

# FINAL PROJECT: SUPERSTORE PRODUCTS DATA RESEARCH ANALYSIS

- ▶ Student: Rina Rafalski
- ▶ Mentor: Shiran Alon
- ▶ Program: DRA





**Rina Rafalski**  
Data Analyst

### Certificates

- Microsoft Data Science Professional Certificate
- IBM Certificate Python for Data Science

### Experience

- Data Science Intern
- 5+ years as Marketing Data Analyst

### Areas of Interest

- Data Analysis
- Data Science
- Business Intelligence

About me

# Table of contents

▶ Introduction .....	3
▶ Data Description .....	4
▶ Descriptive Analysis .....	5
▶ Impact of Discount Strategies .....	11
▶ Product Bundling Opportunities .....	15
▶ A/B Testing .....	18
▶ Dashboard .....	21
▶ Conclusion.....	22







# Introduction

- ▶ This project aims to provide a comprehensive overview of products of Superstore retail business located in the United States which is specialized on selling Office Supplies and Technology products, furniture etc. focusing on products sales, and profitability, with a specific emphasis on high- and low-profit products, impact of discount strategies and uncovering opportunities for product bundling.

- ▶ **Initial Research Question of the project:**

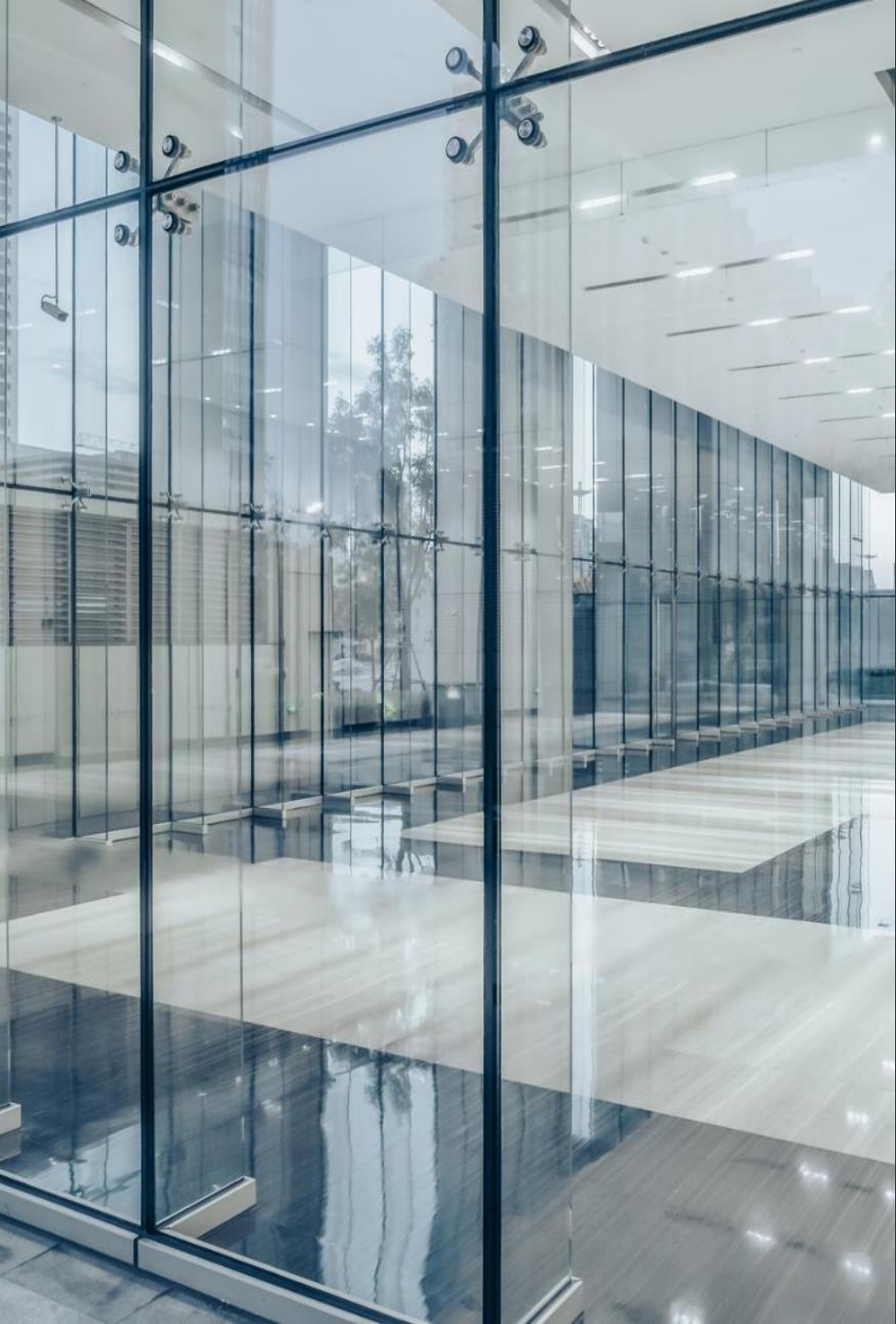
**How can the company optimize its product portfolio by identifying high- and low-profit products, determining the impact of discount strategies on sales and margins, and uncovering potential opportunities for product bundling?**



# Data description

- ▶ The source of the data is [Superstore Sales](https://www.kaggle.com/datasets/santoshdeepak2/superstore) dataset from online community platform for data scientists [www.kaggle.com](https://www.kaggle.com).
- ▶ The dataset contains the orders for 4 years (2011-2014) and consists of 51290 records with 21 features that can be divided in following kinds:
  - **Sales data** - order date, order ID, quantity, and sales
  - **Product information** - product ID, quantity, category and subcategory
  - **Geographical segmentation** - country, state, and city
  - **Profit** - profits generated from sales transactions
  - **Consumer segmentation** - customer ID, segment, and region
- ▶ **Data preprocessing** steps contained:
  - formatting data types
  - dropping rows with missing values of sales
  - treatment outliers using IQR (Interquartile Range)
  - adding derived columns
- ▶ The final dataset shape: 44152 rows, 22 columns, no missing values

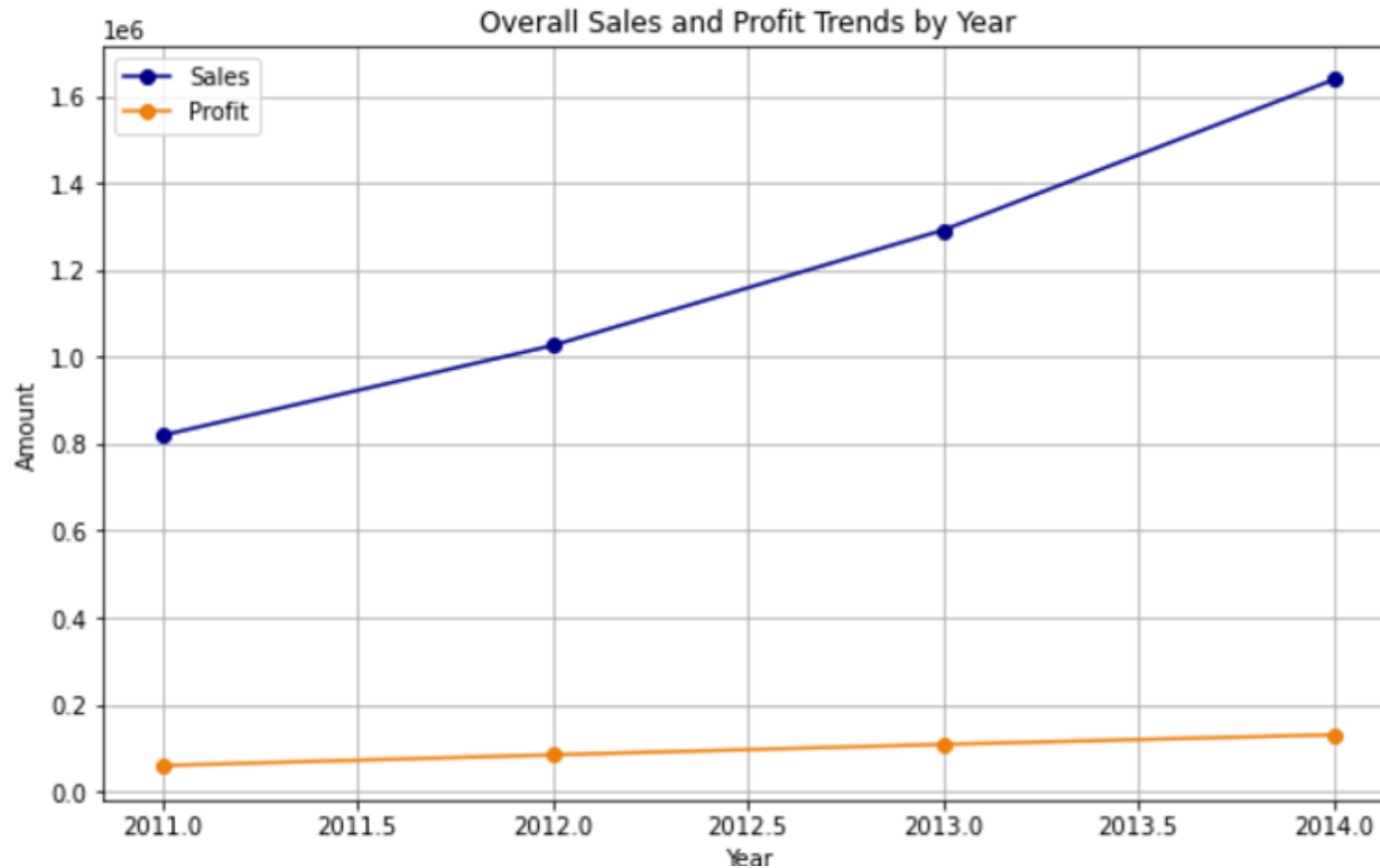




# Descriptive Analysis

# Descriptive Analysis

## Overall Sales and Profit Trends



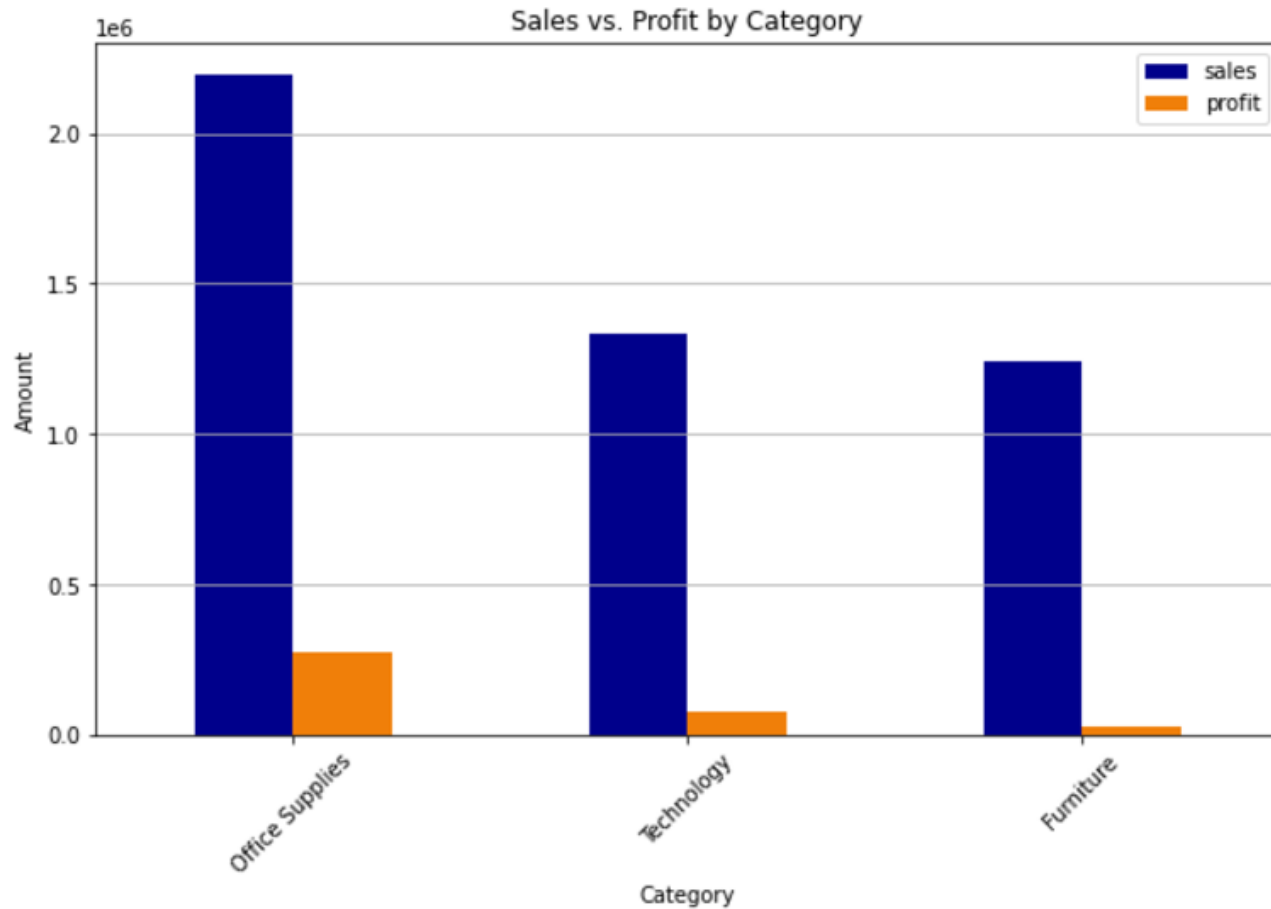
- Sales have consistently increased year over year, showing a positive growth trend.
- Profit also shows a gradual increase but remains significantly lower compared to sales.

### Insights:

- The business is growing in terms of sales volume, but profit margins appear to be relatively small.
- Further analysis could explore why profits are not increasing at the same rate as sales (e.g., high discounts, operational costs, or unprofitable products).

# Descriptive Analysis

## Sales vs. Profit by Category



- Office Supplies has the highest sales and profit among all categories.
- Technology has moderate sales and profit
- Furniture has high sales but low profit, suggesting smaller margins or operational inefficiencies.

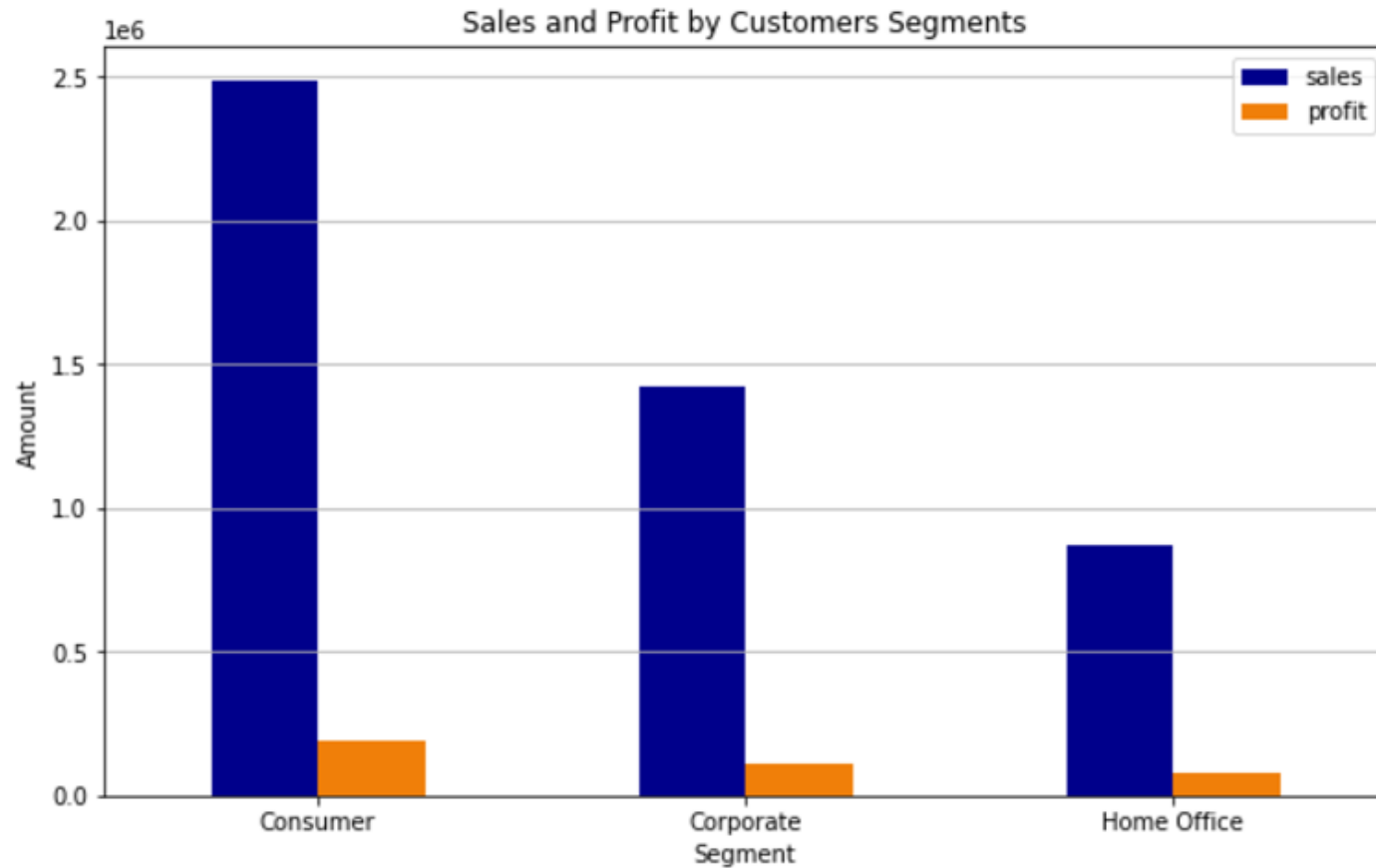
### Insights:

- Focus on improving profitability in the Furniture category.
- Office Supplies is the most profitable category and should be a priority for expansion or marketing.



# Descriptive Analysis

## Sales and Profit by Customers Segments



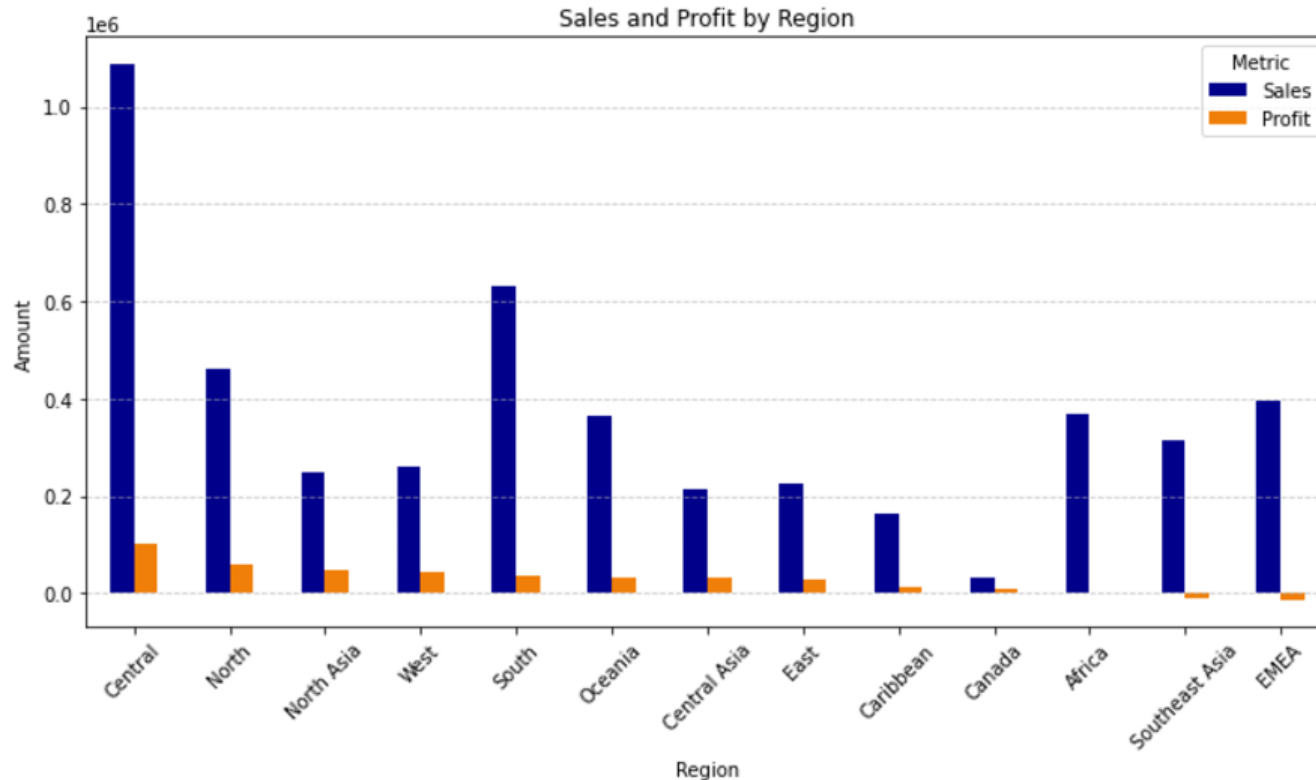
- Consumer Segment has the highest sales and profit.
- Corporate Segment and Home Office show lower sales and profits.

### Insights:

- The Consumer Segment is the key driver of revenue and profit, suggesting that marketing and operational efforts should focus more on this segment.
- Explore opportunities to grow Corporate and Home Office segments to balance revenue streams.

# Descriptive Analysis

## Sales and Profit by Region



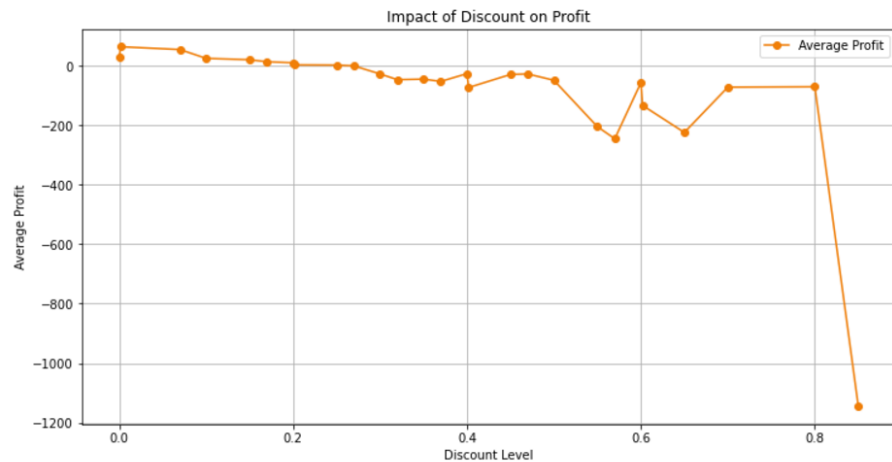
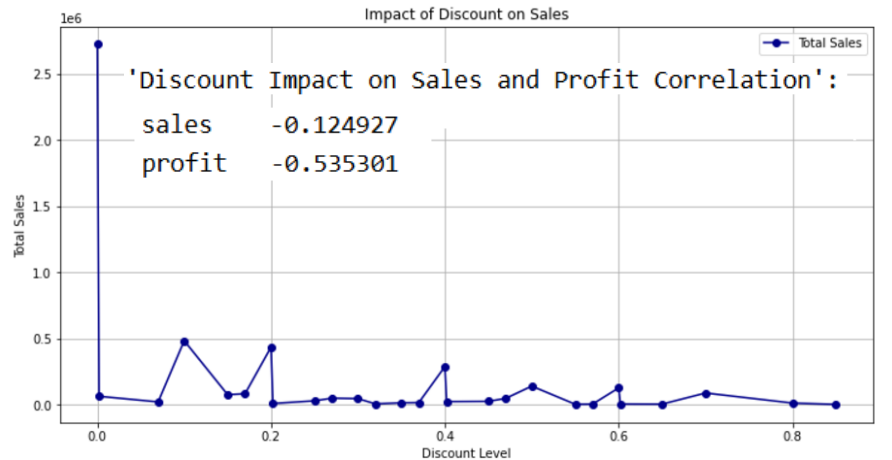
- **Central Region**
  - Dominates both sales and profit.
  - Indicates as the primary revenue generator.
- **South Region**
  - Shows high sales, but profits are comparatively much smaller.
  - This could suggest operational inefficiencies, high discounts, or high costs in this region.
- **Africa, Southeast Asia, and EMEA**
  - Moderate sales and low profits.
  - May not be as profitable due to high costs or low pricing strategies.
- **Canada and Caribbean**
  - Minimal sales and almost negligible profit.
  - Smallest contributors to the business currently.
- **North Asia and Oceania**
  - Moderate sales but very low profit margins.
  - Indicates potential challenges in scaling profitability.



# Impact of Discount Strategies

# Impact of Discount Strategies

## Regression Analysis



- **Impact of Discounts on Sales**
  - Weak negative correlation (-0.12)
  - Sales are highest at a 0% discount level, indicating that a significant portion of revenue comes from products sold at full price.
- **Impact of Discounts on Profit**
  - Moderate negative correlation (-0.54)
  - Profit decreases drastically as discounts increase.
  - At higher discount levels (e.g., 80%+), profit turns negative, indicating significant losses for those transactions.

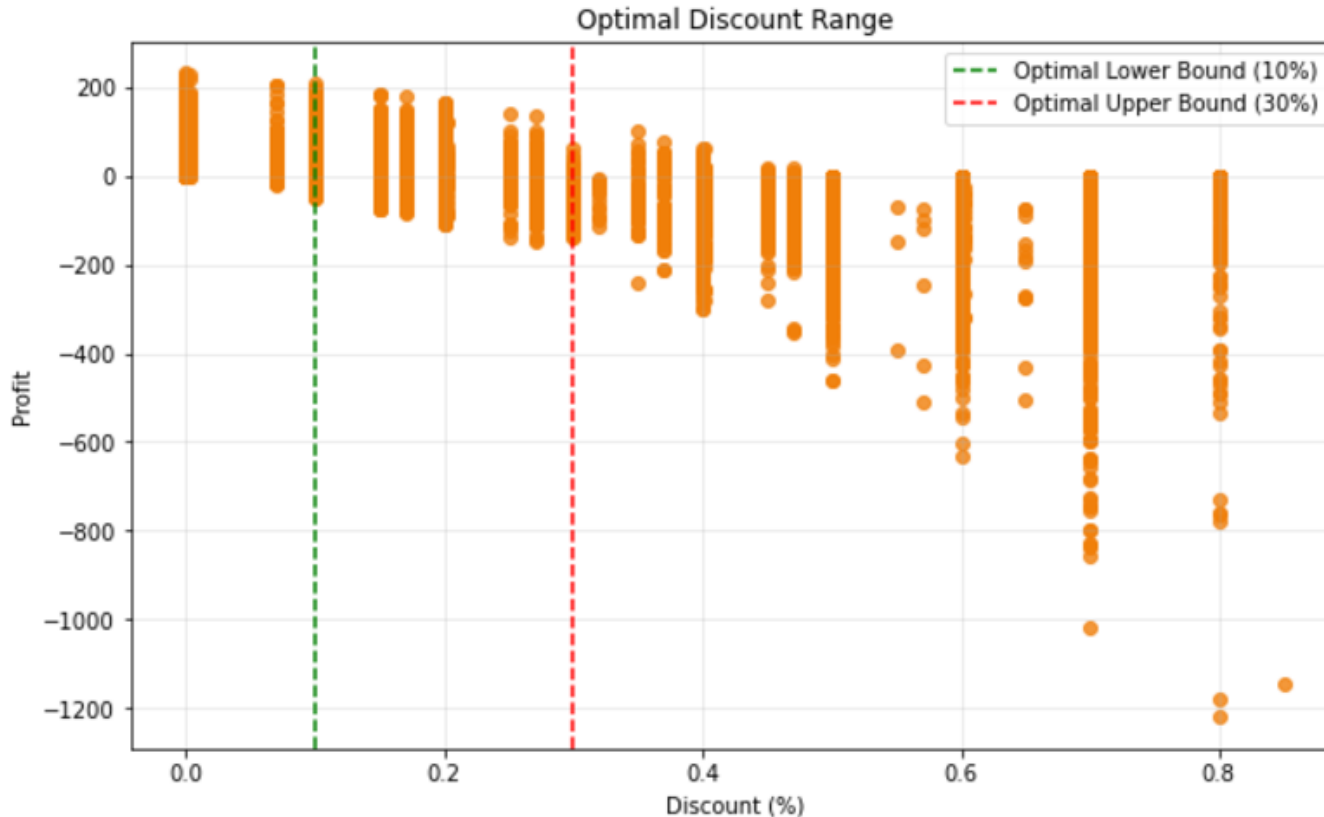
### Insights

- Discounts do not strongly influence sales growth. Other factors, such as product demand, customer loyalty, and market conditions, may have a larger impact.
- Higher discounts lead to a measurable decline in profitability, indicating the importance of balancing discount levels to avoid eroding profit margins.



# Impact of Discount Strategies

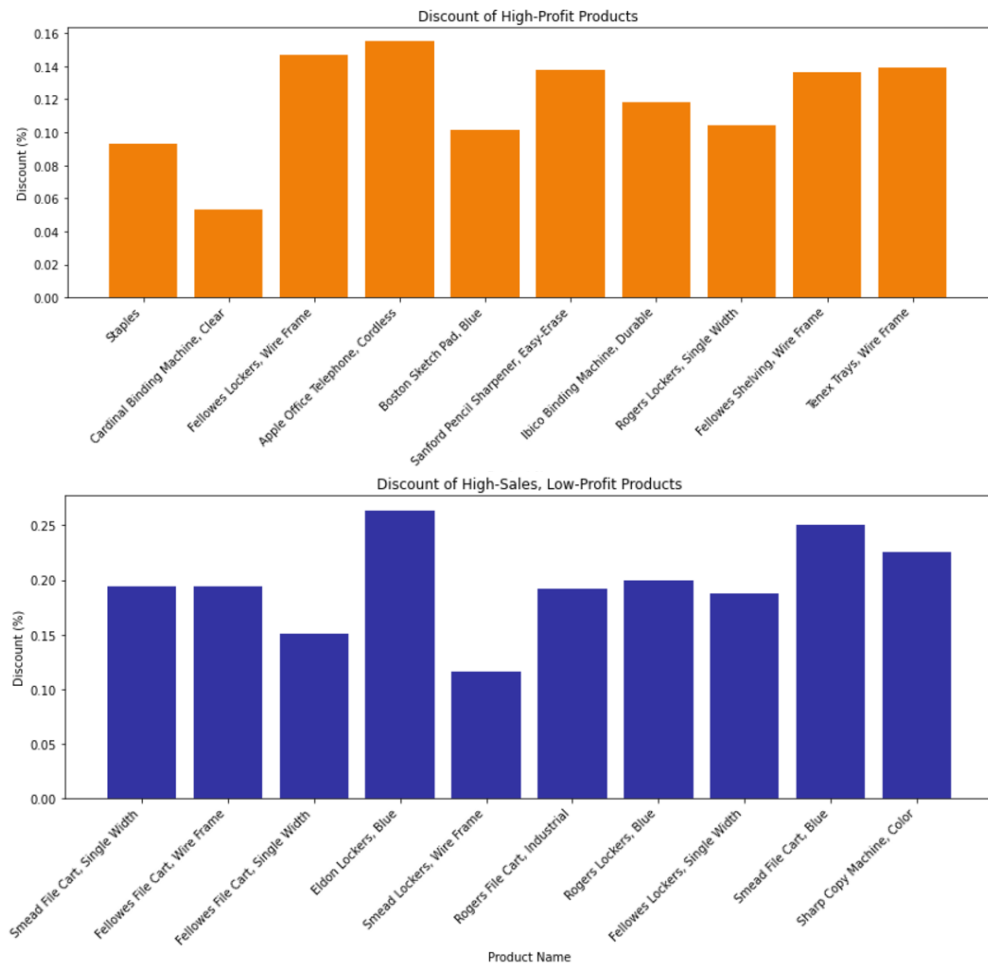
## Optimal Discount Range



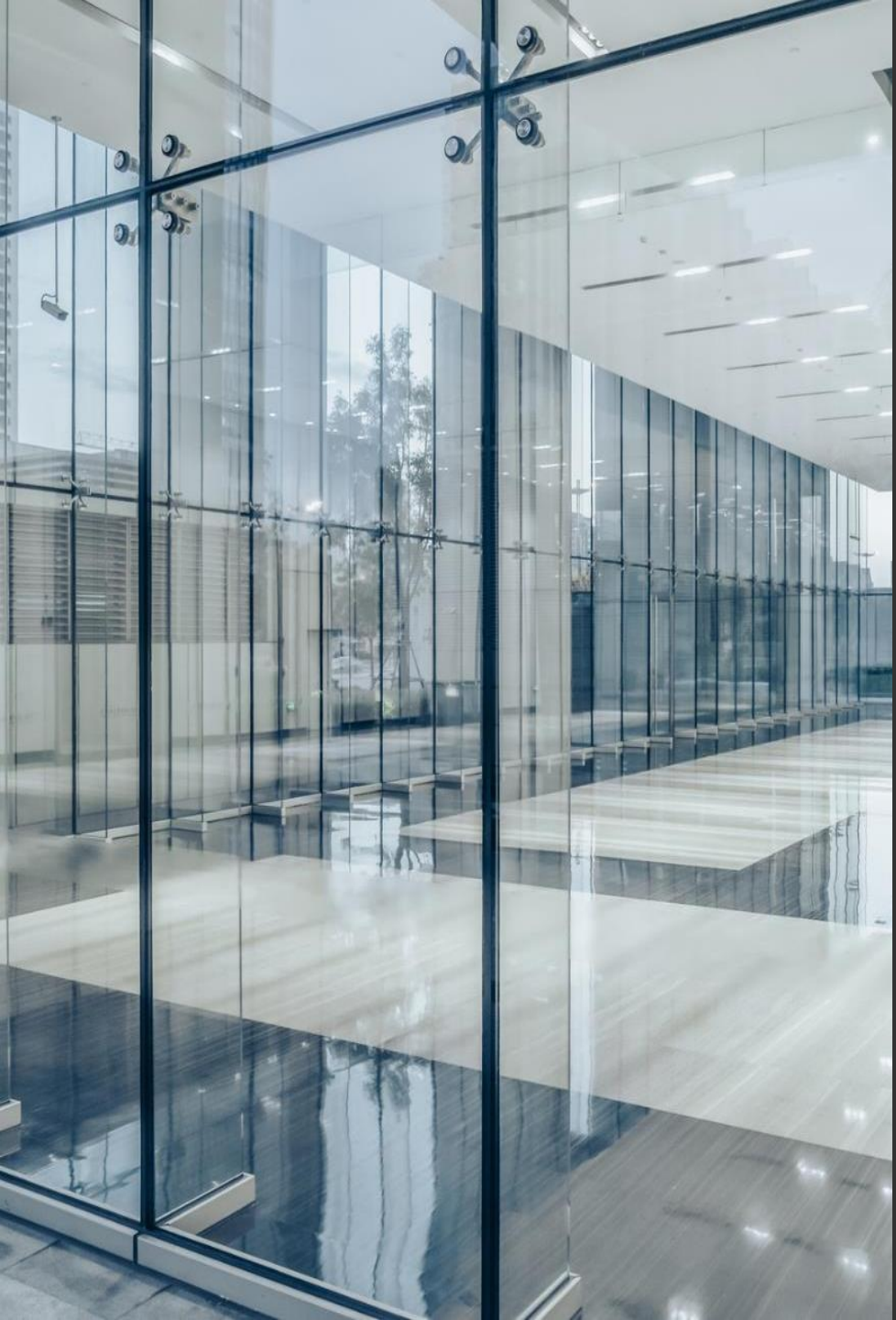
- Within the range 10-30%, most transactions appear to have positive or near-zero profits. This suggests that offering discounts within this range is generally profitable.
- Beyond the upper bound (30%) many transactions resulting in negative profit, that indicates that discounts above 30% tend to erode profitability significantly.
- Transactions with discounts below the lower bound (10%) show less variability in profit, this suggests that minimal discounts may preserve margins but might not drive higher sales volume.

# Impact of Discount Strategies

## Optimal Discount Range for High- / Low-Profit Products



- **High-profit products**
  - The optimal discount percentages for high-profit products are typically below 15%.
  - Offering excessive discounts on these products may not significantly improve their profitability and could potentially reduce the overall profit margins.
- **High-sales, low-profit products**
  - The optimal discount percentages are generally higher compared to high-profit products. They range from around 15% to 25%.
  - Increasing sales volume at these discounts can compensate for low profitability.
  - They are suitable candidates for bundling with more profitable items to improve overall profitability.



# Product Bundling Opportunities

# Product Bundling Opportunities

## Bundling vs Non-Bundling Profitability Comparison

High-sales low-profit products	Bundling Transactions	Non-Bundling Transactions
Total Sales	27,929	1,407,525
Total Profit	1,163.13 \$	-46,288.83 \$
Profit Ratio	4.16%	-3.29%

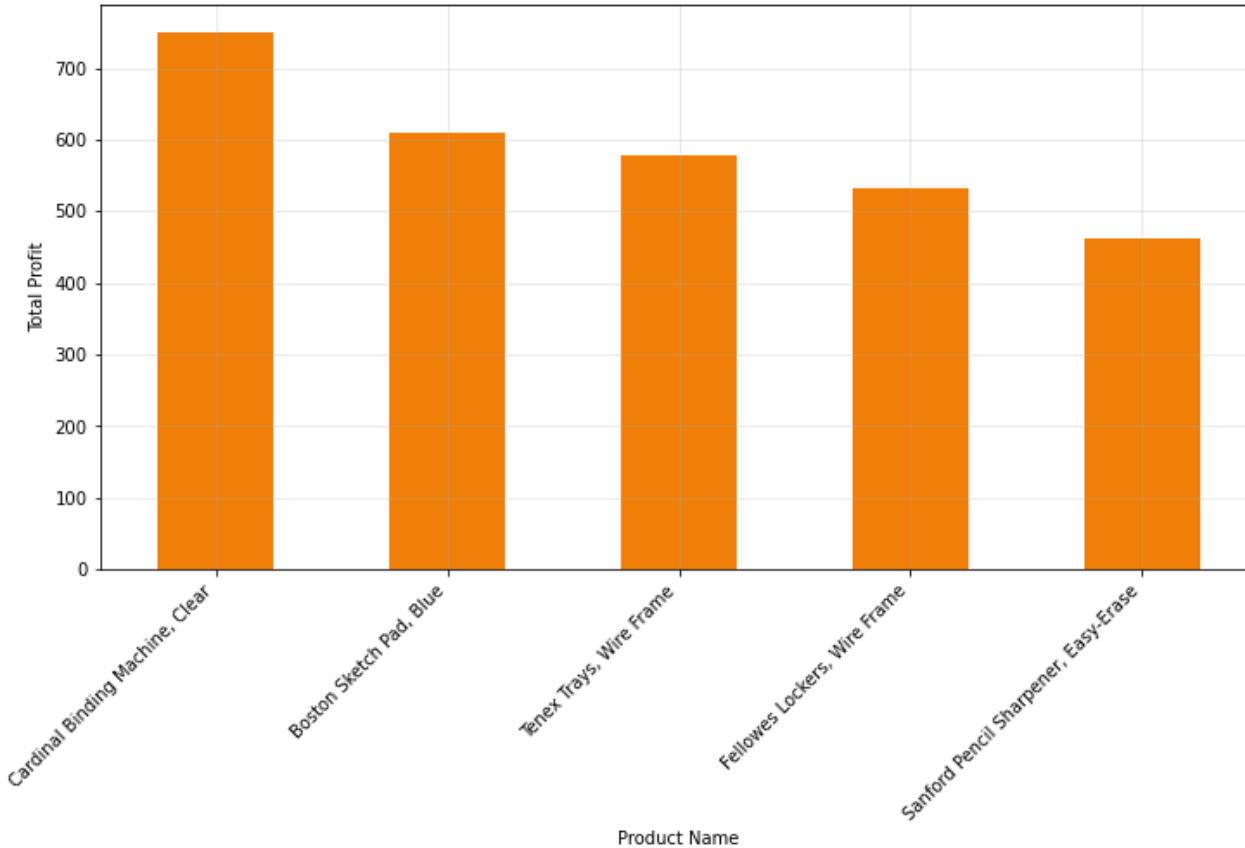
- Bundling high-sales low-profit products with high-profit products has a clear positive impact on profitability, turning a negative profit margin into a positive one.
- Although the sales volume is much lower in bundled transactions compared to non-bundling transactions, the profit margin improves significantly. This indicates that bundling strategically boosts profitability for high-sales, low-profit products.



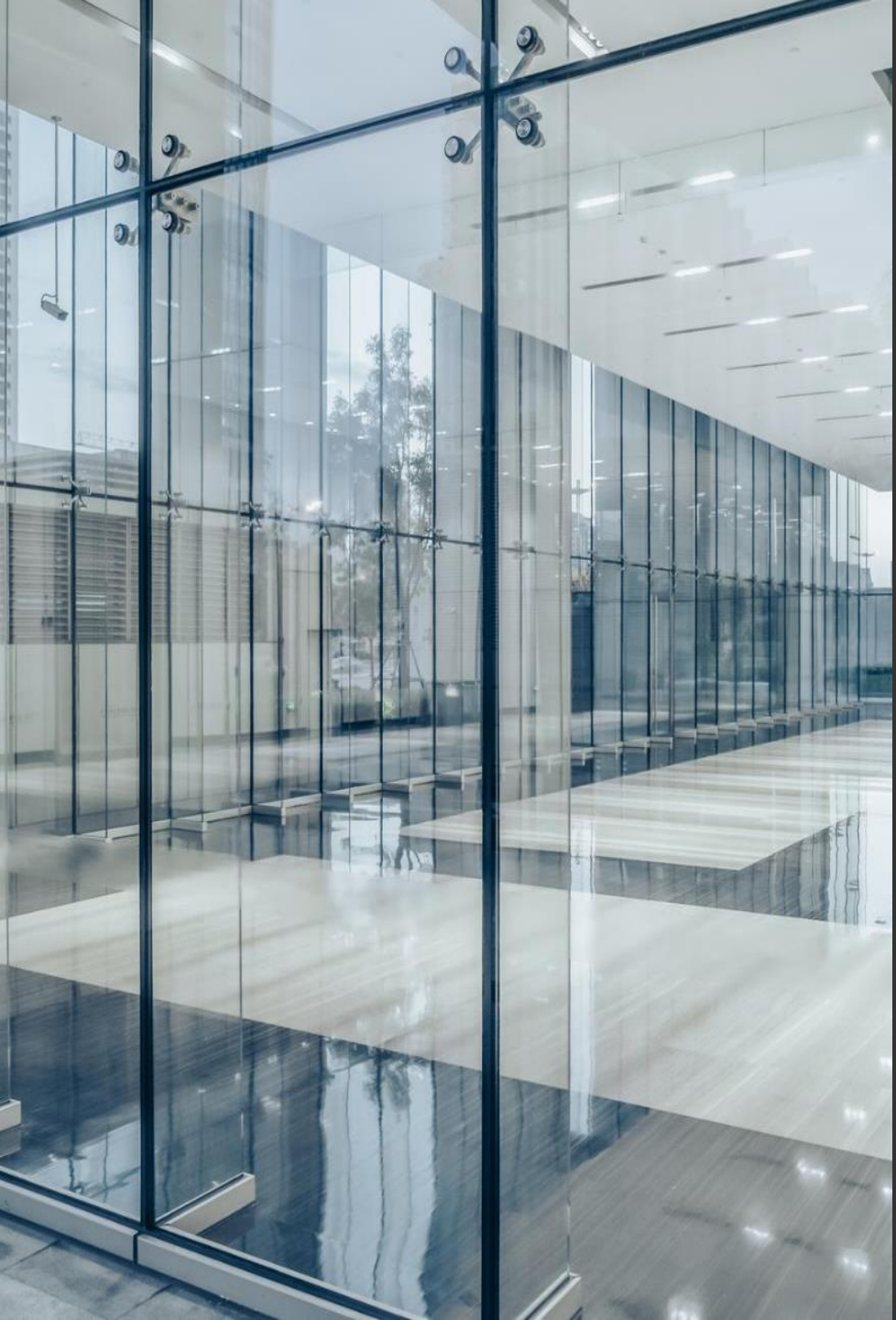
# Product Bundling Opportunities

## Top High-Profit Product Contributors

Top 5 High-Profit Product Contributions



- **Top 3 contributors in boosting bundled transactions:**
  1. Cardinal Binding Machine, Clear: contributes the most to total profit with \$750.60.
  2. Boston Sketch Pad, Blue: generates a profit of \$609.66.
  3. Tenex Trays, Wire Frame: profit contribution of \$578.046.
- **Bundling strategy:**
  - These products exhibit a high total profit and are thus key drivers of profitability when bundled.
  - Including these products in bundles with low-profit, high-sales products can potentially improve the overall profitability of the transactions.



# AB Testing

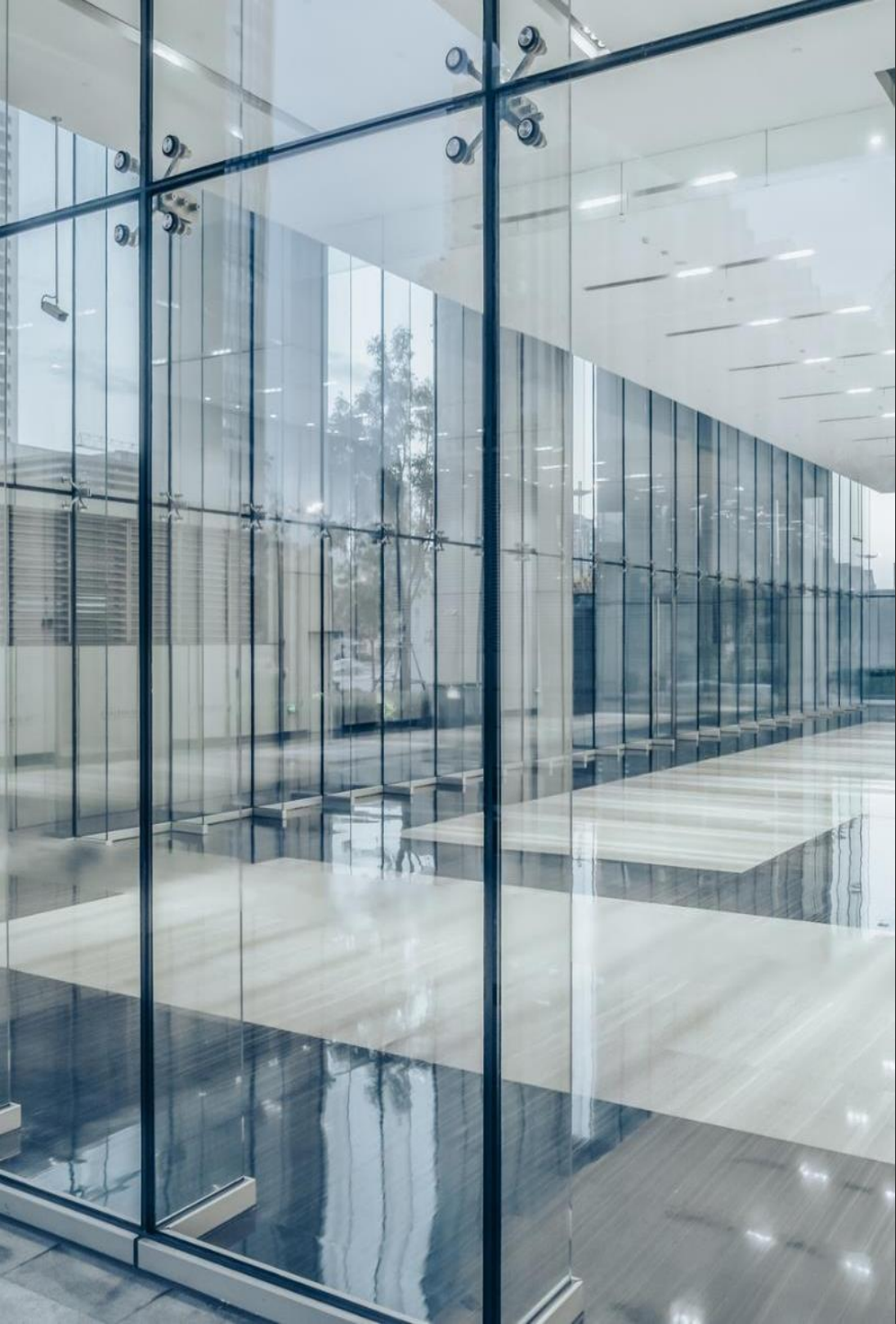
# AB Testing

## Discount and Product Bundling Strategies Testing

	Test 1. Discount Strategy	Test 2. Product Bundling Strategies
Objective	Validate the optimal discount range (10%-30%) for maximizing profitability	Test the effectiveness of bundling high-sales low-profit products with high-profit products
Groups	<ul style="list-style-type: none"><li>- Group A: Transactions with 10%-20% discounts.</li><li>- Group B: Transactions with 20%-30% discounts.</li></ul>	<ul style="list-style-type: none"><li>- Group A: Customers offered standalone high-sales low-profit products.</li><li>- Group B: Customers offered bundles (e.g., high-sales low-profit products + high-profit product).</li></ul>
Metrics	Average profit per transaction	<ul style="list-style-type: none"><li>- Bundle profitability</li><li>- Sales volume of bundled products</li><li>- Profit margin per transaction</li></ul>
User Segmentation	Geographical segmentation	By customer segment (e.g., Consumer, Corporate, Home Office)
Experiment Duration	Minimum: 2-4 weeks (to account for purchasing cycle variations)	Minimum: 4-6 weeks (to capture sufficient transactions across segments).
Statistical Significance	<ul style="list-style-type: none"><li>- Threshold (<math>\alpha</math>): 0.05</li><li>- Confidence interval: 95%.</li></ul>	<ul style="list-style-type: none"><li>- Threshold (<math>\alpha</math>): 0.05</li><li>- Confidence interval: 95%.</li></ul>
Sample Size	64	17686

### Testing instructions

• **Set up:** Randomly assign users to the control or test group | Ensure equal representation across segments to avoid bias • **Run:** Deploy the test and ensure consistent experiences for users in each group • **Monitor and Collect Data:** Continuously track the KPI for both groups • **Analyze Results:** Compare the performance of the two groups using statistical methods. using statistical methods.

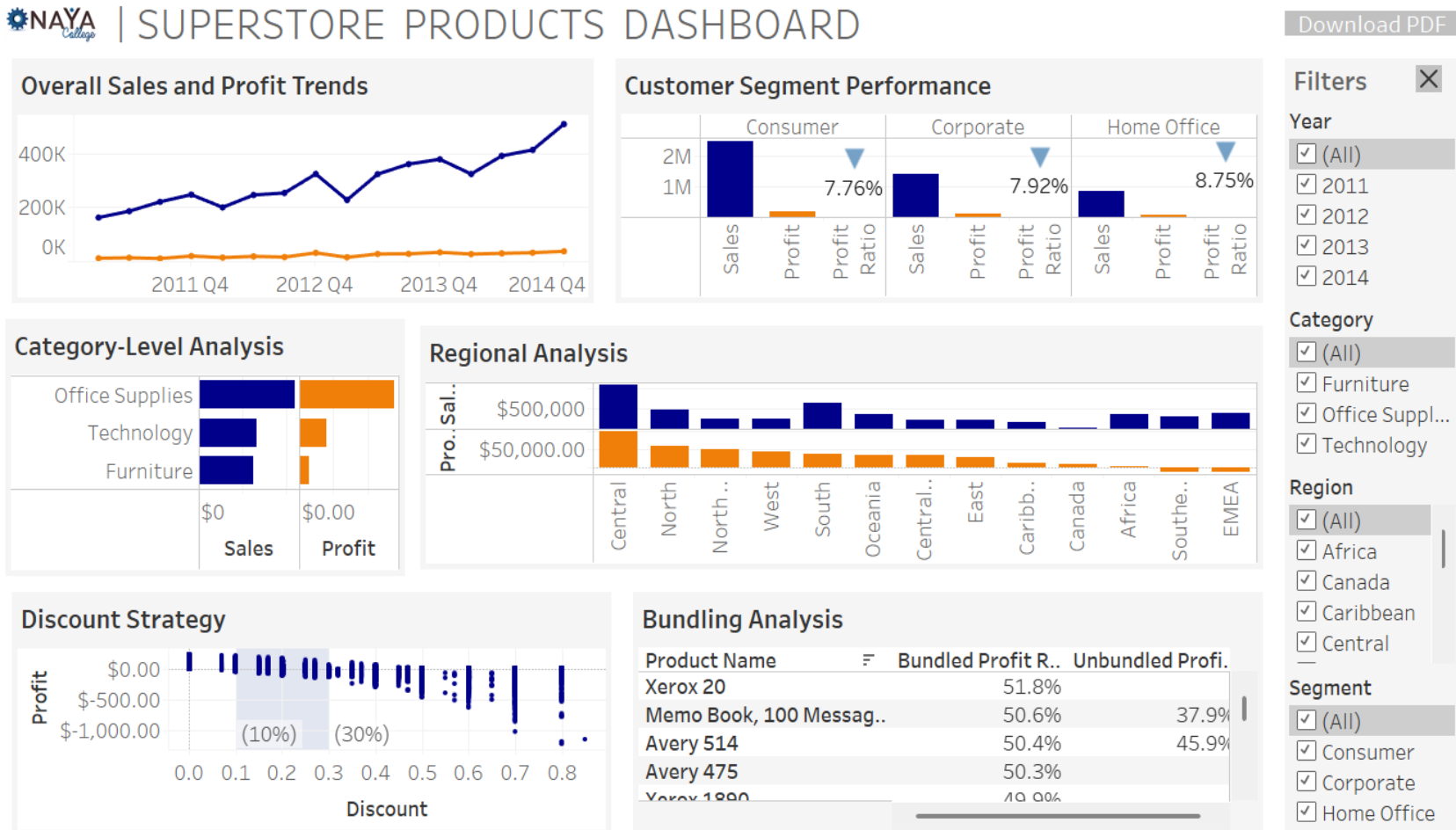


# Superstore Products Dashboard



# Dashboard

## Superstore Products



- ## Key features:
- **Sales and Profit Trends:**  
Track overall growth and identify focus areas.
  - **Category & Segment Analysis:** Assess profitability by product categories and customer segments.
  - **Regional Insights:**  
Highlight high and low-performing regions.
  - **Discount Strategy:**  
Evaluate discount impact and optimal ranges.
  - **Bundling Analysis:**  
Compare bundled vs. unbundled transactions to optimize profitability.
  - **Filters and drill-downs:**  
enhance exploration for strategic insights.



# Conclusions

- ▶ **Category-Level Focus**
  - Expand inventory and marketing for high-profit categories like Office Supplies.
  - Analyze and address inefficiencies in Furniture to improve its profit margins.
- ▶ **Regional Optimization**
  - Invest in marketing and operational improvements in high-potential regions like Central.
  - Reassess strategies for low-profit regions (e.g., Canada, Southeast Asia) to minimize losses.
- ▶ **Discount Strategy**
  - Invest Focus on moderate discounts (10%-30%) for high-sales, low-profit products to balance sales growth and profitability.
  - Limit or eliminate discounts for high-demand products with strong profit margins.
- ▶ **Product Bundling**
  - Prioritize bundling low-profit, high-sales products with the top 3 high-profit products ("Cardinal Binding Machine, Clear", "Boston Sketch Pad, Blue", and "Tenex Trays, Wire Frame").

# THANK YOU



Final Project: Superstore Products Data Analysis

## CONTACT

 Rina Irene Rafalski  
@ [rinaraf@gmail.com](mailto:rinaraf@gmail.com)