

# Data Quality Report

---



# Dimensions



# General Information about the dataset

---

Dataset Name	
Main Data Source	
Receiving Date/Collection Date of the Data	
Total Number of Records in Raw Dataset	
Number of variables in the dataset	
Data Completeness Percentage (including all requested variables)	
Availability of Metadata	
Data Confidentiality Degree	
Data Consistency over time	
Data reference period	
Frequency of data collection/ receiving	

The purposes of  
using the  
collected/received  
dataset

---

**#**      **Purposes**

1

2

3

# Variable Format

---

Checking the variable format and suggesting the suitable format and modifications based on the value of data entry.

Identified Variables	Format	Suitable Format/ Modification

# Variables with Substantial Missing Value

---

Checking the extent of missing values which should be examined carefully whether they are missing at random (MAR) or missing at complete random (MCAR).

Identified Variables	% Missing	MAR, MCAR , or systematic Error	Patterns of Missing Data

# Variables with Implausible Values

---

Identification of variables with irrelevant values which defined as values resulting from data entry mistake.

Identified Variables	Irrelevant Values	Percentage of Irrelevant Values

# Variables Inconsistency

---

Inconsistency with code book includes miscoded values, values with multiple units under the same variable, and inconsistent scale of measurement, also checking the inconsistency with variables' values with other data sources which have the same attribute.

Identified Variables	Inconsistency (Variable, Data Source)	Percentage



# Constrains Violation

---

- **Record type: Uniqueness Violation**

In single observation per record dataset structure, each observation should be tagged by unique identification number.

Identified Variables	Number of Duplicated Values

---

- **Attribute Domain Constraints**

Rules that restrict allowable values in individual data elements ex. acceptable range of values.

Identified Variables	Domain Constraints Violation	Percentage

---

- **Record: Violated Attribute Dependency**

Variables computed based on other variables

Identified Variables	Variables Based on