



TRƯỜNG ĐẠI HỌC KINH TẾ - ĐẠI HỌC ĐÀ NẴNG
KHOA THƯƠNG MẠI ĐIỆN TỬ
NGÀNH KHOA HỌC DỮ LIỆU VÀ PHÂN TÍCH KINH DOANH

BÁO CÁO CUỐI KÌ

PHÂN TÍCH VÀ TỐI ƯU HÓA HOẠT ĐỘNG KINH DOANH CỦA DATACO

GV hướng dẫn:

Nhóm:

Thành viên:

1. Phan Trường Huy
2. Tiêu Thanh Mai Anh
3. Phạm Hoàng Phương Anh

ThS Trần Danh Nhân

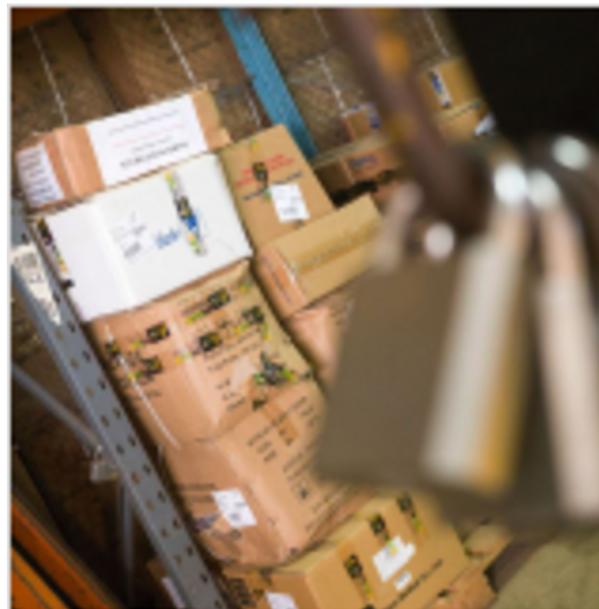
DV103 - 48K29.1

Thành viên nhóm

Thành viên	Mã sinh viên	Lớp
Phạm Hoàng Phương Anh	221124029102	48K29.1
Tiêu Thanh Mai Anh	221124029101	48K29.1
Phan Trường Huy	221124029119	48K29.1

BỐI CẢNH DỮ LIỆU

- Bộ dữ liệu bao gồm hơn 52 cột và 180,000 dòng.
- Tập hợp thông tin chi tiết về các đơn hàng được thực hiện trong một hệ thống bán lẻ. Dữ liệu cung cấp thông tin về giao dịch, khách hàng, sản phẩm, đơn hàng và quá trình vận chuyển. Đây là bộ dữ liệu đa chiều, giúp phân tích hiệu quả hoạt động, tình trạng giao hàng và hành vi mua sắm của khách hàng
- Nguồn dữ liệu:

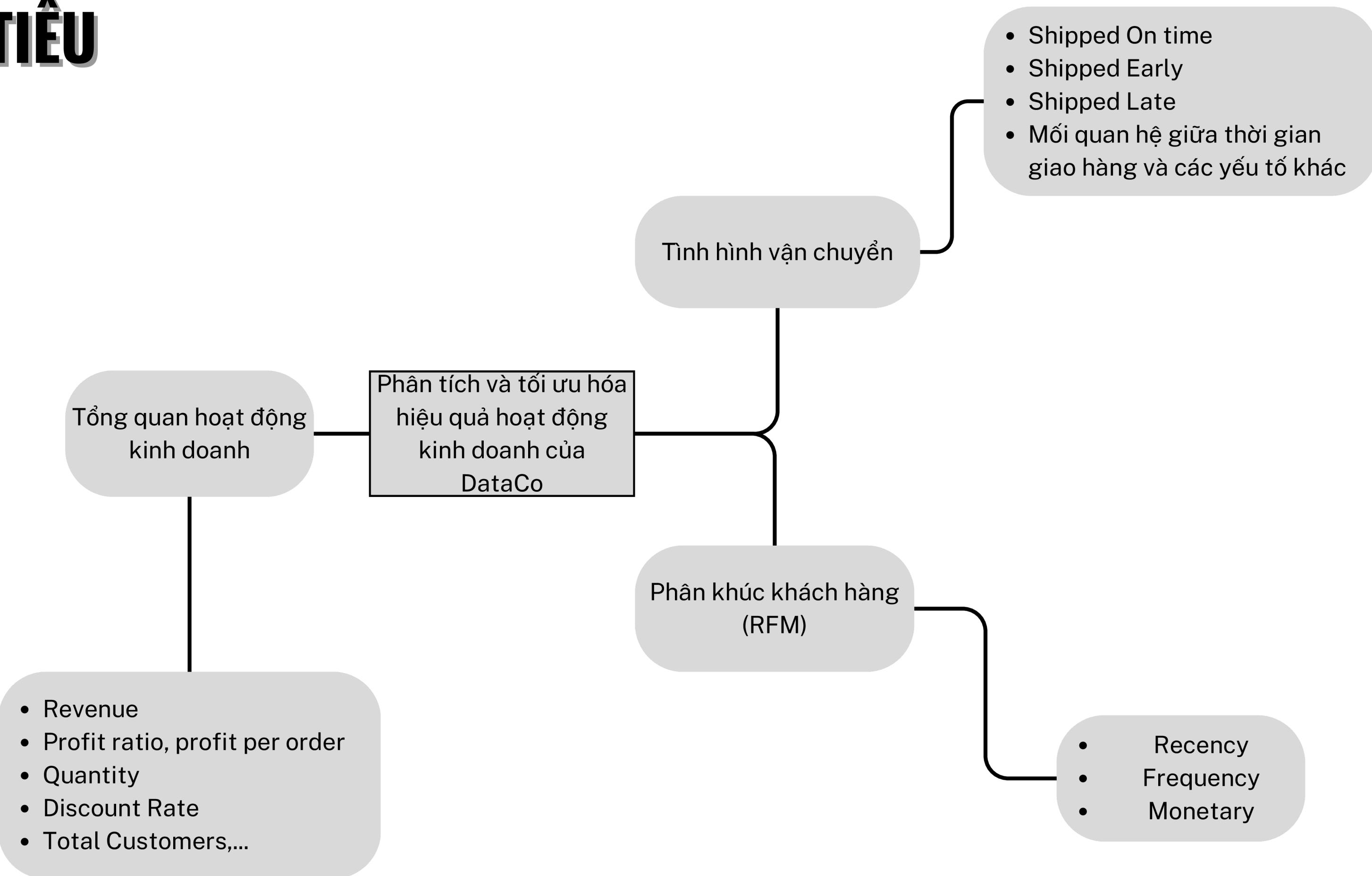


DataCo SMART SUPPLY CHAIN FOR BIG DATA ANALYSIS

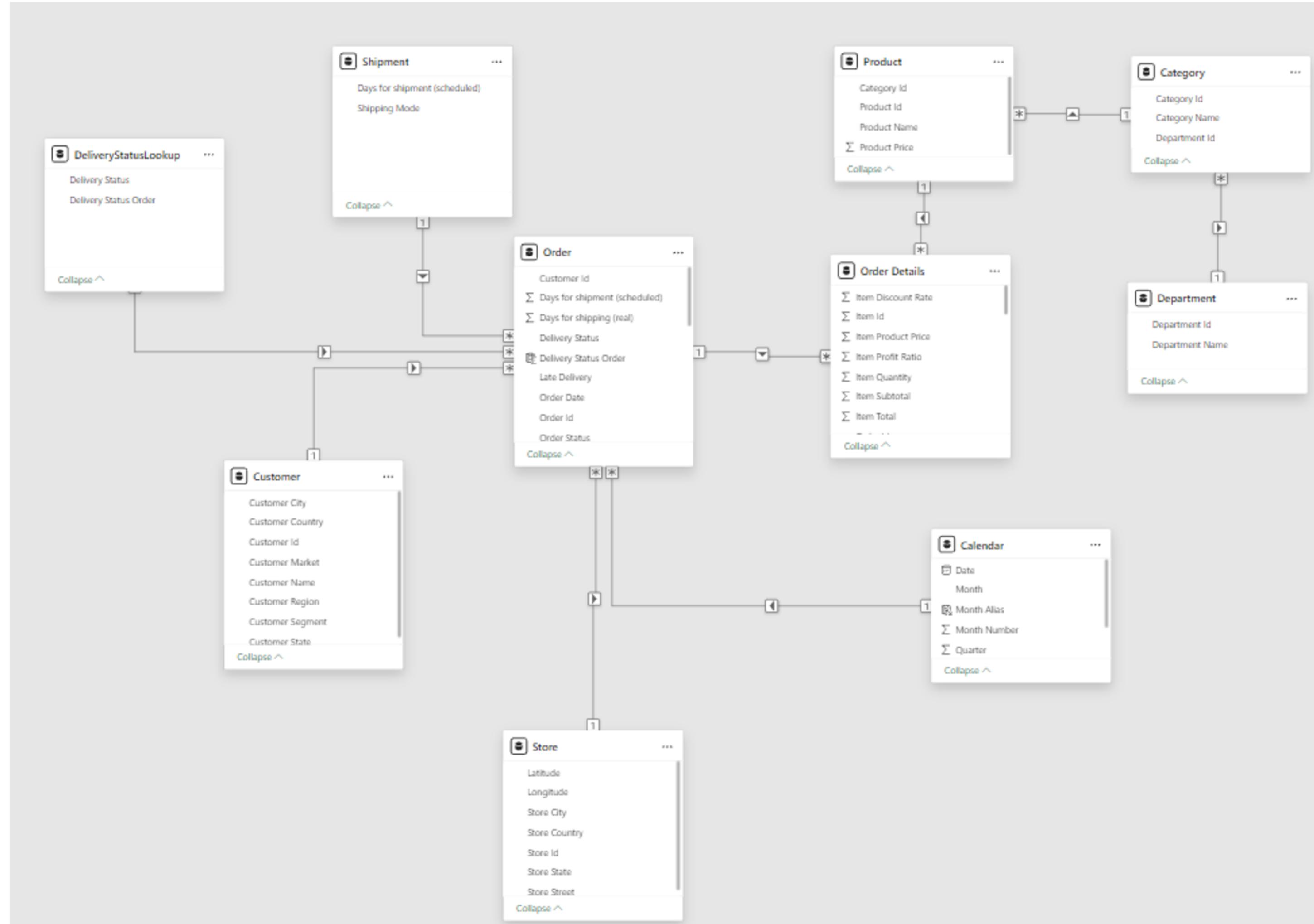
SUPPLY CHAIN FOR BIG DATA ANALYSIS

 kaggle.com

MỤC TIÊU



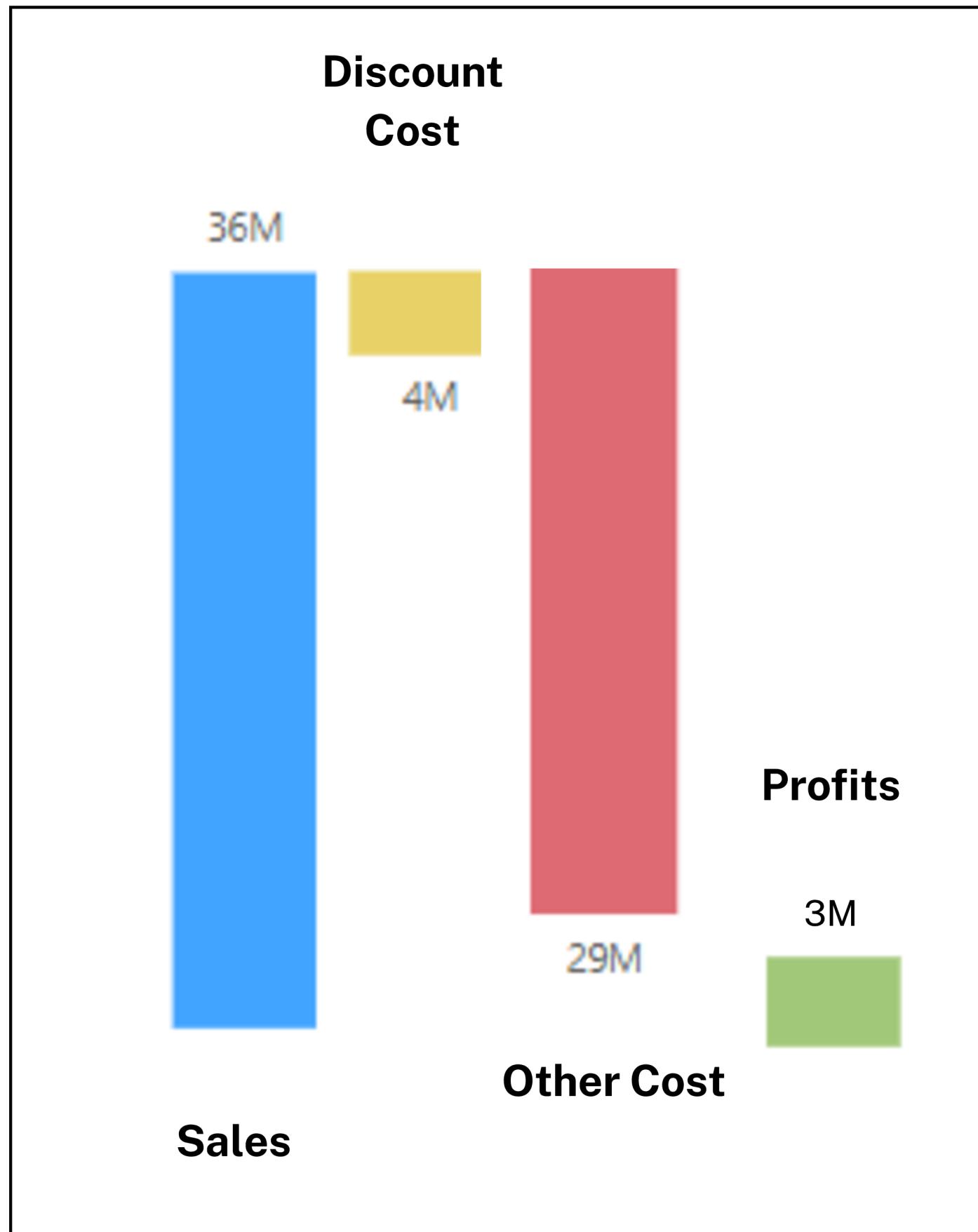
MÔ HÌNH HÓA DỮ LIỆU



TỔNG QUAN TÌNH HÌNH KINH DOANH



Tổng quan cơ cấu doanh thu, chi phí và lợi nhuận

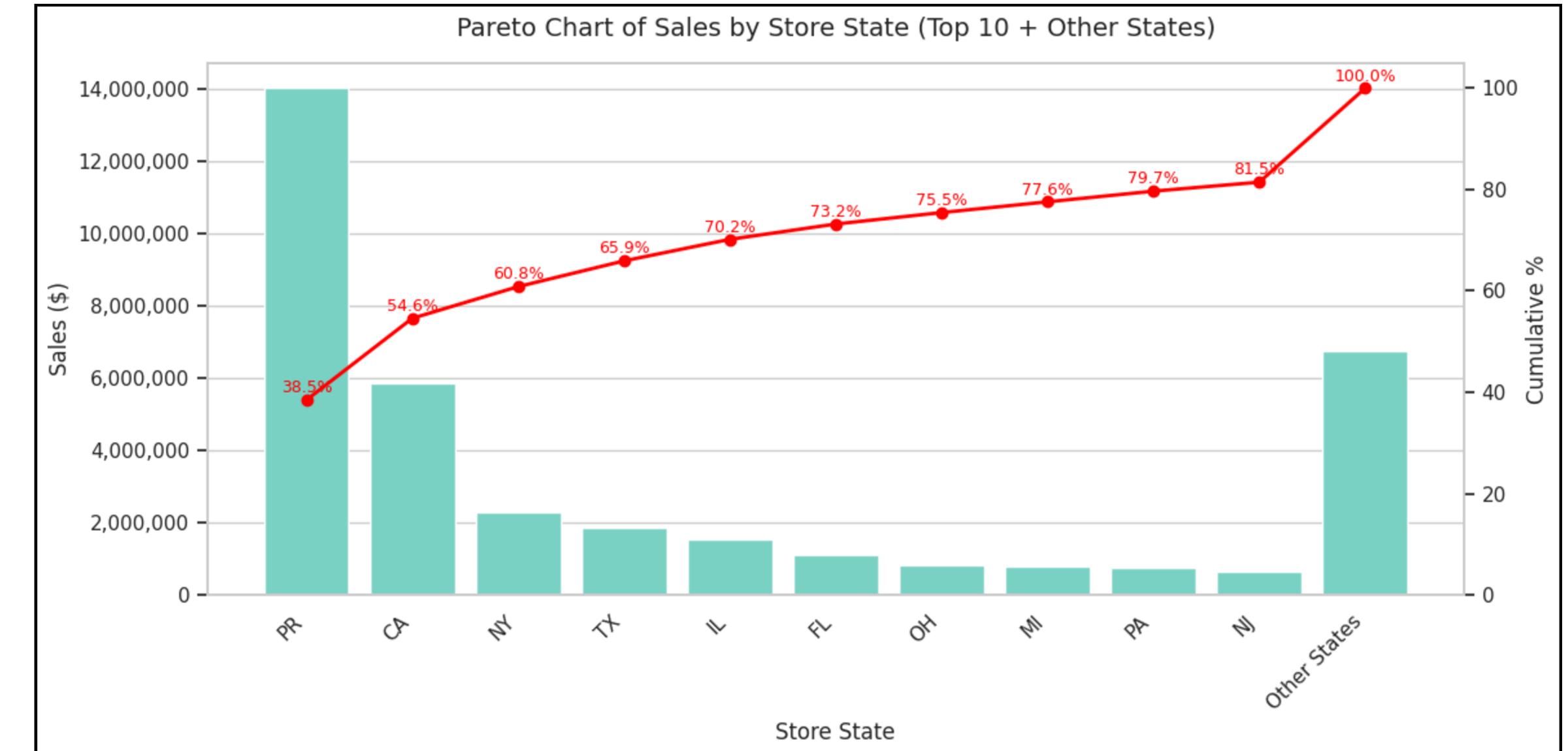
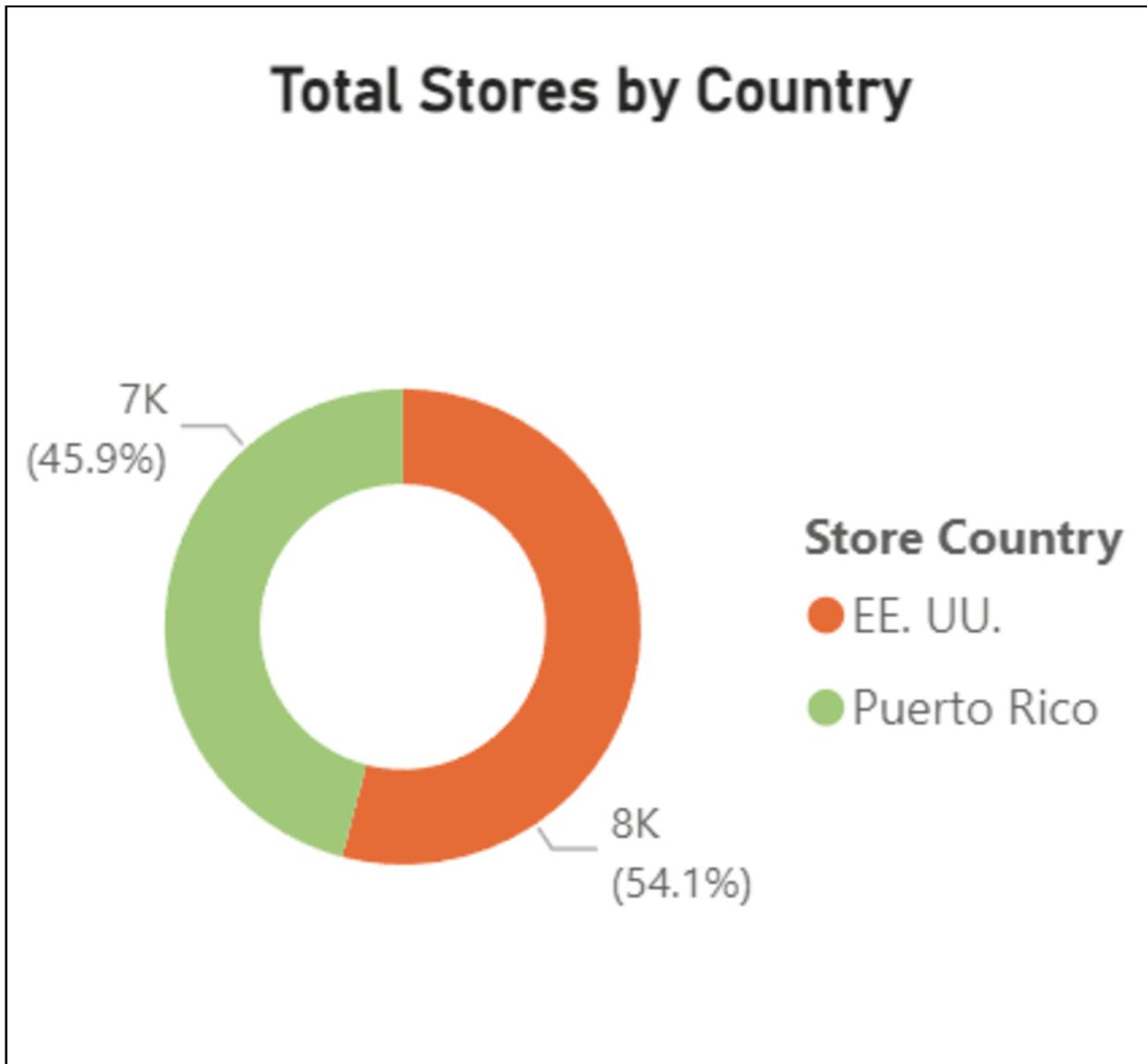


Cơ cấu chi phí so với doanh thu qua 3 năm:

- Chi phí chiết khấu chiếm 11,11%
- Những chi phí khác chiếm 80,6%
- Tỷ suất lợi nhuận trên doanh thu khoảng 8.33%.

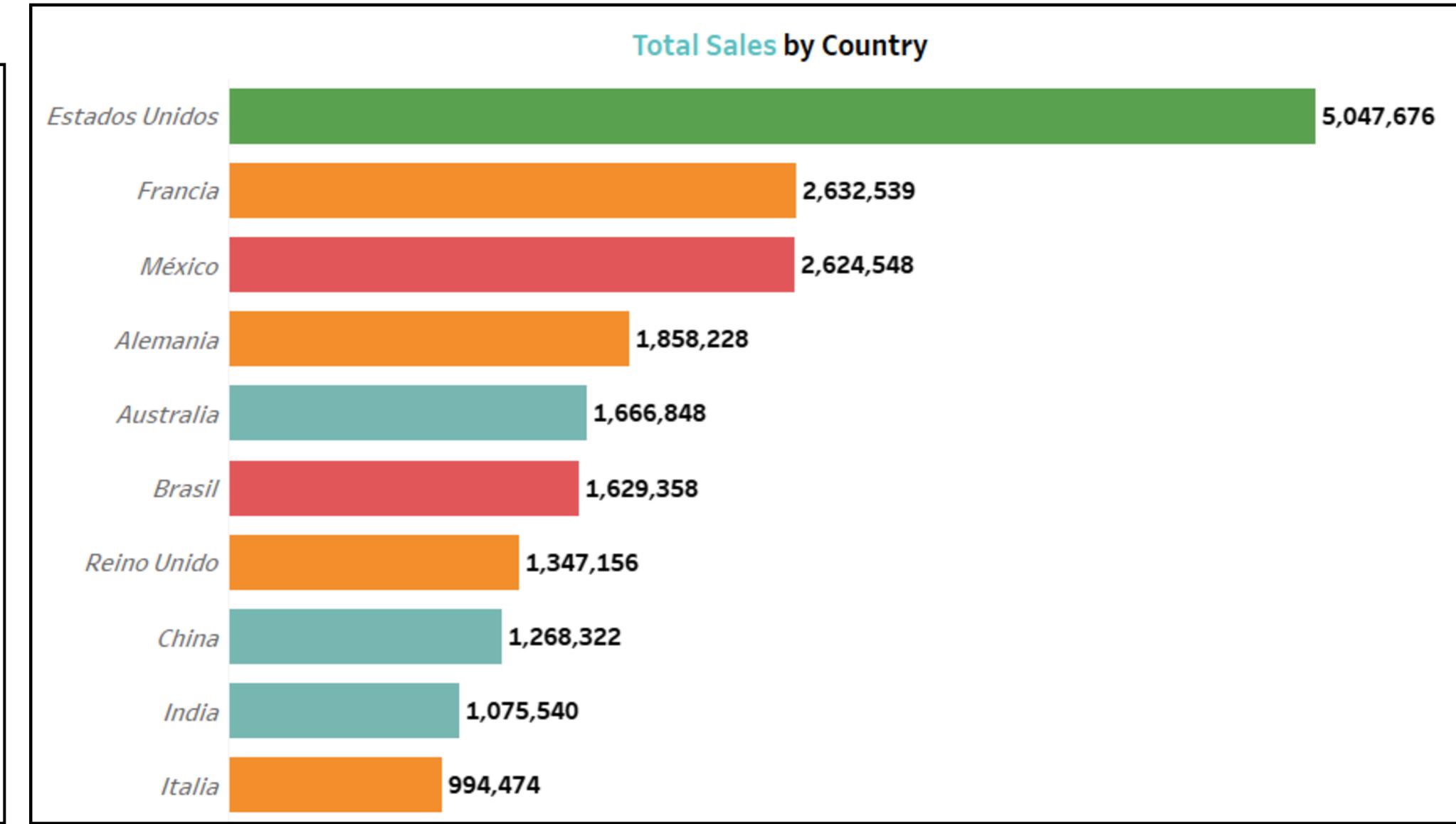
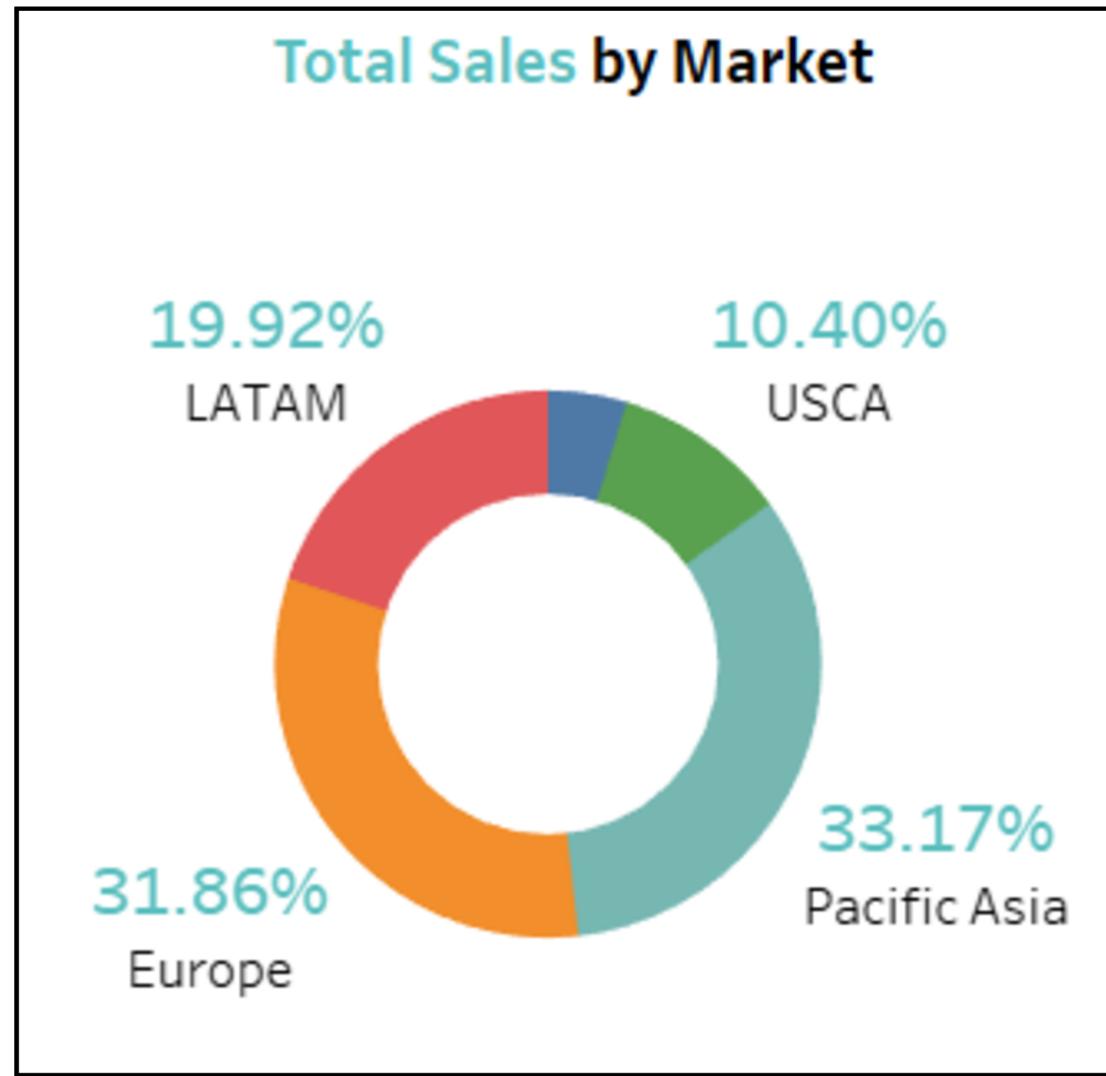
Mức lợi nhuận ròng khá thấp với ngành bán lẻ

Quy mô của doanh nghiệp



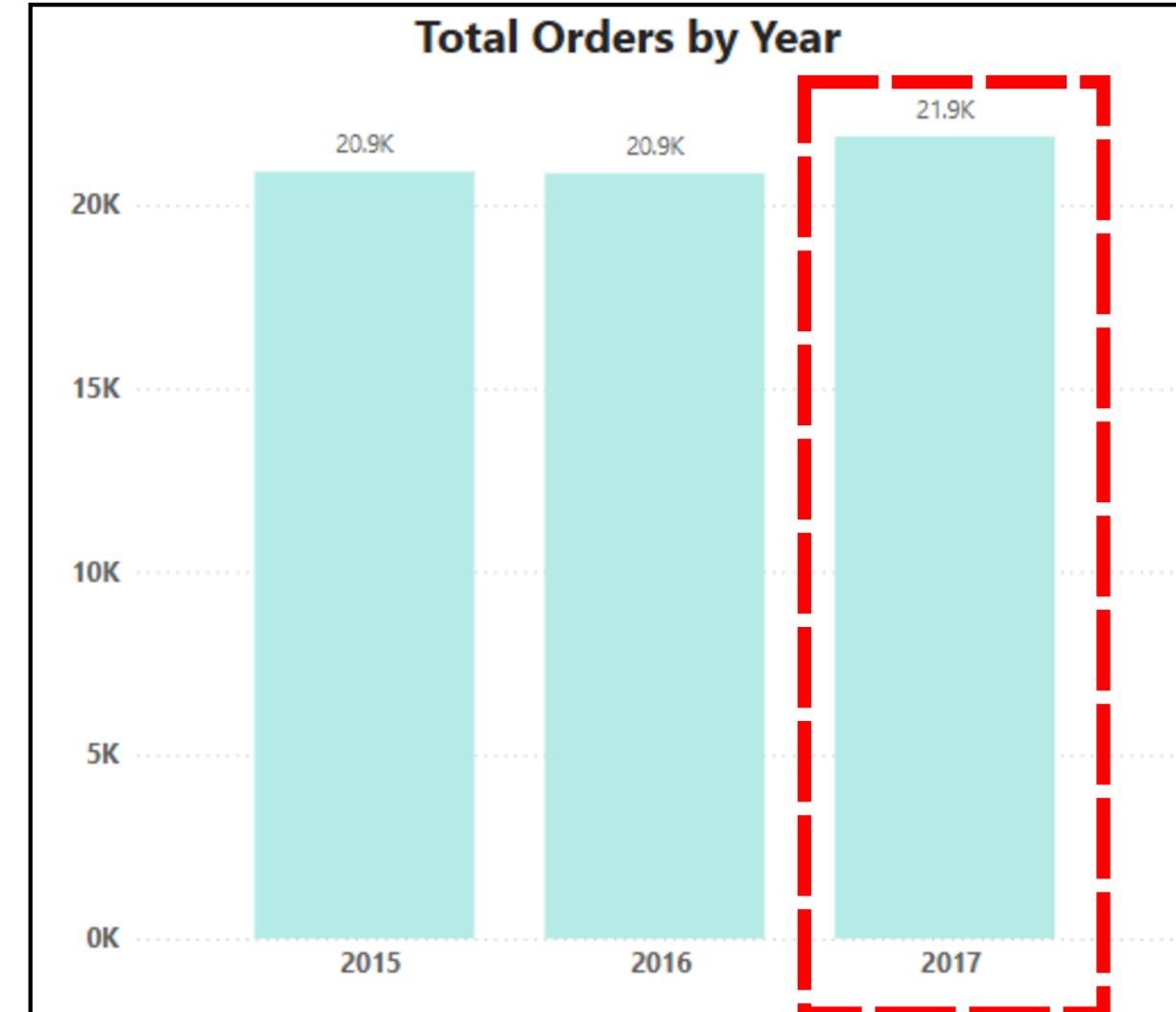
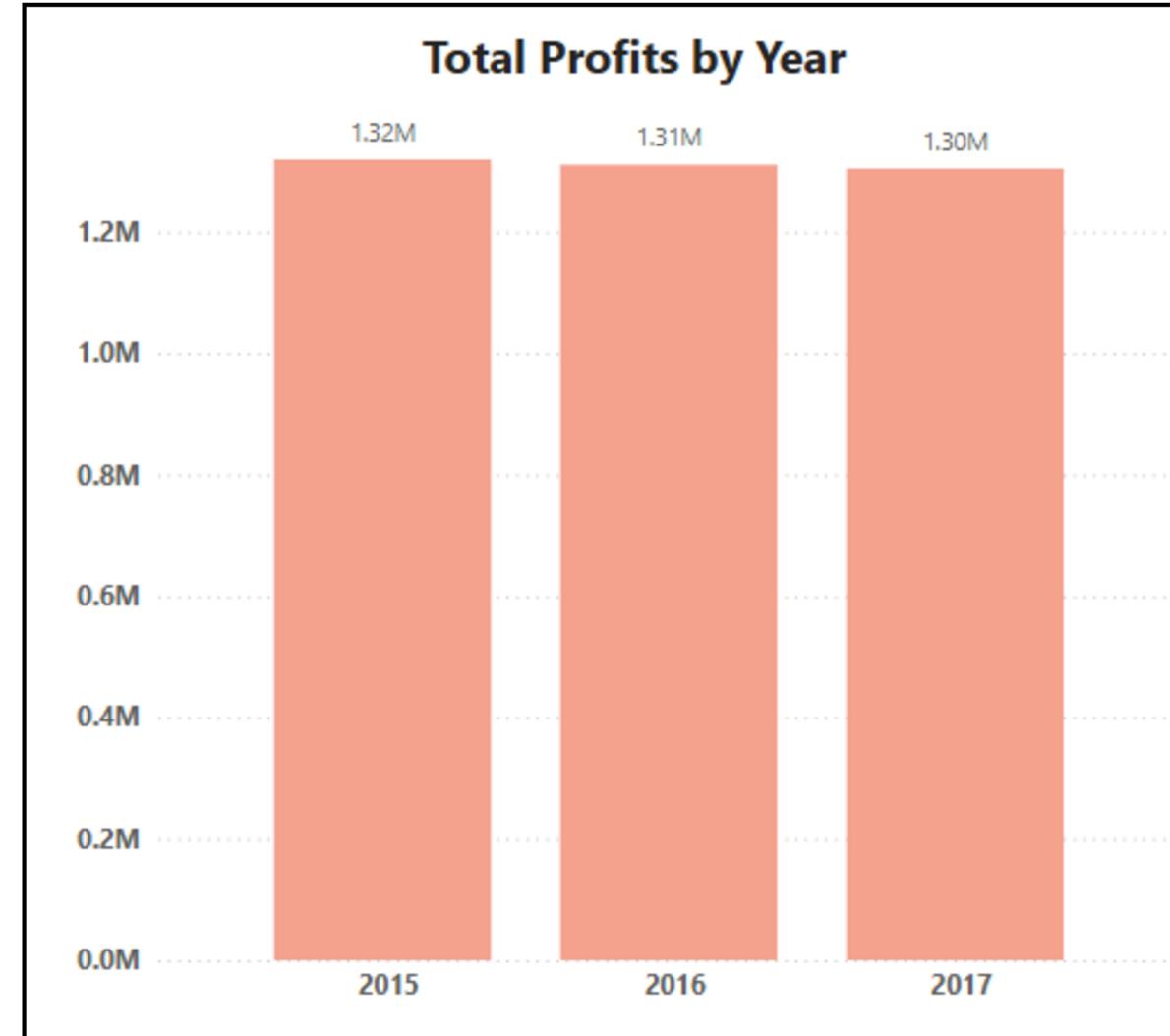
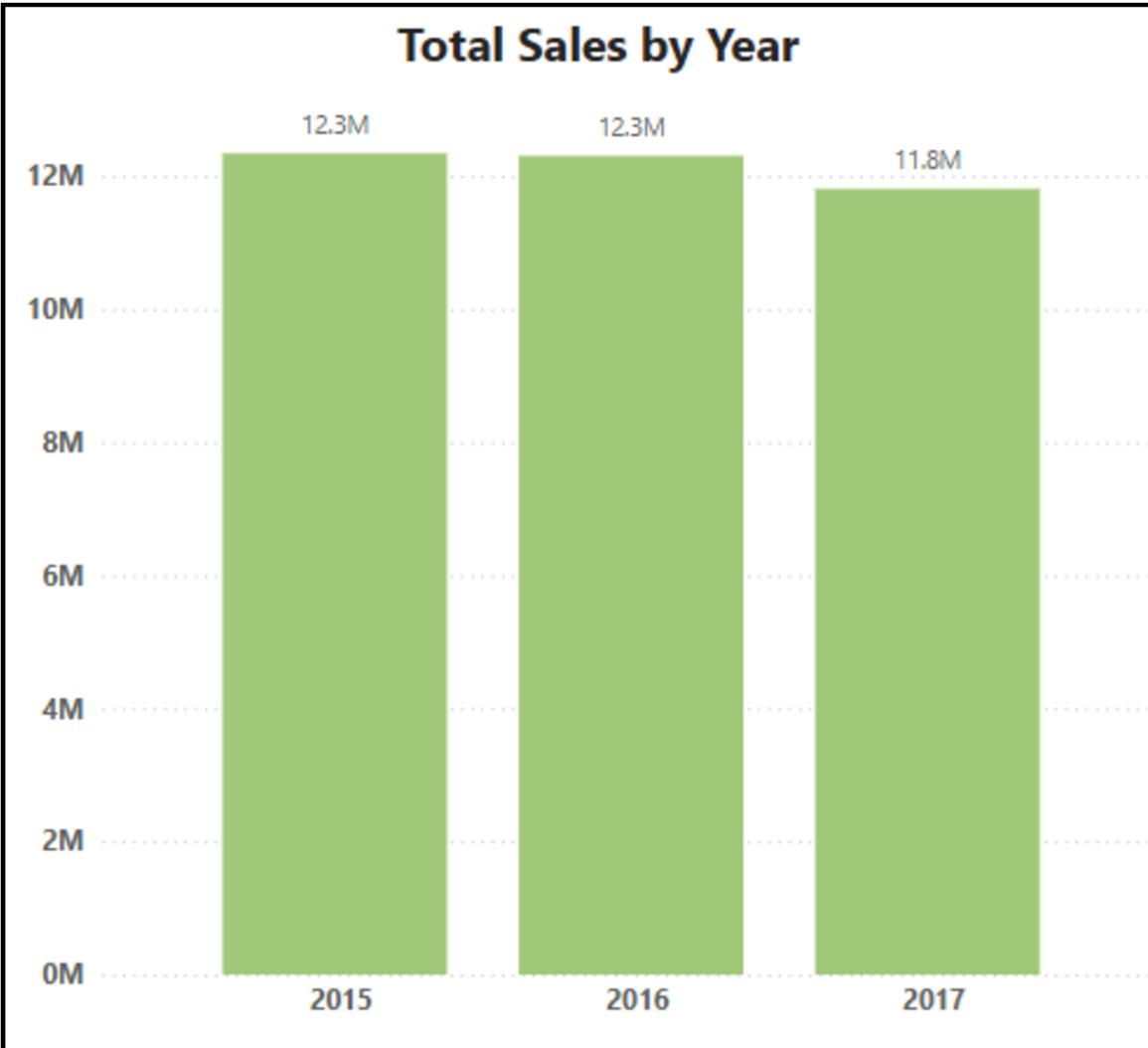
- Đóng góp bởi hơn 15K cửa hàng, trong đó Mỹ chiếm **54.1%** tổng số cửa hàng
- Xét về đóng góp của các tiểu bang, **tiểu bang Puerto Rico** của Puerto Rico đóng góp tỉ lệ vào tổng doanh thu **lớn nhất** với hơn **38.5%** doanh thu
- Xếp sau đó đa số là các tiểu bang của Mỹ: California, New York, Texas...

Giá trị của các thị trường



- **Châu Á-TBD (33.17%) & Châu Âu (31.86%)** dẫn đầu doanh thu theo khu vực.
- Dù **USCA** chỉ chiếm 10.40% thị phần khu vực, Hoa Kỳ (5M) là quốc gia mang lại doanh thu cao nhất tuyệt đối.
 - **Châu Âu:** Dẫn dắt bởi Pháp (2.6M) & Đức (1.8M).
 - **Châu Á-TBD:** Nổi bật với Úc (1.6M), Trung Quốc (1.2M), Ấn Độ (1M).
 - **LATAM:** Mexico (2.6M) & Brazil (1.6M) đóng góp chính

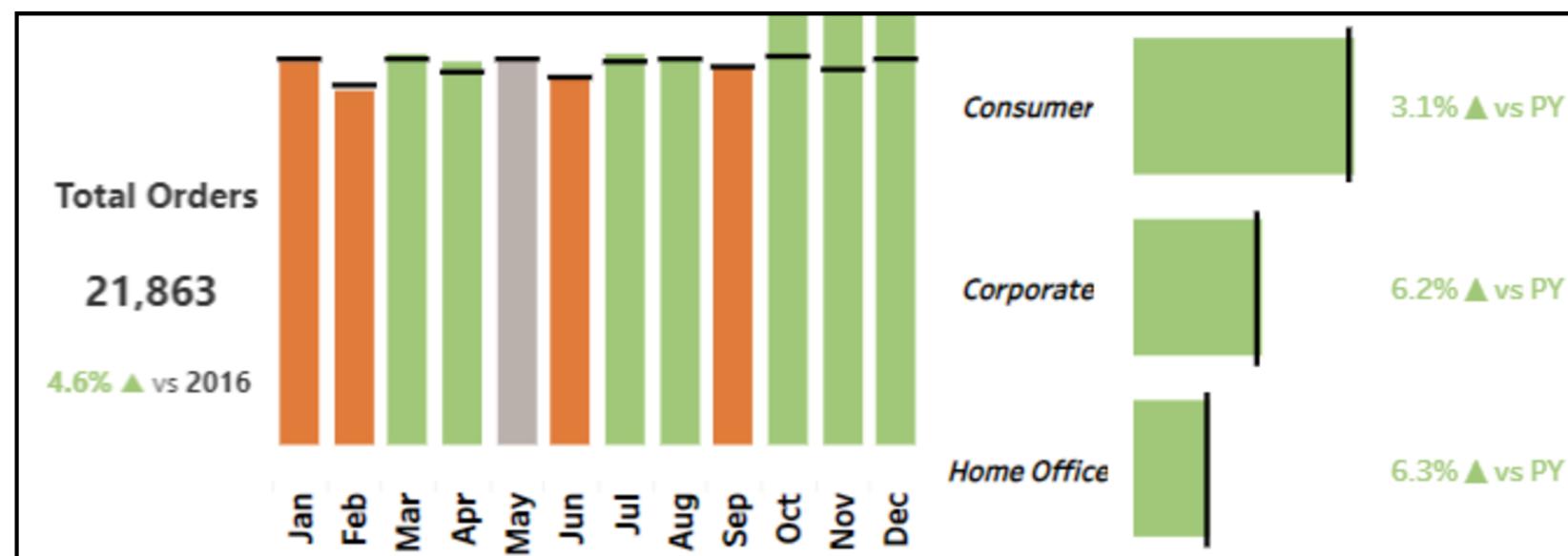
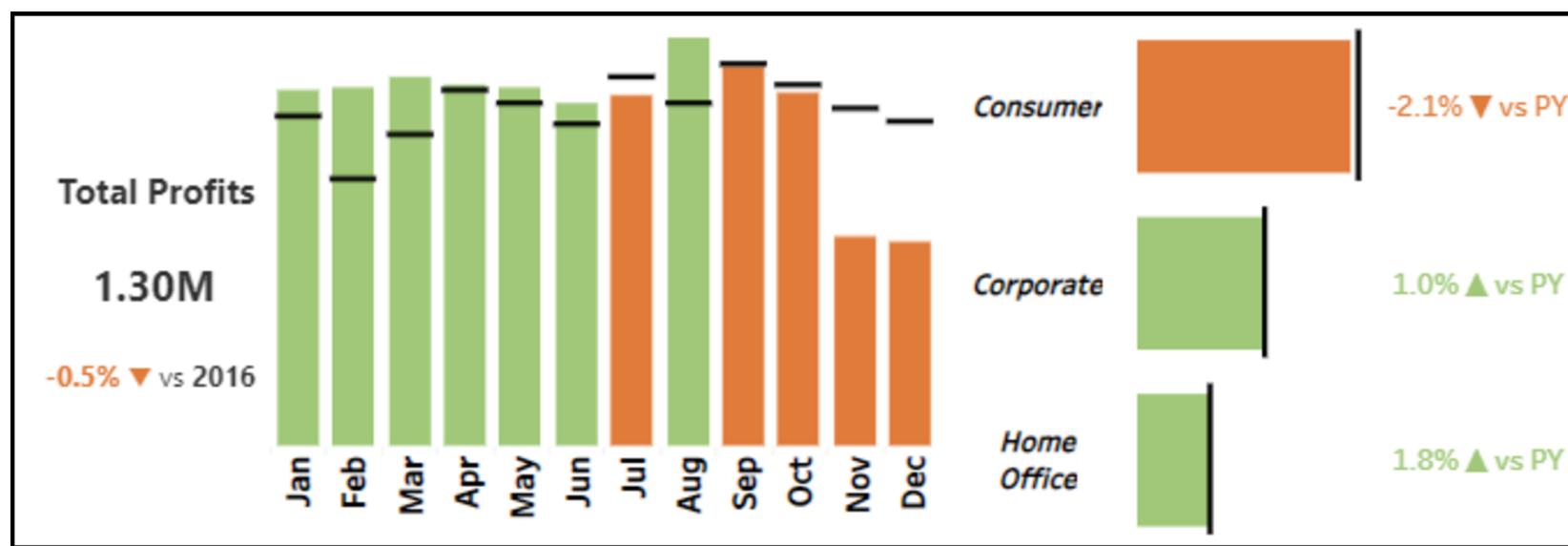
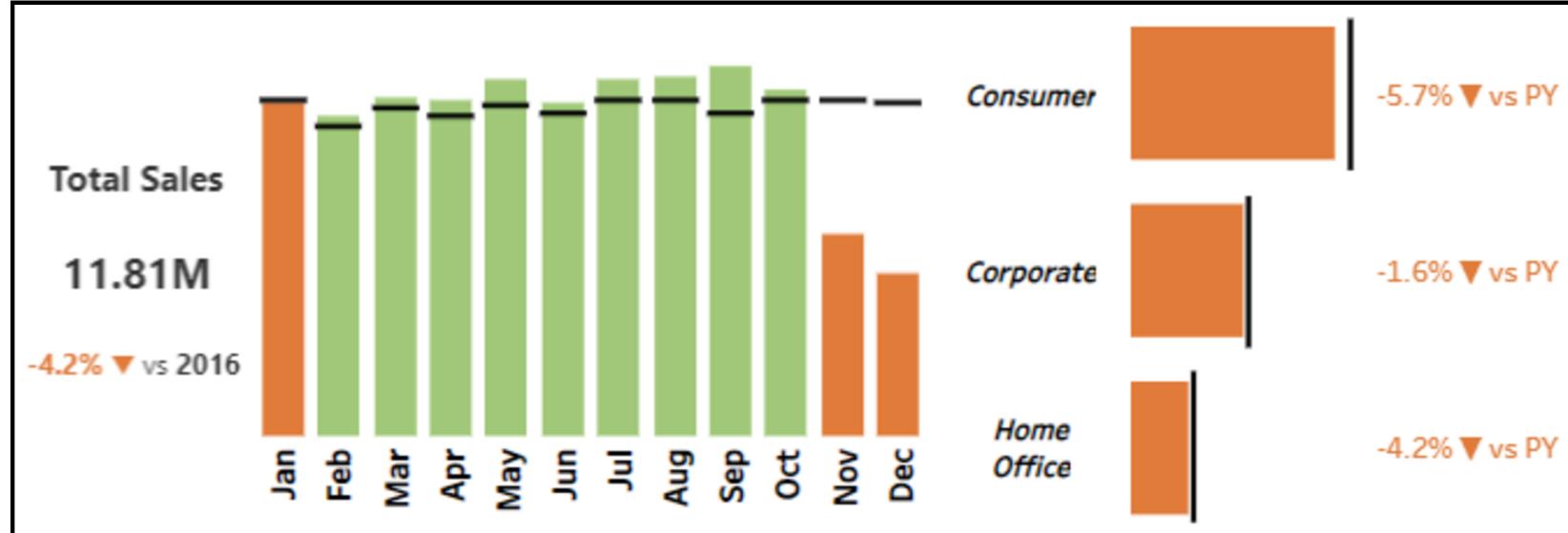
Theo thời gian



Tình hình kinh doanh, bán lẻ của DataCo **không có sự biến động lớn qua các năm**

Tuy vậy, đáng chú ý là năm 2017, doanh thu và lợi nhuận **giảm** so với 2 năm trước, mặc dù số lượng đơn hàng **tăng**

NĂM 2017

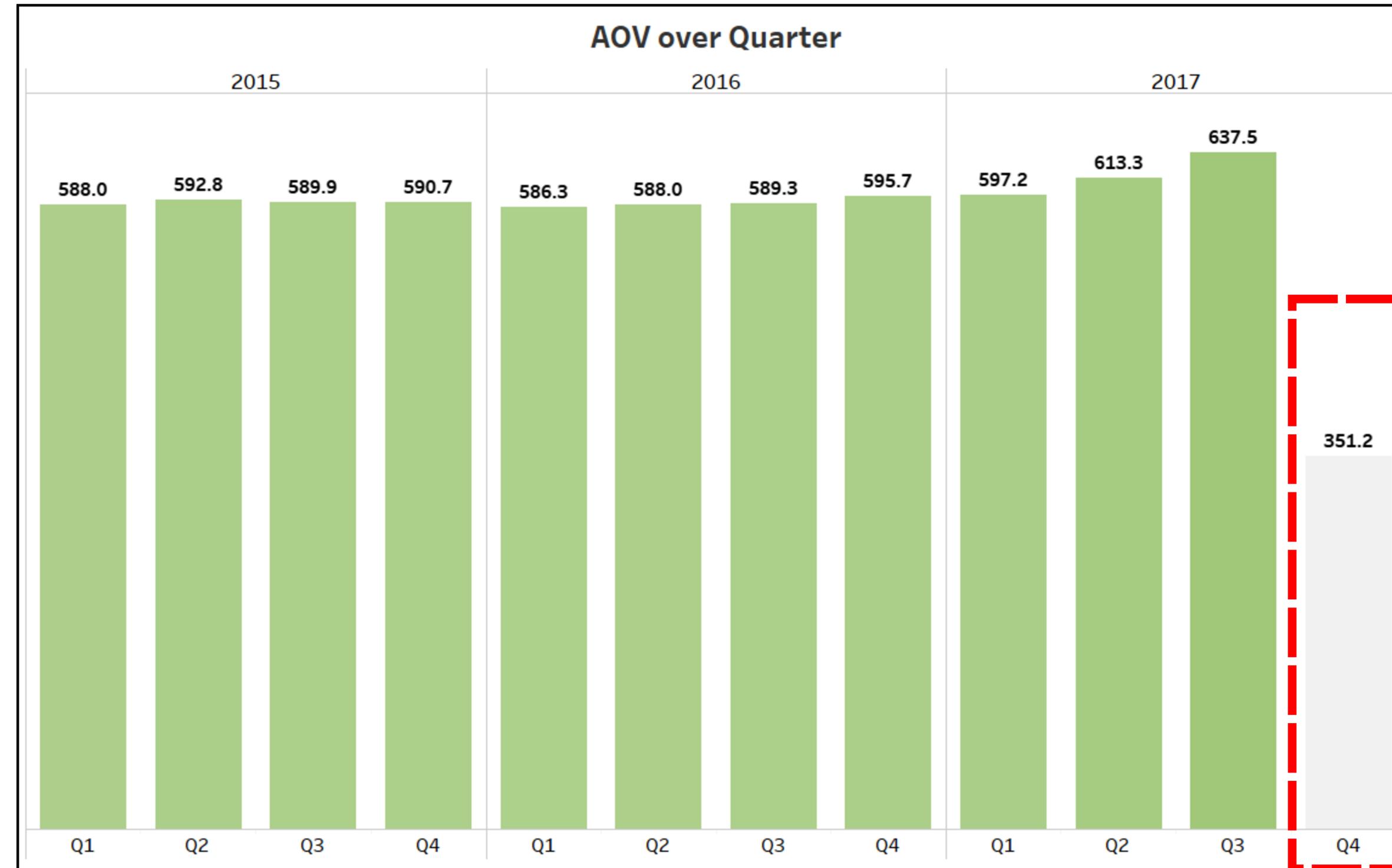


Trong năm 2017, doanh thu và lợi nhuận **giảm** do sự giảm mạnh của 2 tháng cuối năm. Sự sụt giảm này lớn hơn phần tăng của cả các tháng trước đó

Mặc dù vậy, số lượng đơn hàng lại **tăng lên 4.6%**,

- **Tăng mạnh ở Quý 3**
- **Tăng đều ở tất cả nhóm khách hàng**

NĂM 2017

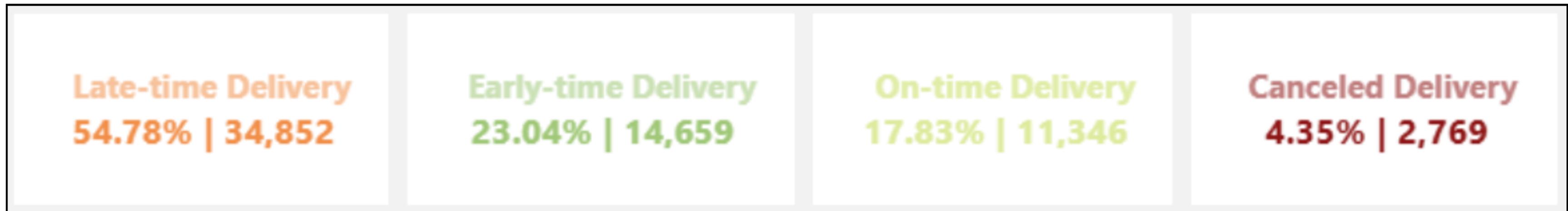


Giá trị trên một đơn hàng đã giảm so với năm trước
Giải thích cho việc giảm của doanh thu và lợi nhuận



**PHÂN TÍCH HIỆU SUẤT
GIAO HÀNG**

Tổng quan tình trạng giao hàng

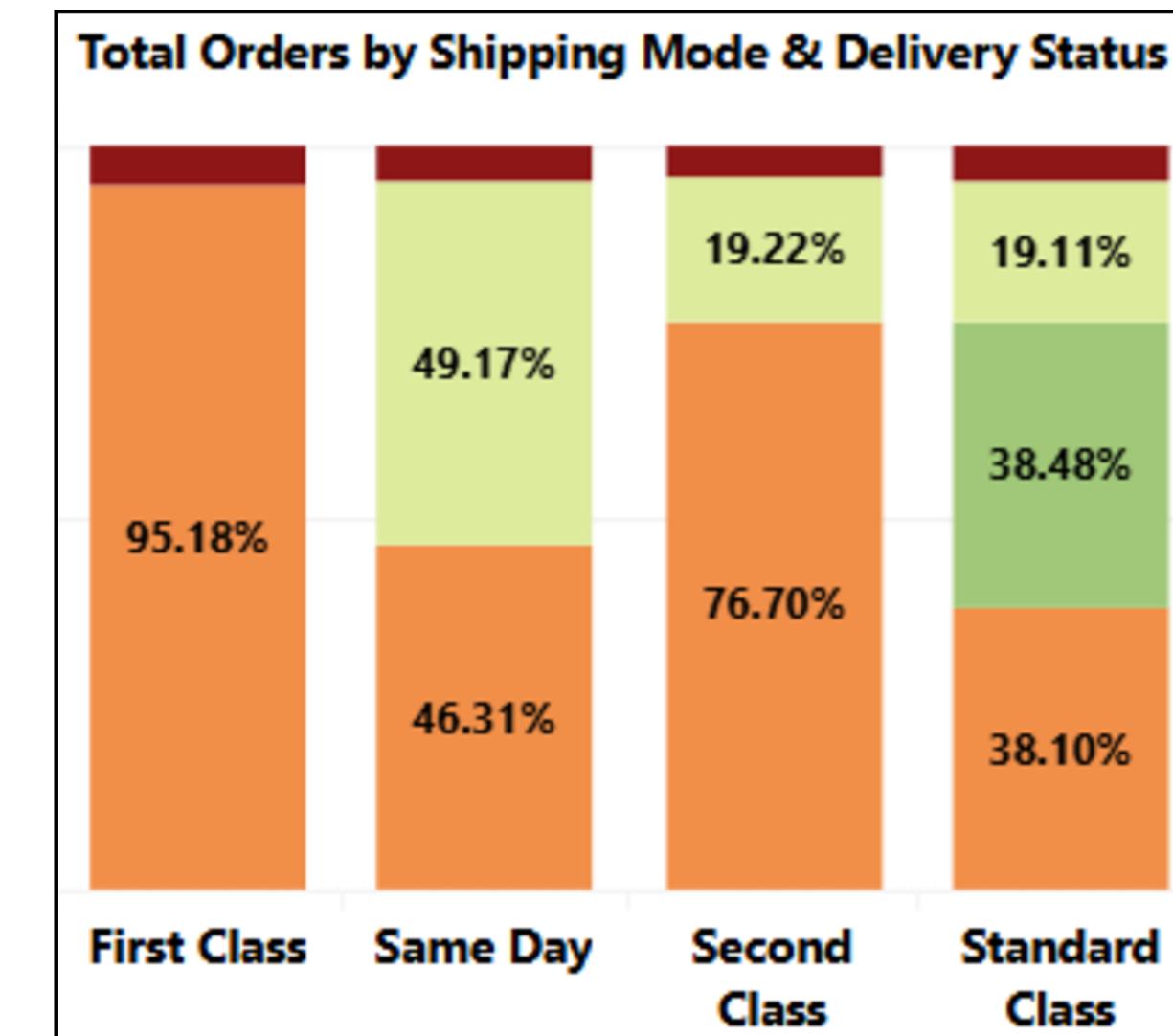
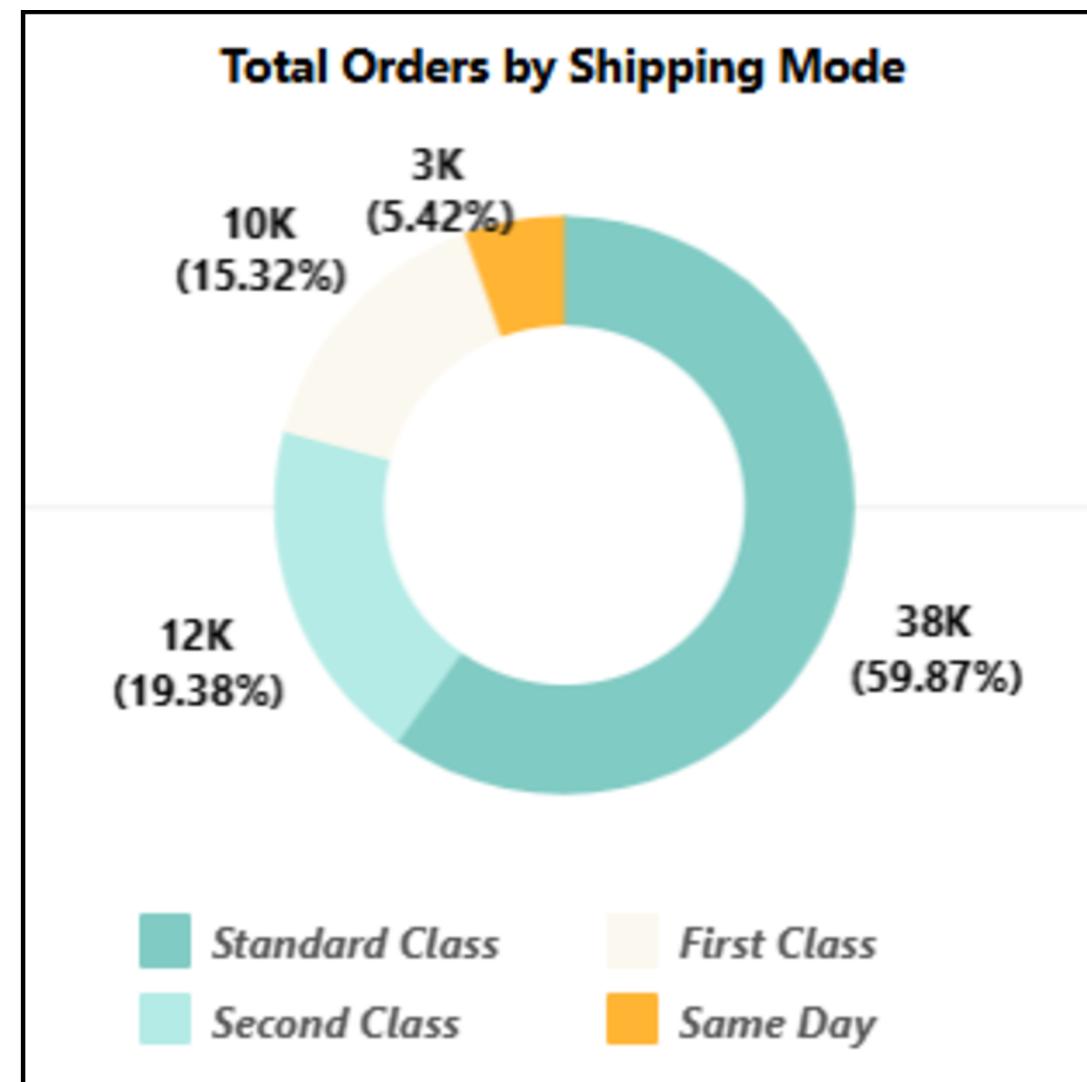


- Tỷ lệ hoàn thành đơn hàng là **95.65%**
- Hiệu suất giao hàng đáng báo động với **54,78% tỉ lệ giao hàng trễ**

Tuy tỷ lệ hoàn thành đơn cao, nhưng tình trạng giao trễ **rất nghiêm trọng**



Tổng quan tình trạng giao hàng



- **Standard Class** chiếm ưu thế với **59.87%**
- **Same Day** ghi nhận số đơn ít nhất, chỉ **5.42%**

Tình trạng giao hàng trễ diễn ra ở tất cả các phương thức:

- Gần như **toàn bộ** đơn hàng của **First Class** đều trễ
- **Ship đúng hạn** cao nhất tại **Same Day** nhưng tình trạng trễ vẫn **đáng báo động**

Tổng quan tình trạng giao hàng

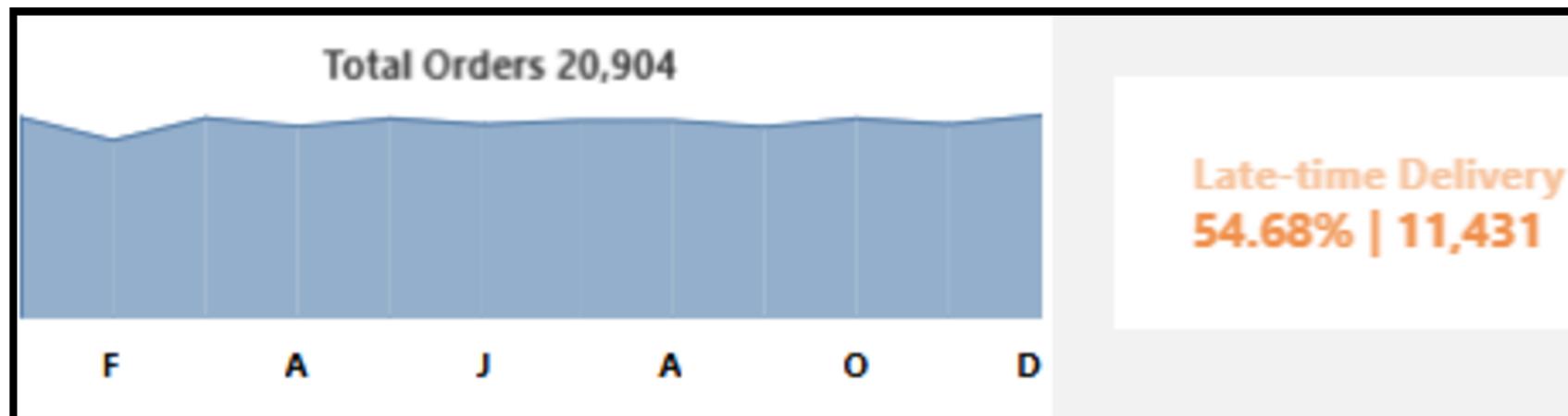


- Thời gian giao hàng dự kiến luôn **cao hơn** so với dự kiến ở **hầu hết các phương thức**
- **Second Class** trễ đến **2 ngày**

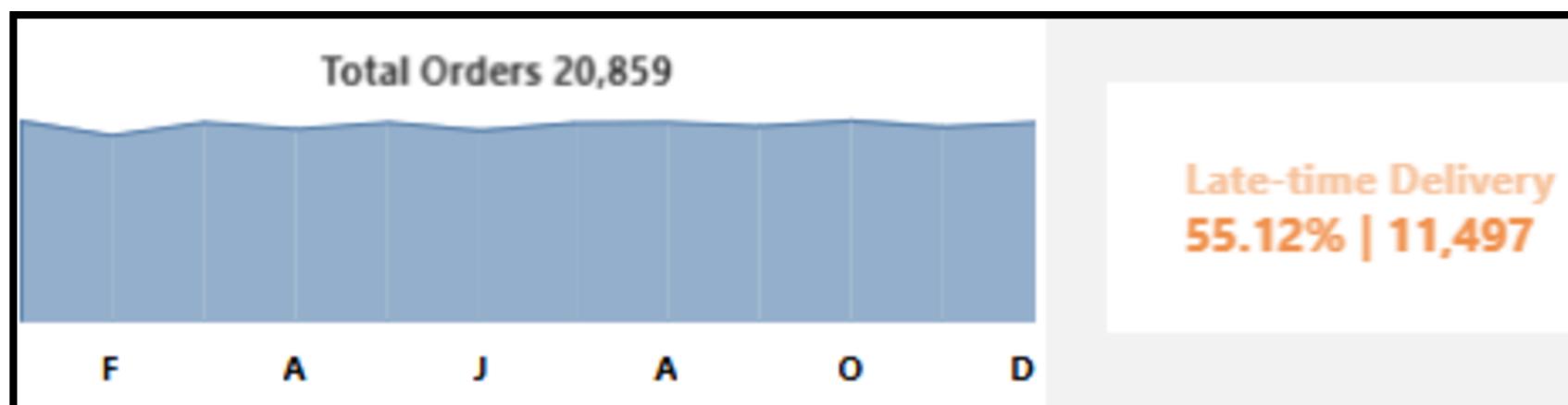
Phản ánh tình trạng giao hàng trễ đáng lo ngại, diễn ra ở mọi thời điểm

Tổng quan tình trạng giao hàng

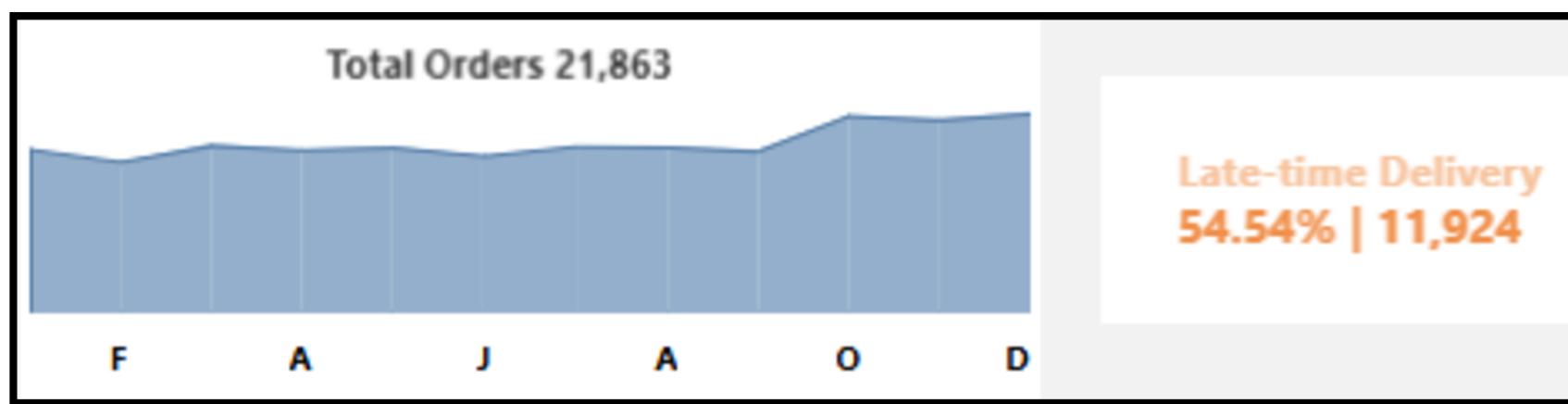
2015



2016

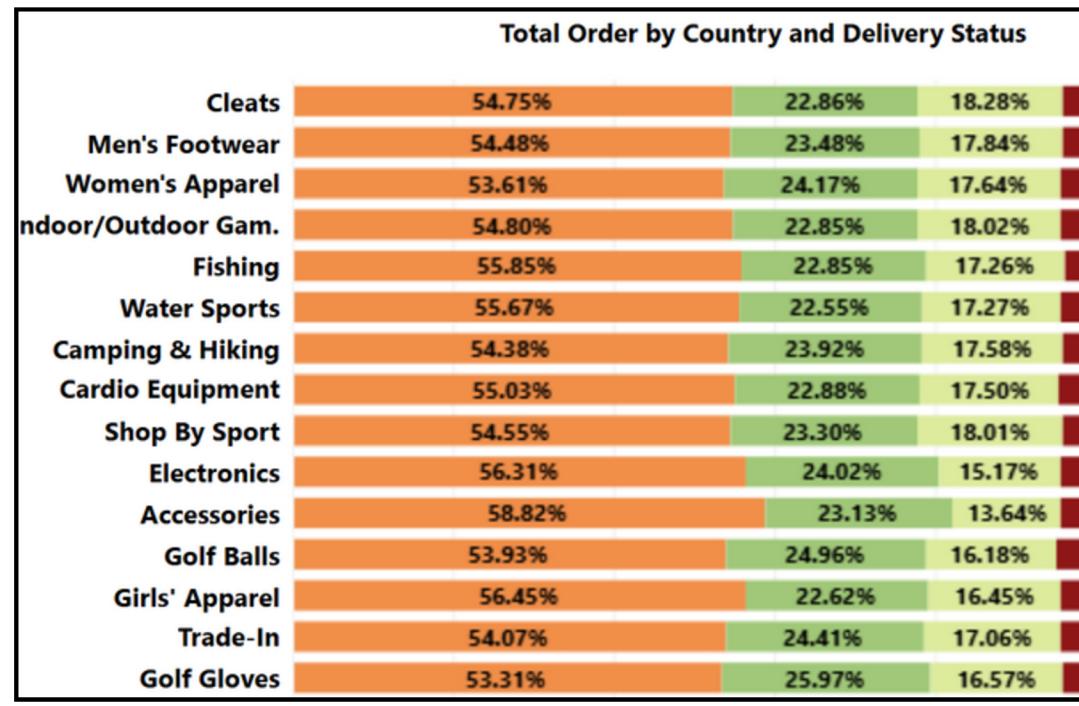


2017

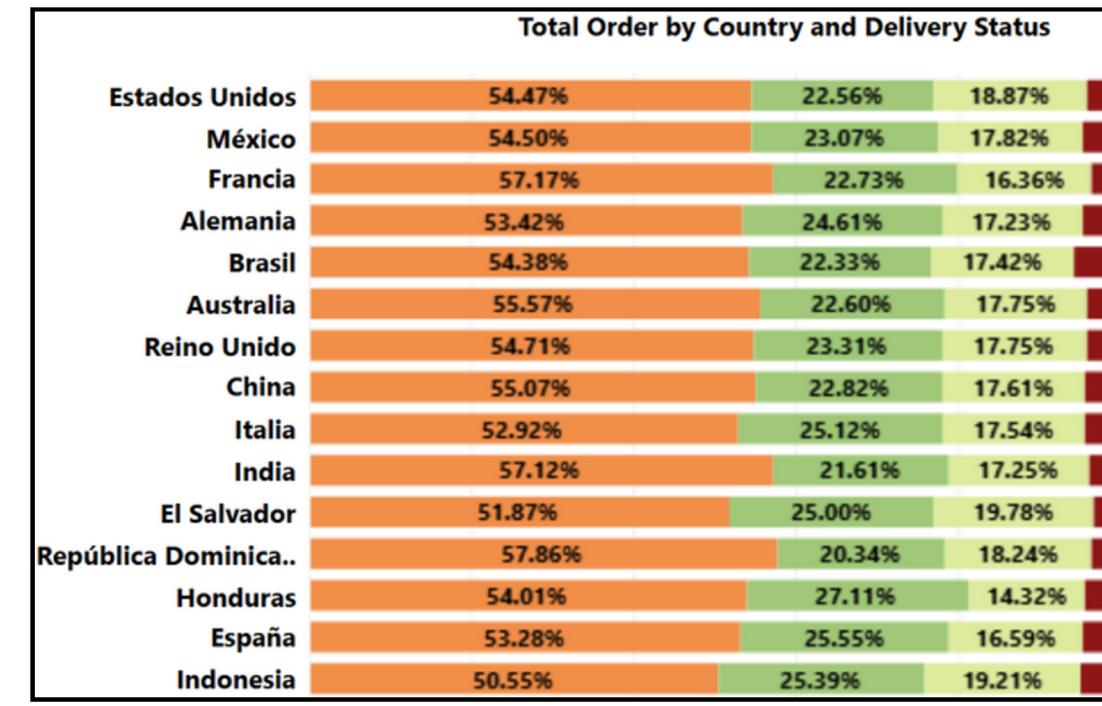


Số đơn hàng tăng thì số lượng đơn hàng trễ cũng tăng theo

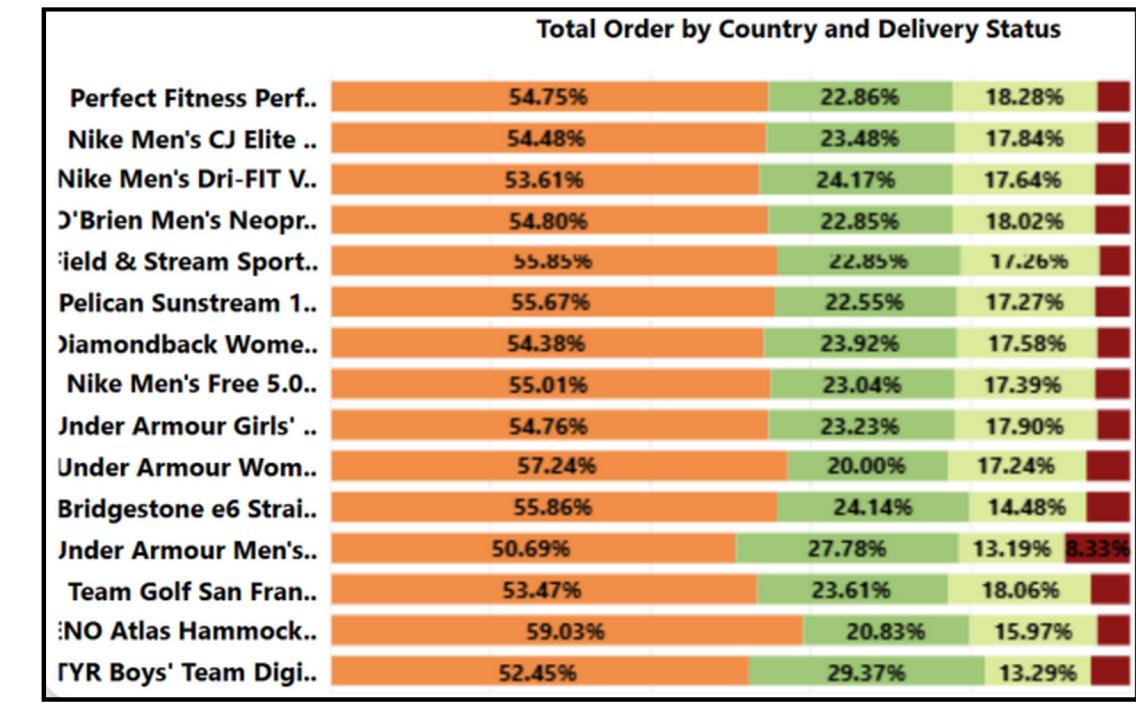
Tổng quan tình trạng giao hàng



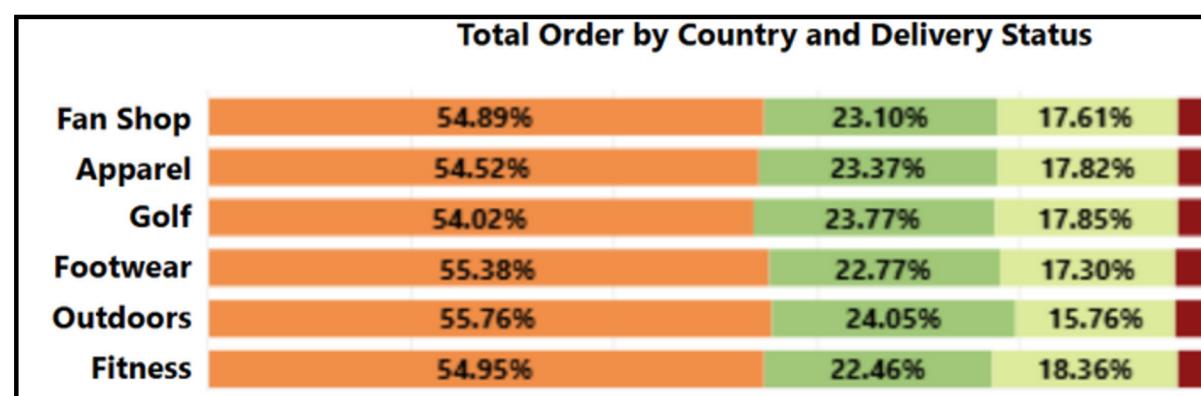
Category



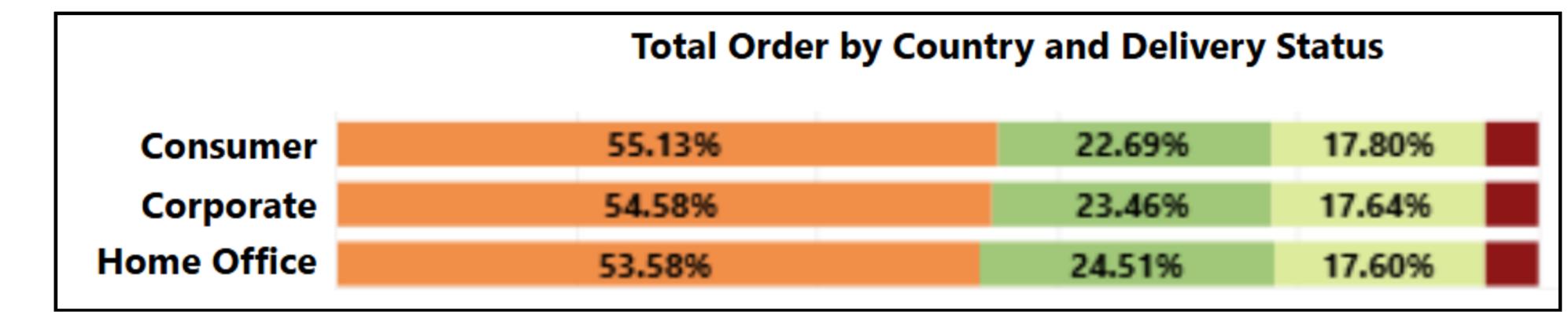
Country



Product



Department



Segment

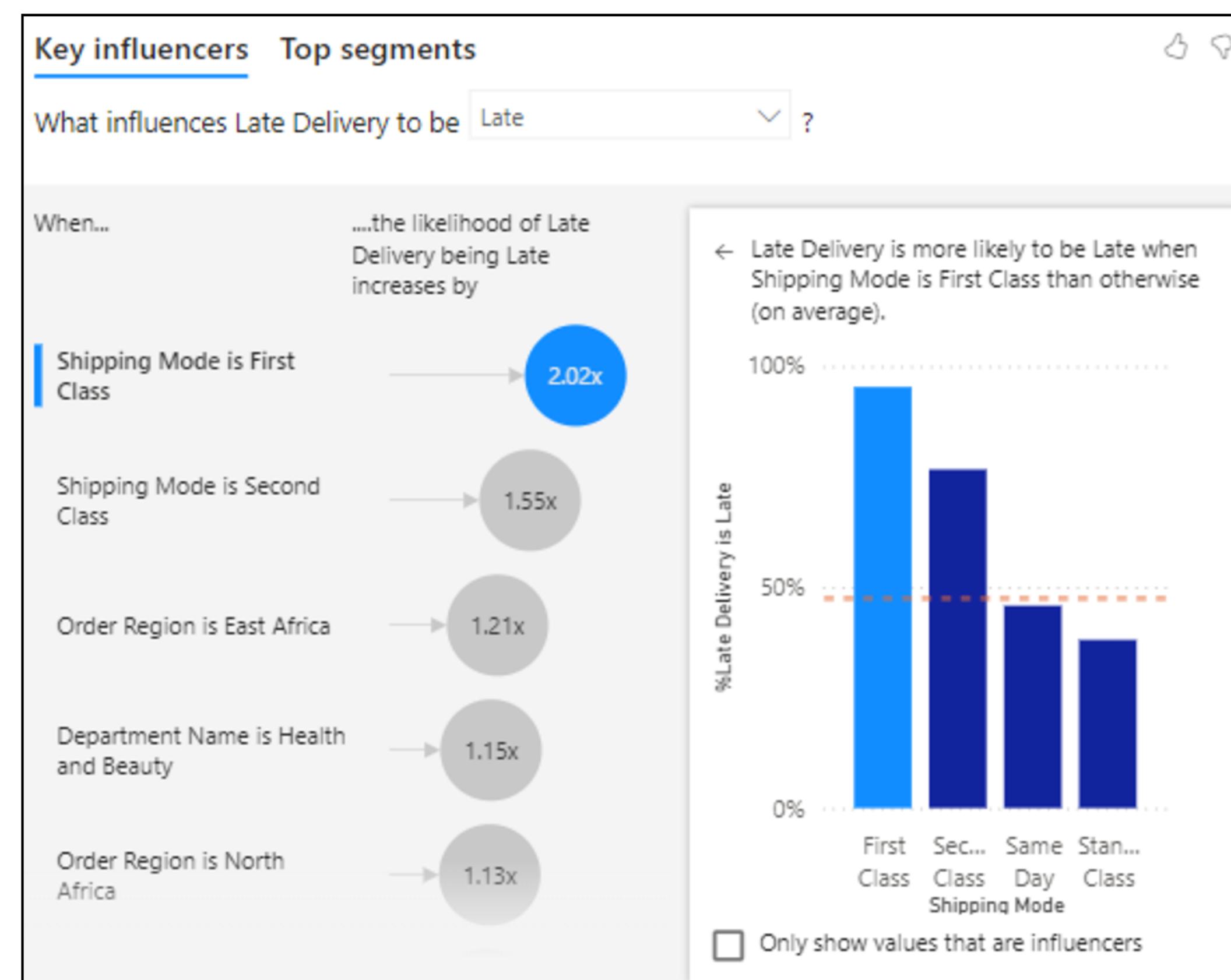


Tình trạng giao hàng trễ diễn ra gần như đồng ở hầu hết các chiều

Nhân tố ảnh hưởng chính đến tình trạng giao hàng trễ

Explain by

- Market
- Shipping Mode
- Order Region
- Department Name

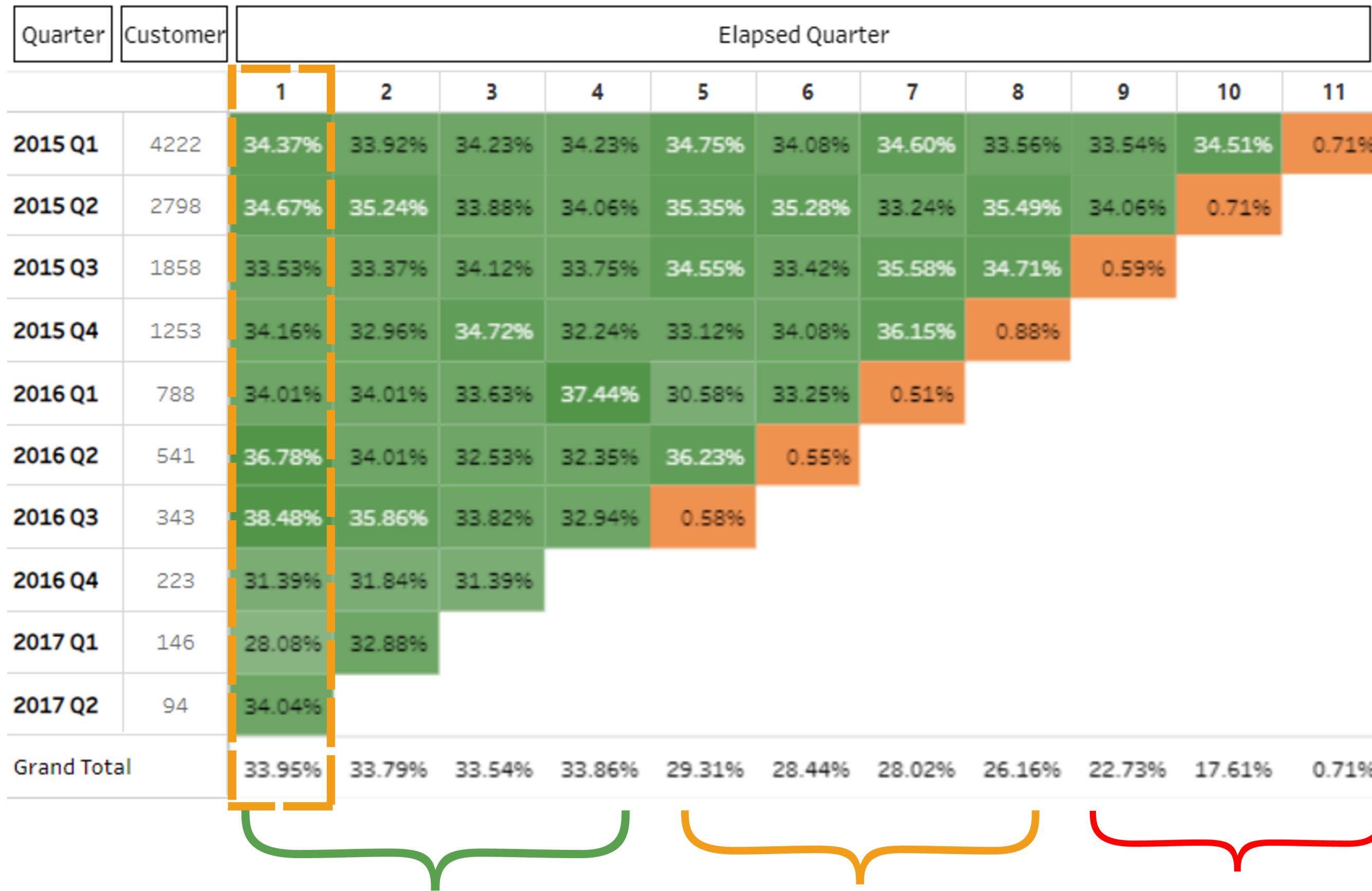


KẾT LUẬN

THỰC TRẠNG & NGUYÊN NHÂN	ĐỀ XUẤT GIẢI PHÁP
<p>"First Class": trễ ~100%, tăng nguy cơ trễ 2.02 LẦN. "Second Class": trễ rất cao, tăng nguy cơ trễ 1.55 LẦN. Mọi mode đều có thực tế > dự kiến.</p>	<ol style="list-style-type: none">1. Giải quyết khẩn cấp "First Class" & "Second Class": Điều tra sâu, cải tổ triệt để hoặc tạm ngưng/loại bỏ.2. Tối ưu quy trình Kho & Điều chỉnh cam kết thực tế: Rà soát quy trình nội bộ. Cập nhật thời gian giao dự kiến cho khách
Khu vực: Đông Phi, Bắc Phi. Ngành hàng: Health & Beauty. (Cũng góp phần tăng tỷ lệ trễ, ảnh hưởng ~1.1x - 1.2x)	<ol style="list-style-type: none">3. Xử lý Điểm nóng Khu vực & Ngành hàng: Đánh giá logistics cho Đông/Bắc Phi. Kiểm tra quy trình ngành "Health & Beauty".



PHÂN TÍCH CÁC CỤM KHÁCH HÀNG



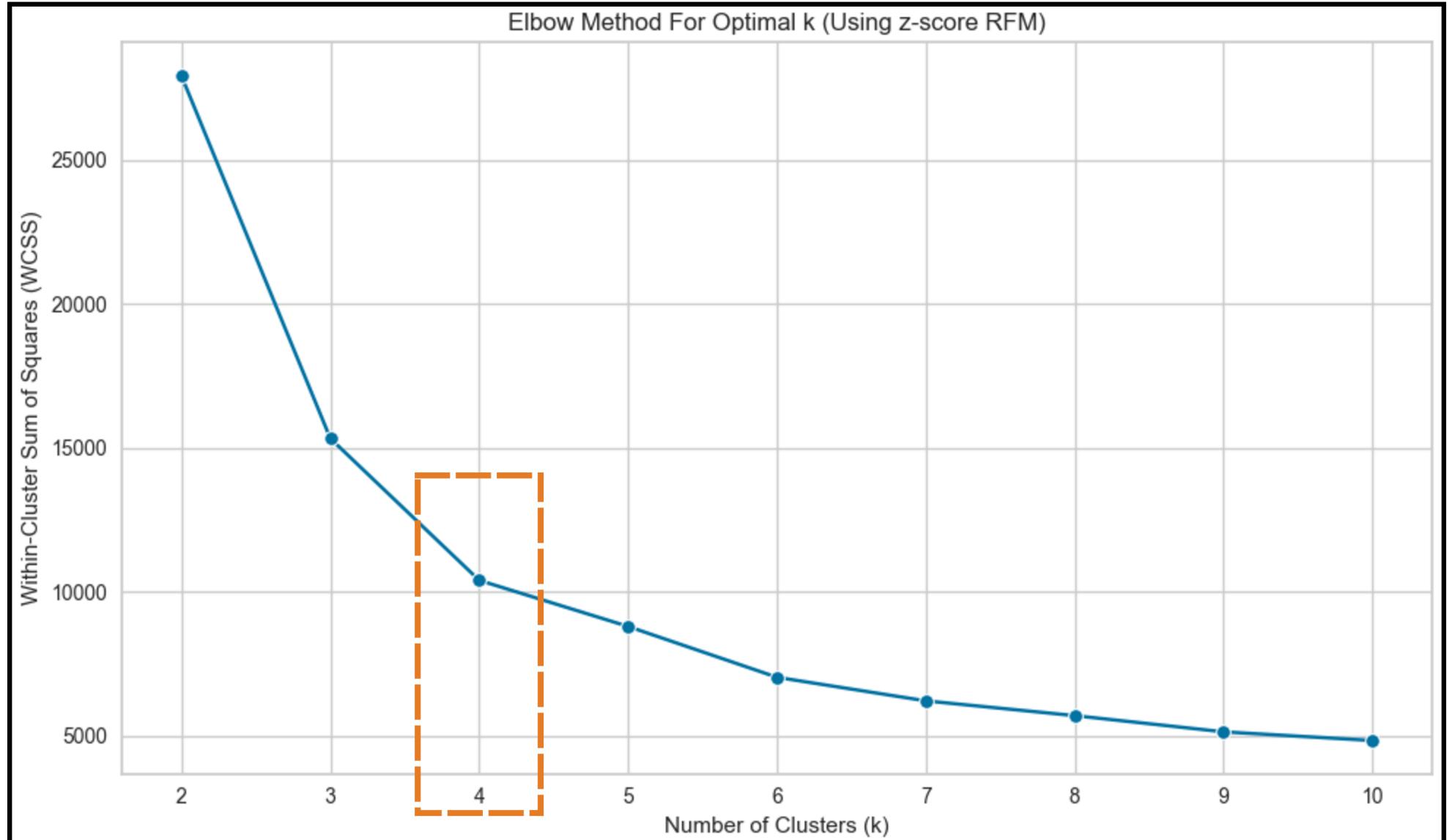
2/3 khách hàng không quay lại trong suốt 4 quý tiếp theo

tỉ lệ khách hàng mua lại tiếp tục giảm nhẹ qua các quý

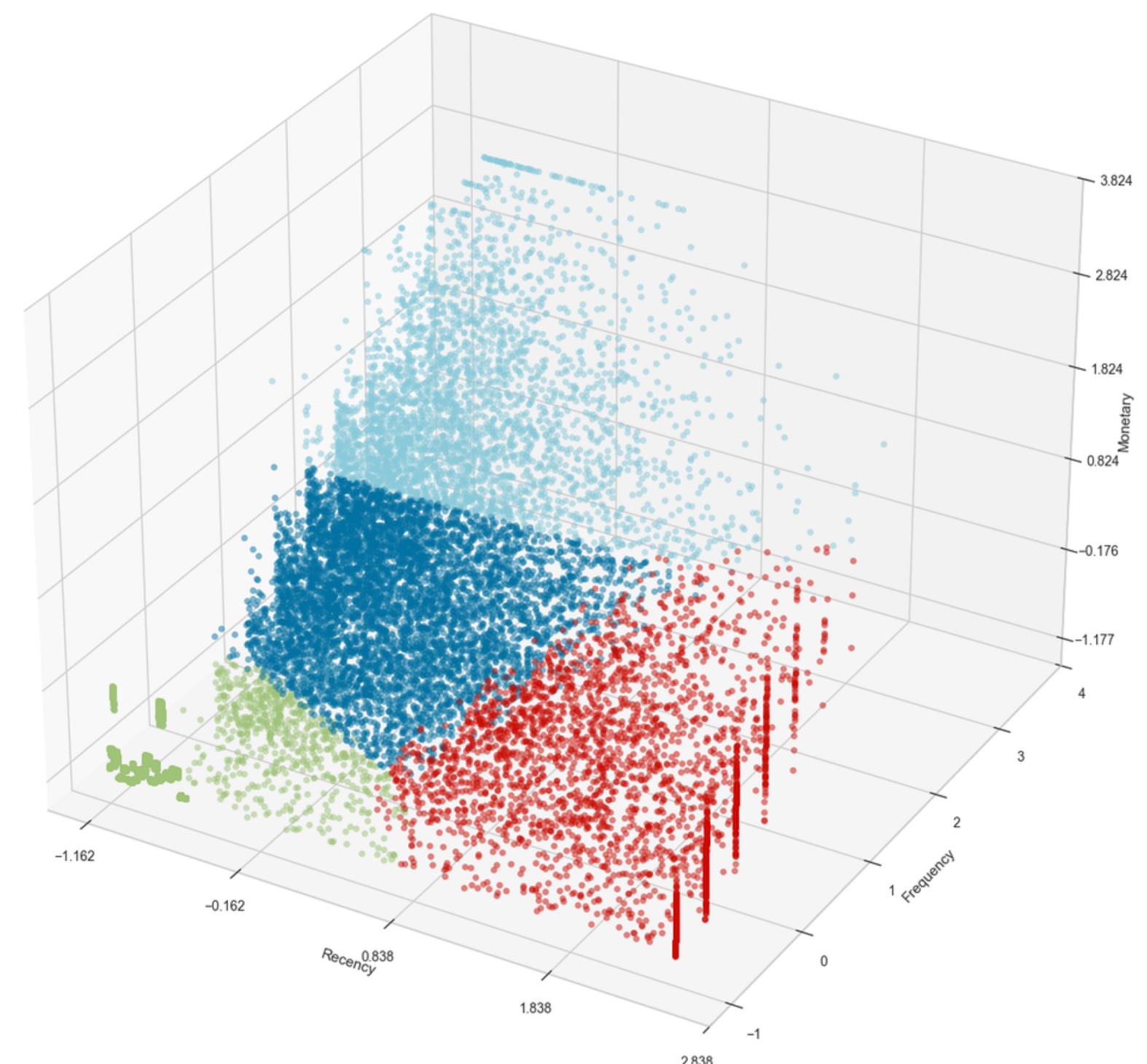
giảm mạnh sau 9 quý trở đi

- Sau quý đầu tiên công ty chỉ giữ được khoảng **1/3** khách hàng
- Tỉ lệ tương đối ổn định trong ngắn và trung hạn
- Đặc biệt, **khả năng giữ chân** khách hàng trên **dài hạn** (trên 2 năm) là **rất thấp**

Elbow Method For Optimal k (Using z-score RFM)



RFM 3D Visualization

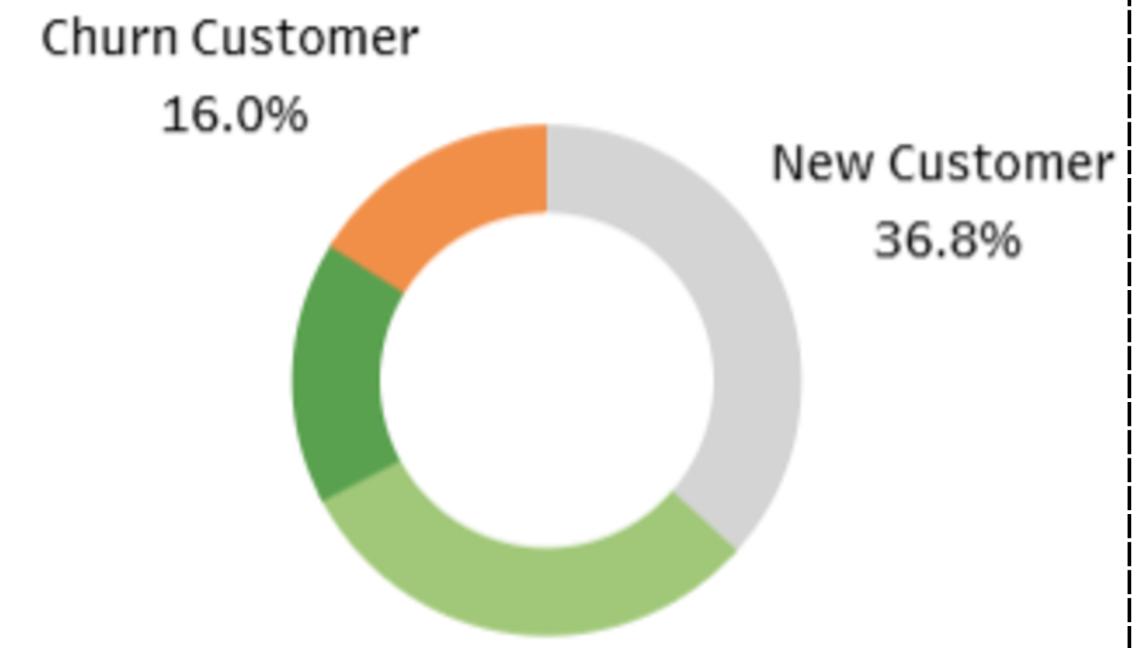


**Phân cụm khách hàng - RFM
bằng KMeans với K = 4**

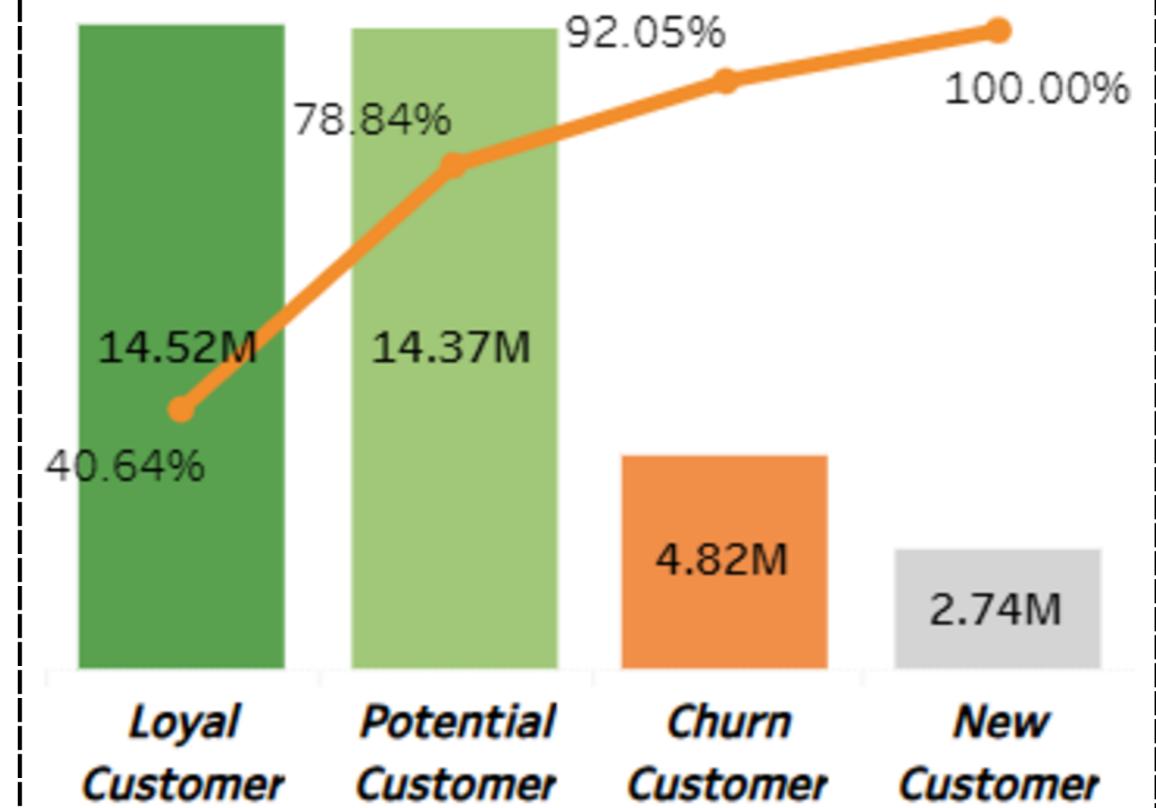
Phân cụm khách hàng - RFM

- **Khách hàng trung thành** và **tiềm năng** mang lại **phần lớn giá trị lợi nhuận (gần 80%)**, dù tổng số lượng **chỉ chiếm 47.2%** tổng khách hàng (Loyal 16.6%, Potential 30.6%)
 - Phản ánh nguyên tắc 80/20 trong kinh doanh: số ít khách hàng tạo ra phần lớn giá trị.
- **Khách hàng mới** chiếm **tỉ trọng lớn nhất theo số lượng (36.8%)** nhưng lại đóng góp chỉ ~8% giá trị.
- Tỷ lệ **khách hàng rời bỏ** chiếm 16% về số lượng và đóng góp 13.2% về giá trị.

Total customers by clusters



Customer Cluster Overview



Mô hình mở rộng RFM-V



- Variety là **số lượng loại sản phẩm khác nhau** mà khách hàng đã mua.
- Chỉ số này cho biết **mức độ đa dạng** trong hành vi mua hàng của từng nhóm khách hàng.



Variety là biến bổ sung cho RFM, giúp làm rõ **mức độ phong phú trong tiêu dùng**, bên cạnh việc đo lường theo tần suất, giá trị và thời gian mua hàng.

Đặc điểm, hành vi của các cụm khách hàng

Tần suất mua hàng và tổng chi tiêu cao nhất, mức độ đa dạng sản phẩm cao, thể hiện gắn bó và sự tin tưởng cao.

Clusters Feature Analysis							
Customer Classify	Recency	Frequency	Monetary	Diversity	Avg Profits Ratio	Avg Discount Ratio	AOV
Loyal Customer	208.8	7.355	4,217	10	12.33%	10.10%	638.5 USD
Potential Customer	217.8	4.415	2,278	8	11.76%	10.22%	574.5 USD
Churn Customer	527.1	2.851	1,461	6	12.13%	10.16%	570.2 USD
New Customer	57.4	1.087	361	1	12.45%	10.23%	369.3 USD

- New Customer có giá trị đơn **thấp nhất**
- Churn và Potential lại có AOV **xấp xỉ nhau** cho thấy **Churn từng là nhóm "có giá trị" trước khi rời đi.**

Tuy mua ít đa dạng và giá trị đơn hàng thấp, nhưng mang lại **tỷ suất lợi nhuận cao nhất**.

Nhận xét

Đặc điểm, hành vi của các cụm khách hàng



Chiến lược chiết khấu phân khúc

Doanh nghiệp đang áp dụng chiết khấu đồng đều, **thiếu cá nhân hóa**
Chưa tối ưu hiệu quả.

Avg Profits Ratio	Avg Discount Ratio
12.33%	10.10%
11.76%	10.22%
12.13%	10.16%
12.45%	10.23%



Không cần thiết với khách hàng Trung thành

- Nhóm Loyal vẫn mang lại **lợi nhuận cao** dù nhận ít chiết khấu
Coi trọng trải nghiệm hơn giá cả
- Cân nhắc ưu đãi **phi tài chính** thay thế



Lãng phí với khách hàng Tiềm năng

Nhóm Potential nhận chiết khấu cao nhất nhưng **lợi nhuận thấp**
Chiến lược hiện tại đang giảm biên lợi nhuận mà không tăng giá trị khách hàng



Không hiệu quả với khách hàng Rời bỏ

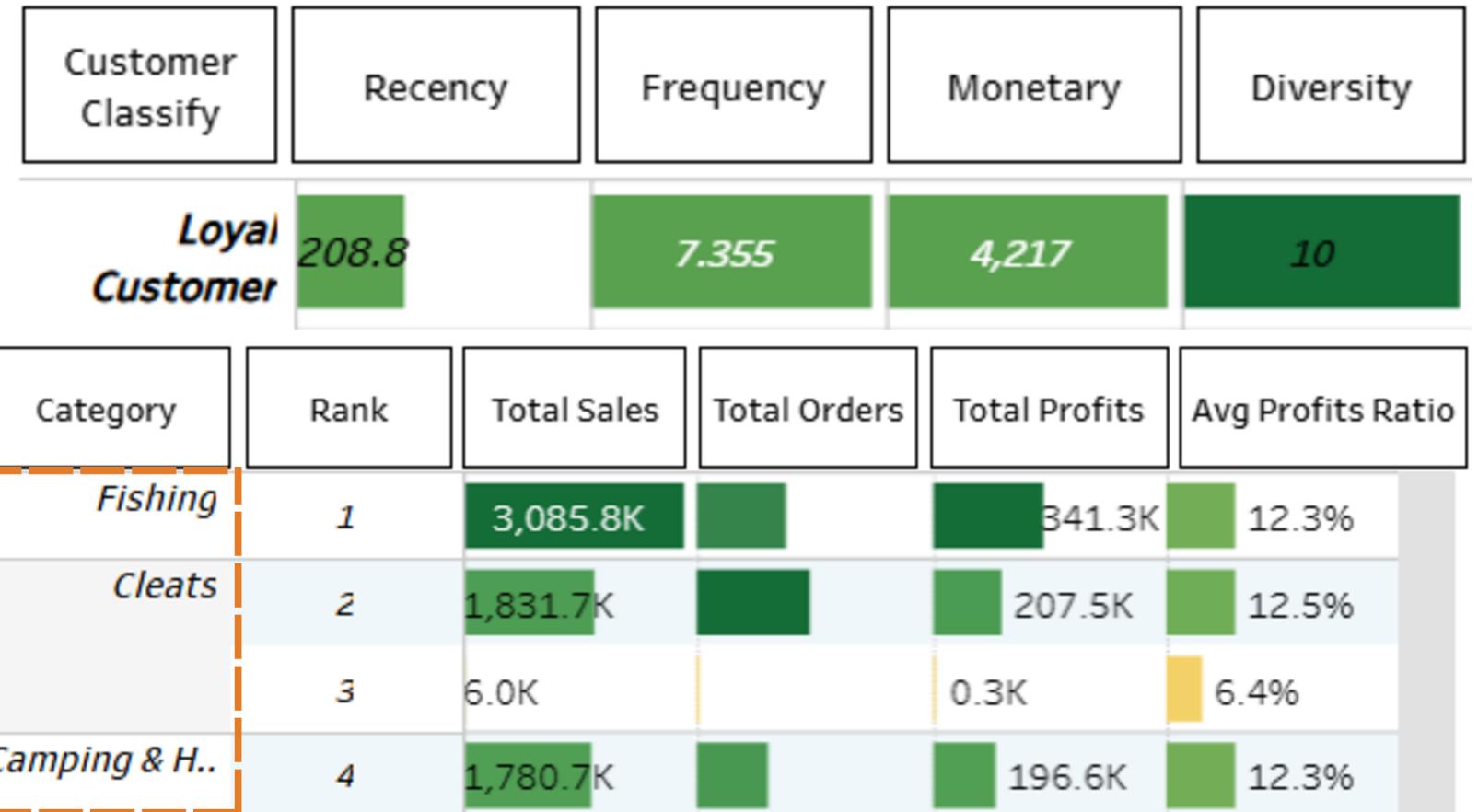
Nhóm Churn vẫn **rời đi** dù được chiết khấu
Cần cải thiện các yếu tố khác như sản phẩm, dịch vụ và trải nghiệm.

Nhận xét

Đặc điểm, hành vi của các cụm khách hàng

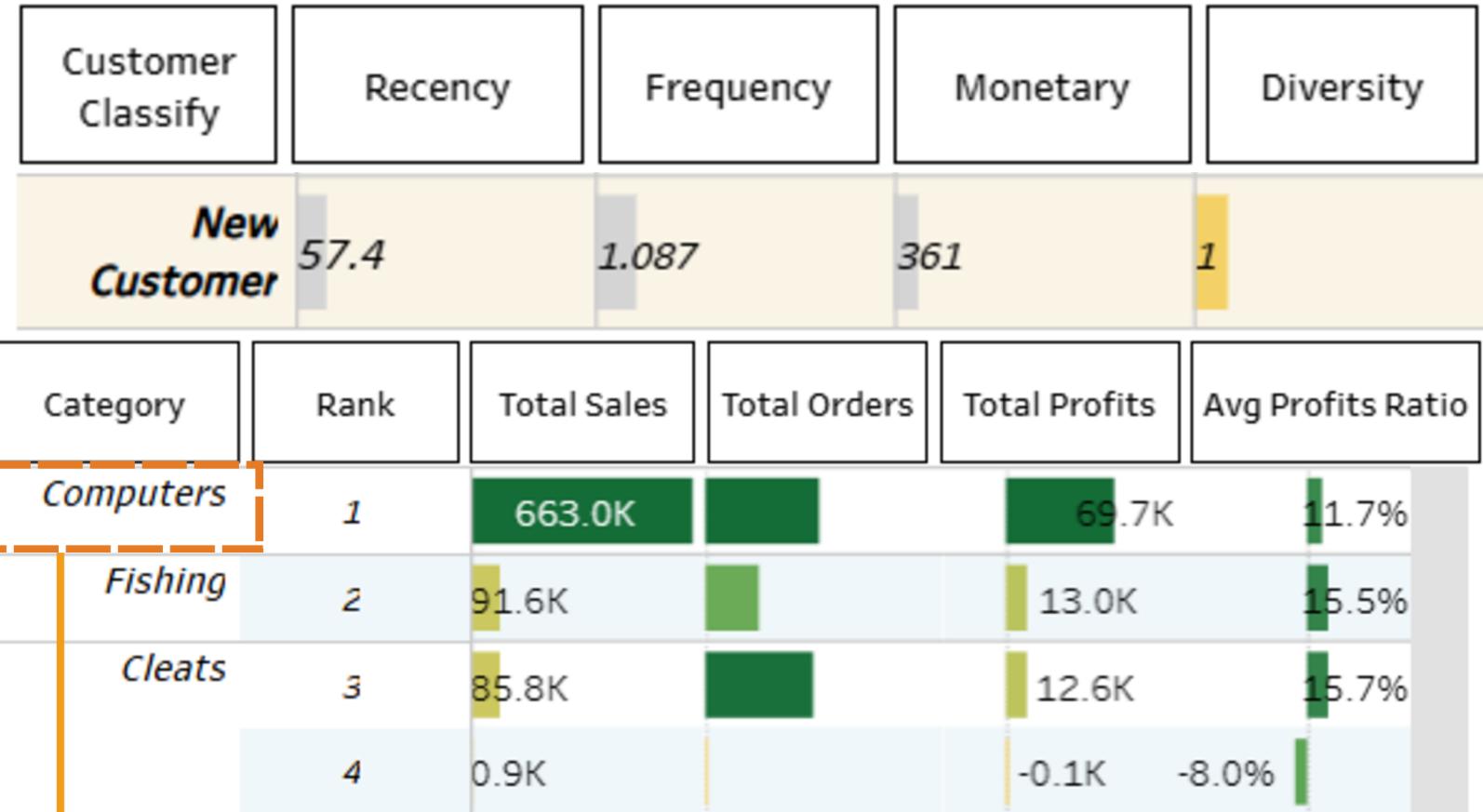


Loyal Customer có độ đa dạng trong giỏ hàng **cao nhất**.



Đứng đầu là **Fishing** và lần lượt là **Cleats** và **Camping & Hiking** với doanh thu đóng góp vào doanh thu tổng cao và tỷ suất lợi nhuận trung bình cao.

New Customer có độ đa dạng trong giỏ hàng **thấp nhất**.



Nhu cầu mới so với các phân khúc khác

➡ Đề xuất áp dụng **chiến lược Up-sell** đối với 3 dòng này

**Thank you
FOR LISTENING**

