**Aim of the project:**

My aim of this data science project was to compare the crimes in the city of Chicago (Illinois) from 2010 and 2015. In this 5 year period I wanted to see how the crime rate has changed over the years (which ones increased and decreased).

The reason behind this specific subject is personal. My family live in Chicago and it is well known that Chicago is dangerous city “murders in Chicago have already topped 700 this year” (Staff, 2016). This report would make a clear visualisation of either an incline or decline of specific crimes.



(The Editors of Encyclopædia Britannica, 2016)

**Information about Chicago’s crimes:**

"From the moment of its incorporation as a city in 1837, Chicago has been systematically seduced, looted, and pilloried by an aeonian horde of venal politicians, mercenary businessmen, and sadistic gangsters. Nothing has changed in more than 130 years” (Demaris, 3).

More than fifty years have passed since the “noise” of machine guns has died on the streets of Chicago, the city’s image of organized crimes. Al Capone, lived only twelve years in Chicago, but he has become the city’s most famous citizen.

"It would be foolish to expect such an environment to produce a moral and law-abiding youth, possessing the right theories of life and of success, when everywhere around him he sees official lawlessness and vice in the saddle; when he sees his hardworking father laboring for a few dollars a day and accumulating nothing, and the bootlegger and the gambler riding in limousines" (Landesco, 7).

**Acquiring the data:**

I found the data relatively easy: most of the cities in the US have public data websites where everyone can download available data. In my case I found the data at Chicago’s Open Data Portal and I used two databases (one from 2010 crimes and another from 2015 crimes).

Links are as follow:

<https://data.cityofchicago.org/Public-Safety/Crimes-2015/vwwp-7yr9> (2015).

<https://data.cityofchicago.org/Public-Safety/Crimes-2010/q4de-h6yq> (2010).

The first data set from 2010 contains 370,127 instances and the second data set from 2015 contains 262,901 instances.

Both of the data sets have same attributes but some of them are not needed for my purpose.

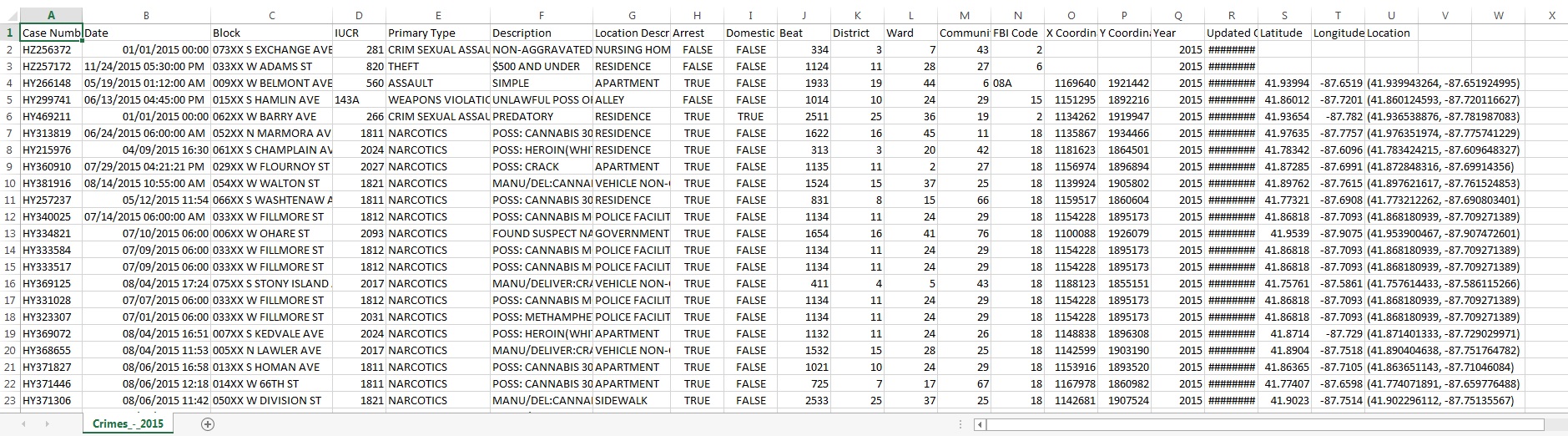
**Here is an image of the data presented in an excel table:**

2010 Crimes table (with all attributes available).



2015 Crimes Table (with all attributes available as well).

**2015 Crimes Table (with all attributes available as well).**



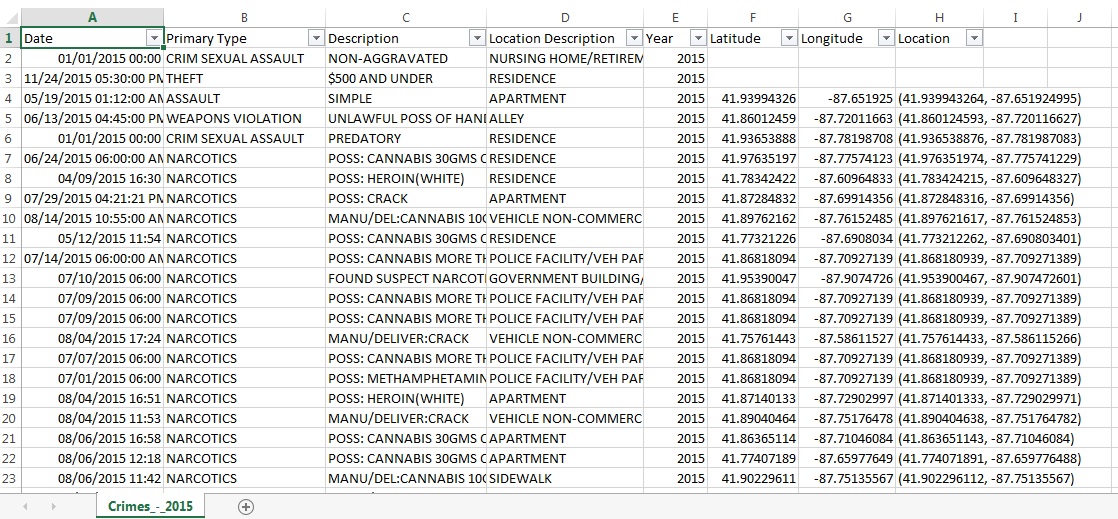
**After filtering the data and deleting the not needed attributes, my table has these attributes:**

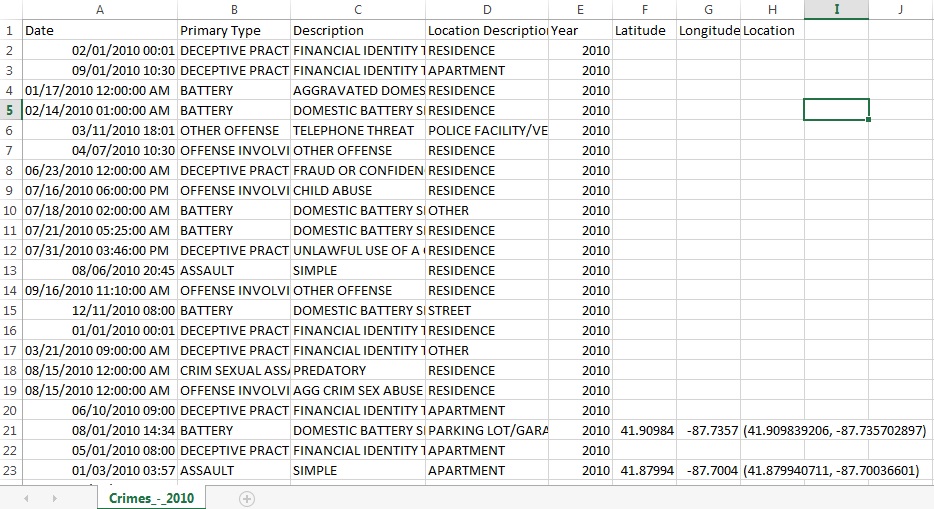
1. Date – Exact data and time when the crime took place
2. Primary Type

* Arson
* Assault
* Battery
* Burglary
* Concealed Carry License Violation
* Crim Sexual Assault
* Criminal Damage
* Criminal Trespass
* Deceptive Practice
* Gambling
* Homicide
* Human Trafficking
* Interference with public officer
* Intimidation
* Kidnapping
* Liquor Law Violation
* Motor vehicle theft
* Narcotics
* Non – criminal
* Obscenity
* Offense involving children
* Other narcotic violation
* Other offense
* Prostitution
* Public Indecency
* Public peace violation
* Robbery
* Sex offense
* Stalking
* Theft
* Weapons Violation

1. Description – short description of the crime
2. Location Description – where the crime was carried out (list is too long I am not going to list them all, but for example most of them are – apartment, residence, vehicle, sidewalk, etc.).
3. Year
4. Latitude
5. Longitude
6. Location – exact coordinates

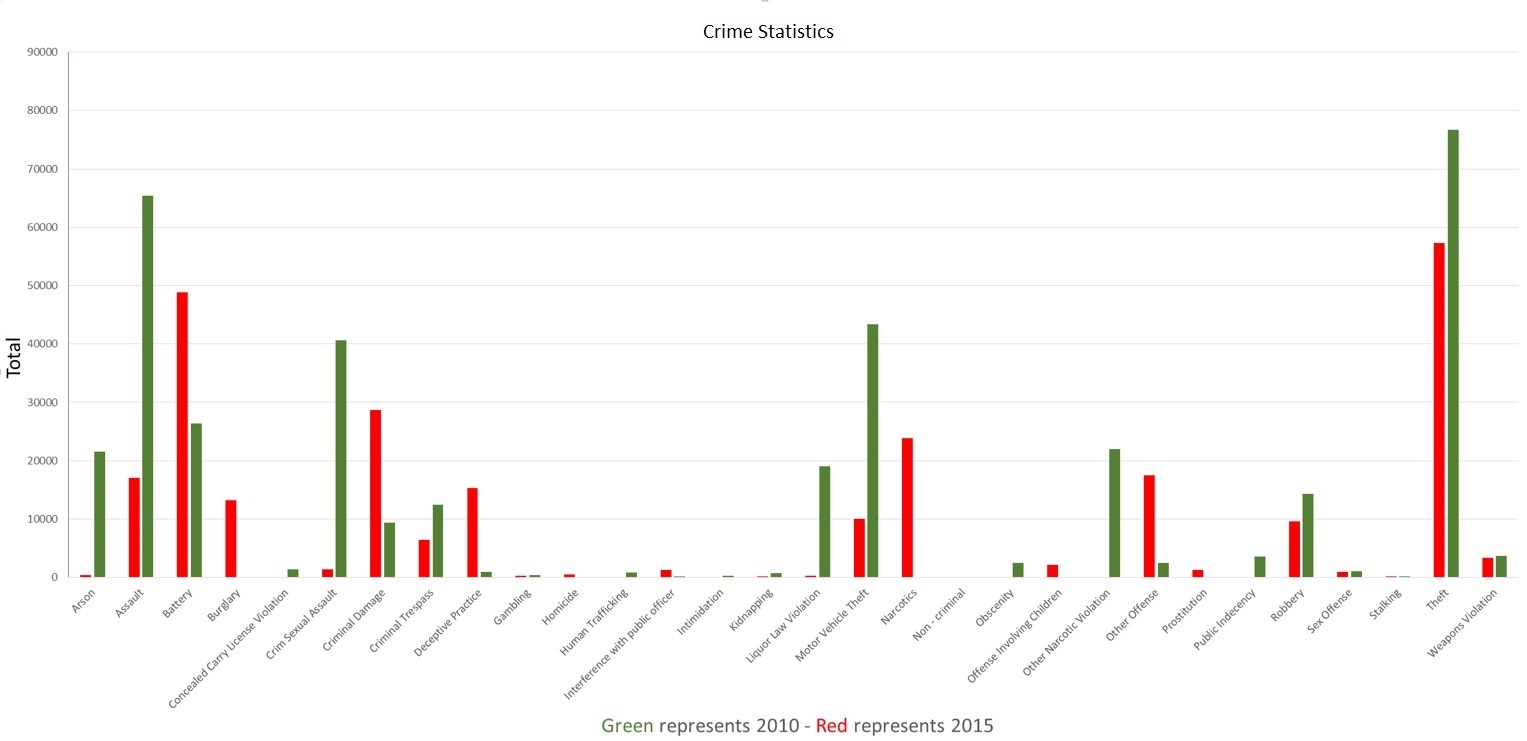
**Below you can see how it has changed:**

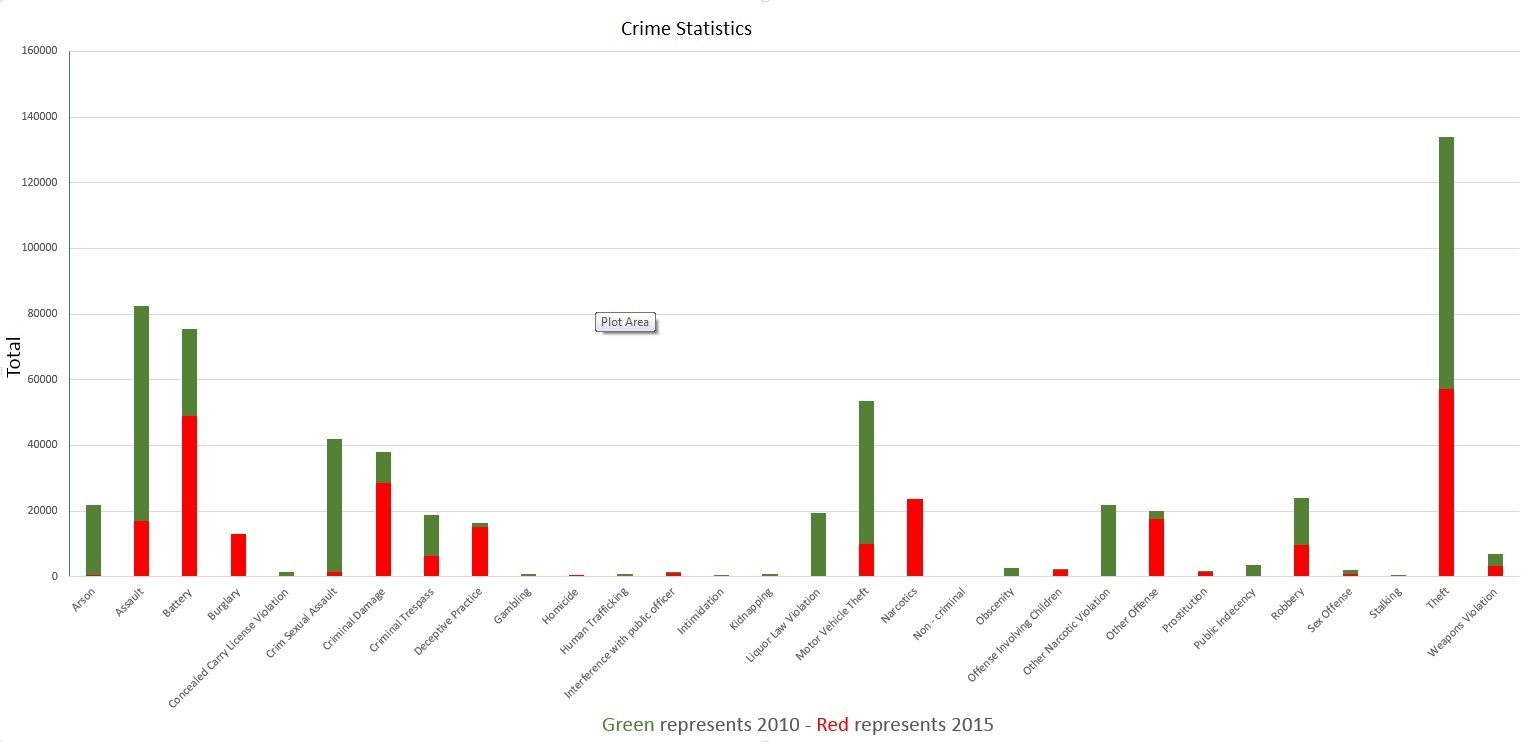




**Crimes Statistics 2015 – 2010:**

I started off with a graph statistics of all crimes. I used Excel for this:

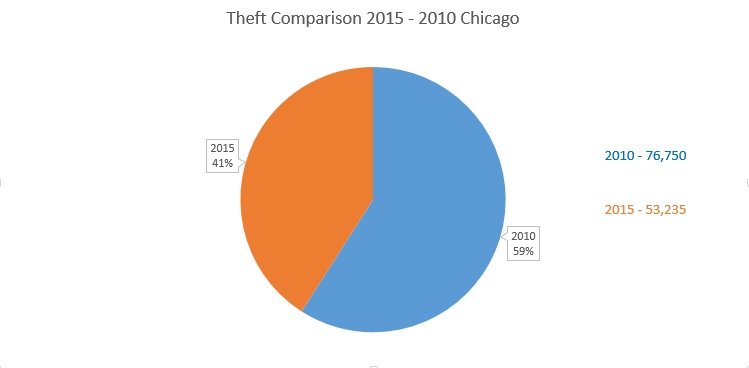




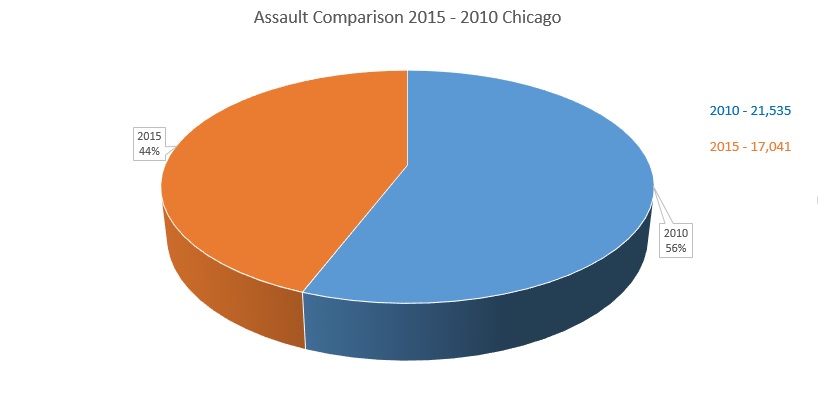
As you can see from the column charts above, green colour which represents 2010 is dominant in most of the crimes.

Now, let’s take a deeper view of the situation:

I chose the top three crimes (theft, assault and motor vehicle theft) and see how they changed over the years. I filtered the data by these crimes with a function from Excel: = COUNTIF(B1:B262901,”Theft”) – that’s just one example of how I took out the data for all of them and then created the pie charts below.

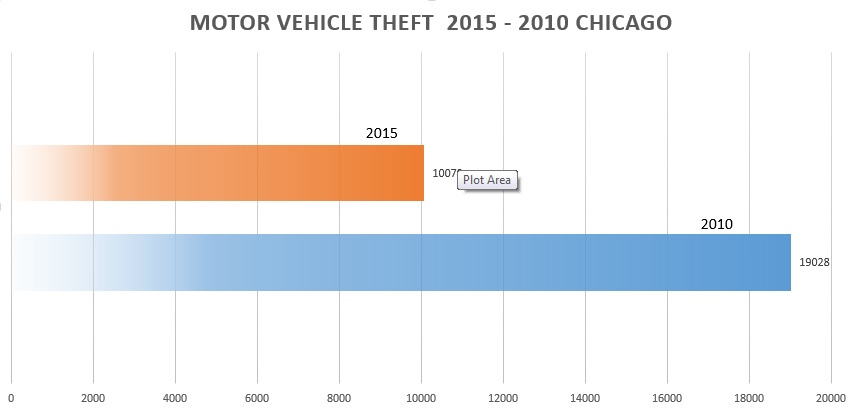


As you can see, in theft crimes there is significant decline with 18% from 2010 to 2015. In total, 2010 76,750 theft crimes and in 2015 53,235 crimes.



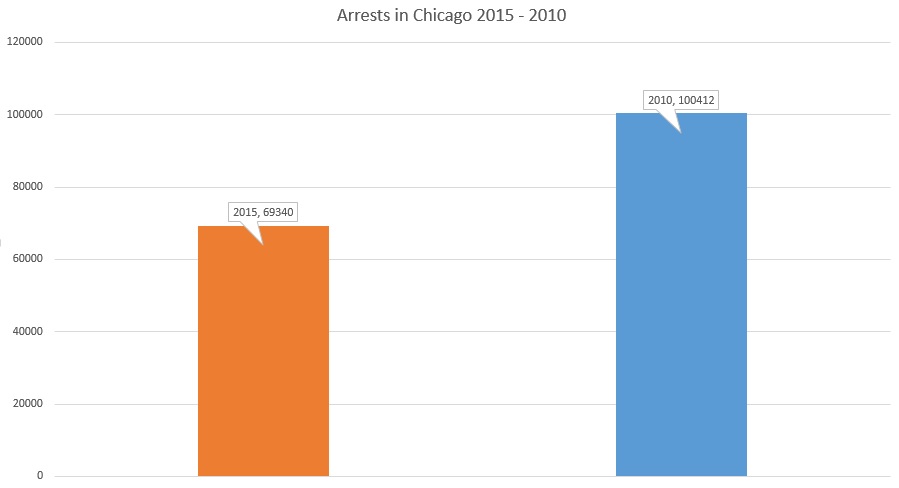
As for the assault crimes, there is also a decline with 12% for the period of five years.

In total, 2010 21,535 crimes and in 2015 17,041 crimes.

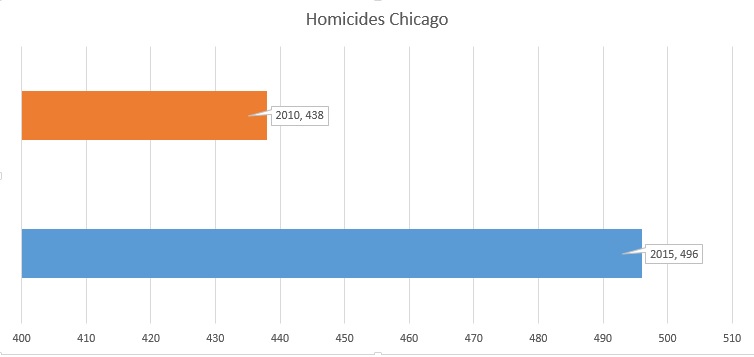


The biggest decline is on motor vehicle theft. In 2010 there were 19,028 crimes but in 2015 only 10,079. That is almost a 50% decline which is probably because of the fact that car security has improved!

I used the attribute “arrest” (I haven’t mentioned it above but I wanted to see the arrest rate so I added it back from the original datasets).

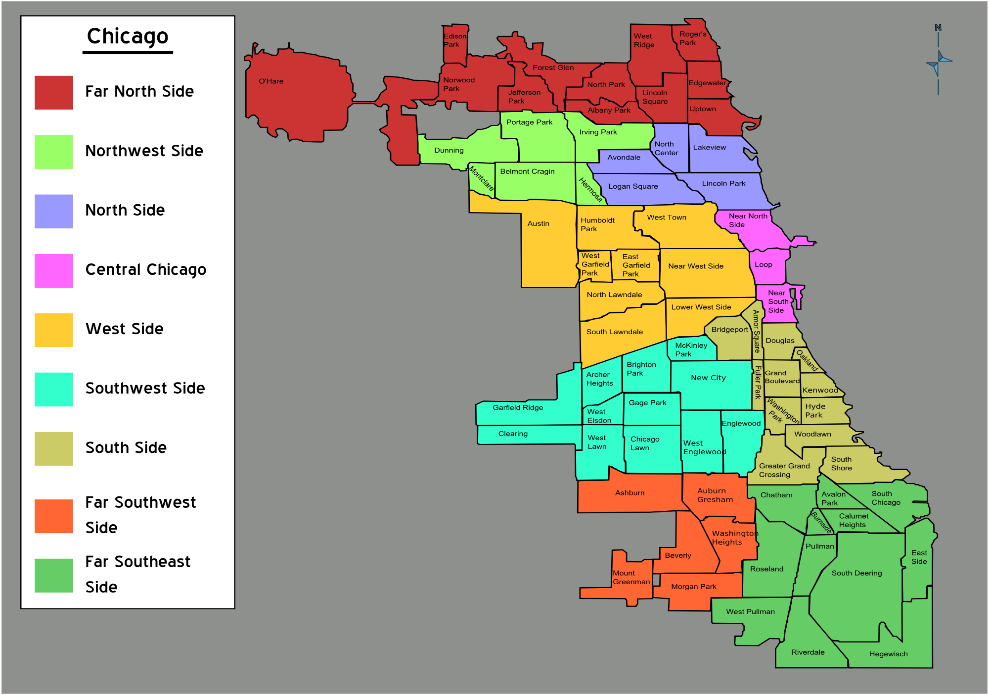


The graph comparison above shows the total number of arrests. But in reality there is not a big difference – in 2010 there were 370,127 crimes and only 100,412 ended up with an arrest (28% for that year). And in 2015 – total crimes were 262,901 and only 69,340 ended up with an arrest (27% for that year).

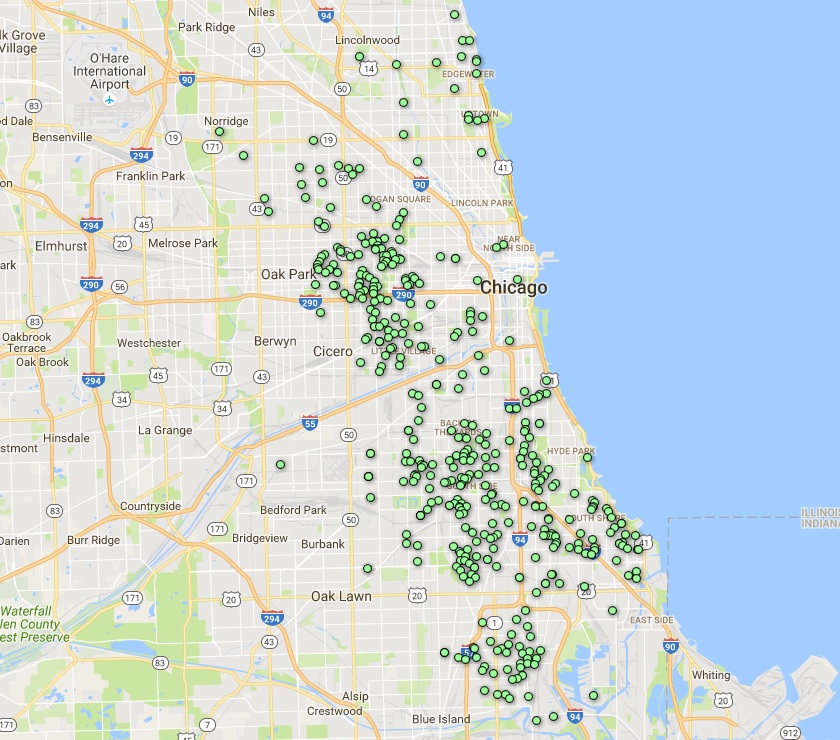
The last comparison (but not the least) is homicide. Chicago is the most dangerous place in the state of Illinois according to the authorities and the media. Some of the most dangerous neighbourhoods are West Englewood, Englewood, Riverdale “You know this neighborhood has a reputation when Google auto-suggests “West Englewood Crime” before you finish typing the name of the neighbourhood” (B and Raiford, 2016). 

The graph above shows that in 2010 there were 438 homicides while in 2015 - 496. Having in mind that the population has increased there is a slight difference in the rate. In five year period the rate hasn’t changed and this fact is scary.

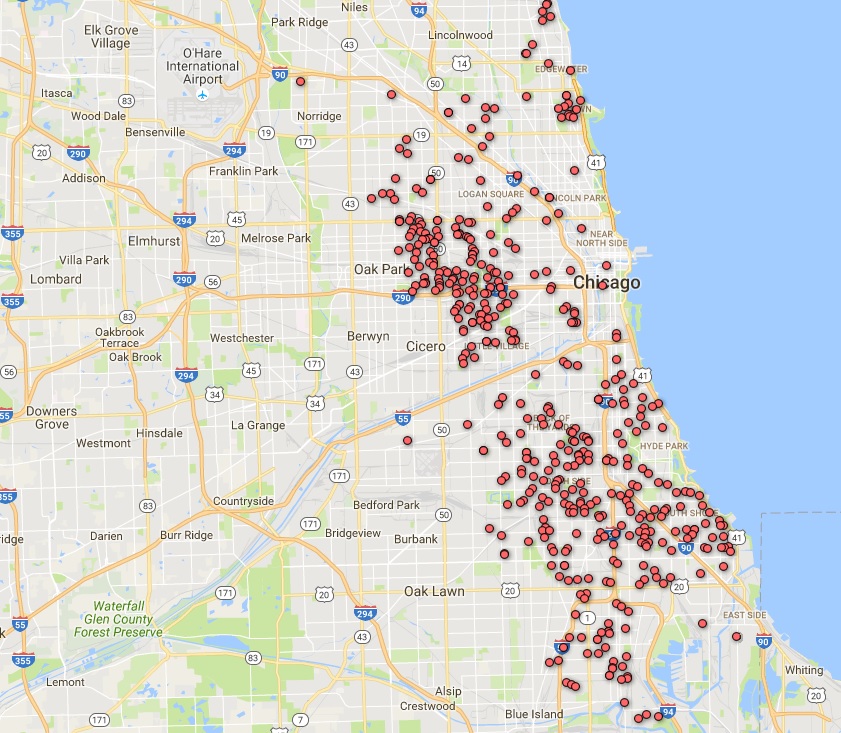
I decided to map these homicides on google fusion tables to find out where each incident was carried out and also to check whether the locations remain the same. The Google Fusion Tables is an amazing and super easy way to visualise these crimes. I used the COUNTIF function again from Excel to take out all the data with its coordinates and then mapped it. Below you can see the result: (I also included a small map for easy navigation).



(*File: Chicago neighborhoods map.png - Wikimedia commons*, 2007)



**Each green dot represents a homicide. This is for 2010 (438 homicides).**



**Each red dot represents a homicide. This is for 2015 (496 homicides).**

**Conclusion:**

From the maps above you can see that the homicide locations remains mostly the same. The South and Southwest Sides where Englewood and West Englewood are located remains the same, West Side as well for both of the years. There is a slight decline in the North Side but is still dangerous. Also, for 2016 up to December, the homicides reached 708 which is an increase with 43% compared to the year before “2016 DEATH TOLL: 708” (DNAinfoSearchSearchpresented, 2016).

Apart from homicides, most of the crimes have declined significantly (Theft – 18%, Assault – 12%, Motor Vehicle Theft – 49%) but unfortunately the arrests percentage remains the same. This leads me to believe that most of the criminals weren’t caught or were left off with just a warning. This fact is unrestful because there is a chance that all this criminals could repeat their crimes.

**Future work:**

This report could be used from authorities to implement safety measures in the high risk areas, such as: higher police presence, brighter street lighting etc. Also, it can be used to aware people of the zones where the most homicides took place or add time restrictions (warning people to avoid certain areas at certain time of the night). Signs could be made and put on certain locations to aware the citizens to walk/drive carefully and to look around for suspicious people/subjects.

**References:**

Staff, A. (2016) *Chicago hits grim milestone with highest murder count in nearly 20 years*. Available at: http://www.aol.com/article/news/2016/12/02/chicago-murders-top-700-for-first-time-in-nearly-two-decades/21619186/ (Accessed: 4 December 2016)

B, N. and Raiford, T. (2016) *Housely*. Available at: http://housely.com/the-10-most-dangerous-neighborhoods-in-chicago/ (Accessed: 8 December 2016).

*File: Chicago neighborhoods map.png - Wikimedia commons* (2007) Available at: https://commons.wikimedia.org/wiki/File:Chicago\_neighborhoods\_map.png (Accessed: 8 December 2016).

*Create: A map - fusion tables help* (2016) Available at: https://support.google.com/fusiontables/answer/2527132?hl=en (Accessed: 8 December 2016).

DNAinfoSearchSearchpresented (2016) *2016 chicago murders - Timeline*. Available at: https://www.dnainfo.com/chicago/2016-chicago-murders (Accessed: 8 December 2016).

*Genesis of organized crime in chicago* (no date) Available at: http://www.laborers.org/Genesis.html (Accessed: 8 December 2016).

(Landesco, 7) .,Hales, T. and Kazmers, N. (no date) *Bibliography*. Available at: http://www.umich.edu/~eng217/student\_projects/nkazmers/bibliography.html (Accessed: 8 December 2016).

The Editors of Encyclopædia Britannica (2016) ‘Flag of the United States of America’, in *Encyclopædia Britannica*. Available at: https://www.britannica.com/topic/flag-of-the-United-States-of-America (Accessed: 9 December 2016).