

Implementasi *Data Mining* Untuk Pengelompokan Aset Perusahaan pada *Regional Office* PT.NAK Dengan Metode *K-Means Clustering*

Jumadi^{1*}, Hidayatullah Al Islami¹

¹Fakultas Ilmu Komputer, Teknik Informatika, Universitas Pamulang, Jl. Raya Puspipetek No. 46, Kel. Buaran, Kec. Serpong, Kota Tangerang Selatan. Banten 15310, Indonesia

Email: ^{1*}didijumadi79@gmail.com, ²dosen02408@unpam.ac.id

(* : coressponding author)

Abstrak– Dalam hal pengadaan dan pendistribusian aset perusahaan di *regional office* PT.NAK dipengaruhi oleh perencanaan yang dilakukan. Agar aset dapat tersedia dengan jenis dan jumlah yang cukup sesuai dengan kebutuhan operasional *project*. Pada saat ini lokasi area *regional office* yang ada belum melakukan kegiatan seperti yang diharapkan, mengingat beberapa kendala antara lain pendataan dan pemantauan aset dengan baik sehingga keberadaan aset tidak diketahui dan dikelola dengan baik, kemampuan tenaga ahli, terbatasnya pengetahuan manajemen *regional office*, kebijakan manajemen head office, dan terbatasnya akses sarana. Akibat dari kondisi ini pengelolaan aset maka masih bersifat konvensional dimana data yang diolah manual dinilai belum efektif dari segi waktu dan belum memberikan informasi yang jelas mengenai data aset yang harus diprioritaskan untuk diganti atau diperbaiki. Penelitian ini akan mengembangkan sebuah sistem *data mining* yang dapat membantu proses pengelompokan data aset *regional office* kedalam beberapa *cluster*, sehingga mengurangi lama waktu proses pengolahan data untuk rencana pembelian aset sesuai kebutuhan, serta mengurangi penumpukan aset yang tidak cepat pakai. Penelitian ini menggunakan metode *K-Means clustering* yaitu sebuah metode yang dapat digunakan untuk mengelompokkan data berdasarkan tingkat kemiripan dan karakteristik data. Penerapan metode *K-Means* pada penelitian ini digunakan untuk mengolah data yang akan dimasukkan kedalam masing-masing 3 *cluster*, *cluster* yang pertama yaitu data aset tidak layak pakai atau harus dilakukan perbaikan maupun pergantian dengan pembelian diatas 5 tahun, dan yang kedua data aset masih layak pakai dengan pembelian maksimal 5 tahun dan ketiga data aset sangat layak pakai dengan tahun pembelian baru maksimal 2 tahun. Data yang digunakan adalah data yang didapatkan dari hasil penelitian dilapangan dimana terdapat 6434 data mentah yang diolah sesuai kebutuhan penelitian menjadi 322 data yang akan diuji pada sistem. Berdasarkan hasil pengujian pada sistem menggunakan metode *K-Means Clustering* didapatkan hasil bahwa 119 aset termasuk kedalam *cluster* pertama, 5 aset termasuk kedalam *cluster* kedua dan 199 aset termasuk kedalam *cluster* ketiga.

Kata Kunci: Aset, *Data Mining*, *K-Means*, *Clustering*, *Regional Office*

Abstract– *In terms of the procurement and distribution of company assets in the PT.NAK regional office, it is influenced by the planning carried out. So that assets can be available with the type and amount that is sufficient according to the operational needs of the project. At present the location of the existing regional office area has not carried out activities as expected, considering several obstacles including proper data collection and monitoring of assets so that the whereabouts of assets are not known and properly managed, the ability of experts, limited knowledge of regional office management, policies head office management, and limited access to facilities. As a result of this condition, asset management is still conventional where data processed manually is considered ineffective in terms of time and does not provide clear information regarding asset data that must be prioritized for replacement or repair. This research will develop a data mining system that can help the process of grouping regional office asset data into several clusters, thereby reducing the time it takes to process data for plans to purchase assets as needed and reduce the accumulation of assets that are not used quickly. This study uses the K-Means clustering method, which is a method that can be used to group data based on the degree of similarity and characteristics of the data.*

The application of the K-Means method in this study is used to process data that will be entered into each of the 3 clusters, the first cluster is asset data that is not suitable for use or must be repaired or replaced with purchases over 5 years, and the second is that asset data is still feasible used with a maximum purchase of 5 years and the three data assets are very suitable for use with a maximum of 2 years of new purchases. The data used is data obtained from research results in the field where there are 6434 raw data that are processed according to research needs into 322 data that will be tested on the system. Based on the test results on the system using the K-Means Clustering method, it was found that 119 assets were included in the first cluster, 5 assets were included in the second cluster and 199 assets were included in the third cluster.

Keywords: Assets, *Data Mining*, *K-Means*, *Clustering*, *Regional Office*

1. PENDAHULUAN

Aset perusahaan merupakan semua sumber ekonomi atau kekayaan yang dimiliki yang diharapkan mampu memberi manfaat usaha di masa mendatang. Dengan adanya aset yang dimiliki perusahaan diharapkan bisnis dapat terus beroperasi dengan lancar, aset perusahaan bisa dalam bentuk uang tunai, tanah, bangunan, peralatan kantor serta aset juga dapat berupa sumber daya manusia seperti karyawan perusahaan. Penggunaan aset perusahaan oleh karyawan juga merupakan basis data besar yang dapat dilakukan dengan pengelompokan secara terpusat melalui algoritma *K-Means Clustering*.

Pada PT.NAK yaitu perusahaan yang bergerak di bidang *Security System*, terdapat proses pembelian aset perusahaan yang terpusat di *Head Office* Jakarta dan selanjutnya disalurkan ke berbagai *Regional Office* perusahaan di seluruh Indonesia. Penggunaan aset perusahaan tersebut tentunya harus diatur dengan baik agar apabila dilakukan pengecekan terdapat data aset yang bisa dipantau penggunaannya. Didalam pengecekannya tentunya dibutuhkan data yang akurat untuk dikelompokkan dengan baik agar mudah dianalisa penggunaannya dengan bobot kriteria yang sudah ditentukan dari masa pemakaiannya.

Permasalahan yang timbul adalah tidak dapat terkontrolnya penggunaan aset di berbagai *Regional Office* perusahaan, pengelompokan data yang masih tidak tertata dengan baik sehingga menyulitkan kantor pusat untuk menganalisa ataupun mengecek status dan keberadaan dari aset yang digunakan. Data aset yang ada belum mencerminkan kebutuhan dari manajemen perusahaan, sehingga pengelompokan aset yang baik dapat membantu perusahaan untuk melakukan pembaharuan aset atau pun penilaian penggantian aset yang digunakan dapat meminimalkan pengeluaran budget dari perusahaan yang tentunya akan membantu dari segi *cash flow* perusahaan.

Berdasarkan uraian diatas dengan demikian dapat disimpulkan bahwa PT.NAK membutuhkan suatu implemetasi dari pengelompokan aset perusahaan yang begitu banyak yang tersebar diberbagai *Regional Office* perusahaan dengan bantuan menggunakan metode *K-Means Clustering* yang diharapkan dapat membantu manajemen perusahaan dalam mengontrol aset perusahaan yang sedang digunakan di seluruh *Regional Office* yang ada.

2. METODOLOGI PENELITIAN

2.1 Data Mining

Data mining merupakan metode dalam ilmu komputer yang biasa digunakan dalam proses pencarian knowledge. Tahapan di dalamnya berguna untuk mencari pola-pola tertentu dari data yang ada pada database. Beberapa pembagian metode data mining berikut yang diketahui adalah sebagai berikut:

- a. **Classification data mining** adalah sebuah proses menemukan definisi kesamaan karakteristik dalam suatu kelompok atau *class*.
- b. **Association** adalah metode data mining yang kedua adalah *Market basket analysis* atau analisa keranjang pasar bisa disebut *association*. Berhubungan dengan pemasaran, metode ini bertujuan untuk mengidentifikasi produk yang sering dibeli bersamaan oleh pelanggan.
- c. **Clustering** data mining adalah nama lain untuk metode *segmentation*. Tujuan dari segmentasi pada metode data mining adalah mengelompokkan suatu *class* ke dalam beberapa segmen berdasarkan atribut yang ditentukan.
- d. **Regression**, merupakan metode yang bertujuan untuk mencari pola nilai *numerik*, bukan kelas. Hasil dari metode *regression* adalah sebuah fungsi sebagai penentu hasil yang didasarkan nilai dari *input*.
- e. **Forecasting** data mining adalah metode yang digunakan untuk memprediksi nilai yang akan dicapai pada satu periode. Dengan menggunakan teknik ini, *noise* data dan nilai pada periode sebelumnya dijadikan dasar bahan prediksi.
- f. **Sequence** adalah sebuah urutan peristiwa, metode *sequence analysis* berfungsi untuk mencari sebuah pola pada serangkaian kejadian atau *sequence*.

- g. **Descriptive**, metode data mining yang satu ini bertujuan untuk memahami lebih dalam mengenai data-data yang masuk dalam pengamatan. Hasil akhirnya adalah mengetahui perilaku dari data itu sendiri.

Dengan melakukan penggalian pola informasi terhadap data yang berguna sebagai memanipulasi data menjadi sebuah informasi baru serta lebih bermanfaat yang didapatkan melalui cara mengekstrasi juga mengetahui beberapa pola yang berharga atau menarik yang didapatkan dari basis data (Dito, 2019). Adapun tahapan dalam data mining adalah sebagai berikut:

a. Data Cleaning

Sebelum diproses dan dibentuk menjadi sebuah *knowledge*, data yang ada harus dibersihkan terlebih dahulu. Jika ada yang mengandung *error*, maka data-data tersebut harus dibuang. Sehingga tersisa data yang ‘bagus’ untuk diolah dalam tahap selanjutnya.

b. Data Integration

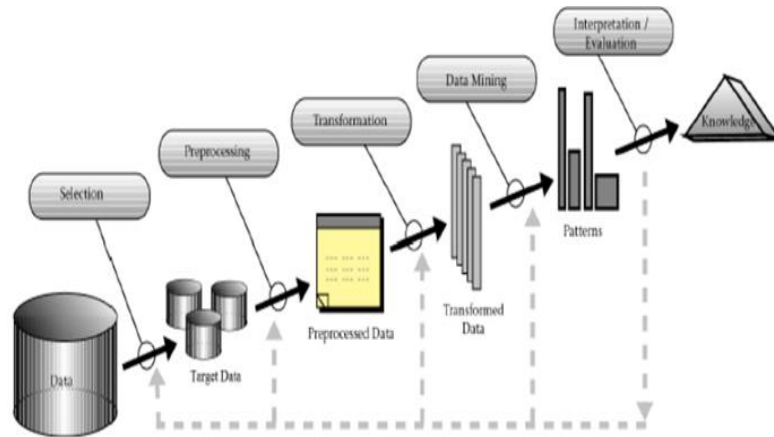
Dalam data mining, data yang berhasil dibersihkan juga akan diintegrasikan.

c. Transformation Data

Dalam tahap ini data akan dipilih dan diubah formatnya agar sesuai dengan teknik atau metode yang dipakai. Pada tahap ini pula kualitas *data mining* akan terlihat.

d. Pattern Data Evaluation

Setelah selesai melakukan proses *data mining*, pola-pola yang dihasilkan dari proses tersebut perlu untuk dievaluasi. Tujuan dari dilakukannya evaluasi adalah untuk menguji hipotesis awal. Setelah teruji data bisa dipresentasikan kepada pengguna.



Gambar 1. Tahapan Pencarian *Data Mining*

Berisi penjelasan tentang tahapan penelitian yang menggambarkan urutan logis untuk mendapatkan hasil penelitian sesuai dengan harapan dan gambaran sistem. Jika ada gambar dan tabel, itu harus disajikan dengan nama tabel dan gambar yang disertai dengan nomor urut.

2.2 Algoritma K-Means

Algoritma *K-Means* merupakan algoritma pengelompokan *iterative* yang melakukan partisi set data kedalam sejumlah *k cluster* yang sudah ditetapkan di awal. Algoritma *K-Means* sederhana untuk diimplemtasikan dan dijalankan, *relative* cepat, mudah beradaptasi, umum penggunaannya dalam praktek.

K-Means merupakan salah satu metode data *clustering* non hirarki yang berusaha mempartisi data yang ada kedalam bentuk satu atau lebih *cluster* atau kelompok. Metode ini mempartisi kedalam cluster atau kelompok sehingga data yang memiliki karakteristik yang sama (*High Intra class similarity*) dikelompokkan kedalam satu *cluster* yang sama dan yang memiliki karakteristk yang berbeda (*Low inter class similarity*) dikelompokkan pada kelompok yang lain. Proses *clustering* dimulai dengan mengidentifikasi data yang akan di *cluster*.

$$X_{ij}(i = 1, \dots, n; j = 1, \dots, m)$$

Dengan n adalah jumlah data yang akan di *cluster* dan m adalah jumlah *variable*. Pada awal iterasi pusat setiap *cluster* ditetapkan secara bebas (sembarang).

$$C_{kj}(k = 1, \dots, k; j = 1, \dots, m)$$

Kemudian dihitung jarak antara setiap data dengan setiap pusat *cluster*. Untuk melakukan jarak data ke- i (x_i) pada pusat *cluster* ke- k (c_k), diberi nama (d_{ik}), dapat digunakan formula *Euclidean Distance*, seperti pada persamaan (1), dibawah ini yaitu:

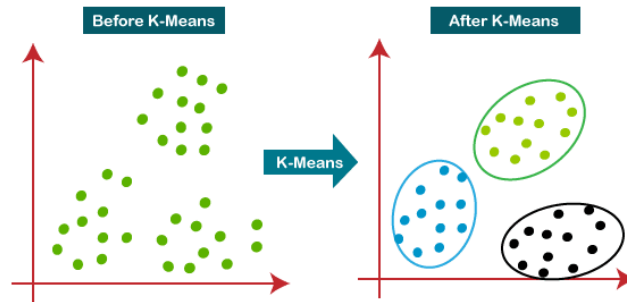
$$d_{ik} = \sqrt{\sum_{j=1}^m (x_{ij} - c_{kj})^2}$$

2.3 Clustering

Metode *clustering* hirarki merupakan satu-satunya metode yang masuk kedalam kategori pengelompokan hirarki. Metode *clustering partitioning* membagi set data kedalam sejumlah kelompok yang tidak tumpang tindih (*overlap*) antara satu kelompok dengan kelompok yang lain artinya setiap data hanya menjadi anggota satu kelompok.

Metode *clustering* menurut keanggotaan dalam kelompok dibagi menjadi dua yaitu eksklusif dan tumpang tindih. Metode tersebut masuk dalam kategori eksklusif jika sebuah data hanya menjadi anggota satu kelompok dan tidak menjadi anggota kelompok yang lain. Metode *clustering* yang masuk dalam kategori ini adalah *K-Means* dan *DBSCAN* sedangkan yang masuk dalam kategori tumpang tindih adalah metode *clustering* yang membolehkan sebuah data menjadi anggota dilebih dari satu kelompok misalnya *Fuzi C-Means*.

Metode *clustering* menurut kategori kekompakan terbagi menjadi dua yaitu komplet dan parsial. Semua data bisa dikatakan kompak menjadi satu kelompok jika semua data bias bergabung menjadi satu dalam konteks penyekatan namun jika ada sedikit data yang tidak ikut bergabung dalam kelompok mayoritas data tersebut dikatakan mempunyai perilaku menyimpang. Data yang menyimpang ini dikenal dengan sebutan *noise*.



Gambar 2. Proses Clustering K-Means

Banyak metode *clustering* yang sudah dikembangkan oleh para ahli. Masing-masing metode mempunyai karakter, kelebihan dan kekurangan. *Clustering* dapat dibedakan juga menurut struktur *cluster*, keanggotaan data dalam *cluster* dan kekompakan data dalam *cluster*. Metode *clustering* menurut strukturnya dibagi menjadi dua yaitu pengelompokan hirarki dan *partitioning*. Pengelompokan hirarki memiliki aturan satu data tunggal bisa dianggap sebagai sebuah kelompok, dua atau lebih kelompok kecil dapat bergabung menjadi kelompok besar dan begitu seterusnya hingga semua data dapat bergabung menjadi satu kelompok.

3. ANALISA DAN PEMBAHASAN

Tahapan-tahapan dari analisa dan perancangan yang dilakukan dalam melakukan penelitian dari awal hingga akhir dan tahapan penelitian berguna agar pelaksanaan penelitian dapat berjalan dengan baik dan sistematis serta memenuhi tujuan yang diinginkan.

3.1 Analisa Penelitian

Analisa atau analisis adalah mengamati aktivitas objek dengan cara mendeskripsikan komposisi objek dan menyusun kembali komponen-komponennya untuk dikaji atau dipelajari secara detail. Analisis merupakan upaya yang dilakukan untuk mengamati sesuatu secara mendalam dan mendetail melalui proses penguraian berbagai komponen pembentuknya atau juga penyusunan komponen tersebut untuk dipelajari atau diselidiki lebih lanjut.

3.1.1 Analisa Sistem Yang Ada

Analisa sistem yang ada adalah kegiatan analisis sistem yang berjalan di lokasi penelitian merupakan kegiatan penguraian suatu sistem informasi yang utuh dan nyata ke dalam bagian-bagian sistem dengan tujuan untuk mengidentifikasi serta mengevaluasi masalah yang muncul dari data sistem tersebut. Tujuan dari analisa sistem yang ada disini adalah untuk mendapatkan data penelitian yang akan digunakan sesuai dengan kebutuhan dari penelitian.



Gambar 3. System Asset Management PT.NAK

3.1.2 Analisa Kebutuhan Penelitian

Analisa Kebutuhan Pengujian merupakan proses untuk mendapatkan kebutuhan pengujian penelitian diantaranya kebutuhan data uji dan kebutuhan software pengujian yang akan digunakan pada penelitian.

a. Data yang digunakan

Data yang akan digunakan melalui beberapa proses seperti tahapan seleksi data yang bertujuan untuk mendapatkan data yang bersih dan siap untuk digunakan dalam penelitian. Adapun data mentah berjumlah 6434 data yang akan diproses melalui beberapa tahap, sample data dapat dilihat pada tabel 1.

Tabel 1. Tabel Data Mentah (*Sample Data*)

TGL PEROLEH	NO.ASSET	NAMA BARANG	MERK	TAHUN	NAMA KOTA	NAMA AREA	JUMLAH	SATUAN
26/10/2003	VMT-1003-0	SEPEDA MOTOR	YAMAHA	2003	KAB.PASURUAN	R.O SURABAYA	1	UNIT
01/02/2006	FUR-0106-00	FILING CABINET	VIP	2006	KAB.BANDUNG	R.O BANDUNG	1	UNIT
01/02/2006	NFU-0106-00	FILING CABINET	VIP	2006	KAB.BANDUNG	R.O BANDUNG	1	UNIT
05/02/2006	FUR-0506-00	MEJA	OLYMPIC	2006	KAB.BANDUNG	R.O BANDUNG	1	UNIT
05/02/2006	NFU-0506-00	MEJA	OLYMPIC	2006	KAB.BANDUNG	R.O BANDUNG	1	UNIT
27/08/2006	LVA-0806-00	MEJA	UNO	2006	KAB.BANDUNG	R.O BANDUNG	1	UNIT
27/08/2006	LVA-0806-00	KURSI	FUTURA	2006	KAB.BANDUNG	R.O BANDUNG	2	UNIT
27/08/2006	FUR-0806-00	LEMARI	OLIMPIC	2006	KAB.BANDUNG	R.O BANDUNG	1	UNIT
29/08/2006	NO PQ-806-0	METAL DETECTOR	GARRET	2006	Kab. Bandung	R.O BANDUNG	4	UNIT

Selanjutnya dari data mentah aset dilakukan percangan untuk 3 cluster yang akan digunakan pada penelitian ini seperti dalam tabel dibawah ini:

Tabel 2. Tabel Cluster Area (Sample Data)

Data ke-	Nama Area	Nama Kota	(X) = Tahun Pembelian Aset	(Y) = Total Jumlah Aset	Cluster
1	RO Balikpapan	Kab Kutai Kertanegara	2018	136	3
2	RO Balikpapan	Kab Kutai Kertanegara	2019	111	3
3	RO Balikpapan	Kab Kutai Kertanegara	2020	114	3
11	RO Balikpapan	Kab Kutai Timur	2016	72	1
12	RO Balikpapan	Kab Kutai Timur	2017	10	1
19	RO Balikpapan	Kab Tanah Laut	2017	4	1
26	RO Balikpapan	Kota Balikpapan	2022	391	2
27	RO Balikpapan	Kota Balikpapan	2022	157	3
28	RO Balikpapan	Kota Balikpapan	2017	64	1

b. Software Data Mining Yang Digunakan

Dari analisa beberapa *software data mining* untuk pengelompokan *K-Means Clustering* penulis akan menggunakan *RapidMiner Studio* karena memiliki karakteristik yang cocok dengan data yang akan dilakukan pengujian. *RapidMinerStudio* menyediakan *GUI (Graphic User Interface)* untuk merancang sebuah *pipeline* analitis. *GUI* ini akan menghasilkan *file XML (Extensible Markup Language)* yang mendefinisikan proses analitis keinginan pengguna untuk diterapkan ke data. *File* ini akan dibaca *RapidMiner* untuk menjalankan analisis secara otomatis.

c. Analisa Metode Penelitian

Metode penelitian menggunakan Algoritma *K-Means Clustering* sebagai implementasi pengelompokan aset perusahaan pada *regional office* sesuai kebutuhan dan tujuan yang ingin dicapai.



Gambar 4. Algoritma Flowchart Clustering K-Means

3.2 Perancangan Penelitian

Perancangan penelitian adalah kerangka kerja yang dibuat oleh peneliti untuk menjelaskan jalannya proses pengujian penelitian guna mendapatkan informasi untuk menyelesaikan dan memecahkan masalah penelitian. Fungsi rancangan penelitian adalah untuk memastikan bahwa data yang diperlukan sesuai dengan masalah yang ada dikumpulkan secara akurat dan rancangan penelitian tergantung pada kebutuhan dari penelitian.

Prosedur penelitian merupakan serangkaian langkah-langkah yang dilakukan dalam penelitian, yaitu tahap persiapan pelaksanaan penelitian dan analisa data serta penyusunan laporan. Prosedur dalam penelitian Implementasi pengelompokan data aset pada *regional office* PT.NAK dimulai dari tahap perencanaan penelitian yaitu pengelompokan aset *regional office* menggunakan metode *K-Means Clustering* dengan menyiapkan instrumen pengumpulan data seperti observasi penelitian, menentukan kriteria dari penelitian hingga pelaksanaan penelitian dilakukan sesuai jadwal yang ditetapkan. Perancangan sistem atau *software data mining* yang akan digunakan disesuaikan dengan analisa digunakan yaitu *Software RapidMiner Studio* dengan alur algoritma *K-Means Clustering*.

4. IMPLEMENTASI

Implementasi *K-Means Clustering* adalah proses dari pelaksanaan penelitian data yang dilakukan melalui tahapan-tahapan yang sudah ditentukan dan dilakukan guna mendukung dari tujuan penelitian yang ingin dicapai.

4.1 Cleaning Data

Sebelum data aset diproses dan dibentuk menjadi sebuah *knowledge*, data mentah yang didapat harus dibersihkan terlebih dahulu. Jika ada yang mengandung *error*, maka data-data tersebut harus dibuang. Sehingga tersisa data yang ‘bagus’ untuk diolah dalam penelitian. Dari 6434 data yang sudah diperoleh selanjutnya dilakukan *cleaning data* dan sebanyak 1278 data yang tidak digunakan, dan menghasilkan data yang bagus yang akan digunakan sebagai bahan penelitian yaitu sebanyak 5155 data aset *regional office* PT.NAK.

4.2 Data Integration

Selanjutnya data aset yang sudah berhasil dibersihkan juga akan dilakukan proses integrasi data, yaitu menggabungkan beberapa data yang sama kedalam satu tampilan yang diinginkan sesuai kebutuhan penelitian. Maka proses integrasi data yang dilakukan dapat terlihat pada tabel berikut.

Tabel 3. Tabel Integrasi Data (*Sample Data*)

Data ke	Nama Area	Nama Kota	Tahun Pembelian Aset	Total Jumlah Aset
1	RO Balikpapan	Kab Kutai Kertanegara	2018	136
2	RO Balikpapan	Kab Kutai Kertanegara	2019	111
3	RO Balikpapan	Kab Kutai Kertanegara	2020	114
4	RO Balikpapan	Kab Kutai Kertanegara	2021	155
5	RO Balikpapan	Kab Kutai Timur	2010	65
318	RO Surabaya	Kota Surabaya	2018	167
319	RO Surabaya	Kota Surabaya	2019	110
320	RO Surabaya	Kota Surabaya	2020	121
321	RO Surabaya	Kota Surabaya	2021	146
322	RO Surabaya	Kota Surabaya	2022	113
323	RO Surabaya	Kota Tuban	2021	138

Selanjutnya dari hasil proses *integration data* yang berjumlah 5155 diolah dan menghasilkan data baru sebanyak 322 data untuk proses tahapan berikutnya yaitu transformasi data.

4.3 Data Transformation

Metode yang digunakan pada pengumpulan data dalam program aplikasi ini adalah sebagai berikut:

Pada tahap transformasi 322 data akan dan diubah formatnya agar sesuai dengan teknik atau metode yang dipakai, kualitas data mining akan terlihat dengan melalui beberapa kali proses *K-Means* mencari pusat *centroid* tiap *cluster*. Perhitungan manual dilakukan menggunakan rumus *Euclidean Distance* dengan memilih data secara acak sebagai pusat *cluster* awal sbb:

$$\text{Centroid 1} = 2014,555 - 19,118$$

$$\text{Centroid 2} = 2019,800 - 346$$

$$\text{Centroid 3} = 2019,990 - 127,583$$

Menghitung jarak tiap data ke pusat *cluster* dimulai dari **data ke-1** sbb:

$$D1 \rightarrow C1 (2014,555 - 19,118) = \sqrt{(2018 - 2014,555)^2 + (136 - 19,118)^2} = 166,933$$

$$D1 \rightarrow C2 (2019,800 - 346) = \sqrt{(2018 - 2019,800)^2 + (136 - 346)^2} = 210,008$$

$$D1 \rightarrow C3 (2019,990 - 127,583) = \sqrt{(2018 - 2019,990)^2 + (136 - 127,583)^2} = 8,649$$

Selanjutnya hitung **data ke-2** sbb:

$$D2 \rightarrow C1 (2014,555 - 19,118) = \sqrt{(2019 - 2014,555)^2 + (111 - 19,118)^2} = 91,990$$

$$D2 \rightarrow C2 (2019,800 - 346) = \sqrt{(2019 - 2019,800)^2 + (111 - 346)^2} = 235,001$$

$$D2 \rightarrow C3 (2019,990 - 127,583) = \sqrt{(2019 - 2019,990)^2 + (111 - 127,583)^2} = 16,612$$

Dan seterusnya, hitung sampai data terakhir → data ke-323 sbb:

$$D323 \rightarrow C1 (2014,555 - 19,118) = \sqrt{(2021 - 2014,555)^2 + (138 - 19,118)^2} = 119,057$$

$$D323 \rightarrow C2 (2019,800 - 346) = \sqrt{(2021 - 2019,800)^2 + (138 - 346)^2} = 208,003$$

$$D323 \rightarrow C3(2019,990 - 127,583) = \sqrt{(2021 - 2019,99)^2 + (138 - 127,583)^2} = 10,466$$

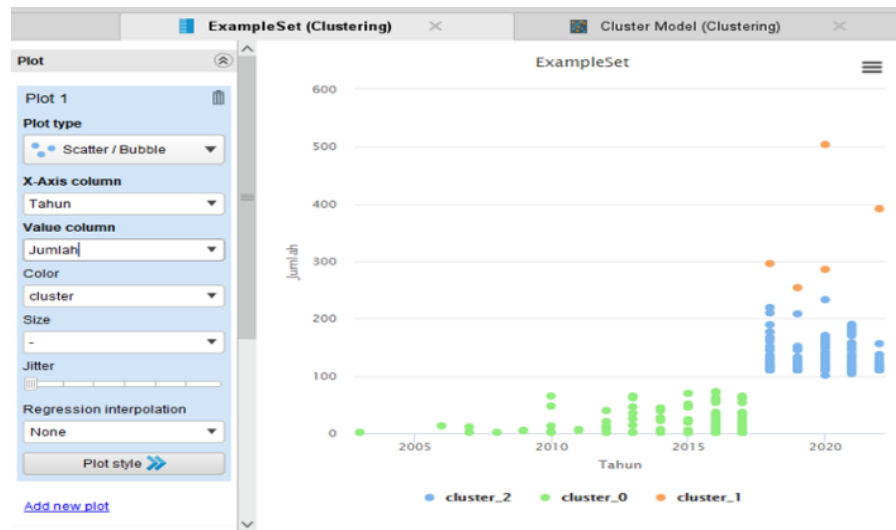
Pada Proses *K-Means* dilakukan sampai menghasilkan data yang baik dan pusat centroid yang tidak mengalami perubahan lagi sehingga menghasilkan cluster yang paling baik sbb:

Jumlah Data: 323

Cluster 1: 119 data

Cluster 2: 5 data

Cluster 3: 199 data



Cluster Model

Cluster 0: 119 items

Cluster 1: 5 items

Cluster 2: 199 items

Total number of items: 323

Gambar 5. Clustering *K-Means* & RapidMiner Studio

5. KESIMPULAN

5.1 Kesimpulan

Berdasarkan uraian yang telah dibahas sebelumnya dan pengujian yang telah dilakukan terhadap data aset *regional office* PT.NAK menggunakan metode *K-Means Clustering* maka dapat diambil kesimpulan sebagai berikut :

1. Dengan menerapkan algoritma *K-Means Clustering* didapatkan sebuah model untuk implementasi pengelompokan aset *regional office* PT.NAK berdasarkan jumlah dan tahun pengadaan, sehingga dapat memudahkan *Dept Asset Management GA* dalam menemukan aset yang dimiliki perusahaan pada *regional office* di seluruh lokasi.
2. Penentuan aset perudahaan berdasarkan tahun pembelian dan jumlah aset dimiliki *regional office* terbagi menjadi 3 *cluster*, yaitu *cluster C1*, *cluster C2*, dan *cluster C3*, yang telah dilakukan penghitungan dengan menggunakan algoritma *K-Means Clustering*, sehingga dapat diketahui pada *cluster* ke berapa bisa direkomendasikan pergantian aset *regional office* dapat berjalan dengan baik.

5.2 Saran

Selanjutnya penulis mempertimbangkan beberapa saran untuk perbaikan, yaitu:

1. Untuk mendapatkan hasil yang lebih variatif penelitian ini dapat juga dikembangkan dengan teknik *data mining* yang lain seperti algoritma *Hierarchical Clustering*, algoritma *K-Nearest Neighbor* atau algoritma *Fuzzy Clustering*.
2. Faktor lain yang menyebabkan sebagian data dalam *cluster* bisa tidak relevan atau berkaitan adalah sifat homogen yang dimiliki oleh algoritma *k-means clustering*, sehingga disarankan penggunaan metode lain yang dapat mengurangi ukuran dimensi yang cukup besar.

REFERENCES

- Jollyta, D., William, R., & Muhammad, Z., 2020, *Konsep Data Mining Dan Penerapannya*, Deepublish, Yogyakarta.
- Siregar, A.M., & Adam, P., 2020, *Pengolahan Data Menjadi Informasi Denga RapidMiner*, Kekata Group CV, Indonesia.
- Najia, S. 2019, *Klasifikasi Barang Menggunakan Metode Clustering K-Means Dalam Penentuan Prediksi Stok Barang UKM Mar'ah Jilbab Kediri*. Skripsi. Universitas Islam Negeri Maulana Malikibrahim, Malang.
- Rega, F.P.D. 2019, *Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Klasifikasi Penyakit ISPA*. Skripsi. Universitas Pelita Bangsa, Bekasi.
- Muhammad, L.H. 2020, *Pengelompokan Data Transaksi Penjualan Menggunakan Algoritma Clustering K-Means Pada Toko AN NUR*. Skripsi. Universitas Semarang, Semarang.
- Efrain, H.B. 2020, *Pengelompokan Data Obat Menggunakan Metode K-Means Clustering Pada UPT Puskesmas Kondoran Kec.Sangala*. Skripsi. Institut Teknologi Nasional, Malang.
- Event. 2021, *Analisis Clustering Dengan K-Means Untuk Pengelompokan Penjualan Produk Pada Hotel Newton*. Skripsi. Universitas Putera Batam, Batam.
- Uun, H. 2022, *Implementasi Metode K-Means Clustering Untuk Klasifikasi Penjualan Parfum*. Skripsi. Universitas Nusantara, Kediri.