



Data Quality

Scorecard

by Olga Maydanchik



DQ-09: Data Quality Scorecard

Module 0. About the Course (3 min)

Module 1. Case Studies (46 min)

- *What is Data Quality Assessment?*
- *What is a Data Quality Scorecard?*
- *Data Quality Scorecard Case Study #1: Improving the Efficiency of the Risk-Weighted Asset (RWA) Calculation Process (Financial Company)*
 - Data Quality Scorecard Goal
 - RWA Calculation Process
 - Solution Steps
 - Case Study 1 DQ Scorecard Components
 - Field Scores
 - Field Score Vs. EAD
 - List of Rules Affecting DQ Score
 - Rule Error List
 - Business Process DQ Score
 - DQ Score Decomposition By Instrument Type
 - DQ Score Decomposition By Asset Delinquency
 - Case Study 1 Summary
- *Data Quality Scorecard Case Study #2: Data Quality Impact on Catastrophe Risk Modeling (Insurance Company)*
 - Catastrophe Insurance
 - Catastrophe Risk Modeling
 - DQ Scorecard Goal
 - Case Study 2 Data Quality Scorecard Components
 - Data Quality Factor Vs. Structure Insured Value
 - The Cost of Poor Data Quality
 - Data Quality Factor Analytics
 - Rules for Data Element
 - Business Case for Data Quality Improvement
 - Case Study 2 Summary

Module 2. Data Quality Score Calculation Methods (37 min)

- *Averages Method For Score Cards*
 - Averages Method Score Calculation Example
 - Averages Score Calculations Drawbacks: Part 1
 - Averages Score Calculations Drawbacks: Part 2
- *Record-Level Score Calculations*
 - Record-Level Score Calculation Formula
 - Record-Level (RL) Close Parenthesis Score Calculation Example
- *Subject-Level Score Calculations*
 - What Are Subjects
 - Subject-Level (SL) Score Calculation Formula
 - Subject-Level (SL) Score Calculation Example 1
 - Subject-Level (SL) Score Calculation Example 2
- *Score Types Comparisons*
- *Score Decomposition By Business Dimensions*
 - Score Decomposition Examples
 - DQ Assessment Granular Results



DQ-09: Data Quality Scorecard

- Scores Calculated From Granular Results
- DQ Scores Visualization
- *Business Dimensions Versus Subjects*
 - Overview
 - HR Example
 - Manufacturing Example
 - Insurance Example
- *Module Summary*

Module 3. Data Modeling Considerations Part 1 (55 min)

- Why DQMDW?
- DQMDW Components
- Case Study DQMDW For Property Insurance Company
- *DQMDW: Critical Data Elements Catalog*
 - Technology Landscape
 - Technology Landscape - Data Example
 - Critical Data Elements - Definition
 - Critical Data Elements Example
 - Critical Data Elements In DQMDW
 - Enterprise Data Model
 - EDM & DQMDW
 - Critical Data Elements - Data Example
 - Business Process Taxonomy And DQMDW
 - Business Process Taxonomy – Data Example
 - CDE Catalog - Putting It All Together
- *DQMDW: Rule Catalog*
 - Rule Catalog - Overview
 - Rules – Threshold
 - Rules – Governance
 - Rules – DQ Dimensions
 - Rules – Life Cycle
 - Rules – Execution Type
 - Rule Error Caption
 - DQ Rule Error Output - Example
 - Rule Error Output - Purpose
 - DQ Rule Error Output - Dashboard Example
 - Analysis Example
 - DQ Rule Error Output - Where To Record
 - Rule Instance - Data Example
 - Rule Catalog - Data Example

Module 4. Data Modeling Considerations Part 2 (68 mins)

- *DQMDW: Subject Master And Business Dimensions Master*
 - Subject Master Lists - Overview
 - Dimensions Master List - Overview
 - Master Lists Vs Reference Data
 - Subject Master – Data Example
 - Business Dimensions Master – Data Example
 - Subject Master - Additional Fields



DQ-09: Data Quality Scorecard

- *DQMDW: Error Catalog*
 - Error Catalog Overview
 - DQ Assessment Results – Assessment Info
 - DQ Assessment Results – Record Info
 - DQ Assessment Results – Subjects And Dimensions Info
 - DQ Assessment Results – Data Example
 - Granular Results – Rule Versioning Decisions
 - Periodic Assessments – How Much History To Keep
 - Hierarchical Subjects – Error Treatment Decisions
- *Error Details – Storage Options*
 - Error Details Table
 - Option 1: XML
 - Option 2: Delimited Fields
 - Option 3: Multiple Rows
- *Rule Error Output – Advanced Examples*
 - One Error Affects Two Records
 - One Error Affects Many Records
 - Two Errors For The Same Subject
- *DQMDW: Score Catalog*
 - DQ Scores - Overview
 - Score History – Score Id
 - Score History – Score Type
 - Score History – Score Value
 - Score History – Object
 - Score History – Business Dimension
 - Score History – Data Example SL
 - Score History – Data Example RL
- *DataMarts For DQ Visualization*
 - DQ Visualization Steps
 - Step 1: Create Visualization Requirements
 - Step 2: To Create Screen Mock Ups
 - Step 3: Create DataMart Supporting Visualization
 - DataMart For Score Hierarchy
 - DataMart - Data Example
 - Score Catalog Summary

Module 5. Building a Data Quality Scorecard Process (61 mins)

- *Process Overview*
- *Step 1: Define The DQ Assessment And DQ Scorecard Scope*
 - How To Define The Scope
 - Why Limit The Data Volume
 - How To Limit The Data Volume
 - Why Not Use Random Data Sample
- *Step 2: Populate The Staging Area*
 - Load Data Into Staging Area
- *Step 3: Prepare The Data*
 - Data Preparation Tasks
 - Why Do We Need To Prepare Data
- *Step 3A: Add Record ID Step*
- *Step 3B: Fill Dataset AsOfDate*



DQ-09: Data Quality Scorecard

- **Step 3C: Create/Update Subject Master List**
 - Subject Master Example
 - Why Create Subject Master
 - How To Create Subject Master
 - Using Subject Master To Speed Up Development
 - Subject Hierarchies
- **Step 3D: Create/Update Business Dimension Master Lists**
 - How To Come Up With Business Dimensions
 - When To Create Business Dimension Master
- **Step 4: Perform Data Profiling**
 - Data Profiling
- **Step 5: Create and Code Data Quality Rules**
 - DQ Rules Sources
 - Coding DQ Rules
 - DQ Rules Logic Template
 - DQ Rules Logit Template Cons 1 & 2
- **Step 6: Run DQ Rules**
 - Rule Execution Tree
 - Rule Execution Tree Examples
 - Rule Execution Process Design
 - Rule Execution and Score Calculation Options
- **Step 7: Move Rule Execution Results into DQMDW**
 - Rule Results Transformation
- **Step 8: Calculate Aggregate Scores**
 - OLAP Cube for Unlimited Number of DQ Scores
 - Calc Engine for Pre-Selected Scores
 - Visualization Tool for Score Calculation
- **Step 9: Examine DQ Scores and DQ Rule Results**
 - Examining Rule Result
 - Control Charts
 - Control Charts in Data Quality
- **Step 10: Fine-Tune Data Quality Rules**
 - Rule Fine-Tuning
 - Using Profiling Techniques
 - Cyclical Nature of the Rule-Fine Tuning

Module 6. Data Quality Scorecard Demo (35 mins)

- **Overall Score Analysis**
 - Screen Look And Feel
 - DQ Score Column – Size Color Sorting Tooltip
 - DQ Score Column – Filtering
 - How Field Errors Affect Business Processes
 - Data Quality Score Prioritization Quadrant
 - DQ Work Prioritization Steps - 1
 - DQ Work Prioritization Steps - 2
 - Score Prioritization Ideas
 - Customer Volume
 - Customer Lifetime Value
 - DQ Score Timeline Column
 - Rule Details Drill Down Functionality
 - Rule Details Error Profiles
 - Demo Summary
- **Summary**



DQ-09: Data Quality Scorecard

- Key Definition
- Key DQ Score Calculation Concepts
- Typical DQ Scorecard Components and Features