CAPSTONE PROJECT

THE BATTLE OF NEIGHBORHOODS

WEEKS 4 & 5

LOCATION ANALYSIS FOR A TRAVEL AGENCY BUSINESS IN SINGAPORE

WRITTEN BY:

MICHAEL

24TH APRIL 2020

Introduction

Singapore is an island nation located in Southeast Asia. For over two decades, Singapore has been a thriving global hub and one of the most successful financial hubs of Asia. It is a multicultural nation with a population of over 5.6 million people and a Gross Domestic Product (GDP) of over 360 billion USD [1-2].

As Singapore faces land scarcity and with the increasing numbers of businesses, economic competitiveness grows among businesses in Singapore [3]. For this project, I am going to focus on finding the best possible location for the main office, with efforts in avoiding competition with other businesses in the same sector.

Suppose an individual wants to start a travel agency business in Singapore. Travel agencies are less common nowadays, but is still a profitable business as starting one is usually inexpensive and there is still demand for tours, affordable travel packages, customized trips, and so on. The main office should be located in a strategic location so that people could drop by the office easily [4-6].

In order to avoid competition in this tight sector, I will avoid areas with many travel agencies and also not to locate the office closely to existing travel agencies. It is also appropriate to locate the office as close as possible to means of transportation, venues, or landmarks of a targeted area.

The project objective is to analyze and determine the optimal area to locate a travel agency business in Singapore depending on factors such as competition, surrounding venues, and accessibility. Other factors such as land pricing and area popularity are not accounted for in this project and will serve as future improvements.

Data

For this project, the following data is required:

- List of regions and areas of Singapore and list of geographical coordinates of each area boundary.
  - Source: Data.gov.sg – Master Plan 2019 Planning Area Boundary (No Sea)
  - Description: This dataset contains area boundary data.
  - Usage: This dataset contains complete area boundary data; however, only the list of regions and areas of Singapore will be extracted first. The area boundary coordinates will be extracted later on to generate a choropleth map [7].

- Geographical coordinates of each area.
  - Source: Google Maps Platform – Geocoding API (via *GeoPy* library)
  - Description: This API facilitates the conversion between addresses and geographical coordinates.
  - Usage: I will be using this API to generate geographical coordinates for a given area. The output coordinates are mostly approximate values. This geocoder is preferred over OpenStreetMap Nominatim as it outputs more accurate results [8].

- Map of Singapore.
  - Source: OpenStreetMap (via *Folium* library)
  - Description: This platform provides the world map.
  - Usage: I will be using this library to generate leaflet maps and choropleth maps of Singapore. The leaflet map with superimposed markers will show areas of Singapore, and the choropleth map will show the variation of the number of travel agencies by area [9].

- List of existing travel agencies in each area and list of surrounding venues of a given area.
  - Source: Foursquare API
  - Description: This API facilitates the interaction with Foursquare platform.
  - Usage: I will be using this API to retrieve the detailed list of existing travel agencies in each area. I will also retrieve the surrounding venues of a given area [10].

- Distance between two geographical coordinates.
  - Source: GeoPy Library – Distance Function
  - Description: This function calculates the geodesic distance between two geographical coordinates.
  - Usage: I will be using this function to retrieve the distance between two geographical coordinates: the approximate coordinates of the business and the existing travel agencies in a specific area [11].

Methodology

Part 1 - Identifying regions and areas of Singapore.

First, I identified the regions and areas of Singapore using the area boundary data. Then, using the *Google Geocoding API,* I extracted the geographical coordinates (latitude and longitude values) of all areas, and then generated a dataframe. Figure 1 below shows the first five rows of the resulting dataframe.

	Area	Region	Latitude	Longitude
0	BEDOK	EAST REGION	1.323604	103.927340
1	BOON LAY	WEST REGION	1.314256	103.709310
2	BUKIT BATOK	WEST REGION	1.349238	103.758483
3	BUKIT MERAH	CENTRAL REGION	1.281905	103.823918
4	BUKIT PANJANG	WEST REGION	1.386664	103.771581

Figure 1: Dataframe consisting of areas and regions and their respective latitude and longitude values.

Second, I plotted the number of areas by region using the above dataframe in the form of a bar graph using *Matplotlib*. Figure 2 on the next page shows that Singapore consists of 5 regions and 55 areas in total.

Third, I retrieved the geographical coordinates of Singapore using the *Google Geocoding API*, and then plotted the map of Singapore using *Folium*. Using the above dataframe, I superimpose markers on top of the map to label each area. Figure 3 on the next page shows the resulting map.

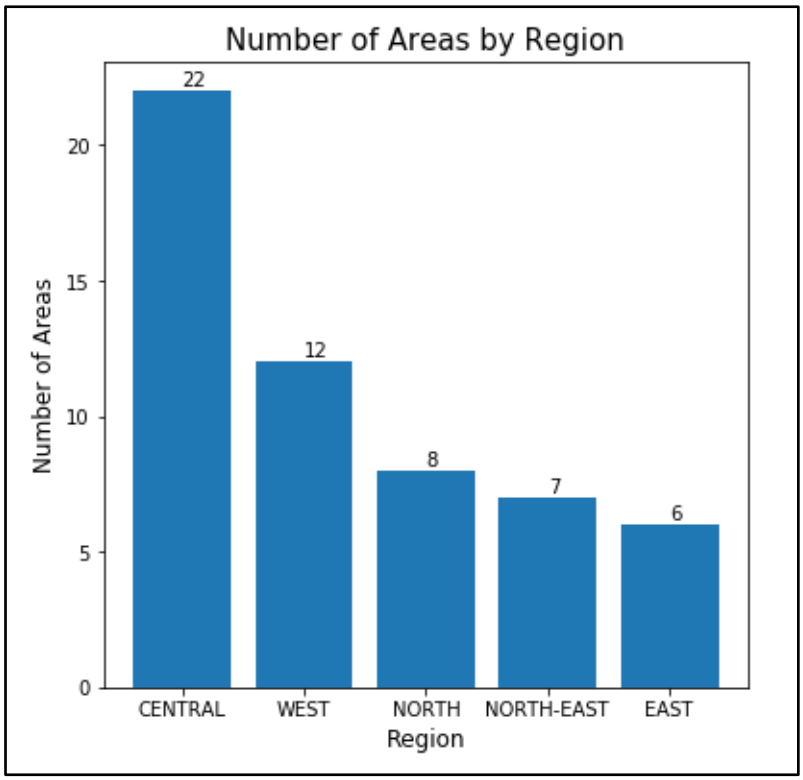

Figure 2: Bar graph showing number of areas by region.

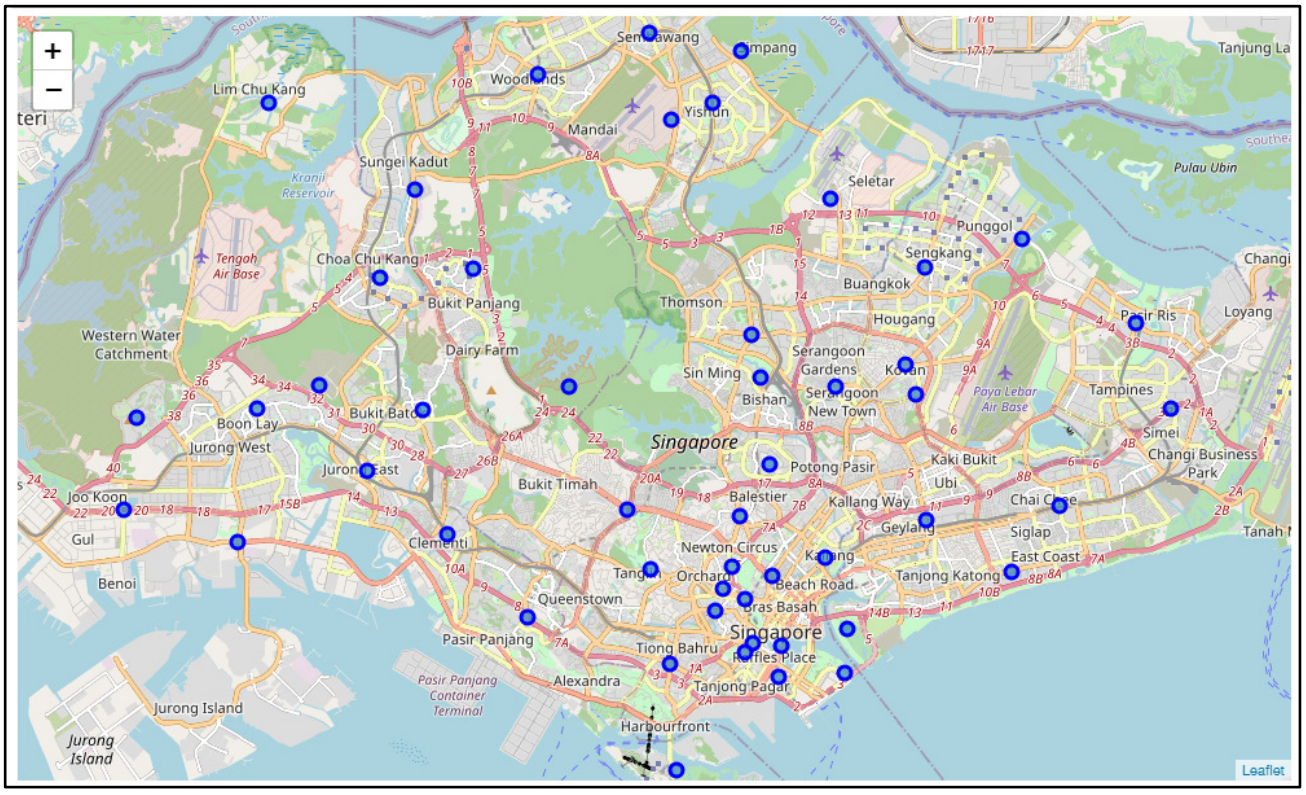

Figure 3: Map (zoomed) with superimposed markers showing the areas of Singapore.

Part 2.1 - Exploring travel agencies in all areas of Singapore using Foursquare API.

Next, I explored travel agencies in all areas of Singapore using the *Foursquare API*. Using the *Foursquare API*, I generated a dataframe consisting of travel agencies within 1500 meters of all areas. Figure 4 below shows the first five rows of the resulting dataframe.

	Area	Region	Area Latitude	Area Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Venue ID
0	BUKIT MERAH	CENTRAL REGION	1.281905	103.823918	Pricebreaker	1.293361	103.832605	Travel Agency	4dafa753a86e9639ef520ffe
1	BUKIT PANJANG	WEST REGION	1.386664	103.771581	WTS Travel	1.379745	103.764582	Travel Agency	4d6cc934fbf0a0933e88f68c
2	BUKIT PANJANG	WEST REGION	1.386664	103.771581	WTS Travel	1.377157	103.767049	Travel Agency	4d2d7249853ff04d1fdac8da
3	CHOA CHU KANG	WEST REGION	1.383980	103.746961	WTS Travel	1.384917	103.744998	Travel Agency	4bc18f614cdfc9b611dd9421
4	CHOA CHU KANG	WEST REGION	1.383980	103.746961	WTS Travel	1.376225	103.752877	Travel Agency	4d941dfd14a65481f14b0a4c

Figure 4: Dataframe consisting of information about existing travel agencies.

The dataframe above consists of 442 rows and 9 columns. After examining the dataframe, I found that most of the entries are duplicated entries, which is caused by some areas located closely to another. I removed the duplicated entries and only kept the first entry from each set of duplicated entries. Subsequently, the dataframe consists of 133 rows and 9 columns as there are 309 duplicated entries.

Then, I plotted the number of travel agencies by area using the above dataframe in the form of a bar graph using *Matplotlib*. Figure 5 on the next page shows that only 27 out of 55 areas have travel agencies. I decided to not instantly pick areas without existing travel agencies as probably the market there is not favorable for the business.

To better visualize, I also plotted a choropleth map indicating the number of travel agencies using *Folium*. Figure 6 on the next page indicates that most areas do not have existing travel agencies (lightest color) and very few areas have many existing travel agencies (darkest color).

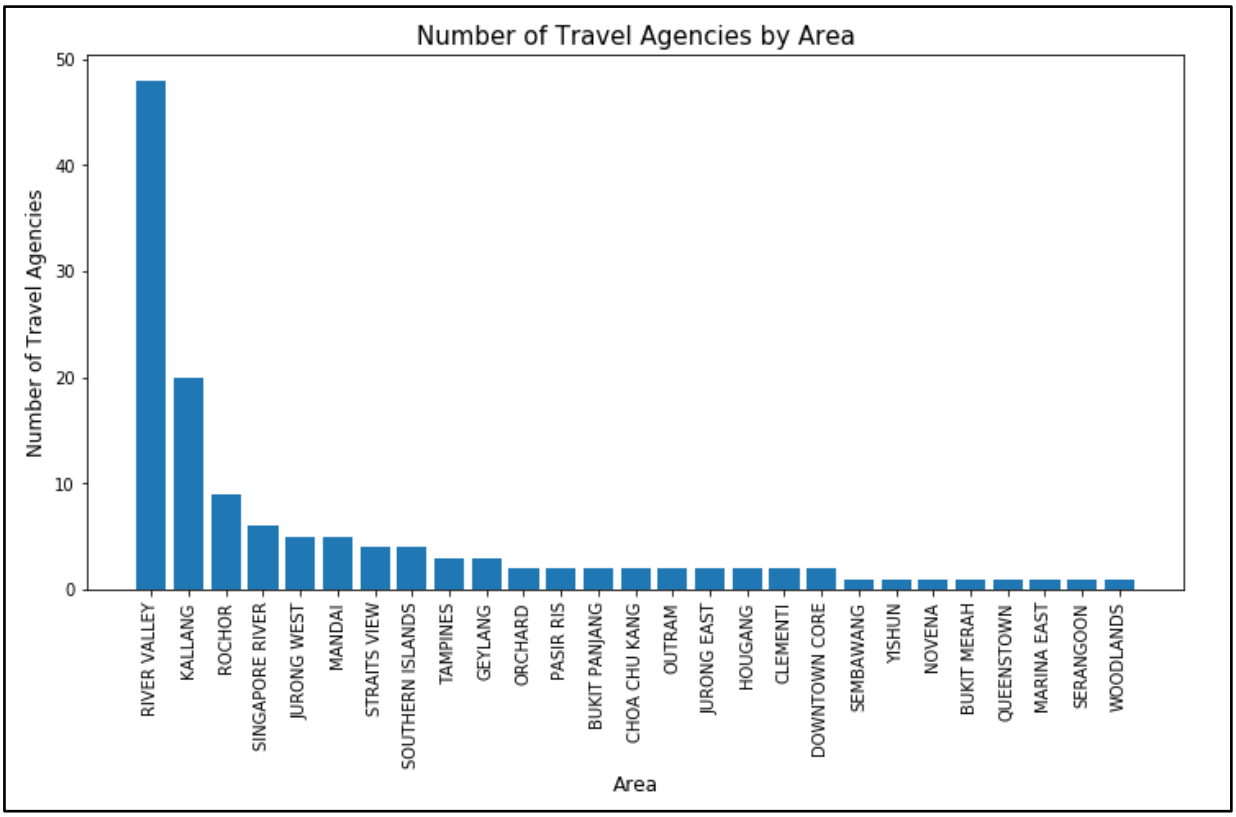

Figure 5: Bar graph showing the number of travel agencies by area.

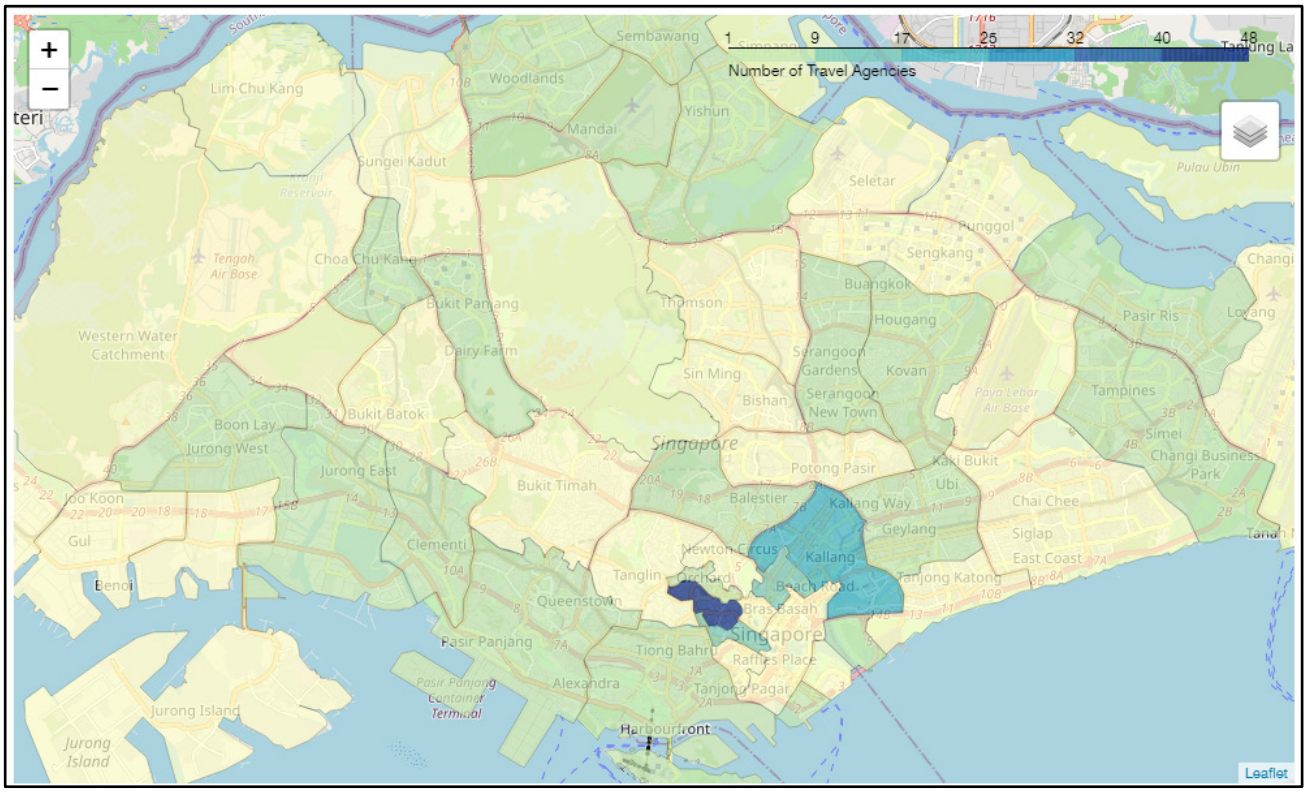

Figure 6: Choropleth map indicating the number of travel agencies by area.

In order to categorize the areas and find out how many travel agencies are present per region, I plotted the number of travel agencies by region in the form of a bar graph using *Matplotlib*. This is shown in Figure 7 below.

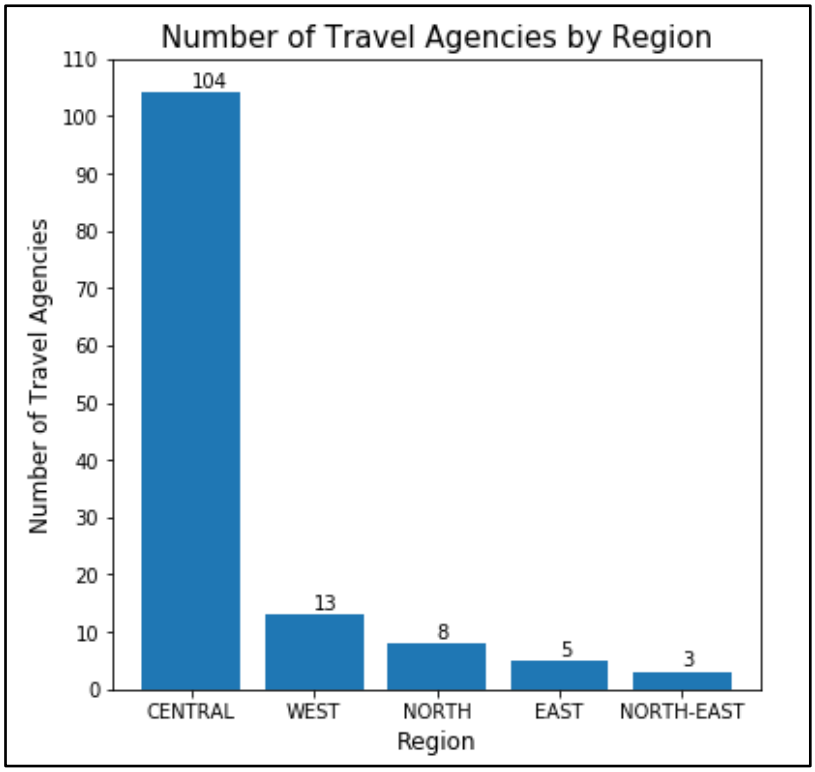

Figure 7: Bar graph showing the number of travel agencies by region.

From the graph above, there are 133 travel agencies in total, and the region packed with the most travel agencies is the Central Region, which has 104 travel agencies and should be avoided. I picked the region with moderate number of travel agencies to consider for the business, which is the West Region. I avoided regions with low numbers of travel agencies as probably the market there is not favorable for the travel agency business.

Part 2.2 - Exploring travel agencies in the West Region of Singapore using Foursquare API.

Next, I explored travel agencies located in the West Region. Using the

previous dataframe, I looked into only the travel agencies in the West Region. Figure 8 below shows the first five rows of the filtered dataframe.

	Area	Region	Area Latitude	Area Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Venue ID
0	BUKIT PANJANG	WEST REGION	1.386664	103.771581	WTS Travel	1.379745	103.764582	Travel Agency	4d6cc934fbf0a0933e88f68c
1	BUKIT PANJANG	WEST REGION	1.386664	103.771581	WTS Travel	1.377157	103.767049	Travel Agency	4d2d7249853ff04d1fdac8da
2	CHOA CHU KANG	WEST REGION	1.383980	103.746961	WTS Travel	1.384917	103.744998	Travel Agency	4bc18f614cdfc9b611dd9421
3	CHOA CHU KANG	WEST REGION	1.383980	103.746961	WTS Travel	1.376225	103.752877	Travel Agency	4d941dfd14a65481f14b0a4c
4	CLEMENTI	WEST REGION	1.316181	103.764938	Global Holidays	1.303585	103.764996	Travel Agency	4ffcde13e4b0279c5b32126a

Figure 8: Filtered dataframe consisting of information about existing travel agencies in the West Region.

Then, in order to categorize those listed travel agencies, I plotted the number of travel agencies by area of the West Region in the form of a horizontal bar graph using *Matplotlib*. This is shown in Figure 9 below.

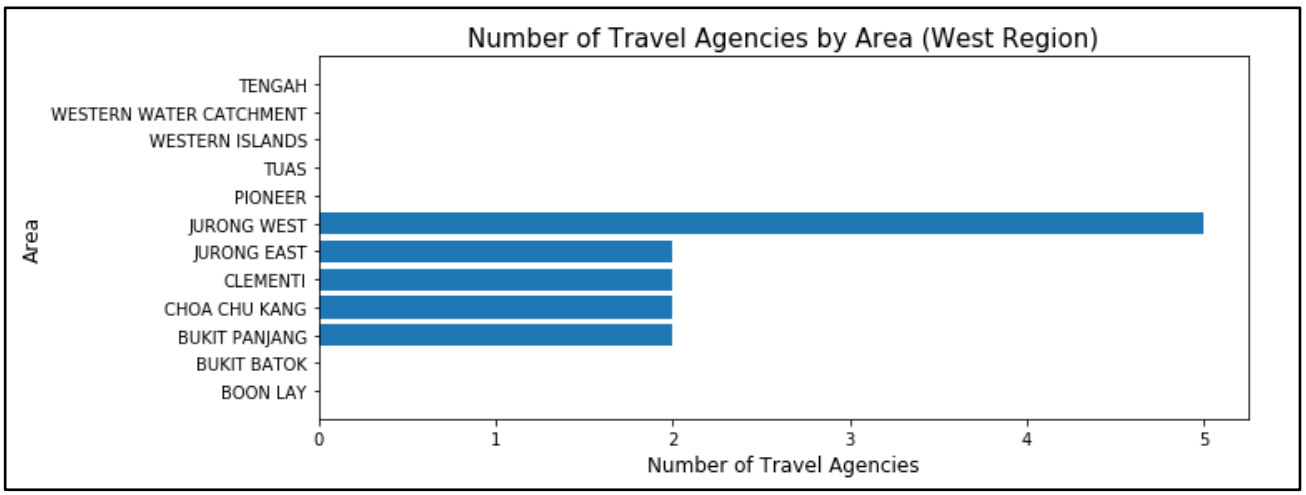

Figure 9: Horizontal bar graph showing the number of travel agencies by area of the West Region.

From the graph above, Jurong West has the greatest number of travel agencies and the other areas have 2 or no travel agencies. Many areas have no existing travel agencies, but before making a decision, I will check the top five venues around those areas. Again, some areas might not be favorable for the business.

Part 3 - Analyzing the best possible location in the West Region.

Next, to compare the different areas, I decided to use the *Foursquare API* to return a list of recommended venues near an area. I searched for the top 5 recommended venues of each area within a radius of 250 meters. However, for some areas, *Foursquare* returned less than 5 venues and in some areas, it returned no venues. In the case of no venues, I consider those areas not suitable for the business. I generated a dataframe consisting of the areas and the top 5 recommended venues around the areas. Figure 10 below shows the resulting dataframe.

	Area (WEST REGION)	1st Recommended Venue	2nd Recommended Venue	3rd Recommended Venue	4th Recommended Venue	5th Recommended Venue
0	BOON LAY	Bus Station	Bus Station	NONE	NONE	NONE
1	BUKIT BATOK	Park	Food Court	Food Court	Fast Food Restaurant	Grocery Store
2	BUKIT PANJANG	Basketball Court	Athletics & Sports	Playground	Candy Store	Coffee Shop
3	CHOA CHU KANG	Portuguese Restaurant	Italian Restaurant	Supermarket	Pharmacy	Furniture / Home Store
4	CLEMENTI	Video Game Store	Soup Place	Dessert Shop	Bakery	Fried Chicken Joint
5	JURONG EAST	Clothing Store	Furniture / Home Store	Chinese Restaurant	Bakery	Movie Theater
6	JURONG WEST	Playground	Bus Stop	NONE	NONE	NONE
7	WESTERN WATER CATCHMENT	Fast Food Restaurant	NONE	NONE	NONE	NONE

Figure 10: Dataframe consisting of areas and its surrounding top 5 recommended venues.

To compare further, I generated a dataframe consisting of the areas and the number of travel agencies present in the respective areas, and then combined with the above dataframe. This is shown in Figure 11 below.

	Area (WEST REGION)	Number of Travel Agencies	1st Recommended Venue	2nd Recommended Venue	3rd Recommended Venue	4th Recommended Venue	5th Recommended Venue
0	BOON LAY	0	Bus Station	Bus Station	NONE	NONE	NONE
1	BUKIT BATOK	0	Park	Food Court	Food Court	Fast Food Restaurant	Grocery Store
2	WESTERN WATER CATCHMENT	0	Fast Food Restaurant	NONE	NONE	NONE	NONE
3	BUKIT PANJANG	2	Basketball Court	Athletics & Sports	Playground	Candy Store	Coffee Shop
4	CHOA CHU KANG	2	Portuguese Restaurant	Italian Restaurant	Supermarket	Pharmacy	Furniture / Home Store
5	CLEMENTI	2	Video Game Store	Soup Place	Dessert Shop	Bakery	Fried Chicken Joint
6	JURONG EAST	2	Clothing Store	Furniture / Home Store	Chinese Restaurant	Bakery	Movie Theater
7	JURONG WEST	5	Playground	Bus Stop	NONE	NONE	NONE

Figure 11: Dataframe consisting of areas, the number of travel agencies present, and its surrounding top 5 recommended venues.

Referring to Figure 11, I can decide the location based on two factors, namely the number of travel agencies and the venues near that area. The business can be located in Boon Lay, Bukit Batok, or Western Water Catchment as in those areas there are no travel agencies present.

Based on the venues, it is appropriate to locate the business in Bukit Batok as the nearby venues are mostly eateries, which is favorable for the business. It is less appropriate to locate the business in Boon Lay as it is near 2 bus stations but there are no interesting nearby venues. I avoided Western Water Catchment as it is only near a fast food restaurant, and there are no interesting venues or means of transportation nearby.

Finally, to locate the business in Bukit Panjang, Choa Chu Kang, Clementi, or Jurong East, where there are only 2 travel agencies in each area, the distances between the business and existing travel agencies must be maximized in order to avoid competition. Suppose the business is going to be located near the Mass Rapid Transit (MRT) station of an area, then the distances between the MRT station of each area and the existing travel agencies must be maximized.

I calculated the geodesic distances between the MRT station and the existing travel agencies of each area using the *GeoPy* distance function. I then discovered that the distances are maximized in Clementi, and minimized in Bukit Panjang. If the business is located in Clementi, the distance between the business and existing travel agencies are greater than 1km, and thus satisfying the condition. Moreover, the recommended venues in Clementi comprise of mostly eateries, which is favorable for the business.

More in-depth explanations are provided in my project notebook here.

Results

- In Singapore, there are only 133 travel agencies in total and no travel agency is present in more than half of the areas. Specifically, travel agencies are only present in only 27 areas out of 55 areas.

- Most of the travel agencies are located in the Central Region and in total, there are 104 travel agencies located in the Central Region.

- Few travel agencies are located in the other regions, namely West, North, East, North-East regions. The number of travel agencies for each region are 13, 8, 5, and 3 respectively.

- Out of all areas with travel agencies, most areas have less than 10 travel agencies.

- In the West Region, 5 travel agencies are located in Jurong West, and the other areas have 2 or no travel agencies present.

- Only 5 out of 12 areas in the West Region have 5 recommended venues.

- Most of the surrounding recommended venues in the West Region are eateries and shops.

- In areas with 2 existing travel agencies in the West Region, the distances between the MRT station and the existing travel agencies is maximized in Clementi, and minimized in Bukit Panjang.

Discussion

Based on the results obtained, there are few travel agencies in Singapore. I avoided the Central Region as it is packed with travel agencies, and chose the West Region as it has a moderate number of existing travel agencies. Other regions have small numbers of existing travel agencies, but while this may avoid competition, the market or demand there may be challenging or unsupportive for the business. Hence, I chose the West Region.

The West Region consists of 12 areas, and most areas have 2 or no travel agencies present. I decided to not easily pick the area with the least number of existing travel agencies, and explore the surrounding venues of each area. I then found out that only 5 out of 12 areas have 5 recommended venues, and most of the surrounding recommended venues are eateries and shops.

Based on the venues, it is appropriate to locate the business in Bukit Batok as the nearby venues are mostly eateries and there are no existing travel agencies in this area, which is favorable for the business. Alternatively, it is acceptable to locate the business in Boon Lay as it is near two bus stations, but note that there are no other nearby interesting venues within 250m.

Finally, I explored areas with 2 existing travel agencies in the West Region as the surrounding recommended venues are favorable for the business. At the end, I found out that it is optimal to locate the business in Clementi as the distances between Clementi MRT Station and existing travel agencies in Clementi are maximized, avoiding competition.

Conclusion

The purpose of this project is to identify optimal areas with low numbers of existing travel agencies and also supportive surrounding venues. This is to avoid competition in this tight sector and enhance customer experience. By using *Google Geocoding API* to generate geographical coordinates, *Foursquare API* to list down and explore venues based on the coordinates, and using *GeoPy* distance function to calculate distance based on geographical coordinates, I analyzed the retrieved data and derived the optimal areas to start the business, namely Bukit Batok and Clementi.

Although I only examined the West Region due to the moderate number of existing travel agencies, the same analysis procedure can be used to examine other regions and derive optimal areas in other regions. This project does not pinpoint exact locations on where to start the business, but rather suggest optimal areas. This analysis is derived from the data provided by *Google Geocoding API*, *Foursquare API*, and *GeoPy* library. Thus, the data accuracy and integrity solely depend on the source of data.

To conclude, the final decision on where to locate the business depends on several other factors such as land or rental prices, land availability, area popularity, demand in each area, and so on. Note that travel agencies are less common nowadays, but is still a potential business idea as the capital requirement to start a travel agency is generally low and can be very profitable, and also there is still demand for affordable tours and travel packages.

References

[1] Wikipedia. 2020. Singapore. [online] Available at: <https://en.wikipedia.org/wiki/Singapore> [Accessed 10 April 2020].

[2] World Bank Data. n.d. Singapore. [online] Available at: <https://data.worldbank.org/country/singapore> [Accessed 10 April 2020].

[3] Wikipedia. 2020. Economy of Singapore. [online] Available at: <https://en.wikipedia.org/wiki/Economy_of_Singapore> [Accessed 10 April 2020].

[4] Muriithi, R., 2019. 7 Profitable Small Business Ideas In Singapore 2019. [online] ZUU Online. Available at: <https://zuuonline.sg/business/startup/top-profitable-small-business-ideas-in-singapore-2019/> [Accessed 10 April 2020].

[5] Morgaine, B., 2019. How to Become a Travel Agent: Starting Your Business. [online] Bplans. Available at: <https://articles.bplans.com/how-to-start-a-travel-agency/> [Accessed 10 April 2020].

[6] Shah, J., 2018. 13 Easy Steps On How To Start A Travel Agency. [online] designhill. Available at: <https://www.designhill.com/design-blog/easy-steps-on-how-to-start-a-travel-agency/> [Accessed 10 April 2020].

[7] Data.gov.sg. 2020. Master Plan 2019 Planning Area Boundary (No Sea). [online] Available at: <https://data.gov.sg/dataset/master-plan-2019-planning-area-boundary-no-sea> [Accessed 10 April 2020].

[8] Google Developers. 2020. Geocoding API Documentation. [online] Available at: <https://developers.google.com/maps/documentation/geocoding/start> [Accessed 16 April 2020].

[9] Rob Story. 2019. Folium 0.10.1 Documentation. [online] Available at: <https://python-visualization.github.io/folium/> [Accessed 15 April 2020].

[10] Foursquare. n.d. Places API Reference. [online] Available at: <https://developer.foursquare.com/docs/api-reference/venues/search/> [Accessed 15 April 2020].

[11] GeoPy. n.d. GeoPy 1.21.0 Documentation. [online] Available at: <https://geopy.readthedocs.io/en/stable/> [Accessed 15 April 2020].