Business opportunities in Seattle, WA

By Suresh Setty, April 18, 2020.

Introduction

Background

Seattle is a seaport city on the West Coast of the United states. According to census data release in 2018, the Seattle metropolitan area population stands at 3.98 million, and ranks as the 15th largest in the US. The Seattle area developed into a technology center from the 1980s onwards with companies like Microsoft becoming established in the region; internet retailer Amazon was also founded in Seattle, and a major airline Alaska Airlines is based in Seattle as well. Additionally, a lot of new software and biotechnology companies have been moving into the area, contributing to the area's vast economic and population growth.

Business Environment and Opportunities

The objective of this study is to identify various Seattle neighborhoods with good potential for new business opportunities. This is done by investigating the popular venues in each neighborhood to gain insights into the current business environment there. These insights in conjunction with house prices in the neighborhood could throw light at potential business opportunities.

Data

Acquisition and Methodology

We will be using the following sources for acquiring data needed for this study.

- There are a lot of sources on the internet identifying and classifying various Seattle neighborhoods. After some research, the data from Zillow.com seemed to be more comprehensive. So, the data from this Zillow page was used - https://www.zillow.com/east-queen-anne-seattle-wa/home-values/ .

- Once we have the list of neighborhoods, we use geopy and Nominatim Geocoding service to obtain latitude and longitude values of these neighborhoods.

- We use developer access to Foursqaure (https://foursquare.com) to explore the venues in all the neighborhoods. All venues in Foursquare are categorized into 10 main categories - Arts & Entertainment, College & University, Event, Food, Nightlife Spot,

Outdoors & Recreation, Professional & Other Places, Residence, Shop & Service, Travel & Transport. For each neighborhood, we will get the number of venues in all categories.

- Once we have the venue data for all neighborhoods, we can use k-means clustering algorithm to segment neighborhoods. We can then analyze the segments for any patterns and look for potential business opportunities.

Data Wrangling

- The neighborhood data from Zillow.com has a lot more information that what's needed. So, only the relevant information was kept.
- It was realized that some ZRI (rental index) values were missing from the Zillow.com data, so those missing values were replaced with the mean of all other ZRI values.
- Nominatim Geocoding service couldn't provide latitude and longitude values for a couple of neighborhoods, so those were entered manually after a google search.
- After obtaining latitude and longitude values for the remaining neighborhoods (using Nominatim), it was realized that there were some duplicates in the coordinate values. In Zillow data, the neighborhoods with the duplicate coordinate values had considerable differences in their ZHVI (home value index) values, so instead of deleting or merging these neighborhoods, choice was made to replace the duplicate coordinate values manually with data from google search just for these neighborhoods.