

New Service from XYZLifestyle – Suggested Neighborhoods for New Residents

Kevin Spradlin

August 14, 2021

Introduction

- XYZLifestyle.zib creates health plans (diet and exercise).
- Increasing numbers of customers have been moving and asking the company for help. They need to find new stores and venues to help them stay on their plans.
- Idea: the company can identify health-related businesses and venues as new service. When customers ask questions, the company will already have the information to answer them.
- Create prototype service focusing on Toronto area.

Data

- Three sources
 - Names of neighborhoods from Wikipedia
 - Locations (latitude and longitude) from Python geocoder API
 - Business and venues from Foursquare
- Combine Wikipedia and geocoder API data → 103 neighborhoods of Toronto

Data

- Combine Wikipedia/geocoder API data and Foursquare data
→ 100 neighborhoods with venue data
- Manually identify Foursquare venue categories for health-related venues.
 - 68 neighborhoods with health-related venues.

Athletics & Sports	Baseball Field
Basketball Court	Climbing Gym
Fruit & Vegetable Store	Gym
Gym / Fitness Center	Health Food Store
Juice Bar	Park
Pool	Salad Place
Soccer Field	Tennis Court
Trail	Vegetarian / Vegan Restaurant
Yoga Studio	

Data Processing Methodology

- Use k-means cluster modeling to consolidate data.
 - Instead of grouping neighborhoods by physical location, can group them by common types of health-related venues in them
- Why use cluster modeling?
 - Are hundreds of health-related venues spread over 68 neighborhoods
 - Need to consolidate the data in order to use it

Results

- Found eight clusters of neighborhoods, grouped by most common types of health-related venues in each.

Most Common Venues	Number of Neighborhoods
Juice Bars	3
Trails	5
Athletic & Sport Venues	5
Parks, Yoga Studios	12
Pools, Baseball Fields	4
Gyms	6
Parks	15
No Common Venue (Residual Cluster)	18

Results

- If customers ask about specific neighborhoods in Toronto, company can tell them
 - If there are any health-related venues in those neighborhoods
 - The most common health-related venues in the neighborhoods
- Will need to regularly update the health-related venue, due to new businesses starting or closing.

Potential Improvements

- Instead of cluster modeling, put information into a database.
 - Company could provide customers with a list of health-related venues in a neighborhood
 - Company could provide customers with a list of neighborhoods with health-related venues
- Try other clustering algorithms – hierarchical, density-based
- Test process with other cities