

Data Science Capstone Project

Toronto Public Libraries - Ensuring Their Success by Establishing Nearby Businesses That Influence People to Visit Neighborhood Libraries

By: Alexander Martinez

1. Introduction

1.1 Background

Based on several studies, the presence of "3rd places" strengthens communities. The home (first place), workspace / school (second place), and the library (or church or coffee shops) so called 3rd places help provide socially accepting and safe environments. The library can be considered as a place where people meet and social relationships made. Neighborhoods that have "famous" 3rd places like libraries tend to add economic and safety values to their neighborhoods.

At the end of this project, we would know if the libraries have increasing visits and if not, encourage the local government and private establishments that contributed to visits of successful libraries to set up shops near the libraries (or inside if local ordinance permit) to promote people traffic into the libraries.

1.2 Problem

This project will determine popularity of Public Libraries based on foot traffic (using Foursquare Check-ins). The stakeholders in this case are people who are living in or around these neighborhoods. And as such, this is not really a business problem but a social issue. It is imperative to ensure that people continue to visit libraries to help strengthen their communities. To further assist in increasing foot traffic to the libraries, we will determine top venues that would possibly contribute nearby to the success of these libraries.

1.3 Interest

As stated earlier, a library success is also a community's success. People who visit libraries come from a wide spectrum of races, financial backgrounds, levels of education and varied

interests. Having diverse people come to a “meeting place” can generate a sense of economic stability and social well-being for its residents.

2. Data acquisition

2.1 Data sources

We will determine which libraries are having considerable foot traffic and those that have low visits or decreasing traffic. We will limit our studies to libraries located within 10 kilometer radius from the geo center of Toronto. We will also get the top 5 venues that are close to the libraries.

Foursquare will be providing the data which would be the library locations, top venues, and check-in statistics. .

2.2 Data processing

Data from Foursquare will be processed for both the libraries and the top venues and determine which libraries are getting consistent and good foot traffic and those that have decreasing and will probably need to be revitalized. We will recommend / suggest establishments of venues that would influence people to "check-in" to its neighbor library.

A list of libraries and venues would have to be retrieved. Together with these, the latitude and longitude information. We would explore the top venues based on these geographic coordinates.

2.3 Feature selection

Initial dataset would contain the geographical coordinates of the libraries. These can be dropped once the top venues have been identified.

Final data would include Library names, addresses, check-in stats, venue names, check-in stats, and venue categories. Check-in stats for 5 years is being considered for this study.