SOP: Data Quality Checks and Validation Rules

This SOP details the procedures for **data quality checks and validation rules**, including the identification of data accuracy, completeness, consistency, and reliability requirements. It covers methods for implementing validation rules, executing automated and manual data quality assessments, managing exceptions, and ensuring data integrity throughout the data lifecycle. The goal is to maintain high-quality data standards to support accurate reporting, analysis, and decision-making processes.

1. Purpose

To define procedures and standards for ensuring the quality and integrity of data through systematic checks and validation rules throughout its lifecycle.

2. Scope

This SOP applies to all data assets managed, processed, or reported by the organization, including structured and unstructured data sources.

3. Responsibilities

- **Data Owners:** Define quality requirements and business rules.
- **Data Stewards:** Implement validation rules and coordinate quality assessments.
- **IT/Data Engineers:** Develop, deploy, and monitor automated data checks.
- **Quality Assurance:** Oversee compliance and manage exception handling.

4. Definitions

Term	Definition
Accuracy	The extent to which data correctly reflects the real-world entities or events.
Completeness	The degree to which all required data is present.
Consistency	The uniformity and coherence of data across different sources or systems.
Reliability	The extent to which data is dependable and free from error.
Validation Rule	A logic-based constraint implemented to confirm the validity of data inputs or updates.

5. Procedures

5.1 Data Quality Requirement Identification

- Engage stakeholders to determine data quality requirements (accuracy, completeness, consistency, reliability).
- Document business rules and quality thresholds for all critical data elements.

5.2 Implementation of Validation Rules

- Develop validation rules based on identified requirements (e.g., value ranges, mandatory fields, reference checks).
- Configure rules within data capture systems (ETL processes, APIs, databases, applications).

5.3 Data Quality Assessments

- **Automated Checks:** Schedule and execute regular automated data quality scans using scripts or tools.
- **Manual Reviews:** Perform periodic manual sampling and review for areas not covered by automation.
- Log results including metrics and identified issues.

5.4 Exception Management

- Document and classify exceptions detected during validation.
- Assign responsibility for resolution and track progress in issue management tools.

- Review root causes and communicate findings to stakeholders for continuous improvement.

5.5 Data Integrity Monitoring

- Continuously monitor data for anomalies and updates that may affect integrity.
- Conduct periodic audits of data and validation processes.

6. Quality Metrics and Reporting

- Track and report data quality metrics including error rates, resolved exceptions, and compliance trends.
- Communicate results to stakeholders regularly.

7. Records and Documentation

- Maintain records of data quality assessments, validation results, exceptions, and corrective actions.
- Ensure documentation is updated whenever validation rules or procedures change.

8. Review and Revision

- Review this SOP at least annually or upon significant changes in systems or business requirements.
- Document revisions with version history and approval signatures.

Version Control

Version	Date	Description	Approved By
1.0	2024-06-01	Initial SOP issued	[Name/Title]