

Cafes in Penang, Malaysia

By HueyLing Lim

December 20, 2020

1. Introduction

1.1 Background

Penang is a Malaysian state located on the northwest coast of Peninsular Malaysia, by the Malacca Strait. Penang's population stood at nearly 1.767 million as of 2018, while its population density rose to 1,684/km² (4,360/sq mi). As of 2020, Penang has the highest gross domestic product per capita among all Malaysian states. In addition, Penang recorded the nation's third highest Human Development Index at 0.838, after the country's Kuala Lumpur and Selangor Correspondingly, the state has a relatively well-educated population, with a youth literacy rate of 99.5% as of 2014.

Penang's colonial history and architecture are second to none, and its 2008 UNESCO heritage listing really put its capital, George Town, on the map. In recent years, Malaysia has developed a café culture that rivals the likes of New York, London, Paris, and Prague. Penang takes its coffee seriously.

1.2 Business Problem

In this project we will list and visualize all major parts of Penang that has great Cafes. Specifically, this report will be targeted to stakeholders interested in opening a Cafe in Penang, Malaysia.

We will use our data science powers to generate a few most promising neighbourhoods based on these criteria. Advantages of each area will then be clearly expressed so that best possible final location can be chosen by stakeholders.

2. Data Acquisition and Cleaning

2.1 Data Sources

By using Google Map, we have collected data of geographic coordinates for all neighborhoods in major parts of Penang. The analysis focuses on the focus development areas in Penang which is the boroughs around city center George Town including Batu Ferringhi, Tanjung Bungah, Tanjung Tokong, Seri Tanjung Pinang, Gurney, Pulau Tikus, George Town, Jelutong, Air Itam, Farlim, Paya Terubong, Gelugor and Bayan Lepas.

For further analysis, we also collected data from Foursquare API for the venue of all existing Cafes and the ranking of each Cafes in each neighborhood.

2.2 Data Cleaning

Data downloaded from multiple sources were combined into one table. First, we interacted with Foursquare API and get top 100 venues within a radius of 1000 metres for each neighbourhood. The Café with duplicate names were removed. Second, our analysis will be calculation and exploration of 'ranking' (rating, tips and like) of each Cafe by using Foursquare API and focus our attention on those areas. Therefore, the Café without rankings were removed.

3. Methodology

In this project we will list and visualize all major parts of Penang that has great Cafes. We will limit our analysis to major part area around city center George Town. In first step we have collected the required data (location) and identified Cafes around city center according to Foursquare categorization. Second step in our analysis will be calculation and exploration of 'ranking' (rating, tips and like) of each Cafe by using Foursquare API and focus our attention on those areas. In final step we will present map of the Neighborhoods around city center with top average Rating which should be a starting point for final exploration and search for optimal venue location by stakeholders.

4. Results and Discussion

Our analysis shows that there is 160 Cafes in major parts of Penang around city center George Town. Although the borough "Bayan Lepas" has highest number of neighborhoods, but the borough "George Town" has the highest number of Cafes with a total count of 60 (around 40%). By using Foursquare API, we found that the Cafe "PIKNIK" at George Town has maximum likes and tips and the Cafe "Meraki Coffee" at Tanjung Tokong has maximum rating. Result of the Foursquare Ranking, we found that the boroughs with top 3 of average rating greater or equal 7.0 which are George Town, Tanjung Tokong and Batu Ferringhi.

5. Conclusion

Purpose of this project was to list and visualize all major parts of Penang that has great Cafes to aid stakeholders in narrowing down the search for optimal location for a new Cafe. By analysis ranking from Foursquare data, we have first identified general boroughs that justify further analysis, and then generated extensive collection of locations which satisfy some basic requirements regarding existing nearby Cafes.

Final decision on optimal restaurant location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location (proximity to park or water), levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.