



E-COMMERCE CUSTOMER SEGMENTATION ANALYSIS

Final Project Data Science Batch 37

Presented by
Jessica Agnesia Tataung



OVERVIEW

- 01 About me**
- 02 My Previous Project**
- 03 Project Background**
- 04 Objective & Business Problem**
- 05 Data Understanding**
- 06 Data Analysis & Insight**
- 07 Dashboard**

Jessica Agnesia Tataung

EDUCATION

- | | |
|-------------------------|--|
| Agust 2025 -
Present | Data Science Bootcamp
dibimbing.id |
| 2020 -
2024 | Bachelor of pharmacy
Institut Teknologi Sumatera |



[Jessica Agnesia Tataung](#)



[Jessica Agnesia Tataung](#)



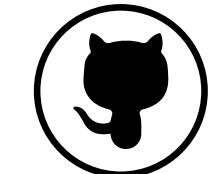
Previous project

These are several projects I explored during my learning journey with Dibimbing.id

01

Exploratory Data Analysis (EDA) on E-commerce sales

This section performs Exploratory Data Analysis (EDA) on the E-Commerce Sales dataset to understand the data structure, identify key patterns, and uncover insights related to customer behavior and sales performance.



[EDA E-commerce sales](#)

02

Advanced SQL Project on DVD Rental Database

This project is focusing on applying advanced SQL techniques in PostgreSQL. Students use subqueries, window functions, CTEs, and CASE WHEN statements to perform complex data analysis and optimize data retrieval for deeper insights.



[Advanced SQL Project](#)

MAIN PROJECT

>>>

Project Background



Context

E-commerce is one of the fastest growing industries in the digital era. Intense competition drives businesses to understand customer behavior, sales trends, and product profitability to make data-driven decisions. Customer segmentation helps identify purchasing patterns and loyalty levels to support more targeted marketing strategies.

Objective

The main objective of this analysis is to perform Exploratory Data Analysis (EDA) and Customer Segmentation using RFM (Recency, Frequency, Monetary) techniques to gain deeper insights into customer behavior.

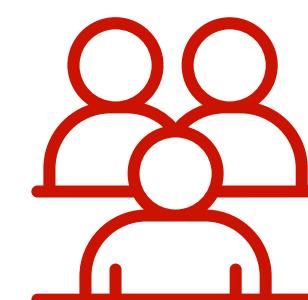
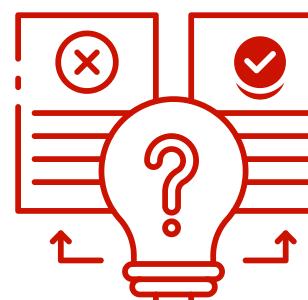
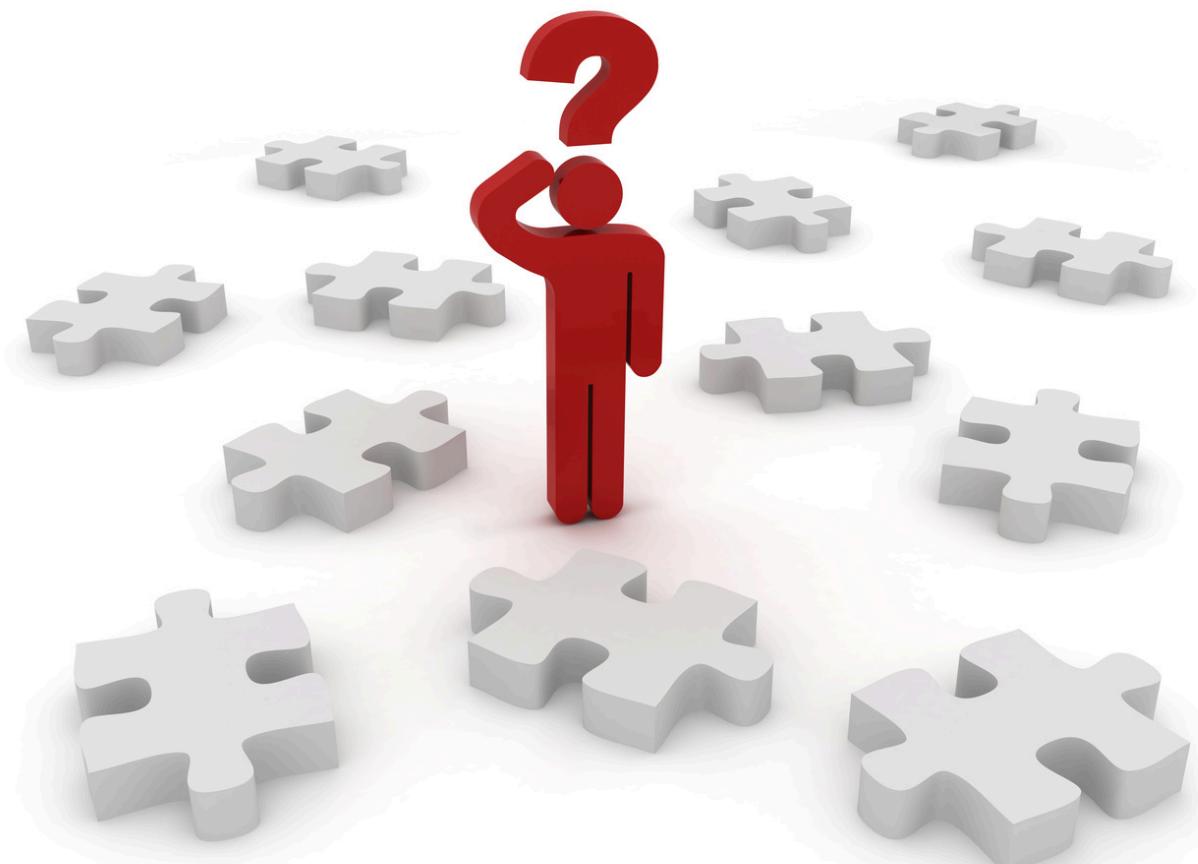
This project aims to:

- Identify customer distribution across different RFM segments
- Analyze profitability and loss by segment, product category, and country.
- Evaluate shipping cost and discount impact on overall performance.
- Provide actionable recommendations to increase customer retention and revenue.

Limitation

This analysis only focuses on historical transaction data available in the dataset. External factors such as market conditions, promotions, or competitor activities are not included within the scope of this study.

Business Problem



Key Question

Which product categories, customer segments, and regions contribute the most to total sales and profit, and how do different customer groups perform based on RFM segmentation?

Hypotheses

H_0 : There is no significant difference in sales or profit among product categories, customer segments, or regions.

H_1 : There is a significant difference, where certain categories, segments, or regions perform better than others.

Stakeholder Needs

Provide insights to management and marketing teams for data-driven decisions by identifying high-performing products, customer segments, and regions to improve revenue, retention, and profitability..

Data Understanding

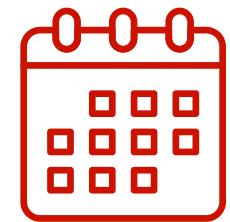


[Link dataset E-commerce Website Sales](#)

kaggle Collected from **Kaggle**, focusing on **e-commerce website sales**.



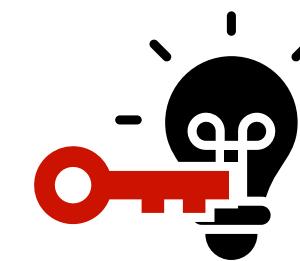
Includes **order details, customer information, shipping data, and product performance**



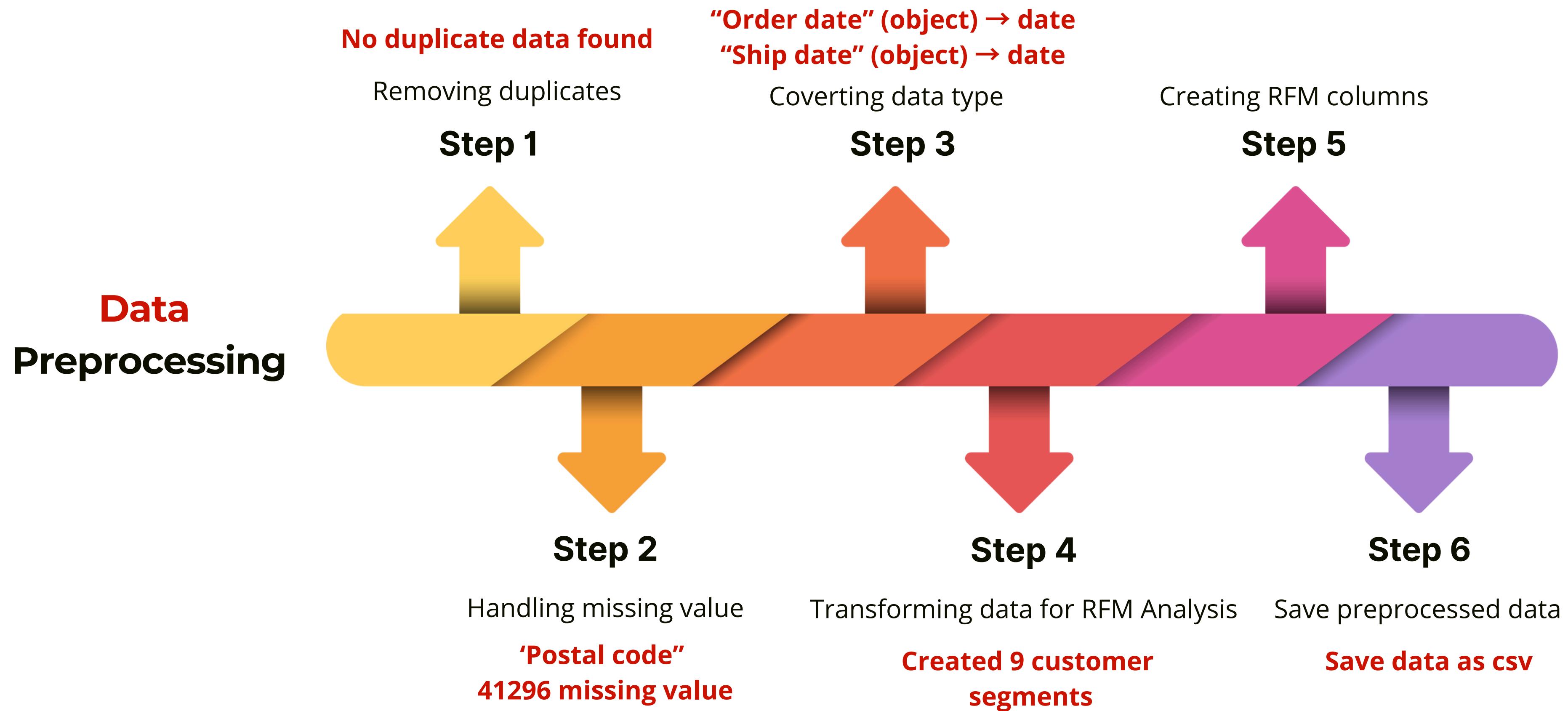
Covers transactions from **2011 - 2014**



Contains **51290 transactions** data and **24 columns**



Key columns : Order ID, Order Date, Ship date, Customer ID, Customer Name, Segemnt, City, State, Country, Category



Finding and Result

Category	Recency Score	Combined Score (Frequency & Monetary)
Champions	4-5	4-5
Loyal Customers	2-5	3-5
Potential Loyalist	3-5	1-3
Recent Customers	4-5	0-1
Promising	3-4	0-1
Customers Needing Attention	2-3	2-3
About to Sleep	2-3	0-2
At Risk	0-2	2-5
Can't Lose Them	0-1	4-5
Hibernating	1-2	1-2

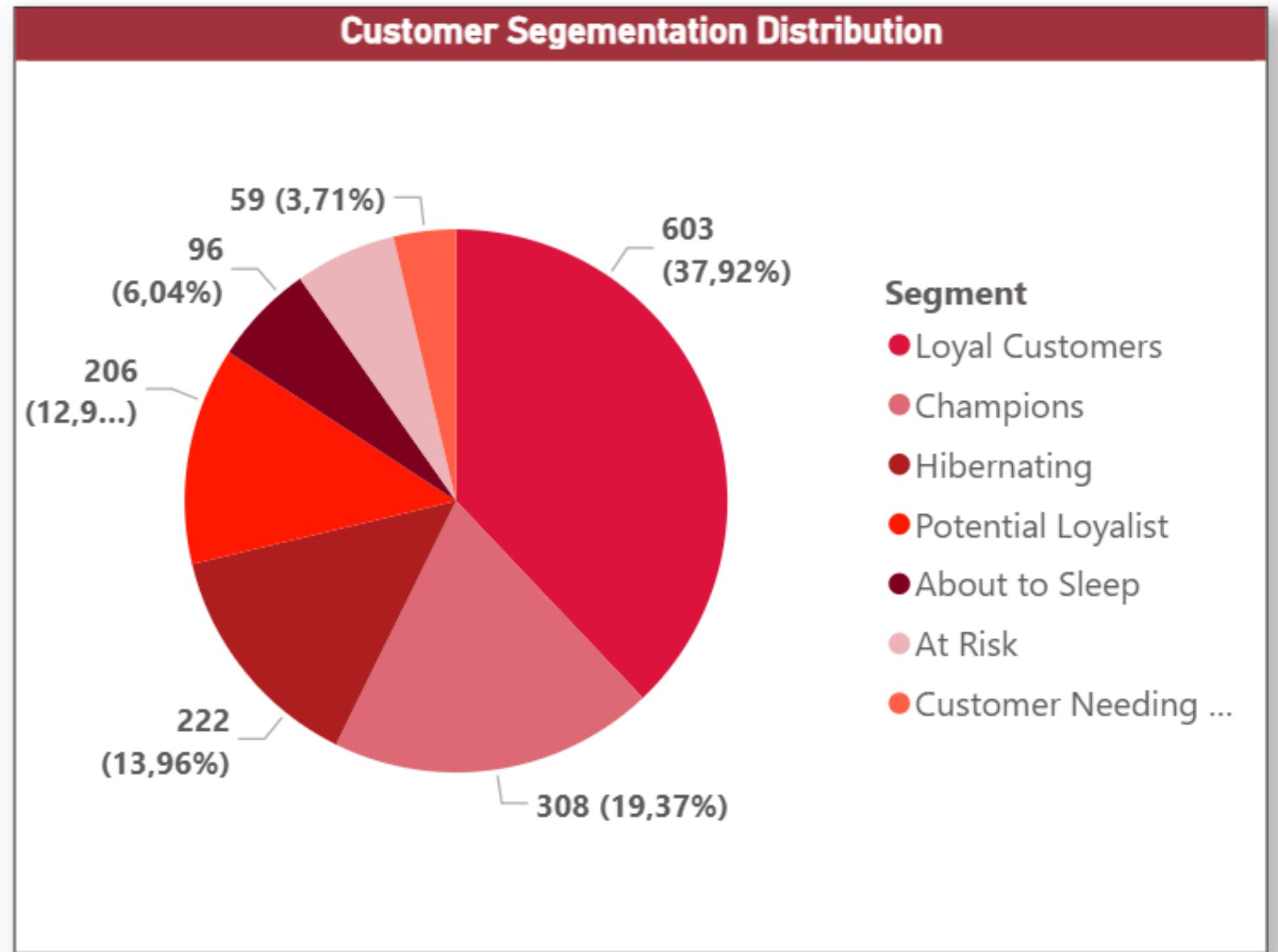
Finding and Result

How is the distribution of customers across RFM segments?

Insight:

Based on the RFM segmentation distribution:

- Majority are **Loyal Customers (37,92%)**, showing strong retention.
- **Champions (19,37%)** are high-value and should be prioritized.
- **Potential Loyalists & Hibernating (13,96% & 12,9%)** show growth opportunity.
- Small segments are At Risk / About to Sleep, needing reengagement



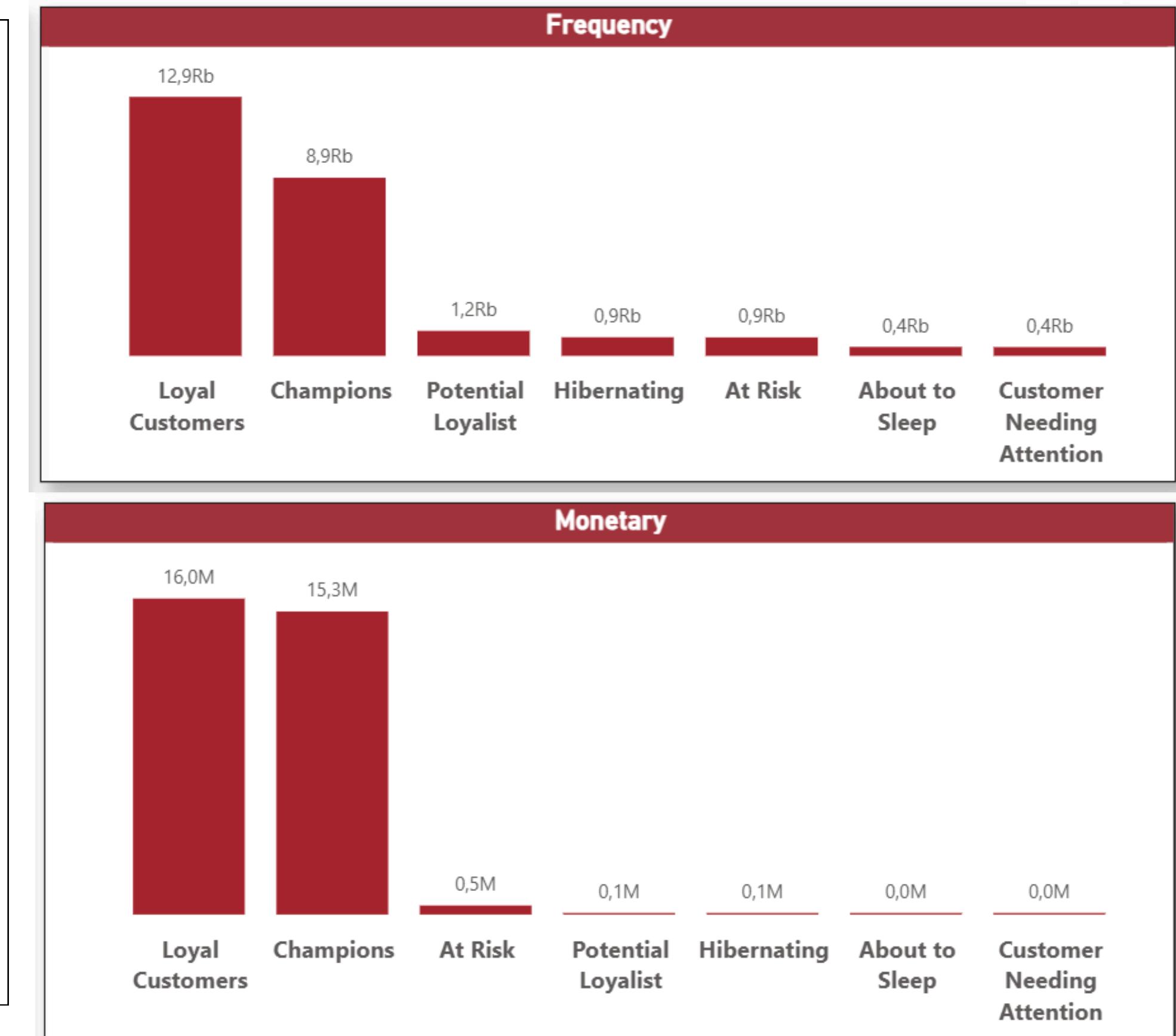
Finding and Result

How are frequency and monetary values distributed across customer segments?

Insight:

Based on the analysis of frequency and monetary values:

- The **Loyal customer and champion** are segment shows the **highest frequency and monetary values**
- The **Potential loyalist** segment ranks third in frequency, however, their **monetary value is relatively low**
- The **At risk** segments demonstrate **consistent purchasing activity** and contribute **amount of spending**
- Other segments such as **Hibernating, about to sleep, and customer needing attention** still did purchasing activity but they spent low monetary values



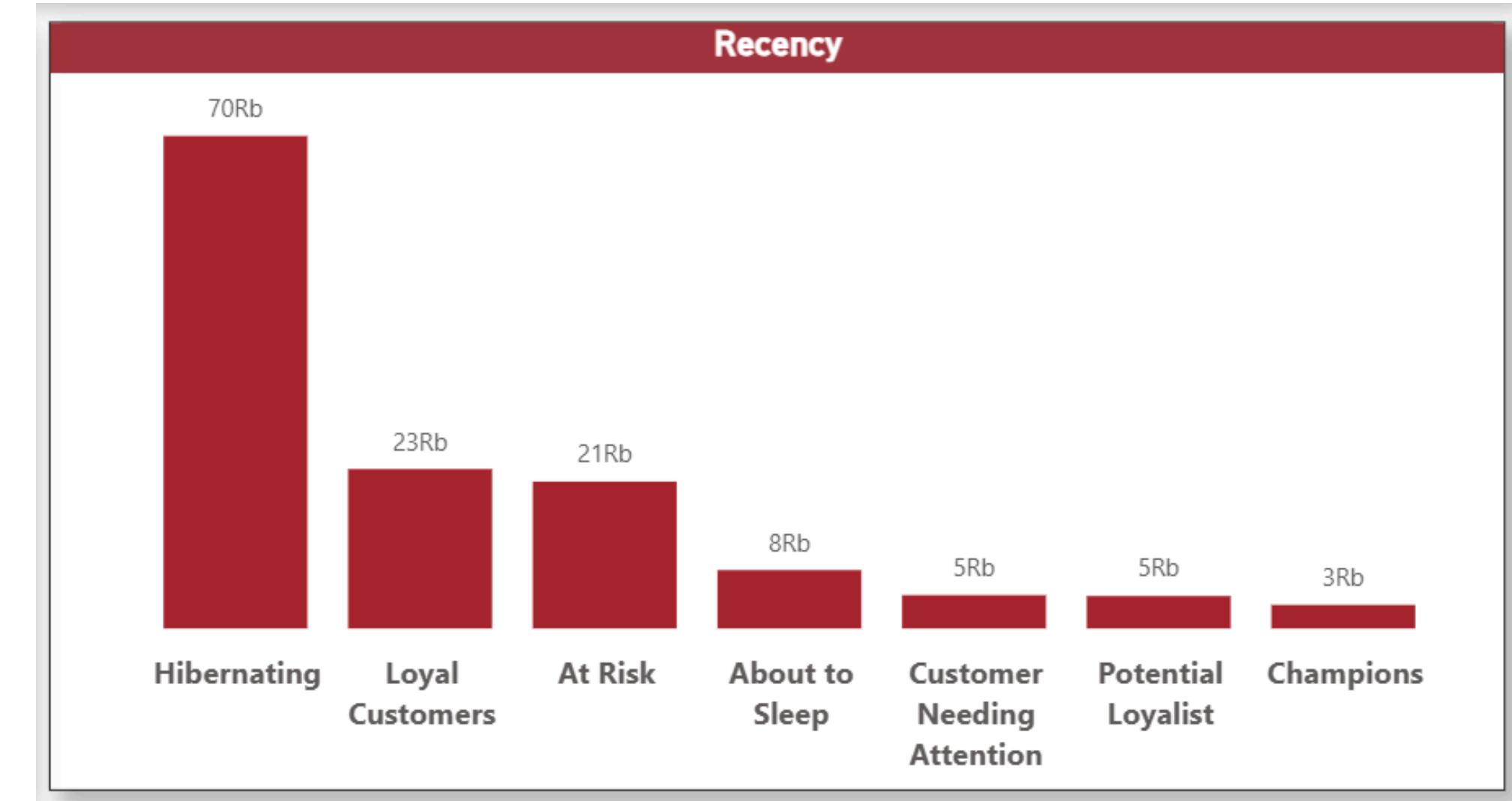
Finding and Result

How is recency values distributed across customer segments?

Insight:

Based on the analysis of recency values:

- Majority is **Hibernating** segment of the longest customer that have not made purchases
- Most segments **have not made purchases for a long time**, suggesting a widespread decline in recent customer activity and indicating the need for re-engagement strategies.



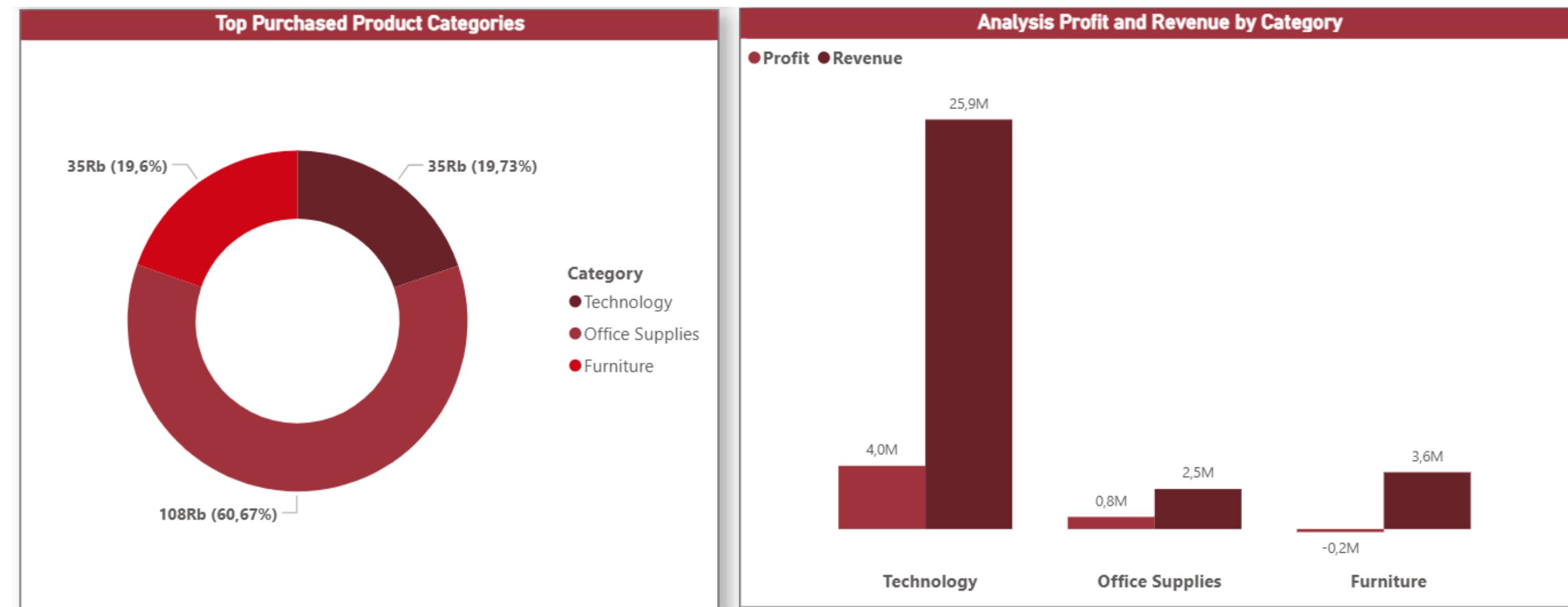
Finding and Result

Which product categories are most in demand, and which categories generate the highest revenue and profit?

Insight:

Based on the analysis:

- **Office Supplies** is the **most purchased product category**, with approximately **108,000** units sold
- **Technology** category generates both the **highest revenue and highest profit**, with approximately revenue around **26M** and profit around **40M**
- **Furniture** category shows a **loss** of approximately **0.2M**

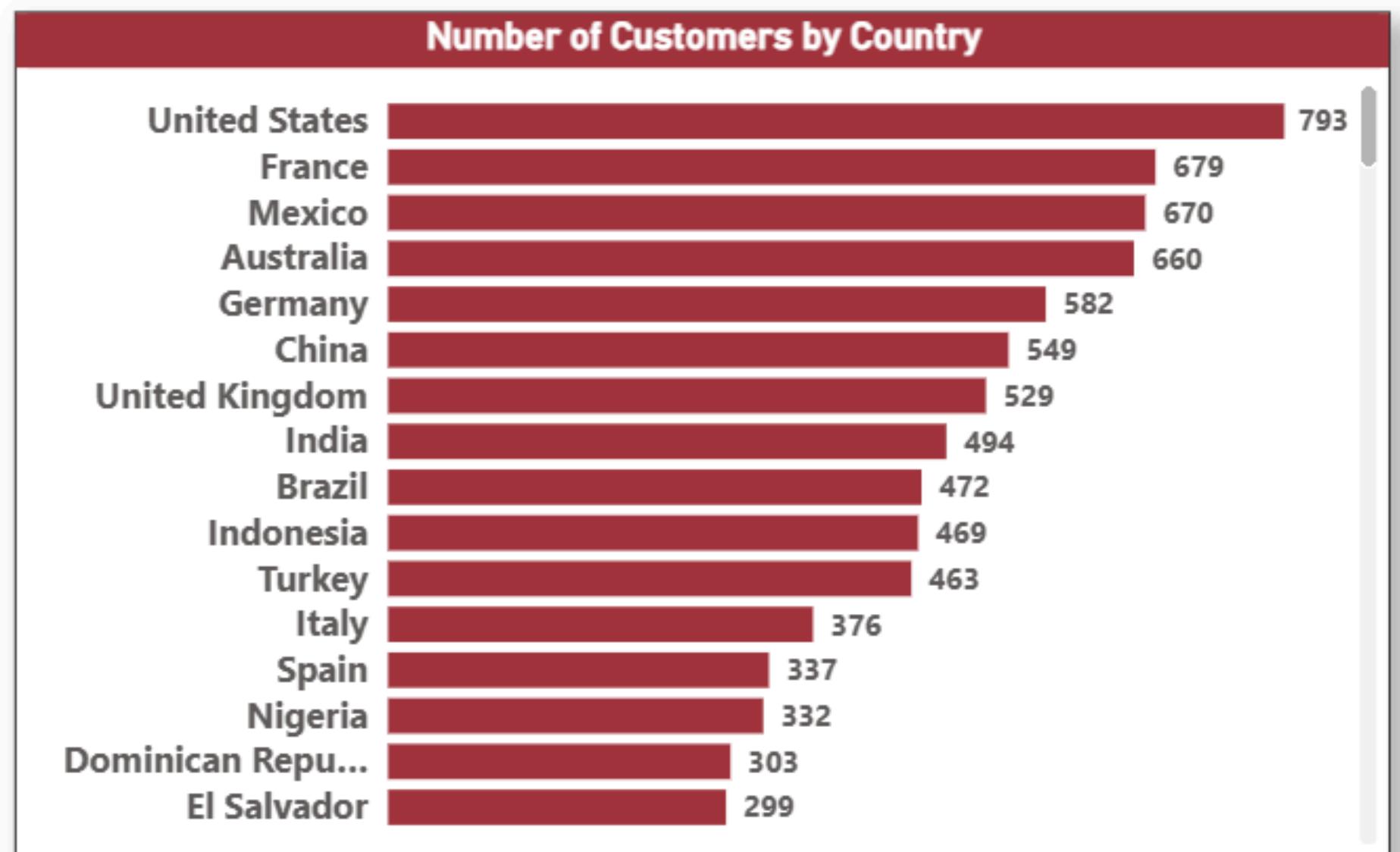


Finding and Result

How is the distribution of customers across countries?

Insight:

- The customer base is distributed across various countries.
- The **majority** of customers are from the **United States**, totaling **793 customers**, indicating a strong market dominance in this region.
- Other countries with relatively **high customer** counts include **France, Mexico, Australia, Germany, and the United Kingdom**, reflecting solid market presence.
- In contrast, countries such as **Guatemala and New Zealand** have a **low number of customers**, suggesting potential opportunities for future market expansion.

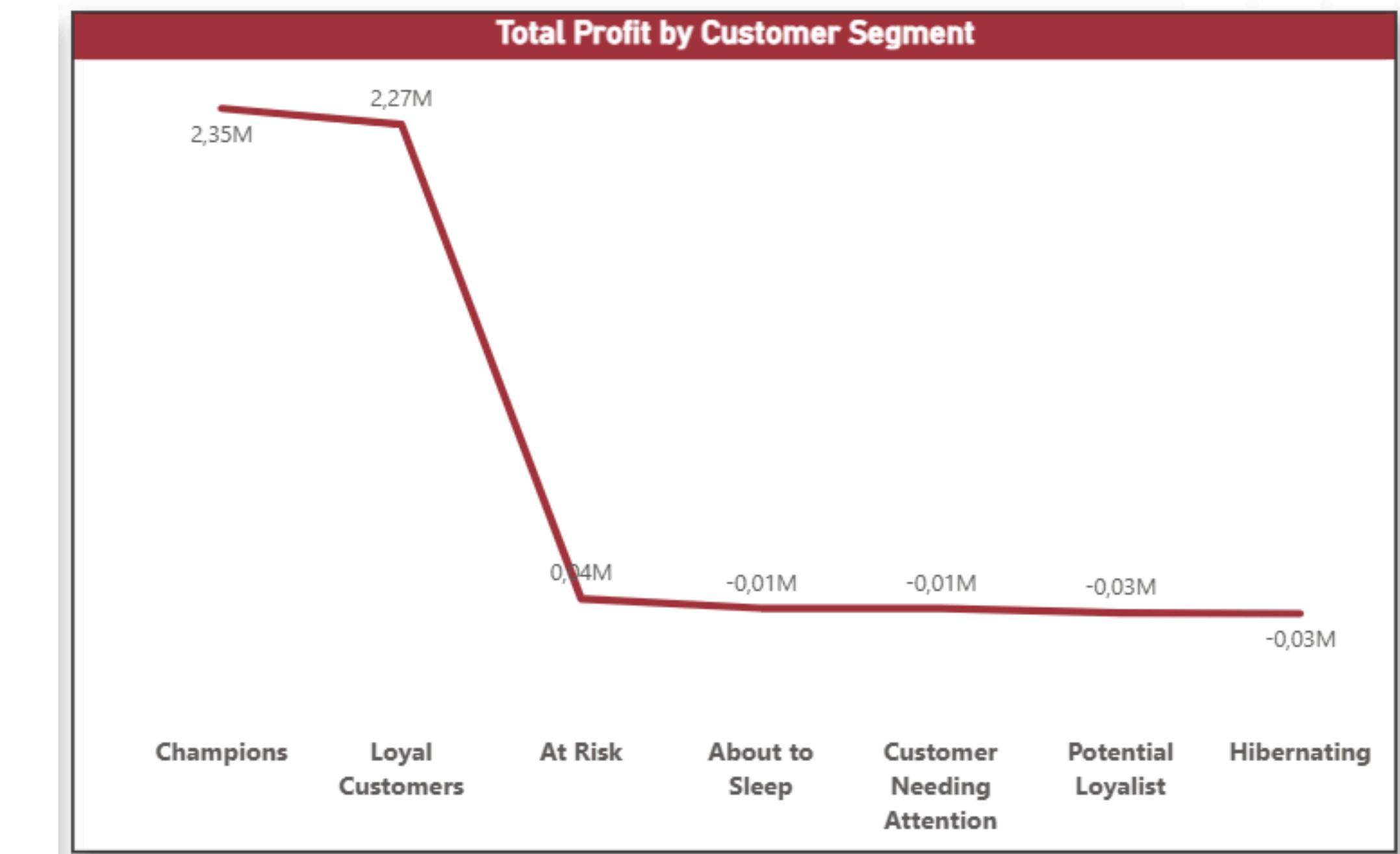


Finding and Result

How is sales profit distributed across customer segments?

Insight:

- **Champions (2.35M)** and **Loyal Customers (2.27M)** are the main profit contributors.
- Other segments show **very low to negative profit**, mainly due to **shipping cost losses**.

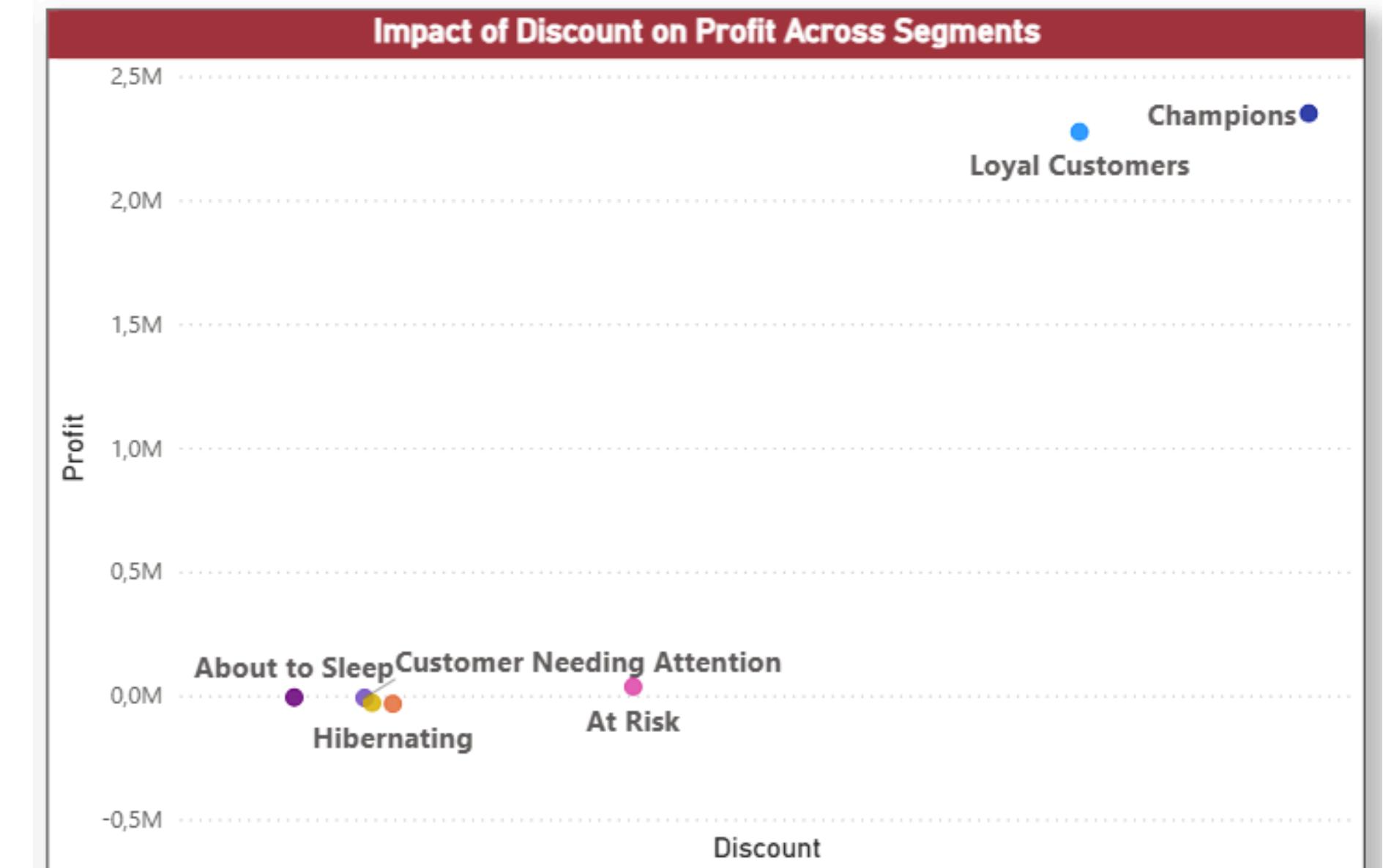


Finding and Result

How does discount level affect profit across customer segments?

Insight:

- **Champions and Loyal Customers**, even with **higher discount levels**, still generate high profit, showing strong buying behavior.
- **Other segments** receive **lower discounts** yet produce **low or even negative profit**, indicating that small discounts are not enough to drive profitability in those groups.

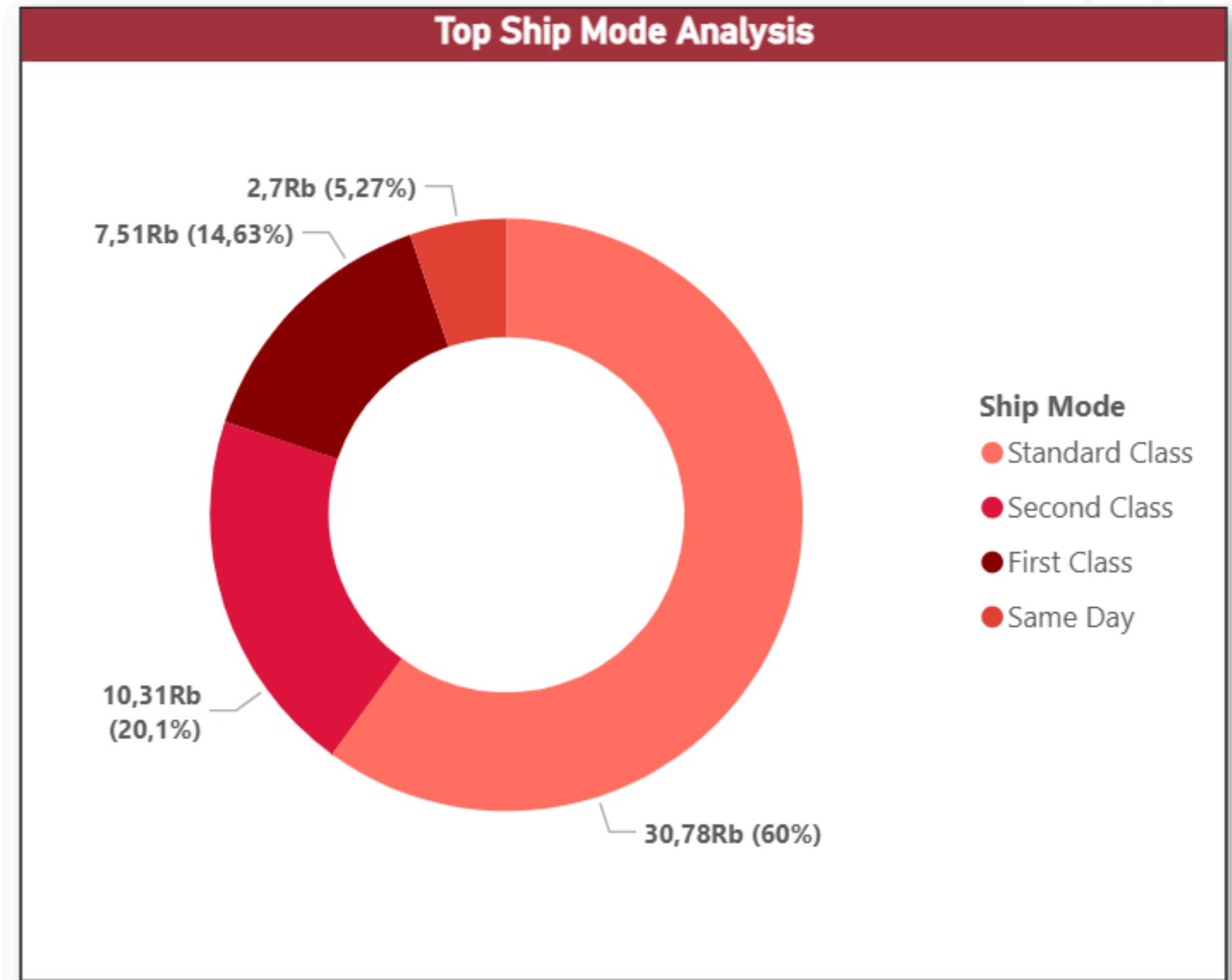


Finding and Result

How do customers prefer shipping methods?

Insight:

- The **majority** of customers prefer **Standard Class** as their primary shipping method, accounting for approximately **60%** of all shipments. Meanwhile, **Second Class** and **First Class** represent smaller portions at **20.1%** and **14.63%**, respectively. **Same Day** delivery has the lowest adoption rate at only **5.27%**.
- These findings indicate that customers **prioritise lower shipping costs over faster delivery** options.

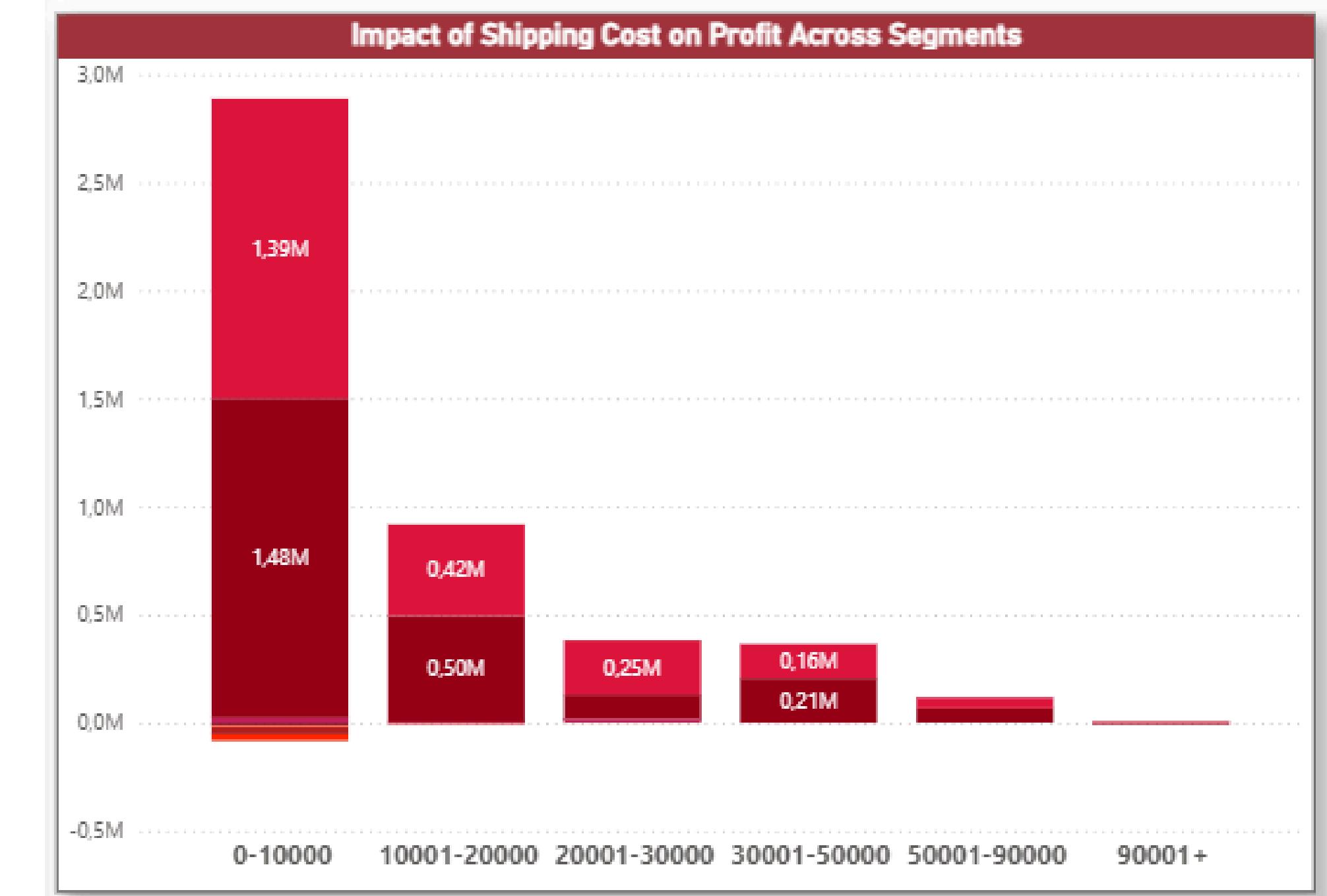


Finding and Result

How does shipping cost affect profit across different customer segments?

Insight:

- Highest profit occurs at **shipping costs** of 0-**10,000**, but most losses are also found in this range.
- As **shipping costs increase, profit** consistently **declines**.
- At **shipping costs > 50,000, profit** becomes **very small and may even turn negative**



Conclusion

- Majority of value comes from **Loyal Customers & Champions**, other segments show **low activity** and **need reengagement**.
- Technology drives the **highest revenue & profit**, Furniture shows **losses**
- Customer base is mostly in the **U.S.**, with room for expansion elsewhere.
- Higher discounts still **profitable** for **top segments**, while others remain low/negative.
- Profit is **highest** at shipping cost 0–10,000, but **losses also occur**
- Mostly **profit declines** and **turns negative** as **shipping costs rise**.
- Most customers prefer **Standard Class** shipping



Recommendation action

- Prioritize and retain Loyal Customers & Champions through exclusive offers and personalized experiences.
- Reengage inactive segments via targeted campaigns (email, vouchers, remarketing).
- Focus on high-profit categories (Technology), review pricing & cost structure for Furniture.
- Expand presence in countries with low customer counts to grow market share.
- Optimize discount strategies for low value segments to improve profitability.
- Manage shipping cost structure to prevent margin erosion at high cost tiers.

DASHBOARD

>>>

E-Commerce Customer Segmentation Analysis

Date
Semua

Month
Semua

Year
Semua

RFM Total

Orders

Ship Mode

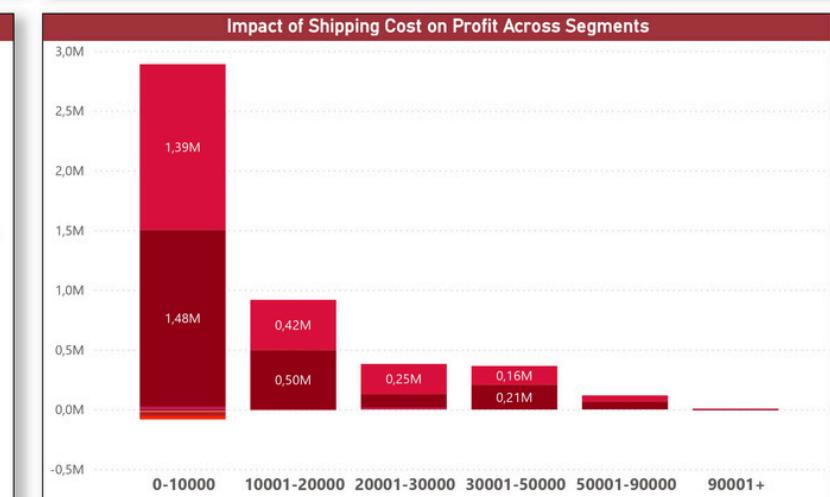
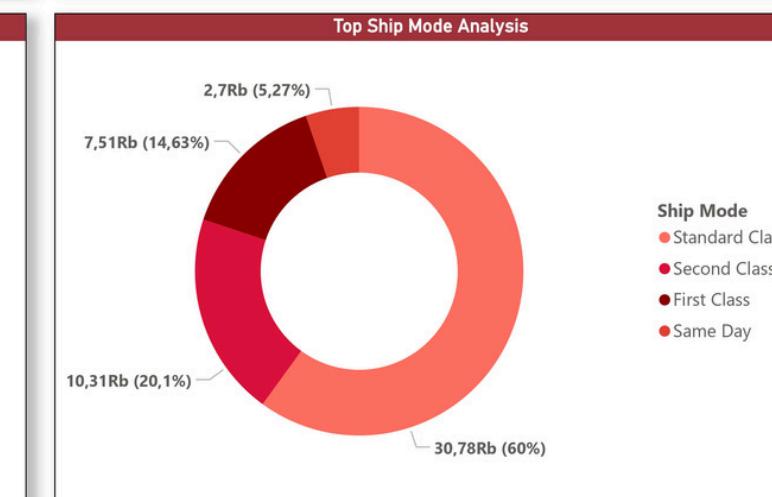
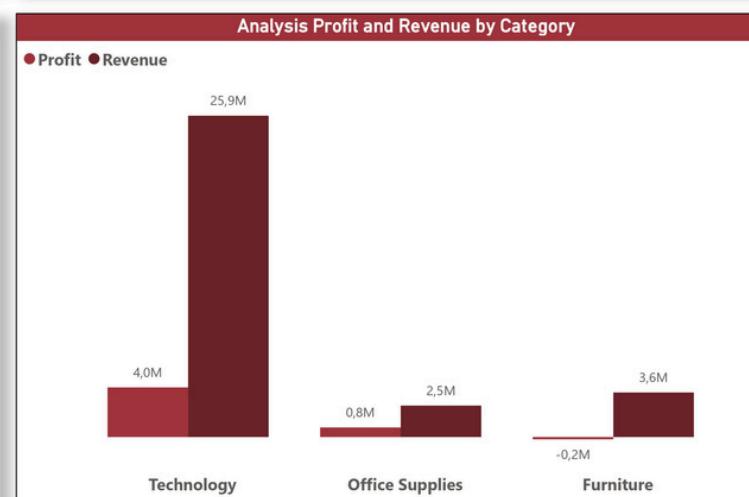
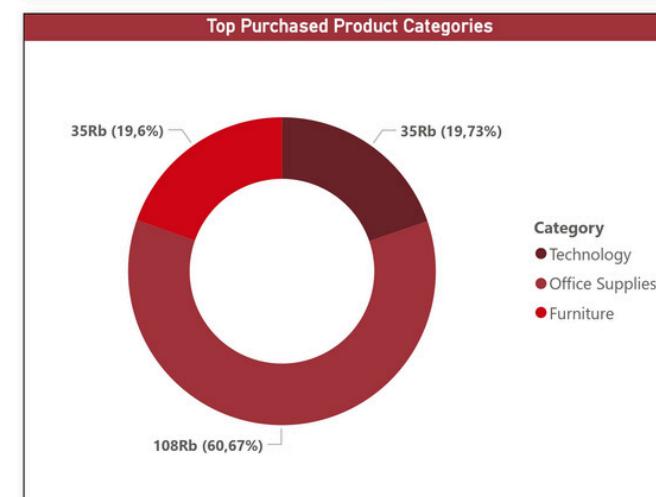
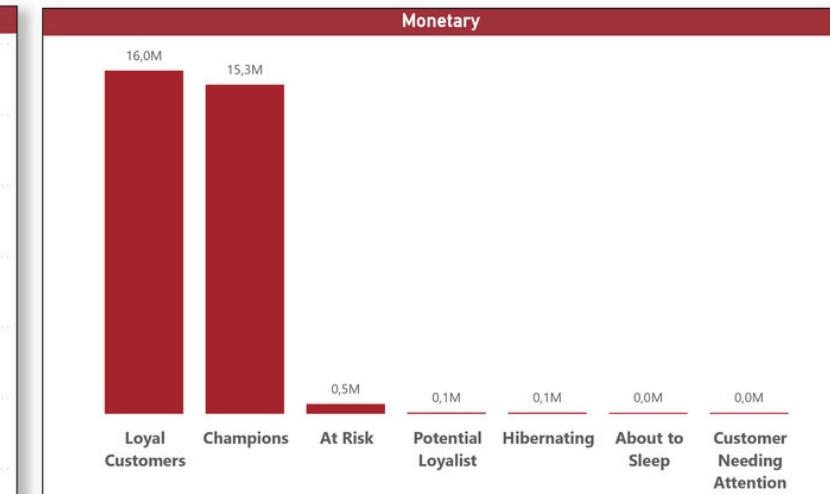
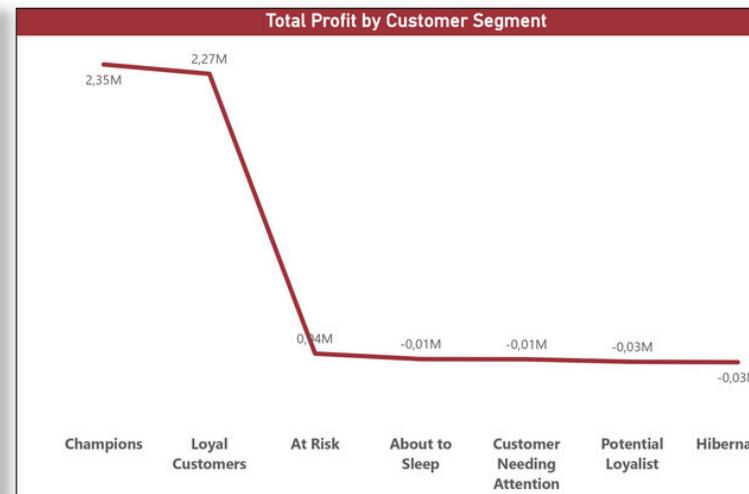
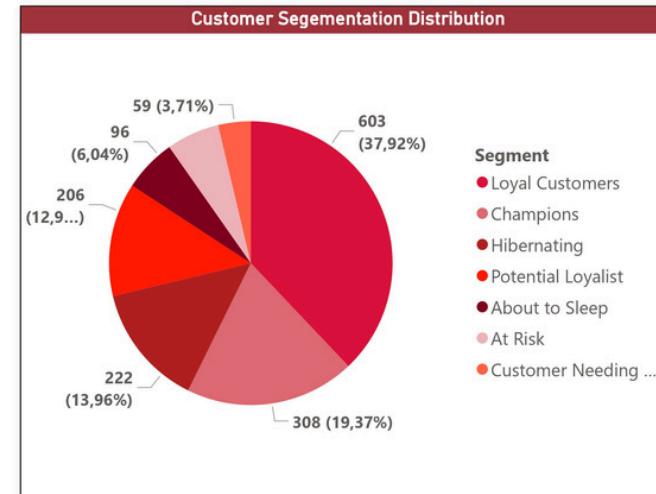
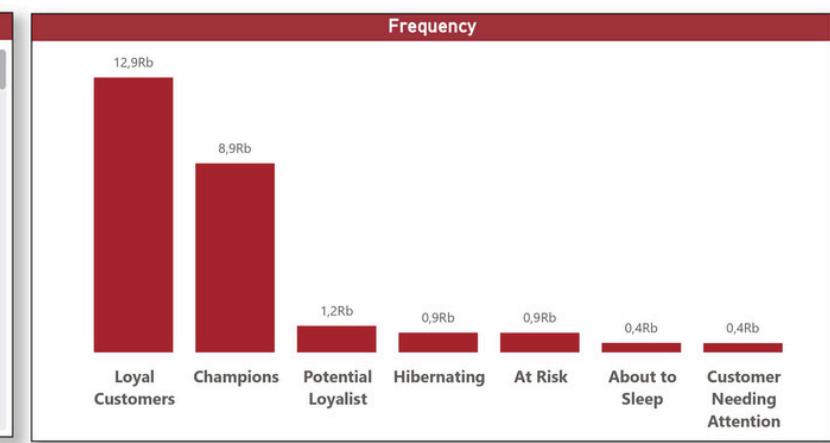
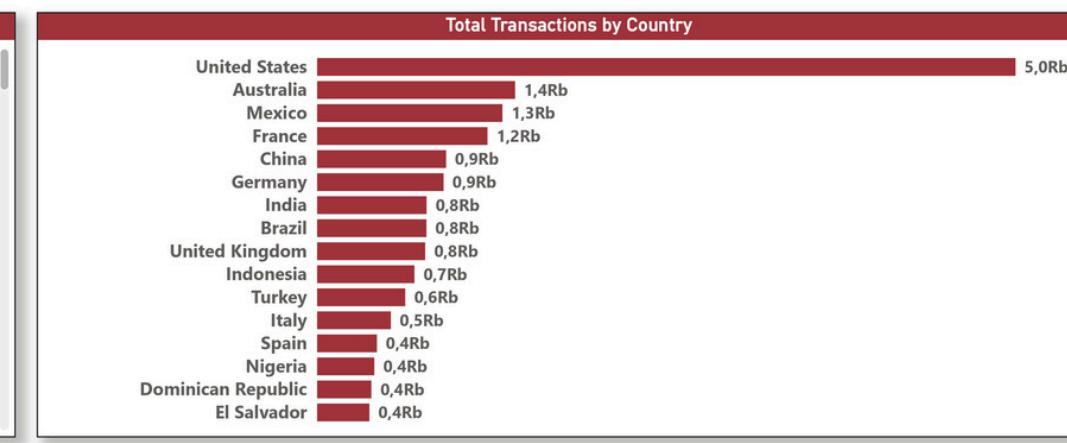
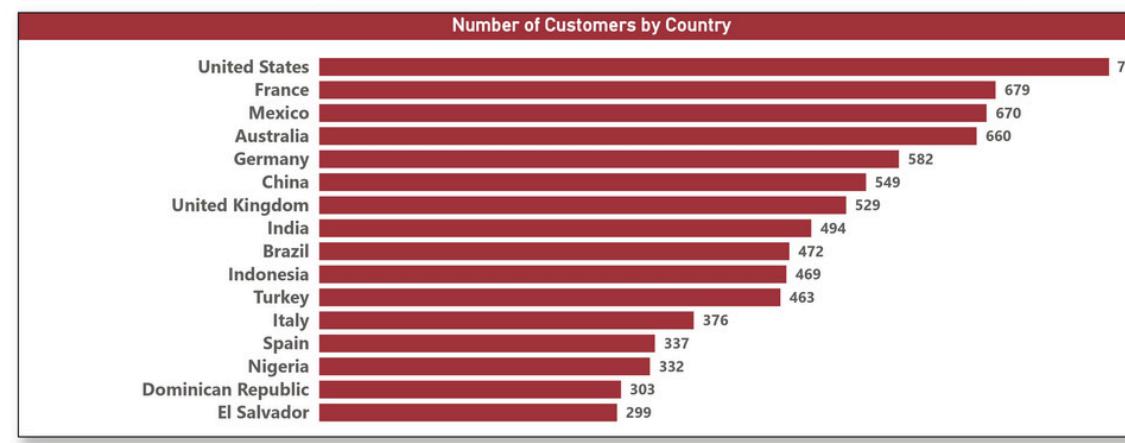
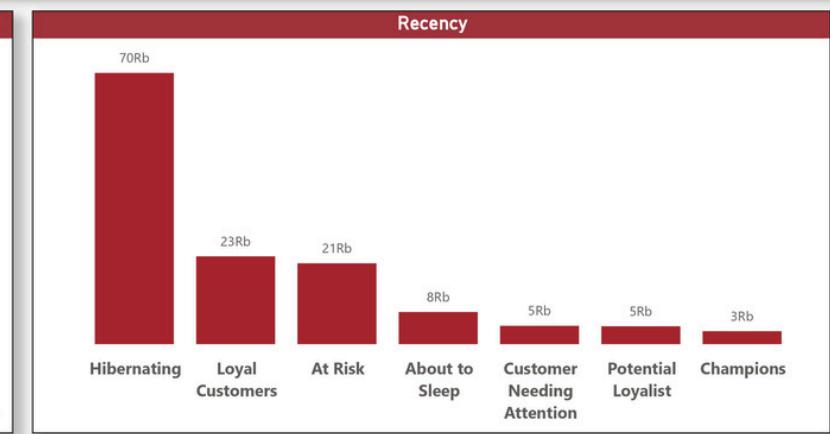
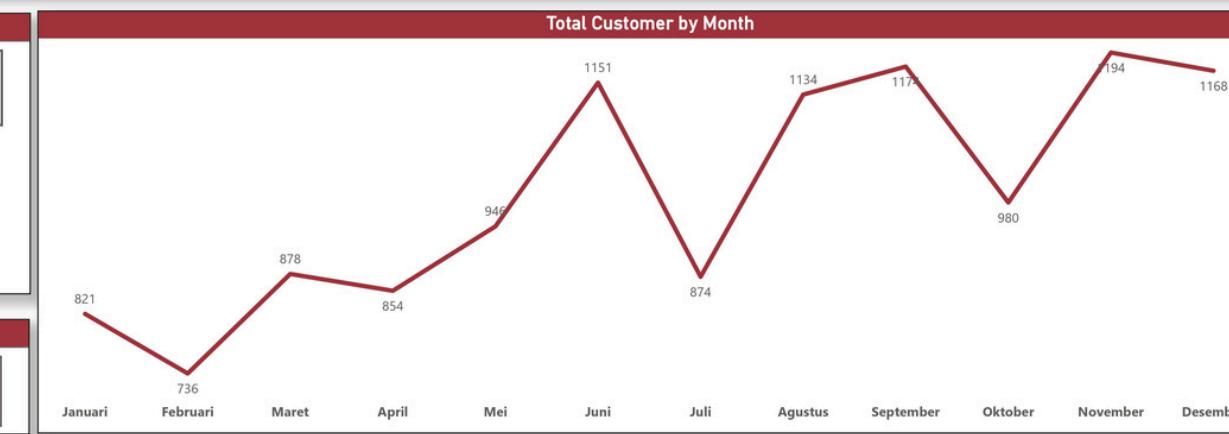
Count of Customer
1590

SEGMENT CUSTOMER

About to Sleep	Customer Needing Attention	Potential Loyalist
At Risk	Hibernating	
Champions	Loyal Customers	

CATEGORY

Furniture	Office Supplies	Technology
-----------	-----------------	------------





THANK YOU

@jessicaagnesiat / +62895351100201
[linkedin.jessicaagnesiat](#)