Opening a Coffee Shop

Jamie Fraser

July 27, 2021

Abstract

This document contains my final report for the Applied Data Science Capstone project, in fulfilment of the requirements for the IBM Data Science Professional Certificate on Coursera. I completed the course during June and July of 2021. The purpose of the project is to describe and analyze a business problem using Foursquare geospatial data in addition to inferential statistics and machine learning techniques that were learned during the course. I am writing from the point of view as a data scientist who was hired as a consultant to help determine the best area to open a new coffee shop in my hometown of Honolulu, Hawaii.

1 Background

The clients have hired me to assess the best location in Honolulu, Hawaii to open a new coffee shop. They are entrepreneurs who want to open a full service cafe and take away coffee shop that caters to visitors and local people. Their aim to create a warm and inviting place where customers can grab a coffee, sandwiches and other small food items, and even bring their computers and get some work done.

Coffee shops are popular in Honolulu. Most neighborhoods have at least one coffee shop already, and many neighboorhoods have multiple coffee shops. Due to the competition, it is important to select the right neighborhood to open this new coffee shop. They want to select an area that has enough traffic to support the coffee shop and produce a good profit, but not be in an area where there are already too many coffee shops and thus the market risks becoming oversaturdated.

Finally, they want to distinguish their coffee shop to make sure that it is a comforable and unique space for customers to enjoy their coffee and work or study. They further want to develop a menu that is unique from other coffee shops in the area and has healthy options for their customers to enjoy.

Given these business requirements, the entrepreneurs have asked me to develop a document that can guide them in making these decisions. I hope to provide enough information to guide them in data-based decision making.

2 Data

The data used for this project will be obtained using the Foursquare API. Specifically, the report requires data on:

1. Number and location of coffee shops in Honolulu, Hawaii

2. Neighborhoods with the highest number of reviews for coffee shops

3. Details on existing coffee shops

The number and location of coffee shops by neighborhood will be used to evaluate which areas are most popular for coffee shops. This can help to chose a location for the shop, as well as identify neighborhoods which may be oversaturated or have too many competing shops. Data on the neighborhoods with the highest number of reviews for coffee shops will help to identify nighborhoods with the highest traffic, and therefore the most customers, for coffee shops. Finally, it is important that the coffee shop stands out and is unique among competitors in the area. Thus, Foursquare data on the details on existing coffee shops, such as photos and reviews, will provide information on competitors and help to design the menu and look of the coffee shop.

Foursquare data is supplemented by data on zip code coordinates in Hawaii, from the open source data project Open Data Delaware on GitHub (https://github.com/OpenDataDE).

3 Methodology

4 Results

5 Discussion

6 Conclusion