E-commerce Analysis Using Product-centric RFM Clustering

DATA

Lazuardi Fauzan Primaputra

Table of Contents

01 Project Background

02 Initial Data Analysis

03 RFM Clustering

04 Actionable Insights

Project Background

01

In e-commerce operations, data is valuable asset that can be utilized to gain insights into behavior, refine strategies, and optimize overall business performance.

This self-learning project analyzes a dummy dataset with the goal of:

- **Identifying different product market segments based on the product behavioral characteristics.**

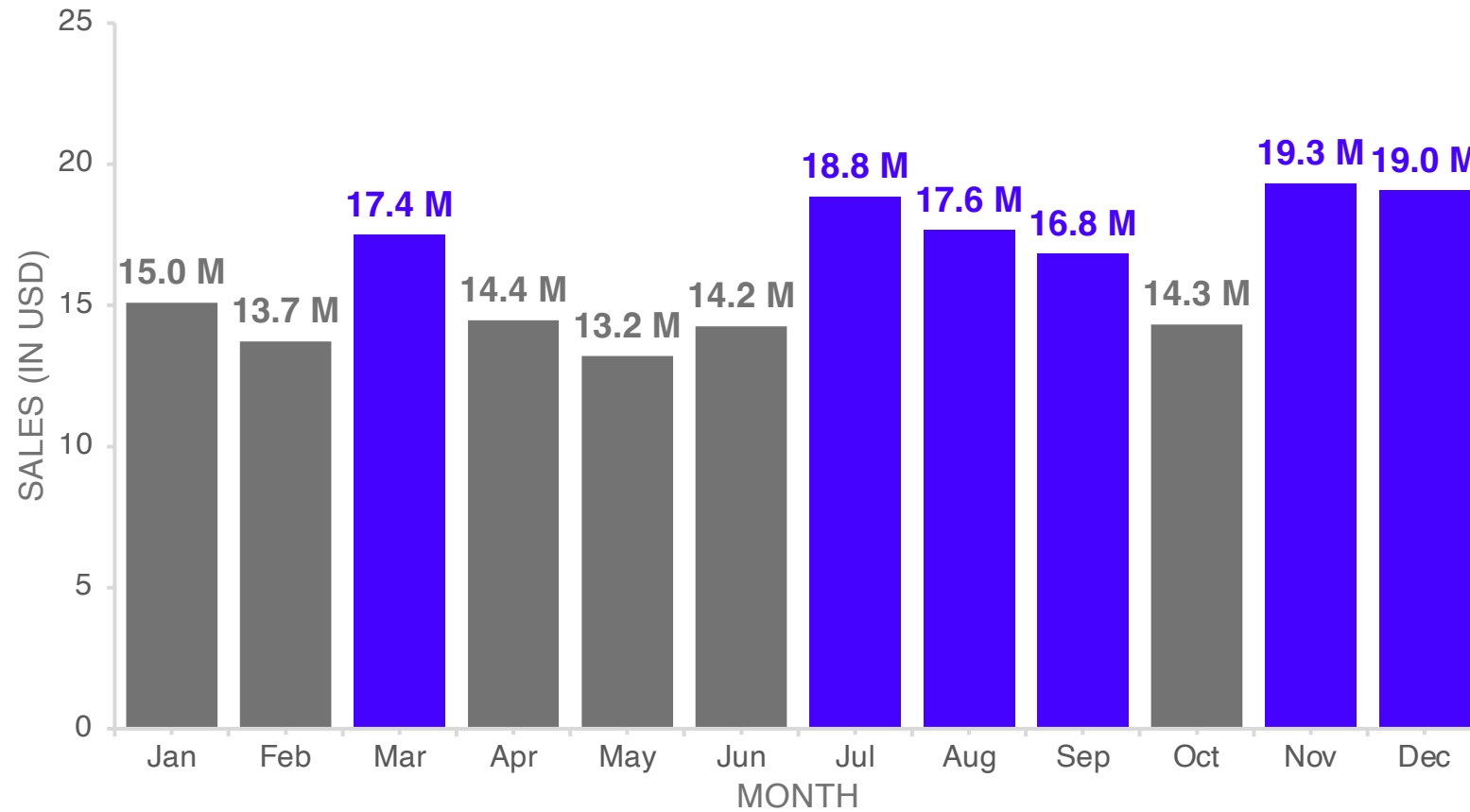- **Create targeted strategies for specific product segments.**

Data used for this project is obtained from Kaggle, with Apache 2.0 License.

Initial Data Analysis

Monthly Average Sales Performance

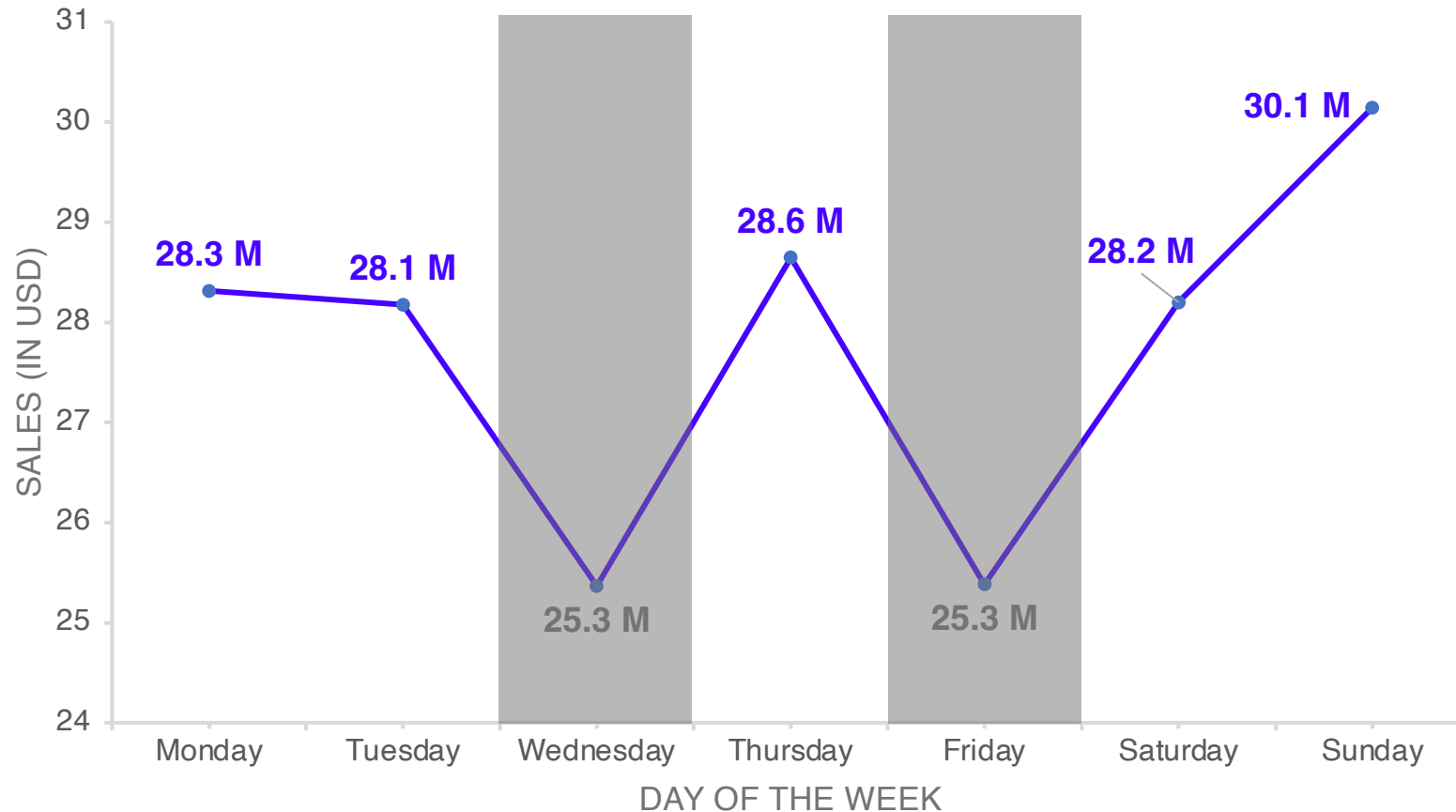Sales on the blue months is higher than the average (mean = $16.1 M)

SALES (IN USD)

Month	Sales
Jan	15.0 M
Feb	13.7 M
Mar	17.4 M
Apr	14.4 M
May	13.2 M
Jun	14.2 M
Jul	18.8 M
Aug	17.6 M
Sep	16.8 M
Oct	14.3 M
Nov	19.3 M
Dec	19.0 M

MONTH

Customers are the most active in **March**, **July**, **August**, **September** **November**, and **December.**

This is attributed to **Spring Sale**, **Summer Sale**, **Black Friday**, and **Christmas and Holiday Sale** occurring in those months.

Daily Average Sales Performance

Average daily sales on the blue days are higher than the average (mean = $ 27.7 M)

31

30.1 M

28.3 M   **28.1 M**   **28.6 M**   **28.2 M**

30

29

28

SALES (IN USD)

27

26

25   25.3 M   25.3 M

24

Monday   Tuesday   Wednesday   Thursday   Friday   Saturday   Sunday

DAY OF THE WEEK

Customers are mostly active throughout the week and **most active** in **Sunday**. **High-quantity transactions** also occurs at this time.

However, there is a **dip** in sales on **Wednesday** and **Friday**. This dip is attributed to people **having less leisure time** in that days.
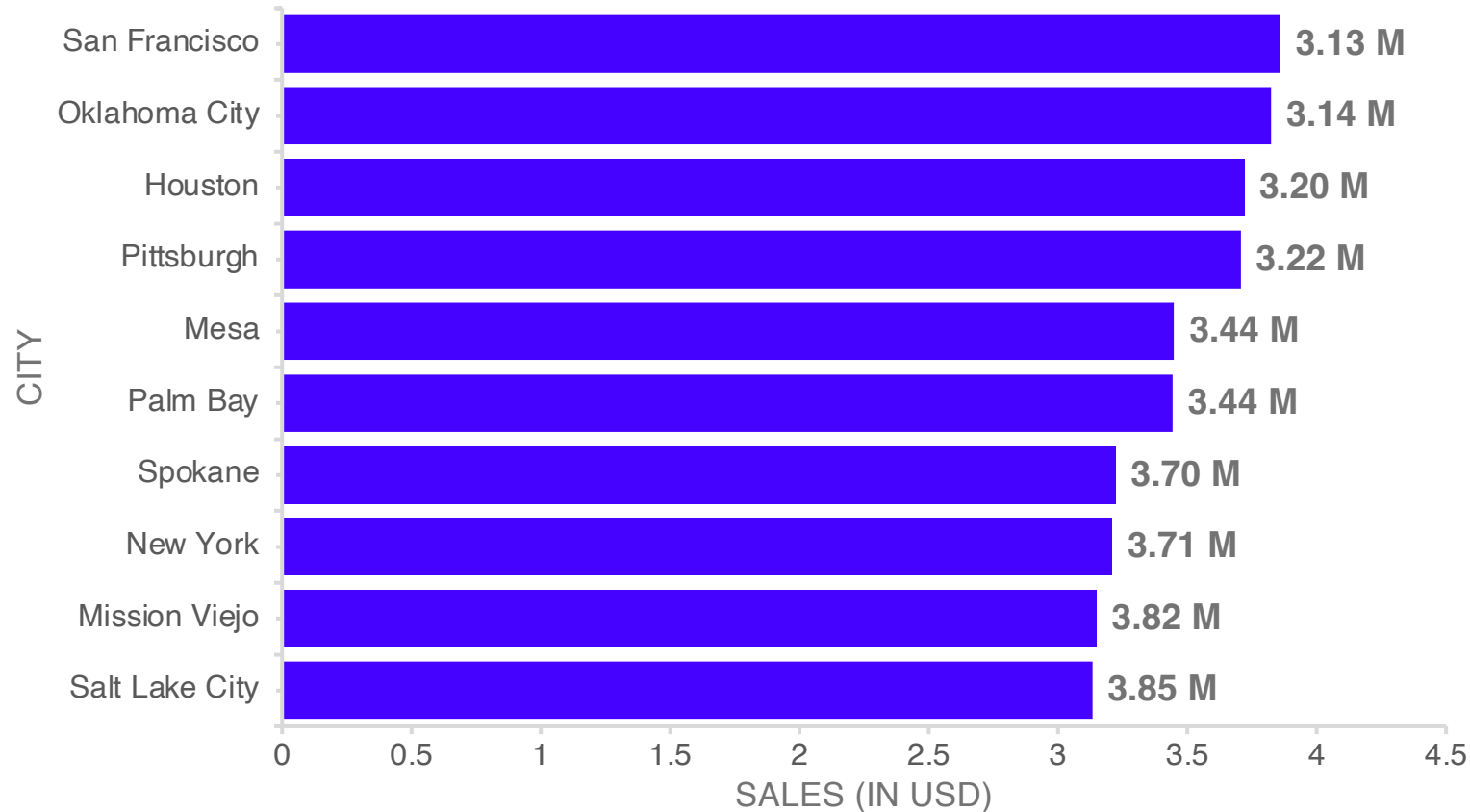
Sales by Geographic Distributions

Top States with the highest average sales (mean = $4.5 M)

Sales

28,536,724.02

5,631,317.73
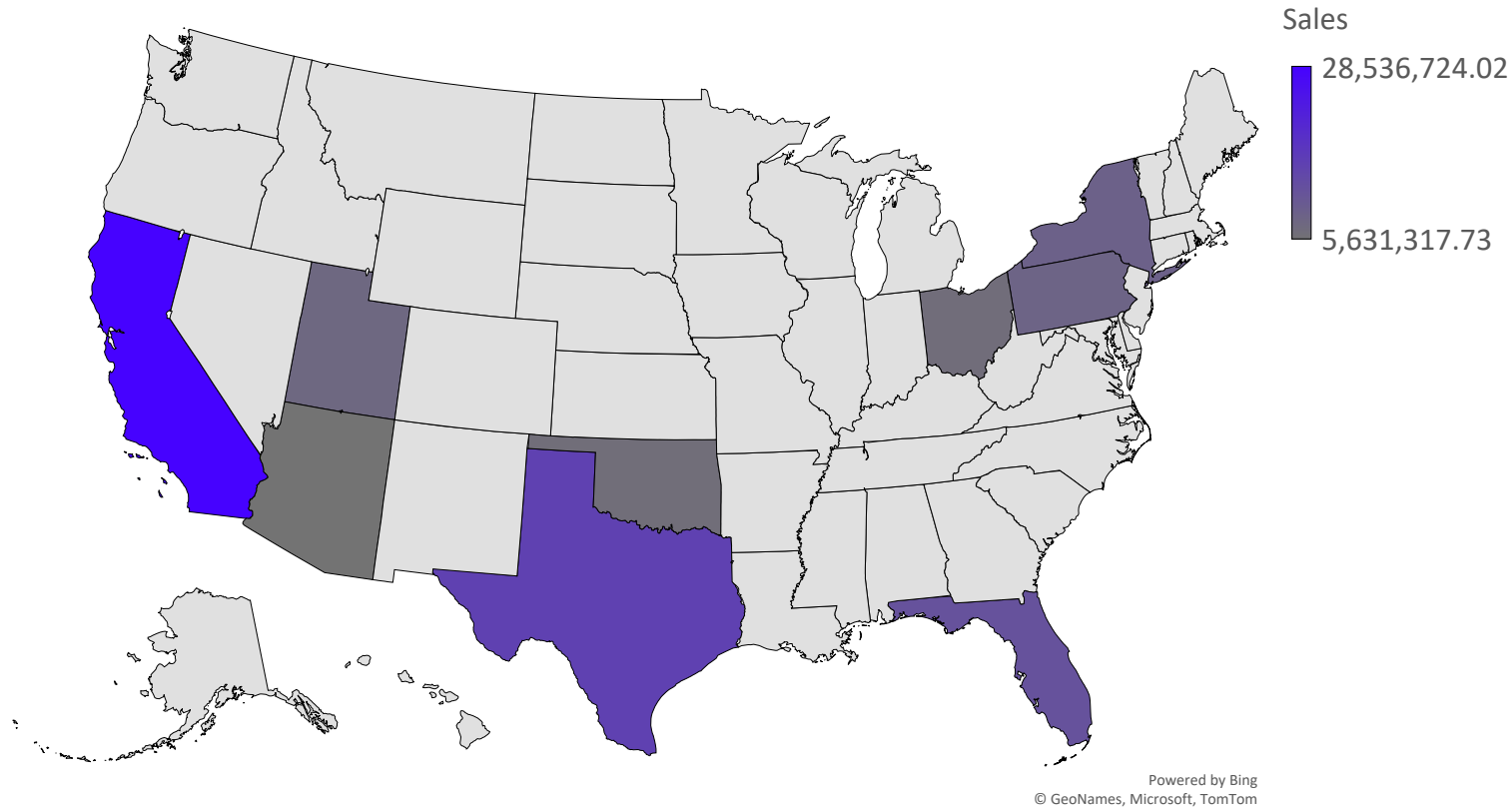
Highest customer base is located in **California, Texas, Florida, New York, Pennsylvania, Utah, Ohio, Oklahoma, and Arizona.**

Powered by Bing
© GeoNames, Microsoft, TomTom

Sales by State Distribution

Top 10 states with the highest sales.

City	Sales (in USD)
San Francisco	3.13 M
Oklahoma City	3.14 M
Houston	3.20 M
Pittsburgh	3.22 M
Mesa	3.44 M
Palm Bay	3.44 M
Spokane	3.70 M
New York	3.71 M
Mission Viejo	3.82 M
Salt Lake City	3.85 M

These cities share a common **demographic characteristics** such as income level and lifestyle.

Other **economic factors**, such as disposable income, population density, and economic growth **influence the higher sales** in these cities.
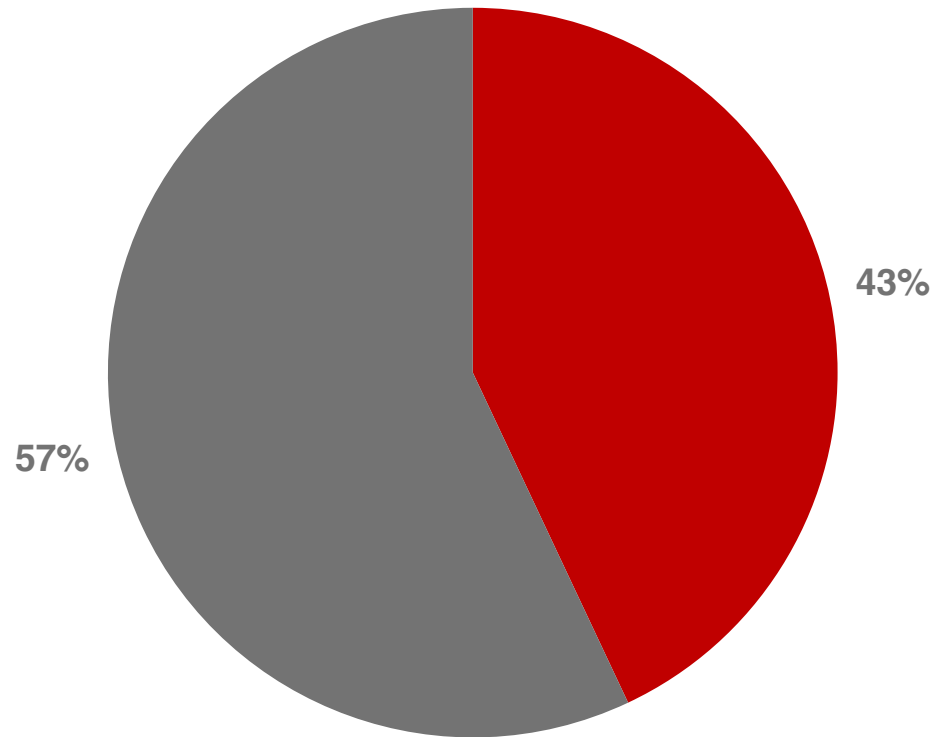
Top 10 Products

Top products consists a mix of household goods, clothing, hobbies, and gadget.

CITY	SALES (IN USD)
Biography	4.92 M
Mountain Bike	4.19 M
Wool Socks	3.57 M
Cleanser	3.54 M
Water Bottle	3.50 M
Denim Jacket	3.45 M
Nail Clippers	3.33 M
Earbuds	3.21 M
Eau de Toilette	3.10 M
Spark Plugs	3.09 M
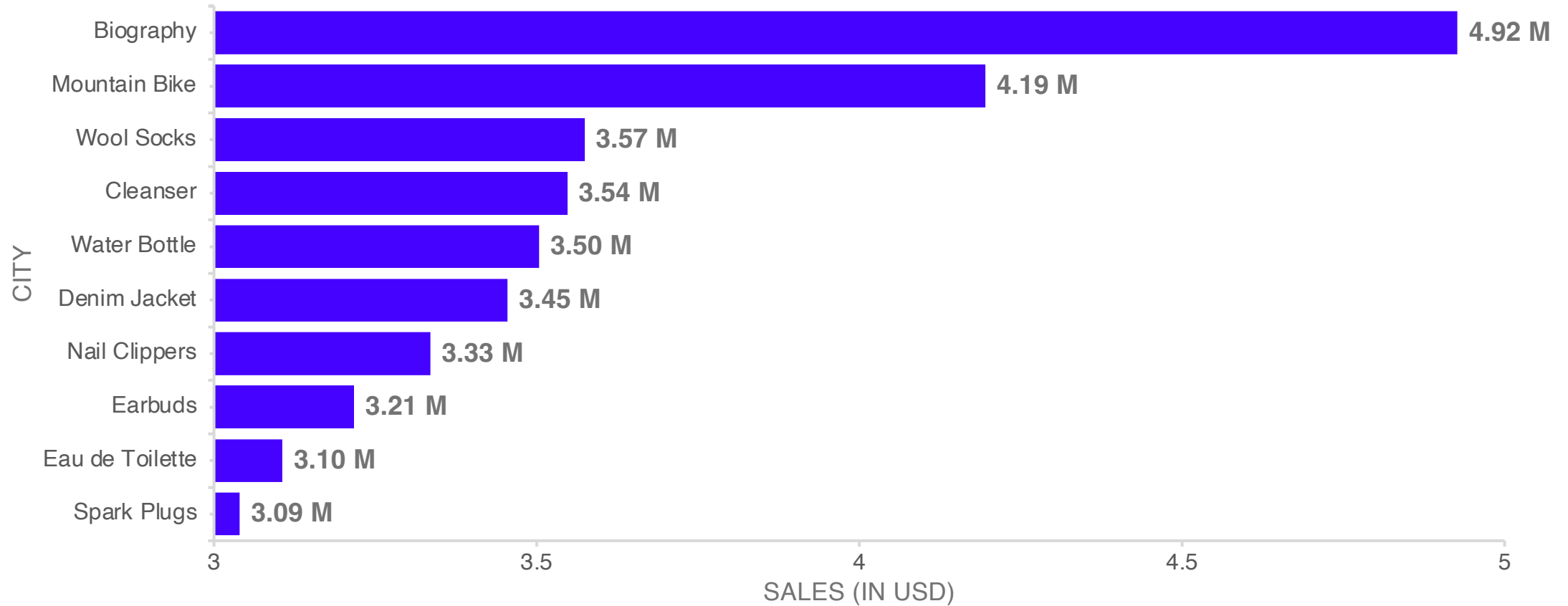
Product Performance

43% of products are performing below average (mean = $1.94 M)

43%

57%

Initial analysis reveals that **nearly half** of the products are performing **below average** in sales, highlighting a critical **area for improvements.**

RFM Clustering

03

To improve sales and more effective product management, Product-centric RFM Clustering helps in understanding product performance, optimizing supply chain and inventory.

RFM Clustering? What?

- **Recency**   How recently a product was purchased?

- **Frequency**   How often a product is purchased?

- **Monetary**   How much revenue generated by a product?

RFM Clustering? Why?

- **Product Segmentation**   Classify product based on performance

- **Sales Insights**   Identify high and low performing products

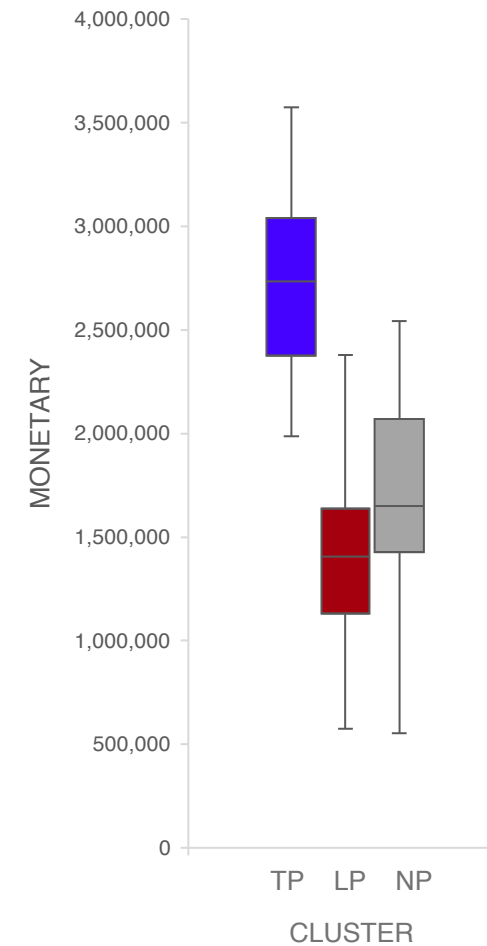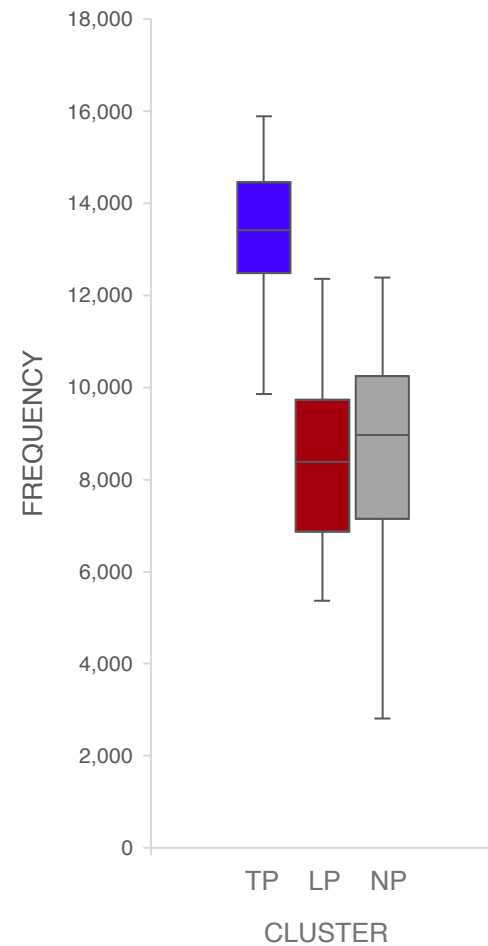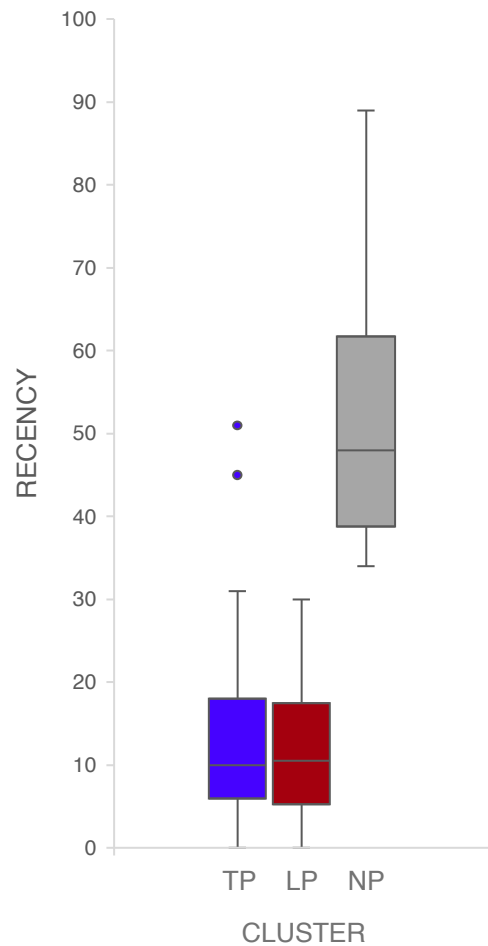- **Inventory Management**   Optimize stock levels and supply chain
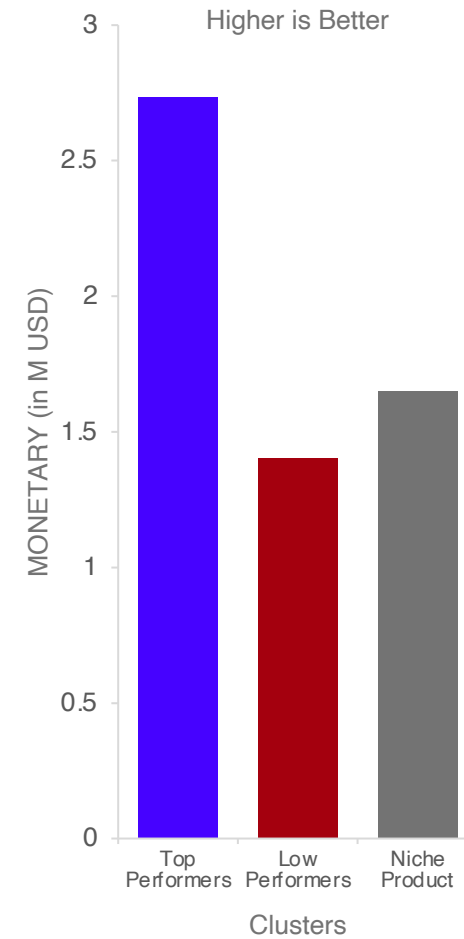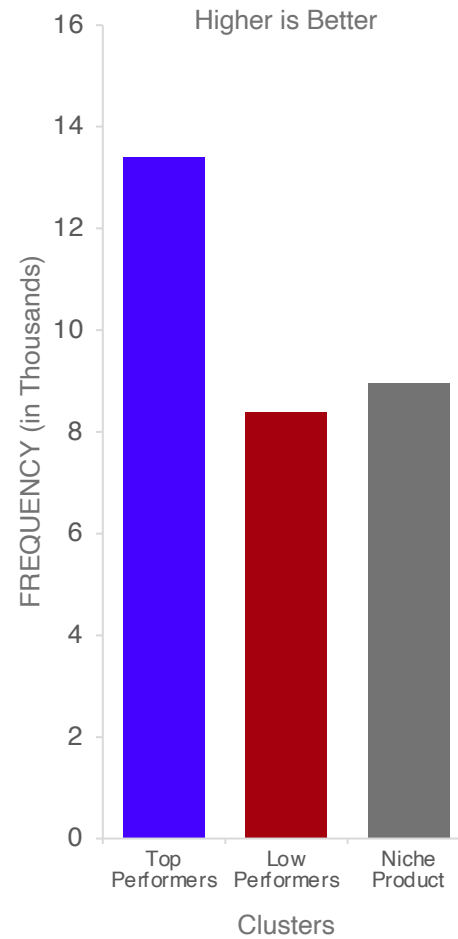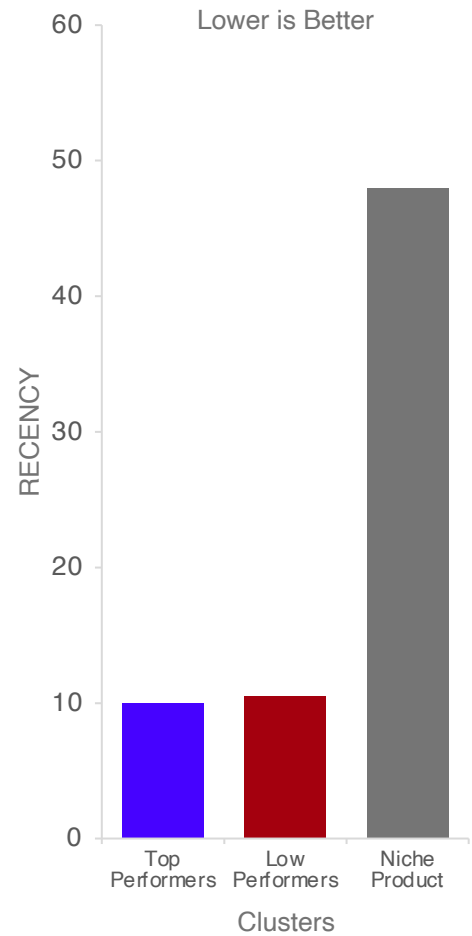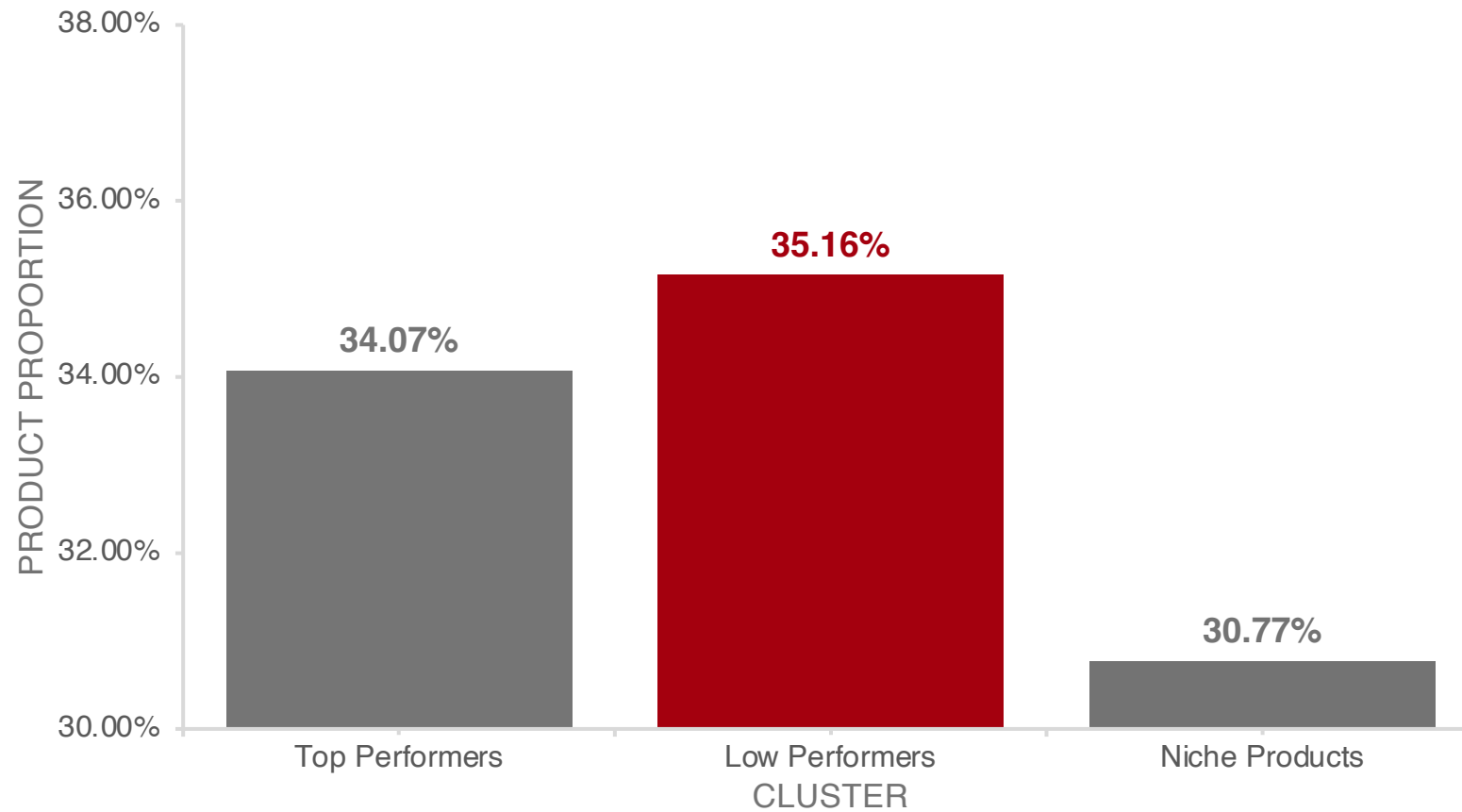
Product Lifetime Value by Cluster

To identify which clusters that brings the most long-term value
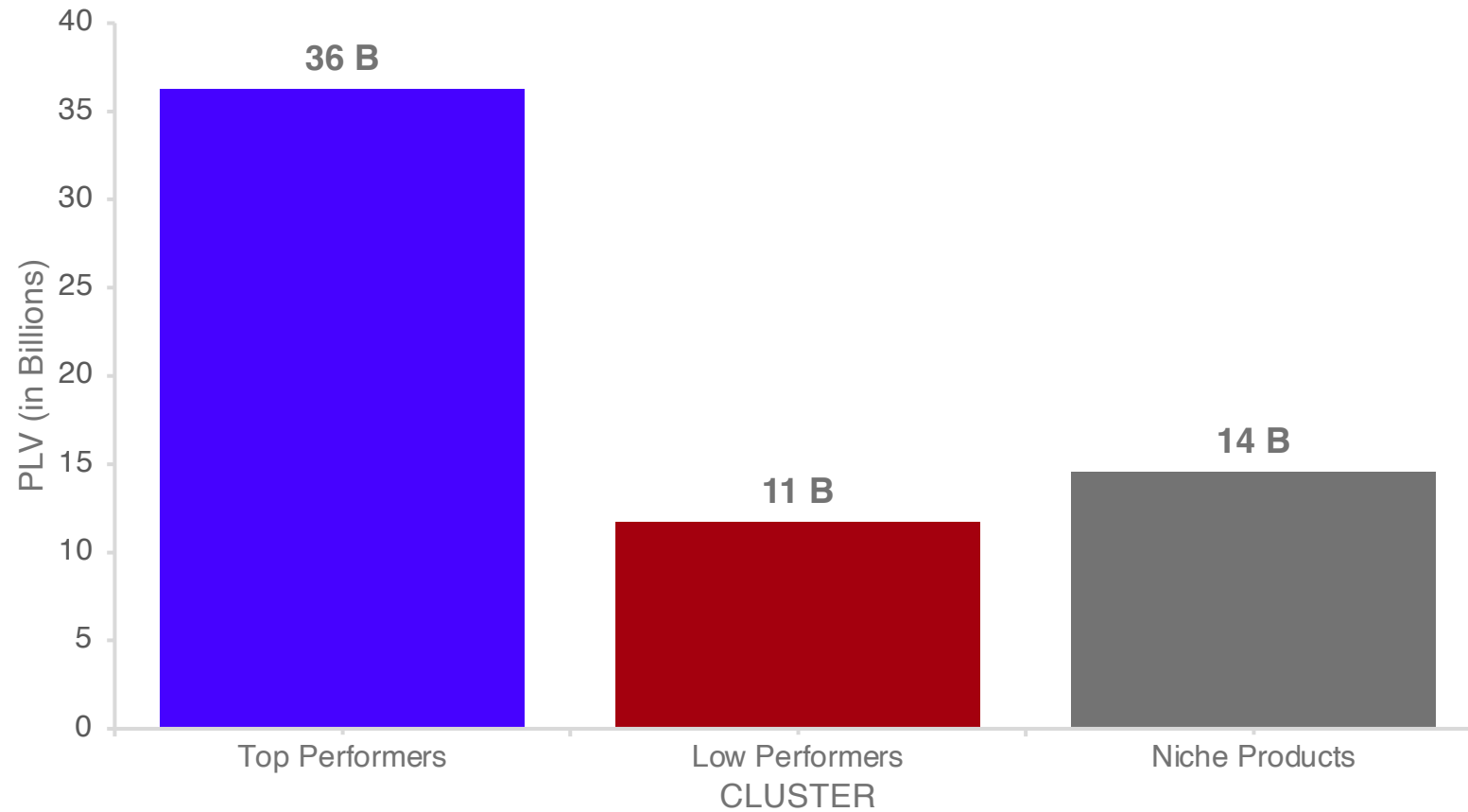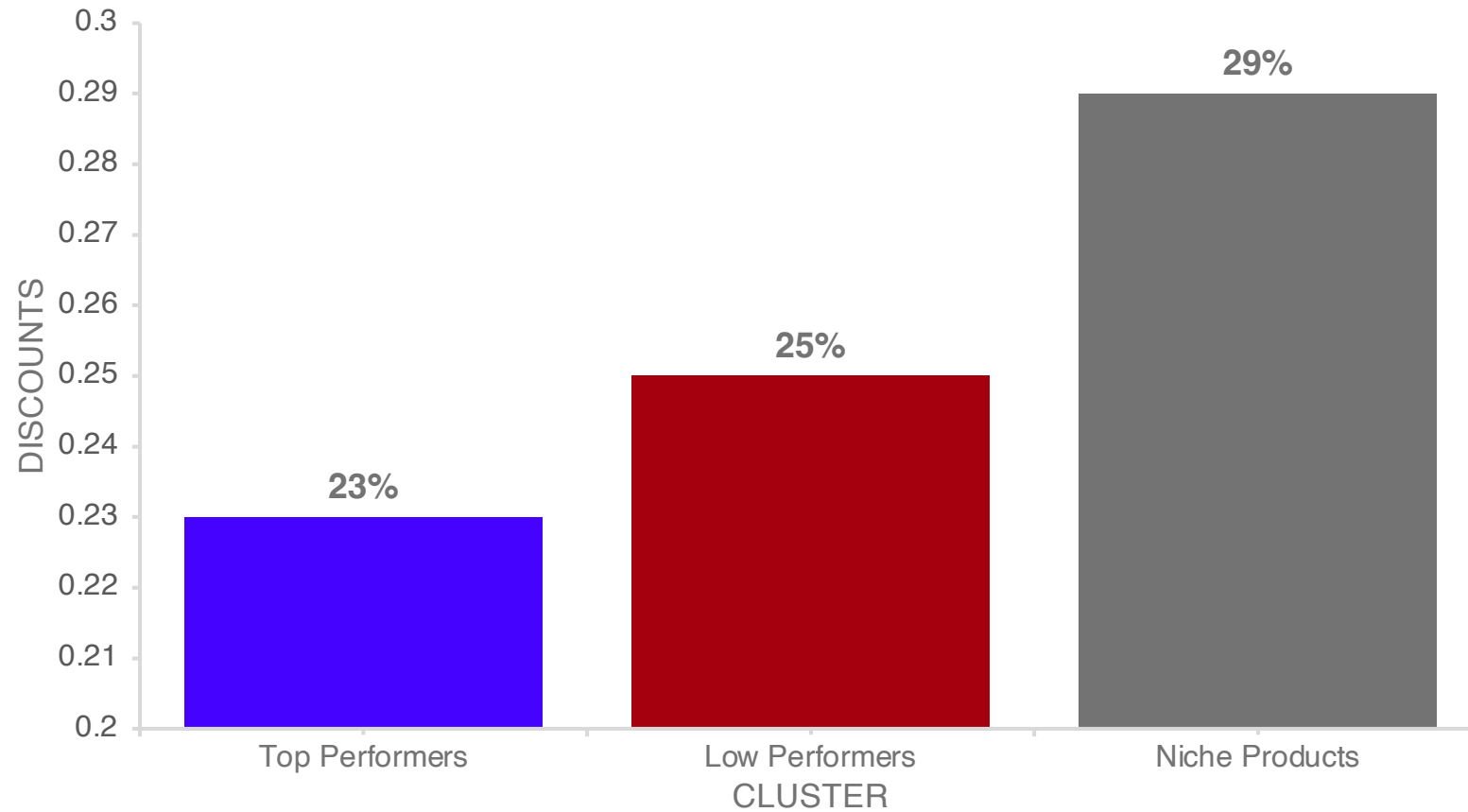

Without surprise, **Top Performers** has the **highest PLV** compared to the other clusters.

Discount Offered by Cluster

Insights into the discounting strategies across different clusters


Top performers do not need significant price reduction to drive sales as it has the **highest** performance.

Low performers have a higher discount to stimulate demand. However, it still **perform poorly** in terms of sales and revenue.

Niche products have the highest discount to attract and retain customers. With the moderate PLV, this discount is **justified**.

Customer Ratings by Cluster

Insights into customer satisfaction across all clusters
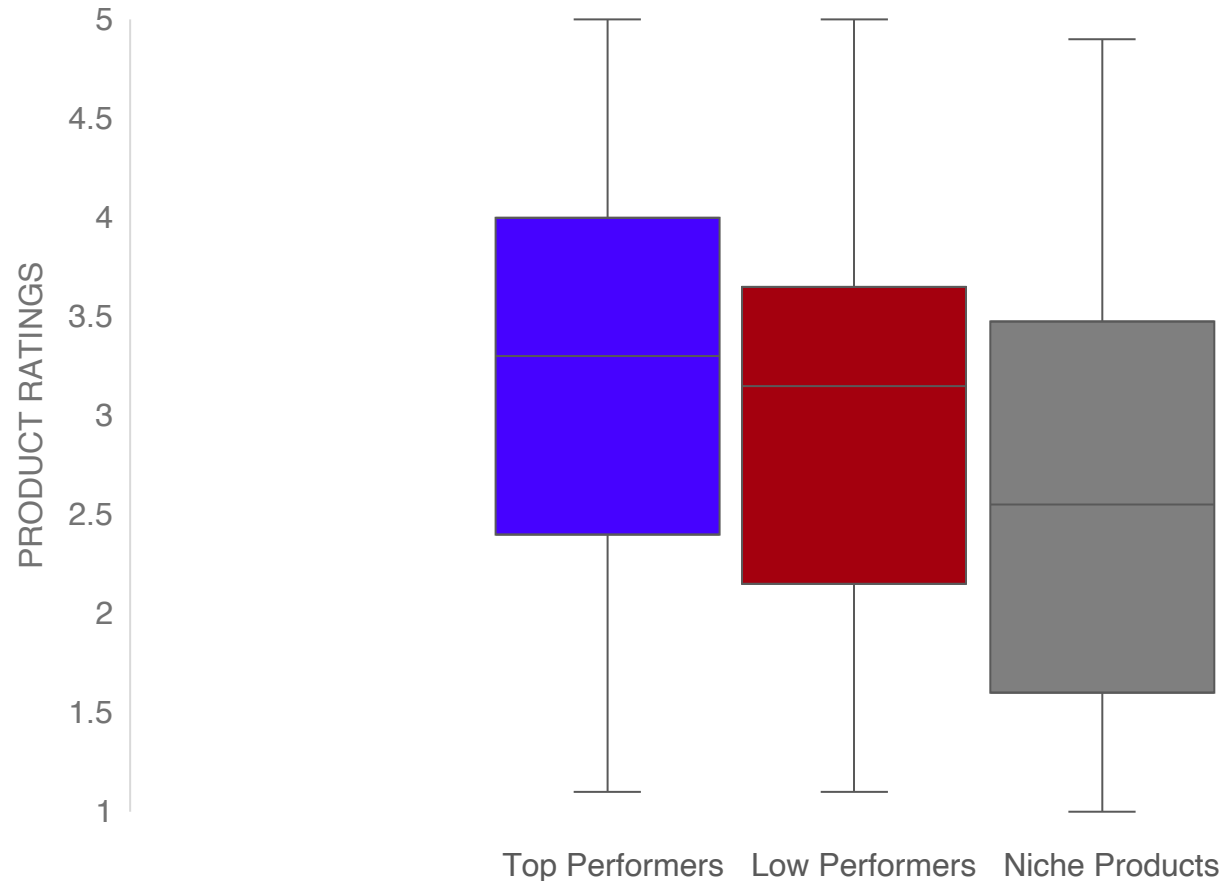
Top performers have a
strong customer satisfaction.
However, this has variability
to some customers.

Low performers have a
steady but not exceptional
customer satisfaction.

Niche products have the
lowest and widest median
rating, which is a **critical**
area for improvements.

Product Performance by Clusters Summary

Cluster	Recency	Frequency	Monetary	Dist.	PLV	Discount	Ratings	
Top Performers	L	H	H	34.07%	36 B	23%	H Score	M Range
Low Performers	L	L	L	35.26%	11 B	25%	M Score	M Range
Niche Product	H	L	L	30.77%	14 B	29%	L Score	H Range

L = Low, M = Moderate, H = High

Actionable Insights

04

Proposed Actions

Sales Optimization

Marketing Strategies

Inventory Management

Customer Experience Improvement

Sales Optimization

Dynamic Pricing Strategy

Benefits: Stimulate demand and increase market reach.

Implement dynamic pricing strategies based on seasonal trends and customer preferences for **all clusters**.

Actions Needed:
- **Develop a pipeline –** create a data pipeline that integrates real-time sales data, seasonal trends, and customer preferences.
- **Algorithm implementation –** utilize machine learning algorithms to adjust prices dynamically.
- **Testing and iteration –** conduct A/B testing to optimize pricing strategies and ensure they are effective

Sales Optimization

Premium Pricing Strategy

Benefits: Maximize revenue per unit sold

Adjust pricing for products in the **Top cluster** to reflect product value and customer perception, **increasing market demand** for top performing products.

Actions Needed:

- **Customer segmentation** – identify and target segments that perceive high value.
- **Price monitoring** – regularly monitor competitor pricing and market trends.
- **Value Communications** – enhance marketing communications to emphasize product value.

Sales Optimization

Sales Tactics

Benefits: Increase overall sales

Bundle **low-performing** products with **high-performing** ones to **create attractive value bundles.**

Actions Needed:
- **Product analysis** – identify complementary products based on purchase data.
- **Bundle creation** – design attractive bundles that offer value and encourage additional purchases.
- **Marketing campaigns** – launch marketing campaigns to promote these bundles.

Inventory Management

Demand Forecasting

Benefits: Better inventory planning

Use forecasting to **predict sales and demand trends** and inform inventory policies.

Actions Needed:
- **Data collection –** gather historical sales data and external factors affecting demand.
- **Model development –** develop and validate machine learning models for accurate forecasting.
- **Integration –** integrate forecast into inventory management systems.

Inventory Management

Inventory Optimization

Benefits: Minimize holding costs and maximize inventory efficiency

Optimize stock levels using **EOQ and demand forecasting** to manage the **supply chain.**

Actions Needed:
- **EOQ Calculation –** determine optimal order quantities and reorder points.
- **Inventory policies –** develop policies based on demand forecast and EOQ calculations.
- **Monitoring –** continuously monitor inventory levels and adjust policies as needed.

Inventory Management

Increase Stock Levels in High-demand Cities

Benefits: Meet high demand and prevent stockouts

Adjust stock based on EOQ and demand forecasting, especially in **high-demand cities**.

Actions Needed:
- **Stock allocation –** allocate inventory to cities with high median order quantities and revenue based on demand forecasts.
- **Stock monitoring –** monitor stock levels and sales trends to prevent stockouts.

Marketing Strategies

Promotion and Visibility

Benefits: Stimulate demand and increase sales conversion rates

Run targeted marketing campaign to promote low-performing products, focusing on USP and benefits.

Actions Needed:
- **Campaign design –** develop targeted campaign highlighting the unique benefits of products in the low-performing clusters.
- **Channel selection –** choose appropriate marketing channels to reach target audiences.
- **Performance tracking –** track campaign performance and adjust strategies as needed.

Marketing Strategies

Highlight Product Benefits

Benefits: Increase customer loyalty and retention

Reinforcing product value by **emphasizing premium features**, customer reviews, and benefits in **marketing** materials.

Actions Needed:
- **Content creation –** create compelling content that highlights product benefit.
- **Customer testimonials –** use customer reviews and testimonials to build trust.
- **Marketing channels –** utilize multiple channels to reach a wider audience.

Marketing Strategies

Targeted Marketing

Benefits: Increase relevance and engagement among specific customer groups

Tailor marketing messages to **niche audience**, highlighting features that **resonates** with specific demographics.

Actions Needed:
- **Audience segmentation –** segment the customer base to identify niche audiences.
- **Personalized messaging –** develop personalized marketing messages for each segment.
- **Campaign execution –** launch targeted campaigns and monitor their effectiveness.

Customer Experience Improvement

Benefits: Enhance overall customer experience

Focus on improving **customer satisfaction** for product in Niche cluster to increase **ratings** and increase customer **loyalty**.

Actions Needed:
- **Quality control –** implement stricter quality control measures.
- **Customer feedback –** collect and analyze customer feedback to identify areas for improvements.
- **Continuous improvements –** continuously improve product quality based on feedback and quality control results.

Thank You!

If you have any questions, feel free to contact me through:

✉ lazuardifauzan19@gmail.com

in Lazuardi Fauzan Primaputra

○ lazuardifp19