



**PENGELOMPOKAN RESELLER POTENSIAL MENGGUNAKAN
ALGORITMA CLUSTERING K-MEANS PADA TOKO PREMIUM
SHOPPING**

SKRIPSI

ASTRI NUR KHUSNA

18670064

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS PGRI SEMARANG**

2022



**PENGELOMPOKAN RESELLER POTENSIAL BERDASARKAN DATA
PENJUALAN MENGGUNAKAN ALGORITMA CLUSTERING K-MEANS
PADA TOKO PREMIUM SHOPPING**

SKRIPSI

**Diajukan kepada Fakultas Teknik dan Informatika
Universitas PGRI Semarang untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Teknik**

ASTRI NUR KHUSNA

18670064

PROGRAM STUDI INFORMATIKA

UNIVERSITAS PGRI SEMARANG

2022

SKRIPSI

**PENGELOMPOKAN RESELLER POTENSIAL BERDASARKAN DATA
PENJUALAN MENGGUNAKAN ALGORITMA CLUSTERING K-MEANS
PADA TOKO PREMIUM SHOPPING**

Disusun dan diajukan oleh

ASTRI NUR KHUSNA

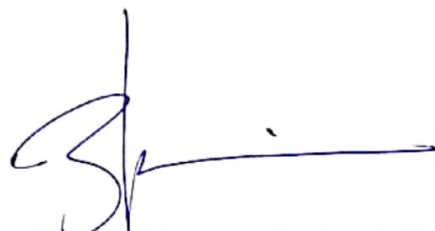
18670064

telah disetujui oleh pembimbing untuk dilanjutkan ujian sidang

dihadapan Dewan Penguji

Semarang, 16 Agustus 2022

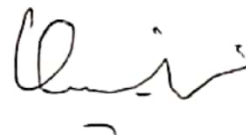
Pembimbing I,



Aris Tri Jaka Harjanta, M.Kom.

NIDN. 0619048202

Pembimbing II,



Khoiriya Latifah, M.Kom.

NIDN. 0617077801

SKRIPSI

PENGELOMPOKAN RESELLER POTENSIAL BERDASARKAN DATA
PENJUALAN MENGGUNAKAN ALGORITMA CLUSTERING K-MEANS
PADA TOKO PREMI M SHOPPING

Disusun dan diajukan oleh

ASTRI NUR KHUSNA

18670064

telah dipertahankan di depan Dewan Penguji
pada tanggal 19 Agustus 2022
dan dinyatakan telah memenuhi syarat

Dewan Penguji



Ketua,


Dr. Slamet Supriyadi, M.Env

NIDN. 0028125901

Sekretaris,


Bambang Agus Herlambang, M.Kom

NIDN. 0601088201

Penguji I,


Bambang Agus Herlambang, M.Kom.

NIDN. 0601088201

Penguji II,


Aris Tri Jaka Harjanta, M.Kom.

NIDN. 0619048202

Penguji III,


Khoiriya Latifah, M.Kom.

NIDN. 0617077801

MOTO DAN PERSEMBAHAN

Moto :

Dimanapun berada jadilah manusia yang berguna bagi banyak orang dan tetaplah menjadi orang baik tanpa mengharapkan kebaikan dari orang lain. Berpikir positif dan bekerja keras karena perjalanan impian belum berakhir.

Persembahan :

Kupersembahkan skripsi ini untuk :

1. Ayah Ibuku tercinta
2. Ibu Bapak Dosen Informatika
Universitas PGRI Semarang
yang selalu memberikan
bimbingannya
3. Almamaterku Universitas
PGRI Semarang

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini :

Nama : Astri Nur Khusna

NPM : 18670064

Progdi : Informatika

Fakultas : Teknik dan Informatika

Menyatakan dengan sebenarnya bahwa skripsi yang saya buat ini benar-benar merupakan hasil karya saya sendiri, bukan plagiarisme. Apabila pada kemudian hari skripsi ini terbukti hasil plagiarisme, saya bersedia menerima sanksi atas perbuatan tersebut.

Semarang, Agustus 2022

Yang membuat pernyataan

Astri Nur Khusna

NPM 18670064

ABSTRAK

Toko Premium Shopping merupakan tempat penjualan suplemen *fitness* yang menyediakan berbagai jenis suplemen. Toko ini memiliki banyak pelanggan dan *reseller* yang tersebar di berbagai kota dan daerah. Dari banyaknya *reseller* yang ada membuat data penjualan *reseller* semakin lama semakin bertambah, disamping itu perusahaan ingin terus mengembangkan dan meningkatkan penjualan produk. Dari permasalahan yang ada, diperlukan sistem yang dapat mengelompokkan *reseller* berpotensi berdasarkan data penjualan yang dilakukan di Toko Premium Shopping. Adanya sistem ini membantu perusahaan dalam menargetkan kota mana saja yang berpotensi dan berpeluang untuk menambah *reseller* baru sebagai upaya untuk meningkatkan daya saing dengan perusahaan lainnya. Sistem ini menggunakan algoritma K-Means untuk menentukan dan melakukan data *clustering reseller* potensial dari Toko Premium Shopping. Hasil *clustering* berupa diagram yang menunjukkan pengelompokkan reseller berdasarkan penjualan sesuai dengan cluster, yakni Cluster 0 (Reseller kurang potensial), Cluster 1 (Reseller cukup potensial), Cluster 2 (Reseller sangat potensial). Sehingga hasil *clustering* dapat digunakan sebagai acuan untuk memperluas target pasar dari Toko Premium Shopping.

Kata Kunci : K-Means, Data Mining, *Clustering*, *reseller*, penjualan.

PRAKATA

Dengan memanjatkan puji syukur kehadirat Tuhan Yang Maha Esa, karena berkat rahmat-Nya, penulis dapat menyusun dan menyelesaikan skripsi dengan baik. Skripsi dengan judul “Pengelompokan Reseller Potensial Berdasarkan Data Penjualan Menggunakan Algoritma Clustering K-Means Pada Toko Premium Shopping” dapat dilaksanakan dengan baik.

Penyusunan skripsi ini tidak lepas dari berbagai hambatan dan rintangan. Namun berkat bimbingan, nasehat, bantuan, dan dorongan dari berbagai pihak, khususnya pembimbing segala hambatan dan permasalahan dapat teratasi. Atas tersusunnya proposal skripsi ini, penulis mengucapkan terima kasih yang sebesar-besarnya kepada :

1. Allah SWT yang telah memberikan limpahan rahmat serta selalu melindungi penulis sehingga dapat menyelesaikan skripsi ini.
2. Ibu dan Ayah atas segala doa dan dukungan moral maupun materi.
3. Ibu Drs. Sri Suciati, M. Hum selaku Rektor Universitas PGRI Semarang telah memberikan kesempatan kepada penulis untuk menuntut ilmu di Universitas PGRI Semarang.
4. Bapak Dr. Slamet Supriyadi, M.Env. St. selaku Dekan Fakultas Teknik dan Informatika Universitas PGRI Semarang yang telah memberikan ijin kepada penulis untuk melakukan penelitian.
5. Bapak Bambang Agus Herlambang, M.Kom. selaku Ketua Program Studi Informatika Fakultas Teknik dan Informatika Universitas PGRI Semarang.
6. Bapak Aris Trijaka Harjanta, M.Kom. selaku Pembimbing I Program Studi Informatika Fakultas Teknik dan Informatika yang telah meluangkan waktu untuk memberikan bimbingan selama proses pembuatan skripsi ini.

7. Ibu Khoiriya Latifah, M.Kom. selaku Pembimbing II Program Studi Informatika Fakultas Teknik dan Informatika yang telah membimbing penulis dengan dedikasi yang tinggi.
8. Seluruh Dosen Pengajar, Staff dan Karyawan Fakultas Teknik dan Informatika Universitas PGRI Semarang.
9. Seluruh Rekan Kerja Toko Premium Shopping yang telah memberikan semangat kepada penulis.
10. Teruntuk teman-teman penulis yang telah menjadi *support system* dan *mood booster* ketika menghadapi masalah.
11. Semua pihak yang tidak dapat disebutkan satu per satu yang terlibat dalam proses penyusunan skripsi ini.
12. *Last but not least, i wanna thank me, i wanna thank me for believing on me, i wanna thank me for doing all this hard work, i wanna thank me for having no days off, i wanna thank me for never quitting.*

Semarang, Agustus 2022

Penulis

DAFTAR ISI

HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
MOTO DAN PERSEMBAHAN	v
PERNYATAAN KEASLIAN TULISAN.....	vi
ABSTRAK	vii
PRAKATA	viii
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
DAFTAR NOTASI.....	xiv
DAFTAR LAMPIRAN.....	xv
BAB I PENDAHULUAN	1
A. Latar Belakang Masalah	1
B. Pembatasan Masalah	2
C. Rumusan Masalah	2
D. Tujuan Penelitian	2
E. Manfaat Penelitian	3
BAB II KAJIAN PUSTAKA	4
A. Tinjauan Pustaka	4
B. Landasan Teori.....	7
BAB III METODE PENELITIAN	16
A. Pendekatan Penelitian.....	16
B. Lokasi Penelitian	23
C. Populasi dan Sampel	23
D. Variabel Penelitian	24
E. Pengumpulan Data	25
F. Kerangka Berpikir	26
BAB IV HASIL DAN PEMBAHASAN	28
A. Pemahaman Bisnis (Business Understanding)	28

B. Tahap Pemahaman Data (Data Understanding).....	28
C. Tahap Persiapan Data (Data preparation).....	34
D. Tahap Pemodelan (Modeling)	36
E. Evaluation	43
F. Deployment.....	46
BAB V SIMPULAN DAN SARAN.....	48
A. Kesimpulan	48
B. Saran.....	48
DAFTAR PUSTAKA	49
LAMPIRAN	50

DAFTAR TABEL

Tabel 2. 1 Tabel Referensi Penelitian	4
Tabel 3. 1 Tabel Contoh Kelompok Data	18
Tabel 3. 2 Tabel Data Ke-1 dan Ke-2	19
Tabel 3. 3 Tabel Hasil Pengelompokan Clustering	19
Tabel 3. 4 Tabel Hasil Perhitungan Data	20
Tabel 3. 5 Tabel Hasil Centroid Baru	20
Tabel 3. 6 Tabel Daftar Produk	24
Tabel 4. 1 Tabel Penjualan Reseller	33
Tabel 4. 2 Kriteria cluster.....	38

DAFTAR GAMBAR

Gambar 2. 1 Tampilan Awal Google Colab.....	13
Gambar 2. 2 Membuat Notebook baru	13
Gambar 2. 3 Tampilan Notebook baru	14
Gambar 2. 4 Menyimpan file Notebook	14
Gambar 2. 5 Mendownload file Notebook.....	14
Gambar 2. 6 Tampilan Streamlite.....	15
Gambar 3. 1 Flowchart Clustering K-Means	17
Gambar 3. 2 Model CRISP-DM.....	21
Gambar 3. 3 Lokasi Penelitian	23
Gambar 3. 4 Diagram Jumlah Reseller	25
Gambar 3. 5 Kerangka Berpikir	27
Gambar 4. 1 Periksa Data Null.....	35
Gambar 4. 2 Periksa Data sama dan hapus kolom.....	36
Gambar 4. 3 Elbow Method	37
Gambar 4. 4 Pusat Cluster.....	39
Gambar 4. 5 Diagram Hasil Clustering.....	41
Gambar 4. 6 Diagram masing-masing cluster	42
Gambar 4. 7 Model Data Disimpan	42
Gambar 4. 8 Evaluasi model	45
Gambar 4. 9 Save Model dan Akurasi	45
Gambar 4. 10 Tab Penjualan	46
Gambar 4. 11 Tab Daftar Produk	46
Gambar 4. 12 Tab Data Mining	47
Gambar 4. 13 Diagram Hasil Clustering.....	47

DAFTAR NOTASI

Rumus 3.1 Euclidean Distance	18
------------------------------------	----

DAFTAR LAMPIRAN

Lampiran 1 Tabel Hasil Clustering.....	50
Lampiran 2 Tabel Hasil Cluster 0.....	54
Lampiran 3 Tabel Hasil Cluster 1.....	57
Lampiran 4 Tabel Hasil Cluster 2.....	58

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Berkembangnya teknologi serta meningkatnya penggunaan internet oleh masyarakat ditandai dengan melonjaknya pelanggan maupun pelaku bisnis *online*. Selain itu mudahnya mengakses internet dan tersedianya berbagai *platform* penjualan memudahkan promosi dan kegiatan bertransaksi. Ini merupakan kesempatan penting bagi pelaku bisnis yakni bagaimana cara memperoleh keuntungan secara aman dan mudah. Beberapa upaya yang dilakukan perusahaan untuk meningkatkan penjualan dan bersaing dengan perusahaan lainnya yakni dengan cara menambah *outlet* atau *reseller* baru diberbagai kota dan daerah. *Reseller* itu sendiri merupakan seseorang yang membeli barang jadi dari sebuah toko atau tempat usaha kemudian dijual kembali kepada pembeli yang lain [1]. *Reseller* juga membantu perusahaan untuk memperluas pangsa pasar. Selain itu perusahaan juga menghemat biaya operasional seperti biaya promosi dan biaya transportasi penjualan. Dengan jumlah *reseller* yang tidak sedikit dan pelanggan yang bertransaksi setiap harinya tidak menutup kemungkinan menghasilkan data penjualan yang sangat banyak. Data penjualan itu sendiri terdiri dari berbagai informasi seperti nama cabang atau *reseller*, kemudian lokasi *reseller*, jenis produk, dan jumlah penjualan. Dari banyaknya data penjualan tersebut nantinya dapat digunakan perusahaan sebagai acuan stok produk, segmentasi pasar, dan prediksi laba dari perusahaan.

Salah satu bidang usaha yang banyak membuka *reseller* adalah bidang usaha kesehatan. Dimana keperluan obat-obatan dan suplemen yang dikonsumsi hampir setiap hari, sehingga permintaan dari pelanggan tidak sedikit. Toko Premium Shopping merupakan salah satu tempat usaha dibidang kesehatan yakni suplemen *fitness* yang menyediakan berbagai

jenis suplemen dan telah memiliki banyak pelanggan baik didalam kota Semarang maupun diluar kota Semarang. Toko Premium Shopping beralamat di jalan Parang Kusumo Raya no 36 Tlogosari Kulon Kecamatan Pedurungan Kota Semarang 50196. Toko Premium Shopping memiliki beberapa cabang *reseller* yang tersebar diberbagai kota. Banyaknya *reseller* tersebut menjadikan perusahaan memiliki data penjualan dari masing-masing *reseller*. Namun dari data penjualan tersebut hanya disimpan sebagai arsip saja tidak diolah dan dimanfaatkan lagi untuk memprediksi pasar dari berbagai kota melalui data tersebut. Untuk itu perusahaan memerlukan sistem yang dapat memudahkan dalam melakukan *clustering reseller* potensial. Proses *clustering* data menggunakan sistem diperlukan metode data mining yakni K-Means *clustering*. K-Means merupakan salah satu metode *clustering* yang berfungsi untuk mengelompokkan data sesuai dengan kriteria yang ditentukan. Sedangkan *clustering* itu sendiri merupakan pengelompokkan suatu grup yang memiliki kesamaan tingkat tinggi [2]. *Clustering* berguna untuk memprediksi dan menganalisa segmentasi pasar, marketing, dan pemetaan zonasi wilayah. *Clustering* berbeda dengan klasifikasi karena *clustering* mengelompokkan data berdasarkan kemiripan objek data dan meminimalkan kemiripan terhadap kelas lain [3]. Dengan adanya sistem ini diharapkan dapat membantu perusahaan dalam memaksimalkan data penjualan *reseller* yang dimiliki perusahaan untuk target pemasaran berdasarkan hasil *clustering reseller* potensial. Sehingga promosi perusahaan lebih terarah, tepat sasaran dan menghemat biaya promosi.

Berdasarkan uraian diatas, maka Toko Premium Shopping memerlukan sebuah model pengelompokan untuk Pengelompokan Reseller Potensial Menggunakan Algoritma Clustering K-Means Pada Toko Premium Shopping.

B. Pembatasan Masalah

Adapun batasan masalah dari penelitian ini adalah :

1. Pemodelan data ini menggunakan algoritma K-Means dengan metode data mining untuk mengelompokkan *reseller* potensial berdasarkan data penjualan.
2. Data penelitian yang akan digunakan diperoleh dari Toko Premium Shopping.
3. Data yang akan digunakan adalah data penjualan *reseller* Toko Premium Shopping.
4. Pemodelan data ini menggunakan bahasa pemrograman Python.

C. Rumusan Masalah

Berdasarkan latar belakang penelitian maka dapat dirumuskan permasalahannya sebagai berikut :

1. Bagaimana mengolah data penjualan *reseller* agar dapat digunakan lebih maksimal?
2. Bagaimana memodelkan data penjualan pada Toko Premium Shopping?
3. Bagaimana agar promosi pada Toko Premium Shopping menjadi lebih terarah?

D. Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dijabarkan sebelumnya tujuan dari penelitian pada Toko Premium Shopping adalah :

1. Mengolah data penjualan *reseller* untuk meningkatkan strategi promosi agar promosi lebih baik dan efisien.
2. Menerapkan metode K-Means *Clustering* untuk memodelkan data penjualan pada Toko Premium Shopping.

E. Manfaat Penelitian

1. Bagi Mahasiswa

- a. Penelitian ini diharapkan dapat melatih dan mengukur daya tangkap penulis dalam mempraktekkan ilmu mengenai sistem informasi yang diperoleh di bangku kuliah dan menjadi referensi penelitian di masa yang akan datang.
- b. Penelitian ini sebagai dasar penulis untuk mewujudkan implementasi di dunia kerja nantinya dan melakukan penerapan ke dalam lingkungan hidup yang membutuhkan.
- c. Dalam pelaksanaan penelitian diharapkan penulis mendapatkan wawasan yang lebih mendalam mengenai ilmu yang dibahas dalam penelitian ini.

2. Bagi Akademik

Penelitian ini berguna untuk melakukan evaluasi dan tolak ukur keberhasilan proses belajar mengajar bagi pihak akademik serta sebagai referensi bagi mahasiswa dalam penelitian lebih lanjut yang berkaitan dengan studi yang dibahas dalam penelitian ini.

3. Bagi Toko Premium Shopping

Dari penelitian ini diharapkan dapat mempermudah Toko Premium Shopping dalam mengelola dan mengelompokkan data *reseller* yang berpotensi berdasarkan data penjualan dan meningkatkan daya saing dengan toko lainnya.

BAB II

KAJIAN PUSTAKA

A. Tinjauan Pustaka

Sebuah penelitian dikatakan baik apabila terdapat tinjauan pustaka sebagai sarana atau referensi dari penelitian yang sedang dikerjakan. Peneliti menggunakan penelitian terdahulu yang pernah dilakukan oleh peneliti lainnya. Berikut penelitian terdahulu yang digunakan peneliti sebagai referensi :

Tabel 2. 1 Tabel Referensi Penelitian

No.	Peneliti	Tahun	Judul	Metode	Hasil
1.	Deni Triyansyah	2018	Analisis Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing.	K-Means Clustering	Dari hasil analisis menggunakan algoritma K-Means Clustering menghasilkan 3 tipe sepatu yang diminati konsumen, serta menghasilkan 1 reseller yang potensial dalam penjualan.
2.	Indra Maulana YK	2021	Aplikasi Clustering Pengembangan Reseller Order Berdasarkan Kota	K-Means Clustering	Hasil penelitian ini memudahkan perusahaan dalam menyajikan laporan secara cepat dan akurat. Sehingga hasil dari clustering data dapat memudahkan dalam meningkatkan strategi penjualan produk.
3.	Event	2021	Analisis Clustering Dengan K-Means Untuk Pengelompokan Penjualan Produk Pada Hotel Newton	K-Means Clustering	Dari hasil penelitian memberikan manfaat untuk perusahaan dalam pengolahan data penjualan yang ada menjadi acuan persediaan barang. agar terhindar dari kekurangan ,maupun

No.	Peneliti	Tahun	Judul	Metode	Hasil
					kelebihan barang persediaan.
4.	Yulia Darmi	2016	Penerapan Metode Clustering K-Means Dalam Pengelompokan Penjualan Produk	K-Means Clustering	Hasil penelitian ini mempermudah pihak perusahaan dalam menemukan informasi dari tumpukan-tumpukan data penjualan agar lebih efektif dan menghemat waktu untuk menggali informasi yang dibutuhkan.
5.	Asrul Sani	2018	Penerapan Metode K-Means Clustering Pada Perusahaan	K-Means Clustering	Penelitian ini menghasilkan informasi dari pengelompokan produk untuk meningkatkan penjualan dan menentukan langkah peningkatan stok pada produk secara tepat.
6.	Novita Lestari Anggreini	2019	Teknik Clustering Dengan Algoritma K-Medoids Untuk Menangani Strategi Promosi Di Politeknik TEDC Bandung	K-Medoids Clustering	Pada penelitian ini menggunakan algoritma K-Medoids untuk pengelompokan data calon mahasiswa dan menghasilkan chart bar yang menunjukkan daerah terbanyak dari pendaftar yang dapat digunakan untuk informasi kegiatan promosi.
7.	Emi Ariska	2018	Implementasi Agglomerative Hierarchical Clustering Pada Data Produksi Dan Data Penjualan Perusahaan	Agglomerative Hierarchical Clustering	Penerapan Agglomerative Hierarchical Clustering menggunakan data produksi dan data penjualan yang menghasilkan dendrogram juga informasi produk terkelompok yang berguna untuk strategi

No.	Peneliti	Tahun	Judul	Metode	Hasil
					pemasaran.
8.	Cakra Ramadhana,dkk	2013	Data Mining Dengan Algoritma Fuzzy C-Means Clustering Dalam Kasus Penjualan Di PT. Sepatu Bata	Fuzzy C-Means Clustering	Dari penelitian ini menghasilkan rancangan aplikasi berbasis Visual Basic untuk mengelola informasi dari data penjualan menggunakan metode Fuzzy C-Means Clustering dengan hasil pengujian 91,2% menyatakan aplikasi berfungsi dengan baik.
9.	Ririn Patmawati Sari	2016	Penerapan Data Mining Pada Data Penjualan Sepatu Untuk Membentuk Segmentasi Distributor Di CV. XYZ Menggunakan Metode Clustering	Agglomerative Hierarchical Clustering	Hasil dari penelitian ini berupa aplikasi pembentukan kelompok distributor sesuai dengan kemiripan penjualan produk yang dapat digunakan untuk acuan pendistribusian barang dan strategi bisnis yang tepat.
10.	Dewi Astria	2017	Penerapan algoritma Fuzzy C-Means untuk Clustering pelanggan pada CV. Mataram Jaya Bawen	Fuzzy C-Means Clustering	Pada penelitian ini menghasilkan data pelanggan yang dibagi ke dalam 4 cluster pelanggan yaitu Golden, Silver, Bronze, dan Iron dengan variabel yang dijadikan acuan adalah tanggal pembelian akhir.

B. Landasan Teori

1. Data Mining

Data Mining adalah serangkaian proses untuk menggali nilai tambah dari suatu data berupa pengetahuan tersembunyi yang tidak diketahui secara normal [4]. Dalam prosesnya data mining menggunakan teknik statistik, matematika, kecerdasan buatan dan *machine learning* yang berfungsi untuk melakukan ekstraksi dan identifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar. Sehingga tujuan utama dari data mining adalah menemukan dan menggali pengetahuan dari data atau informasi yang ada [5].

Data mining tidak hanya melakukan *mining* terhadap data transaksi saja. Penelitian dibidang data mining saat ini sudah merambah ke sistem database lanjut seperti *object oriented database*, *image/spatial database*, *time-series data/ temporal database*, *teks* (dikenal dengan nama text mining), *web* (dikenal dengan nama web mining) dan *multimedia database*. Sedangkan fungsi dari data mining itu sendiri yakni digunakan untuk menemukan pola yang tersembunyi dan memprediksi tren masa depan yang tidak muncul dalam analisis kueri sederhana sehingga memiliki bagian penting dalam hal menemukan pengetahuan dan membuat keputusan.

Dari beberapa definisi data mining merupakan sebuah pola yang sulit. Dimana pola yang ditemukan harus penuh arti dan memberikan keuntungan. Berikut merupakan karakteristik dari data mining :

- a. Data mining berhubungan dengan penemuan sesuatu yang tersembunyi dan pola data tertentu yang tidak diketahui sebelumnya.

- b. Data mining biasa menggunakan data yang sangat besar. Biasanya data yang besar digunakan untuk membuat hasil lebih dipercaya.
- c. Data mining berguna untuk membuat keputusan krisis juga dapat digunakan untuk pengambilan keputusan di masa depan berdasarkan informasi yang diperoleh dari masa lalu [6].

2. Analisis *Clustering*

Clustering merupakan pengelompokan dalam suatu grup yang memiliki tingkat kesamaan/kemiripan yang tinggi dan sebaliknya memiliki perbedaan yang tinggi (kemiripan yang rendah) terhadap kelompok yang berbeda [2]. *Clustering* juga dikenal sebagai data segmentasi karena *clustering* mempartisi banyak data set ke dalam banyak group berdasarkan kesamaannya [7]. Dengan menggunakan *clustering*, kita dapat mengklasifikasi daerah yang padat menemukan pola-pola distribusi data menemukan keterkaitan antara atribut data. Beberapa kebutuhan *clustering* dalam data mining meliputi skalabilitas, kemampuan menangani dimensionalitas yang tinggi, menangani data yang mempunyai *noise*, dan dapat diterjemahkan dengan mudah.

Sedangkan analisis cluster adalah proses pengelompokkan satu set benda-benda fisik atau abstrak ke dalam kelas objek yang sama. Tujuannya adalah untuk menghasilkan pengelompokan objek yang mirip satu sama lain dalam kelompok-kelompok. Semakin besar kemiripan objek dalam suatu cluster dan semakin besar perbedaan tiap cluster maka kualitas analisis cluster semakin baik [6].

3. Algoritma K-Means

K-Means merupakan salah satu algoritma *clustering* yang digunakan untuk membagi data menjadi beberapa kelompok dengan sistem partisi. Algoritma ini menerima masukan berupa data tanpa label kelas. Pada algoritma K-means, komputer mengelompokkan sendiri data-data yang menjadi masukannya tanpa mengetahui terlebih dahulu target kelasnya. Masukan yang diterima adalah data atau objek dan k buah kelompok (cluster) yang diinginkan. Algoritma ini mengelompokkan data atau objek kedalam k buah kelompok tersebut.

Algoritma K-Means sangat terkenal karena kemudahan dan kemampuannya untuk mengkluster data yang besar dan outlier dengan sangat cepat. Dalam algoritma K-Means setiap data harus termasuk ke kluster tertentu dan bisa dimungkinkan bagi setiap data yang termasuk kluster tertentu pada suatu tahapan proses, pada tahapan berikutnya berpindah ke kluster lainnya [8].

Pada dasarnya algoritma K-Means hanya mengambil sebagian dari banyaknya komponen yang didapatkan untuk kemudian dijadikan pusat cluster awal, pada penentuan pusat cluster ini dipilih secara acak dari populasi data. Kemudian algoritma K-Means akan menguji masing-masing dari setiap komponen dalam populasi data tersebut dan menandai komponen tersebut ke dalam salah satu pusat cluster yang telah didefinisikan sebelumnya tergantung dari jarak minimum antar komponen tiap-tiap pusat cluster. Selanjutnya posisi pusat cluster akan dihitung kembali hingga semua komponen data digolongkan ke dalam tiap-tiap cluster dan terakhir akan terbentuk cluster baru [9].

4. Metode Data Mining

Data mining merupakan bidang ilmu yang digunakan untuk menangani masalah pengambilan informasi dari database yang besar dengan menggabungkan teknik statistik pembelajaran mesin, visualisasi data, pengenalan pola, dan database [10]. Sistem ini terdiri dari 3 tahapan yaitu input, proses, dan output. Input dapat berupa set data, prosesnya adalah algoritma atau metode dalam data mining, dan outputnya berupa pola, cluster, decision tree, dan lain-lain.

Tahap-tahap data mining ada 6 yaitu :

1) Pembersihan data (*data cleaning*)

Pembersihan data merupakan proses menghilangkan noise data yang tidak konsisten atau data tidak relevan. Pembersihan data akan mempengaruhi performansi dari teknik data mining karena data yang ditangani akan berkurang jumlah dan kompleksitasnya.

2) Integrasi data (*data integration*)

Integrasi data merupakan penggabungan data dari berbagai database ke dalam satu database baru. Integrasi data dilakukan pada atribut-atribut yang mengidentifikasi entitas-entitas yang unik seperti atribut nama, jenis produk, nomor pelanggan, dan lainnya.

3) Transformasi data (*data transformation*)

Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam data mining beberapa metode data mining membutuhkan format data yang khusus sebelum bisa diaplikasikan. Proses ini disebut transformasi data karena digunakan beberapa metode data mining yang

berbeda seperti estimasi, prediksi, klasifikasi, clustering, dan asosiasi.

4) Proses mining

Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data. Setelah proses transformasi ditemukan mining dari data yang diolah dengan metode tersebut menjadi model.

5) Evaluasi pola (*pattern evaluation*)

Untuk mengidentifikasi pola-pola menarik dikedalam *knowledge based* yang ditemukan. Dalam tahap ini hasil dari data mining berupa pola-pola yang khas maupun model prediksi dievaluasi untuk menilai apakah hipotesa yang ada memang tercapai.

6) Presentasi pengetahuan

Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna. Tahap terakhir ini adalah bagaimana memformulasikan keputusan atau aksi dari hasil analisis yang didapat [11].

5. *Software* Pendukung

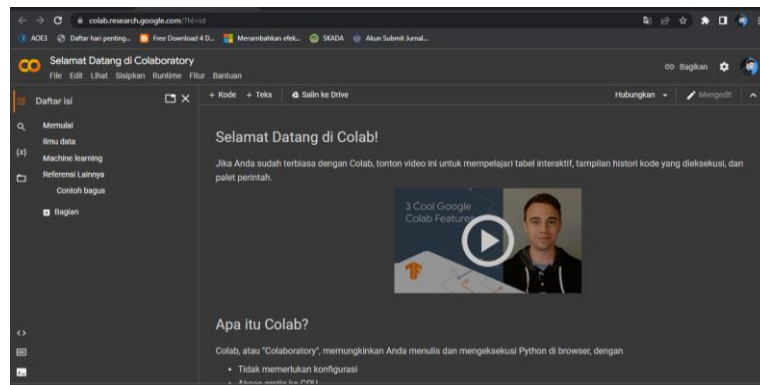
Software yang digunakan oleh peneliti yakni Google Colab yang memiliki tampilan hampir serupa dengan Jupyter Notebook, berbentuk cloud dan memudahkan pengguna untuk menjalankan kode Phyton tanpa harus *setting* dan *adjustment*. Berikut adalah beberapa manfaat menggunakan Google Colab :

- 1) Free GPU, Google Colab memudahkan pengguna untuk menjalankan program pada komputer dengan spek tinggi (GPU Tesla, RAM ~12GB, Disk ~300GB dan running dalam waktu yang lama (Google Colab mengizinkan kita untuk merunning program hingga 12 jam). Hal ini membantu peneliti yang terhalang sulitnya mendapatkan akses komputer dengan spek tinggi.
- 2) Collaborate, Google Colab juga memudahkan pengguna berkolaborasi dengan orang lain dengan cara membagi *file* secara online (mirip Google Doc). Pengguna bisa lebih mudah bereksperimen secara bersamaan atau sekedar membaca hasil code dengan tampilan yang rapi.
- 3) Mudah berintegrasi, Google Colab terbilang sangat fleksibel dalam hal integrasi. Pengguna dapat dengan mudah menghubungkan Google Colab dengan jupyter notebook di komputer lain (local runtime), menggunakan Google Drive, atau dengan Github
- 4) Fleksibel, pengguna dapat dengan mudah melakukan running melalui *smartphone* karena pada esensinya Google Colab hanya perlu running menggunakan browser, pengguna dapat mengawasi proses *training* (atau bahkan coding) via browser *smartphone* selama *smartphone* pengguna terhubung dengan Google Drive yang sama.

Berikut adalah tampilan dari Google Colab yang dapat kita akses melalui browser :

1) Tampilan Awal

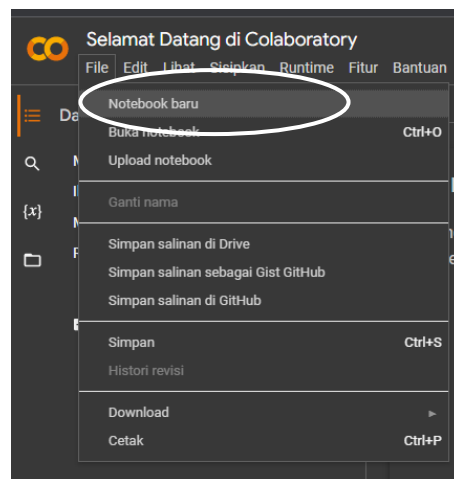
Pada tampilan awal pengguna akan disambut dengan penjelasan mengenai Google Colab tersebut dan berbagai tutorial, contoh, bahkan refensi untuk memulai menggunakan Google Colab.



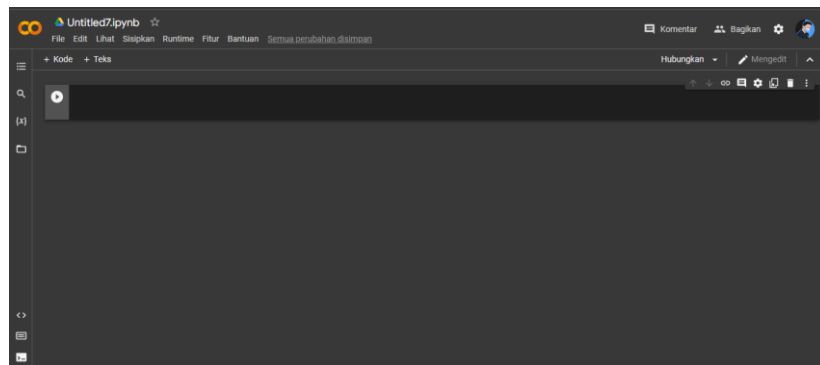
Gambar 2. 1 Tampilan Awal Google Colab

2) Membuat Notebook baru

Untuk memulai menggunakan Google Colab pengguna perlu membuat notebook baru melalui menu file kemudian pilih Notebook baru. Selanjutnya akan muncul tampilan awal untuk notebook baru.



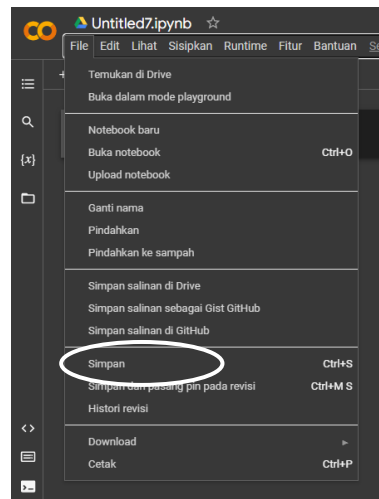
Gambar 2. 2 Membuat Notebook baru



Gambar 2. 3 Tampilan Notebook baru

3) Menyimpan dan download file

Setelah selesai menggunakan Google Colab pengguna perlu menyimpan notebook yang sudah dibuat untuk kemudian di download melalui menu file, kemudian pilih simpan setelah itu file bisa di download melalui menu download dan memilih ekstensi file yang diinginkan.



Gambar 2. 4 Menyimpan file Notebook



Gambar 2. 5 Mendownload file Notebook

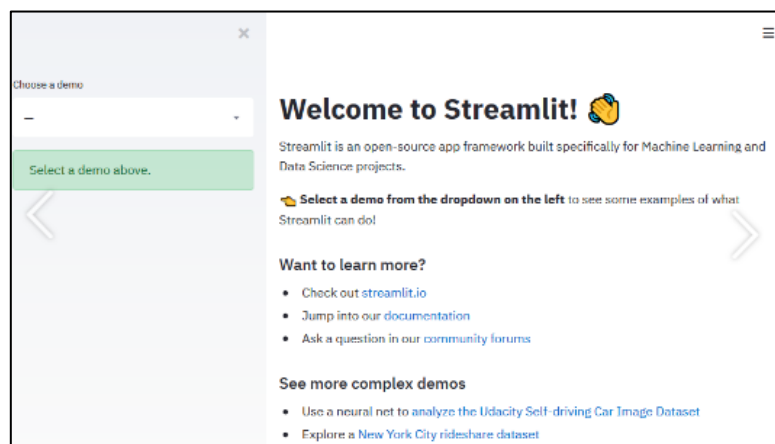
Selain Google Colab, peneliti juga menggunakan framework Streamlite. Streamlite ini merupakan framework berbasis Python dan bersifat open-source untuk memudahkan dalam membangun aplikasi web dibidang sains data dan *machine learning* yang interaktif. Sebelum menggunakan streamlite pastikan sudah meng-*install* Python 3.6 di perangkat yang akan digunakan. Setelah itu pengguna dapat melakukan install Streamlite menggunakan Pip melalui terminal menggunakan perintah berikut ini :

```
pip install streamlit
```

Setelah itu untuk mengecek instalasi aplikasi cukup mengetik perintah seperti berikut ini melalui terminal :

```
streamlit hello
```

Kemudian akan muncul seperti gambare berikut ini, hasilnya akan terbuka ditab baru pada browser seperti berikut ini :



Gambar 2. 6 Tampilan Streamlite

Melalui halaman diatas pengguna dapat melihat beberapa contoh demo dari aplikasi streamlite. Contoh yang ada telah disediakan beserta *source code* nya, beberapa contoh seperti animasi, *plotting*, *mapping*, dan *dataframe*. Hal terpenting sebelum menggunakan streamlit adalah meng-*import* beberapa library yang akan digunakan seperti berikut ini :

```
import streamlit as st
import pandas as pd
import numpy as np
```

BAB III

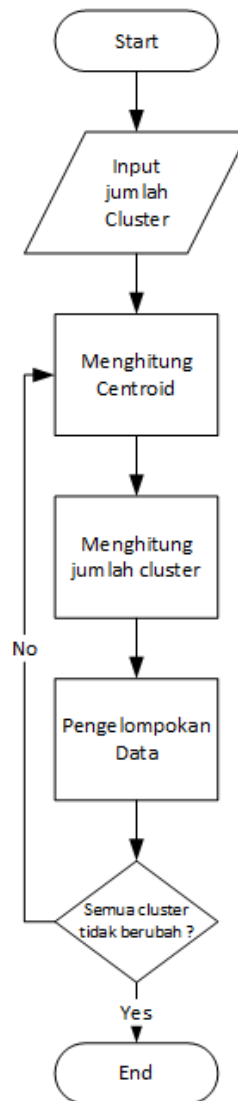
METODE PENELITIAN

A. Pendekatan Penelitian

Dalam pembangunan sebuah aplikasi diperlukan pendekatan untuk mempermudah proses dalam penyelesaian sistem yang diinginkan. Adapun untuk pengembangan sistem menggunakan bahasa pemrograman Python, sedangkan untuk pemodelan data dengan menggunakan metode data mining.

1) Metode Data Mining

Dalam analisis data metode yang dilakukan adalah menggunakan metode data mining algoritma clustering K-Means. Algoritma ini berguna untuk melakukan pengelompokan data yang memiliki kesamaan antara data satu dengan yang lainnya dengan syarat tertentu. Algoritma K-means ini dapat mengelompokkan data yang berskala tidak terlalu besar dengan cepat dan efektif. Algoritma K-Means ini dimulai dengan menentukan terlebih dahulu berapa cluster yang ingin dibuat, selanjutnya menentukan nilai awal untuk masing-masing cluster, lalu menghitung jarak dari masing-masing data dengan nilai awal yang telah ditentukan. Setelah jarak dari tiap data dihitung, data dimasukkan dalam cluster terdekat. Selanjutnya menghitung rata-rata dari tiap cluster, rata-rata tersebut nantinya akan digunakan sebagai nilai awal yang baru dalam menghitung jarak dari masing-masing data. Berikut adalah flowchart clustering k-means dapat dilihat pada gambar 3.1 :



Gambar 3. 1 Flowchart Clustering K-Means

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

1. Menentukan nilai K untuk jumlah cluster yang ingin dibentuk
2. Memilih nilai untuk pusat cluster awal (centroid) sebanyak K
3. Menghitung jarak setiap data input terhadap masing-masing centroid menggunakan rumus jarak Euclidean hingga ditemukan jarak yang paling dekat dari setiap data dengan centroid

Berikut adalah rumus Euclidean Distance :

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

Rumus 3.1 Euclidean Distance

Keterangan :

d = Jarak Euclidean

x_1 = Koordinat x untuk fasilitas 1

y_1 = Koordinat y untuk fasilitas 1

x_2 = Koordinat x untuk fasilitas 2

y_2 = Koordinat y untuk fasilitas 2

Setelah melakukan perhitungan jarak maka dilakukan pengelompokan data yang mendekati titik pusat. Kemudian dilakukan iterasi perhitungan ulang sampai mendapatkan pola yang sama sehingga iterasi selesai dan mendapatkan pengelompokan data yang merupakan hasil dari algoritma. Berikut adalah contoh perhitungan menggunakan rumus *Euclidean Distance* :

Contoh perhitungan menggunakan 6 buah data yang akan dikelompokkan menjadi 2 cluster yaitu data K1 dan K2.

Tabel 3. 1 Tabel Contoh Kelompok Data

Sample Data	X	Y	Cluster
1	100	50	
2	40	60	
3	30	70	
4	90	10	
5	65	40	
6	25	35	

Perhitungan pertama menggunakan data ke-1 dan ke-2 untuk mendapatkan jarak minimum data terhadap centroid.

Tabel 3. 2 Tabel Data Ke-1 dan Ke-2

Cluster	X	Y
K1	100	50
K2	40	60

Jarak cluster 1 ke cluster 1 Custer 1 = (100 ,50)

$$d(x1, c1) = \sqrt{(a1 - c1a)^2} + \sqrt{(b1 - c1b)^2} =$$

$$\sqrt{(100 - 100)^2} + (50 - 50)^2 = 0$$

Jarak dari cluster 2 ke cluster 1 (40 ,60) ↔ (100 ,50) =

$$d(x1, c1) = \sqrt{(a1 - c1a)^2} + \sqrt{(b1 - c1b)^2}$$

$$= \sqrt{(40 - 100)^2} + (60 - 50)^2 = 60.83$$

Jarak dari cluster 1 ke cluster 2 (100 ,50) ↔ (40 ,60) =

$$d(x1, c1) = \sqrt{(a1 - c1a)^2} + \sqrt{(b1 - c1b)^2}$$

$$= \sqrt{(100 - 40)^2} + (50 - 60)^2 = 60.83$$

Jarak cluster 2 ke cluster 2 (40 ,60)

$$d(x1, c1) = \sqrt{(a1 - c1a)^2} + \sqrt{(b1 - c1b)^2}$$

$$= \sqrt{(40 - 40)^2} + (60 - 60)^2 = 0$$

Dari hasil perhitungan pertama mendapatkan hasil sebagai berikut :

Tabel 3. 3 Tabel Hasil Pengelompokan Clustering

Cluster	Centroid		Kelompok Cluster
	X	Y	
K1 (100 ,50)	0	60.83	1
K2 (40 ,60)	60.83	0	2

Dari tabel diatas dapat dilihat jarak minium dari data 1 ke data 1 adalah 0

dan jarak minimum dari data 2 ke data 2 yaitu 0. Sehingga centroid K1 adalah data 1 dan data 2 menjadi centroid K2. Selanjutnya dilakukan perhitungan kedua untuk data ke-3, berikut adalah perhitungannya :

Dimulai dari perhitungan jarak dataset terhadap cluster 1

$$(100, 50) \leftrightarrow (30, 70)$$

$$\begin{aligned} d(x1, c1) &= \sqrt{(a1 - c1a)^2} + \sqrt{(b1 - c1b)^2} \\ &= \sqrt{(30 - 100)^2} + \sqrt{(70 - 50)^2} = 72.80 \end{aligned}$$

Kemudian perhitungan jarak dataset terhadap cluster 2

$$(40, 60) \leftrightarrow (30, 70)$$

$$\begin{aligned} d(x1, c1) &= \sqrt{(a1 - c1a)^2} + \sqrt{(b1 - c1b)^2} \\ &= \sqrt{(30 - 40)^2} + \sqrt{(70 - 60)^2} = 14.14 \end{aligned}$$

Tabel 3. 4 Tabel Hasil Perhitungan Data

Dataset	Euclidean Distance		Kelompok Cluster
	Cluster 1	Cluster 2	
(30, 70)	72.80	14.14	2

Dilihat dari jarak minimum data terhadap centroid, data ke-3 lebih dekat ke cluster 2 dengan nilai 14.14. Selanjutnya adalah membuat centroid baru karena data ke-3 masuk ke K2 maka centroid K2 diperbarui dengan cara :

$$X \text{ centroid baru} = (x_{K2} + x_{data3})/2$$

$$Y \text{ centroid baru} = (y_{K2} + y_{data3})/2$$

Sehingga menghasilkan centroid baru pada tabel berikut ini :

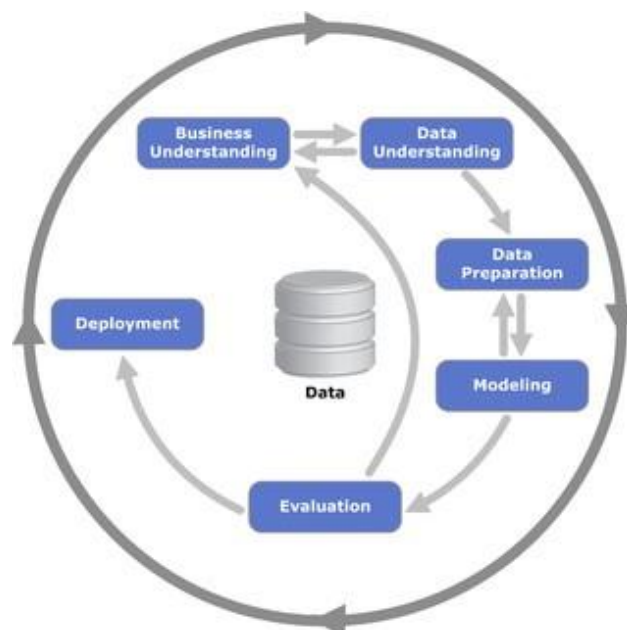
Tabel 3. 5 Tabel Hasil Centroid Baru

Cluster	X	Y
K1	100	50
K2	$= \frac{40 + 30}{2} = 35$	$= \frac{60 + 70}{2} = 65$

Perhitungan selanjutnya untuk dataset ke-4, 5, dan seterusnya mengikuti langkah-langkah seperti pada dataset ke-3. Yang perlu diperhatikan adalah menggunakan centroid yang diperbarui dan juga melakukan perhitungan untuk centroid yang baru. Setelah dilakukan perhitungan menggunakan rumus kemudian data tersebut dibuat kedalam model menggunakan bahasa pemrograman python.

2) Metode CRISP-DM

Metode pemodelan data yang dipakai adalah menggunakan metode CRISP-DM. Proses data mining harus terukur, dapat dipercaya dan memenuhi suatu standar. CRISP-DM (*The Cross Industry Standard Process for Data Mining*) adalah standarisasi data mining yang disusun oleh tiga penggagas data mining market yaitu Daimler Chrysler (Daimler-Benz), SPSS (ISL), NCR. Dalam proses data mining berdasarkan CRISP-DM terdiri dari 6 tahapan sebagai berikut :

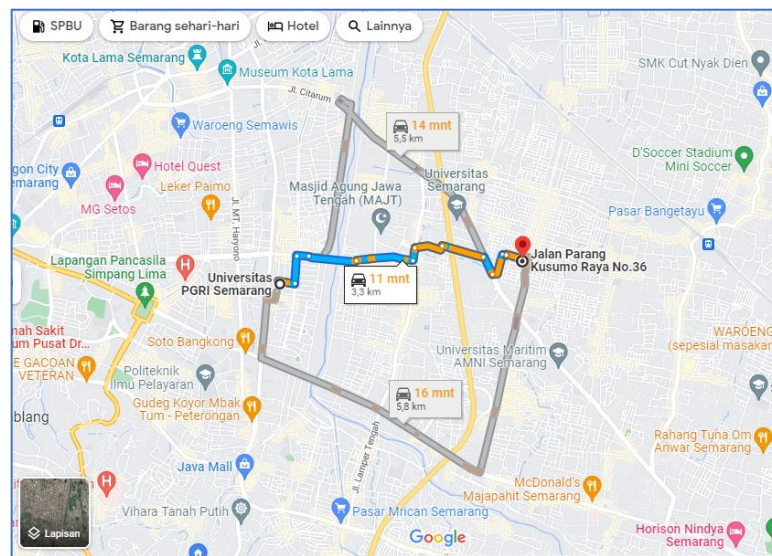


Gambar 3. 2 Model CRISP-DM

- a. *Business Understanding* adalah pemahaman tentang substansi dari kegiatan data mining yang akan dilakukan, kebutuhan dari perspektif bisnis. Kegiatannya antara lain: menentukan sasaran atau tujuan bisnis, memahami situasi bisnis, menentukan tujuan data mining dan membuat perencanaan strategi serta jadwal penelitian.
- b. *Data Understanding* adalah fase mengumpulkan data awal, mempelajari data untuk bisa mengenal data yang akan dipakai, mengidentifikasi masalah yang berkaitan dengan kualitas data, mendeteksi subset yang menarik dari data untuk membuat hipotesa awal.
- c. *Data preparation* sering disebut sebagai fase yang padat karya. Aktivitas yang dilakukan antara lain memilih table dan field yang akan ditransformasikan ke dalam database baru untuk bahan data mining (set data mentah).
- d. *Modeling* adalah fase menentukan tehnik data mining yang digunakan, menentukan tools data mining, teknik data mining, algoritma data mining, menentukan parameter dengan nilai yang optimal.
- e. *Evaluation* adalah fase interpretasi terhadap hasil data mining yang ditunjukkan dalam proses pemodelan pada fase sebelumnya. Evaluasi dilakukan secara mendalam dengan tujuan menyesuaikan model yang didapat agar sesuai dengan sasaran yang ingin dicapai dalam fase pertama.
- f. *Deployment* atau penyebaran adalah fase penyusunan laporan atau presentasi dari pengetahuan yang didapat dari evaluasi pada proses data mining [14].

B. Lokasi Penelitian

Lokasi penelitian adalah Toko Premium Shopping yang beralamat di Jalan Tlogosari Raya no 36, Kelurahan Tlogosari Kulon, Kecamatan Pedurungan, Kota Semarang 50196. Untuk fokus penelitian penulis langsung mewawancarai Kepala Toko dan Karyawan toko untuk mendapat data yang tepat dan akurat.



Gambar 3. 3 Lokasi Penelitian

C. Populasi dan Sampel

Populasi menurut ahli adalah keseluruhan subyek penelitian [12]. Populasi dari penelitian ini adalah seluruh data penjualan reseller Toko Premium Shopping mulai tahun 2019 hingga tahun 2021. Sedangkan sampel penelitian menurut Arikunto (2006) adalah sebagian atau wakil populasi yang akan diteliti. Adapun data yang diperoleh yakni data hasil rekapitulasi penjualan reseller dari Toko Premium Shopping dianggap terlalu banyak dan memakan waktu yang lama dalam melakukan perhitungan.

Maka dari data tersebut ditarik sampel dengan menggunakan metode *purposive sampling*. Teknik *purposive sampling* dilakukan dengan memilih responden berdasarkan kriteria tertentu [13]. Kriteria data yang

diperlukan untuk penelitian ini adalah rekapitulasi data penjualan reseller dalam kurun waktu 1 tahun terakhir. Sehingga sampel yang digunakan adalah hasil rekapitulasi penjualan reseller Toko Premium Shopping pada tahun 2021.

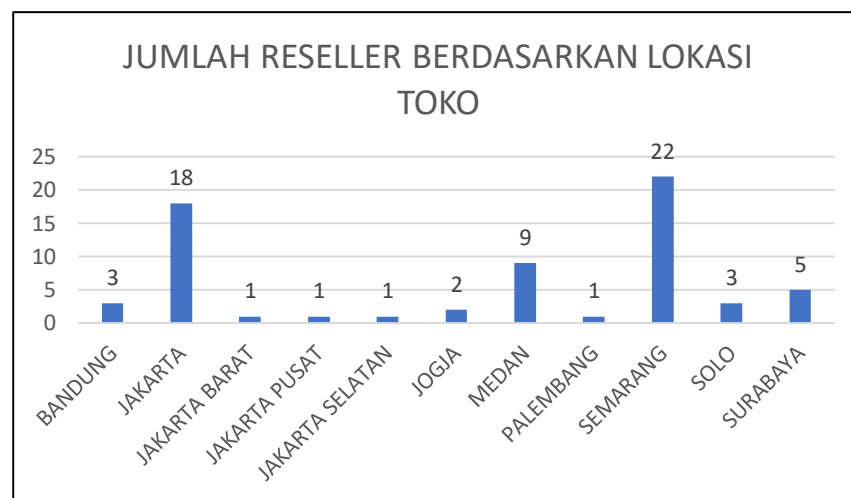
D. Variabel Penelitian

Pada penelitian ini variabel yang akan digunakan adalah barang atau produk yang dijual oleh *reseller* dan data *reseller* dari Toko Premium Shopping, berikut adalah tabel data produk pada tabel 3.6 dan gambar diagram jumlah reseller pada gambar 3.4 dibawah ini :

Tabel 3. 6 Tabel Daftar Produk

NO	KODE PRODUK	NAMA PRODUK
1	AR1	Aros
2	BY1	Byoote
3	BYT2	Byt colla
4	BGLO3	Byt glo oil
5	CREV1	Crevolene
6	EM1	Evomass
7	EW2	Evowhey
8	FBAR1	Flimbar
9	FMEAL2	Flimeal
10	FLIM3	Flimty
11	HLEAN1	Herlean
12	HSLIM2	Herslim
13	ILENE1	Isolene
14	MISO1	Muscle proisolate
15	MW2	Muscle prowhey
16	PREV1	Prevo
17	SOP1	Sopalpa

Dari tabel tersebut dapat diketahui berbagai produk yang dijual oleh reseller Toko Premium Shopping antara lain Aros, Byoote, Byt colla, Byt glo oil, Crevolene, Evomass, Evowhey, Flimbar, Flimeal, Flimty, Herlean, Herslim, Isolene, Muscle proisolate, Muscle prowhey, Prevo dan Sopalpha. Selain itu data reseller dari Toko Premium Shopping juga diperlukan, berikut adalah tampilan data dalam bentuk diagram jumlah reseller :



Gambar 3. 4 Diagram Jumlah Reseller

E. Pengumpulan Data

Untuk mendapatkan data yang benar-benar akurat, relevan, dan valid maka penulis mengumpulkan data dengan cara :

a. Teknik observasi

Teknik observasi yaitu dengan melakukan observasi langsung ke lokasi penelitian yakni Toko Premium Shopping. Tujuan observasi untuk meminta izin kepada pihak manajemen untuk mendapatkan data primer berupa Data Penjualan Reseller Toko Premium Shopping dan melakukan pengamatan secara langsung mengenai permasalahan yang ada pada Toko Premium Shopping.

b. Teknik studi pustaka

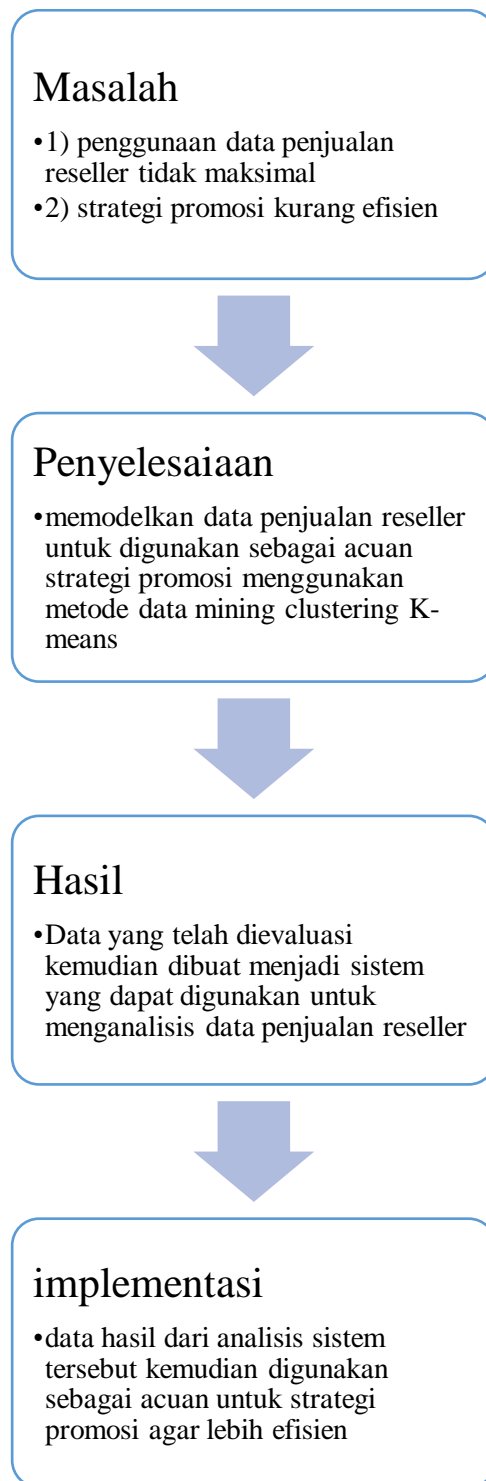
Studi pustaka merupakan teknik pengumpulan data dengan cara mempelajari dan membaca berbagai macam buku referensi atau jurnal yang berkaitan dengan penelitian.

F. Kerangka Berpikir

Beberapa upaya yang dilakukan perusahaan untuk meningkatkan penjualan dan bersaing dengan perusahaan lainnya yakni dengan cara menambah outlet atau reseller baru diberbagai kota dan daerah. Dengan jumlah reseller yang tidak sedikit dan pelanggan yang bertransaksi setiap harinya tidak menutup kemungkinan menghasilkan data penjualan yang sangat banyak. Untuk itu perusahaan memerlukan sistem yang dapat memudahkan dalam melakukan clustering reseller potensial. Proses clustering data menggunakan sistem diperlukan metode data mining yakni K-Means clustering.

Dengan adanya sistem ini dapat membantu perusahaan dalam memaksimalkan data penjualan reseller yang dimiliki perusahaan untuk target pemasaran berdasarkan hasil clustering reseller potensial. Sehingga promosi perusahaan lebih terarah, tepat sasaran dan menghemat biaya promosi.

Berikut adalah kerangka berpikir dalam penelitian ini yang disajikan dalam bentuk diagram pada gambar 3.5 dibawah ini :



Gambar 3. 5 Kerangka Berpikir

BAB IV

HASIL DAN PEMBAHASAN

A. Pemahaman Bisnis (Business Understanding)

Tahapan ini merupakan tahap pertama dalam proses CRISP-DM yang dapat disebut juga sebagai tahap pemahaman bisnis. Tahapan ini dibutuhkan dalam rangka untuk memahami data yang akan dianalisis. Peneliti telah melakukan observasi pada Toko Premium Shopping, dimana hasil yang didapatkan yakni toko tersebut memiliki tujuan bisnis untuk memperluas jangkauan konsumen dengan cara membuka reseller sebanyak-banyaknya. Dari tujuan bisnis yang ingin dicapai oleh Toko Premium Shopping, peneliti memutuskan untuk mengelola data hasil penjualan reseller pada Toko Premium Shopping menggunakan data mining. Data penjualan reseller tersebut didapatkan dari hasil penjualan berbagai reseller dari Toko Premium Shopping yang kemudian di kumpulkan dan di proses oleh admin untuk di rekapitulasi menjadi satu.

Tujuan pengelolaan data hasil penjualan reseller tersebut menggunakan data mining adalah untuk dijadikan sebagai dasar dalam menentukan target wilayah pemasaran dan promosi agar tepat sasaran. Sehingga sebagai langkah awal dalam menerapkan tujuan dilakukannya data mining adalah melakukan permintaan data hasil rekapitulasi penjualan reseller pada Toko Premium Shopping terlebih dahulu ke bagian admin untuk kemudian diproses pada tahapan selanjutnya.

B. Tahap Pemahaman Data (Data Understanding)

Dari dataset yang telah diperoleh dari admin Toko Premium Shopping berupa dokumen excel yang berisi 66 data reseller. Adapun sumber data utama yang digunakan dalam penelitian ini adalah dataset Rekapitulasi Hasil Penjualan Reseller Toko Premium Shopping tahun 2021 dengan format .xlsx karena data yang diberikan berupa dokumen excel.

Dataset hasil penjualan reseller ini terdiri dari atribut kode reseller, nama reseller, dan produk yang dijual yakni Aros, Byoote, Byt colla, Byt glo oil, Crevolene, Evomass, Evowhey, Flimbar, Flimeal, Flimty, Herlean, Herslim, Isolene, Muscle proisolate, Muscle prowhey, Prevo dan Sopalpha. Berikut adalah data yang akan diolah menggunakan algoritma K-means :

Data Rekapitulasi Penjualan Reseller Toko Premium Shopping Tahun 2021

NO	NAMA RESELLER	PRODUK																
		Aros	Byoote	Byt colla	Byt glo oil	Crevolene	Evomass	Evowhey	Flimbar	Flimeal	Flimty	Herlean	Herslim	Isolene	Muscle proisolate	Muscle prowhey	Prevo	Sop alpha
1	ajudansyifa		124	5	126		316	148		248	351	93	122	5			65	139
2	antonwarman1	6		27	27		2137	1100	48		1600	61	147	299	10	33	51	184
3	ARTHAMA97						67							32				
4	bluegreen12										94							
5	boworamadhan44	7		57		30	910	48		35	1209	628	496			7	4	183
6	chocofun13			10							195							
7	citrakumala22		25								45							
8	clarabell48										56			12				
9	clayrizha										112							
10	cokomelon						79	22		15	113							
11	danisgina		107	106	103	124	420	110		101	599		105					197
12	danitadara										119							
13	darengown					186	188	561		173	753			373				187
14	daylinestore			2							569							
15	distyanda										83							
16	elinanoer	31	523	681	210	375	789	788	605	216	660	218	270	659				624

NO	NAMA RESELLER	PRODUK																
		Aros	Byoote	Byt colla	Byt glo oil	Crevolene	Evomass	Evowhey	Flimbar	Flimeal	Flimty	Herlean	Herslim	Isolene	Muscle proisolate	Muscle prowhey	Prevo	Sop alpha
17	emil2435									189								
18	faradella87							35		9								
19	femmitaninda	26	82	26			21	42		513	21		163			24	49	
20	gondarina12			116						395								
21	honeymelon11									341								
22	jellyrainbow11									151								
23	jemimagina		138	276	140	150	574	280		642	1958	142	144			22	709	
24	Jikoshop09								38		105							
25	jionda11									60								
26	kiranaani12				32					76								
27	kristashellania	32	35	28	38	72	520	284		841	31	53		32	54	21	403	
28	lionashim									51								
29	lyodrinefine						188	188		430			240					
30	malikadwin					34				37								
31	mangosticky77									221								
32	melisa.putri7		41		52	34	96			53	535	38	54				63	
33	miskashow			94						366								
34	Monicahantika			12						15								

NO	NAMA RESELLER	PRODUK																
		Aros	Byoote	Byt colla	Byt glo oil	Crevolene	Evomass	Evowhey	Flimbar	Flimeal	Flimty	Herlean	Herslim	Isolene	Muscle proisolate	Muscle prowhey	Prevo	Sop alpha
35	myvirashell										133							
36	nauraindah55			13							45							
37	Nitaliya45										242							
38	nucadwianda					12					168							
39	okeyshopping12		144	54		160	561	613			701	22		473				
40	relaxday66		229	383		194	428	623			892	178		797				
41	revindersell										149						18	
42	rismalinda11										26							
43	ristaqueen								112		176							
44	rosalindaa78									95	18							
45	rownbean						334											
46	safiraoktaa	65	59		35	68	246	58		11	349	43					21	198
47	Sarinieta										93							
48	seablue00									1134	137							
49	setyoaji33	1	57		29	27	435	143		68	793	31	55	79			46	156
50	shienaputri45						104											43
51	shoppingday13						112											
52	showday11										390							

NO	NAMA RESELLER	PRODUK																
		Aros	Byoote	Byt colla	Byt glo oil	Crevolene	Evomass	Evowhey	Flimbar	Flimeal	Flimty	Herlean	Herslim	Isolene	Muscle proisolate	Muscle prowhey	Prevo	Sop alpha
53	simbalion22			29						396								
54	siskazien		178							43								119
55	soopinthe			21						193								
56	stopmotion33									266								
57	tamaninda		255	159		127	440	97		251	881			232				287
58	tatasquirel		9	18	9		20	17		16	45			36				18
59	tigershop17		105							48	189							41
60	todayshop22		65								177							98
61	turtledown_			31	115		120	124		190	361		113					241
62	unicorninshell								113		24							101
63	wendycandy22					11					209							
64	yindadinda									13	210							
65	zahraamanda13			18							92							
66	zanakirana							45		114								

Tabel 4. 1 Tabel Penjualan Reseller

Dari tabel penjualan reseller tersebut dapat diketahui jumlah penjualan dari masing-masing reseller dari Toko Premium Shopping. Kemudian data tabel diatas diubah formatnya menjadi .csv untuk selanjutnya diimpor kedalam google colab dengan perintah berikut ini, sehingga data tersebut akan muncul dalam bentuk tabel, untuk mempermudah dalam proses modeling :

```
df = pd.read_csv('Data Penjualan Reseller.csv')
df
```

C. Tahap Persiapan Data (Data preparation)

Data yang telah diperoleh kemudian dilakukan tahap preprocessing yakni dilakukan evaluasi data yang kosong, data sama, dan menghapus data yang tidak digunakan. Dari dataset yang digunakan peneliti melakukan 4 tahapan sebagai berikut :

a. Memeriksa data yang kosong

Untuk memeriksa apakah ada data yang kosong atau tidak peneliti menggunakan perintah sebagai berikut :

```
df.isnull().sum()
```

Dari dataset tersebut tidak terdapat data yang kosong sehingga ketika dilakukan pengecekan tidak muncul angka pada kolom sebelah kanan yang berarti data yang digunakan sudah terisi seluruhnya. Berikut adalah tampilan hasilnya :

```
[86] df.isnull().sum()
KODE RESELLER      0
NAMA RESELLER      0
Aros                0
Byoote             0
Byt colla          0
Byt glo oil        0
Crevolene          0
Evomass            0
Evowhey            0
Flimbar            0
Flimeal            0
Flimty             0
Herlean            0
Herslim            0
Isolene            0
Muscle proisolate  0
Muscle prowhey     0
Prevo              0
Sop alpha          0
dtype: int64
```

Gambar 4. 1 Periksa Data Null

- b. Memeriksa data yang sama dan menghapus kolom yang tidak digunakan

Selanjutnya peneliti melakukan pemeriksaan apakah ada data yang sama dengan menggunakan perintah berikut :

```
#Memeriksa data yang sama
df.duplicated().sum()
```

Peneliti menggunakan cara berikut untuk menghapus kolom yang tidak digunakan dalam tahap pemodelan nantinya :

```
#Menghapus kolom kode reseller
df.drop(columns=["KODE RESELLER"],axis=1,inplace=True)

#Menghapus kolom NAMA RESELLER
df.drop(columns=["NAMA RESELLER"],axis=1,inplace=True)

df.columns
```

Dari proses pemeriksaan data dan penghapusan kolom tersebut berikut adalah tampilan proses nya ketika dituliskan ke dalam sistem dan menghasilkan kolom yang akan digunakan sebagai berikut :

```
Index(['Aros ', 'Byoote', 'Byt colla', 'Byt glo oil', 'Crevolene',
'Evomass ', 'Evowhey', 'Flimbar', 'Flimeal', 'Flimty', 'Herlean',
'Herslim', 'Isolene', 'Muscle proisolate', 'Muscle prowhey',
'Prevo', 'Sop alpha'], dtype='object')
```

```
[87] #Memeriksa data yang sama
df.duplicated().sum()

0

[88] #Menghapus kolom kode reseller
df.drop(columns=["KODE RESELLER"], axis=1, inplace=True)

[89] #Menghapus kolom NAMA RESELLER
df.drop(columns=["NAMA RESELLER"], axis=1, inplace=True)

[90] df.columns

Index(['Aros ', 'Byoote', 'Byt colla', 'Byt glo oil', 'Crevolene', 'Evomass ',
'Evwhey', 'Flimbar', 'Flimeal', 'Flimty', 'Herlean', 'Herslim',
'Isolene', 'Muscle proisolate', 'Muscle prowhey', 'Prevo', 'Sop alpha'],
dtype='object')
```

Gambar 4. 2 Periksa Data sama dan hapus kolom

Sebelum melakukan tahapan modeling data yang akan digunakan perlu dilakukan proses scaling agar nantinya mendapatkan hasil yang maksimal. Data scaling ini berfungsi agar data yang digunakan berada dalam rentang yang tidak terlalu jauh. Berikut adalah penulisannya :

```
scaled_df = scalar.fit_transform(df)
```

D. Tahap Pemodelan (Modeling)

Setelah data melalui tahapan preprocessing kemudian data akan diolah menjadi sebuah model dengan berbagai tahapan. Berikut adalah tahapan-tahapan yang dilakukan dalam proses clustering :

a. Menentukan jumlah cluster

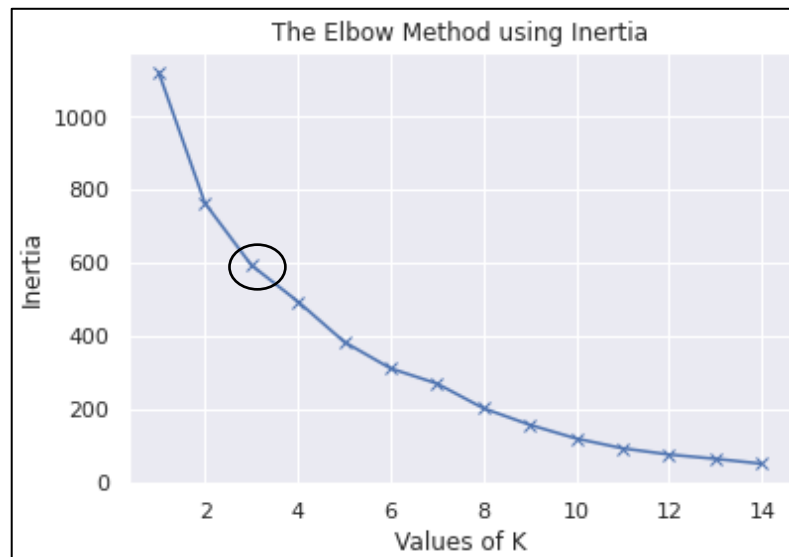
Untuk menentukan jumlah cluster peneliti menggunakan Elbow Method dengan cara sebagai berikut :

```

inertia = []
range_val = range(1,15)
for i in range_val:
    kmean = KMeans(n_clusters=i)
    kmean.fit_predict(pd.DataFrame(scaled_df))
    inertia.append(kmean.inertia_)
plt.plot(range_val,inertia,'bx-')
plt.xlabel('Values of K')
plt.ylabel('Inertia')
plt.title('The Elbow Method using Inertia')
plt.show()

```

Kemudian akan muncul diagram dari perintah `plt.show()` seperti berikut ini :



Gambar 4. 3 Elbow Method

Dari hasil metode elbow menghasilkan diagram seperti gambar diatas yang digunakan untuk menentukan jumlah cluster pada proses clustering. Berdasarkan diagram diatas titik yang hampir membentuk siku berada diantara 2 dan 4 yang berarti jumlah cluster yang akan digunakan pada K-means clustering

adalah 3. Dari hasil tersebut data akan dibagi ke dalam 3 cluster dengan kriteria sebagai berikut :

Tabel 4. 2 Kriteria cluster

Cluster	Nama Cluster	Kriteria
Cluster 0	Kurang potensial	Hasil penjualan mulai 1-6 produk, dengan masing-masing produk terjual lebih dari 200 pcs.
Cluster 1	Cukup potensial	Penjualan mulai dari 4-10 produk, dengan penjualan masing-masing produk lebih dari 200 pcs.
Cluster 2	Sangat potensial	Produk terjual lebih dari 10 jenis produk, masing-masing produk memiliki penjualan lebih dari 200 pcs.

b. Proses clustering

Untuk melakukan proses clustering peneliti menggunakan perintah sebagai berikut :

```
kmeans_model=KMeans(3)
kmeans_model.fit_predict(scaled_df)
df_kmeans= pd.concat([df,pd.DataFrame({'cluster':kmeans_model.labels_})],axis=1)
```

c. Pusat Cluster

Untuk melihat pusat cluster yang digunakan menggunakan perintah berikut ini :

```
cluster_centers = pd.DataFrame(data=kmeans_model.cluster_centers_,columns=[df.columns])
cluster_centers = scalar.inverse_transform(cluster_centers)
cluster_centers = pd.DataFrame(data=cluster_centers,columns=[df.columns])
cluster_centers
```

```

162] cluster_centers = pd.DataFrame(data=kmeans_model.cluster_centers_,columns=[df.columns])

cluster_centers = scalar.inverse_transform(cluster_centers)
cluster_centers = pd.DataFrame(data=cluster_centers,columns=[df.columns])
cluster_centers

```

	Aros	Boote	Byt colla	Byt glo oil	Crevolene	Evomass	Ewohney	Flimbar	Flimeal	Flinty	Herlean	Herslim	Isolene	Muscle proisolate
0	4.181818	99.000000	99.545455	42.090909	97.272727	629.909091	364.272727	4.363636	138.000000	961.636364	107.818182	102.000000	205.272727	3.818182e+0
1	1.685185	10.444444	7.222222	4.500000	2.944444	25.685185	9.833333	4.870370	32.074074	166.074074	2.055556	3.092593	8.944444	8.881784e-1
2	31.000000	523.000000	681.000000	210.000000	375.000000	789.000000	788.000000	605.000000	216.000000	660.000000	218.000000	270.000000	659.000000	0.000000e+0

Gambar 4. 4 Pusat Cluster

d. Tabel Hasil Clustering

Untuk melihat hasil clustering kedalam bentuk tabel menggunakan perintah sebagai berikut :

```

cluster_df = pd.concat([df,pd.DataFrame({'Cluster':kmeans_m
odel.labels_})],axis=1)

cluster_df

```

Tabel hasil clustering dapat dilihat pada lampiran 1 Tabel Hasil Clustering.

e. Tabel Cluster 0

Untuk melihat tabel hasil clustering khusus pada cluster 0 saja, dapat menggunakan cara sebagai berikut :

```

cluster_0_df = cluster_df[cluster_df["Cluster"]==0]

cluster_0_df

```

Dengan menggunakan perintah `cluster_0_df` maka tabel data yang akan muncul hanya cluster 0 saja. Tabel hasil cluster 0 dapat dilihat pada lampiran 2 Tabel Hasil Cluster 0.

f. Tabel Cluster 1

Sedangkan untuk melihat tabel khusus cluster 1 cara yang digunakan adalah berikut ini :

```

cluster_1_df = cluster_df[cluster_df["Cluster"]==1]

cluster_1_df

```

Dengan menggunakan perintah `cluster_1_df` maka data yang akan muncul hanya data yang termasuk ke dalam cluster 1 saja, tabel dapat dilihat pada lampiran ke 3 Tabel Hasil Cluster 1.

g. Tabel Cluster 2

Untuk melihat data khusus pada cluster 2 cara yang digunakan sama dengan cluster sebelumnya hanya angkanya saja yang diubah menjadi cluster 2 seperti berikut ini :

```
cluster_2_df = cluster_df[cluster_df["Cluster"]==2]
cluster_2_df
```

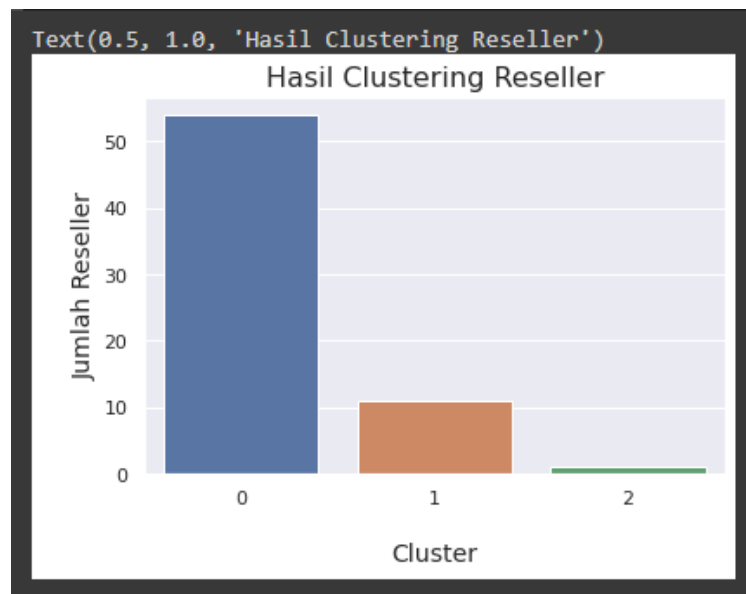
Dan akan muncul tabel hasil cluster 2, tabel dapat dilihat pada lampiran ke 4 Tabel Hasil Cluster 2.

h. Diagram Hasil Clustering

Sedangkan untuk melihat diagram dari hasil clustering peneliti menggunakan perintah berikut :

```
#Visualization
sns.countplot(x='Cluster', data=cluster_df)
sns.set_theme(style="darkgrid")
plt.xlabel('\nCluster', size=14)
plt.ylabel('\nJumlah Reseller', size=14)
plt.title('Hasil Clustering Reseller', size=16)
```

Maka akan muncul diagram batang sesuai dengan hasil clustering seperti pada gambar dibawah ini :



Gambar 4. 5 Diagram Hasil Clustering

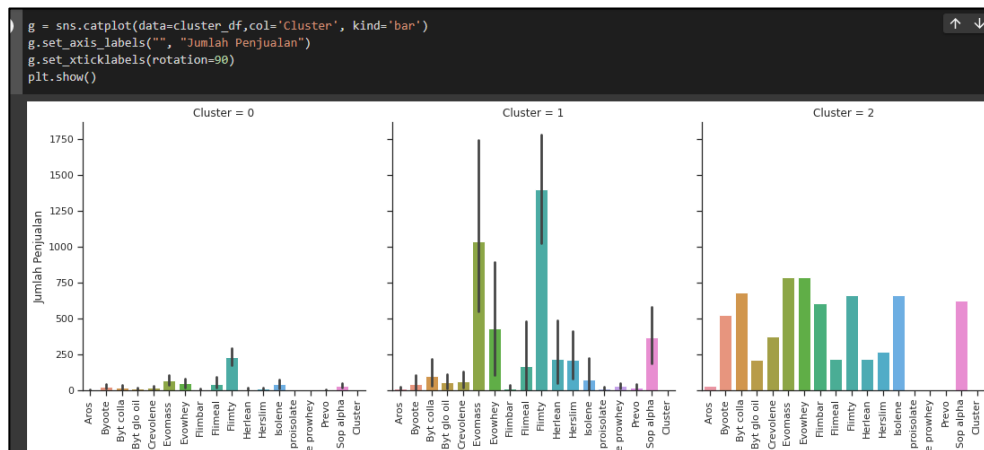
Dari diagram diatas dapat dilihat hasil clustering dimana cluster 0 menjaid cluster dengan jumlah 54 reseller, kemudian cluster 1 dengan jumlah 12 reseller dan cluster 2 dengan 1 reseller.

i. Diagram masing-masing cluster

Untuk menampilkan diagram dari masing-masing cluster peneliti menggunakan perintah berikut :

```
g = sns.catplot(data=cluster_df,col='Cluster', kind='bar')
g.set_axis_labels("", "Jumlah Penjualan")
g.set_xticklabels(rotation=90)
plt.show()
```

Dengan menggunakan perintah `plt.show()` maka akan muncul diagram batang dari masing-masing cluster seperti pada gambar berikut :



Gambar 4. 6 Diagram masing-masing cluster

j. Model data dan Data hasil clustering disimpan

Kemudian setelah proses clustering selesai peneliti menggunakan cara berikut ini untuk menyimpan data hasil clustering ke dalam file dengan format .csv dan menyimpan dalam bentuk save model.

```
#Saving Scikitlearn models
import joblib
joblib.dump(kmeans_model, "kmeans_model.pkl")
cluster_df.to_csv("Clustered_Reseller_Data.csv")
```

```
[169] #Saving Scikitlearn models
import joblib
joblib.dump(kmeans_model, "kmeans_model.pkl")

['kmeans_model.pkl']

[170] cluster_df.to_csv("Clustered_Reseller_Data.csv")
```

Gambar 4. 7 Model Data Disimpan

Dari proses clustering yang telah dilakukan peneliti yang terdiri dari 3 cluster. Hasil proses clustering reseller dapat dilihat dalam bentuk tabel pada Lampiran ke 5, berikut uraian pengelompokan dari masing-masing cluster sebagai berikut :

- 1) Cluster 0 (Reseller kurang potensial) terdapat 54 reseller yang masuk ke dalam cluster ini. Dalam kelompok data ini kurang potensial karena tingkat penjualan yang rendah. Reseller yang termasuk ke dalam cluster ini adalah :
- 2) Cluster 1 (Reseller cukup potensial) terdiri dari 12 reseller, dimana reseller dalam kelompok ini cukup potensial karena hasil penjualan yang cukup tinggi namun perlu ditingkatkan lagi. Reseller yang termasuk ke dalam cluster 1 adalah :
- 3) Cluster 2 (Reseller sangat potensial) pada cluster ini hanya terdapat 1 reseller dimana tingkat penjualan lebih tinggi dibandingkan dengan reseller lainnya dan termasuk reseller yang potensial. Reseller yang termasuk ke dalam cluster 2 yakni sangat potensial adalah elinanoer.

Sehingga dari hasil clustering tersebut perusahaan dapat meningkatkan promosi di area reseller yang kurang potensial dan cukup potensial agar lebih banyak konsumen yang membeli dan menambah reseller baru di area reseller potensial untuk memperluas jangkauan konsumen yang lebih luas.

E. Evaluation

Setelah data selesai diolah dan mendapatkan hasil maka hasil tersebut dapat disajikan dalam bentuk tabel spesifik yang dapat dilihat

secara langsung pengelompokkan data berdasarkan cluster yang dibuat. Kemudian data hasil clustering tersebut akan dievaluasi menggunakan *precision* dan *recall* sebagai berikut :

a. Data hasil clustering di lakukan split data

Sebelum dilakukan proses *accuracy* data hasil clustering perlu dipisahkan menjadi 2 yakni Training-Set dan Test-Set terlebih dahulu. Berikut adalah perintah seperti dibawah ini :

```
#Split Dataset
X = cluster_df.drop(['Cluster'],axis=1)
y= cluster_df[['Cluster']]
X_train, X_test, y_train, y_test =train_test_split(X, y, test_size
=0.3)

X_train
```

Dari perintah split data diatas menghasilkan 2 data yakni data `X_train` dan `X_test`. Data `X_train` adalah data yang digunakan untuk membuat model machine learningnya. Sedangkan data `X_test` untuk menguji performa dari model yang telah dibuat.

b. Menghitung evaluasi model kemudian di simpan

Kemudian data `X_test` yang telah dipisahkan akan digunakan untuk menguji model yang telah dibuat. Seperti pada text dibawah ini dari model yang telah dibuat, peneliti mengukur akurasi nya menggunakan *precision* dan *recall*.

```
model.fit(X_train, y_train)
y_pred = model.predict(X_test)

print(classification_report(y_test, y_pred))
```

```
print(classification_report(y_test, y_pred))
```

	precision	recall	f1-score	support
0	0.93	0.93	0.93	14
1	0.67	0.80	0.73	5
2	0.00	0.00	0.00	1
accuracy			0.85	20
macro avg	0.53	0.58	0.55	20
weighted avg	0.82	0.85	0.83	20

Gambar 4. 8 Evaluasi model

Dapat dilihat pada gambar hasil evaluasi model precision, recall, dan f1-score. Dimana pada cluster 0 mendapatkan nilai precision, recall dan f1-score sebesar 0.93. Sedangkan pada cluster 1 dengan nilai precision 0.67, recall 0.80 dan f1-score 0.73. Namun pada cluster 2 nilai precision, recall dan f1-score bernilai 0 karena hanya terdapat 1 data saja. Kemudian untuk hasil akurasi model yakni 0.85%.

Setelah data berhasil di akurasi kemudian model akhir akan disimpan kedalam format .sav seperti gambar berikut ini :

```
import pickle
filename = 'final_model.sav'
pickle.dump(model, open(filename, 'wb'))

loaded_model = pickle.load(open(filename, 'rb'))
result = loaded_model.score(x_test, y_test)
print(result, '% Accuracy')
```

0.85 % Accuracy

Gambar 4. 9 Save Model dan Akurasi

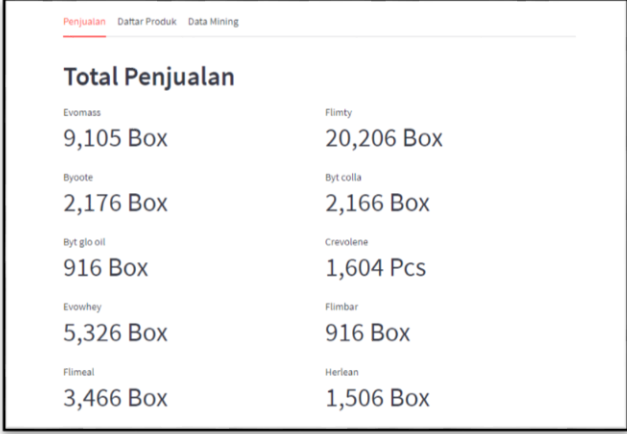
Model disimpan kedalam format sav karena data akan ditampilkan kembali dalam bentuk website. Sehingga hasilnya disimpan kedalam file sav.

F. Deployment

Kemudian setelah data melalui tahap evaluasi, peneliti telah menyusun tampilan sebagai bentuk presentasi dari hasil evaluasi pada proses data mining. Berikut adalah tampilan akhir dari model yang telah disimpan :

a. Tab Penjualan

Tab ini menampilkan data hasil penjualan berbagai produk dari reseller Toko Premium Shopping pada tahun 2021.

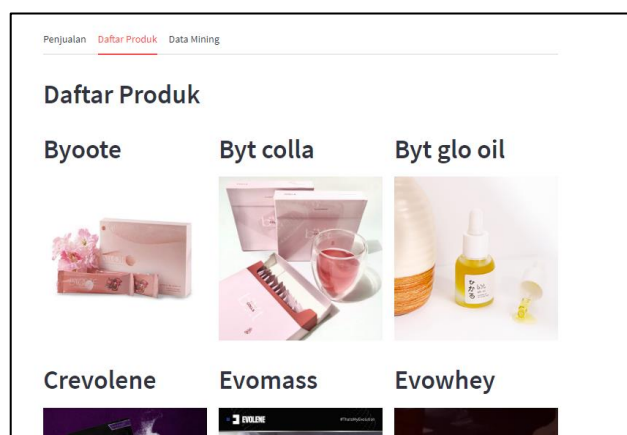


Produk	Jumlah
Evomass	9,105 Box
Flimty	20,206 Box
Byoote	2,176 Box
Byt colla	2,166 Box
Byt glo oil	916 Box
Crevolene	1,604 Pcs
Evowhey	5,326 Box
Flimbar	916 Box
Flimeal	3,466 Box
Herlean	1,506 Box

Gambar 4. 10 Tab Penjualan

b. Tab Daftar Produk

Dari tab ini telah disajikan berbagai gambar dari produk yang dijual oleh reseller Toko Premeium Shopping.



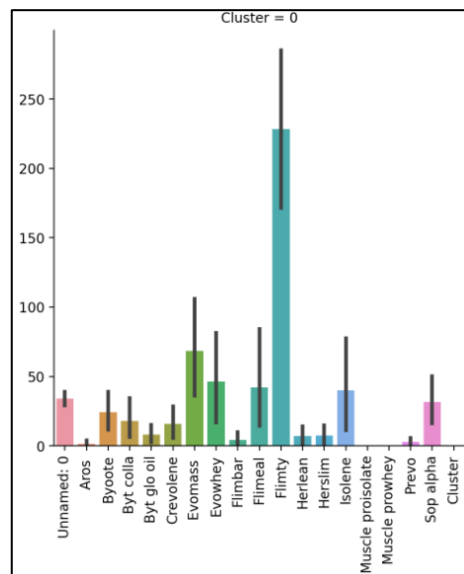
Gambar 4. 11 Tab Daftar Produk

c. Tab Data Mining

Pada Tab ini user dapat mengisi jumlah penjualan reseller sesuai form yang sudah tersedia untuk kemudian dianalisis apakah data tersebut termasuk ke dalam cluster 0,1, atau 2. Berikut ada form yang harus diisi oleh user :

Gambar 4. 12 Tab Data Mining

Kemudian setelah dianalisis akan muncul diagram dibawah form yang menunjukkan hasil clustering dan diagram penjualan produk dari reseller tersebut.



Gambar 4. 13 Diagram Hasil Clustering

BAB V

SIMPULAN DAN SARAN

A. Kesimpulan

Berdasarkan analisis data yang telah dilakukan menggunakan data mining terhadap dataset yang telah ditentukan, maka dapat ditarik kesimpulan sebagai berikut :

1. Setelah dilakukan pengelompokan data reseller melalui rekapitulasi hasil penjualan reseller Toko Premium Shopping berdasarkan potensi penjualan menggunakan K-Means clustering terbentuk tiga cluster yaitu cluster 0 dengan jumlah reseller sebanyak 54 reseller, cluster 1 dengan jumlah 11 reseller dan cluster 2 dengan 1 reseller
2. Strategi promosi bagi calon reseller baru yang tepat sasaran sesuai wilayah reseller berdasarkan cluster yang terbentuk adalah dengan melakukan promosi untuk meningkatkan jumlah penjualan pada cluster 0 dan cluster 1, sedangkan pada cluster 2 perlu dilakukan proses promosi untuk memperbanyak reseller potensial.

B. Saran

Untuk memperbaiki kekurangan penelitian ini dapat diusulkan beberapa perbaikan sebagai berikut :

1. Pengelompokan data reseller pada Toko Premium Shopping sebaiknya dilakukan secara rutin setiap tahun untuk menghindari turunnya tingkat penjualan dan memperluas jangkauan konsumen pada tahun berikutnya.
2. Penelitian ini dapat dijadikan sebagai salah satu referensi bagi pihak Toko Premium shopping untuk mengambil keputusan dalam melakukan strategi promosi bagi calon reseller baru.

DAFTAR PUSTAKA

- [1] S. I, Toko Online Professional dengan Blogger dan Wordpress, Jakarta: Elex Media Komputindo, 2014.
- [2] I. Tahyudin, Data Mining Dan Data Warehouse Menggunakan Aplikasi KNIME, Banyumas: Zahira Media Publisher, 2021.
- [3] P. E, DATA MINING Pengelompokan Data untuk Informasi dan Evaluasi, Duta Media Publishing, 2017.
- [4] A. M. Siregar, DATA MINING : Pengolahan Data Menjadi Informasi dengan RapidMiner, CV Kekata Group, 2020.
- [5] L. Muflikhah, Data Mining, Universitas Brawijaya Press, 2018.
- [6] E. Irwansyah, Advanced Clustering Teori dan Aplikasi, DeePublish, 2015.
- [7] M. Wahyudi, Data Mining: Penerapan Algoritma K-Means Clustering dan K-Medoids Clustering, Yayasan Kita Menulis, 2020.
- [8] W. A, Data Mining : Algoritma dan Implementasi, Yayasan Kita Menulis, 2020.
- [9] M. F. AZIS, Object Oriented Programming Php 5, Jakarta: Elex Media Komputindo, 2005.
- [10] P. H, Pemodelan dan Simulasi untuk Pengelolaan Adaptif Sumber Daya Alam dan Lingkungan, Bogor: PT Penerbit IPB Press, 2019.
- [11] P. D. S. Mulyani, Analisis dan Perancangan Sistem Informasi Manajemen Keuangan Daerah, Bandung: Abdi Sistematika, 2016.
- [12] A. S, Prosedur Penelitian, Suatu Pendekatan Praktik, Jakarta: Rinneka Cipta, 2006.
- [13] U. Sumarwan, Metode riset bisnis dan konsumen edisi revisi, PT Penerbit IPB Press, 2018.
- [14] C. Shearer, The CRISP-DM Model: The New Blueprint for DataMining. JOURNAL of Data Warehousing, Vol. 5, No. 4, p. 13-22, 2000.

LAMPIRAN

Lampiran 1 Tabel Hasil Clustering

NO	NAMA RESELLER	Aros	Byoote	Byt colla	Byt glo oil	Crevolene	Evomass	Evowhey	Flimbar	Flimeal	Flimty	Herlean	Herslim	Isolene	Muscle nroisolate	Muscle nrowhev	Prevo	Sop alpha	Cluster
1	ajudansyifa	0	124	5	126	0	316	148	0	248	351	93	122	5	0	0	65	139	1
2	antonwarman1	6	0	27	27	0	2137	1100	48	0	1600	61	147	299	10	33	51	184	1
3	ARTHAMA97	0	0	0	0	0	67	0	0	0	0	0	0	32	0	0	0	0	0
4	bluegreen12	0	0	0	0	0	0	0	0	0	94	0	0	0	0	0	0	0	0
5	boworamadhan44	7	0	57	0	30	910	48	0	35	1209	628	496	0	0	7	4	183	1
6	chocofun13	0	0	10	0	0	0	0	0	0	195	0	0	0	0	0	0	0	0
7	citrakumala22	0	25	0	0	0	0	0	0	0	45	0	0	0	0	0	0	0	0
8	clarabell48	0	0	0	0	0	0	0	0	0	56	0	0	12	0	0	0	0	0
9	clayrizha	0	0	0	0	0	0	0	0	0	112	0	0	0	0	0	0	0	0
10	cokomelon	0	0	0	0	0	79	22	0	15	113	0	0	0	0	0	0	0	0
11	danisgina	0	107	106	103	124	420	110	0	101	599	0	105	0	0	0	0	197	1
12	danitadara	0	0	0	0	0	0	0	0	0	119	0	0	0	0	0	0	0	0
13	darengsown	0	0	0	0	186	188	561	0	173	753	0	0	373	0	0	0	187	1
14	daylinestore	0	0	2	0	0	0	0	0	0	569	0	0	0	0	0	0	0	0
15	distyanda	0	0	0	0	0	0	0	0	0	83	0	0	0	0	0	0	0	0
16	elinanoer	31	523	681	210	375	789	788	605	216	660	218	270	659	0	0	0	624	2

NO	NAMA RESELLER	Aros	Byoote	Byt colla	Byt glo oil	Crevolene	Evomass	Evowhey	Flimbar	Flimeal	Flimty	Herlean	Herslim	Isolene	Muscle proisolate	Muscle nrowhev	Prevo	Sop alpha	Cluster	
17	emil2435	0	0	0	0	0	0	0	0	0	189	0	0	0	0	0	0	0	0	0
18	faradella87	0	0	0	0	0	0	35	0	0	9	0	0	0	0	0	0	0	0	0
19	femmitaninda	26	82	26	0	0	21	42	0	0	513	21	0	163	0	0	24	49	0	0
20	gondarina12	0	0	116	0	0	0	0	0	0	395	0	0	0	0	0	0	0	0	0
21	honeymelon11	0	0	0	0	0	0	0	0	0	341	0	0	0	0	0	0	0	0	0
22	jellyrainbow11	0	0	0	0	0	0	0	0	0	151	9	0	0	0	0	0	0	0	0
23	jemimagina	0	138	276	140	150	574	280	0	642	1958	142	144	0	0	22	0	709	1	0
24	Jikoshop09	0	0	0	0	0	0	0	38	0	105	0	0	0	0	0	0	0	0	0
25	jionda11	0	0	0	0	0	0	0	0	0	60	0	0	0	0	0	0	0	0	0
26	kiranaani12	0	0	0	32	0	0	0	0	0	76	0	0	0	0	0	0	0	0	0
27	kristashellania	32	35	28	38	72	520	284	0	0	841	31	53	0	32	54	21	403	1	0
28	lionashim	0	0	0	0	0	0	0	0	0	51	0	0	0	0	0	0	0	0	0
29	lyodrinefine	0	0	0	0	0	188	188	0	0	430	0	0	240	0	0	0	0	0	0
30	malikadwin	0	0	0	0	34	0	0	0	0	37	0	0	0	0	0	0	0	0	0
31	mangosticky77	0	0	0	0	0	0	0	0	0	221	0	0	0	0	0	0	0	0	0
32	melisa.putri7	0	41	0	52	34	96	0	0	53	535	38	54	0	0	0	0	63	0	0
33	miskashow	0	0	94	0	0	0	0	0	0	366	0	0	0	0	0	0	0	0	0
34	Monicahantika	0	0	12	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0	0
35	myvirashell	0	0	0	0	0	0	0	0	0	133	0	0	0	0	0	0	0	0	0
36	nauraindah55	0	0	13	0	0	0	0	0	0	45	0	0	0	0	0	0	0	0	0
37	Nitaliya45	0	0	0	0	0	0	0	0	0	242	0	0	0	0	0	0	0	0	0

NO	NAMA RESELLER	Aros	Byoote	Byt colla	Byt glo oil	Crevolene	Evomass	Evowhey	Flimbar	Flimeal	Flimty	Herlean	Herslim	Isolene	Muscle proisolate	Muscle nrowhev	Prevo	Sop alpha	Cluster	
38	nucadwianda	0	0	0	0	12	0	0	0	0	168	0	0	0	0	0	0	0	0	0
39	okeyshopping12	0	144	54	0	160	561	613	0	0	701	22	0	473	0	0	0	0	0	1
40	relaxday66	0	229	383	0	194	428	623	0	0	892	178	0	797	0	0	0	0	0	1
41	revindersell	0	0	0	0	0	0	0	0	0	149	0	0	0	0	0	0	18	0	0
42	rismalinda11	0	0	0	0	0	0	0	0	0	26	0	0	0	0	0	0	0	0	0
43	ristaqueen	0	0	0	0	0	0	0	112	0	176	0	0	0	0	0	0	0	0	0
44	rosalindaa78	0	0	0	0	0	0	0	0	95	18	0	0	0	0	0	0	0	0	0
45	rownbean	0	0	0	0	0	334	0	0	0	0	0	0	0	0	0	0	0	0	0
46	safiraoktaa	65	59	0	35	68	246	58	0	11	349	43	0	0	0	0	21	198	1	
47	Sarinieta	0	0	0	0	0	0	0	0	0	93	0	0	0	0	0	0	0	0	0
48	seablue00	0	0	0	0	0	0	0	0	1134	137	0	0	0	0	0	0	0	0	0
49	setyoaji33	1	57	0	29	27	435	143	0	68	793	31	55	79	0	0	46	156	1	
50	shienaputri45	0	0	0	0	0	104	0	0	0	0	0	0	0	0	0	0	0	43	0
51	shoppingday13	0	0	0	0	0	112	0	0	0	0	0	0	0	0	0	0	0	0	0
52	showday11	0	0	0	0	0	0	0	0	0	390	0	0	0	0	0	0	0	0	0
53	simbalion22	0	0	29	0	0	0	0	0	0	396	0	0	0	0	0	0	0	0	0
54	siskazien	0	178	0	0	0	0	0	0	43	0	0	0	0	0	0	0	0	119	0
55	soopinthe	0	0	21	0	0	0	0	0	0	193	0	0	0	0	0	0	0	0	0
56	stopmotion33	0	0	0	0	0	0	0	0	0	266	0	0	0	0	0	0	0	0	0
57	tamaninda	0	255	159	0	127	440	97	0	251	881	0	0	232	0	0	0	0	287	1
58	atasquirel	0	9	18	9	0	20	17	0	16	45	0	0	36	0	0	0	0	18	0

NO	NAMA RESELLER	Aros	Byoote	Byt colla	Byt glo oil	Crevolene	Evomass	Evowhey	Flimbar	Flimeal	Flimty	Herlean	Herslim	Isolene	Muscle proisolate	Muscle nrowhev	Prevo	Sop alpha	Cluster	
59	tigershop17	0	105	0	0	0	0	0	0	48	189	0	0	0	0	0	0	0	41	0
60	todayshop22	0	65	0	0	0	0	0	0	0	177	0	0	0	0	0	0	0	98	0
61	turtledown_	0	0	31	115	0	120	124	0	190	361	0	113	0	0	0	0	0	241	0
62	unicorninshell	0	0	0	0	0	0	0	113	0	24	0	0	0	0	0	0	0	101	0
63	wendycandy22	0	0	0	0	11	0	0	0	0	209	0	0	0	0	0	0	0	0	0
64	yindadinda	0	0	0	0	0	0	0	0	13	210	0	0	0	0	0	0	0	0	0
65	zahraamanda13	0	0	18	0	0	0	0	0	0	92	0	0	0	0	0	0	0	0	0
66	zanakirana	0	0	0	0	0	0	45	0	114	0	0	0	0	0	0	0	0	0	0

Lampiran 2 Tabel Hasil Cluster 0

NO	NAMA RESELLER	Aros	Byoote	Byt colla	Byt glo oil	Crevolene	Evomass	Evowhey	Flimbar	Flimeal	Flimty	Herlean	Herslim	Isolene	Muscle proisolate	Muscle nrowhev	Prevo	Sop alpha	Cluster	
1	ARTHAMA97	0	0	0	0	0	67	0	0	0	0	0	0	32	0	0	0	0	0	0
2	bluegreen12	0	0	0	0	0	0	0	0	0	94	0	0	0	0	0	0	0	0	0
3	chocofun13	0	0	10	0	0	0	0	0	0	195	0	0	0	0	0	0	0	0	0
4	citrakumala22	0	25	0	0	0	0	0	0	0	45	0	0	0	0	0	0	0	0	0
5	clarabell48	0	0	0	0	0	0	0	0	0	56	0	0	12	0	0	0	0	0	0
6	clayrizha	0	0	0	0	0	0	0	0	0	112	0	0	0	0	0	0	0	0	0
7	cokomelon	0	0	0	0	0	79	22	0	15	113	0	0	0	0	0	0	0	0	0
8	danitadara	0	0	0	0	0	0	0	0	0	119	0	0	0	0	0	0	0	0	0
9	daylinestore	0	0	2	0	0	0	0	0	0	569	0	0	0	0	0	0	0	0	0
10	distyanda	0	0	0	0	0	0	0	0	0	83	0	0	0	0	0	0	0	0	0
11	emil2435	0	0	0	0	0	0	0	0	0	189	0	0	0	0	0	0	0	0	0
12	faradella87	0	0	0	0	0	0	35	0	0	9	0	0	0	0	0	0	0	0	0
13	femmitaninda	26	82	26	0	0	21	42	0	0	513	21	0	163	0	0	24	49	0	0
14	gondarina12	0	0	116	0	0	0	0	0	0	395	0	0	0	0	0	0	0	0	0
15	honeymelon11	0	0	0	0	0	0	0	0	0	341	0	0	0	0	0	0	0	0	0
16	jellyrainbow11	0	0	0	0	0	0	0	0	0	151	9	0	0	0	0	0	0	0	0
17	Jikoshop09	0	0	0	0	0	0	0	38	0	105	0	0	0	0	0	0	0	0	0
18	jionda11	0	0	0	0	0	0	0	0	0	60	0	0	0	0	0	0	0	0	0
19	kiranaani12	0	0	0	32	0	0	0	0	0	76	0	0	0	0	0	0	0	0	0
20	lionashim	0	0	0	0	0	0	0	0	0	51	0	0	0	0	0	0	0	0	0

NO	NAMA RESELLER	Aros	Byoote	Byt colla	Byt glo oil	Crevolene	Evomass	Evowhey	Flimbar	Flimeal	Flimty	Herlean	Herslim	Isolene	Muscle proisolate	Muscle nrowhev	Prevo	Sop alpha	Cluster
21	lyodrinefine	0	0	0	0	0	188	188	0	0	430	0	0	240	0	0	0	0	0
22	malikadwin	0	0	0	0	34	0	0	0	0	37	0	0	0	0	0	0	0	0
23	mangosticky77	0	0	0	0	0	0	0	0	0	221	0	0	0	0	0	0	0	0
24	melisa.putri7	0	41	0	52	34	96	0	0	53	535	38	54	0	0	0	0	63	0
25	miskashow	0	0	94	0	0	0	0	0	0	366	0	0	0	0	0	0	0	0
26	Monicahantika	0	0	12	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0
27	myvirashell	0	0	0	0	0	0	0	0	0	133	0	0	0	0	0	0	0	0
28	nauraindah55	0	0	13	0	0	0	0	0	0	45	0	0	0	0	0	0	0	0
29	Nitaliya45	0	0	0	0	0	0	0	0	0	242	0	0	0	0	0	0	0	0
30	nucadwianda	0	0	0	0	12	0	0	0	0	168	0	0	0	0	0	0	0	0
31	revindersell	0	0	0	0	0	0	0	0	0	149	0	0	0	0	0	18	0	0
32	rismalinda11	0	0	0	0	0	0	0	0	0	26	0	0	0	0	0	0	0	0
33	ristaqueen	0	0	0	0	0	0	0	112	0	176	0	0	0	0	0	0	0	0
34	rosalindaa78	0	0	0	0	0	0	0	0	95	18	0	0	0	0	0	0	0	0
35	rownbean	0	0	0	0	0	334	0	0	0	0	0	0	0	0	0	0	0	0
36	Sarinieta	0	0	0	0	0	0	0	0	0	93	0	0	0	0	0	0	0	0
37	seablue00	0	0	0	0	0	0	0	0	1134	137	0	0	0	0	0	0	0	0
38	shienaputri45	0	0	0	0	0	104	0	0	0	0	0	0	0	0	0	0	43	0
39	shoppingday13	0	0	0	0	0	112	0	0	0	0	0	0	0	0	0	0	0	0
40	showday11	0	0	0	0	0	0	0	0	0	390	0	0	0	0	0	0	0	0
41	simbalion22	0	0	29	0	0	0	0	0	0	396	0	0	0	0	0	0	0	0

NO	NAMA RESELLER	Aros	Byoote	Byt colla	Byt glo oil	Crevolene	Evomass	Evowhey	Flimbar	Flimeal	Flimty	Herlean	Herslim	Isolene	Muscle proisolate	Muscle nrowhev	Prevo	Sop alpha	Cluster
42	siskazien	0	178	0	0	0	0	0	0	43	0	0	0	0	0	0	0	119	0
43	soopinthe	0	0	21	0	0	0	0	0	0	193	0	0	0	0	0	0	0	0
44	stopmotion33	0	0	0	0	0	0	0	0	0	266	0	0	0	0	0	0	0	0
45	tatasquirel	0	9	18	9	0	20	17	0	16	45	0	0	36	0	0	0	18	0
46	tigershop17	0	105	0	0	0	0	0	0	48	189	0	0	0	0	0	0	41	0
47	todayshop22	0	65	0	0	0	0	0	0	0	177	0	0	0	0	0	0	98	0
48	turtledown_	0	0	31	115	0	120	124	0	190	361	0	113	0	0	0	0	241	0
49	unicorninshell	0	0	0	0	0	0	0	113	0	24	0	0	0	0	0	0	101	0
50	wendycandy22	0	0	0	0	11	0	0	0	0	209	0	0	0	0	0	0	0	0
51	yindadinda	0	0	0	0	0	0	0	0	13	210	0	0	0	0	0	0	0	0
52	zahraamanda13	0	0	18	0	0	0	0	0	0	92	0	0	0	0	0	0	0	0
53	zanakirana	0	0	0	0	0	0	45	0	114	0	0	0	0	0	0	0	0	0

Lampiran 3 Tabel Hasil Cluster 1

NO	NAMA RESELLER	Aros	Byoote	Byt colla	Byt glo oil	Crevolene	Evomass	Evowhey	Flimbar	Flimeal	Flimty	Herlean	Herslim	Isolene	Muscle nroisolate	Muscle nrowhev	Prevo	Sop alpha	Cluster
1	ajudansyifa	0	124	5	126	0	316	148	0	248	351	93	122	5	0	0	65	139	1
2	antonwarman1	6	0	27	27	0	2137	1100	48	0	1600	61	147	299	10	33	51	184	1
3	boworamadhan44	7	0	57	0	30	910	48	0	35	1209	628	496	0	0	7	4	183	1
4	danisgina	0	107	106	103	124	420	110	0	101	599	0	105	0	0	0	0	197	1
5	darengown	0	0	0	0	186	188	561	0	173	753	0	0	373	0	0	0	187	1
6	jemimagina	0	138	276	140	150	574	280	0	642	1958	142	144	0	0	22	0	709	1
7	kristashellania	32	35	28	38	72	520	284	0	0	841	31	53	0	32	54	21	403	1
8	okeyshopping12	0	144	54	0	160	561	613	0	0	701	22	0	473	0	0	0	0	1
9	relaxday66	0	229	383	0	194	428	623	0	0	892	178	0	797	0	0	0	0	1
10	safiraoktaa	65	59	0	35	68	246	58	0	11	349	43	0	0	0	0	21	198	1
11	setyoaji33	1	57	0	29	27	435	143	0	68	793	31	55	79	0	0	46	156	1
12	tamaninda	0	255	159	0	127	440	97	0	251	881	0	0	232	0	0	0	287	1

Lampiran 4 Tabel Hasil Cluster 2

No	NAMA RESELLER	Aros	Byoote	Byt colla	Byt glo oil	Crevolene	Evomass	Evowhey	Flimbar	Flimeal	Flimty	Herlean	Herslim	Isolene	Muscle proisolate	Muscle nrowhey	Prevo	Sop alpha	Cluster
1	elinanoer	31	523	681	210	375	789	788	605	216	660	218	270	659	0	0	0	624	2

Lampiran 5 Tabel Hasil Clustering Reseller

NO	NAMA RESELLER	LOKASI TOKO	Cluster
1	ajudansyifa	JOGJA	1
2	antonwarman1	JAKARTA SELATAN	1
3	ARTHAMA97	SURABAYA	0
4	bluegreen12	SEMARANG	0
5	boworamadhan44	BANDUNG	1
6	chocofun13	MEDAN	0
7	citrakumala22	SOLO	0
8	clarabell48	JAKARTA	0
9	clayrizha	JAKARTA	0
10	cokomelon	SEMARANG	0
11	danisgina	JAKARTA	1
12	danitadara	SEMARANG	0
13	darengsown	SEMARANG	1
14	daylinestore	MEDAN	0
15	distyanda	MEDAN	0
16	elinanoer	SEMARANG	2
17	emil2435	JAKARTA BARAT	0
18	faradella87	MEDAN	0
19	femmitaninda	JAKARTA	0
20	gondarina12	JAKARTA	0
21	honeymelon11	JAKARTA	0
22	jellyrainbow11	SEMARANG	0
23	jemimagina	SURABAYA	1
24	Jikoshop09	SEMARANG	0
25	jionda11	JOGJA	0
26	kiranaani12	SURABAYA	0
27	kristashellania	SURABAYA	1
28	lionashim	SEMARANG	0
29	lyodrinefine	JAKARTA	0
30	malikadwin	JAKARTA	0
31	mangosticky77	SEMARANG	0
32	melisa.putri7	PALEMBANG	0
33	miskashow	SEMARANG	0
34	Monicahantika	JAKARTA	0
35	myvirashell	SEMARANG	0
36	nauraindah55	BANDUNG	0
37	Nitaliya45	MEDAN	0

NO	NAMA RESELLER	LOKASI TOKO	Cluster
38	nucadwianda	SEMARANG	0
39	okeyshopping12	JAKARTA	1
40	relaxday66	SEMARANG	1
41	revindersell	SEMARANG	0
42	rismalinda11	JAKARTA	0
43	ristaqueen	SEMARANG	0
44	rosalindaa78	SOLO	0
45	rownbean	SEMARANG	0
46	safiraoktaa	JAKARTA PUSAT	1
47	Sarinieta	JAKARTA	0
48	seablue00	SEMARANG	0
49	setyoaji33	SOLO	1
50	shienaputri45	BANDUNG	0
51	shoppingday13	SURABAYA	0
52	showday11	SEMARANG	0
53	simbalion22	SEMARANG	0
54	siskazien	JAKARTA	0
55	soopinthe	JAKARTA	0
56	stopmotion33	SEMARANG	0
57	tamaninda	JAKARTA	1
58	tatasquirel	MEDAN	0
59	tigershop17	JAKARTA	0
60	todayshop22	MEDAN	0
61	turtledown_	JAKARTA	0
62	unicorninshell	SEMARANG	0
63	wendycandy22	JAKARTA	0
64	yindadinda	JAKARTA	0
65	zahraamanda13	MEDAN	0
66	zanakirana	SEMARANG	0



UNIVERSITAS PGRI SEMARANG
FAKULTAS TEKNIK DAN INFORMATIKA

Kampus : Jalan Sidodadi Timur Nomor 24 Dr. Cipto, Semarang – Indonesia 50125

Telp. (024) 8316377, Faks. (024) 8448217, E-mail : upgrismg@gmail.com, Homepage : www.upgrismg.ac.id

LEMBAR PEMBIMBINGAN SKRIPSI

Nama Mahasiswa : ASTRI NUR KHUSNA
 N P M : 18670069
 Program Studi : INFORMATIKA
 Judul Skripsi : PENGELompokAN RESELLER POTENSIAL
 MENGGUNAKAN ALGORITMA CLUSTERING
 K-MEANS PADA TOKO PREMIUM
 SHOPPING
 Dosen Pembimbing I : ARIS TRI JAKA HARJANTA , M. Kom.
 Dosen Pembimbing II : KHOIRIYA LATIFAH , M. Kom.

No.	Hari, tanggal	Uraian Bimbingan	Paraf
	19 Jun 22	Proposal Fee	<i>[Signature]</i>
	26 Juli 22	Coba Algoritma k-means dan python & RML	<i>[Signature]</i>
	28 Juli '22	REV BAB I	<i>[Signature]</i>
	1 Agst '22	Rev BAB II	<i>[Signature]</i>
	3 Agst '22	ACC BAB I & II	<i>[Signature]</i>
	9 Agst '22	Rev BAB III	<i>[Signature]</i>
	10 Agst '22	ACC BAB III Rev Bab IV	<i>[Signature]</i>
	12 Agst '22	ACC BAB IV	<i>[Signature]</i>
	15 Agustus 22	ACC Uraian	<i>[Signature]</i>

Dosen Pembimbing I,

[Signature]

NIP/NPP 148201443

Mahasiswa,

[Signature]

NPM 18670069



UNIVERSITAS PGRI SEMARANG
FAKULTAS TEKNIK DAN INFORMATIKA

Kampus : Jalan Sidodadi Timur Nomor 24 Dr. Cipto, Semarang – Indonesia 50125

Telp. (024) 8316377, Faks. (024) 8448217, E-mail : upgrismg@gmail.com, Homepage : www.upgrismg.ac.id

LEMBAR PEMBIMBINGAN SKRIPSI

Nama Mahasiswa : ASTRI NUR KHUSNA
N P M : 18670069
Program Studi : INFORMATIKA
Judul Skripsi : PENGELompokkan RESELLER POTENSIAL
MENGGUNAKAN ALGORITMA CLUSTERING
K-MEANS PADA TOKO PREMIUM SHOPPING
Dosen Pembimbing I : ARIS TRI JAKA HARJANTA , M. Kom.
Dosen Pembimbing II : KHOIRIYA LATIFAH , M. Kom.

No.	Hari, tanggal	Uraian Bimbingan	Paraf
1.	19 Jan '22	ACC Proposal	
2.	26 Juli '22	Bimbingan BAB I & Latar belakang	
3.	28 Juli '22	Coba algoritma k-means & python	
4.	1. Agst '22	Rev BAB II	
5.	3 Agst '22	Rev BAB III	
6.	9 Agst '22	ACC BAB I - III	
7.	10 Agst '22	Rev BAB IV	
8.	12 Agst '22	ACC BAB IV	
9.	15 Agst '22	ACC Ujian	

Dosen Pembimbing I

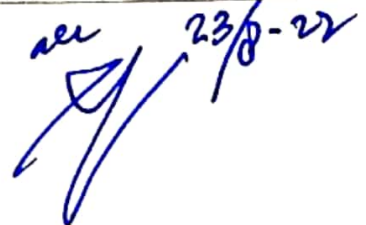
NIP/NPP 147801434

Mahasiswa,

NPM 18670069

LEMBAR REVISI UJIAN SKRIPSI

Nama Mahasiswa : Astri Nur Khusna
 N P M : 18670064
 Judul : PENGELOMPOKAN RESELLER POTENSIAL MENGGUNAKAN ALGORITMA CLUSTERING K-MEANS PADA TOKO PREMIUM SHOPPING

No	Uraian Revisi	Keterangan
1	Perbaiki Abstrak ✓	 23/8-22
2	Perbaiki No. Halaman	
3	Lampirkan Google Docs lengkap ✓	
4	Langkah? / Metode Sampling ✓	
5	Pembahasan lengkap dan Cluster / Kriteria yg sesuai ✓	
6	Penjelasan algoritma K-Means lengkap ✓	
7	Jelaskan y apa itu Gerdan ✓ mengevaluasi hasil olah data	
8	Tampilkan dan Jelas Reseller? sesuai ✓ Cluster nya	
9	Konsistensi font ✓	

10. Cek Ciptaan. ✓

Pengesahan Penguji I



Bambang Agus H. S. Kom, M. Kom
 NIP/NPP. 148201433

✳ Bawa Manuskrip y/ revisi!

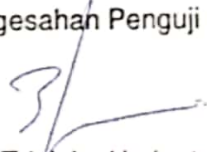
*) Revisi Maksimal 7 Hari Setelah Pelaksanaan Ujian Skripsi

LEMBAR REVISI UJIAN SKRIPSI

Nama Mahasiswa : Astri Nur Khusna
 N P M : 18670064
 Judul : PENGELOMPOKAN RESELLER POTENSIAL MENGGUNAKAN ALGORITMA CLUSTERING K-MEANS PADA TOKO PREMIUM SHOPPING

No	Uraian Revisi	Keterangan
1.	Abstract.	3/
2.	Fungsi waterfall ?	3/
3.	- tambah metode pemodelan k-means / data mining	3/
4.	penjelasan hasil proses	3/
		<p>10 3/</p> <p>2</p> <p>23/27</p> <p>8</p>

Pengesahan Penguji II


 Aris Tri Jaka Harjanta S.Kom., M.Kom
 NIP/NPP. 148201443

*) Revisi Maksimal 7 Hari Setelah Pelaksanaan Ujian Skripsi

LEMBAR REVISI UJIAN SKRIPSI

Nama Mahasiswa : Astri Nur Khusna
N P M : 18670064
Judul : PENGELOMPOKAN RESELLER POTENSIAL MENGGUNAKAN ALGORITMA CLUSTERING K-MEANS PADA TOKO PREMIUM SHOPPING

No	Uraian Revisi	Keterangan
1.	Di bab II tambahkan ttg K-MEANS & contoh perhitungannya manualnya	
2.	Pre processing nya lha detail lagi	
3.	Clustering o tgg kriteria apa 1, 2 spt apa	
4.	Filtering PCA & perbaikan reduksi dimensi nya	
5.	Validasi model sesuai kan	

Pengesahan Penguji III

Khoiriya Latifah, S. Kom, M. Kom
NIP/NPP. 147801434

*) Revisi Maksimal 7 Hari Setelah Pelaksanaan Ujian Skripsi