Standard Operating Procedure (SOP)

Data Aggregation and Analysis Protocols

> **Objective:** To ensure consistent, accurate, and reliable handling and interpretation of data to support informed decision-making and maintain data integrity across projects and departments.

1. Scope

This SOP applies to all personnel involved in data collection, aggregation, analysis, visualization, and reporting within the organization across all projects and departments.

2. Definitions

- **Data Aggregation:** Process of compiling and summarizing data from multiple sources.
- **Data Cleaning:** Identification and correction/removal of errors, inconsistencies, or inaccuracies in data.
- **Data Validation:** Verification of data accuracy and quality.
- **Statistical Analysis:** Application of statistical methods to interpret data and extract insights.
- **Data Visualization:** Graphic representation of data findings.

3. Responsibilities

- **Data Analysts:** Execute aggregation and statistical procedures, visualization, and reporting.
- **Project Leads:** Oversee adherence to protocols and resolve data-related issues.
- **IT/Data Management:** Ensure secure data storage and controlled access.

4. Procedures

4.1. Data Collection

- Identify relevant data sources (internal/external systems, surveys, sensors, etc.).
- Use standardized forms or electronic data capture systems.
- Maintain data integrity via secure transfer protocols.

4.2. Data Cleaning and Validation

- Remove duplicate records; handle missing values per documentation.
- Standardize data formats (e.g., date, numeric fields).
- Run data validation checks to ensure completeness, consistency, and accuracy.
- Document all cleaning and validation steps in a data processing log.

4.3. Data Aggregation

- Define aggregation criteria (e.g., by time, location, category).
- Use approved tools (e.g., SQL, Python, R) for aggregating data.
- Summarize results using descriptive statistics or other relevant indicators.
- Retain raw and aggregated datasets as per data retention guidelines.

4.4. Statistical Analysis

- Select and document appropriate statistical methods (e.g., mean, median, regression, hypothesis testing).
- Use validated statistical software or scripts.
- Document all analytical methods, parameters, and outcomes.
- Review findings with a second analyst when applicable.

4.5. Data Visualization

- Follow organizational standards for charts, tables, and figures.
- Use approved tools (e.g., Excel, Power BI, Tableau, matplotlib, etc.).
- Ensure visual clarity, accuracy, and inclusion of titles, labels, legends, and data sources.

- Maintain consistency in color schemes and visual formats.

4.6. Reporting Guidelines

- Prepare reports according to organizational templates and requirements.
- Include sections for data sources, methodology, analysis, visualizations, interpretation, and conclusions.
- Review and approve reports through relevant stakeholders before dissemination.
- Maintain records of distributed reports as per policy.

5. Documentation and Records

- Maintain all data processing, analytical, and reporting documentation securely.
- Ensure all logs are accessible for audit and review purposes.
- Follow retention and privacy guidelines as per organizational policy.

6. Review and Updates

- This SOP is to be reviewed annually or as needed if protocols or regulations change.
- Record and communicate updates to all relevant personnel.

7. References

- Organizational Data Handling Policy
- Data Management Best Practices
- Relevant Regulatory Guidance (e.g., GDPR, HIPAA, as applicable)