



**DESAIN APLIKASI CLUSTERING PERUSAHAAN DI KOTA
DEPOK MENGGUNAKAN FRAMEWORK STREAMLIT**

SKRIPSI

FIORENT ARIE HERNANTI

2110512116

**PROGRAM STUDI S1 SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN”
JAKARTA**

2026



**DESAIN APLIKASI CLUSTERING PERUSAHAAN DI KOTA
DEPOK MENGGUNAKAN FRAMEWORK STREAMLIT**

SKRIPSI

**Ditujukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar
Sarjana Ilmu Komputer**

FIORENT ARIE HERNANTI

2110512116

PROGRAM STUDI S1 SISTEM INFORMASI

FAKULTAS ILMU KOMPUTER

**UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN”
JAKARTA**

2026

PERNYATAAN ORISINALITAS

PERNYATAAN ORISINALITAS

Tugas Akhir Proyek ini merupakan hasil karya sendiri serta semua sumber referensi yang dikutip maupun yang dirujuk telah saya nyatakan benar.

Nama : Fiorent Arie Hernanti

NIM : 2110512116

Tanggal : 27 Januari 2026

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, maka saya bersedia dituntun dan diproses sesuai dengan ketentuan yang berlaku.

Jakarta, 27 Januari 2026

Yang Menyatakan,



Fiorent Arie Hernanti

PERNYATAAN PUBLIKASI DAN PELIMPAHAN HAK CIPTA

PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul "Desain Aplikasi Clustering Perusahaan di Kota Depok Menggunakan Framework Streamlit" adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Universitas Pembangunan Nasional "Veteran" Jakarta.

Jakarta, 27 Januari 2026



Fiorent Arie Hernanti

NIM 2110512116

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

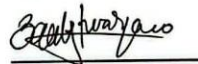
Judul : DESAIN APLIKASI CLUSTERING PERUSAHAAN DI KOTA
DEPOK MENGGUNAKAN FRAMEWORK STREAMLIT
Nama : Fiorent Arie Hernanti
NIM : 2110512116

Disetujui oleh :

Penguji 1:
Iin Ernawati, S.Kom, M.Si.



Penguji 2:
Bambang Tri Wahyono, S.Kom, M.Si.



Pembimbing 1:
Dr. Didit Widiyanto, S.Kom, M.Si.



Pembimbing 2:
Nindy Irzavika, S.Si., M.T.



Diketahui oleh:

Koordinator Program Studi:
Anita Muliawati, S.Kom., MTI.
NIP. 197005212021212002



Dekan Fakultas Ilmu Komputer:
Prof. Dr. Ir. Supriyanto, S.T., M.Sc., IPM
NIP. 197605082003121002



Tanggal Ujian Tugas Akhir :
7 Januari 2026

DESAIN APLIKASI CLUSTERING PERUSAHAAN DI KOTA DEPOK MENGUNAKAN FRAMEWORK STREAMLIT

Fiorent Arie Hernanti

ABSTRAK

Pemerintah Kota Depok memiliki data ketenagakerjaan dan informasi kepemilikan dokumen seperti Peraturan Perusahaan (PP) dan Perjanjian Kerja Bersama (PKB) dari perusahaan-perusahaan yang berada di wilayahnya. Namun, pemanfaatannya masih terbatas. Kondisi ini menyebabkan potensi informasi yang terkandung dalam data belum sepenuhnya tergali. Penelitian ini bertujuan membangun aplikasi *clustering* untuk mengelompokkan perusahaan berdasarkan total tenaga kerja tetap, total tenaga kerja kontrak, serta kepemilikan PP dan PKB. Proses *clustering* pada aplikasi memanfaatkan algoritma K-Means, dengan penentuan jumlah klaster dilakukan secara otomatis menggunakan *Davies-Bouldin Index* (DBI). Pengembangan aplikasi dilakukan dengan metode *waterfall* yang mencakup tahap analisis kebutuhan, desain, implementasi, pengujian, dan penerapan. Hasil penelitian menunjukkan bahwa aplikasi yang dibangun mampu melakukan proses *clustering* dan menyajikan hasil pengelompokkan perusahaan beserta karakteristik tiap klaster secara informatif. Aplikasi dikembangkan menggunakan *framework* Streamlit sehingga memungkinkan pengguna untuk mengakses hasil *clustering* dan visualisasi data secara interaktif. Penelitian ini diharapkan dapat membantu pemanfaatan data ketenagakerjaan secara lebih optimal serta memudahkan pengguna dalam memahami pola pengelompokkan perusahaan di Kota Depok.

Kata Kunci : Algoritma K-Means, Sistem Informasi, Streamlit, *Waterfall*

DESIGN OF A COMPANY CLUSERING APPLICATION IN DEPOK CITY USING THE STREAMLIT FRAMEWORK

Fiorent Arie Hernanti

ABSTRACT

The Depok City Government has employment data and information on documents such as Company Regulations (PP) and Collective Labor Agreements (PKB) from companies located in its territory. However, its use is still limited. This condition means that the potential information contained in the data has not been fully explored. This study aims to develop a clustering application to group companies based on the number of permanent employees, the number of contract employees, and the ownership of PP and PKB. The clustering process in the application utilizes the K-Means algorithm, with the number of clusters determined automatically using the Davies-Bouldin Index (DBI). The application was developed using the waterfall method, which includes the stages of requirements analysis, design, implementation, testing, and deployment. The results of the study show that the application built is capable of performing the clustering process and presenting the results of company grouping along with the characteristics of each cluster in an informative manner. The application was developed using the Streamlit framework, allowing users to access clustering results and data visualization interactively. This research is expected to help optimize the use of employment data and make it easier for users to understand the patterns of company clustering in Depok City.

Keywords: *Clustering, K-Means, Company, Streamlit, Data Visualization*

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Allah SWT karena atas berkat rahmat dan hidayahnya, penulis mendapatkan kesempatan untuk menyelesaikan skripsi yang berjudul “Desain Aplikasi Clustering Perusahaan di Kota Depok Menggunakan Framework Streamlit.” Adapun skripsi ini ditulis dalam rangka memenuhi syarat untuk mendapatkan gelar dan penyelesaian studi pada Program Studi S1 Sistem Informasi Universitas Pembangunan Nasional “Veteran” Jakarta.

Dalam proses penyelesaian studi dan penulisan skripsi ini, penulis banyak memperoleh dukungan, bimbingan, serta dorongan dari berbagai pihak baik secara langsung maupun tidak langsung. Untuk itu, penulis ingin menyampaikan penghargaan dan terima kasih yang sebesar-besarnya kepada:

1. Allah SWT yang selalu memberikan rahmat dan rezeki kepada penulis. Atas berkat rahmat-Nya, penulis mendapatkan kesempatan untuk menyelesaikan skripsi ini.
2. Keluarga penulis, terutama mama yang selalu memberikan dukungan dan mendoakan. Orang tua penulis, yaitu mama dan bapak penulis yang memberikan dukungan moral dan moril dalam proses penyelesaian studi. Serta, adik dan kakak penulis yang terkadang menghibur penulis baik secara langsung maupun tidak langsung.
3. Bapak Dr. Didit Widiyanto, S.Kom, M.Si, selaku Dosen Pembimbing 1 yang dengan baik hati menerima penulis sebagai mahasiswi bimbingannya disaat penulis kesulitan mencari dosen pembimbing. Terima kasih atas arahan, bimbingan, dan kesempatan yang telah diberikan.
4. Ibu Nindy Irzavika, S.Si., M.T, selaku Dosen Pembimbing 2, yang selalu memberikan masukan berharga dan membimbing penulis dengan penuh ketelitian. Terima kasih atas arahan jelas dan ketelitian yang luar biasa.
5. Sahabat dan teman-teman keren penulis, seperti Gracia dan teman-teman MAG juga teman-teman kuliah seperti Angeline, Eliyanti, Matahari, dan Keisha. Terima kasih sudah menghibur, memotivasi, mendukung, dan memvalidasi perasaan penulis.
6. Tim Back Office QR yang asik dan seru, selalu memberikan *support* penulis dalam menyelesaikan skripsi.
7. Idol, karakter, atau orang-orang favorit penulis, yang bahkan mereka mungkin tidak kenal penulis atau tahu bahwa penulis hidup, tetapi penulis melihat dan mengetahuinya secara online. Mereka secara tidak langsung memotivasi dan menghibur penulis dengan sangat baik.
8. Pihak-pihak yang tidak dapat disebutkan satu-persatu yang secara langsung atau tidak langsung pernah membantu penulis.

Penulis menyadari bahwa skripsi ini masih sangat jauh dari kata sempurna. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun dari pembaca. Semoga skripsi ini dapat memberikan manfaat bagi pengembangan ilmu di masa depan.

DAFTAR ISI

PERNYATAAN ORISINALITAS	i
PERNYATAAN PUBLIKASI DAN PELIMPAHAN HAK CIPTA	ii
LEMBAR PENGESAHAN	iii
ABSTRAK	iv
ABSTRACT	v
KATA PENGANTAR	vi
DAFTAR ISI.....	viii
DAFTAR GAMBAR	xi
DAFTAR TABEL.....	xii
DAFTAR LAMPIRAN.....	xiii
DAFTAR SIMBOL.....	xiv
DAFTAR RUMUS	xviii
BAB I 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
1.6. Sistematika Penulisan.....	4
BAB II TINJAUAN PUSTAKA.....	6
2.1. Sistem.....	6
2.2. Sistem Informasi	8
2.3. Pengawasan Ketenagakerjaan	8
2.3.1. Hubungan Kerja	11
2.3.2. Peraturan Perusahaan (PP) dan Peraturan Kerja Bersama (PKB).....	12
2.4. <i>Machine Learning</i>	13
2.5. <i>Clustering</i>	13
2.6. Algoritma K-Means Clustering.....	14
2.7. <i>Davies-Bouldin Index</i>	15
2.8. <i>Unified Modeling Language (UML)</i>	16

2.9.	Waterfall.....	18
2.10.	Python	19
2.11.	Streamlit	20
2.12.	Blackbox Testing	20
2.13.	Penelitian Terdahulu	21
BAB III METODOLOGI PENELITIAN.....		25
3.1.	Alur Penelitian	25
3.2.	Tahapan Penelitian.....	25
3.2.1.	Identifikasi Masalah.....	25
3.2.2.	Analisis Kebutuhan Sistem	26
3.2.3.	Pemodelan Data	26
3.2.4.	Desain Sistem.....	26
3.2.5.	Implementasi Kode	26
3.2.6.	Pengujian Sistem.....	26
3.2.7.	Penerapan Sistem	27
3.3.	Instrumen Penelitian.....	27
3.4.	Jadwal Rencana Penelitian.....	27
BAB IV HASIL DAN PEMBAHASAN		29
4.1.	Profil Dinas Tenaga Kerja Kota Depok	29
4.2.	Deskripsi Objek Penelitian.....	30
4.3.	Analisis Deskripsi	31
4.4.	Analisis Penelitian.....	32
4.4.1.	Deskripsi Aktor	32
4.4.2.	<i>Use Case Diagram</i>	32
4.4.3.	<i>Activity Diagram</i>	36
4.4.4.	<i>Sequence Diagram</i>	39
4.4.5.	Rancangan Antarmuka	41
4.5.	Pemodelan Data	42
4.5.1.	Persiapan Data.....	42
4.5.2.	<i>Pre-processing Data</i>	46
4.5.3.	Penentuan Jumlah Klaster Optimal (K)	49
4.5.4.	Pembuatan Model K-Means.....	50

4.5.5. Analisis Hasil <i>Cluster</i>	51
4.6. Implementasi Aplikasi	52
4.7. Pengujian Aplikasi	55
BAB V PENUTUP	58
5.1. Kesimpulan	58
5.2. Saran.....	59
DAFTAR PUSTAKA	60
LAMPIRAN	62

DAFTAR GAMBAR

Gambar 2. 1 Karakteristik sistem.....	7
Gambar 2. 2 Model Waterfall (Widarti et al. 2024).....	18
Gambar 3. 1 Diagram alur penelitian.....	25
Gambar 4. 1 Struktur organisasi Dinas Tenaga Kerja.....	29
Gambar 4. 2 Use case diagram.....	33
Gambar 4. 3 Activity diagram melakukan clustering otomatis.....	37
Gambar 4. 4 Activity diagram melihat dashboard visualisasi.....	38
Gambar 4. 5 Activity diagram melihat informasi sistem	38
Gambar 4. 6 Sequence diagram beranda.....	39
Gambar 4. 7 Sequence diagram dashboard	40
Gambar 4. 8 Sequence diagram informasi	40
Gambar 4. 9 Rancangan antarmuka halaman beranda	41
Gambar 4. 10 Rancangan antarmuka halaman dashboard	41
Gambar 4. 11 Rancangan antarmuka halaman informasi	42
Gambar 4. 12 Informasi struktur data mentah	43
Gambar 4. 13 Potongan kode dan informasi data setelah drop data duplicate	43
Gambar 4. 14 Potongan kode mengubah kata null menjadi nilai null	44
Gambar 4. 15 Informasi data dengan nilai null.....	44
Gambar 4. 16 Jumlah tenaga kerja berdasarkan klasifikasi perusahaan	45
Gambar 4. 17 Informasi dataset setelah drop data incomplete	45
Gambar 4. 18 Potongan kode feature engineering	46
Gambar 4. 19 Potongan kode feature encoding	47
Gambar 4. 20 Heatmap korelasi antar fitur tenaga kerja	47
Gambar 4. 21 Hasil perhitungan proporsi nilai 0	48
Gambar 4. 22 Potongan kode feature transformation	49
Gambar 4. 23 Hasil perhitungan nilai DBI	50
Gambar 4. 24 Potongan kode penerapan model K-Means.....	51
Gambar 4. 25 Halaman beranda I	53
Gambar 4. 26 Halaman beranda II	53
Gambar 4. 27 Halaman beranda III.....	54
Gambar 4. 28 Halaman dashboard I.....	54
Gambar 4. 29 Halaman dashboard II	55
Gambar 4. 30 Halaman informasi	55

DAFTAR TABEL

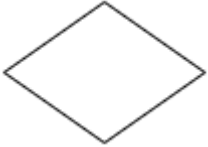
Tabel 4. 1 Analisis PIECES	31
Tabel 4. 2 Deskripsi aktor	32
Tabel 4. 3 Tabel use case skenario upload data	33
Tabel 4. 4 Tabel use case skenario melakukan clustering otomatis.....	34
Tabel 4. 5 Tabel use case skenario melihat hasil clustering	35
Tabel 4. 6 Tabel use case skenario melihat dashboard visualisasi.....	35
Tabel 4. 7 Tabel use case skenario melihat informasi sistem	36
Tabel 4. 8 Deskripsi atribut baru feature engineering.....	46
Tabel 4. 9 Deskripsi atribut yang digunakan dalam pemodelan	48
Tabel 4. 10 Distribusi jumlah perusahaan tiap cluster	51
Tabel 4. 11 Nilai rata-rata atribut per cluster	51
Tabel 4. 12 Scenarion pengujian aplikasi	55

DAFTAR LAMPIRAN

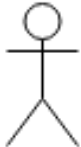
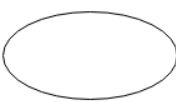
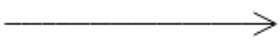
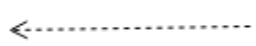
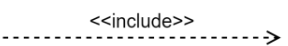
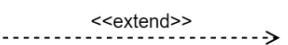
Lampiran 1 Surat Permohonan Riset	62
Lampiran 2 Hasil Wawancara.....	63
Lampiran 3 Source Code Google Colab	64
Lampiran 4 Hasil Turnitin.....	70

DAFTAR SIMBOL

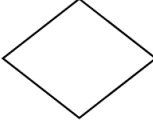
1. *Flowchart* (Khairunnisa et al. 2023)

Simbol	Nama	Fungsi Simbol
	<i>Flow</i>	Simbol digunakan untuk menghubungkan antar simbol.
	<i>Terminator</i>	Simbol yang digunakan untuk menyatakan awal atau akhir dari suatu program.
	<i>Process</i>	Simbol yang digunakan untuk menyatakan proses atau tindakan yang dilakukan.
	<i>Document</i>	Simbol yang digunakan untuk menyatakan bahwa input berasal dari suatu dokumen.
	<i>Input/Output</i>	Simbol yang digunakan untuk menyatakan input/output.
	<i>Decision</i>	Simbol yang digunakan untuk menunjukkan kondisi tertentu yang menghasilkan dua kemungkinan jawaban, yaitu ya dan tidak.
	<i>Predefine Process</i>	Simbol yang digunakan untuk pelaksanaan suatu bagian atau sub-proses.

2. *Use Case Diagram* (Sumirat et al. 2023)

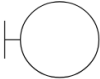
Simbol	Nama	Fungsi Simbol
	Aktor	Simbol digunakan untuk merepresentasikan peran pengguna, sistem lain, atau perangkat yang berinteraksi dengan <i>use case</i> .
	<i>Use Case</i>	Simbol digunakan untuk menggambarkan bagaimana aktor berinteraksi dengan sistem.
	<i>Association</i>	Simbol digunakan untuk menggambarkan hubungan antara aktor dengan <i>use case</i> .
	Generalisasi	Simbol digunakan untuk menggambarkan hubungan spesialisasi aktor dalam berpartisipasi pada <i>use case</i> .
	<i>Include</i>	Simbol digunakan untuk menunjukkan bahwa satu <i>use case</i> menjadi bagian dari <i>use case</i> lain.
	<i>extend</i>	Simbol digunakan untuk menunjukkan penambahan fungsi pada <i>use case</i> lain berdasarkan kondisi tertentu.

3. *Activity Diagram* (Zein et al. 2024)

Simbol	Nama	Fungsi Simbol
	Status Awal	Simbol digunakan untuk menunjukkan awal dari diagram <i>activity</i> .
	Status Akhir	Simbol digunakan untuk menunjukkan akhir dari suatu <i>activity</i> .
	Aktivitas	Simbol digunakan untuk menggambarkan aktivitas yang dilakukan sistem,
	<i>Decision</i>	Simbol digunakan untuk menggambarkan percabangan atau kondisi dalam suatu aktivitas.
	<i>Swimlane</i>	Simbol digunakan untuk memisahkan organisasi yang bertanggung jawab atas setiap aktivitas.
	<i>Control Flow</i>	Simbol digunakan untuk menunjukan aliran aktivitas secara keseluruhan.

4. *Sequence Diagram* (Zein et al. 2024)

Simbol	Nama	Fungsi Simbol
	Aktor	Simbol digunakan untuk menggambar orang yang sedang berinteraksi dengan

		sistem.
	<i>Life line</i>	Simbol digunakan untuk menggambarkan tempat mulai dan berakhirnya <i>message</i> .
	<i>Message</i>	Simbol digunakan untuk menunjukkan pengiriman pesan.
	<i>Entity Class</i>	Simbol digunakan untuk menggambarkan objek data dalam sistem.
	<i>Control Class</i>	Simbol digunakan untuk mengontrol alur interaksi antara <i>entity</i> dengan <i>boundary class</i> .
	<i>Boundary Class</i>	Simbol digunakan untuk menangani interaksi antara pengguna dan sistem.

DAFTAR RUMUS

2. 1. Rumus Elbow Criterion.....	14
2. 2. Rumus Perhitungan Centroid	15
2. 3. Rumus Euclidean Distance	15
2. 4. Rumus Sum of Square Within Cluster	16
2. 5. Rumus Sum Of Square Between Cluster	16
2. 6. Rumus Ratio.....	16
2. 7. Rumus Davies-Bouldin Index.....	16