



## IMPLEMENTASI SISTEM PENGOLAHAN DATA TERINTEGRASI DENGAN ALGORITMA K-MEANS PADA KNIME

Zakiah Nabila<sup>1</sup>, Ahmad Rio Adriansyah<sup>2</sup>

<sup>1,2</sup>Teknik Informatika, Sekolah Tinggi Teknologi Terpadu Nurul Fikri  
Jakarta Selatan, DKI Jakarta, Indonesia 12640  
zaki21071ti@student.nurulfikri.ac.id, arasy@nurulfikri.ac.id

### Abstract

*This research aims to develop and implement an integrated data processing system based on KNIME to analyze employee satisfaction at X School. The methodology involved collecting data via a survey distributed to 125 employees, integrating data from Google Sheets, preparing the data, applying the K-Means algorithm to cluster employees by satisfaction levels, and visualizing the results in an interactive dashboard. The research results indicate that the system was successfully built and can group employees into three clusters: Very Satisfied, Satisfied, and Less Satisfied. User acceptance testing (UAT) showed that the system met 80% of the testing criteria, indicating that most features functioned as expected by users. Evaluation using the Silhouette Coefficient produced an average value of 0.19, indicating less-than-optimal clustering quality, but the system still provided an overview of employee satisfaction levels. This system supports KNIME use for employee satisfaction analysis and provides strategic recommendations for X School to improve employee satisfaction and retention.*

**Keywords:** Clustering, Employee Satisfaction, Integrated Data, K-Means, KNIME

### Abstrak

Penelitian ini bertujuan untuk mengembangkan dan mengimplementasikan sistem pengolahan data terintegrasi berbasis KNIME untuk menganalisis kepuasan karyawan di Sekolah X. Metode yang digunakan meliputi pengumpulan data melalui survei yang didistribusikan kepada 125 karyawan, integrasi data dari Google Sheets, persiapan data, penerapan algoritma K-Means untuk mengelompokkan karyawan berdasarkan tingkat kepuasan, serta visualisasi hasil analisis melalui *dashboard* interaktif. Hasil penelitian menunjukkan bahwa sistem berhasil dibangun dan mampu mengelompokkan karyawan ke dalam tiga *cluster*: Sangat Puas, Puas, dan Kurang Puas. *User acceptance testing* (UAT) menunjukkan bahwa sistem memenuhi 80% dari kriteria pengujian, menunjukkan sebagian besar fitur berfungsi sesuai harapan pengguna. Evaluasi menggunakan Silhouette Coefficient menghasilkan nilai rata-rata 0.19, mengindikasikan kualitas *clustering* yang kurang optimal, namun sistem tetap memberikan gambaran umum mengenai tingkat kepuasan karyawan. Sistem ini memberikan kontribusi dalam pemanfaatan KNIME untuk analisis kepuasan karyawan, serta memberikan rekomendasi strategis bagi Sekolah X untuk meningkatkan kepuasan dan retensi karyawan.

**Kata kunci:** Clustering, Data Terintegrasi, K-Means, Kepuasan Karyawan, KNIME

### 1. PENDAHULUAN

Kepuasan karyawan merupakan salah satu indikator penting dalam keberhasilan suatu organisasi. Dalam era kerja yang semakin kompetitif, memahami kepuasan karyawan bukan hanya berdampak pada produktivitas individu, tetapi juga pada retensi sumber daya manusia[1]. Kondisi lingkungan kerja yang kondusif, baik fisik maupun psikologis, sangat penting untuk kepuasan karyawan. Lingkungan kerja yang sehat, aman, dan nyaman berkontribusi pada peningkatan produktivitas, karena menciptakan suasana yang positif dan meningkatkan motivasi kerja[2].

Tingkat *turnover* karyawan di Indonesia pada tahun 2023 mencapai 41%, dengan industri layanan profesional dan bisnis mencatat tingkat tertinggi, yaitu 57%, diikuti oleh industri konstruksi 54% dan perdagangan, transportasi, serta utilitas 49%[3]. Angka ini menekankan urgensi untuk menciptakan lingkungan kerja yang dapat mempertahankan dan mengembangkan karyawan.

Dalam proses analisis kepuasan karyawan, banyak organisasi menghadapi masalah dalam pengolahan data yang terfragmentasi. Data sering kali berasal dari berbagai

sumber, seperti survei dan wawancara, yang tidak terintegrasi dengan baik. Hal ini mengakibatkan kesulitan dalam mengakses dan menganalisis data secara holistik, yang pada gilirannya dapat menghambat pengambilan keputusan strategis.

Sekolah X, sebagai lembaga pendidikan inklusif di Kota Depok, juga menghadapi tantangan serupa dalam mengelola data kepuasan karyawannya. Data kepuasan karyawan yang ada sering kali berasal dari berbagai sumber seperti survei manual, umpan balik informal, dan data kinerja yang belum terintegrasi. Hal ini menyebabkan kesulitan dalam memperoleh gambaran komprehensif mengenai faktor-faktor yang mempengaruhi kepuasan karyawan, serta menghambat pengambilan keputusan berbasis data yang efektif. Pemilihan Sekolah X sebagai objek penelitian didasarkan pada ketersediaan data empiris, relevansi permasalahan, dan potensi manfaat untuk meningkatkan kepuasan dan retensi karyawan.

Penelitian ini diharapkan dapat memberikan manfaat bagi Sekolah X dalam pengambilan keputusan, peningkatan efisiensi pengolahan informasi, peningkatan kepuasan dan retensi karyawan, serta kontribusi akademis.

Untuk memastikan fokus dan efektivitas penelitian, batasan masalah ditetapkan sebagai berikut: (1) Penelitian dilakukan berdasarkan data yang diambil pada bulan April 2018. (2) Penelitian mengutamakan faktor-faktor kepuasan kerja seperti komunikasi, peluang pengembangan karier, dan lingkungan kerja. (3) *Tools* yang digunakan adalah KNIME *Analytic platform* versi 5.2.3.

## 2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan desain penelitian pengembangan (*research and development* / R&D) untuk membangun dan menguji sistem pengolahan data terintegrasi. Pendekatan ini dipilih untuk menguji efektivitas sistem dalam menganalisis kepuasan karyawan[4].

### 2.1 Metode Pengumpulan Data

Metode pengumpulan data dalam penelitian ini adalah survei yang didistribusikan kepada 125 karyawan di Sekolah X. Kuesioner survei dirancang untuk mengukur tingkat kepuasan karyawan terhadap berbagai aspek pekerjaan, seperti lingkungan kerja, komunikasi, gaji, dan kepemimpinan. Kuesioner dibuat menggunakan Google Forms, yang memungkinkan pengumpulan data secara *online*. Responden diminta memberikan penilaian menggunakan skala *Likert* 1-5, dengan 1 = "Sangat Tidak Setuju" dan 5 = "Sangat Setuju".

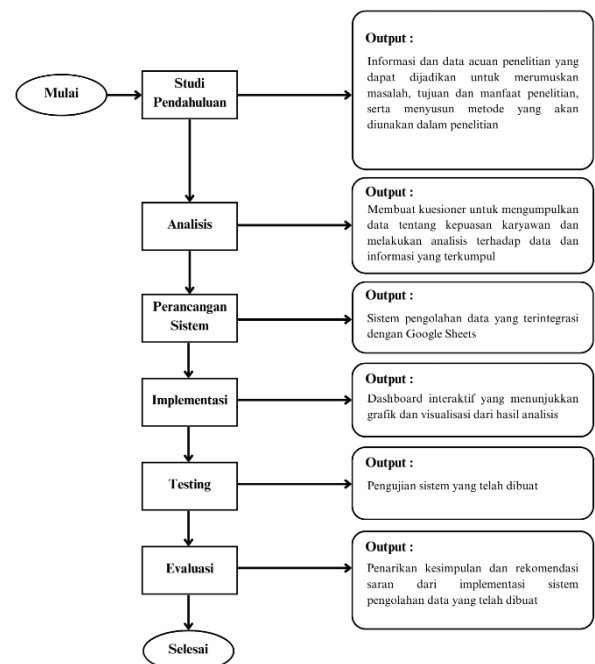
### 2.2 Metode Pengujian

Metode pengujian yang digunakan adalah *User acceptance testing* (UAT). Proses UAT melibatkan pengujian oleh pengguna akhir untuk memastikan sistem berfungsi sesuai dengan persyaratan yang telah ditetapkan. Di Sekolah X,

UAT berfungsi sebagai validasi akhir bahwa sistem pengolahan data terintegrasi telah berhasil memenuhi kebutuhan dan harapan manajemen[5]. Skenario pengujian mencakup tugas-tugas spesifik yang relevan dengan penggunaan sistem, seperti memasukkan data, menjalankan sistem, dan menghasilkan laporan. Umpan balik dari peserta UAT kemudian diambil untuk menilai kenyamanan, kemudahan penggunaan, dan fitur-fitur sistem.

### 2.2 Tahapan Penelitian

Penelitian ini dilakukan melalui beberapa tahapan yang terstruktur, dimulai dari studi pendahuluan hingga evaluasi hasil.



Gambar 1. Alur Penelitian

Alur penelitian dapat dilihat pada Gambar 1, yang mengilustrasikan tahapan penelitian secara sistematis, mulai dari studi pendahuluan, analisis, perancangan sistem, implementasi, pengujian, hingga evaluasi. Tahapan penelitian ini secara rinci meliputi:

#### 2.1.1 Studi Pendahuluan

Tahap ini melibatkan pengumpulan informasi dari dokumen internal Sekolah X, jurnal, dan artikel terkait kepuasan karyawan, analisis data, algoritma K-Means, dan penggunaan KNIME *Analytic platform*.

#### 2.1.2 Analisis

Data dikumpulkan melalui kuesioner yang didistribusikan kepada 125 responden. Ringkasan statistik deskriptif digunakan untuk memahami distribusi jawaban dan tren umum dalam data survei.

### 2.1.3 Perancangan Sistem

Perancangan sistem pengolahan data terintegrasi menggunakan *KNIME Analytic platform*, yang mencakup integrasi data, implementasi algoritma K-Means, dan visualisasi hasil.

### 2.1.3 Implementasi

Sistem diimplementasikan dengan mengintegrasikan data dari Google Sheets, melakukan persiapan data, menerapkan algoritma K-Means, dan membuat visualisasi hasil dalam bentuk *dashboard*.

### 2.1.3 Testing

Sistem diuji menggunakan metode UAT untuk memastikan fungsi dan fitur bekerja sesuai rancangan.

### 2.1.3 Evaluasi

Hasil pengujian UAT dievaluasi untuk menilai tingkat keberhasilan sistem dalam memenuhi kebutuhan pengguna dan mengidentifikasi area yang perlu diperbaiki.

## 3. HASIL DAN PEMBAHASAN

Data yang digunakan dalam penelitian ini diperoleh melalui kuesioner yang didistribusikan kepada 125 karyawan di Sekolah X pada April 2018. Kuesioner dirancang untuk mengukur tingkat kepuasan karyawan terhadap berbagai aspek pekerjaan, seperti komunikasi, lingkungan kerja, gaji, tunjangan, dan kepemimpinan. Responden diminta untuk memberikan penilaian menggunakan skala *Likert* 1-5, dengan 1 menunjukkan "Sangat Tidak Setuju" dan 5 menunjukkan "Sangat Setuju".

Data dari kuesioner yang dikumpulkan melalui Google Form secara otomatis tersimpan di Google Sheets. Setelah pengumpulan data, dilakukan analisis deskriptif untuk memahami distribusi jawaban dan tren umum dalam data survei.

**Tabel 1.** Ringkasan Statistik Deskriptif

Pernyataan	Min	Max	Mean	Std. Dev	Skewness	Kurtosis	N
Sekolah X merupakan salah satu tempat bekerja yang terbaik	1	5	4,010	0,842	0,710	-1,152	523
Sekolah X memperlakukan saya dengan baik	1	5	4,184	0,766	0,587	-1,312	523
Saya bangga mengatakan kepada orang lain bahwa saya bekerja di Sekolah X	1	5	4,192	0,849	0,721	-1,344	524
Mempertimbangkan segala hal, saya	1	5	3,896	0,860	0,739	-0,958	1.487

Pernyataan	Min	Max	Mean	Std. Dev	Skewness	Kurtosis	N
merasa puas bekerja di Sekolah X							
Saya puas dengan pekerjaan saya dan jenis tugas yang saya kerjakan	1	5	3,928	0,795	0,632	-0,947	491
Pekerjaan saya menantang dan menarik	1	5	4,056	0,816	0,666	-0,918	507

Tabel 1 menyajikan contoh ringkasan statistik deskriptif dari beberapa pertanyaan survei, termasuk nilai minimum, maksimum, rata-rata (mean), standar deviasi, skewness, kurtosis, dan jumlah keseluruhan. Dalam analisis data, analisis deskriptif berperan sebagai alat untuk memberikan deskripsi yang jelas dan ringkas mengenai suatu kondisi. Pendekatan ini memanfaatkan berbagai teknik visualisasi, seperti tabel dan grafik, untuk menyajikan data observasi secara efektif, sehingga memudahkan pembaca dalam memahami informasi statistik yang disajikan[6].

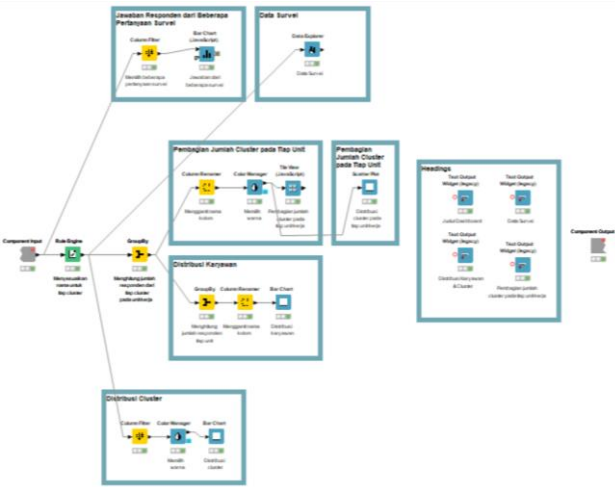
Berdasarkan analisis deskriptif yang telah dilakukan, secara umum karyawan di Sekolah X cenderung merasa puas dengan aspek yang ditanyakan dalam pertanyaan survei. Distribusi jawaban cenderung terkonsentrasi pada nilai-nilai yang lebih tinggi, dengan variasi jawaban yang relatif kecil.

Data survei kepuasan karyawan yang tersimpan di Google Sheets diintegrasikan ke dalam *KNIME Analytic platform* untuk memudahkan analisis. Integrasi dilakukan menggunakan *node* Google Authenticator, Google Sheets Connector, dan Google Sheets Reader. Proses integrasi dimulai dari otentikasi hingga pembacaan data. Keunggulan *KNIME* terletak pada pendekatan *no-code* yang memungkinkan pengguna membangun alur analisis secara modular dan fleksibel[7].

Setelah data berhasil diintegrasikan, langkah selanjutnya adalah persiapan data. Pada tahap ini, data disiapkan untuk analisis lebih lanjut. Proses persiapan data dilakukan dengan menggunakan *node String to Number* untuk transformasi data.

Setelah data siap, algoritma K-Means diterapkan untuk mengelompokkan karyawan berdasarkan tingkat kepuasan. Algoritma K-Means digunakan untuk mengidentifikasi pola dan struktur dalam data yang kompleks, mengelompokkan karyawan berdasarkan karakteristik yang serupa[8]. Kemudian dilakukan evaluasi hasil *clustering* untuk menilai kualitas pengelompokan data menggunakan *Silhouette Coefficient* yang merupakan alat ukur yang mengukur seberapa baik suatu data point berada dalam *cluster*-nya. Hal ini dilakukan dengan membandingkan kemiripan data point tersebut dengan data point lain dalam *cluster* yang sama, dibandingkan dengan data point di *cluster* lain[9].

Hasil evaluasi menunjukkan kualitas *clustering* yang kurang optimal. Selanjutnya hasil analisis divisualisasikan menggunakan *node Scatter Plot*, *Tile View*, *Bar Chart*, dan *Data Explorer* untuk mempresentasikan hasil analisis dalam bentuk *Dashboard*.



Gambar 2. Workflow Dashboard Visualisasi

Gambar 2 menunjukkan *workflow* untuk visualisasi hasil analisis dalam bentuk *dashboard*. Pengembangan *dashboard* difokuskan pada visualisasi data untuk mendukung pengambilan keputusan yang lebih baik. Melalui tampilan data yang terorganisir dan mudah dipahami, *dashboard* memfasilitasi pengguna dalam menganalisis informasi dan mengidentifikasi area-area penting[10]. *Dashboard* yang dihasilkan pada penelitian ini menyajikan informasi yang komprehensif dan mudah dipahami, memungkinkan manajemen untuk mengidentifikasi area yang perlu ditingkatkan.

Metode pengujian yang digunakan adalah *User acceptance testing* (UAT). UAT melibatkan beberapa karyawan dan direktur di Sekolah X. Skenario pengujian mencakup tugas-tugas spesifik yang relevan dengan penggunaan sistem, seperti memasukkan data, menjalankan sistem, dan menghasilkan laporan.

Tabel 2. Pengujian Sistem

Fitur yang Diuji	Data Uji	Hasil yang Diharapkan	Status Hasil
Integrasi Google Drive	Akun Google dengan akses ke file Google Drive berisi data survei	Data dari Google Drive berhasil dibaca dan ditampilkan di KNIME	Google Drive berhasil terintegrasi dengan KNIME dan data terbaca
Pengubahan tipe data String Menjadi Numerik (String to Number)	Data String yang berisi angka dalam format teks	Data berhasil dikonversi menjadi numerik	Data berhasil dikonversi

Fitur yang Diuji	Data Uji	Hasil yang Diharapkan	Status Hasil
Clustering K-Means	Keseluruhan dataset yang telah dipersiapkan sebelumnya	Data dikelompokkan ke dalam 3 cluster	Data berhasil dikelompokkan ke dalam 3 cluster
Evaluasi Silhouette Coefficient	Hasil clustering dari node K-Means	Nilai rata-rata Silhouette Coefficient mendekati 1	Tidak berhasil, hasil Silhouette Coefficient yang muncul hanya 0.19
Dashboard Visualisasi	Dashboard Visualisasi yang telah disusun	Dashboard dapat digunakan secara interaktif yang dapat memberikan wawasan bagi pengguna	Berhasil menampilkan Dashboard visualisasi yang interaktif dengan baik

Tabel 2 merupakan skenario pengujian sistem dengan hasil yang menunjukkan bahwa sistem memenuhi 80% dari kriteria pengujian. Penelitian ini berhasil membangun sistem pengolahan data terintegrasi untuk menganalisis kepuasan karyawan di Sekolah X menggunakan KNIME. Sistem ini mampu mengintegrasikan data dari Google Sheets, melakukan persiapan data, menerapkan algoritma K-Means, dan menghasilkan visualisasi hasil analisis dalam bentuk *Dashboard*.

Hasil penelitian menunjukkan bahwa sistem berhasil mengelompokkan karyawan ke dalam tiga *cluster*: Sangat Puas, Puas, dan Kurang Puas. Hasil UAT menunjukkan bahwa sistem berfungsi sesuai dengan harapan pengguna. Penelitian ini memiliki persamaan dan perbedaan dengan penelitian sebelumnya. Persamaannya adalah penggunaan algoritma K-Means untuk *clustering* dan penerapan metode kuantitatif. Perbedaannya terletak pada konteks penelitian, yaitu analisis kepuasan karyawan di Sekolah X, serta penggunaan KNIME sebagai platform analisis data.

4. KESIMPULAN

Penelitian ini berhasil mencapai tujuan utama, yaitu membangun dan mengimplementasikan sistem pengolahan data terintegrasi untuk menganalisis kepuasan karyawan di Sekolah X menggunakan KNIME. Sistem yang dibangun mampu mengintegrasikan data dari Google Sheets, melakukan persiapan data, menerapkan algoritma K-Means untuk pengelompokan, dan menghasilkan visualisasi hasil analisis dalam bentuk *dashboard* yang informatif.

Hasil penelitian menunjukkan bahwa sistem berhasil mengelompokkan karyawan ke dalam tiga *cluster* utama: Sangat Puas, Puas, dan Kurang Puas. Temuan ini memberikan gambaran yang jelas mengenai variasi tingkat kepuasan karyawan di Sekolah X, yang dapat menjadi dasar bagi pengambilan keputusan strategis. Implementasi sistem

ini juga terbukti efisien dalam proses pengolahan data, yang sebelumnya dilakukan secara manual.

Evaluasi sistem menggunakan *User acceptance testing* (UAT) menunjukkan bahwa sistem berfungsi sesuai dengan harapan pengguna, dengan tingkat pemenuhan kriteria pengujian mencapai 80%. Hal ini mengindikasikan bahwa sistem tidak hanya efektif dalam menghasilkan analisis, tetapi juga mudah digunakan dan memberikan manfaat nyata bagi pihak sekolah. Meskipun evaluasi Silhouette Coefficient menunjukkan kualitas *clustering* yang kurang optimal, sistem tetap memberikan gambaran umum mengenai tingkat kepuasan karyawan.

#### DAFTAR PUSTAKA

- [1] D. Vahera and Onsardi, "Analisis Gaya Kepemimpinan, Motivasi dan Lingkungan Kerja Terhadap Kepuasan Kerja Karyawan," *Jurnal Manajemen Modal Insani dan Bisnis*, Jul. 2021.
- [2] C. Andriani and Onsardi, "Pengaruh Kompensasi, Pelatihan Kerja, dan Kepuasan Kerja Terhadap Kinerja Karyawan Marketing," *Jurnal Etrepreneur dan Manajemen Sains*, Jun. 2020.
- [3] A. R. Nadzirah, "Pengaruh *Work-Life Balance* dan *Employee Engagement* Terhadap *Turnover Intention* Karyawan Pada PT Suitmedia Kreasi Indonesia," 2024. Accessed: May 15, 2025. [Online]. Available: [https://repository.upi.edu/126098/2/S\\_PEM\\_2000994\\_Chapter%201.pdf](https://repository.upi.edu/126098/2/S_PEM_2000994_Chapter%201.pdf)
- [4] M. Waruwu, "Metode Penelitian dan Pengembangan (R&D): Konsep, Jenis, Tahapan dan Kelebihan," *Jurnal Ilmiah Profesi Pendidikan*, vol. 9, no. 2, pp. 1220–1230, May 2024, doi: 10.29303/jipp.v9i2.2141.
- [5] M. Asqia, H. Aditijawijaya, Y. Zulkarnain, A. Fadlila, and Z. Imaduddin, "Pengembangan Sistem Pengajuan Surat Berbasis GSuite Untuk Meningkatkan Kemudahan Akses Layanan Administrasi Akademik Untuk Mahasiswa," *Teknika*, vol. 11, no. 3, pp. 197–207, Oct. 2022, doi: 10.34148/teknika.v11i3.547.
- [6] Y. Mansur, "Analisis Perkembangan Penduduk Miskin, Karakteristik Kemiskinan dan Kedalaman Kemiskinan di Indonesia," *Jurnal EMT KITA*, vol. 8, no. 1, pp. 18–31, Jan. 2024, doi: 10.35870/emt.v8i1.1930.
- [7] V. Palacios *et al.*, *Best of KNIME The COTM Collection*. Switzerland: Knime Press, 2023. Accessed: May 09, 2025. [Online]. Available: <https://www.knime.com/knimepress/cotm>
- [8] B. D. Lund and J. Ma, "A Review of Cluster Analysis Techniques and Their Uses in Library and Information Science Research: K-means and K-medoids Clustering," *Performance Measurement and Metrics*, vol. 22, no. 3, pp. 161–173, Nov. 2021, doi: 10.1108/PM-05-2021-0026.
- [9] R. Hidayati, A. Zubair, A. H. Pratama, and L. Indana, "Analisis Silhouette Coefficient pada 6 Perhitungan Jarak K-Means Clustering," *Techno.Com*, May 2021. Accessed: Mar. 21, 2025. [Online]. Available: <https://eprints.unmer.ac.id/id/eprint/3004/1/Analisis%20Silhouette%20Coefficient%20pada%206%20Perhitungan%20Jarak.pdf>
- [10] A. Rahmah, P. Sukmasetya, M. S. Romadhon, and A. R. Adriansyah, "Developing Distance Learning Monitoring Dashboard with Google Sheet: An Approach for Flexible and Low-Price Solution in Pandemic Era," in *7th International Conference on ICT for Smart Society: AIoT for Smart Society, ICISS 2020 - Proceeding, Institute of Electrical and Electronics Engineers Inc.*, Nov. 2020. doi: 10.1109/ICISS50791.2020.9307558.