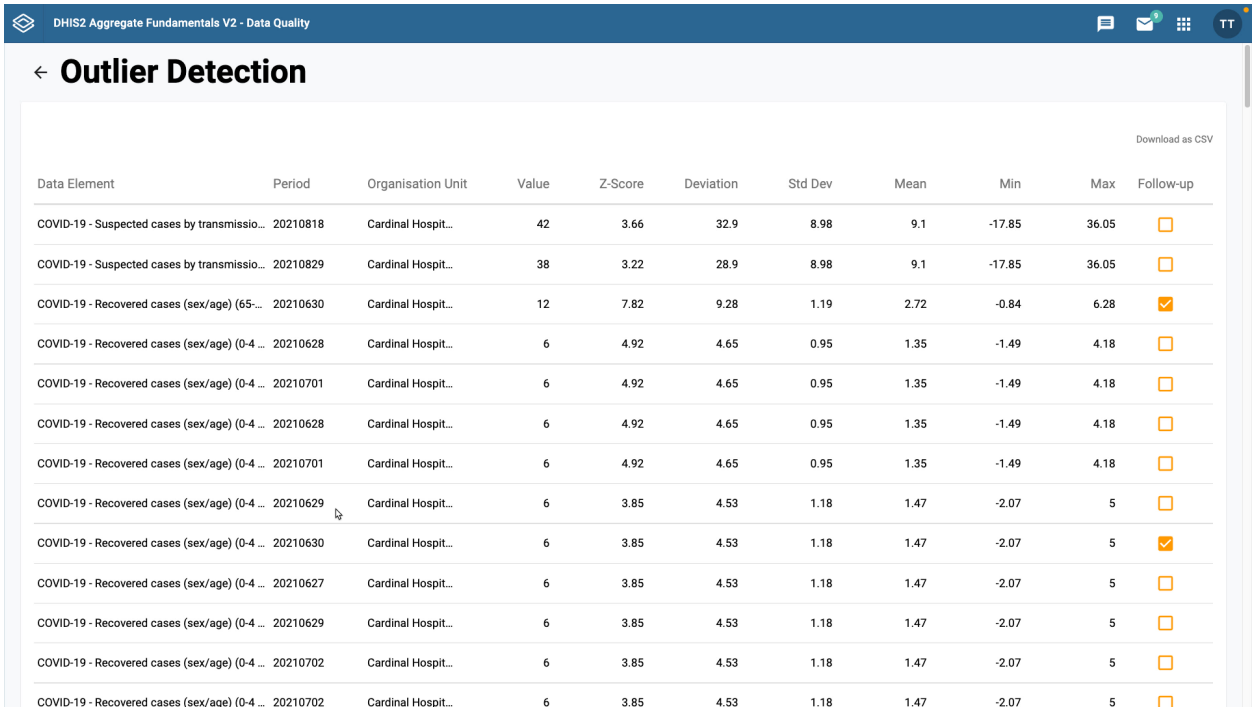
3.2.2 Outlier Detection

Overview

In this activity, you will run an Outlier Detection to check for any outliers in a data set.

Instructions



Data Element	Period	Organisation Unit	Value	Z-Score	Deviation	Std Dev	Mean	Min	Max	Follow-up
COVID-19 - Suspected cases by transmissio...	20210818	Cardinal Hospit...	42	3.66	32.9	8.98	9.1	-17.85	36.05	☐
COVID-19 - Suspected cases by transmissio...	20210829	Cardinal Hospit...	38	3.22	28.9	8.98	9.1	-17.85	36.05	☐
COVID-19 - Recovered cases (sex/age) (65-...	20210630	Cardinal Hospit...	12	7.82	9.28	1.19	2.72	-0.84	6.28	☑
COVID-19 - Recovered cases (sex/age) (0-4 ...	20210628	Cardinal Hospit...	6	4.92	4.65	0.95	1.35	-1.49	4.18	☐
COVID-19 - Recovered cases (sex/age) (0-4 ...	20210701	Cardinal Hospit...	6	4.92	4.65	0.95	1.35	-1.49	4.18	☐
COVID-19 - Recovered cases (sex/age) (0-4 ...	20210628	Cardinal Hospit...	6	4.92	4.65	0.95	1.35	-1.49	4.18	☐
COVID-19 - Recovered cases (sex/age) (0-4 ...	20210701	Cardinal Hospit...	6	4.92	4.65	0.95	1.35	-1.49	4.18	☐
COVID-19 - Recovered cases (sex/age) (0-4 ...	20210629	Cardinal Hospit...	6	3.85	4.53	1.18	1.47	-2.07	5	☐
COVID-19 - Recovered cases (sex/age) (0-4 ...	20210630	Cardinal Hospit...	6	3.85	4.53	1.18	1.47	-2.07	5	☑
COVID-19 - Recovered cases (sex/age) (0-4 ...	20210627	Cardinal Hospit...	6	3.85	4.53	1.18	1.47	-2.07	5	☐
COVID-19 - Recovered cases (sex/age) (0-4 ...	20210629	Cardinal Hospit...	6	3.85	4.53	1.18	1.47	-2.07	5	☐
COVID-19 - Recovered cases (sex/age) (0-4 ...	20210702	Cardinal Hospit...	6	3.85	4.53	1.18	1.47	-2.07	5	☐
COVID-19 - Recovered cases (sex/age) (0-4 ...	20210702	Cardinal Hospit...	6	3.85	4.53	1.18	1.47	-2.07	5	☐

Image 1 - An Outlier Detection table with identified outliers

1. Run Outlier Detection:
    a. Sign in to **DHIS2 Trainingland** using the following URL:
    b. Open the **Data Quality app** from the Apps Menu
    c. Select the **Outlier Detection** tool
    d. Select your **Outlier Detection Dimensions** and other options:

i. Select **COVID-19 Surveillance (Aggregate) Data Set** from the dropdown menu

ii. Select the **Cat District** as the **Organisation Unit**

iii. Select the **Period:** July 1, 2021-September 30, 2021

iv. Select the **Outlier Detection Method: z-score**

v. Select the **threshold** of **3** (this sets how many **standard deviations** the data are allowed to deviate from the mean before they are classified as an outlier; 3 is the standard)

vi. Select the **maximum number** of results you want to view

e. Select **'Start'**

f. Analyze the results of the Outlier Detection Report

i. What organisation units have outliers?

ii. Look at the various values, z-scores and standard deviations

iii. Are there any values marked for follow-up?

g. When you are finished analyzing the results, run another Outlier Detection Report with the same dimensions, except this time change the **threshold** to **4**. How does this impact the results?

Explore DHIS2

If you'd like additional practice in DHIS2, try these optional, exploratory steps below:

1. In the Data Quality app, use the min/max values Outlier Detection Method for 2 or more facility level organisation units in Animal Region. How does the data value

compare to the min/max limits? Which data elements would you mark for follow-up?

2. Use the z-score Outlier Detection Method to look for outliers in the Food Region or facility level organisation units. Try setting a threshold lower than the standard to see how it impacts the results