

# **SKRIPSI**

## **KLASTERISASI POLA PENYEBARAN PENYAKIT PASIEN BERDASARKAN USIA PADA PUSKESMAS KEDUNGTUBAN MENGUNAKAN ALGORITMA *K-MEANS CLUSTERING***



Oleh :

**Dyah Selvi Anggraini**

NIM: 19083000195

**PROGRAM S1 SISTEM INFORMASI  
FAKULTAS TEKNOLOGI INFORMASI  
UNIVERSITAS MERDEKA MALANG**

**2023**

**KLASTERISASI POLA PENYEBARAN PENYAKIT PASIEN  
BERDASARKAN USIA PADA PUSKESMAS KEDUNGTUBAN  
MENGUNAKAN ALGORITMA *K-MEANS CLUSTERING***



**SKRIPSI**

Diajukan kepada  
Program S1 Sistem Informasi Universitas Merdeka Malang  
Untuk Memenuhi Salah Satu Persyaratan  
Dalam Memperoleh Gelar Sarjana Sistem Informasi

**Oleh :**

**Dyah Selvi Anggraini**

NIM: 19083000195

**PROGRAM S1 SISTEM INFORMASI  
FAKULTAS TEKNOLOGI INFORMASI  
UNIVERSITAS MERDEKA MALANG**

**2023**

# LEMBAR PERSETUJUAN

## LEMBAR PERSETUJUAN


Nama : Dyah Selvi Anggraini  
NIM : 19083000195  
Universitas : Universitas Merdeka Malang  
Fakultas : Teknologi Informasi  
Program Studi : S1 Sistem Informasi  
Judul : Klusterisasi Pola Penyebaran Penyakit Pasien  
Berdasarkan Usia pada Puskesmas Kedungtuban  
Menggunakan Algoritma K-Means Clustering.

Malang, 27 Juli 2023

Disetujui Dan Diterima

Ketua Program Studi S1 Sistem Informasi

Dosen Pembimbing

  
Galandaru Swalaganata, S.Si., M.Si  
NIDN.728109104

  
Arif Saivul Affandi, S.Kom., M.M.  
NIDN.728098305

Dekan  
Fakultas Teknologi Informasi

  
Dr. Marsiana Anggrawati, SE., M.Si  
NIDN.716037601

# LEMBAR PENGESAHAN

## LEMBAR PENGESAHAN

### KLUSTERISASI POLA PENYEBARAN PENYAKIT PASIEN BERDASARKAN USIA PADA PUSKESMAS KEDUNGTUBAN MENGUNAKAN ALGORITMA K-MEANS CLUSTERING

Dipersiapkan dan disusun oleh

Dyah Selvi Angraini


19083000195

Telah dipertahankan di depan Dewan Penguji


Pada Tanggal 27 Juli 2023

Susunan Dewan Penguji

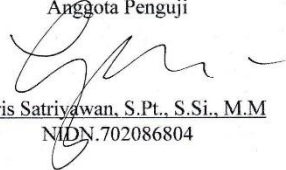
Ketua Penguji

  
Aditya Galih S., S.Kom., M.Kom  
NIDN.714018502

Sekretaris Penguji

  
Arif Saivul Affandi., S.Kom., M.M  
NIDN.728098305

Anggota Penguji

  
Haris Satriyawan, S.Pt., S.Si., M.M  
NIDN.702086804

Skripsi Ini Telah Diterima Sebagai Salah Satu Persyaratan untuk Memperoleh  
Gelar Sarjana Sistem Informasi

Malang 27 Juli 2023  
Dekan Fakultas Teknologi Informasi  
Universitas Merdeka Malang



D. Mardiana Andarwati, SE., M.Si  
NIDN.716037601

# SURAT PERNYATAAN

## SURAT PERNYATAAN

Yang bertanda tangan dibawah ini

Nama : Dyah Selvi Anggraini  
NIM : 19083000195  
Program Studi : S1 Sistem Informasi  
Bidang Kajian Skripsi : Analisis Data  
Judul Tugas Akhir : Klusterisasi Pola Penyebaran Penyakit Pasien Berdasarkan Usia pada Puskesmas Kedungtuban Menggunakan Algoritma K-Means Clustering  
Lokasi Penelitian : Puskesmas Kedungtuban, Jl. Raya Randu Cepu-Randublatung No.Km 12, Kedunggampeng, Kecamatan Kedungtuban, Kabupaten Blora, Jawa Tengah  
Alamat Rumah Asal : Desa Kalen RT 01/RW 02, Kecamatan Kedungtuban, Kabupaten Blora, Jawa Tengah.  
No. Telp/Email : 082336805142 / dyah.anggraini@student.unmer.ac.id

Dengan ini menyatakan bahwa saya benar-benar melakukan penelitian dan penulisan skripsi tersebut diatas benar-benar karya saya dan tidak melakukan plagiasi. Kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam dengan disebutkan nama dan dicantumkan dalam daftar pustaka.

Demikian surat pernyataan ini saya buat dengan sesungguhnya dan apabila dikemudian hari terdapat penyimpangan dan tidak kebenaran dalam pernyataan ini, maka saya bersedia menerima sanksi akademik sesuai dengan aturan yang berlaku.

Malang, 27 Juli 2023



Dyah Selvi Anggraini

## **MOTTO**

“Pastikan kamu tahu apa yang benar-benar ingin kamu lakukan dan lakukanlah. Jika tidak sekarang, kapan lagi?”

## **LEMBAR PERSEMBAHAN**

Skripsi ini penulis persembahkan untuk kedua orang tua, kakak, adik dan keluarga yang sudah memberikan dukungan penuh bagi penulis dari masih kecil hingga sekarang. Yang sudah dengan susah payah dalam pengorbanan, doa, usaha hingga cinta dan kasih sayang yang telah diberikan. Serta tak lupa untuk semua sahabat dekat penulis yang ikut membantu dalam memberi dukungan kepada penulis. Terima kasih banyak penulis ucapkan kepada seluruh orang-orang yang telah memberikan dukungan penuh.

## KATA PENGANTAR

Puji dan Syukur penulis haturkan kehadirat Tuhan Yang Maha Esa, karena berkat Karunia-Nya, tugas akhir ini dapat diselesaikan. Pada kesempatan yang baik ini dan dengan penuh rasa syukur yang mendalam, perkenankan penulis mengucapkan terimakasih yang sebesar-besarnya dan apresiasi yang setinggi-tingginya kepada yang terhormat:

1. Ibu Dr. Mardiana Andarwati, SE., M.Si selaku Ketua Dekan Program Studi Sistem Informasi yang telah memberikan dedikasi selama 4 tahun untuk tercapainya pembelajaran maksimal bagi program studi yang dipimpin.
2. Bapak Galandaru Swalaganata, S.Si., M.Si selaku Ketua Program Studi S1 Sistem Informasi Fakultas Teknologi Informasi Universitas Merdeka Malang.
3. Bapak Ariv Saivul Affandi, S.Kom., M.M selaku dosen pembimbing saya yang telah memberikan arahan, bimbingan, serta motivasi kepada penulis dan terima kasih atas kesabaran beliau kepada penyusun selama proses pengerjaan skripsi ini.
4. Seluruh dosen dan jajaran karyawan di Fakultas Teknologi Informasi
5. Kedua Orangtua saya Bpk. Sakroni dan Ibu Karningsih yang senantiasa memberikan dukungan mental dan materil kepada saya serta Bima adik saya tercinta.
6. Sahabat saya Neni dan Septi serta Mbak Dea dan Ipeh yang telah membantu dan memberi semangat selama proses pengerjaan skripsi ini.
7. Sahabat di tanah rantauan Dini, Della, Iloh, Cyntia, Dewa, Dicky, Deni, Koko, Bagas, Taufik yang sudah berjuang bersama dengan caranya masing-masing membantu penulis selama menyusun skripsi ini.
8. Seluruh sahabat dan sahabati rayon teknologi informasi.
9. Kepada diri saya sendiri, terima kasih sudah bertahan melewati kesulitan dan masalah yang hadir.

## DAFTAR ISI

HALAMAN JUDUL.....	i
LEMBAR PERSETUJUAN.....	ii
LEMBAR PENGESAHAN .....	iii
SURAT PERNYATAAN.....	iv
MOTTO.....	v
KATA PENGANTAR .....	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	x
DAFTAR LAMPIRAN .....	xii
ABSTRAK .....	xiii
ABSTRACT .....	xiv
BAB I PENDAHULUAN .....	1
1.1 Latar belakang .....	1
1.1 Rumusan Masalah .....	4
1.2 Tujuan.....	4
1.3 Manfaat Penelitian.....	4
1.3.1 Manfaat Teoritis .....	4
1.3.2 Manfaat Praktis .....	4
1.4 Batasan Penelitian .....	5
BAB II KAJIAN PUSTAKA .....	6
2.1 Landasan Teori .....	6
2.1.1 Data Mining .....	6
2.1.2 Clustering .....	8
2.1.3 K-Means Clustering .....	9
2.1.4 Microsoft Excel.....	12
2.1.5 Orange .....	12
2.2 Penelitian Terdahulu.....	12
BAB III METODE PENELITIAN.....	18
3.1 Desain Penelitian .....	18
3.2 Lokasi Penelitian .....	23
3.3 Populasi dan Sampel.....	23

3.4	Teknik Pengumpulan Data .....	23
3.5	Teknik Analisis Data .....	24
BAB IV HASIL DAN PEMBAHASAN .....		25
4.1	Hasil Penelitian.....	25
4.1.1	Gambaran Umum Puskesmas Kedungtuban.....	25
4.1.2	Data Penelitian.....	25
4.2	Pembahasan .....	26
4.2.1	Perhitungan K-Means Clustering Secara Manual Menggunakan Microsoft Excel.....	27
4.2.2	Pengolahan data dengan melakukan klusterisasi menggunakan Algoritma pada aplikasi orange .....	38
BAB V PENUTUP.....		54
5.1	Kesimpulan.....	54
5.2	Saran .....	54
DAFTAR PUSTAKA .....		55

## DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu .....	13
Tabel 3.1 Seleksi Data.....	19
Tabel 3.2 Transformasi Usia.....	20
Tabel 3.3 Transformasi Usia.....	20
Tabel 3. 4 Transformasi Penyakit .....	21
Tabel 3.5 Transformasi Seluruh Objek Data.....	22
Tabel 4.1 Data Kunjungan Pasien.....	26
Tabel 4.2 Transformasi Objek Data .....	27
Tabel 4.3 Centroid Secara Random .....	28
Tabel 4.4 Iterasi 2.....	31
Tabel 4.5 Titik Pusat Cluster Baru .....	32
Tabel 4.6 Iterasi 3.....	32
Tabel 4.7 Hasil C1.....	34
Tabel 4.8 Total Penderita Penyakit pada C1 .....	35
Tabel 4.9 Hasil C2.....	36
Tabel 4.10 Total Penderita Penyakit pada C2 .....	37
Tabel 4.11 Hasil C3.....	37
Tabel 4.12 Total Penderita Penyakit C3 .....	38
Tabel 4.13 Hasil C1.....	46
Tabel 4.14 Rentang Usia dan Total Penderita C1 .....	47
Tabel 4.15 Hasil C2.....	47
Tabel 4.16 Rentang Usia dan Total Penderita C2 .....	48
Tabel 4.17 Hasil C3.....	49
Tabel 4.18 Rentang Usia dan Total Penderita pada C3 .....	50
Tabel 4.19 Tabel Hasil Presentase Data Kunjungan Pasien .....	51
Tabel 4.20 Tabel Jumlah Penyakit Yang Diderita Pasien Setiap.....	52
Tabel 4.21 Penyakit yang Paling Dominan Terjadi Pada .....	52
Tabel 4.22 tabel prediksi jumlah pasien berdasarkan penyakit.....	53

## DAFTAR GAMBAR

Gambar 3.1 Desain Penelitian.....	18
Gambar 3.2 Data Kunjungan Pasien Puskesmas Kedungtuban.....	19
Gambar 4.1 Perhitungan Jarak Objek ke Centroid pada Ms.Excel.....	29
Gambar 4.2 Mengelompokkan Cluster .....	30
Gambar 4.3 Menentukan Pusat Cluster Baru di Ms.Excel.....	30
Gambar 4.4 Tampilan Utama Aplikasi Orange.....	39
Gambar 4.5 Tampilan Pengaturan File Data.....	39
Gambar 4.6 Tampilan Menambahkan Fitur Data Tabel .....	40
Gambar 4.7 Tampilan Data-Data Pada Fitur Data Tabel.....	41
Gambar 4.8 Tampilan Menambahkan Fitur K-Means .....	41
Gambar 4.9 Tampilan Pengaturan K-Means.....	42
Gambar 4.10 Menambahkan Silhouette Plot .....	42
Gambar 4.11 Tampilan Silhouette Plot.....	43
Gambar 4.12 Penambahan Scater Plot.....	43
Gambar 4.13 Tampilan Hasil Scater Plot.....	44
Gambar 4.14 Tampilan Hasil Clustering .....	44
Gambar 4.15 Tampilan Hasil Clustering .....	45
Gambar 4.16 Cara Menyimpan Hasil Pengolahan di Aplikasi Orange .....	45
Gambar 4.17 Tampilan Lokasi Penyimpanan .....	46
Gambar 4.18 Diagram Pie Bulan Januari.....	50
Gambar 4.19 Diagram Pie bulan Februari .....	51
Gambar 4.20 Diagram Pie Bulan Maret.....	51
Gambar 4.21 Grafik Trednd Penyakit Beberapa Bulan Kedepan .....	53

## **DAFTAR LAMPIRAN**

Lampiran 1 Halaman Riwayat Hidup .....	57
Lampiran 2 Hasil Plagiasi .....	58
Lampiran 3 Lembar Persetujuan Tugas Akhir .....	59
Lampiran 4 Data Kunjungan Pasien Bulan Januari-Maret .....	60

# KLASTERISASI POLA PENYEBARAN PENYAKIT PASIEN BERDASARKAN USIA PADA PUSKESMAS KEDUNGTUBAN MENGGUNAKAN ALGORITMA *K-MEANS CLUSTERING*

Dyah Selvi Anggraini

Program Studi Teknologi Informasi, Program  
Sistem Informasi Universitas Merdeka Malang

Email : [diahselvi9@gmail.com](mailto:diahselvi9@gmail.com)

## ABSTRAK

Penelitian ini bertujuan untuk mengelompokan penyakit yang diderita masyarakat wilayah Kecamatan Kedungtuban berdasarkan usia pada Puskesmas Kedungtuban dengan algoritma k-means clustering menggunakan aplikasi Ms. Excel dan Orange. Dari perhitungan menggunakan Ms. Excel didapatkan 3 cluster yaitu C1 yang merupakan cluster penderita penyakit tertinggi yang banyak dialami rentang usia lansia yaitu penyakit DM dan HT. C2 merupakan penderita penyakit sedang yang banyak dialami oleh usia dewasa yaitu penyakit lambung dan TBC. C3 merupakan cluster penyakit dengan penderita terendah yang banyak dialami oleh usia remaja dan anak-anak yaitu penyakit stroke dan penyakit syaraf. Dari hasil perhitungan pada Aplikasi Orange didapatkan hasil C1 yang merupakan 14 jenis penyakit yangmana penyakit yang sering diderita kelompok usia dewasa dan lansia ialah DM dan HT. Pada C2 berfokus pada usia remaja, dewasa dan lansia yangmana terdapat 13 jenis penyakit. Penyakit TBC dan penyakit lambung merupakan penyakit yang banyak diderita pada usia ini. Pada C3, pada cluster ini berfokus pada usia anak-anak dan remaja. Yangmana terdapat 18 jenis penyakit. Penyakit CC dan penyakit kulit merupakan penyakit yang paling banyak diderita kelompok usia tersebut. Data hasil klusterisasi menggunakan k-means clustering dapat digunakan sebagai referesi petugas penyuluhan kesehatan dalam memetakan merebaknya epidemi dan menentukan tindakan preventif di berbagai cluster penderita serta dapat membantu penentuan jumlah stok obat kedepannya.

Kata Kunci: *Clustering, Penyakit, Usia, K-Means, Ms. Excel, Orange*

# CLUSTERIZING PATIENT SPREAD PATTERNS BASED ON AGE AT KEDUNGTUBAN HEALTH CENTER USING K-MEANS CLUSTERING ALGORITHM

Dyah Selvi Anggraini

Information Technology Study Program, Program

Information System of Merdeka University Malang

Email : [diahselvi9@gmail.com](mailto:diahselvi9@gmail.com)

## ABSTRACT

This study aims to classify the diseases suffered by the people of the Kedungtuban District based on age at the Kedungtuban Health Center with the k-means clustering algorithm using the Ms. Excel and Orange. From calculations using Ms. Excel obtained 3 clusters, namely C1 which is the highest cluster of disease sufferers that are widely experienced by the elderly, namely DM and HT. C2 is a patient with moderate diseases that are mostly experienced by adults, namely gastric disease and tuberculosis. C3 is a disease cluster with the lowest number of sufferers experienced by adolescents and children, namely stroke and neurological diseases. From the calculation results in the Orange Application, the results of C1 are 14 types of diseases which are the diseases that often affect the adult and elderly age groups, namely DM and HT. C2 focuses on adolescents, adults and the elderly, where there are 13 types of diseases. Tuberculosis and gastric disease are diseases that many suffer from at this age. At C3, this cluster focuses on children and adolescents. There are 18 types of disease. CC disease and skin disease are the diseases most commonly suffered by this age group. Clustered data using k-means clustering can be used as a reference for health extension workers in preventing the spread of epidemics and determining preventive measures in various patient clusters and can help determine the amount of drug stock in the future.

*Keywords: Clustering, Disease, Age, K-Means, Ms. Excel, Orange*