# Dataset Note: ECommerce\_Consumer Behaviour.csv

## Overview

The ECommerce\_Consumer Behaviour.csv dataset is part of the “Sales” category of sample data available from Excel BI Analytics, designed for testing and analysis in Power Query, Power Pivot, and data processing or ETL pipelines. It provides a large-scale, realistic simulation of global e-commerce sales transactions, making it suitable for practicing data cleaning, data transformation, business analytics, and machine learning tasks.

## Dataset Description

File Name: ECommerce\_Consumer Behaviour.csv

File Type: CSV (compressed in ZIP/7z format)

Size: ~5,000,000 rows × 14 columns

Category: Sales / E-commerce Transactions

Data Type Mix: Text, Numbers, and Dates

Purpose: Testing data loading, transformation, aggregation, and modeling processes.

## Column Details

|  |  |  |
| --- | --- | --- |
| Column Name | Description | Data Type |
| Region | Geographic region where the sale occurred | Object (Text) |
| Country | Country name | Object (Text) |
| Item Type | Category of product sold (e.g., Cereal, Snacks, Beverages) | Object (Text) |
| Sales Channel | Mode of sale (Online / Offline) | Object (Text) |
| Order Priority | Priority of the order (H=High, M=Medium, L=Low, C=Critical) | Object (Text) |
| Order Date | Date when the order was placed | Object (Text) |
| Order ID | Unique identifier for each transaction | Integer |
| Ship Date | Date when the order was shipped | Object (Text) |
| Units Sold | Quantity of items sold | Integer |
| Unit Price | Price per unit of item | Float |
| Unit Cost | Cost per unit of item | Float |
| Total Revenue | Total revenue = Units Sold × Unit Price | Float |
| Total Cost | Total cost = Units Sold × Unit Cost | Float |
| Total Profit | Profit = Total Revenue − Total Cost | Float |

## Sample Records

Region | Country | Item Type | Sales Channel | Order Priority | Units Sold | Unit Price | Total Profit  
--------|----------|------------|----------------|----------------|-------------|-------------|---------------  
Australia and Oceania | Palau | Office Supplies | Online | H | 2401 | 651.21 | 303,126.25  
Europe | Poland | Beverages | Online | L | 9,339 | 47.45 | 146,264.40  
North America | Canada | Cereal | Online | M | 103 | 205.70 | 9,124.77  
Europe | Belarus | Snacks | Online | C | 1,414 | 152.58 | 77,967.96  
Middle East and North Africa | Oman | Cereal | Offline | H | 7,030 | 205.70 | 622,521.93

## Key Insights / Analytical Opportunities

• Sales trend analysis across regions, product categories, and channels.

• Profitability and cost efficiency comparisons between product types.

• Regional or country-level performance dashboards.

• Customer behavior simulation for online vs. offline sales.

• Forecasting and machine learning experiments for revenue or profit prediction.

• ETL testing and Power BI / Power Pivot integration for business intelligence workflows.

## Data Quality & Limitations

• The data is synthetic, generated via Excel macros (Country Sales Data Generator).

• May not perfectly represent real-world distributions or business cycles.

• Suitable for testing, practice, and demonstrations — not real customer data.

• Contains no personally identifiable or confidential information.

## Conclusion

The ECommerce\_Consumer Behaviour.csv dataset is a rich, large-scale synthetic dataset that accurately reflects typical sales transaction structures. It serves as an excellent resource for testing ETL pipelines, data modeling, visualization tools, and machine learning models in e-commerce analytics.

## Reference

* https://excelbianalytics.com/wp/downloads-18-sample-csv-files-data-sets-for-testing-sales/