

## **PENERAPAN METODE CRISP-DM UNTUK ANALISA PENDAPATAN BERSIH BULANAN PEKERJA INFORMAL DI PROVINSI JAWA BARAT DENGAN ALGORITMA K-MEANS**

**<sup>1</sup>Farras Salsabila, <sup>2</sup>Ika Fitrianti, <sup>3</sup>Yuyun Umaidah, <sup>4</sup>Nono Heryana**

<sup>1,2,3</sup> Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Singaperbangsa Karawang  
e-mail: <sup>1</sup>2010631250012@student.unsika.ac.id, <sup>2</sup>2010631250053@student.unsika.ac.id,  
<sup>3</sup>yuyun.umaidah@staff.unsika.ac.id, <sup>4</sup>nono@staff.unsika.ac.id

### **ABSTRAK**

Tujuan dari penelitian ini adalah untuk mengenali pola pendapatan dan karakteristik pekerja informal dengan lebih baik dengan menggunakan clusterisasi Algoritma K-Means sehingga pemerintah dapat mengarahkan sumber daya dan program dengan lebih tepat sasaran. Dalam penelitian ini, Indeks Davies Bouldin (DBI) akan digunakan sebagai metode untuk mengoptimalkan jumlah cluster dalam proses pengelompokan wilayah Kabupaten/Kota di Jawa Barat berdasarkan pendapatan bersih pekerja informal. Data yang digunakan adalah data pendapatan dari para pekerja Informal di Jawa Barat. Pengolahan data dilakukan menggunakan tools RapidMiner dengan pemodelan clustering dan untuk metode penelitian yang digunakan adalah CRISP-DM. Hasil pengujian dan pemodelan yang telah adalah terdapat indikasi adanya perbedaan yang cukup besar dalam pendapatan bulanan bersih pekerja informal di daerah kabupaten/kota di Jawa Barat. Cluster 0 menunjukkan pendapatan rendah, Cluster 1 menunjukkan pendapatan sedang, sementara Cluster 2 menunjukkan pendapatan tinggi

**KATA KUNCI:** K-Means, CRISP-DM, RapidMiner, Pekerja Informal, Pendapatan

### **1. PENDAHULUAN**

Peningkatan sektor usaha di Indonesia telah memberikan dampak positif baik bagi ekonomi negara maupun perkembangan pekerja informal. Pekerja informal merupakan pekerja yang memiliki usaha berdiri sendiri atau perseorangan. Di Indonesia sendiri, pekerja informal juga memiliki landasan hukum. Salah satu landasan hukum yang relevan adalah Undang-Undang Nomor 25 Tahun 1997 tentang ketenagakerjaan yang berisi pekerja sektor informal merujuk pada individu yang bekerja dalam lingkup sektor informal dengan menerima pembayaran atau imbalan atas pekerjaan yang dilakukan.

Sebagaimana data yang terdapat dalam BPS Agustus 2022, jumlah pekerja informal mencapai 80,24 juta jiwa atau 59,31% dari total tenaga kerja. Jumlah ini lebih besar dibandingkan dengan pekerja formal yang hanya mencapai 55,06 juta jiwa atau 40,69%. Dengan jumlah yang tidak sedikit ini, maka pekerja informal diharapkan mampu menyokong pertumbuhan ekonomi negara. Pekerja informal berperan penting dalam sektor-sektor perdagangan, jasa, dan pertanian, serta diharapkan dapat mendukung pertumbuhan ekonomi negara. Oleh karena itu, pemerintah perlu memberikan dukungan dan perhatian kepada pekerja informal untuk meningkatkan kesejahteraan dan mendorong pertumbuhan ekonomi yang inklusif dan berkelanjutan.

Disamping itu, pekerja informal menghadapi tantangan utama berupa ketidakstabilan pendapatan. Pendapatan rata-rata mereka masih di bawah persentase yang cukup. Fluktuasi pendapatan yang signifikan terjadi karena ketergantungan pada kondisi pasar, musim, dan permintaan pelanggan yang berubah-ubah. Ketidakstabilan pendapatan ini menyulitkan pekerja informal dalam perencanaan keuangan, memenuhi kebutuhan sehari-hari, dan membangun stabilitas ekonomi jangka panjang.

Untuk dapat mengatasi tantangan diatas, maka dilakukan pengklasifikasian penduduk berdasarkan tingkat pendapatan mereka menggunakan teknik clustering. Salah satu algoritma clustering yang digunakan adalah K-Means. Dengan menganalisis pendapatan rata-rata penduduk menggunakan algoritma K-Means pada pekerja informal, kita dapat memperoleh wawasan tentang pola pendapatan antar kelompok penduduk.

Berdasarkan penelitian yang dilakukan Zulfa Nabila, dkk. membahas mengenai Analisis Kasus Covid di Provinsi Lampung menggunakan clustering dengan algoritma K-Means dan menghasilkan data yang dibagi kedalam 4 kategori yaitu Zona Merah, Zona Orange, Zona Kuning dan Zona Hijau.

Data tersebut juga divalidasi menggunakan DBI dan menghasilkan perhitungan mendekati 0, yang berarti cluster yang dihasilkan merupakan klaster yang baik.[1]

Gilang Fajriansyah, dkk. melakukan penelitian menggunakan algoritma K-Means mengenai daftar pemilih tetap pada hasil rekapitulasi KPU berdasarkan usia di Kota Bandar Lampung. Data dikelompokkan menjadi usia muda, usia dewasa, dan usia lansia. Berdasarkan hasil analisis cluster, usia paling dominan di setiap Kecamatan, termasuk Kecamatan Langkapura, adalah usia muda sebanyak 3379 orang.[2]

Penelitian yang dilakukan Dwi Astuti, dkk. melakukan penelitian terhadap algoritma K-Means Clustering dapat menjadi salah satu metode untuk menentukan strategi promosi yang efektif bagi UMKM. Data dikelompokkan kedalam 3 variabel sesuai karakteristiknya. Hasil penelitian tersebut menghasilkan clustersatu sebanyak 182 UMKM, cluster dua terdiri dari 47 UMKM, dan klaster tiga terdiri dari 55 UMKM.

Berdasarkan beberapa penelitian sebelumnya, maka metode CRISP-DM terbukti efektif untuk melakukan analisis pendapatan bersih bulanan pekerja informal di Provinsi Jawa Barat. Metode CRISP-DM menjadi suatu kerangka kerja yang digunakan dalam proses penggalian data. Kerangka kerja ini terdiri dari serangkaian tahapan yang sistematis dan terstruktur, dimulai dari pemahaman bisnis hingga penyebaran hasil analisis. CRISP-DM telah terbukti efektif dalam berbagai industri, termasuk dalam analisis data untuk pemahaman pendapatan. Pekerja informal di Provinsi Jawa Barat merupakan kelompok yang memiliki karakteristik khusus dalam hal pendapatan. Oleh karena itu, pemahaman yang komprehensif tentang pendapatan bersih bulanan pekerja informal menjadi penting untuk merumuskan kebijakan yang sesuai

Dengan demikian, penerapan metode CRISP-DM dan penggunaan algoritma K-means dalam analisis pendapatan bersih bulanan pekerja informal di Provinsi Jawa Barat merupakan pendekatan yang berpotensi kuat untuk mendapatkan pemahaman yang lebih baik tentang pola pendapatan dan karakteristik pekerja informal, serta mendorong pengambilan kebijakan yang lebih efektif dan berkelanjutan.

Penelitian ini juga dapat memberikan rekomendasi kepada pemerintah, terkhusus kepada pemerintah Jawa Barat dan Kementerian Ketenagakerjaan untuk mengembangkan strategi dan program yang lebih tepat sasaran dalam meningkatkan pendapatan pekerja informal. Sehingga, diharapkan dapat tercipta kondisi yang lebih baik bagi pekerja informal dalam merencanakan keuangan, mengatasi ketidakstabilan pendapatan, dan meningkatkan kesejahteraan ekonomi mereka.

## 2. TINJAUAN PUSTAKA

### 2.1. Pendapatan

Pendapatan merujuk pada jumlah uang atau nilai ekonomi yang diperoleh oleh individu atau perusahaan sebagai hasil dari aktivitas atau penjualan produk atau jasa. Dalam bidang akuntansi atau bisnis, pendapatan dapat dijelaskan sebagai jumlah uang yang diperoleh oleh perusahaan melalui berbagai kegiatan, terutama penjualan produk atau jasa. Pendapatan ini termasuk dalam kegiatan normal perusahaan selama satu periode akuntansi yang meningkatkan ekuitas (modal) dan tidak berasal dari investasi modal. Penting untuk membedakan antara pendapatan dan penghasilan dalam konteks akuntansi. Pendapatan mengacu pada pendapatan kotor atau laba kotor dari suatu usaha sebelum dikurangi dengan beban atau biaya. Secara umum, pendapatan ini juga dikenal sebagai omzet penjualan.

### 2.2. Pekerja Informal

Pekerja informal merujuk pada pekerjaan yang dilaksanakan oleh individu yang tidak mengadakan kesepakatan kerja tertulis sesuai dengan peraturan yang berlaku.[3] Perlindungan hukum bagi pekerja informal berbeda dengan pekerja formal karena mereka tidak memiliki kesepakatan kerja yang tertulis.[4] Pekerja tidak resmi berisiko terkena dampak dari ketidakstabilan ekonomi, contohnya pada masa pandemi COVID-19, yang berpotensi menyebabkan penurunan penawaran dan pendapatan[5] Program jaminan kesehatan yang dimiliki oleh pemerintah Indonesia bertujuan untuk meningkatkan kesejahteraan masyarakat dalam hal kesehatan. Namun, masih ada banyak pekerja informal yang belum mendapatkan perlindungan dari program Jaminan Kesehatan Nasional (JKN) karena mereka merasa tidak mampu untuk membayar iuran.[6]. Perlindungan kesehatan yang memadai sangat penting bagi pekerja informal karena mereka juga memiliki risiko terpapar kecelakaan.[7]

### 2.3. Clustering

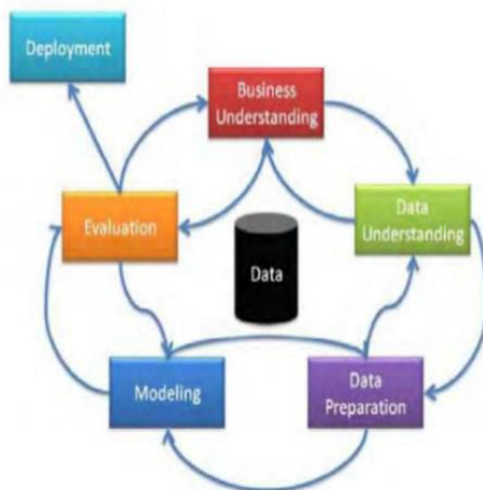
Clustering merupakan metode dalam data mining yang dipakai untuk mengkategorikan sejumlah data atau objek ke dalam kelompok (cluster), dengan tujuan agar setiap cluster berisi data yang serupa dan berbeda dari objek dalam cluster lainnya. [8] Penerapan clustering dapat diterapkan dalam berbagai bidang, termasuk dalam pengolahan citra. [9] Beberapa teknik pengelompokan yang bisa diterapkan meliputi k-means, Power Iteration Clustering (PIC), Self Organizing Maps (SOM), dan metode lainnya.[9] Perbedaan terdapat dalam tingkat keakuratan yang dimiliki oleh setiap teknik clustering dalam menghasilkan modelnya. Harapannya, dengan menggabungkan metode klasifikasi dan clustering dalam data mining, akurasi yang diperoleh dapat ditingkatkan.[10].

### 2.4. Algoritma K-Means

Salah satu teknik clustering yang sering digunakan dalam data mining adalah algoritma k-means. Algoritma ini berguna untuk mengelompokkan sejumlah data atau objek ke dalam cluster, di mana setiap cluster berisi data yang mirip satu sama lain dan berbeda dari objek dalam cluster lainnya.[10] Proses k-means dilakukan dengan menghitung jarak antara setiap objek data dengan pusat (centroid) dari setiap cluster, dan kemudian mengelompokkan objek data ke dalam cluster yang memiliki centroid terdekat. Selanjutnya, centroid dari setiap cluster dihitung ulang berdasarkan rata-rata objek data yang ada di dalam cluster tersebut. [10]Proses pengelompokkan ulang dilakukan berulang kali hingga tidak ada lagi perubahan dalam pengelompokkan objek data. Algoritma k-means dapat diterapkan dalam berbagai bidang, seperti pengolahan citra, analisis keuangan, dan penentuan penerima beasiswa. Terdapat beberapa metode k-means yang dapat digunakan, seperti k-means++, Power Iteration Clustering (PIC), dan lainnya. Setiap teknik clustering memiliki tingkat akurasi yang berbeda-beda tergantung pada model yang dihasilkan. Dengan mengintegrasikan metode klasifikasi dan clustering dalam data mining, diharapkan dapat meningkatkan tingkat akurasi yang dapat dicapai.

## 3. METODE PENELITIAN

Metode penelitian yang digunakan dalam penelitian ini adalah Cross Industry Standard Process for Data Mining (CRISP-DM). Metode CRISP-DM digunakan sebagai kerangka kerja yang terstruktur dan terdokumentasi dengan langkah-langkah yang jelas dalam melakukan analisis data mining. Metode CRISP-DM ini melibatkan 6 tahapan. Sebagai visualisasi, Gambar 1. menampilkan tahapan dari CRISP-DM..



(Sumber : Sastry dan Babu 2013)

Gambar 1. Tahapan CRISP-DM

#### 1. Business Understanding

Pada tahap awal ini, dilakukan pemahaman terhadap kebutuhan dan tujuan bisnis terkait penelitian. Tahap ini melibatkan langkah-langkah sebagai berikut:

- a) Memahami kebutuhan bisnis. Pada proses ini tujuan bisnis dan masalah yang ingin diselesaikan dapat dipahami dengan jelas.

- b) Menentukan tujuan. Setelah pemahaman bisnis yang mendalam, maka ditentukan tujuan yang ingin dicapai sesuai dengan masalah yang diangkat.
- c) Menentukan proses data mining. Proses ini harus memiliki tujuan spesifik dari proyek data mining yang harus memiliki kaitan erat dengan kebutuhan bisnis dan tujuan penelitian.
- d) Perencanaan. Proses ini menyimpulkan rencana penelitian untuk mencapai tujuan yang sudah ditentukan serta tools apa saja yang akan digunakan.

## 2. Data Understanding

Pengumpulan data dari sumber yang relevan dilakukan pada tahap ini yang kemudian data dieksplorasi secara mendalam untuk memahami struktur, kualitas dan karakteristiknya. Hal ini membantu memastikan data memiliki kualitas yang baik sehingga dapat mendukung analisis yang dilakukan. Kriteria yang dipakai adalah Pendapatan Bersih Bulanan Pekerja Informal di Provinsi Jawa Barat pada tahun

## 3. Data Preparation

Setelah data dikumpulkan dan dideskripsikan, maka tahap selanjutnya adalah data preparation yang bertujuan untuk menghasilkan data yang siap diolah memiliki kualitas terbaik. Menurut (Rahma, 2022), beberapa proses dari tahapan ini adalah:

- a) Pemilihan Data. Proses ini dilakukan pemilihan atribut yang akan digunakan sesuai dengan kebutuhan dan rencana penelitian.
- b) Pra-pemrosesan Data. Proses ini sama dengan Data Cleaning yang dilakukan pemeriksaan untuk mengatasi missing values serta duplikasi atau inkonsisten data.
- c) Transformasi Data. Proses ini dilakukan dengan mengubah tipe data yang sesuai.

## 4. Modelling

Pembuatan model data dilakukan untuk mengelompokkan data berdasarkan karakteristik atau atribut-atribut yang relevan. Hal ini dapat digunakan untuk memahami pola dari suatu data yang dapat berguna sebagai pengetahuan di tahap selanjutnya. Dengan menggunakan rumus Euclidean Distance, dapat dihitung jarak antara data dan centroid. Perhitungan pada penelitian ini menggunakan tools RapidMiner untuk membantu menemukan hasil penelitian.

## 5. Evaluation

Tahap berikutnya adalah evaluasi data, yaitu proses mengevaluasi kualitas dan keefektifan model yang telah dihasilkan. Pada penelitian ini, tahapan evaluasi menggunakan metode evaluasi Davies Bouldin Index (DBI). Hasil yang dihasilkan berupa nilai numerik yang mengindikasikan sejauh mana kelompok-kelompok yang terbentuk saling terpisah dan terdiferensiasi satu sama lain. Selanjutnya dari hasil evaluasi tersebut dapat juga menentukan apakah proses dapat dilanjutkan atau harus diulang karena tidak sesuai dengan rencana penelitian.

## 6. Deployment

Tahap penyebaran dibuat berupa laporan akhir mengenai hasil informasi yang telah didapat dari proses sebelumnya. Data yang sudah diolah dan diuji coba hingga mendapatkan hasil akan divisualisasikan sehingga dapat memberikan informasi dan pengetahuan yang lebih mudah dipahami.

# 4. HASIL DAN PEMBAHASAN

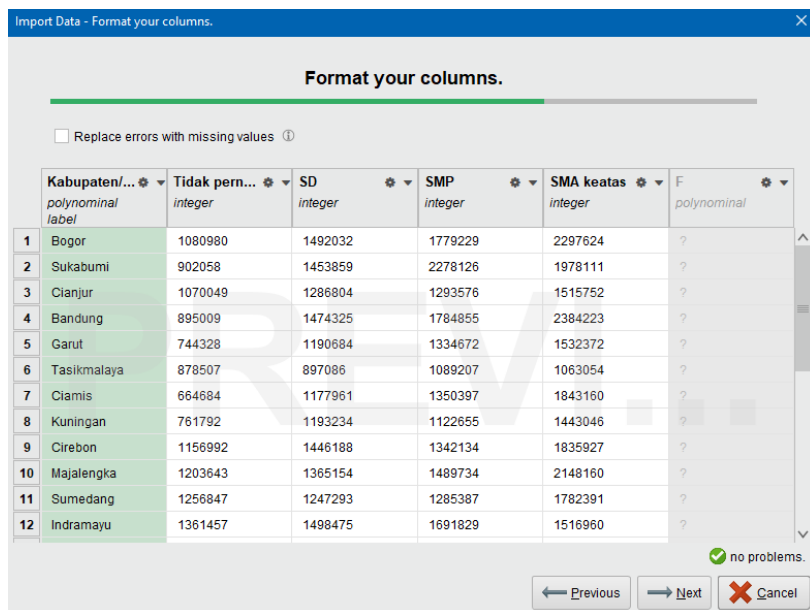
## 1. Pengumpulan Data

Dalam penelitian ini, dataset yang digunakan adalah dataset Rata-rata Pendapatan Bersih Sebulan Pekerja Informal Menurut Kabupaten/Kota dan Pendidikan Tertinggi yang Ditamatkan di Provinsi Jawa Barat (rupiah) pada tahun 2020. Data ini diambil dari website Badan Pusat Statistik (BPS) Indonesia dengan sumber yang dapat diakses melalui [www.bps.go.id](http://www.bps.go.id). Dataset ini terdiri dari sejumlah Kabupaten/Kota di Jawa Barat sebagai label dan terdiri dari 4 atribut yang akan digunakan dalam analisis. Itemset dalam dataset mencakup nama-nama Kabupaten/Kota di Jawa Barat serta atribut tentang pendapatan yang dihasilkan berdasarkan pendidikan terakhir.

## 2. Preprocessing Data

Tahap ini dilakukan untuk memastikan kualitas dan konsistensi data yang digunakan untuk tahap selanjutnya. Maka dalam tahap ini, kolom Kabupaten/Kota dijadikan label sebagai pembeda dan memiliki tipe data polynomial karena itemset lebih dari dua. Sedangkan kolom atribut tipe datanya

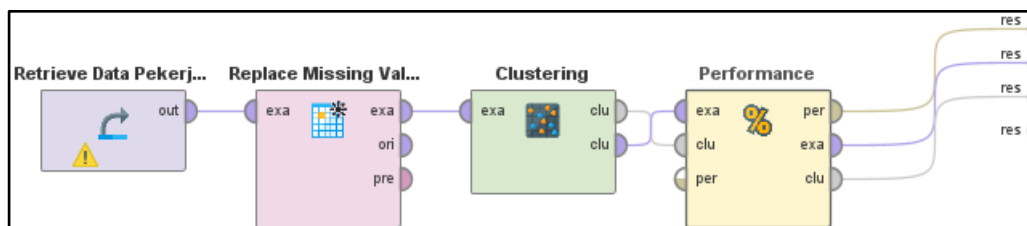
diubah menjadi integer semua karena bersifat numerik. Data yang akan diolah dapat dilihat pada Gambar 2 sebagai berikut:



Gambar 2. Preprocessing Data

### 3. Modelling

Tahapan pembuatan model ini dilakukan menggunakan algoritma K-Means Cluster dengan bantuan tools RapidMiner. Untuk mengecek apakah dataset terdapat nilai yang hilang atau tidak, maka operator replace missing values ditambahkan. Jika pada dataset terdapat nilai yang hilang, maka data tersebut dapat dihapus.



Gambar 2. Modelling Dataset

Langkah-langkah dalam tahapan modelling adalah sebagai berikut:

- a) Memasukkan dataset.
- b) Memasukkan operator Replace Missing Values sebagai operator tambahan.
- c) Menambahkan operator Clustering K-Means sebagai algoritma pemodelan. Jumlah cluster yang dimodelkan adalah 3, 4, 5, 6 dan 7.
- d) Menambahkan operator Performance untuk melihat nilai DBI sebagai bahan evaluasi.

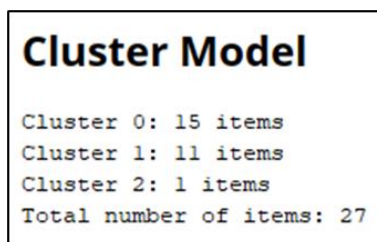
### 4. Evaluasi

Nilai DBI menjadi teori evaluasi yang digunakan dalam penelitian ini. Pada tabel 1. dapat dilihat jumlah cluster yang dimodelkan beserta nilai DBI yang dihasilkan.

Tabel 1. Nilai DBI tiap Cluster

Jumlah Cluster	Nilai DBI
3	-0.690
4	-0.786
5	-0.829
6	-0.939
7	-0.868

Mengacu pada aturan DBI yang mana nilai DBI yang baik dalam data mining adalah nilai yang semakin kecil atau mendekati nol. Dari tabel 1. jumlah cluster 3 memiliki nilai DBI yang paling baik karena nilai DBInya paling mendekati nol dibandingkan nilai DBI cluster lain yaitu sebesar -0.690. Maka, pengelompokkan dilakukan menggunakan 3 cluster. Gambar 2 dibawah ini menampilkan jumlah masing-masing anggota cluster.



Gambar 3. Jumlah Item pada tiap Cluster

Dengan total item sebanyak 27 data, maka item tersebut dikelompokkan menjadi 3 cluster. Cluster 0 beranggotakan 15 Kabupaten/Kota, cluster 1 beranggotakan 11 Kabupaten/Kota, sedangkan cluster 2 beranggotakan 1 Kabupaten/Kota.

Adapun nilai centroid yang dihasilkan dari masing-masing atribut dilihat pada gambar 4 berikut.

Attribute	cluster_0	cluster_1	cluster_2
Tidak pernah sekolah/Belum tamat SD	1210631.600	942150.455	3251355
SD	1647370.467	1283361.909	1976512
SMP	1929236.533	1289841.091	2036816
SMA keatas	2374584.800	1577870.636	2460661

Gambar 4. Nilai centroid

Pengelompokkan 3 cluster ini dapat dibedakan pula menjadi 3 kategori berdasarkan jumlah pendapatan bersih bulanan. Berdasarkan nilai centroid yang dihasilkan, kategori tersebut adalah:

- a. **Cluster 0:** Kabupaten/Kota di Jawa Barat yang memiliki jumlah pendapatan bersih bulanan (pekerja informal) **rendah**.
- b. **Cluster 1:** Kabupaten/Kota di Jawa Barat yang memiliki jumlah pendapatan bersih bulanan (pekerja informal) **sedang**.
- c. **Cluster 2:** Kabupaten/Kota di Jawa Barat yang memiliki jumlah pendapatan bersih bulanan (pekerja informal) **tinggi**.

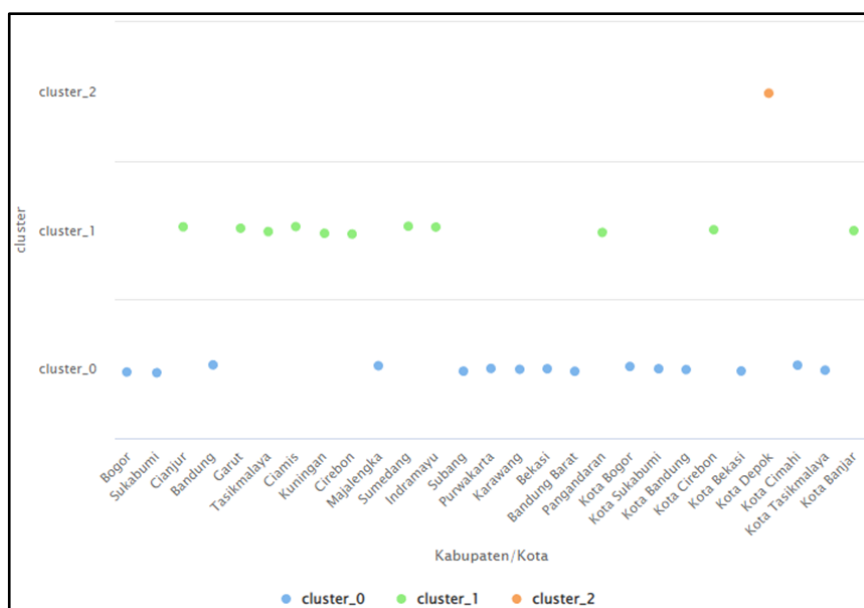
Data Kabupaten/Kota untuk setiap cluster ditampilkan dalam Tabel 2 di bawah ini:

Tabel 2. Data Kabupaten/Kota tiap cluster

Cluster	Jumlah Cluster	Anggota Cluster
0	15	Bogor, Sukabumi, Bandung, Majalengka, Bekasi, Kota Bekasi, Karawang, Subang, Purwakarta, Bandung Barat, Kota Bogor, Kota Sukabumi, Kota Bandung, Kota Cimahi, dan Kota Tasikmalaya.
1	11	Cianjur, Garut, Tasikmalaya, Ciamis, Kuningan, Cirebon, Indramayu, Sumedang, Pangandaran, Kota Cirebon, dan Kota Banjar.
2	1	Kota Depok

Dari data diatas dapat diketahui bahwa Kota Depok menjadi satu-satunya Kabupaten/Kota yang dikelompokkan ke dalam klasifikasi pendapatan bersih bulanan (pekerja informal) tinggi. Cianjur, Garut, Tasikmalaya, Ciamis, Kuningan, Cirebon, Indramayu, Sumedang, Pangandaran, Kota Cirebon, dan Kota Banjar menjadi Kabupaten/Kota yang memiliki jumlah pendapatan bersih bulanan (pekerja informal) sedang. Sedangkan 15 Kabupaten/Kota lainnya menjadi wilayah yang memiliki jumlah pendapatan bersih bulanan (pekerja informal) rendah.

Sebagai visualisasi data, grafik di bawah ini menunjukkan penyebaran anggota cluster berdasarkan hasil clustering:



Gambar 5, Visualisasi Dataset per cluster

## 5. KESIMPULAN

Berdasarkan hasil pengelompokan dan pemodelan clustering yang telah dilakukan diatas, maka didapatkan kesimpulan bahwa hasil menunjukkan terdapat 3 kategori yang berbeda berdasarkan jumlah pendapatan bersih bulanan pekerja informal di Kabupaten/Kota di Jawa Barat yang diantaranya yaitu :

- Cluster 0: Kabupaten/Kota di Jawa Barat yang memiliki jumlah pendapatan bersih bulanan (pekerja informal) rendah dengan jumlah daerah kabupaten/kota di Jawa Barat yang dihasilkan terdapat 15, seperti Bogor, Sukabumi, Bandung, Majalengka, Bekasi, Kota Bekasi, Karawang, Subang, Purwakarta, Bandung Barat, Kota Bogor, Kota Sukabumi, Kota Bandung, Kota Cimahi, dan Kota Tasikmalaya
- Cluster 1: Kabupaten/Kota di Jawa Barat yang memiliki jumlah pendapatan bersih bulanan (pekerja informal) sedang, dengan dengan jumlah daerah kabupaten/kota di Jawa Barat yang dihasilkan terdapat 11, seperti Cianjur, Garut, Tasikmalaya, Ciamis, Kuningan, Cirebon, Indramayu, Sumedang, Pangandaran, Kota Cirebon, dan Kota Banjar.
- Cluster 2: Kabupaten/Kota di Jawa Barat yang memiliki jumlah pendapatan bersih bulanan (pekerja informal) tinggi, dengan jumlah daerah kabupaten/kota di Jawa Barat yang dihasilkan terdapat 1, seperti Kota Depok

Dengan demikian, hasil analisis pengelompokan dan pemodelan clustering mengindikasikan adanya perbedaan yang cukup besar dalam pendapatan bulanan bersih pekerja informal di daerah kabupaten/kota di Jawa Barat. Cluster 0 menunjukkan pendapatan rendah, Cluster 1 menunjukkan pendapatan sedang, sementara Cluster 2 menunjukkan pendapatan tinggi. Karena terbatasnya waktu dan sumber daya yang tersedia, peneliti menyadari bahwa hasil penelitian ini masih jauh dari kesempurnaan. Oleh karena itu, sebagai saran untuk penelitian selanjutnya, dapat dilakukan perbandingan dengan metode clustering lainnya.

## DAFTAR PUSTAKA

- [1] T. Hardiani, "Analisis Clustering Kasus Covid 19 di Indonesia Menggunakan Algoritma K-Means," *J. Nas. Pendidik. Tek. Inform.*, vol. 11, no. 2, pp. 156–165, 2022, doi: 10.23887/janapati.v11i2.45376.
- [2] G. Fajriansyah, "Analisis Daftar Pemilih Tetap Pada Hasil Rekapitulasi Kpu Berdasarkan Usia Menggunakan Algoritma K-Means (Studi Kasus : Kota Bandar Lampung)," *Electrician*, vol. 15, no. 1, pp. 39–53, 2021, doi: 10.23960/elc.v15n1.2147.
- [3] A. Mathematics, "Pekerja Anak "Joki Merpati Balap Sprint" Di Kelurahan Pakis Kecamatan Banyuwangi Kabupaten Banyuwangi," vol. 2, no. 9, pp. 1–23, 2016.
- [4] I. D. A. Dila Pariutami and I. M. Udiana, "Perlindungan Hukum Terhadap Hak Pekerja Rumah

- Tangga Yang Tidak Sesuai Dengan Perjanjian Di Bawah Tangan,” *Kertha Semaya J. Ilmu Huk.*, vol. 8, no. 8, p. 1291, 2020, doi: 10.24843/ks.2020.v08.i08.p15.
- [5] H. Murda, “Dampak Covid-19 Terhadap Penawaran Dan Pendapatan Pekerja Informal Daerah Istimewa Yogyakarta/Impact of Covid-19 on the Supply and Income of Informal Workers in the Special Region of Yogyakarta,” *J. Ekon. dan Pendidik.*, vol. 18, no. 2, pp. 171–182, 2022, doi: 10.21831/jep.v18i2.41237.
- [6] K. Gambut, T. Zonasi, P. Infrastruktur, D. I. Kebun, and R. Sriwijaya, “Publikasi penelitian terapan dan kebijakan,” vol. 1, no. 2, pp. 1–8, 2018.
- [7] N. S. Ramadhani and T. H. Siagian, “Determinan Status Pemanfaatan Jaminan Kesehatan pada Pekerja Informal di Provinsi Lampung Tahun 2020,” *Semin. Nas. Off. Stat.*, vol. 2022, no. 1, pp. 1063–1072, 2022, doi: 10.34123/semnasoffstat.v2022i1.1337.
- [8] N. Putu *et al.*, “Ketepatan Klasifikasi Penerima Beasiswa Stimik Stikom Bali Dengan Hybrid Self Organizing Maps Dan Algoritma K-Means Abstrak Data mining adalah penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar . Salah sa,” 2017.
- [9] M. Caron, P. Bojanowski, A. Joulin, and M. Douze, “Deep clustering for unsupervised learning of visual features,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11218 LNCS, pp. 139–156, 2018, doi: 10.1007/978-3-030-01264-9\_9.
- [10] C. R. Sistiani, Widodo, and B. P. Padhi, “Kinerja Algoritma Kmeans++ pada Pengelompokan Dokumen Teks Pendek pada Abstrak di Jurusan Teknik Elektro Fakultas Teknik UNJ,” *PINTER J. Pendidik. Tek. Inform. dan Komput.*, vol. 2, no. 1, pp. 39–44, 2018, doi: 10.21009/pinter.2.1.6.