

Project Coversheet

Full Name	Alvin Siphosenkosi Moyo
Project Title (Example – Week1, Week2, Week3, Week 4)	Week2: Sales and Customer Behaviour Insights – Green Cart Ltd.

Instructions:

Students must download this cover sheet, use it as the first page of their project, and then save the entire document as a PDF before submission.

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.

YOU CAN START YOUR PROJECT FROM HERE

1. Introduction

This project analyses Q2 sales and customer behaviour data for Green Cart Ltd. to support the company's quarterly performance review. The objective was to understand revenue trends, customer purchasing behaviour, delivery performance, and the effectiveness of discounts across regions and product categories.

The analysis combines transactional sales data with customer and product attributes to identify key revenue drivers and operational challenges. A specific focus was placed on Q2 activity (April–June) to ensure insights aligned with the review period. Findings from this analysis are intended to inform marketing prioritisation, operational improvements, and customer engagement strategies for the upcoming quarter.

2. Data Cleaning Summary

Three datasets (sales, customer, and product information) were cleaned and prepared prior to analysis. Key data types were corrected, including converting order and signup dates to datetime format and ensuring numeric fields such as quantity were properly cast. Categorical fields such as gender, loyalty tier, region, and delivery status contained inconsistent formatting and spelling, which were standardised to ensure accurate grouping and aggregation.

Missing discount values were interpreted as zero discount and filled accordingly, while records missing critical transactional values (such as quantity) were removed. Left joins were used when merging datasets to retain all sales transactions. Post-cleaning validation confirmed consistent categories, appropriate data ranges, and a merged dataset suitable for Q2 analysis.

3. Feature Engineering Summary

Several new features were engineered to enhance analysis depth. Revenue was calculated at transaction level using quantity, unit price, and discount rate, providing a consistent financial metric. Time-based features such as order week and order month enabled trend analysis across Q2. Delivery performance was captured using an is_late flag derived from standardised delivery status values.

Customer-level insights were supported by extracting email domains, while product segmentation was achieved through price bands (Low, Medium, High) based on unit price. Additionally, days to order was calculated to measure the time between product launch and purchase, supporting product lifecycle analysis. These features enabled targeted evaluation of sales performance, customer behaviour, and operational efficiency.

4. Key Findings & Trends

Revenue showed a clear upward trend through mid-Q2, peaking around Weeks 24–26 before declining sharply in the final week. High-priced products accounted for the majority of total revenue, despite lower purchase volumes, indicating strong reliance on premium items. Discount analysis showed limited correlation between higher discounts and increased quantities sold, suggesting discounts alone are not driving bulk purchasing.

Delivery performance varied significantly by region, with Central, North, and East regions experiencing late delivery rates well above the target threshold. Regional heatmap analysis revealed strong category-specific performance differences, highlighting opportunities for targeted regional marketing and inventory optimisation.

5. Business Question Answers

The analysis found that high-priced products generate the largest share of revenue across most regions, particularly in Cleaning and Storage categories. Revenue peaked in the latter half of Q2, indicating seasonal or campaign-driven uplift. Discount levels showed weak correlation with quantities purchased, suggesting current discount strategies may not significantly influence order size.

Delivery delays were a notable issue, with several regions exceeding acceptable late delivery rates, posing an operational risk. Regional analysis showed uneven performance across categories, indicating opportunities to tailor product focus and promotions by location. Customer behaviour patterns suggest revenue is driven more by product value and regional demand than by discount intensity alone.

6. Recommendations

Green Cart Ltd. should prioritise improving delivery performance in high-delay regions to reduce operational risk and improve customer satisfaction. Marketing efforts should focus on high-revenue product categories and regions identified as strong performers, rather than relying heavily on discounting strategies. Given the limited impact of discounts on quantity sold, promotional strategies should be reviewed to emphasise value-based positioning rather than price reductions. Additionally, regional performance insights should inform inventory planning and targeted campaigns to maximise Q3 revenue.

7. Data Issues or Risks

Several data quality issues were identified, including inconsistent formatting of categorical fields and missing values in customer attributes. While these were addressed during cleaning, incomplete customer information limits deeper demographic

analysis. Additionally, delivery performance analysis relies on categorical delivery status rather than precise delivery timestamps, restricting granularity. These issues could be mitigated through stricter data validation at entry and enhanced logistics tracking in future datasets.
