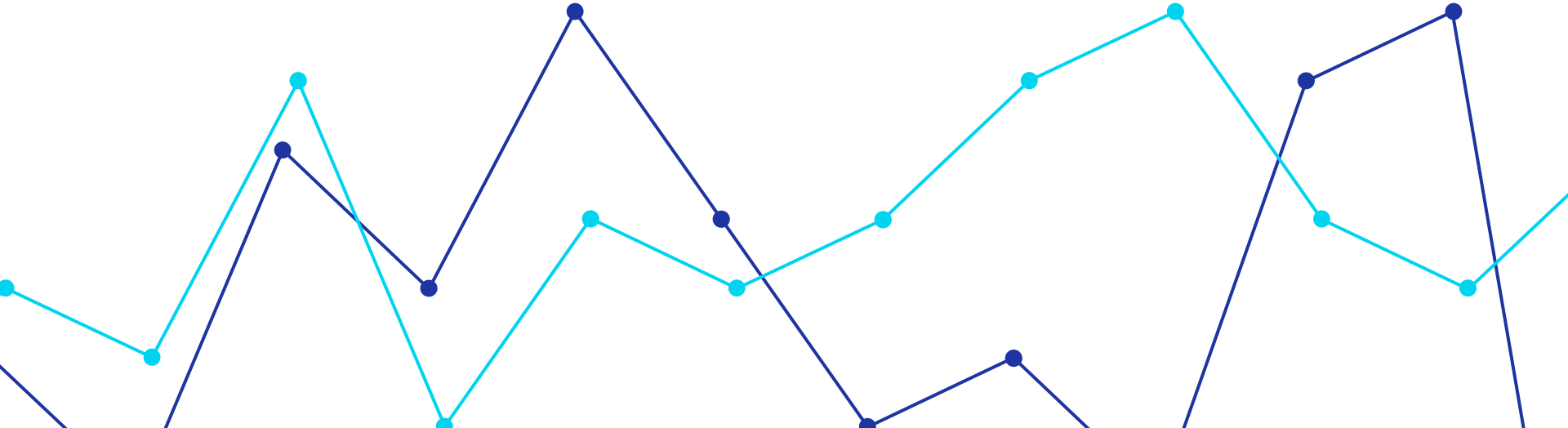
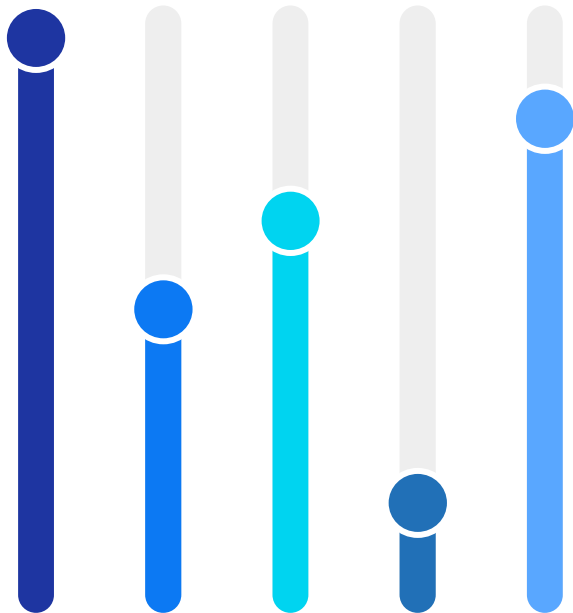


Data Co Supply Chain Analysis

Nguyễn Minh Phát



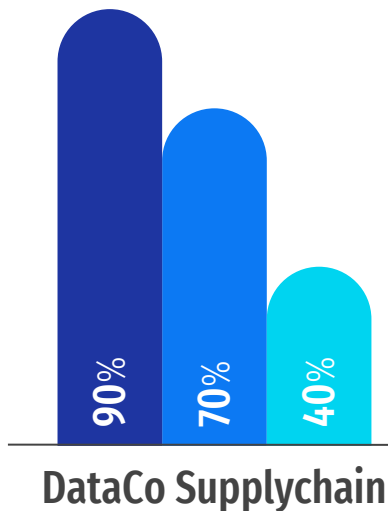
Nội dung



1. Giới thiệu về dữ liệu
 - + Thông tin về dữ liệu
 - + Làm sạch dữ liệu
2. Phân tích dữ liệu
 - + Tổng quan
 - + Vấn đề giao hàng trễ trong chuỗi cung ứng
3. Đề xuất cải thiện

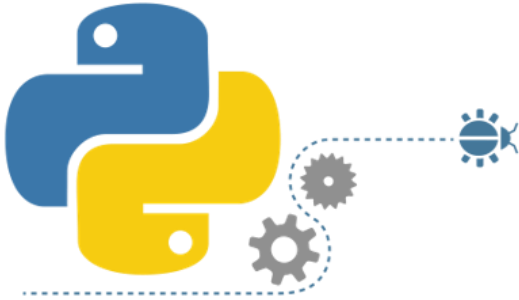
1. Giới thiệu về dữ liệu

Thông tin về dữ liệu



- DataCo supply chains dataset được cung cấp bởi công ty DataCo Global.
- Dữ liệu chứa các thông tin về phương thức thanh toán, phương thức vận chuyển, tình trạng giao hàng, thông tin đơn hàng, khách hàng theo từng khu vực và thời gian, v.v...
- Dataset gồm: 52 cột và 180519 dòng

Làm sạch dữ liệu



```
cor = df.corr()  
top = cor.index  
plt.figure(figsize=(30,20))  
heatm = sns.heatmap(df[top].corr(), annot = True, cmap='bone')  
heatm;
```

- Dùng Python để làm sạch dữ liệu.
- Theo quan sát, trong 52 cột thì sẽ có vài cột trùng nhau => xóa và chỉ giữ lại một cột để phân tích (dùng heatmap trong thư viện seaborn để tìm ra cột trùng dữ liệu)



Làm sạch dữ liệu

Chọn những giá trị là 1 (không nằm trong đường chéo) thì ta có kết quả các cột giống nhau là:

Benefit per order = **Order Profit Per Order**

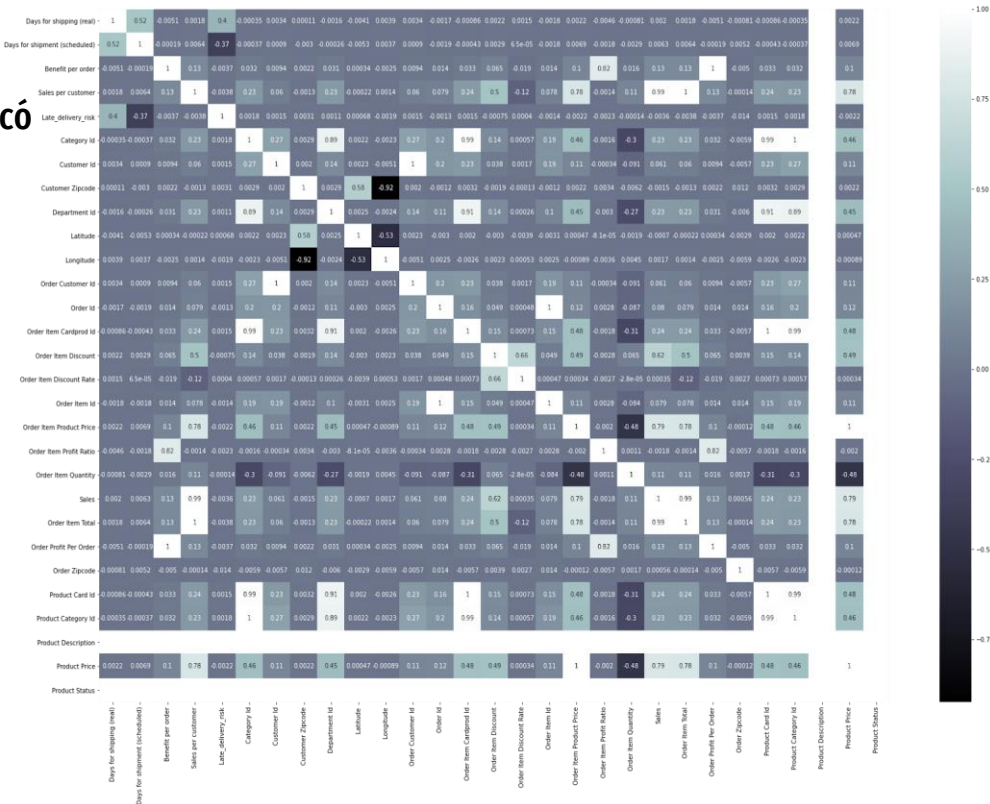
Sales per customer = **Order Item Total**

Category ID = **Product Category Id**

Customer ID = **Order Customer Id**

Product card ID = **Order Item Cardprod Id**

Order Item Product Price = **Product Price**





Làm sạch dữ liệu

Xem xét các cột chứa NULL:

```
for i in df.columns:  
    if df[i].isna().sum()>0:  
        print(i, 'có tổng null là:', df[i].isna().sum())
```

Customer Lname có tổng null là: 8

Customer Zipcode có tổng null là: 3

Order Zipcode có tổng null là: 155679

Product Description có tổng null là: 180519

Xóa cột Order Zipcode và Product Description
(tỉ lệ NULL chiếm hơn 75% tổng records)

2. Tổng quan

Doanh thu, lợi nhuận và số đơn hàng theo khu vực

Doanh thu

36.78M

Lợi nhuận

3.97M

Số đơn hàng

65.75k

Top 1

Western Europe

5.89M

625K

10.01k

Top 2

Central America

5.67M

616K

9396

Top 3

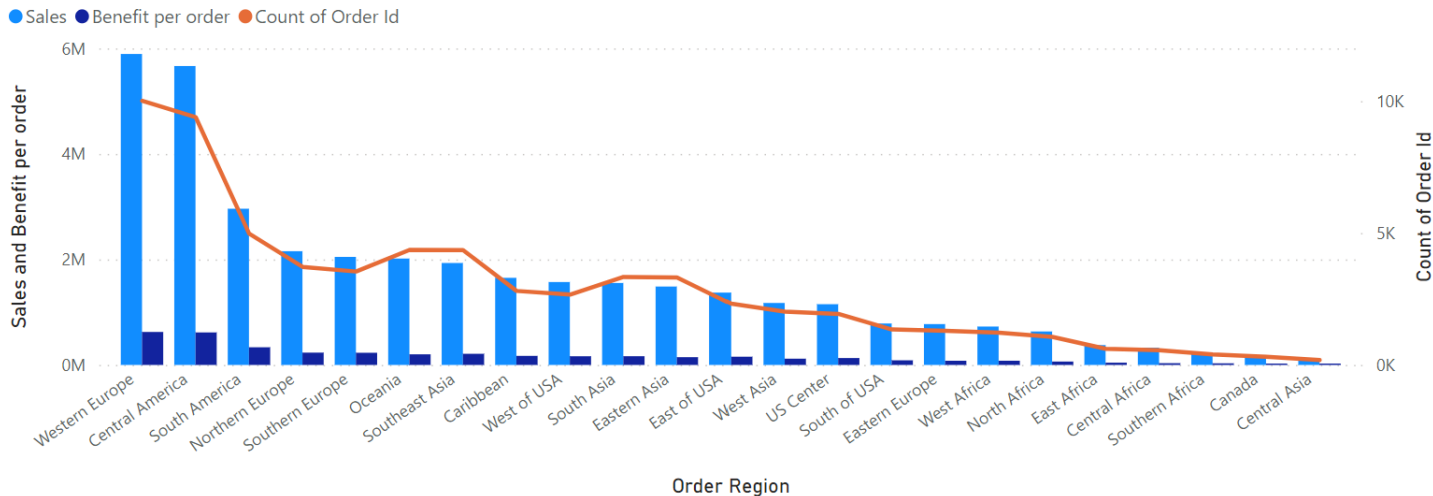
South America

2.96M

335K

4979

Sales, Benefit per order and Count of Order Id by Region



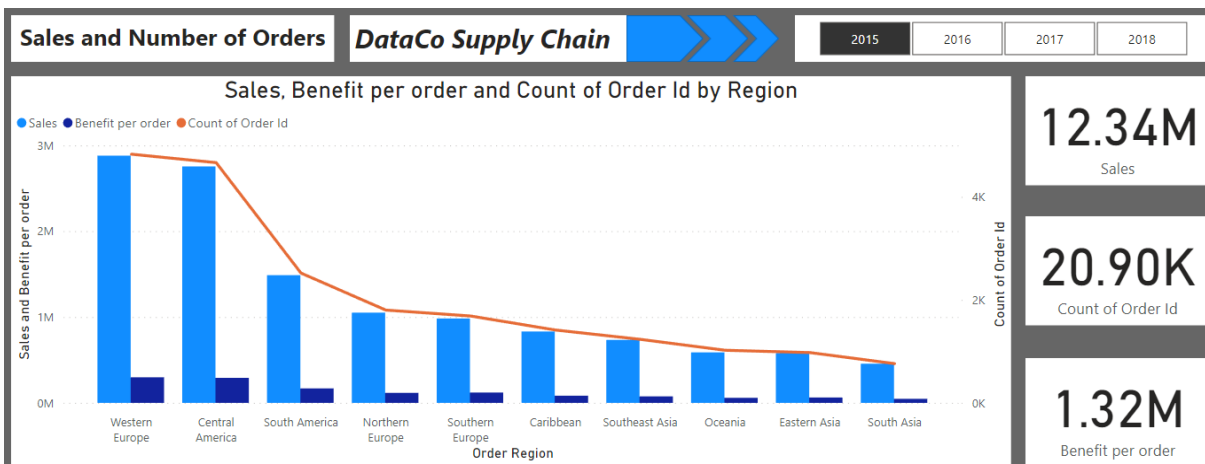
2015



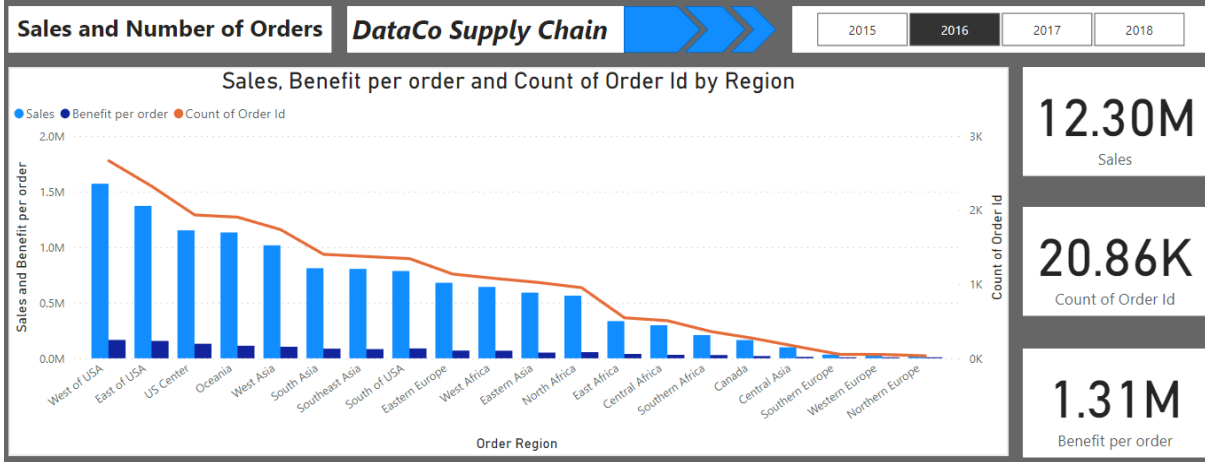
2018

Doanh thu theo khu vực

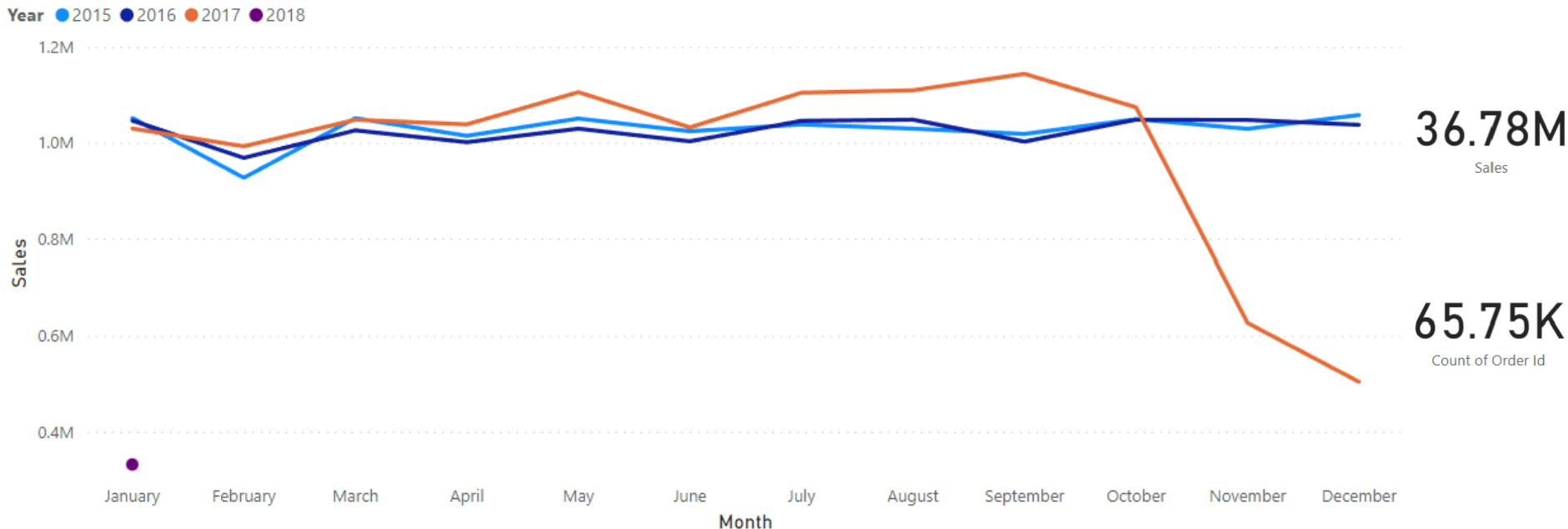
2015



2016

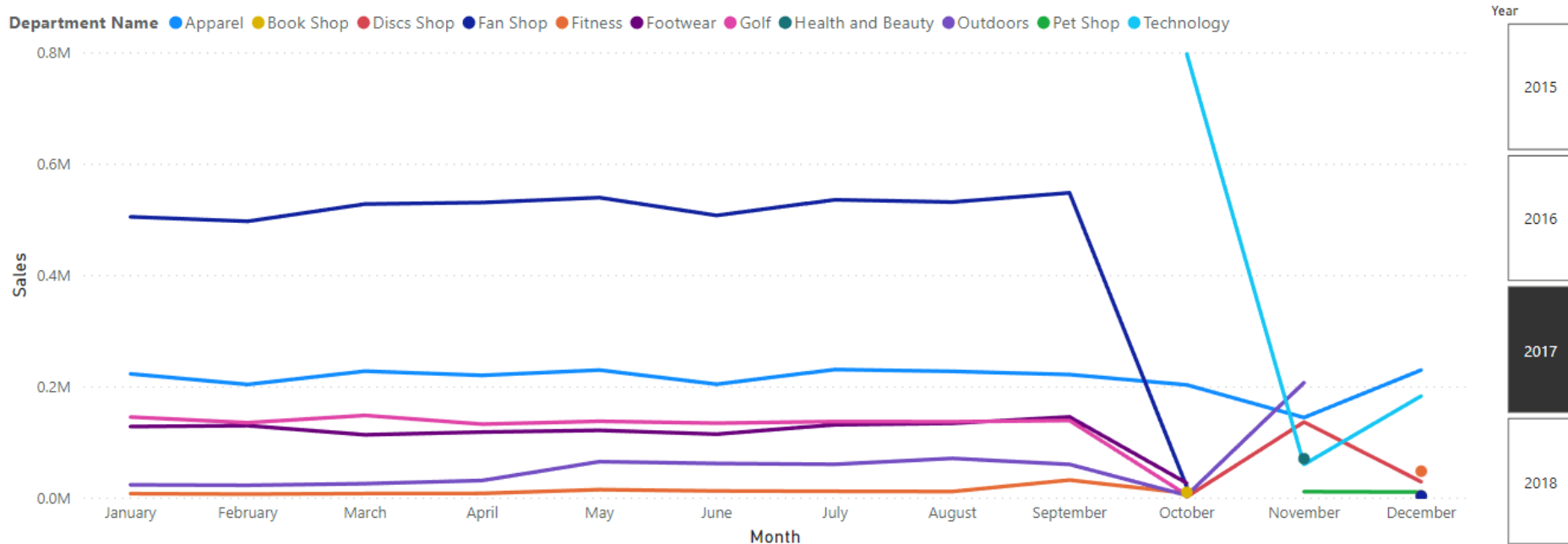


Tổng quan doanh thu qua các năm



Doanh thu năm 2017 **tăng nhẹ** ở các tháng cùng kì so với năm 2015 và 2016.
Tuy nhiên, doanh thu năm 2017 vào **tháng 11 và 12** có sự **giảm mạnh**.

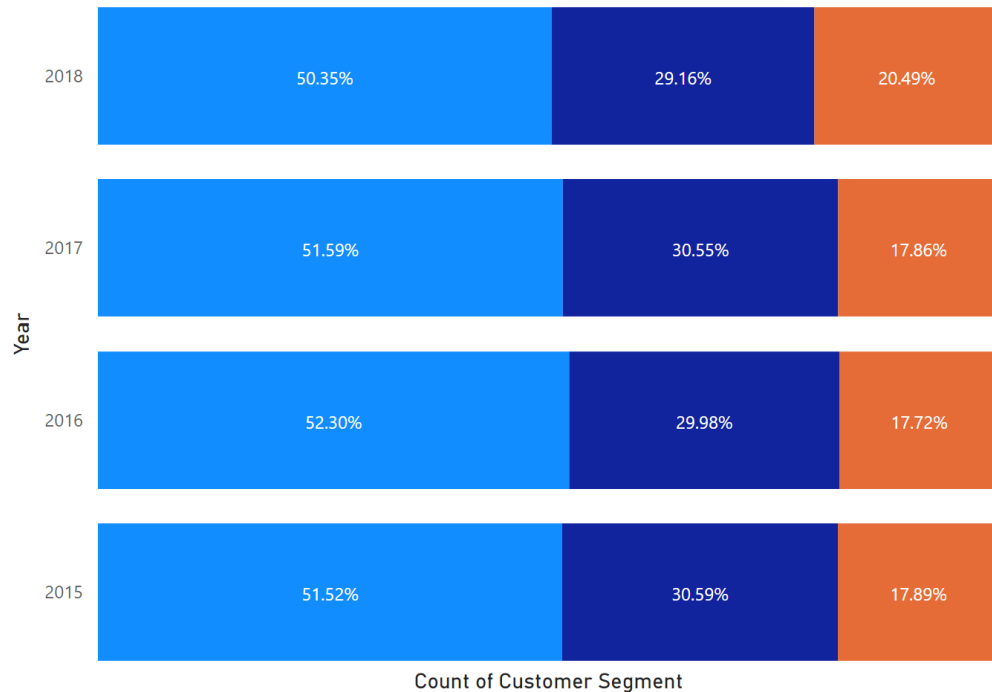
Doanh thu theo phân loại hàng hóa



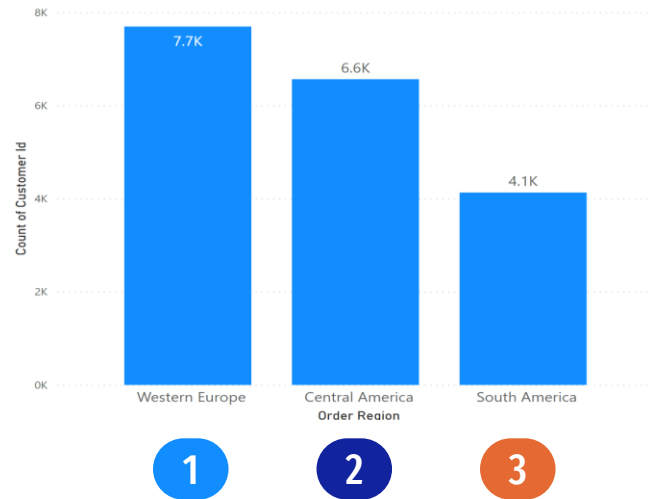
Xét theo năm 2017, doanh thu về mặt hàng **Fan Shop** chiếm tỉ lệ **cao nhất**. Doanh thu **tất cả** mặt hàng **giảm mạnh** vào **tháng 10**, trừ mặt hàng **Apparel**. Mặt hàng **Technology** được phân phối vào tháng 10 nhưng có **sự giảm mạnh** vào **tháng 11**,

Phân khúc khách hàng

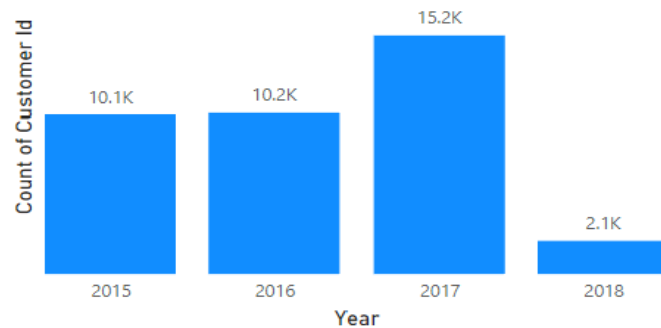
Customer Segment ● Consumer ● Corporate ● Home Office



Consumer chiếm 50% trong phân khúc khách hàng.
Năm 2017, số lượng khách hàng tăng 50% so với năm 2016.
Năm 2018, dữ liệu không đủ 12 tháng -> không thể so sánh.

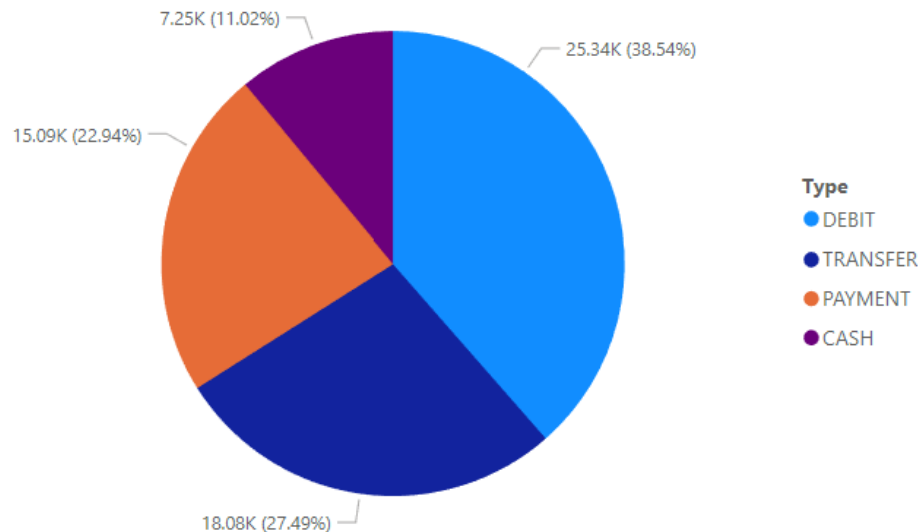


Count of Customer Id by Year

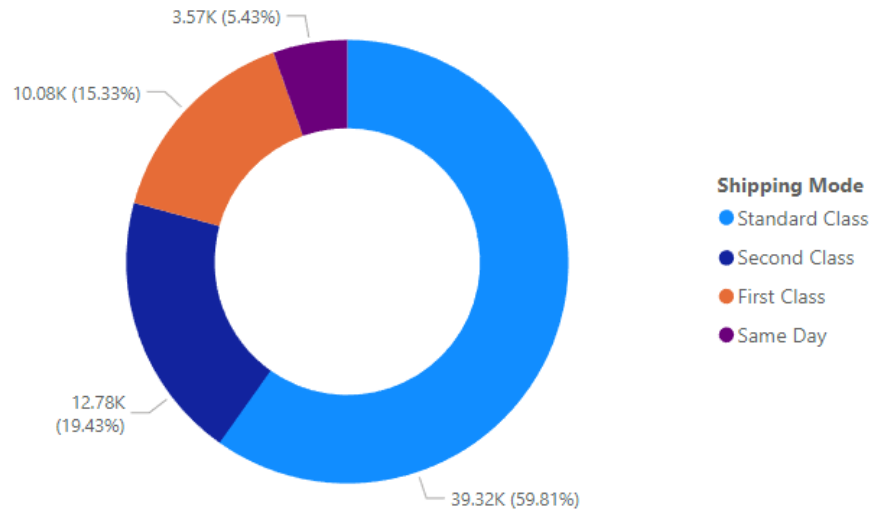


Phương thức thanh toán và vận chuyển

Count of Order Id by Type



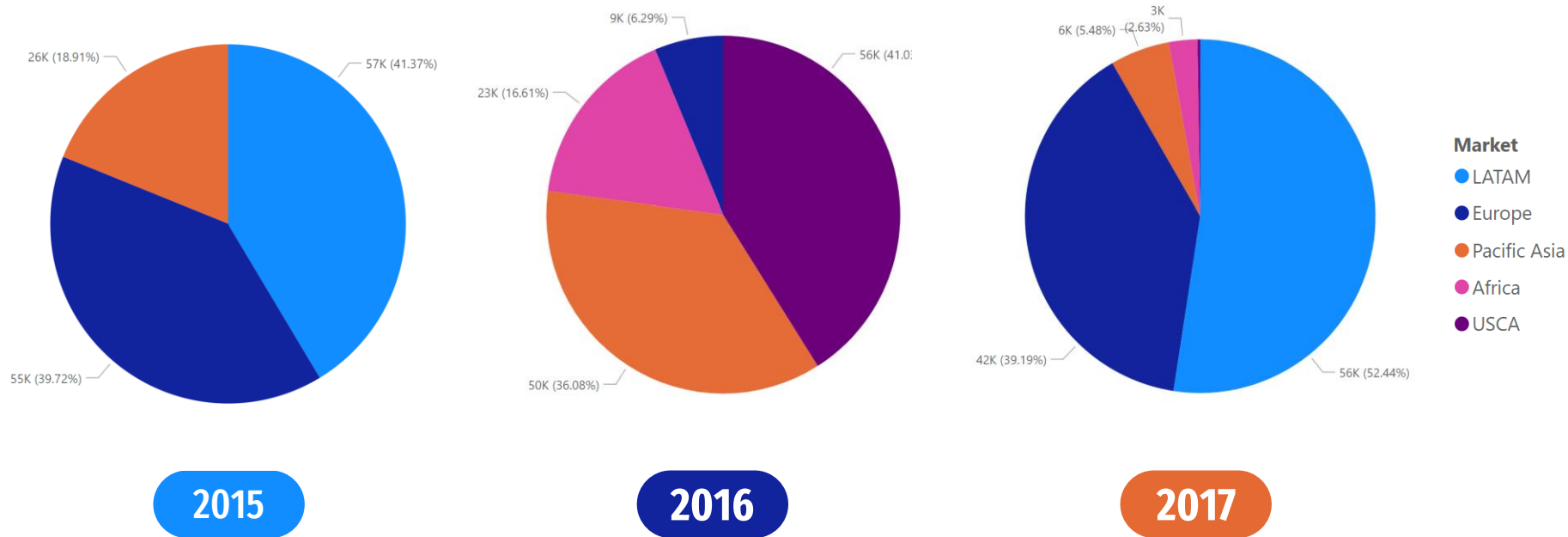
Count of Order Id by Shipping Mode



Hình thức thanh toán **trực tuyến** chiếm phần lớn (89%) so với trả bằng **tiền mặt** (11%)

Phương thức vận chuyển **Standard Class** chiếm gần 60%

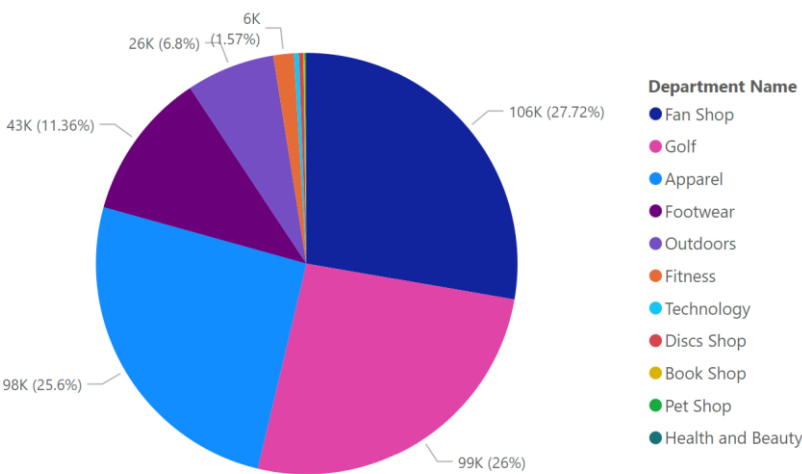
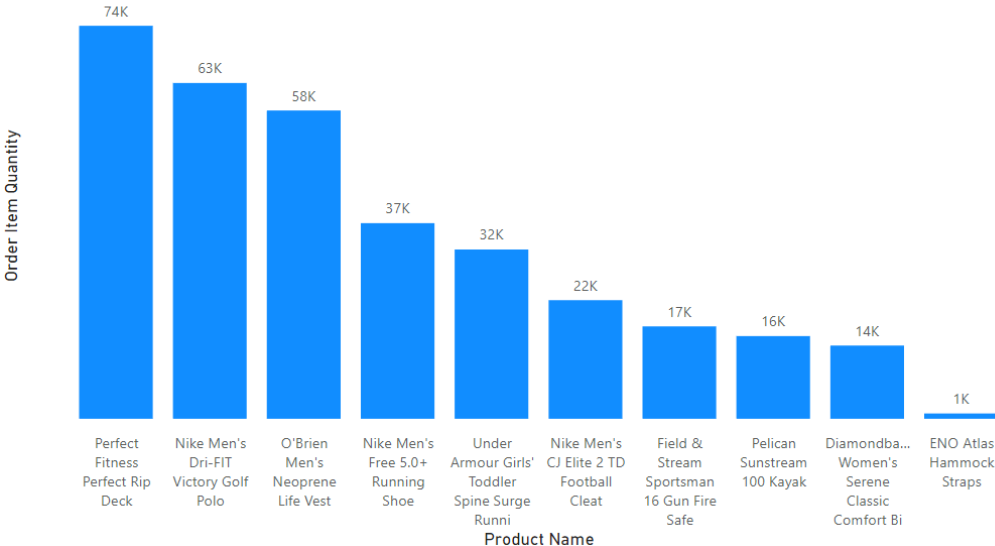
Số lượng sản phẩm theo thị trường



LATAM (Latin America) có số lượng sản phẩm theo đơn vận chuyển **hiều nhất** (2015 – 41.37%, 2017 – 52.44%). Năm 2016, khu vực **USCA** đứng đầu trong danh sách (41.03%)

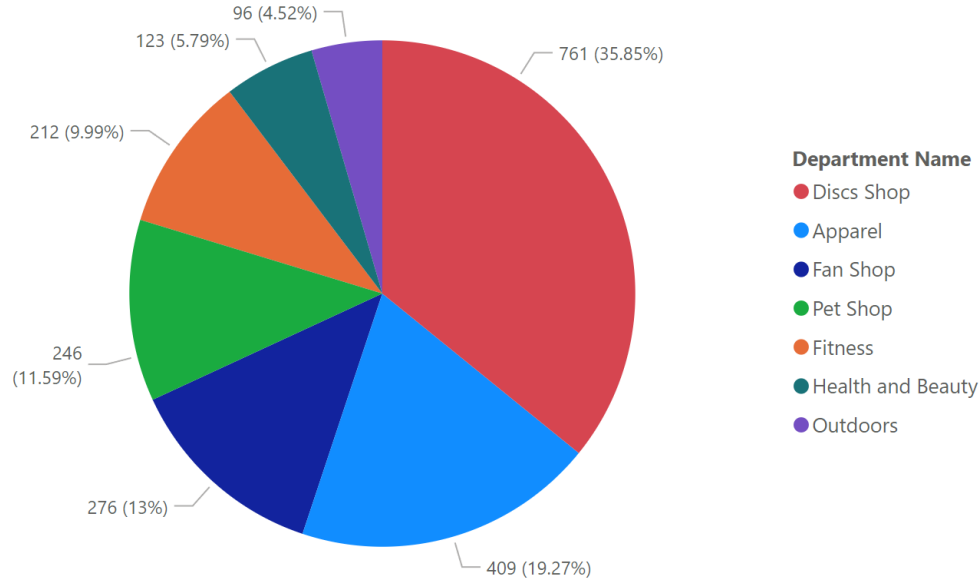
Số lượng đơn hàng theo mã & bộ phận

Order Item Quantity by Product Name



2015 -----> 2017

Số lượng đơn hàng theo mã & bộ phận



Product Name	Order Item Quantity
Fighting video games	610
Summer dresses	378
Toys	276
Adult dog supplies	246
Smart watch	212
Rock music	144
First aid kit	123
Lawn mower	96
Men's gala suit	31
DVDs	7
Total	2123

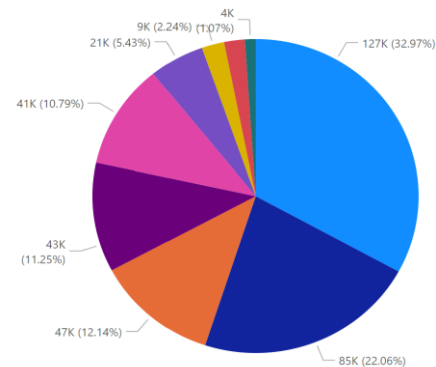
2018

Trạng thái đặt hàng

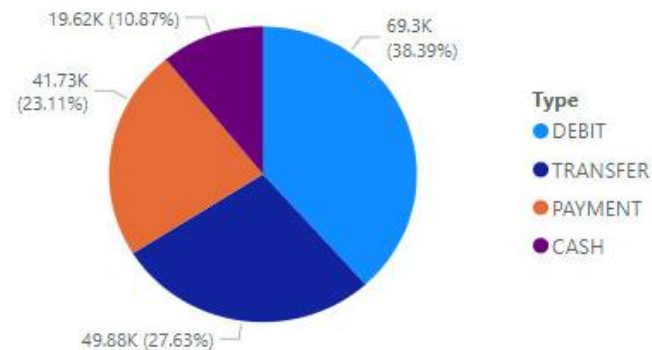
Order Status	CASH	DEBIT	PAYMENT	TRANSFER	Total
CANCELED				1367	1367
CLOSED	7249				7249
COMPLETE		21716			21716
ON_HOLD		3624			3624
PAYMENT_REVIEW			704		704
PENDING				7321	7321
PENDING_PAYMENT			14382		14382
PROCESSING				7901	7901
SUSPECTED_FRAUD				1488	1488
Total	7249	25340	15086	18077	65752

Order Status

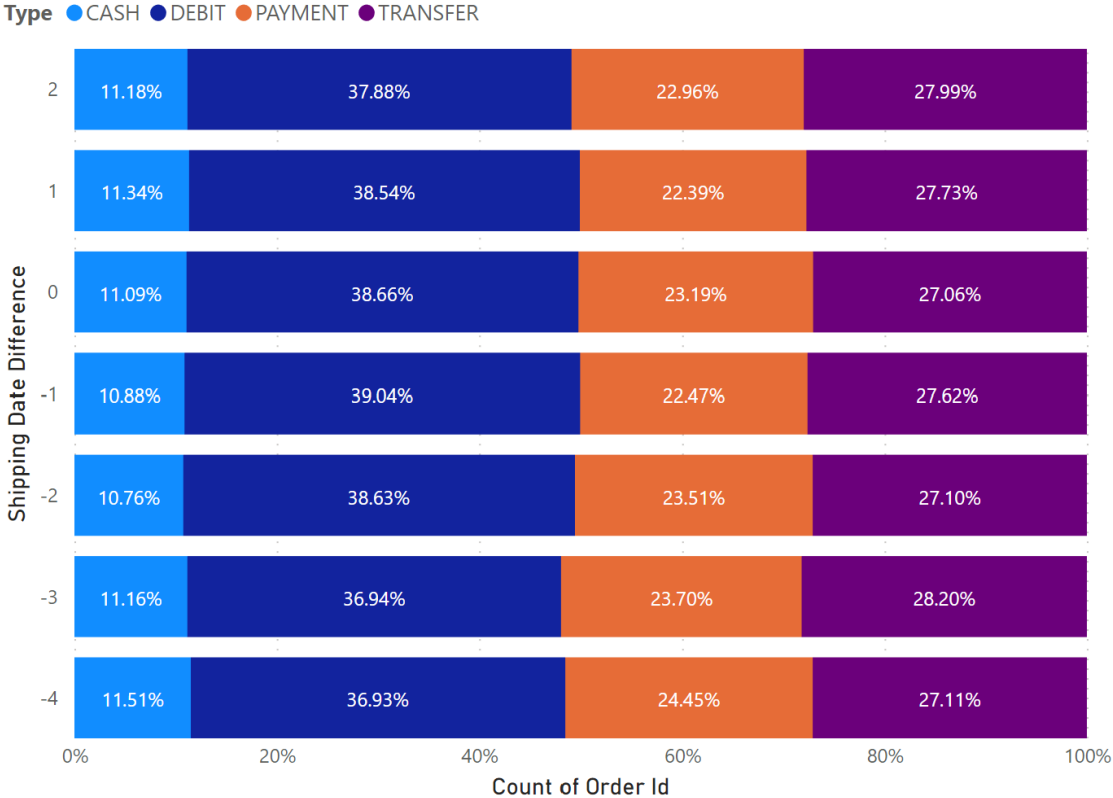
- COMPLETE
- PENDING_PAYMENT
- PROCESSING
- PENDING
- CLOSED
- ON_HOLD
- SUSPECTED_FRAUD
- CANCELED
- PAYMENT_REVIEW



Count of Order State by Type



Trạng thái giao hàng

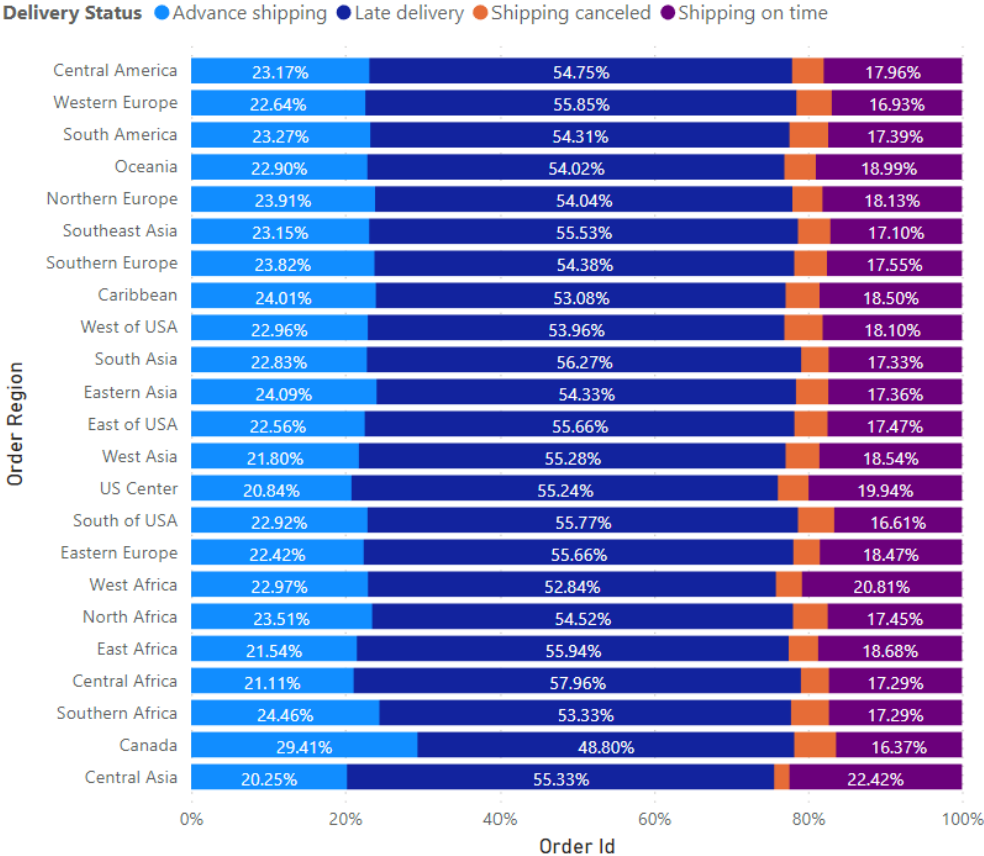


Delivery Status	Count of Order Id	%GT Order Id
Advance shipping	15127	23.04%
Late delivery	36048	54.83%
Shipping canceled	2855	4.30%
Shipping on time	11722	17.84%
Total	65752	100.00%

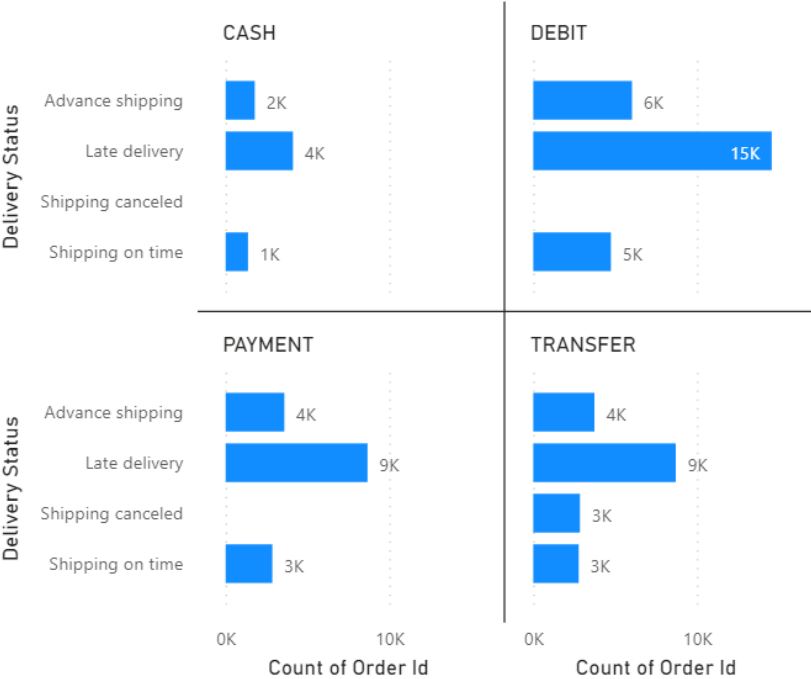
Giao hàng trễ chiếm đến 54.83% tổng đơn

Trạng thái giao hàng

Order Id by Order Region and Delivery Status

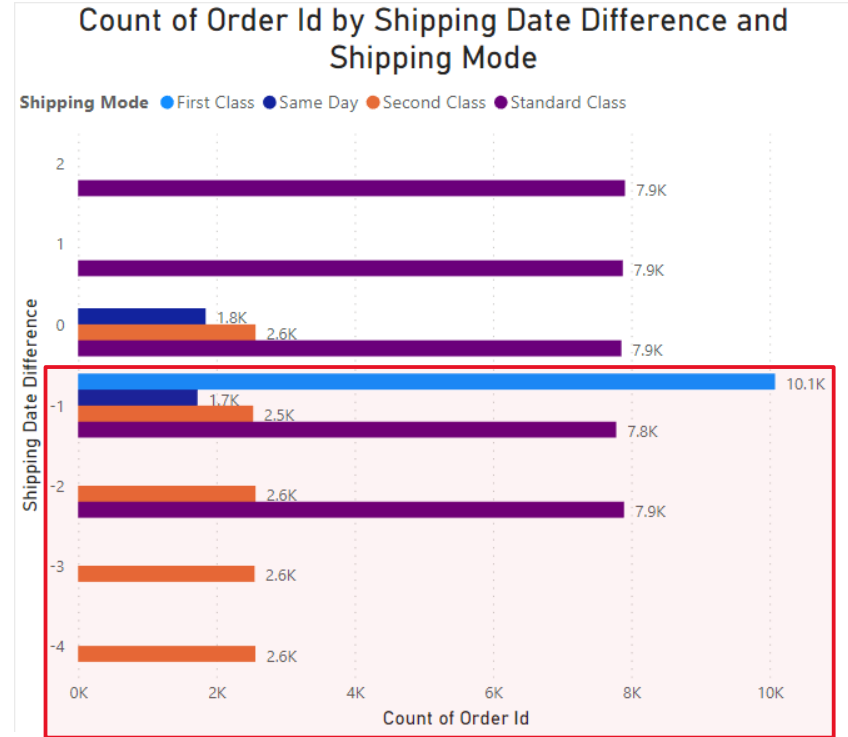
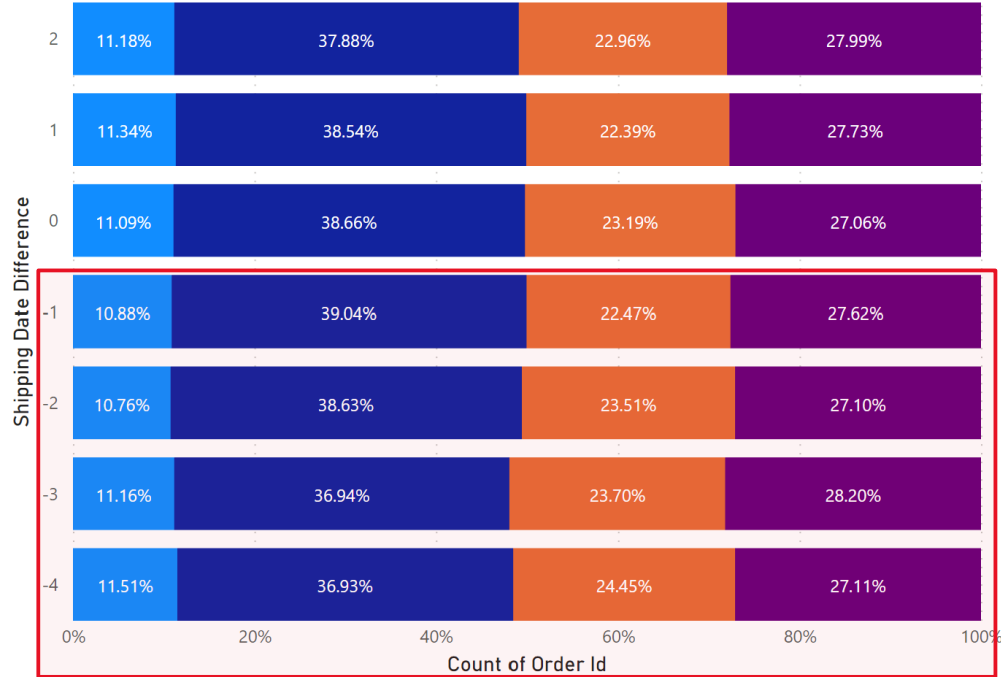


Count of Order Id by Delivery Status and Type



Ngày giao hàng chênh lệch

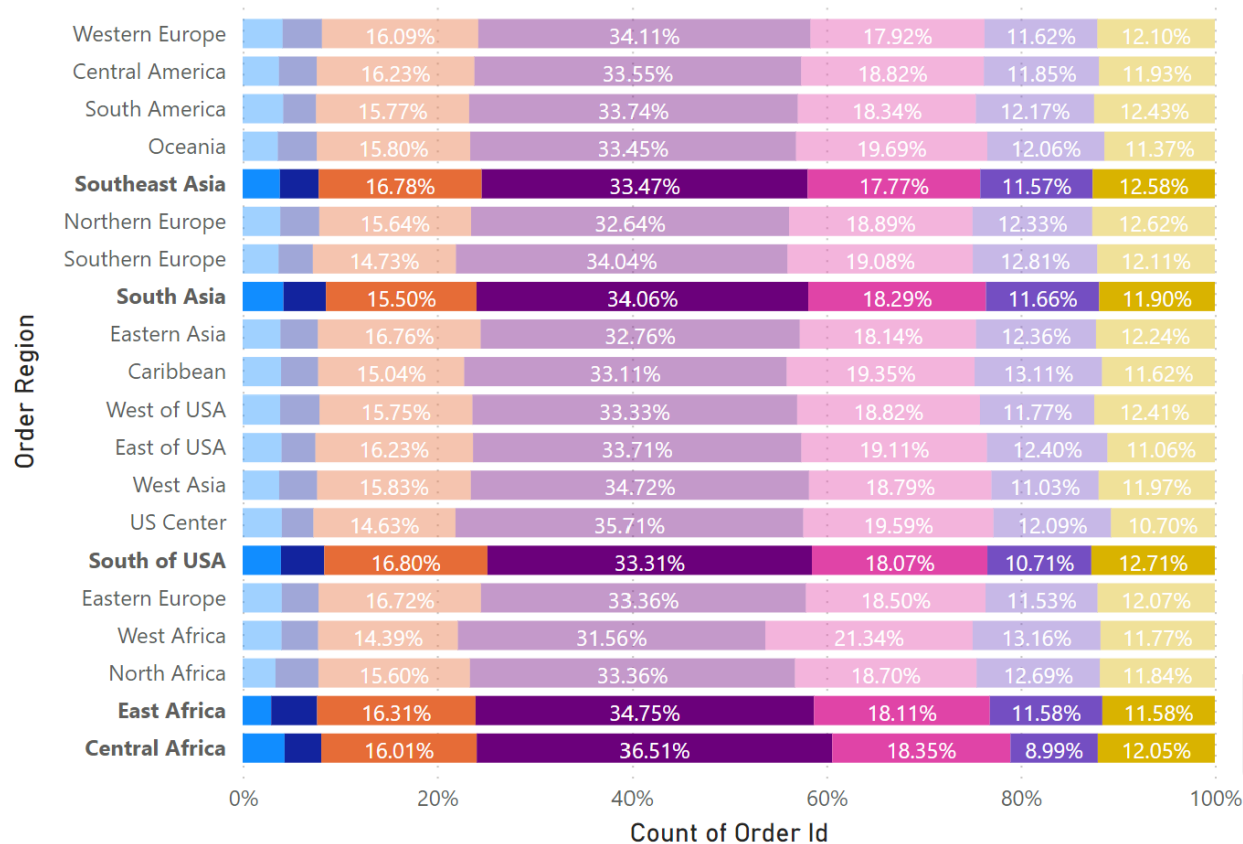
Type ● CASH ● DEBIT ● PAYMENT ● TRANSFER



Shipping date difference 1,2: Giao sớm từ 1,2 ngày 0: Đúng hạn -1 đến -4: Giao trễ từ 1 đến 4 ngày

Ngày giao hàng chênh lệch theo khu vực

Shipping Date Difference ● -4 ● -3 ● -2 ● -1 ● 0 ● 1 ● 2



3. Đề xuất cải thiện

- Phân tích nguyên nhân của phương thức giao hàng First Class và Same Date Class gây ra việc giao hàng trễ.
- Marketing các sản phẩm khác của công ty nhằm tăng lợi nhuận cũng như tăng lượng đơn hàng.
- Cải thiện khâu xử lý đơn hàng và mở rộng thêm các đơn vị vận chuyển nhằm mục đích tăng tỷ lệ completed (order status) ít nhất trên 50% vì hiện tại completed chỉ chiếm 32.97%.