Coursera Capstone – Battle of the Neighbourhoods, Assignment Report

By Lyubomir Ahtapodov
August, 2019

1.   Introduction/Business Problem. Sports are becoming an increasingly popular part of people's everyday life both as a means of improved wellbeing and personal health. It is only natural to expect that this increased popularity of regular training will lead to increased demand for sporting goods. Yet in big cities with a hectic everyday life such as Chicago, it is of crucial importance for commercial venues such as sporting goods shops to be visible and conveniently located in order to maximise their customer base. The present project explores the problem of finding the best location for opening a sporting goods shop in Chicago, based on a data-scientific approach. The outcome of this project will potentially be of significant interest for sporting goods retail businesses looking to either enter the market or expand their activities in the city of Chicago.

2.   Data. The data used in this project will be obtained from two sources. Firstly, the precise geographical coordinates of multiple points along the boundary of each community area will be obtained from the City of Chicago's official website. With some processing, this data will be used to obtain the centre point of each community area, as well as determine the radius of a circular vicinity that best matches the borders of the community area. After that, venue information for the area around each community area centre point within a vicinity of radius as determined above will be requested from Foursquare. Following this step, a venue shortlist containing venue types that tend to occur together with sporting goods shops will be created, which will then be used to perform cluster analysis of all Chicago community areas.