ISSN: 2807-4998 (online)

Buletin Ekonomika Pembangunan
https://journal.trunojoyo.ac.id/bep
Vol. 5 No.1 February 2024, Hal. 8-20

Cluster Analysis of Low, Medium and High Social Welfare Effect Groupings in Malang City

M. Tafana Imania Arifin[1], Azzahra Mubyaring Putri Sayekti[2], Jordan Putra Cahyono[3]*
[1]Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Brawijaya, Malang, Indonesia
[2,3] Faculty of Economics and Business, Master of Economics, Airlangga University, Surabaya, Indonesia
Email: jordan.putra.cahyono-2023@feb.unair.ac.id
DOI: https://doi.org/10.21107/bep.v5i1.25409

ABSTRACT
Malang City is one of the cities in Indonesia that is facing the problem of poverty. One of the efforts made by the Malang City government is to provide Family Hope Program (PKH) assistance aimed at improving community welfare. In order for the allocation of assistance to be evenly distributed and directed, it is necessary to group villages based on the social welfare of the community through PKH data. This study aims to group Malang city villages based on PKH using average linkage clustering, obtained three clusters with low, medium and high social welfare. Clusters with high social welfare consist of 54 urban villages and get a lot of PKH assistance. Clusters with low social welfare consist of 2 urban villages and get little PKH assistance. Meanwhile, there is 1 urban village with low social welfare and little PKH assistance.

Keywords: *Average Linkage, Poverty, Social Welfare, Malang City*

INTRODUCTION

Statistics is the study of data management that is used for problem solving so that decisions and policies can be made. The application of statistics can be utilized in various aspects, such as health, social, economics, administration, and government. As an applied science, statistics can be applied in the form of Field Work Practice (PKL) activities. PKL is a forum for students in an effort to apply the knowledge of statistics that has been obtained during lectures and add insight into the world of work. Students are expected to be able to solve problems that occur in the agency where PKL.

The Family Hope Program is a conditional cash transfer program for very poor families who meet the requirements and have been regulated by the Ministry of Social Affairs. The Indonesian Ministry of Social Affairs explains that the conditions for families who are entitled to become PKH recipients are poor families with one or more PKH requirements, namely: (1) In families with pregnant women / children under five years of age; (2) Having children under 7 years of age who are not attending primary school; (3) Having children between 7 and 21 years of age who have not completed 12 years of compulsory education; (4) Having children with disabilities from 0 to 21 years of age. From this statement, it can be concluded that PKH beneficiary families are families who meet the criteria set by the government.

The success of PKH assistance in Malang City can be supported and

assisted by grouping villages based on the requirements of PKH recipients who will form groups according to the level of social welfare. Thus, it is hoped that the grouping can be a reference in providing PKH assistance so that it is more targeted and right on target according to the level of social welfare in each village. The method used to group urban villages is cluster analysis. In this study, a hierarchical cluster is used, namely the Average Linkage method. Average Linkage Method.

Average Linkage Clustering is very appropriate to apply in this case because the data has many extreme values. This can occur from the state of the data in the field and not because of data collection errors. This research was carried out through the opportunity given to the author through the Field Work Practice program carried out at the Social Service of Malang City, which is one of the applications of Statistics in the social field, namely by grouping villages based on PKH beneficiary families. The results of this research are expected to provide benefits to readers and the P3AP2KB Social Service of Malang City so that the distribution of PKH assistance can be more precise, targeted and equitable according to the level of community welfare.

LITERATURE REVIEW

Poverty

According to the Central Bureau of Statistics, poverty is the inability to meet the minimum standard of basic needs, which includes food and non-food needs. Comparing the consumption level of the population with the poverty line or the amount of rupiah for person consumption per month.

Poverty is seen from the aspect of social inequality, because there are people who have been able to meet the minimum basic needs but are still much lower than the surrounding community. The greater the income inequality between the upper and lower classes, the more the number of people categorized as poor, so poverty is relatively closely related to income distribution problems. Poverty is a problem faced by all countries (Atalay, 2015). Economic growth as one of the indicators in overcoming the problem of poverty, where economic growth is the concept of economic development. Poverty is a situation that cannot fulfill basic needs such as food, shelter, clothing, education, and health (Suryawati, 2010). In a proper sense, poverty is understood as a state of lack of money, goods, and basic needs to ensure and fulfill survival. Relative poverty in terms of social inequality is a situation where a person meets his minimum basic needs but is still lower when compared to the surrounding community (Adji, 2020).

Family Hope Program

The Family Hope Program (PKH) is one of the social protection programs in Indonesia in the form of social assistance. This assistance is provided to poor and vulnerable poor families with certain requirements where they are registered in the Integrated Social Welfare Data (DTKS). Since its launch in 2007, PKH has contributed to reducing poverty and encouraging the independence of social assistance recipients, hereafter referred to as Beneficiary Families. The main objective of the Family Hope Program is to reduce poverty and improve the quality of human resources, especially the poor. PKH is in the form of assistance to pregnant women, school children, people with disabilities, the elderly, and early childhood.

Gap Research

Poverty is one of the most important issues facing the second largest city after Jakarta, Malang. In an effort to reduce the poverty rate, the Malang City Government through the Social Service is committed to alleviating poverty through improving the Integrated Data of Malang Communities. Through this data, the government can distribute Family Hope Program assistance evenly in each urban village. One of the problems in distributing PKH assistance is equity, for this reason a study is needed that can be used to help group urban villages based on data on PKH recipient families so that groups will be formed according to the level of social welfare.

Regarding the problem of many urban villages in Malang City, in order to optimize each environmental program, we can use urban village clustering based on the Family Hope Program. The analysis method used is cluster analysis of the average linkage method which can be used to form several groups based on the characteristics of each village. The next stage is to define the characteristics of each group which is expected to be a solution for the Malang City Government in conducting socialization planning so that many people support and implement the improvement program with the right target in each kelurahan area.

RESEARCH METHODS

Descriptive Statistics

Statistics is a collection of techniques or methods that are useful for making decisions about a process or population based on the analysis of information contained in samples from that population (Montgomery, 2009). Statistical methods are divided into two, namely descriptive statistics and inferential statistics. Descriptive statistics is statistics that talks about data or describes data without making conclusions with analysis in it (Yitnosumarto, 1990). In this descriptive statistics there is no need to compare which variable is better. Data collection can be done with group or non-group data in the population under study.

Data is a set of facts or figures that can be used as the basis of a study or research. According to its nature, data is divided into two types, namely qualitative and quantitative. Both qualitative and quantitative, data must be processed or analyzed in order to obtain useful information. Average is a measure of data centralization. The average ($\bar{y}$) describes the variable value of most objects.

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n} \qquad (2.1)$$

Description:
yi: The value of the i-th object variable.
n: Sample size.

Assumption Test
There are two assumptions in clusters, namely:
The sample can represent the population

1. The Kaiser-Meyer-Olikin (KMO) test is used to see whether the sample can represent the population or not. If the KMO test value ranges from 0.5 to 1, the cluster analysis process can continue. Meanwhile, if it is less than 0.5 then further action is needed. The formula that can be used to find the KMO

ISSN: 2807-4998 (online)

value is found in Equation (2.2).

$$KMO = \frac{\sum_{i=1}^{p}\sum_{j=1}^{p} r_{ij}^2}{\sum_{i=1}^{p}\sum_{j=1}^{p} r_{ij}^2 + \sum_{i=1}^{p}\sum_{j=1}^{p} a_{ij}^2} \quad (2.2)$$

For i ≠ j, i = 1,2, ... ,p; j = 1,2, ...,p
Description:
rij: simple correlation coefficient between the i-th and j-th variables.
aij: partial correlation coefficient between the i-th and j-th variables.

2. Non Multicollinearity
Multicollinearity is the presence of a perfect or definite linear relationship between some or all of the variables (Gujarati, 2009), Multicollinearity should not occur between variables unless there is a time series in the data. However, since cluster analysis does not involve time series, it should first test whether there is multicollinearity between variables.

One way to identify the presence of multicollinearity is to calculate the Variance Inflation factori (VIF) value.

$$VIF_j = \frac{1}{1-R_j^2} \quad (2.3)$$

Description:
R^2_j = coefficient of determination of the jth variable.
Multicollinearity is indicated if the VIF value is> 10. One way that can be done if there is multicollinearity is to remove the correlated variables.

3. Cluster Analysis
Cluster analysis is a technique of grouping a collection of objects that have certain similarities into the same cluster (Han and Kamber, 2006). This grouping is based on the similarity of several objects to then be classified into the same group, so that objects in one cluster are homogeneous. In this technique, it is expected that objects between clusters have very little similarity, in other words, the characteristics of objects between other clusters are heterogeneous (Hair et al., 2006).
There are two basic classifications in cluster analysis, namely hierarchical methods and non-hierarchical methods. The difference between the two methods lies in determining the number of clusters to be formed (Simamora, 2005). In the hierarchical method, the determination of the number of clusters is carried out through a gradual grouping process to form a hierarchical level that resembles a branching tree (Safitri et al., 2012). Whereas in the non-hierarchical method, the number of clusters is determined first manually.
In the hierarchical method, there are two approaches used, namely agglomerative and dispersive. The merging approach consists of 5 methods, namely the single linkage method, the complete linkage method, the average linkage method, the centroid method, and the Ward method (Supranto, 2004). While the method that is classified as a non-hierarchical method is the K-Means method.

4. Average Linkage Method
The average linkage method is used to group two objects based on the average distance obtained by calculating the average of all object distances

first. The similarity value between clusters in the average linkage method is considered more stable and unbiased when compared to other methods. The distance between clusters can be calculated through equation (2.3).

$$d_{(UV)} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij}}{n_U n_V} \quad (2.4)$$

Description:
dij : distance between the i-th object and the j-th object.
nU : number of objects in cluster U.
nV : number of objects in cluster V.

5. Distance Size
Distance measure is a measure of dissimilarity between objects. The calculation is based on the distance between a pair of objects. The larger the distance of a pair of objects, the smaller the similarity of a particular pair of objects. Vice versa, the smaller the distance of a pair of objects, the greater the similarity of a pair of metric-scale objects. There are several distances that can be used to group objects, including Euclidean distance, squared Euclidean distance, Manhattan distance, chebychev distance and power distance (Simamora, 2005).

6. Euclidean distance
Euclidean distance is one of the most commonly used object similarity measurement distances to group objects in cluster analysis. The concept of calculating the Euclidean distance is to measure the root of the sum of the squared differences in the values of a pair of objects that you want to know the distance between. (Simamora, 2005) explains the Euclidean distance calculation formula, which is as follows.

$$d_{ij} = \sqrt{\sum_{k=1}^{n}(x_{ik} - x_{jk})^2} \quad (2.5)$$

Description:
d_{ij} : Distance between object *i* and object *j*.
x_{ik} : the value of object *i* in the *kth* variable.
x_{jk} : the value of object *j* in the *kth* variable.
n : number of variables observed

7. Cluster Validity
Cluster validity is a procedure to evaluate the results of cluster analysis formed whether it is able to explain and represent the population in general (Hair, 2006). Cluster validity can be used to determine the number of groups that can provide the best (optimal) results. The optimum group is a group that has a dense distance or the shortest distance between individual objects in the cluster and a long distance from other clusters (Dubes and Jain, 1988). There are several indices that can be used in cluster validity, including connectivity, dunn, and silhouette indices.
   - Connectivity Index
     The formula for calculating the connectivity index (conn) is as follows:

$$Conn = \sum_{i=1}^{n}\sum_{j=1}^{k} x_{i,nn_{i(j)}} \quad (2.6)$$

Buletin Ekonomika Pembangunan
https://journal.trunojoyo.ac.id/bep
Vol. 5 No.1 February 2024, hal.8-20
ISSN: 2807-4998 (online)

Description:
Conn: *connectivity* index.
$nn_{i(j)}$ *: nearest neighbor* observations of object number one
j the i-th object.
n: number of objects.
k: number of *clusters.*
If objects are in the same cluster, then $\mathrm{nn}j$ will be zero. Meanwhile, if the object is in a cluster that is then nn will be worth 1/j. the smaller the connectivity index value, the better the number of clusters formed (Halim and Widodo, 2017).

- Dunn Index
Dunn's index is a comparison of the minimum distance between observations in different clusters with the maximum distance from each cluster (Irwansyah and Faisal, 2015). The dunn index will produce the most optimal cluster when it has a large value.

$$C = \frac{d_{min}}{d_{max}} \quad (2.7)$$

Description:
$dmin$: the smallest distance between observations in different clusters.
$dmax$: the largest distance in each data cluster.

- Silhouette Index
The usefulness of the sillhoutte index is to measure how well an object can be placed in a cluster. The silhouette index is calculated as the degree of confidence in the grouping of objects in the cluster, which is said to be well formed if the index value is close to 1 and the opposite condition if the index value is close to -1 (Irwansyah, 2015).

$$s(i) = \frac{b(i)-a(i)}{\max\,(a(i),b(i))} \quad (2.8)$$

Description:
$a(i)$: the average distance between i and all other observations in the same cluster.
$b(i)$: average distance between i and observations in the cluster
Nearby

Data and Methods

The data used in this study is data on the PKH Component of Malang City for the October-November 2023 period obtained from the Malang City P3AP2KB Social Service attached in the appendix. The analysis method used in this report is hierarchical cluster analysis of the average linkage method. Data processing is done using Rstudio 4.1.2 software.
The steps used in data analysis are as follows.

1. Collect data obtained from the Social Service P3AP2KB Malang City and determine the variables that will be used for research.
2. Perform descriptive statistical analysis on equation (2.1) in order to determine the average, minimum and maximum values.
3. Testing cluster assumptions, namely the assumption of sample adequacy that

can represent the population using the formula in equation (2.2), namely the Kaiser-Meyer-Olkin (KMO) value is greater than 0.5.
4. Testing cluster assumptions, namely the assumption of non-multicollinearity using the correlation coefficient between variables with a Variance Inflation Factorori (VIF) value smaller than 10 in the equation formula (2.3).
5. Measuring the similarity of observation objects by calculating the euclidean distance to group similarities between objects using the equation formula (2.5).
6. Measuring cluster validity and determining the best number of clusters using the shortest and densest distance clustering with connectivity, dunn, and silhouette indices. The formulas that have been explained are equations (2.6), (2.7), and (2.8).
7. Truncate according to with manybest clusters.
8. Identify the characteristics of each cluster
9. formed to determine the average comparison

RESULTS AND DISCUSSION

Analysis Result

Research Variables

The variables used in this study are as follows:

$_1$: Number of Pregnant Women (Soul).
$_2$: Number of Early Childhood (Life).
$_3$: Number of School Children (Soul).
$_4$: Number of people with disabilities (people).
$_5$: Number of Elderly (Soul).

Descriptive Statistics

The descriptive analysis used is the average, minimum value, and maximum value. The results of the descriptive analysis are presented in Table 3.1.

Table 3.1 Descriptive Statistics Results

Variables	Minimum	Average	Maximum
X_1	0	0,1404	2
X_2	0	10,18	37
X_3	33	150,7	645
X_4	0	8,105	28
X_5	17	83,4	352

Based on Table 3.1, it can be concluded as follows:
1. Most urban villages in Malang City had 0 pregnant women in October-November 2023. Kelurahan Mojolangu had the highest number of pregnant women in the October-November period with 2 pregnant women.
2. Most urban villages in Malang City had 10 children in the October - November period, while Kelurahan Bandunrejo had the highest number of children at 37.
3. Most urban villages in Malang City had 151 school children in the October - November 2023 period. Kelurahan Ketawanggede has the least number of school children, 33. Meanwhile, Kelurahan Tanjungrejo has the highest number of school children at 645.
4. Most urban villages in Malang City had 8 people with disabilities in the

October - November 2023 period. Kelurahan Tanjungrejo has the most people with disabilities with 28 people with disabilities.

5. Most urban villages in Malang City had 85 elderly people in the October - November 2023 period. Kelurahan Kesatrian has the least number of elderly people at 17, while Kelurahan Tanjungrejo has the most elderly people at 352.

Assumption Test

Before conducting cluster analysis, first conduct an assumption test to see whether the data to be analyzed has met the requirements for cluster analysis or not. There are two assumption tests used in cluster analysis.

1. The sample can represent the population.

   The calculation of the sample assumption test can represent the population (representativeness of the sample) is to find the Kaiser-Meyer-Olkin (KMO) value. How to find the Kaiser-Meyer- Olkin (KMO) value can be done using the help of RStudio software. The KMO value of 0.7905609 was obtained, which means that this KMO value is greater than 0.5, so it can be concluded that the assumption that the sample can represent the population (representativeness of the sample) has been fulfilled.

2. Non-Multicollinearity.

   In this study, the calculation of the multicollinearity assumption test is to look for the Variance Inflation Factori (VIF) value whose calculation is assisted by Rstudio software. The results are presented in Table 3.2.

Table 3. 2 Variance Inflation Factori (VIF) Results

Variables	VIF Value
X_1	1,185082
X_2	3,794437
X_3	8,847859
X_4	2,415281
X_5	6,836170

Based on Table 3.2. above shows that all values are less than 10 so it can be concluded that there is no multicollinearity between variables. Because there is no multicollinearity, cluster analysis can be done.

Cluster Validation

To determine the number of clusters, three cluster validity index rules were used, namely the dunn index, connecticity, and sillhouette. The results of determining the number of clusters using theavergae linkage method and the euclidean distance measure based on the cluster validity index are presented in Table 3.3.

Table 3. 3 Cluster Validation Results

Validity Index	Number of Clusters				
	3	4	5	6	7
Connectivity	6,7869	12,0833	14,0833	20,7647	27,0127
Dunn	0,4255	0,1480	0,1480	0,1719	0,1516
Silhouette	0,7045	0,5282	0,5007	0,4471	0,3919

In the connectivity test, the clusters formed can be said to be good when the connectivity index is low or close to zero. The dunn test will produce the most optimal cluster when it has a large value. As for the silhouette test, the clusters formed can be said to be good if the silhouette index value is close to 1. Based on Table 3.3 above, the optimal value is presented in Table 3.4. The following:

Table 3. 4 Cluster Validation Results

Validity Index	Value	Clusters Formed
Connectivity	6,7869	3
Dunn	0,4255	3
Silhouette	0,7045	3

Table 3.4 shows that the connectivity index, silhouette index, and Dunn index form two optimal clusters.

Cluster Analysis Results

After knowing the number of clusters formed, the next step is to form a dendrogram according to the cluster validity results in Table 3.4. The results of cluster analysis with many optimal clusters of 3 will form a dendrogram presented in Figure 3.3 below.

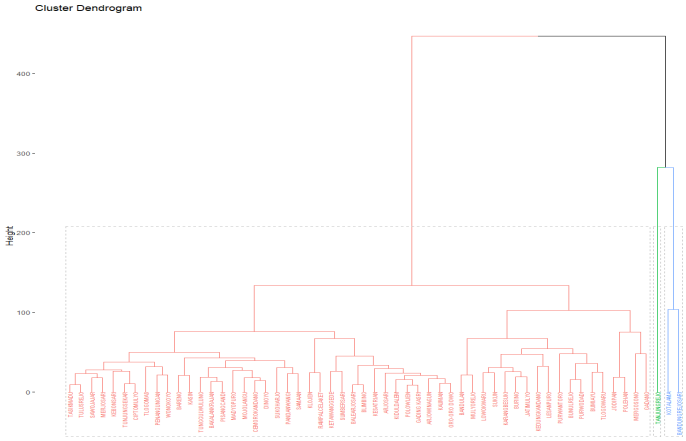

Figure 3. 1 Dendrogram of 3 Clusters

Buletin Ekonomika Pembangunan
https://journal.trunojoyo.ac.id/bep
Vol. 5 No.1 February 2024, hal.8-20
ISSN: 2807-4998 (online)

It can be seen that the cluster with the least members is cluster three with 1 urban village while the cluster with the most members is cluster one with 54 urban villages. In detail, the cluster members in each of the clusters formed are presented in Table 3.5.

Table 3. 5 Clustering Results of 3 Clusters

Cluster	**Village Name**	**Total**
1	Arjosari, Balearjosari, Blimbing, Bunulrejo, Jodipan, Kesatrian, Pandanwangi, Polehan, Polowijen, Purwantoro, Purwodadi, Bareng, Gading Kasri, Kasin, Kauman, Kiduldalem, Klojen, Oro-Oro Dowo, Penangungan, Rampalcelaket, Samaan, Sukoharjo, Arjowinagun, Bumiayu, Buring, Cemorokandang, Kedungkandang, Lesanpuro, Madyopuro, Mergosono, Sawojajar, Tlogowaru, Wonokoyo, Dinoyo, Jatimulyo, Ketawanggede, Lowokwaru, Merjosari, Mojolangu, Sumbersari, Tasikmadu, Tlogomas, Tulusrejo, Tunggulwulung, Tunjungsekar, Bakalankrajan, Bandulan, Ciptomulyo, Gadang, Karangbesuki, Kebonsari, Mulyorejo, Pisangcandi, Sukun	54
2	Kotalama, Bandungrejosari	2
3	Tanjungrejo	1
Total		57

Based on the results obtained, many groups are formed, namely as many as three clusters where each cluster has different characteristics. The method used to compare the two clusters that have been formed is to see the average value of each variable presented in Table 3.6.

Table 3. 6 Average Variables of Each Cluster

Variables	***Cluster***		
	1	2	3
Number of Pregnant Women $(X)_1$	0,141	0	0
Number of Early Childhood $(X)_2$	8,852	32,50	37
Number of School Children $(X)_3$	131,6	419	645
Number of Persons Disability $(X)_4$	7,407	17	28
Number of Elderly $(X)_5$	73,19	225	352

Based on Table 3.6, the characteristics of each cluster can be seen as follows.

a. Cluster One
   In cluster one with 54 kelurahan members, the average value of each variable is mostly below cluster two and cluster three. This means that the villages in cluster one have high social welfare.
b. Cluster Two
   Cluster two with 2 kelurahan members has an average value for each variable that is higher than cluster one but lower than cluster three. This means that the villages in cluster two have moderate social welfare characteristics.
c. Cluster Three
   Cluster Three with 1 urban village as a member has a higher average value for each variable than cluster one and cluster two, meaning that urban villages in cluster three have low social welfare characteristics.

Discussion

Clustering of villages based on the Family Hope Program in Malang city using hierarchical cluster analysis with the average linkage method and using the connectivity index, dunn index, and silhouette index, three optimal clusters are formed and the characteristic results of each group that have been formed are obtained from the average of each variable of the group formed. Table 3.5 shows the groups formed from the variables used. In this report, it can be seen that there are 54 villages that have a high level of welfare, 2 villages have a medium level of social welfare, and 1 village has a low level of social welfare. Table 3.6 shows that the average group of recipients of assistance for pregnant women, early childhood, school age, disability, and the elderly in each kelurahan in cluster 3 received more PKH assistance than cluster 2 and cluster 1 because they have a low level of social welfare, cluster 2 received more PKH assistance than cluster 1 and less than cluster 3 because they have a medium level of social welfare, while the average recipient of assistance for pregnant women, early childhood, school age, disability, and the elderly in each kelurahan in cluster 1 received less PKH assistance because they have a high level of social welfare.

CLOSING

Conclusion

Based on the results and discussion of the cluster analysis results using the average linkage method, it can be concluded that the optimum number of clusters in this report is three clusters. The clustering of villages based on the Family Hope Program (PKH) in the October - November period in 2023 is as follows.

1. Clusters with high levels of social welfare consist of 54 villages, namely Arjosari, Balearjosari, Blimbing, Bunulrejo, Jodipan, Kesatrian, Pandanwangi, Polehan, Polowijen, Purwantoro, Purwodadi, Bareng, Gading Kasri, Kasin, Kauman, Kiduldalem, Klojen, Oro-Oro Dowo, Penangungan, Rampalcelaket, Samaan, Sukoharjo, Arjowinagun, Bumiayu, Buring, Cemorokandang, Kedungkandang, Lesanpuro, Madyopuro, Mergosono, Sawojajar, Tlogowaru, Wonokoyo, Dinoyo, Jatimulyo, Ketawanggede, Lowokwaru, Merjosari, Mojolangu, Sumbersari, Tasikmadu, Tlogomas, Tulusrejo, Tunggulwulung, Tunjungsekar, Bakalankrajan, Bandulan, Ciptomulyo, Gadang, Karangbesuki,

ISSN: 2807-4998 (online)

Kebonsari, Mulyorejo, Pisangcandi, Sukun. So that they get less PKH assistance.
2. Clusters with moderate levels of social welfare consist of 2 villages, namely Kotalama, Bandungrejosari. So that they get more PKH assistance than cluster one and less than cluster three.
3. Clusters with low levels of social welfare consist of 1 Kelurahan, namely Tanjungrejo. Therefore, it receives more PKH assistance than clusters one and two.

Suggestions

Based on the results of the research analysis that has been carried out, the author provides suggestions that are expected to be useful. First, the Malang City Government is expected to pay special attention to urban villages with moderate and low levels of social welfare, so that assistance programs other than PKH can be provided to improve the level of community welfare, but still pay attention to urban villages with high levels of social welfare so that there is no decline. Second, the Malang City Social Service, especially the Social Rehabilitation and Security sector, is expected to be able to distribute PKH assistance and collect data on economic conditions in clusters with low and moderate levels of social welfare in a measured and equitable manner, while in clusters with high levels of social welfare, the distribution of PKH assistance is still carried out by paying attention to the development of the community's social welfare conditions every month. Third, for future research, it is recommended to use other assistance data such as BPNT and BST, because the assistance provided by the social service is not only PKH, and there is still a lack of research by students related to new assistance programs from the social service.

REFERENCES

Adji, A., Hidayat, T., Tuhiman, H., Kurniawati, S., & Maulana, A. (2020). Pengukuran Garis Kemiskinan di Indonesia: Tinjauan Teoritis dan Usulan Perbaikan. 1–36.

Atalay, R., (2015). Science Direct The education and the human capital to get rid of the middle-income trap and to provide the economic development. Procedia - Social and Behavioral Sciences, 174, pp.969–976. Available at: http://dx.doi.org/10.1016/j.sbspro.2015.01.7.20

Barkatin, Lailan, S., dan Hari, W. 2016. Analisis Perilaku Pelajar Terhadap Lingkungan Studi Kasus Pendidikan Menengah di Kabupaten Bogor. Vol. 6 No. 2.

Budiharjo, E. dan Hardjohubojo, S. 1993. Kota Berwawasan Lingkungan. Penerbit Alumni. Bandung.

Dubes dan Jain, A. K. 1988. Algorithm for Clustering Data. Pearson Education Inc. New Jersey.

Gujarati, D. N. 1995. Basic Econometrics. The McGrow Hill Companies Inc. New York.

Hair, J. F., Black, W. C., Babin, Barry J., dan Anderson, R. E. 2006. Multivariate Data Analysis. Pearson Education Inc. New Jersey.

Halim, N. N. dan Widodo, E. 2017. Clustering Dampak Gempa Bumi di Indonesia

Han, J. dan Kamber, M. 2006 Data Mining Concept and Techniques.

Morgan Kaufman Publishers. San Francisco.

Menggunakan Kohenen Self Organizing Maps Prosiding SI MANIS. Seminar Nasional Integrasi Matematika dan Nilai Islami. 1(1): 188-194.

Irwansyah, E. dan Faisal, M. 2015. Advanced Clustering Teori dan Aplikasi. Deepublish. Yogyakarta.

Safitri, D., Widiharih, T., Wilandari, Y., dan Saputra, H. S. 2012. Analisis Cluster pada Kabupaten/Kota di Jawa Tengah Berdasarkan Produksi Palawija. Media Statistika. 5(1): 11-16. Simamora, B. 2005. Analisis Multivariat Pemasaran. PT. Gramedia

Pustaka Umum. Jakarta.

Supranto, J. 2004. Analisis Multivariat Arti dan Interpretasi. Rineka Cipta. Jakarta.

Walpole, R. E. 1995. Pengantar Statistika. PT Gramedia Pustaka Utama. Jakarta.

Widarjono, A. 2010. Analisis Statistika Multivariat Terapan.UPP STIM YKPN. Yogyakarta.

Yitnosumarto, S. 1990. Dasar-Dasar Statistika. Rajawali. Jakarta.