



*Final Project*

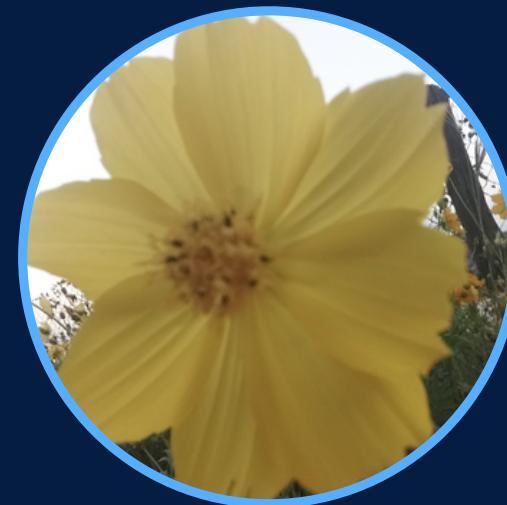
# DATA CO GLOBAL ANALYSIS

*Presented by: Team B*





# THÀNH VIÊN NHÓM



Đoàn Ngọc Tuấn

BI53



Nguyễn Minh Khang

BI53



# NỘI DUNG

TỔNG QUAN VỀ DATASET

PHÂN TÍCH TỔNG QUAN

PHÂN TÍCH CHI TIẾT

KẾT LUẬN VÀ ĐỀ XUẤT



## I. TỔNG QUAN VỀ DATASET



## 1. Dataset

- Đây là bộ dữ liệu về chuỗi cung ứng của công ty DataCo Global, liên quan tới các lĩnh vực cung ứng, sản xuất, bán hàng, phân phối thương mại.
- Bộ dữ liệu gồm 22 cột và 180.511 dòng

- order\_id
- type
- order\_status
- order\_date
- sales
- market
- discount\_rate
- product\_quantity

- category\_id
- category\_name
- day\_shipping\_real
- day\_shipping\_schedule
- delivery\_status
- shipping\_mode
- customer\_id
- customer\_name

- customer\_country
- customer\_segment
- product\_id
- product\_name
- product\_price
- shipping\_id



## 2. Tiền xử lý

# CÔNG CỤ SỬ DỤNG



Lọc, tìm kiếm thông tin  
và tiền xử lý dữ liệu

Trực quan hóa dữ liệu, hỗ trợ  
phân tích và tìm insights



# TỔNG QUAN VỀ DATASET

## 2. Tiền xử lý

- Tạo thêm cột customer\_name

```
df['Customer Name'] = df['Customer Fname'] + ' ' + df['Customer Lname']
```

- Xóa null và giá trị trùng lặp

```
df_res.dropna(inplace=True)
df_res = df_res.drop_duplicates()
```

- Xử lý tên cột

```
df_res = df_res.rename(columns={'Order Item Discount Rate':'discount_rate', 'Order Item Quantity': 'product_quantity',
                                'order date (DateOrders)': 'order_date', 'Days for shipping (real)': 'day_shipping_real',
                                'Days for shipment (scheduled)': 'day_shipment_scheduled', 'Product Card Id': 'product_id'})
df_res.columns = df_res.columns.str.lower().str.replace(' ', '_')
```



## 2. Tiền xử lý

- Kiểm tra cột

```
df_res.columns  
  
Index(['order_id', 'type', 'order_status', 'order_date', 'sales', 'market',  
       'discount_rate', 'latitude', 'longitude', 'product_quantity',  
       'category_id', 'category_name', 'day_shipping_real',  
       'day_shipment_scheduled', 'delivery_status', 'shipping_mode',  
       'customer_id', 'customer_country', 'customer_name', 'customer_segment',  
       'product_id', 'product_name', 'product_price'],  
      dtype='object')
```

- Tạo thêm cột shipping\_id

```
df_res['shipping_id'] = range(1, len(df_res) + 1)
```



## 2. Tiền xử lý

- Tạo bảng Fact và Dimensions

```
df_product = df[['product_id', 'product_name', 'product_price', 'category_id']]
df_product = df_product.drop_duplicates(subset=['product_id'])
df_product.to_excel('dim_product.xlsx', index = False)

df_category = df[['category_id', 'category_name']]
df_category = df_category.drop_duplicates(subset=['category_id'])
df_category.to_excel('dim_category.xlsx', index = False)

df_customer = df[['customer_id', 'customer_country', 'customer_name', 'customer_segment']]
df_customer = df_customer.drop_duplicates(subset=['customer_id'])
df_customer.to_excel('dim_customer.xlsx', index = False)

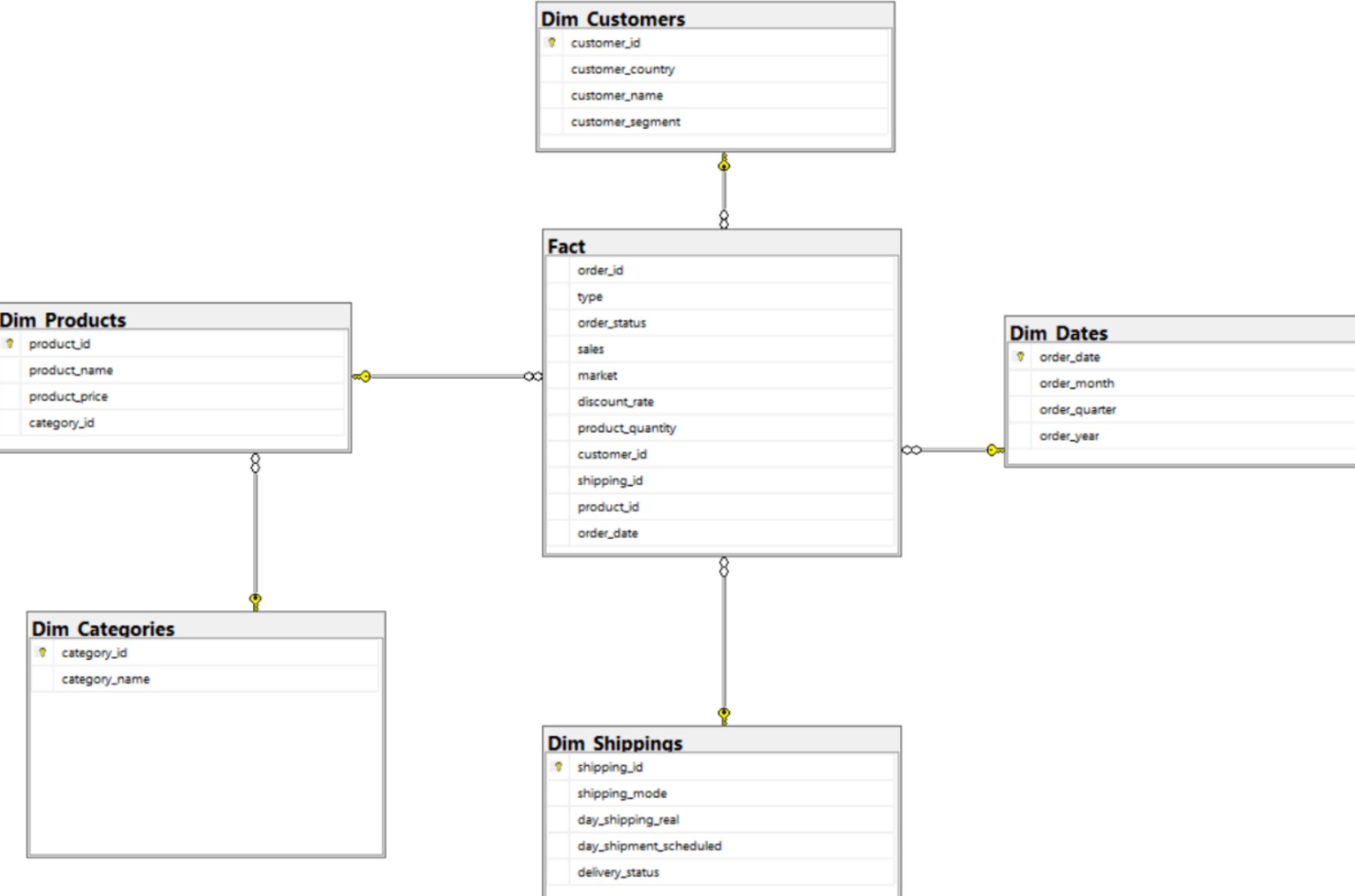
df_shipping = df[['shipping_id', 'shipping_mode', 'day_shipping_real', 'day_shipment_scheduled', 'delivery_status']]
df_shipping.to_excel('dim_shipping.xlsx', index = False)

df.order_date = pd.to_datetime(df.order_date)
df_date = df[['order_date']]
df_date['order_month'] = df.order_date.dt.month
df_date['order_quarter'] = df.order_date.dt.quarter
df_date['order_year'] = df.order_date.dt.year
df_date.to_excel('dim_date.xlsx', index = False)

df_fact = df[['order_id', 'product_id', 'customer_id', 'shipping_id', 'order_date', 'type',
              'order_status', 'sales', 'market', 'discount_rate', 'product_quantity']]
df_fact.to_excel('fact.xlsx', index = False)
```



## 3. Data Model





# TỔNG QUAN VỀ DATASET

## 4. Mục tiêu phân tích

Cải thiện sản phẩm để  
tối ưu hóa doanh thu

Kiểm soát phân khúc  
khách hàng

Mục tiêu 1

Mục tiêu 2

Mục tiêu 3

Mục tiêu 4

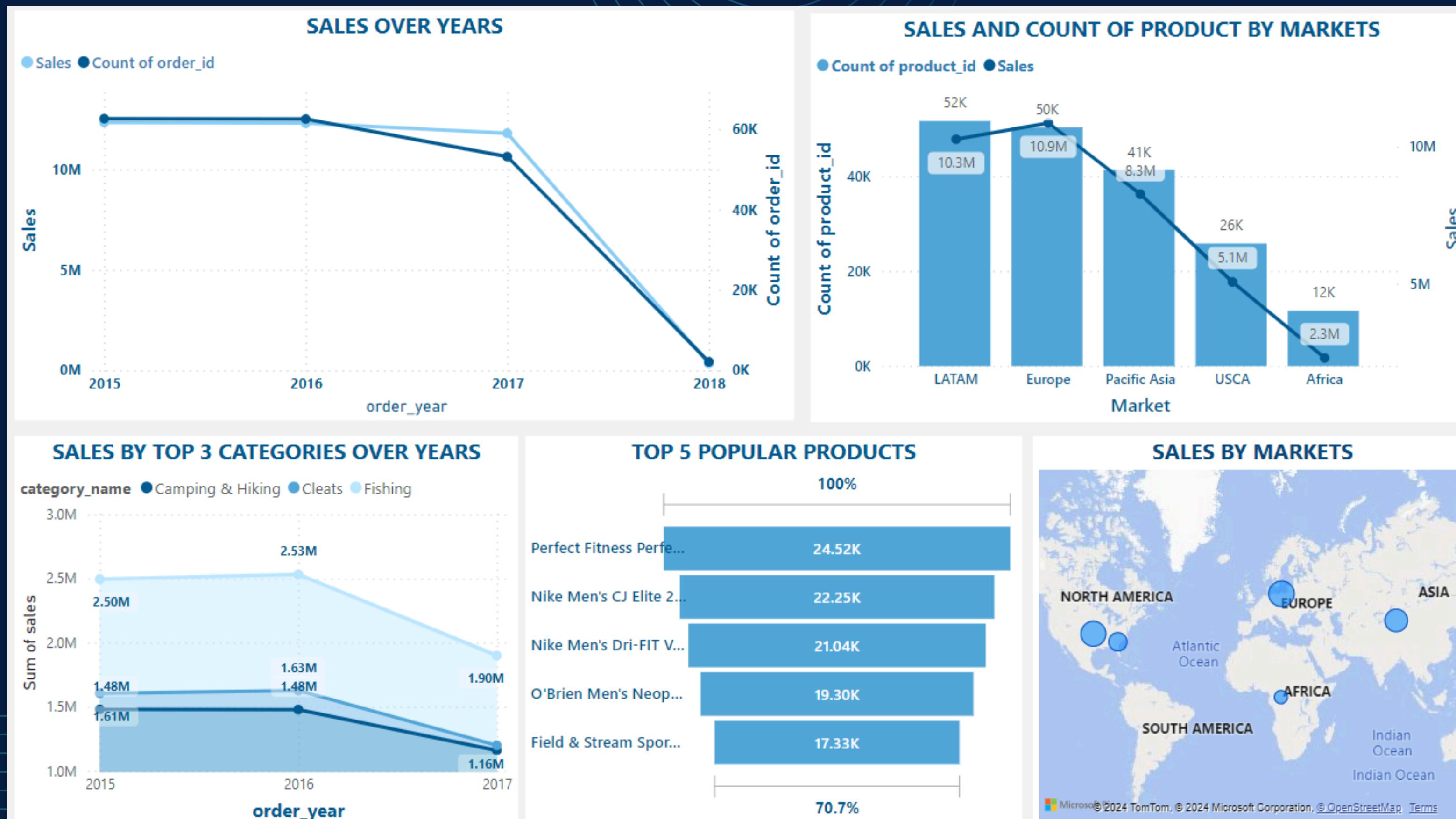
Tối ưu hóa thanh toán

Cải thiện chất lượng  
giao hàng



## II. PHÂN TÍCH TỔNG QUAN

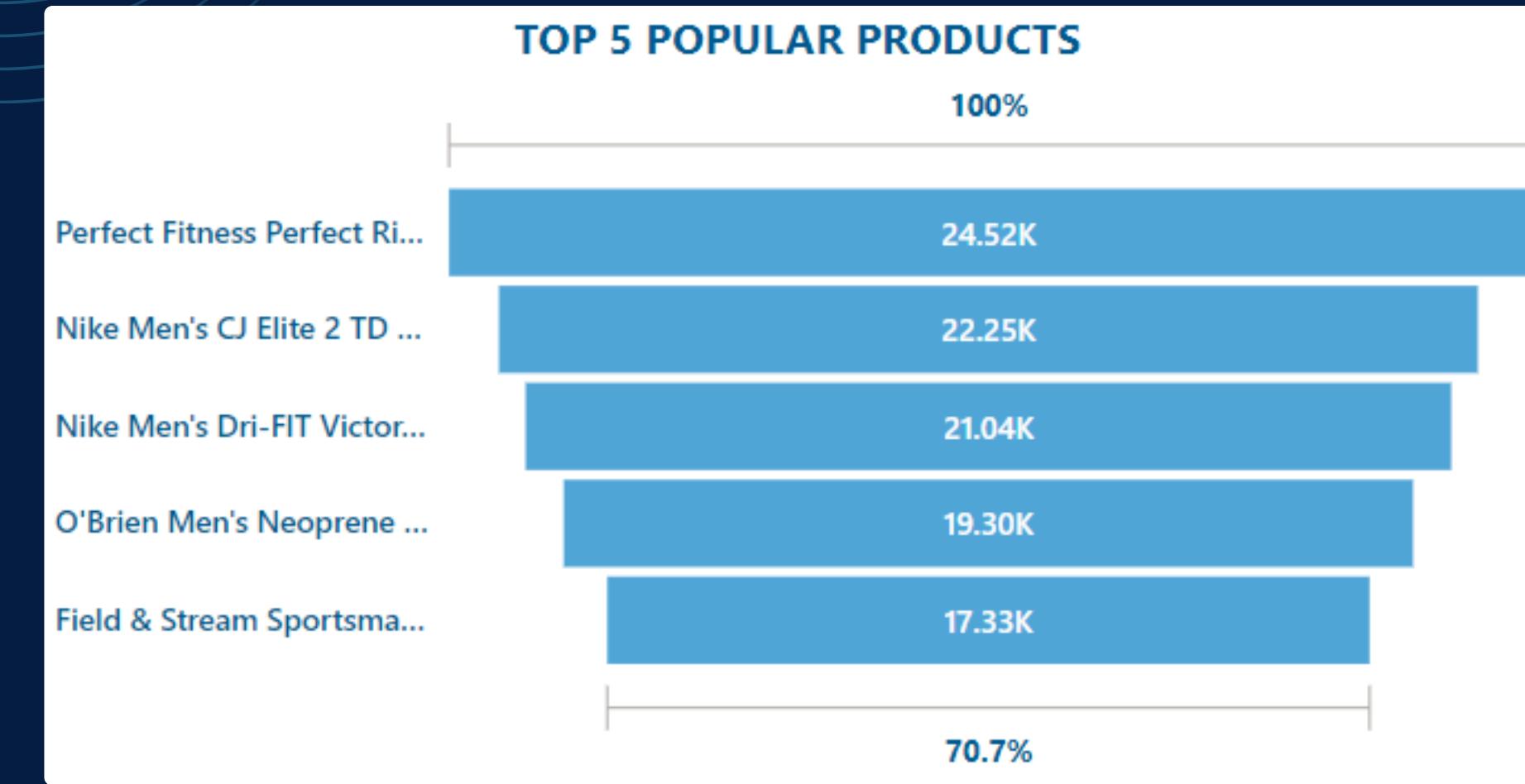
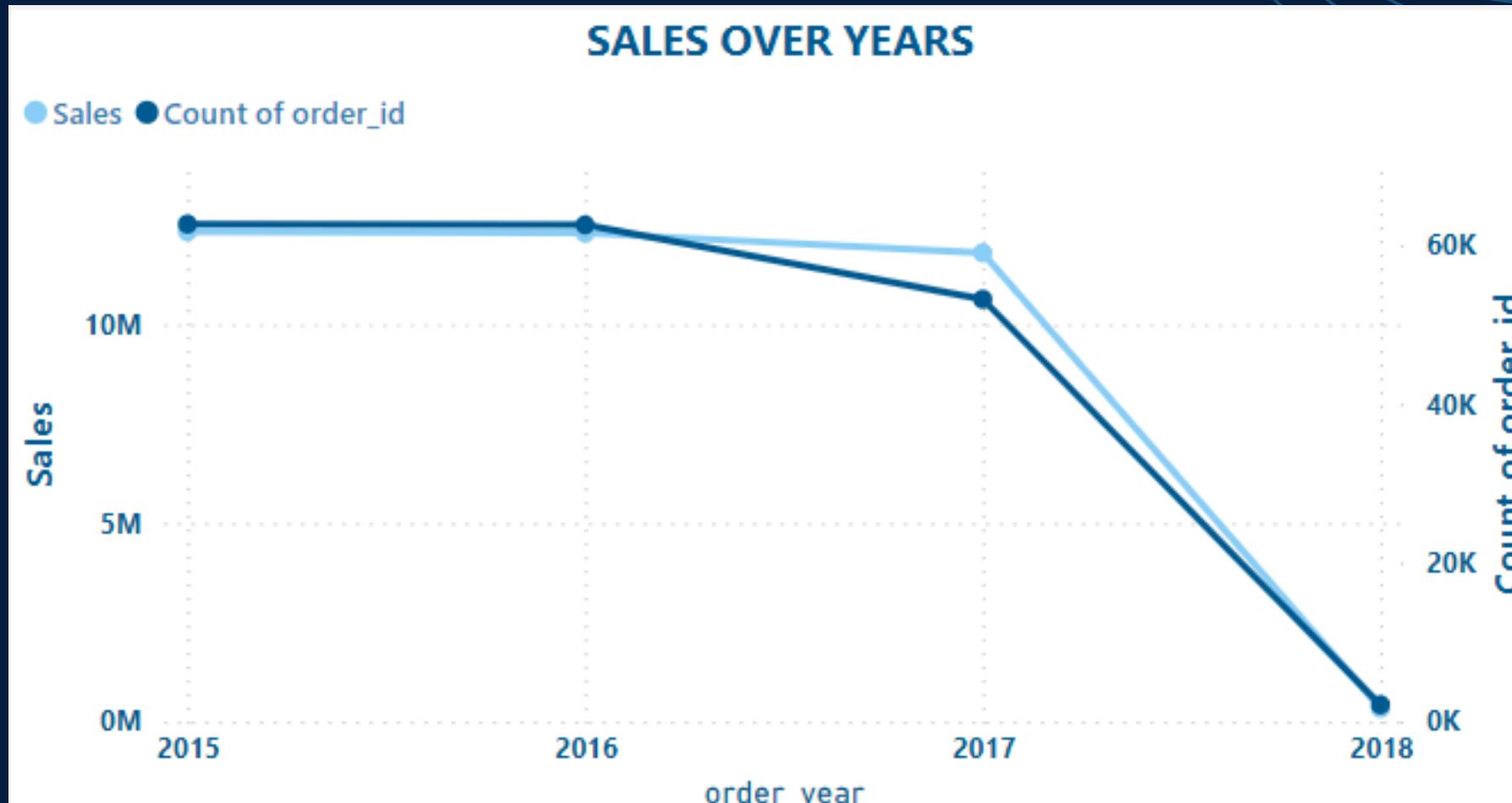
# PHÂN TÍCH TỔNG QUAN



# PHÂN TÍCH TỔNG QUAN



Nhóm B



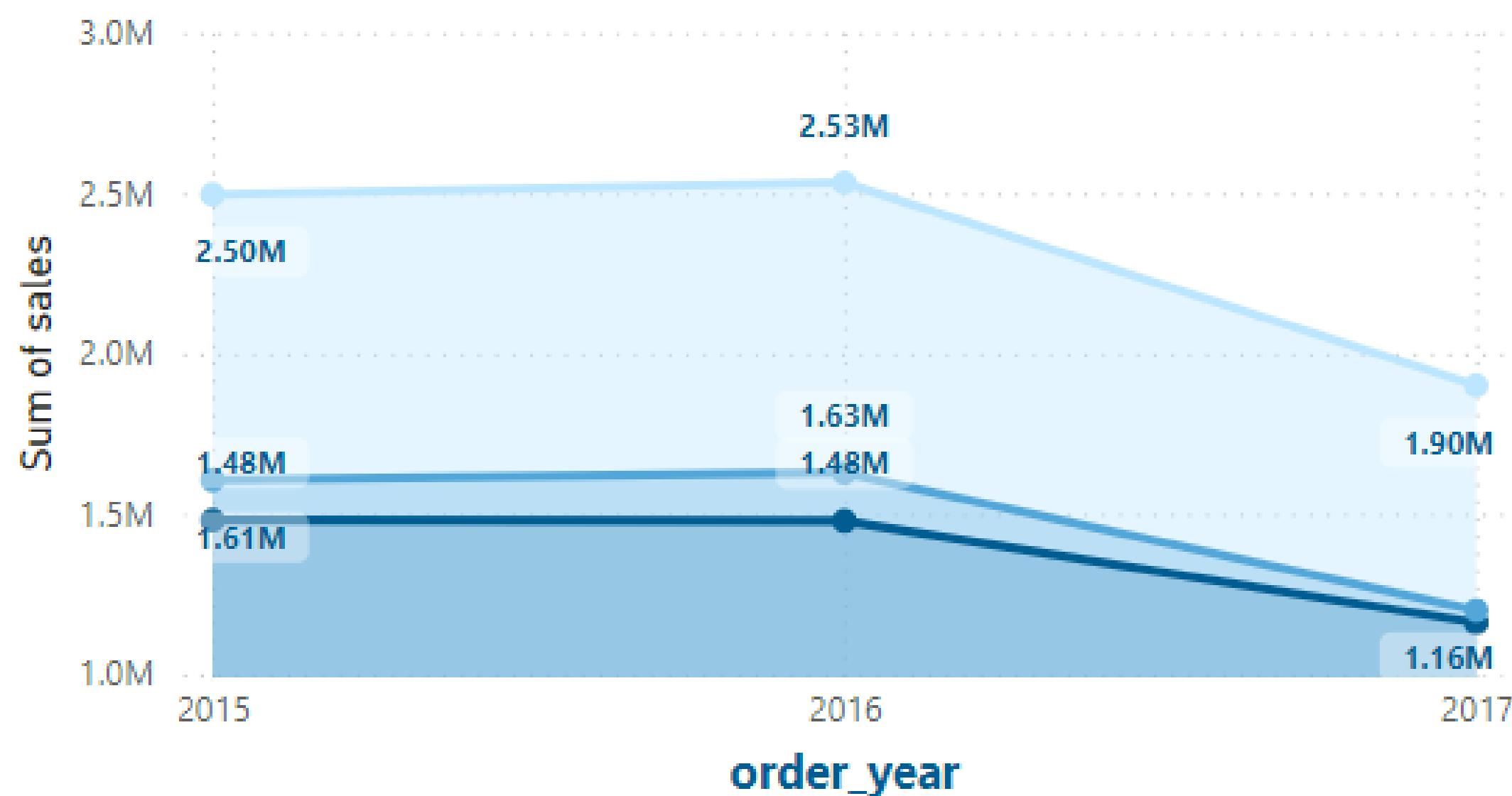
- Doanh số giảm tương đối từ 2016 đến 2017
- Tỉ lệ giữa số lượng đơn hàng và doanh số bán hàng cũng tương đối cân bằng.

- Sản phẩm **Perfect Fitness Rip Deck** và các dòng sản phẩm của **Nike Men's** có số lượng bán được rất cao.



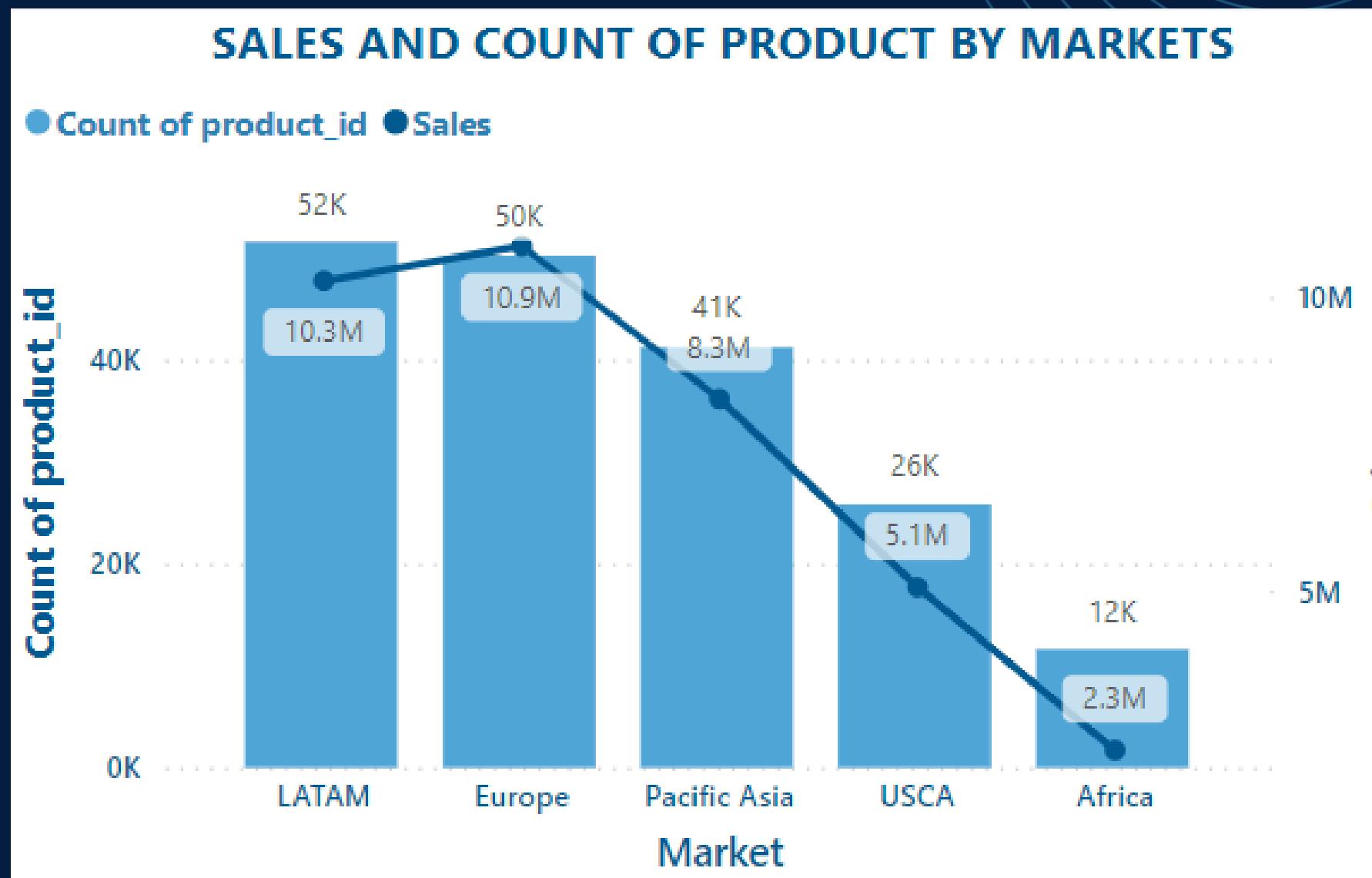
## SALES BY TOP 3 CATEGORIES OVER YEARS

category\_name ● Camping & Hiking ● Cleats ● Fishing



**Camping & Hiking, Cleats, Fishing** là top 3 categories có doanh số đóng góp cao nhất. Trong đó **Camping & Hiking** vượt trội hơn so với 2 categories còn lại.

Doanh số tăng đều từ **2015** đến **2016** nhưng có sự giảm mạnh tới năm **2017**.



- Thị trường **LATAM** và **Europe** có số lượng đơn hàng được mua cũng như doanh thu nằm trong Top đầu.
- Thị trường **Africa** và **USCA** chưa được khai thác triệt để nên doanh số bán hàng cũng như số lượng đơn hàng được mua khá thấp.



### III. PHÂN TÍCH CHI TIẾT THỊ TRƯỜNG LATAM

1. PHÂN TÍCH THEO  
SẢN PHẨM - DANH MỤC

3. PHÂN TÍCH RFM

2. PHÂN TÍCH THANH TOÁN

4. PHÂN TÍCH GIAO HÀNG



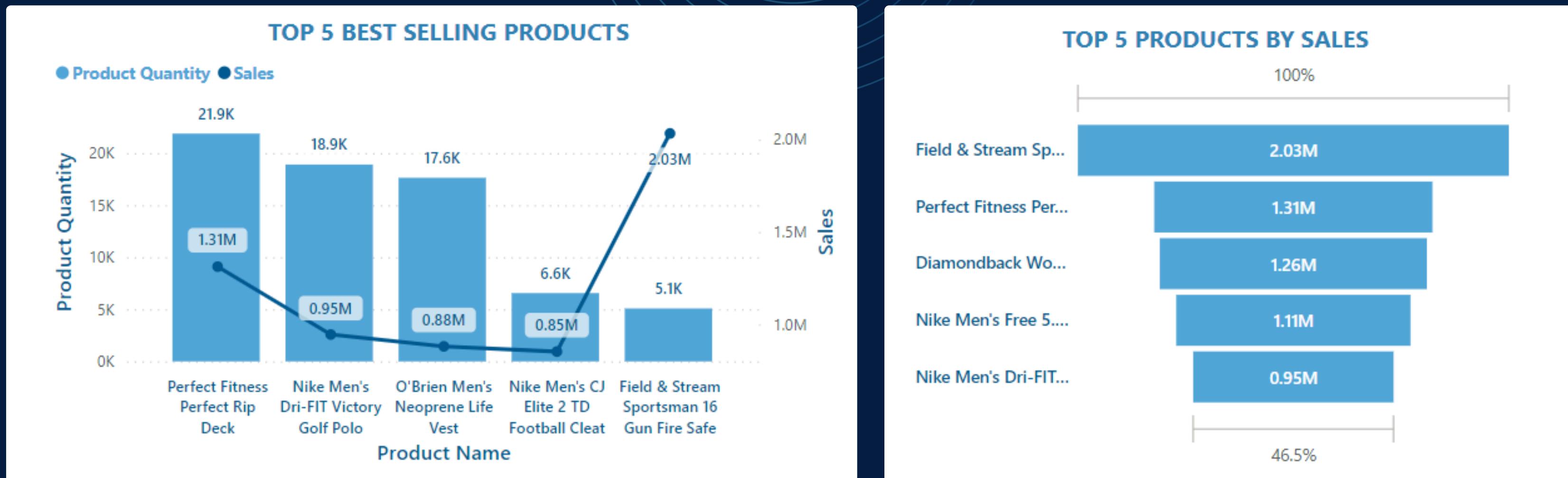
Nhóm B

## PHÂN TÍCH THEO SẢN PHẨM-DANH MỤC

# PHÂN TÍCH THEO SẢN PHẨM - DANH MỤC



Nhóm B



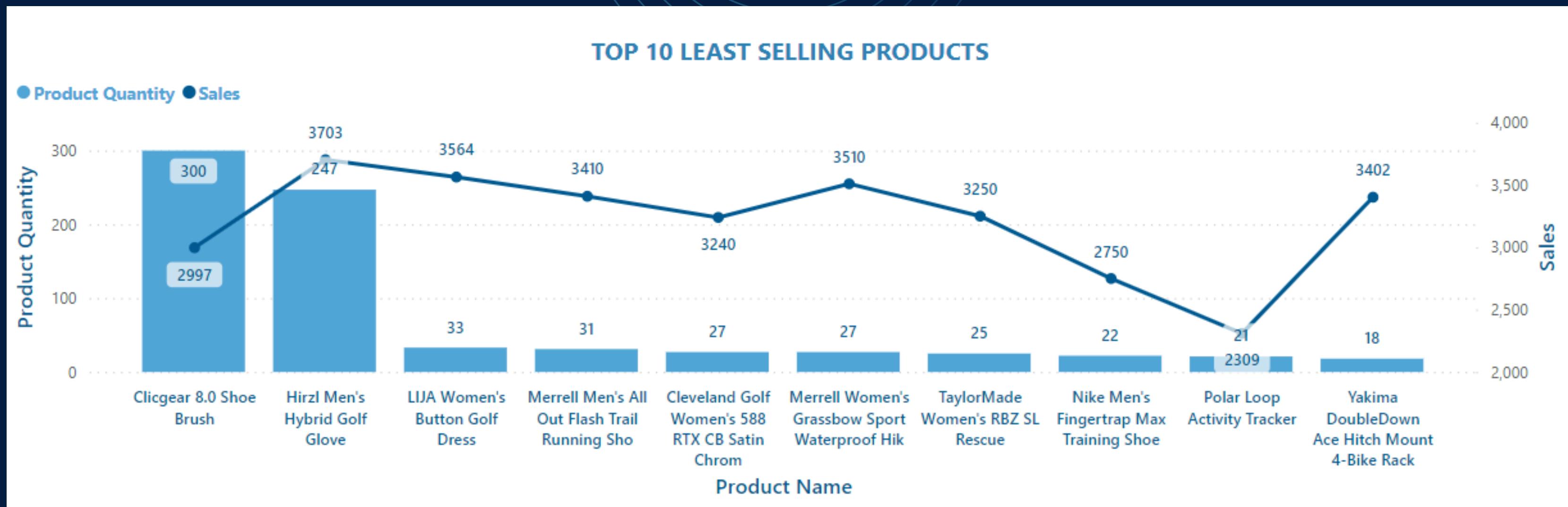
Top 5 sản phẩm có số lượng bán được cao nhất ở LATAM là **Perfect Fitness Rip Deck**, các dòng sản phẩm của **Nike Men's**, **O'Brien Men's** và **Field & Stream Sports**.

- Các dòng sản phẩm **Nike Men's**, **O'Brien** có số lượng bán nằm trong Top đầu nhưng doanh số đóng góp rất thấp. → Xem xét điều chỉnh giá để cải thiện về doanh thu, nhưng cũng nên có những ưu đãi cho khách hàng cũng như điều chỉnh giá hợp lý để không làm mất khách hàng hoặc số lượng bán sản phẩm.
- Sản phẩm **Field & Stream Sports** có số lượng bán đứng thứ 5 nhưng doanh thu lại cao hơn hẳn so với các sản phẩm còn lại. → Đẩy mạnh tiếp thị bán hàng, tối ưu hóa lợi nhuận

# PHÂN TÍCH THEO SẢN PHẨM - DANH MỤC



Nhóm B



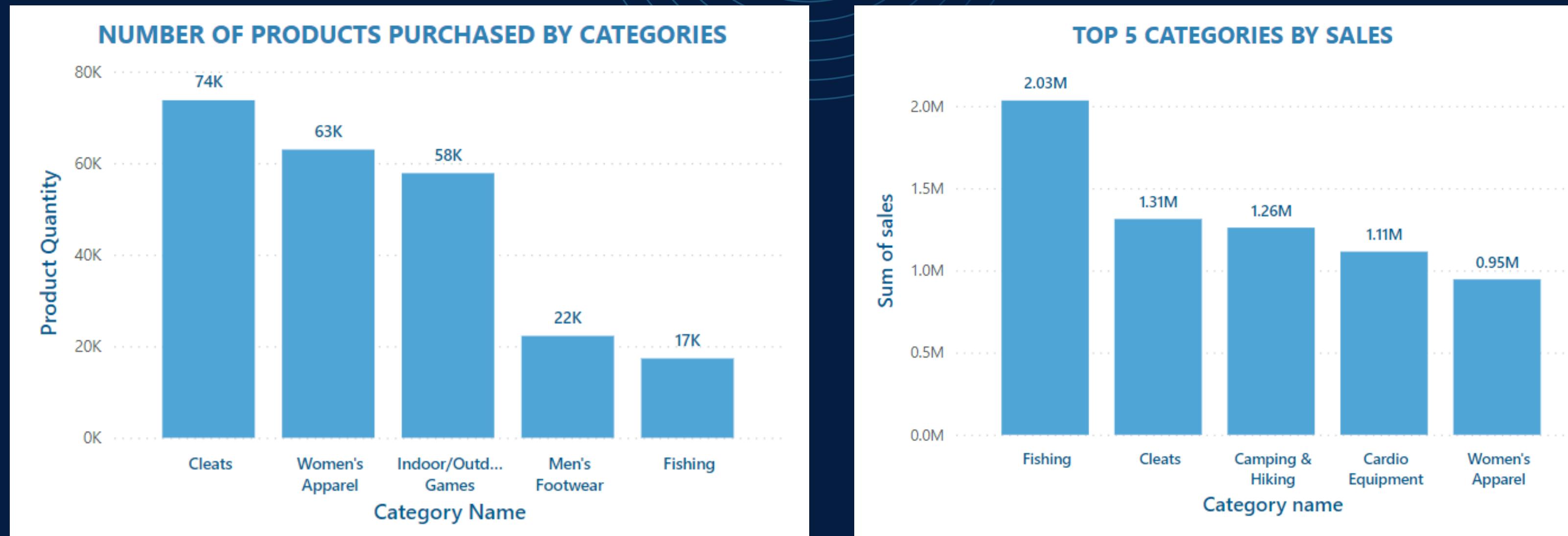
Top 10 sản phẩm có lượt mua thấp nhất

→ Phân tích và tìm hiểu nguyên nhân tại sao các mặt hàng này ko được ưa chuộng ở thị trường **LATAM**. Đồng thời tạo ra các chương trình khuyến mãi đặc biệt để kích thích nhu cầu mua hàng của khách hàng hoặc tạo combo sản phẩm với các sản phẩm bán chạy nhất để tối ưu hóa lợi nhuận.

# PHÂN TÍCH THEO SẢN PHẨM - DANH MỤC



Nhóm B

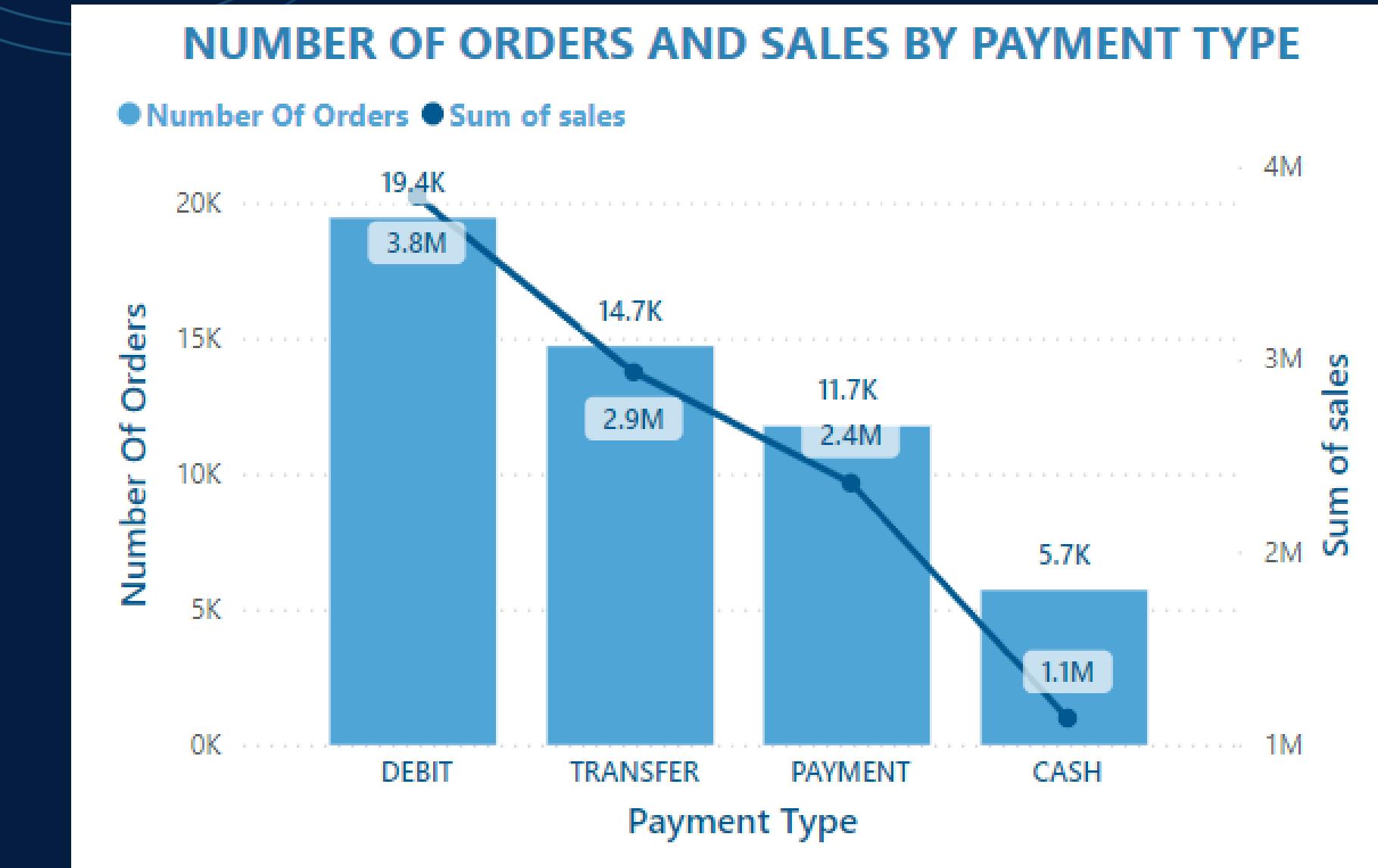
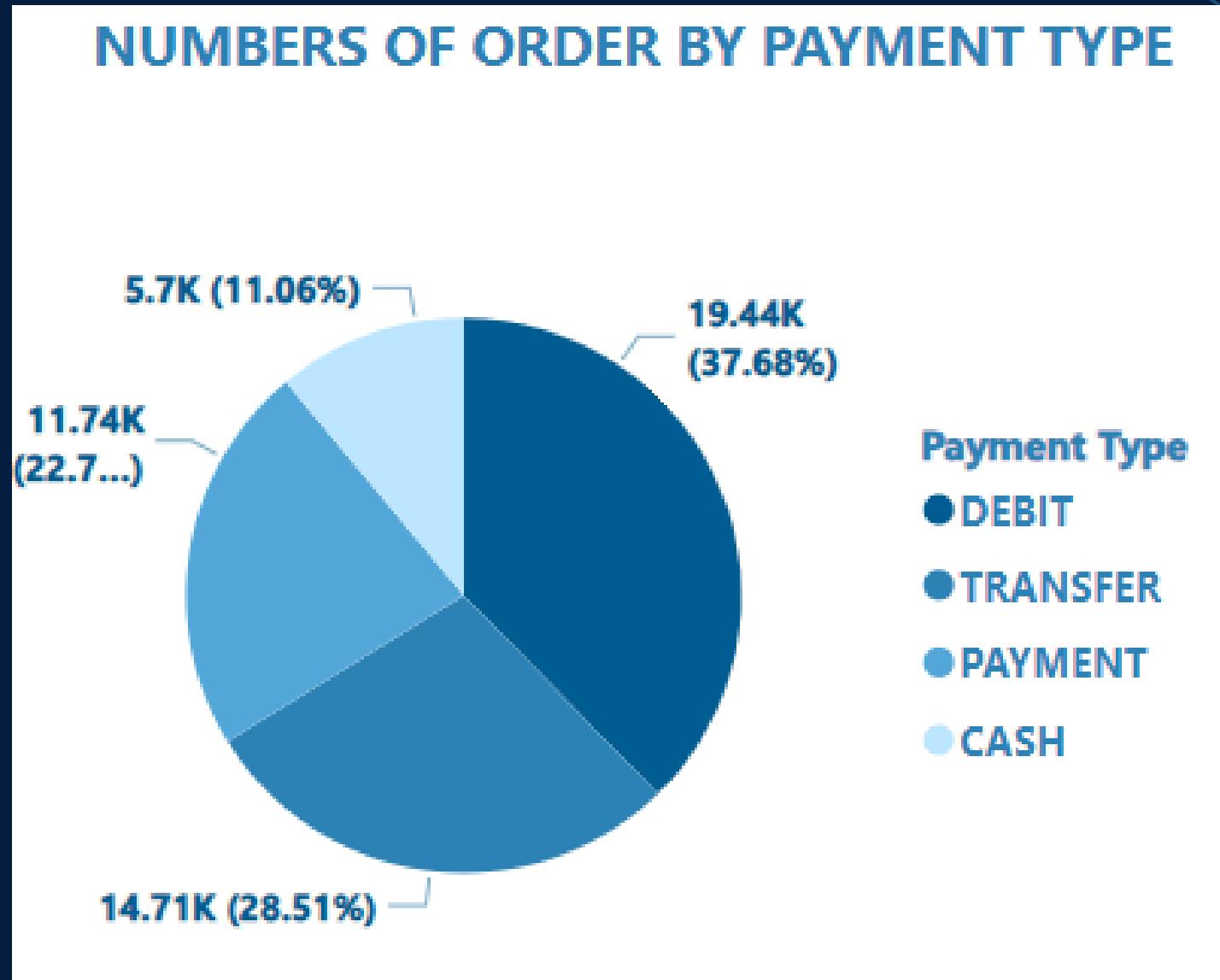


- Top 5 danh mục có số lượng sản phẩm bán được cao nhất: **Cleats, Women's Apparel, Indoor/Outdorr, Men's Footwear và Fishing.**
  - Fishing là danh mục có đóng góp doanh thu cao nhất, vượt trội so với các danh mục còn lại.
- Tăng cường nhập các dòng sản phẩm thuộc **Cleats, Fishing** để tối ưu hóa doanh số bán hàng.

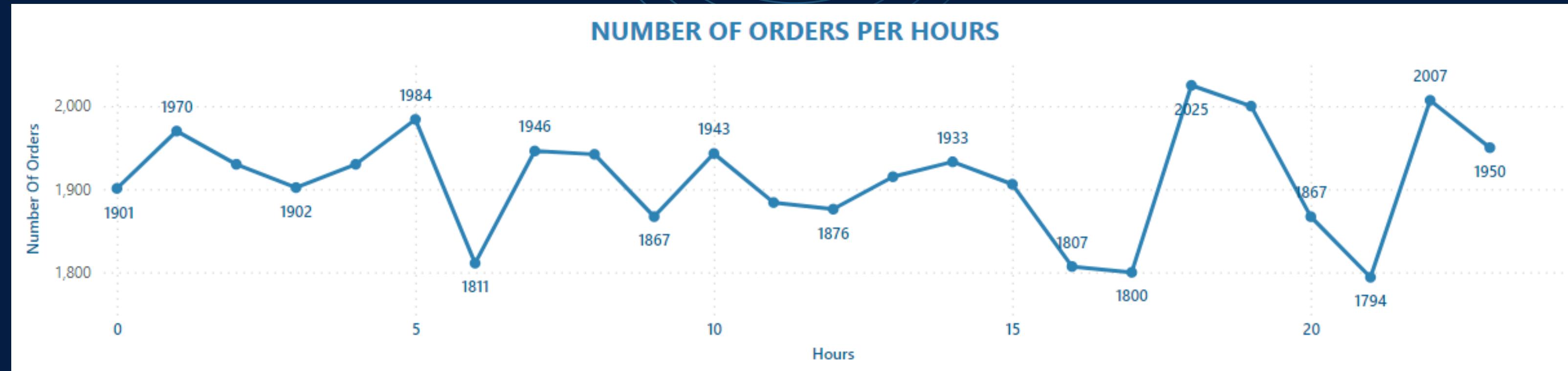


Nhóm B

## PHÂN TÍCH THANH TOÁN



- Khách hàng có xu hướng thích thanh toán bằng **DEBIT** và một số hình thức thanh toán online khác. Thanh toán bằng tiền mặt trở nên ít phổ biến ở LATAM.

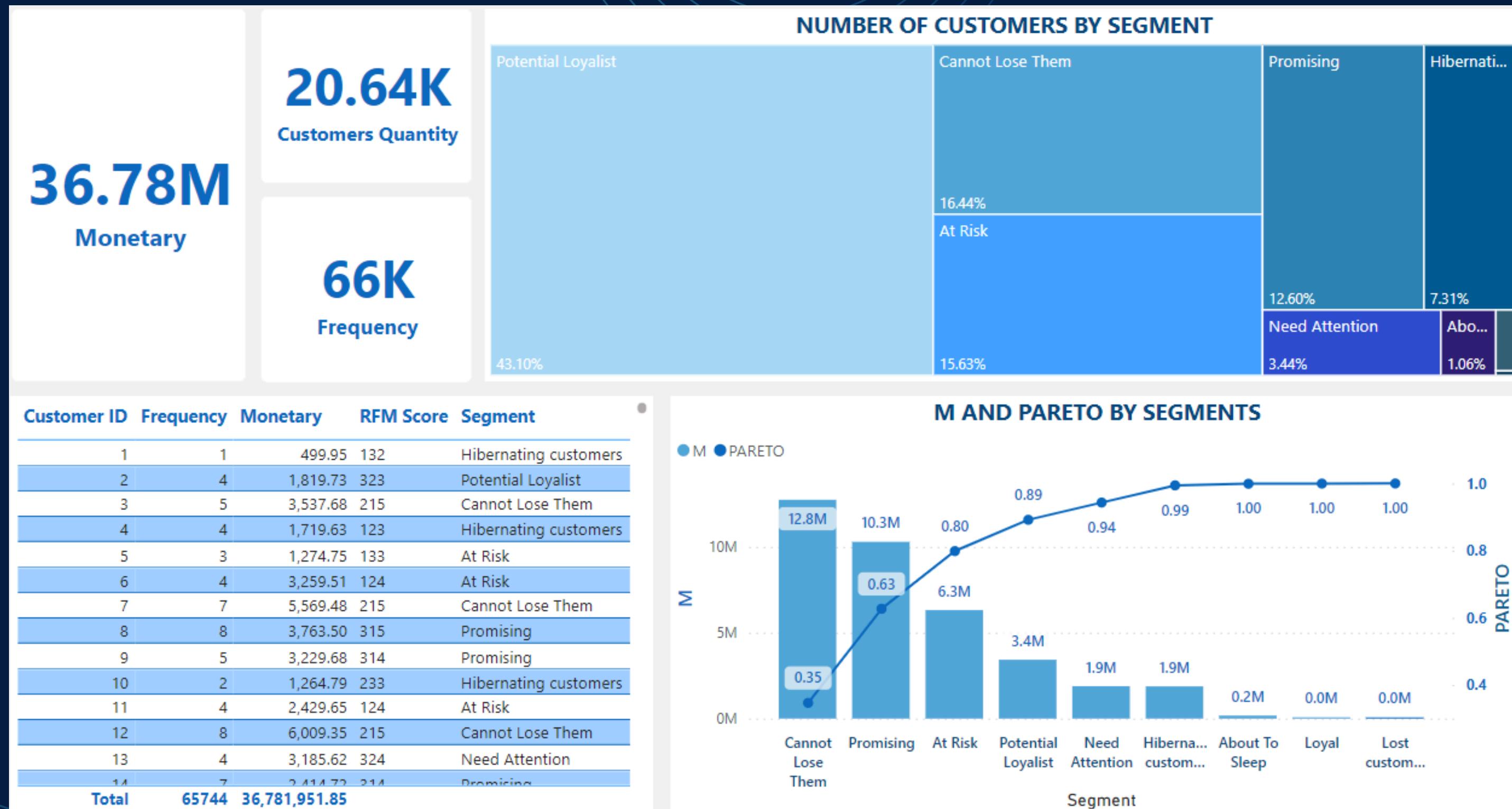


- Số lượng đơn đặt hàng theo giờ của các đơn hàng thanh toán online khá tương đồng nhau.
- **Tạo ra các chương trình ưu đãi và khuyến mãi** giảm giá dành riêng cho các phương thức thanh toán cụ thể để kích thích nhu cầu mua hàng.
- **Xây dựng và tích hợp thêm các hệ thống thanh toán trực tuyến** để thuận tiện cho nhu cầu thanh toán của khách hàng.



Nhóm B

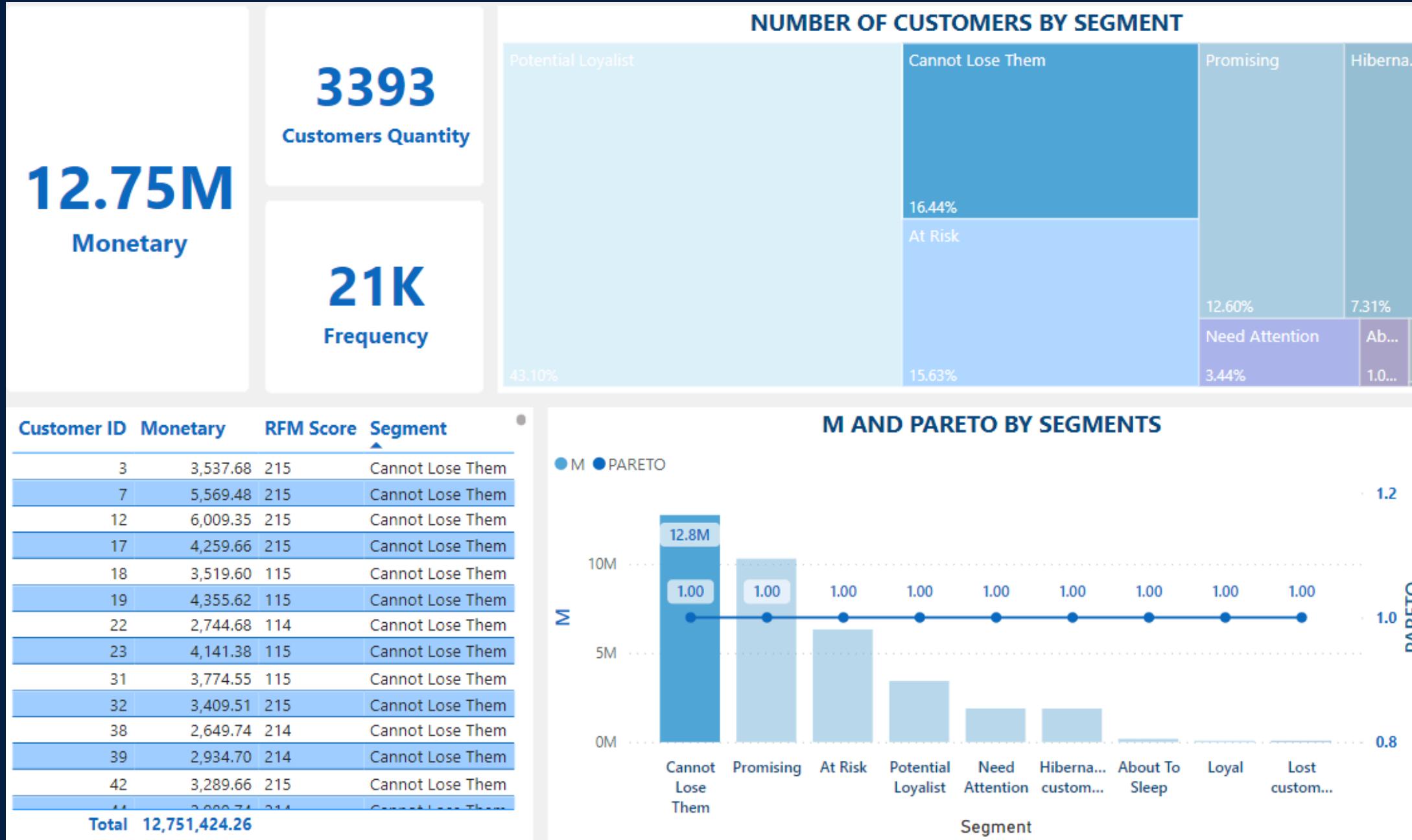
## PHÂN TÍCH RFM



- 3 nhóm khách hàng **Cannot Lose Them, Promising, At Risk** chiếm tới 80% doanh thu.



# Cannot Lose Them



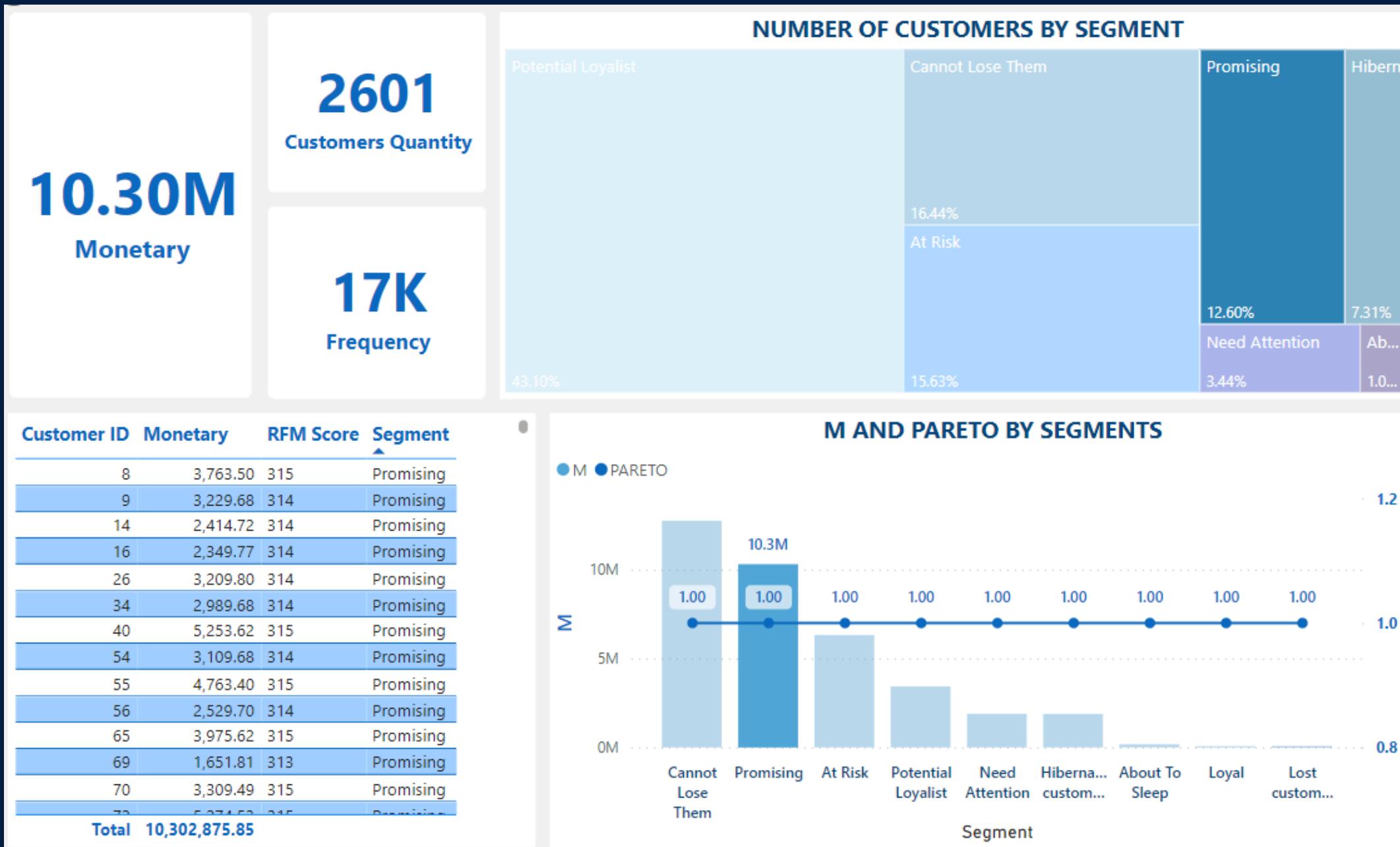
Chỉ chiếm **16,44%** trong tổng số khách hàng nhưng doanh thu đóng góp tới 35%. Họ là nhóm khách hàng thường đặt các đơn hàng lớn nhưng thời gian mua hàng cách đây khá lâu.

→ Nghiên cứu sâu hơn để tìm ra nguyên nhân tại sao thời gian mua hàng của nhóm KH này cách đây khá lâu. **Cải thiện chất lượng** nhóm sản phẩm mà nhóm khách hàng này thường mua.

→ **Tạo ra các chương trình ưu đãi** đặc biệt cho nhóm khách hàng này để khuyến khích họ tiếp tục mua hàng. Ưu tiên hỗ trợ sau khi mua hàng, **thường xuyên liên hệ** để tạo mối quan hệ và niềm tin bán hàng.



# Promising



Chiếm **12,6%** nhưng doanh thu đóng góp **28,1%**. Đây là nhóm khách hàng trung thành, từng mua hàng thường xuyên và chi tiêu rất nhiều nhưng thời gian mua hàng cách đây cũng khá lâu.

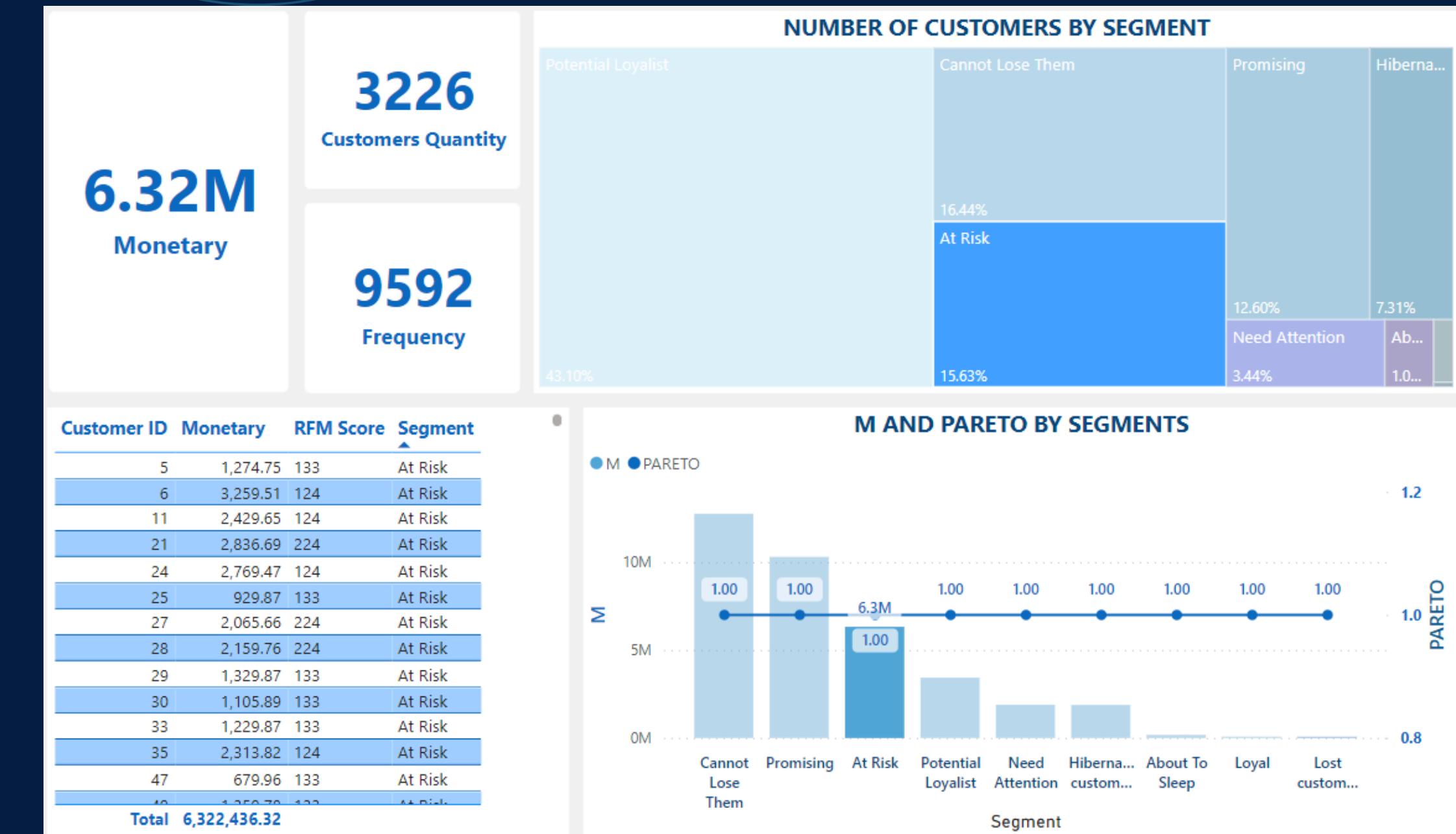
→ Cung cấp các mã ưu đãi, khuyến mãi để kích thích nhu cầu mua hàng. Tăng cường chăm sóc khách hàng sau khi KH mua hàng để giữ chân khách hàng. Xem các nhóm sản phẩm mà nhóm KH này hay mua để trữ tồn kho.



## At Risk

Chiếm **15.63%** trong tổng số khách hàng và có doanh thu đóng góp khoảng **17%**. Mặc dù đặt các đơn hàng giá trị lớn nhưng lần cuối thực hiện mua hàng cách đây rất lâu và cũng không thường xuyên mua hàng.

→ **Tạo ra các chiến dịch** tái kích hoạt dành riêng cho nhóm **At Risk** bằng cách gửi thông điệp nhắc nhở, ưu đãi đặc biệt hoặc các thông tin giới thiệu sản phẩm mới để kích thích họ quay lại mua hàng. **Tìm hiểu nguyên nhân rời bỏ** của nhóm khách hàng này.

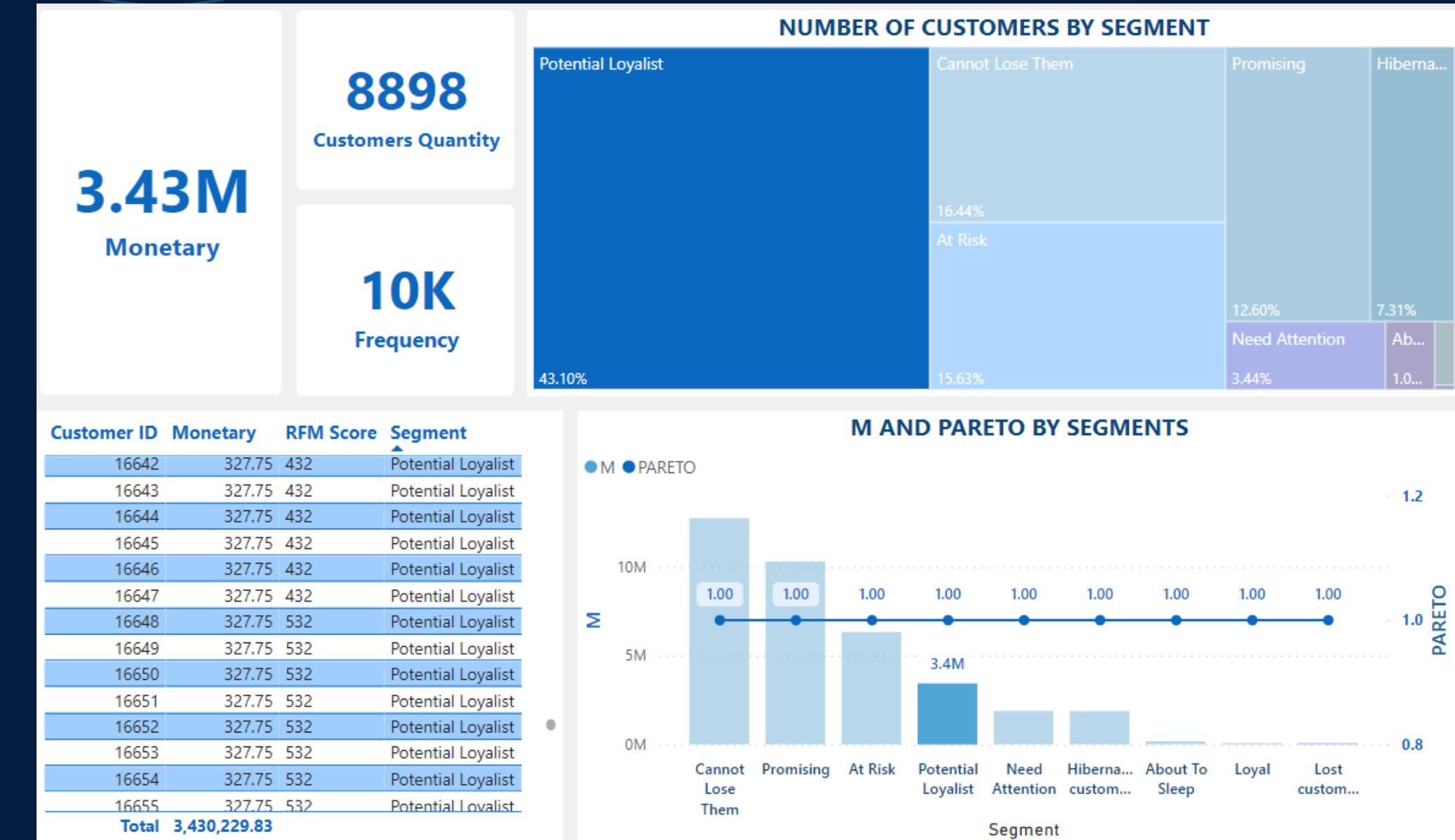




## Potential Loyalist

Chiếm tới **43,1%** trong tổng số KH. Tổng số đơn hàng mua cũng rất lớn nhưng doanh thu của nhóm KH này chỉ chiếm **9,33%**. Đây là nhóm KH mua hàng gần đây, mua hàng thường xuyên nhưng đa phần chỉ mua các đơn hàng có giá trị vừa và nhỏ.

→ Đây là nhóm KH tiềm năng, tương lai có thể phát triển thành Loyal hoặc Champion. **Khuyến khích khách hàng** trong nhóm này mua các sản phẩm có giá trị cao hơn bằng cách tạo ra các gói sản phẩm kết hợp hoặc khuyến mãi khi mua số lượng lớn.



→ **Tạo ra các chương trình khuyến mãi** và ưu đãi đặc biệt dành riêng cho nhóm này để khích lệ họ mua các sản phẩm và dịch vụ cao cấp hơn.

→ Phân tích các dòng sản phẩm được mua thường xuyên bởi nhóm KH này để trữ tồn kho, để đáp ứng kịp nhu cầu mua hàng của KH

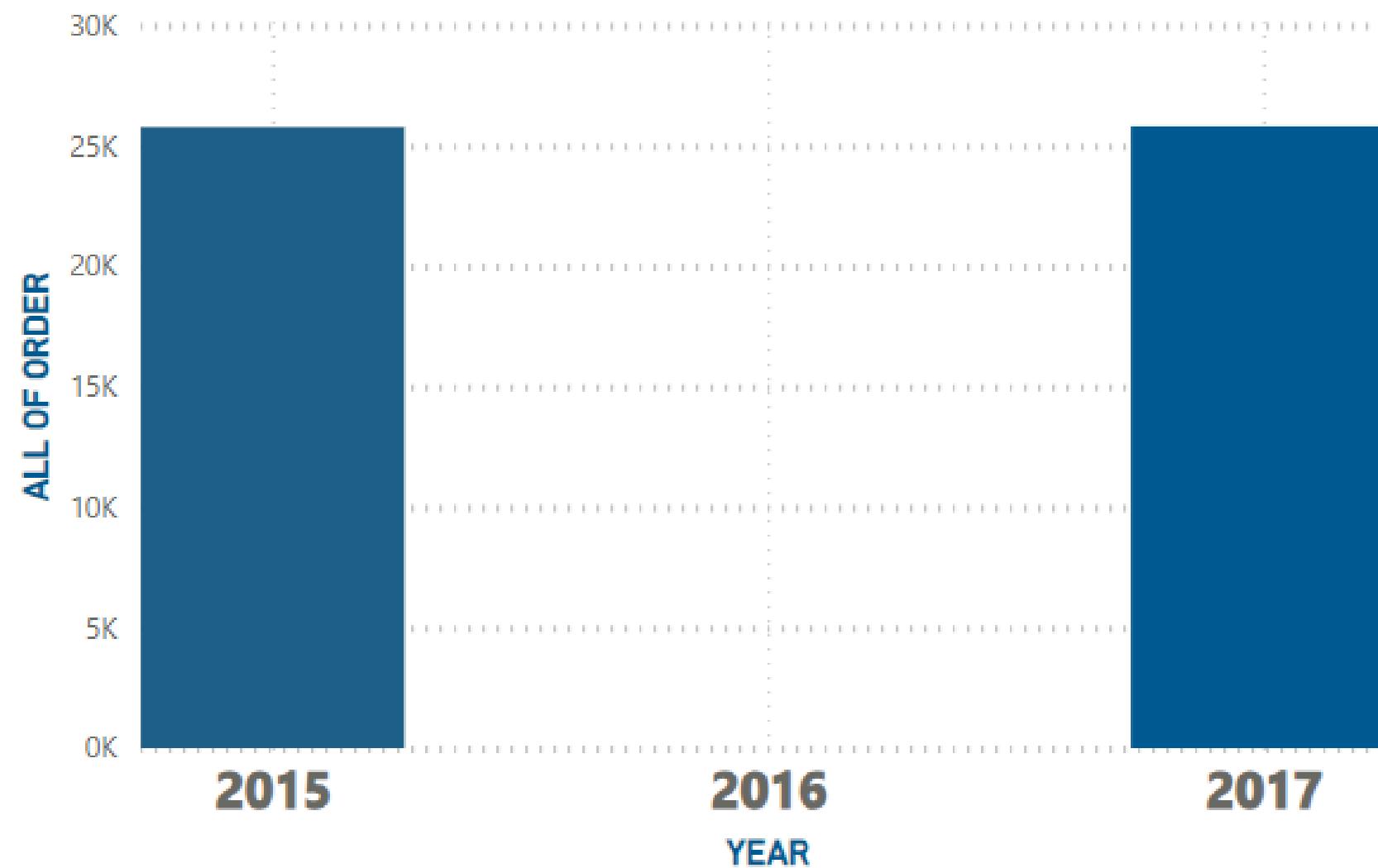


Nhóm B

## PHÂN TÍCH GIAO HÀNG

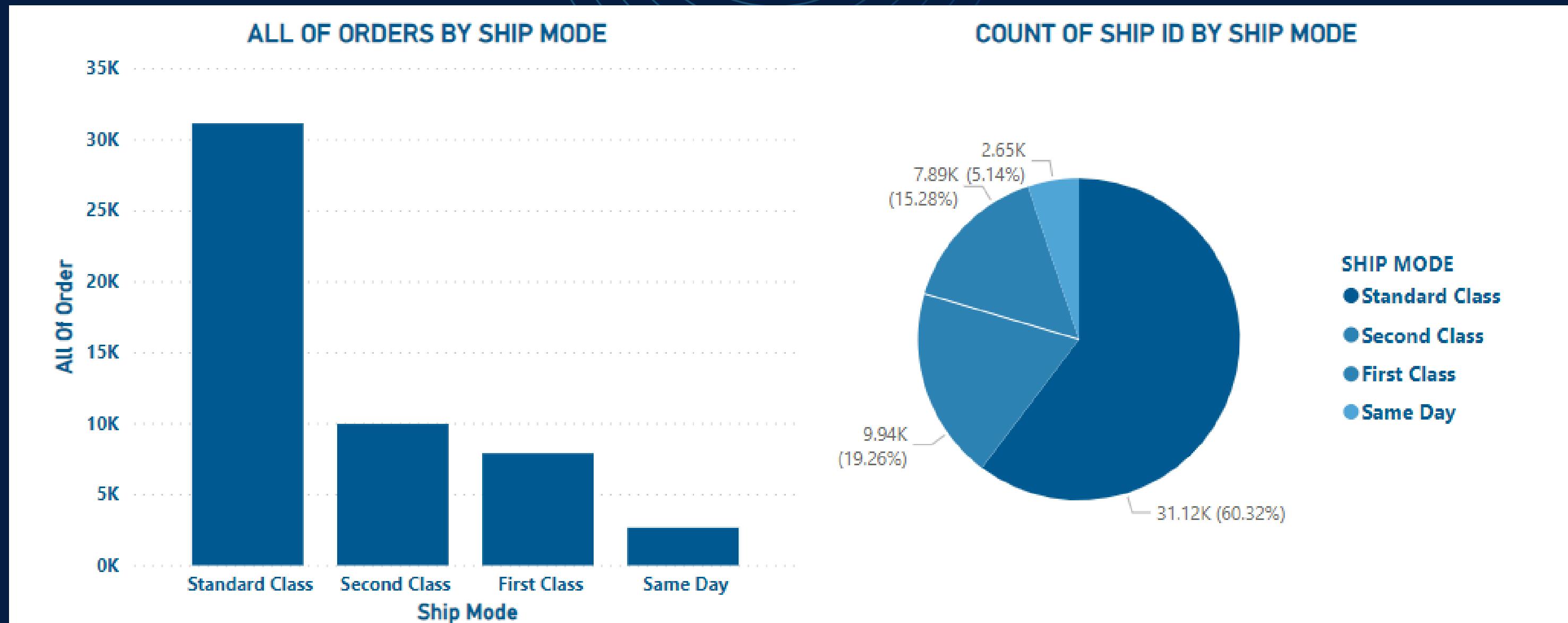


## THE NUMBER OF ORDER BY YEAR



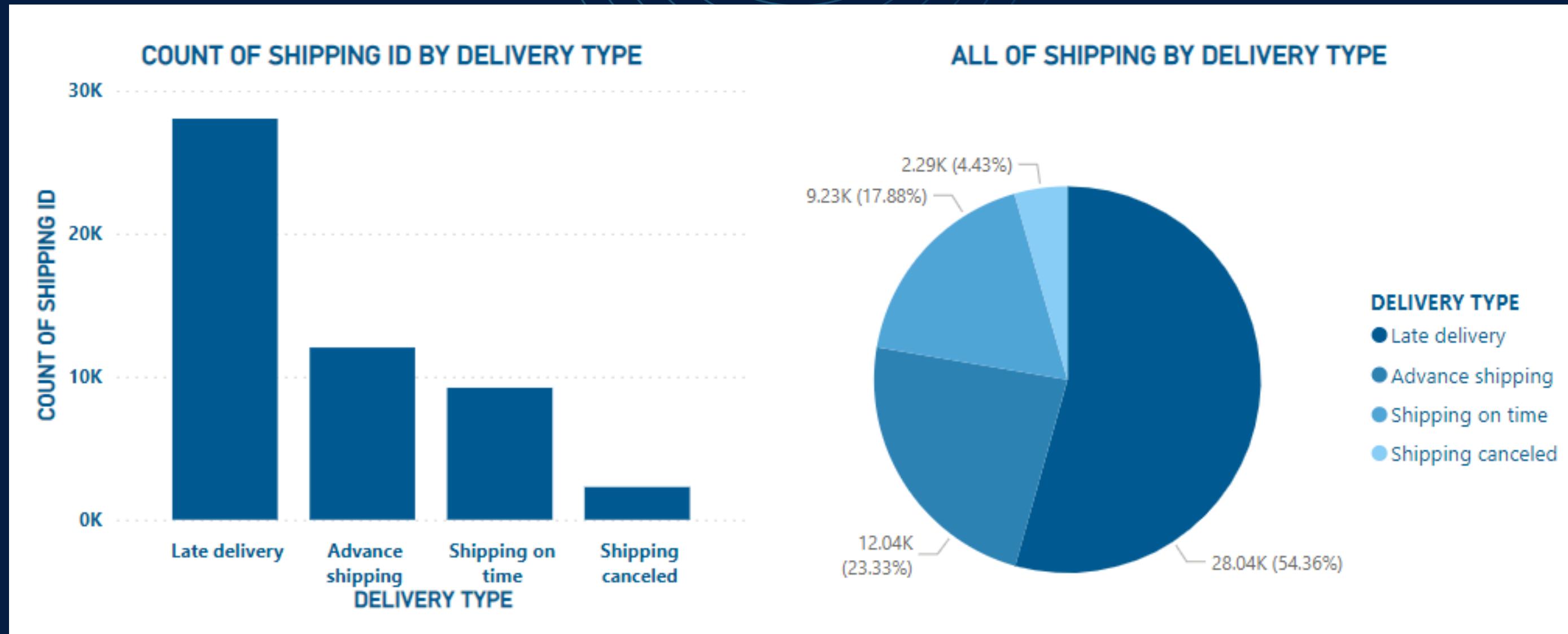
**51.59K**  
ALL OF ORDERS

- Có tổng cộng 51,59K đơn hàng vận chuyển được ghi nhận trong hai năm 2015 và 2017 tại LATAM.
- Với trung bình 25K đơn hàng vận chuyển mỗi năm.
- Trong năm 2016 không có đơn hàng vận chuyển nào được ghi nhận.



- Các đơn hàng vận chuyển được chia thành 4 loại : Standard Class, Second Class, First Class và Same Day.
- Loại vận chuyển được sử dụng nhiều nhất tại đây là **Standard Class** với số lượng chiếm hơn 60% trên tổng số các đơn hàng được ghi nhận tại đây.

→ **Standard Class** là loại vận chuyển **cơ bản, dễ tiếp cận** đối với nhiều đối tượng khách hàng. Nên nhu cầu sử dụng đang ở mức tuyệt đối so với các loại vận chuyển khác.

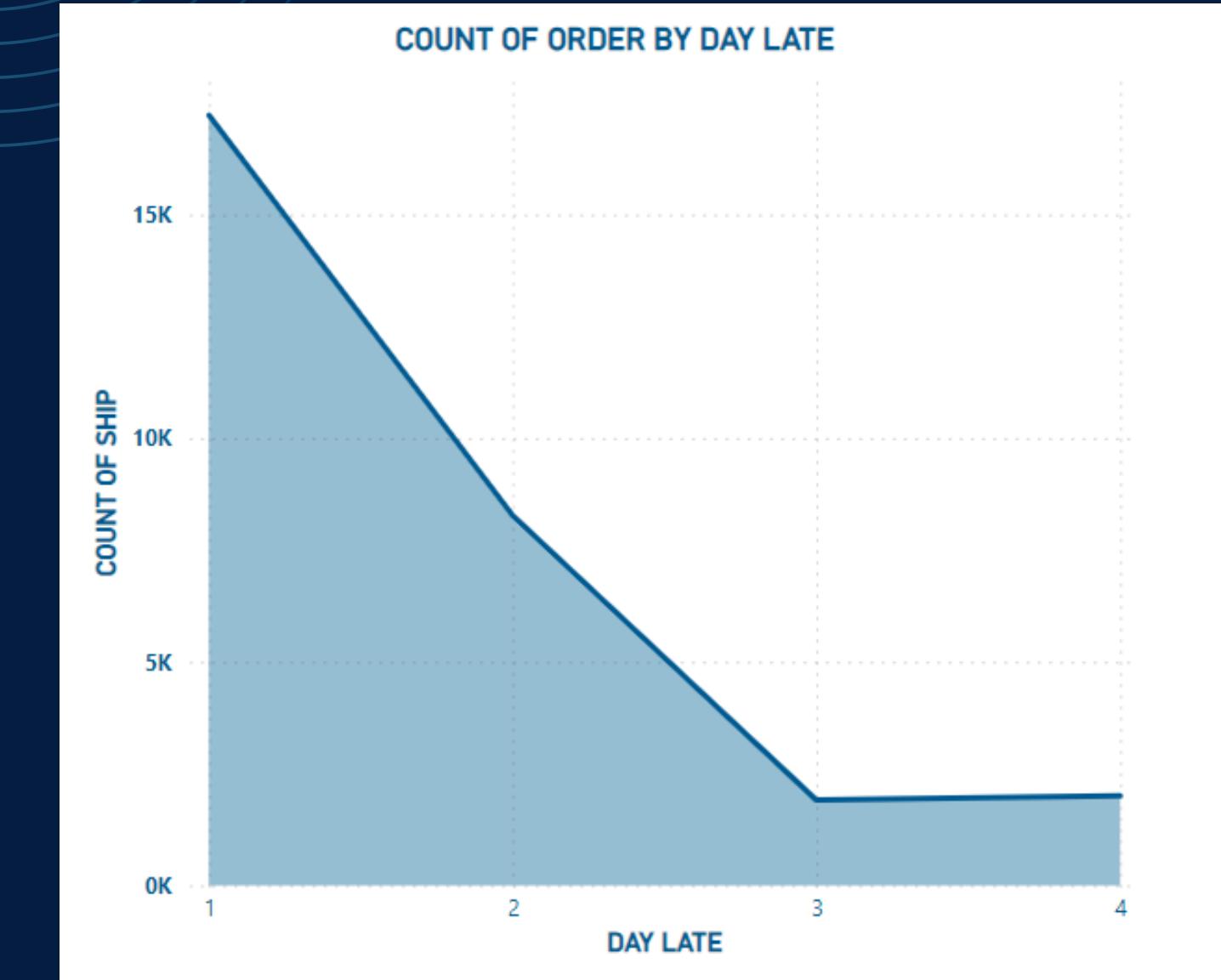


- Các đơn hàng vận chuyển thành công cũng được chia thành 4 loại : **Late Delivery, Advance Ship, Shipping On Time và Shipping Canceled.**
- Các đơn hàng thường được Ship trễ hơn so với ngày dự kiến.

Điều này có thể dẫn đến ấn tượng không tốt trong việc trải nghiệm dịch vụ giao hàng. Và cũng có thể bị đánh giá không tốt bởi các khách hàng bị trễ đơn hàng giao đến.



- Mức độ trễ của các đơn hàng đa số nằm ở mức 1 ngày cho đến 2 ngày.
- Những đơn hàng trễ 3 ngày tới 4 ngày nằm ở khoảng 2,5K đơn mỗi loại.



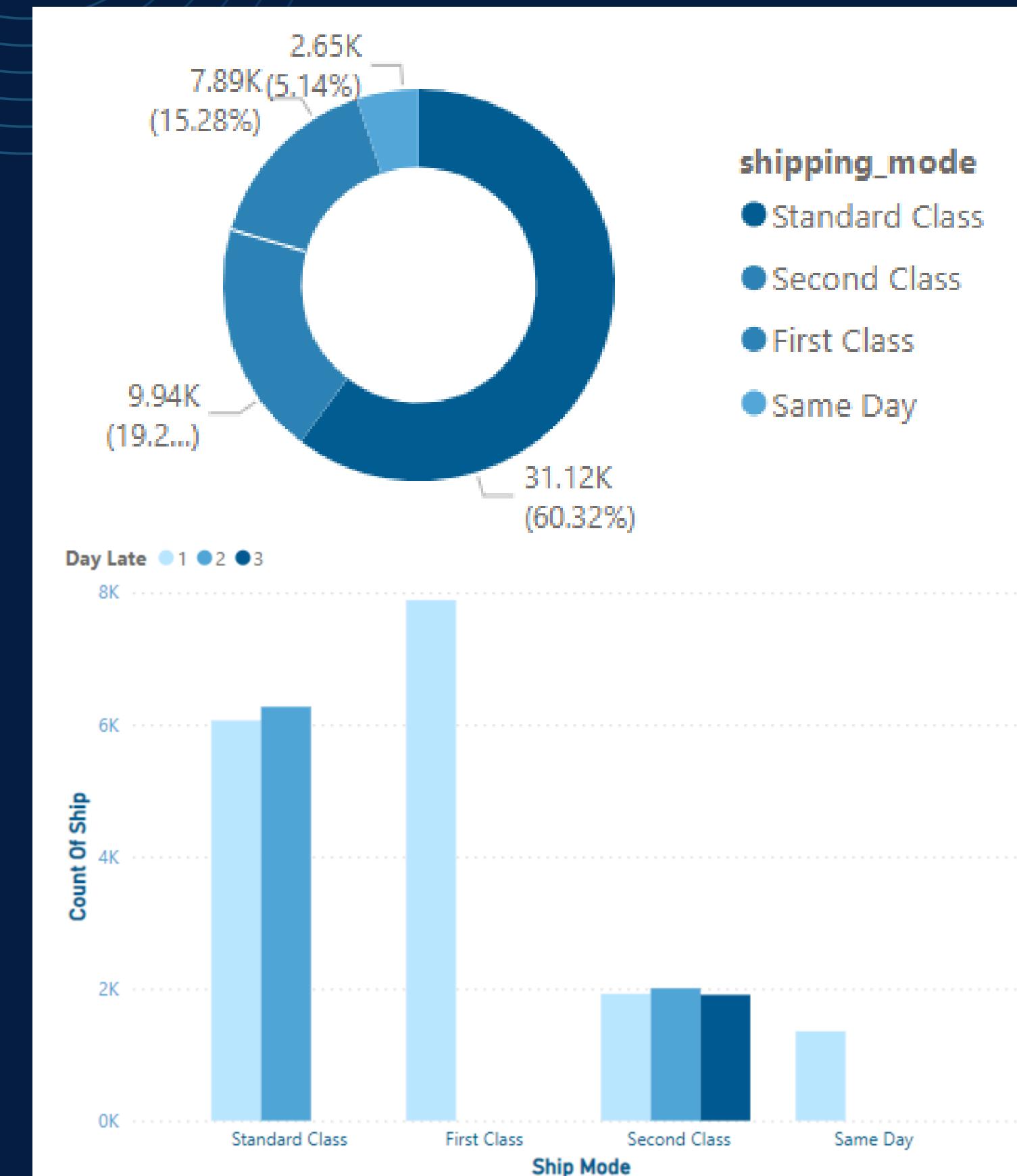
Trong khi hầu hết các đơn hàng được xử lý đúng hạn, có một số lượng đáng kể đơn hàng gãy trễ, đặc biệt là trong khoảng thời gian 3 đến 4 ngày. Điều này cần sự quan tâm và cải thiện từ phía công ty để đảm bảo rằng tất cả các đơn hàng được xử lý và giao hàng đúng hẹn, giúp tăng cường hài lòng của khách hàng và hiệu suất tổng thể của doanh nghiệp.

# PHÂN TÍCH ĐƠN GIAO HÀNG



Nhóm B

- **Standard Class** : 60,32% tổng số lượt giao hàng được vận chuyển với Hạng Tiêu chuẩn. Trong số những lần giao hàng này, 39% là giao hàng trễ. Độ trễ trung bình là khoảng hai ngày.
- **Frist Class** : 15,28% tổng số lần giao hàng được vận chuyển bằng Frist Class. Đó là Phương thức Vận chuyển đáng tin cậy nhất.
- **Second Class** : 19,26% tổng số lần giao hàng được vận chuyển bằng Second Class. Khoảng 76% tổng số lần giao hàng bị trễ. Độ trễ trung bình là khoảng hai ngày
- **Same Day** : 5,14% tổng số lần giao hàng được vận chuyển trong cùng ngày. Gần một nửa (46%) số lần giao hàng bị trễ. Độ trễ trung bình là một ngày





Chế độ giao hàng Standard Class, Second Class, First Class có số lượng đơn hàng giao trễ khá lớn. Trong đó ở First Class, số lượng đơn hàng bị giao trễ cao hơn rất nhiều so với các đơn giao đúng hẹn.

→ Trong khi cả ba chế độ giao hàng đều gặp vấn đề với giao hàng trễ, First Class xuất hiện với tỷ lệ giao hàng trễ cao hơn đáng kể so với các chế độ khác. Điều này đòi hỏi sự tập trung đặc biệt để cải thiện hiệu suất giao hàng trong lớp này để đảm bảo rằng các yêu cầu của khách hàng được đáp ứng đúng hẹn và tăng cường hài lòng của họ.



## PHÂN TÍCH COHORT TẠI LATAM NĂM 2017



COHORT OF SHIP LATE						
Month	New User Late	Month Rate 1	Month Rate 2	Month Rate 3	Month Rate 4	Month Rate 5
1	100%	11.84%	12.09%	14.36%	14.61%	6.80%
2	100%	13.74%	13.74%	12.25%	5.32%	0.00%
3	100%	11.29%	13.88%	8.82%	0.00%	0.00%
4	100%	14.67%	9.10%	0.00%	0.00%	0.00%
5	100%	6.22%	0.00%	0.00%	0.00%	0.00%
6	100%	0.00%	0.00%	0.00%	0.00%	0.00%

Month	New User Late	Month 1	Month 2	Month 3	Month 4	Month 5
1	397	47	48	57	58	27
2	808	111	111	99	43	0
3	850	96	118	75	0	0
4	736	108	67	0	0	0
5	723	45	0	0	0	0
6	360	0	0	0	0	0

- Các đơn hàng mới gia tăng từ tháng 2 - tháng 5.
- Hai tháng đầu và giữa năm có xu hướng giảm.
- Khi có một đơn hàng bị trễ mới vào mỗi tháng thì mức độ người khách hàng đó lại bị trễ tiếp tục nằm khoảng 10% - 15 %.

→ Mặc dù có sự tăng về số lượng đơn hàng từ tháng 2 đến tháng 5, nhưng xu hướng giảm trong hai tháng đầu và giữa năm, cùng với mức độ đơn hàng bị trễ tăng lên, có thể đề xuất một sự không ổn định trong hoạt động kinh doanh và cần phải được quản lý cẩn thận để đảm bảo hiệu suất và sự hài lòng của khách hàng.



COHORT OF SHIP ONTIME						
Month	New User Ontime	Month Rate 1	Month Rate 2	Month Rate 3	Month Rate 4	Month Rate 5
1	100%	11.36%	13.31%	12.34%	11.69%	6.17%
2	100%	12.84%	14.45%	13.32%	7.87%	0.00%
3	100%	14.55%	12.73%	6.82%	0.00%	0.00%
4	100%	13.74%	7.51%	0.00%	0.00%	0.00%
5	100%	7.66%	0.00%	0.00%	0.00%	0.00%
6	100%	0.00%	0.00%	0.00%	0.00%	0.00%

Month	New User Ontime	Month 1	Month 2	Month 3	Month 4	Month 5
1	308	35	41	38	36	19
2	623	80	90	83	49	0
3	660	96	84	45	0	0
4	626	86	47	0	0	0
5	561	43	0	0	0	0
6	307	0	0	0	0	0

- Cũng giống như các đơn hàng Late thì ở các đơn hàng Ontime số lượng khách hàng mới ở các tháng từ 2 tới 5 có xu hướng tăng so với 2 tháng đầu và giữa năm
- Sản lượng khách hàng mới quay lại ở mức trung bình với khoảng 10% đến 15% mỗi tháng.
- Chúng ta có thể thấy % của tháng 6 ở mức thấp hơn so với các tháng còn lại khoảng 3 - 4 %

→ Trong khi số lượng khách hàng mới trong các đơn hàng đúng hạn có xu hướng tăng và tỷ lệ quay lại ổn định, việc đánh giá kỹ lưỡng các vấn đề đặc biệt, như sự giảm của tháng 6, là quan trọng để đảm bảo sự ổn định và tăng trưởng



## IV. KẾT LUẬN, ĐỀ XUẤT

## #1. Cải thiện sản phẩm để tối ưu hóa doanh thu

- Đối với các dòng sản phẩm của **Nike Men's**, **O'Brien**, xem xét điều chỉnh giá để cải thiện về doanh thu, nhưng cũng nên có những ưu đãi cho khách hàng cũng như điều chỉnh giá hợp lý để không làm mất khách hàng hoặc số lượng bán sản phẩm.
- Đẩy mạnh marketing với dòng sản phẩm **Field & Stream Sports**
- Tăng cường nhập các dòng sản phẩm thuộc **Cleats**, **Fishing** để tối ưu hóa doanh số bán hàng.



## #2. Tối ưu hóa thanh toán

- **Tạo ra các chương trình ưu đãi và khuyến mãi** giảm giá dành riêng cho các phương thức thanh toán cụ thể để kích thích mua hàng.
- **Xây dựng và tích hợp các hệ thống thanh toán trực tuyến** như các cổng thanh toán, các cổng thanh toán của bên thứ ba hoặc các cổng thanh toán ngân hàng để thuận tiện cho khách hàng mua hàng trực tuyến.



## #3. Kiểm soát phân khúc khách hàng

### Cannot Lose Them

Nghiên cứu sâu hơn để tìm ra nguyên nhân tại sao thời gian mua hàng của nhóm KH này cách đây khá lâu. **Cải thiện chất lượng** nhóm sản phẩm mà nhóm khách hàng này thường mua.

**Tạo ra các chương trình ưu đãi** đặc biệt cho nhóm khách hàng này để khuyến khích họ tiếp tục mua hàng. Ưu tiên hỗ trợ sau khi mua hàng, **thường xuyên liên hệ** để tạo mối quan hệ và niềm tin bán hàng.

### Promising

Cung cấp các mã ưu đãi, khuyến mãi để kích thích nhu cầu mua hàng. Tăng cường chăm sóc khách hàng sau khi KH mua hàng để giữ chân khách hàng. Xem các nhóm sản phẩm mà nhóm KH này hay mua để trữ tồn kho.

## #3. Kiểm soát phân khúc khách hàng

### At Risk

Tạo ra các chiến dịch tái kích hoạt dành riêng cho nhóm **At Risk** bằng cách gửi thông điệp nhắc nhở, ưu đãi đặc biệt hoặc các thông tin giới thiệu sản phẩm mới để kích thích họ quay lại mua hàng. **Tìm hiểu nguyên nhân rời bỏ** của nhóm khách hàng này.

### Potential Loyalist

**Khuyến khích khách hàng** trong nhóm này mua các sản phẩm có giá trị cao hơn bằng cách tạo ra các gói sản phẩm kết hợp hoặc khuyến mãi khi mua số lượng lớn.

Tạo ra các chương trình khuyến mãi và ưu đãi đặc biệt dành riêng cho nhóm này để khích lệ họ mua các sản phẩm và dịch vụ cao cấp hơn.

Phân tích các dòng sản phẩm được mua thường xuyên bởi nhóm KH này để trữ tồn kho, để đáp ứng kịp nhu cầu mua hàng của KH

## #4. Cải thiện chất lượng về mặt giao hàng

- Tối ưu hóa quy trình giao hàng.
- Nâng cao chất lượng dịch vụ First Class.
- Tăng cường quản lý độ trễ.
- Tăng cường trải nghiệm khách hàng.
- Đầu tư vào hạ tầng và công nghệ.



Nhóm B

# THANK YOU FOR LISTENING!

