# Phân tích dữ liệu kinh doanh từ bộ dữ liệu Online Retail

## 1 – Phân tích phân bố log(Tổng giá trị giao dịch) A graph with blue line AI-generated content may be incorrect.

Biểu đồ cho thấy phân bố log của tổng giá trị giao dịch bị lệch phải mạnh, nghĩa là phần lớn các giao dịch có giá trị nhỏ, chỉ có một số ít giao dịch có giá trị rất lớn. Điều này phản ánh quy luật Pareto thường thấy trong kinh doanh – một số ít giao dịch lớn chiếm phần lớn doanh thu.  
  
Kết luận: Dữ liệu có độ lệch cao, cần chuẩn hóa hoặc log-transform để tránh ảnh hưởng đến các mô hình thống kê. Các giao dịch lớn là đối tượng cần chú trọng phân tích thêm.

## 2 – Kiểm tra các điểm ngoại lai bằng Boxplot (log-transformed) A graph with a blue rectangle and black lines AI-generated content may be incorrect.

Boxplot cho thấy mặc dù đã log-transform, dữ liệu vẫn có nhiều điểm ngoại lai ở phía phải (giá trị log cao). Các outlier này đại diện cho những đơn hàng có giá trị rất cao, khác biệt lớn so với trung bình.  
  
Kết luận: Dữ liệu chứa nhiều outlier, cần xem xét xử lý hoặc phân tích riêng nhóm khách hàng có tổng giá trị giao dịch cao để hiểu nguyên nhân.

## 3 – Phân tích mối tương quan giữa các biến trong dữ liệu A blue and yellow squares with white text AI-generated content may be incorrect.

Có tương quan rất mạnh (0.91) giữa Quantity và TotalPrice, cho thấy tổng giá trị giao dịch phụ thuộc chủ yếu vào số lượng sản phẩm bán ra. Ngược lại, UnitPrice gần như không có tương quan với Quantity (≈ 0), tức giá sản phẩm không ảnh hưởng nhiều đến số lượng mua.  
  
Kết luận: Số lượng hàng bán ra là yếu tố quyết định chính đối với doanh thu. Việc tăng số lượng bán sẽ có tác động lớn hơn so với việc điều chỉnh giá.

## 4 – Phân tích doanh thu theo thời gian (Doanh thu theo tháng) A graph with a line and a line AI-generated content may be incorrect.

Doanh thu dao động theo thời gian, tăng mạnh vào cuối năm (tháng 10–11), trùng với mùa lễ hội và mua sắm Giáng Sinh. Giữa năm (tháng 4–7) doanh thu giảm nhẹ, thể hiện tính mùa vụ rõ ràng.  
  
Kết luận: Hoạt động kinh doanh mang tính thời vụ cao. Nên tập trung chiến dịch marketing vào quý IV để tối ưu doanh thu.

## 5 – Phân tích doanh thu theo quốc gia (Top 10 quốc gia) A graph with green and white lines AI-generated content may be incorrect.

Vương quốc Anh (United Kingdom) chiếm ưu thế tuyệt đối với doanh thu gấp nhiều lần so với các quốc gia khác. Các quốc gia tiếp theo như Netherlands, EIRE, Germany, France đóng góp nhỏ hơn nhiều.  
  
Kết luận: Thị trường trọng điểm là Anh Quốc, nên duy trì tập trung và khai thác sâu hơn, trong khi mở rộng dần sang châu Âu để giảm rủi ro phụ thuộc.

## 6 – Phân tích sản phẩm bán chạy nhất (Top 10 sản phẩm) A graph with different colored bars AI-generated content may be incorrect.

Các sản phẩm quà tặng nhỏ, trang trí, đồ lưu niệm chiếm tỷ trọng cao nhất, đặc biệt là PAPER CRAFT, LITTLE BIRDIE và CERAMIC STORAGE JAR. Điều này gợi ý nhóm sản phẩm chính là đồ trang trí – món quà tặng phổ biến dịp lễ.  
  
Kết luận: Danh mục sản phẩm bán chạy mang tính thẩm mỹ và quà tặng, doanh nghiệp có thể đầu tư vào thiết kế sáng tạo và đa dạng mẫu mã để thu hút khách hàng.

## 7 – Phân tích khách hàng theo tổng doanh thu (Top 10 khách hàng) A graph with different colored bars AI-generated content may be incorrect.

Một số khách hàng (ví dụ ID 14646, 18102) có tổng doanh thu rất lớn, cao hơn đáng kể so với nhóm còn lại. Các khách hàng này có thể là khách buôn hoặc đối tác phân phối.  
  
Kết luận: Nhóm khách hàng lớn (VIP) đóng vai trò quan trọng trong doanh thu. Cần có chính sách chăm sóc riêng (chiết khấu, hỗ trợ đơn hàng, ưu tiên giao hàng...) để duy trì họ lâu dài.

## Kết luận

Dữ liệu thể hiện sự phân bố lệch phải, cần xử lý outlier để đảm bảo tính ổn định khi phân tích.  
Số lượng bán ra là yếu tố chính ảnh hưởng đến doanh thu, trong khi giá sản phẩm ít ảnh hưởng.  
Doanh thu có tính mùa vụ rõ rệt, tập trung vào quý IV.  
Anh Quốc là thị trường chủ lực, nhóm sản phẩm quà tặng, trang trí dẫn đầu doanh số.  
Một nhóm nhỏ khách hàng lớn chiếm phần lớn doanh thu – cần chiến lược chăm sóc đặc biệt.