

**PERENCANAAN PRIORITAS STOK SPAREPART SERVIS
HANDPHONE BERDASARKAN PENGELOMPOKAN KASUS
SERVIS MENGGUNAKAN AGGLOMERATIVE
HIERARCHICAL CLUSTERING PADA KONTER ANJAS
REPARASI**

SKRIPSI

DISUSUN OLEH

MUHAMMAD BAGUS NURCAHYO

NPM. 2209010085



UMSU

Unggul | Cerdas | Terpercaya

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

MEDAN

2026

**Perencanaan Prioritas Stok Sparepart Servis Handphone
Berdasarkan Pengelompokan Kasus Servis Menggunakan
Agglomerative Hierarchical Clustering pada Konter Anjas
Reparasi**

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**

MUHAMMAD BAGUS NURCAHYO

NPM. 2209010085

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
MEDAN
2025**

LEMBAR PENGESAHAN

Judul Skripsi : Perencanaan Prioritas Stok Sparepart Servis Handphone Berdasarkan Pengelompokan Kasus Servis Menggunakan Agglomerative Hierarchical Clustering pada Konter Anjas Reparasi
Nama Mahasiswa : MUHAMMAD BAGUS NURCAHYO
NPM : 2209010085
Program Studi : SISTEM INFORMASI

Menyetujui
Komisi Pembimbing



(Halim Maulana, S.T., M.Kom., MTA)
NIDN. 0121119102

Ketua Program Studi



(Mahardika Abdi Prawira Tanjung,
S.Kom, M.Kom)
NIDN. 0117988902



(Dr. Al-Khwarizmi, S.Kom, M.Kom.)
NIDN. 0127099201

PERNYATAAN ORISINALITAS

**Perencanaan Prioritas Stok Sparepart Servis Handphone
Berdasarkan Pengelompokan Kasus Servis Menggunakan
Agglomerative Hierarchical Clustering pada Konter Anjas
Reparasi**

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



Muhammad Bagus Nurcahyo

NPM. 2209010085

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

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

Nama : Muhammad Bagus Nurcahyo
NPM : 2209010085
Program Studi : Sistem Informasi
Karya Ilmiah : Skripsi

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

**Perencanaan Prioritas Stok Sparepart Servis Handphone Berdasarkan
Pengelompokan Kasus Servis Menggunakan Agglomerative Hierarchical
Clustering pada Konter Anjas Reparasi**

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



Muhammad Bagus Nurcahyo

NPM. 2209010085

RIWAYAT HIDUP

DATA PRIBADI

Nama Lengkap : Muhammad Bagus Nurcahyo
Tempat dan Tanggal Lahir : Medan, 18 Agustus 2004
Alamat Rumah : Jl. Karya Bakti Lingkungan 8, Gg. Sepakat
Ujung, Tanah Enam Ratus, Medan Marelan,
Kota Medan
Telepon/Faks/HP : 081260425607
E-mail : muhammadbagusnurcahyo41@gmail.com
Instansi Tempat Kerja : -
Alamat Kantor : -

DATA PENDIDIKAN

SD : SD Panngeran Antasari TAMAT: 2016
SMP : SMP Pangeran Antasari TAMAT: 2019
SMA : SMK Trittech Informatika TAMAT: 2022

KATA PENGANTAR



Penulis tentunya berterima kasih kepada berbagai pihak dalam dukungan serta doa dalam penyelesaian skripsi. Penulis juga mengucapkan terima kasih 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. Firahti Rizky, S.Kom., M.Kom selaku Wakil Dekan I Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU.
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, S.Kom., M.Kom selaku Kepala Program Studi Sistem Informasi Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU.
6. Bapak Mulkan Azhari, S.Kom., M.Kom selaku Sekretaris Program Studi Sistem Informasi Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU.
7. Bapak Halim Maulana, S.T., M.Kom selaku Dosen Pembimbing, penulis menyampaikan rasa terima kasih yang tulus atas segala bimbingan dan arahan yang berarti selama proses penyusunan skripsi ini.
8. Kepada kedua orang tua saya bapak Sutarto dan Ibu Suswati saya ucapkan terima kasih yang sebesar-besarnya karena dengan sangat tulus Bapak dan Ibu rela mengorbankan dan mempertaruhkan segalanya untuk pendidikan saya. Yang tak kenal lelah mencari biaya untuk mendukung pendidikan saya sampai memotong kebutuhan pokok untuk belasan tahun kedepan. Saya ucapkan rasa terima kasih yang tak terhingga.

9. Kepada kakak tersayang saya Diah Retno Kusumaningrum dan abang ipar saya Azwar Anas saya ucapkan terima kasih yang sebesar-besarnya atas segala dukungan yang telah kalian berikan kepada saya. Yang selalu menolong dan mendukung saya dalam bentuk moral dan materil walaupun dalam keadaan paling sulit kalian. Terima kasih juga untuk selalu mendengarkan segala keluh kesah saya dan menjadi garda terdepan untuk saya dalam hal apapun itu. Saya tidak tahu bagaimana saya bisa melalui semua ini tanpa kalian. Terima kasih yang sebesar-besarnya juga kepada kakak saya Diah Putri Novitasari dan suami yang juga memberikan dukungan kepada saya.
10. Kepada keponakan-keponakan saya, bayi-bayi lucu saya Dzaki, Gibran, El, dan Syafa. Terima kasih sudah hadir di dunia dan hidup saya, Kehadiran kalian mengisi ruang kosong di hidup saya. Senyum, tangis, dan tawa kalian memberi semangat bagi saya.
11. Kepada sahabat saya sedari kecil sekaligus abang saya yang tidak sedarah yang juga sama memperjuangkan pendidikan dan hidup yang layak Muhammad Roy Prayuda, saya ucapkan terima kasih yang sebesar-besarnya atas segala dukungannya, yang selalu menjadi rumah dan mendengarkan segala keluh kesah saya. Yang menutupi segala kesedihan dan kecemasan saya dengan canda tawanya. Terima kasih untuk selalu ada saat keadaan apapun itu.
12. Kepada sahabat-sahabat kos saya semasa program Pertukaran Mahasiswa Merdeka di Universitas Amikom Purwokerto, saya ucapkan terima kasih sudah menjadi bagian kisah dari perjalanan pendidikan saya. Mengenal kalian memberikan pengalaman baru yang saya tidak pernah rasakan sebelumnya. Juga saya ucapkan terima kasih kepada sahabat-sahabat saya semasa magang MSIB di Jakarta, terima kasih sudah selalu menolong dan menemani saya yang anak rantauan sendiri disana. Terima kasih sudah memperkenalkan saya dengan banyak tempat dan makanan enak juga banyak pengalaman yang tak terhingga.

13. Kepada sahabat-sahabat saya di kelas Aricha Olmi dan Nadia Pratiwi sekaligus juga tim projek Sakuya P2MW, Arya Viandra, Muhammad Fauzhan, saya ucapkan terima kasih juga sebesar-besarnya sudah menemani segala proses saya, memberikan banyak sekali dukungan dan pengalaman seru yang tak terhingga, terutama juga keseruan trip-trip kita di Palaruga dan Sabang, Aceh. Terima kasih sudah selalu mengisi tahun-tahun terakhir saya kuliah dengan segala canda dan tawa kalian. Harap paling besar saya kita semua akan mencapai titik kesuksesan dalam hidup masing-masing dan berjumpa lagi.
14. Kepada seluruh orang yang pernah hadir memberikan cinta, kasih sayang dan juga luka di hidup saya. Terima kasih telah menjadi bagian kisah dari perjalanan hidup saya. Terima kasih juga sudah memberikan pengalaman yang terus membentuk saya.
15. Kepada diri saya sendiri, saya ucapkan terima kasih telah bertahan dari segala badai. Terima kasih telah berjuang dan melangkah sejauh ini, melalui segala proses yang pada akhirnya membentuk karakter dan pola pikir. Dibalik segala air mata, insomnia, kecemasan dan rasa sunyi yang kerap memeluk erat, tapi ternyata mampu melewati semuanya walaupun terkadang memaksa untuk terlihat kuat dan tegar padahal tidak. Terima kasih sudah tumbuh menjadi pribadi dewasa yang jauh lebih baik dari diri semasa kecil.
16. Semua pihak yang terlibat langsung ataupun tidak langsung yang tidak dapat penulis ucapkan satu-persatu yang telah membantu penyelesaian skripsi ini.

Perencanaan Prioritas Stok Sparepart Servis Handphone Berdasarkan Pengelompokan Kasus Servis Menggunakan Agglomerative Hierarchical Clustering pada Konter Anjas

ABSTRAK

Manajemen persediaan suku cadang (*sparepart*) yang tidak terencana pada penyedia jasa reparasi telepon seluler sering kali mengakibatkan penumpukan modal (*dead-stock*) atau kekosongan barang saat dibutuhkan. Penelitian ini bertujuan untuk mengembangkan sistem *dashboard* perencanaan prioritas stok *sparepart* pada Konter Anjas Reparasi menggunakan algoritma *Agglomerative Hierarchical Clustering* (AHC). Metodologi yang digunakan mengacu pada kerangka kerja CRISP-DM (*Cross-Industry Standard Process for Data Mining*). Berdasarkan hasil *Exploratory Data Analysis* (EDA) terhadap 1.113 data riwayat servis historis, diperoleh 811 data valid setelah mengeliminasi anomali operasional. Pemodelan AHC menggunakan pendekatan *bottom-up* menghasilkan jumlah kluster optimal sebanyak 4 kelompok ($k = 4$). Pemilihan $k = 4$ didasarkan pada integrasi evaluasi matematis (*Silhouette Score* dan *Davies-Bouldin Index*) serta validasi pakar (*Expert Judgment*). Hasil klastering ini diimplementasikan ke dalam sebuah *Decision Support System* berbasis web (Flask) yang memiliki fitur utama "Smart Stok". Sistem ini menerjemahkan wawasan analitik menjadi rekomendasi tindakan operasional, yaitu: pemetaan prioritas tinggi (*ready stock*) untuk komponen *fast-moving*, penyediaan terbatas (*buffer stock*) untuk suku cadang situasional bernilai tinggi, serta peringatan tanpa penyetokan fisik untuk kasus *indent part* guna menghindari *dead-stock*. Sistem ini juga dilengkapi fitur "Simulasi Servis" untuk mengestimasi biaya dan durasi pengerjaan, sehingga mampu memberikan panduan berbasis data (*data-driven*) dalam meningkatkan efisiensi rantai pasok Konter Anjas Reparasi.

Kata Kunci: Dashboard, Perencanaan Stok, Sparepart, Agglomerative Hierarchical Clustering, CRISP-DM.

MOBILE PHONE REPAIR SPARE PART STOCK PRIORITY PLANNING BASED ON SERVICE CASE CLUSTERING USING AGGLOMERATIVE HIERARCHICAL CLUSTERING AT ANJAS REPARASI COUNTER

ABSTRACT

Unplanned spare part inventory management in mobile phone repair services often leads to capital stagnation (dead-stock) or stockouts when items are needed. This study aims to conduct mobile phone repair spare part stock priority planning at Anjas Counter using the Agglomerative Hierarchical Clustering (AHC) algorithm. The methodology refers to the CRISP-DM (Cross-Industry Standard Process for Data Mining) framework. Based on the Exploratory Data Analysis (EDA) of 1,113 historical service records, 811 valid data points were obtained after eliminating operational anomalies. The AHC modeling using a bottom-up approach yielded an optimal number of clusters of 4 groups ($k = 4$). The selection of $k = 4$ was based on the integration of mathematical evaluation (Silhouette Score and Davies-Bouldin Index) and expert validation (Expert Judgment). These clustering results were then implemented into a Flask-based Decision Support System web dashboard. This system translates analytical insights into actionable operational recommendations: high-priority mapping (ready stock) for fast-moving components, limited supply (buffer stock) for high-value situational spare parts, and a no-physical-stocking alert for indent part cases to avoid dead-stock. The system also includes a "Service Simulation" feature to estimate service costs and durations, thereby providing data-driven guidance to improve the supply chain efficiency of Anjas Counter.

Keywords: Dashboard, Stock Planning, Sparepart, Mobile Phone, Agglomerative Hierarchical Clustering, CRISP-DM.

DAFTAR ISI

LEMBAR PENGESAHAN	ii
PENYATAAN ORISINALITAS	iii
PENYATAAN PERSETUJUAN PUBLIKASI.....	iv
RIWAYAT HIDUP	v
KATA PENGANTAR.....	vi
ABSTRAK	ix
ABSTRACT	x
DAFTAR ISI.....	xi
DAFTAR TABEL	xiv
DAFTAR GAMBAR.....	xv
BAB I. PENDAHULUAN.....	1
1.1. LATAR BELAKANG MASALAH	1
1.2. RUMUSAN MASALAH.....	2
1.3. BATASAN MASALAH.....	3
1.4. TUJUAN PENELITIAN.....	4
1.5. MANFAAT PENELITIAN	5
1.5.1. MANFAAT TEORITIS.....	5
1.5.2. MANFAAT PRAKTIS.....	5
BAB II. LANDASAN TEORI	6
2.1. LANDASAN TEORI.....	6
2.2. KONSEP LAYANAN SERVIS PERANGKAT ELEKTRONIK.....	6
2.3. MANAJEMEN PERSEDIAAN SPAREPART.....	8
2.4. KONSEP PRIORITAS STOK.....	10
2.5. <i>DATA MINING DAN UNSUPERVISED LEARNING</i>	11
2.5.1. KONSEP DASAR <i>DATA MINING</i> DAN <i>CRISP-DM</i>	11
2.5.2. CLUSTERING SEBAGAI TEKNIK SEGMENTASI.....	12
2.6. <i>ALGORITMA AGGLOMERATIVE CLUSTERING</i>	14
2.6.1. KONSEP DASAR DAN TAHAPAN ALGORITMA	14
2.6.2. PENGUKURAN KEDEKATAN JARAK.....	15
2.6.3. METODE PENGGABUNGAN	15
2.6.4. INTERPRETASI DENDOGRAM	16
2.7. EVALUASI KUALITAS KLASTER	17
2.7.1. <i>SILHOUETTE COEFFICIENT</i>	18
2.7.2. <i>DAVIES-BOULDIN INDEX</i>	18
2.8. DASHBOARD ANALITIK	19
2.8.1. DEFENISI DAN PERAN DASHBOARD DALAM PENGAMBILAN KEPUTUSAN	19
2.8.2. DIGITALISASI PERSEDIAAN PADA UMKM	20
2.8.3. VISUALISASI DATA DAN PROFIL KLASTER	20
2.9. PENELITIAN TERDAHULU.....	21
2.10. KERANGKA KONSEPTUAL.....	26
BAB III. METODOLOGI PENELITIAN	29
3.1. JENIS PENELITIAN.....	29
3.2. LOKASI DAN WAKTU PENELITIAN	30

3.3.	SUMBER DATA DAN TEKNIK PENGUMPULAN DATA	30
3.3.1.	SUMBER DATA.....	30
3.3.2.	TEKNIK PENGUMPULAN DATA	32
3.3.3.	STRUKTUR DATA TRANSAKSI.....	32
3.4.	DEFENISI OPERASIONAL VARIABEL DAN INSTRUMEN.....	34
3.4.1.	VARIABEL INPUT	34
3.4.2.	INSTRUMEN PENELITIAN	35
3.4.3.	VARIABEL OUTPUT	36
3.5.	ANALISIS SISTEM	37
3.5.1.	<i>BUSINESS UNDERSTANDING</i>	40
3.5.2.	<i>DATA UNDERSTANDING</i>	42
3.5.3.	<i>DATA PREPARATION</i>	49
3.5.4.	<i>MODELING (AHC)</i>	57
3.5.5.	<i>EVALUATION</i>	64
3.5.6.	<i>DEPLOYMENT</i>	66
3.6.	PERANCANGAN SISTEM DASHBOARD	68
3.6.1.	PERANCANGAN MODUL DASHBOARD	69
3.6.2.	<i>USE CASE DIAGRAM</i>	72
3.6.3.	<i>ACTIVITY DIAGRAM</i>	73
3.7.	ALUR PENELITIAN	74
BAB IV.	HASIL DAN PEMBAHASAN	77
4.1.	GAMBARAN UMUM IMPLEMENTASI PENELITIAN	77
4.2.	HASIL PERSIAPAN DATA UNTUK CLUSTERING	78
4.2.1.	PEMBERSIHAN DATA	78
4.2.2.	SELEKSI DAN TRANSFORMASI FITUR	79
4.2.3.	STANDARDISASI DATA	80
4.3.	HASIL PEMODELAN CLUSTERING MENGGUNAKAN AHC..	82
4.3.1.	KONFIGURASI MODEL EKSEKUSI ALGORITMA	82
4.3.2.	VISUALISASI DENDOGRAM	83
4.4.	EVALUASI MODEL DAN PENENTUAN JUMLAH KLASTER OPTIMAL	84
4.4.1.	PERBANDINGAN METRIK EVALUASI	85
4.4.2.	PENENTUAN JUMLAH KLASTER OPTIMAL	86
4.4.3.	DISTRIBUSI ANGGOTA KLASTER	87
4.5.	PROFILING DAN INTERPRETASI KLASTER.....	87
4.5.1.	ANALISIS VISUAL KARAKTERISTIK LAYANAN.....	88
4.5.2.	INTERPRETASI KARAKTERISTIK TIAP KLASTER	89
4.6.	REKOMENDASI PRIORITAS PENGADAAN SPAREPART	91
4.7.	IMPLEMENTASI DASHBOARD	94
4.7.1.	MODUL <i>EXECUTIVE DASHBOARD</i>	94
4.7.2.	MODUL DATA SERVIS	95
4.7.3.	MODUL <i>SMART STOCK</i>	96
4.7.4.	MODUL SIMULASI SERVIS	97
4.8.	PENGUJIAN SISTEM	98
4.8.1.	KESIMPULAN PENGUJIAN SISTEM	100
4.9.	PEMBAHASAN DAN ANALISIS HASIL	100

4.9.1. ANALISIS METODOLOGIS DAN KUALITAS PENGELOMPOKAN.....	101
4.9.2. ANALISIS PERGESERAN STRATEGI PERSEDIAAN	101
4.9.3. ANALISIS DAMPAK SISTEM PENDUKUNG KEPUTUSAN	102
BAB V. KESIMPULAN DAN SARAN	104
5.1. KESIMPULAN.....	104
5.2. SARAN.....	106
DAFTAR PUSTAKA	108
LAMPIRAN	

DAFTAR TABEL

TABEL	2.1.			PENELITIAN
TERDAHULU.....				21
TABEL	3.1.	SUMBER	DAN	JENIS
DATA.....				31
TABEL	3.2.	STRUKTUR		DATA
TRANSAKSI.....				33
TABEL	3.3.	DEFINISI	OPERASIONAL	VARIABEL
INPUT.....				34
TABEL	3.4.	DEFINISI	OPERASIONAL	VARIABEL
OUTPUT.....				36
TABEL	3.5.	PROFIL	STRUKTUR	DATA
TRANSAKSI.....				42
TABEL	3.6.	STATISTIK	DESKRIPTIF	VARIABEL
KATEGORIKAL.....				44
TABEL	3.7.	STATISTIK	DESKRIPTIF	VARIABEL
NUMERIK.....				45
TABEL	3.8.	CONTOH	HASIL	NORMALISASI
KERUSAKAN.....				53
TABEL	3.9.		TOTAL	AKSI
SERVIS.....				54
TABEL	3.10.	RINGKASAN	HASIL	DATA
PREPARATION.....				55
TABEL	3.11.	PERUBAHAN	DATA	SETELAH
PREPARATION.....				55
TABEL	3.12.		SAMPEL	DATA
CONTOH.....				62
TABEL	3.13.		MATRIKS	JARAK
AWAL.....				63
TABEL	4.1.	RINGKASAN		PENYUSUTAN
DATASET.....				79
TABEL	4.2.	RINCIAN	ALASAN	PENGHAPUSAN
DATA.....				79
TABEL	4.3.		SAMPEL	DATASET
FINAL.....				81
TABEL	4.4.	PERBANDINGAN	HASIL	EVALUASI
AHC.....				85
TABEL	4.5.	PROFILING	DAN	INTERPRETASI
KLASTER.....				87

TABEL 4.6.	RINGKASAN PROFIL KARAKTERISTIK KLASSTER.....	91
TABEL 4.7.	HIERARKI REKOMENDASI PRIORITAS PENGADAAN SPAREPART.....	93
TABEL 4.8.	SKENARIO DAN HASIL <i>BLACK BOX TESTING</i>.....	98

DAFTAR GAMBAR

GAMBAR 2.1.	KERANGKA KONSEPTUAL PENELITIAN.....	26
GAMBAR 3.1.	SIKLUS <i>CRISP-DM</i>.....	38
GAMBAR 3.2.	DISTRIBUSI JENIS KERUSAKAN DAN TIPE HP TERBANYAK.....	44
GAMBAR 3.3.	DISTRIBUSI BIAYA SERVIS.....	46
GAMBAR 3.4.	IDENTIFIKASI POLA: BIAYA VS DURASI.....	47
GAMBAR 3.5.	TREN BULANAN: PRODUKTIVITAS DAN PENDAPATAN KONTER.....	48
GAMBAR 3.6.	ALUR ALGORITMA AHC.....	60
GAMBAR 3.7.	<i>USE CASE DIAGRAM</i> SISTEM DASHBOARD.....	72
GAMBAR 3.8.	<i>ACTIVITY DIAGRAM</i> SISTEM DASHBOARD.....	73
GAMBAR 3.9.	FLOWCHART ALUR PENELITIAN.....	75
GAMBAR 4.1.	VISUALISASI DENDOGRAM HASIL PEMODELAN AHC.....	83
GAMBAR 4.2.	PEMETAAN KARAKTERISTIK LAYANAN: BIAYA VS DURASI.....	88

GAMBAR 4.3. ANALISIS SEBARAN BIAYA PER TINGKAT LAYANAN.....	89
GAMBAR 4.4. TAMPILAN ANTARMUKA DASHBOARD.....	94
GAMBAR 4.5. TAMPILAN MODUL DATABASE RIWAYATSERVIS.....	95
GAMBAR 4.6. TAMPILAN MODUL SMART STOK.....	96
GAMBAR 4.7. TAMPILA OUTPUT SIMULASI ESTIMASI SERVIS.....	97

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Ketersediaan sparepart sangat menentukan kelancaran operasional konter servis handphone. Kesalahan dalam pengadaan barang dapat memicu dua masalah utama: kekurangan stok (stockout) yang membuat pelanggan kecewa, atau penumpukan stok (overstock) yang menahan perputaran modal usaha (Abbey et al., 2024; Adetula & Akanbi, 2023). Umumnya, banyak konter menggunakan pendekatan seperti Economic Order Quantity (EOQ) atau Reorder Point (ROP) untuk mengelola persediaan. Namun, metode ini kurang efektif karena sifat permintaan servis handphone yang acak dan terputus-putus (intermittent demand). Oleh karena itu, tantangan utamanya bukan lagi sekadar menghitung berapa banyak barang yang harus dipesan, melainkan menentukan suku cadang mana yang harus diprioritaskan agar layanan tetap berjalan tanpa membebani biaya operasional (Torre et al., 2025).

Kondisi tersebut dialami langsung oleh Konter Anjas Reparasi. Selama periode Mei–Oktober 2025, konter ini menangani 1.113 transaksi servis dengan tingkat kerusakan dan durasi pengerjaan yang beragam. Karena belum ada pemetaan kasus servis yang terstruktur, pengelola konter kesulitan menentukan prioritas stok sparepart yang krusial. Akibatnya, keputusan pengadaan sering kali reaktif; konter kerap mengalami indent pada komponen yang penting, namun di sisi lain terjadi penumpukan pada komponen yang jarang terpakai.

Untuk menyelesaikan masalah tersebut, data historis transaksi perlu diklasifikasikan agar pola kebutuhan sparepart dapat diidentifikasi secara objektif (Bhalla et al., 2021; Cisse et al., 2022; Kandemir, 2022). Mengingat data transaksi ini tidak memiliki label kategori baku, pendekatan unsupervised learning sangat cocok diterapkan. Penelitian ini menggunakan algoritma Agglomerative Hierarchical Clustering (AHC) untuk mengelompokkan kasus servis. AHC dipilih karena mampu membentuk hierarki cluster dan menampilkan dendrogram, sehingga pemilik konter dapat lebih mudah mengatur tingkat prioritas sparepart, mulai dari komponen paling kritis hingga yang sekadar cadangan (Li et al., 2022).

Sebagai solusi praktis, hasil pengelompokan AHC ini akan diintegrasikan ke dalam sebuah dashboard interaktif. Pemanfaatan dashboard analitik terbukti mempermudah pemantauan persediaan secara real-time dan meningkatkan kualitas keputusan operasional (Abayomi et al., 2023; Akter & Kudapa, 2024; Sharma, 2025). Melalui sistem ini, diharapkan manajemen logistik Konter Anjas Reparasi dapat berubah dari pendekatan yang reaktif menjadi lebih proaktif dan terarah berdasarkan data pengelompokan kasus servis.

1.2. Rumusan Masalah

Berdasarkan latar belakang tersebut, rumusan masalah penelitian adalah sebagai berikut:

1. Bagaimana karakteristik hasil data transaksi servis Konter Anjas Reparasi periode Mei–Oktober 2025 setelah proses preprocessing dan feature engineering?

2. Bagaimana proses pengelompokan kasus servis menggunakan algoritma Agglomerative Hierarchical Clustering (AHC), serta bagaimana penentuan jumlah cluster terbaik dilakukan?
3. Bagaimana profil karakteristik setiap cluster (berdasarkan variabel biaya, durasi, jumlah komponen, dan skor kesulitan) digunakan untuk menentukan prioritas kebutuhan sparepart?
4. Bagaimana hasil pengelompokan kasus servis tersebut diimplementasikan ke dalam sebuah sistem dashboard untuk mendukung keputusan operasional di Konter Anjas Reparasi?

1.3. Batasan Masalah

Agar penelitian terfokus, batasan yang digunakan meliputi:

1. Data yang digunakan terbatas pada data transaksi historis layanan servis Konter Anjas Reparasi periode Mei–Oktober 2025 sebanyak 1.113 transaksi.
2. Variabel utama yang digunakan dalam proses pengelompokan (clustering) mencakup biaya servis, durasi pengerjaan, jumlah komponen, dan skor kesulitan. Atribut aksi servis dan komponen utama digunakan khusus untuk interpretasi profiling cluster.
3. Metode pengelompokan menggunakan algoritma Agglomerative Hierarchical Clustering (AHC) dengan metode linkage Ward.
4. Evaluasi kualitas cluster dilakukan menggunakan parameter pengujian internal, yaitu Silhouette Score dan Davies–Bouldin Index (DBI).

5. Pembobotan skor kesulitan (feature engineering) didasarkan pada instrumen penilaian empiris yang divalidasi langsung oleh teknisi utama Konter Anjas Reparasi.
6. Implementasi hasil analitik dibangun dalam bentuk aplikasi berbasis web (dashboard) menggunakan kerangka kerja Python (Flask) dipadukan dengan HTML, CSS, dan JavaScript.
7. Output penelitian berfokus pada sistem pemetaan dan penentuan prioritas stok sparepart berbasis profil cluster, bukan membangun sistem pembelian otomatis (auto-replenishment) ataupun model peramalan permintaan kuantitatif (seperti ROP/EOQ).

1.4. Tujuan Penelitian

Tujuan penelitian ini adalah:

1. Menyajikan dataset analitik yang valid melalui proses preprocessing dan feature engineering dari data mentah transaksi Konter Anjas Reparasi.
2. Mengimplementasikan algoritma AHC untuk mendapatkan pengelompokan cluster kasus servis yang paling optimal.
3. Menganalisis profil karakteristik tiap cluster sebagai landasan objektif dalam memetakan prioritas kebutuhan sparepart.
4. Mengembangkan sistem dashboard interaktif berbasis hasil clustering yang mendukung manajemen dan pengambilan keputusan stok sparepart.

1.5. Manfaat Penelitian

1.5.1. Manfaat Teoritis

Penelitian ini diharapkan memberikan kontribusi akademis terkait penerapan unsupervised machine learning, khususnya algoritma Agglomerative Hierarchical Clustering (AHC), dalam memecahkan masalah segmentasi kasus pemeliharaan/servis. Selain itu, penelitian ini memperkaya kajian mengenai integrasi model analitik ke dalam sistem dashboard pendukung keputusan pada skala Usaha Mikro Kecil dan Menengah (UMKM).

1.5.2. Manfaat Praktis

1. Memberikan gambaran komprehensif mengenai pola transaksi layanan servis (mencakup biaya, durasi, dan tingkat kesulitan) di Konter Anjas Reparasi.
2. Membantu pemilik usaha mengenali dan mengantisipasi jenis perbaikan dengan risiko keterlambatan tinggi (kasus indent part).
3. Mendukung penentuan prioritas pengadaan sparepart berdasarkan komponen yang paling dominan pada cluster layanan yang kritis.
4. Menyediakan alat bantu interaktif (dashboard) yang memudahkan pengelola dalam melakukan pemantauan dan pengambilan keputusan operasional secara cepat, presisi, dan berbasis data.

BAB II

LANDASAN TEORI

2.1. Landasan Teori

Bab ini memaparkan konsep teoretis, kerangka analisis, serta tinjauan penelitian terdahulu yang menjadi landasan dalam membangun sistem perencanaan prioritas stok sparepart berbasis pengelompokan kasus servis pada Konter Anjas Reparasi menggunakan algoritma Agglomerative Hierarchical Clustering (AHC).

Kerangka teori mencakup manajemen persediaan suku cadang (sparepart management), penentuan prioritas stok pada usaha servis, serta karakteristik permintaan terputus-putus (intermittent demand). Selain itu, dibahas konsep klasifikasi inventaris, unsupervised machine learning, serta AHC meliputi metrik jarak, metode linkage Ward, interpretasi dendrogram, dan penentuan jumlah cluster. Bab ini juga menguraikan metode evaluasi kualitas clustering (Silhouette Score dan Davies–Bouldin Index) serta konsep dashboard interaktif sebagai media pendukung keputusan operasional.

Seluruh teori tersebut menjadi acuan konseptual dan metodologis dalam penentuan variabel, perancangan proses clustering, interpretasi profil cluster menjadi prioritas sparepart, serta implementasinya ke dalam sistem dashboard berbasis data historis.

2.2. Konsep Layanan Servis Perangkat Elektronik

Usaha servis handphone berperan penting dalam pemeliharaan perangkat dan perpanjangan umur pakai elektronik, yang secara tidak langsung turut

mendukung prinsip ekonomi sirkular dan pengurangan limbah elektronik atau e-waste (Soudachanh & Salhofer, 2025; Türkeli et al., 2019). Namun, dalam kegiatan operasional sehari-hari, keberhasilan layanan servis ini sangat bergantung pada ketersediaan sparepart yang tepat waktu.

Ketiadaan komponen perbaikan pada saat dibutuhkan (*stockout*) dapat menyebabkan penumpukan antrean perbaikan dan waktu tunggu servis yang berkepanjangan. (Zhang et al., 2021) menyatakan bahwa kekurangan sparepart berpotensi menimbulkan kerugian finansial akibat hilangnya peluang penjualan serta menurunnya kepuasan pelanggan. Oleh karena itu, pengelolaan ketersediaan komponen menjadi elemen krusial untuk menjaga kelancaran operasional konter servis.

Berbeda dengan barang konsumsi harian, kebutuhan sparepart pada usaha servis memiliki karakteristik permintaan yang sulit diprediksi. (Achetoui et al., 2019) menjelaskan bahwa permintaan sparepart bersifat *intermittent* (terputus-putus) dan *lumpy* (tidak merata atau terkonsentrasi pada periode tertentu), sehingga metode perencanaan dan peramalan kuantitatif tradisional sering kali kurang akurat. Dinamika pasar dan variasi tingkat kerusakan perangkat juga semakin memperbesar kompleksitas permintaan tersebut (Muntaja & Sriani, 2024). Kondisi ini menunjukkan bahwa pengendalian persediaan tidak dapat hanya mengandalkan pendekatan jumlah pesanan, melainkan perlu mempertimbangkan pola historis penggunaan komponen melalui pendekatan pengelompokan (*clustering*) berdasarkan karakteristik kasus servis.

Pada konteks UMKM, pencatatan persediaan dan penggunaan sparepart masih banyak dilakukan secara manual sehingga menyulitkan pemantauan stok

secara akurat (Harahap et al., 2024). Oleh karena itu, diperlukan sebuah sistem informasi yang mampu memetakan kasus servis secara terstruktur dan menyajikannya dalam bentuk informasi prioritas pengadaan yang mudah dipantau.

Dalam penelitian ini, sparepart didefinisikan secara operasional sebagai komponen utama yang diganti atau digunakan dalam proses perbaikan handphone, seperti LCD, baterai, konektor charging, tombol, dan Integrated Circuit (IC). Informasi kebutuhan sparepart direpresentasikan secara langsung melalui atribut data transaksi servis, yaitu variabel jumlah komponen serta jenis komponen utama. Atribut-atribut ini selanjutnya menjadi dasar untuk menyusun hierarki prioritas stok. Prioritas pengadaan tertinggi diarahkan pada komponen yang paling sering muncul pada cluster layanan yang bersifat kritis (misalnya memiliki biaya, durasi pengerjaan, dan skor kesulitan yang lebih tinggi) atau berisiko indent (keterlambatan akibat menunggu ketersediaan komponen). Selanjutnya, konsep manajemen persediaan sparepart dan pendekatan penentuan prioritas stok akan dibahas lebih mendalam pada sub-bab berikutnya.

2.3. Manajemen Persediaan Sparepart

Manajemen persediaan sparepart pada dasarnya bertujuan untuk menjamin tingkat ketersediaan layanan (*service level*) yang tinggi sekaligus meminimalkan total biaya persediaan. Dalam teori manajemen persediaan, terdapat tiga komponen biaya utama yang harus dikendalikan, yaitu biaya pemesanan (*ordering cost*), biaya penyimpanan (*holding cost*), dan biaya kekurangan stok (*shortage cost*). Menurut (Sukma et al., 2019), keseimbangan antara biaya-biaya tersebut

sangat penting agar operasional pemeliharaan dan perbaikan (*corrective maintenance*) dapat berjalan lancar tanpa membebani modal kerja UMKM secara berlebihan.

Pada praktiknya, pengelola konter servis selalu dihadapkan pada situasi *trade-off* (tarik-ulur) antara risiko kekurangan dan kelebihan stok. (Sukrasorn et al., 2025) menjelaskan bahwa ketiadaan komponen krusial dapat memicu *shortage cost* yang tinggi akibat berhentinya aktivitas perbaikan secara mendadak dan hilangnya peluang pendapatan. Di sisi lain, menumpuk stok terlalu banyak tidak hanya meningkatkan *holding cost*, tetapi juga memunculkan risiko keusangan (*obsolescence*). (Zhang et al., 2021) menegaskan bahwa sparepart handphone sangat rentan terhadap *obsolescence* seiring dengan cepatnya pergantian tren dan model perangkat di pasar, yang dapat mengakibatkan komponen tidak terpakai dan nilainya turun drastis.

Kebijakan persediaan konvensional umumnya berfokus pada penentuan kapan harus memesan dan berapa banyak kuantitas yang dipesan. Namun, praktik perencanaan persediaan yang hanya berlandaskan pada peramalan kuantitas teoritis sering kali belum memadai pada kondisi permintaan yang tidak stabil. (Scarf et al., 2024) menegaskan bahwa manajemen persediaan tidak akan berjalan optimal apabila tidak diintegrasikan secara langsung dengan realitas penanganan kasus pemeliharaan atau servis harian di lapangan. Hal ini menuntut adanya pengelolaan stok yang lebih dinamis, yang selaras dengan pola kebutuhan aktual teknisi.

Mengingat kompleksitas layanan perbaikan tersebut, pendekatan evaluasi persediaan yang hanya menggunakan satu parameter (seperti harga atau volume

permintaan tunggal) dinilai tidak lagi memadai. (Torre et al., 2025) menyatakan bahwa evaluasi persediaan yang efektif memerlukan pendekatan multi-kriteria guna menekan biaya operasional sekaligus meningkatkan efisiensi layanan. Sejalan dengan itu, (Achetoui et al., 2019) menekankan perlunya mempertimbangkan berbagai atribut secara bersamaan untuk menilai tingkat kekritisan suatu komponen. Oleh karena itu, diperlukan pendekatan berbasis data historis untuk menyusun prioritas stok secara objektif melalui pengelompokan atau segmentasi kasus servis. Pembahasan mengenai konsep prioritas stok dan metode pengelompokan tersebut akan diuraikan lebih mendalam pada subbab berikutnya.

2.4. Konsep Priotitas Stok

Penentuan prioritas stok secara tradisional melalui Analisis ABC, FSN, atau VED sering kali dinilai kurang memadai karena hanya mengandalkan kriteria tunggal (Jeddou, 2014; Sukma et al., 2019). Seiring meningkatnya kompleksitas layanan, evaluasi persediaan bergeser menjadi masalah *Multi-Criteria Inventory Classification* (MCIC) yang mengintegrasikan berbagai atribut operasional secara komprehensif (Ghorabae et al., 2015; Torre et al., 2025).

Khusus pada industri jasa perbaikan ponsel, durasi pengerjaan dan biaya servis merupakan faktor operasional paling dominan yang menentukan karakteristik layanan (Hasibuan et al., 2024). Karena pemilahan multi-kriteria secara manual bersifat rumit dan rentan bias, teknik *clustering* seperti *Agglomerative Hierarchical Clustering* (AHC) digunakan untuk mengidentifikasi pola tersembunyi dari data historis secara otomatis (Lolli et al., 2017; Raja et al.,

2016). Profil klaster yang terbentuk nantinya menjadi acuan objektif dalam menentukan prioritas pengadaan dan pemantauan persediaan, terutama pada lingkungan dengan pola permintaan yang fluktuatif (Darma & Setiadi, 2025; Priambodo & Jananto, 2022).

Penelitian ini menetapkan prioritas stok melalui segmentasi kasus servis berdasarkan empat variabel utama: biaya, durasi, jumlah komponen, dan skor kesulitan. Prioritas *sparepart* diturunkan dari profil setiap klaster, khususnya pada kelompok yang menunjukkan tingkat kekritisian layanan yang tinggi (seperti klaster dengan durasi dan skor kesulitan di atas rata-rata). Melalui pendekatan ini, alokasi sumber daya di Konter Anjas Reparasi dapat difokuskan pada komponen yang terbukti paling krusial bagi kelancaran operasional harian.

2.5. Data Mining dan Unsupervised Machine Learning

Perkembangan volume data transaksi pada berbagai sektor usaha, termasuk konter servis handphone, menuntut adanya pendekatan analitik yang mampu mengolah data secara otomatis dan sistematis. Data historis transaksi servis yang tersimpan dalam sistem pencatatan sebenarnya memiliki pola yang bisa dimanfaatkan untuk mendukung keputusan operasional. Oleh sebab itu, pendekatan berbasis *data mining* dan *unsupervised machine learning* relevan diterapkan pada penelitian ini guna mengekstrak informasi yang masih tersembunyi dari sekumpulan data transaksi tersebut.

2.5.1. Konsep Dasar Data mining dan CRISP-DM

Data mining adalah proses eksplorasi dan analisis data dalam jumlah besar untuk menemukan pola, anomali, atau hubungan tersembunyi yang bermakna

(Maimon & Rokach, 2009). Dalam praktiknya, sebuah proyek data mining membutuhkan metodologi standar agar langkah-langkah penelitian berjalan sistematis. Salah satu kerangka kerja yang paling sering digunakan adalah Cross-Industry Standard Process for Data Mining (CRISP-DM). Menurut Schröer dkk. (2021), CRISP-DM telah diakui secara luas sebagai standar de facto dan model proses lintas industri yang andal untuk memandu jalannya proyek data mining. Siklus CRISP-DM sendiri terdiri dari enam tahapan utama yang saling berkesinambungan, yaitu pemahaman bisnis (business understanding), pemahaman data (data understanding), persiapan data (data preparation), pemodelan (modeling), evaluasi (evaluation), dan penyebaran (deployment).

Dalam konteks penelitian di Konter Anjas Reparasi, tahapan awal CRISP-DM memegang peran krusial. Pada fase data understanding, teknik Exploratory Data Analysis (EDA) umumnya diterapkan untuk mengenali karakteristik awal, distribusi, dan anomali pada data historis servis (Agarwal et al., 2024). Pemahaman mendalam dari hasil EDA ini kemudian menjadi masukan utama pada fase data preparation. Selanjutnya, hasil tersebut dimanfaatkan untuk mendukung keputusan preprocessing, seperti penanganan missing value, identifikasi outlier, dan standardisasi fitur. Proses preprocessing ini dilakukan pada atribut operasional (biaya, durasi pengerjaan, jumlah komponen, dan skor kesulitan) sebelum nantinya dieksekusi ke tahap pemodelan machine learning. Melalui penggunaan kerangka CRISP-DM, proses data mining bisa dilakukan secara terstruktur mulai dari pengolahan data mentah hingga menjadi dashboard analitik.

2.5.2. Clustering sebagai Teknik Segmentasi

Secara umum, machine learning pada ranah data mining terbagi menjadi dua kategori utama, yakni supervised learning dan unsupervised learning. Permasalahan yang ada pada Konter Anjas Reparasi tidak memiliki label awal atau target klasifikasi terkait kategori prioritas stok. (Widyawati et al., 2020) menjelaskan bahwa segmentasi atau clustering termasuk dalam kelompok unsupervised learning, di mana algoritma bekerja dengan mengelompokkan data murni berdasarkan kemiripan atau jarak antar objek tanpa panduan variabel target apa pun.

Tujuan utama clustering adalah membentuk kelompok data yang homogen. Hal ini dicapai dengan cara meminimalkan jarak antar objek di dalam cluster yang sama (intra-cluster distance) dan memaksimalkan jarak antar cluster yang berbeda (inter-cluster distance). (Hanum et al., 2023) menegaskan bahwa kualitas cluster yang baik ditandai dengan tingginya keseragaman atribut di dalam suatu kelompok dan adanya perbedaan batasan yang jelas antar kelompok.

Melakukan pengelompokan karakteristik layanan secara manual tentu tidak efektif saat jumlah variabel dan datanya semakin banyak. (Immanuel & Alfian, 2025) menyatakan bahwa metode clustering mampu memfasilitasi segmentasi karakteristik secara otomatis berdasarkan kedekatan nilai atributnya. Hasil akhir dari proses clustering pada dasarnya berupa pemberian label cluster untuk setiap objek data. Label-label ini kemudian dianalisis karakteristik ringkasnya (seperti rata-rata atau median) melalui tahapan pemprofilan cluster (cluster profiling). Sebagaimana ditunjukkan oleh (Hicham & Karim, 2022), pendekatan unsupervised efektif untuk menemukan struktur kelompok data secara mandiri. Oleh karena itu, clustering pada penelitian ini difungsikan untuk

memetakan kasus servis ke dalam beberapa segmen tersembunyi. Profil karakteristik dari segmen tersebut nantinya akan diterjemahkan menjadi kategori prioritas stok sparepart secara objektif. Sebagai bentuk implementasi, teknik clustering yang digunakan dalam penelitian ini adalah algoritma Agglomerative Hierarchical Clustering (AHC) yang penjelasannya dibahas lebih mendalam pada subbab berikutnya.

2.6. Algoritma Agglomerative Hierarchical Clustering (AHC)

2.6.1. Konsep Dasar dan Tahapan Algoritma

Hierarchical clustering adalah metode pengelompokan data yang membangun hierarki kelompok dalam bentuk struktur pohon. Algoritma ini memiliki dua pendekatan utama, yakni *divisive (top-down)* dan *agglomerative (bottom-up)*. Pada penelitian ini, metode yang digunakan adalah *Agglomerative Hierarchical Clustering (AHC)*. Menurut (Widyawati et al., 2020), AHC bekerja dengan prinsip *bottom-up* di mana pada tahap awal setiap objek data individu dianggap sebagai satu kluster mandiri. Secara umum, tahapan algoritma AHC terdiri dari langkah-langkah berikut:

1. Inisialisasi setiap data sebagai satu kluster tunggal.
2. Menghitung matriks jarak antar kluster menggunakan metrik kedekatan tertentu.
3. Menggabungkan dua kluster yang memiliki jarak paling dekat berdasarkan metode tautan (*linkage*) yang dipilih.
4. Memperbarui matriks jarak untuk merepresentasikan kedekatan antar kluster baru.

5. Mengulangi langkah 3 dan 4 hingga seluruh objek menyatu menjadi satu kluster besar.

2.6.2. Pengukuran Kedekatan Jarak (Euclidean Distance)

Langkah krusial dalam AHC adalah menentukan tingkat kemiripan atau jarak antar objek data. Pengukuran jarak yang digunakan dalam penelitian ini adalah *Euclidean Distance*. Menurut (Farhana, 2021), *Euclidean Distance* menghitung jarak garis lurus terpendek antara dua titik dalam ruang multidimensi. Persamaan untuk menghitung jarak *Euclidean* antara objek x dan objek y didefinisikan pada Persamaan (2.1):

$$d(x, y) = \sqrt{\sum_{i=1}^p (x_i - y_i)^2} \quad \dots\dots\dots(2.1)$$

Keterangan:

- a. $d(x, y)$ Jarak *Euclidean* antara objek x dan objek y .
- b. p : Jumlah atribut atau variabel data (*features*).
- c. x_i : Nilai atribut ke- i pada objek x .
- d. y_i : Nilai atribut ke- i pada objek y

Penting untuk diperhatikan bahwa *Euclidean Distance* sangat sensitif terhadap perbedaan skala pada setiap atribut. Mengingat atribut penelitian ini memiliki rentang nilai yang berbeda, seperti biaya servis dan durasi pengerjaan, maka data numerik perlu melalui proses standarisasi (misalnya *Z-score*) sebelum proses *clustering*. Hal ini bertujuan agar perhitungan jarak tidak didominasi oleh variabel

dengan skala nilai yang besar sehingga hasil pengelompokan menjadi lebih objektif.

2.6.3. Metode Penggabungan (Ward's Linkage)

Setelah jarak antar objek dihitung, algoritma AHC membutuhkan metode tautan (*linkage method*) untuk menentukan aturan penggabungan kluster. Penelitian ini menggunakan *Ward's Linkage* atau metode Ward. Menurut (Murtagh & Legendre, 2024), metode Ward umumnya digunakan bersama jarak *Euclidean* karena optimasinya berbasis pada minimisasi varians di dalam kluster yang diukur melalui *Sum of Squared Errors* (SSE). Metode ini bekerja dengan menghitung nilai SSE yang didefinisikan pada Persamaan (2.2):

$$SSE = \sum_{i=1}^n (x_i - \bar{x})^2 \dots\dots\dots(2.2)$$

Keterangan:

- a. SSE : *Sum of Squared Errors* (jumlah kuadrat galat atau varians internal).
- b. n : Jumlah anggota atau objek di dalam kluster.
- c. x_i : Nilai data ke-i di dalam kluster.
- d. \bar{x} : Nilai rata-rata (*mean*) atribut di dalam kluster tersebut.

Pada data multivariat, SSE dihitung berdasarkan seluruh atribut numerik setelah melalui tahap standardisasi. Dengan demikian, penggabungan kluster mempertimbangkan variasi gabungan dari semua atribut secara serentak. Pada setiap iterasi, metode Ward akan mengevaluasi seluruh kemungkinan pasangan kluster dan memilih untuk menggabungkan dua kluster yang menghasilkan kenaikan SSE total paling minimum. Karakteristik ini membuat metode Ward cenderung menghasilkan kluster yang kompak (*compact*) dengan variasi internal yang kecil.

2.6.4. Interpretasi Dendrogram

Seluruh proses penggabungan pada AHC divisualisasikan menggunakan diagram pohon yang disebut *dendrogram*. Menurut (Farhana, 2021), *dendrogram* merupakan representasi grafis yang menunjukkan hubungan kedekatan antar objek serta urutan penggabungan klaster. Sumbu vertikal pada *dendrogram* merepresentasikan tingkat ketidakmiripan atau jarak antar klaster saat digabungkan. Adanya lonjakan garis vertikal yang tinggi mengindikasikan bahwa penggabungan klaster tersebut melibatkan dua kelompok yang sudah memiliki perbedaan karakteristik yang signifikan.

Penentuan jumlah kelompok optimal (k) dilakukan dengan memotong *dendrogram* menggunakan garis ambang batas horizontal (*cut-off line*). Pemotongan ini menghasilkan sejumlah klaster kasus servis yang selanjutnya diprofilkan untuk menentukan komponen dominan serta prioritas pengadaan *sparepart*. Namun, penentuan jumlah klaster pada penelitian ini tidak hanya bergantung pada interpretasi dendrogram semata. Menurut (Gere, 2023), hasil pemotongan tersebut perlu dikonfirmasi menggunakan evaluasi internal seperti *Silhouette Score* dan *Davies-Bouldin Index* yang akan dibahas lebih mendalam pada subbab berikutnya.

2.7. Evaluasi Kualitas Klaster

Evaluasi kualitas klaster pada algoritma yang bersifat *unsupervised learning* dilakukan menggunakan metrik internal karena tidak adanya label target sebagai acuan. Metrik ini berfungsi untuk mengukur tingkat kepadatan objek di dalam kelompok serta jarak pemisah antar kelompok yang berbeda (Shutaywi &

Kachouie, 2021). Penggunaan kombinasi beberapa metrik evaluasi sangat disarankan untuk mendapatkan penilaian performa yang konsisten dan objektif, guna meminimalkan bias yang mungkin timbul jika hanya mengandalkan satu metrik tunggal (Hicham & Karim, 2022). Pada penelitian ini, evaluasi dilakukan menggunakan dua metrik utama, yaitu *Silhouette Coefficient* dan *Davies-Bouldin Index* (DBI).

2.7.1. Silhouette Coefficient

Silhouette Coefficient digunakan untuk mengukur seberapa mirip suatu objek dengan klasternya sendiri dibandingkan dengan klaster lainnya (Shutaywi & Kachouie, 2021). Metrik ini mempertimbangkan aspek kohesi dan separasi untuk menentukan ketepatan penempatan suatu objek dalam kelompok. Nilai *Silhouette Coefficient* didefinisikan pada Persamaan (2.3):

$$s(i) = \frac{b(i)-a(i)}{\max(a(i),b(i))} \dots\dots\dots(2.3)$$

Keterangan:

- a. $s(i)$: Nilai koefisien *silhouette* untuk objek i.
- b. $a(i)$: Rata-rata jarak objek i ke semua objek lain dalam klaster yang sama.
- c. $b(i)$: Rata-rata jarak objek i ke semua objek di dalam klaster tetangga terdekat.

Rentang nilai koefisien ini berada di antara -1 hingga 1. Nilai yang mendekati 1 menunjukkan bahwa objek berada di dalam kelompok yang sangat tepat. Sebaliknya, nilai yang mendekati -1 mengindikasikan adanya kemungkinan kesalahan pengelompokan. Rata-rata nilai *silhouette* dari seluruh objek digunakan untuk menentukan jumlah klaster (k) yang paling optimal.

2.7.2. Davies-Bouldin Index (DBI)

Davies-Bouldin Index merupakan metrik evaluasi yang mengukur rasio antara dispersi intra-kluster dengan jarak antar kluster (Poetri et al., 2024). Metrik ini bertujuan untuk memastikan bahwa setiap kluster memiliki kepadatan yang tinggi dan terpisah secara kontras dari kluster lainnya. Perhitungan nilai DBI didefinisikan pada Persamaan (2.4):

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{j \neq i} \left(\frac{S_i + S_j}{M_{ij}} \right) \dots \dots \dots (2.4)$$

Keterangan:

- a. *DBI* : Nilai *Davies-Bouldin Index*.
- b. *k* : Jumlah kluster.
- c. S_i : Rata-rata jarak objek dalam kluster *i* terhadap titik pusat kluster tersebut.
- d. M_{ij} : Jarak antara titik pusat kluster *i* dan kluster *j*.

Berbeda dengan *Silhouette*, kualitas pengelompokan pada metrik DBI dianggap semakin baik apabila nilainya semakin kecil dan mendekati nol. Nilai DBI yang rendah menandakan bahwa kluster yang terbentuk memiliki struktur yang kompak dan jarak pemisah yang lebar antar kelompok. Penggunaan kedua metrik ini secara bersamaan memberikan validasi matematis yang kuat bagi penentuan kategori prioritas stok *sparepart* di Konter Anjas Reparasi.

2.8. Dashboard Analitik untuk Perencanaan Persediaan

Pemanfaatan hasil analisis data memerlukan media penyampaian yang efektif agar informasi tersebut dapat digunakan secara praktis oleh pemilik usaha. Dalam penelitian ini, hasil pengelompokan kasus servis disajikan melalui sebuah

dashboard analitik yang dirancang khusus untuk mendukung perencanaan stok *sparepart* di Konter Anjas Reparasi.

2.8.1. Defenisi dan Peran Dashboard dalam Pengambilan Keputusan

Dashboard analitik merupakan antarmuka pengguna yang menyajikan representasi visual dari metrik utama secara ringkas dan mudah dipantau. Menurut (Awodiji, 2021), *dashboard* merupakan alat pelaporan *business intelligence* yang mengintegrasikan berbagai indikator kinerja dalam satu layar tunggal sehingga mempermudah pemantauan performa secara cepat. Saat ini, fungsi *dashboard* telah bertransformasi dari sekadar alat pelaporan pasif menjadi platform yang menghasilkan kecerdasan yang dapat ditindaklanjuti (*actionable intelligence*) guna mendukung proses pengambilan keputusan yang lebih dinamis (Hooli, 2025). Bagi Konter Anjas Reparasi, keberadaan *dashboard* ini berfungsi untuk menerjemahkan hasil teknis metode *clustering* menjadi informasi hierarki prioritas pengadaan *sparepart* yang mudah dipahami oleh pemilik usaha.

2.8.2. Digitalisasi Persediaan pada UMKM

Penerapan sistem digital merupakan langkah krusial untuk mengatasi hambatan operasional pada usaha skala kecil. Menurut (Tangon et al., 2025), digitalisasi pencatatan persediaan sangat diperlukan bagi UMKM untuk mendukung pemantauan stok secara berkelanjutan. Implementasi sistem ini bertujuan untuk menyajikan informasi persediaan secara akurat sekaligus meminimalkan risiko kesalahan data yang sering terjadi pada pencatatan manual. Sebagai bentuk implementasi dalam penelitian ini, *dashboard* analitik dibangun berbasis web menggunakan kerangka kerja (*framework*) Flask yang diintegrasikan dengan pustaka Bootstrap untuk memastikan antarmuka yang responsif. Menurut

(Banerjee et al., 2025), sistem pemantauan berbasis web memberikan keunggulan dalam hal aksesibilitas, sehingga pengguna dapat memantau kondisi persediaan secara berkala dengan lebih teratur dan efisien.

2.8.3. Visualisasi Data dan Profil Klaster

Tujuan utama dari visualisasi data adalah mempermudah pengguna dalam menangkap pola dan cerita di balik sekumpulan data yang kompleks. (Rathore et al., 2025) menyatakan bahwa teknik visualisasi yang efektif melalui berbagai jenis diagram memungkinkan pengambil keputusan untuk mendapatkan informasi yang akurat dalam mendukung perencanaan bisnis. Dalam *dashboard* perencanaan persediaan ini, data hasil pengelompokan kasus servis ditampilkan melalui elemen grafis seperti diagram batang dan tabel ringkasan.

Dashboard tersebut menampilkan profil karakteristik dari setiap klaster, daftar komponen dominan, serta rekomendasi prioritas pengadaan berdasarkan frekuensi kemunculan kasus pada klaster yang dinilai kritis (misalnya memiliki durasi dan skor kesulitan lebih tinggi). Visualisasi profil klaster ini membantu pemilik konter untuk mengenali jenis kerusakan yang paling sering terjadi serta menentukan kebijakan pengadaan *sparepart* secara objektif. Dengan demikian, penggunaan visualisasi interaktif diharapkan dapat meminimalkan risiko terjadinya kehabisan stok (*stockout*) maupun penumpukan barang yang berlebih tanpa memerlukan perhitungan optimasi stok otomatis yang kompleks (Sultana, 2025).

2.9. Penelitian Terdahulu

Bagian ini menyajikan perbandingan antara penelitian yang dilakukan saat ini dengan beberapa studi terdahulu yang relevan. Fokus perbandingan terletak pada penerapan metode *clustering* untuk segmentasi data operasional serta implementasi *dashboard* analitik untuk mendukung manajemen persediaan.

Tabel 2.1 Penelitian Terdahulu

No	Judul Penelitian dan Tahun	Konteks dan Metode	Hasil Utama	Gap atau Keterbatasan Penelitian
1	Exploratory Data Analysis (EDA) for Banking and Finance: Unveiling Insights and Patterns(Agarwal et al., 2024)	Exploratory Data Analysis (EDA) untuk pola transaksi.	EDA terbukti efektif dalam mengungkap pola, tren, dan anomali data secara visual.	Landasan penggunaan teknik EDA untuk memahami karakteristik data transaksi.
2	Penerapan Agglomerative Hierarchical Clustering Untuk Segmentasi Pelanggan(Widyawati et al., 2020)	Pengelompokan dataset besar menggunakan algoritma AHC.	AHC terbukti efektif dalam menyusun hierarki data berdasarkan kemiripan atribut.	Keterbatasan pada penggunaan variasi metode, di mana pengujian hanya terbatas pada <i>Average Linkage</i> , <i>Manhattan Distance</i> , dan divalidasi secara tunggal menggunakan <i>Koefisien Silhouette</i>
3	Determining Zoning of Areas Affected by Flood Disasters in Medan City Using Silhouette Coefficient and Davies Bouldin Index Analysis(Poetri et al., 2024)	Evaluasi kualitas kluster menggunakan Silhouette dan DBI.	Penggunaan dua metrik evaluasi menghasilkan penilaian jumlah kelompok yang objektif.	Diterapkan pada data spasial bencana, bukan pada data transaksi operasional persediaan.

4	Analisis Kualitas Pelayanan Dengan Menggunakan Metode Service Quality Guna Meningkatkan Kualitas Pelayanan Pada Irham Ponsel (Hasibuan et al., 2024)	Analisis kualitas layanan pada bisnis perbaikan ponsel.	Mengidentifikasi bahwa durasi dan biaya menjadi faktor dominan dalam kepuasan pelanggan.	Hanya melakukan analisis deskriptif, tidak menggunakan algoritma pengelompokan untuk perencanaan stok.
5	Perancangan Konsep Sistem Pencatatan Persediaan Berbasis Digital Pada UMKM di Kota Manado(Tangon et al., 2025)	Digitalisasi sistem pencatatan persediaan pada UMKM.	Sistem digital mampu meminimalkan kesalahan informasi stok dan mempermudah pemantauan persediaan.	Fokus pada pencatatan stok secara digital namun belum menyentuh aspek analisis prioritas menggunakan machine learning.
6	Artificial Intelligence in Data Visualization: Reviewing Dashboard Design and Interactive Analytics for Enterprise Decision-Making(Sultana, 2025)	Implementasi dashboard interaktif untuk keputusan perusahaan.	Dashboard interaktif mampu meminimalkan risiko operasional melalui penyajian data terstruktur.	Diterapkan pada skala perusahaan besar, bukan pada konteks UMKM jasa reparasi dengan keterbatasan sumber daya.
7	Visualisasi Segmentasi Pelanggan Berdasarkan Atribut RFM Menggunakan Algoritma K-Means Untuk Memahami Karakteristik Pelanggan pada Toko Retail Online(Imanuel & Alfian, 2025)	Visualisasi segmentasi data riwayat transaksi operasional.	Visualisasi sistematis mempermudah pemangku kepentingan dalam memahami perilaku operasional.	Tidak mengintegrasikan hasil segmentasi ke dalam rekomendasi prioritas pengadaan sparepart yang aplikatif.
8	An ensemble agglomerative hierarchical clustering algorithm based on clusters	Ensemble AHC	AHC memiliki stabilitas tinggi	Penguatan metode AHC

	clustering technique and the novel similarity measurement(Li et al., 2022)			
9	Ward's Hierarchical Clustering Method: Clustering Criterion and Agglomerative Algorithm(Murtagh & Legendre, 2024)	Studi Ward Linkage	Ward menghasilkan cluster homogen	Dasar pemilihan Ward Linkage
10	Silhouette analysis for performance evaluation in machine learning with applications to clustering(Shutaywi & Kachouie, 2021)	Evaluasi Clustering	Silhouette efektif mengukur kualitas	Metrik evaluasi
12	Cluster Analysis with Complete Linkage and Ward's Method for Health Service Data in Makassar City(Muthahharah & Juhari, 2021)	Analisis Health Service	Ward stabil untuk dataset kecil	Konteks data UMKM
14	Strategy to Improve Service Quality in Service UMKM At "Berkah Rizki Photocopy" Gunung Pasir Jaya, A Village in East Lampung, Lampung Province(Daud et al., 2024)	Kualitas Layanan UMKM	UMKM butuh klasifikasi layanan	Konteks operasional

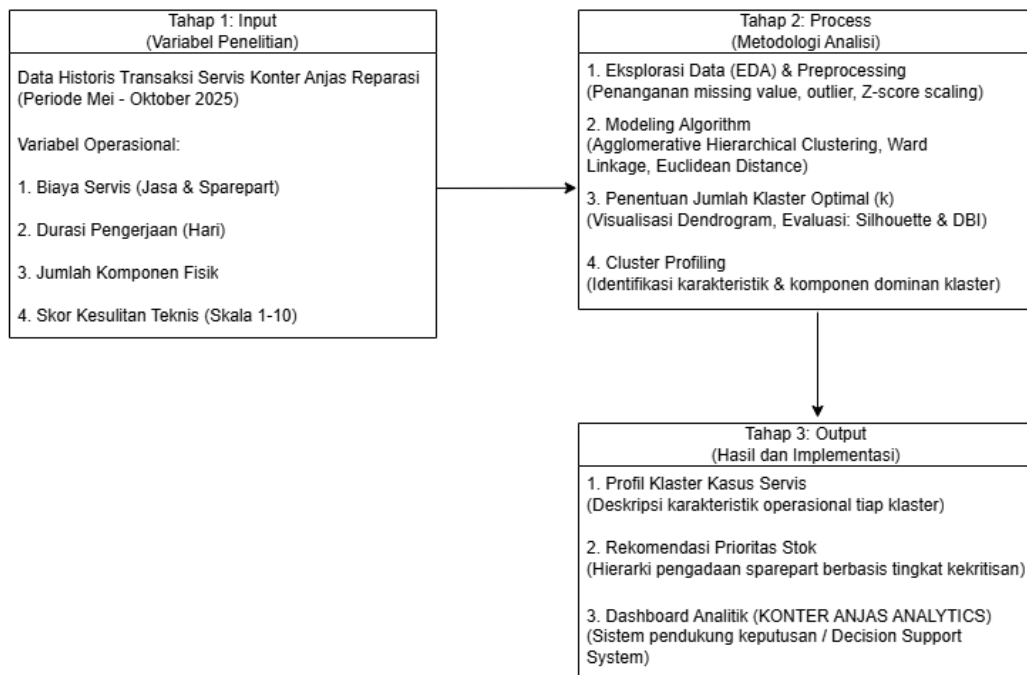
Sebagian besar penelitian terdahulu membahas klasifikasi inventaris berbasis kriteria tertentu atau menerapkan metode *clustering* tanpa mengaitkan hasilnya secara langsung ke prioritas pengadaan pada konteks UMKM jasa. Selain itu, integrasi hasil *clustering* ke dalam *dashboard* operasional yang siap pakai masih terbatas. Oleh karena itu, penelitian ini mengisi celah tersebut dengan:

1. Mengelompokkan kasus servis berbasis multi-kriteria operasional yang mencakup biaya, durasi, jumlah komponen, dan skor kesulitan secara bersamaan.
2. Menurunkan hierarki prioritas sparepart berdasarkan profil klaster dan identifikasi komponen dominan pada setiap kelompok.

Mengimplementasikan hasil analisis ke dalam *dashboard* berbasis web menggunakan *Flask* dan *Bootstrap* untuk mendukung keputusan pengadaan di Konter Anjas Reparasi secara objektif.

2.10. Kerangka Konseptual Penelitian

Kerangka konseptual berikut menggambarkan hubungan antar variabel serta alur logis penelitian dalam menghasilkan Perencanaan Prioritas Stok Sparepart Servis Handphone. Alur ini menitikberatkan pada pengolahan data historis transaksi servis untuk menghasilkan wawasan strategis bagi operasional Konter Anjas Reparasi.



Gambar 2.1 Kerangka Konseptual Penelitian

Hubungan antar elemen dalam kerangka konseptual yang ditunjukkan pada Gambar 2.1 di atas dibagi ke dalam tiga tahap utama, yaitu:

1. Input (Variabel Penelitian)

Bagian ini mencakup data transaksi servis periode Mei hingga Oktober 2025 dengan empat variabel operasional:

- a. **Biaya Servis:** Besaran biaya yang dibayarkan pelanggan untuk jasa dan *sparepart*.
- b. **Durasi Pengerjaan:** Waktu yang dibutuhkan untuk menyelesaikan satu kasus servis.
- c. **Jumlah Komponen:** Total item *sparepart* yang digunakan dalam satu perbaikan.

- d. Skor Kesulitan Teknis: Penilaian tingkat kerumitan pengerjaan berdasarkan kategori servis.
- e. Pemilihan variabel biaya dan durasi didasarkan pada karakteristik utama jasa perbaikan ponsel (Hasibuan dkk., 2024), sedangkan jumlah komponen dan skor kesulitan ditambahkan untuk memperkuat akurasi pemetaan kebutuhan *sparepart*.

2. Process (Metodologi Analisis)

Data diolah melalui tahapan sistematis untuk menjamin validitas hasil pengelompokan:

- a. Eksplorasi Data (EDA) dan *Preprocessing*: Melakukan analisis pola awal serta penyiapan data yang meliputi penanganan *missing value*, pencilan (*outlier*), serta standarisasi fitur (*Z-score*).
- b. *Modeling* (AHC): Penerapan metode *clustering* dengan algoritma *Agglomerative Hierarchical Clustering* menggunakan metode *Ward Linkage* dan *Euclidean Distance*.
- c. Penentuan Jumlah Klaster (*k*): Menentukan jumlah kelompok optimal berdasarkan hasil visualisasi *dendrogram* yang dikonfirmasi dengan memilih konfigurasi yang menghasilkan *Silhouette Score* maksimum dan *DBI* minimum.
- d. *Cluster Profiling*: Tahap analisis untuk mengidentifikasi karakteristik operasional dan komponen *sparepart* dominan yang muncul pada setiap kelompok yang terbentuk.

3. Output (Hasil dan Implementasi)

Hasil akhir dari penelitian ini diterjemahkan ke dalam solusi praktis berupa:

- a. Profil Klaster Kasus Servis: Deskripsi karakteristik operasional setiap klaster berdasarkan nilai ringkas variabel (seperti rata-rata atau median biaya, durasi, jumlah komponen, dan skor kesulitan) beserta interpretasi maknanya.
- b. *Dashboard* Analitik: Antarmuka visual yang menyajikan profil klaster dan daftar komponen dominan secara transparan.
- c. Rekomendasi Prioritas Stok: Penentuan hierarki pengadaan *sparepart* berdasarkan data historis transaksi servis, karakteristik klaster dengan tingkat kekritisian operasional yang tinggi, dan identifikasi komponen dominan yang paling sering muncul guna mendukung keputusan stok secara objektif.

Landasan teori, metode analisis, serta penelitian terdahulu yang telah dibahas menjadi dasar utama dalam penyusunan metodologi penelitian. Pembahasan selanjutnya pada Bab III akan menjelaskan metode penelitian secara rinci sesuai dengan kerangka kerja CRISP-DM.

BAB III

ANALISA DAN PERANCANGAN SISTEM

3.1. Jenis Penelitian

Penelitian ini tergolong ke dalam jenis penelitian kuantitatif terapan (*applied research*) yang dipadukan dengan pendekatan *data mining* eksploratif. Pendekatan ini dipilih untuk menggali pola tersembunyi dari data historis transaksi servis dan menerjemahkannya menjadi solusi praktis berupa Perencanaan Prioritas Stok Sparepart. Metodologi analisis utama yang memandu jalannya penelitian mengacu pada kerangka kerja CRISP-DM (*Cross-Industry Standard Process for Data Mining*). Kerangka ini terdiri dari enam tahapan sistematis, mulai dari pemahaman bisnis hingga penerapan (*deployment*) hasil analisis ke dalam bentuk *dashboard* interaktif (Schröer et al., 2021).

Pada tahap pemodelan data, metode utama yang diterapkan adalah algoritma *Agglomerative Hierarchical Clustering* (AHC). Metode ini membentuk kluster secara bertahap dengan memperlakukan setiap data sebagai kluster tunggal pada awalnya, lalu menggabungkannya berdasarkan tingkat kemiripan atribut operasional. Keunggulan utama AHC adalah kemampuannya dalam menyajikan struktur kedekatan antar-data secara visual melalui *dendrogram*, serta memberikan fleksibilitas dalam menentukan jumlah kluster yang paling optimal di tahap akhir (Rajasekaran et al., 2025). Berbagai kajian juga menegaskan bahwa *clustering* sebagai teknik *unsupervised learning* sangat efektif diterapkan pada data operasional operasional bisnis yang tidak berlabel untuk mengungkap

informasi baru (Rolf et al., 2025), sehingga sangat cocok untuk memetakan prioritas penyediaan komponen di UMKM jasa reparasi (Muthoharoh, 2025).

3.2. Lokasi dan Waktu Penelitian

Penelitian dilakukan pada:

1. Lokasi Penelitian: Penelitian dilakukan di Konter Anjas Reparasi, yang berlokasi di Simpang Mata Pao, Kabupaten Serdang Bedagai, Provinsi Sumatera Utara. Lokasi ini dipilih karena memiliki volume transaksi servis *handphone* yang memadai untuk pemodelan *machine learning* serta telah menerapkan sistem pencatatan riwayat transaksi secara digital.
2. Periode Data: Data sekunder yang digunakan sebagai *dataset* pemodelan merupakan rekapitulasi data historis layanan servis perangkat elektronik (khususnya *handphone*) pada periode operasional Mei hingga Oktober 2025.
3. Waktu Penelitian: Rangkaian kegiatan penelitian yang meliputi tahap pengumpulan data, prapemrosesan, perancangan instrumen, pemodelan algoritma, validasi pakar, hingga implementasi *dashboard* dilakukan pada rentang waktu November 2025 hingga Februari 2026

3.3. Sumber Data dan Teknik Pengumpulan Data

3.3.1. Sumber Data

Sumber data yang digunakan dalam penelitian ini diklasifikasikan menjadi dua jenis, yaitu data primer dan data sekunder, dengan rincian sebagai berikut:

1. Data Primer

Data primer merupakan data yang diperoleh langsung dari sumber asli. Dalam penelitian ini, data primer berupa hasil validasi pakar (*expert judgment*) terkait tingkat kerumitan perbaikan dan penggunaan komponen suku cadang (*sparepart*). Data ini bersumber langsung dari wawancara dengan pemilik sekaligus teknisi ahli di Konter Anjas Reparasi, yang memberikan justifikasi operasional untuk pembentukan instrumen Skor Kesulitan Teknis.

2. Data Sekunder

Data sekunder adalah data yang diperoleh dari sistem arsip yang telah ada. Data sekunder dalam penelitian ini berupa rekam jejak historis transaksi layanan servis di Konter Anjas Reparasi periode Mei hingga Oktober 2025. *Dataset* mentah ini berisi 1.113 baris transaksi dan menjadi *input* utama yang akan diproses menggunakan algoritma *Agglomerative Hierarchical Clustering* (AHC).

Tabel 3.1 Sumber dan Jenis Data

No	Jenis Data	Sumber	Format	Keterangan
1	Validasi Rubrik Kerusakan & Sparepart	Pemilik/Teknisi Ahli Konter Anjas Reparasi	Catatan/Transkrip	Digunakan sebagai dasar konversi teks kerusakan menjadi numerik (Skor Kesulitan)
1	Data Transaksi Servis	Sistem Pencatatan Konter Anjas Reparasi	.csv (per bulan)	Berisi 1.113 baris riwayat transaksi harian
2	Literatur Ilmiah	Jurnal, Prosiding, Buku Teks	PDF/Digital	Landasan teori CRISP-DM, Algoritma AHC, dan Manajemen Rantai Pasok

3.3.2. Teknik Pengumpulan Data

Untuk memperoleh data yang akurat dan relevan dengan tujuan perancangan *dashboard* prioritas stok *sparepart*, teknik pengumpulan data yang dilakukan meliputi:

1. Studi Dokumentasi

Mengunduh dan mengekspor data transaksi layanan servis periode Mei–Oktober 2025 dari sistem komputer konter. Data tersebut kemudian disimpan dan digabungkan ke dalam format .csv untuk mempermudah proses impor data ke dalam bahasa pemrograman Python.

2. Wawancara (Interview)

Teknik wawancara dilakukan secara langsung dengan pemilik dan teknisi di Konter Anjas Reparasi. Wawancara ini difokuskan pada dua tujuan:

- a. Mengidentifikasi masalah operasional terkait pengelolaan stok *sparepart* (seperti risiko *dead-stock* dan *indent part*).
- b. Memvalidasi instrumen pembobotan (rubrik) untuk menerjemahkan teks deskripsi kerusakan menjadi angka metrik Skor Kesulitan Teknis agar sejalan dengan logika teknis di lapangan.

3.3.3. Sruktur Data Transaksaksi

Agar proses pembersihan data (*Data Preparation*) dapat dilakukan dengan tepat, langkah pertama adalah mendefinisikan struktur *database* transaksi mentah yang telah diekstraksi. Berdasarkan hasil pembacaan sistem, data mentah

memiliki 14 atribut/kolom awal. Rincian struktur tersebut disajikan pada Tabel 3.2 berikut:

Tabel 3.2 Struktur Data Transaksi Konter Anjars Reparasi

No	Nama Kolom	Tipe Data	Deskripsi	Contoh
1	NO	Integer	Nomor urut transaksi	1, 2, 3, ...
3	TANGGAL MASUK	Date	Tanggal perangkat masuk	"15/05/2025"
4	NAMA	String	Nama pelanggan	"Budi Santoso"
5	NO HP	String	Nomor telepon pelanggan	"081234567890"
6	ALAMAT	String	Alamat pelanggan	"Jl. Merdeka No. 10"
7	TYPE HP	String	Tipe/model perangkat	"iPhone 12 Pro", "Samsung A52"
8	KERUSAKAN	String	Deskripsi kerusakan	"LCD Pecah", "Mati Total"
9	TEKNISI	String	Nama teknisi pengerjaan	"Teknisi A", "Teknisi D"
10	BIAYA	Integer/Float	Total biaya servis (IDR)	250000, 500000
11	KETERANGAN	String	Status pembayaran	"Lunas", "DP"
12	TANGGAL KELUAR	Date	Tanggal perangkat selesai/diambil	"18/05/2025"
13	Unnamed: 12	String/Object	Kolom kosong/sisa format tabel	NaN
14	Unnamed: 14	String/Object	Kolom kosong/sisa format tabel	NaN

Tabel 3.2 memperlihatkan bahwa data mentah sangat bervariasi dan memuat anomali berupa kolom yang tidak relevan (seperti Unnamed: 12). Dalam penelitian ini, atribut yang mengandung informasi pribadi pelanggan (NAMA, NO HP, ALAMAT), serta kolom yang tidak memiliki nilai informatif, tidak akan

dilibatkan dalam proses *clustering*. Atribut teknis (Kerusakan, Tanggal, Biaya) akan menjadi fokus utama untuk ditransformasi pada tahap selanjutnya.

3.4. Defenisi Operasional Variabel dan Instrumen

3.4.1. Variabel Input

Berdasarkan struktur data mentah yang telah dieksplorasi dan dibersihkan, dipilih sejumlah variabel spesifik untuk dijadikan *input* (masukan) ke dalam algoritma *Agglomerative Hierarchical Clustering* (AHC). Variabel-variabel ini dipilih secara khusus karena memiliki representasi langsung terhadap tingkat kerumitan servis dan kebutuhan operasional suku cadang (*sparepart*). Definisi operasional untuk setiap variabel *input* dijelaskan secara rinci pada Tabel 3.3.

Tabel 3.3 Variabel Input

No	Variabel	Definisi Operasional	Skala Pengukuran
1	Biaya Servis	Total nominal biaya layanan yang dikenakan kepada pelanggan, mencakup jasa teknisi dan harga <i>sparepart</i> , dihitung dalam satuan Rupiah (IDR).	Skala Rasio
2	Durasi Pengerjaan	Waktu aktual yang dihabiskan untuk menyelesaikan servis, dihitung dari selisih tanggal masuk dan tanggal keluar dalam satuan hari.	Skala Rasio
3	Jumlah Komponen	Total akumulasi item fisik <i>sparepart</i> yang digunakan/diganti dalam satu nomor riwayat transaksi servis.	Skala Rasio
4	Skor Kesulitan	Nilai kuantitatif (1-10) yang merepresentasikan tingkat kerumitan dan risiko perbaikan, didapatkan dari konversi data teks menggunakan instrumen rubrik pakar.	Skala Interval

5	Aksi Servis	Kategori tindakan utama yang dilakukan teknisi (misal: 'Ganti', 'Hardware_Repair', 'Software') yang telah ditransformasi menjadi vektor biner (0/1) menggunakan teknik <i>One-Hot Encoding</i> (OHE).	Skala Nominal
6	Komponen Utama	Kategori jenis suku cadang yang paling dominan diganti (misal: 'LCD', 'Baterai', 'IC') yang telah ditransformasi menjadi vektor biner (0/1) menggunakan teknik <i>One-Hot Encoding</i> (OHE).	Skala Nominal

Seluruh variabel numerik pada Tabel 3.3 ini akan melalui tahap *preprocessing* berupa standardisasi (*Z-Score Scaling*) sebelum dieksekusi oleh model AHC untuk memastikan perhitungan metrik jarak berjalan dengan seimbang.

3.4.2. Instrumen Penelitian (Rubrik Skor Kesulitan)

Mengingat algoritma AHC mensyaratkan data dalam bentuk numerik, data tekstual dari catatan konter tidak dapat diproses secara langsung. Oleh karena itu, penelitian ini menggunakan instrumen berupa Rubrik Expert Judgment untuk mengonversi data deskripsi kerusakan menjadi variabel SKOR_KESULITAN.

Instrumen rubrik ini disusun dan divalidasi melalui wawancara langsung dengan teknisi ahli di Konter Anjas Reparasi. Skala penilaian ditetapkan dari angka 1 (paling ringan) hingga 10 (paling berat/berisiko), dengan pedoman klasifikasi sebagai berikut:

1. Skor 1 - 3 (Ringan/Non-Fisik): Pengecekan ringan, pembersihan, perbaikan *software* dasar, atau kasus yang tidak memerlukan pembongkaran perangkat keras secara fatal.
2. Skor 4 - 6 (Menengah/Modular): Pergantian suku cadang *fast-moving* yang bersifat modular (tinggal pasang), seperti penggantian LCD, Baterai, Kamera, atau *Speaker*.
3. Skor 7 - 10 (Berat/Mikro-Komponen): Perbaikan *hardware* tingkat lanjut yang membutuhkan penyolderan, pengangkatan IC (seperti IC *Power* / EMMC), perbaikan jalur *motherboard*, atau kasus perangkat mati total (matot).

Instrumen ini memastikan bahwa variabel masukan bagi model algoritma tetap mewakili kondisi dan logika keteknikan yang riil di lapangan.

3.4.3. Variabel Output

Selain menetapkan variabel *input*, penelitian ini juga mendefinisikan variabel *output* yang merupakan hasil akhir dari proses pengolahan algoritma *data mining*. Variabel *output* ini berfungsi sebagai luaran fungsional untuk *dashboard* serta indikator keberhasilan model. Rincian variabel *output* disajikan pada Tabel 3.4.

Tabel 3.4 Definisi Operasional Variabel Output

No	Variabel	Definisi Operasional	Skala Pengukuran
1	Label Klaster	Identitas kelompok kasus servis yang dihasilkan dari algoritma AHC (<i>k</i> optimal), yang merepresentasikan tingkat prioritas penanganan dan pengadaan stok <i>sparepart</i> .	Skala Nominal
2	Metrik Evaluasi Algoritma	Nilai validasi statistik berupa <i>Silhouette Score</i> dan <i>Davies-</i>	Skala Rasio

		<i>Bouldin Index</i> (DBI) yang memastikan kualitas dan kepadatan pemisahan kluster.	
--	--	--	--

Seperti yang tertera pada Tabel 3.4, *output* utama penelitian ini adalah Label Kluster yang membagi seluruh riwayat transaksi ke dalam beberapa kelompok karakteristik yang paling optimal. Hasil penentuan jumlah kluster terbaik beserta pelabelannya akan dibahas dan dianalisis lebih lanjut pada Bab IV untuk mengekstrak rekomendasi pengadaan *sparepart* spesifik dari masing-masing kelompok.

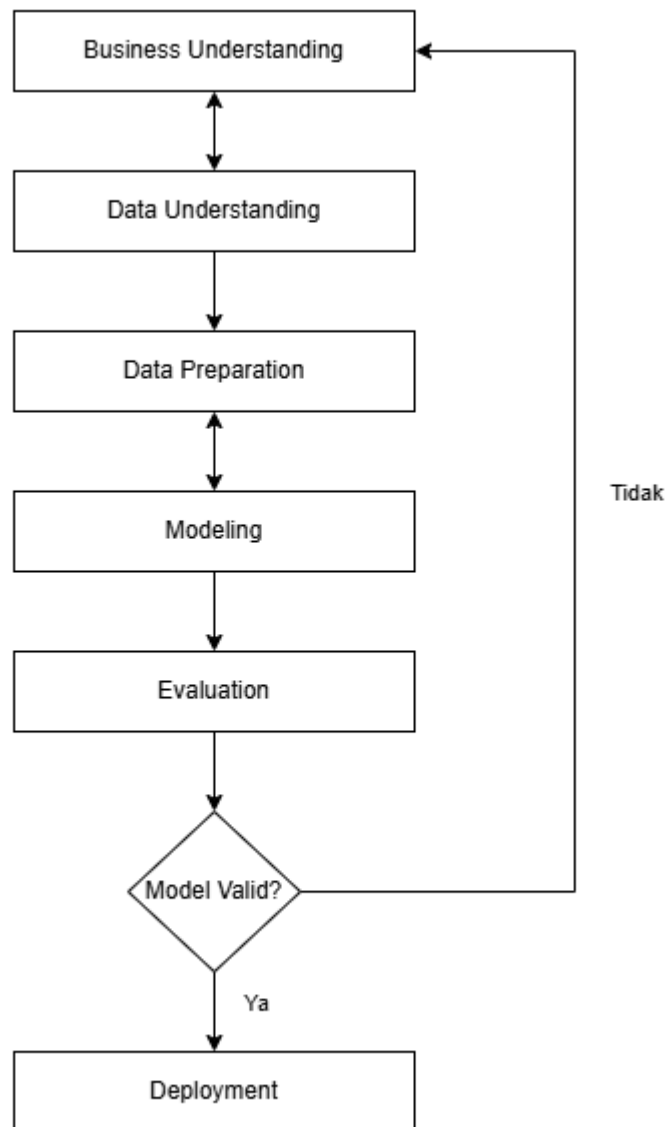
3.5. Analisis Sistem (Metode CRISP-DM)

Prosedur analisis data dalam penelitian ini mengacu pada kerangka kerja CRISP-DM (*Cross-Industry Standard Process for Data Mining*), yang terdiri atas enam fase utama: *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation*, dan *deployment*. Model ini digunakan karena memberikan struktur proses yang sistematis dan iteratif dalam mengubah data transaksi mentah menjadi wawasan bisnis yang dapat diimplementasikan.

Literatur CRISP-DM menekankan bahwa pengumpulan, eksplorasi, serta pemeriksaan kualitas data merupakan tugas-tugas esensial pada fase *data understanding* (Schröer et al., 2021). Selain itu, (Schröer et al., 2021) menyoroti bahwa fase *deployment* sering kali belum sepenuhnya terintegrasi ke lingkungan produksi dalam banyak proyek *data mining*. Oleh karena itu, penelitian ini dirancang untuk menutup celah tersebut dengan mengeksekusi fase *deployment* secara nyata, yaitu melalui perancangan sistem *dashboard* perencanaan stok

sparepart agar hasil pemodelan dapat langsung digunakan dalam operasional sehari-hari di Konter Anjas Reparasi.

Secara garis besar, hubungan antar-fase dalam penelitian ini dapat digambarkan menggunakan diagram sebagai berikut:



Gambar 3.1 Siklus CRISP-DM

Gambar 3.1 di atas mengilustrasikan siklus hidup proyek *data mining* yang diterapkan dalam penelitian ini. Penting untuk dicatat bahwa alur proses ini tidak bersifat linear kaku, melainkan iteratif (berulang) dan adaptif. Penjelasan detail mengenai hubungan antar-fase dalam diagram tersebut adalah sebagai berikut:

1. Siklus Awal (*Business & Data Understanding*): Panah dua arah antara *Business Understanding* dan *Data Understanding* menunjukkan hubungan timbal balik yang erat. Awalnya, penelitian mendefinisikan masalah bisnis utama, yaitu kesulitan penentuan prioritas pengadaan stok *sparepart*. Namun, setelah melihat kondisi data mentah di tahap *Data Understanding* (eksplorasi awal), pemahaman terhadap masalah bisnis dapat dipertajam atau disesuaikan kembali dengan ketersediaan fitur data historis di lapangan.
2. Siklus Teknis (*Data Preparation & Modeling*): Hubungan antara *Data Preparation* dan *Modeling* juga bersifat bolak-balik. Data yang telah dibersihkan dan ditransformasi akan dimasukkan ke dalam model algoritma AHC. Jika hasil *modeling* menunjukkan struktur kluster yang buruk atau tumpang tindih, peneliti harus kembali ke tahap *Data Preparation* untuk memperbaiki metode *cleaning*, melakukan standarisasi fitur (*Z-score*) ulang, atau mengevaluasi kembali instrumen *Skor Kesulitan*.
3. Evaluasi dan Umpan Balik: Tahap *Evaluation* berfungsi sebagai gerbang penentu (*decision gate*). Jika hasil evaluasi (metrik *Silhouette Score* dan DBI) dinilai baik dan menjawab rumusan masalah, proses berlanjut ke

tahap *Deployment*. Namun, panah yang mengarah kembali ke *Business Understanding* menunjukkan bahwa jika hasil profil kluster ternyata tidak dapat diterjemahkan menjadi rekomendasi pengadaan *sparepart* yang relevan secara operasional bagi Konter Anjas Reparasi, maka seluruh siklus harus ditinjau ulang dari awal.

4. *Deployment*: Tahap akhir ini bukanlah akhir dari siklus data, melainkan implementasi hasil model ke dalam lingkungan produksi. Dalam penelitian ini, *deployment* diwujudkan berupa sistem *dashboard* analitik perencanaan stok *sparepart* yang dirancang untuk mendukung pengambilan keputusan dalam operasional sehari-hari pengguna.

3.5.1. Business Understanding

Fase *business understanding* berfokus pada pemahaman konteks bisnis dan permasalahan operasional yang dihadapi Konter Anjas Reparasi terkait manajemen persediaan, sehingga tujuan analisis data dapat ditentukan secara jelas dan terarah.

1. Identifikasi Masalah Operasional

Permasalahan utama yang dihadapi oleh pihak konter dalam mengelola persediaan komponen antara lain:

- a. Belum adanya sistem klasifikasi kasus servis yang terstruktur untuk menentukan prioritas pengadaan *sparepart*.
- b. Penentuan stok *sparepart* masih bergantung pada insting atau estimasi subjektif pemilik, yang berisiko memunculkan

penumpukan barang yang jarang terpakai (*dead-stock*) atau justru kekosongan komponen krusial.

- c. Tingginya variasi waktu pengerjaan akibat kasus komponen kosong (*indent part*) yang dapat menurunkan kepuasan pelanggan.
- d. Belum dimanfaatkannya data historis riwayat transaksi secara optimal untuk memetakan pola kebutuhan komponen berbasis data (*data-driven*).

2. Penentuan Indikator Prioritas Stok

Berdasarkan rumusan masalah dan landasan teori pada Bab I dan Bab II, indikator operasional yang digunakan dalam penelitian ini difokuskan pada atribut teknis pengerjaan. Mengacu pada penelitian (Hasibuan et al., 2024), biaya dan durasi merupakan faktor operasional paling dominan yang menentukan karakteristik layanan perbaikan handphone. Oleh karena itu, penelitian ini menetapkan enam indikator utama, yaitu:

- a. Biaya Servis: Total nominal yang dibayarkan pelanggan untuk jasa perbaikan dan/atau harga sparepart.
- b. Durasi Pengerjaan: Waktu yang dihabiskan dari perangkat masuk hingga selesai diperbaiki (dalam satuan hari).
- c. Jumlah Komponen: Total item fisik sparepart yang diganti dalam satu riwayat transaksi.
- d. Skor Kesulitan Teknis: Tingkat kerumitan dan risiko pengerjaan yang dikuantifikasi menjadi rentang skor (1-10) berdasarkan instrumen pakar.

- e. Aksi Servis: Jenis tindakan yang dilakukan (misalnya: Ganti *Sparepart*, Perbaiki *Hardware*, atau *Software*).
- f. Komponen Utama: Nama suku cadang dominan yang diperbaiki atau diganti (misalnya: LCD, IC, Baterai).

Keenam indikator ini akan ditransformasikan ke dalam format numerik dan digunakan sebagai *input* algoritma AHC.

3.5.2. Data Understanding

Tahap Data Understanding bertujuan memahami kondisi awal dataset transaksi servis sebelum dilakukan pembersihan dan rekayasa fitur. Pada tahap ini dilakukan Exploratory Data Analysis (EDA) menggunakan bahasa pemrograman Python untuk mengevaluasi struktur data, kelengkapan atribut, tipe data, sebaran nilai, serta potensi anomali. Hasil EDA digunakan sebagai dasar keputusan pada fase Data Preparation, agar dataset yang digunakan untuk pengelompokan (clustering) memiliki kualitas yang memadai dan tidak bias oleh kesalahan input maupun outlier ekstrem.

1. Profil dan Struktur Data Mentah (Raw Data)

Dataset mentah diperoleh dari pencatatan Konter Anjas Reparasi dengan total 1.113 transaksi dan 14 kolom atribut. Kolom pada data mentah meliputi: NO, STATUS, TANGGAL_MASUK, NAMA, NO HP, ALAMAT, TYPE HP, KERUSAKAN, TEKNISI, BIAYA, KETERANGAN, TANGGAL_KELUAR, Unnamed: 12, dan Unnamed: 14.

Tabel 3.5 Profil Struktur Data Mentah Transaksi Servis

Atribut Informasi	Keterangan
Total Transaksi (Baris)	1.113 data transaksi
Total Variabel (Kolom)	14 kolom atribut

Kolom Atribut Mentah	NO, STATUS, TANGGAL_MASUK, NAMA, NO HP, ALAMAT, TYPE HP, KERUSAKAN, TEKNISI, BIAYA, KETERANGAN, TANGGAL_KELUAR, Unnamed: 12, Unnamed: 14
----------------------	--

Berdasarkan pemeriksaan tipe data, sebagian besar kolom masih bertipe object, termasuk kolom BIAYA yang seharusnya numerik. Selain itu, ditemukan kolom tidak relevan (Unnamed: 12 dan Unnamed: 14) yang hampir seluruhnya kosong, sehingga perlu dihapus pada tahap Data Preparation.

2. Kualitas Data dan Missing Value

Hasil pemeriksaan kelengkapan data menunjukkan adanya missing value pada beberapa kolom, khususnya NO HP (537 kosong), ALAMAT (330 kosong), TEKNISI (220 kosong), BIAYA (183 kosong), serta TANGGAL_KELUAR (224 kosong). Missing value pada NO HP dan ALAMAT tidak mengganggu tujuan penelitian karena atribut tersebut tidak digunakan sebagai fitur utama untuk pengelompokan kasus servis dan prioritas stok. Namun, missing value pada BIAYA serta TANGGAL_MASUK/TANGGAL_KELUAR berpengaruh terhadap pembentukan variabel numerik turunan (misalnya durasi pengerjaan), sehingga perlu ditangani agar kualitas dataset tetap terjaga.

Temuan ini menjadi dasar keputusan pembersihan data, yaitu: (1) menghapus kolom tidak relevan, (2) mengonversi tipe data BIAYA menjadi numerik, serta (3) menangani nilai kosong pada atribut yang digunakan dalam analisis.

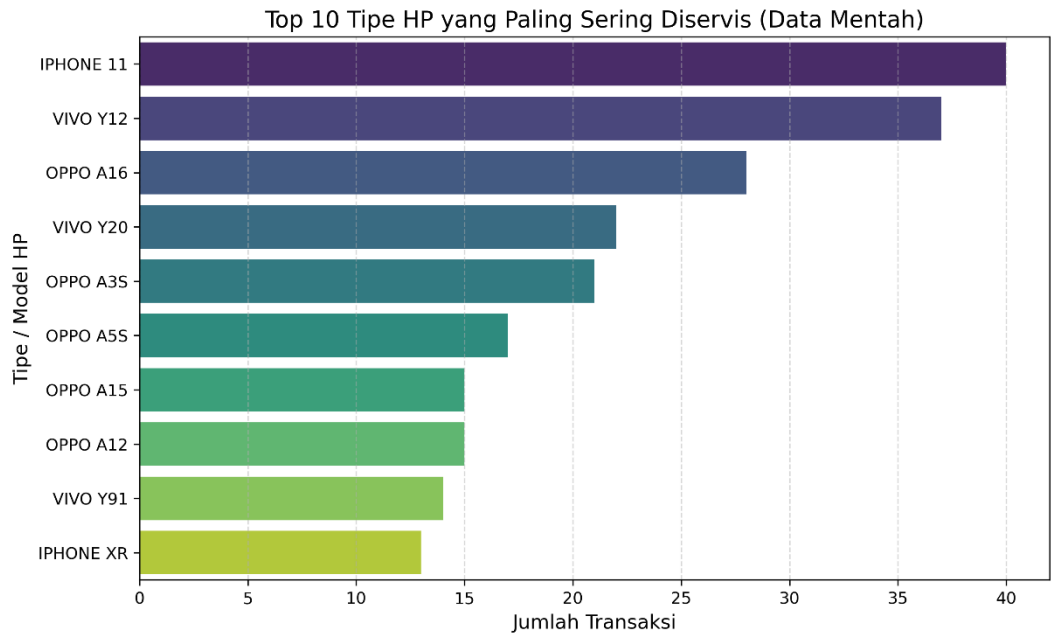
3. Eskplorasi Variabel Kategorikal

Analisis deskriptif dilakukan pada variabel kategorikal untuk memahami keragaman perangkat dan jenis kasus servis yang terjadi.

Tabel 3.6 Statistik Deskriptif Variabel Kategorikal

Variabel	Jumlah Entri (Count)	Nilai Unik (Unique)	Data Terbanyak (Top)	Frekuensi Muncul (Freq)
TYPE HP	962	326	IPHONE 11	40
KERUSAKAN	960	263	GANTI LCD	184
TEKNISI	893	11	Teknisi D	691

Berdasarkan Tabel 3.6, terlihat bahwa variasi penulisan KERUSAKAN sangat tinggi (263 variasi), sehingga diperlukan normalisasi/standarisasi agar variasi teks kerusakan dapat dirangkum menjadi kategori yang lebih seragam. Selain itu, dominasi teknisi tertentu (teknisi D) menunjukkan distribusi pekerjaan yang tidak merata pada data mentah, sehingga pada tahap Data Preparation perlu dipastikan bahwa atribut teknisi digunakan secara tepat (misalnya untuk analisis deskriptif operasional, bukan sebagai penentu utama prioritas stok).



Gambar 3.2 Distribusi Jenis Kerusakan dan Tipe HP Terbanyak

Grafik menunjukkan tipe handphone yang paling sering masuk servis. Dominasi tipe HP tertentu (misalnya IPHONE 11) dan jenis kerusakan dominan (misalnya GANTI LCD) mengindikasikan area fokus utama konter dalam operasional servis, serta menjadi konteks penting dalam interpretasi kebutuhan sparepart. Namun, TYPE HP tidak digunakan sebagai fitur utama clustering karena jumlah kategorinya sangat banyak (326 tipe) dan lebih sesuai ditampilkan sebagai informasi pendukung pada dashboard.

4. Eksplorasi Variabel Numerik dan Deteksi Anomali

Pada tahap ini dilakukan analisis pada variabel numerik utama, yaitu biaya servis dan durasi pengerjaan (hasil turunan dari tanggal masuk dan tanggal keluar).

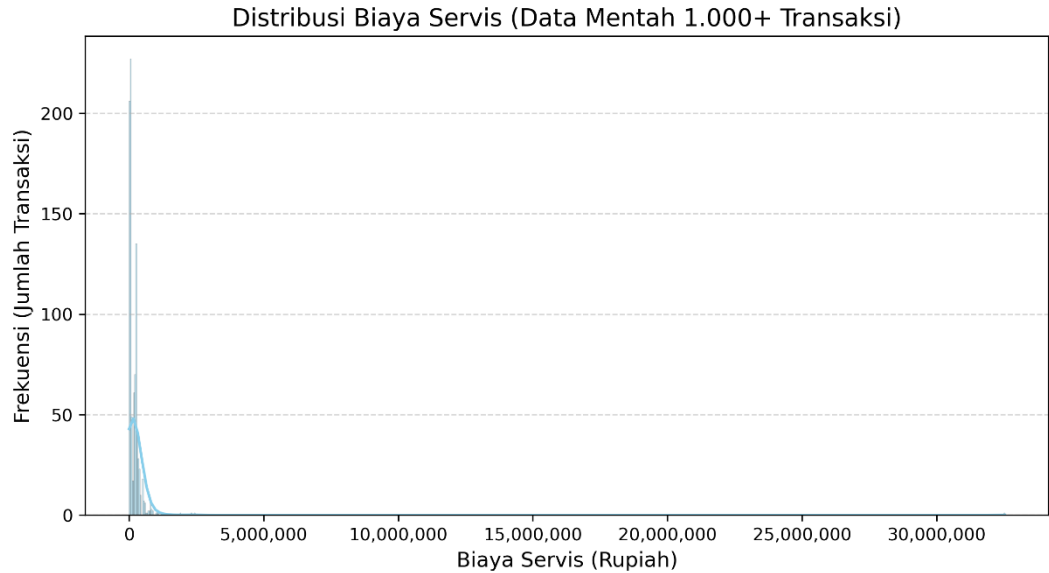
Tabel 3.7 Statistik Deskriptif Variabel Numerik (Data Mentah)

Metrik Statistik	Biaya Servis (Rp)	Durasi Pengerjaan (Hari)
-------------------------	--------------------------	---------------------------------

Rata-rata (Mean)	218.558,30	1,14
Standar Deviasi	1.084.083,60	6,33
Nilai Minimum	0,00	-1,00
Median (Kuartil ke-2/ 50%)	100.000,00	0,00
Nilai Maksimum	32.540.000,00	111,000

Hasil pada Tabel 3.7 menunjukkan adanya kondisi ekstrem (outlier) dan anomali pada data mentah:

- Durasi median 0 hari menunjukkan banyak transaksi selesai di hari yang sama, namun terdapat durasi maksimum 111 hari dan nilai durasi negatif (-1 hari) yang mengindikasikan kesalahan input tanggal.
- Biaya maksimum Rp32.540.000 menunjukkan outlier yang sangat ekstrem dan perlu diverifikasi/ditangani agar tidak mendominasi proses clustering.



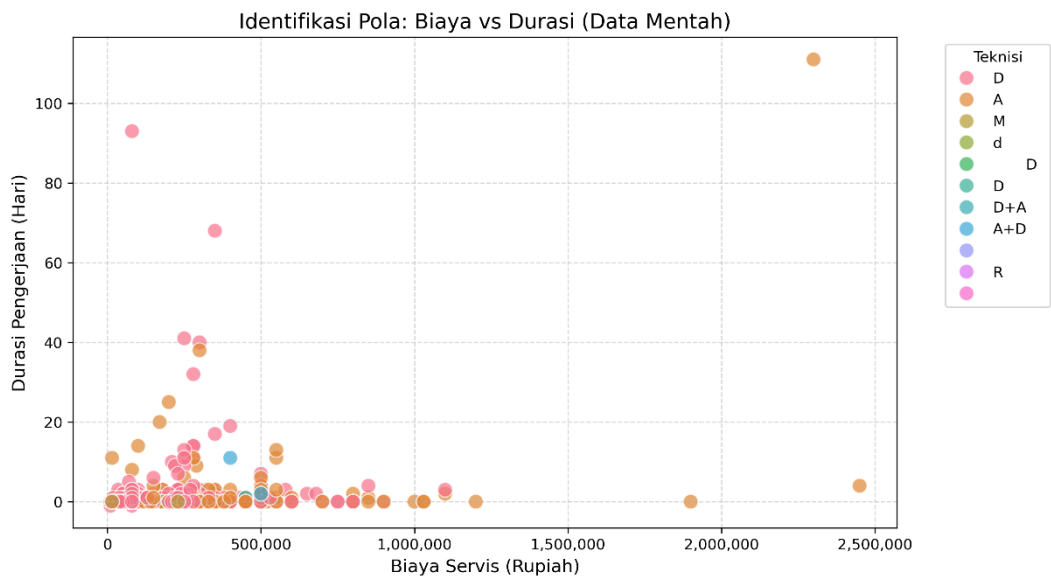
Gambar 3.3 Distribusi Biaya Servis (Data Mentah 1.000+ Transaksi)

Distribusi biaya terlihat sangat condong ke kanan (right-skewed), dengan mayoritas transaksi berada pada rentang biaya rendah–menengah, namun

terdapat beberapa nilai ekstrem yang jauh lebih besar. Pola ini menegaskan perlunya penanganan outlier dan standardisasi skala variabel numerik pada tahap Data Preparation agar proses clustering tidak bias terhadap nilai ekstrem.

5. Identifikasi Pola Biaya vs Durasi

Untuk melihat keterkaitan antara biaya servis dan durasi pengerjaan, dilakukan visualisasi scatter plot.



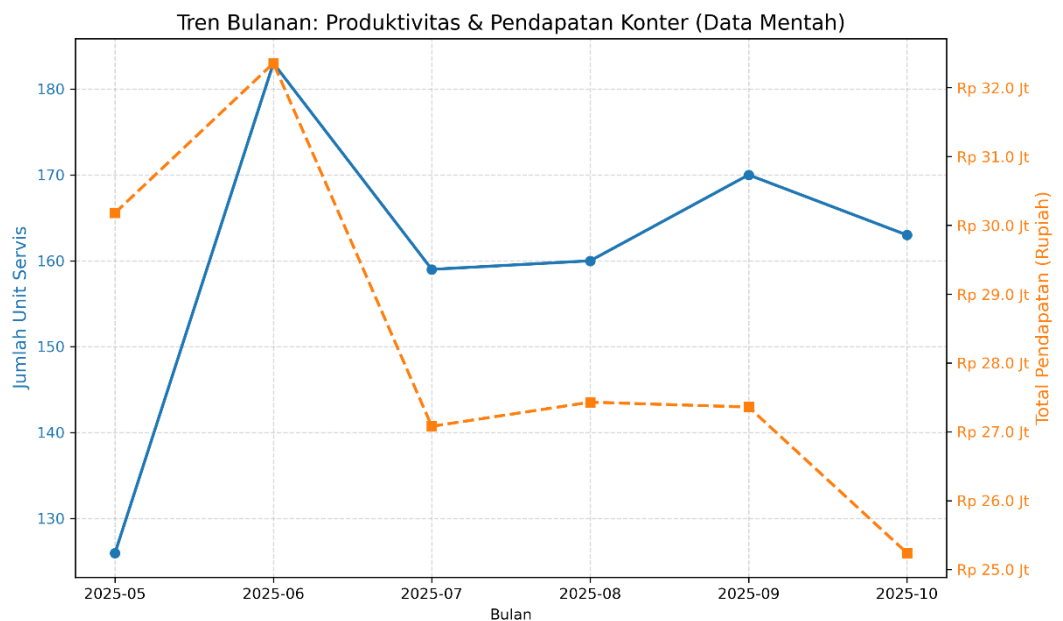
Gambar 3.4 Identifikasi Pola: Biaya vs Durasi (Data Mentah)

Scatter plot memperlihatkan sebagian besar transaksi terkonsentrasi pada durasi pendek dengan biaya rendah–menengah. Namun terdapat kelompok titik dengan durasi pengerjaan sangat panjang yang menyimpang dari mayoritas transaksi. Kelompok ini kemungkinan merepresentasikan kasus indent atau hambatan pasokan sparepart, sehingga relevan dalam pemetaan prioritas stok dan memperkuat alasan penggunaan durasi sebagai variabel penting dalam pembentukan cluster. Temuan ini juga menegaskan perlunya

validasi durasi negatif serta penanganan outlier biaya agar hasil clustering lebih stabil dan dapat diinterpretasikan.

6. Eksplorasi Tren Waktu (Produktivitas & Pendapatan)

Selain karakteristik transaksi per baris, dilakukan analisis tren bulanan untuk memahami dinamika jumlah transaksi dan pendapatan konter.



Gambar 3.5 Tren Bulanan: Produktivitas & Pendapatan Konter (Data Mentah)

Grafik menunjukkan fluktuasi jumlah unit servis dan total pendapatan per bulan pada periode pengamatan. Fluktuasi ini mengindikasikan adanya dinamika musiman atau kejadian operasional tertentu yang perlu diwaspadai dalam perencanaan operasional dan dapat ditampilkan sebagai informasi pemantauan pada dashboard. Analisis tren ini bersifat pendukung dan tidak digunakan sebagai fitur utama clustering.

7. Ringkasan Temuan Data Understanding dan Implikasi ke Data Preparation

Berdasarkan hasil EDA pada data mentah, diperoleh temuan utama sebagai berikut:

- a. Kualitas data bervariasi, ditandai missing value pada beberapa atribut serta adanya kolom tidak relevan (Unnamed).
- b. Keragaman teks KERUSAKAN sangat tinggi (263 variasi), sehingga perlu normalisasi dan rekayasa fitur agar representasi kasus servis lebih seragam.
- c. Terdapat anomali/outlier pada variabel numerik (durasi negatif, durasi ekstrem, dan biaya ekstrem) yang berpotensi mengganggu hasil clustering.
- d. Pola biaya–durasi menunjukkan adanya kelompok kasus yang menyimpang (durasi sangat panjang) sehingga variabel durasi relevan sebagai pembeda cluster kasus servis.

Temuan-temuan tersebut menjadi dasar penerapan pembersihan data dan rekayasa fitur pada tahap Data Preparation, meliputi penghapusan kolom tidak relevan, konversi tipe data biaya ke numerik, penanganan missing value pada atribut kunci, koreksi durasi negatif, penanganan outlier, serta standardisasi skala variabel sebelum pemodelan clustering.

3.5.3. Data Preparation

Tahap Data Preparation bertujuan menyiapkan dataset transaksi servis agar layak digunakan pada pemodelan Agglomerative Hierarchical Clustering (AHC). Berdasarkan hasil Data Understanding, data mentah memiliki beberapa isu utama, yaitu adanya kolom tidak relevan, missing value pada atribut kunci, tipe data yang belum sesuai (khususnya BIAYA), variasi penulisan kerusakan yang tinggi, serta

anomali pada variabel durasi dan biaya. Oleh karena itu, tahap ini dilakukan melalui proses pembersihan data (data cleaning), transformasi, dan rekayasa fitur (feature engineering) sehingga dihasilkan dataset akhir yang siap digunakan untuk pengelompokan kasus servis dan pemetaan prioritas stok sparepart.

1. Seleksi Atribut dan Penghapusan Kolom Tidak Relevan

Langkah awal adalah menyaring atribut yang relevan terhadap tujuan penelitian. Kolom yang tidak relevan dihapus, terutama:

- a. Unnamed: 12 dan Unnamed: 14 (kolom kosong/tidak memiliki makna operasional),
- b. atribut identitas pelanggan seperti NAMA, NO HP, dan ALAMAT, karena tidak diperlukan untuk pengelompokan kasus servis maupun pemetaan sparepart.

Kolom operasional yang dipertahankan sebagai dasar pengolahan meliputi TANGGAL_MASUK, TANGGAL_KELUAR, TYPE HP, KERUSAKAN, TEKNISI, BIAAYA, dan KETERANGAN.

2. Penanganan Missing Value dan Validasi Data

Penanganan data kosong dilakukan dengan mempertimbangkan kebutuhan pembentukan fitur:

- a. Baris yang tidak memiliki informasi minimum untuk pengolahan (misalnya tanggal tidak valid/ tidak bisa diproses) dikeluarkan dari dataset.

- b. Atribut yang bersifat opsional dan tidak digunakan sebagai fitur utama (misalnya TEKNISI untuk analisis operasional) dapat diberi label “Tidak diketahui” atau dibiarkan sesuai kebutuhan dashboard.

Pada tahap ini dilakukan pula validasi awal agar nilai-nilai ekstrem dapat ditelusuri pada tahap berikutnya (misalnya biaya dan durasi yang tidak wajar).

3. Konversi Tipe Data dan Standardisasi Format

Agar dapat diproses secara komputasional, dilakukan transformasi tipe data:

- a. Konversi Tanggal:

TANGGAL_MASUK dan TANGGAL_KELUAR dikonversi menjadi format datetime.

- b. Konversi Biaya:

Kolom BIAYA yang semula bertipe object dibersihkan dari karakter non-numerik, kemudian dikonversi menjadi numerik.

- c. Normalisasi teks:

Variabel teks seperti TYPE HP, KERUSAKAN, dan KETERANGAN dinormalisasi (misalnya kapitalisasi, penghapusan spasi ganda) untuk mempermudah proses ekstraksi fitur.

4. Rekayasa Fitur (Feature Engineering)

- a. Pembentukan DURASI_HARI

Durasi pengerjaan dihitung dari selisih tanggal keluar dan tanggal masuk (DURASI_HARI = TANGGAL_KELUAR - TANGGAL_MASUK).

Berdasarkan hasil eksplorasi data (EDA), ditemukan dua jenis anomali pada

variabel ini yang memerlukan aturan penanganan (*rule-based handling*) secara tegas:

- a) Durasi Negatif (< 0 hari): Kasus dengan nilai durasi negatif diidentifikasi sebagai murni *human error* (kesalahan input tanggal oleh penjaga konter) atau transaksi yang dibatalkan. Aturan penanganannya adalah dihapus (*drop baris*) secara permanen dari *dataset* pemodelan agar tidak merusak perhitungan jarak *Euclidean*.
- b) Durasi Ekstrem (Puluhan hingga ratusan hari): Berbeda dengan durasi negatif, kasus dengan durasi sangat panjang (misalnya 90 hari) dipertahankan (tidak dibuang atau di-*winsorizing*). Keputusan ini didasarkan pada justifikasi bisnis (*expert judgment*) bahwa nilai ekstrem tersebut bukanlah *outlier* statistik yang salah, melainkan representasi dari kasus nyata "Indent Part" (menunggu ketersediaan *sparepart* langka dari *supplier* luar daerah). Mempertahankan data ini krusial agar algoritma AHC dapat menangkap sinyal bisnis tersebut dan mengelompokkannya menjadi kluster prioritas rendah/non-stok.

b. Penanganan Anomali pada Variabel BIAYA

Berdasarkan temuan *Exploratory Data Analysis* (EDA), terdapat anomali pada nilai biaya servis berupa nilai Rp0 dan nilai yang sangat ekstrem (mencapai puluhan juta rupiah). Aturan penanganannya ditetapkan secara tegas sebagai berikut:

- a) Biaya Rp0: Dihapus (*drop baris*) dari *dataset* karena diidentifikasi sebagai kasus klaim garansi, pembatalan servis, atau kesalahan *input*. Data ini tidak mencerminkan beban operasional dan pendapatan riil konter.
- b) Biaya Ekstrem (Outlier Atas): Tetap dipertahankan (*retained*) dan tidak dilakukan pemotongan (*winsorizing*). Berdasarkan validasi operasional, biaya yang sangat tinggi tersebut merupakan representasi kasus nyata dari perbaikan *hardware* tingkat berat (seperti penggantian mesin utama/IC) pada perangkat *flagship*. Mempertahankan data ini sangat krusial agar algoritma dapat menangkap sinyal bisnis untuk membentuk kluster "Perbaikan *Hardware* Intensif".
- c. Normalisasi Deskripsi Kerusakan (KERUSAKAN_NORM)
- Karena variasi penulisan kerusakan pada data mentah tinggi, dilakukan normalisasi teks untuk menyatukan variasi istilah. Normalisasi menghasilkan kolom KERUSAKAN_NORM sebagai representasi konsisten dari kerusakan.

Tabel 3.8 Contoh Hasil Normalisasi Kerusakan

KERUSAKAN_RAW	KERUSAKAN_NORM
PADAM MASUK AIR	MATI TOTAL TOTAL MASUK AIR
CEK PROGRAM	CEK FLASH PROGRAM
LCD	GANTI LCD
LCD DAN BATERAI	GANTI LCD + BATERAI
T.POWER+BOBOL PIN	TOMBOL POWER+BOBOL PASSWORD

Normalisasi ini diperlukan agar informasi kerusakan lebih seragam dan dapat dipetakan ke kategori aksi servis serta komponen utama.

d. Standardisasi Aksi Servis (AKSI_SERVIS_STD)

Untuk merangkum kerusakan ke kategori operasional yang lebih sederhana, KERUSAKAN_NORM dipetakan menjadi aksi servis standar AKSI_SERVIS_STD (misalnya SOFTWARE, JASA, GANTI, HARDWARE_REPAIR). Hasil distribusi menunjukkan kategori GANTI sebagai yang paling dominan, diikuti HARDWARE_REPAIR, SOFTWARE, dan JASA.

Tabel 3.9 Total Aksi Servis

Aksi_Servis	Total
GANTI	475
HARDWARE_REPAIR	141
SOFTWARE	115
JASA	80

e. Ekstraksi Komponen Utama dan Jumlah Komponen

Untuk mendukung perencanaan stok, dilakukan ekstraksi komponen dari kerusakan/keterangan servis:

- a) KOMPONEN_UTAMA: komponen dominan yang digunakan/diganti dalam transaksi (contoh: LCD, MESIN_IC, KONEKTOR_CAS, BATERAI, SISTEM_SOFTWARE).
- b) JUMLAH_KOMPONEN: jumlah komponen fisik yang digunakan/diganti; transaksi SOFTWARE/JASA umumnya bernilai 0.

c) Hasil ekstraksi menunjukkan komponen yang paling dominan adalah LCD, diikuti SOFTWARE, TOMBOL, MESIN_IC, dan KONEKTOR_CAS. Proses ekstraksi memastikan tidak ada lagi label komponen “LAINNYA”, sehingga seluruh transaksi terklasifikasi ke komponen yang bermakna untuk inventaris.

f. Penyusunan Skor Kesulitan (SKOR_KESULITAN)

Setelah pembersihan dan rekayasa fitur, dataset akhir berisi atribut utama:

BIAYA, DURASI_HARI, JUMLAH_KOMPONEN, SKOR_KESULITAN, AKSI_SERVIS_STD, KOMPONEN_UTAMA, serta atribut pendukung untuk dashboard (misalnya TYPE_HP, TEKNISI, KETERANGAN).

Pada penelitian ini, dataset yang memenuhi kriteria kelengkapan dan validasi menghasilkan **811 baris** yang digunakan pada tahap transformasi (encoding dan scaling) dan pemodelan clustering.

Tabel 3.10 Ringkasan Hasil Data Preparation

Aspek	Kondisi Data Mentah	Setelah Data Preparation
Kolom tidak relevan	Ada (Unnamed)	Dihapus
BIAYA	Object (teks)	Numerik
Tanggal	Object	Datetime
Durasi	Ada nilai negatif & ekstrem	Divalidasi/koreksi
KERUSAKAN	Variasi teks tinggi	KERUSAKAN_NORM (seragam)
Fitur operasional	Belum terstruktur	AKSI_SERVIS_STD, KOMPONEN_UTAMA, JUMLAH_KOMPONEN
Kompleksitas	Tidak ada	SKOR_KESULITAN
Dataset modeling	1.113 baris	811 baris siap modeling

	(mentah)	
--	----------	--

Tabel 3.11 Perubahan Data Setelah Data Preparation

(Mentah) KERUSA KAN	(Hasil) KERUSAKAN _NORM	AKSI_SERVI S_STD	JUMLAH_KO MPONEN
PADAM MASUK AIR	MATI TOTAL TOTAL MASUK AIR	HARDWARE_ REPAIR	1
CEK PROGRA M	CEK FLASH PROGRAM	SOFTWARE	0
LCD DAN BATERA I	GANTI LCD + BATERAI	GANTI	2
SERVICE	JASA	SMALL_PART S	1

g. Transformasi untuk Pemodelan (One-Hot Encoding dan Z-Score Scaling)

Pada tahap transformasi, fitur kategorikal yang bersifat nominal, yaitu AKSI_SERVIS_STD dan KOMPONEN_UTAMA, dikonversi menjadi representasi biner (nilai 0 dan 1) menggunakan teknik *One-Hot Encoding* (OHE).

Penggunaan teknik OHE pada penelitian ini dipilih untuk mengatasi kelemahan representasi angka berurutan (*ordinal*). Karena kategori jenis kerusakan atau komponen (seperti LCD, Baterai, atau IC) murni bersifat nominal dan tidak memiliki tingkatan hierarki alami, pemberian kode bilangan bulat (0, 1, 2, dst.) berisiko menimbulkan bias berupa "jarak matematis palsu" pada algoritma berbasis jarak seperti *Euclidean*. Dengan menerapkan OHE, setiap kategori dipecah menjadi kolom independen

tersendiri, sehingga representasi jarak antar-kategori di dalam algoritma menjadi 100% valid dan objektif secara matematis.

Langkah selanjutnya adalah melakukan standardisasi skala (*scaling*) menggunakan metode *Z-Score* (*StandardScaler*) untuk mencegah variabel dengan rentang nilai yang sangat besar (seperti BIAYA) mendominasi perhitungan jarak *Euclidean*. Perlu ditegaskan bahwa dalam penelitian ini, proses standardisasi hanya diterapkan secara eksklusif pada fitur numerik (BIAYA, DURASI_HARI, JUMLAH_KOMPONEN, dan SKOR_KESULITAN). Sementara itu, kolom-kolom baru hasil OHE tidak diikutsertakan dalam proses *scaling* guna mempertahankan integritas makna biner aslinya (0 mewakili ketiadaan, dan 1 mewakili keberadaan atribut tersebut).

Setelah melalui proses transformasi ini, dimensi fitur pada *dataset* meluas menjadi kombinasi antara variabel numerik yang telah distandardisasi dan matriks biner kategorikal. Setelah data dipersiapkan melalui proses pembersihan, normalisasi, rekayasa fitur, serta transformasi (*encoding* dan *scaling*), *dataset* akhir yang dihasilkan telah tervalidasi dan siap digunakan sebagai masukan (*input*) pemodelan.

Tahap selanjutnya adalah melakukan pemodelan *clustering* menggunakan algoritma *Agglomerative Hierarchical Clustering* (AHC) untuk membentuk kelompok kasus servis sebagai dasar penentuan prioritas stok *sparepart*, yang akan dibahas secara rinci pada bab berikutnya.

3.5.4. Modeling (AHC)

Tahap *modeling* bertujuan untuk membentuk pengelompokan kasus servis berdasarkan kemiripan karakteristik transaksi. Hasil pengelompokan ini akan menjadi landasan objektif dalam penyusunan prioritas pengadaan stok *sparepart*. Karena data historis transaksi tidak memiliki label kategori baku sebelumnya, penelitian ini menggunakan pendekatan *unsupervised learning* melalui algoritma *clustering* yang terbukti efisien untuk menemukan struktur dan pola tersembunyi pada data tanpa label (Kashyap, 2024; Li et al., 2022; Rolf et al., 2025).

1. Pemilihan Metode Clustering

Model utama yang digunakan dalam penelitian ini adalah *Agglomerative Hierarchical Clustering* (AHC). Pemilihan metode ini didasarkan pada pertimbangan berikut:

- a. AHC sangat efektif diterapkan pada *dataset* berskala kecil hingga menengah dan mampu memvisualisasikan struktur hierarki pemisahan data secara transparan melalui *dendrogram*, sehingga memudahkan interpretasi operasional (Rajasekaran et al., 2025).
- b. Berbeda dengan metode partisional seperti K-Means yang mengharuskan penentuan jumlah kluster di awal, AHC memungkinkan peneliti untuk menentukan jumlah kluster yang optimal setelah melihat keseluruhan struktur hierarki data, serta terbukti lebih *robust* terhadap variasi data untuk segmentasi layanan (Bansal, 2023).

2. Dataset Input Pemodelan

Dataset yang dimasukkan ke dalam model algoritma adalah hasil akhir dari tahap *Data Preparation* yang telah dibersihkan sebanyak 811 baris data

transaksi. Fitur yang digunakan sebagai *input* pemodelan terdiri dari gabungan variabel numerik (BIAYA, DURASI_HARI, JUMLAH_KOMPONEN, SKOR_KESULITAN) dan kolom biner hasil *One-Hot Encoding* (OHE) dari atribut kategorikal.

Sebelum dieksekusi, proses standardisasi menggunakan *Z-Score* hanya diterapkan secara eksklusif pada fitur numerik untuk menyetarakan skala. Sementara itu, fitur hasil OHE dipertahankan murni dalam bentuk biner (0 dan 1) agar integritas maknanya tidak terdistorsi.

3. Konfigurasi Algoritma

Parameter teknis yang dikonfigurasi pada model AHC adalah pendekatan *Agglomerative* (Bottom-up) menggunakan metrik jarak *Euclidean* dan *Ward Linkage*. Walaupun terdapat fitur biner hasil OHE dalam *dataset*, penggunaan kombinasi *Ward-Euclidean* tetap dipertahankan dan valid secara metodologis. Hal ini dikarenakan seluruh fitur telah berada pada skala yang sebanding (numerik distandardisasi, OHE bernilai 0/1), dan tujuan utama penelitian ini adalah membentuk kluster operasional yang kompak untuk *profiling* kebutuhan stok barang. Validitas dari keputusan metodologis ini selanjutnya diuji secara kuantitatif melalui metrik *Silhouette Score* dan *Davies-Bouldin Index* (DBI).

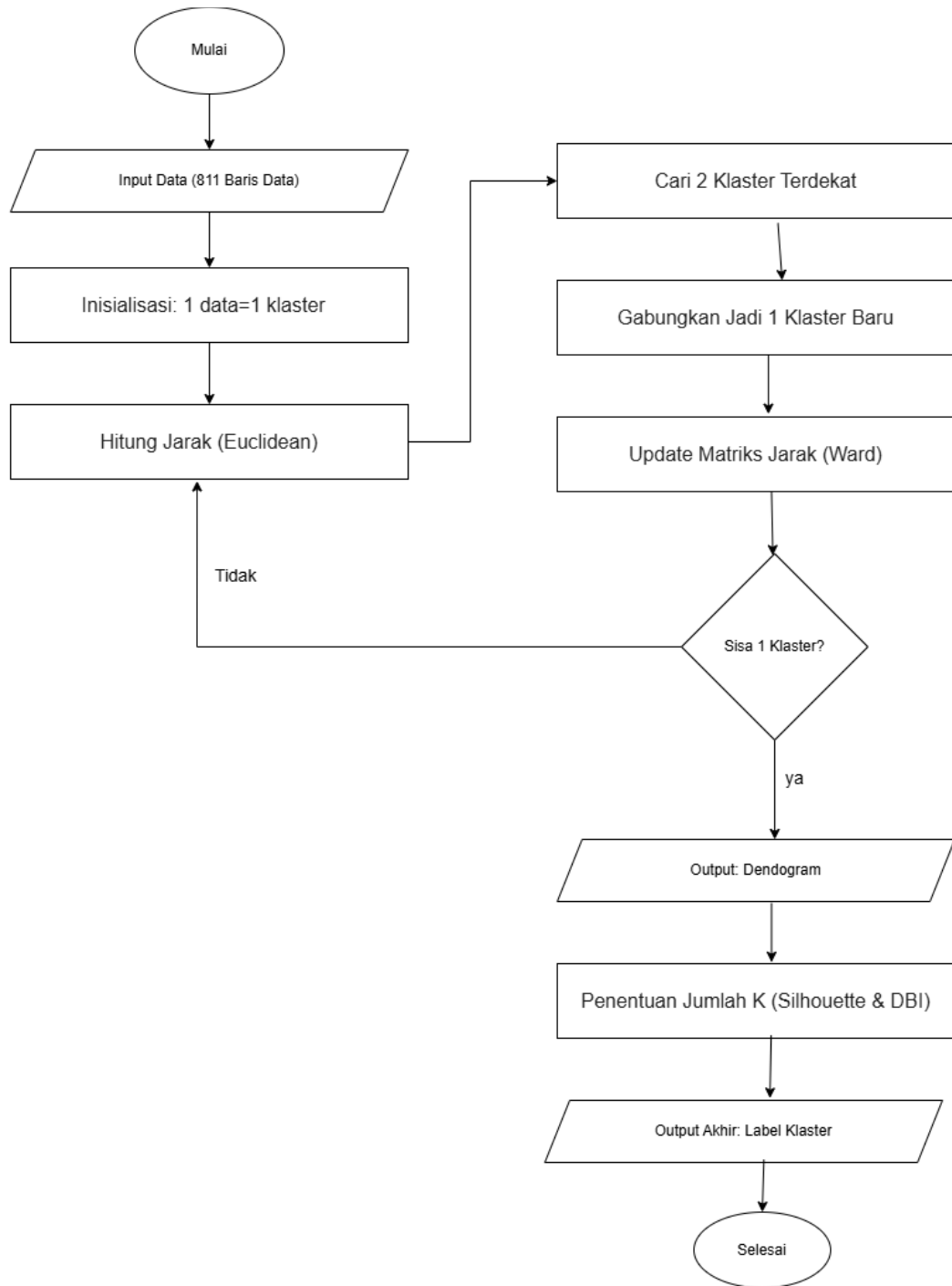
4. Prosedur Pembentukan Kluster

Secara sistematis, proses pembentukan kluster melalui AHC dilakukan dengan tahapan:

- a. Setiap baris data (811 baris) awalnya dianggap sebagai satu klaster tunggal.
- b. Algoritma menghitung kedekatan antar-klaster menggunakan metrik jarak *Euclidean*.
- c. Dua klaster yang paling mirip digabungkan menjadi satu klaster baru berdasarkan kriteria *Ward Linkage*.
- d. Proses berulang hingga seluruh titik data bergabung membentuk satu struktur hierarki utuh (*dendrogram*).

5. Alur Proses Algoritma

Secara sistematis, tahapan pemrosesan data menggunakan algoritma AHC dalam penelitian ini digambarkan pada Gambar 3.6 berikut:



Gambar 3.6 Alur Algoritma Agglomerative Hierarchical

Clustering

Diagram di atas menunjukkan bahwa proses dimulai dengan menganggap setiap data sebagai satu cluster, kemudian secara iteratif menggabungkan dua cluster

terdekat berdasarkan jarak Euclidean dan metode Ward hingga seluruh data membentuk satu hierarki utuh.

1. Penentuan Jumlah Cluster

- a. Menggunakan dendrogram untuk mengamati struktur penggabungan cluster dan menentukan cut-off yang relevan.
- b. Menguji beberapa kandidat jumlah cluster (misalnya 2, 3, 4 cluster) dan membandingkannya menggunakan metrik evaluasi (Silhouette dan DBI).
- c. Mempertimbangkan aspek operasional dan justifikasi pakar (*expert judgment*) terkait manajemen persediaan suku cadang.

Selain menggunakan pendekatan matematis, pemilihan jumlah klaster (k) yang paling optimal juga sangat mempertimbangkan justifikasi logika operasional bisnis (*expert judgment*). Metrik evaluasi tertinggi tidak akan dipilih jika hasil klasternya tidak dapat diinterpretasikan. Sebagai contoh proyeksi, jumlah klaster yang dibentuk harus mampu merepresentasikan kategori kebutuhan manajemen persediaan di lapangan, seperti pemisahan antara layanan yang membutuhkan barang wajib ada (*ready stock*), barang cadangan terbatas (*buffer stock*), kasus kelangkaan barang (*indent/pre-order*), maupun layanan yang murni tanpa stok fisik.

2. Visualisasi Cluster

- a. Melakukan reduksi dimensi menggunakan PCA untuk memproyeksikan data ke dalam ruang dua dimensi.

- b. Menampilkan scatter plot yang memperlihatkan pemisahan antar cluster berdasarkan kombinasi fitur utama.

3.5.4.1 Contoh Perhitungan Manual AHC

Guna memberikan pemahaman mendalam mengenai mekanisme kerja algoritma AHC, bagian ini menyajikan simulasi perhitungan manual menggunakan sampel data sederhana yang merepresentasikan karakteristik dataset layanan servis.

1. Penentuan Sampel Data

Sebanyak empat data sampel dipilih dengan dua variabel utama, yaitu Durasi (X) dalam satuan hari dan Biaya (Y) dalam satuan ratus ribu rupiah.

Tabel 3.12 Sampel Data Contoh

Data (Titik)	Durasi (X)	Biaya (Y)
P1	1	1
P2	2	1
P3	4	5
P4	5	5

(Keterangan: Nilai 1 merepresentasikan Rp 100.000, Nilai 5 merepresentasikan Rp 500.000, dst).

2. Perhitungan Matriks Jarak (Euclidean Distance)

Langkah pertama adalah menghitung jarak antar titik data menggunakan rumus Euclidean Distance:

$$d(x, y) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Berikut adalah contoh perhitungan jarak antara titik P1 dan P2:

$$d(P1, P2) = \sqrt{(1 - 2)^2 + (1 - 1)^2} = \sqrt{(-1)^2 + 0} = 1$$

Berikut adalah contoh perhitungan jarak antara titik P1 dan P3:

$$d(P1, P3) = \sqrt{(1-4)^2 + (1-5)^2} = \sqrt{(-3)^2 + (-4)^2} = \sqrt{9+16} = 5$$

Berdasarkan perhitungan tersebut untuk seluruh pasangan data, diperoleh matriks jarak awal (Iterasi 1) sebagai berikut:

Tabel 3.13 Matriks Jarak Awal

Jarak	P1	P2	P3	P4
P1	0	1	5	5.66
P2	1	0	4.47	5
P3	5	4.47	0	1
P4	5.66	5	1	0

3. Tahap Penggabungan Awal (Iterasi 1)

Algoritma mengidentifikasi nilai jarak terkecil di dalam matriks jarak. Berdasarkan Tabel Matriks Jarak Awal, nilai minimum adalah 1, yang terdapat pada pasangan (P1, P2) dan (P3, P4). Oleh karena itu, dilakukan penggabungan iterasi pertama: a. Titik P1 dan P2 digabungkan membentuk Kluster 1 (C1). b. Titik P3 dan P4 digabungkan membentuk Kluster 2 (C2).

4. Pembaruan Matriks Jarak (Ward Linkage)

Pada tahap ini, terbentuk dua kluster, yaitu C1 dan C2. Karena algoritma menggunakan pendekatan *Ward Linkage*, jarak antar-kluster dihitung berdasarkan peningkatan *Sum of Squared Error* (SSE). Rumus jarak *Ward* antara dua kluster adalah kombinasi dari jarak *centroid* (titik pusat) yang dikalikan dengan rasio jumlah anggotanya:

$$a. \text{ Titik Pusat C1 (P1, P2)} = \left(\frac{1+2}{2}, \frac{1+1}{2} \right) = (1.5, 1)$$

$$b. \text{ Titik Pusat C2 (P3, P4)} = \left(\frac{4+5}{2}, \frac{5+5}{2} \right) = (4.5, 5)$$

Selanjutnya, hitung jarak Euclidean antar *centroid* tersebut:

$$d(\text{Centroid}) = \sqrt{(1.5 - 4.5)^2 + (1 - 5)^2} = \sqrt{(-3)^2 + (-4)^2} = 5$$

Langkah terakhir, masukkan ke dalam rumus Ward Linkage, di mana n_1 adalah jumlah anggota C1 (2 data) dan n_2 adalah jumlah anggota C2 (2 data):

$$d_{\text{Ward}}(C1, C2) = \sqrt{\frac{n_1 \times n_2}{n_1 + n_2}} \times d(\text{Centroid})$$

$$d_{\text{Ward}}(C1, C2) = \sqrt{\frac{2 \times 2}{2 + 2}} \times 5 = \sqrt{\frac{4}{4}} \times 5 = 1 \times 5 = 5$$

Hasil perhitungan menunjukkan jarak antara Klaster 1 dan Klaster 2 menggunakan metode *Ward* adalah 5.

5. Hasil Akhir (Dendogram)

Contoh perhitungan ini mendemonstrasikan kemampuan algoritma AHC dengan *Ward Linkage* dalam mengelompokkan data berdasarkan kedekatan karakteristiknya. Pada simulasi sederhana di atas, data dengan durasi singkat dan biaya rendah (P1 dan P2) secara alami memisahkan diri membentuk kelompok kasus sederhana, sedangkan data dengan nilai variabel tinggi (P3 dan P4) membentuk kelompok kasus kompleks. Logika pemisahan matematis inilah yang kemudian diaplikasikan oleh sistem pada 811 *dataset* riil untuk membentuk empat klaster prioritas pengadaan *sparepart* yang lebih spesifik dan terperinci.

3.5.5. Evaluation

Tahap evaluasi (*evaluation*) merupakan fase krusial dalam CRISP-DM yang berfungsi sebagai gerbang penentu (*decision gate*) sebelum model diimplementasikan. Evaluasi pada penelitian ini dilakukan melalui dua

pendekatan: evaluasi kuantitatif secara algoritmik menggunakan metrik pengukuran kluster, dan evaluasi kualitatif secara operasional untuk memastikan hasil *clustering* relevan dengan kebutuhan bisnis Konter Anjas Reparasi.

1. Evaluasi Kuantitatif (Metrik Internal)

Karena penelitian ini menggunakan pendekatan *unsupervised learning*, tidak ada label target kebenaran (*ground truth*) yang bisa dijadikan acuan. Oleh karena itu, evaluasi dilakukan menggunakan metrik internal yang mengukur tingkat kepadatan (*cohesion*) dan keterpisahan (*separation*) kluster. Dua metrik yang digunakan adalah:

a. Silhouette Coefficient (Skor Silhouette)

Metrik ini digunakan untuk mengukur seberapa mirip suatu objek dengan klusternya sendiri dibandingkan dengan kluster tetangga terdekatnya. Nilai *Silhouette* berkisar antara -1 hingga +1. Nilai yang semakin mendekati +1 mengindikasikan bahwa objek berada di dalam kluster yang sangat tepat dan jarak antar-kluster terpisah dengan sangat baik (Shutaywi & Kachouie, 2021).

b. Davies–Bouldin Index (DBI)

Metrik ini mengukur rasio penyebaran (dispersi) intra-kluster berbanding dengan jarak antar-kluster. Berbeda dengan *Silhouette*, nilai DBI yang semakin rendah (mendekati 0) menunjukkan bahwa kluster yang terbentuk semakin kompak dan saling terpisah dengan tegas. Penggunaan *Silhouette Coefficient* dan DBI secara bersamaan sangat direkomendasikan untuk memvalidasi jumlah kluster optimal dan mengurangi bias evaluasi algoritma (Poetri et al., 2024).

2. Evaluasi Kualitatif dan Interpretasi Bisnis (*Profiling*)

Selain evaluasi matematis, literatur terbaru oleh (Yin et al., 2024) menekankan bahwa algoritma *clustering* memiliki peran krusial dalam aplikasi dunia nyata, seperti pengorganisasian data dan analisis keputusan operasional. Oleh karena itu, evaluasi bisnis wajib dilakukan pada tahap akhir dengan membedah profil karakteristik dari masing-masing kluster yang terbentuk (k optimal). Evaluasi ini bertujuan memastikan hasil pengelompokan tidak hanya valid secara statistik, tetapi juga logis untuk menjawab rumusan masalah penentuan prioritas stok sparepart.

Interpretasi profil akan difokuskan pada pengamatan nilai agregasi statistik (rata-rata dan nilai maksimum) dari atribut biaya dan durasi pengerjaan, serta modus dari jenis komponen fisik yang paling dominan di tiap kelompok. Hasil interpretasi analitik ini kemudian divalidasi dan dicocokkan dengan logika operasional nyata (*expert judgment*) oleh pihak manajemen (Ahli/Pemilik Konter). Jika profil karakteristik kluster dinilai relevan dan representatif, maka tahapan CRISP-DM dinyatakan berhasil secara operasional dan siap dilanjutkan ke fase terakhir, yaitu *Deployment*.

3.5.6. Deployment

Tahap *deployment* merupakan fase pemungkas dalam metodologi CRISP-DM. Sebagaimana disoroti oleh (Schröer et al., 2021), fase ini sering kali menjadi titik lemah dalam banyak proyek *data mining* karena model analitik gagal diintegrasikan ke lingkungan produksi operasional. Penelitian ini secara khusus menjembatani celah tersebut dengan mengimplementasikan hasil klusterisasi AHC ke dalam sebuah sistem *dashboard* interaktif berbasis *web* (menggunakan

framework Flask dan antarmuka Bootstrap) yang dapat diakses secara langsung oleh pengambil keputusan di Konter Anjas Reparasi.

Fokus utama pada tahap *deployment* ini meliputi tiga aspek operasional:

1. Penerapan Aturan Bisnis (*Business Rules Integration*)

Mengonversi profil klaster operasional yang telah tervalidasi pada fase evaluasi menjadi aturan sistemik. Hasil pengelompokan dijadikan basis logika bagi sistem untuk mengkategorikan data transaksi baru dan menentukan prioritas layanan di masa depan.

2. Visualisasi Data dan Analitik Interaktif

Dashboard dirancang untuk menyajikan representasi visual yang intuitif dari hasil data mining. Sistem akan menampilkan metrik penting seperti persentase distribusi klaster kasus servis, tren produktivitas, serta peta persebaran komponen (sparepart) yang paling sering mengalami pergantian. Modul ini mengaitkan setiap klaster dengan faktor-faktor utama seperti durasi, biaya, dan jenis kerusakan.

3. Rekomendasi Prioritas Stok Sparepart

Ini merupakan fitur inti dan tujuan akhir dari sistem. *Dashboard* akan menerjemahkan wawasan analitik menjadi rekomendasi tindakan operasional (*actionable insights*) bagi pemilik konter, di antaranya:

1. Memetakan daftar komponen kategori *fast-moving* yang berstatus prioritas tinggi (*ready stock*) untuk mempercepat waktu layanan.
2. Mengidentifikasi suku cadang situasional bernilai tinggi yang memerlukan penyediaan terbatas (*buffer stock*) agar pemilik dapat mengatur alokasi anggaran belanja secara presisi.

3. Memberikan peringatan (*alert*) terkait suku cadang langka yang sering menyebabkan kasus *indent part* dengan durasi pengerjaan ekstrem, sehingga konter tidak perlu melakukan penyetokan fisik dan menghindari *dead-stock*.

Melalui eksekusi fase *deployment* ini, *output* penelitian bertransformasi dari sekadar laporan analitik matematis menjadi sebuah *Decision Support System* (Sistem Pendukung Keputusan) yang aplikatif. Sistem ini diharapkan mampu memberikan panduan berbasis data (*data-driven*) untuk meningkatkan efisiensi rantai pasok dan kualitas layanan di Konter Anjas Reparasi.

3.6. Perancangan Sistem Dashboard

Perancangan sistem *dashboard* dilakukan untuk mengimplementasikan hasil analisis algoritma AHC ke dalam sebuah antarmuka yang interaktif. Sistem yang diberi nama KONTER ANJAS ANALYTICS ini dirancang sebagai *Decision Support System* (DSS) berbasis *web* yang responsif (*responsive web design*) untuk memfasilitasi pemilik konter dalam pengambilan keputusan operasional.

Secara arsitektur perangkat lunak, sistem ini dibangun menggunakan pendekatan tiga lapis (*Three-Tier Architecture*):

1. Lapis Antarmuka (*Presentation Layer*): Dibangun menggunakan HTML dan *framework* Bootstrap untuk menyajikan visualisasi data yang interaktif kepada pengguna.
2. Lapis Aplikasi (*Application Layer*): Dibangun menggunakan *framework* Flask (Python) yang menangani logika rute (*routing*), agregasi data, serta pemrosesan *script* pencocokan kluster berbasis kemiripan profil.

3. Lapis Data (*Data Layer*): Mengelola basis data transaksi dalam bentuk *file* terstruktur (.csv) yang memuat riwayat servis beserta label klaster hasil pemodelan AHC. Sistem membaca data secara langsung dari direktori lokal aplikasi, sehingga fitur unggah (*upload*) *dataset* maupun proses otentikasi (*login*) tidak menjadi fokus penelitian, mengingat sistem dirancang murni untuk penggunaan internal pada lingkungan perangkat yang terbatas.

3.6.1. Perancangan Modul Dashboard

Sistem ini membagi fungsionalitasnya ke dalam empat modul utama. Untuk mendeskripsikan alur kerja secara sistematis, setiap modul dijabarkan menggunakan format *Input - Proses - Output* (IPO).

1. Modul *Executive Dashboard* (Monitoring Operasional) Modul ini berfungsi sebagai halaman ringkasan (*snapshot*) kondisi operasional konter.
 - a. **Input:** *Dataset* riwayat transaksi servis (format .csv) yang dimuat dari direktori lokal sistem.
 - b. **Proses:** Sistem melakukan agregasi data (penjumlahan, rata-rata, dan pencarian frekuensi tertinggi) berdasarkan periode waktu tertentu serta menghitung proporsi distribusi masing-masing label klaster AHC.
 - c. **Output:** Visualisasi *KPI Cards* yang menampilkan metrik ringkasan operasional (seperti total transaksi, total omzet, rata-rata durasi pengerjaan, dan kasus perbaikan terbanyak), grafik tren fluktuasi omzet bulanan, *donut chart* distribusi persentase

transaksi per klaster kasus servis, serta diagram batang beban performa teknisi.

2. Modul Data Servis (*Database Terstruktur*) Modul ini bertujuan menyajikan integrasi langsung antara data operasional harian dengan hasil analitik algoritma AHC pada tingkat baris transaksi.

- a. **Input:** *Dataset* lengkap hasil *Data Preparation* beserta kolom atribut CLUSTER.
- b. **Proses:** Sistem memuat tabel basis data, mengaplikasikan fitur penomoran halaman (*pagination*), serta mengaktifkan fungsi pencarian dan pengurutan karakter teks/angka.
- c. **Output:** Tabel riwayat servis interaktif yang menampilkan detail setiap transaksi (tanggal, durasi, tipe HP, biaya, teknisi) yang dilengkapi dengan *badge* visual penanda kategori/label klasternya.

3. Modul Smart Stok (Rekomendasi Prioritas) Modul ini merupakan representasi fungsional dari luaran utama penelitian, di mana rekomendasi prioritas pengadaan *sparepart* dipetakan berdasarkan karakteristik pemisahan klaster.

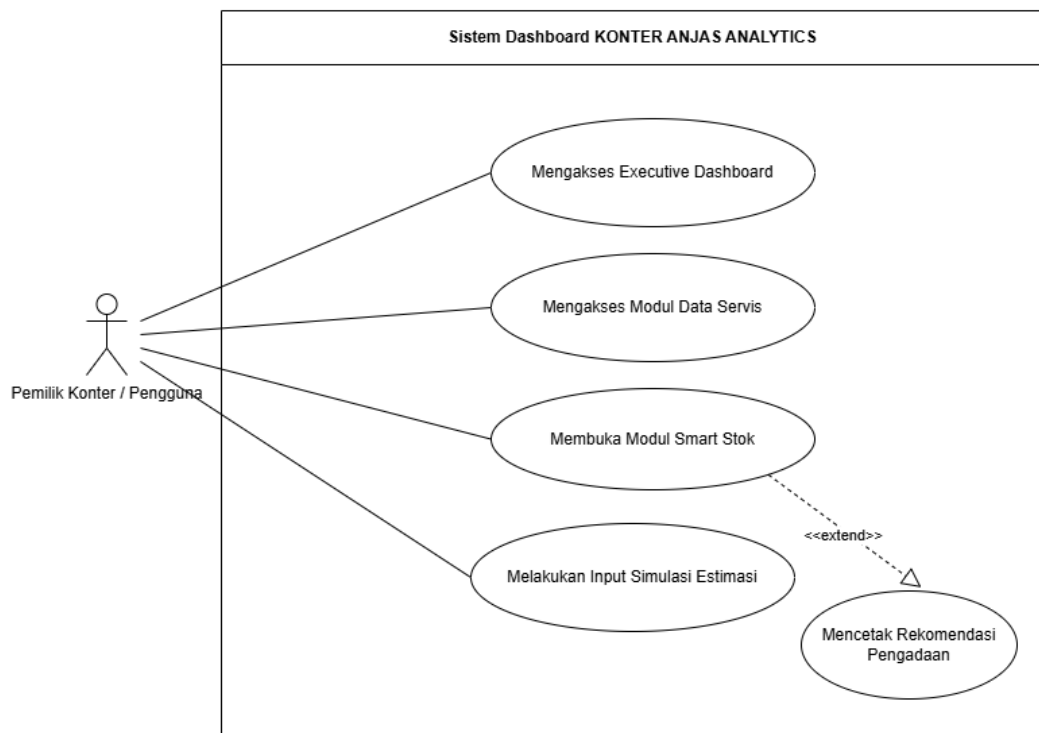
- a. **Input:** Label klaster hasil algoritma AHC dan atribut KOMPONEN_UTAMA dari *dataset* historis.
- b. **Proses:** Sistem melakukan agregasi frekuensi kemunculan komponen utama pada setiap label klaster. Selanjutnya, sistem menyoroti komponen dominan pada klaster yang memiliki tingkat kekritisasi operasional lebih tinggi (misalnya klaster

dengan durasi dan skor kesulitan yang lebih tinggi) sebagai dasar prioritas pengadaan *sparepart*.

- c. **Output:** Tabel rekomendasi prioritas pengadaan *sparepart* per klaster beserta frekuensi kemunculan historisnya, serta tombol fungsional untuk mengeksport/mencetak daftar belanja.
4. Modul Simulasi (Estimasi Berbasis Profil Historis): Modul ini dirancang untuk memberikan perkiraan beban layanan bagi perangkat pelanggan yang baru masuk.
- a. **Input:** Parameter spesifikasi dari pengguna, yaitu Tipe Handphone dan deskripsi Jenis Kerusakan.
 - b. **Proses:** Sistem melakukan penyaringan (*filtering*) terhadap *input* pengguna pada basis data riwayat transaksi untuk menemukan rekam jejak kasus dengan karakteristik perbaikan serupa. Sistem tidak menggunakan perhitungan jarak algoritmik maupun model prediksi terawasi (*supervised learning*), melainkan murni melakukan pencocokan data historis (*historical data matching*) untuk mengekstrak kelompok kasus yang relevan.
 - c. **Output:** Tampilan estimasi kategori klaster layanan, lengkap dengan perkiraan durasi pengerjaan dan rentang biaya servis. Estimasi ini diperoleh dari perhitungan agregasi statistik dasar (seperti nilai rata-rata atau median) dari kumpulan data historis yang berhasil disaring.

3.6.2. Use Case Diagram

Untuk menggambarkan fungsionalitas sistem dari sudut pandang pengguna, dirancang sebuah Use Case Diagram. Diagram ini mendefinisikan interaksi antara aktor (pengguna sistem) dengan fitur-fitur utama yang tersedia di dalam dashboard analitik. Rancangan interaksi tersebut disajikan pada Gambar 3.7.



Gambar 3.7 Use Case Diagram Sistem Dashboard

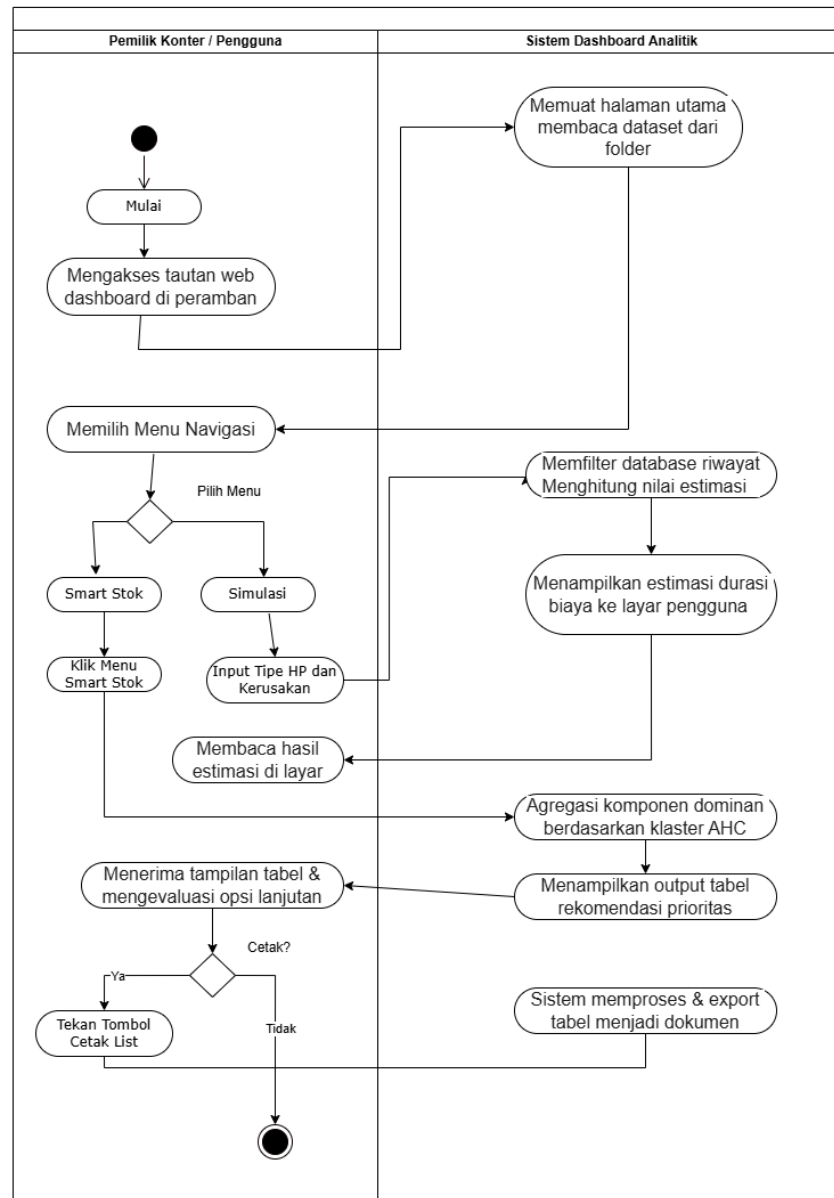
Berdasarkan *Use Case Diagram* di atas, terdapat satu aktor utama, yaitu Pemilik Konter / Pengguna, yang berinteraksi langsung dengan sistem. Mengingat *dataset* diakses secara otomatis dari direktori lokal, aktor berfokus pada lima fungsionalitas utama, yaitu:

1. Mengakses *Executive Dashboard* untuk memantau ringkasan visual.
2. Mengakses modul Data Servis untuk menelusuri riwayat terstruktur.

3. Membuka modul Smart Stok untuk melihat prioritas *sparepart*.
4. Mencetak/Mengekspor rekomendasi Smart Stok menjadi daftar belanja.
5. Melakukan *input* pada fitur Simulasi untuk mendapatkan estimasi servis berbasis histori.

3.6.3. Activity Diagram

Alur kerja sistem secara prosedural digambarkan menggunakan *Activity Diagram*. Diagram ini merinci urutan interaksi antara aksi pengguna dan pemrosesan sistem.



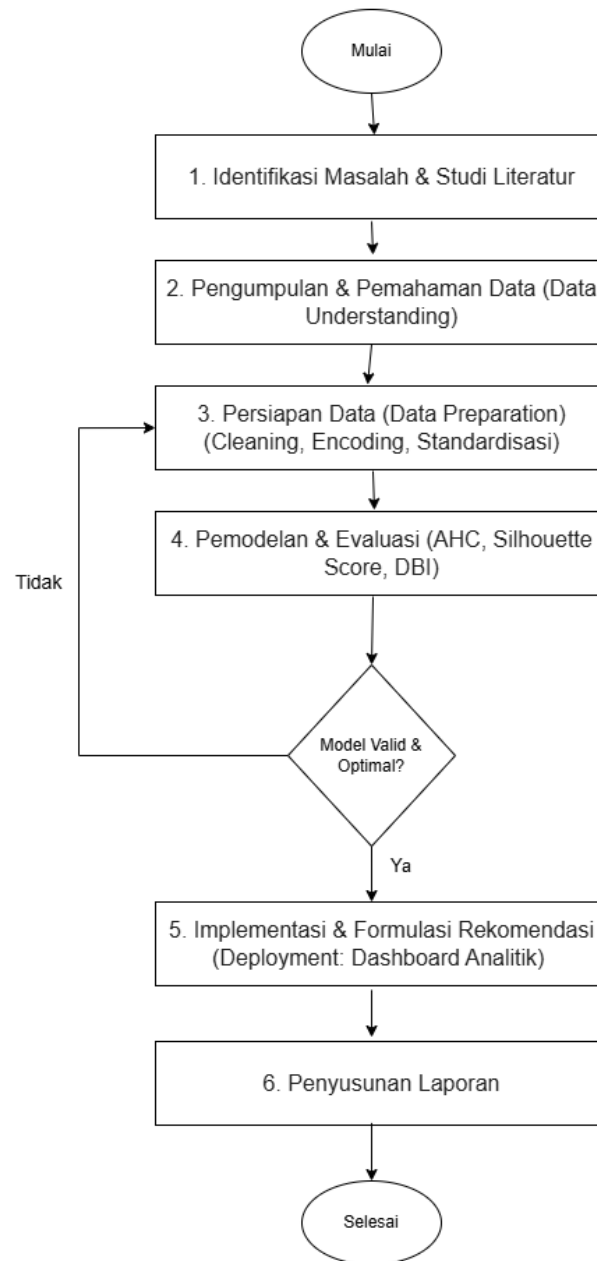
Gambar 3.8 Activity Diagram Sistem Dashboard

Berdasarkan *Activity Diagram* di atas, alur operasional dimulai saat Pemilik Konter mengakses tautan *web dashboard* di peramban. Sistem memuat halaman utama dan membaca *dataset* dari direktori lokal. Jika pengguna menavigasi ke menu Smart Stok, sistem akan memproses agregasi data komponen berdasarkan kluster AHC dan mengembalikan *output* tabel rekomendasi, di mana pengguna diberikan pilihan aktivitas lanjutan untuk menekan tombol cetak dokumen.

Sebaliknya, jika pengguna memilih menu Simulasi, pengguna diwajibkan melakukan *input* deskripsi kerusakan, kemudian sistem merespons dengan memfilter basis data riwayat untuk mencari rekam jejak kasus serupa dan menghitung nilai estimasi statistiknya, sebelum menampilkan hasil tersebut ke layar pengguna.

3.7. Alur Penelitian

Untuk memberikan gambaran menyeluruh mengenai pelaksanaan penelitian dari awal hingga akhir, disusun bagan alir (*flowchart*) penelitian. Bagan ini merangkum tahapan logis dan prosedural yang dilakukan peneliti untuk menjawab rumusan masalah secara sistematis. Alur lengkap penelitian disajikan pada Gambar 3.9.



Gambar 3.9 Flowchart Alur Penelitian

Berdasarkan Gambar 3.9, alur penelitian dilaksanakan melalui tahapan berikut:

1. Identifikasi masalah dan studi literatur: Observasi di Konter Anjas Reparasi dilakukan untuk merumuskan permasalahan prioritas stok

sparepart. Studi literatur digunakan untuk memperkuat landasan teori terkait manajemen persediaan, algoritma *clustering* (AHC), evaluasi klaster, dan *dashboard* analitik.

2. Pengumpulan dan pemahaman data (*Data Understanding*): Data historis transaksi servis periode Mei–Oktober 2025 dikumpulkan dan dianalisis menggunakan EDA untuk melihat sebaran, kelengkapan atribut, serta mendeteksi anomali.
3. Persiapan data (*Data Preparation*): Data mentah dibersihkan dan ditransformasi melalui penghapusan kolom tidak relevan, penanganan *missing value*, normalisasi teks, pembentukan fitur (misalnya skor kesulitan), *encoding* atribut kategorikal, serta standardisasi skala variabel.
4. Pemodelan dan evaluasi (*Modeling & Evaluation*): *Dataset* diproses menggunakan *Agglomerative Hierarchical Clustering* dengan *Ward Linkage*. Kualitas klaster dievaluasi menggunakan *Silhouette Score* dan *Davies–Bouldin Index* (DBI). Apabila nilai *Silhouette* belum memadai atau DBI masih tinggi, serta interpretasi klaster belum selaras dengan kondisi operasional, proses diiterasi kembali ke tahap *Data Preparation*.
5. Implementasi dan formulasi rekomendasi (*Deployment*): Hasil *profiling* klaster dan komponen dominan diintegrasikan ke dalam sistem *dashboard* (KONTER ANJAS ANALYTICS) untuk menghasilkan rekomendasi prioritas pengadaan *sparepart*.
6. Penyusunan laporan: Seluruh hasil analisis, implementasi perangkat lunak, dan kesimpulan yang diperoleh disusun dalam laporan skripsi.

BAB IV

IMPLEMENTASI DAN PENGUJIAN SISTEM

4.1. Gambaran Umum Implementasi Penelitian

Bab ini menguraikan hasil dan pembahasan dari tahapan inti penelitian yang dijalankan berdasarkan kerangka kerja CRISP-DM, yang meliputi tahap persiapan data (*data preparation*), pemodelan algoritma (*modeling*), evaluasi (*evaluation*), hingga implementasi perangkat lunak (*deployment*). Implementasi pengelompokan (*clustering*) ini diterapkan pada data historis transaksi servis Konter Anjas Reparasi selama periode enam bulan, yaitu mulai bulan Mei hingga Oktober 2025.

Fokus utama dari implementasi ini adalah memproses data mentah transaksi harian menjadi wawasan strategis menggunakan algoritma *Agglomerative Hierarchical Clustering* (AHC). Proses analisis dilakukan untuk menemukan struktur tersembunyi dan mengidentifikasi pola karakteristik dari setiap kasus perbaikan yang masuk ke konter. Pendekatan ini dipilih untuk mengubah pengambilan keputusan yang sebelumnya mengandalkan intuisi menjadi berbasis data (*data-driven*).

Hasil keluaran (*output*) yang disajikan pada bab ini tidak hanya berfokus pada metrik evaluasi matematis dan pembentukan profil kluster semata. Lebih dari itu, hasil akhir analisis bermuara pada perumusan rekomendasi prioritas pengadaan *sparepart* yang secara langsung diintegrasikan ke dalam *dashboard* analitik berbasis *web* bernama KONTER ANJAS ANALYTICS. Sistem tersebut diuji dan diimplementasikan sebagai *Decision Support System* (DSS) untuk

membantu pemilik konter dalam merencanakan manajemen persediaan stok operasional secara lebih efisien, objektif, dan tepat sasaran.

4.2. Hasil Persiapan Data untuk Clustering

Tahap persiapan data bertujuan memastikan algoritma Agglomerative Hierarchical Clustering (AHC) menerima masukan data yang valid, relevan, dan memiliki skala yang sebanding. Pada tahap ini dilakukan pembersihan data (*data cleaning*), pembentukan fitur (*feature engineering*), transformasi atribut kategorikal menjadi numerik (*encoding*), serta standardisasi skala (*scaling*).

4.2.1. Pembersihan Data (Data Cleaning)

Dataset historis transaksi servis Konter Anjas Reparasi periode Mei hingga Oktober 2025 pada awalnya berjumlah 1.113 baris data mentah. Berdasarkan hasil eksplorasi data awal (EDA), ditemukan beberapa isu kualitas data seperti nilai kosong (*missing value*) pada atribut kunci, kolom tidak relevan, serta anomali input (misalnya durasi bernilai negatif). Oleh karena itu dilakukan penyaringan data agar dataset yang masuk ke tahap pemodelan benar-benar memenuhi kriteria kelayakan.

Secara umum, data yang dikeluarkan dari dataset mencakup:

1. Transaksi dengan *missing value* pada atribut kunci yang diperlukan untuk pembentukan fitur (misalnya tanggal masuk/keluar dan biaya).
2. Transaksi dengan nilai durasi tidak valid (misalnya durasi negatif) yang mengindikasikan kesalahan input tanggal.

3. Transaksi yang tidak memenuhi kriteria analisis operasional (misalnya status dibatalkan/tidak selesai diperbaiki) sesuai aturan pembersihan yang ditetapkan pada tahap Data Preparation.

Setelah proses pembersihan, dataset menyisakan 811 baris data transaksi yang valid untuk dianalisis lebih lanjut.

Tabel 4.1 Ringkasan Penyusutan Dataset

Keterangan	Jumlah Baris Data
Data mentah awal (<i>raw data</i>)	1.113
Data terhapus pada proses cleaning	302
Dataset akhir siap pemodelan	811

Tabel 4.2 Rincian Alasan Penghapusan Data

Tahap/Alasan Penghapusan	Jumlah Baris
Hapus duplikasi data	149
Missing pada kolom kritis (KERUSAKAN/TEKNISI/BIAYA)	74
Tanggal masuk/keluar tidak valid (gagal parsing)	6
Nilai BIAYA tidak valid (gagal konversi numerik)	2
Durasi negatif (anomali input tanggal)	2
Durasi null (tanggal keluar belum tercatat)	23
Filter transaksi non-handphone (LAPTOP)	37
Pembersihan Type HP tidak valid + audit final (hapus baris bermasalah)	9
Total data terhapus	302

4.2.2. Seleksi dan Transformasi Fitur (Feature Engineering & Encoding)

Algoritma AHC dengan jarak Euclidean membutuhkan seluruh variabel input dalam bentuk numerik. Oleh karena itu, dilakukan pembentukan dan seleksi fitur agar kasus servis dapat direpresentasikan secara terstruktur.

Fitur numerik utama yang digunakan meliputi:

1. **BIAYA**: biaya transaksi servis yang telah dikonversi ke numerik,
2. **DURASI_HARI**: selisih tanggal keluar dan tanggal masuk,
3. **JUMLAH_KOMPONEN**: jumlah item sparepart yang digunakan pada transaksi,
4. **SKOR_KESULITAN**: skor kompleksitas teknis (skala 1–10) berdasarkan rubrik dan validasi pakar.

Selain itu, dilakukan transformasi terhadap atribut kategorikal nominal (**AKSI_SERVIS_STD** dan **KOMPONEN_UTAMA**) menggunakan teknik *One-Hot Encoding* (OHE). Teknik ini memecah setiap kategori menjadi kolom biner baru (nilai 0 dan 1) untuk mencegah munculnya jarak matematis palsu yang biasa terjadi pada *ordinal encoding*. Proses ini menghasilkan perluasan dimensi fitur dari 6 kolom awal menjadi 19 kolom fitur (4 fitur numerik dan 15 kolom biner hasil OHE)..

Dengan demikian, dataset input pemodelan terdiri dari enam fitur operasional berikut:

BIAYA, **DURASI_HARI**, **JUMLAH_KOMPONEN**, **SKOR_KESULITAN**, **AKSI_ENC**, dan **KOMPONEN_ENC**.

4.2.3. Standardisasi Data (Z-Score Scaling)

Setelah seluruh fitur ditransformasikan ke dalam bentuk numerik, langkah terakhir dalam persiapan data adalah melakukan standardisasi. Keenam fitur operasional yang terpilih memiliki rentang skala yang sangat berbeda, di mana variabel **BIAYA** memiliki nilai hingga jutaan rupiah, sementara variabel lainnya berada pada rentang angka kecil.

Untuk mencegah variabel dengan rentang nilai besar mendominasi perhitungan jarak *Euclidean*, dilakukan proses *scaling* menggunakan metode *Z-Score (StandardScaler)*. Penting untuk dicatat bahwa proses standardisasi ini hanya diterapkan pada 4 fitur numerik tersebut. Sementara itu, fitur hasil *One-Hot Encoding* dipertahankan dalam nilai biner aslinya (0 dan 1) agar tidak kehilangan makna eksistensi atribut kategorikalnya.

Berikut adalah 5 sampel baris pertama dari *dataset* final yang telah melalui seluruh proses persiapan data dan siap untuk dimasukkan ke dalam tahap pemodelan algoritma AHC:

Tabel 4.3 Sampel Dataset Final (Input Pemodelan)

Fitur/Variabel	Sampel 1	Sampel 2
BIAYA	29,876,616,456,289,900	- 12,882,132,986,040,700
DURASI_HARI	15,635,065,151,104,700	- 3,146,679,778,838,700
JUMLAH_KOMPONEN	15,397,697,004,668,600	- 1,580,402,512,006,960
SKOR_KESULITAN	17,315,937,252,293,800	- 13,254,348,704,953,000
AKSI_SERVIS_STD_GANTI	0	0
AKSI_SERVIS_STD_HARDWARE_REPAIR	1	0
AKSI_SERVIS_STD_JASA	0	0
AKSI_SERVIS_STD_SOFTWARE	0	1
KOMPONEN_UTAMA_BATERAI	0	0
KOMPONEN_UTAMA_BODY_FISIK	0	0
KOMPONEN_UTAMA_FLEXIBEL	0	0
KOMPONEN_UTAMA_KAMERA	0	0
KOMPONEN_UTAMA_KONEKTOR_CAS	0	0
KOMPONEN_UTAMA_LCD	0	0
KOMPONEN_UTAMA_MESIN_IC	1	0
KOMPONEN_UTAMA_SISTEM_SOFTWARE	0	1
KOMPONEN_UTAMA_SMALL_PART	0	0

S		
KOMPONEN_UTAMA_SPEAKER_AU DIO	0	0

Tabel 4.3 di atas menunjukkan representasi nilai yang telah siap diproses oleh algoritma Agglomerative Hierarchical Clustering (AHC). Seluruh fitur numerik (BIAYA, DURASI_HARI, JUMLAH_KOMPONEN, dan SKOR_KESULITAN) kini berada pada skala yang seragam hasil dari proses Z-Score Scaling, di mana nilai positif menunjukkan posisi data di atas rata-rata dataset, sedangkan nilai negatif menunjukkan posisi data di bawah rata-rata.

Sementara itu, fitur kategorikal yang telah melalui proses One-Hot Encoding direpresentasikan dalam nilai biner 0 dan 1. Angka 1 menunjukkan bahwa transaksi tersebut memiliki karakteristik kategori terkait (misalnya jenis komponen tertentu), sedangkan angka 0 menunjukkan ketiadaan karakteristik tersebut. Kombinasi antara data numerik yang terstandarisasi dan data biner ini memastikan perhitungan jarak Euclidean dalam pemodelan AHC berjalan secara objektif dan seimbang.

4.3. Hasil Pemodelan *Clustering* Menggunakan AHC

Tahap pemodelan dilakukan dengan menerapkan algoritma *Agglomerative Hierarchical Clustering* (AHC) pada 811 baris data transaksi yang telah melalui proses persiapan. Proses pemodelan ini bertujuan untuk membangun struktur pengelompokan kasus servis secara hierarkis berdasarkan kemiripan nilai fitur operasional.

4.3.1. Konfigurasi Model dan Eksekusi Algoritma

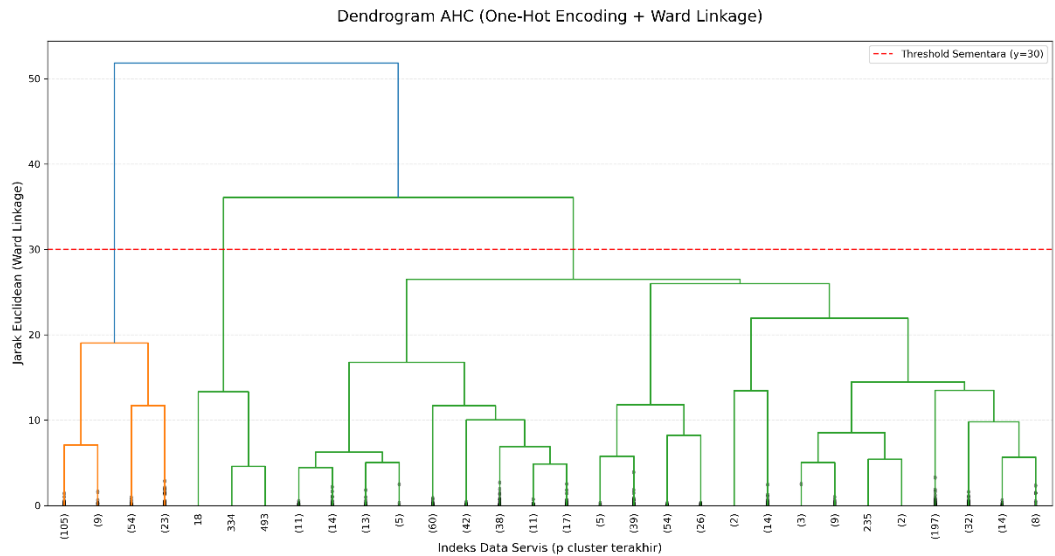
Dalam penelitian ini, algoritma AHC dikonfigurasi dengan parameter teknis utama sebagai berikut:

1. Metode Linkage: *Ward Linkage*. Pemilihan metode ini didasarkan pada kemampuannya dalam menghasilkan klaster yang kompak dengan cara meminimalkan total variansi di dalam klaster (*within-cluster sum of squares*).
2. Metrik Jarak: *Euclidean Distance*. Metrik ini digunakan untuk menghitung jarak geometris antar-titik data dalam ruang multidimensi dari 19 dimensi fitur (kombinasi fitur numerik dan kolom biner hasil *One-Hot Encoding*) yang telah ditentukan.

Secara sistematis, algoritma memulai dengan menganggap setiap baris data sebagai satu klaster tunggal, kemudian secara iteratif menggabungkan pasangan klaster yang memiliki peningkatan variansi terkecil hingga seluruh data membentuk satu kesatuan hierarki.

4.3.2. Visualisasi Dendrogram

Hasil dari proses penggabungan hierarki tersebut direpresentasikan melalui visualisasi dendrogram. Dendrogram berfungsi sebagai peta visual yang menunjukkan urutan penggabungan data dan jarak disimilaritas antar-objek.



Gambar 4.1 Visualisasi Dendrogram Hasil Pemodelan AHC

Dendrogram tersebut merepresentasikan struktur organisasi dari 811 data transaksi servis Konter Anjas Reparasi yang dikelompokkan berdasarkan kemiripan fitur operasionalnya. Sumbu vertikal (kiri) menunjukkan jarak *Euclidean* yang dihitung menggunakan metode *Ward Linkage*, di mana tinggi garis vertikal mencerminkan tingkat ketidakmiripan (*dissimilarity*) antar-objek atau kluster. Semakin tinggi garis penghubung, semakin besar perbedaan karakteristik antar-data tersebut.

Untuk menentukan jumlah kelompok yang paling representatif, dilakukan pengamatan terhadap struktur cabang hierarki. Peneliti menetapkan garis ambang batas (*threshold*) pada nilai jarak $y = 30$ (ditunjukkan dengan garis putus-putus merah pada Gambar 4.1). Pemotongan pada tingkat ini menghasilkan pembentukan empat kluster utama ($k = 4$).

Keputusan untuk melakukan pemotongan pada level tersebut didasarkan pada pengamatan visual terhadap cabang-cabang dendrogram yang memiliki rentang vertikal cukup panjang (jarak pemisahan yang lebar) sebelum bergabung

kembali. Hal ini mengindikasikan bahwa pada konfigurasi $k = 4$, setiap kelompok memiliki profil karakteristik yang kontras dan unik secara operasional. Validasi secara kuantitatif untuk menguji kualitas pemisahan pada jumlah kluster ini akan dibahas lebih mendalam menggunakan metrik evaluasi pada subbab berikutnya.

4.4. Evaluasi Model dan Penentuan Jumlah Kluster Optimal

Tahap evaluasi dilakukan untuk menguji kualitas pengelompokan yang dihasilkan oleh algoritma AHC dan menentukan jumlah kluster (k) yang paling optimal secara matematis maupun operasional. Pengujian dilakukan dengan membandingkan dua metrik utama, yaitu *Silhouette Score* untuk mengukur kekuatan pemisahan antar-kluster dan *Davies-Bouldin Index* (DBI) untuk mengukur kepadatan serta jarak pemisahan kluster.

4.4.1. Perbandingan Metrik Evaluasi

Peneliti melakukan serangkaian eksperimen dengan menguji beberapa kandidat jumlah kluster untuk menemukan konfigurasi terbaik. Mengingat *dataset* telah ditransformasi menggunakan *One-Hot Encoding* (OHE), dimensi fitur meluas menjadi 19 kolom, yang secara alami memberikan tantangan lebih tinggi dalam pemisahan ruang vektor. Hasil perbandingan nilai evaluasi disajikan pada Tabel 4.4 berikut:

Tabel 4.4 Perbandingan Hasil Evaluasi Model AHC

Jumlah Kluster (k)	Silhouette Score	DBI Score
$k = 2$	0.4342	0.9199
$k = 3$	0.4479	0.7592
$k = 4$	0.3402	1.1867

$k = 5$	0.4258	1.0516

Berdasarkan Tabel 4.4, nilai *Silhouette Score* tertinggi dan nilai DBI terendah (paling optimal) secara matematis dicapai pada $k = 3$ dengan skor 0.4479. Skor ini menunjukkan bahwa struktur kluster berada dalam kategori pemisahan yang cukup (*fair structure*). Meskipun terdapat penurunan nilai dibandingkan pengujian sebelumnya, hal ini merupakan konsekuensi logis dari penerapan *One-Hot Encoding* yang menghilangkan asumsi jarak palsu antar-fitur nominal, sehingga menghasilkan struktur kluster yang lebih objektif.

4.4.2. Penentuan Jumlah Kluster Optimal

Meskipun secara matematis $k = 3$ memiliki performa metrik yang paling unggul, penelitian ini menetapkan $k = 4$ sebagai jumlah kluster optimal yang diimplementasikan pada sistem *dashboard* KONTER ANJAS ANALYTICS. Keputusan ini diambil berdasarkan prinsip integrasi antara validasi statistik dengan kebutuhan operasional bisnis (*expert judgment*), dengan pertimbangan sebagai berikut:

1. Isolasi Kasus Indent Part: Pada konfigurasi $k = 4$, algoritma berhasil memisahkan 3 data transaksi dengan durasi pengerjaan ekstrem (rata-rata 90,67 hari) ke dalam kelompok tersendiri yaitu "Perbaikan *Hardware (Indent Part)*". Jika menggunakan $k = 3$, kasus-kasus kritis ini akan melebur ke dalam kluster perbaikan *hardware* lainnya, sehingga pemilik konter kehilangan kemampuan untuk mengidentifikasi hambatan rantai pasok secara spesifik.

2. Kebutuhan Prioritas Stok: Pembagian ke dalam 4 kelompok memberikan segmentasi kebijakan stok yang lebih tajam: Layanan tanpa stok fisik (Software), stok utama (*Ready Stock*), stok cadangan terbatas (*Buffer Stock*), dan kategori non-stok/pre-order (*Indent Part*).
3. Toleransi Metrik: Penurunan *Silhouette Score* pada $k = 4$ (menjadi 0.3402) dinilai sebagai kompromi yang dapat diterima demi mendapatkan fungsionalitas sistem pendukung keputusan yang lebih aplikatif dan bermanfaat bagi pengguna akhir.

4.4.3. Distribusi Anggota Klaster

Setelah menetapkan $k = 4$, langkah selanjutnya adalah melihat sebaran data pada masing-masing klaster. Distribusi akhir dari 811 data transaksi servis menggunakan data OHE menunjukkan pergeseran proporsi yang lebih mendalam pada tingkat kesulitan hardware, Sebagaimana disajikan pada Tabel 4.5 berikut:

Tabel 4.5 Distribusi Kasus Servis per Tingkat Layanan

Kategori Tingkat Layanan	Jumlah Transaksi	Persentase (%)
Perbaikan Hardware Intensif	406	249,729%
Pergantian Komponen Fisik	211	26,02%
Layanan Jasa & Software	191	23,55%
Perbaikan Hardware (Indent Part)	3	0,37%
Total	811	100,00%

Data menunjukkan bahwa operasional Konter Anjas Reparasi saat ini didominasi oleh klaster Perbaikan *Hardware* Intensif (50,06%). Hal ini

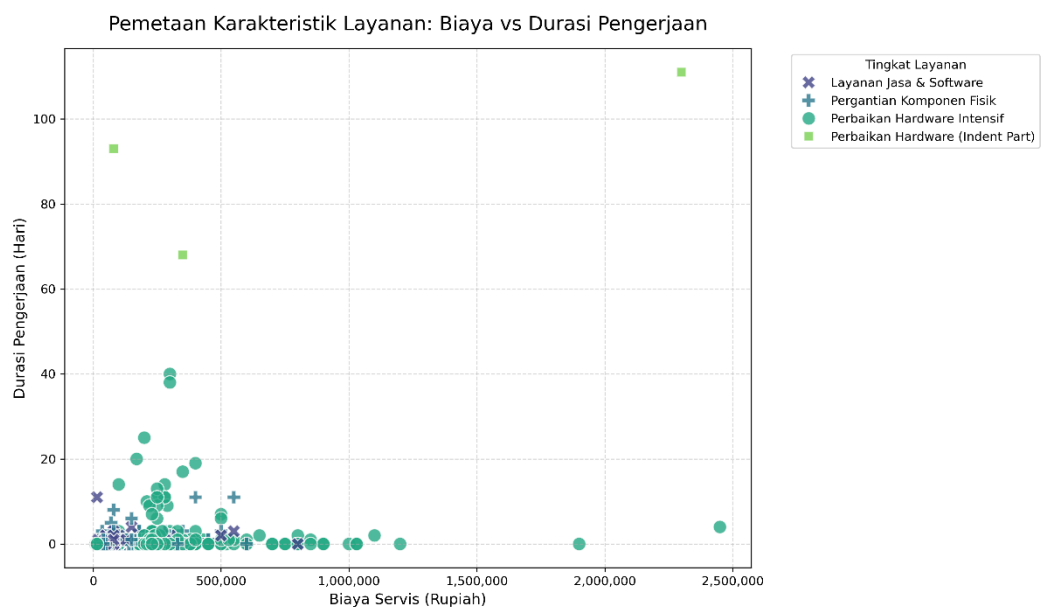
mengindikasikan bahwa mayoritas pengerjaan di konter memiliki tingkat risiko dan kesulitan yang tinggi. Sementara itu, kelompok *Indent Part* meskipun hanya berjumlah 3 transaksi (0,37%), memiliki pengaruh biaya dan durasi yang sangat signifikan, sehingga tetap dipertahankan sebagai kluster mandiri demi akurasi perencanaan stok jangka panjang.

4.5. Profiling dan Interpretasi Kluster

Setelah menetapkan jumlah kluster optimal $k = 4$, dilakukan tahap *profiling* untuk memahami karakteristik unik dari setiap kelompok. Tahap ini bertujuan untuk memberikan interpretasi bisnis terhadap hasil pengelompokan berdasarkan nilai rata-rata (*mean*) dan sebaran data pada setiap variabel operasional.

4.5.1. Analisis Visual Karakteristik Layanan

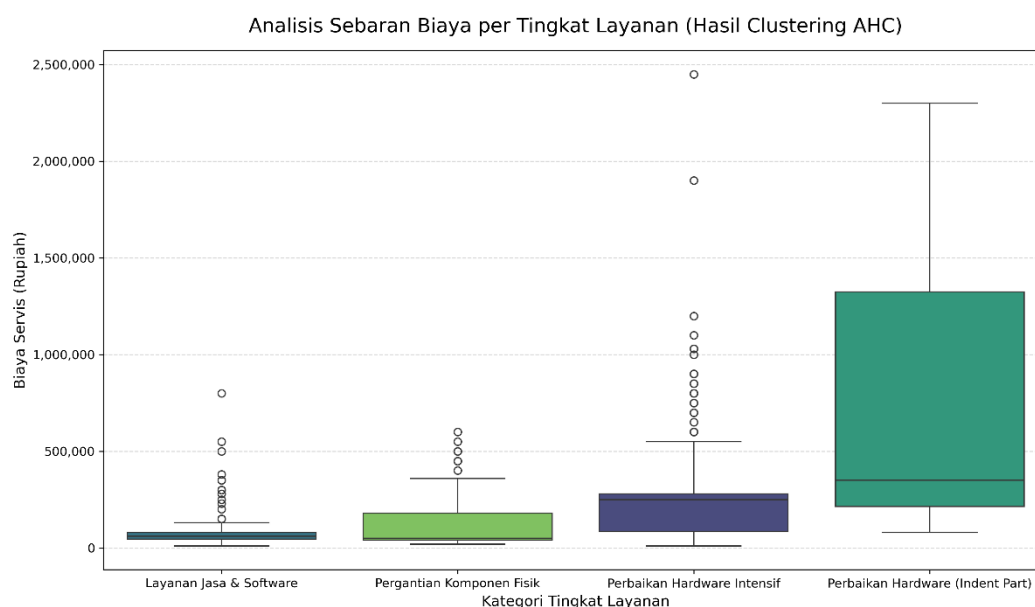
Secara visual, perbedaan karakteristik antar-kluster dapat diamati melalui pemetaan antara variabel Biaya dan Durasi Pengerjaan, serta sebaran biaya pada setiap kategori layanan.



Gambar 4.2 Pemetaan Karakteristik Layanan: Biaya vs Durasi

Pengerjaan

Berdasarkan Gambar 4.2, terlihat pemisahan yang sangat kontras pada kluster Perbaikan Hardware (Indent Part) yang memiliki durasi pengerjaan ekstrem mencapai puluhan hingga ratusan hari. Di sisi lain, kluster Perbaikan Hardware Intensif mendominasi spektrum biaya menengah ke atas, mencerminkan kompleksitas teknis yang lebih tinggi dibandingkan layanan reguler.



Gambar 4.3 Analisis Sebaran Biaya per Tingkat Layanan

Gambar 4.3 memperkuat temuan tersebut melalui visualisasi *boxplot*. Kluster Perbaikan Hardware (Indent Part) memiliki rentang biaya yang luas dengan nilai rata-rata tertinggi, disusul oleh kluster Perbaikan Hardware Intensif yang memiliki banyak pencilan (*outlier*) biaya di atas Rp 500.000. Sebaliknya, kluster Layanan Jasa & Software terkonsentrasi pada rentang biaya paling rendah dan stabil.

4.5.2. Interpretasi Karakteristik Tiap Klaster

Berdasarkan ringkasan statistik dan profil data hasil transformasi OHE, berikut adalah interpretasi mendalam untuk masing-masing klaster:

1. Layanan Jasa & Software
 - a. Profil: Memiliki rata-rata biaya paling rendah (Rp 82.592) dengan durasi pengerjaan tercepat (0,27 hari). Skor kesulitan pada kelompok ini sangat minim (1,60).
 - b. Interpretasi: Klaster ini merepresentasikan servis ringan non-fisik (seperti penanganan sistem *software*) atau jasa pengecekan. Secara operasional, kelompok ini tidak memerlukan manajemen stok *sparepart* fisik.
2. Pergantian Komponen Fisik
 - a. Profil: Memiliki rata-rata biaya menengah (Rp 111.588) dengan durasi yang sangat efisien (0,46 hari). Skor kesulitan berada pada tingkat moderat (4,93).
 - b. Interpretasi: Ini merupakan layanan penggantian komponen *fast-moving* yang bersifat modular (seperti konektor *cas* atau tombol). Kelompok ini menjadi prioritas tinggi untuk ketersediaan stok harian karena pengerjaannya yang cepat dan frekuensinya yang stabil.
3. Perbaikan Hardware Intensif
 - a. Profil: Menjadi klaster dengan populasi terbesar (50,1%). Memiliki rata-rata biaya yang cukup tinggi (Rp 249.729) dengan skor kesulitan teknis tertinggi mencapai 7,56.

- b. Interpretasi: Kelompok ini mencakup perbaikan *hardware* berat (seperti LCD atau komponen mesin/IC). Dominasi klaster ini menunjukkan bahwa operasional konter berfokus pada pengerjaan berisiko tinggi yang membutuhkan ketersediaan stok komponen utama.

4. Perbaikan Hardware (Indent Part)

- a. Profil: Memiliki karakteristik durasi pengerjaan yang sangat ekstrem (rata-rata 90,67 hari) dengan rata-rata biaya paling tinggi (Rp 910.000).
- b. Interpretasi: Klaster ini merepresentasikan kasus di mana komponen (seperti *body* fisik langka) harus dipesan secara khusus. Identifikasi kelompok ini sangat penting untuk memberikan transparansi estimasi waktu kepada pelanggan dan menghindari penumpukan stok mati (*dead-stock*).

Tabel 4.6 Ringkasan Profil Karakteristik Klaster

Atribut (Rata-rata)	Jasa & Software	Komponen Fisik	Hardware Intensif	Indent Part
Biaya (Rp)	82,592	111,587	249,729	910,000
Durasi (Hari)	0.27	0.46	1	90.67
Skor Kesulitan	1.6	4.93	7.56	7
Jml. Komponen	0.12	1.23	1.11	1.67

4.6. Rekomendasi Prioritas Pengadaan Sparepart

Hasil akhir dari pemodelan dan *profiling* klaster ditransformasikan menjadi wawasan bisnis (*business insight*) yang aplikatif. Rekomendasi ini dirancang untuk mengatasi masalah utama di Konter Anjas Reparasi, yaitu kesulitan dalam memprioritaskan belanja stok *sparepart* yang berujung pada penumpukan modal atau kekosongan barang operasional.

Berdasarkan analisis tingkat kekritisannya (durasi dan biaya) serta volume transaksi pada tiap klaster, hierarki prioritas pengadaan *sparepart* dirumuskan sebagai berikut:

1. Prioritas Strategis (Tinggi): Klaster Perbaikan Hardware Intensif
 - a. Klaster ini merupakan pusat gravitasi operasional konter dengan volume mencapai 50,06% dari total transaksi. Karena biaya perbaikannya yang cukup tinggi dan skor kesulitannya yang maksimal, komponen pada klaster ini menjadi prioritas utama untuk selalu tersedia.
 - b. Komponen Dominan: Berdasarkan data historis, komponen yang paling wajib tersedia (*ready stock*) untuk mendukung klaster ini adalah LCD dan Baterai. Menjamin ketersediaan LCD akan secara signifikan mengamankan omzet harian konter.
2. Prioritas Operasional (Menengah): Klaster Pergantian Komponen Fisik
 - a. Meskipun volumenya berada di angka **26,02%**, klaster ini memiliki durasi pengerjaan yang sangat cepat (rata-rata 0,46 hari). Komponen di kategori ini bersifat *fast-moving* dan modular.
 - b. Komponen Dominan: Komponen utama yang terdeteksi pada klaster ini meliputi Konektor Cas dan Tombol. Rekomendasinya

adalah menyediakan stok cadangan (*buffer stock*) yang memadai agar pelayanan bisa selesai di hari yang sama tanpa menunggu pengiriman barang.

3. Prioritas Rendah (On-Demand): Klaster Perbaikan Hardware (Indent Part)
 - a. Klaster ini ditandai dengan durasi pengerjaan yang sangat ekstrem (rata-rata 90,67 hari), yang mengindikasikan kelangkaan *sparepart* di pasar lokal.
 - b. Komponen Dominan: Kasus langka yang terdeteksi meliputi Body Fisik khusus. Untuk komponen ini, konter direkomendasikan untuk tidak melakukan penyetokan guna menghindari risiko modal mengendap (*dead-stock*). Pengadaan hanya dilakukan berdasarkan pesanan khusus (*pre-order*).
4. Tidak Membutuhkan Stok (N/A): Klaster Layanan Jasa & Software
 - a. Perbaikan pada klaster ini (seperti *flashing program* atau lupa *password*) murni mengandalkan keahlian teknisi dan perangkat lunak. Oleh karena itu, tidak ada alokasi anggaran belanja fisik untuk kategori ini.

Ringkasan rekomendasi prioritas pengadaan *sparepart* disajikan pada Tabel 4.7 berikut:

Tabel 4.7 Hierarki Rekomendasi Prioritas Pengadaan *Sparepart*

Prioritas	Kategori Klaster	Rekomendasi Tindakan	Contoh Komponen Dominan
Tinggi	Perbaikan Hardware Intensif	Sediakan ready stock harian (Wajib Ada)	LCD, Baterai
Menengah	Pergantian	Sediakan buffer	Konektor Cas,

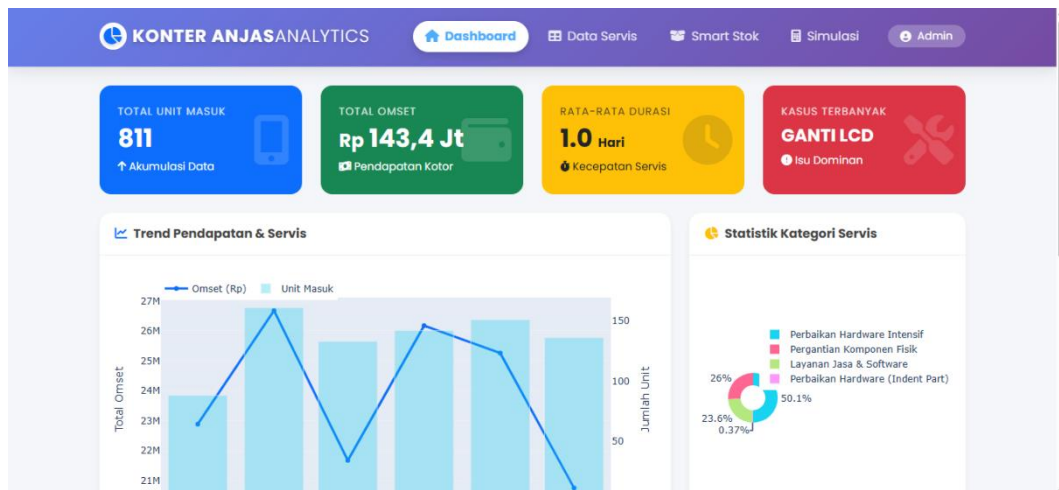
	Komponen Fisik	stock proporsional	Tombol
Rendah	Perbaikan Hardware (Indent)	Tidak perlu stok. Lakukan pemesanan on-demand.	Body Fisik, IC Langka
N/A	Layanan Jasa & Software	Fokus pada update tools teknisi.	Sistem/Software

4.7. Implementasi Dashboard (Deployoment)

Tahap akhir dari kerangka kerja CRISP-DM adalah *Deployment*, yaitu mengimplementasikan model yang telah dievaluasi ke dalam lingkungan produksi. Dalam penelitian ini, hasil pengelompokan algoritma AHC diintegrasikan ke dalam sebuah *Decision Support System* (DSS) berbasis *web* bernama KONTER ANJAS ANALYTICS.

Sistem ini dibangun menggunakan bahasa pemrograman Python dengan *framework* Flask sebagai *backend*, antarmuka Bootstrap untuk *frontend* yang responsif, serta *library* Plotly untuk visualisasi data interaktif. Sistem ini terbagi menjadi empat modul utama yang terintegrasi langsung dengan 811 data transaksi hasil klasterisasi.

4.7.1. Modul Executive Dashboard



Gambar 4.4 Tampilan Antarmuka *Executive Dashboard*

Modul ini berfungsi sebagai pusat kendali visual yang menyajikan ringkasan (*snapshot*) operasional konter.

1. KPI Cards (Key Performance Indicators): Di bagian atas, sistem secara dinamis menghitung metrik utama dari *dataset* historis. Terdapat informasi Total Unit Masuk (811 transaksi), Total Omset kotor (Rp 143,4 Juta), Rata-rata Durasi kecepatan servis (1.0 Hari), serta Isu Dominan kerusakan terbanyak (Ganti LCD).
2. Visualisasi Analitik Terintegrasi: Sistem menampilkan kombinasi grafik *Trend* Pendapatan bulanan, grafik batang *Top* Performa Teknisi (didominasi Teknisi D dengan 634 kasus), dan distribusi *Top* 5 Jenis Kerusakan. Selain itu, terdapat *Donut Chart* persentase kategori servis yang secara interaktif memperlihatkan dominasi kluster Perbaikan *Hardware* Intensif (50,1%).

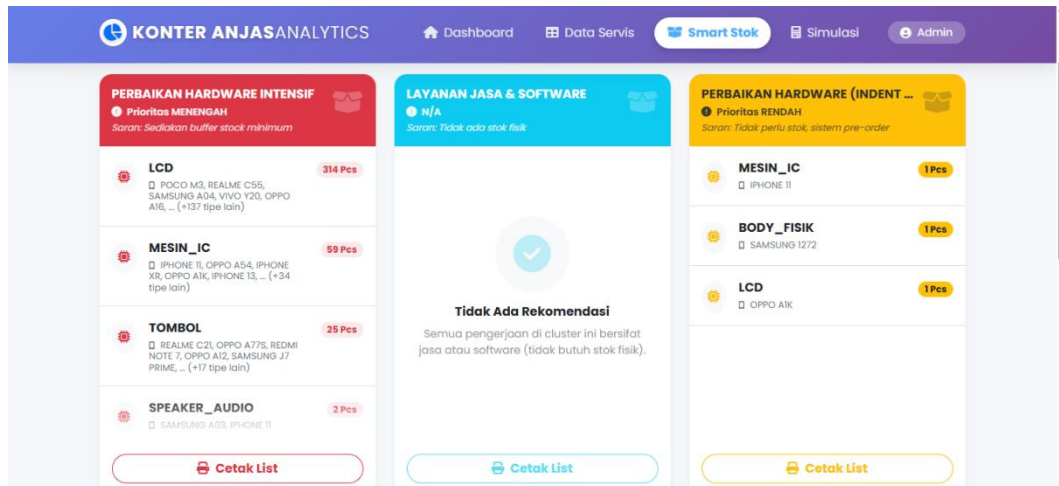
4.7.2. Modul Data Servis

NO	TGL MASUK	TGL KELUAR	DURASI	TIPE HP	KERUSAKAN	CLUSTER	BIAYA (RP)	TEKNISI
1	02-05-2025	04-05-2025	2 Hari	IPHONE 11	MATI TOTAL TOTAL MASUK AIR	PERBAIKAN HARDWARE INTENSIF	800.000	A
2	02-05-2025	03-05-2025	1 Hari	POCO X3 PRO	CEK FLASH PROGRAM	LAYANAN JASA & SOFTWARE	150.000	D
3	02-05-2025	03-05-2025	1 Hari	POCO M3	GANTI LCD	PERBAIKAN HARDWARE INTENSIF	250.000	D
4	03-05-2025	03-05-2025	0 Hari	REALME C55	GANTI LCD	PERBAIKAN HARDWARE INTENSIF	300.000	D
5	03-05-2025	03-05-2025	0 Hari	REDMI 9	SERVICE	LAYANAN JASA & SOFTWARE	80.000	D
6	03-05-2025	03-05-2025	0 Hari	SAMSUNG A04	GANTI LCD	PERBAIKAN HARDWARE INTENSIF	250.000	D

Gambar 4.5 Tampilan Modul *Database Riwayat Servis*

Modul ini merupakan representasi *database* terstruktur yang menampilkan riwayat transaksi servis secara lengkap. Tabel ini dilengkapi dengan fitur penomoran halaman (*pagination*) dan kolom pencarian (*search bar*) yang responsif. Nilai tambah dari modul ini adalah penyematan label visual (*badge* berwarna) pada setiap baris data untuk menunjukkan kategori klaster kasus tersebut (misalnya, warna merah muda/merah untuk "Perbaikan *Hardware* Intensif" dan biru muda/cyan untuk "Layanan Jasa & Software"), sehingga pemilik konter dapat langsung mengenali tingkat kompleksitas setiap rekam jejak transaksi.

4.7.3. Modul Smart Stock (Rekomendasi Prioritas)



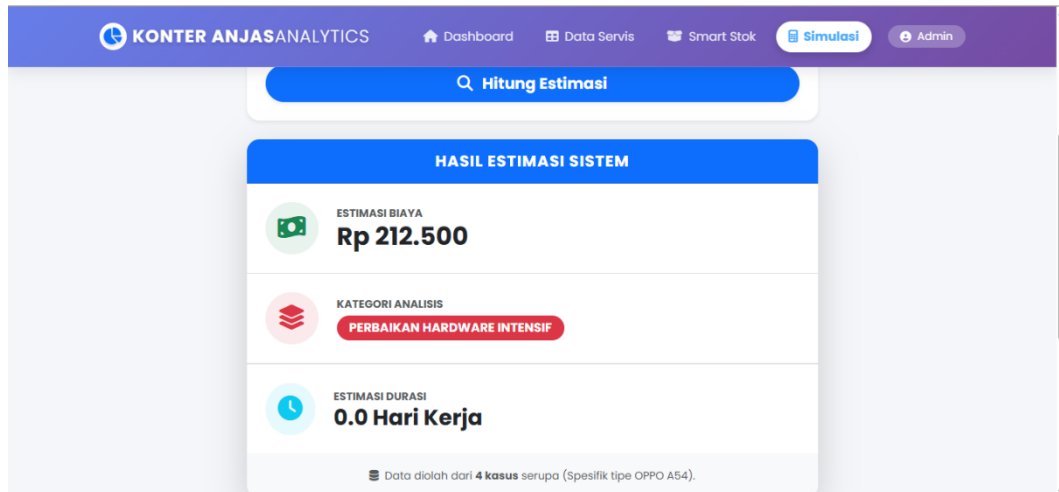
Gambar 4.6 Tampilan Modul *Smart Stok* (Rekomendasi Prioritas)

Modul ini adalah luaran fungsional utama dari penelitian, di mana wawasan analitik AHC diterjemahkan menjadi rekomendasi pengadaan *sparepart*. Sistem secara cerdas menyaring komponen fisik dan memetakannya ke dalam empat kartu prioritas (*Priority Cards*):

1. Pergantian Komponen Fisik (Warna Biru - Prioritas Tinggi):
Menguatkan penyediaan *ready stock* harian, menampilkan komponen *fast-moving* terbanyak seperti Konektor Cas (69 Pcs) dan Baterai (47 Pcs).
2. Perbaikan Hardware Intensif (Warna Merah - Prioritas Menengah):
Menguatkan pengadaan *buffer stock* minimum, didominasi oleh LCD (314 Pcs) dan Mesin/IC (59 Pcs).
3. Perbaikan Hardware Indent Part (Warna Kuning - Prioritas Rendah):
Memberikan peringatan untuk tidak melakukan stok fisik melainkan sistem *pre-order* (misal: Mesin_IC tipe spesifik).
4. Layanan Jasa & Software (Warna Cyan - N/A): Mengonfirmasi ketiadaan rekomendasi barang fisik.

Modul ini juga dilengkapi dengan fitur cetak rekapitulasi (Tombol "Cetak List"), yang menghasilkan dokumen rincian kebutuhan stok lengkap dengan daftar spesifik tipe *handphone* yang kompatibel, sehingga siap digunakan sebagai daftar belanja riil oleh pemilik konter.

4.7.4. Modul Simulasi Servis



Gambar 4.7 Tampilan *Output* Simulasi Estimasi Servis

Modul ini dirancang sebagai alat bantu estimasi operasional untuk kasus servis baru di masa depan. Berdasarkan arsitektur yang telah dirancang pada Bab 3, modul ini tidak menggunakan model prediksi terawasi (*supervised learning*), melainkan menerapkan metode pencocokan data historis (*historical data matching*).

Pengguna cukup memasukkan Tipe Handphone (contoh: OPPO A54) dan Jenis Kerusakan (contoh: Ganti LCD). Sistem kemudian menyaring riwayat transaksi dengan kasus serupa dan menghitung nilai agregasinya (*mean* dan *modus*). Hasil akhirnya menampilkan rentang Estimasi Biaya (misal: Rp 212.500), Estimasi Durasi pengerjaan (0.0 Hari), label Kategori Analisis (*Cluster*), serta tabel histori 5 kasus terakhir yang dijadikan acuan dasar perhitungan tersebut.

4.8. Pengujian Sistem

Setelah tahap implementasi selesai, tahap selanjutnya adalah melakukan pengujian perangkat lunak untuk memastikan bahwa sistem *dashboard* KONTER ANJAS ANALYTICS berjalan sesuai dengan fungsionalitas yang diharapkan. Pengujian pada penelitian ini menggunakan metode *Black Box Testing* dengan pendekatan *Equivalence Partitioning*.

Metode *Black Box* berfokus pada spesifikasi fungsional dari perangkat lunak, di mana pengujian dilakukan dengan memberikan *input* tertentu untuk melihat apakah *output* yang dihasilkan oleh sistem sudah sesuai dengan perancangan tanpa perlu melihat struktur kode program di dalamnya.

Skenario pengujian difokuskan pada empat modul utama sistem, yaitu *Executive Dashboard*, Modul Data Servis, Modul *Smart Stok*, dan Modul Simulasi. Hasil pengujian fungsionalitas sistem disajikan pada Tabel 4.8 berikut:

Tabel 4.8 Skenario dan Hasil *Black Box Testing*

No	Modul Skenario Pengujian	Input Tindakan	Output yang Diharapkan	Hasil Pengujian	Kesimpulan
1	Executive Dashboard: Memuat halaman utama saat aplikasi dijalankan.	Pengguna membuka rute (/ atau /dashboard) di browser.	Sistem menampilkan KPI Cards (Total Transaksi 811, Total Omset), grafik pendapatan, persentase klaster, dan performa teknis secara akurat dari dataset.	Sesuai	Valid
2	Data Servis: Menguji fitur pencarian (Search)	Pengguna mengetikkan kata kunci "GANTI	Tabel secara reaktif hanya menampilkan baris transaksi	Sesuai	Valid

	pada tabel riwayat transaksi.	LCD" pada kolom Search.	yang memuat teks "GANTI LCD", lengkap dengan badge warna klasternya.		
3	Smart Stok: Menguji integrasi tabel rekomendasi sparepart.	Pengguna membuka menu Smart Stok (/stok).	Sistem memisahkan komponen ke dalam kartu klaster yang tepat (misal: LCD di Prioritas Menengah/Tinggi) sesuai perhitungan agregasi algoritma AHC.	Sesuai	Valid
4	Smart Stok: Menguji fitur Ekspor/Cetak (Print List).	Pengguna menekan tombol "Cetak List" pada salah satu klaster.	Sistem membuka tab baru berisi dokumen format cetak (print-ready) daftar belanja komponen beserta jumlah (Qty) dan model handphone terkait.	Sesuai	Valid
5	Simulasi Servis: Menguji fitur estimasi dengan input data yang valid / riwayat tersedia.	Pengguna memasukkan Tipe: "OPPO A54" dan Kerusakan: "LCD", lalu klik "Hitung Estimasi".	Sistem berhasil menemukan kecocokan data historis dan menampilkan Estimasi Biaya, Estimasi Durasi, kategori Klaster, dan histori 5 transaksi terakhir.	Sesuai	Valid
6	Simulasi Servis: Menguji validasi penanganan error (Error Handling) pada input kosong atau data tidak	Pengguna memasukkan kerusakan aneh yang tidak ada di riwayat (misal: "MESIN CUCI").	Sistem tidak mengalami crash, melainkan menampilkan pesan peringatan (Pesan Error): "Maaf, data untuk kerusakan tersebut belum tersedia."	Sesuai	Valid

	ditemukan.				
--	------------	--	--	--	--

4.8.1. Kesimpulan Pengujian Sistem

Berdasarkan hasil *Black Box Testing* yang dilakukan pada 6 skenario utama di atas, seluruh fitur operasional pada sistem *dashboard* KONTER ANJAS ANALYTICS telah berjalan 100% sesuai dengan fungsionalitas yang dirancang. Sistem mampu menangani alur pembacaan data *One-Hot Encoding* hasil *clustering* AHC, melakukan agregasi agregasi visual, serta memproses estimasi historis tanpa adanya *error* atau kegagalan sistem (*system crash*). Hal ini membuktikan bahwa tahap *Deployment* pada siklus CRISP-DM telah berhasil diwujudkan menjadi perangkat lunak yang layak pakai.

4.9. Pembahasan dan Analisis Hasil

Tahap pembahasan dan analisis hasil merupakan sintesis dari seluruh rangkaian pemodelan data transaksi servis di Konter Anjas Reparasi. Berdasarkan evaluasi algoritma dan implementasi sistem pendukung keputusan yang telah diuraikan pada sub-bab sebelumnya, terdapat beberapa temuan krusial yang dapat dianalisis baik dari perspektif metodologi *data science* maupun dari perspektif strategis manajemen persediaan bisnis.

4.9.1. Analisis Metodologis dan Kualitas Pengelompokan

Secara metodologis, penerapan teknik *One-Hot Encoding* (OHE) untuk mentransformasi atribut nominal (Aksi Servis dan Komponen Utama) terbukti memberikan landasan perhitungan jarak *Euclidean* yang lebih objektif dibandingkan representasi ordinal. Pemecahan fitur kategorikal menjadi matriks biner (0/1) berhasil menghilangkan bias jarak matematis palsu, sehingga algoritma

Agglomerative Hierarchical Clustering (AHC) dapat mengelompokkan data murni berdasarkan kemiripan beban operasional yang direpresentasikan melalui 19 dimensi fitur.

Selain itu, keputusan penelitian untuk mempertahankan jumlah klaster optimal sebanyak 4 kelompok ($k = 4$) meskipun konfigurasi 3 kelompok ($k = 3$) memiliki skor *Silhouette* yang sedikit lebih tinggi (0.4479 berbanding 0.3402) terbukti merupakan langkah yang sangat esensial secara operasional. Analisis menunjukkan bahwa konfigurasi 4 klaster mampu mengisolasi kasus anomali berdurasi ekstrem (rata-rata 90,67 hari) ke dalam kelompok mandiri yaitu "Perbaikan *Hardware (Indent Part)*". Pengorbanan minor pada metrik statistik tersebut terbayar dengan tingginya nilai interpretasi bisnis, di mana sistem dapat membedakan secara tegas antara barang yang harus distok dan barang yang murni mengandalkan *pre-order*.

4.9.2. Analisis Pergeseran Strategi Persediaan (Wawasan Bisnis)

Hasil *profiling* klaster mengungkap fakta operasional yang sangat berharga bagi manajemen Konter Anjas Reparasi. Sebelum adanya sistem ini, pengadaan barang mungkin dilakukan secara merata atau hanya berdasarkan intuisi. Namun, data faktual dari 811 transaksi menunjukkan realitas yang berbeda:

1. Dominasi Kasus Kompleks: Sebanyak 50,06% dari total pengerjaan ternyata masuk ke dalam klaster "Perbaikan *Hardware Intensif*". Hal ini mengindikasikan bahwa konter lebih banyak menangani kasus dengan tingkat kesulitan tinggi dan biaya mahal dibandingkan servis ringan. Oleh karena itu, prioritas belanja modal perusahaan harus secara strategis

dialokasikan lebih besar untuk mengamankan komponen utama seperti LCD dan mesin (IC).

2. Efisiensi Kasus Ringan: Klaster "Pergantian Komponen Fisik" (26,02%) seperti konektor *cas* atau tombol, memiliki perputaran waktu yang sangat cepat (0,46 hari). Komponen di klaster ini tidak memerlukan modal besar namun menuntut ketersediaan fisik setiap hari (*buffer stock*) guna menjaga tingkat kepuasan pelanggan atas layanan *same-day service*.
3. Reduksi Risiko *Dead-Stock*: Adanya klaster "Layanan Jasa & Software" (23,55%) membuktikan bahwa hampir seperempat pendapatan konter dihasilkan murni dari keahlian teknisi tanpa melibatkan perpindahan barang fisik. Analisis ini mencegah pemilik konter membuang anggaran untuk komponen yang sebenarnya tidak relevan dengan tren perbaikan *software*.

4.9.3. Analisis Dampak Sistem Pendukung Keputusan (*Deployment*)

Transformasi hasil *clustering* ke dalam bentuk aplikasi *dashboard* interaktif (KONTER ANJAS ANALYTICS) berhasil menjembatani kesenjangan antara kerumitan algoritma *machine learning* dengan kebutuhan praktis pengguna akhir. Modul *Smart Stok* membebaskan pemilik dari keharusan membaca data mentah, dengan langsung menyajikan rekomendasi tindakan (Prioritas Tinggi, Menengah, Rendah) yang siap dieksekusi (*actionable insights*).

Di sisi lain, penerapan Modul Simulasi Servis yang menggunakan pendekatan *historical data matching* (pencocokan data historis) terbukti menjadi solusi yang lebih aman dan transparan dibandingkan membangun model prediksi terawasi (*supervised learning*). Sistem mengedukasi pengguna dengan

memberikan estimasi biaya dan durasi beserta bukti jejak 5 kasus serupa di masa lalu. Secara keseluruhan, integrasi operasional ini menandai transisi Konter Anjas Reparasi dari manajemen persediaan tradisional yang reaktif menjadi sistem operasional berbasis data (*data-driven*) yang efisien.

BAB V

PENUTUP

5.1. Kesimpulan

Berdasarkan serangkaian tahapan penelitian yang telah dilakukan menggunakan metodologi CRISP-DM (*Cross-Industry Standard Process for Data Mining*) mulai dari pemahaman bisnis hingga implementasi sistem pada data transaksi servis Konter Anjas Reparasi, dapat ditarik beberapa kesimpulan sebagai berikut:

1. Penerapan Algoritma dan Evaluasi Model:

Penerapan algoritma *Agglomerative Hierarchical Clustering* (AHC) dengan metode *Ward Linkage* dan jarak *Euclidean* berhasil mengelompokkan 811 data transaksi secara objektif. Penggunaan teknik *One-Hot Encoding* (OHE) untuk mentransformasi atribut nominal terbukti mencegah bias jarak matematis, memperluas dimensi data menjadi 19 fitur yang seimbang setelah distandardisasi dengan *Z-Score*. Meskipun secara matematis jumlah kluster $k = 3$ memiliki *Silhouette Score* tertinggi (0.4479), penelitian ini membuktikan bahwa konfigurasi $k = 4$ (dengan *Silhouette Score* 0.3402) adalah model yang paling optimal. Hal ini dikarenakan konfigurasi tersebut mampu mengakomodasi justifikasi operasional bisnis (*expert judgment*) untuk mengisolasi kasus-kasus anomali berdurasi ekstrem (*indent part*).

2. Karakteristik Klaster dan Prioritas Pengadaan Stok:

Hasil *profiling* berhasil merumuskan empat klaster tingkat layanan yang secara langsung mendefinisikan hierarki prioritas pengadaan *sparepart*:

- a. Prioritas Tinggi (Strategis): Klaster *Perbaikan Hardware Intensif* (mendominasi 50,06% populasi). Memiliki kompleksitas dan biaya tinggi, sehingga komponen utama seperti LCD dan mesin wajib berstatus *ready stock*.
- b. Prioritas Menengah (Operasional): Klaster *Pergantian Komponen Fisik* (26,02%). Pengerjaannya sangat cepat, sehingga komponen *fast-moving* seperti konektor *cas* direkomendasikan sebagai *buffer stock*.
- c. Prioritas Rendah (*On-Demand*): Klaster *Perbaikan Hardware (Indent Part)* (0,37%). Memiliki durasi pengerjaan ekstrem (rata-rata 90 hari), sehingga direkomendasikan untuk tidak dilakukan penyetokan fisik, melainkan melalui sistem *pre-order*.
- d. Tidak Membutuhkan Stok: Klaster *Layanan Jasa & Software* (23,55%). Layanan murni mengandalkan keahlian teknisi sehingga tidak ada anggaran untuk stok fisik.

3. Implementasi Sistem Pendukung Keputusan (*Deployment*):

Hasil analitik algoritma AHC telah berhasil ditransformasikan menjadi perangkat lunak *Decision Support System* (DSS) berbasis *web* bernama KONTER ANJAS ANALYTICS menggunakan *framework* Flask. Sistem ini mengintegrasikan wawasan data menjadi rekomendasi tindakan melalui modul *Smart Stok* dan fitur *historical data matching* pada modul Simulasi. Berdasarkan hasil pengujian *Black Box Testing*, seluruh fitur sistem beroperasi 100% valid tanpa *error*, membuktikan bahwa sistem ini siap dan layak digunakan untuk membantu pemilik konter mengubah

pengambilan keputusan dari berbasis intuisi menjadi berbasis data (*data-driven*).

5.2. Saran

Meskipun penelitian ini telah berhasil mencapai tujuannya, masih terdapat ruang untuk pengembangan lebih lanjut. Adapun saran yang dapat diberikan untuk penelitian dan pengembangan sistem ke depannya adalah:

1. Pengembangan Algoritma dan Metodologi: Penelitian selanjutnya dapat membandingkan algoritma AHC dengan algoritma *clustering* yang memang dirancang khusus untuk menangani tipe data campuran (*mixed data* numerik dan kategorikal) seperti algoritma *K-Prototypes*. Penggunaan metrik jarak *Gower Distance* juga dapat dieksplorasi untuk memvalidasi apakah struktur klaster yang dihasilkan bisa lebih padat (*compact*).
2. Pengembangan Arsitektur Sistem: Sistem *dashboard* yang saat ini berjalan secara lokal (*localhost*) disarankan untuk diunggah ke peladen komputasi awan (*cloud hosting*). Selain itu, sistem dapat diintegrasikan secara *real-time* melalui API (*Application Programming Interface*) dengan aplikasi kasir (POS) Konter Anjas, sehingga *dataset* pemodelan dapat diperbarui secara otomatis tanpa harus melakukan *import* data CSV secara manual.
3. Pengayaan Fitur (*Feature Engineering*): Pada pengumpulan data operasional di masa depan, konter disarankan untuk mencatat variabel-variabel baru yang lebih detail, seperti tingkat pengalaman teknisi yang menangani, status garansi servis, atau sentimen keluhan pelanggan.

Variabel tambahan ini dapat memperkaya dimensi data untuk menghasilkan klusterisasi persediaan stok yang jauh lebih komprehensif.

DAFTAR PUSTAKA

- Abayomi, A. A., Mgbame, A. C., Akpe, O. E., Ogbuefi, E., & Adeyelu, O. O. (2023). Building competitive advantage through customized bi solutions in manufacturing SMEs. *International Journal of Social Science Exceptional Research*, 2(2), 48–65. <https://doi.org/10.54660/ijsser.2023.2.2.48-65>
- Abbey, A. B. N., Olaleye, I. A., Mokogwu, C., Olufemi-Phillips, A. Q., & Adewale, T. T. (2024). Developing inventory optimization frameworks to minimize economic loss in supply chain management. *International Journal of Advanced Economics*, 6(12), 826–836. <https://doi.org/10.51594/ijae.v6i12.1774>
- Achetoui, Z., Mabrouki, C., & Mousrij, A. (2019). A review of spare parts supply chain management. *Jurnal Sistem Dan Manajemen Industri*, 3(2), 67. <https://doi.org/10.30656/jsmi.v3i2.1524>
- Adetula, A.-F. A., & Akanbi, T. D. (2023). Beyond guesswork: Leveraging AI-driven predictive analytics for enhanced demand forecasting and inventory optimization in SME supply chains. *International Journal of Science and Research Archive*, 10(2), 1389–1406. <https://doi.org/10.30574/ijra.2023.10.2.0988>
- Agarwal, A., Prabha, S., & Yadav, R. (2024). *Exploratory Data Analysis (EDA) for Banking and Finance: Unveiling Insights and Patterns*.
- Akter, M., & Kudapa, S. P. (2024). A COMPARATIVE ANALYSIS OF ARTIFICIAL INTELLIGENCE-INTEGRATED BI DASHBOARDS FOR REAL-TIME DECISION SUPPORT IN OPERATIONS. *International Journal of Scientific Interdisciplinary Research*, 05(02), 158–191. <https://doi.org/10.63125/47jjv310>
- Awodiji, T. O. (2021). *Interactive Dashboard Design for Manager, Data Analyst and Data Scientist Perspective*. 175–185. <https://doi.org/10.5121/csit.2021.111914>
- Banerjee, S., Fullerton, C. E., Gaharwar, S. S., & Jaselskis, E. J. (2025). Strategic Web-Based Data Dashboards as Monitoring Tools for Promoting Organizational Innovation. *Buildings*, 15(13). <https://doi.org/10.3390/buildings15132204>
- Bansal, A. (2023). OPTIMIZING CUSTOMER SEGMENTATION FOR ENHANCED RECOMMENDATION SYSTEMS THROUGH COMPARATIVE ANALYSIS OF K-MEANS, HIERARCHICAL CLUSTERING, AND DBSCAN ALGORITHMS. *International Journal of Core Engineering & Management*, (7), 6. <https://www.researchgate.net/publication/384604526>
- Bhalla, S., Alfnes, E., Hvolby, H. H., & Sgarbossa, F. (2021). Advances in spare parts classification and forecasting for inventory control: A literature review. *IFAC-PapersOnLine*, 54(1), 982–987. <https://doi.org/10.1016/j.ifacol.2021.08.118>

- Cisse, S. I., Xue, J., & Sali, M. (2022). Research on the Comparison between the Different Policies by Service Level and Inventory Level Performance of Auto Parts in N.A.C.C. (North Automobile Components Company). *Journal of Management Science & Engineering Research*, 5(2), 1–16. <https://doi.org/10.30564/jmsr.v5i2.4593>
- Darma, F. S., & Setiadi, T. (2025). Clustering and Forecasting Implementation for Medical Consumables Stock Recommendation. In *Journal of Applied Informatics and Computing (JAIC)* (Vol. 9, Number 4). <http://jurnal.polibatam.ac.id/index.php/JAIC>
- Daud, S., Putri Dewita, K., Lediestiani, S., & I Ketut Bobiyana, P. (2024). Strategy to Improve Service Quality in Service UMKM At “Berkah Rizki Photocopy” Gunung Pasir Jaya, A Village in East Lampung, Lampung Province. *Economic Education and Entrepreneurship Journal*, 7(2), 241–250. <https://doi.org/10.23960/e3j/v7.i2.241-250>
- Poetri, N. E., Husein, I., & Dur, S. (2024). Determining Zoning of Areas Affected by Flood Disasters in Medan City Using Silhouette Coefficient and Davies Bouldin Index Analysis. *Jurnal Pijar Mipa*, 19(3), 558–563. <https://doi.org/10.29303/jpm.v19i3.6707>
- Farhana, D. (2021). Knowledge Discovery in Protein Sequence Analysis Using Hierarchical Clustering Method. *International Journal of Current Research and Review*, 13(19), 14–16. <https://doi.org/10.31782/ijcrr.2021.131907>
- Gere, A. (2023). Recommendations for validating hierarchical clustering in consumer sensory projects. *Current Research in Food Science*, 6. <https://doi.org/10.1016/j.crfs.2023.100522>
- Ghorabae, M. K., Zavadskas, E. K., Olfat, L., & Turskis, Z. (2015). Multi-Criteria Inventory Classification Using a New Method of Evaluation Based on Distance from Average Solution (EDAS). *Informatica (Netherlands)*, 26(3), 435–451. <https://doi.org/10.15388/Informatica.2015.57>
- Hanum, N., Prianto, C., Rahayu, W. I., & Kishendrian, H. D. (2023). *Penerapan Metode Clustering Dalam Segmentasi Pelanggan Perusahaan Logistik*. <https://doi.org/10.31598>
- Harahap, F., Amanda Putri, R., & Hasyifah Sibarani, F. (2024). IMPLEMENTATION OF SUPPLY CHAIN MANAGEMENT AT GRAHA PHONE SERVICE WEB-BASED. *Antivirus: Jurnal Ilmiah Teknik Informatika*, 18(2), 280–293. <https://doi.org/10.35457/antivirus.v18i2.3794>
- Hasibuan, D., Khairunnatsri, I., Oktaviani, R., & Rahman, Z. (2024). Analisis Kualitas Pelayanan Dengan Menggunakan Metode Service Quality Guna Meningkatkan Kualitas Pelayanan Pada Irham Ponsel. *Fusion: Journal of Research in Engineering*, 1.
- Hicham, N., & Karim, S. (2022). Analysis of Unsupervised Machine Learning Techniques for an Efficient Customer Segmentation using Clustering Ensemble and Spectral Clustering. In *IJACSA) International Journal of Advanced Computer Science and Applications* (Vol. 13, Number 10). www.ijacsa.thesai.org
- Hooli, M. (2025). Enterprise dashboards and decision support systems: Engineering for Actionable Intelligence. *World Journal of Advanced*

- Engineering Technology and Sciences*, 15(3), 1901–1908.
<https://doi.org/10.30574/wjaets.2025.15.3.1062>
- Imanuel, D. A., & Alfian, G. (2025). Visualisasi Segmentasi Pelanggan Berdasarkan Atribut RFM Menggunakan Algoritma K-Means Untuk Memahami Karakteristik Pelanggan pada Toko Retail Online. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 12(2), 283–292.
<https://doi.org/10.25126/jtiik.2025128619>
- Jeddou, M. Ben. (2014). Multi-Criteria ABC Inventory Classification- A Case of Vehicles Spare Parts Items. *Journal of Advanced Management Science*, 181–185. <https://doi.org/10.12720/joams.2.3.181-185>
- Kandemir, B. (2022). A Methodology for Clustering Items with Seasonal and Non-seasonal Demand Patterns for Inventory Management. *Journal of Advanced Research in Natural and Applied Sciences*, 8(4), 753–761.
<https://doi.org/10.28979/jarnas.1112146>
- Kashyap, G. (2024). Unsupervised Learning for High-Dimensional Data: Advancements in Unsupervised Learning Techniques like Clustering, Anomaly Detection, and Dimensionality Reduction. *International Journal of Leading Research Publication (IJLRP) IJLRP24121298*, 5(12). www.ijlrp.com
- Li, T., Rezaeipanah, A., & Tag El Din, E. S. M. (2022). An ensemble agglomerative hierarchical clustering algorithm based on clusters clustering technique and the novel similarity measurement. *Journal of King Saud University - Computer and Information Sciences*, 34(6), 3828–3842.
<https://doi.org/10.1016/j.jksuci.2022.04.010>
- Lolli, F., Ishizaka, A., Gamberini, R., Balugani, E., & Rimini, B. (2017). Decision Trees for Supervised Multi-criteria Inventory Classification. *Procedia Manufacturing*, 11, 1871–1881.
<https://doi.org/10.1016/j.promfg.2017.07.326>
- Maimon, O., & Rokach, L. (2009). Introduction to Knowledge Discovery and Data Mining. In *Data Mining and Knowledge Discovery Handbook* (pp. 1–15). Springer US. https://doi.org/10.1007/978-0-387-09823-4_1
- Muntaja, N., & Sriani, S. (2024). Penerapan Algoritma Logika Fuzzy Mamdani Untuk Optimalisasi Stok Dari Berbagai Jenis Spareparts Handphone. *Journal of Computer System and Informatics (JoSYC)*, 5(4), 1023–1032.
<https://doi.org/10.47065/josyc.v5i4.5836>
- Murtagh, F., & Legendre, P. (2024). *Ward's Hierarchical Clustering Method: Clustering Criterion and Agglomerative Algorithm*.
<https://doi.org/10.1007/s00357-014-9161-z>
- Muthahharah, I., & Juhari, A. (2021). *Cluster Analysis with Complete Linkage and Ward's Method for Health Service Data in Makassar City*.
- Muthoharoh, L. (2025). Text Mining Customer Feedback: An Agglomerative Clustering Approach to Service Optimization. *International Journal of Electronics and Communications Systems*, 5(1), 31–51.
<https://doi.org/10.24042/ijecs.v5i1.27188>
- Priambodo, E., & Jananto, A. (2022). *Perbandingan Analisis Cluster Algoritma K-Means Dan AHC Dalam Perencanaan Persediaan Barang Pada Perusahaan Manufaktur*.

- Raja, A. M. L., Ai, T. J., & Astanti, R. D. (2016). A Clustering Classification of Spare Parts for Improving Inventory Policies. *IOP Conference Series: Materials Science and Engineering*, 114(1). <https://doi.org/10.1088/1757-899X/114/1/012075>
- Rajasekaran, M. D., Krishna Mohan, D., Gowtham Sai, C., & Chetan, C. (2025). CUSTOMER SEGMENTATION USING AGGLOMERATIVE CLUSTERING. In *International Journal for Research Trends and Innovation* (Vol. 10). www.ijrti.org
- Rathore, S., Nawkhare, R., Sharma, N., Chaudhary, N., Chakole, S., & Vishwakrama, B. (2025). Effective Data Visualization Techniques for Business Decision-Makers. *International Research Journal on Advanced Engineering and Management (IRJAEM)*, 3(05), 2029–2036. <https://doi.org/10.47392/irjaem.2025.0318>
- Rolf, B., Beier, A., Jackson, I., Müller, M., Reggelin, T., Stuckenschmidt, H., & Lang, S. (2025). A review on unsupervised learning algorithms and applications in supply chain management. In *International Journal of Production Research* (Vol. 63, Number 5, pp. 1933–1983). Taylor and Francis Ltd. <https://doi.org/10.1080/00207543.2024.2390968>
- Scarf, P., Syntetos, A., & Teunter, R. (2024). Joint maintenance and spare-parts inventory models: a review and discussion of practical stock-keeping rules. *IMA Journal of Management Mathematics*, 35(1), 83–109. <https://doi.org/10.1093/imaman/dpad020>
- Schröer, C., Kruse, F., & Gómez, J. M. (2021). A systematic literature review on applying CRISP-DM process model. *Procedia Computer Science*, 181, 526–534. <https://doi.org/10.1016/j.procs.2021.01.199>
- Sharma, A. (2025). MANAGEMATE-An Integrated Platform for Inventory, Financial, and Business Management. *INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT*, 09(04), 1–9. <https://doi.org/10.55041/ijrem45677>
- Shutaywi, M., & Kachouie, N. N. (2021). Silhouette analysis for performance evaluation in machine learning with applications to clustering. *Entropy*, 23(6). <https://doi.org/10.3390/e23060759>
- Soudachanh, S., & Salhofer, S. (2025). Unveiling Environmental Potential in Smartphone Repair Practices in Vientiane Capital, Laos. *Sustainability (Switzerland)*, 17(2). <https://doi.org/10.3390/su17020711>
- Sukma, P. A., Sahara, A. A., & Eunike, A. (2019). *Minimizing Corrective Maintenance Cost through Spare Parts Classification and Inventory Control*.
- Sukrasorn, R., Horadee, J., Hormjun, T., & Santisookrat, S. (2025). Increasing Efficiency in Spare Parts Management:A Case Study of B.T. MINING Co., Ltd. *Engineering and Technology Horizons*, 42(3). <https://doi.org/10.55003/ETH.420302>
- Sultana, R. (2025). ARTIFICIAL INTELLIGENCE IN DATA VISUALIZATION: REVIEWING DASHBOARD DESIGN AND INTERACTIVE ANALYTICS FOR ENTERPRISE DECISION-MAKING. *International Journal of Business and Economics Insights*, 5(3), 01–29. <https://doi.org/10.63125/cp51y494>

- Tangon, J., Tuerah, R. H., Ratag, J., Maryadi, D. A., Maluw, F., Rombot, R. F., & Toweula, A. (2025). Perancangan Konsep Sistem Pencatatan Persediaan Berbasis Digital Pada UMKM di Kota Manado. *SENTRI: Jurnal Riset Ilmiah*, 4(8), 1017–1028. <https://doi.org/10.55681/sentri.v4i8.4320>
- Torre, N. M. de M., Salomon, V. A. P., & Florek-Paszowska, A. K. (2025). Multi-criteria classification of spare parts in the steel industry. *Brazilian Journal of Operations and Production Management*, 22(1). <https://doi.org/10.14488/BJOPM.2344.2025>
- Türkeli, S., Huang, B., Stasik, A., & Kemp, R. (2019). Circular economy as a glocal business activity: Mobile phone repair in the Netherlands, Poland and China. *Energies*, 12(3). <https://doi.org/10.3390/en12030498>
- Widyawati, W., Saptomo, W. L. Y., & Utami, Y. R. W. (2020). Penerapan Agglomerative Hierarchical Clustering Untuk Segmentasi Pelanggan. *Jurnal Ilmiah SINUS*, 18(1), 75. <https://doi.org/10.30646/sinus.v18i1.448>
- Yin, H., Aryani, A., Petrie, S., Nambissan, A., Astudillo, A., & Cao, S. (2024). A Rapid Review of Clustering Algorithms. <http://arxiv.org/abs/2401.07389>
- Zhang, S., Huang, K., & Yuan, Y. (2021). Spare parts inventory management: A literature review. *Sustainability (Switzerland)*, 13(5), 1–23. <https://doi.org/10.3390/su13052460>

Perencanaan Prioritas Stok Sparepart Servis Handphone Berdasarkan Pengelompokan Kasus Servis Menggunakan Agglomerative Hierarchical Clustering pada Konter Anjas Reparasi

ORIGINALITY REPORT

18%

SIMILARITY INDEX

17%

INTERNET SOURCES

9%

PUBLICATIONS

12%

STUDENT PAPERS

PRIMARY SOURCES

1	repository.umsu.ac.id Internet Source	2%
2	Submitted to Universitas Muhammadiyah Sumatera Utara Student Paper	1%
3	ejournal.scmdigi.id Internet Source	<1%
4	text-id.123dok.com Internet Source	<1%
5	kc.umn.ac.id Internet Source	<1%
6	Submitted to The Independent Institute of Education (IIE) Student Paper	<1%
7	Submitted to Universitas Tarumanagara Student Paper	<1%
8	Eunike Charina Ibrena Tarigan, Angela Wita Simanullang, Daniel S. Simbolon, Sardo Sipayung. "Clustering Penjualan Toko Retail Menggunakan Algoritma K-Means dalam Proses Penambangan Data", remik, 2026 Publication	<1%



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/BAN-PT/Ak.Ppy/PT/III/2024

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

www.umsumedan.ac.id

fas@umsumedan.ac.id

[umsumedan](https://www.facebook.com/umsumedan)

[umsumedan](https://www.instagram.com/umsumedan)

[umsumedan](https://www.linkedin.com/company/umsumedan)

[umsumedan](https://www.youtube.com/channel/UC...)

PENETAPAN DOSEN PEMBIMBING
PROPOSAL/SKRIPSI MAHASISWA
NOMOR : 979/II.3-AU/UMSU-09/F/2025

Assalamu'alaikum Warahmatullahi Wabarakatuh

Dekan Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara, berdasarkan Persetujuan permohonan judul penelitian Proposal / Skripsi dari Ketua / Sekretaris.

Program Studi : Sistem Informasi
Pada tanggal : 27 Oktober 2025

Dengan ini menetapkan Dosen Pembimbing Proposal / Skripsi Mahasiswa.

Nama : Muhammad Bagus Nurcahyo
NPM : 2209010085
Semester : VII (Tujuh)
Program studi : Sistem Informasi
Judul Proposal / Skripsi : Pemetaan Faktor yang Mempengaruhi Tingkat Layanan Servis Perangkat Elektronik Menggunakan Exploratory Data Analysis dan Agglomerative Clustering (Unsupervised Learning) pada Konter Anjas Reparasi

Dosen Pembimbing : Halim Maulana, S.Kom.,M.Kom.

Dengan demikian di izinkan menulis Proposal / Skripsi dengan ketentuan

1. Penulisan berpedoman pada buku panduan penulisan Proposal / Skripsi Fakultas Ilmu Komputer dan Teknologi Informasi UMSU
2. Pelaksanaan Sidang Skripsi harus berjarak 3 bulan setelah dikeluarkannya Surat Penetapan Dosen Pembimbing Skripsi.
3. **Proyek Proposal / Skripsi dinyatakan " BATAL "** bila tidak selesai sebelum Masa Kadalua tanggal : 27 Oktober 2026
4. Revisi judul.....

Wassalamu'alaikum Warahmatullahi Wabarakatuh.

Ditetapkan di : Medan
 Pada Tanggal : 05 Jumadil Awwal 1447 H
 27 Oktober 2025M



Dr. Alif Kowarizmi, M.Kom.
 NIDN : 0127099201

Cc. File



05/05/26, 21:11

Gmail - [syntax] Submission Acknowledgement



Muhammad Bagus Nurcahyo <muhammadbagusnurcahyo41@gmail.com>

[syntax] Submission Acknowledgement

1 message

Ibnu Rusydi <ojs@dharmawangsa.ac.id>

Tue, May 5, 2026 at 9:11 PM

To: Muhammad Bagus Nurcahyo <muhammadbagusnurcahyo41@gmail.com>

Muhammad Bagus Nurcahyo:

Thank you for submitting the manuscript, "Perencanaan Prioritas Stok Sparepart Servis Handphone Berdasarkan Pengelompokan Kasus Servis Menggunakan Agglomerative Hierarchical Clustering pada Konter Anjaks Reparasi" to Syntax : Journal of Software Engineering, Computer Science and Information Technology. With the online journal management system that we are using, you will be able to track its progress through the editorial process by logging in to the journal web site:

Manuscript URL:

<https://jurnal.dharmawangsa.ac.id/index.php/syntax/author/submission/8750>

Username: bagusnurcahyo_18

If you have any questions, please contact me. Thank you for considering this journal as a venue for your work.

Ibnu Rusydi

Syntax : Journal of Software Engineering, Computer Science and Information Technology

Terima Kasih

Syntax:

Journal of Software Engineering, Computer Science and Information Technology

<http://jurnal.dharmawangsa.ac.id/index.php/syntax>



MAJLIS 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/2019
 Pusat Administrasi: Jalan Mukhlis Basri No. 3 Medan 20238 Telp. (061) 6622400 - 66224567 Fax. (061) 6622474 - 6631003
 Email: info@umsu.ac.id | www.umsu.ac.id | @umsu | @umsu | @umsu | @umsu

Berita Acara Pembimbingan Skripsi

Nama Mahasiswa : Muhammad Bagus Program Studi : Sistem Informasi
 Nurcahyo
 NPM : 2209010085 Judul Penelitian : Perencanaan
 Nama Dosen Pembimbing : Halim Maulana, S.T., M.Kom
 Prioritas Stok Sparepart Servis Handphone
 Berdasarkan Pengelompokan Kasus Servis
 Menggunakan Agglomerative Hierarchical
 Clustering pada Konter Anjars Reparasi

Tanggal Bimbingan	Hasil Evaluasi	Paraf Dosen
4 November 2025	Diskusi terkait data yang akan digunakan dan rancangan Bab 1.	
15 November 2025	Revisi Bab 1 dan penambahan Jurnal Referensi untuk menguatkan masalah di Bab 1 dan diskusi terkait data cleaning dan EDA untuk di Bab 3.	
17 November 2025	Membahas rancangan penulisan Bab 3. Memasukkan gambar hasil EDA di Bab 3.	
24 November 2025	Diskusi Bab 3. Menambahkan diagram algoritma, Menambahkan perhitungan manual algoritma, dan menambahkan penjelasan gambar hasil EDA.	
16 Desember 2025	ACC Proposal	
4 Februari 2026	Diskusi revisi seminar proposal	
10 Februari 2026	Memetakan fokus penelitian, algoritma, dan dashboard	
23 Februari 2026	Diskusi terkait dataset, modeling dan dashboard	
25 Februari 2026	Revisi naskah penelitian di skripsi	





UMSU
Unggul | Cerdas | Terpercaya

Sila kunjungi kami di page dan buatkan nomor dan bergabung

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. 69/SK/BAN-PT/Akred/PT/II/2019
Pusat Administrasi: Jalan Mukhtar Basri No. 3 Medan 20238 Telp. (061) 6622400 - 66224567 Fax. (061) 6625474 - 6631003

<https://id.umhu.ac.id>

ftsi@umhu.ac.id

[umsumedan](#)

[umsumedan](#)

[umsumedan](#)

[umsumedan](#)

26 Februari 2026	Diskusi terkait lembar validasi instrumen skor kesulitan	
27 Februari 2026	Revisi bab 3 dan bab 4, memasukkan hasil dendogram di bab 4	
28 Februari 2026	Jelaskan insight dari EDA di bab 3 untuk tahapan Data Understanding	
2 Februari 2026	Revisi teknik encoding dari ordinal ke one hot encoding	
3 Februari 2026	Ace Sidang	

Medan, 3 Maret 2025

Diketahui oleh :
Ketua Program Studi

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

Disetujui oleh :
Dosen Pembimbing

Halim Maulana, S.T., M.Kom



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/BAN-PTIAK-Ppy/PT/2024
 Pusat Administrasi: Jalan Mukhtar Basri No. 3 Medan 20238 Telp. (061) 6622400 - 66224567 Fax. (061) 6625474 - 6631003
 Website: www.umsu.ac.id Email: fas@umsu.ac.id Instagram: @umsuamedan Facebook: umsumedan YouTube: umsumedan

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 : Muhammad Bagus Nurcahyo
 NPM : 2209010085
 Program Studi : Sistem Informasi
 Judul Proposal : Perencanaan Prioritas Stok Sparepart Servis Handphone Berdasarkan Pengelompokan Kasus Servis Menggunakan Agglomerative Hierarchical Clustering Pada Konter Anjlas Reparasi

Materi/Point yang Diperbaiki :

		Paraf
Dr. Al-Khowarizmi, M.Kom.	Masukin proses di sistem	
Rizaldy Khair, M.Kom	<ul style="list-style-type: none"> - perbaiki catatan jumlah yg benar - size image keci font preparasi - siklus 3.1 perbaiki - perbaiki yang saya tawar di paper 	
Halim Maulana, S.T.,M.Kom.		

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

SURAT KETERANGAN SELESAI PENELITIAN

Nomor: 012/KAR-SR/III/2026

Yang bertanda tangan di bawah ini:

Nama Lengkap : Azwar Anas

Jabatan : Pimpinan / Pemilik Konter Anjas Reparasi

Alamat Instansi : Jl. Simpang Matapao, Liberia, Kec. Tlk. Mengkudu, Kabupaten Serdang Bedagai, Sumatera Utara

Dengan ini menerangkan dengan sesungguhnya bahwa:

Nama Lengkap : Muhammad Bagus Nurcahyo

NPM : 2209010085

Program Studi : Sistem Informasi

Fakultas : Ilmu Komputer dan Teknologi Informasi

Perguruan Tinggi : Universitas Muhammadiyah Sumatera Utara (UMSU)

Adalah benar telah selesai melaksanakan penelitian/riset dan pengambilan data di Konter Anjas Reparasi terhitung sejak tanggal **4 November 2025** sampai dengan **3 Maret 2026**.

Penelitian tersebut dilakukan dalam rangka penyusunan tugas akhir (Skripsi) mahasiswa yang bersangkutan dengan judul:

"Perencanaan Stok Sparepart Berdasarkan Pengelompokan Kasus Servis Menggunakan Algoritma Agglomerative Hierarchical Clustering pada Konter Anjas Reparasi"

Selama proses riset dan observasi berlangsung, mahasiswa tersebut telah menunjukkan sikap yang disiplin, kooperatif, dan hasil luaran penelitiannya berupa Sistem Dashboard sangat memberikan manfaat nyata dalam membantu manajemen prioritas stok di Konter Anjas Reparasi.

Demikian surat keterangan ini dibuat dengan sebenar-benarnya agar dapat dipergunakan sebagaimana mestinya untuk melengkapi persyaratan administrasi akademik.

Medan, 5 Maret 2026

Pimpinan Konter Anjas Reparasi,



Azwar Anas