



Analisis Kepuasan Pengguna Transjakarta Koridor B13 (Summarecon Bekasi - Blok M) Menggunakan Metode K-Means Clustering

Pandu Endra Pramuja¹, M. Hafiz Al-azfa², Indra Syah Putra³, Dennis Surya Fadhilah^{4*}

^{1,2,3}Program Studi Sistem Informasi, Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika, Jl. Kaliabang Tengah No 8, Kota Bekasi, Jawa Barat, Indonesia

Email: pandudruu@gmail.com¹, hafizkun31@gmail.com², dennissurya.f@gmail.com³, indrasyahputra026@gmail.com⁴

ABSTRACT

Public transportation plays an important role in supporting urban community mobility. This study aims to analyze the level of user satisfaction with Transjakarta Corridor B13 (Summarecon Bekasi-Blok M) using the K-Means Clustering method. The research subjects consist of active Transjakarta users, including students and workers. Data were collected through a questionnaire based on a 1-5 Likert scale covering service quality and facility indicators. Data processing was conducted using the RapidMiner application to cluster user satisfaction levels. The results indicate the formation of three main clusters representing satisfied, moderately satisfied, and dissatisfied users. Timeliness and comfort were identified as the most influential factors affecting user satisfaction. This study is expected to serve as an evaluation reference for Transjakarta management in improving service quality on Corridor B13.

Keywords: *Clustering, User Satisfaction, Corridor B13, K-Means, Transjakarta.*

ABSTRAK

Transportasi publik memiliki peran penting dalam mendukung mobilitas masyarakat perkotaan. Penelitian ini bertujuan untuk menganalisis tingkat kepuasan pengguna Transjakarta pada Koridor B13 (Summarecon Bekasi - Blok M) menggunakan metode K-Means Clustering. Subjek penelitian mencakup pengguna aktif Transjakarta yang terdiri dari pelajar, mahasiswa, dan pekerja. Data diperoleh melalui penyebaran kuesioner berbasis skala Likert 1-5 yang mencakup indikator kualitas pelayanan dan fasilitas. Pengolahan data dilakukan menggunakan aplikasi RapidMiner untuk mengelompokkan tingkat kepuasan pengguna. Hasil analisis menunjukkan terbentuknya tiga cluster utama yang merepresentasikan kategori puas, cukup, dan tidak puas. Faktor ketepatan waktu dan kenyamanan menjadi aspek yang paling berpengaruh terhadap tingkat kepuasan pengguna. Penelitian ini diharapkan dapat menjadi bahan evaluasi bagi pengelola Transjakarta dalam meningkatkan kualitas layanan di Koridor B13.

Kata Kunci: *Clustering, Kepuasan Pengguna, Koridor B13, K-Means, Transjakarta.*

Article History

Received: Desember 2025

Reviewed: Desember 2025

Published: Desember 2025

Plagiarism Checker No 235

Prefix DOI :

[10.8734/Kohesi.v1i2.365](https://doi.org/10.8734/Kohesi.v1i2.365)

Copyright : Author

Publish by : Kohesi



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



1. PENDAHULUAN

Transportasi publik memegang peran penting dalam mendukung aktivitas masyarakat perkotaan, terutama di wilayah dengan tingkat mobilitas tinggi seperti Jakarta. Keberadaan sistem transportasi yang efisien dapat membantu mengurangi kemacetan, polusi udara, dan ketergantungan terhadap kendaraan pribadi (Administrator, 2019). Salah satu layanan transportasi publik yang menonjol di Jakarta adalah Transjakarta, yang dikenal dengan sistem Bus Rapid Transit (BRT). Layanan ini telah menjadi tulang punggung mobilitas warga karena jangkauannya luas dan integrasi antarkoridor yang memadai.

Transjakarta Koridor B13 (Summarecon Bekasi - Blok M) merupakan salah satu rute baru yang menghubungkan wilayah Bekasi dan Jakarta Selatan, dua kawasan dengan tingkat aktivitas ekonomi tinggi. Pengoperasian koridor ini diharapkan mampu mempercepat waktu tempuh pengguna serta meningkatkan kenyamanan perjalanan. Namun, peningkatan jumlah penumpang tidak selalu diiringi dengan peningkatan kualitas layanan. Berdasarkan data dari Databoks Katadata (Anuur, 2022), Transjakarta mencatat lebih dari 250 juta perjalanan pada tahun 2021, menunjukkan kebutuhan untuk terus mengevaluasi kualitas pelayanan agar tetap sesuai dengan ekspektasi pengguna.

Penelitian mengenai kepuasan pengguna sangat penting dilakukan karena tingkat kepuasan menjadi indikator utama keberhasilan suatu layanan publik. Hasil penelitian ini dapat menjadi acuan bagi pengelola Transjakarta untuk meningkatkan mutu pelayanan dan pengalaman pengguna. Oleh karena itu, penelitian ini bertujuan untuk menganalisis tingkat kepuasan pengguna Transjakarta pada Koridor B13 dengan menggunakan metode K-Means Clustering. Melalui pendekatan ini, diharapkan dapat diperoleh gambaran objektif mengenai persepsi pengguna terhadap pelayanan dan fasilitas yang tersedia.

Fokus penelitian ini diarahkan pada analisis kepuasan pengguna yang mencakup tiga faktor utama, yaitu kualitas pelayanan, fasilitas, dan kenyamanan. Ketiga faktor ini merupakan indikator penting dalam menilai kinerja layanan publik. Menurut (Vanessa & Prabantari, 2020), persepsi pengguna terhadap kualitas pelayanan sangat dipengaruhi oleh kemampuan penyedia jasa dalam memenuhi ekspektasi pelanggan. Di sisi lain, fasilitas dan kenyamanan turut menentukan pengalaman pengguna secara keseluruhan, terutama bagi pengguna yang melakukan perjalanan rutin.

Beberapa penelitian sebelumnya juga telah mengkaji tingkat kepuasan pengguna transportasi umum menggunakan metode K-Means Clustering. Penelitian oleh (Al Atros et al., 2022), menunjukkan bahwa metode ini efektif dalam mengelompokkan tingkat kepuasan

pengguna berdasarkan pola respon kuesioner. Penelitian lain oleh (Ulya et al., 2023), juga membuktikan bahwa K-Means dapat digunakan untuk menemukan pola tersembunyi dalam data survei yang bersifat numerik, sehingga hasil analisis lebih akurat dan mudah diinterpretasikan.

Berbeda dengan penelitian terdahulu yang berfokus pada Koridor 10 Transjakarta (Akbar, n.d.) maupun sektor lain seperti perpustakaan digital atau transportasi umum secara umum, penelitian ini memiliki kebaruan pada objek penelitian, yaitu Koridor B13 yang belum pernah dianalisis sebelumnya. Selain itu, penelitian ini juga menekankan pada kombinasi variabel pelayanan, fasilitas, dan kenyamanan yang mencerminkan kondisi faktual pengguna Transjakarta rute Bekasi-Blok M. Dengan demikian, hasil dari penelitian ini diharapkan dapat memberikan kontribusi nyata bagi pengembangan kualitas layanan Transjakarta ke depannya.

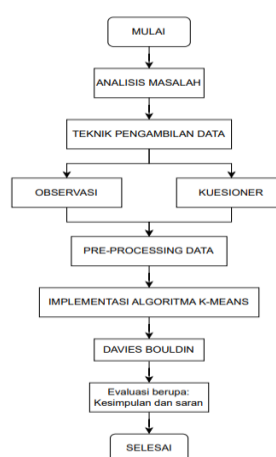
2. METODE PENELITIAN

2.1 Pendekatan Penelitian

Penelitian ini menggunakan data primer yang diperoleh langsung dari responden. Pengumpulan data dilakukan melalui kuesioner yang dibagikan kepada pengguna Transjakarta Koridor B13 serta observasi lapangan pada halte dan armada bus. Kuesioner digunakan untuk mengetahui persepsi dan pengalaman pengguna terkait kualitas layanan, sedangkan observasi dilakukan untuk melihat kondisi aktual di lapangan. Data primer yang terkumpul kemudian diolah dan dianalisis sebagai dasar dalam penelitian ini.

2.2 Tahapan Penelitian

Tahapan penelitian ini disusun untuk memberikan gambaran yang jelas mengenai proses analisis tingkat kepuasan pengguna Transjakarta Koridor B13 (Summarecon Bekasi - Blok M) terhadap kualitas pelayanan, fasilitas, kenyamanan, dan ketepatan waktu. Tahapan ini juga menjadi dasar dalam penyusunan alur kerja penelitian sehingga setiap proses dapat berjalan secara sistematis dan terarah.



Gambar2. 1 Tahap Penelitian

**1. Analisis Masalah**

Pada tahap ini peneliti mengidentifikasi permasalahan terkait tingkat kepuasan pengguna Transjakarta Koridor B13 serta menentukan elemen-elemen yang diperlukan dalam proses clustering. Analisis dilakukan untuk memahami aspek apa saja yang memengaruhi kepuasan, seperti kualitas pelayanan, fasilitas, kenyamanan, dan ketepatan waktu.

2. Teknik Pengumpulan Data

Pengumpulan data dilakukan melalui observasi langsung di halte dan armada bus serta penyebaran kuesioner kepada pengguna Koridor B13. Observasi digunakan untuk melihat kondisi nyata layanan, sedangkan kuesioner dirancang berdasarkan indikator yang telah disesuaikan melalui pengamatan sebelumnya.

3. Pre-processing Data

Pada tahap ini data yang diperoleh diperiksa dan dibersihkan melalui proses cleaning serta transformasi agar siap diproses. Tahapan ini mencakup pengecekan data tidak lengkap, data tidak valid, data duplikat, serta penyesuaian format agar sesuai dengan kebutuhan algoritma. Proses ini memastikan dataset berada dalam kondisi stabil sebelum dilakukan pengelompokan.

4. Implementasi Algoritma K-Means

Data yang telah diproses kemudian dianalisis menggunakan algoritma K-Means. Pengelompokan dilakukan berdasarkan pola kemiripan antar responden untuk mengetahui kecenderungan tingkat kepuasan. Proses implementasi dilakukan menggunakan aplikasi RapidMiner sebagai alat bantu.

5. Davies-Bouldin Index

Davies-Bouldin Index digunakan untuk mengevaluasi kualitas cluster yang terbentuk. Nilai yang dihasilkan menunjukkan tingkat pemisahan antar cluster, sehingga semakin kecil nilai indeksnya semakin baik hasil pengelompokan yang diperoleh.

6. Evaluasi

Tahap evaluasi dilakukan untuk membandingkan hasil pengolahan data. Proses ini mencakup pengujian data sebelum dan sesudah dilakukan transformasi untuk melihat perbedaan nilai evaluasi serta memastikan model yang diperoleh memberikan hasil yang konsisten.

7. Hasil

Tahap ini menampilkan hasil perhitungan dan pengujian yang telah dilakukan, mulai dari pembentukan cluster, nilai evaluasi, hingga interpretasi kategori kepuasan pengguna. Hasil tersebut menggambarkan seberapa efisien proses analisis dalam melihat pola kepuasan pengguna Transjakarta Koridor B13 berdasarkan variabel yang diteliti.

2.3 Variabel dan Indikator Penelitian

Variabel penelitian digunakan sebagai dasar dalam menyusun pernyataan kuesioner yang diberikan kepada responden. Pada penelitian ini terdapat dua variabel utama, yaitu Kualitas Pelayanan dan Fasilitas. Setiap variabel dijabarkan menjadi beberapa indikator yang digunakan

untuk mengukur tingkat kepuasan pengguna Transjakarta Koridor B13. Adapun indikator pada setiap variabel disajikan dalam tabel berikut:

Tabel 2. 1Pernyataan Kuesioner

Variabel	Pertanyaan	index
Kualitas Pelayanan	Seberapa puas Anda dengan kenyamanan armada Transjakarta?	P1
	Seberapa puas Anda dengan ketepatan waktu kedatangan dan keberangkatan bus?	P2
	Seberapa puas Anda dengan ketersediaan informasi rute dan perjalanan?	P3
	Seberapa jelas informasi yang diberikan oleh petugas pramusapa?	P4
	Seberapa sering terjadi kepadatan di dalam halte?	P5
	Bagaimana penilaian Anda terhadap kebersihan halte?	P6
	Seberapa tepat waktu bus saat jam sibuk?	P7
	Seberapa puas Anda melakukan perjalanan menggunakan Transjakarta pada jam sibuk?	P8
Fasilitas	Seberapa memadai jumlah armada yang disediakan pada rute koridor B13?	P1
	Bagaimana kenyamanan sirkulasi udara pada halte?	P2
	Bagaimana penilaian Anda terhadap fasilitas AC di bus?	P3
	Seberapa lengkap fasilitas keamanan seperti CCTV, palu darurat, dan petugas pramusapa?	P4
	Seberapa puas Anda dengan ketersediaan tempat duduk di dalam bus?	P5
	Seberapa puas Anda dengan fasilitas tempat duduk di halte?	P6
	Seberapa mudah Anda naik bus pada jam sibuk?	P7
	Apakah tarif Transjakarta sesuai dengan fasilitas yang Anda dapatkan?	P8

2.4 Pengumpulan Data

Teknik pengumpulan data merupakan metode yang digunakan peneliti untuk memperoleh informasi yang relevan dengan tujuan penelitian. Pada penelitian ini, proses pengambilan data dilakukan menggunakan dua metode utama, yaitu observasi dan kuesioner.

1. Observasi Observasi dilakukan dengan cara peneliti hadir secara langsung sebagai pengguna Transjakarta Koridor B13 (Summarecon Bekasi - Blok M). Melalui keterlibatan ini, peneliti dapat melihat dan merasakan sendiri kondisi pelayanan yang diberikan, mulai dari fasilitas halte, kebersihan armada, kenyamanan perjalanan, hingga ketepatan waktu bus. Pengamatan langsung ini memberikan gambaran yang lebih nyata terkait pengalaman penumpang serta membantu peneliti mengidentifikasi aspek-aspek yang perlu dianalisis lebih lanjut.
2. Kuesioner Kuesioner digunakan sebagai sumber informasi utama dalam penelitian ini. Instrumen ini disebarkan kepada pengguna atau individu yang pernah menggunakan layanan Transjakarta Koridor B13. Pertanyaan dalam kuesioner dirancang menggunakan skala Likert (1-5) untuk mengukur tingkat kepuasan responden terhadap berbagai aspek layanan, seperti pelayanan petugas, fasilitas, kenyamanan, dan ketepatan waktu. Hasil dari kuesioner ini menjadi dasar analisis untuk mengelompokkan pengguna berdasarkan

tingkat kepuasan dan sebagai bahan utama dalam proses pengolahan data menggunakan algoritma K-Means.

2.5 Metode Pengolahan Data

Pengolahan data dilakukan untuk memastikan bahwa data hasil kuesioner siap digunakan dalam proses analisis. Tahapan pengolahan data meliputi:

1. Data Cleaning Proses ini dilakukan untuk mengidentifikasi dan menghapus data yang tidak lengkap, tidak konsisten, atau tidak relevan dari dataset. Data yang kosong, duplikat, ataupun jawaban yang tidak logis akan dieliminasi. Pembersihan data sangat penting agar hasil analisis yang diperoleh lebih akurat dan tidak terpengaruh oleh kesalahan input.
2. Data Integration Tahap ini merupakan proses penggabungan seluruh data dari kuesioner dan observasi menjadi satu dataset yang utuh. Meskipun penelitian menggunakan data primer, integrasi tetap diperlukan agar seluruh variabel dan indikator berada dalam satu struktur data yang koheren. Integrasi yang baik membantu mempermudah proses analisis di tahap selanjutnya.
3. Data Transformation Pada tahap ini data diubah ke dalam format yang sesuai dengan kebutuhan analisis, seperti mengonversi skala Likert menjadi nilai numerik dan menata ulang struktur tabel. Transformasi dilakukan agar data siap diproses oleh sistem seperti RapidMiner, serta memastikan setiap atribut dapat terbaca dengan benar.

2.6 Metode Analisis Data

Setelah proses pengolahan data selesai, langkah berikutnya adalah melakukan analisis untuk menentukan visualisasi data serta jumlah cluster yang sesuai dengan dataset. Analisis ini dilakukan dengan beberapa tahap untuk memahami pola kepuasan pengguna Transjakarta Koridor B13 berdasarkan data kuesioner yang telah diterima.

Langkah pertama: Data dianalisis menggunakan beberapa jenis diagram seperti diagram donat, diagram batang horizontal, dan diagram batang vertikal. Visualisasi ini dilakukan untuk melihat distribusi nilai pada setiap indikator, sehingga peneliti dapat memahami persentase kategori kepuasan responden secara lebih jelas.

Langkah kedua: Tahap berikutnya adalah pembuatan heatmap. Warna pada heatmap berfungsi untuk menunjukkan intensitas nilai pada setiap variabel, sehingga memudahkan identifikasi pola serta perbedaan antarresponden. Heatmap juga membantu peneliti mengenali variabel yang memiliki kecenderungan nilai tinggi ataupun rendah sebelum masuk ke proses clustering.



Setelah itu, proses analisis dilanjutkan dengan penerapan algoritma K-Means, yang memiliki empat proses utama:

1. Menetapkan titik pusat awal (centroid) secara acak untuk setiap cluster.
2. Mengelompokkan setiap data ke cluster dengan centroid terdekat berdasarkan jarak Euclidean.
3. Menghitung ulang posisi centroid berdasarkan rata-rata nilai dari seluruh anggota dalam cluster.
4. Mengulangi langkah pengelompokan dan perhitungan centroid hingga posisi centroid tidak mengalami perubahan.

Analisis yang dilakukan menggunakan RapidMiner akan menghasilkan dua pendekatan utama:

1. Pengujian akurasi menggunakan DaviesBouldin Index (DBI). Metode ini digunakan untuk mengevaluasi sejauh mana cluster yang terbentuk dapat dipisahkan dengan baik. Nilai DBI yang lebih kecil menunjukkan kualitas pengelompokan yang lebih baik dan lebih jelas antar cluster.
2. Visualisasi Cluster dalam bentuk diagram batang. Visualisasi ini digunakan untuk menunjukkan jumlah data dalam setiap cluster, sehingga peneliti dapat memahami distribusi tingkat kepuasan responden berdasarkan hasil pengelompokan. Setelah seluruh proses analisis selesai dan hasilnya diperoleh, data kemudian diekspor ke dalam format Excel menggunakan operator Write Excel pada RapidMiner. Tahap ini memudahkan penyimpanan hasil analisis serta pengelolaan data lebih lanjut dalam bentuk yang lebih terstruktur dan mudah diakses.

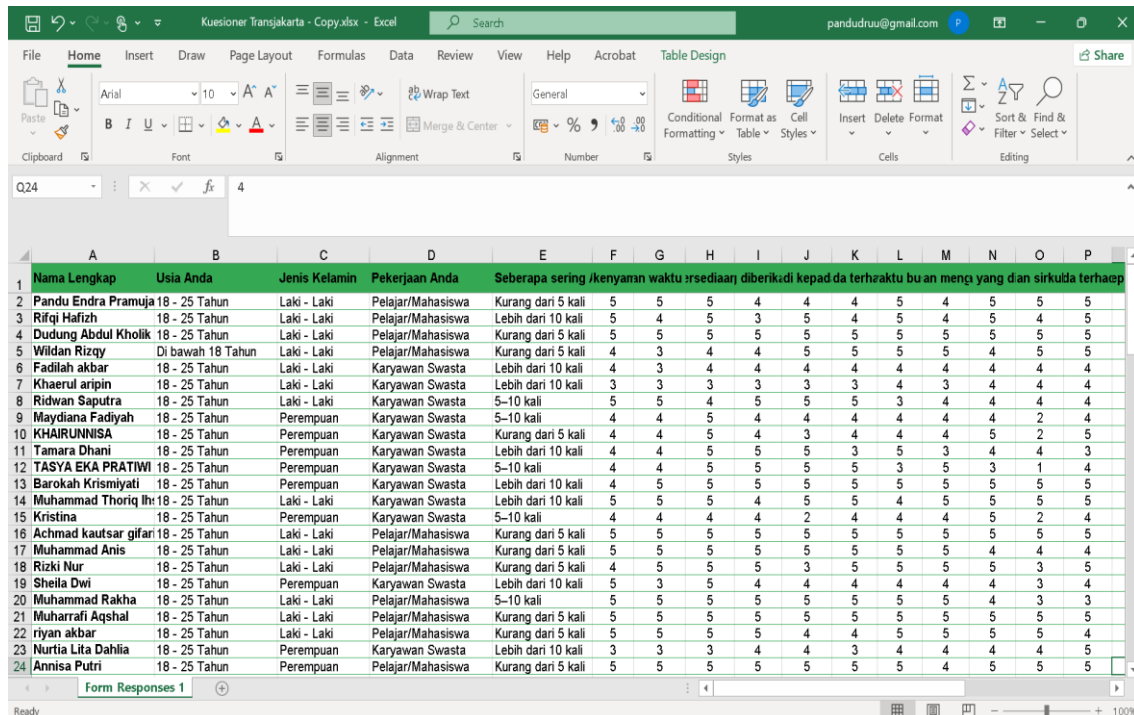
3. HASIL DAN PEMBAHASAN

Data yang digunakan dalam penelitian ini berupa data numerik sehingga termasuk dalam kategori data kuantitatif. Pengolahan data dilakukan dengan bantuan perangkat lunak RapidMiner untuk menganalisis tingkat kepuasan pengguna layanan Transjakarta Koridor B13 terhadap kualitas pelayanan dan fasilitas yang tersedia. Hasil dari pengolahan data ini bertujuan untuk mengelompokkan responden berdasarkan tingkat kepuasannya.

3.1 Spesifikasi

Sumber data utama dalam penelitian ini diperoleh dari hasil penyebaran kuesioner kepada pengguna yang pernah menggunakan layanan Transjakarta Koridor B13. Kuesioner disebarakan secara daring dan data yang terkumpul kemudian direkap secara otomatis dalam bentuk tabel

Microsoft Excel, sehingga memudahkan peneliti dalam melakukan pengolahan data.



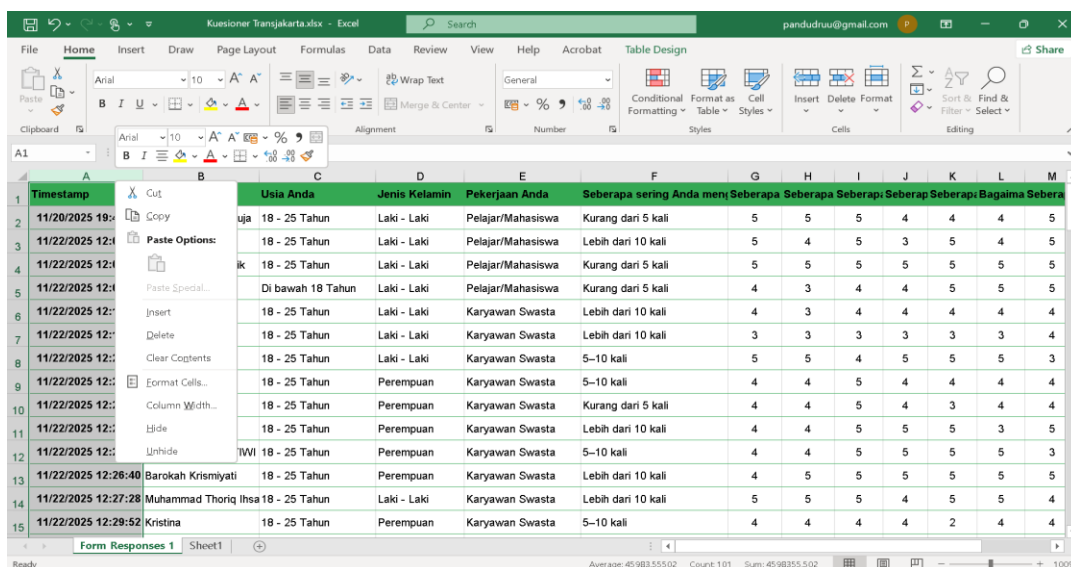
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Nama Lengkap	Usia Anda	Jenis Kelamin	Pekerjaan Anda	Seberapa sering jkenyaman waktu rsediaan diberikan kepada terhaaktu bu an mengi yang dan sirkutla terhasp											
2	Pandu Endra Pramuja	18 - 25 Tahun	Laki - Laki	Pelajar/Mahasiswa	Kurang dari 5 kali	5	5	5	4	4	4	5	4	5	5	5
3	Rifqi Hafizh	18 - 25 Tahun	Laki - Laki	Pelajar/Mahasiswa	Lebih dari 10 kali	5	4	5	3	5	4	5	4	5	4	5
4	Dudung Abdul Kholik	18 - 25 Tahun	Laki - Laki	Pelajar/Mahasiswa	Kurang dari 5 kali	5	5	5	5	5	5	5	5	5	5	5
5	Wildan Rizqy	Di bawah 18 Tahun	Laki - Laki	Pelajar/Mahasiswa	Kurang dari 5 kali	4	3	4	4	5	5	5	4	5	5	5
6	Fadilah akbar	18 - 25 Tahun	Laki - Laki	Karyawan Swasta	Lebih dari 10 kali	4	3	4	4	4	4	4	4	4	4	4
7	Khaerul aripin	18 - 25 Tahun	Laki - Laki	Karyawan Swasta	Lebih dari 10 kali	3	3	3	3	3	3	4	3	4	4	4
8	Ridwan Saputra	18 - 25 Tahun	Laki - Laki	Karyawan Swasta	5-10 kali	5	5	4	5	5	5	3	4	4	4	4
9	Maydiana Fadiah	18 - 25 Tahun	Perempuan	Karyawan Swasta	5-10 kali	4	4	5	4	4	4	4	4	4	2	4
10	KHAIRUNNISA	18 - 25 Tahun	Perempuan	Karyawan Swasta	Kurang dari 5 kali	4	4	5	4	3	4	4	4	5	2	5
11	Tamara Dhani	18 - 25 Tahun	Perempuan	Karyawan Swasta	Lebih dari 10 kali	4	4	5	5	5	3	5	3	4	4	3
12	TASYA EKA PRATIWI	18 - 25 Tahun	Perempuan	Karyawan Swasta	5-10 kali	4	4	5	5	5	5	3	5	3	1	4
13	Barokah Krismiyati	18 - 25 Tahun	Perempuan	Karyawan Swasta	Lebih dari 10 kali	4	5	5	5	5	5	5	5	5	5	5
14	Muhammad Thoriq Ihsa	18 - 25 Tahun	Laki - Laki	Karyawan Swasta	Lebih dari 10 kali	5	5	5	4	5	5	4	5	5	5	5
15	Kristina	18 - 25 Tahun	Perempuan	Karyawan Swasta	5-10 kali	4	4	4	4	2	4	4	4	5	2	4
16	Achmad kautsar gifar	18 - 25 Tahun	Laki - Laki	Pelajar/Mahasiswa	Kurang dari 5 kali	5	5	5	5	5	5	5	5	5	5	5
17	Muhammad Anis	18 - 25 Tahun	Laki - Laki	Pelajar/Mahasiswa	Kurang dari 5 kali	5	5	5	5	5	5	5	5	4	4	4
18	Rizki Nur	18 - 25 Tahun	Laki - Laki	Pelajar/Mahasiswa	Kurang dari 5 kali	4	5	5	5	3	5	5	5	5	3	5
19	Sheila Dwi	18 - 25 Tahun	Perempuan	Karyawan Swasta	Lebih dari 10 kali	5	3	5	4	4	4	4	4	4	3	4
20	Muhammad Rakha	18 - 25 Tahun	Laki - Laki	Pelajar/Mahasiswa	5-10 kali	5	5	5	5	5	5	5	5	4	3	3
21	Muharafi Aqshal	18 - 25 Tahun	Laki - Laki	Pelajar/Mahasiswa	Kurang dari 5 kali	5	5	5	5	5	5	5	5	5	5	5
22	riyan akbar	18 - 25 Tahun	Laki - Laki	Pelajar/Mahasiswa	Kurang dari 5 kali	5	5	5	5	4	4	5	5	5	5	4
23	Nurtia Lita Dahlia	18 - 25 Tahun	Perempuan	Karyawan Swasta	Lebih dari 10 kali	3	3	3	4	4	3	4	4	4	4	5
24	Annisa Putri	18 - 25 Tahun	Perempuan	Pelajar/Mahasiswa	Kurang dari 5 kali	5	5	5	5	5	5	5	4	5	5	5

Gambar3. 1 Data Kuesioner

Data hasil kuesioner selanjutnya dipersiapkan untuk melalui beberapa tahapan pengolahan, yaitu seleksi data, pembersihan data, dan transformasi data agar data siap digunakan pada proses analisis menggunakan RapidMiner.

3.2 Penyeleksian Data

Tahap seleksi data dilakukan untuk menentukan atribut yang relevan dan akan digunakan dalam proses pengolahan data. Pada tahap ini, atribut yang tidak memiliki pengaruh terhadap proses clustering dihapus dari dataset.



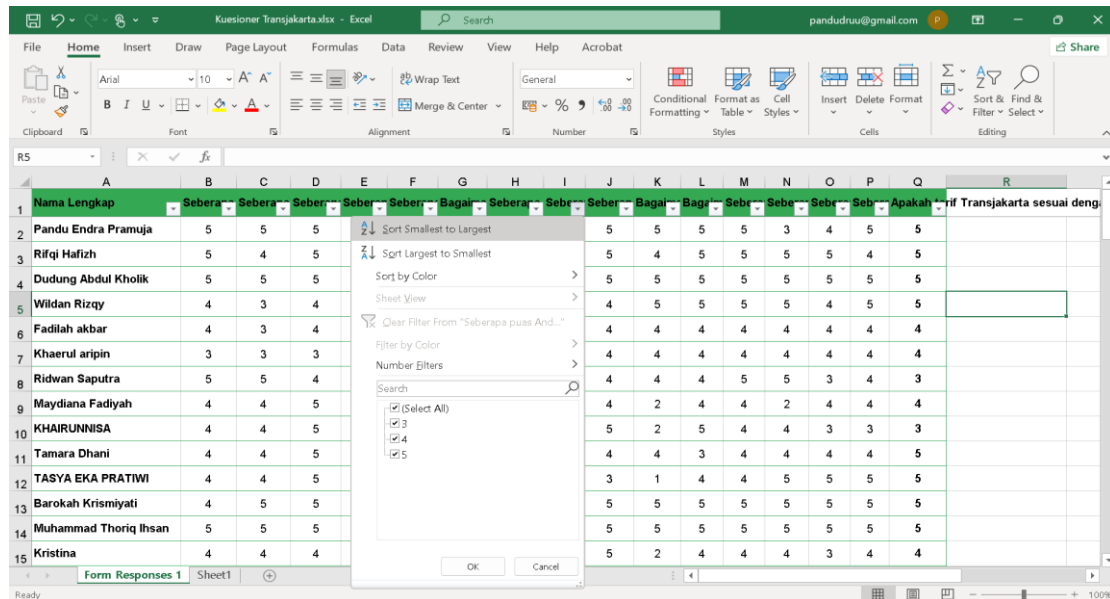
	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Timestamp	Usia Anda	Jenis Kelamin	Pekerjaan Anda	Seberapa sering Anda meny	Seberapa	Seberapa	Seberapa	Seberapa	Seberapa	Seberapa	Seberapa	Seberapa
2	11/20/2025 19:	18 - 25 Tahun	Laki - Laki	Pelajar/Mahasiswa	Kurang dari 5 kali	5	5	5	4	4	4	4	5
3	11/22/2025 12:	18 - 25 Tahun	Laki - Laki	Pelajar/Mahasiswa	Lebih dari 10 kali	5	4	5	3	5	4	5	5
4	11/22/2025 12:	18 - 25 Tahun	Laki - Laki	Pelajar/Mahasiswa	Kurang dari 5 kali	5	5	5	5	5	5	5	5
5	11/22/2025 12:	Di bawah 18 Tahun	Laki - Laki	Pelajar/Mahasiswa	Kurang dari 5 kali	4	3	4	4	5	5	5	5
6	11/22/2025 12:	18 - 25 Tahun	Laki - Laki	Karyawan Swasta	Lebih dari 10 kali	4	3	4	4	4	4	4	4
7	11/22/2025 12:	18 - 25 Tahun	Laki - Laki	Karyawan Swasta	Lebih dari 10 kali	3	3	3	3	3	3	3	4
8	11/22/2025 12:	18 - 25 Tahun	Laki - Laki	Karyawan Swasta	5-10 kali	5	5	4	5	5	5	5	3
9	11/22/2025 12:	18 - 25 Tahun	Perempuan	Karyawan Swasta	5-10 kali	4	4	5	4	4	4	4	4
10	11/22/2025 12:	18 - 25 Tahun	Perempuan	Karyawan Swasta	Kurang dari 5 kali	4	4	5	4	3	4	4	4
11	11/22/2025 12:	18 - 25 Tahun	Perempuan	Karyawan Swasta	Lebih dari 10 kali	4	4	5	5	5	3	5	5
12	11/22/2025 12:	18 - 25 Tahun	Perempuan	Karyawan Swasta	5-10 kali	4	4	5	5	5	5	5	3
13	11/22/2025 12:26:40	Barokah Krismiyati	18 - 25 Tahun	Perempuan	Karyawan Swasta	Lebih dari 10 kali	4	5	5	5	5	5	5
14	11/22/2025 12:27:28	Muhammad Thoriq Ihsa	18 - 25 Tahun	Laki - Laki	Karyawan Swasta	Lebih dari 10 kali	5	5	5	4	5	5	4
15	11/22/2025 12:29:52	Kristina	18 - 25 Tahun	Perempuan	Karyawan Swasta	5-10 kali	4	4	4	4	2	4	4

Gambar3. 2 Penyeleksian Data

Dalam penelitian ini, hanya atribut yang berasal dari pernyataan kuesioner yang berkaitan dengan variabel kualitas pelayanan dan fasilitas yang digunakan dalam proses analisis. Atribut identitas seperti alamat email tidak disertakan karena tidak berpengaruh terhadap hasil pengelompokan.

3.3 Pre-Processing Data

Setelah dilakukan seleksi atribut, data kemudian melalui tahap pre-processing atau pembersihan data. Pada tahap ini dilakukan pemeriksaan terhadap data yang tidak lengkap, data kosong (*missing value*), serta data yang tidak konsisten.



Gambar3. 3 Cleaning Data

3.4 Uji PCA (Principal Components)

Pada tahap transformasi data, data hasil kuesioner diubah ke dalam bentuk yang sesuai agar dapat digunakan dalam proses analisis. Data yang semula berupa data kualitatif dari skala Likert dikonversi menjadi data kuantitatif dengan memfokuskan pada dua variabel utama, yaitu kualitas pelayanan dan fasilitas. Proses transformasi ini dilakukan menggunakan bantuan Microsoft Excel dengan rumus `=COUNTIF(range_data, kriteria_nilai_skala_Likert)` untuk menghitung jumlah kemunculan setiap kategori jawaban responden.

Hasil perhitungan tersebut kemudian dikelompokkan ke dalam kategori sangat puas, puas, netral, tidak puas, dan sangat tidak puas. Selanjutnya dilakukan proses penggabungan kategori, dimana kategori sangat puas digabungkan ke dalam kategori puas, sedangkan kategori sangat tidak puas digabungkan ke dalam kategori tidak puas, sementara kategori netral tetap dipertahankan. Penggabungan ini bertujuan untuk menyederhanakan data sehingga hanya digunakan tiga kategori utama, yaitu puas, netral, dan tidak puas.

Data hasil penggabungan tersebut ditampilkan dalam Gambar 3.4 yang selanjutnya digunakan sebagai input utama dalam proses pengelompokan menggunakan algoritma K-Means pada tahap pengolahan data berikutnya.

No.	Nama Responden	KP PUAS	KP NETRAL	KP TIDAK PUAS	F PUAS	F NETRAL	F TIDAK PUAS
1	Pandu Endra Pramuja	8	0	0	7	1	0
2	Rifqi Hafizh	7	1	0	8	0	0
3	Dudung Abdul Kholik	8	0	0	8	0	0
4	Wildan Rizqy	7	1	0	8	0	0
5	Fadilah akbar	7	1	0	8	0	0
6	Khaerul aripin	1	7	0	8	0	0
7	Ridwan Saputra	7	1	0	6	2	0
8	Maydiana Fadiyah	8	0	0	6	0	2
9	KHAIRUNNISA	7	1	0	4	3	1
10	Tamara Dhani	6	2	0	7	1	0
11	TASYA EKA PRATIWI	7	1	0	6	1	1
12	Barokah Krismiyati	8	0	0	8	0	0
13	Muhammad Thoriq Ihsan	8	0	0	8	0	0
14	Kristina	7	0	1	6	1	1
15	Achmad kautsar gifari	8	0	0	8	0	0
16	Muhammad Anis	8	0	0	7	1	0
17	Rizki Nur	7	1	0	7	1	0
18	Sheila Dwi	7	1	0	7	1	0
19	Muhammad Rakha	8	0	0	4	4	0
20	Muharrafi Aqshal	8	0	0	8	0	0
21	riyan akbar	8	0	0	7	1	0
22	Nurtia Lita Dahlia	4	4	0	8	0	0
23	Annisa Putri	8	0	0	8	0	0
24	Diva	8	0	0	6	1	0

Gambar3. 4 Hasil Penggabungan Puas, Netral dan Tidak Puas

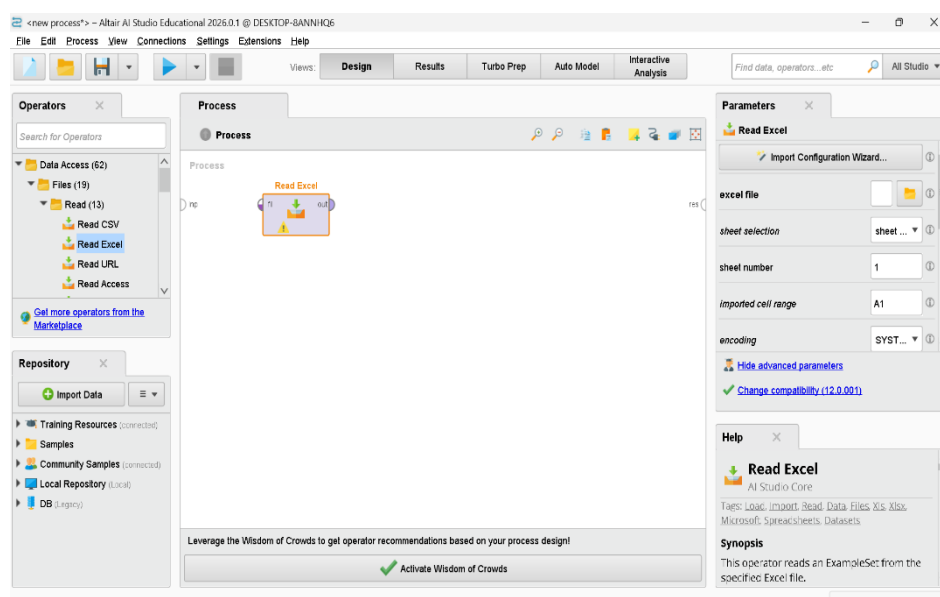
3.5 Pengolahan Data

Data yang digunakan pada tahap ini merupakan data hasil pembersihan dan seleksi. Proses pengolahan data dilakukan menggunakan metode K-Means dengan bantuan aplikasi RapidMiner.

A. Pengelompokan Data Menggunakan RapidMiner

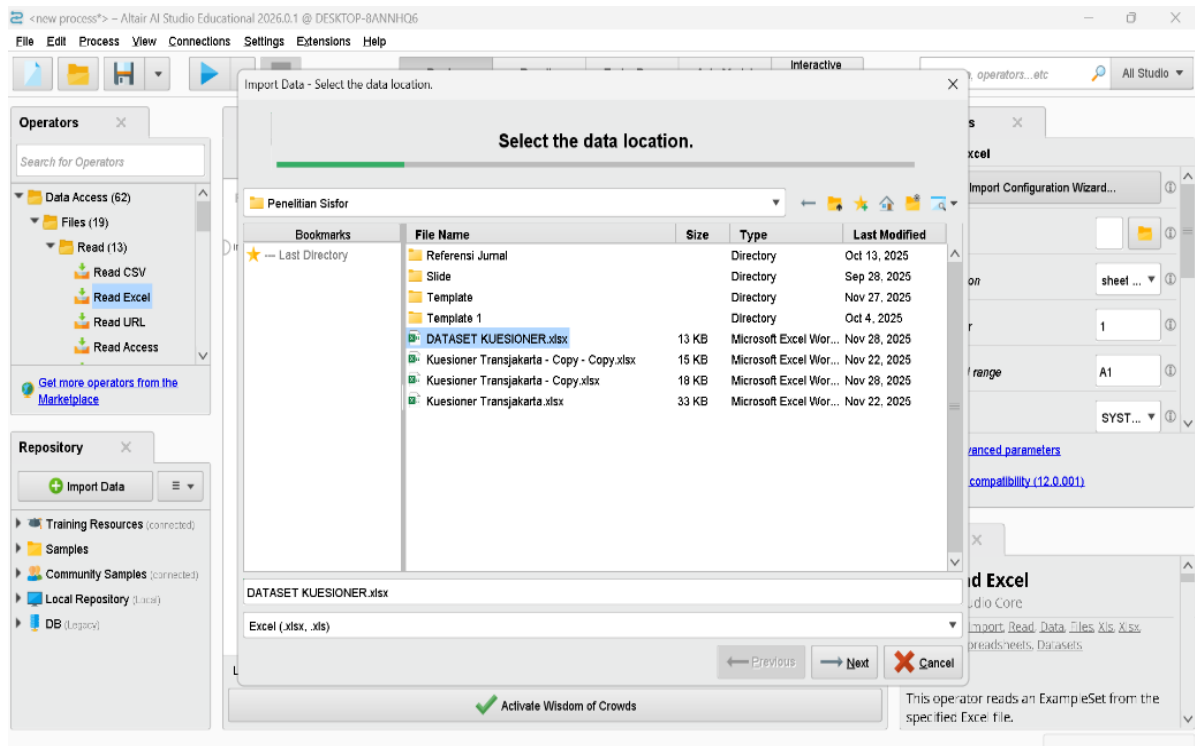
Data yang telah siap diolah dianalisis menggunakan aplikasi RapidMiner. Atribut yang digunakan terdiri dari tujuh atribut, yaitu nama responden sebagai ID, serta atribut puas, netral, dan tidak puas pada variabel kualitas pelayanan dan fasilitas.

1. Langkah awal dilakukan dengan menambahkan operator Read Excel ke dalam canvas RapidMiner untuk mengimpor file dataset yang akan digunakan.



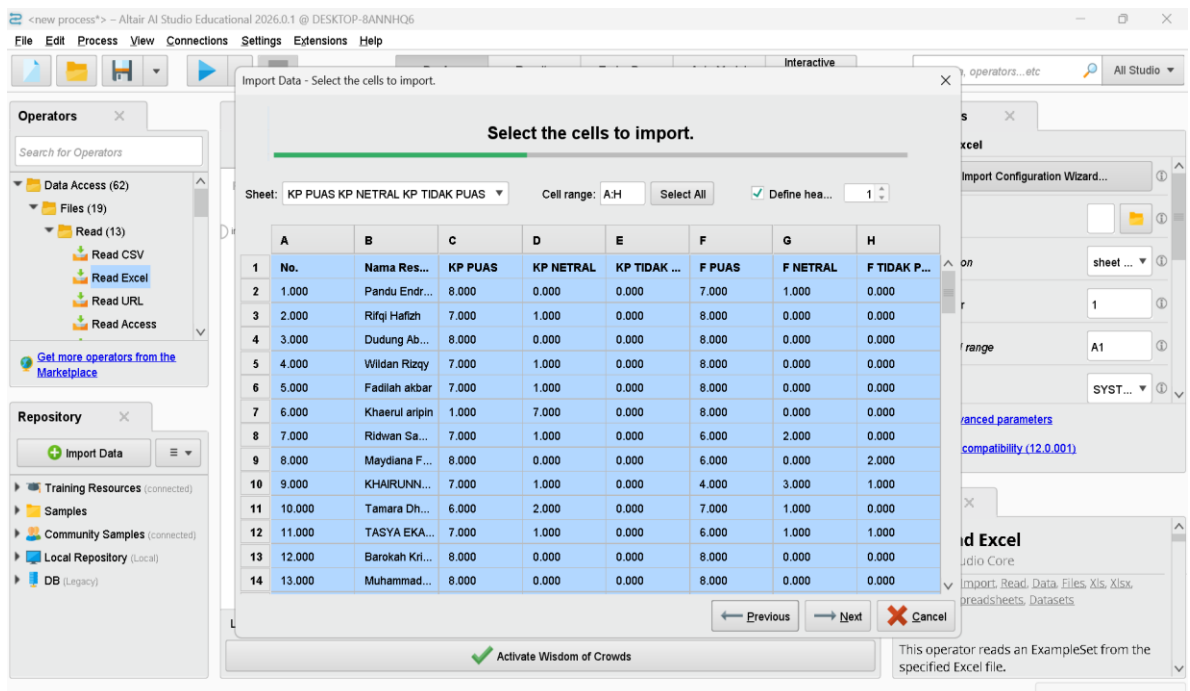
Gambar3. 5 Penambahan Oprator Read Excel

2. Pada tahap berikutnya dilakukan pengaturan file melalui menu Import Configuration Wizard untuk memilih dataset dan atribut yang akan digunakan.



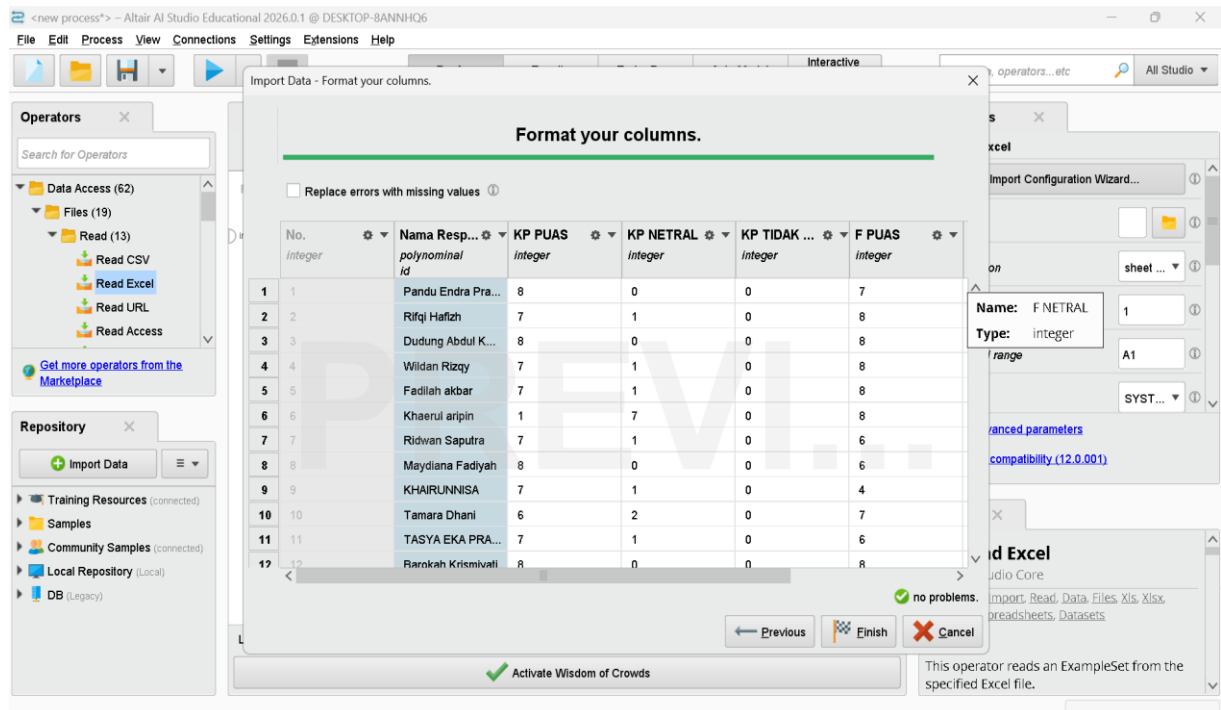
Gambar3. 6 Proses Input Dataset

3. Selanjutnya dilakukan pemilihan sheet dan sel data yang akan digunakan agar sesuai dengan proses clustering.



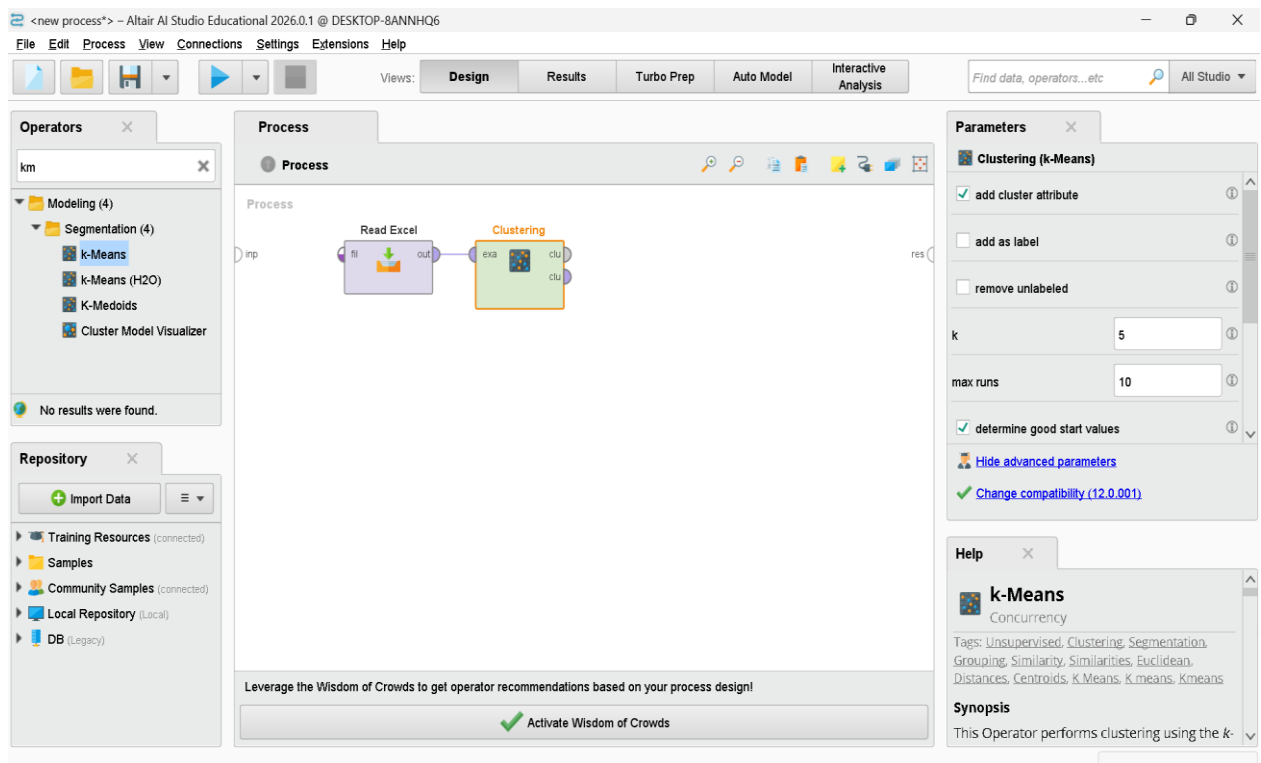
Gambar3. 7 Select Cells

4. Setelah itu dilakukan pelabelan data, dimana atribut nomor di-exclude, atribut nama sebagai ID, dan atribut lainnya ditetapkan sebagai data bertipe numerik.



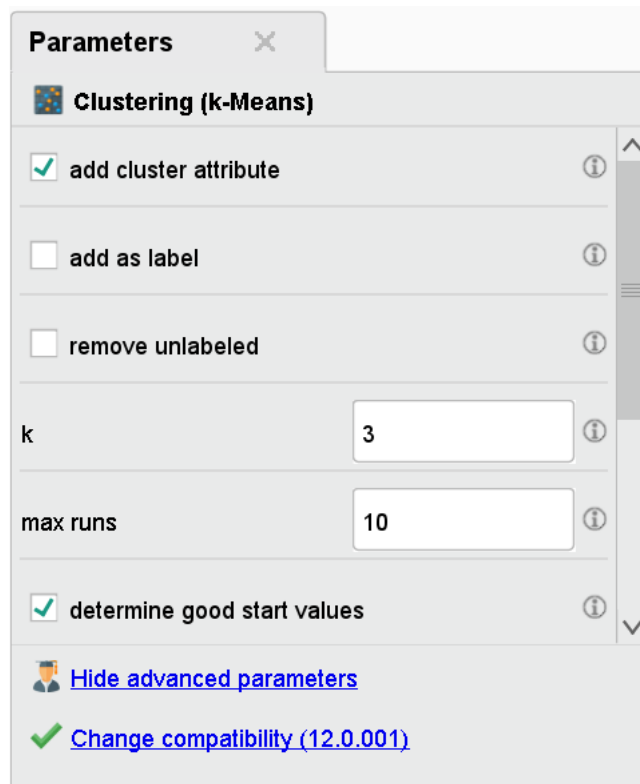
Gambar3. 8 Labeling dan Exclude Columns

5. Operator K-Means kemudian ditambahkan dan dihubungkan dengan operator Read Excel.



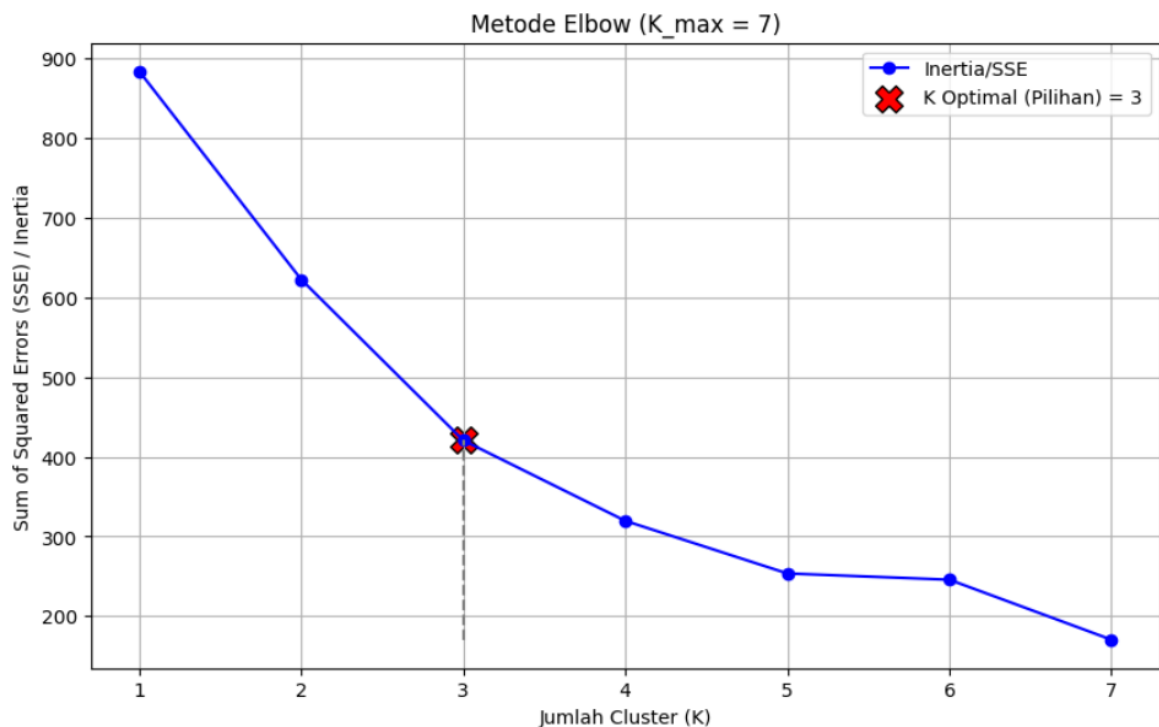
Gambar3. 9Menambahkan Operator K-means

6. Parameter K-Means diatur dengan menentukan jumlah cluster (k) dan nilai maksimal iterasi.



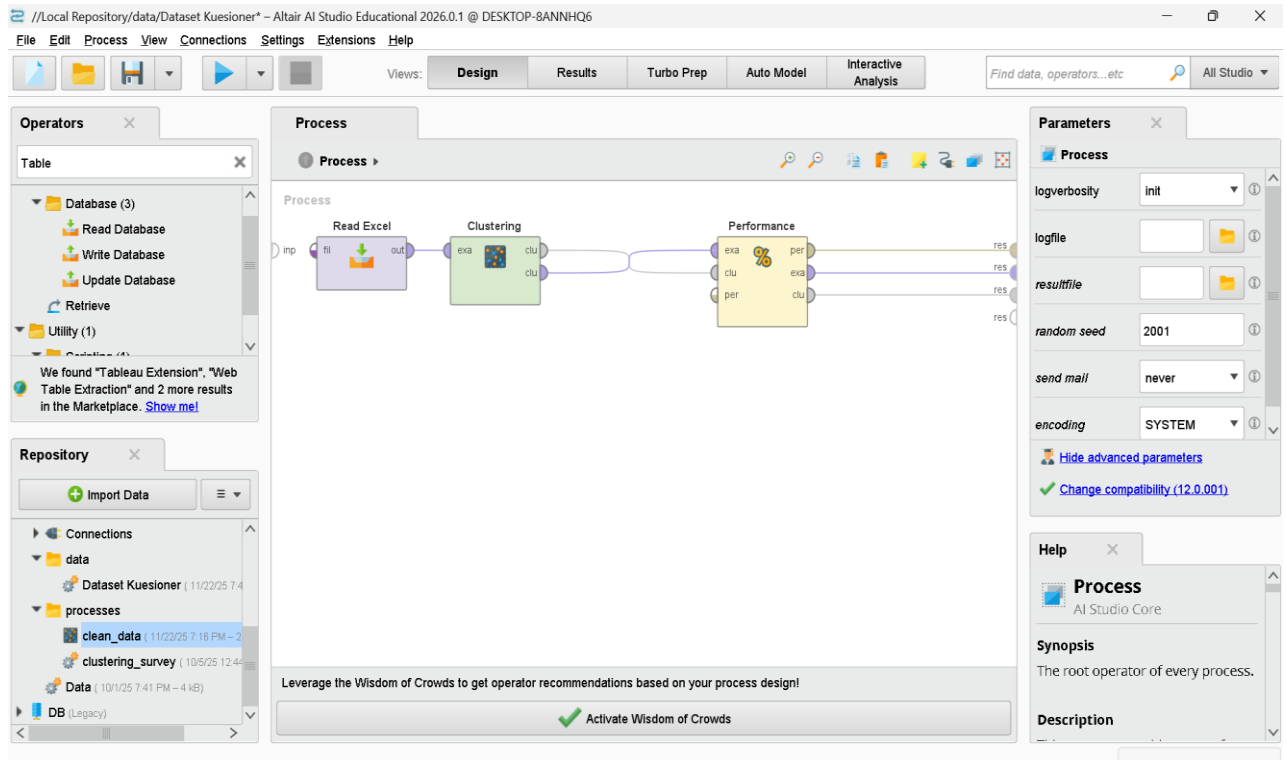
Gambar3. 10Parameters K-means

7. Penentuan jumlah cluster optimal dilakukan menggunakan metode Elbow, dimana hasil pengujian menunjukkan bahwa nilai k terbaik berada pada k = 3.



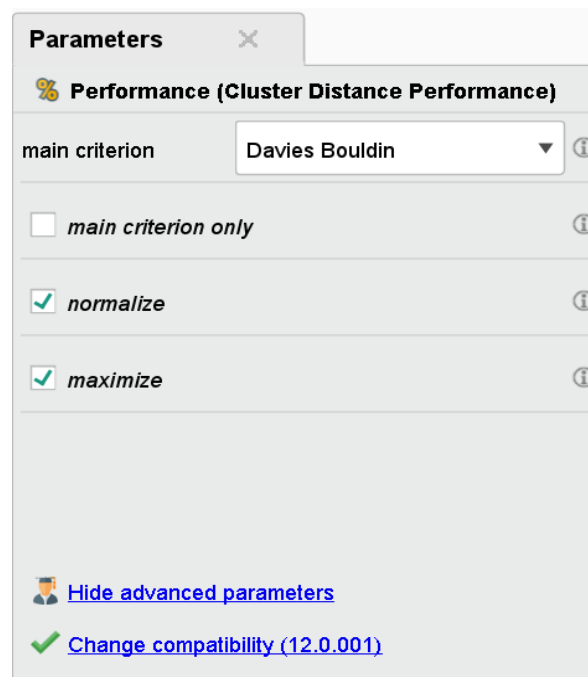
Gambar3. 11 Metode Elbow

8. Untuk mengukur kualitas hasil clustering, ditambahkan operator Cluster Distance Performance dengan kriteria Davies-Bouldin Index (DBI).



Gambar3. 12Operator Cluster Distance Performance

9. Parameter pada operator Cluster Distance Performance diatur dengan memilih DBI sebagai kriteria utama.



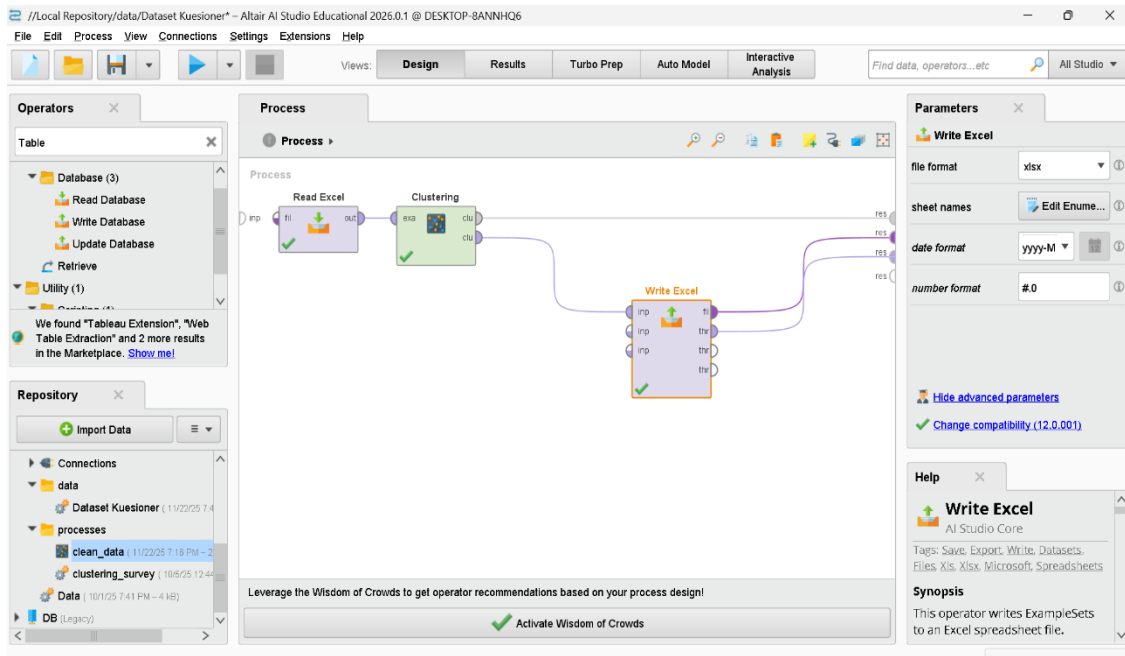
Gambar3. 13Parameters Cluster Distance Performance

Setelah seluruh proses selesai, data dijalankan (*running*) dan hasil clustering ditampilkan pada menu *Result*.

B. Pemindahan Hasil Clustering ke Excel

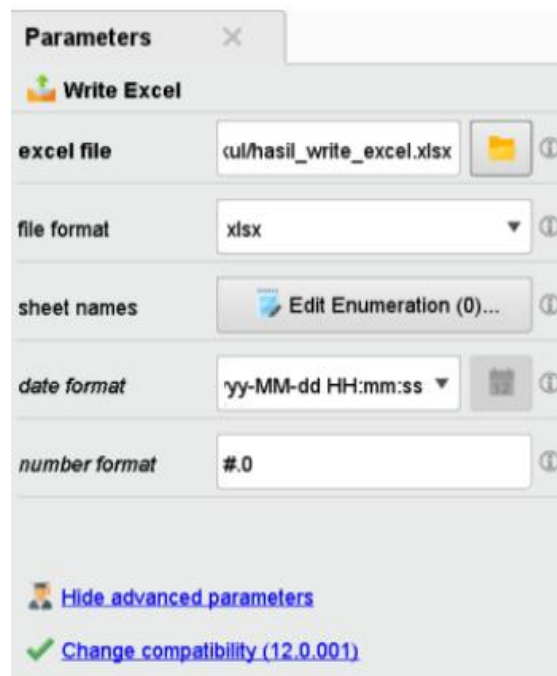
Setelah mendapatkan hasil cluster dan correlation matrix dari RapidMiner, langkah selanjutnya adalah menggunakan operator Write Excel untuk mentransfer data yang telah diproses ke format Excel.

1. *Operator Write Excel* digunakan untuk mengeksport hasil clustering dari RapidMiner ke dalam file Excel.



Gambar3. 14Penambahan Operator Write Excel

2. Parameter file tujuan diatur sesuai dengan lokasi penyimpanan yang diinginkan.



Gambar3. 15 Parameters Write Excel

Setelah proses running selesai, hasil clustering otomatis tersimpan ke dalam file Excel.

3.6 Hasil Pengujian

Pengujian dilakukan menggunakan algoritma K-Means dengan bantuan RapidMiner berdasarkan data hasil transformasi. Tujuan pengujian ini untuk mengetahui pengelompokan tingkat kepuasan pengguna Transjakarta terhadap kualitas pelayanan dan fasilitas.

3.7 Hasil Pengelompokan (*Clustering*)

Proses clustering menggunakan metode K-Means menghasilkan tiga cluster,

- Cluster 0: 31 data
- Cluster 1: 66 data
- Cluster 2: 3 data

Cluster Model

```
Cluster 0: 31 items
Cluster 1: 66 items
Cluster 2: 3 items
Total number of items: 100
```

Gambar3. 16Hasil Cluster Model

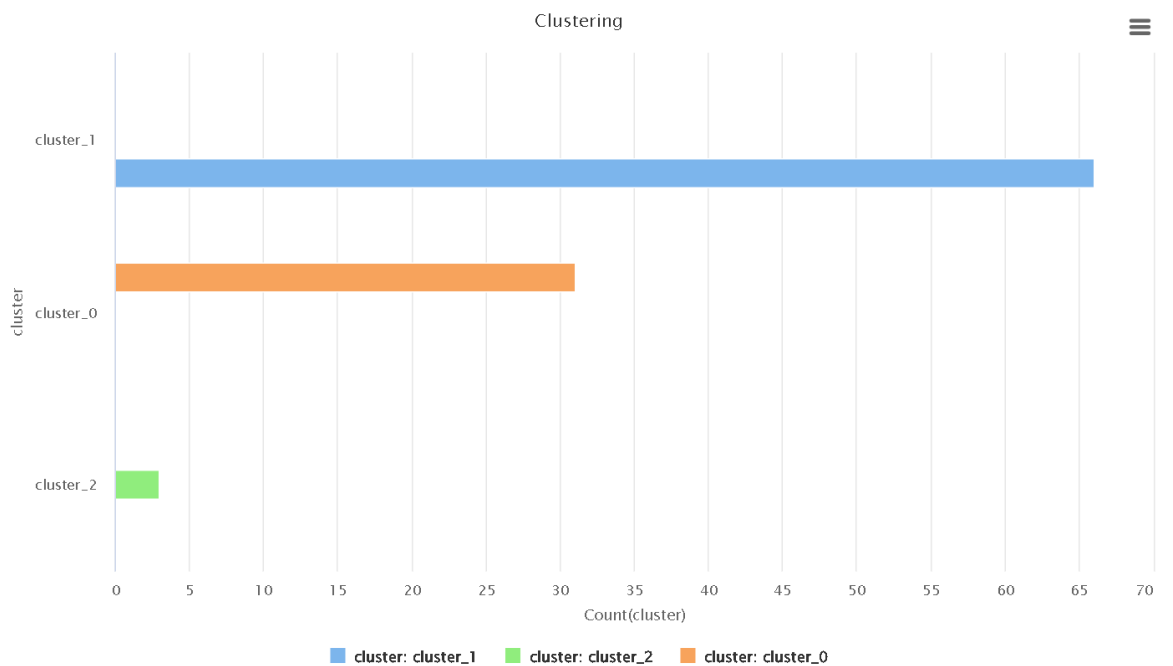
Dari total 100 responden, cluster 1 merupakan cluster dengan tingkat kepuasan tertinggi, cluster 0 mewakili kategori netral, dan cluster 2 termasuk dalam kategori tidak puas.

Attribute	cluster_0	cluster_1	cluster_2
KP PUAS	7.452	7.409	0.667
KP NETRAL	0.548	0.576	6
KP TIDAK PUAS	0	0.015	1.333
F PUAS	4.968	7.409	5
F NETRAL	2.935	0.455	1.667
F TIDAK PUAS	0.097	0.121	1.333

Gambar3. 17Nilai Jarak Centroid

Jika dilihat dari nilai centroid, cluster 1 memiliki nilai tertinggi pada Kualitas Pelayanan Puas (7,409) dan Fasilitas Puas (7,409) sehingga dikategorikan sebagai cluster puas. Cluster 0 juga menunjukkan nilai tinggi pada Kualitas Pelayanan Puas (7,452) namun masih memiliki nilai netral pada fasilitas, sehingga termasuk cluster cukup puas. Sementara itu, cluster 2 memiliki nilai tertinggi pada Kualitas Pelayanan Netral (6) serta nilai Tidak Puas yang lebih besar dibanding cluster lain, sehingga cluster 2 dikategorikan sebagai cluster netral hingga tidak puas.

Berdasarkan visualisasi grafik batang horizontal, dapat diketahui bahwa jumlah responden terbanyak berada pada cluster 1, sedangkan jumlah paling sedikit berada pada cluster 2



Gambar3. 18Bar (Horizontal)

3.7 Hasil Akurasi Davies Bouldin Index

Setelah proses pembentukan cluster dilakukan, tahap selanjutnya adalah mengevaluasi kualitas pengelompokan menggunakan Davies Bouldin Index (DBI) melalui perangkat lunak RapidMiner. DBI digunakan untuk mengukur tingkat keseragaman data dalam satu cluster dan tingkat pemisahan antar cluster. Semakin kecil nilai DBI yang dihasilkan, maka semakin baik kualitas hasil clustering karena menunjukkan cluster yang semakin homogen dan terpisah dengan jelas.

PerformanceVector

```
PerformanceVector:
Avg. within centroid distance: 0.653
Avg. within centroid distance_cluster_0: 0.672
Avg. within centroid distance_cluster_1: 0.486
Avg. within centroid distance_cluster_2: 4.148
Davies Bouldin: 0.143
```

Gambar3. 19Hasil Davies Bouldin Index

Berdasarkan hasil pengujian diperoleh nilai DBI sebesar 0,143 pada $k = 3$, sehingga dapat disimpulkan bahwa jumlah cluster terbaik dalam penelitian ini adalah 3 cluster. Nilai tersebut menunjukkan bahwa hasil pengelompokan data sudah termasuk dalam kategori baik.

KESIMPULAN

Berdasarkan hasil pengolahan data menggunakan algoritma K-Means Clustering, maka diperoleh beberapa kesimpulan sebagai berikut



1. Pengumpulan data dilakukan melalui penyebaran kuesioner dengan variabel kualitas pelayanan dan fasilitas sebagai indikator kepuasan pengguna. Algoritma K-Means Clustering digunakan untuk mengelompokkan data responden guna mengetahui pola tingkat kepuasan pengguna Transjakarta.
2. Penentuan jumlah cluster terbaik dilakukan menggunakan metode Elbow, sedangkan kualitas hasil clustering diuji menggunakan Davies Bouldin Index (DBI). Hasil pengujian menunjukkan nilai DBI sebesar 0,143 pada $k = 3$, sehingga menandakan bahwa hasil pengelompokan data sudah optimal dan memiliki kualitas yang baik.
3. Hasil pengolahan data menunjukkan bahwa tingkat kepuasan pengguna lebih dominan pada kategori puas dibandingkan netral dan tidak puas. Terbentuk 3 cluster, yaitu cluster tidak puas sebanyak 3 responden, cluster netral sebanyak 31 responden, dan cluster puas sebanyak 66 responden. Hasil ini dapat menjadi bahan evaluasi bagi pihak Transjakarta dalam meningkatkan kualitas pelayanan dan fasilitas.

Dengan demikian, penelitian ini berhasil mengelompokkan tingkat kepuasan pengguna Transjakarta Koridor B13 (Summarecon Bekasi - Blok M) menjadi 3 cluster menggunakan metode K-Means Clustering dengan pengujian kualitas menggunakan Davies Bouldin Index (DBI). Hasil clustering menunjukkan bahwa cluster 1 merupakan kategori puas dengan jumlah 66 responden, cluster 0 termasuk kategori netral atau cukup puas dengan 31 responden, dan cluster 2 merupakan kategori tidak puas dengan 3 responden. Hasil ini dapat menjadi dasar bagi pihak Transjakarta untuk melakukan evaluasi dan peningkatan kualitas pelayanan serta fasilitas secara lebih terarah.

DAFTAR PUSTAKA

- Akbar, T. (n.d.). *ANALISIS TINGKAT KEPUASAN PENGGUNA TRANSJAKARTA KORIDOR 10 (TJ PRIOK-PGC 2) TERHADAP KUALITAS PELAYANAN DAN FASILITAS MENGGUNAKAN METODE K-MEANS CLUSTERING*.
- Al Atros, K., Padri, A. R., Nurdiawan, O., Faqih, A., & Anwar, S. (2022). Model Klasifikasi Analisis Kepuasan Pengguna Perpustakaan Online Menggunakan K-Means dan Decision Tree. *JURIKOM (Jurnal Ris. Komputer)*, 8(6), 323-329.
- Anuur, C. M. (2022). *Jumlah Penumpang Bus Transjakarta Berdasarkan Koridor/Rute (2021)*. Databoks Katadata. <https://databoks.katadata.co.id/transportasi-logistik/statistik/9d26bc18b58a174/ini-rute-bus-transjakarta-dengan-jumlah-penumpang-terbanyak-pada-2021>
- Azwan, M., Kurniawan, R., Insani, F., Informatika, T., Islam, U., Syarif, S., Riau, K., & Dengue, D. B. (2022). *Penerapan Algoritma K-Means Clustering dan Correlation Matrix Untuk Menganalisis Risiko Penyebaran Demam Berdarah di Kota Pekanbaru*. 6(3), 1-6.
- Daffauzan, M. F. (2023). Analisa Penumpang Bus Transjakarta dengan Metode K-Means. *Journal Computer Science and Information Systems: J-Cosys*, 3(2), 61-64.
- Maulana, R., Raihan, M., & Santoso, I. (2023). Komparasi Algoritma Naive Bayes Dan K-Nearest Neighbor Pada Analisis Sentimen Terhadap Ulasan Pengguna Aplikasi Tokopedia. *Jurnal Teknologi Informasi: Jurnal Keilmuan Dan Aplikasi Bidang Teknik Informatika*, 17(2), 177-189.
- Purnandika, R. A., & Septiana, H. (2024). Purnandika, R., and Hana Septiana. "Kualitas layanan terhadap kepuasan pelanggan pada pengguna transportasi umum di Jakarta." *Jurnal Ilmiah Manajemen Ekonomi Dan Akuntansi (JIMEA)* 1.2 (2023): 25-32. *Jurnal Ilmiah Manajemen Ekonomi Dan Akuntansi*, 1(Februari), 25-32.
- Ulya, F. A., Abdullah, A. N., Hanan, T. A., & Nur, I. M. (2023). Pengelompokan Tingkat



Pengangguran Terbuka Di Jawa Tengah Menggunakan Metode K-Means Clustering: Grouping Open Unemployment Rates in Central Java Using the K-Means Clustering Method. *Journal Of Data Insights*, 1(2), 71-80.

Vanessa, B., & Prabantari, K. (2020). Analysis of the Relationship between Service Quality and Customer Satisfaction Level of Transjakarta Transportation. *Transaction Journal*, 12(1), 25-39.