

## **The Use Of K-Means Clustering Algorithm to Determine Potential Region for Zakat Distribution Based on Social and Economic Data in Indonesia 2023**

**Farah Dibah<sup>1</sup>, Ilvinda Rezki<sup>2</sup>**

<sup>1</sup>Sepuluh Nopember Institute of Technology

<sup>2</sup>Sriwijaya University

### **ABSTRACT**

*Poverty remains a significant challenge in many developing nations, including Indonesia. Poverty is also considered a problem caused by uneven economic growth. An effective approach to reducing poverty is through the equitable distribution of zakat, which can help narrow the disparity between the wealthy and those in need. Equitable distribution of zakat can be done by determining the level and potential of poverty in each province in Indonesia. To find out the poverty potential in each province in Indonesia, it is necessary to conduct a cluster analysis by looking at several poverty and economic growth variables. This study applies k-means clustering analysis to classify similar objects or individuals based on multiple criteria. The results reveal three clusters: provinces with high poverty and low economic levels (mainly in Java), provinces with moderate conditions, and provinces with low poverty and strong economies (mostly in eastern Indonesia). These findings highlight the potential for zakat institutions and government bodies to target zakat distribution more effectively, prioritising high-poverty clusters while promoting economic balance and welfare across regions.*

**Keywords:** *Poverty, Zakat, K-means Clustering, Cluster Analysis.*

### **INTRODUCTION**

Poverty remains one of the biggest problems for developing countries around the world. Poverty is also regarded as a development issue arising from the adverse effects of uneven economic growth, which exacerbates income disparities among individuals and across regions [Harahap, 2006].

According to the Central Bureau of Statistics (BPS), the poverty rate in Indonesia reached 9.3% as of March 2024. Although this figure has decreased slightly from the previous year. In 2024, Indonesia remains ranked 90th in terms of the highest number of people living in poverty [Statistik, 2023].

In Islam, one way to alleviate poverty is with zakat. Islamic jurisprudence mandates that Muslims allocate a prescribed portion of their

wealth, known as Zakat, to designated beneficiaries in accordance with established religious parameters. When viewed from the macro level, zakat not only aims to purify the wealth of muzaki and help the economic needs of muzaki, zakat can also be used as an instrument to reduce poverty. Based on national zakat collection data from 2018 to 2023, the nominal zakat collection always increases from year to year with an average growth of 39% [BAZNAS, 2024]. The progressive annual increment in zakat contributions is anticipated to serve as a complementary mechanism in supporting the Indonesian government's poverty alleviation initiatives.

Based on Law Number 23 Year 2011, BAZNAS (Badan Amil Zakat Nasional) functions as an autonomous governmental body

vested with the authority to orchestrate and administer zakat management operations at the national level. BAZNAS implements a systematic allocation framework for zakat disbursement predicated on hierarchical priorities, while adhering to fundamental principles of equitable distribution, social justice, and geographical considerations [BAZNAS, 2024]. In this case, BAZNAS has the authority to distribute zakat according to the economic and population conditions in an area. To find out the feasibility of a distribution area, it can be done by clustering.

Cluster analysis constitutes a statistical methodology employed to categorize and aggregate entities or subjects based on multiple predetermined parametric criteria into homogeneous groupings. Cluster analysis is performed by grouping similar elements as research objects into different and mutually exclusive groups [Sari, 2024].

To divide regions in Indonesia according to their poverty level, this study conducted a clustering analysis by dividing each province into groups with similarities in economic growth variables.

This research employs dual clustering methodologies specifically partitioning and hierarchical approaches. The partitioning methodology necessitates a priori determination of cluster quantities and centroid selection, whereas the hierarchical methodology

implements a sequential, structured algorithmic process.

## METHODOLOGY

This research is a quantitative method to determine the cluster in the distribution of zakat, the data used in this study is secondary data obtained from the Central Bureau of Statistics known as BPS with variables that include the poverty rate, namely the Number of Poor People (X1), Poverty Depth Index (X2), Child Mortality Rate (X3), Life Expectancy Rate (X4), and economic variables include Average income (X5), Inflation per Province (X6), GDP Per Capita (X7). The research method used is cluster analysis with k-means. This method is used to determine clusters and map provinces in Indonesia to a particular cluster according to the poverty rate and economy in the province.

## RESULT AND DISCUSSION

The results of this study include descriptive statistics, principal component analysis (PCA), k-means clustering model, and cluster analysis.

### *Descriptive Statistics*

A method used to collect, classify, and present data in a descriptive or explanatory. The Descriptive Statistics results of each variable are shown in the table below.

Table 2. Descriptive Statistics

Index	Number of Poor Population (Thousand)	Index of Poverty Depth	Child Mortality Rate	Average Income (Thousand Rupiah)	Inflation	Life Expectancy Rate	GDP Per Capita (Rp)
mean	761.8235	1.625294	3.848235	1865.199	2.09	70.75191	81948.41
std	1063.095	1.007365	2.204561	444.8451	0.62296	2.422573	61906.96
min	48	0.5	1.64	934.3	1.02	66.055	23078
median	346.5	1.31	2.98	1778.15	2.09	70.8925	64109.5

max	4189	4.35	10.88	2776.1	4.29	75.215	322615
-----	------	------	-------	--------	------	--------	--------

Table 2 shows that the average number of poor people from 34 provinces in Indonesia is 761 thousand people per province with the highest number of poor people in the province of East Java with 4.1 million poor people. The average poverty index in Indonesia is 1.62 with an average child mortality rate in Indonesia of 3.84 and an average life expectancy of 70.75. The average income of the Indonesian population is 1.8 million with the lowest income of 934 thousand rupiah in the province of East Nusa Tenggara and the highest income of 2,7 million in East Kalimantan Province. The average inflation per province is 2.09%, as many as 50% of provinces in Indonesia with GDP per capita of 81 thousand rupiah with the lowest inflation value of 1.02% in Bangka Belitung Islands Province.

*Data Distribution*

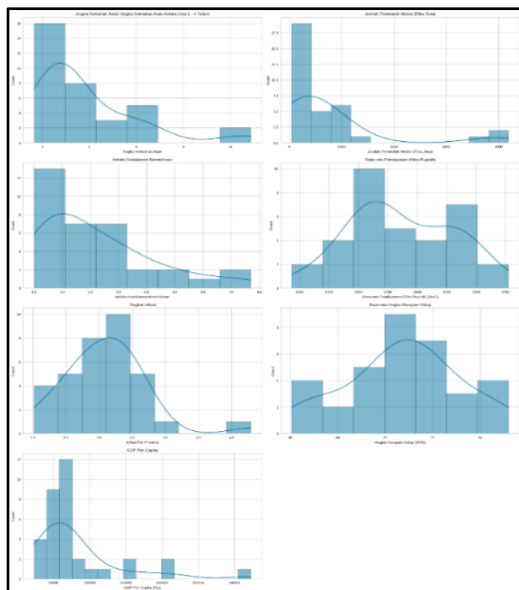


Figure 1. Data Distribution

Figure 1 shows that there are potential outliers in the variables of child mortality rate, number of poor people, poverty depth index, inflation per province, and GDP per capita. Meanwhile, the average income and life

expectancy variables have normally distributed data and there are no suspected outliers in the data. In this study, outlier data will not be eliminated as the purpose of this study, outlier data can be one indication that a province has poverty inequality compared to other provinces and requires special attention to reduce poverty in the province, one of which is by equalizing zakat, infaq and sadaqah.

*Kendall Correlation*

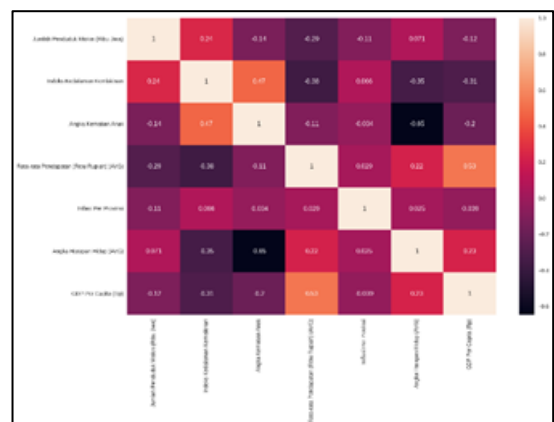


Figure 2. Kendall Correlation

Figure 2 shows that there is a high correlation between life expectancy and child mortality with a correlation value of -0.65. In this study, one of these variables will not be eliminated because each variable has a uniqueness that is important in conducting Cluster analysis. However, the data will be scaled to reduce bias due to the high distance and variance between variables. Scaling the data is done by transforming the data into a normal distribution using the minmax method.

*Principal Component Analysis (PCA)*

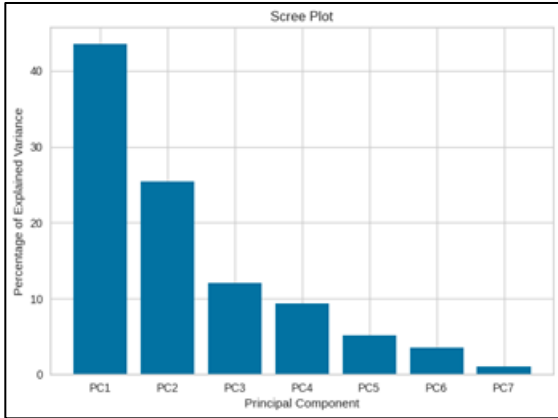


Figure 3. Principal Component Analysis (PCA)

Figure 3 is a Principal Component Analysis (PCA). It shows that component 1 and component 2 explain 68.9% of the variance of all components. In this study, the number of components that will be used to perform Cluster analysis is 3 components. The 3 components represent 82% of the variance of all components. This number of components will be used to create a k-means model.

*K-Means Clustering Model*

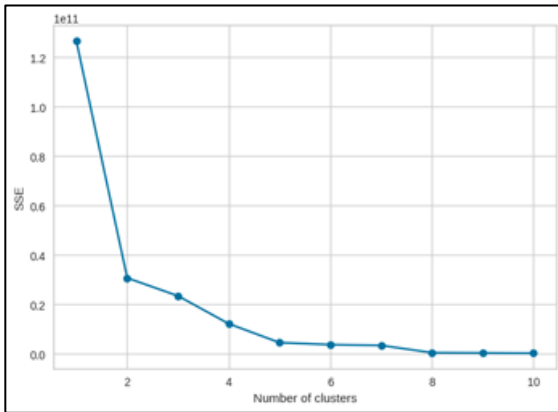


Figure 4. K-means Clustering Model

Figure 4 shows that the effective number of clusters is 3 clusters according to the number of components in the PCA analysis. The effectiveness of the number of clusters can be seen that there is a significant decrease in the first to third clusters. After the third Cluster, it can be seen in the figure that the Sum of Square Error (SSE) value tends to be constant so that

there is a similarity between one Cluster and another Cluster.

*Cluster Analysis*

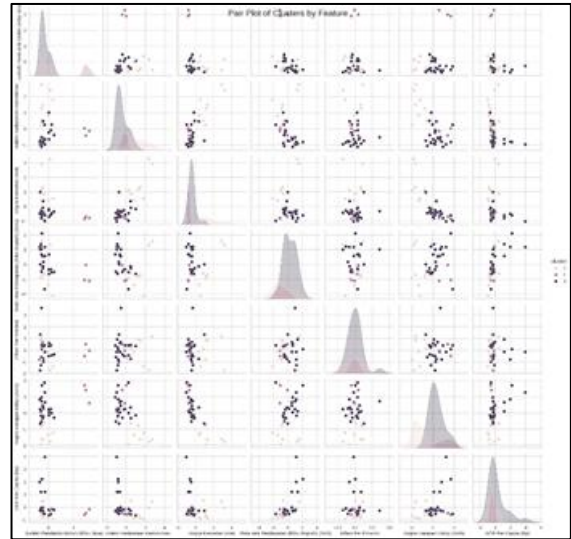


Figure 5. Cluster Analysis

Figure 5 shows that Cluster 0 is a cluster that generally has high GDP per capita, low poverty indices, and high life expectancy. Cluster 1 has extreme outlier values in several matrices such as a high number of poor people, a fairly low average income, a high poverty depth index in some data points, and a fairly high child mortality rate in some data points. While Cluster 2 is the Cluster with the largest number of provinces where this Cluster is a Cluster with matrix values that are on the center line. From Figure 5, it can be concluded that the Cluster division is as follows:

1. First Cluster (Cluster 0)

This Cluster is the best cluster with a low poverty rate and high economic value. The provinces included in this Cluster can be seen in table 3 below.

Table 3. First Cluster

No	First Cluster (Cluster 0)
1	West Nusa Tenggara
2	East Nusa Tenggara

3	Central Sulawesi
4	Gorontalo
5	West Sulawesi
6	Maluku
7	West Papua
8	Papua

Table 3 shows that Cluster 0 consists of 8 provinces in Indonesia. The 8 provinces are West Nusa Tenggara, East Nusa Tenggara, Central Sulawesi, Gorontalo, West Sulawesi, Maluku, West Papua, and Papua.

## 2. Second Cluster (Cluster 1)

This is a cluster with a high poverty rate and low economic value. This cluster has a high population with a high poverty rate. The provinces in this Cluster require further attention and analysis to determine the factors influencing the high poverty rate and low economic level. The provinces in the second cluster are shown in table 4 below.

Table 4. Second Cluster

No	Second Cluster (Cluster 0)
1	East Java
2	Central Java
3	West Java

Table 4 shows that the second cluster consists of 3 provinces. The provinces included in this cluster are West Java, Central Java, and East Java.

## 3. Third Cluster (Cluster 2)

The third cluster is a cluster with poverty and economic values that represent the midpoint or close to the average value. The provinces in this cluster are shown in the table 5 below.

Table 5. Third Cluster

No	Third Cluster (Cluster 2)
1	Aceh
2	North Sumatera
3	West Sumatera

4	Riau
5	Jambi
6	South Sumatera
7	Bengkulu
8	Lampung
9	Bangka Belitung Island
10	Riau Island
11	DKI Jakarta
12	DI Yogyakarta
13	Banten
14	Bali
15	West Kalimantan
16	Central Kalimantan
17	South Kalimantan
18	East Kalimantan
19	North Kalimantan
20	North Sulawesi
21	South Sulawesi
22	Southeast Sulawesi
23	Maluku Utara

Table 5 shows that the third cluster results consist of 23 provinces namely Aceh, North Sumatera, West Sumatera, Riau, Jambi, South Sumatera, Bengkulu, Lampung, Bangka Belitung Islands, Riau Islands, DKI Jakarta, D.I. Yogyakarta, Banten, Bali, West Kalimantan, Central Kalimantan, South Kalimantan, East Kalimantan, North Kalimantan, North Sulawesi, South Sulawesi, Southeast Sulawesi, and North Maluku.

The results of the cluster analysis above can be visualized with the geo chart shown in Figure 6 below.



Figure 6. Cluster Visualization

Figure 6 above shows that the eastern and central parts of Indonesia tend to be in the first cluster and the third cluster. While the western part of Indonesia, especially Java, is dominated by the second cluster and the western part of Indonesia, especially Sumatra and Kalimantan, is dominated by the third cluster.

The specific management and distribution of zakat is in line with the grand design of the 2025-2045 road map that is being compiled by the government with one of its

## CONCLUSION

One of the most crucial aspects of zakat management is its distribution, as it directly determines the effectiveness of zakat in achieving social welfare. The results of this study show that provinces on the island of Java, specifically East Java, Central Java, and West Java, form a high-poverty cluster with limited economic performance despite their large populations. Therefore, zakat institutions such as BAZNAS should prioritize these provinces as the main targets for zakat distribution and empowerment programs. Efforts should focus on strengthening local economic productivity through skills training, microenterprise funding, and community-based economic initiatives to reduce poverty sustainably. Provinces in the medium cluster should receive preventive

goals being community welfare. BAZNAS, which is a zakat institution under the auspices of the government, has a role to harmonize data and methodology in mapping zakat. This can be done by collaborating and harmonizing the data owned by BAZNAS with data owned by other government institutions such as the Central Statistics Agency (BPS), the Ministry of Social Affairs, or the Ministry of Economy. With the collaboration and harmonization of data, it is expected that the methodology in the distribution of zakat can help the government in equalizing the Indonesian economy and achieving community welfare. From the results of the research that has been done by considering social and economic variables, the distribution of zakat that has been done so far has been quite right on target but to maximize equitable distribution it is necessary to do further analysis of the economic conditions of each field in a province.

assistance such as entrepreneurship support and education programs to enhance economic resilience. Meanwhile, wealthier provinces in the low-poverty cluster should be encouraged to increase zakat collection and donations to support redistribution across poorer regions. This targeted, data-driven approach can ensure that zakat contributes effectively to reducing poverty and promoting equitable economic growth throughout Indonesia.

## REFERENCES

- Agusta, Y. (2007). K-Means Penerapan, Permasalahan, dan Metode Terkait. *Jurnal Sistem dan Informatika* Vo. 3, 47-60.
- BAZNAS. (2023, 10 2). Get to know 8 Asnaf Zakat. Retrieved from Badan Amil Zakat

- Website: <https://baznas.go.id/artikel-show/Mengenal-8-Asnaf-Zakat>.
- BAZNAS, P. K. (2024). *Outlook Zakat 2024*. Jakarta: Pusat Kajian Strategis BAZNAS.
- Benbrahim, S., Chaouèche, N. S., & Toumache, R. (2022). Countries Economic Segmentation using K-Means Clustering for the Year 2021. *Economics Researcher Journal* Vol. 9, 512-528.
- Buslim, N., Iswara, R. P., & Agustian, F. (2021). The Modeling of Mustahiq Data Using K-Means Clustering Algorithm and Big Data Analysis (Case Study: LAZ). *Jurnal Teknik Informatika*, 213-230.
- Gudono. (2014). *Analisis Data Multivarit Edisi Tiga*. Yogyakarta: BPFE Yogyakarta.
- Hadi, M. (2010). *Problematika Zakat Profesi dan Solusinya*. Yogyakarta: Pustaka Pelajar.
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining Concepts and Techniques 3rd Edition*. Massachusetts: Morgan Kaufmann.
- Harahap, S. S. (2006). *Critical Analysis of Financial*. Jakarta: PT. Raja Grafindo Perkasa.
- James, G. (2019). *Introductions to Macroeconomics*. Colorado: Gilad James Mystery School.
- Johnson, R. A., & Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis (6th Ed)*. New Jersey: Prentice International Inc.
- Kartikasari, E. (2006). *Introduction to Zakat and Waqf*. Jakarta: PT Grasindo.
- Monica, I. S., & Abidah, A. (2021). Konsep Asnaf Berdasarkan Pemikiran Yusuf Al-Qardawi dan Al-Zuhayli. *Jurnal Antologi Hukum*, 109-124.
- Mufraini, A. (2005). *Zakat Management and Accountant*. Jakarta: Pernada Media.
- Purba, B. (2020). Analisis Pertumbuhan Ekonomi Indonesia Periode 2009 - 2018. *Jurnal Humaniora* Vol. 4, 244-255.
- Rachmatin, D. (2014). Aplikasi Metode-Metode Agglomerative dalam Analisis Klaster pada data Tingkat Polusi Udara. *Jurnal Ilmiah Studi Matematika STKIP Siliwangi Bandung*, 133-149.
- Sari, R. M. (2024). *Clustering Method for Dangerous Disease*. Sumatera Barat: PT Serasi Media Teknologi.
- Statistik, B. P. (2023, July 17). BPS: Berita Resmi Statistik. Retrieved from Profil Kemiskinan di Indonesia Maret 2023: <https://www.bps.go.id/id/pressrelease/2023/07/17/2016/profil-kemiskinan-di-indonesia-maret-2023.html>
- Turnando, G., & Zein, A. S. (2019). Analisis Pengaruh Zakat Terhadap Peningkatan Kesejahteraan Mustahiq. *Al-Masharif: Jurnal Ilmu Ekonomi dan Keislaman*, 162-175.