

The background of the slide is a photograph of a city skyline at sunset. The sky is a mix of orange, yellow, and blue. In the foreground, there are several tall buildings, some with lights on. A large blue geometric shape, composed of several triangles, is overlaid on the top right corner of the image. The title text is centered in the upper half of the image, overlaid on a dark blue rectangular background.

# 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

Gather the required data from various sources

Clean up the collected data set for further analysis

## Exploring Data

Summary statistics

Visualizations (bar graph & boxplot)

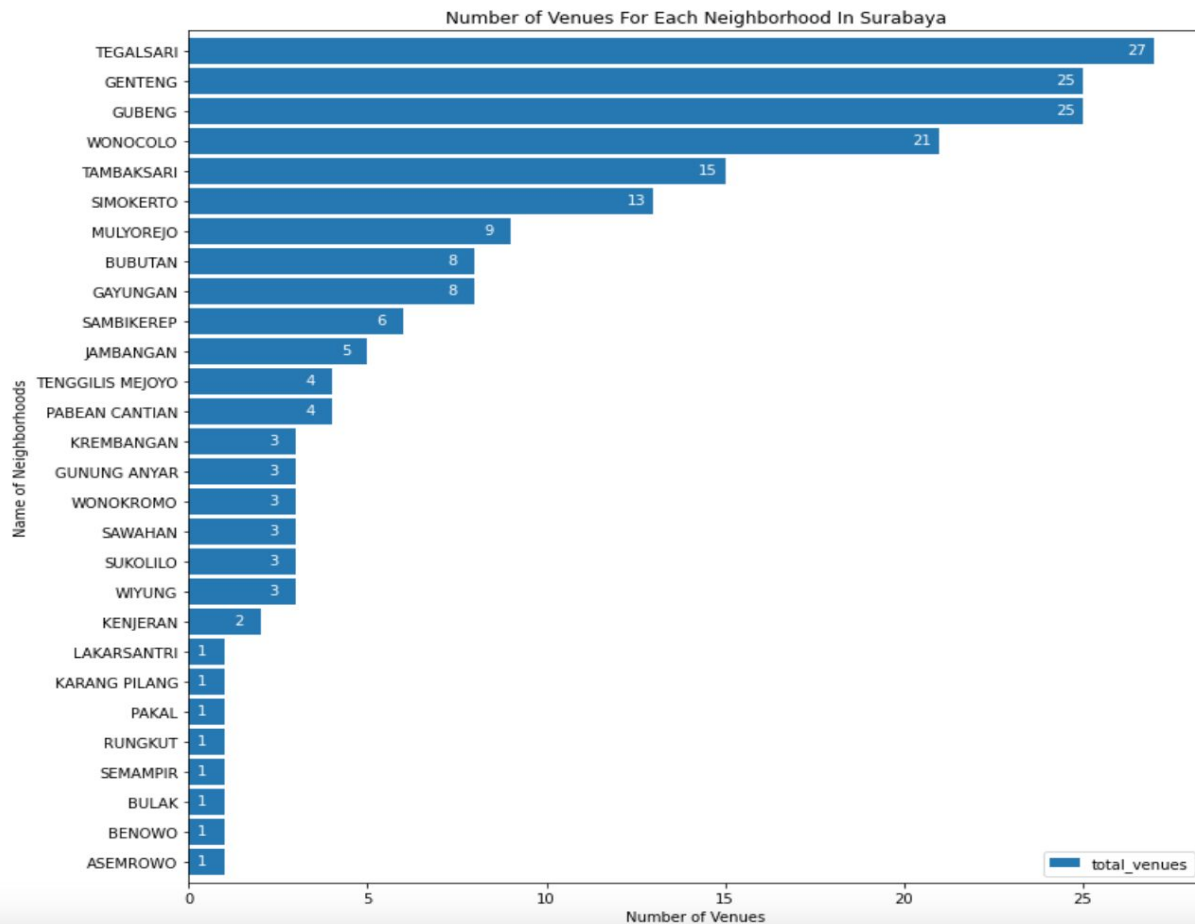
Map of Surabaya

## Clustering

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

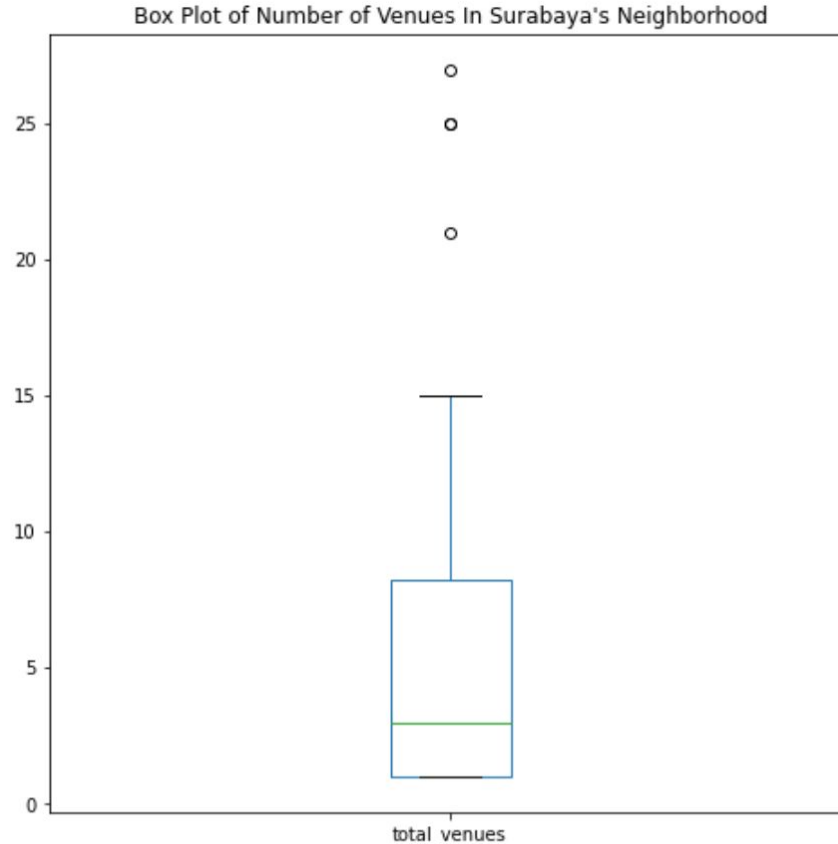


# Bar Graph of Total Number of Venues For Each Neighborhood in Surabaya





# Box Plot of Total 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

The map displays the districts of Surabaya, Indonesia, with various colored circles representing different data points. The city center is marked with a red circle labeled 'Surabaya'. Other districts shown include Kenjeran, Sukolilo, and Wonorejo. The map also includes a legend for distance ranges and a scale bar.

Legend:

- 0 - 1000
- 1000 - 2000
- 2000 - 3000
- 3000 - 4000
- 4000 - 5000
- 5000 - 6000
- 6000 - 7000
- 7000 - 8000
- 8000 - 9000
- 9000 - 10000
- 10000 - 11000
- 11000 - 12000
- 12000 - 13000
- 13000 - 14000
- 14000 - 15000
- 15000 - 16000
- 16000 - 17000
- 17000 - 18000
- 18000 - 19000
- 19000 - 20000
- 20000 - 21000
- 21000 - 22000
- 22000 - 23000
- 23000 - 24000
- 24000 - 25000
- 25000 - 26000
- 26000 - 27000
- 27000 - 28000
- 28000 - 29000
- 29000 - 30000
- 30000 - 31000
- 31000 - 32000
- 32000 - 33000
- 33000 - 34000
- 34000 - 35000
- 35000 - 36000
- 36000 - 37000
- 37000 - 38000
- 38000 - 39000
- 39000 - 40000
- 40000 - 41000
- 41000 - 42000
- 42000 - 43000
- 43000 - 44000
- 44000 - 45000
- 45000 - 46000
- 46000 - 47000
- 47000 - 48000
- 48000 - 49000
- 49000 - 50000
- 50000 - 51000
- 51000 - 52000
- 52000 - 53000
- 53000 - 54000
- 54000 - 55000
- 55000 - 56000
- 56000 - 57000
- 57000 - 58000
- 58000 - 59000
- 59000 - 60000
- 60000 - 61000
- 61000 - 62000
- 62000 - 63000
- 63000 - 64000
- 64000 - 65000
- 65000 - 66000
- 66000 - 67000
- 67000 - 68000
- 68000 - 69000
- 69000 - 70000
- 70000 - 71000
- 71000 - 72000
- 72000 - 73000
- 73000 - 74000
- 74000 - 75000
- 75000 - 76000
- 76000 - 77000
- 77000 - 78000
- 78000 - 79000
- 79000 - 80000
- 80000 - 81000
- 81000 - 82000
- 82000 - 83000
- 83000 - 84000
- 84000 - 85000
- 85000 - 86000
- 86000 - 87000
- 87000 - 88000
- 88000 - 89000
- 89000 - 90000
- 90000 - 91000
- 91000 - 92000
- 92000 - 93000
- 93000 - 94000
- 94000 - 95000
- 95000 - 96000
- 96000 - 97000
- 97000 - 98000
- 98000 - 99000
- 99000 - 100000

Scale: 0 1000 2000 3000 4000 5000 6000 7000 8000 9000 10000 11000 12000 13000 14000 15000 16000 17000 18000 19000 20000 21000 22000 23000 24000 25000 26000 27000 28000 29000 30000 31000 32000 33000 34000 35000 36000 37000 38000 39000 40000 41000 42000 43000 44000 45000 46000 47000 48000 49000 50000 51000 52000 53000 54000 55000 56000 57000 58000 59000 60000 61000 62000 63000 64000 65000 66000 67000 68000 69000 70000 71000 72000 73000 74000 75000 76000 77000 78000 79000 80000 81000 82000 83000 84000 85000 86000 87000 88000 89000 90000 91000 92000 93000 94000 95000 96000 97000 98000 99000 100000

Leaflet

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

**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 Angka

(January 2021)