

Klasterisasi dan Analisis Tren Tenaga Kerja Berdasarkan Gender, Golongan, dan Sektor Usaha Menggunakan Metode *K-Means*

Muhammad Deraya Kautsar¹, Muhammad Bimo Krisyono², Muhammad Hafidz Diaz Alghany³, Rizky Wahyu Saputra⁴, Zurnan Alfian⁵, Jupron⁶

^{1,2,3,4,5,6} Universitas Pamulang, Indonesia

E-mail: alghany0509@gmail.com

Article History:

Received: 24 Juni 2025

Revised: 03 September 2025

Accepted: 22 September 2025

Keywords: *Data Mining, K-Means Clustering, Employee Clustering, Gender Inequality, Business Sector.*

Abstract: *Studi ini membahas analisis dan pengelompokan data ketenagakerjaan di Indonesia berdasarkan gender, golongan karyawan, dan sektor usaha dengan menggunakan pendekatan data mining. Studi ini dilatarbelakangi oleh ketimpangan distribusi tenaga kerja serta perbedaan struktur dan pendapatan antarkelompok kerja. Tiga studi kasus dianalisis, yaitu: 1) distribusi karyawan berdasarkan golongan dan gender; 2) perbandingan pendapatan pria dan wanita menurut jenis pekerjaan; dan 3) tren pertumbuhan tenaga kerja pada sektor Usaha Mikro, Kecil, dan Menengah (UMKM) dan usaha besar. Metode pengelompokan K-Means digunakan untuk melakukan segmentasi data, sedangkan analisis deskriptif digunakan untuk mengevaluasi pertumbuhan angkatan kerja. Data diperoleh dari buku referensi dan dikonversi ke format digital untuk keperluan analisis. Evaluasi hasil pengelompokan menggunakan Indeks Davies-Bouldin menunjukkan bahwa kualitas kluster bervariasi. Hasil studi mengungkap ketimpangan distribusi tenaga kerja berbasis gender dan menegaskan peran sektor UMKM yang dominan dan stabil dalam penyerapan tenaga kerja. Temuan ini diharapkan dapat mendukung kebijakan pengelolaan sumber daya manusia dan strategi pengembangan ketenagakerjaan di Indonesia.*

PENDAHULUAN

Perkembangan dunia kerja di Indonesia saat ini menunjukkan dinamika yang kompleks dari waktu ke waktu. Ketidakseimbangan distribusi tenaga kerja yang tidak hanya terjadi pada aspek kuantitas, tetapi juga dalam struktur komposisinya yang melibatkan gender, golongan jabatan, pendapatan, dan sektor usaha. Misalnya, dalam banyak kasus, perbedaan jumlah tenaga kerja antara laki-laki dan perempuan yang masih cukup signifikan, baik dalam sektor formal maupun informal. Distribusi tenaga kerja yang berdasarkan golongan jabatan atau tingkat pendapatan sering kali tidak merata, mencerminkan adanya kesenjangan struktural yang berpotensi menghambat efisiensi dan keadilan dalam sistem ketenagakerjaan nasional (Harjanto & Ariyani, 2019). Ketimpangan tersebut juga tercermin dalam pertumbuhan sektor usaha, khususnya usaha mikro, kecil, dan menengah (UMKM) dengan usaha besar, di mana UMKM seringkali kurang terfasilitasi aspek produktivitas dan akses tenaga kerja yang kompeten (Nugraha dkk., 2022).

Melalui permasalahan tersebut, maka dibutuhkannya pendekatan yang mampu memberikan

pemetaan serta pemahaman yang lebih mendalam terhadap struktur dan tren ketenagakerjaan saat ini. Salah satu pendekatan yang semakin relevan dan banyak digunakan dalam era data besar saat ini adalah teknik *data mining* (Arhami & Nasir, 2020), khususnya metode klusterisasi. Teknik ini yang memungkinkan mengelompokkan data berdasarkan pada kesamaan karakteristik tertentu, sehingga mengungkap pola-pola tersembunyi ini yang tidak dapat terlihat melalui analisis secara konvensional. Metode *K-Means Clustering* ini menjadi salah satu algoritma yang populer karena kemampuannya dalam membentuk klaster dari data yang multidimensional dengan efisiensi tinggi (Kusuma & Ellyana, 2018). Dengan menerapkan metode ini, peneliti dapat menyusun gambaran yang lebih terstruktur mengenai kelompok-kelompok tenaga kerja berdasarkan atribut seperti jenis kelamin, tingkat jabatan, pendapatan, dan sektor usaha.

Penelitian ini yang menggunakan data tenaga kerja yang dikumpulkan dari berbagai sumber literatur dan telah diolah ke dalam bentuk digital agar sesuai dengan format analisis klaster. Fokus utama penelitian adalah mengelompokkan data tenaga kerja berdasarkan atribut yang relevan menggunakan metode *K-Means*, kemudian menganalisis hasil klusterisasi untuk mengidentifikasi pola distribusi maupun potensi ketimpangan dalam pasar tenaga kerja. Tidak hanya itu, penelitian ini juga menyajikan visualisasi dari hasil klaster yang diperoleh serta tren pertumbuhan tenaga kerja dalam berbagai sektor, sehingga mampu menjadi dasar bagi penyusunan strategi pengelolaan sumber daya manusia secara lebih terarah dan berbasis data.

Melalui penelitian ini, yang diharapkan dapat diperoleh gambaran secara menyeluruh tentang kondisi distribusi tenaga kerja di Indonesia. Temuan dari klusterisasi ini yang dapat menjadi acuan penting bagi para pemangku kebijakan dalam merancang intervensi atau program yang lebih efektif dan berkeadilan, baik di dalam hal penempatan tenaga kerja, pengembangan kapasitas, maupun peningkatan produktivitas sektor usaha. Selain itu, penelitian ini juga memberikan kontribusi nyata dalam menunjukkan bagaimana metode *data mining* seperti *K-Means* yang dapat diterapkan secara langsung dalam isu-isu sosial ekonomi seperti ketenagakerjaan. Sehingga hasil penelitian ini tidak hanya bernilai akademik saja, tetapi juga praktis dan strategis dalam mendukung pengambilan keputusan di bidang ketenagakerjaan nasional.

LANDASAN TEORI

Pada dunia digitalisasi ini, *data mining* menjadi pendekatan utama dalam mengungkap pola-pola tersembunyi yang ada pada kumpulan data besar (Setiawan dkk., 2023). Proses ini melibatkan penggunaan teknik statistik, matematika, dan algoritma *machine learning* untuk mampu menggali informasi yang sebelumnya tidak teridentifikasi. Melalui *data mining*, berbagai organisasi dan peneliti dapat memanfaatkan data yang ada untuk memperoleh wawasan yang dapat diterapkan dalam pengambilan keputusan strategis (Mardiani dkk., 2023). *Data mining* yang memungkinkan untuk memahami struktur pasar kerja, pola distribusi tenaga kerja, hingga ketimpangan yang ada dalam berbagai dimensi seperti gender, golongan jabatan, dan sektor usaha.

Salah satu teknik yang sering digunakan dalam *data mining* adalah klusterisasi, berfungsi untuk mengelompokkan data berdasarkan kesamaan karakteristik. Dalam klusterisasi, data yang memiliki kemiripan dikelompokkan bersama, sementara data berbeda berada dalam kelompok terpisah. Proses ini memungkinkan identifikasi pola yang mungkin tidak terlihat secara langsung. Misalnya, dalam analisis tenaga kerja, klusterisasi dapat mengungkapkan adanya kecenderungan tertentu, seperti perbedaan distribusi antara tenaga kerja laki-laki dan perempuan di berbagai sektor usaha. Teknik ini juga dapat memberikan wawasan mengenai ketimpangan dalam hal

jabatan dan pendapatan, yang seringkali tersembunyi dalam data besar yang kompleks.

Di antara berbagai algoritma klasterisasi yang ada, *K-Means* adalah salah satu yang paling banyak digunakan karena kesederhanaannya dan efektivitasnya dalam mengelompokkan data-data (Anjarwati, 2025). *K-Means* ini bekerja dengan menentukan jumlah klaster (*K*) dahulu, kemudian mengelompokkan data berdasarkan kedekatannya dengan pusat klaster atau *centroid*. Proses ini akan diulang berulang kali hingga posisi *centroid* tidak berubah lagi atau iterasi mencapai batas yang telah ditentukan. Salah satu keunggulan *K-Means* adalah kemampuannya menangani *dataset* besar dan mengelompokkan data secara efisien (Maulidhia dkk., 2025). Namun, pemilihan jumlah klaster yang tepat tetap menjadi tantangan yang perlu diperhatikan untuk mampu menghasilkan klaster yang bermakna dan representatif.

Klasterisasi tidak berhenti pada tahap pembentukan kelompok data saja. Evaluasi terhadap kualitas klaster yang dihasilkan sangat penting untuk memastikan bahwa hasil klasterisasi benar-benar mencerminkan pola yang ada dalam data. Berbagai metode evaluasi digunakan untuk menilai seberapa baik data dikelompokkan, antara lain dengan menggunakan *Silhouette Score*, *Calinski-Harabasz Index*, dan *Davies-Bouldin Indeks* (Harani dkk., 2020). *Silhouette Score*, misalnya, mengukur seberapa baik setiap objek dalam klaster cocok dengan kelompoknya yang dibandingkan dengan klaster lainnya (Rousseeuw, 1987). Indeks lain, *Calinski-Harabasz* dan *Davies-Bouldin*, memberikan ukuran terkait seberapa jauh antar klaster terpisah dan seberapa kompak klaster tersebut. Melalui evaluasi ini, memastikan klasterisasi yang dilakukan tidak hanya akurat, tetapi juga relevan dalam menganalisis tren dan ketimpangan distribusi tenaga kerja di Indonesia.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif deskriptif untuk menganalisis data tenaga kerja di Indonesia dengan fokus pada penerapan metode klasterisasi *K-Means*. Pendekatan ini yang bertujuan untuk menggambarkan pola distribusi tenaga kerja berdasarkan atribut seperti gender, golongan jabatan, pendapatan, dan sektor usaha. Dengan menggunakan metode klasterisasi, maka penelitian ini yang akan memberikan gambaran yang lebih terstruktur dan terperinci mengenai kelompok-kelompok dalam data tenaga kerja, sehingga memudahkan analisis terhadap potensi ketimpangan dan tren yang ada (Azzam dkk., 2024). Proses analisis ini dilakukan komputasional untuk mampu mendapatkan pemetaan yang lebih objektif dan akurat.

Data yang digunakan dalam penelitian ini berasal dari berbagai sumber literatur yang relevan, yang telah dikumpulkan dan dikonversi ke dalam format digital (CSV) untuk bisa mempermudah analisis secara lebih lanjut. Sumber data yang digunakan mencakup tiga kategori utama, yaitu data distribusi pegawai berdasarkan golongan dan gender dari buku ekonomi, data pendapatan pria dan wanita berdasarkan jenis pekerjaan dari buku manajemen SDM, serta data pertumbuhan tenaga kerja sektor UMKM dan usaha besar yang diperoleh dari buku hukum koperasi UMKM. Data ini yang dipilih karena mencakup berbagai dimensi penting yang berkaitan dengan struktur dari ketenagakerjaan di Indonesia dan dapat memberikan gambaran mengenai dinamika pasar kerja.

Teknik pengumpulan data dilakukan dengan cara mengidentifikasi dan menyalin tabel data yang relevan dari buku-buku sumber tersebut, kemudian mentransformasikannya ke dalam format digital (CSV) agar dapat diproses lebih lanjut. Setelah data dikonversi, tahap berikutnya adalah pembersihan dan penyesuaian data agar siap untuk proses klasterisasi. Selanjutnya, dalam analisis data, dilakukan pra-pemrosesan yang meliputi normalisasi fitur numerik seperti pendapatan, jumlah pegawai, dan rasio gender, yang memastikan setiap atribut memiliki skala

yang sebanding. Klasterisasi ini dilakukan dengan menggunakan algoritma *K-Means*, di mana langkah-langkah meliputi penentuan jumlah klaster (*K*), pengelompokan data berdasarkan kedekatannya dengan *centroid*, dan evaluasi hasil klasterisasi menggunakan beberapa ukuran seperti *Silhouette Score*, *Calinski-Harabasz Index*, serta visualisasi *scatter plot* untuk menggambarkan distribusi data dalam klaster. Terakhir, hasil klasterisasi dianalisis untuk memberikan interpretasi yang relevan dalam konteks ketenagakerjaan di Indonesia, yang mana dapat digunakan untuk mendukung pengambilan keputusan terkait kebijakan dan strategi pengelolaan tenaga kerja. Analisis ini dilakukan dengan menggunakan perangkat lunak *Python* dan berbagai *library* seperti *Pandas*, *NumPy*, *Matplotlib*, dan *scikit-learn* untuk mempermudah pemrosesan dan visualisasi data.

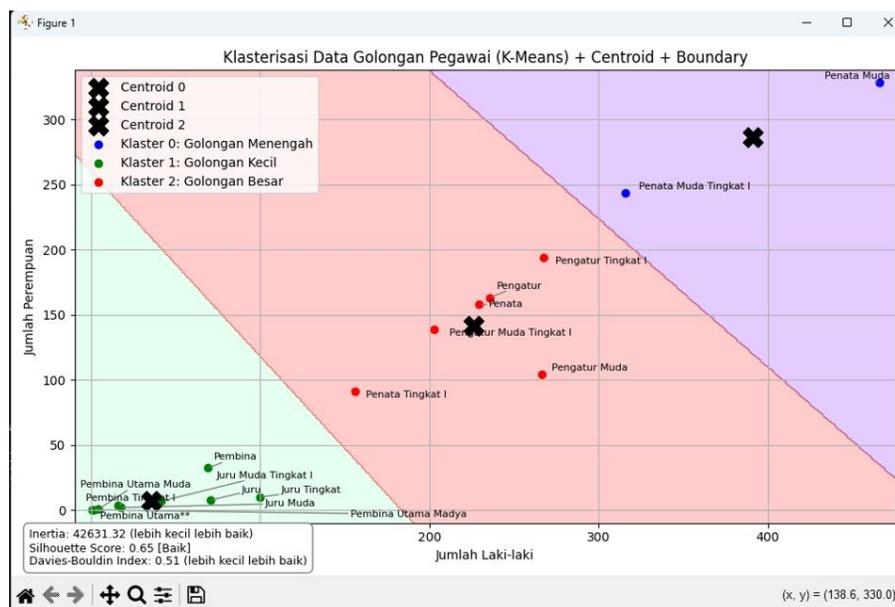
HASIL DAN PEMBAHASAN

1. Klasterisasi Data Pegawai Berdasarkan Golongan dan Gender

Hasil analisis klasterisasi menggunakan metode *K-Means* pada data jumlah pegawai laki-laki dan perempuan berdasarkan golongan jabatan menghasilkan tiga klaster dengan karakteristik yang berbeda. Klaster 0, yang dapat digolongkan sebagai klaster besar, mencakup golongan jabatan seperti "Pengatur" dan "Penata Muda", yang memiliki jumlah pegawai yang tinggi. Klaster ini yang menggambarkan golongan dengan tingkat jabatan yang lebih rendah, di mana jumlah pegawai relatif lebih banyak dan menunjukkan kebutuhan yang besar pada posisi-posisi di level bawah dalam struktur organisasi.

Klaster 1, yang juga disebut sebagai klaster menengah, terdiri dari golongan jabatan seperti "Penata" dan "Juru". Kelompok ini memiliki jumlah pegawai yang lebih sedikit dibandingkan dengan Klaster 0, tetapi masih mencerminkan jabatan dengan jumlah pegawai yang masih cukup signifikan. Golongan ini yang berada di tengah-tengah hierarki organisasi, dengan posisi jabatan yang lebih tinggi dari golongan di Klaster 0, tetapi belum mencapai pangkat yang tinggi seperti yang ada di Klaster 2.

Klaster 2, yang disebut sebagai klaster kecil, mencakup golongan jabatan dengan pangkat tinggi seperti "Pembina Utama", yang memiliki jumlah pegawai paling sedikit. Klaster ini yang mencerminkan jabatan-jabatan tinggi dalam organisasi yang hanya diisi oleh sejumlah kecil pegawai. Hasil ini menunjukkan fenomena yang umum dalam struktur organisasi, yaitu semakin tinggi pangkat jabatan, semakin sedikit jumlah pegawai yang mendudukinya. Visualisasi hasil klasterisasi ini dapat dilihat pada Gambar 1, yang menggambarkan distribusi jumlah pegawai dalam setiap klaster.



Gambar 1. Hasil Klasterisasi Data Golongan Pegawai

Hasil klasterisasi ini menggambarkan struktur organisasi yang berbentuk piramida, di mana semakin tinggi pangkat jabatan, semakin sedikit jumlah pegawai yang mendudukinya. Temuan ini sesuai dengan pola umum yang terlihat dalam organisasi di Indonesia, di mana posisi-posisi jabatan rendah memiliki jumlah pegawai yang lebih banyak, sementara posisi jabatan tinggi lebih terbatas dalam jumlah pegawainya. Struktur ini juga menunjukkan bahwa tantangan dalam distribusi tenaga kerja di tingkat atas, di mana jumlah posisi yang tersedia lebih sedikit, meskipun persaingan dan kesempatan untuk naik ke jabatan yang lebih tinggi sangat terbatas. Hasil analisis ini memberikan wawasan tentang ketimpangan distribusi tenaga kerja berdasarkan pangkat jabatan yang perlu diperhatikan dalam merancang kebijakan ketenagakerjaan yang lebih adil dan merata.

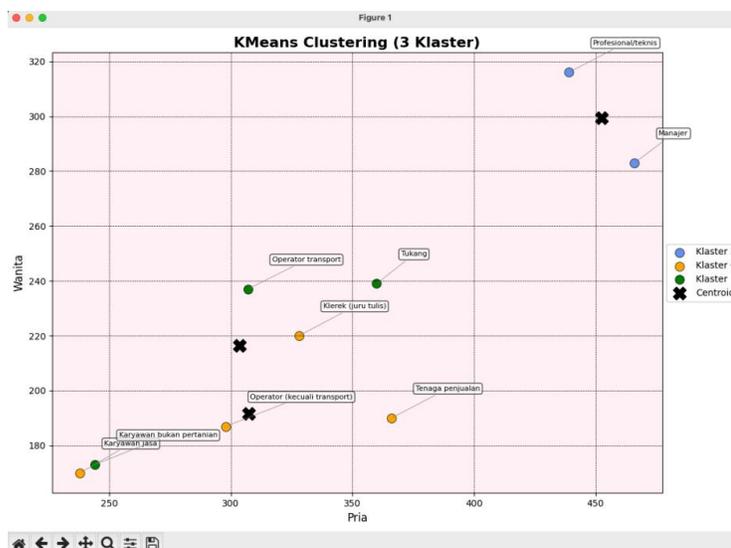
2. Klasterisasi Jenis Pekerjaan Berdasarkan Gender dan Pendapatan

Hasil analisis klasterisasi yang menggunakan metode *K-Means* pada data pendapatan rata-rata pria dan wanita per jenis pekerjaan ini menghasilkan tiga klaster dengan karakteristik yang berbeda. Klaster 0 ini mencakup pekerjaan yang didominasi pria, seperti "Operator", "Tukang", dan "Tenaga Penjualan". Pekerjaan-pekerjaan ini didominasi oleh pria dengan proporsi wanita yang sangat kecil, serta memiliki pendapatan yang relatif lebih rendah. Klaster ini menunjukkan sektor pekerjaan yang lebih banyak diisi oleh pria dan seringkali terkait dengan pekerjaan yang lebih teknis atau fisik.

Klaster 1 yang mencakup pekerjaan dengan proporsi wanita yang besar, seperti sektor "Jasa" dan "Non-Pertanian". Golongan pekerjaan ini yang menunjukkan adanya dominasi wanita, namun dengan pendapatan yang lebih rendah dibandingkan dengan pekerjaan di Klaster 2 itu. Meskipun wanita mengisi lebih banyak posisi dalam sektor ini, ketimpangan dalam hal pendapatan masih tetap terlihat, meskipun dalam proporsi yang lebih seimbang antara pria dan wanita.

Klaster 2 ini yang mencakup pekerjaan bergaji tinggi, seperti "Profesional" dan "Manajer", tetapi dengan ketimpangan gender yang signifikan. Meskipun jenis pekerjaan ini menawarkan gaji yang lebih tinggi, mayoritas pengisi posisi ini adalah pria. Klaster ini menggambarkan posisi-posisi dengan pendapatan yang tinggi, tetapi ketimpangan gender tetap terlihat di posisi-posisi

penting dan bergaji tinggi. Visualisasi hasil klasterisasi ini dapat dilihat pada Gambar 2, menggambarkan distribusi pendapatan dan proporsi gender dalam setiap klaster.



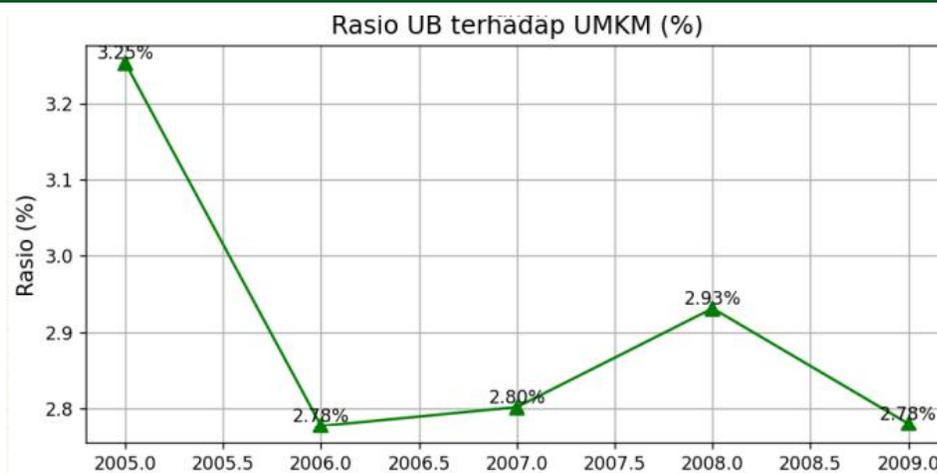
Gambar 2. Klasterisasi Berdasarkan Gender dan Pendapatan

Temuan ini yang menunjukkan bahwa ketimpangan pendapatan berdasarkan gender masih signifikan di berbagai jenis pekerjaan. Klaster 0 menunjukkan dominasi pria dalam pekerjaan dengan pendapatan lebih rendah, sementara Klaster 1 menunjukkan dominasi wanita di pekerjaan dengan pendapatan rendah. Di sisi lain, Klaster 2 mencerminkan pekerjaan bergaji tinggi, tetapi tetap memperlihatkan ketimpangan gender yang jelas, dengan mayoritas posisi-posisi bergaji tinggi diisi oleh pria. Hal ini yang mengindikasikan bahwa meskipun terdapat kesetaraan dalam jumlah wanita di beberapa sektor, ketimpangan pendapatan masih menjadi tantangan besar, terutama di sektor pekerjaan yang lebih tinggi dan profesional.

3. Analisis Tren Tenaga Kerja UMKM Besar (2005-2009)

Hasil analisis terhadap data pertumbuhan tenaga kerja sektor UMKM dan usaha besar selama lima tahun menunjukkan pola yang berbeda. Sektor UMKM mengalami pertumbuhan yang relatif stabil, dengan angka pertumbuhan berkisar antara +2% hingga +5% per tahun. Hal ini yang mana mencerminkan kestabilan sektor UMKM dalam menyerap tenaga kerja, meskipun dengan laju pertumbuhan yang lebih moderat dibandingkan sektor usaha besar. Di sisi lain, sektor usaha besar menunjukkan fluktuasi yang cukup tajam dalam pertumbuhannya, dengan penurunan signifikan pada tahun 2006 yang tercatat sebesar -10.22%. Penurunan ini yang mencerminkan ketidakstabilan sektor usaha besar, yang mungkin dipengaruhi oleh kondisi ekonomi makro maupun perubahan kebijakan yang berdampak pada sektor industri besar.

Selain itu, rasio tenaga kerja di sektor usaha besar terhadap UMKM ini menunjukkan tren penurunan signifikan, dari 3.25% pada tahun 2005 menjadi 2.78% pada tahun 2009. Penurunan rasio ini menunjukkan adanya pergeseran dominasi dalam struktur ketenagakerjaan nasional, di mana sektor UMKM semakin mendominasi dalam penyediaan lapangan pekerjaan dibandingkan sektor usaha besar. Hal ini yang mencerminkan perubahan signifikan dalam pola ketenagakerjaan di Indonesia, di mana sektor UMKM, meskipun memiliki laju pertumbuhan yang lebih lambat, cenderung lebih stabil dan mampu menyerap tenaga kerja lebih banyak dibandingkan sektor usaha besar. Grafik tren pertumbuhan ini dapat dilihat pada Gambar 3, menggambarkan perbandingan tren pertumbuhan tenaga kerja antara kedua sektor tersebut.



Gambar 3. Grafik Tren Pertumbuhan Tenaga Kerja UMKM dan Usaha Besar

Temuan ini menunjukkan sektor UMKM semakin memainkan peran penting dalam struktur ketenagakerjaan Indonesia. Meskipun pertumbuhannya lebih stabil namun moderat, UMKM lebih mampu menyerap tenaga kerja dalam jumlah yang lebih besar, terutama ketika dibandingkan dengan sektor usaha besar yang cenderung mengalami fluktuasi. Penurunan rasio tenaga kerja sektor usaha besar terhadap UMKM menandakan berkurangnya dominasi sektor besar dalam pasar kerja, yang mungkin mencerminkan adanya ketidakpastian ekonomi atau pengurangan jumlah tenaga kerja yang diperlukan perusahaan besar. Hal ini menunjukkan pentingnya sektor UMKM dalam menyediakan lapangan pekerjaan yang lebih stabil dan berkelanjutan, yang seharusnya menjadi fokus dalam kebijakan pengembangan ketenagakerjaan di Indonesia.

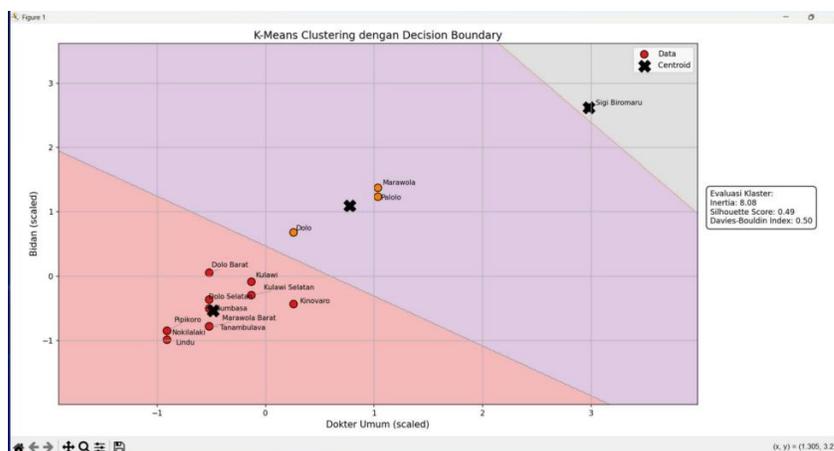
4. Klasterisasi Data Pegawai Berdasarkan Rasio Gender dan Total Pegawai

Analisis ini yang dilakukan untuk mengelompokkan wilayah kecamatan berdasarkan jumlah tenaga medis, yaitu dokter umum dan bidan, dengan menggunakan algoritma *K-Means*. Sebelum dilakukan klasterisasi, data yang digunakan telah melalui standardisasi (*scaling*) agar perbandingan antar fitur menjadi lebih seimbang, sehingga hasil klasterisasi lebih akurat. Algoritma *K-Means* dijalankan dengan parameter jumlah klaster (*n_clusters*) yang ditetapkan sebanyak 3 (tiga). Hasil klasterisasi ini menunjukkan bahwa kecamatan-kecamatan terbagi ke dalam tiga kelompok yang memiliki karakteristik berbeda, tergantung pada jumlah dokter umum dan bidan yang tersedia di masing-masing kecamatan.

Hasil klasterisasi ini divisualisasikan dalam bentuk *scatter plot* dua dimensi, di mana sumbu X mewakili jumlah dokter umum, sumbu Y mewakili jumlah bidan, dan warna titik menunjukkan klaster masing-masing kecamatan. Selain itu, tanda silang besar (X) menandakan posisi *centroid* atau pusat dari tiap klaster, sedangkan warna *background* pada plot ini menggambarkan *decision boundary*, yaitu batas wilayah keputusan antar klaster. Visualisasi ini memberikan gambaran yang jelas mengenai distribusi tenaga medis di setiap kecamatan dan memungkinkan untuk melihat pola distribusi yang ada secara lebih mudah.

Salah satu temuan menarik dalam analisis klasterisasi ini adalah munculnya Klaster 2, yang hanya diisi oleh satu kecamatan, yaitu Sigi Biromaru (Katharina dkk., 2016). Hal ini menunjukkan bahwa kecamatan tersebut memiliki pola jumlah tenaga medis yang sangat berbeda dibandingkan dengan kecamatan lainnya, sehingga secara statistik dianggap sebagai kelompok yang terpisah. Temuan ini bisa jadi menunjukkan adanya kebutuhan atau ketidakmerataan distribusi tenaga medis di wilayah tersebut yang membutuhkan perhatian khusus dalam kebijakan

pengelolaan sumber daya medis. Visualisasi lengkap dari hasil klasterisasi ini dapat dilihat pada Gambar 4, yang memberikan representasi visual dari distribusi tenaga medis antar kecamatan.



Gambar 4. Klasterisasi Berdasarkan Rasio Gender dan Jumlah Pegawai

Temuan ini menunjukkan adanya ketidakmerataan distribusi tenaga medis antar kecamatan, dengan satu kecamatan yang menonjol karena perbedaan yang signifikan dalam jumlah tenaga medisnya. Klasterisasi ini membantu mengidentifikasi wilayah yang memiliki pola-pola distribusi tenaga medis yang serupa, serta wilayah yang memiliki kebutuhan yang berbeda atau tidak merata, seperti yang terlihat pada kecamatan Sigi Biromaru. Hal ini yang memberikan wawasan penting bagi pengambil kebijakan dalam merencanakan distribusi tenaga medis agar lebih merata dan sesuai dengan kebutuhan di tiap daerah. Sebagai contoh, kecamatan seperti Sigi Biromaru mungkin memerlukan intervensi khusus dalam hal penempatan atau redistribusi tenaga medis untuk mampu meningkatkan kualitas pelayanan kesehatan di wilayah tersebut.

Evaluasi klasterisasi menggunakan *Davies-Bouldin Index* (DBI) menunjukkan nilai sebesar 0.51, yang cukup tinggi dan mengindikasikan bahwa kualitas pemisahan antar klaster masih kurang optimal. Nilai ini menunjukkan bahwa klaster yang terbentuk belum sepenuhnya terpisah dengan jelas, yang dapat disebabkan oleh beberapa faktor, seperti ukuran *dataset* yang relatif kecil dan kurangnya jumlah fitur pendukung yang digunakan dalam analisis. Meskipun demikian, hasil dari klasterisasi ini tetap memberikan wawasan awal yang berguna dalam memahami distribusi tenaga kerja dan dapat dijadikan sebagai dasar untuk analisis lebih lanjut dengan memperbaiki aspek-aspek teknis yang ada.

Berdasarkan analisis yang dilakukan, bahwa pertama, pola distribusi pegawai menunjukkan struktur hierarki organisasi, di mana semakin tinggi pangkat, semakin sedikit jumlah pegawai yang mendudukinya. Kedua, ketimpangan gender masih terlihat baik dalam hal pendapatan maupun proporsi tenaga kerja, yang menunjukkan adanya kesenjangan yang perlu diperhatikan dalam kebijakan ketenagakerjaan. Ketiga, sektor UMKM memainkan peran dominan dalam menyerap tenaga kerja nasional, meskipun pertumbuhan yang stabil namun lebih moderat dibandingkan sektor usaha besar yang lebih fluktuatif. Keempat, hasil evaluasi klasterisasi menunjukkan bahwa masih ada ruang untuk perbaikan dalam hal data dan metode yang digunakan agar klasterisasi yang dihasilkan lebih valid dan representatif. Terakhir, klasterisasi dan analisis tren ini dapat dijadikan dasar untuk merumuskan rekomendasi kebijakan yang lebih baik dalam pemerataan sumber daya manusia, pengembangan pelatihan kerja, dan penyusunan strategi pembangunan ekonomi berbasis data yang lebih efektif.

KESIMPULAN

Bahwa terdapat pola distribusi tenaga kerja yang mencerminkan struktur organisasi yang berbentuk piramidal, di mana semakin tinggi pangkat jabatan, semakin sedikit jumlah pegawai yang mendudukinya. Selain itu, penelitian ini juga mengungkapkan adanya ketimpangan gender yang signifikan dalam hal pendapatan dan distribusi tenaga kerja, dengan wanita masih kurang terwakili di beberapa sektor dan jabatan bergaji tinggi. Di sisi lain, sektor UMKM mendominasi struktur dari ketenagakerjaan nasional, dengan pertumbuhan yang stabil meskipun lebih moderat dibandingkan sektor usaha besar. Penelitian lebih lanjut sebaiknya dilakukan dengan menggunakan dataset yang lebih besar dan lebih bervariasi, agar hasil yang diperoleh lebih representatif dan dapat mencakup aspek ketenagakerjaan yang lebih luas. Selain itu, eksplorasi metode klusterisasi lain seperti DBSCAN atau *Hierarchical Clustering* dapat memberikan perspektif yang berbeda dan lebih akurat dalam mengelompokkan data tenaga kerja. Untuk ke depannya, pemerintah dapat menggunakan hasil analisis ini sebagai dasar untuk mendukung kebijakan-kebijakan yang lebih baik dalam pengelolaan sumber daya manusia (SDM), seperti dalam hal pelatihan kerja, upaya pemerataan kesempatan kerja, dan pengurangan kesenjangan gender di tempat kerja.

DAFTAR REFERENSI

- Anjarwati, D. (2025). Clustering Penindakan Kasus Korupsi di Indonesia Menggunakan K-Means: Analisis Data KPK Tahun 2004-2024. *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, 12(2). <https://doi.org/10.35957/jatisi.v12i2.11375>
- Arhami, M., & Nasir, M. (2020). *Data mining - Algoritma dan implementasi*. Penerbit Andi.
- Azzam, A., Purnamasari, A. I., & Ali, I. (2024). Implementasi Algoritma K-Means Clustering untuk Analisis Persebaran UMKM di Jawa Barat. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(3), 3062-3070. <https://doi.org/10.36040/jati.v8i3.8450>
- Harani, N. H., Prianto, C., & Nugraha, F. A. (2020). Segmentasi pelanggan produk digital service Indihome menggunakan algoritma K-Means berbasis Python. *Jurnal Manajemen Informatika (JAMIKA)*, 10(2), 133-146. <https://doi.org/10.34010/jamika.v10i2.2683>
- Harjanto, T., & Ariyani, R. M. (2019). Distribusi Pendapatan di Indonesia. *Cendekia Jaya*, 1(2), 29-41. <https://doi.org/10.47685/cendekia-jaya.v1i2.33>
- Katharina, R., Pahlevi, I., Lindyawayaty, D. S., & Sendhikasari, D. (2016). *Evaluasi terhadap pemekaran daerah dan potensi penggabungan daerah (Kasus Kabupaten Sigi dan Kabupaten Tanjung Jabung Timur)*. PT Balai Pustaka (Persero) bekerja sama dengan Pusat Penelitian Badan Keahlian DPR RI. Kusuma, A. W., & Ellyana, R. L. (2018). Penerapan Citra Terkompresi pada Segmentasi Citra Menggunakan Algoritma K-Means. *Jurnal Terapan Teknologi Informasi*, 2(1), 65-74. <https://doi.org/10.21460/jutei.2018.21.65>
- Mardiani, E., Rahmansyah, N., Handayani, E. T. E., Hidayatullah, D., Wulandana, N. P., Lombu, A. A., & Budyarti, S. (2023). Analisis Kinerja Tren Penjualan untuk Mendapatkan Strategi Penjualan secara Global Menggunakan Tableau Data Mining. *Innovative: Journal of Social Science Research*, 3(3), 2238-2248. <https://j-innovative.org/index.php/Innovative/article/view/2175>
- Maulidhia, A. N. A., Widyastuti, I. I., Sukarno, F. I., Tsany, R. B. S., & Brian, T. (2025). Implementasi Perbandingan Algoritma k-Means dan DB-Scan pada Beban Listrik Rumah Tangga. *INTEGER: Journal of Information Technology*, 10(1). <https://doi.org/10.31284/j.integer.2024.v10i1.7479>

- Nugraha, E. R., Amin, M. N., Rachmawati, S., & Sugiyarti, L. (2022). Pemberdayaan Pengusaha UMKM Bidang Kuliner di Jabodetabek dalam Pengelolaan Cash Flow dan Perpajakannya. *Jurnal Abdikaryasakti*, 2(1), 13-24.
<https://doi.org/10.25105/ja.v2i1.12805>
- Rousseuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53–65.
[https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)
- Setiawan, Z., Fajar, M., Priyatno, A. M., Perdana Putri, A. Y., Aryuni, M., Yuliyanti, S., Widiputra, H., Meilani, B. D., Ibrahim, R. N., Azdy, R. A., Junaidi, S., & Wijaya, A. (2023). *Buku ajar data mining*. PT. Sonpedia Publishing Indonesia.