

Dataset Summary

Source: Online Retail Dataset from UCI Machine Learning Repository

Rows Analyzed: First 10,000 records

Date Range: December 1–5, 2010

Description: Transactional data from a UK-based online retailer, including invoice info, products, quantities, prices, and customer countries.

Data Cleaning Steps

Removed rows with missing CustomerID values (2,291 rows).

Flagged returns using a new column: Returned? (based on negative Quantity).

Extracted clean invoice dates into a new column: IDate (from InvoiceDate).

Calculated TotalPrice as $\text{Quantity} \times \text{UnitPrice}$ (excluding returned items).

Key Notes

Only non-returned transactions were used for revenue analysis.

UK accounted for most transactions in this subset.

Data used for Tableau dashboard analysis and insights.