

VIETNAM NATIONAL UNIVERSITY - HO CHI MINH CITY

UNIVERSITY OF ECONOMICS AND LAW

FACULTY OF INFORMATION SYSTEMS

BUSINESS INTELLIGENCE SEASON 8

TRÍ TUỆ KINH DOANH

REPORT BI8

PROJECT EXECUTION TEAM: BEST THU

HO CHI MINH CITY - 2024

Câu 1: EDA

a. Summary Statistics, Data Types and Missing Values

The data contains 53 columns with 180519 data rows in which 29 columns have numeric data.

Data Types:

Column Name	Data type
Type	String
Days for shipping (real)	Integer
Days for shipment (scheduled)	Integer
Benefit per order	Float
Sales per customer	Integer
Delivery Status	String
Late_delivery_risk	Integer
Category Id	Integer
Category Name	String
Customer City	String
Customer Country	String
Customer Email	String
Customer Fname	String
Customer Id	Integer

Customer Lname	String
Customer Password	String
Customer Segment	String
Customer State	String
Customer Street	String
Customer Zipcode	Integer
Department Id	Integer
Department Name	String
Latitude	Float
Longitude	Float
Market	String
Order City	String
Order Country	String
Order Customer Id	Integer
order date (DateOrders)	Datetime
Order Id	Integer
Order Item Cardprod Id	Integer
Order Item Discount	Float

Order Item Discount Rate	Float
Order Item Id	Integer
Order Item Product Price	Float
Order Item Profit Ratio	Float
Order Item Quantity	Integer
Sales	Float
Order Item Total	Float
Order Profit Per Order	Float
Order Region	String
Order State	String
Order Status	String
Order Zipcode	Integer
Product Card Id	Integer
Product Category Id	Integer
Product Description	String
Product Image	String
Product Name	String
Product Price	Float

Product Status	Integer
Shipping date (DateOrders)	Datetime
Shipping Mode	String

Missing value:

- Customer Email and Customer Password columns have been masked because of user privacy
- Order Zipcode have the most missing data with 155.679 missing rows
- Product Description has no data as the entire column is empty
- Customer Zipcode have 3 missing rows
- Customer Lname have 8 missing rows

As these above data is non-important to the requirements, the team will not process the missing value and choose to ignore them instead.

b. Data Distribution

Numerical features

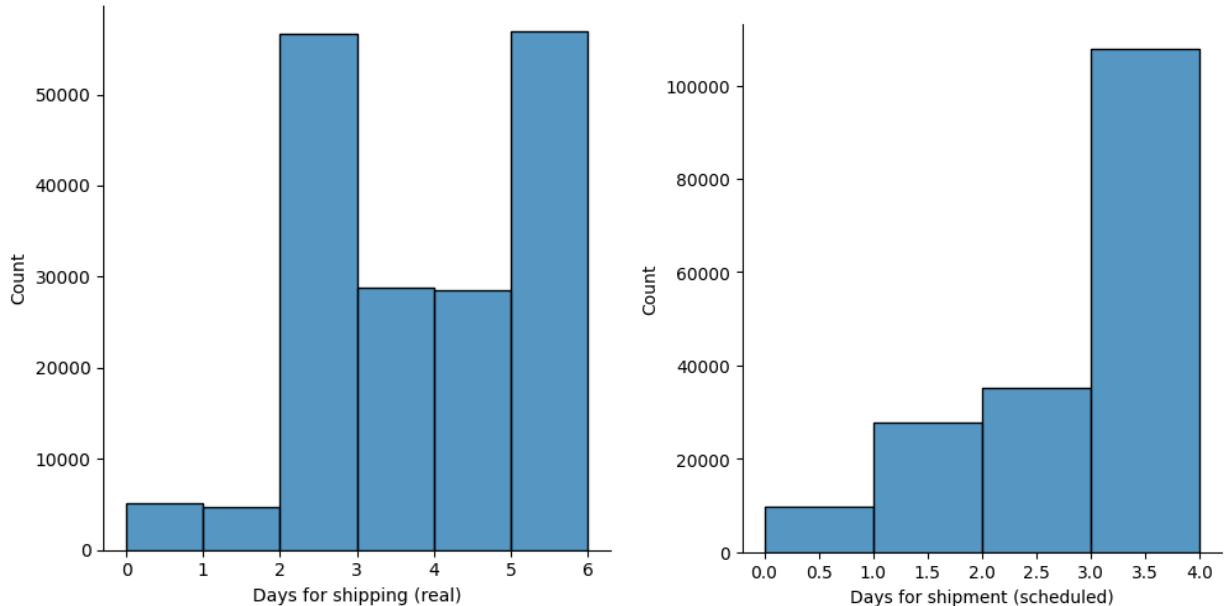


Figure 1.1: Distribution between scheduled and real day for shipping

These figures show that while most shipments are scheduled for 4 days, most of the shipments usually take 2 and 6 days while few take under 2 days to be delivered.

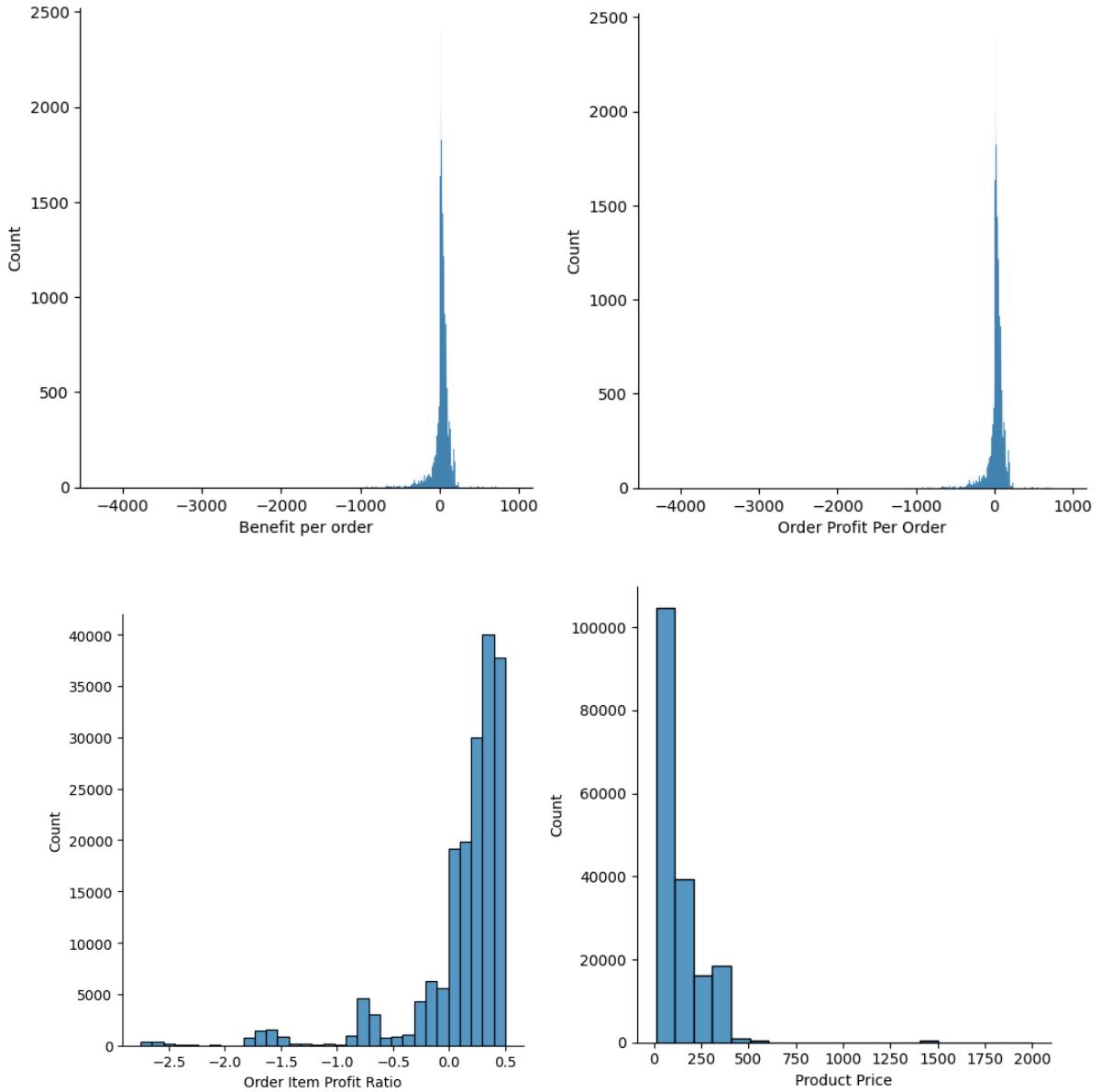


Figure 1.2: Distribution of benefit per order, profit per order, product price and profit ratio

The data shows that the mean profit per order is about 21.9 currency, with more than 25% of the product achieving more than 7 currency units of profit per sale with the average product price is 141.2 currency unit. In addition, the item profit ratio average is about 12% with a small amount of items losing money from selling products.

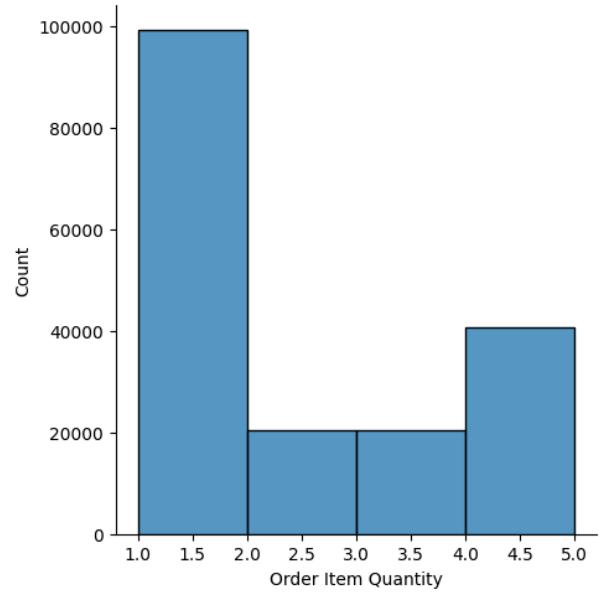


Figure 1.3: Distribution of order item quantity

More than 50% of consumers order 1 item when making purchases with consumers ordering 4 items when purchasing products.

Non-numerical features

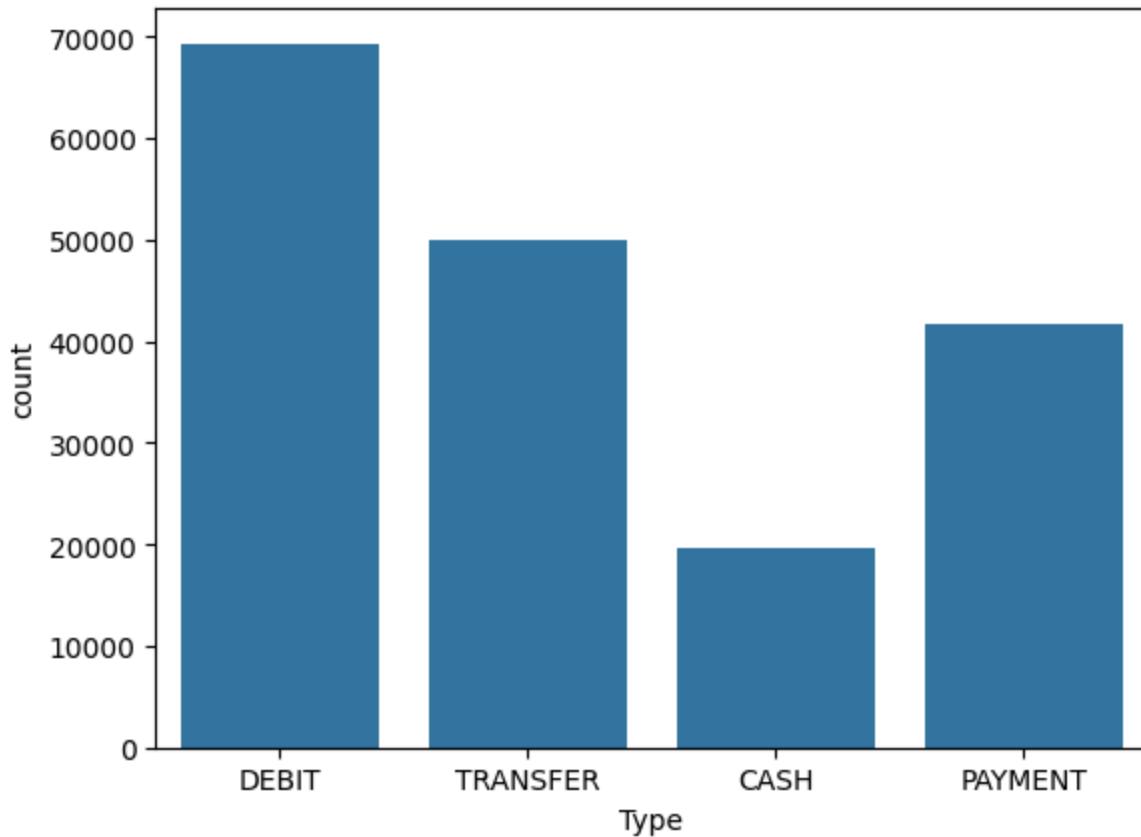


Figure 1.4: Numbers of transaction types

Most consumers prefer using debit cards and transferring when making purchases while cash transactions are the least popular method.

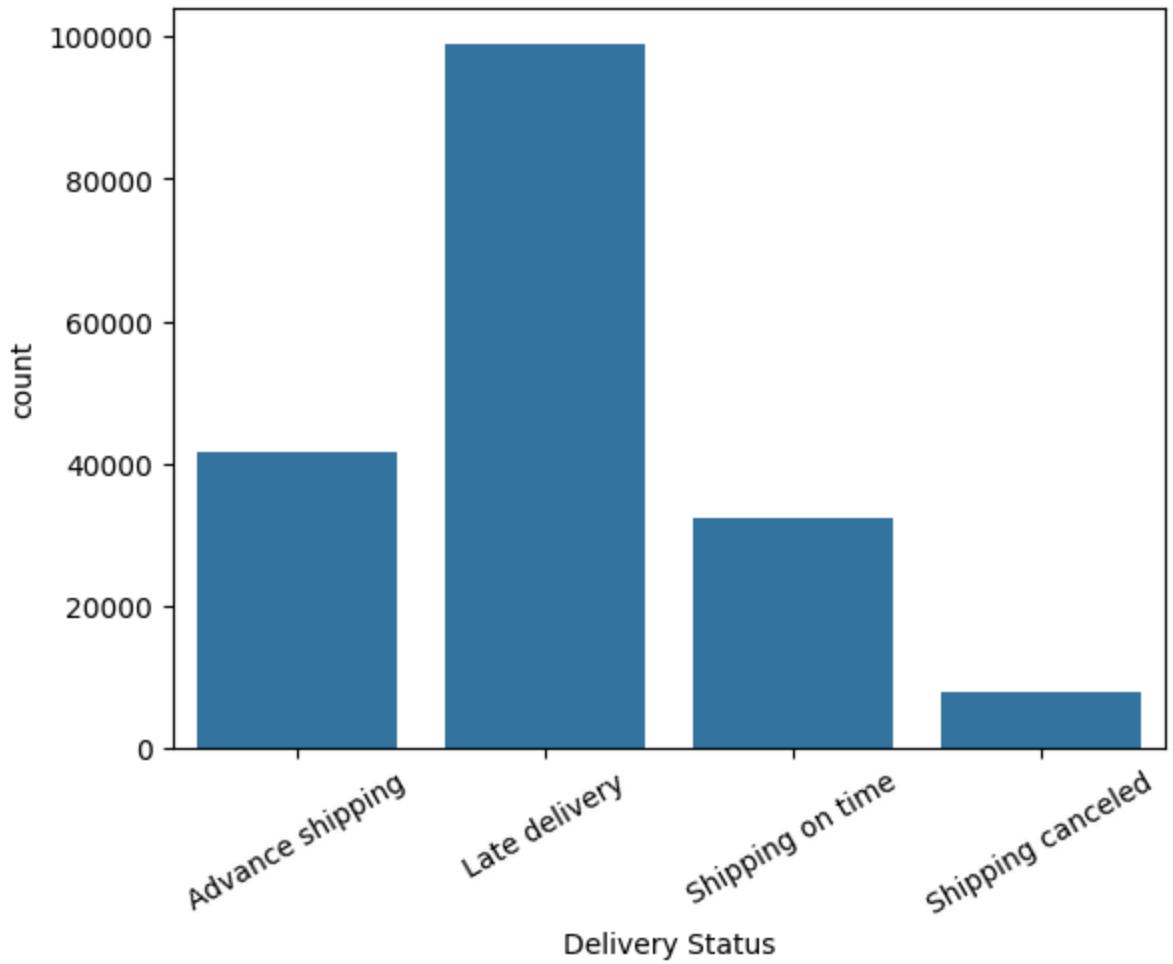


Figure 1.5: Type of delivery status

Compared to figure 1.1 about a real shipping date, this plot shows that most deliveries are late, as shown by figure 1.1 where orders are shipped after 4 days. Whereas some products are shipped early and on time with little being cancelled.

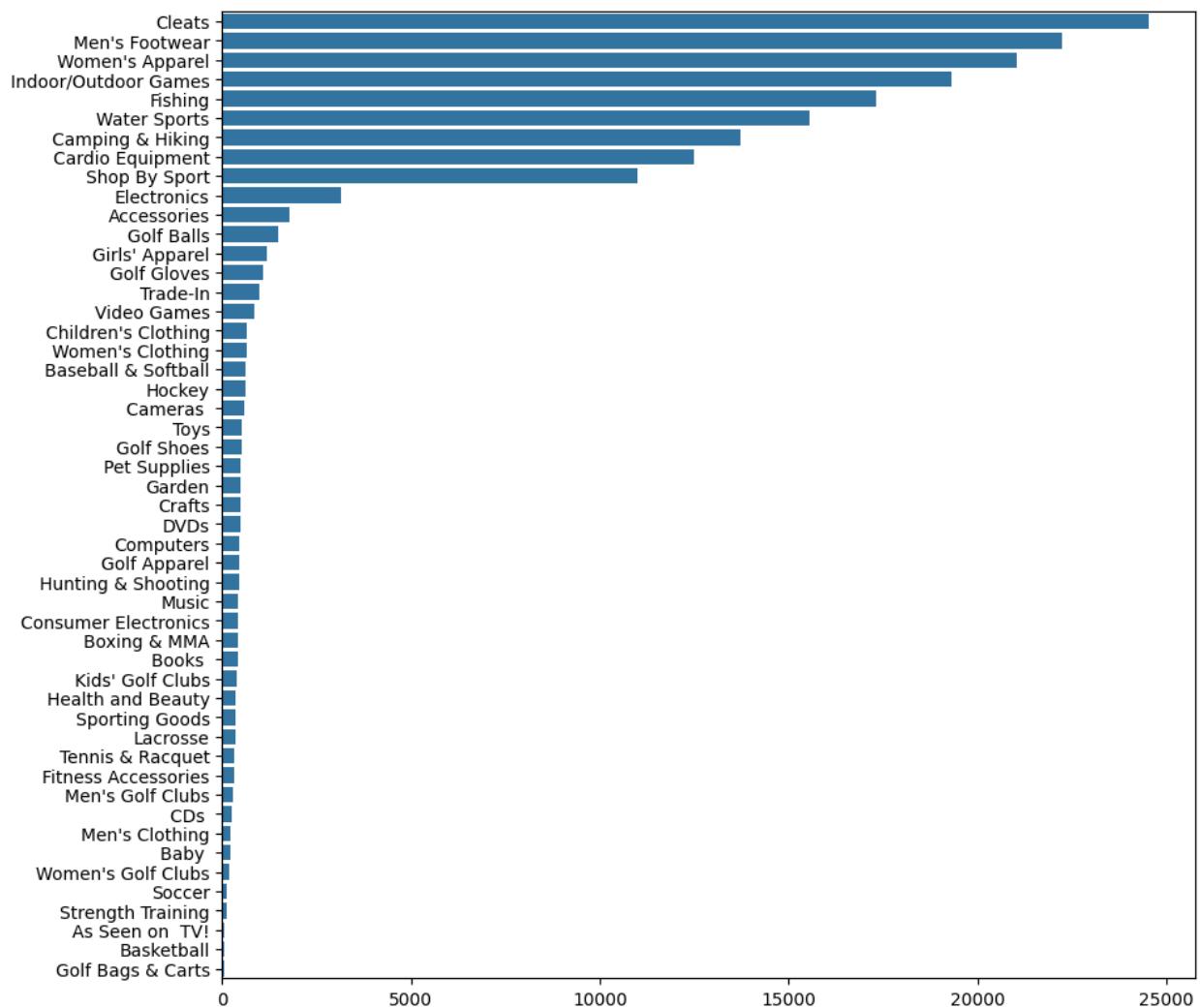


Figure 1.6: Type of product category

The data show that Cleats category is the most popular

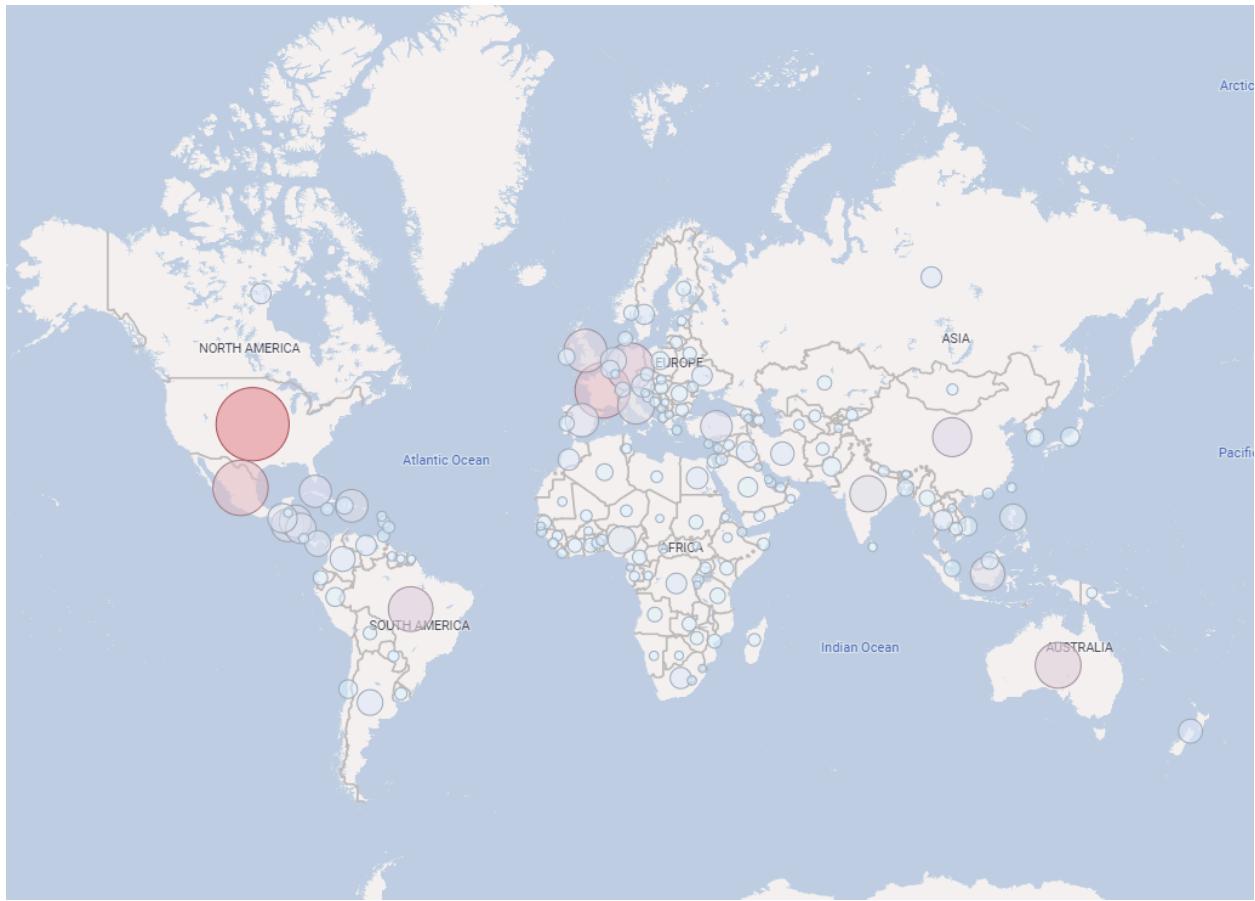


Figure 1.8: World Map of Order Destination

The map shows that most orders are shipped to the US, with Mexico, France and Australia taking the next 3 ranking. In addition, this shows us that products are shipped to most of the world with few exceptions such as Greenland, North Korea,...

c. Correlation Analysis

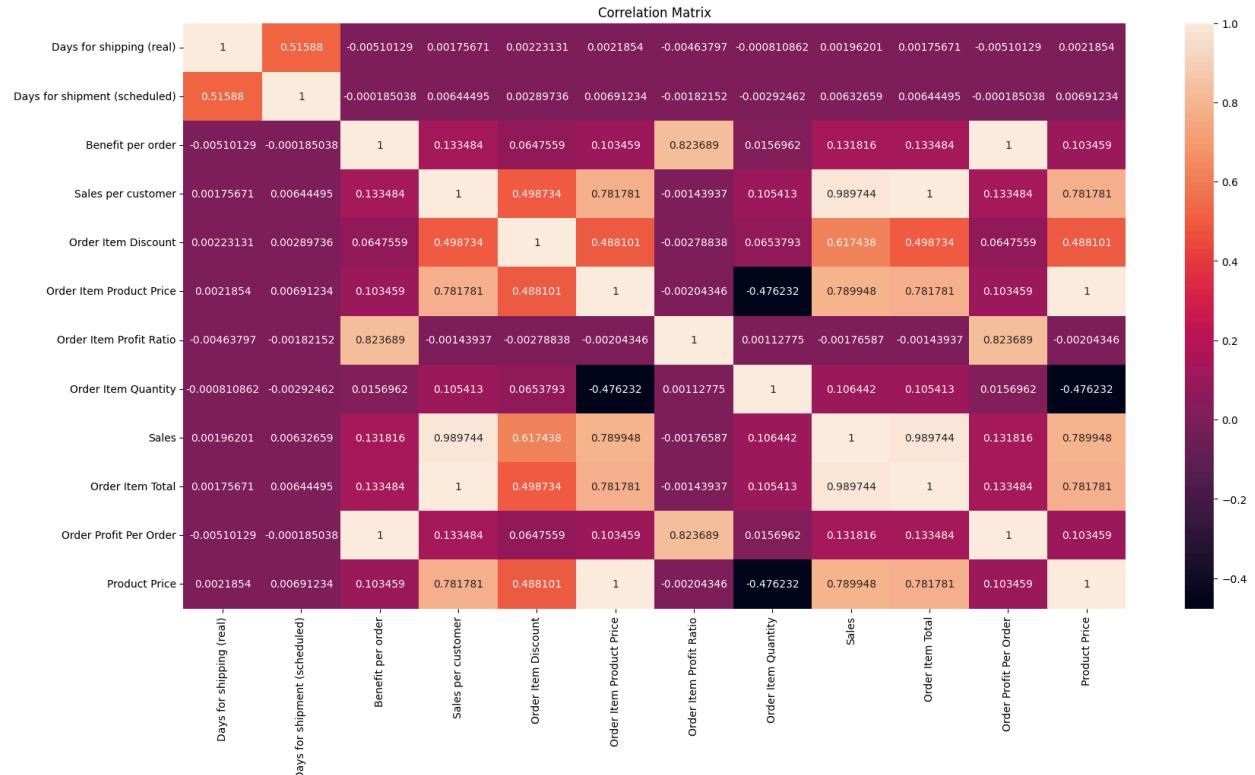


Figure 1.9: Correlation matrix of numeric columns

From the correlation matrix, we can see that there are some factors that are related to each other such as:

- “Sales per customer” and “Sales” show that the more each customer makes purchases, the more the sales raise.
- “Day for shipment (scheduled)” and “Day for shipping (real)” have some correlation but are not very high, showing that some of the shipments did not match the scheduled shipment date.
- “Order Item Discount” and “Sales per customer” show that product discounts will gather more sales from customers as the cheaper price will appeal to more customers.
- “Order Item Product Price” and “Sales per customer” show that the higher the price of items, the higher the sales number will go up, reflecting the income of the business.
- “Order Item Quantity” and “Order Item Product Price” have a negative correlation score, meaning that the more expensive the product, the lower the amount of the product that the customer will want to order.

e. Relationships Between Variables

During the analysis of the data, the team realized that some of the variables have the same data, they can also be seen in the correlation matrix (figure 1.9). These variables are:

- “Order Profit Per Order” and “Benefit per order”
- “Order Item Total” and “Sales per customer”
- “Order Item Product Price” and “Product Price”

After comparing and based on the Data Description provided by the organizers. The team came to the conclusion that these data variables are the same but came from different sources when making this data set. In this report, the team will use these data interchangeably.

Câu 2:

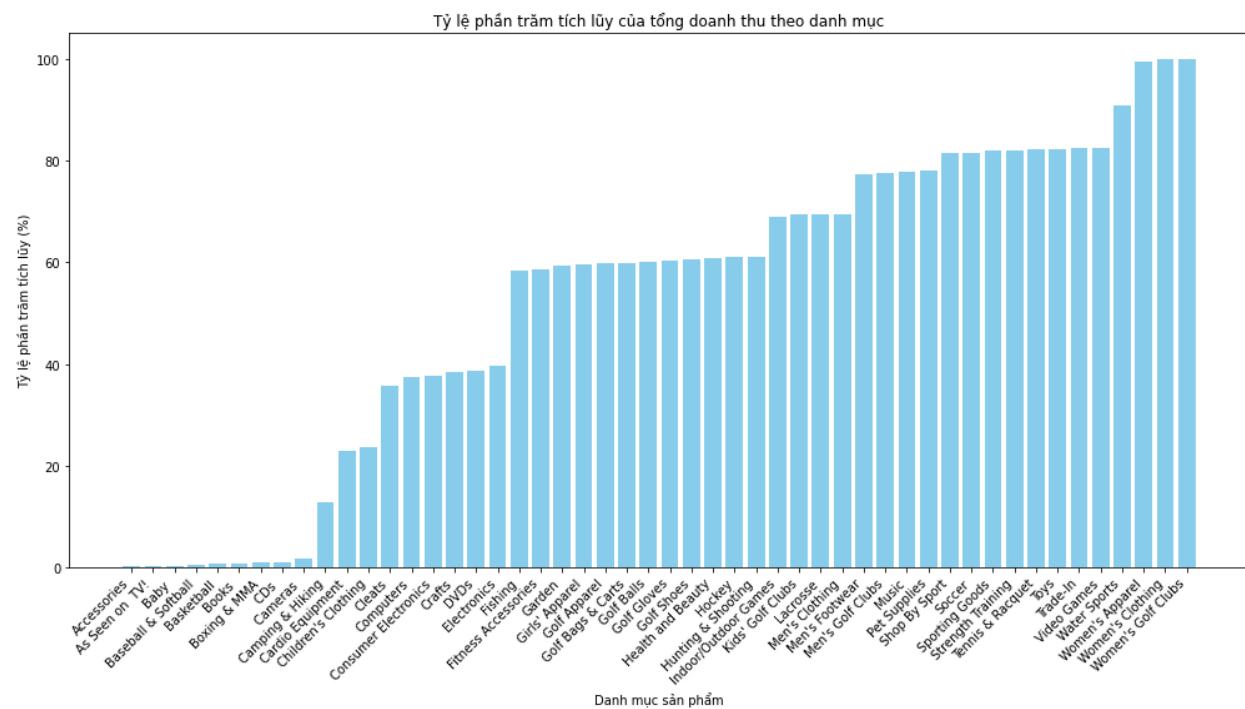


Figure 2.1. A column chart displays the cumulative percentage ratio of total revenue by product category

A. INTRODUCE

- The percentage ratio accumulates the contribution of each product category to total revenue.

- **Percentage Ratio (%)**: Each product category has a percentage of revenue contribution to total revenue.
 - **Accumulation**: The cumulative sum of percentage ratios of product categories in sequence, showing the total revenue from those categories.
 - **X-Axis**: Category Name. Each column represents a specific product category.
 - **Y-Axis**: The value represents the cumulative percentage of revenue from product categories.
 - Cumulative Percentage:
 - First Column: Percentage of revenue from the first product category in the list, has the lowest percentage value.
 - Next Columns: Represent the cumulative percentage of revenue from product categories, adding up the values from previous columns.
 - Purpose
 - Identify which product category contributes the most to total revenue.
 - Analyze the distribution of revenue among product categories.
- B. ANALYZE**
- **Product Category**: Top 3 categories with the highest percentage of total revenue: ***Fishing, Cleats, Camping and Hiking*** ⇒ Identify the high potential customer: Women and women's products.

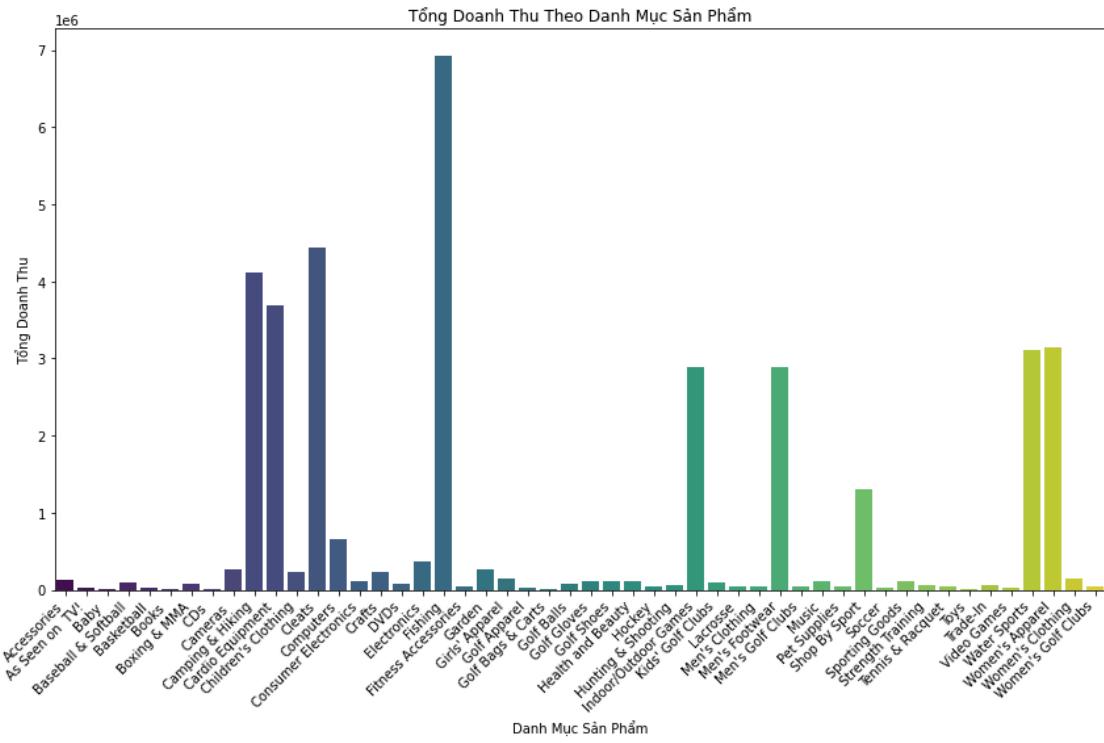


Figure 2.2. The column chart compares the revenue contribution of each product category to the total revenue.

- **X-Axis:** Product Category
- **Y-Axis:** Total Revenue from Each Product Category. The values on the Y-axis indicate the extent to which each product category contributes to the total revenue
- **Product Category Contributing Most to TR:** The product category with the tallest column. In this case, the "Fishing" column is the tallest ⇒ the "Fishing" category has the highest revenue among all product categories.
- **Product Category Contributing Least to TR:** The product category with the shortest column.
- **Trend:** Some product categories have significantly higher revenue compared to others. These are the best-selling or highest-value products.

Câu 3:

Compare FDE of Category appear in both phase in 2017

(First Six Month and Last Six Month)

Category Name	FDE First 6 Month	FDE Last 6 Month
As Seen on TV!	02.63	2.5
Baseball & Softball	-27.4	0.6
Boxing & MMA	-99.9	2.3
Camping & Hiking	0	0.3
Cardio Equipment	-3.1	2.7
Cleats	0	2.2
Fishing	0	0.25
Fitness Accessories	8.93	0
Golf Apparel	-12.1	0.8
Golf Bags & Carts	0.6	0.6
Golf Gloves	-4.87	2.2
Golf Shoes	0	2.54
Hockey	13.3	0
Hunting & Shooting	-5.2	2.4
Indoor/Outdoor Games	0	2.1
Kids' Golf Clubs	2.7	2.4

Lacrosse	6.7	2.58
Men's Footwear	0	0.77
Men's Golf Clubs	2.5	2.6
Shop By Sport	-5.8	2.1
Soccer	3.1	2.3
Strength Training	0.5	0.5
Trade-In	0.4	0.4
Video Games	-5.6	2.5
Water Sports	0	0.5
Women's Apparel	0	2.14
Women's Golf Clubs	2.4	2.5

Elastic vs. Inelastic Products:

- + Elastic Demand ($PED > 1$): Consumers are sensitive to price changes.
- + Inelastic Demand ($PED < 1$): Consumers are less sensitive to price changes.

Identify Products with Significant Changes

To analyze the Price Elasticity of Demand (PED) for each product and provide strategic pricing recommendations, we'll compare the FDE (which I assume stands for "Final Demand Elasticity" and is equivalent to PED in this context) values between the first and last six months of 2017. I'll focus on significant changes and their implications for pricing strategy.

1. Significant changes in demand responsiveness:

- Baseball & Softball:

- + First 6 months: -27.4
- + Last 6 months: 0.6

-> This shows a dramatic shift from highly elastic (price-sensitive) to relatively inelastic demand.

- Boxing & MMA:

- + First 6 months: -99.9
- + Last 6 months: 2.3

-> Another significant shift from extremely elastic to relatively inelastic demand.

- Fitness Accessories:

- + First 6 months: 8.93
- + Last 6 months: 0

-> Changed from highly inelastic to perfectly inelastic demand.

- Hockey:

- + First 6 months: 13.3
- + Last 6 months: 0

-> Similar to Fitness Accessories, shifted from highly inelastic to perfectly inelastic.

- Hunting & Shooting:

- + First 6 months: -5.2
- + Last 6 months: 2.4

-> Changed from elastic to relatively inelastic demand.

2. Strategic pricing recommendations:

- Baseball & Softball:

- + First half: Keep prices low to stimulate demand.

- + Second half: Gradually increase prices as demand becomes less sensitive.

- Boxing & MMA:

- + First half: Maintain very low prices to attract customers.
- + Second half: Implement moderate price increases.

- Fitness Accessories:

- + First half: Consider premium pricing strategy.
- + Second half: Focus on non-price factors (e.g., quality, marketing) as demand is not price-sensitive.

- Hockey:

- + Similar to Fitness Accessories, shift from premium pricing to emphasizing non-price factors.

- Hunting & Shooting:

- + First half: Keep prices competitive.
- + Second half: Implement moderate price increases.

General recommendations:

- + Seasonal pricing: For categories with significant changes between periods, implement dynamic pricing strategies that align with demand elasticity changes.
- + Bundle pricing: For categories with low elasticity in the second half (e.g., Golf Clubs, Soccer), consider bundling with complementary products to increase overall revenue.
- + Promotions timing: Schedule major promotions during periods of high elasticity to maximize impact on demand.
- + Cross-selling: For categories with low elasticity, focus on cross-selling and upselling strategies rather than price adjustments.

- + Inventory management: Align inventory levels with expected demand changes based on elasticity patterns.
- + Customer segmentation: Develop targeted pricing strategies for different customer segments based on their price sensitivity.
- + Monitor competitors: Keep track of competitors' pricing strategies, especially for categories with high elasticity.
- + Data-driven decision-making: Continuously analyze elasticity data to refine pricing strategies throughout the year.

By implementing these strategies, the company can optimize revenue by aligning pricing with changing demand responsiveness across different product categories and time periods.

Câu 4:

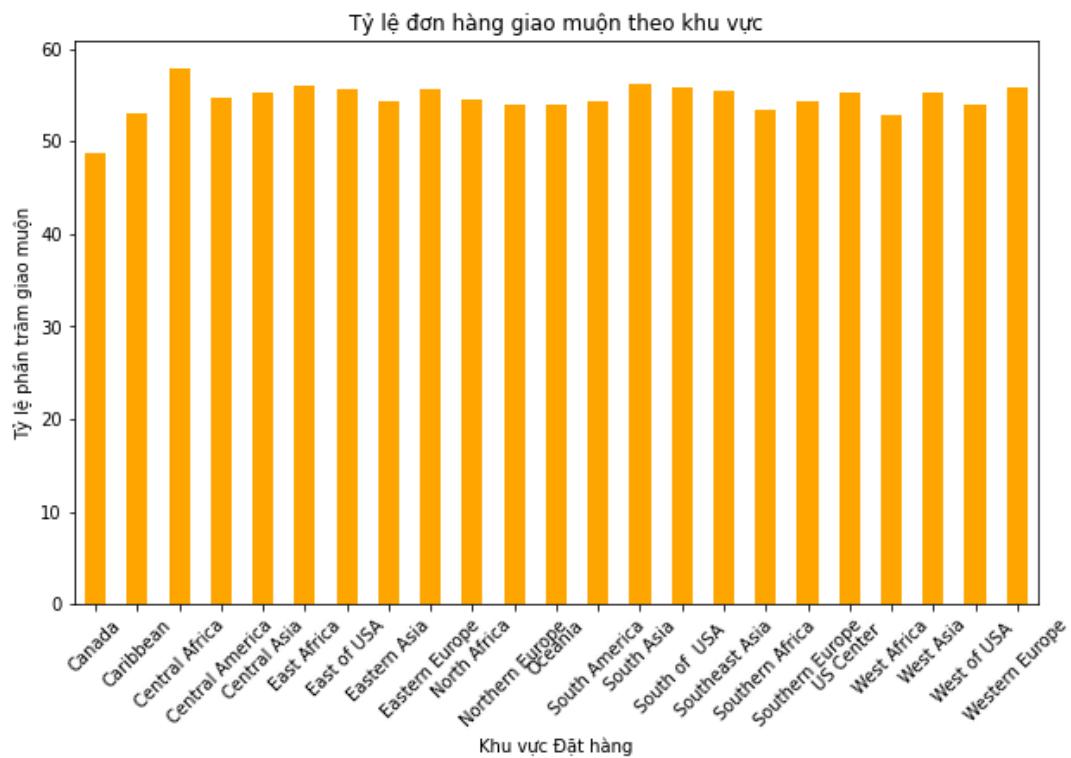


Figure 4.1. The columns chart shows the percentage of orders delivered late for each order region

- This chart compares the percentage of late deliveries across different regions.

- The Central Africa region has the highest rate of late deliveries: reflecting delays in the shipping process and worse delivery service in those areas.
- However, the difference in the rate of late deliveries among regions is not very far.

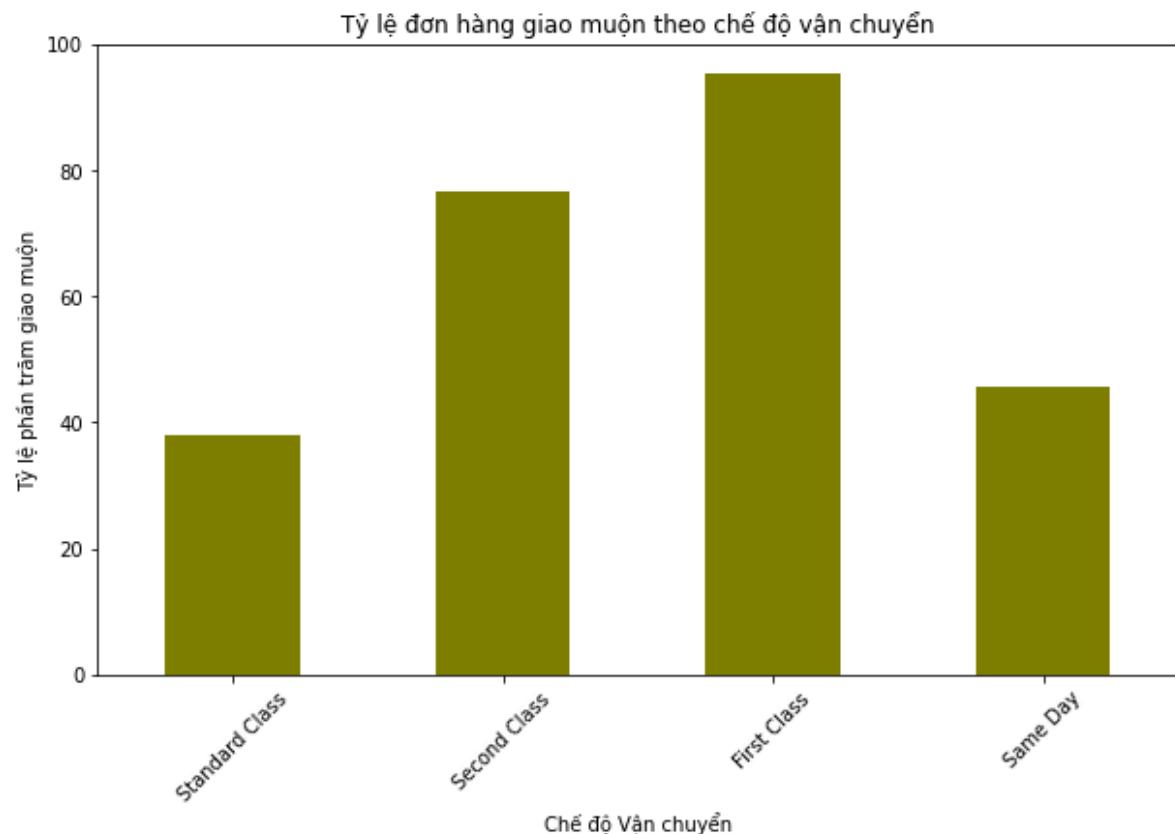


Figure 4.2. The columns chart shows the percentage of orders delivered late for each shipping mode

- This chart compares the percentage of late deliveries across different shipping modes: Standard.
- First Class is the shipping mode with the highest rate of late deliveries, next is Second Class. Same Day and Standard Class are the shipping modes with significantly lower rates of late deliveries.

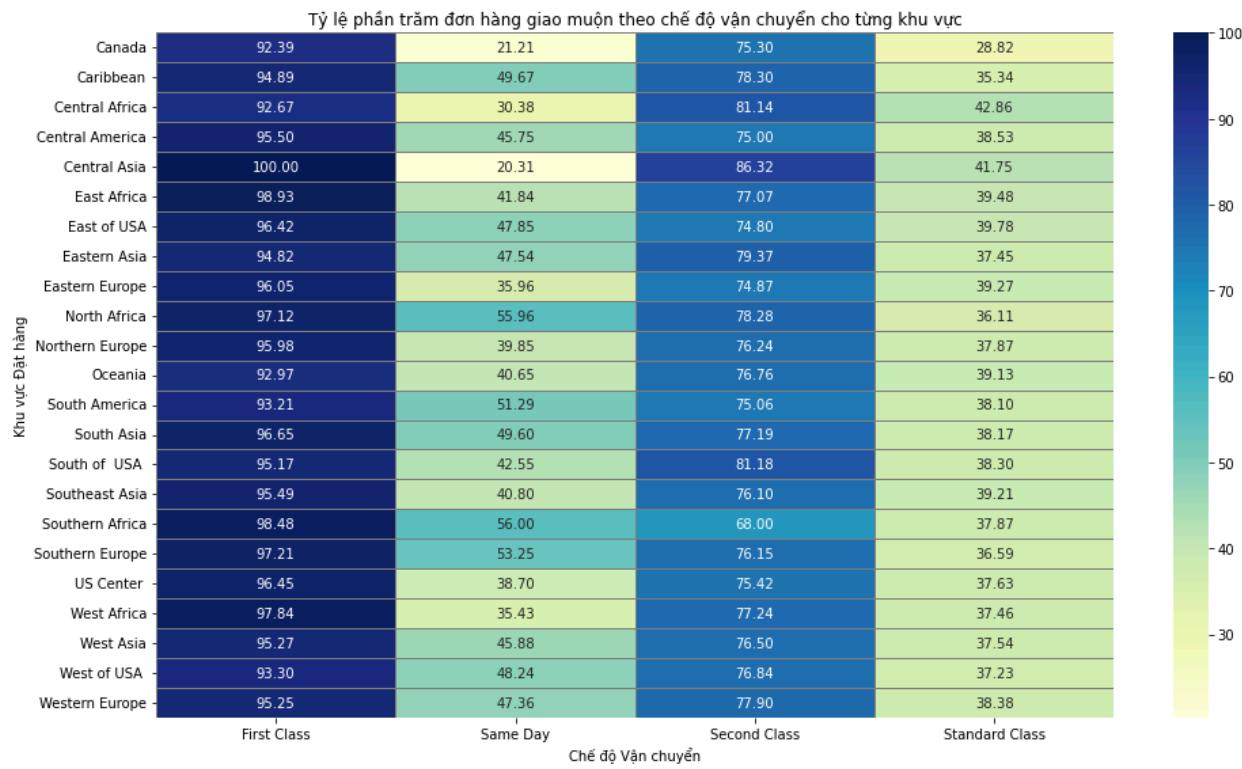


Figure 4.3. The combined chart shows the percentage of late deliveries based on both region and shipping mode

- The darkest-colored cells indicate the specific combinations of regions and shipping modes with the highest percentage of late deliveries.
- **First Class** has the highest rate of late deliveries in the regions of **Central Asia and East Africa**. Specifically, in **Central Asia**, the rate is at 100%.
- **Second Class** also shows a significant rate of late deliveries but is lower compared to First Class in most regions. The rate is particularly high in **Central Asia, Central Africa, and South of the USA**.
- **Standard Class** has the highest rate of late deliveries in various regions, especially in **Central Africa and Central Asia**.
- **Same Day** has the lowest rate of late deliveries, particularly in **Canada and Central Asia**, indicating a higher punctuality in deliveries compared to other methods.

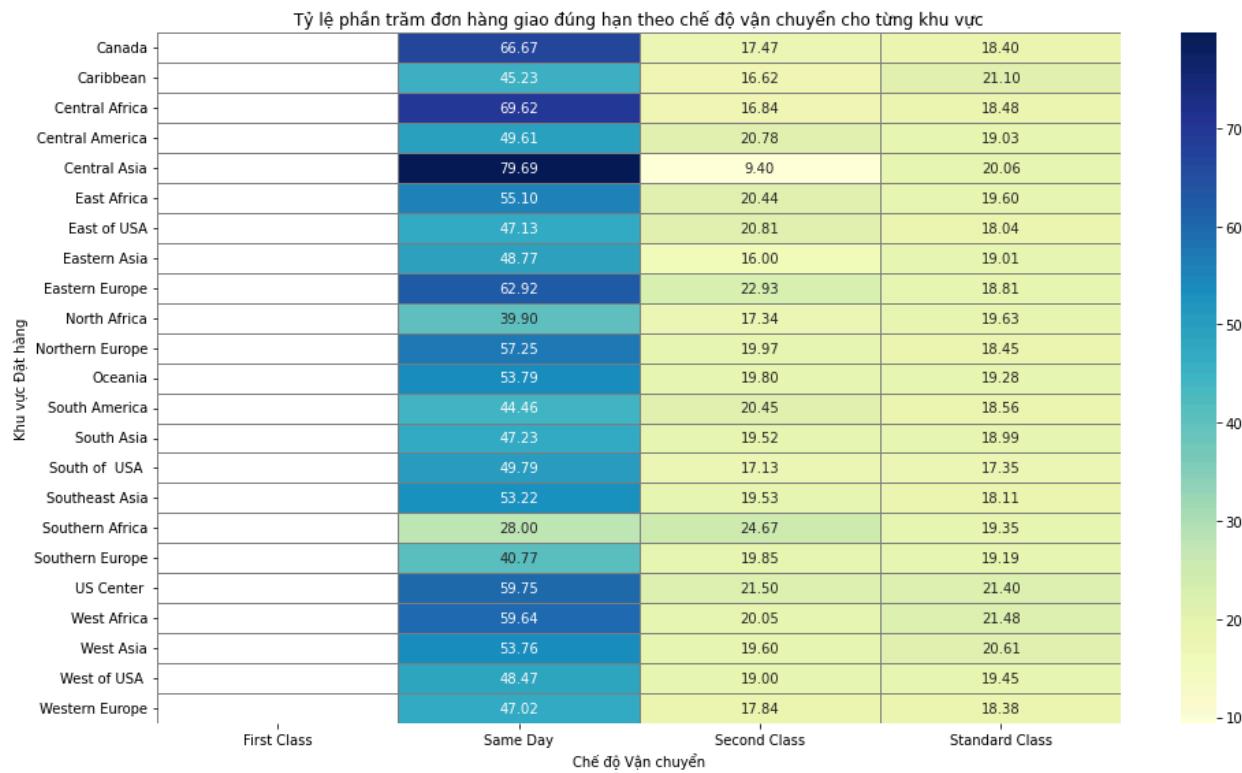


Figure 4.4. The combined chart shows the percentage of on-time deliveries based on both region and shipping mode

- The lightest colours indicate specific combinations of region and shipping mode with the highest on-time delivery rates.
- **First Class** has no on-time delivery rate, indicating a weakness in delivery service.
- **Same Day** has the highest on-time delivery rate, showing that this method is more reliable in ensuring timely deliveries.
- **Second Class** accounts for a significant number of on-time deliveries in **Southern Africa and Eastern Europe**.
- **Standard Class** accounts for a large number of on-time deliveries in **West Africa, US Center, and the Caribbean**.

Compare relationship between Figure 4.3 and 4.4 in some prominent regions

- **Central Africa:**

- **Late Delivery Rate:** The highest compared to other regions, indicating delays in the shipping process and poor delivery service in this area.
- **On-time delivery rate:** The on-time delivery rate is only average. This indicates that despite the high rate of late deliveries, some orders are still delivered on time, but not enough to significantly mitigate the impact of the late delivery rate.
- **Shipping methods:** Shipping methods like First Class and Standard Class have high rates of late deliveries in this region. Same Day is a better option but still not sufficient to significantly reduce the issue of late deliveries.
- **Central Asia:**
 - **On-time delivery rate:** Highest among all regions with Same Day shipping, indicating that this method can improve reliability in this region.
 - **Late delivery rate:** This region has a very high rate of late deliveries, especially with First Class and Second Class. Both methods have the highest late delivery rates across all regions, with First Class showing no on-time deliveries at all. This reflects significant delays and inefficiencies in the shipping process in this region.
 - **On-time delivery rate for First Class:** There are no recorded on-time deliveries for First Class, indicating a major weakness in this shipping method's service.
 - **Overall:** Same-day shipping has a relatively lower late delivery rate and the highest on-time delivery rate among all regions, indicating better performance in ensuring timely deliveries compared to other shipping modes.

Compare the relationship between Figure 4.3 and 4.4 in shipping modes

- **First Class:**
 - **On-time Delivery Rate:** No orders were delivered on time.
 - **Late Delivery Rate:** There is a high late delivery rate in many regions, particularly in Central Asia and East Africa. In Central Asia, the late delivery rate reaches 100%, indicating severe delays in delivery service.
- **Second Class:**

- **Late Delivery Rate:** Also a significant late delivery rate, but it is generally lower compared to First Class. The high rates in Central Asia, Central Africa, and the South of the USA indicate that the service still faces challenges, but not as important as First Class.
- **On-time Delivery Rate:** A large number of on-time deliveries in Southern Africa and Eastern Europe, indicating that Second Class performs better in ensuring timely deliveries compared to First Class, though it still needs improvement.
- **Same Day:**
- **Late Delivery Rate:** The late delivery rate is the lowest across all regions, especially in Canada and Central Asia. This indicates that Same Day has a higher performance in minimizing late deliveries.
- **On-time Delivery Rate:** Same Day is the most reliable method in ensuring on-time deliveries compared to other shipping modes.
- **Standard Class:**
- **Late Delivery Rate:** There is a high late delivery rate in many regions, especially in Central Africa and Central Asia. Although it is not as high as First Class, the rate is still notable.
- **On-time Delivery Rate:** There is a significant number of on-time deliveries in West Africa, the US Center, and the Caribbean. This indicates that Standard Class is better at ensuring timely deliveries compared to First Class and Second Class, but it is less effective than Same Day.

Câu 5:

To analyze the factors associated with a higher likelihood of delivery delays, we'll focus on the variables that seem most relevant to shipping and delivery status. The key field we're interested in is "Late_delivery_risk", which indicates whether a shipment is late (1) or not (0).

Key Factors: Shipping Mode, Order Status, Geographic Location, Product Categories, Seasonal Trends, Discrepancy Between Scheduled and Actual Shipping Days.

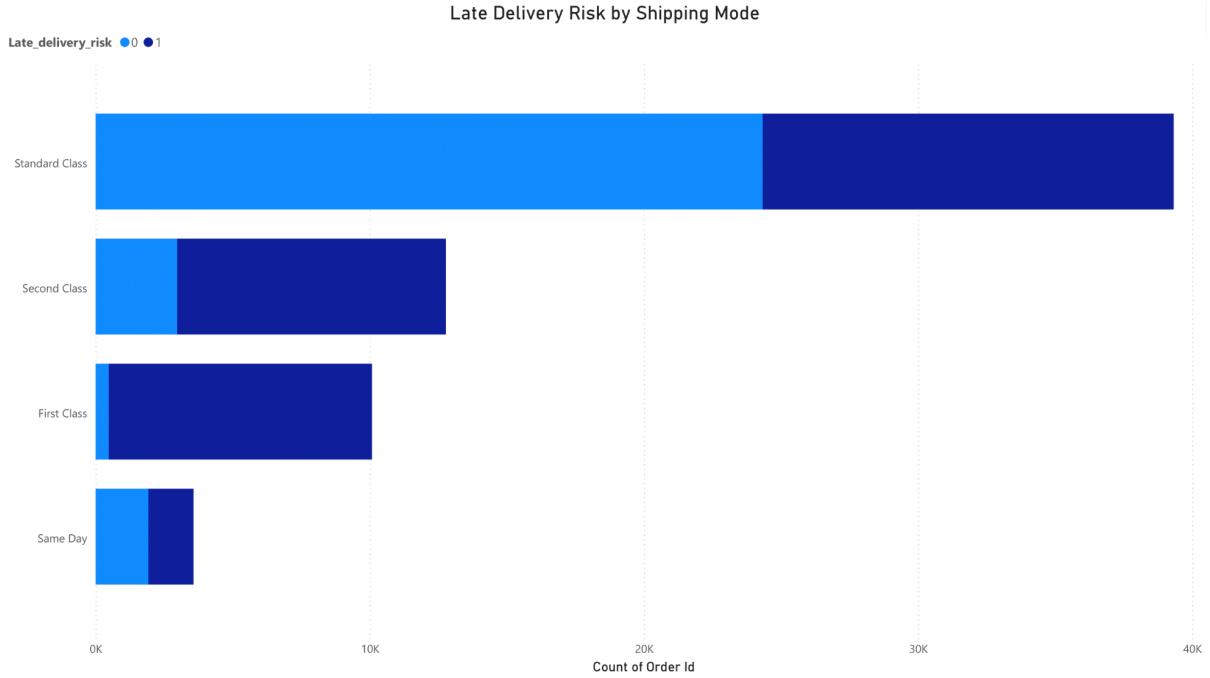


Figure 5.1: Correlation matrix of numeric columns

- Visualization: Stacked Bar Chart
- X-axis: Shipping Mode
- Y-axis: Count of Orders
- Color: Late_delivery_risk (0 or 1)

Standard Mode has the most Orders. Besides that, it has the most Orders which are sent late so it is true that the risk of shipping late is the highest rate.

Standard shipping is typically the default or most common shipping option for several reasons:

- Cost: It's usually the most economical option for customers.
- Balance: It offers a reasonable balance between cost and delivery time.
- Default option: Many e-commerce platforms set it as the default choice.
- Wide applicability: It's suitable for a broad range of products and situations.

Since Standard Mode has the most orders overall, it's likely to have the highest absolute number of late orders as well.

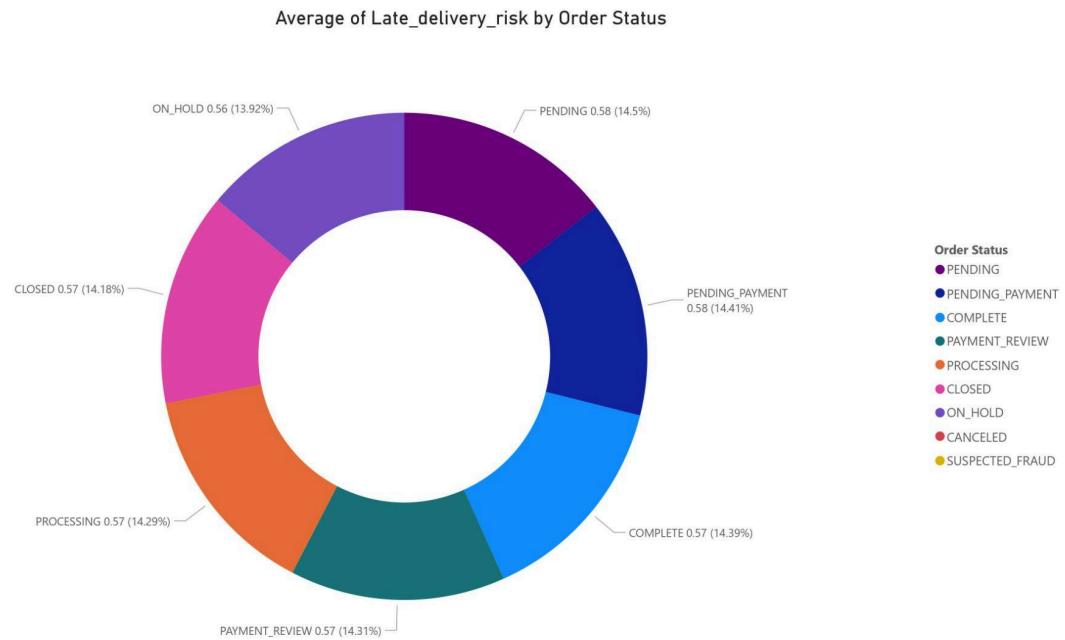


Figure 5.2: Late Delivery Risk by Order Status

- Visualization: Donut chart
- Legend: Order Status
- Values: Late_delivery_risk (Average)

While this might be more of an effect than a cause, orders with statuses like "**PENDING**" or "**PENDING_PAYMENT**" have the higher risk of delays than the other status

The reason for this problem are:

- Payment Issues:
- + Orders in "PENDING_PAYMENT" status are waiting for payment to be completed or verified.
- + This delay at the beginning of the order process can cascade into shipping delays.
- Inventory Problems:

- + "PENDING" status might indicate that the item is back-ordered or there are inventory issues.
- + If an item is not immediately available, it naturally leads to shipping delays.
- Order Verification:
 - + Some orders may be held in a pending status for additional verification (e.g., high-value orders, new customers).
 - + This extra step in processing can contribute to delays.
- System Issues:
 - + Technical problems in order processing systems can cause orders to become stuck in pending statuses.
 - + Resolving these issues takes time, leading to delays.
- Resource Allocation:
 - + Companies might prioritize the processing of confirmed orders over those in pending statuses.
 - + This can lead to pending orders accumulating delays.



Figure 5.3: Geographic Heat Map of Late Deliveries

- Visualization: Map
- Location: Customer Country or State
- Color intensity: Average of Late_delivery_risk

About the country: **The USA** is the country with higher rates of late deliveries.

About region: **North America, Western Europe and Southeast Asia** are the regions with higher rates of late deliveries.

1. The USA has higher rates of late deliveries:

Several factors could contribute to this:

- Volume of Orders:
 - + The USA is likely one of the largest markets, with a high volume of orders.
 - + More orders can mean more opportunities for delays.
- Geographic Size:
 - + The USA is a large country with diverse geography, which can complicate logistics.
 - + Long distances between distribution centres and customers can increase delivery times.

- Complex Supply Chain:
- + The US market might involve more complex supply chains with multiple touchpoints.
- + More complexity can lead to more opportunities for delays.

North America, Western Europe, and Southeast Asia have higher rates of late deliveries:

- High E-commerce Adoption:
- + These regions tend to have high e-commerce penetration, leading to more online orders.
- Urban Concentration:
- + High population density in urban areas can lead to traffic congestion and delivery complications.
- Complex Regulatory Environments:
- + Especially in Western Europe, varying regulations between countries can slow down cross-border shipments.
- Infrastructure Challenges:
- + In parts of Southeast Asia, infrastructure development might not keep pace with the growth in e-commerce.

Market	2015	2016	2017	2018	Total
■ Africa	0.55	0.54	0.55		
Central Africa	0.58	0.58	0.58		
East Africa	0.56	0.55	0.56		
North Africa	0.54	0.56	0.55		
Southern Africa	0.53	0.58	0.53		
West Africa	0.53	0.50	0.53		
■ Europe	0.55	0.56	0.55	0.55	
Eastern Europe	0.56	0.56	0.56		
Northern Europe	0.55	0.69	0.53	0.54	
Southern Europe	0.54	0.58	0.55	0.54	
Western Europe	0.56	0.44	0.55	0.56	
■ LATAM	0.54	0.54	0.54		
Caribbean	0.54	0.52	0.53		
Central America	0.55	0.55	0.55		
South America	0.54	0.54	0.54		
■ Pacific Asia	0.54	0.56	0.55	0.57	0.55
Central Asia	0.53	0.71	0.55		
Eastern Asia	0.53	0.55	0.53	0.58	0.54
Oceania	0.51	0.55	0.54	0.56	0.54
South Asia	0.57	0.56	0.55	0.56	0.56
Southeast Asia	0.55	0.56	0.55	0.57	0.56
West Asia	0.55	0.54			0.55
■ USCA	0.55	0.42	0.55		
Canada	0.50	0.42	0.49		
East of USA	0.56		0.56		
South of USA	0.56		0.56		
US Center	0.55		0.55		
West of USA	0.54		0.54		
Total	0.55	0.55	0.54	0.57	0.55



Figure 5.4: Market and Order Region

Visualization: Map

- Latitude field: Latitude.
- Longitude field: Longitude.
- Color saturation: Late Delivery Rate.

Visualization: Matrix

- Rows: Order Region and Market.
- Values: Late Delivery Rate.

The destination markets like **Central Africa, Eastern Europe, Central America, South - Southeast Asia, and East - South of America** could impact delivery times. These regions are prone to delays due to logistics challenges or distance.

Besides that, there are:

- Infrastructure Challenges:
 - + Limited infrastructure in Central Africa and parts of Southeast Asia
 - + Variable infrastructure quality in Central America and rural areas of East-South America
 - + Urban congestion in densely populated areas of South - Southeast Asia
- Geographical Factors:
 - + Large distances in USA and East-South America
 - + Island logistics in Southeast Asia
 - + Mountainous terrain in Central America
- Political and Regulatory Factors:
 - + Complex customs procedures in Central Africa
 - + Varying regulations between EU and non-EU countries in Eastern Europe
 - + Political instability in some Central African regions

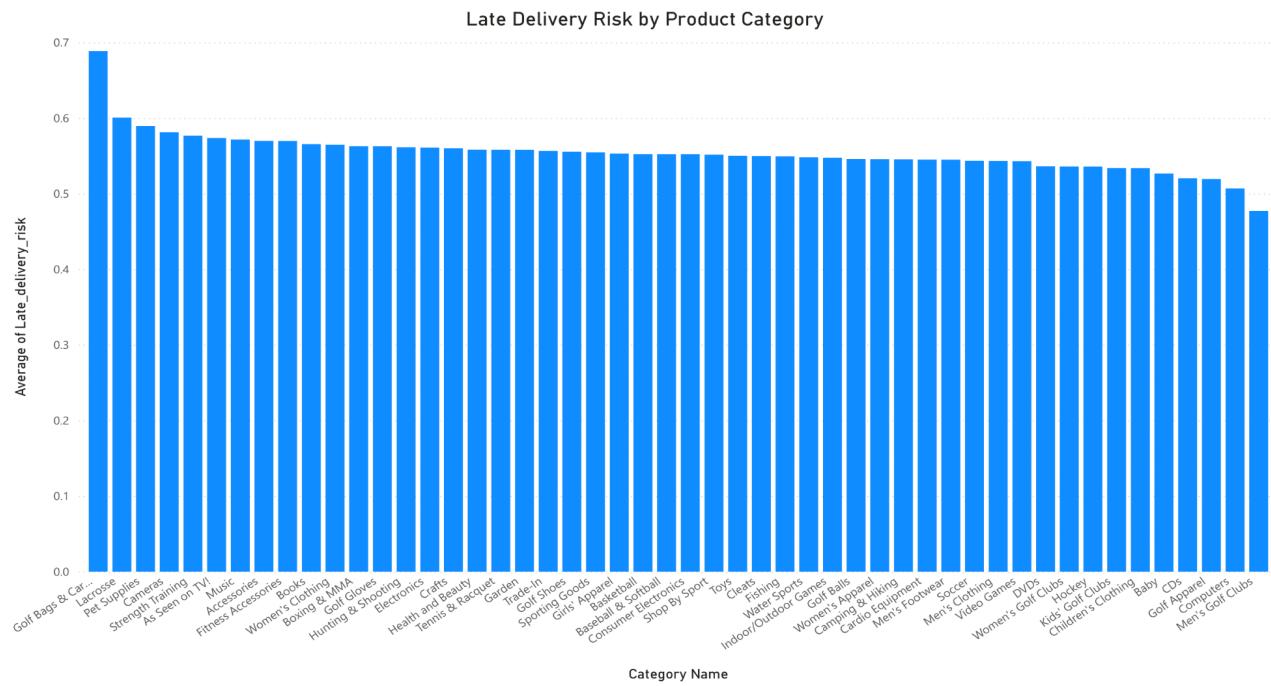


Figure 5.5: Late Delivery Risk by Product Category

- Visualization: Column Chart
- X-axis: Category Name
- Y-axis: Average of Late_delivery_risk

Product categories are the most prone to delivery delays: **Golf Bags and Carts**

Some categories like **Lacrosse, Pet Supplies, Cameras and Strength Training** also have a high rate of late delivery.

Here's the translation:

Product categories that are most prone to delivery delays, especially golf bags and carts, are affected for several reasons:

- + Volume and Size: Golf bags and carts are often large, making transportation more challenging.
- + Order Processing Time: These products may require longer processing times due to packaging and quality inspection procedures.
- + Supply Issues: Shortages of materials or products can lead to delays in availability.
- + Special Shipping Requirements: Some products may need special shipping methods, extending delivery times.

Additionally, categories like lacrosse, pet supplies, cameras, and strength training equipment also experience high rates of late delivery for similar reasons, such as complexities in the shipping process or unstable supply chains.

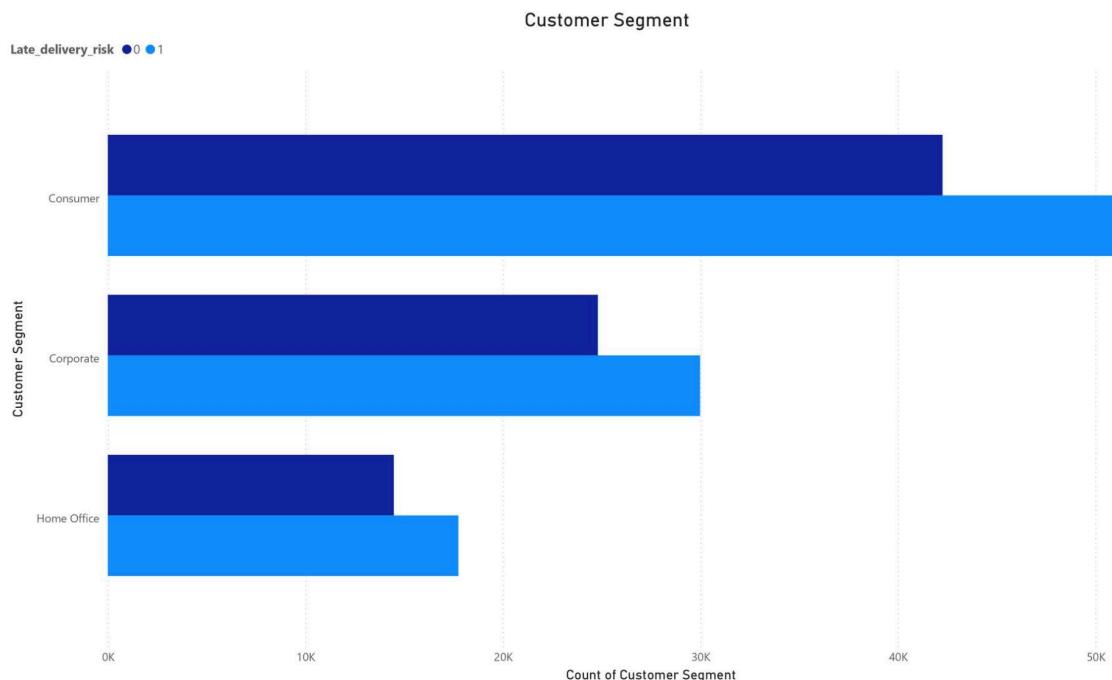


Figure 5.6: Late Delivery Risk by Customer Segment

- Visualization: Clustered Column chart
- Axis field: Customer Segment
- Value field: Late Delivery Rate

There are differences in delivery performance between Consumer, Corporate, and Home Office segments. **Consumer** is the customer segment experience which is the most delayed.

The differences in delivery performance between Consumer, Corporate, and Home Office segments can be attributed to several factors:

- + Order Volume: Consumer orders are typically smaller and more frequent, which can lead to inefficiencies in processing and shipping, resulting in delays.
- + Complexity of Orders: Corporate orders often involve larger quantities and may have specific requirements, allowing for more streamlined processing compared to the varied and unpredictable nature of consumer orders.
- + Shipping Priorities: Corporations may have established relationships with logistics providers, often leading to prioritized shipping options that consumers do not receive.
- + Delivery Locations: Consumer deliveries can be to residential areas that may be harder to access or have less predictable traffic patterns, contributing to delays.
- + Returns and Exchanges: The consumer segment often sees higher rates of returns and exchanges, complicating logistics and potentially slowing down delivery times.

Overall, the consumer segment experiences the most delays due to these factors, making it less efficient compared to corporate and home office segments.

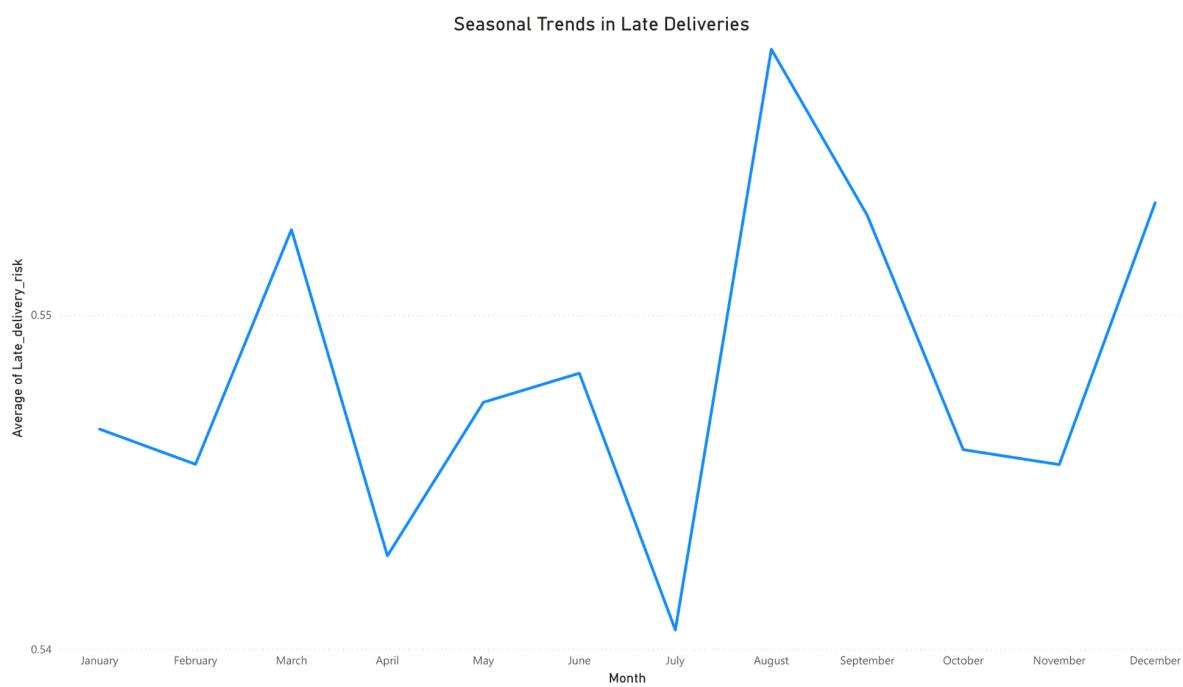


Figure 5.7: Seasonal Trends in Late Deliveries

- Visualization: Line Chart
- X-axis: Order Date (by month)
- Y-axis: Average of Late_delivery_risk

By analyzing the "order date" field, we could identify that the delivery is the most delayed in **March and from August to September.**

In **July**, it seems that Shipping is smooth and hassle-free.

The trends in late deliveries observed in March and from August to September can be attributed to several factors:

- March:
- + Seasonal Demand: March often sees increased demand due to events like spring break, leading to higher order volumes that can overwhelm logistics systems.

- + Weather Conditions: In many regions, March can still experience winter weather, which may disrupt transportation and lead to delays.
- + End of Quarter: Companies may rush to fulfil orders by the end of the fiscal quarter, creating congestion in shipping and handling processes.
- August to September:
 - + Back-to-School Season: The transition from summer to fall typically involves a surge in orders for school supplies, clothing, and electronics, which can strain delivery systems.
 - + Increased Shipping Volumes: Many retailers ramp up their inventory in preparation for the holiday season, leading to a spike in shipping demands.
 - + Weather Factors: Late summer can also bring storms and hurricanes in certain regions, impacting logistics and causing delays.
- July:
 - + The smooth shipping experience in July may be due to lower order volumes as it is typically a quieter month for retail, allowing logistics operations to function more efficiently without the pressure of seasonal spikes.

Scheduled vs Actual Shipping Days

Days for shipment (scheduled)	0	1	2	3	4	5	6	Total
0	5080	4657						9737
1			27814					27814
2			7138	7065	6978	7052	6983	35216
4			21666	21700	21535	21111	21740	107752
Total	5080	4657	56618	28765	28513	28163	28723	180519

Figure 5.8: Scheduled vs Actual Shipping Days

- Visualization: Matrix
- Rows: Days for shipment (scheduled).
- Columns: Days for shipping (real).

- Values: Count of orders.

A significant discrepancy between these two values could indicate systemic issues leading to delays.

This shows where most orders fall in terms of actual shipping days vs scheduled.

A significant discrepancy between actual shipping days and scheduled shipping days can indicate systemic issues for several reasons:

- + Inefficient Processes: If there is a consistent gap between expected and actual delivery times, it may suggest inefficiencies in the order processing, packing, or shipping stages.
- + Logistical Challenges: Discrepancies could point to problems in the supply chain, such as delays in transportation, inadequate carrier performance, or issues with inventory management.
- + High Demand Fluctuations: A sudden increase in order volume that is not matched by capacity can lead to delays, indicating that the system is not equipped to handle peak times effectively.
- + Inaccurate Scheduling: If the scheduled shipping days are not based on realistic assessments of shipping conditions, this could lead to a pattern of missed deadlines.
- + Communication Gaps: Poor communication between departments (e.g., sales, warehousing, and shipping) can lead to misunderstandings about order statuses and expected delivery times.
- + External Factors: Factors such as weather events, political situations, or global supply chain disruptions can contribute to delays that are not accounted for in scheduled shipping times.

By analyzing these discrepancies, businesses can identify and address underlying issues, improve their logistics processes, and enhance overall customer satisfaction. Understanding where most orders fall in terms of actual vs. scheduled shipping days allows companies to make data-driven decisions to optimize their operations.

Câu 6:

Challenges faced by ITBLogistic:

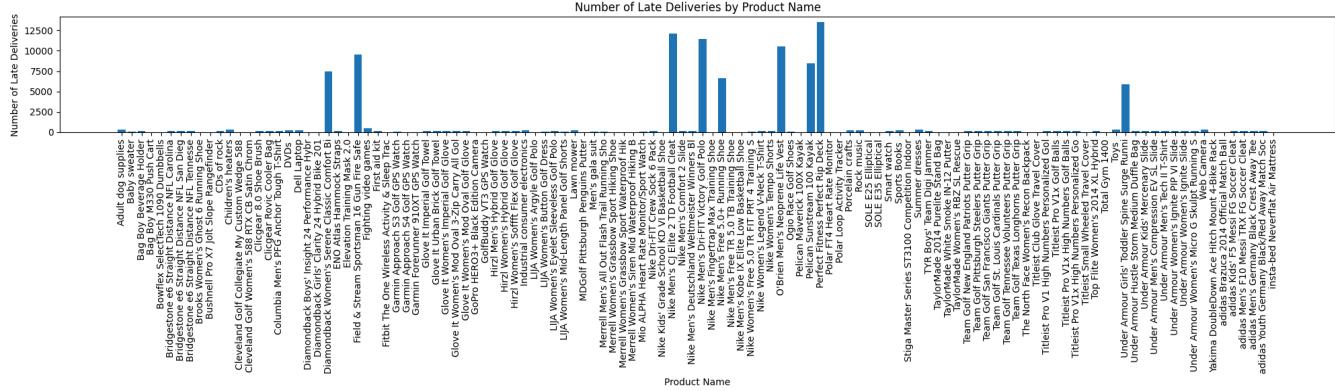


Figure 6.1: Number of Late Deliveries by Product Name

The graph above shows that some of the products have a higher frequency of having late deliveries than others. This means that ITBLogistic has problems in logistics by having a high rate of late deliveries as shown in question 4, some late deliveries might have been caused by problems with inventory management and supply chain where the seller or supplier does not have product available when the consumers purchase it.

- **Transportation Logistics: Delays and Delivery Times:** The delivery is the most delayed in **March** and **from August to September** as seen in Figure 5.7. This shows where most orders fall in terms of actual shipping days vs scheduled.
- **Supply Chain:** The Figure also pinpoints where the supply chain experiences the most disruptions. This data underscores the importance of enhancing supply chain visibility to mitigate late deliveries and improve overall efficiency.

Solution:

- Integrate a supply chain visibility platform that offers real-time tracking and analytics to monitor the movement of goods from suppliers to customers. Adopting JIT principles can help ITBLogistic reduce inventory holding costs by receiving goods only as they are

needed for production or customer orders. Partnering with reliable suppliers who can deliver goods quickly can ensure that inventory is replenished just in time to meet demand.

- Develop a collaborative network with key supply chain partners, utilizing technology to share data and improve coordination across the supply chain. Additionally, we use software that optimizes inventory levels by considering factors such as lead times, and demand variability SAP can help balance inventory across multiple locations and integrating a Transportation Management System can help automate routes and shipment tracking.
- Leveraging data analytics and machine learning to predict customer demand more accurately. For example, analyzing historical sales data, seasonal trends, and market conditions can help forecast future demand to solve the problem with the changing demand throughout the year. Data Analysis can also identify the most efficient routes for delivery trucks. For example, analyzing traffic patterns, weather conditions, and historical delivery data can help optimize routes.