UJM 14 (1) 2025
UNNES Journal of Mathematics
http://journal.unnes.ac.id/sju/index.php/ujm

Regional Prioritization for Free Nutritious Food Programs through Social Data Integration and Public Sentiment Analysis Using K-Means and NLP

Ratna Nur Mustika Sanusi*, Galih Kusuma Wijaya, Nur Achmey Selgi Harwanti

Department of Mathematics, Faculty of Mathematics and Natural Science,
Universitas Negeri Semarang, Semarang, 50299, Indonesia

Article Info

Article History:
Recived ....
Accepted .....
Published .....

Keywords:
Clustering, sentiment analysis public health program evaluation, regional disparities

Abstract

This study evaluates Indonesia's Free Nutritious Meal Program (MBG) using two innovative approaches: geospatial clustering and sentiment analysis. It combines quantitative data from five social indicators from 38 provinces, including literacy rates, stunting prevalence, and HDI-from the Badan Pusat Statistik, with qualitative data from Twitter posts related to the MBG program. Cluster analysis using the K-means method identified three priority zones: high (Eastern Indonesia), medium (Central Indonesia), and low (Jawa-Bali-West Sumatra), revealing stark geographic disparities. Sentiment analysis using Natural Language Processing (NLP) on 1,358 social media posts found that 76.6% of public sentiment was negative, driven by concerns over food safety (e.g., "poisoning," "toxic"). In contrast, 23.4% expressed positive views emphasizing nutritional benefits. The study makes three key contributions: First, it demonstrates the disconnect between regional needs and implementation quality. Second, it introduces an integrated monitoring framework combining cluster mapping with real-time sentiment tracking. Third, it proposes actionable solutions including a rapid-response task force and targeted communication strategies. These findings provide policymakers with evidence-based tools to simultaneously address geographical inequities and improve program execution in nutrition interventions.
.

How to cite:

Sanusi, R. N. M., Wijaya, G. K., & Harwanti, N. A. S. 2025. Regional Prioritization for Free Nutritious Food Programs through Social Data Integration and Public Sentiment Analysis Using K-Means and NLP. *UNNES Journal Of Mathematics*. 14(1): 43-52.

*Correspondence Address
E-mail: rnmustika@mail.unnes.ac.id

p-ISSN 2252-6943
e-ISSN 2460-5859

1. Introduction

The Free Nutritious Food Program (Program Makanan Bergizi Gratis, MBG) is a strategic policy within Indonesia's long-term development framework toward achieving "Golden Indonesia 2045," aligning with the government's vision outlined in Law No. 59 of 2024 on the National Long-Term Development Plan for 2025–2045. This initiative is designed to realize one of the core pillars of *Asta Cita* (Eight Priorities): strengthening human resources through integrated nutritional interventions (President of Indonesia, 2024). The MBG focuses on addressing stunting, malnutrition, and supporting child development, maternal health during pregnancy and lactation, as well as improving access to quality education (Badan Gizi Nasional, 2025).

The 2023 Indonesian Health Survey revealed a stunting prevalence of 21.5%, which has declined over the past decade. Approximately 15 of Indonesia's 38 provinces exhibit stunting rates below the national average. The provinces with the highest stunting rates are Central Papua (39.4%), East Nusa Tenggara (37.9%), and Papua Highlands (37.3%) (Ministry of Health, 2024). Meanwhile, data from the Central Statistics Agency (BPS, 2024) highlights significant regional disparities in food poverty lines, ranging from IDR 352,533 per capita per month in Java to over IDR 800,000 in eastern regions such as Papua Highlands. These findings underscore the necessity for regionally tailored interventions to ensure the MBG program is implemented effectively and equitably.

Nevertheless, during the Free Nutritious Meal Program (MBG) implementation, public responses recorded through social media reflected sharp dynamics of opinion—ranging from support to criticism—regarding the program's effectiveness, distribution accuracy, and potential risks. Amid this controversy, it is crucial to ensure that public perception does not become the sole reference for policy decisions, but is balanced with objective urgency based on macro-level social data.

Previous research by (Kiftiyah *et al*., 2025) examined the Free Nutritious Meal Program (MBG) through the lens of social justice and socio-political dynamics using a Systematic Literature Review (SLR) approach. The study emphasized the importance of prioritizing underdeveloped, frontier, and outermost (3T) regions and identifying factors influencing the program's sustainability. However, it did not incorporate macro-social quantitative data in determining priority areas. The novelty of the present study lies in its dual-method approach, which integrates cluster analysis based on five macro-social indicators from the Badan Pusat Statistik namely, illiteracy rate, expected years of schooling, stunting prevalence, access to improved sanitation, and the Human Development Index (HDI) with sentiment analysis using Natural Language Processing (NLP) on social media data. Cluster analysis was employed to group Indonesia's 38 provinces into intervention priority categories based on varying degrees of social vulnerability. These clustering results were then compared with public sentiment derived from social media to evaluate whether public perception aligns with actual needs indicated by the data. This comparison is essential in balancing subjective perception and objective necessity, ultimately supporting the design of more equitable and effective public policies.

Cluster analysis was conducted to categorize 38 provinces in Indonesia into intervention priority levels based on social vulnerability. These clusters were then compared to public sentiment derived from social media to assess whether public perception aligns with actual needs based on empirical data. This comparison balances subjective perception with objective necessity in designing effective and equitable public policy.

Therefore, this study addresses two main research questions: 1. How are MBG intervention priorities spatially mapped when determined objectively through macro-social indicators? Moreover, 2. To what extent does public sentiment on social media reflect the actual needs of each region? By comparing data-driven spatial mapping with public sentiment analysis, this research aims to provide policy recommendations that are socially just and responsive to public aspirations.

2. Method

This study employs two primary methods to analyze social conditions and public responses in Indonesia. The analysis aims to examine interprovincial development disparities and public perceptions of specific policies, such as the MBG Program (e.g., Government Assistance). This multidimensional approach integrates clustering techniques to categorize regions based on socioeconomic indicators with sentiment analysis to evaluate public opinion derived from social media. The findings are expected to provide data-driven recommendations for formulating more targeted policies.

2.1. K-Means

Cluster analysis was employed to group Indonesian provinces based on their social conditions. Clustering serves as a valuable tool in data science and is a method for identifying inherent cluster structures within datasets, where each cluster is characterized by the highest similarity among its members and the greatest dissimilarity with members of other clusters (Sinaga & Yang, 2020). One of the simplest and most widely used clustering algorithms is *K-means*, which has been extensively adopted as a standard clustering method in machine learning research (Handayanna & Sunarti, 2024). The *K*-means clustering method was selected due to its ability to form clusters based on similarity patterns, facilitating the identification of regions with varying levels of social development priority (Saputra *et al*., 2024). This approach ensures that the analysis results can serve as a foundation for more targeted and efficient policy formulation.

The objective of the *K*-means algorithm is to determine a set of cluster centroids while minimizing the sum of squared distances between each sample and its nearest cluster center. In a recent study, (Nie *et al*., 2022) proposed a modification to the *K*-means algorithm by reformulating the classical *K*-means objective function as a trace maximization problem, subsequently replacing it with a novel formulation. The proposed algorithm eliminates the need for iterative centroid recalculations and reduces intermediate variable computations during optimization, thereby improving computational efficiency.

In this study, clustering was performed based on five social indicators, including: illiteracy rate, expected years of schooling, stunting prevalence, access to improved sanitation, and Human Development Index (HDI) indicators were selected to comprehensively capture socioeconomic disparities across regions. Prior to clustering, data normalization was performed to standardize measurement scales, ensuring each variable contributed equally to the cluster formation process (Wongoutong, 2024).

The *K*-means algorithm aims to minimize the total distance between data points and their respective cluster centroids. The fundamental formula used in this method is:

$$J = \sum_{i=1}^{k} \sum_{x \in C_i} ||x - \mu_i||^2 \quad (1)$$

Notations:

k	= number of clusters
C_i	= i-th cluster
x	= data point
μ_i	= centroid of the i-th cluster
$\|\|x - \mu_i\|\|$	= Euclidean distance between a data point and its centroid

The cluster analysis identified three distinct provincial groupings: high-priority regions (red) exhibiting the most critical social conditions, medium-priority regions (blue) demonstrating intermediate social development levels, and low-priority regions (green) showing relatively favorable social indicators. To enhance interpretability, these clustering results were spatially visualized through a cluster map that clearly delineates the geographical distribution of provinces across these priority categories. This spatial representation serves as a valuable policy tool by explicitly highlighting regional disparities in social development, thereby enabling more targeted and effective intervention strategies.

2.2. Natural Languange Processing (NLP)

To complement the quantitative analysis, sentiment analysis was conducted using Natural Language Processing (NLP) techniques to examine public perceptions regarding the MBG program. This approach involved collecting and analyzing text data (tweets) gathered through relevant keywords, with the objective of identifying dominant public sentiments toward the policy.

The study employed a **lexicon-based sentiment analysis** method, wherein each word in the text was compared against a pre-classified Indonesian sentiment dictionary categorized into positive and negative terms. Each word in a tweet was assigned a sentiment score (+1 for positive, -1 for negative), and the aggregate score determined the overall sentiment polarity. The fundamental sentiment calculation formula is as follows:

$$\boldsymbol{Sentiment_{tweet}} = \sum_{i=1}^{n} \boldsymbol{Score}(\boldsymbol{word_i}) \quad (2)$$

Notations:

n	= total number of words in the tweet
Score($word_i$)	= sentiment score of the i-th word
Total score > 0	= Positive tweet sentiment
Total score < 0	= Negative tweet sentiment

This method provides a systematic framework for quantifying public opinion, enabling comparison between data-driven policy priorities and societal perceptions as expressed on social media platforms. The findings are presented in two primary formats:

Hasil analisis disajikan dalam dua bentuk utama:

1. Percentage distribution of positive versus negative tweets, revealing the dominant sentiment polarity among public responses; and
2. Word cloud visualizations displaying the most frequently occurring terms in both positive and negative tweets, providing intuitive graphical representations of prevailing public concerns and thematic patterns.

This dual-mode presentation enables comprehensive interpretation of sentiment analysis outcomes, combining quantitative metrics with qualitative visual analytics to capture nuanced dimensions of public opinion regarding the MBG policy.

3. Results and discussions

3.1. Descriptive Statistics

This analysis examines data from 38 Indonesian provinces across five key indicators: Illiteracy Rate, Expected Years of Schooling (HLS), Stunting Prevalence, Sanitation Access, and Human Development Index (HDI). Descriptive statistics were employed to understand the distribution, variation, and underlying patterns in the data prior to cluster modeling. The results will facilitate the identification of provinces with similar characteristics to support more targeted nutritional intervention programs. Below is a summary of the descriptive statistics for these five variables:

Table 1. Descriptive Statistics Data

	N	Min	Q1	Median	Q3	Max	Mean
Illiteracy Rate	38	0.04	0.09	0.24	0.50	22.74	1.35
HLS	38	9.63	12.86	13.25	13.72	15.70	13.21
Stunting Prevelence	38	7.20	19.23	23.65	27.35	39.40	23.54
Sanitation Access	38	12.61	19.38	83.36	87.18	98.83	81.14
HDI	38	53.42	71.08	72.18	74.34	83.08	72.39

Data from 38 Indonesian provinces reveal significant variations in key development indicators, particularly those related to nutrition and welfare. The illiteracy rate ranges from exceptionally low (0.04%) to alarmingly high (22.74%), with a national average of 1.35%. Central Papua exhibits the highest illiteracy rate (13.01%), while the Special Region of Yogyakarta (DIY) and Riau report the lowest rates. These disparities underscore substantial gaps in basic education access, which may influence public understanding of nutrition and health.

Expected Years of Schooling (EYS) remains relatively stable across most provinces (median: 13.25 years), with DIY being a notable exception (15.7 years), reflecting its superior education quality. However, stunting remains a critical public health challenge, with the highest prevalence observed in Central Papua (39.4%) and West Sulawesi (30.3%). These elevated stunting rates correlate with poor sanitation coverage—for instance, Central Papua's sanitation access stands at only 41.44%, significantly below the national average (81.14%).

The Human Development Index (HDI) also varies markedly, ranging from 53.42 (lowest) to 83.08 (DIY, highest). Provinces with low HDI scores typically exhibit compounded challenges: high illiteracy, inadequate sanitation, and severe stunting. In contrast, high-HDI provinces such as Riau (74.79) and DIY (83.08) demonstrate better performance across all indicators.

Descriptive statistical analysis identifies potential provincial clusters based on shared characteristics. A "high-priority" cluster emerges, characterized by high illiteracy and stunting rates coupled with low sanitation coverage. Conversely, a "moderately developed" cluster displays more favorable indicators. Stunting and sanitation variables emerge as key determinants in this clustering model, given their strong correlation. To further elucidate data patterns, subsequent steps include data normalization to standardize measurement scales.

3.2. *Clustering Results*

3.2.1. *Elbow Mehod*

The Elbow Method was employed to identify the optimal number of clusters for the K-Means modeling approach (Anggreani *et al*., 2024). This technique evaluates the within-cluster sum of squares (WCSS) against different cluster numbers, with the "elbow point" - where the rate of WCSS reduction markedly decreases - indicating the most appropriate cluster count. The results of the elbow analysis revealed:

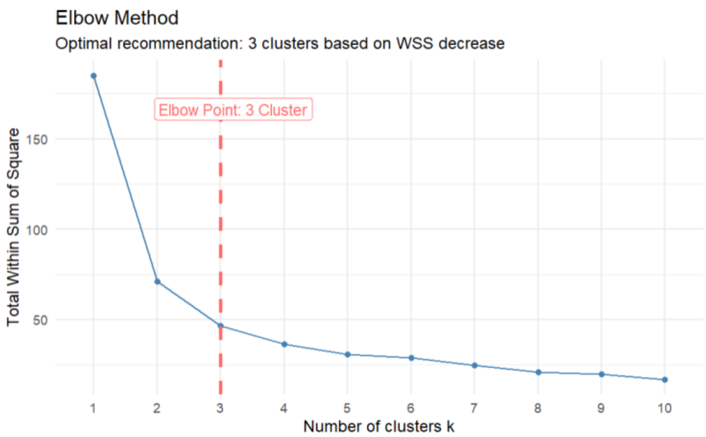

Figure 1. Elbow Method

As evidenced by the graph above, the total Within-Cluster Sum of Squares (WSS) demonstrates a significant reduction up to 3 clusters, beyond which the rate of decrease substantially diminishes. This inflection point (elbow point) indicates that 3 clusters represent the optimal choice, effectively balancing cluster precision with model parsimony. Consequently, the analysis recommends employing 3 clusters for provincial-level data segmentation in Indonesia, which will facilitate clearer identification of spatial patterns in nutrition and development indicators.

3.2.2. Cluster Profiles

The K-Means modeling results with 3 clusters reveal distinct provincial groupings based on socio-economic and nutritional indicators. Table 2 presents the mean profile of each cluster derived from the analysis:

Table 2. Cluster Profiles

Cluster	Illiteracy Rate (%)	HLS	Stunting	Sanitation	IPM
1	0.70	13.45	27.25	80.57	71.13
2	17.88	9.80	38.35	27.03	56.59
3	0.16	13.35	18.18	87.73	75.40

Cluster Profile Analysis

1. Cluster 1
    - Relatively low illiteracy rate (0.70%) paired with moderately high Expected Years of Schooling (EYS: 13.45 years)
    - Intermediate stunting prevalence (27.25%), supported by adequate sanitation coverage (80.57%)
    - Moderate Human Development Index (HDI: 71.13), indicating satisfactory human development progress
2. Cluster 2
    - Exceptionally high illiteracy rate (17.88%) and the lowest EYS (9.80 years)
    - Highest stunting prevalence (38.35%) coupled with critically low sanitation access (27.03%)
    - Lowest HDI (56.59), reflecting compounded developmental challenges
3. Cluster 3
    - Optimal performance with minimal illiteracy (0.16%) and high EYS (13.35 years)
    - Lowest stunting prevalence (18.18%) and superior sanitation coverage (87.73%)
    - Highest HDI (75.40), demonstrating advanced human development outcomes

Priority Intervention Categorization

Based on these profiles, clusters were classified by intervention urgency:

- Highest Priority: Cluster 2, due to critical deficits across all indicators
- Medium Priority: Cluster 1, requiring targeted improvements in specific domains
- Lowest Priority: Cluster 3, maintaining strong performance across metrics

The following table presents the classification of 38 Indonesian provinces into three intervention priority categories for the Free Nutritious Meal Program (MBG), based on cluster analysis using macro-social indicators. Each province is grouped into Cluster 1 (medium priority), Cluster 2 (high priority), or Cluster 3 (low priority).

Table 3. Clustering Results of 38 Indonesian Provinces

No.	Province	Cluster	No.	Province	Cluster
1	Sulawesi Tengah	1	20	Papua Pegunungan	2
2	Sulawesi Barat	1	21	Riau	3
3	Sulawesi Selatan	1	22	Daerah Istimewa Yogyakarta	3
4	Papua Barat	1	23	Dki Jakarta	3
5	Gorontalo	1	24	Bengkulu	3
6	Papua Selatan	1	25	Lampung	3
7	Sumatera Barat	1	26	Kepulauan Riau	3
8	Maluku	1	27	Jambi	3
9	Papua	1	28	Bali	3
10	Nusa Tenggara Barat	1	29	Jawa Timur	3
11	Papua Barat Daya	1	30	Sumatera Utara	3
12	Sulawesi Tenggara	1	31	Kepulauan Bangka Belitung	3
13	Nusa Tenggara Timur	1	32	Sumatera Selatan	3
14	Kalimantan Selatan	1	33	Banten	3
15	Aceh	1	34	Sulawesi Utara	3
16	Kalimantan Tengah	1	35	Kalimantan Utara	3
17	Maluku Utara	1	36	Kalimantan Timur	3
18	Kalimantan Barat	1	37	Jawa Tengah	3
19	Papua Tengah	2	38	Jawa Barat	3

To provide a more comprehensive visual representation of the clustering results, the following map of Indonesia is color-coded based on the intervention priority categories of the Free Nutritious Meal Program (MBG).

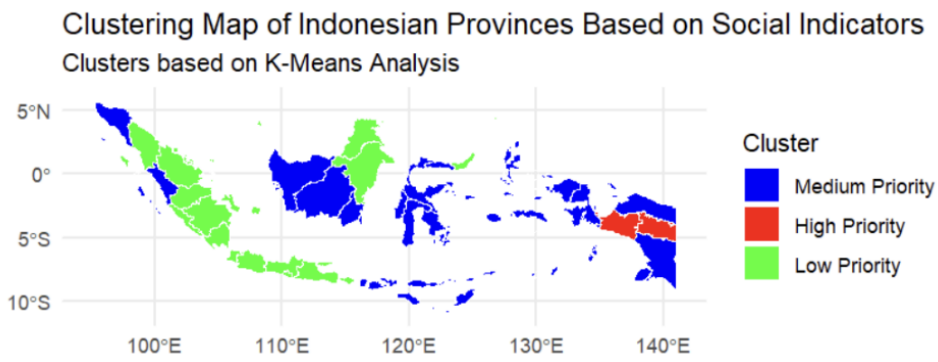

Figure 2. Clustering Map

The distribution of provinces across each cluster reveals a stark contrast in geographic and development patterns. Cluster 2, which represents the highest priority for intervention, consists entirely of provinces in Eastern Indonesia—specifically, Central Papua and Highland Papua. This highlights the persistent developmental lag in these regions, particularly in education, health, and basic infrastructure. These findings align with previous research (Efendi, 2025), which emphasizes the significant development disparity between Indonesia's eastern and western regions.

Cluster 1 includes a broad range of Eastern and Central Indonesia provinces, notably most areas in Papua, Sulawesi, and Kalimantan. These provinces face moderate challenges: indicators such as the Human Development Index (HDI) and sanitation access are relatively adequate, yet stunting prevalence remains high. This suggests that intervention efforts should focus on nutrition and basic health services.

In contrast, Cluster 3 comprises western and urbanized provinces, such as Jakarta (DKI Jakarta), West Java, Central Java, and North Sumatra. These areas generally show strong performance on key indicators. Interestingly, provinces like Riau and Bali are also in this cluster despite common public perceptions of persistent development issues. This could imply that indicators relevant to the MBG program—such as illiteracy rates and sanitation coverage—are already being addressed. However, intra-provincial disparities may still exist and warrant further evaluation.

This cluster distribution underscores the importance of regionally tailored and indicator-specific policies, rather than relying solely on administrative classifications. Each cluster reflects a distinct developmental typology, indicating that a one-size-fits-all approach would be ineffective in addressing Indonesia's development disparities.

3.3. Natural Languange Processing (NLP)

This study examines public sentiment toward the Free Nutritious Food Program (Program Makanan Bergizi Gratis, MBG) using data extracted from platform X. A total of 1,358 tweets were collected and systematically processed through data cleaning to visualization stages. The analysis reveals two dominant sentiment polarities regarding the program.

Table 3. Representative Positive Tweets

No.	Positive Tweet	Key Positive Terms
1	"Pemenuhan HAM melalui makanan bergizi gratis dukung hak anak untuk tumbuh sehat cerdas #makananbergizi #makananbergizigratis"	"sehat", "dukung"
2	"Bandung nggak main-main soal kesehatan! Tim BPOM kerja keras demi program MBG sukses."	"kesehatan", "sukses"
3	"Makan bergizi bikin anak-anak Indonesia lebih fokus di sekolah. Yuk dukung makanan bergizi gratis buat mereka! #makananbergizigratis"	"dukung", "bergizi"

As shown in Table 3, supportive tweets emphasized the program's benefits for child health, educational outcomes, and employment generation. Key positive terms included "healthy," "support," and "success," reflecting public appreciation for the program's nutritional objectives. Representative tweets highlighted:

1. The program's role in fulfilling children's rights to healthy development
2. Local government efforts to ensure program success
3. Improved student concentration through better nutrition

Table 4. Representative Negative Tweets

No.	Negative Tweet	Key Negative Terms
1	Kasus pelajar keracunan usai memakan makanan program MBG kerap terjadi belakangan ini. Badan Gizi Nasional (BGN) memperketat prosedur.	"meracuni", "keracunan", "lalainya"
2	*MBG tuh makanan bergizinya bukan buat yang dapat proyek bancakannya dong! Soalnya anak sekolahnya kok dapatnya makanan beracun ya? Bukan makanan bergizi. Ga mungkin makanan beracun gratis kann?*	"beracun"
3	*[Barito dari CNN Indonesia] Ratusan pelajar SMP di Bandung keracunan setelah makan makanan bergizi gratis (MBG). Program MBG tu pamarintah punyo ngasih makanan gratis bergizi ka pelajar, ibu hamil, jo ibu menyusui. Tapi di SMP Negeri 35 Bandung, 342 pelajar jo 2 guru alami diare, sakit perut, muntah, pusing, jo demam pasca makan MBG pada 29 April 2025. Semuanya dibawa ka puskesmas jo rumah sakit, tapi Alhamdulillah tak ado yang dirawat inap. Penyebabnyo masih diselidiki, tapi dugaan ado masalah jo kualitas makanan.*	"keracunan", "diare", "sakit perut", "muntah"

Table 4 documents critical tweets focusing on food poisoning incidents and quality control failures. Predominant negative terms included "toxic," "poisoning," and "diarrhea," indicating public concerns about implementation. Major criticisms involved:

1. Recurrent student poisoning cases
2. Allegations of program mismanagement
3. Inadequate food quality supervision

This analysis provides empirical evidence for program evaluation, highlighting both public support for MBG's objectives and critical concerns regarding its implementation. The findings suggest that while the program's conceptual framework receives approval, operational execution requires significant improvement to maintain public trust and achieve nutritional goals.

3.3.1. Sentiment Analysis

Based on the data visualization, the sentiment distribution indicates a significant imbalance.

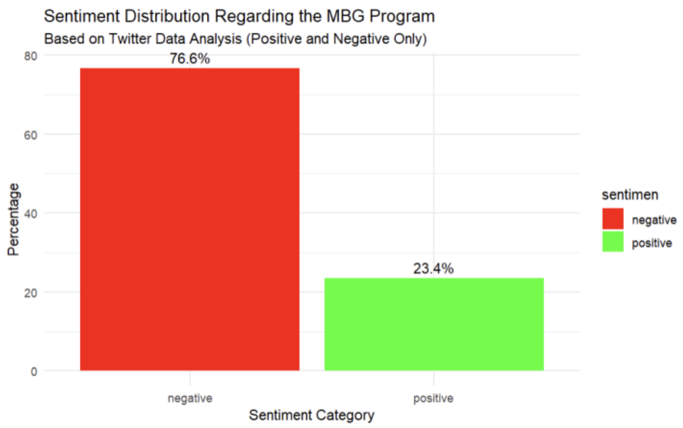

Figure 3. Persentase Sentimen

The sentiment distribution analysis reveals a dominant negative sentiment at 76.6%, which is qualitatively supported by the occurrence of negative keywords in the textual analysis. The term *"keracunan"* ("food poisoning") appears most frequently (128 times), followed by *"beracun"* ("toxic") (36 times) and *"salah"* ("wrong") (23 times), highlighting serious concerns related to food safety within the program. Other negative terms such as *"masalah"* ("problem"), *"korupsi"* ("corruption"), and *"kritik"* ("criticism") (22, 15, and 13 occurrences respectively) indicate fundamental issues in program governance Interestingly, the frequent mention of *"evaluasi"* ("evaluation") (13 times) in public tweets suggests an emerging public awareness of the need for systemic improvement. The prevalence of these negative terms aligns with the earlier quantitative findings of sentiment dominance, while also providing deeper insights into the root causes. Words like *"sakit"* ("illness") (9 times) and *"buruk"* ("bad") (8 times) further emphasize the program's negative impact on beneficiaries. These findings confirm that public dissatisfaction is not merely general but grounded in specific concerns related to food safety and transparency. The prominence of the terms *"keracunan"* and *"beracun"* underscores an urgent need for program evaluation, especially as they pertain to basic public health. The distribution pattern of negative keywords, which clusters around health and governance issues, should be prioritized in drafting program improvement recommendations.

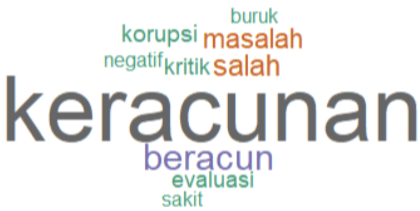

Figure 4. Wordcloud Sentimen Negatif

The negative sentiment word cloud, featuring keywords such as *"corruption"*, *"problem"*, *"poisoning"*, and *"toxic"*, visually reinforces the previously identified negative sentiment patterns. The prominence of *"keracunan"* in the visual confirms earlier frequency-based analyses, validating that food safety remains the primary concern among the public.

Although negative sentiment is predominant, a deeper examination of positive sentiment reveals certain appreciated aspects of the program. The word *"sehat"* ("healthy") appears 70 times, followed by *"baik"* ("good") 53 times and *"dukung"* ("support") 51 times, indicating public appreciation for the program's core goal of providing nutritious food. The terms *"lanjutkan"* ("continue") (40 times) and *"kesehatan"* ("health") (34 times) reflect the hope that the program continues with necessary improvements.

Keywords such as *"keamanan"* ("safety") (28 times), *"manfaat"* ("benefit") (22 times), and *"positif"* ("positive") (20 times) suggest that some beneficiaries experience tangible benefits from the program. Additionally, *"semangat"* ("spirit") (20 times) and *"cerdas"* ("smart") (19 times) indicate that psychological and educational impacts have been positively received. These findings confirm that despite implementation challenges, the program's fundamental objectives to improve nutrition and health remain relevant and acknowledged.

This balanced perspective illustrates that the MBG program has a sound conceptual foundation and enjoys a degree of public support, although significant improvements are needed in its implementation. The broader variety of positive keywords compared to negative ones (10 vs. 8 key terms) suggests that support for the program stems from multiple dimensions. This forms a valuable asset for driving improvements while preserving the program's successful aspects.

Figure 5. Wordcloud Sentimen Positif

The positive sentiment word cloud further supports previous findings by illustrating an engaging distribution of key terms. The word *"sehat"* appears as the most visually dominant term, aligning with frequency data that identifies it as the most frequently used positive word (70 times). The large font size of *"baik"* and *"dukung"* is consistent with their high frequencies (53 and 51 times respectively) in the quantitative dataset.

The findings of this study are consistent with those of (Rahamatullah *et al*., 2025), which revealed that negative perceptions toward the Free Nutritious Meal Program (MBG) are highly prevalent in digital spaces. An analysis of comments on the State Secretariat's official YouTube channel found that **1,403 out of 1,470 comments (approximately 95.4%)** expressed negative sentiments toward the program. These results underscore the critical importance of managing public perception, particularly in the disconnect between objective needs and social media-driven public opinion.

3.4. Integration of Findings and Implementation Issues

The cluster analysis results show that regions with the highest levels of social vulnerability such as Maluku, East Nusa Tenggara, and Papua are consistently classified within the medium- to high-priority intervention categories. Interestingly, sentiment analysis reveals that negative public perception originates from low-priority regions, such as West Java, highlighting a misalignment between data-driven needs and public sentiment on social media.

Moreover, data from the Indonesian Food and Drug Authority (BPOM) as of May 20, 2025, recorded 17 food poisoning cases and eight non-poisoning incidents directly linked to MBG implementation across 10 provinces. These findings are significant, as they demonstrate that public concerns over food safety are not merely perceptual, but are supported by factual events. However, not all reported cases occurred in medium- or high-priority regions many were in low-priority areas, where implementation quality, rather than need, becomes the central issue.

This reinforces the argument that, in addition to targeting regions based on objective need, implementation quality and food distribution oversight systems must become central concerns in future MBG policies.

Integrating research findings with real-world implementation issues highlights the need for holistic monitoring of nutrition intervention programs not only in distribution logistics but also from the perspective of public reception and on-the-ground effectiveness. As such, the proposed integration of spatial mapping and online sentiment monitoring offers a novel framework for identifying policy gaps and enhancing the government's responsiveness to evolving social dynamics.

4. Conclusion

The clustering and sentiment analyses of the MBG Program offer two equally important yet independently derived perspectives. Clustering analysis based on geographic indicators identified three priority zones: red zones (high priority) in Eastern Indonesia requiring urgent intervention, yellow zones (medium priority) in Central Indonesia needing focused improvement, and green zones (low priority) in Java, Bali, and West Sumatra, which can serve as models of successful implementation. Meanwhile, sentiment analysis revealed that 76.6% of public responses were negative, primarily focusing on poisoning incidents and distribution issues, while 23.4% were positive, highlighting health benefits.

These two analyses complement each other without implying direct causality—clustering aids in objective regional prioritization, while sentiment analysis reflects public perception of program implementation, with no direct correlation between cluster categories and sentiment ratios.

The independent findings from sentiment analysis underscore the urgency of addressing food poisoning incidents and uneven distribution, which dominate public complaints. In response, we recommend establishing a Complaint Response Task Force capable of responding to reports within 24 hours and implementing a real-time data-driven distribution system. Meanwhile, positive sentiments appreciating the nutritional benefits of the program should be leveraged through testimonial-based educational campaigns.

To ensure effective implementation, an integrated monitoring system that combines clustering maps with sentiment analysis dashboards is essential. Communication strategies should be adapted to regional characteristics, utilizing traditional media in underdeveloped areas and social media for rapid complaint resolution. Quarterly evaluations based on up-to-date data are key to maintaining program responsiveness to real needs on the ground while ensuring public accountability.

References

Anggreani, D., Nurmisba, N., Setiawan, D., & Lukman, L. (2024). Optimization of K-Means Clustering Method by Using Elbow Method in Predicting Blood Requirement of Pelamonia Hospital Makassar. *Internet of Things and Artificial Intelligence Journal*, *4*, 541–550. https://doi.org/10.31763/iota.v4i3.755

Badan Gizi Nasional. (2025, Mei 15). *Pentingnya program MBG: Tingkatkan SDM dan pencegahan stunting*. https://www.bgn.go.id/news/artikel/pentingnya-program-mbg-tingkatkan-sdm-dan-pencegahan-stunting_tanggal_15_Mei_2025

Efendi, L. O. (2025). Analisis Pertumbuhan Ekonomi Regional dan Ketimpangan Antarwilayah. *Circle Archive*, *1*(7).

Handayanna, F., & Sunarti, S. (2024). Penerapan Algoritma K-Means Untuk Mengelompokkan Kepadatan Penduduk Di Provinsi DKI Jakarta. *Journal of Applied Computer Science and Technology*, *5*, 50–55. https://doi.org/10.52158/jacost.v5i1.477

Indonesia. (2024). *Undang-Undang Nomor 59 Tahun 2024 tentang Rencana Pembangunan Jangka Panjang Nasional Tahun 2025–2045*. Sekretariat Negara.

Kementerian Kesehatan Republik Indonesia. (2024). *Stunting di Indonesia dan determinannya*. Badan Kebijakan Pembangunan Kesehatan.

Kiftiyah, A., Palestina, F. A., Abshar, F. U., & Rofiah, K. (2025). Program Makan Bergizi Gratis (MBG) dalam Perspektif Keadilan Sosial dan Dinamika Sosial–Politik. *Pancasila: Jurnal Keindonesiaan*, *5*(1), 101-112.

Nie, F., Li, Z., Wang, R., & Li, X. (2022). An effective and efficient algorithm for K-means clustering with new formulation. *IEEE Transactions on Knowledge and Data Engineering,* *35*(4), 3433–3443. https://doi.org/10.1109/TKDE.2022.3155450

Rahmatullah, B., Saputra, S. A., Budiono, P., & Wigandi, D. P. (2025). Sentimen Analisis Makan Bergizi Gratis Menggunakan Algoritma Naive Bayes. *Journal of Information Technology*, *5*(1).

Saputra, A. K., Erlangga, E., & Thamrin, T. (2024). Implementasi K-Means pada Klasterisasi Data Intervensi Prioritas Penerima Bantuan Sosial. *EXPERT: Jurnal Manajemen Sistem Informasi Dan Teknologi*, *14*, 23. https://doi.org/10.36448/expert.v14i1.3717

Sinaga, K. P., & Yang, M. S. (2020). Unsupervised K-means clustering algorithm. *IEEE Access, 8*, 80716–80727. https://doi.org/ 10.1109/ACCESS.2020.2988796

Wongoutong, C. (2024). The impact of neglecting feature scaling in k-means clustering. *PLoS ONE*, *19*(12). https://doi.org/10.1371/journal.pone.0310839