

LẬP TRÌNH KHOA HỌC DỮ LIỆU

BIÊN BẢN HỌP LẦN 2

NHÓM 15

1. **Thời gian:** từ 11h00 đến 12h00 ngày 6/12/2024.
2. **Hình thức:** trực tiếp.
3. **Nội dung:**
 - a. **Kiểm tra giai đoạn tiền xử lý:**
 - b. **Thống nhất các câu hỏi để thực hiện phân tích.**
 - c. **Đưa ra cách thực hiện cho từng câu hỏi.**
4. **Kết quả:**
 - a. **Giai đoạn tiền xử lý:**
 - Đã thực hiện loại bỏ các cột không có giá trị trong quá trình phân tích.
 - Đã điều chỉnh kiểu dữ liệu của các cột phù hợp.
 - Không có cột dữ liệu nào bị thiếu.
 - Dữ liệu hầu hết các cột đã đồng nhất.
 - Riêng cột category (mô tả thể loại) chứa nhiều thể loại khác nhau nên sẽ tiếp tục được xử lý trong giai đoạn phân tích.
 - Phân phối giá trị của các cột bình thường.
 - b. **Câu hỏi:** *“Xu hướng lựa chọn sách theo tác giả của khách hàng như thế nào?”*

Lợi ích: nhận diện tác giả nổi bật để thực hiện chiến dịch quảng bá cũng như điều chỉnh các chiến lược với các tác giả ít được chú ý.

Cách thực hiện:

- + Tính doanh thu của từng sách.
- + Nhóm dữ liệu theo tác giả.
- + Phân tích các tác giả mang lại doanh thu cao:
 - *Vẽ biểu đồ top 10 tác giả có doanh thu cao nhất.*
 - *Vẽ biểu đồ tỷ lệ doanh thu của các tác giả trong top 10.*
- + Phân tích các tác giả có lượng sách bán chạy và không bán chạy:
 - *Vẽ biểu đồ top 10 tác giả có sách bán chạy và tỷ lệ giảm giá.*
 - *Vẽ biểu đồ top 10 tác giả có sách ít bán chạy và tỷ lệ giảm giá.*
- c. **Câu hỏi:** *“Phân tích đặc điểm và độ phổ biến của những Nhà xuất bản sách trên nền tảng Tiki”.*

Lợi ích:

- + Từ góc nhìn của Tiki: hiểu rõ mức độ yêu thích của nhà xuất bản và phát hiện được xu hướng thị trường để từ đó tập trung hợp tác với các Nhà xuất bản có tiềm năng, tối ưu chiến lược nhập hàng
- + Từ góc nhìn của người đọc: hiểu được đặc điểm của những Nhà xuất bản để yên tâm khi mua sắm và tiết kiệm thời gian chọn sách, biết được những Nhà xuất bản có chất lượng cao, khám phá được những Nhà xuất bản sách phù hợp với sở thích cá nhân.

Cách thực hiện:

- + Nhóm dữ liệu theo số lượng sách của mỗi Nhà xuất bản.
- + Tính các chỉ số bổ sung như tổng số lượng sách bán ra, điểm đánh giá trung bình, tổng lượt đánh giá của mỗi Nhà xuất bản.
- + Phân tích sự khác biệt về độ phổ biến giữa các Nhà xuất bản dựa vào các chỉ số bổ sung để có được cái nhìn tổng quan về nhiều yếu tố.
- + Tìm ra những đầu sách nổi bật của mỗi Nhà xuất bản và phân tích các đặc điểm chung của những đầu sách này như thể loại, giá cả để tìm ra những yếu tố ảnh hưởng đến sự thành công.
- + Phân tích các Nhà xuất bản có tiềm năng nhưng hiệu suất còn thấp để từ đó đề xuất những cải thiện để tận dụng hết tiềm năng của những Nhà xuất bản này.

d. Câu hỏi: “Tiki nên có chiến lược kinh doanh như thế nào đối với việc lựa chọn thể loại sách?”

Lợi ích:

- + Có cái nhìn tổng quát về thị trường các thể loại sách và hình thành chiến lược kinh doanh như có cách chiến dịch ưu đãi và khuyến mãi, nhập những thể loại được độc giả ưa chuộng,...
- + Tăng doanh thu nhờ có chiến lược hiệu quả.
- + Gia tăng lòng tin của khách hàng đối với Tiki: Cung cấp các đầu sách đúng với sở thích, giúp tạo trải nghiệm mua sắm cá nhân hóa, khiến khách hàng quay lại.
- + Hạn chế tồn kho đối với các đầu sách không phổ biến nhờ việc lựa chọn danh mục dựa trên dữ liệu phân tích.
- + Nâng cao hình ảnh thương hiệu và tăng khả năng cạnh tranh trên thị trường.

Cách thực hiện:

- + Xem xét các thể loại phổ biến dựa trên số lượng sách nhiều.
- + Tìm các thể loại sách bán chạy và xem xét mối liên hệ với mức giảm giá.
- + Xem xét giá hiện tại trung bình của từng thể loại sách bán chạy để tìm hiểu xem mức giá có mối liên hệ như thế nào đối với số lượng bán ra.
- + Điểm đánh giá và lượt đánh giá của top 10 thể loại bán chạy nhất.
- + Các sản phẩm theo từng thể loại có mức giá cao có thì có số lượng bán ra như thế nào? -> Để tìm ra những thể loại có giá cao nhưng có số lượng bán ra nhiều.