Universiti Teknologi MARA
Cawangan Melaka

In partnership with
Tadulako University

i-JaMCSIIX
2023

EXTENDED ABSTRACT BOOK

Publication Date: 30 March 2024
ISBN: 978-967-15337-0-3

https://jamcsiix.uitm.edu.my

i-JaMCSIIX 2023

INTERNATIONAL JASIN MULTIMEDIA & COMPUTER SCIENCE INVENTION AND INNOVATION EXHIBITION (I-JaMCSIIX) 2023

EXTENDED ABSTRACT

ISBN: 978-967-15337-0-3
i-JaMCSIIX
Universiti Teknologi MARA Cawangan Melaka Kampus Jasin 77300, Merlimau, Melaka

Web: https://jamcsiix.uitm.edu.my

In partnership with
Tadulako University

ORGANIZING COMMITTEE

PATRON	PM DR ISMADI MD BADARUDIN
ADVISOR I	TS DR JAMALUDDIN HJ JASMIS
ADVISOR II	DATO' DR MOHD HAJAR HASROL JONO
PROGRAM DIRECTOR	**DR. NUR SUHAILAYANI SUHAIMI**
DEPUTY DIRECTOR	TS DR NURUL HIDAYAH BINTI MAT ZAIN
SECRETARY I	**ANIS SHOBIRIN ABDULLAH SANI**
SECRETARY II	FAIQAH HAFIDZAH HALIM
TREASURER I	**SITI AISYAH ABD KADIR**
TREASURER II	UMMU MARDHIAH JALIL
PUBLICATION	**NURBAITY BINTI SABRI** DR. SITI FEIRUSZ AHMAD FESOL DR. AHMAD FIRDAUS BIN AHMAD FADZIL SITI NURAMALINA BINTI JOHARI ROSNIZA ROSLAN
JURY	**Ts DR. ALYA GEOGIANA BUJA** NORBAHIYAH AWANG Ts. DR. NOR AFIRDAUS ZAINAL ABIDIN DR. RAIHAH AMINUDDIN NOOR AFNI DERAMAN
REGISTRATION	**SITI FAIRUS BINTI FUZI** BUSHRA BINTI ABDUL HALIM NORDIANAH BINTI JUSOH@HUSSAIN AINON SYAZANA BINTI AB HAMID SITI NURSYAHIRA BINTI ZAINUDIN
SYSTEM	**FADILAH EZLINA SHAHBUDIN** HAJAR IZZATI MOHD GHAZALLI FADHLINA IZZAH SAMAN NOR AZIDA MOHAMED NOH
INVITATION AND PROMOTION	**SHAHITUL BADARIAH SULAIMAN** IZNI SYAMSINA SAARI NOR ADILA KEDIN

	ADI HAKIM BIN TALIB
	MOHD AMIRUL BIN ATAN
MULTIMEDIA	**Ts. NURUL NAJWA ABDUL RAHID@ABDUL RASHID**
	NOOR ASHITAH ABU OTHMAN
	ANWAR FARHAN ZOLKEPLAY
AWARD	**ANITA BINTI MOHD YASIN**
	NURUL EMYZA ZAHIDI
	FATIMAH HASHIM
	SITI RAMIZAH JAMA
	DR NURUL HUDA NIK ZULKIFLI
	MARIATHY BINTI KARIM
CERTIFICATE	**KHAIRUL NURMAZIANNA ISMAIL**
	NUR NABILAH ABU MANGSHOR
	ZUHRI ARAFAH ZULKIFLI
	HAZRATI ZAINI
INTERNATIONAL RELATIONS	**Ts. DR. SITI RAHAYU ABDUL AZIZ**
	ALBIN LEMUEL KUSHAN
	SHAHADAN SAAD
LIAISON OFFICER	**SYAFNIDAR ABDUL HALIM**
	AJK WAKIL UNTAD
SPONSORSHIP	**ANIS AMILAH SHARI**
	MOHD RAHMAT MOHD NOORDIN
	DR YUZAIMI YUNUS
	DR SURYAEFIZA KARJANTO
SECRETARIAT & APPRECIATION BANQUET	**RAIHANA MD SAIDI**
	NUR SYUHADA BINTI MUHAMMAT PAZIL
	ANIS AFIQAH SHARIP
	SITI MAISARAH MD ZAIN
	HAZWA HANIM MOHAMED HAMZAH

UNTAD'S COMMITTEE FOR I-JAMCSIIX 2023:

PROF. IR. MARSETYO, M.AG., PH.D.

PROF. I WAYAN SUDARSANA, S.SI., M.SI.

PROF. JUNAIDI, S.SI., M.SI., PH.D.

ELISA SESA, S.SI., M.SI., PH.D.

MUKRIM, M.ED., PH.D.

ZARKIANI HASYIM, S.PD., M.ED.

DR. HJ. ANI SUSANTI, M.SI.

DR. ISKANDAR, M.HUM.

DR. IR. ROIS., MP.

SYARIFUL ANAM, S.SI., M.SI., PH.D.

DR. NAHARUDDIN, S.PD, M.SI.

DR. DRG. ELLI YANE BANGKELE, M.KES.

HERMAN, SKM., M.MED.ED.

DR. IR. SAMLIOK NDOBE, M.SI.

DR. RAHMAT BAKRI, S.H., M.H.

DR. HAERUL ANAM, SE., M.SI.

DR. IR. BAKRI, S.T., PG. DIPL. ENG., M.PHIL.

DR. IR. MUHAMMAD YAZDI PUSADAN, S.KOM., M.ENG.

IR. SYAIFUL HENDRA, S.KOM., M.KOM.

RIZANA FAUZI S.T., M.T.

MOHAMMAD FAJRI, S.SI., M.SI.

NURUL FISKIA GAMAYANTI, S.SI., M.SI.

DR. NUR'ENI, S.SI., M.SI.

IMAN SETIAWAN, S.SI., M.SI.

FADJRIYANI, S.SI., M.SI.

LIST OF SPONSORS

External Company Sponsors

Klinik Dr Jamaluddin
Klinik Mawar Jasin
Nasi Ayam Ala Cina Zul
ADS Oasis Enterprise
Noorys Enterprise
Che Ramli bin Che Ismail
Beria Maju Enterprise
Rintiz rezeki
H&K food cafe
HS Gerak Wawasan

Individual Sponsors

En. Muhammad Hanif bin Abdul Aziz
Nor Suhaida binti Karjanto

Table of Contents

Universiti Teknologi MARA Cawangan Melaka
i-JaMCSIIX 2023
International Jasin Multimedia & Computer Science Invention and Innovation Exhibition
In Partnership with: Tadulako University

Clustering Regencies/Cities in Central Sulawesi Province Based on Poverty Level Using the Average Linkage Method with Principal Component Analysis (PCA)

Paskal Immanuel Kontoro[1], Virga Damayanti[2], Arditya Sulistya Ningsih Apusing [3], Alsya Putri Sigandhia[4], and Nurul Fiskia Gamayanti[5]

[1,2,3,4,5] Tadulako University, Indonesia

paskalimmanuel280302@gmail.com, virghadamayanti26@gmail.com, adityasulistya6@gmail.com, sigandhiaal@gmail.com, nurulfiskia@gmail.com

***Abstract*— Poverty is a chronic problem that has haunted Indonesia throughout its history and become a central focus of national development because poverty is the root of various problems. In 2022, Central Sulawesi is one of Indonesia's ten provinces with the highest percentage of poor people. To determine poverty alleviation policies on target, the Government needs to pay attention to the characteristics of each region because it has different characteristics. Therefore, this research aims to group 13 regencies and cities in Central Sulawesi based on poverty levels in 2022 using the Average Linkage method with Principal Component Analysis (PCA) to support the government's efforts to reduce poverty rates. The factors used in measuring poverty levels as a basis for grouping are the number of poor people, poverty depth index, human development index, Gini ratio, poverty severity index, and open unemployment rate. Two clusters were formed from the results of this research. The first cluster is the cluster with the highest poverty rate, which consists of 12 regencies, namely Banggai Kepulauan, Banggai Laut, Tojo Una-Una, Buol, Morowali Utara, Parigi Moutong, Banggai, Poso, Donggala, Toli-Toli, Morowali, and Sigi. Meanwhile, the second cluster has the lowest poverty level, consisting of one city, Palu City.**

Keywords*— *Central Sulawesi Province, Average Linkage Method, Principal Component Analysis, Poverty Level

I. INTRODUCTION

Indonesia has a population reaching 277,749,853 people in 2022, making it the fourth most populous country in the world. However, the significant population in Indonesia has led to various issues, including poverty. Poverty in Indonesia results from injustice and inequality in distributing the well-being of the people. Data from BPS (2023) indicates that in September 2022, the number of poor people in Indonesia reached 26.36 million, with the Central Sulawesi Province being one of the ten provinces with the highest percentage of impoverished population [1].

Central Sulawesi faces severe challenges in terms of poverty. According to a report from the Indonesian Minister of Human Development and Culture, the level of extreme poverty in the Central Sulawesi Province is relatively high, with the number of poor people reaching 389,710 in September 2022, experiencing a significant increase. This phenomenon highlights the need to improve poverty alleviation strategies, especially in Central Sulawesi, which has distinct characteristics [2].

Efforts to address poverty issues in Central Sulawesi can be done with this approach cluster analysis with the average linkage method. Cluster analysis is a multivariate technique that allows for grouping objects based on their characteristics. To conduct cluster analysis, the PCA method is employed to address the violation of multicollinearity assumptions. The results of this grouping will provide better insights into tackling poverty issues in Central Sulawesi, which presents unique challenges in achieving poverty reduction targets by 2030 [3]. Therefore, the researcher is grouping regencies/cities in the Central Sulawesi Province based on the poverty level using the Average Linkage Method with Principal Component Analysis (PCA) for this research.

II. MATERIALS

A. Average Linkage Method

The Average Linkage Method is a clustering process based on the average distance between objects [4]. The average linkage method uses the closest distance, and this method can be used to group objects or variables. The algorithm in this method is as follows.

1) Determines the object corresponding to the closest distance in the distance matrix $D = \{d_{ik}\}$. After that, combining the corresponding U and V objects is the obtained cluster (UV). Next, calculate the distance between the cluster (UV) and the unconnected object (W) using the following formula.

$$d_{(UV)W} = \frac{\sum_i \sum_k d_{ik}}{N_{UV} N_W} \qquad (1)$$

dimana :

$d_{(UV)W}$: Number of objects in the cluster (UV)

d_{ik} : Distance between objects i in the (UV) cluster dan k in the W *cluster*

N_{UV} : Number of objects in the cluster (UV)

N_W : Number of objects in the cluster (W)

After that, recalculate the new distance matrix by deleting rows and columns corresponding to clusters U and V, adding rows and columns for the distance between clusters (UV) and the remaining clusters.

2) Repeat step 2 until it forms into 1 cluster.

B. Principal Component Analysis

Principal component analysis, or PCA, is a multivariate analysis technique that transforms original correlated variables into new variables that do not correlate by reducing them so that they have fewer dimensions but absorb some of large amount of variance from the initial data without losing the information contained [5]. Selecting the main components can be done in 3 ways: selecting the main components with a cumulative diversity value of 80%, selecting eigen values greater than 1, and finally by looking at the scree plot. However, some experts recommend that when selecting a main component, it would be better to look at the eigenvalue that has a value greater than 1 [6].

III. METHODS

A. Research Data and Variables

The data used in this study are secondary data sourced from the Central Statistics Agency's website and the publication "Provinsi Sulawesi Tengah Dalam Angka 2023" in the form of a book. This research utilizes poverty-level data for the year 2022 based on districts/cities in the Central Sulawesi Province. The population and sample in this study consist of 13 regencies/cities in the Central Sulawesi Province. The variables used include the number of poor people (X_1), poverty depth index (X_2), human development index (X_3), gini ratio (X_4), poverty severity index (X_5), and open unemployment rate (X_6).

B. Data Analysis

The data analysis in this research uses the Average Linkage method with Principal Component Analysis (PCA) assisted by RStudio software. The following are the analysis steps conducted in this research: (1) Data collection and input into RStudio. (2) Describing the data using descriptive statistical analysis. (3) Standardizing data by transforming it into z-scores. (4) Testing the assumptions of sample sufficiency and multicollinearity. (5) Applying the PCA method to address multicollinearity issues. (6) Determining similarity measures using the Euclidean distance. (7) Implementing the average linkage algorithm. (8) Determining the optimal number of clusters using silhouette width. (9) Obtaining cluster results from the average linkage analysis with PCA. (10) Visualizing the clustered data on a map. (11) Interpretation and concluding.

IV. RESULTS AND FINDINGS

A. Cluster Analysis Assumption Test

1) Sample Adequacy Test

Table 1. Kaiser Meyer Olkin (KMO)

Variable	X_1	X_2	X_3	X_4	X_5	X_6	***Overall***
KMO	0,55	0,69	0,70	0,65	0,66	0,81	0,68

Table 1 shows that the overall KMO value is 0.68. From the results of each variable and overall KMO> 0.5, it can be concluded that the sample we used is representative, which means our sample represents the population, so it is appropriate to carry out cluster analysis.

2) Multicollinearity Test

Table 2. *Bartlett's Test of Sphericity*

Chi-Square	***Df***	***P-Value***
45,11	15	$7{,}355 \times 10^{-5}$

Based on Table 2, the results of the multicollinearity test show p-value = $7{,}355 \times 10^{-5}$, which means that with a significance level of 5%, the p-value ($7{,}355 \times 10^{-5}$) < $\alpha(0.05)$. It indicates that there are symptoms of multicollinearity in the variables X_1, X_2, X_3, X_4, X_5, X_6. Therefore, we can proceed to principal component analysis (PCA) testing to overcome multicollinearity violations.

B. Principal Component Analysis

Table 3. Principal Component Analysis (PCA)

Component	***1***	***2***	***3***	***4***	***5***	***6***
Eigen values	3,51	1,45	0,55	0,30	0,13	0,07
Total Cumulative Variance (%)	58,48	82,67	91,75	96,72	98,85	100

The criteria used in determining PCA are seen from eigen values of more than $1(\lambda > 1)$. Based on the results of the principal component analysis in Table 3, it was obtained that the number of PCs formed was two PCs, namely PC1 and PC2 with variance of the six variables by 82.67%. After obtaining the formed PC, the distance is measured to carry out cluster analysis.

C. Distance Measures

Table 4. Euclidean distance between pairs of objects

Regency/City	***Banggai Island***	***Banggai***	...	***Palu City***
Banggai Island	0	2,30	...	5,79
Banggai	2,30	0	...	3,51
⋮	⋮	⋮	⋱	⋮
Banggai Laut	0,24	2,36	...	5,86
Morowali Utara	0,80	2,14	...	5,60
Palu City	5,79	3,51	...	0

Based on Table 4, it can be seen that one of the values we marked in red indicates that Banggai Laut Regency and Banggai Kepulauan Regency have the closest distance with a distance of 0,24. It indicates that Banggai Laut Regency and Banggai Kepulauan Regency have similar characteristics in terms of poverty levels.

D. Clustering Metode Average Linkage

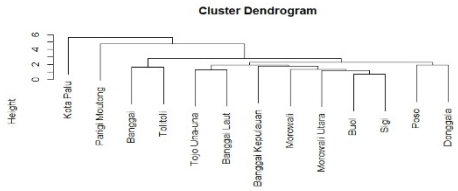

Fig. 1. Dendrogram *Clustering Result*

Based on Figure 1, it can be seen that a combination of objects that are the closest in distance will be combined to form a group containing objects with similar characteristics. From the dendrogram of clustering results using the average linkage method, it is possible to form 2 clusters, 4 clusters, or even 6 clusters, so it is necessary to determine the optimum number of clusters to be formed.

E. Determination of the Optimum Number of Clusters

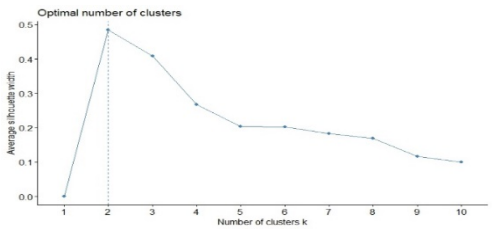

Fig. 2. The Optimum Number of Clusters with Silhouettes Method

Based on Figure 1, it can be seen that the average silhouette width obtained has the most significant value at number 2. This shows that the optimum number of clusters formed in our research is 2 clusters

F. Cluster Division with the Optimum Number of Clusters

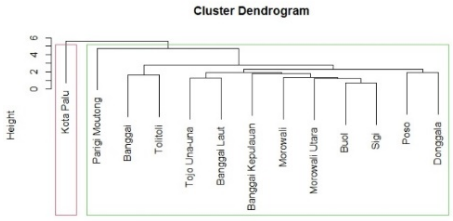

Fig. 3. Dendrogram of Cut-Off Results

Based on Figure 3, it can be formed based on the cut-off results we have carried out; 2 clusters were formed. Cluster 1 consists of 12 regencies, namely Banggai Island, Banggai Laut, Tojo Una-Una, Buol, North Morowali, Parigi Moutong, Banggai, Poso, Donggala, Toli-Toli, Morowali, and Sigi, which are coloured green in the box, and Cluster 2 consists of 1 city, Palu City, which is coloured pink in the box.

G. Group Profiling

Table 5. Characteristics of Each Cluster Based on Average

Cluster	X_1	X_2	X_3	X_4	X_5	X_6	***The Average of All Variables except X_3***
1	30,13	2,10	68,28	0,27	0,59	2,69	7,16
2	26,75	0,94	82,02	0,36	0,25	6,15	6,89

Based on Table 5, the characteristics of each cluster are obtained, wherein the table, the red color indicates the highest value. Overall, this average does not involve the human development index (X_3), because X_3 has a negative relationship with poverty. Meanwhile, X_1, X_2, X_4, X_5, and X_6 have a positive relationship. Briefly, Cluster 1 consists of 12 regencies whose poverty level is at the highest level at 7.16. Meanwhile, Cluster 2 comprises one city with the lowest poverty level of 6.89.

V. CONCLUSIONS

Based on the results of cluster analysis using the average linkage method with principal component analysis. The characteristics of cluster 1 show that this cluster has the highest poverty level compared to cluster 2. Cluster 1 consists of 12 regencies which have variables for the number of poor people (X_1), poverty depth index (X_2), and poverty severity index (X_5) with the highest average compared to cluster 2. Meanwhile, cluster 2 only includes Palu City with the highest average human development index (X_3), Gini ratio (X_4), and open unemployment rate (X_6) compared to cluster 2.

ACKNOWLEDGMENT

First of all, our team would like to express our gratitude to UiTM Cawangan Melaka Campus Jasin for hosting the International Jasin Multimedia & Computer Science Invention and Innovation Exhibition (i-JaMCSIIX). Next, we would like to sincerely thank Badan Pusat Statistik (BPS) for providing data. We also wish to convey our heartfelt appreciation to Ms. Nurul Fiskia Gamayanti, S.Si., M.Stat, our supervisor, for her guidance, advice, and support throughout the research process. Last, we want to thank all the team members for their hard work, dedication, and extraordinary contributions in completing this abstract on time.

REFERENCES

[1] [BPS] Badan Pusat Statistik. (2023). *Badan Pusat Statistik Republik Indonesia*. Jakarta.

[2] Farachi, A. N., Ni'matillah, T. B., Putra, N. A., & Widodo, E. (2019*). Analisis Pengelompokan Kabupaten/Kota di Provinsi Papua Berdasarkan Indikator Kemiskinan Tahun 2018*. Seminar Nasional Multimedia & Artificial Intelligence, Vol.2. 16-20.

[3] Iis, Yahya, I., Wibawa, G. N., Baharuddin, Ruslan, & Laome, L. (2022). *Penggunaan Korelasi Cophenetic Untuk Pemilihan Metode Cluster Berhierarki Pada Mengelompokkan Kabupaten/Kota Berdasarkan Jenis Penyakit Di Provinsi Sulawesi Tenggara Tahun 2020*. PROSIDING SEMINAR NASIONAL SAINS DAN TERAPAN (SINTA) VI.

[4] Rachmatin, D. (2014). Aplikasi Metode Agglomerative Dalam Analisis Klaster Pada Data Tingkat Polusi Udara. *Jurnal Ilmiah Program Studi Matematika STKIP Siliwangi Bandung*, Vol 3, No. 2.

[5] Utami, N. D. (2017). Perbandingan Hasil Pengelompokkan Antara Metode Average Linkage, Ward, Complete Linkage, Dan Single Linkage (Studi Kasus: Indikator Kesehatan Indonesia Tahun 2015). Yogyakarta: Skripsi. Fakultas Matematika Dan Ilmu Pengetahuan Alam Universitas Islam Indonesia.

[6] Sriningsih, M., Hatidja, D., & Prang, J. D. (2018). Penanganan Multikolinieritas Dengan Menggunakan Analisis Regresi Komponen Utama Pada Kasus Impor Beras di Provinsi Sulut. *Jurnal Ilmiah Sains,* Vol.18 No.1, 18-24.

i-JaMCSIIX 2023

PUBLISHED BY:
i-JaMCSIIX
Universiti Teknologi MARA Cawangan Melaka
Kampus Jasin
77300 Merlimau, Melaka

Tel: 062645000
Email: jamcsiix@uitm.edu.my
Web: https://jamcsiix.uitm.edu.my/