

Implementasi Algoritma K-Means Clustering dalam Pengelompokan Produk Pigeon Pada PT. Digital Niaga Indonesia Berdasarkan Analisis Recency, Frequency, Monetary (RFM)

Ahmad Pudoli¹, Radityo Nugroho², Yulianawati³, Dewi Kusumaningsih⁴

^{1,2,4}Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Budi Luhur, Indonesia

³Program Studi Sistem Informasi, Fakultas Teknologi Informasi, Universitas Budi Luhur, Indonesia

Email: ¹ahmad.pudoli@budiluhur.ac.id, ²1911510350@student.budiluhur.ac.id, ³yulianawati@budiluhur.ac.id,

⁴dewi.kusumaningsih@budiluhur.ac.id

Abstrak - Persaingan bisnis dalam era global saat ini semua perusahaan yang bergerak dalam bidang perdagangan produk harus cermat dan tepat dalam menemukan konsep serta pola di penjualan yang dilakukan agar dapat meningkatkan angka penjualan dan penggunaan metode pemasaran di perusahaan, salah satunya adalah dengan pemanfaatan dan pengelolaan data penjualan yang dimiliki. PT. Digital Niaga Indonesia merupakan perusahaan yang menjalankan bisnisnya di online shop seperti Shopee, Tokopedia, serta Lazada dari Toko Pigeon Indonesia yang menjual perlengkapan bayi mulai dari botol susu, dot maupun perlengkapan bayi lainnya. Namun demikian dari semua produk yang ditawarkan tentunya tidak semua diminati dan terjual dengan lancar, ada yang terjual sangat laris, yang laris, serta kurang laris. Data dan informasi penjualan yang dimiliki perusahaan jika hanya tersimpan secara digital di dalam database tentunya tidak dapat dimanfaatkan dan diolah untuk kebutuhan dari pengembangan strategi pemasaran perusahaan. Tujuan dari penelitian yang dilakukan merupakan untuk mengetahui Algoritma K-Means fungsionalitasnya dalam menghasilkan menentukan Cluster atau disebut juga kelompok untuk produk paling laris, laris dan kurang laris untuk menjadi dasar pembuatan promosi di setiap marketplace pada PT. Digital Niaga Indonesia. Dalam menentukan peminatan barang menggunakan laporan penjualan 4 bulan yaitu Agustus sampai November 2022, kemudian data akan diproses dengan Algoritma K-Means Clustering. Hasil akhir adalah didapatkan 557 produk dan menghasilkan nilai di akhir 3 Cluster atau kelompok berikut : Cluster 1 atau ditulis dengan C1 sebanyak 100 data, kemudian untuk Cluster 2 ditulis dengan lambang C2 dengan hasil 20 data serta dalam Cluster 3 atau kelompok tiga dilambangkan dengan C3 menghasilkan 437 data yang didapat dari 124473 data transaksi. Metode RFM dapat menentukan bahwa Cluster 1 adalah produk yang kurang laris, Cluster 2 paling laris, dan Cluster 3 laris. Evaluasi Cluster yang dibuat dihitung dengan DBI dan mengeluarkan hasil 0.520.

Kata Kunci - k-means, marketplace, Recency, Frequency, Monetary

Abstract - Business competition in the current global era, all companies engaged in product trading must be careful and precise in finding concepts and patterns in sales in order to increase sales figures and use marketing methods in the company, one of which is by utilizing and managing sales data. owned. PT. Digital Niaga Indonesia is a company that runs its business in online shops such as Shopee, Tokopedia, and Lazada from Toko Pigeon Indonesia which sells baby equipment ranging from milk bottles, pacifiers and other baby equipment. However, of all the products offered, of course not all are in demand and sell smoothly, some sell very well, some sell well, and some sell less well. If the company's sales data and information is only stored digitally in a database, it certainly cannot be utilized and processed for the needs of developing the company's marketing strategy. The aim of the research carried out is to find out the functionality of the K-Means Algorithm in producing clusters or also known as groups for the best-selling, best-selling and least-selling products to become the basis for creating promotions in each marketplace at PT. Indonesian Digital Commerce. In determining demand for goods, use a 4 month sales report, namely August to November 2022, then the data will be processed using the K-Means Clustering Algorithm. The final result was that 557 products were obtained and produced a value at the end of the following 3 Clusters or groups: Cluster 1 or written with C1 with 100 data, then for Cluster 2 written with the symbol C2 with a result of 20 data and in Cluster 3 or group three it was symbolized with C3 resulting 437 data obtained from 124473 transaction data. The RFM method can determine that Cluster 1 is the least selling product, Cluster 2 is the best selling product, and Cluster 3 is the best selling product. The cluster evaluation created was calculated using DBI and produced a result of 0.520

Keywords - k-means, marketplace, Recency, Frequency, Monetary

I. PENDAHULUAN

Menemukan ide dan pola dari konsep penjualan adalah salah satu kunci dalam keberlangsung penjualan dimasa global sekarang. Dalam satu dekade ini, kecanggihan informasi dan teknologi merupakan suatu keniscayaan yang menjadi simbol perubahan kehidupan. Bisa dipastikan saat ini sebagian besar sektor penunjang kehidupan mengalami

perubahan yang signifikan, termasuk di dalamnya lingkup pemasaran [1]. Persaingan bisnis di masa sekarang, kita harus pintar dalam menemukan pola penjualan untuk bisa meningkatkan hasil penjualan dan marketing pada perusahaan dalam menawarkan produk ke masyarakat [2]. Hal yang belum dimanfaatkan oleh PT. Digital Niaga Indonesia dalam menentukan pola penjualan dan pemasaran agar bisnisnya berjalan dengan lancar. PT. Digital Niaga Indonesia merupakan perusahaan yang menjalankan bisnisnya di online shop seperti Shopee, Tokopedia, serta Lazada dari Toko Pigeon Indonesia yang menjual perlengkapan bayi mulai dari botol susu, dot maupun perlengkapan bayi lainnya. Namun demikian dari beberapa produk yang ditawarkan oleh perusahaan pastinya belum tentu semua produk dapat terserap oleh konsumen. Ada yang paling diminati dengan pola diskon ataupun sistem paket/*bundling*. Produk tersebut dikelompokkan nantinya dengan jenis yang kurang laris, laris dan juga sangat laris. Data produk yang terdapat dalam transaksi penjualan yang hanya tersimpan di dalam sistem tentunya belum dapat diolah dan hasilnya dimanfaatkan yang hasilnya digunakan dalam menentukan strategi pemasaran produk dalam rangka pengembangan strategi penjualan produk tersebut. Data yang dimiliki dapat dimaksimalkan untuk digunakan menentukan pola yang nantinya dapat diberikan pola paket atau diskon dan *bunling* sebagai salah satu penentuan pengambilan keputusan di pihak management untuk solusi dari bisnis yang dijalankan.

Pigeon Indonesia mempercayakan bisnis onlinenya kepada PT Digital Niaga Indonesia sejak tahun 2017. Rata-rata penjualan dari Pigeon Indonesia yaitu 300 transaksi per harinya dari semua *marketplace* yang menjual produk pigeon. Transaksi yang banyak dan *marketplace* yang banyak juga membuat tim marketing PT Digital Niaga Indonesia kebingungan dalam memberi keputusan untuk mempromosikan produk apa saja yang laku dan tidak laku di setiap *marketplace*. Dalam menyelesaikan masalah yang dijabarkan tersebut, dapat dimanfaatkan implementasi model kluster atau clustering dengan menerapkan suatu algoritma K-means guna produk yang dijual dapat digolongkan dalam kelompok kurang laris, atau laris amupun sangat laris dan membuat tim marketing PT Digital Indonesia dapat memutuskan cara mempromosikan produknya.

Penelitian yang sudah diselesaikan Normah dengan adalah menerapkan konsep Kluster/*Clustering* yaitu K-Means. Riset yang dilakukan memakai Dataset yang digolongkan yaitu pasokan awal, produ yang terjual, dan sisa produk akhir. Konklusi yang dilaporkan mengeluarkan hasil sebanyak tiga kluster/*cluster*, 34 artikel kurang laris dan 55 artikel yang masuk dalam kluster laris dan ada 11 artikel yang masuk sebagai artikel sangat laris. Adapun *output* yang dihasilkan pada riset/penelitian yang dilakukan Normah adalah menggolongkan kesimpulan pada tiga *Cluster*/kluster ialah 55 artikel laris dan juga 11 artikel sangat laris, serta terakhir 34 artikel kurang laris. Masih belum ada parameter data yang digunakan sebagai parameter faktor dalam analisis untuk melakukan penentuan *Cluster* penjualan. Oleh karena itu, peneliti menggunakan

metode model *Recency, Frequency, Monetary* (RFM) tambahan pada riset yang dilakukan ini dikarenakan metode tersebut mengeluarkan konklusi yang sangat efektif dalam menganalisis segmentasi pelanggan [3].

Penelitian lain dilakukan oleh Tamba, dalam penelitiannya menggunakan hasil *Dataset* stok awal dan *stock* terjual. Penelitian tersebut mengeluarkan hasil tiga kluster produk *spare part* Toyota yaitu laris, tiga belas barang yang kurang laris, lima belas barang yang paling laris dan 45 barang yang laris, Namun validitas hasil *Clustering* tidak dapat diverifikasi karena tidak dilakukan uji validitas/validasi pada riset yang Tamba lakukan. Dikarenakan hal tersebut, maka penulis memulai konsep *Davies-Bouldin Index* atau DBI dalam melakukan validasi kluster/*Cluster* pada hasil metode K-Means, yang menghasilkan kesimpulan bahwa *Cluster* tersebut dinyatakan valid [4].

Oleh karena itu, pengolahan data harus dianalisis menggunakan model *Recency, Frequency, Monetary* (RFM). RFM dapat dipergunakan untuk menampilkan perilaku konsumen, transaksi terakhir konsumen atau disebut juga *Recency*, jumlah atau frekuensi pelanggan dalam melakukan proses transaksi atau disebut juga *Frequency* serta nilai atau angka yang dikeluarkan pelanggan dalam melakukan transaksi atau disebut juga *Monetary*. Sehingga dibutuhkan proses analisis dalam melakukan identifikasi pola penjualan yang didapat dari pengolahan data tersebut. Kebutuhan tersebut dapat ditangani dengan model *data mining* yang berfokus pada *clustering*. [5]

Data yang didapat pada penelitian yang dilakukan Kurniawan, konsep *Clustering* K-Means didapat melalui metode ini, data digolongkan menjadi beberapa *Cluster* yang dibuat atas dasar nilai dari kesamaan data. Data dengan nilai karakteristik sama akan digolongkan dalam satu *Cluster* dan juga data yang memiliki nilai karakteristik berbeda dibuat kelompok *cluster* yang lain. Data yang telah dikelompokkan tersebut, mempunyai nilai karakteristik yang tentunya sama. Konsep *preprocessing* pada data kerap diimplementasikan pada kelompok data agar data yang dimiliki menjadi lebih bersih, dan juga konsisten, serta tidak *random*. Proses normalisasi tentunya dipakai guna menghapus lebih serta akan mengeluarkan *Cluster* dengan tinggi kualitas, tentunya mampu meningkatkan nilai dari efisiensi Algoritma *Clustering*. [6]. Teknik *data mining* adalah pengamatan terhadap data yang ada untuk membuat model. Model yang digunakan untuk mengidentifikasi model data lain yang tidak disimpan dalam *database* [7]. Berbagai pengolahan data dilakukan secara komputerisasi, mulai dari penyimpanan data, arsip, membuat laporan, serta menghasilkan informasi, balik yang dibutuhkan secara perorangan maupun perusahaan. *Data mining* diaplikasikan menguasai proses analisis dalam data yang sangat besar ataupun dalam jumlah yang sangat banyak. Dalam penerapannya terdapat konsep atau metode yang digunakan yaitu K-means *Clustering*. Algoritma K-Means *Clustering* adalah salah satu proses penganalisisan data ataupun informasi ataupun tahapan *datamining* yang melakukan pemaksimalan proses pemodelan tanpa supervisi [8].

Menjaga produk penjualan agar selalu tersedia atau tidak dalam stok kosong yang merupakan salah satu cara pemeliharaan pelanggan. Untuk itu, pihak manajemen Poultry Shop wajib melakukan analisis produk yang mana paling diminati baik yang sangat laku/laris, kurang laris atau laris/laku di bagian penjualan. Didasari akan banyaknya produk yang dimiliki sehingga sangat dibutuhkan adanya rekayasa komputerisasi yang diperuntukan dalam penyelesaian masalah penjualan. Algoritma K-Means *Clustering* digunakan dalam menyelesaikan masalah penggolongan produk yang terjual bukan hanya dari jumlah pembelian saja, namun digolongkan produk yang terjual tersebut kedalam beberapa golongan atau *cluster* yang terbentuk. Berdasarkan hasil riset tersebut didapatkan tiga *Cluster* yang terpola, yaitu *Cluster* satu berisi 2 produk dan *Cluster* dua terdapat 9 produk serta terdapat 25 produk yang masuk dalam *Cluster* 3. Berdasar dari kesimpulan tersebut, maka manajemen Poultry Shop bisa mengimplementasikan *cluster* tersebut guna menaikkan nilai dari persediaan atau stok produk serta menerapkan strategi penjualan.[9]. Pengelompokan pengunjung guna menaikkan pendapatan yang sempat turun karena pandemi. Kelanjutan dari hasil *Cluster* yang terbentuk tersebut mampu diketahui nilai karakteristik dari pengunjung mall. Karakteristik dari pengunjung mall yang ada nantinya akan digunakan dalam menambah pendapatan dari mall tersebut. Adapun *dataset* yang dipakai dilabeli pengunjung mall dengan memiliki format CSV. *Dataset* tersebut akan diolah dengan menerapkan bahasa pemrograman python dalam mengimplementasikan Algoritma K-Means. Guna mendapatkan kepastian dalam nilai akurasi algoritma K-Means tentunya diterapkan proses optimasi dengan algoritma *Particle Swarm Optimization* (PSO). Langkah lebih lanjut Selanjutnya dihasilkan nilai dengan evaluasi menggunakan Davies Boudin Index (DBI) pada microsoft excel. Hal itu dilakukan guna mengetahui seberapa baik dari *Clustering* yang dihasilkan serta mendapatkan pelanggan dengan tingkat penghasilan tinggi serta mempunyai nilai atau skor dengan pengeluaran tinggi yang menjadi target dengan tingkat prioritas utama bagi mall[10].

Model RFM merupakan model yang membantu proses *Clustering*/pengelompokan dengan menggunakan tiga variabel yaitu Recency, Frequency, Monetary sebagai berikut[11]:

- 1) Recency, merupakan rentang transaksi terakhir konsumen. Semakin kecil rentangnya, maka skor R semakin besar.
- 2) Frequency, merupakan seberapa seringnya konsumen melakukan transaksi dalam satu periode. Semakin banyak Frequency, maka skor F semakin besar.
- 3) Monetary, merupakan seberapa besar jumlah uang dari konsumen atas transaksinya. Semakin besar jumlah uang yang dikeluarkan pada periode tersebut maka skor M semakin besar.

Pengumpulan data transaksi penjualan dilakukan dengan memeriksa data dan membersihkan data data yang tidak valid serta data yang tidak diperlukan. Kemudian analisis RFM digunakan untuk menentukan nilai recency,

dan recency pada masing –masing produk. Selanjutnya nilai RFM masing –masing produk dinormalisasi menggunakan normalisasi standar. Untuk proses selanjutnya adalah *clustering* data penjualan produk menggunakan k-means *clustering*, dari hasil *clustering* didapatkan data segmentasi produk, hasil *clustering* tersebut diuji menggunakan rumus Silhouette Coefficient[12].

Segmentasi pelanggan dapat melakukan proses segmentasi menggunakan model RFM untuk mengetahui karakteristik dari pelanggan dan metode K-means yang digunakan untuk melakukan segmentasi kepada pelanggan. Berdasarkan hasil *clustering* dengan metode K-Means jumlah segmen pelanggan yang dimiliki oleh PT. Coversuper Indonesia Global adalah 4 *cluster*/segmen pelanggan dari keseluruhan jumlah pelanggan sebanyak 736. Hasil olah data pelanggan menggunakan algoritma K-Means dengan metode *clustering* dan dengan menggunakan atribut RFM diimplementasikan ke dalam proses pengelompokan pelanggan potensial pada perusahaan yang menghasilkan kelompok pelanggan Consumers 226, Ordinary 186, Big Consumers 299 dan Top Class 25. Berdasarkan hasil User Acceptance Testing dari 29 kali pengujian, terdapat 29 pelanggan (100%) dengan karakteristik yang dihasilkan sistem sesuai dengan pengetahuan user, 0 pelanggan (0%) tidak sesuai dengan pengetahuan user[13].

Digitalisasi dan otomatisasi dalam pelayanan mahasiswa di Perguruan Tinggi dapat menghasilkan big data. Amanat pemerintah dalam Peraturan Menteri Riset Teknologi dan Pendidikan Tinggi agar besaran Uang Kuliah Tunggal (UKT) di Perguruan Tinggi Negeri dibagi ke dalam 5 kelompok berdasarkan tingkatan kondisi sosial ekonomi orang tua. Dalam proses menetapkan UKT begitu banyak indikator sosial ekonomi orang tua yang harus dijadikan acuan sehingga menyulitkan dalam mengidentifikasi dan mencari formula yang tepat. Untuk mengelompokkan data mahasiswa ini dilakukan dengan teknik data mining menggunakan metode K-Means *Clustering*. Metode ini mengelompokkan besaran UKT mahasiswa berdasarkan pola atau kemiripan data sosial ekonomi orang tua. Data yang digunakan dalam penelitian ini adalah data calon mahasiswa baru Universitas Negeri Padang. Pengelompokan ini bertujuan untuk membantu menetapkan besaran UKT calon mahasiswa baru pada Perguruan Tinggi Negeri. Hasil dari penelitian diperoleh 5 kelompok besaran UKT, terdiri dari UKT kategori 1 Rp. 500.000, UKT kategori 2 Rp. 1.000.000, UKT kategori 3 Rp. 2.000.000, UKT kategori 4 Rp. 3.000.000 dan UKT kategori 5 Rp. 4.000.000[14].

Didasari atas penelitian sebelumnya yang telah dijelaskan dalam latar belakang diatas, sehingga dapat dihasilkan rumusan masalah pada penelitian ini yaitu bagaimana Algoritma K-Means dapat diimplementasikan guna membuat dan menentukan strategi promosi berdasarkan produk paling laris, laris dan kurang laris di setiap marketplace berdasarkan analisis recency, frequency, monetary serta bagaimana Algoritma K-Means dapat diimplementasikan pada pengujian aplikasi[15].

Data yang akan dipakai untuk penelitian atau riset merupakan nilai data pada bulan Agustus hingga November 2022. Dan tanggal penelitian 7 Januari 2023.

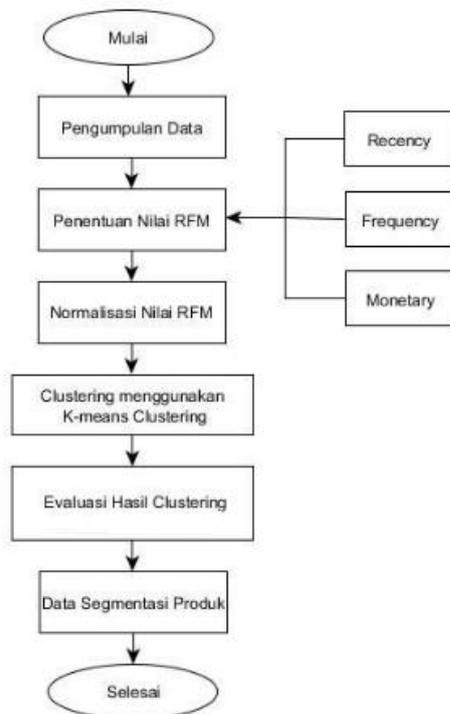
Tujuan dalam penelitian ini adalah membuat pengelompokan atau penggolongan guna menghasilkan kelompok data produk yang paling laris, laris dan kurang laris dan melakukan analisis pada data produk yang ada di data transaksi penjualan dengan menerapkan model analisis RFM serta memakai algortime dari metode K-Means Clustering dengan menggunakan paramater nilai RFM.

II. METODE PENELITIAN

Penelitian pada PT Digital Niaga Indonesia yang merupakan perusahaan yang menaungi brand Pigeon Indonesia yang teretak di daerah Alam Sutera Tangerang Selatan. Sebagai brand yang besar Pigeon Indonesia memiliki banyak sekali data, data transaksi yang ada di setiap marketplace seperti Shopee, Tokopedia, Lazada, dan Blibli, yang berupa file excel, dengan 124473 transaksi dari bulan Agustus sampai November 2022.

Adapun data yang diambil dari penelitian pada brand Pigeon Indonesia meliputi atribut-atribut berikut:

- 1) *Created_date*
- 2) *Market_place*
- 3) *Market_place_order_number*
- 4) *Brand*
- 5) *Product_type*
- 6) *Item_Code*
- 7) *Item_Name*
- 8) *Selling_price*
- 9) *Qty*
- 10) *Total_amount*



Gambar 1. Alur Metode Penelitian

A. Metode Pembandingan

Metode pembandingan adalah cara yang digunakan dalam membandingkan pada penerapan dari metode yang dipakai dengan membandingkan beberapa kelebihan serta kekurangan dari metode yang dipakai. Pada laporan penelitian ini, penulis menggunakan algortime k-means yang merupakan salah satu cara dari metode atau Algoritma *Clustering* data *non-hierarchical* yang melakukan analisis guna mengelompokkan data yang dimiliki sehingga menghasilkan satu atau lebih *Cluster* atau kelompok.

Kelebihan yang dimiliki pada *Algoritma K-Means Clustering* adalah mampu membuat pengelompokkan dari objek yang besar serta titik data yang jauh dari titik data yang lain dengan waktu yang sangat cepat, hingga dapat mempercepat proses pengelompokkan data.

Sedangkan kekurangan dari *Algoritma K-Means Clustering* adalah titik awalnya harus dilakukan proses inisialisasi secara random, hingga proses *Clustering* dari data yang dihasilkan kemungkinan akan memiliki nilai berbeda. Namun jika pada nilai acak/random yang didapatkan pada inisialisasi kurang baik maka pengelompokkan yang dihasilkan akan tidak optimal.

B. Penerapan Metode

Berdasarkan dari nilai daya yang dimiliki oleh Pigeon Indonesia, maka data tersebut bisa dianalisis dan diolah guna membuat peminatan atas barang atau produk dengan menggunakan konsep data mining dengan Algoritma K-Means Clustering. Adapun tujuannya adalah menghasilkan produk yang mana saja yang termasuk dalam kelompok sangat laris, laris dan kurang laris pada setiap marketplace.

C. Tahapan Cross Industry Standard for Data Mining (CRISP-DM)

1) Pemahaman Bisnis

Mendapatkan pemahaman dari tujuan serta kebutuhan yang berada pada lingkup proses bisnis berjalan atau juga lembaga riset penelitian, serta melakukan penjabarkan pengetahuan kedalam konsep inti dalam permasalahan data mining. Langkah selanjutnya melakukan penentuan strategi guna menggapai tujuan tersebut. Penelitian ini berada pada PT. Digital Niaga Indonesia yang merupakan penyedia layanan atau jasa untuk brand yang membutuhkan mitra dalam mengembangkan bisnis online mulai dari penjualan produk maupun memasarkan secara digital. Data transaksi yang tersimpan digunakan untuk melakukan pengelompokan atau segmentasi produk di setiap *marketplace*. Guna meminimlisir proses yang dilakukan pada kegiatan pengelompokkan produk, maka pada riset yang dilakukan ini memakai parameter *recency* (yaitu terakhir dalam melakukan transaksi), *frequency* (yaitu seringnya dalam melakukan transaksi), serta *monetary* (yaitu berapa banyak uang dalam proses transaksi). Di lain hal tersebut memudahkan pengelompokan data serta penamaan produk, pada penelitian ini akan diterapkan pada platform berbasis web guna proses pengelompokkan pelanggan secara otomatis serta dapat mengetahui apa saja produk yang ada di setiap *marketplace*.

2) Pemahaman Data

Adapun yang dilakukan pada tahapan ini yaitu: persiapan proses mengidentifikasi dari kualitas data, serta melakukan pemeriksaan dari data dan melakukan pembersihan data yang tidak valid atau disebut juga data *cleansing*. Pada riset ini menggunakan data transaksi pigeon dari bulan Agustus 2022 sampai dengan November 2022. Data diambil dari dokumentasi PT. Digital Niaga Indonesia yang berformat xls. Jumlah data transaksi terdiri dari 124.473 transaksi.

3) Persiapan Data

Pada proses yang dilakukan, tahap ini digunakan data yang telah tersimpan untuk dipersiapkan untuk memudahkan kegiatan penambangan (*mining*). Terdapat beberapa langkah ataupun proses yang wajib dilakukan dalam melakukan persiapan data yaitu diantaranya adalah melakukan pemilihan variabel yang nantinya akan dilakukan analisis, kemudian dilakukan pembersihan data, selanjutnya disiapkan data awal sehingga siap untuk dilakukan tahap selanjutnya yaitu menyiapkan data *transformation*.

Segmentasi pada produk yang dilakukan pada riset ini didasari atas model *recency*, *frequency*, *monetary* (RFM), untuk itu pemilihan data pada model RFM merupakan rentang jangka waktu terakhir pembeli melakukan kegiatan transaksi dengan menggunakan metode analisis, jumlah frekuensi (atau seberapa sering) transaksi tersebut dilakukan pembeli dalam masa periode penelitian ini, serta dibutuhkan jumlah transaksi dalam bentuk angka (nominal) untuk setiap pembeli selama masa riset/penelitian yaitu selama bulan Agustus 2022 sampai dengan November 2022. Tabel 1 menunjukkan hasil dari data *selection*.

Tabel 1. Hasil Seleksi Data

Data Awal		Data Akhir
Tanggal transaksi	akhir dari	<i>Recency (type: number)</i>
Jumlah dari transaksi		<i>Frequency</i>
Jumlah nominal transaksi	dari	<i>Monetary</i>

Setelah dilakukannya proses seleksi data, maka tahapan yang dilakukan selanjutnya yaitu melakukan data *preprocessing*. *Preprocessing* data adalah proses yang memiliki fungsi dalam memberikan kepastian atau validitas data yang terpilih yang dilakukan pada tahap seleksi data atau data *selection*. Di tahap *preprocessing* maka dilakukanlah proses pembersihan data yang merupakan salah satu tahap penting atau disebut juga data *cleansing*. Data *cleansing* yang dilakukan adalah dengan cara melakukan pembersihan data *recency* (yaitu terakhir dalam melakukan transaksi), *frequency* (yaitu seringnya dalam melakukan proses transaksi), serta *monetary* dari *noise* atau *missing value*, dengan melakukan pembuangan dari *noise* serta nilai *missing value* yang ada pada data. Proses yang dilakukan selanjutnya yaitu melakukan transformasi data (data

transformation) yang harus dilakukan dengan menggunakan cara normalisasi. Sedangkan normalisasi dilakukan guna menentukan skala data agar tidak berbeda terlalu jauh.

4) Pemodelan Data

Dalam riset ini, cara ataupun tahapan dari metode yang diterapkan dalam proses pemodelan data untuk penelitian ini adalah *Algoritma K-Means*. Data yang diinput dalam proses *Clustering* adalah data RFM pada transaksi yang dilakukan pelanggan yang sebelumnya telah dilakukan normalisasi. Data pada transaksi telah dilakukan perubahan menjadi data dengan densitas yang lebih tinggi, serta mendapatkan hasil data yang terdiri dari *recency* (tanggal terakhir transaksi), *frequency* (jumlah produk yang terjual), serta juga *monetary* (jumlah pendapatan dari produk yang terjual).

5) Evaluasi

Proses pada evaluasi adalah proses yang melakukan interpretasi pada pemodelan data mining yang didasari domain pengetahuan sesuai dengan pemahaman bisnis. Proses evaluasi diterapkan setelah klaster-klaster atau kelompok-kelompok terbentuk dari proses hasil *Clustering* sebelumnya. Pada proses di tahapan ini digunakanlah rumus *Davies-Bouldin Index* (DBI) guna dihitungnya nilai dari rata-rata di setiap titik pada kelompok data. *Davies-Bouldin Index* (DBI) merupakan langkah yang dipakai guna menghitung akurasi pengelompokan pada metode *clustering*. Kohesi dijabarkan sebagai hasil dari penjumlahan antara kedekatan data dengan pusat cluster dari cluster yang diamati. Proses Pemisahan dibuat berdasarkan pada jarak antara pusat cluster dengan cluster [10]

6) Pengembangan

Dalam proses ini knowledge atau informasi yang sudah didapatkan akan dilakukan penerapan pada sebuah laporan dan aplikasi pada platform website hingga nantinya dapat mudah dipahami serta dipakai dan diterapkan oleh pihak manajemen perusahaan dan stakeholder terkait.

D. Rancangan Pengujian

Pada penelitian ini metode pengujian yang digunakan adalah metode *black box test*. Pengujian *black box* juga dikenal dengan pengujian fungsional, yaitu pengujian yang hanya mengamati hasil eksekusi melalui data pengujian dan memeriksa aplikasi secara fungsional. Pengujian program pada penelitian ini dilakukan oleh pengembang aplikasi untuk menggunakan informasi masukan yang sesuai melalui program yang dikembangkan. Pengujian aplikasi berfokus pada fungsionalitas setiap elemen yang tersedia dan hasil eksekusi yang ditampilkan oleh aplikasi. Pengujian User Interface (UI) aplikasi bertujuan untuk melihat apakah elemen yang terdapat pada aplikasi dapat berjalan dengan baik.

Tabel 2. Pengujian Fungsi Dasar Aplikasi

	Pengujian	Hasil yang diharapkan
1	Tombol <i>Login</i>	Menampilkan halaman menu utama

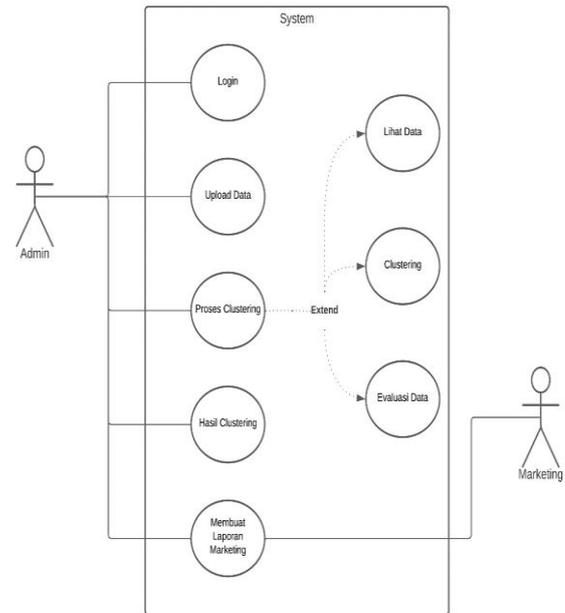
2	Tombol Menu	Menampilkan halaman <i>Home Home</i>
3	Tombol Menu	Menampilkan halaman data Data
4	Tombol Tambah	Menampilkan halaman tambah data
5	Tombol <i>Save</i>	Dapat menyimpan data ke <i>database</i>
6	Tombol <i>Cancel</i>	Dapat reset semua data yang telah diketik
7	Tombol Hapus Semua	Dapat menghapus semua data tabe data pada database
8	Tombol <i>Edit</i>	Menampilkan halaman edit
9	Tombol <i>Delete</i>	Dapat menghapus data yang dipilih pada <i>database</i>
10	Tombol Pilih File	Akses lokasi file excel
11	Tombol <i>Import</i>	Dapat mengunggah isi file excel ke <i>database</i>
12	Tombol Menu Proses <i>Clustering</i>	Menampilkan halaman proses <i>Clustering</i>
13	Tombol Proses	Dapat memproses data pusat <i>Clustering</i> , dan mengakses ke halaman hasil <i>Clustering</i>
14	Tombol Menu Hasil <i>Clustering</i>	Menampilkan halaman hasil <i>Clustering</i>
15	Tombol <i>Detail</i>	Menampilkan halaman <i>detail</i> barang
16	Tombol Menu Laporan	Menampilkan halaman Laporan
17	Tombol Cetak	Menampilkan halaman cetak, dan mencetak hasil Laporan
18	Tombol <i>Logout</i>	Dapat mengakses kembali ke halaman <i>Login</i>

E. Rancangan Basis Data

Rancangan basis data adalah proses penentuan isi dan urutan data sesuai dengan model basis data yang diperlukan untuk mendukung berbagai rancangan sistem.

F. Use Case Diagram

Use Case Diagram dapat digunakan selama proses analisis untuk menangkap *requirement* atau permintaan terhadap perhitungan penentuan banyaknya barang terjual di produk Pigeon di PT. Digital Niaga Indonesia. Berikut ini diagram Use Case Diagram:

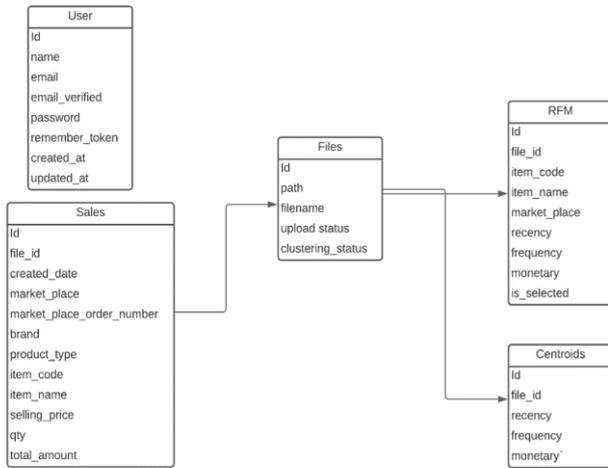


Gambar 1. Use Case Diagram

Pengguna sistem dalam *Use Case Diagram* dapat juga disebut *actor*. *Actor* pada sistem ini akan ada 2 *actor* yaitu admin, dan *marketing*. Admin dapat melakukan 3 kegiatan utama yaitu melakukan unggah data yang telah dibersihkan dan telah ditata dengan menggunakan format xls.menuju sistem. Langkah lebih lanjut user admin akan mendapatkan hasil dari proses apriori serta proses akhir user akan mendapatkan hasil akhir dari proses *Clustering* yang telah dilakukan. Setelah nilai paling akhir yang dari proses *Clustering*, tim marketing membuat hasil dalam bentuk laporan guna menentukan produk yang akan dijual sehingga bisa tercapai target penjualannya

G. Logical Record Structure (LRS)

Logical Record Structure adalah representasi dari struktur record-record pada tabel-tabel yang terbentuk dari hasil relasi antar himpunan entitas. Menentukan kardinalitas, jumlah tabel, dan Foreign Key (FK). Berikut ini diagram LRS:



Gambar 2. Logical Record Structure

III. HASIL DAN PEMBAHASAN

Data yang digunakan untuk pemodelan *Clustering* adalah data transaksi online di marketplace Shopee pada periode Agustus 2022 sampai dengan November 2022. Data diambil secara manual yang dituliskan dalam Microsoft Excel. Tabel 3 dan Tabel 4 merupakan cuplikan data transaksi yang terdiri *created date*, *marketplace*, *marketplace order number*, *brand*, *product type*, *item code*, *item name*, *selling price*, *quantity* dan *total amount*

Tabel 3. Cuplikan Data Transaksi Field 1

<i>created_date</i>	<i>market_place</i>	<i>market_place_order_number</i>	<i>brand</i>	<i>product_type</i>
8/1/2022	PIGEON INDONESIA - SHOPEE	220731UW8P669C	PGN	PGN
8/1/2022	PIGEON INDONESIA - SHOPEE	220731VEDC8FU2	PGN	PGN
8/1/2022	PIGEON INDONESIA - SHOPEE	220731VEDC8FU2	PGN	PGN
8/1/2022	PIGEON INDONESIA - SHOPEE	220731VEDC8FU2	PGN	PGN
8/1/2022	PIGEON INDONESIA - SHOPEE	220731T6CA CH1G	PGN	BLOT
8/1/2022	PIGEON INDONESIA - SHOPEE	220730RM04B1XJ	PGN	BLOT
8/1/2022	PIGEON INDONESIA - SHOPEE	220730RM04B1XJ	PGN	BLOT

Tabel 4. Cuplikan Data Transaksi Field 2

<i>item_code</i>	<i>item_name</i>	<i>selling_price</i>	<i>qty</i>	<i>total_amount</i>
APGNRFED014	PIGEON MAG MAG STEP 3 - STRAW CUP	64700	1	64700
APGNRBOT072	PIGEON BOTOL SUSU KACA WN 160ML - ORANGE	189900	2	379800

APGNRBOT073	PIGEON BOTOL SUSU KACA WN 160ML - GREEN	189900	2	379800
APGNRNIP001	PIGEON DOT PERISTALTIC PLUS SIZE SS ISI 1 PCS	43415	1	43415
APGNRBOT296	PIGEON BTL PP RP 240 ML RUSA W/ S TYPE SILICONE NI	50500	1	50500
APGNRBOT301	PIGEON BOT T ESTER WN 200ML LG	169900	1	169900
APGNRBOT302	PIGEON BOT T ESTER WN 200ML LO	169900	1	169900

Mengekstraksi data produk berdasarkan tiga variabel RFM setiap pelanggan 4 bulan terakhir (bulan Agustus sampai bulan November) di marketplace Shopee. Setiap pelanggan diakumulasi masing-masing nilai RFM dalam periode 4 bulan terakhir.

Yang pertama adalah mencari nilai Recency. Recency merupakan waktu transaksi terakhir suatu produk. Dengan menghitung waktu transaksi terakhir dari produk yang dibeli dan dibandingkan dengan tanggal penelitian ini dilakukan. Jumlah hari dari selisih tanggal tersebut akan dijadikan nilai Recency.

Tabel 5. Tabel Recency

<i>No</i>	<i>Nama Item</i>	<i>Transaksi Terakhir</i>	<i>Recency</i>
1	PIGEON SPONGE BRUSH	29/11/2022	39
2	PIGEON BOTTLE AND NIPPLE BRUSH	30/11/2022	38
3	PIGEON NIPPLE CLEANING BRUSH	23/11/2022	45
4	PIGEON MAG MAG STRAW BRUSH	30/11/2022	38
5	PIGEON COTTON BALL	30/11/2022	38
....
568	PIGEON TEENS JELLY TINT RASPBERRY MACCHIATO	30/11/2022	38
569	PIGEON TEENS JELLY TINT CRANBERRY MOUSSE	30/11/2022	38

Tahap yang dilakukan selanjutnya yaitu melakukan pencarian nilai Frequency. Nilai Frequency adalah jumlah produk yang terjual dalam satu kali periode data Nilai.

Tabel 6. Tabel Frequency

<i>No</i>	<i>Nama Produk</i>	<i>Frequency</i>
1	PIGEON SPONGE BRUSH	51

2	PIGEON BOTTLE AND NIPPLE BRUSH	186
3	PIGEON NIPPLE CLEANING BRUSH	28
4	PIGEON MAG MAG STRAW BRUSH	54
5	PIGEON COTTON BALL	415
....
568	PIGEON TEENS JELLY TINT RASPBERRY MACCHIATO	28
569	PIGEON TEENS JELLY TINT CRANBERRY MOUSSE	20

Tahap selanjutnya adalah mencari nilai monetary yang dihasilkan oleh produk yang dijual dalam satu periode data, sesuai data pada berikut :

Tabel 7. Tabel Monetary

No	Nama Produk	Monetary
1	PIGEON SPONGE BRUSH	6.226.280
2	PIGEON BOTTLE AND NIPPLE BRUSH	12.694.803
3	PIGEON NIPPLE CLEANING BRUSH	2.823.760
4	PIGEON MAG MAG STRAW BRUSH	3.152.820
5	PIGEON COTTON BALL	21.169.350
....
568	PIGEON TEENS JELLY TINT RASPBERRY MACCHIATO	1.021.800
569	PIGEON TEENS JELLY TINT CRANBERRY MOUSSE	701.460

Proses Perhitungan *Algoritme K-Means* yang dilakukan setelahnya ialah melakukan proses hitung dengan *Algoritme K-Means*. Adapun tahap awal yang dilakukan ialah mencari serta menunjuk jumlah *Cluster*. Pada riset penelitian ini *Cluster* ditentukanlah sebanyak tiga *Cluster* diantaranya yaitu produk yang paling laris, produk laris, dan produk tidak laris. Pada tiap *Cluster* yang telah ditentukan sebelumnya akan memiliki pusat *Cluster* atau disebut juga *Centroid*. Nilai yang ada pada nilai *Centroid* tersebut akan dilakukan proses hitung secara berulang-ulang hingga nantinya akan didapatkan nilai *Centroid* yang paling sempurna ataupun juga *Centroid* yang nilainya tidak akan berubah-ubah, hingga *Cluster* yang didapatkan tetap ataupun tidak terdapat produk yang masih berpindah *Cluster*. Guna melakukan proses hitung iterasi *Cluster* yang pertama, maka nilai *Centroid* awal akan diberikan secara acak ataupun random yang dijelaskan pada tabel 8:

Tabel 8. Tabel Proses Perhitungan Clustering

No	Nama Produk	Recency	Frequency	Monetary
1	PIGEON SPONGE BRUSH	39	51	6.226.280

2	PIGEON BOTTLE AND NIPPLE BRUSH	38	186	12.694.803
3	PIGEON NIPPLE CLEANING BRUSH	45	28	2.823.760
4	PIGEON MAG MAG STRAW BRUSH	38	54	3.152.820
5	PIGEON COTTON BALL	38	415	21.169.350
....
568	PIGEON TEENS JELLY TINT RASPBERRY MACCHIATO	38	28	1.021.800
569	PIGEON TEENS JELLY TINT CRANBERRY MOUSSE	38	20	701.460

Dalam melakukan penentuan dari nama *Cluster* maka diperlukan nilai dari *Centroid* akhir tersebut. Tahapan selanjutnya merupakan proses mengurutkan atau pembuatan ranking dalam nilai RFM dengan rumus sesuai berikut :

1. nilai rangking akan semakin besar jika semakin kecil nilai *Recency*
2. nilai rangkingnya akan semakin besar jika semakin besar nilai *Frequency*
3. nilai rangkingnya semakin besar jika semakin besar nilai *Monetary*
4. Nilai rata-rata terbesar dari rangking milik *Cluster* adalah *Cluster* paling laris, sedangkan nilai rata-rata terkecil adalah *Cluster* kurang laris, sisanya adalah *Cluster* laris.

Berikut adalah nilai *ranking* yang didapatkan oleh setiap *Cluster* berdasarkan *Centroid* akhir adalah sebagai berikut:

Tabel 9. Tabel Penentuan Nama Cluster

Cluster	R	F	M	Ranking	Nama Cluster
C1	38.81	403.79	Rp20,633,478	2	Laris
C2	38.35	1528.7	Rp63,420,698	3	Paing Laris
C3	54.764	57.501	Rp3,053,692	1	Kurang Laris

Dari setiap anggota data yang dimiliki setiap kluster, dapat diberikan informasi yaitu setiap anggota data pada dalam kluster tersebut mempunyai kedekatan dalam kriteria. Rata-rata pada anggota kluster dengan nilai kurang laris terinformasi memiliki nilai *recency* yang lebih besar dan nilai dari *frequency* yang kecil, serta nilai *monetary* yang juga kecil. Adapun yang terdapat pada kluster yang paling

laris yaitu memiliki nilai dari *recency* yang kecil, dan juga nilai *frequency* yang besar, serta nilai dari *monetary* yang besar.

Dalam melakukan proses pengukuran pada nilai validitas Cluster yang telah didapatkan pada Algoritma K-Means dipenelitian ini, maka penulis menerapkan metode Davies-Boudin Index (DBI). Tahapan awal yang telah diterapkan untuk metode DBI adalah melakukan proses hitung pada nilai Sum of Square Within Cluster (SSW) ke seluruh data. Berikut ini merupakan proses hitung Sum of Square Within Cluster (SSW) pada suatu data dengan memakai rumus dibawah ini:

PIGEON SPONGE BRUSH

SSW1

$$= \sqrt{(39 - 38.81)^2 + (51 - 403.79)^2 + (6226280 - 20633478)^2}$$

$$= 14407198$$

SSW2

$$= \sqrt{(39 - 38.35)^2 + (51 - 1528.7)^2 + (6226280 - 63420698)^2}$$

$$= 57194418$$

SSW3

$$= \sqrt{(39 - 54.764302)^2 + (51 - 57.501144)^2 + (6226280 - 3053692)^2}$$

$$= 3172588$$

Setelah menghitung SSW pada seluruh data. Selanjutnya adalah menghitung rata-rata nilai SSW1 sampai SSW3.

Langkah selanjutnya yang dilakukan adalah menghitung rasio, dengan cara menghitung jarak antara SSW agar mendapatkan nilai rasio terbesar (R-max), yang selanjutnya akan digunakan untuk mendapatkan nilai DBI menggunakan rumus diatas:

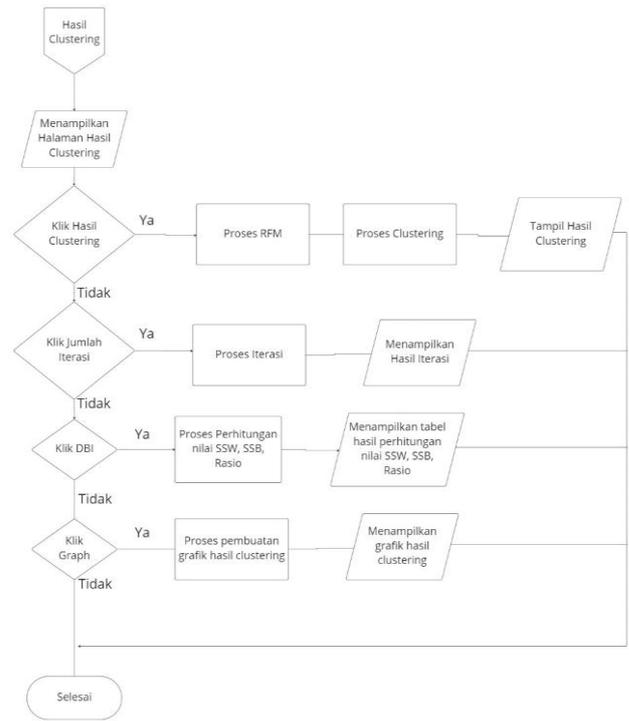
$$R1,2 = \frac{5844981.782+17455336.43}{42787219.62} = 0.5446$$

$$R1,3 = \frac{5844981.782+2435693.779}{17579786.01} = 0.4710$$

$$R2,3 = \frac{17455336.43+2435693.779}{60367005.63} = 0.3295$$

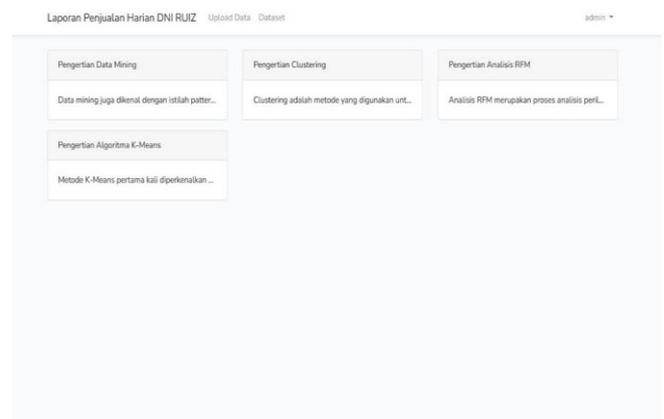
Jika pada perhitungan didapatkan nilai dengan rasio terbesar (R-Max) sudah muncul, maka proses atau tahapan selanjutnya yaitu melakukan proses hitung dari nilai DBI nya dengan cara menghitung rata-rata dari nilai R-Max. berdasarkan pada proses tersebut didapatkanlah nilai DBI sebesar 0.520. Adapun nilai dari DBI akan dapat diberikan label baik jika nilai yang dihasilkan mendekati 0.

Flowchart dari hasil proses Clustering ialah alur dari halaman hasil Clustering yang juga pada halaman ini menampilkan tabel hasil RFM setiap produk dan tabel hasil setiap proses Clustering. Berikut adalah Flowchart hasil Clustering yang dapat dilihat pada gambar 10:



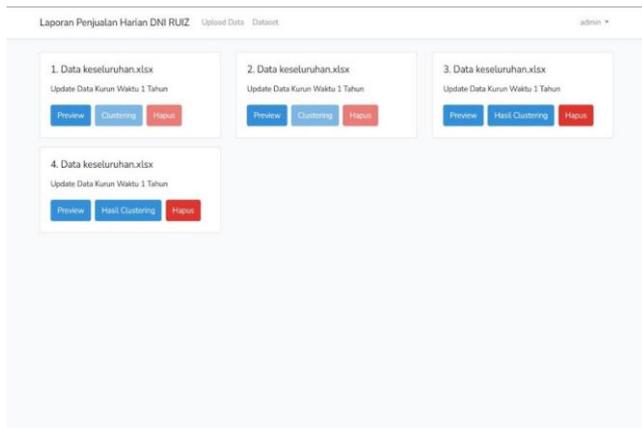
Gambar 10. Flowchart Hasil Clustering

Tampilan layar Home ini akan tampil pertama kali setelah berhasil masuk melewati halaman login. Pada halaman ini pengguna akan ditampilkan profil PT. Digital Niaga Indonesia dan keterangan tentang data mining Clustering dan RFM, di bagian atas bar terdapat berbagai pilihan menu, seperti menu Upload data, menu Dataset, dan log out. Tampilan layar Home dapat dilihat pada gambar 11:



Gambar 11. Tampilan Layar Home

Tampilan layar Dataset menampilkan data yang sudah diunggah. Tampilan layar Dataset dapat dilihat pada Gambar 12:



Gambar 12. Tampilan Layar Dataset

Tampilan layar *hasil Clustering* menampilkan halaman clustering.

No	Kode Item	Nama Item	Market Place	Reccency	Frequency	Monetary	Cluster
1	APONRACC002	PIGION SPONGE BRUSH	TOKOPEDIA	41	72	Rp 9.093.220	Cluster 1
2	APONRACC003	PIGION BOTTLE AND NIPPLE BRUSH	TOKOPEDIA	44	73	Rp 5.203.660	Cluster 1
3	APONRACC004	PIGION NIPPLE CLEANING BRUSH	TOKOPEDIA	41	48	Rp 5.248.404	Cluster 1
4	APONRACC005	PIGION MAG MAG STRAW BRUSH	TOKOPEDIA	41	38	Rp 1.984.845	Cluster 2
5	APONRACC006	PIGION COTTON BALL	TOKOPEDIA	39	161	Rp 14.762.971	Cluster 2
6	APONRACC007	PIGION COTTON SWAB ISI 200 PCS - SMALL TIP	TOKOPEDIA	39	892	Rp 20.922.769	Cluster 2
7	APONRACC008	PIGION COTTON SWAB ISI 200 PCS	TOKOPEDIA	39	86	Rp 2.930.363	Cluster 3
8	APONRACC009	PIGION COTTON SWAB ISI 200 PCS - MIX 2 SIZE	TOKOPEDIA	40	74	Rp 2.686.509	Cluster 3

Gambar 13. Tampilan Layar hasil Clustering

IV. KESIMPULAN DAN SARAN

Berdasarkan hasil pengujian dan evaluasi dari aplikasi yang dibuat menggunakan Dataset dan Algoritma yang telah diusulkan, maka dapat disimpulkan bahwa:

- Evauasi *Cluster* menggunakan perhitungan DBI menghasilkan 0.520 yang dimana nilai ini bagus karena jika nilai mendekati 0 maka hasil yang diperoleh akan semakin bagus.
- Pengujian aplikasi didapatkan 557 produk dengan hasil akhir akhir 3 Cluster yaitu Cluster 1 (C1) 100 data, Cluster 2 (C2) sebanyak 20 data, dan Cluster 3 (C3) sebanyak 437 data dari 124473 data transaksi.
- Metode RFM dapat menentukan bahwa Cluster 1 adalah produk yang kurang laris, Cluster 2 paling laris, dan Cluster 3 laris.
- Dengan adanya pengelompokan data ini diharapkan, manajemen dan tim marketing PT. Digital Niaga Indonesia dapat mengetahui barang yang laris, kurang laris oleh pembeli di setiap marketplace dengan mudah.

Adapun saran yang dapat dibuat untuk pengembangan program lebih lanjut dengan harapan dapat menghasilkan penelitian yang lebih baik lagi, berikut saran yang dapat diberikan:

- Diharapkan pada penelitian selanjutnya bisa ditambahkan parameter atau field lain untuk proses pengelompokan.
- Dalam penelitian berikutnya diharapkan aplikasi dapat menerima file jenis lain seperti .csv, dll
- Diharapkan ada metode lain untuk pengelompokan data.

DAFTAR PUSTAKA

- [1] S. P. Sari, "Strategi Meningkatkan Penjualan Di Era Digital," *Sci. J. Reflect. Econ. Accounting, Manag. Bus.*, vol. 3, no. 3, pp. 291–300, 2020.
- [2] Y. Dharma Putra, M. Sudarma, and I. B. A. Swamardika, "Clustering History Data Penjualan Menggunakan Algoritma K-Means," *Maj. Ilm. Teknol. Elektro*, vol. 20, no. 2, p. 195, 2021.
- [3] Normah, B. Rifai, S. Vambudi, and R. Maulana, "Analisa Sentimen Perkembangan Vtuber Dengan Metode Support Vector Machine Berbasis SMOTE," *J. Tek. Komput. AMIK BSI*, vol. 8, no. 2, pp. 174–180, 2022.
- [4] . F., F. T. Kesuma, and S. P. Tamba, "Penerapan Data Mining Untuk Menentukan Penjualan Sparepart Toyota Dengan Metode K-Means Clustering," *J. Sist. Inf. dan Ilmu Komput. Prima (JUSIKOM PRIMA)*, vol. 2, no. 2, pp. 67–72, 2020.
- [5] N. Afiasari, N. Suarna, and N. Rahaningsi, "Implementasi Data Mining Transaksi Penjualan Menggunakan Algoritma Clustering dengan Metode K-Means," *J. SAINTEKOM*, vol. 13, no. 1, pp. 100–110, 2023.
- [6] R. A. Kurniawan, M. S. Hasibuan, P. Piramida, and R. S. Ramadhan, "Penerapan Algoritma K-Means Untuk Clustering Tempat Makan Di Batubara," *J. Comput. Sci. Informatics Eng.*, vol. 01, no. 1, pp. 10–18, 2022.
- [7] I. K. Juni Arta, G. Indrawan, and G. R. Dantes, "Data Mining Rekomendasi Calon Mahasiswa Berprestasi Di Stmik Denpasar Menggunakan Metode Technique for Others Reference By Similarity To Ideal Solution," *JST (Jurnal Sains dan Teknol.)*, vol. 5, no. 2, pp. 11–21, 2017.
- [8] T. K. Titus and M. Jajuli, "Clustering Data Kecelakaan Lalu Lintas di Kecamatan Cileungsi Menggunakan Metode K-Means," *Gener. J.*, vol. 6, no. 1, pp. 1–12, 2022.
- [9] F. Nurdiyansyah and I. Akbar, "Implementasi Algoritma K-Means untuk Menentukan Persediaan Barang pada Poultry Shop," *J. Teknol. dan Manaj. Inform.*, vol. 7, no. 2, pp. 86–94, 2021.
- [10] T. M. Dista and F. F. Abdulloh, "Clustering Pengunjung Mall Menggunakan Metode K-Means dan Particle Swarm Optimization," *J. Media Inform. Budidarma*, vol. 6, no. 3, p. 1339, 2022.
- [11] A. D. Savitri, F. A. Bachtiar, and N. Y. Setiawan,

- “Segmentasi Pelanggan Menggunakan Metode K-Means Clustering Berdasarkan Model RFM Pada Klinik Kecantikan (Studi Kasus: Belle Crown Malang),” *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 2, no. 9, pp. 2957–2966, 2018.
- [12] M. Fithriyah, M. A. Yaqin, and S. Zaman, “K-Means Clustering Untuk Segmentasi Produk Berdasarkan Analisis Recency, Frequency, Monetary (RFM) Pada Data Transaksi Penjualan,” *Ilk. J. Comput. Sci. Appl. Informatics*, vol. 3, no. 2, pp. 151–164, 2021.
- [13] A. T. Widiyanto and A. Witanti, “Segmentasi Pelanggan Berdasarkan Analisis RFM Menggunakan Algoritma K-Means Sebagai Dasar Strategi Pemasaran (Studi Kasus PT Coversuper Indonesia Global),” *KONSTELASI Konvergensi Teknol. dan Sist. Inf.*, vol. 1, no. 1, pp. 204–215, 2021.
- [14] Haris Kurniawan, Sarjon Defit, and Sumijan, “Data Mining Menggunakan Metode K-Means Clustering Untuk Menentukan Besaran Uang Kuliah Tunggal,” *J. Appl. Comput. Sci. Technol.*, vol. 1, no. 2, pp. 80–89, 2020.
- [15] D. Astuti, “Penentuan Strategi Promosi Usaha Mikro Kecil Dan Menengah (UMKM) Menggunakan Metode CRISP-DM dengan Algoritma K-Means Clustering,” *J. Informatics, Inf. Syst. Softw. Eng. Appl.*, vol. 1, no. 2, pp. 60–72, 2019.