IBM **SkillsBuild**

HACKTIV8 CAPSTONE PROJECT

BY : SION YEHUDA RADJA ABEDNEGO

Email
sionyehuda05@gmail.com

Website
LinkedIn

14 June 2025

About Me

A Business Management graduate with strong interest in business and data analytics.
I'm passionate about exploring trends and transforming them into meaningful insights to support decision-making, backed by quantitative reasoning, structured research, and data visualization.

Tools & Skills:

- Programming & Analytics: Python, Foursquare API, pandas, scikit-learn
- Data Platforms: Power BI, Excel, Google Colab
- Visualization & Dashboarding: Tableau, Folium, matplotlib
- Currently exploring: IBM Granite (LLM & AI-driven analytics)

TABLE OF CONTENT

PROJECT PITCH

01 Background

02 Business Problem

03 Target Audiences

04 Data Used

05 Tools Used

06 Extraction Methodology

07 Findings

08 Project Decision

HACKTIV8

IBM SkillsBuild

BACKGROUND

BUILD WHERE IT'S CROWDED— OR WHERE IT'S NEEDED?

- **Singapore** = Asia's Shopping Capital
- Malls are not just stores — they're food, leisure & social destinations.
- **Retail Saturation**
  → Orchard, Marina, Bugis: overcrowded with malls, intense competition.
  → Suburbs are Rising
- Areas like Punggol & Sembawang are attracting footfall with:
  → New housing
  → MRT lines
  → Community-driven planning
- **The Challenge** : Where should the next shopping mall be built?
- **Why It's Critical**
  → Mall development = high-cost + long-term.
  → A wrong location = years of underperformance.
- **Project's Mission** : Leverage data + machine learning to guide mall site decisions — not guesswork.

HACKTIV8

IBM **SkillsBuild**

BUSINESS PROBLEM

WHERE IN SINGAPORE SHOULD A NEW SHOPPING MALL BE OPENED TO MAXIMIZE RETURNS AND MINIMIZE COMPETITION RISK?

- Singapore’s urban density makes land scarce and development expensive
- Many districts are already saturated with malls
- Poor site selection = high vacancy, low returns, long-term risk
- Traditional location decisions based on intuition or legacy patterns are no longer reliable

HACKTIV8

IBM SkillsBuild

IBM SkillsBuild

TARGET AUDIENCES

WHO NEEDS THIS INSIGHT?

WHO BENEFITS FROM SMARTER MALL LOCATION DECISIONS?

- Property Developers
  → Identify low-saturation, high-opportunity zones
- Retail Investors & REITs
  → Back investments with data, not just sentiment
- Urban Planners & Authorities
  → Guide growth and prevent retail oversaturation
- Retail Tenants & Franchises
  → Find untapped markets to anchor their brands

DATA USED IN THIS PROJECT

What data did we use to find where malls are truly needed in Singapore?

16 Place Categories from Wikipedia

Selected diverse locations (e.g. Orchard, Punggol, Sentosa) to represent Singapore's urban zones

Geolocation via Geopy & Geocoder

Converted each location into latitude/longitude coordinates for spatial analysis

Venue Data from Foursquare API

Queried venues within 1 km radius, 50 venues per place Retrieved name, location, and category for each

Filtered for "Shopping Mall" Venues

Retained only shopping mails for comparison of retail presence

Computed Mall Frequency

Mall Frequency (%)
=
(# of Shopping Malls / 50) x 100

This normalized metric enabled fair compari-son acros locations

HACKTIV8

IBM **SkillsBuild**

Tools

HACKTIV8

IBM **SkillsBuild**

ANALYSIS

What tools I use?

🐍 **Python as the core language**
→ Entire analysis coded in Python using open-source libraries

🌐 **Data Collection**
→ Requests, BeautifulSoup for scraping Wikipedia

📍 **Geocoding**
→ Geopy, geocoder to convert place names to coordinates

☁️ **API Integration**
→ Foursquare API + requests for venue data

🧹 **Data Wrangling & Cleaning**
→ Pandas, numpy to format and transform data

🔢 **Clustering Analysis**
→ Scikit-learn (KMeans) for machine learning-based segmentation

🗺️ **Visualization**
→ Folium for maps, matplotlib for plots

💻 **Platform**
→ Developed in Google Colab, published via GitHub

HACKTIV8

IBM **SkillsBuild**

Extraction Methodology

How did I extract and analyze pattern?

IBM **SkillsBuild**

📚 Defined 16 key areas
Based on Wikipedia place categories (e.g., Orchard, Punggol, Sentosa)

📍 Geocoded all locations
Used geopy and geocoder to get latitude/longitude

📡 Queried Foursquare API
Retrieved 50 venues per location within 1 km radius

🛍️ Filtered for shopping malls only
Cleaned names & calculated Mall Frequency (%)

🧠 Applied K-Means clustering (k=3)
- Labeled areas as:
- Cluster 2: Saturated
- Cluster 0: Competitive
- Cluster 1: Underserved

🗺️ Mapped clusters using Folium
Color-coded interactive map for strategic visualization

IBM **SkillsBuild**

Findings

SINGAPORE'S

MALL LANDSCAPE CLUSTERING

DATA REVEAL??

TURN TO NEXT PAGE >>>

Cluster 1 – Underserved Residential Zones

→ E.g. Punggol, Sembawang, Bukit Panjang
→ 0–1% mall presence, high housing growth
→ MRT-connected, but lacking retail infrastructure

Cluster 0 – Competitive Heartlands

→ E.g. Toa Payoh, Bishan, Tampines
→ 2–3% mall frequency, existing community loyalty
→ Success depends on unique or themed concepts

Cluster 2 – Saturated Urban Core

→ E.g. Orchard, Marina Bay, Bugis
→ 4–6% mall density, premium but overbuilt
→ Needs radical innovation to justify new entry

HACKTIV8

IBM **SkillsBuild**

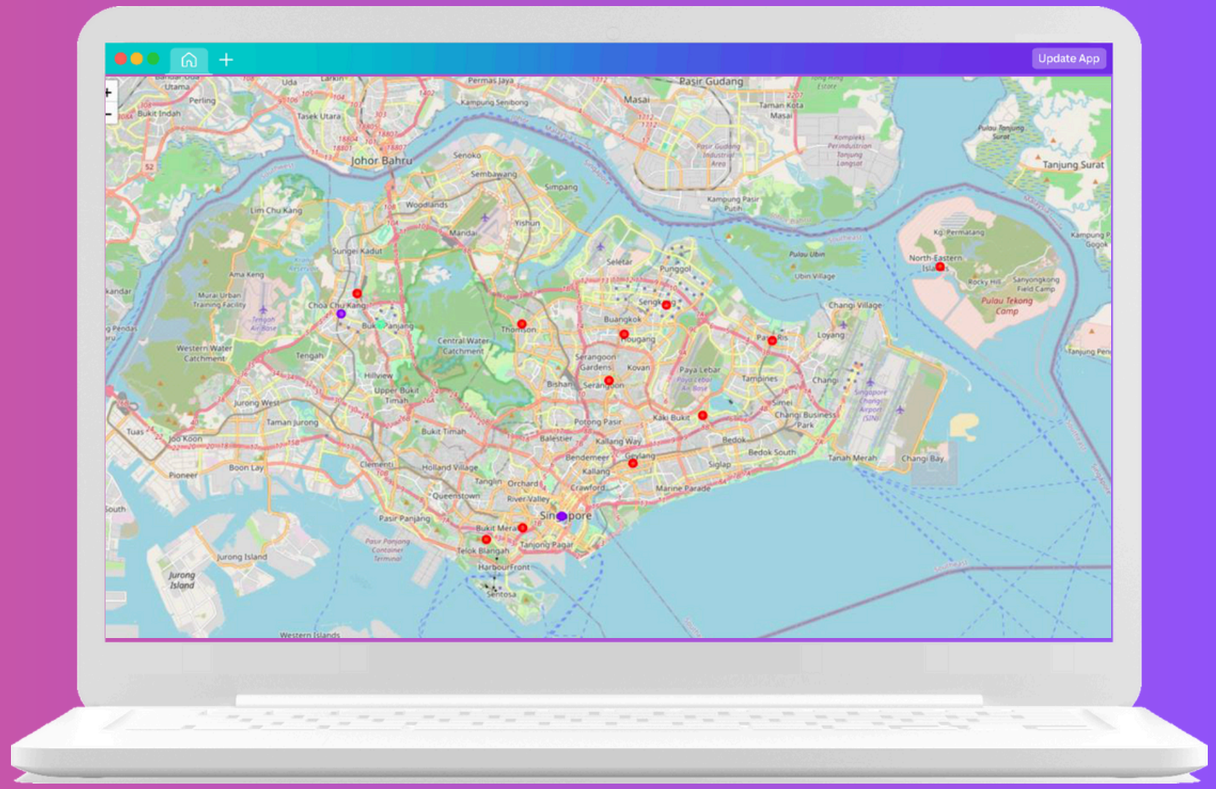
Decisions!!

HACKTIV8

IBM **SkillsBuild**

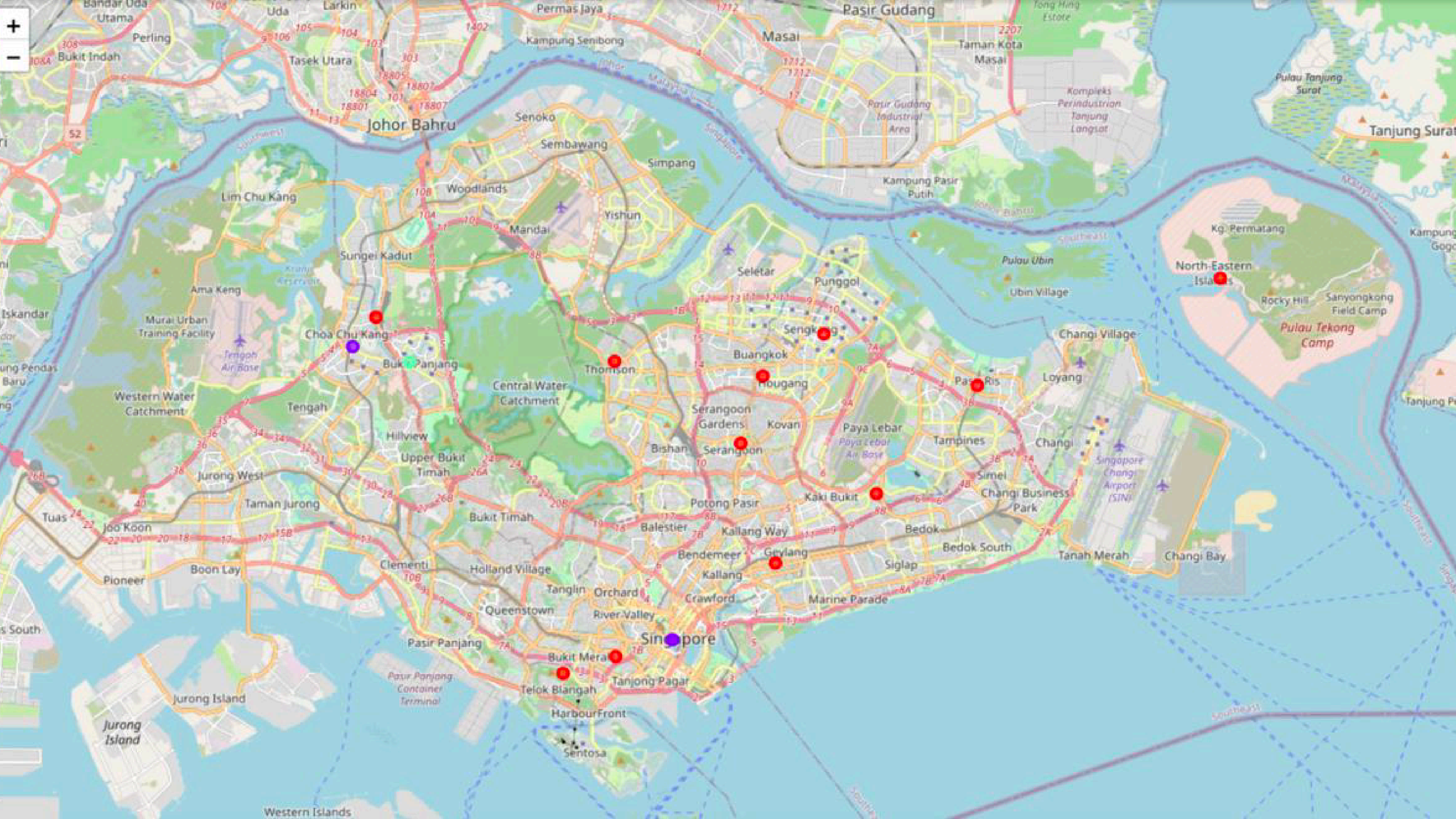
WHERE TO **INVEST**, WHERE TO **AVOID**

Zone Type	Example Areas	Decision
🟣 Underserved Growth	Punggol, Sembawang, Pasir Ris	✅ Build – High potential, low risk
🔴 Competitive Heartlands	Toa Payoh, Tampines, Serangoon	⚠️ Build only if concept is differentiated
🟢 Saturated Core	Orchard, Bugis, Marina Bay	❌ Avoid – Unless radically innovative



HACKTIV8

IBM **SkillsBuild**


Johor Bahru
Singapore
Pasir Gudang
Masai
Taman Kota Masai
Tanjung Surat
Pulau Tanjung Surat
Kompleks Perindustrian Tanjung Langsat
Pasir Gudang Industrial Area
Kampung Pasir Putih
Kampung Senibong
Permas Jaya
Bandar Uda Utama
Perling
Bukit Indah
Tasek Utara
Uda
Larkin
Senoko
Sembawang
Simpang
Woodlands
Yishun
Mandai
Lim Chu Kang
Sungei Kadut
Ama Keng
Murai Urban Training Facility
Tengah Air Base
Choa Chu Kang
Bukit Panjang
Western Water Catchment
Central Water Catchment
Thomson
Seletar
Punggol
Sengkang
Buangkok
Hougang
Serangoon Gardens
Kovan
Serangoon
Bishan
Paya Lebar
Paya Lebar Air Base
Tampines
Changi
Changi Village
Loyang
Pasir Ris
Pulau Ubin
Ubin Village
Kg Permatang
North-Eastern Islands
Rocky Hill
Sanyongkong Field Camp
Pulau Tekong Camp
Tanjung Pe
Hillview
Upper Bukit Timah
Jurong West
Taman Jurong
Tuas
Joo Koon
Pioneer
Boon Lay
Clementi
Bukit Timah
Holland Village
Tanglin
Orchard
Queenstown
River Valley
Balestier
Potong Pasir
Kaki Bukit
Simei
Changi Business Park
Kallang Way
Bendemeer
Geylang
Kallang
Crawford
Marine Parade
Siglap
Bedok
Bedok South
Tanah Merah
Changi Bay
Singapore Changi Airport (SIN)
Pasir Panjang
Bukit Merah
Tanjong Pagar
Telok Blangah
HarbourFront
Sentosa
Jurong Island
Pasir Panjang Container Terminal
Western Islands
Tengah Air Base

This strategic recommendation is based on data, not speculation —and it supports an urban decentralization approach aligned with Singapore's long-term development goals.

IBM **SkillsBuild**

THANK YOU

FOR YOUR NICE ATTENTION

Email
sionyehuda05@gmail.com

Website
LinkedIn

14 June 2025