*Miguel Vidal*

*Soccer Store, NYC*

An Analysis of Where to Put a Soccer Store in New York.

**Introduction**

New York City is one of the most diverse cities in the world. Being a focal point of the USA for any individual in other countries since at least the 1880s, the "Big Apple" has been more than welcoming of other cultures [5]. With roughly 36.8% of the population born outside of the USA today, it is no coincidence that 800 different languages are spoken in the city [1][2]. The diversity in the city has had an influence on food, fashion, pop culture, and sport.

With soccer being the most popular sport in the world and the vast amount of cultures in NYC and the surrounding cities, it makes opening a new soccer shop in the city a realistic opportunity. The question is where exactly. Would it be better in Queens, where the most languages are spoken in the world, or in Manhattan where a professional soccer team plays, or somewhere else entirely? That is what this project will attempt to answer by observing where soccer is the most popular in the city and determine if there are any stores associated with soccer in the vicinity.

**Data**

Foursquare API will be utilized to establish where in the city are soccer fields and sports bars most dense, with the assumption that sports bars will cater to soccer fans. Sporting goods stores will then be located to discern what competition is around the high-traffic areas [4]. Latitude and longitude data of New York City and the surrounding cities will be used to visualize the findings [3].

**Methodology**

In order to discover which neighborhood to put a soccer shop, one must use the Foursquare API to first determine which neighborhood has the most soccer fields and/or sports bars. This is going under the assumption that the amount of soccer fields and sports bars means there is a big soccer following in that neighborhood. The venues found were put in a data frame that is found in Table 1.

Table

Description automatically generated

*Table 1 Data frame of soccer fields and sports bars around NYC.*

The venues were clustered based on latitude and longitude by k-means clustering to see which area of the city would be the ideal spot for a soccer shop. In order to determine the correct number of clusters, silhouette scores were calculated. Based on the score as shown in Figure 1, 4 clusters were used. Although 2 clusters had the best silhouette score, four clusters were chosen due to size of New York City and its surrounding areas. The clusters were then be compared to the number of sporting goods shops in the surrounding area.

Chart, line chart

Description automatically generated

*Figure 1 Silhouette Score of soccer fields and sports bars.*

**Results**

Once the number of clusters was decided, the k-means calculation separated the soccer venues. The results of k-means clustering showed that the densest area of the soccer venues was in Manhattan, as indicated in Figure 2.

Map

Description automatically generated

*Figure 2 Clusters of Soccer Venues*

Due to the size of Manhattan, and the number of venues around that area, k-means clustering was utilized again to narrow down the best spot to put a soccer store. Prior to calculating k-means a silhouette score was calculated using Manhattan cluster in the same way as described in the methodology section. This resulted in four clusters used in k-means clustering for Manhattan.

After the Manhattan cluster was separated into four sections, detailing the number of soccer venues south of Manhattan, the locations of sporting goods shops were found. By observing the visualization found in Figure 3, one can ascertain that a good spot to put a soccer shop would either be in Chelsea or East Village.

Map

Description automatically generated

*Figure 3 Best area to put a soccer shop, indicted in green. The red and*

*teal dots are soccer venues, the blue dots are sporting good shops.*

**Discussion**

The results found in this study had limitations. The max number of venues that would populate per venue category was 50 due to the restrictions of Foursquare [4]. Because of this, it cannot be certain that all the venues were represented in the study. Another limitation is how each venue was categorized. When creating a list of soccer venues, certain venues were filtered out due to their category name (i.e., Baseball Field), although soccer may be played at such venues. The same occurred when creating a sporting goods shop list. Due to leaving out certain venues, the clusters could be skewed and not represent a realistic picture.

To move this study forward, it would be wise to include more venue categories to ensure there is realistic representation. As well, to really pinpoint the best spot for a soccer store, one could analyze the popularity of each soccer venue to see what venue drives in large numbers of people.

**Conclusion**

By observing the soccer venues and the lack of sporting goods stores around, it would be good to open a soccer store in either Chelsea or East Village. However, this is just the beginning of the analysis process. The next steps would be to analyze the real estate market to see what is feasible in the above-mentioned neighborhoods. Next would be to determine the popularity of the soccer venues. One neighborhood may have more venues, but do not drive in a high crowd often. It may be better to go with the neighborhood with less venues but are far more popular.

**References**

<https://www.businessinsider.com/queens-languages-map-2017-2?r=UK>

<https://www.census.gov/quickfacts/newyorkcitynewyork>

<https://cocl.us/new_york_dataset>

https://developer.foursquare.com/developer/

<https://study.com/academy/lesson/history-role-of-immigration-migration-in-new-york.html#:~:text=The%201880s%20saw%20the%20beginning,war%2C%20among%20many%20other%20reasons>.