The Battle of the Neighbourhoods in the Islands of the Gods

Prepared by: William Gozali for IBM Data Science Professional Certificate

August 29, 2019

1.1 Introduction

Bali, which has been called the island of the gods, is one of the most captivating and mesmerizing place to visit. There are many things that you can do in Bali such as: surfing, snorkeling, diving, walking on the endless paddy fields, visiting ancient temples and the monkey forest, witnessing sunrise and sunset at the beautiful beach, etc.

In 2016, there are over 12 million of foreign tourists that visits Bali with average stay of 7.5 nights. This becomes a very lucrative place to open up a business.

This report was compiled by the request of an investor who plans to open a Restaurant and Bar in Bali. I am very excited and am going to use the knowledge that I have learnt from IBM-Coursera to help this investor find the most legitimate place for his new bussiness.

1.2 Business Problem

As per the investor request, the location must be in prime area and preferably located near the beach as he and his family are very fond of surfing. So, I will only use venues that are within 1km by the beach.

1.3 Target Audience

Bali is one of the hottest holiday destinations in the world. This report is prepared for investors or anyone who are looking forward to pursue food and beverages bussinesses in Bali. Hopefully, this report will provide a better clarity and help investors in making informed decisions.

2. Data Section

In Data section, I use the following data parameter:

- List of venues in each sites

In order to obtain the list of the venues, I use Foursquare APIs.

It is important to notice that there are only a few beach in Bali and they are:

- Kuta beach
- Nusa Dua beach
- Pandawa beach
- Tanjung Benoa beach

How the data will be used to solve the problem:

- Use Foursquare and geopy data to map top 100 venues in the islands of Bali.
- Use foursquare and geopy data to map the location of hotels, villas, and tourist attractions in the islands.

There are only several beaches options to choose from. Focus on each one of them and compare between each other.

3. Methodology

In this part, I will determine which location is the most suitable to open a Bar and Restaurant.

Some of the criterias that will be used are:

- Location must be near the beach.
- I will check the diversity (unique values) of the businesses of each of the area.
- The location that will be chosen is the one with the least amount of restaurant nearby but has high businesses diversity (unique values) which means the place is vibrant and lively for doing business, as it is more likely for tourist to visit the area.

The reference I use in this project are:

- Google map to locate each site and then by selecting one beach at a time.
- Foursquare venues on Bali
- Wikipedia (https://en.wikipedia.org/wiki/Tourism_in_Indonesia)

4 Data Analysis and Visualizations

4.1 Visualizations

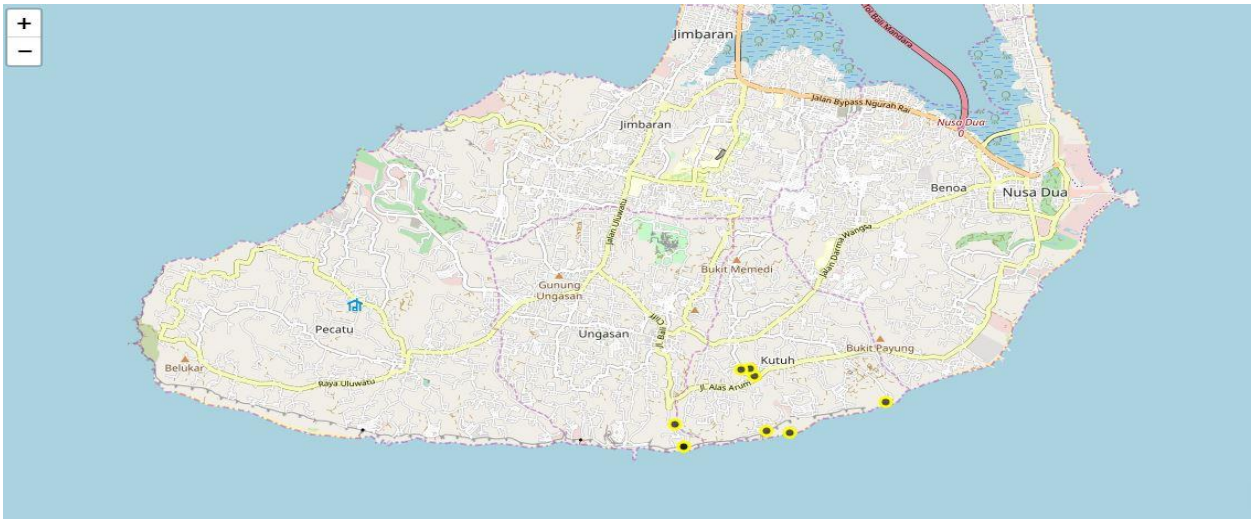
In this section, I analyze and compare each sites.

The 3 sites are:

- Kuta beach
- Nusa Dua beach
- Tanjung Benoa beach

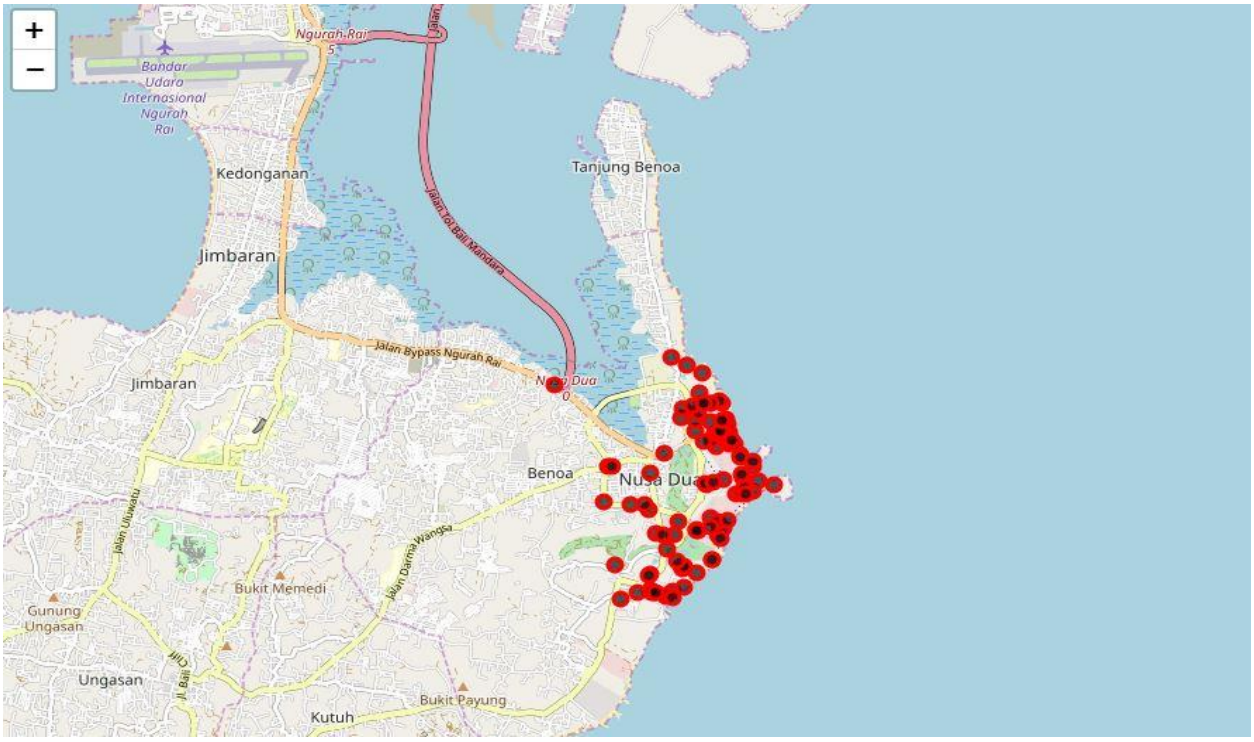
Pandawa beach was dropped from my recommendation list because there are too little venues there, perhaps the area is not too happening for tourists.

Pandawa beach. As we can see there are not much happening here in Pandawa, I proceed to drop this from my recommendation list.



(Pandawa beach)

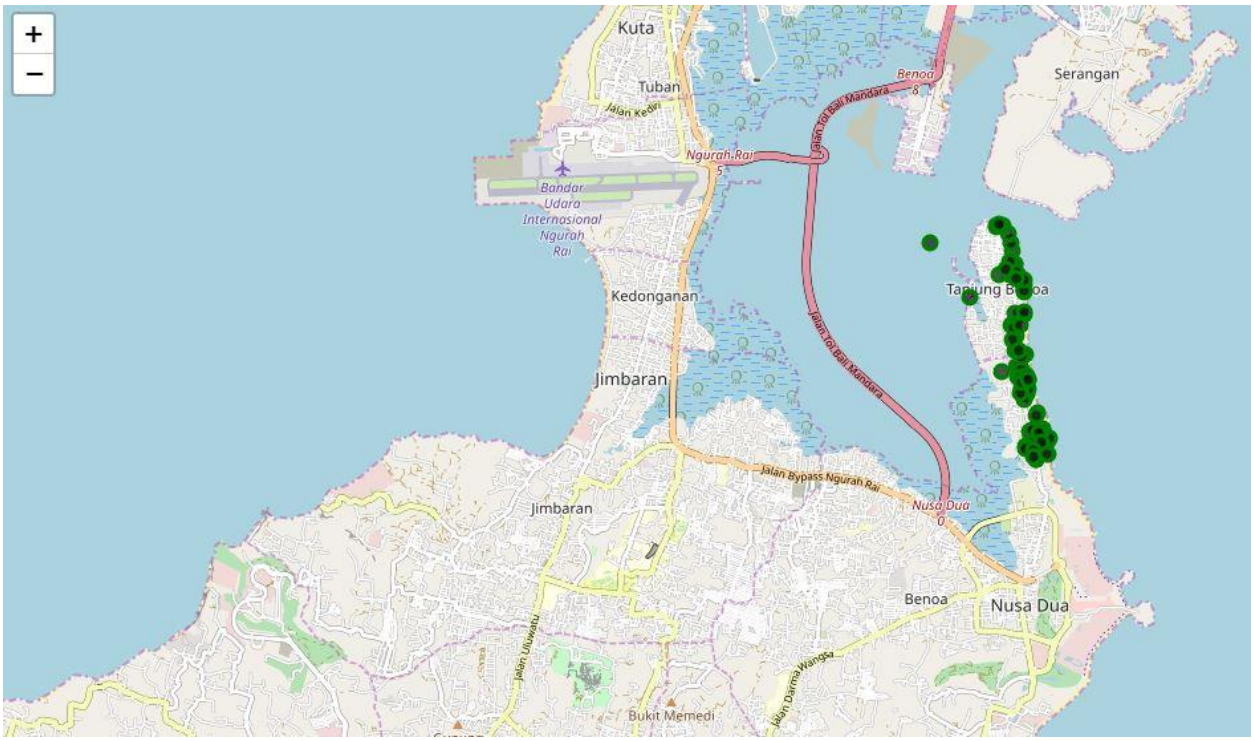
Nusa Dua beach



(Nusa Dua beach)

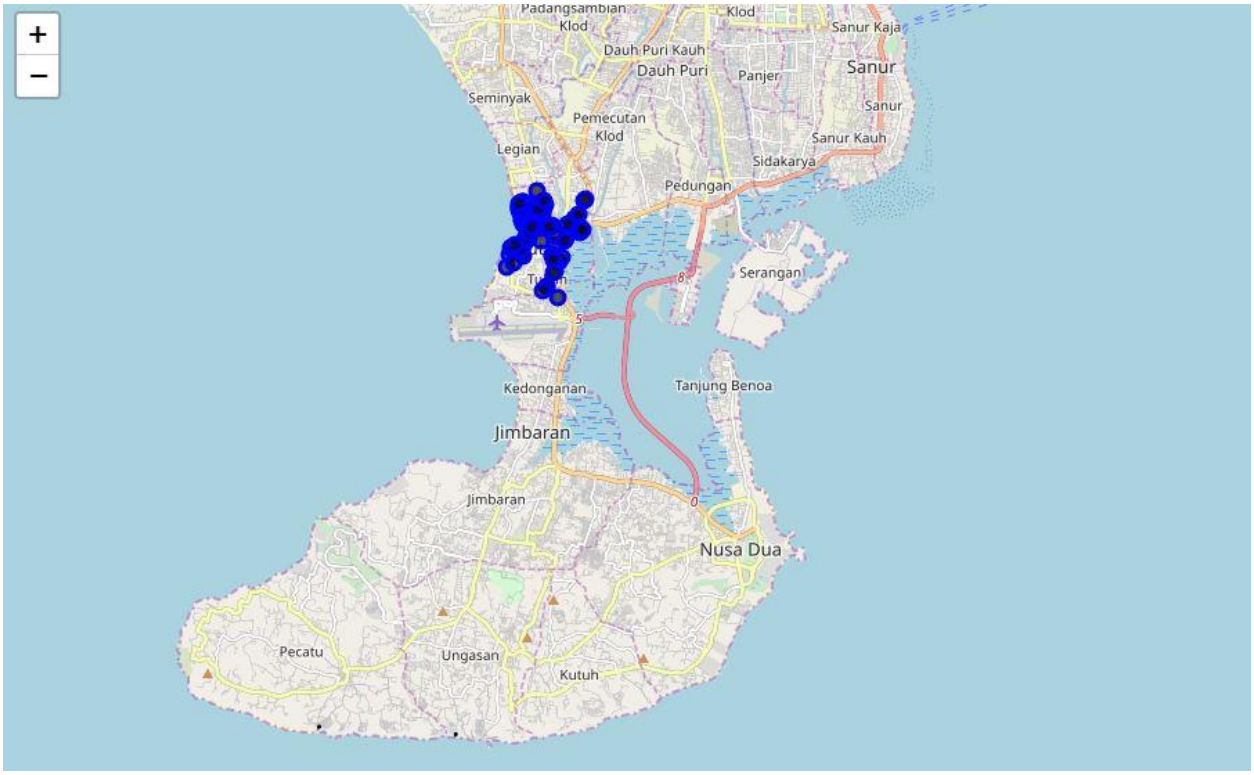
From the plot, we can see that Nusa Dua and Tanjung Benoa form a very long strip of venues. Nusa Dua is definitely one of the favorite go to place for tourists. There are 38 unique types of business in Nusa Dua generated from Foursquare.

Tanjung Benoa beach



Eventhough Tanjung Benoa is located the farthest from the airport, secluded in the far east of the Bali island, the location is definitely preferred by the tourists. As we can see there are a lot of venues clustered at the most outer skirt of the island. Another interesting fact about Tanjung Benoa despite its location is bottlenecked by Nusa Dua beach shows that it is still a favorite and vibrant place for doing business.

Kuta beach



(Kuta beach)

In terms of location, it is certainly closest to the airport and Kuta is sandwiched by 2 beachfront which make this place very lively and vibrant with businesses and tourists. This is proven by having the most number of unique values (52 different kind of business out of 100 venues from Foursquare) in our dataset. That is to say, Kuta has the most diverse type of businesses amongst the other. This is certainly a criteria required by my investor.

In [44]:
```
# 2 Check the diversity of the businesses in each sites
# lets find out how many unique categories on each sites
print('There are {} uniques categories in Kuta beach.'.format(len(BLnearby_venues['categories'].unique())))
print('There are {} uniques categories in Nusa Dua beach.'.format(len(NDnearby_venues['categories'].unique())))
print('There are {} uniques categories in Tanjung Benoa beach.'.format(len(TBnearby_venues['categories'].unique())))
```

```
There are 52 uniques categories in Kuta beach.
There are 38 uniques categories in Nusa Dua beach.
There are 37 uniques categories in Tanjung Benoa beach.
```

4.2 Data Analysis

In this section, I started by cleaning the datas and then merge all the three dataframes. Then I find the unique categories in each of the site. From here, I proceed to use get_dummies function on the data frame and calculate the means of each of the value fields.

Then, I select the only the ones that are only related (containing) to bar, food, and restaurant. The result is as follow:

```
In [49]: # Picked only food related datas
merge_all_food = merge_all.loc[[36,17,8,11,41,6,28,18,45,72,16,73,15,12,48,31,46,53,75,76,52,77,25,7,13,9,79,80,26,64,4,14,20,22,66,67,42,68]]
merge_all_food
```

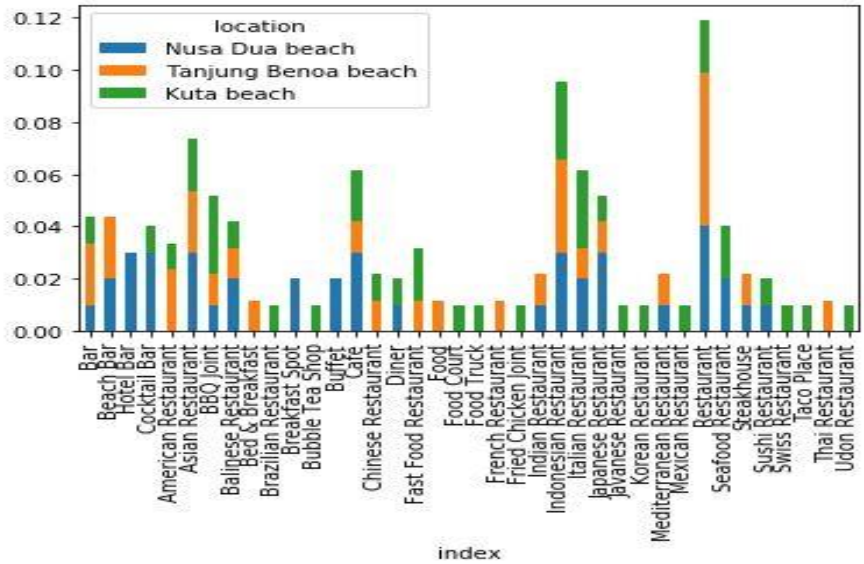
Out[49]:

location	index	Nusa Dua beach	Tanjung Benoa beach	Kuta beach
36	Bar	0.01	0.023529	0.01
17	Beach Bar	0.02	0.023529	NaN
8	Hotel Bar	0.03	NaN	NaN
11	Cocktail Bar	0.03	NaN	0.01
41	American Restaurant	NaN	0.023529	0.01
6	Asian Restaurant	0.03	0.023529	0.02
28	BBQ Joint	0.01	0.011765	0.03
18	Balinese Restaurant	0.02	0.011765	0.01
45	Bed & Breakfast	NaN	0.011765	NaN
72	Brazilian Restaurant	NaN	NaN	0.01
16	Breakfast Spot	0.02	NaN	NaN
73	Bubble Tea Shop	NaN	NaN	0.01
15	Buffet	0.02	NaN	NaN

And then I plot the dataframe above into a stacked bar graph to better understand the data. By looking at the graph below, we can make a general assumption of the overall performance of each of the beach.

```
In [52]: merge_all_food.plot.bar(x='index', stacked=True)
```

```
Out[52]: <matplotlib.axes._subplots.AxesSubplot at 0x7fcbf2927780>
```



By looking at the describe method, we can conclude that in Kuta beach there are more diversity in the types of foods and bars as the mean value is the lowest despite the higher count.

```
In [51]: # Find the most diverse type of food business in each sites
         merge_all_food.describe()
```

Out[51]:

location	Nusa Dua beach	Tanjung Benoa beach	Kuta beach
count	20.000000	19.000000	26.000000
mean	0.020500	0.018389	0.014615
std	0.009445	0.012292	0.007060
min	0.010000	0.012048	0.010000
25%	0.010000	0.012048	0.010000
50%	0.020000	0.012048	0.010000
75%	0.030000	0.024096	0.020000
max	0.040000	0.060241	0.030000

To further confirm this, I summed up all the mean value and it shows that there are 38% of businesses in Kuta beach that are related to food and bars, which means the competition rate is lower than 41% that of Nusa Dua beach. Also, notice from the bar graph that Kuta beach does not have a beach bar yet.

```
In [52]: # Sum the total value
         merge_all_food.sum()
```

```
Out[52]: location
         index                 BarBeach BarHotel BarCocktail BarDive SpotAsia...
         Nusa Dua beach                                                    0.41
         Tanjung Benoa beach                                           0.349398
         Kuta beach                                                        0.38
         dtype: object
```

5. Conclusions

In the beginning, I take 4 beahches in Bali for considerations, they are:

- Kuta beach
- Nusa Dua beach
- Tanjung Benoa beach
- Pandawa beach (dropped)

Pandawa beach analysis: I proceed to drop Pandawa beach upon pulling the datas from Foursquare because there were too few of venues there. This is an indication that the place is not on the same league as the other three.

Nusa Dua beach analysis: There are 38 unique types of businesses in Nusa Dua area. 41% of them are food related businesses. This would seem to be saturated market to open a new Bar/Restaurant in the area as it also means the competition is very high in the area.

Tanjung Benoa beach analysis: There are 37 unique types of businesses in Tanjung Benoa area. Approximately 35% of them are food related business.

Kuta beach analysis: There are 52 unique types of businesses in Kuta area, this by far is the most diverse amongst all the other beaches (many choice of restaurants). Approximately 38% of them are food related business which means there are still rooms to grow and less competition compared to Nusa Dua beach(41% of 38 unique types). Another obvious indicator is the category "Beach Bar", as we can see that Kuta beach does not currently has any beach bar yet and this is one requirement by the investor. It is clear that I should recommend Kuta area to my investor.

Winner: Kuta beach

Thank you