

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi yang sangat pesat di era globalisasi saat ini telah memberikan banyak manfaat dalam kemajuan diberbagai aspek sosial. Penggunaan teknologi oleh manusia dalam membantu menyelesaikan pekerjaan merupakan hal yang menjadi keharusan dalam kehidupan. Perkembangan teknologi ini juga harus diikuti dengan perkembangan pada Sumber Daya Manusia (SDM) [1].

Salah satu aspek penting dalam dunia usaha ritel adalah kemampuan untuk menganalisis data penjualan sebagai dasar pengambilan keputusan strategis. Dengan volume transaksi yang tinggi dan variasi produk yang sangat beragam, pengelolaan dan pemanfaatan data penjualan menjadi tantangan tersendiri bagi pelaku usaha. Berdasarkan hasil wawancara, Habib *Stationery* merupakan salah satu toko retail yang menjual berbagai jenis barang kebutuhan harian dan umum, yang dalam operasionalnya telah mengumpulkan data penjualan dari berbagai jenis produk. Namun, hingga saat ini, pengolahan data penjualan di Habib *Stationery* masih dilakukan secara konvensional dan belum dimanfaatkan secara maksimal untuk menghasilkan informasi yang mendukung strategi bisnis.

Pemilik toko belum memiliki gambaran yang jelas tentang kelompok produk yang memiliki pola penjualan serupa, sehingga sering kali terjadi penumpukan stok pada barang yang kurang diminati dan kekurangan stok pada

barang yang laris. Masalahnya lebih rumit ketika jumlah produk yang dikelola meningkat dan ada fitur penjualan yang tidak linier atau seragam dalam jenis

produk. Beberapa produk yang penjualan terbaik hanya pada waktu tertentu (musim), sementara yang lain biasanya konsisten, tetapi dalam jumlah kecil. Sehingga diperlukan suatu sistem berbasis web yang tidak hanya memungkinkan manajemen data, tetapi juga memungkinkan klasifikasi produk yang akurat dan adaptif berdasarkan kemampuan penjualan secara cerdas menjadi salah satu kendala dalam pemanfaatan data penjualan secara optimal.

Salah satu teknik yang dapat dimanfaatkan untuk menganalisis data penjualan tersebut adalah *data mining*. *Data mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan dan pembelajaran mesin (*machine learning*) mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database yang terkait [2]. Pada *data minning* terdapat proses bentuk pengelompokan (*clustering*).

Clustering adalah metode penganalisis data, yang sering dimasukkan sebagai salah satu metode *data mining*, yang bertujuan adalah untuk mengelompokkan data dengan karakteristik yang sama ke suatu ‘wilayah’ yang sama dan data dengan karakteristik yang berbeda ‘wilayah’ yang lain [3]. Salah satu algoritma *clustering* yang dapat digunakan adalah K-Medoids, Algoritma k-medoids merupakan teknik partisi *clustering* untuk membuat pengelompokkan data sejumlah n objek yang akan menghasilkan sebuah *cluster*. Metode ini memakai sebuah bahan di sekumpulan objek untuk mewakili sebuah *cluster* dari data untuk digunakan [4].

Beberapa penelitian telah melakukan penelitian tentang metode K-Medoids. Seperti yang dilakukan oleh Anggi Ayu Dwi Sulistyawati dan Mujiono

Sadikin pada tahun 2021 dengan judul penerapan algoritma K-Medoids untuk menentukan segmentasi pelanggan. Dari hasil penelitian dapat disimpulkan segmentasi pelanggan dengan menerapkan algoritma K-Medoids menghasilkan jumlah *cluster* optimal adalah 3 (tiga). Secara umum dapat disimpulkan bahwa segmentasi paling baik untuk kasus dataset Perum BULOG ini adalah 3 (tiga) segmen yaitu *Lost Customer*, *Core Customer*, dan *New Customer*. Segmentasi pelanggan dapat mengidentifikasi tipe dan karakteristik pelanggan berdasarkan perilaku atau kebiasaan yang sama. Hasil segmentasi pelanggan dapat menjadi acuan dalam menyusun rekomendasi strategi pemasaran [5].

Penelitian lainnya yang telah diteliti oleh Sukma Sindi, Weni Ratnasari Orktapia Ningse, dkk pada tahun 2020 dengan judul analisis algoritma K-Medoids *clustering* dalam pengelompokan penyebaran covid-19 di Indonesia. Dari hasil penelitian dapat disimpulkan algoritma K-Medoids dapat melakukan pengelompokan data covid-19 mana saja wilayah yang terinfeksi di wilayah masing-masing-masing dengan pengklasteran terbaik dilakukan dengan 3 *cluster*. Dari 34 *record* diperoleh 1 *record* pada *cluster* pertama, 2 *record* pada *cluster* kedua, 31 *record* pada *cluster* ketiga [6].

Dari pemaparan diatas, penelitian ini bertujuan untuk merancang dan membuat sebuah sistem untuk pengelompokan penjualan barang berbasis *web* khususnya di *Habib Stationery*. Penerapan algoritma K-Medoids diharapkan mampu mengelompokkan data penjualan berdasarkan pola tertentu seperti kategori produk, frekuensi penjualan, dan tingkat permintaan. Hasil pengelompokan ini nantinya dapat digunakan oleh pihak manajemen untuk

menentukan strategi bisnis seperti pengadaan barang, promosi, dan penyusunan display produk di toko. Selain itu, dengan membangun sistem ini dalam bentuk aplikasi berbasis *web*, maka proses analisis dapat dilakukan dengan lebih mudah, cepat, dan dapat diakses oleh pihak manajemen di berbagai lokasi secara *real time*.

1.2 Rumusan Masalah

Dari latar belakang yang telah dikemukakan, adapun rumusan masalah dalam penelitian ini adalah bagaimana merancang dan membangun sistem berbasis *web* yang dapat mengimplementasikan algoritma K-Medoids untuk menampilkan hasil pengelompokan data penjualan pada Habib *Stationery* ?

1.3 Tujuan Penelitian

Adapun tujuan penelitian ini adalah merancang dan membangun sistem berbasis *web* yang dapat mengimplementasikan algoritma K-Medoids untuk menampilkan hasil pengelompokan data penjualan pada Habib *Stationery*.

1.4 Batasan Masalah

Agar penelitian ini lebih terarah dan sesuai dengan tujuan yang ingin dicapai, maka adapun batasan masalah dalam penelitian ini adalah sebagai berikut :

1. Metode yang digunakan dalam pengelompokan penjualan barang pada Habib *Stationery* adalah metode K-Medoids.
2. Data penjualan yang digunakan dalam penelitian ini adalah data penjualan tahun 2024.

3. Pada penelitian ini terdiri dari 3 *cluster* yaitu barang laku keras, jarang laku dan tidak laku.
4. Sistem yang dikembangkan berbasis *web* dengan menggunakan bahasa pemrograman PHP dan MySQL sebagai *database* dan hanya difokuskan pada fungsi utama pengelompokan dan visualisasi hasil *clustering*, tanpa mencakup fitur manajemen stok, laporan keuangan, atau integrasi sistem kasir.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah membantu pihak manajemen Habib Stationery dalam memahami pola penjualan barang melalui hasil pengelompokan, sehingga dapat digunakan sebagai dasar dalam pengambilan keputusan bisnis seperti pengadaan stok, penataan produk, dan strategi promosi.

1.6 Sistematika Penulisan

Sistematika penulisan dari skripsi ini terdiri sebagai berikut :

BAB 1 PENDAHULUAN

Bab ini berisi latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

BAB 2 LANDASAN TEORI

Bab ini berisi teori-teori yang digunakan pada penelitian ini. Teori-teori yang berhubungan dengan *data mining*, *clustering*, K-Medoids dan penjualan barang.

BAB 3 METODOLOGI PENELITIAN

Bab ini berisi tahapan-tahapan dalam pengumpulan data, perancangan sistem perumusan masalah dan Analisis.

BAB 4 ANALISIS DAN PERANCANGAN

Bab ini berisi Analisis dan perancangan aplikasi pengelompokan penjualan barang pada Habib *Stationary* dengan metode K-Medoids.

BAB 5 IMPLEMENTASI DAN PENGUJIAN

Bab ini berisi implementasi dari Analisis dan perancangan dan pengujian pada aplikasi yang berhasil dibangun.

BAB 6 PENUTUP

Bab ini berisi rangkuman dari hasil penelitian yang telah dilakukan dan saran-saran untuk pengembangan aplikasi atau penelitian selanjutnya.

BAB 2

LANDASAN TEORI

2.1 *Data Mining*

Data mining merupakan suatu proses pencarian teknik analisis data yang besar untuk proses yang diperoleh dari berbagai macam basis data seperti data relasional data berorientasi objek dan data transaksi untuk mengetahui informasi yang baru didalam *database*. *Data mining* merupakan proses interaktif untuk menemukan pola atau model baru yang sempurna, bermanfaat dan dapat dimengerti dalam suatu *database* yang sangat besar. *Data mining* berisi pencarian *trend* atau pola yang diinginkan dalam *database* besar untuk membantu pengambil keputusan yang akan datang, pola-pola ini dikenali perangkat tertentu yang dapat memberikan suatu Analisis data yang berguna dan berwawasan yang kemudian dapat dipelajari dengan lebih teliti, yang mungkin saja menggunakan perangkat pendukung keputusan yang lain [7].

Sebuah pengetahuan dan informasi-informasi penting yang bisa menaikkan keuntungan, menekan biaya pengeluaran, bisa digali dengan proses *data mining*. Beberapa teknik yang digunakan dalam data mining antara lain matematika, statistik, *artificial intelligence*, dan *machine learning* yang fungsinya mengekstrak dan mengidentifikasi informasi dan pengetahuan berguna yang saling terhubung di beragam *database* [8].

Data mining adalah gabungan sejumlah disiplin ilmu komputer yang mendefinisikan sebagai proses penemuan pola-pola baru dari kumpulan - kumpulan

data sangat besar, meliputi metode-metode yang merupakan irisan dari *artificial intelligence*, *machine learning*, *statistics*, dan *database systems* [9].

2.1.1 **Knowledge Discovery in Database (KDD)**

Definisi *Knowledge Discovery in Database (KDD)* adalah metode dan cara mendapatkan sebuah informasi melalui basis data yang telah tersedia [10]. *KDD* adalah metode yang digunakan dalam penelitian ini. *KDD* adalah teknik untuk mendapatkan informasi berupa data dari basis data yang sudah ada. Ada tabel yang terhubung satu sama lain dalam basis data. Informasi dari data yang dikumpulkan melalui proses *KDD* dapat berfungsi sebagai *database* untuk pengambilan keputusan, *Knowledge Discovery in Database (KDD)* dan *data mining* adalah kata-kata yang sering digunakan secara bergantian untuk menggambarkan proses pengambilan informasi tersembunyi dari basis data yang besar. Istilah *KDD* dan *data mining* memiliki keterkaitan satu sama lain, tetapi memiliki konsep yang berbeda [11].

Knowledge Discovery in Database (KDD) adalah keseluruhan proses untuk mencari dan mengidentifikasi pola dalam data, dimana pola yang ditemukan bersifat sah, baru, dapat bermanfaat dan dapat dimengerti. *KDD* berhubungan dengan teknik integrasi dan penemuan ilmiah, interpretasi dan visualisasi dari pola-pola sejumlah kumpulan data. *KDD* dan *data mining* sering kali digunakan secara bergantian untuk menjelaskan proses penggalian informasi [12].

Beberapa tahapan dalam proses *KDD* adalah sebagai berikut [12] :

1. Pemilihan Data (*Data Selection*)

Sering kali terdapat data yang tidak terpakai dalam *database*. Hanya data yang sesuai untuk dianalisis yang diperlukan. Sebagai contoh, untuk meneliti kebiasaan meminjam, tidak perlu mengambil data nama pengunjung, cukup dengan id pengunjung saja.

2. Pembersih Data (*Data Cleaning*)

Pada kenyataannya, data yang didapat dari suatu *database* belum tentu memiliki kualitas yang cukup baik. Misalnya data tersebut tidak lengkap atau ada informasi yang hilang, maupun data tidak valid, juga terdapat atribut-atribut data yang tidak relevan terhadap teknik *data mining* yang digunakan. *Data cleaning* bertujuan untuk membuang data-data yang tidak konsisten, menghilangkan *noise* dan melengkapi data yang kehilangan informasi, sehingga performansi dari *data mining* dapat meningkat.

3. Transformasi Data (*Data Transformation*)

Pada tahapan ini, data diubah atau ditransformasikan menjadi format data yang sesuai untuk diproses dalam *data mining*, sebab ada metode-metode *data mining* yang memerlukan format data tertentu untuk diolah. Proses mentransformasikan data yang telah dipilih sehingga sesuai untuk *data mining* adalah *coding*. Proses *coding* dalam *KDD* sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

4. *Data Mining*

Metode dan algoritma yang telah ditentukan mulai diterapkan untuk mencari pola dan menemukan informasi berharga yang tersembunyi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses *KDD* secara keseluruhan.

5. Integritasi Data (*Data Integration*)

Data yang akan diproses dalam *data mining* dapat berasal dari berbagai *database*, dan bukan hanya dari satu *database*. Integrasi data diperlukan untuk menggabungkan data dari berbagai sumber data kedalam satu *database* baru. Integrasi yang teliti dapat mengurangi dan menolak redundansi data, sehingga dapat meningkatkan akurasi dan kecepatan dari proses *data mining*.

6. Evaluasi Pola (*Pattern Evaluation*)

Pola informasi yang dihasilkan dari proses *data mining* mungkin berbeda dan tidak sesuai dengan hipotesa. Bila hal ini terjadi, hasil tersebut dapat dijadikan umpan balik untuk memperbaiki proses *data mining*. Solusi lain adalah dengan mengubah metode yang digunakan, atau menerima hasil yang ada sebagai pengetahuan baru yang mungkin dapat bermanfaat.

7. Presentasi Pengetahuan (*Knowledge Presentation*)

Knowledge presentation merupakan tahapan akhir dalam proses *data mining*. Bagaimana pengetahuan yang telah ditemukan akan disajikan kepada *user*. Tidak semua *user* memahami *data mining*, karenanya penting untuk menyusun dengan baik penyajian hasil *data mining* dalam bentuk

yang dapat dipahami oleh *user*. Dalam hal ini, visualisasi juga dapat digunakan untuk membantu menyampaikan hasil *data mining*.

2.2 Penjualan

Penjualan (*sales*) adalah aktivitas atau bisnis dalam menjual produk atau jasa. Pengertian penjualan secara umum adalah kegiatan jual beli dijalankan oleh dua belah pihak atau lebih dengan alat pembayaran yang sah [13].

Penjualan sebagai salah satu fungsi pemasaran yang menentukan dalam usaha mencapai tujuan perusahaan. penjualan sebagai suatu kegiatan yang ditujukan untuk mencari pembeli, memengaruhi, dan memberikan petunjuk agar pembeli dapat menyesuaikan kebutuhannya dengan produk yang ditawarkan serta mengadakan perjanjian mengenai harga yang menguntungkan bagi kedua belah pihak [14].

Penjualan merupakan pembelian sesuatu (barang atau jasa) dari suatu pihak kepada pihak lainnya dengan mendapatkan ganti uang dari pihak tersebut. Penjualan juga merupakan suatu sumber pendapatan perusahaan, semakin besar penjualan maka semakin besar pula pendapatan yang diperoleh perusahaan [15].

2.3 Clustering

Pengertian *clustering* dalam *data mining* adalah mengelompokkan sekumpulan data atau objek ke dalam klaster (kelompok) sehingga semua data dalam suatu klaster berisi data yang semirip mungkin dengan objek di klaster lain saja. *Clustering* adalah proses membagi data ke dalam kelas atau cluster berdasarkan kesamaan [16].

Clustering merupakan proses partisi satu set objek data ke dalam himpunan bagian yang disebut dengan *cluster*. Objek yang di dalam *cluster* memiliki kemiripan karakteristik antar satu sama lainnya dan berbeda dengan *cluster* yang lain. Partisi tidak dilakukan secara manual melainkan dengan suatu algoritma *clustering*. Oleh karena itu, *clustering* sangat berguna dan bisa menemukan group atau kelompok yang tidak dikenal dalam data. *Clustering* banyak digunakan dalam berbagai aplikasi seperti misalnya pada *business intelligence*, pengenalan pola citra, *web search*, bidang ilmu biologi, dan untuk keamanan (*security*) [17].

Clustering adalah sekelompok atau sekumpulan objek-objek data yang similar satu sama lain dalam *clustering* yang sama dan disimilar terhadap objek-objek yang berbeda. Objek akan dikelompokkan ke dalam satu atau lebih *clustering* sehingga objek-objek yang berada dalam satu *clustering* akan mempunyai kesamaan yang tinggi antara satu dengan yang lainnya [18].

2.4 Algoritma K-Medoids

Metode K-Medoids merupakan sebuah algoritma pada *clustering* yang sangat efektif digunakan pada dataset yang berukuran kecil. Langkah awal dalam metode ini adalah dengan mencari titik representatif pada sebuah dataset dengan menghitung jarak antara titik pada setiap kelompok dari semua kemungkinan kombinasi medoids. Tujuannya adalah agar jarak pada titik dalam suatu kelompok menjadi kecil, sedangkan jarak pada kelompok menjadi besar. Keuntungan dari algoritma K-Medoids adalah kemampuannya dalam mengatasi kelemahan algoritme K-Means, yaitu sensitif terhadap *noise* dan *outlier* [19].

Algoritma K-Medoids berfungsi untuk mengurangi jumlah perbedaan antara titik data di setiap kluster yang memiliki titik data yang dipilih sebagai pusat (medoid). Ciri khas algoritma ini adalah pusat kluster terletak di antara titik datanya. Metode jarak geometris digunakan untuk menghitung jarak antara objek dan pusat (medoid). Metode ini memastikan bahwa objek yang dipilih secara acak memiliki kemiripan yang sangat besar dengan objek medoidnya [20].

Algoritma K-Medoids lebih baik diterapkan pada data yang sangat besar. Untuk data yang berbeda parameternya, sebelum dihitung jaraknya terlebih dahulu lakukan normalisasi data agar data seimbang. Cara kerja K-Medoids adalah sebagai berikut [21] :

1. Menentukan jumlah pada kluster yang diinginkan untuk inisialisasi (k)
2. Alokasi setiap objek ke kluster terdekat dengan menggunakan rumus perhitungan jarak *Euclidean Distance* seperti yang ditunjukkan dalam rumus berikut :

$$\sqrt{(x_i - X)^2 + \dots + (y_n - Y)^2} \dots \dots \dots (1)$$

3. Mengambil objek secara acak dari setiap kluster untuk diposisikan sebagai calon medoid yang baru.
4. Menentukan jarak antara masing-masing objek dalam setiap kluster dan calon medoid yang baru.
5. Nilai total simpangan (S) diperoleh dengan persamaan berikut :

$$S = b - a \dots \dots \dots (2)$$

a = jumlah jarak terdekat antara objek ke medoid awal

b = jumlah jarak terdekat antara objek ke medoid baru

Jika $S < 0$, maka tukar objek dengan data untuk menghasilkan sekumpulan k baru sebagai medoid.

6. Lakukan proses pengambilan objek secara acak berulang sampai perhitungan total simpangan tidak mengalami perubahan, sehingga dapat membentuk kelompok anggota kluster yang bersifat stabil.

2.5 Database

Database adalah suatu aplikasi yang menyimpan sekumpulan data. Setiap *database* mempunyai API tertentu untuk membuat, mengakses, mengatur, mencari, menyalin data yang ada di dalamnya. *Database* yaitu kumpulan *file-file* yang berhubungan satu dengan yang lainnya, diatur sedemikian rupa sehingga dapat digunakan oleh beberapa program aplikasi *database* [22].

Database adalah himpunan kelompok data (arsip) yang saling berhubungan, yang di organisasi sedemikian rupa agar kelak dapat di dimanfaatkan kembali dengan cepat dan mudah, *Database* adalah suatu kumpulan data terhubung (*interrelated data*) yang disimpan secara bersama-sama pada suatu media, tanpa mengantap satu sama lain atau tidak perlu satu kerangkapan data (*controlled redudancy*) dengan cara tertentu sehingga mudah digunakan atau ditampilkan kembali, dapat digunakan satu atau lebih program aplikasi secara optimal, data disimpan tanpa mengalami ketergantungan pada program yang akan menggunakannya, data disimpan sedemikian rupa sehingga penambahan, pengambilan dan modifikasi dapat dilakukan dengan mudah dan terkontrol [23].

Database merupakan dasar atau tempat di mana data disimpan dengan cara yang terorganisir dan terstruktur sehingga memudahkan pengguna dalam

mengakses dan mengelola data tersebut. Penggunaan kata ini mencerminkan fungsi utama dari sebuah *database*, yaitu sebagai wadah penyimpanan data yang memungkinkan akses dan manipulasi data secara efisien dan sistematis [24].

Basis data dapat diartikan sebagai kumpulan file / table yang saling berhubungan (dalam sebuah basis data di sebuah sistem komputer), dan sekumpulan program (*DBMS/Database Management System*) yang memungkinkan beberapa *user* (pemakai), dan / atau program lain untuk mengakses dan memanipulasi *file (table)* tersebut [25].

2.6 MySQL

Basis data yang paling digemari dikalangan programmer *web*, dengan alasan bahwa program ini merupakan basis data yang sangat kuat dan cukup stabil untuk digunakan sebagai media penyimpanan data. Sebagai sebuah basis data *server* yang mampu untuk manajemen basis data dengan baik, MySQL terhitung merupakan basis data yang paling digemari dan paling banyak digunakan dibanding basis data lainnya.[26]. MySQL adalah nama *database server*. *Database server* adalah *server* yang berfungsi untuk menangani *database*. *Database* adalah suatu pengorganisasian data dengan tujuan memudahkan penyimpanan dan pengaksesan data [27].

MySQL merupakan suatu jenis database server yang sangat terkenal. MySQL termasuk jenis RDBMS (*Relational Database Manajement System*). MySQL mendukung bahasa pemrograman PH, bahasa permintaan yang terstruktur, karena pada penggunaannya SQL memiliki beberapa aturan yang telah distandarkan oleh asosiasi yang bernama ANSI. [28]. *MySQL* termasuk jenis

RDBMS (*Relational Database Manajement System*). *MySQL* mendukung bahasa pemrograman *PHP*, bahasa permintaan yang terstruktur, karena pada penggunaannya SQL memiliki beberapa aturan yang telah distandarkan oleh asosiasi yang Bernama ANSI. *MySQL* merupakan RDBMS (*Relational Database Management System*) server [28].

MySQL merupakan software yang tergolong sebagai DBMS (*Database Managemen System*) yang bersifat *Open Source*. *Open source* menyatakan bahwa *software* ini dilengkapi dengan *source code* (kode yang dipakai untuk membuat *MySQL*), selain tentu saja bentuk *excutable*-nya atau kode yang dapat dijalankan secara langsung dalam sistem operasi [29].

2.7 *PHP (Hypertext Preprocessor)*

PHP merupakan sebagai bahasa *script server-side* dalam pengembangan *web* yang disisipkan pada dokumen *HTML* [30]. *PHP (PHP: Hypertext Preprocessor)* adalah suatu bahasa pemrograman yang digunakan untuk menterjemahkan basis kode program menjadi kode mesin yang dapat dimengerti oleh komputer yang bersifat *serverside* yang ditambahkan ke *HTML*. *PHP* merupakan bahasa pemrograman untuk pembuatan *website* dinamis, yang mampu berinteraksi dengan pengunjung atau penggunanya [29].

PHP adalah bahasa pemrograman *script* yang paling banyak dipakai saat ini. *PHP* banyak dipakai untuk memrogram situs *web* dinamis, walaupun tidak tertutup kemungkinan digunakan untuk pemakaian lain. Contoh terkenal dari aplikasi *PHP* adalah forum (*PHP BB*) dan *MediaWiki* (*software* di belakang *Wikipedia*). [31]. *PHP* adalah bahasa pelengkap *HTML* yang memungkinkan

dibuatnya aplikasi dinamis yang memungkinkan adanya pengolahan data dan pemrosesan data. Semua syntax yang diberikan akan sepenuhnya dijalankan pada *server* sedangkan yang dikirimkan ke browser hanya hasilnya saja. Kemudian merupakan bahasa berbentuk script yang ditempatkan dalam server dan diproses di server. Hasilnya akan dikirimkan ke *client*, tempat pemakai menggunakan *browser* [28]

PHP adalah suatu bahasa pemrograman yang digunakan untuk menterjemahkan basis kode program menjadi kode mesin yang dapat dimengerti oleh komputer yang bersifat serverside yang ditambahkan ke HTML [29].

2.8 HTML (*HyperText Markup Language*)

HTML adalah bahasa *markup* yang digunakan untuk membuat halaman *web*. [32]. HTML merupakan salah satu format yang digunakan dalam pembuatan dokumen dan aplikasi yang berjalan di halaman *web*. Oleh karena itu agar dapat membuat program aplikasi di atas halaman *web* anda terlebih dahulu harus mengenal dan menguasai HTML [33].

HTML adalah sekumpulan simbol-simbol atau *tag-tag* yang dituliskan dalam sebuah *file* yang digunakan untuk menampilkan halaman pada *Web browser*. *Tag-tag HTML* selalu diawali dengan *<* dan diakhiri dengan *>* dimana *x tag HTML* itu seperti *b*, *i*, *u* dll [34]. HTML merupakan bahasa yang digunakan untuk mendeskripsikan struktur sebuah halaman web[26].

HTML adalah suatu bahasa pemrograman yang digunakan untuk pembuatan halaman website agar dapat menampilkan berbagai informasi baik tulisan maupun gambar pada sebuah web browser [35].

Hasil pemaparan jurnal diatas dapat disimpulkan HTML adalah bahasa *markup* standar yang digunakan untuk membuat dan mendesain halaman *web*. HTML menyediakan struktur dasar untuk konten *web* dengan menggunakan elemen-elemen (*tag*) yang mengelilingi teks, gambar, video, dan konten lainnya. HTML bekerja bersama dengan CSS dan *JavaScript* untuk menciptakan halaman *web* yang dinamis dan interaktif.

2.9 Web

Web adalah sebuah kumpulan halaman yang diawali dengan halaman muka yang berisikan informasi, iklan, serta program aplikasi. [34]. *Website* adalah sekumpulan halaman-halaman yang saling terhubung satu sama lain yang didalamnya memuat berbagai informasi yang dinamis maupun statis yang dapat diakses dan digunakan oleh pengguna. *Website* merupakan media informasi yang baik dalam penyampaian dilakukan secara digital yang dimanfaatkan dengan tujuan untuk memudahkan dalam menyampaikan informasi kepada khalayak ramai [36].

Website adalah kumpulan halaman dalam suatu domain yang memuat tentang berbagai informasi agar dapat dibaca dan dilihat oleh pengguna atau pemakai internet melalui sebuah mesin pencari atau *search engine* [37]. Sebuah situs *web* adalah sebutan bagi sekelompok halaman web, yang umumnya merupakan bagian dari suatu nama domain atau subdomain di *World Wide Web* (WWW) di Internet [26].

Website adalah fasilitas internet yang menghubungkan dokumen dalam lingkup lokal maupun jarak jauh. Dokumen pada *website* disebut dengan *web*

page dan *link* dalam *website* memungkinkan pengguna bisa berpindah dari satu *page* ke *page* lain (*hypertext*), baik diantara *page* yang disimpan dalam *server* yang sama maupun *server* di seluruh dunia [35].

Dari pemaparan diatas dapat disimpulkan *web* atau *World Wide Web* (*WWW*) adalah sistem informasi global yang memungkinkan pengguna mengakses dan berbagi berbagai jenis konten melalui internet menggunakan *browser web*. *Web* terdiri dari halaman-halaman yang terhubung satu sama lain melalui tautan (*hyperlink*) dan biasanya ditulis menggunakan bahasa markup seperti HTML, serta didukung oleh teknologi seperti CSS dan JavaScript untuk tampilan dan interaktivitasnya.

2.10 XAMPP

XAMPP adalah salah satu paket instalasi *apache*, PHP, dan MySQL secara instant yang dapat digunakan untuk membantu proses instalasi ketiga produk tersebut [38]. XAMPP adalah perangkat lunak bebas, yang mendukung banyak sistem operasi, merupakan kompilasi dari beberapa program. XAMPP merupakan tool yang menyediakan paket perangkat lunak ke dalam satu buah paket (Agustini and Kurniawan 2021).

XAMPP adalah sebuah *software* yang berfungsi untuk menjalankan *website* berbasis PHP dan menggunakan pengolah data MySQL dikomputer lokal. XAMPP berperan sebagai *server web* pada komputer lokal. XAMPP juga dapat disebut sebuah *Cpanel server virtual*, yang dapat membantu melakukan *preview* sehingga dapat dimodifikasi *website* tanpa harus *online* atau terakses dengan internet (Fp-growth 2021). XAMPP adalah salah satu paket instalasi *apache*, PHP,

dan MySQL secara instant yang dapat digunakan untuk membantu proses instalasi ketiga produk tersebut . Pengertian XAMPP sendiri adalah perangkat lunak (*free software*) bebas, yang mendukung untuk banyak sistem operasi, yang merupakan kompilasi dari beberapa program [41].

XAMPP merupakan media atau *web server localhost* yang bisa digunakan secara *offline*. Melalui XAMPP, pengguna dapat mengelola Database yang berada di localhost tanpa memerlukan akses internet sehingga jika koneksi internet terganggu dan tidak dapat mengakses *web server*. XAMPP merupakan paket PHP berbasis *open source* yang dikembangkan oleh sebuah komunitas *Open Source* [42].

Dapat disimpulkan XAMPP adalah perangkat lunak yang menyediakan paket instalasi lengkap untuk membangun dan menjalankan *server web* secara lokal.

2.11 UML (Unified Modeling Language)

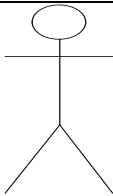
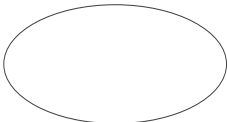
UML adalah bahasa untuk visualisasi, spesifikasi, membangun sistem perangkat lunak, serta dokumentasi, *UML* menyediakan model-model yang tepat, tidak ambigu, dan lengkap, secara khusus *UML* menspesifikasi langkah-langkah penting dalam pengembangan keputusan analisis, perancangan, serta implementasi dalam sistem perangkat lunak [43].

Berikut beberapa diagram-diagram pada *UML (Unified Modeling Language)* [43] :

a) *Use Case Diagram*

Use case pada dasarnya merupakan gambaran dari proses sistem secara keseluruhan yang melibatkan actor dalam hal pengguna, *Use case* adalah cara untuk menunjukan stake holder sistem akan berinteraksi dengan sistem mengembangkan *use case* membantu memahami persyaratan sistem secara detail”.

Tabel 2.1 Usecase Diagram

Gambar	Nama	Keterangan
	<i>Actor</i>	Peran pengguna yang memainkan ketika berinteraksi dengan <i>use case</i>
	<i>Association</i>	Permulaan / pengakhiran data
	<i>Use Case</i>	Urutan aksi-aksi yang ditampilkan sistem dan menghasilkan

b) *Class Diagram*

Class Diagram adalah diagram struktur statis yang menggambarkan struktur sistem dengan menunjukkan kelas sistem, atributnya, operasi (atau metode), dan hubungan antar kelas.

Tabel 2.2 Class Diagram

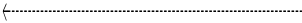
Gambar	Nama	Keterangan
	<i>Class</i>	Kelas pada struktur system

	<i>Generalization</i>	Relasi antar kelas dengan generalisasi-spesialisasi (umum-khusus)
---	-----------------------	---

c) *Sequence Diagram*

Diagram urutan menunjukkan interaksi objek yang diatur dalam urutan waktu menggambarkan objek dan kelas yang terlibat dalam skenario dan ukuran pesan yang dipertukarkan antara objek yang diperlukan untuk melaksanakan fungsi skenario”.

Tabel 2.3 *Sequence Diagram*

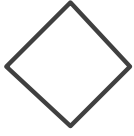
Gambar	Nama	Keterangan
	<i>Object lifeline</i>	Objek <i>entity</i> antarmuka yang saling berinteraksi
	Activation	Objek dalam keadaan aktif dan berinteraksi pesan
	<i>Message (call)</i>	Pesan antar dua objek
	<i>Message (return)</i>	Pengembalian diri pemanggilan prosedur

d) *Activity Diagram*

Activity Diagram adalah sebuah cara untuk memodelkan aliran kerja (*workflow*) dari *use case* dalam bentuk grafik, diagram ini menunjukkan

langkah-langkah di dalam aliran kerja, titik-titik keputusan didalam aliran kerja, siapa yang bertanggung jawab menyelesaikan masing-masing aktivitas, dan objek-objek yang digunakan dalam aliran kerja

Tabel 2.4 Activity Diagram

Gambar	Nama	Keterangan
	<i>Initial State</i>	Bagaimana objek dibentuk dan diawali
	<i>Line Connector</i>	Digunakan untuk menghubungkan satu simbol dengan simbol lainnya
	<i>Action State</i>	Bagaimana masing-masing kelas antarmuka saling berinteraksi yang mencerminkan eksekusi dari suatu aksi
	<i>Activity final</i>	Bagian objek dibentuk dan diakhiri
	<i>Decision</i>	Digunakan untuk menggambarkan suatu keputusan / tindakan yang harus diambil pada kondisi tertentu.
	<i>Join</i>	Asosiasi penggabungan lebih dari satu aktivitas.

2.12 Penelitian Terdahulu

Adapun penelitian terdahulu dapat dilihat pada gambar berikut :

Tabel 2.5 Penelitian Terdahulu

No	Nama Peneliti	Judul	Metode	Kesimpulan
1	Refi Yanti, Wawan Catur Hadi (2024)	Optimasi Jumlah Cluster untuk Analisis Penjualan	K-Medoids	Jumlah cluster optimal adalah 3: produk murah

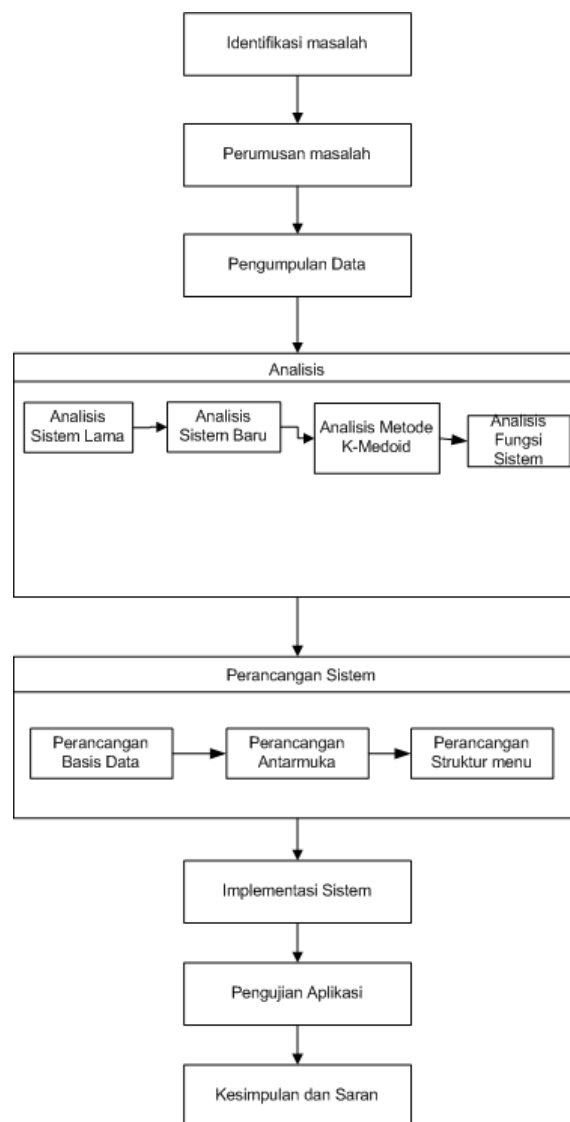
		Barang Kosmetik Menggunakan K-Medoids		& laris, murah & kurang laris, mahal & laris. Hasil bantu optimasi stok & strategi pemasaran.
2	Fitria Budi Rahmah, Wahyudin (2024)	Perbandingan Algoritma K-Means dan K-Medoids Untuk Clustering Harga Beras di Provinsi Jawa Tengah (2024)	K-Means dan K-Medoids	K-Means menghasilkan nilai DBI lebih baik dari K-Medoids, artinya performa klasterisasi K-Means lebih unggul untuk data ini.
3	Deli Agustina, Dede Muhardi (2024)	Klasterisasi Tingkat Penjualan Kedai Kopi Hallo Burjois Menggunakan Algoritma K-Medoids	Algoritma K-Medoids	Terbentuk 3 klaster menu: tinggi, sedang, rendah. Menu dengan penjualan rendah perlu evaluasi. Nilai DBI = 0.95 menunjukkan klasterisasi baik.
4	Wahyuni (2021)	Pengelompokan Data Penjualan Mie Berdasarkan Bulan Dengan Menggunakan Algoritma K-Medoids	K-Medoids,	Klasterisasi per bulan menghasilkan pola musiman. Penjualan tinggi pada bulan Januari dan Desember. Kelompok rendah mendominasi (75%).
5	Dicky Syahputra, Feri Sutanto (2023)	Analisis Penjualan Menggunakan Algoritma K-Medoids Untuk Mengoptimalkan Penjualan Barang	K-Medoids	Terbentuk 2 klaster: barang laris & tidak laris. Hasil klasterisasi bantu

				manajemen dalam penyediaan stok & strategi penjualan.
--	--	--	--	---

BAB 3

METODOLOGI PENELITIAN

Tahapan penelitian yang akan dilakukan dan penyelesaian masalah terhadap penerapan metode K-Medoids untuk pengelompokan penjualan barang pada Habib *Stationery*. Adapun tahapan metodologi yang dilakukan selama penelitian dapat dilihat pada Gambar 3.1 yang mana merupakan proses yang dimulai dari studi literatur hingga diperoleh kesimpulan.



Gambar 3. 1 Tahapan Metodologi Penelitian

3.1 Identifikasi Masalah

Adapun identifikasi masalah terkait pengelompokan penjualan barang dengan metode K- Medoids adalah sebagai berikut :

1. Banyaknya jenis barang yang dijual di Habib stationery menyebabkan kesulitan dalam menganalisis performa penjualan setiap produk secara manual, terutama untuk menentukan barang yang laris, sedang, dan tidak laku.
2. Pengelompokan barang berdasarkan penjualan selama ini masih dilakukan secara subjektif dan tidak berbasis pada data historis yang akurat, sehingga rawan kesalahan dalam pengambilan keputusan manajemen.
3. Kesulitan dalam menentukan prioritas stok dan strategi pemasaran akibat kurangnya informasi yang terstruktur tentang kelompok barang berdasarkan tingkat penjualannya.

3.2 Perumusan Masalah

Berdasarkan hasil dari tahapan pengamatan pendahuluan sebelumnya, maka tahapan selanjutnya adalah tahapan perumusan masalah, rumusan masalah dalam penelitian ini didapatkan dari penelitian terkait dari data pengamatan pendahuluan sebelumnya, solusi yang didapatkan pada tahapan perumusan masalah ini yang akan menjadi judul penelitian ini yaitu “Penerapan Algoritma K- Medoids Untuk Pengelompokkan Penjualan Barang Di Habib stationery Berbasis Web”.

3.3 Pengumpulan Data

Pengumpulan data adalah tahapan-tahapan yang bertujuan dalam memperoleh data-data informasi yang berhubungan dengan penelitian ini, pada tahapan pengumpulan data ini juga berguna untuk mengumpulkan semua kebutuhan data yang akan diproses nantinya menggunakan metode “Metode K-Medoids”. Pengumpulan data tersebut dilakukan dengan beberapa cara, antara lain :

1. Observasi

Peneliti melakukan pengamatan langsung di Habib stationery untuk melihat proses transaksi penjualan, pencatatan barang, serta bagaimana pihak toko melakukan analisis penjualan selama ini. Observasi ini bertujuan untuk memahami alur kegiatan operasional toko serta menemukan permasalahan yang relevan dengan pengelompokan data penjualan barang.

2. Wawancara

Wawancara dilakukan secara langsung dengan pihak pengelola toko, seperti kepala toko dan staf bagian penjualan. Wawancara ini bertujuan untuk menggali informasi mengenai kebutuhan sistem analisis penjualan, cara pengambilan keputusan terhadap barang yang akan di-restock atau dipromosikan, serta tantangan yang dihadapi dalam mengelola berbagai jenis produk.

3. Studi Kepustakaan

Pengumpulan data juga dilakukan melalui penelusuran berbagai literatur yang relevan, seperti jurnal ilmiah, artikel, skripsi terdahulu, buku teks, dan dokumen yang membahas tentang *data mining*, algoritma K-Medoids, dan sistem pengelompokan data penjualan. Studi ini bertujuan untuk memperkuat dasar teori serta mendukung validitas pendekatan yang digunakan dalam penelitian ini.

3.4 Analisis

Tahapan selanjutnya adalah melakukan Analisis metode sistem dari penelitian skripsi ini, adapun tahapan Analisis dalam penelitian Skripsi ini adalah sebagai berikut :

3.4.1 Analisis Sistem Lama

Analisis sistem lama merupakan langkah penting dalam proses pengembangan atau peningkatan aplikasi, fungsi utama dari analisis sistem lama adalah untuk memahami kondisi saat ini dari sistem yang ada sebelum melakukan perubahan atau pengembangan baru, dengan melakukan analisis sistem lama secara menyeluruh, organisasi dapat membuat keputusan yang lebih baik mengenai bagaimana cara mengembangkan atau memperbarui sistem yang ada, hal ini juga membantu dalam meminimalkan risiko dan memastikan bahwa sistem baru atau yang ditingkatkan dapat memenuhi kebutuhan bisnis dan pengguna dengan lebih efektif.

3.4.2 Analisis Sistem Baru

Analisis sistem baru adalah langkah penting dalam pengembangan atau implementasi aplikasi baru, fungsi utama dari analisis sistem baru adalah untuk memastikan bahwa sistem yang akan dikembangkan atau diimplementasikan sesuai dengan kebutuhan dan tujuan organisasi, dengan melakukan analisis sistem baru yang komprehensif, organisasi dapat memastikan bahwa sistem yang dikembangkan atau diimplementasikan akan memenuhi kebutuhan bisnis, meningkatkan efisiensi operasional, dan memberikan nilai tambah yang signifikan bagi organisasi dan pengguna.

3.4.3 Analisis Metode K-Medoids

Tahapan ini adalah proses dimana langkah-langkah pengolahan data menggunakan metode K-Medoids dijalankan, K-Medoids adalah salah satu algoritma dalam *data mining* yang digunakan untuk teknik *clustering* (pengelompokan data). Algoritma ini mirip dengan K-Means, namun lebih robust terhadap *outlier* dan data ekstrem. Berbeda dengan K-Means yang menggunakan rata-rata (*centroid*) sebagai pusat cluster, K-Medoids menggunakan salah satu objek nyata dari dataset sebagai pusat cluster (*medoid*).

3.4.4 Analisis Fungsi Sistem

Setelah melakukan tahapan analisis terhadap metode K-Medoid maka selanjutnya adalah Analisis fungsional sistem yang akan dibangun, adapun tahapan-tahapan analisis fungsional yaitu dalam pembuatan *usecase diagram*, *Activity Diagram*, *Sequence Diagram* dan *Class Diagram*.

Kelebihan algoritma K-Medoids adalah memberikan hasil *cluster* yang lebih akurat dalam konteks penjualan yang mungkin memiliki *outlier* (produk yang sangat laris atau sangat tidak laku).

3.5 Perancangan Sistem

Setelah tahapan Analisis selesai dilakukan, maka tahapan selanjutnya adalah perancangan system, tahapan perancangan sistem terdiri dari :

3.5.1 Perancangan Basis Data

Setelah dilakukannya Analisis sistem yang akan dibuat, maka tahap berikutnya ialah Analisis dan perancangan basis data yang kita lakukan untuk melengkapi komponen dalam pembuatan sistem, dalam perancangan basis data menggunakan *Class Diagram*.

3.5.2 Perancangan Antar Muka (*Interface*)

Dalam mempermudah komunikasi antara sistem dengan pengguna, maka perlu di rancang antar muka (*interface*), dalam perancangan *interface* hal terpenting yang harus dilakukan ialah bagaimana menciptakan tampilan yang baik dan mudah dimengerti oleh pengguna.

3.5.3 Perancangan Struktur Menu

Rancangan struktur menu ini kita perlu untuk memberikan gambaran terhadap menu-menu atau fitur pada sistem yang akan dibuat.

3.6 Implementasi Sistem

Beberapa komponen pendukung yang memiliki peran yang sangat penting dalam implementasi sistem diantaranya adalah perangkat keras (*hardware*) dan

perangkat lunak (*software*) adapun spesifikasi dari perangkat keras (*hardware*) dan perangkat lunak (*software*) yang di gunakan sebagai berikut :

1. Perangkat keras (*hardware*), antara lain :

Prosesor : Intel Core i3
 Memory (RAM) : 4 GB RAM
 System type : 64-bit Operating System
 Harddisk : 500 GB

2. Perangkat Lunak (*Software*), antara lain :

Sistem operasi : Windows 10

3.7 Pengujian Aplikasi

Pengujian sistem ini dilakukan dengan cara menggunakan *Black Box* dalam Pengujian *Black Box* ini berfokus pada perangkat lunak untuk mendapatkan hasil yang baik, apabila terjadi *error* atau tidak sesuai dengan tujuan yang diharapkan maka dilakukan pengAnalisisa sistem kembali hingga tidak ditemukan *error*.

3.7.1 Black Box Testing

Black box testing merupakan pengujian kualitas perangkat lunak yang berfokus pada fungsionalitas perangkat lunak, pengujian black box testing bertujuan untuk menemukan fungsi yang tidak benar kesalahan antarmuka kesalahan pada struktur data, kesalahan performansi, kesalahan inisialisasi dan terminasi, dalam pengujian *black box testing* digunakan alat untuk pengumpulan data yang disebut dengan *user acceptance test*, dokumen ini

terdiri deskripsi indikator dari prosedur-prosedur pengujian fungsionalitas dari perangkat lunak

3.8 Kesimpulan dan Saran

Tahapan terakhir adalah menarik kesimpulan dari hasil penelitian yang didapatkan dalam pengelompokkan penjualan barang dengan metode K-Medoids, pada tahapan ini juga berisikan saran peneliti bagi pembaca untuk melakukan pengembangan terhadap penelitian ini kedepannya.