

Project Pemodelan Statistika

# Optimasi Strategi Pemasaran dengan Segmentasi Pelanggan

# PROJECT BACKGROUND

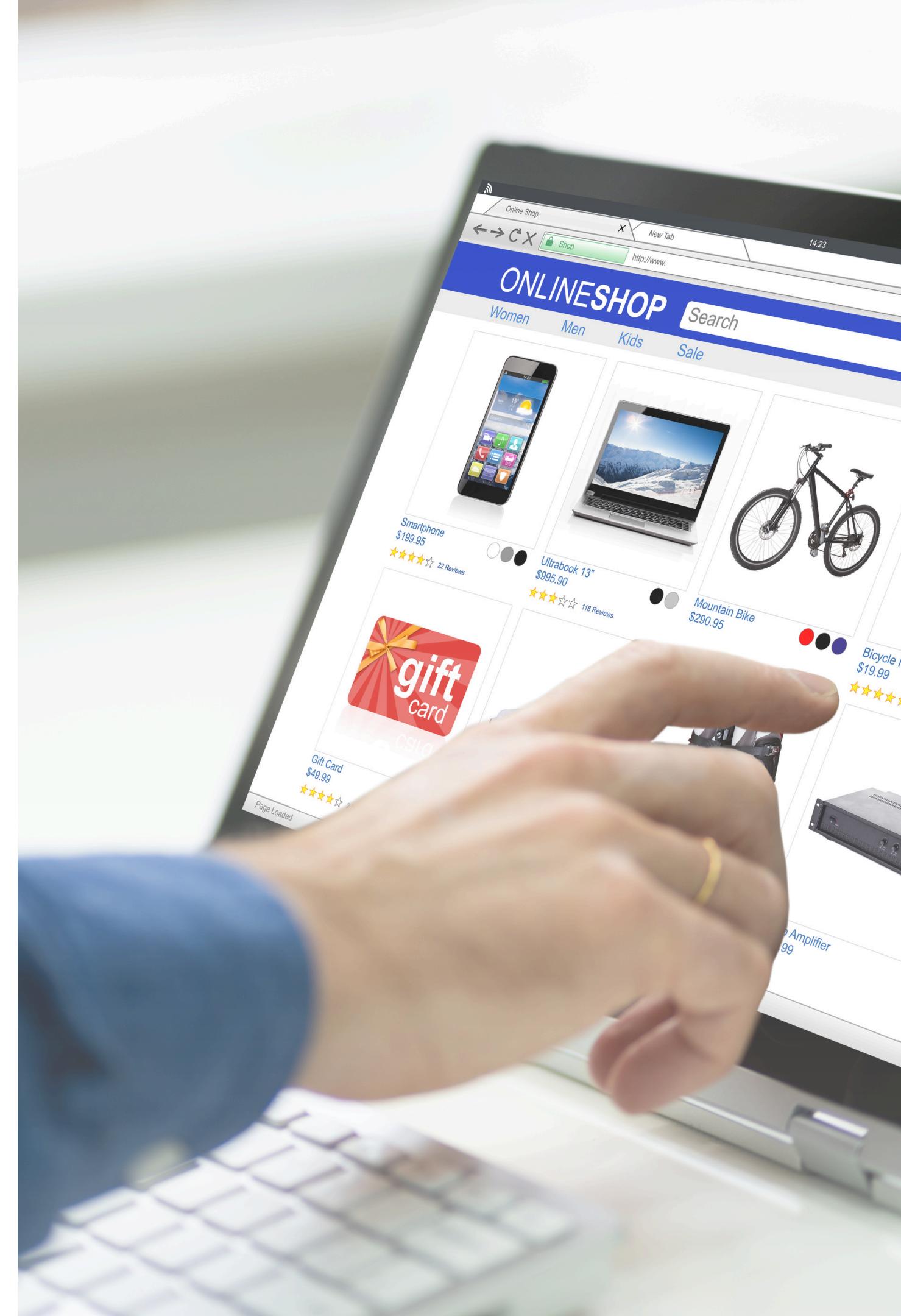
- Proyek ini menggunakan K-Means Clustering untuk mengelompokkan pelanggan berdasarkan perilaku pembelian dan karakteristiknya.
- Mengidentifikasi segmen pelanggan guna membantu perusahaan mengoptimalkan penjualan

## TUJUAN / OBJEKTIF BISNIS

Mengidentifikasi kelompok pelanggan berdasarkan perilaku dan karakteristik pembelian, sehingga perusahaan **dapat mengoptimalkan strategi pemasaran, meningkatkan loyalitas pelanggan, dan memaksimalkan profitabilitas.**

## TUJUAN TEKNIS DATA SCIENCE

Mengelompokan pelanggan menggunakan K-Means Clustering berdasarkan perilaku pembelian.



# RISIKO DAN ALTERNATIF PELAKSANAAN

Risk	Mitigasi
Kualitas Data yang Buruk	Melakukan pembersihan dan validasi data secara menyeluruh sebelum analisis, serta menggunakan teknik imputasi untuk menangani data yang hilang.
Pemilihan Jumlah Cluster yang Tidak Tepat	Menggunakan metode elbow dan silhouette score untuk menentukan jumlah cluster yang optimal sebelum menerapkan K-Means Clustering
Perusahaan enggan untuk menerapkan strategi baru berdasarkan hasil analisis	Melakukan evaluasi berkala dan validasi hasil klasterisasi dengan pemangku kepentingan
Kluster sulit dipahami, sehingga kurang memberikan wawasan untuk keputusan.	Menggabungkan hasil clustering dengan analisis deskriptif untuk memberikan konteks yang lebih baik dan memvisualisasikan hasil melalui grafik dan diagram.

# KRITERIA KESUKSESAN PROYEK

Hasil klaster yang berbeda **memiliki karakteristik unik masing-masing klaster**.

# TELAAH DATA - 1

## DESKRIPSI DATASET

Superstore data merupakan data penjualan retail dari perusahaan Superstore di Amerika Serikat.

<b>Identitas Pelanggan</b>	Customer ID, Customer Name, Segment, Country, City, State, Postal Code, Region
<b>Data pesanan dan pengiriaman</b>	Order ID, Order Date, Ship Date, Ship Mode
<b>Data produk</b>	Product ID, Product Name, Category, Sub-Category
<b>Data penjualan</b>	Sales, Quantity, Discount, Profit

**15** atribut

**9.994**  
record data

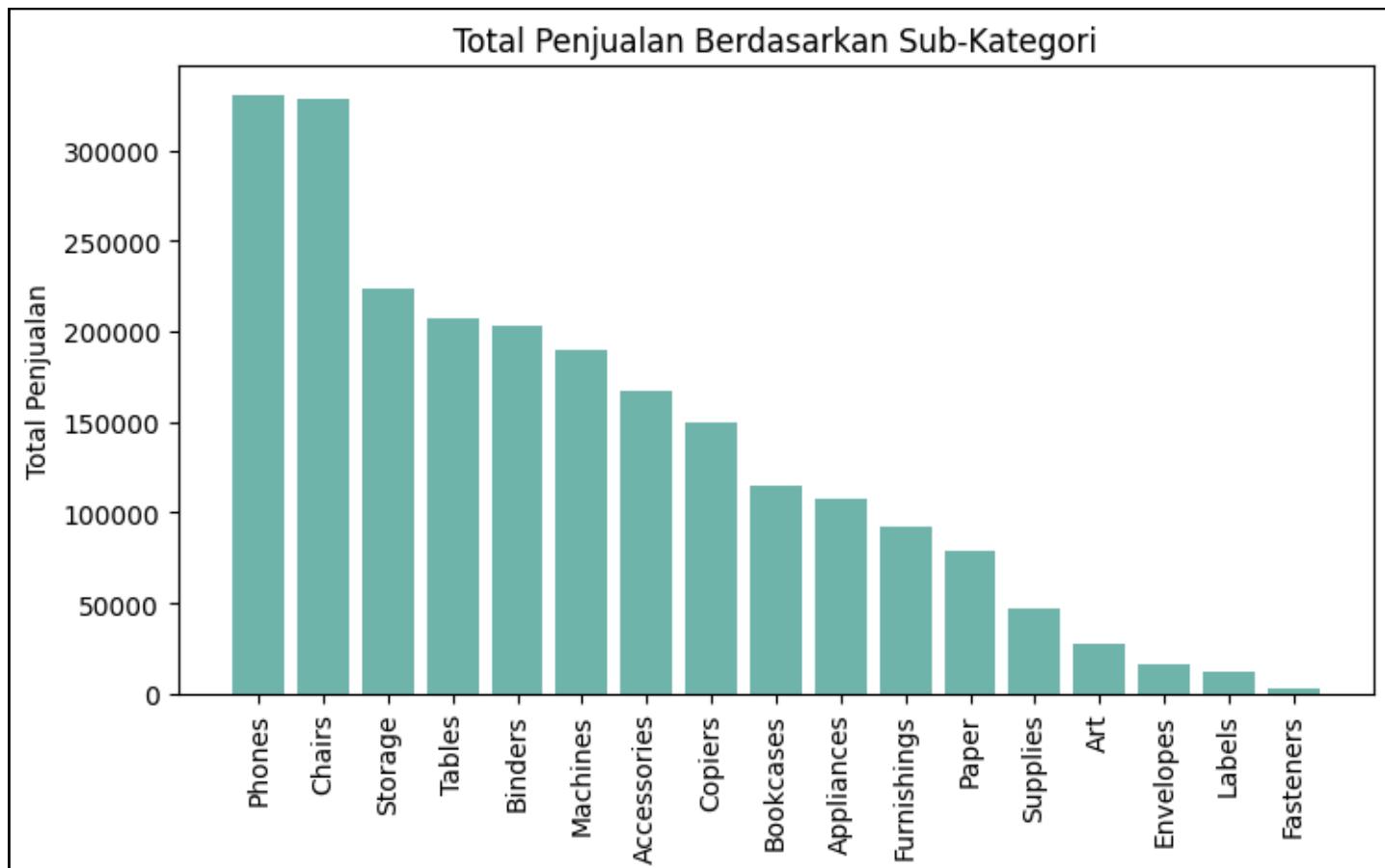
Sumber: <https://www.kaggle.com/datasets/vivek468/superstore-dataset-final>

## STATISTIKA DESKRIPTIF

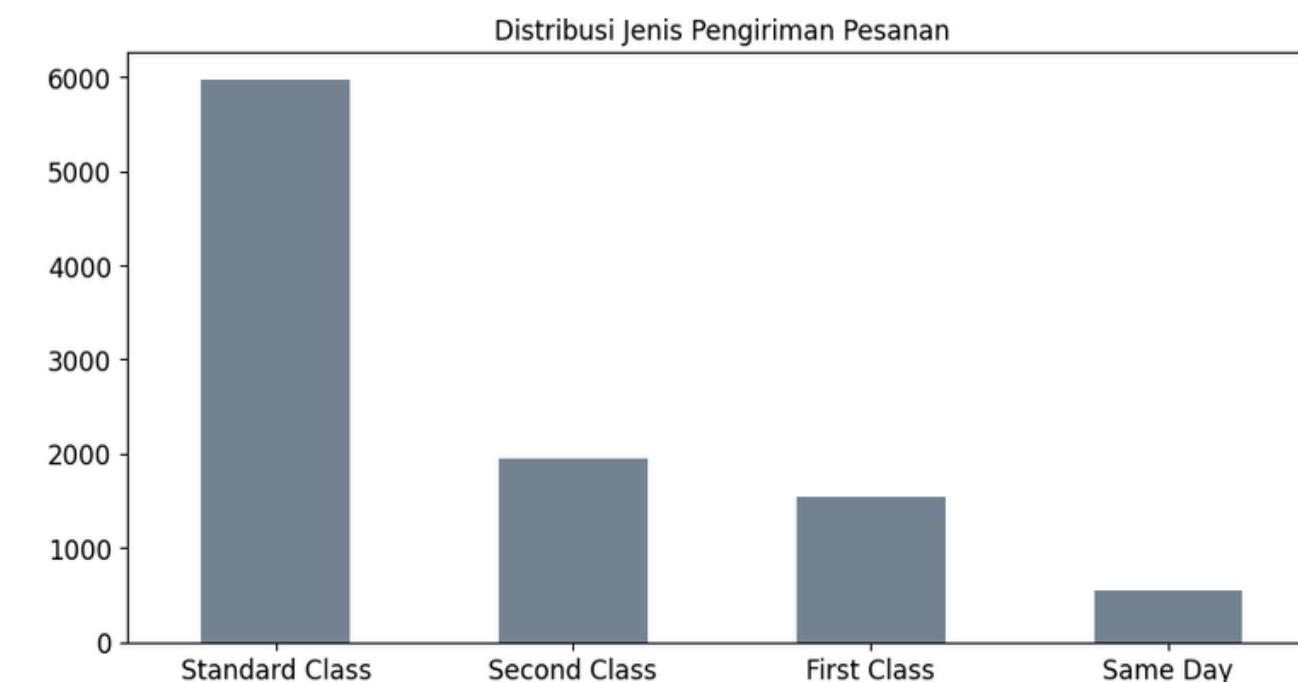
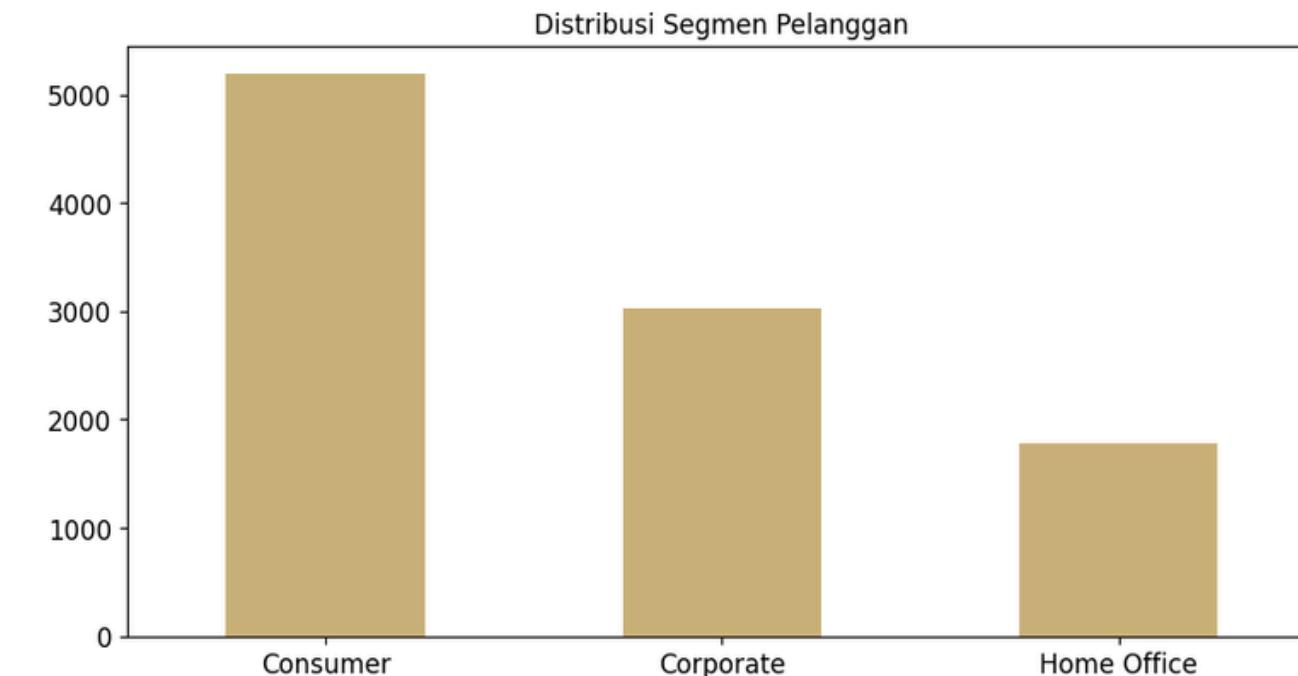
Index	Sales (US\$)	Quantity (unit)	Discount	Profit (US\$)
N	9994	9994	9994	9994
Mean	229.86	3.79	0.16	28.66
Std	623.25	2.23	0.21	234.26
Min	0.44	1.00	0.00	-6599.98
Max	22638.48	14.00	0.80	8399.98

# TELAAH DATA - 2

## EXPLORATORY DATA ANALYSYS



HP dan Kursi mencatat penjualan tertinggi, sementara Fasteners terendah. Keuntungan terbesar dari penjualan Mesin Fotocopy.



# VALIDASI DATA PREPROCESSING DAN PENENTUAN OBJEK DATA

**Missing value:**

0 data

**Duplikat data:**

0 data

**Feature engineering:**

- One Hot Encoding
- Transformasi data - Standarscaler

**Outlier:**

Penjualan 6 produk teknologi dalam 1 transaksi teridentifikasi sebagai outlier,  
namun wajar karena harga per unitnya tinggi.

**Hasil Penentuan Objek Data:**

- Ship Mode
- Segment
- State
- Discount
- Sales
- Category\_Furniture
- Category\_Office Supplies
- Category\_Technology
- Profit

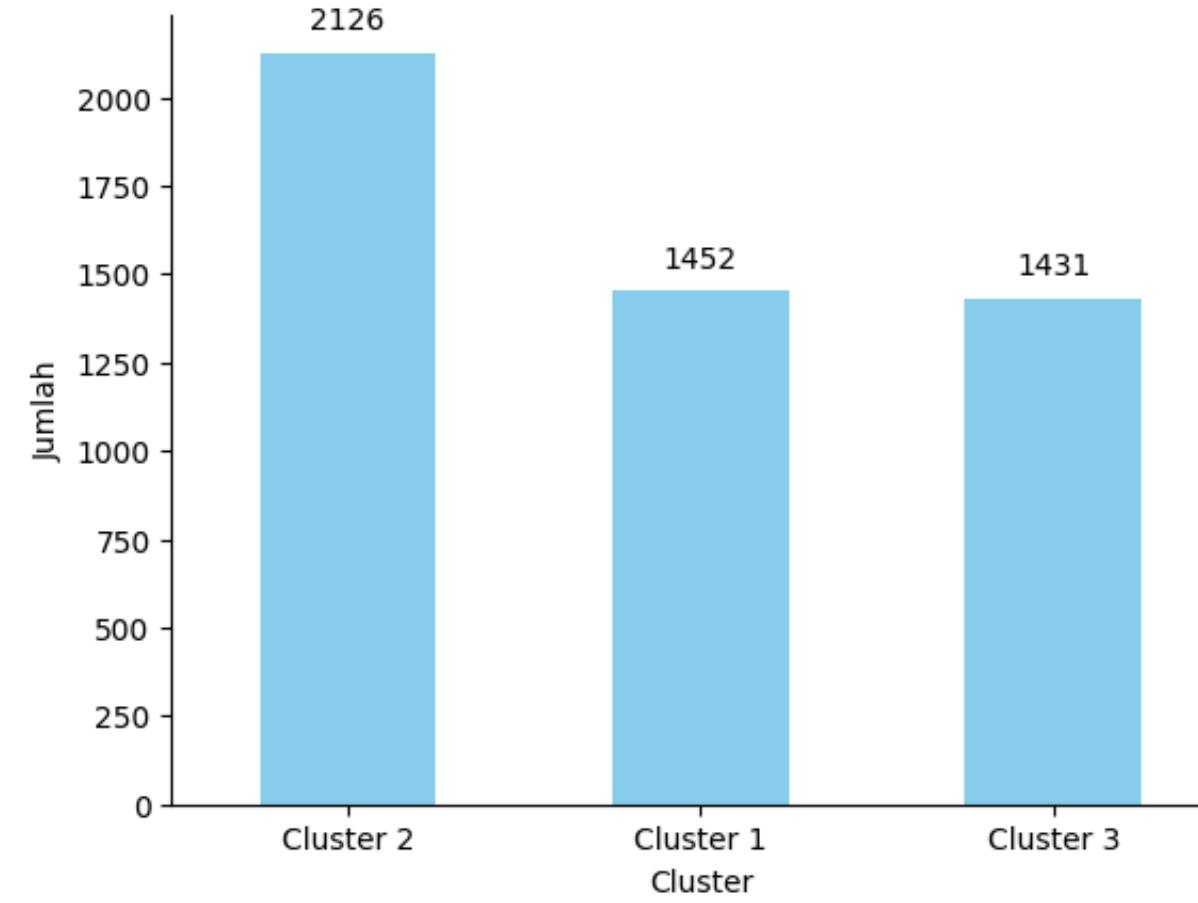
# PEMODELAN

## SKENARIO PEMODELAN

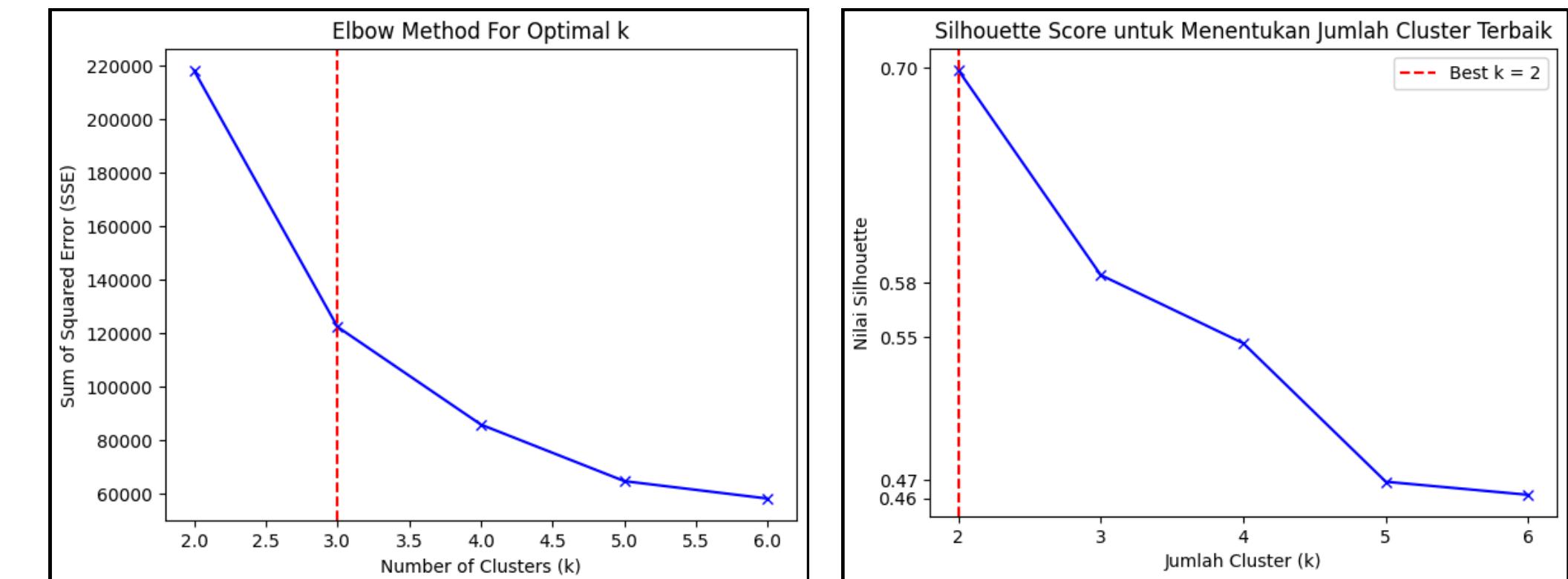


## HASIL CLUSTERING

Jumlah cluster yang ditentukan sebanyak 3 (K=3), dengan hasil silhouette score sebesar 0.58.



## PEMILIHAN K-OPTIMAL

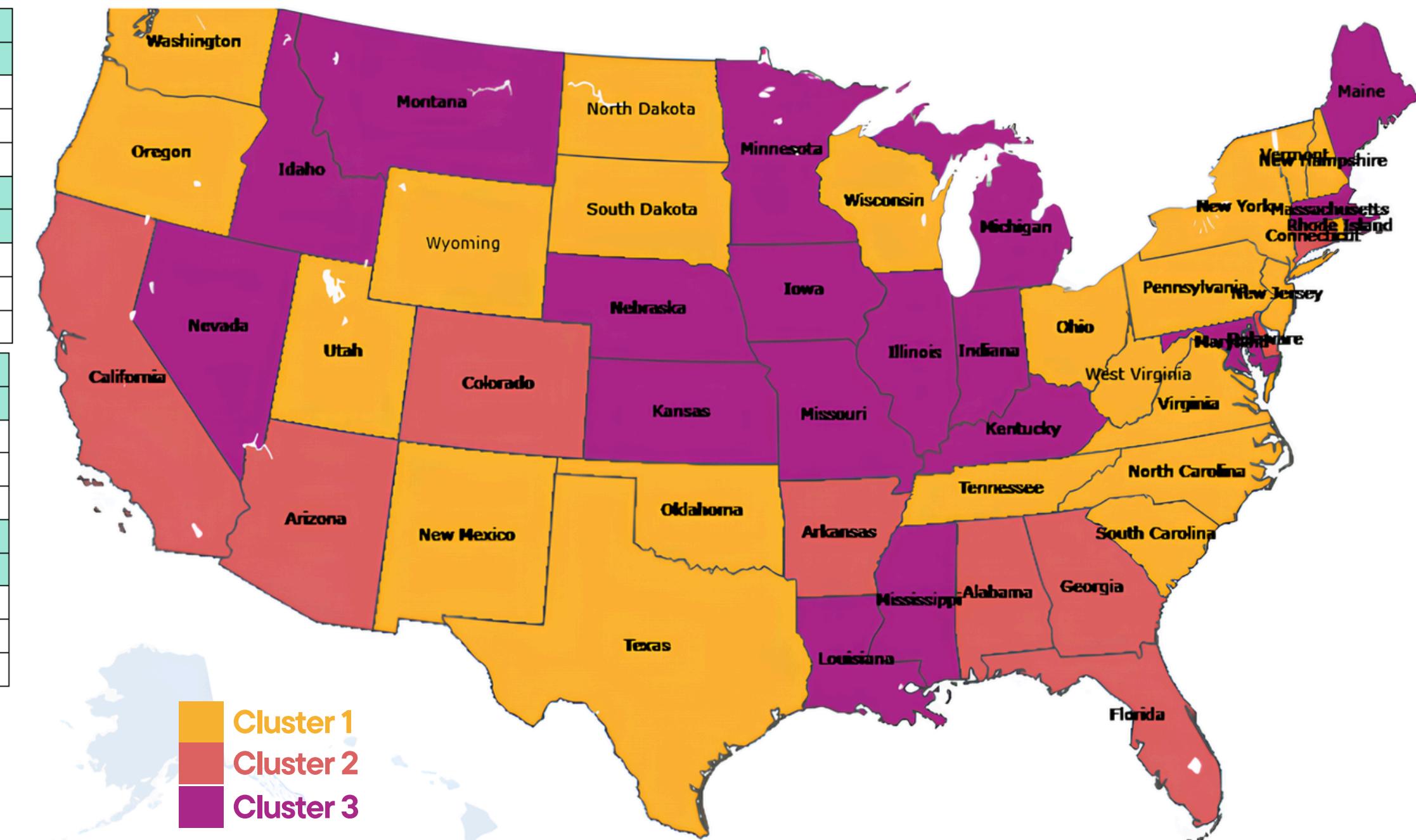


Analisis menunjukkan 3 cluster optimal berdasarkan Elbow Method, sesuai tujuan awal untuk membagi data ke dalam minimal 3 segment strategis. Meski SC menunjukkan 2 cluster lebih seragam, keputusan memilih 3 cluster guna mendukung kebutuhan bisnis.

# PEMODELAN

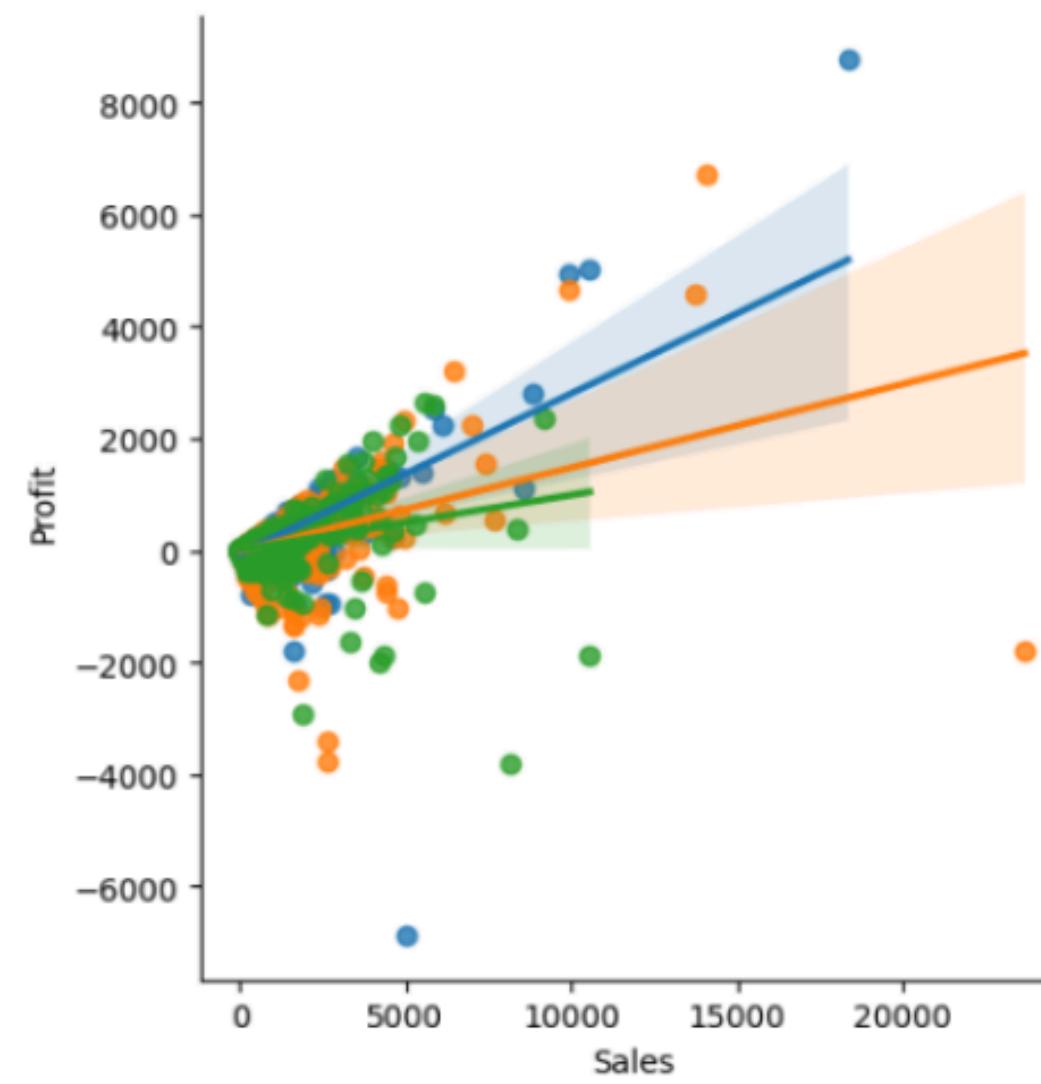
## EXPLORATORY DATA ANALYSYS HASIL KLASTERISASI

Sales (\$)				
Cluster	Mean	Median	Max	Min
Cluster 1	450.677	155.942*	18,336.740	1.080*
Cluster 2	470.431*	149.965	23,661.228*	0.556
Cluster 3	449.114	152.590	10,539.896	0.852
Quantity				
Cluster	Mean	Median	Max	Min
Cluster 1	7.452	6	43	1
Cluster 2	7.614*	6	52*	1
Cluster 3	7.613	6	45	1
Profit (\$)				
Cluster	Mean	Median	Max	Min
Cluster 1	66.713*	16.981	8,762.389*	-6,892.375
Cluster 2	54.656	15.538	6,734.472	-3,791.163*
Cluster 3	51.244	17.220*	2,640.480	-3,825.339
Discount				
Cluster	Median	Max	Min	
Cluster 1	0.2	3.7	0	
Cluster 2	0.2	4.0*	0	
Cluster 3	0.2	3.2	0	

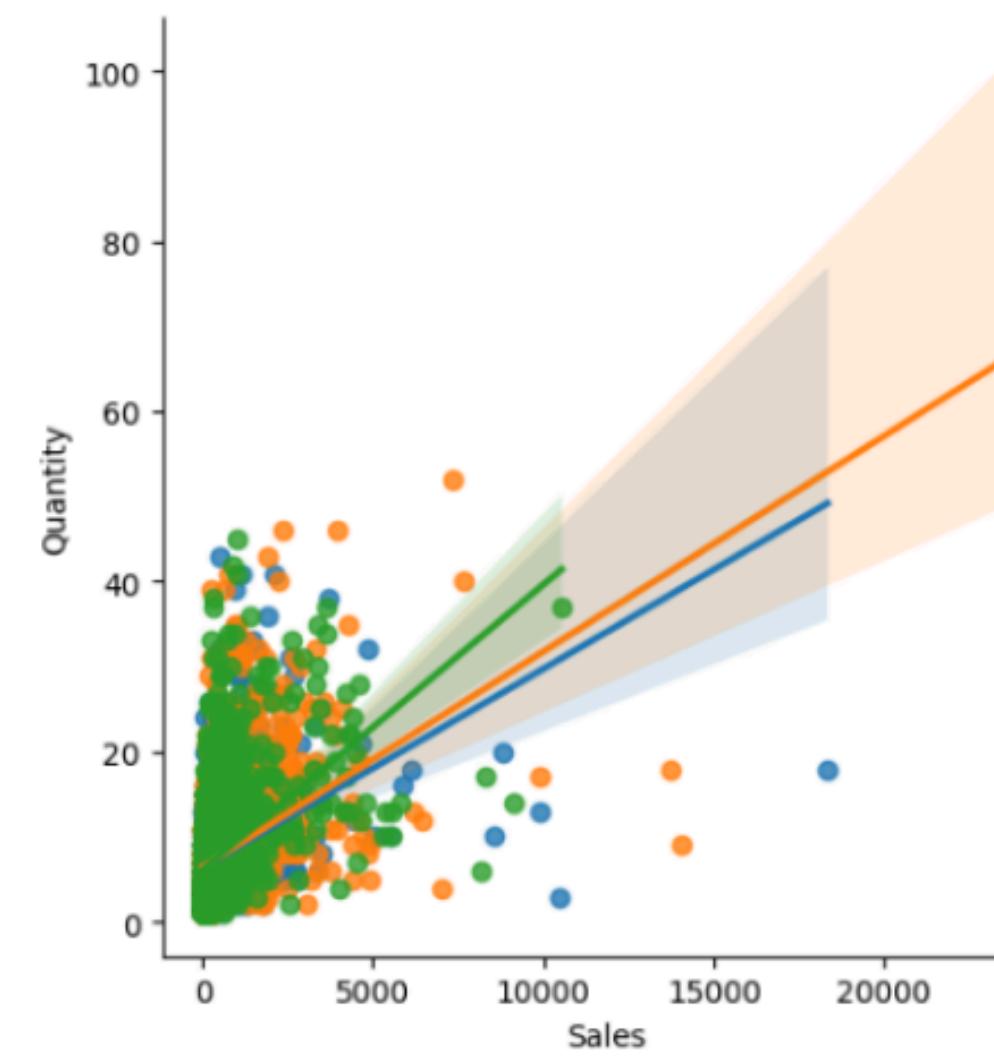


# PEMODELAN EXPLORATORY DATA ANALYSYS HASIL KLASTERISASI

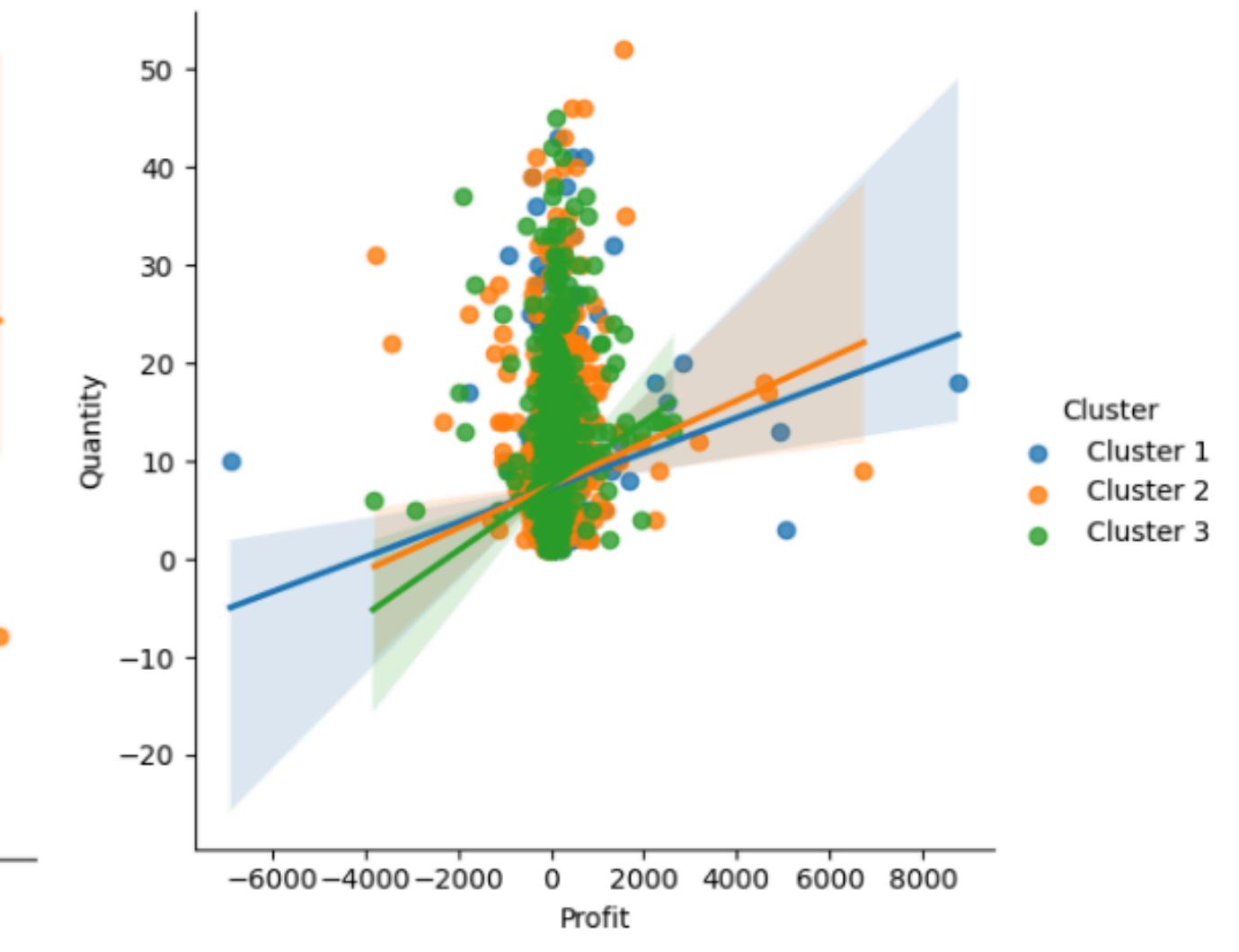
## Sales vs Profit



## Sales vs Quantity



## Profit vs Quantity

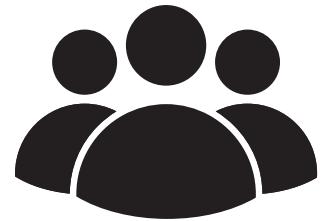


# PEMODELAN

## EXPLORATORY DATA ANALYSYS



# PEMODELAN KARAKTERISTIK CLUSTER



## Cluster 1

- Jumlah pelanggan relatif sedikit, mayoritas dari segmen **Consumer** dengan preferensi pengiriman **Standard Class**.
- Sub-category **Office Supplies** paling banyak dibeli.
- Klaster ini memiliki karakteristik pelanggan yang berkontribusi signifikan pada laba tetapi dengan potensi risiko yang tinggi.

## Cluster 2

- Jumlah pelanggan terbesar, segmen didominasi **Corporate** dan **Home Office** dominan. Pengiriman **Same Day** sering dipilih untuk pesanan besar.
- Sub-category produk **Office Supplies** paling utama sering dibeli, namun pembelian Sub-Category **Technology** dan **Furniture** cukup signifikan juga.
- Klaster ini memiliki pola pembelian yang intensif dengan volume besar serta responsif terhadap penawaran harga khusus, seperti diskon, yang dapat mendorong peningkatan volume pembelian.

## Cluster 3

- Jumlah pelanggan cukup besar, segmen pelanggan tertinggi **Consumer** dan **Corporate**, dengan jenis pengiriman sering digunakan **Standard Class** dengan pesanan dengan kuantitas kecil.
- Sub-category terjual utama **Furniture** dan **Technology**.
- Klaster ini memiliki karakteristik dengan variasi laba dan kerugian yang paling kecil, menunjukkan pelanggan yang lebih stabil namun kurang berkontribusi besar pada laba.

# PENUTUP

## REKOMENDASI STRATEGI PEMASARAN

### Cluster 1

- Pelanggan memberikan laba tinggi tetapi dengan risiko kerugian besar.
- Fokus pada strategi untuk meminimalkan risiko kerugian, seperti meningkatkan loyalitas melalui program point reward untuk setiap pembelian.
- Terapkan strategi upselling, misalnya dengan bundling produk Office Supplies (contoh: "Beli 3 gratis 1" untuk perlengkapan kantor kecil).

### Cluster 2

- Pelanggan memiliki potensi penjualan terbesar dengan jumlah item tinggi dan diskon besar.
- Fokus pada strategi diskon dan promosi untuk menjaga loyalitas dan meningkatkan keuntungan.
- Targetkan perusahaan dengan promosi kontrak jangka panjang untuk pasokan rutin.
- Berikan layanan eksklusif seperti pengiriman gratis untuk pesanan besar.
- Tingkatkan penjualan kategori produk lain seperti Furniture atau Technology melalui paket diskon.

### Cluster 3

- Pelanggan memiliki pola pembelian lebih konsisten dengan risiko kerugian lebih kecil.
- Lakukan kampanye pemasaran digital dengan fokus pada wilayah dominan (misalnya Nevada, Kentucky, Montana, dll).
- Maksimalkan profitabilitas dengan promosi diskon bundle untuk kategori Furniture dan Technology.

**TERIMA  
KASIH**



**[HTTPS://GITHUB.COM/ADHKARAA/PROJECT](https://github.com/adhkaraa/project)**