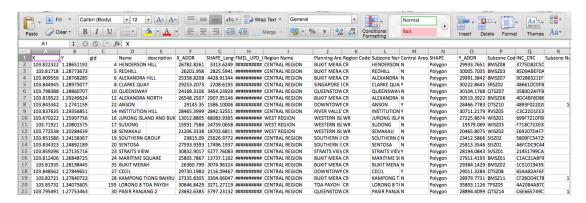Capstone Project – The Battle of Neighbourhoods

I.   Introduction / Business Problem

The aim of this project is to identify new market opportunities/ potential businesses in Singapore using Four Square location data. This project would utilize Four Square data to identify the popular venues in Singapore, what businesses are already available and how saturated the market is. The target audience would be the potential investors

II.   Data

I will be using Singapore boundary data to identify the regions and sub-zones in Singapore. The source of the data is from Data.gov.sg. I have downloaded the data and converted them to a CSV file. Below is the snapshot of the CSV files:

The latitude and longitudes are given by Y and X columns respectively. The sub-zones fall under 5 regions, namely: Central Region, West Region, North-East Region, East Region and North Region. There is also a field called Central Area Indicator, which indicates whether the sub-zone can be found in the central area. This would be useful if we are only interested in the neighbourhoods that belong to the central districts.

Given the location of each sub-zone, we can now do queries using Four Square API to get the businesses or venues for each sub-zone and determine which are the most popular type of establishments.