

# **ANALISIS TREN SEGMENTASI PELANGGAN DALAM MENYESUAIKAN STRATEGI BISNIS YANG BERORIENTASI PADA KONSUMEN**



# ANGGOTA KELOMPOK



Irisaliya Irhabiyah  
Banat

Universitas Airlangga



Arnetta Dea  
Suryanti

Universitas Airlangga



Havinka Angel  
Salsabilla

Universitas  
Muhammadiyah Semarang

# DAFTAR ISI

<b>Daftar Isi</b>	<b>03</b>
<b>Business Understanding</b>	<b>04</b>
<b>Business Problem</b>	<b>05</b>
<b>Data Understanding</b>	<b>06</b>
<b>Data Preparation</b>	<b>08</b>
<b>Feature Engineering</b>	<b>09</b>
<b>Exploratory Data Analysis (EDA)</b>	<b>10</b>
<b>Data Scaling</b>	<b>12</b>
<b>Principal Component Analysis (PCA)</b>	<b>13</b>
<b>Modeling</b>	<b>14</b>
<b>Model Terbaik</b>	<b>17</b>
<b>Business Solution</b>	<b>18</b>

# BUSINESS UNDERSTANDING



## Pemanfaatan Data Rantai Pasok untuk Analisis Prediktif

Penggunaan data dari DataCo Smart Supply Chain memungkinkan analisis mendalam terhadap aktivitas produksi, penjualan, dan distribusi guna memprediksi performa rantai pasok.



## Segmentasi Pelanggan untuk Strategi yang Lebih Personal

Analisis segmentasi pelanggan membantu perusahaan memahami perilaku dan preferensi pelanggan, sehingga strategi pemasaran dan operasional bisa lebih terarah dan relevan.



## Optimalisasi Kinerja Bisnis dan Kepuasan Pelanggan

Pendekatan berbasis data membantu perusahaan merumuskan strategi bisnis yang adaptif terhadap perubahan pasar, meningkatkan kepuasan pelanggan, dan loyalitas merek.

# BUSINESS PROBLEM

...



**Banyak perusahaan kesulitan memahami perilaku pelanggan secara menyeluruh dari data rantai pasok yang kompleks**



**Strategi pemasaran masih bersifat umum dan kurang tepat sasaran karena tidak adanya segmentasi pelanggan yang akurat**



**Pengambilan keputusan bisnis sering tidak berbasis data sehingga menghambat efisiensi dan loyalitas pelanggan**

**“Bagaimana perusahaan dapat memanfaatkan segmentasi pelanggan dari data rantai pasok untuk meningkatkan strategi yang tepat sasaran dan profitabilitas jangka panjang?”**

# DATA UNDERSTANDING

Data diperoleh dari DataCo Smart Supply Chain untuk tahun 2018, terdapat 11 kolom dan total 2123 baris data secara keseluruhan

Data	Keterangan	Tipe Data
Customer Id	Kode kategori produk	Numerik
Order Country	Negara tujuan pesanan	Kategorik
Order City	Kota tujuan pesanan	Kategorik
Order Item Quantity	Jumlah produk dalam setiap pesanan	Numerik
Days for shipping (real)	Jumlah hari aktual pengiriman produk yang dibeli	Numerik

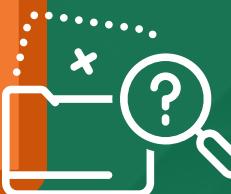
# DATA UNDERSTANDING

Data	Keterangan	Tipe Data
order date (DateOrders)	Tanggal pesanan yang dibuat	Numerik
Order Status	Tanggal pesanan yang dibuat	Kategorik
Order Item Discount Rate	Persentase diskon item dalam pesanan	Numerik
Product Price	Harga produk	Numerik
Order Profit Per Order	Keuntungan per pesanan	Numerik
Category Name	Deskripsi kategori produk	Kategorik

# DATA PREPARATION

## Missing Value

Dataset ini tidak memiliki nilai yang hilang (missing value), sehingga seluruh data dapat digunakan secara utuh untuk analisis tanpa perlu proses imputasi atau penanganan nilai yang kosong.



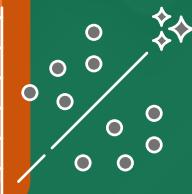
## Duplikat Data

Pada dataset ini juga tidak terdapat data yang duplikat.



## Deteksi Outlier

Terdapat outlier dalam dataset ini yaitu pada variabel Profit Per Order, namun outlier tidak dihapus.



# FEATURE ENGINEERING



## Fitur 1: Customer Lifetime Value (CLV)

Fitur ini menghasilkan nilai numerik yang bertujuan untuk mengelompokkan pelanggan berdasarkan nilai keuntungan jangka panjang yang mereka hasilkan.

## Fitur 2: Average Order Value (AOV)

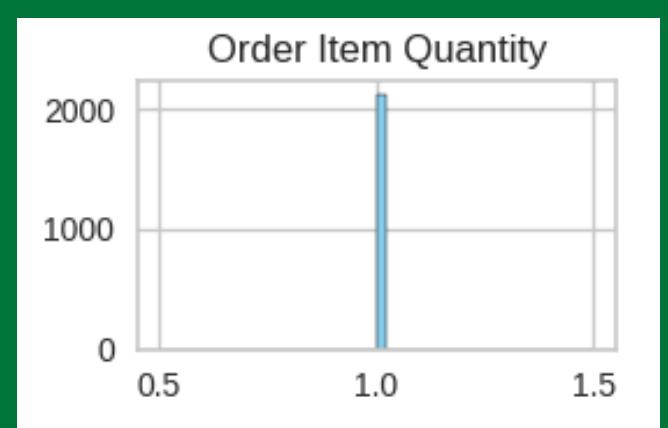
Fitur ini menghasilkan nilai numerik yang bertujuan untuk mengelompokkan pelanggan berdasarkan seberapa besar mereka biasanya membelanjakan uang per transaksi.

## Fitur 3: Discount Sensitivity

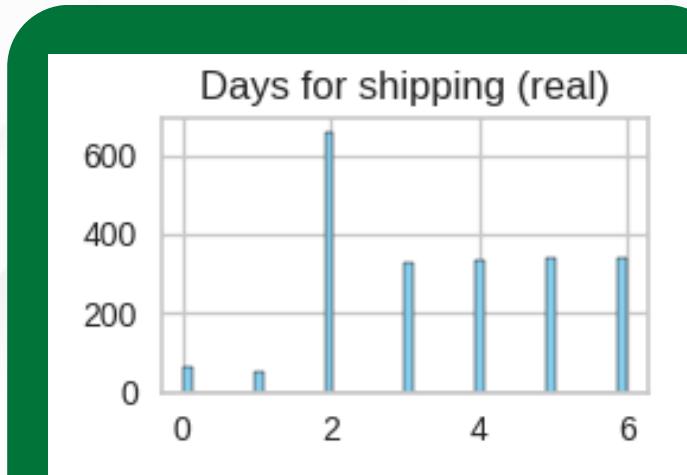
Fitur ini menghasilkan nilai numerik yang bertujuan untuk mengidentifikasi pelanggan yang lebih cenderung berbelanja saat ada diskon.



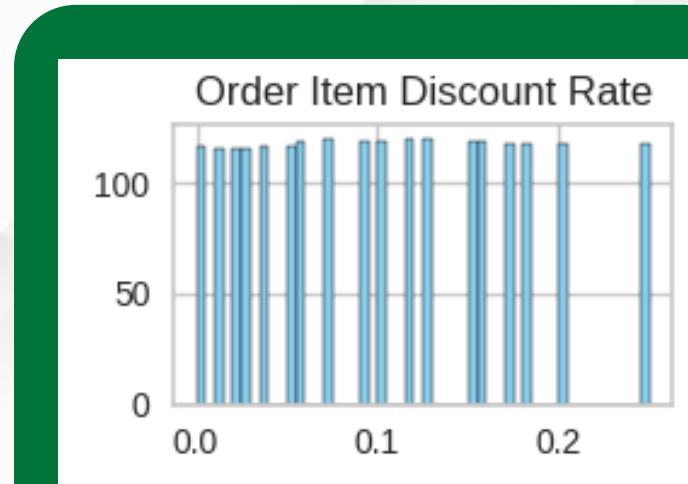
# EXPLORATORY DATA ANALYSIS (EDA)



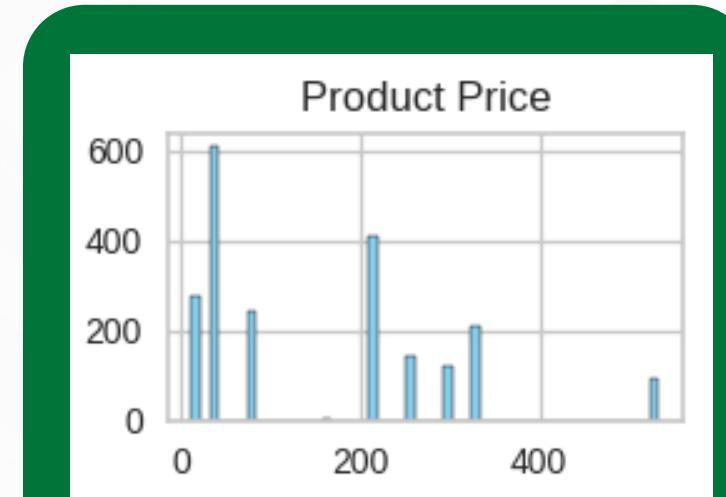
**Order Item Quantity**  
Sebagian besar pesanan hanya memiliki 1 item, menunjukkan konsentrasi tinggi pada pembelian tunggal per transaksi.



**Days for shipping (real)**  
Mayoritas pengiriman dilakukan dalam 2–5 hari, dengan sedikit pengiriman yang memakan waktu lebih lama.

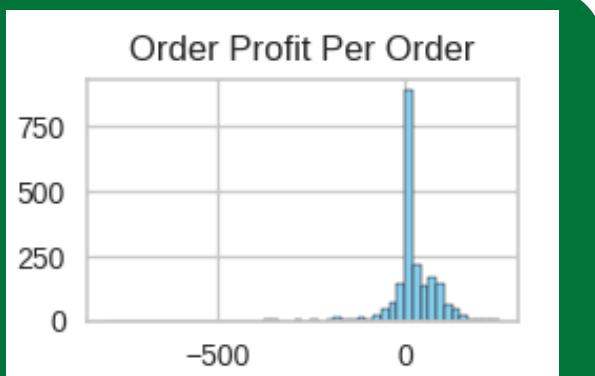


**Order Item Discount Rate**  
Sebagian besar transaksi memiliki tingkat diskon tertentu yang sering muncul, menunjukkan pola pemberian diskon yang konsisten.

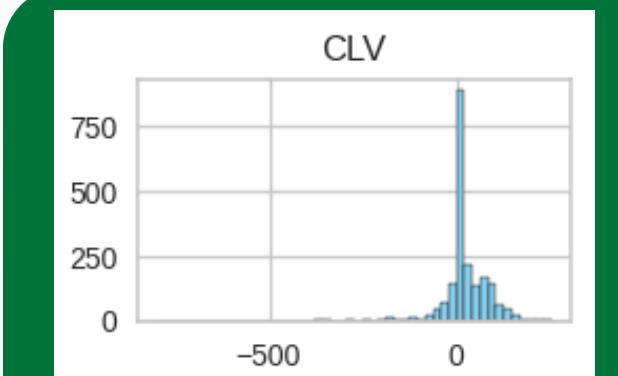


**Product Price**  
Mayoritas produk memiliki harga rendah, dengan hanya sedikit produk berharga tinggi (distribusi miring ke kanan).

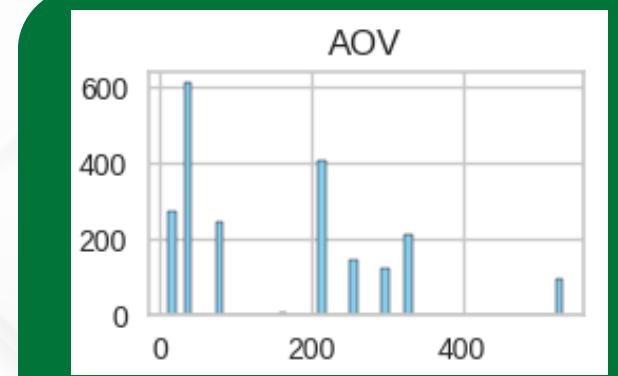
# EXPLORATORY DATA ANALYSIS (EDA)



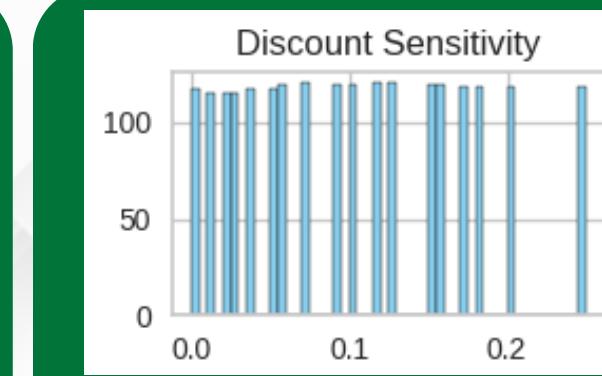
**Order Profit Per Order**  
Sebagian besar pesanan memiliki keuntungan kecil atau nol, dengan beberapa mengalami kerugian besar.



**Customer Lifetime Value (CLV)**  
Majoritas pelanggan memiliki CLV rendah, hanya sedikit pelanggan bernilai tinggi yang berkontribusi besar.



**Average Order Value (AOV)**  
Sebagian besar pesanan bernilai rendah hingga menengah, dengan sedikit pesanan bernilai sangat tinggi.



**Discount Sensitivity**  
Tingkat sensitivitas terhadap diskon tersebut merata, tanpa pola dominan di antara pelanggan.



**Korelasi Antar-Fitur**  
Sebagian besar fitur memiliki korelasi lemah, menunjukkan bahwa fitur-fitur tersebut berdiri sendiri dan tidak saling memengaruhi secara signifikan.

# DATA SCALING



**Data Scaling adalah proses transformasi data untuk memastikan bahwa semua fitur dalam dataset berada pada skala yang sama atau serupa.**

## RobustScaler



Hasil transformasi dengan RobustScaler menunjukkan bahwa Order Profit Per Order memiliki variasi besar, termasuk nilai ekstrem. Setelah normalisasi, data memiliki median sekitar nol, menjadikannya lebih tahan terhadap outlier. Contohnya, nilai 91.25 menjadi 1.46, sementara -249.09 menjadi -4.89.

## StandardScaler



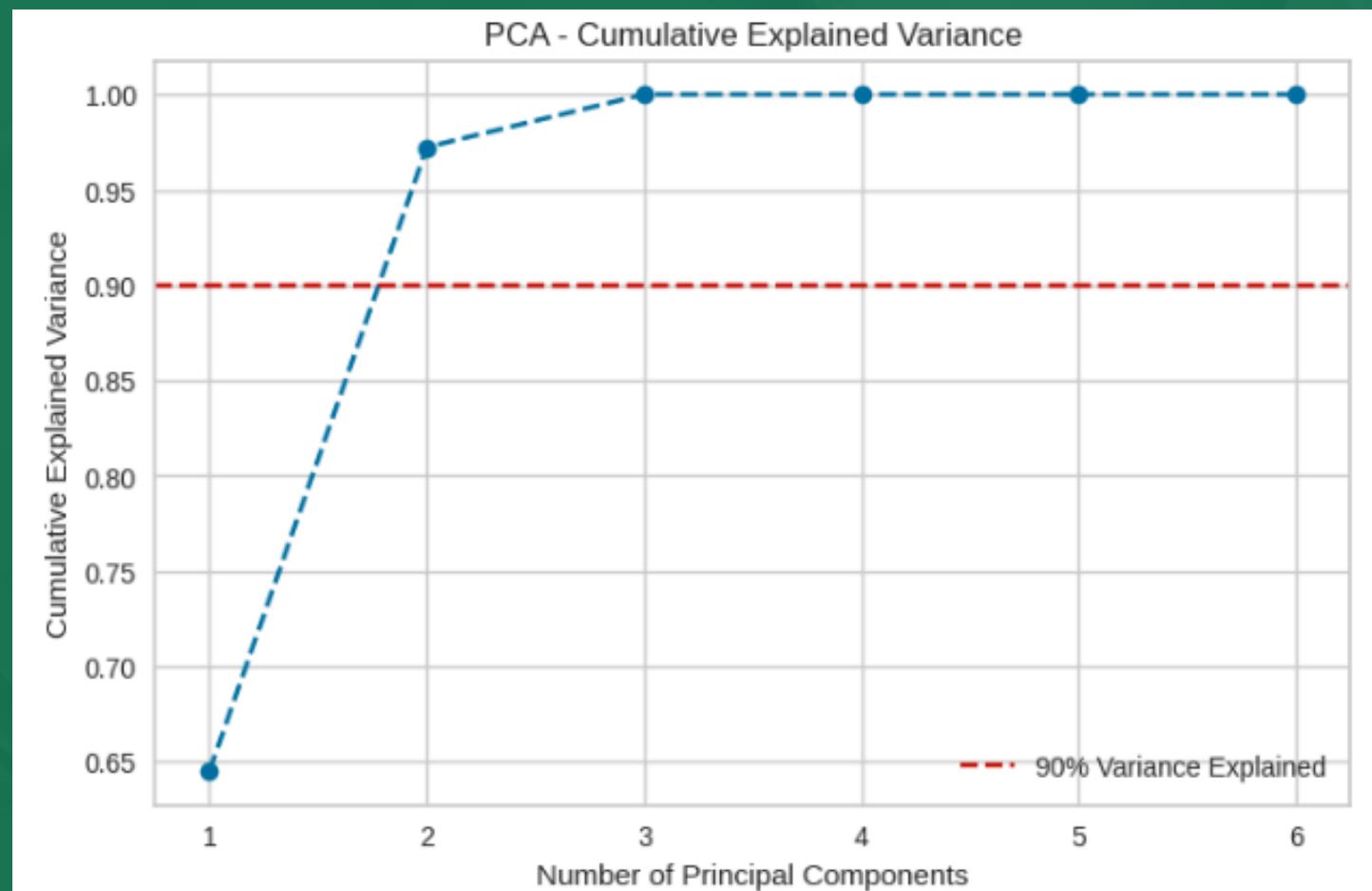
Hasil transformasi dengan StandardScaler menunjukkan bahwa perbedaan skala antar fitur seperti Product Price, CLV, dan AOV berhasil diatasi dengan menyamakan rata-rata menjadi nol dan standar deviasi satu. Contohnya, Product Price 327.75 menjadi 1.24, dan CLV 91.25 serta -249.09 menjadi 0.88 dan -3.10. AOV tetap konstan karena nilainya sama di semua baris.

## MinMaxScaler



MinMaxScaler dipakai untuk Order Item Discount Rate dan Discount Sensitivity, mengubah nilai ke rentang 0 hingga 1. Hasil dari transformasi ini mempertahankan perbandingan antar nilai. Misalnya, nilai 0.04 diubah menjadi 0.16, mencerminkan peningkatan proporsional dari nilai asli.

# PRINCIPAL COMPONENT ANALYSIS (PCA)



## PC1 (Komponen Utama 1)

Komponen ini dipengaruhi oleh dua variabel utama yaitu Order Profit per Order dan Customer Lifetime Value (CLV). Artinya, **komponen ini menggambarkan tingkat profitabilitas pelanggan**. Semakin tinggi nilainya, semakin menguntungkan pelanggan tersebut bagi perusahaan.

## PC2 (Komponen Utama 2)

Komponen ini dipengaruhi oleh dua variabel utama yaitu Harga Produk dan Average Order Value (AOV). Artinya, PC2 lebih merepresentasikan kebiasaan belanja pelanggan, khususnya terkait dengan harga produk dan jumlah pengeluaran rata-rata.

Dua komponen utama (PC1 dan PC2) cukup representatif untuk menggambarkan pola utama dalam perilaku pelanggan. Komponen selanjutnya (PC3 - PC6) memiliki kontribusi yang sangat kecil terhadap variasi data, sehingga tidak digunakan dalam analisis lebih lanjut.

# MODELING K-MEANS

Hasil clustering dengan k-means:

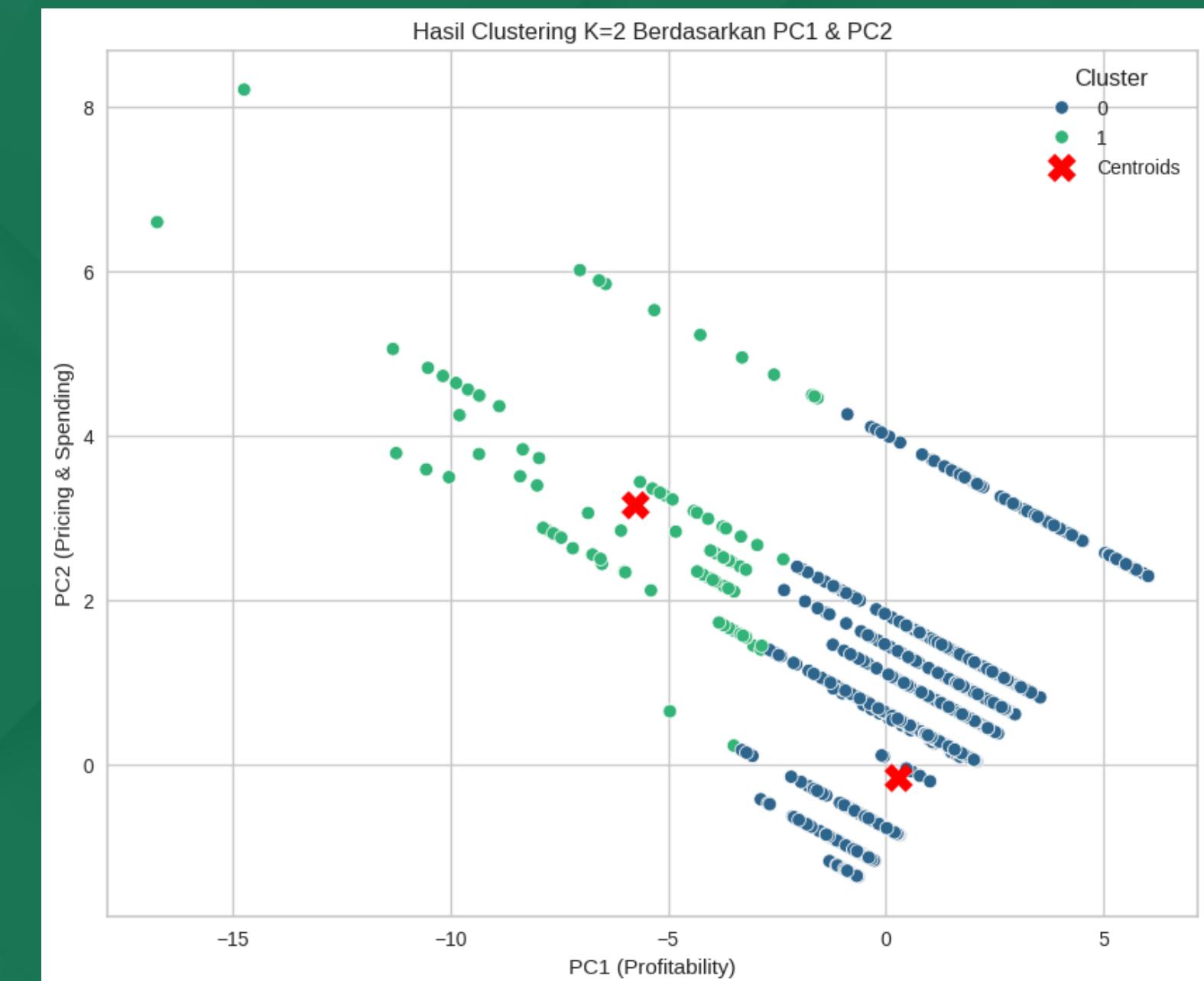
- Score Silhouette: 0,7071
- Cluster yang terbentuk: 2

## ➤ CLUSTER 0:

Pelanggan dengan profitabilitas tinggi dan pengeluaran stabil.

## ➤ CLUSTER 1:

Pelanggan dengan profitabilitas rendah dan pengeluaran tinggi.



# MODELING AGGLOMERATIVE HIERARCHICAL

Hasil clustering dengan Agglomerative Hierarchical :

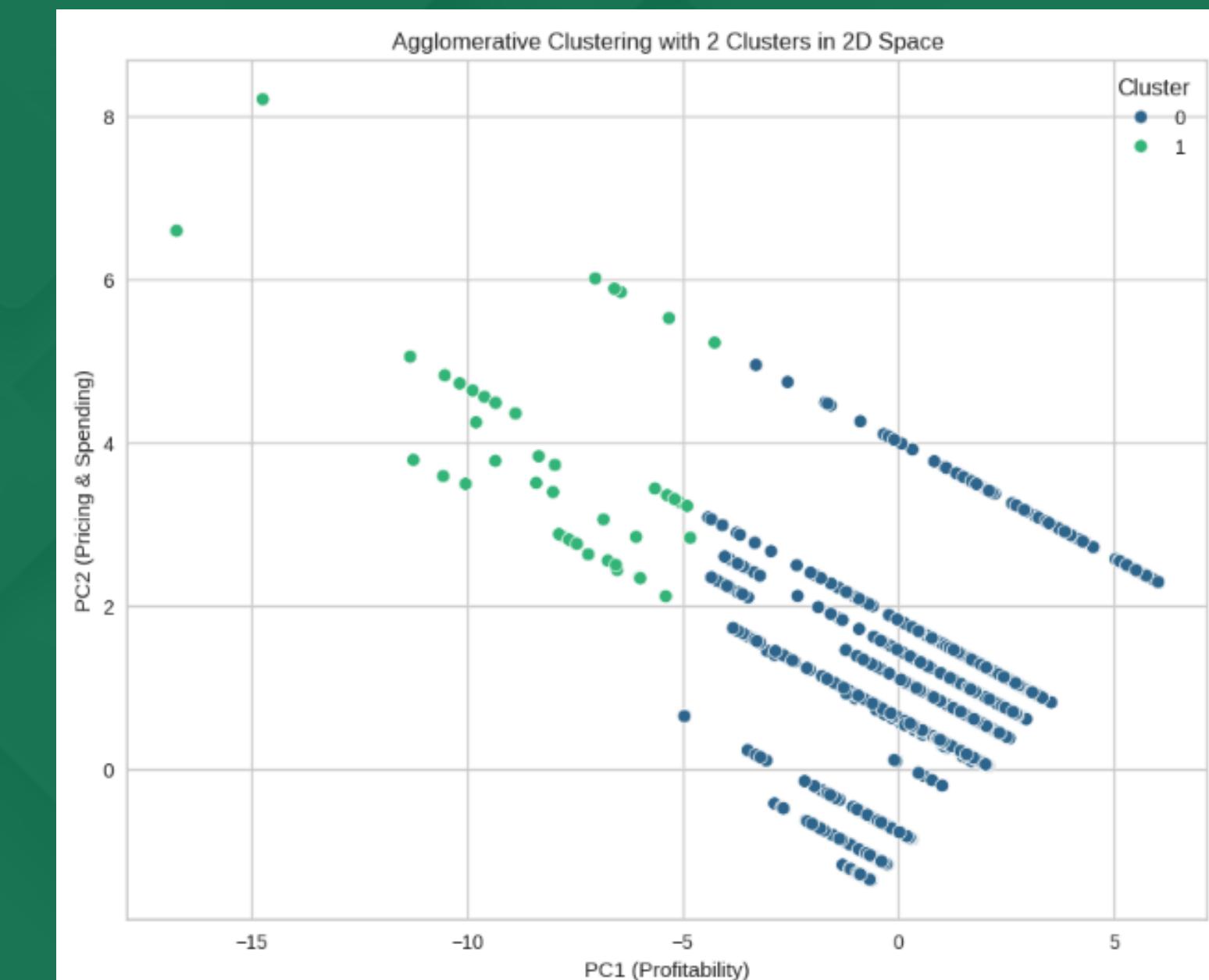
- Silhouette score : 0.7592
- Cluster yang terbentuk : 2 Cluster

## ➤ CLUSTER 0:

Pelanggan dengan profitabilitas tinggi dan pengeluaran stabil.

## ➤ CLUSTER 1:

Pelanggan dengan profitabilitas rendah dan pengeluaran tinggi.

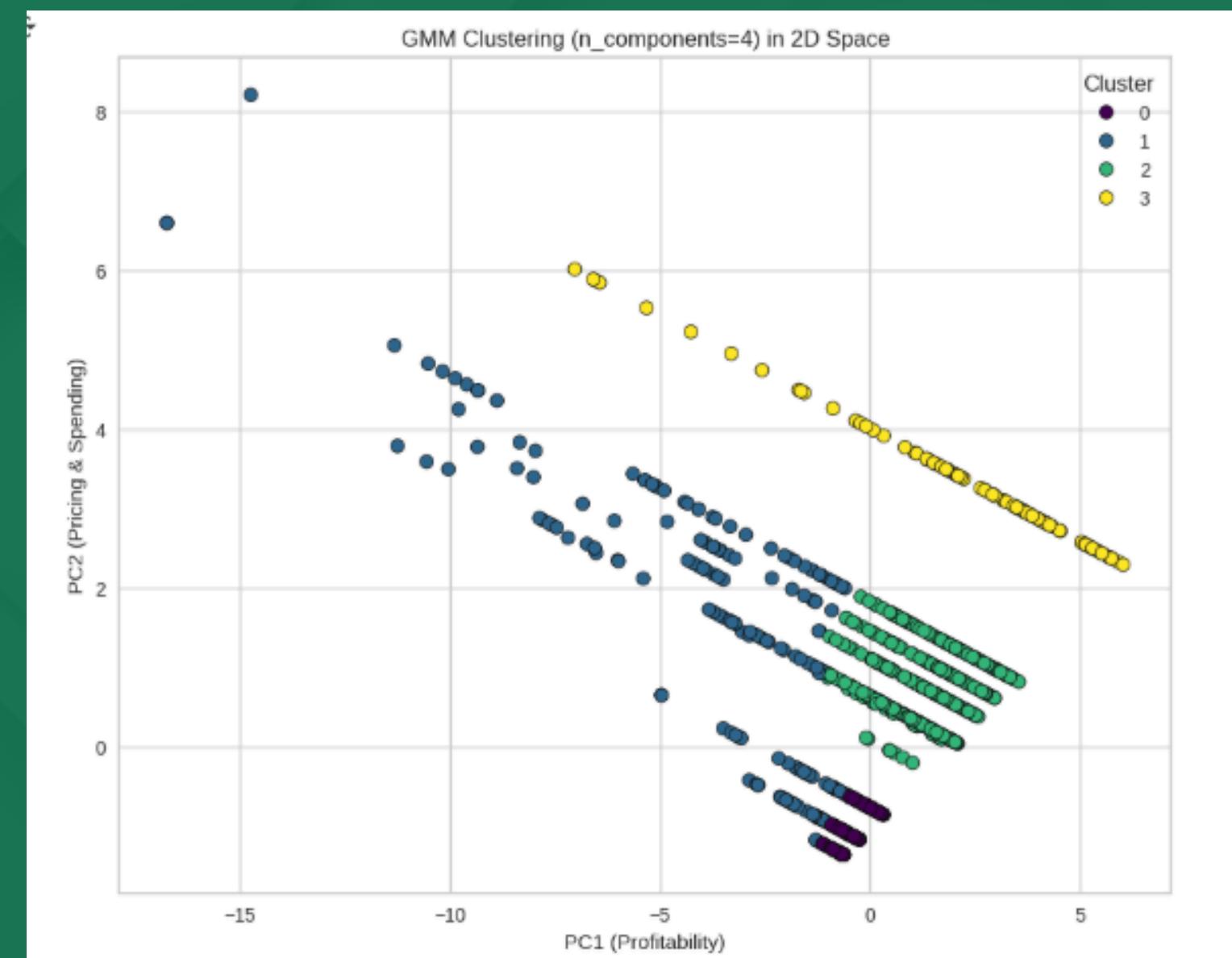


# MODELING GAUSSIAN MIXTURE MODEL (GMM)

Hasil clustering dengan GMM :

- Silhouette score : 0,57006
- Cluster yang terbentuk : 4 cluster

- Cluster 0:  
Pelanggan dengan profitabilitas rendah dan pengeluaran sedang.
- Cluster 1:  
Pelanggan dengan profitabilitas rendah dan pengeluaran tinggi.
- Cluster 2:  
Pelanggan dengan profitabilitas menengah dan pengeluaran rendah.
- Cluster 3:  
Pelanggan dengan profitabilitas tinggi dan pengeluaran tinggi.



# MODEL TERBAIK

**Hasil clustering dengan Agglomerative Hierarchical :**

- Silhouette score: 0.7592
- Cluster yang terbentuk : 2 cluster
- Karakteristik setiap cluster :

Cluster	PC1	PC2
0	Tinggi	Rendah
1	Rendah	Tinggi

**Alasan memilih Agglomerative Hierarchical sebagai model terbaik:**

- Hasil Silhouette score lebih baik dibanding 2 model lainnya, yaitu 0.7592
- Tidak membutuhkan jumlah cluster awal
- Fleksibilitas dalam bentuk cluster
- Informasi hierarkis
- tidak sensitif terhadap outlier
- cocok untuk data dengan struktur kompleks

# BUSINESS SOLUTION



**“COSTAR SOLUTION”**



## C Customer Profitability Optimization

Mengoptimalkan profitabilitas pelanggan dengan strategi berbasis data: loyalitas ditingkatkan untuk pelanggan profit tinggi (Cluster 0) dan edukasi serta insentif finansial untuk pelanggan profit rendah (Cluster 1).

## O ffering Personalized Value Propositions

Memberikan penawaran yang disesuaikan: layanan VIP & cashback untuk pelanggan loyal (Cluster 0), serta diskon dinamis & promosi berbasis perilaku untuk pelanggan sensitif harga (Cluster 1).

## S mart Pricing & Spending Control

Mengatur strategi harga berdasarkan kebiasaan belanja: harga eksklusif & akses prioritas untuk pelanggan loyal (Cluster 0), serta opsi cicilan & bundling hemat untuk pelanggan dengan kontrol belanja ketat (Cluster 1).

## T echnology-Driven Customer Engagement

Memanfaatkan teknologi seperti AI untuk meningkatkan interaksi: layanan AI eksklusif & rekomendasi personal bagi pelanggan loyal (Cluster 0), serta notifikasi harga & aplikasi keuangan untuk pelanggan dengan kontrol belanja (Cluster 1).

## A daptive Marketing Strategies

Menyesuaikan strategi pemasaran sesuai karakteristik cluster: storytelling & influencer untuk pelanggan brand-loyal (Cluster 0), promosi sosial media & referral program untuk pelanggan sensitif harga (Cluster 1).

## R etention & Relationship Building

Membangun hubungan jangka panjang: program loyalitas eksklusif bagi pelanggan setia (Cluster 0), dan komunitas pelanggan aktif untuk engagement dan rekomendasi hemat (Cluster 1).

# **"COSTAR" SOLUTION**

Mengerti Konsumen, Driving Better Business Decisions

**SMART  
BUSINESS**



# THANK YOU



# **ANALISIS TREN SEGMENTASI PELANGGAN DALAM MENYESUAIKAN STRATEGI BISNIS YANG BERORIENTASI PADA KONSUMEN**

