

## Analysis of Financial Messaging to Mitigate Trade Breaks

## **MFS Investment: Team 2 Capstone Project Report, Fall 2025**

Graduate MSBA students at the Manning School of Business:

*Shane Gindel,* [shane\_gindel@student.uml.edu](mailto:shane_gindel@student.uml.edu)

*Jennifer Levine,* [jennifer\_levine@student.uml.edu](mailto:jennifer_levine@student.uml.edu)

*Kheang Meng,* [kheang\_meng@student.uml.edu](mailto:kheang_meng@student.uml.edu)

*Pedro Henrique Grano Lula Pagani,* [*Pedrohenrique\_GranoLulaPagani@student.uml.edu*](mailto:Pedrohenrique_GranoLulaPagani@student.uml.edu)

**Abstract**  
*In securities trading, a trade break occurs when either party fails to deliver securities or cash by the Settlement Date. Trade breaks may result in financial loss, reputational damage, and operational inefficiencies. Managing intraday trade breaks at MFS is labor-intensive and costly. This project develops data-driven recommendations to reduce both the frequency and cost of trade breaks, focusing on MT548 SWIFT messages, which alert to trade settlement failure.*

*Two mitigation strategies are proposed.* ***Strategy One*** *targets process improvement by analyzing trades at the aggregated level, classifying them by MT548 frequency, and using a predictive model to identify custodians and brokers driving high-frequency failures. Permutation feature importance produces ranked lists of top agents, guiding targeted process-improvement recommendations to reduce MT548 frequency.*

***Strategy Two*** *addresses the operational costs of managing trade settlement failure by analyzing message-level patterns. Insights inform staffing recommendations and the development of a dashboard to filter message “noise,” improving efficiency and mitigating the cost of failure management.*

Table of Contents

**1. BUSINESS UNDERSTANDING 3**

Company Overview3

Project Scope3

Benefits & Insights for the Company3

**2. DATA UNDERSTANDING 4**

**3. DATA PROCESSING 4**

**4. METHODOLOGY 5**

a. Strategy One5

b. Strategy Two7

**5. CHALLENGES & ADDITIONAL INSIGHTS** **8**

**6. SUMMARY, RECOMMENDATION, NEXT STEPS9**

**1. BUSINESS UNDERSTANDING**

**Company Overview**  
Founded in 1924 in Boston, MFS created the first open-end mutual fund and now manages approximately $644 billion in assets. Owned by Sun Life Financial, MFS emphasizes long-term, active investing over short-term trading, with CEO Ted Maloney and President Carol Geremia advocating active management as a means to stronger oversight.

The firm is also known for active operational and technological innovation, including its proprietary Global Portfolio Modeler (GPM) system for trading and compliance. Consistent with its hands-on approach, MFS maintains rigorous oversight of the trade lifecycle, with operational teams monitoring MT548 messages to prevent trade breaks.

**Project Scope**  
This project supports the prevention of trade breaks by identifying opportunities to reduce both trade settlement failures and their associated management costs.

The project covers international trades that were successfully executed between October 2024 and September 2025.

Business understanding was critical to developing an approach to the business problem. The process by which an MT548 message leads to trade break was a topic the team returned to frequently and worked closely with sponsors to understand.

Strategy 1

Strategy 2

Mapping the mechanisms of trade break resulted in the development of two distinct strategies. Strategy one focuses on decreasing the frequency of MT548 messages by targeting upstream issues. Strategy two mitigates the labor costs of managing failure through staffing recommendations and a dashboard.

1. **Reduce MT548 Frequency**
   * Analyze drivers of high-frequency failure (HFF) trades using a predictive classification model.
   * Deliverables: ranked list of influential custodians and brokers, metrics for KPI tracking, process improvement recommendations
2. **Mitigate Trade Settlement Failure Costs**
   * Conduct time-series analysis to identify patterns in MT548 occurrences.
   * Deliverables: staffing recommendations, a dashboard for offshore teams to filter and prioritize MT548 surges

All visualizations, insights, and recommendations for this project will be compiled in a Tableau story (see Appendix B).

**Benefits & Insights for the Company**

The project highlights key custodians and brokers driving high-frequency failure trades, enabling MFS to target upstream issues and reduce settlement failures. Predictive insights support process improvements, KPI tracking, and more efficient allocation of operational resources. Temporal and dashboard analyses help to manage surges in MT548 messages, optimizing staffing and minimizing costs.

# **2. DATA UNDERSTANDING**

The data used for this project was collected by MFS from an internal SWIFT messaging system. After merging seven separate files, the dataset contained 623,394 messages, representing the communications for 163,678 international trades. The dataset consisted predominantly of equity securities.

In addition MFS provided one file containing MT548 Reason Codes and Operation Types with descriptions. This data gave context to the failure message and points to when (during which Operation) and why (the failure reason) the MT548 occurs.

A data dictionary was created to examine the initial dataset as shown in Appendix A (Table A1). From the data dictionary, the cardinality of attributes was observed and informed the decision to drop ‘Security Identifier Type’. The message date attribute contained a time parameter, which was important to note for feature engineering. Missing values were also noted for data processing.

The relationships among key attributes in the initial dataset were examined to establish the underlying data structure. Notably, a one-to-many relationship exists between *GroupID* and *MessageID*, as well as between *GroupID* and the *MT548* indicator. Each trade, represented by a unique *GroupID*, may contain multiple associated messages, and any number of these messages may be classified as MT548.

**3. DATA PROCESSING**

Initial data processing steps included data standardization, removal of out-of-scope records, and imputation of missing values. Data standardization involved the removal of leading and trailing whitespaces. Certain attribute values were identified as outside the scope of the project and were removed at the sponsor’s request. Missing values were determined to be **Missing Not at Random (MNAR)** and therefore required imputation rather than deletion (see Appendix A, Table A2).

Members of the sponsor team contributed domain expertise of the underlying dataset, which was used to develop a method of imputation. This method enabled missing values to be interpolated using related attribute values within the same observation, thereby improving overall data quality. This process underscored the value of close collaboration with sponsors to strengthen the analytic integrity of the work—a recurring theme throughout the project.

Subsequent stages of data processing involved feature reduction, feature engineering, and aggregation of the dataset to support analysis at the trade level. The Security Identifier attribute was removed due to a lack of information gain. The Message Date attribute was parsed to derive new temporal features, including *Hour of Day* and *Day of Week*. Finally, message-level records were aggregated to the trade level by counting the number of MT548 messages associated with each transaction.

**4. METHODOLOGY**

Given the dual objectives the analytic framework was structured around two complementary strategies. The CRISP-DM framework shown in Figure 1 was used to develop each strategy independently. The final CRISP-DM stage of *Deployment* was outside the scope of the project, replaced in the framework with *Deliverables*.

| **CRISP-DM Stage** | **Strategy One** | **Strategy Two** |
| --- | --- | --- |
| Business Understanding | Reduce failures | Reduce cost |
| Data Understanding | Trade-level aggregation | Message-level patterns |
| Data Preparation | Feature engineering, binning | Temporal filtering |
| Modeling | Classification | Not applicable |
| Evaluation | F1 (HFF) | Operational Improvement |
| Deliverables | Ranked List | Dashboard |

*Figure 1: Dual Strategy CRISP-DM Framework*

1. **Strategy One: Reducing MT548 Frequency at the Source**

***Objective***

Reduce the frequency of MT548 messages by identifying upstream drivers of High-Frequency Failure (HFF) trades and developing targeted process-improvement recommendations.

***Data Preparation***

Analysis for Strategy One was conducted at the trade level using the aggregated dataset created during data preparation. Message-level records were aggregated by trade, and each trade was classified into one of three categories based on the number of associated MT548 messages: no failure, low-frequency failure (1-3 MT548 messages), or high-frequency failure (4+ MT548 messages).

***Data exploration***

A histogram of MT548 frequency revealed a long tail of one-off trades with high MT548 counts (see Appendix A, Table A3). Outliers were defined by Z-scores greater than three, identifying trades which contained ten or more MT548 messages. At the sponsor’s request, these outliers were excluded from message-level analyses but retained in the predictive model, as binning rendered the model insensitive to extreme values.

Associations among categorical attributes were examined, revealing a strong relationship between Portfolio Code and Custodian (Cramer’s V = 0.837). Recognition of this multicollinearity informed the selection of tree-based models, which are less sensitive to correlated predictors.

***Descriptive Analysis***

Descriptive analytics were used to contextualize and validate model outputs. Visualizations of custodians and brokers were compared to the ranked lists generated by the predictive model. This comparison provided insight into the model’s superior ability to separate agent behavior from trade characteristics which were difficult to understand from descriptive analytics alone.

Visualizations also identified an individual broker executing a disproportionately high percentage of HFF trades relative to peers with similar trading volumes. Examination of Reason Codes showed a significantly higher proportion of *PEND//PRCY* failures. Drawing on sponsor expertise, this pattern was interpreted to indicate trades being placed on hold by the broker to avoid penalties, leading to the recommendation for a targeted process review.

Bar chart visualizations of the top ten Exchanges by trade volume were shaded by HFF percentage. This highlighted regional differences between Asian and European markets. These visualizations provided additional insights regarding the implications of penalties levied by Asian exchanges in decreasing MT548 frequency.

***Predictive Modeling***

Predictive modeling was used to identify drivers of HFF trades. Models were trained to classify trades into zero, low, or high-frequency failure bins. Class imbalance was addressed using SMOTE and weighted-average performance metrics. As the modeling objective became clear, optimization focused on maximizing the F1 score for the HFF (4+ MT548) class.

Four tree-based models demonstrated comparable performance as shown in Appendix A, Table A4. Although a Random Forest model was initially selected, feature importance analysis showed heavy reliance on the Portfolio attribute, which sponsors identified as an area with limited capacity for operational change. Consequently, a decision tree classifier was selected due to its strong performance, interpretability, and actionable permutation feature importance (see Appendix A, Table 5). To reduce computational cost, a subsample of 10,000 observations was used for hyperparameter tuning via randomized grid search, with F1 score optimization specified for the HFF class. Optimal hyperparameters are provided in Appendix A, Table 5.

Custodian emerged as the most influential feature of the final decision tree classifier—an area where MFS has greatest leverage to implement change.

***Operational Deliverables***

The predictive model generated ranked lists of custodians and brokers, showing each agent’s contribution to trades in the 4+ class. Lists were then augmented with additional relevant metrics such as the number of trades and number of MT548 messages for each agent. Further KPIs were developed including *Percentage of HFF Trades*, and *Percentage of MT548 Messages*. These rankings help to identify the primary drivers of high-frequency settlement failures, independent of other trade features. MFS can use these deliverables during annual reviews to collaborate with the identified agents, institute KIPs, and address upstream issues causing elevated MT548 frequency.

1. **Strategy Two: Reducing the Operational Cost of Managing MT548 Messages**

***Objective***

Mitigate the labor and resource costs associated with managing trade settlement failure through analysis of temporal patterns in MT548 message occurrence.

***Data Preparation***

Strategy Two analyses were conducted at the message level. Exploratory analysis revealed substantially lower message volumes during the first four months of the dataset, indicating incomplete coverage. Accordingly, all message-level analyses were restricted to the March 2025–September 2025 subset. As described in Data Processing, removal of outliers was also performed during this stage.

***Descriptive & Time-Series Analysis***

Visualizations were generated using the engineered temporal features *Hour of Day* and *Day of Week*, to identify periods of elevated MT548 frequency. These insights informed staffing recommendations.

Minute-level time-series analysis identified message peaks, defined as minutes containing 120 or more MT548 messages. Over the six-month period, 24 peaks were observed, 14 of which occurred on Sundays at 20:01. This finding was unexpected given prior temporal distribution and the general expectation that Sunday evenings represent a period of low trading activity.

Further analysis showed that approximately 40% of messages during peak periods originated from the Swiss Exchange (SIX), despite the SIX accounting for less than seven percent of MT548 messages overall. External research revealed that MT548 messages associated with the Swiss Exchange are transmitted in scheduled batch cycles rather than continuous real-time streams (SIX SIS AG, n.d.).

Additionally, SIX SIS is integrated with TARGET2-Securities (T2S), which conducts night-time settlement cycles beginning prior to the European business day (Clearstream Banking S.A. (n.d.). Preparation for Monday settlement begins on Sunday evenings between 20:00 and 22:00 CET, providing an explanation for the observed message peaks.

Further analysis revealed batch cycles contain substantial noise, including duplicates and informational MT548 messages that do not require action. Messages arrive when secondary offshore teams monitor settlement failures increasing the need to streamline the process.

***Operational Deliverables***

Based on these findings, a dashboard was developed to improve efficiency, filtering, and prioritization of messages. For example, a peak of 206 messages was reduced to 123 within the dashboard. The dashboard aggregates MT548 messages by Reason Code and trade (GroupID), flagging actionable items as “Elevate,” thereby streamlining oversight and enhancing the management of MT548 surges.

Additional deliverables include staffing recommendations based on temporal distribution analyses.

**5. CHALLENGES & ADDITIONAL INSIGHTS**

High-frequency MT548 occurrences are driven by a combination of agent behaviors and trade characteristics, complicating root-cause analysis. Predictive models can identify agent-level drivers of these failures; however, additional KPI metrics—such as those used to augment rankings—should be incorporated when implementing operational changes.

Through cross-functional collaboration, a dedicated dashboard may be developed for primary MFS operations teams, incorporating rule-based logic to distinguish actionable from informational MT548 messages.

**6. SUMMARY, RECOMMENDATION, NEXT STEPS**

**Summary**

This project addressed two related issues in MFS's trade settlement operations: reducing the frequency of MT548 messages and reducing the operational cost of managing them. Using 623,394 SWIFT messages representing 163,678 international trades (October 2024 to September 2025), analysis was structured into two complementary strategies.

The first strategy aggregated messages to the trade level, then classified trades into failure-frequency bins, and then applied tree-based classification to identify the features that were most strongly associated with High-Frequency Failure (4+ MT548) trades.

Because both custodians and brokers are actionable levers for MFS, the final approach used permutation feature importance to create a ranked list of custodians and brokers that are most often associated with HFF trades, which was then augmented with supporting KPIs to guide targeted process improvement discussions.

The second strategy approached the problem at the message level. Using temporal features and timeseries analysis patterns and peaks in messaging were identified. For this strategy, a dashboard was designed to support the offshore teams during the recurring Sunday evening MT548 peaks.

**Recommendations**

The recommendations follow the dual-strategy framework used throughout the analysis:

**1. Strategy One: Reduce MT548 Frequency**  
Use the ranked custodian and broker lists in conjunction with annual reviews, prioritizing agents driving high-frequency-failure (HFF) trades. Conduct targeted process reviews where failure patterns indicate controllable behavior, such as elevated proportions of PEND//PRCY failures.

**2. Strategy Two: Mitigate the Cost of Managing MT548 Messages**  
Align staffing with the temporal distribution of MT548 messages and deploy the dashboard to support offshore monitoring teams.

All recommendations have been integrated with visualizations into a Tableau story, presented in **Appendix B**.

**Next Steps:**

To enhance the impact of this work, we recommend integrating the dashboard with live data so teams can move seamlessly from insight to action. The dashboard may be expanded to support primary operations teams, incorporating rule-based logic to separate informational and duplicate MT548 messages from those requiring intervention. Strategy Two time-series findings should be validated as additional messaging becomes available to confirm whether observed peaks and exchange-driven patterns persist. Finally, a regular cadence should be established to refresh ranked lists and KPIs and to monitor whether targeted agent improvement reduces high-frequency-failure trade rates and MT548 volumes over time.

**References:**

SIX SIS AG. (n.d.). *IOSCO disclosure: Disclosure of SIX SIS AG regarding the CPMI–IOSCO principles for financial market infrastructures*. [Disclosure of SIX SIS AG regarding the CPMI-IOSCO principles for FMI in accordance with "Disclosure framework and Assessment methodology", December 2012](https://www.six-group.com/dam/download/securities-services/clearing/regulatory/consultation-responses/six-sis-iosco-en.pdf?utm_source=chatgpt.com)

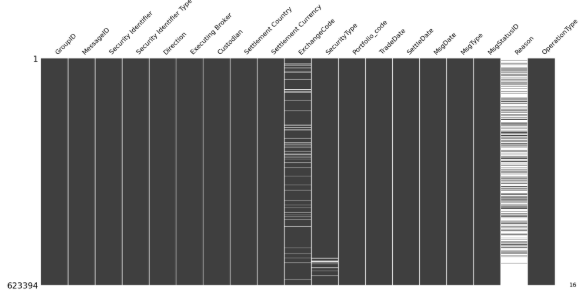
Clearstream Banking S.A. (n.d.). *Settlement process in T2S*. [CASCADE/T2S settlement cycles](https://clearstream.com/clearstream-en/securities-services/custody-and-investor-solutions/cascade-t2s-settlement-cycles-1276176)

**Appendix A**

**Table A1**: Data Dictionary

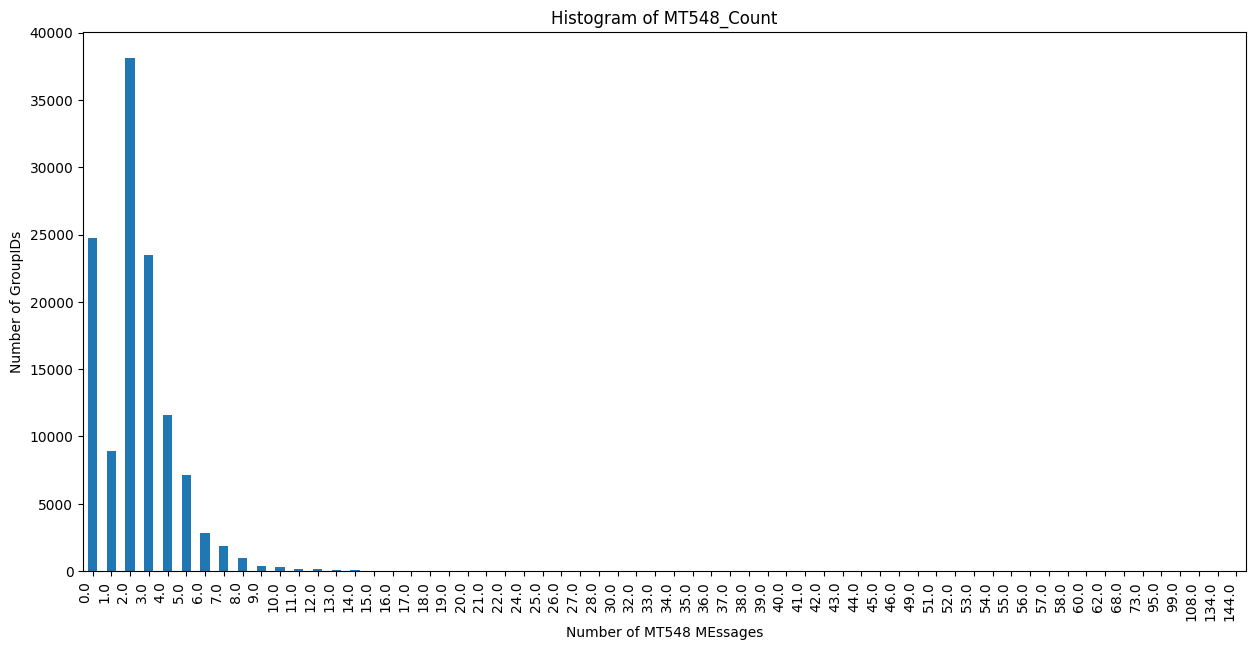


**Table A2:** Map of Missing Values



**Table A3:** Class Bin Frequency of Trades and Histogram of MT548 frequency bins

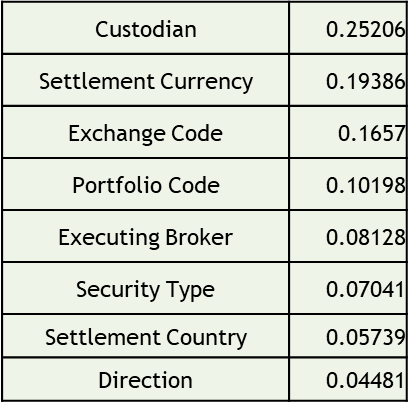
|  |  |  |  |
| --- | --- | --- | --- |
| **Failure Frequency** |  | **MT548 Bin** | **Trade Frequency** |
| No Failure |  | 0 | 24722 |
| Low Frequency Failure (LFF) |  | 1-3 | 70515 |
| High Frequency Failure (HFF) |  | 4+ | 25656 |
|  |  |  | 120893 |



**Table A4:** Model Evaluation Table

|  |  |  |  |
| --- | --- | --- | --- |
| Performance Metrics | | | |
|  | Accuracy | Precision | F-1 |
| Decision Tree | 88.46% | 88.54% | 88.48% |
| Random Forest | 89.11% | 89.21% | 89.15% |
| LightGBM | 85.02% | 86.08% | 85.22% |
| XGBoost | 86.47% | 87.25% | 86.64% |

**Table A5:** Permutation Feature Importance of Decision Tree Classifier and optimal hyperparameters



The optimal hyperparameters identified for the selected decision tree model were as follows:

* *Criterion*: entropy
* *Maximum depth*: 71
* *Minimum samples per leaf*: 1
* *Minimum samples per split*: 13