IBM Applied Data Science Capstone Project

The Battle of Neighborhoods

Authored by Varianto Angga

Image Shown Is For Illustration Purpose Only

Introduction

- Surabaya is the 2nd largest city in Indonesia and is the capital of East Java province

- Surabaya has a fast-growing economy, marked by the rapid development of infrastructures and venues, as well as a growing number of businesses operating in the city

- Surabaya has various venues for travelers to explore, and this opportunity has attracted the attention of investors and businessmen to open hotels, especially luxury ones

- This project aims to facilitate exploration of venues in each neighborhoods in the city of Surabaya, with final outcome of determining the best possible spot for establishing a 5-star hotel

Target Audience

Hotel owners, managers, and investors interested in setting up a new 5-star hotel in the city of Surabaya

Business Problems

- Analyze the neighborhoods and venues data

- Visualize each neighborhoods as well as its total of venues in the city of Surabaya on a map

- Cluster and profile each neighborhoods into some different groups

- Determine the neighborhood(s) that is most likely be the best spot to setup a 5-star hotel in Surabaya

Data Acquisition & Cleaning

Neighborhood Information

Scraped from a [Wikipedia Page](#)

The Wikipedia page contains the neighborhood basic information, such as unique code and name in the form of HTML

Neighborhood Coordinates

[A CSV file](#) containing all coordinates of places in Indonesia

A portion of the data set will be selected for obtaining the coordinates of places in Surabaya

Venues Information

From Foursquare API

Contains a list of venues with information, such as names, categories and coordinates

Methodology

Collecting & Cleaning Data

Exploring Data

Clustering

Gather the required data from various sources

Clean up the collected data set for further analysis

Summary statistics

Visualizations (bar graph & boxplot)

Map of Surabaya
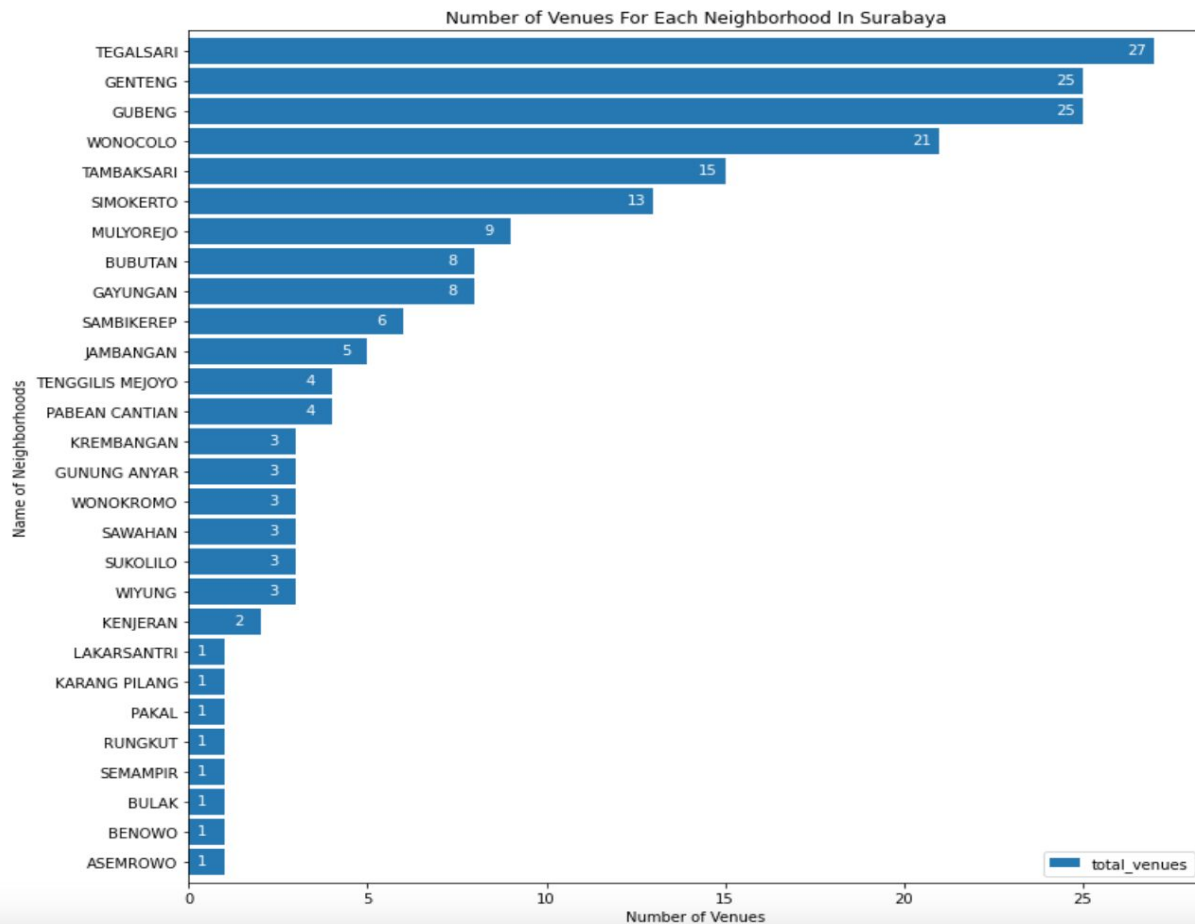
Group the neighborhoods into 5 different groups using **k-Means** algorithm

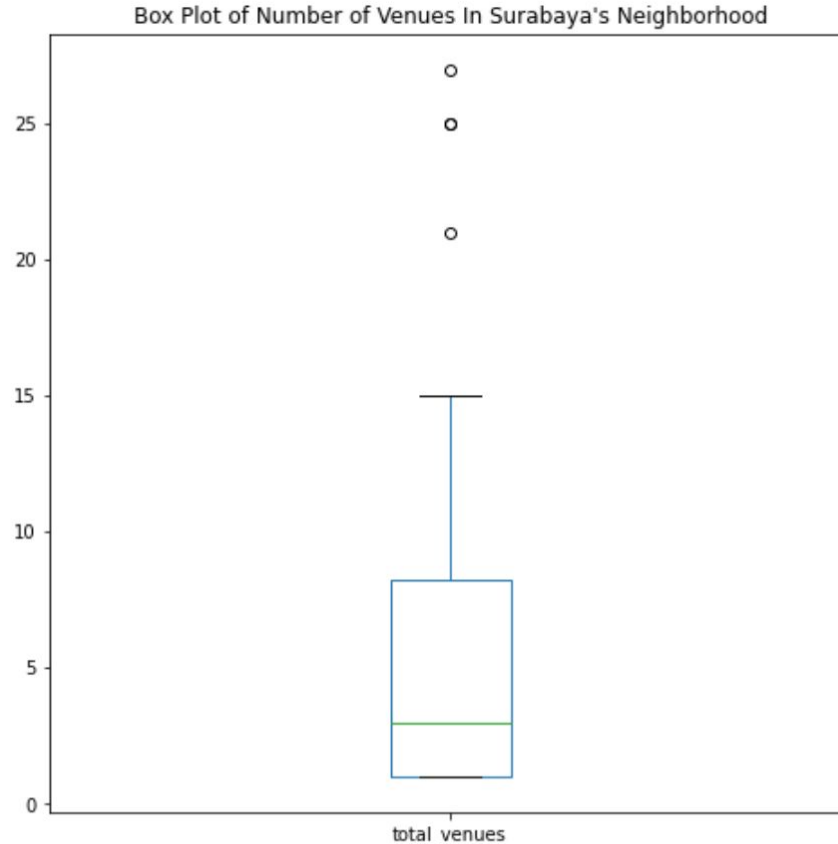# Surabaya Map With Its Neighborhoods & Venues

Bar Graph of Total Number of Venues For Each Neighborhood in Surabaya

Number of Venues For Each Neighborhood In Surabaya

Name of Neighborhoods	Number of Venues
TEGALSARI	27
GENTENG	25
GUBENG	25
WONOCOLO	21
TAMBAKSARI	15
SIMOKERTO	13
MULYOREJO	9
BUBUTAN	8
GAYUNGAN	8
SAMBIKEREP	6
JAMBANGAN	5
TENGGILIS MEJOYO	4
PABEAN CANTIAN	4
KREMBANGAN	3
GUNUNG ANYAR	3
WONOKROMO	3
SAWAHAN	3
SUKOLILO	3
WIYUNG	3
KENJERAN	2
LAKARSANTRI	1
KARANG PILANG	1
PAKAL	1
RUNGKUT	1
SEMAMPIR	1
BULAK	1
BENOWO	1
ASEMROWO	1

total_venues

Box Plot of Total Number of Venues In Surabaya's Neighborhood

Box Plot of Number of Venues In Surabaya's Neighborhood

	total_venues
count	28.000000
mean	7.071429
std	8.123713
min	1.000000
25%	1.000000
50%	3.000000
75%	8.250000
max	27.000000

Surabaya Map With Its Neighborhoods & Venues

Cluster 1 = Neighborhoods with a high number of venues

Cluster 2 = Neighborhoods with a low number of venues (Gift Shop)

Cluster 3 = Neighborhoods with a low number of venues (Toll Booth)

Cluster 4 = Neighborhoods with a low number of venues (Food Truck)

Cluster 5 = Neighborhoods with a low number of venues (Pool)

Neighborhoods in Cluster 1 Sorted By Total Number of Venues

	neighborhood_name	label	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	total_venues
23	TEGALSARI	0	Indonesian Restaurant	Bakery	Indonesian Meatball Place	Karaoke Bar	Coffee Shop	Health & Beauty Service	Bookstore	Gluten-free Restaurant	27
5	GENTENG	0	Coffee Shop	Indonesian Restaurant	Fried Chicken Joint	Soup Place	Furniture / Home Store	Bed & Breakfast	Japanese Restaurant	Multiplex	25
6	GUBENG	0	Indonesian Restaurant	Food Truck	Bakery	Chinese Restaurant	Multiplex	Kids Store	Convenience Store	Dumpling Restaurant	25
26	WONOCOLO	0	Bakery	Hotel	Pool Hall	Restaurant	Convenience Store	Coffee Shop	Café	Bubble Tea Shop	21
22	TAMBAKSARI	0	Convenience Store	Indonesian Restaurant	Asian Restaurant	Noodle House	Food Truck	Food	Bakery	Balinese Restaurant	15
20	SIMOKERTO	0	Seafood Restaurant	Food Truck	Cosmetics Shop	Arcade	Convenience Store	Soup Place	Bakery	Basketball Court	13
13	MULYOREJO	0	Indonesian Restaurant	Convenience Store	Bakery	Coffee Shop	Photography Lab	Fast Food Restaurant	Mobile Phone Shop	Supermarket	9
4	GAYUNGAN	0	Boutique	Convenience Store	Steakhouse	Café	Bakery	Food Truck	Indonesian Restaurant	Seafood Restaurant	8
2	BUBUTAN	0	Convenience Store	Bookstore	Donut Shop	Market	Fast Food Restaurant	Food Truck	Shopping Mall	Food Court	8
17	SAMBIKEREP	0	Bistro	Miscellaneous Shop	Housing Development	Market	Indonesian Restaurant	Flea Market			6

Discussion & Recommendation

Most venues are concentrated in top 23 neighborhoods from Cluster 1

Other clusters have a low number of venues

Tegalsari, Genteng, Gubeng (all from Cluster 1) are the recommended spots for establishing a new 5-star hotel in Surabaya

Conclusion

This project made use of Foursquare API to obtain the venues information

Outcomes from this project can assist the target audience in finding out the best spot to set up a 5-star hotel in Surabaya

This project also displays clustered neighborhoods on a map to facilitate exploration of the area in Surabaya

_____

Thank You

Varianto Angga

(January 2021)