



---

## **K-Means Algorithm for District/City Classification in West Java Based on Stunting Toddler Data**

**Ade Bani Riyan, Ikhsan Nendi**

Politeknik Siber Cerdika Internasional, Indonesia

Email: adebani@polteksci.ac.id, nendi@polteksci.ac.id

---

### **Abstract**

Health issues, especially related to stunting, remain a major concern in Indonesia. Stunting is a condition where toddlers have a height that is too short for their age, which is a sign of chronic malnutrition. The purpose of this study is to group the number of stunting by district/city in West Java. The dataset was obtained from data on the number of stunting toddlers by district/city in West Java Province from 2014 to 2022. This dataset related to the topic of Health is produced by the health Office which is issued every 1 year. The results showed that there are two main groups in the classification of stunting numbers in West Java. The first group (cluster 0) includes Ciamis Regency, Kuningan Regency, Majalengka Regency, Sumedang Regency, Subang Regency, Purwakarta Regency, Bekasi Regency, West Bandung Regency, Pangandaran Regency, Bogor City, Sukabumi City, Bandung City, Cirebon City, Depok City, Cimahi City, Tasikmalaya City, Banjar City. The second group (cluster 1) consists of Bogor Regency, Sukabumi Regency, Cianjur Regency, Bandung Regency, Garut Regency, Tasikmalaya Regency, Cirebon Regency, Indramayu Regency, Karawang Regency, Bekasi City

**keywords: Algoritma K-Means, Data Mining, Stunting**

---

### **INTRODUCTION**

The building sector has contributed one-third of global energy consumption of which 38% comes from Heating Ventilation and Air Conditioning Systems (HVAC)(González-Torres et al., 2022). Building energy conservation is required to combat climate change in accordance with the Paris Agreement 2030. Local climate conditions and solar radiation exposure affect the thermal comfort of the people living in the building (Dong et al., 2020; Marino et al., 2017). In regions with four seasons, buildings need energy for cooling and heating. Meanwhile, in tropical areas around the equator, buildings require cooling energy throughout the year. It is important to consider local climate factors based on climate zones to design buildings with passive cooling techniques for energy efficiency in the future. Indonesia already has eight climate zones for passive cooling, namely equatorial, sub-equatorial, highland tropical, very highland tropical, monsoonal, sub-monsoonal, savanna, and sub-savanna (Putra et al., 2022).

One of the most important climate variables for the design and simulation of energy-efficient buildings is solar radiation (Hashemi & Khatami, 2017; D. Setiawan & Azmi, 2021; Zhou et al.,

2006). Studying the impact of solar radiation on the building envelope is one of the earliest judgments when starting to design a building. This relates to the indoor layout design and its relationship to the outdoors through windows and openings to determine heat gain. A widely adopted measure in many countries to improve energy-efficient building design is the Overall Thermal Transfer Value (OTTV) of solar radiation (Chan & Chow, 2013). The OTTV value is a measure of the average heat gain into the building through the building envelope which can be used to determine building insulation.

Solar radiation has several components, including global horizontal irradiance and global vertical irradiance with four Stunting is a condition in which the growth of children under five years old is stunted due to chronic malnutrition, causing them to have a shorter height than they should for their age. This nutritional deficiency can begin to occur since the fetus is in the womb and continue in the early days after the baby is born. However, stunting is only seen after the child reaches the age of two. Stunted toddlers have a body length (PB) or height (TB) that is lower than the standard set by WHO-MGRS (Multicentre Growth Reference Study) 2006 for their age, or according to the definition of the Ministry of Health (MoH), stunting occurs when the z-score of children under five is less than  $-2SD$  (stunted) and even less than  $-3SD$  (severely stunted) (Kemiskinan, 2017).

The age of 24-59 months is considered a critical period in the effort to ensure quality human resources, especially since the first two years of a child's life are a golden period for optimal brain growth and development (E. Setiawan et al., 2018). Therefore, serious attention is required in this period. The problem of stunting, also known as short children, is one of the most serious challenges, especially in countries with low levels of poverty and development. According to a report from the United Nations Children's Emergency Fund (UNICEF), more than half of stunted children, about 56%, live in Asia, while more than a third, or about 37%, come from the African continent. Indonesia itself still faces major challenges related to child nutrition and growth problems. UNICEF data shows that around 80% of stunted children are spread across 24 developing countries in Asia and Africa. Indonesia ranks fifth in the list of countries with the highest stunting prevalence, after India, China, Nigeria, and Pakistan. Currently, the prevalence of stunting in children under the age of 5 in the South Asia region reaches around 38%.

Nutrition issues remain a major focus in Indonesia, especially in the context of nutrition in toddlers. The health condition and nutritional status of toddlers are important indicators of overall public health. This is due to the impact caused by cases of malnutrition, undernutrition, stunting (growth delay), and other nutritional problems that are a burden for families, communities, and countries (Unicef, 2012). Some factors that are suspected of causing stunting include the mother's pregnancy history, including the mother's short posture, too close pregnancy distance, too many births, the age of the mother during pregnancy who is too old or too young (under 20 years), and insufficient nutritional intake during pregnancy. In addition, factors such as non-implementation of Early Breastfeeding Initiation (IMD), failure of exclusive breastfeeding, and early weaning

process also play a role. Economic and sanitation factors also have a correlation with the incidence of stunting (Indah, 2018). The impact of stunting includes cognitive, motor, and verbal development that is not optimal in children, increased morbidity and mortality, body posture that is not optimal in adulthood (shorter than average), and learning capacity and performance that is less than optimal at school (WHO, 2020)

Cluster analysis is one of the important methods in the field of Data Mining. Data mining is a process that uses various statistical, mathematical, artificial intelligence, and machine learning techniques to extract and identify valuable information from various large databases (Sitepu et al., 2011). Cluster analysis in data mining is a method used to group a series of data into groups based on predetermined similarities (Matdoan & Van Delsen, 2020). Furthermore, according to (Talakua et al., 2017) in their study, they explain that clustering is a grouping process aimed at generating partitions where each member shares similarity based on a specific matrix. Among the various cluster analysis methods available, two of them are K-Means and K-Medoids Clustering. Both of these methods are types of partitioning clustering methods that have interrelated algorithms. Such methods tend to be faster than hierarchical methods and more advantageous especially when the number of data objects is very large.

The K-Means clustering algorithm plays an important role in the data mining domain and is relatively simple to implement and execute. The K-Means algorithm is a distance-based clustering method that divides data into clusters, and it works primarily on numerical attributes. In practice, this algorithm is often used because of its relative speed and ability to adapt easily (Sangga, 2018).

## METHODS

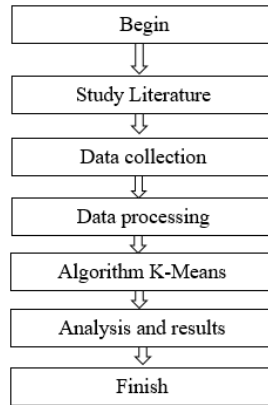
The dataset was obtained from data on the number of stunting toddlers by district/city in West Java Province from 2014 to 2022. This dataset related to the topic of Health is produced by the Health Office which is issued every 1 year.

**Table 1 Research Variables**

| Variabel | Information        |
|----------|--------------------|
| X1       | Data Stunting 2015 |
| X2       | Data Stunting 2016 |
| X3       | Data Stunting 2017 |
| X4       | Data Stunting 2018 |
| X5       | Data Stunting 2019 |
| X6       | Data Stunting 2020 |
| X7       | Data Stunting 2021 |
| X8       | Data Stunting 2022 |

The stages of research that have been carried out can be seen in Figure 1 below. The image provides a visual picture of the steps taken in conducting this research.

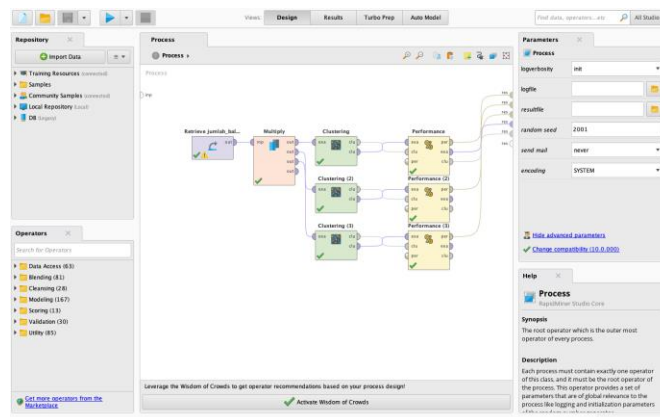
# K-Means Algorithm for District/City Classification in West Java Based on Stunting Toddler Data



**Figure 1. Research Flow**

## RESULTS AND DISCUSSION

The K-Means algorithm is a non-hierarchical method in data clustering that aims to partition data into one or more clusters or clusters. The main purpose is to group data that has similar characteristics into one cluster, while data that has different characteristics are grouped into different clusters.



**Figure 2. The process of clustering using the K-Means algorithm in RapidMiner Studio.**

Based on the results of Data Mining testing using the K-Means Clustering algorithm in the RapidMiner 10 software application, the conclusion that can be drawn is that the centroid value obtained is different due to differences in the amount of data, but the value in the resulting cluster remains the same. Details of the final centroid values can be seen in Figure 2 below:

| Attribute  | cluster_0 | cluster_1 |
|------------|-----------|-----------|
| TAHUN 2015 | 6984.294  | 20593.200 |
| TAHUN 2016 | 6154.647  | 23151     |
| TAHUN 2017 | 4704.176  | 18400.100 |
| TAHUN 2018 | 4730      | 14411.500 |
| TAHUN 2019 | 4881.176  | 14345.600 |
| TAHUN 2020 | 6080.118  | 17270.700 |
| TAHUN 2021 | 4348.294  | 13259.300 |
| TAHUN 2022 | 3318.412  | 12294.900 |

Figure 3. Centroid Value

## Cluster Model

Cluster 0: 17 items

Cluster 1: 10 items

Total number of items: 27

Figure 3. K-Means Clusterization

Based on Figure 3, it can be seen that the classification of districts/cities based on the number of stunting in West Java Province is divided into two clusters. Cluster 0 consists of 17 districts/cities, and cluster 1 consists of 10 districts/cities. More detailed information can be found in Table 2 below.

Table 2. Cluster List

| Group     | Number of Provinces | List of District   |
|-----------|---------------------|--|
| Cluster 0 | 17                  | Ciamis Regency, Kuningan Regency, Majalengka Regency, Sumedang Regency, Subang Regency, Purwakarta Regency, Bekasi Regency, West Bandung Regency, Pangandaran Regency, Bogor City, Sukabumi City, Bandung City, Cirebon City, Depok City, Cimahi City, Tasikmalaya City, Banjar City |
| Cluster 1 | 10                  | Kabupaten Bogor, Kabupaten Sukabumi, Kabupaten Cianjur, Kabupaten Bandung, Kabupaten Garut, Kabupaten Tasikmalaya, Kabupaten Cirebon, Kabupaten Indramayu, Kabupaten Karawang, Kota Bekasi   |

## CONCLUSION

Based on the findings and analysis in this study, it can be concluded that there are two main groups in the classification of stunting numbers in West Java. The first group (cluster 0) includes Ciamis Regency, Kuningan Regency, Majalengka Regency, Sumedang Regency, Subang Regency, Purwakarta Regency, Bekasi Regency, West Bandung Regency, Pangandaran Regency, Bogor City, Sukabumi City, Bandung City, Cirebon City, Depok City, Cimahi City, Tasikmalaya City, Banjar City. The second group (cluster 1) consists of Bogor Regency, Sukabumi Regency, Cianjur Regency, Bandung Regency, Garut Regency, Tasikmalaya Regency, Cirebon Regency, Indramayu Regency, Karawang Regency, Bekasi City.

## REFERENCES

- Chan, A. L. S., & Chow, T. T. (2013). Evaluation of Overall Thermal Transfer Value (OTTV) for commercial buildings constructed with green roof. *Applied Energy*, *107*, 10–24. <https://doi.org/10.1016/j.apenergy.2013.02.010>
- Dong, Q., Li, S., & Han, C. (2020). Numerical and experimental study of the effect of solar radiation on thermal comfort in a radiant heating system. *Journal of Building Engineering*, *32*(May), 101497. <https://doi.org/10.1016/j.jobee.2020.101497>
- González-Torres, M., Pérez-Lombard, L., Coronel, J. F., Maestre, I. R., & Yan, D. (2022). A review on buildings energy information: Trends, end-uses, fuels and drivers. *Energy Reports*, *8*, 626–637. <https://doi.org/10.1016/j.egy.2021.11.280>
- Hashemi, A., & Khatami, N. (2017). Effects of Solar Shading on Thermal Comfort in Low-income Tropical Housing. *Energy Procedia*, *111*(September 2016), 235–244. <https://doi.org/10.1016/j.egypro.2017.03.025>
- Indah, M. (2018). Pusat Data dan Informasi Kementerian Kesehatan RI: Tuberkulosis. *Jakarta: Kementerian Kesehatan RI*.
- Kemiskinan, T. N. P. P. (2017). Pemilihan 10 Desa Prioritas di 100 Kabupaten/Kota Prioritas Penanganan Kemiskinan dan Stunting. *Jakarta: TNPPK*.
- Marino, C., Nucara, A., & Pietrafesa, M. (2017). Thermal comfort in indoor environment: Effect of the solar radiation on the radiant temperature asymmetry. *Solar Energy*, *144*, 295–309. <https://doi.org/10.1016/j.solener.2017.01.014>
- Matdoan, M. Y., & Van Delsen, M. S. N. (2020). Penerapan Analisis Cluster Dengan Metode Hierarki Untuk Klasifikasi Kabupaten/Kota Di Provinsi Maluku Berdasarkan Indikator Indeks Pembangunan Manusia. *Statmat J. Stat. Dan Mat*, *2*(2), 20.
- Putra, I. D. G. A., Nimiya, H., Sopaheluwakan, A., Kubota, T., Lee, H. S., Pradana, R. P., Alfata, M. N. F., Perdana, R. B., Permana, D. S., & Riama, N. F. (2022). Development of climate zones for passive cooling techniques in the hot and humid climate of Indonesia. *Building and Environment*, *226*(2), 109698. <https://doi.org/10.1016/j.buildenv.2022.109698>
- Sangga, V. A. P. (2018). *Perbandingan algoritma K-Means dan algoritma K-Medoids dalam pengelompokan komoditas peternakan di provinsi Jawa Tengah tahun 2015*.

- Setiawan, D., & Azmi, M. A. (2021). Solar Radiation Study of Office Building Shapes at a Bussiness District in South Jakarta. *IOP Conference Series: Earth and Environmental Science*, 794(1). <https://doi.org/10.1088/1755-1315/794/1/012181>
- Setiawan, E., Machmud, R., & Masrul, M. (2018). Faktor-faktor yang berhubungan dengan kejadian stunting pada anak usia 24-59 bulan di Wilayah Kerja Puskesmas Andalas Kecamatan Padang Timur Kota padang Tahun 2018. *Jurnal Kesehatan Andalas*, 7(2), 275–284.
- Sitepu, R., Irmeilyana, I., & Gultom, B. (2011). Analisis cluster terhadap tingkat pencemaran udara pada sektor industri di Sumatera Selatan. *Jurnal Penelitian Sains*, 14(3).
- Talakua, M. W., Leleury, Z. A., & Taluta, A. W. (2017). Analisis cluster dengan menggunakan metode k-means untuk pengelompokkan Kabupaten/Kota di provinsi maluku berdasarkan indikator indeks pembangunan manusia tahun 2014. *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, 11(2), 119–128.
- Unicef. (2012). *The state of the world's children 2012: children in an urban world*. Esocialsciences.
- WHO. (2020). *Progress on the health related sustainable development goals and targets in the Eastern Mediterania Region*.
- Zhou, J., Wu, Y., & Yan, G. (2006). Generation of typical solar radiation year for China. *Renewable Energy*, 31(12), 1972–1985. <https://doi.org/10.1016/j.renene.2005.09.013>

---

**Copyright holder:**

Ade Bani Riyan, Ikhsan Nendi (2024)

**First publication right:**

Journal Transnational Universal Studies (JTUS)

**This article is licensed under:**

