

Sales and Profit Analysis for Superstore Data

Overview

In this project, I analyzed sales and profitability data from the Superstore dataset, using Python for data cleaning and Power BI for visualization. The goal was to uncover insights that support better decision-making in product management and sales strategies.

About the Dataset

The dataset contains 9994 records of customer orders from 2018 to 2021. Each row represents an order line, with details such as order date, customer location, product category, sales, and profit.

Before the analysis, I cleaned the data using Python by checking for missing values, correcting data types, and creating new columns like **Profit Ratio**. I also created a **MonthYear** column to help with time-based analysis.

Key KPIs and Metrics

- **Total Sales:** \$2,297,200.86
- **Total Profit:** \$286,397.02
- **Average Profit Ratio:** 12.47%

Category Analysis: Furniture vs. Office Supplies

One interesting finding was that **Furniture** had **higher total sales** than **Office Supplies**, but generated **less total profit**.

This suggests that the **profit margin for Furniture is lower**, which may be due to:

- Higher cost of goods sold (COGS)
- More discounts or promotions applied
- Logistical or shipping costs related to bulky items

This insight could help the company **re-evaluate pricing strategies** or explore ways to **optimize operational costs** for the Furniture category.

One possible reason for this could be losses in some of the sub-categories under Furniture, such as Tables and Bookcases, which will be discussed in the next section.

Sub-Category Profitability

During the analysis, it was found that some **sub-categories consistently generated negative profit**, despite having moderate or even high sales.

For example:

- **Tables** had relatively high sales but resulted in a **net loss**.
- Other sub-categories like **Bookcases** or **Supplies** also showed negative profit in certain years.

This could be due to:

- Low pricing or **aggressive discounts**
- High shipping or manufacturing costs
- Inventory or supply chain inefficiencies

These findings highlight the need to **review pricing strategies, optimize cost structure**, or consider **repositioning or replacing** underperforming products.

Monthly Sales Trends

Another key insight from the analysis was the **seasonal variation in sales**. Across the years, months like **September, November, and December** consistently showed **higher sales**, while **January and February** recorded the **lowest**.

This recurring pattern suggests that the business experiences **stronger sales performance during the last quarter of the year**, possibly due to:

- Back-to-school season (September)
- Holiday and end-of-year shopping (December)
- Black Friday discounts and promotions in November

On the other hand, **Q1 shows a slowdown**, which might reflect post-holiday consumer behavior or budgeting constraints early in the year.

This trend is important for:

- **Inventory planning**
- **Marketing campaigns**
- **Sales forecasting**

Recommendations

Based on the analysis, here are a few suggestions that could help improve business outcomes:

- **Review pricing and cost structure** in low-profit categories like Furniture, especially sub-categories with losses.
- **Investigate the impact of heavy discounts**, particularly during sales spikes like Black Friday, to ensure profitability is maintained.
- **Optimize marketing and inventory planning** based on seasonal sales patterns to take full advantage of high-demand periods.
- For future analysis, consider exploring:
 - Customer segmentation
 - Regional performance trends
 - Discount and promotion effectiveness

Conclusion

This project provided hands-on experience in combining Python and Power BI to transform raw data into actionable insights. It helped me identify trends, inefficiencies, and opportunities that could support better business decisions.