Data Quality Dashboard Metrics for Master Data in Agentic MDM

Category: Reference | **Model:** BDWP-006

Table of Contents

- 1. Introduction
- 2. Metrics Overview
- 3. Metric Calculations and Segmentation
- 4. Key Error Codes and Modules
- 5. Supporting Documentation and Validation
- 6. Monitoring Data Quality Trends
- 7. Remediation and Action Procedures
- 8. Examples and Practical Scenarios
- 9. Appendices and Resources

1. Introduction

This reference document delineates the key metrics monitored within the Agentic Master Data Management (MDM) quality dashboard. The purpose of this document is to serve as a comprehensive guide for analysts and automated retrieval-augmented generation (RAG) agents when evaluating data quality in master datasets, particularly focusing on identifier conflict rates, duplicate candidate ratios, lineage completeness, and commentary engagement levels.

Effective monitoring of these metrics enables proactive data stewardship, supports compliance with data governance standards, and facilitates timely remediation of data quality issues. The document also details the segmentation of metrics by geographic region and sub-product classifications, providing granular insights into operational performance across diverse portfolios.

Included are descriptions of supporting documents such as vendor taxonomy

updates, MDM validation checklists, error codes, and module-specific procedures, ensuring alignment with operational workflows and validation standards.

2. Metrics Overview

The following core metrics constitute the backbone of the Agentic MDM quality dashboard:

- **Identifier Conflict Rate**: Measures the frequency of conflicting or duplicate identifiers (e.g., CUSIP, ISIN, SEDOL) for the same entity across datasets.
- **Duplicate Candidate Ratio**: Quantifies the proportion of candidate records flagged as potential duplicates during de-duplication processes.
- **Lineage Completeness Score**: Evaluates the extent to which lineage information (origin, transformations, ownership) is complete and accurately documented.
- **Commentary Engagement Level**: Tracks the volume and quality of steward comments and annotations attached to master records to facilitate context sharing and issue resolution.

These metrics are computed periodically and segmented by region (e.g., NA, EMEA, APAC) and sub-product lines such as ETFs, ADRs, and mutual funds.

3. Metric Calculations and Segmentation

3.1 Identifier Conflict Rate

This metric is calculated as the ratio of conflicting identifier instances to the total number of identifiers observed in the dataset within a defined period:

```
Identifier Conflict Rate = (Number of conflicting identifiers)
```

Example: In region 'EMEA', out of 10,000 CUSIPs, 50 exhibit conflicts, resulting in a conflict rate of 0.5%.

3.2 Duplicate Candidate Ratio

Calculated as:

```
Duplicate Candidate Ratio = (Number of duplicate candidates fl
```

Example: A dataset of 100,000 records yields 2,000 duplicate candidates, for a ratio of 2%.

3.3 Lineage Completeness Score

Assessed by evaluating the presence of mandatory lineage attributes per record, with a scoring mechanism:

Lineage Attribute	Mandatory / Optional	Scoring Criteria
Source Origin	Mandatory	Present — 1 point, Missing — 0 points
Transformation History	Mandatory	Present — 1 point, Missing — 0 points
Ownership Record	Optional	Present — 0.5 points, Missing — 0 points

The overall lineage score is the average across all evaluated records, with scores closer to 1 indicating higher completeness.

3.4 Commentary Engagement Level

Tracked as the number of steward comments per record, normalized over total records, with thresholds defined for engagement levels:

- Low Engagement: < 0.5 comments per record
- Moderate: 0.5 - 1.5 comments per record
- High: > 1.5 comments per record

These segmentation methods provide actionable insights for data stewardship prioritization.

4. Key Error Codes and Modules

The following error codes are critical for identifying specific data quality

issues, each associated with modules within the MDM system:

Error Code	Description	Modules Affected	Symptoms	Root Causes	R
MDM-ID-3102	Identifier Conflict	ADR Mapping, Security Master	Duplicate CUSIP/ISIN entries	Data entry errors, inconsistent data sources	

MDM-LIN-2207	Lineage Gap	Lineage Module	Missing transformation history for records	Incomplete data ingestion procedures, manual entry omissions	
MDM-COM-1905	Commentary Required	Stewardship Portal	Records flagged for review pending comments	Lack of steward action or oversight	

Modules Affected

- ADR (American Depositary Receipt) Mapping Module
- ETF Constituent Reconciliation Module

- Security Master Data Module
- Lineage and Transformation Tracking Module

5. Supporting Documentation and Validation Checklists

Detailed validation procedures and checklists are maintained for each module and data quality dimension. These documents serve as authoritative sources for guide steps in data validation, troubleshooting, and remediation.

5.1 Vendor Taxonomy Updates

- **Bloomberg Schema (2025-08-19):** Update suffix pattern recognition for securities (.LN, .GR, .HK).
- **Moody's Sector Changes:** Zone-specific sector codes (MDY-SEC-24xx) reflecting recent classification shifts.

5.2 FactSet Concordance Rules

Rules for mapping CUSIP, ISIN, and SEDOL codes, including validation steps such as cross-referencing with authoritative external sources like *FactSet*, and procedures for handling mismatches or ambiguous matches.

5.3 Steward Override Procedures

Includes approval matrices that define who can authorize overrides, the process for requesting override approval, and documentation standards to ensure traceability.

5.4 Lineage Documentation Templates

Standardized forms and data input schemas for capturing transformation history, source origin, and ownership transfer events.

6. Monitoring Data Quality Trends

Data quality metrics are analyzed over regular intervals (daily, weekly, monthly) to identify trends and anomalies. Dashboards visualize metric trajectories segmented by region and sub-product.

6.1 Trend Analysis Procedures

- Aggregate metrics into time series.
- Apply statistical process control (SPC) techniques to detect shifts or

A mutual fund master record lacks transformation history after recent data ingestion. The lineage module flags this as a 'Lineage Gap' (error code MDM-LIN-2207). The remediation involves sourcing transformation logs from the data provider, updating the lineage record, and retesting for completeness.

8.3 Commentary Engagement Improvement

Steward comments on a security master record are sparse. An alert prompts increased steward engagement, resulting in detailed annotations, which enhance context for future audits and data trustworthiness.

9. Appendices and Resources

9.1 Glossary of Terms

- **Master Data:** Authoritative core data entities like securities, issuers, and funds.
- **Identifier Conflict:** Multiple records with inconsistent identifiers for the same entity.
- **Lineage:** The history and source tracking of data transformations.

9.2 Related Documents

- Vendor Taxonomy Updates (Bloomberg & Moody's)
- Validation Checklists for Modules
- Steward Override Procedures

9.3 External Resources

- Bloomberg Schema Updates
- Moody's Sector Codes
- FactSet Concordance Rules