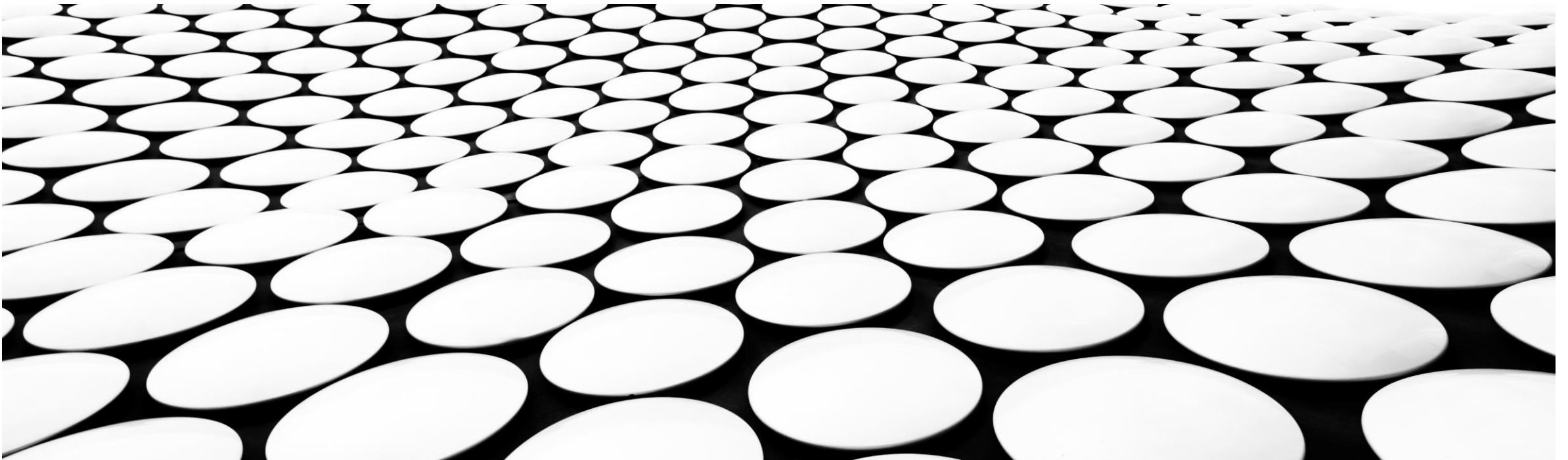# LOCATION BASED MARKET ANALYSIS TO IDENTIFY OPTIMAL BUSINESS LOCATION IN COLOMBO DISTRICT AND SUBURBS

SHAZLY SHANAWAZ

SELECTING THE OPTIMAL LOCATION IS ESSENTIAL FOR BUSINESSES

- The business gains many opportunities when placed in a location withing close proximity of the target audience

- Provides a competitive advantage and business value

- Enables the business to easily reach out to current and potential customers

- Helps manage logistics efficiently and reduce operational costs

DATA COLLECTION AND PREPROCESSING

- Initial data collected from Wikipedia page containing postal codes of Sri Lanka

- The coordinates of the neighborhoods would be extracted using the GeoPy Python package

- The venue details for the coordinates was extracted using the Foursquare API

- The frequency of the venue category was used to find out the most common venues for the model

District	City	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
Colombo	Battaramulla	6.902181	79.919578	1.0	Bakery	Fast Food Restaurant	Clothing Store	Asian Restaurant	Department Store	Food	Electronics Store	Women's Store
Colombo	Bambalapitiya	6.902486	79.854597	0.0	Hotel	Bakery	Chinese Restaurant	Dessert Shop	Coffee Shop	Restaurant	Thai Restaurant	Lingerie Store
Colombo	Wellawatte	6.874384	79.859118	0.0	Hotel	Fast Food Restaurant	Clothing Store	Seafood Restaurant	Asian Restaurant	Indian Restaurant	Café	Women's Store
Colombo	Kollupitiya	6.913526	79.850813	0.0	Sri Lankan Restaurant	Coffee Shop	Pub	Fast Food Restaurant	Hotel	Food Court	Shopping Mall	Clothing Store
Colombo	Narahenpita	6.905727	79.882130	2.0	Golf Course	IT Services	Gift Shop	Women's Store	Fast Food Restaurant	Convenience Store	Cosmetics Shop	Cricket Ground

UNDERSTANDING THE DISTRIBUTION OF VENUE DATA FOR EACH OF THE NEIGHBORHOOD

The visualizations provides insight to the distribution of the entire venue dataset into each neighborhood
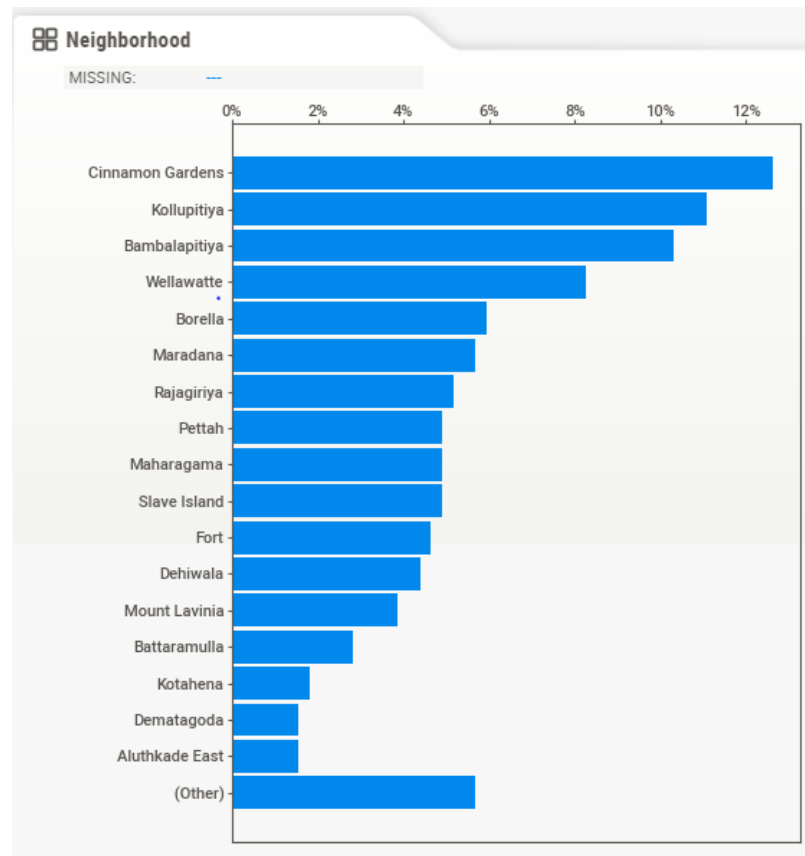
TOP CATEGORIES		
Cinnamon Gardens	49	13%
Kollupitiya	43	11%
Bambalapitiya	40	10%
Wellawatte	32	8%
Borella	23	6%
Maradana	22	6%
Rajagiriya	20	5%
Pettah	19	5%
Maharagama	19	5%
Slave Island	19	5%
Fort	18	5%
Dehiwala	17	4%
Mount Lavinia	15	4%
Battaramulla	11	3%
Kotahena	7	2%
Dematagoda	6	2%
Aluthkade East	6	2%
(Other)	22	6%
ALL	388	100%

UNDERSTANDING THE DISTRIBUTION OF THE VENUE CATEGORY IN THE DATASET

		Venue Category
	MISSING:	---
21	5%	Bakery
15	4%	Clothing Store
14	4%	CafÃ©
14	4%	Asian Restaurant
14	4%	Hotel
13	3%	Restaurant
12	3%	Coffee Shop
11	3%	Chinese Restaurant
10	3%	Fast Food Restaurant
9	2%	Bus Station
9	2%	Pizza Place
8	2%	Women's Store
8	2%	Train Station
8	2%	Bookstore
7	2%	Department Store
7	2%	Indian Restaurant
7	2%	Cosmetics Shop
6	2%	Seafood Restaurant
6	2%	Vegetarian / Vegan Restaurant
6	2%	Supermarket
6	2%	Gym
6	2%	Platform
6	2%	Bar
5	1%	Sri Lankan Restaurant
5	1%	IT Services
5	1%	Convenience Store
5	1%	Pub
5	1%	Shopping Mall
5	1%	Movie Theater
4	1%	Food Court
4	1%	Men's Store
4	1%	Cricket Ground
4	1%	Dessert Shop
4	1%	Market
4	1%	Theater
3	1%	Bubble Tea Shop
3	1%	Electronics Store
3	1%	Casino
3	1%	Pool
3	1%	Italian Restaurant
3	1%	Office
3	1%	Thai Restaurant
3	1%	Boutique
3	1%	Nightclub
2	1%	Multiplex
2	1%	Juice Bar

THE VISUALIZATIONS PROVIDES INSIGHT TO THE MOST COMMON VENUES IN EACH NEIGHBORHOOD

ALSO, IT WAS IDENTIFIED THAT A CONSIDERABLE PORTION OF THE VENUE CATEGORIES WAS LABELLED UNDER "OTHER" WHICH COULD INCLUDE MANY TYPES OF BUSINESSES NOT DEFINED IN THE CATEGORY LIST.

K- MEANS CLUSTERING MODEL RESULT

THE FOLLOWING CLUSTER RESULT WAS
OBTAINED FROM THE K – MEANS
CLUSTERING MODEL AND WAS
VISUALIZED USING FOLIUM
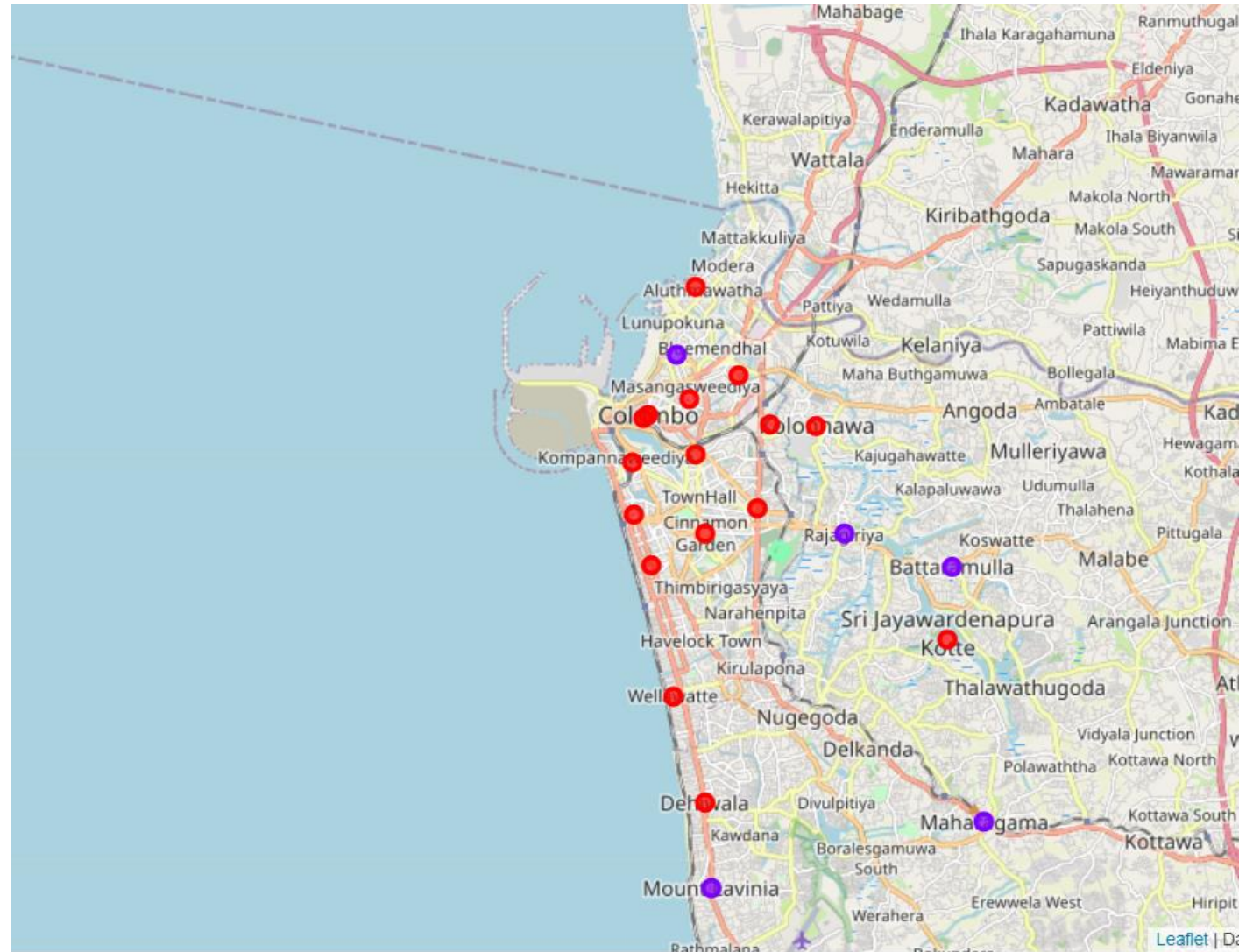
CLUSTER SEGMENTATION

- The first cluster consists of the consumer target market with almost all businesses targeted at residential consumers, with the most common venues being restaurants, coffee shops, supermarkets, consumer stores and other recreational venues for customers. This cluster group is best suited for businesses looking to sell their end product or services to everyday consumers.

- The second cluster consists of a combination of consumer market and business market with restaurants, super markets and also IT companies, electronic stores and other service companies, with a mix of businesses which cater to consumers and also businesses which cater to business users and industries. These neighborhoods are optimal if the business looks to sell its products or services to both target markets

- The third cluster consists of the business and industrial target market where most of the businesses cater towards other businesses or industries, and have relatively a smaller number of consumer-focused businesses compared to the other clusters