SUPPLY CHAIN ANALYTICS

Delivery Performance & Late Delivery Prediction

Executive Business Analysis

180,518	57.3%	1 day	Decision Tree
ORDERS ANALYZED	DELAY RATE	MEDIAN DEVIATION	BEST MODEL

Prepared from the DataCo Supply Chain dataset

SUPPLY CHAIN ANALYTICS

Contents

01 Executive Summary

02 Business Context & Objectives

03 Performance Snapshot

04 Delivery Performance

05 Factors Associated with Delay

06 Financial Performance

07 Time-Based Patterns

08 Predictive Model Results

09 Recommendations

10 Conclusion

Appendix Methodology

01 Executive Summary

The analysis covers 180,518 cleaned orders from the DataCo Supply Chain dataset. The main business question is straightforward: where are delivery delays concentrated, what patterns are associated with them, and can late-delivery risk be predicted from information available at order level?

Delivery performance is the main concern. 103,399 orders, or 57.3% of the dataset, exceeded their scheduled shipping time. The median schedule deviation is 1 day and the 90th percentile is 2 days.

The clearest operational difference is by shipping mode. First Class has a 100.0% delay rate and Second Class 79.7%, compared with 39.8% for Standard Class. Regional, departmental, and customer-segment differences are narrower.

Financial measures also differ by delivery status. Average Benefit per Order is $21.58 for delayed orders versus $22.50 for on-time orders. These figures show association, not a proven causal loss.

For predictive analysis, three baseline models were compared. The Decision Tree produced the highest F1 score at 0.763. This provides a useful analytical baseline for identifying orders with higher late-delivery risk.

MANAGEMENT FOCUS Review shipping-mode performance first, then investigate the highest-delay regions and departments. Use the predictive model as a supporting prioritization tool.

02 Business Context & Objectives

The dataset represents an e-commerce supply chain with order, customer, product, regional, shipping, timing, and financial information. Delivery performance is assessed by comparing actual shipping days with scheduled shipping days.

Business question

The analysis is designed to understand the current delivery-performance position, identify business dimensions associated with higher delay rates, and evaluate whether order-level information can support early identification of late-delivery risk.

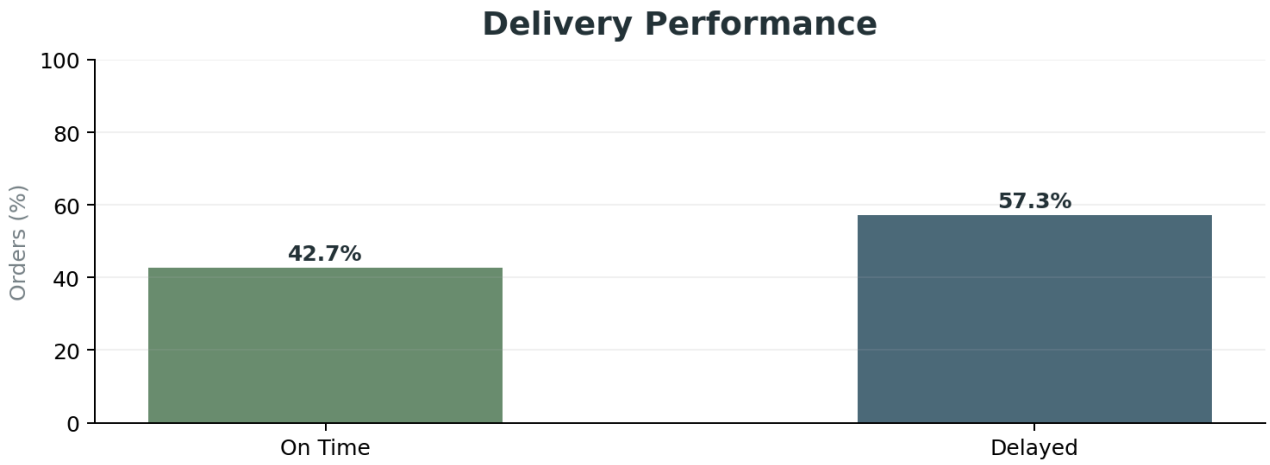
Objectives

- Establish a delivery-performance baseline.
- Compare delays across shipping mode, region, customer segment, department, order type, and order status.
- Identify recurring time-based patterns.
- Compare selected financial measures for delayed and on-time orders.
- Build and compare baseline classification models.
- Translate the findings into practical areas for management review.

03 Performance Snapshot

180,518	103,399	57.3%	42.7%
TOTAL ORDERS	DELAYED ORDERS	DELAY RATE	ON-TIME RATE

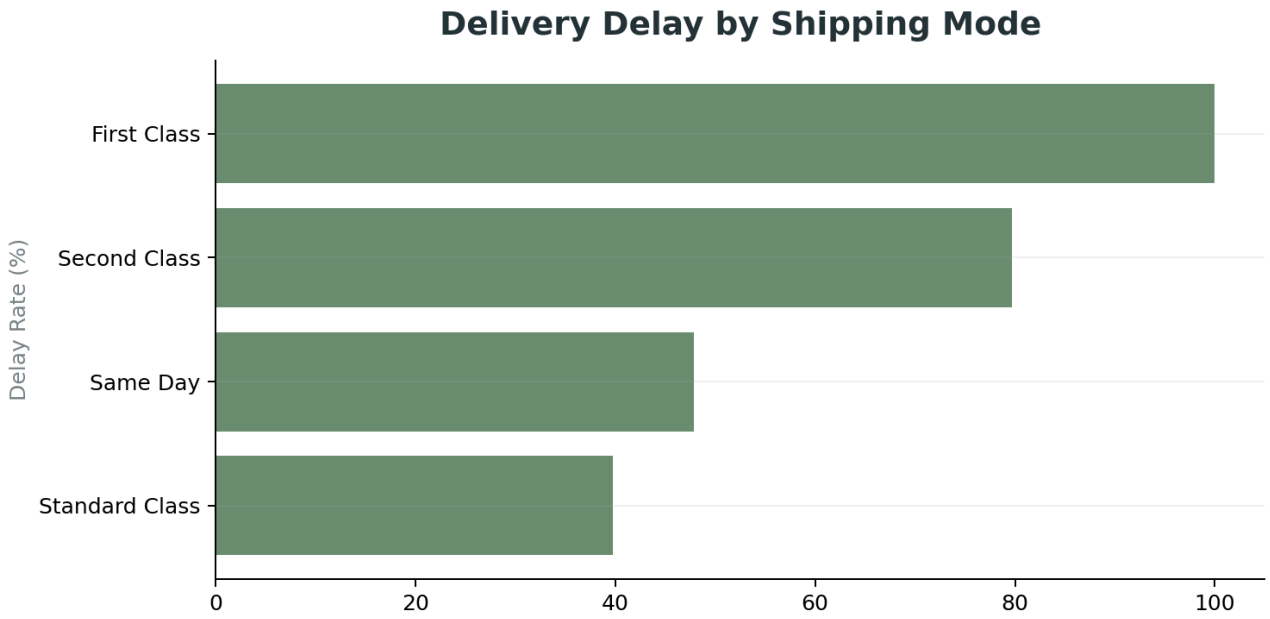
1 day	2 days	$183.11	$21.97
MEDIAN DEVIATION	90TH PERCENTILE	AVG SALES / CUSTOMER	AVG BENEFIT / ORDER



PERFORMANCE READING The delivery problem is frequent rather than isolated. More than half of all orders exceed their scheduled shipping time.

04 Delivery Performance

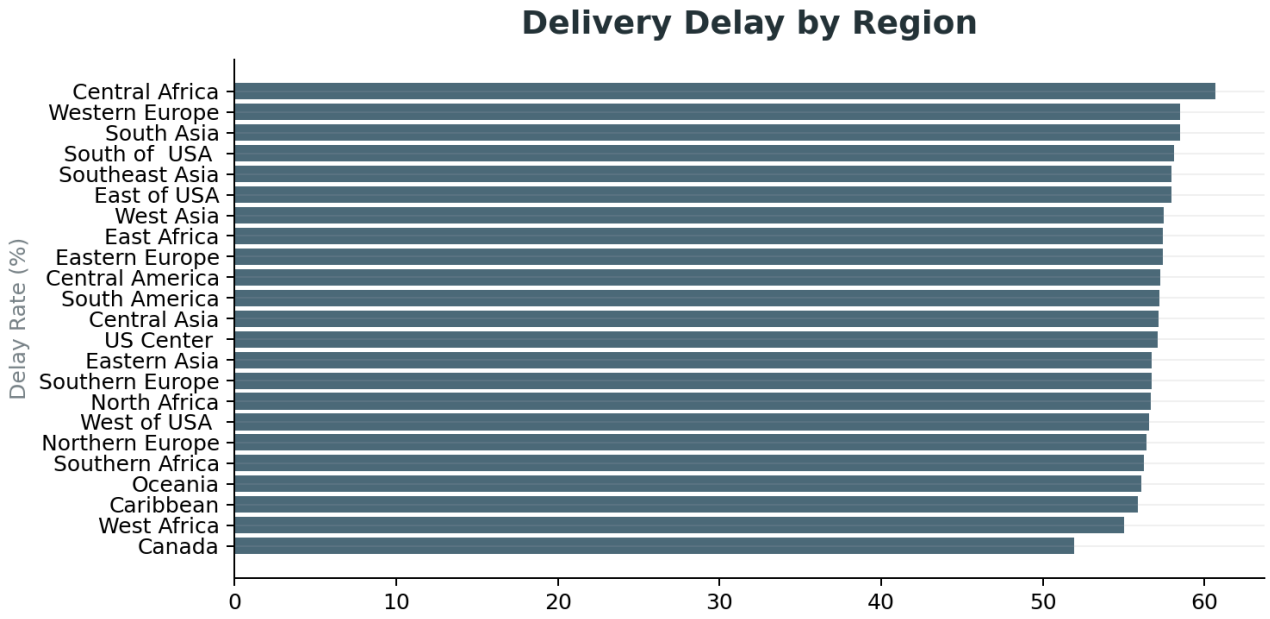
Shipping mode is the clearest differentiator



First Class has the highest observed delay rate at 100.0%, followed by Second Class at 79.7%. Standard Class is lower at 39.8%. Same Day is also elevated at 47.8%.

MANAGEMENT IMPLICATION Premium service promises should be reviewed against actual delivery performance, carrier allocation, and routing.

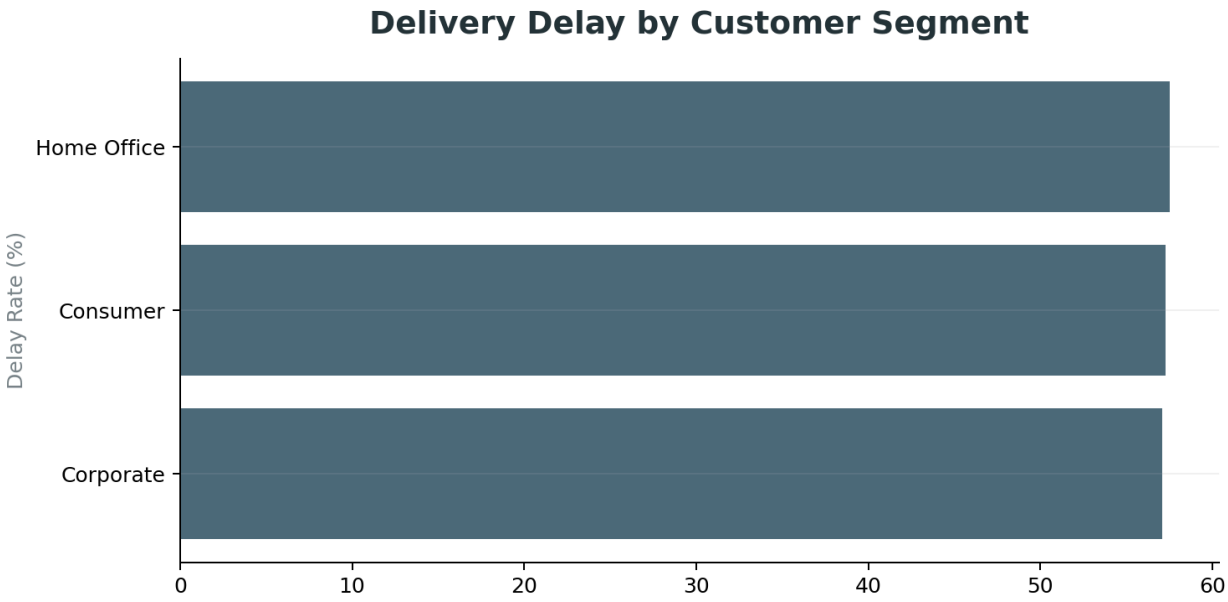
Regional performance



Central Africa records the highest regional delay rate at 60.7%. Most other regions sit within a relatively narrow range, so the regional pattern is important but less pronounced than the shipping-mode pattern.

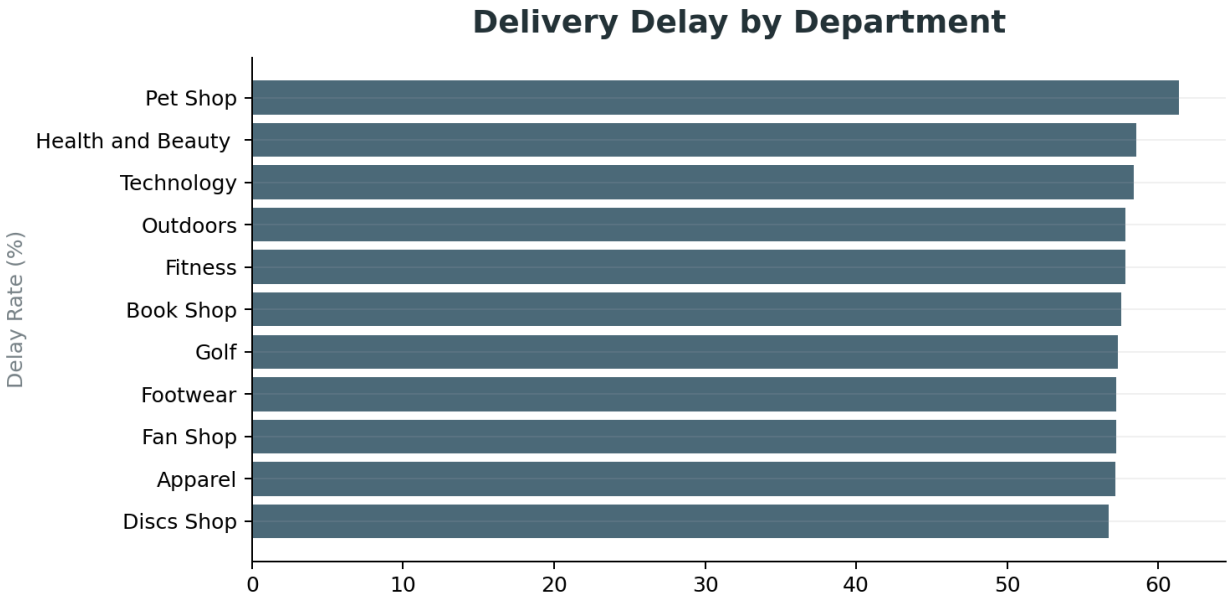
05 Factors Associated with Delay

Customer segment



Customer-segment differences are limited. Home Office has the highest observed rate at 57.5%, followed by Consumer at 57.3% and Corporate at 57.1%.

Department



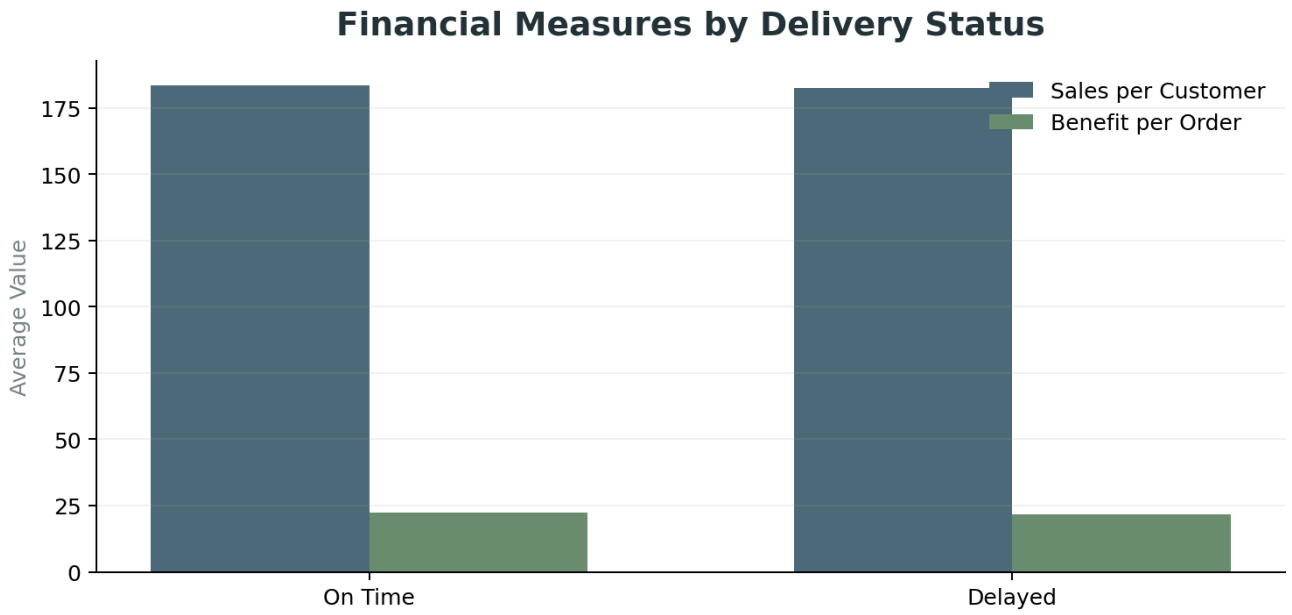
Pet Shop has the highest department delay rate at 61.4%. The next highest departments are Health and Beauty (58.6%) and Technology (58.4%).

Order type and status

Order type and order status show smaller differences than shipping mode. Their observed ranges are 56.6%-57.5% and 55.6%-57.9%, respectively.

INTERPRETATION These dimensions identify patterns for operational review. They should not be described as proven root causes without additional causal analysis.

06 Financial Performance



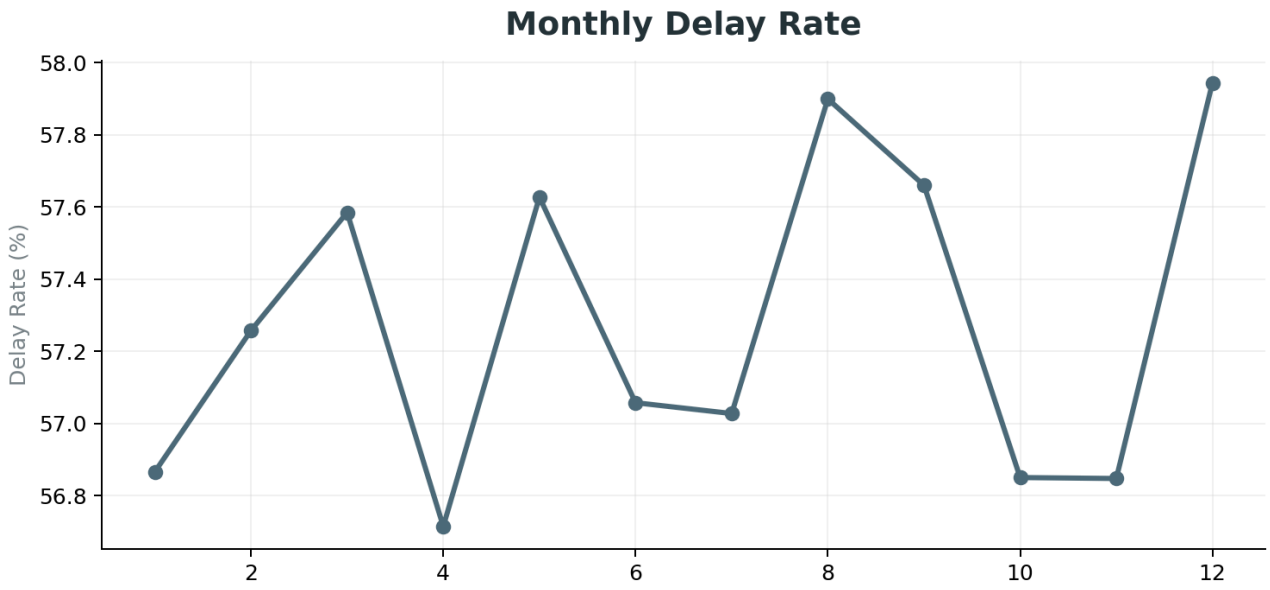
Average Benefit per Order is $21.58 for delayed orders and $22.50 for on-time orders. Average Sales per Customer is $182.69 for delayed orders and $183.67 for on-time orders.

The differences are relatively small at the average order level. The current dataset analysis therefore supports an association between delivery status and selected financial measures, but it does not quantify a causal financial loss.

NEXT ANALYTICAL STEP A richer profitability analysis should combine revenue, profit, discount, shipping cost, and customer-impact measures at order level.

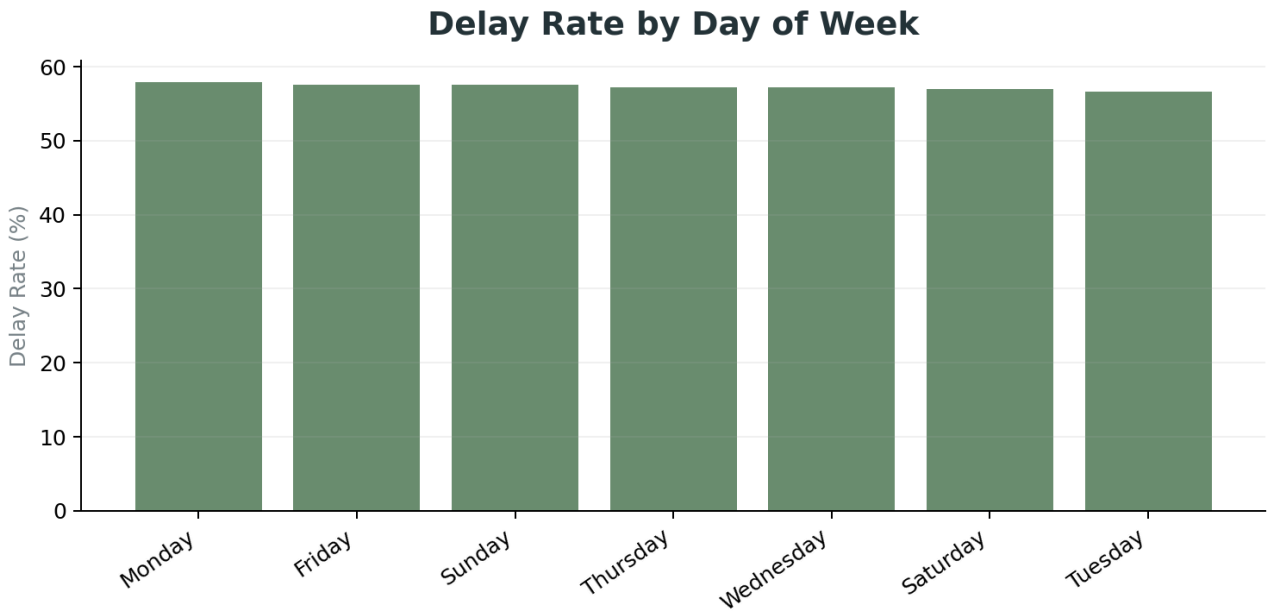
07 Time-Based Patterns

Monthly pattern



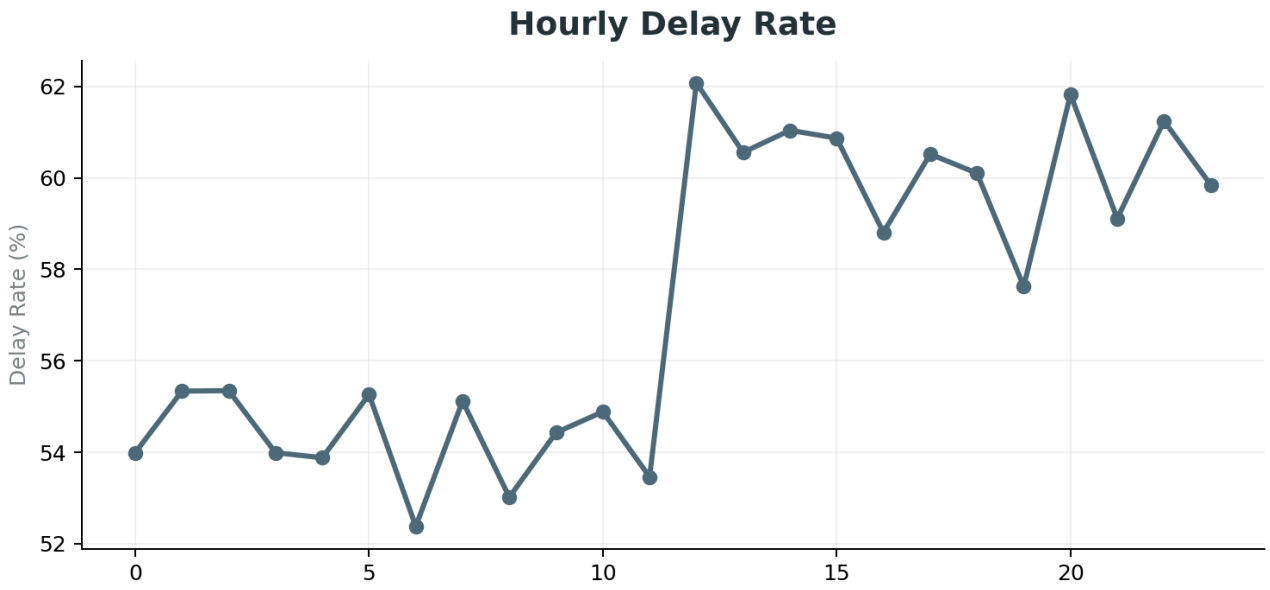
Month 12 has the highest observed delay rate at 57.9%, while month 4 is the lowest at 56.7%. The overall spread is modest.

Day of week



Monday has the highest day-of-week rate at 57.9%, while Tuesday has the lowest at 56.6%. The small spread suggests day of week is not a major management lever.

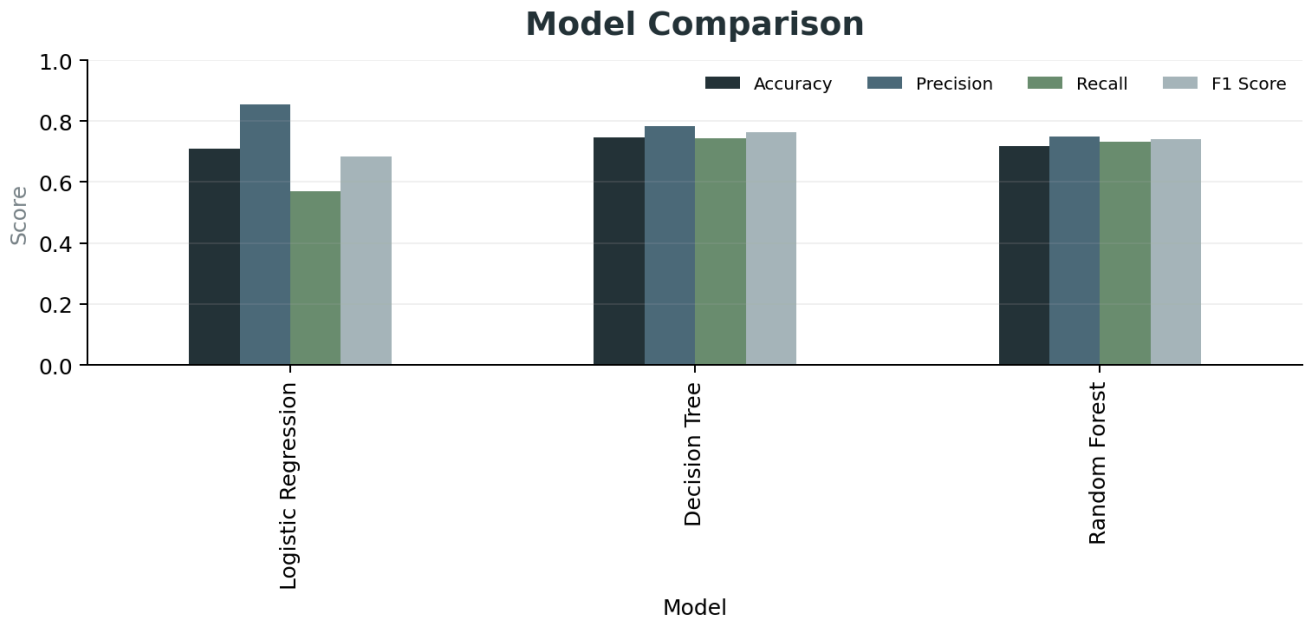
Hour of day



The three highest observed order hours are 12:00 (62.1%), 20:00 (61.8%), 22:00 (61.2%). These periods can be reviewed alongside workload, cutoff times, and processing capacity.

08 Predictive Model Results

The machine-learning task predicts the dataset's Late_delivery_risk target. The model uses order type, scheduled shipping days, category, customer segment, department, region, shipping mode, order month, and order hour. Actual shipping days and the engineered delay flag are not used as predictors.



Model	Accuracy	Precision	Recall	F1
Logistic Regression	0.710	0.854	0.569	0.683
Decision Tree	0.746	0.783	0.743	0.763
Random Forest	0.718	0.748	0.731	0.740

The Decision Tree is the strongest baseline under the selected F1 criterion, with an F1 score of 0.763 and accuracy of 0.746.

Selected model performance

For late-risk class 1, precision is 0.78, recall is 0.74, and F1 is 0.76. The model should be treated as a decision-support baseline rather than a production system.

09 Recommendations

01 Review shipping-mode performance

First Class and Second Class show delay rates of 100.0% and 79.7%. Validate carrier allocation, routing, and promised service levels.

02 Investigate regional constraints

Central Africa has the highest regional delay rate at 60.7%. Review carrier coverage, processing time, and regional service assumptions.

03 Review high-delay departments

Pet Shop has the highest department rate at 61.4%. Compare inventory availability, fulfillment time, and carrier assignment.

04 Use time patterns for capacity planning

Review month 12 and the highest hourly periods against workload and processing capacity.

05 Use predictive risk as a prioritization tool

The Decision Tree provides an F1 score of 0.763. Use risk scoring to prioritize operational review while further validating the model.

06 Expand the financial view

Add order-level revenue, profit, discount, shipping cost, and customer-impact measures to quantify the business cost of delivery performance.

Priority

The first management review should focus on shipping-mode performance and the operating processes behind the highest-delay regions. Predictive scoring can then be used to support order-level prioritization.

10 Conclusion

The analysis shows a delivery operation with a 57.3% observed delay rate across 180,518 cleaned orders. Shipping mode provides the clearest separation in delivery performance, while customer segment and day-of-week differences are comparatively small.

Regional and department patterns provide additional areas for operational review. The financial comparison shows modest differences in selected average measures between delayed and on-time orders, but the current analysis does not establish causal financial impact.

The predictive analysis provides a useful baseline. The Decision Tree achieved the highest F1 score among the three tested models. Further improvement would require additional operational features, validation, and model tuning.

MANAGEMENT TAKEAWAY The delivery issue is broad enough to require process-level attention, not only case-by-case intervention. The analysis provides a practical starting point for prioritizing that review.

Appendix Methodology

Step	Approach
Dataset	DataCo Supply Chain dataset
Selection	14 relevant delivery, business, financial, and date fields
Cleaning	1 duplicate row removed; dates converted to datetime
Delay measure	Actual shipping days minus scheduled shipping days
Delayed flag	Delay greater than zero
Time features	Month, day of week, and hour
Encoding	One-hot encoding for categorical ML features
Split	80% train / 20% test, stratified
Models	Logistic Regression, Decision Tree, Random Forest
Selection criterion	Highest F1 score