# Coursera Capstone: The battle of Neighborhoods

**Introduction: Business Problem:**

This project identifies and recommends the nearest schools and school categories in neighborhoods. People especially families may move from one settlement to the other without considering the status of schools and their categories within the city and its neighborhoods.

The project targets families and students who will need to consider schools and their categories within the neighborhoods of a city before making their choices of settlement.

With regards to development and building of school structures in a city, the government will need to analyze schools and their categories within the neighborhoods to make the best choices for residents.

Private sectors, investors and business institutions are also beneficiaries. These entities can secure the right locations to invest by making the right decision on which categories of schools will be of a high demand in various neighborhoods.

**Data Requirement:**

Based on the problem definition above, factors to be considered are:

* Number of schools and their categories in a neighborhood
* Radius distance of neighborhoods nearest to schools

The following data sources will be required to extract relevant data:

* A list of neighborhoods of Manchester from Wikipedia
* A geocoder library for extracting latitude and longitude coordinates
* FourSquare API for searching for schools within these neighborhoods and providing their various latitudes and longitude coordinates

**Preprocessing the Data:**

The data had some irrelevant records in the dataset which I removed. For instance, I searched for schools using the FourSquare API and the results included school cafeteria and entertainment parks. I also eliminated duplicates which came up as a result of schools present in neighborhoods within a 1000 meter radius.

Below is an image of my pre processed data for analysis:

