The Guide

# What this guide is intended to achieve

This guide is designed to enable an appropriately trained individual to prepare an extract report (file) exported from a Spectrum system to a ‘DHIS2 ready’ status and then initiate the DHIS2 import.

It achieves this by outlining a detailed series of ‘steps’ supported by screenshots of an actual live import process. A video has also been produced and is accessible via a link in the support section of the App.

# Protocol for making a DHIS2 import-ready file

## Step 1 – Open the extract file

There are a number of Spectrum extract file formats. The one most closely resembling the format the DHIS2 needs for import is known as the list format. The filename will normally have *DataList* appended to it.

The file will have the Spectrum Version Number in the first cell (A1), and consist of six columns of data. Each row is a single estimate for a specific location, gender and age-bracket combo (where appropriate).

|  |
| --- |
| Figure - Initial opening of a Spectrum HV Estimate extract file |

## Step 2 – Copy the original sheet

It is important to be able to refer to the original import data. To assist in the ‘preparation’ process, create a copy of the original sheet. Call the sheets:

1. Raw Data (for the original sheet);
2. ‘Data Elimination’ (for the copied sheet);

Expand the columns of the copied sheet so the values are readable. Note that the ‘Indicator’ column includes the details of the age-bracket and gender, while the ‘population’ is represented by two columns – ‘Country’ and ‘Subnational region’.

|  |
| --- |
| Figure – Create a copy of the original sheet |

## Step 3 – Eliminate Totals

The standard Spectrum extract format format typically includes Totals for each combination of disaggregation’s (e.g. the totals for male + female for all age-brackets for a give indicator). Totals are also included for each sub-region as well as a national total.

Since DHIS2 only requires data at the lowest level of granularity (i.e. the disaggregated values), it is not necessary to import indicator totals. DHIS2 will automatically generate aggregates as part of its standard data warehouse capability.

Make sure to identify and delete the Totals!

|  |
| --- |
| Figure – Delete all ‘Totals’ (e.g. Male+Female) as disaggregated values will be summed in DHIS2. |

In the sample file used for this Guide, the following Population rows were also identified and eliminated:

* Population aged 0-4 Total Region; Male
* Population aged 0-4 Total Region; Female
* Population aged 5-14 Total Region; Male
* Population aged 5-14 Total Region; Female
* Population aged 15-24 Total Region; Male
* Population aged 15-24 Total Region; Female
* Population aged 15-64 Total Region; Male
* Population aged 15-64 Total Region; Female
* Population aged 65+ Total Region; Male
* Population aged 65+ Total Region; Female

Leaving

* HIV population (15-49) Total Region; Female
* HIV population (15-49) Total Region; Male
* HIV population (15+) Total Region; Female
* HIV population (15+) Total Region; Male
* Mothers needing PMTCT Total Region; Female
* Calculated number needing adult ART (Dec 31) Female
* Calculated number needing adult ART (Dec 31) Male

Note that while the word ‘Total’ is still in the label for the indicator, these are in fact disaggregated estimates (by age-bracket, gender and subnational grouping).

## Step 4 – Copy the Data Elimination Sheet

As a precursor to the next step, copy the Data Elimination Sheet (after elimination of Totals rows in the previous Step) to a new sheet called Data UIDs.

|  |
| --- |
| Figure – Copy Data Elimination sheet (after removal of total rows) to a new sheet labelled Data UIDs. |

## Step 5 – Replace Indicator labels with UIDs

Each ‘indicator’ in the Spectrum file should match with one of the DHIS2 ‘data elements’ that were setup during the Bootstrapping initiation process. Each of the DHIS2 data elements has a unique UID. The preferred import process for DHIS2 is to identify each item in the import file with a UID to ensure data is properly linked to the right data element (chapter 21 of the DHIS2 User Guide)[[1]](#footnote-1).

### Step 5a – Identify the data element UIDs.

The process below shows the manual process of identifying the data element UIDs. While some of the UIDs are already known (since they were standardized as part of the bootstrapping process), other UIDs are unique to the DHIS2 instance and must be identified manually.

|  |
| --- |
| Figure - Select the Data-Element/Indicators DHIS2 App |
| Figure - Select the Data Element option |

|  |
| --- |
| Figure - Find in the list of data elements those beginning with 'UNAIDS: ...' and click on the appropriate data-element. From the displayed dropdown menu, select 'Show details' |
| Figure - Note down the ID for the specific data element. It will be used in the import spreadsheet. |

Note: The default HIV Estimate Denominators implemented by the App in DHIS2 will use the same UIDs. These are listed below for convenience. It will be necessary to use the above method to identify local organizational subnational region UIDs.

|  |  |
| --- | --- |
| Denominator | UID |
| HIV Den: Estimated number of HIV+ pregnant women | PjLBZcVwRnr |
| HIV Den: Estimated number of people (15+) eligible for ART according to national guidelines | LEdbdocVhvx |
| HIV Den: Estimated number of people (15+) living with HIV | xAihkVcBj4F |
| HIV Den: Estimated number of people (15-49) living with HIV | rhXstKVfvvj |
| HIV Den: Estimated number of people (<15) eligible for ART according to national guidelines | khe0fQWys0p |
| HIV Den: Estimated number of people (<15) living with HIV | F4KpUnnx0S5 |

### Step 5b – Insert 2 new columns

Each Indicator in the spreadsheet is represented in DHIS2 by at least 1 and sometimes more UIDS in DHIS2. One UID must always be present (the one representing the indicator) while the other UID will represent disaggregation (e.g. gender).

To support these UIDS, insert two new columns into the ‘Data UIDs’ sheet:

1. ‘dataelement’; and
2. ‘categoryoptioncombo’

These labels are important as they are used as placeholders to identify which columns hold the data needed by DHIS2, and the meaning of the columns.

|  |
| --- |
| Figure - Inserting two new placeholder columns for dataelement and categorryoptioincombo |

### Step 5c – Insert the matching ‘dataelement’ UIDs for all Indicators

Note that the same UID will be used across various combinations of disaggregation (e.g. in the example below, the same UID has been used across the subnational regions and for the male/female gender), as the same conceptual ‘indicator’ is being referred to in all rows.

Other columns (e.g. the categoryoptioncombo) will be used to further identify the specific gender disaggregation being referred to.

|  |
| --- |
| Figure - Insert into the ‘dataelement’ column the relevant UID representing the specified Indicator |

### Step 5d – Insert the local gender UIDs (specific to the local instance)

Several of the Spectrum data elements in the import file are disaggregated by gender. It is necessary to identify the UID for the Male and Female Combinations in DHIS2 and insert them into the categoryoptioncombo column as appropriate (using the indicator ‘label’ as the guide to which UID goes in which row).

|  |
| --- |
| Figure – Select the Category Option Combination menu item under ‘Data Elements’ in DHIS2 |
| Figure – 1. ‘Show Details’ of the Male/Female Category Option Combination; and 2. Copy the IDs |
| Figure - Insert into the ‘categoryoptioncombo’ column the relevant UID representing the specified gender |

## Step 6 – Insert an Organizational Unit Column, identify and insert the appropriate UID for the given Country/Sub-national Level

The Spectrum extract files represents location for the given estimate by Country and Subnational Region columns. In DHIS2, a single object (UID) represents a location, and they are arranged into what is called an organizational hierarchy. It should be noted that in most DHIS2 country instances, the breakdown of the geographical hierarchies mirrors the typical geopolitical hierarchies of the country (e.g. national, regional, district etc.).

It is assumed that the Spectrum extract will also use an identical breakdown, though the ‘labels’ between the two systems are likely to have different spellings. The task at this stage is to ‘match’ the Subnational Region in the Spectrum file with the equivalent organizational unit in the DHIS2 hierarchy, and extract the UIDs to ensure correct ‘linking’ of estimates during import.

The task is similar to previous steps (i.e. find UIDs and insert as appropriate in the Spectrum extract file).

### Step 6a – Insert an ‘orgunit’ column.

|  |
| --- |
| Figure – Insert an ‘orgunit’ column between ‘Subnational Region’ and ‘Country’ |

### Step 6b – Select the ‘Organisational Units’ App and Menu Item.

|  |
| --- |
| Figure – Select the ‘Organisation Units’ app from the Apps menu in DHIS2 |
| Figure – Select the Organisational Unit menu item to display the Organisational Hierarchy |

### Step 6c – Identify UIDs for each Subnational Region listed in the Extract

In our example, we’ve used the ‘TrainingLand’ demo country. On the left-hand side of the screen highlight the ‘top’ item in the geographical tree (normally the country name). When you do this, the item you clicked on will turn orange and list he ‘children’ in the main working area. When you click on one of these ‘children’, the ‘show details’ menu item will allow you to obtain the UID for that specific child (region).

|  |
| --- |
| Figure – Highlight the ‘country’ item at the left (turns orange), then ‘Show details’ of the subnational region of interest in the main area of the screen. |
| Figure – Copy the desired ID |

|  |
| --- |
| Figure – Insert the appropriate UIDs into the ‘Orgunit’ Column, using the ‘Subnational Region’ labels as a guide. |

## Step 7 – Sort columns into the appropriate order ready for import.

The following sub-steps should be executed resulting in a final DHIS2 import-ready extract file.

### Step 7a – Do a final check then copy to new sheet

After a final check that all ‘Indicator’ labels have an equivalent ‘data element’ ID, that ‘Indicator’ labels that have gender have an equivalent ‘categoryoptioncombo’ ID, and that every ‘Subnational region’ label has an equivalent ‘Orgunit’ ID, copy the worksheet you’ve been working to a new sheet and call it ‘4. Data Order’.

### Step 7b – Change the following column names

* ‘Year’ becomes ‘period’
* ‘Estimate’ becomes ‘value’
* ‘Filename’ becomes ‘comment’

### Step 7c - Delete Extraneous Columns

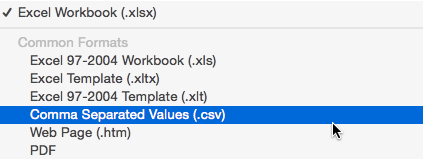
* Delete ‘Subnational region’ and ‘Country’ columns
* Delete the ‘Indicator’ column

### Step 7d - Arrange columns in the following order (insert blank columns as necessary):

1. dataelement,
2. period,
3. orgunit,
4. categoryoptioncombo,
5. attributeoptioncombo *(blank)*,
6. value,
7. storedby *(blank)*,
8. timestamp *(blank)*,
9. comment,
10. followup *(blank)*

|  |
| --- |
| Figure – Import ready version of the Spectrum Extract file |

## Step 8 – Save the final sheet as a CSV formatted file



Importing and Testing

# Protocol for importing and testing the Spectrum

## Step 1 – Go to the Import-Export App in DHIS2

DHIS2 has a comprehensive data and metadata import and export capability built into the Administrative front-end of the system. Imports of data into DHIS2 can occur in several formats (JSON, XML, CSV etc.), and CSV is the format we’ve selected in this Guide (because of the ease and familiarity most people have with MS Excel).

### Step 1a – Go to the Import-Export App and select Data Import

|  |
| --- |
| Figure – Use the DHIS2 ‘Apps’ menu item to locate the ‘Import-Export’ App |
| Figure – Select the ‘Data Import’ item |

The Data Import screen will then appear.

## Step 2 – Do a ‘Dry Run’ import test

The ‘Dry Run’ option allows you to run the full process of import (with all it’s internal checks and balances) without actually importing the data. This is an important step, because it highlights any rows of data that will be rejected amongst other errors in the data.

The most common error will be an incorrect UID (i.e. the value doesn’t exist or could not be found in DHIS2). This row would be rejected because it cannot be successfully linked to the specified Organisation Unit, Data Element, or Disaggregation (i.e. gender ID).

### Step 2a – Adjust the Data Import Parameters

The starting screen should look similar to the image below.

|  |
| --- |
| Figure – Starting Point for Data Import Screen |

From the starting point highlighted above, do the following

1. Select the CSV file created as a result of the steps in Section 2;
2. Change the format to CSV;
3. Select Yes for the ‘Dry Run’ parameter.

Here’s how it should look just before running the import!

|  |
| --- |
| Figure – Example screenshot of the parameters filled in and ready for a ‘dry run’! |

### Step 2b – Run the Import

1. Run the import (leave the Strategy as ‘New and Updates’).

As the import process is running, a process log will indicated progress.

## Step 3 – Review the ‘Dry Run’ Errors

When the ‘dry run’ import process is complete, the last item in the progress log will say ‘import done’ with an option to ‘Display Import Summary’

### Step 3a – Display Import Summary

From the ‘top’ of the process list, click on the link ‘Display Import Summary’.

*Note that the ‘import parameters’ are reset once the ‘Import’ button is clicked.*

|  |
| --- |
| Figure – The progress log with completed process steps after a ‘Dry Run’ has been initiated. Note the ‘Display Import Summary’ link. |

### Step 3b – Review Import Summary

The Import Summary provides two types of information:

1. Counts of the type of import status (i.e. ‘imported’, ‘updated’ or ‘ignored’); and
2. A list of the ‘conflicts’ for ‘ignored’ rows of data.

Note: that the ‘header row’ in the CSV file (i.e. the row that included the names of the columns) will show up as part of the ‘ignored’ count and will have an associated ‘conflict’ (e.g. the dataelement label in cell A1 … being head label rather than the UID of a data element will generate a conflict *of ‘Data Element not found or not accessible’*).

Note: In some instances of DHIS2, countries decide not to record ‘zero’ counts. If this is the case, values of zero in the import data will also be ignored, and generate a conflict message (e.g. ‘*Value is zero, and not significant…*’).

|  |
| --- |
| Figure – The Import Summary of a ‘Dry Run’ showing 11 ignored rows, in this case pertaining to the header row and the zero values in the import data set. |

### Step 3c – Correct conflicts and repeat ‘Dry Run’

If there are any other types of conflict (e.g. a ‘not found’ conflict message), then these should be investigated and corrected in the import file.

Repeat the ‘Dry Run’ process until you get a single ‘ignored’ conflict (the header row) and/or all conflicts are related to ‘zero’ values (if this is how the DHIS2 instance has been set to deal with ‘zero’ counts).

## Step 4 – Run the ‘Import Data’ function for real

Once all conflicts are resolved, and you are happy with the ‘Dry Run’ results, repeat Steps 1 and 2 ‘above’ but this time **set the ‘Dry Run’ message to ‘No’**.

All progress status updates and import summary screens that are displayed will be identical to those outlined in Step 3 above.

## Step 5 – Update the Analytics Tables

DHIS2 needs to ‘generate’ totals (aggregates) for the disaggregated data being imported. This is either run as a nightly process, or can be initiated manually. [[2]](#footnote-2)

To be able to ‘see and test’ the data that has just been imported (Step 4 above), it will be necessary to either wait for the scheduled update or to initiate it manually.

The following figure highlights the progress report of manually updating the analytics tables from the ‘Reports’ App within DHIS2.

|  |
| --- |
| Figure – Running the ‘Analytics tables update’ process to generate ‘totals’ or aggregate values based on disaggregated data loaded during import. |

Reminder: This Guide is intended for DHIS2 Administrators that are already familiar with importing and managing a DHIS2 instance. Therefore, some of the detail is not included, as it is assumed that the user of this Guide knows where to go to invoke some of these functions.

## Step 6 – Verifying the data

It is important to verify that the expected data has been imported correctly. To do that, we have outlined how to use the ‘Pivot Table’ App to generate and display some of the data that has just been imported.

The verification is a ‘spot test’ of the data, and it is assumed that if the values are correct for the selected parameters, then it will be correct for all of them.

### Step 6a – Select the Pivot Table App

|  |
| --- |
| Figure – Screenshot of selecting the ‘Pivot Table’ App, directly after completing Step 5 (UpdatingAnalytics Tables). |

### Step 6b – Select the three Pivot Table parameters (what, when, & where)

All DHIS2 Pivot Tables need three key parameters in order to generate a table of results – what (data element), when (the period), and where (the organisational level.

The following figures highlight capture this process (the figures are a little small, but the caption indicates the parameters selected).

|  |  |  |
| --- | --- | --- |
|  |  |  |
| Figure – What: ‘UNAIDS: Estimated number of people (15+) living with HIV’ | Figure –When: ‘Last 5 years’ | Figure –Where: ‘Animal Region and Food Region’ |

### Step 6b – Generate the ‘Pivot Table’

Once the ‘What’, ‘When’ and ‘Where’ parameters have been selected, click on the ‘Update’ menu item. This will result in a pivot table being generated.

|  |
| --- |
| Figure – Results table after clicking on the ‘Update’ menu link with the selected parameters… What: ‘UNAIDS: Estimated number of people (15+) living with HIV’, When: ‘Last 5 years’, and Where: ‘Animal Region and Food Region’ |

### Step 6c – Check the disaggregates (for gender and/or region)

As a final verification check, manipulated the layout of the pivot table to disaggregate by gender or organisational unit. The example below demonstrates disaggregation by gender.

Note: The displayed values should match the values you have in Spectrum for the given gender, period and sub-national region.

|  |
| --- |
| Figure – Viewing Disaggregate Data: 1.) Click on the ‘Layout’ menu item; 2.) Drag ‘Assigned categories’ from ‘Excluded dimensions’ to ‘Column dimensions’; 3.) Click on ‘Update’. |

Note: The resultant pivot table will include disaggregated columns for Male and Female for the selected data element IF the data element was one that included this disaggregation!

|  |
| --- |
| Figure – Result table after applying disaggregation. Use as a verification check on the imported data. |

Note: One could similarly ‘drag’ the ‘Organizational Units’ into the ‘Row Dimensions’ Area to further disaggregate by Regions. Such and example is also demonstrated below.

|  |
| --- |
| Figure – Viewing Disaggregate Data: 1.) Drag ‘Organisational units’’ from ‘Report Filter’’ to ‘’Row dimensions’; 2.) Click on ‘Update’. |

The resulting pivot table should now display disaggregation by both gender and subnational region.

|  |
| --- |
| Figure – Result table after applying further disaggregation. Use as a verification check on the imported data. |

1. See <http://dhis2.github.io/dhis2-docs/master/en/user/html/dhis2_user_manual_en_full.html> [↑](#footnote-ref-1)
2. See <http://dhis2.github.io/dhis2-docs/master/en/user/html/dhis2_user_manual_en_full.html> [↑](#footnote-ref-2)