Proposed Location for New Shopping Mall in Bangkok, Thailand

IBM Applied Data Science Cousera Capstone Project

By Audrey Cheang

August 2021

Business Problem

- Location in which shopping malls are located is a key deciding factor on whether it will be profitable

- Objective: To analyze the different districts in Bangkok, Thailand and determine on the optimal location to build a new mall

- Business Question: Where in Bangkok, Thailand, would you recommend the construction of a new shopping mall?

Data Science Methodology

- Required Data:
  - Districts in Bangkok and their coordinate values
  - Data on existing venues and their categories
  - Map of Bangkok

- Data Sources:
  - Wikipedia for district data
  - Foursquare for venues and categories
  - Folium package for visualisation

Data Science Methodology

- Web scrapping of Wikipedia page to obtain table of districts
- Call the Foursquare API to get data on venues and their categories
- Data is grouped by their categories and 'shopping mall' categories are extracted
- Cluster the districts into 3 levels of shopping mall concentration
- Visualize the collected data using Folium

Results

- Cluster 0 (red): Moderate concentration of shopping malls in district

- Cluster 1 (purple): High consternation of shopping malls in district

- Cluster 2 (green): Low concentration of shopping malls in district
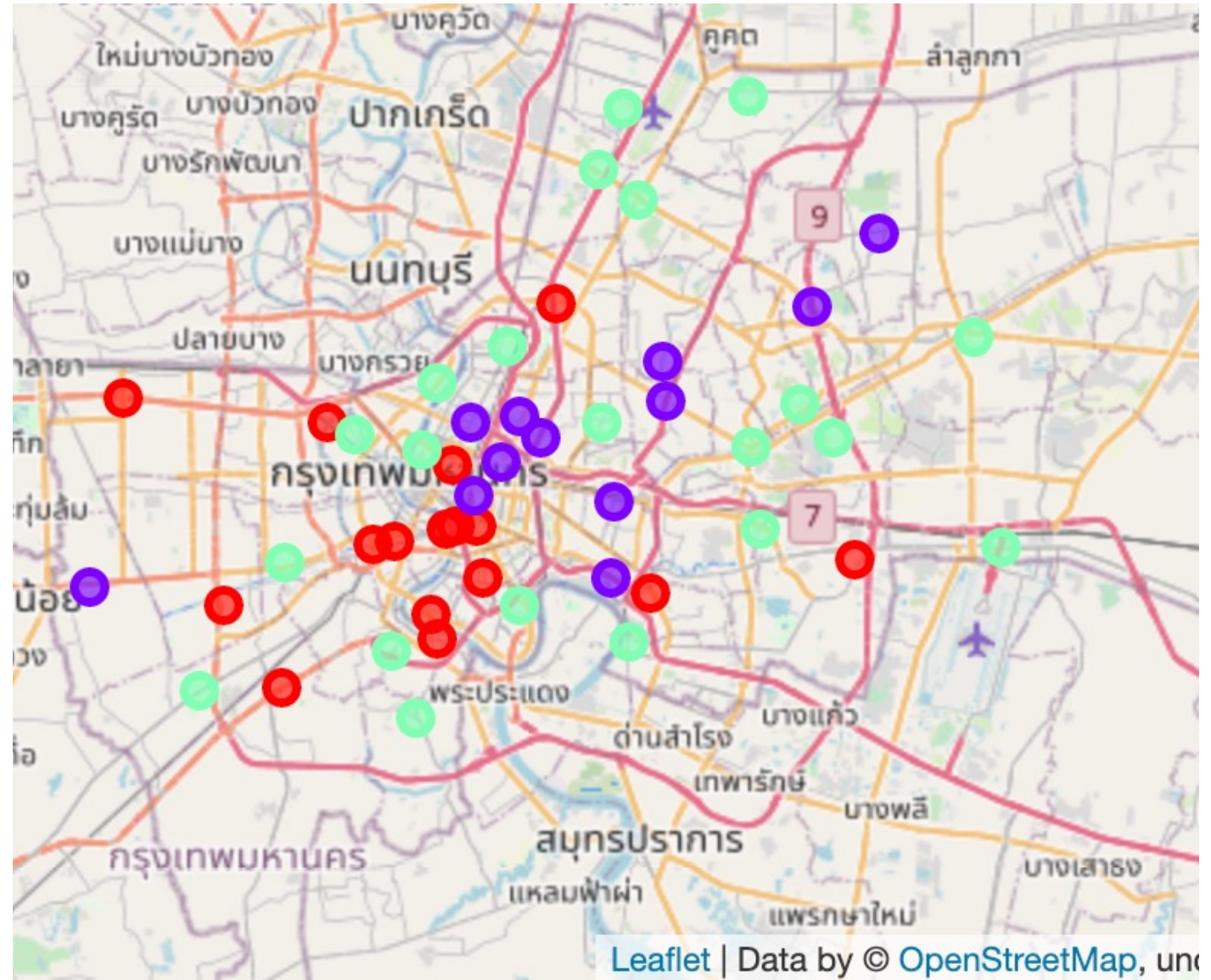
Discussions

- Green cluster: districts with the lowest concentration of shopping malls
  - potential areas to open new shopping malls
  - residential buildings more spread out → imply a lower shopping mall concentration
  - presents an opportunity to enter the market as the low concentration of shopping malls coupled with the lack of options for consumers in the area would make a new shopping mall complex in these areas very attractive
- Red cluster: districts with moderate concentration of shopping malls
- Purple cluster: districts with the highest concentration of shopping malls
  - populated with shopping malls and constructing a new shopping mall would imply that there would be intense competition in the area
  - purple districts are areas that are more highly populated, with many residential buildings observed, along with several essential services like hospitals and public transport stations → developers would anticipate high volumes of people, resulting in the correspondingly higher concentration of shopping malls

Limitations and Recommendations

- Project only considers the concentration of shopping malls in each district to identify potential new areas to build a new shopping mall
- Other factors that influence whether a shopping mall should be built:
  - accessibility to shoppers
  - average wealth and income levels of a particular district
  - presence of places of attraction in the district
- Further research should delve into these areas and identify the reasons that could have contributed to the existing high and low concentration in these districts

Conclusion

- In this project, we have undergone the process of solving a business problem by identifying the data that we needed, extracting that preparing the data, analysing, and visualising the data and finally, utilising machine learning to cluster the data, in order to provide a solution to the business problem 'Where in the city of Bangkok will you recommend a property developer set up a new shopping mall?'.

- Upon analysis, it was concluded that that low concentration districts, as identified from our study, are potential areas where developers can consider building a new shopping mall complex. However, it was also acknowledged that this study did not cover certain factors that could have possibly affected the status of the data that was obtained and analysed, and deeper analysis into these areas are required to come to a more comprehensive conclusion.