Analyzing Optimum Location for Starbucks

Feroz Khan

August, 2019

1. Introduction

1.1. Background

Starbucks, a coffee company, is seeking to open another shop in San Francisco. Which will help the company increase its local impression and gain in revenue.

1.2. Problem

Despite having adequate budget for the new shop, the company does not know what location would be optimal. The team is struggling to find out the optimal location for a new shop in San Francisco. According to the company the optimal location for the new shop must have comparable or similar qualities as one of their existing and most visited outlets surroundings.

1.3. Interest

The explanation will serve the business and support their unit to pick one of the optimal locations for the new shop.

2. Data Acquisition and Cleaning

2.1. Data Sources

Foursquare.com, a data provider via API calls. The data is organized and in JSON format. Foursquare.com collected these data from the internet by their user activity i.e. location check-in, comments, tips, etc.

2.2. Data Processing and Cleaning

Even though data collected from the Foursquare API is in the JSON format, I did not use all of it. The data required simple cleaning to extract the essence from it as it was already in JSON format.

2.2.1. Steps I performed to collect and to handle the data:

2.2.1.1. Crate an account on Foursquare.com to get API credentials.
2.2.1.2. I used a Python library for Foursquare as a wrapper for API endpoints to collect data.

2.2.2. Below are the API endpoints I used:

- Venue method: to collect Starbucks' location.
- Venue Explore method: to explore venues around Starbucks location.
- Venue Search method: to search a venue and to locate its geographic points i.e. longitude and latitude.

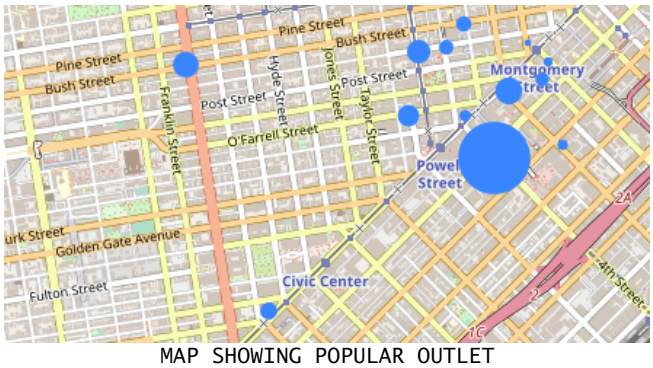
3. Methodology

To find the solution for the problem, I explored the Starbucks venue data provided by the Foursquare. During exploration, I find out the most popular Starbucks outlet in San Francisco. To find out the outlet with the highest traffic, I explore the public visit data and tips provided by them for each shop.

category	category_id	lat	lng
Gardens	4bf58dd8d48988d15a941735	37.784809	-122.402549
Gym	4bf58dd8d48988d16c941735	37.784816	-122.403537
Wine Shop	4bf58dd8d48988d1dc931735	37.784253	-122.402179
Boutique	4bf58dd8d48988d18f941735	37.785484	-122.402451

SAMPLE VENUE DATA

To find out the reason for the most popular outlets, I looked into their nearby venues. Interestingly, the top few popular outlets have these venues in common: *Gyms or Fitness Center, Boutique, Wine shops and Hotels*.

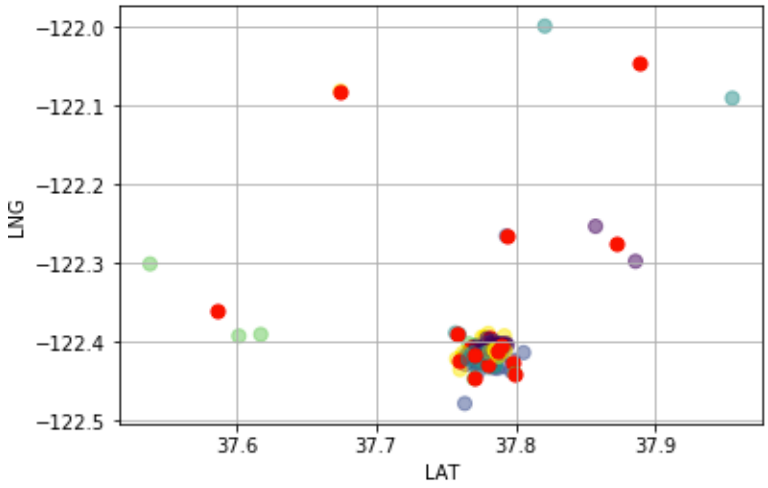


MAP SHOWING POPULAR OUTLET

Based on above finding, I can say that a Starbucks shop along with *Gyms or Fitness Center, Boutique, Wine shops and Hotels* might be a better location for a new shop.

To find out the geographic points of a location, which has popular venues, makes a neighborhood popular, I used K-Means machine learning

algorithm. I used the K-Means algorithm because it helped me create a cluster of venues, specifically clusters of those four venues.



VENUE CLUSTERED DATA

4. **Conclusion**

During data exploration, I discovered that in San Francisco, a popular venue is around Gyms or Fitness Center, Boutique, Wine shops and Hotels within 250 meters of its radius. So, as an optimum location for a new Starbucks shop, I would recommend the space with similar characteristics.