Shape

Description automatically generated with medium confidence



**Essay / Assignment Title: Transforming Data into Insight: Building Interactive Dashboards and Stories with Tableau**

**Programme title: Visualization and Story Telling using Tableau**

**Name: Pagadala Abhishek Raj**

**Year: 2026**

# CONTENTS

[**CONTENTS**](#) **2**

[**ABSTRACT**](#) **5**

[**INTRODUCTION**](#) **6**

[**CHAPTER 1: DATA FOUNDATION AND ANALYTICAL PREPARATION**](#) **7**

[1.1 Dataset Overview](#_tm2iygz816pk) 7

[1.2 Calculated Fields: Building Analytical Intelligence](#_azt0r798qf7) 7

[1.2.1 On-Time Delivery Percentage](#_62bo7zxp90vm) 8

[1.2.2 KPI Status (Good / Average / Poor)](#_bhauakbqeat3) 8

[1.3 Binning Fields (Weight, Discount, Prior Purchases)](#_nduhhxblgsq4) 9

[**CHAPTER 2: EXCLUSIVE SUMMARY DASHBOARD – “THE PULSE OF LOGISTICS PERFORMANCE”**](#) **11**

[2.1 Purpose and Strategic Role](#_qufji1ieo1gk) 11

[2.2 Key Performance Indicator (KPI): On-Time Delivery Rate](#_dbvkhg4yvr1b) 11

[2.3 Delivery Status Distribution Visualization](#_tc91gvbjb199) 12

[2.4 Integrated Global Filters](#_9lfekfsq9ncg) 13

[2.5 Diagnostic Tooltip System – “Why Was This Delayed?”](#_j2qv3x5po9l) 13

[2.6 Story Interpretation Of Exclusive Summary Dashboard](#_x1du30tr4hft) 15

[**CHAPTER 3: CUSTOMER & PRICING IMPACT DASHBOARD – “THE HUMAN AND MONETARY SIDE OF DELAYS”**](#_7n0si9bhz6ci) **17**

[3.1 Purpose and Analytical Motivation](#_su7ogtozkjlg) 17

[3.2 Customer Rating vs On-Time Delivery](#_l6kgejcegwxj) 17

[3.3 Discount Offered vs On-Time Delivery](#_vyidu8mb05vl) 18

[3.4 Prior Purchases vs On-Time Delivery](#_f0405048c9uh) 20

[3.5 Integrated Insights from the Dashboard](#_aszyq743ktif) 21

[3.6 Managerial Implications](#_oww6bnw2icj5) 21

[3.7 Visualization Design Considerations](#_bv139fgy031g) 21

[3.8 Story Interpretation Of Customer And Impact Dashboard](#_94t84qfh9hx3) 22

[**CHAPTER 4: OPERATIONS & COST ANALYSIS DASHBOARD – “WHERE LOGISTICS MEETS PHYSICAL AND ECONOMIC CONSTRAINTS”**](#_r0t5268w8d0f) **24**

[4.1 Purpose and Analytical Motivation](#_51dkuu9mgyvn) 24

[4.2 Cost of Product vs On-Time Delivery](#_a0fw4qw0mtyv) 24

[4.3 Weight Band vs On-Time Delivery Rate](#_5grbwrz7n5hh) 26

[4.4 Integrated Operational Insights](#_ev162ew06cah) 28

[4.5 Managerial Implications](#_qsgahdfc0rfd) 29

[4.6 Visualization Design Considerations](#_cj1dtawxyopd) 29

[4.7 Contribution to Analytical Narrative](#_aaa2cprim6us) 29

[4.8 Story Interpretation Of Operations and Cost Analysis Dashboard](#_dp1b09jhpgy6) 30

[**CHAPTER 5: OPERATIONAL ANALYTICS DASHBOARD – “INSIDE THE LOGISTICS ENGINE ROOM”**](#_qm1kbnakodoz) **32**

[5.1 Purpose and Analytical Motivation](#_vtd1xfg668o3) 32

[5.2 Delivery Performance by Product Importance](#_mjkn9huxkczk) 32

[5.3 Delivery Performance by Warehouse Block](#_ipazl2qdy5xr) 33

[5.4 Delivery Performance by Mode of Shipment](#_ngl682yhn66h) 35

[5.5 Customer Care Calls vs Delivery Performance](#_vl8aojeoczvk) 36

[5.6 Managerial Implications](#_c2eiilk22l9h) 38

[5.7 Contribution to the Overall Analytical Framework](#_ro0l7qw6bsrx) 38

[5.8 Story Interpretation Of Operational Analytics Dashboard](#_av9z3zvviak4) 39

[**CHAPTER 6: MACHINE LEARNING – BASED DELIVERY RISK PREDICTION DASHBOARD**](#_aq6shfpbr6v1) **41**

[6.1 Purpose of the Dashboard](#_74pjx4sk5mnv) 41

[6.2 Model Design and Feature Selection](#_qkpz9tr1frxx) 41

[6.3 Integration with Tableau Using TabPy](#_7c0b1380wvq) 42

[6.4 Logistic Regression Visualization](#_ptfbqb96qv3s) 42

[6.5 Risk Classification Logic](#_7yh73526dnrm) 43

[6.6 Business Value and Practical Application](#_t2fi8rn39cfz) 44

[6.7 Story Interpretation Of Logistic Prediction](#_rgnct7bmou5u) 44

[**CHAPTER 7: Real-Time Industry Application of Predictive Logistics Analytics**](#_9cv3m6ho2e5n) **46**

[7.1 Motivation for Real-Time Deployment](#_55cej5pg5ad6) 46

[7.2 Industry Practices](#_yl8p4oxi6def) 46

[7.3 Alignment with the Proposed System](#_ufm583jtzakk) 46

[7.4 Practical Benefits](#_alm6a85obnsy) 47

[7.5 Summary](#_p5l7j44vnd3r) 47

[**CONCLUSION**](#_9ycz5g5ka46p) **48**

[**BIBLIOGRAPHY**](#) **49**

[**SOURCE CODE**](#_abv7kbd2s9e) **51**

[Google Drive](#_rpylysoar73y) 51

[Github Link](#_jahjpijvpci9) 51

**Statement of compliance with academic ethics and the avoidance of plagiarism**

I honestly declare that this dissertation is entirely my own work and none of its part has been copied from printed or electronic sources, translated from foreign sources and reproduced from essays of other researchers or students. Wherever I have been based on ideas or other people texts I clearly declare it through the good use of references following academic ethics.

(In the case that is proved that part of the essay does not constitute an original work, but a copy of an already published essay or from another source, the student will be expelled permanently from the program).

Name and Surname (Capital letters):

.............Abhishek Raj Pagadala.........................................................

Date: .....30/01/2026.........

# ABSTRACT

Delivery on time, in the context of modern logistics operations, has become an increasingly critical determinant factor for customer satisfaction, operational efficiency, and organizational competitiveness. It remains an analytically challenging task, however, to understand what causes these delays on many different levels. Applying Interactive Data Visualization by Tableau, this project explores key drivers of delivery performance within a large logistics dataset-analyzes and communicates the results.

It develops multiple analytical dashboards, calculated performance metrics, and interactive features such as global filters, tooltips, and dashboard actions that identify key relationships existing between delivery status and operational, financial, and customer-related variables. The dashboards will cater to high-level performance monitoring and granular diagnostic capabilities that enable various stakeholders to detect inefficiencies, optimize resource allocation, and improve service quality.

It presents just how data visualization turns raw transactional data into strategic knowledge and actionable business intelligence.

# INTRODUCTION

Logistical processes interact and intersect at the point of customer demand, inventory management, transportation networks, and operational activities. Even small problems at these interrelated processes can create problems of delivery delays, escalating expenses, and eroding customer confidence.

Conventional reporting tools allow the reporting of performance data in tables, but such tables do not facilitate the exploration of data variables or the discovery of hidden patterns. Data visualization tools like Tableau eliminate these shortcomings by allowing data exploration.

Tableau has also been widely used in academic research for analytical visualization and multi-criteria decision support studies (Kose and Yildiz, 2024).

The project aimed at achieving the following:

* In order to measure the performance of the delivery process with a key performance indicator (KPI).
* Identifying customer, price, and operational variables affecting delivery time.
* To provide insightful dashboards to facilitate exploratory and explanatory analytics.
* To provide interactivity in managerial decision support.

Supply chain analytics platforms such as Tableau are increasingly adopted to improve operational visibility and support data-driven decision making (Almatarneh, Al-Daoud and Alshurideh, 2023).

The rest of the report will feature the design, framework, establishment of the dashboard, as well as the findings of the study.

# CHAPTER 1: DATA FOUNDATION AND ANALYTICAL PREPARATION

## 

## 1.1 Dataset Overview

The dataset contains approximately 10,000 shipment records, each describing a completed delivery transaction. Each record includes information related to product attributes, customer interaction, warehouse operations, shipment logistics, and delivery outcome.

The dataset used in this study was obtained from an open-source logistics and customer analytics repository hosted on Kaggle [Kaggle, 2023].

Key variables include:

* Delivery Status
* Cost of Product
* Weight Band
* Product Importance
* Warehouse Block
* Mode of Shipment
* Customer Rating
* Customer Care Calls
* Discount Offered
* Prior Purchases
* Delay Reason

## 1.2 Calculated Fields: Building Analytical Intelligence

To enable advanced analysis, several calculated fields were created in Tableau.

### 1.2.1 On-Time Delivery Percentage

**Purpose:**To measure the proportion of shipments delivered on time.

**Logic:**Assigns value 1 for "On Time" and 0 for "Delayed", then computes the average.

**Usefulness:**

* Serves as the primary KPI across all dashboards refer figure 1.
* Enables consistent comparison across different categories such as weight band, warehouse block, or shipment mode.

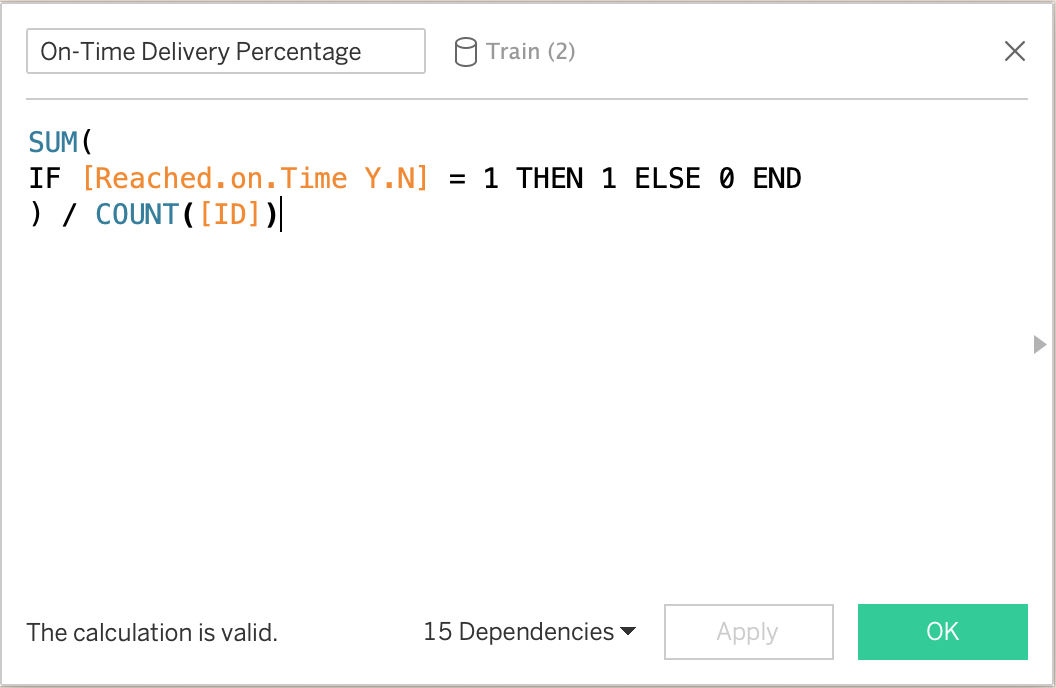


Figure 1 : Calculated Field Reached On Time Y.N

### 1.2.2 KPI Status (Good / Average / Poor)

**Purpose:**To classify delivery performance into qualitative categories. Refer figure 2

**Logic:**Based on threshold ranges of the on-time delivery percentage.

**Usefulness:**

* Enables visual performance segmentation.
* Supports conditional formatting and color encoding.
* Improves interpretability for non-technical users.



Figure 2 : Calculated Field KPI Status

## 1.3 Binning Fields (Weight, Discount, Prior Purchases)

**Purpose:**To group continuous variables into meaningful ranges.

**Usefulness:**

* Simplifies visualization.
* Enables comparative analysis between operational categories.
* Enhances clarity of bar and trend charts.

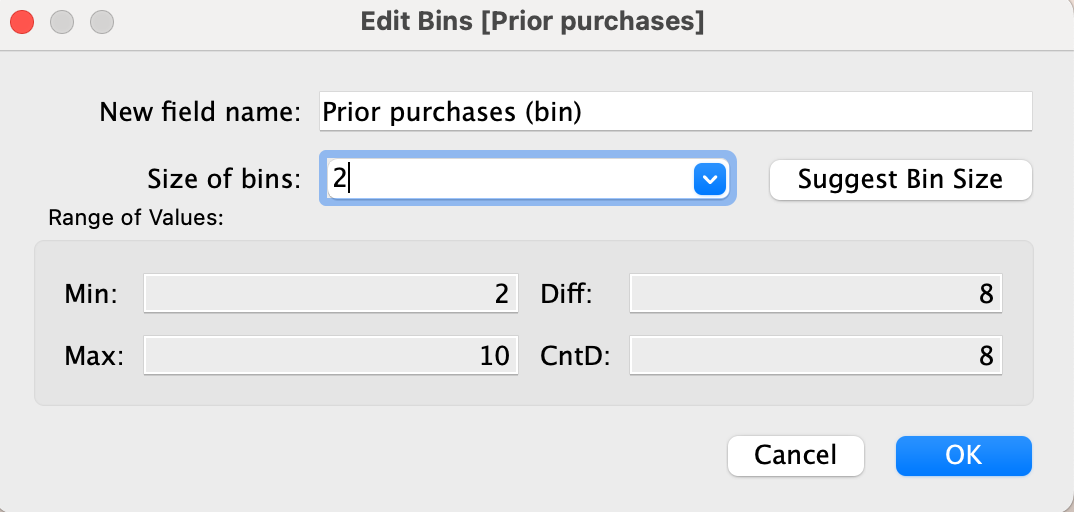


Figure 3 : Prior Purchases Bins

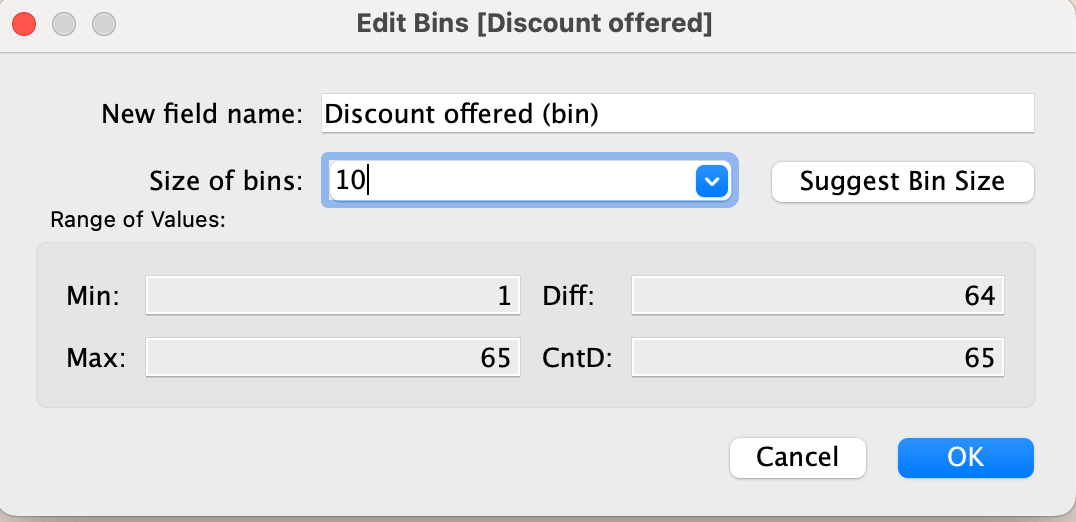


Figure 4 : Discount Offered Bins

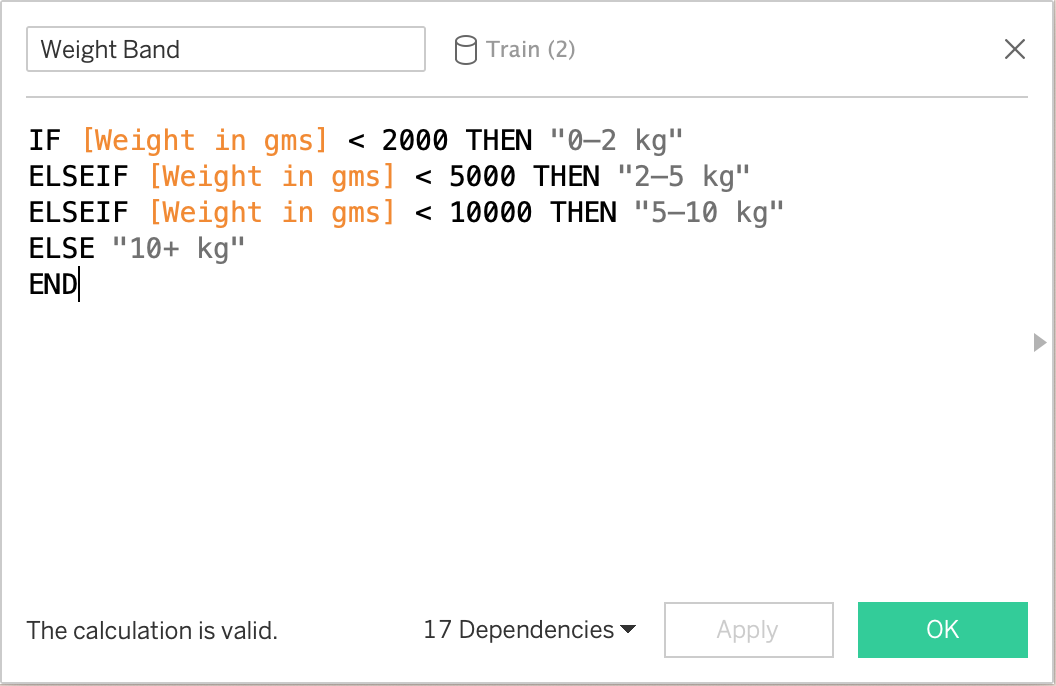


Figure 5 : Calculated Field Weight Band

# CHAPTER 2: EXCLUSIVE SUMMARY DASHBOARD – “THE PULSE OF LOGISTICS PERFORMANCE”

## 2.1 Purpose and Strategic Role

The Executive Summary Dashboard represents the interface of central monitoring for the entire analytical system developed in the project. Its main goal is the provision of a concise, high-impact view of delivery performance that allows for quick identification of systemic inefficiencies.

In operational contexts, senior managers and logistics coordinators have to refer directly to key indicators without going through several detailed reports. This dashboard fulfills that requirement of consolidating critical performance metrics into a single, visually intuitive view.

Interactive dashboards, calculated fields, filters, and dashboard actions were implemented using Tableau Desktop and Tableau Server functionalities (Tableau Software, 2024).

The design philosophy emphasizes:

* Clarity over complexity
* Immediate understandability
* Minimal cognitive load
* Decision-oriented presentation

Thus, the dashboard fulfills the purpose of both the performance thermometer and early warning system in case delivery failures occur.

## 2.2 Key Performance Indicator (KPI): On-Time Delivery Rate

On-Time Delivery Rate is calculated as below:

On-Time Delivery Rate=Total DeliveriesNumber of On-Time Deliveries​×100

This indicator is highlighted through a huge numeric KPI card to effectively portray the efficiency of the system directly.

In the analyzed data, the KPI is: 59.67%. Refer figure 6.



Figure 6 : On-Time Delivery Rate

It can be noted that nearly four out of every ten deliveries experience a failure in the expected time of delivery.

The prominence of this measure is sure to guarantee that:

* Performance deviations are immediately noticeable
* Trend changes can be quickly tracked over time
* Organizational benchmarks can be monitored consistently

## 2.3 Delivery Status Distribution Visualization

The dashboard includes a bar chart illustrating the proportional distribution of:

* On-Time deliveries
* Delayed deliveries

This visualization complements the KPI by providing structural context to the percentage value.

Where the KPI answers: *“How well are we performing overall?”*

The distribution chart answers: *“How is performance divided between success and failure?”*

This dual representation prevents misinterpretation and supports evidence-based evaluation.

## 2.4 Integrated Global Filters

The Executive Summary Dashboard incorporates several global filters, including:

* Delivery Status
* Cost of Product
* Weight Band
* Product Importance
* Customer Care Calls
* Mode of Shipment
* Warehouse Block

These filters transform the dashboard from a static reporting tool into a dynamic analytical environment.

Users can:

* Isolate specific operational scenarios
* Compare performance across different product categories
* Identify high-risk delivery profiles
* Investigate operational bottlenecks interactively

This capability enables both strategic oversight and exploratory investigation from the same interface.

## 2.5 Diagnostic Tooltip System – “Why Was This Delayed?”

One significant addition to the dashboard is the addition of sophisticated tooltip analytics features.

When users hover over the “Delayed” category, the system will display:

* Average number of customer care calls
* Average Product Cost
* Most frequent shipment weight band
* Dominant block of warehouses and delays

"This design enables descriptive visualizations to be used for diagnosis and to generate diagnostic

Instead of being aware of the delays, the users can understand: Under what operational conditions are failures most frequently encountered?

This shortens substantially the interval between finding the problem and conducting root cause analysis.

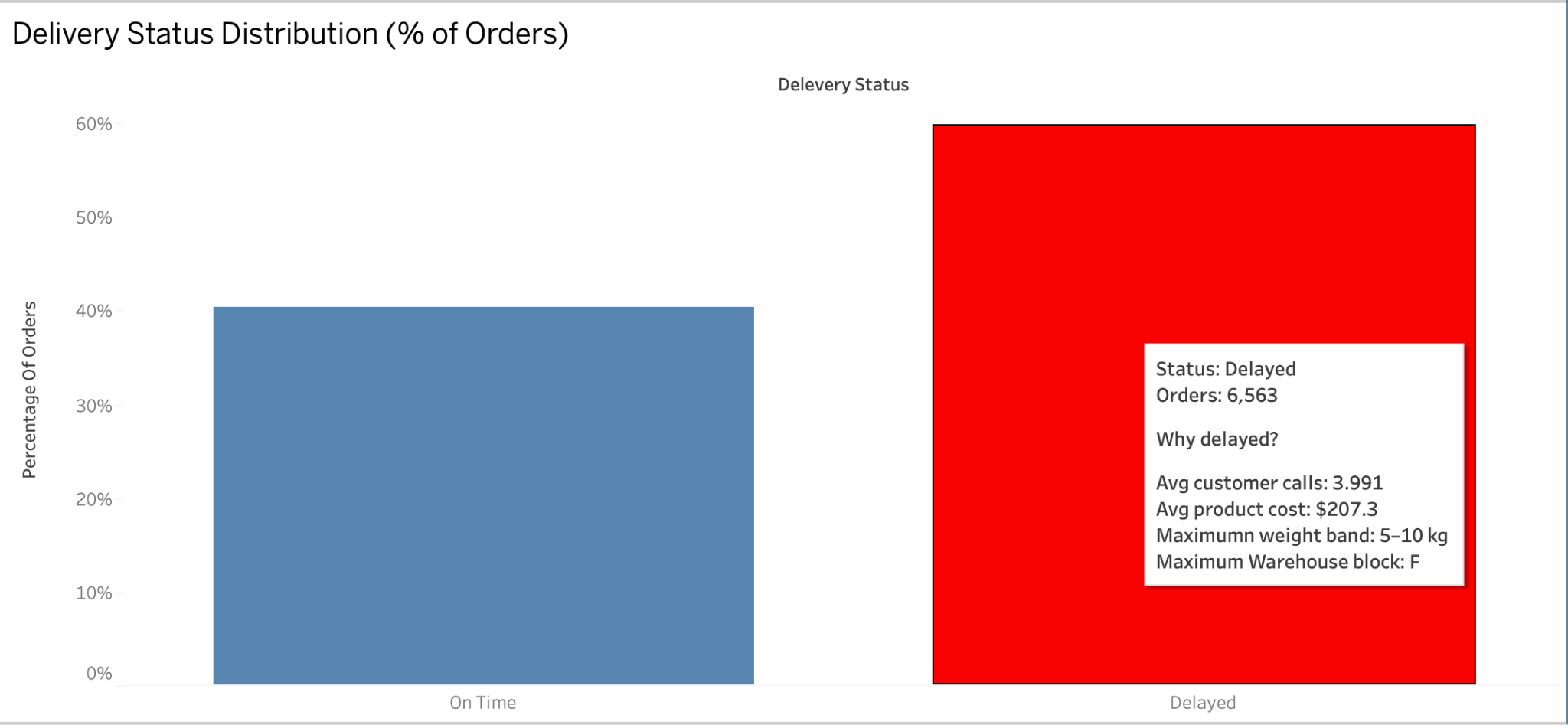


Figure 7 : Delivery Status Distribution

The Dashboard navigates to other four dashboards with a navigation bar object and responds to ctrl + click (Windows) or option + click (Mac) refer figure 8. Furthermore, all dashboards are connected with filters: if one filter is modified, it will be modified for all other dashboards.



Figure 8 : Navigation Bars



Figure 9 : Executive Summary Dashboard

The dashboard (refer figure 9) will provide a consolidated view of performance metrics, which will be particularly useful for a quick evaluation of system health and will also be useful for reporting to executives.

## 2.6 Story Interpretation Of Exclusive Summary Dashboard

The dashboard represents the beginning of the story, providing an overall view of the logistics performance. It enables the stakeholders to get an overview of the overall delivery reliability, as well as identify if the delay is a business concern. The overall view helps in establishing the context, providing the foundation for deeper analysis in subsequent views.

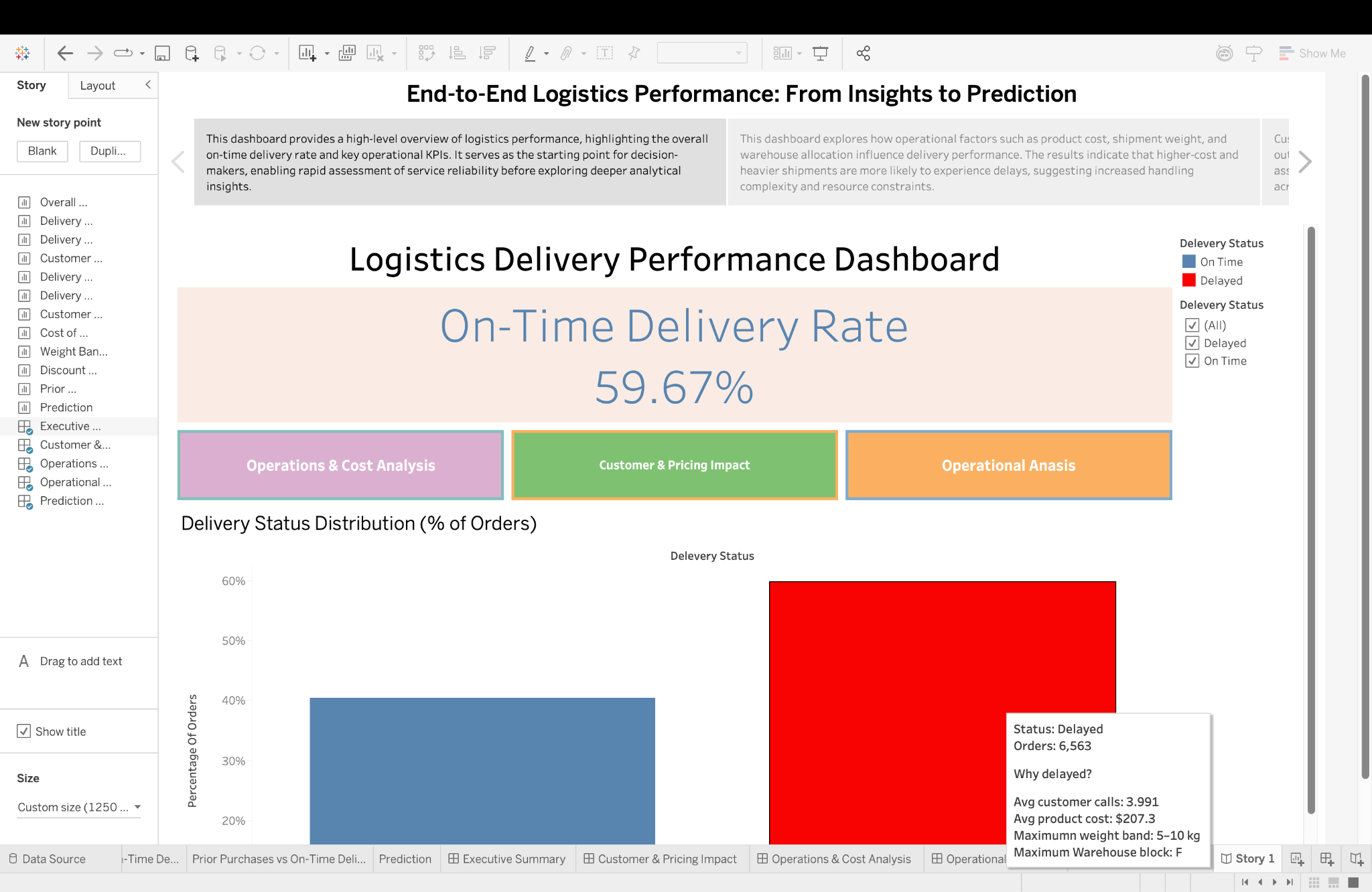


Figure 10 : Story Exclusive Summary Dashboard

# CHAPTER 3: CUSTOMER & PRICING IMPACT DASHBOARD – “THE HUMAN AND MONETARY SIDE OF DELAYS”

## 3.1 Purpose and Analytical Motivation

Though issues such as warehouse efficiency and weight of shipments remain some of the significant factors that influence delivery delays, customer-related aspects, such as price strategy, influence delivery performance, though indirectly. Customers of high commercial value, high purchase history, or even high service needs can sometimes be given high priority.

The Customer & Pricing Impact dashboard was created to test this hypothesis, and this dashboard helps to test whether delivery performance is uniformly distributed across customers or is optimized over any segment related to customers and revenues.

## 3.2 Customer Rating vs On-Time Delivery

**Design Rationale**

Customer ratings are an indirect representation of perceptions of service quality and satisfaction. In this visualization, customer rating is plotted against the on-time delivery rate to explore whether higher-rated customers receive better delivery performance or whether it is the operational outcome that influences customer perception. Refer to figure 11.

**Analytical Findings**

The scatterplot shows a moderate positive relationship between customer rating and delivery reliability. Consistently higher service ratings provided by customers match up with marginally better on-time performance.

**Business Interpretation**

This pattern suggests two possible mechanisms:

* High-quality service experiences may lower the friction of delivery, such as fewer address problems or escalations of complaints.
* Alternatively, customers with premium service history are entitled to preferential treatment in terms of operational handling.

Whatever the causality, it indicates the relationship of sustaining service standards, not only for satisfaction but also for operational efficiency.

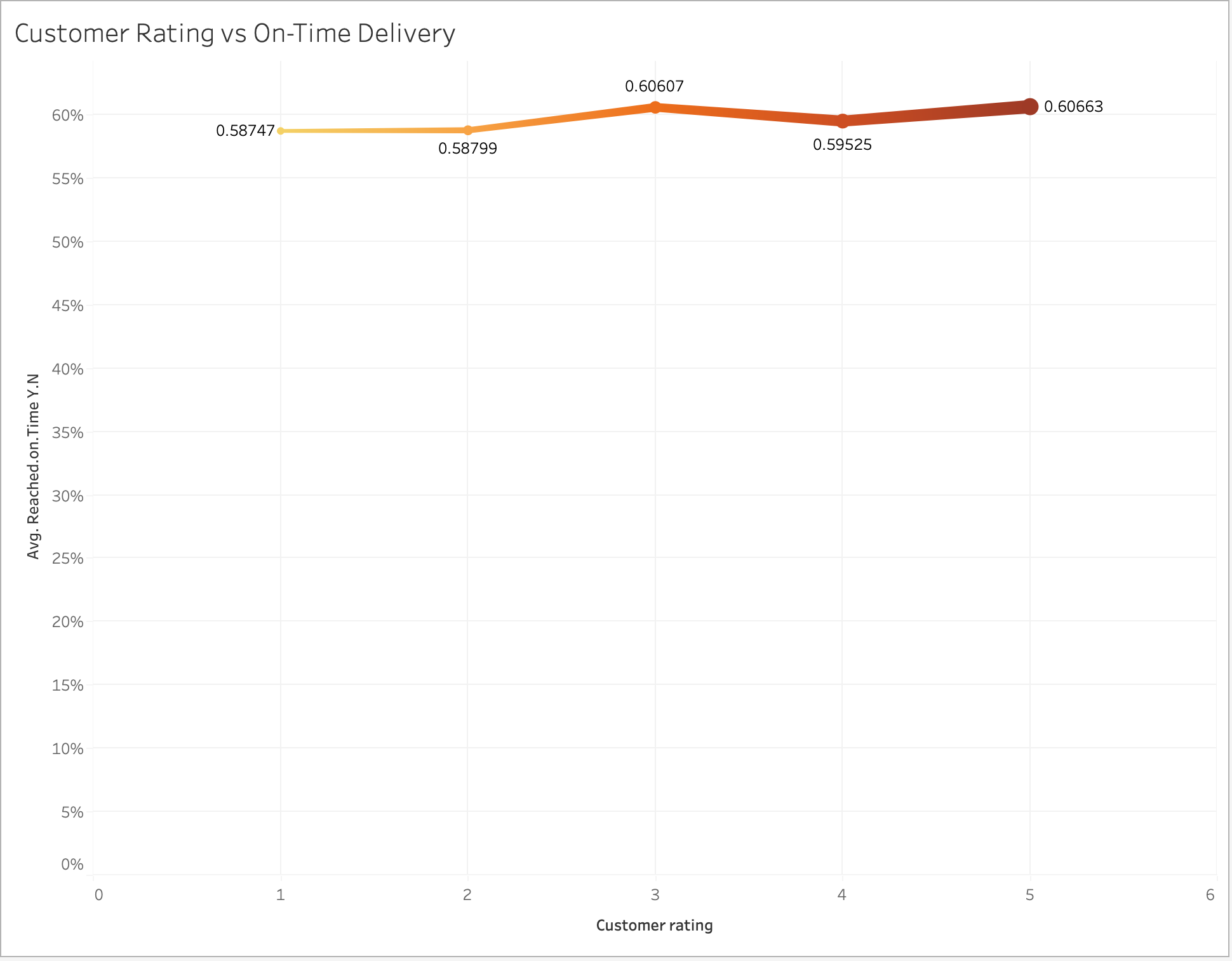


Figure 11 : Customer Rating vs On-Time Delivery

## 3.3 Discount Offered vs On-Time Delivery

**Design Rationale**

Discounts are intentional financial compromises to capture or retain a customer (refer to figure 12). This chart examines whether delivery reliability moves in with price concessions, possibly reflecting strategic prioritization of discounted orders.

**Analytical Outcomes**

The chart has shown that there is a considerable positive relationship between the discount magnitude and the probability of timely deliverables. Orders associated with higher discounts have appreciably better delivery outcomes.

**Business Interpretation**

This result indicates that:

* Discounted shipments can then be operationally prioritized in order to preserve brand reputation.
* Companies may be compensating customers for their potential inconvenience caused by the faster delivery.
* Pricing strategy and logistics planning are interrelated and not independent processes.

This insight is particularly useful for revenue management teams assessing the full cost-benefit implications of promotional campaigns.

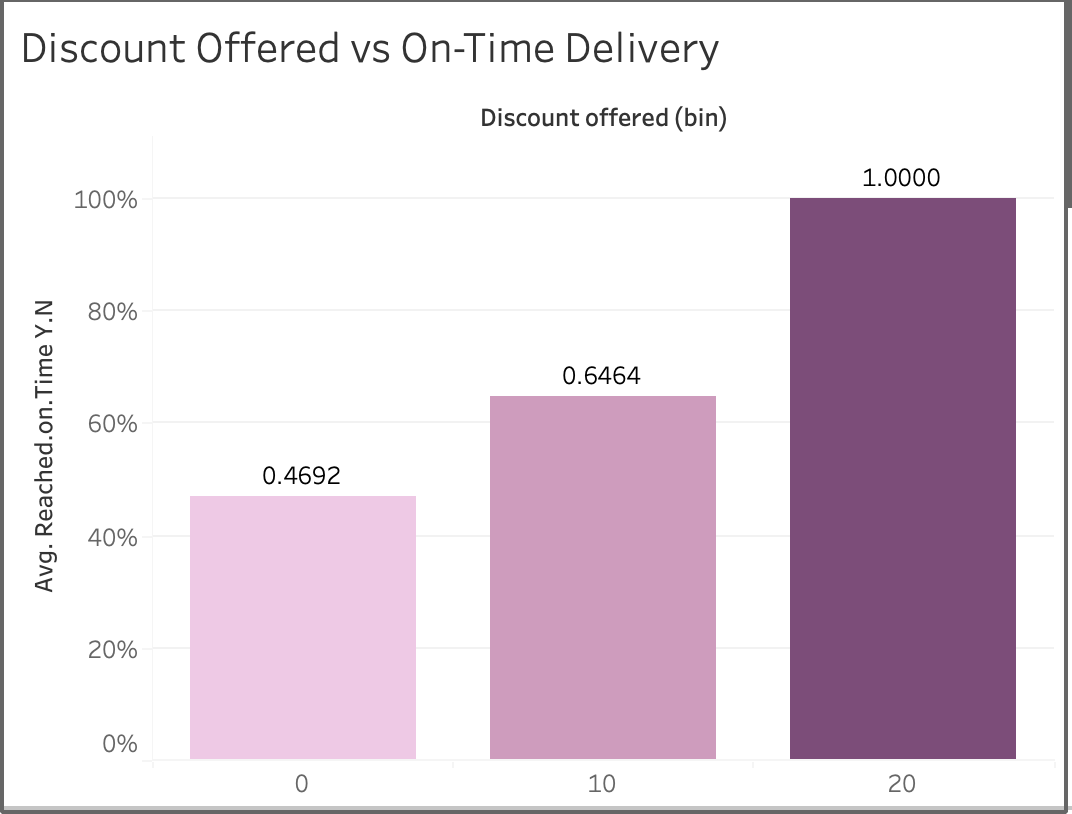


Figure 12 : Discount Offered vs On-Time Delivery

## 3.4 Prior Purchases vs On-Time Delivery

**Design Rationale**

"Customer loyalty is often rewarded with special service treatment." This is a data visualization to check if customers who return after buying a product get any special preference in terms of timely delivery as opposed to customers buying the same product for the first time. Refer to figure 13.

**Analytical Findings**

Although the overall pattern shows a tendency towards the idea that the more a customer has purchased, the higher their rate of on-time deliveries will be, the last bin seems to contradict this fact. The key reason for this may be that even though high-volume purchasers have their deliveries done more frequently, they may also order higher value items and this may serve as another type of disrupting factor.

**Business Interpretation**

which affirms the following:

* Repeat customers could also be categorized as valued consumers within a firm’s internal systems.
* Operational decision-making may be affected by long-term customer relations.
* Loyalty programs can indirectly affect logistics efficiency.

Such findings also support the strategic importance of customer retention, independent of marketing aspects.

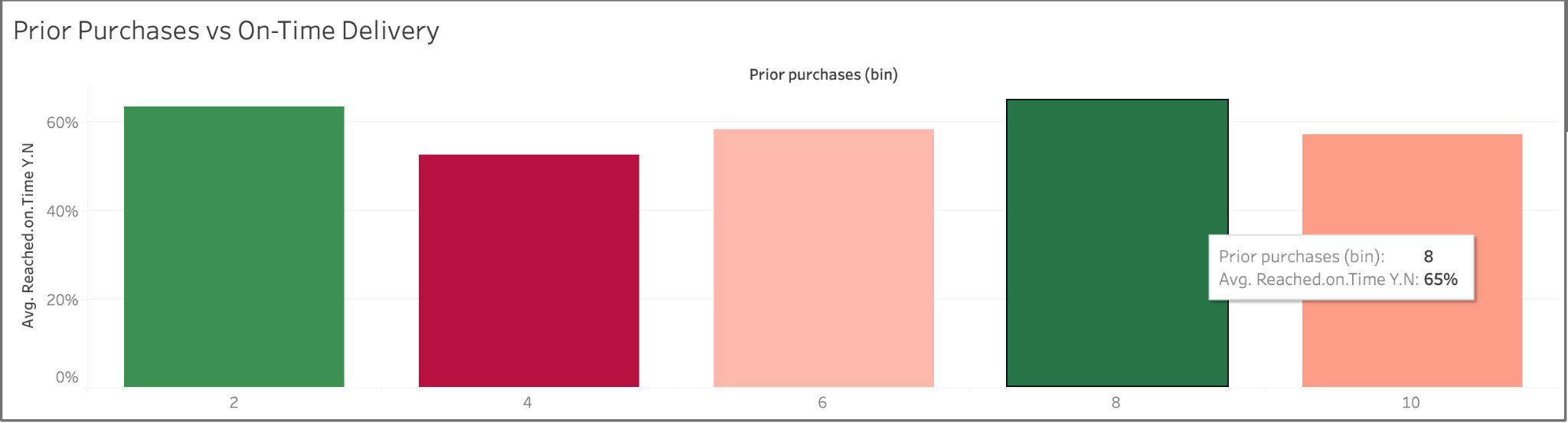


Figure 13 : Prior Purchases vs On-Time Delivery

## 3.5 Integrated Insights from the Dashboard

These three analyses taken together reveal the delivery performance is not simply the result of physical logistics infrastructure, but to an extent influenced by:

* Customer Commercial Value
* Pricing strategy
* Customer relationship history

Basically, it shows that there is an implicit service segments basis, whereby various categories of customers are accorded different degrees of operational priority.

## 3.6 Managerial Implications

The insights derived from this dashboard support several strategic actions:

* Align discount strategy with logistics capacity planning.
* Monitor service equity across customer segments to avoid reputational risk.
* Design loyalty programs that explicitly incorporate service-level guarantees.
* Identify customers at risk of dissatisfaction due to delivery delays.

## 3.7 Visualization Design Considerations

The dashboard employs:

* Consistent color encoding for delivery status
* Bar and line charts for trend comparability
* Global filters for segmentation
* Click-to-filter actions for rapid drill-down

These features enhance analytical flexibility while preserving clarity.

"The below dashboard displays the effect of metrics related to consumer satisfaction, financial benefits, and purchase records upon the accomplishment of on-time deliveries." Also when clicked on any metrics there is a navigation link placed which helps return to the exclusive summary dashboard. Refer to figure 14.

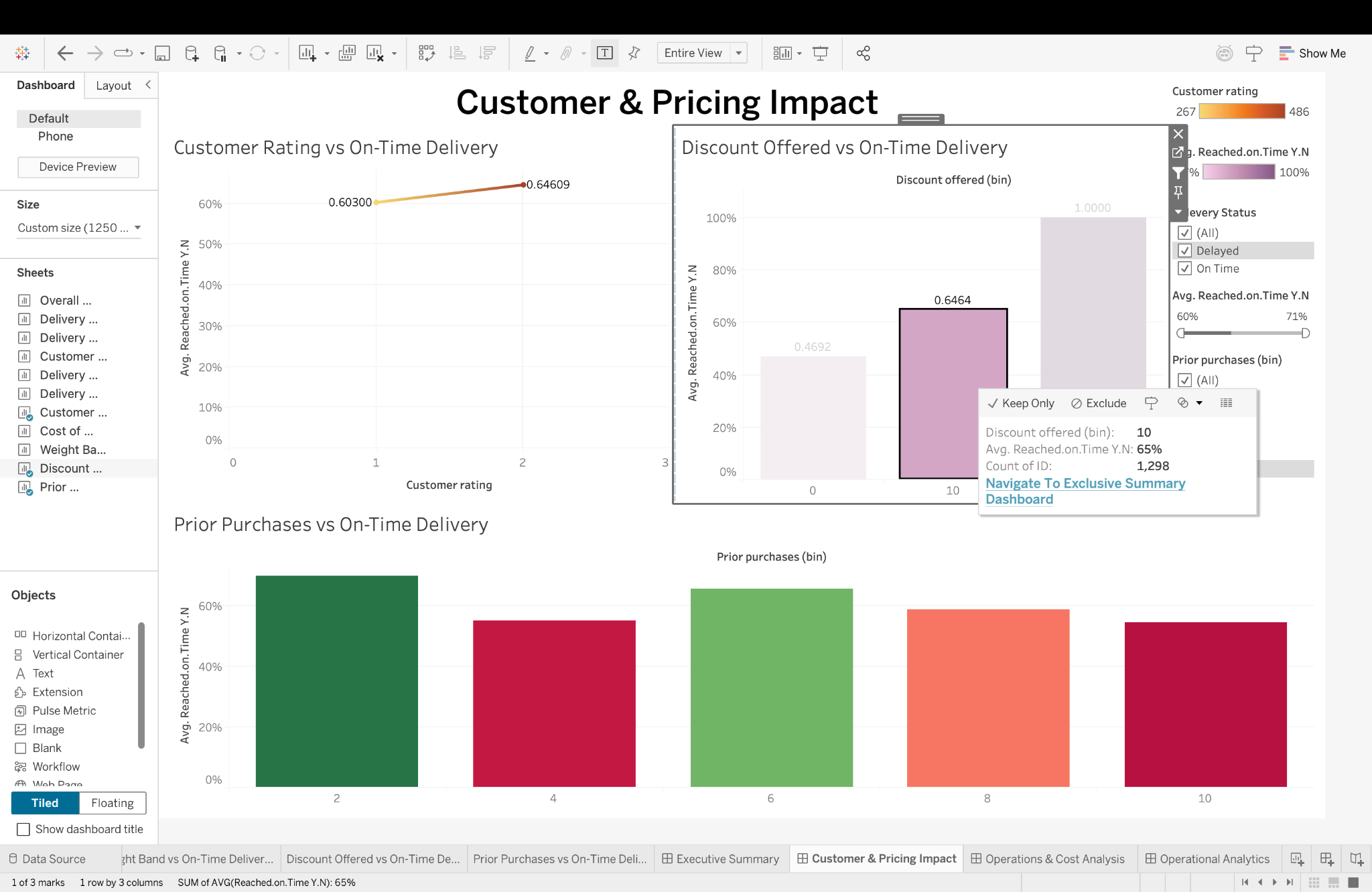


Figure 14 : Customer & Pricing Impact Dashboard

## 3.8 Story Interpretation Of Customer And Impact Dashboard

This dashboard extends the story to customer-centric and pricing-related dimensions, illustrating that the performance of the delivery is not necessarily driven by physical constraints. The differences in the performance of the delivery for customer ratings, discounts, and purchase history imply that there is a form of implicit prioritization.

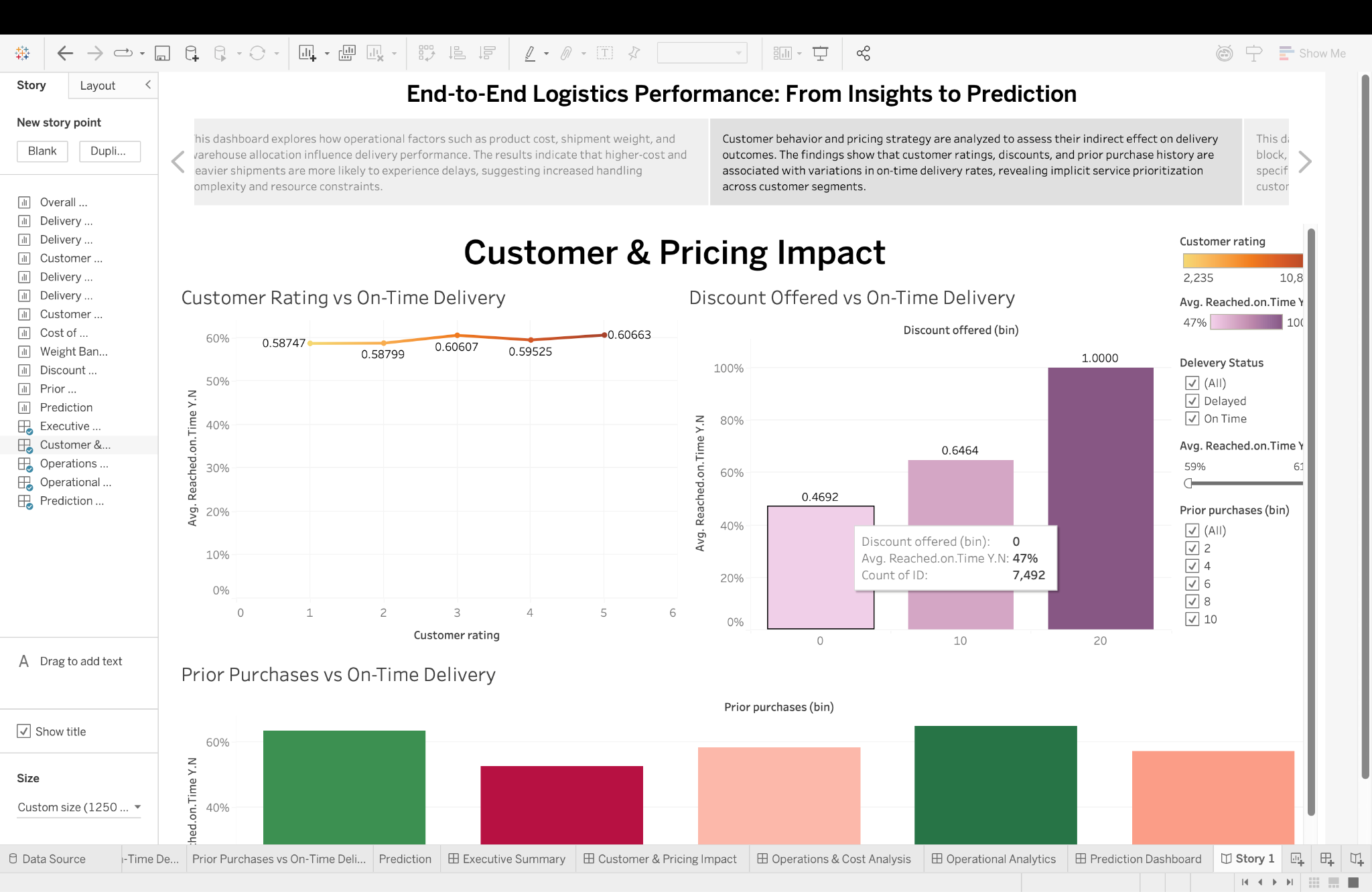


Figure 15 : Story Customer & Pricing Impact Dashboard

# CHAPTER 4: OPERATIONS & COST ANALYSIS DASHBOARD – “WHERE LOGISTICS MEETS PHYSICAL AND ECONOMIC CONSTRAINTS”

## 4.1 Purpose and Analytical Motivation

Even though customer practices and pricing strategies can influence delivery outcomes, the primary factors that affect delivery outcomes directly relate to capacity and shipment characteristics. For instance, product costs, weight, and other factors can directly affect delivery outcomes.

The Operations & Cost Analysis Dashboard was specifically created to assess how this structural constraint influences delivery reliability. The main goal of this analysis is to ascertain if the causes of delay are largely driven by economic risk (high-value goods) rather than physical risk (weight of goods), and by how much.

The focus of analysis for this dashboard therefore changes from customer-related aspects to **infrastructure**, **workload**, and **operational feasibility**.

## 4.2 Cost of Product vs On-Time Delivery

**Design Rationale**

The cost of a product serves as a proxy for:

* Financial risk associated with shipment damage or loss
* Security requirements
* Insurance handling procedures
* Additional verification and documentation

A scatter plot with trend lines segmented by product importance was employed to capture both:

* Individual shipment-level variation
* Overall directional relationships between cost and delivery success

Trend lines allow identification of systematic performance patterns rather than isolated fluctuations.

**Analytical Findings**

The visualization reveals a consistent negative relationship between product cost and on-time delivery probability. As product cost increases, the likelihood of timely delivery declines.

* High-importance products exhibit marginally better performance than low-importance goods at similar cost levels.
* However, even high-priority products experience performance degradation at extreme cost values.

This suggests that operational complexity increases disproportionately with shipment value.

**Business Interpretation**

Several operational mechanisms may explain this pattern:

* **Additional security protocols** increase processing time.
* **Manual verification procedures** delay dispatch.
* **Special handling requirements** reduce routing flexibility.
* **Risk aversion** leads to conservative scheduling choices.

Consequently, expensive products introduce friction into logistics workflows, increasing the probability of schedule deviation. Refer to figure 16.

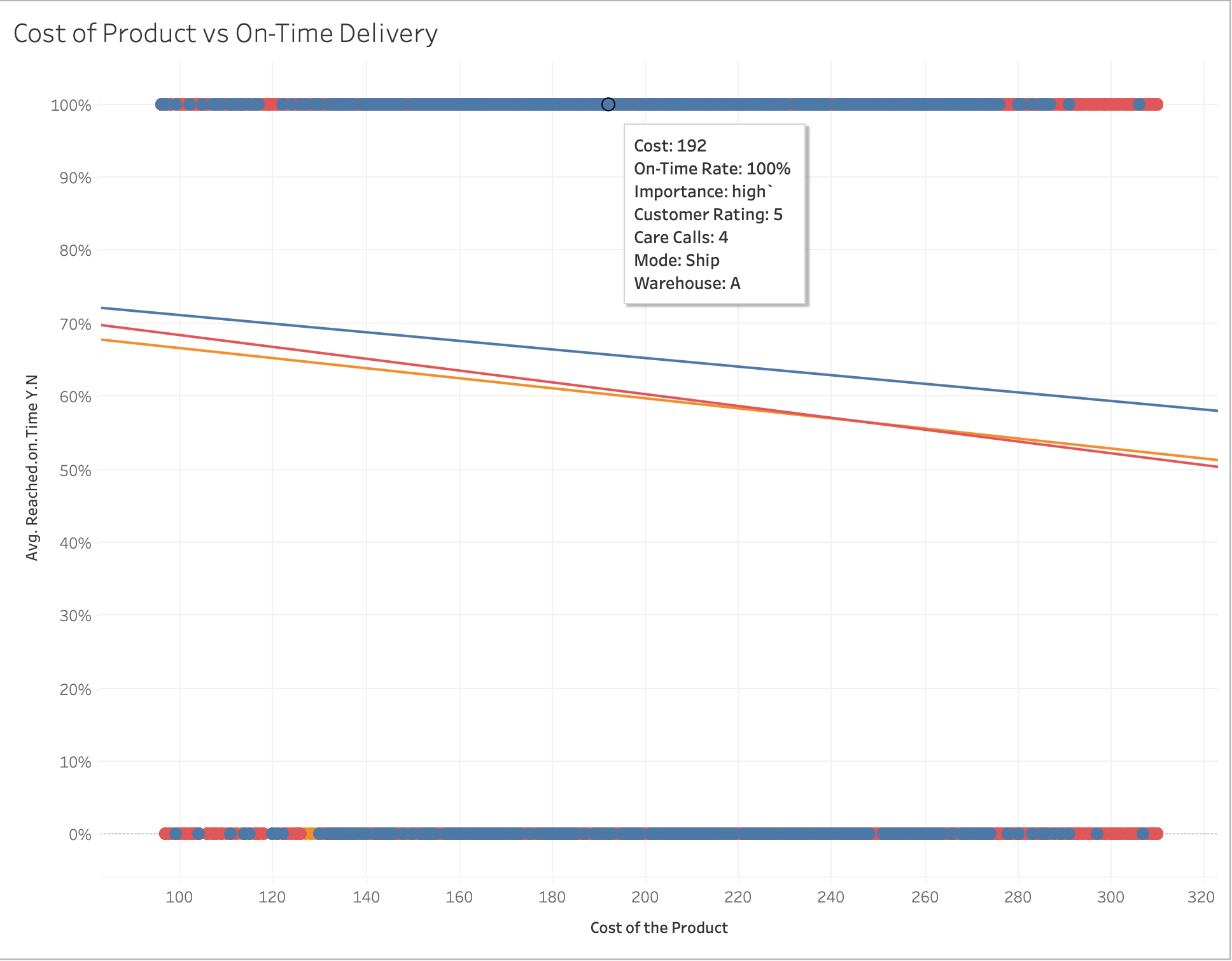


Figure 16 : Cost Of Product vs On-Time Delivery

## 4.3 Weight Band vs On-Time Delivery Rate

**Design Rationale**

Shipment weight directly influences:

* Vehicle capacity utilization
* Loading and unloading time
* Sorting complexity
* Route optimization flexibility
* Fuel consumption and speed constraints

To examine this relationship, shipments were grouped into weight bands using binning techniques to facilitate comparison across discrete operational categories. Refer to figure 17.

**Analytical Findings**

| Weight Band | Approximate On-Time Rate |
| --- | --- |
| 0-2 kg | ~68% |
| 2-5 kg | ~55% |
| 5-10 kg | ~43% |

The bar chart demonstrates a strong monotonic decline in delivery performance as weight increases:This pattern indicates that heavy shipments are significantly more prone to delay.

**Business Interpretation**

Heavy shipments impose several operational penalties:

* Reduced vehicle capacity leads to fewer deliveries per route.
* Increased loading time causes departure delays.
* Limited vehicle availability constrains scheduling flexibility.
* Higher probability of mechanical or handling-related disruptions.

As a result, heavy goods act as structural bottlenecks within the delivery system.

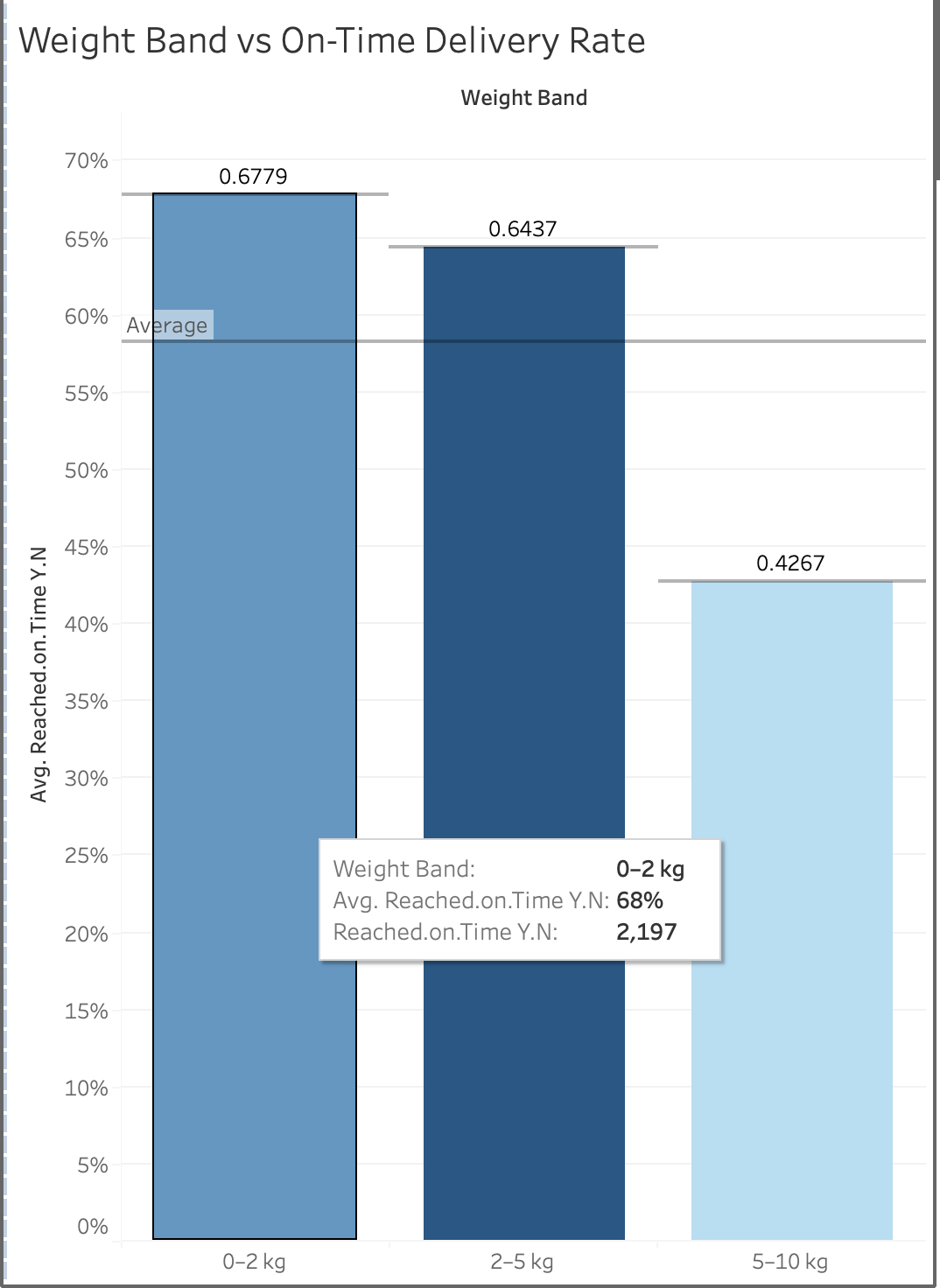


Figure 17 : Weight Band vs On-Time Delivery Rate

## 4.4 Integrated Operational Insights

If we examine product cost and shipment weights together, we find a compounding effect:

* High-cost and high-weight shipments have the worst level of reliability.
* These shipments also signify financial risk.

This interaction indicates that: The timing of delays is neither random nor accidental. There are systematic relationships between delays and strain.

## 4.5 Managerial Implications

The insights generated by this dashboard can be used for a number of evidence-based interventions:

* Segregation of high-value goods into specialized logistics pipelines
* Deployment of dedicated vehicles for heavy shipments
* Investment in automated loading systems
* Risk-adjusted scheduling buffers for costly products
* Dynamic pricing strategies to offset operational burden

## 4.6 Visualization Design Considerations

The dashboard utilizes:

* Scatter plots can be constructed in order to inspect the characteristics of
* Trend lines to expose systematic relationships
* To differentiate the significance levels of a product, color encoding can be used
* Binned Bar Charts: Achieving Categoric Clarity
* Interactive filters for scenario-based analysis
* Click-to-Filter Actions to Support Cross-Dashboard

All these features help in exploratory analysis as well as explanatory stories.

## 4.7 Contribution to Analytical Narrative

This dashboard constitutes the **structural backbone** of the project’s analytical narrative by demonstrating that:

* Operational constraints exert stronger influence on delivery performance than customer behavior alone.
* Physical logistics capacity fundamentally limits service reliability.
* Performance optimization must therefore address infrastructure design, not merely customer management.

"This dashboard (refer to figure 18)examines how economic risk (product cost) and physical burden (shipment weight) influence on-time delivery probability, revealing structural constraints within logistics operations."

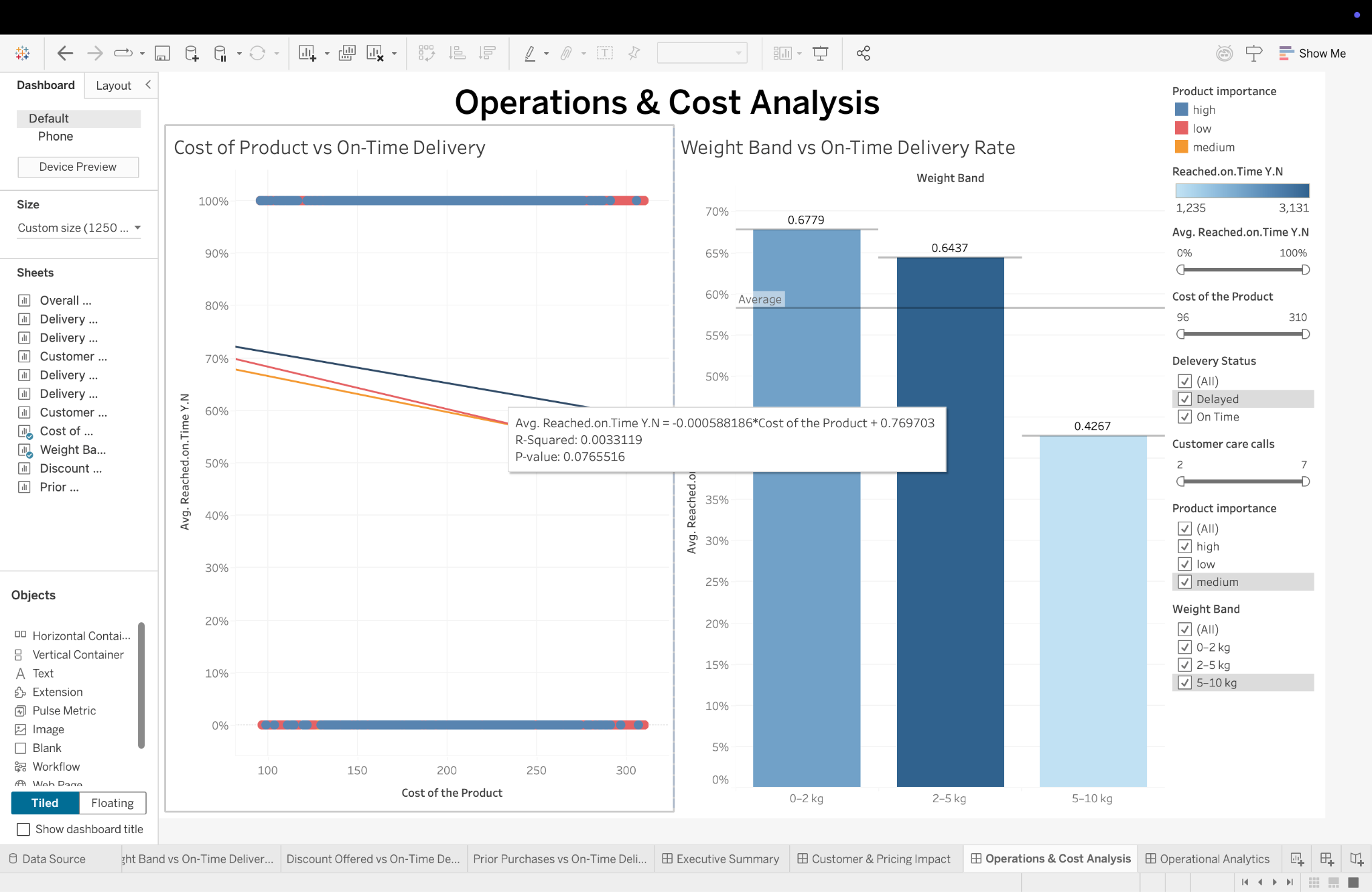


Figure 18 : Operations and Cost Analysis Dashboard

## 4.8 Story Interpretation Of Operations and Cost Analysis Dashboard

This dashboard moves the story forward by exploring the impact of internal operation-related variables on delivery performance. By considering product cost, shipment details, and warehouse allocation, this dashboard points to areas where the complexity of operations is causing delivery delays. This changes the story from what is happening to where the delivery delays are occurring in the logistics chain.

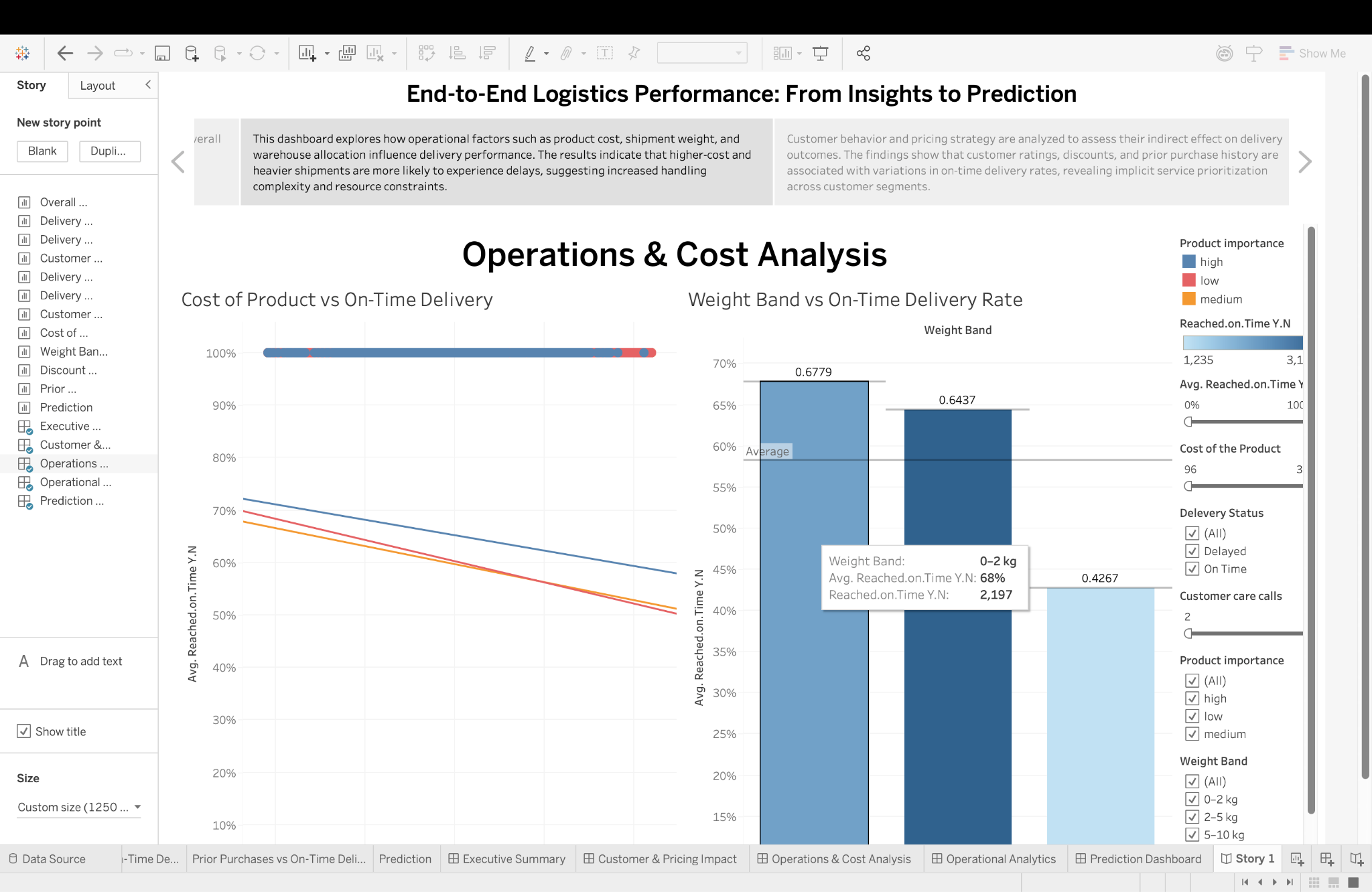


Figure 19 : Story Operations and Cost Analysis Dashboard

# CHAPTER 5: OPERATIONAL ANALYTICS DASHBOARD – “INSIDE THE LOGISTICS ENGINE ROOM”

## 5.1 Purpose and Analytical Motivation

However, while earlier dashboards focused on customer trends and internal shipment limitations, for a comprehensive understanding of delivery outcomes, internal operational processes need to be directly measured. These factors include internal warehouse management efficiency, shipment routing, prioritization, and customer service.

This Operational Analytics Dashboard was constructed with the intention of creating a detailed view of the inner workings of the above mechanisms. Its aim is to be able to see where within the logistics system the delays actually start, rather than reassessing the broader policies being used to inform the response to the delay itself.

"This dashboard serves as the equivalent of looking in the 'engine compartment' beneath the surface of the logistics system and seeing how it affects the delivery."

## 5.2 Delivery Performance by Product Importance

**Design Rationale**

Product categories in this dataset include an importance factor, which defines a level of criticality such as low, medium, or high.

The goal behind this visualization was to assess the consequences of internal policies in the actual delivery. Refer to figure 20.

**Analytical Findings**

The analysis implies that:

* The highest on-time deliveries apply to the high-importance products.
* Medium-importance products display moderate performance.
* The highest delay frequency occurs for low-importance products.

**Business Interpretation**

The findings imply the presence of intentional prioritization within the operation, with scarce logistics resources being allocated to more important goods.

This is in line with how such prioritization is done in industries where Service Level Agreements and Revenue Protection are well applied.

However, this secures the danger of service inequities, which can influence customer satisfaction among less privileged customers.

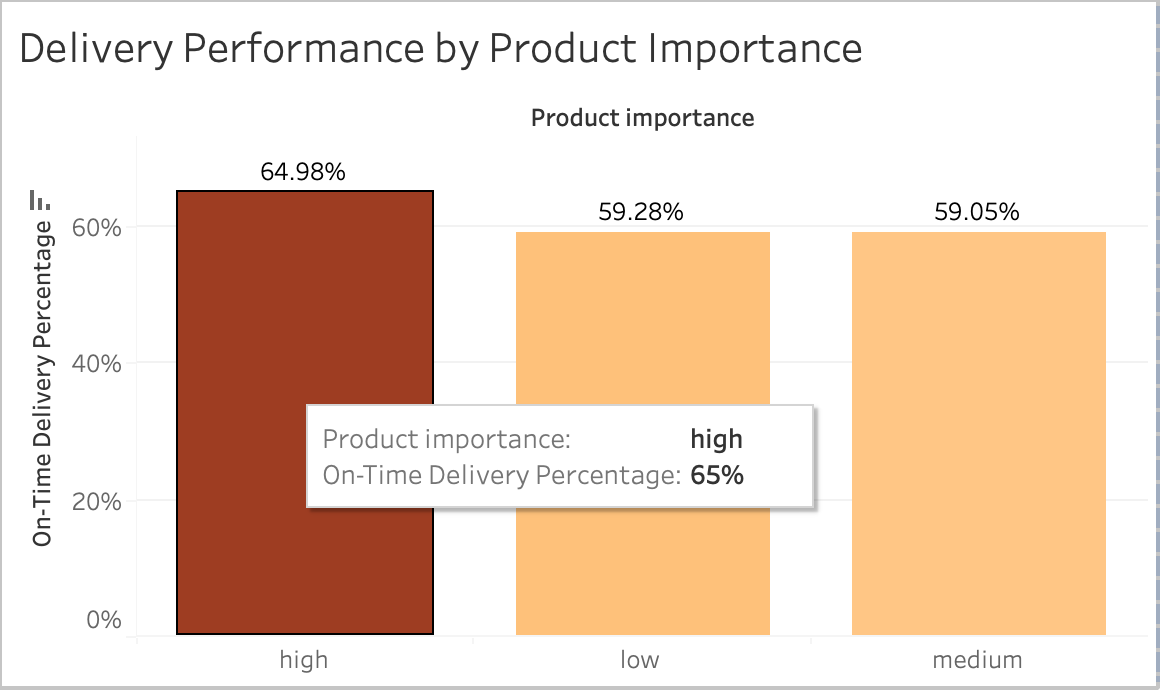


Figure 20 : Delivery Performance By Product Importance

## 5.3 Delivery Performance by Warehouse Block

**Design Rationale**

Blocks representing the warehouse have been used to show the physical storage and processing capacity in the distribution network, where varying staff levels, equipment, and layout might cause variability in performance. Refer to figure 21.

This visualization analyzes how different blocks of a warehouse are performing in terms of deliveries.

**Analytical Findings**

The dashboard shows that significant performance diversely:

* The Warehouse Block F has always recorded the least on-time delivery rate.
* Other blocks operate at a fairly even pace at higher efficiencies.

**Business Interpretation**

Persistent underperformance of a particular block of a warehouse can thus be explained by:

* Congestion and capacity overload
* Suboptimal layout design
* Workforce shortages or skill gaps in a company can
* Equipment Limitations
* Poor coordination with transport schedules

Such bottlenecks can be highly costly since they create delays in the system of distribution.

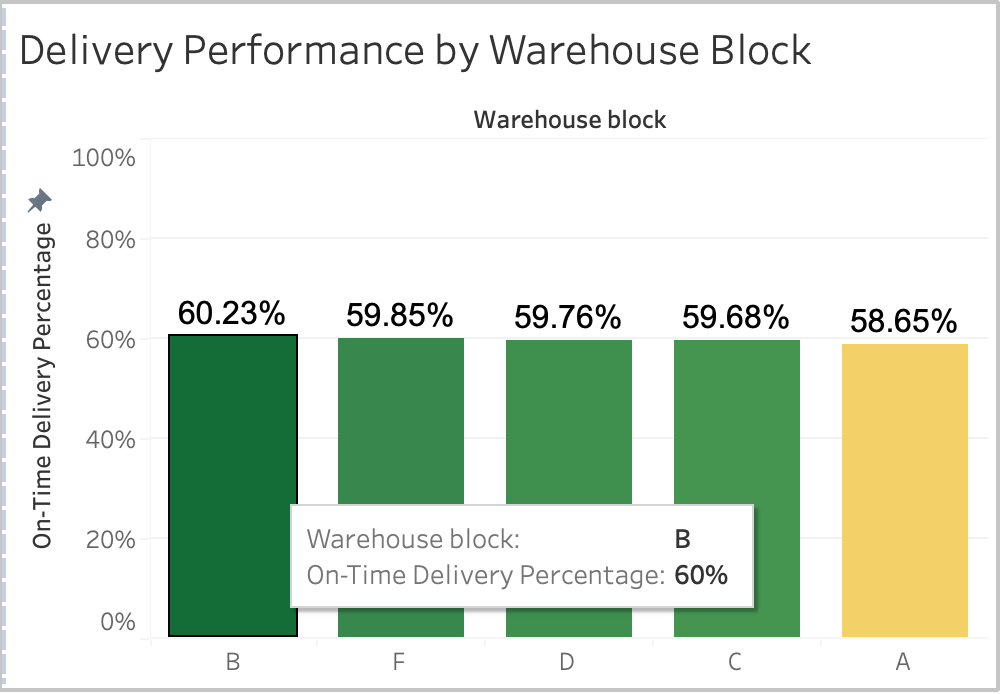


Figure 21 : Delivery Performance By Warehouse Block

## 5.4 Delivery Performance by Mode of Shipment

**Design Rationale**

Speed, reliability, and associated costs are called shipment mode. There are modes of flight, ships, and roads involved too.

The following is a visualization of the performance of different modes in satisfying their respective delivery deadlines.

**Analytical Findings**

The data indicates:

* Types of consignments that have the highest probability of being delivered on time include flights.
* The state of road transport is moderate.
* The mode most subject to variation and longest delay is ship transport.

**Business Interpretation**

These results demonstrate several well-known trade-offs:

* Secondly, air transport ensures a faster service but at a cost.
* The trade-off between flexibility and cost in road transport.
* Though efficient, maritime shipping is also vulnerable to schedules being delayed.

There is a basis through empirical justification, specifically through a particular mode of transportation with respect to the urgency level of deliveries, as can be visualized through the provided information. Refer to figure 22.

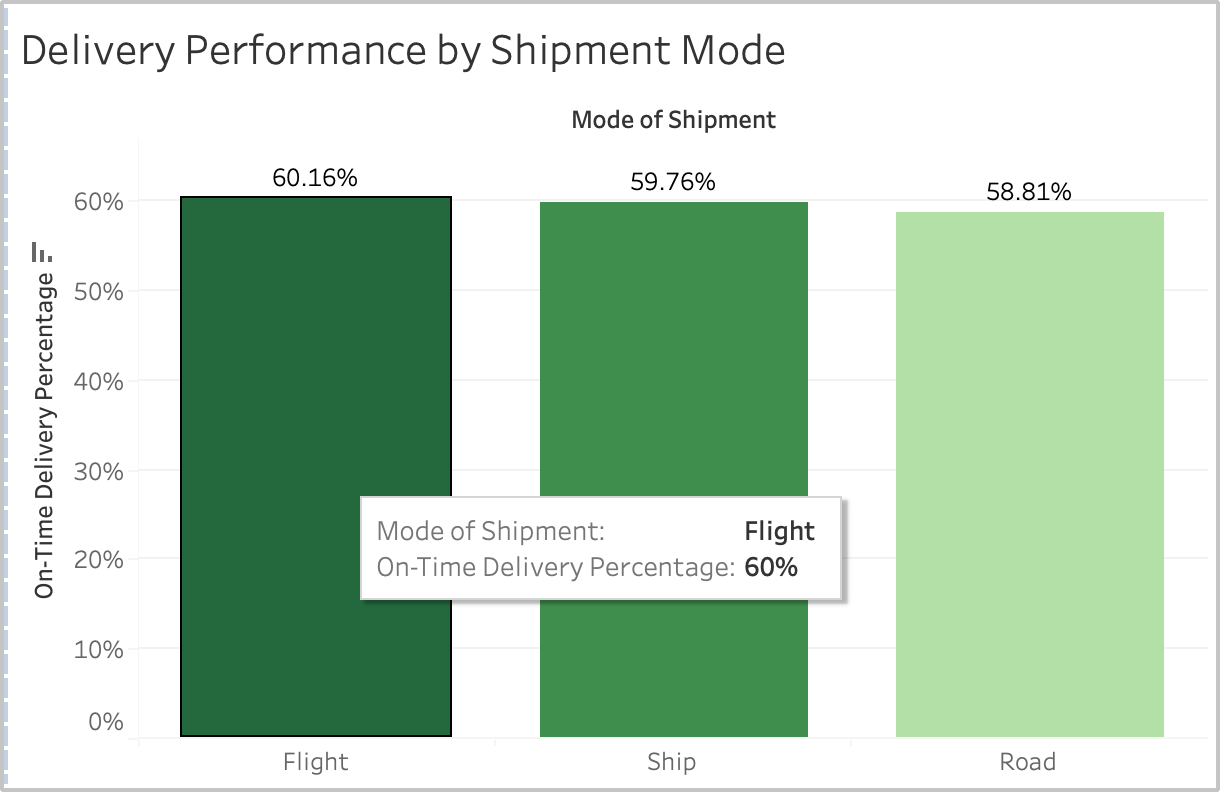


Figure 22 : Delivery Performance By Shipment Mode

## 5.5 Customer Care Calls vs Delivery Performance

**Design Rationale**

Customer care calls are an indirect proxy for the level of operational friction, dissatisfaction with services, and complexity of logistics involved in the firm’s activities.

The discussion will include an analysis of how more interaction with the customer support service and its subsequent correlation with failed deliveries can be determined. Refer figure 23.

**Analytical Findings**

The figure shows that there exists a strong negative relationship:

* Orders with few or no customer care associated have high order reliability.
* The more calls there are, however, the smaller is the chance of on-time deliveries.

**Business Interpretation**

Several causal mechanisms can be suggested:

* Calls may also relate to issues with addresses and scheduling.
* Complaints can be a symptom of underlying delivery problems.
* Redundant communication can lead to delays, depending on human intervention.

Therefore, customer care workload can be considered both a symptom of operational inefficiency and a causal factor.

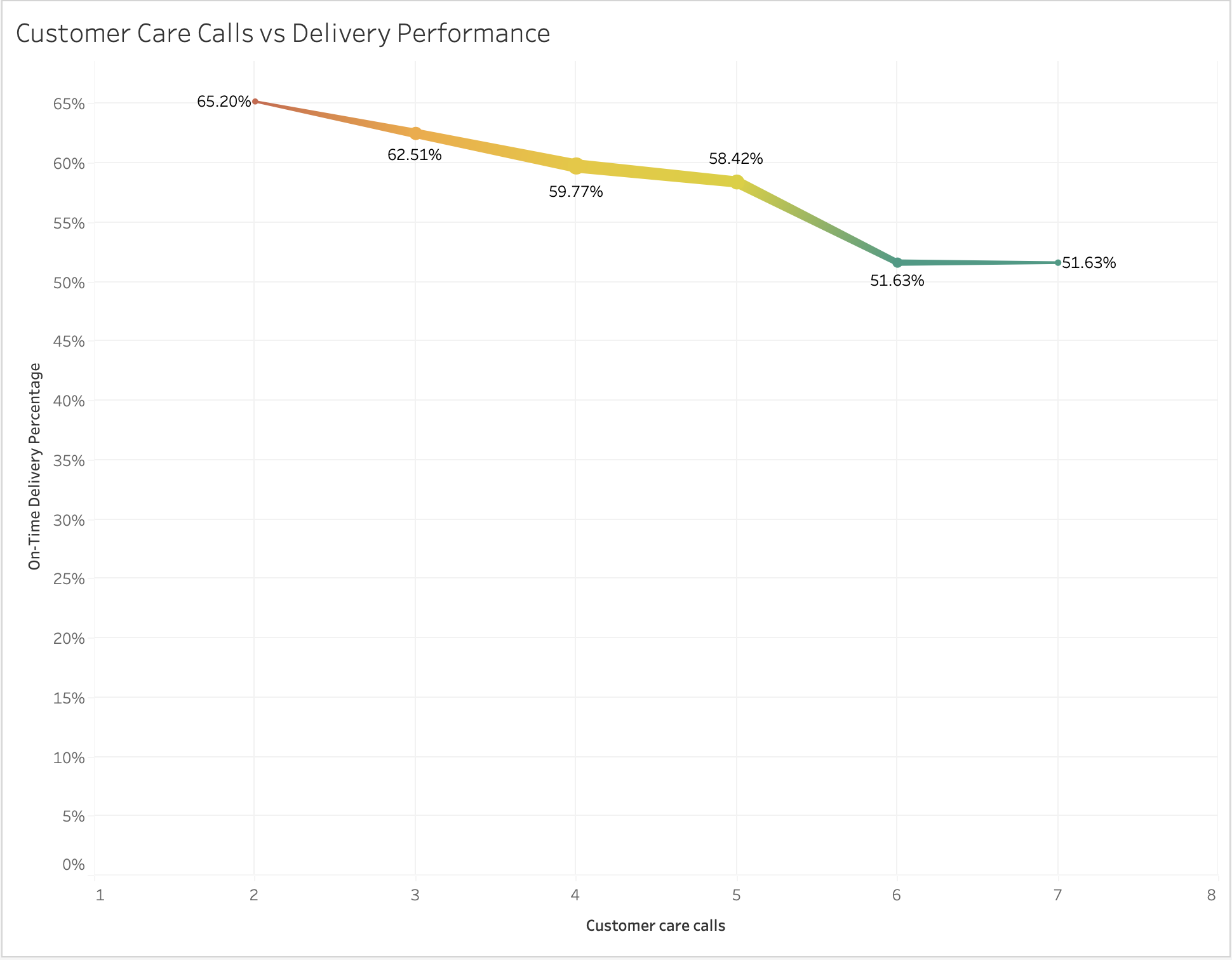


Figure 23 : Customer Care Calls vs Delivery Performance

## 5.6 Managerial Implications

The findings support targeted interventions:

* Conduct an audit and redesign an ineffective warehousing block.
* Workload balancing across the warehouse.
* Defining Service Level Tiers Transparently
* Optimizing the choice of transport mode based on risks.
* Automate process to eradicate customer care burden.

## 5.7 Contribution to the Overall Analytical Framework

This dashboard builds on previous analysis in several ways:

* Translating abstract performance metrics into actionable operating information.
* Linking customer-facing outcomes to internal system behavior.
* Justifying infrastructure development and policy revision with evidence-based reasoning.
* Hence, it includes the diagnostic core of the visualization system.

Research studies have demonstrated that, in addition, artificial intelligence has the potential to improve efficiency in logistics through predictive delay detection, route optimization, and warehouse automation (Patel & Singh, 2024).

"This dashboard provides a granular operational diagnosis of delivery performance, highlighting the internal process and infrastructure factors driving delays." (figure 24).

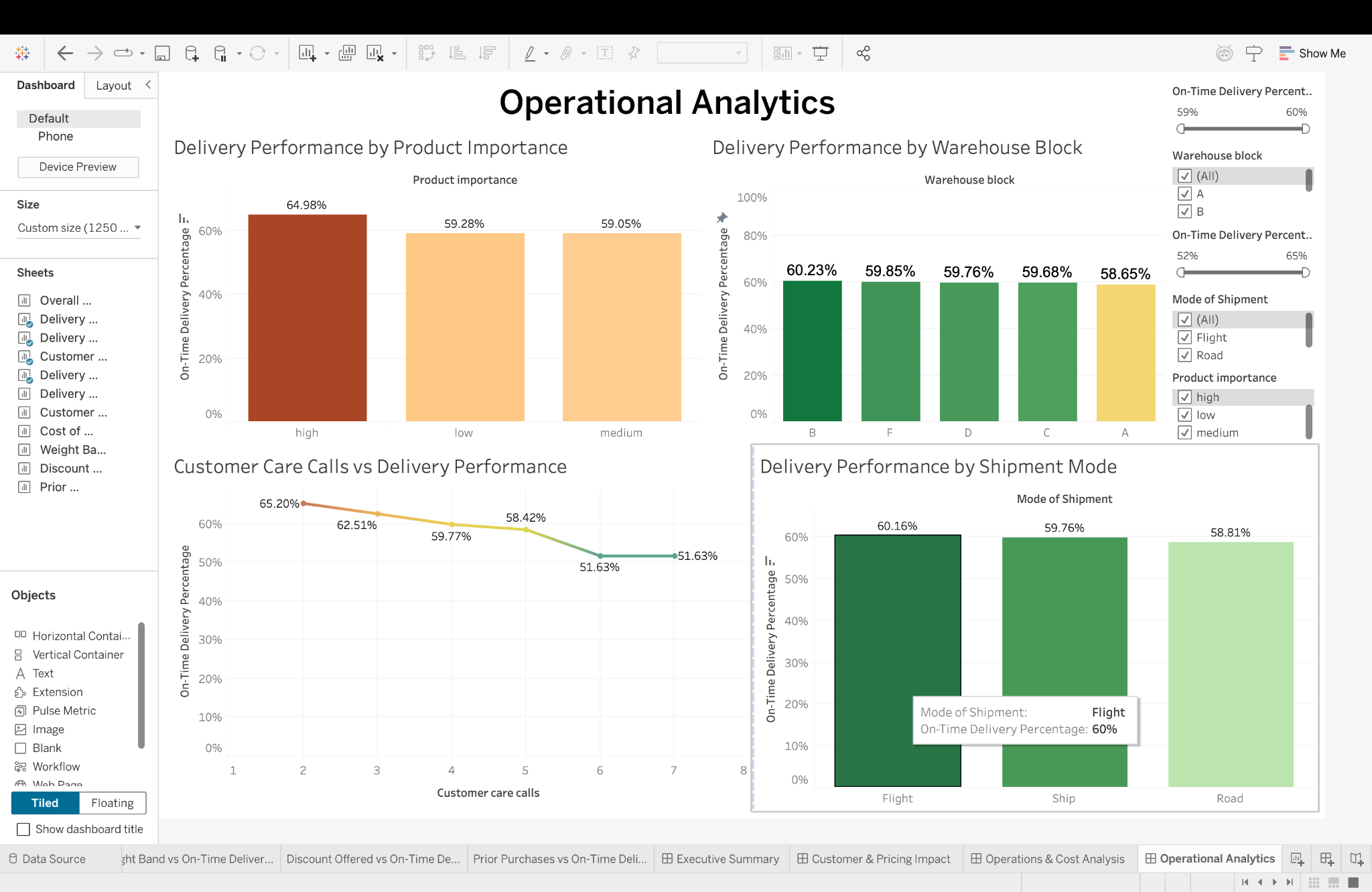


Figure 24 : Operational Analytics Dashboard

## 5.8 Story Interpretation Of Operational Analytics Dashboard

The dashboard offers a diagnostic layer of detail by drilling down into blocks of the warehouse, modes of shipment, product importance, and customer care interactions. It identifies operational bottlenecks and inefficiencies, which can be addressed rather than taking general actions. At this stage, the story begins to move towards operational insights.

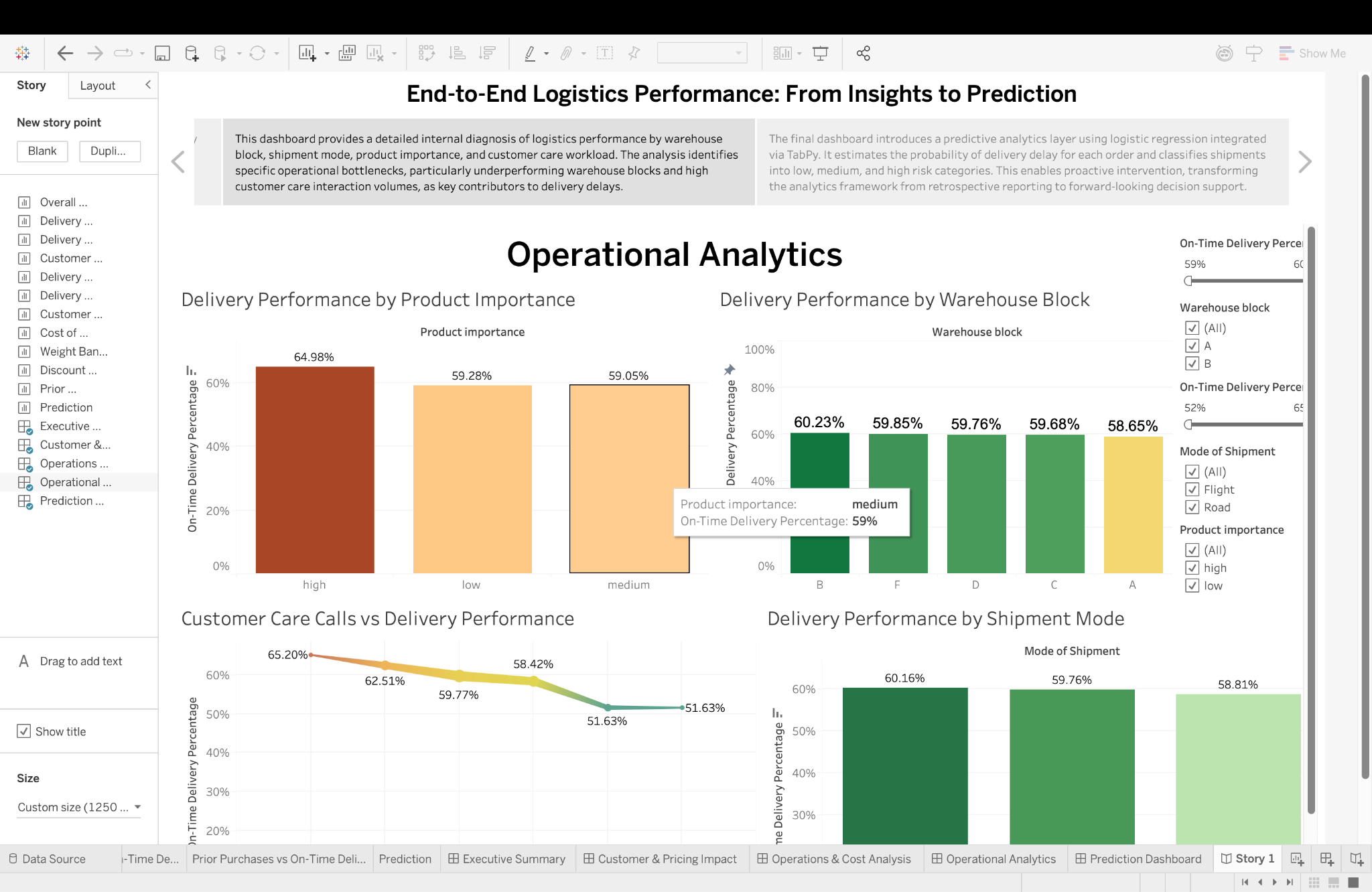


Figure 25 : Story Operational Analysis

# 

# CHAPTER 6: MACHINE LEARNING – BASED DELIVERY RISK PREDICTION DASHBOARD

## 6.1 Purpose of the Dashboard

In order to expand the scope of the dashboard in terms of analysis, a predictive analytics component was created using logistic regression, which was integrated with Tableau using TabPy, also known as Tableau Python Server. This component predicts the probability of delay for individual orders based on certain attributes.

The purpose of the prediction component is to aid in better decision-making by identifying potential risks for individual orders before dispatch.

## 6.2 Model Design and Feature Selection

A binary logistic regression model was implemented to classify whether an order is likely to be delayed (1) or delivered on time (0). The model was trained using the following predictor variables:

* Cost of the Product
* Customer Care Calls
* Customer Rating
* Discount Offered
* Prior Purchases

These variables were selected due to their strong relationship with customer behavior, order complexity, and service workload, which are known to influence delivery performance.

The target variable was defined as:

* **Reached on Time (Y/N)**
  + 1 → Delayed
  + 0 → On time

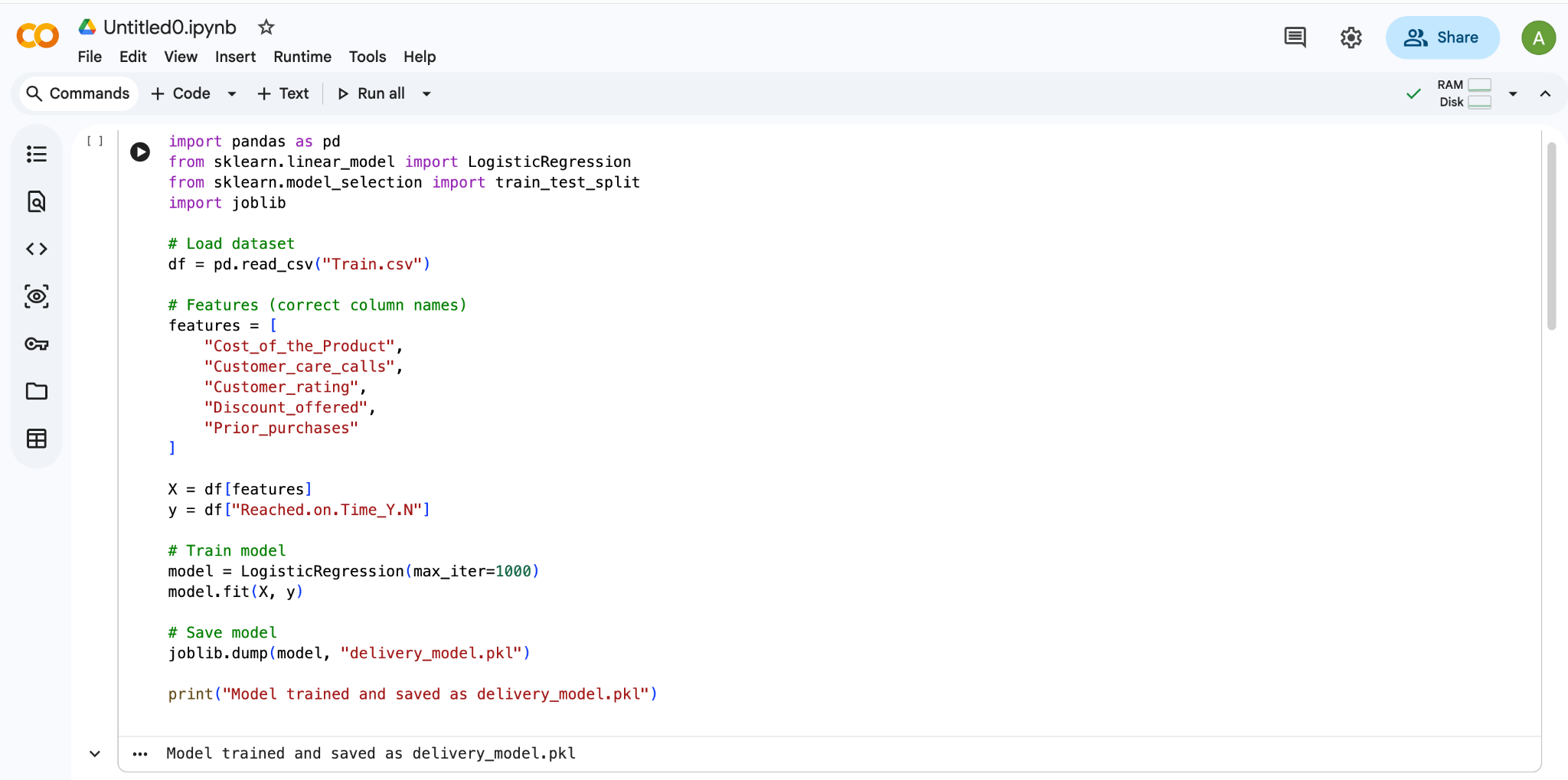


Figure 26 : Delivery\_Model.pkl File

## 6.3 Integration with Tableau Using TabPy

The trained model was deployed on the TabPy server and accessed directly within Tableau using the Script\_real() function. This enabled real-time computation of delay probabilities for each order based on the values selected or filtered in the dashboard.

The output of the model is a continuous value between 0 and 1, referred to as: **Predicted Delay Probability**

This value represents the likelihood that a shipment will experience a delivery delay.

## 6.4 Logistic Regression Visualization

The prediction results were visualized using a continuous curve showing the relationship between **Cost of the Product** and the **Predicted Delay Probability**.

The curve demonstrates a nonlinear relationship between product cost and delivery risk, illustrating how higher-value shipments tend to exhibit increased probability of delay due to handling complexity, prioritization constraints, or security requirements.

Although the curve does not appear as a perfectly symmetrical S-shape, it correctly reflects the probabilistic behavior learned from the dataset and the logistic function applied during model training.

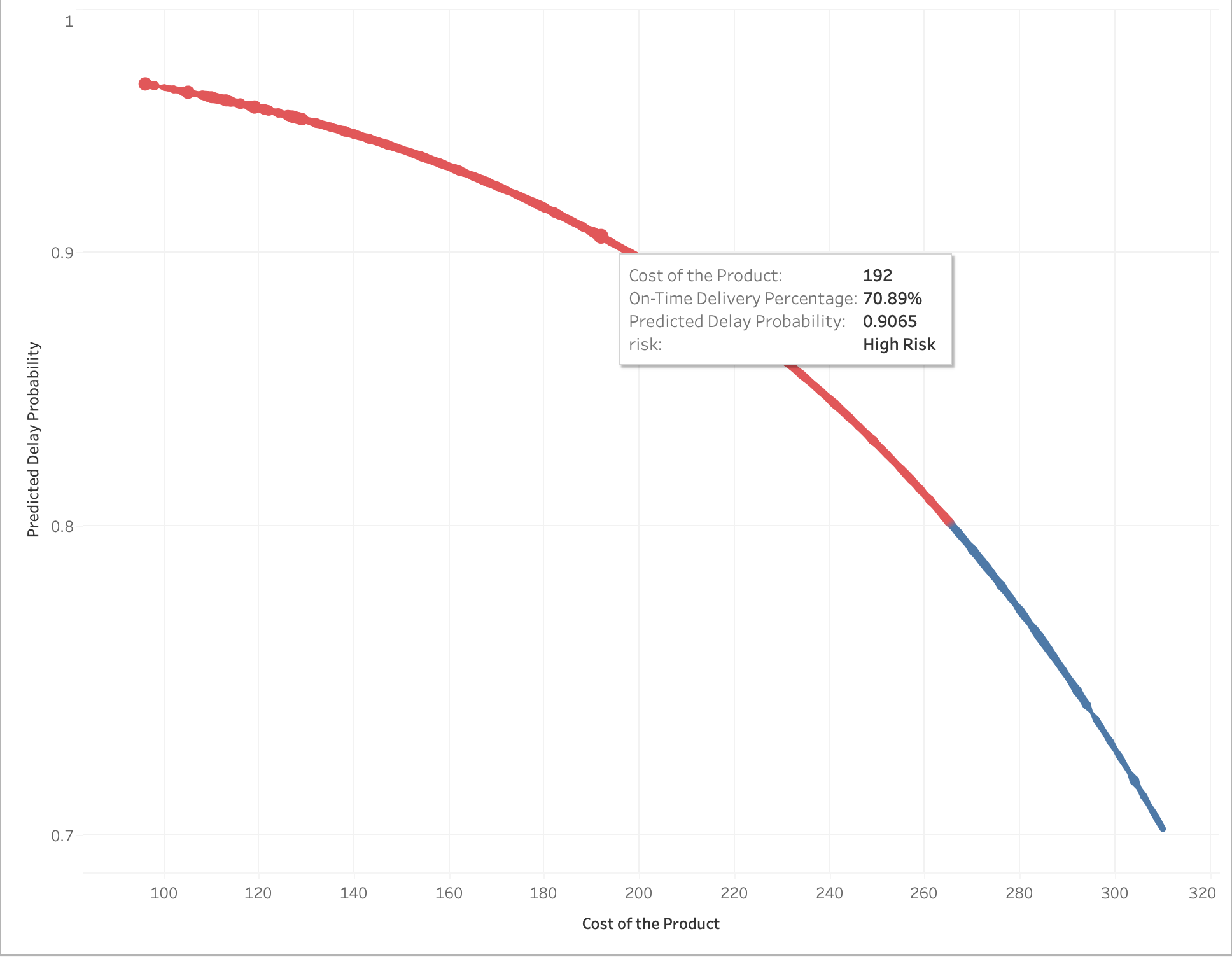


Figure 27 : Logistic Prediction

## 6.5 Risk Classification Logic

To enhance interpretability for business users, predicted probabilities were transformed into categorical risk levels using the following calculated field:

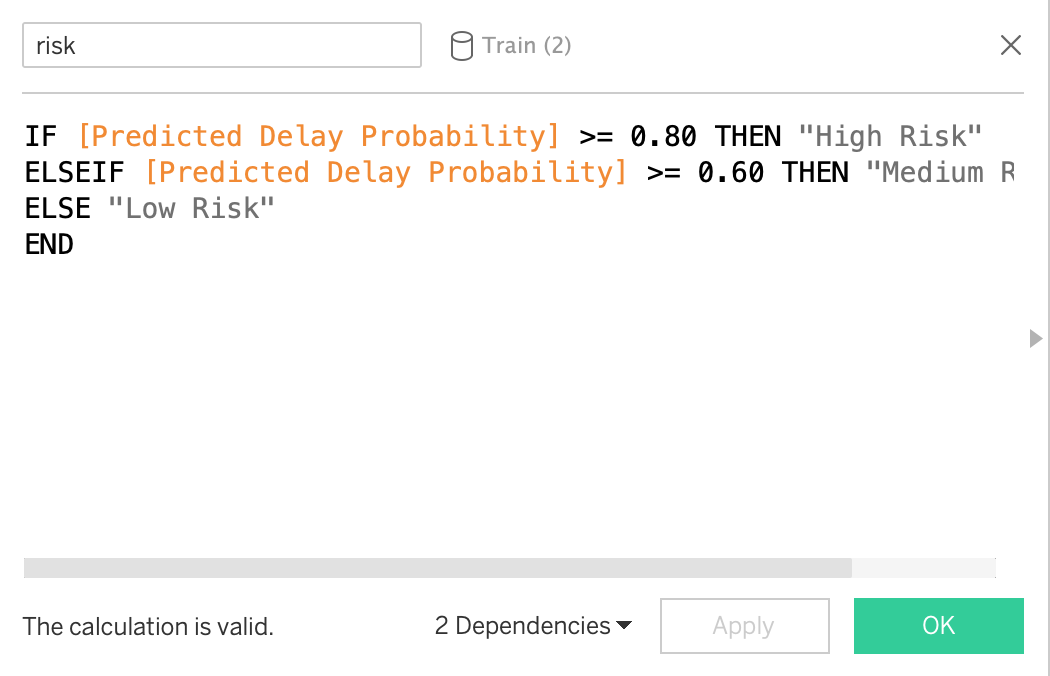


Figure 28 : Calculated Field Risk Severity

This classification allows logistics managers to quickly identify shipments requiring immediate attention refer to figure 28 above.

## 6.6 Business Value and Practical Application

The predictive dashboard enables real-time operational intelligence by allowing organizations to:

* Identify high-risk orders before dispatch
* Prioritize shipments with elevated delay probability
* Adjust warehouse workload dynamically
* Improve customer communication for at-risk deliveries
* Reduce service failures and penalty costs

This predictive capability transforms the dashboard from a passive reporting tool into an active decision-support system.

## 6.7 Story Interpretation Of Logistic Prediction

The final dashboard introduces a predictive perspective to the analytical story. By estimating delivery delay probabilities using logistic regression, this view enables proactive risk identification before shipment completion. The prediction dashboard transforms the analysis from retrospective evaluation to forward-looking decision support, completing the analytical journey from insight discovery to risk-based action planning.

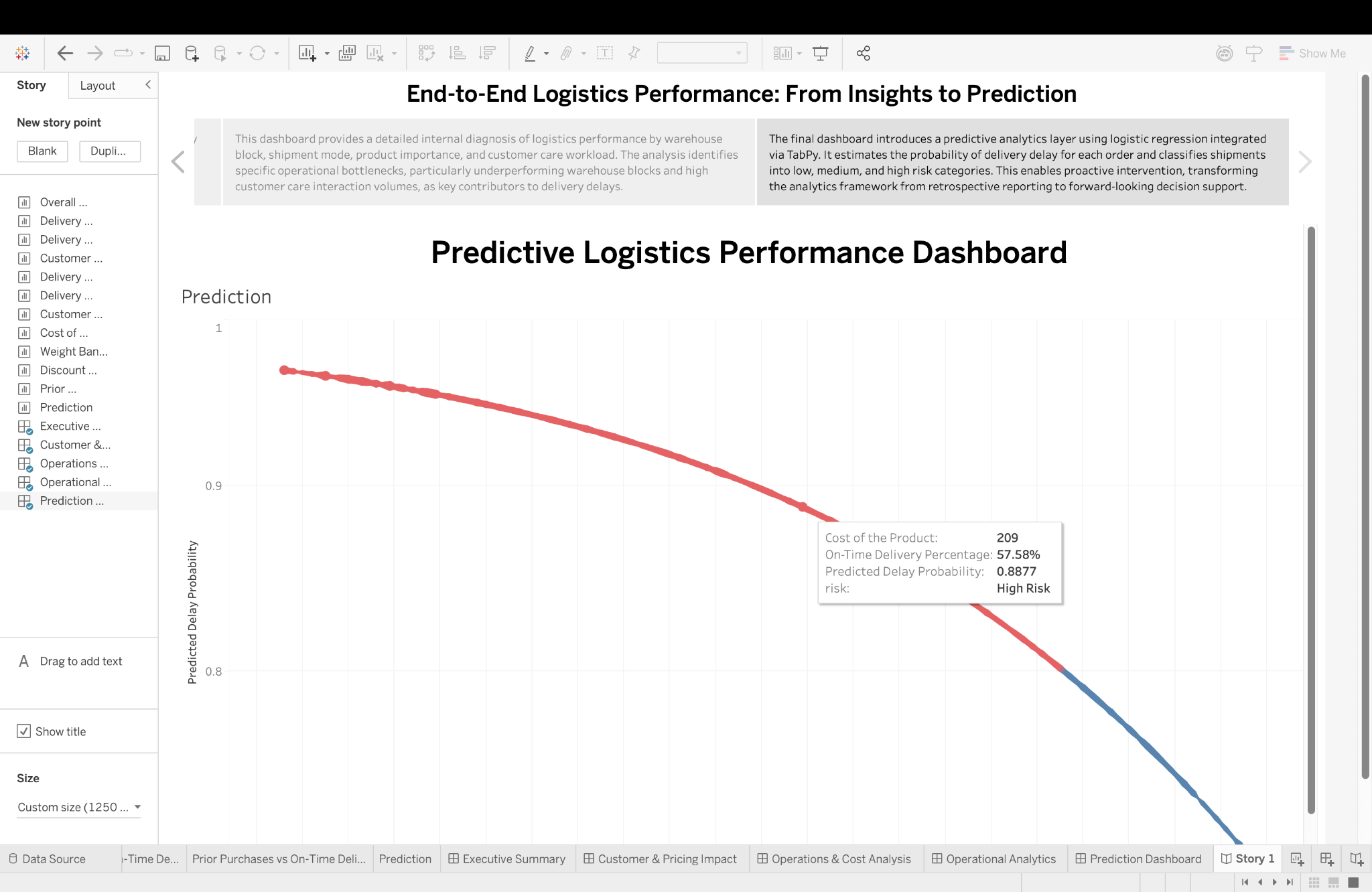


Figure 29 : Story Logistic Prediction

# CHAPTER 7: REAL-TIME INDUSTRY APPLICATION OF PREDICTIVE LOGISTICS ANALYTICS

## 7.1 Motivation for Real-Time Deployment

Although the current system is able to estimate probabilities of delivery delay using historical data, the modern-day logistics environment is increasingly adopting the use of real-time predictive analytics for operational decision-making. Real-time systems enable organizations to proactively respond to risks while they are being processed.

Real-time systems can continuously process data streams from order management systems, warehouses, transportation networks, and customer interactions.

## 7.2 Industry Practices

The major logistics companies and tech companies have started using real-time analytics frameworks that include machine learning models within their operational processes.

According to (IBM ,2023), there is a growing trend among supply chain analytics platforms to combine real-time transactional data with predictive models to predict delivery disruptions, optimize routing decisions, and allocate resources dynamically. Similarly, (McKinsey & Company, 2022) states that artificial intelligence is being widely used in logistics to predict delivery delays, optimize last-mile delivery accuracy, and increase customer satisfaction.

Industry analysts also state that predictive analytics is shifting from offline reporting to continuous event-driven decision-making systems that are integrated within logistics infrastructure (Gartner, 2025).

## 7.3 Alignment with the Proposed System

The proposed framework based on the machine learning model implemented in the project shares the same design principles:

* Logistic regression is employed to determine the likelihood of delayed delivery.
* Dynamic predictions are produced using the Python models, which are linked using TabPy.
* Interactive visualizations are used to display the predictions using Tableau dashboards.

Although the current implementation is based on batch data, the proposed framework can be extended to support real-time data processing using the following approaches:

* Establishing the connection between Tableau and data sources that support streaming data (e.g., using message queues or database systems),
* Refreshing the predictions using the TabPy server, and
* Refreshing the visualizations to reflect the latest predictions.

## 7.4 Practical Benefits

The proposed extension of the system, which would run in real-time, would provide the following benefits:

* Identification of high-risk shipments,
* Dynamic prioritization of critical orders,
* Escalation of delayed deliveries,
* Better communication with customers,
* Reduced costs through preventive actions.

In this way, the proposed system would not only serve as a monitoring tool but would become a decision-support tool.

## 7.5 Summary

Real-time predictive analytics represents a natural evolution of the machine learning framework implemented in this project. By integrating live data streams with probabilistic delay prediction models, logistics organizations can shift from reactive problem resolution to proactive risk management, improving both efficiency and service reliability.

The proposed Tableau–TabPy architecture provides a scalable foundation for such future real-time deployment.

# CONCLUSION

In this project, the power of interactive data visualization through the use of Tableau Software could be employed to transform complex logistics data into meaningful business information through the design of various dashboards with strategic, client-centered, and operational views of delivery performance evaluations.

Future logistics systems are expected to be shaped by increased automation, digital integration, and sustainability-driven transformation (Novak, Horak and Dvorak, 2024).

The analysis indicated that factors that mainly influence delivery delays are operational factors, including shipment weight, cost of products, warehouse operations, choice of transportation, and service priorities, as well as customer value and price strategy. Calculated fields, global filters, dashboard navigation, and diagnostic tooltips were utilized in service monitoring as well as in investigating roots of deviations.

Hence, in total, the visualized system offers a feasible decision-support tool, which can be utilized in identifying performance bottlenecks, planning operations, and increasing service reliability. The current project clearly demonstrates the potential of Tableau, which is generally considered a tool in data visualization, yet can in fact be utilized in analytical storytelling, a concept important in modern logistics.

# BIBLIOGRAPHY

Kaggle, “Logistics and Delivery Performance Dataset,” 2023. [Online]. Available: <https://www.kaggle.com/datasets/prachi13/customer-analytics/data>. [Accessed: Jan. 16, 2026].

Almatarneh, R., Al-Daoud, K. and Alshurideh, M. (2023) *Supply chain analytics in Tableau*. Available at:<https://www.scribd.com/document/669445836/Supply-chain-analytics-in-Tableu> (Accessed: 19 January 2026).

Kose, M. and Yildiz, T. (2024) ‘Research trends in criteria importance through intercriteria correlation (CRITIC method): A visual analysis of bibliographic data using the Tableau software’, *Journal of Information Science and Engineering*, Advance online publication. Available at:<https://www.researchgate.net/publication/382139963> (Accessed: 25 January 2026).

Patel, R. and Singh, A. (2024) ‘Artificial intelligence in logistics’, *International Journal of Logistics Management*, Advance online publication. Available at:<https://www.researchgate.net/publication/399863387> (Accessed: 20 January 2026).

Tableau Software, “Tableau Desktop and Server Documentation,” 2024. [Online]. Available: <https://www.tableau.com/learn/training>. [Accessed: Jan. 21, 2026].

Novak, P., Horak, J. and Dvorak, M. (2024) ‘Key logistics trends up to 2040 and their impact on host nation support logistics’, *Defence Logistics Journal*, Advance online publication. Available at:<https://www.researchgate.net/publication/398789735> (Accessed: 25 January 2026).

IBM (2023) *Supply chain analytics*. IBM Corporation. Available at: <https://www.ibm.com/topics/supply-chain-analytics> (Accessed: 28 January 2026).

McKinsey & Company (2022) *How AI is transforming the logistics industry*. Available at: <https://www.mckinsey.com/capabilities/operations/our-insights/digital-logistics-into-the-express-lane>(Accessed: 27 January 2026).

Gartner (2025) *Supply chain analytics trends*. Gartner, Inc. Available at: <https://www.gartner.com/en/supply-chain/insights/supply-chain-analytics> (Accessed: 29 January 2026).

# 

# 

# 

# SOURCE CODE

## Google Drive

[DRIVE](https://drive.google.com/drive/folders/19Ti2LP12-oYo4zLbWY4a74HrKhrxhdZC?usp=sharing)

## Github Link

[GITHUB](https://github.com/AbhishekAbsy0710/Tableau-logistics)