

DATA SCIENCE ASSIGNMENT TASK 1

TIMING

Jan 25, 2025 1:00 PM GMT+5:30 to

Jan 27, 2025 11:45 PM GMT+5:30

STATUS **COMPLETED**

OWNERS Vipra

TABLE OF CONTENTS

INTRODUCTION

EDA FINDINGS

BUSINESS INSIGHTS

CONCLUSION

→ INTRODUCTION

This report presents key insights derived from an exploratory data analysis (EDA) of the eCommerce dataset comprising three files: Customers.csv, Products.csv, and Transactions.csv. The analysis focuses on standard as well as customised EDA techniques and their findings. Insights include customer distribution, sales performance by product category, and transaction trends over time.

RESOURCES

PYTHON NOTEBOOK

📄 File

EDA FINDINGS

→ TRANSACTION ANALYSIS

The task's main focus. The key insights from the exploratory data analysis output.

INSIGHTS

- The dataset (after merging all three csv files) contains 1000 unique records.
- The most frequent Customer ID is C0109, appearing 11 times.
- Product ID P059 - Soundwave Jeans, appears 19 times, making it the most frequently purchased product.
- There are 66 unique products, divided across 4 major categories.

METRICS

- The average transaction quantity is 2.53 units.
- The mean total expenditure is approximately 689.99.
- The average price per unit is 272.55.
- The standard deviation for product price is 140.73, indicating significant price variation

BUSINESS INSIGHTS

→ CUSTOMER DISTRIBUTION BY REGION

- South America has the highest number of customers, indicating a strong market presence in this region.
- Takeaway: Businesses should consider introducing products and tailoring advertising strategies specific to South America to capitalize on this customer base.

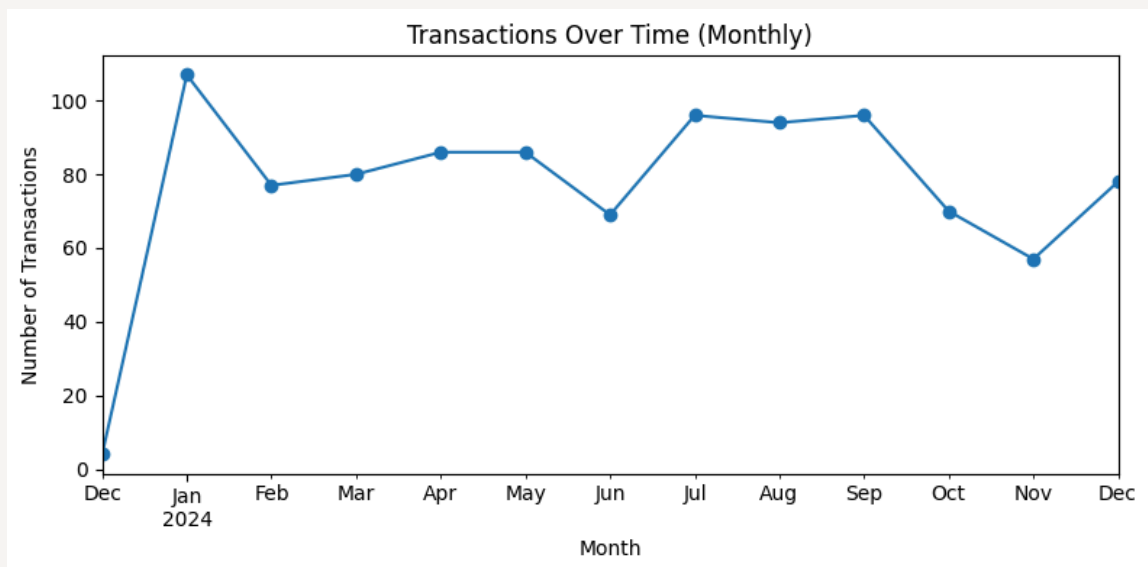
→ TOTAL SALES BY PRODUCT CATEGORY

- The Books category leads in total sales, while Home Decor is the lowest-performing category.

- Takeaway: Increasing marketing efforts for Books could sustain growth, while Home Decor may require re-evaluation of pricing or product offerings.

→ TRANSACTION TRENDS OVER TIME

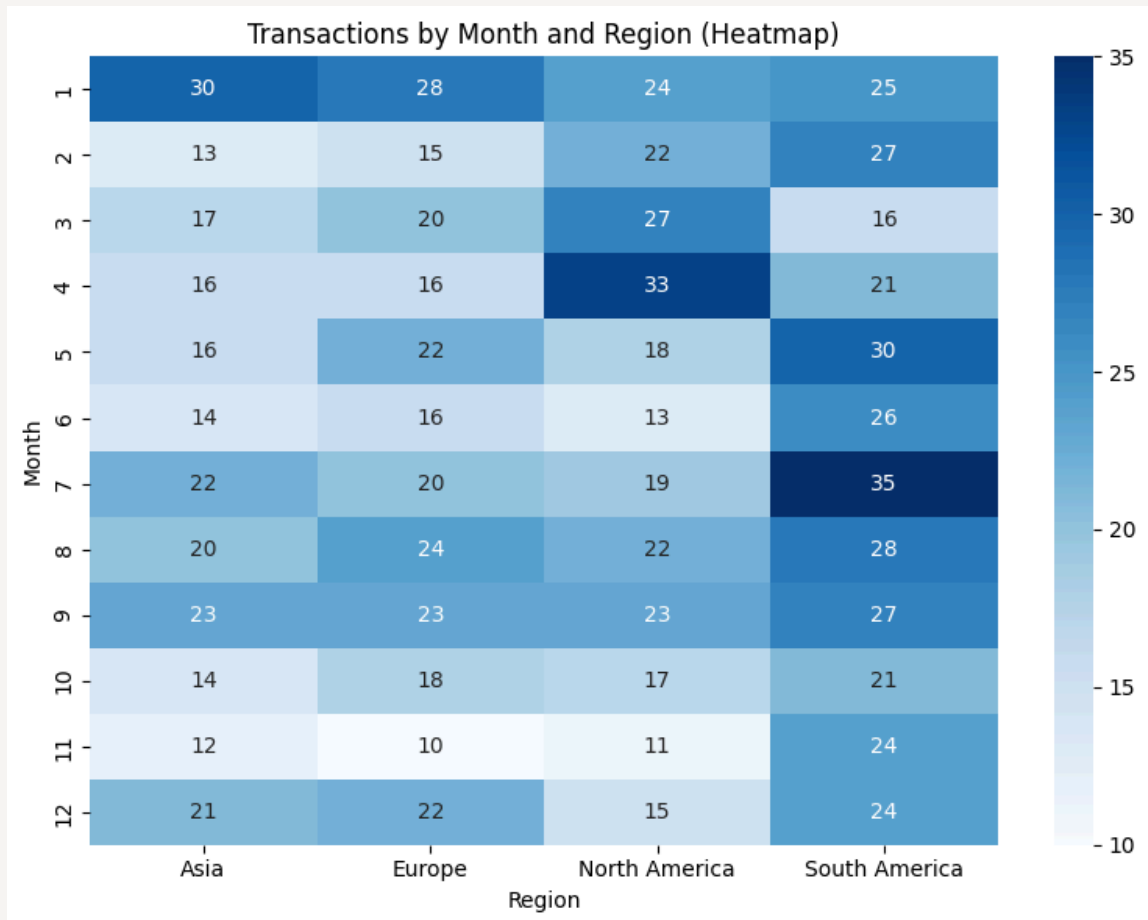
- January consistently shows the highest number of transactions, suggesting a peak sales period. (Due to New Year and various other festivals across the world in this month).
- Takeaway: Businesses should leverage January for promotional activities to maximize revenue.



→ REGION WISE HIGHEST TRN MONTHS

- Stats :

- Asia and Europe transaction counts peak in January.
- North America peaks in April.
- South America peaks in July.
- Takeaway: Regional marketing campaigns should be timed according to these peak months to optimize sales efforts.



→ REGION WISE TOP CATEGORIES

- "Books" dominate transactions in South America, while "Clothing" leads in Europe and Asia. North America shows balanced interest across categories.

- Takeaway: This suggests tailoring inventory and marketing strategies by region, such as promoting books in South America and clothing in Europe and Asia, to align with regional preferences. Also , businesses need to aggressively promote the lowest popular category products.



CONCLUSION

The dataset reveals significant regional and temporal variations in customer behavior and sales performance. By focusing on high-demand regions and peak months, businesses can enhance their marketing strategies and operational planning to drive growth and profitability. This report provides a concise overview of the key findings, offering actionable insights for eCommerce businesses to refine their strategies and improve market performance.