

About Dataset

This synthetic dataset represents **E-commerce sales transactions** containing **34,500 records across 17 features**. It has been carefully designed to simulate realistic online shopping data and can be used for multiple data science and machine learning tasks, including:

- **Sales Analysis** – revenue trends, profit margins, regional performance, category-wise sales
- **Customer Segmentation** – analyzing customer demographics, purchase behavior, loyalty scores
- **Churn Prediction** – identifying customers likely to stop purchasing
- **Product Performance** – tracking returns, pricing impact, and demand across categories

Data Columns

- order_id → Unique identifier for each order
- customer_id → Unique identifier for each customer
- product_id → Unique identifier for each product
- category → Product category (Electronics, Fashion, Home, Beauty, Sports, Toys, Grocery)
- price → Unit price of the product
- discount → Discount applied (%)
- quantity → Number of items purchased
- payment_method → Payment type (Credit Card, Debit Card, UPI, PayPal, COD, Wallet)
- order_date → Date of purchase
- delivery_time_days → Days taken to deliver the order
- region → Geographic region of the customer
- returned → Whether the product was returned (Yes/No)
- total_amount → Final bill amount after discounts
- shipping_cost → Delivery charges
- profit_margin → Profit earned from the order
- customer_age → Age of the customer (18–70)
- customer_gender → Gender of the customer (Male/Female/Other)

Link to Data Set:

<https://www.kaggle.com/datasets/miadul/e-commerce-sales-transactions-dataset>