



FPT POLYTECHNIC

BÁO CÁO DỰ ÁN

**PHÂN TÍCH HÀNH VI NGƯỜI TIÊU DÙNG VÀ HIỆU QUẢ KINH
DOANH CỦA CỬA HÀNG BÁN ĐỒ THỂ THAO**

Chuyên Ngành: Xử Lý Dữ Liệu

Lớp: DP1931

Nhóm: DP06

Các thành viên:

1. Nguyễn Trường Giang - Trưởng nhóm
2. Nguyễn Văn Huỳnh - Thành viên
3. Nguyễn Công Việt Huy - Thành viên
4. Nguyễn Tuấn Đức - Thành viên
5. Đỗ Minh Hiếu - Thành viên

Giảng viên: Chu Thị Ngân

MỤC LỤC

I. GIỚI THIỆU DỰ ÁN	7
1. MÔ TẢ DỮ LIỆU	7
- Hiện trạng:	7
2. YÊU CẦU CỦA CÔNG TY	7
3. LẬP KẾ HOẠCH DỰ ÁN	8
II. PHÂN TÍCH KHÁCH HÀNG	9
1. PHÂN TÍCH YÊU CẦU	9
2. CÂU CHUYỆN DỮ LIỆU	9
3. ĐẶT VẤN ĐỀ	10
4. XÁC ĐỊNH CÂU CHUYỆN	12
5. XÁC ĐỊNH ĐỐI TƯỢNG	13
6. XÁC ĐỊNH CÂU CHUYỆN CHI TIẾT	13
7. TRÌNH BÀY DỮ LIỆU	14
III. KIẾN TRÚC HỆ THỐNG	14
1. KIẾN TRÚC	14
2. GIẢI THÍCH	15
3. GIẢI THÍCH BỘ DỮ LIỆU	15
3.1. CÁC TRƯỜNG DỮ LIỆU	15
IV. LÀM SẠCH VÀ CHUYỂN ĐỔI DỮ LIỆU	18
1. CHUẨN BỊ DỮ LIỆU	18
1.1 GIẢI PHÁP LƯU TRỮ DỮ LIỆU	18
1.2. GIẢI PHÁP PHÂN BỐ DỮ LIỆU	18
2. LÀM SẠCH DỮ LIỆU	21
2.1. CÁC VẤN ĐỀ ẢNH HƯỞNG DỮ LIỆU	21
2.2. CÁC TIÊU CHÍ ĐÁNH GIÁ CHẤT LƯỢNG DỮ LIỆU	21
2.3. CÁC BƯỚC LÀM SẠCH DỮ LIỆU	22
3. CHUYỂN ĐỔI DỮ LIỆU	23

3.1. CÁC TRƯỜNG HỢP CẦN CHUYỂN ĐỔI	23
3.2. CÁC KỸ THUẬT CHUYỂN ĐỔI	23
V. XỬ LÝ DỮ LIỆU.....	24
1. CHUẨN HÓA DỮ LIỆU	24
1.1. TRÌNH BÀY CÁC BƯỚC CHUẨN HÓA.....	24
2. MÔ HÌNH HÓA DỮ LIỆU.....	24
2.1. CÁC LOẠI MÔ HÌNH DỮ LIỆU	24
2.2. CÁC TIÊU CHÍ ĐÁNH GIÁ MÔ HÌNH DỮ LIỆU	25
3. XỬ LÝ DỮ LIỆU DAX.....	25
3.1. MEASURE.....	25
3.2. CALCULATED COLUMN.....	27
3.3. FILTER	27
VI. TRỰC QUAN DỮ LIỆU	29
1. CÁC KỸ THUẬT TRỰC QUAN HÓA	29
1.1. Biểu đồ cơ bản:.....	29
1.2. Biểu đồ nâng cao:.....	30
1.3. Kỹ thuật hiển thị phân cấp:	30
1.4. Bản đồ :	31
1.5. Bảng và KPI:	31
1.6. Biểu đồ tương tác và lọc:	32
1.7. Biểu đồ dòng thời gian :.....	32
2. CÁC NGUYÊN TẮC TRỰC QUAN HÓA	33
3. TRÌNH BÀY TẠO CÁC VISUAL CHO DỰ ÁN.....	33
VII. XÂY DỰNG BÁO CÁO.....	44
1. YÊU CẦU DASHBOARD & REPORT.....	44
2. MỤC TIÊU XÂY DỰNG BÁO CÁO	44
3. SO SÁNH DASHBOARD VÀ REPORT.....	45
4. DASHBOARD CẦN THIẾT KẾ.....	47
5. REPORT.....	50
6. BOOKMARK.....	60

VIII. KẾT LUẬN	66
1. BÁO CÁO	66
2. KHÓ KHĂN.....	66
3. THUẬN LỢI	67
4. HƯỚNG PHÁT TRIỂN	67
4.1. Mở rộng quy mô dữ liệu.....	67
4.2. Nâng cao năng lực dự báo và phân tích	67
4.3. Tối ưu hóa quản lý danh mục sản phẩm	68
4.4. Tích hợp báo cáo và trực quan hóa dữ liệu	68
5. TỔNG KẾT	68

MỤC LỤC HÌNH ẢNH

Hình ảnh 1.1: Kiến trúc hệ thống	13
Hình ảnh 2.1: Mô hình dữ liệu	23
Hình ảnh 3.1.1: Biểu đồ Tổng đơn hàng theo nhóm tuổi.....	32
Hình ảnh 3.1.2: Biểu đồ Gauge tỉ lệ 2 nhóm có con và không có con	33
Hình ảnh 3.1.3: Tỉ lệ Nam/Nữ trên tổng khách hàng.....	33
Hình ảnh 3.1.4: Tỉ lệ nghề nghiệp khách hàng	34
Hình ảnh 3.1.5: Thống kê đặt hàng theo tháng	34
Hình ảnh 3.1.6: Thống kê đặt hàng theo khu vực	35
Hình ảnh 3.2.1: Tổng doanh thu và đơn hàng theo tháng.....	35
Hình ảnh 3.2.2: Tổng doanh thu theo tháng và quý.....	36
Hình ảnh 3.2.3: Tổng doanh thu theo năm.....	36
Hình ảnh 3.2.4: Tỉ lệ Subcategory	37
Hình ảnh 3.2.5: Tỉ lệ Category.....	37
Hình ảnh 3.2.6: Tổng số lượng theo Category và Subcategory	37
Hình ảnh 3.2.7: Tổng doanh thu theo Country	38
Hình ảnh 3.3.1: Tổng doanh thu theo mùa.....	38
Hình ảnh 3.3.2: Tổng doanh thu theo khu vực.....	39
Hình ảnh 3.3.3: Tỷ lệ phần trăm doanh thu theo mùa.....	39
Hình ảnh 3.3.4: Tỷ lệ doanh thu tăng trưởng theo năm	40
Hình ảnh 3.4.1: Top 10 sản phẩm doanh thu cao nhất.....	40
Hình ảnh 3.4.2: Top 3 màu sắc ưa chuộng.....	41
Hình ảnh 3.4.3: Tỉ lệ doanh thu theo khoảng giá.....	41
Hình ảnh 3.4.4: Tổng doanh thu theo mùa.....	42
Hình ảnh 5.1: Report Tổng quan.....	48
Hình ảnh 5.2: Report Category	49
Hình ảnh 5.3: Report Customer	51

Hình ảnh 5.4: Report Product	54
Hình ảnh 5.5: Report Customer Return.....	56
Hình ảnh 6.1.1: Bookmark Slicer theo khu vực.....	58
Hình ảnh 6.1.2: Bookmark filter theo điều kiện Is Parent	58
Hình ảnh 6.1.3: Bookmark hiển thị theo giới tính	59
Hình ảnh 6.1.4: Bookmark hiển thị theo Customer Type	59
Hình ảnh 6.2.1: Bookmark hiển thị theo bản đồ	60
Hình ảnh 6.2.2: Bookmark hiển thị theo tháng	60
Hình ảnh 6.2.3: Bookmark hiển thị theo năm	61
Hình ảnh 6.3.1: Bookmark hiển thị theo Product Type	61
Hình ảnh 6.4.1: Bookmark hiển thị Return Data	62

I. GIỚI THIỆU DỰ ÁN

1. MÔ TẢ DỮ LIỆU

- Hiện trạng:
 1. Trong 3 năm từ 2020 đến 2022 , doanh thu tổng thể của công ty tăng nhẹ, nhưng một số khu vực và dòng sản phẩm không đạt hiệu suất mong đợi.
 2. Lợi nhuận có được qua từng tháng của các năm có sự biến động và chênh lệch chỉ sau 1 tháng, điều này cần phân tích rõ hơn.
 3. Tỷ lệ trả hàng có dấu hiệu tăng cao ở một số danh mục sản phẩm, cần phân tích để biết được nguyên nhân và cách khắc phục.
- Thông tin bộ dữ liệu: Bộ dữ liệu được lấy và tổng hợp trên Kaggle

Link data: <https://www.kaggle.com/datasets/deepujawarkar/adventure-works-raw-data>

2. YÊU CẦU CỦA CÔNG TY

- Về mặt dữ liệu: Cần chính xác và minh bạch đã được qua kiểm chứng.
- Quản lý và lưu trữ: Power BI Service, SQL Server , Google Drive.
- Mục tiêu: Chỉ ra các thông số từ nguồn dữ liệu thu thập được, trực quan hoá và đưa ra báo cáo dễ đọc dễ hình dung.
- Các công cụ và ngôn ngữ được sử dụng: Python, SQL Server, Power BI, Trello...

3. LẬP KẾ HOẠCH DỰ ÁN

TT	Hạng mục	Ngày bắt đầu	Ngày kết thúc	Tiến độ
1	Giới thiệu dự án	30-10-2024	01-11-2024	100%
2	Chốt đề tài dự án	01-11-2024	01-11-2024	100%
3	Lựa chọn dữ liệu	01-11-2024	01-11-2024	100%
4	Xây dựng câu chuyện dữ liệu	02-11-2024	02-11-2024	100%
5	Lưu trữ dữ liệu	02-11-2024	03-11-2024	100%
6	Làm sạch và chuyển đổi dữ liệu	02-11-2024	04-11-2024	100%
7	Thiết kế mô hình dữ liệu	02-11-2024	05-11-2024	100%
8	Thiết kế cơ sở dữ liệu	02-11-2024	06-11-2024	100%
9	Xác định yêu cầu	02-11-2024	06-11-2024	100%
10	Thiết kế khung Dashboard	07-11-2024	08-11-2024	100%
11	Thiết kế dashboard Tổng quát	09-11-2024	12-11-2024	100%
12	Thiết kế dashboard Customer	09-11-2024	12-11-2024	100%
13	Thiết kế dashboard Category	09-11-2024	12-11-2024	100%
14	Thiết kế dashboard Return	09-11-2024	12-11-2024	100%
15	Thiết kế dashboard Product	09-11-2024	12-11-2024	100%
16	Xây dựng dữ liệu dự đoán	13-11-2024	15-11-2024	100%
17	Xây dựng mô hình dự đoán	15-11-2024	20-11-2024	100%
18	Đánh giá mô hình	25-11-2024	28-11-2024	100%
19	Hoàn thiện tài liệu báo cáo	30-10-2024	30-11-2024	100%
20	Hoàn thiện slide báo cáo	30-10-2024	30-11-2024	100%

II. PHÂN TÍCH KHÁCH HÀNG

1. PHÂN TÍCH YÊU CẦU

- Dữ liệu: Lấy từ trang Kaggle
- Quản lý và lưu trữ: Quản lý bằng file Excel, Power BI, lưu trữ tại SQL Server

* Giải thích

- Dữ liệu: Đây là nguồn dữ liệu được lấy trên trang web uy tín đã được qua xác minh và có độ chính xác cao

2. CÂU CHUYỆN DỮ LIỆU

Nhà đầu tư khi xem xét một mặt hàng tiềm năng, họ quan tâm đến các yếu tố trong phân tích thị trường và cạnh tranh như sau:

A. Hiệu suất theo khu vực:

- Từ bảng **Territory Lookup** và **Product Category Sales**, doanh thu khu vực Bắc tăng trưởng tốt hơn so với khu vực Trung và Nam, nhưng vẫn thấp hơn kỳ vọng.
- Ví dụ: Khu vực Trung chiếm 35% tổng lượng trả hàng, cao nhất trong tất cả các khu vực, cho thấy vấn đề có thể nằm ở chất lượng sản phẩm hoặc dịch vụ giao hàng.

B. Phân tích danh mục sản phẩm:

- Dựa trên dữ liệu từ **Product Categories Lookup** và **Returns Data**, các sản phẩm thuộc danh mục "Điện tử" có tỷ lệ trả hàng cao nhất (12%), vượt xa mức trung bình là 5%.
- Đồng thời, doanh thu từ dòng sản phẩm "Gia dụng" giảm 10% so với cùng kỳ năm ngoái.

C. Khách hàng và hành vi mua sắm:

- Dữ liệu từ **Customer Lookup** cho thấy nhóm khách hàng thuộc "Income Level: Cao" đóng góp tới 50% tổng doanh thu nhưng chỉ chiếm 20% tổng số khách hàng.
- Tuy nhiên, nhóm "Income Level: Trung bình" đang có xu hướng giảm mua sắm. Điều này có thể liên quan đến việc giá sản phẩm không còn cạnh tranh.

D. Hiệu quả trả hàng theo thời gian:

- Từ **Calendar Lookup** và **Returns Data**, tỷ lệ trả hàng tăng mạnh vào quý 2, đặc biệt là trong tuần thứ 3 và thứ 4 của tháng. Điều này trùng khớp với thời điểm các chương trình khuyến mãi lớn được thực hiện.

3. ĐẶT VẤN ĐỀ

3.1. MÔ TẢ THỰC TRẠNG:

A. Thị trường phát triển mạnh mẽ

- Tăng trưởng doanh số rõ rệt trong các sản phẩm như xe đạp địa hình, quần áo chuyên dụng và các thiết bị hỗ trợ tập luyện tại nhà
- Các thương hiệu lớn và nhỏ đều tận dụng xu hướng chuyển dịch sang bán hàng trực tuyến để mở rộng kênh phân phối

B. Chuyển đổi hành vi người tiêu dùng

- Người tiêu dùng ưu tiên các sản phẩm bền vững, thân thiện với môi trường
- Sự gia tăng của thương mại điện tử giúp các doanh nghiệp tiếp cận được nhiều khách hàng hơn nhưng cũng tạo ra được áp lực về chi phí vận hành bất kịp xu hướng của sản phẩm

C. Sự phát triển của công nghệ

- Công nghệ đang được ứng dụng mạnh mẽ trong sản xuất và kinh doanh đồ dùng thể thao, từ thiết kế sản phẩm thông minh đến quản lý chuỗi cung ứng.

D. Xu hướng tiêu dùng gia tăng

- Hiện tại ý thức về sức khỏe của người tiêu dùng ngày càng cao thúc đẩy nhu cầu tiêu thụ các sản phẩm thể thao đặc biệt là xe đạp, quần áo thể thao, linh kiện, phụ kiện. Các hoạt động ngoài trời như chạy bộ, đạp xe tập gym đang trở thành xu hướng phổ biến dẫn đến việc gia tăng nhu cầu các sản phẩm liên quan
- Người tiêu dùng có xu hướng chuộng các sản phẩm giá hợp lý và dễ tiếp cận nhưng vẫn đảm bảo chất lượng.

3.2. DỮ LIỆU LIÊN QUAN:

- Dữ liệu thông tin khách hàng : Bảng **Customer Lookup**
- Dữ liệu thời gian: Bảng **Calendar Lookup**
- Dữ liệu về sản phẩm: Bảng **Product Lookup**
- Dữ liệu doanh số danh mục của sản phẩm: Bảng **Product Category Sales**
- Dữ liệu về danh mục sản phẩm: Bảng **Product Categories Lookup**
- Dữ liệu về thông tin danh mục sản phẩm con: Bảng **Product Subcategories Lookup**
- Dữ liệu về số lượng sản phẩm trả về: Bảng **Returns Data**
- Dữ liệu về khu vực bán hàng : Bảng **Territory Lookup**

3.3. MỤC TIÊU:

Truyền đạt dữ liệu một cách **rõ ràng, hấp dẫn** giúp người nghe hiểu được vấn đề **chính xác và nhanh chóng**.

Thể hiện **sinh động, dễ hiểu, thuyết phục**.

4. XÁC ĐỊNH CÂU CHUYỆN

Từ bộ dữ liệu trên có thể đưa ra những nhận xét và phân tích như sau:

- 1) Tần suất mua hàng:
 - Dựa trên ‘Order Quantity’ và ‘month’ để xác định được số lượng đơn hàng mua theo từng tháng
- 2) Lợi nhuận của cửa hàng qua từng năm:
 - Dựa vào ‘Total Revenue’ và ‘Total Cost’ để xác định được lợi tổng lợi nhuận và sử dụng ‘Year’ để xem lợi nhuận qua từng năm
- 3) Phân khúc khách hàng theo nhóm tuổi và giới tính:
 - Dựa vào ‘BirthDate’ và ‘Gender’ để xác định nhóm khách hàng mục tiêu
 - Loại sản phẩm mà nhóm khách hàng đó hướng đến
- 4) Ảnh hưởng của vị trí địa lý
 - Phân tích dữ liệu bán hàng theo khu vực ‘Region’ để xem khu vực nào có tiềm năng nhất
 - Loại sản phẩm nào được ưa chuộng nhất của từng khu vực
 - Các tác động ảnh hưởng đến doanh số của từng khu vực
- 5) Phân tích các sản phẩm thể thao trả lại
 - Tính tổng lượng đơn hàng hoàn về dựa vào ‘Return quantity’,
 - Xác định được loại sản phẩm nào được trả về nhiều nhất ‘Product Key’

Đưa ra các đề xuất:

- Xác định các khách hàng tiềm năng
- Đánh giá lợi nhuận thu được qua 3 năm
- Theo dõi xu hướng của sản phẩm qua thời gian để cần cải tiến

- Cải thiện và quản lý sản phẩm
- Chú trọng và trải nghiệm khách hàng
- Đánh giá lại các sản phẩm bị hoàn về
- Cải thiện chính sách về giá cả phù hợp với thu nhập người dùng

5. XÁC ĐỊNH ĐỐI TƯỢNG

● Đối tượng hướng đến:

- Những người yêu thích thể thao hoặc cần dụng cụ thể thao để cho mục đích sức khỏe
- Nhà bán lẻ và các đại lý phân phối đồ dùng thể thao
- Nhà sản xuất và cung cấp nguyên liệu bán đồ thể thao
- Các tổ chức thể dục và chăm sóc sức khỏe
- Các trường học và tổ chức giáo dục có nhu cầu mua đồ dùng thể thao cho sinh viên

6. XÁC ĐỊNH CÂU CHUYỆN CHI TIẾT

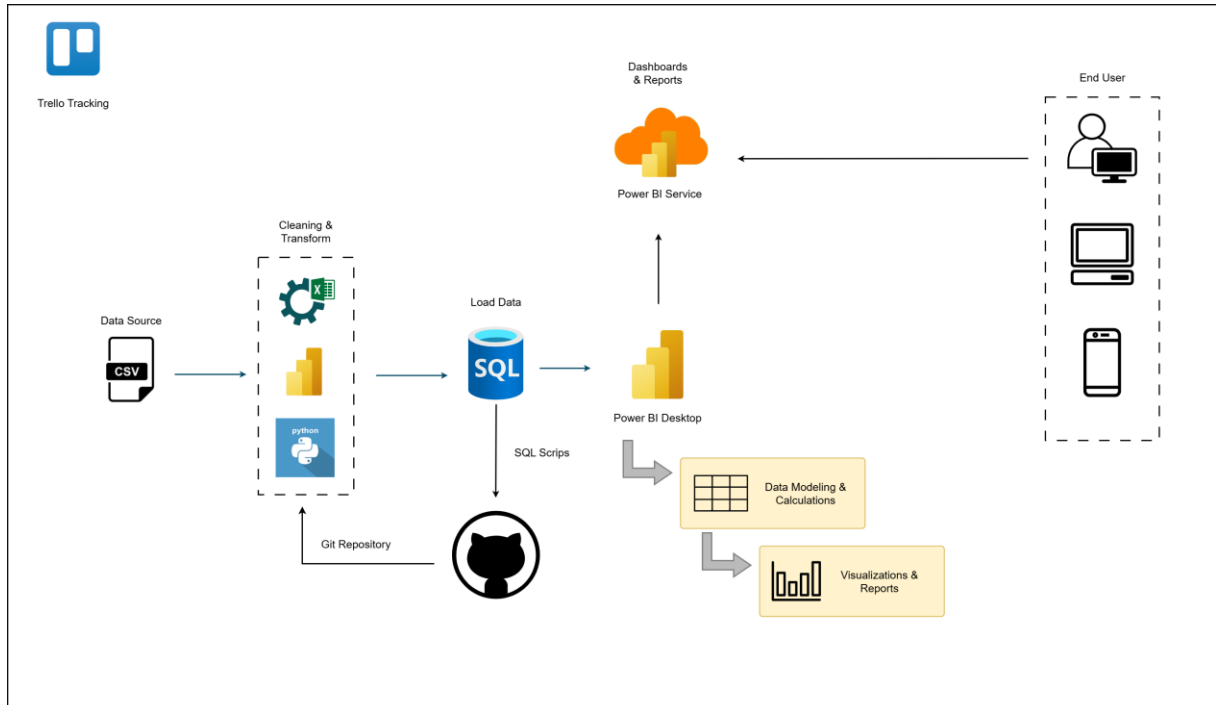
Cung cấp cho nhà kinh doanh một cái nhìn trực quan hơn về hành vi, nhu cầu của người dùng ở các mặt hàng mà cửa hàng đang bán. Từ đó có được chiến lược tối ưu lợi nhuận từ các sản phẩm đang có tại cửa hàng.

7. TRÌNH BÀY DỮ LIỆU

- Dùng hình ảnh để trình bày:
 - Biểu đồ cột: So sánh số lượng giao dịch, số lượng sản phẩm bán được
 - Biểu đồ đường: Biểu diễn lượng giao dịch qua từng giai đoạn thời gian
 - Biểu đồ tròn: Biểu diễn tỷ lệ doanh thu (của danh mục, của thành phố)
 - Biểu đồ map: Biểu diễn số lượng đơn hàng của từng thành phố

III. KIẾN TRÚC HỆ THỐNG

1. KIẾN TRÚC



Hình ảnh 1.1

2. GIẢI THÍCH

- Từ file Csv
- Làm sạch và chuẩn hóa dữ liệu qua Excel và Python
- Lưu trữ dữ liệu trên Database
- Trực quan hóa dữ liệu và tạo dashboard trên PowerBI

3. GIẢI THÍCH BỘ DỮ LIỆU

3.1. CÁC TRƯỜNG DỮ LIỆU

- Dữ liệu thông tin khách hàng : Bảng **Customer Lookup**
 - CustomerKey: Mã khách hàng
 - Prefix: Tiền tố (Mr, Mrs, Ms)
 - First Name : Tên
 - Last Name: Họ
 - BirthDate: Ngày sinh
 - MaritalStatus: Tình trạng hôn nhân
 - Gender: Giới tính
 - EmailAddress: Địa chỉ email
 - AnnualIncome: Thu nhập hàng tháng
 - TotalChild: Số con
 - EducationLevel: Bậc học
 - Occupation: Nghề nghiệp
 - HomeOwner: Chủ sở hữu nhà ở
 - Full Name: tên đầy đủ

- IsParent: có đang là cha mẹ
- IncomeLevel: Mức thu nhập
- Customer Income Level: Loại khách hàng
- Education Category: Loại tốt nghiệp
- Dữ liệu thời gian: Bảng **Calendar Lookup**
 - Date: Ngày
 - Day Name: Thứ
 - Start of Week: Ngày bắt đầu của tuần
 - Start of Month: Ngày bắt đầu của tháng
 - Start of Quarter: Ngày bắt đầu của quý
 - Month: Tháng
 - Year: Năm
- Dữ liệu về sản phẩm: Bảng **Product Lookup**
 - ProductKey: Mã sản phẩm
 - ProductSubcategoryKey: Mã danh mục con
 - ProductSKU: Mã lưu kho
 - ProductName: Tên sản phẩm
 - ModelName: Tên phiên bản
 - ProductionDiscription: Miêu tả sản phẩm
 - ProductColor: Màu sản phẩm
 - ProductStyle: Phong cách của sản phẩm
 - ProductCost: Giá gốc sản phẩm
 - ProductPrice: Giá bán sản phẩm
 - SKU Type: Loại lưu kho
 - Price Group: Phân loại giá
- Dữ liệu doanh số danh mục của sản phẩm: Bảng **Product Category Sales**
 - Date: Ngày
 - ProductCategory: Danh mục sản phẩm

- North Region: Vùng phía Bắc
- Central Region: Vùng Trung tâm
- South Region: Vùng phía Nam
- Dữ liệu về danh mục sản phẩm: Bảng **Product Categories Lookup**
 - ProductCategoryKey: Mã danh mục lớn
 - CategoryName: Tên danh mục
- Dữ liệu về thông tin danh mục sản phẩm con: Bảng **Product Subcategories Lookup**
 - ProductSubcategoryKey: Mã danh mục con
 - SubcategoryName: Tên danh mục con
 - ProductCategoryKey: Mã danh mục lớn
- Dữ liệu về số lượng sản phẩm trả về: Bảng **Returns Data**
 - Return Date: Ngày hoàn trả
 - TerritoryKey: Mã lãnh thổ
 - ProductKey: Mã sản phẩm
 - ReturnQuantity: Số lượng hoàn trả
 - CustomerKey: Mã khách hàng
- Dữ liệu về khu vực bán hàng : Bảng **Territory Lookup**
 - SaleTerritoryKey: Mã lãnh thổ bán hàng
 - Region: Khu vực
 - Country: Đất nước
 - Continent: Lục địa

IV. LÀM SẠCH VÀ CHUYỂN ĐỔI DỮ LIỆU

1. CHUẨN BỊ DỮ LIỆU

1.1 GIẢI PHÁP LƯU TRỮ DỮ LIỆU

- Tiêu chí so sánh: Chi phí và đầu tư ban đầu, Quản lý và bảo trì, Khả năng mở rộng và linh hoạt, Bảo mật và tuân thủ quy định, Hiệu suất và khả năng truy cập
- Quyết định : Sử dụng ứng dụng lưu trữ tại chỗ (on-premise) như SQL Server.

1.2. GIẢI PHÁP PHÂN BỐ DỮ LIỆU

1.2.1 Ý NGHĨA CỦA VIỆC PHÂN BỐ DỮ LIỆU

1. Tăng cường hiệu suất
2. Khả năng mở rộng dữ liệu
3. Tăng tính sẵn sàng và độ tin cậy của dữ liệu
4. Tối ưu hóa chi phí lưu trữ
5. Phân tích và khai thác dữ liệu hiệu quả
6. Giảm độ phức tạp của dữ liệu

1.2.2. TRÌNH BÀY CÁCH PHÂN BỐ DỮ LIỆU

1.2.2.1. PHÂN BỐ THEO SẢN PHẨM

- **Mục tiêu:** Phân tích sự phân bố doanh thu theo các loại sản phẩm thể thao.
- **Phân tích:**

- Xem xét các loại sản phẩm (giày thể thao, quần áo thể thao, dụng cụ thể thao, v.v.) và so sánh doanh thu của từng loại.
- Xác định sản phẩm bán chạy nhất và ít bán.
- Các sản phẩm mang lại lợi nhuận cao hoặc thấp.
- **Phương pháp:**
 - Sử dụng **biểu đồ phân phối (histogram)** để nhìn thấy phân bố doanh thu cho từng loại sản phẩm.
 - **Biểu đồ boxplot** giúp phát hiện các giá trị ngoại lai (outliers) trong doanh thu của các sản phẩm.
- **Kết quả phân tích:** Đưa ra kết luận về sản phẩm nào chiếm tỷ lệ cao trong doanh thu và các nhóm sản phẩm có xu hướng bán chạy.

1.2.2.2. PHÂN BỐ THEO KHÁCH HÀNG

- **Mục tiêu:** Xem xét phân bố khách hàng theo các đặc điểm như độ tuổi, giới tính, và khu vực.
- **Phân tích:**
 - Phân tích khách hàng theo **địa lý** (quốc gia, khu vực) để xác định các thị trường chính.
 - Phân tích theo **độ tuổi** và **giới tính** để hiểu nhóm khách hàng nào chủ yếu tiêu dùng sản phẩm thể thao.
- **Phương pháp:**
 - Dùng **biểu đồ tròn (pie chart)** để phân bố khách hàng theo quốc gia hoặc giới tính.
 - Sử dụng **biểu đồ cột (bar chart)** để so sánh số lượng khách hàng theo độ tuổi và khu vực.
- **Kết quả phân tích:** Kết luận về nhóm khách hàng chủ yếu và các phân khúc có tiềm năng phát triển.

1.2.2.3. PHÂN BỐ THEO THỜI GIAN

- **Mục tiêu:** Xem xét doanh thu và số lượng sản phẩm bán ra theo các khoảng thời gian khác nhau (tháng, quý).
- **Phân tích:**
 - Phân tích sự biến động doanh thu theo **tháng/quý/năm** để tìm ra mùa cao điểm và mùa thấp điểm.
 - Xem xét tác động của các **chương trình khuyến mãi** hoặc **sự kiện thể thao lớn** đối với doanh thu.
- **Phương pháp:**
 - Sử dụng **biểu đồ đường (line chart)** để biểu diễn sự thay đổi doanh thu theo thời gian.
 - Dùng **heatmap** để phân tích doanh thu theo ngày trong tuần hoặc theo tháng.
- **Kết quả phân tích:** Cung cấp thông tin về thời điểm nào trong năm có doanh thu cao nhất và thấp nhất.

1.2.2.4. PHÂN BỐ THEO GIÁ SẢN PHẨM

- **Mục tiêu:** Phân tích giá của các sản phẩm và mức độ tiêu thụ.
- **Phân tích:**
 - Xem xét phân bố **giá sản phẩm** để xác định mức giá phù hợp với khách hàng mục tiêu.
 - Phân tích mối quan hệ giữa giá và số lượng bán.
- **Phương pháp:**
 - **Boxplot** để xem sự phân bố giá của các sản phẩm.
 - **Scatter plot** để kiểm tra mối quan hệ giữa giá và số lượng bán.
- **Kết quả phân tích:** Kết luận về mức giá nào có xu hướng bán được nhiều và mức giá nào có thể giảm để tăng doanh thu.

1.2.2.5. PHÂN BỐ THEO LỢI NHUẬN

- **Mục tiêu:** Phân tích lợi nhuận thu được từ các sản phẩm bán ra.
- **Phân tích:**
 - Xác định các sản phẩm nào mang lại lợi nhuận cao và lợi nhuận thấp.
 - Phân tích lợi nhuận theo **danh mục sản phẩm** (ví dụ: giày thể thao có lợi nhuận cao hơn quần áo thể thao).
- **Phương pháp:**
 - **Biểu đồ cột** để so sánh lợi nhuận giữa các sản phẩm.
 - **Biểu đồ đường** để phân tích lợi nhuận theo thời gian.
- **Kết quả phân tích:** Đưa ra các khuyến nghị về chiến lược tối ưu lợi nhuận (tăng giá, giảm chi phí, v.v.).

2. LÀM SẠCH DỮ LIỆU

2.1. CÁC VẤN ĐỀ ẢNH HƯỞNG DỮ LIỆU

- Thiếu dữ liệu tại các cột
- Kiểu dữ liệu không đồng nhất
- Các dữ liệu ngoại lệ và các giá trị bất thường
- Các dữ liệu trùng lặp
- Độ dài và kích thước của dữ liệu

2.2. CÁC TIÊU CHÍ ĐÁNH GIÁ CHẤT LƯỢNG DỮ LIỆU

- Độ chính xác của dữ liệu
- Dữ liệu phải đầy đủ không được thiếu và có tính nhất quán
- Dữ liệu phải tuân thủ các quy tắc và giới hạn đã được định sẵn

- Các bản ghi trong dữ liệu phải là duy nhất không được trùng lặp
- Độ an toàn và khả năng truy xuất của dữ liệu
- Mức độ truyền đạt của dữ liệu

2.3. CÁC BƯỚC LÀM SẠCH DỮ LIỆU

2.3.1. TRÌNH BÀY CÁC BƯỚC LÀM SẠCH

- Xử lý missing values
- Xử lý dữ liệu trùng lặp
- Gộp các dữ liệu để tạo data mới
- Chuẩn hóa dữ liệu
- Xử lý các ngoại lệ
- Xử lý các dữ liệu được phân loại
- Tạo hoặc tính toán các trường dữ liệu mới
- Kiểm tra tính chính xác của dữ liệu

2.3.2. TRÌNH BÀY CÁC BƯỚC LÀM SẠCH TRONG PHẠM VI DỰ ÁN

- Xóa các cột không cần thiết
- Replace các dữ liệu trống
- Xóa các giá trị lỗi
- Xóa các ký tự đặc biệt và bỏ các từ thừa

3. CHUYỂN ĐỔI DỮ LIỆU

3.1. CÁC TRƯỜNG HỢP CẦN CHUYỂN ĐỔI

- Dữ liệu bị thiếu và không nhất quán
- Dữ liệu thời gian cần phân tích theo chu kì, xu hướng
- Dữ liệu quá lớn
- Khi làm sạch các dữ liệu văn bản
- Khi cần gộp hoặc tách các trường dữ liệu
- Khi cần tính toán các trường dữ liệu mới

3.2. CÁC KỸ THUẬT CHUYỂN ĐỔI

- Loại bỏ giá trị thiếu
- Chuẩn hóa dữ liệu
- Thay đổi kiểu dữ liệu
- Phân nhóm dữ liệu
- Chuẩn hóa văn bản
- Ghép hoặc chia cột dữ liệu
- Tính toán các trường mới
- Các biến đổi được dùng trong dự án
 - Loại bỏ giá trị thiếu
 - Chuẩn hóa văn bản
 - Thay đổi kiểu dữ liệu

V. XỬ LÝ DỮ LIỆU

1. CHUẨN HÓA DỮ LIỆU

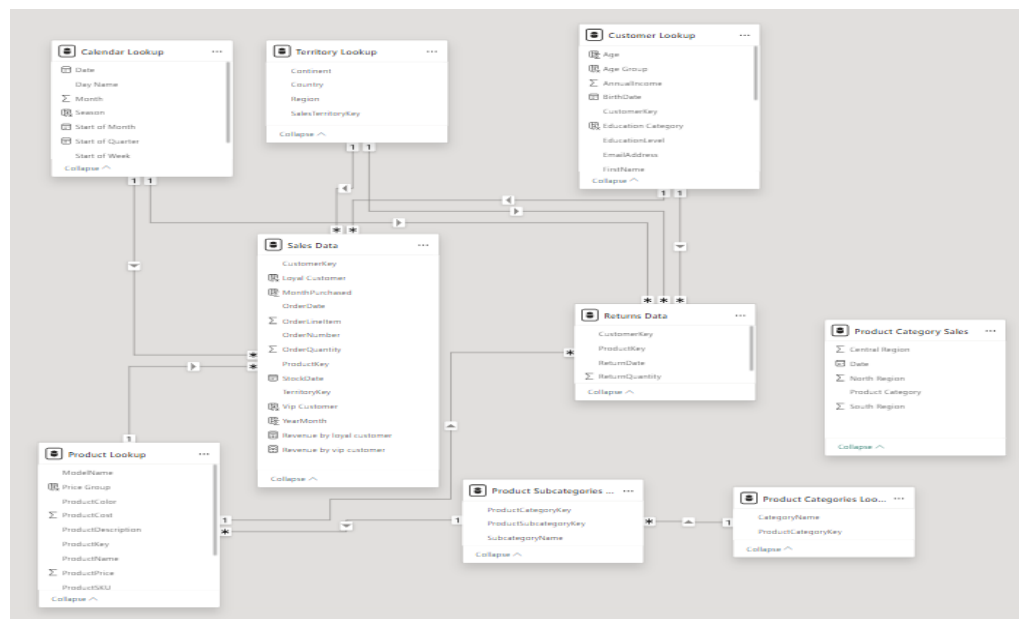
1.1. TRÌNH BÀY CÁC BƯỚC CHUẨN HÓA

- Xử lý dữ liệu trống
- Xóa dữ liệu trùng lặp, không dùng đến
- Định dạng lại dữ liệu

2. MÔ HÌNH HÓA DỮ LIỆU

2.1. CÁC LOẠI MÔ HÌNH DỮ LIỆU

- 1) Mô hình Star Schema
- 2) Mô hình Snowflake Schema
- 3) Mô hình Galaxy Schema
- 4) **Mô hình được dùng trong dự án**
 - Mô hình Galaxy Schema



Hình 2.1

2.2. CÁC TIÊU CHÍ ĐÁNH GIÁ MÔ HÌNH DỮ LIỆU

- Tính chính xác của mô hình
- Tính đầy đủ của dữ liệu
- Tính nhất quán
- Khả năng truyền đạt của mô hình
- Mô hình có tính cập nhật
- Khả năng đo lường và tính khách quan

3. XỬ LÝ DỮ LIỆU DAX

3.1. MEASURE

- Các measure được sử dụng
 - % All Returns: Tỷ lệ sản phẩm trả về
 - Revenue Growth (%): Tỷ lệ phần trăm doanh thu
 - Revenue Previous Year: Tổng doanh thu năm ngoái
 - Percentage of customer with children: Tỷ lệ khách hàng có con
 - PERCENTAGE OF PARENT: Lọc theo Is Parent
 - Previous month order: Số lượng sản phẩm order cùng kỳ tháng trước
 - Previous month return : Số lượng sản phẩm cùng kỳ trả về tháng trước
 - Previous year profit: Lợi nhuận cùng kỳ năm ngoái
 - Previous year revenue: Doanh thu cùng kỳ năm ngoái
 - Quantity return: Số lượng sản phẩm trả về
 - Quantity sold: Số lượng sản phẩm bán ra
 - Return quantity: Số lượng sản phẩm trả về

- Return rate: Tỷ lệ hoàn trả hàng
- Total customer: Tổng số lượng khách hàng
- Total profit: Tổng lợi nhuận
- Total return value: Tổng giá trị trả về
- Total revenue per customer: Doanh thu theo mỗi khách hàng
- Total return customer: Tổng số lượng khách hàng hoàn hàng
- Total revenue: Tổng doanh thu
- AVG revenue per customer: Trung bình doanh thu trên mỗi khách hàng
- Revenue by loyal customer: Doanh thu của khách hàng trung thành
- Revenue by vip customer: Doanh thu của những khách hàng vip

3.1.1. DASHBOARD TỔNG QUÁT

3.1.2. DASHBOARD CUSTOMER

- Total customer: Tổng số lượng khách hàng
- Total revenue: Tổng doanh thu
- AVG revenue per customer: Trung bình doanh thu trên mỗi khách hàng
- Percentage of customer with children: Tỷ lệ khách hàng có con
- Revenue by loyal customer: Doanh thu của khách hàng trung thành
- Revenue by vip customer: Doanh thu của những khách hàng vip

3.1.3. DASHBOARD CATEGORY & SUBCATEGORY

- Total revenue: Tổng doanh thu
- Previous Month Revenue: Doanh thu tháng trước
- Previous Year Revenue: Doanh thu năm trước
- Quantity sold: Số lượng bán ra

3.1.4. DASHBOARD PRODUCT

- Total Style: Số lượng sản phẩm Nam, Nữ, Unisex
- Color: **Dashboard Return**
- Return rate: Tỷ lệ hoàn trả
- Total return value: Doanh thu tháng trước
- Total revenue per customer: Doanh thu theo mỗi khách hàng
- Total return customer: Tổng số lượng khách hàng hoàn hàng

3.1.5. DASHBOARD TỔNG KẾT

3.2. CALCULATED COLUMN

- Các calculated column được áp dụng trong dự án
 - Với bảng Customer Lookup
 - Income Level: Phân loại mức thu nhập
 - Type customer: Phân loại khách hàng
 - Education category: Phân loại trình độ học vấn
 - Age: Tính tuổi của khách hàng
 - Age Group: Phân loại theo nhóm tuổi
 - Is Parent: Phân tích hành vi khách hàng
 - Loyal Customer: Phân loại các khách hàng trung thành
 - Vip Customer: Phân loại các khách hàng vip

3.3. FILTER

- Các filter được áp dụng trong dự án
- Với bảng Home
 - Year: Bộ lọc theo năm
 - Season: Bộ lọc theo mùa
 - Region: Bộ lọc theo khu vực

- Với bảng Customer Lookup
 - Gender: Bộ lọc theo giới tính khách hàng
 - Type Customer: Bộ lọc phân loại khách hàng
 - Continent: Bộ lọc theo khu vực
 - Is Parent: Bộ lọc so sánh giữa khách hàng có con và không có con
- Với bảng Category and Subcategory
 - Country: Bộ lọc theo đất nước
 - Time: Bộ lọc theo thời gian
 - Category: Bộ lọc theo Category
 - Subcategory: Bộ lọc theo Subcategory
- Với bảng Product:
 - Month: Bộ lọc theo Tháng
 - Year: Bộ lọc theo Năm
 - Color: Bộ lọc theo màu sắc
 - Category: Bộ lọc theo danh mục

VI. TRỰC QUAN DỮ LIỆU

1. CÁC KỸ THUẬT TRỰC QUAN HÓA

1.1. Biểu đồ cơ bản:

- **Bar Chart (Biểu đồ cột):**
 - Dùng để so sánh dữ liệu giữa các danh mục (categories).
 - Ví dụ: So sánh doanh thu theo từng sản phẩm hoặc quốc gia.
 - **Column Chart (Biểu đồ cột dọc):**
 - Hiển thị các giá trị qua danh mục, tốt cho dữ liệu thời gian hoặc dữ liệu phân nhóm.
 - **Line Chart (Biểu đồ đường):**
 - Dùng để phân tích xu hướng (trend) theo thời gian.
 - Ví dụ: Số lượng khách hàng mỗi tháng.
 - **Pie Chart (Biểu đồ tròn):**
 - Biểu thị tỷ lệ phần trăm của các phần trong tổng thể.
 - Ví dụ: Tỷ lệ doanh thu đóng góp của từng sản phẩm.
-

1.2. Biểu đồ nâng cao:

- **Stacked Bar/Column Chart (Biểu đồ cột xếp chồng):**
 - Hiện thị tổng giá trị của các danh mục, đồng thời phân chia thành các thành phần con.
 - Ví dụ: Doanh thu từng năm, chia theo từng khu vực.
 - **Area Chart (Biểu đồ vùng):**
 - Giống biểu đồ đường nhưng làm nổi bật phần diện tích dưới đường.
 - Dùng để biểu diễn xu hướng với sự nhấn mạnh vào tổng giá trị.
 - **Scatter Plot (Biểu đồ phân tán):**
 - Hiện thị mối quan hệ giữa hai biến số.
 - Ví dụ: Mối quan hệ giữa chi tiêu quảng cáo và doanh thu.
 - **Bubble Chart (Biểu đồ bong bóng):**
 - Giống Scatter Plot, nhưng thêm một biến nữa được biểu diễn qua kích thước bong bóng.
-

1.3. Kỹ thuật hiển thị phân cấp:

- **Treemap (Biểu đồ cây):**
 - Biểu diễn dữ liệu phân cấp bằng các hình chữ nhật với kích thước tương ứng giá trị.
 - Ví dụ: Tỷ lệ đóng góp doanh thu theo danh mục sản phẩm.
 - **Sunburst Chart (Biểu đồ ánh dương):**
 - Biểu diễn dữ liệu phân cấp bằng các vòng tròn đồng tâm.
-

1.4. Bản đồ :

- **Map (Bản đồ cơ bản):**
 - Hiện thị dữ liệu trên bản đồ địa lý.
 - Ví dụ: Doanh số theo khu vực hoặc quốc gia.
 - **Filled Map (Bản đồ vùng):**
 - Làm nổi bật các khu vực theo giá trị (VD: dân số, doanh thu).
-

1.5. Bảng và KPI:

- **Table (Bảng dữ liệu):**
 - Hiện thị thông tin chi tiết với các hàng và cột.
 - Thích hợp để cung cấp dữ liệu ở dạng thô.
 - **Matrix (Ma trận):**
 - Giống bảng, nhưng hỗ trợ phân cấp dữ liệu, giúp người dùng drill down chi tiết hơn.
 - **KPI (Chỉ số hiệu suất chính):**
 - Biểu thị một giá trị đo lường cụ thể với ngữ cảnh (VD: đạt/không đạt mục tiêu doanh thu).
-

1.6. Biểu đồ tương tác và lọc:

- **Slicers (Bộ lọc):**
 - Dùng để lọc dữ liệu trực tiếp trên báo cáo, giúp người dùng tập trung vào một phần dữ liệu cụ thể.
- **Chỉ báo trạng thái (Gauge/Indicator):**
 - Thể hiện mức độ hoàn thành so với mục tiêu.

1.7. Biểu đồ dòng thời gian :

- **Gantt Chart:**
 - Dùng để quản lý dự án và biểu diễn thời gian thực hiện của từng nhiệm vụ.

2. CÁC NGUYÊN TẮC TRỰC QUAN HÓA

1. Sự phù hợp và rõ ràng

- Cần đảm bảo các dashboard và report phù hợp với đối tượng được truyền tải
- Các thông tin đưa ra cần được trình bày một cách chính xác, rõ ràng

2. Đơn giản, có tính thiết kế

- Nên sử dụng các thiết kế đơn giản cho các chi tiết trên dashboard và cần chú ý vào cách chia bố cục, cách trình bày sao cho hợp lý
- Không đưa các dữ liệu vào một cách ngẫu nhiên mà các dữ liệu đó phải có mục đích cụ thể.

3. Có thông điệp kể chuyện trên dữ liệu

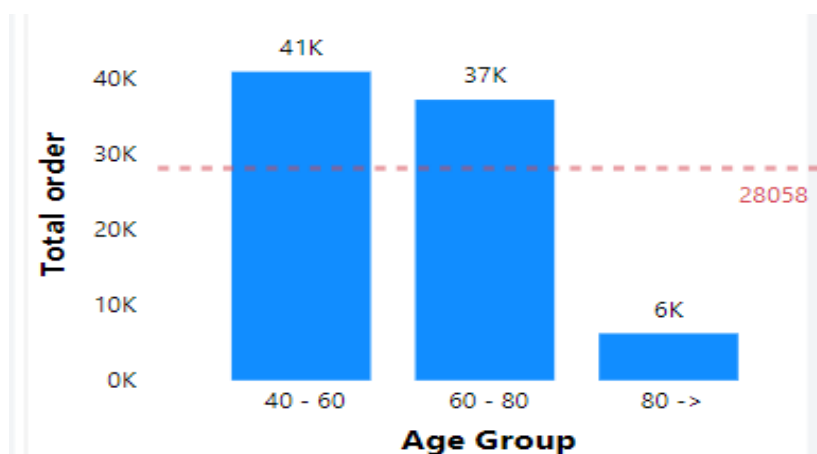
- Các dashboard/report đưa ra cần phải có thông điệp cụ thể, hỗ trợ đưa ra quyết định dựa trên số liệu, tránh mang các tính chất liệt kê
- Tránh việc làm dashboard nhưng chỉ đơn thuần là ghép các biểu đồ lại với nhau trên trang tổng thể

3. TRÌNH BÀY TẠO CÁC VISUAL CHO DỰ ÁN

3.1. TẠO VISUAL THỐNG KÊ CUSTOMER

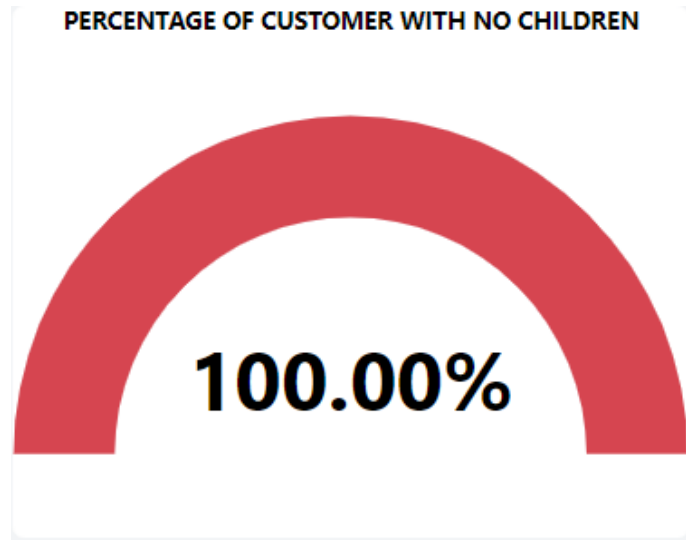
● Page Customer

1. Tạo visual thống kê tổng số đơn hàng theo nhóm tuổi



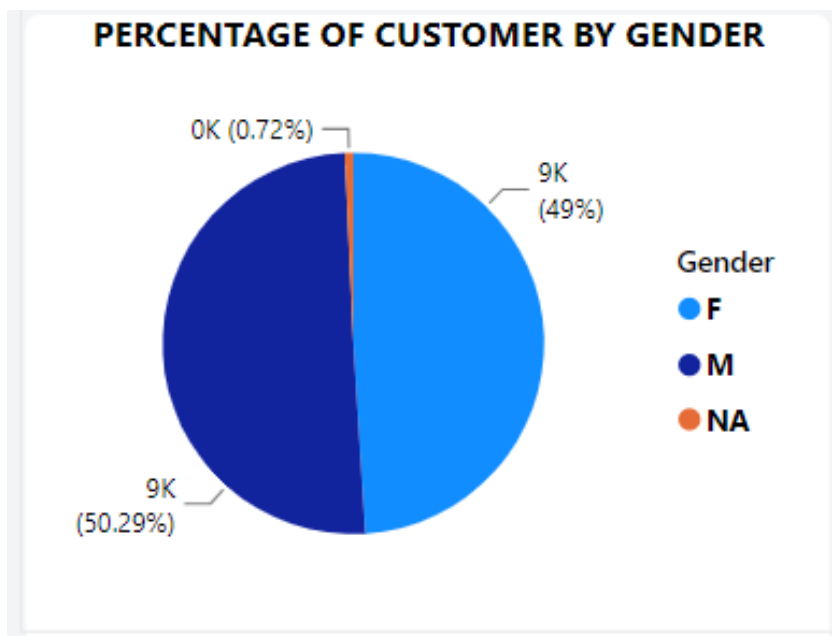
Hình ảnh 3.1.1:
Biểu đồ Tổng đơn hàng theo nhóm tuổi

2. Tạo visual thống kê tỷ lệ khách hàng có con hay không



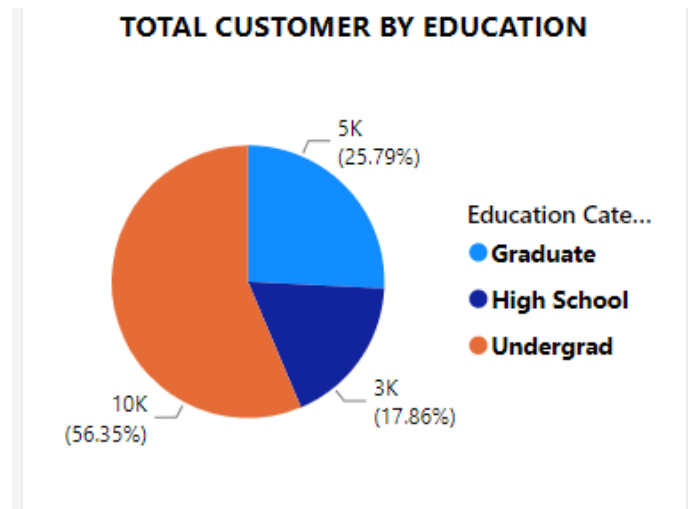
Hình ảnh 3.1.2: Biểu đồ Gauge tỉ lệ 2 nhóm có con và không có con

3. Tạo visual thống kê tỷ lệ khách hàng theo giới tính



Hình ảnh 3.1.3: Tỷ lệ Nam/Nữ trên tổng khách hàng

4. Tạo visual thống kê số lượng đơn hàng theo nghề nghiệp



Hình ảnh 3.1.4: Tỷ lệ nghề nghiệp của khách hàng

5. Tạo visual thống kê số lượng đơn hàng được đặt theo tháng



Hình ảnh 3.1.5: Thống kê số lượng đơn đặt hàng theo tháng

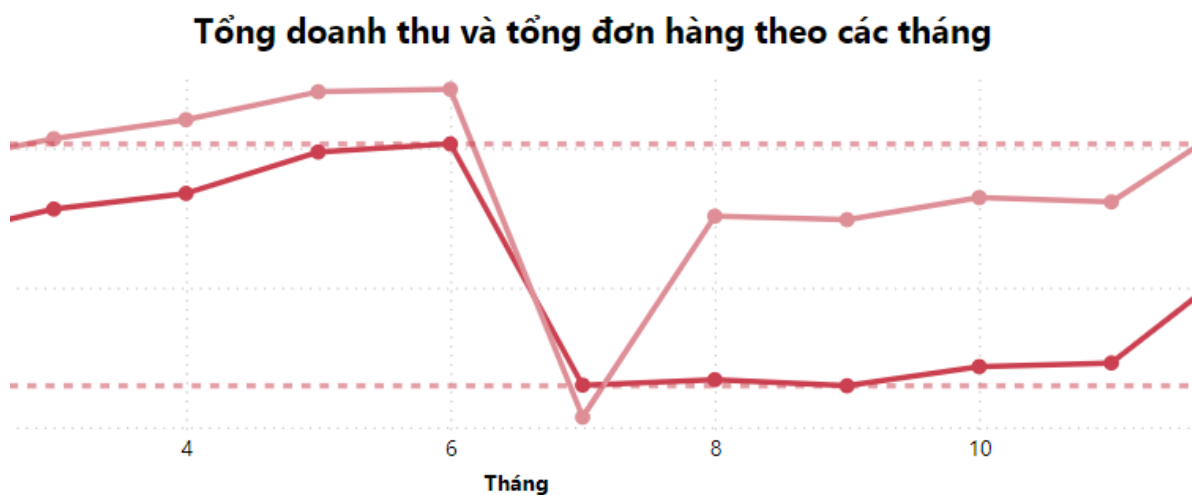
6. Tạo visual thống kê số lượng đơn hàng theo từng khu vực



Hình ảnh 3.1.6: Thống kê số lượng đơn đặt hàng theo từng khu vực

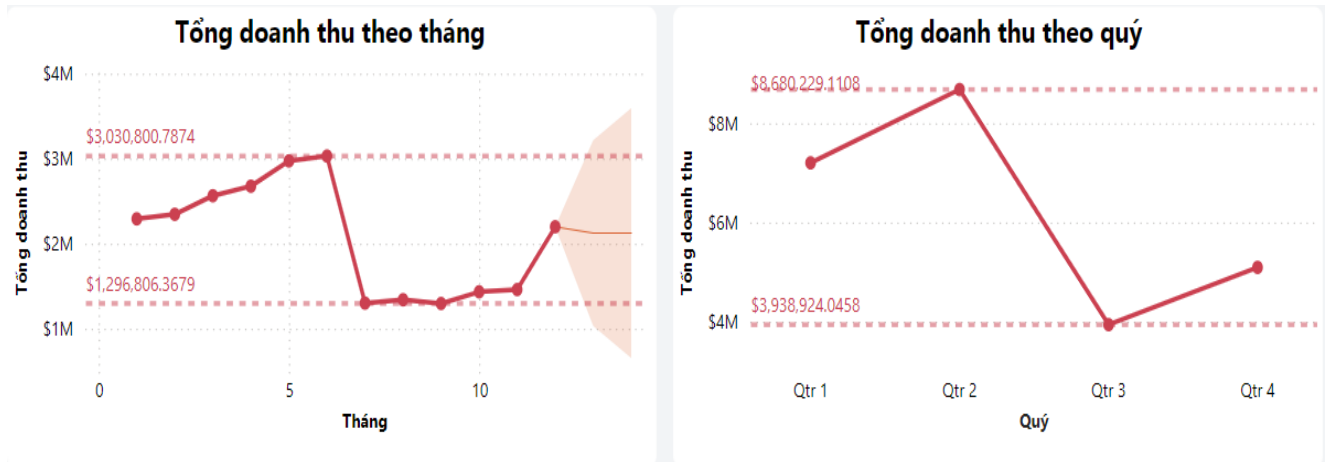
3.2. TẠO VISUAL THỐNG KÊ CATEGORY & SUBCATEGORY

- Tạo visual Doanh thu và Đơn hàng theo tháng giao dịch



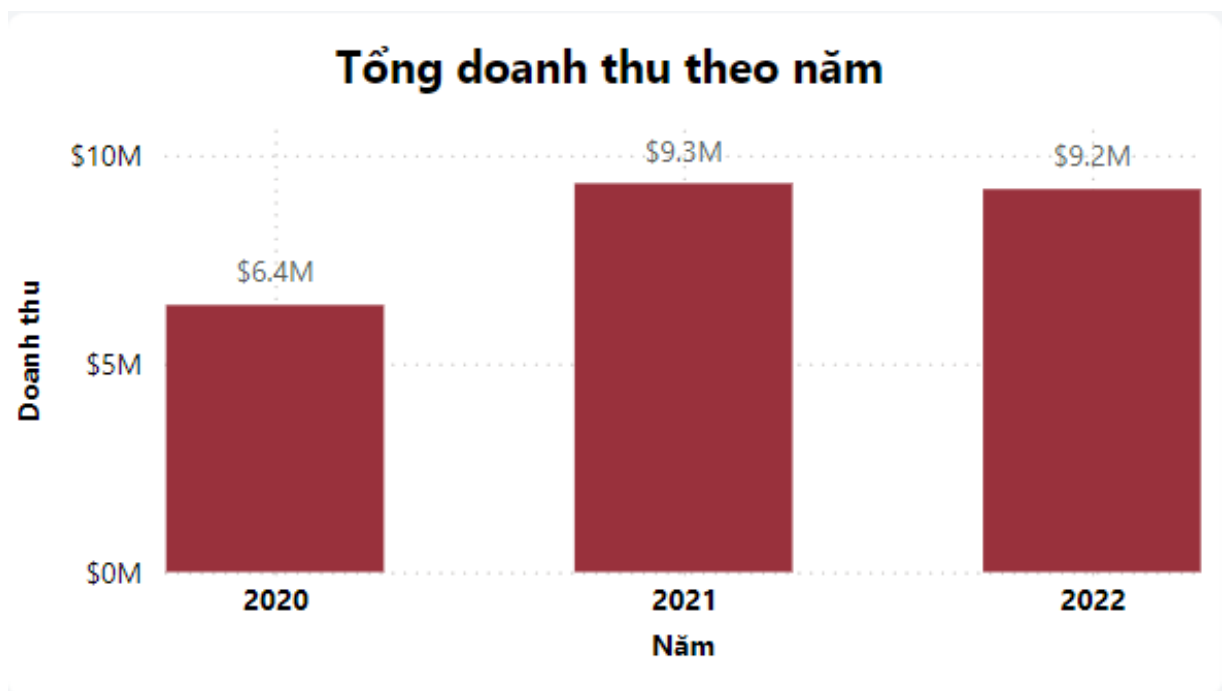
Hình ảnh 3.2.1: Tổng doanh thu và đơn hàng theo tháng

- Tạo visual thống kê Tổng Doanh thu theo tháng và quý



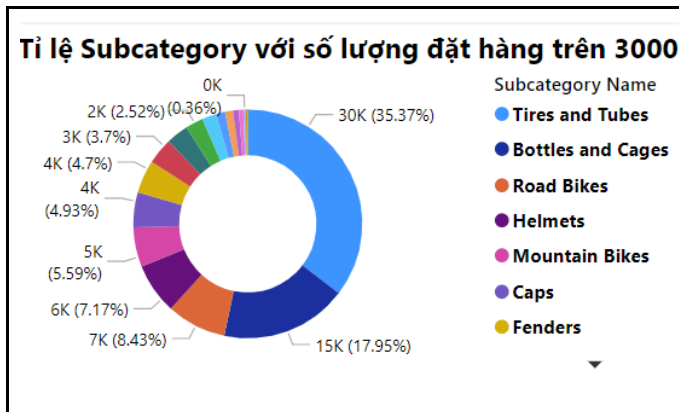
Hình ảnh 3.2.2: Thống kê tổng doanh thu theo tháng và quý

- Tạo visual thống kê Tổng Doanh thu theo Năm

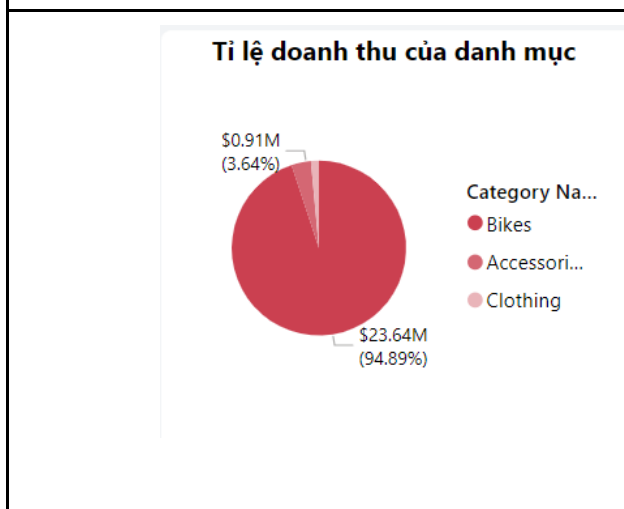


Hình ảnh 3.2.3: Tổng doanh thu theo Năm

- Tạo visual thông kê tỷ lệ Category và Subcategory theo Doanh thu

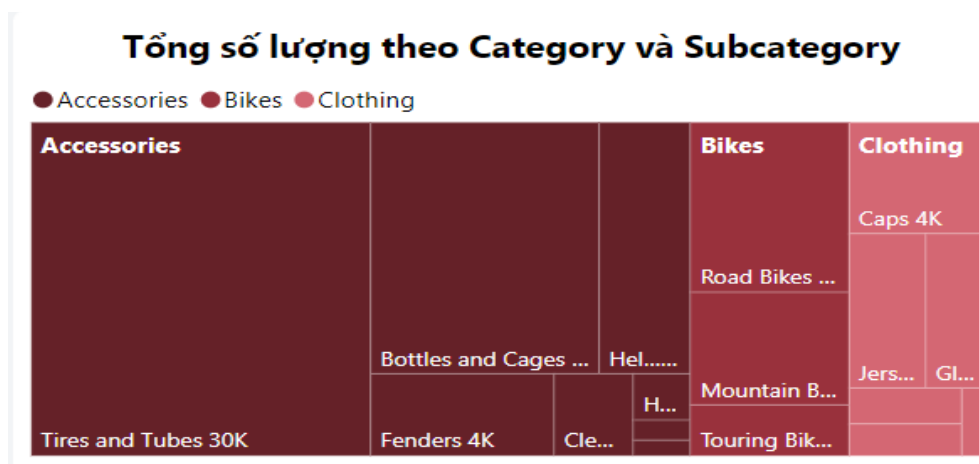


Hình ảnh 3.2.4: Tỉ lệ Subcategory



Hình ảnh 3.2.5: Tỉ lệ Category

- Tổng số lượng Category và Subcategory



Hình ảnh 3.2.6: Tổng số lượng theo Category và Subcategory

- **Tạo visual Map theo Country**

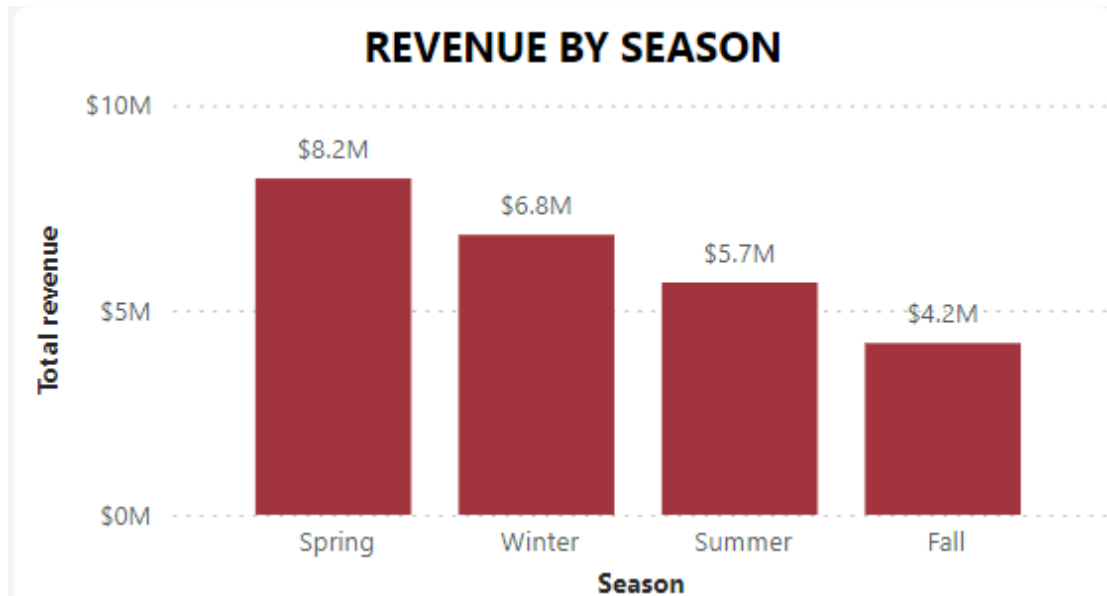


Hình ảnh 3.2.7: Tổng doanh thu theo quốc gia

3.3. TẠO VISUAL TỔNG QUÁT

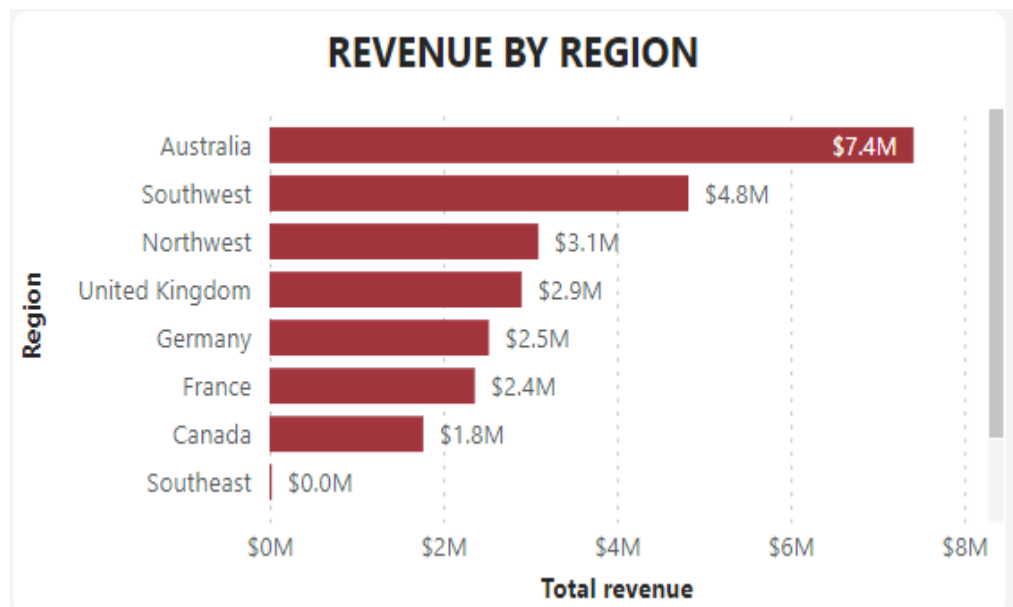
- **Page Home**

1. Tạo visual thống kê doanh thu theo mùa



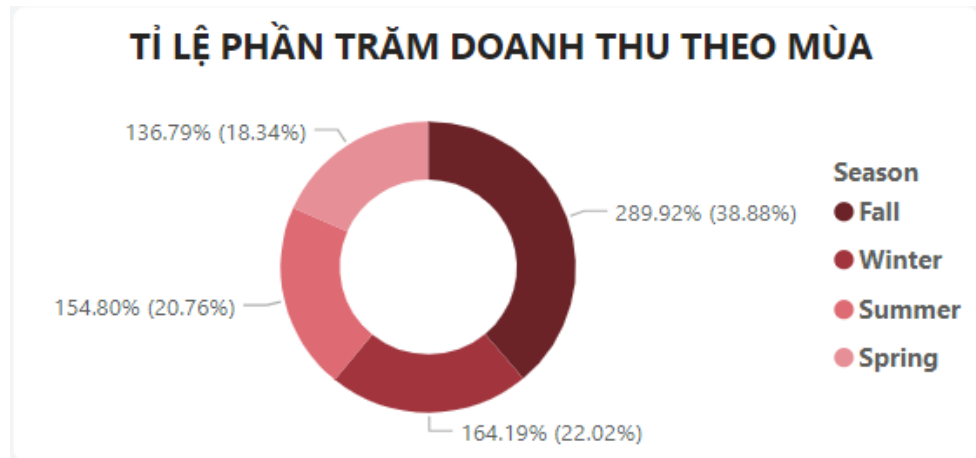
Hình ảnh 3.3.1: Tổng doanh thu theo mùa

2. Tạo visual thống kê doanh thu theo khu vực



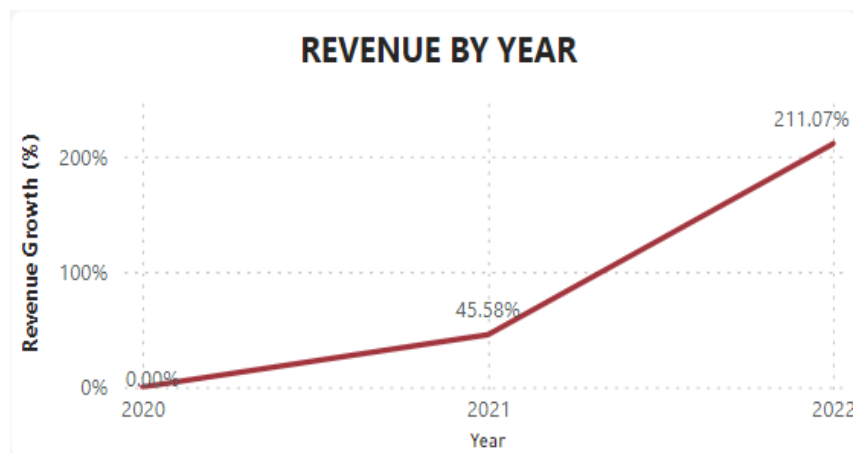
Hình ảnh 3.3.2: Tổng doanh thu theo khu vực

3. Tạo visual thống kê tỷ lệ phần trăm doanh thu theo mùa



Hình ảnh 3.3.3: Tỷ lệ phần trăm doanh thu theo mùa

4. Tạo visual thống kê mức độ tăng trưởng doanh thu theo năm

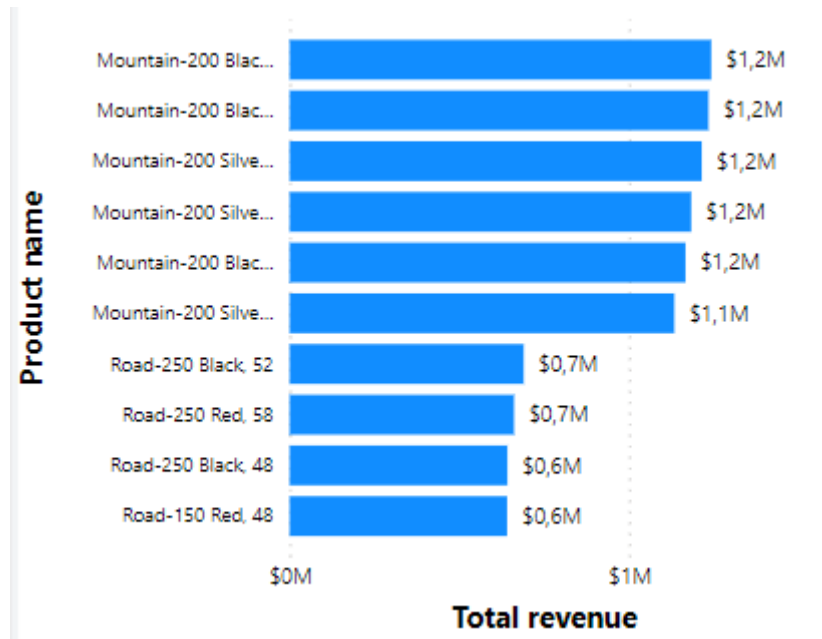


Hình ảnh 3.3.4: Tỷ lệ tăng trưởng doanh thu theo năm

3.4. TẠO VISUAL THỐNG KÊ PRODUCT

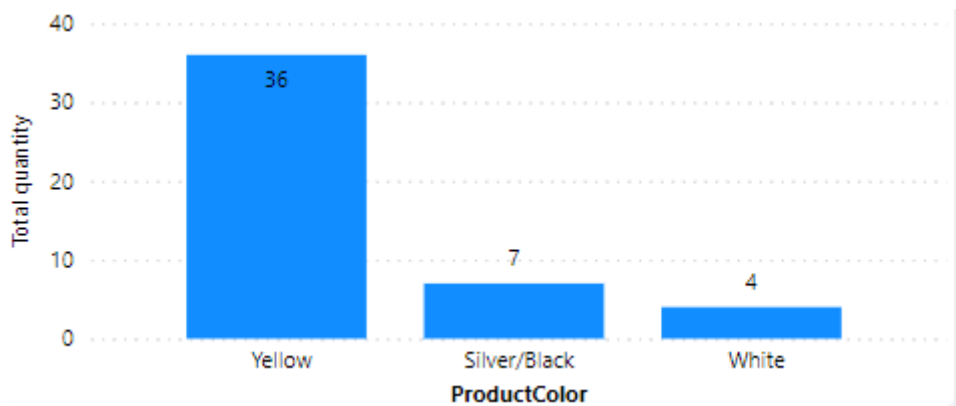
● Page Product

1. Visual thống kê top 10 sản phẩm có doanh thu cao nhất



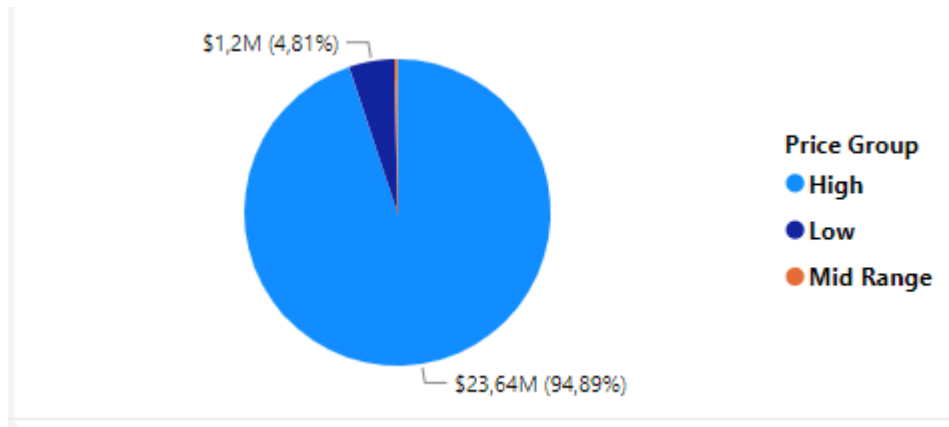
Hình ảnh 3.4.1

2. Visual thống kê top 3 màu sắc được ưa chuộng nhất



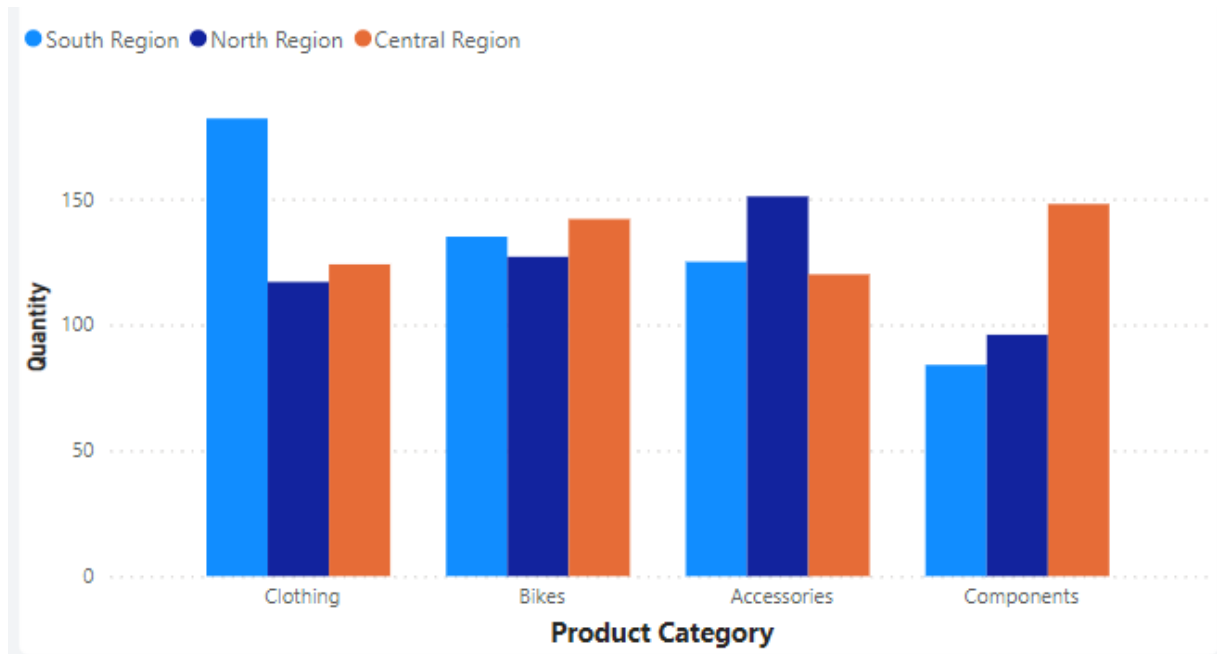
Hình ảnh 3.4.2

3. Visual thể hiện tỷ lệ doanh thu theo khoảng giá



Hình ảnh 3.4.3

4. Visual so sánh danh mục sản phẩm theo khu vực



Hình ảnh 3.4.4

VII. XÂY DỰNG BÁO CÁO

1. YÊU CẦU DASHBOARD & REPORT

- **Dashboard:**
 - Thiết kế đúng theo yêu cầu
 - Tạo vị trí visual trực quan rõ ràng
 - Thiết kế gọn gàng
 - Tương tác giữa các slicer, filter
- **Report:**
 - Trình bày rõ ràng
 - Các trang dashboard có sự liên kết với các nút và bookmark
 - Các biểu đồ thay đổi dữ liệu hiển thị khi sử dụng slicer
 - Có khả năng chia sẻ (PowerBi)

2. MỤC TIÊU XÂY DỰNG BÁO CÁO

Dashboard:

- Cung cấp cái nhìn tổng quan, theo dõi các chỉ số và thông tin quan trọng một cách trực quan.
- Tập trung vào việc **giám sát hiệu quả** và **ra quyết định nhanh chóng** dựa trên các chỉ số và xu hướng hiện tại.

Report:

- Cung cấp một cách phân tích chi tiết và sâu hơn về dữ liệu.
- Hỗ trợ các báo cáo chi tiết, phân tích lịch sử, so sánh dữ liệu, và khám phá các mẫu dữ liệu phức tạp.

3. SO SÁNH DASHBOARD VÀ REPORT

Item	Dashboard	Report
Pages	Chỉ một trang.	Có thể tạo một hoặc nhiều trang.
Data sources	Tạo từ nhiều dataset hoặc reports.	Tạo từ một dataset.
Visualization	Xây dựng thông tin chi tiết về dữ liệu bằng cách sử dụng biểu đồ, hình ảnh sinh động, lưu đồ.	Xây dựng trang tổng hợp thông tin, cung cấp góc nhìn tổng quan.
Available in Power BI Desktop	Không thể tạo.	Có thể tạo và xem.

Filters and Slicers	Không thể dung, bị giới hạn trên trang đơn.	Có thể dùng các loại filter, highlight và slicer khác nhau.
User Interactivity	Cho phép pin visuals từ report và datasets trên canvas đơn, làm cho nó đơn giản để nhóm những gì cần thiết cho người dùng.	Tập trung hơn vào khả năng trực quan hóa, áp dụng các phép biến đổi trên một tập dữ liệu.
Favourite	Có thể gán nhiều dashboards	Có thể gán nhiều dashboards
Q&A feature	Có	Có thể phân quyền chỉnh sửa report.
Alerts	Gửi email cảnh báo khi đáp ứng điều kiện, tiêu chí cụ thể hoặc vượt qua giới hạn.	Không thể
Subscribe	Có	Có
See underlying dataset tables and fields	Không thể xem dataset nhưng có thể xuất dữ liệu này.	Có thể xem dữ liệu trong tab dữ liệu.
Purpose	Được sử dụng để giám sát cấp cao, thường theo thời gian thực hoặc gần thời gian thực, cung cấp cái nhìn tổng hợp về hiệu quả kinh doanh.	Được sử dụng để phân tích chuyên sâu và khám phá dữ liệu nhằm trả lời các câu hỏi kinh doanh phức tạp.

4. DASHBOARD CẦN THIẾT KẾ

- Tổng quát
- Customer
- Category
- Customer Return
- Product

Lý do chọn các dashboard này:

- **Tổng quát:**
 - Tạo dashboard Tổng quát để nhìn được các thông số tổng quát của dữ liệu: Doanh thu, Lợi nhuận, Số lượng bán, Lượng khách hàng,...
 - Biết được các chỉ số tổng quát mà cửa hàng có được, để biết được cửa hàng đang có các thuận lợi/khó khăn nào.
- **Customer:**
 - Tạo dashboard Customer để theo dõi và phân tích hành vi khách hàng như: Tần suất mua hàng, nhóm tuổi mua hàng và các chỉ số khác, nếu có sự thay đổi đột ngột trong các chỉ số thì cửa hàng có thể phản ứng kịp thời và đưa ra được giải pháp
 - Giúp cửa hàng có thể cá nhân hóa dịch vụ và cung cấp các chương trình khuyến mãi phù hợp, phát hiện các vấn đề trong quá trình tương tác với khách hàng từ đó cải thiện trải nghiệm của khách hàng
 - Theo dõi được hiệu quả của chiến dịch marketing, theo dõi các chỉ số như: Tổng doanh thu mang lại, doanh thu theo các nhóm tuổi,

tổng đơn hàng theo từng tháng để đánh giá mức độ hiệu quả của chiến dịch

- **Category & Subcategory:**

- Tạo dashboard Category & Subcategory để theo dõi và phân tích dựa trên Category và Subcategory như: Số lượng bán hàng, Doanh thu theo các Danh mục, nếu có sự thay đổi đột ngột trong các chỉ số thì cửa hàng có thể phản ứng kịp thời và đưa ra được giải pháp
- Giúp cửa hàng có thể thống kê được các danh mục có ấn tượng với người dùng, để có các chiến lược kinh doanh hiệu quả. Theo dõi doanh thu của các danh mục cũng cho biết được tổng doanh thu của các tháng, các quý, các năm đến từ danh mục nào, giúp tối ưu chi phí quảng cáo, chi phí nhập hàng.
- Theo dõi được độ phủ của các mặt hàng tại các quốc gia đang có cửa hàng bán, giúp nhận định rõ về các nhu cầu của người tiêu dùng, kết hợp với bảng Customer để có được các chiến dịch truyền thông, marketing hiệu quả hơn đến từng đối tượng khách hàng và từng mục đích của khách hàng.

- **Customer Return:**

- Tạo dashboard Customer Return để theo dõi và phân tích dựa trên số lượng đơn hàng trả về và các thông số liên quan như: Số lượng khách trả về, các mẫu mã được trả về, các vấn đề liên quan đến sản phẩm,...
- Giúp cửa hàng có thể thống kê được các sản phẩm được khách hàng trả về nhiều nhất, từ đó tìm ra nguyên nhân dẫn đến việc khách hàng trả lại các mặt hàng đã mua.

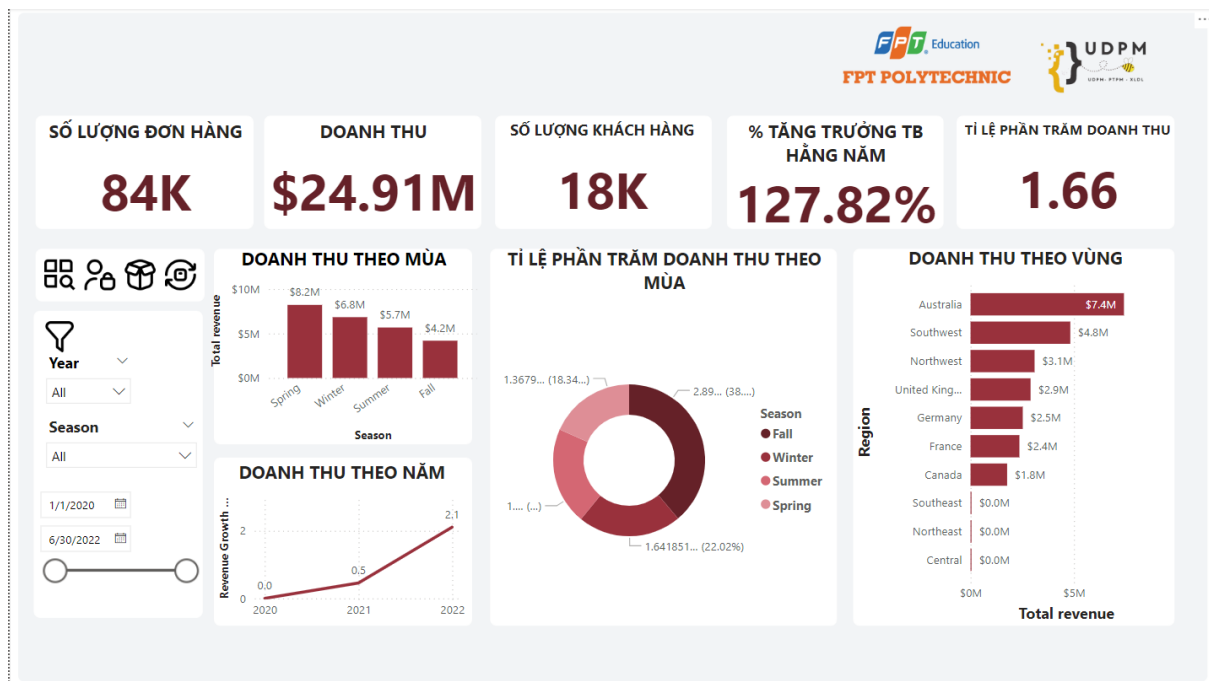
- Theo dõi theo vị trí địa lý để nắm được khu vực nào đang có số lượng trả về nhiều, từ đó tìm ra nguyên nhân liên quan.

- **Product:**

- Việc tạo dashboard cho bảng Product là một phần quan trọng trong phân tích dữ liệu và quản lý thông tin sản phẩm, bởi nó giúp cửa hàng trực quan hóa và hiểu rõ hơn về hiệu suất sản phẩm trong hệ thống
- Tạo dashboard Product là để doanh nghiệp quản lý sản phẩm hiệu quả hơn, đưa ra các quyết định kinh doanh dựa trên dữ liệu, và tối ưu hóa lợi nhuận.
- Mỗi sản phẩm có đóng góp khác nhau vào doanh thu, lợi nhuận, hoặc số lượng bán hàng. Dashboard cho phép bạn dễ dàng theo dõi những sản phẩm nào đang **bán chạy**, mang lại **lợi nhuận cao**, hoặc ngược lại, để đưa ra quyết định kịp thời.

5. REPORT

5.1. TẠO REPORT TỔNG QUAN



Hình ảnh 5.1

Giải thích:

- Thanh điều hướng: lần lượt: Category, Customers, Product, Customer Return
- Visual Card: Hiển thị thông số tổng thu nhận được: Tổng doanh thu, Tổng khách hàng, Tổng số lượng đơn hàng, Phần trăm tăng trưởng TB, Tỉ lệ phần trăm doanh thu
- Visual Bar chart: Hiển thị Doanh thu theo các vùng (Trục X: Tổng doanh thu, Trục Y: Các vùng kinh doanh)
- Visual Column Chart: Doanh thu theo các mùa (Trục X: Mùa, Trục Y: Doanh thu)
- Visual Pie Chart: Hiển thị Tỉ lệ phần trăm doanh thu theo mùa

- Bộ lọc:

Giải thích:

YEAR

All

Lọc theo năm

SEASON

All

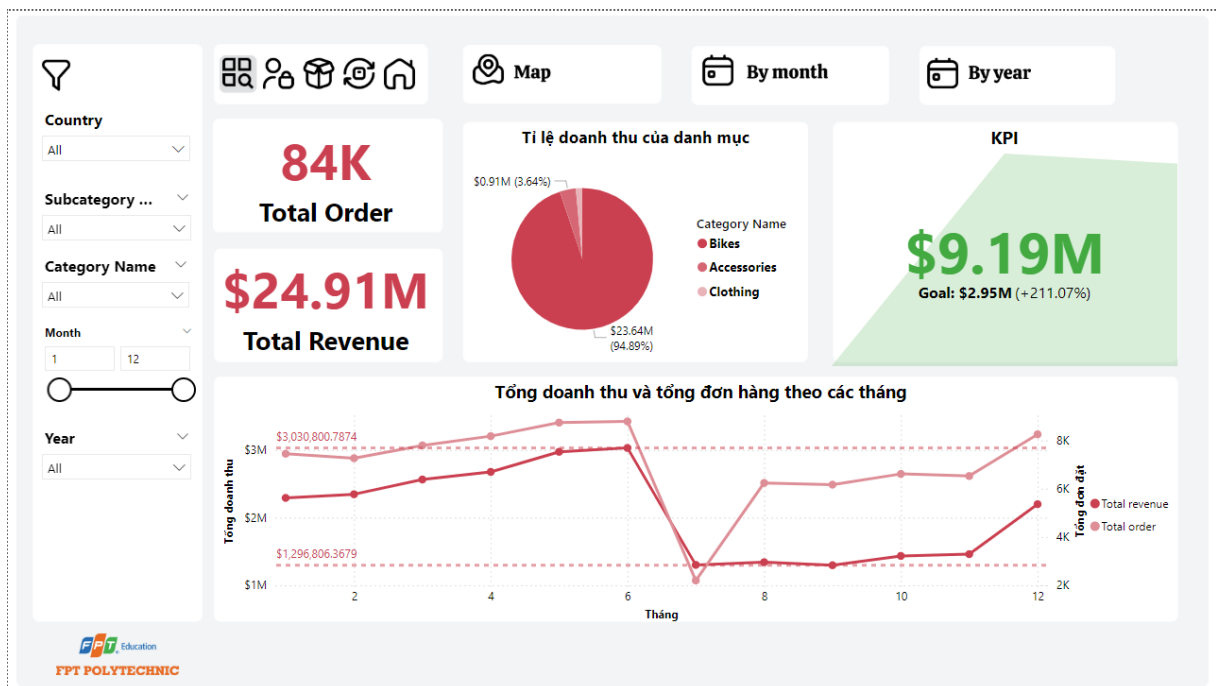
Lọc theo mùa(Xuân, Hạ, Thu, Đông)

REGION

All

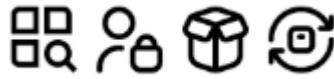
Lọc theo khu vực

5.2. TẠO REPORT CATEGORY



Hình ảnh 5.2

Giải thích:



- Thanh điều hướng: lần lượt:

Category, Customers, Product, Customer Return

- Các bookmark lọc hiển thị



lần lượt: Hiển thị bản đồ, Hiển thị theo tháng, Hiển thị theo năm

- Visual Card: Hiển thị thông số tổng thu nhận được: Tổng đơn hàng, Tổng doanh thu.
- Visual KPI: Hiển thị KPI theo năm
- Visual Line chart: Hiển thị Doanh thu và Tổng đơn hàng theo thời gian
- Bộ lọc:

Giải thích:

Country

Lọc theo quốc gia

Category Name

Lọc theo Category

Subcategory Name

Lọc theo Subcategory khi đã lọc theo Category

Month

1 12

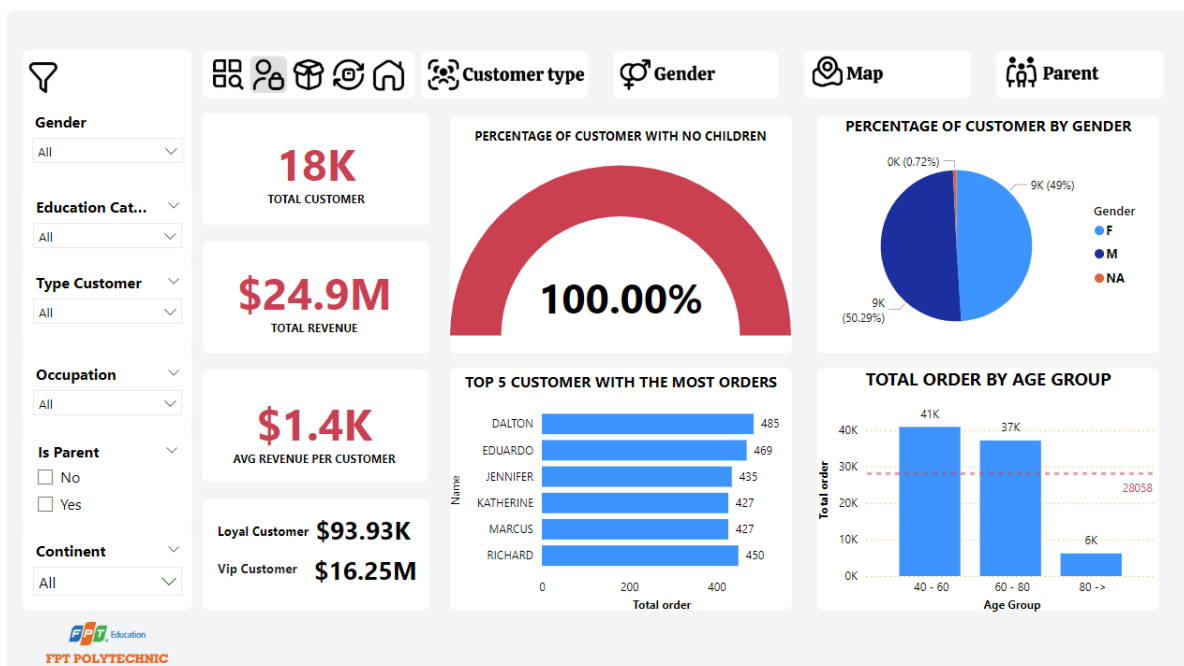
○ — ○

Year

All ▼

Lọc theo thời gian các năm và khoảng thời gian các tháng

5.3. TẠO REPORT CUSTOMER



Hình ảnh 5.3


Giải thích:

- Thanh điều hướng: lần lượt: Category, Customers, Product, Customer Return
- Các bookmark lọc hiển thị lần lượt

Customer type Gender Map Parent

Hiển thị loại khách hàng, Hiển thị theo giới tính, Hiển thị theo bản đồ,
Hiển thị theo trạng thái sinh con

- Visual Card: Hiển thị thông số thu nhận được: Tổng khách hàng, Tổng doanh thu, Doanh thu TB dựa trên khách hàng, Số khách hàng Loyal, Số khách hàng Vip
- Visual Gauge Chart: Hiển thị tỉ lệ giữa 2 thành phần: Có con và Không có con
- Visual Pie chart: Hiển thị tỉ lệ giữa 2 giới tính
- Visual Bar Chart: Hiển thị top 5 khách hàng có nhiều đơn nhất
- Visual Column Chart: Hiển thị tổng đơn theo các nhóm tuổi
- Bộ lọc:



Gender

All

Education Cat...

All

Type Customer

All

Occupation

All

Is Parent

☐ No

☐ Yes

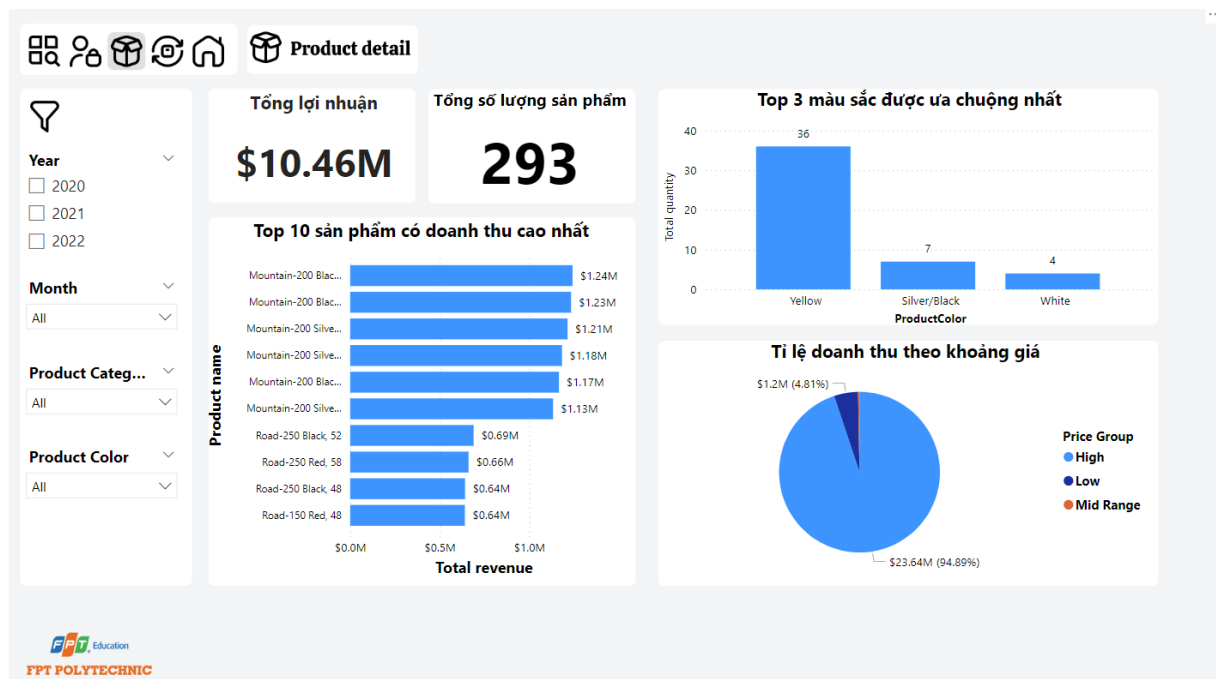
Continent

All

Giải thích:



- Gender: Lọc theo giới tính
- Education Category: Lọc theo trình độ giáo dục
- Type Customer: Lọc theo loại khách hàng
- Occupation: Lọc theo nghề nghiệp của khách hàng
- Is Parent: Bộ lọc Yes/No, thể hiện các khách hàng có con hay không
- Continent: Lọc theo lục địa mà khách hàng đang sinh sống

5.4. TẠO REPORT PRODUCT




Hình ảnh 5.4:

Giải thích:

- Thanh điều hướng:  lần lượt: Category, Customers, Product, Customer Return
- Các bookmark lọc hiển thị  sản phẩm chi tiết
- Visual Card: Hiển thị thông số thu nhận được: Tổng lợi nhuận, Tổng số lượng sản phẩm, Số lượng sản phẩm theo Style, Tỉ lệ 1 sản phẩm so với tổng sản phẩm.
- Visual Pie chart: Hiển thị tỷ lệ doanh thu theo khoảng giá
- Visual Bar Chart: Hiển thị top 10 sản phẩm có doanh thu cao nhất
- Visual Column Chart: Hiển thị top 3 màu sắc được ưa chuộng nhất và so sánh danh mục sản phẩm giữa 3 miền Bắc - Trung - Nam

- Bộ lọc:




Year


☐ 2020


☐ 2021


☐ 2022


Month

All 

Product Categ... 

All 

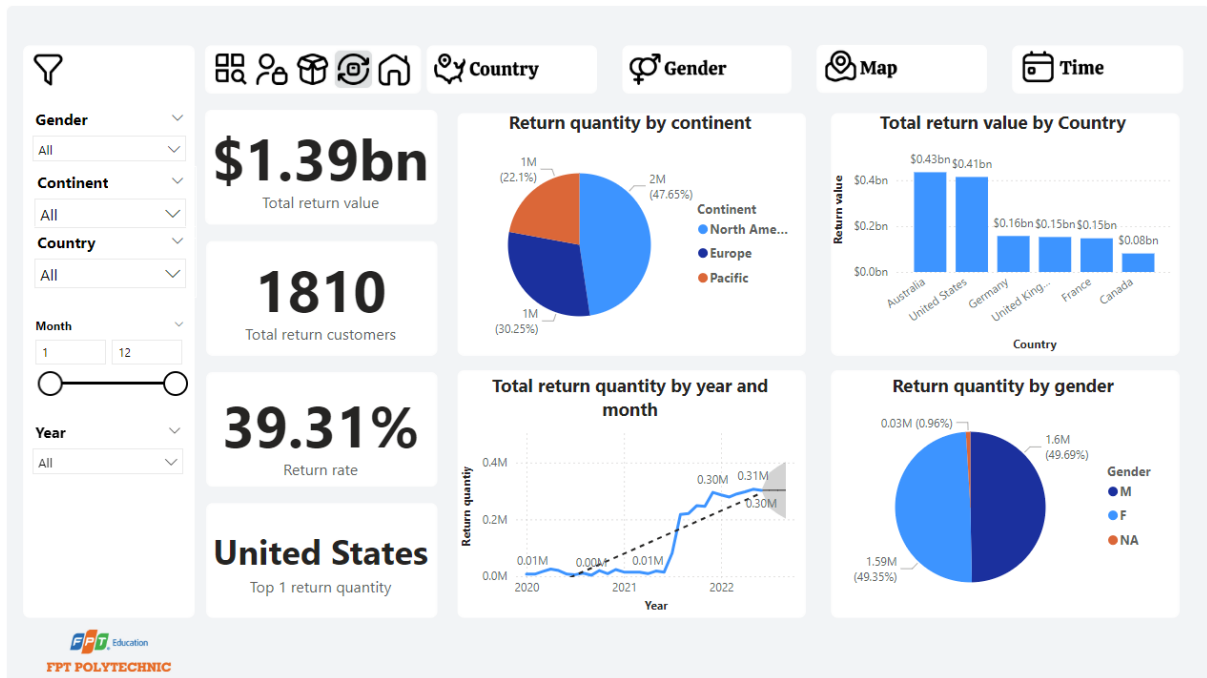
Product Color 

All 

Giải thích:

- Year: Lọc theo thời gian (Năm)
- Month: Lọc theo thời gian(Tháng)
- Product Category: Lọc theo danh mục sản phẩm
- Product Color: Lọc theo màu của sản phẩm

5.5. TẠO REPORT CUSTOMER RETURN



Hình ảnh 5.5:

Giải thích:

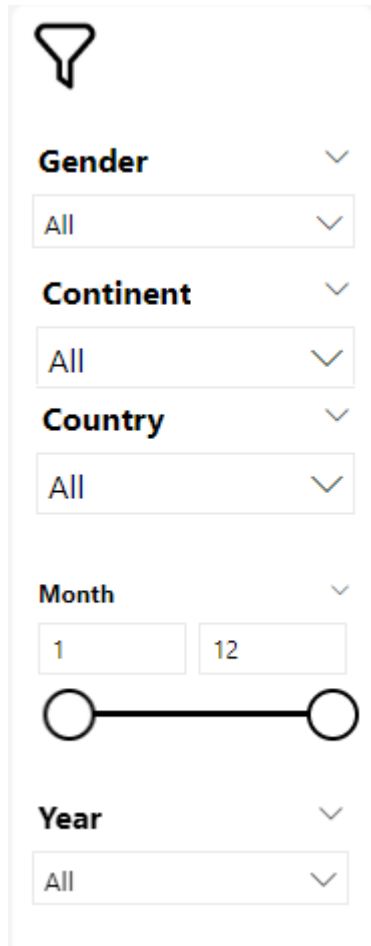
- Thanh điều hướng: lần lượt: Category, Customers, Product, Customer Return
- Các bookmark lọc hiển thị lần lượt



Hiển thị theo quốc gia, Hiển thị theo giới tính, Hiển thị theo bản đồ, Hiển thị theo thời gian

- Visual Card: Hiển thị thông số thu nhận được: Tổng khách hàng trả hàng, Tổng giá trị trả hàng, Phần trăm trả hàng
- Visual Pie chart: Hiển thị tỉ lệ giữa các lục địa, Hiển thị tỉ lệ giới tính trả hàng
- Visual Line Chart: Hiển thị tổng lượng trả hàng theo tháng và năm

- Visual Column Chart: Hiển thị tổng đơn hàng trả theo quốc gia
- Bộ lọc:



The filter panel is located on the left side of the chart area. It contains a funnel icon at the top. Below it are several filter sections: 'Gender' with a dropdown menu set to 'All'; 'Continent' with a dropdown menu set to 'All'; 'Country' with a dropdown menu set to 'All'; 'Month' with two input boxes for '1' and '12', and a range slider below them; and 'Year' with a dropdown menu set to 'All'.

Giải thích:

- Gender: Lọc theo giới tính
- Continent: Lọc địa
- Country: Quốc gia
- Lọc theo khoảng thời gian các tháng và năm

6. BOOKMARK

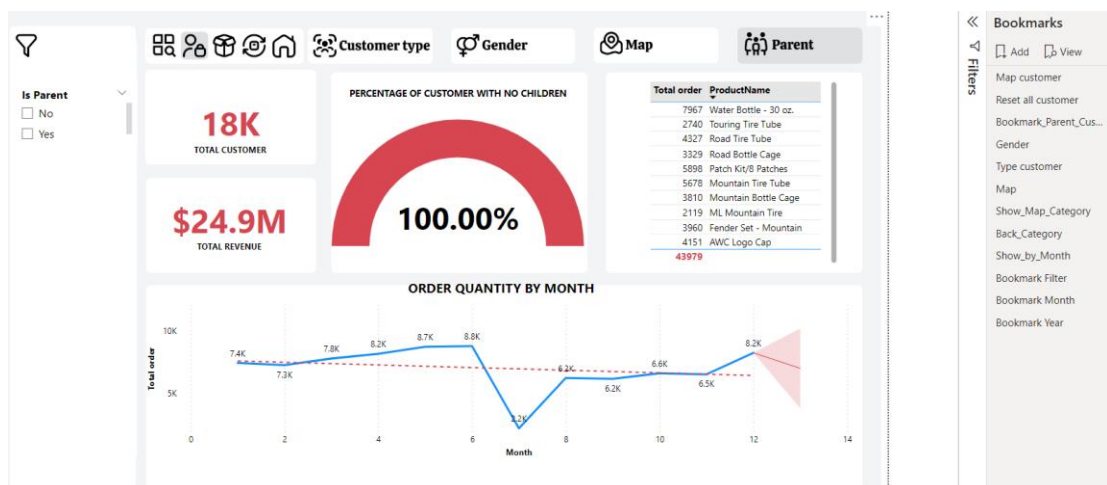
6.1. PAGE CUSTOMER

- Tạo bookmark slicer tổng số lượng đơn hàng theo khu vực



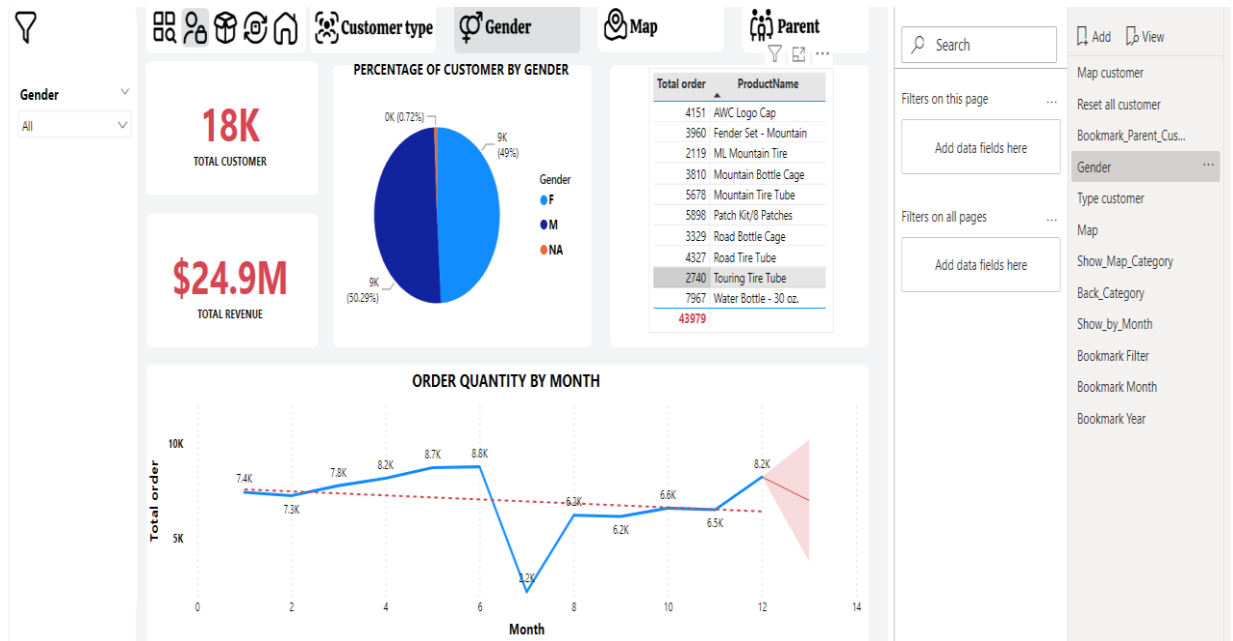
Hình ảnh 6.1.1

- Tạo bookmark hiển thị các visual filter theo Is Parent



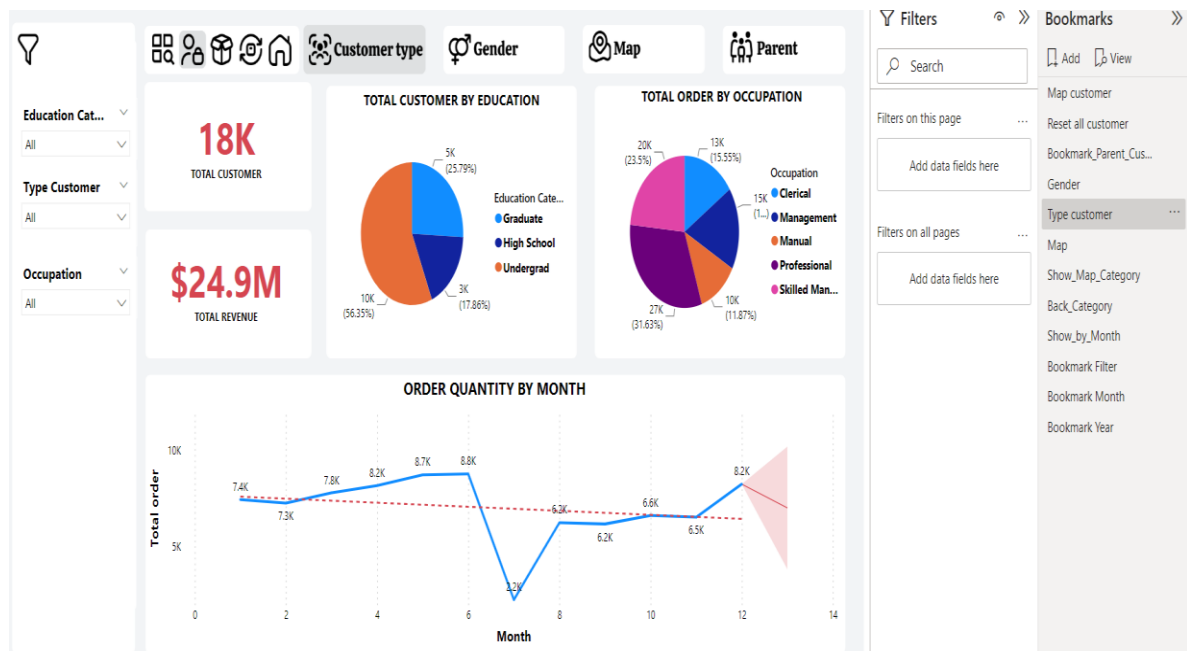
Hình ảnh 6.1.2

- Tạo bookmark hiển thị các visual và filter theo giới tính



Hình ảnh 6.2.3

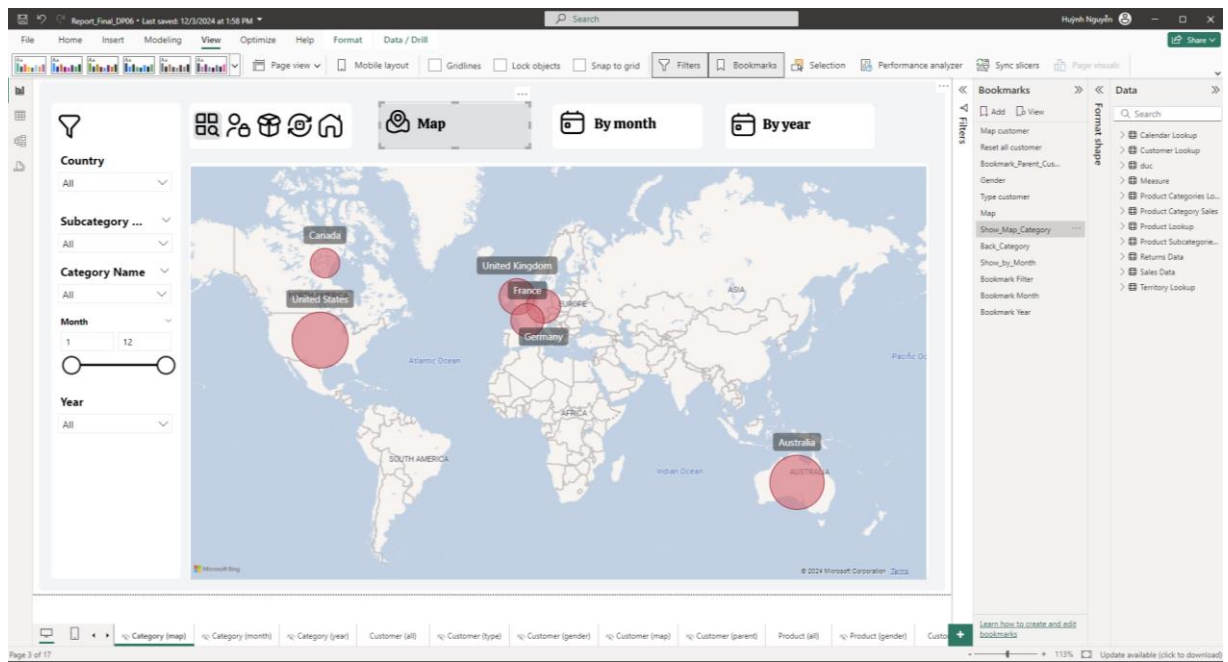
- Tạo bookmark hiển thị các visual thống kê theo loại khách hàng



Hình ảnh 6.2.4

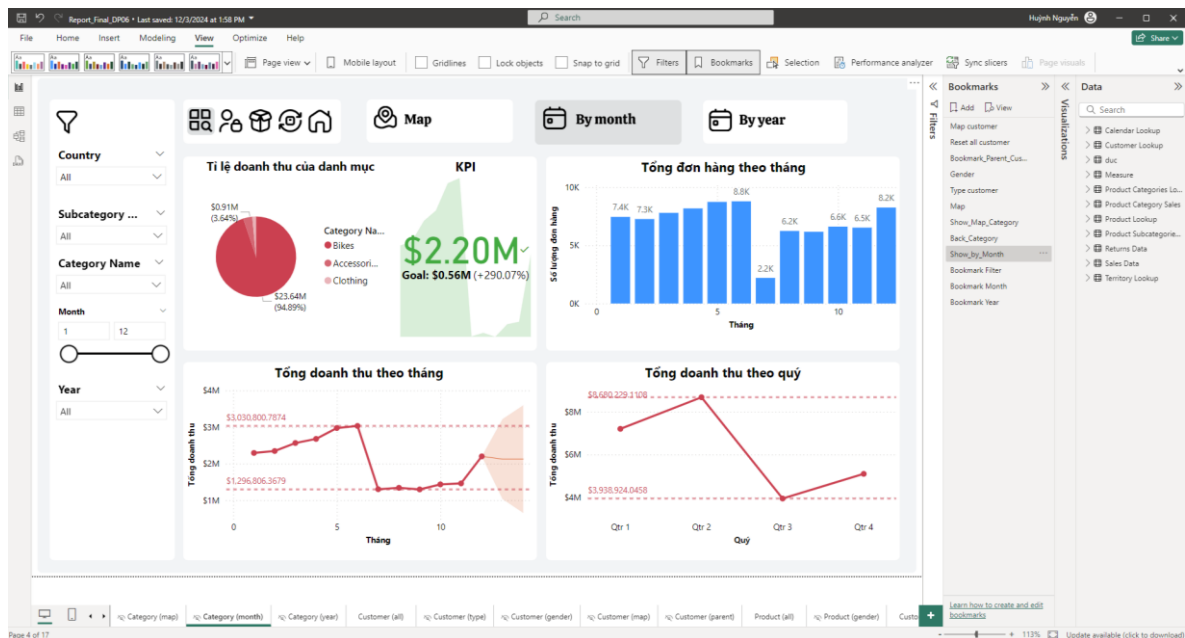
6.2. PAGE CATEGORY

- Bookmark hiển thị bản đồ



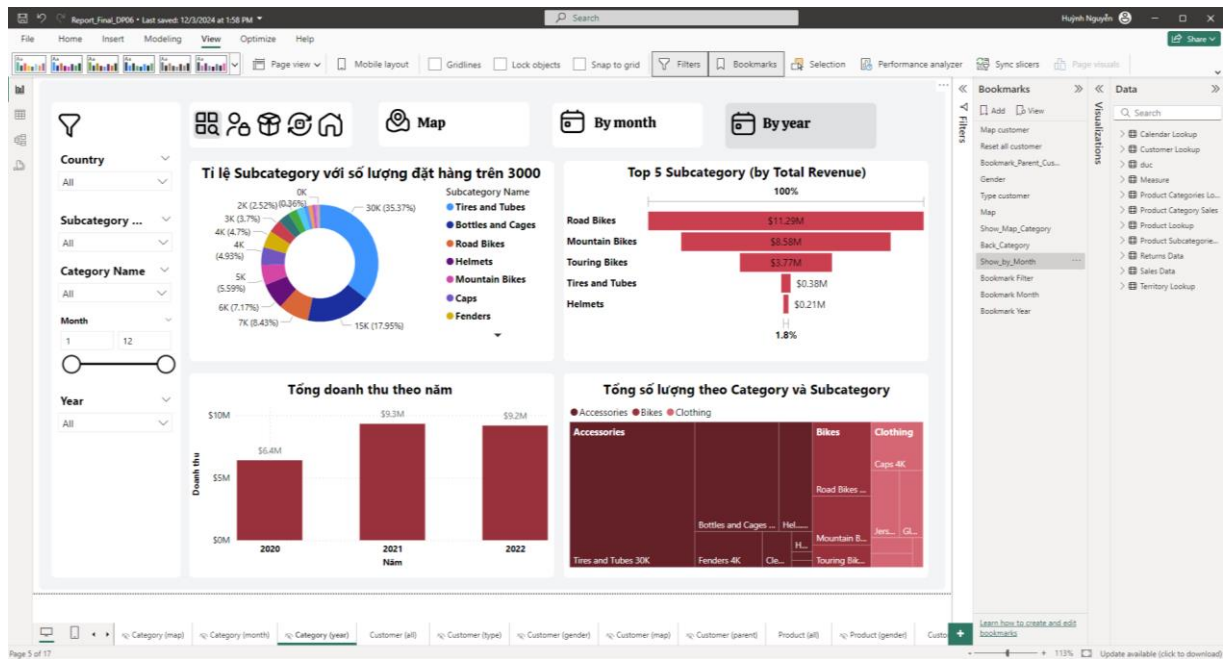
Hình ảnh 6.2.1

- Bookmark hiển thị theo Tháng



Hình ảnh 6.2.2

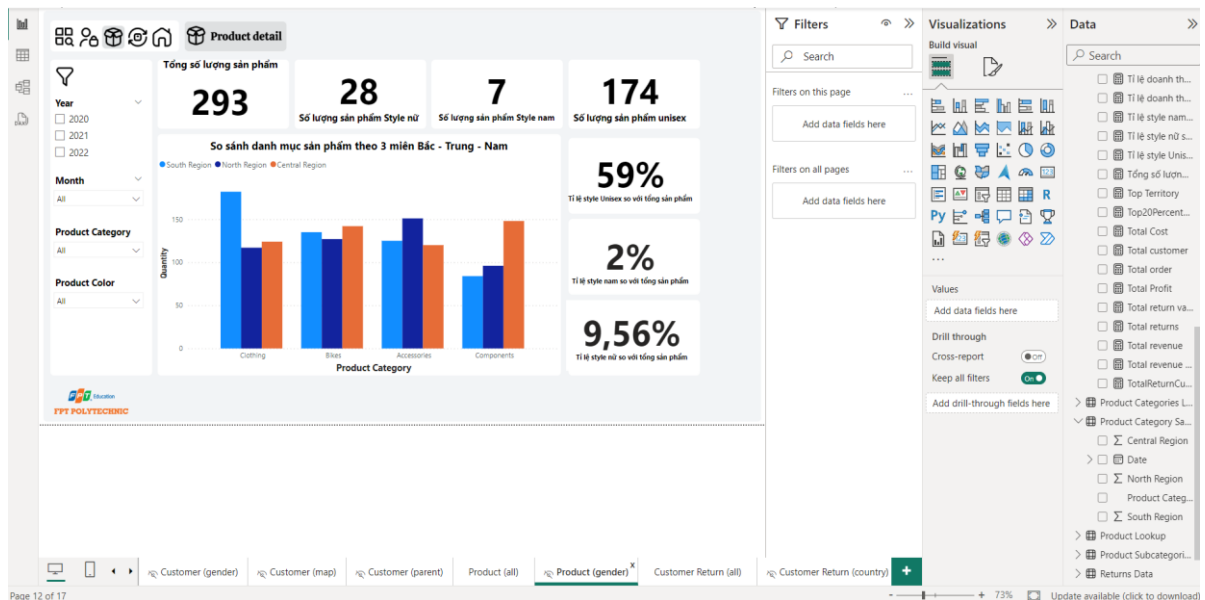
- Bookmark hiển thị theo Năm



Hình ảnh 6.2.3

6.3. PAGE PRODUCT

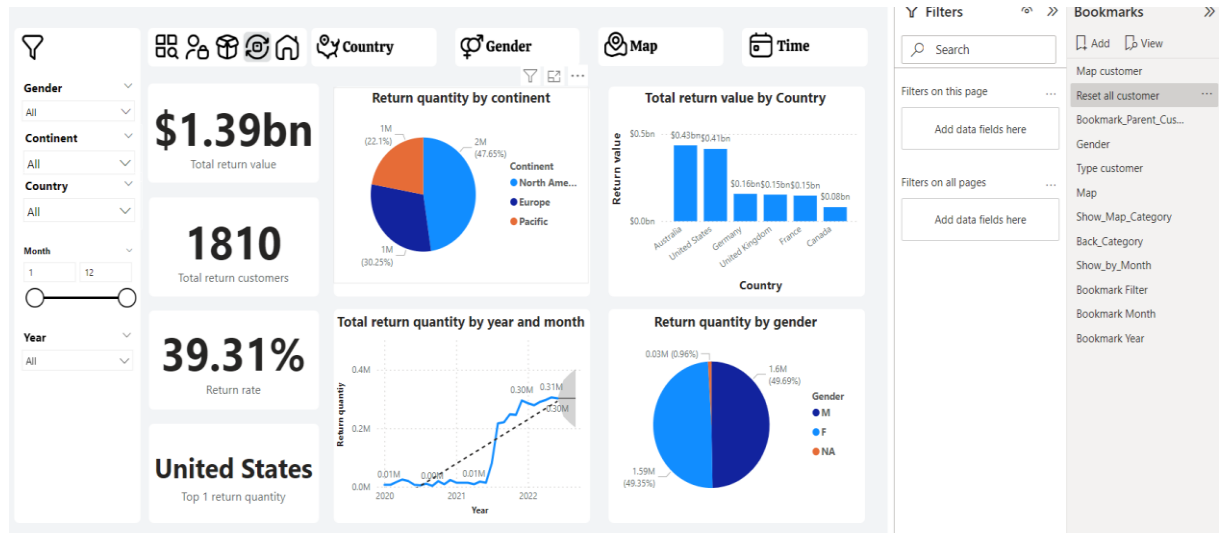
- Bookmark hiển thị theo thể loại sản phẩm (Style)



Hình ảnh 6.3.1

6.4. PAGE RETURN CUSTOMER

- Bookmark hiển thị Return Data



VIII. KẾT LUẬN

1. BÁO CÁO

- **Các bước viết báo cáo**

- Giới thiệu đề tài và câu chuyện dữ liệu
- Thiết kế khung báo cáo
- Tạo mục lục dự án cho báo cáo
- Lập kế hoạch
- Trình bày về bộ dữ liệu
- Các bước làm sạch và trực quan dữ liệu
- Trình bày dashboard và biểu đồ đã sử dụng
- Tạo bookmark
- Nêu lên những khó khăn và thuận lợi trong quá trình làm dự án
- Hướng phát triển và kết luận

2. KHÓ KHĂN

- Lựa chọn data
- Lên kế hoạch hoạt động dự án
- Chuẩn hóa dữ liệu
- Thiết kế mô hình dữ liệu
- Tạo các page dashboard sao cho phù hợp
- Thiết kế mô hình dự đoán

3. THUẬN LỢI

- Có nhiều tài liệu hỗ trợ
- Các vấn đề gặp phải đều dễ dàng xử lý
- Thành viên nhóm tích cực
- Giảng viên hỗ trợ kịp thời

4. HƯỚNG PHÁT TRIỂN

4.1. Mở rộng quy mô dữ liệu

Việc tích hợp thêm dữ liệu lịch sử và dữ liệu từ các nguồn bên ngoài sẽ làm phong phú thêm mô hình, giúp nó phản ánh đầy đủ hơn bối cảnh kinh doanh. Cụ thể, dữ liệu lịch sử về giá bán, doanh số, và tồn kho có thể được thêm vào để tạo nên bức tranh toàn diện về xu hướng sản phẩm qua thời gian. Đồng thời, việc tích hợp dữ liệu thị trường từ các nguồn bên ngoài, như báo cáo thị trường, thông tin đối thủ cạnh tranh, hoặc chỉ số kinh tế (lạm phát, thuế suất) sẽ giúp doanh nghiệp dự đoán các thay đổi trong hành vi người tiêu dùng và lên kế hoạch phù hợp.

4.2. Nâng cao năng lực dự báo và phân tích

Mô hình có thể phát triển thành một công cụ dự báo mạnh mẽ. Với dữ liệu lịch sử và các thuật toán phân tích, mô hình có thể dự đoán doanh thu, nhu cầu sản phẩm, hoặc lợi nhuận trong tương lai. Điều này rất hữu ích trong việc lập kế hoạch sản xuất, quản lý hàng tồn kho, và điều chỉnh chiến lược kinh doanh.

Bên cạnh đó, mô hình cũng có thể tập trung vào phân tích hành vi khách hàng. Bằng cách kết hợp dữ liệu từ bảng khách hàng, doanh số bán hàng, và tỷ lệ hoàn trả, doanh nghiệp có thể hiểu rõ hơn về thói quen mua sắm, mức độ trung thành,

và các yếu tố dẫn đến việc hoàn trả sản phẩm. Thông qua phân tích này, doanh nghiệp có thể cá nhân hóa trải nghiệm mua sắm và cải thiện chất lượng dịch vụ.

4.3. Tối ưu hóa quản lý danh mục sản phẩm

Dashboard sản phẩm có thể được nâng cấp để cung cấp cái nhìn chi tiết về hiệu suất của từng sản phẩm và danh mục sản phẩm. Thông qua việc so sánh doanh thu, lợi nhuận, và tỷ lệ hoàn trả, doanh nghiệp có thể xác định những sản phẩm cần tập trung, loại bỏ hoặc tái cấu trúc.

4.4. Tích hợp báo cáo và trực quan hóa dữ liệu

Các dashboard này có thể cung cấp thông tin theo thời gian thực, giúp đội ngũ quản lý đưa ra quyết định nhanh chóng và chính xác hơn. Ví dụ, một dashboard hiển thị doanh thu theo từng khu vực hoặc danh mục sản phẩm sẽ giúp doanh nghiệp nhanh chóng nhận ra những khu vực có hiệu suất kém và điều chỉnh chiến lược phân phối.

5. TỔNG KẾT

- **Doanh thu theo mùa :**

- Mùa Xuân đóng góp doanh thu cao nhất với \$8.2M.
- Mùa Thu có doanh thu thấp nhất (\$4.2M).
- Điều này cho thấy bạn có thể tập trung vào các chiến lược tăng trưởng doanh thu vào mùa Thu.

- **Tỷ lệ phần trăm doanh thu theo mùa:**

- Doanh thu mùa Xuân chiếm tỷ lệ lớn nhất (38.8%), tiếp theo là mùa Đông (22.0%).
- Mùa Thu có tỷ lệ nhỏ nhất (15.4%), cần xây dựng chiến lược cải thiện.

- **Doanh thu theo khu vực :**

- Australia dẫn đầu về doanh thu với \$7.4M, trong khi Canada và một số khu vực khác đóng góp thấp hơn (\$1M trở xuống).
- Khu vực có doanh thu thấp có thể được xem xét để đẩy mạnh marketing hoặc cải thiện các hoạt động kinh doanh.

- **Tăng trưởng doanh thu qua các năm:**

- Tăng trưởng vượt bậc từ năm 2021 (45.58%) đến năm 2022 (211.07%).
- Đây là tín hiệu tích cực, cho thấy các chiến lược gần đây đang hoạt động hiệu quả





- **Khách hàng và sản lượng đặt hàng :**

- Sản lượng đặt hàng đạt 84K, với tổng số khách hàng là 18K.
- Trung bình mỗi khách hàng đặt khoảng 4.67 đơn hàng (84K/18K). Điều này cho thấy mức độ tương tác tốt, nhưng bạn có thể phân tích thêm để cải thiện giá trị đơn hàng trung bình.

- **Tỷ lệ tăng trưởng hàng năm :**

- Tỷ lệ tăng trưởng hàng năm đạt 127.82%, vượt trội so với kỳ vọng thông thường trong ngành.

QR DỰ ÁN:

Slide báo cáo	PowerBI Service
	
Github Repo	Trello quản lý
	
Streamlit	