

Multidimensional Cluster Analysis to Evaluate the Unevenness of Social Welfare Decile Change Policy in Indonesia: A Data Science Approach

Rizal Purnama¹, Aldi Algino², Septien Dwi Savandha³

^{1,2} STMIK IKMI Cirebon, Indonesia

³Universidad Tecnológica Latinoamericana en Línea (UTEL), Mexico City, Mexic

*Corresponding author: alginoaldi88@gmail.com

Keywords:

data science; K-Means clustering; welfare decile; social assistance policy

ABSTRACT

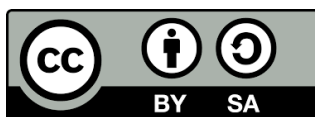
Background: Social assistance targeting policy in Indonesia relies on a decile system within the Integrated Social Welfare Data (DTKS) to classify community welfare levels, yet this classification is frequently criticized for producing uneven exclusion and inclusion errors across regions.

Objective: This study evaluates that criticism from a data science perspective. Five province-level socioeconomic indicators from Statistics Indonesia (BPS) for 2021, namely the Human Development Index, poverty percentage, Gini ratio, open unemployment rate, and GRDP growth rate, were collected for 32 Indonesian provinces.

Method: The data were standardized and grouped using the K-Means algorithm, with the optimal number of clusters determined through the elbow method and silhouette score.

Results: The resulting multidimensional cluster classification was then compared with a single-indicator poverty-based quintile classification, used as a simplified proxy for the single-indicator logic of the government decile, not as a reconstruction of the decile itself. The results show that five clusters produced the highest silhouette score (0.2802), and 59.38% of provinces were classified differently between the two approaches, with eight provinces experiencing a severe shift of two levels or more. The Adjusted Rand Index between the two classifications was only 0.0521, indicating very low agreement.

Conclusion: These findings demonstrate substantial disagreement between multidimensional and poverty-based classification; they do not observe targeting errors, since household-level data were not accessible. The study is a province-level proof of concept; provincial aggregates cannot establish household-level exclusion or inclusion error. Multidimensional clustering may be considered as a complementary analytical tool for the further validation of welfare targeting. **Keywords:** data science; K-Means clustering; welfare decile; social assistance policy



This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

Copyright © 2026 by Author. Published by Pusat Publikasi Nusantara

INTRODUCTION

Mis-targeting in social assistance is not a distinctively Indonesian problem. Wherever governments must identify poor households without reliable income records, they fall back on a proxy means test, which predicts household consumption from observable assets and

demographic characteristics and then applies a cut-off. The method is administratively practical and has been adopted across low and middle income countries, but its accuracy has been contested for as long as it has been in use. Because the prediction explains only part of the variance in actual consumption, a substantial share of poor households falls above the cut-off while a substantial share of non-poor households falls below it, so exclusion and inclusion error are produced simultaneously rather than traded off against one another. Evidence from nine African countries shows that a conventional proxy means test misses a large fraction of the poorest households even when the model is well specified (Brown et al., 2018), and targeting error rises further once measurement error in the underlying consumption data is taken into account (Gazeaud, 2020). The scale of what is at stake is set out by Hanna and Olken (2018), who compare targeted transfers with universal alternatives using nationwide programme data from Indonesia and Peru. The error is not random noise: it is structured by whatever the proxy happens to measure, so regions and households whose deprivation takes a form the instrument does not capture are systematically disadvantaged. A parallel literature makes the same point from the measurement side, arguing that poverty consists of simultaneous deprivations across health, education, and living standards that no single monetary threshold can represent (Alkire & Foster, 2011), an argument that recent Indonesian applications have taken to the provincial and regional level (Pratiwi & Joshi, 2025; Wibisono & Sirait, 2024). Indonesia's decile system sits inside both debates at once: it is a proxy means test applied at national scale, and it compresses a multidimensional condition into a single ordering.

Poverty alleviation through social assistance programs is one of the main priorities of the Indonesian government. To ensure that aid is distributed on target, the government uses the Integrated Social Welfare Data (DTKS), which is now transitioning to the National Social and Economic Single Data (DTSEN), as the main database of beneficiaries. This system classifies households into ten groups or deciles based on welfare level, with deciles 1-4 prioritized as recipients of assistance (Dinas Sosial Provinsi Jawa Tengah, n.d.). The decile classification is updated regularly using the Proxy Mean Test approach jointly managed by the Central Statistics Agency (BPS) and the Ministry of Social Affairs.

Although this mechanism is designed to be more objective, various reports show that decile changes often do not reflect the real economic conditions of the beneficiaries, resulting in exclusion errors and inclusion errors that are uneven between regions (Achmadi, 2024). Purbaningrum and Adinugraha (2024) found that the inaccuracy of social assistance targets at the village level is still a recurring problem that requires more accurate data-based policy transformation. This kind of inequality is important to be studied not only from the point of view of public policy, but also from the perspective of informatics science, considering that the decile system is basically a problem of multidimensional data classification that can be analyzed using a data science approach.

Various informatics studies have applied clustering algorithms, especially K-Means, to group regions in Indonesia based on poverty and welfare levels. Bahauddin et al. (2021) grouped provinces in Indonesia based on poverty levels using K-Means and found that provinces in eastern Indonesia are consistently in high poverty clusters. Nasution et al. (2020) and Istiqamah et al. (2021) show that K-Means is effectively used to group poor people according to administrative regions with fairly good accuracy, while Niswatin (2021) applies a similar approach at the district/city level. More specifically, Zuhendra et al. (2024) applied

data mining to clustering poverty levels based on the DTKS data itself, and Rady Putra and Anggrawan (2021) grouped community social assistance recipients using the K-Means method, both of which emphasized the direct relevance of the clustering approach to the problem of social assistance targeting. Ahmed et al. (2020) in their comprehensive review emphasized that K-Means remains one of the most widely used clustering algorithms due to its simplicity and interpretability, although it is sensitive to the selection of cluster numbers and data scales. Beyond the Indonesian informatics literature, the methodological choices made in this study have an established basis. The silhouette index used here to select the number of clusters was introduced by Rousseeuw (1987), and comparative work on cluster validity indices cautions that no single index is decisive on its own (Arbelaitz et al., 2013). Alternatives to K-Means that relax its assumption of spherical, comparably sized clusters are well developed, including K-Medoids (Park & Jun, 2009), model-based Gaussian mixtures (Fraley & Raftery, 2002), Ward hierarchical agglomeration (Murtagh & Legendre, 2014), and density-based methods (Schubert et al., 2017); Andayani et al. (2025) compare two of these directly on Indonesian provincial welfare indicators drawn from the same BPS series used here.

Although there is a considerable amount of informatics literature on poverty clustering, most of these studies stop at the regional grouping stage without explicitly comparing the results with the single-indicator classification system that is the basis of government policies, such as the DTKS/DTSN decile system. There have not been many studies that quantitatively measure the potential mismatch between multi-indicator-based welfare classifications and single-indicator-based classifications of poverty, even though these differences directly have implications for the equitable distribution of social assistance between regions. Therefore, this study aims to (1) construct a classification of provincial welfare levels in Indonesia using K-Means clustering based on five socio-economic indicators, and (2) compare the results of these classifications with quintile classifications based on single poverty indicators to quantitatively measure the level of inequality that has the potential to arise due to decile determination approaches that rely on only one welfare dimension. These two objectives are measurable: the first produces a cluster assignment for each province, and the second produces a mismatch rate, a count of severe shifts, and an index of agreement between the two classifications.

The urgency of the question can be put in figures rather than in principle. Simulating the targeting of a major Indonesian cash transfer programme, Sartirano et al. (2023) report an exclusion error of 32.82 percent, meaning that roughly one in three households that should have qualified did not appear among the intended recipients. Field studies of the operating database describe the mechanisms behind such figures: deceased recipients still carried on the list, migration not reflected in records, and comparatively well-off households retained across update cycles (Wijayanti et al., 2025). The consequences are not confined to administrative tidiness, since crisis-responsive transfers in Indonesia have measurable effects on poverty when they reach the intended households and none when they do not (Haryaningsih et al., 2024). With the transition from DTKS to DTSN now under way, the classification rule that decides who is counted as deserving is being rebuilt, which makes this the moment at which evidence about the rule itself is most useful.

The study offers one practical and one methodological contribution, which should be distinguished. Practically, it gives policymakers a quantitative estimate of how much a welfare ranking changes when the definition of welfare is widened from one dimension to five, and it

supplies a cheap, reproducible diagnostic that can be run on public aggregates before any decision is taken about household-level instruments. For the agencies rebuilding DTSEN, that diagnostic indicates where a single-indicator rule is most likely to disagree with a broader one, and therefore where validation effort is best spent. Methodologically, the study contributes a comparison design rather than another clustering result. The Indonesian poverty clustering literature has established repeatedly that regions can be grouped by socioeconomic indicators; what it has not done is treat the government's own classification logic as a second partition and measure the agreement between the two with a formal index. Framing the question as the comparison of two partitions, rather than as the evaluation of one, is transferable to any targeting system that reduces a multidimensional condition to a single ranking, and it is the element of this study most readily reused elsewhere.

METHODS

This study used a data science-based quantitative approach with an unsupervised clustering method. The data were sourced from official data published by the Central Statistics Agency (BPS) in 2021 and included five provincial-level socioeconomic indicators: the Human Development Index (HDI), percentage of the poor population, Gini ratio, open unemployment rate (TPT), and Gross Regional Domestic Product (GRDP) growth rate. The year 2021 was selected because it was the most recent year for which complete data were available for all five indicators at the time of data collection. After the national aggregate data rows and rows with incomplete values were removed, the final dataset comprised 32 of Indonesia's 34 provinces. Two provinces were excluded because data for one of the indicators were unavailable for that year. The dataset is the published BPS provincial series (Central Statistics Agency, 2023). The five indicators were chosen to span the dimensions that the multidimensional poverty literature treats as distinct: human capability through the Human Development Index, income deprivation through the poverty percentage, distribution through the Gini ratio, labour market access through the open unemployment rate, and economic momentum through GRDP growth. They are also the only five indicators published annually by BPS for every province on a consistent definition, which makes the design reproducible by any reader from the public release. It should be noted that Indonesia had 34 provinces in 2021, the data year used here; the four provinces created in Papua during 2022 did not exist when these data were collected, so the figure of 32 of 34 is correct for the data year although the country now has 38 provinces.

Before clustering, all indicators were standardized using z-scores (StandardScaler) to ensure that they were measured on an equivalent scale. The direction of the three negative indicators—poverty percentage, Gini ratio, and unemployment rate—was first reversed so that, after standardization, higher values consistently represented higher levels of well-being across all indicators. The K-Means algorithm was then applied to the standardized data, with the optimal number of clusters determined using a combination of the elbow (inertia) and silhouette score methods for values of k ranging from 2 to 7. The final clustering solution consisted of five groups, in accordance with the highest silhouette score. The clusters were labeled from 0 (lowest level of well-being) to 4 (highest level of well-being) based on the ordering of the average composite welfare scores for each cluster. Equal weighting was applied deliberately rather than by default, and its limits should be stated. The five indicators are not independent: the Human Development Index and the poverty percentage are strongly and negatively

correlated by construction, since both are driven in part by household consumption, and an unweighted z-score composite therefore gives the income dimension more influence than the single poverty variable would suggest. Equal weighting was nevertheless retained for three reasons. It imposes no prior judgement about which dimension of welfare matters more, which is the judgement the single-indicator approach makes implicitly and which this study is examining rather than replacing. It keeps the procedure transparent and reproducible, since any alternative weighting would require an external normative source. And it is the convention in the comparable clustering literature on Indonesian provincial poverty, which makes the results comparable with earlier work. A correlation-aware alternative, such as clustering on principal components or applying a correlation-adjusted distance, is a natural robustness check and is noted in the limitations as a direction for further work.

For comparison, a quintile classification based on a single poverty indicator was also constructed by dividing the 32 provinces into five groups of approximately equal size according to the percentage of the poor population alone. This classification was used as a simple proxy for the logic of determining DTKS/DTSSEN deciles, which, in practice, also relies heavily on relatively limited household economic indicators through the Proxy Means Test. The two classification systems were then compared using three measures: (1) the number and percentage of provinces that exhibited differences in classification levels (mismatch); (2) the number of provinces that experienced severe level shifts, defined as a difference of two or more levels; and (3) the Adjusted Rand Index (ARI), a statistical measure of the degree of agreement between two grouping results, where a value close to 1 indicates a high level of agreement and a value close to 0 indicates agreement equivalent to random grouping. All analyses were conducted using Python 3 and the scikit-learn library, with a fixed random seed (42) to ensure the reproducibility of the results. Because the entire mismatch argument rests on this substitution, its validity limits are stated explicitly rather than left implicit. The quintile constructed here is a simplified proxy for a single-indicator welfare classification; it is not the DTKS/DTSSEN decile, and the analysis neither reconstructs nor audits that decile. Three differences matter. First, the official Proxy Means Test uses a regression-weighted index of many household assets and characteristics, not a single poverty rate, so the proxy here is a deliberately stricter version of single-indicator logic than the system actually applies. Second, the official system ranks households into ten deciles, whereas this study ranks provinces into five quintiles, chosen to match the number of clusters so that the two classifications are directly comparable. Third, and most importantly, the unit differs: provincial aggregates conceal the within-province heterogeneity on which household targeting operates. The proxy is therefore defensible for the narrow question asked here, namely whether ranking units by one welfare dimension produces a different ordering from ranking them by several, and it is not defensible as evidence about how the DTKS/DTSSEN decile performs in practice. Readers should treat the mismatch statistics below as a property of the two classification rules, not as a measured error rate of the government system.

The main limitation of this study was that the unit of analysis was at the provincial level rather than the household level, as in the actual DTKS/DTSSEN. Therefore, the findings should be interpreted as an illustration and proof of concept regarding potential inequalities arising from a single-indicator approach, rather than as a direct audit of the DTKS/DTSSEN database, which is not publicly accessible.

RESULT AND DISCUSSION

The analysis was carried out on 32 provinces in Indonesia using five socio-economic indicators in 2021. Figure 1 shows the results of the evaluation of the optimal number of clusters using the elbow and silhouette score methods.

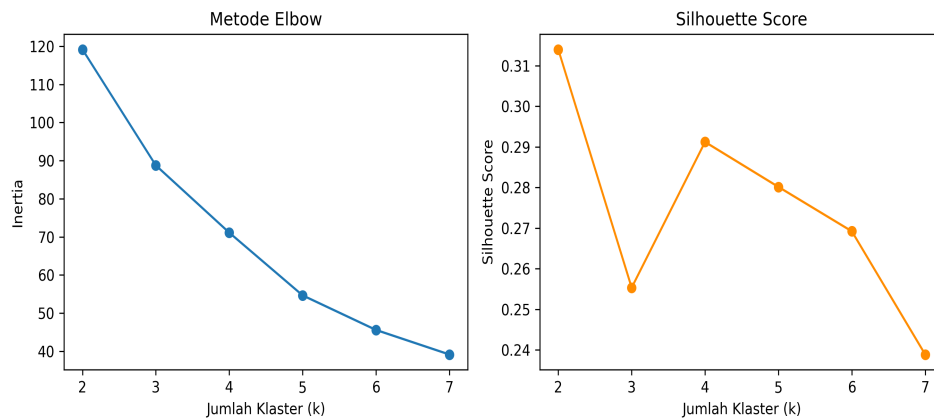


Figure 1. Elbow (left) and silhouette score (right) methods to determine the optimal number of clusters.

The highest silhouette score value was obtained at $k = 5$ (0.2802). A silhouette in the range of 0.25 to 0.30 is weak rather than moderate, and this should be read as a substantive result rather than explained away by precedent. On the interpretive bands proposed when the index was introduced (Rousseeuw, 1987), a value near 0.28 means that the average province sits only slightly closer to its own cluster centre than to the nearest competing one, so the five-cluster partition describes a continuum with soft boundaries rather than five well-separated types of province. Three consequences follow for the argument of this article. The first is that individual cluster assignments near a boundary are not robust, and no claim should be made about any single province on the strength of its label alone; the province-level illustrations below are offered as interpretation of the structure, not as verdicts on those provinces. The second is that the low agreement reported in Table 2 is, if anything, understated by the weak separation rather than produced by it: a poorly separated multidimensional partition that nonetheless disagrees almost completely with the poverty ranking indicates that the two rules order provinces on genuinely different grounds, since noise in cluster assignment would tend to move the Adjusted Rand Index toward, not away from, the value expected by chance. The third is that the weak separation is itself substantive evidence about the object of study: Indonesian provincial welfare does not fall into naturally discrete strata, which is precisely the assumption that any decile or quintile cut imposes on the data. Comparative work on validity indices cautions against reading any single index as decisive (Arbelaitz et al., 2013), which is a further reason to treat the silhouette as descriptive of structure rather than as a verdict on the solution. Where separation is weak, the appropriate response is to assess how stable the partition is under resampling rather than to rely on the index alone (Hennig, 2007). A classification system that draws hard boundaries across a continuum will place similar units on opposite sides of a threshold regardless of which indicator it uses, and that is a structural property of the ranking exercise, not an artefact of this particular clustering. Based on

considerations of interpretability and equality of the number of groups with the comparative quintile approach, $k = 5$ was chosen as the final cluster number.

Table 1 summarizes the characteristics of the data and the results of clustering, while Table 2 presents the results of a comparison between the multidimensional classification (K-Means) and the classification of the single poverty quintile.

Table 1. Summary of data and clustering results

Features	Value
Number of provinces analyzed	32
Data year	2021
Number of indicators	5 (HDI, Poverty, Gini, TPT, GDP)
Optimal cluster count (k)	5
Silhouette score (k=5)	0.2802

Table 2. Comparison of multidimensional classifications (K-Means) and single poverty quintiles

Comparison Size	Value
Provinces with mismatch	19 provinces (59.38%)
Provinces with severe shifts (≥ 2 -level difference)	8 provinces
Adjusted Rand Index (ARI)	0.0521

The results in Table 2 show that more than half of the provinces (59.38%) experienced differences in classification levels between the multidimensional clustering approach and the single poverty quintile approach, with eight provinces (25% of the total) experiencing severe level shifts, i.e. differences of two or more levels. The Adjusted Rand Index value of 0.052 was well below the threshold of good fit (generally above 0.5), which statistically suggests that the two classification systems result in almost uncorrelated clustering. These findings empirically support the argument that the use of a single poverty indicator, as the basic logic of the Proxy Mean Test in the DTKS/DTSEN decile system, risks resulting in a welfare classification that differs significantly from the real conditions of a region when viewed multidimensionally. The Adjusted Rand Index is reported here as 0.0521, which is the value given in Table 2; the figure of 0.052 used in earlier drafts was a rounding of the same quantity. One qualification should be attached to this comparison before it is interpreted. A near-zero Adjusted Rand Index establishes that the two rules produce orderings that are close to statistically independent, which is a strong result about the rules themselves. It does not establish that one ordering is correct and the other mistaken, because the study has no external criterion of true welfare against which either could be scored. The finding is therefore one of disagreement, and the policy weight it carries comes from the fact that a classification used to allocate public resources should not depend so heavily on which single defensible definition of welfare happens to be chosen.

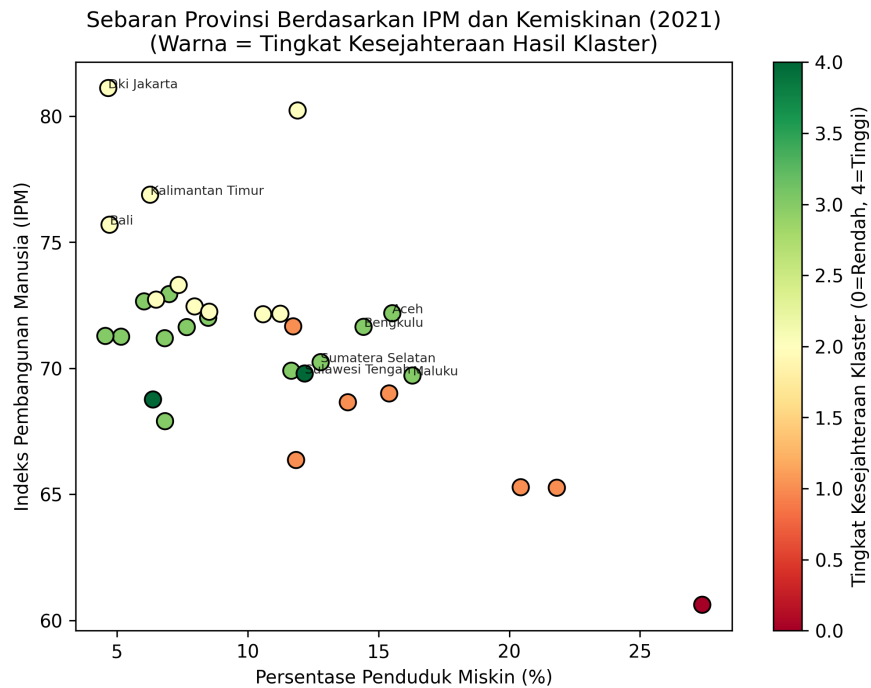


Figure 2. The provincial distribution based on HDI and poverty percentage in 2021 is colored by the level of welfare of cluster results.

Figure 2 shows that provinces with relatively low poverty rates do not always have a high level of cluster welfare, and vice versa. For example, Central Sulawesi has a medium poverty rate (12.18%) but is placed in the highest welfare cluster due to a very high GDP growth rate (11.70%), while based on the single poverty quintile this province is actually in the low welfare category. In contrast, the Special Region of Yogyakarta has a very high HDI but also a high Gini ratio (0.436), so multidimensional clustering places it at a medium welfare level, two levels different from the classification of the single poverty quintile which places it at a low level.

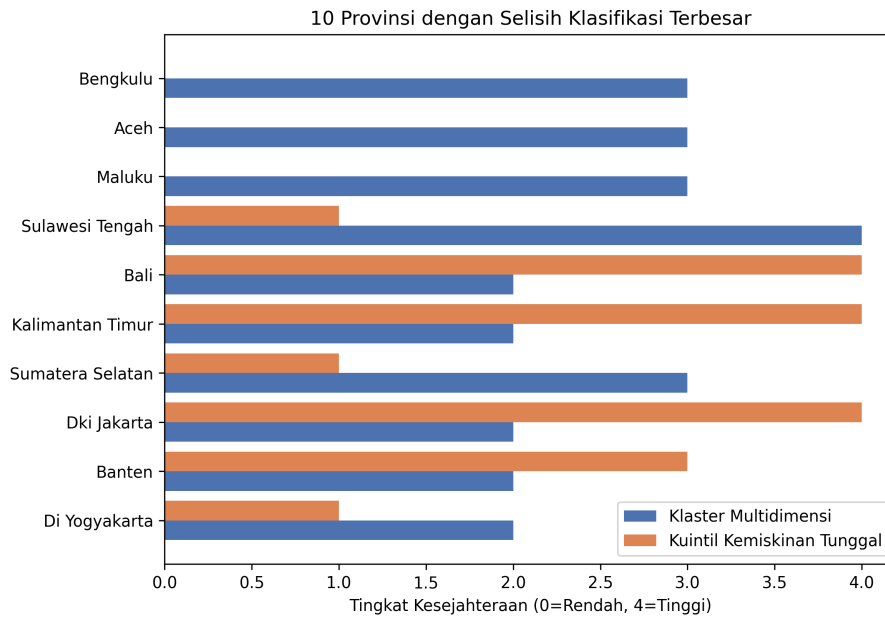


Figure 3. The ten provinces with the largest classification difference between the multidimensional cluster and the single poverty quintile.

Figure 3 highlights the ten provinces with the largest classification differences. Provinces such as Aceh, Bengkulu, and Maluku are rated as having low welfare levels by the single poverty quintile, but multidimensional clustering places them at the intermediate level because other indicators, such as economic growth or unemployment rates, are relatively better than other provinces in the same poverty group. In contrast, DKI Jakarta, East Kalimantan, and Bali are rated as very prosperous by the single poverty quintile because of their low poverty percentage, but multidimensional clustering downgrades the welfare rankings of these three provinces due to high Gini ratios (an indication of high internal inequality) or even negative economic growth in the case of Bali due to the impact of the pandemic on the tourism sector. This pattern is consistent with the findings of Zuhendra et al. (2024) and Rady Putra and Anggrawan (2021) who emphasize that social assistance recipients should be grouped based on a combination of indicators, not a single indicator, in order for more accurate targeting.

From a policy perspective, these findings provide quantitative support for the concerns that have been qualitatively conveyed by the Achmadi (2024), writing for the Ombudsman of the Republic of Indonesia, regarding the accuracy of DTKS, as well as by Purbaningrum and Adinugraha (2024) regarding the inaccuracy of social assistance targets at the village level. It is tempting to extrapolate from the provincial inconsistency to the household level, on the reasoning that households vary more than provincial aggregates and that disagreement should therefore be sharper still. That extrapolation is plausible but speculative, and this study provides no evidence for it. Provincial aggregates and household records differ not only in variance but in the mechanism by which they are classified: the official Proxy Means Test weights many household variables through an estimated model, whereas the comparison made here sets a five-indicator composite against a single rate. The direction of any household-level divergence therefore cannot be inferred from these results, and the household case remains an open question that this design cannot settle. From the perspective of informatics science, these results

confirm the relevance of data science approaches, especially K-Means-based multidimensional clustering (Ahmed et al., 2020; Bahauddin et al., 2021), as a tool that can complement or even improve the Proxy Mean Test mechanism which currently still relies heavily on the weight of relatively static indicators.

This study has a number of limitations that need to be considered in the interpretation of the results. First, the unit of analysis is at the provincial level and not at the household level, so these results are illustrative to show potential structural inequality, not a direct audit of the DTKS/DTSEN database which is closed and individual/family scale. Second, the number of indicators used (five indicators) is still much smaller than the variables actually used in the official Proxy Mean Test. Third, the data used is 2021 data due to the limited availability of HDI data in the public repositories used, so it does not reflect the current conditions after the policy change towards DTSEN. Further research is recommended to use more granular household data, more complete indicators, and other clustering algorithms such as K-Medoids or Gaussian Mixture Model to test the consistency of these findings (Ahmed et al., 2020). Fourth, the five-cluster solution has weak separation, with a silhouette of 0.2802, so individual cluster assignments near a boundary should not be treated as firm. Fifth, the indicators were weighted equally despite being correlated, in particular the Human Development Index and the poverty percentage, which gives the income dimension more influence than the nominal weighting suggests. Further research is accordingly recommended to use more granular household data, a fuller indicator set, a correlation-aware alternative such as clustering on principal components, and other clustering algorithms, including K-Medoids (Park & Jun, 2009), Gaussian mixture models (Fraley & Raftery, 2002), Ward hierarchical clustering (Murtagh & Legendre, 2014), and density-based methods such as DBSCAN (Schubert et al., 2017), none of which assumes spherical clusters of comparable size and which may describe a welfare continuum more faithfully than K-Means; Andayani et al. (2025) provide a directly comparable Indonesian provincial application, in order to test the consistency of these findings.

CONCLUSION

This study applied a data science approach, namely K-Means clustering based on five socio-economic indicators, to evaluate the inequality that has the potential to arise from the policy of changing the decile of social welfare in Indonesia. The results showed that 59.38% of the 32 provinces analyzed experienced differences in classification levels between the multidimensional approach and the single poverty quintile approach, with eight provinces experiencing severe level shifts, as well as a very low Adjusted Rand Index (0.0521) indicating almost no conformity between the two classification systems. Two statements should be kept apart here. As an empirical finding, the results demonstrate substantial disagreement between a multidimensional classification and a poverty-based one at the provincial level; they do not demonstrate exclusion or inclusion error in the DTKS/DTSEN system, a claim that would require household-level evidence this study does not hold. As a policy implication, the disagreement indicates that a welfare classification used to allocate public resources is sensitive to the choice of welfare definition, which is a reason to test that sensitivity rather than a finding of misallocation. This study is a province-level proof of concept, built on public aggregates because household DTKS/DTSEN records are not publicly accessible, and its scope should be read accordingly. On that basis, K-Means-based multidimensional classification may be

considered as a complementary analytical tool for the further validation of welfare targeting rather than as a replacement for the Proxy Means Test. The concrete next step follows directly: pilot the same multidimensional clustering on household-level DTKS or DTSEN microdata for a single province, under a data-sharing agreement with the Ministry of Social Affairs, and compare the resulting household groupings against the deciles the official instrument assigns to the same households. That design would convert the disagreement demonstrated here into a measurable statement about targeting accuracy, which the present design cannot provide.

REFERENCES

- Zuhendra, M. I., Hidayat, R., & Hendrawaty. (2024). Application of data mining for poverty clustering based on Integrated Social Welfare Data (DTKS). *SKANIKA: Computer Systems and Informatics Engineering*, 7(1), 32-41. <https://doi.org/10.36080/skanika.v7i2.3193>
- Wijayanti, C. D., Soegianto, S., Sudarmanto, K., & Arifin, Z. (2025). Implementasi bantuan sosial berbasis DTKS untuk akurasi penerima di Kota Semarang. *Journal Juridisch*, 3(3), 221-235. <https://doi.org/10.26623/jj.v3i3.12886>
- Wibisono, M. A. M., & Sirait, T. (2024). Multidimensional poverty modeling in Central Java, DI Yogyakarta, and East Java provinces. *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, 18(3), 1939-1954. <https://doi.org/10.30598/barekengvol18iss3pp1939-1954>
- Schubert, E., Sander, J., Ester, M., Kriegel, H. P., & Xu, X. (2017). DBSCAN revisited, revisited: Why and how you should (still) use DBSCAN. *ACM Transactions on Database Systems*, 42(3), 1-21. <https://doi.org/10.1145/3068335>
- Sartirano, D., Kalimeri, K., Cattuto, C., Delamónica, E., Garcia-Herranz, M., Mockler, A., Paolotti, D., & Schifanella, R. (2023). Strengths and limitations of relative wealth indices derived from big data in Indonesia. *Frontiers in Big Data*, 6, Article 1054156. <https://doi.org/10.3389/fdata.2023.1054156>
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53-65. [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)
- Rady Putra, L. G., & Anggrawan, A. (2021). Grouping of community social assistance recipients using the K-Means method. *Mathematics: Journal of Management, Informatics Engineering and Computer Engineering*, 21(1), 205-214.
- Purbaningrum, D., & Adinugraha, H. H. (2024). Public policy transformation to overcome the inaccuracy of social assistance targets in Pringsurat village for better community welfare. *Journal of Development and Public Policy*, 15(2), 31-44. <https://doi.org/10.36624/jpkp.v15i2.116>
- Pratiwi, I. E., & Joshi, N. P. (2025). Exploring the regional multidimensional poverty pattern in Indonesia: Does climate matter? *Etikonomi*, 24(2), 635-652. <https://doi.org/10.15408/etk.v24i2.39523>
- Park, H.-S., & Jun, C.-H. (2009). A simple and fast algorithm for K-medoids clustering. *Expert Systems with Applications*, 36(2), 3336-3341. <https://doi.org/10.1016/j.eswa.2008.01.039>

- Niswatin, K. (2021). Clustering of districts and cities in Indonesia based on poverty indicators using the K-Means method. *Proceedings of the International Conference on Computing, Mathematics and Statistics*.
- Nasution, I., Windarto, A. P., & Fauzan, M. (2020). Application of K-Means algorithm in the grouping of poor population data by province. *Building of Informatics, Technology and Science*, 2(2), 76-83.
- Murtagh, F., & Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: Which algorithms implement Ward's criterion? *Journal of Classification*, 31(3), 274-295. <https://doi.org/10.1007/s00357-014-9161-z>
- Istiqamah, N., Soesanto, O., & Anggraini, D. (2021). Application of the K-Means algorithm to determine poverty status in Hulu Sungai Tengah. *Journal of Physics: Conference Series*, 2106(1), 012027. <https://doi.org/10.1088/1742-6596/2106/1/012027>
- Hennig, C. (2007). Cluster-wise assessment of cluster stability. *Computational Statistics & Data Analysis*, 52(1), 258-271. <https://doi.org/10.1016/j.csda.2006.11.025>
- Haryaningsih, S., Andriani, F. D., Pambudi, A., Seran, S. T., & Sihalo, N. T. P. (2024). Crisis-responsive social protection: Counterfactual evidence on poverty reduction from Indonesia's response to COVID-19. *Natapraja*, 12(1), 58-74. <https://doi.org/10.21831/napapraja.v12i1.88630>
- Hanna, R., & Olken, B. A. (2018). Universal basic incomes versus targeted transfers: Anti-poverty programs in developing countries. *Journal of Economic Perspectives*, 32(4), 201-226. <https://doi.org/10.1257/jep.32.4.201>
- Gazeaud, J. (2020). Proxy means testing vulnerability to measurement errors? *The Journal of Development Studies*, 56(11), 2113-2133. <https://doi.org/10.1080/00220388.2020.1715942>
- Fraley, C., & Raftery, A. E. (2002). Model-based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical Association*, 97(458), 611-631. <https://doi.org/10.1198/016214502760047131>
- Dinas Sosial Provinsi Jawa Tengah. (n.d.). *Cari BDT: Daya tampung panti*. Retrieved October 17, 2026, from <https://caribdt.dinsos.jatengprov.go.id/>
- Central Statistics Agency. (2023). *Human Development Index, percentage of poor population, Gini ratio, open unemployment rate, and provincial GDP growth rate 2010-2022 [Dataset]*. <https://bps.go.id>
- Brown, C., Ravallion, M., & van de Walle, D. (2018). A poor means test? Econometric targeting in Africa. *Journal of Development Economics*, 134, 109-124. <https://doi.org/10.1016/j.jdeveco.2018.05.004>
- Bahauddin, A., Fatmawati, A., & Permata Sari, F. (2021). Analysis of provincial clustering in Indonesia based on poverty level using the K-Means algorithm. *Journal of Informatics and Information Systems Management*, 4(1), 1-8. <https://doi.org/10.36595/misi.v4i1.216>
- Arbelaitz, O., Gurrutxaga, I., Muguerza, J., Pérez, J. M., & Perona, I. (2013). An extensive comparative study of cluster validity indices. *Pattern Recognition*, 46(1), 243-256. <https://doi.org/10.1016/j.patcog.2012.07.021>
- Andayani, S., Retnani, N., Yusri, T. A. S., & Marwoto, B. S. H. (2025). A comparative analysis of DBSCAN and Gaussian Mixture Model for clustering Indonesian provinces based

- on socioeconomic welfare indicators. *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, 19(3), 2039-2056. <https://doi.org/10.30598/barekengvol19iss3pp2039-2056>
- Alkire, S., & Foster, J. (2011). Counting and multidimensional poverty measurement. *Journal of Public Economics*, 95(7), 476-487. <https://doi.org/10.1016/j.jpubeco.2010.11.006>
- Ahmed, M., Seraj, R., & Islam, S. M. S. (2020). The k-means algorithm: A comprehensive survey and performance evaluation. *Electronics*, 9(8), 1295. <https://doi.org/10.3390/electronics9081295>
- Achmadi, M. (2024, June 27). *Problematika bantuan sosial dan DTKS*. Ombudsman Republik Indonesia. <https://ombudsman.go.id/artikel/r/pwkinternal--problematika-bantuan-sosial-dan-dtks>