

ON-TIME EVERY-TIME

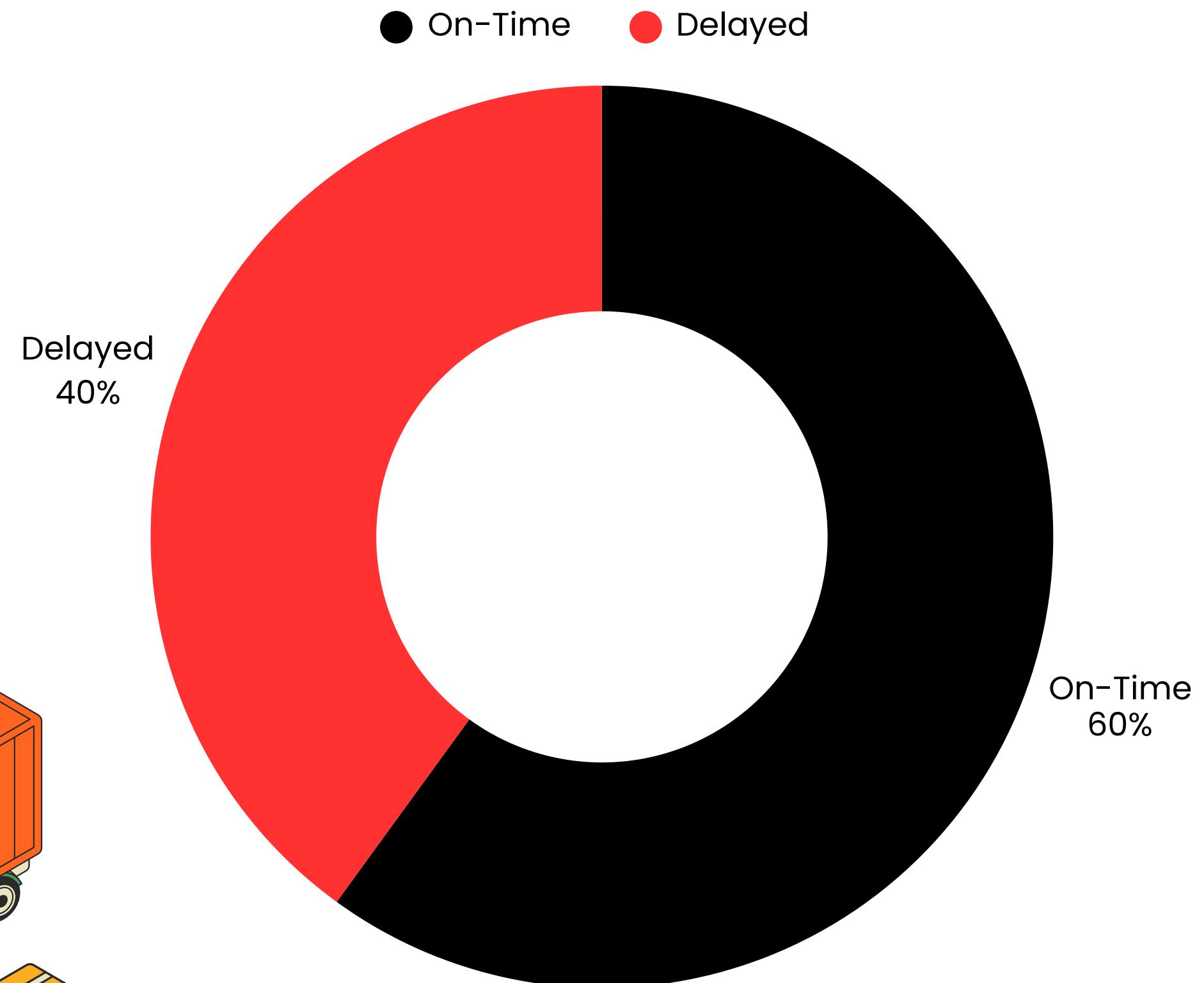
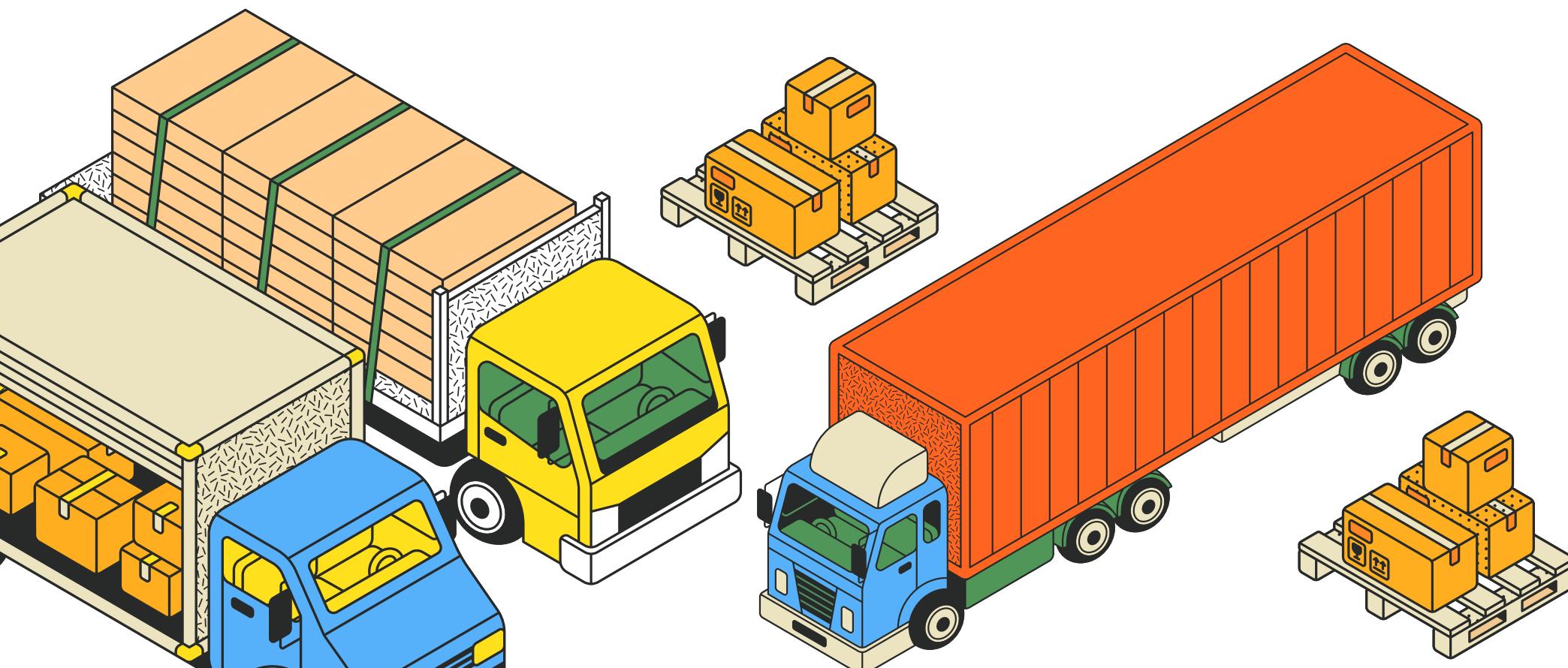
E-COMMERCE PRODUCT DELIVERY PREDICTION

Presented By - Shruti Rathod

Capstone Project - BIA Feb 2025 Batch

THE PROBLEM: LATE DELIVERIES IN E-COMMERCE

- Late deliveries significantly impact customer satisfaction and brand reputation.
- Our data shows a critical issue: **40%** of our deliveries were delayed, totaling 4,436 out of 10,999 shipments.





WHY IT MATTERS

Timely deliveries are crucial for **building customer satisfaction and loyalty**. Conversely, late shipments directly lead to increased complaints and higher customer care call volumes.

This impacts our business by eroding trust, reducing repeat purchases, and significantly increasing operational costs.

PROJECT OBJECTIVE

PREDICTIVE MODEL DEVELOPMENT

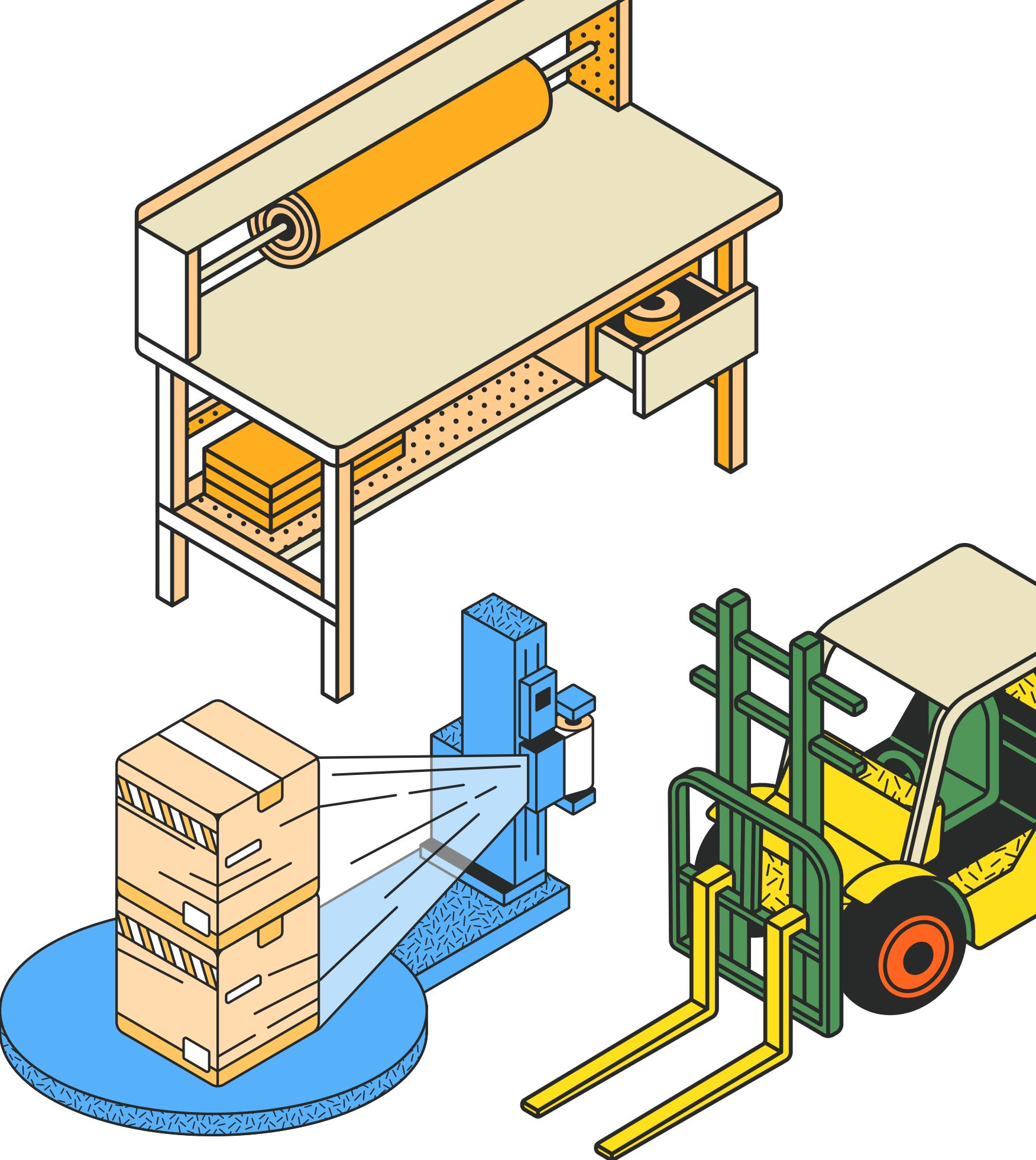
Build a robust model to forecast on-time or delayed deliveries based on historical data.

LOGISTICS OPTIMIZATION

Utilize data insights from the model to enhance shipping modes, routing, and warehouse efficiency.

CUSTOMER SATISFACTION ALIGNMENT

Bridge the gap between customer expectations and actual delivery performance, building trust and loyalty.



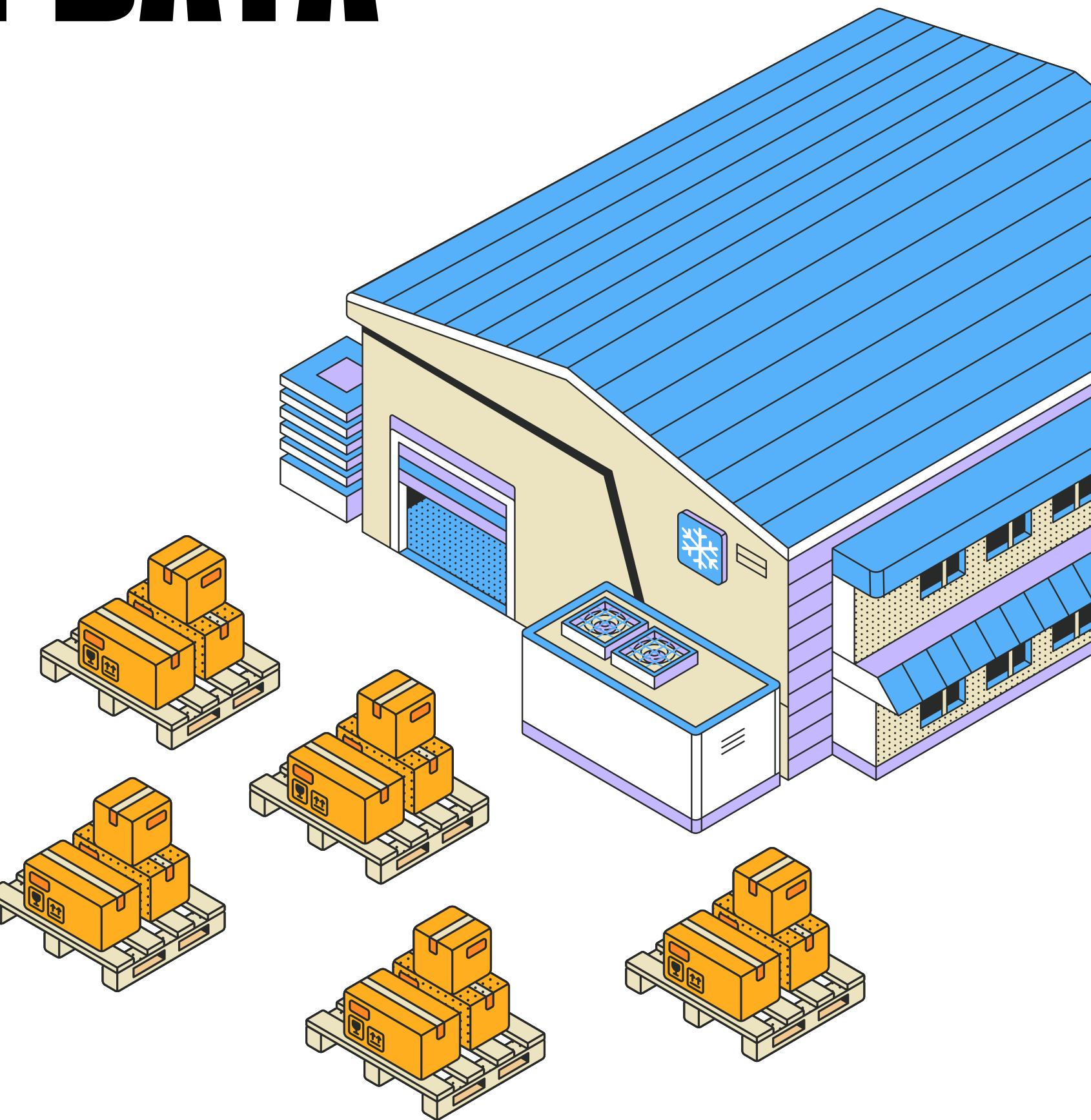
UNDERSTANDING THE DATA

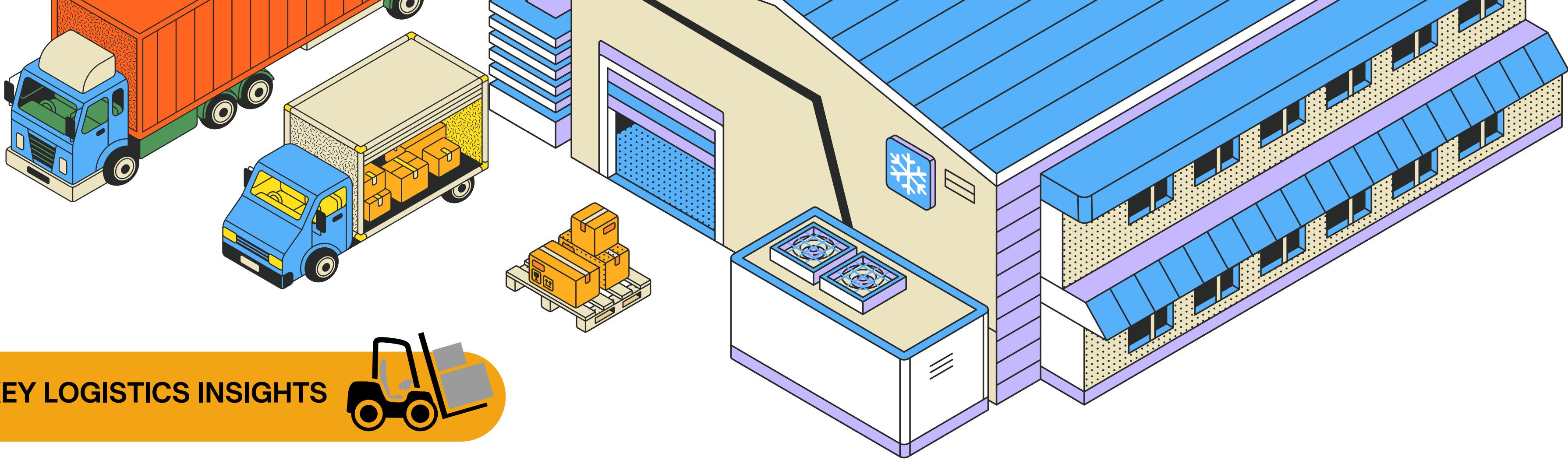
Our predictive model leverages a comprehensive dataset of 10,999 ecommerce delivery records.

Each record includes **12** distinct features, covering critical aspects such as:

- Warehouse location
- Shipping mode
- Product weight
- Discount offered
- Customer care calls

The model's target is to accurately predict if a delivery will be **on-time (1)** or **delayed (0)**.





KEY LOGISTICS INSIGHTS

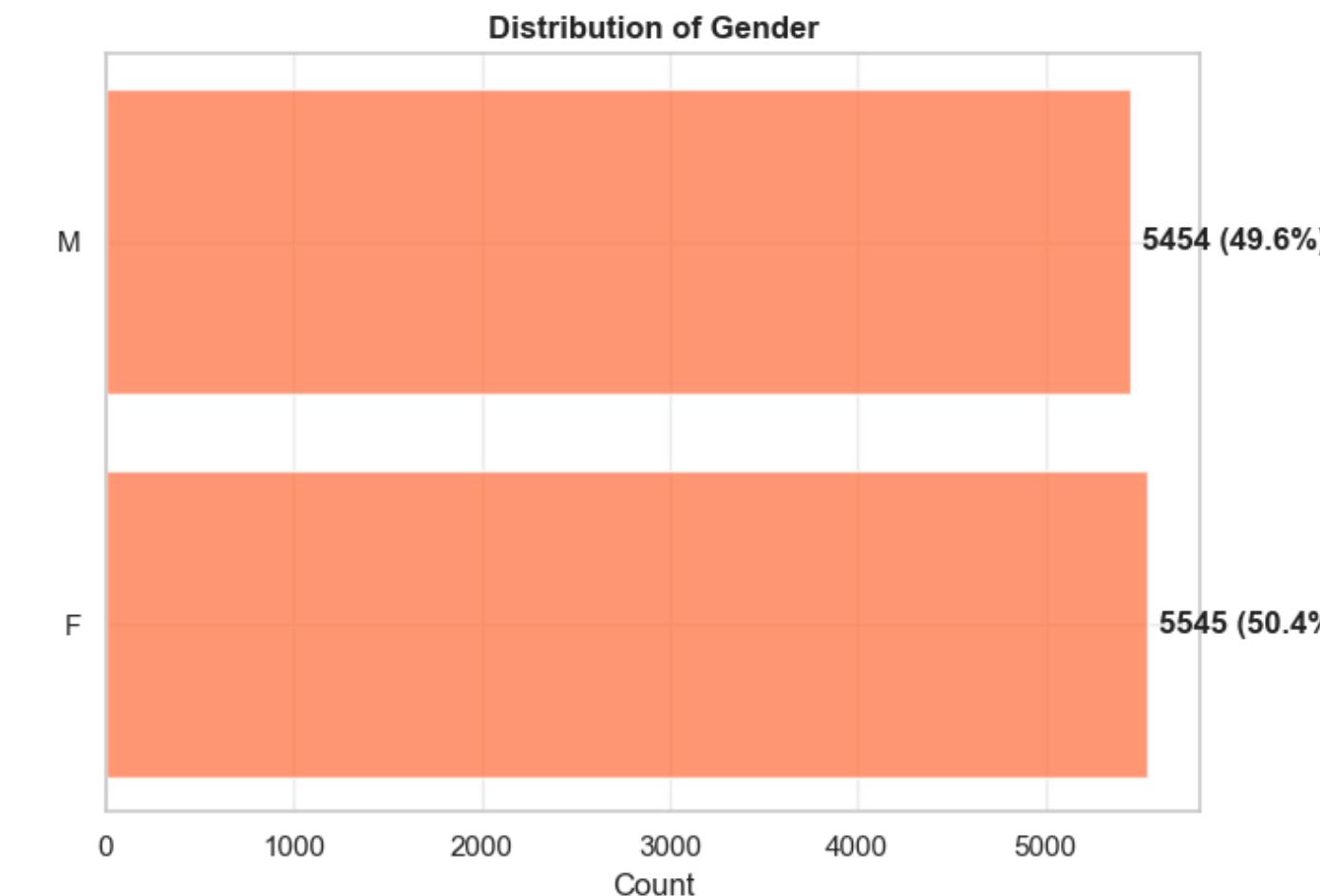
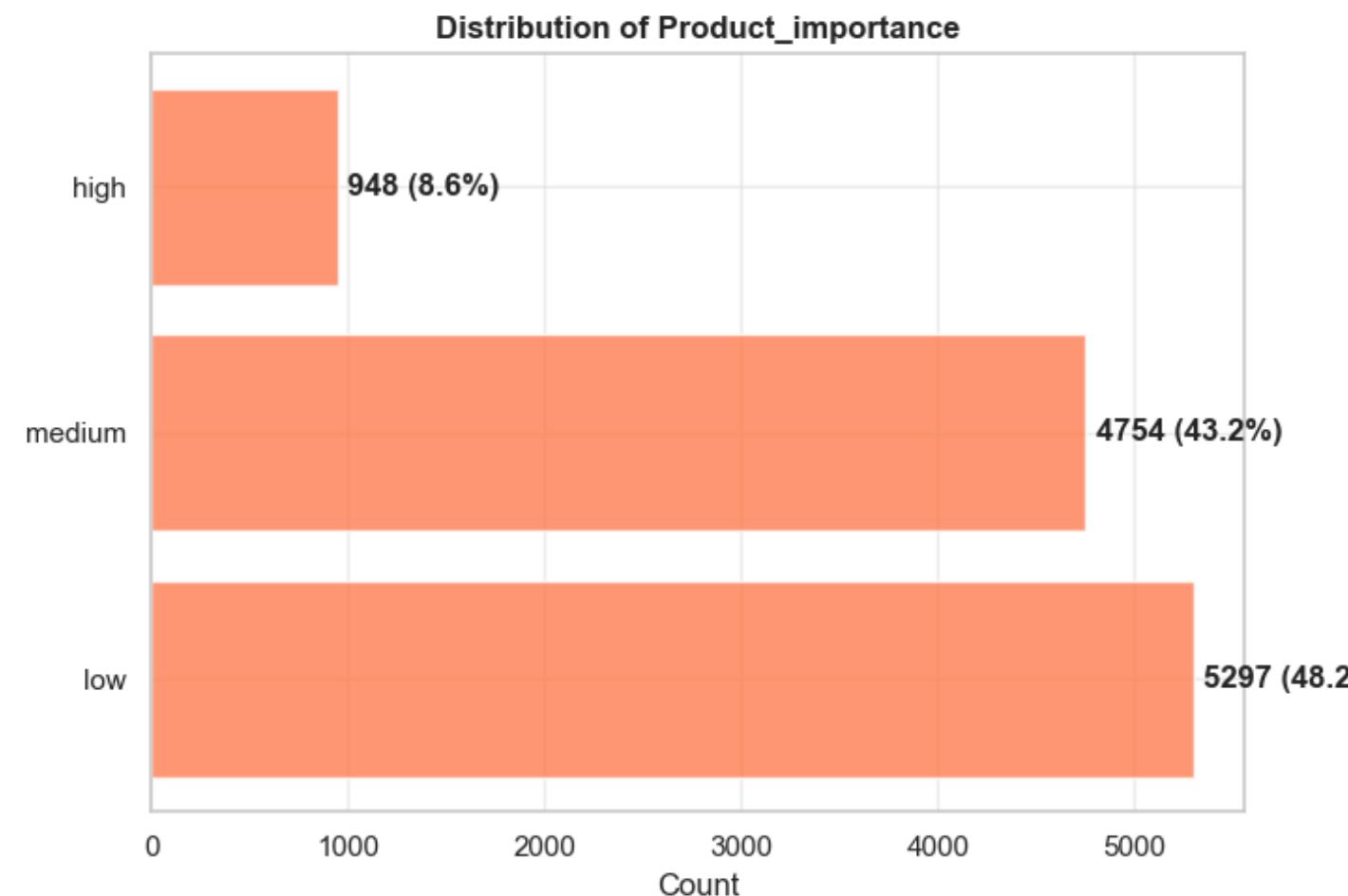
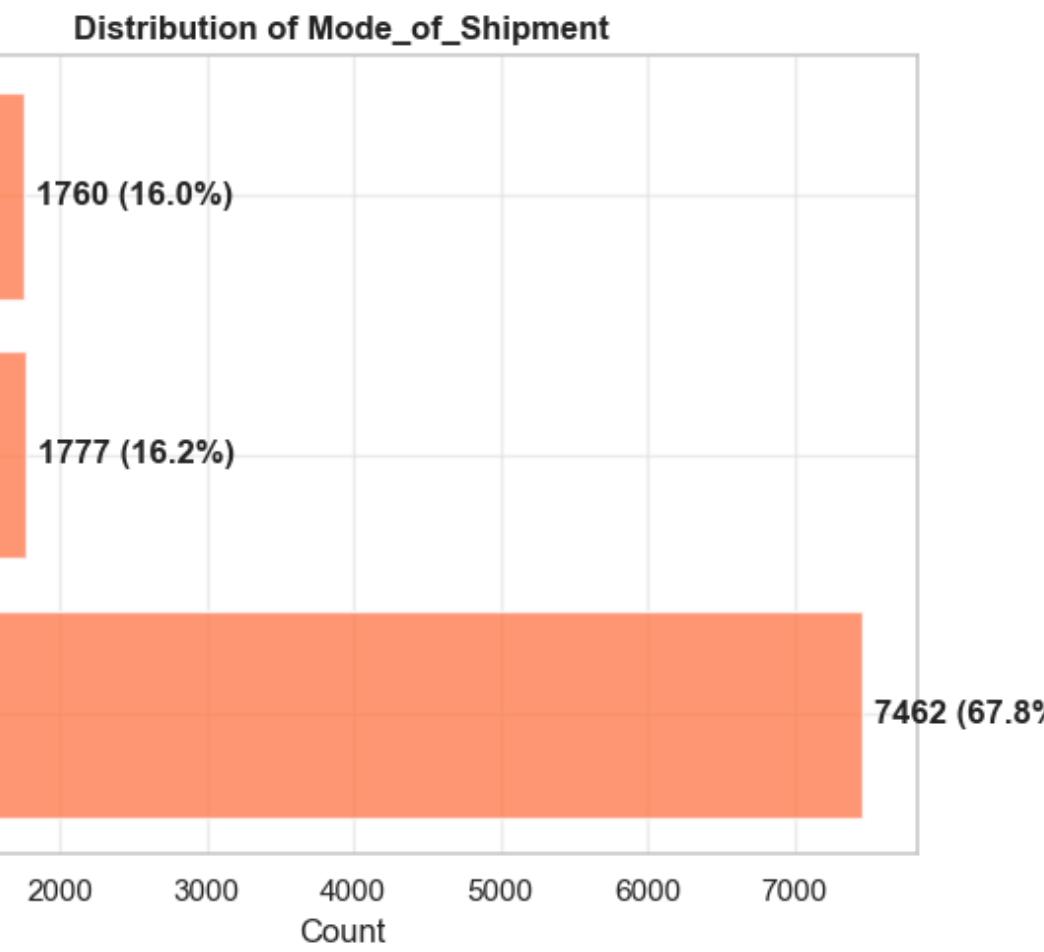
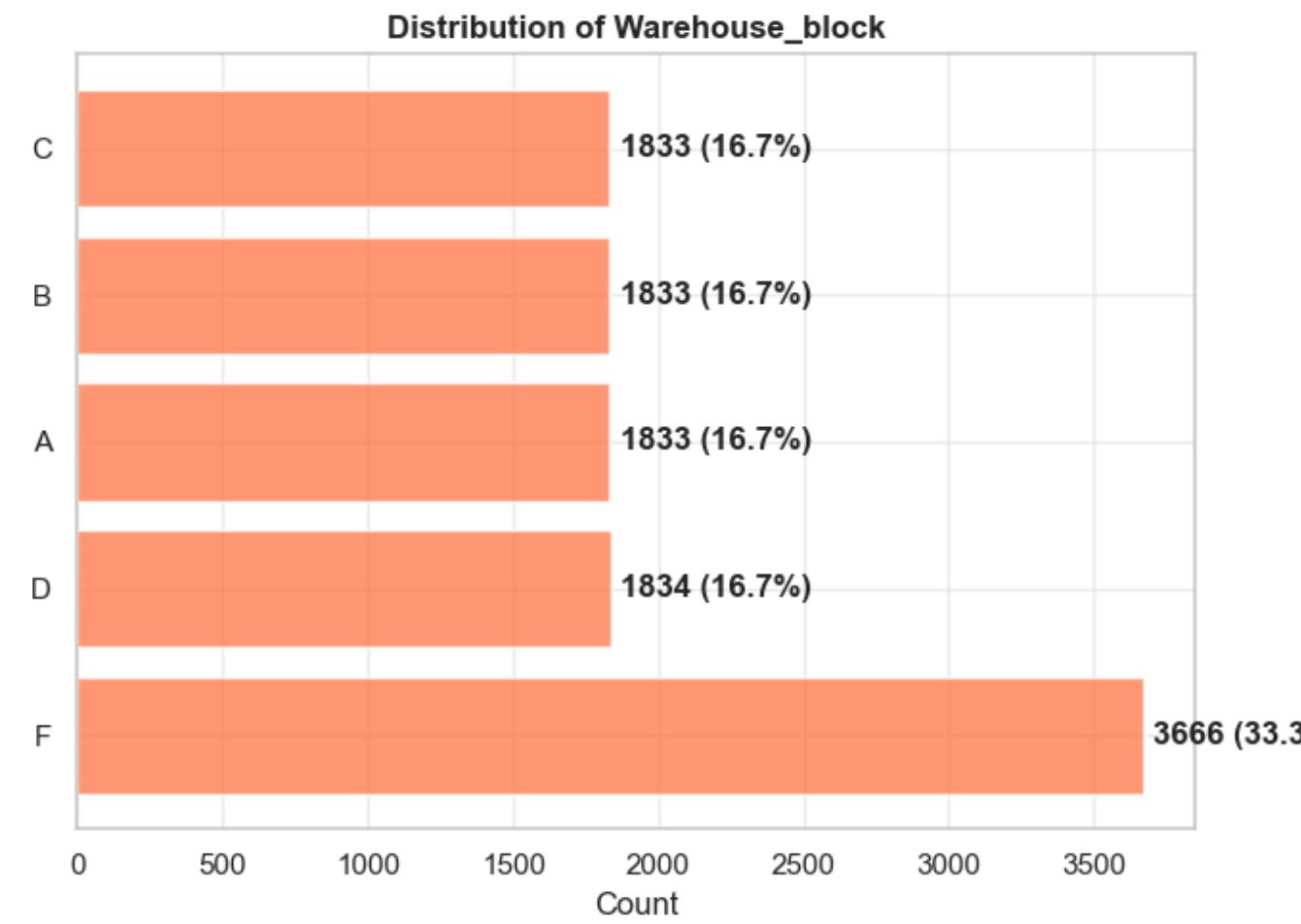


Heavier products (**high Weight_in_gms**) and significant discounts (**Discount_offered**) are key drivers increasing the risk of delivery delays.

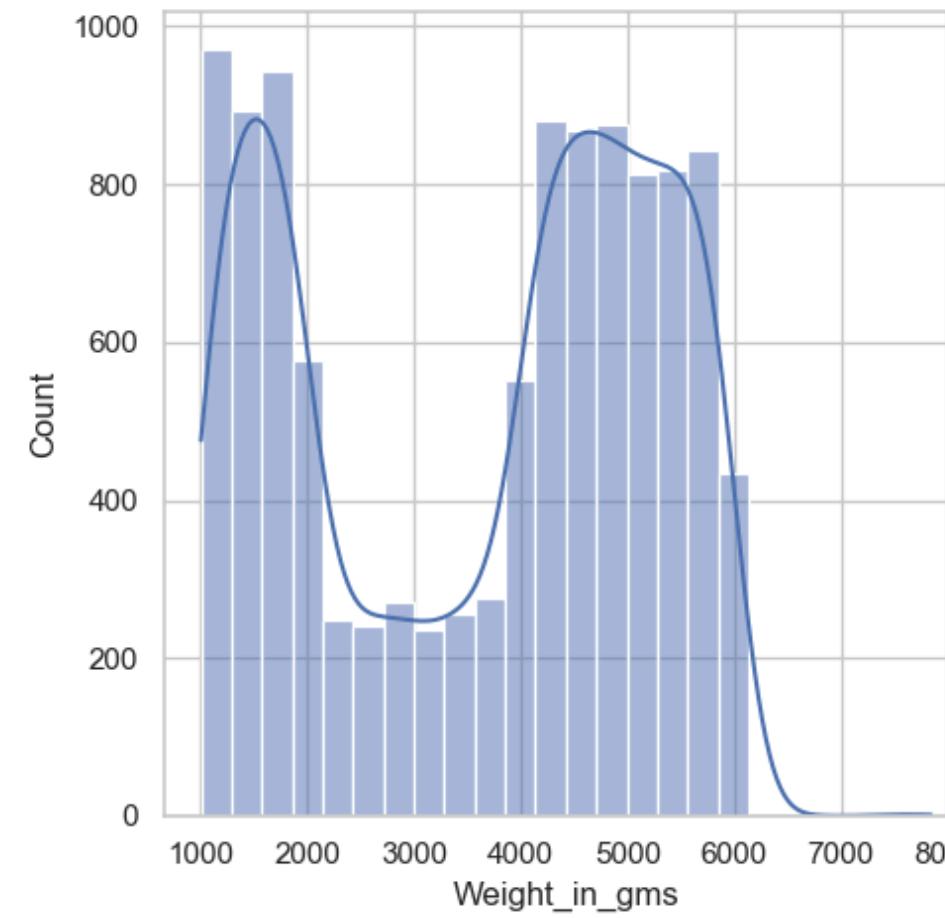
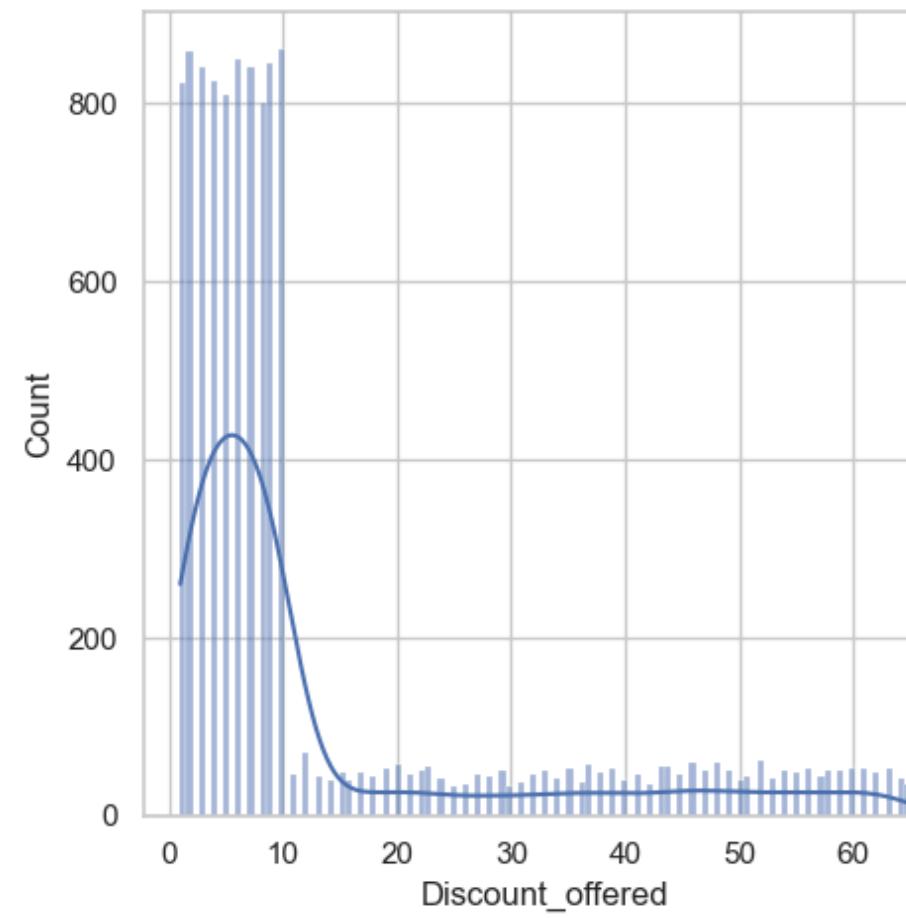
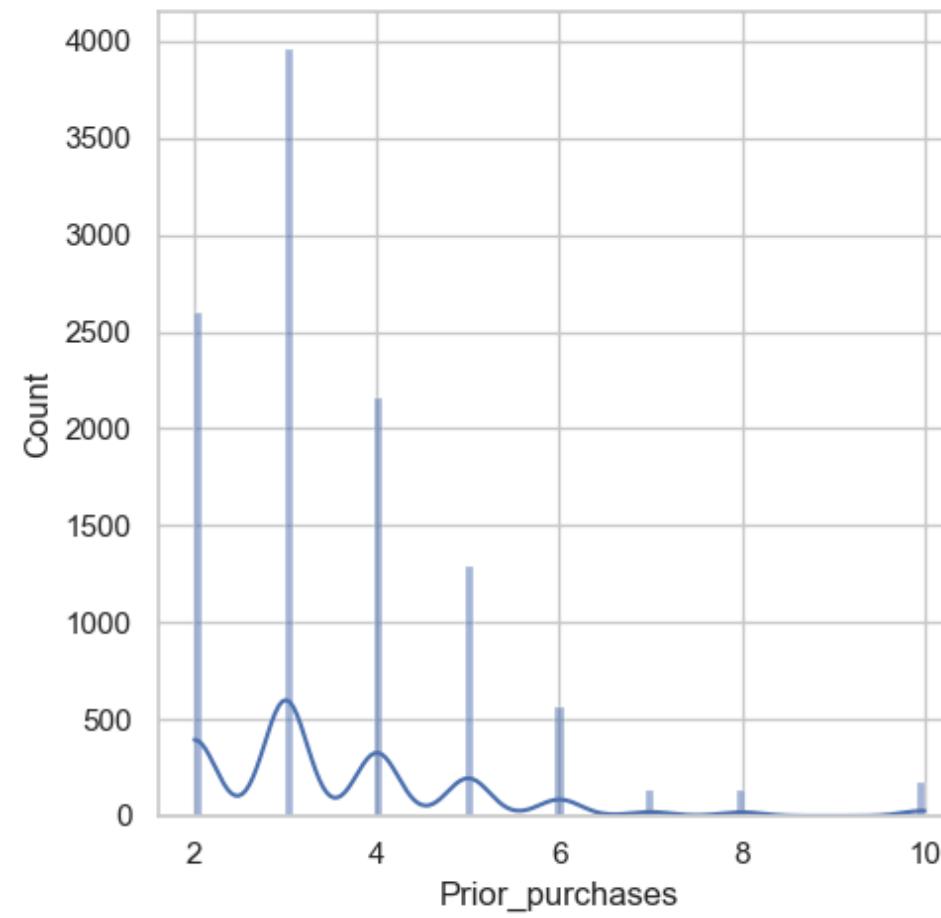
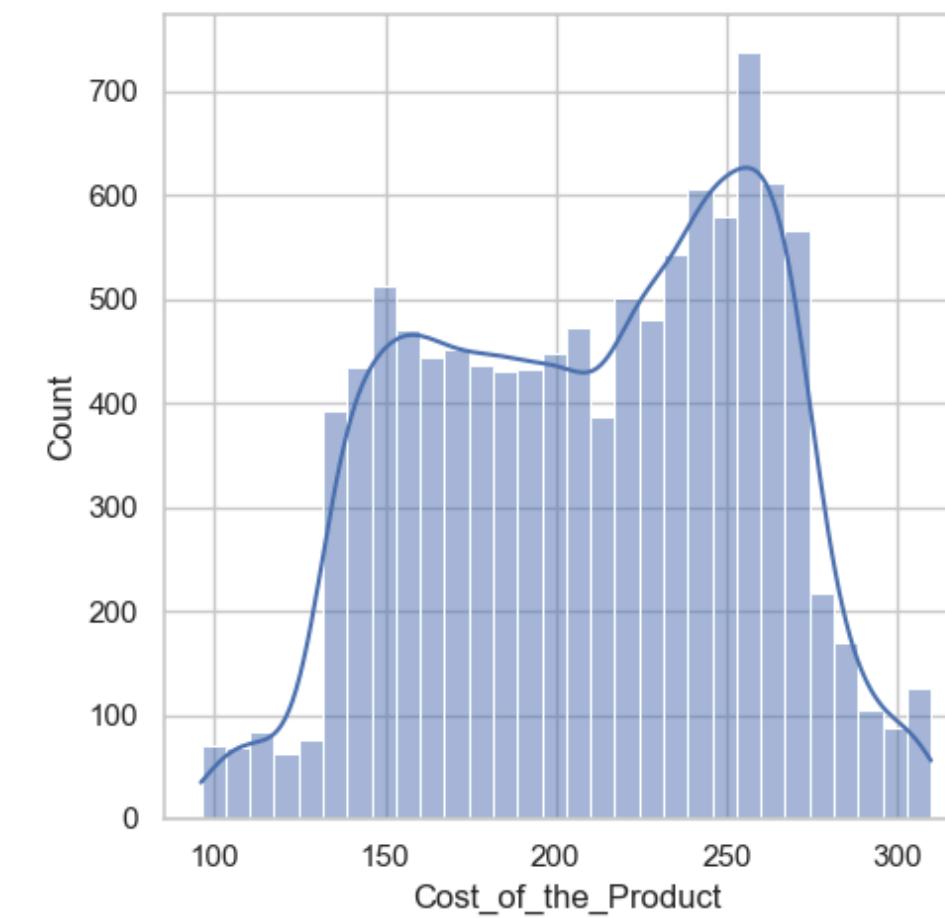
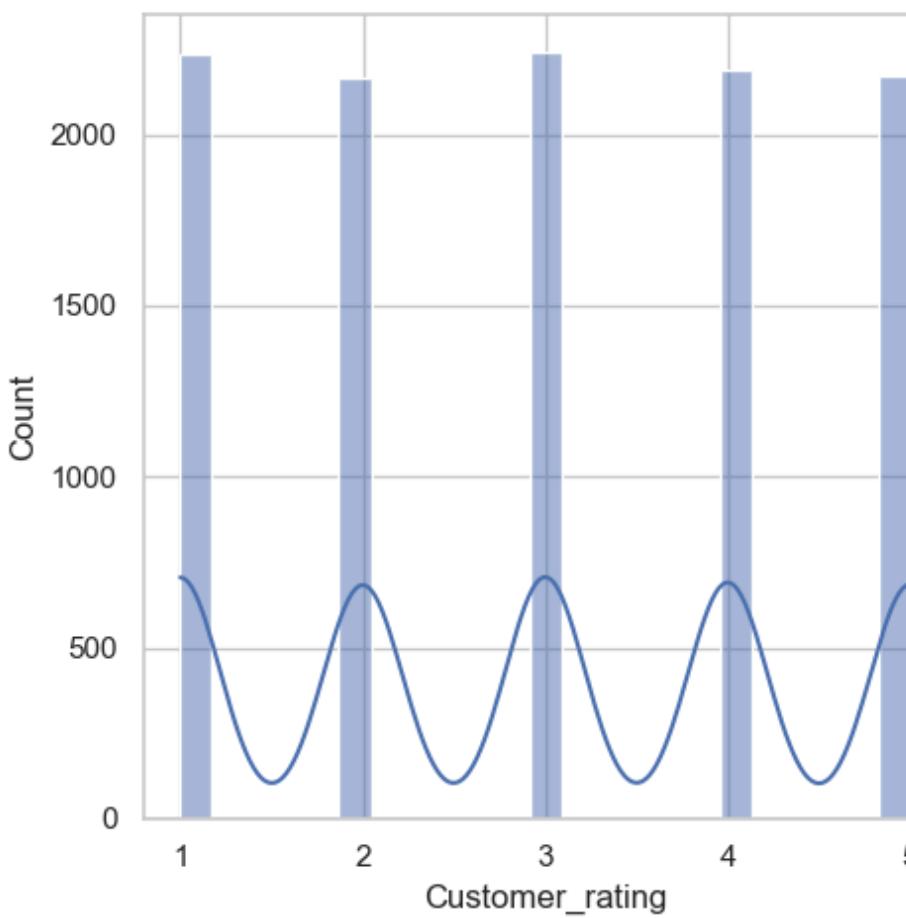
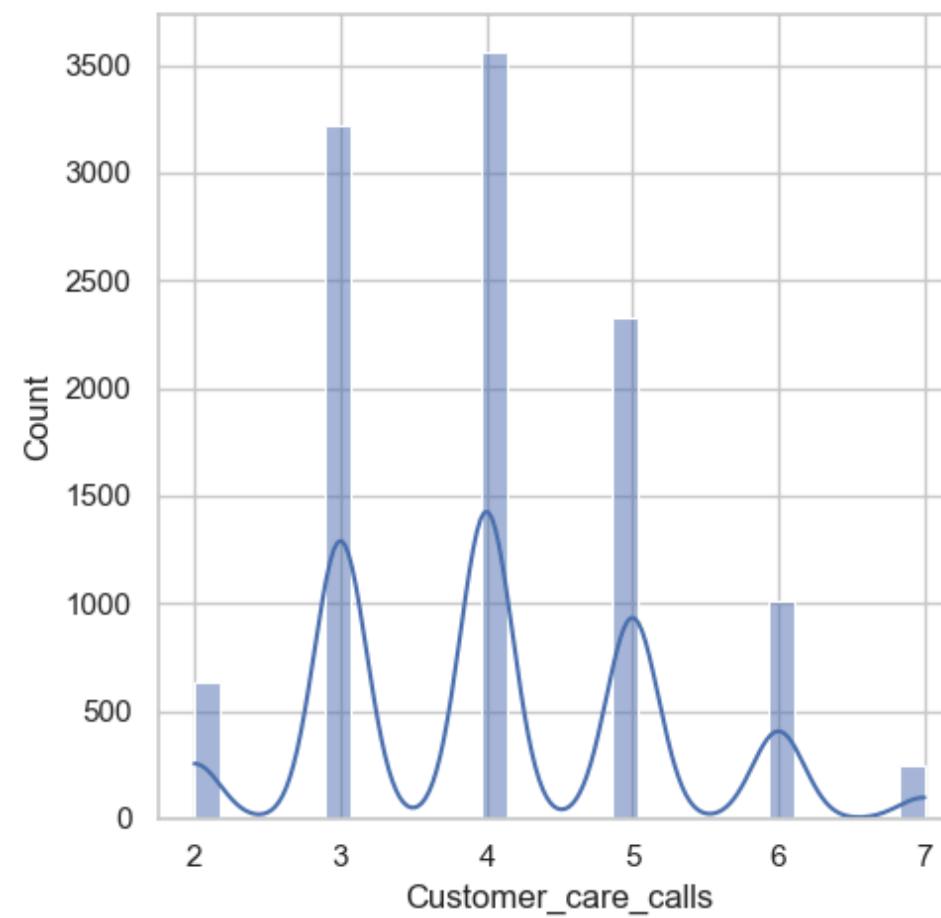
Warehouse F processes the highest volume of shipments, handling 5,545 records, making it a critical point for operational efficiency.

DATA INSIGHTS WHAT WE FOUND

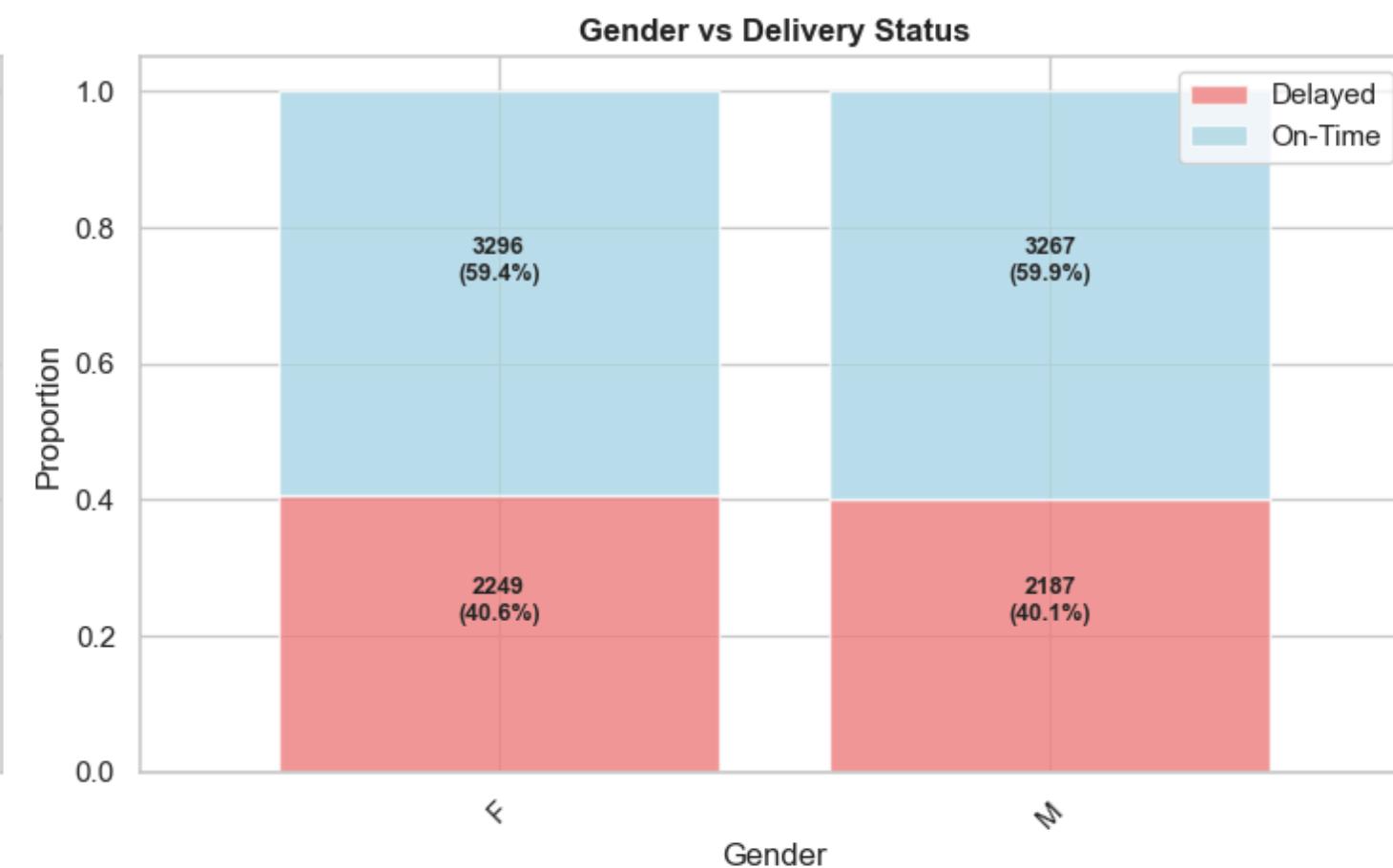
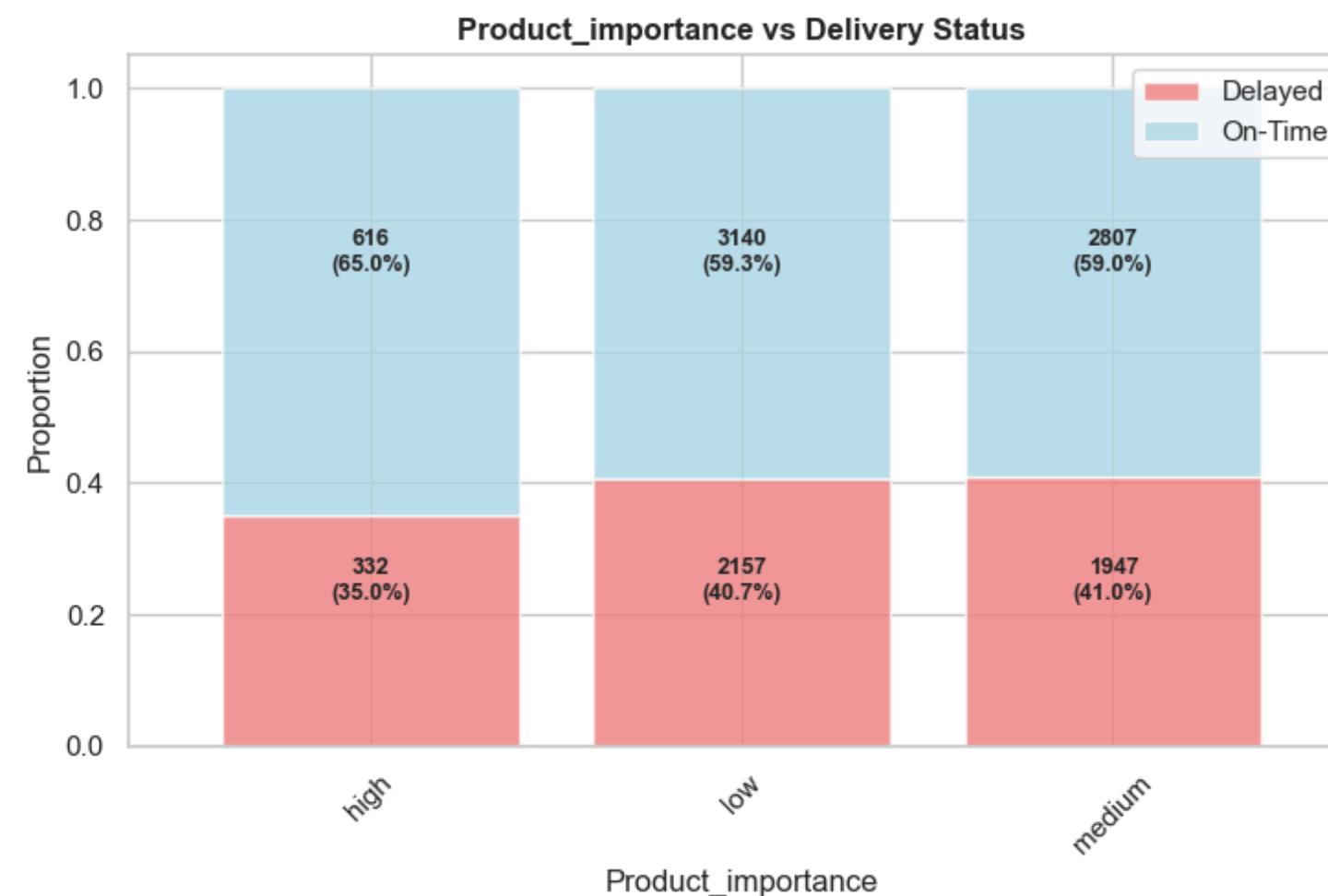
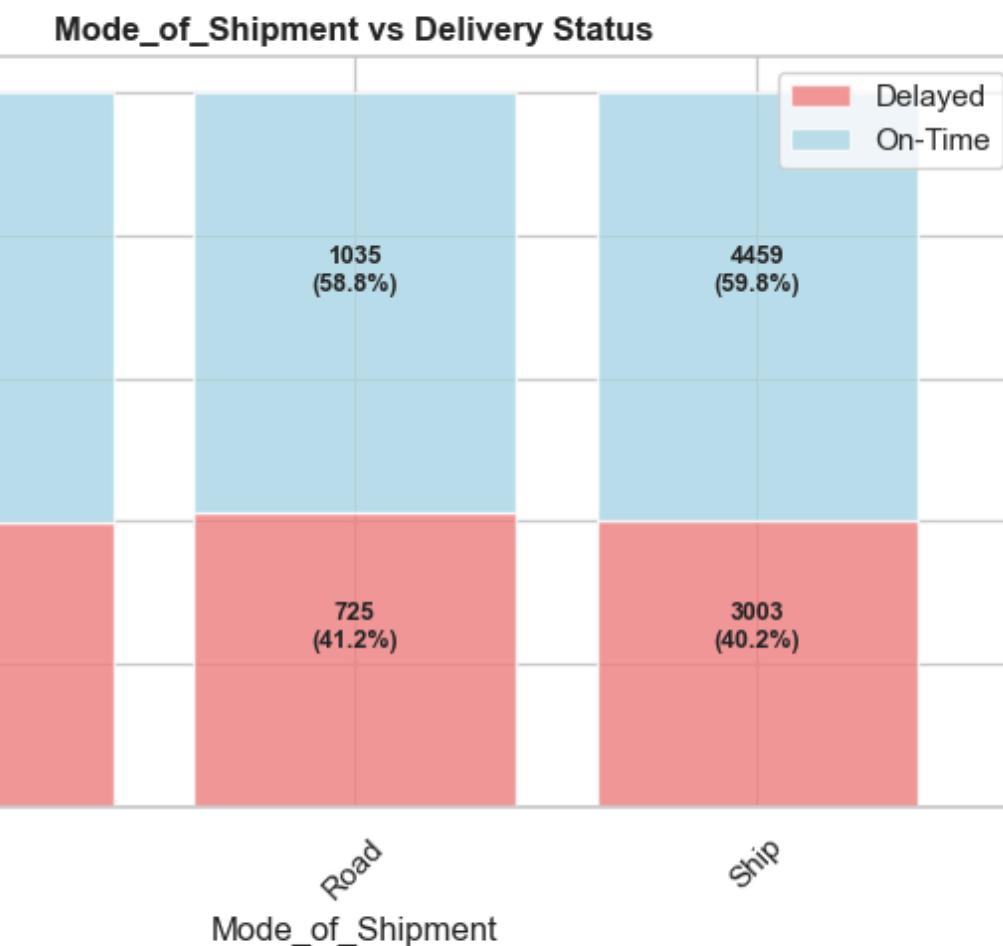
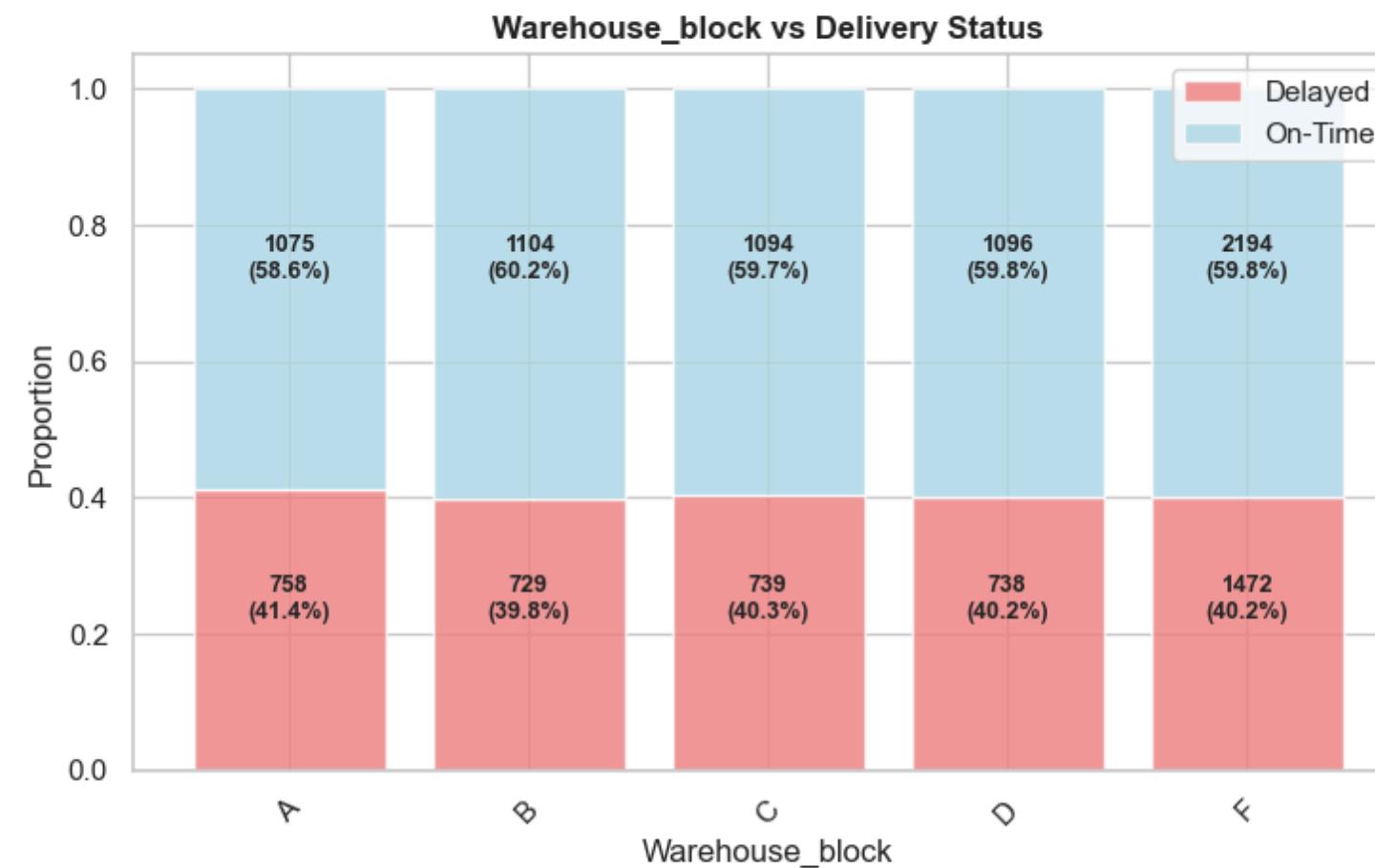
Categorical Variables Distribution

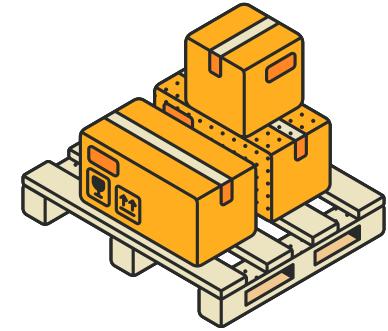


Numerical Variables Distribution



Categorical Features vs Target Variable





OUR LOGISTICS-DRIVEN APPROACH



Data Analysis

Thoroughly analyzed historical delivery data to identify critical delay patterns and root causes.



Data Preparation & Splitting

Cleaned and encoded logistics features, then split data into training (8,799) and testing (2,200) sets.



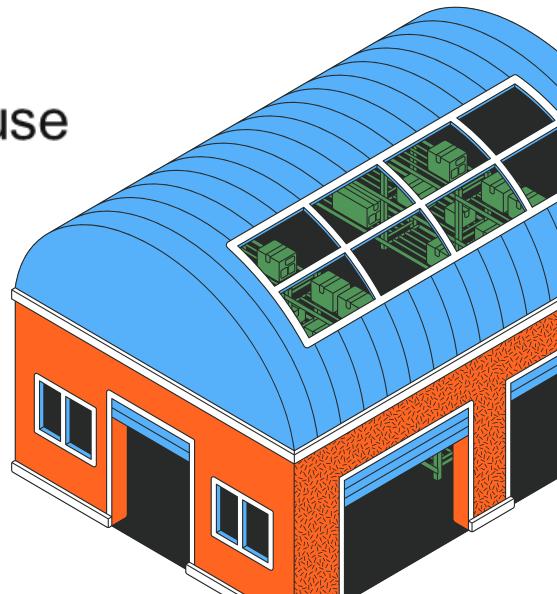
Predictive Model Building

Developed robust models to accurately forecast delivery outcomes (on-time or delayed).



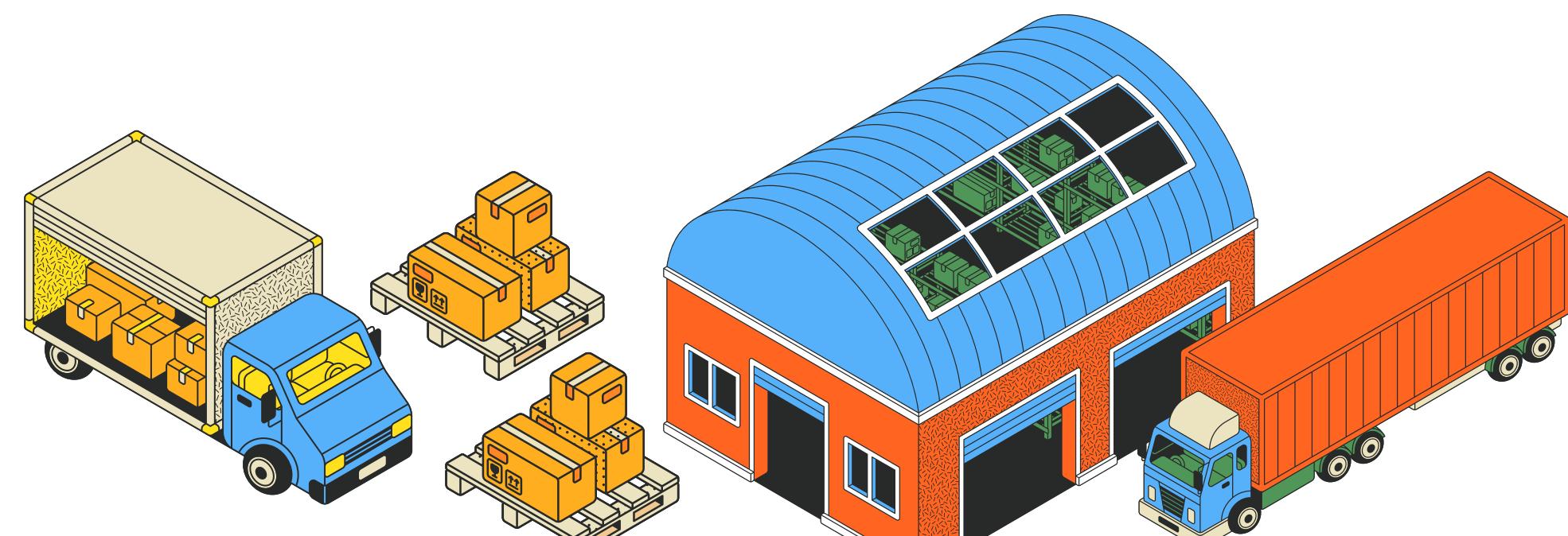
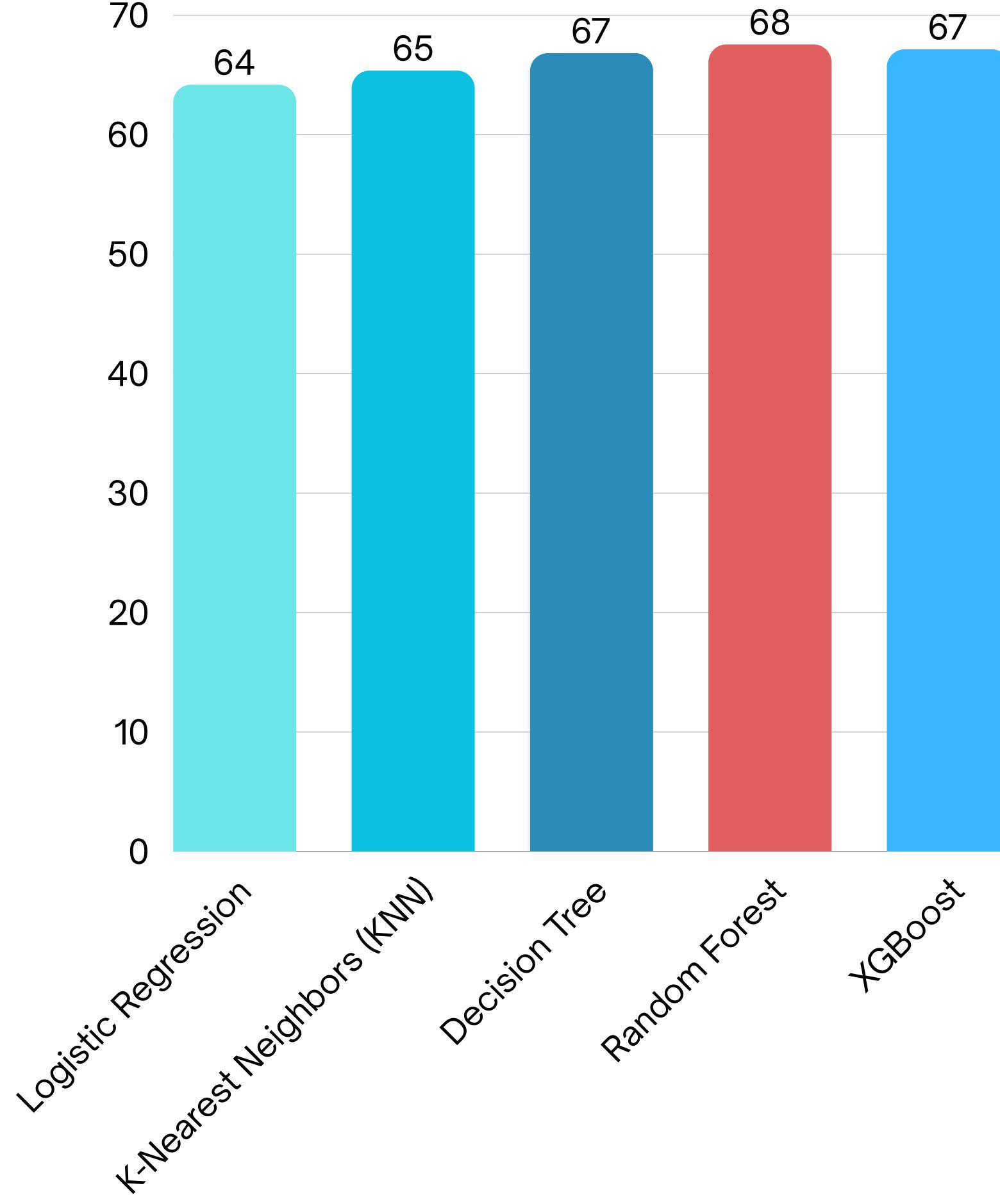
Logistics Decision Optimization

Applied model insights to guide optimized warehouse allocation and shipping choices.

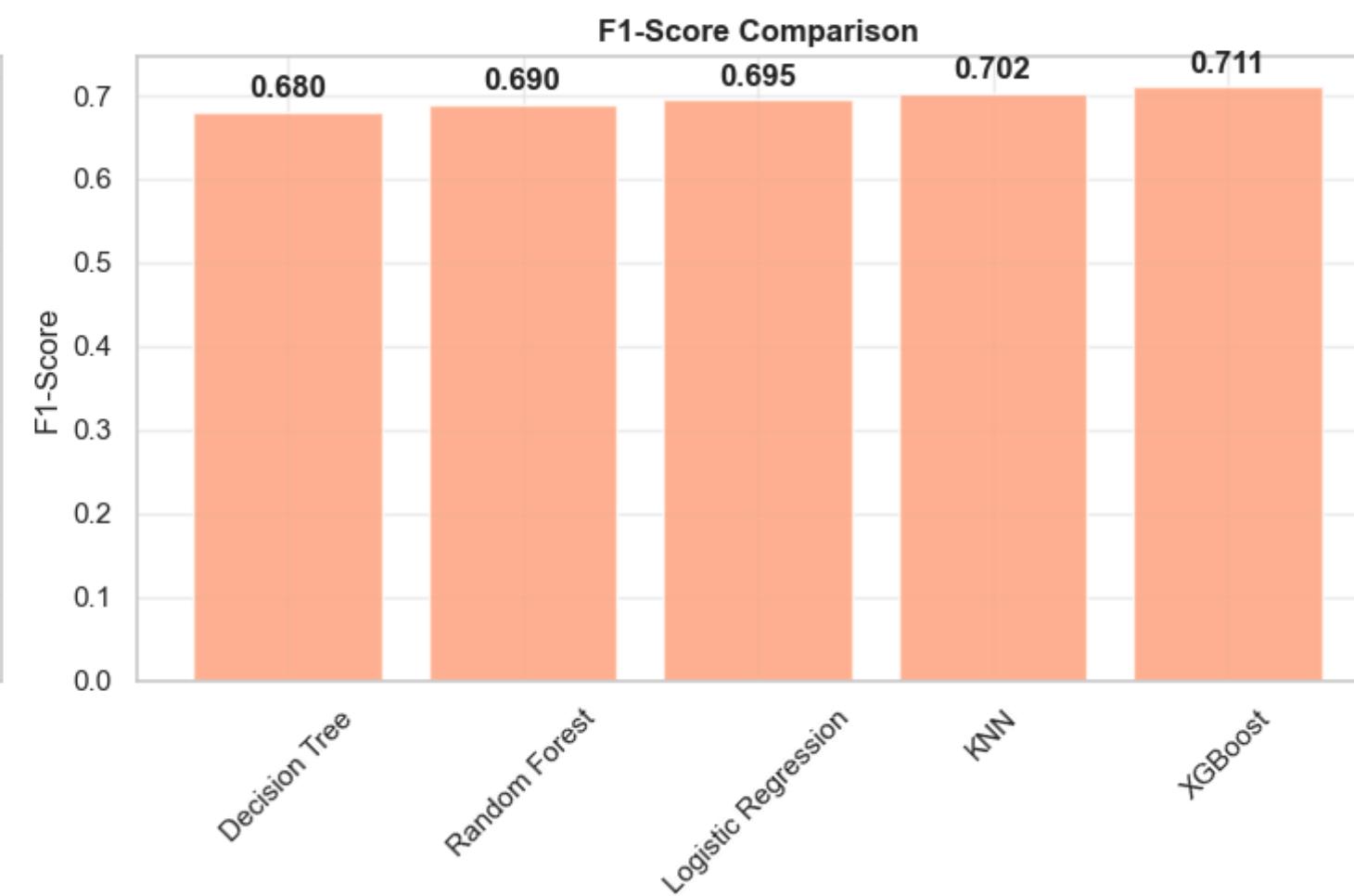
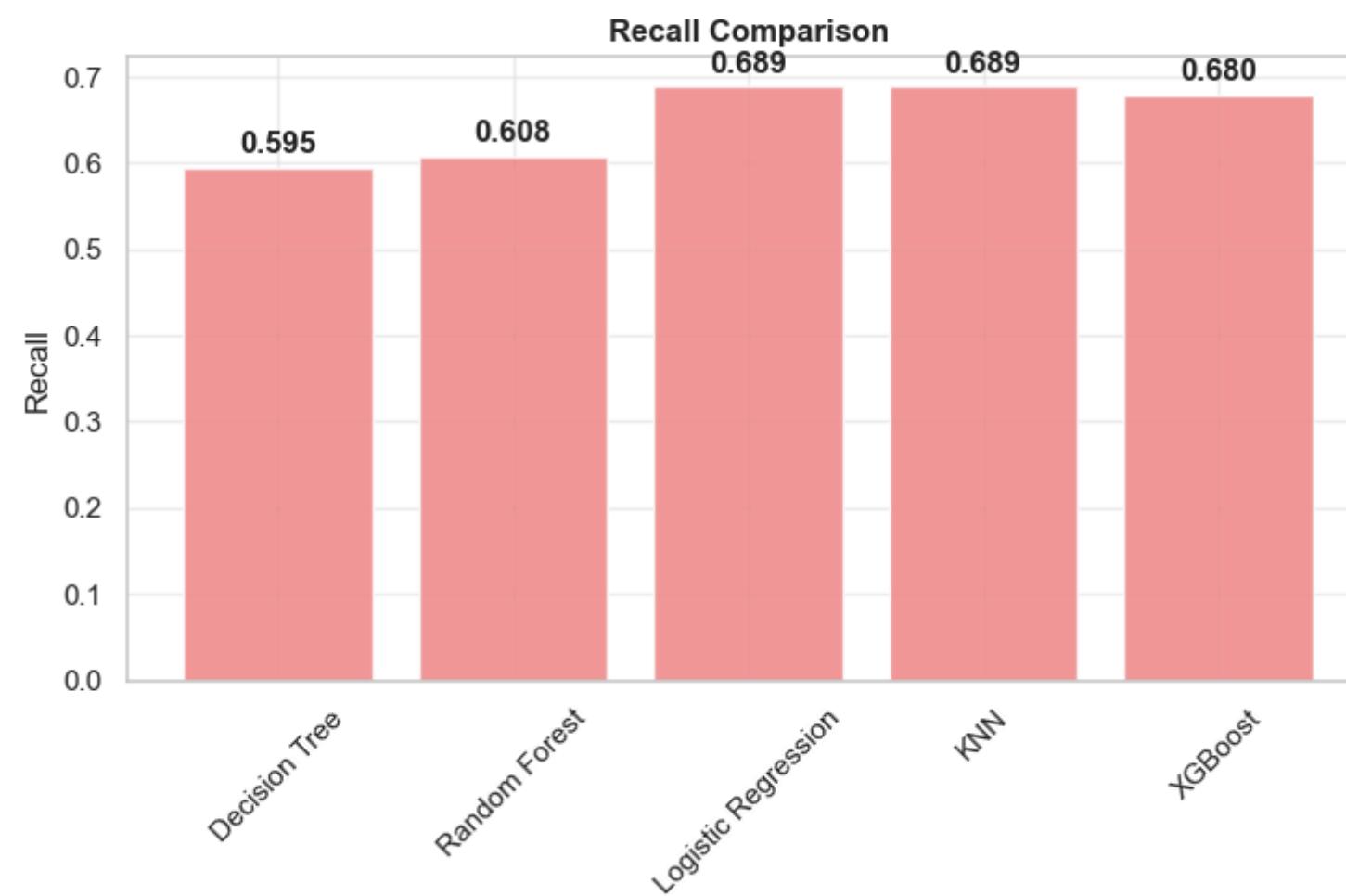
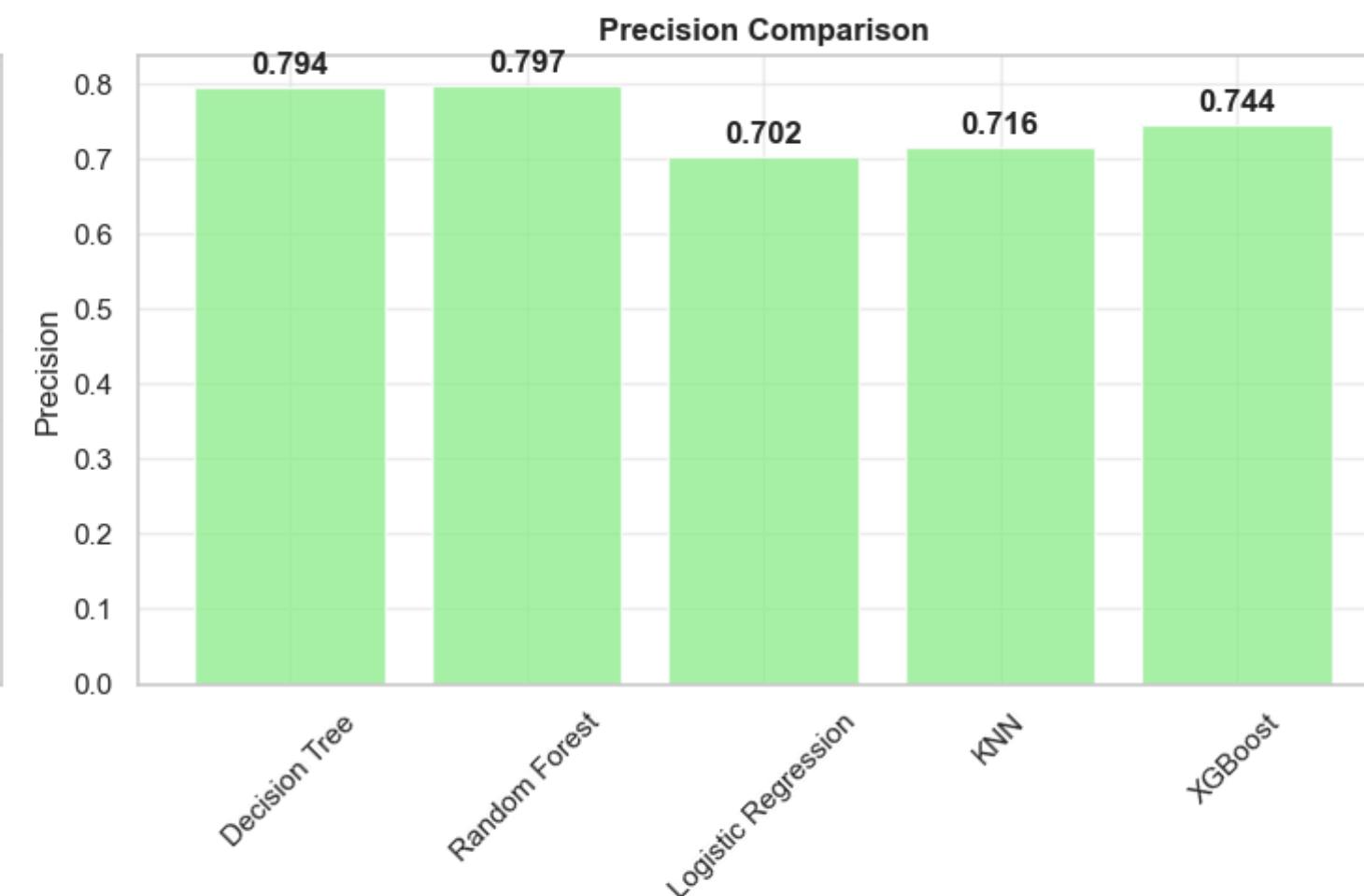
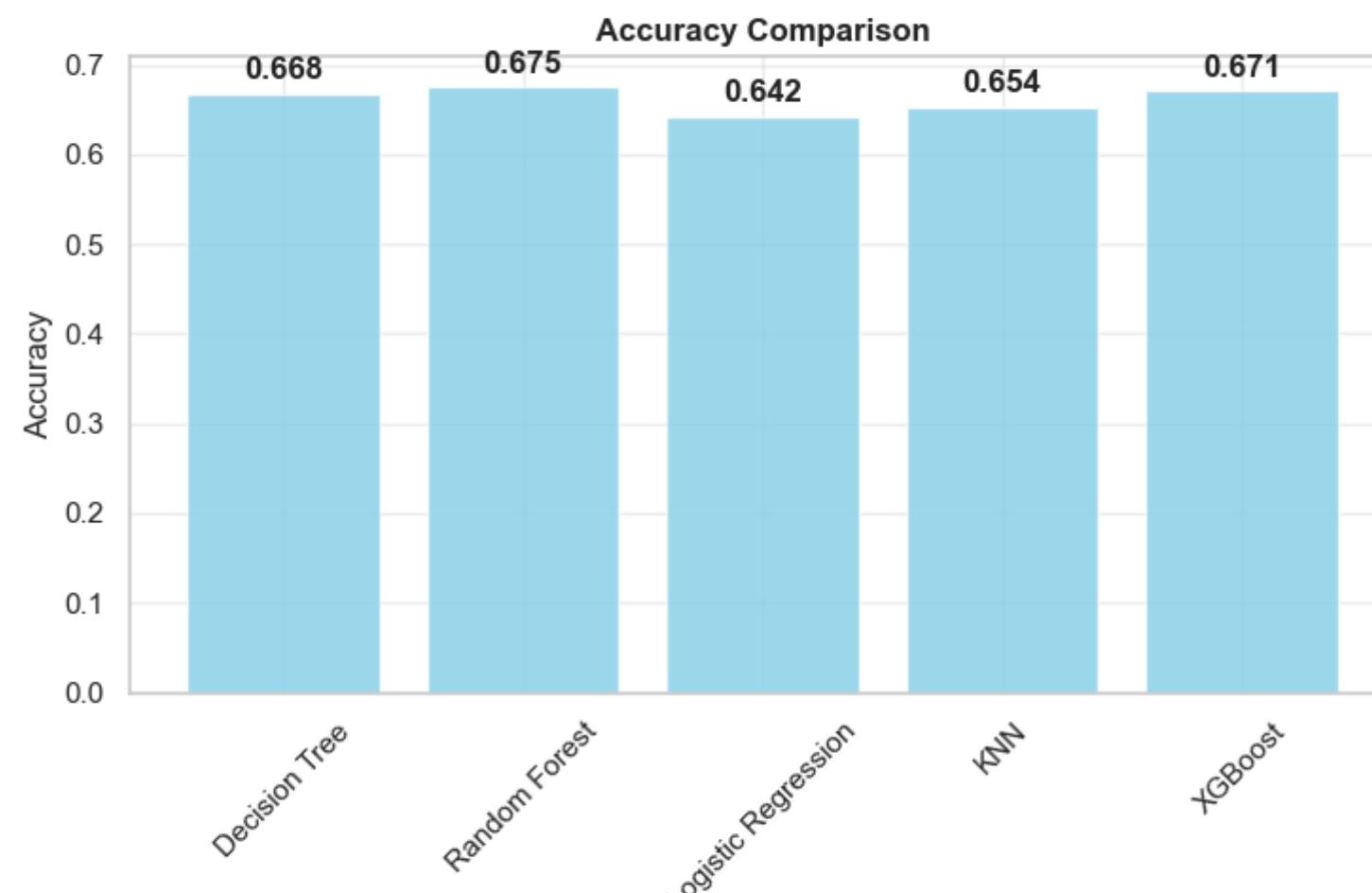


THE PREDICTIVE MODELS

- To accurately forecast delivery outcomes, we evaluated five distinct machine learning models: **Logistic Regression, K-Nearest Neighbors (KNN), Decision Tree, Random Forest, and XGBoost.**
- Among these, **the Random Forest** model demonstrated the highest predictive capability, achieving an accuracy of **68%**. These models are designed to predict whether a delivery will be on-time or delayed based on the input features we've identified.



Model Performance Comparison



KEY PREDICTORS OF DELIVERY SUCCESS

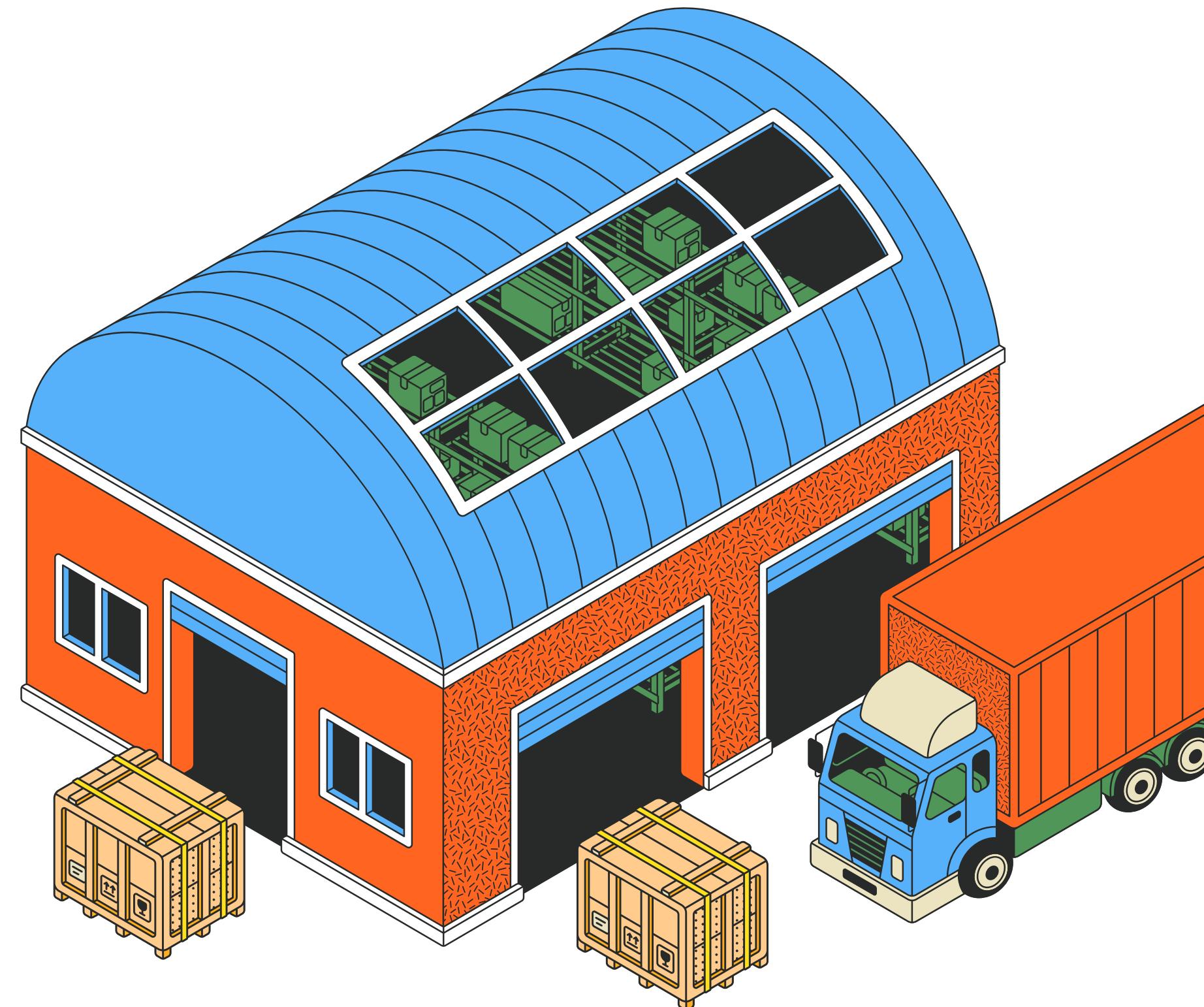
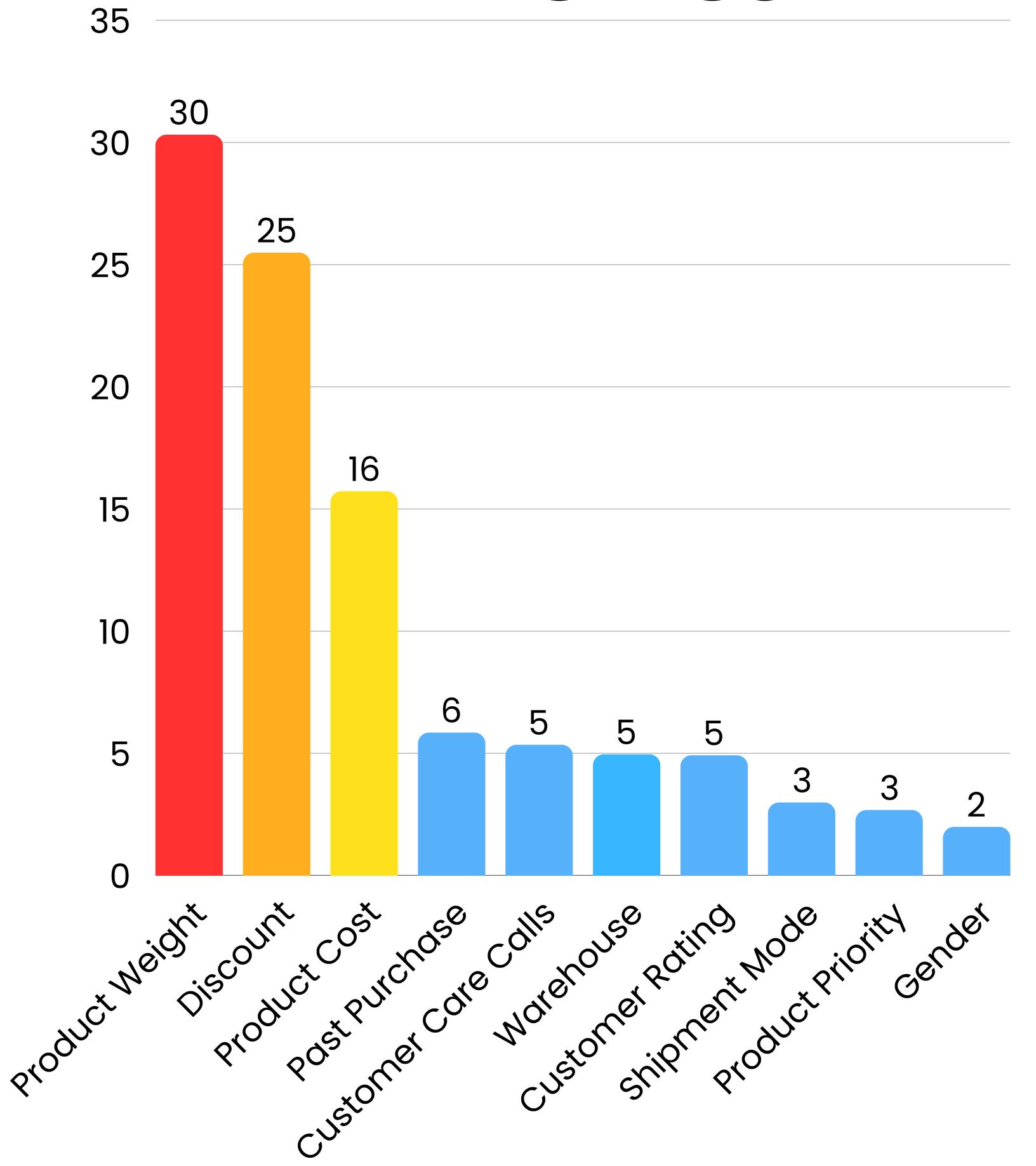
Understanding which factors most influence delivery timeliness is crucial for effective logistics optimization. Our predictive model highlights the following as top drivers of delivery outcomes:

- **Product Weight (30.3% Importance)**: Heavier products significantly increase the likelihood of delays, often requiring specialized handling or longer processing times.
- **Discount Offered (25.5% Importance)**: Shipments associated with discounts are statistically more prone to delays, possibly due to higher volume or different fulfillment workflows.
- **Cost of Product (15.7% Importance)**: Higher value items also contribute to delay risk, potentially linked to enhanced security protocols or specific delivery requirements.

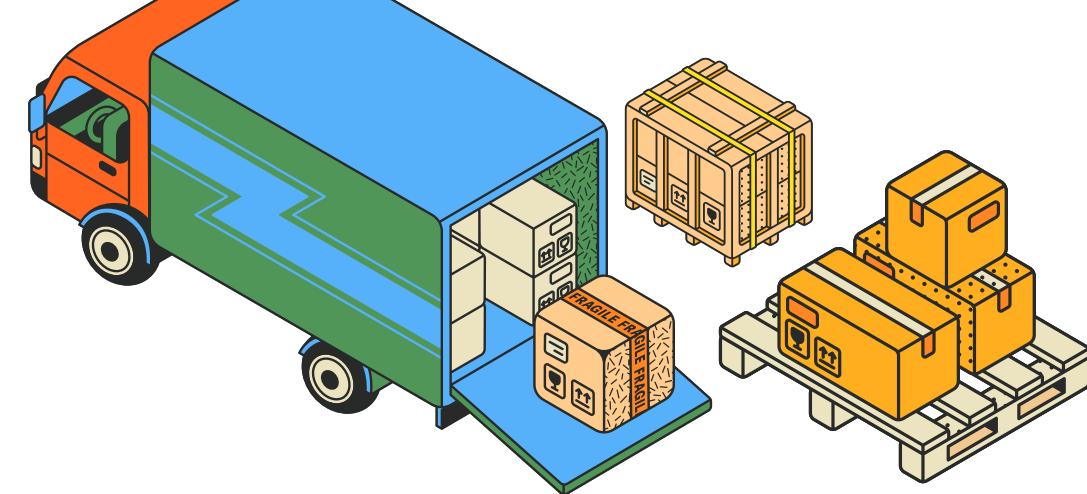
This insight confirms that heavy products and high discounts represent key risk factors that demand proactive management to ensure on-time delivery.



LET'S VISUALIZE THE FEATURE IMPORTANCE



HOW THE MODEL BRIDGES THE GAP



Our predictive model is designed to be a proactive tool, enabling timely interventions and ensuring customer satisfaction.



Predict Delivery Risks

Model identifies potential delays **before dispatch**, flagging high-risk shipments (e.g., heavy products from Warehouse F).



Enable Logistics Adjustments

Logistics teams proactively adjust **warehouse assignments** or switch **shipping modes** for flagged shipments.



Ensure On-Time Delivery

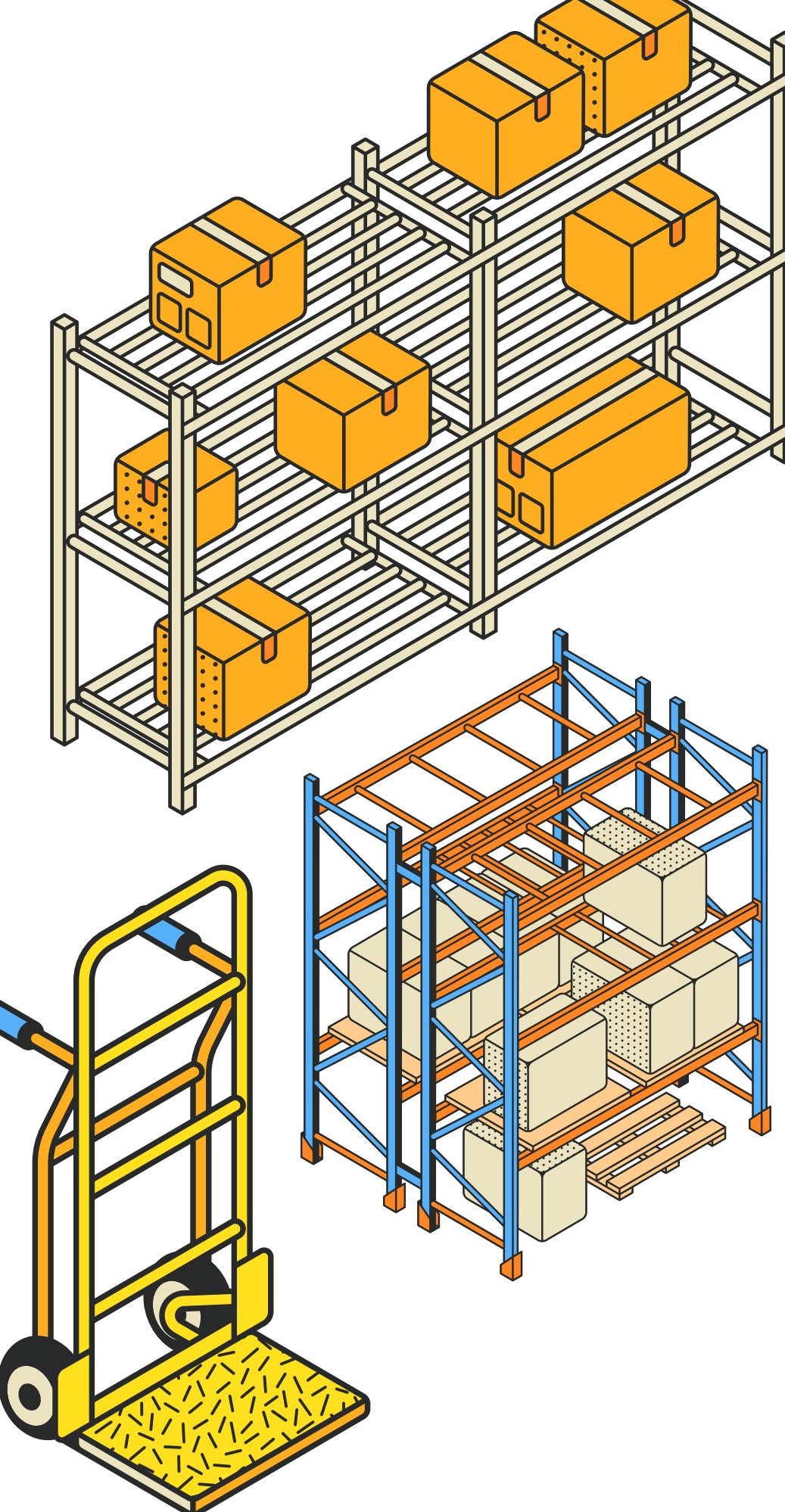
Proactive interventions result in **reliable, on-time service**, enhancing customer satisfaction and operational flow.

LOGISTICS BENEFITS

Fewer Delays: Optimized warehouse allocation and shipping strategies lead to significantly reduced delivery times.

Reduced Operational Costs: Decreased customer care calls and lower reshipping expenses directly impact the bottom line.

Improved Reliability & Satisfaction: Consistent on-time deliveries enhance brand reputation and build strong customer trust.



REAL-WORLD LOGISTICS INTEGRATION

Our powerful Random Forest model is designed for real-time applications, seamlessly integrating into operations to provide instant delivery predictions.

INPUT KEY DATA

Simply input critical shipment details like product weight, shipping mode, and discount offered.

PREDICTIVE INSIGHTS

Input critical shipment data like product weight, warehouse, and shipping mode to predict "On-Time" or "Delayed" delivery status.

PREDICTIVE INSIGHTS

Proactively adjust logistics plans, for example, by prioritizing faster shipping for high-risk, heavy items flagged for potential delays.





LOGISTICS RECOMMENDATIONS

Optimize warehouse operations for **heavy, high-demand products** to reduce processing time and minimize delays.

Adjust shipping modes based on **product weight and importance**, selecting faster or more secure options for critical deliveries.

Monitor **customer care calls** to flag potential delays early, enabling proactive communication and resolution.

FUTURE IMPROVEMENTS

Periodic Model Retraining

Continuously update the model with new delivery data to maintain high accuracy and adapt to evolving trends.

1

Predictive Time Windows

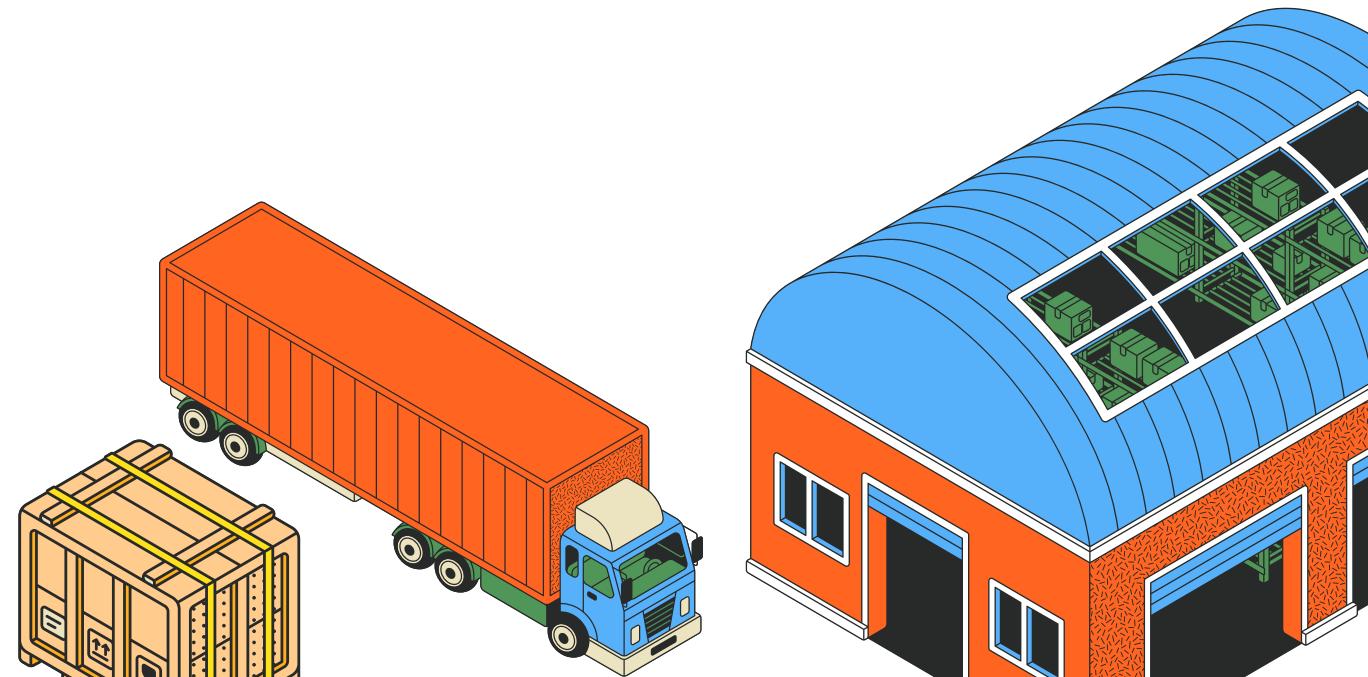
Expand the model's capability to forecast specific delivery time windows, enabling better planning and customer communication.

2

Enhanced Feature Integration

Incorporate external data sources such as real-time weather and traffic conditions for even more precise predictions.

3





CONCLUSION DELIVERING FUTURE SUCCESS

PREDICTIVE POWER

Our Random Forest model achieves **68% accuracy**, directly addressing the 40% e-commerce logistics delay challenge.

OPERATIONAL IMPACT

Expect streamlined operations, significantly **reduced delays**, and **enhanced customer trust**.



THANK YOU