Online Retail Sales Analysis

Identifying Bestselling Products, Customer Behavior & Peak Shopping Times

Business Problem

The business aims to better understand customer purchasing behavior in order to improve marketing effectiveness, inventory planning, and revenue growth.

Key questions included:

- Which products are the bestsellers?

- When do customers shop the most?

- Do repeat customers contribute significantly to revenue?

Dataset

Source: Online Retail Dataset (Kaggle)
Scope: United Kingdom transactions only
Size: ~46,000 rows after filtering
Period: 2010–2011

Data Quality Assessment

Negative quantities due to returns, cancelled invoices starting with 'C', and missing CustomerID values were identified. Cancelled invoices were removed, while missing CustomerIDs were retained and documented.

Data Preparation

Filtered UK data, removed cancelled invoices, corrected data types, and created TotalSales = Quantity × UnitPrice.

Analysis Performed

• Bestselling Products by quantity and revenue
• Sales by hour and day
• Repeat customer analysis
• Average Order Value analysis

Key Insights

A small number of products drive most revenue.

Sales peak during midday hours.

Repeat customers generate higher lifetime value.

Business Recommendations

1. Implement a loyalty program to increase retention.
2. Schedule promotions before peak shopping hours.
3. Focus upselling during high AOV periods.

Tools Used

Power BI, DAX, Power Query, Excel

Outcome

This project demonstrates an end-to-end data analytics workflow with actionable business insights applicable to real-world retail scenarios.