# SELECTING PARK BUILD SITE - SYRACUSE, NY

Donald Keeney

## 1. Introduction

### 1.1 Background and Problem

City parks play a vital role in the social, economic, and physical well-being of America's cities and their residents. Building a community park is a time intensive and costly process; it is therefore important to select the proper location to ensure its use and positive impact are maximized. The purpose of this analysis is to identify neighborhoods in the city of Syracuse, NY that are:

1. Currently underserved from a public park perspective, yet similar enough in nature to neighborhoods containing parks, in order to leverage past city design decisions.

2. Residentially dense enough to minimize travel time for users, and with a significant population of families with children.

### 1.2 Interest

The target audience receiving this analysis would be local politicians or community boards who could facilitate the funding and community support for the proposed build site

## 2. Data Acquisition and Cleaning

### 2.1 Data Sources

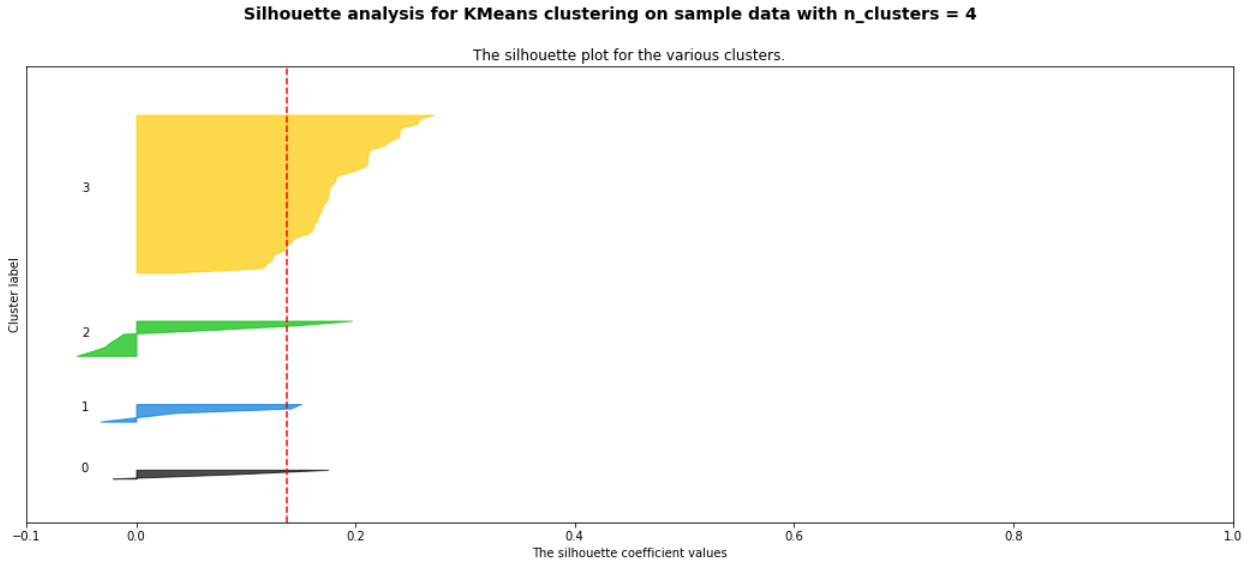
Syracuse neighborhoods will be defined according to census tracts obtained from a geojson file from the U.S. Census Bureau to include those tracts in the greater Syracuse Metropolitan Area within Onondaga County, N.Y. The center points of these neighborhoods will be used to obtain venue data via a Foursquare API call. Additionally, a separate U.S. Census Bureau file will be used to identify a demographic subset within these census tracts / neighborhoods; namely, children aged 0-14.

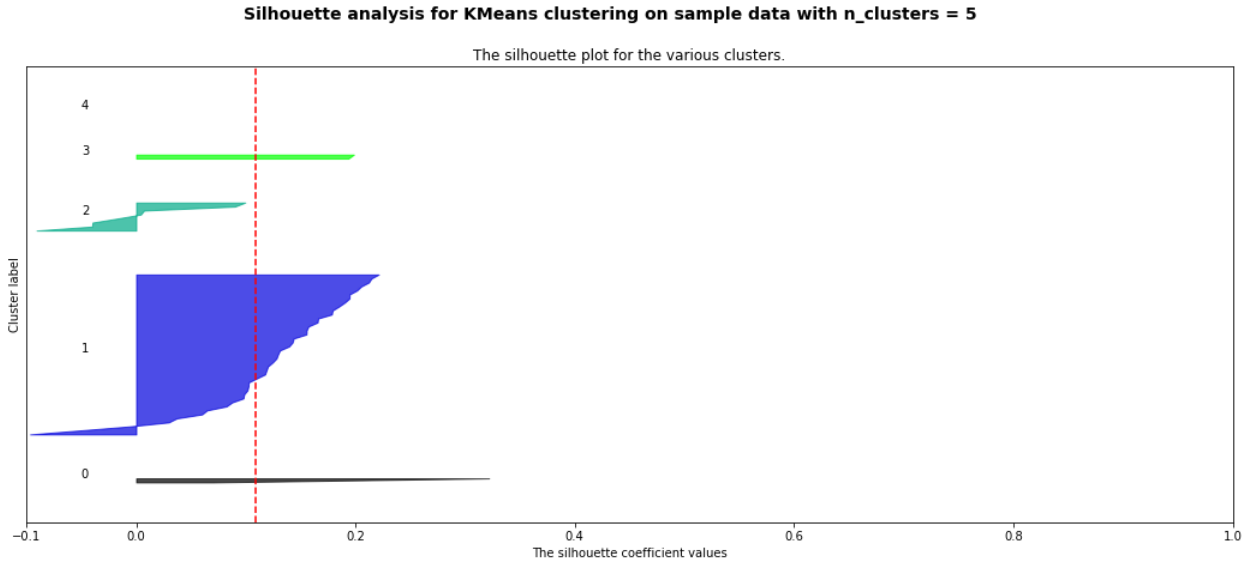
### 2.2 Data Cleaning

Despite being robust, data from two Census Bureau files obtained was not particularly well formatted or segmented. It was necessary to reduce the number of tracts included in this analysis which was done by parsing the original geojson file for selected census tracts comprising the Greater Syracuse Metropolitan Area. The demographic data was also cumbersome; it included a csv file with raw data, as well as a json file with readable column names corresponding to the ambiguous names in the csv file. The json file was iterated through to rename the columns in the csv file to be understandable in nature. In addition, given the relatively small population in the city, venue data was often not complete, with one neighborhood containing no Foursquare venues within the specified radius. This single neighborhood was ignored for the purposes of this analysis.

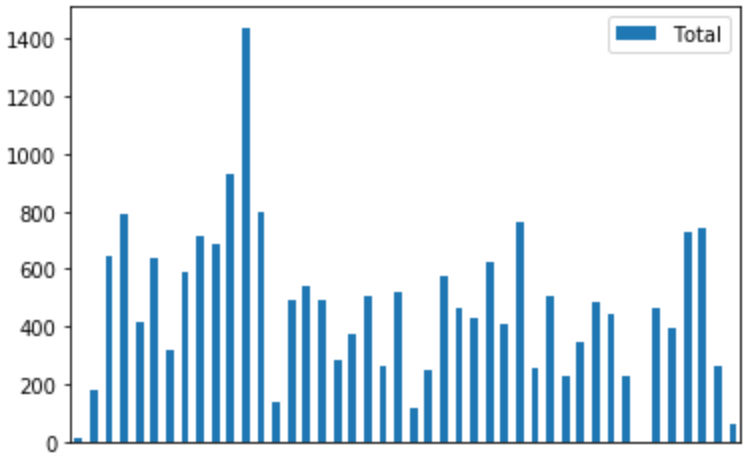
### 2.3 Feature and Model Selection

Venues for each neighborhood were queried within a 500 meter radius of the specified neighborhood centroid. Again, given the somewhat sparsely populated nature of the city, limiting the number of venues to 5 was ideal (up to 10 locations were queried, resulting in only 5 results, and sometimes fewer). A k-means clustering algorithm was then used to identify clusters of neighborhoods defined by the frequency of venue types. Silhouette analysis was performed to select an ideal number for k, which was set to 4.




Additionally, male and female children aged 0-14 were plotted for each neighborhood to aid with final selection of build site. The assumption was made that an adequate childhood aged population would maximize park use.




## 3. Results

Resulting clusters were analyzed to determine suitability for a park build site

Cluster 0, below, was almost defined by popularity of parks, so this cluster was excluded from consideration.

| | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue |
|---|---|---|---|---|---|
| **11** | Pool Hall | Park | Zoo | Discount Store | Electronics Store |
| **16** | Flower Shop | Convenience Store | Park | Eastern European Restaurant | Zoo |
| **34** | Garden | Gym / Fitness Center | Park | Theater | Zoo |
| **39** | Park | Gym / Fitness Center | Deli / Bodega | Scenic Lookout | Music Venue |
| **40** | Concert Hall | Park | Discount Store | Convenience Store | Dive Shop |
| **44** | Bakery | Park | Dive Bar | Exhibit | Electronics Store |

Cluster 1, below, was selected as the cluster to contain the proposed park site.

| | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue |
|---|---|---|---|---|---|
| **0** | Sandwich Place | Coffee Shop | Pizza Place | Sushi Restaurant | Bar |
| **1** | Sandwich Place | Coffee Shop | Pizza Place | Pharmacy | Donut Shop |
| **2** | Donut Shop | Motorcycle Shop | Grocery Store | Discount Store | Gay Bar |
| **3** | Convenience Store | Bakery | Liquor Store | Smoke Shop | Candy Store |
| **4** | Zoo | Bar | Convenience Store | Park | Other Nightlife |
| **5** | Herbs & Spices Store | Pizza Place | Convenience Store | Sports Bar | Donut Shop |
| **6** | Construction & Landscaping | Deli / Bodega | Business Service | Intersection | Dance Studio |
| **7** | Bar | Playground | Vietnamese Restaurant | Italian Restaurant | Zoo |
| **8** | Convenience Store | Vietnamese Restaurant | Drugstore | Supermarket | Zoo |
| **10** | Bar | Pool | Convenience Store | Deli / Bodega | Rock Climbing Spot |

It contained some park venues, as well as other outdoor attractions (zoos, rock climbing, and pools) and was determined to be a suitable neighborhood type to build a park. The neighborhoods

contained in this cluster were further refined to select those neighborhoods that had a relatively high youth population, as indicated by the "Total" column below:

| | Neighborhood | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue | Total |
|---|---|---|---|---|---|---|---|
| 0 | Census Tract 43.02 | Sandwich Place | Coffee Shop | Pizza Place | Sushi Restaurant | Bar | 12 |
| 1 | Census Tract 43.01 | Sandwich Place | Coffee Shop | Pizza Place | Pharmacy | Donut Shop | 181 |
| 2 | Census Tract 21.01 | Donut Shop | Motorcycle Shop | Grocery Store | Discount Store | Gay Bar | 647 |
| 3 | Census Tract 5.01 | Convenience Store | Bakery | Liquor Store | Smoke Shop | Candy Store | 791 |
| 4 | Census Tract 29.01 | Zoo | Bar | Convenience Store | Park | Other Nightlife | 418 |
| 5 | Census Tract 2 | Herbs & Spices Store | Pizza Place | Convenience Store | Sports Bar | Donut Shop | 635 |
| 6 | Census Tract 3 | Construction & Landscaping | Deli / Bodega | Business Service | Intersection | Dance Studio | 319 |

## 4. Discussion

Following the cluster analysis, emphasizing neighborhoods that were similar to neighborhoods with parks, but not defined by parks, as well as population considerations, the following census tracts were selected: tract 14, tract 10, and tract 61.01

## 5. Conclusion

Given the time and financial cost involved in building a park, it is necessary to select the most ideal location possible. Through use of machine learning clustering and population analysis, this report has summarized 3 ideal, safe, and practical locations to develop the next public park in the city of Syracuse, NY.