

**ANALISIS PENGELOMPOKAN MENU GOFOOD
BERDASARKAN HARGA DAN DISKON MENGGUNAKAN
ALGORITMA K-MEANS CLUSTERING**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 INFORMATIKA



disusun oleh

MUHAMMAD SYAHRIZAL

21.11.3968

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

**ANALISIS PENGELOMPOKAN MENU GOFOOD
BERDASARKAN HARGA DAN DISKON MENGGUNAKAN
ALGORITMA K-MEANS CLUSTERING**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 INFORMATIKA



disusun oleh

MUHAMMAD SYAHRIZAL

21.11.3968

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

**ANALISIS PENGELOMPOKAN MENU GOFOOD BERDASARKAN
HARGA DAN DISKON MENGGUNAKAN ALGORITMA K-MEANS
CLUSTERING**

yang disusun dan diajukan oleh

Muhammad Syahrizal

21.11.3968

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 19 November 2025

Dosen Pembimbing,



Nur Aini, A.Md., S.Kom., M.Kom.

NIK. 190302066

HALAMAN PENGESAHAN

SKRIPSI

**ANALISIS PENGELOMPOKAN MENU GOFOOD BERDASARKAN
HARGA DAN DISKON MENGGUNAKAN ALGORITMA K-MEANS
CLUSTERING**

yang disusun dan diajukan oleh

Muhammad Syahrizal
21.11.3968

Telah dipertahankan di depan Dewan Penguji
pada tanggal 19 November 2025

Susunan Dewan Penguji

Nama Penguji

Anna Baita, S.Kom., M.Kom.
NIK. 190302290

Uyock Anggoro Saputro, S.Kom., M.Kom.
NIK. 190302419

Nur Aini, A.Md., S.Kom., M.Kom.
NIK. 190302066

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 19 November 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusrini, M.Kom.

NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Muhammad Syahrizal
NIM : 21.11.3968

Menyatakan bahwa Skripsi dengan judul berikut:

ANALISIS PENGELOMPOKAN MENU GOFOOD BERDASARKAN HARGA DAN DISKON MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING

Dosen Pembimbing : Nur Aini, A.Md., S.Kom., M.Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 19 November 2025

Yang Menyatakan,



Muhammad Syahrizal



HALAMAN PERSEMBAHAN KATA PENGANTAR

Alhamdulillah Rabbil ‘Alamin, segala puji dan syukur penulis panjatkan ke hadirat Allah Subhanahu wa Ta'ala atas limpahan rahmat, hidayah, dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “*Analisis Pengelompokan Menu GoFood Berdasarkan Harga dan Diskon Menggunakan Algoritma K-Means Clustering*” dengan baik. Skripsi ini disusun untuk memenuhi salah satu syarat memperoleh gelar Sarjana pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta.

Dalam proses penyusunan skripsi ini, penulis menyampaikan ucapan terima kasih kepada:

1. Bapak Prof. Dr. M. Suyanto, M.M., selaku Rektor Universitas Amikom Yogyakarta.
2. Ibu Prof. Dr. Kusri, S.Kom., M.Kom., selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Ibu Eli Pujastuti, M.Kom., selaku Ketua Program Studi S1 Informatika Universitas Amikom Yogyakarta.
4. Ibu Nur Aini, A.Md., S.Kom., M.Kom., selaku dosen pembimbing yang telah memberikan bimbingan, arahan, serta masukan yang sangat berarti selama proses penelitian hingga penyusunan skripsi ini.
5. Bapak dan Ibu orang tua penulis beserta keluarga besar, atas doa, dukungan moral maupun material, serta motivasi yang senantiasa menguatkan penulis dalam menyelesaikan studi ini.

Penulis berharap skripsi ini dapat memberikan manfaat bagi pembaca, peneliti selanjutnya, serta pihak-pihak yang berkepentingan dalam pengembangan analisis data pada layanan pemesanan makanan daring.

Yogyakarta, 19 November 2025

Penulis

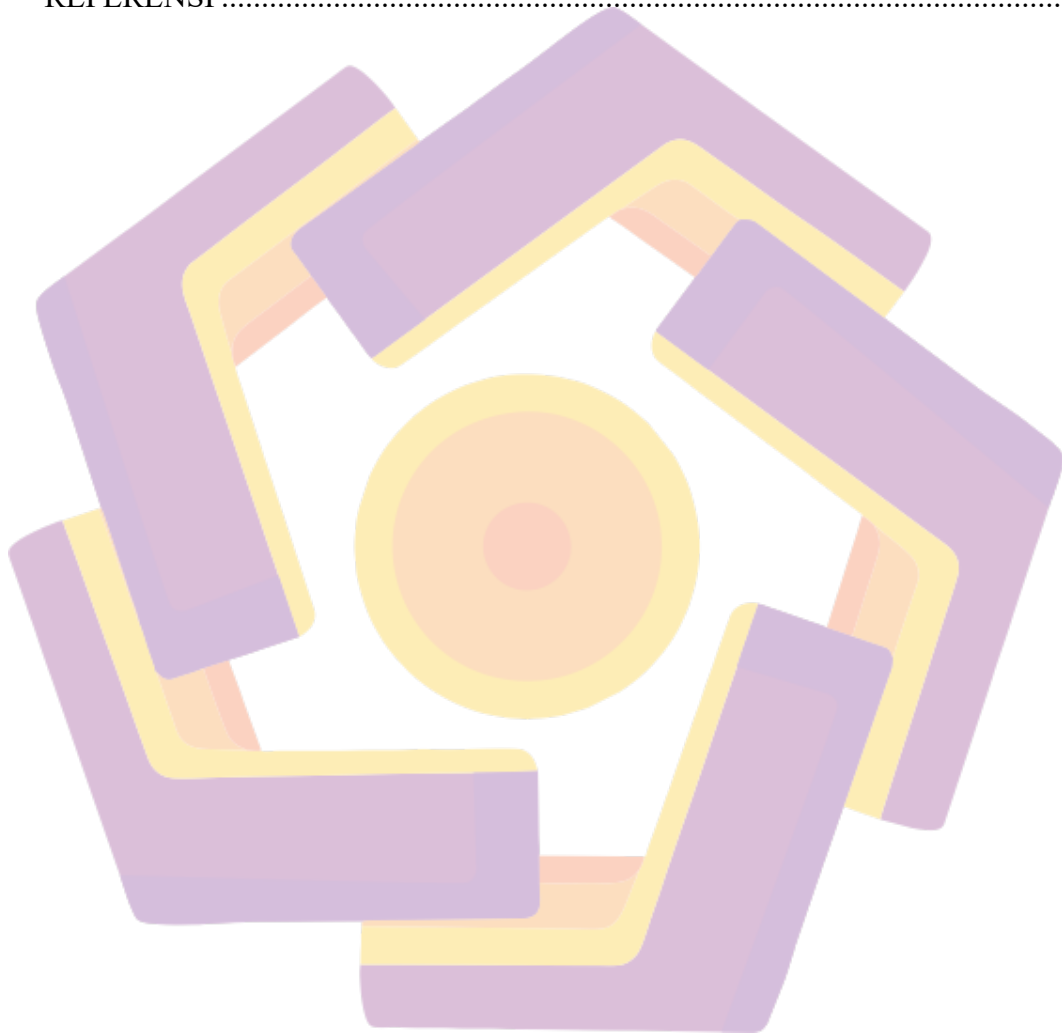
DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	v
DAFTAR ISI.....	vi
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMBANG DAN SINGKATAN	xii
DAFTAR ISTILAH	xiii
INTISARI	xvi
ABSTRACT.....	xvii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA	7
2.1 Studi Literatur	7
2.2 Dasar Teori.....	12

2.2.1	K-Means Clustering	12
2.2.2	Harga (Price)	13
2.2.3	Diskon Harga (Discount Rate)	13
2.2.4	Pre-processing Data	14
2.2.5	Standardisasi Data dengan StandardScaler	14
2.2.6	Elbow Method	15
2.2.7	Evaluasi Kualitas Kluster	16
2.2.8	Silhouette Score	16
2.2.9	Davies Bouldin Index (DBI)	17
2.2.10	Interpretasi Kluster	17
2.2.11	Python	18
BAB III METODE PENELITIAN		19
3.1	Objek Penelitian	19
3.2	Alur Penelitian	20
3.3	Alat dan Bahan	22
3.3.1	Data Penelitian	22
3.3.2	Alat dan Perangkat Lunak	23
3.3.3	Spesifikasi Perangkat Keras	24
3.4	Alur Proses Penelitian	24
3.4.3	Pra-Pemrosesan Data (Data Preprocessing)	26
3.4.4	Pemilihan dan Penskalaan Fitur (Feature Selection & Scaling)	27
3.4.5	Penentuan Jumlah Kluster Optimal (Elbow Method)	27
3.4.6	Penerapan Algoritma K-Means	28
3.4.7	Evaluasi Hasil Klusterisasi (Clustering Evaluation)	28
3.4.8	Analisis dan Interpretasi Kluster	29

BAB IV HASIL DAN PEMBAHASAN	30
4.1 Gambaran Umum Dataset.....	30
4.2 Hasil Eksplorasi Data (EDA).....	32
4.2.1 Pemeriksaan struktur dan tipe data	32
4.2.2 Analisis Deskriptif Statistik	33
4.2.3 Handling missing values (Pemeriksaan Nilai Kosong).....	35
4.2.4 Analisis Distribusi Kategorikal.....	37
4.3 Preprocessing Data.....	43
4.3.1 Handling Missing Values.....	43
4.3.2 Data Cleaning.....	44
4.3.3 Penskalaan Fitur (Feature Scaling)	47
4.4 Penentuan Jumlah Klaster (Elbow Method)	48
4.5 Implementasi K-Means Clustering	50
4.5.1 Rata-Rata Harga dan Diskon per Klaster.....	51
4.5.2 Visualisasi Hasil Klasterisasi.....	52
4.6 Evaluasi Hasil Klasterisasi.....	53
4.6.1 Silhouette Score	54
4.6.2 Davies–Bouldin Index (DBI).....	54
4.7 Analisis Klasterisasi.....	55
4.7.1 Distribusi Merchant Area Akhir	56
4.7.2 Distribusi Kategori Produk Akhir.....	57
4.7.3 Distribusi Nilai Display Akhir	59
4.8 Analisis dan Interpretasi Klaster	60
4.8.1 Klaster 0 – Produk Ekonomis	61
4.8.2 Klaster 1 – Produk Menengah dengan Diskon Moderat	61

4.8.3	Klaster 2 – Produk Premium.....	62
BAB V PENUTUP		64
5.1	Kesimpulan	64
5.2	Saran	65
REFERENSI		66



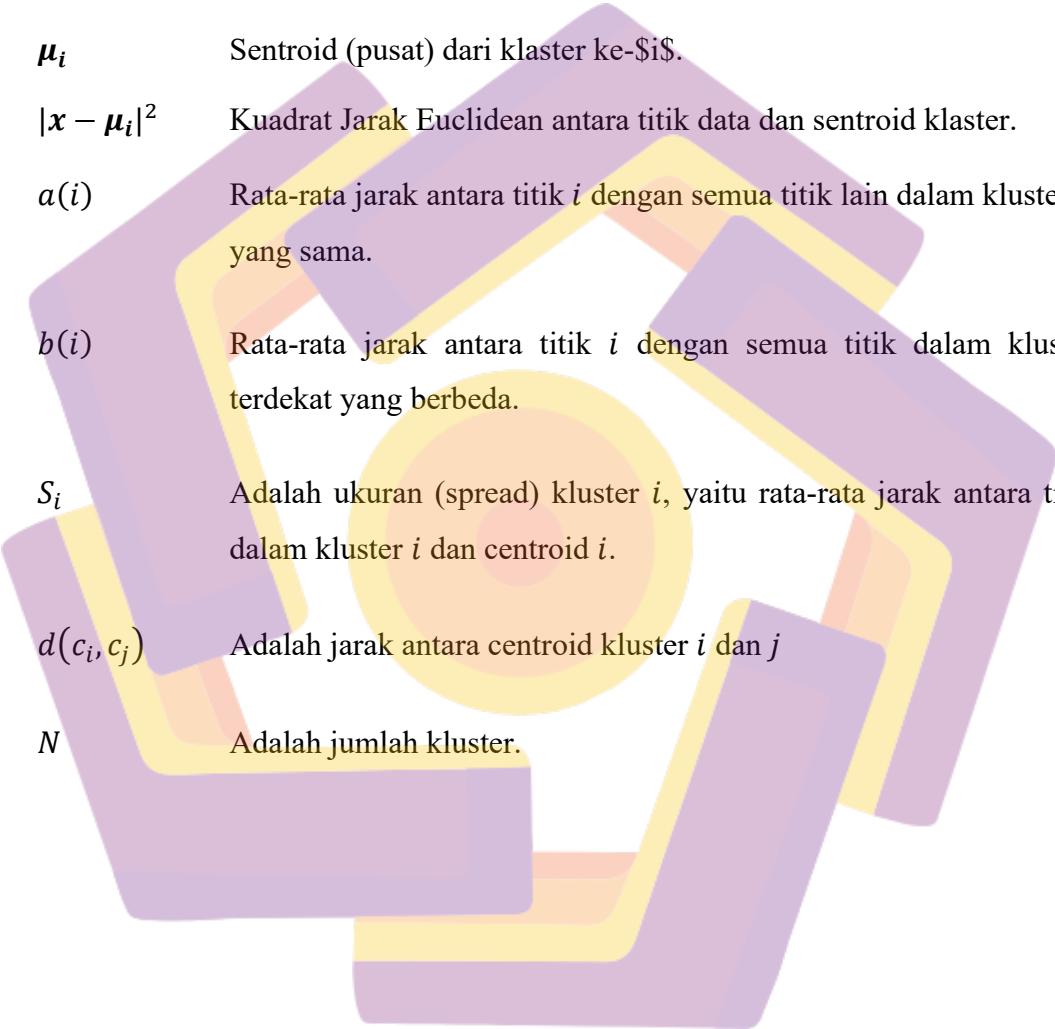
DAFTAR TABEL

Tabel 2. 1 Keaslian Penelitian	10
Tabel 3. 1 Karakteristik dataset	22
Tabel 3. 2 Alat dan Perangkat Lunak.....	23
Tabel 3. 3 Spesifikasi Perangkat Keras.....	24
Tabel 4. 1 Keterangan Fitur	30
Tabel 4. 2 Dataset awal	31
Tabel 4. 3 Cluster Rata-Rata Harga	51
Tabel 4. 4 Cluster Rata-Rata Harga Diskon.....	51
Tabel 4. 5 Evaluasi Silhouette Score	54
Tabel 4. 6 Davies–Bouldin Index (DBI).....	54
Tabel 4. 7 Ringkasan Interpretasi	60
Tabel 4. 8 Cluster 0	61
Tabel 4. 9 Cluster 1	62
Tabel 4. 10 Cluster 2	62

DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian	20
Gambar 3. 2 Alur Proses	25
Gambar 4. 1 Integrasi Google Drive.....	31
Gambar 4. 2 Struktur dan Tipe Data Awal	32
Gambar 4. 3 Descriptive Numerik	34
Gambar 4. 4 Descriptive Kategorikal	35
Gambar 4. 5 Code Missing Value	35
Gambar 4. 6 Missing Value awal.....	36
Gambar 4. 7 Distribusi Merchant Area.....	37
Gambar 4. 8 Distribusi Category	38
Gambar 4. 9 Distribusi Display	39
Gambar 4. 10 Statistik Distribusi Harga	40
Gambar 4. 11 Visualisasi Distribusi Harga.....	41
Gambar 4. 12 Statistik Distribusi Harga	42
Gambar 4. 13 Visualisasi Distribusi Harga Diskon	43
Gambar 4. 14 Handling Missing Value	44
Gambar 4. 15 Data Cleaning.....	44
Gambar 4. 16 Code Type Checking	45
Gambar 4. 17 Missing Value Akhir	45
Gambar 4. 18 Menghapus Duplikat	45
Gambar 4. 19 Hasil Menghapus duplikat	46
Gambar 4. 20 Feature Scaling.....	47
Gambar 4. 21 Elbow Method	48
Gambar 4. 22 Visualisasi Elbow Method	49
Gambar 4. 23 Visualisasi Hasil Klasterisasi	52
Gambar 4. 24 Evaluasi Klasterisasi	53
Gambar 4. 25 Merchant Area Akhir	56
Gambar 4. 26 Kategori Produk Akhir.....	57
Gambar 4. 27 Nilai Display Akhir	59

DAFTAR LAMBANG DAN SINGKATAN



k	Jumlah kluster
x	Titik data individu.
C_i	Kluster ke- i .
μ_i	Sentroid (pusat) dari kluster ke- i .
$ x - \mu_i ^2$	Kuadrat Jarak Euclidean antara titik data dan centroid kluster.
$a(i)$	Rata-rata jarak antara titik i dengan semua titik lain dalam kluster yang sama.
$b(i)$	Rata-rata jarak antara titik i dengan semua titik dalam kluster terdekat yang berbeda.
S_i	Adalah ukuran (spread) kluster i , yaitu rata-rata jarak antara titik dalam kluster i dan centroid i .
$d(c_i, c_j)$	Adalah jarak antara centroid kluster i dan j
N	Adalah jumlah kluster.

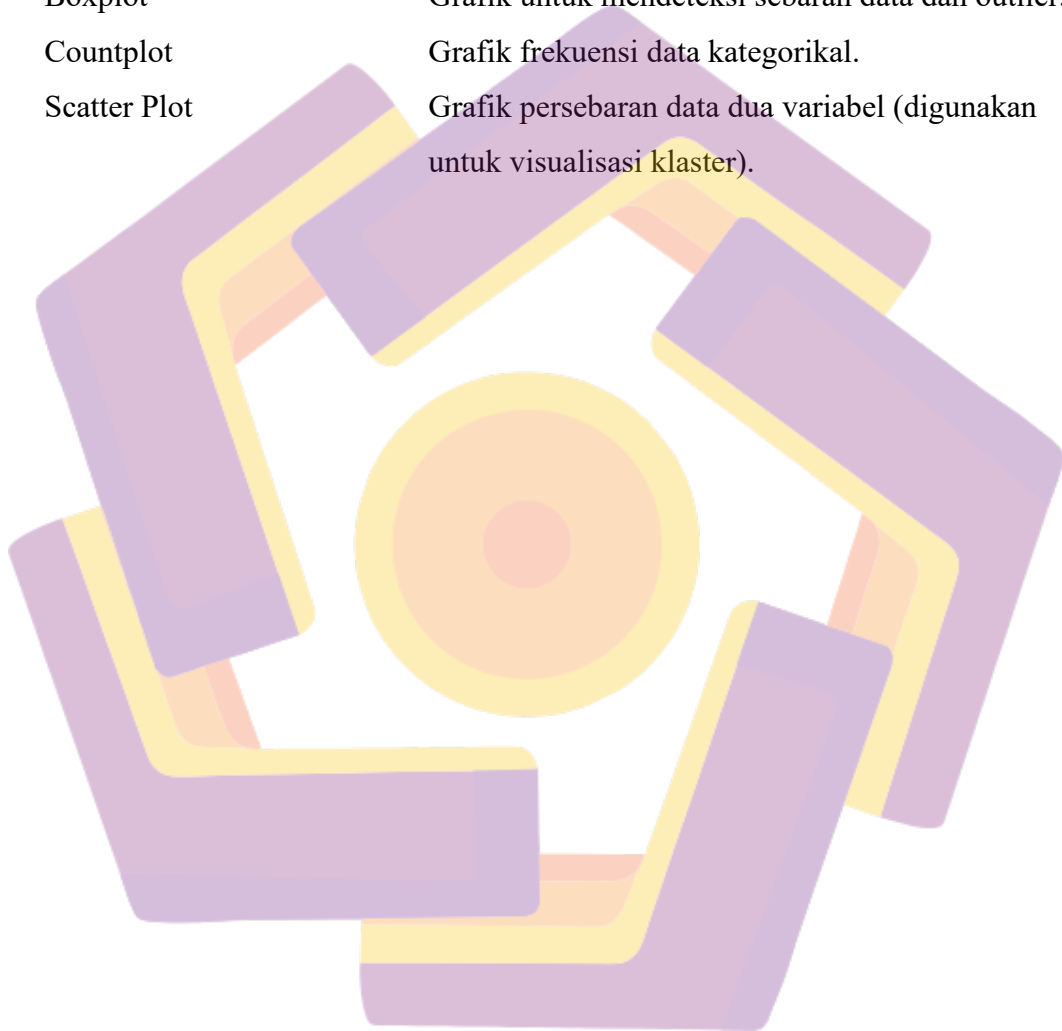
DAFTAR ISTILAH

K-Means Clustering	Algoritma unsupervised learning untuk mengelompokkan data berdasarkan jarak antar titik dan centroid.
Centroid	Titik pusat suatu klaster yang dihitung dari rata-rata seluruh anggota klaster.
Cluster (Klaster)	Kelompok data yang memiliki kemiripan karakteristik berdasarkan fitur tertentu.
Unsupervised Learning	Metode pembelajaran mesin tanpa label; model menemukan pola secara mandiri.
Elbow Method	Teknik untuk menentukan jumlah klaster optimal dengan menganalisis nilai WCSS.
Silhouette Score	Metrik yang menilai kualitas klaster berdasarkan kedekatan data dengan klasternya dibandingkan klaster lain.
Davies–Bouldin Index (DBI)	Metrik evaluasi klaster yang mengukur rasio jarak antar pusat klaster dengan penyebaran dalam klaster.
Euclidean Distance	Jarak geometris antar dua titik dalam ruang multidimensi.
WCSS (Within-Cluster Sum of Squares)	Jumlah total kuadrat jarak dari setiap titik data ke centroid klasternya.
Standard Deviation (Standar Deviasi)	Ukuran penyebaran data dari nilai rata-rata.
Mean, Median, Kuartil	Ukuran-ukuran statistik deskriptif untuk memahami sebaran data.
Outlier	Nilai ekstrem yang jauh dari pola data mayoritas.
Data Cleaning	Proses pembersihan data dari duplikasi, nilai kosong, atau kesalahan tipe.
Handling Missing Values	Penanganan nilai hilang, misalnya mengisi dengan 0 pada <code>discount_price</code> .
Feature Scaling	Teknik menyetarakan skala fitur numerik agar tidak mendistorsi perhitungan jarak.



StandardScaler	Metode penskalaan data menjadi distribusi dengan mean = 0 dan standar deviasi = 1.
Feature Selection	Pemilihan fitur penting yang digunakan dalam proses analisis (price & discount_price).
Merchant	Penjual atau restoran yang menawarkan produk pada platform GoFood.
Merchant Area	Lokasi geografis tempat merchant beroperasi (Jakarta, Medan, Surabaya).
Category	Kategori produk seperti Roti, Minuman, Aneka Nasi, dll.
Display	Tipe tampilan promosi menu di aplikasi GoFood (misal: Paket Hemat, New Varian).
Product	Menu makanan atau minuman yang dijual.
Price (Harga)	Harga asli produk sebelum potongan.
Discount Price (Harga Diskon)	Harga setelah diskon; 0 menunjukkan produk tanpa diskon.
isDiscount	Variabel biner (0/1) yang menunjukkan apakah produk memiliki diskon.
Python	Bahasa pemrograman utama untuk analisis data dan klusterisasi.
Pandas	Library Python untuk manipulasi data tabular.
NumPy	Library untuk komputasi numerik.
Matplotlib & Seaborn	Library visualisasi data.
Scikit-Learn	Library machine learning yang menyediakan fungsi K-Means, StandardScaler, dan evaluasi klaster.
Google Colab	Lingkungan komputasi cloud untuk menjalankan kode Python.
Dataset	Kumpulan data yang digunakan untuk analisis.
EDA (Exploratory Data Analysis)	Analisis eksploratif untuk memahami struktur dan distribusi data.

Interpretasi Klaster	Penjelasan mengenai karakteristik tiap klaster untuk menghasilkan insight.
Segmentasi Produk	Pengelompokan produk berdasarkan kemiripan karakteristik (harga & diskon).
Histogram	Grafik distribusi frekuensi data numerik.
Boxplot	Grafik untuk mendeteksi sebaran data dan outlier.
Countplot	Grafik frekuensi data kategorikal.
Scatter Plot	Grafik persebaran data dua variabel (digunakan untuk visualisasi klaster).



INTISARI

Layanan pesan antar makanan seperti GoFood menghadirkan jutaan produk dari ribuan merchant dengan variasi harga dan diskon yang sangat beragam, sehingga menimbulkan tantangan dalam memahami struktur penawaran dan menentukan strategi pemasaran yang efektif. Kurangnya segmentasi produk berbasis karakteristik harga dan diskon dapat menghambat merchant maupun pihak platform dalam menyusun kebijakan promosi dan strategi harga yang tepat sasaran. Untuk mengatasi masalah tersebut, penelitian ini mengimplementasikan algoritma K-Means Clustering pada dataset menu GoFood tahun 2022 - 2024 dengan fokus pada atribut harga (price) dan harga diskon (discount_price). Tahapan penelitian meliputi pra-pemrosesan data, pengisian nilai hilang, normalisasi fitur, penentuan jumlah kluster optimal menggunakan metode Elbow, serta evaluasi hasil kluster dengan Silhouette Score dan Davies–Bouldin Index. Hasil pengelompokan menunjukkan terbentuknya tiga kluster utama: produk berharga rendah tanpa diskon signifikan, produk menengah dengan potongan harga sedang hingga besar, dan produk premium dengan harga tinggi namun jarang didiskon. Nilai Silhouette Score sebesar 0.8977 dan Davies–Bouldin Index 0.4888 menunjukkan kualitas kluster yang baik. Temuan ini memberikan kontribusi praktis bagi merchant dan pengelola GoFood dalam merancang strategi promosi serta penetapan harga yang lebih terarah berdasarkan segmen produk. Penelitian selanjutnya dapat memperluas analisis dengan menambahkan fitur perilaku konsumen atau ulasan pelanggan untuk meningkatkan akurasi segmentasi pasar.

Kata Kunci: K-Means Clustering, GoFood, Segmentasi Produk, Harga, Diskon

ABSTRACT

Food delivery services such as GoFood offer millions of products from thousands of merchants, posing challenges in understanding product diversity and developing effective marketing strategies. The absence of product segmentation based on price and discount characteristics can hinder both merchants and platform providers in designing targeted promotional and pricing strategies. To address this research gap, this study implements the K-Means Clustering algorithm on the 2022 - 2024 GoFood menu dataset, focusing on the attributes of price and discount_price. The research process involved data preprocessing, missing value imputation, feature normalization, determining the optimal number of clusters using the Elbow Method, and evaluating the results with the Silhouette Score and Davies–Bouldin Index. The clustering analysis revealed three main clusters: low-priced products with minimal discounts, mid-range products with moderate to high discounts, and premium products with high prices but rarely discounted. The obtained Silhouette Score of 0.8977 and Davies–Bouldin Index of 0.4888 indicate good clustering quality. These findings provide practical insights for GoFood merchants and managers in designing more precise promotional and pricing strategies based on product segmentation. Future studies are recommended to enhance segmentation accuracy by incorporating customer behavior data and review-based sentiment features.

Keyword: K-Means Clustering, GoFood, Product Segmentation, Price, Discount