

The Battle of Neighborhoods: NYC Families

Eric Jones

June 19, 2020

1. Introduction

1.1 Background

New York City, NY is the most populated city in the United States of America. As an international hub for travel, food, culture, art, business, and much more, a large number of people work and live in the area. As an area where so many people's lives happen, a growing amount of people are looking to bring families to live within the city's limits. However, with New York City being a largely urban environment, parents must determine neighborhoods in the area that cater do the needs or raising children in a city environment. Specifically, neighborhoods with nearby parks, civic centers, childcare facilities, and schools would be much more desirable for people living in the city with children. The identification of these neighborhoods would provide families the confidence to successfully raise children in the large-city environment of New York.

1.2 Problem

Living in New York City while raising kids requires finding an environment that is suitable for children and the responsibilities of the parents. A way to assess the suitability of a neighborhood in New York is to determine the number of parks, civic centers, childcare facilities, schools, and other similar venues that exist in the area. The goal of this project is to leverage the Foursquare location data to determine the best neighborhoods for families with children in New York City.

1.3 Audience

The target audience for this project is families moving to New York or living in New York already. The analysis performed in this project determines the best neighborhoods for these families to live with their children. Another stakeholder that could make use of this analysis is retail companies. Since these companies assist families finding a place to live within the city, the list of family-friendly neighborhoods determined by this project can help real estate agents determine the best locations for families.

2. Data

2.1 Data Sources

In order to determine the solution to the problem stated above, neighborhood and location data must be gathered to determine the venues that exist within a neighborhood. To do this, the Foursquare API is used to gather venue information based on latitudes and longitudes.

Data regarding the neighborhoods, including the latitude and longitude of each neighborhood, is obtained from an online data source: https://cocl.us/new_york_dataset

2.2 Data Usage

The dataset containing the New York neighborhoods will be used to generate the list of neighborhoods with their respective latitude and longitude. This data will be the information that is passed to the Foursquare API.

The Foursquare API will be used to generate venue category data in the neighborhoods in New York City. Using the explore function within the Foursquare API will enable the collection of all venue types within a certain radius of each neighborhood, based on the neighborhood's latitude and longitude. Once this data is collected, the results will be filtered to include only relevant venue types, namely childcare locations, parks, civic centers, and others. The frequency of each of the venue types will be determined for each neighborhood, and the neighborhoods will be clustered based on these frequencies. Once clustered, the set containing the highest frequency of these particular venues can be selected, thus determining the list of best neighborhoods for families in New York City.