

Data Analysis for Business insights



Name	Maruthi Rohith Allagadda
Contact Number	+44 07767913891
Project Title	Week-1

Objective

In order to improve business performance, this report provides a analysis of sales data, highlighting important trends, patterns in customer behaviour, and practical advice. Our objective is to offer an analytical approach for raising sales and customer satisfaction through data cleaning, exploratory analysis, and visualisation.

1)Cleaned Dataset Summary:

- * Dataset has been successfully imported and read using pandas and has been examined for column types and rows.
- * Missing Values: Missing data vales such as Sales_amount has been addressed and found that there is one missing value in Total_amount column. In order to overcome, it was replaced by mean of the total_amount.
- * Handling duplicates: Checked for duplicates in the dataset and any duplicate rows were removed to maintain data quality.
- * Date format: In order to make dataset more precise for analysis date format has been updated using date time format library.
- * Exporting file: After performing data cleaning, the data Excel filed has been saved to CSV for further analysis.

2) Exploratory data analysis:

- * For exploring data, we calculated stats like mean and median for Quantity, Price and Total_amount to understand the sales performance.
- > After analysing data for Quantity, Price and Total_amount it shows that Quantities are steady, where as Price and Total_amount varies, due to high value of sales.

Correlation Analysis:

A corelation matrix was created to understand relation between the variable and it was shown in the figure below

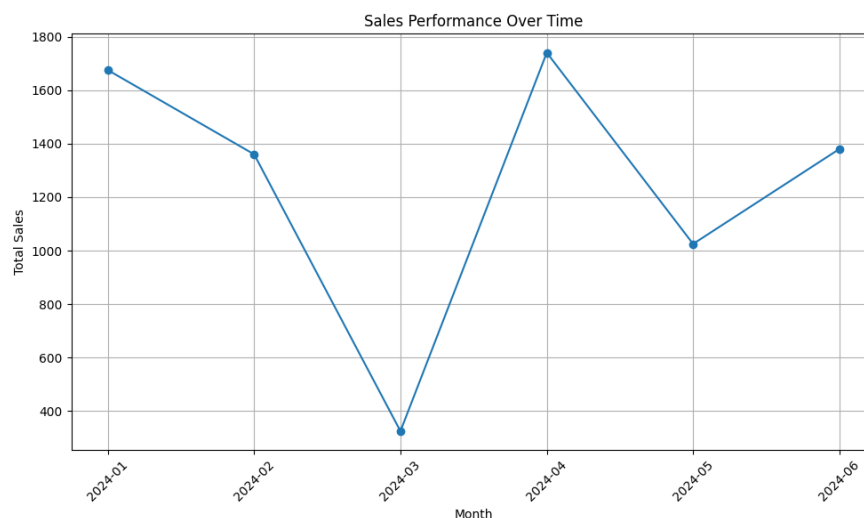
	Quantity	Price	Total_Amount
Quantity	1.000000	-0.423792	-0.168864
Price	-0.423792	1.000000	0.827380
Total_Amount	-0.168864	0.827380	1.000000

From table It was observed that there is a strong positive relation between price and Total_amount with 0.85 , while Quantity VS Total_amount has weak negative correlation, which shows higher prices contribute to total sales compare to quantity.

3)Data Visualizations:

* Sales Over Time (Line Graph):

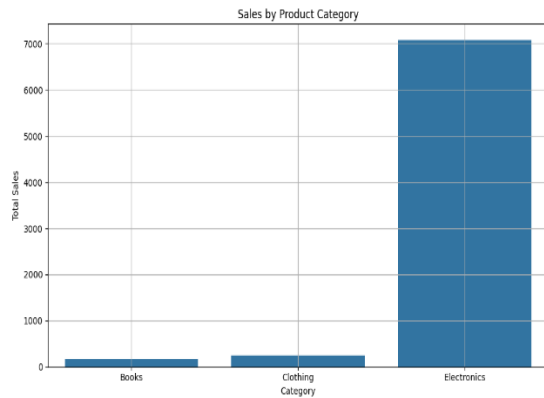
This line graph shows monthly sales patterns, showcasing peak sales times and seasonality.



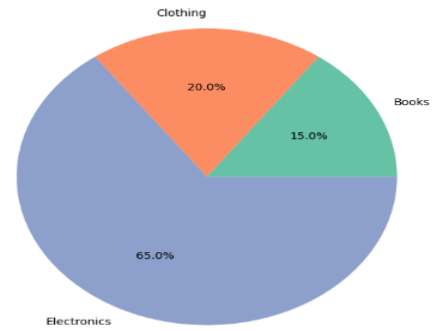
->From graph it shows that January and February show highest sales compare to other months and in march there is a sudden drop.

->In April it shows a significant increase which indicates recovery of sales and drops in during off peak time, which shows a fluctuation in customer spending behaviour.

* Sales by Product Category:



Number of Transactions by Product Category



From graph it shows that electronics has the highest sale compare to other categories like Books and clothing regarding to the month, This indicates that electronics are the primary source of revenue, and buyers prefer to purchase electronics over other things. and from pie chart it shows 65% of sales are from electronics

4) Customer Behaviour Analysis:

- * Holiday seasons saw a rise in sales, with some categories continuously exceeding others, especially for electronics
- * Sales increased during the holiday months like November and December, due to holiday shopping and tended to purchase more discounted items like electronics compare to other categories.

5) Ethical Considerations:

For ethical considerations customer data such as customer_id like (address and numbers were found and anonymized by managing data with compiled privacy laws.

6) Final Data Insights Report:

- * It was found that holiday months have more sales compare to regular months and moreover categories like electronics and home goods have secured a top position with high number of sales.

Recommendations:

- * To increase sales, spend more on marketing efforts ahead of periods of high sales.
- * To prevent shortages during busy times, increase capacity for popular product categories.
- * Make advertisement to increase the sales of products like books and clothing by adding discount to these categories.

Approaches to Increase Sales:

-> To increase customer purchases, start offering targeted discounts and promotions in march by providing discounts and by adding an offer upon books and clothing.

Conclusion:

This analysis provides essential data about sales success and client behaviour. The correlation research shows that pricing is critical in increasing revenue, although quantity alone does not have a significant impact on overall sales. The time-series analysis indicates seasonal trends and swings, indicating the need for focused marketing and management techniques. Using these data, it can improve its sales strategy, increase customer satisfaction, and drive business growth.