# Potential location for new resort investment in Thailand

Wirach Leelakiatiwong

June 02, 2020

## Introduction

Tourism is an economic contributor to Thailand (with more than 20% of Thailand's GDP), Travel restriction due to COVID-19 pandemic will for sure causes tremendous shrink in Thailand's GDP. But this situation will not stay forever. When the pandemic is over, lots of travelers will begin their journey to Thailand again. Therefore, this is a good opportunity to find an optimal location to start new resort in Thailand.

There are lots of traveling province in Thailand. Start from mountainous area in Northern part like Chaing mai, Chiang rai down to the Southern part where tons of beautiful tropical beach lay there (such as Koh samui : Surat Thani, Krabi, Phuket). So in beginning part of our project we'll try to figure out which province will be focused based on their traveling income per number of hotel rooms.

For focused province, we will try to detect locations that are not already crowded with resorts. We are also particularly interested in areas with filled with community and facilities. We would also prefer location as close to city center as possible, assuming that first two conditions are met.

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

## Data

Based on definition of our problem, factors that will influence our decision are:

* Traveling income per number of hotel rooms of each province
* Number of existing hotels in the neighborhood
* Number of and distance to resorts in the neighborhood, if any
* Distance of neighborhood from city center

We decided to use regularly spaced grid of locations, centered around city center, to define our neighborhoods.

Following data sources will be needed to extract/generate the required information:

* Travelling income and number of hotel rooms by province from **National Statistical Office of Thailand**
* Centers of candidate areas will be generated algorithmically and approximate addresses of centers of those areas will be obtained using **ARCGIS geocoding API**
* Number of hotels and their type and location in every neighborhood will be obtained using **Foursquare API**
* Coordinate of focused province city center will be obtained using **ARCGIS reverse geocoding API**

### Data cleansing and features selection

Data downloaded from NSI were in form of excel spreadsheet which were merged together and they were in format that was not ready to be use so we decide to clean some section before proceed on further process.

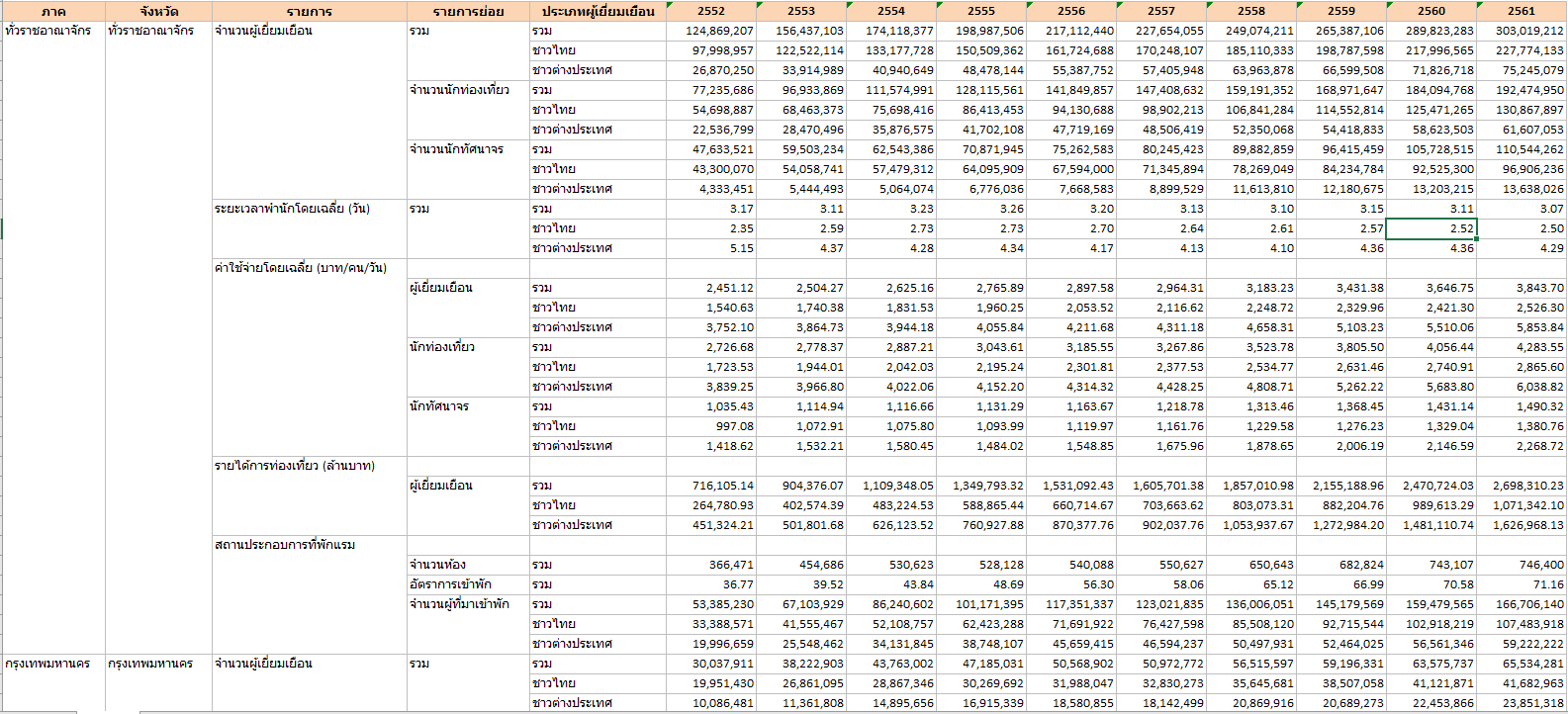


Figure 1 Data table before processing

First of all, we decide to keep only data from most recent year (2561 or 2018). And select only information we need for analysis



Figure 2 Data table after processing

## Exploratory Data Analysis

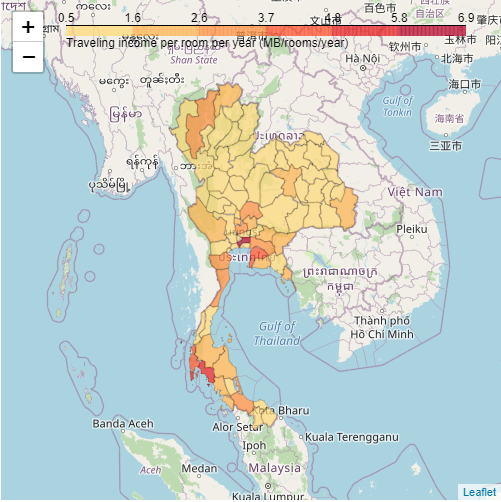
### Calculation of income/room/year for potential province selection

With our data we can calculate income per room per year by divide traveling income in 2018 by number of room available in 2018. As a result we’ll know which province have potential for our investment in building new resort.



Figure 3 Data table sorted by income/room/year

For visualization, we decide to plot choropleth map centered on Thailand.



**Phuket,Thailand**

Figure 4 Choropleth map of Thailand

From above map, potential provinces are spread along southern part of Thailand. Apart from Bangkok, the most potential province we'll focus on our further study will be **Phuket**

### Explore for neighborhood location

As we look deep into Phuket, we can get neighborhood information (Geographical coordinates) by passing data of borough and neighborhood and send request to ARCGIS geocoder API. As result we’ll get Latitude and Longitude of each neighborhood as show in Figure 5.



Figure 5 Geographical location of neighborhoods in Phuket

### Searching for candidate neighborhood

As we know geographical location of each neighborhood we can pass this information to Foursquare API. We look forward to get back information about various venues in each neighborhood so we can further analyze which neighborhoods are good for our investment of new resort.

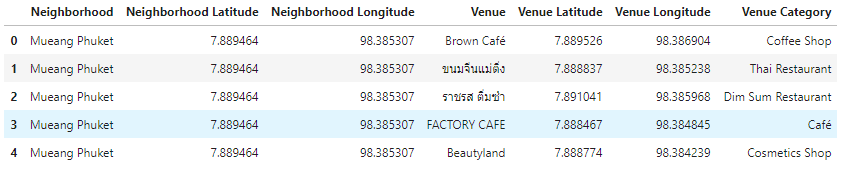


Figure 6 Venue list

We put further effort on grouping venue to each neighborhood and find out top 10 most common venues for each neighborhood.

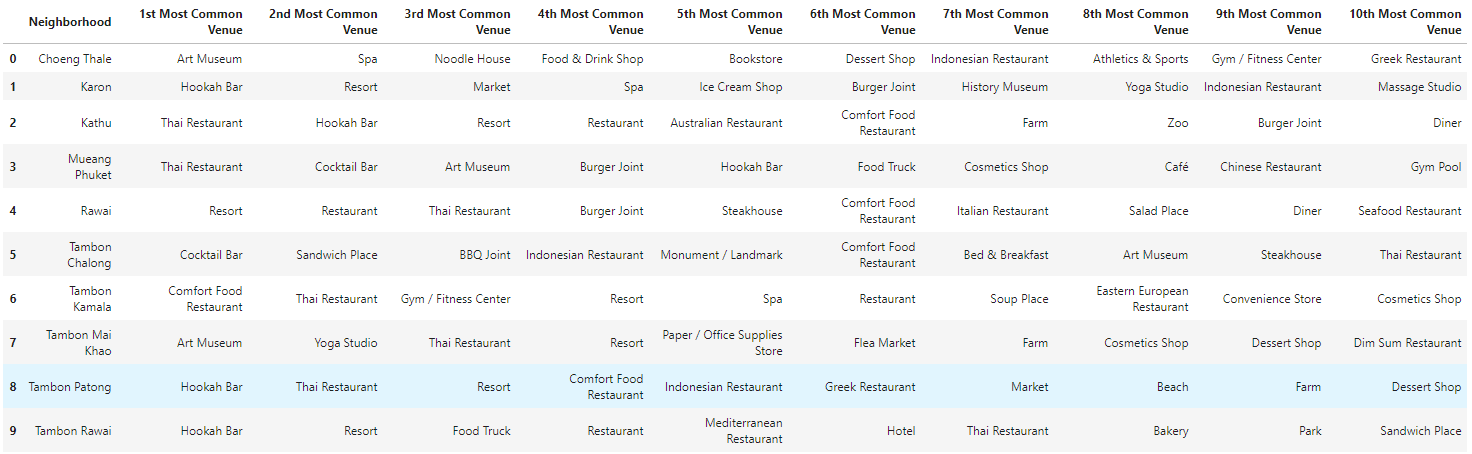


Figure 7 Top 10 venue in each neighborhood

### Clustering neighborhood

First of all, to determine optimal cluster number, we run through elbow method and find out that “6” is optimum number of cluster. Then we cluster our neighborhoods in to 6 cluster as follow.

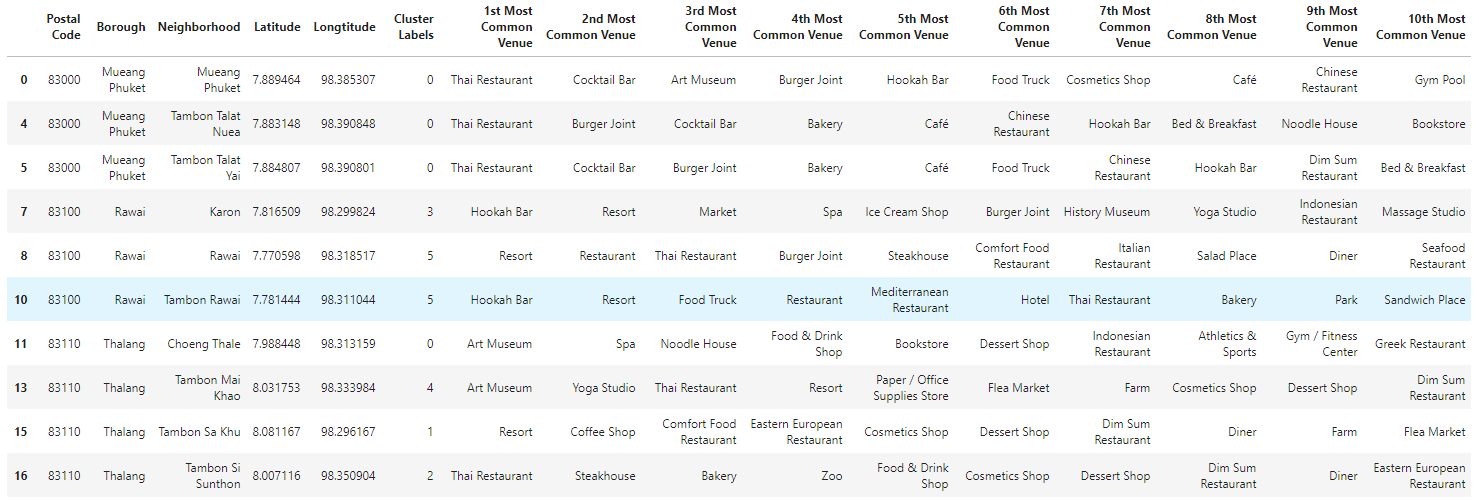


Figure 8 Cluster of neighborhood in Phuket

## Result and Discussion

To help us visualize characteristic of each cluster, we take top 3 venues in each cluster and accumulate their frequency as show in figure 9

Among 6 neighborhood cluster in Phuket, Cluster number 3 is the most suitable for our investment in new resort for following reason

* Moderate amount of restaurants in the area compare to other clusters
* Cluster 3 has market which other clusters don’t have.
* Phuket is multicultural province fill with foreign travelers, cluster 3 have high number of bars which provide vibrance to its neighborhood
* Advantage of geographical location, with reference to figure 10, cluster 3 located near Phuket city center and near the most famous tropical beach (Patong beach)

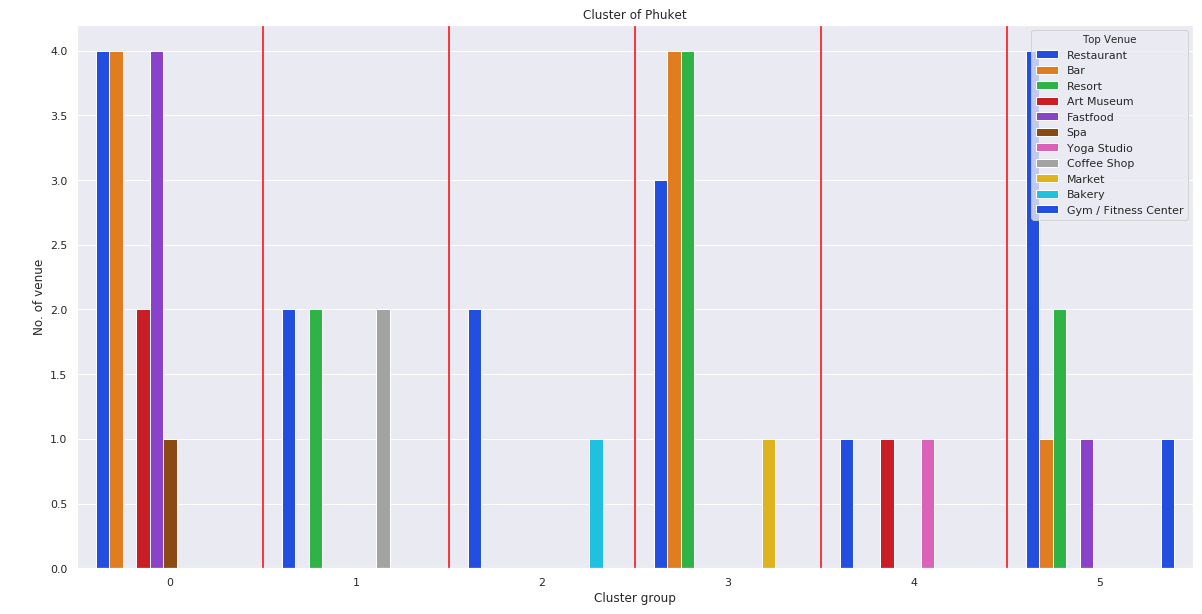


Figure 9 Characteristic of each cluster

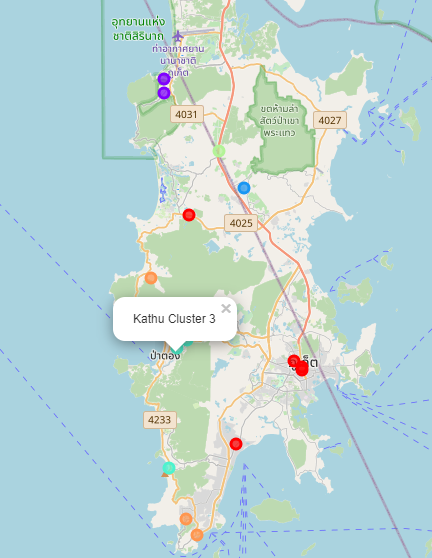


Figure 10 Cluster distribution in Phuket

One major drawback of cluster 3 is it fills with lots of hotels and resorts as show in figure 11, However if we look into heat map of number of resort in cluster 3 (figure 12) we’ll found that there is still some room for settlement of new resort in south of Patong beach and north of Patong beach.

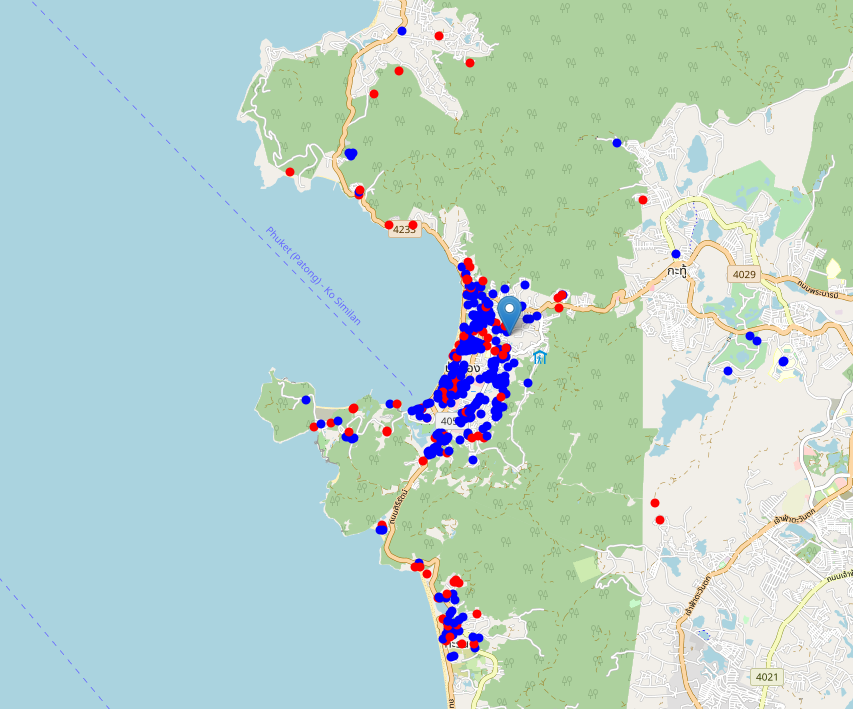


Figure 11 Hotels and Resorts distribution in cluster 3 : blue dots represent ‘Hotel’ while red dots represent ‘Resort’

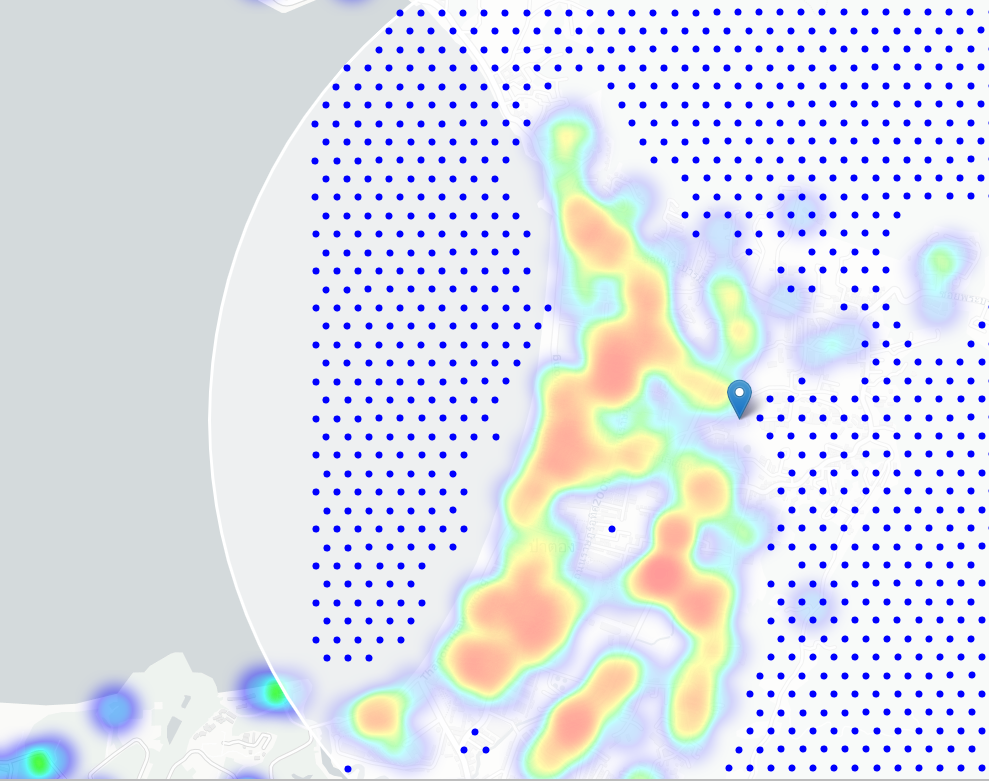


Figure 12 Heatmap of resort in cluster 3 : blue dot represent potential location for new resort investment

## Conclusion

Purpose of this project was to identify suitable location in Thailand in order to aid stakeholders in narrowing down the search for optimal location for a new resort. By calculating traveling income per available room (money making potential) of each province and with help from resort and hotel density distribution from Foursquare data we have first identified general boroughs that justify further analysis (Kathu and Patong), and then generated extensive collection of locations which satisfy some basic requirements regarding existing nearby resort. Clustering of those locations was then performed in order to create major zones of interest (containing greatest number of potential locations) and addresses of those zone centers were created to be used as starting points for final exploration by stakeholders.

Final decision on optimal resort 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, levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.

## Special thank to

* National Statistical Office of Thailand : <http://statbbi.nso.go.th/staticreport/page/sector/en/17.aspx>
* ARCGIS API : [https://developers.arcgis.com/python/guide/](https://developers.arcgis.com/python/guide/reverse-geocoding/)
* Foursquare API : <https://developer.foursquare.com/>
* Thailand GeoJSON file creator : <https://github.com/apisit/thailand.json/blob/master/thailandWithName.json>
* List of postal code and neighborhood of Phuket : <https://postal-codes.cybo.com/thailand/>