

**KLASTERISASI DATA PENJUALAN UNTUK MENENTUKAN
BARANG TERLARIS DI DESA LUBANG BUAYA, KABUPATEN
BEKASI (STUDI KASUS: AGEN SEMBAKO DEWI SRI)**



**RANGGA ADITYA RIZALDI
2110512157**

**S1 SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” JAKARTA
JAKARTA
2025**

PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Klasterisasi Data Penjualan Untuk Menentukan Barang Terlaris Di Desa Lubang Buaya, Kabupaten Bekasi (Studi Kasus: Agen Sembako Dewi Sri)” 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 proposal ini.

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

Jakarta, Oktober 2025



Rangga Aditya Rizaldi
2110512157

**KLASTERISASI DATA PENJUALAN UNTUK MENENTUKAN
BARANG TERLARIS DI DESA LUBANG BUAYA, KABUPATEN
BEKASI (STUDI KASUS: AGEN SEMBAKO DEWI SRI)**

RANGGA ADITYA RIZALDI

2110512157

Proposal Skripsi

sebagai salah satu syarat untuk melaksanakan
penelitian oleh mahasiswa pada

Program Studi S1 Sistem Informasi

**S1 SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN” JAKARTA
JAKARTA
2025**

LEMBAR PERSETUJUAN

LEMBAR PERSETUJUAN

Yang bertanda tangan dibawah ini:

Nama : Rangga Aditya Rizaldi
NIM : 2110512157
Program Studi : Sistem Informasi Program Sarjana
Judul Tugas Akhir : Klasterisasi Data Penjualan Untuk Menentukan Produk Terlaris Di
Desa Lubang Buaya, Kabupaten Bekasi (Studi Kasus: Agen Sembako
Dewi Sri)

Dinyatakan telah memenuhi syarat dan menyetujui untuk mengikuti ujian sidang Tugas Akhir.

Jakarta, 24 Oktober 2025

Menyetujui,

Dosen Pembimbing I,

Dosen Pembimbing II,



Nur Hafifah Matondang, S.Kom., MM., MTI
NIDN. 0010108804



Anita Muliawati, S.Kom., M.TI
NIDN. 0321057001

Mengetahui,

Koordinator Program Studi,



Anita Muliawati, S.Kom., MTI.
NIDN. 0321057001

PERNYATAAN ORISINALITAS

PERNYATAAN ORISINALITAS

Skripsi ini merupakan hasil karya sendiri serta semua sumber referensi yang di kutip maupun yang di rujuk telah saya nyatakan benar.

Nama : Rangga Aditya Rizaldi

NIM : 2110512157

Tanggal : 15 Januari 2026

Bilamana di kemudian hari di temukan ketidak sesuaian dengan pernyataan ini, maka saya bersedia di tuntutan dan di proses sesuai ketentuan yang berlaku.

Jakarta, 15 Januari 2026

Yang menyatakan,

A handwritten signature in black ink is written over a yellow 10,000 Rupiah stamp. The stamp features the Garuda Pancasila emblem and the text 'REPUBLIK INDONESIA', '10000', 'METARUPIAH', and 'TEMPER'. The signature is stylized and covers the right side of the stamp.

Rangga Aditya Rizaldi

PERNYATAAN PERSETUJUAN PUBLIKASI

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan di bawah ini:

Nama : Rangga Aditya Rizaldi

NIM : 2110512157

Fakultas : Ilmu Komputer

Program Studi : S1 Sistem Informasi

Demi pembangunan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Eksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

Klasterisasi Data Penjualan Untuk Menentukan Barang Terlaris Di Desa
Lubang Buaya, Kabupaten Bekasi (Studi Kasus: Agen Sembako Deswi
Sri)

Beserta perangkat yang ada (jika diperlukan). Dengan hak bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*Database*), merawat, dan mempublikasikan Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sesungguhnya.

Dibuat di : Jakarta

Pada Tanggal : 15 Januari 2026

Yang menyatakan,



Rangga Aditya Rizaldi

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

Judul : Klasterisasi Data Penjualan Untuk Menentukan Barang Terlaris Di Desa
Lubang Buaya, Kabupaten Bekasi (Studi Kasus: Agen Sembako Dewi Sri)
Nama : Rangga Aditya Rizaldi
NIM : 2110512157

Disetujui oleh :

Penguji 1:

Ika Nurlaili, S.Kom, M.Sc.

Penguji 2:

Sarika M.Kom.

Pembimbing 1:

Nur Hafifah Matondang, S.Kom, MM, M.T.I

Pembimbing 2:

Anita Muliawati, S.Kom., MTI.

Diketahui oleh:

Koordinator Program Studi:

Anita Muliawati, S.Kom., MTI.

NIP. 197005212021212002/0321057001

Dekan Fakultas Ilmu Komputer:

Prof. Dr. Ir. Supriyanto, ST., M.Sc., IPM

NIP. 197605082003121002

Tanggal Ujian Skripsi:

5 Januari 2026



ABSTRAK

Di tengah persaingan bisnis ritel yang semakin ketat, Agen Sembako Dewi Sri menghadapi tantangan dalam pengelolaan stok dan strategi penjualan yang masih dilakukan secara manual. Ketidadaan analisis data yang mendalam menyebabkan pemilik usaha kesulitan mengidentifikasi pola penjualan barang secara akurat, yang berdampak pada efisiensi persediaan. Penelitian ini bertujuan untuk menerapkan teknik data mining menggunakan algoritma K-Means Clustering guna mengelompokkan data transaksi penjualan periode Juli 2024 hingga Juli 2025 ke dalam kategori tingkat kelarisan barang. Penelitian ini dilakukan melalui tahapan persiapan data yang meliputi pembersihan data (data cleaning), transformasi, dan normalisasi data menggunakan Min-Max Scaler. Proses klasterisasi dilakukan menggunakan algoritma K-Means dengan atribut total jumlah barang terjual dan harga rata-rata. Penentuan jumlah klaster optimal divalidasi menggunakan metode Elbow dan Silhouette Score. Hasil pengujian menunjukkan bahwa jumlah klaster paling optimal adalah 3 ($K=3$), yang merepresentasikan kategori barang Laris (15 produk), Sedang (97 produk), dan Kurang Laku (40 produk). Evaluasi model menghasilkan nilai Sum of Squared Error (SSE) sebesar 2.24 dan Silhouette Score sebesar 0.4490, yang menunjukkan bahwa pengelompokan memiliki tingkat pemisahan yang cukup baik. Hasil klasterisasi ini diimplementasikan dalam bentuk dashboard visualisasi yang dapat digunakan sebagai rekomendasi bagi Agen Sembako Dewi Sri dalam menentukan prioritas restock dan strategi promosi yang lebih tepat sasaran.

Kata Kunci: Data Mining, K-Means Clustering, Silhouette Score, Agen Sembako, Strategi Penjualan.

ABSTRACT

Amid increasingly intense competition in the retail sector, Dewi Sri Grocery Agent faces challenges in inventory management and sales strategies that are still carried out manually. The absence of in-depth data analysis makes it difficult for the business owner to accurately identify sales patterns, which affects inventory efficiency. This study aims to apply data mining using the K-Means Clustering algorithm to group sales transaction data from July 2024 to July 2025 into product salability categories. This research follows data preparation stages including data cleaning, transformation, and normalization using Min-Max Scaling. The clustering process is performed with K-Means using two attributes: total quantity sold and average price. The optimal number of clusters is validated using the Elbow method and Silhouette Score. The results indicate that the most suitable number of clusters is 3 ($K=3$), representing three categories: Best-selling (15 products), Moderate (97 products), and Low-selling (40 products). Model evaluation yields a Sum of Squared Errors (SSE) of 2.24 and a Silhouette Score of 0.4490, indicating a reasonably good separation between clusters. The clustering results are implemented in a visualization dashboard to support recommendations for restocking priorities and more targeted promotional strategies for Dewi Sri Grocery Agent.

Keywords: Data Mining, K-Means Clustering, Silhouette Score, Grocery Agent, Sales Strategy.

KATA PENGANTAR

Puji dan syukur penulis panjatkan kehadirat Allah SWT atas segala rahmat dan karunianya sehingga penulis dapat menyelesaikan skripsi ini yang berjudul " Klasterisasi Data Penjualan Untuk Menentukan Barang Terlaris Di Desa Lubang Buaya, Kabupaten Bekasi (Studi Kasus: Agen Sembako Dewi Sri)". Penyelesaian skripsi ini tidak lepas dari bimbingan dan dorongan dari berbagai pihak. Dalam kesempatan ini, penulis ingin menyampaikan ucapan terima kasih kepada:

1. Allah SWT, atas segala rahmat, petunjuk, dan perlindungan-Nya dalam setiap langkah kehidupan penulis.
2. Papa dan Mama, selaku orang tua yang telah membimbing dengan penuh kasih serta tanpa henti mendoakan keberhasilan peneliti.
3. Bapak Prof. Dr. Ir. Supriyanto, ST., M.Sc., IPM selaku Dekan Fakultas Ilmu Komputer.
4. Ibu Anita Muliawati, S.Kom.,M.TI, selaku Kepala Program Studi S-1 Sistem Informasi dan dosen Pembimbing II yang membimbing dalam proses penelitian ini.
5. Ibu Catur Nugrahaeni Puspita Dewi, S.Kom., M.Kom. selaku dosen pembimbing akademik yang telah memberikan bimbingan dan masukkan selama masa studi.
6. Ibu Nur Hafifah Matondang, S.Kom, MM, selaku Dosen Pembimbing I, yang memberikan bimbingan dan pengetahuan yang sangat berharga hingga peneliti dapat menyelesaikan skripsi ini.
7. Seluruh dosen Fakultas Ilmu Komputer UPN “Veteran” Jakarta atas ilmu dan bimbingan yang telah diberikan selama masa studi.
8. Havzy Lavayet selaku anak dari pengelola Agen Sembako Dewi Sri atas kerja sama dan bantuan selama proses penelitian.
9. Genta Mahardika Budiman, Adam, Fatih, Vico, dan seluruh teman kuliah atas kebersamaan dan dukungan yang berarti sepanjang perjalanan ini.
10. Windah Basudara, peneliti juga berterima kasih kepada seorang streamer favorit yang kontennya menjadi teman dan penyemangat selama proses penyusunan skripsi ini.
11. Risa Puspa Rini yang setia mendampingi proses ini, terima kasih atas doa, dukungan, dan semangat yang tak pernah putus.
12. Dan yang tak kalah penting, terima kasih untuk diri sendiri, terima kasih telah bertahan, berjuang, dan menyelesaikan perjalanan ini dengan penuh keyakinan.

Penyusun menyadari bahwa penyusunan skripsi ini masih jauh dari sempurna, oleh karena itu penyusun mengharapkan kritik dan saran yang bersifat membangun untuk kesempurnaan skripsi ini. Semoga skripsi ini dapat bermanfaat bagi semua pihak.

Jakarta, 15 Januari 2026

A handwritten signature in black ink, appearing to read 'Rangga Aditya Rizaldi', with a stylized, cursive script.

Rangga Aditya Rizaldi

DAFTAR ISI

LEMBAR PERSETUJUAN	iii
PERNYATAAN ORISINALITAS	iv
PERNYATAAN PERSETUJUAN PUBLIKASI.....	v
LEMBAR PENGESAHAN	v
ABSTRAK	vii
ABSTRACT	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xiii
DAFTAR SIMBOL	xv
DAFTAR LAMPIRAN	xvi
BAB I. PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Dan Manfaat Penelitian.....	4
1.4.1 Tujuan Penelitian.....	4
1.4.2 Manfaat Penelitian.....	4
1.5 Sistematika Penulisan.....	4
BAB II. TINJAUAN PUSTAKA	6
2.1 Kajian Teoritis	6
2.1.1 Usaha Mikro, Kecil, dan Menengah (UMKM)	6
2.1.2 Sembako	7
2.1.3 Penjualan	7
2.1.4 Barang.....	8
2.1.5 Data Mining.....	9
2.1.6 Klasterisasi	10
2.1.7 Klasterisasi K-Means.....	11
2.1.8 Python.....	12
2.1.9 Metode Elbow	12
2.1.10 Streamlit	13
2.2 Model Konseptual	14
2.3 Penelitian Terdahulu	16
BAB III. METODE PENELITIAN	21
3.1 Tahapan Penelitian.....	21
3.1.1 Mulai.....	22

3.1.2 Identifikasi Masalah dan Rumusan Masalah.....	22
3.1.3 Studi Literatur.....	22
3.1.4 Wawancara.....	23
3.1.5 Pengumpulan Data	23
3.1.6 Preprocessing Data	24
3.1.7 Tentukan Jumlah Klaster	25
3.1.8 Klasterisasi K-Means.....	26
3.1.9 Jika Objek Yang Berpindah Kelompok	26
3.1.10 Evaluasi Model.....	27
3.1.11 Profiling klaster	27
3.1.12 Visualisasi Hasil (Dashboard Streamlit).....	27
3.1.13 Rekomendasi Hasil.....	28
3.1.14 Selesai.....	28
3.2 Alat dan Bahan Penelitian	28
3.2.1 Perangkat Keras (Hardware)	29
3.2.2 Perangkat Lunak (Software).....	29
3.3 Waktu dan Tempat Penelitian	29
3.4 Jadwal Penelitian.....	29
BAB IV. HASIL DAN PEMBAHASAN	31
4.1 Profil Perusahaan.....	31
4.2 Deskripsi Dataset Penelitian.....	32
4.3 Data Selection.....	33
4.4 Data Pre Processing.....	35
4.4.1 Data Cleaning	36
4.5 Data Transformation.....	38
4.6 Data Mining.....	41
4.6.1 Persiapan Data	42
4.6.2 Menentukan Jumlah Cluster (K)	43
4.6.3 Pelatihan model K-Means	44
4.6.4 Evaluasi K	45
4.7 Inisialisasi Centeroid	47
4.8 Hasil Klasterisasi	48
4.9 Rekomendasi bisnis	50
4.9.1 Distribusi Jumlah Nama Barang per Jenis Barang per Klaster	52

4.9.2 Distribusi Jumlah Nama Barang per Tipe Pembayaran per Klaster	53
4.10 Penyimpanan hasil klaster	55
4.11 Visualisasi Data	55
4.11.1 Heatmap Korelasi Antar Fitur.....	56
4.11.2 Bar Chart Distribusi.....	56
4.12 Dashboard Dewi Sri	57
4.12.1 Fitur Filter Data	59
4.12.2 Output Fitur Filter Data	60
4.12.3 Barchart Top 10 Barang.....	60
4.12.4 Tabel Klaster Laris	61
4.12.5 Tabel Klaster Sedang	62
4.12.6 Tabel klaster rendah.....	63
4.12.7 Tabel Jumlah Nama Barang per tipe barang per Klasterisasi.....	64
4.12.8 Tabel Jumlah Nama Barang per Metode Pembayaran Per Klaster.....	65
4.12.9 Tabel Data Klasterisasi	66
4.12.10 Tabel Profil Klaster	67
BAB V. PENUTUP.....	69
5.1 Kesimpulan.....	69
5.2 Saran.....	70
DAFTAR PUSTAKA.....	71
LAMPIRAN.....	6

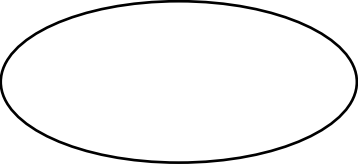
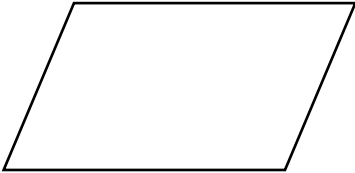
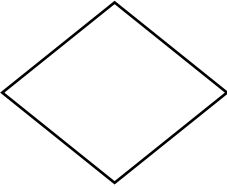
DAFTAR GAMBAR

Gambar 2.1 Tahapan Data Mining	9
Gambar 2.2 Penerapan Aplikasi Python (Suharto, 2023).....	12
Gambar 2.3 Model Konseptual Penelitian	15
Gambar 3.1 Diagram Alur Penelitian.....	21
Gambar 4.1 Struktur organisasi agen dewi sri	32
Gambar 4.2 Source code proses penyesuaian nama kolom	34
Gambar 4.3 Source code data selection	35
Gambar 4.4 Kolom mengandung null value	36
Gambar 4.5 Source code drop column	36
Gambar 4.6 Kolom setelah pre processing.....	37
Gambar 4.7 Source Code Pembersihan Kolom Harga _Jual.....	38
Gambar 4.8 Source code proses agregasi data	40
Gambar 4.9 Source code Membuat atribut menjadi numerik dan pengurangan skewed	41
Gambar 4.10 Source code proses normalisasi data	41
Gambar 4.11 Source Code Data yang akan digunakan	42
Gambar 4.12 Output Penentuan Jumlah Klaster	44
Gambar 4.13 Inisialisasi K Means	44
Gambar 4.14 Evaluasi K	47
Gambar 4.15 Source Code inisialisasi centroid	48
Gambar 4.16 Source code Visualisasi dan outputnya	49
Gambar 4.17 Source code rekomendasi bisnis.....	51
Gambar 4.18 Output rekomendasi bisnis	52
Gambar 4.19 Distribusi per Jenis Barang.....	53
Gambar 4.20 Distribusi per Tipe Pembayaran	54
Gambar 4.21 Source code simpan dataset & profile klaster	55
Gambar 4.22 Heatmap korelasi antar fitur	56
Gambar 4.23 Visualisasi Bar Chart Distribusi	56
Gambar 4.24 Dashboard streamlit dewi sri.....	58
Gambar 4.25 Fitur filter data.....	59
Gambar 4.26 Tabel filter	60
Gambar 4.27 Bar chart 10 teratas barang laris	60
Gambar 4.28 Tabel klaster laris.....	61
Gambar 4.29 Tabel klaster sedang	62
Gambar 4.30 Tabel klaster rendah.....	63
Gambar 4.31 Jumlah Nama Barang per Tipe Barang per Klaster.....	64
Gambar 4.32 Tabel Jumlah Nama Barang per Metode Pembayaran Per Klaster.....	65
Gambar 4.33 Tabel dataset klasterisasi	66
Gambar 4.34 Tabel profil klaster.....	67
Gambar 5.1 Foto wawancara dan identifikasi masalah.....	6
Gambar 5.2 Foto PC penyimpanan transaksi dewi sri	6
Gambar 5.3 Foto wawancara dan pengambilan data.....	7
Gambar 5.4 Foto data transaksi dewi sri	7

DAFTAR TABEL

Tabel 2.1 Ringkasan penelitian terdahulu	16
Tabel 3.1 Informasi Atribut Dataset	24
Tabel 3.2 Jadwal Penelitian	29
Tabel 4.1 Informasi Atribut Dataset	32
Tabel 4.2 Informasi Atribut Dataset	33
Tabel 4.3 Hasil Perhitungan SSE	45
Tabel 4.4 Hasil perhitungan Silhouette Score	46
Tabel 4.5 Score Centroid awal	47
Tabel 4.6 Score Centroid akhir	48

DAFTAR SIMBOL

SIMBOL	KETERANGAN
	Oval: Menyatakan awal atau akhir suatu proses
	Panah: Menunjukkan arah proses dari satu langkah ke langkah berikutnya
	Persegi Panjang: Menyatakan suatu proses atau kegiatan yang dilakukan
	Jajar genjang: Menyatakan proses input (masukan) atau output (keluaran) data
	Belah ketupat: Menyatakan proses pengambilan keputusan

DAFTAR LAMPIRAN

Lampiran 1. Proses wawancara dan identifikasi masalah.	6
Lampiran 2. Dokumentasi PC dewi sri.	6
Lampiran 3. Proses permintaan data kepada anak dari pengelola agen	7
Lampiran 4. Proses pengambilan data.....	7
Lampiran 5. Script wawancara dengan anak dari pengelola agen sembako dewi sri.	7