GROUP 3

28th June 2024

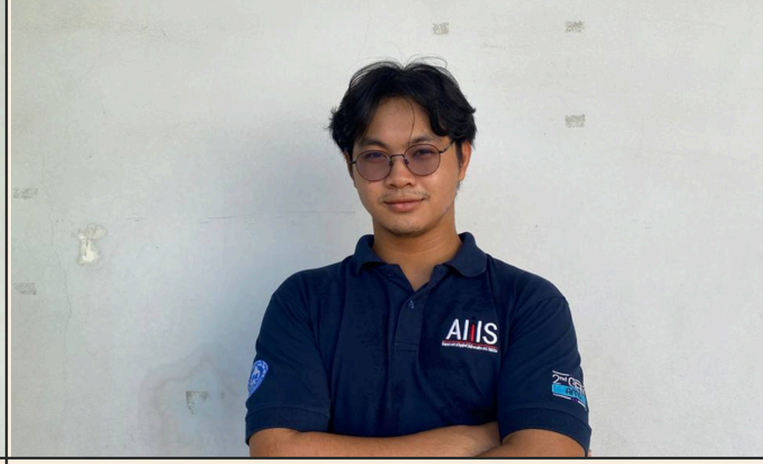
EduCafé

Optimizing Cafe Placement in Phnom Penh Through K-Means Clustering

Information available in audio.

Our Team Members

PEN Virak
e20211572



PEL Bunkloem
e20201314

KHUN Sithanut
e20211527

PAV Limseng
e20211548

PEANG Rattanak
e20210072

Agenda

1 Introduction

2 Methodology

3 Result and Finding

4 Discussion and Conclusion

Introduction

Objective: Identify optimal cafe locations in Phnom Penh targeting educational institutions using k-means clustering.

Rationale: Educational buildings are hubs of consistent daily activity, making them ideal focal points for a cafe due to the high density of potential customers, including students and staff.

Approach:

- **Data Collection:** Gather data on educational buildings in Phnom Penh.
- **Clustering:** Apply k-means clustering to identify geographic clusters of educational institutions.
- **Optimization:** Use the elbow method to determine the optimal number of clusters, pinpointing areas underserved by existing cafes.

Goal: Strategically position a new cafe to maximize accessibility and profitability by serving the educational community, enhancing customer retention and business success.

Back to Agenda

Methodology

Briefly elaborate on what you want to discuss.

Data Source

- Google Map

Data Collection

- Scraping from Google map by using Google map API Key.
- Main Target: Educational buildings.

Back to Agenda

Target Area

Location:

Phnom Penh

Center point:

Stueng Mean Chey

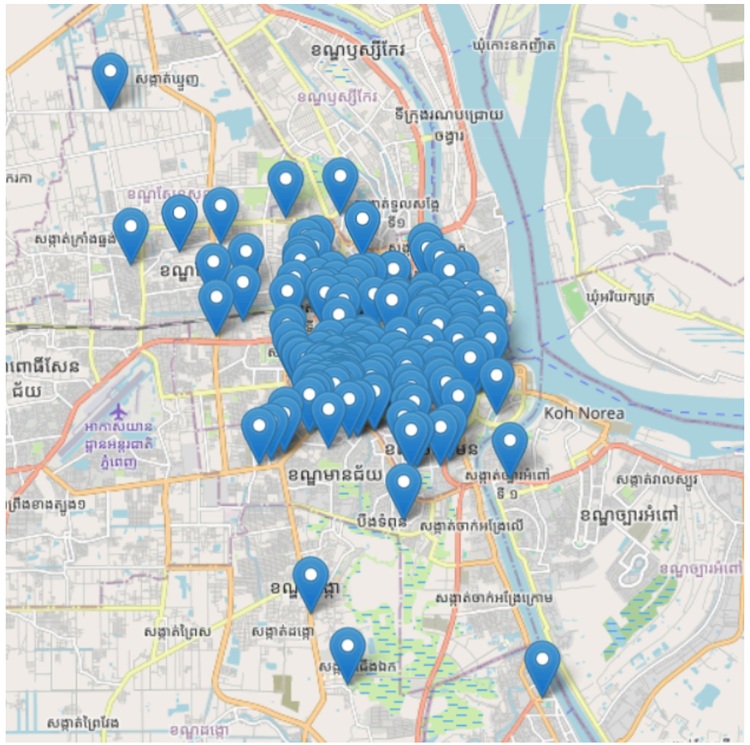
Radius around center point:

10 KM

Back to Agenda

Data Preprocessing

Raw Data



183 Educational Buildings

	Name	Address	Latitude	Longitude	Type
0	CJCC (Cambodia-Japan Cooperation Center)	Rupp-CJCC, មហាវិថី សហព័ន្ធរុស្ស៊ី (១90), ភ្នំពេញ	11.568929	104.893694	school, point_of_interest, establishment
1	Angkor Computer Center	#95E0, Saint 164, Phnom Penh	11.563684	104.912819	school, point_of_interest, establishment
2	SAS Santhormuk - Stanford American School	#197, St.146, Teuk laok 2, Phnom Penh	11.564989	104.899767	school, primary_school, secondary_school, poin...
3	Sovannaphumi School, Tep Phan Campus	6A Oknha Tep Phan St. (182), Phnom Penh	11.563052	104.900276	school, point_of_interest, establishment
4	Aii Language Center (Aii), Mao Tse Tong (QLH B...	217 ABCD Mao Tse Tong Blvd, ភ្នំពេញ	11.546374	104.907986	school, point_of_interest, establishment

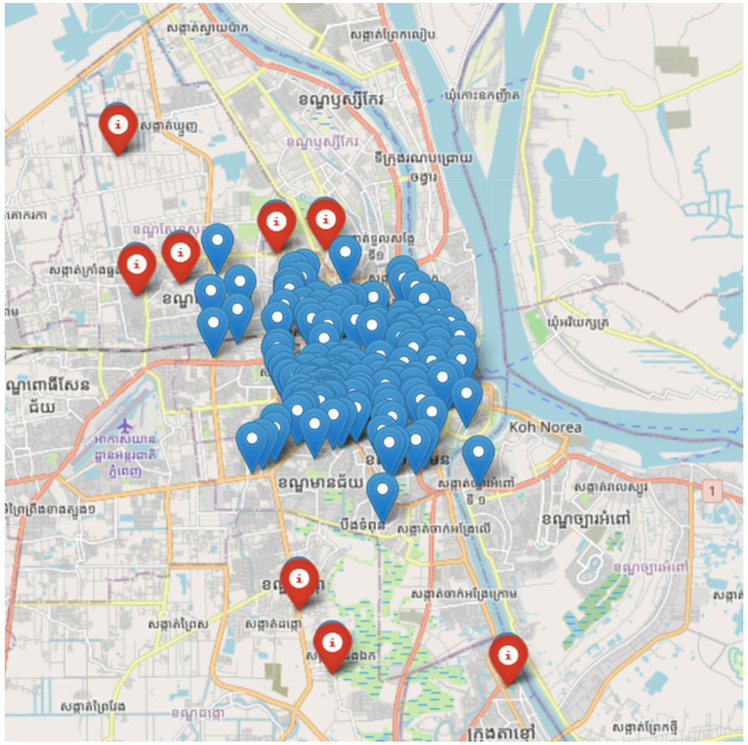
Dataset Head

Back to Agenda

Data Preprocessing

Detect Outlier

After confirming these outliers, we excluded them from our dataset to maintain the integrity and relevance of our analysis, focusing on core areas where our target demographic is concentrated. After removing the 8 outliers, we were left with a refined dataset of 175 addresses,

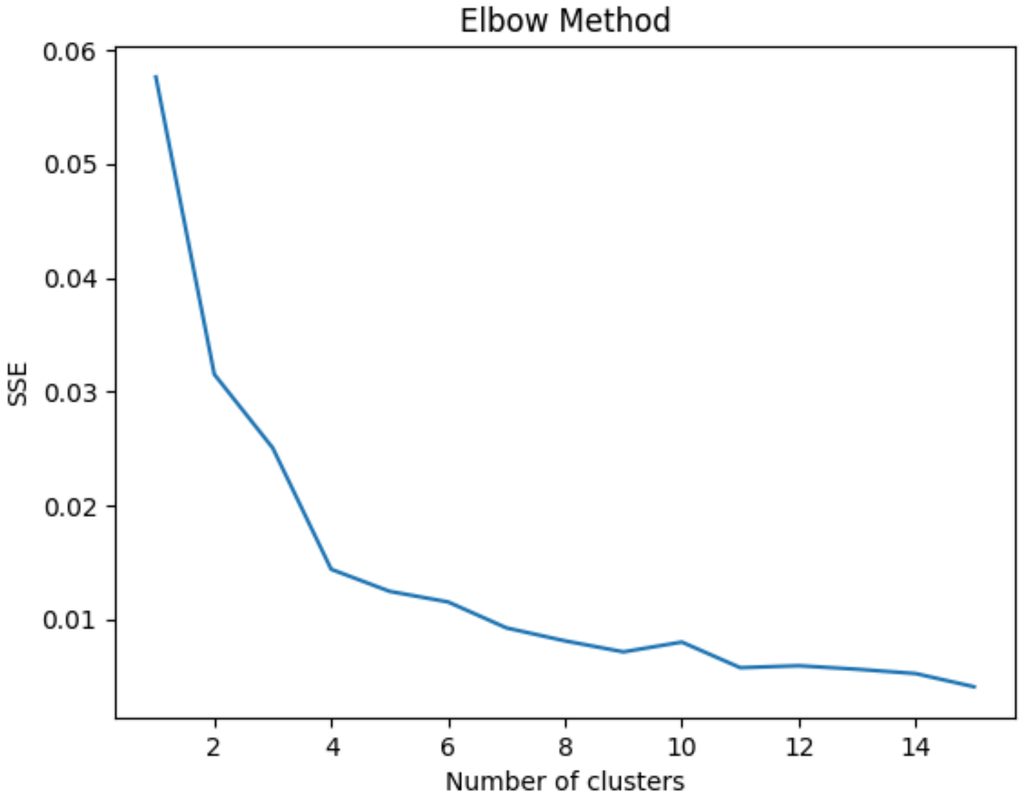


Outlier Locations

Back to Agenda

Data Preprocessing

Appying Elbow Method

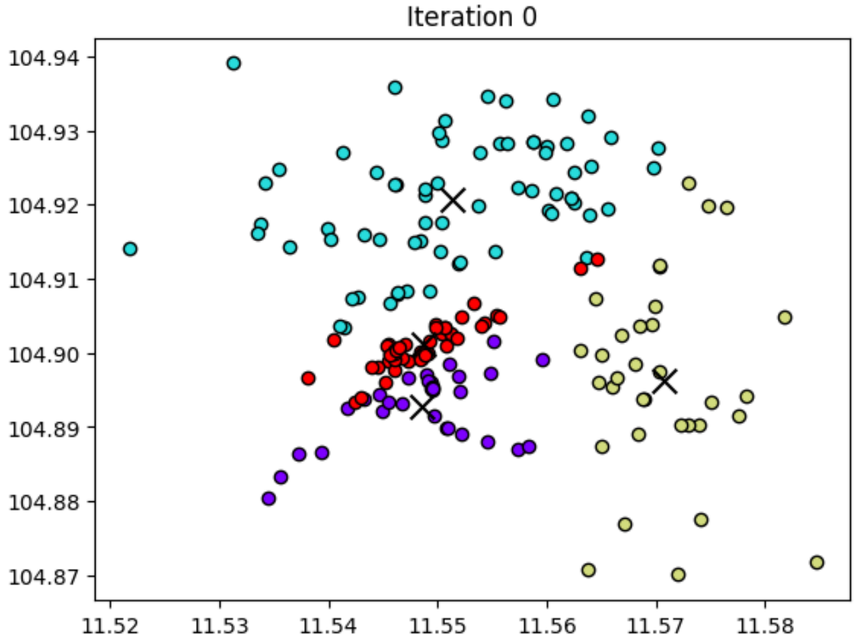


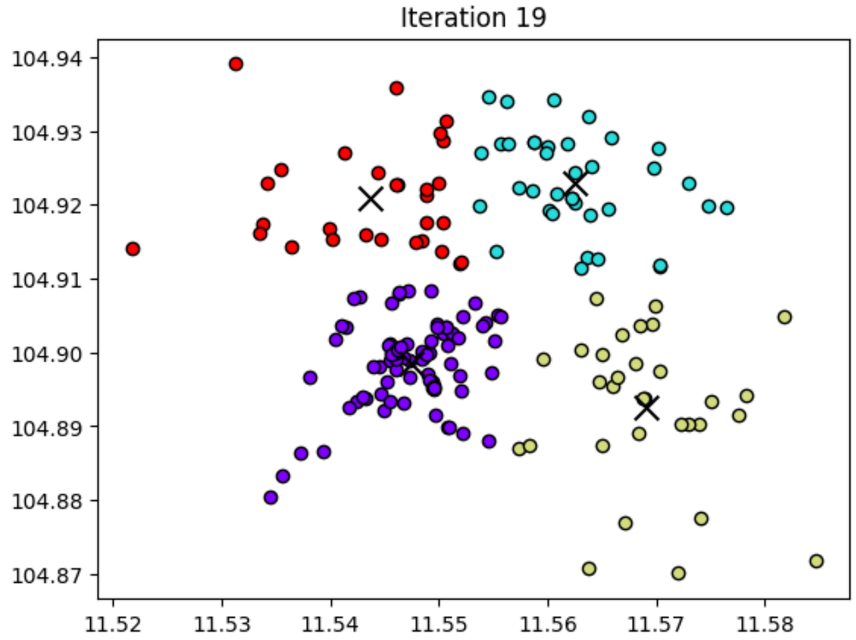
In our analysis, we plotted the WCSS for different values of K and observed the "elbow" point. The plot indicated a distinct bend at K=4, suggesting that 4 clusters are optimal for our dataset.

Back to Agenda

3. Result

• Cluster 1 (Yellow): This cluster has a higher density and smaller spread, indicating a concentrated area of potential customers.
• Cluster 2 (Cyan): Points in this cluster are more dispersed, suggesting a larger geographic area with potential customers spread out.
• Cluster 3 (Purple): This cluster shows moderate density and spread, providing a balance between customer concentration and area coverage.
• Cluster 4 (Red): Similar to Cluster 1, this cluster has a high density of points but in a different location.





4. Discussion and Conclusion

The k-means clustering model and elbow method were used to identify optimal locations for a cafe business in Phnom Penh, focusing on proximity to educational buildings. The optimal number of clusters was determined to be 4, with clusters strategically positioned for accessibility and convenience. The locations have significant potential for attracting students and staff from nearby educational institutions, aligning with the target demographic. The clustering analysis and elbow method identified four optimal locations for sustainable growth and profitability.

GROUP 3

Thank You !

Email

Group3@itc.com

Social Media

@Group 3

Contact

168-168-168