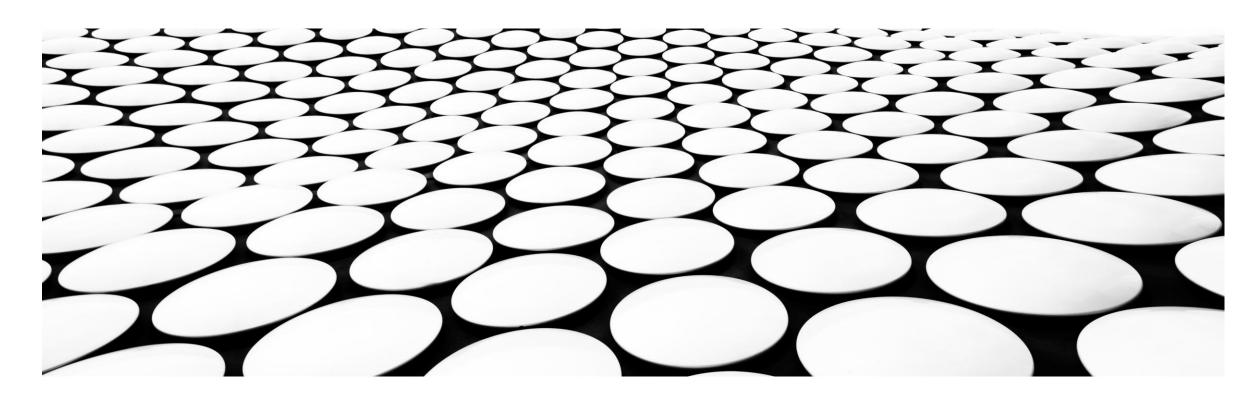
E-Commerce and Retail B2B Case Study

SUBMITTED BY

K R K PRASANNA KUMAR INDRANI SONAR ABHINAV



PROBLEM STATEMENT & GOALS

Problem Statement:

- Schuster, a multinational retail company, deals with vendors having credit arrangements.
- Some vendors make late payments, causing financial impact and inefficiency.
- Employees spend time chasing payments, leading to non-value-added activities.
- Schuster wants to understand payment behavior and predict late payments.

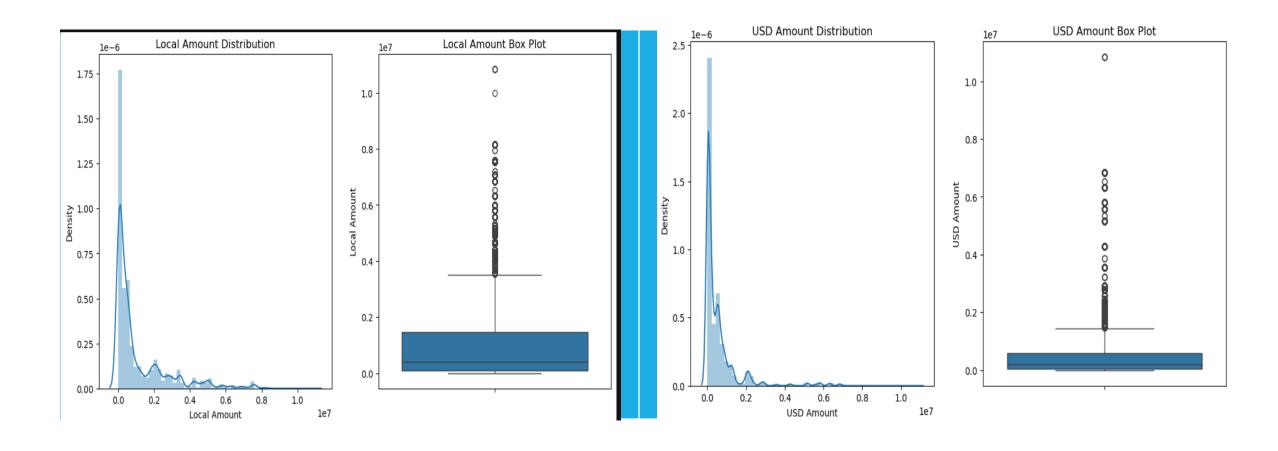
Goals:

- Analyze Payment Behavior: Segment customers based on past payment patterns.
- Predict Late Payments: Use historical data to predict delayed payments on open invoices.
- Improve Collection Efficiency: Help collectors prioritize vendors for timely payment.
- Develop a Model: Identify key factors influencing late payments and recommend a classification model for production deployment.

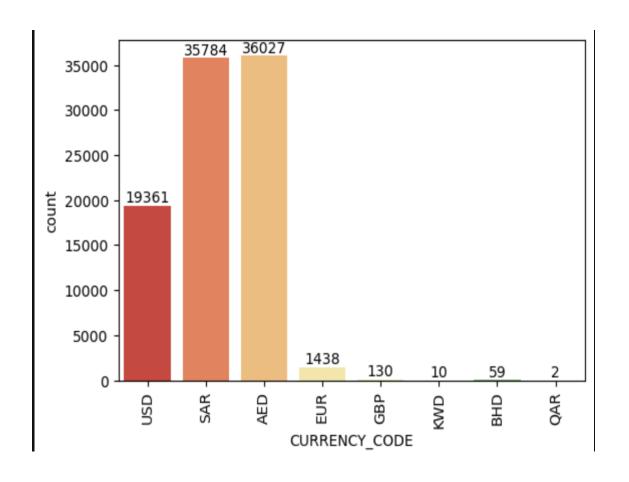
PROBLEM & PPRO&CH

- 1.Start
- 2.Data Reading and Understanding
- 3.EDA
- 4. Feature Engineering
- 5.K- Means Clustering
- 6.Model Building
- 7. Feature Tuning
- 8. Model Testing
- 9.Model Finalization
- 10.Predicting & Recommendations
- **11.End**

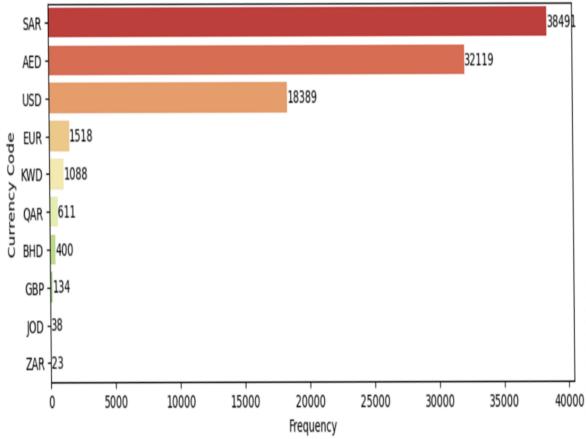
LOCAL AMOUNT & USD



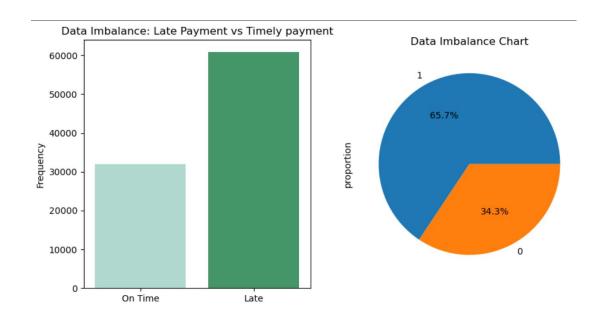
CURRENCY COUNT AND INVOICE DISTRIBUTION

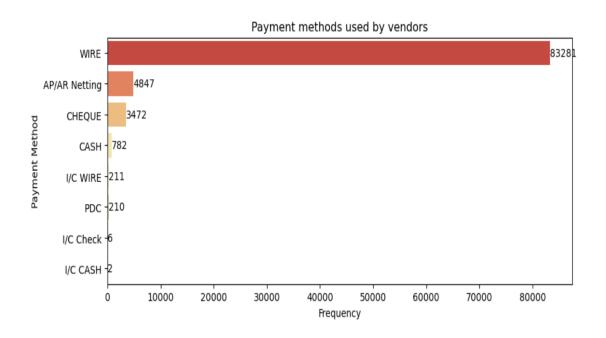




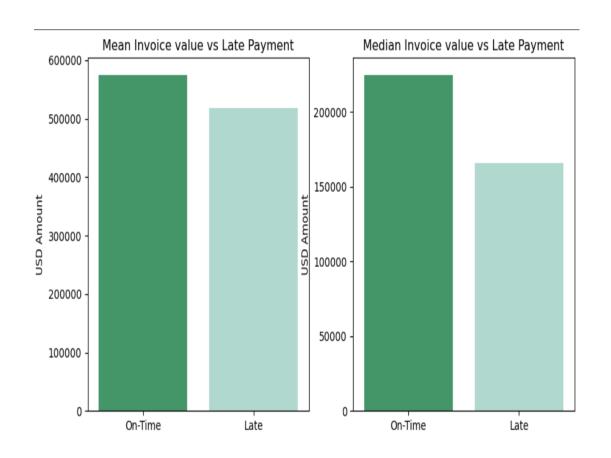


DATA IMBALANCE & PAYMENT METHOD





USD ONTIME AND LATE PAYMENT COMPARISON



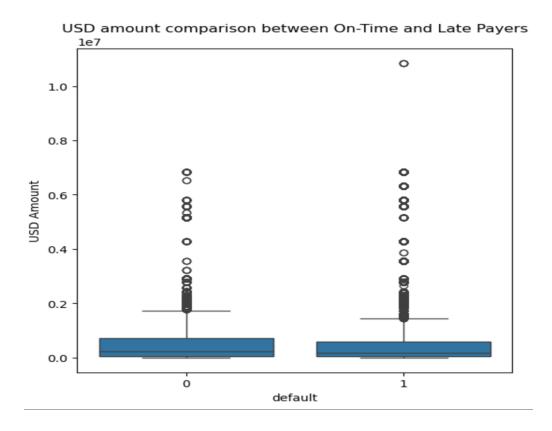
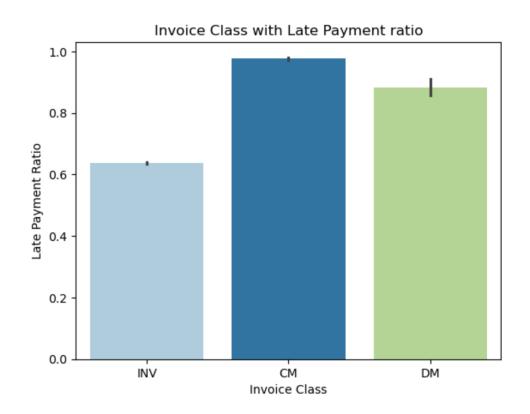
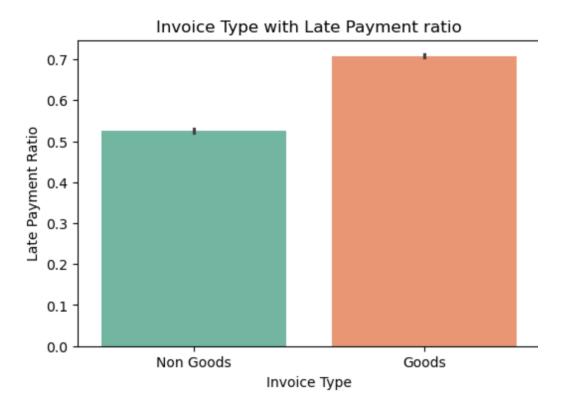


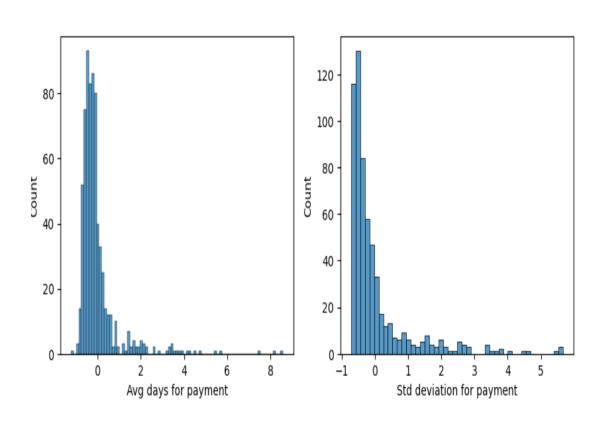
FIG 1: THE LATE PAYMENT RATIO IS HIGHEST FOR CREDIT NOTE TRANSACTIONS, FOLLOWED BY DEBIT NOTE AND INVOICE, INDICATING THAT CREDIT AND DEBIT NOTE TRANSACTION TYPES CARRY A HIGHER RISK OF PAYMENT DELAYS

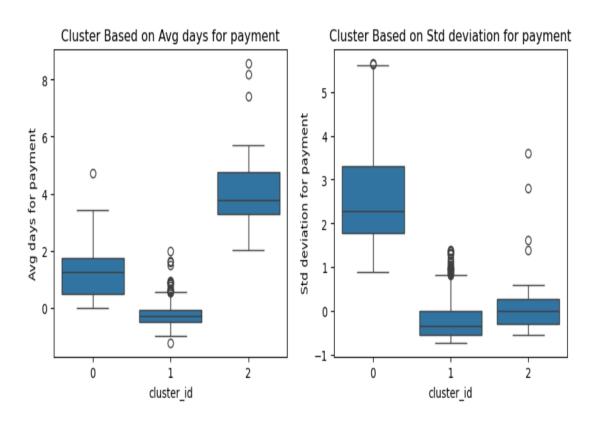
FIG 2 :GOODS-TYPE INVOICES EXHIBIT A HIGHER LATE PAYMENT RATIO COMPARED TO NON-GOODS INVOICES, INDICATING AN INCREASED LIKELIHOOD OF PAYMENT DELAYS FOR GOODS-RELATED TRANSACTIONS.





AVG DAYS & STANDARD DEVIATION OF PAYMENTS

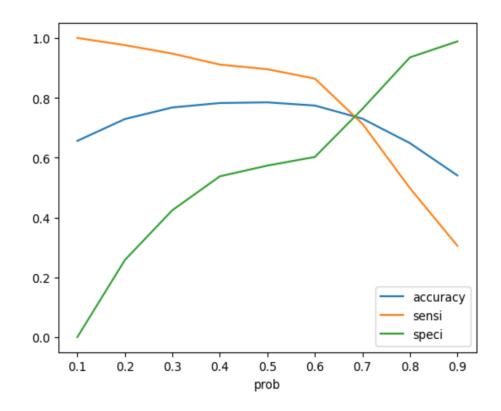


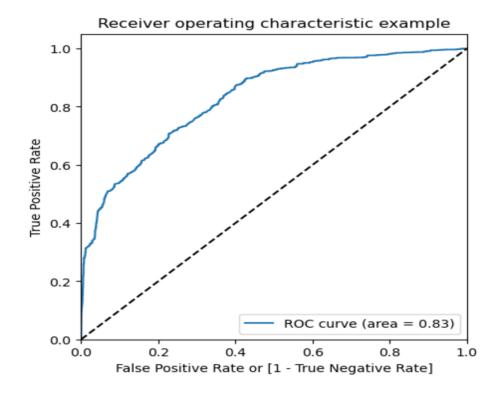


HEAT MAP OF TRAINING DATA

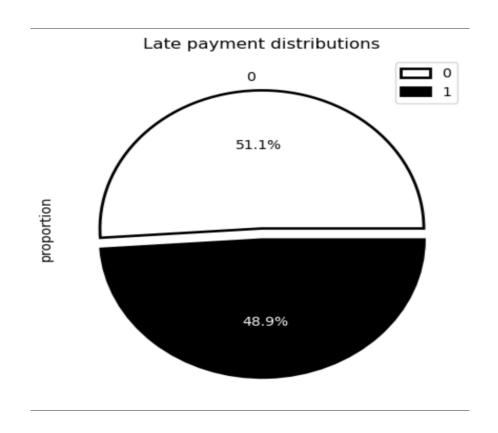
USD Amount -	1	-0.09	-0.11	-0.05	-0.056	-0.078	0.16	0.093	-0.088	-0.025	0.15	0.16	-0.023	
15 Days from EOM -	-0.09	1	-0.11	-0.12	-0.062	-0.059	-0.12	-0.15	-0.062	-0.047	-0.11	-0.39	-0.0079	
30 Days from EOM -	-0.11	-0.11	1	-0.16	-0.078	-0.074	-0.15	-0.19	-0.078	-0.059	-0.13	0.055	-0.027	
30 Days from Inv Date -	-0.05	-0.12	-0.16	1	-0.088	-0.084	-0.17	-0.22	-0.089	-0.067	-0.15	0.14	0.013	
45 Days from EOM -	-0.056	-0.062	-0.078	-0.088	1	-0.042	-0.083	-0.11	-0.044	-0.033	-0.074	-0.2	-0.00053	
45 Days from Inv Date -	-0.078	-0.059	-0.074	-0.084	-0.042	1	-0.079	-0.1	-0.042	-0.031	-0.071	0.017	-0.028	
60 Days from EOM -	0.16	-0.12	-0.15	-0.17	-0.083	-0.079	1	-0.2	-0.083	-0.062	-0.14	0.00043	-0.023	
60 Days from Inv Date -	0.093	-0.15	-0.19	-0.22	-0.11	-0.1	-0.2	1	-0.11	-0.082	-0.18	0.065	0.017	
90 Days from EOM -	-0.088	-0.062	-0.078	-0.089	-0.044	-0.042	-0.083	-0.11	1	-0.033	-0.075	0.091	0.056	
90 Days from Inv Date -	-0.025	-0.047	-0.059	-0.067	-0.033	-0.031	-0.062	-0.082	-0.033	1	-0.056	0.048	0.044	
Immediate Payment -	0.15	-0.11	-0.13	-0.15	-0.074	-0.071	-0.14	-0.18	-0.075	-0.056	1	0.053	-0.0066	
cluster_id -	0.16	-0.39	0.055	0.14	-0.2	0.017	0.00043	0.065	0.091	0.048	0.053	1	-0.0027	
Invoice_Month -	-0.023	-0.0079	-0.027	0.013	-0.00053	-0.028	-0.023	0.017	0.056	0.044	-0.0066	-0.0027	1	
	USD Amount -	15 Days from EOM -	30 Days from EOM -	30 Days from Inv Date -	45 Days from EOM -	45 Days from Inv Date -	60 Days from EOM -	60 Days from Inv Date -	90 Days from EOM -	90 Days from Inv Date -	Immediate Payment -	duster_id -	hvoice_Month -	

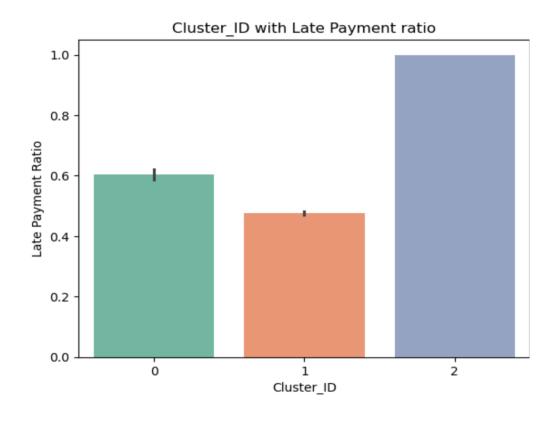
- •The **Logistic Regression Model** was developed by eliminating multicollinearity and unnecessary variables, resulting in significant features with a **good ROC curve** (0.83).
- •The **optimal probability cutoff of ~0.6** was identified from the trade-off analysis between accuracy, sensitivity, and specificity, used to predict delayed payment transactions.





- •The final model predicts that **51.1% of transactions** are likely to experience payment delays, which could significantly disrupt business operations.
- •Customers with a history of **prolonged payment days** are expected to have a **near 100% delay rate**, compared to those with historically early or medium payment days, aligning with previous historical trends.





- •Predictions indicate that **customers with the highest delay probabilities** are those with the most frequent delayed payments and total payments.
- •The companies listed in the table on the left show the **maximum probability of default**, indicating the highest risk for delayed payments.

	Delayed_Payment	Total_Payments	Delay%
Customer_Name			
IL G Corp	13	13	100.0
RNA Corp	9	9	100.0
SHIS Corp	8	8	100.0
ALSU Corp	7	7	100.0
LVMH Corp	4	4	100.0
FINA Corp	4	4	100.0
V PE Corp	4	4	100.0
TRAF Corp	3	3	100.0
MAYC Corp	3	3	100.0
VIRT Corp	3	3	100.0

RECOMMENDATIONS

1. Tighten Payment Policies for Credit Note Payments:

- Credit Note payments experience the greatest delay rates compared to Debit Note or Invoice types.
- Enforce stricter payment collection policies for Credit Note invoices to reduce delays.

2.Stricter Policies for Goods-Type Invoices:

- Goods-type invoices have significantly higher payment delay rates.
- Implement stricter payment policies for goods-type invoices to address higher delay tendencies.

3. Focus on Lower-Value Payments:

- Lower-value payments make up most transactions and are associated with higher delay rates.
- Prioritize smaller payments, and consider applying penalties for late payments, especially for lower-value invoices.

4. Special Attention to Cluster 1 Customers (Prolonged Payment Duration):

- Cluster 1 customers, with prolonged payment delays, show much higher delay rates than early or medium-payers.
- Give special attention to Cluster 1 customers to reduce payment delays.

5. Prioritize Companies with High Delayed Payments and Probability of Future Delays:

- Companies with the highest total and delayed payment counts should be prioritized.
- Closely monitor and intervene early with companies that show high delay probabilities.

thank