

# Setting a coffee shop business in Bali Using Foursquare API and unsupervised algorithm

**Mickael Gratia**

Capstone Project

IBM Data Science Professional Certificate

# 1. Introduction

- Who doesn't like coffee? Who doesn't like Bali?
- Simplicity of the business in a place called paradise
- Which location would be best to maximise the probability of success? How can we know where the competition is less?

## 2. Data acquisition and cleaning/scraping

- Bali and its districts
  - Data source: Wikipedia - [https://en.wikipedia.org/wiki/Category:Districts\\_of\\_Bali](https://en.wikipedia.org/wiki/Category:Districts_of_Bali)
- Geographical coordinates
  - Data source: Python geocoded package
- Venue data: knowing where and what type of business
  - Data source: Foursquare API

### 3. Methodology - Data collection

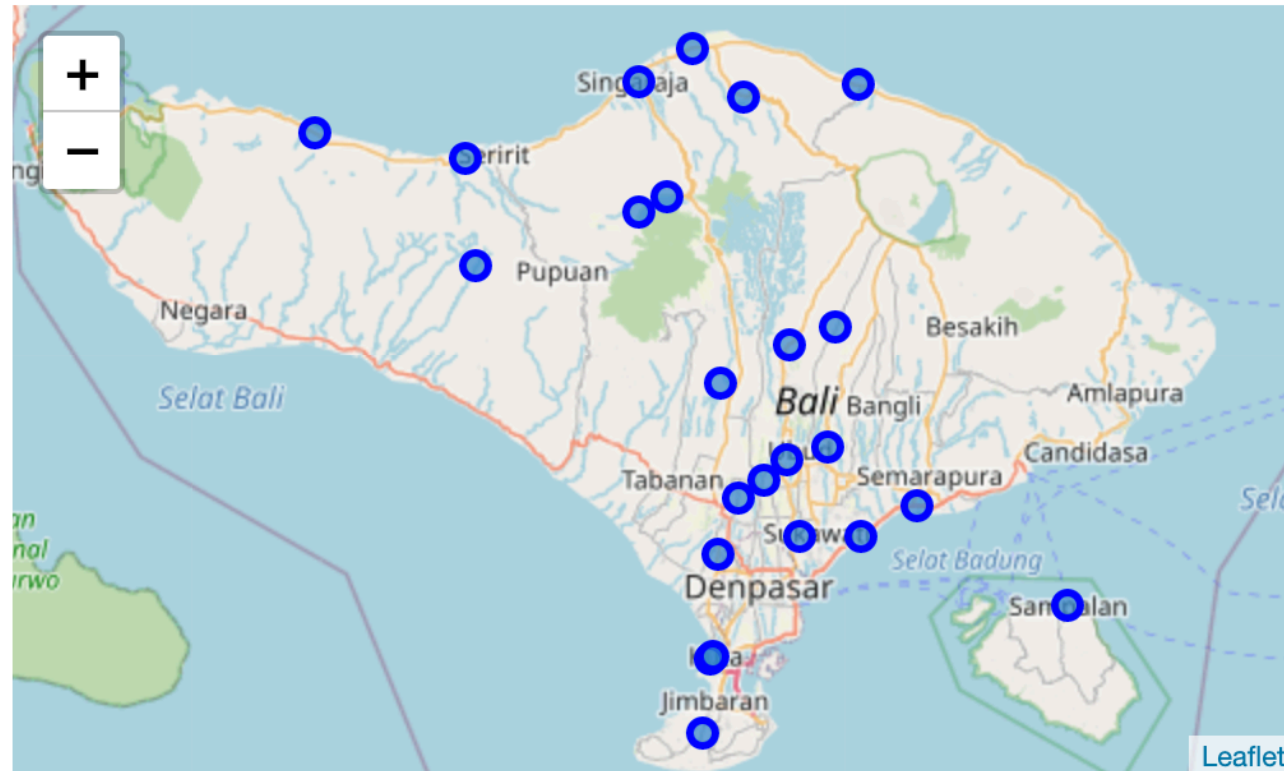
- Data scraping by collecting the list of districts
- Using Geocoder package provides coordinates

Neighborhood	
19	Sukasada, Buleleng
20	Sukawati
21	Tegallalang
22	Tejakula, Buleleng
23	Ubud District

	Neighborhood	Latitude	Longitude
0	List of districts of Bali	-8.72969	115.16812
1	Abiansemal District	-8.54097	115.22325
2	Banjar, Buleleng	-8.25522	115.09030
3	Banjarangkan	-8.56762	115.39053
4	Blahbatuh	-8.59972	115.32788
5	Buleleng, Bali	-8.11591	115.09037
6	Busung Biu, Buleleng	-8.31146	114.91358
7	Gerokgak, Buleleng	-8.17042	114.74011
8	Tampaksiring	-8.50366	115.29287

### 3. Methodology - mapping

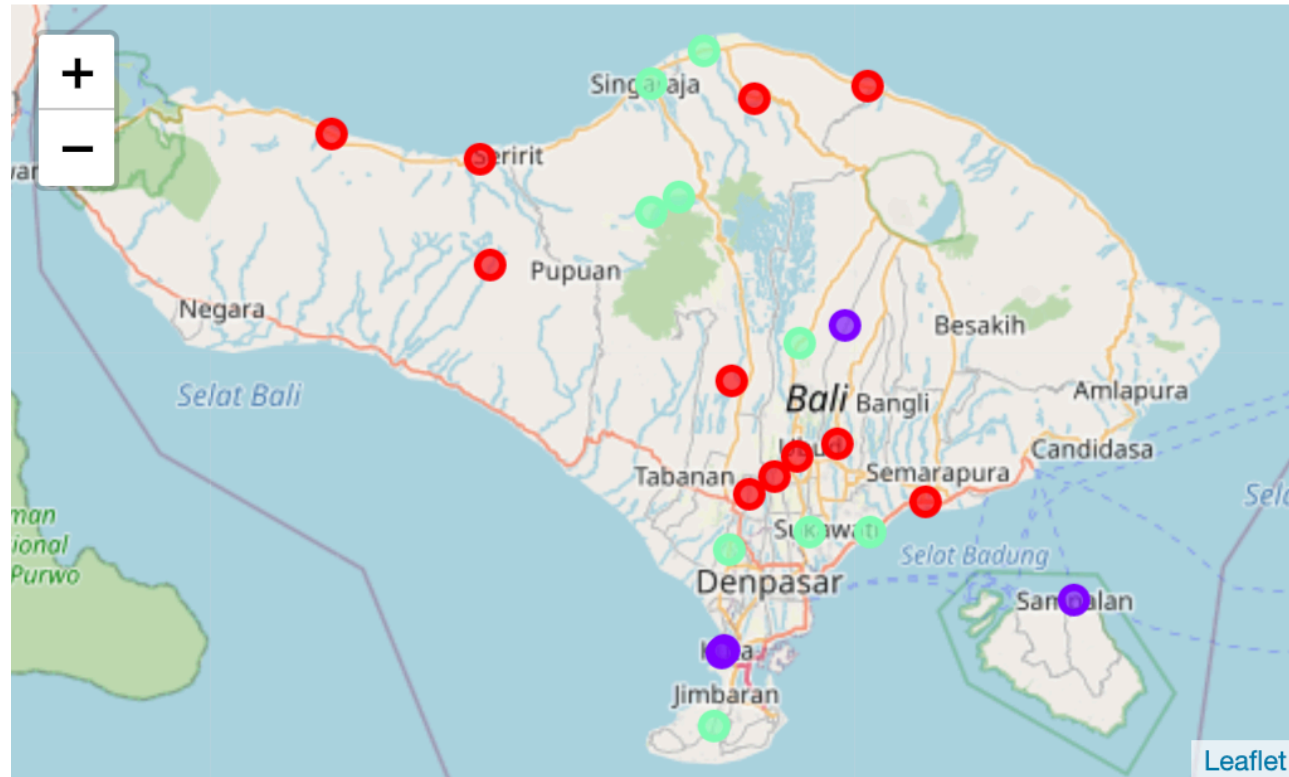
- Plotting districts - Bali key places



### 3. Methodology - plotting k-means clusters

- Coffee shop in Bali - clustering

- red cluster - Low competition
- green cluster - Medium competition
- purple cluster - High competition



# 4 Results

We can observe a high concentration of coffee shop in well known places such as Kuta or the rice fields north of Ubud. Interesting enough the island of Nusa Penida has seen a huge business increase over the past 2 to 3 years and indeed the number of coffee shops is now quite important.

Green markers refer to either expanding Denpasar area due to the new highway opened and that links the airport to the north part of the main city of Denpasar but also to the fact people are now taking houses more towards the east part of the island.

To note the north with black sand beaches of Lovina gets a medium concentration of coffee shops

Less touristic areas marked in red have a low concentration of coffee shops which makes sense. Also access to some of those location can be a bit time consuming. However interesting enough is the area of Tabanan which is marked as low in terms of competition and is a good tourist area, well known for a Unesco rice fields, temple and strawberries producer area

# 5. Conclusion

For this project we would therefore recommend the entrepreneur to go for Tabanan if the focus is on tourism. North of Denpasar otherwise due to the medium competition.