1

Introduction

"Guide: to show or indicate the way"

The prerequisites

This guide is part of a support package for HIS Managers who want to import and utilise Spectrum derived HIV/ADIS estimates in DHIS2.

The guide assumes that:

1. The HIV-Spectrum Bootstrap Application, available in the DHIS2 Apps store, has been run; and

2. That the 'post-initiation' guideline has been followed.[1]

Without the above prerequisites, the instructions in this document are not applicable.
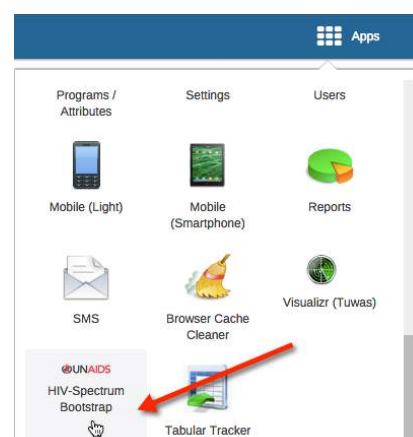
Figure 1: Screen capture of the DHIS2 HIV-Spectrum Bootstrap Application

Who this guide is for

The configuration and administration of DHIS2 is a specialist topic that involves significant prior knowledge. This import guide is not intended for use by those unfamiliar with DHIS2 nor is it intended for the end-users of the HIV/AIDS estimates.

It is intended for DHIS2 administrators who are already familiar with:

- The preparation and import of data into DHIS2;
- The configuration and use of DHIS2 Organizational Hierarchies, Data Elements and Indicators; and
- The identification and extraction of DHIS2 object UIDs.[2]

---

[1] The guideline "UNAIDS_DHIS2_Spectrum_Bootstrap_PostInitiation_Guide_V002_EN.pdf"

[2] Unique Identifiers. Each object in DHIS2 (e.g. an indicator) has a unique ID that can be used by the import mechanism to correctly link imported data with the correct object.

What this guide is intended to achieve

This guide is designed to enable an appropriately trained individual to prepare an extract report (file) 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 part of the associated 'support' package accompanying this guide.

Setting the context

UNAIDS wishes to make its Spectrum[3] derived HIV/AIDS estimates more easily available in DHIS2,[4] an open source analytics platform used widely by Ministries of Health in LIMCs.[5] Access to these estimates will enable Ministries of Health and other interested parties to visualise them (e.g. on maps, charts and graphs) and to combine them with local data already collected in DHIS2 to generate useful HIV/AIDS indicators.

While the creation of indicators is a uniquely localised process in DHIS2, the fact that Spectrum derived HIV/AIDS estimates are standardised provides an opportunity to partially automate the setup and import of these estimates into DHIS2.

To support the HIV/AIDS estimates setup process, a Bootstrapping App has been made available to the DHIS2 platform that:

- Instantiates a set of standard Spectrum HIV/AIDS data elements
- Instantiates a set of placeholder Indicators that use the Spectrum HIV/AIDS data elements (which need further localisation effort before they can be used)
- Provides access to relevant documentation (including this guideline) from within the DHIS2 platform.

It should be noted that the UIDs of the data elements and indicators instantiated by the Bootstrapping App will be common to all DHIS2 instances where it has been run. This provides opportunities in the future to leverage these indicators due to them being standardised and identifiable objects within any DHIS2 instance. This guide takes advantage of these 'known' UIDs as part of the preparation process necessary for importing data into DHIS2.

---

[3] See http://www.unaids.org/en/dataanalysis/datatools/spectrumepp2013

[4] See https://www.dhis2.org/

[5] Low and Middle Income Countries

The data

The Spectrum HIV/AIDS Estimates consist of a number of 'data elements' that are disaggregated by standardised age-brackets, gender, and sub-national geographic breakdowns. To support the widest possible import potential, the 'data elements' instantiated during the Bootstrapping (where age brackets are applicable) are handled as separate objects rather than as a single equivalent data element.

Examples of the DHIS2 instantiated data elements include:

- UNAIDS: Estimated number of HIV+ pregnant women
- UNAIDS: Estimated number of people (15+) eligible for ART according to national guidelines
- UNAIDS: Estimated number of people (15-49) eligible for ART according to national guidelines
- UNAIDS: Estimated number of people (<15) eligible for ART according to national guidelines
- UNAIDS: Estimated number of people (15+) living with HIV
- UNAIDS: Estimated number of people (15-49) living with HIV
- UNAIDS: Estimated number of people (<15) living with HIV

The indicators that were instantiated during the Bootstrapping process utilise these data elements as part of their definition along with local data elements specific to a country DHIS2 instance. Examples include:

- UNAIDS: Percentage of all people (15+) living with HIV who are receiving ART
- UNAIDS: Percentage of all people (15-49) living with HIV who are receiving ART
- UNAIDS: Percentage of all people (<15) living with HIV who are receiving ART
- UNAIDS: Percentage of HIV+ pregnant women who receive antiretrovirals to reduce the risk of MTCT
- UNAIDS: Percentage of HIV-exposed infants who received an HIV test
- UNAIDS: Percentage of people (15+) eligible for HIV treatment according to national criteria that are receiving ART
- UNAIDS: Percentage of people (15+) living with HIV who are enrolled in HIV care
- UNAIDS: Percentage of people (15-49) eligible for HIV treatment according to national criteria that are receiving ART
- UNAIDS: Percentage of people (15-49) living with HIV who are enrolled in HIV care
- UNAIDS: Percentage of people (<15) eligible for HIV treatment according to national criteria that are receiving ART
- UNAIDS: Percentage of people (<15) living with HIV who are enrolled in HIV care

Using the guide

The steps outlined in the next section 'follow' a real-life data preparation session based on a Spectrum HIV/AIDS estimates extract. The data used in the guide is real data, but has been anonymized by replacing the country label in the extract file with the fictional label of the training country used by DHIS2 trainers.

The DHIS2 training country, called 'TrainingLand' is a fully articulated country with sub-national regions, districts and facilities and a serious of standardized WHO recommended routine indicators.

All screenshots and data in this and the accompanying documents use TrainingLand as the nominal country for demonstration purposes.

This Guide is equally applicable to:

1. List extracts from versions of ~~the~~ Spectrum ~~application~~ up to and including September 2015[6]; and

2. DHIS2 compatible Spectrum extracts that implement the newer DHIS2 export functionality

For extracts of type 1, the reader should start at Step 1 of the Guide, while for extracts of type 2, the reader can start at Step 5d (since this new extract function sets up common UIDs and correctly specifies DHIS2 compatible column names and column orders).

---

[6] Build Number xx, Version Number yy

2

The Guide

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 2 – Initial opening of a Spectrum HIV/AIDS Estimate extract file (list format)

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 3 – Create a copy of the original sheet

Step 3 – Eliminate Totals

The standard Spectrum extract format ~~format~~ includes Population Totals, as well as 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.

~~Step 3a – Delete all population rows~~

~~Population data placeholders are not initiated in the initial release of the bootstrapping application, therefore these population estimates can be eliminated.~~

~~Delete all Total Population rows (subnational and national).~~

Figure 4 – Search for and delete all rows containing Total Population data

Make sure to identify and delete the disaggregated Totals as well!
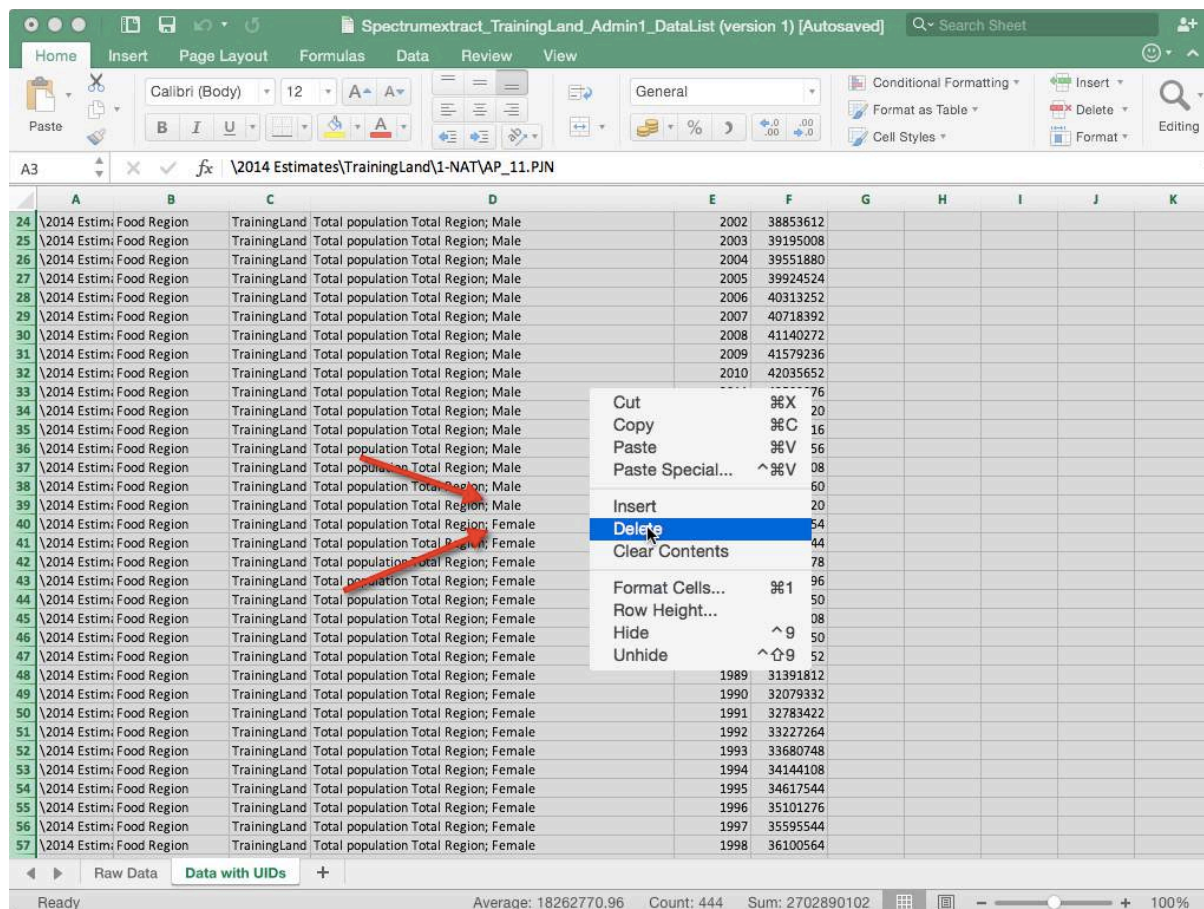
Figure 5 – Delete all disaggregated 'Total' population data as well, e.g. Male, Female totals.

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 6 – 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.

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.
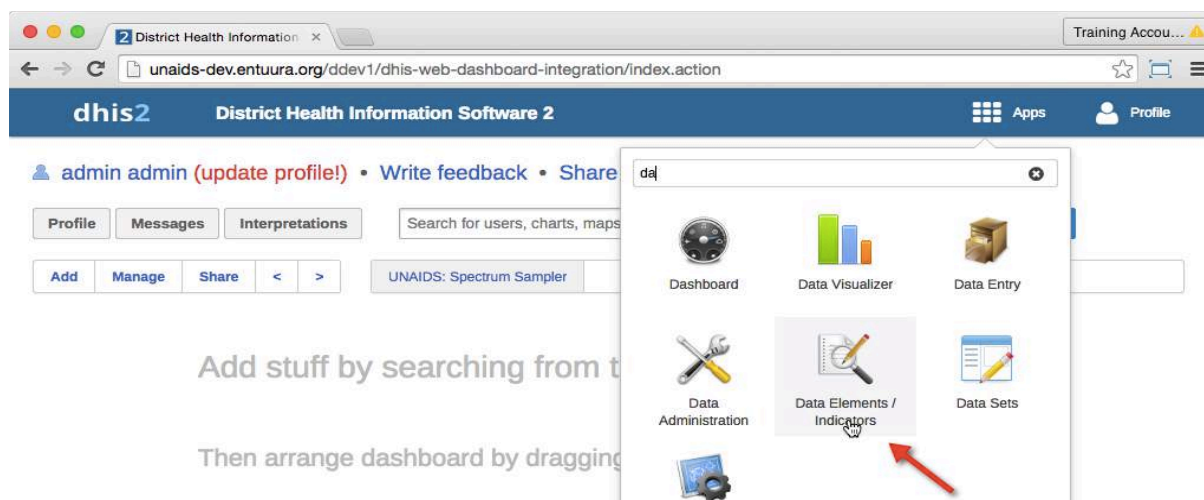
Figure 7 - Select the Data-Element/Indicators DHIS2 App
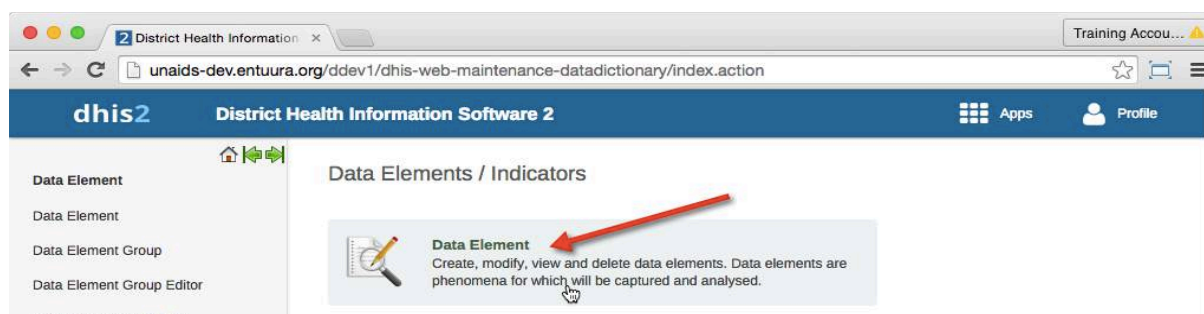
Figure 8 - Select the Data Element option
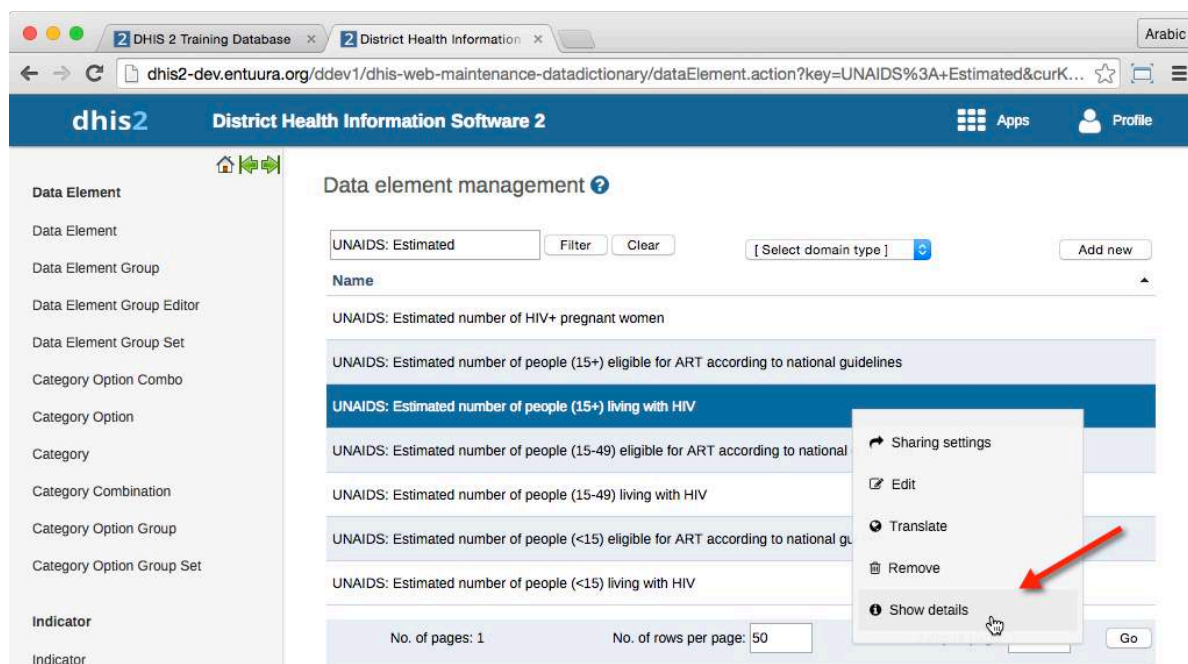
Figure 9 - Find in the list of data elements those beginning with 'UNAIDS: ...' and click on the appropriate data-element. From the displayed dropdown menu, select 'Show details'
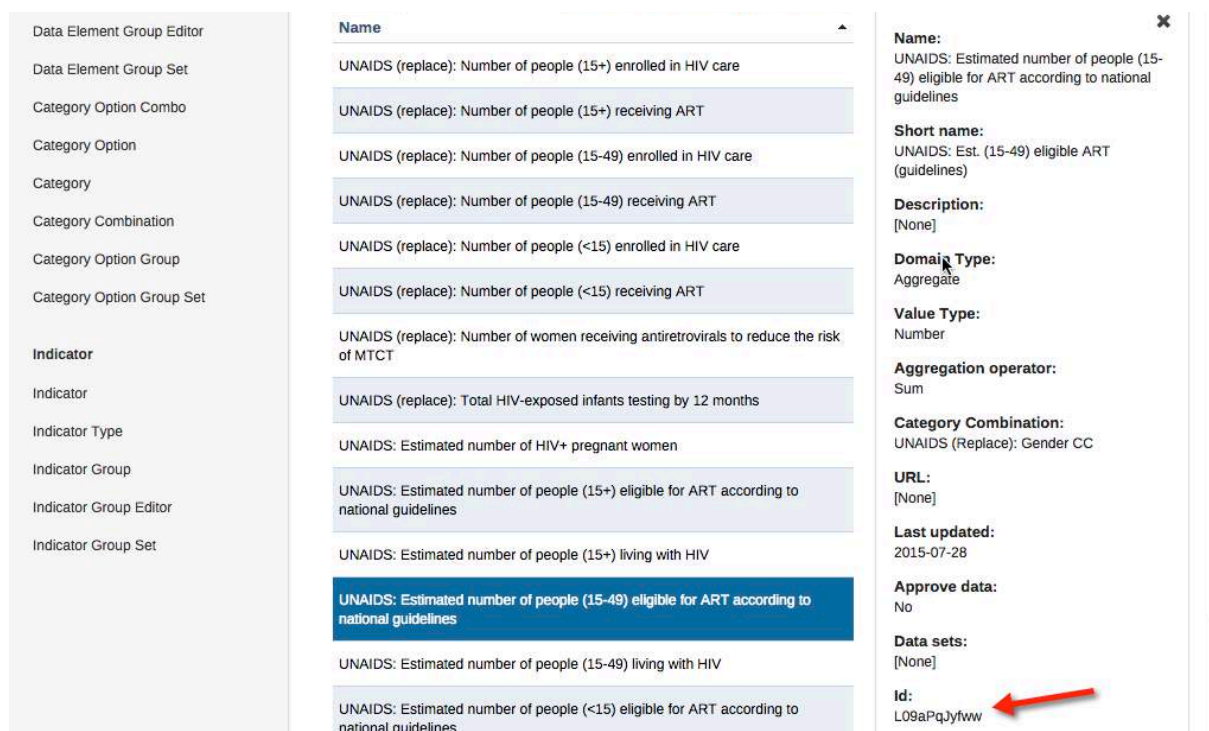
Figure 10 - Note down the Id for the specific data element. It will be used in the import spreadsheet.

Note: The known Spectrum Data Element UIDs are available in the appendix. It will be necessary to still use the above method to identify local country instance gender and subnational region UIDs.

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.

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 11 - 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).
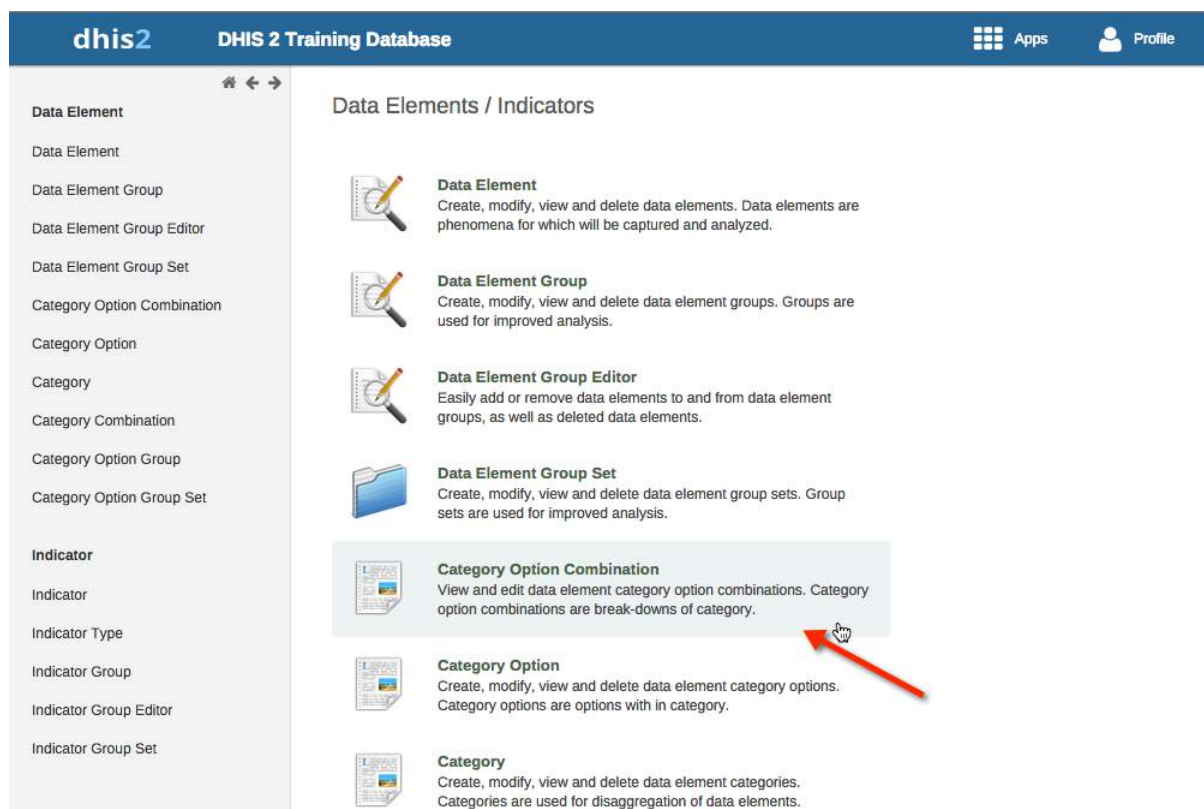
Figure 12 – Select the Category Option Combination menu item under 'Data Elements' in DHIS2
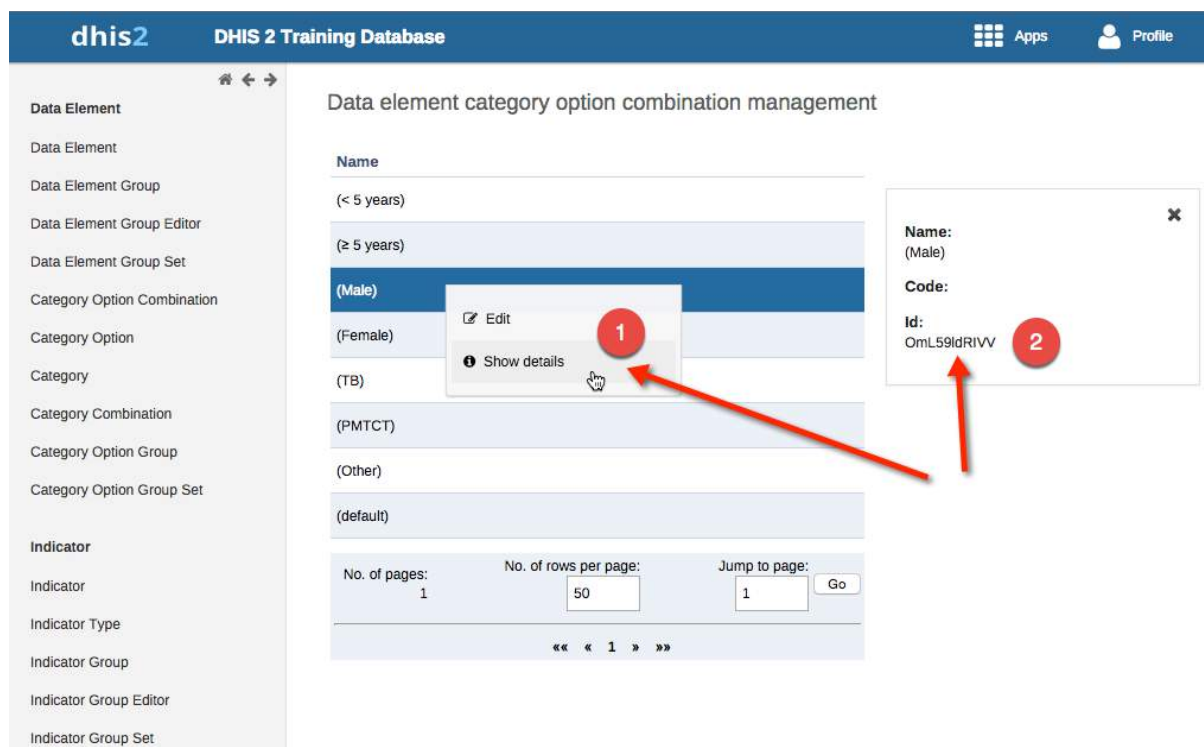
Figure 13 – 1. 'Show Details' of the Male/Female Category Option Combination; and 2. Copy the IDs

Figure 14 - 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 15 – Insert an 'orgunit' column between 'Subnational Region' and 'Country'

Step 6b – Select the 'Organisational Units' App and Menu Item.

Figure 16 – Select the 'Organisation Units' app from the Apps menu in DHIS2

Figure 17 – 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 the '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).
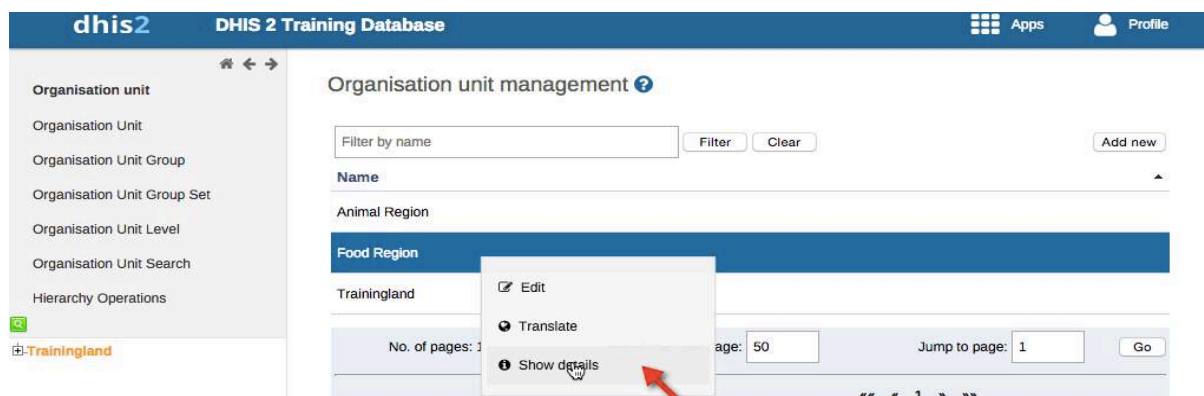
Figure 18 – Highlight the 'country' item at the left (turns orange), then 'Show details' of the subnational region of interest in the main area of the screen.
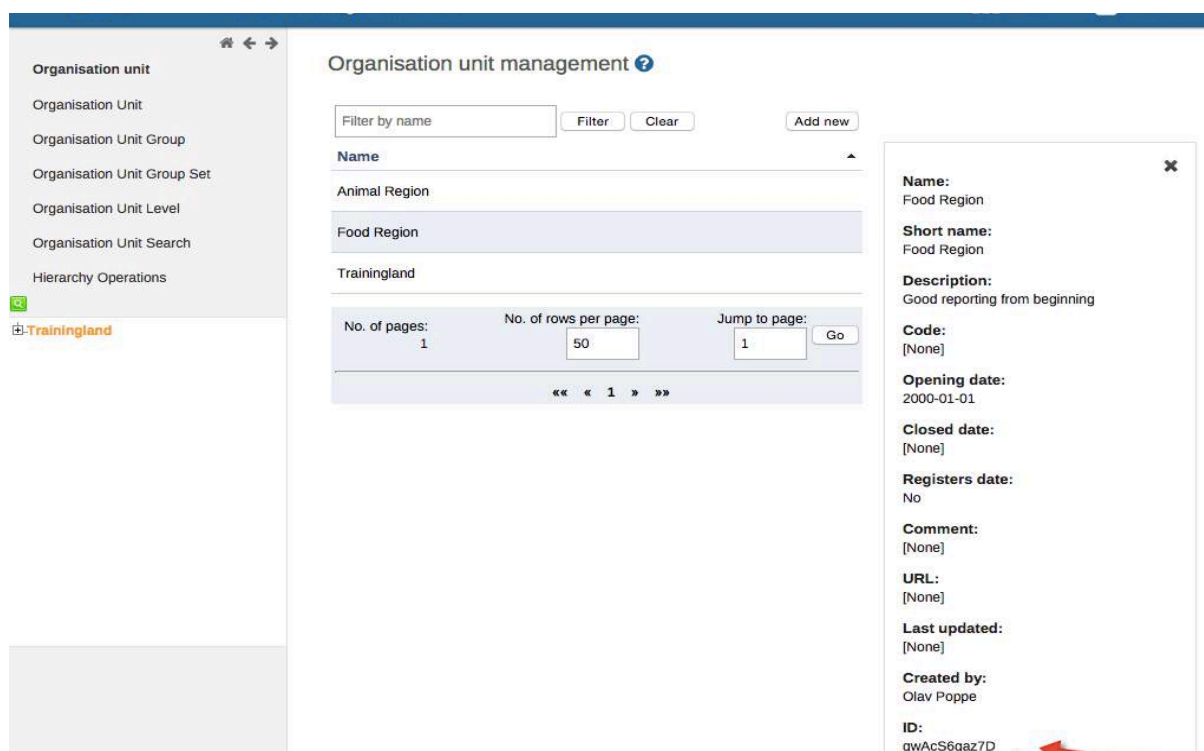
Figure 19 – Copy the desired ID

Figure 20 – 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 on 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 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 21 – Import ready version of the Spectrum Extract file

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

3
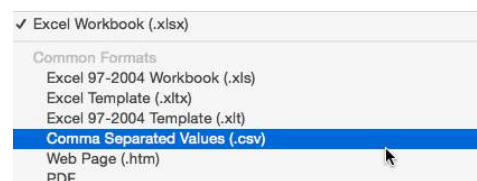
Importing and Testing

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
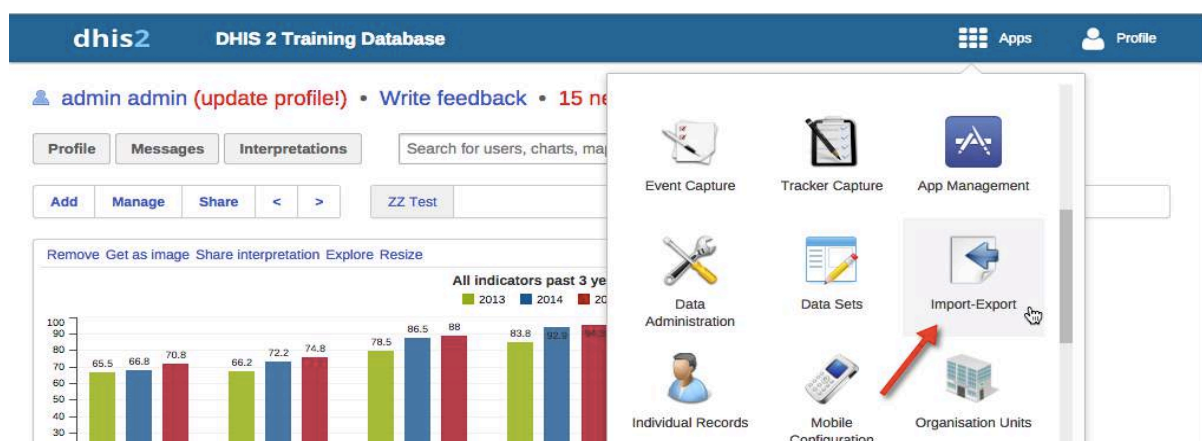
Figure 22 – Use the DHIS2 'Apps' menu item to locate the 'Import-Export' App

Figure 23 – 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.
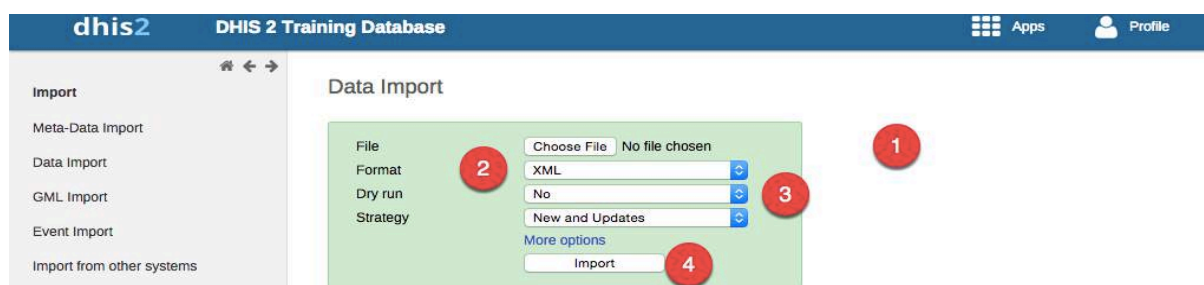
Figure 24 – Starting Point for Data Import Screen

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!
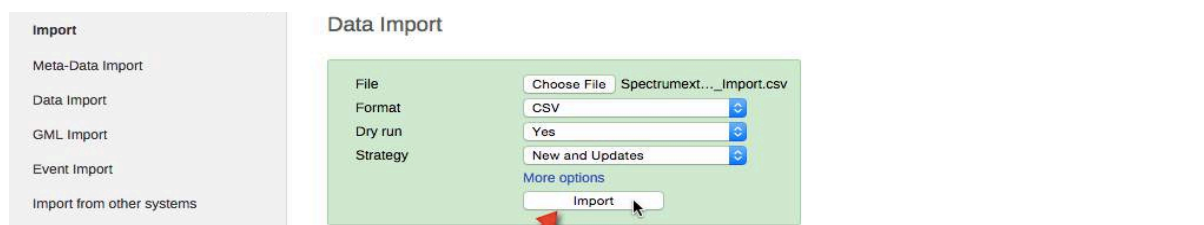
Figure 25 – Example screenshot of the parameters filled in and ready for a 'dry run'!

Step 2b – Run the Import

4. Run the import (leave the Strategy as 'New and Updates').

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

21

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.
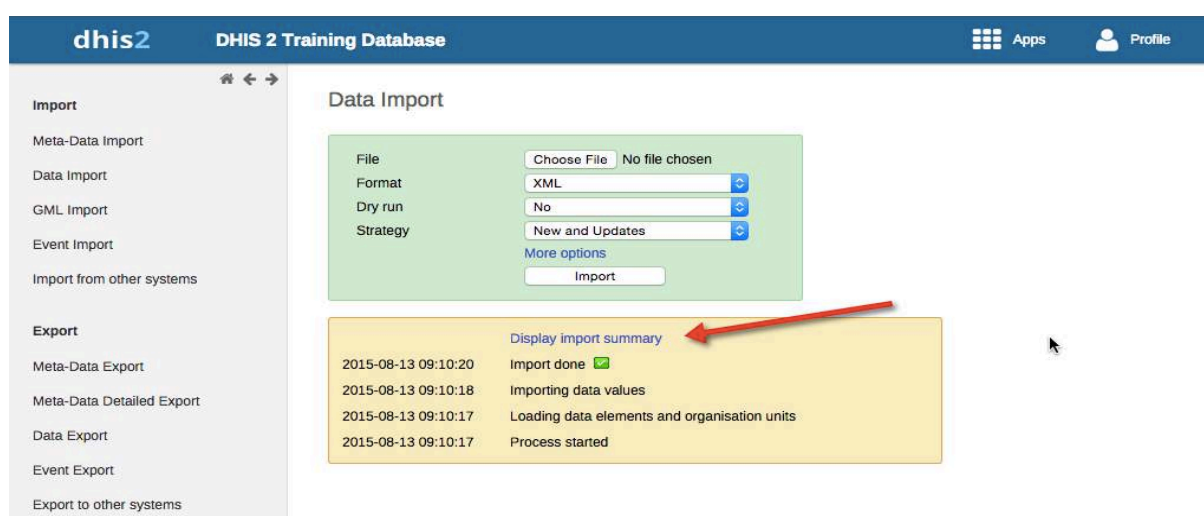
Figure 26 – The progress log with completed process steps after a 'Dry Run' has been initiated. Note the 'Display Import Summary' link.

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…*').
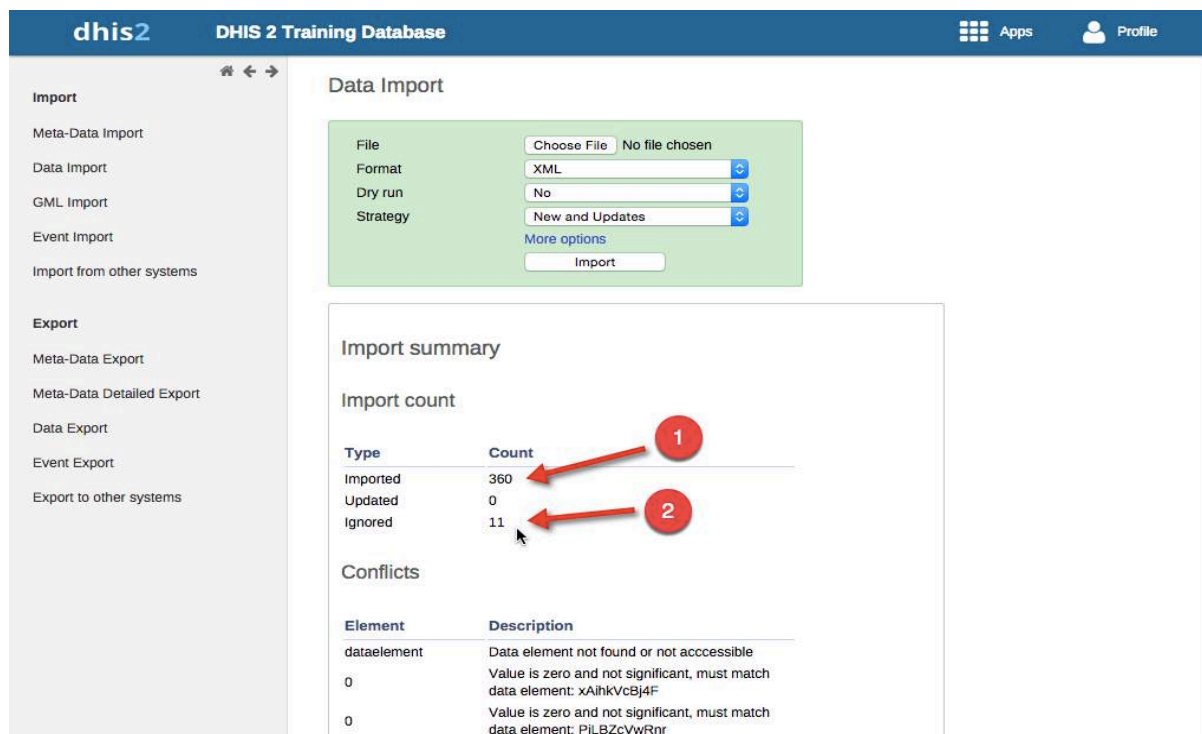
Figure 27 – The Import Summary of a 'Dry Run' showing 11 ignored rows, in this case pertaining to the header row and the zero values in the import data set.

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 processed, or can be initiated manually.

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.
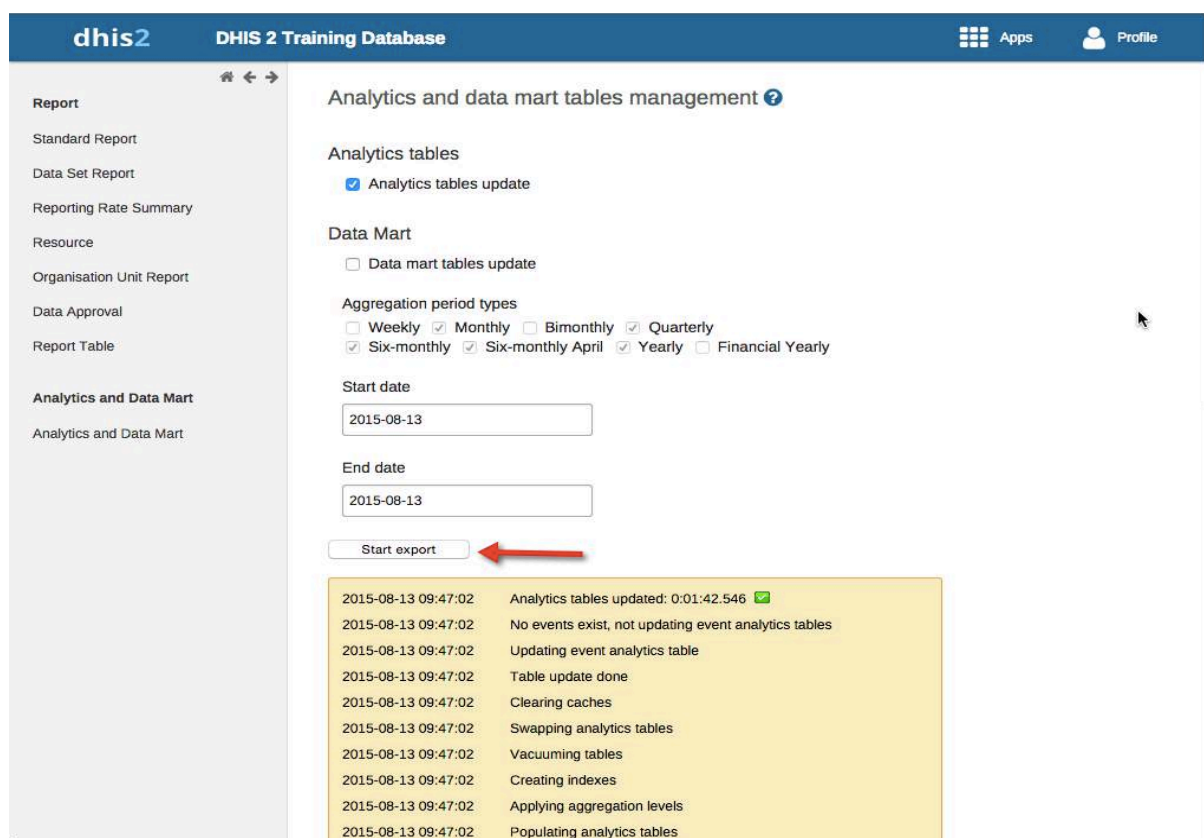
Figure 28 – Running the 'Analytics tables update' process to generate 'totals' or aggregate values based on disaggregated data loaded during import.

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
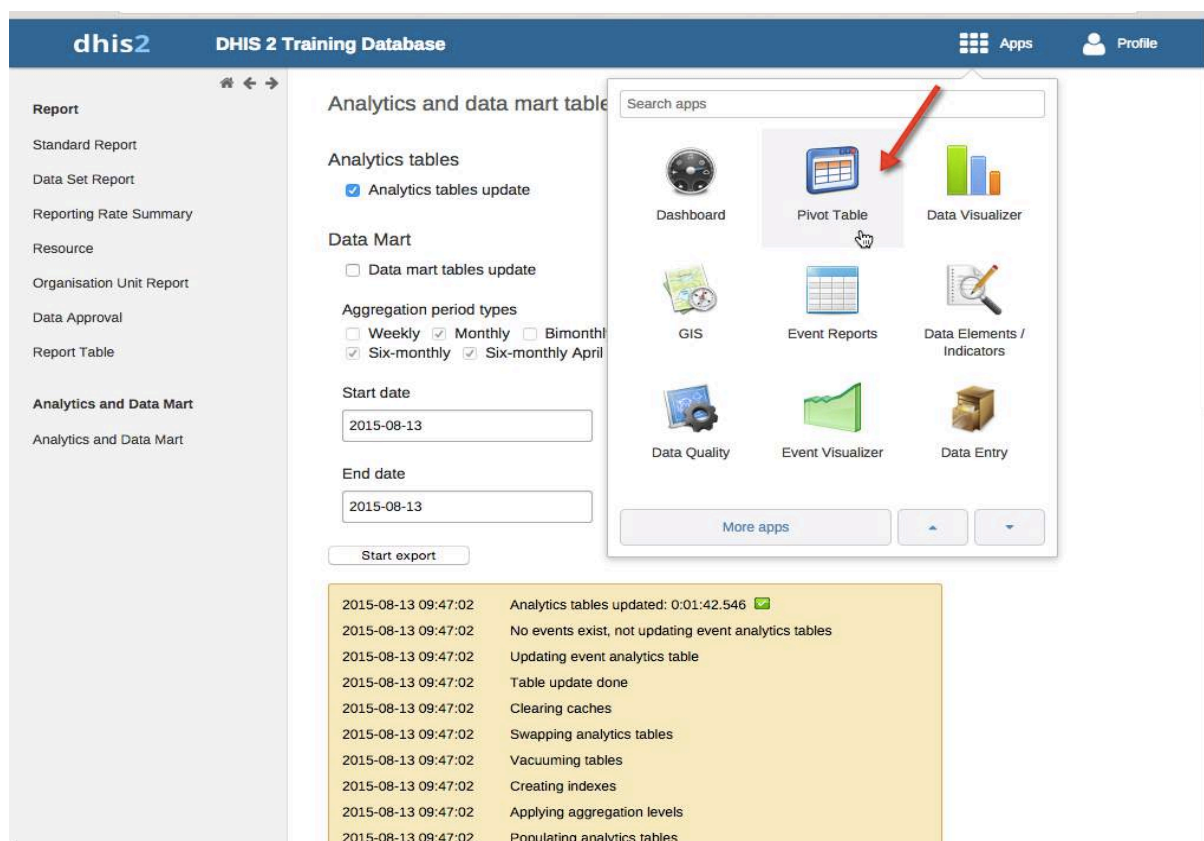
Figure 29 – Screenshot of selecting the 'Pivot Table' App, directly after completing Step 5 (UpdatingAnalytics Tables).

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).
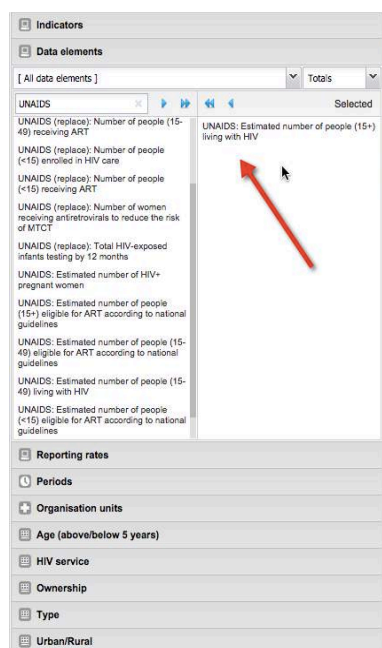
Figure 30 – What: 'UNAIDS: Estimated number of people (15+) living with HIV'
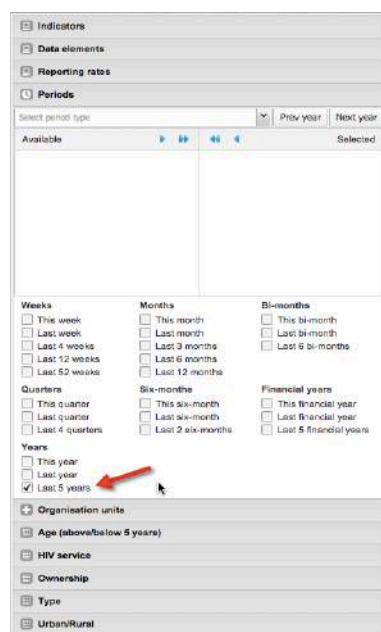
Figure 31 –When: 'Last 5 years'
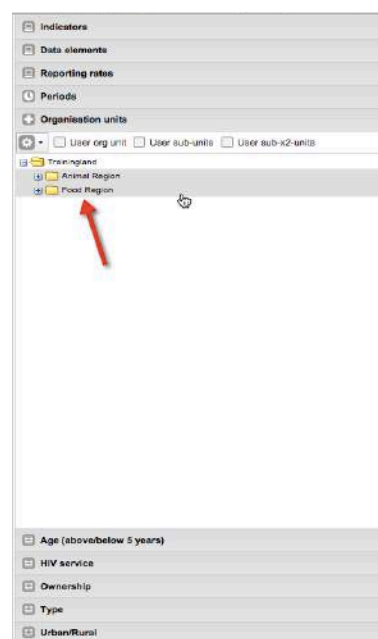
Figure 32 –Where: 'Animal Region and Food Region'

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.

Note: The result will be the 'aggregate' of the parameters selected by default, and can therefore ACT AS A VERIFICATION of the aggregation (i.e. the totals displayed should match the totals in Spectrum)

Figure 33 – 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.
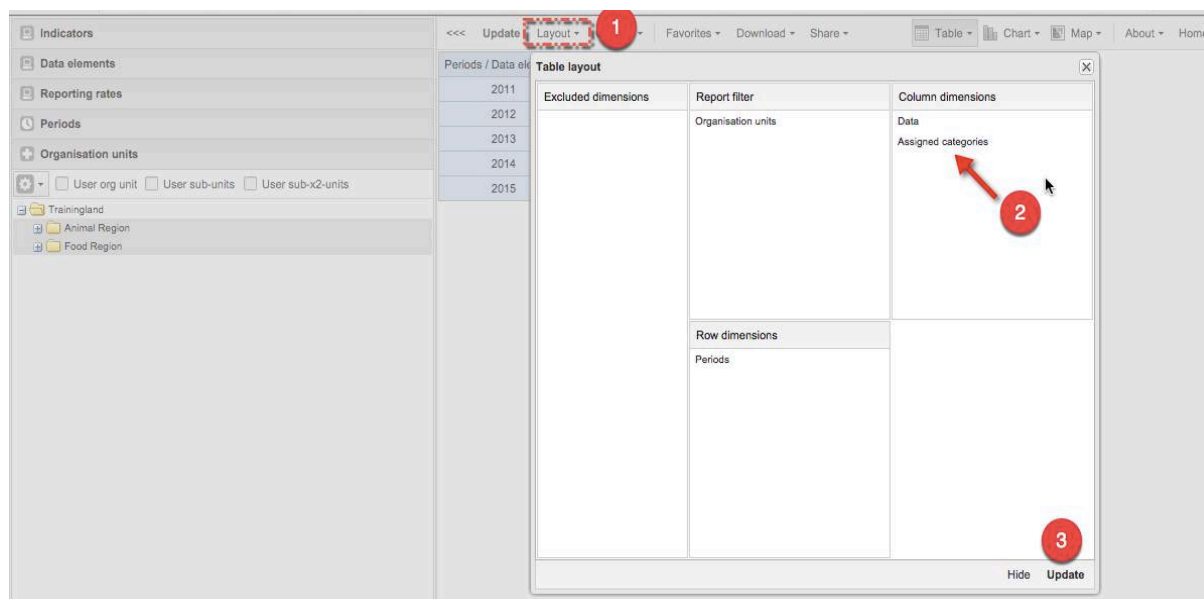

Figure 34 – Viewing Disaggregate Data: 1.) Click on the 'Layout' menu item; 2.) Drag 'Assigned categories' from 'Excluded dimensions' to 'Column dimensions'; 3.) Click on 'Update'.

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 35 – 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.
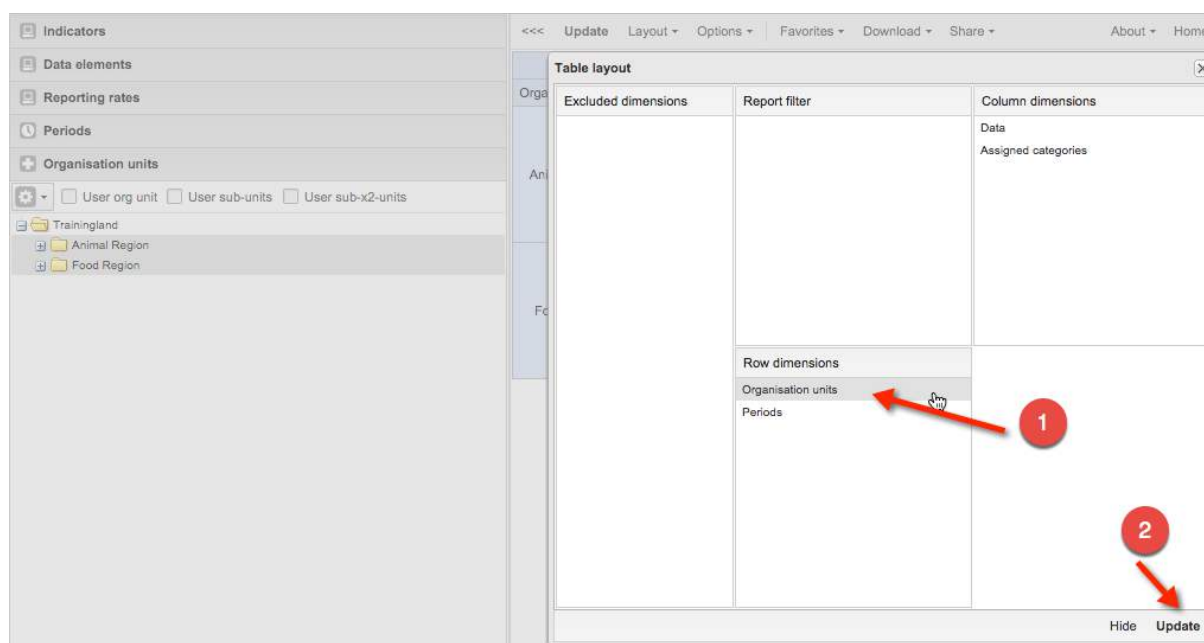
Figure 36 – Viewing Disaggregate Data: 1.) Drag 'Organisational units'' from 'Report Filter'' to ''Row dimensions'; 2.) Click on 'Update'.

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

Figure 37 – Result table after applying further disaggregation. Use as a verification check on the imported data.

A

Appendices

Appendix A – List of Data Elements used in DHIS2 for handling Spectrum estimates.

- UNAIDS (replace): Number of people (15+) enrolled in HIV care
- UNAIDS (replace): Number of people (15+) receiving ART
- UNAIDS (replace): Number of people (15-49) enrolled in HIV care
- UNAIDS (replace): Number of people (15-49) receiving ART
- UNAIDS (replace): Number of people (<15) enrolled in HIV care
- UNAIDS (replace): Number of people (<15) receiving ART
- UNAIDS (replace): Number of women receiving antiretrovirals to reduce the risk of MTCT
- UNAIDS (replace): Total HIV-exposed infants testing by 12 months
- UNAIDS: Estimated number of HIV+ pregnant women
- UNAIDS: Estimated number of people (15+) eligible for ART according to national guidelines
- UNAIDS: Estimated number of people (15+) living with HIV
- UNAIDS: Estimated number of people (15-49) eligible for ART according to national guidelines
- UNAIDS: Estimated number of people (15-49) living with HIV
- UNAIDS: Estimated number of people (<15) eligible for ART according to national guidelines
- UNAIDS: Estimated number of people (<15) living with HIV