

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
KHOA TOÁN – TIN HỌC



BÁO CÁO ĐỒ ÁN CUỐI KỲ

Môn học: Trực quan hóa dữ liệu

Đề tài: Project 3 - National Clothing Chain.

Nhóm 7

Họ và tên sinh viên	Mã số sinh viên
Đào Huy Hoàng	21110297
Nguyễn Vũ Tiến Dũng	21110273
Lê Hồng Cát	21110249
Nguyễn Minh Hùng	21110301
Nguyễn Trung Đức	21110269
Phạm Ngọc Trai	21110196

Ngày 17 Tháng 01 Năm 2025, Thành phố Hồ Chí Minh

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
KHOA TOÁN – TIN HỌC

BÁO CÁO ĐỒ ÁN CUỐI KỲ

Môn học: Trắc quan hóa dữ liệu

Đề tài: Project 3 - National Clothing Chain.

Họ và tên sinh viên	Mã số sinh viên	Phân công
Đào Huy Hoàng	21110297	Làm Dashboard 1, viết báo cáo, thiết kế outline, tổng hợp.
Nguyễn Vũ Tiến Dũng	21110273	Làm Dashboard 3, viết báo cáo
Lê Hồng Cát	21110249	Viết báo cáo, đóng góp ý kiến, thuyết trình.
Nguyễn Minh Hùng	21110301	Làm Dashboard 2, viết báo cáo.
Nguyễn Trung Đức	21110269	Làm Dashboard 2, viết báo cáo.
Phạm Ngọc Trai	21110196	Viết báo cáo, đóng góp ý kiến.

MỤC LỤC

I.	Giới thiệu	2
1.	Giới thiệu về bộ dữ liệu.....	2
2.	Mục tiêu hướng đến	2
II.	Phân tích thị trường	2
III.	Phân tích hiệu suất sản phẩm.....	7
IV.	Phân tích các nhóm khách hàng	14
V.	Kết luận và kiến nghị	18

I. Giới thiệu

1. Giới thiệu về bộ dữ liệu

Bộ dữ liệu National Clothing Chain là bộ dữ liệu về các chuỗi cung ứng quần áo trực tuyến quốc gia tại Mỹ cung cấp các thông tin hữu ích cho 1 chuỗi cửa hàng quần áo nghiên cứu tìm ra giải pháp nhằm tìm ra chiến lược kinh doanh mới. Các file dữ liệu bao gồm:

- census-data.xlsx: dữ liệu được thu thập bởi US Census Bureau, bao gồm, average income, location, population, industry;
- customer-list.xlsx: danh sách khách hàng;
- purchase-list.xlsx: danh sách sản phẩm mua;
- state-list.xlsx: danh sách bang của chuỗi cửa hàng;
- Ngoài ra, dữ liệu về thời tiết có thể được truy xuất từ **Weather US**.

2. Mục tiêu hướng đến

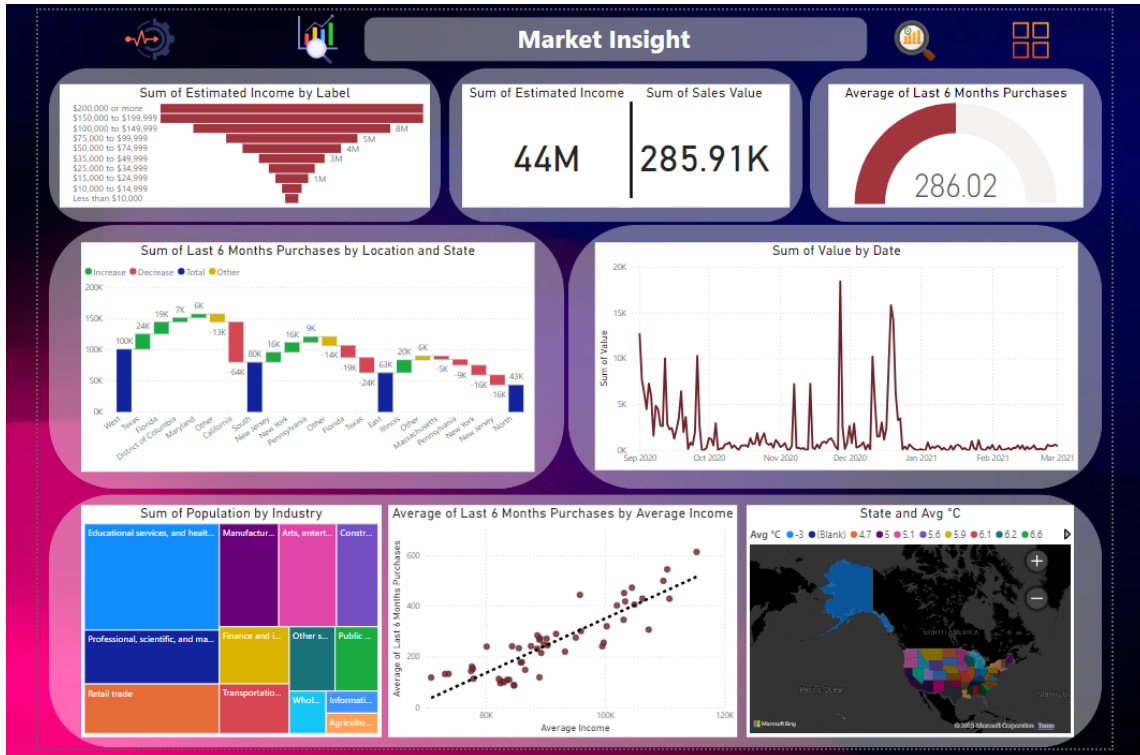
Trong dự án này, nhóm đặt ra 3 mục tiêu chính:

- Phân tích các yếu tố liên quan đến các khu vực địa lý (tiểu bang) và các yếu tố xã hội (dân số, thu nhập, ngành nghề) để hiểu rõ về tình hình thị trường, từ đó xây dựng chiến lược tiếp cận khách hàng và tối ưu hóa các hoạt động tiếp thị của từng khu vực.
- Đánh giá hiệu suất của các sản phẩm tồn kho, từ đó xác định các sản phẩm bán chạy, sản phẩm có tỷ lệ trả lại cao hoặc thấp hay các yếu tố khác như giá cả và đánh giá của khách hàng, giúp tối ưu hoá chiến lược sản phẩm.
- Hiểu rõ về các nhóm khách hàng theo độ tuổi, thu nhập và địa lý để tối ưu hóa chiến lược tiếp cận và bán hàng. Phân tích các yếu tố này giúp xây dựng các chiến dịch tiếp thị phù hợp với nhu cầu và khả năng chi trả của từng nhóm khách hàng.

Để cụ thể hóa các mục tiêu này, nhóm sẽ lần lượt đi qua từng mục tiêu trong các đề mục.

II. Phân tích thị trường

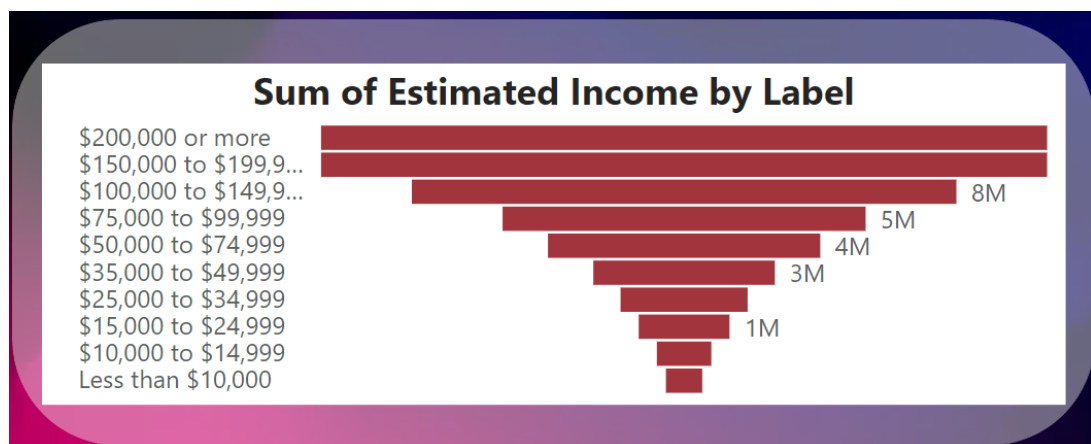
Ở phần này, nhóm tập trung phân tích các yếu tố liên quan đến các khu vực địa lý (tiểu bang) và các yếu tố xã hội (dân số, thu nhập, ngành nghề) để hiểu rõ về tình hình thị trường, từ đó có thể đề xuất các cách tiếp cận khách hàng và tối ưu hóa các hoạt động tiếp thị của từng khu vực cho thích hợp.



Tổng thu nhập ước tính theo nhãn (Sum of estimated Income by label)

Biểu đồ thanh này thể hiện số lượng thu nhập theo từng nhóm thu nhập của khách hàng. Một số nhận xét như:

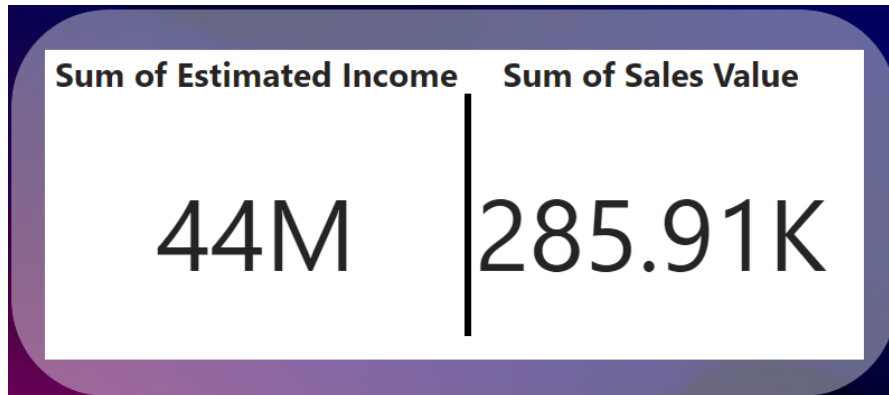
- Phần lớn khách hàng có thu nhập nằm trong nhóm từ **\$25,000 đến \$99,999**
- Tỷ lệ khách hàng có thu nhập trên \$100,000 khá thấp, cho thấy nhóm khách hàng mục tiêu của chuỗi cửa hàng là những người thuộc tầng lớp trung lưu.



Tổng thu nhập và tổng giá trị bán hàng (Sum of estimated Income & sum of Sales value)

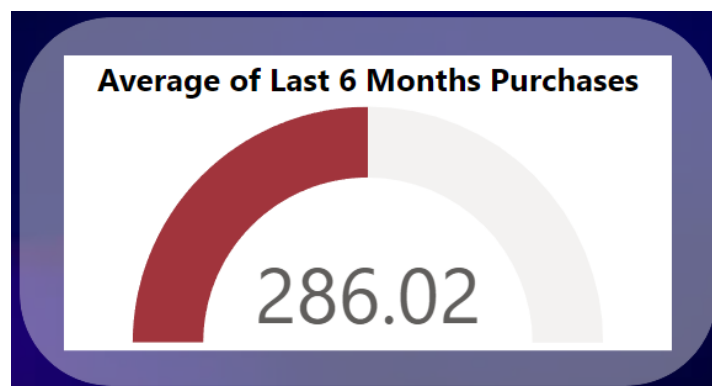
Biểu đồ này thể hiện 2 giá trị KPI quan trọng: Tổng thu nhập ước tính và tổng giá trị bán hàng. Từ đó, ta có thể thấy được một vài kết luận sau:

- Tổng thu nhập ước tính đạt khoảng \$44 triệu, trong khi đó tổng giá trị bán hàng đạt khoảng \$285,910.
- Như vậy, ta có thể thấy được rằng doanh thu còn khiêm tốn so với tiềm năng thu nhập của khách hàng. Điều này có thể phản ánh các chiến lược tiếp thị chưa đạt hiệu quả tối đa.



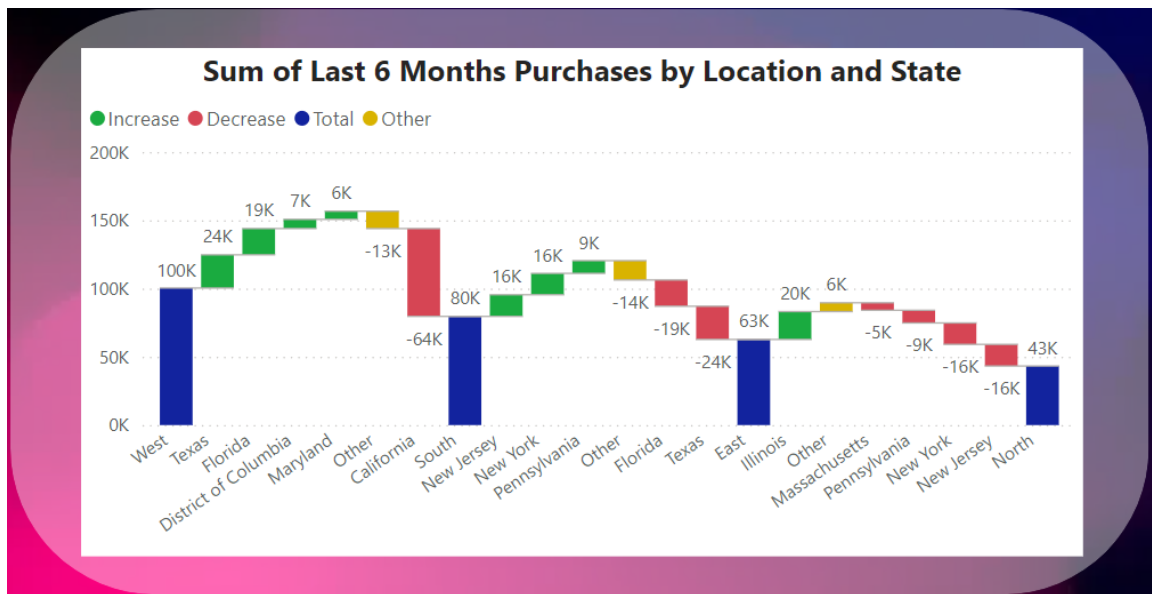
Trung bình số lượng mua sắm trong 6 tháng qua (Average of last 6 months purchases)

Ta thấy giá trị trung bình lượng mua sắm trong 6 tháng vừa qua đạt khoảng 286.02 sản phẩm/bang. Cần tập trung phân tích các khu vực có giá trị trung bình mua thấp để tìm cách cải thiện.



Tổng số lượng mua sắm 6 tháng qua theo tiểu bang (Sum of last 6 months purchases by Location and State)

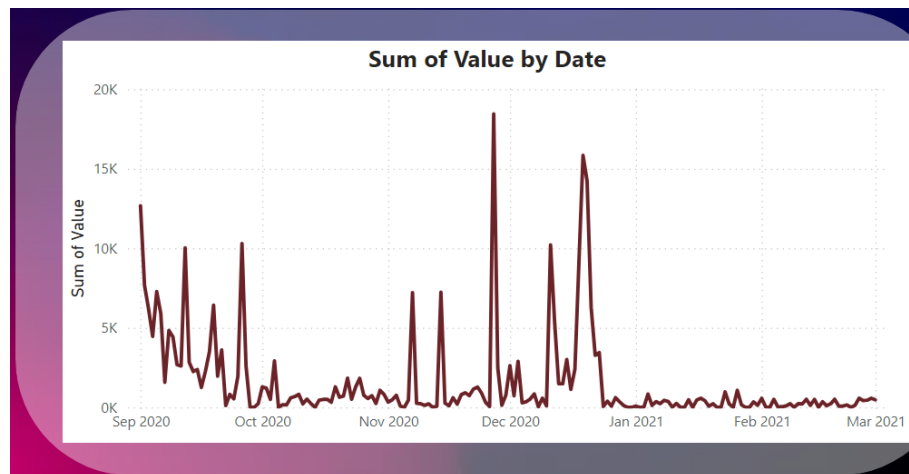
Ta sử dụng biểu đồ Waterfall để phân tích chi tiết tăng giảm số lượng mua sắm giữa các bang. Các bang như **California** và **Texas** có doanh số cao nhất. Một số bang như **New York** và **Michigan** cần được chú trọng cải thiện.



Tổng giá trị bán hàng theo ngày (Sum of Value by Date)

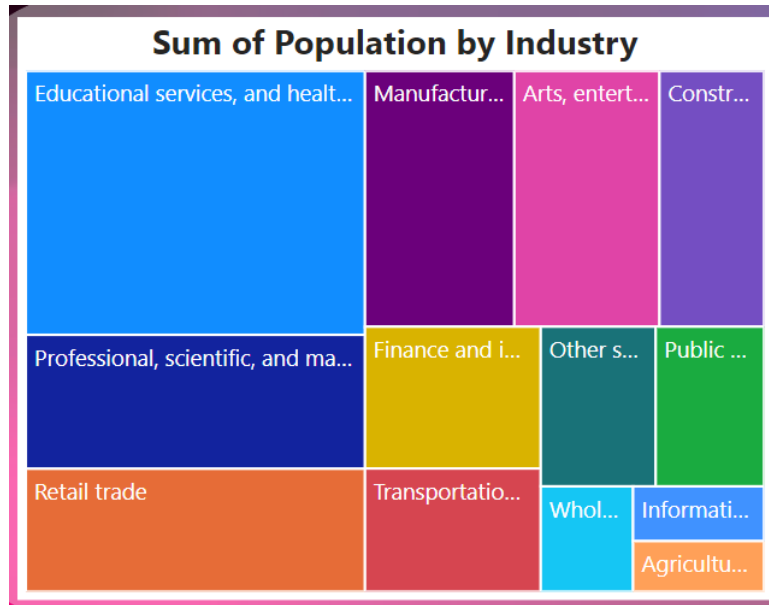
Biểu đồ đường hiển thị dao động giá trị bán hàng theo thời gian. Một số nhận xét như sau:

- Doanh số có xu hướng tăng vào các tháng lễ như **tháng 12 (Giáng Sinh)** và giảm dần sau đó.
- Điều này gợi ý cần xây dựng các chương trình khuyến mãi để kích thích mua sắm vào các thời điểm thấp điểm.



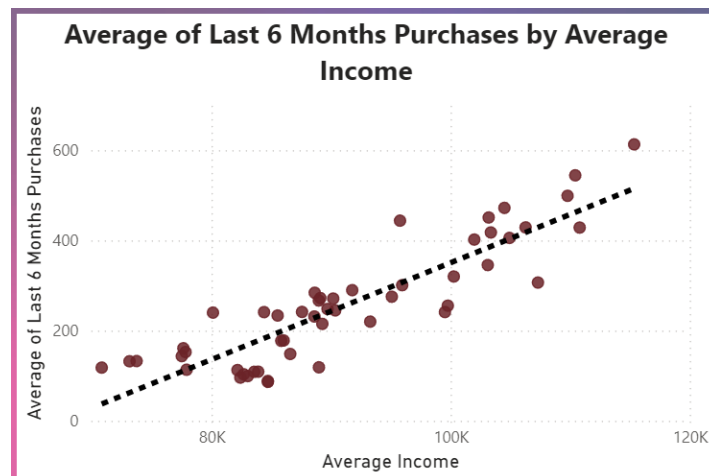
Tổng dân số theo ngành nghề (Sum of Population by Industry)

Biểu đồ Tree Map cho thấy phân bố dân số theo ngành nghề: Ngành giáo dục và y tế chiếm phần lớn, tiếp theo đó là các ngành sản xuất và bán lẻ. Chuỗi cửa hàng nên tập trung xây dựng sản phẩm và dịch vụ phù hợp với nhóm khách hàng thuộc các ngành giáo dục, y tế và sản xuất.



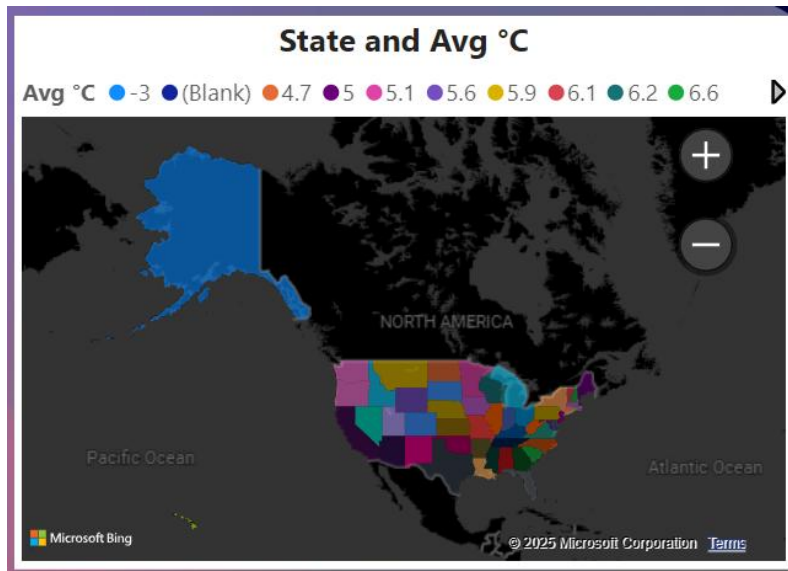
Mối quan hệ giữa thu nhập trung bình và số lượng mua sắm (Average of Last 6 Months Purchases by Average Income)

Biểu đồ scatter plot cho thấy mối tương quan tích cực giữa thu nhập trung bình và số lượng mua sắm. Các bang có thu nhập cao hơn thường mua sắm nhiều hơn. Chiến lược tiếp thị nên ưu tiên các khu vực có thu nhập cao.



Bản đồ nhiệt độ trung bình theo bang (State and Avg °C)

Bản đồ biểu diễn nhiệt độ trung bình của các bang, một số khu vực lạnh hơn (phía Bắc) có thể yêu cầu các sản phẩm mùa đông nhiều hơn. Ngược lại, các bang phía Nam cần ưu tiên sản phẩm thời trang mùa hè.



Đề xuất chiến lược

Dựa trên các phân tích, sau đây là các đề xuất chiến lược của nhóm:

1. Tiếp thị theo địa lý và ngành nghề:

- Tăng cường chiến dịch quảng cáo tại các bang như California, Texas.
- Đưa ra ưu đãi cho các ngành giáo dục, y tế và sản xuất.

2. Tối ưu hóa sản phẩm theo thời tiết

- Phân phối sản phẩm phù hợp với điều kiện khí hậu từng khu vực.

3. Chương trình khuyến mãi theo mùa:

- Triển khai chiến dịch khuyến mãi lớn vào các dịp lễ như Giáng Sinh, Black Friday.

4. Nâng cao trải nghiệm khách hàng:

- Phân tích thêm dữ liệu để cải thiện trải nghiệm mua sắm, tập trung vào nhóm khách hàng trung lưu.

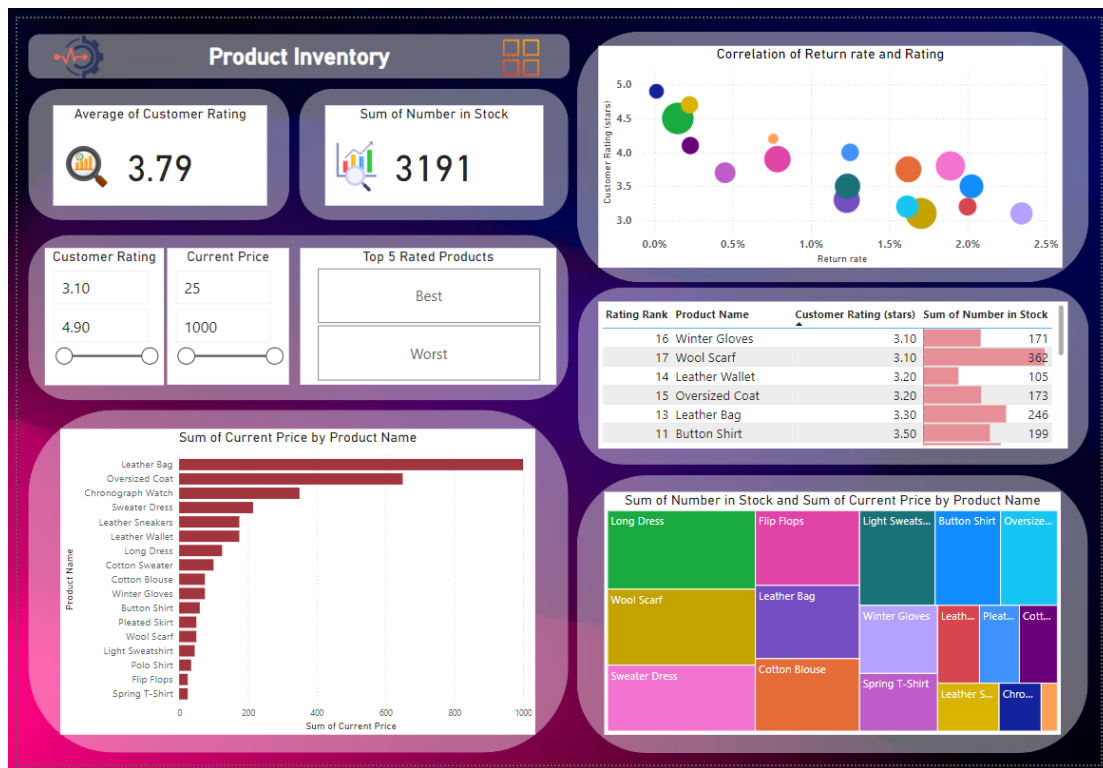
5. Tăng cường nghiên cứu thị trường:

- Thu thập thêm dữ liệu về hành vi mua sắm để xây dựng các chiến lược phù hợp hơn.

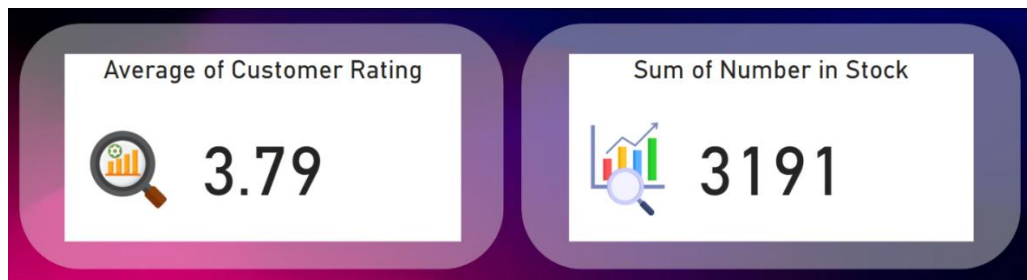
III. Phân tích hiệu suất sản phẩm

Mục tiêu: Nhóm sẽ cố gắng đánh giá hiệu suất của các sản phẩm tồn kho, từ đó xác định các sản phẩm bán chạy, sản phẩm có tỷ lệ trả lại cao hoặc thấp hay các yếu tố khác như giá cả và đánh giá của khách hàng, giúp tối ưu các đề xuất.

Dashboard được thiết lập với các biểu đồ nhằm phân tích sản phẩm, với các Tooltips phù hợp để cung cấp thêm thông tin khi tương tác:



Hai thẻ (card) về **Trung Bình Đánh Giá Khách Hàng** (Average of Customer Rating (stars)) và **Tổng Số Hàng Trong Kho** (Sum of Number in Stock) sẽ cung cấp cái nhìn tổng quan về chất lượng sản phẩm dựa trên đánh giá của khách hàng và tình trạng tồn kho hiện tại.

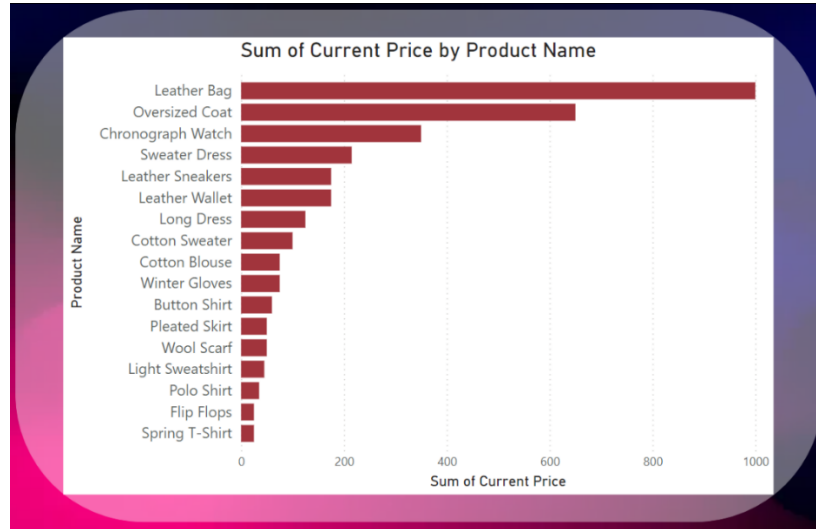


Tổng giá hiện tại của từng sản phẩm (Sum of Current Price by Product Name)

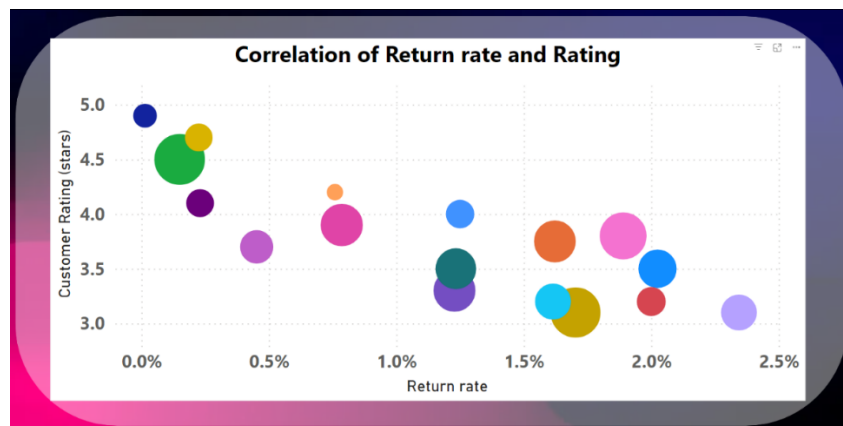
- **Biểu đồ**: Nhóm đề xuất biểu đồ cột hoặc thanh (Bar/Column Chart) nằm ngang
- **Mục tiêu**: Cho cái nhìn tổng quan cũng như dễ dàng so sánh giữa giá của các sản phẩm.
- **Nhận xét**:
 - + Leather Bag là sản phẩm có giá cao nhất, tiếp theo là Oversized Coat. Hai sản phẩm này có giá trị cao đáng kể so với các sản phẩm còn lại. Nhưng mà rating lại xếp hạng lần lượt ở 13 và 15 (nằm trong top 5 các sản phẩm được đánh giá tệ nhất), chứng tỏ sản phẩm mắc tiền thì đi kèm với sự kỳ vọng cao hơn từ người

dùng, nên Rating có độ biến động lớn hơn, khiến cho Rating có thể bị giảm đi đáng kể.

- + Sản phẩm Chronograph Watch đắt thứ ba, có chênh lệch giá lớn với 2 sản phẩm mắc nhất nhưng lại có Rating cao.



Sự tương quan của các sản phẩm giữa tỷ lệ hoàn trả sản phẩm và xếp hạng theo đánh giá khách hàng (Correlation of Return rate and Rating)



Để có thể đánh giá được sự tương quan của các sản phẩm giữa tỷ lệ hoàn trả và đánh giá xếp hạng của khách hàng, nhóm đề xuất sử dụng biểu đồ và có kết luận như sau:

- **Biểu đồ:** Phân tán (Scatter Chart) với trục X là Return Rate (Tỷ lệ hoàn trả sản phẩm) và trục Y là Customer Rating (Điểm đánh giá của khách hàng). Với màu sắc của từng sản phẩm được đồng bộ với biểu đồ treemap ở dưới, với kích thước bong bóng khác nhau đại diện cho số lượng hàng của từng sản phẩm khác nhau, cùng với việc cài đặt giới hạn khoảng cho trục X và trục Y để tập trung vào điểm phân bố của dữ liệu, giúp dễ dàng so sánh giữa các sản phẩm.

- **Mục tiêu:** Biểu đồ nhằm xác định mối quan hệ giữa tỷ lệ hoàn trả và đánh giá khách hàng cùng với số lượng hàng trong kho.
- **Nhận xét:** Qua biểu đồ cho thấy, các sản phẩm có tỷ lệ hoàn trả hàng thấp thường nhận được điểm đánh giá cao hơn, tức có sự tương quan giữa tỷ lệ hoàn trả sản phẩm và đánh giá của khách hàng.

Tổng số lượng hàng trong kho và tổng giá hiện tại theo sản phẩm (Sum of Number in Stock and Sum of Current Price by Product Name)

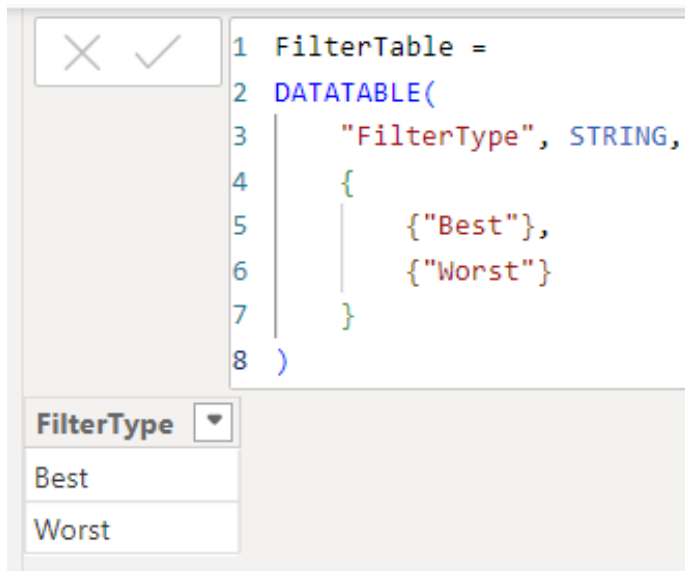


Đối với tác vụ này, nhóm đề xuất như sau:

- **Biểu đồ:** Biểu đồ cây (Tree Map) với Category là Product Name và Values là Sum of Number in Stock.
- **Mục tiêu:** Biểu đồ giúp trực quan hoá tỷ lệ hàng trong kho và giá hiện tại giữa các sản phẩm, giúp dễ dàng so sánh mức độ tồn kho và giá của các sản phẩm.
- **Nhận xét:**
 - + Long Dress chiếm ô lớn nhất trong biểu đồ, cho thấy sản phẩm này có giá trị đáng kể về số lượng tồn kho hoặc giá cao. Tiếp theo là Wool Scarf, Sweater Dress, Flip Flops, cho thấy đây là những sản phẩm còn tồn kho nhiều, cần phải tối ưu hoá chiến lược để bán được những sản phẩm này. Đặc biệt là Wool Scarf khi nó có Rating thấp và tỷ lệ hoàn trả cao.

Ngoài ra, nhóm tạo thêm bảng và xây dựng các **DAX functions** để xếp hạng các sản phẩm dựa trên đánh giá khách hàng (đặc biệt dễ dàng thấy được **top 5 sản phẩm được đánh giá tốt nhất và tệ nhất**) và số lượng hàng trong kho:

- *Filter Table*
 - + Tạo bảng FilterTable chứa các nội dung là Best và Worst.
 - + **Mục tiêu:** Bảng sẽ hỗ trợ bộ lọc trong việc tạo hai nút Best và Worst.



- *Rating Rank Column*
- + Viết hàm để tạo cột Rating Rank trong bảng Product Inventory.
- + **Mục tiêu:** Xếp hạng các sản phẩm dựa trên đánh giá khách hàng và số sản phẩm tồn trong kho.

```
1 Rating Rank =  
2 RANKX(  
3     ALL('Product Inventory'),  
4     (  
5         'Product Inventory'[Customer Rating (stars)] * 1000000  
6         + (1 / 'Product Inventory'[Number in Stock])  
7     ),  
8     ,  
9     DESC  
10 )
```

- + Quy trình:
 - Xếp hạng các sản phẩm theo Customer Rating (được nhân với 1,000,000 để ưu tiên hơn các yếu tố khác).
 - Nếu các sản phẩm có điểm Rating Rank bằng nhau, hàm sẽ dùng nghịch đảo của Number in Stock để làm yếu tố phụ cho việc xếp hạng (giữa hai sản phẩm cùng điểm đánh giá, sản phẩm nào ít hàng tồn kho hơn sẽ được xếp hạng cao hơn).
 - Các thao tác trên sẽ cho vào RANKX để cho ra kết quả được thao tác giảm dần (DESC).
 - **Nhận xét:** Các sản phẩm sẽ được xếp hạng theo các tiêu chí trên.

Rating Rank	Product Name	Customer Rating (stars)	Sum of Number in Stock
1	Chronograph Watch	4.90	65
2	Leather Sneakers	4.70	95
3	Long Dress	4.50	373
4	Polo Shirt	4.20	25
5	Cotton Sweater	4.10	95
6	Pleated Skirt	4.00	100
7	Flip Flops	3.90	250
8	Sweater Dress	3.80	314
9	Cotton Blouse	3.75	243
10	Spring T-Shirt	3.70	145
11	Button Shirt	3.50	199
12	Light Sweatshirt	3.50	230
13	Leather Bag	3.30	246
14	Leather Wallet	3.20	105
15	Oversized Coat	3.20	173
16	Winter Gloves	3.10	171
17	Wool Scarf	3.10	362
Total			3191

- ShowTopBottom5 Column:
- + Viết hàm để tạo cột ShowTopBottom5 trong bảng Product Inventory.
- + **Mục tiêu:** Hàm sẽ hỗ trợ trong việc thể hiện Top 5 sản phẩm tốt nhất hoặc tệ nhất từ kết quả xếp hạng.

```

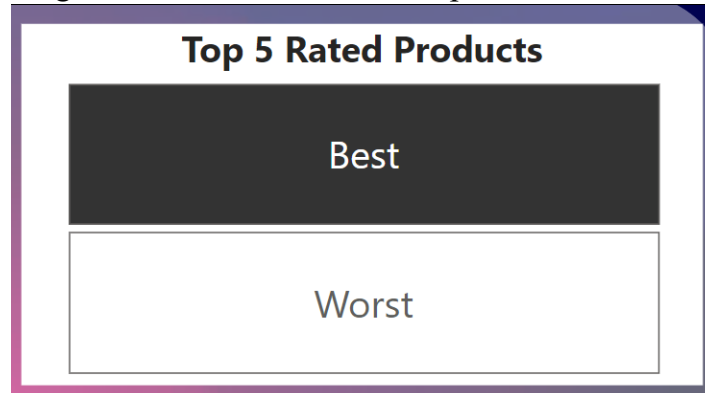
1 ShowTopBottom5 =
2 VAR SelectedFilter = SELECTEDVALUE(FilterTable[FilterType], "All")
3 VAR RankTop =
4     RANKX(
5         ALL('Product Inventory'),
6         CALCULATE(MAX('Product Inventory'[Rating Rank])),
7         ,
8         ASC
9     ) -- Xếp hạng tăng dần cho Top 5
10 VAR RankBottom =
11     RANKX(
12         ALL('Product Inventory'),
13         CALCULATE(MAX('Product Inventory'[Rating Rank])),
14         ,
15         DESC
16     ) -- Xếp hạng giảm dần cho Bottom 5
17 RETURN
18     SWITCH(
19         TRUE(),
20         SelectedFilter = "Best" && RankTop <= 5, 1, -- Hiển thị Top 5
21         SelectedFilter = "Worst" && RankBottom <= 5, 1, -- Hiển thị Bottom 5
22         SelectedFilter = "All", 1,
23         0 -- Mặc định không hiển thị sản phẩm nào nếu không thoả điều kiện
24     )

```

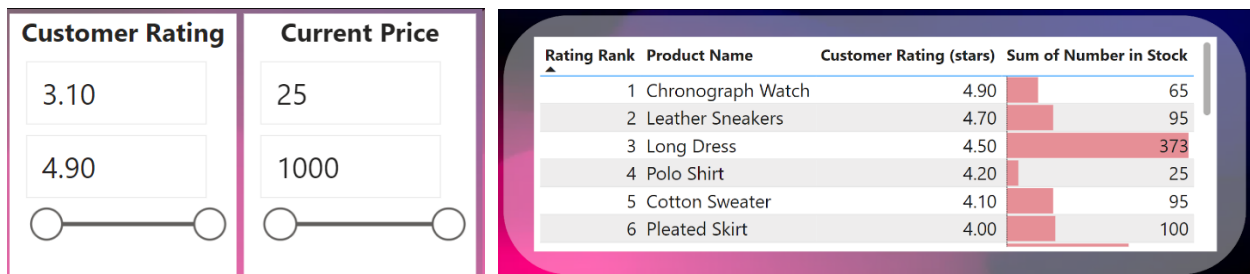
- + Quy trình:
 - VAR SelectedFilter: Lấy giá trị được chọn từ Slicer TableFilter[FilterType] với mặc định hiện tất cả sản phẩm.
 - VAR RankTop: Xếp hạng các sản phẩm dựa trên cột Rating Rank đã tạo ở trên theo thứ tự tăng dần (ASC).
 - VAR RankBottom: Tương tự khi xếp hạng các sản phẩm theo thứ tự giảm dần (DESC).
 - SWITCH logic: Trả về giá trị 1 nếu thoả điều kiện: Nếu Slicer chọn Best, hiển thị các sản phẩm nằm trong Top 5 (RankTop <= 5). Nếu Slicer chọn Worst, hiển

thì các sản phẩm nằm trong Bottom 5 ($\text{RankBottom} \leq 5$). Nếu không chọn nút lọc nào, mặc định hiện tất cả sản phẩm.

- **Nhận xét:** Hàm kết hợp với FilterTable để tạo một Slicer dạng nút bấm. Khi chọn Best, 5 sản phẩm tốt nhất sẽ được hiển thị trên tất cả các biểu đồ trong Dashboard. Tương tự với Worst đối với 5 sản phẩm tệ nhất.



Ngoài ra, nhóm tạo thêm các Slicer dạng kéo các giá trị của Customer Rating (stars) và Current Price để tăng tính tương tác cho dashboard, cũng như Bảng (Table), tích hợp với xếp hạng của sản phẩm theo Rating, để cho biết thêm thông tin tổng quan của các sản phẩm.



Đề xuất chiến lược:

Từ các phân tích trên, ta có thể đưa ra các định hướng sau:

1. Tối ưu hóa các sản phẩm có giá trị cao:

- Cần đảm bảo chất lượng sản phẩm để duy trì đánh giá tốt từ khách hàng.
- Xem xét chiến lược tiếp thị để tăng sức mua, đẩy mạnh doanh số các sản phẩm này.

2. Giảm thiểu tỷ lệ hoàn trả:

- Tập trung vào các sản phẩm có tỷ lệ hoàn trả cao. Từ đó phân tích lý do hoàn trả (lỗi kỹ thuật, sai mô tả hay vấn đề vận chuyển) từ đó cải thiện quy trình kiểm tra chất lượng sản phẩm trước khi đưa ra thị trường.
- Ngoài ra, có thể xem xét chính sách đổi trả và chăm sóc khách hàng để hỗ trợ khách hàng tốt hơn.

3. Tăng cường theo dõi, quản lý hàng tồn kho:

- Phân tích tốc độ bán hàng tránh tình trạng tồn kho quá lâu đối với các sản phẩm có số lượng hàng tồn kho nhiều.
- Thay đổi chiến lược tiếp thị hoặc chương trình khuyến mãi phù hợp để giảm lượng hàng tồn kho.

4. Nâng cao trải nghiệm khách hàng:

- Phân tích các phản hồi của khách hàng để cải thiện sản phẩm, nhất là các sản phẩm có điểm đánh giá thấp.
- Khảo sát nhận góp ý từ khách hàng, từ đó theo dõi được xu hướng và sở thích của khách hàng

5. Tập trung vào các sản phẩm tiềm năng:

- Đẩy mạnh chiến lược marketing đối với các sản phẩm có tỷ lệ hoàn trả thấp và được đánh giá cao để tăng doanh số.
- Duy trì số lượng hàng tồn kho đủ để đáp ứng nhu cầu khách hàng.

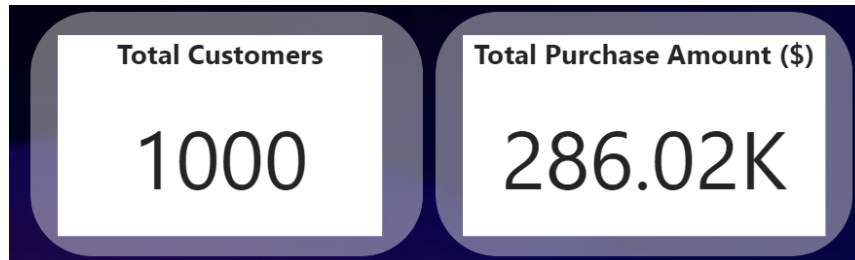
IV. Phân tích các nhóm khách hàng

Mục tiêu của nhóm là cố gắng hiểu rõ về các nhóm khách hàng theo độ tuổi, thu nhập và địa lý để tối ưu hóa chiến lược tiếp cận và bán hàng. Phân tích các yếu tố này giúp xây dựng các chiến dịch tiếp thị phù hợp với nhu cầu và khả năng chi trả của từng nhóm khách hàng. Dashboard được thiết kế với các biểu đồ và thẻ tổng kết, giúp hỗ trợ phân tích và cung cấp cái nhìn tổng quan về thị phần thị trường.



Thẻ Số lượng khách hàng (Customers) và Tổng số tiền thực hiện mua bán (Total purchase amount):

Cho biết số lượng khách hàng có trong tệp dữ liệu được dùng để phân tích và tổng số tiền các khách hàng này đã thực hiện mua bán.

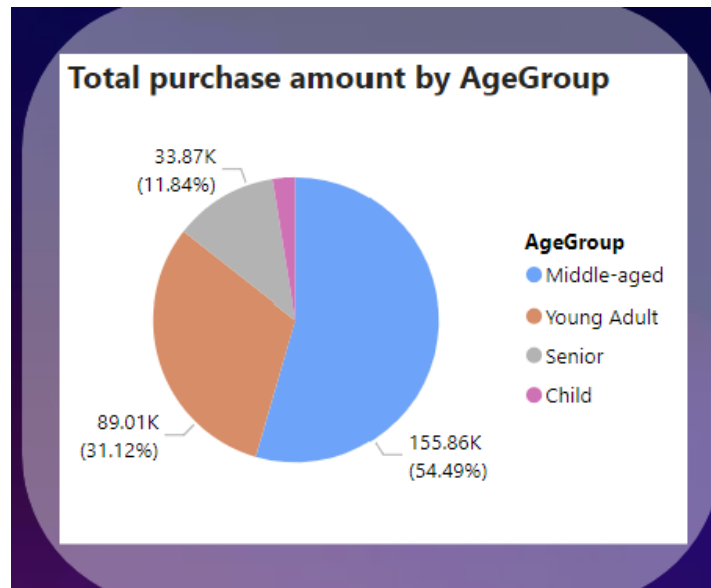


Tổng số tiền thực hiện mua bán theo nhóm độ tuổi (Total purchase amount (\$) by AgeGroup):

Nhóm đề xuất sử dụng biểu đồ tròn (Pie chart).

Mục tiêu: Trực quan hóa tỷ lệ đóng góp của từng nhóm độ tuổi vào tổng số tiền mua bán. Biểu đồ tròn giúp dễ dàng nhận diện các nhóm độ tuổi có mức chi tiêu cao hoặc thấp, từ đó hỗ trợ việc phân tích hành vi người tiêu dùng và đưa ra các chiến lược marketing phù hợp cho từng đối tượng khách hàng.

Nhận xét: Doanh thu tập trung chủ yếu ở Middle-aged và Young Adult, đây là 2 nhóm tuổi thuộc độ tuổi từ 18 - 55 tuổi, ta có thể thấy nhu cầu mua sắm của những bạn trẻ và người trong độ tuổi đi làm rất cao vì họ cần ăn mặc đẹp để phục vụ cho mục đích đi học đại học cũng như đi làm. Từ đó các doanh nghiệp nên tập trung đánh mạnh vào 2 nhóm tuổi này để tối đa hóa lợi nhuận.

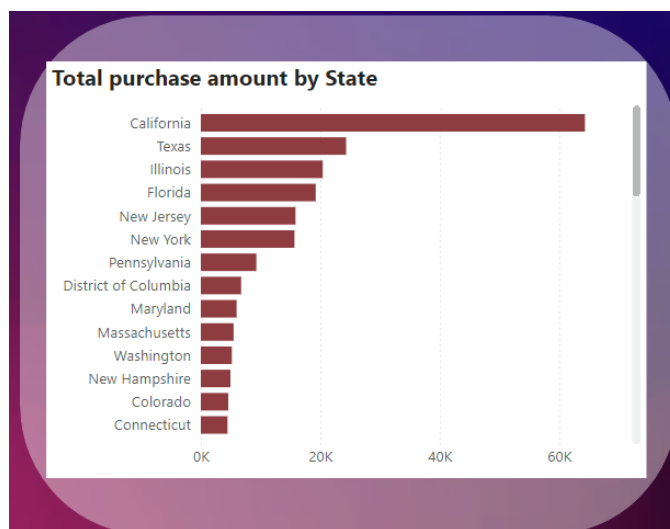


Tổng số tiền thực hiện mua bán theo nhóm các bang (Total purchase amount (\$) by State):

Để thực hiện tác vụ này, nhóm đề xuất sử dụng biểu đồ cột nhóm (Clustered bar chart).

Mục tiêu: Trực quan hóa số tiền đóng góp của từng tiểu bang. Xác định được những tiểu bang có đóng góp lớn để tập trung phát triển tối đa hóa dòng tiền, bên cạnh đó xác định những tiểu bang có doanh số thấp để tìm ra giải pháp, khắc phục vấn đề để cải thiện hơn.

Nhận xét: Ta có thể thấy tiểu bang California, Texas và Illinois là 1 trong những tiểu bang có doanh số cao nhất, nên tập trung khai thác tối đa những tiểu bang này để tối đa hóa dòng tiền. Thấp nhất là South Dakota và Montana, các doanh nghiệp, cửa hàng ở những nơi đây cần khám phá, tìm ra vấn đề vì sao lại có doanh thu thấp như vậy để từ đó cải thiện, thay đổi chiến lược kinh doanh, các chương trình ưu đãi.

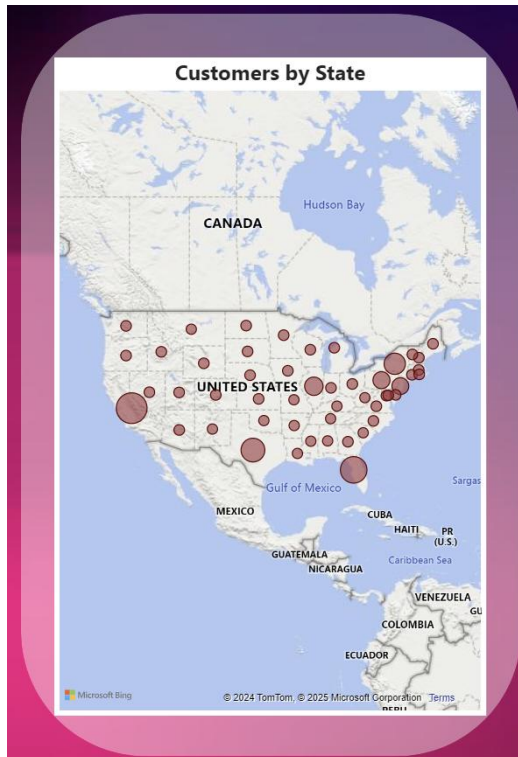


Số lượng khách hàng theo tiểu bang (Customers by State):

Để phân tích khách hàng theo tiểu bang thì nhóm sẽ sử dụng biểu đồ bản đồ để thực hiện tác vụ này.

Mục tiêu: Trực quan hóa số lượng khách hàng theo từng khu vực. Từ biểu đồ Map và Clustered bar chart ở trên ta có thể rút ra được mối liên hệ đó là số lượng khách hàng ở mỗi tiểu bang càng cao thì doanh thu càng lớn.

Nhận xét: Từ biểu đồ Map và Clustered bar chart ở trên ta có thể rút ra được mối liên hệ đó là số lượng khách hàng ở mỗi tiểu bang càng cao thì doanh thu càng lớn.



Mười khách hàng hàng chi tiêu nhiều nhất (Top 10 Customers):

Nhóm sử dụng biểu đồ Matrix để phân tích tác vụ này.

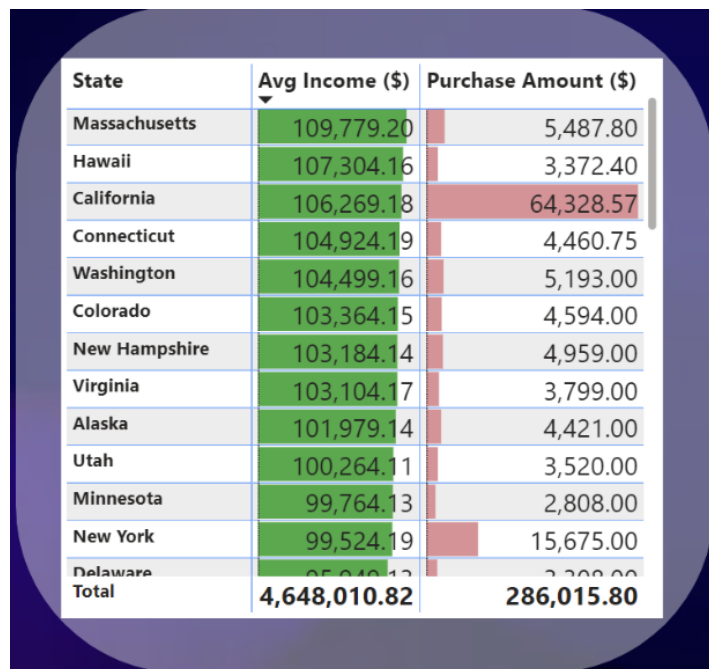
Mục tiêu: Xác định những khách hàng tiềm năng, có nhu cầu mua sắm cao. Từ đó xây dựng mối quan hệ, kết nối tương tác nhiều hơn với những khách hàng này để tạo ra sự tin tưởng, từ những khách hàng tiềm năng trở thành những khách hàng trung thành.

Full Name	Purchase amount (\$)
Jon Little	5,250.00
Jody Foster	2,500.00
Tonya Rodriguez	2,350.00
Edna Barnett	1,773.52
Larry Cain	1,736.00
Floyd Quinn	1,484.00
Margarita Walters	1,466.64
Sue Burgess	1,450.96
Total	20,816.48

Tổng hợp: Sử dụng biểu đồ Matrix.

Mục tiêu: Tìm hiểu xem thu nhập của khách hàng ở từng tiểu bang có ảnh hưởng đến số tiền mua sắm của họ hay không. Từ matrix trên ta có thể thấy không phải thu nhập của khách hàng càng cao thì nhu cầu mua sắm của họ càng cao. Doanh nghiệp, cửa hàng ở các tiểu bang nên xem xét lại chiến lược marketing, các chương trình ưu đãi của mình có đang thật sự hiệu quả hay không để từ đó có các hướng giải pháp để có thể kích cầu người tiêu dùng.

Nhận xét: Từ matrix trên ta có thể thấy không phải thu nhập của khách hàng càng cao thì nhu cầu mua sắm của họ càng cao. Doanh nghiệp, cửa hàng ở các tiểu bang nên xem xét lại chiến lược marketing, các chương trình ưu đãi của mình có đang thật sự hiệu quả hay không để từ đó có các hướng giải pháp để có thể kích cầu người tiêu dùng.



State	Avg Income (\$)	Purchase Amount (\$)
Massachusetts	109,779.20	5,487.80
Hawaii	107,304.16	3,372.40
California	106,269.18	64,328.57
Connecticut	104,924.19	4,460.75
Washington	104,499.16	5,193.00
Colorado	103,364.15	4,594.00
New Hampshire	103,184.14	4,959.00
Virginia	103,104.17	3,799.00
Alaska	101,979.14	4,421.00
Utah	100,264.11	3,520.00
Minnesota	99,764.13	2,808.00
New York	99,524.19	15,675.00
Delaware	95,010.13	3,300.00
Total	4,648,010.82	286,015.80

V. Kết luận và kiến nghị

Dự án đã thành công trong việc tạo ra ba dashboard, mỗi dashboard tập trung vào một khía cạnh khác nhau của hoạt động kinh doanh: thông tin thị trường (market insight), hàng tồn kho (product inventory) và phân tích khách hàng (customer analysis). Các dashboard này cung cấp dữ liệu trực quan và phân tích có giá trị để hỗ trợ việc ra quyết định kinh doanh.

Những phát hiện chính: Các dashboard đã cho thấy một số phát hiện quan trọng:

- Đa số khách hàng thuộc tầng lớp trung lưu.

- Doanh số bán hàng cao nhất ở các bang như California và Texas, nhưng một số bang khác lại có sự sụt giảm.
- Độ tuổi, địa lý và thu nhập của khách hàng là các yếu tố quan trọng để tiếp thị.

Đề xuất chiến lược:

- Tập trung vào các chiến dịch tiếp thị ở các khu vực có thu nhập cao và các ngành cụ thể.
- Tối ưu hóa việc phân phối sản phẩm dựa trên khí hậu của từng khu vực.
- Triển khai các chương trình khuyến mãi theo mùa.
- Cải thiện trải nghiệm của khách hàng, đặc biệt là nhóm khách hàng trung lưu.
- Tập trung cải thiện chất lượng các sản phẩm có giá trị cao.
- Giảm tỷ lệ hoàn trả sản phẩm bằng cách cải thiện mô tả sản phẩm và kiểm soát chất lượng.
- Quản lý hàng tồn kho hiệu quả và quảng bá các sản phẩm có lượng tồn kho cao.
- Phát triển các chiến dịch tiếp thị cá nhân hóa dựa trên độ tuổi, thu nhập và địa lý của khách hàng.
- Xem xét lại chiến lược tiếp thị ở các bang có doanh số thấp.

Những điều có thể cải thiện:

- Phân tích sâu hơn về hành vi mua sắm của khách hàng có thể dẫn đến các chiến lược hiệu quả hơn.
- Thu thập và phân tích thêm dữ liệu để cải thiện hiệu quả của các chiến lược tiếp thị.
- Tinh chỉnh các dashboard để tập trung vào các thông tin hữu ích nhất.
- Cần có tài liệu mô tả rõ hơn về các bước xử lý dữ liệu.

Ý kiến: Dự án đã chứng minh được giá trị của việc trực quan hóa và phân tích dữ liệu trong việc đưa ra các quyết định chiến lược. Các dashboard đã cung cấp một cái nhìn toàn diện về hoạt động kinh doanh, và các khuyến nghị đưa ra là cơ sở vững chắc để cải thiện hiệu suất của công ty. Bằng cách tập trung vào các thông tin chính và tiếp tục tinh chỉnh phân tích, chuỗi cửa hàng có thể tối ưu hóa hoạt động và nâng cao sự hài lòng của khách hàng.