



INDEPTH Network

iSHARE2 Core Micro Dataset Quality

Tathagata Bhattacharjee
iSHARE2 Support Team

INDEPTH Data Repository

iSHARE 2 
an **INDEPTH** project

Outline

- Understanding the concept of Data Quality
- Data Quality Rules for Core Micro Dataset

Session duration: 15 minutes

Data Quality

- Data are of high quality "if they are fit for their intended uses in operations, decision making and planning" (J. M. Juran, evangelist for quality)
- The measurement of quality of data can be expressed by its completeness, validity, consistency, timeliness and accuracy that will make the data appropriate for a specific use
- The processes and technologies involved in ensuring the conformance of data values to a specific requirements and acceptance criteria is quality assurance



Quality Checks done in iSHARE2 ETL Process

- Ensuring the first event is legal
 - Like the first event must be **Enumeration, Birth or Inmigration.**
- Ensuring the last event is legal
 - Like the last event must be **End of Observation, Death or Outmigration.**
- Ensuring the event transitions are legal

Transition Checks

After Event (Followed By)

	ENU	BTH	IMG	EXT	ENT	DTH	OMG	OBE	None
ENU	Illegal	Illegal	Illegal	Legal	Illegal	Legal	Legal	Legal	Illegal
BTH	Illegal	Illegal	Illegal	Legal	Illegal	Legal	Legal	Legal	Illegal
IMG	Illegal	Illegal	Illegal	Legal	Illegal	Legal	Legal	Legal	Illegal
EXT	Illegal	Illegal	Illegal	Illegal	Legal	Illegal	Illegal	Illegal	Illegal
ENT	Illegal	Illegal	Illegal	Legal in case of multiple movements	Illegal	Legal	Legal	Legal	Illegal
DTH	Illegal	Illegal	Illegal	Illegal	Illegal	Illegal	Illegal	Illegal	Legal
OMG	Illegal	Illegal	Legal in case of multiple movements	Illegal	Illegal	Illegal	Illegal	Illegal	Legal
OBE	Illegal	Illegal	Illegal	Illegal	Illegal	Illegal	Illegal	Illegal	Legal

	Legal Transition
	Illegal Transition
	Legal in case of multiple movements

Quality Metric

- Quality Metric 1: % of all start events that are illegal
- Quality Metric 2: % of all end events that are illegal
- Quality Metric 3: % of all transitions that are illegal



Thank
you

