**STRATEGIC**

**BRANCH**

**ESTABLISHMENT:**

Establishing a New Coffee Shop

Branch in Yogyakarta, Indonesia

**Author :** Kevin Elfri Yodia Shaputra

**Date Created :** June, 17th 2021

**Contact :** [kevinelfri@ymail.com](mailto:kevinelfri@ymail.com)

**Preface**

Indonesia and coffee has close relationship among each other, where as Indonesia is holding the title of fourth largest coffee producer in the world and Indonesia hold the seventh position title as the world largest coffee consumer bringing supply and demand for coffee in the country has important role on its economy’s growth. In other words, both of the facts tell us explicitly that coffee has great value to Indonesian businesses and economy in many ways of business point of view.

Coffee as business idea yielding vast amount of business type, broad of product variety and business model, scalable business valuation and/or organization size, and vast target consumer and/or market. In fact, Indonesian Ministry of Industry release a publication that said even though amid coronavirus global pandemic impact, coffee related business trade balance still made a surplus of USD 211.05 million in 2020 first semester.

Indonesia in the era of digital transformation has impact on shaping business behavior due to people higher literation on technology. Businesses are likely to thrive in when using agile and intelligent strategy on their decision making processes. One of the most case of making decision is ‘where to establish the business operation?’, based on the case stated in prior sentence this publication will elaborate and demonstrate the process of decision making based on data.

**Chapter I**

**Introduction**

* 1. **Background**

Establishing a business capital is not an easy task to do for every stakeholders in any business size and there are lot of things to consider and decision to be made in this process. One of the major problem arise from is ‘where do the place that suit best?’.

Choosing a business spot is an initial step to businesses operation apart from concepting business model and business operation planning. When choosing an establishment spot we might consider potential benefit or potential risk in the future.

* 1. **Business Problem**

This research is an attempt to find a strategic place to establish a coffee shop within Yogyakarta, Indonesia. This section will discuss the business problems that will shape the research direction and scope, therefore below will point out the business problem that occurred.

1. Want to establish a coffee shop somewhere in Yogyakarta
2. Provide establishment location option in village level
3. Preferably presented in geological map
4. Target customers are college university students and office workers
5. Success criteria : locations with lowest competitors and highest potential customers
6. Based on domain knowledge, working place and college university mainly available in two main regency : Sleman and Kota Yogyakarta
7. Cannot afford third party data vendor
8. The data required might be scattered somewhere in open public
9. To reach highest potential customers, the location should be near to working places (office) and college universities
10. To reach lowest competitors potential, the location should have least coffee shop nearby

Analytics approach conducted in this research:

1. Geospatial analysis and visualization
2. Unsupervised task: Clustering

**Chapter II**

**Data Definition**

**2.1 Data Understanding**

Based on the business problem stated in prior section; location point of interest in this research will be the regions of Yogyakarta and its subdistrict and village. Therefore to generate geo visualization format, geo coordinate information is needed such as latitude and longitude of a particular location. Since no data provided at the first time, external sources needed to gather all data requirements. One of the external sources for data gathering is foursquare places API an open-source with the sandbox free tier services and Nominatim openstreetmap geocoders. Data requested from the API using version updated on July, 1st 2020.

**2.2 Data Requirements**

Data required to conduct research:

1. Yogyakarta’s subdistrics name
2. Yogyakarta’s villages
3. College universities in Yogyakarta
4. Office buildings in Yogyakarta
5. Coffee shop establishments in Yogyakarta
6. Geo location (latitude,longitude) from above data

**2.3 Data Collection**

Web scraping sources:

1. <https://id.wikipedia.org/wiki/Daftar_kapanewon,_kemantren,_kalurahan,_dan_kelurahan_di_Daerah_Istimewa_Yogyakarta>
2. <https://kependudukan.jogjaprov.go.id/statistik/penduduk/golonganusia/15/produktif/3/71/34.ez>
3. <https://kependudukan.jogjaprov.go.id/statistik/penduduk/golonganusia/15/produktif/07/71/34.ez>
4. <https://kependudukan.jogjaprov.go.id/statistik/penduduk/golonganusia/15/produktif/8/04/34.ez>
5. <https://kependudukan.jogjaprov.go.id/statistik/penduduk/golonganusia/15/produktif/6/71/34.ez>
6. <https://kependudukan.jogjaprov.go.id/statistik/penduduk/golonganusia/15/produktif/01/71/34.ez>

data from web scraping source containing informations of Yogyakarta subdistrics, villages name, and number of population data.

Nominatim-Openstreetmap requests:

1. Using geopy.Nominatim python package to get geo coordinate of a location. With this python package, latitude and longitude of Yogyakarta’s subdistricts and villages are retrieved.

Foursquare API requests:

1. <https://api.foursquare.com/v2/venues/explore>

parameters:

* Client id
* Client secret
* Radius = 700 meters
* ll (latitude, longitude)
* version = 20200701
* limit =100
* q (query) = College%20%26%20University
* categoryId (college places) = 4d4b7105d754a06372d81259
* categoryId (office places) = 4bf58dd8d48988d124941735
* categoryId (coffee shop venues) = 4bf58dd8d48988d1e0931735

with foursquare API request, data containing coffee shop venues, college universities, and office buildings are retrieved. The format the data retrieved is JSON (Javascript Object Notation) containing geo coordinate of a venues, list of nearby venues, and other informations.

Openstreetmap Geojson:

1. Yogyakarta Province Geojson

Geojson used for showing boundary of a particular location

**2.4 Data Quality**

**a. Actuality**

- Population data are taken with the update on 2020 2nd semester or June 2020

- Foursquare API places with the update on July, 1st 2020

**b. Factuality**

- Population data are taken from official Yogyakarta government sites

- Places date are taken from Foursquare places endpoint

- Geo coordinate are taken from Nominatim Openstreetmap opensource map provider

**Chapter III**

**Methodology**

**3.1 Data Preparation**

**Source: webscraping**

|  |  |  |
| --- | --- | --- |
| **Raw Data** | **Feature Engineering/Action Taken** | **Result Data** |
| **Source:**  **webscraping urls [b-f]**   * Not productive age yet * Productive age * Not productive age * Number of woman within categories per village * Number of man within categories per village * Grand total woman within categories per subdistricts * Grand total man within categories per subdistricts * Grand total man/woman within villages per village * Village name categories | * Feature selection * Extracting information | * Productive age * Grand total man/woman within villages per village * Village name categories * Dataframe name: **top\_5\_demographics** |
| **Source:**  **webscraping urls [a]**   * Yogyakarta’s territory * Regencies table * Subdistricts table * Villages table | * Feature selection * Extracting information * String data type method * List data type manipulation * converting into pandas dataframe * concatenating two main regions dataframe * correcting wrong villages name | * Sleman and Kota Yogyakarta Region * Two main region’s subdistricts and villages * Dataframe name: **full\_regency** |

**Source : Nominatim API request**

|  |  |  |
| --- | --- | --- |
| **Source:**  **Nominatim Requests**   * **full\_regency** villages name data | * Endpoint request * Combine resulting data into **full\_regency** | * Geo coordinate (latitude,longitude) for villages data |
| **Source:**  **full\_regency** | * Recalibrating geo coordinate latitude longitude | * Correct coordinate |

**Source: Foursquare API endpoint requests**

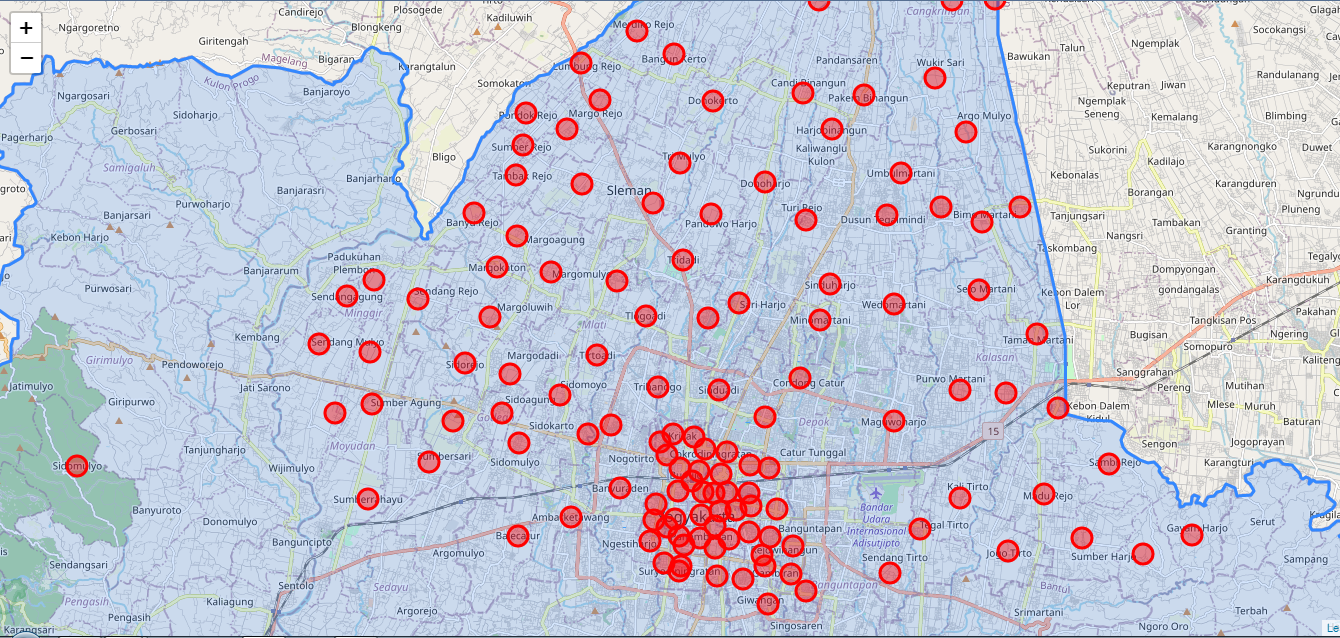
|  |  |  |
| --- | --- | --- |
| **Source:**  **Explore endpoint requests**   * **full\_regency** villages latitude longitude data   Response in JSON   * venues name * venues category * venues distance * venues lat,lon * venues unique id * venues city * venues postal * venues country * venues city * venues icon * etc | * Endpoint request * Feature selection * Aggregated sum of venues in latitude, longitude given | * Total coffee shop nearby coordinate given * Total office building nearby coordinate given * Total college university nearby coordinate given |

**3.2 Data Modeling**

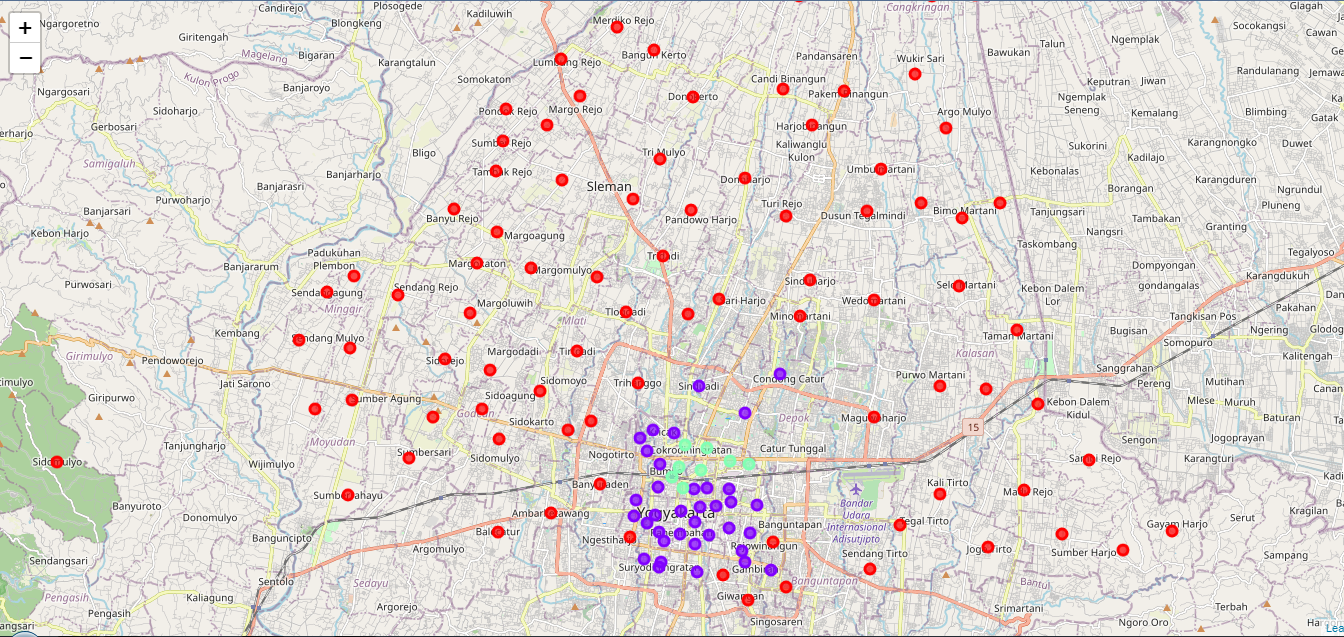
Clustering

|  |  |  |
| --- | --- | --- |
| **Source:**  **Full\_regency dataframe**   * Total number of college universities * Total\_number of office buildings * Total coffee shop venues | * Unsupervised Machine Learning Task: Clustering. Algorithm used: Kmeans * Attempt to find best K cluster using elbow method. Result : n\_cluster = 3 | * Cluster label for villages |

Spatial Visualization of Yogyakarta



Spatial Visualization of Cluster



**Chapter IV**

**Results**

**4.1 Clustering Results**

**Cluster 1**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Regency** | **Sub\_District** | **Village** | **Latitude** | **Longitude** | **Number\_College\_Venues** | **Number\_Office\_Venues** | **Number\_Coffee\_Venues** | **Cluster\_Label** |
| Sleman | Berbah | Jogotirto | -7.8129504 | 110.4720599 | 2 | 0 | 1 | 0 |
| Sleman | Berbah | Kalitirto | -7.7951322 | 110.4553582 | 4 | 0 | 0 | 0 |
| Sleman | Berbah | Sendangtirto | -7.82044215 | 110.4313663 | 8 | 0 | 2 | 0 |
| Sleman | Berbah | Tegaltirto | -7.8054225 | 110.4417249 | 13 | 0 | 0 | 0 |
| Sleman | Cangkringan | Argomulyo | -7.6706097 | 110.4574592 | 0 | 0 | 0 | 0 |
| Sleman | Cangkringan | Glagaharjo | -7.62522955 | 110.4676811 | 0 | 0 | 0 | 0 |
| Sleman | Cangkringan | Kepuharjo | -7.6254682 | 110.4525964 | 1 | 0 | 1 | 0 |
| Sleman | Cangkringan | Wukirsari | -7.6520334 | 110.4468055 | 0 | 0 | 0 | 0 |
| Sleman | Depok | Maguwoharjo | -7.7686938 | 110.4327364 | 7 | 1 | 4 | 0 |
| Sleman | Gamping | Ambarketawang | -7.8014489 | 110.3218737 | 11 | 0 | 5 | 0 |
| Sleman | Gamping | Balecatur | -7.8078411 | 110.3037379 | 8 | 1 | 1 | 0 |
| Sleman | Gamping | Banyuraden | -7.79162 | 110.3388925 | 2 | 0 | 2 | 0 |
| Sleman | Gamping | Nogotirto | -7.7702585 | 110.3358139 | 7 | 1 | 4 | 0 |
| Sleman | Gamping | Trihanggo | -7.7571187 | 110.3517176 | 12 | 1 | 5 | 0 |
| Sleman | Godean | Sidoagung | -7.76618085 | 110.298346 | 3 | 1 | 3 | 0 |
| Sleman | Godean | Sidoarum | -7.7732136 | 110.3276793 | 3 | 1 | 10 | 0 |
| Sleman | Godean | Sidokarto | -7.77627705 | 110.304159 | 11 | 0 | 6 | 0 |
| Sleman | Godean | Sidoluhur | -7.76895085 | 110.2815757 | 3 | 0 | 2 | 0 |
| Sleman | Godean | Sidomoyo | -7.7600099 | 110.3182508 | 0 | 0 | 0 | 0 |
| Sleman | Godean | Sidomulyo | -7.7839938 | 110.152508 | 0 | 0 | 0 | 0 |
| Sleman | Godean | Sidorejo | -7.7492391 | 110.2856145 | 1 | 0 | 0 | 0 |
| Sleman | Kalasan | Purwomartani | -7.75818705 | 110.4555707 | 2 | 0 | 0 | 0 |
| Sleman | Kalasan | Selomartani | -7.72417035 | 110.4619782 | 1 | 0 | 0 | 0 |
| Sleman | Kalasan | Tamanmartani | -7.7391953 | 110.4817909 | 0 | 0 | 3 | 0 |
| Sleman | Kalasan | Tirtomartani | -7.75929095 | 110.4712909 | 9 | 2 | 3 | 0 |
| Sleman | Minggir | Sendangagung | -7.726428 | 110.2451237 | 1 | 0 | 0 | 0 |
| Sleman | Minggir | Sendangarum | -7.74538965 | 110.2528654 | 0 | 0 | 0 | 0 |
| Sleman | Minggir | Sendangmulyo | -7.7426899 | 110.2353842 | 0 | 0 | 0 | 0 |
| Sleman | Minggir | Sendangrejo | -7.7272268 | 110.2694529 | 0 | 0 | 1 | 0 |
| Sleman | Minggir | Sendangsari | -7.7210165 | 110.2543945 | 0 | 0 | 1 | 0 |
| Sleman | Mlati | Sendangadi | -7.7337165 | 110.3689073 | 3 | 2 | 6 | 0 |
| Sleman | Mlati | Sumberadi | -7.7212094 | 110.3378185 | 2 | 0 | 3 | 0 |
| Sleman | Mlati | Tirtoadi | -7.7464001 | 110.3310165 | 2 | 0 | 1 | 0 |
| Sleman | Mlati | Tlogoadi | -7.7329816 | 110.3475649 | 4 | 0 | 3 | 0 |
| Sleman | Moyudan | Sumberagung | -7.7629521 | 110.2536653 | 5 | 0 | 7 | 0 |
| Sleman | Moyudan | Sumberarum | -7.7661535 | 110.2408898 | 3 | 0 | 1 | 0 |
| Sleman | Moyudan | Sumberrahayu | -7.795372 | 110.2523454 | 0 | 0 | 0 | 0 |
| Sleman | Moyudan | Sumbersari | -7.78265475 | 110.2732611 | 0 | 0 | 0 | 0 |
| Sleman | Ngaglik | Donoharjo | -7.6875954 | 110.3886959 | 4 | 0 | 2 | 0 |
| Sleman | Ngaglik | Minomartani | -7.7345297 | 110.4075393 | 3 | 2 | 2 | 0 |
| Sleman | Ngaglik | Sardonoharjo | -7.7005804 | 110.4026509 | 1 | 0 | 4 | 0 |
| Sleman | Ngaglik | Sariharjo | -7.7287731 | 110.3798034 | 4 | 3 | 6 | 0 |
| Sleman | Ngaglik | Sinduharjo | -7.7221769 | 110.4109855 | 3 | 1 | 0 | 0 |
| Sleman | Ngaglik | Sukoharjo | -7.6986524 | 110.4302895 | 10 | 3 | 8 | 0 |
| Sleman | Ngemplak | Bimomartani | -7.70115725 | 110.463005 | 5 | 0 | 2 | 0 |
| Sleman | Ngemplak | Sindumartani | -7.695977 | 110.4759792 | 0 | 0 | 2 | 0 |
| Sleman | Ngemplak | Umbulmartani | -7.68444 | 110.4352525 | 1 | 1 | 3 | 0 |
| Sleman | Ngemplak | Wedomartani | -7.7291658 | 110.4328328 | 6 | 0 | 0 | 0 |
| Sleman | Ngemplak | Widodomartani | -7.69616745 | 110.4490018 | 1 | 1 | 2 | 0 |
| Sleman | Pakem | Candibinangun | -7.65739945 | 110.4017249 | 0 | 0 | 0 | 0 |
| Sleman | Pakem | Hargobinangun | -7.59551805 | 110.4307599 | 6 | 1 | 9 | 0 |
| Sleman | Pakem | Harjobinangun | -7.6694491 | 110.4114568 | 2 | 1 | 2 | 0 |
| Sleman | Pakem | Pakembinangun | -7.6579966 | 110.4224758 | 4 | 2 | 4 | 0 |
| Sleman | Pakem | Purwobinangun | -7.6255927 | 110.4072227 | 1 | 0 | 0 | 0 |
| Sleman | Prambanan | Bokoharjo | -7.76428525 | 110.4892454 | 10 | 0 | 2 | 0 |
| Sleman | Prambanan | Gayamharjo | -7.80754245 | 110.5353072 | 0 | 0 | 0 | 0 |
| Sleman | Prambanan | Madurejo | -7.79371975 | 110.484334 | 4 | 0 | 2 | 0 |
| Sleman | Prambanan | Sambirejo | -7.7835024 | 110.5066974 | 0 | 0 | 2 | 0 |
| Sleman | Prambanan | Sumberharjo | -7.80863675 | 110.4973449 | 0 | 0 | 2 | 0 |
| Sleman | Prambanan | Wukirharjo | -7.8142312 | 110.5181983 | 0 | 0 | 0 | 0 |
| Sleman | Seyegan | Margoagung | -7.70585145 | 110.3034038 | 3 | 0 | 1 | 0 |
| Sleman | Seyegan | Margodadi | -7.73338605 | 110.2940325 | 0 | 0 | 1 | 0 |
| Sleman | Seyegan | Margokaton | -7.7165055 | 110.296628 | 1 | 0 | 2 | 0 |
| Sleman | Seyegan | Margoluwih | -7.75283635 | 110.3009377 | 0 | 0 | 0 | 0 |
| Sleman | Seyegan | Margomulyo | -7.71811375 | 110.3152181 | 1 | 0 | 0 | 0 |
| Sleman | Sleman | Caturharjo | -7.688233 | 110.3257783 | 1 | 0 | 0 | 0 |
| Sleman | Sleman | Pandowoharjo | -7.69831385 | 110.3698756 | 3 | 0 | 0 | 0 |
| Sleman | Sleman | Tridadi | -7.7139519 | 110.3605194 | 3 | 6 | 6 | 0 |
| Sleman | Sleman | Triharjo | -7.6945594 | 110.3502065 | 5 | 4 | 4 | 0 |
| Sleman | Sleman | Trimulyo | -7.68120925 | 110.3594467 | 0 | 0 | 1 | 0 |
| Sleman | Tempel | Banyurejo | -7.69810745 | 110.2886183 | 0 | 0 | 1 | 0 |
| Sleman | Tempel | Lumbungrejo | -7.6470831 | 110.3254578 | 3 | 0 | 5 | 0 |
| Sleman | Tempel | Margorejo | -7.6594994 | 110.3319138 | 1 | 0 | 2 | 0 |
| Sleman | Tempel | Merdikorejo | -7.63608625 | 110.3447612 | 0 | 0 | 0 | 0 |
| Sleman | Tempel | Mororejo | -7.6693605 | 110.3205163 | 1 | 0 | 0 | 0 |
| Sleman | Tempel | Pondokrejo | -7.66402455 | 110.3065501 | 2 | 0 | 0 | 0 |
| Sleman | Tempel | Sumberejo | -7.67500355 | 110.3056002 | 0 | 0 | 0 | 0 |
| Sleman | Tempel | Tambakrejo | -7.68496605 | 110.3032194 | 0 | 0 | 0 | 0 |
| Sleman | Turi | Bangunkerto | -7.6440249 | 110.3573944 | 0 | 0 | 1 | 0 |
| Sleman | Turi | Donokerto | -7.6597927 | 110.3707183 | 0 | 0 | 2 | 0 |
| Sleman | Turi | Girikerto | -7.60950435 | 110.4003064 | 0 | 0 | 0 | 0 |
| Sleman | Turi | Wonokerto | -7.6148047 | 110.3772611 | 0 | 0 | 0 | 0 |
| Kota Yogyakarta | Kotagede | Purbayan | -7.82675905 | 110.4026556 | 5 | 3 | 8 | 0 |
| Kota Yogyakarta | Kotagede | Rejowinangun | -7.8114664 | 110.3982926 | 5 | 8 | 3 | 0 |
| Kota Yogyakarta | Umbulharjo | Sorosutan | -7.82241405 | 110.380925 | 8 | 3 | 2 | 0 |
| Kota Yogyakarta | Umbulharjo | Giwangan | -7.8310545 | 110.3895136 | 3 | 0 | 4 | 0 |
| Kota Yogyakarta | Wirobrajan | Patangpuluhan | -7.8097895 | 110.3489382 | 8 | 5 | 6 | 0 |

**Aggregated Cluster 1 to Subdistrict level**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sub\_District** | **Number\_College\_Venues** | **Number\_Office\_Venues** | **Number\_Coffee\_Venues** | **Customer\_to\_Competitor\_r** |
| Berbah | 27 | 0 | 3 | 9 |
| Gamping | 40 | 3 | 17 | 2.529411765 |
| Kalasan | 12 | 2 | 6 | 2.333333333 |
| Umbulharjo | 11 | 3 | 6 | 2.333333333 |
| Wirobrajan | 8 | 5 | 6 | 2.166666667 |
| Sleman | 12 | 10 | 11 | 2 |
| Depok | 7 | 1 | 4 | 2 |
| Kotagede | 10 | 11 | 11 | 1.909090909 |
| Prambanan | 14 | 0 | 8 | 1.75 |
| Ngemplak | 13 | 2 | 9 | 1.666666667 |
| Ngaglik | 25 | 9 | 22 | 1.545454545 |
| Seyegan | 5 | 0 | 4 | 1.25 |
| Pakem | 13 | 4 | 15 | 1.133333333 |
| Godean | 21 | 2 | 21 | 1.095238095 |
| Mlati | 11 | 2 | 13 | 1 |
| Moyudan | 8 | 0 | 8 | 1 |
| Cangkringan | 1 | 0 | 1 | 1 |
| Tempel | 7 | 0 | 8 | 0.875 |
| Minggir | 1 | 0 | 2 | 0.5 |
| Turi | 0 | 0 | 3 | 0 |

**Cluster 2**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Regency** | **Sub\_District** | **Village** | **Latitude** | **Longitude** | **Number\_College\_Venues** | **Number\_Office\_Venues** | **Number\_Coffee\_Venues** | **Cluster\_Label** |
| Sleman | Cangkringan | Umbulharjo | -7.8143777 | 110.3873737 | 20 | 9 | 12 | 1 |
| Sleman | Depok | Caturtunggal | -7.76739765 | 110.3884916 | 28 | 16 | 30 | 1 |
| Sleman | Depok | Condongcatur | -7.7541293 | 110.4006194 | 4 | 13 | 13 | 1 |
| Sleman | Mlati | Sinduadi | -7.7582322 | 110.3727371 | 32 | 5 | 13 | 1 |
| Kota Yogyakarta | Danurejan | Bausasran | -7.79936195 | 110.3731806 | 15 | 27 | 9 | 1 |
| Kota Yogyakarta | Danurejan | Tegalpanggung | -7.7934268 | 110.3711352 | 13 | 26 | 17 | 1 |
| Kota Yogyakarta | Gedongtengen | Pringgokusuman | -7.792505 | 110.3586344 | 8 | 16 | 17 | 1 |
| Kota Yogyakarta | Gondokusuman | Baciro | -7.7934704 | 110.3829725 | 27 | 18 | 4 | 1 |
| Kota Yogyakarta | Gondomanan | Ngupasan | -7.80468305 | 110.3712478 | 16 | 18 | 12 | 1 |
| Kota Yogyakarta | Gondomanan | Prawirodirjan | -7.80084825 | 110.3664578 | 17 | 24 | 21 | 1 |
| Kota Yogyakarta | Jetis | Bumijo | -7.7848352 | 110.3594767 | 18 | 27 | 15 | 1 |
| Kota Yogyakarta | Kotagede | Prenggan | -7.8207434 | 110.3974193 | 9 | 6 | 15 | 1 |
| Kota Yogyakarta | Kraton | Panembahan | -7.8086288 | 110.3661821 | 12 | 15 | 17 | 1 |
| Kota Yogyakarta | Kraton | Kadipaten | -7.80783795 | 110.3591871 | 17 | 10 | 13 | 1 |
| Kota Yogyakarta | Kraton | Patehan | -7.81090445 | 110.3607098 | 17 | 10 | 9 | 1 |
| Kota Yogyakarta | Mantrijeron | Gedongkiwo | -7.8171586 | 110.3540101 | 16 | 3 | 9 | 1 |
| Kota Yogyakarta | Mantrijeron | Suryodiningratan | -7.81975655 | 110.3591406 | 14 | 5 | 19 | 1 |
| Kota Yogyakarta | Mantrijeron | Mantrijeron | -7.81806685 | 110.3597312 | 14 | 6 | 20 | 1 |
| Kota Yogyakarta | Mergangsan | Brontokusuman | -7.8216755 | 110.3720775 | 9 | 9 | 18 | 1 |
| Kota Yogyakarta | Mergangsan | Keparakan | -7.81190505 | 110.3713592 | 39 | 17 | 13 | 1 |
| Kota Yogyakarta | Mergangsan | Wirogunan | -7.8088261 | 110.376306 | 41 | 7 | 9 | 1 |
| Kota Yogyakarta | Ngampilan | Ngampilan | -7.80218335 | 110.3576152 | 27 | 10 | 5 | 1 |
| Kota Yogyakarta | Ngampilan | Notoprajan | -7.80476795 | 110.3549974 | 30 | 9 | 6 | 1 |
| Kota Yogyakarta | Pakualaman | Gunungketur | -7.7992594 | 110.3786366 | 27 | 25 | 13 | 1 |
| Kota Yogyakarta | Pakualaman | Purwokinanti | -7.79298975 | 110.37545 | 6 | 23 | 12 | 1 |
| Kota Yogyakarta | Tegalrejo | Bener | -7.77613145 | 110.3523934 | 18 | 5 | 4 | 1 |
| Kota Yogyakarta | Tegalrejo | Karangwaru | -7.77426865 | 110.3641123 | 28 | 12 | 11 | 1 |
| Kota Yogyakarta | Tegalrejo | Kricak | -7.773404 | 110.3569328 | 13 | 8 | 5 | 1 |
| Kota Yogyakarta | Tegalrejo | Tegalrejo | -7.78045495 | 110.355073 | 24 | 14 | 2 | 1 |
| Kota Yogyakarta | Umbulharjo | Pandeyan | -7.81831055 | 110.3884081 | 14 | 7 | 9 | 1 |
| Kota Yogyakarta | Umbulharjo | Warungboto | -7.80821305 | 110.3902059 | 26 | 7 | 15 | 1 |
| Kota Yogyakarta | Umbulharjo | Muja muju | -7.7986723 | 110.3927106 | 29 | 30 | 3 | 1 |
| Kota Yogyakarta | Umbulharjo | Semaki | -7.7978461 | 110.3837372 | 25 | 28 | 5 | 1 |
| Kota Yogyakarta | Umbulharjo | Tahunan | -7.8067426 | 110.3831715 | 13 | 9 | 7 | 1 |
| Kota Yogyakarta | Wirobrajan | Pakuncen | -7.79696465 | 110.3511308 | 20 | 9 | 3 | 1 |
| Kota Yogyakarta | Wirobrajan | Wirobrajan | -7.8026238 | 110.3504467 | 22 | 4 | 3 | 1 |

**Aggregated Cluster 2 to Subdistrict level**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sub\_District** | **Number\_College\_Venues** | **Number\_Office\_Venues** | **Number\_Coffee\_Venues** | **Customer\_to\_Competitor\_r** |
| Gondokusuman | 27 | 18 | 4 | 11.25 |
| Wirobrajan | 42 | 13 | 6 | 9.166666667 |
| Ngampilan | 57 | 19 | 11 | 6.909090909 |
| Tegalrejo | 83 | 39 | 22 | 5.545454545 |
| Umbulharjo | 107 | 81 | 39 | 4.820512821 |
| Pakualaman | 33 | 48 | 25 | 3.24 |
| Danurejan | 28 | 53 | 26 | 3.115384615 |
| Mergangsan | 89 | 33 | 40 | 3.05 |
| Jetis | 18 | 27 | 15 | 3 |
| Mlati | 32 | 5 | 13 | 2.846153846 |
| Cangkringan | 20 | 9 | 12 | 2.416666667 |
| Gondomanan | 33 | 42 | 33 | 2.272727273 |
| Kraton | 46 | 35 | 39 | 2.076923077 |
| Depok | 32 | 29 | 43 | 1.418604651 |
| Gedongtengen | 8 | 16 | 17 | 1.411764706 |
| Mantrijeron | 44 | 14 | 48 | 1.208333333 |
| Kotagede | 9 | 6 | 15 | 1 |

**Cluster 3**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Regency** | **Sub\_District** | **Village** | **Latitude** | **Longitude** | **Number\_College\_Venues** | **Number\_Office\_Venues** | **Number\_Coffee\_Venues** | **Cluster\_Label** |
| Kota Yogyakarta | Danurejan | Suryatmajan | -7.79283925 | 110.3673405 | 13 | 34 | 44 | 2 |
| Kota Yogyakarta | Gedongtengen | Sosromenduran | -7.78933775 | 110.3633763 | 12 | 33 | 46 | 2 |
| Kota Yogyakarta | Gondokusuman | Demangan | -7.78487125 | 110.3900292 | 34 | 31 | 35 | 2 |
| Kota Yogyakarta | Gondokusuman | Klitren | -7.783821 | 110.3833898 | 28 | 28 | 29 | 2 |
| Kota Yogyakarta | Gondokusuman | Kotabaru | -7.7868301 | 110.3734326 | 33 | 42 | 31 | 2 |
| Kota Yogyakarta | Gondokusuman | Terban | -7.7792947 | 110.3755437 | 18 | 50 | 50 | 2 |
| Kota Yogyakarta | Jetis | Cokrodiningratan | -7.7784096 | 110.3679234 | 17 | 26 | 34 | 2 |
| Kota Yogyakarta | Jetis | Gowongan | -7.7859455 | 110.3659155 | 22 | 42 | 50 | 2 |

**Cluster 3 aggregated to subdistrict level**

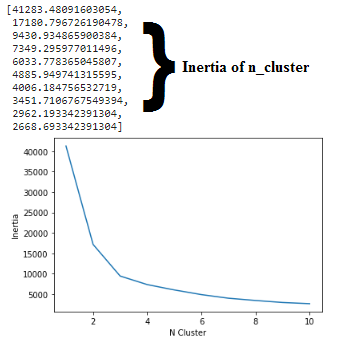
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sub\_District** | **Number\_College\_Venues** | **Number\_Office\_Venues** | **Number\_Coffee\_Venues** | **Customer\_to\_Competitor\_r** |
| Gondokusuman | 113 | 151 | 145 | 1.820689655 |
| Jetis | 39 | 68 | 84 | 1.273809524 |
| Danurejan | 13 | 34 | 44 | 1.068181818 |
| Gedongtengen | 12 | 33 | 46 | 0.97826087 |

**Top villages to consider**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Sub\_District\_x** | **Village** | **Latitude** | **Longitude** | **Number\_College\_Venues** | **Number\_Office\_Venues** | **Number\_Coffee\_Venues** | **Cluster\_Label** | **Customer\_to\_Competitor\_r** | **Total\_Population x1000** |
| Tegalrejo | Tegalrejo | -7.78045495 | 110.355073 | 24 | 14 | 2 | 1 | 19 | 6.54 |
| Berbah | Tegaltirto | -7.8054225 | 110.4417249 | 13 | 0 | 0 | 0 | 14 | 8.833 |
| Gondokusuman | Baciro | -7.7934704 | 110.3829725 | 27 | 18 | 4 | 1 | 11.25 | 8.927 |
| Wirobrajan | Pakuncen | -7.79696465 | 110.3511308 | 20 | 9 | 3 | 1 | 9.666666667 | 7.712 |
| Wirobrajan | Wirobrajan | -7.8026238 | 110.3504467 | 22 | 4 | 3 | 1 | 8.666666667 | 6.749 |
| Ngampilan | Ngampilan | -7.80218335 | 110.3576152 | 27 | 10 | 5 | 1 | 7.4 | 7.367 |
| Ngampilan | Notoprajan | -7.80476795 | 110.3549974 | 30 | 9 | 6 | 1 | 6.5 | 5.798 |
| Tegalrejo | Bener | -7.77613145 | 110.3523934 | 18 | 5 | 4 | 1 | 5.75 | 3.529 |
| Berbah | Kalitirto | -7.7951322 | 110.4553582 | 4 | 0 | 0 | 0 | 5 | 9.692 |
| Tegalrejo | Kricak | -7.773404 | 110.3569328 | 13 | 8 | 5 | 1 | 4.2 | 9.443 |
| Berbah | Sendangtirto | -7.82044215 | 110.4313663 | 8 | 0 | 2 | 0 | 4 | 12.59 |
| Berbah | Jogotirto | -7.8129504 | 110.4720599 | 2 | 0 | 1 | 0 | 2 | 7.8 |

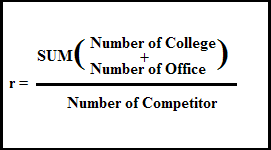
**Elaboration**

To solve the business problem, this research using unsupervised machine learning taks named clustering technique with Kmeans algorithm as the approach. Kmeans is one of unsupervised algorithm that works by assigning similarity of data characteristics to the cluster center (centroid), kmeans is highly dependant on distance metrics to determine the similarity between data points and when assigning data points to its centroid. Kmeans algorithm need number of clusters as essential parameter to get the clustering done. Since there is no initial business problem stated on how many cluster to find, this research using elbow method from inertia (density) of a cluster to find best cluster.



From the elbow method conducted above, the best number of cluster for this dataset is 3, therefore Kmeans with n\_clusters of 3 will be implied to the dataset. Dataset contain three features of every villages, which are number of college university, number of office building, and number for coffee shop (competitors) nearby.

Result of clustering that applied to are shown above table. To get better understanding to analyze cluster label results, a particular heuristic method used which is calculating sum of potential customer to competitor ratio.



* ratio of potential customers to competitors in cluster 1 = 17 : 10
* ratio of potential customers to competitors in cluster 2 = 29 : 10
* ratio of potential customers to competitors in cluster 3 = 3 : 2

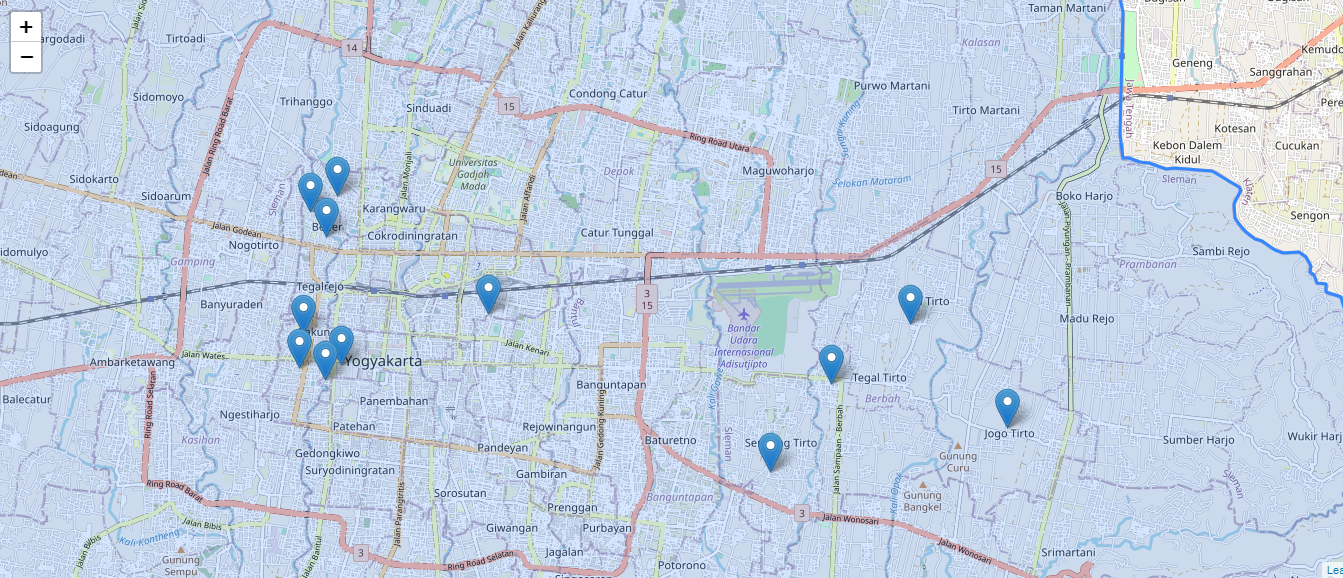
In summary, the **formula means the higher the ratio is the better will be** for the business because the business will have more potential customers and less competitors nearby location. Cluster 1 has moderate customer to competitor ratio, cluster 2 has most customer to competitor ratio, and cluster 3 has least customer to competitor ratio. At a glance cluster 2 and cluster 1 village location is worth taking, but further analysis might have more clarity and more satisfy the problem solution.

Using the customer to competitor ratio and from the table shown above, we may conclude **best sub-districts** which are;

* Gondokusuman with 11.25
* Wirobrajan with 9.17
* Berbah 9.0
* Ngampilan 6.9
* Tegalrejo 5.6

Due to the business objective to provide location options at the village level, we will deep dive from the best subdistricts into village level. Above table is showing the top villages with highest customer to competitor ratio with addition of total population in the village.

Yogyakarta is a province where the population is mostly immigrants like college students, entrepreneurs, workers, and others. **The total population in the village data can be a gross minimum of potential customers to the establishment**.



**(Recommended Location for Establishment)**

**Chapter V**

**Discussion**

**5.1 Discussion**

The top ten villages recommendation are Tegalrejo, Tegaltirto, Baciro, Pakuncen, Wirobrajan, Ngampilan, Notoprajan, Bener, Kalitirto, Kricak, Sendangtirto, Jogotirto. The results choosen from several consideration such as finest cluster customer to competitor ratio, finest subdistrict customer to competitor ratio, and finally the finest village customer to competitor ratio. High value of customer to competitor ratio is one of the factors leading to business success and generating provit, other factors of successful business such as quality of product, operation management, asset management, etc are beyond consideration of this research.

Despite the results might have answered the business problem and questions, the results may be accepted or rejected by the decision maker of the business stakeholders due to other factors such as the location captialization is too expensive or beyond affordable or the location is just not suit to open coffee shop business.

**5.2 Possible Flaw**

* Foursquare API isnt that powerful in Indonesia, hence the location of venues are not fully show its potentials.
* Venues response of Foursquare API might not relevant to the query’s categoryId
* Consideration factors to determine a cluster too simple, only three (college venues, office venues, and coffee shop venues)
* Location with better customer to competitor ratio might be exist and not choosen due to angle of point of view.
* Different Kmeans parameter might yield different results. Like n\_init, centroid algorithm, and n\_cluster

**Chapter VI**

**Conclusion**

**6.1 Conclusion**

Attempting to find strategic location to establish coffee shop within Yogyakarta in village level position can be done with geo spatial visualization combined with unsupervised machine learning task: clustering technique. The results of this research paper are: Tegalrejo with customer to competitor ratio 19.0, Tegaltirto with customer to competitor ratio 14.0, Baciro with customer to competitor ratio 11.25, Pakuncen with customer to competitor ratio 9.7, Wirobrajan with customer to competitor ratio 8.7, Ngampilan with customer to competitor ratio 7.4, Notoprajan with customer to competitor ratio 6.5, Bener with customer to competitor ratio 5.75, Kalitirto with customer to competitor ratio 5.0, Kricak with customer to competitor ratio 4.2, Sendangtirto with customer to competitor ratio 4, and Jogotirto with customer to competitor ratio 2.

**6.2 Further Research**

* using google places API
* increase clustering consideration factor (variables)
* further research on different clustering algorithms
* add more risk factor beside number of competitor

**References**

Urls:

1. <https://id.wikipedia.org/wiki/Daftar_kapanewon,_kemantren,_kalurahan,_dan_kelurahan_di_Daerah_Istimewa_Yogyakarta>
2. <https://kependudukan.jogjaprov.go.id/statistik/penduduk/golonganusia/15/produktif/3/71/34.ez>
3. <https://kependudukan.jogjaprov.go.id/statistik/penduduk/golonganusia/15/produktif/07/71/34.ez>
4. <https://kependudukan.jogjaprov.go.id/statistik/penduduk/golonganusia/15/produktif/8/04/34.ez>
5. <https://kependudukan.jogjaprov.go.id/statistik/penduduk/golonganusia/15/produktif/6/71/34.ez>
6. <https://kependudukan.jogjaprov.go.id/statistik/penduduk/golonganusia/15/produktif/01/71/34.ez>

tools:

1. Python 3
2. Jupyter environment
3. Google colabs
4. Foursquare API
5. Nominatim openstreetmap
6. Github

Python libraries and packages:

1. Pandas
2. Numpy
3. Matplotlib
4. Folium
5. Beautiful soup
6. Json
7. Requests
8. Scikit-learn
9. Joblib
10. Pickle

Project file:

1. Strategic\_Branch\_Estabilishment.ipynb
2. Top5\_subdistricts\_demographics\_webscrape.ipynb
3. Coffee\_NewBranch\_Kmeans\_cluster.pkl
4. YK\_geojson.geojson
5. Strategic\_place\_to\_establish.html
6. Summary villages recommendation.csv
7. Village Cluster Map.html
8. Productive\_population\_15to64\_of\_top5\_subdistricts.csv
9. Complete\_data\_strategic\_establishment.csv
10. Final\_Data\_SLE\_YK.csv
11. Sleman\_Yogyakarta\_village.html