Chapter13
Data Quality Management

Managing Data Quality Rules

➢ Data Quality rules should be **documented consistently, tied to business impact, backed by data analysis, and accessible to all data consumers.**

• Clear documentation ensures that rules can be implemented and monitored effectively.

Reference: Data Quality rules should align with organizational goals and be easily understood by all stakeholders.

Data Quality Dimension

➢ A **Data Quality Dimension is a measurable feature or characteristic of data.**

• Examples include accuracy, timeliness, completeness, and consistency.

Reference: Data dimensions help evaluate and monitor the quality of data in a structured way.

Data Quality Management Cycle

➢ The four stages of the Data Quality Management cycle are:
- **Plan**, **Monitor**, **Act**, and **Deploy**.
- "Improve" focuses on refining processes to address identified issues.

Reference: Continuous improvement ensures sustained Data Quality over time.

Non-Deliverable of Data Quality Management

➢ **Data attribute definitions** are **not** a primary deliverable of Data Quality Management.

• Primary deliverables include Data Quality strategies, frameworks, and reports.

Reference: Attribute definitions are part of data modeling, not Data Quality deliverables.

Non-Stage in the Data Quality Management Cycle

➢ **Intervene** is **not** a stage in the Data Quality Management cycle.

• The correct stages are Plan, Check, Act, and Do.

Reference: The Shewhart-Deming cycle guides the improvement lifecycle for Data Quality.

Non-Typical Activity in Data Quality Management

➢ **Enterprise Data Modeling** is **not** a typical activity in Data Quality Management.

• Typical activities include defining business rules, analyzing data, and identifying issues.

Reference: Data modeling is part of database design, separate from Data Quality processes.

Shewhart-Deming Cycle

➢ The stages of the Shewhart-Deming Cycle are:

- **Plan, Do, Check, Act.**
- "Investigate" is **not** a stage.

Reference: This cycle is foundational for continuous improvement in Data Quality.

True Statement About Data Quality Management

➢ **Data Quality Management is a continuous process.**
• It involves ongoing monitoring, measurement, and improvement to maintain high-quality data.

Reference: Continuous processes ensure long-term Data Quality sustainability.

Goal of Collecting Business Rules

➢ The goal of collecting and documenting business rules is **to identify the requirements for Data Quality.**

• Clear rules ensure data meets organizational needs and compliance standards.

Reference: Business rules guide the creation and monitoring of Data Quality processes.

Defining Data Quality Indicators

➢ Data Quality indicators should have **Measurability, Relevance, and Acceptability.**

- Indicators must be actionable and tied to organizational objectives.

Reference: Well-defined indicators enable effective Data Quality monitoring.

Business Relevance in Data Quality Metrics

➢ The value of a metric is limited unless **it can be linked to some aspect of a business.**

• Business relevance ensures that metrics are meaningful and actionable for stakeholders.

Reference: Aligning Data Quality metrics with business goals improves their utility and impact.

Top-Down and Bottom-Up Analysis

➤ These methods work well together because **they balance business relevance and the actual state of the data.**

• Top-down aligns with organizational goals, while bottom-up reveals the current state of data.

Reference: Combining both approaches provides a holistic view of Data Quality.

Non-Role of Data Quality Oversight Board

➢ **Data Profiling and Analysis** is **not** an expected role of a Data Quality Oversight Board.

• Their role includes setting priorities, establishing feedback mechanisms, and ensuring compliance.

Reference: Oversight boards focus on governance, not operational activities.

Non-Dimension of Data Quality

➢ **Relevance** is **not** a dimension of Data Quality according to DMBoK.

• Valid dimensions include Timeliness, Completeness, Currency, and Reasonableness.

Reference: Data Quality dimensions provide structured criteria for evaluation.

Purpose of a Data Lineage Tool

➢ A data lineage tool helps **track historical changes to a dataset.**
• It shows how data moves and transforms across systems.
Reference: Data lineage ensures transparency and traceability in data processes.

Non-Feature of Data Quality Improvement Tools

➢ **Data Modeling** is **not** usually a feature of Data Quality improvement tools.

• Features typically include profiling, parsing, transformation, and standardization.

Reference: Data Quality tools focus on improving existing data rather than designing schemas.

Data Quality SLA

➢ A Data Quality SLA includes **roles and responsibilities for Data Quality.**

• SLAs define accountability and service expectations for maintaining data standards.

Reference: SLAs are formal agreements to maintain Data Quality.

Data Quality Measurement Granularity

➢ Measurements can be taken at **data element value, data instance or record, and data set** levels.

• These levels provide detailed, intermediate, and high-level assessments of Data Quality.

Reference: Granularity ensures comprehensive Data Quality evaluation.

Data Quality Program vs. Project

➢ Data Quality Management is a **program** because it involves **project and maintenance work along with communications and training.**

• Programs address long-term goals and continuous improvement.

Reference: Data Quality programs sustain ongoing processes beyond project lifecycles.

Responsibility for Identifying Data Defects

➢ **Any employee** should identify and report data defects.
• Data Quality is everyone's responsibility, not just IT or compliance teams.

Reference: Involving all employees promotes a culture of Data Quality awareness.

Manual Directed Data Quality Correction

➢ Manual directed correction involves **using automated cleanse and correction tools with results manually checked before committing outputs.**

• This ensures that corrections are accurate and reviewed.

Reference: Combining automation with manual checks improves Data Quality outcomes.

Key Process in Defining Data Quality Business Rules

➢ The key process is **separating data that does not meet business needs from data that does.**

- This ensures that only high-quality data is used for business operations.

Reference: Defining business rules helps enforce standards for acceptable data.

Non-Goal of Data Quality

➢ The delivery of a Data Quality Strategy and framework is not a goal.

• Goals focus on defining, measuring, and maintaining Data Quality standards.

Reference: Advocacy is a direct goal of Data Quality Management.

Definition of a Root Cause

➤ A root cause is **a factor, that if eliminated, removes the problem itself.**

• Root cause analysis identifies the fundamental issue behind defects.

Reference: Addressing root causes prevents recurring Data Quality problems.

Example of Consistency in Data Quality

➢ **All the records in the CRM have been accounted for in the data warehouse.**

- Consistency ensures data is uniform and reconciled across systems.

Reference: Consistent data aligns with expectations across datasets.

Purpose of Referential Integrity

➢ Referential integrity ensures **data validity**.

• It enforces relationships between data, such as foreign key constraints in relational databases.

Reference: Referential integrity prevents orphaned or mismatched data entries.

Shewhart-Deming Cycle Sequence

➢ The sequence is **Plan, Do, Check, Act.**
• This cycle underpins continuous improvement processes in Data Quality.

Reference: The PDCA model is widely used for iterative improvement efforts.

Thank You