dhis2 Academy

Data Elements and Categories

DHIS2 Academy Level 1, Goa (India)

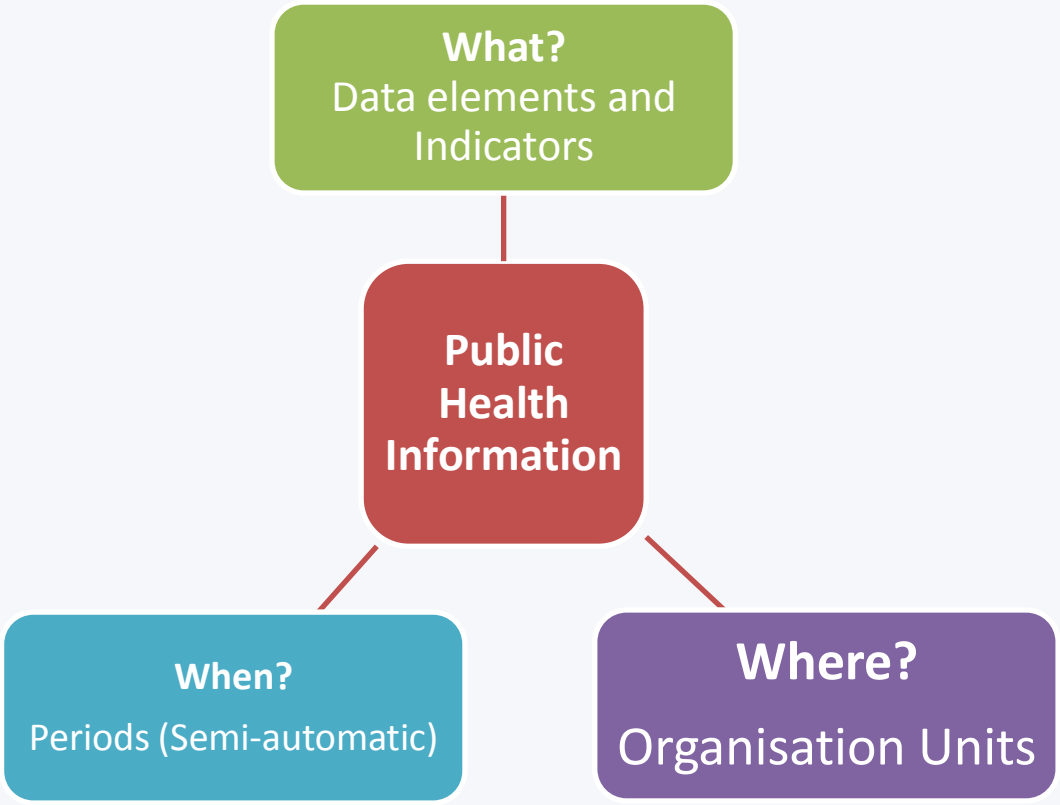
November 16 – November 24 2015

Objectives

- **Understand concepts of data elements and categories**
- **Learn how to design and customize data elements and categories in DHIS2**
- **Learn best practices and standards in design and customizing data elements and categories.**

The 3 core dimensions describing the public health information

Number of infants born to HIV Positive mothers registered in Ngelehun CHC in April 2014



Data Elements

- **Explains *what* we collect and analyse**
  - e.g. "Malaria new cases", "Total Population", "1st ANC visits", "Measles doses given".
- **Describes the raw data**, e.g. the counts and not the coverage
- **Can be further disaggregated using categories**, e.g. "Measles doses given" + <1y / >1y
- **Sometimes referred to as *Indicators*** in other contexts
  - in DHIS data elements and indicators are not the same!

Data Element Types

Routine Data Element: is data that is collected at regular intervals through mechanisms designed to meet predictable information needs: Such as Total OPD cases / total delivery in a month.

Semi-permanent Data Element: is data that changes at after relatively longer intervals of time (say annually) as compared with routine health data. For example, data on population in a district.

Infrastructure Data Element: data related to infrastructure such as number of beds in a facility, Nature of power supply, etc.

Data Element Types

Routine Data Element: is data that is collected at regular intervals through mechanisms designed to meet predictable information needs: Such as Total OPD cases / total delivery in a month.

Semi-permanent Data Element: is data that changes at after relatively longer intervals of time (say annually) as compared with routine health data. For example, data on population in a district.

Infrastructure Data Element: data related to infrastructure such as number of beds in a facility, Nature of power supply, etc.

Data Elements – Best Practices

- **The data element name should be distinctive and with meaningful description** of the data (outside the context of a form or excel table)
  - Bad examples: “Attendance”, “Number of tests”, “Positive”
  - Remember to include the information in the headings to provide full meaning of the data collected.
- **Names should not be too long** as the names are used in reports, various lists, selection boxes etc.
  - Bad example. “Number of women that are pregnant who attended their first visit of antenatal care at the facility”
    - Could be “ANC 1st visit”

Data Elements – Best Practices

- **Some times its necessary to specify where data comes from.**
  E.g. “Measles cases IDSR”, “Measles cases OPD” to separate weekly and monthly data of the same cases

- **Follow a standard naming convention**
  - E.g. don’t mix “under”, “below”, “less than”, and “<“ across data elements
  - Avoid special characters
  - Start the name with essential information, makes easier to browse information

Data Element Categories

- **Disaggregation** (dimensions to data values)

- Deciding the detail on how data values are collected and stored
  - e.g "Malaria new cases" + <5y + Male
  -

- A data element can be linked to one category combination

- A data value is linked to a categoryoption combination

Data Element Category Model

The category model has four “levels”:

1) The ***category combination***
   - “age and gender”

2) The ***category***
   - "age" , “gender”

3) The ***category option***:
   - “< 5 y” and “> 5 y” and “female”, “male”

4) The **category option combination**
   - “<5 y, male”, “<5 y, female”, “>5 y, male”, “>5 y, female”

Data Element Category Model

	<5yrs		>=5yrs	
	Male	Female	Male	Female
Malaria New cases				
Malaria Positive Cases				

DE Categories Best Practices

- **The big question to ask: What goes into the data element name and what goes into data element categories** (dimensions)?

- **Dimensions that are repeated for many data elements should be categories** (such as gender, age groups, place of service)

- **It should be possible to add together the categories/options and get a meaningful total for the data element**, e.g.

Malaria new cases = Malaria new cases (<5y, male) + Malaria new cases (<5y, female) + Malaria new cases (>5, male) + Malaria new cases (>5y, female).

DE Categories Best Practices

- **Design your dimensions with data use in mind**, not data collection.
- **Think integrated data repository and not forms or programs** when designing the metadata model and revising forms.
- **Reuse dimensions** as much as possible as it increases the ability to compare disaggregated data (e.g. age groups, fixed/outreach, gender).
- **Be consistent in your data element** + category strategy across forms, it makes it easier to combine data from different forms in the same reports etc.

DE Categories Best Practices

- Sometimes it can be tempting to break the “meaningful total rule”, and may be worth it if you can reduce the number of data elements by hundreds (e.g. large stock forms)

- The most important is to understand how categories are aggregated and only use the data element totals that make sense in reports

- A data element can only have one category combination