Project Cover Sheet

Full Name	Yogalakshmi Shanmuga Jothi
Email	yogalakshmishanmugajothi@gmail.com
Contact Number	+44 7587814559
Project Title (Example – Week1, Week2, Week3)	Week 1

Project Guidelines and Rules

1. Formatting and Submission

- Format: Use a readable font (e.g., Arial/Times New Roman), size 12, 1.5 line spacing.
- Title: Include Week and Title (Example - Week 1: Travel Ease Case Study.)
- File Format: Submit as PDF or Word file
- Page Limit: 4–5 pages, including the title and references.

2. Answer Requirements

- Word Count: Each answer should be within 100–150 words; Maximum 800–1,200 words.
- Clarity: Write concise, structured answers with key points.
- Tone: Use formal, professional language.

3. Content Rules

- Answer all questions thoroughly, referencing case study concepts.
- Use examples where possible (e.g., risk assessment techniques).
- Break complex answers into bullet points or lists.

4. Plagiarism Policy

- Submit original work; no copy-pasting.
- Cite external material in a consistent format (e.g., APA, MLA).

5. Evaluation Criteria

- Understanding: Clear grasp of business analysis principles.
- Application: Effective use of concepts like cost-benefit analysis and Agile/Waterfall.
- Clarity: Logical, well-structured responses.
- Creativity: Innovative problem-solving and examples.
- Completeness: Answer all questions within the word limit.

6. Deadlines and Late Submissions

- Deadline: Submit on time; trainees who fail to submit the project will miss the “Certificate of Excellence”

7. Additional Resources

- Refer to lecture notes and recommended readings.
- Contact the instructor or peers for clarifications before the deadline.

WEEK 1- DATA ANALYSIS FOR BUSINESS INSIGHTS

Introduction

An analysis of transactional sales data from online retailer ShopEase is presented in this report which aims to extract actionable business insights through data cleaning, exploration, and visualization. By uncovering customer behavior patterns, product sales trends, and time-based fluctuations, the findings will support ShopEase's decision-making for future growth.

Dataset Overview

The dataset includes 20 sales transactions over six months in 2024. A transaction ID, date, customer ID, product name, category, quantity, price per item, total transaction amount, payment method, and region are all included in each record. After initial inspection, the dataset was found to be mostly complete and well-structured. One missing value in the `Total_Amount` column was corrected by recalculating `Quantity × Price`. All date formats were standardized, and new columns for Month, Year, and Month_Name were created to support temporal analysis.

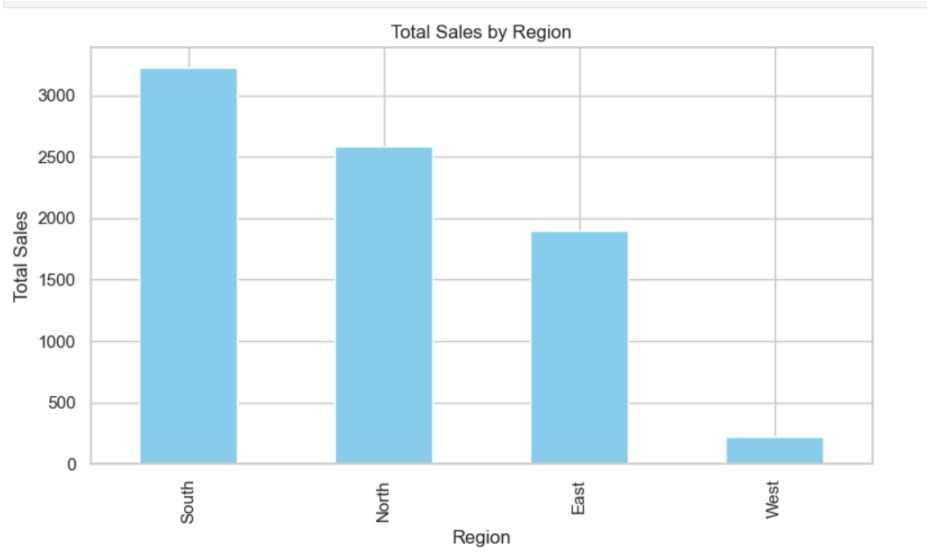
Exploratory Data Analysis (EDA) Summary

- Highest Sales Month: January recorded the highest total sales, mainly driven by high-value electronics purchases.
- March was the month with the lowest sales, suggesting that there is room for focused improvement.
- Top-Selling Category: Electronics consistently outperformed other categories, with laptops, smartphones, and tablets as leading products.
- Repeated customers is directly Proportional to Early sign of loyalty customers.
- Top Customers: The top 10 customers contributed a significant share of revenue, making them key targets for retention.
- Regional Trends: The North and South regions led in total sales, indicating strong market presence.
- Payment Preferences: Credit cards were preferred for electronics, while cash was more common for clothing.
- Correlations: Quantity × Price and Total_Amount was a strong relationship, indicating that the data was in line.

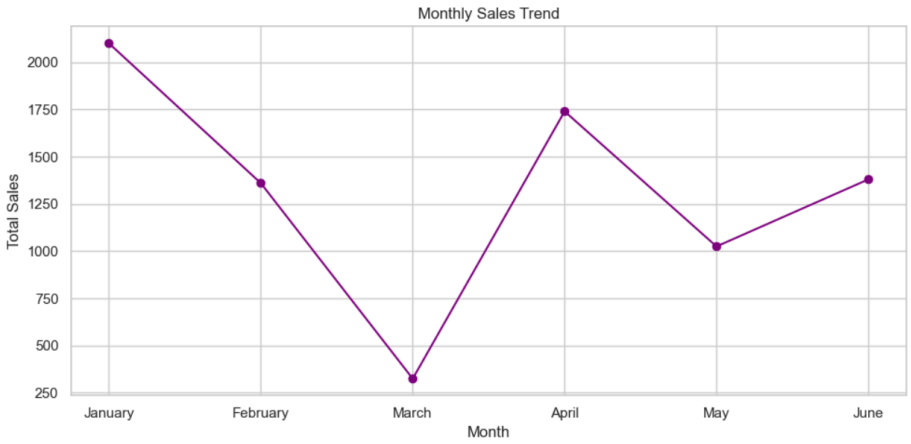
Data Visualisations

- Sales by Region (Bar Chart): The best-performing regions felt the North and South.

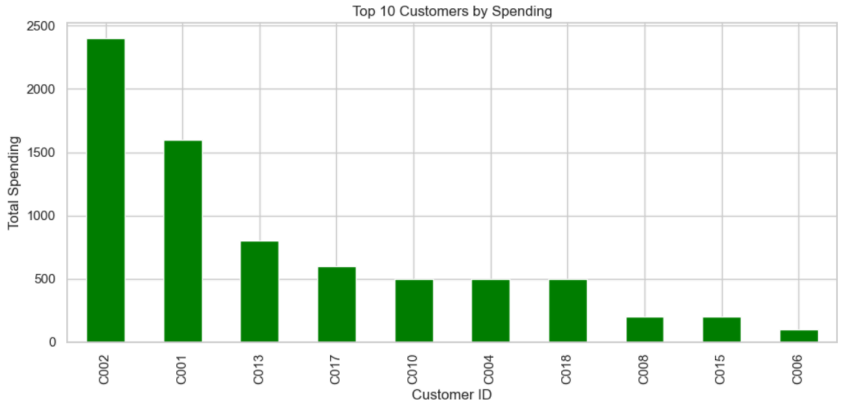
  .



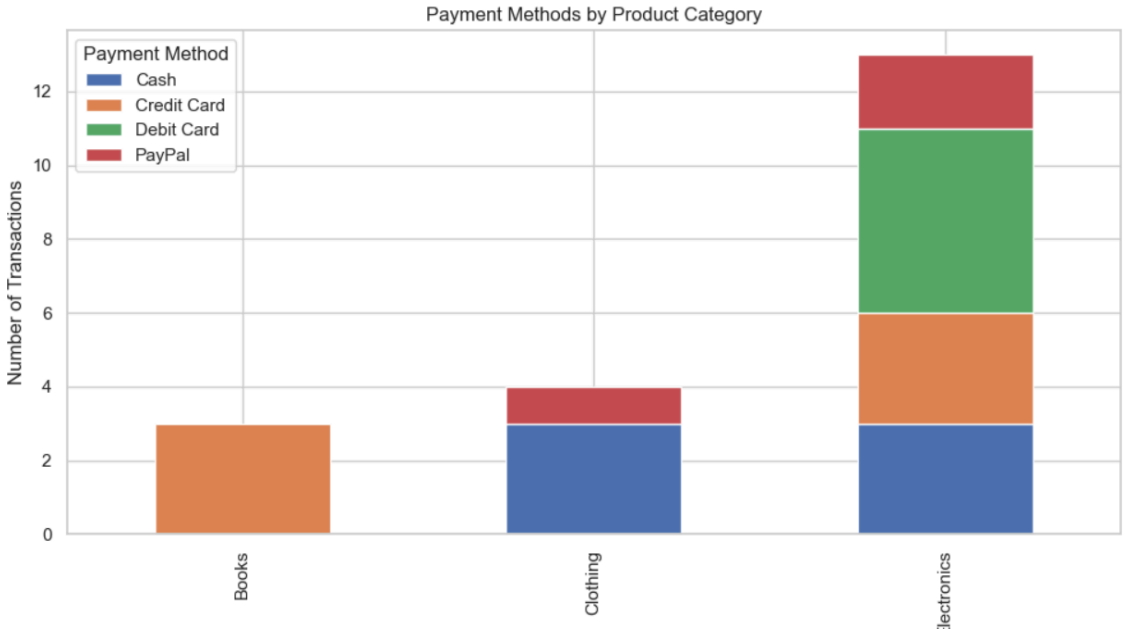
- Monthly Sales (Line Chart): January saw the highest sales, and March saw a decrease in sales.



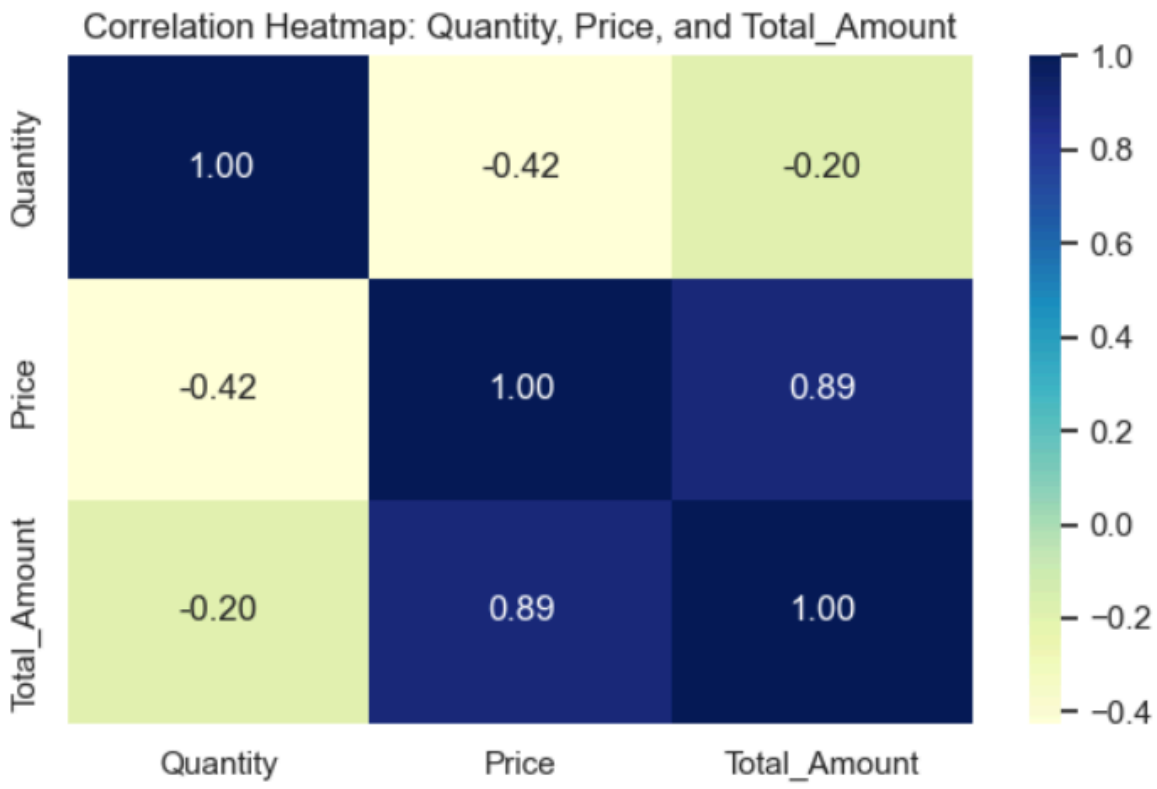
- Top 10 Customers (Bar Chart): Highlighted the highest spenders.



- Payment by Category (Stacked Bar): Showed variation in payment methods by product category.



- Correlation Heatmap: Visualized relationships between numerical fields.



Key Findings

1. Best-Selling Months:
   January has the highest sales , largely due to electronics. March was the weakest month, suggesting a need for seasonal promotions.
2. Top-Selling Product Categories:
   Electronics dominated the list, followed by books and apparel.
3. Repeat customers and a small number of heavy spenders generated much revenue, reflecting the worth of loyalty and retention..
4. Customer Behavior:
   Repeat customers and a small group of top spenders drove a large share of revenue, indicating the value of loyalty and retention.
5. Regional Sales:
   The North and South regions were the strongest performers, making them ideal for testing new campaigns.
6. Payment Preferences:
   Credit cards dominated for electronics, while cash was more common for lower-priced items like clothing and books.

Actionable Recommendations

- Launch targeted campaigns and new products during high-sales months like January to maximize engagement and revenue.
- Maintain strong inventory for electronics, as they consistently drive sales.
- Use seasonal discounts, flash sales, or loyalty rewards to boost sales during low-performing months.
- Implement personalized offers and loyalty programs for repeat and high-value customers.
- To maximize the advantage of the present consumer base, concentrate marketing and inventory efforts in the North and South.
- To simplify the checkout process and improve client satisfaction, highlight the preferred methods of payment for each product category rates of conversion.

Ethical Considerations

All customer data was anonymized, with only customer IDs used in the analysis. No personally identifiable information (PII) was present. The project adhered to GDPR and best practices by ensuring data privacy, limiting access, and recommending secure storage for any sensitive information in future datasets.

Bonus Challenge: Strategies for Low-Performing Months

1. Targeted Promotions:
   Launch time-limited campaigns in March (e.g., “March Magic deal”) to create urgency and re engage customers.
2. Bundle or Discount Offers:
   To boost visibility and sales across segments, connect low-selling stuff, like clothing or books, with devices at a discounted price.

Limitations and Considerations

While the analysis and visualizations provide valuable insight, there are several of important limitations and variables special to this ShopEase dataset and techniques:

- Small Sample Size:
  The dataset contains only 20 transactions over six months. Given the small sample size, trends (like top categories or best-selling months) may not apply to

the company's true customers or future sales. Any conclusions should be considered preliminary.

- Short Time Frame:
  The analysis covers just half a year. Such as a full year or more might change seasonal trends or patterns (such as January's peak sales). A longer data period is needed for a more thorough understanding of seasonality.
- Missing Data and Imputation:
  There was one missing value in the `Total_Amount` column, which was filled using `Quantity × Price`. While logical, this assumes there were no discounts, returns, or errors in quantity/price, which may not always be true in real-world scenarios.
- No Advertising or Marketing Data:
  Campaign, promotion, and advertising spend columns have been excluded from the dataset. This limiting our ability to suggest data-driven marketing strategies as we were not able to analyze the impact of marketing initiatives on sales.
- Limited Customer Behavior Insights:
  With only 20 transactions and anonymized customer IDs, the analysis of repeat customers and loyalty is very limited. Patterns such as customer retention or lifetime value cannot be reliably established from such a small sample.
- Product Category Breadth:
  Some categories (e.g., Electronics) may appear dominant simply because of higher-value items, not necessarily higher volume. The analysis may not distinguish between revenue-driven and volume-driven trends due to the small dataset.
- Regional and Payment Method Trends:
  Very few data points support the observed regional and payment preference differences.. These trends could easily change with more data.
- Data Cleaning Assumptions:
  The code assumes that the date and category fields are filled in correctly and that the Total_Amount is always calculated correctly from Quantity × Price. Any data entry errors in the original file could affect results.
- Visualization Interpretation:
  With such a small dataset, charts like bar graphs or line plots may exaggerate minor differences. It is most effective to see visual trends as illustrative rather than conclusive.

In summary: The report's conclusions and graphics give an appropriate place to start when analyzing ShopEase's sales data, but because of the small sample size, brief.Larger and more thorough datasets should be used in future analyses to provide more reliable business insights.