P-ISSN: 2656-1743 | E-ISSN: 2656-1735 **JURNAL RISET INFORMATIKA**
DOI: https://doi.org/10.34288/jri.v7i3.374 Vol. 7, No. 3. Juni 2025
Accredited rank 4 (SINTA 4), excerpts from the decision of the DITJEN DIKTIRISTEK No. 230/E/KPT/2023

MODELING THE DISTRIBUTION OF HIV CASES WITH K-MEANS CLUSTERING CASE STUDY OF WEST JAVA PROVINCE

Muhammad Difa Prakoso Fuadi[1], Tukino[2], Agustia Hananto[3], Fitria Nurafriani[4]

Sistem Informasi, Fakultas Ilmu Komputer
Universitas Buana Perjuangan
Jl. HS.Ronggo Waluyo, Puseurjaya, Telukjambe Timur, Karawang, Jawa Barat 41361
Si22.muhammadfuadi@mhs.ubpkarawang.ac.id[1], tukino@ubpkarawang.ac.id[2],
agustia.hananto@ubpkarawang.ac.id[3], fitria.apriani@ubpkarawang.ac.id[4]

Abstract

In Indonesia, the problem of HIV/AIDS is a serious concern because the trend of cases tends to increase in several regions, including in West Java Province, 2018 data from the Health Office shows a significant variation in the number of HIV cases among districts and cities in the province, in this journal, a visualization process is carried out using Google Colaboratory (Google Colab) to provide an overview of the distribution pattern of cases based on the results of the K-Means Clustering algorithm. The results showed the existence of three main clusters, namely areas with low, medium, and high numbers of cases. Large cities such as Bandung and Bekasi were in the group with the highest number of cases, while peripheral and rural areas showed lower numbers of cases. This finding is expected to be the basis for formulating more effective health policies, especially in education programs, early detection, and community-based interventions to support the goal of eliminating HIV by 2030, then what can be done is to carry out intervention strategies or steps to prevent the spread of HIV tailored to the risk level of each cluster resulting from clustering analysis. Local governments are expected to utilize the results of this mapping to develop more detailed prevention strategies according to the characteristics of each region.

Keywords: *HIV; Jawa Barat; K-Means; Clustering; Visualization*

Abstrak

Di Indonesia, permasalahan HIV/AIDS menjadi perhatian serius karena tren kasus yang cenderung meningkat di beberapa wilayah, termasuk di Provinsi Jawa Barat, Data tahun 2018 dari Dinas Kesehatan menunjukkan adanya variasi jumlah kasus HIV yang signifikan di antara kabupaten dan kota di provinsi ini, dalam jurnal ini dilakukan proses visualisasi dilakukan menggunakan Google Colaboratory (Google Colab) tujuannya memberikan gambaran pola sebaran kasus berdasarkan hasil algoritma K-Means Clustering. Hasil penelitian menunjukkan adanya tiga klaster utama, yaitu daerah dengan jumlah kasus rendah, sedang, dan tinggi, Kota-kota besar seperti Bandung dan Bekasi berada dalam kelompok dengan jumlah kasus terbanyak, sedangkan daerah pinggiran dan pedesaan menunjukkan angka kasus yang lebih rendah. Penemuan ini diharapkan dapat menjadi landasan dalam merumuskan kebijakan kesehatan yang lebih efektif, terutama dalam program pendidikan, deteksi awal, dan intervensi berbasis komunitas untuk mendukung tujuan penghapusan HIV pada tahun 2030, maka yang dapat dilakukan adalah melakukan strategi intervensi atau langkah upaya pencegahan penyebaran HIV yang disesuaikan dengan tingkat risiko masing-masing klaster hasil dari analisis clustering. Pemerintah daerah diharapkan memanfaatkan hasil pemetaan ini untuk menyusun strategi pencegahan yang lebih terperinci sesuai karakteristik masing-masing wilayah.

Kata kunci: HIV; Jawa Barat; K-Means; Clustering; Visualisasi

INTRODUCTION

Human Immunodeficiency Virus or commonly known as HIV, is a virus that attacks the human immune system, especially CD4 cells (a type of white blood cell that is important for fighting infections). If left untreated, HIV can develop into AIDS (Acquired Immunodeficiency Syndrome), which is the final stage of HIV infection where the body is very weak and vulnerable to various diseases. HIV is still one of the world's public health problems. According to the UNAIDS report, despite a global decline in new cases, there are still millions of people living with HIV, especially in developing countries (UNAIDS, 2023). In Indonesia, the problem of HIV/AIDS is a serious concern because the trend of cases tends to increase in several

regions, including in West Java Province (Indonesian Ministry of Health, 2023)..

Data from (West Java Provincial Health Office, 2018) states that West Java, as the province with the largest population in Indonesia, has high social dynamics and population mobility. These conditions increase the risk of HIV transmission, especially in urban and peri-urban areas. The 2018 data from the Health Office shows significant variation in the number of HIV cases among districts and cities in the province, indicating the need for a data-driven approach to understanding the distribution. Without accurate mapping, interventions are often generalized and poorly targeted.

In recent years, data mining approaches have proven effective in analyzing health data, especially in detecting hidden patterns and making data-driven predictions. One of the commonly used methods in spatial data exploration is the K-Means Clustering algorithm, which is able to group data based on the similarity of its characteristics without supervision. With this method, areas with a high number of HIV cases can be identified as specific clusters, allowing intervention policies to be focused more efficiently.(Ismayati, Rifai, & Rahayu, 2023).

This journal aims to model the distribution of HIV cases in West Java Province using the K-Means Clustering algorithm. This study contributes to the local literature on the application of data mining in epidemiology and supports the achievement of HIV elimination targets by 2030 as stated in the Sustainable Development Goals..

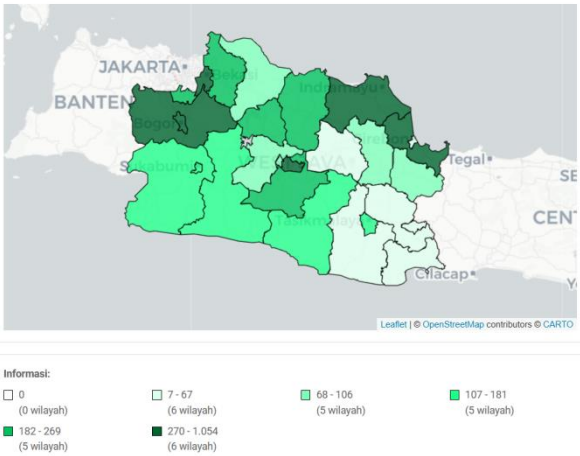

Figure 1. West java map
(sumber : https://opendata.jabarprov.go.id/id/dataset/jumlah-kasus-hiv-berdasarkan-kabupatenkota-di-jawa-barat)

.

RESEARCH METHODS

Visualisasi Data Methods

Data visualization is used to clarify the results of clustering the number of HIV cases by district / city in West Java Province. This visualization aims to provide an overview of the distribution pattern of cases based on the results of the K-Means Clustering algorithm.

The K-Means Clustering method was chosen because the process is simple, fast, and suitable for numerical data such as the number of HIV cases. The clustering results are easy to understand because they group areas based on the similarity of the number of cases. This method allows for the objective identification of areas with different case rates, supporting the formulation of more targeted intervention strategies.

The visualization process was conducted using Google Colaboratory (Google Colab), a cloud platform that supports Python programming and integration of various data analysis and visualization libraries such as Pandas, Matplotlib, and Seaborn. The choice of Google Colab was based on its advantages in ease of access, not requiring local installation, supporting online collaboration, and providing enough free computing resources for the needs of this research..

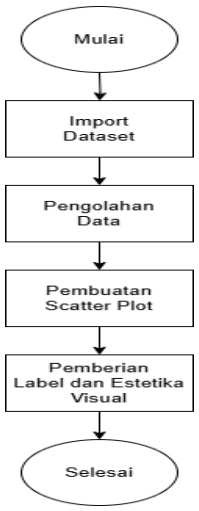

Figure 2. Flowchart visualisation data

The data visualization steps are as follows:

1. Import Dataset

Datasets containing district names, number of HIV cases, and cluster results were uploaded to Google Colab in CSV format and read using the Pandas library.

2. Data Processing

The data was processed to ensure there were no empty values and all columns were available, namely: District Name, Number of Cases, and Cluster.

3. Scatter Plot Creation

A scatter plot visualization was created using Matplotlib. This scatter plot maps:

a. X-axis: District Name/City
b. Y-axis: Number of HIV Cases
c. Dot Color: Indicates the cluster results that have been determined from the K-Means algorithm.

4. Labeling and Visual Aesthetics

The visualization is equipped with titles, axis labels, cluster legends, as well as X-axis label rotations to make the data easy to read. In this way, complex data can be understood more intuitively through graphical representations, thus supporting more effective analysis of clustering results..

Data Preprocessing

Before the clustering process, data pre-processing is carried out to ensure data quality. The pre-processing stages include:

1. Attribute Selection: Only the attributes "name_district_city" and "number_cases" were used in the analysis.
2. Missing Value Check: The data was checked to ensure that there were no blank values. The check results showed that all entries were complete.
3. Data Normalization: To avoid bias due to different data scales, the number of HIV cases was normalized using the Min-Max Scaling method.

Normalization was done using the formula:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$

where X is the original value, and $X1$ is the normalized value.

Determination of Optimal k

Determining the optimal number of clusters (k) is a crucial step in the K-Means algorithm. In this study, the Elbow Method was used to determine the most appropriate k value..

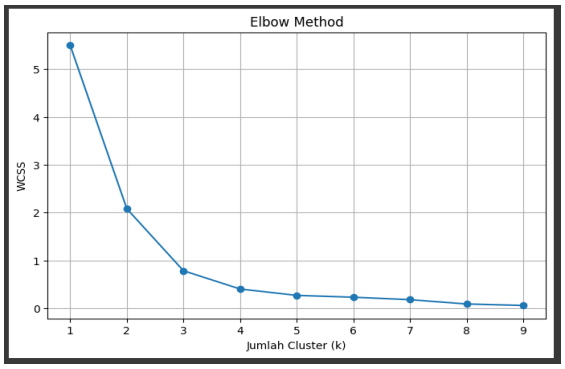

Figure 3. Metode elbow

Figure 2 is a visualization of the Elbow method used to determine the optimal number of clusters in the K-Means Clustering algorithm. In this graph, the X-axis shows the number of clusters (k) tried, while the Y-axis shows the WCSS (Within-Cluster Sum of Square) value, which is the sum of squares of the distances between data in a cluster. This method involves calculating the Within-Cluster Sum of Squares (WCSS) values for various values of k, and then selecting the point where the decrease in WCSS starts to slow down significantly ("elbow shape").

RESULTS AND DISCUSSION

Data Collection

Data collection came from the website https://opendata.jabarprov.go.id/id. This website is the official portal of public data for the West Java provincial government and contains data from the local equipment of the West Java provincial government to quickly and responsibly address the needs of local governments and to fulfill responsible information and information. On the website, researchers recorded information on HIV patients in West Java by district/city 2018-2021. By government/city, data on HIV patients in West Java. Researchers show the results of the data in the following table.

Table 1. HIV Case Data in West Java

Name Kabupaten/Kota	Number of Cases			
	2018	2019	2020	2021
Kab. Bogor	318	475	417	430
Kab. Sukabumi	112	112	110	117
Kab. Cianjur	124	0	189	111
Kab. Bandung	219	179	176	225
Kab. Garut	107	0	5	172
Kab. Tasikmalaya	40	77	95	68
Kab. Ciamis	66	95	62	58

P-ISSN: 2656-1743 |E-ISSN: 2656-1735
DOI: https://doi.org/10.34288/jri.v7i3.374
Accredited rank 4 (SINTA 4), excerpts from the decision of the DITJEN DIKTIRISTEK No. 230/E/KPT/2023

Kab. Kuningan	78	91	48	113
Kab. Cirebon	278	257	251	232
Kab. Majalengka	92	97	87	123
Kab. Sumedang	61	104	100	120
Kab. Indramayu	407	151	275	113
Kab. Subang	230	339	247	222
Kab. Purwakarta	194	197	234	130
Kab. Karawang	92	255	315	244
Kab. Bekasi	203	222	134	239
Kab. Bandung Barat	71	36	67	67
Kab. Pangandaran	48	26	16	4
Kota Bogor	446	443	364	333
Kota Sukabumi	161	57	41	43
Kota Bandung	1054	357	82	43
Kota Cirebon	65	189	342	254
Kota Bekasi	360	335	322	390
Kota Depok	227	247	220	199
Kota Cimahi	104	44	361	342
Kota Tasikmalaya	121	105	130	99
Kota Banjar	7	47	68	40

Figure 4 explains that cases in 2018 are visualized in blue, cases in 2019 are green, and cases in 2020 are orange, and the last case in 2021 is black. The visualization feature makes it easy for us to see the number of cases in each district / city by only looking at the line.

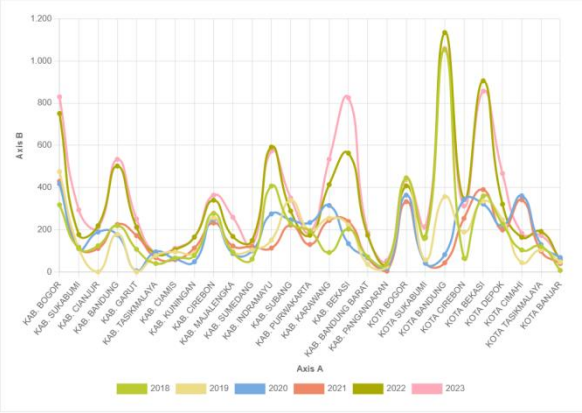

Figure 4. Graphic line chart

Determining the Centroid of the Initial Data

Based on the clustering results using the K-Means algorithm with an optimal number of clusters of three (k = 3), three groups of districts/cities in West Java Province based on the number of HIV cases were obtained.

a) Cluster 0 consists of districts with relatively low to moderate numbers of HIV cases. Examples include Sukabumi district (112 cases), Garut district (107 cases), Tasikmalaya district (40 cases), and Banjar city (61 cases). Regions in this cluster generally have a more controlled distribution of cases than other regions..
b) Cluster 1 includes areas with very high numbers of HIV cases. These include Bandung City (1054, 1135, and 1059 cases in different data sets), Bekasi City (906 and 857 cases), and Bogor District (751 and 830 cases). Regions in this cluster are the top priority for handling and intervention because the number of cases far exceeds the provincial average.
c) Cluster 2 contains districts/municipalities with medium HIV case numbers. For example, Bogor City (446 cases), Indramayu District (407 cases), Subang District (339 cases), and Cimahi City (361 cases). Regions in this cluster indicate that although there are a large number of cases, they have not reached the very high levels of cases found in cluster 1.

With this number of clusters, the K-Means model successfully grouped the data into three clusters.

Centroid	
Cluster 0	The number of cases is low
Cluster 1	The number of cases is high
Cluster 2	The number of cases is moderate

Implementation on Google Colab App

Researchers used the k-means method to group data based on the attributes of district name, number of cases and year. In the k-means method, the selected attributes become the basis for data clustering. Based on these attributes, k-means will form groups (clusters) by calculating the distance between each data and the centroid that represents the cluster. This process will group data into clusters that have similar attributes or are close to each other. By using the k-means method, researchers can gain a better understanding of the patterns and relationships in the data based on the selected attributes, as well as facilitate further analysis of each cluster formed.

Here is the code to display the results of data view clustering using google colab:

```
# Terapkan KMeans dengan k yang dipilih
k = 3  # misal hasil elbow method
kmeans = KMeans(n_clusters=k, random_state=42)
data['cluster'] = kmeans.fit_predict(X)

# Gabungkan hasil clustering ke dataframe awal
result = data[['nama_kabupaten_kota',
'jumlah_kasus', 'cluster']]
```

```
# Tampilkan tiap cluster
for i in range(k):
  print(f"\nCluster {i}:\n")
  print(result[result['cluster'] == i])
```

Explanation of the code diatass as follows:

Part 1

```
k = 3
kmeans = KMeans(n_clusters=k, random_state=42)
data['cluster'] = kmeans.fit_predict(X)
```

a. k = 3: Determine the number of clusters, here 3 is chosen, probably based on the elbow method.
b. KMeans(...): Create a KMeans object to perform clustering with 3 clusters.
c. fit_predict(X): Train the KMeans model using X data and then cluster each row of data into one of the clusters (0, 1, atau 2).
d. data['cluster'] = ...: The cluster results are inserted into a new column called cluster in the DataFrame data.

Part 2

```
Result = data[['nama_kabupaten_kota', 'jumlah_kasus', 'cluster']]
```

a. Memilih hanya kolom yang relevan: nama daerah (nama_kabupaten_kota), jumlah kasus, dan hasil cluster.

Part 3

```
for i in range(k):
  print(f"\nCluster {i}:\n")
  print(result[result['cluster'] == i])
```

a. Melakukan loop untuk setiap cluster (dari 0 sampai k-1).
b. Menampilkan data yang masuk ke masing-masing cluster menggunakan filter result['cluster'] == i.

Results Data View Clustering

The data presented is the result of the clustering process that I have done using Google Colab. This process involves processing HIV case data which is then grouped based on the similarity of certain characteristics using an unsupervised learning algorithm, namely K-Means. With the help of Google Colab, the whole process from data cleaning, to clustering modeling can be run..

Table 2. Summary of Literature Review results

No	Name Kabupaten/Kota	Number of Cases	Cluster
1	Kabupaten Sukabumi	112	0
2	Kabupaten Cianjur	124	0
3	Kabupaten Bandung	219	0
4	Kabupaten Garut	107	0
5	Kabupaten Tasikmalaya	40	0
6	Kabupaten Pangandaran	54	0
7	Kota Sukabumi	215	0
8	Kota Cimahi	183	0
9	Kota Tasikmalaya	173	0
10	Kota Banjar	61	0
11	Kota Bandung	1054	1
12	Kabupaten Bogor	751	1
13	Kota Bandung	1135	1
14	Kota Bekasi	906	1
15	Kabupaten Bogor	830	1
16	Kabupaten Bekasi	826	1
17	Kota Bandung	1059	1
18	Kota Bekasi	857	1
19	Kabupaten Bogor	318	2
20	Kabupaten Cirebon	278	2
21	Kabupaten Indramayu	407	2
22	Kota Bogor	446	2
23	Kota Bekasi	360	2
24	Kabupaten Bogor	475	2
25	Kabupaten Cirebon	257	2
26	Kabupaten Subang	339	2
27	Kabupaten Karawang	255	2
28	Kota Bogor	443	2
29	Kota Bandung	357	2
30	Kota Bekasi	335	2
31	Kota Depok	247	2
32	Kabupaten Bogor	417	2
33	Kabupaten Cirebon	251	2
34	Kabupaten Indramayu	275	2
35	Kabupaten Subang	247	2
36	Kabupaten Karawang	315	2
37	Kota Bogor	364	2
38	Kota Cirebon	342	2
39	Kota Bekasi	322	2
40	Kota Cimahi	361	2
41	Kabupaten Bogor	430	2
42	Kabupaten Karawang	244	2
43	Kota Bogor	333	2
44	Kota Cirebon	254	2
45	Kota Bekasi	390	2
46	Kota Cimahi	342	2
47	Kabupaten Bandung	502	2

48	Kabupaten Cirebon	340	2
49	Kabupaten Indramayu	592	2
50	Kabupaten Subang	289	2
51	Kabupaten Karawang	414	2
52	Kabupaten Bekasi	563	2
53	Kota Bogor	408	2
54	Kota Cirebon	346	2
55	Kota Depok	321	2
56	Kabupaten Sukabumi	294	2
57	Kabupaten Bandung	534	2
58	Kabupaten Garut	251	2
59	Kabupaten Cirebon	364	2
60	Kabupaten Majalengka	259	2
61	Kabupaten Indramayu	573	2
62	Kabupaten Subang	351	2
63	Kabupaten Karawang	534	2
64	Kota Bogor	438	2
65	Kota Cirebon	312	2
66	Kota Depok	467	2

Data Visualization Analysis

The scatter plot between district_code and number of cases shows two large clusters by administrative code range. The data indicate an unequal distribution of HIV cases, with most districts having relatively low numbers of cases, but some districts with significant spikes in cases. This pattern reinforces the clustering results, where significant differences in case numbers are a strong reason for further segmentation.

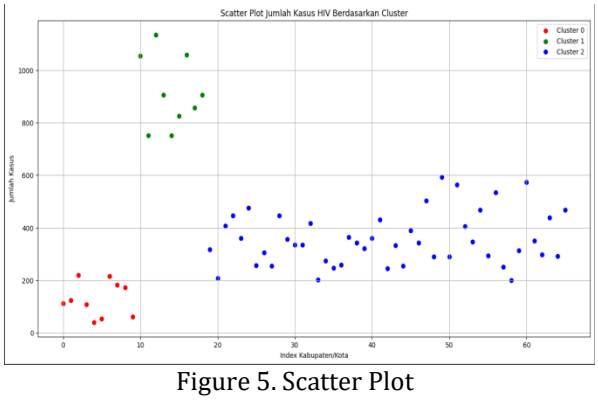

Figure 5. Scatter Plot

Additional Insights from Cluster Visualization

From the cluster mapping results, we can see that areas with high HIV case numbers are mostly metropolitan cities or urban areas such as Bandung City, Bekasi City, and Bogor City. This can be explained by several factors:

a. High mobility and urbanization in big cities leads to an increased risk of HIV transmission.
b. Population density in big cities makes transmission easier.
c. The more dynamic social, economic, and cultural activities in big cities allow for more intensive social contact, including risky behaviors.
d. Access to health services is better in big cities, so more cases are detected.

Meanwhile, districts/cities with moderate and low numbers of cases tend to be in suburban or rural areas, where the level of social interaction and urbanization is lower.

Intervention Recommendations

Based on the distribution of these clusters, intervention strategies or measures to prevent the spread of HIV can be tailored to the risk level of each cluster resulting from the K-Means clustering analysis. The three main categories of clusters are High Risk, Moderate Risk, and Low Risk, each requiring a different but integrated approach.

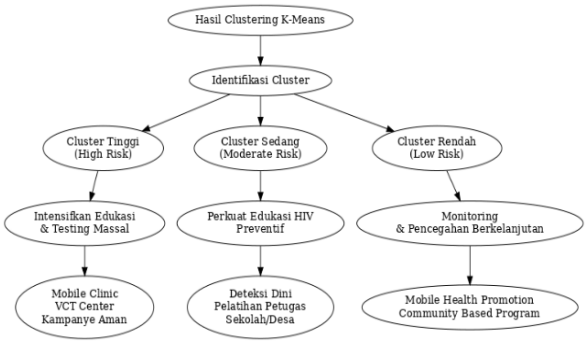

Figure 6. flowchart of hiv intervention strategies based on clustering

1) Intervention Focus in Big Cities (High Case Cluster)
   a. Expansion of HIV/AIDS education programs in vulnerable communities (youth, migrant workers, LGBTQ+ communities, etc.).
   b. Intensification of condom use campaigns and routine HIV testing.
   c. Increase Voluntary Counseling and Testing (VCT) clinic services in crowded centers.
2) Preventive Approach in Medium Case Areas (Medium Cluster)

a. Integrate HIV education in schools and workplaces.
b. Strengthening of HIV reporting and early detection networks at health centers.
c. Training of health workers to expand the reach of HIV counseling..

3) Anticipatory Efforts in Low Case Areas (Low Cluster)
a. Routine monitoring to detect rising trends in new cases.
b. Community-based health promotion campaigns with a cultural approach.
c. Mobile clinic program for hard-to-reach areas.

CONCLUSION AND SUGGESTIONS

Conclusion

This journal successfully modeled the distribution of HIV cases in West Java Province using the K-Means Clustering method, K-Means was chosen as the main method in this study because it is able to provide simple, fast, and appropriate cluster modeling for numerical data such as the number of HIV cases. With three clusters, the regions were divided into low, medium, and high case categories. The results show that HIV cases are highest in big cities such as Bandung City and Bekasi City, which are influenced by urbanization, population mobility, and better access to health services.

Based on the cluster mapping, different intervention strategies are suggested according to the level of cases in each region, ranging from intensive education in high-case areas, strengthening early detection in medium-case areas, to routine monitoring in low-case areas. This data-driven approach is expected to improve the effectiveness of HIV programs in West Java and support the achievement of HIV elimination targets by 2030..

Suggestions

For better results, this research can be further developed by using the latest data and considering other factors such as education level, income, and population mobility, so that the clustering results are more accurate. In addition, the application of other clustering algorithms such as DBSCAN or Hierarchical Clustering can be compared to see which method is most effective in mapping HIV cases.

Local governments are expected to utilize the results of this mapping to develop prevention strategies that are more specific to the characteristics of each region. Collaboration across sectors, including education, local communities, and the private sector, is also important to strengthen HIV/AIDS awareness campaigns and health services throughout West Java..

REFERENCES

UNAIDS, “Global HIV & AIDS statistics — Fact sheet,” UNAIDS, 2023. [Online]. Available: https://www.unaids.org/en/resources/fact-sheet

Kementerian Kesehatan RI, Profil Kesehatan Indonesia Tahun 2022. Jakarta: Kemenkes RI, 2023.

Dinas Kesehatan Provinsi Jawa Barat, Jumlah Kasus HIV Berdasarkan Kabupaten/Kota Tahun 2018, 2018. [Dataset].

Ariyanto, D. (2022). Data Mining Menggunakan Algoritma K-Means Untuk Klasifikasi Penyakit Infeksi Saluran Pernafasan Akut. *Jurnal Sistim Informasi dan Teknologi*, 13-18.

Dewi, D. P. (2023). Penerapan Algoritma K-means Pada Kasus HIV Di Jawa Barat Untuk Pengelompokan Berdasarkan Tingkat Penyebaran Di Setiap Kabupaten/Kota MenggunakanRapid Miner. *JTSI*, 289-300.

Gustientiedina. (2022). *Clustering Data Menggunakan Self Organizing Maps dan Algoritma K-Means.* Samudra Biru.

Heriana, C. (2015). DISTRIBUSI SPASIAL DAN DETERMINAN KEJADIAN HIV/AIDS DI PROPINSI JAWA BARAT TAHUN 2014. *JURNAL ILMU-ILMU KESEHATAN*, 1-10.

Hidayat, T. (2022). Klasifikasi Data Jamaah Umroh Menggunakan Metode K-means clustering. *Jurnal Sistim Informasi dan Teknologi*, 19-24.

Hutagalung, J. (2021). *Buku Kombinasi K-Means Clustering dan Metode MOORA.* Deepublish.

Irtawaty, A. S. (2017). Klasifikasi Penyakit Ginjal dengan Metode K-Means. *JURNAL TEKNOLOGI TERPADU*, 49-53.

Ismayati, N. (2023). Media Informasi Kesehatan untuk Pencegahan HIV/AIDS Yang Disukai Generasi Z: Upaya Penurunan Kasus HIV/AIDS di Kalangan Remaja di Indonesia . *Jurnal Ilmu Perpustakaan dan Informasi* , 54-66.

Likas, A. (2001). The global k-means clustering algorithm. *Technical Reports*, 1-13.

Masita, M. W. (2020). *Data Mining: Penerapan Algoritma K-Means Clustering dan K-Medoids Clustering.* Yayasan Kita Menulis.

Maulana, M. F. (2024). Melonjaknya Kasus HIV Dikalangan Remaja Indonesia. *AMSIR COMMUNITY SERVICE JOURNAL*, 1-7.

P.S.Bradley. (2000). Constrained K-Means Clustering. 1-9.

Penerapan Algoritma K-means Pada Kasus HIV Di Jawa Barat Untuk Pengelompokan Berdasarkan Tingkat Penyebaran Di Setiap Kabupaten/Kota MenggunakanRapid Miner. (Dica Parameswari Syifa Dewi). *JTSI*, 289-300.

Pramesti, A. A. (2024). *Penerapan Teknik Clustering dengan Metode K-Means untuk Mengelompokkan Buku di Perpustakaan AMIKOM.* Universitas AMIKOM Yogyakarta.

Romli, I. (2021). PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA K-MEANS UNTUK KLASIFIKASI PENYAKIT ISPA. *Indonesian Journal of Business Intelligence*, 10-15.

Rosida, W. (2023). Klasterisasi Penyakit HIV/AIDS di Jawa Barat Menggunakan Algoritma K-Means Clustering. *BLEND SAINS JURNAL TEKNIK* , 306-315.

Shaluhiyah, Z. (2015). Stigma Masyarakat terhadap Orang dengan HIV/AIDS . *Artikel Penelitian*, 333-339.

SINAGA, K. P. (2020). Unsupervised K-Means Clustering Algorithm. *Digital Object Identifier*, 80716-80727.

Sugiharti. (2014). GAMBARAN KEPATUHAN ORANG DENGAN HIV-AIDS (ODHA) DALAM MINUM OBAT ARV DI KOTA BANDUNG, PROVINSI JAWA BARAT. 1-11.

Suprawoto, T. (2016). KLASIFIKASI DATA MAHASISWA MENGGUNAKAN METODE K-MEANS UNTUK MENUNJANG PEMILIHAN STRATEGI PEMASARAN. *Jurnal Informatika dan Komputer*, 12-18.

Tanjung, A. A. (2024). *Perancangan Segmentasi Pasar Menggunakan K-Means Clustering untuk Pembukaan Kedai Kopi pada Bisnis Rintisan Kopi Mangandrew.* Universitas Telkom.