



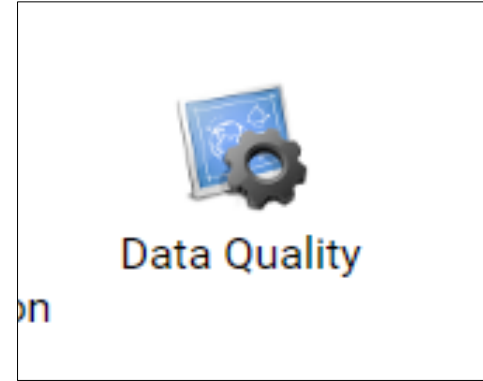
# dhis2 academy

## Data quality app – a hidden gem

Data Quality Academy On line

# DQ app and some questions

- Have you seen this app?
- Have you used this app?
- Do ALL analytic users have access to this app?

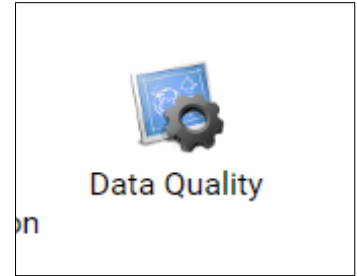


# Data Quality App - Learning Objectives

- Most common use of DQ App
- The value of validation rules
- How to translate VR into actual use
- Issues associated with DQ App

Contains the following

- Validation rule Analysis
- Standard deviation Analysis
- Min-Max outlier Analysis
- Follow-up analysis



Who is responsible ?

- ❑ They are not important - so it does not matter
- ❑ Nobody runs them any way, and if they run them, no one does anything anyway

❑ Who is responsible for **thinking** about them and then **defining** them

- HMIS DHIS2 configuration expert
- Program manager
- District Information Officer

Who is responsible for **running** them

- Data Entry clerk
- Program manager
- FIO/DIO or higher level

# Understanding the Logic of Validation Rules dhiss2

- A method to assess the quality of data entered in DHIS2
- Based on a set of predefined rules set for the data
- This does NOT allow you to FULLY assess whether or not reported numbers are completely accurate

# Validation rules – how to define

- An expression that defines a relationship between a number of data elements
- Expression has Left side, Right side and an Operator
- Common operators used
  - less than  $<=$
  - equal to  $==$
  - greater than  $>=$
- **Validation Rules** are created based on what you know to be true e.g. 'Malaria RDT tested' CANNOT be more than 'Malaria RDT positive'
- For instructions on how to set up a validation rule – look at the DHIS2 manual






# Validation rules

- Some simple guides to VR
  - Validation rules are assigned to groups
  - Easy to use and understand
  - Signs like  $\leq$  or  $=$  or  $\geq$  - does **everybody** understand them
  - Use of Importance field (Low/Medium/High)
  - Consistent guide to setup, i.e. always have smaller number first
- These **SHOULD** be run after data entry before the Complete button is ticked
- Validation rules should be run after most/all monthly data has been captured
- Who is responsible for this (F/DIO, Program manager?)
- Is there an SOP to guide the process of running these rules and then taking action?

# Validation rules - and then?

- How to run validation rules
  - Select OrgUnit
  - Select time period
  - Select Validation rule group
  - Click Validate
- Gives a report which can be downloaded into different formats
- Download into Excel

# Validation rule Output

<a href="#">DOWNLOAD AS PDF</a> <a href="#">DOWNLOAD AS XLS</a> <a href="#">DOWNLOAD AS CSV</a>							
Organisation Unit	Period	Importance	Validation Rule	Value	Operator	Value	Details
Beetle Health ...	March 2020	MEDIUM		136.00	==	133.00	
Beetle Health ...	March 2020	MEDIUM		30.00	==	32.00	
Beetle Health ...	March 2020	MEDIUM		0.00	>=	11.00	

Detail icon

VALIDATIONS RESULT DETAILS	
<b>VALIDATION RULE</b>	
Name	RMNCAH Child 5-9 with acute respiratory illness (ARI)>= child 5-9 with pneumonia
Description	
<b>LEFT SIDE</b>	
<b>DATA ELEMENT</b>	<b>VALUE</b>
RMNCAH - Child with symptoms of acute respiratory illness (ARI) 5-9 years	-
<b>RIGHT SIDE</b>	
<b>DATA ELEMENT</b>	<b>VALUE</b>
RMNCAH - Child diagnosed with pneumonia 5-9 years	11.00

# Validation rules - output

Data Quality Report							
Source	Period	Validation rule	Left side description	Value	Operator	Value	Right side description
Afghan Dispensary	March 2020	RMNCAH Women received prophylactic uterotonic immediately after birth <= facility deliveries (by age)	Received uterotonic	67	<=	0	Delivery in facility (age)
Afghan Dispensary	March 2020	RMNCAH Contraception first time user (by age and sex) = Total Contraception first time user (by method)	Contraception first time user (by age and sex)	18	==	0	Contraception first time user by method
Alsatian Primary Health Centre	March 2020	RMNCAH Total deliveries <= Live births + still births	Delivery in facility	197	<=	196	Sum of live births and still births

- Sort according to OrgUnits
- Give to supervisors/program managers for action

# Data Quality App – the rest...

- Standard deviation Analysis – now mainly WHO DQ Tool is used
- Min-Max outlier Analysis – used if Min/Max values have been set
  - If not – set first before running Min-Max outlier analysis
  - See next presentation
- Follow-up analysis
  - Used to mark data that is correct but does not fit the pattern, or triggers a validation rule

# Marked for follow-up

- Mark problematic data for follow up
- Highlight star
- SAVE - (done at Data Entry)
- Useful to explain when a validation rule has been triggered
- Report can be run on Follow-up stars

Comment 

Closure of other delivery centres, all deliveries referred here

Save comment

- What are the barriers to running Validation Rules
- Have most possible validation rules been set
- Do validation rules pick up **all** the data mistakes
- When obvious data entry errors are identified, why are they not corrected
- Who is responsible for chasing up the data entry errors
- What happens when they are not corrected

# Exercises

## **Access to own country DHIS2 OR DEMO database for this training**

- Confirm access to Data Quality App – in MOH DHIS2
- Obtain list of all Validation rules and VR Groups
- Review all Validation rules
- Review for consistency (smaller value first)
- Run at least 2 VR Groups for last 6 months
- Write a short report on rules violated
- Consider an action plan
- Do some rules need to be re-written to make them easier to understand
  - Are some rules triggered consistently
  - Is the problem with the data quality or the rule?
- Think of 1 rule to add to the current list