

**IMPLEMENTASI METODE AGGLOMERATIVE HIERARCHICAL
CLUSTERING DENGAN ALGORITMA AVERAGE LINKAGE
PADA DATA KEMISKINAN DI SULAWESI SELATAN**

SKRIPSI



AZMAN

H12116313

**PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN**

MAKASSAR

MARET 2020



**IMPLEMENTASI METODE AGGLOMERATIVE HIERARCHICAL
CLUSTERING DENGAN ALGORITMA AVERAGE LINKAGE
PADA DATA KEMISKINAN DI SULAWESI SELATAN**



SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains pada
Program Studi Statistika Departemen Statistika Fakultas Matematika dan Ilmu
Pengetahuan Alam Universitas Hasanuddin

AZMAN

H12116313

**PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN**

MAKASSAR

MARET 2020



LEMBAR PERNYATAAN KEOTENTIKAN

Saya yang bertanda tangan di bawah ini menyatakan dengan sungguh-sungguh
bahwa skripsi yang saya buat dengan judul:

**“Implementasi Metode *Agglomerative Hierarchical Clustering*
dengan Algoritma *Average Linkage* pada data kemiskinan
di Sulawesi Selatan”**

adalah benar hasil karya saya sendiri, bukan hasil plagiat dan belum pernah
dipublikasikan dalam bentuk apapun.

Makassar, 03 Maret 2020



Optimized using
trial version
www.balesio.com

iii



Scanned with
CamScanner

IMPLEMENTASI METODE AGGLOMERATIVE HIERARCHICAL CLUSTERING DENGAN ALGORITMA AVERAGE LINKAGE PADA DATA KEMISKINAN DI SULAWESI SELATAN

Disetujui oleh:



NIP. 19730227 199802 2001

Pembimbing Pertama

Sri Astuti Thamrin, S.Si., M.Stat, Ph.D

NIP. 19731228 200003 1001

Pada Tanggal: 03 Maret 2020



Optimized using
trial version
www.balesio.com

iv



Scanned with
CamScanner

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh :

Nama : Azman
NIM : H12116313
Program Studi : Statistika
Judul Skripsi : Implementasi Metode *Agglomerative Hierarchical Clustering* dengan Algoritma *Average Linkage* pada data kemiskinan di Sulawesi Selatan

Telah berhasil dipertahankan di hadapan Dewan Pengaji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Sains pada Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

DEWAN PENGUJI

1. Ketua : Anisa, S.Si., M.Si. (.....)
2. Sekretaris : Sri Astuti Thamrin, S.Si., M.Stat, Ph.D (.....)
3. Anggota : Dr. Anna Islamiyati, S.Si., M.Si. (.....)
4. Anggota : Dr. Dr. Georgina M. Tinungki, M.Si. (.....)



di : Makassar

: 03 Maret 2020

Optimized using
trial version
www.balesio.com

v



Scanned with
CamScanner

KATA PENGANTAR

Segala puji bagi Allah Rabb semesta alam, shalawat serta salam semoga selalu dilimpahkan kepada Nabi Muhammad SAW dan kepada para keluarga serta sahabat beliau. Alhamdulillah wasy-syukurillah, berkat pertolongan Allah akhirnya skripsi dengan judul "**Implementasi Metode Agglomerative Hierarchical Clustering dengan Algoritma Average Linkage pada data kemiskinan di Sulawesi Selatan**" yang disusun sebagai salah satu syarat akademik untuk meraih gelar sarjana pada Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin ini dapat dirampungkan. Penulis berharap skripsi ini dapat memberikan tambahan pengetahuan baru bagi para pembelajar statistika.

Penulis menyadari bahwa penyelesaian tugas akhir ini tidak luput dari bantuan dari berbagai pihak. Oleh karena itu, penulis ucapan terima kasih yang sebesar-besarnya kepada kedua orang tua penulis, Ayahanda **Amiruddin** dan Ibunda **Mareamang** yang telah menjadi inspirasi, mendidik dan membesarkan penulis dengan bertabur cinta, kasih sayang, serta dengan ikhlas telah mengiri setiap langkah penulis dengan doa dan restunya. Untuk kakak tersayang **Aswan Amiruddin** dan adik tersayang **Azniaty**, terima kasih atas segala bentuk motivasi dan semangat tiada henti yang telah diberikan kepada penulis. Serta untuk keluarga besarku, terima kasih atas doa dan dukungannya selama ini. Penghargaan yang tulus dan ucapan terima kasih juga penulis ucapkan kepada:

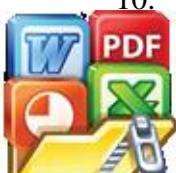
1. **Ibu Prof. Dr. Dwia Aries Tina Palubuhu, MA**, Selaku Rektor Universitas Hasanuddin beserta seluruh jajarannya.
2. **Bapak Dr. Eng. Amiruddin**, selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin beserta seluruh jajarannya
3. **Ibu Dr. Nurtiti Sunusi, S.Si, M.Si**, selaku Ketua Departemen Statistika serta Segenap dosen pengajar dan staf Departemen Statistika yang telah membekali ilmu dan kemudahan-kemudahan kepada penulis dalam berbagai hal selama adi mahasiswa di Departemen Statistika.

Anisa, S.Si, M.Si selaku Pembimbing Utama dan **Ibu Sri Astutinrin, S.Si, M.Stat, Ph.D** selaku Pembimbing Pertama yang telah



meluangkan waktu dan pemikirannya untuk memberikan arahan, dorongan, dan motivasi kepada penulis mulai dari awal hingga selesainya penulisan tugas akhir ini.

5. **Alm. Pak Drs. M. Shaleh A.F., M.Si.,** dan **Pak Andi Khresna Jaya, S.Si., M.Si**, selaku Penasehat Akademik atas segala masukan, bantuan, nasehat serta motivasi yang diberikan dari penulis masih berstatus Mahasiswa Baru sampai penulis menyelesaikan studi di Departemen Statistika.
6. **Ibu Dr. Anna Islamiyati, S.Si, M.Si**, selaku Ketua Tim Penguji dan **Ibu Dr. Georgina M. Tinungki, M.Si**, selaku Sekretaris Tim Penguji yang telah memberikan kritikan dan pembelajaran yang membangun dalam penyempurnaan tugas akhir ini serta atas waktu yang di luangkan kepada penulis.
7. Sahabat Pondok Hisyam : **Muhammad Naim, Muhammad Ashar, Samsir Aditya Ania, Muhammad Idman, Kak Gio, Arkan, Kak hadi, Kak abu .** *Jazaakumullah khairan* karena masih selalu ada untuk mendoakan, mengajak pada kebaikan, tidak pernah meninggalkan penulis walaupun sedang *futur*, Selalu menemani sejak maba dan selalu berbagi cerita serta saling berbagi makanan di kala akhir bulan. Semoga Allah Azza wajala menjaga kita semua!
8. Teman-teman bacotku yang selalu nge-gas tanpa kenal waktu dan tempat **Nur Alfianingsih Raja (Inci), Nurul Afdhaliah (Dilla), Nurazrawiyah Amin (Wiya)**. Terima kasih selalu setia untuk mendengar dan berbagi kisah kehidupan perkuliahan dan menemani penulis selama mengerjakan skripsi Terima kasih atas warna warni kehidupan selama ini. Semangat yah kerja Skripsinya!
9. Teman XII MIA 1 (DUBELEX ONE) : **Rika Riana, Alda Muchtar, Syahril A., Arham Akramullah** dan yang tidak sempat disebutkan. Terima kasih karena masih menjaga komunikasi hingga saat ini.
10. Teman FLP Unhas : **Novita, Muliana, Afi, Kak Shaleh, Mila** dan yang tidak dapat disebutkan. Terima kasih atas ilmu kepensilisannya dan rsamaannya selama kepengurusan 2018-2019.



11. Sepupu-sepupu yang Selalu siap membantu kapanpun dan dimanapun : **Hisyam, Rahmat, Wahyudi, Kak Kartika**. Terima kasih, karena tidak pernah bosan pasang badan untuk penulis.
12. Teman-teman sebagai *support system* : **Fita, Arsyillah Mughni R., Kak Nurwati, Akh Fachmi** dan **Fadli Arif**. Terima kasih karena telah memberikan pengalaman berharga bagi penulis dan selalu mendukung dan memberi saran bagi penulis.
13. Sobat minor : **Andis, Imma, Aten, Eja** dan **kawan-kawan**. Walaupun penulis bukan anggota dalam geng ini, Tapi penulis mengucapkan terima kasih karena bantuan dan kerecehannya yang *unlimited*.
14. Saudara seperjuangan di Statistika 2016 terkhusus **Sultan, Owi, Endang** dan **Ibnu** dan seluruh **STATISTIKA 2016** yang tidak sempat disebutkan. Terima kasih atas kebersamaannya selama ini dan Telah mau menjadi teman pertama penulis sejak masih maba. Semangat kerja Skripsi juga!. Terkhusus **Iis Cendrah kasih**, Terima kasih karena telah berjuang bersama mulai Proposal sampai Sidang.
15. Keluarga besar **HIMATIKA FMIPA UNHAS** dan **HIMASTAT FMIPA UNHAS** terkhusus saudara seperjuangan **A16ORITMA 2016** yang tidak sempat disebutkan satu persatu, terima kasih untuk cerita suka dan duka perjuangannya.
16. Teman-Teman KKN Posko Desa Pa'Bentengang : **Ari, Mawa, Inu, Inar, Diat, Sakkir, Muda', Rama, Zizah dan Afni**. Terima kasih untuk cerita 1 Bulannya di Lokasi KKN, Semoga cerita kita takkan berakhir.
17. Semua pihak yang telah membantu secara langsung maupun tidak langsung dalam penyusunan Tugas Akhir ini yang tak sempat penulis sebutkan satu per satu.

Penulis menyadari bahwa masih banyak kekurangan dalam skripsi ini. Oleh karena itu, dengan segala kerendahan hati penulis memohon maaf. Akhir kata alisan ini memberi manfaat kepada pembaca.



Makassar, 03 Maret 2020

Azman

PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Hasanuddin, saya yang bertandatangan di bawah ini:

Nama : Azman
NIM : H12116313
Program Studi : Statistika
Departemen : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam
Jenis karya : Skripsi

demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Hasanuddin **Hak Bebas Royalti Noneksklusif (Non-exclusive Royalty-Free Right)** atas karya ilmiah saya yang berjudul:

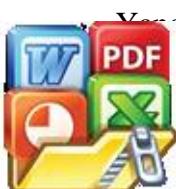
“Implementasi Metode Agglomerative Hierarchical Clustering dengan Algoritma Average Linkage pada data kemiskinan di Sulawesi Selatan”

Beserta perangkat yang ada (jika diperlukan). Terkait dengan hal di atas, maka pihak universitas berhak menyimpan, mengalih-media/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di Makassar pada tanggal 03 Maret 2020

Yang menyatakan



ABSTRAK

Kemiskinan merupakan salah satu persoalan mendasar yang menjadi pusat perhatian pemerintah di negara manapun. Salah satu aspek penting untuk mendukung Strategi Penanggulangan Kemiskinan adalah tersedianya data kemiskinan yang akurat dan tepat sasaran. Analisis *Cluster* merupakan analisis yang digunakan untuk mengelompokkan kabupaten/kota di Sulawesi selatan berdasarkan informasi data kemiskinan. Tujuan dari penelitian ini adalah untuk mengelompokkan dan mengidentifikasi karakteristik setiap kelompok dengan menggunakan Metode *Agglomerative Hierarchical Clustering*. Algoritma yang digunakan dalam Metode *Agglomerative Hierarchical Clustering* adalah Algoritma *average linkage*. Untuk mengatasi masalah multikolinearitas variabel predictor pada data kemiskinan menggunakan Analisis Komponen Utama (AKU). Interpretasi *Cluster* menggunakan nilai *Centroid* setiap cluster yang terbentuk. Implementasi Algoritma *average linkage* pada data kemiskinan di Sulawesi selatan tahun 2017, menunjukkan pengelompokan kabupaten/kota di Sulawesi Selatan menghasilkan 3 cluster dengan cluster 1 yang berjumlah 21 kabupaten/kota, cluster 2 berjumlah 2 kabupaten/kota dan cluster 3 berjumlah 1 kabupaten/kota.

Kata Kunci : Kemiskinan, Analisis *Cluster*, *Agglomerative Hierarchical Clustering*, Algoritma *average linkage*, Analisis Komponen Utama.



ABSTRACT

Poverty is one of the fundamental problems that is the center of attention of governments in any country. One of important aspect to support the Poverty Reduction Strategy is the availability of accurate and targeted poverty data. Cluster Analysis is an analysis used to group districts / cities in South Sulawesi based on poverty data information. The purpose of this research is to classify and identify the characteristics of each group by using the Agglomerative Hierarchical Clustering Method. The algorithm used in the Agglomerative Hierarchical Clustering Method is the average linkage algorithm. To overcome the multicollinearity problem of predictor variables in poverty data using Principal Component Analysis (PCA). Cluster Interpretation uses the Centroid value of each cluster formed. Implementation of the average linkage algorithm on poverty data in South Sulawesi in 2017, indicating the groups of regencies/cities in South Sulawesi produces 3 clusters with cluster 1 totaling 21 regencies/cities, cluster 2 totaling 2 regencies / cities and cluster 3 totaling 1 regency/cities.

Keywords : Poverty, Cluster Analysis, Agglomerative Hierarchical Clustering, average linkage algorithm, Principal Component Analysis.



DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN JUDUL.....	ii
LEMBAR PERNYATAAN KEOTENTIKAN	iii
LEMBAR PERSETUJUAN PEMBIMBING	iv
HALAMAN PENGESAHAN.....	v
KATA PENGANTAR	vi
PERSETUJUAN PUBLIKASI KARYA ILMIAH.....	ix
ABSTRAK	x
ABSTRACT	xi
DAFTAR ISI.....	xii
DAFTAR TABEL.....	xv
DAFTAR GAMBAR	xvi
DAFTAR LAMPIRAN.....	xvii
BAB I	1
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	4
BAB II.....	5
TINJAUAN PUSTAKA	5
alisis Multivariat.....	5
alisis <i>Cluster</i>	5



2.3 Metode Hirarki	6
2.4 Metode Penggabungan	6
2.5 Algoritma <i>Average Linkage</i>	7
2.6 Matriks Varians-Kovarians dan Matriks Korelasi.....	8
2.7 Pencilan	9
2.8 Jarak Mahalanobis	10
2.9 Uji Kaiser Meyer Olkin (KMO)	10
2.10 Uji Multikolinearitas	11
2.11 Analisis Komponen Utama	11
2.12 Ukuran Jarak <i>Euclidean</i>	13
2.13 Interpretasi <i>Cluster</i> Yang Telah Terbentuk.....	13
2.14 Kemiskinan.....	14
 BAB III	16
 METODE PENELITIAN.....	16
3.1 Sumber Data	16
3.2 Deskripsi Variabel	16
3.3 Metode Analisis.....	16
3.4 Alur Penelitian.....	19
 BAB IV	20
 HASIL DAN PEMBAHASAN.....	20
4.1 Statistik Deskriptif.....	20
4.1.1 Deskripsi Variabel Jumlah Penduduk.....	20
4.1.2 Deskripsi Variabel Laju Pertumbuhan PDRB	21
4.1.3 Deskripsi Variabel Indeks Pembangunan Manusia (IPM)	21
4.1.4 Deskripsi Variabel Rata-rata Pengeluaran Perkapita.....	22



4.1.5 Deskripsi Variabel Tingkat Pengangguran	22
4.1.6 Deskripsi Variabel Rata-rata Lama sekolah	23
4.1.7 Deskripsi Variabel RT yang menggunakan air layak	23
4.1.8 Deskripsi Variabel RT yang Pernah Membeli Beras Raskin.....	24
4.2 Menghitung Jarak Mahalanobis	24
4.3 Mendeteksi Pencilan.....	25
4.4 Uji Kecukupan Sampel.....	26
4.5 Uji Multikolinearitas	27
4.6 Analisis Komponen Utama	27
4.7 Menentukan Ukuran Jarak.....	30
4.8 Proses <i>Cluster</i> dengan Algoritma <i>Average Linkage</i>	31
4.9 Jumlah <i>Cluster</i> dan Anggotanya	34
4.10 Interpretasi <i>Cluster</i>	35
BAB V	39
PENUTUP	39
5.1 Kesimpulan.....	39
5.2 Saran	40
DAFTAR PUSTAKA	41



DAFTAR TABEL

Tabel 3. 1 Variabel Penelitian	16
Tabel 4. 1 Deskripsi Variabel Jumlah Penduduk	20
Tabel 4. 2 Deskripsi Variabel Laju Pertumbuhan PDRB.....	21
Tabel 4. 3 Deskripsi Variabel Indeks Pembangunan Manusia (IPM)	21
Tabel 4. 4 Deskripsi Variabel Rata-rata Pengeluaran Perkapita	22
Tabel 4. 5 Deskripsi Variabel Tingkat Pengangguran	22
Tabel 4. 6 Deskripsi Variabel Rata-rata Lama sekolah.....	23
Tabel 4. 7 Deskripsi Variabel Rumah tangga yang menggunakan air layak.....	23
Tabel 4. 8 Deskripsi Variabel Rumah tangga yang Pernah Membeli Beras Raskin	24
Tabel 4. 9 Jarak Mahalanobis Kuadrat	25
Tabel 4. 10 Deteksi Pencilan	25
Tabel 4. 11 Nilai KMO.....	26
Tabel 4. 12 Matriks Korelasi antar Variabel	27
Tabel 4. 13 Pasangan Nilai Eigen dan Vektor eigen.....	28
Tabel 4. 14 Proporsi variansi dan proporsi variansi kumulatif	29
Tabel 4. 15 Nilai atau Score Komponen Utama.....	30
Tabel 4. 16 Jarak Euclid Kabupaten/Kota.....	31
Tabel 4. 17 Cluster Membership	34
Tabel 4. 18 Cluster beserta anggotanya dengan Algoritma Average Linkage	35
Tabel 4. 19 Rata-Rata Variabel (Centroid) Setiap Cluster	36

