**CSE703014 – ĐỒ ÁN LIÊN NGÀNH**

**ĐỀ XUẤT DỰ ÁN**

**TÊN DỰ ÁN**: Thu thập, xử lý và phân tích dữ liệu từ các sàn thương mại điện tử

**NHÓM SINH VIÊN**: Nguyễn Quốc Trường - 21012325

**THỜI GIAN THỰC HIỆN**: 9/2023 – 3/2024

**PHẠM VI CÔNG VIỆC:**

Nhiệm vụ 1. Thu thập dữ liệu

Sử dụng python để thu thập dữ liệu từ sàn tiki, lazada

Dữ liệu thu về để trong các tệp theo sơ đồ: loại mặt hàng -> danh mục mặt hàng -> Ngày -> Sàn thương mại điện tử

A diagram of a flowchart

Description automatically generated

Nhiệm vụ 2. Xử lý dữ liệu

Dữ liệu thu về loại bỏ các dữ liệu trùng lặp, NaN, …

Ghép nối dữ liệu lần lượt từ loại mặt hàng, danh mục mặt hàng, ngày.

Nhiệm vụ 3. Phân tích dữ liệu

Dữ liệu sau khi đã làm sạch lấy ra tổng doanh số theo ngày, theo mặt hàng.

Trực quan hóa bằng các loại biểu đồ khác nhau.

**KẾT QUẢ DỰ ÁN:** Các kết quả dự án sẽ có

1. Tổng doanh số 5 mặt hàng của sàn tiki trong 7 ngày
2. Tổng sản phẩm có số lượt bán
3. Tổng sản phẩm đã bán
4. Phân tích thị trường theo doanh số
5. Phân tích thị trường theo số lượng ngành hàng
6. 10 loại mặt hàng bán chạy nhất
7. 10 loại mặt hàng doanh số cao nhất

|  |  |
| --- | --- |
|  | Hà Nội, Ngày: 05/3/2023  Xác nhận của GVHD |