

Penerapan Data Mining dalam Mengelompokkan Jumlah UMKM Kota Jambi Menggunakan *K-Means Clustering*

Muhammad Aji Triatama¹, Jasmir², Benni Purnama³

Fakultas Ilmu Komputer, Magister Sistem Informasi, Universitas Dinamika Bangsa, Jambi, Indonesia

Email: ¹ajitriatama@gmail.com, ²ijay_jasmir@yahoo.com, ³bennipurnama@stikom-db.ac.id

Email Penulis Korespondensi: ajitriatama@gmail.com

Submitted :
14 April 2025

Revision :
24 Juni 2025

Accepted:
25 September 2025

Published:
30 September 2025

Abstrak—Usaha Mikro, Kecil, dan Menengah (UMKM) memiliki kontribusi yang signifikan dalam mendukung perkembangan ekonomi di tingkat lokal, termasuk di Kota Jambi. Namun, banyak UMKM menghadapi kendala dalam pengelolaan data dan perumusan strategi bisnis yang efektif. Penelitian ini bertujuan untuk mengelompokkan UMKM di Kota Jambi dengan memanfaatkan metode *K-Means Clustering* sebagai bagian dari teknik data mining. Data penelitian mencakup 1.331 UMKM di bidang *fashion*, dengan variabel karakteristik seperti kepemilikan, modal awal, penghasilan per bulan dan per tahun, jumlah karyawan, serta aset. Hasil penelitian menunjukkan bahwa perhitungan manual menghasilkan tiga kelompok: 627 UMKM (*Cluster 1*) direkomendasikan menerima bantuan peralatan, 358 UMKM (*Cluster 2*) mendapatkan fasilitas kelembagaan, dan 346 UMKM (*Cluster 3*) diarahkan untuk pelatihan. Sementara itu, perhitungan dengan *tools Rapid Miner* menghasilkan pembagian berbeda akibat variasi pengambilan *centroid*. *Cluster 0* berisi 840 UMKM, *Cluster 1* berisi 9 UMKM, dan *Cluster 2* berisi 482 UMKM. Penelitian ini membuktikan bahwa *K-Means Clustering* dapat memberikan wawasan berharga tentang distribusi dan kebutuhan UMKM, sehingga membantu pengambil kebijakan dalam merancang program yang lebih tepat sasaran. Dengan demikian, metode ini diharapkan dapat berkontribusi pada pengembangan UMKM secara lebih terarah dan berkelanjutan di Kota Jambi.

Kata Kunci: Data Mining; *K-Means Clustering*; Analisis Data; Pengelompokan UMKM; *Rapid Miner*

Abstract—*Usaha Mikro, Kecil, dan Menengah (UMKM) plays a significant role in supporting local economic development, including in Kota Jambi. However, many UMKM face challenges in data management and the formulation of effective business strategies. This study aims to cluster UMKM in Kota Jambi using the K-Means Clustering method as part of data mining techniques. The research data includes 1,331 UMKM in the fashion sector, with characteristics such as ownership, initial capital, monthly and annual income, number of employees, and assets. The results show that manual calculation produced three clusters: 627 UMKM (Cluster 1) are recommended to receive equipment assistance, 358 UMKM (Cluster 2) should receive institutional facilities, and 346 UMKM (Cluster 3) should be directed for training. Meanwhile, calculations using the Rapid Miner tools resulted in a different division due to variations in centroid selection. Cluster 0 contains 840 UMKM, Cluster 1 contains 9 UMKM, and Cluster 2 contains 482 UMKM. This study proves that K-Means Clustering can provide valuable insights into the distribution and needs of UMKM, thus helping policymakers design more targeted programs. Therefore, this method is expected to contribute to the more focused and sustainable development of UMKM in Kota Jambi.*

Keywords: Data Mining; *K-Means Clustering*; Data Analysis; UMKM Clustering; *Rapid Miner*

1. PENDAHULUAN

Dalam era informasi digital saat ini, data menjadi salah satu aset terpenting yang dapat memengaruhi keputusan strategis di berbagai bidang. Penggunaan teknik data mining, yang berfokus pada penambangan informasi berharga dari kumpulan data yang besar, telah berkembang pesat karena kemampuannya untuk mengidentifikasi pola-pola tersembunyi dan memberikan wawasan yang mendalam [1]. Data mining adalah suatu metode untuk menyelesaikan permasalahan dengan menganalisis data yang tersimpan dalam *database* secara elektronik. Proses ini melibatkan pencarian dan pengekstrakan informasi dari data yang ada secara otomatis, biasanya menggunakan perangkat komputer [2]. Di sektor ekonomi, terutama bagi usaha mikro, kecil, dan menengah (UMKM) yang sering kali menjadi tulang punggung dari perekonomian lokal, data mining menawarkan peluang besar untuk meningkatkan efektivitas pengelolaan dan pengembangan bisnis [3].

UMKM adalah jenis usaha kecil yang memiliki peran penting dalam meningkatkan dan mendorong pertumbuhan ekonomi masyarakat [4]. UMKM memiliki ketahanan yang kuat dalam berbagai situasi, sehingga dapat berkontribusi pada kesejahteraan masyarakat. UMKM memainkan peran yang sangat penting dalam perekonomian di Kota Jambi, karena berkontribusi terhadap pertumbuhan ekonomi dan penciptaan lapangan pekerjaan. Namun, banyak UMKM menghadapi tantangan dalam hal pengelolaan data dan pemahaman pasar yang lebih luas. Tanpa adanya analisis yang tepat, banyak potensi bisnis yang tidak dapat dioptimalkan secara maksimal. Pengelompokan UMKM berdasarkan karakteristik tertentu dapat memberikan gambaran yang lebih jelas tentang segmen-semen pasar dan kebutuhan spesifik dari berbagai kelompok UMKM [5].

K-Means Clustering sebagai teknik perhitungan dalam data mining dapat menawarkan solusi yang efektif [6]. *Clustering* merupakan pendekatan dalam data mining yang digunakan untuk mengelompokkan data berdasarkan kesamaan atau kedekatan antara data yang ada. Teknik ini dikenal dengan sifat tanpa arahan (*unsupervised*), yang berarti bahwa *clustering* dapat diterapkan tanpa memerlukan data pelatihan, label, atau *output* target [7]. *K-Means Clustering* adalah teknik pengelompokan yang dapat mengelompokkan suatu data ke dalam beberapa kelompok yang memiliki kesamaan karakteristik. Dengan menerapkan metode ini, kita dapat mengidentifikasi kelompok-kelompok UMKM di Kota Jambi berdasarkan berbagai variabel. Hal ini memungkinkan pengambil kebijakan dan pemangku kepentingan untuk merancang strategi yang lebih tepat sasaran dan efisien, serta untuk menyediakan dukungan yang lebih sesuai dengan kebutuhan setiap kelompok UMKM [8].

Algoritma *K-Means Clustering* adalah teknik yang sederhana namun efisien dalam menemukan kluster-kluster dalam data. Proses algoritma ini terdiri dari beberapa langkah, yaitu langkah pertama menentukan jumlah kluster yang diinginkan, langkah kedua menentukan posisi pusat kluster awal, langkah ketiga menghitung pusat kluster terdekat untuk setiap data *point*, langkah keempat untuk setiap kluster, menghitung *centroid* dan memperbarui posisi pusat kluster tersebut, dan langkah kelima mengulangi proses dari langkah ketiga hingga kelima hingga mencapai konvergensi atau penghentian, di mana posisi pusat kluster tidak lagi berubah secara signifikan [9].

Beberapa penelitian sejenis menunjukkan bahwa penerapan metode pengelompokan dalam konteks UMKM dapat memberikan wawasan yang lebih mendalam mengenai dinamika pasar. Penelitian oleh Ahmad Azzam Arrosyad dkk. [10] menganalisis persebaran UMKM di Jawa Barat menggunakan *K-Means Clustering*, mengelompokkan wilayah menjadi tiga kluster yaitu rendah, sedang, dan tinggi. Hasil penelitian ini membantu merumuskan kebijakan ekonomi dan strategi pengembangan UMKM di wilayah tersebut.

Penelitian yang dilakukan oleh Iin dkk. [11] menggunakan *K-Means Clustering* untuk mengelompokkan jumlah UMKM berdasarkan kabupaten/kota di Jawa Barat. Penelitian ini menemukan dua *cluster*, satu dengan jumlah UMKM tinggi dan satu lagi dengan jumlah sedang. Hasil ini dapat mendukung pengambilan keputusan terkait pengembangan UMKM di wilayah tersebut.

Selanjutnya, penelitian yang dilakukan oleh Eko Nur Wahyudi dkk. [12] menerapkan *K-Means* untuk memetakan wilayah di Semarang berdasarkan potensi industri mikro, kecil, dan menengah. Fokus penelitian ini adalah pembinaan dan pendampingan usaha berdasarkan karakteristik wilayah, berbeda dengan penelitian lainnya yang lebih berfokus pada distribusi jumlah UMKM di berbagai daerah.

Penelitian lain yang dilakukan oleh Hafshoh Habiballoh dkk. [13] mengelompokkan wilayah potensi objek daya tarik wisata di Jawa Barat menggunakan *K-Means*. Hasilnya mengidentifikasi tiga klaster utama yang dapat membantu dalam pengembangan kebijakan pariwisata, meskipun tidak berfokus langsung pada UMKM.

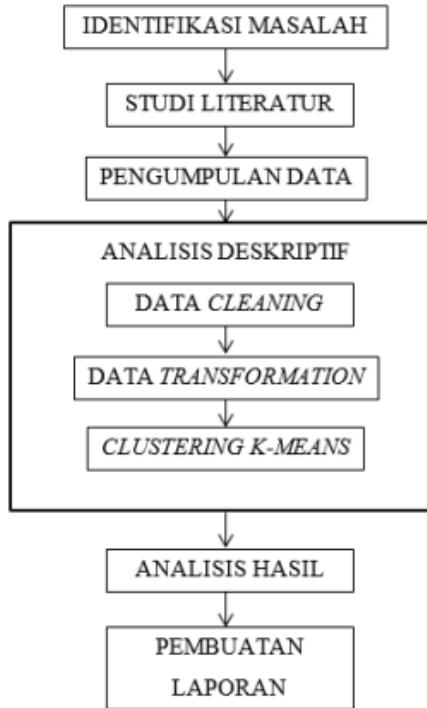
Penelitian yang dilakukan oleh Hendrik dkk. [14] menggunakan *K-Means* dan metode Elbow untuk mengklaster UMKM di Jawa Barat. Hasil penelitian ini menghasilkan lima kluster yang memberikan gambaran lebih rinci tentang keragaman jenis usaha, yang membantu dalam merencanakan strategi pengembangan UMKM secara lebih tepat.

Berdasarkan beberapa penelitian sejenis yang telah disebutkan, penulis menyimpulkan bahwa penerapan metode *K-Means Clustering* untuk mengelompokkan UMKM di Kota Jambi diharapkan dapat memberikan wawasan yang lebih mendalam mengenai distribusi dan potensi bisnis di wilayah tersebut. Berdasarkan permasalahan yang telah diuraikan, penelitian ini bertujuan untuk mengimplementasikan metode *K-Means Clustering* dalam mengelompokkan UMKM di Kota Jambi serta menganalisis hasil *clustering* yang diperoleh melalui algoritma tersebut. Penelitian ini diharapkan dapat memberikan manfaat yaitu pertama, memungkinkan pemerintah dan lembaga terkait merancang kebijakan serta program dukungan yang lebih spesifik dan sesuai dengan kebutuhan masing-masing kelompok UMKM. Kedua, mempermudah alokasi bantuan dan sumber daya dengan mengetahui kebutuhan spesifik dari setiap kelompok UMKM, sehingga dukungan yang diberikan lebih tepat sasaran. Ketiga, penelitian ini diharapkan dapat menjadi referensi bagi peneliti lain sebagai sumber informasi, perbandingan, dan acuan dalam bidang penelitian yang serupa.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Tahapan dalam penelitian ini akan dilakukan secara berkelanjutan hingga tujuan penelitian tercapai. Kerangka kerja penelitian digunakan untuk mempermudah penyusunan dan pelaksanaan penelitian. Kerangka kerja penelitian dapat dilihat pada gambar berikut :



Gambar 1. Tahapan Penelitian

Berdasarkan tahapan penelitian yang telah digambarkan pada gambar 1, maka dapat diuraikan pembahasan dari masing-masing beberapa tahapan penelitian yaitu :

1. Identifikasi Masalah

Dalam konteks ini, diharapkan dapat diperoleh analisis yang tepat mengenai permasalahan, yaitu cara mengidentifikasi kelompok UMKM di Kota Jambi sesuai dengan data yang ada, sehingga menghindari kesalahan dalam penentuan kelompok UMKM. Berdasarkan masalah tersebut, penulis mencari solusi yang relevan.

2. Studi Literatur

Studi literatur adalah proses mencari referensi teori yang relevan dengan kasus atau permasalahan yang dihadapi. Pada tahap ini, dilakukan pencarian sumber teori dari berbagai buku dan internet untuk memperkaya pemahaman konsep dan teori, sehingga dapat membangun landasan ilmiah yang kuat dan sesuai.

3. Pengumpulan Data

Pada langkah ini, penulis mengumpulkan informasi dan data yang diperlukan untuk penelitian. Proses pengumpulan data dengan menggunakan berbagai teknik, yaitu :

a. Observasi

Teknik ini melibatkan pengamatan langsung terhadap objek yang diteliti. Penulis mengamati data UMKM di Kota Jambi untuk memperoleh salinan data yang diperlukan dalam penelitian ini.

b. Wawancara

Dalam pendekatan ini, penulis melakukan wawancara langsung dengan pihak terkait di Kota Jambi untuk mendapatkan informasi dan data yang lebih tepat mengenai data UMKM.

c. Penelitian Kepustakaan (*Library Research*)

Untuk mendukung penelitian yang akan dilakukan, penulis melakukan *library research* dengan mencari data dari buku dan jurnal penelitian yang relevan, terutama yang berkaitan dengan metode yang digunakan dalam penelitian ini, yaitu metode *K-Means*.

4. Analisis Deskriptif

Analisis Deskriptif yaitu menganalisa sebuah data yang dilakukan untuk menilai sebuah karakteristik dari beberapa data yang sedang di proses. Berikut penjelasan alur proses dari analisis deskriptif :

a. *Data Cleaning*

Data cleaning merupakan operasi dasar yang dilakukan seperti penghapusan *noise*. Proses *cleaning* ialah membuang duplikasi data, memeriksa data yang inkonsisten atau ketidakserasian, dan memperbaiki kesalahan pada sebuah data, seperti kesalahan cetak dan lain sebagainya [15]. Data bisa diperkaya dengan suatu informasi ekternal yang relevan. Pada tahap ini peneliti melakukan persiapan data UMKM, setelah itu peneliti menyeleksi atribut yang berpengaruh dalam proses *Clustering*, guna menentukan *cluster* pada UMKM di KOTA JAMBI.

b. *Data Transformation*

Data Transformation adalah proses pengolahan data yang telah dipilih agar sesuai untuk analisis data mining. Proses ini sangat bergantung pada tipe atau pola informasi yang ingin dicari dalam basis data [16].

c. *Clustering* Algoritma *K-Means*

Pada langkah ini, dilakukan suatu proses pengklasteran dan dibantu dengan menggunakan alat bantu (*tools*) yaitu *Rapid Miner* terhadap data UMKM [17].

5. Evaluasi dan Analisis Hasil

Pada langkah ini peneliti menganalisis data UMKM pada Kota Jambi untuk dinilai keakuratan data model yang telah didapatkan dengan perhitungan *Clustering K-Means*. Hasil yang telah didapatkan berupa *cluster* dalam menentukan Penerima bantuan untuk kelompok UMKM di Kota Jambi

6. Pembuatan Laporan

Pembuatan laporan bertujuan sebagai dokumentasi dari penelitian yang telah dilakukan. Setelah semua tahapan penelitian telah dilakukan, maka laporan tersebut akan menjadi hasil penting dari penelitian agar dapat dimanfaatkan pada waktu yang akan datang.

2.2 Bahan Penelitian

Adapun bahan yang digunakan penulis dalam penelitian ini menggunakan data UMKM yang berlokasi di Kota Jambi, khususnya yang beroperasi di bidang *fashion*, sebagai bahan utama dalam analisis. Data ini mencakup berbagai aspek yang berkaitan dengan usaha kecil dan menengah dalam industri *fashion*, termasuk faktor-faktor lain yang relevan dengan tujuan penelitian.

3. HASIL DAN PEMBAHASAN

3.1 Dataset

Data yang digunakan penelitian ini adalah data terkait penduduk dan UMKM di Kota Jambi dari Dinas Tenaga Kerja, Koperasi, dan Usaha Kecil Menengah, yang mencatat sebanyak 1331 data UMKM. Data ini memberikan gambaran tentang jumlah, jenis usaha, serta perkembangan sektor UMKM di kota ini. Informasi tersebut sangat penting untuk menganalisis kondisi ekonomi lokal, tantangan yang dihadapi pelaku UMKM, dan potensi yang bisa dikembangkan untuk mendukung kemajuan UMKM di Kota Jambi.

Tabel 1. Variabel Data

No	Nama Variabel/Atribut
1	NIK
2	Nama
3	Alamat
4	Usaha
5	Tahun Berdiri
6	Kepemilikan
7	Modal Awal
8	Penghasilan Perbulan
9	Penghasilan Pertahun
10	Jumlah Karyawan
11	Aset

NO	NIK	NAMA	ALAMAT	USAHA	TAHUN BERDIRI	KEPEMILIKAN	MODAL AWAL	PENGHASILAN PERBULAN	PEN PE
1	15710XXX	A. AZIZ	Rt. 25 Kel. Talang B	Jual Pakaiian	2001	Keluarga	Rp 53.619.000	Rp 7.900.000	Rp
2	15710XXX	A. Khusairi	Jl. Kh. Hasan Anang	Batik	2018	Kemitraan	Rp 28.272.000	Rp 8.390.000	Rp
3	15710XXX	A. RAHMAN DWI	JL. UTAMA RT.009	JUALAN PAKAIA	2005	Keluarga	Rp 20.000.000	Rp 8.960.000	Rp
4	15710XXX	A. agus	RT.07 Kel.Mudung I	Jual Baju BJ	2000	Pribadi	Rp 15.416.000	Rp 6.670.000	Rp
5	15710XXX	A.FAIZAL WAHYU	MAYANG MANGU	JUAL PAKAIA	2023	Kemitraan	Rp 25.189.000	Rp 7.360.000	Rp
6	15710XXX	Abdul Kudus	Rt. 08 Kel. Olak Kem	Pengrajin Batik	2015	Keluarga	Rp 53.558.000	Rp 7.890.000	Rp
7	15710XXX	Abdullah Kasihani	Rt.01 Kel. Ulu Gedor	Batik Jambi	2004	Pribadi	Rp 13.760.000	Rp 5.840.000	Rp
8	15710XXX	Absiah	Rt.08 Kel. Ulu Gedor	Batik Jambi	2005	Pribadi	Rp 16.607.000	Rp 7.270.000	Rp
9	15710XXX	ACHMAD NASRY	JL. KH. A. Tomi RT	Albaita Syar'i C	2019	Kemitraan	Rp 16.369.000	Rp 7.150.000	Rp
10	15710XXX	ACHMAD RIFA'I H.	JL. ADITYAWARM	FASHION	2017	Kemitraan	Rp 16.403.000	Rp 7.160.000	Rp
11	15710XXX	ADA LAILA	RT.09 KEL. ARAB	PENGRAJIN BAJU	2016	Kemitraan	Rp 13.013.000	Rp 5.470.000	Rp
12	15710XXX	ADAM	Perum. Bougenville	Menjual Pakaiian	2013	Kemitraan	Rp 21.202.000	Rp 6.030.000	Rp
13	15710XXX	ADAM YOSEPH	JL. RADEN FATAH	JUAL SEPATU	2017	Kemitraan	Rp 15.051.000	Rp 6.490.000	Rp
14	15710XXX	ADAM YOSEPH	JLN. RADEN FATAH	DAGANG SEPATU	2019	Keluarga	Rp 25.331.000	Rp 7.410.000	Rp
15		ADE PUTRI	ALAM BARAJO	Fashion	2018	Kemitraan	Rp 32.664.000	Rp 7.130.000	Rp
16	15710XXX	ADE ULI KURNIATI	Jl. Bhayangkara	Rt. Jual Baju	2003	Keluarga	Rp 15.450.000	Rp 6.690.000	Rp
17	15710XXX	ADEL MUSKAR	RT20 KEL PASIR PU	JUAL BELI BAJU	2002	Kemitraan	Rp 19.942.000	Rp 8.930.000	Rp
18	15710XXX	ADELIA	RT 40 KEL LEGOK	JUAL TAS	2019	Keluarga	Rp 41.666.000	Rp 7.300.000	Rp
19	15710XXX	aden simanjuntak	kenali asam atas	online shop kos	2022	Pribadi	Rp 40.620.000	Rp 7.090.000	Rp
20	15710XXX	ADRISON	RT. 29 KEL. TAMBAH	DAGANG PAK	2013	Keluarga	Rp 39.405.000	Rp 8.810.000	Rp
21	15710XXX	AFLIANTI	JL. RAJA YAMIN	JUALAN KOSMETIK	2004	Pribadi	Rp 16.887.000	Rp 7.410.000	Rp
22	15710XXX	AFRI ALDI	JLN. GADO-GADO	USAHA KACA	2018	Pribadi	Rp 19.617.000	Rp 8.770.000	Rp
23	15710XXX	Agung Darmanto	Jl.Kol. Amir Hamzah	Baju Gamis	2009	Keluarga	Rp 23.088.000	Rp 6.660.000	Rp
24	15710XXX	AGUNG WIBOWO	PASIR PUTIH	BUSANA	2005	Kemitraan	Rp 27.706.000	Rp 8.200.000	Rp
25	15710XXX	AGUS MULYADI	SELAMAT RT. 027	JUAL PAKAIA	2000	Pribadi	Rp 19.453.000	Rp 8.690.000	Rp
26	15710XXX	AGUS PEBRIYANTI	KELURAHAN KEN	JUAL BAJU	2004	Kemitraan	Rp 40.619.000	Rp 7.090.000	Rp
27	15710XXX	AGUS SURAPATI	SELAMAT RT. 019	JUAL PAKAIA	2006	Kemitraan	Rp 19.573.000	Rp 8.750.000	Rp
28	15710XXX	AGUS WATI	JL. SOEMANTRI	BR. ACCESSORIES	2009	Keluarga	Rp 12.965.000	Rp 5.450.000	Rp
29	15710XXX	AGUSTINI	RT. 23 KEL. LEGOK	BATIK	2019	Kemitraan	Rp 14.490.000	Rp 6.210.000	Rp

Gambar 2. Data Mentah

3.2 Data Selection

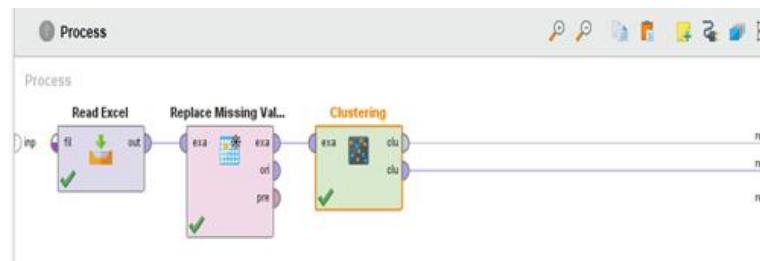
Selanjutnya, dilakukan proses pembersihan data (*data cleaning*) untuk menghilangkan data yang mengandung *noise*, seperti redundansi dan inkonsistensi. Setelah proses pembersihan data, diperoleh sebanyak 1331 data. Data penduduk ini menggunakan enam atribut yang dianggap paling berpengaruh dalam penentuan, yaitu Kepemilikan, Modal Awal, Penghasilan Per Bulan, Penghasilan Per Tahun, Jumlah Karyawan, dan Aset yang dimiliki. Data yang telah dibersihkan ini akan digunakan untuk perhitungan *Clustering K-Means*, seperti terlihat pada tabel berikut :

Tabel 2. Data Perhitungan *K-Means*

No	Nama	Kepemilikan	Modal Awal	Penghasilan Perbulan	Penghasilan Pertahun	Jumlah Karyawan	Aset
1	A. Aziz	1	53619000	7900000	94800000	4	94738000
2	A. Khusairi	2	28272000	8390000	100680000	3	100588000
3	A. R. D. Saputra	1	20000000	8960000	107520000	2	107500000
4	A. Agus	3	15416000	6670000	80040000	2	79996000
5	A. F. Wahyu S.	2	25189000	7360000	88320000	3	88256000
6	Abdul Kudus	1	53558000	7890000	94680000	4	94616000
7	A. Kasihan	3	13760000	5840000	70080000	2	70060000
8	Absiah	3	16607000	7270000	87240000	2	87142000
9	A. Nasry	2	16369000	7150000	85800000	2	85714000
10	A. Rifa'i H. M.	2	16403000	7160000	85920000	2	85918000
11	A. Laila	2	13013000	5470000	65640000	2	65578000
12	Adam	2	21202000	6030000	72360000	3	72308000
13	A. Yosefh	2	15051000	6490000	77880000	2	77806000
14	A. Yosefh	1	25331000	7410000	88920000	3	88824000
15	Ade Putri	2	32664000	7130000	85560000	4	85492000
16	A. Uli K. S.	1	15450000	6690000	80280000	2	80200000
17	Adel Muskar	2	19942000	8930000	107160000	2	107152000
18	Adelia	1	41666000	7300000	87600000	5	87498400
19	A. Simanjuntak	3	40620000	7090000	85080000	5	84988000
20	Adrison	1	39405000	8810000	105720000	4	105715000
...
1331	Zuraida	2	18320000	8120000	97440000	2	97420000

3.3 Data Preprocessing

Visualisasi proses metode *K-Means* pada tools *Rapid Miner* dapat dilihat pada gambar berikut :



Gambar 3. Preprocessing

Pada gambar 3, dijelaskan proses *clustering K-Means* menggunakan *Rapid Miner* melalui beberapa langkah yaitu pertama pilih import data untuk memilih data yang akan diproses kemudian cari data di My Computer dan lanjutkan hingga finish. Selanjutnya pilih list Local Repository pada kolom Repository dan tarik data yang telah diimpor ke kolom process. Pada kolom Operators cari metode *K-Means* dan tarik metode tersebut ke kolom process. Kemudian hubungkan proses read excel, replace missing value, dan *clustering* dengan menghubungkan output read excel ke example set di proses *clustering*. Fungsi *cluster* model dihubungkan ke result 1 dan *clustered* set ke result 2. Setelah semua proses terhubung jalankan untuk memulai *clustering*. Jika ditemukan missing value ubah operator dengan memilih replace missing value di folder cleansing lalu sambungkan data retrieve dengan replace missing value dan *clustering*. Setelah itu jalankan kembali proses *clustering*.

3.4 Data dan Pembagian Cluster pada Proses Clustering

Visualisasi letak data dan klasernya pada *Rapid Miner* dapat dilihat pada gambar berikut :

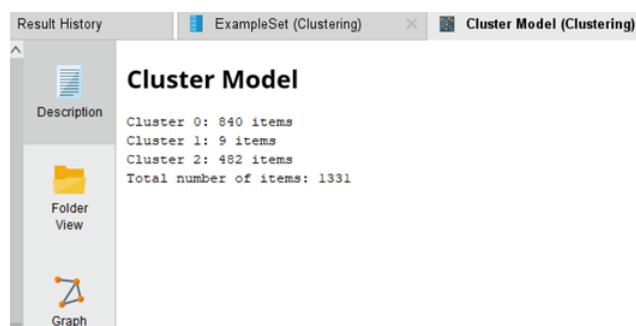
Row No.	NAMA	USAHA	cluster	KEPEMILIKAN	MODAL AWAL	PENGHASIL...	PENGHASIL...	JUMLAH KA...	ASET
1	A. AZIZ	Jual Pakaiian	cluster_0	1	53619000	7900000	94800000	4	94738000
2	A. Khusairi	Batik	cluster_0	2	28272000	8390000	100680000	3	100580000
3	A. RAHMAN D...	JUALAN PAK...	cluster_0	1	2000000	8960000	107520000	2	107500000
4	A.agus	Jual Baju BJ	cluster_2	3	15416000	6670000	80040000	2	79996000
5	A.FAIZAL WA...	JUAL PAKAIAN	cluster_0	2	25189000	7360000	88320000	3	88256000
6	Abdul Kudus	Pengrajin Batik	cluster_0	1	53558000	7890000	94680000	4	94616000
7	Abdullah Kas...	Batik Jambl	cluster_2	3	13760000	5840000	70080000	2	70060000
8	Absiah	Batik Jambl	cluster_2	3	16607000	7270000	87240000	2	87142000
9	ACHMAD NA...	Albalty Syari'	cluster_2	2	16369000	7150000	85800000	2	85714000
10	ACHMAD RIF...	FASHION	cluster_2	2	16403000	7160000	85920000	2	85918000
11	ADA LAILA	PENGRAJIN...	cluster_2	2	13013000	5470000	65540000	2	65578000
12	ADAM	Menjual Pakai...	cluster_2	2	21202000	6030000	72360000	3	72308000
13	ADAM YOSEPH	JUAL SEPATU	cluster_2	2	15051000	6490000	77880000	2	77806000
14	ADAM YOSEPH...	DAGANG SE...	cluster_0	1	25331000	7410000	88920000	3	88824000
15	ADE PUTRI	Fashion	cluster_0	2	32664000	7130000	85560000	4	85492000

Gambar 4. Letak Data Dan Klasernya

Pada gambar 4, menunjukkan hasil dari proses *clustering K-Means* yang dilakukan menggunakan *Rapid Miner*, di mana kolom-kolom muncul bersama dengan pembagian setiap nomor ke dalam *cluster* masing-masing. Sebagai contoh, Nomor 1 berada pada *cluster* 0, sementara Nomor 2 juga tergolong dalam *cluster* 0, dan seterusnya untuk nomor-nomor berikutnya.

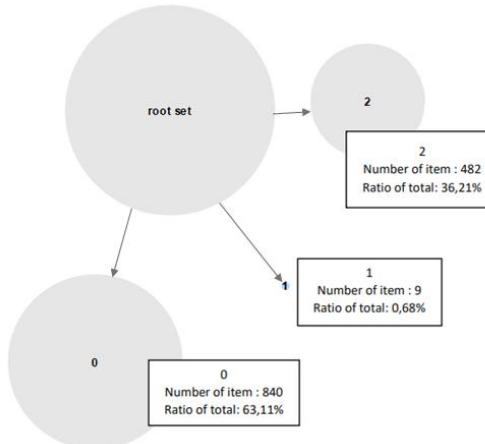
3.5 Evaluasi

Pada tahapan *Evaluation* ini, pada tools *Rapid Miner* dapat menggunakan Operator Performance.



Gambar 5. Cluster Model

Pada gambar 5, dijelaskan mengenai visualisasi *Cluster Model* yang menunjukkan hasil dari proses *clustering* yang dilakukan dengan *Rapid Miner*. Setiap *cluster* memiliki jumlah item yang berbeda, di mana *Cluster 0* memiliki 840 item, *Cluster 1* hanya memiliki 9 item, dan *Cluster 2* memiliki 482 item. Total keseluruhan data yang diolah dalam proses *clustering* ini adalah sebanyak 1331 item. Hasil ini memberikan gambaran mengenai distribusi data dalam masing-masing *cluster*, yang bisa digunakan untuk menganalisis pola atau karakteristik yang terkandung dalam setiap kelompok berdasarkan metode *K-Means* yang diterapkan. Visualisasi ini juga memudahkan untuk mengevaluasi sejauh mana data terdistribusi dan seberapa representatif hasil *cluster* tersebut dalam menggambarkan kelompok data ada.



Gambar 6. Grafik Hasil *Cluster*

Pada gambar 6, terlihat hasil visualisasi grafik *clustering* yang dihasilkan oleh *Rapid Miner*, yang memberikan gambaran jelas mengenai distribusi data pada masing-masing *cluster* beserta rasio dari total data. *Cluster 0*, yang terdiri dari 840 item, memiliki rasio sebesar 63,11% dan dikategorikan sebagai kelompok yang "berhak mendapatkan bantuan peralatan/barang." *Cluster 1*, yang hanya terdiri dari 9 item, memiliki rasio sebesar 0,68% dan dikategorikan sebagai kelompok yang "berhak mendapatkan bantuan fasilitas kelembagaan." Sementara itu, *Cluster 2*, dengan jumlah item sebanyak 482, memiliki rasio sebesar 36,21% dan dikategorikan sebagai kelompok yang "berhak mendapatkan pelatihan." Visualisasi ini membantu dalam memahami bagaimana data terbagi dalam setiap kategori dan rasio setiap *cluster*, yang dapat digunakan untuk perencanaan dan pengambilan keputusan yang lebih terarah terkait pemberian bantuan atau dukungan kepada setiap kelompok.

Attribute	cluster_0	cluster_1	cluster_2
KEPEMILIKAN	1.982	1.667	2.052
MODAL AWAL	38628384.524	-12290777.778	17905558.091
PENGHASILAN PERBULAN	8019214.286	11254444.444	6258132.780
PENGHASILAN PERTAHUN	96230571.429	135053333.333	75097593.361
JUMLAH KARYAWAN	3.777	2	2.442
ASET	96172477.636	134989333.333	75039136.929

Gambar 7. Nilai Setiap Atribut

Pada gambar 6, yang menunjukkan hasil visualisasi nilai dari setiap atribut, terdapat tiga *cluster* yang dihasilkan dari perhitungan menggunakan *Rapid Miner* dari 1331 data UMKM di Kota Jambi. Setiap *cluster* memiliki nilai atribut yang berbeda berdasarkan kategori seperti Kepemilikan, Modal Awal, Penghasilan Perbulan, Penghasilan Pertahun, Jumlah Karyawan, dan Aset. Pada *Cluster 0*, nilai Kepemilikan tercatat sebesar 1.982, Modal Awal sebesar 38.628.384.524, Penghasilan Perbulan sebesar 8.019.214.286, Penghasilan Pertahun sebesar 96.230.571.429, Jumlah Karyawan sebesar 3.777, dan Aset sebesar 96.172.477.636. *Cluster* ini memiliki nilai yang lebih tinggi pada sebagian besar atribut, mencerminkan usaha dengan potensi besar dalam hal modal dan penghasilan.

Pada *Cluster 1*, nilai Kepemilikan tercatat sebesar 1.667, Modal Awal sebesar -12.290.777.778, Penghasilan Perbulan sebesar 11.254.444.444, Penghasilan Pertahun sebesar 135.053.333.333, Jumlah Karyawan sebesar 2, dan Aset sebesar 134.989.333.333. *Cluster* ini menunjukkan adanya nilai negatif pada Modal Awal yang mungkin menandakan usaha yang masih dalam tahap awal atau mengalami kerugian. Sedangkan pada *Cluster 2*,

nilai Kepemilikan tercatat sebesar 2.052, Modal Awal sebesar 17.905.558.091, Penghasilan Perbulan sebesar 6.258.132.780, Penghasilan Pertahun sebesar 75.097.593.361, Jumlah Karyawan sebesar 2.442, dan Aset sebesar 75.039.136. *Cluster* ini memiliki nilai Modal Awal yang positif dan penghasilan yang cukup tinggi, namun jumlah karyawan sedikit lebih rendah dibandingkan dengan *Cluster* 0, mencerminkan perbedaan dalam struktur operasional dan sumber daya.

4. KESIMPULAN

Kesimpulan yang dapat diambil dari hasil penelitian ini adalah bahwa penelitian berhasil mengelompokkan data UMKM di Kota Jambi, khususnya di bidang *fashion*, dengan menggunakan 1331 data yang memiliki berbagai karakteristik seperti kepemilikan, modal awal, penghasilan per bulan, penghasilan per tahun, jumlah karyawan, dan aset. Dalam penelitian ini, perhitungan *clustering* dilakukan menggunakan dua metode, yaitu perhitungan manual dan penggunaan *tools Rapid Miner*. Hasil perhitungan manual menunjukkan bahwa *Cluster* 1 terdiri dari 627 data yang berhak menerima bantuan peralatan/barang, *Cluster* 2 terdiri dari 358 data yang direkomendasikan untuk menerima bantuan fasilitas kelembagaan, dan *Cluster* 3 berjumlah 346 data yang berhak mendapatkan pelatihan. Namun, hasil yang diperoleh menggunakan *tools Rapid Miner* menunjukkan perbedaan, di mana *Cluster* 0 berjumlah 840 item, *Cluster* 1 berjumlah 9 item, dan *Cluster* 2 berjumlah 482 item. Perbedaan hasil ini disebabkan oleh perbedaan dalam pengambilan pusat *centroid*, yang mempengaruhi cara pengelompokan data dalam masing-masing *cluster*. Meskipun demikian, kedua metode ini menunjukkan bahwa analisis menggunakan *K-Means Clustering* dapat memberikan gambaran yang jelas tentang pembagian kategori UMKM yang ada di Kota Jambi, yang nantinya dapat digunakan untuk merancang kebijakan yang lebih tepat sasaran dalam memberikan bantuan kepada masing-masing kelompok UMKM sesuai dengan kebutuhan dan karakteristik mereka. Hasil penelitian ini juga memberikan kontribusi dalam meningkatkan pemahaman tentang penerapan metode *clustering* dalam konteks analisis data ekonomi lokal, khususnya dalam mendukung pengembangan UMKM di Kota Jambi.

REFERENCES

- [1] A. Rahmawati, "Penerapan Data Mining dalam Pengembangan UMKM di Era Digital," *Jurnal Ekonomi dan Bisnis Digital Indonesia*, vol. 7, no. 2, pp. 112-125, 2019.
- [2] I. Z. Maimon and L. Rokach, "Data Mining and Knowledge Discovery Handbook," Springer, 2020.
- [3] H. Ramdlan grum, N. Ismah, D. Mawesti, C. N. Aidha, F. Armintasari, and D. R. Ningrum, Pemberdayaan ekonomi untuk UMKM yang dipimpin perempuan di pedesaan Indonesia : Pembelajaran dari CSO Indonesia. 2022.
- [4] Firdausya, Lily Zahra, and Dicky Perwira Ompusunggu. 2023. "Usaha Mikro Kecil Dan Menengah (UMKM) Di Era Digital Abad 21." *Tali Jagad Journal* 1(1): 16–20
- [5] D. Susanto, "Analisis Pengelompokan UMKM untuk Optimalisasi Pasar di Kota Jambi," *Jurnal Ekonomi dan Manajemen Indonesia*, vol. 6, no. 2, pp. 123-135, 2020.
- [6] S. Hadi, "Implementasi Metode *K-Means Clustering* untuk Pengelompokan Data UMKM," *Jurnal Teknologi dan Sistem Informasi*, vol. 9, no. 1, pp. 77-89, 2021.
- [7] A. S. Kumar and R. K. Singh, "A Survey on *Clustering Techniques in Data Mining*," *International Journal of Computer Applications*, vol. 178, no. 5, pp. 10-15, 2022.
- [8] B. S. S. Rahman and F. A. Hossain, "Data Mining Techniques: A Survey," *Journal of Computer and Communications*, vol. 9, no. 3, pp. 45-55, 2021.
- [9] Khormarudin, A. N. (2016). Teknik data mining: Algoritma *K-Means clustering*. *J. Ilmu Komput*, 1-12.
- [10] 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.
- [11] Iin, I., Fadila, R., Rinaldi, A. R., & Fathurrohman, F. (2024). Penerapan Data Mining Dalam Mengelompokan Jumlah UMKM Berdasarkan Kabupaten Kota Menggunakan *K-Means Clustering*. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(2), 1446-1450.
- [12] Wahyudi, E. N., Utomo, A. P., & Mariana, N. (2019). Pengelompokan Jenis Usaha UMKM Kota Semarang Dalam Rangka Proses Pembinaan Dan Pendampingan Untuk Pengembangan Usaha Dengan Teknik Data Mining. *Dinamik*, 24(1), 13-20.
- [13] Habiballoh, H., Faqih, A., & Suprapti, T. (2024). Implementasi Algoritma *K-Means* Dalam Mengelompokan Kabupaten/Kota Di Jawa Barat Berdasarkan Jenis Dan Jumlah Potensi Objek Daya Tarik Wisata. *Jurnal Informatika dan Teknik Elektro Terapan*, 12(2).
- [14] Hendrik, H., & Mulyana, T. M. S. (2023). Pemetaan Daerah Berdasarkan Jenis Usaha Umkm Dengan Algoritma *K-Means* Di Jawa Barat. *Jurnal Teknologi Informasi, Komputer, dan Aplikasinya (JTIKA)*, 5(2), 164-172.
- [15] H. Hanafi, "Data Cleaning dalam Big Data," no. December, 2023, pp. 1–5.
- [16] F. Yunita, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Pada Penerimaan Mahasiswa Baru," *J. Sist.*, vol. 7, no. 3, p. 238, 2018, doi: 10.32520/stmsi.v7i3.388.
- [17] M. Rafi Nahjan, Nono Heryana, and Apriade Voutama, "Implementasi Rapidminer Dengan Metode Clustering K-Means Untuk Analisa Penjualan Pada Toko Oj Cell," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 101–104, 2023, doi: 10.36040/jati.v7i1.6094.