

MUNICIPAL RESOURCE DISTRIBUTION IN ISRAEL

A COMPREHENSIVE ANALYSIS OF LOCAL
AUTHORITIES IN 2018

Course: Advanced Topics in Machine Learning

Instructor: Chen Hajaj

Team Members: Odeya Hazani

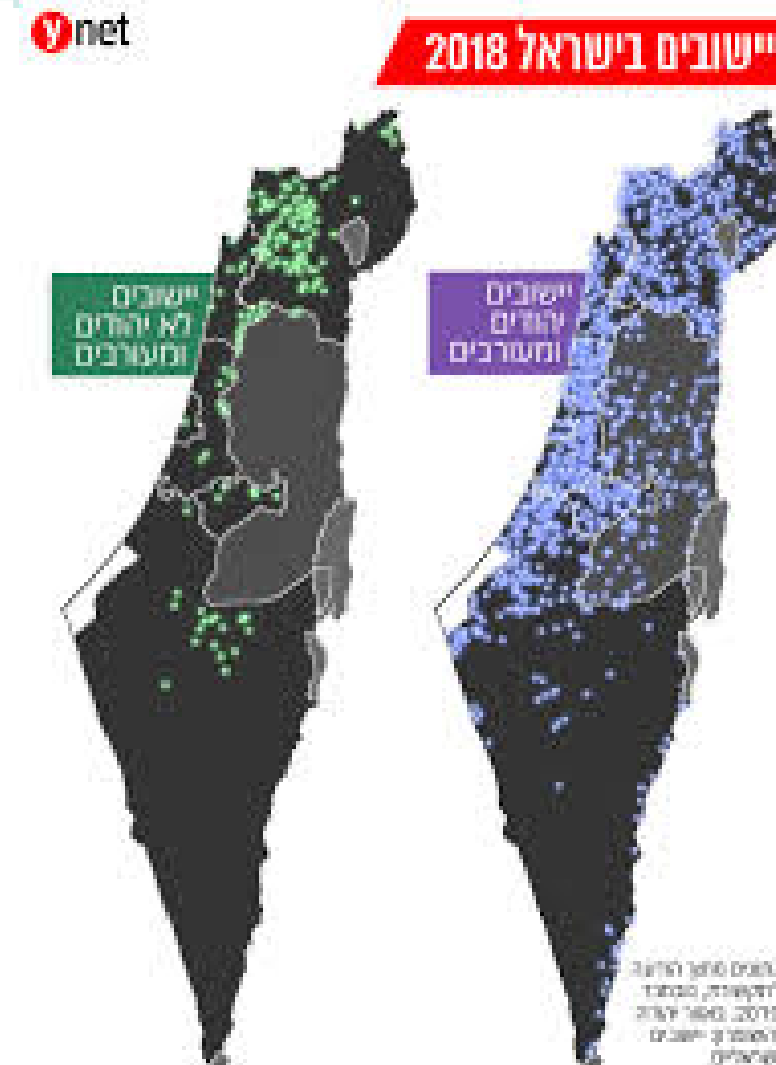
Roni Epstein

-
-
-
-

Abstract

This study explores resource distribution among 255 local authorities in Israel using unsupervised learning techniques.

By analyzing financial, demographic, and infrastructure-related data, the research identifies significant disparities and patterns that can inform policymakers and improve resource allocation strategies.



Introduction

- **Problem Statement:** A lack of detailed understanding of resource allocation patterns across Israel's municipal authorities hinders effective policy-making.
- **Research Objective:** To uncover patterns in municipal governance and derive actionable insights for policy decisions.



Dataset and Features

- **Source:** Data.gov.il.
- **Details:** 255 authorities (77 municipalities, 124 local councils, 54 regional councils).
- **397 features**, including demographics, education, welfare, infrastructure, and finance.

Examples of features:

- authority_status: The type of authority (regional or local council).
- general_district: The district of the settlement.
- Settlement Religion: The religion associated with the settlement.

נתונים פיזיים ונתוני אוכלוסייה		1
עיריות	רשימת רשויות לפי מעמד מוניציפלי	2
מועצות מקומיות		3
מועצות אזוריות		4
סמל הרשות	כללי	5
מחוז		6
מעמד מוניציפלי	תקבולים	7
מרחק מגבול מחוז תל אביב	תקבולים בתקציב הרגיל	8
שנת קבלת מעמד מוניציפלי	לפי פרקי תקציב	9
מרחב ימי, רשויות מקומיות הגובלות בחוף הים		10
מספר חברי מועצה		11
סמל ועדת תכנון ובנייה		12
שם ועדת תכנון ובנייה		13
שטח בקמ"ר	תקבולים בתקציב הרגיל	14
צפיפות אוכלוסייה לקמ"ר	לפי סיווג כלכלי	15
סה"כ אוכלוסייה בסוף השנה		16
יהודים ואחרים (אחוזים)		17
יהודים (אחוזים מתוך יהודים ואחרים)		18
ערבים (אחוזים)		19
מוסלמים (אחוזים מתוך האוכלוסייה הערבית)		20
נוצרים (אחוזים מתוך האוכלוסייה הערבית)		21
דרוזים (אחוזים מתוך האוכלוסייה הערבית)		22
סה"כ גברים בסוף השנה (אלפים)	תקבולים בתקציב הבלתי רגיל	23
סה"כ נשים בסוף השנה (אלפים)		24
אחוז באוכלוסייה לפי גיל בסוף השנה		25
יחס תלות (לאלף תושבים בלתי תלויים)		26
לידות חי'		27
פטירות תינוקות (ממוצע)		28
פטירות תינוקות (ממוצע)		
שיעור פטירות תינוקות לאלף לידות חי'		
פטירות		
רביי טבעי		
רביי טבעי לאלף תושבים		
אחוז גידול האוכלוסייה לעומת שנה קודמת		
אחוז עולי 1990+ מסך האוכלוסייה		
השתקעות ראשונה של עולים		
אחוז עולים בהשתקעות ראשונה לפי גיל		
מאזן הגירה ביישוב		
מאזן הגירה פנימית		
נכנסים מיישובים אחרים לפי מין וגיל		
יוצאים ליישובים אחרים לפי מין וגיל		
הגירה בתוך הרשות המקומית		
יישאים		
ותגרשים		
אוכלוסייה במוסדות לניה (אחוז מתוך אוכלוסיית הרשות)		
שיעור פטירות מתוקן ל-1,000 תושבים	בראות	
תוחלת חיים בלידה סך הכל		
שיעור פרוץ כולל		
מספר מקרי סוכרת ממוצע לשנה		
שיעור מקרי סוכרת מתוקן ל-1,000 תושבים		
עודף משקל והשמנה כיתה א		
עודף משקל והשמנה כיתה ז		
מספר מקרי סרטן ממוצע לשנה, גברים		
שיעור סרטן מכל הסוגים מתוקן ל-100,000 תושבים, גברים		

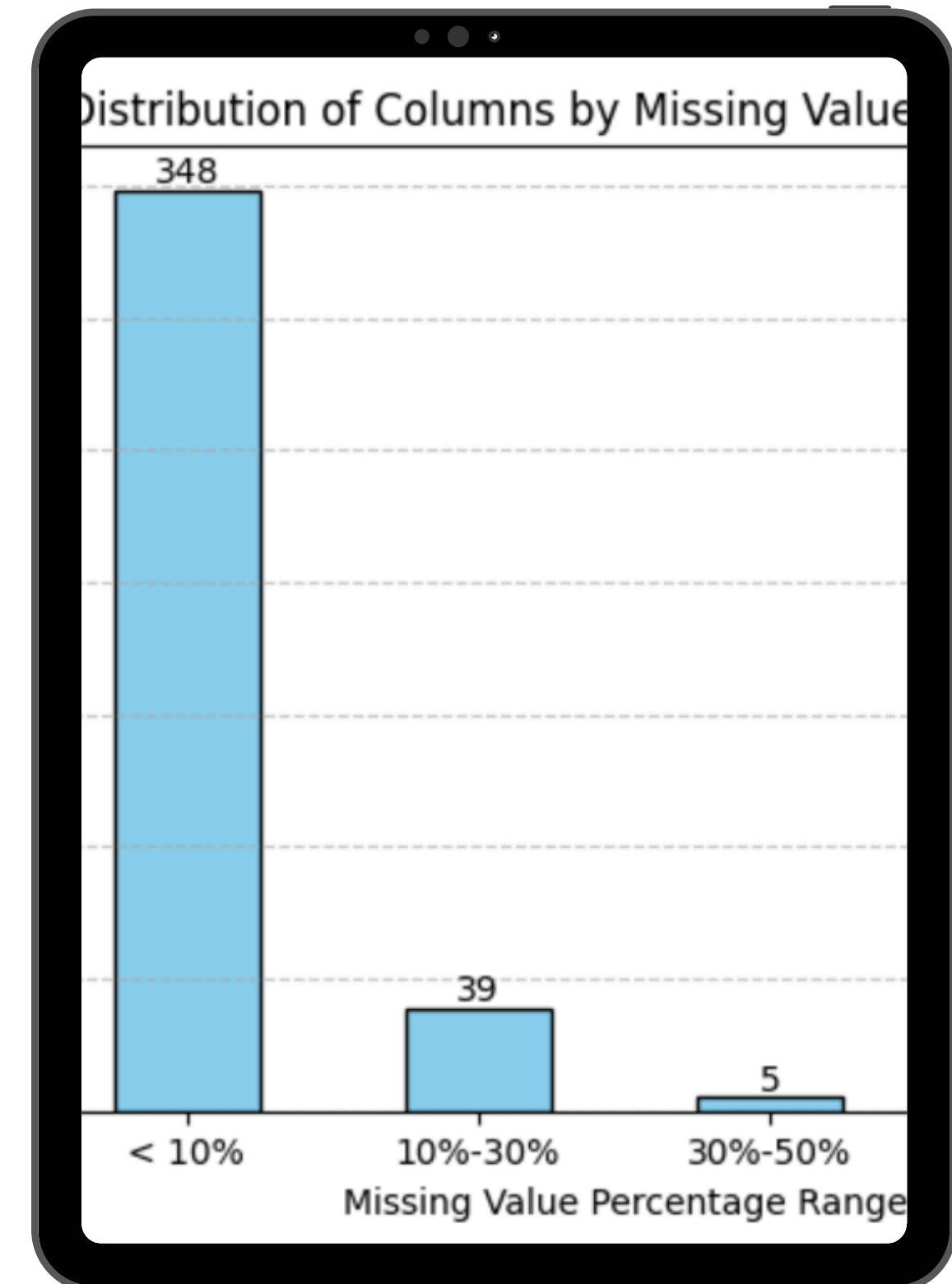
Data Preprocessing

Handled missing data

Addressed outliers with Isolation Forest, retaining meaningful cases.

Reducing Redundancy

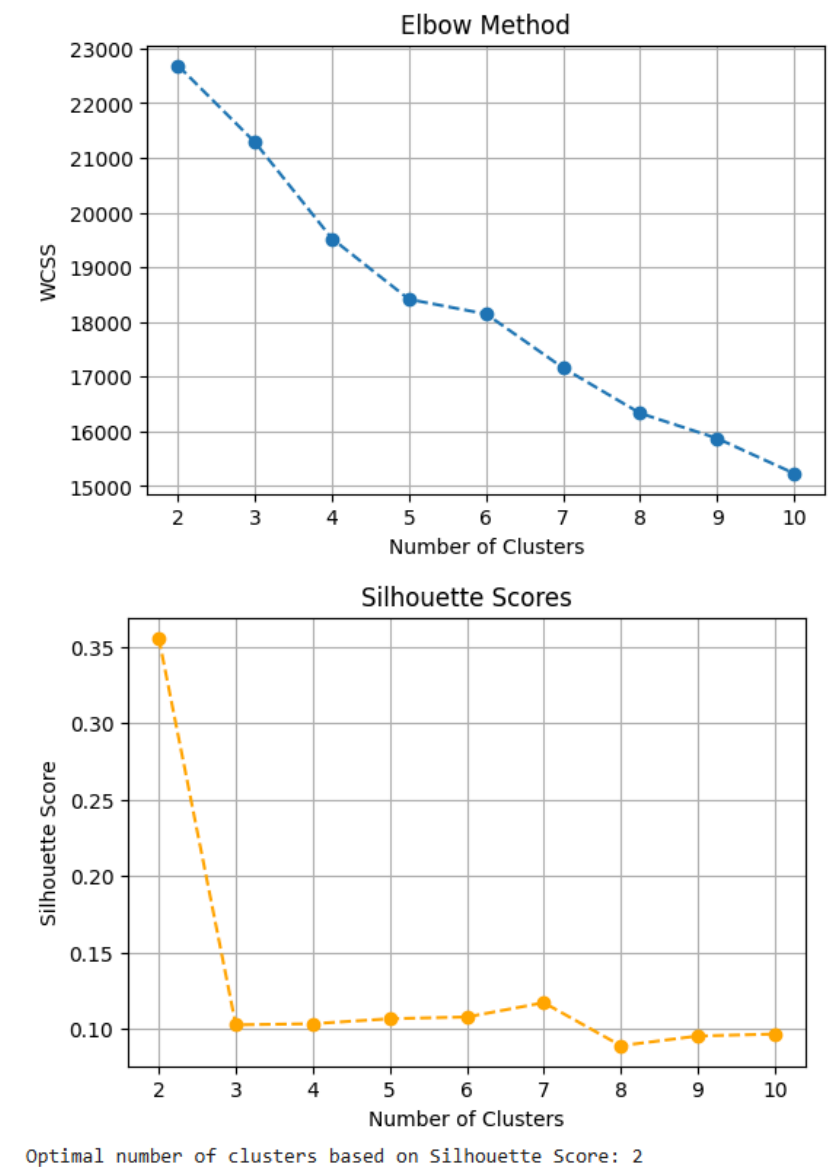
Encoded categorical columns and Scaled numeric columns



Methodology

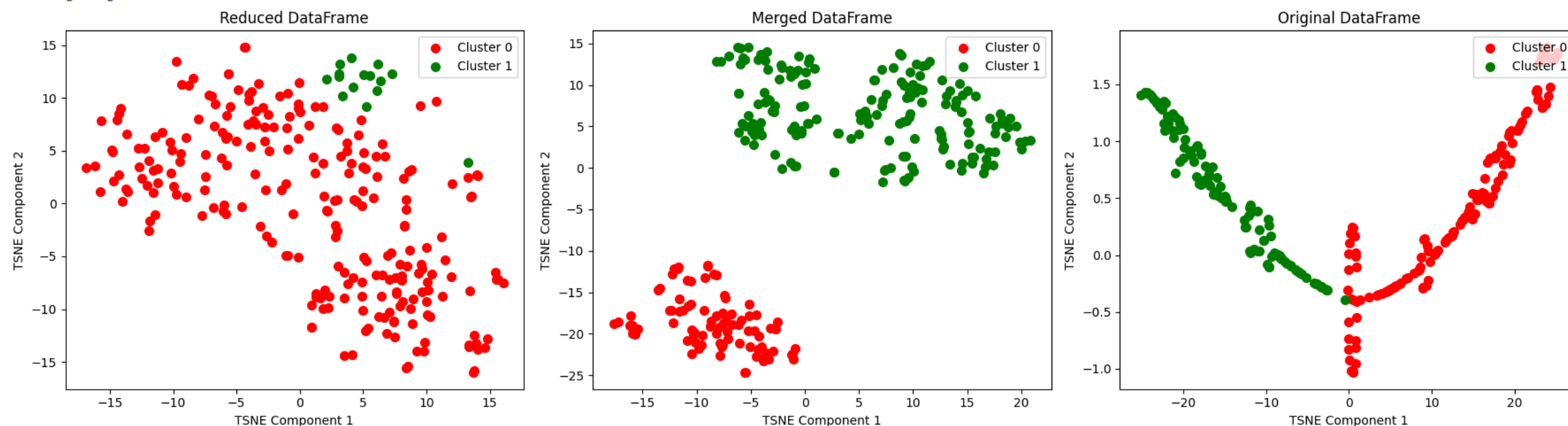
DIMENSIONALITY REDUCTION AND INITIAL ANALYSIS

- **Initial Approach:** Divided data into 22 topic-based groups (e.g., economy, demographics) to handle dataset complexity, Applied PCA with Standardization on each group, retaining 90% variance and visualizing results.
- **Unified Dataset:** Combined PCA components into 108 columns.
- **Cluster Number Optimization:** The optimal number of clusters was determined using a combination of the Elbow Method, Silhouette Score and TSNE visualization



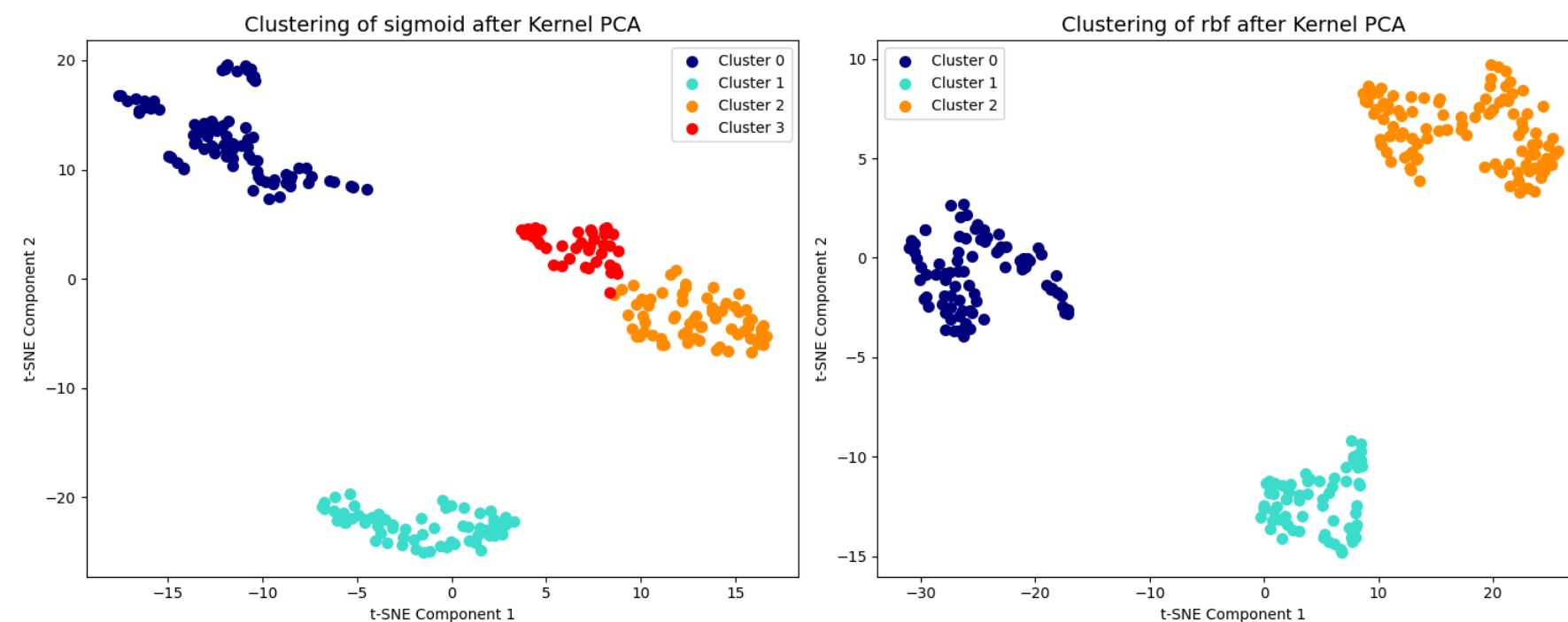
Methodology PCA

- **PCA** was applied despite its known limitations with linear and high-dimensional data.
- **PCA Analysis:** Reduced the dataset to 49 principal components, capturing 90% of the variance without affecting clustering quality.
- **Clustering Analysis:** Achieved two clusters with a Silhouette Score of 0.355.
- **t-SNE Visualization:** the distribution appeared uneven and suboptimal.



Methodology Kernel PCA

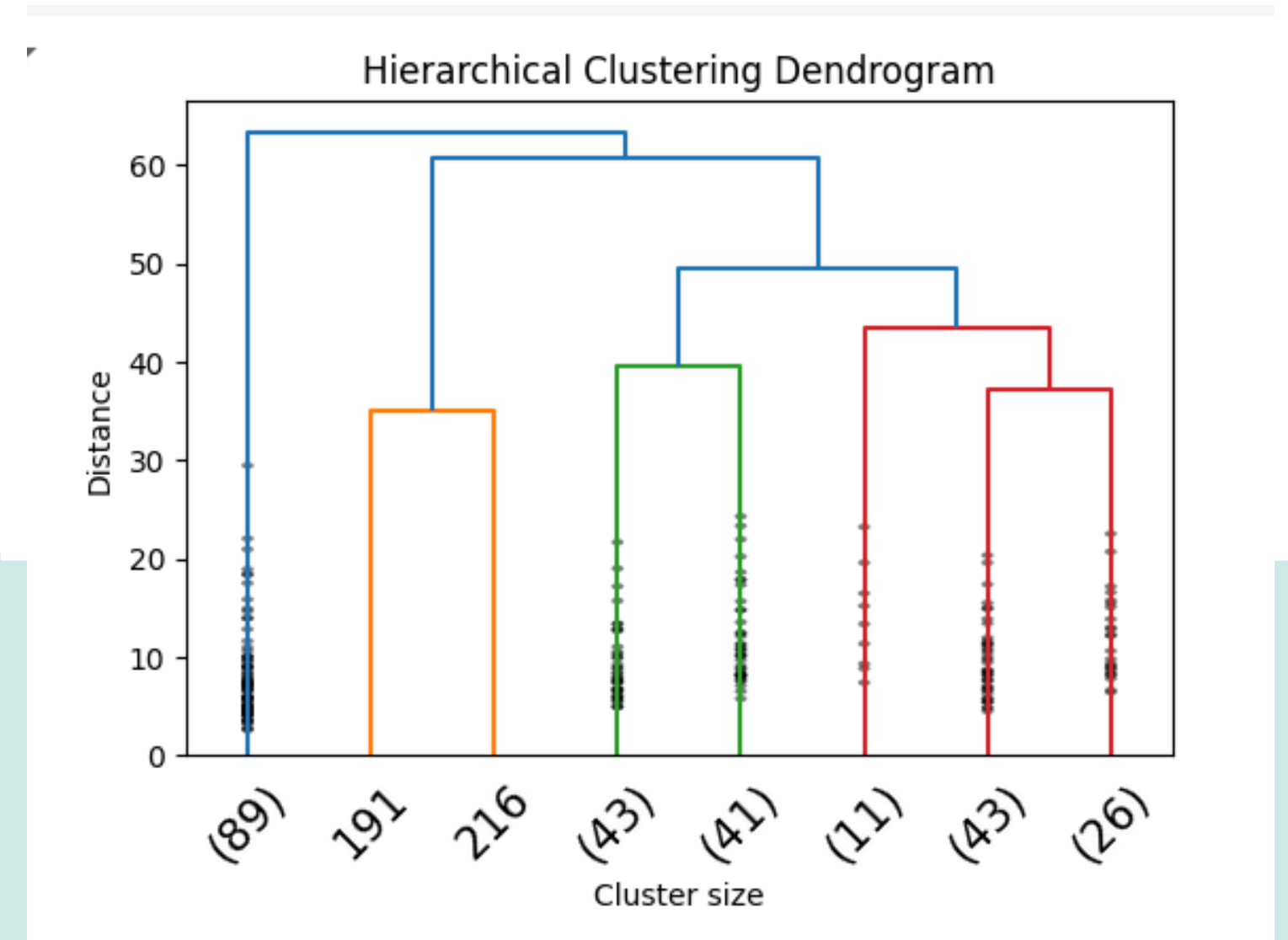
- **Kernel PCA:** RBF and Sigmoid Used to explore non-linear relationships.
- **Clustering Results:** Clustering was performed with $k=4$ for both kernels.
- **Silhouette Scores:** RBF (0.508), Sigmoid (0.544).
- **t-SNE Visualization:** Both kernels produced clearer and better-defined clusters, the RBF kernel, while recommending $k=3$, provided superior visual clarity.



Methodology

Hierarchical Trees

- **Hierarchical Tree Validation:** Validated the Three-Cluster Hypothesis, ensuring stable and meaningful clusters.
- **Ward's Linkage Method:** Chosen to minimize Within-Cluster Variance, suitable for data without assumptions.
- **Silhouette Score:** Achieved 0.466, indicating moderate fit and some cluster separation.
- **Cluster Selection:** The tree was cut to form 3 clusters based on the dendrogram.



Comparison

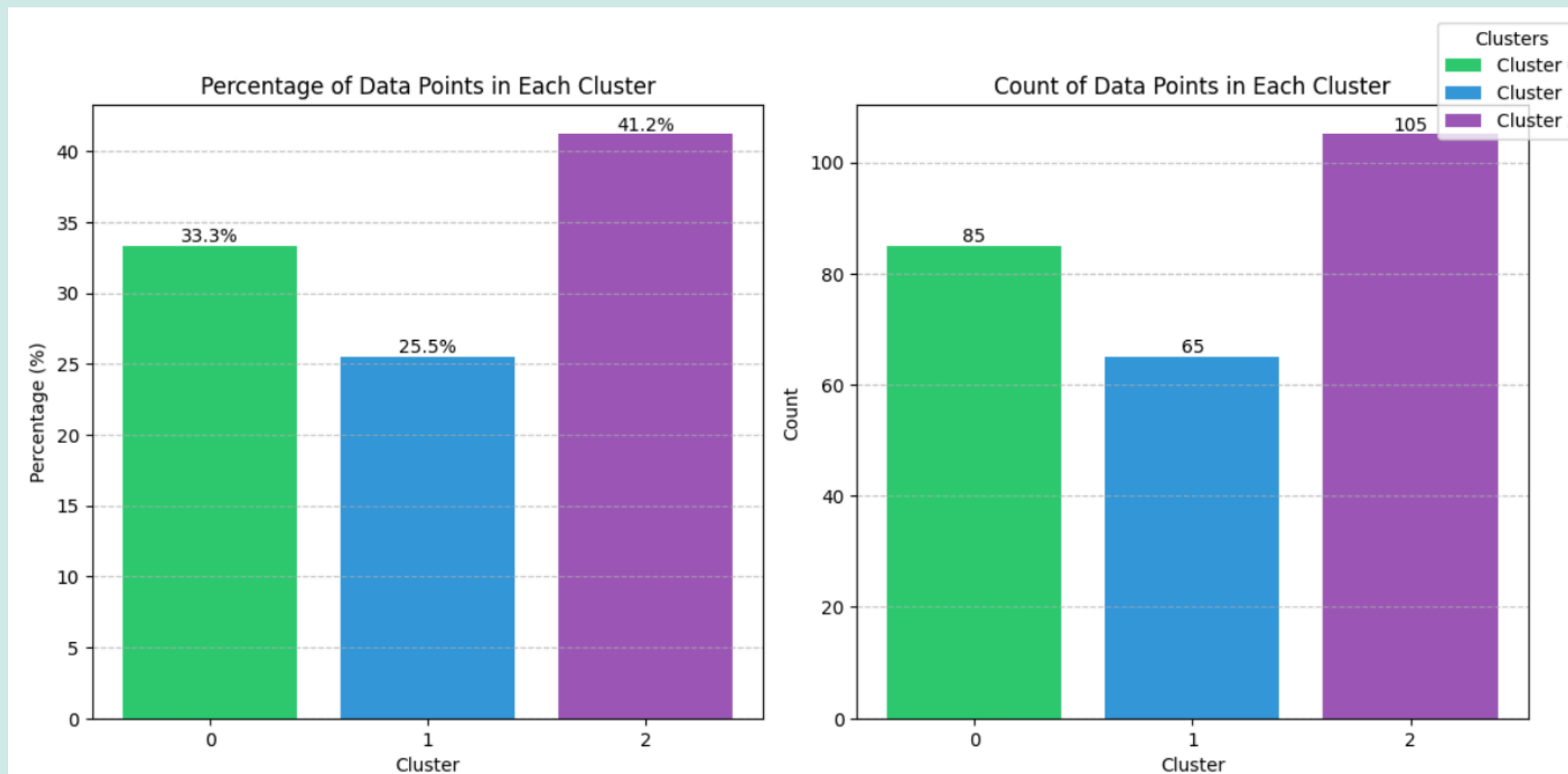
Method	Silhouette Score	Davies-Bouldin Index	Calinski-Harabasz Index	Inertia	t-SNE Visualization
PCA	0.355				nice
Kernel PCA (Sigmoid)	0.544	0.686	348.34	61.94	good
Kernel PCA (RBF)	0.508	0.868	303.32	11.51	great

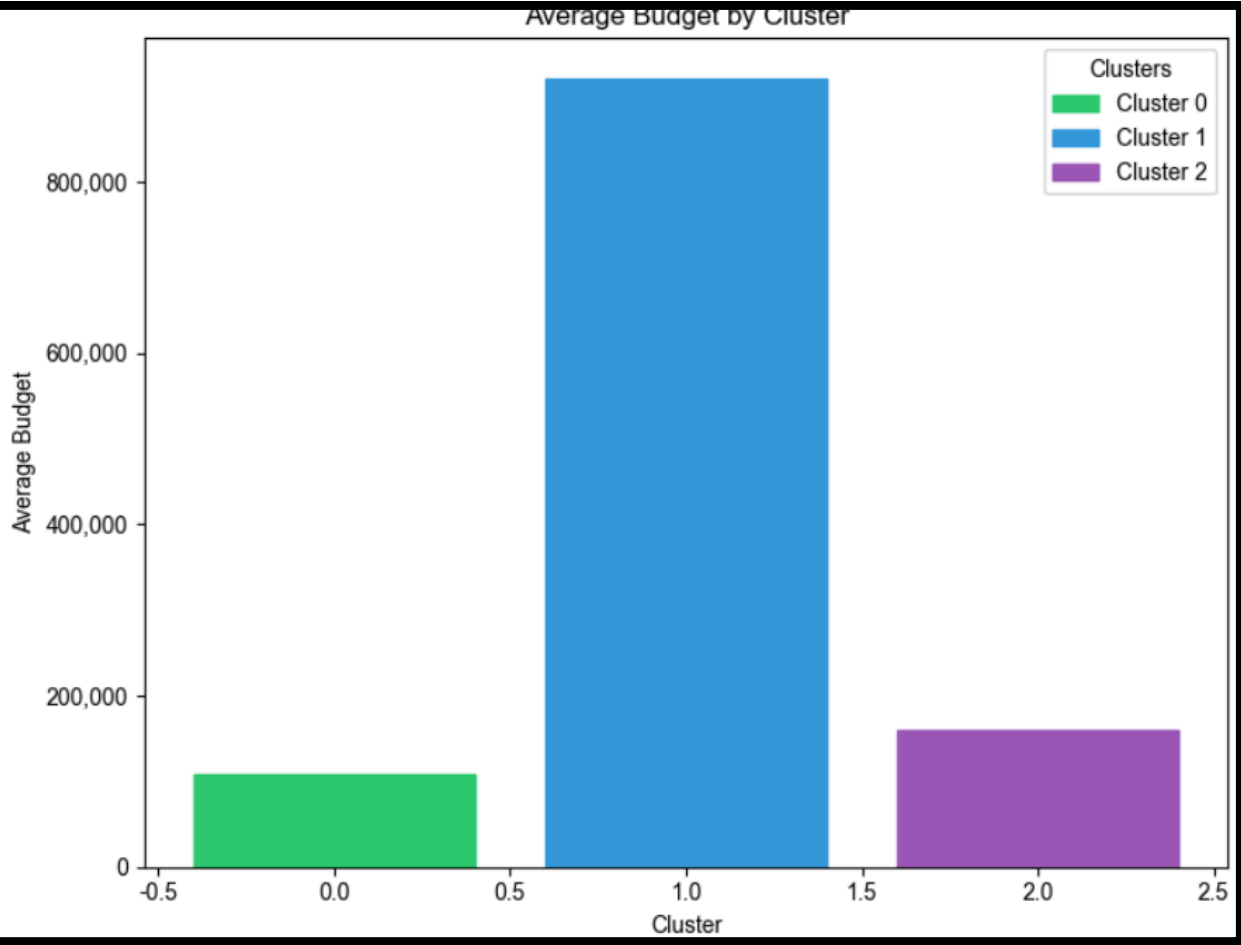
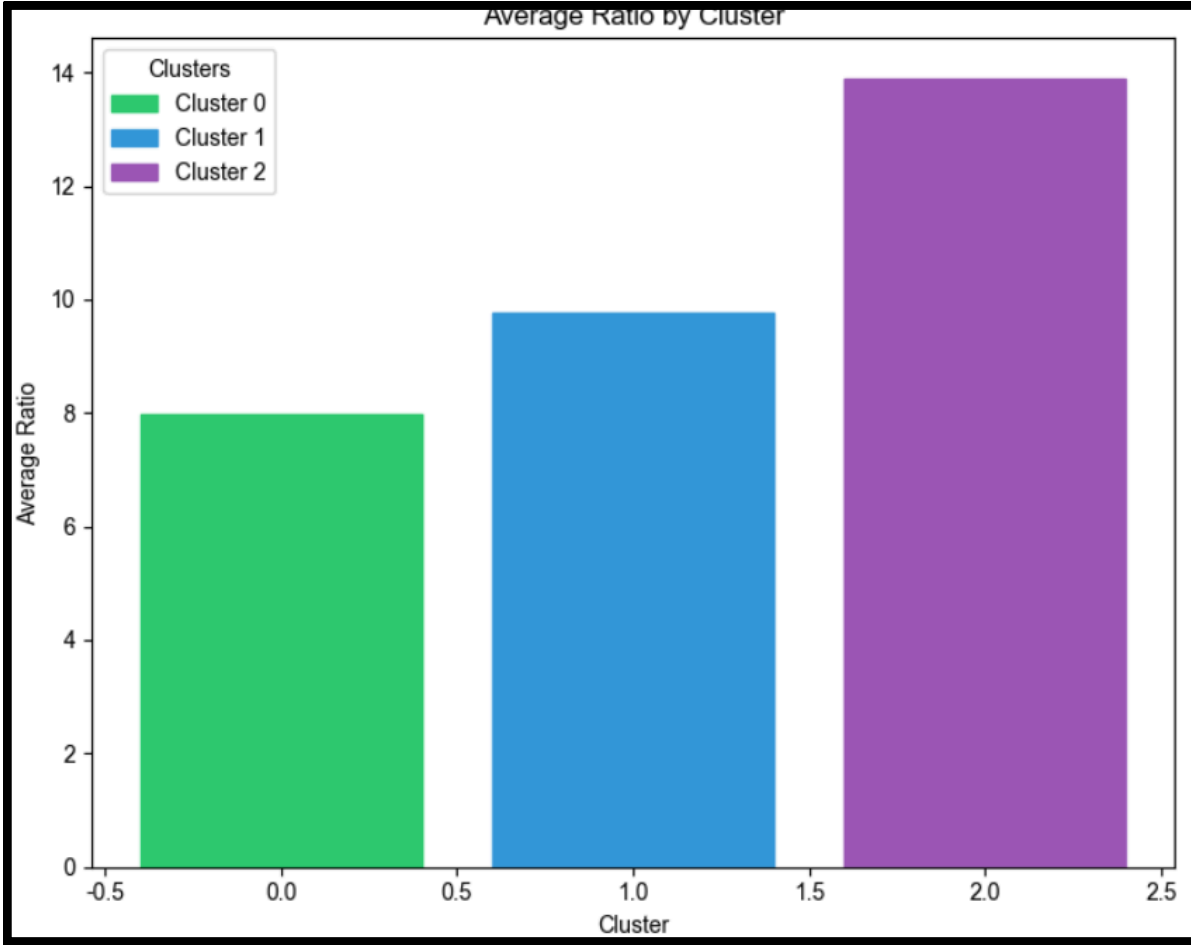
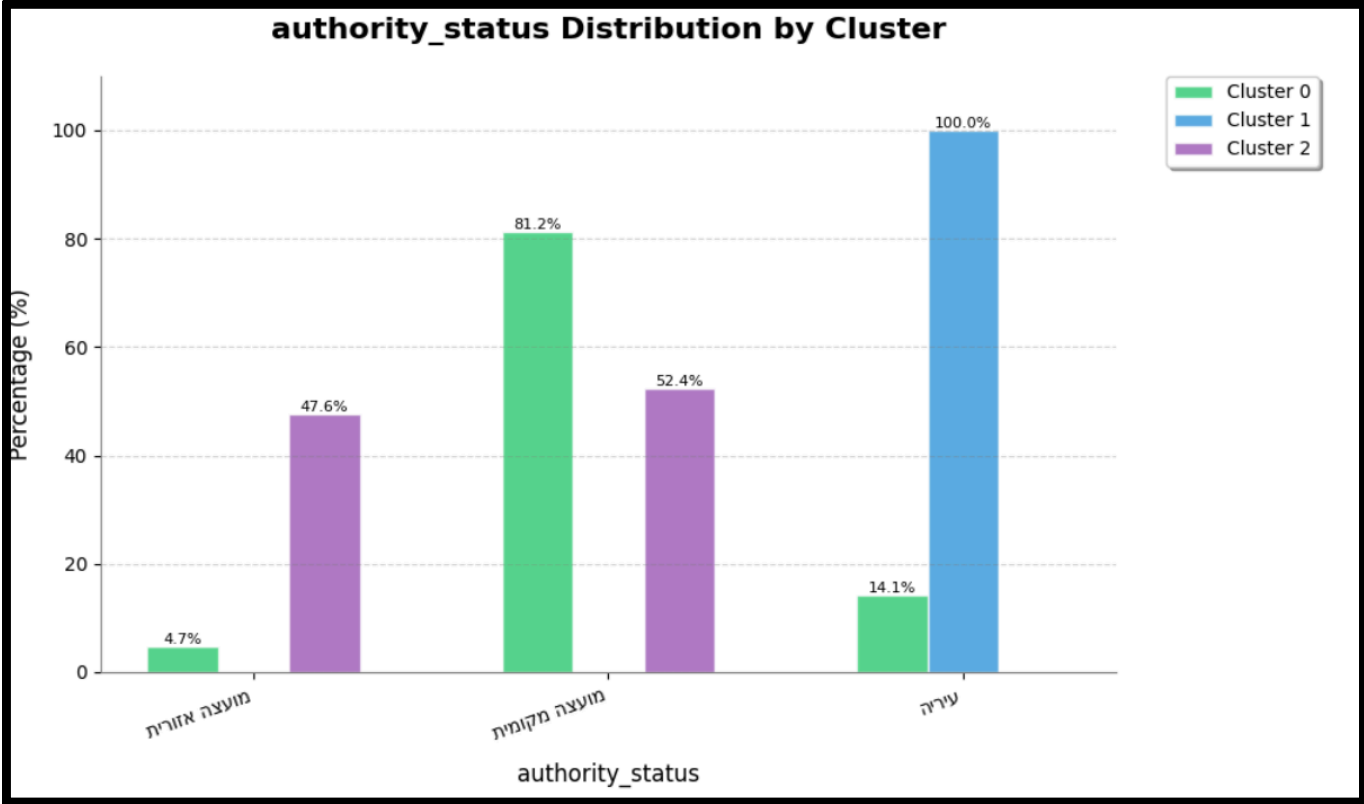
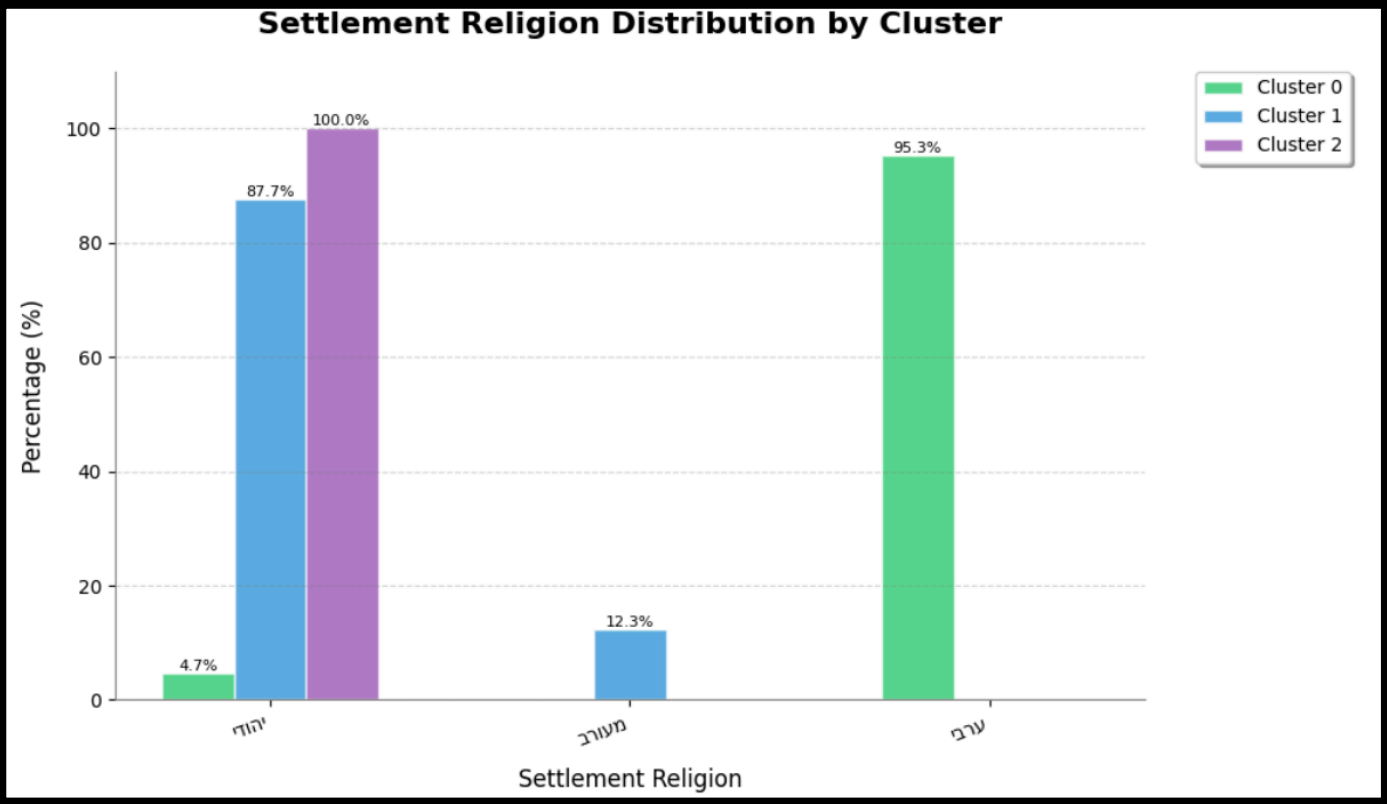
Results and Discussion

Model: Kernel PCA with RBF and clustering into 3 groups.

Justification:

- Best performance across quantitative metrics.
- Clear separation in TSNE visualizations.
- Robust interpretability of clustering results.





CONCLUSION

Key Findings:

- Cluster 0: Peripheral Arab municipalities face severe resource constraints and infrastructure challenges.
- Cluster 1: Urban mixed centers show growth but require infrastructure support.
- Cluster 2: Established Jewish municipalities demonstrate stability and efficiency.

Policy Implications:

- Target aid for under-resourced Cluster 0.
- Focus on infrastructure for growing Cluster 1.
- Maintain stability in Cluster 2.

Insight:

- This clustering analysis highlights resource disparities and development needs, offering a clear roadmap for informed policy-making.

Thank You !