

## Seleksi Penerimaan Beasiswa Dengan Metode K-Means Clustering Menggunakan Orange

### *Selection of Scholarship Recipients Using the K-Means Clustering Method Using Orange*

Silvana Nazuah<sup>1</sup>, Shofa Shofia Hilabi<sup>2</sup>, Agustia Hananto<sup>3</sup>, Baenil Huda<sup>4</sup>, Tukino<sup>5</sup>

<sup>1,2,3,4,5</sup> Prodi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Buana Perjuangan Karawang

Email: <sup>1</sup>[Si19.silvananazuah@mhs.ubpkarawang.ac.id](mailto:Si19.silvananazuah@mhs.ubpkarawang.ac.id), <sup>2</sup>[shofa.hilabi@ubpkarawang.ac.id](mailto:shofa.hilabi@ubpkarawang.ac.id)

<sup>3</sup>[Agustia.hananto@ubpkarawang.ac.id](mailto:Agustia.hananto@ubpkarawang.ac.id)

\*Koresponden

Diterima: 31 Desember 2023 | Direvisi: 30 Januari 2023 | Disetujui: 15 Februari 2023



This work is licensed under  
a Creative Commons Attribution 4.0  
International License.  
Copyright (c) 2023 JUSTINDO

#### ABSTRAK

Beasiswa adalah insentif keuangan yang ditawarkan oleh pemerintah, industri swasta, kedutaan ataupun Lembaga pendidikan. Proses penyeleksian penerimaan beasiswa secara tidak terintegrasi sistem butuh waktu yang lama, tidak efektif dan butuh ketelitian yang tinggi. Oleh karena itu diperlukan suatu metode untuk membantu pihak sekolah dan staf TU dalam menentukan siswa yang layak menerima beasiswa pada tahun berikutnya. Penelitian ini menggunakan metode K-Means clustering dengan tools orange data mining dan mengelompokkannya menjadi 3 cluster. Data yang digunakan adalah data sampel sebanyak 120 data yang telah direkap dari setiap program keahlian. Hasil yang diperoleh penelitian ini adalah setiap cluster akan memiliki anggota masing-masing yaitu cluster 1 dengan presentase 48% memiliki 58 anggota, cluster 2 dengan presentase 33% memiliki 39 anggota dan cluster 3 dengan presentase 19% memiliki 23 anggota. Penelitian ini bertujuan untuk mengetahui jumlah siswa yang termasuk berhak menerima beasiswa, tidak berhak menerima, dan dipertimbangkan.

**Kata kunci:** *Seleksi Penerimaan Beasiswa, K-Means Clustering, Orange Data Mining*

#### ABSTRACT

*Information Scholarships are financial incentives offered by the government, private industry, embassies or educational institutions. The process of selecting scholarship recipients in a system that is not integrated takes a long time, is not effective and requires high accuracy. Therefore a method is needed to assist the school and TU staff in determining students who are eligible to receive scholarships in the following year. This study used the K-Means clustering method with orange data mining tools and grouped them into 3 clusters. The data used is sample data of 120 data that have been recapitulated from each expertise program. The results obtained from this study are that each cluster will have its own members, namely cluster 1 with a percentage of 48% has 58 members, cluster 2 with a percentage of 33% has 39 members and cluster 3 with a percentage of 19% has 23 members. This study aims to determine the number of students who are eligible to receive scholarships, are not eligible to receive, and are considered*

**Keywords:** *Scholarship Selection, K-Means Clustering, Orange Data Mining*

## 1. Pendahuluan

Penelitian lain yang dilakukan oleh Syasya Sahira (Sahira et al., 2020), menggunakan metode ISO/IEC 27005:2018 untuk pengujian Aplikasi *E-Office* yang dikelola oleh PT Telkom Indonesia. ISO/IEC 27005:2018 digunakan untuk melakukan identifikasi jenis ancaman dan kemungkinan risiko yang terjadi pada Aplikasi *E-Office*. Dari hasil penelitian didapatkan bahwa ditemukan risiko dengan berbagai macam tingkatan mulai dari rendah, sedang, dan tinggi pada Aplikasi *E-Office*.

Beasiswa merupakan insentif keuangan yang ditawarkan oleh pemerintah, industri swasta, kedutaan ataupun Lembaga pendidikan untuk diberikan kepada siswa- siswi yang berprestasi ataupun orang tua yang berpenghasilan rendah. (Rahmah and Antares, 2022). data penerima beasiswa dipilih untuk memastikan siswa siswi yang berhak menerima beasiswa

Melihat data penerimaan siswa dari tahun ke tahun pasti berbeda pada setiap tahunnya apakah bertambah atau berkurang berdasarkan info yang diperoleh terdapat 855 data siswa pada tahun ajaran 2021/2022

Dalam Penelitian ini teknik mengumpulkan data yaitu berdasarkan hasil wawancara dan survey yang dilakukan dengan kepala sekolah dengan tujuan untuk dapat mendalami permasalahan secara keseluruhan sehingga penelitian ini memiliki konsep untuk memecahkan masalah. Selain wawancara teknik pengambilan data ini dengan melakukan survey kepada kepala sekolah SMKN 1 Karawang. Proses penyeleksian penerimaan beasiswa secara tidak terintegrasi sistem butuh waktu yang lama, tidak efektif dan butuh ketelitian yang tinggi. Sehingga diperlukan suatu metode untuk memudahkan pihak guru dan staf TU dalam mengidentifikasi apakah siswa yang layak menerima beasiswa pada tahun berikutnya, mengingat beasiswa tersebut sangat membantu melanjutkan pendidikan oleh keluarga yang berpenghasilan rendah dan anak yang putus sekolah karena keterbatasan ekonomi. (Darlinda and Utamajaya, 2022) Syarat yang harus dipenuhi seperti nilai rapor, jarak rumah, penghasilan orang tua, status orang tua, dan kondisi orang tua. (Dewi, Aryni and Umaidah, 2022)

Penelitian terdahulu yang dilakukan oleh Agustin Ely Rahayu & khoiril hikmah yaitu dengan judul "Penerapan *K-Means Clustering* untuk penentuan klasterisasi beasiswa bidikmisi mahasiswa" dengan data pengujian yang digunakan sebanyak 25 data sintesis dengan enam atribut yaitu nis, nama, IPK, jarak rumah, kehadiran mahasiswa, dan penghasilan orang tua. (Rahayu et al., 2019)

Nur afriani Manihuruk serta Muhammad Zarlis dengan penelitian yang berjudul "Penerapan Data mining dalam mengelompokkan calon penerima beasiswa dengan menggunakan algoritma *K-Means*" data set yang digunakan sebanyak 128 dan di olah menggunakan aplikasi *Rapid Miner 5.3*. Hasil penelitian menunjukkan C1 sebanyak 73Item, C2 sebanyak 30Item, C3 sebanyak 25Item. Dari hasil analisis diharapkan dapat membantu pihak sekolah memberikan keputusan yang tepat. (Manihuruk et al., 2020)

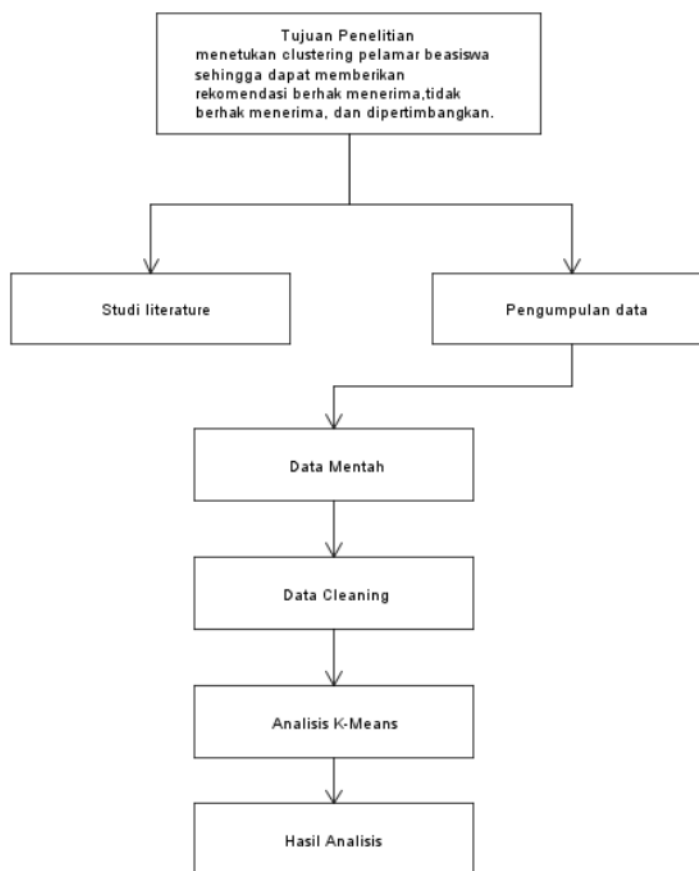
Penelitian sebelumnya mengenai *Clustering* data calon siswa baru menggunakan metode *K-Means* di SMKN Wahidin kota Cirebon" Oleh Asmana, Yudhistira dan Martanto. Berdasarkan data yang di terima pada tahun pelajaran 2020/2021 yaitu sebanyak 774 siswa. Mendapatkan hasil *Davies Bouldin Index* terbaiknya 0,168 measure types menggunakan operator *Mixed Measures* dan nilai  $k = 2$  (0,168),  $k = 3$  (0,248),  $k = 4$  (0,232),  $k = 5$  (0,26),  $k = 6$  (0,197). (Wijaya et al., 2022)

Penelitian terdahulu Oleh Darlinda dan Joy Nashar Utamajaya yang berjudul "Sistem Pendukung Keputusan Penerima beasiswa Program Indonesia Pintar Menggunakan Metode *K-Means Clustering*" Penelitian ini dilakukan dengan menggunakan perhitungan metode *K-means* dan alat *Rapid Miner Studio*. Temuan metode studi yang dikumpulkan dari 236 dataset menunjukkan bahwa 118 data mengikuti *cluster 0* yang berarti disarankan untuk disetujui atau mendapatkan beasiswa, sedangkan 118 data lainnya mengikuti *cluster 1* yang menunjukkan tidak direkomendasikan atau tidak memperoleh beasiswa (Darlinda and Utamajaya, 2022)

Febry Sandrian Sagala, Mugiarto, dan Wowon Priatna dengan penelitian yang berjudul "Algoritma *K-means clustering* untuk rekomendasi pemberian beasiswa bagi siswa berprestasi" pengolahan data tersebut menggunakan perhitungan manual *clustering* menggunakan algoritma *K-means* dengan beberapa tahapan. Jumlah *cluster* yang ditetapkan adalah 2 *cluster*. *Cluster 1* sebanyak 113 siswa untuk diberikan rekomendasi tidak layak dan *Cluster 2* sebanyak 32 siswa diberikan rekomendasi layak. (Sagala, Mugiarto and Priatna, 2021)

## 2. Metode Penelitian

Tahapan yang dilakukan untuk membuat rekomendasi berdasarkan data calon siswa yang akan menerima beasiswa maka memerlukan mekanisme untuk mengubah data tersebut menjadi hasil yang diharapkan penerima beasiswa. (Darlinda and Utamajaya, 2022) seperti pada gambar 1.1



Gambar 1. Tahapan Penelitian

### 1) Studi literature

Bagian yang dilakukan untuk membaca memahami dan mempelajari sebanyak banyaknya paper penelitian terdahulu yang relevan sebagai dasar referensi penelitian ini

### 2) Pengumpulan data

Dataset yang diperoleh adalah 855 record dan yang akan digunakan untuk diuji adalah data sampel sebanyak 120 data mewakili setiap program keahlian pada tahun ajaran 2021/2022 di SMKN 1 Karawang.

### 3) Data Cleaning

Data yang diperoleh baik dari database PPDB SMKN 1 Karawang maupun survei, terdapat data yang tidak valid atau salah ketik. Proses data cleaning ini yaitu membuang data yang tidak relevan dan tidak terformat karena dapat mengurangi kualitas akurasi hasil data mining.

### 4) Analisis *K-Means*

Data uji diolah dengan Algoritma *K-Means* menggunakan tools *Orange* data mining.

### 5) Hasil Analisis

Pada bagian ini akan memperoleh hasil berupa akurasi dan beberapa anggota yang akan menerima hasil untuk yang berhak menerima, tidak berhak menerima dan dipertimbangkan.

### 2.1. Algoritma K-Means

*K-Means* adalah metode yang sering digunakan untuk mengidentifikasi suatu kelompok dari kasus.(Nawawi, Sembiring and Erfina, 2021). Kelebihan dari penerapan *K-Means* yaitu mampu mengelompokkan objek besar serta meningkatkan kecepatan proses pengelompokan.(Parlambang and Fauziah, 2020)

Adapun di bawah ini merupakan uraian dari algoritma menggunakan metode K-Means clustering.

- 1) Tentukan sebanyak k (*cluster*) kelompok yang dibutuhkan.
- 2) Inisiasi *centroid* untuk setiap kelompok yang mempresentasikan kelompok.
- 3) Hitung jarak kedekatan antara suatu data dengan *centroid*, kemudian masukkan data tersebut pada kelompok *centroid* yang memiliki jarak terdekat.
- 4) Pilih *centroid* baru untuk masing-masing kelompok dari anggota kelompok itu sendiri dan lakukan perhitungan kedekatan antara suatu data dengan *centroid*.
- 5) Lakukan perulangan langkah-langkah sebelumnya hingga tidak ada perubahan anggota untuk semua kelompok(Triandini, Defit and Nurcahyo, 2021)

### 2.2. Orange

*Orange* adalah alat visualisasi data, platform pembelajaran mesin, dan alat peringkasan data. Fitur program visual front-end ini memungkinkan analisis eksplorasi data kuantitatif yang cepat serta visualisasi data interaktif. Untuk menganalisis dan memvisualisasikan data, melihat tabel data, dan menjadi lebih produktif, *Orange* adalah platform yang sangat direkomendasikan karena menyediakan platform sumber terbuka dengan keamanan yang kuat dan potensi untuk memperluas fungsionalitas di seluruh bidang sains.(Sujaini, 2019)

## 3. Hasil dan Pembahasan

Penelitian ini menggunakan data sampel untuk pengolahan data sampel yaitu menggunakan *tools orange* data mining meliputi tahap-tahap yang dijabarkan sebagai berikut:

### 3.1. Tahap Pengumpulan data

Sumber data yang diperoleh dari nilai rapor siswa SMKN 1 Karawang kelas 11 Jurusan Teknik Elektronika Industri tahun ajaran 2021/2022. Dilakukan dengan menggunakan teknik wawancara secara langsung kepada pihak sekolah ataupun guru.

### 3.2. Cleaning Data

Data siswa kelas 11 tahun ajaran 2021/2022 hanya diambil data dengan kategori nilai rapor, jarak rumah ke sekolah, pendapatan orang tua, dan tanggungan orang tua sebagai atribut yang akan digunakan meliputi *meta* data dan target. Berikut adalah data hasil proses cleaning.

Tabel 1 Data Hasil Proses Cleaning

No	Nama	NR	JR	PO	TO
1	Siswa 01	80	500	2500000	4
2	Siswa 02	80	200	1200000	5
3	Siswa 03	85	1000	1600000	4
4	Siswa 04	79	200	1500000	3
5	Siswa 05	85	200	1200000	4
6	Siswa 06	77	300	2000000	5
7	Siswa 07	83	460	1500000	4
8	Siswa 08	78	670	2000000	4
9	Siswa 09	82	300	1800000	4
10	Siswa 10	85	200	1500000	5

11	Siswa 11	70	400	1800000	3
12	Siswa 12	87	500	1600000	5
13	Siswa 13	88	100	1200000	4
14	Siswa 14	86	200	1500000	5
15	Siswa 15	85	350	1500000	4
16	Siswa 16	90	700	1600000	4
17	Siswa 17	89	670	1800000	3
18	Siswa 18	80	560	2000000	5
19	Siswa 19	81	430	3000000	5
20	Siswa 20	75	674	1600000	3
.....	.....	.....	.....	.....	.....
120	Siswa 120	80	740	1500000	4

Pada tabel 1 terdapat empat (4) atribut yaitu Nama yang di inisialisasikan mulai dari siswa 01 sampai siswa 120, Nilai rapot (NR), Jarak rumah (JR), Penghasilan orang tua (PO), Kondisi orang tua yang diambil dari siswa-siswi di SMKN 1 Karawang.

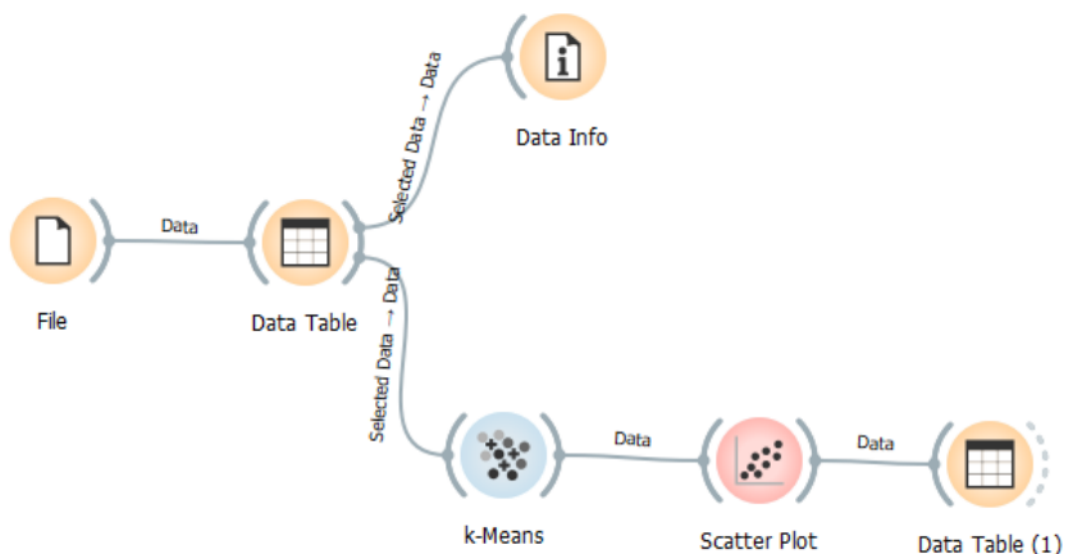
### 3.3. Tahap K-Means Clustering

Dengan jumlah data sampel yaitu 120 data yang sudah dipaparkan sebelumnya maka pada tahap proses *K-means* clustering yaitu sebagai berikut:

- 1) Jumlah *k* (*Cluster*) yang ditentukan secara random yaitu 3 *cluster* dengan kategori berhak menerima, tidak berhak menerima dan dipertimbangkan.
- 2) Penentuan *centroid* awal dilakukan secara acak otomatis oleh mesin pada *tools Orange*. Data yang digunakan adalah data sampel yaitu 120 dan sudah otomatis di proses pada aplikasi *Orange*.

### 3.4. Pengolahan Data

Pada pengolahan data yang menggunakan *tools orange* data mining. proses pengolahan serta workflow terdapat pada gambar 1.2



Gambar 2. Proses Pengolahan Data Melalui Orange

1) Data Table

Data table adalah representasi dalam memori dari table database tunggal yang memiliki kumpulan kolom dan baris. Data table hanya mengambil satu *TableRow* pada satu waktu *Dataset*.

2) Scatter Plot

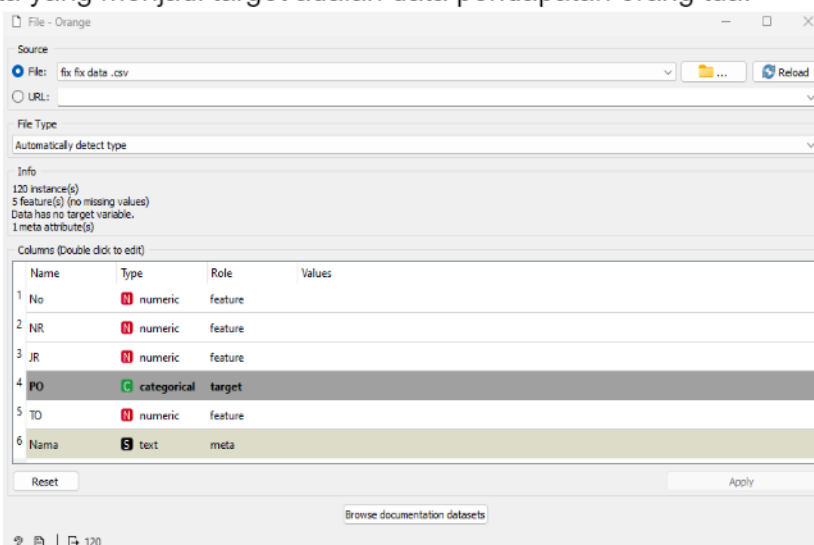
*Scatter plot* merupakan penggunaan titik titik untuk merepresentasikan nilai dari dua variabel numerik terpisah. Posisi setiap titik pada sumbu X dan Y menunjukkan nilai untuk titik data individual. *Scatter plot* juga digunakan untuk memeriksa hubungan antar variabel.

3) Data Info

Data info adalah *widget* yang menyajikan informasi tentang ukuran dataset size, *feature*, *target*, *atribut meta*, dimana data disimpan dan menghasilkan report.

3.5. Penginputan Data

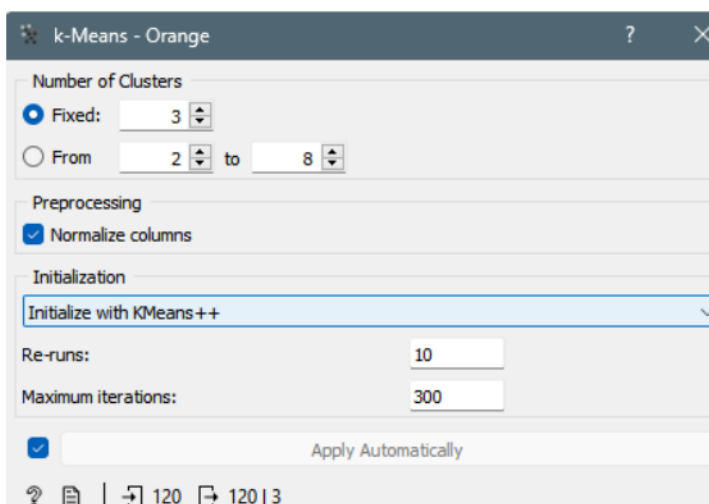
Data sampel yang digunakan adalah data yang sudah di rekap dari keseluruhan program keahlian sebanyak 120 data yang menjadi target adalah data pendapatan orang tua.



Gambar 3. Penginputan Data di Orange

3.6. Menentukan Jumlah Cluster

Penentuan nilai hasil cluster dilihat dari jarak terdekat antara objek data. Sebelum menghitung jarak terdekatnya, harus menentukan terlebih dahulu jumlah *cluster* centroid. Pada penelitian ini menggunakan *Cluster* 1 (Berhak menerima) *cluster* 2 (tidak berhak menerima) dan *cluster* 3 (dipertimbangkan).



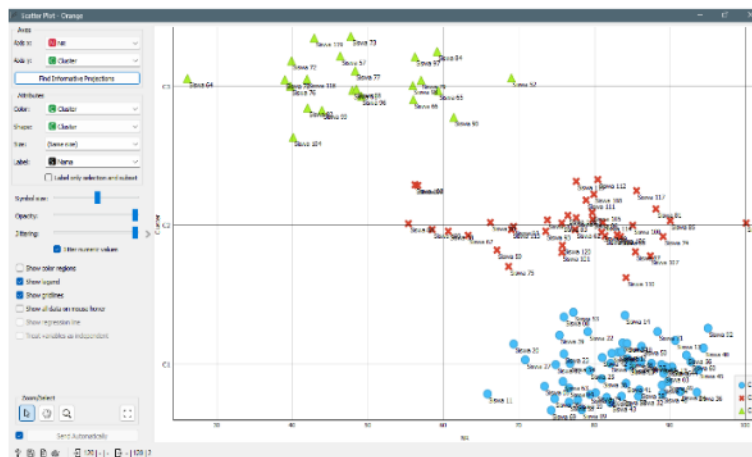
Gambar 4. Proses Menentukan Cluster

PO	Nama	Cluster	Silhouette	No	NR	JR	IO
2290000	Siswa 01	C1	0,643681	1	30	500	4
1200000	Siswa 02	C1	0,652133	2	30	1200	15
1690000	Siswa 03	C1	0,320670	3	35	1000	4
1590000	Siswa 04	C1	0,644223	4	79	1200	13
1200000	Siswa 05	C1	0,661870	5	35	500	4
2000000	Siswa 06	C1	0,640770	6	77	500	1
1500000	Siswa 07	C1	0,55344	7	35	600	4
2000000	Siswa 08	C1	0,368440	8	78	670	4
1800000	Siswa 09	C1	0,668441	9	32	500	4
1500000	Siswa 10	C1	0,556107	10	35	200	1
1200000	Siswa 100	C2	0,504637	100	85	600	13
1500000	Siswa 101	C2	0,383379	101	70	496	5
2500000	Siswa 102	C2	0,523954	102	80	633	5
2000000	Siswa 103	C2	0,307094	103	37	733	4
1500000	Siswa 104	C1	0,394900	104	38	754	3
2000000	Siswa 105	C2	0,504965	105	75	630	5
2000000	Siswa 106	C2	0,522372	106	70	678	4
1000000	Siswa 107	C2	0,505512	107	86	436	12
2800000	Siswa 108	C2	0,300370	108	70	500	7
2000000	Siswa 109	C2	0,548365	109	37	540	17
1800000	Siswa 11	C1	0,610332	11	70	400	2
2500000	Siswa 110	C2	0,506240	110	85	777	3
2000000	Siswa 111	C2	0,5882	111	75	676	5
4000000	Siswa 112	C2	0,524464	112	70	676	4
2500000	Siswa 113	C2	0,578677	113	74	600	12
1500000	Siswa 114	C2	0,543164	114	70	700	2
1000000	Siswa 115	C2	0,528620	115	76	674	12
1500000	Siswa 116	C2	0,307907	116	70	546	7
2000000	Siswa 117	C2	0,506622	117	80	614	4

Gambar 5. Hasil Cluster 1, 2, 3

### 3.7. Melakukan Plotting

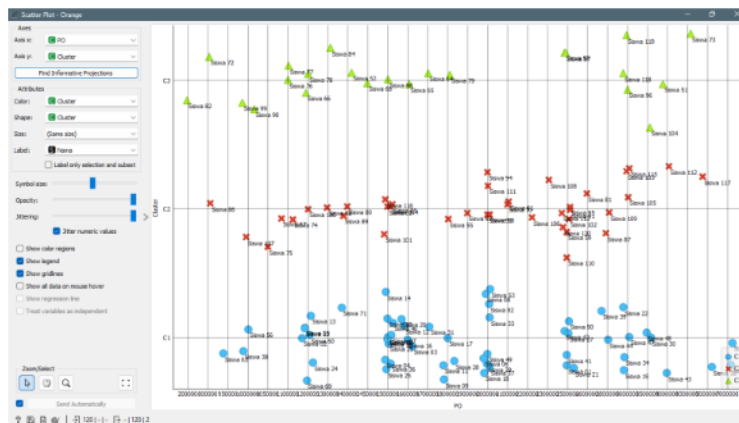
Setelah menentukan jumlah cluster, kemudian akan melakukan metode *Plotting*. Plotting yang dilakukan adalah membandingkan nilai rapor (NR) dengan pendapatan orang tua (PO) pada masing-masing *cluster*



Gambar 6. ScatterPlot Widget Berdasarkan NR

Berdasarkan hasil plotting nilai rapor menggunakan scatterplot widget yang menyediakan visualisasi scatterplot 2 dimensi untuk atribut yang kontinu dan bernilai diskrit. Untuk atribut jittering yang berguna menghindari tumpang tindih titik yang memiliki nilai yang sama. Kemudian, kedua sumbu dapat dilihat dari gambar diatas, sumbu axis x diterapkan untuk nilai rapor dan sumbu axis y untuk ditempatkan di cluster yang telah ditentukan.

Berdasarkan hasil plotting nilai rapor menggunakan scatterplot widget yang menyediakan visualisasi scatterplot 2 dimensi untuk atribut yang kontinu dan bernilai diskrit. Untuk atribut jittering yang berguna menghindari tumpang tindih titik yang memiliki nilai yang sama. Kemudian, kedua sumbu dapat dilihat dari gambar diatas, sumbu axis x diterapkan untuk nilai rapor dan sumbu axis y untuk ditempatkan di cluster yang telah ditentukan.



Gambar 7. ScatterPlot Widget berdasarkan PO

Setelah berakhir proses iterasi yang disebabkan karena sudah tidak ada lagi nilai yang berubah, maka proses K-Means dinyatakan selesai dan mendapatkan hasil sebagai berikut:

Tabel 2. Pengelompokkan Cluster 1

No	Nama	Cluster 1
1	Siswa 01	C1
2	Siswa 02	C1
3	Siswa 03	C1
4	Siswa 04	C1
5	Siswa 05	C1
6	Siswa 06	C1
7	Siswa 07	C1
8	Siswa 08	C1
9	Siswa 09	C1
10	Siswa 10	C1
....	.....	...
56	Siswa 69	C1
57	Siswa 71	C1
58	Siswa 92	C1

Tabel 2 Pengelompokkan Cluster 2

No	Nama	Cluster 2
1	Siswa 120	C2
2	Siswa 105	C2
3	Siswa 106	C2
4	Siswa 107	C2
5	Siswa 108	C2
6	Siswa 109	C2
7	Siswa 110	C2
8	Siswa 111	C2
9	Siswa 112	C2
10	Siswa 113	C2
...	...	...



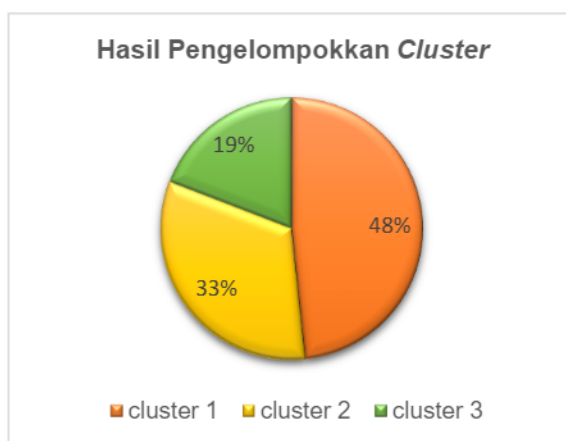
37	Siswa 58	C2
38	Siswa 59	C2
39	Siswa 55	C2

Tabel 3 Pengelompokan Cluster 3

No	Nama	Cluster 3
1	Siswa 51	C3
2	Siswa 52	C3
3	Siswa 57	C3
4	Siswa 64	C3
5	Siswa 65	C3
6	Siswa 66	C3
7	Siswa 68	C3
8	Siswa 72	C3
9	Siswa 73	C3
10	Siswa 76	C3
...	...	...
21	Siswa 104	C3
22	Siswa 118	C3
23	Siswa 119	C3

Berdasarkan hasil pengelompokan dengan metode *K-means clustering* menggunakan *tools orange* data mining didapatkan hasil *Cluster 1* yang berhak menerima berdasarkan pendapatan orang tua sebanyak 58 siswa, *Cluster 2* tidak berhak menerima sebanyak 39 siswa dan *Cluster 3* untuk dipertimbangkan berdasarkan variabel lainnya sebanyak 23 siswa.

Hasil diagram lingkaran menjadi irisan yang mewakili nilai numerik suatu data guna untuk merepresentasikan data dengan variabel numerik tunggal seperti gambar 1.8



Gambar 8. Diagram Lingkaran

#### 4. Kesimpulan

Analisis menggunakan aplikasi *orange* menunjukkan bahwa terdapat 58 anggota *cluster 1* dengan presentase sebesar 48% yang berhak menerima beasiswa, 39 anggota *cluster 2* dengan presentase sebesar 33% yang tidak berhak menerima beasiswa, dan 23 anggota yang dipertimbangkan pada *cluster 3* dengan presentase sebesar 19%. Berdasarkan hasil analisis dapat disimpulkan bahwa terdapat beberapa anggota disetiap *cluster* untuk direkomendasikan berhak menerima beasiswa, tidak berhak menerima dan untuk dipertimbangkan.

## Daftar Pustaka

- Darlinda and Utamajaya, N. (2022) 'Sistem Pendukung Keputusan Penerima Beasiswa Program Indonesia Pintar Menggunakan Metode Algoritma K-Means Clustering', *JURIKOM (Jurnal Riset Komputer)*, 9(2), pp. 167–175. doi:10.30865/jurikom.v9i2.3971.
- Dewi, F.P., Aryni, P.S. and Umaidah, Y. (2022) 'Implementasi Algoritma K-Means Clustering Seleksi Siswa Berprestasi Berdasarkan Keaktifan dalam Proses Pembelajaran', *JISKA (Jurnal Informatika Sunan Kalijaga)*, 7(2), pp. 111–121. doi:10.14421/jiska.2022.7.2.111-121.
- Manihuruk, N.A. et al. (2020) 'Penerapan Data Mining Dalam Mengelompokkan Calon Penerima Beasiswa Dengan Menggunakan Algoritma K-Means', *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, 4(1), pp. 29–34. doi:10.30865/komik.v4i1.2575.
- Nawawi, M.S., Sembiring, F. and Erfina, A. (2021) 'Implementasi Algoritma K-Means Clustering Menggunakan Orange Untuk Penentuan Produk Busana Muslim Terlaris', ... *Teknologi Informasi dan ...*, pp. 789–797. Available at: <http://prosiding.unipma.ac.id/index.php/SENATIK/article/view/1837%0Ahttp://prosiding.unipma.ac.id/index.php/SENATIK/article/viewFile/1837/1723>.
- Parlambang, B. and Fauziah (2020) 'Implementasi Algoritma K-Means Dalam Proses Penilaian Kuesioner Kepada Dosen Guna Mendukung Kepuasan Mahasiswa Terhadap Dosen', *Jurnal Ilmiah Teknologi dan Rekayasa*, 25(2), pp. 161–173. doi:10.35760/tr.2020.v25i2.2719.
- Rahayu, A.E. et al. (2019) 'Penerapan K-Means Clustering Untuk Penentuan Klasterisasi Beasiswa Bidikmisi Mahasiswa', *ILKOMNIKA: Journal of Computer Science and Applied Informatics*, 1(2), pp. 82–86. doi:10.28926/ilkomnika.v1i2.23.
- Rahmah, S.A. and Antares, J. (2022) 'Klasterisasi Seleksi Mahasiswa Calon Penerima Beasiswa Yayasan Menggunakan K-Means Clustering', *I N F O R M a T I K a*, 13(2), p. 25. doi:10.36723/juri.v13i2.282.
- Sagala, F.S., Mugiarto, M. and Priatna, W. (2021) 'Algoritma K-Means Clustering Untuk Menentukan Rekomendasi Pemberian Beasiswa Bagi Siswa Berprestasi', *Journal of Students' Research in Computer Science*, 2(2), pp. 111–120. doi:10.31599/jsrscs.v2i2.840.
- Sujaini, H. (2019) 'Klasifikasi Citra Alat Musik Tradisional dengan Metode k-Nearest Neighbor, Random Forest, dan Support Vector Machine', *Jurnal Sistem Informasi Bisnis*, 9(2), p. 185. doi:10.21456/vol9iss2pp185-191.
- Triandini, M., Defit, S. and Nurcahyo, G.W. (2021) 'Data Mining dalam Mengukur Tingkat Keaktifan Siswa dalam Mengikuti Proses Belajar pada SMP IT Andalas Cendekia', *Jurnal Informasi dan Teknologi*, 3, pp. 167–173. doi:10.37034/jidt.v3i3.120.
- Wijaya, Y.A. et al. (2022) 'K-Means Di Sekolah Menengah Kejuruan Wahidin Kota Cirebon', 6(2), pp. 552–559.