

**SISTEM INFORMASI ANALISIS KLASIFIKASI PENJUALAN
UMKM BERBASIS WEB DI KELURAHAN TANJUNG REJO
MENGUNAKAN K-MEANS CLUSTERING DAN
VISUALISASI DATA**

SKRIPSI DISUSUN

OLEH:

THAQIF ANDIKA PUTRA
NPM: 2209010233



UMSU

Unggul | Cerdas | Terpercaya

**PROGRAM STUDI SISTEM INFORMASI FAKULTAS
ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
MEDAN**

2026

**SISTEM INFORMASI ANALISIS KLASIFIKASI PENJUALAN UMKM
BERBASIS WEB DI KELURAHAN TANJUNG REJO MENGGUNAKAN
K-MEANS CLUSTERING DAN VISUALISASI DATA**

SKRIPSI

**Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana
Komputer (S.Kom) dalam Program Studi Sistem Informasi pada Fakultas
Ilmu Komputer dan Teknologi Informasi, Universitas Muhammadiyah
Sumatera Utara**

**THAQIF ANDIKA PUTRA
NPM. 2209010233**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA**

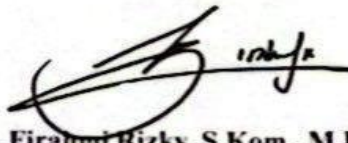
MEDAN

2026

LEMBAR PENGESAHAN

Judul Skripsi : SISTEM INFORMASI ANALISIS KLASIFIKASI
PENJUALAN UMKM BERBASIS WEB DI
KELURAHAN TANJUNG REJO MENGGUNAKAN K-
MEANS CLUSTERING DAN VISUALISASI DATA
Nama Mahasiswa : THAQIF ANDIKA PUTRA
NPM : 2209010233
Program Studi : SISTEM INFORMASI

Menyetujui Komisi Pembimbing



(Dr. Firahmi Rizky, S.Kom., M.Kom)
NIDN. 0116079201

Ketua Program Studi



(Mahardika Abdi Prawira Tanjung, S.Kom., M.Kom)
NIDN. 0117088902

Pekan



(Dr. Al Khowrizmi, S.Kom., M.Kom.)
NIDN. 0127099201

PERNYATAAN ORISINALITAS

**SISTEM INFORMASI ANALISIS KLASIFIKASI PENJUALAN UMKM
BERBASIS WEB DI KELURAHAN TANJUNG REJO MENGGUNAKAN
K-MEANS CLUSTERING DAN VISUALISASI DATA**

SKRIPSI

Saya menyatakan bahwa karya tulis ini adalah hasil karya sendiri, kecuali beberapa kutipan dan ringkasan yang masing-masing disebutkan sumbernya.

Medan, Maret 2026
Yang membuat pernyataan



THAQIF ANDIKA PUTRA
NPM. 2209010233

**PERNYATAAN PERSETUJUAN
PUBLIKASI KARYA ILMIAH UNTUK
KEPENTINGAN AKADEMIS**

Sebagai sivitas akademika Universitas Muhammadiyah Sumatera Utara, saya bertanda tangan dibawah ini:

Nama : Thaqif Andika Putra
NPM : 2209010233
Program Studi : Sistem Informasi
Karya Ilmiah : Skripsi

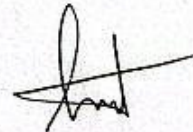
Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Muhammadiyah Sumatera Utara Hak Bebas Royalti Non-Eksekutif (*Non- Exclusive Royalty free Right*) atas penelitian skripsi saya yang berjudul:

**SISTEM INFORMASI ANALISIS KLASIFIKASI PENJUALAN UMKM
BERBASIS WEB DI KELURAHAN TANJUNG REJO MENGGUNAKAN
K-MEANS CLUSTERING DAN VISUALISASI DATA**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non-Eksekutif ini, Universitas Muhammadiyah Sumatera Utara berhak menyimpan, mengalih media, memformat, mengelola dalam bentuk database, merawat dan mempublikasikan Skripsi saya ini tanpa meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemegang dan atau sebagai pemilik hak cipta.

Demikian pernyataan ini dibuat dengan sebenarnya.

Medan, Maret 2026
Yang membuat pernyataan



THAQIF ANDIKA PUTRA
NPM. 2209010233

RIWAYAT HIDUP

DATA PRIBADI

Nama Lengkap : Thaqif Andika Putra Tempat dan
Tanggal Lahir : Medan, 28 Februari 2004
Alamat Rumah : Jalan Murni Gg Setia Kawan No.15D
Telepon/Faks/HP : 085762052176
E-mail : trys.xc21@gmail.com
Instansi Tempat Kerja : -
Alamat Kantor : -

DATA PENDIDIKAN

SD : SD Negeri Percobaan TAMAT: 2016
SMP : SMP Kemala Bhayangkari 1 Medan TAMAT: 2019
SMA : SMA Kemala Bhayangkari 1 Medan TAMAT: 2022

KATA PENGANTAR



Alhamdulillah, puji Syukur penulis panjatkan kehadiran Allah SWT atas segala limpahan taufiq, rahmat hidayah, serta inayah-Nya, sehingga penulis dapat menyusun dan menyelesaikan laporan skripsi dengan judul “Sistem Informasi Analisis Klasifikasi Penjualan UMKM Berbasis Web di Kelurahan Tanjung Rejo Menggunakan K-Means Clustering dan Visualisasi Data”. Skripsi ini disusun untuk memenuhi syarat dalam menyelesaikan studi dan memperoleh gelar sarjana untuk program studi Sistem Informasi Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara.

Selama proses penelitian hingga penyusunan skripsi ini, saya telah memperoleh banyak pelajaran berharga saat menyusun laporan tugas akhir. Demikian pula, berbagai tantangan yang muncul ternyata memberikan pengalaman yang bermanfaat untuk kehidupan mendatang. Semua pencapaian ini tidak mungkin terwujud tanpa kontribusi dari orang-orang di sekitar saya, yang terus memberikan dorongan dan semangat dalam menyelesaikan laporan tugas akhir ini. Melalui momen ini, saya ingin menyampaikan rasa syukur yang mendalam serta penghormatan yang tulus kepada:

1. Bapak Prof. Dr. Akrim, M.Pd., Selaku Rektor Universitas Muhammadiyah Sumatera Utara (UMSU).
2. Bapak Dr. Al-Khowarizmi, S.Kom., M.Kom. Selaku Dekan Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU.
3. Ibu Dr. Firahmi Rizky, S.Kom., M.Kom., selaku Wakil Dekan I Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU, dan Juga

Selaku Dosen Pembimbing yang telah meluangkan waktu, tenaga, dan pikiran dalam memberikan arahan, bimbingan, serta masukan yang sangat berarti selama proses penyusunan skripsi ini hingga selesai dengan baik.

4. Bapak Mhd. Basri, S.Si., M.Kom., selaku Wakil Dekan III Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU.
5. Bapak Mahardika Abdi Prawira Tanjung, S.Kom., M.Kom., Selaku Ketua Program Studi Sistem Informasi yang selalu memberikan dukungan
6. Bapak Mulkan Azhari, S.Kom., M.Kom., selaku Sekretaris Program Studi Sistem Informasi yang telah membantu dan memberikan dukungan selama proses perkuliahan dan penyusunan skripsi.
7. Teruntuk para Bapak/Ibu Dosen FIKTI (UMSU) atas ilmu pengetahuan yang telah diberikan kepada saya selama perkuliahan.
8. Seluruh Staff dan Keanggotaan Biro Kemahasiswaan yang mendukung dalam proses pengerjaan penelitian ini.
9. Kepada kedua orang tua tercinta, Bapak Prof. Dr. Ir. Suwarno. M.T dan Ibu Elva Rosidah yang telah memberikan kasih sayang kepada anaknya tanpa henti dan memberikan dukungan baik emosional maupun materi. Semoga ini menjadi langkah awal untuk membuat papa dan mama bahagia, terimakasih untuk selalu menuliskan namaku disetiap doa yang kalian langitkan, semoga hidup ku kedepan nya sesuai dengan apa yang kalian impi.
10. Teruntuk seluruh saudara saya, Resha Nufitria Adini, Bahana Qolbi Annaufal, Faris Bagaswara, dan Naura Alandra Auni, terimakasih telah menjadi penyemangat dan bagian besar dalam hidup penulis serta menjadi saudara terbaik yang selalu menemani penulis dalam suka

maupun duka.

11. Teruntuk seseorang yang saya temui dari awal kuliah yaitu Elsa Ramadhani, terimakasih telah hadir dan menjadi bagian dari perjalanan hidup penulis. Berkontribusi banyak dalam penulisan karya tulis ini, baik waktu maupun tenaga kepada penulis serta telah mendukung, menghibur, dan mendengarkan keluh kesah, serta memberikan semangat pantang menyerah untuk menyelesaikan studi ini.
12. Teruntuk diri saya sendiri, Thaqif Andika Putra, terimakasih dan apresiasi sebesar- besarnya karena telah bertanggung jawab untuk menyelesaikan apa yang telah dimulai. Terimakasih karena selalu berusaha dan tidak menyerah, serta senantiasa menikmati setiap prosesnya yang bisa dibilang tidak mudah. Semoga bisa menjadi bermanfaat bagi banyak orang dimanapun berada. Dalam menyelesaikan tugas akhir ini, sebagai peneliti, saya sadar bahwa ada banyak kekurangan dan kesalahan yang mungkin terjadi, baik yang sengaja maupun tidak. Oleh sebab itu, saya sangat menghargai setiap saran dan umpan balik yang diberikan, karena itu akan membantu saya berkembang dan menjadi lebih baik di masa depan. Untuk menutup, saya meminta maaf yang tulus atas segala kekurangan, serta berharap bahwa tugas akhir ini dapat memberi nilai tambah dan manfaat bagi pembaca.

SISTEM INFORMASI ANALISIS KLASIFIKASI PENJUALAN UMKM BERBASIS WEB DI KELURAHAN TANJUNG REJO MENGGUNAKAN K-MEANS CLUSTERING DAN VISUALISASI DATA

ABSTRAK

Perkembangan teknologi digital mendorong pemanfaatan data sebagai dasar pengambilan keputusan, termasuk pada sektor Usaha Mikro, Kecil, dan Menengah (UMKM). Namun, sebagian besar UMKM di Kelurahan Tanjung Rejo masih mengalami keterbatasan dalam mengelola dan menganalisis data penjualan secara sistematis sehingga sulit mengidentifikasi performa usaha secara objektif. Penelitian ini bertujuan untuk merancang sistem informasi berbasis web yang mampu melakukan analisis dan klasifikasi penjualan UMKM menggunakan algoritma K-Means Clustering serta menyajikan hasilnya dalam bentuk visualisasi data yang interaktif dan mudah dipahami. Penelitian ini menggunakan pendekatan kuantitatif dengan teknik data mining melalui tahapan Knowledge Discovery in Database (KDD), yang meliputi pengumpulan data, preprocessing (data cleaning dan normalisasi Min-Max), proses clustering menggunakan algoritma K-Means, serta interpretasi hasil. Variabel yang digunakan meliputi omzet bulanan, jumlah transaksi per bulan, jam operasional per hari, lokasi strategis, dan jumlah hari operasional per bulan. Sistem dirancang menggunakan Unified Modeling Language (UML) dan diimplementasikan dengan arsitektur three-tier berbasis web. Hasil yang diharapkan adalah terbentuknya sistem yang mampu mengelompokkan UMKM ke dalam tiga kategori, yaitu belum berkembang, kurang berkembang, dan sudah berkembang, berdasarkan kemiripan karakteristik data penjualan. Sistem ini diharapkan dapat membantu pelaku UMKM dan pihak kelurahan dalam melakukan pemetaan kondisi usaha serta mendukung pengambilan keputusan berbasis data secara lebih efektif dan terarah.

Kata Kunci: Sistem Informasi; UMKM; K-Means Clustering; Data Mining; Visualisasi Data; Web-Based System.

WEB-BASED INFORMATION SYSTEM FOR MSME SALES CLASSIFICATION ANALYSIS IN TANJUNG REJO VILLAGE USING K- MEANS CLUSTERING AND DATA VISUALIZATION

ABSTRACT

The rapid development of digital technology encourages the utilization of data as a basis for decision-making, including in the Micro, Small, and Medium Enterprises (MSMEs) sector. However, most MSMEs in Tanjung Rejo Village still face limitations in managing and analyzing sales data systematically, making it difficult to objectively identify business performance. This study aims to design a web-based information system capable of analyzing and classifying MSME sales performance using the K-Means Clustering algorithm and presenting the results through interactive and easy-to-understand data visualization. This research employs a quantitative approach using data mining techniques through the Knowledge Discovery in Database (KDD) process, which includes data collection, preprocessing (data cleaning and Min-Max normalization), clustering using the K-Means algorithm, and result interpretation. The variables used in this study consist of monthly revenue, number of transactions per month, daily operating hours, business location strategy, and number of operating days per month. The system is designed using Unified Modeling Language (UML) and implemented using a web-based three-tier architecture. The expected outcome of this research is the development of a system capable of grouping MSMEs into three categories: underdeveloped, moderately developed, and well-developed, based on similarities in sales performance characteristics. This system is expected to assist MSME owners and local authorities in mapping business conditions and supporting more effective and data-driven decision-making.

Keywords: Information System; MSMEs; K-Means Clustering; Data Mining; Data Visualization; Web-Based System.

DAFTAR ISI

LEMBAR PENGESAHAN	i
PERNYATAAN ORISINALITAS.....	ii
PERNYATAAN PERSETUJUAN	iii
RIWAYAT HIDUP	iv
KATA PENGANTAR.....	v
ABSTRAK	viii
ABSTRACT	ix
DAFTAR ISI.....	x
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah.....	6
1.3 Batasan Masalah	7
1.4 Tujuan Penelitian	7
1.5 Manfaat Penelitian	8
BAB II LANDASAN TEORI	9
2.1. Sistem Informasi Berbasis Web	9
2.1.1. UMKM.....	11
2.2. Data Mining	13
2.2.1 Tahapan Data Mining.....	13
2.3 Clustering.....	15
2.4 Algoritma K-Means Clustering	15
2.5 Visualisasi Data	16
2.6 Unified Modeling Language (UML)	18
2.6.1 Use Case Diagram.....	18
2.6.2 Activity Diagram.....	19
2.6.3 Class Diagram	19
2.7 Arsitektur Sistem Three-Tier.....	19
2.7.1 SQLite	21

2.8 User Interface (UI) dan User Experience (UX).....	22
BAB III ANALISA DAN PERANCANGAN SISTEM.....	24
3.1 Analisis Permasalahan.....	24
3.2 Jenis dan Sumber Data	25
3.3 Variabel dan Indikator Penelitian	26
3.4 Algoritma Sistem.....	27
3.5 Pemodelan dan Perancangan Sistem	31
BAB IV IMPLEMENTASI DAN PENGUJIAN SISTEM	40
4.1. Kebutuhan Sistem	40
4.1.1. Kebutuhan Perangkat Keras	40
4.1.2. Kebutuhan Perangkat Lunak.....	41
4.1.3. Kebutuhan Pengguna	42
4.2 Implementasi Sistem.....	43
4.2.1 Implementasi Halaman Login.....	43
4.2.2 Implementasi Dashboard.....	44
4.2.3 Implementasi Halaman Data UMKM.....	46
4.2.4 Implementasi Halaman Tambah Data Manual	47
4.2.5 Implementasi Halaman Import Excel	47
4.2.6 Implementasi Halaman Hasil dan Analisis	48
4.3 Pengujian Sistem.....	50
4.3.1 Pengujian Fungsional (Black Box Testing)	50
4.3.2 Pengujian Akurasi Algoritma	51
4.3.3 Pengujian Kinerja Sistem.....	51
4.4 Analisis dan Interpretasi Hasil.....	51
4.4.1 Interpretasi Cluster 1 (UMKM Belum Berkembang).....	52
4.4.2 Interpretasi Cluster 2 (UMKM Kurang Berkembang / Sedang Berkembang)	52
4.4.3 Interpretasi Cluster 3 (UMKM Sudah Berkembang).....	53
4.5 Kekurangan Sistem.....	53
4.5.1 Ketergantungan pada Algoritma Tunggal.....	54
4.5.2 Subjektivitas dalam Penentuan Jumlah Cluster (K).....	54
4.5.3 Mekanisme Pembaruan Analisis Belum Otomatis	54

4.5.4 Cakupan Variabel Analisis yang Terbatas	55
BAB V PENUTUP.....	56
5.1. Kesimpulan.....	56
5.2. Saran	57
DAFTAR PUSTAKA.....	60

DAFTAR TABEL

Tabel 3.1 Variabel Indikator Pengelompokan.....	26
Tabel 3.2. Tabel Data Asli	28
Tabel 3.3. Tabel Hasil Normalisasi	29
Tabel 3.4. Tabel Hasil Iterasi 1	30
Tabel 3.5. Tabel Hasil Normalisasi	31

DAFTAR GAMBAR

Gambar 3.1 Flowchart Sistem.....	33
Gambar 3. 2 Rancangan Use Case Diagram Sistem.....	33
Gambar 3. 3 Rancangan Activity Diagram.....	34
Gambar 3. 4 Rancangan Class Diagram.....	35
Gambar 3. 5 Rancangan Halaman Login	37
Gambar 3. 6 Rancangan Dashboard	37
Gambar 3. 7 Rancangan Data UMKM	38
Gambar 3. 8 Rancangan Tambah Data UMKM	38
Gambar 3. 9 Rancangan Input Data	39
Gambar 3. 10 Rancangan Hasil Analisis.....	39
Gambar 4.1 Tampilan Halaman Login Admin	44
Gambar 4.2 Tampilan Dashboard	45
Gambar 4.3 Tampilan Halaman Data	46
Gambar 4.4 Tampilan Halaman Tambah Data.....	47
Gambar 4.5 Tampilan Import Excel	48
Gambar 4.6 Tampilan Distribusi dan Karakteristik	49
Gambar 4.7 Tampilan Rata Rata Nilai dan Detail Cluster	49

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Perkembangan teknologi digital pada berbagai sektor kehidupan dalam beberapa tahun terakhir telah meningkatkan jumlah data yang dihasilkan secara signifikan, menciptakan fenomena yang dikenal dengan *Big Data* yang ditandai oleh volume data yang besar, keragaman tipe data, dan kecepatan pembaruan yang tinggi (Siregar & Harahap, 2025). Fenomena ini terjadi karena meningkatnya penggunaan aplikasi digital, sistem informasi daring, transaksi elektronik, serta aktivitas operasional modern yang semuanya menghasilkan data dalam jumlah besar setiap harinya (Gemawaty & Yuliani, 2023). *Big Data* kemudian menjadi isu penting bagi organisasi karena metode pengolahan tradisional tidak lagi mampu menangani kompleksitas dan skala data tersebut secara efektif untuk memanfaatkan nilai informasi di dalamnya (Prasetyo et al., 2024).

Dalam konteks tersebut, Data mining hadir sebagai pendekatan ilmu komputer yang bertujuan menemukan pola, struktur, atau informasi bermakna dari kumpulan data besar melalui teknik analisis yang sistematis (Sari, Al-Khowarizmi, & Batubara, 2021). Salah satu teknik penting dalam data mining adalah clustering yang mengelompokkan data berdasarkan kemiripan karakteristik tanpa memerlukan label kelas sebelumnya (Sari et al., 2021). Metode clustering semakin banyak digunakan untuk menganalisis data numerik dalam berbagai domain seperti pendidikan, kesehatan, dan bisnis karena mampu mengungkap struktur tersembunyi dalam data secara eksploratif (IP Sari, Al-Khowarizmi, Sulaiman, & Apdilah, 2023).

Dalam konteks UMKM, data penjualan seperti omzet bulanan, jumlah transaksi, jam operasional, lokasi usaha, dan jumlah hari operasional merupakan indikator kuantitatif yang dapat dianalisis untuk mengukur performa usaha secara objektif (Sari & Apridonal, 2025). Namun, sebagian besar UMKM khususnya UMKM di Kelurahan Tanjung Rejo masih melakukan pencatatan secara manual sehingga sulit dilakukan analisis berbasis komputasi (Haryanti et al., 2024). Kondisi tersebut menyebabkan pelaku UMKM tidak dapat mengidentifikasi posisi perkembangan usahanya secara terukur dan berbasis data (Marcelina et al., 2023). bersifat umum, kurang terarah, dan seringkali tidak efektif.

Namun, permasalahan utama pada UMKM di Kelurahan Tanjung Rejo bukan terletak pada ketersediaan data, melainkan pada keterbatasan metode analisis yang digunakan untuk mengolah data tersebut menjadi informasi bermakna (Hartomi et al., 2022). Oleh karena itu, diperlukan metode analitik yang mampu mengelompokkan data penjualan UMKM berdasarkan kemiripan karakteristik sehingga terbentuk segmentasi performa usaha yang objektif (Bagaskoro et al., 2025).

Untuk menjawab permasalahan tersebut metode yang paling relevan dan banyak digunakan dalam pengelompokan data numerik adalah algoritma K-Means Clustering (Chong, 2021). Algoritma K-Means merupakan metode unsupervised learning yang bekerja dengan mempartisi data ke dalam sejumlah klaster berdasarkan kedekatan jarak terhadap titik pusat (centroid) menggunakan perhitungan Euclidean Distance (Ikotun et al., 2022). K-Means dinilai efektif karena meminimalkan variasi dalam cluster dan memaksimalkan perbedaan antar cluster sehingga menghasilkan pengelompokan yang optimal (Setyaningtyas et al.,

2022). Dalam penelitian terkait UMKM, K-Means terbukti mampu mengidentifikasi segmentasi performa usaha secara lebih objektif dibandingkan evaluasi manual (Wijaya & Daulay, 2025).

Keunggulan utama K-Means terletak pada kesederhanaan algoritma, efisiensi komputasi, serta kemampuannya menangani data numerik dalam jumlah besar secara cepat (Ikotun et al., 2022). Selain itu, metode ini sangat sesuai dengan pendekatan Knowledge Discovery in Database (KDD) yang mencakup tahapan data cleaning, normalisasi Min-Max, proses clustering, hingga interpretasi hasil sebagaimana diterapkan dalam penelitian ini (Laurenso et al., 2024). Dengan demikian, penerapan K-Means dalam klasifikasi penjualan UMKM menjadi solusi yang rasional dan metodologis.

Namun penerapan K-Means masih memiliki beberapa keterbatasan seperti sensitivitas terhadap pemilihan jumlah cluster yang dapat memengaruhi kualitas hasil pengelompokan (Ikotun et al., 2022). Perbedaan skala antar variabel juga dapat menyebabkan ketidakakuratan dalam perhitungan jarak jika tidak dilakukan proses normalisasi terlebih dahulu (Laurenso et al., 2024). Oleh karena itu, tahapan preprocessing seperti data cleaning dan normalisasi Min-Max Scaling menjadi langkah penting sebelum proses clustering dilakukan (Setyaningtyas et al., 2022). Dengan penerapan tahapan tersebut, proses clustering menggunakan algoritma K-Means diharapkan mampu menghasilkan pengelompokan data yang lebih akurat dan representatif sehingga dapat memberikan dasar analisis yang lebih objektif dalam mengevaluasi performa UMKM (Laurenso et al., 2024). Selain itu, hasil pengelompokan tersebut juga dapat membantu pihak terkait dalam memahami pola perkembangan usaha serta mendukung proses pengambilan

keputusan yang lebih tepat dalam pengembangan UMKM (Setyaningtyas et al., 2022).

Penelitian terdahulu oleh Manurung (2024) dengan judul “Pengelompokan UMKM Kota Binjai Menggunakan Metode *Clustering K-Means* untuk Mengidentifikasi Pola Perkembangan Bisnis”. Penelitian ini lebih berfokus pada segmentasi UMKM (jenis usaha) dan pola perkembangan secara agregat, tetapi tidak menganalisis data penjualan kuantitatif (misalnya volume penjualan per produk) atau menampilkan visualisasi interaktif berbasis sistem informasi.

Penelitian lain oleh Bagaskoro et al. (2025) dengan judul “Optimalisasi Pemetaan UMKM di Wilayah DKI Jakarta melalui *K-Means Clustering dan PCA*”. Penelitian ini lebih berfokus pada pemetaan geografis (clustering wilayah), bukan analisis performa penjualan produk per UMKM. Visualisasi cenderung peta geografis dan kluster wilayah, bukan *dashboard* penjualan dengan *insight* per produk.

Selanjutnya penelitian oleh Maulidya et al. (2025) dengan judul “Integrating *K-Means Clustering dan Apriori* untuk Strategi Pemasaran Digital UMKM: Studi Kasus Kota Stabat”. Penelitian ini memiliki kekurangan seperti; meskipun ada analisis segmen pelanggan, penelitian ini difokuskan pada data transaksi (pelanggan) bukan analisis penjualan produk UMKM secara keseluruhan atau pengelompokan produk berdasarkan tren penjualan. Sistem yang dikembangkan lebih bersifat analitik (strategi pemasaran), bukan sistem informasi web “*dashboard*” untuk pelaku UMKM.

Berdasarkan penelitian terdahulu diatas, ditemukan beberapa research gap

yang dapat diidentifikasi seperti, minimnya studi *clustering* penjualan produk UMKM di skala Kelurahan (lokal) dibanding klaster wilayah atau jenis usaha, kurangnya sistem informasi berbasis web yang menggabungkan *clustering* dan visualisasi untuk UMKM agar analisis menjadi *actionable*, serta keterbatasan dalam *dashboard* atau antarmuka visual yang ramah bagi pemilik UMKM non-teknis. Menjawab kesenjangan tersebut, penelitian ini dikembangkan untuk menghasilkan sistem informasi analisis klasifikasi penjualan UMKM berbasis web yang menerapkan algoritma K-Means Clustering dengan visualisasi dashboard yang mudah dipahami.

Penelitian ini menggunakan lima variabel utama yaitu omzet bulanan, jumlah transaksi per bulan, jam buka per hari, lokasi strategis, dan hari operasional per bulan. (Sari & Apridonol, 2025), (Manurung, 2024), (Bagaskoro et al., 2025), (Adek Maulidya et al., 2024). Kemudian kelima variabel tersebut akan diproses menggunakan K-Means Clustering untuk dikelompokkan kedalam 3 klaster yaitu, Klaster 1 UMKM belum berkembang, Klaster 2 UMKM kurang berkembang, Klaster 3 UMKM sudah berkembang. Tiap klaster menggambarkan kondisi UMKM tersebut berdasarkan kelima variabel utama dalam penelitian ini.

Penelitian ini bertujuan untuk membangun sebuah sistem informasi berbasis web yang mampu menganalisis dan mengklasifikasikan performa penjualan UMKM secara akurat berdasarkan lima variabel penentu tersebut. Secara teoritis, penelitian ini memberikan kontribusi penting dalam pengembangan *data mining* dan penerapannya pada sektor UMKM. Secara praktis, hasil penelitian ini diharapkan dapat membantu pemerintah kelurahan, pelaku UMKM, dan pihak terkait lainnya dalam memetakan kondisi usaha,

memprioritaskan program pembinaan, serta merumuskan strategi pengembangan yang efektif dan tepat sasaran.

Berdasarkan latar belakang diatas, maka penulis mengangkat sebuah topik yang menghasilkan judul “Sistem Informasi Analisis Klasifikasi Penjualan UMKM Berbasis Web di Kelurahan Tanjung Rejo Menggunakan K-Means Clustering dan Visualisasi data”.

1.2 Rumusan Masalah

Berdasarkan latar belakang tersebut, maka rumusan masalah penelitian ini yaitu:

1. Bagaimana menganalisis kebutuhan sistem informasi yang diperlukan untuk mengelola dan mengolah data penjualan UMKM di Kelurahan Tanjung Rejo?
2. Bagaimana merancang sistem informasi berbasis web yang mampu melakukan klasifikasi penjualan UMKM menggunakan *algoritma K-Means Clustering* berdasarkan lima variabel utama (omzet bulanan, jumlah transaksi, jam operasional per hari, lokasi usaha, dan jumlah hari operasional per bulan)?
3. Bagaimana menguji sistem yang menggunakan algoritma *K-Means* dan memastikan bahwa hasil klasifikasi penjualan dapat ditampilkan secara akurat dalam bentuk visualisasi data yang mudah dipahami?
4. Bagaimana mengimplementasikan sistem informasi berbasis web yang terintegrasi dengan proses *clustering* sehingga dapat digunakan oleh pelaku UMKM maupun pihak kelurahan?

1.3 Batasan Masalah

Batasan masalah dalam penelitian ini adalah:

1. Data penelitian diperoleh melalui kuesioner yang diberikan kepada pelaku UMKM di Kelurahan Tanjung Rejo.
2. Variabel yang digunakan terbatas pada lima indikator penjualan, yaitu omzet bulanan, jumlah transaksi, jam buka per hari, lokasi usaha, dan hari operasional per bulan.
3. Sistem yang dibangun hanya mencakup fitur input data, pengolahan clustering, manajemen data UMKM, serta visualisasi cluster dalam bentuk grafik dan ringkasan analisis.
4. Sistem dikembangkan berbasis web menggunakan Python dan SQLite dengan visualisasi dashboard interaktif.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah:

1. Menganalisis kebutuhan sistem informasi yang diperlukan untuk memetakan performa penjualan UMKM di Kelurahan Tanjung Rejo.
2. Merancang sistem informasi berbasis web yang mampu mengolah data penjualan UMKM dan melakukan klasifikasi menggunakan algoritma *K-Means Clustering*.
3. Menguji akurasi dan kinerja metode *K-Means* dalam mengelompokkan penjualan UMKM berdasarkan kelima variabel utama pada penelitian ini.
4. Mengimplementasikan sistem informasi berbasis web yang menampilkan hasil klasifikasi penjualan dalam bentuk visualisasi data interaktif sehingga mudah dipahami oleh pengguna.

1.5 Manfaat Penelitian

Berikut merupakan manfaat yang diharapkan pada penelitian yang akan dilakukan:

1. Manfaat bagi kampus:
 - a. Menjadi referensi akademik bagi mahasiswa lain yang ingin mengembangkan penelitian pada bidang UMKM, *data mining*, atau sistem informasi.
 - b. Meningkatkan kontribusi akademik kampus dalam pengembangan teknologi untuk pemberdayaan UMKM.
2. Manfaat bagi penulis:
 - a. Peneliti dapat meningkatkan keterampilan dalam analisis data, penerapan algoritma *data mining*, dan pengembangan sistem informasi berbasis web.
 - b. Peneliti mendapatkan pengalaman praktis dalam merancang sistem yang dapat digunakan untuk pengambilan keputusan di dunia nyata.
3. Manfaat bagi UMKM :
 - a. Membantu pelaku UMKM di Kelurahan Tanjung Rejo dalam memahami posisi penjualan mereka berdasarkan klasifikasi klaster yang terukur dan berbasis data.
 - b. Menjadi dasar bagi UMKM di Kelurahan Tanjung Rejo untuk menentukan strategi promosi, perbaikan layanan, dan pengembangan usaha yang lebih tepat sasaran.
 - c. Membantu pihak kelurahan dalam memetakan kondisi UMKM sehingga program pembinaan dapat dilakukan secara lebih efektif.

BAB II

LANDASAN TEORI

2.1. Sistem Informasi Berbasis Web

Sistem informasi adalah sekumpulan komponen yang saling terkait untuk mengumpulkan, memproses, menyimpan, dan mendistribusikan informasi untuk mendukung pengambilan keputusan dan kontrol dalam suatu organisasi . Pengembangan sistem informasi berbasis *web* memungkinkan akses data dan analisis secara *real-time* dari mana saja, yang penting untuk mendukung operasional UMKM (Arla & Sutabri, 2024). Arsitektur tiga lapis (*three-tier architecture*) merupakan kerangka umum dalam perancangan sistem informasi modern karena menawarkan skalabilitas dan pemisahan fungsionalitas yang jelas antara antarmuka, logika bisnis, dan basis data (Abdillah et al., 2024). Lapisan presentasi (*front-end*) berperan penting dalam menyajikan hasil analisis secara intuitif dan ramah pengguna, terutama bagi pemilik UMKM non-teknis (Putra, R. A., 2020). Untuk analisis penjualan UMKM, sistem berbasis web memungkinkan pemilik UMKM mengakses dan menganalisis data penjualan mereka kapan saja dan di mana saja, tanpa terikat pada perangkat atau lokasi tertentu (Ternando & Mulyono, 2022). Karakteristik penting sistem informasi berbasis web untuk analisis penjualan:

- a. Aksesibilitas Tinggi: Dapat diakses melalui berbagai perangkat dengan koneksi internet (Panjaitan & Pakpahan, 2022).
- b. Interaktivitas: Memungkinkan pengguna untuk berinteraksi dengan data dan visualisasi secara dinamis (Lestari & Henderi, 2021).
- c. Skalabilitas: Mampu menangani peningkatan jumlah data dan pengguna

seiring pertumbuhan UMKM.

- d. Integrasi Fitur Analisis: Kemampuan untuk mengintegrasikan algoritma *data mining* (seperti K-Means) dan alat visualisasi data untuk menghasilkan wawasan (Lestari & Henderi, 2021).

Dalam konteks UMKM, integrasi analisis dan *dashboard* dalam sistem informasi penjualan menjadi krusial. *Dashboard* menyediakan tampilan visual yang ringkas dan interaktif tentang metrik penjualan kunci, memungkinkan pemilik UMKM memantau kinerja secara *real-time* (Lestari & Henderi, 2021). Integrasi dengan modul analisis, seperti *K-Means Clustering*, memungkinkan UMKM untuk tidak hanya melihat data, tetapi juga memahami pola tersembunyi, mengidentifikasi segmen pelanggan, dan merumuskan strategi penjualan yang lebih efektif (Lestari & Henderi, 2021). Kebutuhan ini muncul karena pemilik UMKM seringkali memiliki keterbatasan waktu dan sumber daya untuk melakukan analisis data yang mendalam secara manual (Prabowo et al., 2023). Sistem terintegrasi memberikan kemudahan akses terhadap informasi yang dapat ditindaklanjuti (Prabowo et al., 2023).

Penelitian oleh Kuswidyawan et al. (2023) yang berjudul “Customer Segmentation with K-Means Clustering Algorithm for E-Commerce Sales Visualization” membuktikan bahwa penerapan sistem berbasis web yang dikombinasikan dengan algoritma K-Means mampu meningkatkan pemahaman pengguna terhadap pola penjualan melalui visualisasi data. Namun, penelitian tersebut masih berfokus pada segmentasi pelanggan e-commerce dan belum mengarah pada klasifikasi performa penjualan

2.1.1. UMKM

Usaha Mikro, Kecil, dan Menengah adalah entitas bisnis yang memiliki peran vital dalam perekonomian suatu negara, termasuk Indonesia. Menurut Undang-Undang Nomor 20 Tahun 2008, UMKM dikategorikan berdasarkan kriteria aset dan omzet tahunan, di mana usaha mikro memiliki kekayaan bersih paling banyak Rp 50 juta dan omzet tahunan paling banyak Rp 300 juta, usaha kecil memiliki kekayaan bersih antara Rp 50 juta hingga Rp 500 juta dan omzet tahunan antara Rp 300 juta hingga Rp 2,5 miliar, dan usaha menengah memiliki kekayaan bersih antara Rp 500 juta hingga Rp 10 miliar dan omzet tahunan antara Rp 2,5 miliar hingga Rp 50 miliar (Amalia et al., 2023).

UMKM seringkali menghadapi berbagai tantangan dalam pengelolaan data penjualan, yang dapat menghambat pertumbuhan dan pengembangan usaha mereka. Tantangan-tantangan ini antara lain (Santi et al., 2024):

1. Keterbatasan Teknologi dan Infrastruktur: Banyak UMKM belum memiliki akses atau kemampuan untuk memanfaatkan teknologi digital secara optimal dalam pencatatan penjualan (Fauzan et al., 2022).
2. Kurangnya Pengetahuan dan Keterampilan Digital: Pemilik UMKM mungkin kurang familiar dengan *software* atau sistem untuk mengelola data penjualan, sehingga pencatatan masih manual atau tidak terstruktur (Santi et al., 2024).
3. Volume Data yang Besar dan Tidak Terorganisir: Seiring dengan pertumbuhan usaha, data penjualan dapat menjadi sangat besar dan sulit dikelola jika tidak ada sistem yang memadai (Firdhaus & Akbar, 2022).

4. Keterbatasan Anggaran: Implementasi sistem informasi penjualan yang canggih seringkali memerlukan investasi yang besar, yang mungkin tidak terjangkau bagi sebagian UMKM (Tjandra, 2022).
5. Keamanan Data: Masalah keamanan dan privasi data juga menjadi perhatian, terutama bagi UMKM yang beroperasi secara *online* (Prasetya & Utomo, 2023).

Klasifikasi performa UMKM penting untuk mengidentifikasi UMKM dengan kinerja baik, sedang, atau kurang, sehingga intervensi atau strategi yang tepat dapat diterapkan (Marcelina et al., 2023). Dengan mengelompokkan UMKM berdasarkan performa penjualan, pemerintah, lembaga keuangan, atau pemangku kepentingan lainnya dapat memberikan dukungan yang lebih terarah, seperti pelatihan, akses modal, atau program pemasaran (Puntoriza & Fibriani, 2020). Bagi UMKM itu sendiri, klasifikasi ini membantu mereka memahami posisi relatif mereka di pasar dan mengidentifikasi area yang memerlukan perbaikan (Syifa & Fahmi, 2021). Klasifikasi performa juga dapat menjadi dasar untuk segmentasi strategi pemasaran yang lebih efektif (Ardana et al., 2024).

Penelitian sebelumnya yang dilakukan oleh Rahmawati et al. (2025) dengan judul “Optimasi stok dengan Clustering Data Transaksi Penjualan Menggunakan Algoritma K-Means di Konter Agung Cell”. Penelitian ini menggunakan K-Means Clustering untuk mengelompokkan produk berdasarkan tingkat penjualan guna mendukung pengambilan keputusan stok. Hasil penelitian ini membuktikan bahwa K-Means efektif dalam mengelompokkan produk berdasarkan pola transaksi. Namun, penelitian ini belum mengembangkan sistem informasi berbasis web yang terintegrasi dengan visualisasi dashboard analitik.

2.2. Data Mining

Data mining adalah proses penemuan pola, hubungan, anomali, dan struktur yang valid dan potensial berguna dalam kumpulan data yang besar (Hartomi et al., 2022). Tujuannya adalah untuk mengubah data mentah menjadi informasi yang dapat ditindaklanjuti untuk mendukung pengambilan keputusan (Hartomi et al., 2022). Tujuan utama data mining meliputi:

- a. Deskripsi: Menggambarkan pola dan karakteristik data yang sudah ada, misalnya, karakteristik rata-rata pelanggan atau penjualan UMKM.
- b. Prediksi: Meramalkan nilai atau kecenderungan masa depan berdasarkan pola historis, contohnya memprediksi omzet penjualan bulan depan (Hartomi et al., 2022).
- c. Klasifikasi: Mengkategorikan data ke dalam kelas-kelas yang telah ditentukan, seperti mengklasifikasikan UMKM ke dalam kategori berkinerja tinggi atau rendah (Nuryaman et al., 2022).
- d. Clustering: Mengelompokkan data yang memiliki kemiripan karakteristik ke dalam segmen-segmen, misalnya mengelompokkan UMKM dengan profil penjualan serupa (Hartomi et al., 2022).
- e. Asosiasi: Mengidentifikasi hubungan atau korelasi antar item data, seperti mengetahui produk apa yang sering dibeli bersamaan (Hartomi et al., 2022).

2.2.1 Tahapan Data Mining

Cross-Industry Standard Process for Data Mining adalah metodologi yang paling banyak digunakan dalam proyek data mining, terdiri dari enam tahapan yang saling berhubungan (Maskanah, 2020).

- a. Business Understanding: Tahap awal untuk memahami tujuan bisnis atau penelitian dari perspektif bisnis. Ini melibatkan identifikasi masalah, tujuan, dan kriteria keberhasilan proyek (Laurenso et al., 2024).
- b. Data Understanding: Pengumpulan data awal dan pemeriksaan kualitas data untuk mengidentifikasi masalah, memahami struktur data, dan melakukan eksplorasi data awal. Ini juga melibatkan verifikasi relevansi data dengan tujuan bisnis (Laurenso et al., 2024).
- c. Data Preparation: Tahap paling memakan waktu, melibatkan pembersihan data (menangani nilai hilang, *outlier*), transformasi data (normalisasi, agregasi), pemilihan fitur yang relevan, dan integrasi data dari berbagai sumber (Maskanah, 2020).
- d. Modeling: Pemilihan dan penerapan teknik *data mining* yang sesuai (misalnya, K-Means Clustering) untuk membangun model analitik. Tahap ini juga mencakup pemilihan parameter model dan validasi awal (Maskanah, 2020).
- e. Evaluation: Penilaian terhadap model yang telah dibangun untuk memastikan validitas, keandalan, dan relevansinya dalam mencapai tujuan bisnis yang telah ditetapkan. Jika model belum memenuhi kriteria, proses mungkin kembali ke tahap sebelumnya (Maskanah, 2020).
- f. Deployment: Penerapan model yang telah dievaluasi ke dalam lingkungan operasional untuk menghasilkan manfaat nyata. Ini bisa berupa implementasi sistem, pembuatan laporan, atau integrasi dengan aplikasi lain (Maskanah, 2020).

2.3 Clustering

Clustering merupakan teknik data mining yang mengelompokkan data berdasarkan tingkat kemiripan karakteristik tertentu (Sari & Nugroho, 2021). Metode clustering bersifat *unsupervised learning* karena tidak membutuhkan label kelas sebelumnya (Manurung, 2024). Clustering sangat sesuai digunakan untuk mengelompokkan UMKM berdasarkan performa penjualan secara objektif (Bagaskoro et al., 2025).

2.4 Algoritma K-Means Clustering

Algoritma K-Means adalah metode *clustering* non-hierarki yang paling populer dan sering digunakan untuk mengelompokkan data ke dalam sejumlah *cluster* (Chong, 2021). Tujuan utama K-Means adalah mempartisi observasi ke dalam *cluster*, di mana setiap observasi termasuk ke *cluster* dengan *mean* terdekat (Chong, 2021). Algoritma K-Means Clustering merupakan metode *unsupervised learning* dalam bidang data mining yang bertujuan untuk mengelompokkan data ke dalam sejumlah (k) kluster berdasarkan kesamaan karakteristik (*similarity features*). Setiap kluster memiliki pusat data (*centroid*) yang mewakili rata-rata dari titik-titik dalam kluster tersebut. Tujuan utama K-Means adalah meminimalkan variansi intra-kluster dan memaksimalkan variansi antar-kluster agar hasil pengelompokan menjadi optimal.

Algoritma K-Means adalah salah satu algoritma *clustering* yang paling populer dan banyak digunakan karena kesederhanaan dan efisiensinya (Chong, 2021). K-Means merupakan metode *partitioning clustering* yang mengelompokkan data ke dalam *cluster* yang telah ditentukan sebelumnya (Setyaningtyas et al.,

2022). Tujuannya adalah meminimalkan jarak antara titik data dan *centroid cluster*-nya, sekaligus memaksimalkan jarak antar *centroid* Ikotun et al., 2022. K- Means umumnya menggunakan *Euclidean Distance* sebagai metrik untuk mengukur kedekatan antar titik data dan *centroid* Sari et al., 2023.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Gambar 2.1 Rumus *Euclidean Distance*.

2.5 Visualisasi Data

Visualisasi data merupakan bentuk penyajian informasi dan data secara grafis dengan memanfaatkan elemen visual seperti grafik, diagram, dan peta guna memudahkan pemahaman terhadap tren, pola, serta anomali yang terdapat dalam kumpulan data yang bersifat kompleks (Balan, 2024). Barrocas et al. (2022) menjelaskan bahwa visualisasi data berperan dalam mentransformasikan data numerik ke dalam bentuk visual yang lebih mudah dipersepsikan oleh indera manusia, sehingga memungkinkan proses interpretasi hasil analisis dilakukan secara cepat, akurat, dan mendalam.

Dalam konteks Usaha Mikro, Kecil, dan Menengah (UMKM), visualisasi data tidak hanya berfungsi sebagai elemen estetika tampilan sistem, tetapi juga sebagai instrumen strategis untuk memahami kinerja usaha, pola penjualan, serta efektivitas kegiatan promosi. Visualisasi data memiliki beberapa fungsi utama, antara lain sebagai sarana eksplorasi data, media komunikasi informasi yang efektif, alat identifikasi pola dan kecenderungan, validasi hipotesis, serta pendukung pengambilan keputusan yang lebih cepat dan tepat. Bagi pelaku UMKM, visualisasi data memberikan manfaat signifikan dalam membantu

pemahaman intuitif terhadap performa penjualan, mendeteksi peluang pasar potensial, serta menyajikan data bisnis secara informatif dan menarik kepada pihak investor maupun pemerintah daerah (Makhnun & Agussalim, 2023).

Jenis visualisasi yang umum digunakan dalam analisis penjualan meliputi bar chart untuk membandingkan kategori omzet dan jumlah transaksi, line chart dan area chart untuk menampilkan tren penjualan berdasarkan waktu, radar chart untuk membandingkan kinerja antar variabel seperti omzet, jumlah pelanggan, dan intensitas promosi, pie chart untuk menunjukkan proporsi penjualan setiap produk, serta scatter plot untuk menggambarkan hubungan atau korelasi antar variabel, misalnya antara frekuensi promosi dan peningkatan omzet (Addepalli et al., 2023). Penerapan visualisasi data pada sistem informasi berbasis web dilakukan dengan memanfaatkan teknologi modern, seperti Chart.js sebagai library utama dalam penyajian grafik interaktif (Dogadina & Voronin, 2024), Python dan framework Flask sebagai backend sistem untuk pengelolaan data UMKM secara terstruktur (Panjaitan & Pakpahan, 2022), serta JavaScript yang berfungsi menghubungkan data dari server dan menampilkannya secara real-time pada antarmuka pengguna (Arwani et al., 2022). Integrasi teknologi tersebut memungkinkan sistem informasi yang dikembangkan bersifat dinamis, interaktif, dan mampu menyajikan hasil analisis K-Means Clustering dalam bentuk visual yang mudah dipahami oleh pengguna dengan latar belakang yang beragam.

Penelitian sebelumnya yang dilakukan oleh Maurits dan Ghania (2024) berjudul “Visualisasi Data Usaha Kecil dan Menengah Provinsi DKI Jakarta Menggunakan Database” menekankan pentingnya dashboard visualisasi data dalam membantu pemangku kepentingan memahami kondisi UMKM.

Penelitian ini berhasil menyajikan data UMKM secara informatif melalui dashboard visual. Akan tetapi, penelitian tersebut belum mengintegrasikan metode data mining seperti K-Means Clustering untuk klasifikasi performa penjualan UMKM.

2.6 *Unified Modeling Language (UML)*

Unified Modeling Language (UML) merupakan bahasa pemodelan standar yang digunakan untuk memvisualisasikan, merancang, dan mendokumentasikan sistem informasi secara sistematis (Putra, 2020). UML menyediakan notasi grafis yang mampu mempresentasikan struktur dan perilaku sistem perangkat lunak secara terintegrasi (Yusuf & Sitorus, 2021). Penggunaan UML membantu pengembang dalam memahami kebutuhan sistem sebelum proses implementasi dilakukan (Prasetyo et al., 2022). Dalam pengembangan sistem informasi berbasis web, UML berperan sebagai alat komunikasi antara pengembang dan pemangku kepentingan (Hartana & Wijaya, 2024). UML juga membantu mengurangi kesalahan perancangan dengan mendefinisikan model sistem secara jelas sejak tahap awal pengembangan (Triatama et al., 2025).

2.6.1 Use Case Diagram

Use case diagram merupakan salah satu diagram UML yang digunakan untuk menggambarkan interaksi antara aktor dan sistem melalui fungsi-fungsi yang disediakan sistem (Putra, 2020). Aktor pada use case diagram mempresentasikan pengguna atau entitas eksternal yang berinteraksi langsung dengan sistem (Yusuf & Sitorus, 2021). Use case diagram berfokus pada kebutuhan fungsional sistem dari sudut pandang pengguna tanpa menampilkan detail teknis implementasi (Prasetyo et al., 2022). Diagram ini membantu mengidentifikasi peran

pengguna serta batasan ruang lingkup sistem yang dikembangkan (Hartana & Wijaya, 2024). Dalam sistem informasi berbasis web, use case diagram memudahkan pemetaan fungsi utama seperti pengelolaan data, proses analisis, dan penyajian informasi (Triatama et al., 2025).

2.6.2 Activity Diagram

Activity diagram merupakan diagram UML yang digunakan untuk menggambarkan alur aktivitas atau proses bisnis yang terjadi dalam sistem secara berurutan (Yusuf & Sitorus, 2021). Diagram ini menunjukkan urutan aktivitas dari awal hingga akhir serta kemungkinan percabangan proses yang terjadi (Prasetyo et al., 2022). Activity diagram sangat efektif untuk memodelkan proses dinamis seperti pengolahan data dan eksekusi algoritma dalam sistem informasi (Triatama et al., 2025). Dalam sistem informasi berbasis web, activity diagram digunakan untuk menggambarkan alur proses mulai input data hingga penyajian output kepada pengguna (Hartana & Wijaya, 2024). Activity diagram membantu memastikan bahwa setiap proses sistem berjalan sesuai dengan logika dan kebutuhan yang telah dirancang (Putra, 2020).

2.6.3 Class Diagram

Class Diagram merupakan model data yang digunakan untuk memodelkan struktur data dan objek dalam sistem dengan menampilkan kelas beserta relasi antar kelas tersebut (Ian Sommerville, 2021). Class diagram juga menggambarkan entitas-entitas utama dan hubungan antar entitas yang menjadi acuan dalam implementasi perangkat lunak (Walayah et al., 2025).

2.7 Arsitektur Sistem Three-Tier

Arsitektur *three-tier* membagi sistem informasi menjadi tiga lapisan logis:

Presentation Layer, Application Layer, dan Data Layer (Hartana & Wijaya, 2024). Lapisan presentasi (*front-end*) bertanggung jawab atas antarmuka pengguna, menampilkan *dashboard* dan grafik interaktif menggunakan teknologi seperti HTML, CSS, dan Chart.js (Gemawaty & Yuliani, 2023). Lapisan aplikasi (*back-end*) menjalankan logika bisnis, termasuk eksekusi algoritma *K-Means Clustering* menggunakan teknologi seperti PHP dan Python. (Gemawaty & Yuliani, 2023). Lapisan data bertanggung jawab terhadap penyimpanan dan pengelolaan data UMKM serta hasil kluster. Pada penelitian ini, lapisan data diimplementasikan menggunakan sistem manajemen basis data relasional (*Relational Database Management System/RDBMS*) berbasis file, yaitu SQLite, yang terintegrasi langsung dengan aplikasi melalui library `sqlite3` pada Python. Python merupakan bahasa pemrograman yang dirancang untuk mendukung pengembangan perangkat lunak secara cepat dan efisien (Putra, 2020). Python memiliki sintaks yang sederhana dan mudah dipahami sehingga banyak digunakan dalam pengembangan sistem informasi berbasis web (Yusuf & Sitorus, 2021). Bahasa python mendukung paradigma pemrograman prosedural, objektif, dan fungsional yang memberikan fleksibilitas dalam pengembangan sistem (Prasetyo et al., 2022). Python banyak digunakan dalam bidang data mining dan analisis data karena memiliki pustaka pendukung yang lengkap seperti Numpy, Pandas, dan Scikit-learn (Saputri, 2025). Penggunaan python dalam implementasi algoritma k-means clustering dinilai efektif karena mampu menangani pengolahan data numerik secara efisien (Ikotun et al., 2022). Dalam pengembangan sistem informasi berbasis web, Python sering digunakan pada sisi server untuk mengelola logika bisnis dan proses analisis data (Hartana & Wijaya, 2024). Penggunaan

python dalam sistem analisis penjualan UMKM memungkinkan proses pengolahan data, clustering, dan visualisasi dilakukan secara terintegrasi dalam satu sistem (Revi Setiawan et al., 2025).

2.7.1 SQLite

SQLite adalah sistem manajemen basis data relasional yang dikenal karena arsitekturnya yang *self-contained*, *serverless*, dan *zero-configuration*, dirancang untuk penyimpanan dan pengelolaan data terstruktur (Lang et al., 2025; Yu, 2020; Zakaria et al., 2024). Berbeda dengan sistem basis data client-server, SQLite beroperasi langsung dalam aplikasi, berinteraksi dengan file disk lokal untuk mengelola data (Lang et al., 2025; Suraya & Sholeh, 2021). Sistem ini mendukung *Structured Query Language* untuk operasi data, termasuk penyimpanan, pengambilan, dan manipulasi, serta sebagian besar mematuhi standar SQL-92 (Lang et al., 2025; Yu, 2020). Karena sifatnya yang *open-source*, stabil, dan efisien, SQLite merupakan pilihan populer untuk mengembangkan berbagai aplikasi, terutama untuk proyek berskala kecil hingga menengah, *rapid prototyping*, serta sistem seluler dan tersemat (*embedded systems*) (Lang et al., 2025; Pavlova & Tikhonovsky, 2020; Saxe et al., 2025; Yu, 2020). Desainnya, di mana seluruh basis data berada dalam satu file yang portabel, menyederhanakan pengelolaan basis data, pencadangan, dan transfer (Yu, 2020; Zakaria et al., 2024). Data di SQLite disimpan dalam tabel-tabel relasional, dan dukungannya terhadap transaksi ACID menjamin integritas dan keandalan data yang kuat, bahkan jika terjadi kegagalan sistem (Lang et al., 2025). Sistem ini menjaga konsistensi data melalui penerapan *primary key* dan *foreign key*, yang menegakkan integritas referensial antar tabel (Soto, 2025; Weidmann, 2023; Yu,

2020; Zhao et al., 2024). Untuk sistem informasi, seperti aplikasi analisis penjualan UMKM, SQLite mampu menyimpan data penting, termasuk profil UMKM, catatan penjualan, informasi pengguna, dan hasil klasifikasi kluster (Lang et al., 2025; Saxe et al., 2025; Suraya & Sholeh, 2021). Sebagai RDBMS yang mendukung *query* SQL, SQLite dapat memfasilitasi proses pengambilan dan persiapan data untuk algoritma analitik seperti K-Means Clustering (Esmailpour & Sintos, 2024; Villarroya & Baumann, 2022). Algoritma K-Means, yang sering digunakan untuk mempartisi kumpulan data ke dalam kelompok yang bermakna, dapat diintegrasikan secara efektif dengan basis data relasional untuk menganalisis data dan menemukan pola tersembunyi (Amer, 2020; Khamis et al., 2020; Olteanu, 2020).

2.8 User Interface (UI) dan User Experience (UX)

User Interface (UI) merujuk pada tampilan visual sistem, termasuk elemen interaktif, tata letak, dan tipografi, yang digunakan pengguna untuk berinteraksi dengan perangkat lunak (Pamungkas, 2024). Tujuan utama perancangan UI adalah menciptakan tampilan yang intuitif, menarik, dan sesuai dengan identitas visual sistem (Nifail Zazhemi & Marcos, 2025).

User Experience (UX) adalah keseluruhan pengalaman pengguna saat berinteraksi dengan produk, meliputi aspek kemudahan, efisiensi, dan kepuasan emosional yang dirasakan (Ihsan & Fadilah, 2020). Perancangan UI/UX yang optimal merupakan elemen krusial dalam keberhasilan sistem berbasis *web* modern (Gunawan, 2021). Dalam konteks UMKM, perancangan UI/UX yang baik sangat penting karena mayoritas pengguna (pelaku UMKM) mungkin memiliki tingkat literasi digital yang beragam (Latifah & Prasetyo, 2023).

Berdasarkan landasan teori dan hasil penelitian terdahulu diatas, dapat disimpulkan bahwa penerapan data mining, khususnya algoritma K-Means, serta visualisasi data telah banyak digunakan dalam berbagai konteks penelitian. Namun, masih terdapat keterbatasan penelitian secara khusus mengintegrasikan analisis klasifikasi penjualan UMKM berbasis data kuantitatif ke dalam sebuah sistem informasi berbasis web dengan vvisualisasi dashboard yang mudah dipahami pada tingkat kelurahan. Oleh karena itu, penelitian ini dikembangkan untuk mengisi celah tersebut dengan membangun sistem informasi analisis klasifikasi penjualan UMKM berbasis web di Kelurahan Tanjung Rejo menggunakan algoritma K-Means Clustering dan visualisasi data.

BAB III

ANALISA DAN PERANCANGAN SISTEM

3.1 Analisis Permasalahan

Analisis permasalahan dilakukan untuk mengidentifikasi hambatan-hambatan yang muncul dalam pengelolaan data penjualan UMKM di Kelurahan Tanjung Rejo serta kendala dalam penerapan pendekatan analisis berbasis data. Permasalahan yang ditemukan adalah sebagai berikut:

1. Data penjualan UMKM tidak terdokumentasi secara sistematis.

Banyak UMKM tidak menyimpan data penjualan dalam format digital sehingga sulit dilakukan analisis berbasis data. Kondisi ini sesuai dengan temuan Pratama & Wibowo (2022) bahwa UMKM sering memiliki pencatatan manual yang menyebabkan proses analisis data tidak efektif.

2. Tidak adanya sistem informasi berbasis web yang dapat mengolah data penjualan secara otomatis.

Minimnya penggunaan sistem terintegrasi membuat proses analisis memakan waktu dan rentan kesalahan. Hal ini sejalan dengan Yusuf & Sitorus (2021) yang menyatakan bahwa ketidakhadiran sistem digital menyebabkan UMKM kesulitan melakukan pengolahan data.

3. UMKM tidak dapat mengidentifikasi pola dan performa penjualan secara objektif.

Tanpa adanya metode analisis seperti clustering, UMKM tidak dapat mengetahui kategori perkembangan usaha mereka. Sari & Nugroho (2021)

menyebutkan bahwa klasifikasi performa usaha membutuhkan metode analisis agar menghasilkan evaluasi yang objektif.

4. Tidak tersedia visualisasi data yang mudah dipahami oleh pelaku UMKM.

Visualisasi penting agar pemilik UMKM tanpa kemampuan teknis tetap dapat memahami hasil analisis (Putra, 2020). Dari temuan tersebut dapat disimpulkan bahwa UMKM membutuhkan sistem berbasis web yang mampu mengelola data penjualan, mengelompokkan performa usaha menggunakan K- Means Clustering, dan menampilkan hasilnya dalam bentuk visualisasi interaktif.

& Devi, 2023; Smajic et al., 2022; Vivek & Nanthagopan, 2021).

3.2 Jenis dan Sumber Data

Penelitian ini akan menggunakan kombinasi data primer dan sekunder untuk mencapai tujuan identifikasi klasifikasi penjualan UMKM menggunakan algoritma K- Means clustering dan visualisasi data. Pendekatan metode campuran ini diyakini dapat memberikan pemahaman yang komprehensif dengan memanfaatkan kekuatan dari kedua jenis data tersebut, sehingga menghasilkan analisis yang lebih akurat dan mendalam (Saraswati & Devi, 2023; Smajic et al., 2022; Vivek & Nanthagopan, 2021).

3.2.1 Data Primer

Data primer berasal langsung dari responden (pelaku UMKM) di Kelurahan Tanjung Rejo melalui metode survei dengan kuesioner sesuai 5 variabel yang telah ditentukan, yaitu:

1). Berapa total penjualan usaha anda dalam 1 bulan terakhir?

- 2). Berapa jumlah transaksi penjualan yang terjadi dalam 1 bulan terakhir?
- 3). Berapa rata-rata jam operasional usaha per hari?
- 4). Bagaimana kategori lokasi usaha anda?
- 5). Berapa jumlah hari usaha beroperasi dalam 1 bulan terakhir?

3.2.2 Data Sekunder

- a. Literatur Ilmiah: Jurnal, prosiding, dan buku teks yang membahas tentang sistem informasi, analisis penjualan, K-Means *clustering*, visualisasi data, dan UMKM (Angellin et al., 2023; Arla & Sutabri, 2024; Febima & Magdalena, 2024).
- b. Data Pendukung Lainnya: Informasi umum mengenai kondisi ekonomi lokal atau tren pasar yang relevan dari lembaga kredibel.

3.3 Variabel dan Indikator Penelitian

Variabel dalam ini ditetapkan berdasarkan tujuan penelitian, yaitu untuk mengelompokkan (mengklasifikasikan) UMKM berdasarkan kinerja penjualannya menggunakan pendekatan data mining dengan algoritma K-Means Clustering. Seluruh variabel bersifat kuantitatif, sehingga setiap indikator dapat diukur menggunakan data numerik yang diperoleh dari hasil kuesioner dan dokumentasi penjualan UMKM.

Tabel 3.1 Variabel Indikator Pengelompokan

Kode Variabel	Nama Variabel	Jenis Data	Indikator Pengukuran	Sumber Data	Skala Pengukuran
1	Omzet Bulanan	Kuantitatif	Total pendapatan usaha per bulan (dalam juta rupiah)	Data UMKM (Excel)	Rasio
2	Jumlah Transaksi per Bulan	Kuantitatif	Total jumlah transaksi penjualan dalam satu bulan	Data UMKM (Excel)	Rasio

Kode Variabel	Nama Variabel	Jenis Data	Indikator Pengukuran	Sumber Data	Skala Pengukuran
3	Jam Buka per Hari	Kuantitatif	Jumlah jam operasional usaha dalam satu hari	Data UMKM (Excel)	Rasio
4	Lokasi Strategis	Kuantitatif	Skor tingkat strategis lokasi usaha (skala 1–10)	Data UMKM (Excel)	Interval
5	Hari Operasional per Bulan	Kuantitatif	Jumlah hari usaha beroperasi dalam satu bulan	Data UMKM (Excel)	Rasio

Kelima variabel di atas merupakan variabel input (independen) dalam proses analisis *K-Means clustering*. Data yang dihasilkan dari masing-masing indikator akan dinormalisasi agar setiap variabel memiliki skala yang seragam sebelum dilakukan proses klusterisasi. Tujuan utama pengelompokan ini adalah untuk mengidentifikasi pola atau karakteristik penjualan UMKM berdasarkan performa ekonominya, seperti kelompok dengan penjualan tinggi, sedang, atau rendah.

Pemilihan indikator tersebut sejalan dengan penelitian oleh Bagustio et al. (2024) dan Kuswidyawan et al. (2023), yang menggunakan variabel omzet, volume transaksi, dan jumlah pelanggan sebagai parameter utama dalam analisis performa UMKM berbasis data mining. Dengan demikian, seluruh variabel dan indikator yang digunakan dalam penelitian ini memiliki relevansi empiris dan metodologis untuk mendukung hasil klasifikasi yang akurat dan aplikatif dalam sistem informasi berbasis web.

3.4 Algoritma Sistem

Algoritma yang digunakan dalam penelitian ini adalah K-Means Clustering yang berfungsi mengelompokkan UMKM berdasarkan lima variabel

utama, yaitu: omzet bulanan, jumlah transaksi perbulan, jam buka perhari, lokasi strategis, dan hari operasional perbulan. Tahapan algoritma sistem adalah sebagai berikut:

1. Data Collection

Data diperoleh dari hasil kuesioner terhadap 6 UMKM di Kelurahan Tanjung Rejo.

Tabel 3.2. Tabel Data Asli

UMKM	Omzet	Transaksi	Jam	Lokasi	Hari
Munchy Bar	4.000.000	250	5	2	22
Garasi Haircut	2.600.000	127	11	2	26
Rumah Jempol	5.000.000	500	11	2	20
Kedai Racil	4.000.000	180	13	2	26
Kue Pukis W9	4.000.000	400	8	3	27
Rumah Laundry	2.500.000	120	8	2	26

2. Data Preprocessing

a. Data Cleaning

Tidak ditemukan data kosong atau duplikasi.

b. Normalisasi Min-Max

Rumus: $X_{norm} = (X - X_{min}) / (X_{max} - X_{min})$

c. Penyusunan dataset dalam bentuk format numerik Hasil normalisasi:

Tabel 3.3. Tabel Hasil Normalisasi

UMKM	Omzet	Trans	Jam	Lokasi	Hari
Munchy Bar	0.600	0.342	0.000	0.000	0.286
Garasi Haircut	0.040	0.018	0.750	0.000	0.857
Rumah Jempol	1.000	1.000	0.750	0.000	0.000
Kedai Racil	0.600	0.158	1.000	0.000	0.857
Kue Pukis W9	0.600	0.737	0.375	1.000	1.000
Rumah Laundry	0.000	0.000	0.375	0.000	0.857

3. Penentuan Jumlah Cluster (k = 3)

Cluster ditentukan:

1. Cluster 1 → Belum Berkembang
2. Cluster 2 → Kurang Berkembang
3. Cluster 3 → Sudah Berkembang

4. Inisialisasi Centroid Awal

Dipilih secara acak dari data:

1. C1 = Munchy Bar (0.600, 0.342, 0.000, 0.000, 0.286)
2. C2 = Rumah Jempol (1.000, 1.000, 0.750, 0.000, 0.000)
3. C3 = Kue Pukis W9 (0.600, 0.737, 0.375, 1.000, 1.000)

5. Perhitungan Jarak (Euclidean Distance)

Rumus: $d = \sqrt{\sum (x_i - c_i)^2}$

Contoh Perhitungan:

Garasi Haircut ke C1

$$d = \sqrt{(0.040 - 0.600)^2 + (0.018 - 0.342)^2 + (0.750 - 0)^2 + (0 - 0)^2 + (0.857 - 0.286)^2}$$

$$\begin{aligned}
&= \sqrt{(0.040-0.600)^2 + (0.018-0.342)^2 + (0.750-0)^2 + (0-0)^2 +} \\
&(0.857- \\
&0.286)^2\}d=(0.040-0.600)^2+(0.018-0.342)^2+(0.750-0)^2+(0-0)^2+(0.857-0. \\
&286)^2 \\
&=0.314+0.105+0.563+0+0.326= \sqrt{0.314 + 0.105 + 0.563 + 0 +} \\
&0.326\}=0.314+0.105+0.563+0+0.326 =1.308= \sqrt{1.308}=1.308 \\
&=1.144= 1.144=1.144
\end{aligned}$$

Perhitungan dilakukan untuk seluruh data ke C1, C2, dan C3.

6. Pengelompokan Data (Iterasi 1)

Setelah seluruh jarak dihitung, setiap UMKM dimasukkan ke cluster dengan jarak terkecil.

Tabel 3.4. Tabel Hasil Iterasi 1

Cluster	Anggota
Cluster 1	Munchy Bar, Rumah Jempol
Cluster 2	Garasi Haircut, Kedai Racil, Rumah Laundry
Cluster 3	Kue Pukis W9

7. Update Centroid

1. Centroid Baru C1

Rata-rata (Munchy Bar & Rumah Jempol):

$$\text{Omzet} = (0.600 + 1.000)/2 = 0.800$$

$$\text{Trans} = (0.342 + 1.000)/2 = 0.671 \text{ Jam} = (0 + 0.750)/2 = 0.375$$

$$\text{Lokasi} = 0$$

$$\text{Hari} = (0.286 + 0)/2 = 0.143$$

$$C1 \text{ Baru} = (0.800, 0.671, 0.375, 0, 0.143)$$

2. Centroid Baru C2

$$\text{Omzet} = (0.040 + 0.600 + 0)/3 = 0.213$$

$$\text{Trans} = (0.018 + 0.158 + 0)/3 = 0.059$$

$$\text{Jam} = (0.750 + 1.000 + 0.375)/3 = 0.708$$

$$\text{Lokasi} = 0$$

$$\text{Hari} = (0.857 + 0.857 + 0.857)/3 = 0.857$$

$$C2 \text{ Baru} = (0.213, 0.059, 0.708, 0, 0.857)$$

3. Centroid Baru C3

$$\text{Tetap karena hanya 1 anggota: } C3 = (0.600, 0.737, 0.375, 1, 1)$$

8. Iterasi Hingga Konvergen

Langkah 5–7 diulang kembali menggunakan centroid baru. Hasilnya tidak ada perubahan anggota cluster, sehingga algoritma berhenti (konvergen).

9. Visualisasi dan Interpretasi Hasil

Tabel 3.5. Tabel Hasil Normalisasi

Cluster	Anggota	Interpretasi
Cluster 1	Munchy Bar, Rumah Jempol	UMKM dengan omzet & transaksi tinggi
Cluster 2	Garasi Haircut, Kedai Racil, Rumah Laundry	UMKM stabil / menengah
Cluster 3	Kue Pukis W9	UMKM lokasi sangat strategis & operasional maksimal

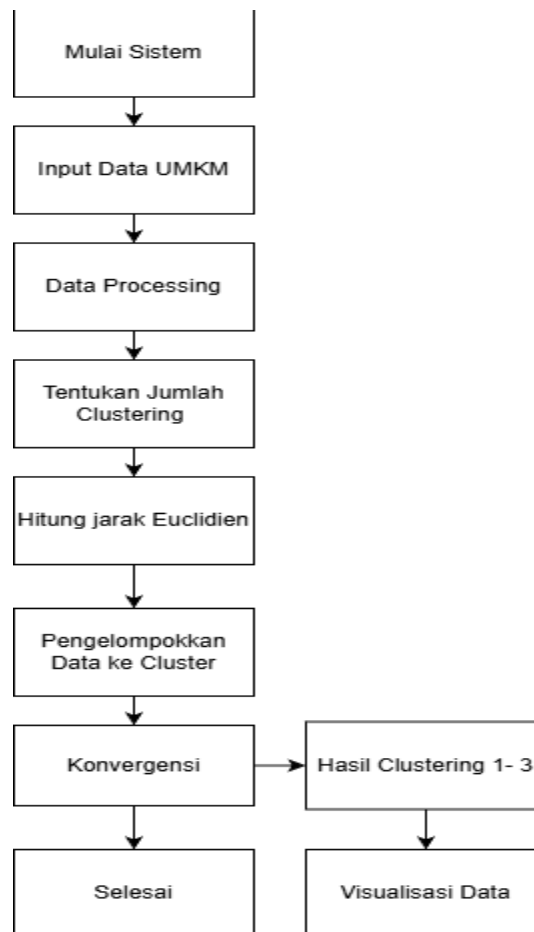
3.5 Pemodelan dan Perancangan Sistem

Bagian ini menyajikan diagram UML (Unified Modeling Language) untuk menggambarkan desain sistem secara lebih formal. Diagram UML digunakan untuk

memodelkan aspek struktural dan perilaku sistem, termasuk interaksi pengguna dan alur proses.

3.5.1 Flowchart Sistem

Flowchart sistem digunakan untuk menggambarkan alur kerja sistem informasi secara terstruktur mulai dari proses input data hingga output informasi yang dihasilkan oleh sistem (Yusuf & Sitorus, 2021). Flowchart berfungsi sebagai representasi visual dari logika proses sistem sehingga memudahkan pemahaman hubungan antar tahapan proses yang saling terintegrasi (Triatama et al., 2025).



Gambar 3.1 Flowchart Sistem.

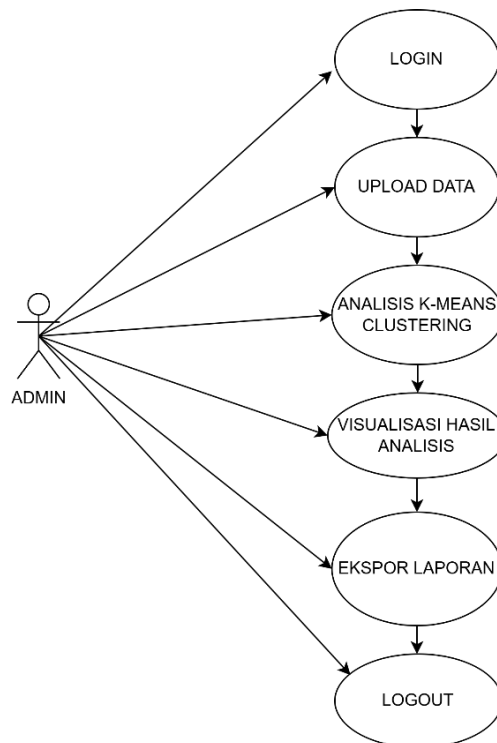
3.4.2 Perancangan Unified Modeling Language (UML)

A. Use Case Diagram

Use Case Diagram digunakan untuk menggambarkan interaksi antara pengguna dan sistem. Pada sistem informasi analisis klasifikasi UMKM berbasis web ini, terdapat satu aktor utama yaitu Admin yang memiliki hak akses penuh.

1. Admin

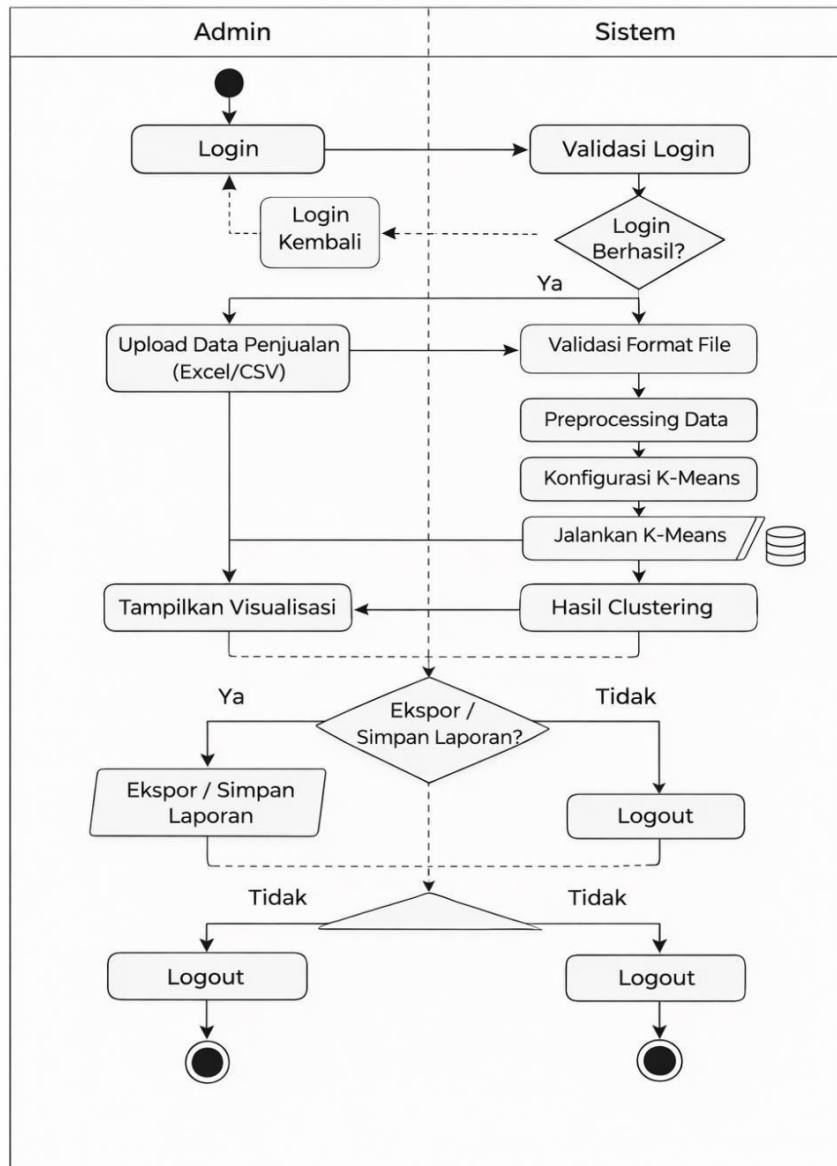
- a. Melakukan login ke sistem
- b. Mengunggah (upload) data UMKM dalam format Excel
- c. Melakukan proses analisis clustering menggunakan algoritma K-Means
- d. Melihat hasil klasifikasi UMKM
- e. Mengakses visualisasi dashboard analitik
- f. Mengunduh (download) hasil clustering dalam format Excel
- g. Logout dari sistem



Gambar 3. 2 Rancangan Use Case Diagram Sistem.

B. Activity Diagram

Activity Diagram digunakan untuk menggambarkan urutan aktivitas dari awal hingga akhir pada proses analisis data UMKM. Diagram ini menegaskan bahwa sistem bekerja secara berurutan:

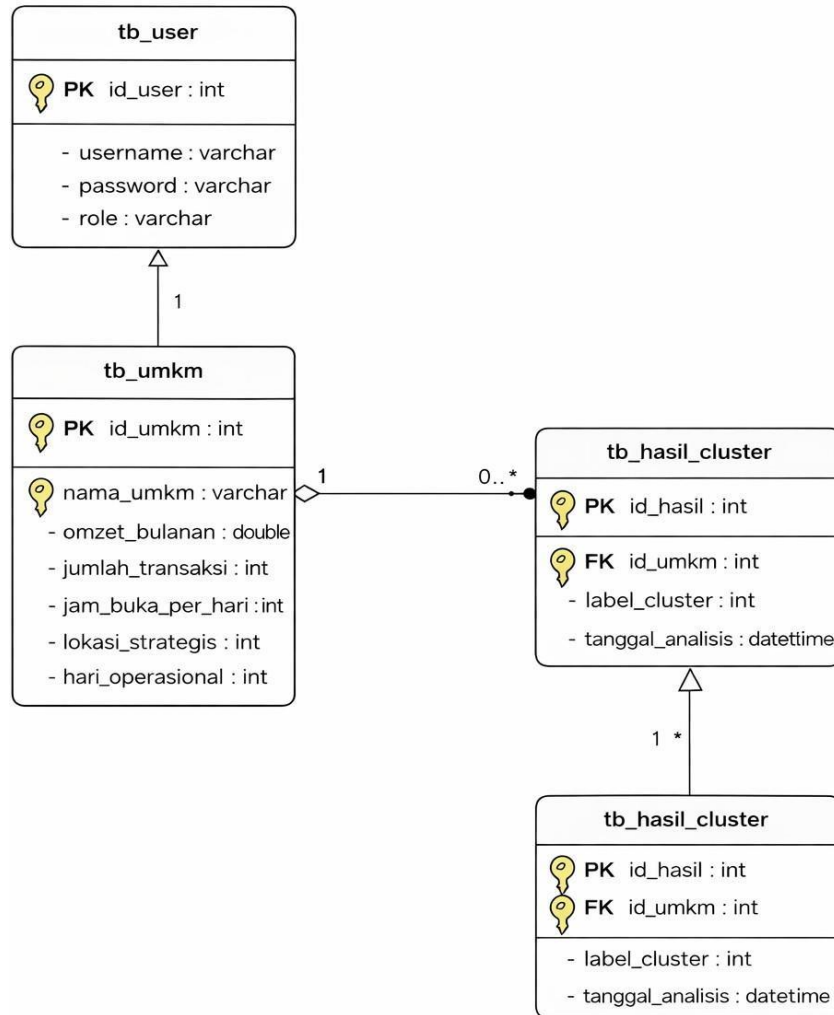


Gambar 3. 3 Rancangan Activity Diagram

C. Class Diagram

Model data pada sistem dirancang menggunakan *Unified Modeling Language (UML)* dalam bentuk Class Diagram untuk menggambarkan

struktur kelas, atribut, serta hubungan antar kelas dalam sistem informasi analisis klasifikasi UMKM. Pada sistem informasi analisis klasifikasi UMKM ini terdapat tiga tabel utama, yaitu :



Gambar 3. 4 Rancangan Class Diagram

3.5.2 Arsitektur Sistem

Arsitektur sistem pada penelitian ini dirancang menggunakan pendekatan berbasis web dengan tiga lapisan utama (three-tier architecture), yaitu Presentation Layer, Application Layer, dan Data Layer. Pendekatan ini digunakan untuk memisahkan tampilan antarmuka, logika proses, dan pengelolaan data agar sistem

lebih terstruktur dan mudah dikembangkan.

1. **Presentation Layer (Front-End)** Lapisan ini menyediakan antarmuka pengguna dalam bentuk halaman login, fitur upload data UMKM, dashboard visualisasi hasil clustering, serta fitur ekspor laporan. Teknologi yang digunakan pada lapisan ini meliputi: HTML, CSS, Bootstrap, JavaScript, Chart.js
2. **Application Layer (Back-End)** Lapisan ini menjalankan logika bisnis sistem, termasuk proses validasi data, preprocessing (data cleaning dan normalisasi), serta eksekusi algoritma K-Means Clustering. Teknologi yang digunakan pada lapisan ini adalah: Python, Framework Flask, Library Pandas untuk pengolahan data, Scikit-learn untuk implementasi algoritma K-Means .
3. **Data Layer (Database)** Lapisan ini bertugas menyimpan data pengguna (admin), data UMKM, serta hasil clustering dalam bentuk terstruktur menggunakan basis data relasional. Struktur tabel utama yang digunakan dalam sistem adalah: tb_user , tb_umkm, tb_hasil_cluster.

Arsitektur three-tier ini memberikan pemisahan tanggung jawab yang jelas antara tampilan, proses, dan penyimpanan data, sehingga sistem lebih modular, fleksibel, dan mudah dikembangkan di masa mendatang.

3.5.3 Rancangan Antarmuka (UI/UX)

Rancangan antarmuka dibuat untuk memberikan gambaran awal mengenai tampilan sistem. Desain antarmuka meliputi:

a. Halaman Login

UMKM CLUSTERING

Text

Text

Gambar 3. 5 Rancangan Halaman Login

b. Dashboard Analysis

UMKM ANALYSIS	<input type="button" value="DASHBOARD"/>	<input type="button" value="DATA UMKM"/>	<input type="button" value="INPUT DATA"/>	<input type="button" value="HASIL & ANALISIS"/>
TOTAL DATA UMKM				
TERAKHIR DIANALISIS		ANALISIS TERAKHIR		
KATEGORI CLUSTER				
CLUSTER 1 BELUM BERKEMBANG		CLUSTER 2 KURANG BERKEMBANG		CLUSTER 3 SUDAH BERKEMBANG

Gambar 3. 6 Rancangan Dashboard

c. Form Data UMKM

UMKM ANALYSIS DASHBOARD DATA UMKM INPUT DATA HASIL & ANALISIS

DAFTAR DATA UMKM TAMBAH MANUAL IMPORT EXCEL

nama umkm	omzet	transaksi	jam buka	lokasi	hari operasional
-----------	-------	-----------	----------	--------	------------------

Gambar 3. 7 Rancangan Data UMKM

d. Halaman Tambah Data UMKM

UMKM ANALYSIS DASHBOARD DATA UMKM INPUT DATA HASIL & ANALISIS

TAMBAH UMKM

nama umkm

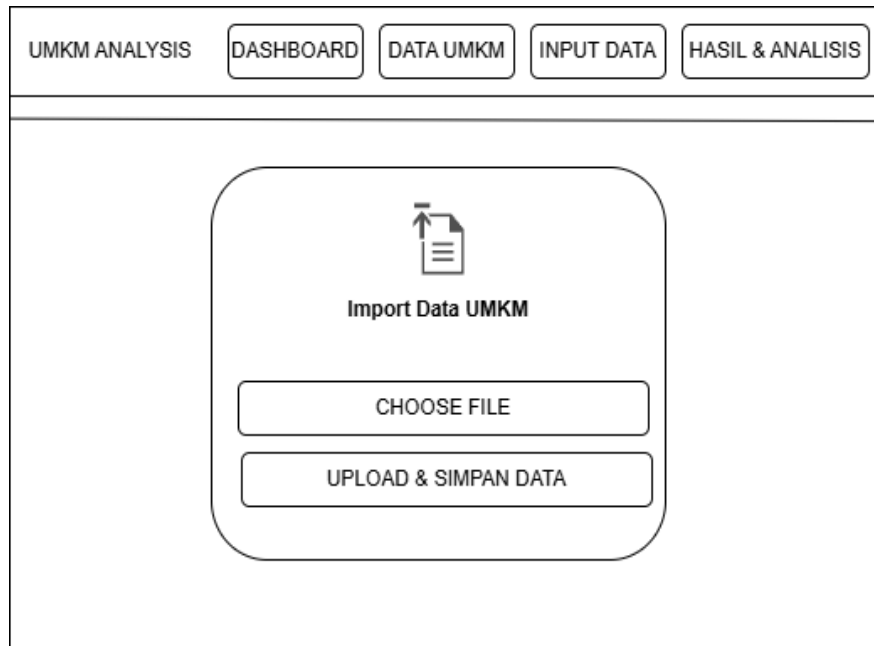
omzet transaksi

jam buka skor lokasi hari operasional

SIMPAN

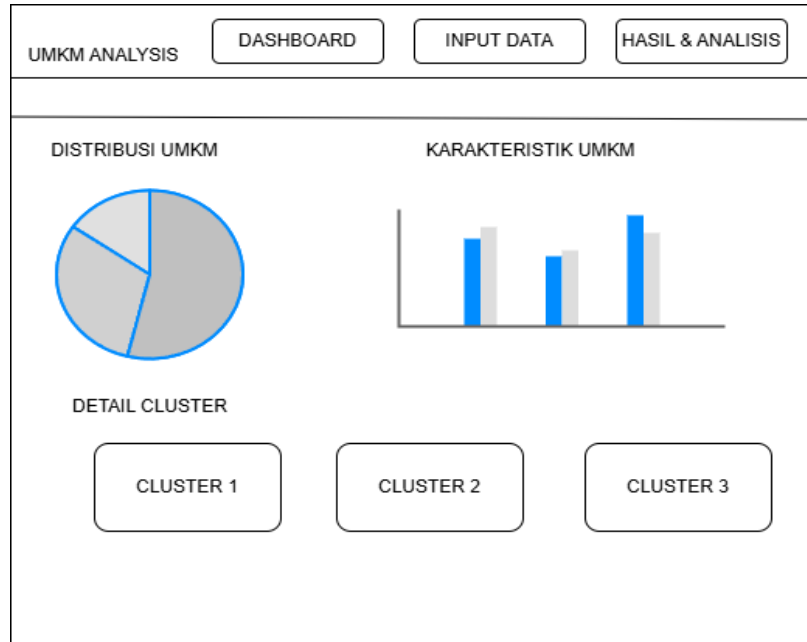
Gambar 3. 8 Rancangan Tambah Data UMKM

e. Halaman Input Data



Gambar 3. 9 Rancangan Input Data

f. Halaman Hasil Analisis



Gambar 3. 10 Rancangan Hasil Analisis

BAB IV

IMPLEMENTASI DAN PENGUJIAN SISTEM

4.1 Kebutuhan Sistem

Bagian ini menjelaskan spesifikasi lingkungan pengembangan dan operasional yang digunakan untuk membangun dan menjalankan Sistem Informasi Analisis Klasifikasi Penjualan UMKM Berbasis Web di Kelurahan Tanjung Rejo menggunakan algoritma *K- Means Clustering*. Integrasi antara perangkat keras, perangkat lunak, dan sumber daya manusia sangat penting untuk memastikan sistem mampu menghasilkan pengelompokan data industri kecil yang optimal dan stabil (Afiasari et al., 2023; Santoso et al., 2022). Perancangan kebutuhan ini bertujuan agar sistem dapat menangani seluruh tahapan *data mining*, mulai dari *preprocessing* hingga visualisasi hasil secara efisien (Imron et al., 2020).

Dengan dukungan spesifikasi perangkat keras dan perangkat lunak yang memadai, sistem ini mampu menjalankan proses klasifikasi penjualan UMKM secara efisien, menghasilkan pengelompokan ke dalam tiga kategori utama yaitu UMKM belum berkembang, kurang berkembang, dan sudah berkembang, serta menampilkan hasil analisis dalam bentuk visualisasi yang mudah dipahami oleh pengguna.

4.1.1 Kebutuhan Perangkat Keras

Perangkat keras yang digunakan dalam pengembangan sistem ini dipilih untuk menghindari hambatan komputasi (*bottleneck*) selama proses iterasi algoritma *clustering* (Sutramiani et al., 2024). Spesifikasi minimal yang

digunakan adalah sebagai berikut:

- a. Processor: Minimal Intel Core i3 generasi ke-8 atau setara. Spesifikasi ini memadai untuk menangani beban komputasi *K-Means* pada dataset skala kecil hingga menengah tanpa memerlukan akselerator perangkat keras tingkat tinggi (Badawi & Bilal, 2019).
- b. RAM: Minimal 8 GB. Kapasitas ini diperlukan untuk mendukung kelancaran *iterative grouping* dan manipulasi dataset besar di memori menggunakan Python (Sutramiani et al., 2024).
- c. Media Penyimpanan: Minimal 256 GB SSD. Penggunaan SSD secara signifikan mempercepat proses pembacaan data dan waktu muat aplikasi web berbasis Flask dibandingkan HDD konvensional (Anggoro & Aziz, 2021).
- d. Koneksi Internet: Diperlukan untuk proses *deployment*, *hosting*, serta instalasi pustaka (*library*) pendukung melalui *package manager*.

4.1.2 Kebutuhan Perangkat Lunak

Perangkat lunak yang dipilih dirancang untuk mendukung ekosistem *data science* yang ringan namun kuat:

- e. Sistem Operasi: Windows 10/11.
- f. Bahasa Pemrograman: Python 3.x. Dipilih karena memiliki ekosistem pustaka analisis data yang sangat lengkap untuk implementasi algoritma *clustering* (Alam & Everhard, 2024; AlShammari, 2024).
- g. Framework Web: Flask. Digunakan sebagai *micro-framework* yang efisien untuk membangun aplikasi web karena sifatnya yang ringan dan hemat sumber daya (Anggoro & Aziz, 2021).

h. Library Pendukung:

1. *Pandas & NumPy*: Untuk manipulasi data tabular dan operasi numerik (Alam & Everhard, 2024).
2. *Scikit-learn*: Pustaka utama untuk implementasi algoritma *K-Means* (AlShammari, 2024).
3. *Openpyxl & Matplotlib*: Untuk manajemen berkas Excel dan evaluasi model.

i. Frontend: HTML, CSS, Bootstrap, JavaScript, dan *Chart.js*. Komponen ini berfungsi untuk menyajikan visualisasi interaktif yang memudahkan pemilik UMKM dalam memahami segmentasi pasar mereka (Kuswidyawan et al., 2023).

j. SQLite (*sqlite3*): Digunakan sebagai sistem manajemen basis data relasional berbasis file (*file-based database*) untuk menyimpan data UMKM, data pengguna, serta hasil clustering. SQLite dipilih karena ringan, tidak memerlukan server terpisah, dan terintegrasi langsung dengan Python melalui library bawaan *sqlite3*, sehingga sesuai untuk pengembangan sistem berbasis web skala kecil hingga menengah.

4.1.3 Kebutuhan Pengguna

Sistem dirancang dengan antarmuka yang ramah pengguna, mengingat target pengguna mungkin tidak memiliki latar belakang teknis yang mendalam (Kuswidyawan et al., 2023). Aktor utama sistem ini adalah Admin Sistem dengan hak akses sebagai berikut:

- k. Keamanan: Melakukan *login* untuk menjaga kerahasiaan informasi sensitif dan data UMKM (Hole, 2024).

- l. Manajemen Data: Mengunggah data penjualan, melakukan *preprocessing* otomatis, dan mengelola *master data* (Octiva et al., 024).
- m. Proses Analisis: Menjalankan algoritma *clustering* dan memantau hasilnya melalui dasbor visualisasi (Oladele et al., 2021).
- n. Output: Melihat, mengunduh hasil klasifikasi untuk keperluan laporan, dan melakukan *logout* untuk keamanan sesi (Octiva et al., 2024).

4.2. Implementasi Sistem

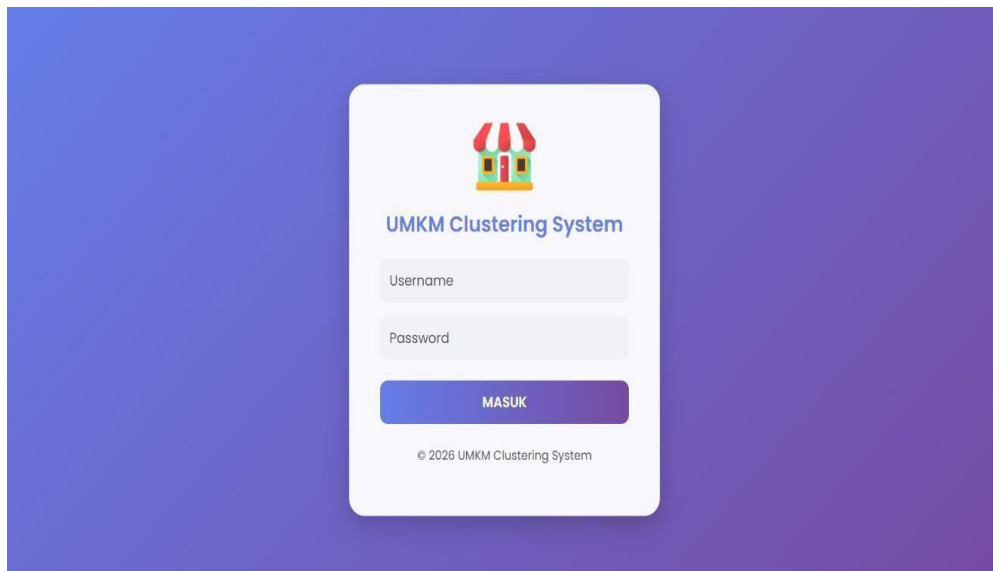
Tahap implementasi merupakan realisasi fisik dari desain sistem yang telah dirancang sebelumnya. Implementasi ini mencakup pembangunan antarmuka pengguna, modul pemrosesan data, integrasi algoritma *K-Means*, serta penyajian hasil analisis dalam bentuk visualisasi interaktif (Яшина & Щербак, 2019). Sistem ini dikembangkan menggunakan Three-Tier Architecture yang memungkinkan pemisahan logika bisnis dengan tampilan untuk meningkatkan efisiensi pengembangan dan kemudahan pemeliharaan sistem (Wang & Jiang, 2018; Yu & Wang, 2022).

4.2.1. Implementasi Halaman Login

Halaman login merupakan gerbang awal sistem yang berfungsi sebagai mekanisme autentikasi pengguna sebelum dapat mengakses fitur utama dalam aplikasi. Halaman ini dirancang dengan tampilan yang sederhana dan informatif, terdiri dari logo *UMKM Clustering System*, kolom input untuk username dan password, serta tombol “MASUK” sebagai aksi untuk melakukan proses autentikasi. Desain antarmuka dibuat minimalis agar memudahkan pengguna dalam melakukan proses login tanpa kebingungan.

Ketika pengguna memasukkan username dan password, sistem akan

mengirimkan data tersebut ke backend untuk dilakukan proses validasi. Validasi dilakukan dengan mencocokkan kredensial yang dimasukkan dengan data autentikasi yang telah ditentukan dalam kode program. Apabila data yang dimasukkan sesuai, sistem akan membuat session login sebagai tanda bahwa pengguna telah terautentikasi, kemudian secara otomatis mengarahkan pengguna menuju halaman Dashboard. Sebaliknya, apabila username atau password tidak sesuai, sistem akan menampilkan pesan kesalahan sebagai bentuk umpan balik kepada pengguna.



Gambar 4.1 Tampilan Halaman Login Admin

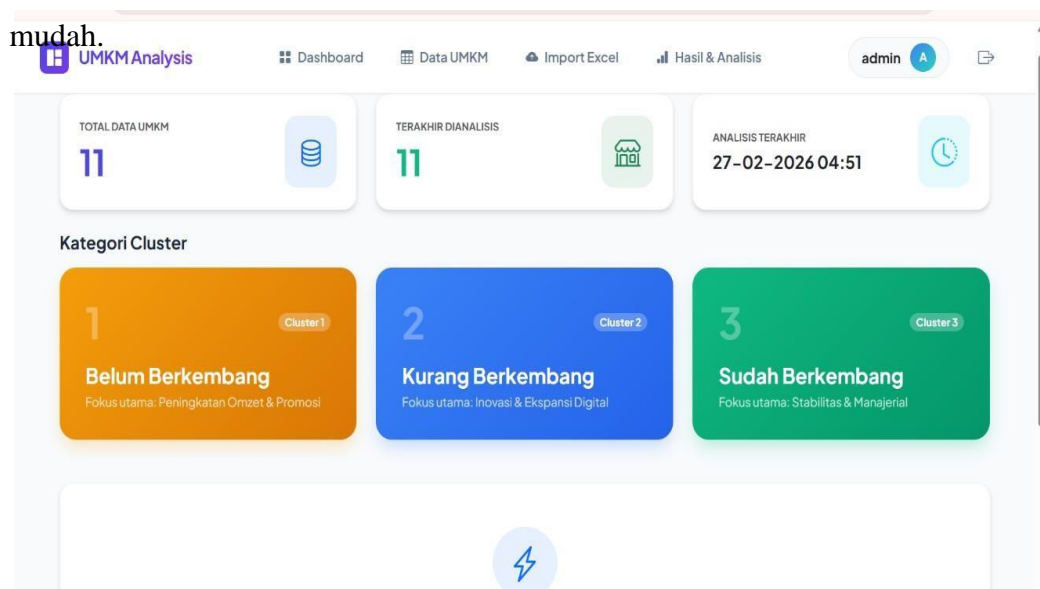
4.2.2. Implementasi Dashboard

Halaman Dashboard merupakan halaman utama yang ditampilkan setelah admin berhasil melakukan login ke dalam sistem UMKM Clustering System. Halaman ini berfungsi sebagai pusat informasi yang menampilkan ringkasan data dan hasil analisis clustering.

Pada bagian atas dashboard terdapat tiga kartu ringkasan, yaitu Total Data UMKM yang menampilkan jumlah keseluruhan data UMKM yang tersimpan di

dalam database, Terakhir Dianalisis yang menunjukkan jumlah data yang diproses pada analisis terakhir, serta Analisis Terakhir yang menampilkan tanggal dan waktu terakhir proses clustering dijalankan. Informasi ini terhubung langsung dengan database sehingga selalu menampilkan data terbaru.

Di bawah bagian ringkasan terdapat Kategori Cluster yang terbagi menjadi tiga kelompok berdasarkan hasil algoritma K-Means Clustering, yaitu Cluster 1 (Belum Berkembang), Cluster 2 (Kurang Berkembang), dan Cluster 3 (Sudah Berkembang). Setiap cluster ditampilkan dalam bentuk kartu berwarna yang berbeda untuk memudahkan identifikasi visual. Masing-masing cluster juga dilengkapi dengan fokus pengembangan, seperti peningkatan omzet dan promosi, inovasi dan ekspansi digital, serta stabilitas dan manajerial. Menu navigasi pada bagian atas terdiri dari Dashboard, Data UMKM, Import Excel, dan Hasil & Analisis yang memungkinkan admin berpindah antar halaman sistem dengan mudah.

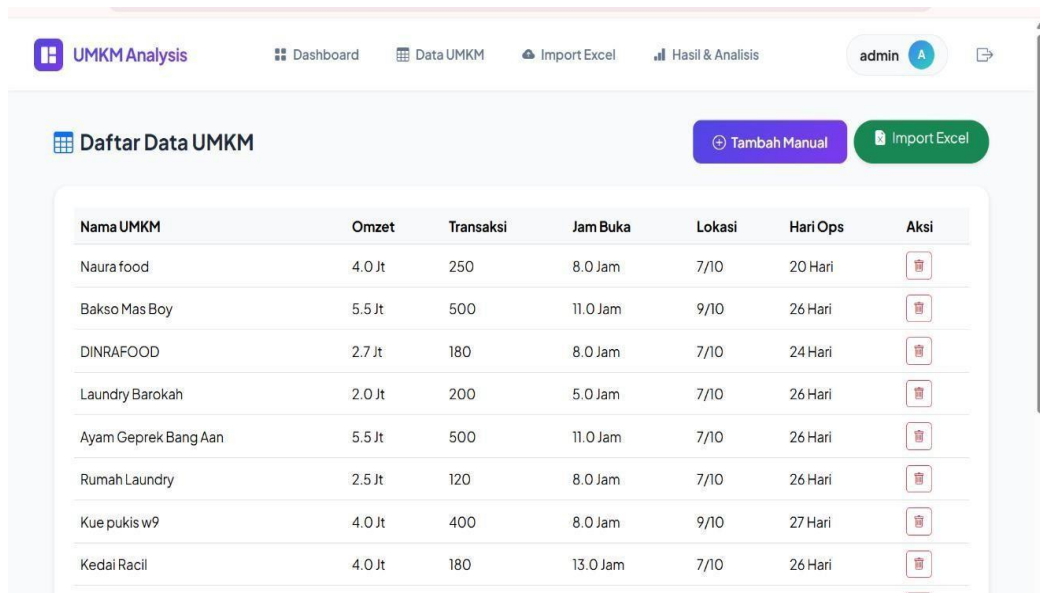










Gambar 4.2 Tampilan Dashboard

4.2.3. Implementasi Halaman Data UMKM

Halaman Data UMKM berfungsi untuk menampilkan seluruh data UMKM yang tersimpan di dalam database dalam bentuk tabel. Data yang ditampilkan meliputi nama UMKM, omzet bulanan, jumlah transaksi per bulan, jam buka per hari, skor lokasi, dan jumlah hari operasional per bulan.

Pada halaman ini tersedia tombol Tambah Manual yang mengarahkan admin ke halaman input data secara manual serta tombol Import Excel untuk menambahkan data dalam jumlah besar. Selain itu, terdapat fitur aksi berupa tombol hapus pada setiap baris data yang memungkinkan admin menghapus data tertentu dari database. Tabel ini terintegrasi langsung dengan database sehingga setiap perubahan data akan langsung tersimpan dan diperbarui secara otomatis pada sistem.



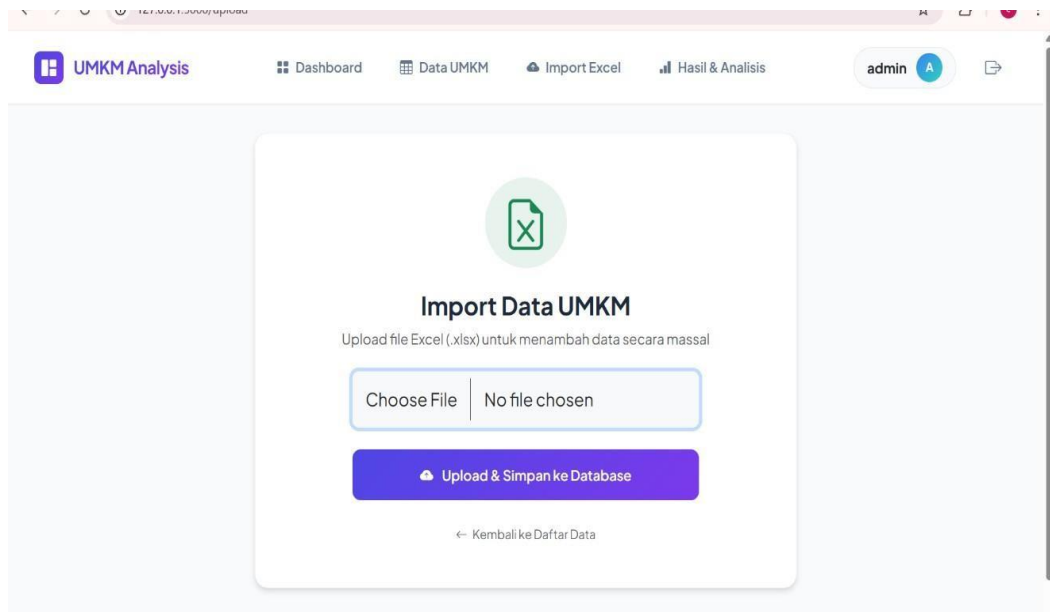
Nama UMKM	Omzet	Transaksi	Jam Buka	Lokasi	Hari Ops	Aksi
Naura food	4.0 Jt	250	8.0 Jam	7/10	20 Hari	
Bakso Mas Boy	5.5 Jt	500	11.0 Jam	9/10	26 Hari	
DINRAFOOD	2.7 Jt	180	8.0 Jam	7/10	24 Hari	
Laundry Barokah	2.0 Jt	200	5.0 Jam	7/10	26 Hari	
Ayam Geprek Bang Aan	5.5 Jt	500	11.0 Jam	7/10	26 Hari	
Rumah Laundry	2.5 Jt	120	8.0 Jam	7/10	26 Hari	
Kue pukis w9	4.0 Jt	400	8.0 Jam	9/10	27 Hari	
Kedai Racil	4.0 Jt	180	13.0 Jam	7/10	26 Hari	

Gambar 4.3 Tampilan Halaman Data

4.2.4. Implementasi Halaman Tambah Data Manual

Halaman Tambah UMKM Secara Manual digunakan untuk memasukkan data UMKM satu per satu melalui form input. Form ini terdiri dari beberapa field yaitu nama UMKM, omzet bulanan (dalam juta), jumlah transaksi per bulan, jam buka per hari, skor lokasi (1–10), dan jumlah hari operasional per bulan.

Admin dapat mengisi seluruh data sesuai dengan kondisi UMKM yang akan dianalisis. Setelah data diinput, admin dapat menekan tombol Simpan Data untuk menyimpan data ke dalam database atau tombol Batal untuk membatalkan proses input. Data yang berhasil disimpan akan langsung muncul pada halaman Data UMKM.



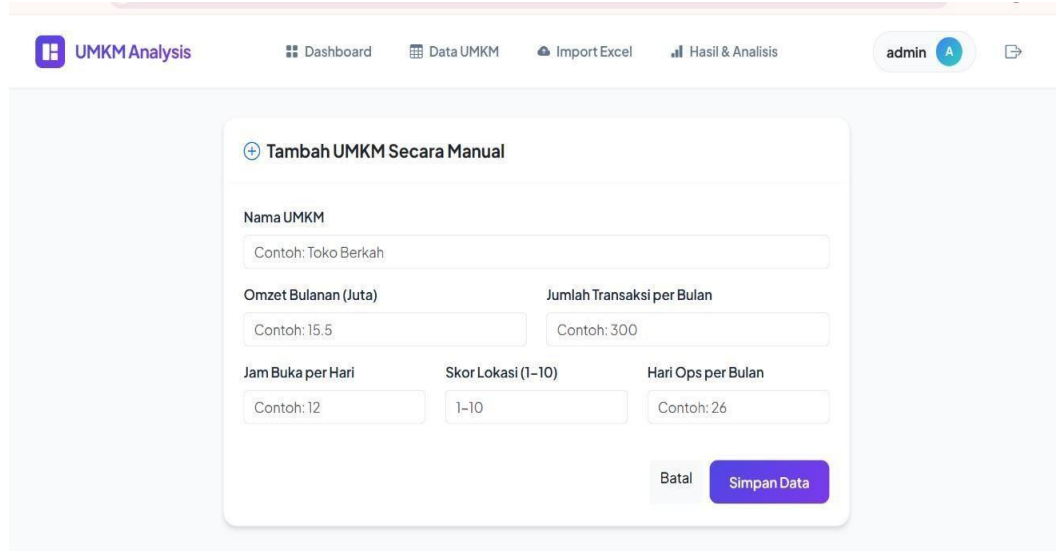
Gambar 4.4 Tampilan Halaman Tambah Data

4.2.5. Implementasi Halaman Import Excel

Halaman Import Data UMKM digunakan untuk menambahkan data secara massal melalui file Excel dengan format .xlsx. Pada halaman ini tersedia fitur Choose File untuk memilih file dari perangkat, serta tombol Upload & Simpan ke

Database untuk memproses dan menyimpan data ke dalam database.

Setelah proses upload berhasil, seluruh data dari file Excel akan otomatis tersimpan dan dapat dilihat pada halaman Data UMKM. Fitur ini memudahkan admin dalam menginput data dalam jumlah besar tanpa harus memasukkan satu per satu secara manual.



Gambar 4.5 Tampilan Import Excel

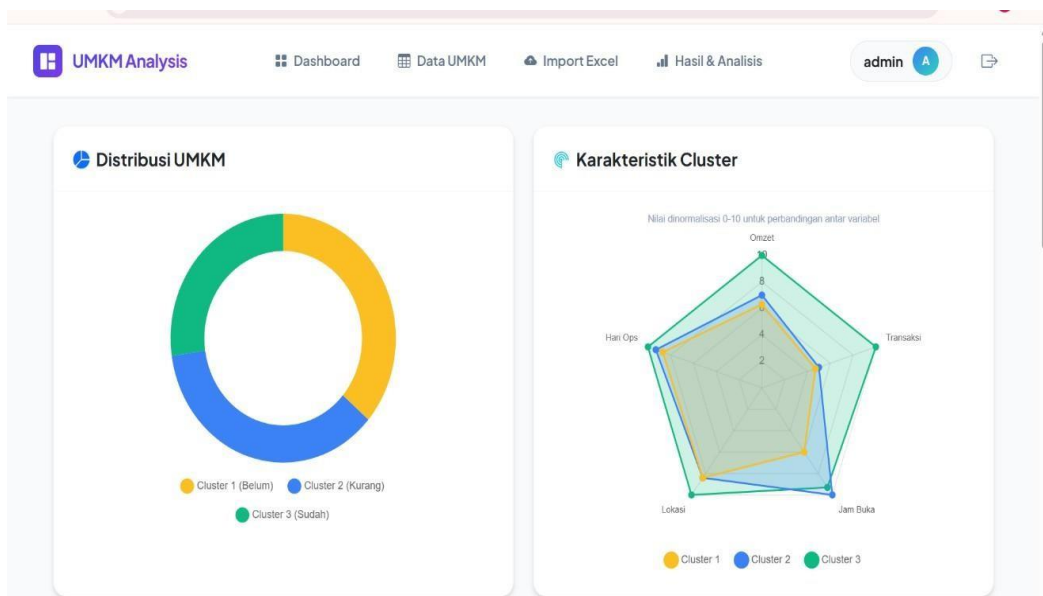
4.2.6. Implementasi Halaman Hasil dan Analisis

Halaman Hasil & Analisis menampilkan hasil proses clustering yang telah dilakukan menggunakan algoritma K-Means. Pada bagian atas ditampilkan grafik rata-rata nilai per cluster yang menunjukkan perbandingan variabel seperti omzet, transaksi, jam buka, lokasi, dan hari operasional pada masing-masing cluster.

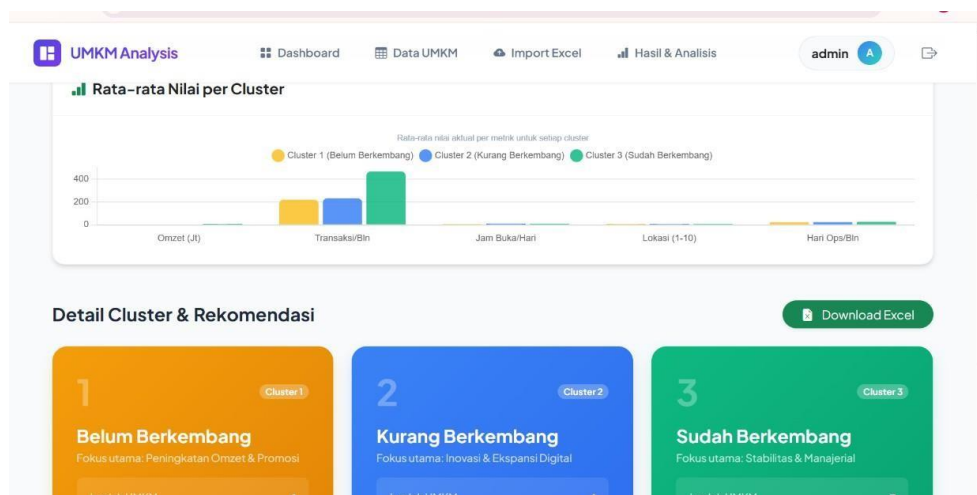
Selanjutnya terdapat bagian Detail Cluster & Rekomendasi yang menampilkan ringkasan setiap cluster, termasuk jumlah UMKM dalam masing-masing cluster serta karakteristik utamanya. Pada halaman ini juga tersedia tombol Download Excel yang memungkinkan admin mengunduh hasil clustering dalam bentuk file Excel.

Di bagian bawah halaman ditampilkan grafik Distribusi UMKM dalam

bentuk diagram lingkaran yang menunjukkan proporsi jumlah UMKM pada masing-masing cluster. Selain itu, terdapat grafik radar Karakteristik Cluster yang menampilkan nilai yang telah dinormalisasi untuk membandingkan kekuatan masing-masing variabel pada setiap cluster. Seluruh hasil analisis ini diambil dari data yang telah diproses dan disimpan dalam database, sehingga hasil yang ditampilkan konsisten dengan data yang tersedia dalam sistem.



Gambar 4.6 Tampilan Distribusi dan Karakteristik



Gambar 4.7 Tampilan Rata Rata Nilai dan Detail Cluster

4.3. Pengujian Sistem

Pengujian sistem dilakukan untuk memastikan bahwa seluruh fitur yang telah diimplementasikan berjalan sesuai dengan kebutuhan sistem. Pengujian ini bertujuan untuk membuktikan bahwa proses input data, preprocessing, perhitungan algoritma K-Means Clustering, serta penyajian hasil dalam bentuk visualisasi dapat berjalan secara akurat dan konsisten.

4.3.1. Pengujian Fungsional (Black Box Testing)

Pengujian fungsional dilakukan tanpa melihat kode program secara langsung, melainkan dengan menguji setiap fitur berdasarkan skenario penggunaan.

- a. Pengujian Login: Sistem diuji dengan memasukkan username dan password yang benar serta salah. Hasil pengujian menunjukkan bahwa sistem hanya mengizinkan akses jika kredensial sesuai, dan menampilkan pesan kesalahan jika tidak sesuai.
- b. Pengujian Upload Data: Sistem diuji dengan mengunggah file berformat .xlsx yang sesuai dan tidak sesuai. Hasil menunjukkan bahwa sistem hanya menerima file dengan format dan struktur kolom yang benar.
- c. Pengujian Preview Data: Setelah file berhasil diunggah, sistem mampu menampilkan isi data dalam bentuk tabel. Data yang tidak sesuai format tidak dapat diproses ke tahap berikutnya.
- d. Pengujian Proses Clustering: Sistem diuji dengan dataset valid dan menghasilkan pembagian cluster tanpa error. Proses berjalan hingga konvergen dan menghasilkan label cluster untuk setiap UMKM.
- e. Pengujian Visualisasi Hasil: Grafik batang, grafik donut, dan

grafik radar ditampilkan sesuai dengan hasil perhitungan cluster. Tidak ditemukan kesalahan dalam pemetaan data ke grafik.

4.3.2. Pengujian Akurasi Algoritma

Pengujian akurasi dilakukan untuk memastikan bahwa hasil clustering sesuai dengan perhitungan algoritma K-Means.

- a. Dilakukan pengecekan nilai normalisasi untuk memastikan seluruh variabel berada pada rentang 0–1.
- b. Dilakukan perhitungan jarak Euclidean secara manual pada beberapa sampel data.
- c. Dibandingkan hasil perhitungan manual dengan output sistem.

Hasil pengujian menunjukkan bahwa label cluster yang dihasilkan sistem sesuai dengan perhitungan manual. Nilai centroid mengalami perubahan pada beberapa iterasi awal dan berhenti ketika kondisi konvergen tercapai.

4.3.3. Pengujian Kinerja Sistem

Pengujian kinerja dilakukan untuk melihat stabilitas dan kecepatan sistem saat memproses data.

- a. Sistem mampu memproses dataset uji dalam waktu relatif singkat.
- b. Tidak terjadi error atau crash selama proses upload dan clustering.
- c. Visualisasi ditampilkan secara responsif pada browser.

Karena sistem tidak menggunakan database dan hanya memproses data dalam memori, performa sistem relatif ringan dan stabil.

4.4. Analisis dan Interpretasi Hasil

Berdasarkan hasil pengelompokan menggunakan algoritma K-Means Clustering dengan jumlah cluster (k) sebanyak tiga, data UMKM di Kelurahan

Tanjung Rejo terbagi menjadi tiga kelompok utama dengan karakteristik yang berbeda. Perbedaan tersebut terlihat dari rata-rata nilai variabel yang digunakan dalam proses clustering, yaitu omzet bulanan, jumlah transaksi, jam buka, lokasi strategis, dan hari operasional.

4.4.1. Interpretasi Cluster 1 (UMKM Belum Berkembang)

Cluster 1 menunjukkan nilai rata-rata terendah pada sebagian besar variabel yang dianalisis. UMKM dalam kelompok ini memiliki omzet bulanan yang relatif rendah serta jumlah transaksi yang masih terbatas. Selain itu, jam operasional usaha cenderung lebih singkat dan hari operasional belum optimal. Nilai lokasi strategis pada cluster ini juga menunjukkan bahwa sebagian besar usaha berada pada posisi yang kurang mendukung peningkatan penjualan.

Berdasarkan karakteristik tersebut, Cluster 1 dikategorikan sebagai UMKM belum berkembang. Kondisi ini menunjukkan bahwa pelaku usaha pada kelompok ini masih menghadapi keterbatasan dalam aspek pemasaran, manajemen operasional, maupun daya saing produk.

Dari hasil interpretasi ini, rekomendasi yang dapat diberikan adalah peningkatan strategi promosi, optimalisasi jam operasional, serta pelatihan manajemen usaha untuk meningkatkan performa penjualan.

4.4.2. Interpretasi Cluster 2 (UMKM Kurang Berkembang / Sedang Berkembang)

Cluster 2 memiliki nilai rata-rata yang berada pada tingkat menengah untuk sebagian besar variabel. Omzet dan jumlah transaksi menunjukkan kestabilan, meskipun belum mencapai kategori tinggi. Jam buka dan hari operasional relatif lebih baik dibandingkan Cluster 1, serta lokasi usaha cukup

mendukung aktivitas penjualan.

Kelompok ini dapat dikategorikan sebagai UMKM kurang berkembang atau sedang berkembang. Artinya, UMKM dalam cluster ini telah memiliki struktur usaha yang cukup stabil, namun masih memiliki ruang untuk peningkatan performa.

Rekomendasi yang dapat diberikan pada cluster ini adalah penguatan strategi pemasaran digital, inovasi produk, peningkatan kualitas layanan, serta pemanfaatan teknologi dalam pencatatan transaksi agar dapat meningkatkan efisiensi dan daya saing.

4.4.3. Interpretasi Cluster 3 (UMKM Sudah Berkembang)

Cluster 3 menunjukkan nilai rata-rata tertinggi pada variabel omzet bulanan dan jumlah transaksi. Jam operasional dan hari operasional juga cenderung lebih optimal dibandingkan dua cluster lainnya. Selain itu, lokasi usaha dalam cluster ini umumnya berada pada kategori strategis yang mendukung peningkatan penjualan.

Berdasarkan karakteristik tersebut, Cluster 3 dikategorikan sebagai UMKM sudah berkembang. UMKM dalam kelompok ini telah menunjukkan performa usaha yang baik dan relatif stabil. Strategi yang dapat diterapkan pada cluster ini bukan lagi pada tahap pembinaan dasar, melainkan pada pengembangan usaha seperti ekspansi pasar, peningkatan kapasitas produksi, serta penguatan branding dan kemitraan.

4.5. Kekurangan Sistem

Meskipun sistem analisis dan clustering UMKM berbasis algoritma K-Means yang dikembangkan telah berjalan sesuai dengan perancangan dan mampu

menghasilkan pengelompokan data secara otomatis, namun masih terdapat beberapa keterbatasan yang perlu diperhatikan. Adapun kekurangan sistem yang teridentifikasi adalah sebagai berikut:

4.5.1. Ketergantungan pada Algoritma Tunggal

Sistem saat ini sepenuhnya bergantung pada algoritma *K-Means* untuk proses segmentasi. Meskipun efisien secara komputasi, algoritma ini memiliki kelemahan inheren berupa sensitivitas tinggi terhadap *outlier* dan asumsi bentuk *cluster* yang sferis (Khan et al., 2024; Spoor, 2022). Selain itu, *K-Means* seringkali kesulitan dalam menangani data yang bersifat kategorikal tanpa transformasi mendalam (Khotimah et al., 2024). Penggunaan metode alternatif seperti *DBSCAN* dapat memberikan hasil yang lebih akurat dalam mengidentifikasi pola data yang kompleks serta memisahkan data kebisingan (*noise*) secara otomatis (Sutramiani et al., 2024).

4.5.2. Subjektivitas dalam Penentuan Jumlah Cluster (K)

Jumlah *cluster* dalam sistem ini ditetapkan secara statis menjadi tiga kategori. Ketiadaan fitur evaluasi otomatis seperti *Elbow Method* atau *Silhouette Score* menyebabkan penentuan ini bersifat subjektif dan berisiko menghasilkan segmentasi yang tidak sesuai dengan struktur alami data (Lungu et al., 2024, 2024). Penggunaan skor siluet sangat krusial untuk memberikan ukuran kuantitatif mengenai kualitas klasifikasi dan memastikan bahwa setiap UMKM ditempatkan dalam kelompok yang paling representatif (Kasem et al., 2023; Zhernov et al., 2021).

4.5.3. Mekanisme Pembaruan Analisis Belum Otomatis

Walaupun data disimpan secara permanen dalam SQLite, proses *clustering*

belum berjalan secara otomatis (*real-time*) setiap kali terdapat perubahan data. Dalam arsitektur manajemen informasi modern, kemampuan pemrosesan data secara *real-time* sangat penting bagi organisasi untuk merespons perubahan kondisi pasar secara instan (Kisielnicki & Markowski, 2021). Ketergantungan pada proses manual oleh admin untuk menjalankan ulang analisis dapat menyebabkan keterlambatan dalam pengambilan keputusan strategis yang berbasis data (Ogbuefi et al., 2024).

4.5.4. Cakupan Variabel Analisis yang Terbatas

Variabel yang digunakan saat ini (omzet, transaksi, lokasi, dll.) baru mencakup aspek operasional dasar. Penelitian terbaru menunjukkan bahwa kinerja dan pertumbuhan UMKM dipengaruhi oleh dimensi yang lebih luas, termasuk adopsi teknologi digital, kapasitas sumber daya manusia, kemampuan inovasi, hingga dukungan jaringan sosial (Ciampi et al., 2021; Malesu & Syrovátka, 2025; Mohsin et al., 2025). Penambahan variabel non-finansial seperti tingkat kepuasan pelanggan dan produktivitas karyawan akan memberikan gambaran yang lebih holistik mengenai tingkat perkembangan usaha (Hendri, 2025; ZAITSEV, 2023).

BAB V

PENUTUP

5.3. Kesimpulan

Berdasarkan hasil penelitian, perancangan, implementasi, dan pengujian sistem informasi analisis klasifikasi penjualan UMKM berbasis web di Kelurahan Tanjung Rejo menggunakan algoritma K-Means Clustering, maka dapat disimpulkan sebagai berikut:

1. Analisis kebutuhan sistem informasi telah dilakukan dengan mengidentifikasi permasalahan utama dalam pengelolaan data penjualan UMKM, yaitu belum terdokumentasinya data secara sistematis, tidak adanya sistem terintegrasi, serta belum adanya metode analisis performa usaha. Hasil analisis tersebut menjadi dasar dalam perancangan sistem berbasis web yang sesuai dengan kebutuhan pengguna.
2. Perancangan sistem informasi berbasis web berhasil dilakukan menggunakan pendekatan UML dan arsitektur three-tier. Sistem mampu mengolah lima variabel utama, yaitu omzet bulanan, jumlah transaksi per bulan, jam operasional per hari, lokasi strategis, dan jumlah hari operasional per bulan, serta mengimplementasikan algoritma K-Means Clustering untuk mengelompokkan UMKM ke dalam tiga kategori, yaitu belum berkembang, kurang berkembang, dan sudah berkembang.
3. Pengujian sistem dan algoritma K-Means menunjukkan bahwa proses clustering berjalan dengan baik melalui tahapan preprocessing, normalisasi Min-Max, perhitungan Euclidean Distance, hingga iterasi konvergen. Hasil klasifikasi dapat ditampilkan secara akurat dalam bentuk visualisasi data

berupa grafik dan dashboard interaktif yang mudah dipahami oleh pengguna.

4. Implementasi sistem informasi berbasis web telah terintegrasi dengan proses clustering dan visualisasi data, sehingga sistem dapat digunakan oleh pelaku UMKM maupun pihak kelurahan sebagai alat bantu dalam memetakan kondisi usaha serta mendukung pengambilan keputusan berbasis data secara lebih efektif dan objektif.

Secara keseluruhan, sistem yang dibangun telah mampu mendukung proses klasifikasi UMKM berbasis data dan dapat digunakan sebagai alat bantu dalam pengambilan keputusan terkait pembinaan dan pengembangan UMKM di Kelurahan Tanjung Rejo.

5.4. **Saran**

Berdasarkan hasil implementasi dan evaluasi sistem yang telah dilakukan, terdapat beberapa saran untuk pengembangan sistem di masa mendatang:

1. Sistem yang dikembangkan saat ini masih bergantung sepenuhnya pada algoritma K-Means dalam proses segmentasi data. Meskipun metode ini efisien dan mudah diimplementasikan, K-Means memiliki keterbatasan dalam menangani data yang mengandung outlier, bentuk cluster yang tidak sferis, serta variabel kategorikal tanpa transformasi yang kompleks. Oleh karena itu, penelitian selanjutnya disarankan untuk mengimplementasikan metode clustering alternatif seperti DBSCAN atau mengembangkan pendekatan hybrid clustering. Dengan adanya perbandingan beberapa algoritma, sistem dapat menghasilkan segmentasi yang lebih adaptif terhadap struktur data yang kompleks serta mampu mengidentifikasi noise secara otomatis.

2. Penetapan jumlah cluster (K) dalam sistem ini masih dilakukan secara statis tanpa dukungan metode evaluasi kuantitatif. Hal tersebut berpotensi menimbulkan subjektivitas dalam menentukan jumlah kelompok yang paling representatif. Untuk pengembangan selanjutnya, disarankan agar sistem dilengkapi dengan fitur evaluasi seperti Elbow Method dan Silhouette Score untuk menentukan nilai K secara otomatis dan objektif. Dengan penerapan metode evaluasi tersebut, kualitas hasil klasifikasi dapat diukur secara kuantitatif sehingga segmentasi UMKM menjadi lebih akurat dan sesuai dengan pola alami data.
3. Meskipun data telah tersimpan secara permanen dalam database SQLite, proses analisis clustering belum berjalan secara otomatis ketika terjadi perubahan data. Ketergantungan pada eksekusi manual oleh admin dapat menyebabkan keterlambatan dalam pembaruan hasil klasifikasi. Oleh karena itu, penelitian selanjutnya disarankan untuk mengembangkan sistem pemrosesan data secara real-time atau berbasis event-driven, sehingga setiap perubahan data secara otomatis memicu proses perhitungan ulang clustering. Pengembangan ini akan meningkatkan responsivitas sistem dalam mendukung pengambilan keputusan berbasis data secara cepat dan akurat.
4. Variabel yang digunakan dalam penelitian ini masih terbatas pada aspek operasional dasar seperti omzet, jumlah transaksi, lokasi, dan jam operasional. Padahal, perkembangan dan daya saing UMKM dipengaruhi oleh faktor multidimensional yang meliputi adopsi teknologi digital, kapasitas sumber daya manusia, inovasi produk, kepuasan pelanggan, serta jejaring usaha. Oleh karena itu, penelitian selanjutnya disarankan untuk

menambahkan variabel non-finansial dan indikator strategis lainnya agar hasil segmentasi tidak hanya menggambarkan kondisi operasional, tetapi juga mencerminkan tingkat kematangan dan potensi pertumbuhan usaha secara lebih menyeluruh.

DAFTAR PUSTAKA

- Adek Maulidya, Selfira, Gomgom Sidabutar, Reyya Ryn Al Hafiz. (2025). Integrating K- Means Clustering and Apriori for Data Mining-Based Digital Marketing Strategy For Increase UMKM: Study Case Stabat City. *Sinkron*, 9(4), 3190–3196. <https://doi.org/10.33395/sinkron.v9i4.15299>
- Aditya Hari Wibowo, Khatin Faisah, Yudo Devianto. (2022). Analisa Dan Visualisasi Data Penjualan Menggunakan Exploratory Data Analysis Pada PT. Telkominfra. *urnal Teknik Informatika dan Sistem Informasi*, 9(3), 2292–2304. <http://jurnal.mdp.ac.id/>
- Aji Putro Prasetyo, Ahmad Fauzi, Andi Dwi Pangestu. (2024). Aplikasi Manajemen Biro Wisata & Travel Berbasis Java Pada Pt Global Service Wisata. *Jurnal Riset Sistem Informasi*, 1(1), 32–38. <https://journal.smartpublisher.id/index.php/jissi>
- Cut Gemawaty dan Yuce Yuliani. (2023). Rancangan Sistem Informasi Penjualan Suku Cadang dan Pelayanan Pada Bengkel Hiba Mobil. *Remik: Riset dan E-Jurnal Manajemen Informatika Komputer*, 7(1), 123-. <http://doi.org/10.33395/remik.v7i1.11963>
- Dayu Ternando dan Herry Mulyono. (2022). Sistem Informasi Promosi Berbasis Web Pada UMKM Jajanan Aulia. *Manajemen Sistem Informasi*, 7(3), 487–500.
- Dinda Dewi Rahma Wijaya dan Syahfitri Ronauli Daulay. (2025). Strategi Pemasaran Berbasis Segmentasi: Menentukan, Menganalisis, Dan Meramalkan Pasar Sasaran Untuk Peningkatan Kinerja Bisnis. *Jurnal Ilmiah Ekonomi dan Manajemen*, 3(2), 464–478. <https://doi.org/10.61722/jiem.v3i2.3877>
- Ela Khairani Br Siregar dan Lailan Sofinah Harahap. (2025). Tinjauan Literatur Peran Big Data Dalam Pengembangan Konten Digital. *Jurnal Media Akademik (JMA)*, 3(1), 1–12.
- Faridatul Faridatul dan M Bambang Purwanto. (2025). Pemanfaatan Artificial Intelligence (AI) dalam Pemasaran Digital: Tantangan dan Peluang. *Optimal: Jurnal Ekonomi dan Manajemen*, 5(3), 607–622. <https://researchhub.id/index.php/optimal>
- Fauzan Abdillah. Achmad Sirojudin, M. Yusril Amin, Ery Setiawan Jullel Atmadji. (2024). Analisis skalabilitas aplikasi smart parking melalui penerapan clean architecture. *Methomika*, 8(1), 96–104. <https://doi.org/10.46880/jmika.Vol8No1.pp96-104>
<http://ejournal.sisfokomtek.org/index.php/jumin>

- Ikotun, A. M., Ezugwu, A. E., Abualigah, L., Abuhaija, B., & Heming, J.. K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences*, 622, 178-210. <https://doi.org/10.1016/j.ins.2022.11.139>
- Indriyani, I Putu Bagus Wiranata, Steven Hiu. (2024). Strategi Peningkatan Efisiensi Operasional UMKM di Era Digital: Pendekatan Kualitatif dengan Business Intelligence dalam Implementasi E-Commerce. *Informatics for Educators And Professionals: Journal of Informatics*, 9(1), 22–32.
- Kiran Basavaraju and Omid Fatahi Vallilai. (2025). Developing a demand planning strategy for joint forecasting and employing analytical tool in an empirical case study. *Discover Applied Science*, 7. <https://doi.org/10.1007/s42452-025-06740-9>
- Lakshmi Haritha Medida and Kumar. (2024). Addressing Challenges in Data Analytics: A Comprehensive Review and Proposed Solutions. In *Chapter 2* (p. 18).
- Laurenso, J., Jiustian, D., Fernando, F., & Wijaya, A.. Implementation of K-Means, Hierarchical, and BIRCH clustering algorithms to determine marketing targets for vape sales in Indonesia. *Journal of Applied Informatics and Computing*, 8, 64-71. <https://doi.org/10.30814/jaic.v8i1.631>
- Manurung, H. (2024). Pengelompokan Umkm Kota Binjai Menggunakan Metode Clustering K-Means Untuk Mengidentifikasi Pola Perkembangan Bisnis. *Jurnal Informatika Paputama (JIK)*, 8(2), 93–99.
- Marcelina, D., Kurnia, A. I., & Terttiaavini, T.. Analisis kluster kinerja usaha kecil dan menengah menggunakan algoritma K-Means Clustering. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 10, 785-794. <https://doi.org/10.25126/jtiik.2023106365>
- Maulidya, et al.. Integrating K-Means clustering and Apriori association rule mining to develop data-driven marketing strategies for MSMEs in Stabat City. *Proceedings of the International Conference on Informatics and Computing*. <https://doi.org/10.1109/ICIC59021.2023.10299512>
- Mawarni Saputri, Evendi, Rouzatul Jannah. (2025). Exploring The Effect of Discussion Based Learning in The Flipped Classroom on Physics Concept Mastery. *Jurnal Pendidikan MIPA*, 26(1), 327–339. <http://jurnal.fkip.unila.ac.id/index.php/jpmipa/>
- Mayang Fadillah Haryanti, Achmad Fauzi, Alvina Arum Jelita, Anggi Setyowati, Azzahra Octarina, Edo Putra Edina, Reva Zahra Aulia, S., & Fitriana. (2024). Pengaruh Data Mining, Strategi Perusahaan Terhadap Laporan Kinerja Perusahaan. *Portofolio Juranl Manajemen dan Bisnis*, 3(1), 71–90.

- Muh. Junaidi, Pita Penengah, Danda Bagaskoro, M. I. A. (2025). Optimalisasi Pemetaan UMKMDi Wilayah Dki Jakarta Melalui K-Means Clustering dan PCA. *Jurnal Nusantara of Engineering*, 8(2), 297–306.
<https://ojs.unpkediri.ac.id/index.php/noe>
- Muhammad Noval, Windarsyah, Finki Dona Marleny. (2025). Implementasi Algoritma K- Means Untuk Analisis Pola Penjualan Pada Toko Monisa. *Jurnal Media Informatika (JUMIN)*, 6(3), 1996–2002.
- Nirwana Hendrastuty. (2024). Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa. *Jurnal Ilmiah Informatiak dan Ilmu Komputer (JIMA-ILKOM)*, 3(1), 46–56.
<https://doi.org/10.58602/jima-ilkom.v3i1.26>
- Parisius Ando Duran, Anik Vega Vitianingsih, Moch. Syaiful Riza. (2024). Data Mining Untuk Prediksi Penjualan Menggunakan Metode Simple Linear Regression. *TEKNIKA*, 13(1), 27–34.
<https://10.0.133.100/teknika.v13i1.712>
- Priyo Ari Wibowo dan Raka Fajar Fitra Pratama. (2023). Analisis Kepuasan Pelanggan Untuk Meningkatkan Penjualan Ukm Coffee Shop Cangkir Pertama di Purwakarta. *Jurnal Cahaya Mandalika*, 203–214.
- Rahayu Mayang Sari dan Yori Apridonan M. (2025). Data Mining Untuk Memprediksi Penjualan. *Jurnal Minfo Polgan*, 14(2), 3164–3172.
<https://doi.org/10.33395/jmp.v14i2.15690>
- Sari, I. P., Al-Khowarizmi, A., & Batubara, I. H.. Cluster analysis using K-Means algorithm and Fuzzy C-Means clustering for grouping students' abilities in online learning process. *Journal of Computer Science, Information Technology and Telecommunication Engineering*, 2, 139-144.
<https://doi.org/10.31289/jite.v5i2.5641>
- Sutisna dan Muhammad Nurur Raudhan. (2019). Strategi Pengambilan Keputusan Berbasis Data Analytics Pada UMKM Untuk Meningkatkan Penjualan. *Data Sciebce Indonesia*, 5(2), 38–49.
<https://doi.org/10.47709/dsi.v5i2.6768>
- Veri Arinal dan Irma Rusmarhadi. (2024). Implementasi Data Mining Untuk Menentukan Strategi Penjualan Produk Umkm Raja Geprek Pada Pola Pembelian Konsumen Menggunakan Algoritma Apriori. *Jurnal of Information Technology and Computer Science (INTECOMS)*, 7(5), 1–13.
- Wikarta, N. C. A., & Defiyanti, S.. Pengelompokan UMKM berdasarkan kategori usaha dan sebaran wilayah di Jawa Barat dengan clustering K-Means. *Jurnal Riset Komputer*, 12, 15-25.
<https://doi.org/10.30865/jurikom.v12i1.7820>

Yunda Lestari, RRudi Hartono, Issy Yuliasri, Hendi Pratama. (2023). Problem-Based Learning Implementation for Students' Speaking Achievement. *International Conference on Science, Education and Technology (ISET)*, 886–891. <https://proceeding.unnes.ac.id/index.php/iset>

Zalwanda Vadiissa Arla dan Tata Sutabri. (2024). Analisis Produk Terlaris Menggunakan Metode K-Means Clustering Pada “Toko Hartati”. *Uranus : Jurnal Ilmiah Teknik Elektro, Sains dan Informatika*, 2(4). <https://doi.org/10.61132/uranus.v2i4.513>

TA Final

ORIGINALITY REPORT

25%

SIMILARITY INDEX

22%

INTERNET SOURCES

13%

PUBLICATIONS

11%

STUDENT PAPERS

PRIMARY SOURCES

1	repository.umsu.ac.id Internet Source	7%
2	Submitted to Universitas Muhammadiyah Sumatera Utara Student Paper	1%
3	eprints.upj.ac.id Internet Source	1%
4	docplayer.info Internet Source	1%
5	123dok.com Internet Source	1%
6	www.markombur.com Internet Source	1%
7	Submitted to Universitas Putera Batam Student Paper	<1%
8	etd.uinsyahada.ac.id Internet Source	<1%
9	etheses.uin-malang.ac.id Internet Source	<1%
10	repo.palcomtech.ac.id Internet Source	<1%
11	Rahmat Hidayat Sanusi, Audy A. Kenap, Efraim R. S. Moningkey. "Implementasi Algoritma Dijkstra Untuk Menentukan Rute	<1%

CS Dipindai dengan CamScanner



MAJELIS PENDIDIKAN TINGGI PENELITIAN & PENGEMBANGAN PIMPINAN PUSAT MUHAMMADIYAH
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

UMSU Terakreditasi Unggul Berdasarkan Keputusan Badan Akreditasi Nasional Perguruan Tinggi No. 174/SK/BAK-PT/Ak.Ppy/PT/18/2024

Pusat Administrasi: Jalan Mukhtar Basri No. 3 Medan 20238 Telp. (061) 6622400 - 66224567 Fax. (061) 6625474 - 6631003

www.umsu.ac.id info@umsu.ac.id [umsu.medan](https://www.facebook.com/umsu.medan) [umsu.medan](https://www.instagram.com/umsu.medan) [umsu.medan](https://www.youtube.com/channel/UC...) [umsu.medan](https://www.tiktok.com/@umsu.medan)

FORMULIR PERBAIKAN UJIAN SKRIPSI

Pada hari ini, Jumat, 13 Maret 2026 telah dilaksanakan Ujian Skripsi bagi mahasiswa Fakultas Ilmu Komputer Dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara. Sbb:

Nama Mahasiswa : Tha'qif Andika Putra
NPM : 2209010233
Program Studi : Sistem Informasi
Judul Skripsi : Sistem Informasi Analisis Klasifikasi Penjualan Umkm Berbasis Web Di Kelurahan Tanjung Rejo Menggunakan K-Means Clustering Dan Visualisasi Data

Materi/Point yang Diperbaiki :

	Paraf
Fatma Sari Hutagalung, S.Kom., M.Kom	
Martiano S.Kom., M.Kom	
Dr. Firahmi Rizky, M.Kom	

Berita acara ini ditandatangani setelah skripsi diperbaiki sesuai petunjuk/arahan dari Pembimbing dan Penguji/Pembahas.





UMSU
Unggul | Cerdas | Terpercaya

MAJELIS PENDIDIKAN TINGGI PENELITIAN & PENGEMBANGAN PIMPINAN PUSAT MUHAMMADIYAH
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

UMSU Terakreditasi A Berdasarkan Keputusan Badan Akreditasi Nasional Perguruan Tinggi No. 89/SK/BAN-PT/Akred/PT/III/2019
Pusat Administrasi: Jalan Mukhtar Basri No. 3 Medan 20238 Telp. (061) 6622400 - 66224567 Fax. (061) 6625474 - 6631003
Telp: 061-6622400-401-402-403-404-405-406-407-408-409-410-411-412-413-414-415-416-417-418-419-420-421-422-423-424-425-426-427-428-429-430-431-432-433-434-435-436-437-438-439-440-441-442-443-444-445-446-447-448-449-450-451-452-453-454-455-456-457-458-459-460-461-462-463-464-465-466-467-468-469-470-471-472-473-474-475-476-477-478-479-480-481-482-483-484-485-486-487-488-489-490-491-492-493-494-495-496-497-498-499-500
Email: info@umsu.ac.id | f@umsu.ac.id | i@umsu.ac.id | t@umsu.ac.id | u@umsu.ac.id | v@umsu.ac.id | w@umsu.ac.id | x@umsu.ac.id | y@umsu.ac.id | z@umsu.ac.id

Berita Acara Pembimbingan Proposal

Nama Mahasiswa : Tha'qif Andika Putra Program Studi : Sistem Informasi
NPM : 2209010233
Judul Penelitian : Sistem Informasi Analisis klasifikasi penjualan UMKM Berbasis Web di kelurahan tanjung rejo Menggunakan K-Means Clustering dan Visualisasi Data.
Nama Dosen Pembimbing : Dr. Firahmi Rizky, S.Kom., M.Kom

Tanggal Bimbingan	Hasil Evaluasi	Paraf Dosen
2/12/2025	Bimbingan bab I	<i>[Signature]</i>
4/12/2025	bimbingan bab II dan III	<i>[Signature]</i>
11/12/2025	Perbaikan bab II dan III acc sempit	<i>[Signature]</i>
3/2/2026	bimbingan bab IV	<i>[Signature]</i>
10/2/2026	Perbaikan bab IV	<i>[Signature]</i>
17/2/2026	Bimbingan bab V	<i>[Signature]</i>
20/2/2026	Perbaikan bab V	<i>[Signature]</i>
4/03/2026	Acc sidang	<i>[Signature]</i>

Diketahui oleh :
Ketua Program Studi

[Signature]

(Maharoka Abdi Prawira Tanjung, S.Kom., M.Kom)

Medan, Maret 2026

Disetujui oleh :
Dosen Pembimbing

[Signature]

(Dr. Firahmi Rizky, S.Kom., M.Kom)





**LEMBAGA PENELITIAN DAN
PENGEMBANGAN PUSTAKA CENDEKIA**
**Ambidextrous : Journal of Innovation, Efficiency and
Technology in Organization**
HP : 0895404313613 / 081274887803 ; E-Mail : admin@jurnalpustakacendekia.com ; E-ISSN :
3031-7002



SURAT KETERANGAN PUBLIKASI ARTIKEL JURNAL

Letter Of Accepted (LoA)

Nomor : 2260/JIFTO/Vol4/No1/2026

Kepada YTH.
THAQIF ANDIKA PUTRA

Terimakasih telah mengirimkan artikel terbaik anda untuk diterbitkan pada **Ambidextrous : Journal of Innovation, Efficiency and Technology in Organization** dengan judul:

**"SISTEM INFORMASI ANALISIS KLASIFIKASI PENJUALAN UMKM
BERBASIS WEB DI KELURAHAN TANJUNG REJO MENGGUNAKAN
K-MEANS CLUSTERING DAN VISUALISASI DATA"**

Berdasarkan hasil review dan keputusan tim editor, maka artikel tersebut dinyatakan **DITERIMA** untuk dipublikasikan pada **Ambidextrous : Journal of Innovation, Efficiency and Technology in Organization** edisi **Volume 4 Nomor 1**.

Demikian surat keterangan ini kami sampaikan untuk dipergunakan sebagaimana mestinya, kami ucapkan terimakasih.

No Reg : LOA20260501055129

Palembang, 01 May 2026
Editor In Chief



Keaslian LOA Dapat
Diperiksa Dengan
Memindai QR Code
Disamping !

LOA20260501055129



Ilzar Daud, Ph.D

Penerbit :

Lembaga Penelitian dan Pengembangan Pustaka Cendekia

Jl. Sei Sahang, NO. 5296, Kec. IB 1, Kel. Lorok Pakjo, Kota Palembang

0895404313613 / 081274887803

admin@jurnalpustakacendekia.com

Dipindai dengan CamScanner

Submissions

My Queue 1

Archives

Help

My Assigned

Filters

New Submission

498 **Thaqif Andika Putra**

Web-Based MSME Sales Classification Analysis Information System in Tanjung Rejo Village Using ...

Submission

View

