

The Battle of Neighborhoods

Minshen Li

September 23, 2019

1. Introduction

When people have decided to travel to a specific city, for jobs or tours, they want to decide on which neighborhood to live. For people who do not want surprises, they are interested in finding the similar neighborhoods to their current neighborhood. For people who want a change, they want to know which neighborhoods are different from their current living environment and what is surprising about those new neighborhoods. Even from the time when they start searching for jobs, they may like to know what cities they desire so they could plan in accordingly. In this case study, I analyze the similarity among the different neighborhoods of the three cities, Philadelphia (Center City), New York City, and Toronto, and further analyze which pairs of cities share higher similarity than other pairs.

2. Data

I define *similarity between two neighborhoods* as *the similarity in the distribution of venue categories between the two neighborhoods*. I use Foursquare location data as it has category information for each venue. Foursquare database provides records on the categories of the Top 100 venues within the radius of 1,000 meters of the CenterPoint of each neighborhood. Using the data, I construct the frequency of occurrence of each venue category for each neighborhood (i.e., the distribution of venue categories). For each city, I measure similarity based on the constructed venue category distributions.