Retail Analysis and Inventory Management

Exploratory Data Analysis

Problem Statement

The challenges of optimizing operations in the retail industry have become increasingly complex as businesses strive to meet consumer demands and maintain profitability. Inaccurate sales forecasts can lead to issues such as overstocking, understocking, and missed revenue opportunities. These challenges highlight the need for effective tools and strategies to predict sales trends accurately. The Store Sales Forecasting Dataset provides an opportunity to address these challenges by analyzing historical sales data from various stores. By exploring this data, we can uncover patterns, seasonal trends, and other factors that influence sales performance. These insights can then be utilized to develop strategies for improving inventory management, refining pricing tactics, and allocating resources more efficiently, ultimately enhancing the company's operational success.

Business Impact

By analyzing historical sales and product performance data, the project seeks to provide actionable insights that:
- Reduce overstocking and understocking issues.
- Improve procurement decision-making by forecasting demand.
- Identify high-performing products and categories for strategic prioritization.
- Increase customer satisfaction by ensuring product availability.
- Optimize discounting strategies to maximize profitability.

General Dataset Information

File Name: stores_sales_forecasting.csv

Description: Retail Sales, procurement and profit enhancement in furniture sector (Can use the same approach for other sectors)

Dataset Details: 2121 Rows & 21 Columns

Size: 486 KB

Source: Kaggle - Dataset Link

Target Features

The focus of the project is to analyze sales trends across multiple sub-categories within the furniture sector. The objective is to evaluate profit margins and understand how they are influenced by key factors such as seasonality trends, discount values, and customer segments. Additionally, a thorough study of seasonal sales trends for these sub-categories will enable better procurement planning, optimize marketing strategies, and enhance discounting periods to minimize losses and maximize profitability. The insights derived from this analysis will support strategic decision-making in inventory management, ensuring the right products are available at the right time while aligning with demand fluctuations.

Let's take a look on columns information and statistics:
- Total of 26 columns divided into:
- 8 Numerical Variable
- 12 Categorical Variable
- 2 date-time Columns
- 4 Text Columns

Descriptive Statistics of Numerical Columns

|	Sales	Quantity	Discount	Profit	Order Ship Difference	Unit Price	Profit Ratio	Unit Profit Ratio
Count	2121.000	2121.000	2121.000	2121.00	2121.000	2121.000	2121.000	2121.000
Mean	349.835	3.785	0.174	8.69	3.917	90.854	0.039	0.014
Std	503.061	2.251	0.182	136.02	1.755	98.912	0.344	0.149
min	1.892	1.000	0.000	-1862.31	0.000	1.164	-2.100	-1.567
25%	1.892	1.000	0.000	-1862.31	0.000	15.980	-2.100	-0.024
50%	182.220	3.000	0.200	7.77	4.000	60.784	0.111	0.031
75%	435.168	5.000	0.300	33.73	5.000	126.686	0.275	0.083
max	4416.174	14.000	0.700	1013.13	7.000	880.980	0.500	0.480

Sales Distribution
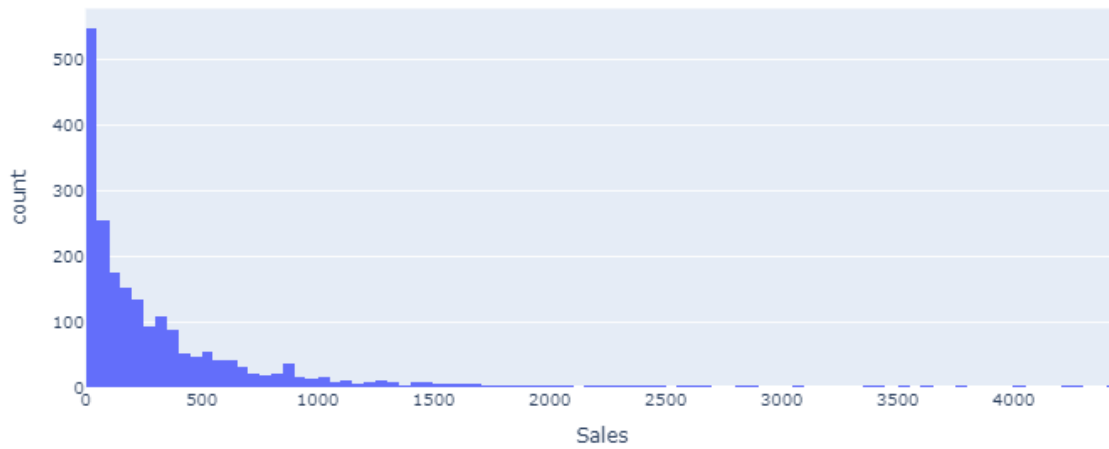
Unit Price Distribution

✓ Sales Distribution is right skewed showing that most sales orders values less than 500$, with 75% of values less than 435$ as shown in the table above. We can explain that later while we see that high percentage of orders are from individual consumers who they buy furnatures for their own personal use which should be limited. The distribution of 'Sales' and 'Unit Price' follow a similar pattern, indicating that lower-priced products are sold more frequently. And since 75% of unit prices values are below 126.7$, pricing and marketing strategies should focus on optimizing profit margins in this range to capture the majority of demand.

Quantity Distribution
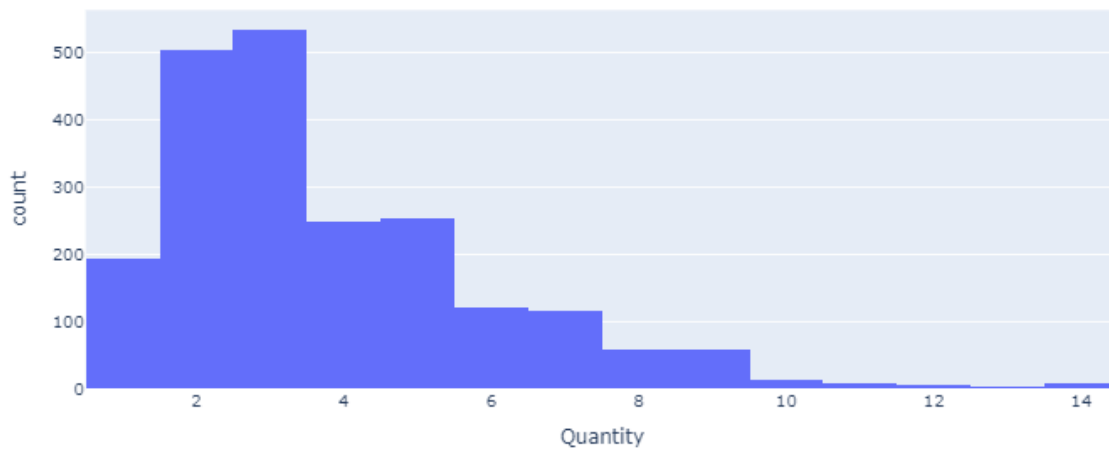
✓ Quantity distribution also is right skewed, and most orders' quantities contain less than 6 items with a max of 14 per order. (Noted that some orders contain more than one sub-category so could be represented in more than one row, therefore the maximum number of items per order could be more than 14 items but will be very small number of orders).

Discount Distribution
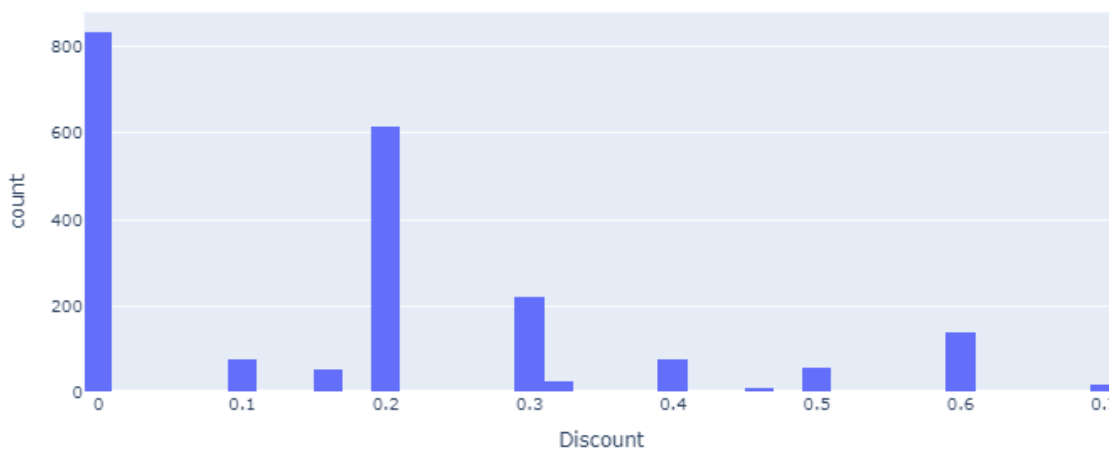
✓ From the Discount Distribution plot and statistics, we can see that most orders sold with no or just 20% discount with about 1400 over 2121 row.
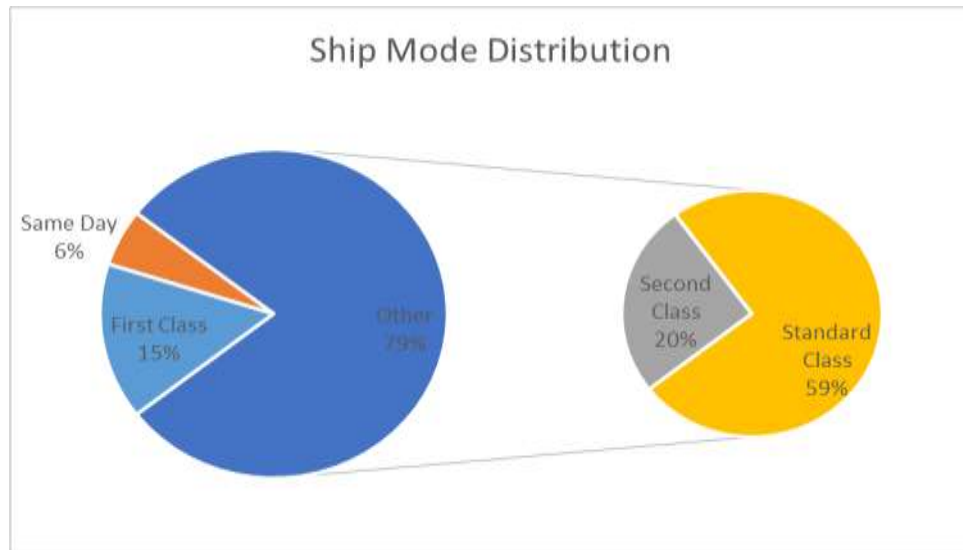
Profit Distribution
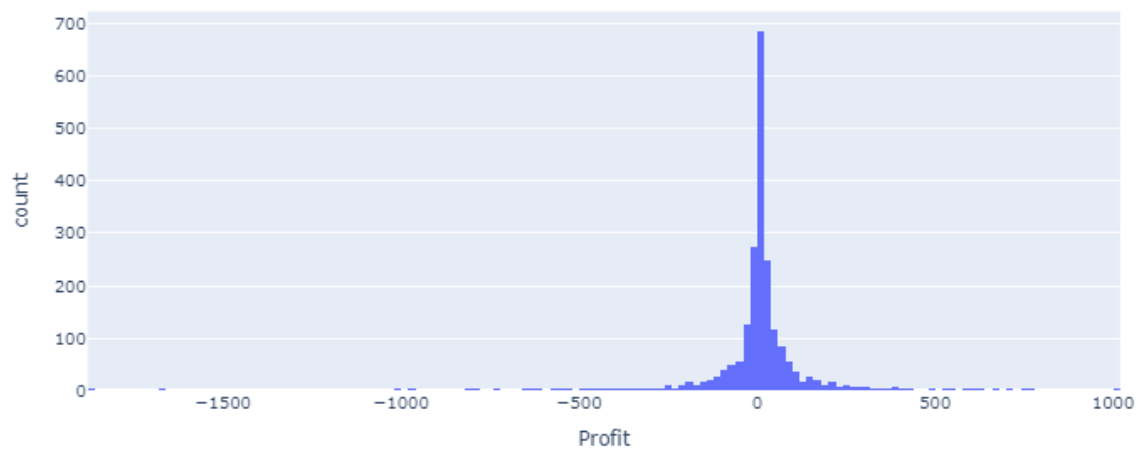
✓ Profit distribution appears to be nearly normal but slightly right-skewed around a median value of $7.78, with a mean of $8.70. Some extreme values are present, which should be investigated further to determine their validity.

Order Ship Difference Distribution

Ship Mode Distribution

- ✓ The left-skewed order ship difference distribution indicates that most orders take longer times to ship. This trend suggests that customers, particularly those in the standard or second-class shipping segments, prioritize cost savings over faster delivery. It may be beneficial to study whether offering lower prices for slower shipping options influences customer preferences and impacts profitability. This can be linked to customer segments, where the majority of customers prefer standard or second-class shipping methods, leading to extended delivery times and potential impacts on customer satisfaction.

Profit Ratio Distribution
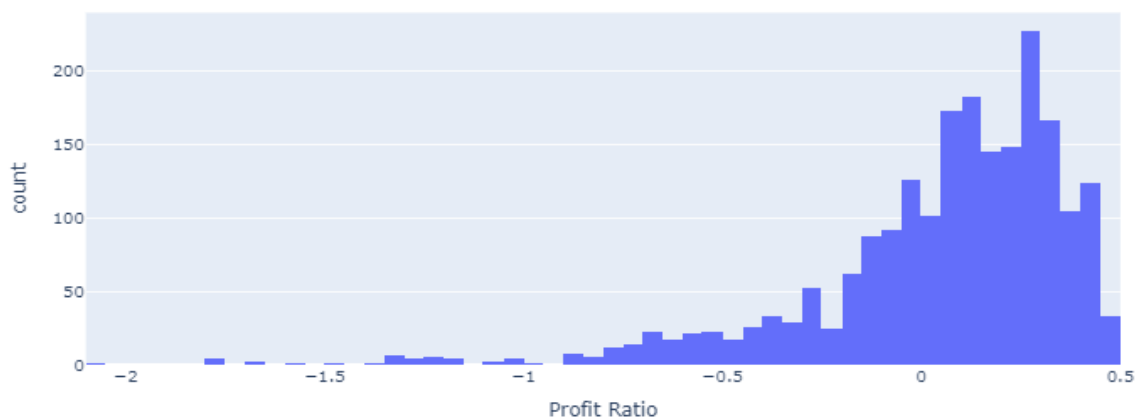
- ✓ The profit ratio distribution is left-skewed with a median of 0.111, revealing that some orders have loss values exceeding the product's cost (profit ratios below -1). This suggests potential issues such as defective products that required costly replacements, excessive discounting, or operational inefficiencies. Identifying the root causes of these losses is crucial for improving pricing strategies, cost management, and overall profitability.

Quarterly Sales Performance Over Time

Quarterly Sales Trends by Product Sub-Category
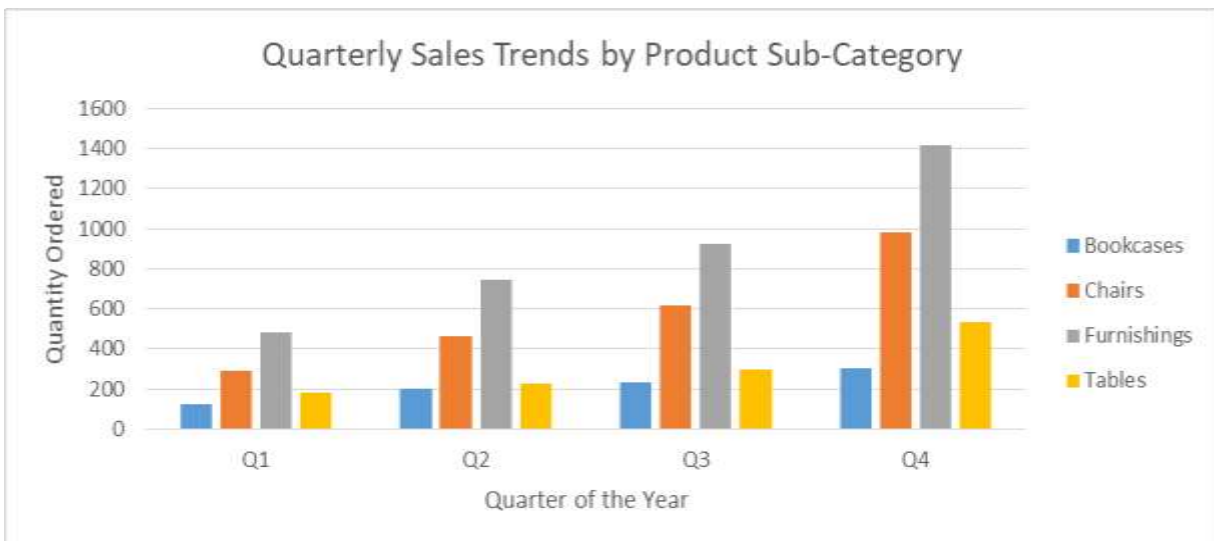
1. Seasonal Demand Trends:

The first chart shows an overall increasing trend in sales over time, with significant spikes in Q4 each year. This suggests that demand peaks towards the end of the year, likely due to seasonal factors such as holidays or year-end purchases. Procurement should focus on ramping up inventory levels ahead of Q4 to meet anticipated demand and avoid stockouts.

2. Product-Specific Demand Fluctuations:

The second chart indicates that different product sub-categories exhibit varying sales patterns across quarters. Furnishings consistently have the highest demand across all quarters, especially in Q4, while Chairs and Tables also see significant growth. Bookcases, though lower in volume, maintain steady

demand. Procurement strategies should prioritize high-demand products like Furnishings and Chairs in Q4 while maintaining steady stock levels for Bookcases throughout the year.

3. Inventory Optimization:

Since Q1 and Q2 show lower sales volumes across all product categories, procurement should adopt a just-in-time approach during these periods to minimize holding costs and excess inventory. Adjusting procurement schedules to align with these seasonal fluctuations can help balance stock levels and reduce storage costs.

4. Demand Forecasting:

The data highlights the importance of forecasting based on historical trends. Leveraging sales data to predict future demand and adjust procurement plans accordingly will help prevent overstocking during low-demand periods and shortages during peak seasons.

5. Supplier Coordination:

Given the cyclic nature of demand, procurement teams should work closely with suppliers to establish flexible contracts that allow for scaling orders up or down based on quarterly demand fluctuations. This flexibility will help ensure timely deliveries without excess storage costs.

| Count of Profit Category by Region				
Region	Loss	Low Profit	Medium Profit	Grand Total
Central		324	48	109	481
East		193	174	234	601
South		65	107	160	332
West		165	275	267	707
Grand Total		**747**	**604**	**770**	**2121**
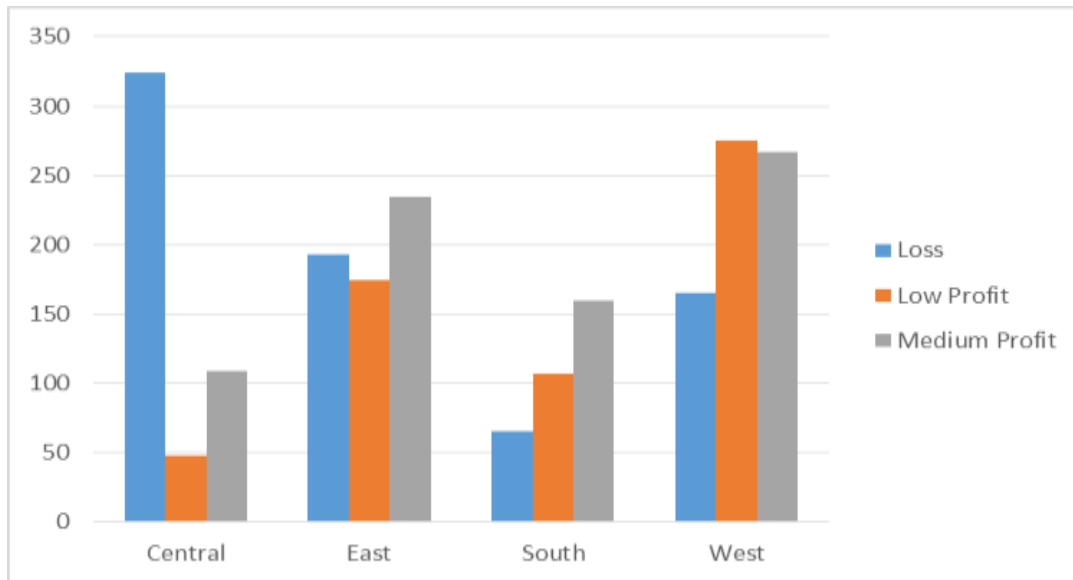
✓ The table and chart illustrate the distribution of profit categories across the different regions. The Central region has the highest number of loss-making orders (324), while the West region has the highest overall sales volume (707), with a more balanced distribution among loss, low, and medium profit categories. The South region, despite having the lowest total orders (332), shows a relatively higher proportion of medium-profit orders compared to other regions. To enhance sales and profitability, focusing on reducing losses in the Central region through better quality control, improved pricing strategies, and targeted promotional efforts can help optimize performance. Additionally, analyzing customer preferences and demand patterns in high-performing regions like the West can provide insights to replicate success in other areas.

| Product Sub-Category	Quantity Sold	Average of Discount	Average of Unit Profit Ratio	Average of Profit Ratio
Bookcases	868	0.211140351	-0.037221334	-0.126640068
Chairs	2356	0.170178282	0.015897104	0.04389963
Furnishings	3563	0.138349007	0.050341975	0.137066353
Tables	1241	0.261285266	-0.06019682	-0.147726526
Grand Total	**8028**	**0.173922678**	**0.01428408**	**0.038783533**
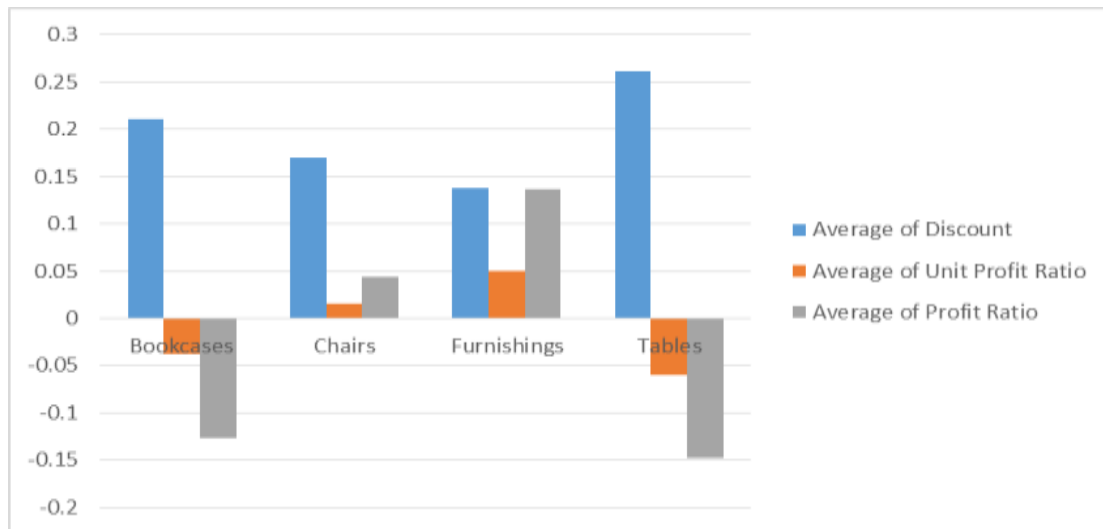
✓ The table and chart provide insights into the relationship between product sub-categories, discounts, and profit ratios. Among the product sub-categories, Tables have the highest average discount (0.2618), yet they exhibit the lowest profit ratios, both at the unit level (-0.0602) and overall (-0.1748), indicating significant losses. Bookcases also show negative profitability despite receiving substantial discounts. On the other hand, Furnishings have the highest overall profit ratio (0.1370) with the lowest discount rate (0.1383), suggesting that lower discounts contribute positively to profitability. To enhance profitability, it is recommended to re-evaluate discount strategies, particularly for Tables and Bookcases, and focus on optimizing pricing and cost control while leveraging the success of Furnishings as a benchmark for other categories.
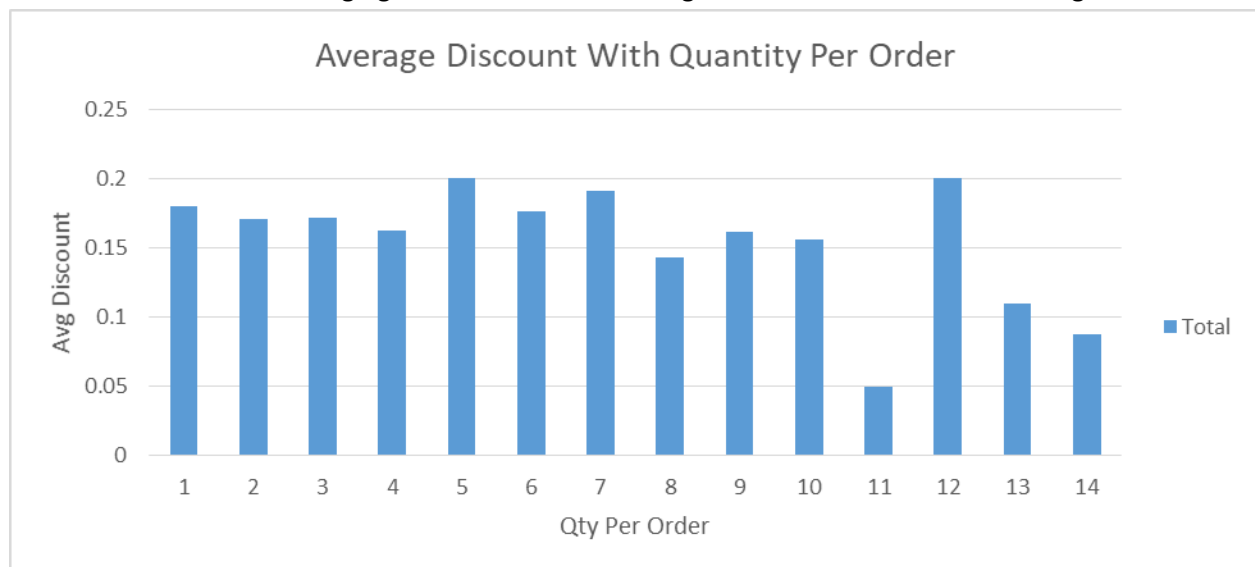
✓ The chart displays the average discount applied across different order quantities, showing fluctuations without a clear pattern or trend. This indicates that there is no significant relationship between the quantity per order and the average discount, suggesting that discounts are applied inconsistently rather than following a structured volume-based pricing strategy.

Customer Segment	Quantity Sold	Average of Discount	Average of Profit	Sum of Profit
Consumer	4166	0.176738544	6.281292543	6991.0786
Corporate	2495	0.174071207	11.74120093	7584.8158
Home Office	1367	0.165	10.70546519	3875.3784
Grand Total	**8028**	**0.173922678**	**8.69932711**	**18451.2728**
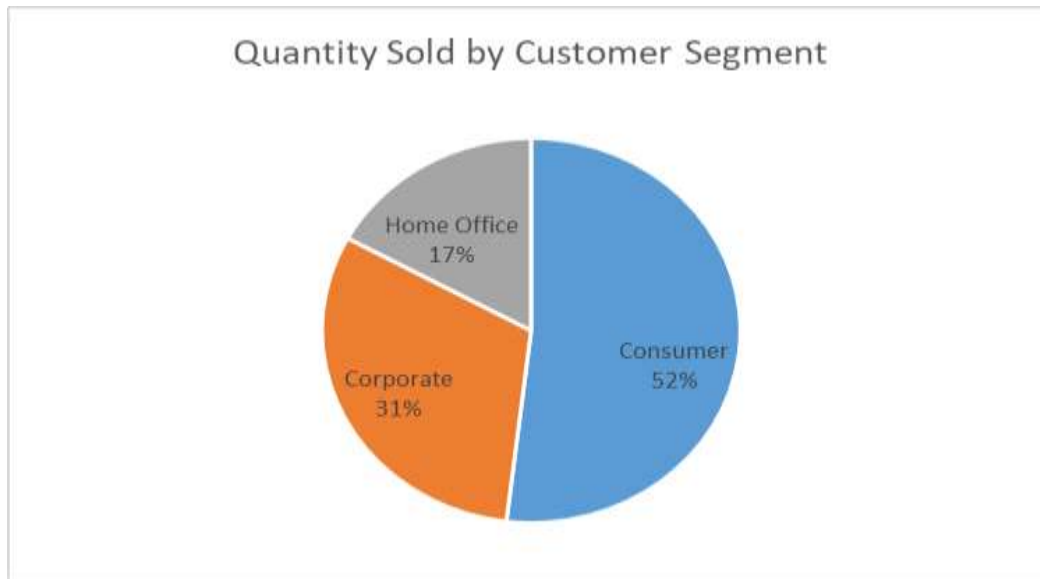
✓ The Consumer segment contributes the highest sales volume at 52%, followed by Corporate at 31% and Home Office at 17%, indicating that Consumers are the primary market driver. Efforts should focus on retaining and expanding the Consumer base while exploring strategies to boost sales in the Corporate and Home Office segments to achieve a more balanced distribution.

Discount Value by Customer Segment

Segment	Discount
Home Office	0.165
Corporate	0.174
Consumer	0.177

✓ Consumers receive the highest average discount (0.177), followed closely by Corporate (0.174) and Home Office (0.165). Although the Consumer segment generates the second-highest total profit, it has the lowest average profit per order, suggesting that high discount rates are driving higher sales volumes but at the expense of profitability. Reducing discounts might negatively impact sales, so a balanced approach is needed—such as targeted promotions or loyalty programs—to maintain volume while gradually improving margins.

Average and Total Profit by Customer Segment
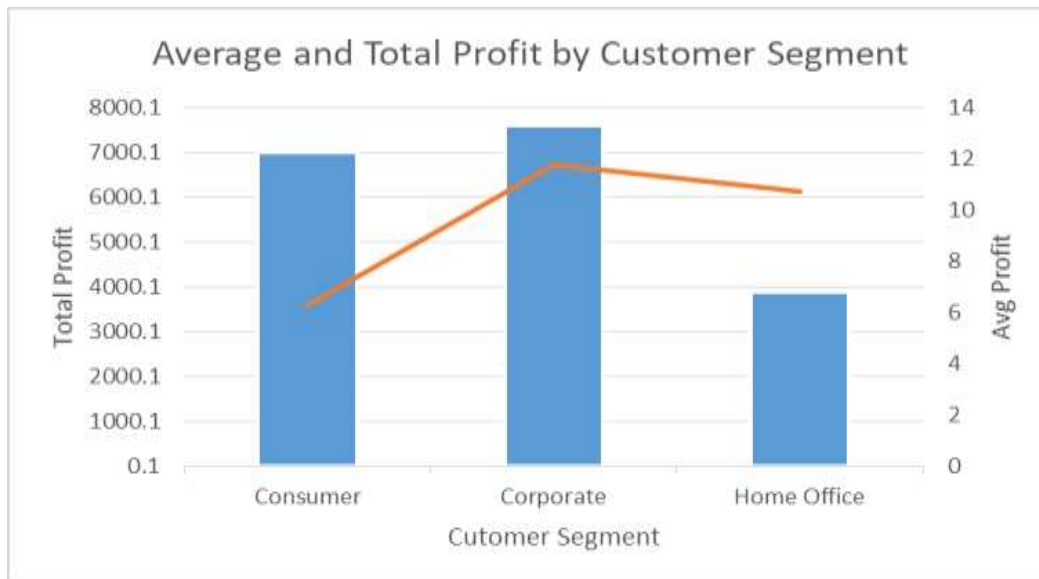
✓ The Corporate segment, despite having fewer sales than the Consumer segment, generates the highest total profit (7,584) and the highest average profit per order, whereas the Home Office segment lags in both metrics. Efforts should focus on increasing Corporate sales, given their higher profitability, and exploring ways to improve profitability in the Home Office segment through tailored marketing strategies and cost optimization.