

REPORT ASSIGNMENT TOPIC 1

BUSINESS CASE 2: SALE MARKETING PROJECT

Executive Summary

The primary goal of this sales data project is to analyse historical sales data to gain insights into sales trends, optimise marketing strategies, and enhance overall business performance. By examining sales records, we aim to improve decision-making, boost revenue, and enhance customer satisfaction. This report outlines the methodology, exploratory data analysis (EDA), model development, and conclusions drawn from the data.

Introduction

The sales data project addresses the need for in-depth analysis of sales performance to inform business strategies. The project focuses on identifying sales trends, optimising product performance, understanding regional and country-specific insights, and improving sales channel effectiveness. By leveraging the dataset, we seek to refine marketing campaigns, forecast future sales, and increase operational efficiency.

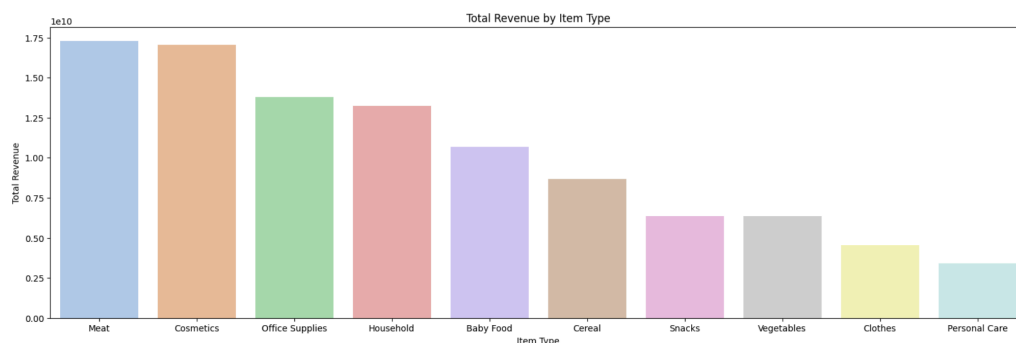
Methodology

The project follows a structured methodology involving data collection, preprocessing, exploratory data analysis (EDA), model training, and evaluation. Python libraries such as pandas, numpy, matplotlib, seaborn, and scikit-learn are employed for data manipulation, visualisation, and modelling.

Exploratory Data Analysis (EDA)

Review Total Revenue by different groups such as Item Type, Sales Channel, Region, Country.

1. Total Revenue by Item Type

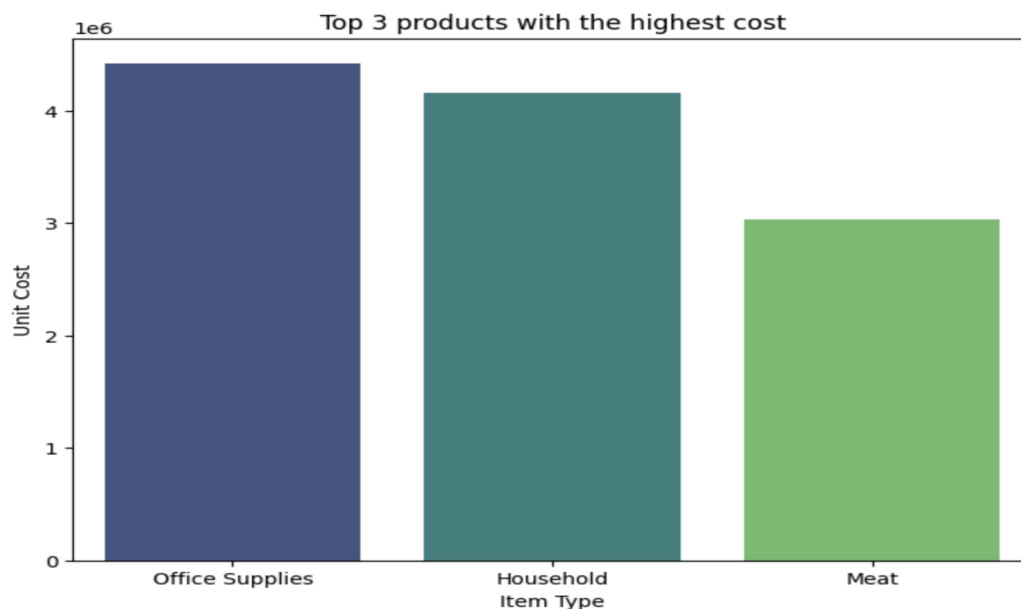


The chart reveals the following insights into product revenues:

- **Highest Revenue Product:**
 - **Meat** leads with the highest revenue.
- **Other High-Revenue Products:**
 - **Cosmetics, Office Supplies, and Household** items also generate significant revenue.
- **Average Revenue Products:**
 - **Baby Food, Cereal, Snacks, Clothes, and Personal Care** fall into the average revenue category.

Conclusion: Understanding revenue distribution helps identify key revenue drivers and average performers, guiding strategies to boost sales and optimise product offerings.

2. Top 3 products with the highest cost



The chart highlights the top three product categories with the highest unit costs:

A. Office Supplies and Household Appliances:

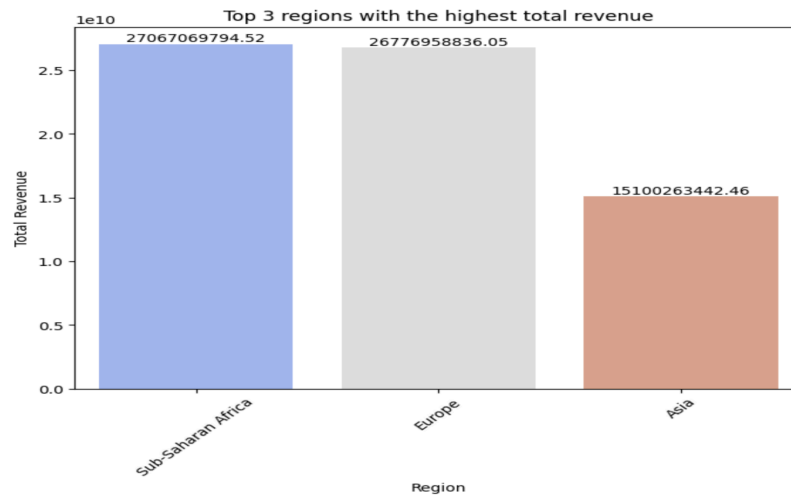
- a. Both categories have exceptionally high unit costs, suggesting significant expenses in production or procurement.

B. Meat:

- a. While less expensive than the top two, meat still ranks among the highest unit costs, indicating notable cost implications.

Conclusion: Focusing on managing costs for Office Supplies, Household Appliances, and Meat is crucial for optimising profitability. Effective cost control in these areas can significantly enhance financial performance.

3. Top 3 regions with the most revenue:

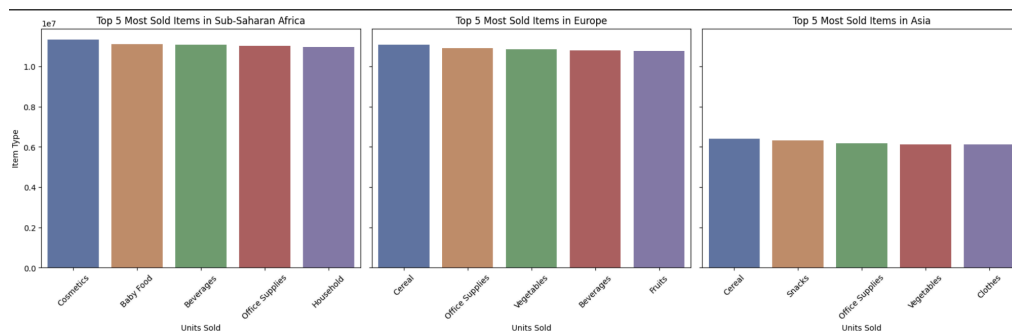


The chart highlights the top three regions with the most revenue

- Sub-Saharan Africa leads, indicating significant market potential.
- Europe shows strong and stable revenue, close to Sub-Saharan Africa.
- Asia, while lower, remains an important market.

Conclusion: The chart highlights key revenue regions. Businesses should focus on Sub-Saharan Africa and Europe while exploring growth opportunities in Asia.

4. Top 5 most sold items in top 3 Regions



Sub-Saharan Africa:

- Top items: Cosmetics, Baby Food, Beverages, Office Supplies, Household.
- Sales are approximately 10 million units per item, showing similar sales volumes.

Europe:

- Top items: Cereal, Office Supplies, Vegetables, Beverages, Fruits.
- Sales are around 10 million units per item, indicating balanced sales.

Asia:

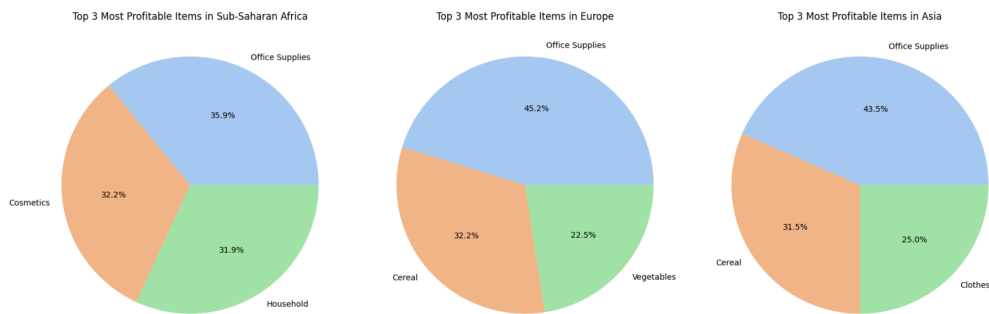
- Top items: Cereal, Snacks, Office Supplies, Vegetables, Clothes.

- Sales are about 7 million units per item, with relatively even distribution.

Conclusion:

- Office Supplies is the only item appearing in the top 5 across all three regions.
- Each region has distinct top-selling items, reflecting different consumption patterns.
- Within each region, sales volumes of top items are generally balanced, indicating steady demand.

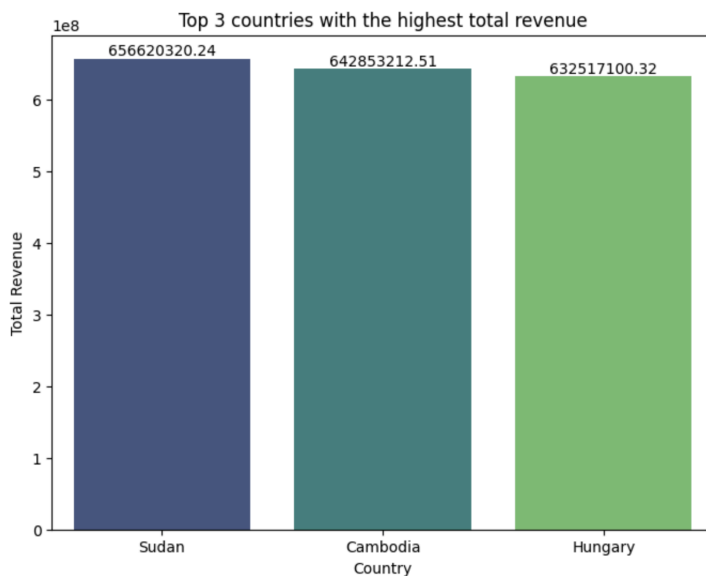
5. Top 3 most profitable products



Conclusion:

- Office Supplies is the top profitable item across all three regions.
- Profit share of other items varies by region, reflecting differing market demands.

6. Top 3 Countries with the most revenue



Top Countries:

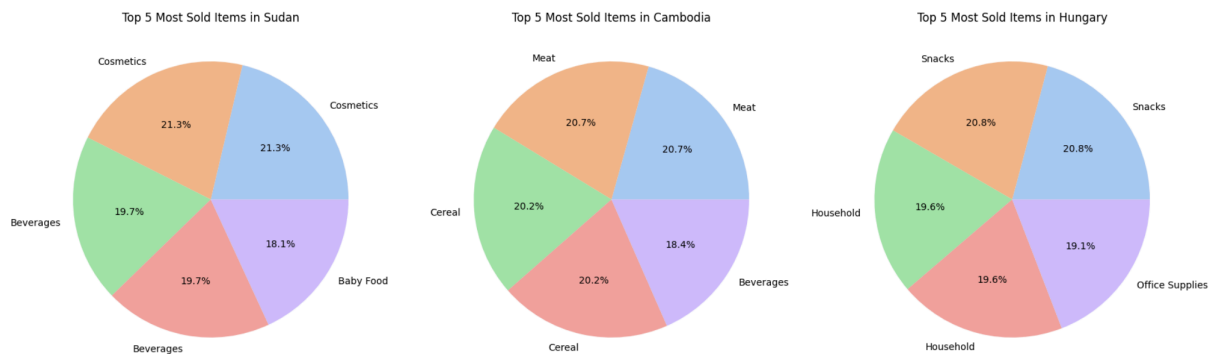
- **Sudan:** Highest revenue
- **Cambodia:** Second highest

- **Hungary:** Third highest

Conclusion:

- Revenue differences among these top three countries are minimal, with all exceeding \$630 million.
- This indicates a competitive market among these nations, with similar revenue levels.

7. Top 5 most sold items in top 3 countries

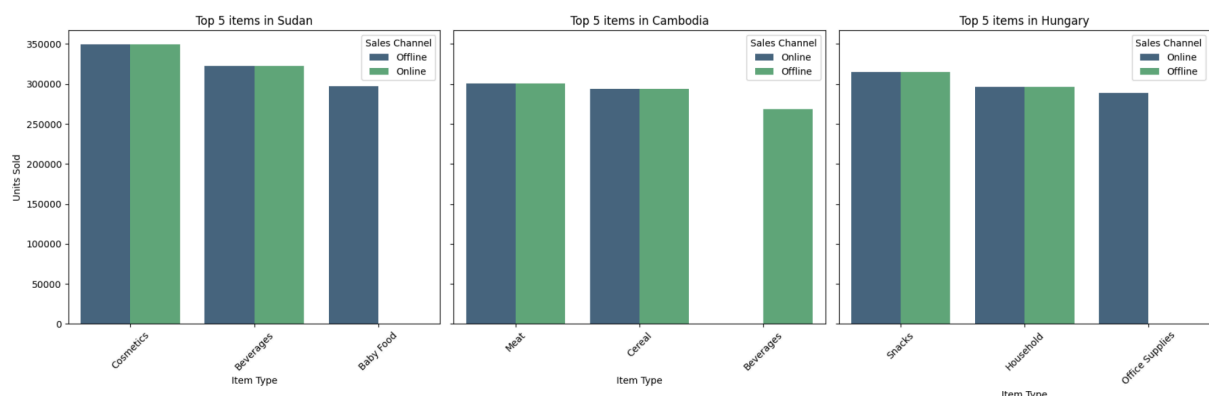


- **Sudan:** Items like Cosmetics, Snacks, and Beverages have significant shares, showing diverse product preferences.
- **Cambodia:** Office Supplies and Clothes dominate, reflecting distinct market demand.
- **Hungary:** Household and Cereal items are prominent, highlighting different sales priorities compared to other countries.

Conclusion:

- Each country shows a unique distribution of top-selling items, with distinct preferences and market trends evident in the pie charts.

8. Top 3 items of each countries in Sales Channel

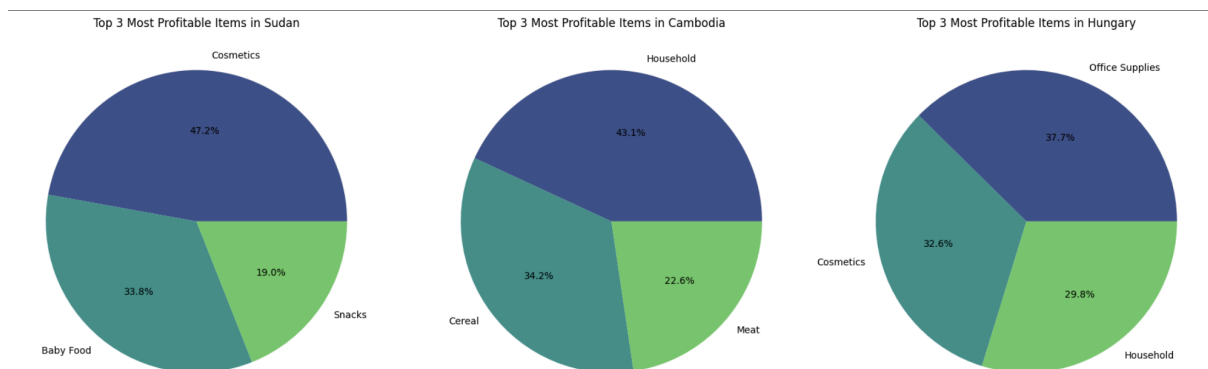


- **Sudan:** Cosmetics lead in sales with a balanced distribution between online and offline. Beverages and Baby Food follow with similar sales patterns.

- **Cambodia:** Meat and Cereal are top sellers, with nearly equal sales across both channels. Beverages have lower sales compared to Meat and Cereal.
- **Hungary:** Snacks are the top-selling item with balanced sales across channels. Household items also show similar sales distribution, while Office Supplies lag behind.

Conclusion: All three countries exhibit balanced sales between online and offline channels, with varying top-selling items reflecting different consumer preferences.

9. Top 3 most profitable Items in top 5 most sold Items of top 3 countries



Conclusion: The charts show the top 3 most profitable items in Sudan, Cambodia, and Hungary, with Cosmetics leading in Sudan, Household items in Cambodia, and Office Supplies in Hungary.

Model Development and Evaluation

-Models:

1. Linear Regression: Predicts outcomes using linear relationships.
2. Decision Tree Regressor: Splits data to predict outcomes.
3. Random Forest Regressor: Combines multiple decision trees for better accuracy.

-Parameters Tuned:

1. Linear Regression: Regularization (Lasso, Ridge) to prevent overfitting.
2. Decision Tree Regressor: Max Depth and Min Samples Split to control tree complexity.
3. Random Forest Regressor: Number of Estimators and Max Features for robustness and performance.

-Training and Testing:

Process:

1. **Training:** Model learns patterns from the training data.

2. **Testing:** Model's performance is evaluated on unseen test data.

Metrics:

- **Accuracy:** Proportion of correct predictions.
- **Precision/Recall:** Measures of classification performance.
- **MSE/R-squared:** Metrics for regression performance.

-Metrics for Evaluation:

- **Evaluation Metrics:** Metrics such as mean absolute error (MAE), root mean squared error (RMSE), and R-squared were used to evaluate model performance. The Random Forest Regressor demonstrated superior accuracy compared to other models.

Conclusion

This project demonstrates the application of advanced analytics to sales data, providing actionable insights into sales trends, product performance, and regional market dynamics. The findings can help the business optimise marketing strategies, forecast sales, and improve overall efficiency. Future work should focus on deploying these models in real-world scenarios and continuously updating them with new data.