Data Cleaning and Validation Techniques in M&E Training Course

Professional Development Training Course Outline

Category	Monitoring and Evaluation	**Duration**	10 Days
Base Fee	$3,000 USD	**Delivery Mode**	Online / On-site Hybrid

Course Introduction

In the modern Monitoring and Evaluation (M&E) landscape, accurate, reliable, and clean data is critical for informed decision-making, program optimization, and organizational accountability. Data Cleaning and Validation Techniques in M&E Training Course equips participants with practical skills and hands-on strategies to systematically detect, correct, and validate errors in datasets. Through this course, learners will master advanced data cleaning frameworks, error-checking algorithms, and validation protocols to enhance data quality and support evidence-based reporting.

The course emphasizes a combination of theoretical insights and applied exercises, enabling participants to implement data integrity measures, automate quality checks, and perform robust validation techniques in real-world M&E contexts. Using case studies from development programs, health surveys, and social impact initiatives, learners will gain actionable skills in data standardization, anomaly detection, and data reconciliation, ensuring the delivery of high-quality, trustworthy datasets.

Programme Curriculum

Data Cleaning and Validation Techniques in M&E Training Course

Introduction

In the modern Monitoring and Evaluation (M&E) landscape, accurate, reliable, and clean data is critical for informed decision-making, program optimization, and organizational accountability. Data Cleaning and Validation Techniques in M&E Training Course equips participants with practical skills and hands-on strategies to systematically detect, correct, and validate errors in datasets. Through this course, learners will master advanced data cleaning frameworks, error-checking algorithms, and validation protocols to enhance data quality and support evidence-based reporting.

The course emphasizes a combination of theoretical insights and applied exercises, enabling participants to implement data integrity measures, automate quality checks, and perform robust validation techniques in real-world M&E contexts. Using case studies from development programs, health surveys, and social impact initiatives, learners will gain actionable skills in data standardization, anomaly detection, and data reconciliation, ensuring the delivery of high-quality, trustworthy datasets.

Course Duration

10 days

Course Objectives

By the end of this training, participants will be able to:

1. Apply advanced data cleaning techniques to large and complex datasets.
2. Identify and correct missing, duplicate, and inconsistent data errors.
3. Implement data validation protocols to ensure reliability and accuracy.
4. Conduct quality assurance checks in M&E systems.
5. Use data transformation methods for standardization and normalization.
6. Detect outliers and anomalies using statistical and software tools.
7. Utilize automated validation scripts to streamline data cleaning processes.
8. Design data validation dashboards for continuous monitoring.
9. Apply error reporting and tracking frameworks to improve program data.
10. Integrate data cleaning best practices into M&E workflows.
11. Evaluate software tools for data validation such as Excel, R, Python, and ODK.
12. Apply case-based problem-solving for real-world M&E data challenges.
13. Ensure compliance with ethical standards and data protection regulations during cleaning and validation.

Target Audience

1. M&E Officers and Specialists
2. Data Analysts and Data Managers
3. Program Coordinators and Monitoring Staff
4. Research Assistants and Field Data Collectors
5. NGO and Development Practitioners
6. Health Information Officers
7. Policy Analysts and Evaluation Consultants
8. Students and Professionals interested in data quality and validation

Course Modules

Module 1: Introduction to Data Cleaning in M&E

- Importance of clean data for program evaluation
- Common data errors in M&E datasets
- Data quality dimensions: accuracy, completeness, consistency
- Case study: Health survey dataset cleaning
- Identifying errors in sample data

Module 2: Data Quality Assurance Frameworks

- Data quality standards and protocols

- Developing data quality checklists
- Integrating QA into M&E workflows
- Case study: QA in nutrition monitoring programs
- Building a QA checklist

Module 3: Handling Missing and Duplicate Data

- Techniques for identifying missing values
- Strategies to impute or remove missing data
- Detecting duplicates and inconsistencies
- Case study: Education program survey data
- Removing duplicates using Excel and R

Module 4: Standardization and Normalization of Data

- Data formatting standards
- Transforming text, dates, and numeric fields
- Scaling and normalization techniques
- Case study: Standardizing multi-country datasets
- Normalizing survey responses

Module 5: Data Validation Principles and Techniques

- Defining validation rules and thresholds
- Cross-validation with reference datasets
- Logical and consistency checks
- Case study: Water sanitation monitoring dataset
- Implementing validation rules in Excel

Module 6: Automated Data Cleaning Tools

- Overview of R, Python, and Excel automation
- Using macros and scripts for error detection
- Automating repetitive cleaning tasks
- Case study: Automating health survey data cleaning
- Writing a Python script for duplicate removal

Module 7: Detecting Outliers and Anomalies

- Statistical methods for outlier detection
- Graphical techniques: boxplots, scatterplots
- Handling extreme values in M&E datasets
- Case study: Outlier detection in agricultural data
- Identifying anomalies using R

Module 8: Data Reconciliation Techniques

- Comparing multiple datasets for consistency
- Error tracking and correction workflows
- Case study: Reconciling household survey vs. administrative data
- Reconciliation in Excel
- Documentation for audit trails

Module 9: Data Validation Dashboards

- Designing interactive dashboards
- Key indicators for data quality monitoring
- Visualizing errors and anomalies
- Case study: Dashboard for program monitoring
- Building a dashboard in Power BI

Module 10: Ethical Considerations in Data Cleaning

- Data privacy and confidentiality
- Compliance with GDPR and local data protection laws
- Ethical handling of sensitive information
- Case study: Health data privacy in M&E projects
- Applying ethical cleaning principles

Module 11: Data Cleaning in Field Data Collection

- Field-level error detection techniques
- Training field teams on validation rules
- Tools for real-time error correction
- Case study: Mobile data collection in rural programs
- Setting up validation rules in ODK

Module 12: Integrating Cleaning with Data Analysis

- Preparing cleaned data for analysis
- Ensuring consistency in merged datasets
- Data profiling for quality assurance
- Case study: M&E report preparation from cleaned data
- Data profiling in R

Module 13: Quality Assurance for Longitudinal Datasets

- Maintaining consistency over time
- Techniques for tracking changes and updates
- Case study: Long-term health monitoring programs
- QA checks for longitudinal data
- Version control in datasets

Module 14: Problem-Solving with Real-World Data

- Identifying systemic data issues
- Designing correction strategies
- Case study: NGO program evaluation dataset
- Error correction plan
- Documentation for transparency

Module 15: Advanced Validation Techniques

- Statistical and machine learning approaches
- Predictive error detection

- Cross-dataset validation using AI tools
- Case study: Predictive validation in large-scale surveys
- Applying ML techniques for validation

Training Methodology

This course employs a participatory and hands-on approach to ensure practical learning, including:

- Interactive lectures and presentations.
- Group discussions and brainstorming sessions.
- Hands-on exercises using real-world datasets.
- Role-playing and scenario-based simulations.
- Analysis of case studies to bridge theory and practice.
- Peer-to-peer learning and networking.
- Expert-led Q&A sessions.
- Continuous feedback and personalized guidance.

Register as a group from 3 participants for a Discount

Send us an email: info@fineskilltrainingcenter.org or call *+254769199797*

Certification

Upon successful completion of this training, participants will be issued with a globally- recognized certificate.

Tailor-Made Course

We also offer tailor-made courses based on your needs.

Key Notes

a. The participant must be conversant with English.

b. Upon completion of training the participant will be issued with an Authorized Training Certificate

c. Course duration is flexible and the contents can be modified to fit any number of days.

d. The course fee includes facilitation training materials, 2 coffee breaks, buffet lunch and A Certificate upon successful completion of Training.

e. One-year post-training support Consultation and Coaching provided after the course.

f. Payment should be done at least a week before commence of the training, to Fineskill Training Center account, as indicated in the invoice so as to enable us prepare better for you.

Upcoming Scheduled Sessions

Start Date	End Date	Location	Price
23 Nov 2026	04 Dec 2026	Online	$2,000

Start Date	End Date	Location	Price
23 Nov 2026	04 Dec 2026	Physical	$3,000
23 Nov 2026	04 Dec 2026	Physical	$6,000
23 Nov 2026	04 Dec 2026	Physical	$6,000
23 Nov 2026	04 Dec 2026	Physical	$12,000
23 Nov 2026	04 Dec 2026	Physical	$9,000
23 Nov 2026	04 Dec 2026	Physical	$12,000
23 Nov 2026	04 Dec 2026	Physical	$12,000

Ready to enroll or request a customized program? Book online or contact us:

Register Online Now

Email: info@fineskilltrainingcenter.com | Website: www.fineskilltrainingcenter.com