



## *Cluster Analysis of Small Medium Enterprise Performance with K-Means Clustering Algorithm*

### **Analisis Klaster Kinerja Usaha Kecil dan Menengah Menggunakan Algoritma K-Means Clustering**

**Dona Marcelina<sup>1\*</sup>, Annisa Kurnia<sup>2</sup>, Terttiaavini<sup>3</sup>**

<sup>1,3</sup>Program Studi Sistem Informasi, Universitas Indo Global Mandiri

<sup>2</sup>Program Studi Survei dan Pemetaan, Universitas Indo Global Mandiri

E-Mail: <sup>1</sup>donamarcelina@uigm.ac.id, <sup>2</sup>annisaks@uigm.ac.id, <sup>3</sup>avini.saputra@uigm.ac.id

*Received Sep 06th 2023; Revised Oct 25th 2023; Accepted Nov 5th 2023*  
*Corresponding Author: Dona Marcelina*

#### **Abstract**

*The focus of this research is to solve the problems faced by the cooperative and SMEs office of South Sumatra Province, namely the difficulty in implementing business development programs for SMEs. So far, the Cooperatives and SMEs Office of South Sumatra Province has conducted various activities related to improving the quality of SME management. However the data collection of SMEs is incomplete, it is difficult to determine the best program for SMEs that can accelerate business development in SMEs. The purpose of this study is to complete the SME data through the distribution of questionnaires and group SMEs based on SME performance. This grouping will be used to develop appropriate and targeted SME development strategies. This research uses the K-Means Clustering method with indicators, namely financial performance, product sales, and marketing strategies. The KNIME application is used as a tool for easy and accurate data analysis, data processing, data modelling, and model visualization. This research produced three clusters, namely independent SMEs, developing SMEs and assisted SMEs. It is hoped that the results of this grouping will provide useful input for the Department of Cooperatives and SMEs to implement more specific strategy development programs that suit the characteristics of each cluster.*

*Keyword: Algorithm, Clustering, K-Means Clustering, Knime application, Small and Medium Enterprises (SMEs)*

#### **Abstrak**

Fokus penelitian ini adalah untuk menyelesaikan masalah yang dihadapi oleh dinas koperasi dan UKM Provinsi Sumatera Selatan, yaitu kesulitan dalam menerapkan program pengembangan usaha bagi UKM. Selama ini dinas koperasi dan UKM Provinsi Sumatera selatan telah melakukan berbagai kegiatan yang berhubungan dengan peningkatan kualitas pengelolaan UKM. Namun karena pendataan UKM kurang lengkap, maka sulit untuk menentukan program terbaik bagi UKM yang dapat mempercepat pengembangan usaha di UKM. Tujuan dari penelitian ini adalah untuk melengkapi data UKM melalui penyebaran kuesioner dan melakukan mengelompokkan UKM berdasarkan kinerja UKM. Pengelompokan ini nantinya akan digunakan untuk menyusun strategi pengembangan UKM yang sesuai dan tepat sasaran. Penelitian ini, menggunakan metode K-Means Clustering dengan indikator, yaitu kinerja keuangan, penjualan produk, dan strategi pemasaran sebagai dasar pengelompokkan. Aplikasi KNIME digunakan sebagai alat untuk analisis data, pemrosesan data, pemodelan data, dan visualisasi model yang mudah dan akurat. Hasil analisis data menunjukkan UMKM terbagi menjadi tiga kelompok atau klaster, yaitu UKM mandiri, UKM berkembang, dan UKM binaan. Hasil pengelompokkan ini diharapkan dapat memberikan masukan yang berguna bagi Dinas Koperasi dan UKM untuk menerapkan program pengembangan strategi yang lebih spesifik yang sesuai dengan karakteristik dari masing-masing klaster.

Kata Kunci: Algoritma, Klasterisasi, K-Means Clustering, Knime, Usaha Kecil dan Menengah (UKM)

#### **1. PENDAHULUAN**

Usaha Kecil dan Menengah (UKM) memegang peran penting dalam perekonomian nasional dengan potensi besar untuk menciptakan lapangan kerja dan memperkuat pertumbuhan ekonomi di daerah. Meskipun memiliki potensi yang signifikan, banyak UKM masih menghadapi berbagai kendala dalam mengembangkan bisnis mereka, seperti masalah terkait pengembangan produk, pemasaran, dan persaingan di pasar yang

semakin ketat. Oleh karena itu, diperlukan strategi yang tepat untuk mengembangkan produk unggulan UKM agar mampu bersaing [1]. Berdasarkan data dari Dinas Koperasi dan UKM Kota Palembang disingkat DinKop-UKM mencatat, bahwa saat ini hanya ada 1.103 UKM terdaftar, padahal jumlah sebenarnya diperkirakan mencapai 1,6 juta UKM. Pendataan yang belum optimal menciptakan kesenjangan antara data yang tercatat dengan jumlah UKM yang sebenarnya [2], [3]. Data yang belum lengkap ini tidak dapat digunakan untuk pengambilan keputusan.

Dalam hal ini, rendahnya kesadaran UKM untuk memberikan data dan proses pendataan kurang maksimal menjadi penyebab permasalahan ini. Data yang tidak terintegrasi dengan baik akan sulit digunakan untuk menentukan kebutuhan UKM. Dampak dari kondisi ini adalah efektivitas dari program yang dicanangkan oleh DinKop-UKM tidak tercapai dengan maksimal, sehingga tidak memberikan manfaat yang signifikan bagi UKM [3], [4]. DinKop-UKM telah berupaya mengundang UKM di kota tersebut untuk berpartisipasi dalam berbagai kegiatan yang diadakan setiap tahun. Namun, sayangnya, upaya ini tidak selalu berhasil dan hanya diikuti oleh sebagian kecil UMK di Kota Palembang [5].

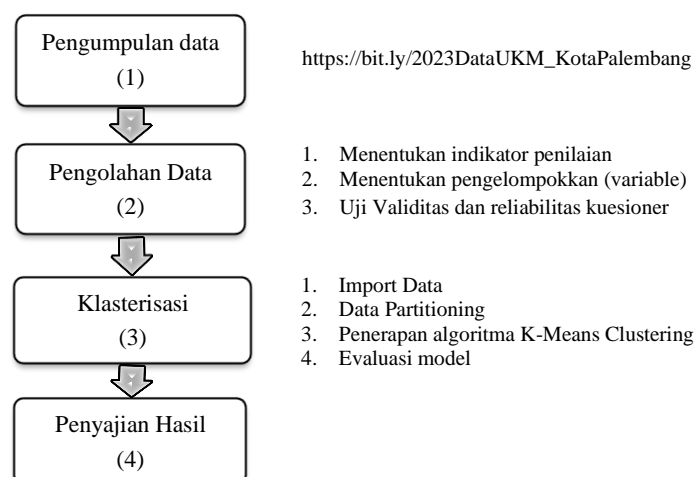
Penelitian ini bertujuan mengatasi permasalahan yang di hadapi oleh DinKop-UKM yaitu melengkap data UKM melalui penyebaran kuesioner dengan indikator yang dapat digunakan untuk menentukan program yang tepat bagi UKM dengan melakukan mengelompokkan UKM berdasarkan kinerja UKM. Pengelompokan ini nantinya akan digunakan untuk menyusun strategi pengembangan UKM yang sesuai dan tepat sasaran.

Teknik pengelompokan menggunakan metode *K-Means Clustering* [6]–[8]. K-Means Clustering dapat digunakan dengan berbagai jenis data, termasuk data numerik dan data kategorikal [9]. Hal ini dapat mempermudah dalam menganalisis data UKM yang mungkin memiliki beragam jenis data. Selain itu penelitian ini juga menggunakan KNIME sebagai alat analisis data, pemrosesan data, pemodelan data, dan visualisasi model yang mudah dan akurat. KNIME dapat mentransformasi, menggabungkan dan membersihkan data dengan mudah, selain itu hasil pengelompokan K-Means dapat divisualisasikan dengan baik sehingga dapat memberikan wawasan yang lebih komprehensif.

Hasil penelitian ini diharapkan dapat memberikan kontribusi penting dalam mengatasi masalah pendataan dan pengembangan UKM di Kota Palembang. Klasterisasi UKM berdasarkan kinerja UKM dapat ditindaklanjuti untuk pengembangan strategi pengelolaan UKM yang efektif sehingga dapat meningkatkan kesuksesan usaha UKM [10]. Selain itu juga diharapkan dapat membantu Dinas Koperasi dan UKM Kota Palembang dalam pengambilan keputusan yang lebih efektif dan berdampak positif pada pertumbuhan ekonomi dan peningkatan kesejahteraan masyarakat di Kota Palembang [3], [11].

## 2. METODOLOGI PENELITIAN

Penelitian ini menerapkan pendekatan analisis klaster dengan tujuan mengelompokkan sejumlah UKM di Kota Palembang [12]. Data primer dikumpulkan dari UKM di kota Palembang dan akan diolah menjadi informasi yang memiliki nilai signifikan [13]. Penelitian ini terdiri dari empat tahap, dijelaskan pada gambar 1 sebagai berikut:



**Gambar 1.** Tahapan penelitian

## 2.1 Pengumpulan Data

Pengolahan Data merupakan tahap awal dalam proses penelitian yang sangat kritis. Pada tahap ini, data dikumpulkan melalui distribusi kuesioner secara online yang dapat diakses melalui URL: [https://bit.ly/2023DataUKM\\_KotaPalembang](https://bit.ly/2023DataUKM_KotaPalembang). Pengumpulan data berlangsung mulai bulan Juni hingga Agustus 2023 dan mencakup berbagai aspek kinerja UKM yang memiliki signifikansi penting.

Untuk memastikan keandalan dan kualitas data, uji validitas dan reliabilitas kuesioner penilaian kinerja melibatkan partisipasi dari 100 responden yang merupakan pemilik atau pengelola UKM. Data yang diperoleh dalam penelitian ini sangat berharga, karena akan menjadi dasar analisis klusterisasi yang akan dilakukan. Proses pengumpulan data yang cermat dan melibatkan berbagai elemen kunci merupakan fondasi kuat untuk penelitian ini dan memungkinkan hasil yang akurat dan bermakna dalam mengembangkan strategi pemasaran yang lebih efektif bagi UKM di Kota Palembang [6].

Dari total 360 kuesioner yang terisi, ditemukan bahwa 10 di antaranya tidak lengkap, sehingga data tersebut harus dihapus dari analisis. Pemilik atau pengelola UKM yang telah mengisi kuesioner ini berperan penting dalam memberikan wawasan yang diperlukan untuk mengevaluasi kinerja UKM di Kota Palembang [14].

Untuk mengevaluasi kinerja UKM, digunakan kriteria yang tercantum dalam Tabel 1. Tabel ini berisi indikator-indikator penilaian yang digunakan untuk menilai kinerja UKM [15]. Indikator-indikator ini mencakup berbagai aspek yang relevan untuk menganalisis kinerja UKM di berbagai bidang [16].

**Tabel 1.** Indikator Penilaian UKM

No	Indikator	Kriteria
1.	Tingkat Pendidikan	S2 / S3; S1 / K4; D3; D1 / SMA; SMP/ SD
2.	Status Kepemilikan Usaha	Individu; Kelompok
3.	Status tempat usaha	Sewa; Milik Sendiri; Pinjam Gratis
4.	Alamat tempat Usaha	Alamat tempat usaha
5.	Skala Usaha	Usaha Mikro; Usaha Kecil; Usaha Menengah
6.	Jenis Usaha	Berbagai jenis usaha di UKM
7.	Jumlah Karyawan	Orang
8.	Omset Penjualan Perbulan	Jumlah total pendapatan yang dihasilkan
9.	Biaya Operasional	Biaya yang dikeluarkan untuk keperluan operasional sehari-hari.
10.	Labar / keuntungan Kotor	Jumlah pendapatan bersih setelah dikurangkan biaya
11.	Rata Jumlah Produksi / per Bulan	Angka rata-rata jumlah barang yang dihasilkan.
12.	Jumlah produk terjual	Total unit produk yang terjual.
13.	Kategori Pembeli	Anak-anak; Remaja Putri; Remaja Laki-laki; Orang Dewasa
14.	Jenis Usaha	Individu; Kelompok
15.	Jumlah Karyawan	Usaha Mikro; Usaha Kecil; Usaha menengah
16.	Metode Penjualan	Total jumlah orang yang bekerja

Berdasarkan data kriteria UKM, data tersebut telah dikelompokkan menjadi tiga variabel utama, yaitu:

1. Kinerja Keuangan, yaitu Variabel ini mencakup atribut Omset (OM), Biaya Operasional (BO), dan Laba (LB).
2. Penjualan Produk, yaitu Variabel ini mencakup atribut Rata Jumlah Produksi (RJP) dan Jumlah Produk Terjual (JPT).
3. Strategi Pemasaran, yaitu Variabel ini mencakup atribut Status Kepemilikan Usaha (SKU), Jenis Usaha (JU), Jumlah Karyawan (JK), dan Metode Penjualan (MP).

Pengelompokan data ini memungkinkan analisis yang lebih terfokus terhadap faktor-faktor yang memengaruhi produktivitas UKM. Dengan memisahkan variabel-variabel ini ke dalam tiga kelompok utama, penelitian dapat menganalisis masing-masing aspek secara terpisah, mengeksplorasi hubungan antar variabel, dan mengidentifikasi pola atau tren yang mungkin muncul [17], [18]. Dengan demikian, penelitian ini dapat memberikan wawasan yang lebih mendalam tentang faktor-faktor yang memengaruhi kinerja UKM di Kota Palembang.

## 2.2 Pengolahan Data

Pengolahan data merupakan serangkaian langkah yang dilakukan untuk membersihkan, merapikan, dan mengorganisasi data mentah menjadi format yang dapat digunakan untuk analisis [19]. Dalam konteks penelitian ini, pengolahan data mencakup proses mempersiapkan data kinerja UKM dan kualitas produk untuk analisis klusterisasi menggunakan algoritma *K-Means Clustering* [20]. Adapun tahapan pengolahan data adalah

### 1. Uji Validitas dan Reliabilitas Kuesioner.

Tahap awal melibatkan uji validitas dan reliabilitas kuesioner penilaian. Hal ini melibatkan partisipasi dari 100 responden. Uji validitas bertujuan untuk memastikan bahwa kuesioner benar-benar mengukur apa yang seharusnya diukur. Sementara uji reliabilitas digunakan untuk memeriksa konsistensi dalam kuesioner. Hasil dari uji ini akan memberikan keyakinan bahwa kuesioner yang digunakan adalah alat yang valid dan dapat diandalkan dalam mengumpulkan data.

### 2. Proses Analisis Kluster.

Setelah mendapatkan data yang valid dan terpercaya, tahap berikutnya adalah analisis kluster. Proses ini dilakukan dengan menggunakan aplikasi KNIME, yang merupakan alat analisis data yang kuat. Dengan algoritma K-Means Clustering, data kinerja UKM dan kualitas produk akan dikelompokkan ke dalam kluster-kluster yang sesuai.

Dengan menjalani tahapan Pengolahan Data ini, penelitian ini memastikan bahwa data yang digunakan dalam analisis selanjutnya adalah data yang valid, terorganisir, dan memiliki dasar yang kuat untuk diproses. Hasil analisis kluster nantinya akan menjadi landasan untuk mengidentifikasi faktor-faktor yang memengaruhi produktivitas UKM berdasarkan kinerja.

## 3. HASIL DAN PEMBAHASAN

Pada tahap ini akan menguraikan temuan utama dari penelitian ini dan membahas implikasinya terhadap tujuan penelitian serta literatur yang relevan serta mengeksplorasi hasil dari analisis kluster UKM berdasarkan kinerja dan kualitas produk menggunakan algoritma K-Means Clustering.

### 3.1 Uji Validitas dan Reliabilitas

Uji Validitas dan Reliabilitas dilakukan agar dapat memastikan bahwa instrumen pengukuran atau kuesioner yang digunakan dalam penelitian benar-benar valid. Tujuan dari uji validitas dan reliabilitas adalah untuk mengukur dan memastikan bahwa kuesioner atau instrumen yang digunakan dalam penelitian memiliki tingkat validitas (mengukur apa yang seharusnya diukur) dan reliabilitas (konsistensi dalam mengukur) yang memadai. Uji validitas membantu menilai apakah instrumen benar-benar mengukur konstruk atau variabel yang dimaksud dalam penelitian, sementara uji reliabilitas menilai apakah instrumen tersebut konsisten dalam memberikan hasil yang sama jika digunakan secara berulang. Pengujian ini menggunakan app SPSS. Pengujian ini dilakukan dengan menggunakan perangkat lunak SPSS (*Statistical Package for the Social Sciences*). SPSS merupakan salah satu alat analisis data yang umum digunakan dalam penelitian ilmiah untuk melakukan berbagai jenis analisis statistik. Pemilihan SPSS sebagai alat analisis data yang diperlukan dalam penelitian ini didasarkan pada kehandalan dan kemampuannya dalam mengolah data secara statistik. Adapun Hasil uji validitas kuesioner ditampilkan pada Tabel 2.

**Tabel 2.** Hasil Uji Validitas

Variabel	R-Hitung	R-Tabel	Keterangan
Kinerja Keuangan	0,676	0,463	Valid
Penjualan Produk	0,952	0,463	Valid
Strategi Pemasaran	0,927	0,463	Valid

Sumber : Data diolah (2023)

Hasil uji validitas menunjukkan bahwa nilai R-hitung lebih besar dari R-tabel, mengindikasikan bahwa kuesioner yang digunakan untuk menilai kualitas produk berdasarkan empat dimensi kualitas sudah terbukti valid [21]. Dalam hal ini, nilai R-tabel diperoleh dengan menggunakan rumus korelasi product moment dengan jumlah responden sebanyak 100 orang dan tingkat signifikansi sebesar 1%, yang menghasilkan nilai sebesar 0,1654. Oleh karena itu, dapat disimpulkan bahwa semua pernyataan dalam kuesioner dinilai layak dan dapat diandalkan untuk keperluan penelitian ini. Langkah selanjutnya adalah uji reliabilitas, yang akan dijelaskan lebih lanjut dalam Tabel 3.

**Tabel 3.** Hasil Uji Reliabilitas

Variabel	Cronbach Alpha	R-Tabel	Keterangan
Kinerja Keuangan	0,710	0,1654	Reliabel
Penjualan Produk	0,888	0,1654	Reliabel
Strategi Pemasaran	0,827	0,1654	Reliabel

Berdasarkan dari hasil uji reliable dapat ditunjukkan bahwa kuesioner kualitas produk berdasarkan 3 dimensi kualitas memiliki nilai cronbach alpha  $\geq 0,1654$ . Dengan demikian, alat ukur yang digunakan dalam

penelitian ini telah memenuhi persyaratan reliabilitas yang diperlukan untuk mendapatkan data yang konsisten dan dapat dipercaya. Ini membantu memastikan bahwa analisis yang dilakukan lebih akurat dan valid.

### 3.2 Penerapan algoritma K-Means Clustering

Pada tahap ini menjelaskan penerapan algoritma K-Means Clustering. Penerapan algoritma ini merupakan langkah kunci dalam analisis data, di mana kegiatan ini berfokus pada pengelompokan data terkait kinerja UKM [22]. Metode ini dapat menggolongkan data ke dalam kelompok yang saling berhubungan, mengidentifikasi pola yang mungkin tidak terlihat pada iterasi pertama, dan menghasilkan wawasan yang kuat yang dapat mendukung pengambilan keputusan yang lebih cerdas dalam pengembangan strategi pemasaran yang lebih efektif [9].

Penelitian ini memanfaatkan aplikasi KNIME sebagai alat utama untuk melakukan analisis data dan menerapkan algoritma K-Means Clustering [23]. KNIME adalah platform analisis data yang bersifat sumber terbuka dan telah terbukti efektif dalam mengelola, mengolah, dan menganalisis data secara komprehensif. Pemanfaatan KNIME dalam penelitian ini memberikan keunggulan dalam merapikan, mengintegrasikan, serta memodelkan data kinerja UKM dan kualitas produk dengan cara yang sistematis dan terstruktur. Penggunaan KNIME memberikan pendekatan yang lebih terorganisir untuk menganalisis data yang relevan.

Untuk pembuatan model klustering dengan aplikasi KNIME, diimplementasikan serangkaian node yang masing-masing memiliki peran dan fungsi khusus. Prosesnya dimulai dengan pengumpulan dan persiapan data kinerja UKM dan kualitas produk. Data tersebut kemudian diimpor ke dalam lingkungan KNIME. Proses pengolahan data dimulai dengan langkah awal yaitu pengumpulan data UKM yang telah dijelaskan sebelumnya. Tahapannya dalam penggunaan aplikasi KNIME dengan algoritma K-Means adalah sebagai berikut:

#### 1. Import Data

Data kinerja UKM yang telah dikumpulkan dari Kota Palembang diimpor ke dalam lingkungan KNIME menggunakan node CSV Reader.

#### 2. Data Partitioning

Dataset akan dibagi menjadi dua subset utama, yaitu set pelatihan (*training set*) dan set pengujian (*testing set*) [24]. Pembagian ini bertujuan untuk menguji dan mengevaluasi sejauh mana model yang telah dibangun mampu memprediksi dengan baik. Rasio tipikal adalah 85-15, di mana 85% data digunakan untuk pelatihan model, dan 15% data digunakan untuk pengujian.

#### 3. Penerapan algoritma K-Means Clustering

Algoritma *K-Means Clustering* digunakan untuk mengelompokkan UKM berdasarkan karakteristik tertentu yang memungkinkan identifikasi pola dan relasi antar UKM. Adapun tahapan algoritma yang digunakan adalah sebagai berikut:

##### 1) Inisialisasi

Pilih jumlah kluster (K) yang akan dibentuk. Pada penelitian ini Jumlah kluster = 3. Tentukan secara acak K titik sebagai pusat awal kluster (centroid). Pusat ini mewakili titik awal setiap kluster.

##### 2) Pengelompokan

Setiap data point (pengamatan) diberikan kepada kluster yang memiliki pusat terdekat. Perhitungan setiap data point dengan kluster terdekat menggunakan rumus *euclidean distance* yaitu:

$$d(x, y) = \sum_{i=1}^n |x_i - y_i| \quad (1)$$

Dimana:

d = jarak antara titik pada data training x dan titik data testing y yang akan diklasifikasi, dimana

$x = x_1, x_2, \dots, x_i$  dan  $y = y_1, y_2, \dots, y_i$

d = Jarak antara x dan y

x = data pusat cluster

y = data pada atribut

i = setiap data

n = banyaknya objek

$x_i$  = data pada pusat cluster ke i

$y_i$  = data pada setiap data ke i

##### 3) Pembaruan Pusat Kluster.

Hitung rata-rata (centroid) dari semua data point yang termasuk dalam kluster tersebut.

- 4) Iterasi:  
Langkah 2 dan 3 diulangi hingga tidak ada perubahan dalam pengelompokan data atau hingga batasan iterasi yang ditentukan tercapai. Dalam setiap iterasi, data mungkin berpindah ke kluster yang berbeda, dan pusat kluster dihitung ulang.
- 5) Hasil Akhir.  
Setelah konvergensi (ketika tidak ada lagi perubahan dalam pengelompokan), algoritma akan menghasilkan kluster yang stabil. Setiap data point akan termasuk dalam satu kluster yang sesuai dengan pusat kluster terdekat.
- 6) Evaluasi.  
Hasil pengelompokan dapat dievaluasi menggunakan berbagai metrik, seperti SSE (*Sum of Squared Errors*) atau indeks validitas kluster, untuk mengukur kualitas kluster yang dihasilkan. Proses klusterisasi dengan Algoritma *K-Means Clustering* ini dilakukan dengan menggunakan node *K-Means* pada KNIME.

#### 4. Evaluasi model

Evaluasi model dilakukan untuk menilai performa model klustering UKM yang telah dibentuk oleh node *Regression Predictor*, dengan tujuan menganalisis dan mengukur sejauh mana model klustering ini dapat memberikan prediksi atau estimasi yang akurat terhadap karakteristik UKM. Selain itu, pengujian model juga melibatkan penggunaan indeks validitas kluster untuk mengukur kualitas dari kluster yang dihasilkan. Tabel 4 menampilkan hasil indeks validitas kluster menggunakan *node Numeric Score*.

**Tabel 4.** Analisa *Node Numeric Score*

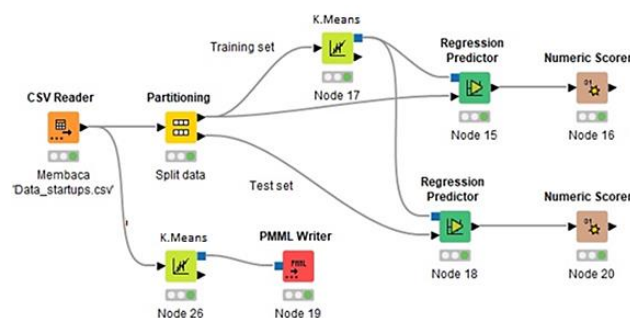
Name	Value
Cohen's kappa	0.8756246812850586
#False	8
#Correct	47
Error	0.14545454545454545
Accuracy	0.8545454545454545

Berdasarkan analisis menggunakan *node score*, diperoleh nilai Cohen's Kappa sebesar 0.87. Hal ini menunjukkan bahwa hasil prediksi model sesuai dengan hasil pengukuran aktual. Nilai Cohen's Kappa berkisar antara -1 hingga 1. Semakin mendekati 1, semakin baik konsistensi antara prediksi dan hasil aktual. Nilai positif menunjukkan kesesuaian antara prediksi dan data aktual, sedangkan nilai negatif menunjukkan ketidaksesuaian. Nilai Error sebesar 0.145 mengindikasikan bahwa rata-rata selisih antara nilai prediksi dan nilai aktual adalah sekitar 0.145. Nilai Error diukur dalam bentuk persentase.

Dalam persentase, nilai Error 0.145 setara dengan 14.5%. Ini berarti model memiliki tingkat kesalahan sekitar 14.5% dalam memprediksi data uji yang digunakan. Sedangkan nilai Accuracy sebesar 0.85 mengindikasikan bahwa model mampu memprediksi dengan benar dalam sekitar 85% dari total kasus yang diuji. Nilai ini menggambarkan sejauh mana model berhasil memisahkan kelas atau kategori dengan benar dalam prediksi.

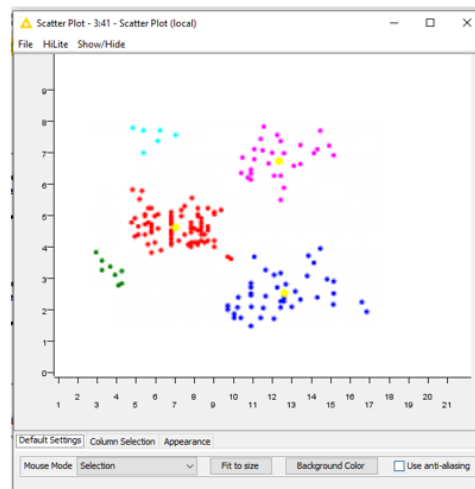
#### 5. Visualisasi Hasil.

Visualisasi data yang dihasilkan dalam KNIME bertujuan untuk memberikan gambaran yang lebih jelas tentang pola kluster yang dihasilkan dari data setelah menjalankan algoritma *K-Means Clustering*. Hasil dari workflow data ini dapat ditemukan pada Gambar 2, yang akan memperlihatkan dengan lebih terperinci kluster-kluster yang terbentuk serta distribusi data dalam masing-masing kluster.



**Gambar 2.** Visualisasi Model UKMK menggunakan aplikasi Knime

*Node* yang digunakan untuk melihat hasil dalam bentuk pola atau visualisasi adalah *Scatter Plot*. *Node* ini memungkinkan untuk memvisualisasikan data dalam bentuk grafik titik yang dapat membantu melihat pola-pola dalam data dan hubungan antar variabel dengan lebih jelas. Dengan menggunakan *Scatter Plot*, dapat dengan mudah mengidentifikasi kelompok data atau kluster yang terbentuk setelah proses analisis kluster, seperti yang dihasilkan dari algoritma *K-Means Clustering* pada gambar 3 berikut.



**Gambar 3.** Visualisasi dengan menggunakan *Scatter Plot*

Dalam visualisasi *Scatter Plot* pada gambar 3 di atas, hubungan antara tiga variabel kunci dalam konteks UKM, yaitu Kinerja Keuangan, Penjualan Produk, dan Strategi Pemasaran menggunakan warna berbeda untuk tiga kelompok kluster, yaitu UKM Mandiri, UKM Berkembang dan UKM Binaan. Setiap titik pada plot mewakili satu UKM. Pada sebaran data pada plot. UKM Mandiri, diwakili oleh titik-titik berwarna merah, terletak di sepanjang sumbu X (Kinerja Keuangan) dengan penjualan produk yang bervariasi. Hal ini menunjukkan kecenderungan UKM tersebut memiliki strategi pemasaran yang beragam, tercermin dari sebaran data yang lebih luas di sumbu Y (Penjualan Produk).

Sementara itu, UKM Berkembang, yang ditunjukkan dengan titik-titik berwarna biru, cenderung memiliki kinerja keuangan yang lebih kuat, terutama di sepanjang sumbu X, dengan penjualan produk yang meningkat secara konsisten. Hal ini kecenderungan UKM tersebut memiliki strategi pemasaran yang lebih fokus, ditunjukkan dengan sebaran data yang lebih terkonsentrasi di sepanjang sumbu Y.

UKM Binaan, yang diwakili oleh titik-titik berwarna hijau, cenderung memiliki kinerja keuangan yang bervariasi, dengan sebaran data di sepanjang sumbu X yang lebih luas. Penjualan produk mereka bervariasi, dan mereka memiliki variasi dalam strategi pemasaran mereka, seperti yang tercermin dari sebaran data di sepanjang sumbu Y.

Dengan menggunakan visualisasi ini, kita dapat melihat perbedaan dalam perilaku dan karakteristik antara tiga kelompok UKM, yang dapat membantu dalam pengambilan keputusan dan perencanaan strategi lebih lanjut dalam mendukung pertumbuhan dan keberlanjutan UKM. Penelitian ini menyajikan sebuah pendekatan yang komprehensif dalam menganalisis UKM dan memberikan wawasan berharga untuk pengembangan lebih lanjut.

#### 4. KESIMPULAN

Hasil penelitian ini menunjukkan bahwa pendekatan ini dapat membantu dalam mengelompokkan UKM berdasarkan karakteristik tertentu. Beberapa temuan utama dari penelitian ini yang dapat disimpulkan adalah Penerapan algoritma *K-Means Clustering* menggunakan aplikasi KNIME memberikan pendekatan terstruktur dalam mengelola data. Hal ini memungkinkan pengelompokan data kinerja UKM menjadi kelompok yang saling berhubungan dan identifikasi pola yang mungkin tidak terlihat sebelumnya. Evaluasi model dengan menggunakan *node Regression Predictor* dan *node Numeric Score* memberikan gambaran tentang sejauh mana model klustering UKM dapat memprediksi secara akurat. Visualisasi data menggunakan *Scatter Plot* memungkinkan pemahaman yang lebih baik tentang pola kluster yang dihasilkan. Hasil dari klasterisasi ini menghasilkan 3 (tiga) kelompok data yang saling berdekatan dengan inisialisasi, yaitu UKM mandiri, UKM berkembang dan UKM binaan. Pengelompokan ini menjadi masukan bagi DinKop-UKM untuk menerapkan program pengembangan strategi yang lebih tepat sasaran.

Berdasarkan hasil penelitian ini, beberapa saran dapat diajukan diantaranya adalah Hasil penelitian ini dapat digunakan sebagai dasar untuk pengembangan strategi pemasaran yang lebih efektif bagi UKM di Kota Palembang. Analisis klustering memungkinkan untuk menyesuaikan strategi sesuai dengan karakteristik

masing-masing kelompok UKM. UKM perlu lebih memanfaatkan teknologi informasi untuk meningkatkan efisiensi operasional dan meningkatkan visibilitas produk mereka. DinKop-UKM dapat memberikan dukungan dalam hal ini. Diperlukan upaya untuk memastikan bahwa seluruh UKM di Kota Palembang terlibat dalam berbagai program dan kegiatan yang diselenggarakan oleh DinKop-UKM. Kolaborasi yang kuat dapat membantu dalam pertumbuhan ekonomi dan kesejahteraan masyarakat. Penelitian ini memberikan kontribusi yang signifikan dalam mengatasi masalah pendataan dan pengembangan UKM di Kota Palembang. Semoga temuan dan rekomendasi dari penelitian ini memberikan manfaat yang nyata bagi UKM dan pertumbuhan ekonomi daerah.

## REFERENSI

- [1] Terttiaavini, L. Marnisah, Y. Yulius, and T. Setiawan Saputra, "Pengembangan Kewirausahaan 'Kemplang Tunu' Sebagai Produk Cemilan Khas Kota Palembang," *Abdimas Mandiri*, vol. 3, no. 1, pp. 63–72, 2019.
- [2] A. Perdana, H. H. Lee, S. K. Koh, and D. Arisandi, "Data analytics in small and mid-size enterprises: Enablers and inhibitors for business value and firm performance," *Int. J. Account. Inf. Syst.*, vol. 44, no. November 2021, p. 100547, 2022, doi: 10.1016/j.accinf.2021.100547.
- [3] T. Terttiaavini, F. Zamzam, M. Ramadhan, and T. S. Saputra, "Design a Decision Support System to Evaluate The Performance of Indonesian Lecturers by Developing a Simple Additive Weighting Method," *Test Eng. Manag.*, vol. 28, no. 11, pp. 36–41, 2019, [Online]. Available: <http://sersc.org/journals/index.php/IJAST/article/view/1038/903>
- [4] A. F. Sari, R. H. Sampurna, and D. Meigawati, "Strategi Dinas Koperasi, Ukm, Perdagangan Dan Perindustrian Dalam Pemberdayaan Umkm Di Kota Sukabumi," *J. Inov. Penelit.*, vol. 2, no. 10, pp. 3353–3360, 2022.
- [5] Y. Abdullah and I. Gultom, "Dinas Koperasi dan UMKM Palembang bentuk tim data usaha kecil," *Antara News*. <https://sumsel.antaraneews.com/berita/643913/dinas-koperasi-dan-umkm-palembang-bentuk-tim-data-usaha-kecil> (accessed Apr. 12, 2023).
- [6] Terttiaavini *et al.*, "Clustering Analysis of Premier Research Fields," *Int. J. Eng. Technol.*, vol. 7, no. 4.44, p. 43, 2018, doi: 10.14419/ijet.v7i4.44.26860.
- [7] Y. T. Chen and D. M. Witten, "Selective inference for k-means clustering," *J. Mach. Learn. Res.*, vol. 24, pp. 1–20, 2022, [Online]. Available: <http://arxiv.org/abs/2203.15267>
- [8] A. Heryati and M. I. Herdiansyah, "The Application of Data Mining by using K Means Clustering Method in Determining New Students' Admission Promotion Strategy," *Int. J. Eng. Adv. Technol.*, vol. 9, no. 3, pp. 824–833, 2020, doi: 10.35940/ijeat.c5414.029320.
- [9] D. Abdullah, S. Susilo, A. S. Ahmar, R. Rusli, and R. Hidayat, "The application of K-means clustering for province clustering in Indonesia of the risk of the COVID-19 pandemic based on COVID-19 data," *Qual. Quant.*, vol. 56, no. 3, pp. 1283–1291, 2022, doi: 10.1007/s11135-021-01176-w.
- [10] E. N. Wahyudi, A. P. Utomo, and N. Mariana, "Pengelompokan Jenis Usaha Umkm Kota Semarang Dalam Rangka Proses Pembinaan Dan Pendampingan Untuk Pengembangan Usaha Dengan Teknik Data Mining," *Dinamik*, vol. 24, no. 1, pp. 13–20, 2019, doi: 10.35315/dinamik.v24i1.7840.
- [11] S. A. Waqas and I. Khan, "Development Of A Sustainable Business Model For Small Businesses Using Data Mining," *J. Islam Econ. Financ. Islam Ekon. Ve Finans. Derg.*, vol. 7, no. 1, pp. 111–144, 2021.
- [12] E. Rouza, Basorudin, and Efrida, "Identifikasi dan Klasifikasi UMKM di Kabupaten Rokan Hulu Menggunakan Metode K-Means," *J. Ilm. Univ. Pengaraian*, vol. 7, no. 01, pp. 32–40, 2021.
- [13] Hartatik *et al.*, *Data Science - Data Science*, no. September 2016. 2023. [Online]. Available: [https://www.data-science.ruhr/about\\_us/](https://www.data-science.ruhr/about_us/)
- [14] Terttiaavini, L. Marnisah, Y. Yulius, and T. S. Saputra, "Evaluating the Kemplang Tunu Production Training for Low Income and Education Communities Using the Kirkpatrick Model," in *Atlantis Press, Fakultas Ekonomi Trisakti*, 2020, pp. 238–242. doi: 10.2991/aebmr.k.200915.055.
- [15] Terttiaaviani *et al.*, "Data Profil di Desa Pemulutan Ilir Organ Ilir Sumatera Selatan," *JMM (Jurnal Masy. Mandiri)*, vol. 5, no. 6, pp. 3–11, 2021, [Online]. Available: <http://journal.ummat.ac.id/index.php/jmm/article/view/5608>
- [16] T. Terttiaavini, Sofian, and T. S. Saputra, "Pendampingan Penyusunan Program Rencana Kerja Badan Usaha Milik Desa Dalam Rangka Optimalisasi Potensi Desa Serjabo Organ Ilir Sumatera Selatan," *JMM (Jurnal Masy. Mandiri)*, vol. 5, no. 6, pp. 3–11, 2022, [Online]. Available: <http://journal.ummat.ac.id/index.php/jmm/article/view/5608>
- [17] T. M. Ghazal *et al.*, "Performances of k-means clustering algorithm with different distance metrics," *Intell. Autom. Soft Comput.*, vol. 30, no. 2, pp. 735–742, 2021, doi: 10.32604/iasc.2021.019067.
- [18] M. Ahmed, R. Seraj, and S. M. S. Islam, "The k-means algorithm: A comprehensive survey and performance evaluation," *Electron.*, vol. 9, no. 8, pp. 1–12, 2020, doi: 10.3390/electronics9081295.
- [19] S. A. N. Alexandropoulos, S. B. Kotsiantis, and M. N. Vrahatis, *Data preprocessing in predictive data*



- mining*, vol. 34, no. April 2020. 2019. doi: 10.1017/S026988891800036X.
- [20] S. P. Tamba, M. D. Batubara, W. Purba, M. Sihombing, V. M. Mulia Siregar, and J. Banjarnahor, "Book data grouping in libraries using the k-means clustering method," *J. Phys. Conf. Ser.*, vol. 1230, no. 1, 2019, doi: 10.1088/1742-6596/1230/1/012074.
- [21] H. K. Mohajan, "Two Criteria for Good Measurements in Research: Validity and Reliability," *Ann. Spuru Haret Univ. Econ. Ser.*, vol. 17, no. 4, pp. 59–82, 2017, doi: 10.26458/1746.
- [22] E. G. Unus, "Teknologi Data Mining Berbasis Metode Clustering Sebagai Ujung Tombak Perkembangan Umkm Di Indonesia Dalam Era Revolusi Industri 4.0," *J. Repos.*, vol. 3, no. 3, pp. 315–320, 2021, doi: 10.22219/repositor.v2i3.1266.
- [23] S. O'Hagan and D. B. Kell, "Software review: the KNIME workflow environment and its applications in genetic programming and machine learning," *Genet. Program. Evolvable Mach.*, vol. 16, no. 3, pp. 387–391, 2015, doi: 10.1007/s10710-015-9247-3.
- [24] K. P. Sinaga and M. S. Yang, "Unsupervised K-means clustering algorithm," *IEEE Access*, vol. 8, pp. 80716–80727, 2020, doi: 10.1109/ACCESS.2020.2988796.