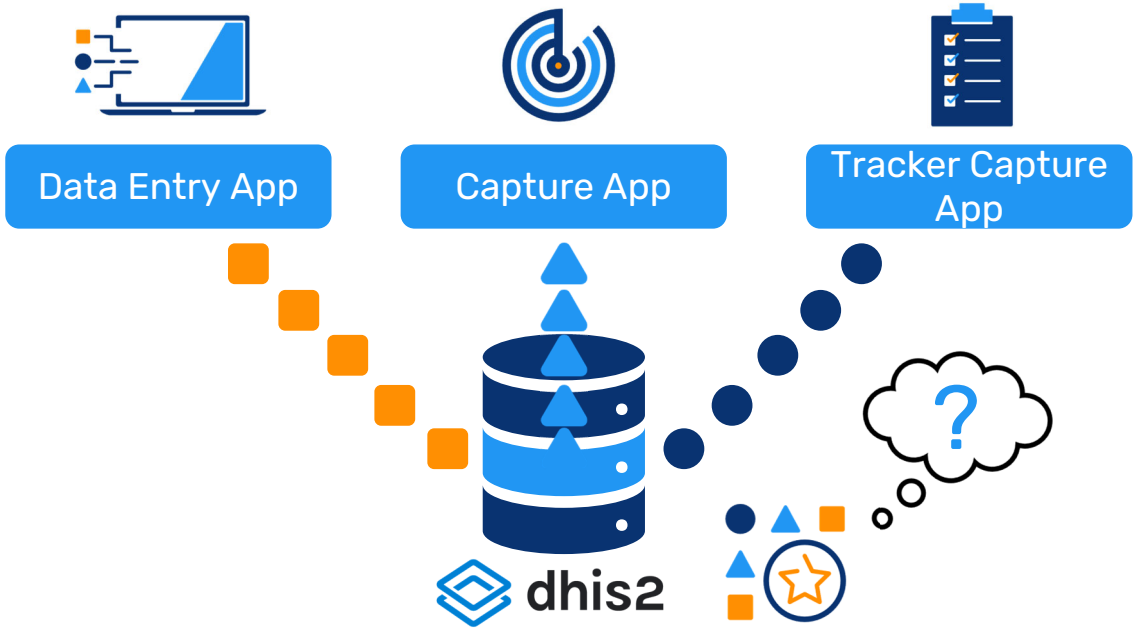
My notes

Narration



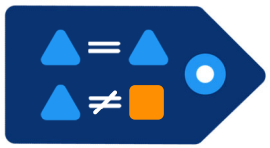
My notes

Narration

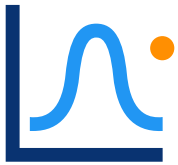
Once we have our data entered into the platform, how can we check the data's accuracy and reliability?

In this video, you will see...

Data Quality app



Validation rule analysis



Standard deviation outlier

Minimum maximum outlier analysis

Follow-up analysis

My notes

Narration

In this video, we will explore several tools in the DHIS2 data quality app that will help us to review and improve the quality of our data

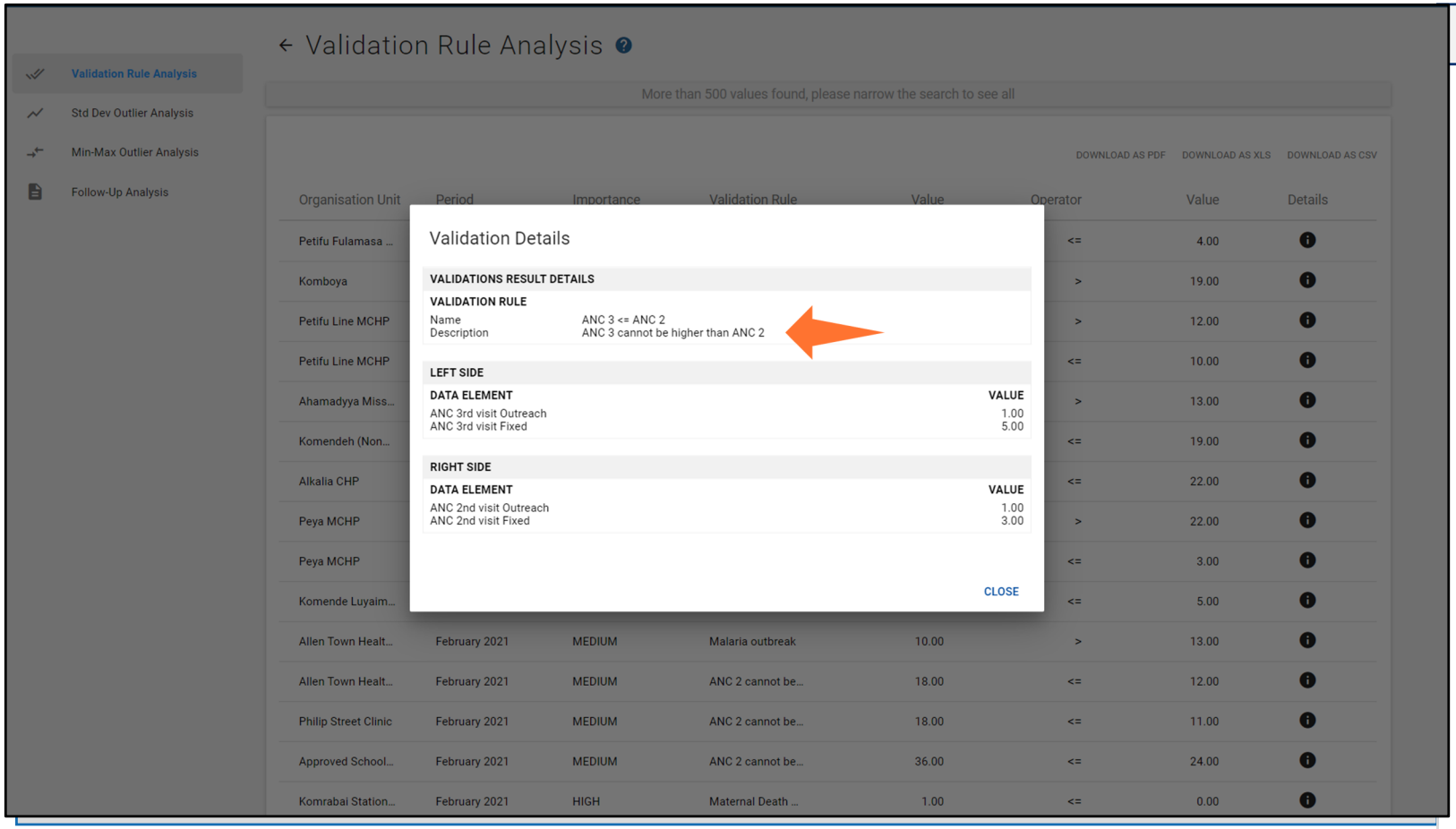
Data quality app

Validation rule analysis

My notes

Narration

First, in the Data Quality app, you can run a Validation rule analysis, which tests validation rules against the data registered in the system.



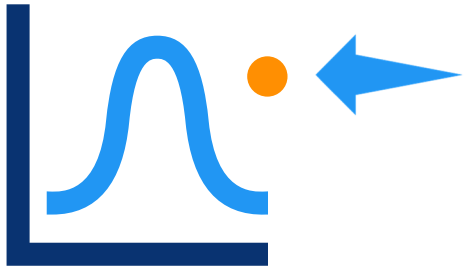
My notes

Narration

After running the check, you will get a report with a list of data that need to be checked. For example, if we look at the first row, our validation rule says that the value registered for the 3rd Antenatal care visit, or ANC, cannot be higher than the value registered for the second ANC visit. But the value registered of ANC 3 is 6, while the value of ANC2, is 4. We can follow up with this location to determine the specific source of the error.
Note that you can also run the validation rule analysis during the data entry process.

Standard deviation outlier

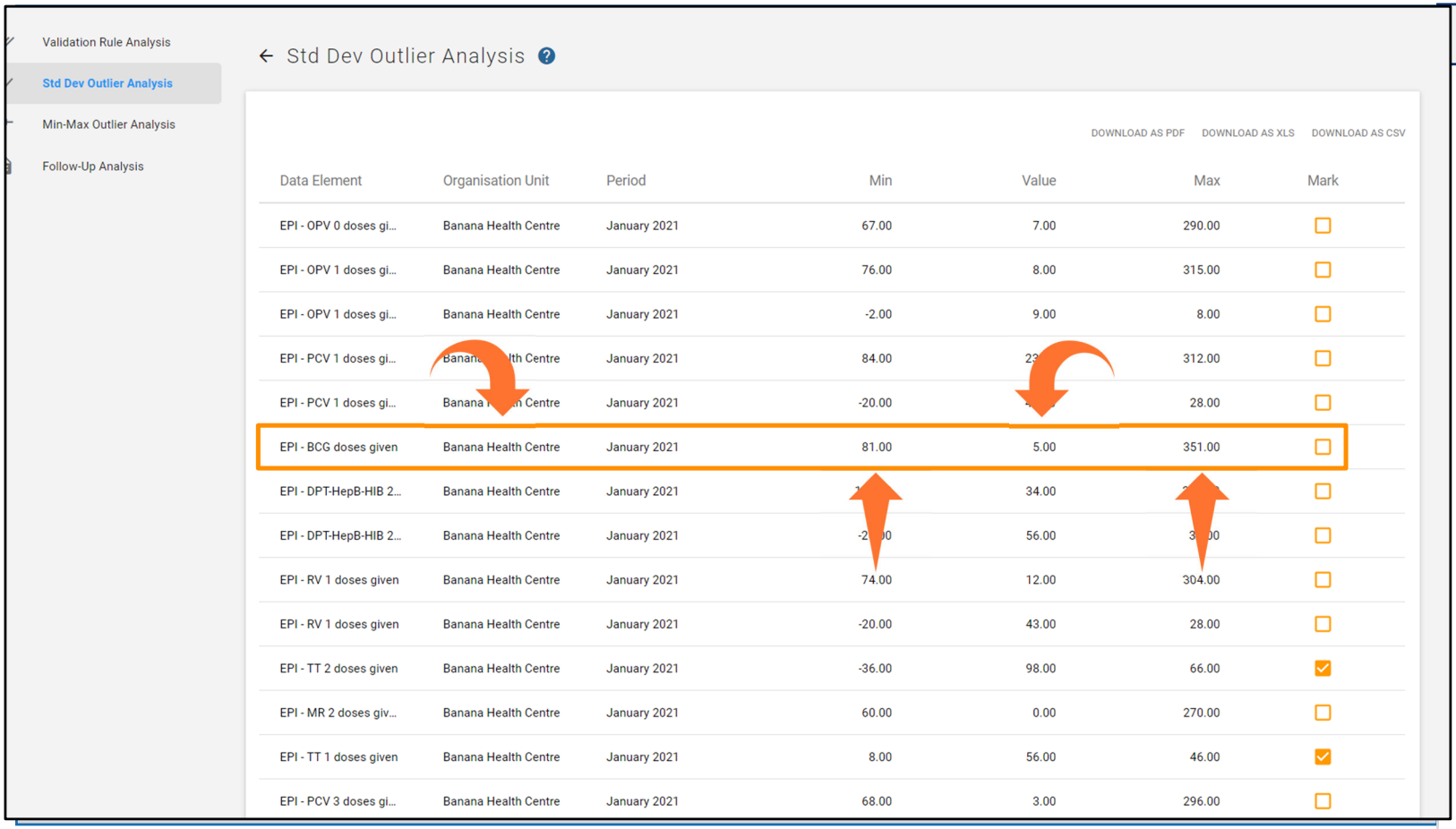
Standard deviation outlier



My notes

Narration

The Standard deviation outlier analysis is another tool we can use to check data quality, identifying those values that are potential outliers when compared to the standard normal distribution of the data we are reviewing.



My notes

Narration

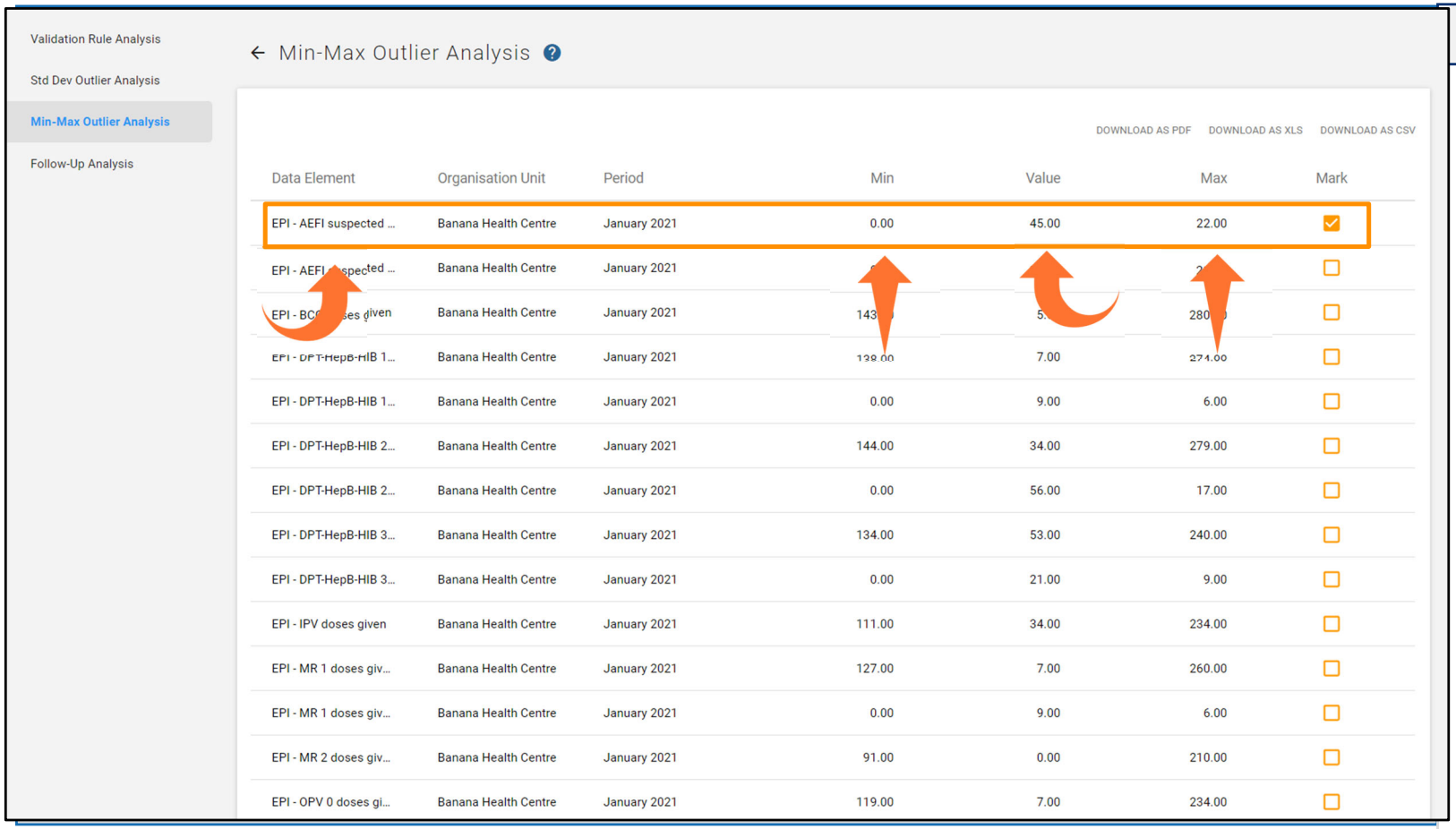
Like in the validation rule analysis, when we run a standard deviation outlier analysis, we get a report like this with all the data that need to be checked.
For example, in this row, the data indicate that only 5 BCG vaccines were given in this health facility, this value of 5 falls outside of the normal distribution of this data and has been identified as a potential source of error.

Minimum maximum outlier analysis

My notes

Narration

Another tool is the Min/max outlier analysis, which checks if the data are within a predefined range that has been specified for them.



My notes

Narration

This lists the values that are out of this predefined range. We can see in the example a value of 45 AEFI suspected cases, which is out of the predefined range of a minimum of 0 to a maximum of 22.

Follow-up analysis

Follow-up analysis

List of data values marked for follow-up

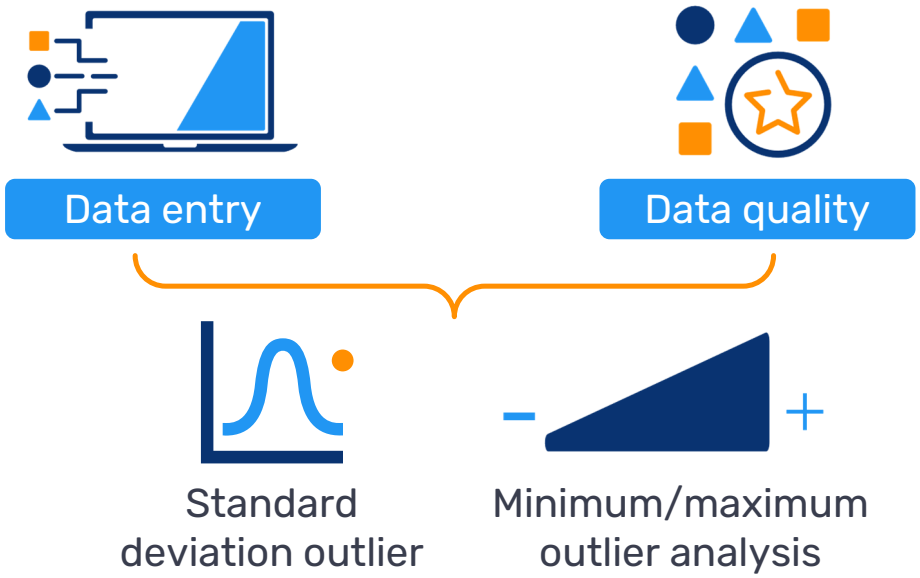
My notes

Narration

The fourth and last tool available in the Data Quality App is the Follow-up analysis tool.Follow-up analysis creates a list of all data values marked for follow-up.

Follow-up analysis

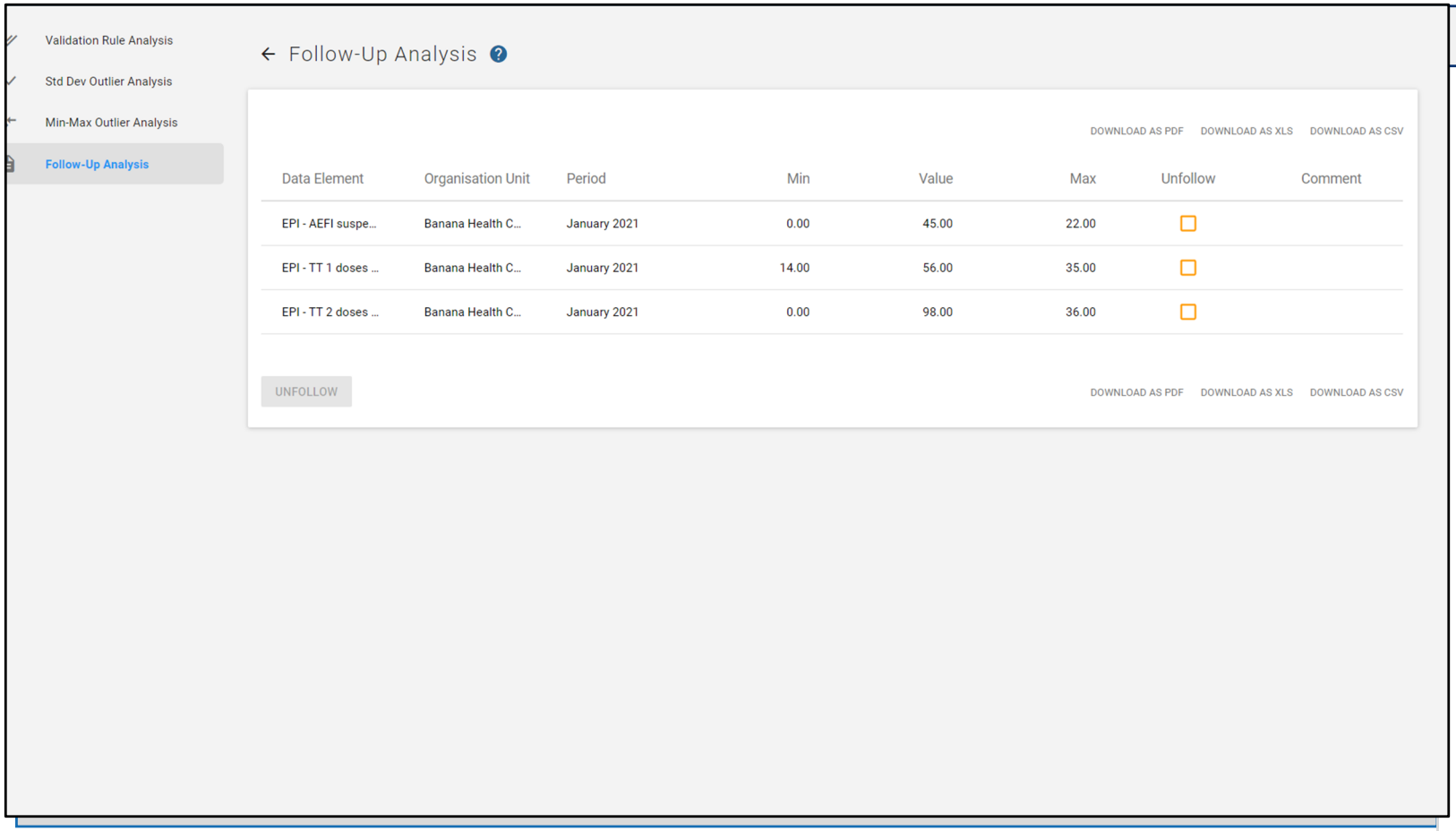
List of data values marked for follow-up:



My notes

Narration

We can mark a data value for follow-up in the Data Entry app and the Data Quality app via the reports from the standard deviation outlier and min/max outlier analysis tools.



My notes

Narration

This is an example of the report obtained after running a follow-up analysis in DHIS2.

WHO Data Quality Tool

My notes

Narration

In coordination with WHO, a Data Quality tool for DHIS2 has also been created. This app generates findings on data quality following WHO's Data Quality Review Framework.

WHO Data Quality Tool

WHO Data Quality Review Framework

- Completeness
- Timeliness
- Internal consistency
- External consistency

My notes

Narration

This includes completeness, that is, if all the expected data are recorded;
timeliness, or if the data was received on time,
internal consistency, which compares internally submitted data with one another,
and external consistency, which compares the data with other sources such as surveys

DEMO

My notes

Narration

As an example, we can review internal consistency. The WHO Data Quality tool allows us to identify outliers within our data. We can see an example of this in District C-1, where the number of Measles vaccines given in January 2020 is much higher in comparison to the other values reported within the same district in the same year.

Example of internal consistency

														Weight		
Region	Unit	Data	Jan 19	Feb 19	Mar 19	Apr 19	May 19	Jun 19	Jul 19	Aug 19	Sep 19	Oct 19	Nov 19	Dec 19	Missing	Outlier	Total
Region C	District C-1	Measles vaccine given	17555.0	1772.0	2304.0	1983.0	2126.0	1875.0	2193.0	2206.0	2424.0	1890.0	3182.0	2789.0	0	15306	15306
Region C	District C-5	BCG given < 1	4310.0	4342.0	4396.0	11665.0	3460.0	3294.0	3438.0	3199.0	3118.0	2941.0	2986.0	7570.0	0	12138	12138
Region C	District C-6	BCG given < 1	1350.0	1069.0	1434.0	1084.0	1047.0	1583.0	1223.0	1354.0	1356.0	1326.0	1407.0	4045.0	0	2751	2751
Region D	District D-1	OPV 2 given < 1	9183.0	7411.0	7840.0	6964.0	6706.0	6115.0	6686.0	6995.0	6958.0	7262.0	5998.0	7790.0	0	2208	2208
Region D	District D-1	OPV 1 given < 1	9213.0	7492.0	7866.0	7054.0	6859.0	6254.0	6805.0	7102.0	7053.0	7261.0	5907.0	7990.0	0	2155	2155
Region D	District D-1	OPV 3 given < 1	8362.0	6807.0	7172.0	6397.0	6132.0	5673.0	6180.0	6432.0	6375.0	6551.0	5463.0	7148.0	0	1968	1968
Region D	District D-1	Measles vaccine given	6508.0	6327.0	6925.0	6408.0	8139.0	6436.0	6303.0	6335.0	7217.0	5867.0	5142.0	7290.0	0	1706	1706
Region C	District C-5	OPV 1 given < 1	5421.0	4438.0	4036.0	4186.0	3892.0	2717.0	3610.0	4168.0	4031.0	3103.0	3184.0	3870.0	0	1672	1672
Region C	District C-5	OPV 2 given < 1	5327.0	4436.0	3949.0	4200.0	3886.0	2647.0	3524.0	4115.0	3966.0	3059.0	3127.0	3814.0	0	1625	1625
Region A	District A-2	Penta 3 given < 1	1493.0	1599.0	1483.0	1434.0	1139.0	1483.0	1701.0	1751.0	1417.0	1318.0	3002.0	1166.0	0	1549	1549
Region C	District C-5	OPV 3 given < 1	4906.0	4051.0	3642.0	3805.0	3517.0	2426.0	3257.0	3767.0	3655.0	2812.0	2862.0	3506.0	0	1515	1515
Region D	District D-1	Rotavirus vaccine 1 given < 1	8313.0	7409.0	7495.0	7134.0	7238.0	6774.0	6844.0	7046.0	7006.0	7233.0	5855.0	7798.0	0	1444	1444
Region D	District D-1	Penta 1 given < 1	8544.0	7743.0	8065.0	7268.0	7556.0	6982.0	6895.0	7259.0	7265.0	7551.0	6201.0	7991.0	0	1355	1355
Region C	District C-1	OPV 3 given < 1	2405.0	973.0	2169.0	2056.0	2005.0	2072.0	2473.0	2432.0	2576.0	2290.0	2213.0	2473.0	0	1315	1315
Region D	District D-1	Penta 2 given < 1	8234.0	7284.0	7587.0	7033.0	7122.0	6731.0	6655.0	7003.0	6937.0	7258.0	5914.0	7635.0	0	1311	1311

My notes

Narration

This helps us easily visualize what data need to be double-checked before running any analysis.

Summary

- Tools in the DHIS2 Data Quality app
  - ✓ Validation rules
  - ✓ Standard deviation outlier analysis
  - ✓ Minimum/maximum outlier analysis
  - ✓ Follow-up analysis
- WHO Data Quality Tool:
  - ✓ Completeness
  - ✓ Timeliness
  - ✓ Internal consistency
  - ✓ External consistency

My notes

Narration

In summary, in DHIS2 there are several tools that help us to check the quality of the data entered, such as validation rules, standard deviation outlier analysis, minimum/maximum outlier analysis, and follow-up analysis.
DHIS2 and WHO have also collaborated to create a Data Quality tool that runs checks to validate completeness, timeliness, internal consistency, and external consistency.

www.dhis2.org/academy

My notes

Narration