

Applied Data Science Capstone

Final Report

OPENING A COFFEE SHOP IN HO CHI MINH CITY (VIETNAM)

Name: VO VAN VIET

Overview

1/ Introduction

2/ Business Problem

3/ Data

4/ Methodology

5/ Results

6/ Conclusion



INTRODUCTION

HO CHI MINH CITY:

- The largest city of Vietnam
- Located in southeastern Vietnam
- Economic and financial center of Vietnam
- Plays an important role in the country's culture and scientific developments



Business Problem

A Local Coffee Shop wants to open their first store in Ho Chi Minh City. The new store would be located in an ideal place with crowded customers, acceptable housing prices, and low density of other coffee shops.



DATA

To find the ideal place for the new coffee shop, we need the following data:

- List of Districts in Ho Chi Minh City contains district names, and their population, area, density from wikipedia (https://en.wikipedia.org/wiki/Ho_Chi_Minh_City)
- List of Housing Price in Ho Chi Minh City from Mogi (<https://mogi.vn/gia-nha-dat>)
- Foursquare API to get the venues.

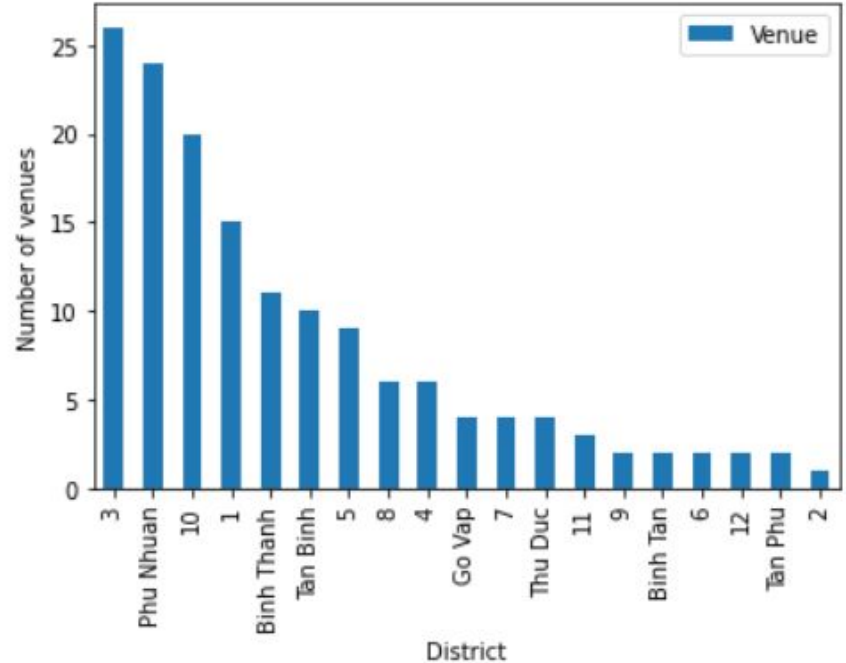
METHODOLOGY

- First, by using the library BeautifulSoup to scrape the data from wikipedia, the list of districts had been collected.
- Then, the density had been calculated by the subtraction of population and area.
- geopy.geocoders.Nominatim had been used to get the coordinates of districts and add them to the main data frame.
- The library Folium had been used to sketch the map of the districts.
- Foursquare API to explore the venues in each district and segment the districts based on them.
- Clustering the venues "Coffee Shop" by the algorithm K-Mean Clustering from the library Scikit-Learn.

RESULT

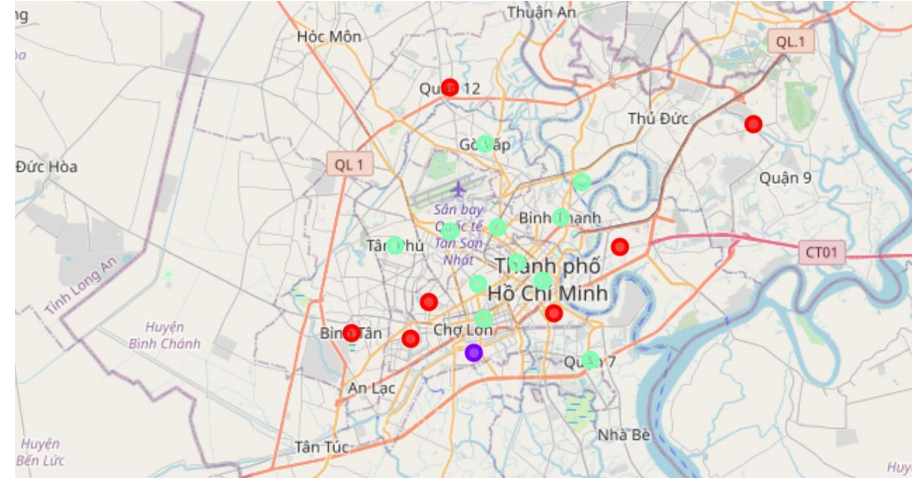
District 3, 10, 1, Phu Nhuan had the biggest number of coffee shops.

District 2, 12, 6, 9, Tan Phu, Binh Tan had the lowest number of shops.



RESULT

- Cluster 0: There are not many coffee shops in these districts. (red)
- Cluster 1: The number of coffee shops in these districts is high. (blue)
- Cluster 2: The number of coffee shops in these districts is medium. (green)



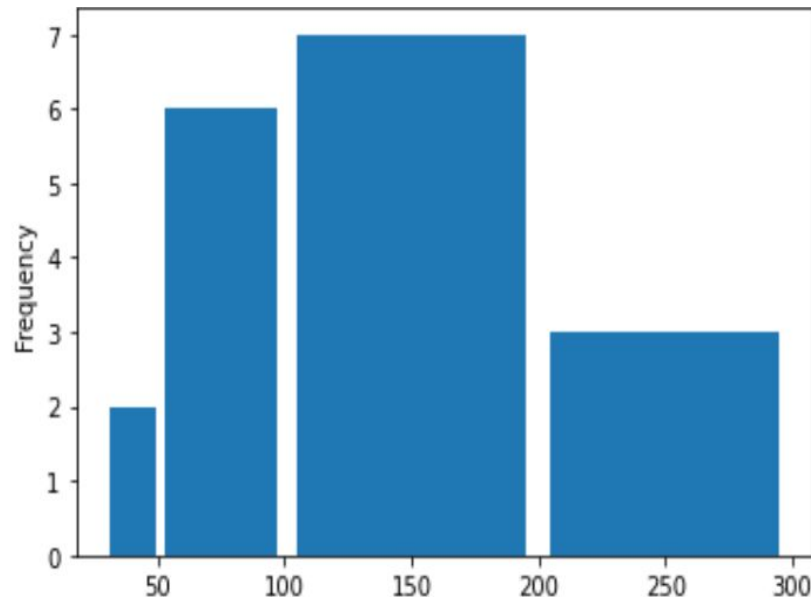
HOUSING PRICES

Low: $30\text{mVND} < \text{AHP} < 50\text{mVND}$.

Medium: $50\text{mVND} < \text{AHP} < 100\text{mVND}$.

High: $100\text{mVND} < \text{AHP} < 200\text{mVND}$.

Very High: $\text{AHP} > 200\text{mVND}$.



CONCLUSION

From all above results, we conclude that, the best place for us to set up a new coffee shop is in Go Vap district because there are a lot of people living there (high density), the number of already-working coffee shop is medium (cluster 2) and the average housing price is medium.

THANKS