PaDi UMKM
Pasar Digital UMKM Indonesia

CUSTOMER SEGMENTATION

MILESTONE 2 - VIRTUAL INTERNSHIP

by – Ahmad Bukhari

CONTENTS

- Executive Summary
- Background & Objectives
- Data Overview
- Processing Steps
- Current Performance Evaluation
- Segmentation Results
  - Buyers
  - Sellers
- Recommendations

BACKGROUND

At PaDi UMKM, we're using data to fuel our mission of connecting Indonesian MSMEs with government buyers.

By analysing platform activity, we can forge stronger bonds with buyers through personalized experiences driven by customer segmentation, leading to more procurement and greater engagement.

Simultaneously, we aim to empower MSME sellers to thrive by understanding their unique needs and providing the right support at the right time, creating a productive environment for success.

OBJECTIVE

Leveraging data analysis at PaDi UMKM, we aim to achieve the following:

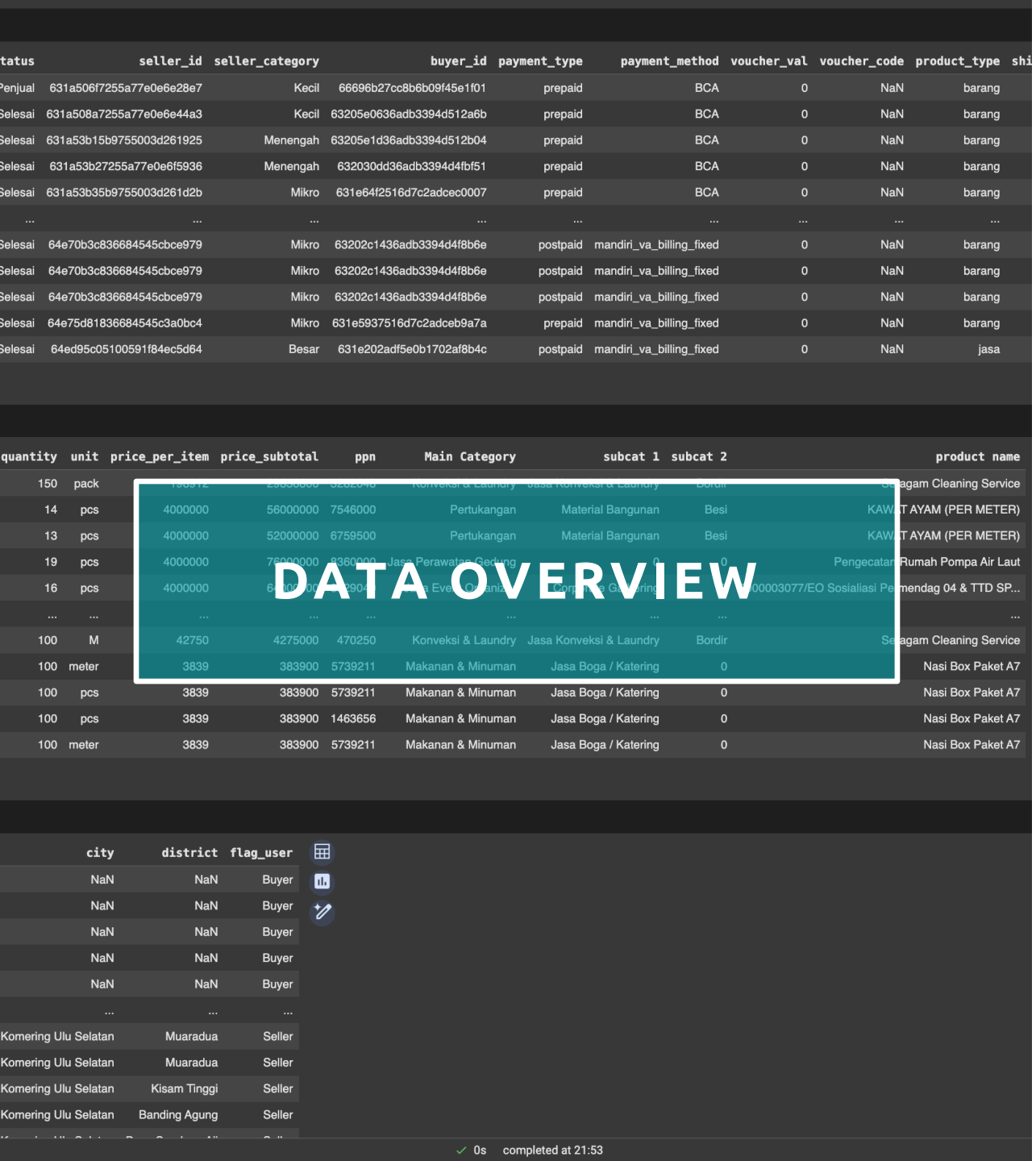
- **Enhance buyers relationships** through targeted approaches based on customer segmentation, leading to increased procurement and platform engagement.
- **Empower MSME sellers** by understanding their unique needs and providing optimized support, fostering a productive environment for their success on the platform.

Business Problem

- Leverage customer and seller segmentation to tailor strategies that improve engagement and business performance.

Goals

- Evaluate overall performance.
- Analyse performance of different buyer and seller groups.
- Propose actionable strategies to improve business outcomes.



Datasets

- **df_order**: Orders ID, GMV, seller and buyer IDs, payment and shipping details.
- **df_order_complete**: Product-level details of orders.
- **df_users**: Geographic and user type information.

Key Metrics Derived

- **Buyers**: GMV, Revenue, Orders, Recency.
- **Sellers**: GMV, Revenue, Orders Fulfilled, Unique Buyers.

The data provided is a representation of the original data, the amount of which has been adjusted.

PROCESSING STEPS

Data Preparation:

- Cleaning, merging, and preprocessing data from four datasets: orders, order details, users, and products.
- Create Performance overview charts

Feature Engineering:

- **Buyers**: Total GMV, total revenue, total orders, average order value, unique categories, voucher usage, recency.
- **Sellers**: Total GMV, total revenue, number of orders fulfilled, unique buyers served, return rate, recency.

Clustering:

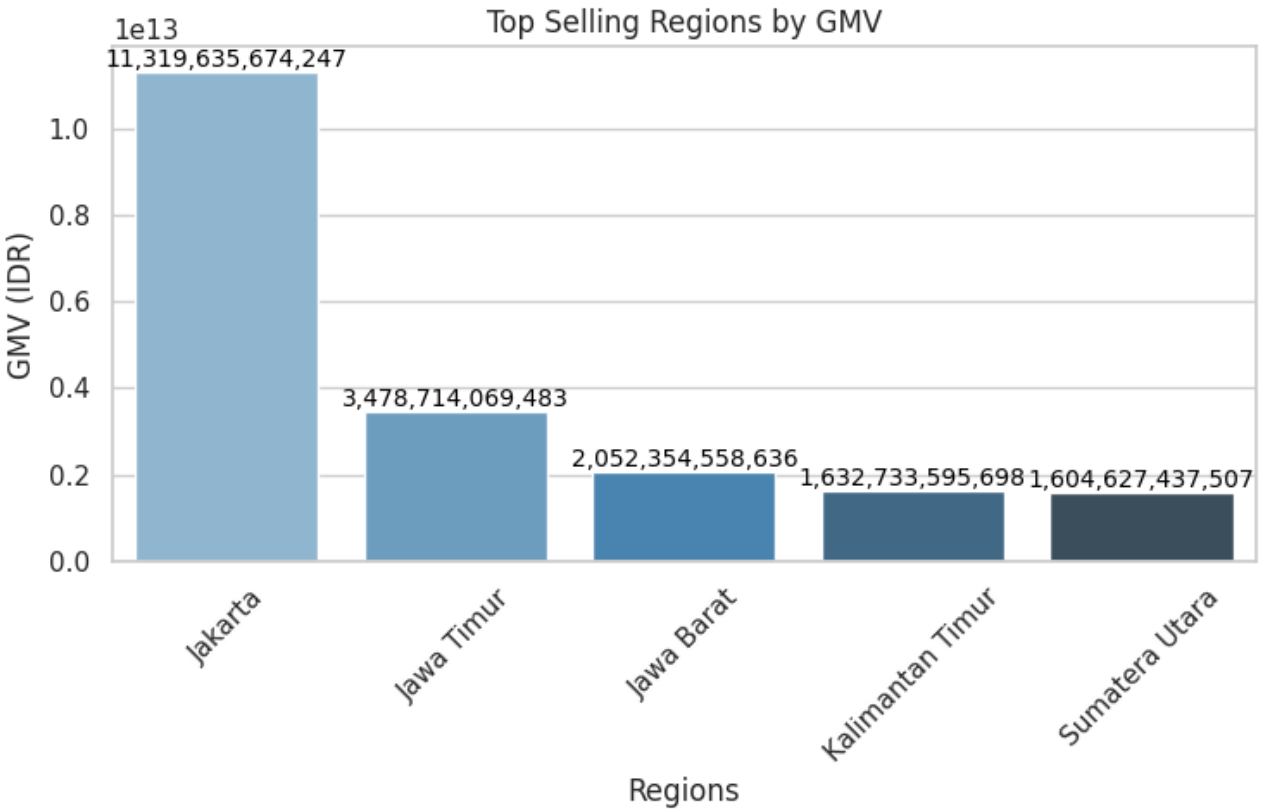
- Applied K-Means clustering to segment customers and sellers.

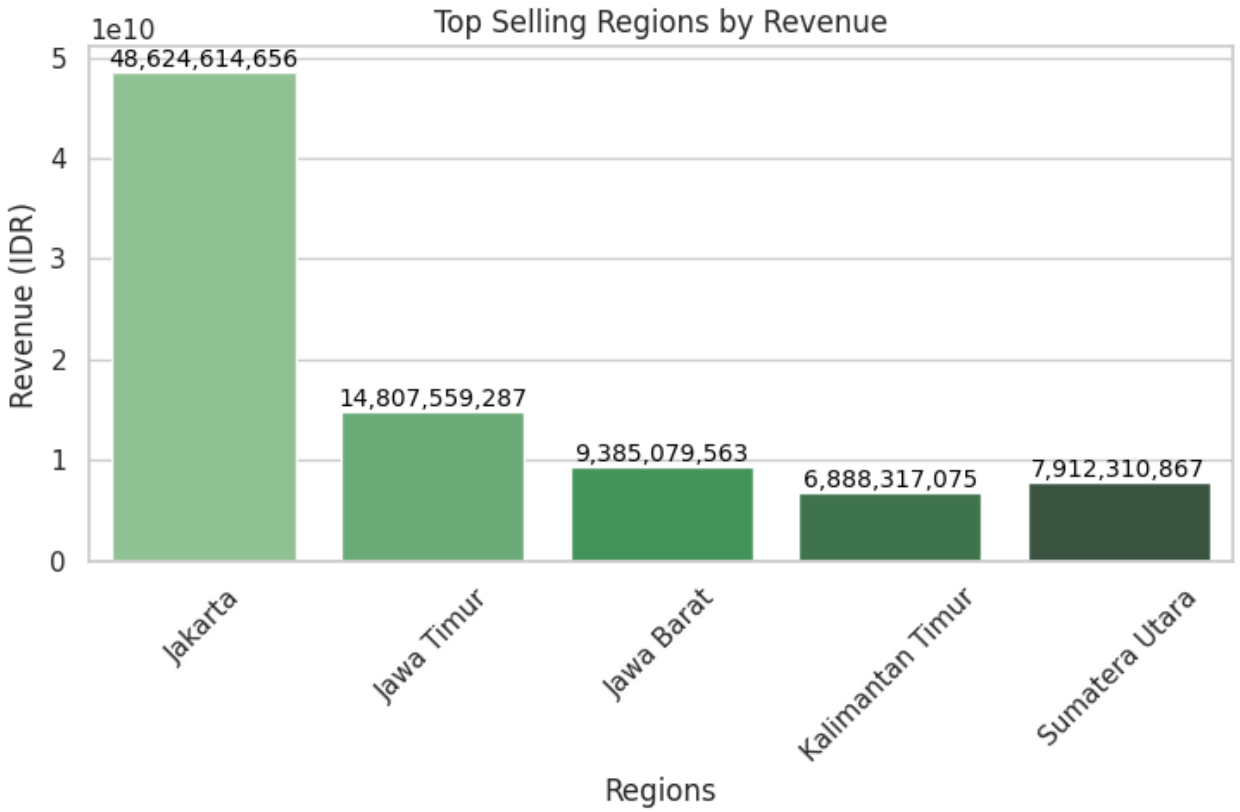
CURRENT PERFORMANCE

Business Metrics

- **Total GMV**: IDR 27,540,670,420,148
- **Total Revenue**: IDR 122,766,236,154
- **Total Orders**: 325,074

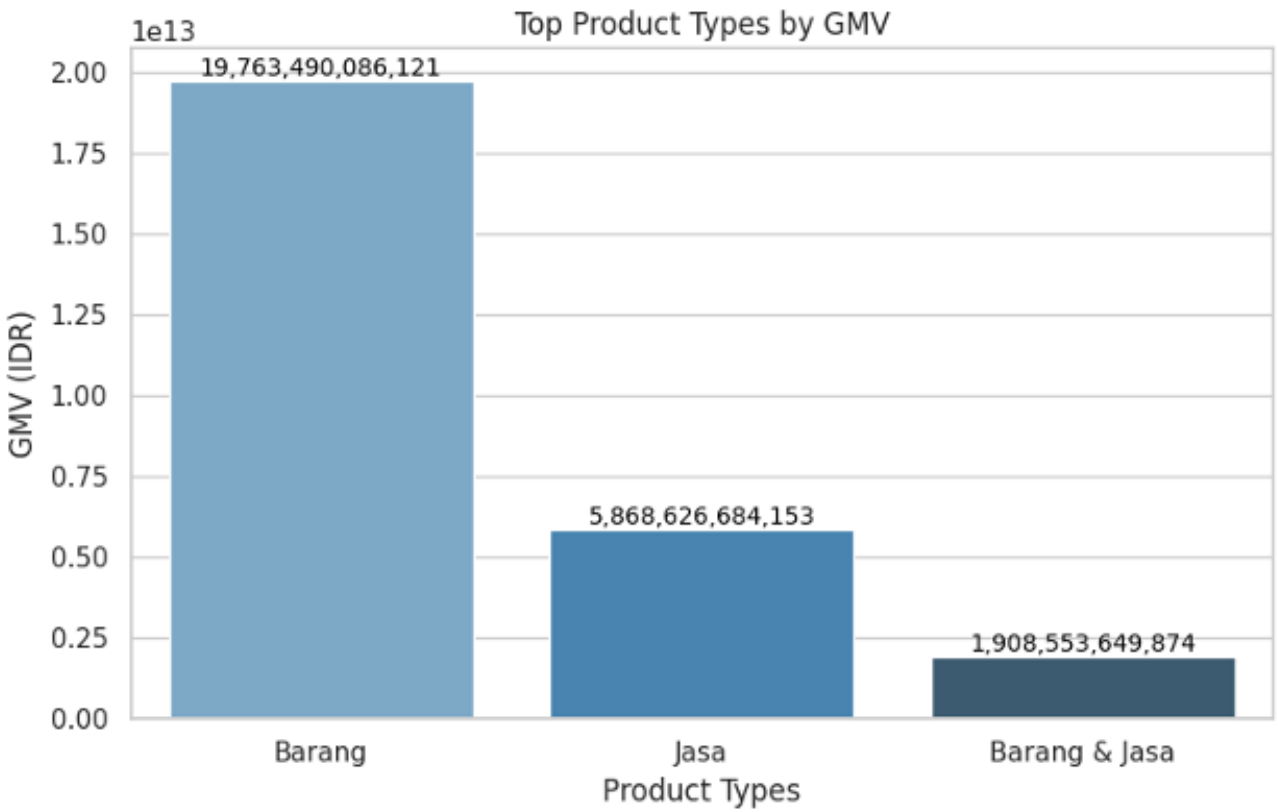
TOP REGION

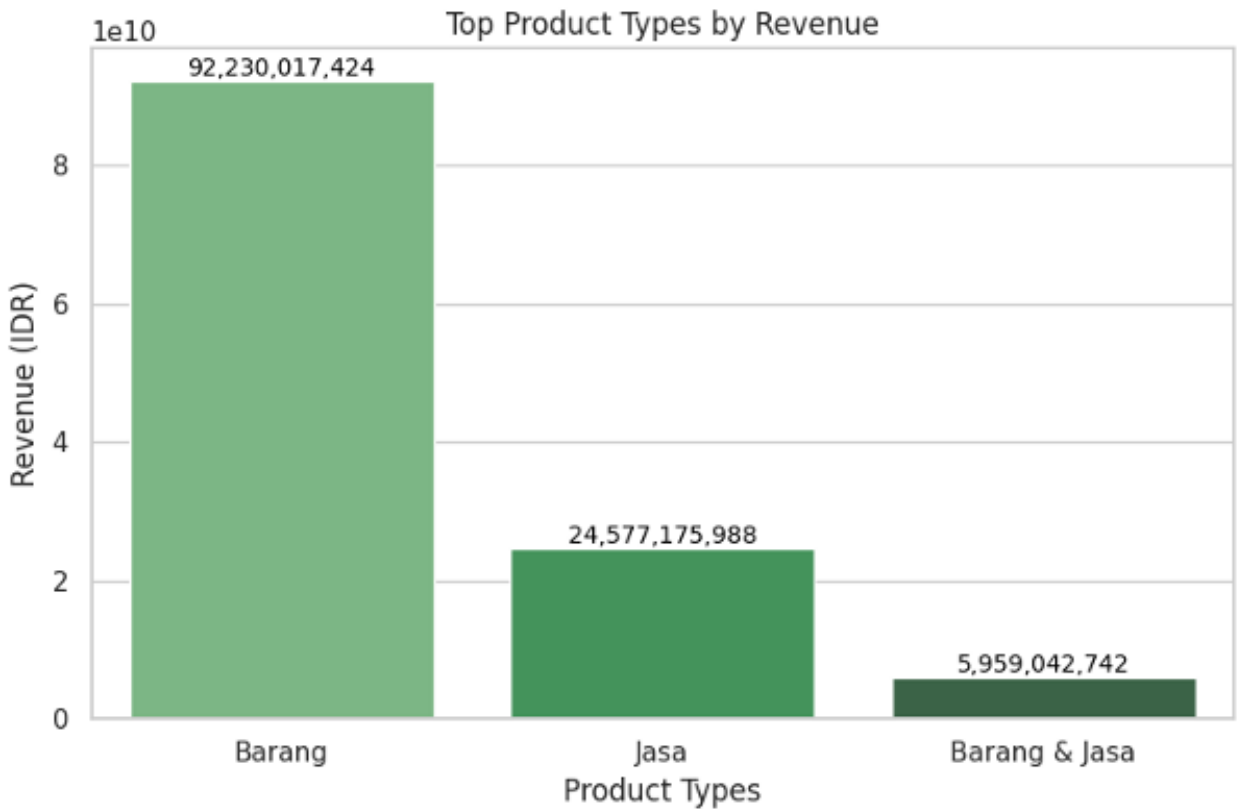




Jakarta is the highest-performing region, contributing nearly 41% of GMV and 40% of revenue.

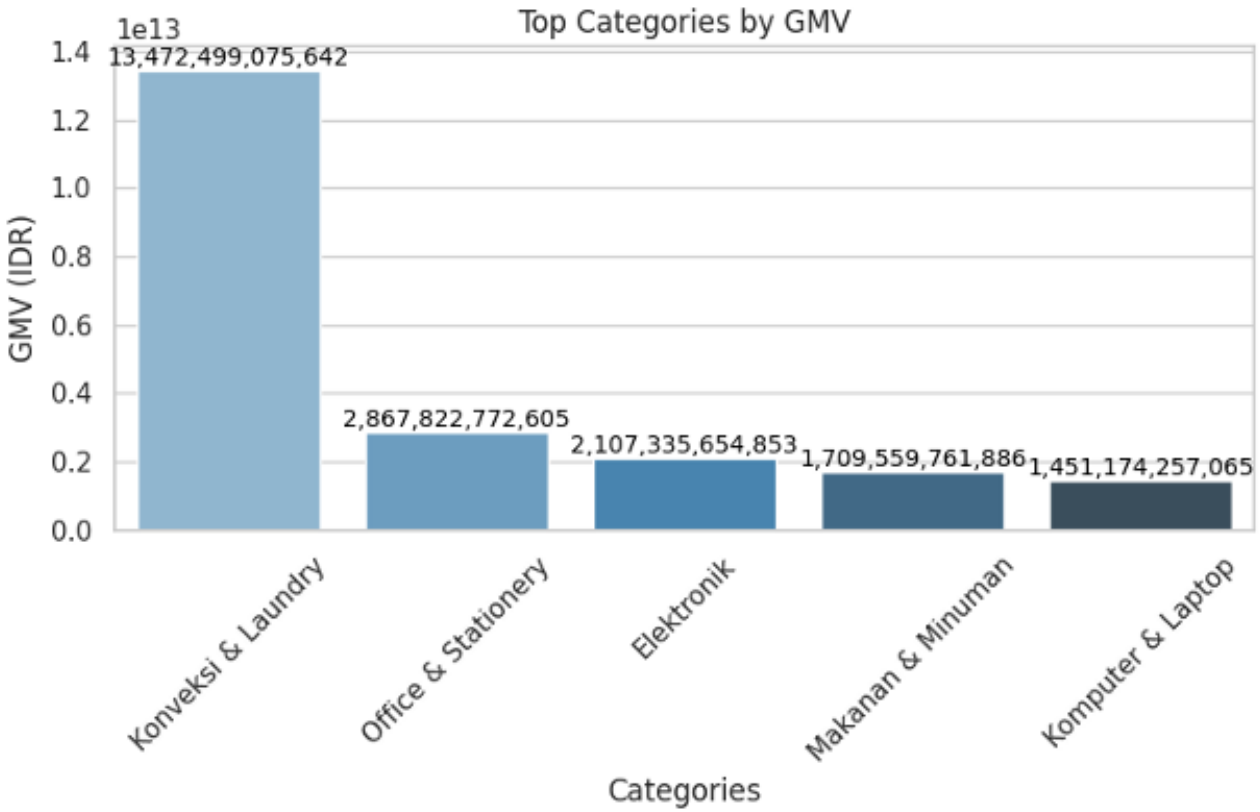
TOP PRODUCT TYPES

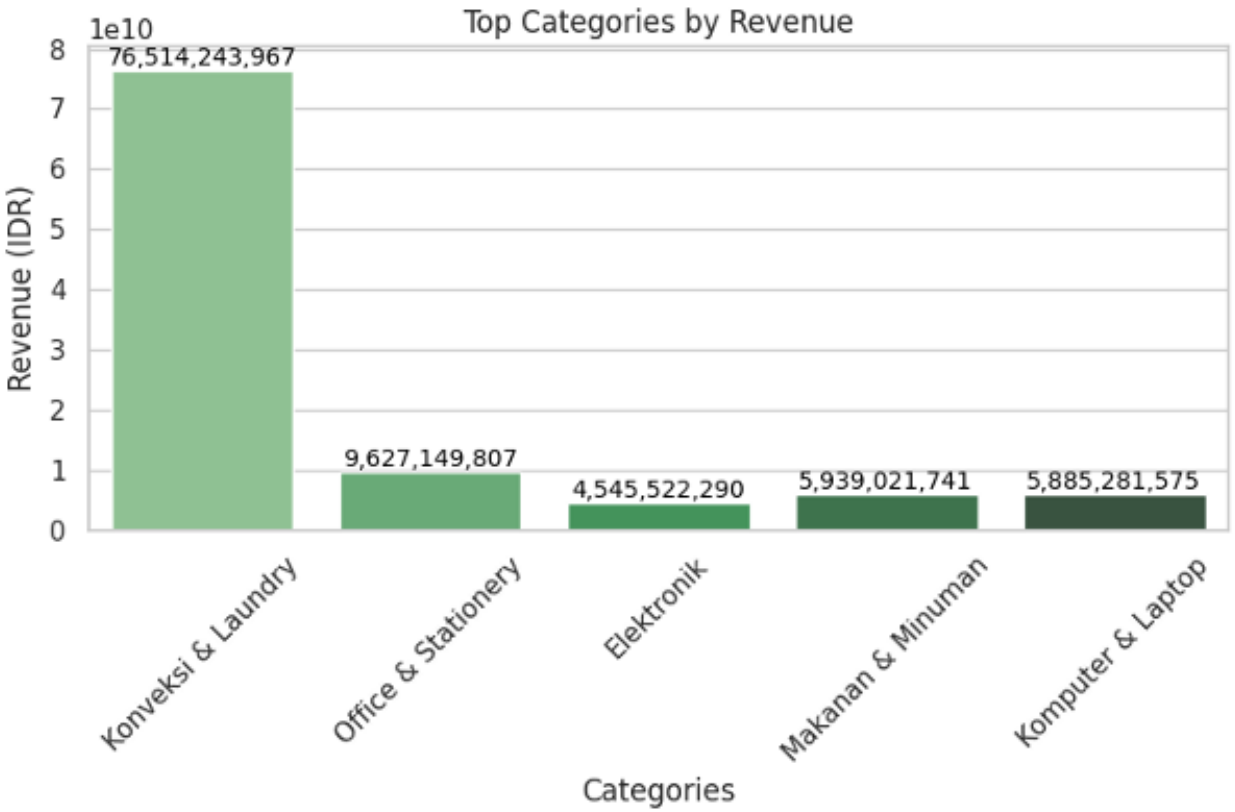




Barang dominates both GMV and revenue, accounting for a significant portion of business activity.

TOP 5 PRODUCT CATEGORIES





Konveksi & Laundry dominates both GMV and revenue, indicating a high-demand category that drives overall marketplace performance.

RECOMMENDATIONS

1. For Buyers

- **Focus on retention and engagement**: Develop loyalty programs or exclusive offers targeting frequent buyers in the **Barang** segment.
- **Expand regional campaigns**: Promote localized marketing campaigns in **emerging regions** like Kalimantan Timur and Sumatera Utara.
- **Target high-potential categories**: Incentivize buyers to explore underperforming categories, such as **Barang & Jasa**, through discounts or bundle offers.

2. For Sellers

- **Seller enablement programs**: Provide training, resources, and analytics tools to help mid-tier sellers improve their performance and bridge the revenue gap with top sellers.
- **Region-specific support**: Assist sellers in less competitive regions to expand their reach and optimize logistics to grow their presence.
- **Boost underperforming categories**: Encourage sellers to diversify their portfolios by exploring **Barang & Jasa** opportunities or offering complementary products.

RECOMMENDATIONS

3. General Marketplace Strategies

- Category-specific promotions: Prioritize growth in categories like Konveksi & Laundry, while developing strategies to boost profitability in Office & Stationery and Elektronik.
- Region-wise expansion plans: Leverage the success of Jakarta and replicate winning strategies in other high-potential regions.
- Optimize revenue streams: Focus on maximizing commission from high-demand categories and regions while balancing operational costs.

RESULTS

Datasets

- df_order: Orders, GMV, seller and buyer IDs, payment and shipping details.
- df_order_complete: Product-level details of orders.
- df_users: Geographic and user type information.
- df_products: Product-level attributes (not used beyond preprocessing).

Key Metrics Derived

- Buyers: GMV, Revenue, Orders, Recency.
- Sellers: GMV, Revenue, Orders Fulfilled, Unique Buyers.

BUYERS' CLUSTER CHARACTERISTICS

	segment	count_orders	avg_order_value	voucher_val	avg_recency
0	0	170.626508	1.884012e+07	16.893315	11.961119
1	1	5.087536	2.167863e+05	233970.761775	196.224746
2	2	13.223103	1.676430e+07	0.000000	160.905755

Cluster 0

High-Spending Regulars

Average Orders: 170.63
Average Order Value: 18,840,121
Voucher Usage: Minimal (16.89)
Recency: 11.96 days
(Very recent activity)

Cluster 1

Low-Spending Sporadic Buyers

Average Orders: 5.09
Average Order Value: 216,786
Voucher Usage: High (233,971)
Recency: 196.22 days
(Infrequent buyers)

Cluster 2

Moderate-Spending Dormant Buyers
Average Orders: 13.22
Average Order Value: 16,764,295
Voucher Usage: None
Recency: 160.91 days
(Occasionally active)

SELLERS' CLUSTER CHARACTERISTICS

	segment	count_orders	avg_order_value	unique_buyers	total_vouchers	recency
0	0	10.672523	7.398821e+07	1.657068	0.000000e+00	155.782007
1	1	248.663937	5.217293e+07	7.073121	5.849664e+01	10.117578
2	2	70.722277	1.134343e+06	18.500491	1.515993e+06	140.199215

Cluster 0

Niche Low-Volume Sellers

Average Orders: 10.67
Average Order Value: 73,988,206
Unique Buyers: 1.66
Voucher Usage: None
Recency: 155.78 days
(Occasionally active)

Cluster 1

High-Volume Established Sellers

Average Orders: 248.66
Average Order Value: 52,172,932
Unique Buyers: 7.07
Voucher Usage: Moderate (58.50)
Recency: 10.12 days
(Highly active)

Cluster 2

High-Potential Voucher-Dependent Sellers
Average Orders: 70.72
Average Order Value: 1,134,343
Unique Buyers: 18.50
Voucher Usage: Very High (1,515,993)
Recency: 140.20 days
(Moderate activity)

INSIGHTS

Buyer Insights:

- Cluster 0 buyers are highly engaged, making frequent purchases with minimal voucher dependence.
- Cluster 1 buyers rely heavily on vouchers but purchase infrequently and at low value.
- Cluster 2 buyers occasionally spend large amounts but are less engaged over time.

Seller Insights:

- Cluster 0 sellers operate in niche segments with low volume but high average order values.
- Cluster 1 sellers are the backbone of the marketplace, highly active with consistent performance.
- Cluster 2 sellers depend significantly on voucher campaigns and have high potential with diverse buyer bases.

RECOMMENDATIONS

For Buyers

- Cluster 0: Maintain engagement through loyalty programs.
- Cluster 1: Introduce reactivation campaigns and incentives to reduce churn.
- Cluster 2: Target with occasional high-value promotions to reengage.

For Sellers

- Cluster 0: Offer growth-focused support for niche segments.
- Cluster 1: Continue fostering high-volume sellers with tailored promotions.
- Cluster 2: Provide marketing assistance and optimize voucher campaigns for sustainable growth.

NEXT STEPS

Next Steps

1. Implement tailored marketing campaigns based on segmentation..
2. Explore alternative clustering methods that are better suited to the specific shapes and distributions of the datasets.
3. Revisit segmentation periodically for updates.

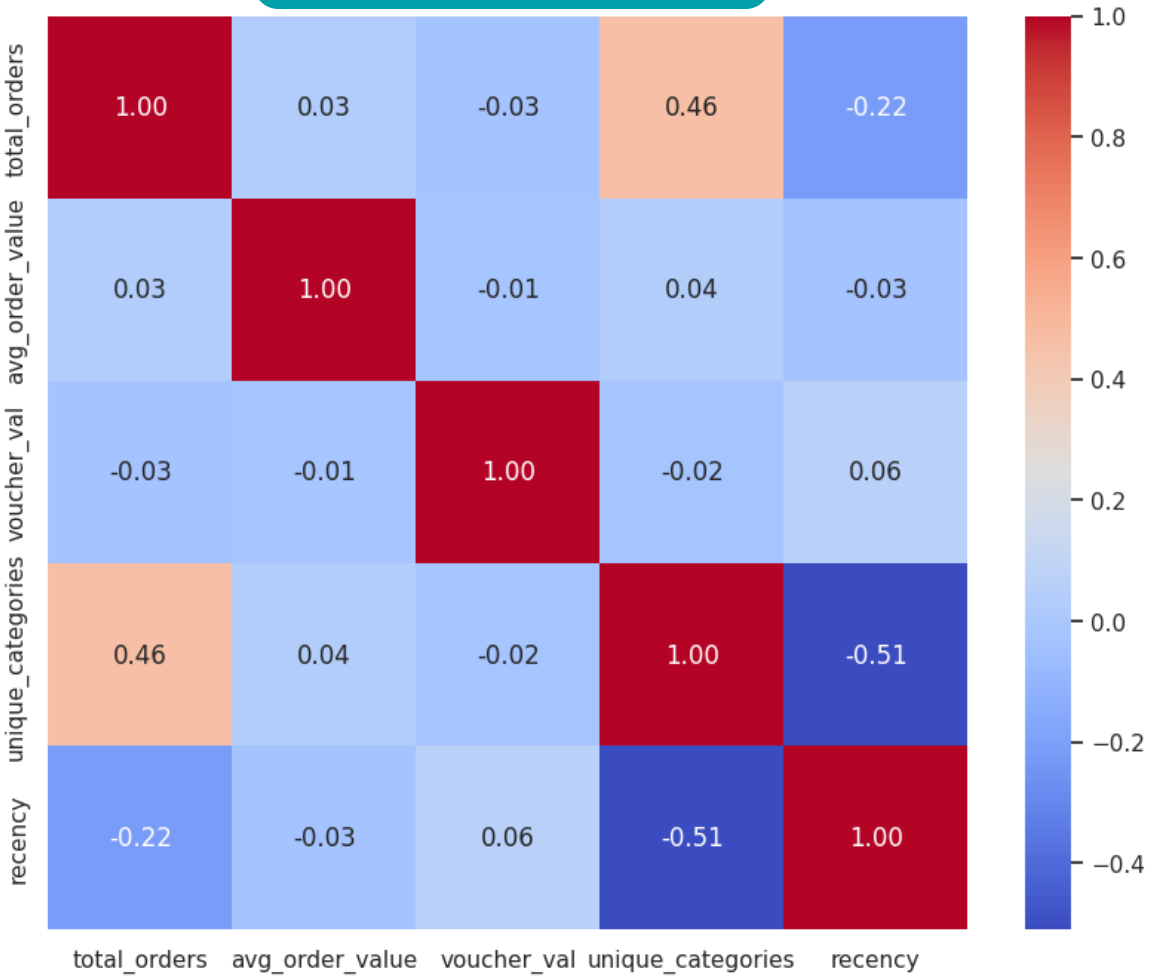
TERIMA KASIH!

APPENDIX

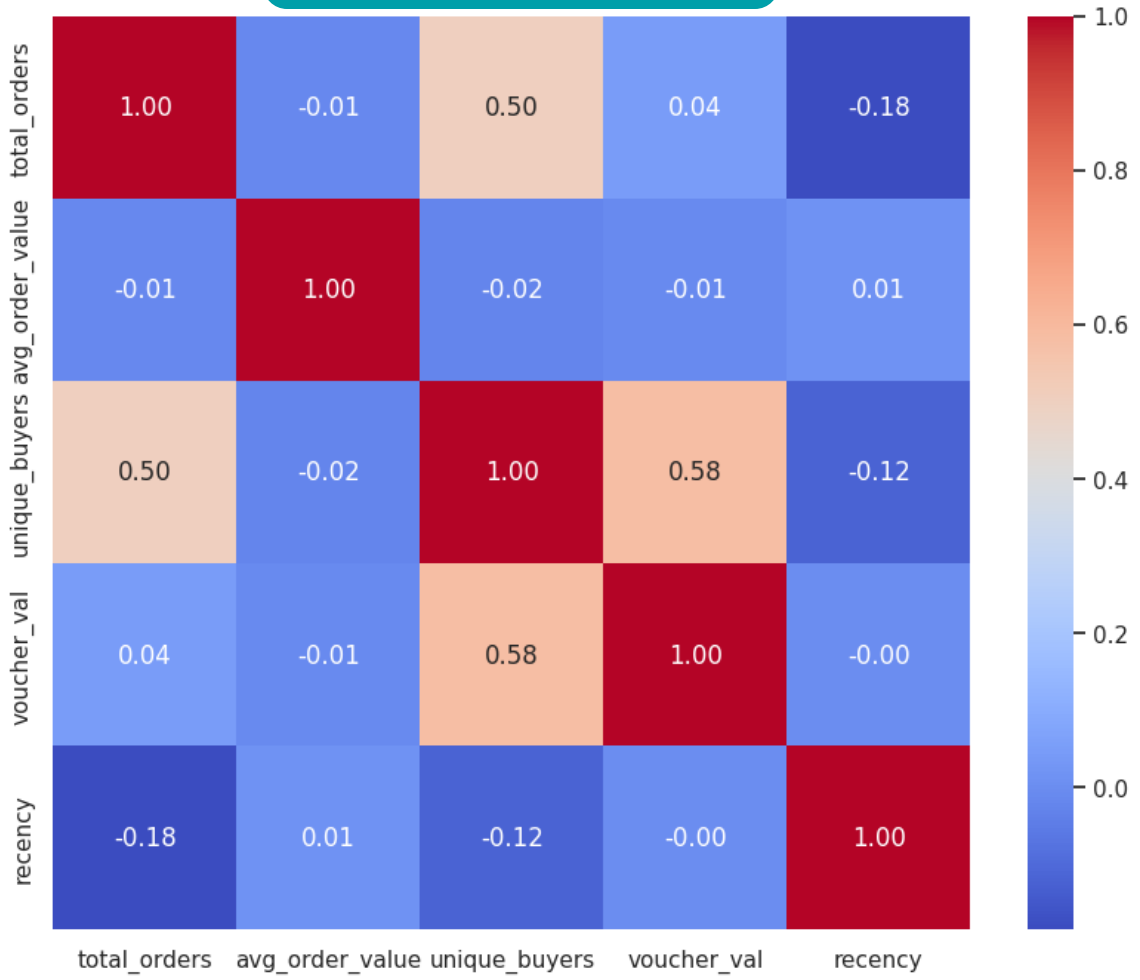
- Feature Correlation Matrix: heatmap or table of feature correlations.
- Elbow Analysis Diagram: diagram of elbow method.
- Silhouette Analysis: diagram of silhouette score and grid of silhouette visualizations.

FEATURE CORRELATION MATRIX

Buyers'

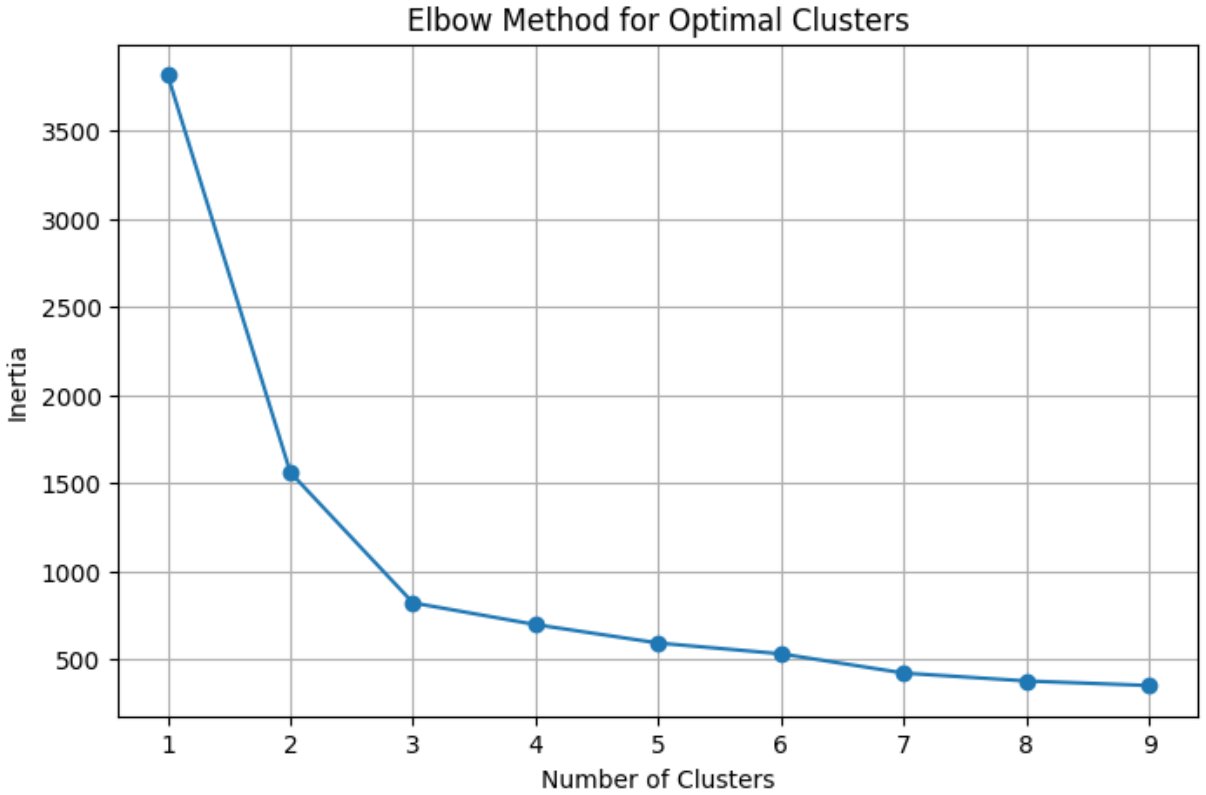


Sellers'

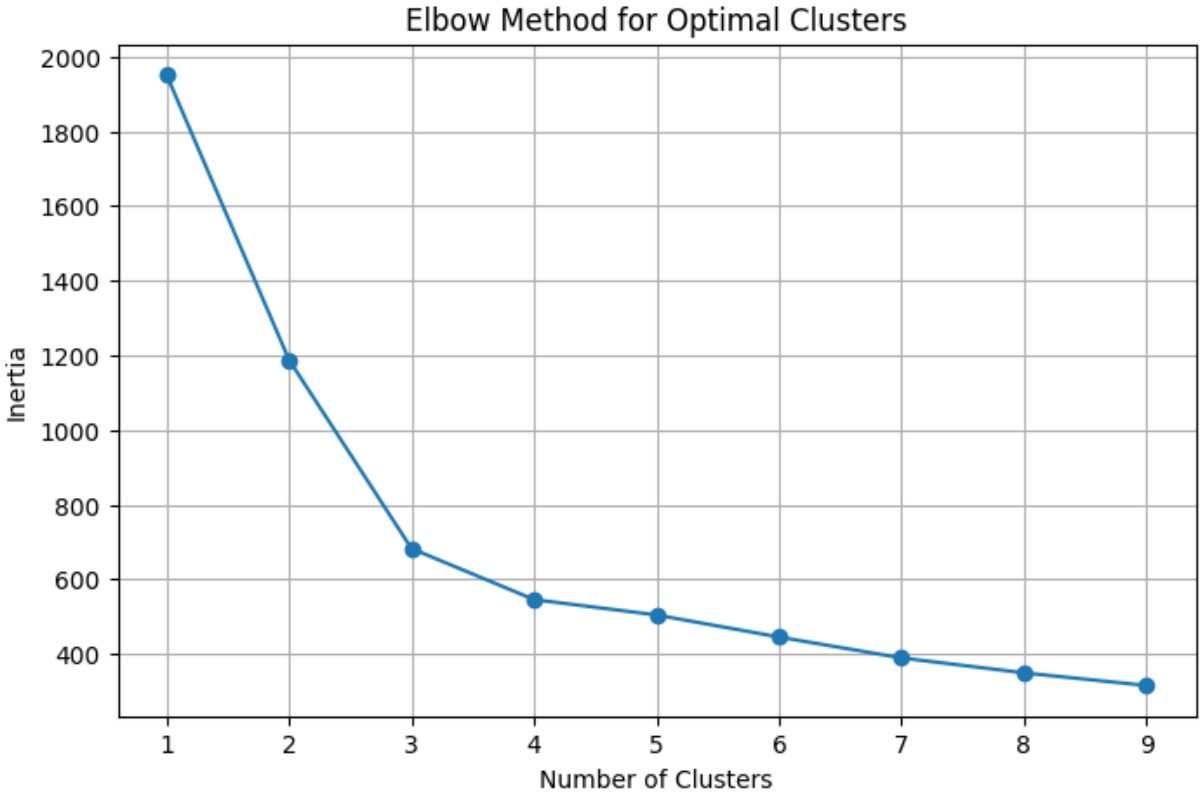


ELBOW METHOD

Buyers'

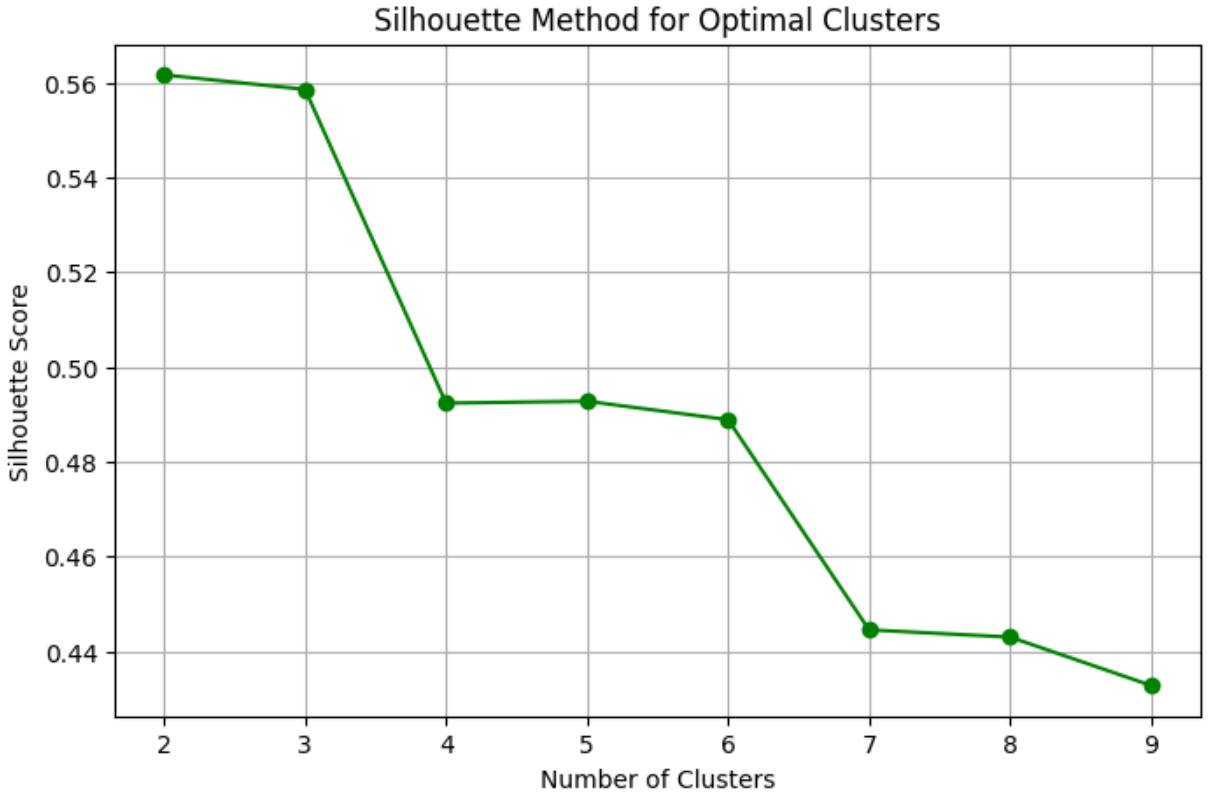


Sellers'

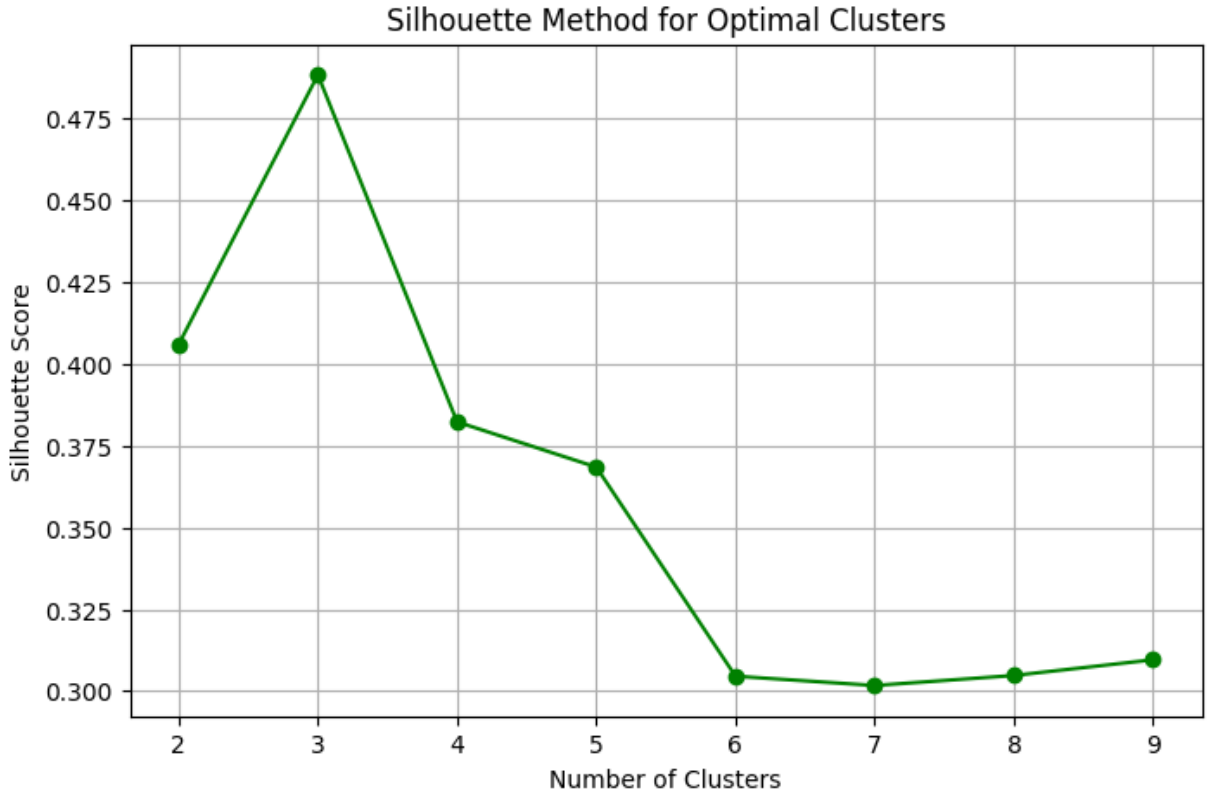


SILHOUETTE SCORE

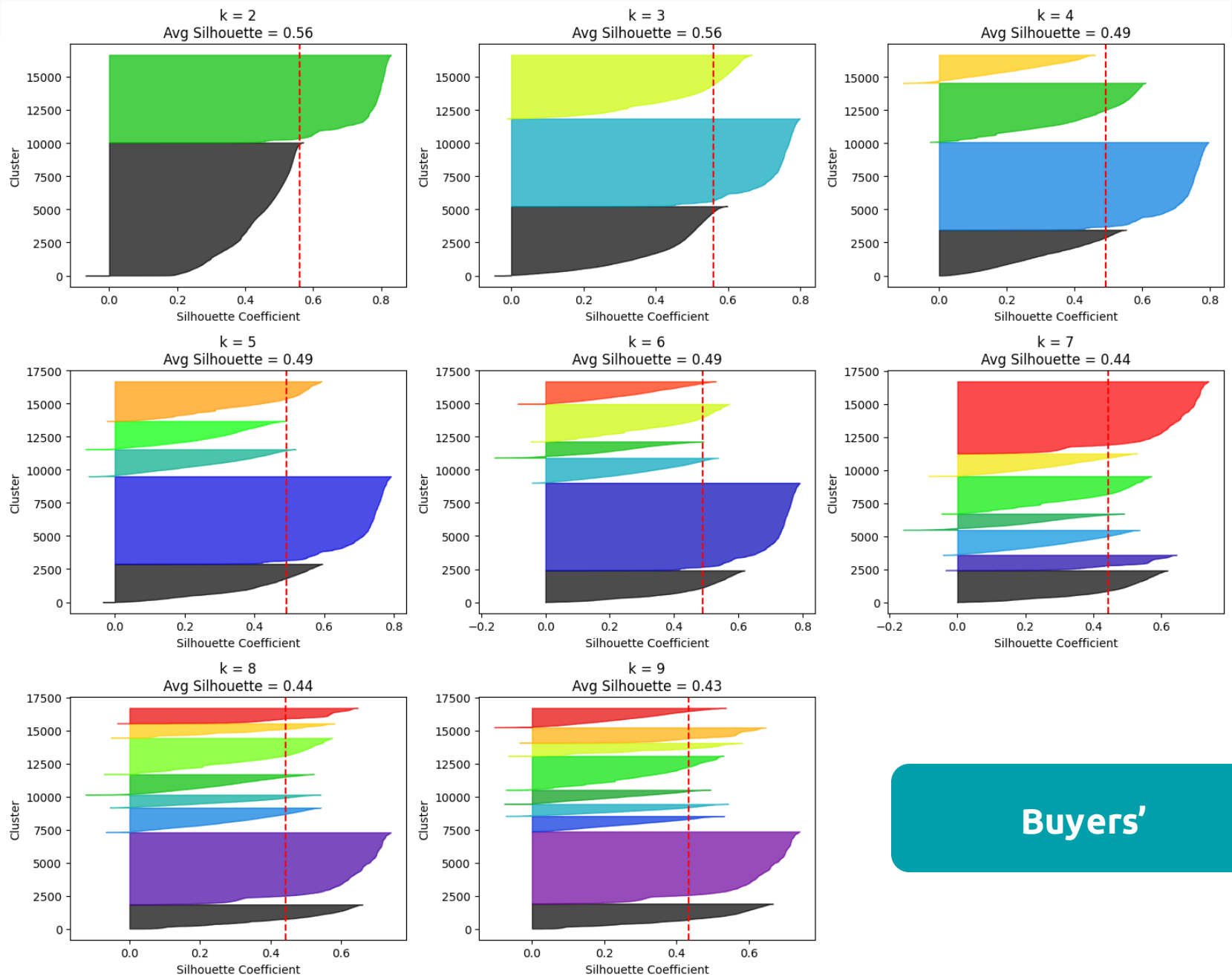
Buyers'



Sellers'

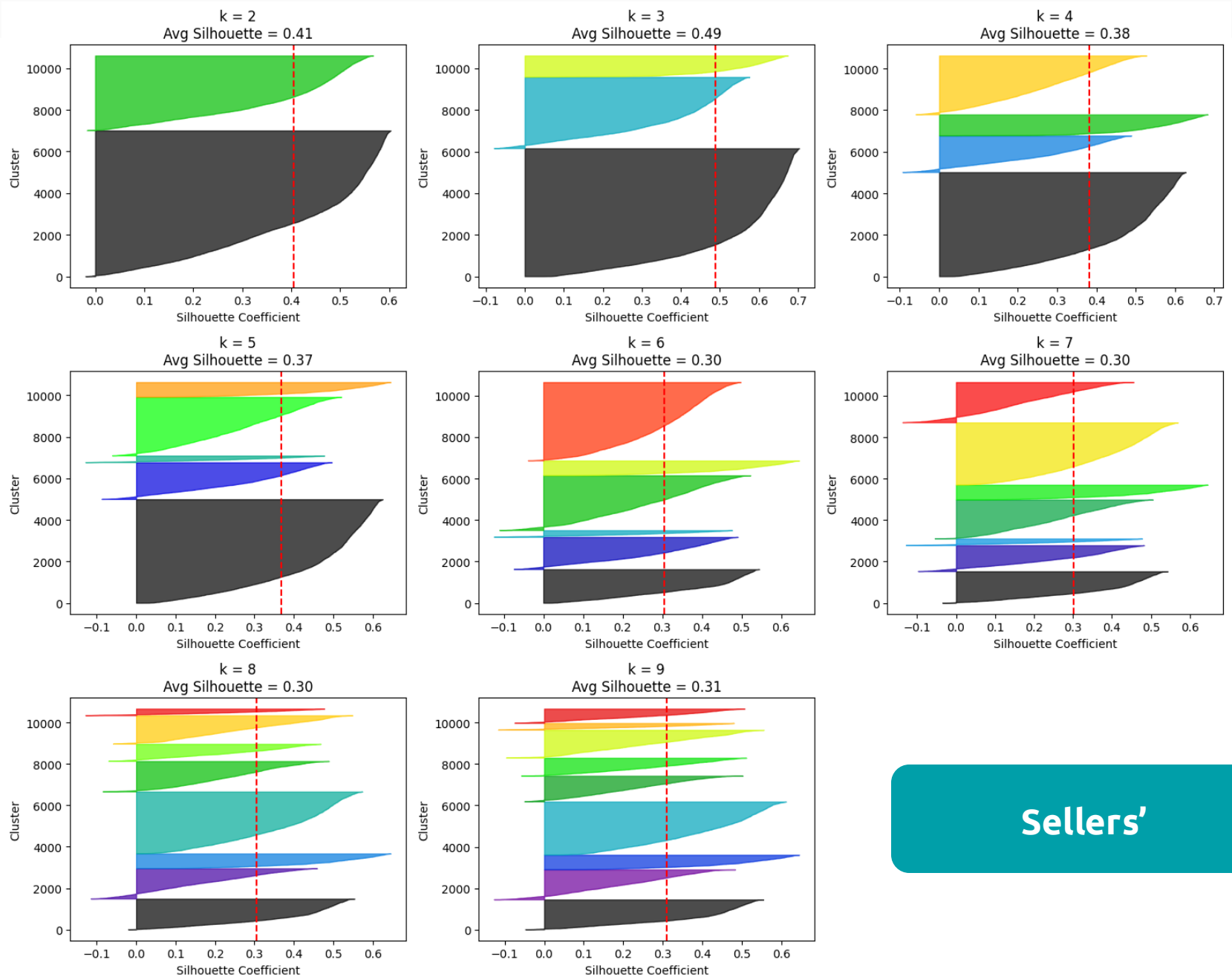


SILHOUETTE PLOT



Buyers'

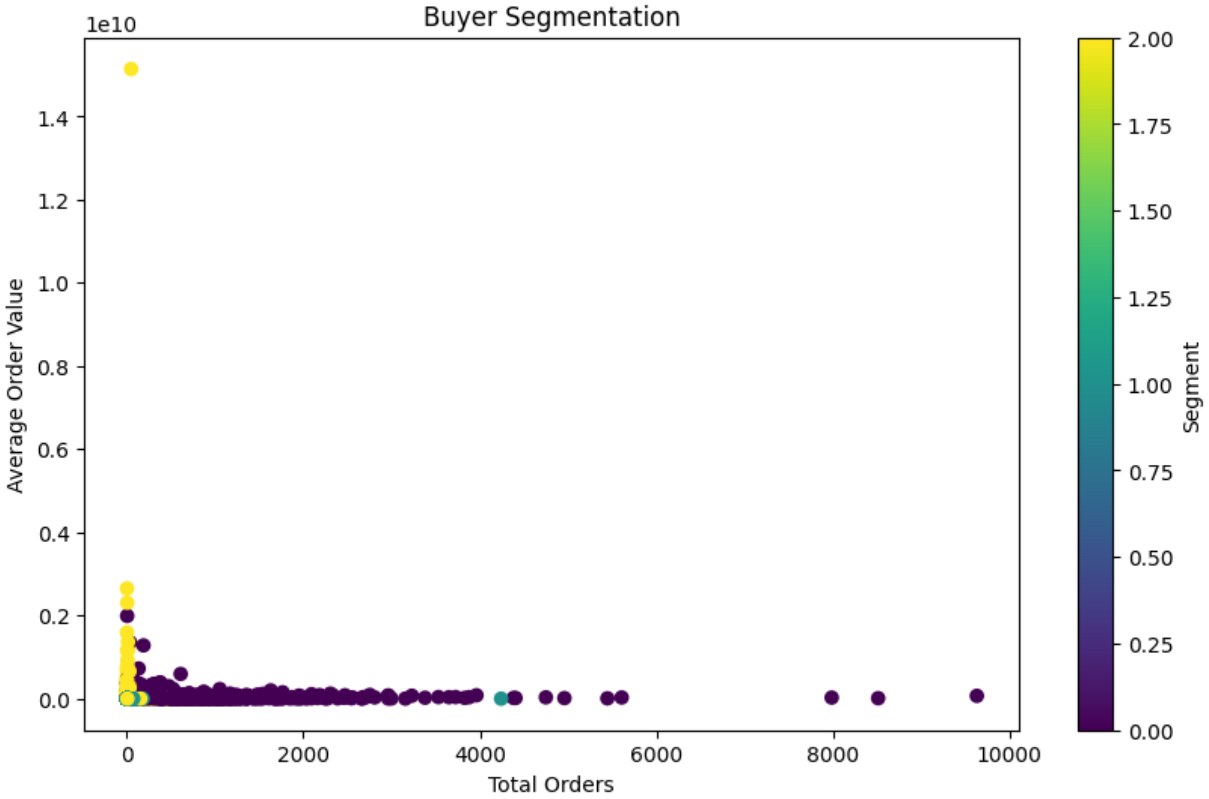
SILHOUETTE PLOT



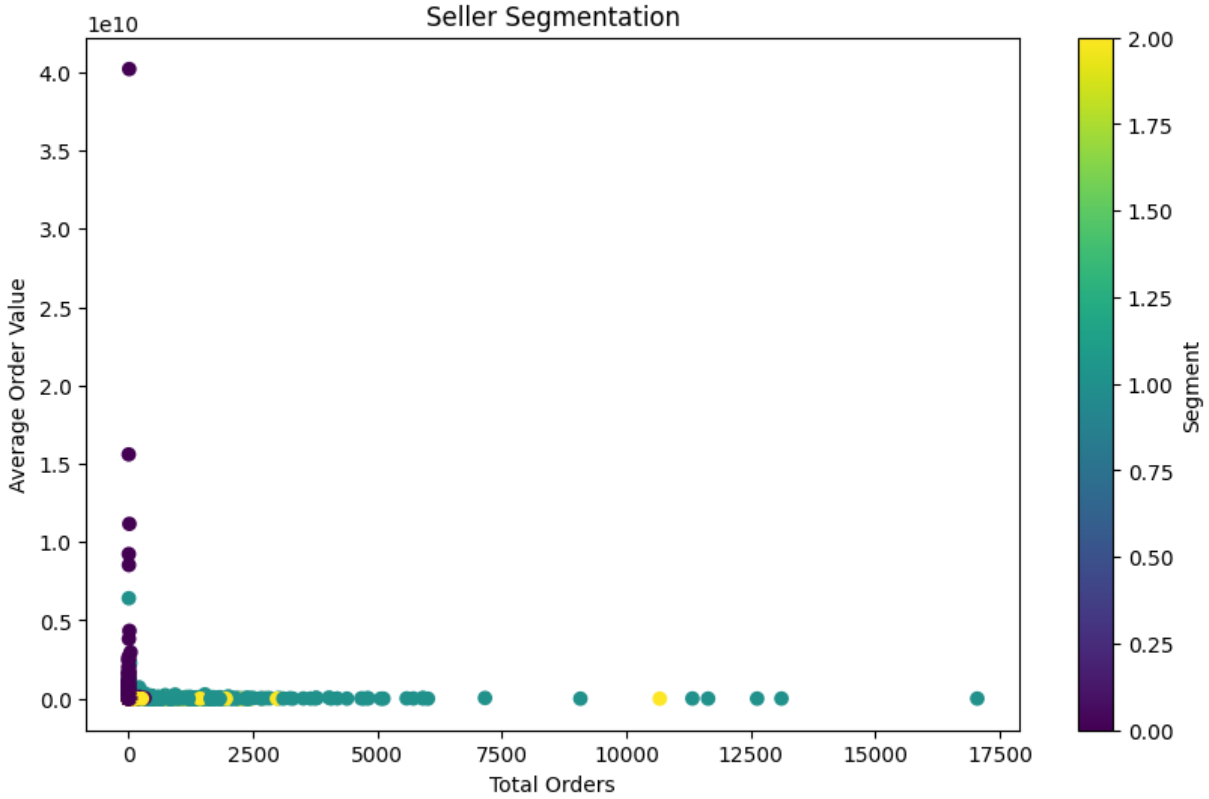
Sellers'

DATASETS SHAPES

Buyers'

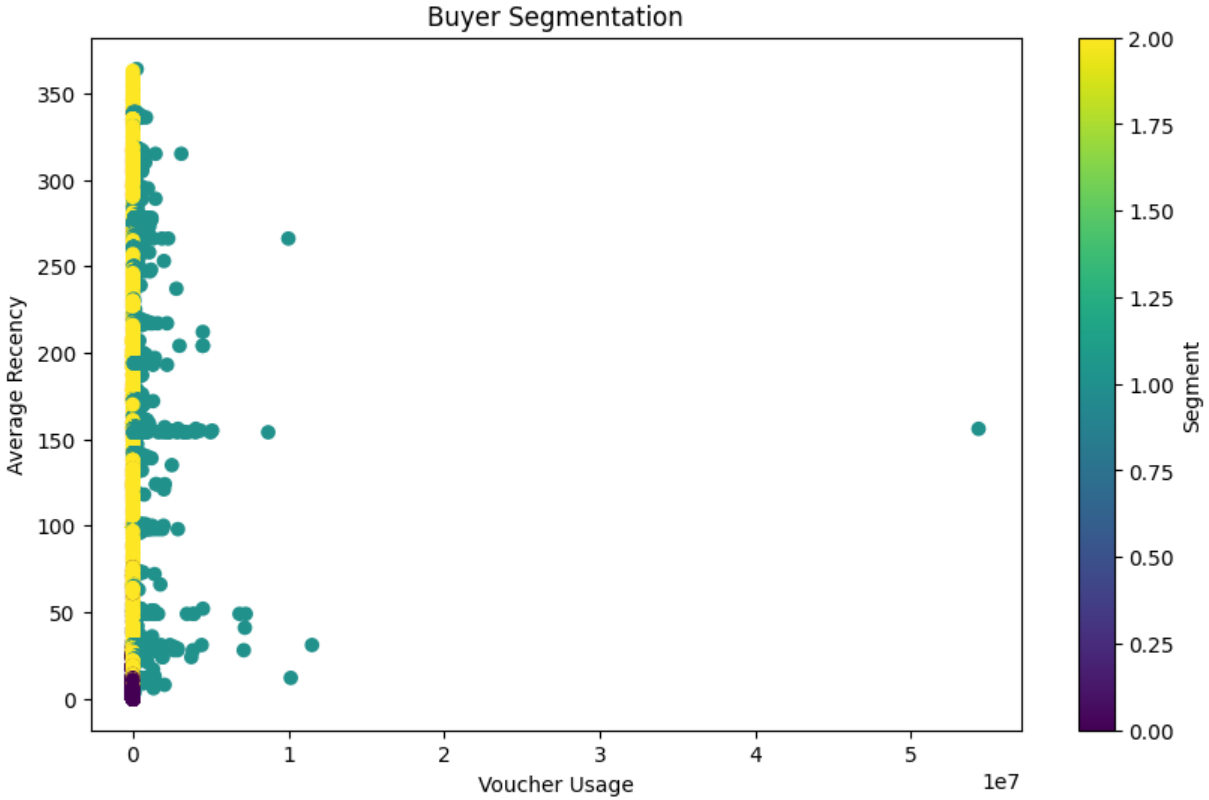


Sellers'

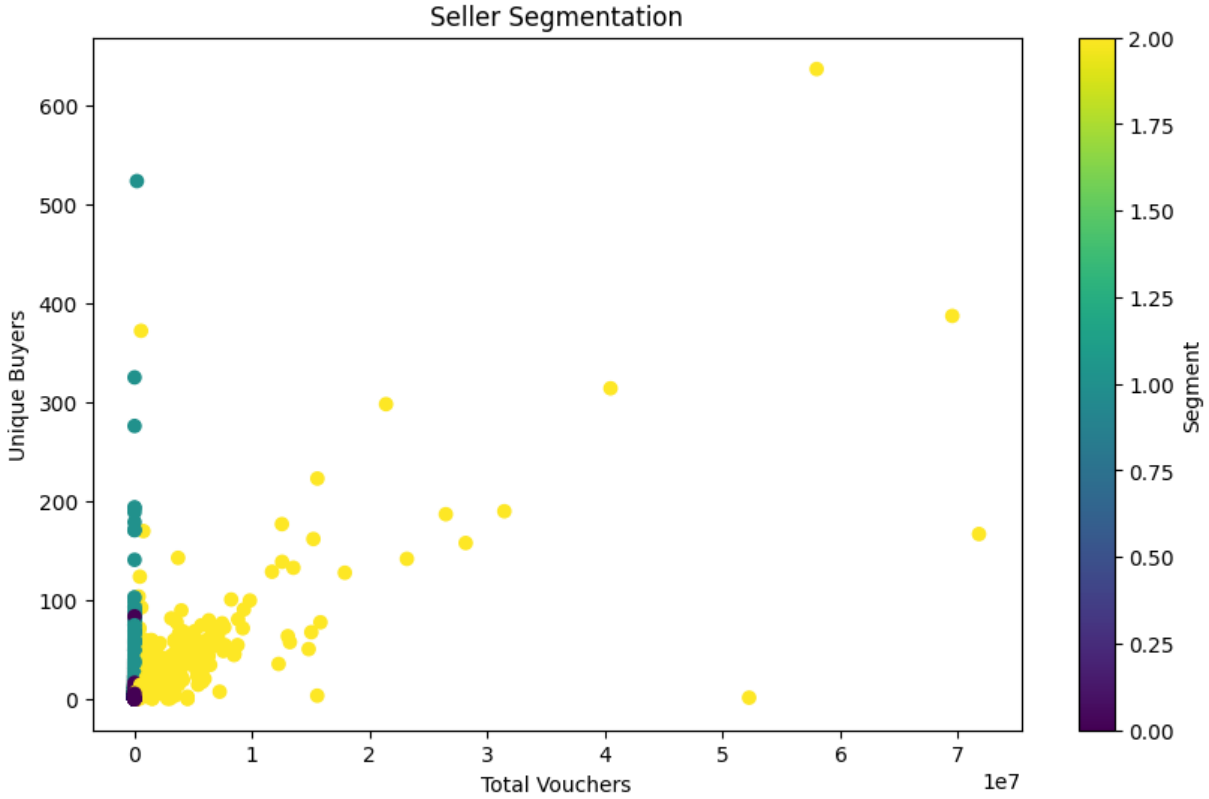


DATASETS SHAPES

Buyers'

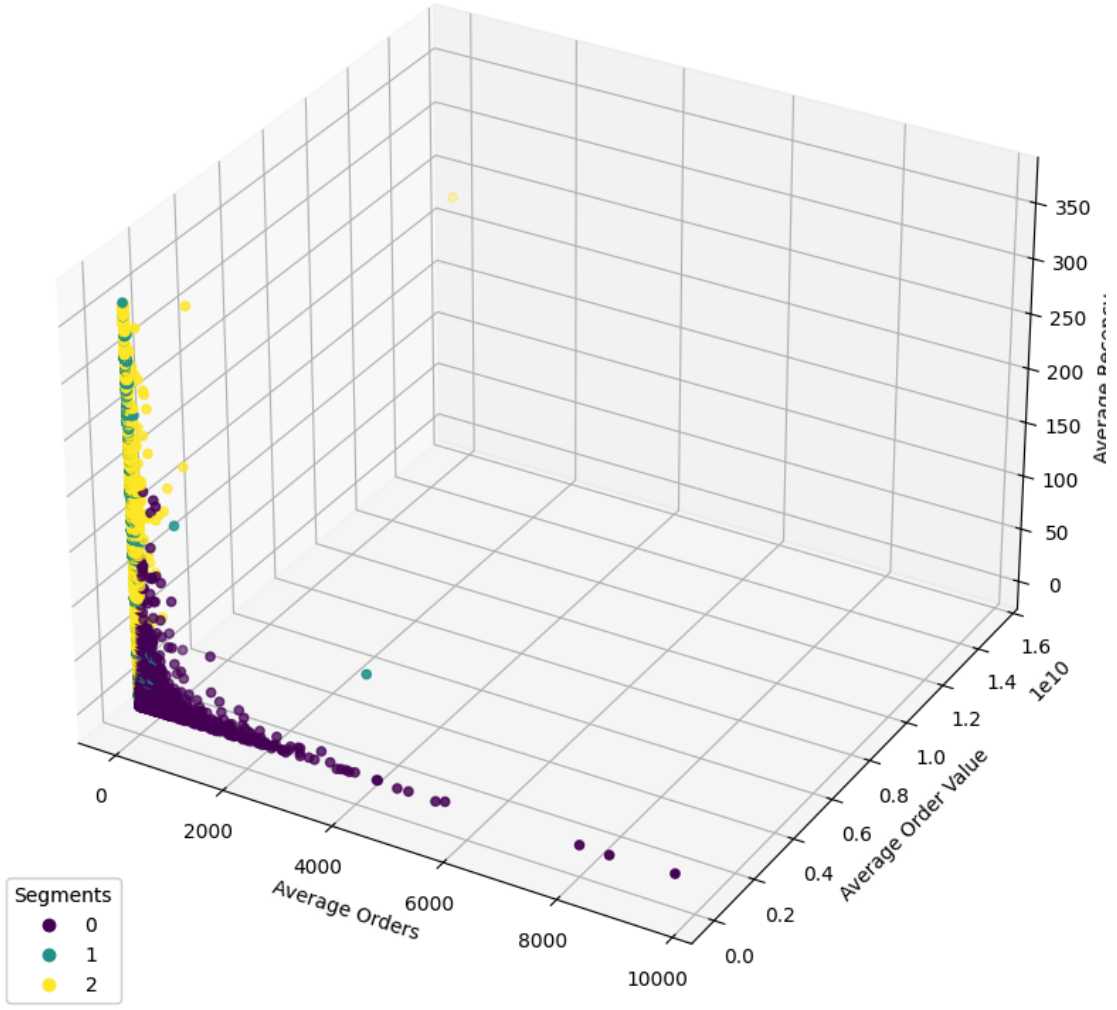
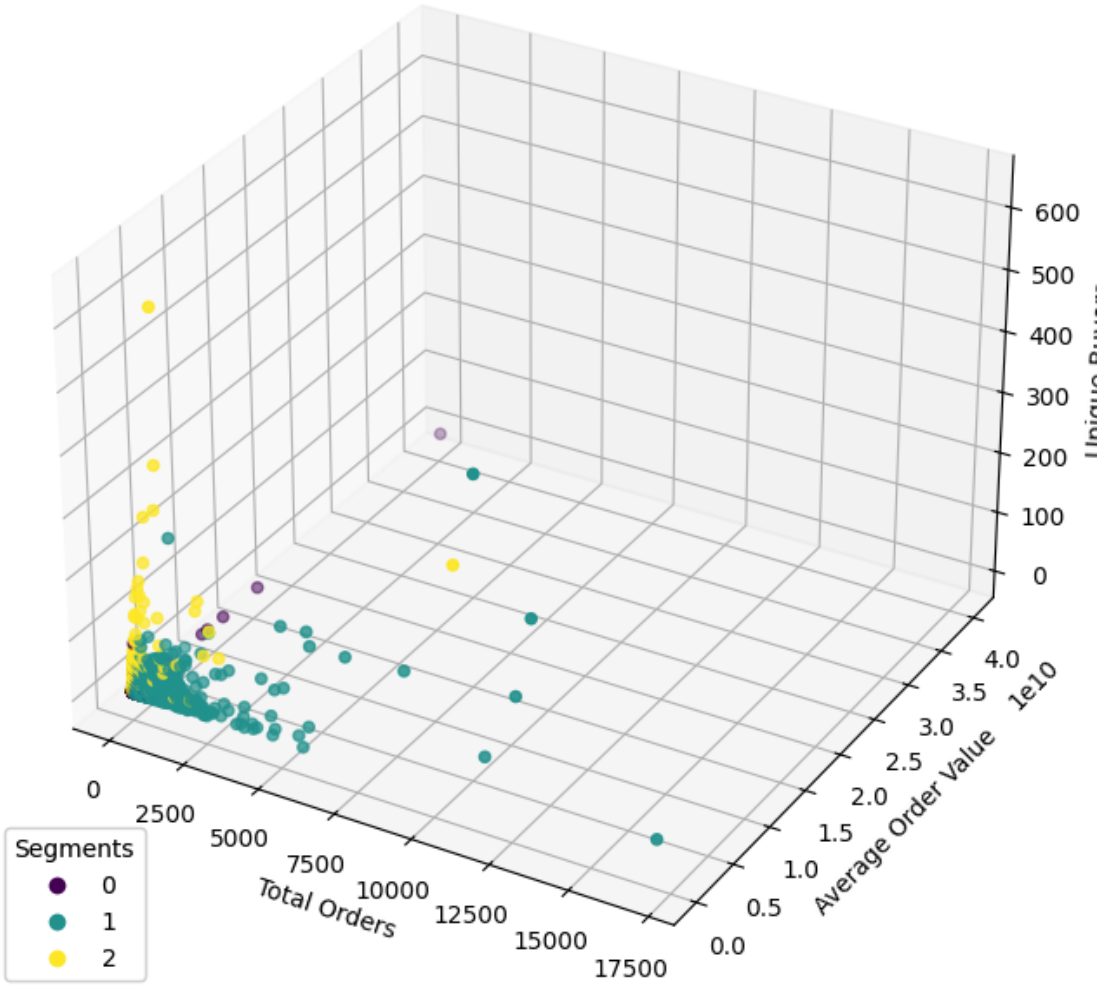


Sellers'



DATASETS SHAPES

Buyers'



Sellers'