

IMPLEMENTASI K-MEANS UNTUK MELAKUKAN SEGMENTASI PRODUK BERDASARKAN DATA TRANSAKSI RETAIL

Fatia A. Maresti¹, Woro I. Rahayu², Maria B. B. C. Lustin³, Thiodas H. Pakpahan⁴

Universitas Logistik dan Bisnis Internasional

Jl. Sariosih No.54, Sarijadi, Kec. Sukasari, Kota Bandung, Jawa Barat 40151

Email: ¹fatiaamalia@ulbi.ac.id, ²woroisti@ulbi.ac.id, ³cintanylstn@gmail.com,

⁴thiodaspakpahan111@gmail.com

Abstract

Retail companies are one form of business that has experienced significant development. Significant development can lead to many business competitors in the same field. Therefore, every company must be able to find strategies to get the attention of the target market. Fast service and good prices are included in the factors that influence customer interest in making transactions in the retail business. This research aims to classify products into several categories. The method used to perform segmentation is the data mining method with the K-Means algorithm and the elbow method to find the optimal number of clusters. The obtained results of many optimal clusters are 3, namely cluster 0, cluster 1, and cluster 2. The variable that has the most influence on cluster characteristics is the price variable, while the variable that has little influence on cluster characteristics is the quantity variable.

Keywords: Elbow Method, K-Means, Product Segmentation, Retail Company

PENDAHULUAN

Sebuah perusahaan dapat berjalan dengan baik apabila proses bisnisnya menghasilkan keuntungan. Untuk mendapatkan keuntungan maksimal setiap perusahaan menggunakan cara agar dapat bertahan di antara persaingan bisnis yang cukup sengit. Perusahaan ritel merupakan salah satu bentuk bisnis yang mengalami perkembangan cukup signifikan (Ramadhanty, 2023). Perkembangan yang cepat juga memunculkan banyak pesaing bisnis yang mengharuskan perusahaan ritel mampu bertahan di industri yang memiliki banyak pesaing dan cenderung dinamis (Nisa et al., 2023). Salah satu permasalahan yang dihadapi oleh perusahaan ritel adalah tidak seimbangnya antara persediaan produk dengan kebutuhan pelanggan. Oleh sebab itu, diperlukan adanya strategi untuk meningkatkan keuntungan perusahaan ritel dengan menyesuaikan penawaran produk sesuai dengan pola perilaku pelanggan. Salah satu bentuk strategi bisnisnya adalah dengan melakukan pengelompokan produk ke dalam beberapa kategori sesuai dengan pola perilaku pelanggan.

Segmentasi merupakan proses untuk mengetahui perilaku pelanggan dan memutuskan strategi pemasaran yang efektif dan tepat sasaran sehingga menghasilkan keuntungan yang diharapkan oleh pemilik bisnis (Harani et al., 2020). Untuk melakukan proses pengelompokan dibutuhkan proses *clustering* yang membagi data ke dalam beberapa bagian yang masing-masing memiliki kemiripan. K-Means merupakan salah satu algoritma yang digunakan untuk membagi data ke dalam beberapa partisi berdasarkan adanya kesamaan karakteristik. K-Means adalah salah satu algoritma klasterisasi yang populer dan efisien untuk mengelompokkan data ke dalam beberapa klaster berdasarkan kesamaan karakteristik.

Implementasi K-Means Untuk Melakukan Segmentasi Produk Berdasarkan Data Transaksi Retail

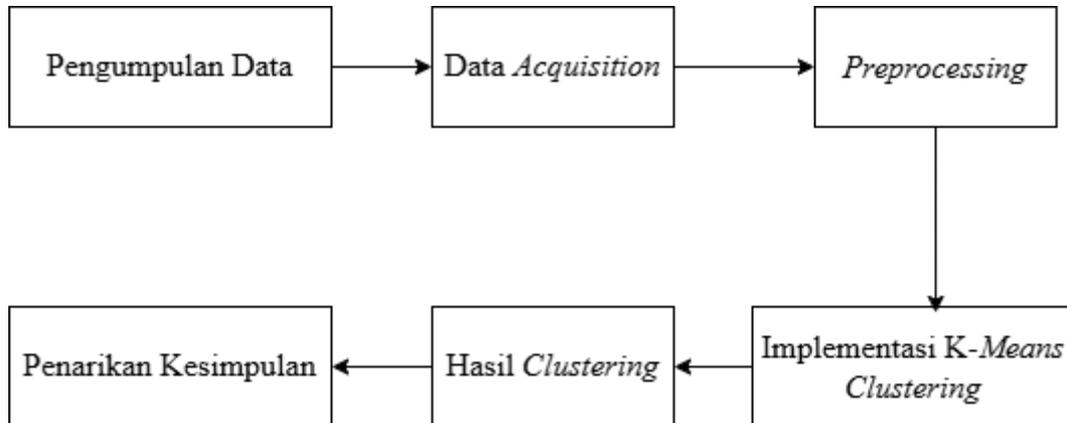
Penelitian yang dilakukan Kartika Zahretta Wijaya, Arif Djunaidy dan Faizal Mahananto menggunakan algoritma K-Means untuk menganalisis data 1443 pelanggan yang melakukan transaksi pada periode pengambilan data berhasil mendapatkan 4 cluster menggunakan model RFM yaitu pelanggan *superstar*, pelanggan *everyday*, pelanggan *occasional* dan pelanggan *dormant* (Wijaya et al., 2021). Pada penelitian berjudul Optimalisasi Strategi Pemasaran dengan Segmentasi Pelanggan Menggunakan Penerapan K-Means Clustering pada Transaksi *Online Retail*, dilakukan segmentasi pelanggan menggunakan data transaksi online retail berjumlah 54910. Penelitian ini menggunakan algoritma K-Means dan berhasil mendapatkan 4 kluster yaitu kluster 0 dengan rata-rata kuantitas produk 7.20 dan rata-rata harga satuan 1.70, kluster 1 dengan rata-rata kuantitas produk 5.30 dan rata-rata harga satuan 1.70, kluster 2 dengan rata-rata kuantitas produk 23.00 dan rata-rata harga satuan 1.10, kluster 3 dengan rata-rata kuantitas produk 3.60 dan rata-rata harga satuan 14.8 (Awalina & Woro Isti Rahayu, 2023). Selanjutnya, Muhammad Hilman Fakhriza dan Khaerul Umam melakukan penelitian Analisis Produk Terlaris Menggunakan Metode K-Means Clustering Pada PT. Sukanda Djaya untuk mengetahui produk paling laris di PT. Sukanda Djaya yang menghasilkan 3 *cluster* dengan *cluster* 2 sebagai produk paling laris dengan nilai *cluster* 86,5 (Fakhriza & M. Khaerul Umam, 2021).

Pada penelitian lainnya yang berjudul K-Means Clustering Untuk Segmentasi Produk Berdasarkan Analisis RFM pada Data Transaksi Penjualan melakukan segmentasi produk menggunakan metode K-Means. Pada proses segmentasi, dilakukan 5 kali proses *clustering* dengan nilai k yang berbeda, kemudian didapatkan nilai *silhouette* tertinggi yaitu 0,4314 pada *clustering* ke 2. Hasil *clustering* mendapatkan 3 kluster yaitu produk dengan nilai R, F, M paling baik (C1) berjumlah 8 produk, produk dengan nilai R, F baik namun M yang buruk (C2) berjumlah 18 produk, produk dengan nilai R, F, M cukup baik (C3) berjumlah 18 produk (Fithriyah et al., 2021). Penelitian yang dilakukan Heri Yansah, Tina Tri Wulansari, Faza Alameka pada Penerapan Algoritma K-Means Dalam Clustering Produk Terlaris Pada Fr Parfum menggunakan 1655 data penjualan parfum pada tahun 2021. Penelitian ini menghasilkan 3 *cluster* dengan *cluster* 0 merupakan *cluster* terlaris dengan total penjualan sebanyak 884 parfum, *cluster* 1 merupakan *cluster* yang cukup laris dengan total penjualan sebanyak 498 parfum, dan *cluster* 2 merupakan *cluster* yang kurang laris dengan total penjualan sebanyak 273 parfum (Yansah et al., 2022).

Berdasarkan penjelasan penelitian sebelumnya, tujuan utama dari penelitian ini adalah untuk mengidentifikasi pola perilaku pelanggan yang berbeda dan mengelompokkan produk sesuai dengan pola tersebut, sehingga setiap produk dapat dipasarkan dengan cara yang paling efektif untuk segmen yang tepat. Melalui analisis data, penelitian ini akan mengungkap segmen pasar yang potensial dan memberikan rekomendasi tentang bagaimana perusahaan ritel dapat menyesuaikan penawaran produknya untuk meningkatkan penjualan dan keuntungan. Berdasarkan hasil identifikasi segmentasi produk maka diharapkan penelitian ini tidak hanya akan memberikan wawasan tentang perilaku pelanggan tetapi juga menyediakan panduan praktis bagi perusahaan ritel dalam mengimplementasikan strategi segmentasi produk yang efektif.

METODE PENELITIAN

Dalam penelitian ini, metode yang digunakan adalah K-Means dengan tahapan penelitian seperti pada gambar di bawah.



Gambar 1. Alur Tahapan Penelitian

Pengumpulan Data

Tahapan awal dalam penelitian adalah pengumpulan data yang bersumber dari situs web *kaggle* dengan judul *Retail Transaction Dataset*.

Data Acquisition

Setelah mengumpulkan data, tahapan selanjutnya adalah tahap *data acquisition*. Tahapan ini merupakan awal dari proses analisis data. Pada tahapan ini dilakukan pengenalan data dan analisis nilai yang ada pada setiap variabel. Tahapan ini bertujuan untuk menggali informasi penting dalam data.

Preprocessing

Tahap ini merupakan tahapan setelah *data acquisition*. Pada tahapan ini dilakukan proses pengecekan dan pembersihan data dengan memastikan tidak ada data duplikat, missing value, dan outlier pada dataset. Tujuan tahapan ini adalah menyiapkan data untuk siap diolah sehingga proses implementasi metode dapat berjalan dengan lancar.

Implementasi K-Means Clustering

Tahap berikutnya setelah data *pre-processing* adalah implementasi model *k-means clustering*. Dalam tahap ini, akan dilakukan *clustering* data yang nantinya menghasilkan beberapa segmen produk sesuai dengan kesamaan karakteristik yang diperoleh dari hasil *clustering*. Berikut merupakan penjelasan langkah-langkah algoritma *k-means clustering*.

1. Menentukan jumlah *cluster* (*k*) optimal dengan menggunakan metode *elbow* untuk memperoleh jumlah *k* optimal.
2. Menentukan titik pusat data (*centroid*) secara acak.
3. Menghitung jarak setiap titik data ke titik pusat data (*centroid*) dengan formula berikut.

$$d(x_i, C_j) = \sqrt{\sum_{m=1}^n (x_{im} - C_{jm})^2} \quad (1)$$

Dimana,

x_i : titik data.

C_j : titik pusat data (*centroid*).

m : data awal (data pertama).

Implementasi K-Means Untuk Melakukan Segmentasi Produk Berdasarkan Data Transaksi Retail

- n : data ke-n.
4. Mengumpulkan nilai minimum dari setiap titik data untuk mencari tahu cluster dari data yang digunakan.
 5. Menginterpretasikan hasil prediksi berupa total dan karakteristik dari setiap *cluster*. Namun, bila terdapat perbedaan titik pusat *cluster*, proses akan diulangi dari menentukan *centroid* baru.
 6. Menentukan *centroid* baru dengan mencari rata-rata setiap *cluster* dari pemetaan iterasi sebelumnya dengan formula berikut.

$$C_j = \frac{1}{|S_j|} \sum_{x_i \in S_j} x_i \quad (2)$$

Dimana,

S_j : kumpulan titik data dalam cluster j.

C_j : titik pusat data (centroid).

x_i : titik data.

HASIL DAN PEMBAHASAN

a. Persiapan Data

Penelitian yang dilakukan mengambil *dataset* yang bersumber dari situs Kaggle dengan judul “Retail Transaction Dataset” yang dapat diakses melalui tautan berikut <https://www.kaggle.com/datasets/fahadrehman07/retail-transaction-dataset>. Berdasarkan *dataset* tersebut, terdapat 100.000 baris data dengan dari 10 variabel yang tersedia. Namun, pada penelitian ini hanya akan diambil 1000 baris data yang diambil pada tanggal 25 April 2024 hingga 28 April 2024. Variabel yang digunakan untuk melakukan analisis selanjutnya hanya diambil 3 variabel dari total 10 variabel yang tersedia.

CustomerID	ProductID	Quantity	Price	TransactionDate	PaymentMethod	StoreLocation	ProductCategory	DiscountApplied(%)	TotalAmount
368213	A	7	24.13081147	9/30/2023 9:53	PayPal	6831 Tiffany Rue Apt. 791 Lesliefort, OH 49412	Home Decor	17.327925	139.6460979
920607	B	9	45.07803681	9/30/2023 9:52	Cash	USS Stout FPO AE 54528	Clothing	1.804263092	398.3823959
371778	A	3	69.98345715	9/30/2023 9:46	PayPal	PSC 0682, Box 2301 APO AA 39262	Books	18.60577974	170.8874678
240452	C	4	61.62944567	9/30/2023 9:43	PayPal	760 Samuel Common Apt. 458 South Emilyburgh, OK 30074	Clothing	17.76792771	202.7166813
641281	C	9	42.00161977	9/30/2023 9:37	Credit Card	42020 Lori Rapid Jasonborough, TN 86158	Books	17.41853321	312.1699831
845675	D	5	65.075185	9/30/2023 9:35	Cash	737 Wu Ferry Jamesberg, UT 45720	Books	16.51349119	271.6450003
371439	B	8	65.83328424	9/30/2023 9:33	PayPal	98609 Tracy Circles Suite 113 South Teresa, TX 24932	Clothing	0.429553077	524.4039627
129521	C	2	58.21331118	9/30/2023 9:24	PayPal	051 Julie Roads South Benjaminton, LA 67967	Home Decor	9.539816166	105.3197366
467177	A	9	85.05263288	9/30/2023 9:20	Debit Card	5552 Velazquez Harbor South Tannerfort, MA 60396	Home Decor	1.1834675	756.4145635
391645	B	6	44.51039809	9/30/2023 9:19	Debit Card	748 Brian Prairie Georgefort, AL 69387	Clothing	5.178816911	253.2317164

Gambar 2. Dataset Tampilan Awal

b. Pre-Processing

Tahapan ini bertujuan untuk mempersiapkan data sebelum dilakukan pemodelan dan menghindari kesalahan pada data. Pada tahapan ini dilakukan perubahan format data dan penghapusan data pada urutan 1001 hingga 100.000. Penelitian ini hanya menggunakan 3 kolom

yaitu *Price*, *Quantity*, dan *DiscountApplied(%)* yang berjumlah 1000 baris. Data yang telah melalui proses *pre-processing* ini dapat ditemukan pada gambar di bawah ini:

	A	B	C
1	Quantity	Price	TransactionDate
2	2	47.31868527	2024
3	8	27.59375726	2024
4	5	61.23959324	2024
5	3	55.35471479	2024
6	1	94.39225898	2024
7	7	21.23646427	2024
8	9	53.52190556	2024
9	2	39.67339637	2024
10	3	17.93914008	2024
11	9	55.74080004	2024
12	5	35.27004451	2024
13	4	49.22860159	2024
14	3	95.44666168	2024
15	9	49.16142535	2024
16	8	95.74351264	2024
17	7	12.08831843	2024
18	1	40.72858457	2024
19	2	73.41590534	2024
20	2	29.19928045	2024
21	1	34.98166577	2024

Gambar 3. Data Setelah Hasil *Pre-Processing*

c. Statistika Deskriptif

Berdasarkan hasil analisis menggunakan *software* SPSS diperoleh statistika deskriptif untuk setiap variabel yang digunakan pada penelitian seperti pada gambar berikut.

```
#mendapatkan deskripsikan tiap atribut
df.describe()
```

	CustomerID	Quantity	Price
count	1000.00	1000.00	1000.00
mean	513168.47	4.88	52.95
min	754.00	1.00	10.00
25%	272720.25	3.00	29.30
50%	513643.00	5.00	51.62
75%	770105.50	7.00	76.04
max	999860.00	9.00	99.96
std	289569.17	2.55	26.47

Gambar 4. Statistika Deskriptif

d. Implementasi K-Means Clustering

Untuk memulai program *python* maka terlebih dahulu mengimport *library* yang dibutuhkan untuk menjalankan kode pemrograman. Media yang digunakan untuk menjalankan proses implementasi kode program ini adalah *Google Colab*. Setelah tahap *pre-processing* telah selesai dengan bukti bahwa data tidak memiliki duplikat, data kosong dan outlier serta tipe data variabel yang digunakan pada proses *clustering* sesuai, maka data siap diolah untuk tahap implementasi *k-means clustering*.

Implementasi K-Means Untuk Melakukan Segmentasi Produk Berdasarkan Data Transaksi Retail

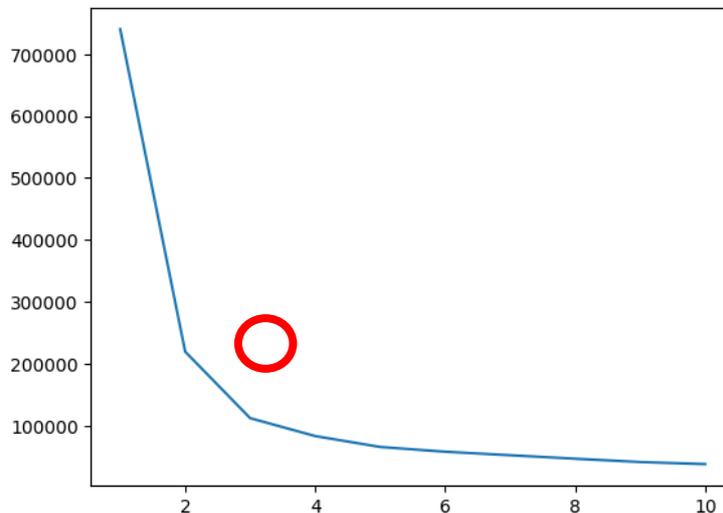
1. Membuat *Subset* Data Baru

```
#membuat subset data baru untuk menampung 3 variabel yang menjadi  
#variabel dari proses cluster  
X = df.loc[:, ['Price', 'Quantity', 'DiscountApplied(%)']]
```

Gambar 5. Membuat Subset Data

Untuk melakukan proses *clustering*, maka variabel yang dianalisis dalam implementasi akan dimasukkan ke dalam *subset* data yang baru untuk menampung variabel-variabel tersebut sehingga tidak mempengaruhi variabel lainnya. Variabel yang digunakan dalam proses *clustering* adalah 'Price', 'Quantity' dan 'DiscountApplied'. Atribut tersebut dijadikan sebagai variabel dalam proses klasterisasi karena memiliki data *numeric* yang merupakan salah satu syarat pada algoritma K-Means serta menjadi komponen penting dalam rekomendasi keputusan pemasaran bisnis retail.

2. Metode *Elbow*



Gambar 6. Grafik Metode *Elbow*

Metode *elbow* merupakan metode yang digunakan untuk menentukan jumlah k optimal dalam klasterisasi pada metode K-Means. Jumlah k yang optimal dalam proses *clustering* dapat diketahui melalui grafik metode *elbow* yang menunjukkan sudut berupa siku. Sudut siku tersebut merupakan bentuk dari penurunan signifikan sebelum daerah dengan penurunan yang tidak signifikan. Berdasarkan grafik metode *elbow* di atas, diketahui bahwa jumlah k optimal adalah 3.

3. Clustering K-Means

```
#melakukan proses clustering
nclust = 3
kmean = KMeans(n_clusters=nclust).fit(X)
X['Labels'] = kmean.labels_
```

Gambar 7. Inisiasi Model K-Means

```
from mpl_toolkits.mplot3d import Axes3D
fig = plt.figure()
ax = fig.add_subplot(111, projection='3d')

# Warna untuk tiap kluster
colors = ['r', 'g', 'b']

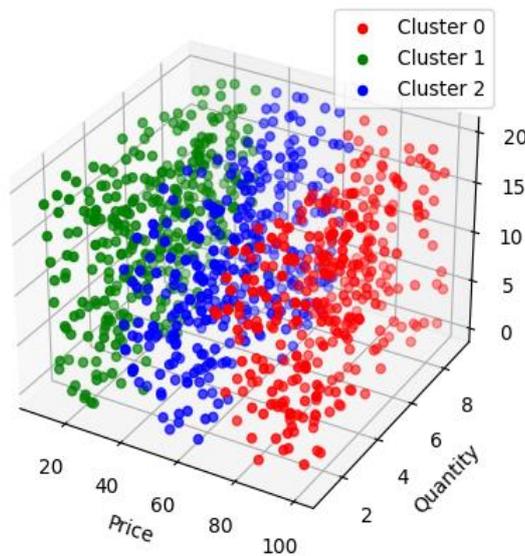
for i in range(nclust):
    cluster = X[X['Labels'] == i]
    ax.scatter(cluster['Price'], cluster['Quantity'], cluster['DiscountApplied(%)'], c=colors[i], label=f'Cluster {i}')

ax.set_xlabel('Price')
ax.set_ylabel('Quantity')
ax.set_zlabel('DiscountApplied(%)')
ax.legend()

plt.show()
```

Gambar 8. Proses Clustering dan Pembuatan Plot 3D Cluster

Tahap ini merupakan proses klasterisasi yang akan membagi data berdasarkan variabel dan jumlah k yang telah diinput. Proses klasterisasi akan menghasilkan label yang menjadi keterangan setiap class yang dimiliki oleh data.



Gambar 9. Plot Hasil Clustering

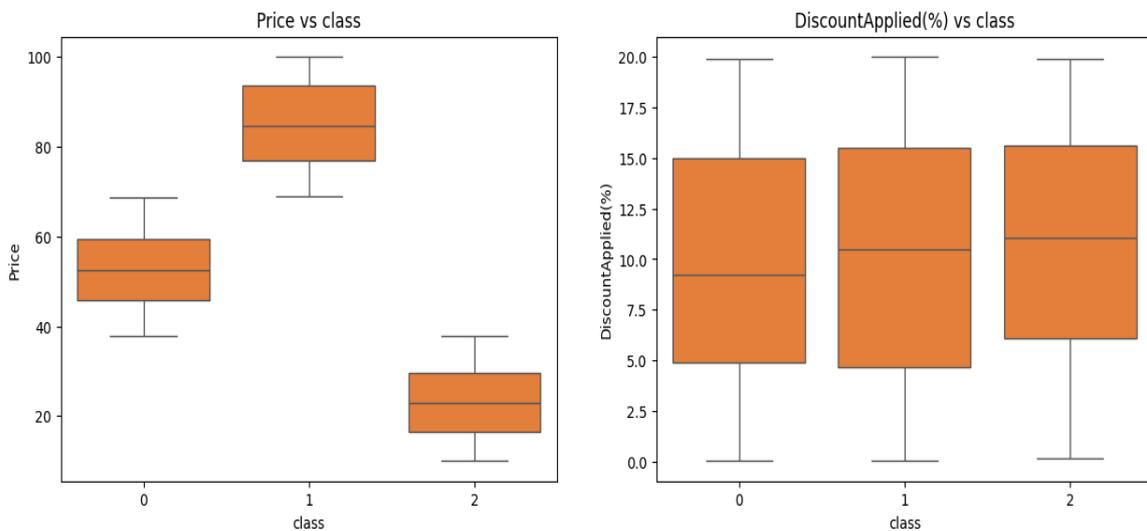
Implementasi K-Means Untuk Melakukan Segmentasi Produk Berdasarkan Data Transaksi Retail

Plot hasil *clustering* diatas menunjukkan visualisasi segmentasi berdasarkan atribut 'Price', 'Quantity' dan 'DiscountApplied(%)'. Proses klasterisasi menghasil 3 *cluster* atau segmen yaitu klaster 0, 1 dan 2.

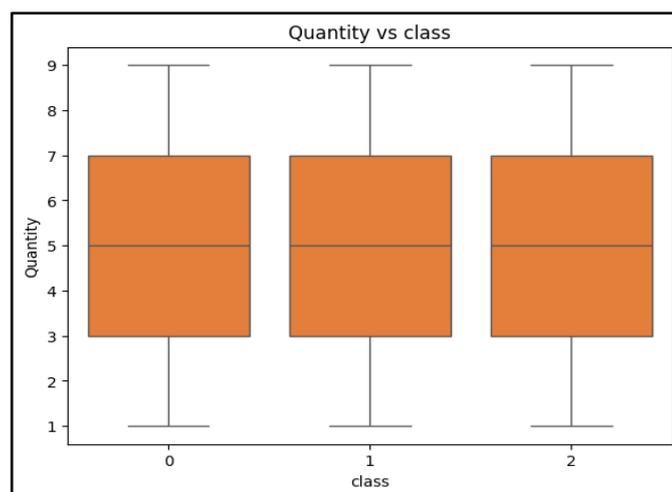
Proses *clustering* menghasilkan label yang menunjukkan keterangan *cluster* yang dimiliki oleh setiap data. Label ini menjadi atribut atau kolom baru di subset yang diinput untuk proses klasterisasi, sehingga kolom label ini akan ditambahkan ke data utama dengan kolom baru yaitu *class*.

4. Visualisasi Data

Setelah melakukan proses *clustering* yang membagi data menjadi beberapa segmen menggunakan implementasi K-Means, maka dilakukan visualisasi untuk mempermudah memahami *output* dari program menggunakan beberapa grafik yang menjelaskan terkait 3 variabel yang dianalisis.



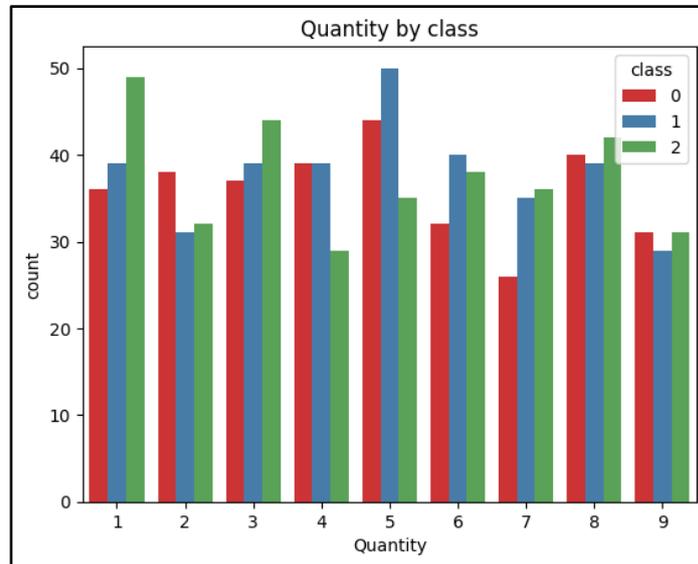
Gambar 10. Boxplot Price VS Class dan DiscountApplied(%) VS Class



Gambar 11. Boxplot Quantity VS Class

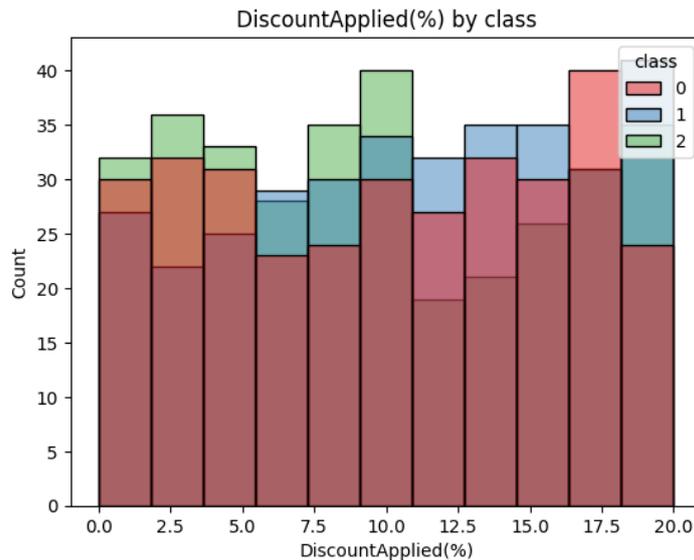
Berdasarkan ketiga *boxplot* diatas dapat diketahui bahwa variabel *price* menunjukkan perbedaan yang signifikan dibandingkan variabel *discountapplied(%)* dan *quantity*. Hal ini

dikarenakan *range* data untuk variabel *discountapplied(%)* dan *quantity* tidak terlalu besar, terkhususnya variabel *quantity* yang *rangennya* hanya dari 0 - 9 serta rata-rata 5 yang mengakibatkan variabel *quantity* tidak terlalu berpengaruh dalam segmentasi sehingga tidak menjadi karakteristik yang membedakan setiap *cluster*.



Gambar 12. Quantity Berdasarkan Class

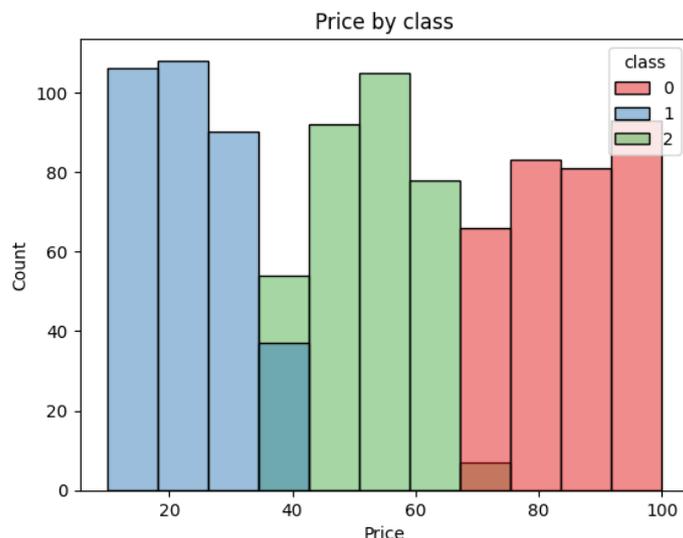
Grafik diatas menunjukkan visualisasi atribut *quantity* berdasarkan *class*. Pada grafik dapat diketahui jumlah banyak barang pada *cluster* 0 ada pada jumlah *quantity* 5, jumlah banyak barang pada *cluster* 1 ada pada jumlah *quantity* 5 dan jumlah banyak barang pada *cluster* 2 ada pada jumlah *quantity* 1.



Gambar 13. DiscountApplied(%) Berdasarkan Class

Implementasi K-Means Untuk Melakukan Segmentasi Produk Berdasarkan Data Transaksi Retail

Grafik diatas menunjukkan visualisasi atribut *discountapplied(%)* berdasarkan *class*. Pada grafik dapat diketahui distribusi diskon tertinggi pada *cluster* 0 ada pada diskon 10%, distribusi diskon tertinggi pada *cluster* 1 ada pada diskon 20% dan jumlah distribusi diskon tertinggi pada *cluster* 2 ada pada diskon 10%.



Gambar 14. Price Berdasarkan Class

Grafik diatas menunjukkan visualisasi atribut *price* berdasarkan *class*. Pada grafik dapat diketahui distribusi harga produk pada *cluster* 0 ada pada rentang 0 hingga 30, distribusi harga produk pada *cluster* 1 ada pada rentang 40 hingga 70 dan distribusi harga produk pada *cluster* 2 ada pada rentang 80 hingga 100.

e. Analisis Hasil Clustering

Hasil implementasi metode K-Means yang membagi produk menjadi 3 segmen yaitu *cluster* 0,1 dan 2. Pengelompokan ini dilakukan berdasarkan data transaksi retail menggunakan variabel *price* (harga produk), *quantity* (jumlah barang yang dibeli) dan *discountapplied(%)* (persentase diskon setiap produk). Dari hasil *clustering*, maka berikut merupakan penjelasan masing-masing *cluster*.

Cluster 0 (Label = 0) : kelompok produk ini memiliki rata-rata harga satuan adalah 52 dolar, rata-rata kuantitas adalah 5 *pieces* dan rata-rata persentase diskon adalah 9 persen. Kelompok ini memiliki pelanggan yang membeli produk dengan harga menengah dan mendapatkan diskon barang yang rendah. Rekomendasi pemasaran yang dapat dilakukan pada kelompok pelanggan ini adalah memberikan diskon untuk pembelian berikutnya dalam jumlah besar sehingga ada insentif untuk belanja lebih banyak serta menawarkan langganan dan rekomendasi produk yang sesuai dengan pembelian produk harga menengah.

Cluster 1 (Label = 1) : kelompok produk ini memiliki rata-rata harga satuan adalah 85, rata-rata kuantitas adalah 5 dolar dan rata-rata persentase diskon adalah 10 persen. Kelompok ini memiliki pelanggan yang membeli produk dengan harga tinggi dan mendapatkan diskon barang yang menengah. Rekomendasi pemasaran yang dapat dilakukan pada kelompok pelanggan ini adalah dengan menawarkan program kualitas premium seperti *VIP membership* untuk mendapatkan harga dan pelayanan eksklusif serta memberikan diskon bertingkat untuk meningkatkan jumlah pembelian produk.

Cluster 2 (Label = 2) : kelompok produk ini memiliki rata-rata harga satuan adalah 21 dolar, rata-rata kuantitas adalah 5 *pieces* dan rata-rata persentase diskon adalah 11 persen. Kelompok ini memiliki pelanggan yang membeli produk dengan harga rendah dan mendapatkan diskon barang yang menengah. Rekomendasi pemasaran yang dapat dilakukan pada kelompok pelanggan ini adalah pemberian diskon bertingkat serta menawarkan paket hemat sehingga mendorong pembelian dalam jumlah besar.

KESIMPULAN

Dalam penelitian ini ditemukan 3 *cluster* (segmen) pelanggan. Segmen pertama (*cluster 0*) merupakan kelompok produk yang memiliki persentase pembeli sebesar 34,1%. Segmen kedua (*cluster 1*) merupakan kelompok produk yang memiliki persentase pembeli sebesar 33,7%. Segmen ketiga (*cluster 2*) merupakan kelompok produk yang memiliki persentase pembeli sebesar 32,2%. Dari ketiga *cluster*, diperoleh bahwa *cluster 0* (segmen pertama) merupakan kelompok produk potensial yang memiliki ciri berupa produk dibeli dengan harga menengah dan mendapatkan diskon produk yang rendah. Sehingga rekomendasi pemasaran yang dilakukan adalah memberikan insentif untuk belanja lebih banyak serta menawarkan langganan dan rekomendasi produk yang sesuai dengan pembelian produk harga menengah. Variabel yang berpengaruh dalam penelitian ini adalah variabel *price* (harga) yang memiliki persebaran data 50 dollar hingga 100 dollar sehingga menjadi karakteristik yang membedakan setiap *cluster*.

Implementasi K-Means Untuk Melakukan Segmentasi Produk Berdasarkan Data Transaksi Retail

SARAN

Berdasarkan hasil penelitian dengan *dataset retail transaction* dengan metode *k-means clustering*, berikut merupakan saran yang dapat dipertimbangkan untuk penelitian yang akan datang.

1. Penelitian Lanjutan

Peneliti selanjutnya dapat mengidentifikasi serta menambahkan variabel-variabel tertentu guna memperoleh hasil segmentasi produk yang lebih banyak dan beragam. Selain itu peneliti juga dapat mencoba menggunakan metode lain untuk memperoleh hasil segmentasi produk yang lebih baik.

2. Pengimplementasian Metode

Perusahaan *retail* dapat mencoba mengaplikasikan metode *k-means clustering* guna menghasilkan keputusan bisnis yang tepat. Penggunaan metode *k-means clustering* dapat membantu perusahaan dalam mengidentifikasi dan menyesuaikan produk terhadap pasar guna menghasilkan strategi pemasaran yang tepat.

DAFTAR PUSTAKA

- Nisa, F. D.(dkk). (2023). EVALUASI RENCANA STRATEGIS DALAM MEMBANGUN KEUNGGULAN BERSAING DI INDUSTRI RETAIL: KASUS STUDI PADA BEBERAPA PERITEL TERKEMUKA. *Musyteri Neraca Manajemen Ekonomi*, 1(11), 1-13.
- Ramadhanty, S. (2023). *Aprindo Ungkap Pertumbuhan Ritel Nasional Hingga Tutup Tahun 2023 Bisa Capai 4,2%*. Diakses tanggal 29 Juli 2024 dari <https://industri.kontan.co.id/news/aprindo-ungkap-pertumbuhan-ritel-nasional-hingga-tutup-tahun-2023-bisa-capai-42>
- Fithriyah, M. (dkk). (2021). K-Means Clustering Untuk Segmentasi Produk Berdasarkan Analisis Recency, Recency, Recency (RFM) pada Data Transaksi Penjualan. *ILKOMNIKA : Journal of Computer Science and Applied Informatics*, 3(2), 151-164.
- Harani, N. (dkk). (2020). Segmentasi Pelanggan Produk Digital Service Indihome Menggunakan Algoritma K-Means Berbasis Python. *Jurnal Manajemen Informatika (JAMIKA)*, 10(2), 133-146.
- Amna. (dkk). (2023). *DATA MINING*. Padang: PT GLOBAL EKSEKUTIF TEKNOLOGI.
- Rahayu, P. (dkk). (2024). *Buku Ajar Data Mining*. Jambi: PT. Sonpedia Publishing Indonesia.
- Astuti, S. (2020). ALGORITMA K-MEANS DALAM MENENTUKAN PENERIMA BEASISWA UPZ (Unit Pengumpulan Zakat) PADA MAHASISWA UIN SUMATERA UTARA MEDAN. Tugas Akhir Sarjana. FAKULTAS SAINS DAN TEKNOLOGI UIN SUMATERA UTARA.
- Salsabila, N. (2019). KLASIFIKASI BARANG MENGGUNAKAN METODE CLUSTERING & K-MEANS DALAM PENENTUAN PREDIKSI STOK BARANG. Tugas Akhir Sarjana. FAKULTAS SAINS DAN TEKNOLOGI UIN MAULANA MALIK IBRAHIM.
- Achray, A. (2020). Implementasi Algoritma *K-Means* untuk Menentukan Mengelompokkan Data Penjualan Mobil di PT Honda Arista Mangga Dua. Tugas Akhir Sarjana. FT UNIVERSITAS SATYA NEGARA INDONESIA.
- Rahman, S. (dkk). (2023). *PYTHON : DASAR DAN PEMROGRAMAN BERORIENTASI OBJEK*. PENERBIT TAHTA MEDIA GROUP.
- Harani, N. Nugraha, F. (2020). *Segmentasi Pelanggan Menggunakan Python*. Bandung: Kreatif Industri Nusantara.
- Sunyoto, D. Mulyono, A. (2021). *MANAJEMEN BISNIS RITEL*. Purbalingga: CV. EUREKA MEDIA AKSARA.

- Wijaya, K. (dkk). (2021). Segmentasi Pelanggan Menggunakan Algoritma K-Means dan Analisis RFM di Ova Gaming E-Sports Arena Kediri. *JURNAL TEKNIK ITS*, 10(2), 230-237.
- Awalina, E. Woro Isti Rahayu. (2023). Optimalisasi Strategi Pemasaran dengan Segmentasi Pelanggan Menggunakan Penerapan K-Means Clustering pada Transaksi Online Retail. *Jurnal Teknologi dan Informasi (JATI)*, 13(2), 122-137.
- Fakhriza, M. Khaerul Umam. (2021). Analisis Produk Terlaris Menggunakan Metode K-Means Clustering Pada PT. Sukanda Djaya. *JIKA (Jurnal Informatika)*, 5(1), 8-15.
- Yansah, H, (dkk). (2022). Penerapan Algoritma K-Means Dalam Clustering Produk Terlaris Pada Fr Parfum. *SAKTI : Sains, Aplikasi, Komputasi dan Teknologi Informasi*, 4(2), 83-90.