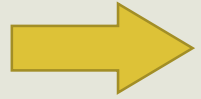# Module 9: Price Data Validation

Training Workshop on Price Statistics for the Production of Purchasing Power Parities and the Integration between CPI and PPPs in the Arab Region

07-08 June 2021

Objectives of **Module 9:** Price Data Validation

➡️ | **Purpose of Price Data Validation** |

➡️ | **Stages of Price Data Validation** |

➡️ | **Steps and Tools for each Validation Stage** |

Sampling and Non-sampling Errors

Similar to any other statistical practice, *2 types of errors* can occur during the data collection process:

1. Sampling Errors

2. Non-sampling Errors

Sampling and Non-sampling Errors
(cont'd)

1. Sampling Errors	▪ **Occurs due to the design of a sampling frame** • Should be controlled <u>before</u> price collection
2. Non-sampling Errors	▪ **Occurs during the collection and processing of price data** • Should be handled <u>before</u> and <u>during</u> price collection through survey design and management • Also, <u>after</u> price collection through a validation process of checking and correcting prices.

Price Data Validation

The purpose of the validation process is to minimize the incidence of non-sampling errors through:

1. Editing

2. Verification

Editing is the process of checking the prices for possible non-sampling errors.

Verification is the process of either confirming prices or correcting those prices flagged as possibly incorrect or incomparable.

Price Data Validation

Price data validation focuses on two types of non-sampling errors:

1. Price Error

occurs when:
- data collectors price items that match the specifications, but they record the price incorrectly.
- they record the price correctly and an error is introduced later in the process of reporting and transmitting the price.
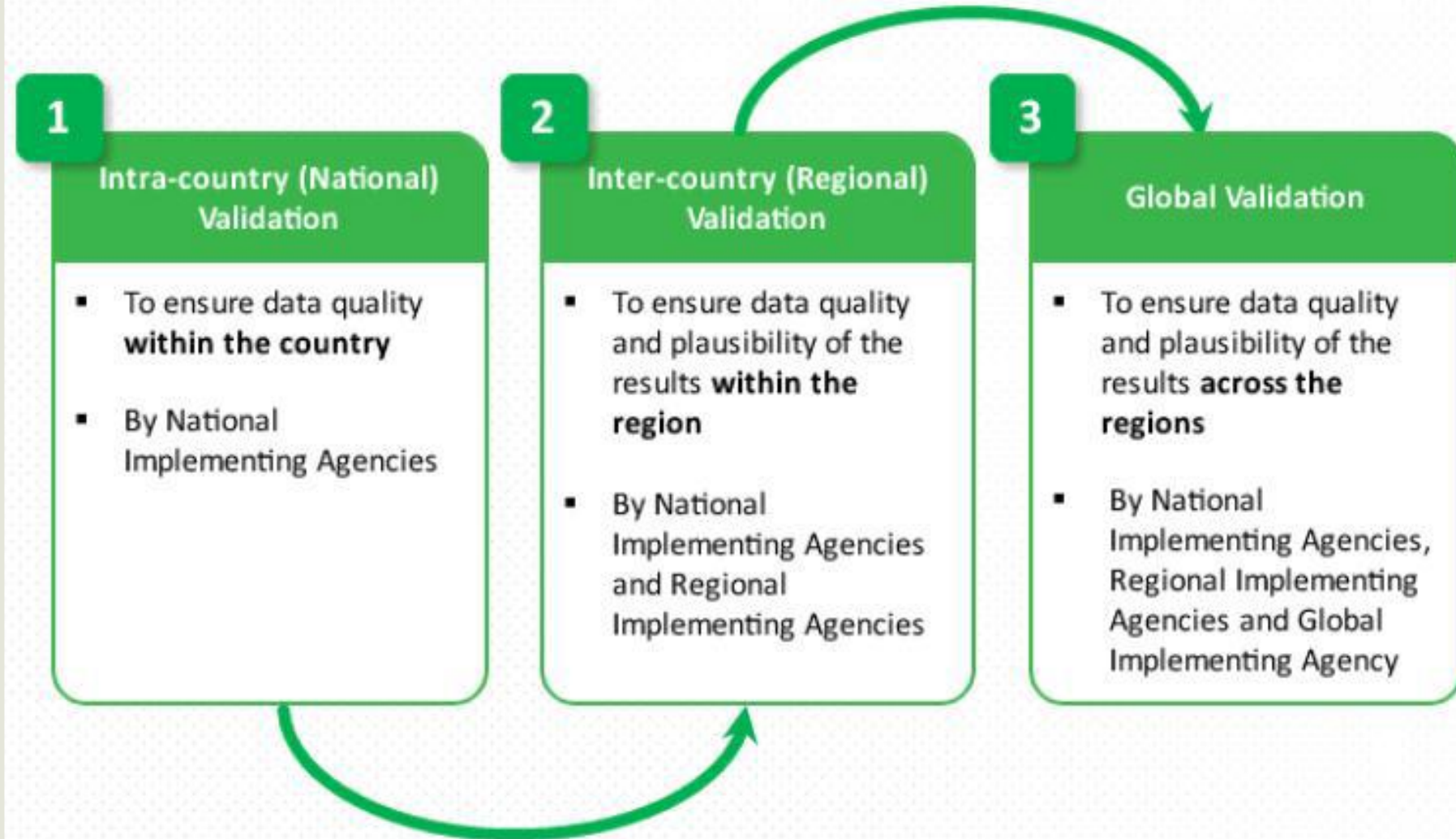
2. Product Error

occurs when:
- data collectors price items that do not match the specifications.
- they neglect to report having done so.

Data Validation Stages

Price Data Validation has *3 distinct stages.*

Data Validation Stages *(cont'd)*

1. Intra-country (National) Validation

This stage consists of *3 Validation Stages:*

- Initial Data Validation
- Statistical Tests
- Finalization of Price Data

1. Initial Data Validation
- Review of merged datasets: item codes; observed quantities and units of measurement; sub-national price variations; temporal price variations

2. Statistical Tests
- Review of average price measures: coefficients of variation; max/min ratios
- Review of individual price observation measures: standard deviations; t-values

3. Finalization of Price Data
- Preparation of average prices; number of observations; coefficients of variation

Data Validation Stages *(cont'd)*

2. Inter-country (Regional) Validation

This stage consists of *5 Validation Stages:*

- **Initial Data Validation**
- **Validation at the Basic Heading Level**
- **Aggregate Level**
- **Temporal Analysis**
- **Finalization of Data**

1.	Initial Data Validation	▪ Review for extreme data errors
2.	Validation at Basic Heading Level	▪ Comparison of averages prices and related measures across the countries (main phase of the validation stage)
3.	Aggregate Level	▪ Comparison of averages prices and related measures across the countries above a single basic heading
4.	Temporal Analysis	▪ Comparison of PPP and price level indices across time
5.	Finalization of Finalization of Data	▪ Preparation of final regional results and data for submission (after iterative rounds)

Data Validation Stages *(cont'd)*

3. Global Validation

This final stage consists of *4 Validation Stages:*

- **Analysis of the Regional Validation Tables**
- **Validation of the Global Core List Items**
- **Validation of the Global PPPs**
- **Validation of Plausibility of Global Results**

Global Validation

1.	**Analysis of the Regional Validation Tables**	▪ Review of regional quality thresholds ▪ Review for extreme data errors
2.	**Validation of the Global Core list Items**	▪ Review of representativity of global prices among regional prices; quality of global prices across the countries; variability of inter-regional linking factors
3.	**Validation of the Global PPPs**	▪ Review of degree of similarity of aggregated bilateral PPPs
4.	**Validation of Plausibility of Global Results**	▪ Review against historical results

Thank You