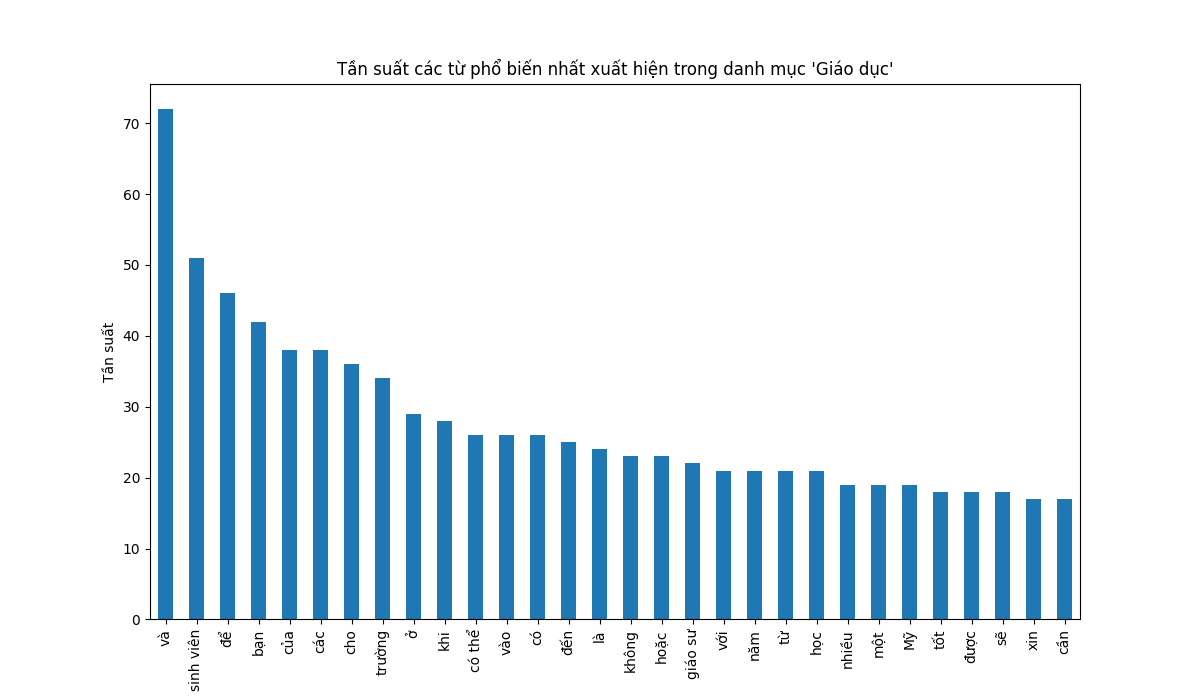
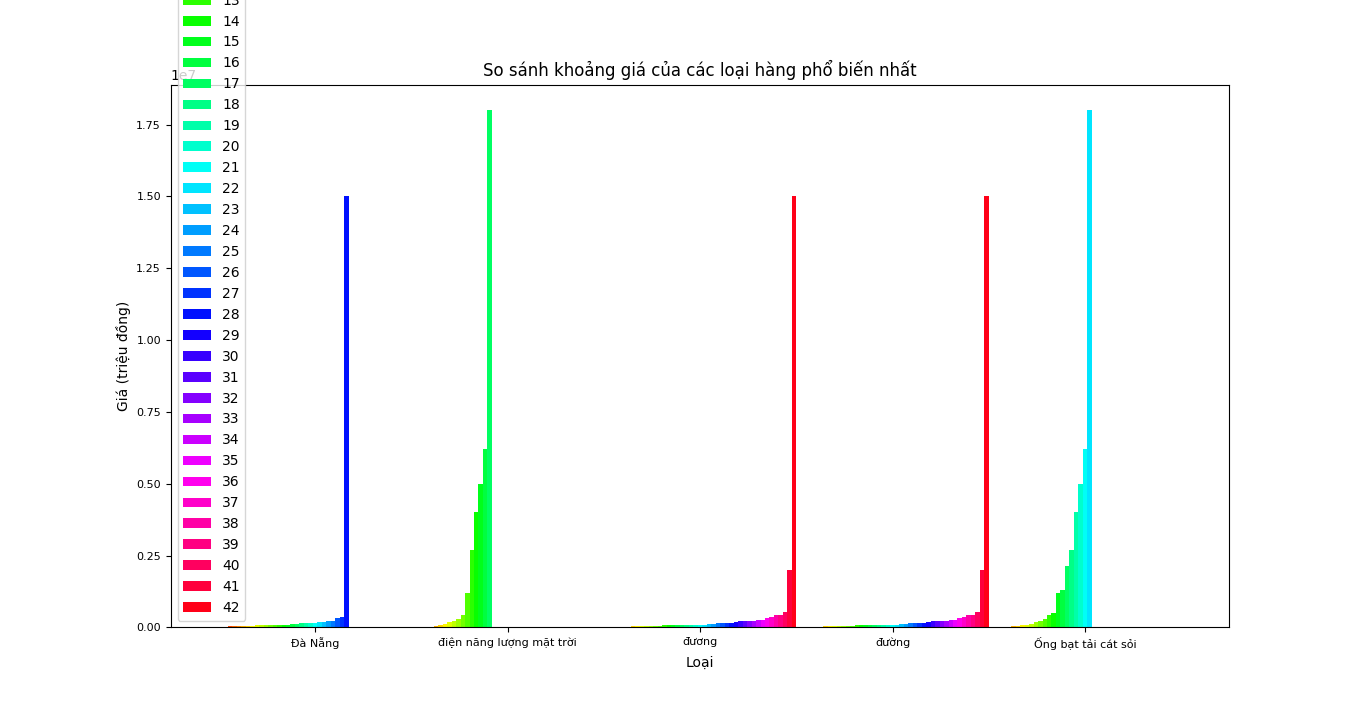
**Báo cáo thực hành tuần 4**

1. **Bộ dữ liệu vnexpress**  
   *Mục tiêu*: Phân tích văn bản nội dung của danh mục xuất hiện nhiều nhất trong bộ dữ liệu đã thu thập được.  
   *Đã xử lý*: Thống kê ra những từ/cụm từ xuất hiện nhiều nhất của danh mục ‘Giáo dục’ 
2. **Bộ dữ liệu raovat***Mục tiêu*: So sánh giá cả dựa trên phân loại sản phẩm  
   *Đã xử lý*: Lấy ra những loại hàng(dựa theo tags) phổ biến nhất trong bộ dữ liệu đã thu thập,  
   thống kê và so sánh khoảng giá dao dộng của các loại hàng đó