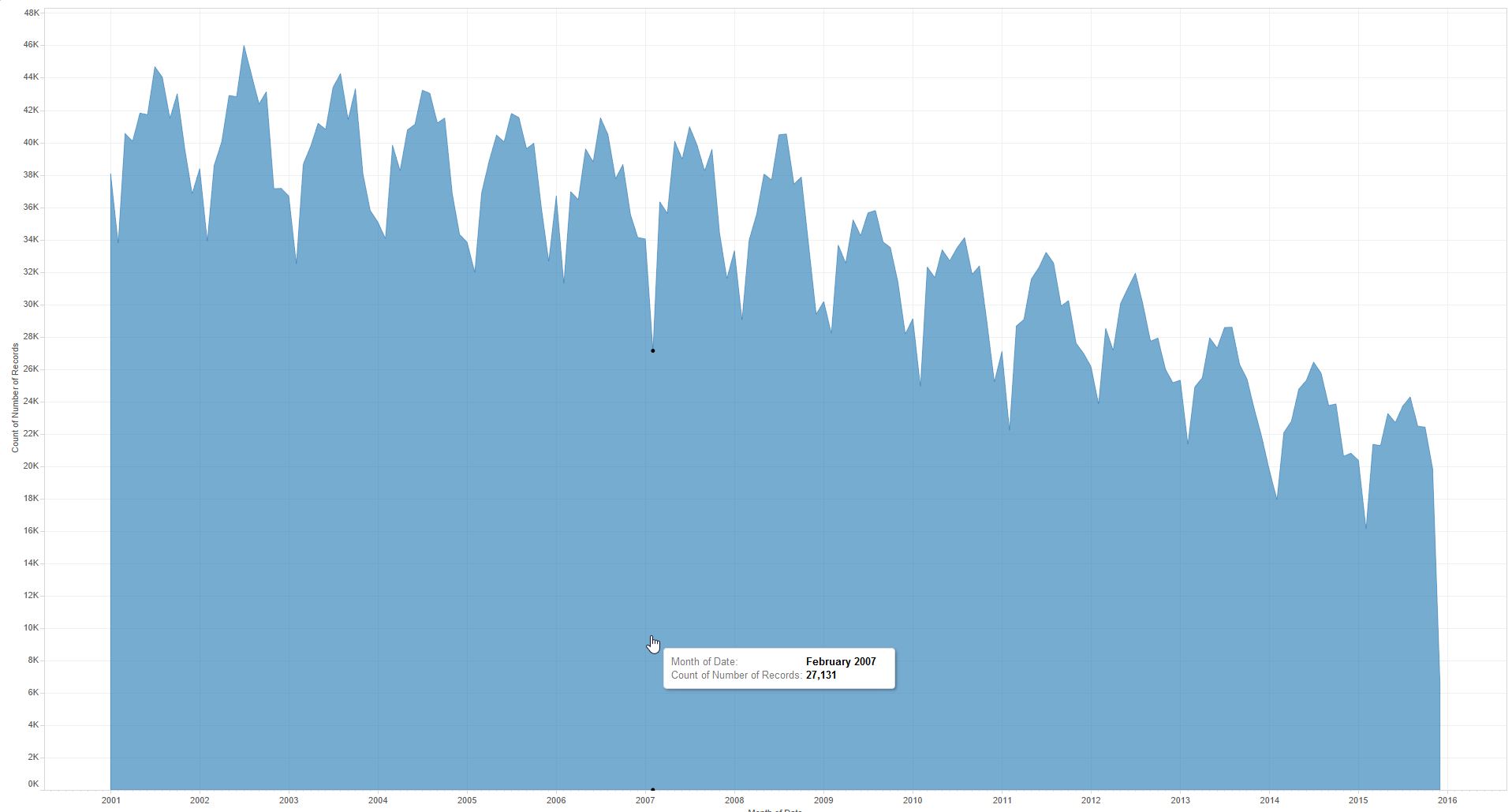
This is a simple graph showing the total crime data trend throughout the years. The graph shows that crime total in Chicago is still decreasing every year. This follows suit and seems that since the 1992, Chicago is seemingly becoming an ever safer city to live in.

**Crime By Type**



This next graph shows the different classifications of crime that the Chicago Police Department uses. The larger the circles the more incidents of that crime, and the smaller circles means less incidents.

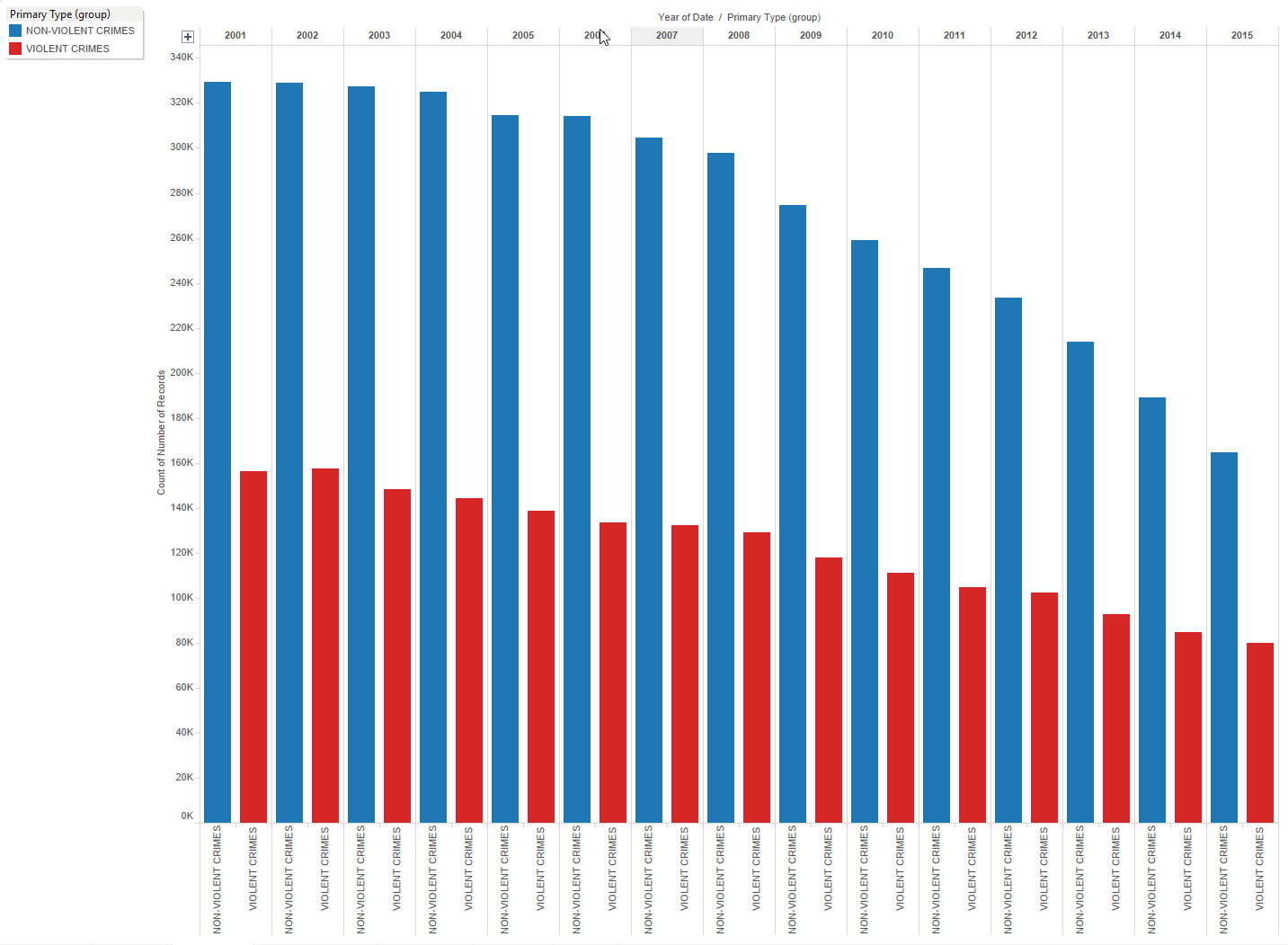
**Crime Peaks By Year**

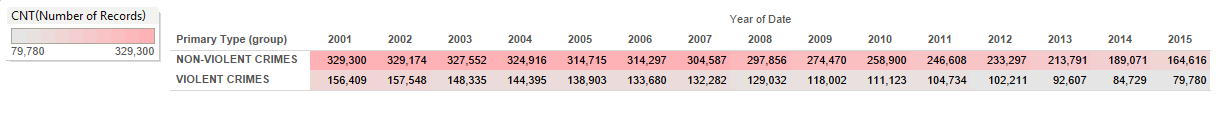


This is a graph showing peak times during the year when crime was highest. Every year since 2001, crime has been highest between quarter 2 and quarter 3, so during the summer months of June, July, and August.

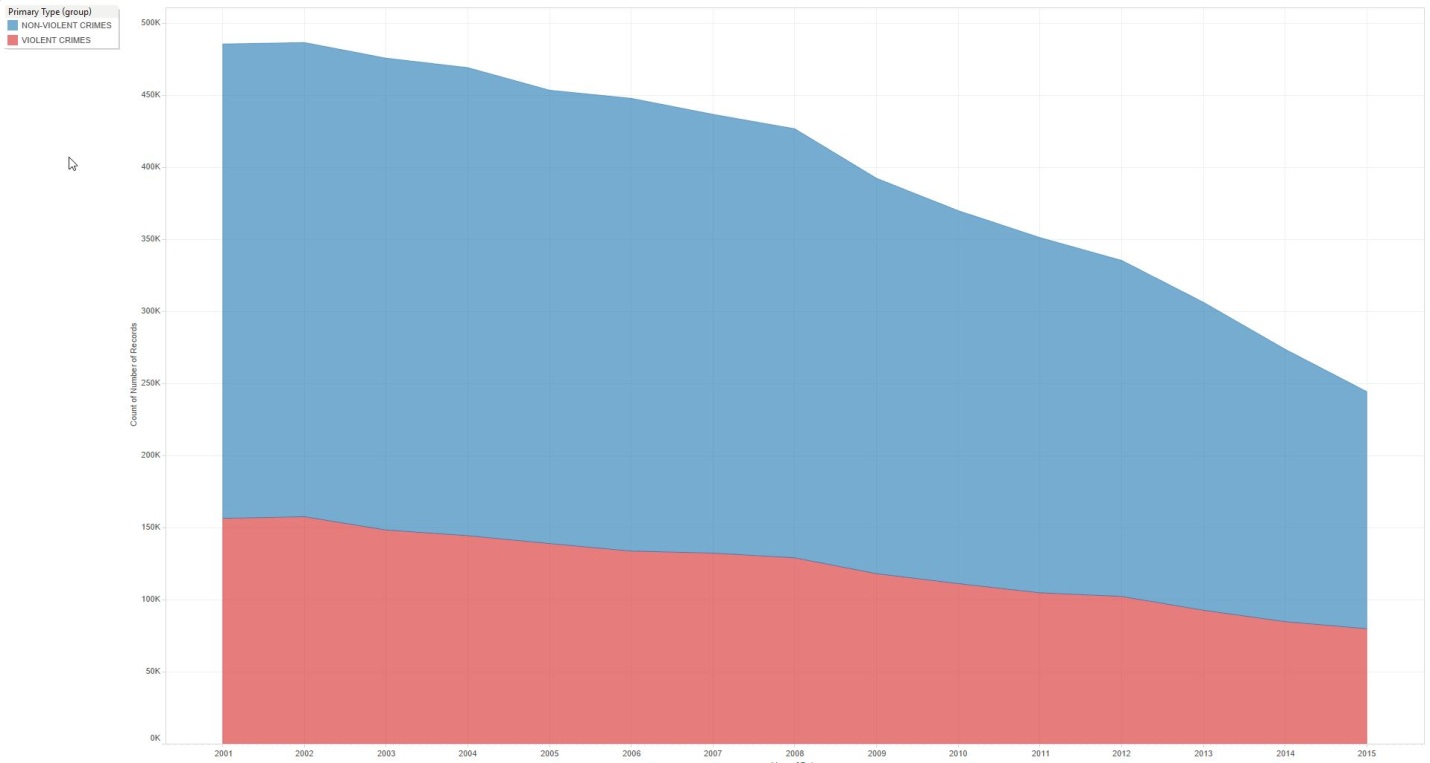
The crimes were classified as violent or nonviolent as defined by the following chart. The classifications were based on a description of the crimes that were present in the original dataset.

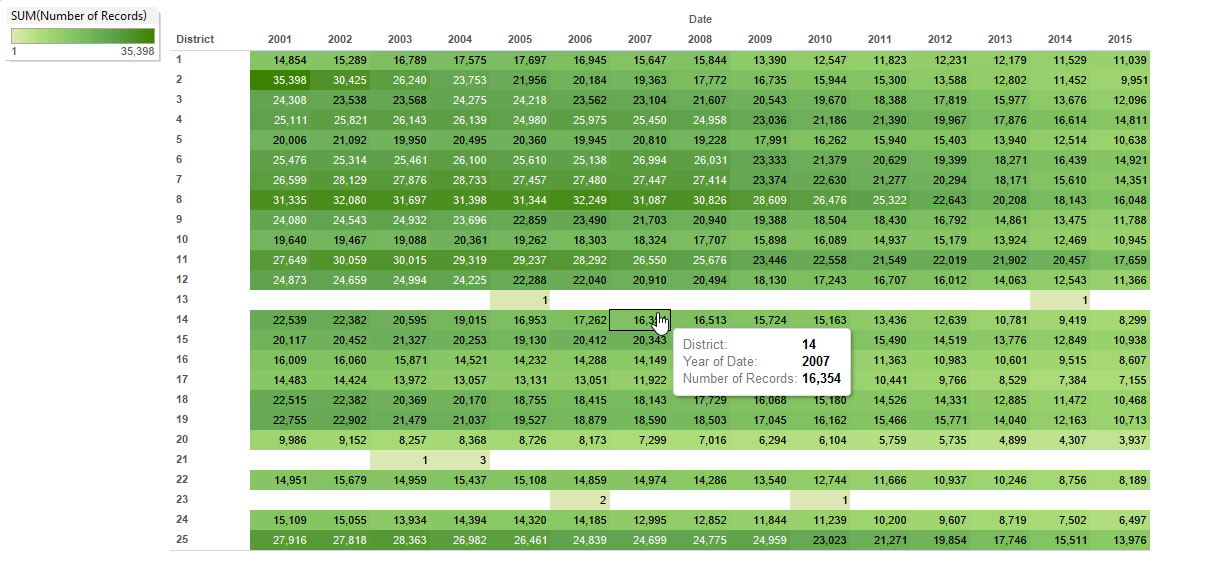
|  |  |
| --- | --- |
| **Violent** | **NonViolent** |
| Assault | Theft |
| Battery | Motor Vehicle Theft |
| Robbery | Criminal Damage |
| Arson | Burglary |
| Weapons Violation | Prostitution |
| Offense Involving Children | Narcotic Possession |
| Kidnapping | Deceptive Practice |
| Criminal Sexual Assault | Criminal Trespass |
| Concealed Carry | Interfering with Police |
| Homicide | Intimidation |
| Criminal Sexual Offense | Non-Criminal Offense |
| Human Trafficking | Liquor Violation |
| Domestic Violence | Stalking |
|  | Public Indecency |
|  | Gambling |
|  | Obscenity |
|  | Ritualism |

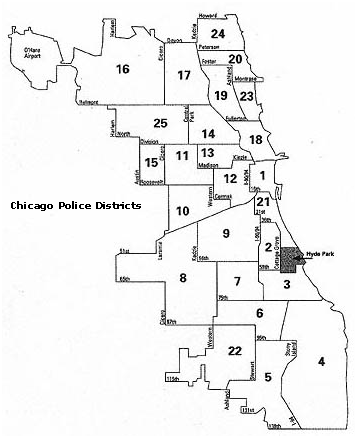




The previous two graphs show how difference in violent crimes vs nonviolent crimes throughout the years. While both categories of crime are decreasing, it seems that violent crime is decreasing at a much lower rate than nonviolent crime. Below is another chart that shows this trend. The blue area is the incidents of nonviolent crime throughout the year and red is the incidents of violent crime. This graph clearly shows how violent crime is decreasing at a much lower pace than nonviolent crime.







The previous graph and map shows how crime is distributed throughout Chicago based on districts. I was planning on plotting all the data on a map, but with around 6 million data entries it was not feasible. Even trying to plot data for a single year ended up being around 400,000 entries. So the only thing showing up on the map was one large blob, not separate points.

A possible improvement on this project would be if Tableau would be able to access Google Earth API directly to plot the points on that. Also if we could convert Latitude and Longitude to area zip codes, which could be done through Google’s Geolocation API. This is not cost effective at the moment, at least for students, since Google has a restriction of 1,000 API per day and to exceed it there was additional costs. Since this data set has around 6 million entry points, this would not be feasible at all.

The development environment was Windows 10 64 bit edition running Tableaeu for Students version 9.2 64-bit edition. TO turn the software you have to request for a student access through the Tableau website, use the 30 day trial, or just buy the program. Tableau at the moment is a tool that is only designed to run on Windows. The next step is to import the data set into the program. From there you can use a variety of tools offered on the program to run analytics on the dataset. The data set used for this project was “Crimes-2001 to Present” that is available on the City of Chicago data website. It is a very large file at around 1.4 GB and around 6 million data points.