

**FINAL REPORT**

**Capstone Project - The Battle of Neighbourhoods**

**Introduction:**

New York City's demographics show that it is a large and ethnically diverse metropolis. It is the largest city in the United States with a long history of international immigration. New York City was home to nearly 8.5 million people in 2014, accounting for over 40% of the population of New York State and a slightly lower percentage of the New York metropolitan area, home to approximately 23.6 million. Over the last decade the city has been growing faster than the region. The New York region continues to be by far the leading metropolitan gateway for legal immigrants admitted into the United States.

New York City has also been a major point of entry for immigrants; the term "melting pot" was coined to describe densely populated immigrant neighbourhoods on the Lower East Side. As many as 800 languages are spoken in New York, making it the most linguistically diverse city in the world. English remains the most widely spoken language, although there are areas in the outer boroughs in which up to 25% of people speak English as an alternate language, and/or have limited or no English language fluency. English is least spoken in neighbourhoods such as Flushing, Sunset Park, and Corona.

With its diverse culture, comes diverse food items. There are many restaurants in New York City, each belonging to different categories like Chinese, Indian, and French etc.

**Problem:**

One of the main problems when ones move to a new city is where to find a good neighbourhood. This work will be focused on analysing and visualizing some community of Chicago. A good community will be ones with low rate of crime, healthy schools, and high number of public schools. Hopefully, the analysis can give valuable insights about the life of a good community in Chicago (SOUTH LAWNDALE), helping people to make decision when they want to move to Chicago. Popular venues in SOUTH LAWNDALE are also recommend to the users.

The main business accompanied in this work is what are the best areas in Chicago in term of safety and schools as well as what are different type of venues in those area that one can expect for.

## Data

Two datasets are used:

#### Chicago Public Schools

This dataset shows all school level performance data used to create CPS School Report Cards for the 2011-2012 school year. This dataset is provided by the city of Chicago's Data Portal. For this work I only use a snapshot of this dataset downloaded from: <https://ibm.box.com/shared/static/0g7kbanvn5l2gt2qu38ukooatnjqyuys.csv>

A detailed description of this dataset and the original dataset can be obtained from the Chicago Data Portal at: <https://data.cityofchicago.org/Education/Chicago-Public-Schools-Progress-Report-Cards-2011-/9xs2-f89t>

#### Chicago Crime Data

This dataset reflects reported incidents of crime (with the exception of murders where data exists for each victim) that occurred in the City of Chicago from 2001 to present, minus the most recent seven days.

For this assignment I will use a smaller sample of this dataset which can be downloaded from: <https://ibm.box.com/shared/static/svflyugsr9zbqy5bmowgswqemfpm1x7f.csv>

A detailed description of this dataset and the original dataset can be obtained from the Chicago Data Portal at: <https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-present/ijzp-q8t2>