



**ANALISIS PROFIL PROGRAM BANTUAN SOSIAL
MENGUNAKAN K-MEANS DENGAN METODOLOGI
CRISP-DM (STUDI KASUS: DESA BENGLE, KABUPATEN
KARAWANG)**

**TUGAS AKHIR
SKRIPSI**

**WITARLINA ZAI
41822110006**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA JAKARTA
2026**



**ANALISIS PROFIL PROGRAM BANTUAN SOSIAL
MENGUNAKAN K-MEANS DENGAN METODOLOGI
CRISP-DM (STUDI KASUS: DESA BENGLE, KABUPATEN
KARAWANG)**

**TUGAS AKHIR
SKRIPSI**

**WITARLINA ZAI
41822110006**

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA JAKARTA
2026**

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Witarlina Zai
NIM : 41822110006
Program Studi : Sistem Informasi

Menyatakan dengan sebenar-benarnya bahwa Tugas Akhir berjudul: “Analisis Profil Program Bantuan Sosial Menggunakan K-Means dengan Metodologi CRISP-DM (Studi Kasus: Desa Bengele, Kabupaten Karawang)” adalah hasil karya saya sendiri, tidak mengandung unsur plagiarisme, pelanggaran hak cipta, atau konten ilegal dalam bentuk apapun dan tidak melanggar hukum atau hak pihak manapun.

Apabila di kemudian hari ditemukan pelanggaran terhadap pernyataan ini, saya bersedia menanggung seluruh konsekuensi hukum dan membebaskan Universitas Mercu Buana dari segala bentuk tuntutan hukum dan saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Demikian pernyataan ini saya buat dengan sebenar-benarnya untuk digunakan sebagaimana mestinya.

Jakarta, 2 Agustus 2026



Witarlina Zai

PERNYATAAN SIMILARITY CHECK
/SIMILARITY CHECK STATEMENT

Saya yang bertanda tangan dibawah ini menyatakan, bahwa karya ilmiah yang ditulis oleh
/The undersigned, hereby declare that the scientific work written by

Nama /Name : Witarlina Zai
NIM /Student ID Number : 41822110006
Program Studi /Study Program : Sistem Informasi

Dengan Judul Tugas Akhir

/The title:

“ANALISIS PROFIL PROGRAM BANTUAN SOSIAL MENGGUNAKAN K-MEANS DENGAN METODOLOGI CRISP-DM (STUDI KASUS: DESA BENGLE, KABUPATEN KARAWANG)”

telah dilakukan pengecekan *similarity* dengan sistem Turnitin pada tanggal:

/Similarity checks have been carried out with the Turnitin system on the date:

22 Juli 2026

dengan nilai persentase sebesar :

/with a percentage value of:

10%

dinyatakan memenuhi standar sesuai dengan ketentuan berlaku di **Fakultas Ilmu Komputer** Universitas Mercu Buana. */declared to meet standards in accordance with applicable regulations at the Faculty of Computer Science, Universitas Mercu Buana.*

File hasil cek similarity turnitin:

/Turnitin similarity report file

https://drive.google.com/file/d/1HRdQeWDNNktC3VRGIGoU30SRwZVHzM2r/view?usp=drive_link



Jakarta, 22 Juli 2026
Admin Turnitin Fasilkom UMB



Agung Prawoto

Agung Prawoto, S.Kom., B.Sc

NIK : 322970503

HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Witarlina Zai

NIM : 41822110006

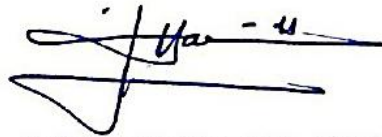
Program Studi : Sistem Informasi

Judul Laporan Skripsi : Analisis Profil Program Bantuan Sosial Menggunakan K-Means Dengan Metodologi Crisp-Dm (Studi Kasus: Desa Benge, Kabupaten Karawang).

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing,



Lukman Hakim, S.T., M.Kom

NIDN: 0327107701

Jakarta, 3 Agustus 2026

Mengetahui,

Dekan Fakultas Ilmu Komputer



Dr. Bambang Jokonowo, S.Si., M.T.I.

NIDN: 0320037002

Ketua Program Studi
Sistem Informasi



Wawan Gunawan, S.Kom., M.T., M.Kom.

NIDN: 0424108104

KATA PENGANTAR

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan Laporan Skripsi ini. Penulisan Laporan Skripsi ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer pada Fakultas Ilmu Komputer Universitas Mercu Buana. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan Laporan Skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Prof. Dr. Andi Adriansyah, M.Eng selaku Rektor Universitas Mercu Buana
2. Dr. Bambang Jokonowo, S.Si., MTI selaku Dekan Fakultas Ilmu Komputer
3. Wawan Gunawan, S.Kom., M.T selaku Ketua Program Studi Sistem Informasi
4. Lukman Hakim, S.T., M.Kom selaku Dosen Pembimbing yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan saya dalam penyusunan skripsi ini;
5. Ibu Sitilia Zalukhu dan Bapak Atoaro Zai selaku orang tua kandung penulis yang senantiasa memberikan doa dan dukungan, selama proses perkuliahan hingga penyusunan skripsi ini;
6. Ibu Esther Esti dan Bapak Harry Nenobais selaku orang tua angkat penulis yang telah memberikan fasilitas, dukungan, dan bantuan selama proses perkuliahan hingga penyusunan skripsi ini;
7. Mawar, Teri, Iter, Eli, Hati, dan Abraham selaku saudara-saudari penulis yang senantiasa memberikan dukungan, dan semangat selama proses perkuliahan hingga penyusunan skripsi ini;
8. William Yohanes Hutubessy yang telah memberikan doa, dukungan, bantuan, motivasi, dan semangat serta senantiasa menemani penulis selama proses perkuliahan hingga penyusunan skripsi ini;
9. Aulia, Aspiyah, Sheba, Febi, Fachri, Anwar, Bayu, Ridho, Rifky, Dias, dan Rayhan selaku rekan kuliah yang telah memberikan dukungan, semangat, dan kebersamaan selama proses perkuliahan hingga penyusunan skripsi ini;
10. Ailsa, Cherin, Rafly, Pieter, Zhafran, dan Sydqy selaku rekan kerja yang telah memberikan dukungan, semangat, dan kebersamaan selama proses penyusunan skripsi ini.

Akhir kata, saya berharap Tuhan Yang Maha Esa berkenan membalas segala kebaikan semua pihak yang telah membantu. Semoga Laporan Skripsi ini membawa manfaat bagi pengembangan ilmu.

Jakarta, 2 Agustus 2026

Witarlina Zai

**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI
TUGAS AKHIR DI REPOSITORI UMB**

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama : Witarlina Zai

NIM : 41822110006

Fakultas/Program Studi : Ilmu Komputer/Sistem Informasi

Judul Tugas Akhir : Analisis Profil Program Bantuan Sosial Menggunakan K-Means dengan Metodologi CRISP-DM (Studi Kasus: Desa Bengle, Kabupaten Karawang)

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Tugas Akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 2 Agustus 2026

Yang menyatakan,



Witarlina Zai

ANALISIS PROFIL PROGRAM BANTUAN SOSIAL MENGUNAKAN K-MEANS DENGAN METODOLOGI CRISP-DM (STUDI KASUS: DESA BENGLE, KABUPATEN KARAWANG)

WITARLINA ZAI

ABSTRAK

Ketepatan sasaran program bantuan sosial di tingkat desa masih menjadi tantangan karena perangkat desa kesulitan mengevaluasi cakupan program secara sistematis. Penelitian ini bertujuan menerapkan algoritma K-Means *Clustering* dengan metodologi *Cross-Industry Standard Process for Data Mining* (CRISP-DM) untuk menganalisis profil kependudukan dan cakupan tiga program bantuan sosial (PKH, Sembako, dan PBI-JK) di Desa Bengle, Kecamatan Majalaya, Kabupaten Karawang. Data yang digunakan mencakup 1.074 penduduk hasil integrasi data kependudukan tingkat RT tahun 2024 dengan data DTSEN tahun 2026 melalui pencocokan NIK, menghasilkan 773 penduduk terdaftar DTSEN (72,0%). Enam atribut demografis digunakan sebagai *fitur clustering* setelah preprocessing dengan *LabelEncoder* dan *StandardScaler*. Jumlah kluster optimal ditentukan menggunakan kombinasi Metode *Elbow* dan *Silhouette Score*. Hasil penelitian memperoleh $k = 6$ dengan *Silhouette Score* 0,5767 (kategori *strong structure*) dan *inertia* 1.084,37. Enam kluster yang terbentuk meliputi Kepala Keluarga Senior Laki-laki (3,26%), Anak-Remaja Perempuan (15,64%), Ibu Rumah Tangga (27,56%), Balita Laki-laki (16,11%), Karyawan Swasta Laki-laki (24,21%), dan Anak-Remaja Laki-laki (13,22%). Analisis cakupan bantuan sosial terhadap penduduk eligible berdasarkan desil kesejahteraan menunjukkan cakupan PKH 9,8%, Sembako 11,0%, dan PBI-JK 61,3%. Konsentrasi penerima PKH dan Sembako berada pada kluster Ibu Rumah Tangga (75,0% dan 77,8%), Sementara itu, kluster laki-laki dewasa dan anak-remaja tidak memiliki penerima PKH maupun Sembako meskipun terdapat penduduk eligible. Hasil analisis diimplementasikan pada *dashboard* interaktif berbasis *Streamlit* dengan empat halaman utama dan *fitur drill-down*, yang dirancang mendukung alur pelaporan bertingkat dari Ketua RT, Ketua RW, hingga Kepala Desa. Penelitian ini bersifat monitoring dan evaluasi cakupan, bukan penentuan kelayakan individu penerima bantuan sosial.

Kata Kunci: K-Means *Clustering*, CRISP-DM, Bantuan Sosial, DTSEN, *Dashboard*, Desa Bengle.

SOCIAL ASSISTANCE PROGRAM PROFILE ANALYSIS USING K-MEANS WITH CRISP-DM METHODOLOGY (CASE STUDY: BENGLE VILLAGE, KARAWANG REGENCY)

WITARLINA ZAI

ABSTRACT

The targeting accuracy of social assistance programs at the village level remains a challenge because village officials struggle to systematically evaluate program coverage. This study aims to apply the K-Means Clustering algorithm with the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology to analyze demographic profiles and the coverage of three social assistance programs (PKH, Sembako, and PBI-JK) in Bengle Village, Majalaya Subdistrict, Karawang Regency. The dataset consists of 1,074 residents obtained by integrating the RT-level 2024 population data with the 2026 DTSEN database through NIK matching, resulting in 773 residents registered in DTSEN (72.0%). Six demographic attributes were used as clustering features after preprocessing with LabelEncoder and StandardScaler. The optimal number of clusters was determined using a combination of the Elbow Method and Silhouette Score. The results yielded $k = 6$ with a Silhouette Score of 0.5767 (strong structure category) and an inertia of 1,084.37. The six clusters formed include Senior Male Household Heads (3.26%), Female Teenagers (15.64%), Housewives (27.56%), Male Toddlers (16.11%), Male Private Employees (24.21%), and Male Teenagers (13.22%). Coverage analysis against eligible residents based on welfare deciles shows PKH coverage of 9.8%, Sembako 11.0%, and PBI-JK 61.3%. PKH and Sembako recipients are concentrated in the Housewives cluster (75.0% and 77.8%, respectively), while three clusters of adult and adolescent males have no PKH or Sembako recipients despite containing eligible residents. The analysis results are implemented in an interactive Streamlit-based dashboard with four main pages and a drill-down feature, designed to support a hierarchical reporting flow from the RT Head, RW Head, to the Village Head. This research is intended for monitoring and coverage Evaluation, not for determining the eligibility of individual recipients.

Keywords: *K-Means Clustering, CRISP-DM, Social Assistance, DTSEN, Dashboard, Bengle Village.*

DAFTAR ISI

HALAMAN SAMPUL	0
HALAMAN JUDUL.....	i
HALAMAN PERNYATAAN KARYA SENDIRI.....	ii
HALAMAN SURAT KETERANGAN HASIL UJI TURNITIN	iii
HALAMAN PENGESAHAN	iv
KATA PENGANTAR	v
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR DI REPOSITORI UMB.....	vii
ABSTRAK.....	viii
ABSTRACT	ix
DAFTAR ISI.....	x
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR	xv
DAFTAR LAMPIRAN	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah.....	4
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Penelitian.....	5
1.5 Batasan Masalah.....	6
BAB II TINJAUAN PUSTAKA.....	8
2.1 Teori Utama.....	8
2.1.1 Metodologi CRISP-DM	8

2.1.2 Algoritma K-Means <i>Clustering</i>	9
2.1.3 Metode <i>Elbow</i> dan <i>Silhouette Score</i>	11
2.2 Penelitian Terdahulu.....	12
2.3 Gap Penelitian	29
2.4 Teori Pendukung	30
2.4.1 Data Kependudukan	30
2.4.2 Data Tunggal Sosial dan Ekonomi Nasional (DTSEN)	31
2.4.3 Program Bantuan Sosial di Indonesia.....	31
2.4.4 Desil Kesejahteraan dan Proxy Means Test	33
BAB III METODE PENELITIAN.....	34
3.1 Jenis Penelitian	34
3.2 Tahapan Penelitian	34
3.2.1 Fase <i>Business Understanding</i>	36
3.2.2 Fase <i>Data Understanding</i>	36
3.2.3 Fase <i>Data Preparation</i>	37
3.2.4 Fase <i>Modeling</i>	37
3.2.5 Fase <i>Evaluation</i>	38
3.2.6 Fase <i>Deployment</i>	38
3.3 Instrumen Penelitian	39
3.4 Subjek Penelitian	39
3.5 Desain Penelitian	40
3.5.1 Desain Arsitektur Sistem Analisis.....	40
3.5.2 Penerapan Metodologi CRISP-DM.....	41
3.5.3 Penerapan Algoritma K-Means <i>Clustering</i>	42
3.5.4 Perancangan Sistem dengan UML	44

3.6 Struktur Program	55
3.7 Teknik Pengumpulan Data	57
BAB IV HASIL DAN PEMBAHASAN	59
4.1 Persiapan Data	59
4.2 Pengumpulan dan Deskripsi Data	60
4.3 Preprocessing Data	64
4.3.1 Pembersihan Data	64
4.3.2 Pencocokan NIK dan <i>Feature engineering</i>	65
4.3.3 Encoding dan Standardisasi untuk K-Means	65
4.4 Hasil Wawancara dengan Ketua RT Desa Benge	66
4.5 Penerapan Algoritma K-Means <i>Clustering</i>	67
4.5.1 Penentuan Jumlah Klaster Optimal	68
4.5.2 Hasil <i>Clustering</i> Final	69
4.5.3 Visualisasi Klaster dalam Ruang PCA 2D	71
4.6 Profiling dan Analisis Cakupan Bantuan Sosial	71
4.6.1 Profiling Karakteristik Demografis per Klaster	72
4.6.2 Analisis Cakupan Program Bantuan Sosial per Klaster	74
4.6.3 Analisis Silang Klaster, Desil Kesejahteraan, dan Status Penerima Bansos	76
4.6.4 Temuan Analitik	81
4.7 Implementasi <i>Dashboard</i> Interaktif	83
4.7.1 Halaman Ringkasan Umum	84
4.7.2 Halaman Profil Kependudukan	84
4.7.3 Halaman Analisis Klaster K-Means	86
4.7.4 Halaman Tabel Data Detail	87

4.7.5 Implementasi Autentikasi dan Otorisasi Berbasis Peran.....	87
4.7.6 Halaman Upload Data Penduduk (Ketua RT).....	88
4.7.7 Halaman Catatan Verifikasi Warga Eligible (Ketua RW).....	89
4.7.8 Halaman Kelola Akun Pengguna (Kepala Desa)	90
4.8 Validasi Hasil	91
4.8.1 Validasi Hasil <i>Clustering</i>	92
4.8.2 Validasi Fungsional <i>Dashboard</i>	93
BAB V KESIMPULAN DAN SARAN.....	95
5.1 Kesimpulan.....	95
5.2 Saran.....	96
DAFTAR PUSTAKA	99
LAMPIRAN	103

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu.....	12
Tabel 3. 1 Struktur Direktori Program Dashboard	56
Tabel 4. 1 Deskripsi Atribut Data Sumber	62
Tabel 4. 2 Aturan Pembersihan Data.....	64
Tabel 4. 3 Distribusi Anggota per Klaster.....	69
Tabel 4. 4 Profiling Karakteristik Demografis per Klaster	73
Tabel 4. 5 Cakupan Program Bantuan Sosial per Klaster	74
Tabel 4. 6 Cakupan Program Bantuan Sosial: Dua Denominator	75
Tabel 4. 7 Distribusi Kategori Desil per Klaster (Penduduk Terdaftar DTSEN) .	77
Tabel 4. 8 Analisis Kesenjangan Cakupan Program Bantuan Sosial terhadap Penduduk Eligible per Klaster.....	78
Tabel 4. 9 Hasil Validasi Fungsional Dashboard	93

DAFTAR GAMBAR

Gambar 3. 1 Tahapan Penelitian Berbasis Metodologi CRISP-DM	35
Gambar 3. 2 Entity Relationship Diagram Basis Data Sistem Dashboard	54
Gambar 4. 1 Fungsi Pemuatan dan Pengayaan Data (Data Loading & Enrichment)	62
Gambar 4. 2 Fungsi Preprocessing, Encoding, Standardisasi, dan Fitting K-Means	66
Gambar 4. 3 Grafik Elbow Method dan Silhouette Score untuk Penentuan k Optimal	68
Gambar 4. 4 Visualisasi Klaster K-Means dalam Ruang PCA 2D	71
Gambar 4. 5 Heatmap Cakupan Program Bantuan Sosial per Klaster (%)	76
Gambar 4. 6 Komposisi Kategori Desil Kesejahteraan per Klaster (Penduduk Terdaftar DTSEN, n=773).....	77
Gambar 4. 7 Analisis Kesenjangan: Penduduk Eligible (Desil 1-4/5) vs Penerima Aktual per Klaster(%).....	79
Gambar 4. 8 Heatmap Tingkat Cakupan Program Bansos terhadap Penduduk Eligible per Klaster (%).....	79
Gambar 4. 9 Halaman Ringkasan Umum Dashboard	84
Gambar 4. 10 Halaman Profil Kependudukan Dashboard	85
Gambar 4. 11 Halaman Analisis Klaster K-Means pada Dashboard	86
Gambar 4. 12 Halaman Tabel Data Detail dengan Filter Interaktif	87
Gambar 4. 13 Halaman login pengguna	88
Gambar 4. 14 Halaman upload data penduduk	89
Gambar 4. 15 Halaman Catatan Verifikasi Warga Eligible	90
Gambar 4. 16 Halaman Kelola Akun Pengguna	91

DAFTAR LAMPIRAN

Lampiran 1 Dokumentasi kunjungan desa Bengle.....	103
Lampiran 2 Kartu Asistensi Tugas Akhir.....	104
Lampiran 3 Curriculum Vitae	106
Lampiran 4 Bukti telah mengikuti BNSP.....	107
Lampiran 5 Hasil cek turnitin.....	108