Grouping of Regencies and Cities in East Java Based on Adequacy of Health Human Resources with Cluster Analysis

Antonio Nikolas Manuel Bonar Simamora[a], Rudy Hardiyanto[b], Sediono[c], Azizah Andzar Ridwanah[d], Saraswati Retno Gumilang[e]

[a]*antonio.nikolas.manuel-2019@fst.unair.ac.id*
[a,b,c]*Statistics Study Program, Mathematics Department, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia*
[d]*East Java Provincial Health Office, Surabaya, Indonesia*
[e]*Faculty of Medicine, Universitas Airlangga, Surabaya, Indonesia*

Abstract

One part of the Sustainable Development Goals (SDGs) is achieving a healthy and prosperous life. In the health sector, the SDGs are realized through the Healthy Indonesia Program with three pillars: a healthy paradigm, health services, and national health insurance. One of the things that support the achievement of these three pillars is the quality of health human resources. Improving health services, optimizing management governance, and evenly distributing Human Resources for Health (HRH) are important to be considered. The insufficient availability of HRH, both in number, type, and qualifications, and uneven distribution has an impact on the low access of the community to quality health services. East Java is one of the provinces with the second largest population in Indonesia, which has a high need for HRH. An analysis of the distribution of HRH in each area of East Java Province was carried out in this study. This study classified regencies and cities in East Java based on the type of HRH. This research was conducted using hierarchical and non-hierarchical analysis methods. The optimal number of clusters was determined with the greatest Pseudo-F value, while the best cluster analysis was obtained by looking at the smallest internal cluster dispersion rate (ICD rate) and greatest R^2. After grouping, the results of the K-Means method, which included non-hierarchical cluster analysis, were able to perform the most optimal grouping with ICD rate 0.0835 and R^2 0.9165. By implementing policy strategies that are in accordance with regional clusters, efforts to achieve good health and welfare could take place efficiently and comprehensively in East Java.

Keywords: Cluster Analysis; SDGs; Human Resources for Health (HRH)

1. Introduction

Health issues are vital for a country. For this reason, all health issues are included in goal number three of the Sustainable Development Goals, namely Good Health and Well-being. The health sector has 38 SDGs to achieve [1]. Of course, in achieving this, the active role of all stakeholders, both central and regional governments, is needed. In realizing the SDGs in the health sector, the Healthy Indonesia Program is carried out with three pillars namely a healthy paradigm, health services, and national health insurance. The health service points are strengthening in terms of increasing access to health services, optimizing the referral system, management governance and improving quality [2]. In the context of improving health services and optimizing management governance, equity in Human Resources for Health (HRH) is one of the things that need attention.

According to Government Regulation of the Republic of Indonesia, Number 67 of 2019, a health worker or HRH is any person who has devoted himself to the health sector and has knowledge and or skills through education in the health sector which for certain types requires an authority to carry out health efforts. Law Number 36 of 2009 concerning Health mandates that the government is responsible for the availability of resources in the health sector that are fair and equitable for all people to obtain the highest degree of health. The insufficient availability of HRH, both in number, type, and qualifications as well as uneven distribution has an impact on the low access of the community to quality health services [3].

East Java is one of the provinces with the largest population in Indonesia, ranking second after West Java. Based on the results of the 2020 population census, the population reaches 40.67 million people or around 15 percent of Indonesia's total population [4]. However, the large number of residents does not guarantee an even distribution of facilities and HRH in each region because shortages and disparities in the distribution of health facilities between regions still occur. The availability of HRH in Community Health Centre and the types of Community Health Centre staff still do not meet the minimum standards. The development of health services also varies between regions depending on commitment, fiscal capacity, and fulfillment of facilities, human resources, facilities, and infrastructure. In this case, the most visible problem is the vacancy or shortage of health human resources [5].

Based on the Coordinating Minister for People's Welfare No. 54 of 2013 which is stated in the 2011-2025 Health Workforce Development Plan, the standard of manpower based on the ratio of staff to population is as follows:

Table 1.1. HRH Requirement Ratio Target Standard

No.	Type of Health Personnel	Target Ratio per 100,000 population		
		2014	2019	2025
1.	Specialist Doctors	10	11	12
2.	General Practitioners	40	45	50
3.	Dentists	12	13	14
4.	Nurses	158	180	200
5.	Midwives	100	120	130
6.	Dental Nurses	15	18	21
7.	Pharmacist	9	12	15
8.	Pharmacy Technician	18	24	30
9.	SKM	13	15	18
10.	Sanitarian	15	18	20
11.	Nutritionist	10	14	18
12.	Physical Therapy	4	5	6
13.	Medical Technician	14	16	18

This standard is used to calculate the HRH needs for the 2019-2025 Medium Term Projection. Therefore, to find out the distribution of HRH in each region of East Java Province, a study was conducted to classify regencies and cities in East Java to optimize the distribution of HRH to improve the quality of health services in East Java and reduce the unequal distribution of HRH, especially in remote areas.

analysis methods cluster hierarchical and non-hierarchical. This was done concerning research conducted by Mardianto in 2015 which compared various types of clusters both hierarchical and non-hierarchical [6]. Number of clusters optimal Pseudo-F analysis method cluster is obtained by looking at the internal cluster dispersion (ICD) rate smallest and largest in this study, district/city grouping was carried out based on 12 types of health workers including specialist doctors, general practitioners, dentists, nursing staff, midwifery staff, pharmacists, pharmaceutical technical staff, public health workers, environmental health workers, the ratio of nutrition workers, ratio physical therapists, and medical technicians.

The grouping of regencies and cities in East Java based on HRH is a novelty in this research. In addition, this research will result in the formulation of policy recommendations to improve the quality of health services in East Java Province. Thus, the acceleration of achieving the SDGs goals can be realized and the healthy and prosperous life of the community can be guaranteed.

2. Literature Review and Research Methods

2.1. Human Resources for Health

According to Minister of Health Regulation No. 33 of 2015, Human Resources for Health (HRH) are people who work actively in the health sector, whether they have formal education in health or not, who for certain types require authority in carrying out health efforts. HRH are also a major factor in increasing the competitiveness of health services. HRH are the backbone of health services that are directed at growing the number and proportion of the population of productive age and the elderly in the future [7]. HRH aims to produce health human resources who have the following competencies:

1. Able to develop and update science and technology in the field of health promotion by mastering and understanding scientific approaches, methods and principles as well as their application skills in the development and management of health human resources.

2. Able to identify and formulate HRH development and management solutions through research activities.
3. Develop/improve professional performance through the development and advocacy of health programs and policies in the context of developing and managing health human resources through a sharp analysis of health problems.

2.2. Hierarchical Cluster

Multivariate analysis is a statistical method which aims to analyze data consisting of many variables. Multivariate analysis is widely used to overcome various problems such as data reduction, group formation, to hypothesis testing [8]. Cluster analysis is a part of multivariate analysis. It consists of hierarchical and non-hierarchical grouping. In this case, the grouping will be based on a hierarchical approach. The distance measure used is the Euclidean. Euclidean distance is determined from object-*i* to object-*j*. It is defined as follows [9].

$$d_{ij} = \sqrt{\sum_{k=1}^{p} (x_{ik} - x_{jk})^2} \tag{1}$$

Cluster formation through a hierarchical approach can be displayed through a tree diagram which is often referred to as a dendogram. This method can be applied to single linkage, complete linkage, and average linkage [10].

1. Single Linkage (Nearest Neighbor)
   Single *linkage* performs cluster formation by combining the smallest distances.

$$d_{(i,j)k} = min(d_{ik}, d_{jk}) \tag{2}$$

2. Complete Linkage (Farthest Neighbor)
   Complete linkage of all objects in the *cluster* that is farthest from each other.

$$d_{(i,j)k} = max(d_{ik}, d_{jk}) \tag{3}$$

3. Average Linkage (Unweighted Pair-Group Average)
   Average linkage is based on the distance between two *clusters* which is equal to the average distance between all pairs of objects in the cluster.

$$d_{(i,j)k} = average(d_{ik}, d_{jk}) \tag{4}$$

2.3. Non-hierarchical Cluster Analysis

One of the non-hierarchical clustering methods is K-Means. This data grouping technique is simple and fast [11]. K-Means clustering or non-hierarchical clustering uses iteration to select temporary centers (centroids) that are continuously updated until criteria are met [12]. The K-means clustering is directionless because it does not have a definite cluster. Objects that have entered a certain cluster still could be moved to another cluster. It is more suitable to mark each object entering the cluster that has the closest cluster center [13].

$$K - Means\ (J) = \sum_{i=1}^{m} \sum_{k=1}^{K} w_{ik} \left\| x^i - \mu_k \right\|^2 \tag{5}$$

Where,

$$w = \begin{cases} 1 \; if \; k = argmin_j \left\| x^i - \mu_k \right\|^2 \\ 0 \; if \; k \; does \; not \; have \; cluster \end{cases}$$

2.4. Selection Size of the Optimal Number of Clusters

Selection of the optimal number of clusters is done by looking at the value of the Pseudo-F statistic [14]. The high value of the Pseudo-F statistic for a certain number of clusters indicates that the number of clusters is optimal. If R^2 is the coefficient of determination, n is the number of samples, c is the number of clusters, then the Pseudo-F statistic is formulated as follows:

$$Pseudo - F = \frac{(n-c)R^2}{(c-1)(1-R^2)} \tag{6}$$

2.5. Criteria for Selection of the Best Cluster Method

The data used in this study is secondary data on the number of HRH from the publication entitled Health Profile of East Java Province 2020 which was published by the East Java Provincial Health Office in 2021 [16]. The data was collected from 29 regencies and 9 cities in East Java Province which were then converted into a ratio by dividing the amount of each health resource by the total population in each district or city in East Java. A total of 12 types of HRH were used as research variables which are presented briefly in Table 2.1.

$$R^2 = \frac{SST - SSW}{SST} \tag{7}$$

With

$$SST = \sum_{i=1}^{n_c} \sum_{j=1}^{c} \sum_{k=1}^{p} (x_{ij}^k - \overline{x^k})^2 \tag{8}$$

$$SSW = \sum_{i=1}^{n_c} \sum_{j=1}^{c} \sum_{k=1}^{p} (x_{ij}^k - \overline{x_j^k})^2 \tag{9}$$

The results of grouping using cluster will be better if the ICD rate is getting smaller. Mathematically, the ICD rate can be written as follows [14].

$$\text{ICD } rate = 1 - R^2 \tag{10}$$

2.6. Data and Research Variables

The data used in this study is secondary data on the number of HRH from the publication entitled Health Profile of East Java Province 2020 which was published by the East Java Provincial Health Office in 2021 [16]. The data was collected from 29 regencies and 9 cities in East Java Province which were then converted into a ratio by dividing the amount of each health resource by the total population in each district or city in East Java. A total of 12 types of HRH were used as research variables which are presented briefly in Table 2.1.

Table 2.1. Research Variables

Variable	Human Resources for Health (Ratio)
X_1	Specialist
X_2	Doctor General
X_3	Practitioner Dentist
X_4	Nursing
X_5	Worker
X_6	Pharmacist Pharmacy
X_7	Technical Staff
X_8	Community
X_9	Environmental Health
X_{10}	Worker Nutrition
X_{11}	Worker Physical
X_{12}	Therapy Worker Medical Technician Worker

2.7. Analysis Procedure

The statistical method used to analyze the data is cluster hierarchical and non-hierarchical. The analytical procedures carried out to achieve the research objectives are as follows:

1. Collect observational data.
2. Describe the characteristics of the research variables.
3. Grouping regencies and cities in East Java based on the adequacy of health human resources with the following steps:
   a. Hierarchical grouping using the single linkage, complete linkage, and average linkage.
   b. Non-hierarchical grouping using the K-Means
   c. Select clusters for each clustering method based on the Pseudo-F.
   d. Choose the cluster based on the size of the coefficient of determination and the ICD rate.
4. Visualize the best clustering results in the form of a map chart and proceed with describing the characteristics of each cluster that has been formed.
5. Formulate policy recommendations to improve the quality and optimize the distribution of health human resources in East Java based on the cluster results.

3. Results and Discussion

In this study, the analysis and discussion include the characteristics of the research variables which are presented descriptively as well as a comparison of the results of the optimal grouping of districts and cities in East Java based on the adequacy of health human resources using cluster hierarchical and non-hierarchical.

3.1. Characteristics of Research Variables

The following describes the characteristics of the data on the ratio of health human resources in East Java in 2020, which include specialist doctors (X_1), general practitioners (X_2), dentists (X_3), nursing staff (X_4), midwifery staff (X_5), pharmacists (X_6), medical staff pharmaceutical technician (X_7), public health personnel (X_8), environmental health personnel (X_9), ratio of nutrition personnel (X_{10}), ratio of physical therapy personnel (X_{11}), and medical technical personnel (X_{12}). The size of the descriptive statistics used is in the form of maximum values, minimum values, and averages, each of which is taken by one province from each research variable as shown in Table 3.1 as follows:

Table 3.1. Characteristics Variable Human Resources for Health in East Java in 2020

Variable	Max	City	Min	Regency	Average
X_1	0.00089306	City of Mojokerto	0.00002105	Regency of Sumenep	0.00019659
X_2	0.00087536	City of Blitar	0.00007988	Regency of Sampang	0.00026797
X_3	0.00029401	City of Kediri	0.00001921	Regency of Sampang	0.00007642
X_4	0.00620635	City of Madiun	0.0004732	Regency of Sampang	0.00179916
X_5	0.00156145	City of Madiun	0.00032901	Regency of Malang	0.0007497
X_6	0.00072717	City of Madiun	0.00000809	Regency of Sampang	0.00012779
X_7	0.00126269	City of Madiun	0.00001112	Regency of Sampang	0.00018357
X_8	0.00024803	City of Madiun	0.00000954	Regency of Malang	0.00005877
X_9	0.00018038	City of Madiun	0.00001138	Regency of Jember	0.000049
X_{10}	0.00032131	City of Madiun	0.00003219	Regency of Bangkalan	0.00009024
X_{11}	0.00026176	City of Mojokerto	0	Regency of Sampang	0.00003679
X_{12}	0.00051297	City of Madiun	0.00002224	Regency of Sampang	0.00012224

Table 3.1 shows that the adequacy of health human resources in City of Madiun ranks highest in most of the research variables, while the lowest order in research variables is dominated by Regency of Sampang. In fact, in 2020 there were no physical therapy personnel in Regency of Sampang. Based on this, it can be proposed that efforts to increase the adequacy of health human resources in Regency of Sampang will be a top priority compared to other cities or districts in East Java.

Before carrying out cluster, some assumptions must be fulfilled, namely that there is no detectable multicollinearity between research variables [6]. If the p-value correlation test results Pearson between the research variables is more than α (0.05), it can be concluded that the research data does not contain multicollinearity. Based on the test results, the p-value correlation Pearson for all research variables is more than α (0.05). This means that there are no symptoms of multicollinearity in the data. Thus, the next stage of analysis is hierarchical grouping using single linkage, complete linkage, and average linkage and non-hierarchical grouping using K-Means.

3.2. Hierarchical Cluster Analysis

3.2.1. Single Single Linkage

Single Linkage method forms groupings by combining the closest distance or greatest similarity of objects based on the Euclidean. Computational calculations are carried out in determining the agglomeration schedule and obtaining Pseudo-F for each cluster. The results the agglomeration schedule the single linkage are briefly presented in Table 3.2 as follows:

Table 3.2. Results of Agglomeration Schedule Method Single Linkage

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	14	23	0.000	0	0	7
2	3	8	0.000	0	0	14
3	20	22	0.000	0	0	5
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
35	1	32	0.000	33	34	36
36	1	30	0.000	35	31	37
37	1	35	0.000	36	32	0

Based on Table 3.2, at stage formed cluster consisting of the 14th and 23rd regions with a distance of 0.000. Because the agglomeration schedule starts from the 2 objects that have the highest similarity, then this distance can be said to be the closest of the 38 object combinations. The last column in stage 1 (next stage) shows the number 7. This means that the

grouping process for the next stage is carried out by observing stage 7. The process continues until the stage. After the grouping results are obtained, the Pseudo-F values are calculated for 2 to 5 clusters, the results of which are shown in Table 3.3 as follows:

Table 3.3. Pseudo-F Value for Each Many Clusters Method Single Linkage

Number of *Cluster*	Pseudo-F
2	35.91254806032
3	71.505186958832
4	105.999581609317
5	81.253514307374

Based on Table 3.3, 4 *clusters* are determined as *of clusters* optimal *single linkage* with members as follows:

a. *Cluster* 1 consists of all regencies and cities in East Java which are not members of *Clusters* 2 to 4.
b. *Cluster* 2 consists of the City of Kediri and the City of Blitar.
c. *Cluster* 3 consists of Malang City and Surabaya City.
d. *Cluster 4* consists of the City of Mojokerto and the City of Madiun.

The dendrogram in Figure 1 which illustrates the performance results of the *single linkage* in grouping is as follows:

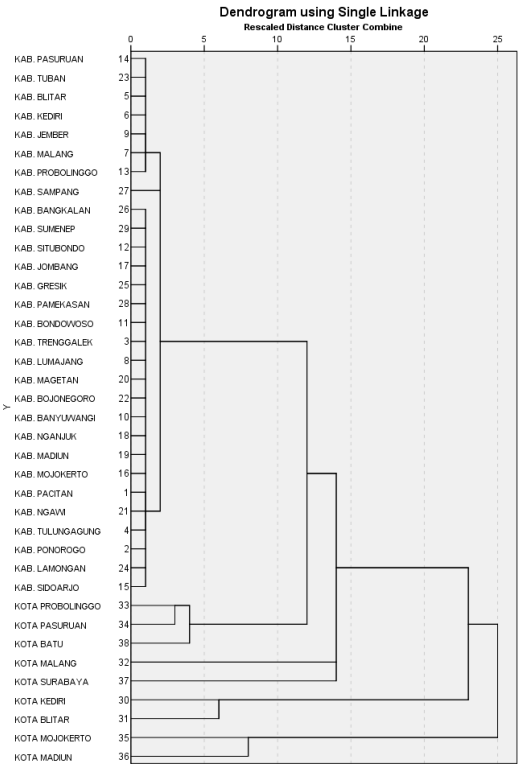

Fig. 1. Single Linkage Method Dendogram

3.2.2. Complete Linkage

The complete linkage forms groupings by combining the farthest distances between objects based on the Euclidean. Computational calculations are carried out in determining the agglomeration schedule and obtaining Pseudo-F for each cluster. The results the agglomeration schedule the complete linkage are briefly presented in Table 3.4 as follows:

Table 3.4. Agglomeration Schedule Method Complete Linkage Results

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	14	23	0.000	0	0	15
2	3	8	0.000	0	0	13
3	20	22	0.000	0	0	18
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
35	32	33	0.000	32	29	36
36	1	32	0.000	33	35	37
37	1	30	0.000	36	34	0

Based on Table 3.4, at stage formed cluster consisting of the 14th and 23rd regions with a distance of 0.000. Because the agglomeration schedule starts from the 2 objects with the highest similarity, this distance can be said to be the closest of the 38 object combinations. The last column in stage 1 (next stage) shows the number 15. This means that the grouping process for the next stage is carried out by reviewing stage 15. The process continues until the stage. After the grouping results are obtained, the Pseudo-F values are calculated for 2 to 5 clusters, the results of which are shown in Table 3.5 as follows:

Table 3.5. Pseudo-F Value for Each Many Clusters Method Complete Linkage

Number of Cluster	Pseudo-F
2	118.972455100934
3	192.109320801092
4	154.143110654385
5	170.085342025095

Based on Table 3.5, 3 clusters are determined as of clusters optimal complete linkage with members as follows:

a. Cluster 1 consists of all regencies and cities in East Java which are not members of cluster 2 and cluster 3.
b. Cluster 2 consists of the City of Kediri, City of Blitar, City of Mojokerto, and City of Madiun.
c. Cluster 3 consists of Malang City, Probolinggo City, Surabaya City, and Batu City.

The dendrogram in Figure 2 which illustrates the performance results of the complete linkage in grouping is as follows:

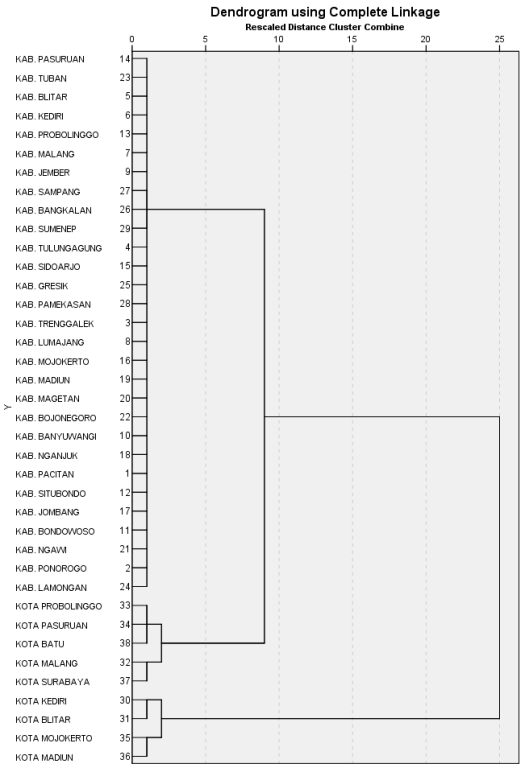

Fig. 2. Complete Linkage Method Dendogram

3.2.3. Average Linkage

The average linkage treats the distance between two clusters as the average distance between all pairs of objects where one member of the pair belongs to each cluster. Computational calculations are carried out in determining the agglomeration schedule and obtaining Pseudo-F for each cluster. The results of the agglomeration schedule method average linkage are briefly presented in Table 3.6 as follows:

Table 3.6. Results Agglomeration Schedule Method Average Linkage

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	14	23	0.000	0	0	5
2	3	8	0.000	0	0	7
3	20	22	0.000	0	0	15
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
35	1	32	0.000	34	0	37
36	30	35	0.000	31	32	37
37	1	30	0.000	35	36	0

Based on Table 3.6, at stage formed cluster consisting of the 14th and 23rd regions with a distance of 0.000. Because the agglomeration schedule starts from the 2 objects with the highest similarity, this distance can be said to be the closest of the 38 object combinations. The last column in stage 1 (next stage) shows the number 5. This means that the grouping process for the next stage is carried out by reviewing stage 5. The process continues until the stage. After the grouping

results are obtained, the Pseudo-F values are calculated for 2 to 5 clusters, the results of which are shown in Table 3.7 as follows:

Table 3.7. Pseudo-F Value for Each Many Clusters Method Average Linkage

Number of Cluster	Pseudo-F
2	118.972455100934
3	71.5051869588324
4	74.1254833961492
5	81.2535143073741

Based on Table 3.7, 2 clusters are determined as of clusters optimal average linkage with members as follows:

a. Cluster 1 consists of all regencies and cities in East Java apart from members of cluster 2.
b. Cluster 2 consists of the City of Kediri, City of Blitar, City of Mojokerto, and City of Madiun.

The dendrogram in Figure 3 which illustrates the performance results of the average linkage in grouping is as follows:

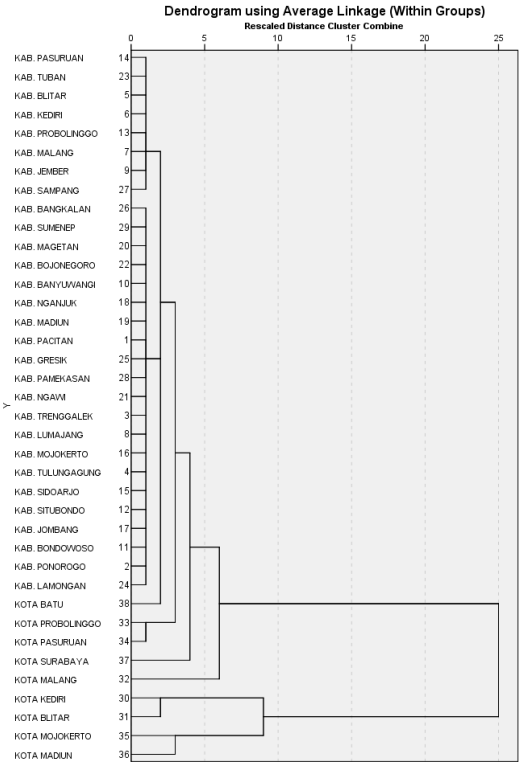

Fig. 3. Average Linkage Method Dendogram

3.3. Results of Non-Hierarchical Cluster Analysis

Non-hierarchical grouping was carried out using the K-Means method. The procedure for the K-Means is determined clusters first. Analysis algorithm cluster using the K-Means implemented in this study is as follows:

1. Determine *k* as the number of clusters to be formed.
2. Generate *k* centroids (cluster) randomly.
3. Calculate the distance of each observation to each centroid and proceed by grouping each observation to the nearest centroid
4. Calculates the average of observations located at centroids to determine the location of centroids.
5. The iteration continues until there are no more changes in the grouping.
6. Calculations to obtain Pseudo-F values can also be performed using the K-Means to determine of clusters the optimal table 3.8 displays the Pseudo-F values for 2 to 5 clusters, namely as follows:

Table 3.8 Value *Pseudo-F* for Each Many *Clusters* Method *K-Means*

Number of Cluster	Pseudo-F
2	157.037428190796
3	192.109320801094
4	176.165461929765
5	170.085342025097

Based on Table 3.8, 3 clusters are determined as of clusters optimal K-Means with the following members:

1. Cluster 1 consists of the City of Kediri, City of Blitar, City of Mojokerto and City of Madiun.
2. Cluster 2 consists of Malang City, Probolinggo City, Pasuruan City, Surabaya City, and Batu City.
3. Cluster 3 consists of all regencies and cities in East Java that are not grouped in cluster 1 and cluster 2.

3.4. Selection of the Best Cluster Analysis Method

In this analysis phase, the best cluster method is determined which gives the most optimal grouping results for regencies and cities in East Java based on the adequacy of health human resources. The ICD rate value is used to determine the goodness of clustering produced by a cluster method. If the ICD rate value is smaller, it can be said that the method is getting better at grouping. The value of R^2 and the ICD rate which shows the goodness of each hierarchical and non-hierarchical cluster analysis method are presented in Table 3.9 as follows:

Table 3.9 Value R^2 and ICD Rate in Each Cluster

Types of Cluster Analysis	Methods	Number of Cluster	R^2	ICD Rate
Hierarchical	Single Linkage	4	0.903408746481132	0.0965912535188681
	Complete Linkage	3	0.916511346284040	0.0834886537159603
	Average Linkage	2	0.767700653793262	0.2322993462067380
Non-Hierarchical	K-Means	3	0.916511346284041	0.0834886537159595

Table 3.9 shows that the K-Means analysis cluster is capable of grouping regencies and cities in East Java the most optimal among the methods other. This is evidenced by the ICD rate in the K-Means, which has the lowest value (0.083) with the highest value, which is 0.917. Thus, it was decided to choose the clustering results from the K-Means, which formed 3 clusters for further factual interpretation.

3.5. Characteristics of District and City Clusters Formed in East Java

Each cluster in East Java, based on the adequacy of health resources, has various characteristics. Optimal grouping results are visualized in the form of a map presented in Figure 4 as follows:

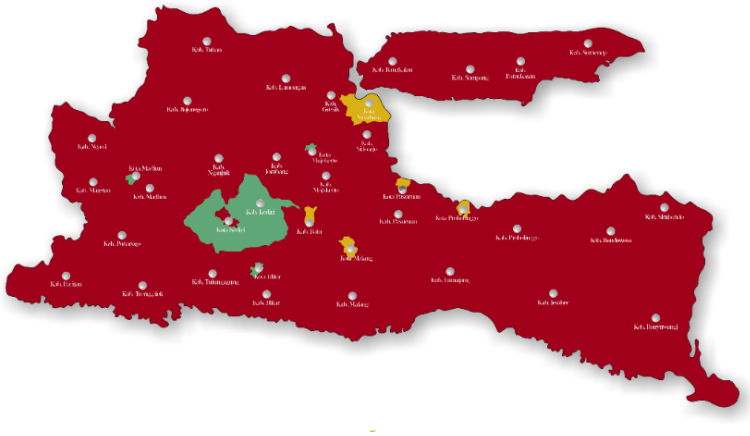

Fig. 4. Mapping of Regencies and Cities in East Java Based on Adequacy of Health Resources

The mapping graph shown in Figure 1 shows that as many as three colors are used to indicate the level of adequacy of health resources in East Java according to the grouping results using the K-Means method. Starting from green for cluster 1, which indicates regencies and cities in East Java with a high HRH adequacy level, up to maroon for cluster 3, which indicates regencies and cities in East Java with a low HRH adequacy level.

Efforts to overcome the problem of HRH adequacy in East Java must be carried out in accordance with the problems faced by each region. Therefore, identification of the characteristics of each district and city cluster in East Java is carried out so that efforts to distribute and improve the quality of human and health resources in East Java can be carried out on target. The characteristics of each cluster are reviewed based on the Decree of the Coordinating Minister for People's Welfare of the Republic of Indonesia Number 54 of 2013 concerning the Health Workforce Development Plan for 2011-2025 and is presented in Table 3.10.

Table 3.10 Characteristics Clusters in East Java Based on Adequacy of Health Resources

***Cluster* 1: Regions with a High Level of Adequacy of Human Health Resources** Members: City of Kediri, City of Blitar, City of Mojokerto, and City of Madiun. **Characteristics *of Cluster* 1** a. The HRH ratio in *cluster* is relatively high *clusters* . b. Based on the results of calculating the number of HRH per 100,000 population which is then compared with the target of the 2025 Health Workforce Development Plan, most of the HRH in *cluster* are available in adequate numbers, except for community health workers and environmental health workers.
***Cluster* 2: Regions with Adequacy Level of Human Health Resources Classified as Moderate** Members: Malang City, Probolinggo City, Pasuruan City, Surabaya City, and Batu City. **Characteristics *of Cluster* 2** a. The HRH ratio in this cluster is in the middle category among the other clusters. b. Some of the HRH in this cluster are available in adequate numbers, but this is lower when compared to cluster 1. This can be seen from the number of HRH per 100,000 residents in this cluster who have reached the target of the 2025 Health Workforce Development Plan, except for community health workers, environmental health workers , obstetrics, and nutritionists.
***Cluster* 3: Regions with a Low Adequacy Level of Human Health Resources** Members: Pacitan Regency, Ponorogo Regency, Trenggalek Regency, Tulungagung Regency, Blitar Regency, Malang Regency, Lumajang Regency, Jember Regency, Banyuwangi Regency, Bondowoso Regency, Situbondo Regency, Probolinggo Regency , Pasuruan Regency, Sidoarjo Regency, Mojokerto Regency, Jombang Regency, Nganjuk Regency, Madiun Regency, Magetan Regency, Ngawi Regency, Bojonegoro Regency, Tuban Regency, Lamongan Regency, Gresik Regency, Bangkalan Regency, Sampang Regency, Pamekasa Regency, Sumenep Regency, and Kediri Regency. **Characteristics *of Cluster* 3** a. The HRH ratio in *cluster* is the lowest among *clusters* . b. Regencies and cities in *cluster* have not met sufficient health resources. This can be seen from the number of HRH per 100,000 population in *cluster* which has not yet reached the target of the 2025 Health Workforce Development Plan.

3.6. Policy Recommendations to Improve the Quality of Human Resources for Health in East Java

The performance of the health sector can increase if each district and city in East Java maximizes the HRK they have. By considering various research variables as a whole, the characteristics of HRH in each region can be clearly described. Therefore, policy recommendations can be proposed based on the role that can be played by the government and SDMK in realizing good health and prosperity in East Java. In this case, policy recommendations are intended for each cluster that has been formed from the results of the K-Means method and are presented in Table 3.11.

Table 3.11 Policy Recommendations related to Fulfillment of HRH Needs

***Cluster*: Regions with a High Adequacy Level of HRH** *a.* Improving HRH quality supervision and development in providing services at every health facility. *b.* Increasing the fulfillment of HRH, especially for public health workers and environmental health workers. This is necessary in order to increase the degree of public health by promoting health-oriented development and making *cluster* 1 an area with a healthier environment.
***Cluster* 2: Regions with Moderate Adequacy of HRH** a. Increase the fulfillment of HRH, especially public health workers, environmental health workers, midwifery workers, and nutrition workers. b. Optimizing the implementation of subsidies, scholarships, and official bonds for the education of public health workers, environmental health workers, and nutrition workers.
***Cluster* 3: Regions with a Low Adequacy Level of HRH** a. Optimizing the implementation of adequate financial and non-financial incentive systems to maintain HRH working in areas within the *cluster*. b. Increase outreach and advocacy from the Ministry of Health to regional governments in *clusters* to increase HRH formation and recruitment. c. Strengthening regulations to ensure procurement, quality, and equitable distribution of HRH.

By implementing policy strategies that are in accordance with regional clusters, efforts to realize good health and prosperity in Indonesia can take place efficiently and comprehensively in Indonesia, especially in East Java.

4. Conclusion

Research shows that by using the best cluster analysis method, the most optimal K-Means method is obtained for grouping districts and cities in East Java. This is evidenced by the value of the ICD rate which has the lowest value and the highest R2 value with values of 0.835 and 0.9165 respectively. By using K-Means, 3 optimal clusters were obtained covering areas with high, medium and low HRH adequacy levels. Furthermore, by knowing the district and city groups in East Java and having formulated recommendations for each cluster, this research is expected to be useful as material for consideration in the evaluation of the local government, especially the East Java Provincial Health Office to improve health services including the distribution of HRH in every region in East Java.

Acknowledgements

Authors express their deepest gratitude to Faculty of Science and Technology Universitas Airlangga, East Java Provincial Health Office, and Faculty of Medicine Universitas Airlangga who have contributed to the implementation of research and the preparation of this scientific article.

References

[1] Kementerian PPN/Bappenas, “Kehidupan Sehat dan Sejahtera”, diambil dari SDGs Bappenas: http://sdgs.bappenas.go.id/laporan-sdgs-2019/, 2019.
[2] Kementerian Kesehatan Republik Indonesia, “Program Indonesia Sehat dengan Pendekatan Keluarga”, Diambil dari Kemkes Web Site: https://www.kemkes.go.id/article/view/17070700004/program-indonesia-sehat-dengan-pendekatan-keluarga.html, 2017.
[3] BPPHRH Kementerian Kesehatan, “Rencana Aksi Kegiatan Pusat Perencanaan dan Pendayagunaan SDM Kesehatan 2015-2019”, diambil dari: http://bppHRH.kemkes.go.id/pusdikHRH/wp-content/uploads/2017/10/RAK-2015-2019.pdf, 2017.
[4] V. B. Kusnandar, “Empat Wilayah di Jawa Timur yang Berpenduduk di Atas 2 Juta Jiwa”, diambil dari Databoks Katadata: https://databoks.katadata.co.id/datapublish/2021/07/13/empat-wilayah-di-jawa-timur-yang-berpenduduk-di-atas-2-juta-jiwa#:~:text=Selain%20Surabaya%2C%20Jawa%20Timur%20terdapat,Sidoarjo%202%2C08%20juta%20jiwa.
[5] Kementerian PPN/Bappenas, “Buku Putih Reformasi Sistem Kesehatan Nasional”, diambil dari Perpustakaan Bappenas: https://databoks.katadata.co.id/datapublish/2021/07/13/empat-wilayah-di-jawa-timur-yang-berpenduduk-di-atas-2-juta-jiwa#:~:text=Selain%20Surabaya%2C%20Jawa%20Timur%20terdapat,Sidoarjo%202%2C08%20juta%20jiwa, 2022.
[6] M. F. F. Mardianto, Kuzairi, T. Yulianto dan R. Amalia, “Pengelompokan Optimal Kabupaten dan Kota Rawan Kriminalitas di Jawa Timur dengan Metode Analisis Kluster Terbaik”, Zeta-Math Journal, vol 1, no 1, pp. 22-29, 2015.
[7] A. Putri, “Kesiapan Sumber Daya Manusia Kesehatan dalam Menghadapi Masyarakat Ekonomi Asean (MEA)”, Jurnal Medicoeticolegal dan Manajemen Rumah Sakit, vol 6, no 1, pp. 55-60, 2017.

[8] J. F. Hair, W. C. Black, B. J. Babin, dan R. E. Anderson, “Multivariate Data Analysis”, 8th ed., vol. 34, Gewerbestrasse: Springer, 2019.
[9] A. Afifi, S. May, R. A. Donatello, dan V. A. Clark, “Practical Multivariate Analysis”, 6th ed., Boca Raton: Taylor & Francis Group, 2020.
[10] S. T. Wierzchon dan M. A. Klopotek, “Modern Algorithms of Cluster Analysis”, Studies in Big Data, vol. 34, Gewerbestrasse: Springer, 2018.
[11] Rosmini, A. Fadlil, dan Sunardi, “Implementasi Metode K-Means Dalam Pemetaan Kelompok Mahasiswa Melalui Data Aktivitas Kuliah”, IT Journal Research and Development, pp. 22-31, 2018.
[12] C. P. Kappe, M. Böttinger, H. Leitte, “Analysis of Decadal Climate Predictions with User-guided Hierarchical Ensemble Clustering”, In Computer Graphics Forum, vol. 38, no. 3, pp. 505-515, 2019.
[13] D. M. Strauss, “Clustering Techniques in Climate Analysis”, In Oxford Research Encyclopedia of Climate Science, 2019.
[14] S. P. Rahayu, S. Febryani, H. B. Kusuma, Suhartono, dan D. D. Prasetyo, “The Approach of Fuzzy C-Means Cluster and Factor Analysis on Economic Potential Mapping of Regency/City in East Java Province”, International Conference on Science and Applied Science (ICSAS), 020082, pp. 1-8, 2018.
[15] R. J. Hayes dan L. H. Moulton, “Cluster Randomised Trials”, 2nd ed., London: Chapman & Hall/CRC Biostatistics Series, 2017.
[16] Dinas Kesehatan Provinsi Jawa Timur, “Profil Kesehatan 2020”, Surabaya: Dinas Kesehatan Provinsi Jawa Timur, 2021.