

Analysis Done by: Sindio-Apaun Mac-John

Date: June 12th 2025

Dataset: Cleaned Sales_Data.xlsx

1. Cleaning and Standardizing

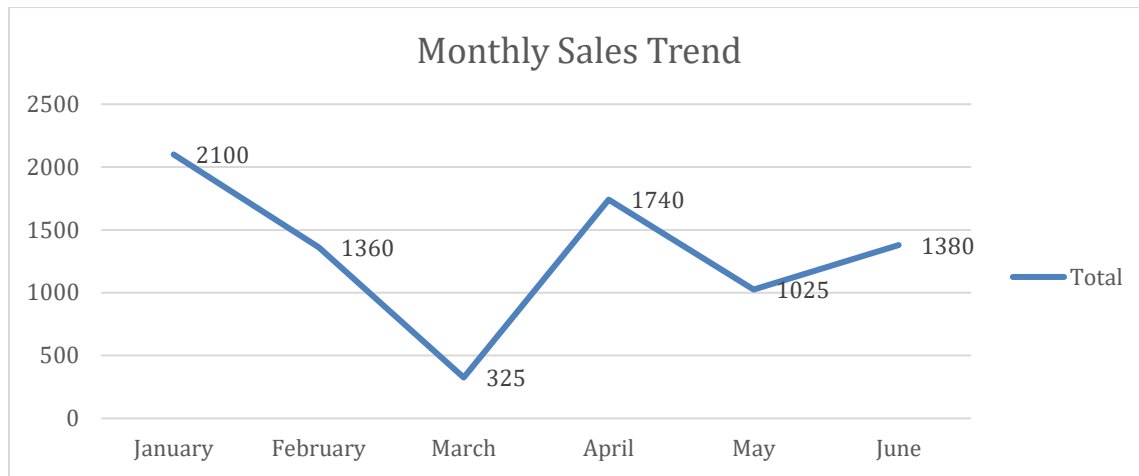
- I removed 2 duplicate entries using the *“remove duplicate”* function in excel.
- Filled missing values in *“Total_Amount”* using the existing pattern which is *“Total_Amount = Price * Quantity”*.
- Standardized date format to YYYY-MM-DD.
- Ensured numerical columns had consistent data types.
- Added a new column called the *“Month”* Column to facilitate my EDA Summary and Visualization
- Exported cleaned file as 'Clean sales_data.xlsx'.

2. Exploratory Data Analysis Summary

- Peak sales occurred in January, with the highest sale coming from the sales of smartphones.
- Electronics category had the highest total sales with approximately 94% of General Sales.
- There Existed a Positive correlation (≈ 0.96) between Advertising Spend and Sales Total.
- Customer preference leaned heavily towards credit card payments with Paypal being the least used.
- North and South regions dominated overall sales with each of the regions providing 28% each of the entire sales.

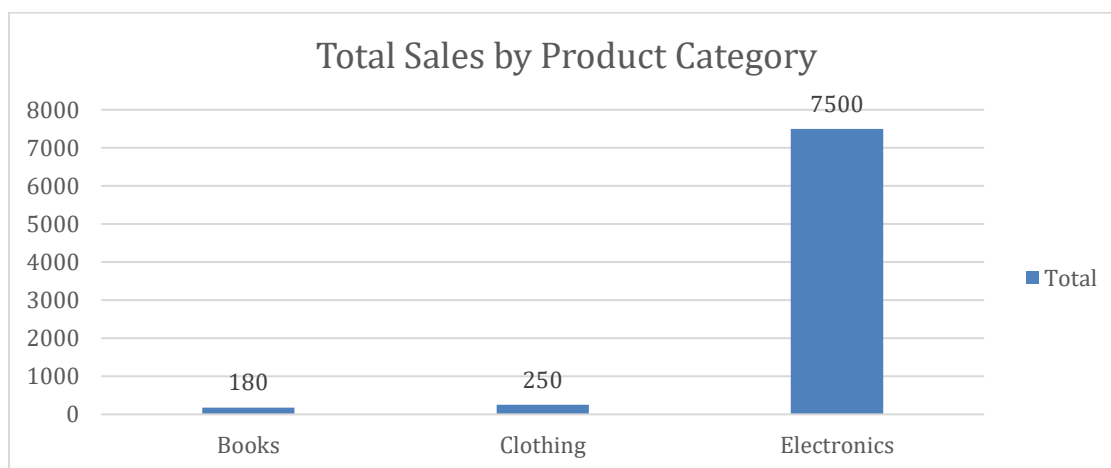
3. Data Visualizations

- **Monthly Sales Trend**



Based on the Monthly Sales Trend, we can observe the year started with a peak sale and the steady decrease in sale over the next two months.

- **Total Sales by Product Category**



The Chart above revealed Electronics as the dominant category, with figures far exceeding those of Clothing and Books combined."

• **Correlation Matrix**

*	ADVERTISING_SPEND	QUANTITY	TOTAL_SALES
ADVERTISING_SPEND	1	-0.432686952	0.957753649
QUANTITY	-0.432686952	1	-0.268289939
TOTAL_SALES	0.957753649	-0.268289939	1

The heatmap reveals a strong positive correlation (≈ 0.96) between Advertising Spend and Total Sales, highlighting the significant impact of marketing on revenue. Other variables showed weak or negligible correlations.

4. Final Data Insights Report

Major Observations

Between January and June 2024, ShopEase experienced a shaky sales momentum, the year started with a crazy sale record in January, which topped the charts in terms of total revenue, followed by a steady decrease within the next 2 months and a fluctuating movement from April to June.

Electronics, particularly smartphones, drove this success in January, accounting for nearly all revenue generated during the period although Books sold more pieces in total and sold more especially in June, which one would assume is as a result of the resumptions of schools.

Advertising spends proved to be a major sales driver. I observed a near-perfect correlation coefficient of 0.96 between Ad Spend and Sales, showing the importance of marketing campaigns.

In terms of geography, the North and South regions contributed significantly, each covering around 28% of all transactions.

Customer preference skewed toward credit card payments, with PayPal showing minimal traction.

Recommendations

After much analysis I recommend we take some of the following suggestions to increase sales.

- **Boost Ads During Low Months:** Increase advertising efforts in March and May specifically to address the sales decline observed during those periods.
- **Stock & Highlight Bestsellers:** Ensure steady availability of high-demand products like smartphones and Books which exceeded others in quantity sold.
- **Regional Loyalty Campaigns:** Launch reward programs or discount sales in underperforming regions like West to balance geographic sales distribution.
- **Payment Incentives:** Introduce small discounts (e.g., 5%) for hardly used payment methods like PayPal to diversify the different method used.

Ethical Considerations

- I ensured to anonymize all customers identifiers
- There wasn't any personal sensitive information (e.g Phone Number) included in the data set
- I followed all GDPR and Data Privacy Law

Conclusively:

In order to increase sales in the two lowest performing months (March and June) based on my data insights,

- **Increase Ad. Spend:** based on our correlation analysis we observed that sales and ad. spend are almost perfectly correlated.

This means that an increase in the ad spend on March and June would statistically increase sale.

- **Bundling Offer:** As per our mentor Madam Natasza, she highlighted the importance of bundling the bestselling products with other products with a little increase in price which would generate more sales of different products.

Example: Buy a smartphone and get 50% off a book, or buy 2 clothing items and get a get one free.