**Introduction**

The retail industry is one of the business fields that can easily and effectively be evolved with the aid of data science and analytics. Information about customer needs, historical sales figures, competitor and market analysis and other relevant data can assist in the development of a certain business, by allowing investors to take optimal decisions. One of the specific retails sectors that is of great interest to the author is the sports and fitness field. There is an extremely wide range of sports and activities that are being exercised daily by numerous people.

**Business problem/ Shareholder Interest**

The current social distancing measures worldwide have maybe urged a wider audience to actively engage with sports exercised by less than two people at a time. At the same time, it might be expected that after the lockdown measures are lifted, there will be a massive interest in participating in group sports again, or simply participate in athletic activities using shared space and equipment, such as gyms, pools and basketball courts. For these reasons, investing in supplying sports equipment appears like a reasonable consideration. This report aims to develop a statistical methodology of deciding in which neighborhood of New York it would be best for a group of shareholders to invest in a brand-new sports equipment retail store.

**Data**

In order to analyze the various factors that affect the outcome of a hypothetical investment, Foursquare location data will be utilized in order to compare between neighborhoods. More specifically, for reasons which will become clear in the methodology section, the following attributes of a neighborhood will suffice for our analysis:

* Total number of sports venues: This is the sum of all enlisted sports and fitness facilities in which people can exercise. The Foursquare database included a variety of venues, such as basketball, baseball and badminton courts
* Neighborhood Median Salary: The median salary for each NYC neighborhood
* Total population for each neighborhood: The population of each neighborhood

The total number of sports facilities can be found by utilising the Foursquare API, while the total population of each neighborhood an be found on the NYC OpenData website: <https://data.cityofnewyork.us/City-Government/New-York-City-Population-By-Neighborhood-Tabulatio/swpk-hqdp>. The median salary data were retrieved from <https://ny.curbed.com/2017/8/4/16099252/new-york-neighborhood-affordability>.

Further data that were used throughout the analysis included location data for the sports venues, such as latitude and longitude, as well as venue ID. All the relevant information is available by Foursquare.