

## **PEMANFAATAN RAPID MINER STUDIO 8.2 UNTUK PENGELOMPOKAN DATA PENJUALAN AKSESORIS MENGUNAKAN ALGORITMA K-MEANS**

**Mardalius**

Program Studi Sistem Informasi, STMIK Royal Kisaran

Jl. Prof. M. Yamin 173 Kisran, Sumatera Utara 21222

Email : [mardalius@royal.ac.id](mailto:mardalius@royal.ac.id)

**Abstract:** At Rafadel Store Acc various types of accessories available that are sold in the Store. Of the various types of accessories are sold certainly not all of them sold and also there are less salable and some are never sold at all. With this problem then we need to do calculations to determine or classify which categories of accessories that sell, sell and sell unsold, in the process of grouping it will be used a method of clustering using K-Means Clustering Algorithm as a method of calculation manually and in the implementation then used a Data Mining software using RapidMiner Studio version 8.2. With the RapidMiner Studio application store owners can see the results of the grouping which accessories are the most salable, sellable and less salable. So, if there are products that do not sell, shop owners can look for other alternatives to unsold accessories that can be sold. The method used in data collection is observation and interview to Rafadel Acc shop owner.

**Keywords:** *Data Mining, K-Means, Clustering, Store, Accessories*

**Abstrak:** Pada toko Rafadel Acc menjual berbagai jenis aksesoris yang tersedia yang dijual di toko tersebut. Dari berbagai jenis aksesoris yang dijual tentu tidak semuanya yang laku terjual dan juga ada yang kurang laku serta ada juga yang tidak pernah terjual sama sekali. Dengan adanya masalah ini maka kita perlu melakukan perhitungan untuk menentukan atau mengelompokkan mana kategori aksesoris yang laku, kurang laku dan tidak laku terjual, dalam proses pengelompokan maka akan digunakan sebuah metode pengelompokan menggunakan Algoritma K-Means Clustering sebagai metode perhitungan secara manual dan dalam implementasinya maka digunakan sebuah software Data Mining menggunakan RapidMiner Studio versi 8.2. Dengan adanya aplikasi RapidMiner Studio ini pemilik toko dapat melihat hasil pengelompokan aksesoris mana yang paling laku, laku dan kurang laku. Maka, bila terdapat produk yang tidak laku, pemilik toko dapat mencari alternative lain agar aksesoris yang tidak laku dapat menjadi laku. Metode yang digunakan dalam pengumpulan data adalah observasi dan wawancara kepada pemilik toko Rafadel Acc.

**Kata kunci:** *Data Mining, K-Means, Clustering, Penjualan, Aksesoris*

## PENDAHULUAN

Rafadel Acc merupakan salah satu bidang usaha yang bergerak di bidang penjualan aksesoris mulai dari aksesoris laki-laki, perempuan serta anak-anak juga tersedia pada toko tersebut. Di mana toko ini setiap harinya harus memenuhi kebutuhan konsumen yang lagi trend saat ini. Untuk dapat melakukan hal tersebut, maka membutuhkan sumber informasi yang cukup banyak untuk dapat dianalisis lebih lanjut. Pada Toko Rafadel Acc terdapat beberapa permasalahan yang selalu muncul mengenai penjualan. Berdasarkan hal tersebut maka diharapkan bisa mempermudah pihak Toko Rafadel Acc dalam menemukan informasi dari kumpulan-kumpulan data yang berguna bagi Toko Rafadel Acc. Tidak hanya itu penulis juga mengharapkan bisa memberikan rangsangan kepada pengelola data supaya lebih aktif dalam menggali informasi yang dibutuhkan oleh Toko Rafadel Acc, sehingga bisa menjawab kebutuhan dari Toko tersebut. Penulis berharap dapat membantu memberikan informasi mengenai hasil Clustering data penjualan. Dengan demikian, dibutuhkan analisis yang komputerisasi menggunakan software data mining yang menunjang arus data dan informasi sesuai dengan kebutuhan dari proses-proses tersebut.

Knowledge discovery in Database (KDD) didefinisikan sebagai ekstraksi informasi potensial, implisit dan tidak dikenal dari sekumpulan data. Proses knowledge discovery melibatkan hasil dari proses data mining (proses mengekstrak kecenderungan pola suatu data), kemudian mengubah hasilnya secara akurat menjadi informasi yang mudah dipahami (Ronald, 2015). Proses knowledge discovery in database (KDD) secara garis besar terdiri dari Data Selection, Pre-processing/Cleaning, Transformation, Data mining, dan Interpretation/Evaluation.

Algoritma K-Means merupakan algoritma non hirarki yang berasal dari metode data clustering. Menurut Eko Prasetyo (2012) mengatakan bahwa metode K-Means ini mempartisi data kedalam kelompok sehingga data berkarakteristik sama dimasukkan kedalam satu kelompok yang sama dan data yang berkarakteristik berbeda dikelompokkan kedalam kelompok yang lain. Adapun tujuan dari pengelompokan data ini adalah untuk meminimalkan fungsi objektif yang diset dalam proses pengelompokan, yang pada umumnya berusaha meminimalkan variasi didalam suatu kelompok dan memaksimalkan variasi antar kelompok.

Pemilihan metode K-Means dikarenakan metode ini harus menggunakan data fisik tidak abstrak dan bersifat jelas, hal ini sesuai dengan data yang akan digunakan pada permasalahan di dalam pengelompokan penjualan aksesoris. Selain itu, metode ini bersifat fleksibel sebab pengguna dapat menentukan jumlah cluster yang akan dibuat.

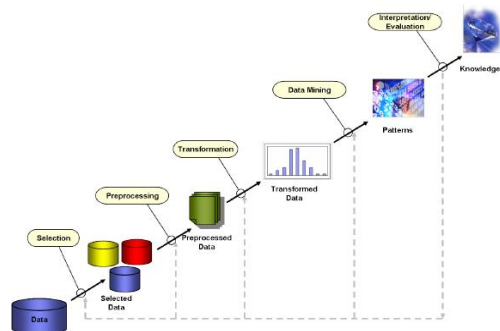
## METODOLOGI

Metodologi dipergunakan oleh penulis untuk menganalisa, mengerjakan dan mengatasi masalah yang dihadapi. Kerangka teoritis atau kerangka ilmiah merupakan metode-metode ilmiah yang akan diterapkan dalam pelaksanaan penelitian. Pada kerangka kerja penelitian yang digunakan yaitu mempelajari literatur, mengumpulkan data, menganalisa data, menganalisa metode K-Means, implementasi software, menguji hasil, dan menganalisa hasil.

### 1. Knowledge Discovery in Database (KDD)

*Knowledge discovery in database* (KDD) adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan pola atau

hubungan dalam *set* data berukuran besar (Buulolo, 2013).



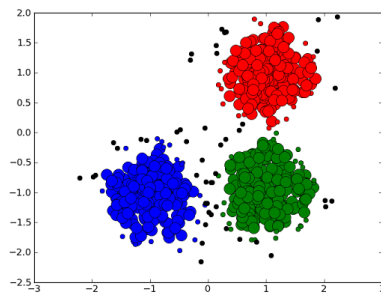
Gambar 1. Aliran Informasi Dalam Data Mining

#### a. Data Mining

Menurut Widodo (2013) *Data Mining* adalah analisa terhadap data untuk menemukan hubungan yang jelas serta menyimpulkannya yang belum diketahui sebelumnya dengan cara terkini dipahami dan berguna bagi pemilik data tersebut.

#### b. Clustering

Menurut Widodo (2013) *Clustering* atau klasifikasi adalah metode yang digunakan untuk membagi rangkaian data menjadi beberapa group berdasarkan kesamaan-kesamaan yang telah ditentukan sebelumnya.



Gambar 2. Clustering

#### c. Algoritma K-Means

Menurut Eko Prasetyo (2012) mengatakan bahwa metode *K-Means* ini mempartisi data kedalam kelompok sehingga data berkarakteristik sama dimasukkan kedalam satu kelompok yang

sama dan data yang berkarakteristik berbeda dikelompokkan kedalam kelompok yang lain.

Langkah-langkah melakukan clustering dengan metode *K-Means* adalah sebagai berikut:

1. Tentukan nilai  $k$  sebagai jumlah kluster yang ingin dibentuk.
2. Inisialisasi  $k$  pusat *cluster* ini bisa dilakukan dengan berbagai cara, namun yang paling sering dilakukan adalah dengan cara random yang diambil dari data yang ada.
3. Menghitung jarak setiap data *input* terhadap masing – masing *centroid* menggunakan rumus jarak *Euclidean* (*Euclidean Distance*) hingga ditemukan jarak yang paling dekat dari setiap data dengan *centroid*. Berikut adalah persamaan *Euclidean Distance* :

$$d(xi, \mu_j) = \sqrt{\sum (xi - \mu_j)^2}$$

Dimana :

$d$  : titik dokumen

$xi$  : data kriteria

$\mu_j$  : *centroid* pada cluster ke- $j$

4. Mengklasifikasikan setiap data berdasarkan kedekatannya dengan *centroid* (jarak terkecil).
5. Memperbaharui nilai *centroid*. Nilai *centroid* baru diperoleh dari rata-rata *cluster* yang bersangkutan dengan menggunakan rumus:

$$\mu_j(t+1) = \frac{1}{N_{sj}} \sum_{j \in s_j} x_j$$

Dimana:

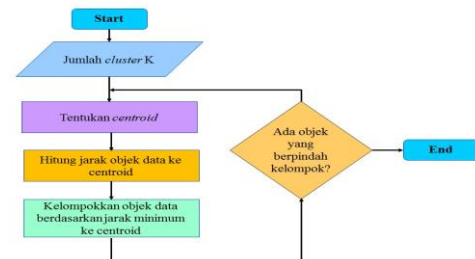
$\mu_j(t+1)$  : *centroid* baru pada iterasi ke  $(t+1)$

$N_{sj}$  : banyak data pada cluster  $s_j$ ,

6. Melakukan perulangan dari langkah 2 hingga 5, sampai anggota tiap *cluster* tidak ada yang berubah.

Jika langkah 6 telah terpenuhi, maka nilai pusat *cluster* ( $\mu_j$ ) pada iterasi terakhir

akan digunakan sebagai parameter untuk menentukan klasifikasi data.



Gambar 3. Flowchart Algoritma K-Means Clustering

## HASIL DAN PEMBAHASAN

Pada proses pengumpulan data ada tiga parameter yang akan digunakan dalam pengolahan data yaitu stok awal, terjual dan stok akhir. Pada penelitian ini data yang dikumpulkan yaitu data penjualan aksesoris pada tahun 2017 Data tersebut diambil dari Toko Rafadel Acc.

Tabel 1. Sampel Data Penjualan Aksesoris

KD	AKSESORIS	SA	ST	SAK
AC01	Gelang	90	58	32
AC02	Anting Cewek	40	33	7
AC03	Anting Xuping	35	12	23
AC04	Bandana Anak	70	47	23
AC05	Aseton	58	28	30
AC06	Bandana Jilbab	48	45	3
AC07	Bros Kain	36	32	4
AC08	Bulu Mata Palsu	144	87	57
AC09	Bando	48	12	36
AC10	Boneka	36	23	13
AC11	Bros	132	97	35
AC12	Bunga	40	36	4
AC13	Cincin Anak	38	29	9
AC14	Celengan	24	15	9
AC15	Cincin Titanium	45	40	5
AC16	Gelang Anak	120	78	42
AC17	Gelang Dewasa	50	47	3
AC18	Gantungan HP	60	27	33
AC19	Hena	58	15	43
AC20	Gantungan Kunci	36	35	1
AC21	Gelang Kaki	48	15	33
AC22	Ikat Pinggang	36	17	19
AC23	Jepitan	180	115	65
AC24	Jam Tangan	65	27	38
AC25	Jam Weker	75	27	48
AC26	Kotak Cincin	96	39	57

Ket : SA = Stok Awal

ST = Stok Terjual

SAK = Stok Akhir

Data yang sudah dijadikan sampel akan dilakukan pengolahan data dengan proses *clustering* dengan menggunakan algoritma *K-Means* sehingga didapatkanlah hasil pengelompokan data yang diinginkan. Adapun langkah dalam *cluster* dengan algoritma *K-Means* yaitu :

### 1. Menentukan Jumlah Cluster

Menentukan jumlah *cluster* yang digunakan pada data penjualan aksesoris sebanyak 3 *cluster* diantaranya Paling laku, laku, kurang laku berdasarkan data penjualan dalam 1 tahun

### 2. Menentukan Centroid

Penentuan pusat awal *cluster* (*centroid*) ditentukan secara *random* atau acak yang diambil dari data yang ada. Nilai *cluster* 0 diambil dari baris ke-20, nilai *cluster* 1 pada baris ke-4, nilai *cluster* 2 pada baris ke-8.

Tabel 2. Centroid Awal

C	AKSESORIS	SA	ST	SAK
C0	Gantungan Kunci	36	35	1
C1	Bandana Anak	70	47	23
C2	Bulu Mata Palsu	144	87	57

### 3. Menghitung Jarak dari Centroid

Menghitung jarak antara titik *centroid* dengan titik tiap objek dengan menggunakan *Euclidian Distance*. Adapun penghitungan *centroid* awal secara manual. Perhitungannya adalah sebagai berikut :

$$D(i,f) = \sqrt{(X_{i1} - X_{1j})^2 + (X_{i2} - X_{2j})^2 + \dots + (X_{ik} - X_{kj})^2}$$

Tabel 3. Perhitungan Jarak dan Pengelompokan Data Iterasi Ke-1

KD	DC0	DC1	DC2	C0	C1	C2
AC01	66.38	24.54	66.20	0	1	0
AC02	7.48	36.77	127.40	1	0	0
AC03	31.84	49.50	136.61	1	0	0
AC04	42.24	0.00	90.73	0	1	0
AC05	37.07	23.54	107.73	0	1	0
AC06	15.75	29.80	117.88	1	0	0
AC07	4.24	41.74	132.28	1	0	0
AC08	132.30	90.73	0.00	0	0	1
AC09	43.57	43.34	123.62	0	1	0
AC10	16.97	42.80	133.03	1	0	0
AC11	119.23	80.55	26.98	0	0	1

AC12	5.10	37.18	127.38	1	0	0
AC13	10.20	39.29	130.02	1	0	0
AC14	24.66	57.76	147.95	1	0	0
AC15	11.05	31.59	121.30	1	0	0
AC16	102.89	61.82	29.70	0	0	1
AC17	18.55	28.28	115.55	1	0	0
AC18	40.79	24.49	105.98	0	1	0
AC19	51.46	39.60	113.03	0	1	0
AC20	0.00	42.24	132.30	1	0	0
AC21	39.60	40.10	122.38	1	0	0
AC22	25.46	45.52	134.19	1	0	0
AC23	176.73	135.97	46.30	0	0	1
AC24	47.69	25.50	101.00	0	1	0
AC25	61.60	32.40	91.88	0	1	0
AC26	82.17	43.54	67.88	0	1	0

Dari hasil perhitungan di atas didapatkan pusat *cluster* baru seperti Tabel sebagai berikut :

Tabel 4. Centroid Baru Iterasi Ke-1

C	SA	ST	SAK
C0	39.38	29.15	10.23
C1	68.89	31.11	37.78
C2	144.00	94.25	49.75

Iterasi selanjutnya melakukan perhitungan lagi dengan menggunakan titik centroid yang baru Iterasi Ke-1

Tabel 5. Perhitungan Jarak dan Pengelompokan Data Iterasi Ke-2

KD	DC0	DC1	DC2	C0	C1	C2
AC01	62.20	34.67	67.42	0	1	0
AC02	5.06	42.26	128.04	1	0	0
AC03	21.83	41.62	139.15	1	0	0
AC04	37.67	21.73	91.78	0	1	0
AC05	27.18	13.74	110.34	0	1	0
AC06	19.44	42.88	117.59	1	0	0
AC07	7.64	47.16	132.79	1	0	0
AC08	128.37	95.58	10.25	0	0	1
AC09	32.13	28.37	127.16	0	1	0
AC10	7.54	41.97	134.50	1	0	0
AC11	117.45	91.28	19.21	0	0	1
AC12	9.28	44.72	127.68	1	0	0
AC13	1.85	42.27	130.97	1	0	0
AC14	20.94	55.70	149.47	1	0	0
AC15	13.29	41.52	121.44	1	0	0
AC16	99.47	69.49	30.00	0	0	1
AC17	21.99	42.65	115.13	1	0	0
AC18	30.79	10.90	108.90	0	1	0
AC19	40.26	20.13	117.14	0	1	0
AC20	11.44	49.49	132.48	1	0	0
AC21	28.16	26.81	125.61	0	1	0
AC22	15.36	40.42	136.30	1	0	0
AC23	173.62	141.86	44.26	0	0	1
AC24	37.84	5.66	104.41	0	1	0
AC25	51.96	12.60	96.37	0	1	0
AC26	74.10	34.16	73.55	0	1	0

Dari *centroid* baru iterasi ke-2, dilakukan perhitungan kembali, sehingga di dapatkan hasil Tabel Perhitungan Jarak dan Pengelompokan Data Iterasi Ke-1 dimana C0 memiliki 12 data, C1 memiliki 10 data, C2 memiliki 4 data.

Berdasarkan dari perhitungan di atas bahwa pusat *cluster* baru iterasi ke-1 seperti Tabel sebagai berikut :

Tabel 6. Centroid Baru Iterasi Ke-2

C	SA	ST	SAK
C0	38.67	30.33	8.33
C1	66.80	29.50	37.30
C2	144.00	94.25	49.75

Dari hasil *centroid* yang didapatkan pada iterasi ke-2, kemudian lakukan lagi perhitungan yang sama sampai data tiap *cluster* tidak ada lagi yang berubah.

Tabel 7. Perhitungan Jarak dan Pengelompokan Data Iterasi Ke-3

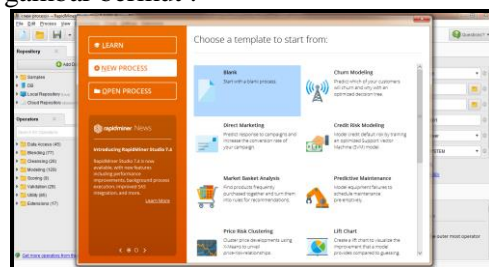
KD	DC0	DC1	DC2	C0	C1	C2
AC01	62.93	37.13	67.42	0	1	0
AC02	3.27	40.60	128.04	1	0	0
AC03	23.76	39.01	139.15	1	0	0
AC04	38.40	22.82	91.78	0	1	0
AC05	29.13	11.53	110.34	0	1	0
AC06	18.18	42.07	117.59	1	0	0
AC07	5.35	45.43	132.79	1	0	0
AC08	129.13	98.26	10.25	0	0	1
AC09	34.48	25.72	127.16	0	1	0
AC10	9.09	39.77	134.50	1	0	0
AC11	117.76	93.88	19.21	0	0	1
AC12	7.26	43.24	127.68	1	0	0
AC13	1.63	40.38	130.97	1	0	0
AC14	21.23	53.32	149.47	1	0	0
AC15	12.03	40.36	121.44	1	0	0
AC16	100.10	72.14	30.00	0	0	1
AC17	20.85	42.01	115.13	1	0	0
AC18	32.78	8.42	108.90	0	1	0
AC19	42.55	17.89	117.14	0	1	0
AC20	9.09	47.92	132.48	1	0	0
AC21	30.51	24.13	125.61	0	1	0
AC22	17.28	37.94	136.30	1	0	0
AC23	174.23	144.54	44.26	0	0	1
AC24	39.81	3.16	104.41	0	1	0
AC25	53.89	13.71	96.37	0	1	0
AC26	75.70	36.48	73.55	0	1	0

Karena pada Iterasi Ke-2 dan Ke-3 posisi cluster tidak berubah maka iterasi dihentikan dan hasil akhir yang diperoleh yaitu:

1. C0 memiliki 12 data yang diartikan bahwa kelompok pertama adalah kategori aksesoris kurang laku.
2. C1 memiliki 10 data yang diartikan bahwa kelompok kedua adalah kategori aksesoris laku.
3. C2 memiliki 4 data yang diartikan bahwa kelompok ketiga adalah kategori aksesoris paling Laku.

Pada Implementasi Dan Pengujian disini kita menggunakan sebuah Software RapidMiner Studio Versi 8.2, dengan pengujian data menggunakan software kita akan membandingkan bagaimana hasil pengolahan data secara manual dengan hasil pengolahan data menggunakan sebuah software.

Pada gambar 4 merupakan halaman utama dari aplikasi *RapidMiner Studio v.8.2* saat pertama kali membuka aplikasi setelah proses loading selesai terlihat pada gambar berikut :



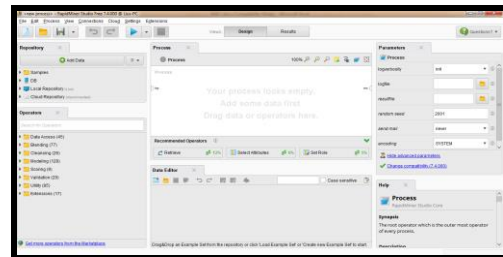
Gambar 4. Halaman Utama *RapidMiner Studio v.8.2*

Pada tampilan halaman utama ada lima menu yang akan digunakan yaitu :

1. *Icon New Process* berfungsi untuk membuat halaman proses kerja *Data Mining* yang baru.
2. *Open Process* berfungsi untuk membuka proses yang sudah ada direpository sebelumnya.
3. *Learn* berisikan petunjuk-petunjuk menggunakan *RapidMiner Studio v.8.2*.

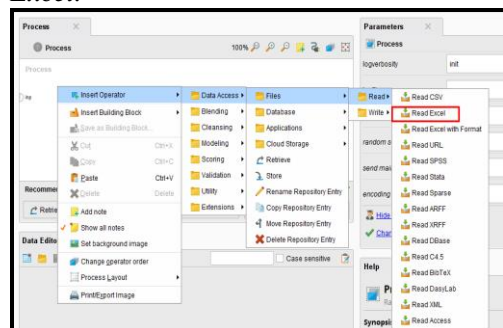
Berikut ini adalah tampilan *New Process* yang akan digunakan untuk

memulai proses kerja baru terlihat pada gambar 5.



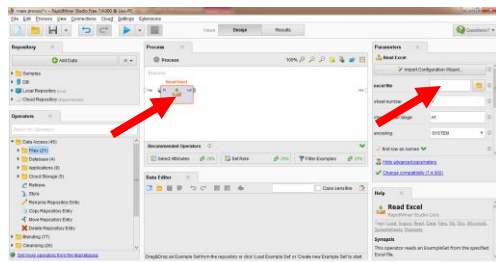
Gambar 5. *New Process RapidMiner Studio v.8.2*

Tampilan *new process* adalah untuk membuat halaman kerja pada *RapidMiner Studio v.7.4*. *Import data* dilakukan untuk memasukkan data yang akan diuji dalam bentuk format *.xls* atau *.xlsx*. Berikut adalah cara untuk melakukan import file Microsoft Excel. Untuk membuat mengimport data yang akan diproses, maka dilakukan *New process*, dengan cara klik kanan → *Insert Operator* → *Data Access* → *Files* → *Read* → *Read Excel*.



Gambar 6. Tampilan *Import Data* Dengan Membuat Operator Baru

Setelah selesai maka akan tampil *read excel* pada *RapidMiner Studio v.8.2*, Seperti terlihat pada Gambar 8. Setelah itu lanjutkan dengan mengklik Tab *Import Configuration Wizard* untuk melakukan menginputkan data dari *Microsoft Excel* (dengan format *.xls*, atau *.xlsx*).

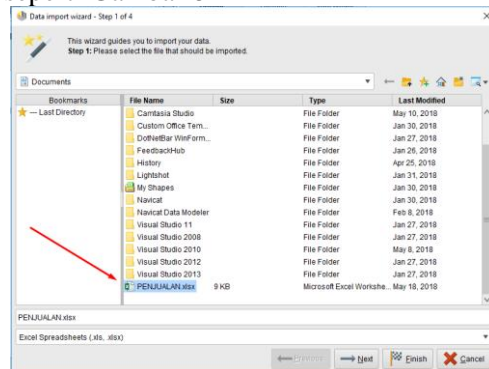


Gambar 7. Tampilan Operator Baru Read Excel.

Dalam melakukan *import* data pada aplikasi *RapidMiner Studio v.8.2* terdiri dari 4 tahap, adapun langkah-langkahnya sebagai berikut:

#### 1. Data Import Wizard – Step 1 of 4

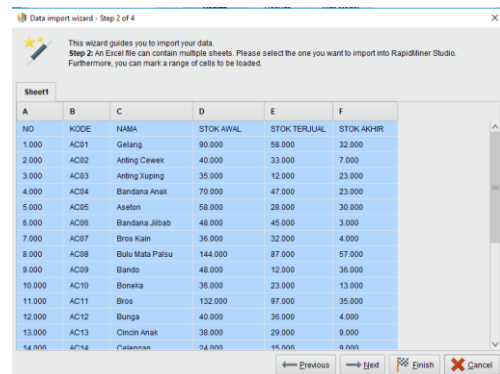
Tahap pertama adalah mencari lokasi file yang sudah dibuat sebelumnya dengan format *.xlsx* atau *.xls*, pilih dan kemudian simpan. Pada pengujian ini data yang akan diuji disimpan dengan nama *PENJUALAN.xlsx* lalu pilih diklik, seperti Gambar 8



Gambar 8. Tampilan Data Import Wizard - Step 1 of 4

#### 2. Data Import Wizard – Step 2 of 4

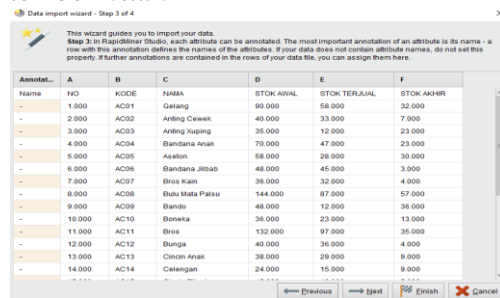
Kemudian pilih *next* maka lanjut ke tahap berikutnya yaitu *import wizard step 2 of 4*. Pada tahap ini akan dipilih *Sheet* yang berisikan sumber data pada *Microsoft Excel* yang digunakan. Maka akan terlihat data *import wizard* seperti terlihat pada gambar 9.



Gambar 9. Tampilan Data Import Wizard - Step 2 of 4

#### 3. Data Import Wizard – Step 3 of 4

Tahap ini merupakan tahap pemberian anotasi, jika data kita tidak memiliki nama *attribute*, tidak usah melakukan apa-apa pada step 3 ini. Kemudian klik tombol *Next*.

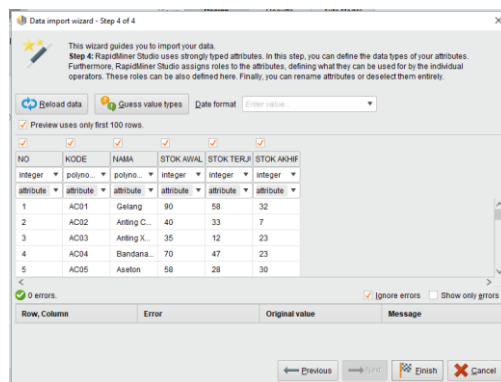


Gambar 10. Tampilan Data Import Wizard - Step 3 of 4

#### 4. Data Import Wizard - Step 4 of 4

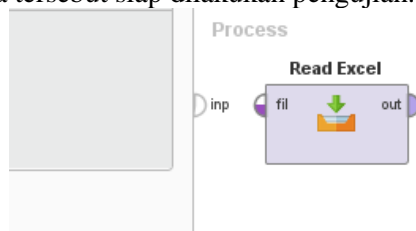
Pada *Step 4*, tahap ini merupakan tahap penentuan tipe data dan atribut. Sebenarnya *RapidMiner Studio* akan memberikan tipe data yang tepat secara otomatis. Namun, jika kita merasa tipe data yang diberikan *RapidMiner Studio* tidak cocok, kita bisa mengubahnya. Kemudian klik tombol *Next*.





Gambar 11. Tampilan Data Import Wizard - Step 4 of 4

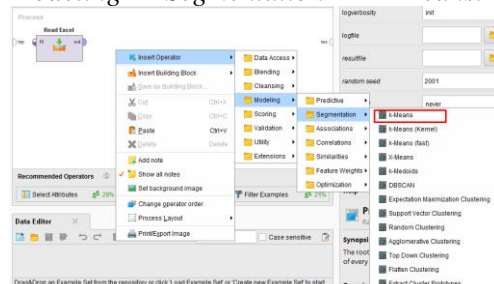
Tahap *import* data selesai dengan memilih *Finish*, maka pada *Main Process* terlihat sebuah operator baru *Read Excel* yang sudah berisi file *PENJUALAN.xlsx* yang di *import* langsung dari file *excel*. Data tersebut siap dilakukan pengujian.



Gambar 12. Tampilan Operator Data Siap Import

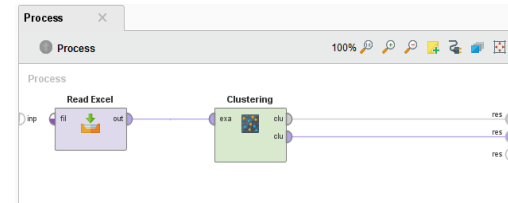
Pada gambar 12 menampilkan data yang sudah di import, sehingga tombol pada operator *read excel* tidak ada lagi tanda seru warna kuning (menandakan operator telah berisi data dan siap diolah).

Langkah selanjutnya adalah menambahkan operator *K-Means* dengan cara klik kanan → *Insert Operator* → *Modeling* → *Segmentation* → *K-Means*.



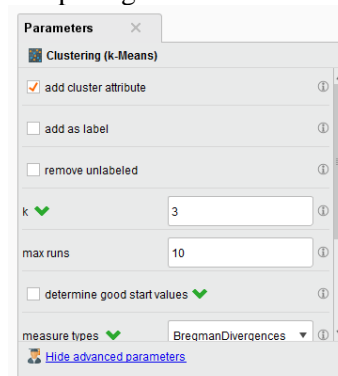
Gambar 13. Menambahkan Operator K-Means

Hubungkan data dengan *clustering k-means* untuk mengetahui output ke arah *result*. Apabila sudah terhubung maka tentukan jumlah *cluster* dan *maksimal runs* yang diberikan. Terakhir klik *Tombol Play*, seperti pada Gambar 14.



Gambar 14. Tampilan Untuk Proses K-Means

Selanjutnya melakukan pengaturan algoritma *K-Means*, yang diatur pada menu *Parameter Clustering K-Means*, seperti terlihat pada gambar 15.



Gambar 15. Penentuan Jumlah Cluster

Pada gambar 15 di atas dilakukan pengaturan nilai *k*, di mana *k* merupakan nilai yang digunakan untuk menentukan jumlah *cluster* yang akan dibentuk. Di sini jumlah *cluster* yang akan dibentuk adalah sebanyak 3 *cluster* sesuai tingkatan bahaya yaitu rendah, sedang dan tinggi.

Setelah proses *running*, akan tampil *Example Set (Read Excel)*, seperti yang terlihat pada Gambar 16.



Row No.	NAMA	cluster	STOK AWAL	STOK TERJ.	STOK AKHIR
1	Gelang	cluster_1	90	58	32
2	Anting Cewek	cluster_0	40	33	7
3	Anting Taping	cluster_0	35	12	23
4	Bandana Anak	cluster_1	70	47	23
5	Aseton	cluster_1	58	28	30
6	Bandana Jilb...	cluster_0	48	45	3
7	Bros Kam	cluster_0	36	32	4
8	Buku Mata Pal...	cluster_2	144	87	57
9	Bando	cluster_0	48	12	36
10	Boneka	cluster_0	36	23	13
11	Bros	cluster_2	132	97	35
12	Bunga	cluster_0	40	36	4
13	Cincon Anak	cluster_0	38	29	9
14	Celengan	cluster_0	24	15	9

Gambar 16. Tampilan *Running Data*

Pada bagian *Example Set* terdapat beberapa bentuk hasil *cluster* yaitu: *Data View*, *Statistics View*, *Charts*, dan *Annotations*. Untuk *Cluster model* terdapat *Description*, *Folder View*, *Graph View*, *Centroid Table*, *Centroid Plot View* dan *Annotations*.

Pada Hasil Pengujian Data terdapatlah beberapa output yang dihasilkan oleh software RapidMiner v.8.2 yaitu :

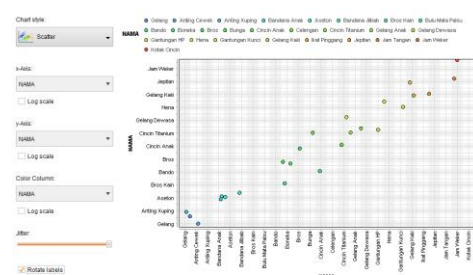
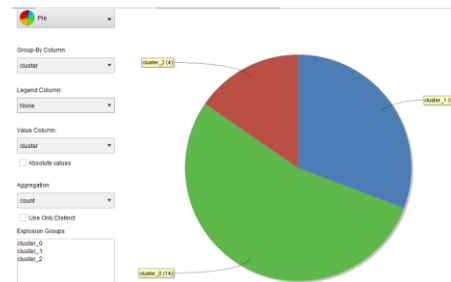
### 1. ExampleSet

Pada *ExampleSet* dapat dilihat beberapa tampilan hasil *cluster*, yaitu *Data View*. *Data View* merupakan tampilan hasil *cluster* data secara keseluruhan sesuai dengan data yang telah diinputkan

Row No.	NAMA	cluster	STOK AWAL	STOK TERJ.	STOK AKHIR
1	Gelang	cluster_1	90	58	32
2	Anting Cewek	cluster_0	40	33	7
3	Anting Taping	cluster_0	35	12	23
4	Bandana Anak	cluster_1	70	47	23
5	Aseton	cluster_1	58	28	30
6	Bandana Jilb...	cluster_0	48	45	3
7	Bros Kam	cluster_0	36	32	4
8	Buku Mata Pal...	cluster_2	144	87	57
9	Bando	cluster_0	48	12	36
10	Boneka	cluster_0	36	23	13
11	Bros	cluster_2	132	97	35
12	Bunga	cluster_0	40	36	4
13	Cincon Anak	cluster_0	38	29	9
14	Celengan	cluster_0	24	15	9

Gambar 17. Tampilan Hasil *Cluster* Pada *Data View*

*Chart* merupakan tampilan grafik hasil pengelompokan atau *cluster* sampel data nilai mata pelajaran ujian nasional dengan 3 *cluster*.

Gambar 18. Tampilan *Scatter* pada *Chart*Gambar 19. Tampilan *Pie* pada *Chart*

Pada gambar 19. dapat dilihat untuk *cluster 0* kategori kurang laku dengan jumlah 14, *cluster 1* kategori laku dengan jumlah 8, sedangkan *cluster 2* kategori paling laku dengan jumlah 4 orang.

### 2. Cluster Model (Clustering)

Pada *Cluster Model (Clustering)* dapat dilihat beberapa tampilan hasil *cluster*, yaitu *Text View* yang merupakan tampilan hasil pengelompokan berdasarkan *cluster* dan jumlah anggotanya. Tampilan *Text View* akan terlihat pada gambar 5.17.

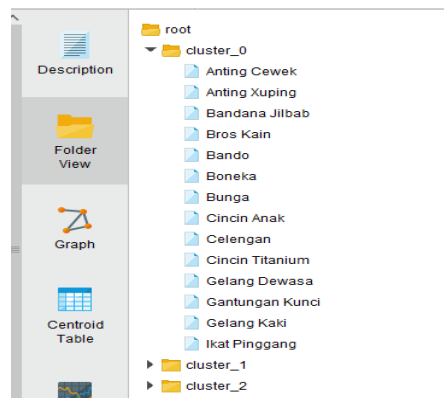
```

Cluster Model

Cluster 0: 14 items
Cluster 1: 8 items
Cluster 2: 4 items
Total number of items: 26

```

Gambar 20. Tampilan *Text View*

Gambar 21. Tampilan *Folder View*

*Folder View* merupakan tampilan data bagian-bagian *cluster* secara keseluruhan, dimana masing – masing anggota cluster menampilkan nama siswa.

## SIMPULAN

Berdasarkan uraian yang sudah dikemukakan pada sebelumnya, maka dapat ditarik beberapa kesimpulan Metode *Clustering* dengan algoritma *K-Means* dapat digunakan untuk mengelompokkan data penjualan aksesoris berdasarkan jumlah penjualannya, yaitu Paling Laku, Laku, Dan Kurang Laku. Sehingga pihak toko dapat mengantisipasi aksesoris yang kurang laku ini bagaimana supaya

menjadi laku, Metode *Clustering* dapat digunakan untuk membantu Toko dalam melakukan pengelompokan aksesoris, Berdasarkan hasil perbandingan dengan cara manual dengan software terdapat sedikit perbedaan yang hasil manual cluster 0 berjumlah 12, cluster 1 berjumlah 10 dan cluster 2 berjumlah 4. Sedangkan hasil dari softwarenya cluster 0 berjumlah 14, cluster 1 berjumlah 8, dan cluster 2 berjumlah 4.

Peneliti menyadari adanya kekurangan dalam penulisan ini, karena keterbatasan penulis baik dalam hal waktu maupun pengetahuan. Dalam rangka memperbaiki kekurangan dan untuk penyempurnaan penelitian ini penulis memberikan beberapa saran Untuk penelitian selanjutnya dapat dilakukan dengan data yang lebih banyak dan parameter yang lebih banyak lagi serta Untuk memaksimalkan waktu proses *clustering* dengan algoritma *K-Means*, penentuan *centroid* awal yang baik akan membuat proses *clustering* dapat dilakukan dengan lebih cepat dan Untuk penelitian selanjutnya dalam hal pengekrasian data penjualan menerapkan metode *Data Mining* dengan algoritma yang berbeda.

## DAFTAR PUSTAKA

- Sri, T.S. (2016). “Analisa Penerapan Data Mining Untuk Menenyukan Kubikasi Air Terjual Berdasarkan Pengelompokan Pelanggan Menggunakan Algoritma K-Means Clustering”, 9(1).
- Bendri, M.M., & Herlina, L.S. (2015). “Analisis Clustering Menggunakan Metode K-Means Dalam Pengelompokan Penjualan Produk Pada Swalayan Fadhila”, 11(2).
- Asroni & Ronald, A. (2015). “Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interfase Studi Kasus Pada Jurusan Teknik Informasi UMM Magelang”, 18(1).
- Nurul, R.W. (2015). “Implementasi Algoritma K-means Dalam Pengklasteran Mahasiswa Pelamar Beasiswa”, 1(2).
- Sibuea, M. L., & Safta, A. (2017). Pemetaan Siswa Berprestasi Menggunakan Metode K-Means Clustering. *JURTEKSI*, 4(1), 85-92.
- Risnawati, R., & Handayani, M. (2017). penerapan Jaringan Saraf Tiruan Untuk Proyeksi Logistik Berdasarkan Prediksi Pasien Menggunakan Algoritma Backpropagation. *Jurteks*, 4(1)

