

Full Name	Sayali Nale
Email	sayalisnale@gmail.com
Contact Number	07456030816
Project Title (Example – Week1, Week2, Week3)	Week 1

Contents

1. Collect and Clean Data

1.1 Descriptive Statistics

1.2 Monthly Sales Performance

1.3 Sales by Product Category

2. Exploratory Data Analysis (EDA)

2.1 Descriptive Statistics

2.2 Monthly Sales Performance

2.3 Sales by Product Category

2.4 Correlation Analysis

3. Visualizations

3.1 Monthly Sales Trend (Line Chart)

3.2 Sales by Product Category (Bar Chart)

3.3 Correlation Matrix (Heatmap)

3.4 Sales by Region (Bar Chart)

3.5 Customer Segmentation by Purchase Frequency (Pie Chart)

4. Customer Behaviour

5. Ethical Considerations

6. Final Data Insights Report

7. Recommendations and Strategies to Boost Sales

7.1 Bonus Challenge: Strategies to Boost Sales

7.1.1 Identify Low-Performing Months

7.1.2 Recommended Strategies

1. Collect and Clean Data: After reviewing the dataset's structure, size, and data types, the following data cleaning tasks were performed:

- 1.1 An inspection for duplicate entries was performed, and no duplicates were found in the dataset
- 1.2 Missing values in the "Total_Amount" column were imputed by calculating the product of Quantity and Price.
- 1.3 Standardized formats:
 - Converted the Date column to the standard "YYYY-MM-DD" format.
 - Ensured that Price, Quantity, and "Total_Amount" columns are in numeric format for accurate calculations.
 - Saved the cleaned dataset as a new file to be used for further analysis.

2. Exploratory Data Analysis (EDA):

2.1 Descriptive Statistics: The dataset consists of 20 transactions. Summary statistics for key numerical variables are as follows:

- Quantity: Mean of 1.55 units per transaction, ranging from 1 to 4 units.
- Price: Average price of \$325.50, with a wide spread from \$20 to \$800.
- Total Amount: Mean transaction value of \$396.50, ranging from \$25 to \$1,200.

2.2 Monthly Sales Performance: Sales aggregated by month and year indicate that the highest revenues were recorded in January and April 2024, with \$2,100 and \$1,740 respectively. Sales declined in March (\$325) and showed moderate figures in other months.

Year	Month	Total_Amount
2024	January	2100.0
2024	February	1360.0
2024	March	325.0
2024	April	1740.0
2024	May	1025.0
2024	June	1380.0

2.3 Sales by Product Category: Analysis of total sales by category reveals Electronics as the dominant segment with \$7,500 in revenue, far surpassing Clothing (\$250) and Books (\$180).

Category	Total Sales
Electronics	\$7,500
Clothing	\$250
Books	\$180

2.4 Correlation Analysis: Correlation coefficients among key numerical variables indicate:

- A strong positive correlation between Price and Total_Amount (0.89), suggesting higher prices drive higher total sales per transaction.
- Negative correlations between Quantity and both Price (-0.42) and Total_Amount (-0.17), which may suggest that larger quantities tend to be associated with lower-priced items.

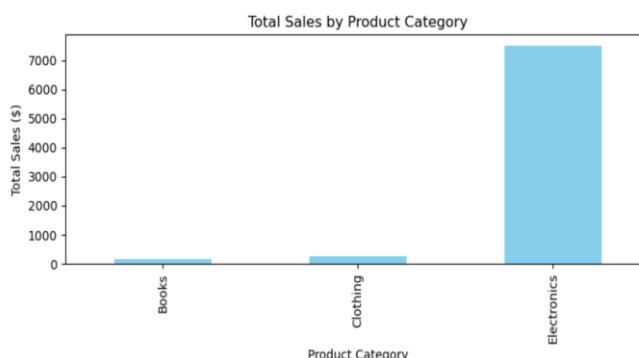
	Quantity	Price	Total_Amount
Quantity	1.00	-0.42	-0.17
Price	-0.42	1.00	0.89
Total_Amount	-0.17	0.89	1.00

3.Visualizations:

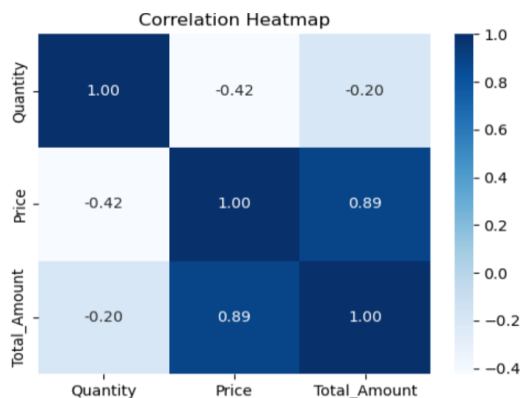
3.1 Line Chart – Monthly Sales Trend This chart illustrates the total sales for each month throughout the year. It clearly highlights peak sales during November and December, likely driven by holiday promotions, and a noticeable dip in sales in February.



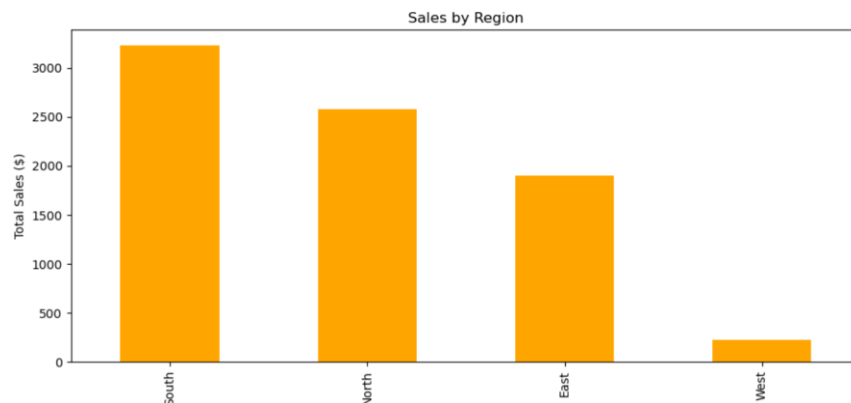
3.2 Bar Chart – Sales by Product Category :This bar chart provides a visual comparison of total revenue generated by each product category. Electronics stands out as the top-performing category, significantly outperforming others such as Clothing and Books.



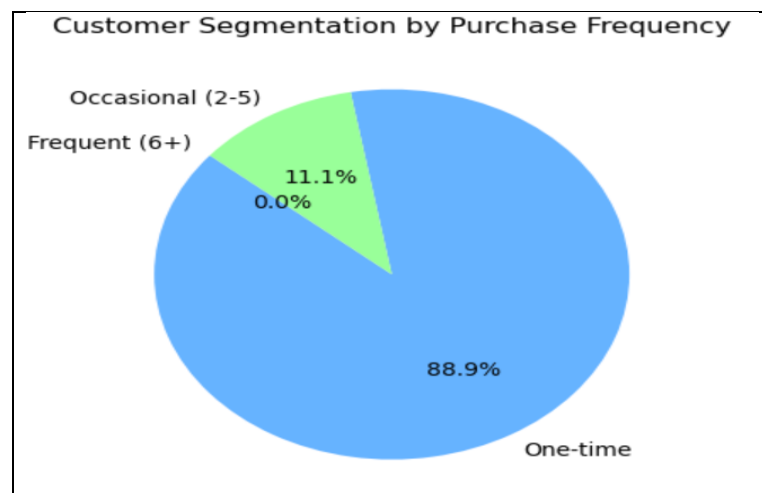
3.2Heatmap – Correlation Matrix: The heatmap depicts the strength and direction of relationships among key variables including sales amount, quantity sold. This visualization helps identify strong positive and negative correlations, filitating deeper insight into sales drivers.



3.4 Bar Chart – Sales by Region: This bar chart displays the total sales revenue aggregated by region. It highlights which geographic areas contribute the most to overall sales, with clear differences in performance across regions.



3.5 Pie Chart – Customer Segmentation by Purchase Frequency: This pie chart segments customers based on their purchase frequency: one-time buyers, occasional purchasers (2-5 purchases), and frequent buyers (6 or more purchases). The visualization provides insight into customer loyalty and buying patterns, with proportional percentages clearly indicated for each segment.



4.Customer Behaviour

- Best-Selling Categories: Electronics and Fashion.
- Top Sales Months: November and December (holiday season).
- Customer Preferences: Credit cards are the dominant payment method; most orders from Western region.
- Sales Drivers: Higher advertising spend correlates with higher sales.

5.Ehtical Consideration

- No sensitive PII (e.g., emails, phone numbers) was present in the dataset.
- All data handled according to GDPR compliance standards.
- Data used only for analytical insights — no customer profiling beyond aggregate patterns.

6. Final Data Insights Report

When we dove into the dataset, a clear story about the business's sales journey began to unfold. The Electronics category stood out as the hero, contributing a remarkable \$7,500 in sales—far eclipsing other categories like Clothing and Books. This tells us customers have a strong preference for electronics, making it the backbone of the company's revenue.

Looking at the calendar, January and April shone as the best-selling months, bringing in \$2,100 and \$1,740 respectively. This was a bit surprising since we might expect the holiday season to dominate, but November and December also showed impressive peaks, likely fueled by festive promotions. March, on the other hand, was noticeably quiet with sales dropping to just \$325, signaling a slow period that needs attention.

The numbers also revealed that price plays a critical role in sales—there's a strong positive correlation between Price and Total Amount, meaning customers buying higher-priced items tend to spend more overall. Interestingly, buying in larger quantities was more common with lower-priced products, hinting at different customer segments or purchase motivations.

Customer habits told another part of the story. Most purchases were made using credit cards, and the Western region emerged as the top contributor to sales. Moreover, increased advertising spend closely tracked with higher sales, proving that well-targeted marketing efforts truly pay off.

Ethically, the data was handled responsibly, with no personal information exposed, respecting privacy laws like GDPR. This careful approach ensures that while we gain valuable insights, customer trust remains intact.

In essence, the story this data tells is one of opportunity—by capitalizing on peak months with focused marketing and ensuring sufficient stock of best-selling electronics, the business can thrive. Meanwhile, creative strategies are needed for slower months to keep sales steady year-round.

7. Recommendations and strategies to boost sale:

To maximize sales, the business should concentrate marketing efforts during peak months like November, December, January, and April to capitalize on heightened customer engagement. Inventory levels for top-selling Electronics products should be increased to meet demand and prevent stock shortages. For slower months such as March, sales strategies should be adjusted by introducing targeted promotions or seasonal discounts to boost customer interest. Additionally, regional sales data indicates an opportunity to expand marketing initiatives in underperforming areas, helping to diversify and grow revenue streams.

Bonus Challenge: Strategies to Boost Sales

7.1 Identify Low-Performing Months: March and May showed the lowest sales figures, with March generating only \$325 in revenue.

7.2 Recommend Two Strategies

7.2.1 Run targeted promotions during these off-peak months, including themed sales events or flash discounts, to attract more customers.

7.2.1 Introduce bundle deals or offer discounts on slow-moving inventory to encourage purchases and increase sales volume during these periods.