Journal of Information System and Technology Research

journal homepage: https://journal.aira.or.id/index.php/jistr/

Implementation Of Data Mining Grouping Of Old Age Guarantee (Jht) Based On Region In Pandemic Period

Tata Mustika Dewi*, Rusmin Saragih*, Siswan Syahputra*

Progam Studi Sistem Informasi STMIK KAPUTAMA Binjai Jl. Veteran No. 4A-9A, Binjai 20714, North Sumatra

ARTICLE INFO

Article history:

Received Mar 31, 2022
Revised May 4, 2022
Accepted May 28, 2022
Available online May 31, 2022

Keywords:

Data Mining
Clustering
Old Age Security
Region

ABSTRACT

During the COVID-19 pandemic, many companies experienced a decline or went bankrupt, so they had to reduce the number of workers and even close the company. BPJS Ketenagakerjaan is a public legal entity that is responsible to the president and functions to administer four programs, namely Work Accident Insurance (JKK), Death Insurance (JKM), Old Age Security (JHT), with the addition of the Pension Guarantee program (JP). One of them is the submission of claims from too many participants of the Old Age Security program from various regions, especially the Langkat sub-district, so that it becomes a big problem to provide good service or information for the participants. For this reason, the author tries to create a system to support a computerized grouping process that can help automatically classify JHT claims by region, so there is an opportunity to design a grouping data mining system in it. Data mining is a process of mining data in very large amounts of data using statistical, mathematical methods, to utilize the latest artificial intelligence technology. Clustering is a method that is applied in creating a grouping data mining system to make it easier for employees to group JHT by region. Based on the analysis that has been done in the grouping of old-age insurance data using the clustering method, it is necessary to do the cluster process several times to get the same results according to the process that was first carried out, namely in cluster 1 : 2 3 2 cluster 2 : 2 8 2, cluster 3: 2 13 2 with 545 data in cluster 1, 308 data in cluster 2 and 421 data in cluster 3.

Corresponding Author:

Tata Mustika Dewi
Progam Studi Sistem Informasi STMIK KAPUTAMA Binjai Jl. Veteran No. 4A-9A, Binjai 20714, North Sumatra
Email: tatamustika09@gmail.com

1. INTRODUCTION

Manpower has an important role as a supporting element for the success of national development and in the context of implementing Indonesia's development. Workers who have a working relationship with the company where he works have the potential to increase productivity at the company. For this reason, it is only natural that they are given the right to protection, maintenance of welfare and development of survival, so that social security for workers is needed.

BPJS is a public legal entity that is responsible to the president and functions to administer four programs, namely Work Accident Insurance (JKK), Death Security (JKM), Old Age Security (JHT), with the addition of the Pension Guarantee program (JP). One of the government's efforts to guarantee the workforce, namely the program held which is currently the most beneficial for participants is the old age insurance (JHT)[1].

With this program, of course, the number of workers who disburse JHT funds is increasing. This is because during the COVID-19 pandemic, many companies experienced a decline or went bankrupt, so they had to reduce the number of workers and even close the company. Especially in the Langkat area, there are a lot of workers who claim JHT because of this impact.[1]

A claim is a claim from the insured party in connection with the existence of a guarantee with the insured party, where each party binds itself to guarantee payment by the insurer if payment has been made by the insured party, when there is a risk by the insured party. The JHT program is a cash benefit that is given when the participant no longer works or dies. JHT can be accepted when the participant stops working with a waiting period of 1 (one) month after stopping work, and no longer has to wait for membership of at least 10 (ten) years.

https://doi.org/xx.xxxx/xxxxxxx.xxxxxx

The problem that hinders the BPJS Employment Langkat Branch office is the submission of claims from too many Old Age Security program participants from various regions, especially the Langkat sub-district, so that it becomes a big problem to provide good service or information for participants and the large amount of Old Age Security data results in JHT data keeps adding more and more which are usually only stored in an archive or database. Without realizing from the pile of data - the data can still be retrieved an information that can help the BPJS Employment. Very efficient data grouping can be done by grouping based on the specified variable, namely region. One of the data grouping processes that can be used is the application of data mining.[2]–[6]

Data mining is a process of mining data in very large amounts of data using statistical, mathematical methods, and utilizing the latest artificial intelligence . Data mining in the process of grouping data can use a grouping method, namely the Clustering.method Clustering is the process of analyzing data objects into subsets called clusters. Objects in the cluster have similar characteristics to each other and are different from clusters . By utilizing the data mining process using the Clustering , it is hoped that the system to be built can solve the problem of grouping BPJS Employment participant data.[1], [2], [5]

This research is reinforced by previous studies related to the title, the first research conducted by (Achmad Fikri Sallaby and Eko Suryana) with the title "Application of Data Mining to Determine the Number of Registered Job Seekers Based on Age and Education Using K-Means Clustering (Case Study in Department of Manpower and Transmigration Bengkulu Province)" concluded that the results of processing with Data Mining facilitate grouping the number of registered job seekers based on age and education, making it easier for job seekers to obtain information.[1], [4], [7]

2. RESEARCH METHOD

Research methodology is a scientific process or method for obtaining data to be used for research. In conducting this research, the authors followed the methodological stages in this study, namely as follows:

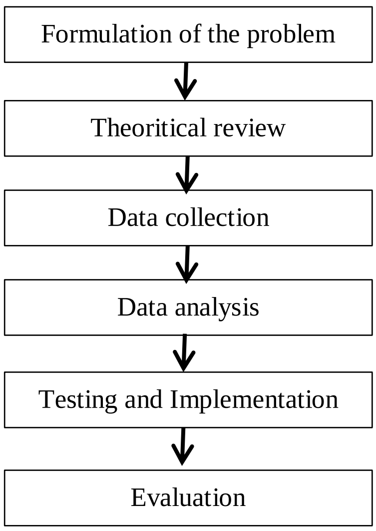


Figure 1 Research Workflow

Based on the picture above, it can be seen that there are several stages in completing the research, namely:

1. Problem Identification This stage is the initial stage in research, namely by determining the background of the problem, objectives and benefits so as not to get out of the problem and discussion or preparation of the thesis made.
2. Theoretical Study This stage is looking for information, sources related to the problems faced, both from literature studies, journals and the internet as a support and basic foundation for thesis writing.

3. Data collection
This stage is the collection of data needed in making a thesis such as interviews, observations which can then be processed to the next stage.

4. Data analysis

This stage is the stage of managing and analyzing the data that has been obtained so that the data can be grouped according to predetermined variables.

5. Testing and implementation This stage is the stage that performs the validation and implementation of the data that has been previously analyzed and the preparation of the program.

6. Evaluation

This stage is the stage of taking conclusions and suggestions that can be made in the preparation of the thesis. With the conclusion, the results of the entire thesis will be known and it is hoped that with suggestions there will be improvements and benefits for others.[3], [4], [8]

3. RESULTS AND DISCUSSION

The following will explain the results and discussion that has been obtained **:**

Calculation:
The number *of clusters* (K) is 3 *clusters*.centers *cluster* (Centroids) used, which are as follows:
Centroid1 is taken from the 1st data, namely alternative F: C1 = (3,7,1)
Centroid2 is taken from the 7th data, namely alternative H: C2 = (4, 6,3)
Centroid3 is taken from the 20th data, namely alternative J: C3 = (1,5,2)
equation. *Euclidean distance*
Calculation of the value of the closest distance between data using the equation
$\mathrm{D}(ij) = \sqrt{(X1i - X1j)^2 + (X2i - X2j)^2 + (X3i - X3j)^2}$,
process is as follows:

ITERATION PROCESS I

- A (1,17,1)

Centroid : C1 = (3,7,1); C2 = (4,6,3) dan C3 = (1,2,1)
C1 = $\sqrt{(1-3)^2 + (17-7)^2 + (1-1)^2}$ =10,198
C2 = $\sqrt{(1-4)^2 + (17-6)^2 + (1-3)^2} = 11{,}576$
C3 = $\sqrt{(1-1)^2 + (17-2)^2 + (1-1)^2} = 12{,}042$

- B (2,3,2)

Centroid : C1 = (3,7,1); C2 = (4,6,3) dan C3 = (1,2,1)
C1 = $\sqrt{(2-3)^2 + (3-7)^2 + (2-1)^2} = 4{,}243$
C2 = $\sqrt{(2-4)^2 + (3-6)^2 + (2-3)^2} = 3{,}742$
C3 = $\sqrt{(2-1)^2 + (3-2)^2 + (2-1)^2} = 2{,}236$

- C (1,6,1)

Centroid : C1 = (3,7,1); C2 = (4,6,3) dan C3 = (1,2,1)
C1 = $\sqrt{(1-3)^2 + (6-7)^2 + (1-1)^2} = 2{,}236$
C2 = $\sqrt{(1-4)^2 + (6-6)^2 + (1-3)^2} = 3{,}606$
C3 = $\sqrt{(1-1)^2 + (6-2)^2 + (1-1)^2} = 1{,}414$

- D (4,3,1)

Centroid : C1 = (3,7,1); C2 = (4,6,3) dan C3 = (1,2,1)
C1 = $\sqrt{(1-3)^2 + (2-7)^2 + (2-1)^2} = 5{,}477$
C2 = $\sqrt{(1-4)^2 + (2-6)^2 + (2-3)^2} = 5{,}099$
C3 = $\sqrt{(1-1)^2 + (2-2)^2 + (2-1)^2} = 3{,}000$

- E (2,1,1)

Centroid : C1 = (3,7,1); C2 = (4,6,3) dan C3 = (1,2,1)
C1 = $\sqrt{(2-3)^2 + (1-7)^2 + (1-1)^2} = 6{,}083$
C2 = $\sqrt{(2-4)^2 + (1-6)^2 + (1-3)^2} = 5{,}745$
C3 = $\sqrt{(2-1)^2 + (1-2)^2 + (1-1)^2} = 2{,}243$

- F (3,7,1)

Centroid : C1 = (3,7,1); C2 = (4,6,3) dan C3 = (1,2,1)
C1 = $\sqrt{(3-3)^2 + (7-7)^2 + (1-1)^2} = 0{,}000$
C2 = $\sqrt{(3-4)^2 + (7-6)^2 + (1-3)^2} = 2{,}449$
C3 = $\sqrt{(3-1)^2 + (7-2)^2 + (1-1)^2} = 3{,}000$

- G (4,6,3)

Centroid : C1 = (3,7,1); C2 = (4,6,3) dan C3 = (1,2,1)
C1 = $\sqrt{(4-3)^2 + (6-7)^2 + (3-1)^2} = 2{,}449$
C2 = $\sqrt{(4-4)^2 + (6-6)^2 + (3-3)^2} = 0{,}000$

C3 = $\sqrt{(4-1)^2+(6-2)^2+(3-1)^2} = 3{,}317$

- H (4,6,3)

Centroid : C1 = (3,7,1); C2 = (4,6,3) dan C3 = (1,2,1)

C1 = $\sqrt{(4-3)^2+(6-7)^2+(3-1)^2} = 2{,}449$

C2 = $\sqrt{(4-4)^2+(6-6)^2+(3-3)^2} = 0{,}000$

C3 = $\sqrt{(4-1)^2+(6-2)^2+(3-1)^2} = 3{,}317$

- I (4,6,3)

Centroid : C1 = (3,7,1); C2 = (4,6,3) dan C3 = (1,2,1)

C1 = $\sqrt{(4-3)^2+(6-7)^2+(3-1)^2} = 2{,}449$

C2 = $\sqrt{(4-4)^2+(6-6)^2+(3-3)^2} = 0{,}000$

C3 = $\sqrt{(4-1)^2+(6-2)^2+(3-1)^2} = 3{,}317$

- J (1,5,2)

Centroid : C1 = (3,7,1); C2 = (4,6,3) dan C3 = (1,2,1)

C1 = $\sqrt{(3-1)^2+(7-2)^2+(1-1)^2} = 3{,}000$

C2 = $\sqrt{(1-4)^2+(5-6)^2+(2-3)^2} = 3{,}317$

C3 = $\sqrt{(1-1)^2+(5-2)^2+(1-1)^2} = 0{,}000$

- K (1,5,2)

Centroid : C1 = (3,7,1); C2 = (4,6,3) dan C3 = (1,2,1)

C1 = $\sqrt{(3-1)^2+(7-2)^2+(1-1)^2} = 3{,}000$

C2 = $\sqrt{(1-4)^2+(5-6)^2+(2-3)^2} = 3{,}317$

C3 = $\sqrt{(1-1)^2+(5-2)^2+(1-1)^2} = 0{,}000$

- L (1,2,1)

Centroid : C1 = (3,7,1); C2 = (4,6,3) dan C3 = (1,2,1)

C1 = $\sqrt{(1-3)^2+(2-7)^2+(1-1)^2} = 5{,}385$

C2 = $\sqrt{(1-4)^2+(2-6)^2+(1-3)^2} = 5{,}385$

C3 = $\sqrt{(1-1)^2+(2-2)^2+(1-1)^2} = 3{,}162$

- M (4,5,3)

Centroid : C1 = (3,7,1); C2 = (4,6,3) dan C3 = (1,2,1)

C1 = $\sqrt{(4-3)^2+(5-7)^2+(3-1)^2} = 3{,}000$

C2 = $\sqrt{(4-4)^2+(5-6)^2+(3-3)^2} = 1{,}000$

C3 = $\sqrt{(4-1)^2+(5-2)^2+(3-1)^2} = 3{,}162$

- N (2,4,1)

Centroid : C1 = (3,7,1); C2 = (4,6,3) dan C3 = (1,2,1)

C1 = $\sqrt{(2-3)^2+(4-7)^2+(1-1)^2} = 3{,}162$

C2 = $\sqrt{(2-4)^2+(4-6)^2+(1-3)^2} = 3{,}464$

C3 = $\sqrt{(2-1)^2+(4-2)^2+(1-1)^2} = 1{,}732$

- O (1,8,2)

Centroid : C1 = (3,7,1); C2 = (4,6,3) dan C3 = (1,2,1)

C1 = $\sqrt{(1-3)^2+(8-7)^2+(2-1)^2} = 2{,}449$

C2 = $\sqrt{(1-4)^2+(6-6)^2+(2-3)^2} = 3{,}742$

C3 = $\sqrt{(1-1)^2+(8-2)^2+(2-1)^2} = 3{,}000$

- P (1,2,1)

Centroid : C1 = (3,7,1); C2 = (4,6,3) dan C3 = (1,2,1)

C1 = $\sqrt{(1-3)^2+(2-7)^2+(1-1)^2} = 5{,}385$

C2 = $\sqrt{(1-4)^2+(2-6)^2+(1-3)^2} = 5{,}385$

C3 = $\sqrt{(1-1)^2+(2-2)^2+(1-1)^2} = 3{,}162$

- Q (1,3,2)

Centroid : C1 = (3,7,1); C2 = (4,6,3) dan C3 = (1,2,1)

C1 = $\sqrt{(1-3)^2+(3-7)^2+(2-1)^2} = 4{,}583$

C2 = $\sqrt{(1-4)^2+(3-6)^2+(2-3)^2} = 4{,}359$

C3 = $\sqrt{(1-1)^2+(3-2)^2+(2-1)^2} = 2{,}000$

- R (4,15,3)

Centroid : C1 = (3,7,1); C2 = (4,6,3) dan C3 = (1,2,1)

C1 = $\sqrt{(4-3)^2+(15-7)^2+(3-1)^2} = 8{,}307$

C2 = $\sqrt{(4-4)^2+(15-6)^2+(3-3)^2} = 9{,}000$

C3 = $\sqrt{(4-1)^2+(15-2)^2+(3-1)^2} = 10{,}488$

- S (1,7,2)

Centroid : C1 = (3,7,1); C2 = (4,6,3) dan C3 = (1,2,1)

 P ISSN 2828-3864; E ISSN: 2828-2973

C1 = $\sqrt{(1-3)^2+(7-7)^2+(2-1)^2} = 2{,}236$
C2 = $\sqrt{(1-4)^2+(7-6)^2+(2-3)^2} = 3{,}317$
C3 = $\sqrt{(1-1)^2+(7-2)^2+(2-1)^2} = 2{,}000$
- T (4,9,2)
Centroid : C1 = (3,7,1); C2 = (4,6,3) dan C3 = (1,2,1)
C1 = $\sqrt{(4-3)^2+(9-7)^2+(2-1)^2} = 2{,}449$
C2 = $\sqrt{(4-4)^2+(9-6)^2+(2-3)^2} = 3{,}162$
C3 = $\sqrt{(4-1)^2+(9-2)^2+(2-1)^2} = 5{,}000$

The results of the calculation of the value of the *Euclidean distance* in iteration I above can be seen in the following table:

Table 1. Value *of Euclidean Distance* in Iteration I

No	Alternatif	C1	C2	C3	Grup
1	A	10,198	11,576	12,042	1
2	B	4,243	3,742	2,236	3
3	C	2,236	3,606	1,414	3
4	D	5,477	5,099	3,000	3
5	E	6,083	5,745	4,243	3
6	F	0,000	2,449	3,000	1
7	G	2,449	0,000	3,317	2
8	H	2,449	0,000	3,317	2
9	I	2,449	0,000	3,317	2
10	J	3,000	3,317	0,000	3
11	K	3,000	3,317	0,000	3
12	L	5,385	5,385	3,162	3
13	M	3,000	1,000	3,162	2
14	N	3,162	3,464	1,732	3
15	O	2,449	3,742	3,000	1
16	P	5,385	5,385	3,162	3
17	Q	4,583	4,359	2,000	3
18	R	8,307	9,000	10,488	1
19	S	2,236	3,317	2,000	3
20	T	2,449	3,162	5,000	1

Table the results of the iteration process I grouping above to determine the grouping of data can be determined with the following conditions:
Group 1 is obtained by comparing the smallest value of 3 centroids, if the smallest value is in C1 (Centroid1), then the data is entered in group 1.
Group 2 obtained by comparing the smallest value of 3 centroids, if the smallest value of 5
is in C2 (Centroid2), then the data is entered in the .
Group 3 is obtained by comparing the smallest value of 3 centroids, if the smallest value is in C3 (Centroid3), then the data is entered in group 3.
Next, do the calculation for iteration II, before calculating iteration II, it needs to be made first for the center of the Centroid of the three Clusters.
Here are the new 3 Centroids :

- C1 (5 data) :

X = $\frac{(1+3+1+4+4)}{5} = \frac{13}{5}$ = 2,600
Y = $\frac{(17+7+8+15+9)}{5} = \frac{56}{5}$ = 11,200
Z = $\frac{(1+1+2+3+2)}{5} = \frac{9}{5}$ = 1,800
Result Centroid1 : C1(2,600; 11,200; 1,800)

- C2 (4 data) :

X = $\frac{(4+4+4+4)}{5} = \frac{16}{4}$ = 4,000
Y = $\frac{(6+6+6+5)}{4} = \frac{23}{4}$ = 5,750
Z = $\frac{(3+3+3+3)}{4} = \frac{12}{4}$ = 3,000
Result Centroid2 : C2(4,000; 5,750; 3,000)

- C3 (3 data):
X = $\frac{(2+1+1+2+1+1+1+2+1+1+1)}{11} = \frac{14}{11}$ =1,273
Y = $\frac{(3+6+2+1+5+5+2+4+2+3+7)}{11} = \frac{40}{11}$ =3,636
Z = $\frac{(2+1+2+1+2+2+1+1+1+2+2)}{11} = \frac{17}{11}$ =1,545

Result Centroid3 : C3(1,273; 2,636; 1,545)

ITERATION PROCESS II
- A (1,17,1)
Centroid : C1(2,600; 11,200; 1,800); C2(4,000; 5,750; 3,000); C3(1,273; 3,636; 1,545)
C1 = $\sqrt{(1-2{,}600)^2+(17-11{,}200)^2+(1-1{,}800)^2}$ = 10,229
C2 = $\sqrt{(1-4{,}000)^2+(17-5{,}750)^2+(1-3{,}000)^2}$ = 11,576
C3 = $\sqrt{(1-1{,}273)^2+(17-3{,}636)^2+(1-1{,}545)^2}$ = 12,042
- B (2,3,2)
Centroid : C1(2,600; 11,200; 1,800); C2(4,000; 5,750; 3,000); C3(1,273; 3,636; 1,545)
C1 = $\sqrt{(2-2{,}600)^2+(3-11{,}200)^2+(2-1{,}800)^2}$ = 4,128
C2 = $\sqrt{(2-4{,}000)^2+(3-5{,}750)^2+(2-3{,}000)^2}$ = 3,742
C3 = $\sqrt{(2-1{,}273)^2+(3-3{,}636)^2+(2-1{,}545)^2}$ = 2,236
- C (1,6,1)
Centroid : C1(2,600; 11,200; 1,800); C2(4,000; 5,750; 3,000); C3(1,273; 3,636; 1,545)
C1 = $\sqrt{(1-2{,}600)^2+(6-11{,}200)^2+(1-1{,}800)^2}$ = 2,375
C2 = $\sqrt{(1-4{,}000)^2+(6-5{,}750)^2+(1-3{,}000)^2}$ = 3,606
C3 = $\sqrt{(1-1{,}273)^2+(6-3{,}636)^2+(1-1{,}545)^2}$ = 1,414
- D (1,2,2)
Centroid : C1(2,600; 11,200; 1,800); C2(4,000; 5,750; 3,000); C3(1,273; 3,636; 1,545)
C1 = $\sqrt{(1-2{,}600)^2+(2-11{,}200)^2+(2-1{,}800)^2}$ = 5,389
C2 = $\sqrt{(1-4{,}000)^2+(2-5{,}750)^2+(2-3{,}000)^2}$ = 5,099
C3 = $\sqrt{(1-1{,}273)^2+(2-3{,}636)^2+(2-1{,}545)^2}$ = 3,000
- E (2,1,1)
Centroid : C1(2,600; 11,200; 1,800); C2(4,000; 5,750; 3,000); C3(1,273; 3,636; 1,545)
C1 = $\sqrt{(2-2{,}600)^2+(1-11{,}200)^2+(1-1{,}800)^2}$ = 6,135
C2 = $\sqrt{(2-4{,}000)^2+(1-5{,}750)^2+(1-3{,}000)^2}$ = 5,745
C3 = $\sqrt{(2-1{,}273)^2+(1-3{,}636)^2+(1-5{,}545)^2}$ = 4,243
- F (3,7,1)
Centroid : C1(2,600; 11,200; 1,800); C2(4,000; 5,750; 3,000); C3(1,273; 3,636; 1,545)
C1 = $\sqrt{(3-2{,}600)^2+(7-11{,}200)^2+(1-1{,}800)^2}$ = 0,800
C2 = $\sqrt{(3-4{,}000)^2+(7-3{,}223)^2+(1-3{,}000)^2}$ = 2,449
C3 = $\sqrt{(3-1{,}273)^2+(7-3{,}636)^2+(1-1{,}545)^2}$ = 3,000
- G (4,6,3)
Centroid : C1(2,600; 11,200; 1,800); C2(4,000; 5,750; 3,000); C3(1,273; 3,636; 1,545)
C1 = $\sqrt{(4-2{,}600)^2+(6-11{,}200)^2+(3-1{,}800)^2}$ = 1,855
C2 = $\sqrt{(4-4{,}000)^2+(6-5{,}750)^2+(3-3{,}000)^2}$ = 0,000
C3 = $\sqrt{(4-1{,}273)^2+(6-\ 3{,}636)^2+(3-1{,}545)^2}$ = 3,217
- H (4,6,3)
Centroid : C1(2,600; 11,200; 1,800); C2(4,000; 5,750; 3,000); C3(1,273; 3,636; 1,545)
C1 = $\sqrt{(4-2{,}600)^2+(6-11{,}200)^2+(3-1{,}800)^2}$ = 1,855
C2 = $\sqrt{(4-4{,}000)^2+(6-5{,}750)^2+(3-3{,}000)^2}$ = 0,000
C3 = $\sqrt{(4-1{,}273)^2+(6-\ 3{,}636)^2+(3-1{,}545)^2}$ = 3,217
- I (4,6,3)
Centroid : C1(2,600; 11,200; 1,800); C2(4,000; 5,750; 3,000); C3(1,273; 3,636; 1,545)
C1 = $\sqrt{(4-2{,}600)^2+(6-11{,}200)^2+(3-1{,}800)^2}$ = 1,855
C2 = $\sqrt{(4-4{,}000)^2+(6-5{,}750)^2+(3-3{,}000)^2}$ = 0,000

C3 = $\sqrt{(4-4{,}000)^2+(6-5{,}750)^2+(3-3{,}000)^2}=$ 3,217

- J (1,5,2)

Centroid : C1(2,600; 11,200; 1,800); C2(4,000; 5,750; 3,000); C3(1,273; 3,636; 1,545)

C1 = $\sqrt{(1-2{,}600)^2+(5-11{,}200)^2+(2-1{,}800)^2}$ = 2,835

C2 = $\sqrt{(1-4{,}000)^2+(5-5{,}750)^2+(2-3{,}000)^2}$ = 3,317

C3 = $\sqrt{(1-1{,}273)^2+(5-3{,}636)^2+(2-1{,}545)^2}$ = 0,000

- K (1,5,2)

Centroid C1(2,600; 11,200; 1,800); C2(4,000; 5,750; 3,000); C3(1,273; 3,636; 1,545)

C1 = $\sqrt{(1-2{,}600)^2+(5-11{,}200)^2+(2-1{,}800)^2}$ = 2,835

C2 = $\sqrt{(1-4{,}000)^2+(5-5{,}750)^2+(2-3{,}000)^2}$ = 3,317

C3 = $\sqrt{(1-1{,}273)^2+(5-3{,}636)^2+(2-3{,}636)^2}$ = 0,000

- L (2,1,2)

Centroid : C1(2,600; 11,200; 1,800); C2(4,000; 5,750; 3,000); C3(1,273; 3,636; 1,545)

C1 = $\sqrt{(2-2{,}6005)^2+(1-11{,}200)^2+(2-1{,}800)^2}$ = 5,444

C2 = $\sqrt{(2-4{,}000)^2+(1-5{,}750)^2+(2-3{,}000)^2}$ = 5,385

C3 = $\sqrt{(2-1{,}273)^2+(1-3{,}636)^2+(2-1{,}545)^2}$ = 3,162

- M (4,5,3)

Centroid : C1(2,600; 11,200; 1,800); C2(4,000; 5,750; 3,000); C3(1,273; 3,636; 1,545)

C1 = $\sqrt{(4-2{,}600)^2+(5-11{,}200)^2+(3-1{,}800)^2}$ = 2,538

C2 = $\sqrt{(4-4{,}000)^2+(5-5{,}750)^2+(3-3{,}000)^2}$ = 1,000

C3 = $\sqrt{(4-1{,}273)^2+(5-3{,}636)^2+(3-1{,}545)^2}$ = 3,162

- N (2,4,1)

Centroid : C1(2,600; 11,200; 1,800); C2(4,000; 5,750; 3,000); C3(1,273; 3,636; 1,545)

C1 = $\sqrt{(2-(2{,}600)^2+(4-11{,}200)^2+(1-1{,}800)^2}$ = 3,262

C2 = $\sqrt{(2-4{,}000)^2+(4-5{,}750)^2+(1-3{,}000)^2}$ = 3,464

C3 = $\sqrt{(2-1{,}273)^2+(4-3{,}636)^2+(1-1{,}545)^2}$ = 1,732

- O (1,8,2)

Centroid : C1(2,600; 11,200; 1,800); C2(4,000; 5,750; 3,000); C3(1,273; 3,636; 1,545)

C1 = $\sqrt{(1-(2{,}600)^2+(8-11{,}200)^2+(2-1{,}800)^2}$ = 2,245

C2 = $\sqrt{(1-4{,}000)^2+(8-5{,}750)^2+(2-3{,}000)^2}$ = 3,742

C3 = $\sqrt{(1-1{,}273)^2+(8-3{,}636)^2+(2-1{,}545)^2}$ = 3,000

- P (1,2,1)

Centroid : C1(2,600; 11,200; 1,800); C2(4,000; 5,750; 3,000); C3(1,273; 3,636; 1,545)

C1 = $\sqrt{(1-(2{,}600)^2+(2-11{,}200)^2+(1-1{,}800)^2}$ = 5,444

C2 = $\sqrt{(1-4{,}000)^2+(2-5{,}750)^2+(1-3{,}000)^2}$ = 5,385

C3 = $\sqrt{(1-1{,}273)^2+(2-3{,}636)^2+(1-1{,}545)^2}$ = 3,162

- Q (1,3,2)

Centroid : C1(2,600; 11,200; 1,800); C2(4,000; 5,750; 3,000); C3(1,273; 3,636; 1,545)

C1 = $\sqrt{(1-2{,}600)^2+(3-11{,}200)^2+(2-1{,}800)^2}$ = 4,477

C2 = $\sqrt{(1-4{,}000)^2+(3-5{,}750)^2+(2-3{,}000)^2}$ = 4,359

C3 = $\sqrt{(1-1{,}273)^2+(3-3{,}636)^2+(2-1{,}545)^2}$ = 2,000

- R (4,15,3)

Centroid : C1(2,600; 11,200; 1,800); C2(4,000; 5,750; 3,000); C3(1,273; 3,636; 1,545)

C1 = $\sqrt{(4-2{,}600)^2+(15-11{,}200)^2+(3-1{,}800)^2}$ = 8,151

C2 = $\sqrt{(4-4{,}000)^2+(15-5{,}750)^2+(3-3{,}000)^2}$ = 9,000

C3 = $\sqrt{(4-1{,}273)^2+(15-3{,}636)^2+(3-1{,}545)^2}$ = 10,488

- S (1,7,2)

Centroid : C1(2,600; 11,200; 1,800); C2(4,000; 5,750; 3,000); C3(1,273; 3,636; 1,545)

C1 = $\sqrt{(1-2{,}600)^2+(7-11{,}200)^2+(2-1{,}800)^2}$ = 2,010

C2 = $\sqrt{(1-4{,}000)^2+(7-5{,}750)^2+(2-3{,}000)^2}$ = 3,317

C3 = $\sqrt{(1-1{,}273)^2+(7-3{,}636)^2+(2-1{,}545)^2}$ = 2,000

- T (4,9,2)

Centroid : C1(2,600; 11,200; 1,800); C2(4,000; 5,750; 3,000); C3(1,273; 3,636; 1,545)

C1 = $\sqrt{(4-2{,}600)^2+(9-11{,}200)^2+(2-1{,}800)^2}$ = 2,245

C2 = $\sqrt{(4-4{,}000)^2+(9-5{,}750)^2+(2-3{,}000)^2}$ = 3,162

C3 = $\sqrt{(4-1{,}273)^2+(9-3{,}636)^2+(2-1{,}545)^2}$ = 5,000

The results of the calculation of the value of the *euclidean distance* in the second iteration above can be seen in the following table:

Table 2. **The value *Euclidean distance* in iteration II**

No	Alternatif	C1	C2	C3	Grup
1	A	10,229	11,576	12,042	1
2	B	4,128	3,742	2,236	3
3	C	2,375	3,606	1,414	3
4	D	5,389	5,099	3,000	3
5	E	6,135	5,745	4,243	3
6	F	0,800	2,449	3,000	1
7	G	1,855	0,000	3,317	2
8	H	1,855	0,000	3,317	2
9	I	1,855	0,000	3,317	2
10	J	2,835	3,317	0,000	3
11	K	2,835	3,317	0,000	3
12	L	5,444	5,385	3,162	3
13	M	2,538	1,000	3,162	2
14	N	3,262	3,464	1,732	3
15	O	2,245	3,742	3,000	1
16	P	5,444	5,385	3,162	3
17	Q	4,477	4,359	2,000	3
18	R	8,151	9,000	10,488	1
19	S	2,010	3,317	2,000	3
20	T	2,245	3,162	5,000	1

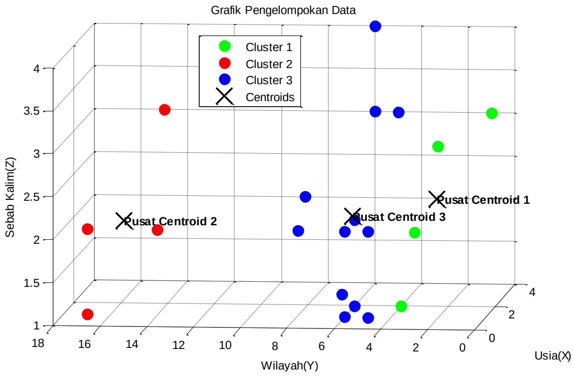
Furthermore, to ensure the similarity between the initial group and the new group, namely the data group iteration I with data in iteration II, this process is to check whether there is a change in the cluster group or not, the following is a comparison of similarities between the initial group and the new group:

Table 3. Comparison of Groups Iteration I and Iteration II

No	Alternative	Groups in Iteration I	Groups in Iteration II
1	A	1	1
2	B	3	3
3	C	3	3
4	D	3	3
5	E	3	3
6	F	1	1
7	G	2	2
8	H	2	2
9	I	2	2
10	J	3	3
11	K	3	3
12	L	3	3
13	M	2	2
14	N	3	3
15	O	1	1
16	P	3	3
17	Q	3	3
18	R	1	1

No	Alternative	Groups in Iteration I	Groups in Iteration II
19	S	3	3
20	T	1	1

The comparison results from the initial group (iteration I) and the new group (iteration II) are the same, the iteration calculation is complete.
Creating a Clustering Graph:
Clustering graph is a graph that displays the coordinates of a data point after the data grouping process. By using the MATLAB programming application, the following is the result of the Clustering graph of data grouping[9]–[12]



lab

Figure 2. Graph of BPJS Employment JHT Data Clustering

From 20 participant data which was processed into alternative data grouping using the Clustering in this study, 3 groups were obtained, group 1 (Cluster 1) contained 4 data, group 2 (Cluster 2) contained 4 data and group 12 (Cluster 3) there are 3 data.

1. Cluster 1 (2,600;11,200;1,800):

based on the above calculation, it can be seen that in cluster 1, the Amount of Old Age Security by Region During the Pandemic Period is 41-50 years old in the Binjai area and the reason for the layoff claim.

2. Cluster 2 (4,000; 5,750; 3,000):

based on the above calculation, it can be seen that in cluster 2 the Amount of Old Age Security Based on Regions In Non-Pandemic Periods, age more than 50 years, Completed areas and reasons for Pension claims.

3. Cluster 3 (1,273;3,636;1,545):

Based on the above calculation, it can be seen that in cluster 3, the Amount of Old Age Security Based on Regions In Non-Pandemic Periods aged 20-30
years in Kuala area due to claims to resign.

3.2 Discussion

The system built is a JHT data grouping system which is grouped by region during the pandemic. Old Age Security is intended as a substitute for the loss of income for workers due to death, disability, or old age and is implemented with an old age savings system. The Old Age Security Program provides certainty of income receipts that are paid when the workforce reaches the age of 56 years or has met certain requirements. Old Age Security will be returned/paid in the amount of the collected contributions plus the results of the development (at least equal to the government bank counter rate deposit interest rate). After implementing the method in the system design of the data grouping system, to find out the results of the implementation of the system, it is necessary to test the system that has been designed. The test will be carried out by processing the system input data, the data grouping process to be carried out is as follows:

1. Clustering Data

The JHT data grouping stage can be done on the "PROCESS" menu. The following is the process of grouping JHT data with a system that has been designed with the application of the Clustering using the MATLAB R2014a programming application:

a. Input JHT

data Participant data can be input into the system by using the “Input” button on the clustering in the system. After the data is input into the system, the data will be displayed in the input data table and the system will display the amount of data that has been inputted. The following is a page display after the data has been successfully input into the system:

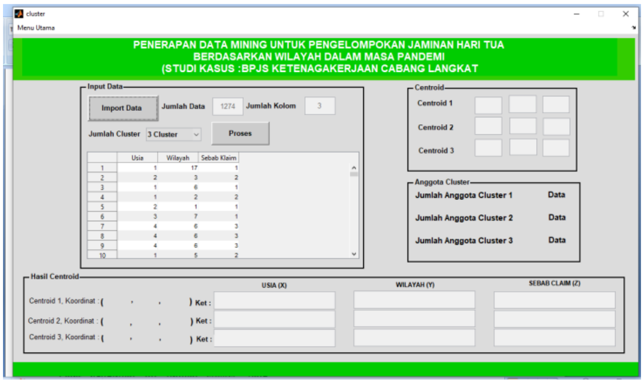


Figure 3. Display After Data Input In The System

b. Determine the data cluster on the system In this test the number of clusters used is 3 clusters.

c. Data clustering process

At this stage the user will carry out the process of grouping the data that has been inputted into the system. The process of grouping data on the system can be done by pressing the "Process" button on the data cluster process page. The process will utilize the data that has been inputted and the number of clusters that have been determined. After the data grouping process is complete, then on this display the system will display a description of the data cluster center in the data processing result table according to the number of clusters used. The following is the display after the data grouping process on the system:

Figure 4. Display After the Data Grouping Test Process

It can be seen in the picture above that in the first test process of grouping data, the data cluster center is generated as follows:

- Centroid 1: (1.9266 2.8807 1.9431)
- Centroid 2 : (1.7987 7.6948 1.9286)
- Centroid 3: (1.8337 12.8076 1.9074)

This display is a display of the results of the data grouping process that has been carried out on the data grouping system with the Clustering method that has been processed, this page is on the "RESULT CLUSTERING DATA" menu on the system, in this view the system will also display a description of the results of the data cluster that is the center data grouping clusters.

d. Results of the Data Grouping Process

The resultant data grouping which provides information in the form of the coordinates of the data grouping points and the data cluster center, the display of the results as the grouping of the data is as follows:

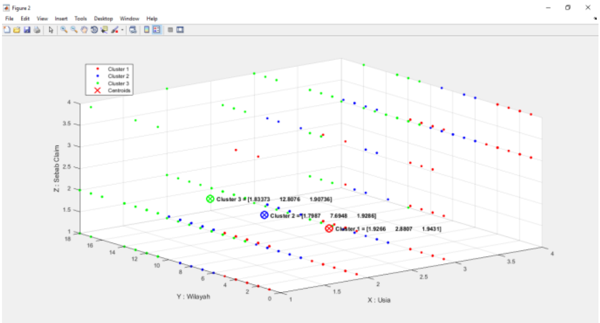


Figure 5. Graph of the Data Grouping Process

Figure 6. Results of the Data Grouping Process

4. CONCLUSION

The author concludes that the grouping of Old Age Security by region during the pandemic is the application of data mining for grouping old age insurance by region during the pandemic using the clustering method that can be applied and generates information using the Matlab GUI. From the 1247 data, there are 3 clusters which are grouped. Based on the calculations, the results obtained are the number of members of group 1 is 545 data, the number of members of group 2 is 308 data and the number of members of group 3 is 421 data. With the results centroid 1: 1.9266 2.8807 1.9431, centroid 2: 1.7987 7.6948 1.9286 and centroid 3: 1.8337 12.8076 1.9074. Based on the cluster process that has been carried out, the grouping of old-age insurance data on C1, C2 and C3 is with the age of 31 - 40 T with wampu, finished, and opposite palm oil areas and the cause of the claim is due to layoffs.

REFERENCES

[1] J. Informasi, J. Wandana, and S. Defit, "Klasterisasi Data Rekam Medis Pasien Pengguna Layanan BPJS Kesehatan Menggunakan Metode K-Means," vol. 2, pp. 4–9, 2020, doi: 10.37034/jidt.v2i4.73.

[2] S. Rahayu, D. T. Nugrahadi, and F. Indriani, "CLUSTERING PENENTUAN POTENSI KEJAHATAN DAERAH DI KOTA BANJARBARU DENGAN METODE K-MEANS," vol. 01, no. 01, pp. 33–45, 2014.

[3] A. Ikhwan, "Analisis Cluster Jarak Antar Kecamatan Dengan Menggunakan Metode Euclidean Di Untuk Penetapan Zona Pengoperasian Mobil Mplik Di Kota Medan," no. August, 2016.

[4] A. Ikhwan and N. Aslami, "Implementasi Data Mining untuk Manajemen Bantuan Sosial Menggunakan Algoritma K-Means," *JurTI (Jurnal Teknol. Informasi)*, vol. 4, no. 2, 2020, [Online]. Available: http://jurnal.una.ac.id/index.php/jurti/article/view/2103.

[5] R. Setiawan and N. Tes, "PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING UNTUK MENENTUKAN STRATEGI PROMOSI MAHASISWA BARU (Studi Kasus : Politeknik LP3I Jakarta)," vol. 3, no. 1, pp. 76–92, 2016.

[6] A. Manaor and H. Pardede, "ANALISIS PENGELOMPOKKAN PERFORMANCE DOSEN DENGAN METODE," no. January 2012, 2017.

[7] G. Gustientiedina, M. H. Adiya, and Y. Desnelita, "Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan," *J. Nas. Teknol. dan Sist. Inf.*, vol. 5, no. 1, pp. 17–24, 2019, doi: 10.25077/teknosi.v5i1.2019.17-24.

[8] A. Almira, S. Suendri, and A. Ikhwan, "Implementasi Data Mining Menggunakan Algoritma Fp-Growth pada Analisis Pola Pencurian Daya Listrik," *J. Inform. ...*, vol. 6, no. 2, pp. 442–448, 2021, [Online]. Available: http://www.openjournal.unpam.ac.id/index.php/informatika/article/view/12278.

[9] B. Kuniadi, D. Puspitaningrum, and F. Farady Coastera, "Perancangan Dan Pembuatan Aplikasi Steganografi Pesan Teks

Pada Audio Digital Dengan Metode Least Significant Bit," vol. 5, no. 3, pp. 285–297, 2017.

[10] A. Ikhwan, "Penerapan Fuzzy Mamdani Untuk Sistem Pendukung Keputusan Pemilihan Laptop," *J. Fasilkom*, vol. 9, no. 2, pp. 476–483, 2019, doi: 10.37859/jf.v9i2.1407.

[11] E. Prasetyo, *Data Mining Konsep dan Aplikasi Menggunakan Matlab*. Yogyakarta: Andi, 2012.

[12] M. Firdaoussi, H. Ghennioui, M. El Kamili, and M. Lamrini, "Performance evaluation of new blind OFDM signal recognition based on properties of the second-order statistics using universal software radio peripheral platform," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 23, no. 2, pp. 1227–1236, 2021, doi: 10.11591/ijeecs.v23.i2.pp1227-1236.