

# Exploratory Data Analysis (EDA) Report: Supply Chain Dataset

## Dataset Overview

The dataset contains 1,999 records with 15 columns, covering key aspects of supply chain operations such as order information, delivery status, shipping methods, customer segments, product categories, and geographic dimensions.

	Column	Type	Nulls	Unique Values
0	Order_Id	int64	0	1999
1	Type	object	0	4
2	Order_Date(DateOrders)	object	0	24
3	Shipping_Date(DateOrders)	object	0	30
4	Days_for_Shipment(scheduled)	int64	0	4
5	Days_for_Shipping(real)	int64	0	7
6	Delivery_Status	object	0	4
7	Late_Delivery_Risk	int64	0	2
8	Shipping_Mode	object	0	4
9	Order_Item_Quantity	int64	0	1
10	Order_Item_Discount_Rate	float64	0	18
11	Product_Category_Id	int64	0	10
12	Order_Region	object	0	4
13	Customer_Country	object	0	2
14	Customer_Segment	object	0	3

Figure 1 Data Summary Table

## Insight 1: Delivery Status Distribution

### Analysis:

A significant portion of orders fall under the ‘Late delivery’ category, highlighting inefficiencies or delays in the supply chain.

### Recommendation:

Investigate causes for these delays: shipping modes, product types, or regional bottlenecks. Consider implementing monitoring tools to proactively flag high-risk orders.

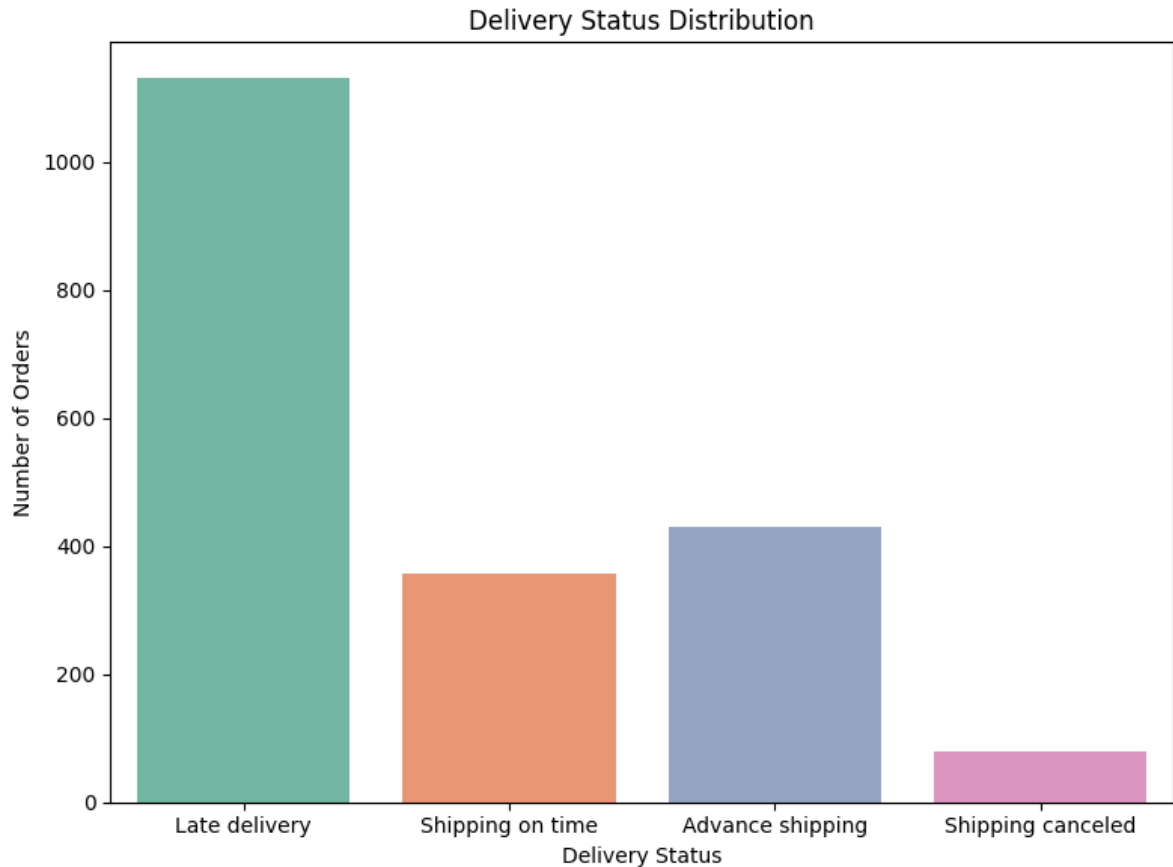


Figure 2 Delivery Status Distribution

## Insight 2: Order Type Breakdown

### Analysis:

The majority of orders are through **Payments and Expenses**, with **Cash transactions** being minimal. This aligns with a digital-first customer base.

### Recommendation:

Ensure order management policies support digital transaction preferences. Explore expansion of digital payment methods.

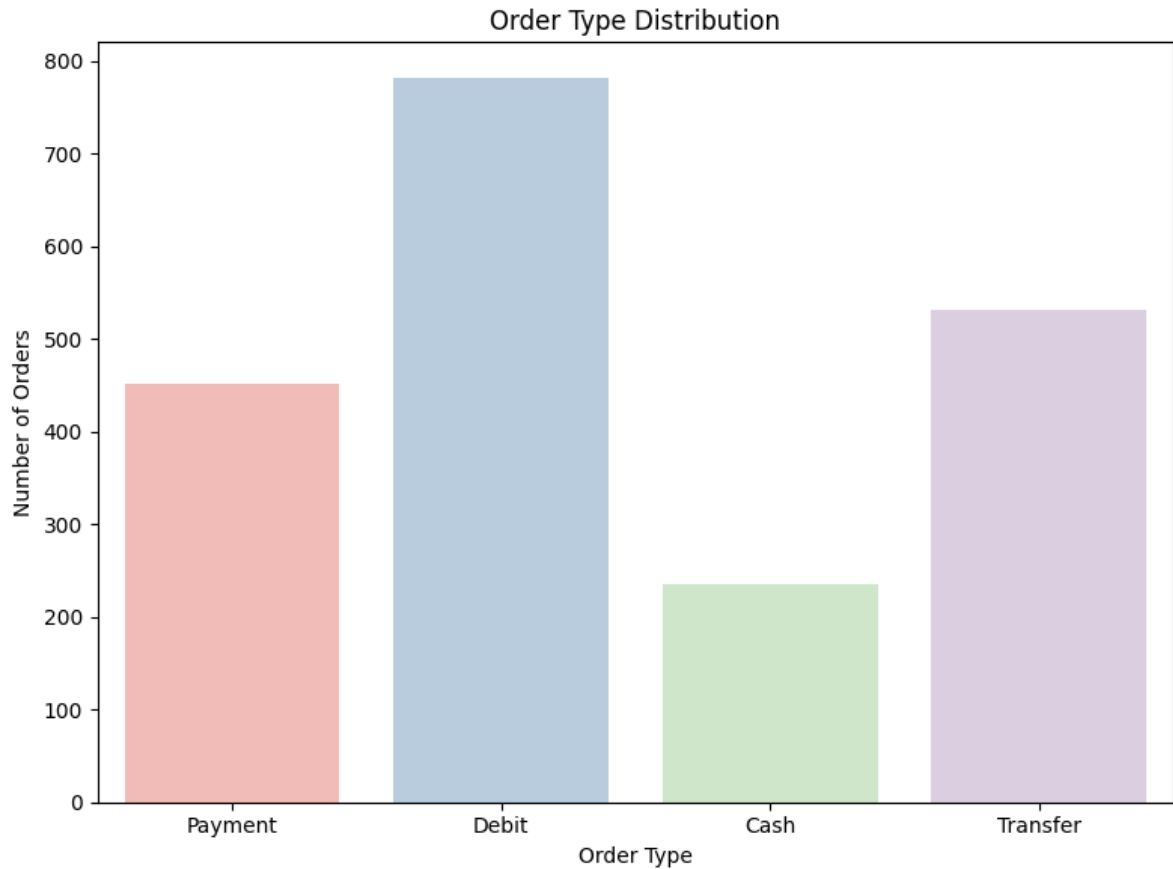


Figure 3 Order Type Distribution

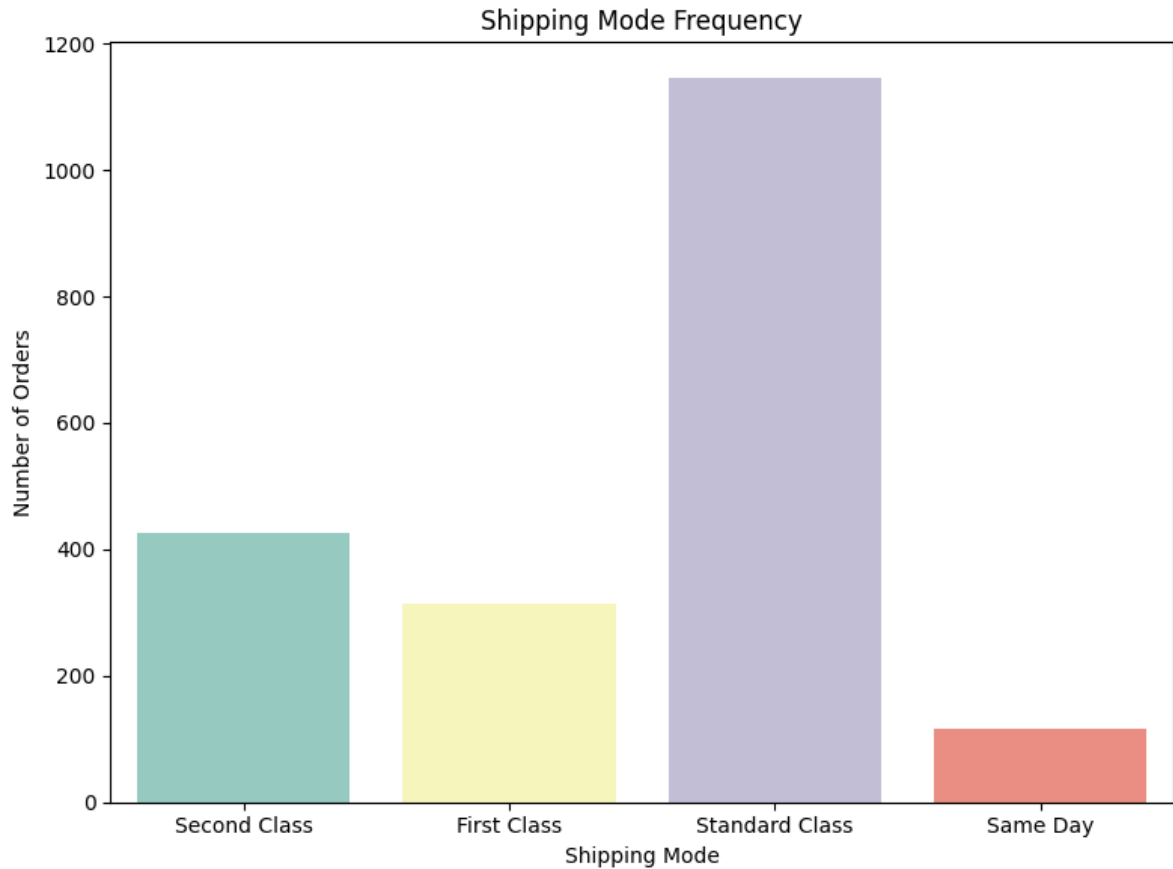
### Insight 3: Shipping Method Trends

#### Analysis:

**Standard Class** dominates, indicating a cost-focused delivery strategy. **1st and 2nd Class** methods are used less frequently, probably for urgent deliveries.

#### Recommendation:

Review how shipping modes affect delivery speed and customer satisfaction. Adjust shipping options where needed.



*Figure 4 Shipping Mode Frequency*

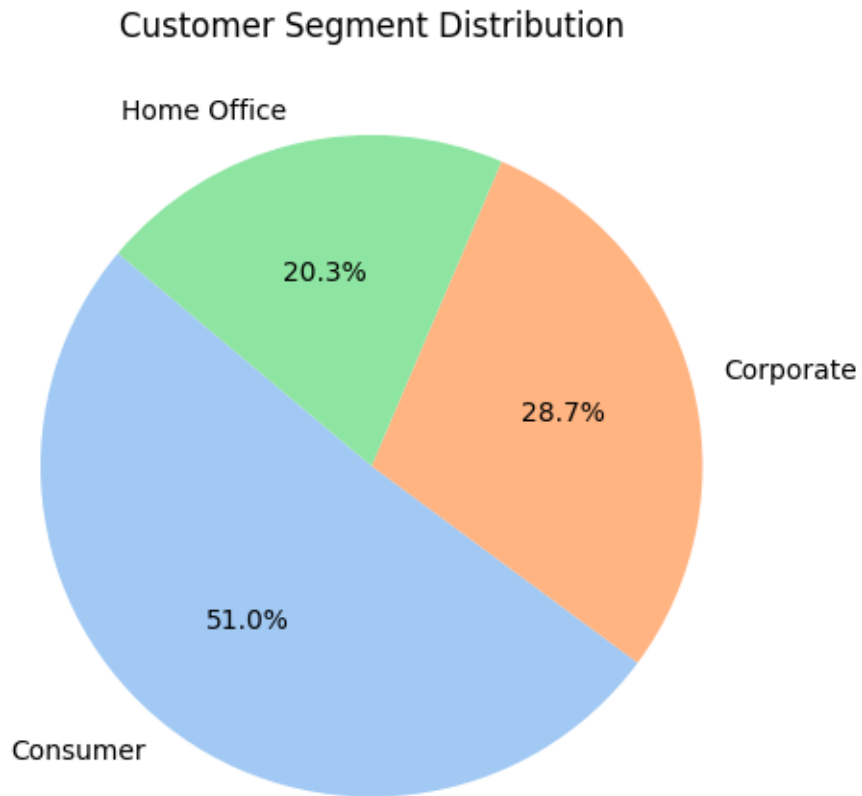
## **Insight 4: Customer Segment Distribution**

### **Analysis:**

The customer base mainly consists of **Consumer** and **Corporate** segments, with **Home Office** being a smaller share.

### **Recommendation:**

Continue focusing on Consumer and Corporate clients. Launch targeted campaigns to grow the Home Office segment.



*Figure 5 Customer Segment Distribution*

## Insights 5: Correlation Insights

### Analysis:

The **correlation heatmap** shows relationships between variables like **delivery times**, **risk**, and **discounts**.

### Recommendation:

Leverage these insights for predictive modelling (e.g., regression) to forecast risks and optimize operations.

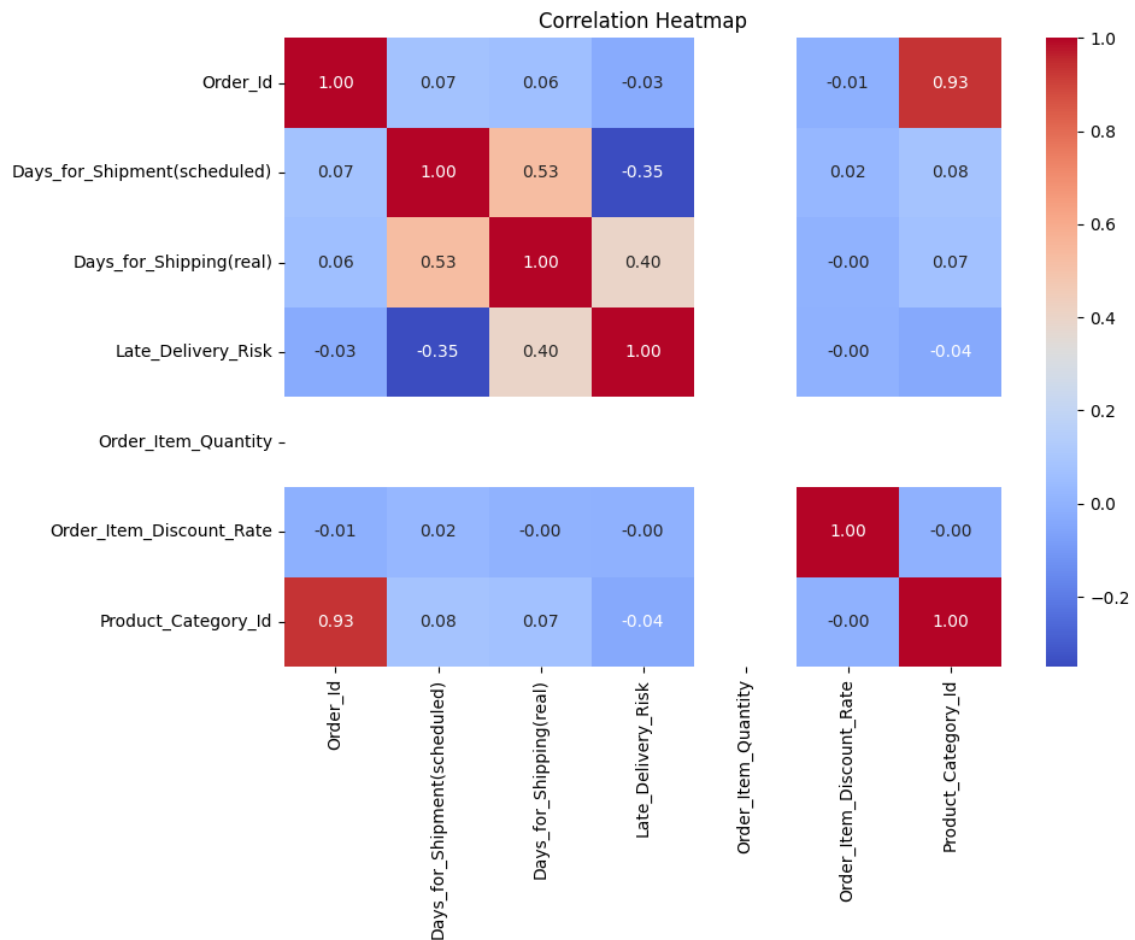


Figure 6 Correlation Heatmap

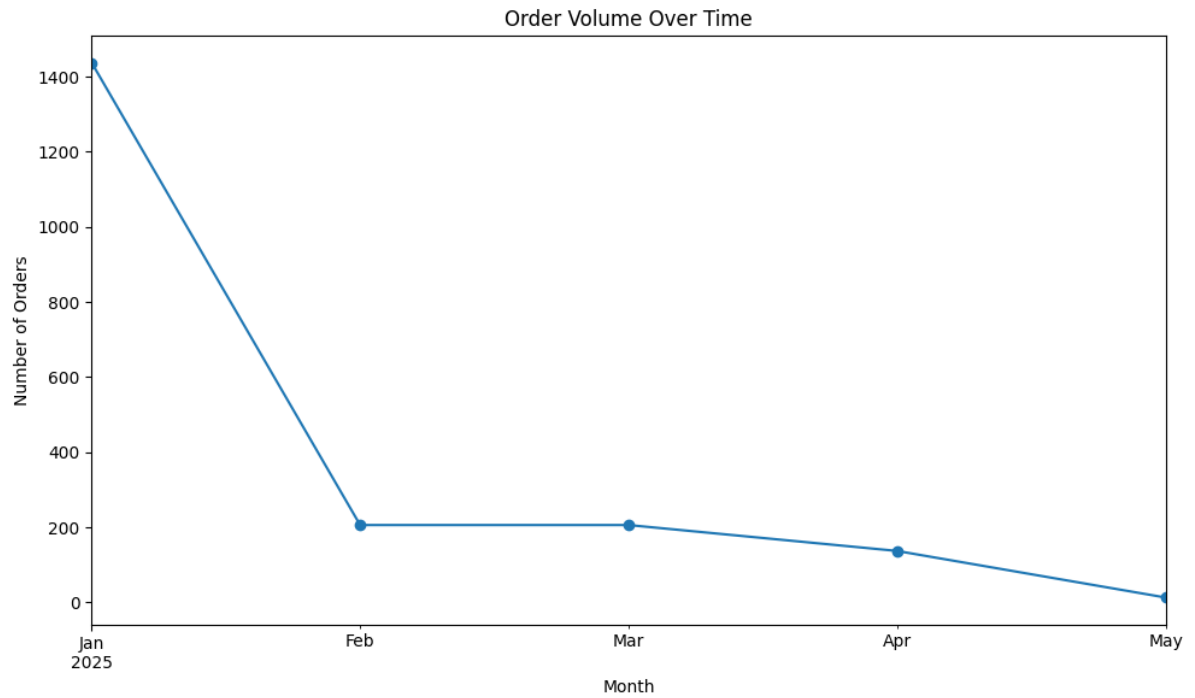
## Insight 6: Order Volume Over Time

### Analysis:

Monthly sales show fluctuations, hinting at **seasonal trends** or **operational bottlenecks**.

### Recommendation:

Align **inventory planning** and **staffing** with these demand cycles to ensure better fulfilment and reduce overstock or under-capacity.



*Figure 7 Oder Volume Over Time*

## Conclusion

This analysis highlights key opportunities to enhance supply chain efficiency. Reducing delivery delays, optimizing shipping methods, and addressing customer segment trends can significantly improve performance. The growing reliance on digital transactions and seasonal order patterns offers further insights to support strategic decision-making.