Deans Dashboard Implementation Guide

2.17

Revision 1383
Version 2.17 2015-03-27 08:24:47

About this guide

The Dean's Dashboard Implementation Guide was developed by the USAID Capacty*Plus*.

Dean's Dashboard

The Dean's Dashboard is an adpation of the use of the DHIS2 software in the Medical Education Sector. The Dean's Dashboard is distributed under an *appropriate* Creative Common's License.

1. DHIS2 Documentation

The DHIS2 documentation is a collective effort and has been developed by the development team and users. While the guide strives to be complete, there may be certain functionalities which have been omitted or which have yet to be documented. This section explains some of the conventions which are used throughout the document.

DHIS2 is a browser-based application. In many cases, screenshots have been included for enhanced clarity. Shortcuts to various functionalities are displayed such as "Maintenance->Data administration". The "->" character indicates that you should choose "Maintenance" and then click on "Data administration" in the menu which appears through the browser.

Part Volume I: What is the Dean's Dashboard?. What is the Dean's Dashboard ?

Chapter 1. What is the Dean’s Dashboard ?

1.1. Overview

(Contextualization/Overview of using DHIS2 as Dean’s Dashboard - adapted from materials provided by CapPlus)

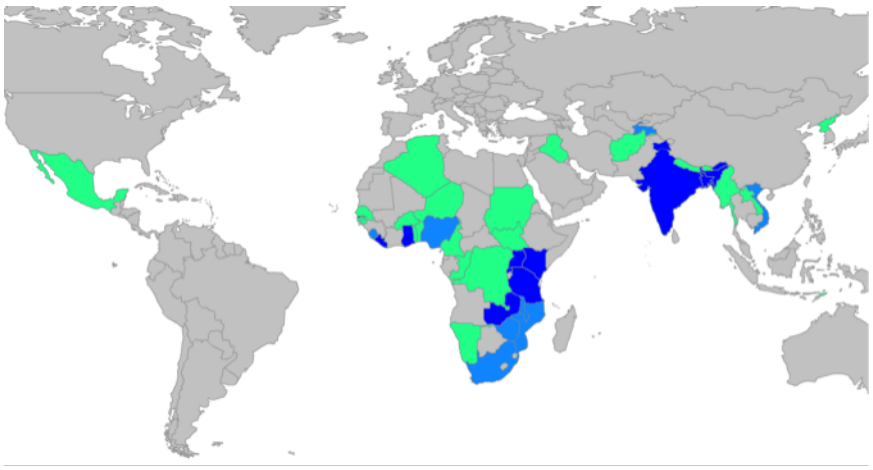
The Dean’s Dashboard is most effective as part of a strategic planning process. Before a school can implement a dashboard, it needs access to high-quality, routinely collected, individual-level data. While data are usable in any easily accessed format, having data in a digital system is ideal, even in electronic spreadsheets. Pilot institutions found that setting up the dashboard’s data and reports was relatively easy once the DHIS 2 interface was mastered. After a representative of an institution has been trained to configure and utilize DHIS 2, few resources are required to continue and expand its usefulness.

Health workforce educational institutions around the globe are struggling to meet the increasing demand for health workers capable of providing quality health services to expanding populations. These institutions must find ways to produce more health workers within current, marginally expanding, or even decreasing budgets. More effective and efficient management of education and training institutions can play a major role in a country’s ability to scale up its production of competent and qualified health workers.

The Dean’s Dashboard is a computerized tool to guide a school’s management team in defining management goals and tracking progress toward goal achievement. The Dean’s Dashboard is meant to measure progress and chart trends, indicating either improvement over time or need for strategic action.

1.1.1. Getting Support

Getting connected to DHIS2 community (for deployment support, academies, etc).



Chapter 2. School Performance

2.1. Performance Areas

The primary use case of the Dean’s Dashboard is tool to monitor key school management performance areas by school leaders. This should be achieved by supporting these main requirements:

- The design of indicators to track school performance areas
- Visualization and trend analysis of the designed indicators
- Data entry for the periodic numeric data used to calculate chosen indicators

Secondary requirements include:

- Limiting access to edit and view data based on user’s role and organizational affiliation within the school
- Ability to monitor data entry statistics
- Ability to import data from secondary systems or other devices (e.g. mobile phones)
- Ability to embed visuals into external applications or web pages

Chapter 3. Organizational Readiness

3.1. Resources and Infrastructure

In terms of infrastructure needed, two options exist—in the cloud or local to the institution. Small institutions may prefer to pay for a cloud-based implementation (that is, to access and use the software entirely online) as purchasing and maintaining servers may be prohibitive. Institutions with strong ICT units may prefer to install the software locally in their data centers. Configuring and implementing the system is not overly complex, but support will be essential to new users.

3.2. Organizational Readiness

Is Your Organization Ready for the Dean’s Dashboard?

Answering these questions can help your institution understand its level of readiness to initiate the Dean’s Dashboard program. You may find it helpful to have a colleague review your responses or to answer the questions with a larger group (e.g., senior leaders). Has your institution clearly defined the need to use the Dean’s Dashboard?

1. Have you clearly defined the need that is driving your institution to consider implementing the Dean’s Dashboard?

   > An organization is more likely to be ready to undertake a Dean’s Dashboard initiative when it has objective information to support the need for improving specific areas.

2. Who in your organization uses information to guide decision making and process improvement?

   > Dean’s Dashboard is an information system that offers informational reports on a number of indicators important to your institution. Understanding than one individual or a team of persons using that information can affect how successful it can be for addressing your situation.

3. Is now the right time for implementing an information tool such as the Dean’s Dashboard (i.e., it will not compete with other major changes currently being made at your institution)?

   > If your organization is currently experiencing many changes, it may not be the ideal moment to begin implementing the Dean’s Dashboard initiative. Attempting to manage multiple change efforts at one time may degrade your institution's ability and employees' willingness to implement and sustain the Dean’s Dashboard effort. The program may be viewed as a distraction rather than a solution.

4. Will your institution's leaders support the effort required to implement and sustain the Dean’s Dashboard initiative?

   > It is essential that the leaders of your institution actively support and champion Dean’s Dashboard. Leaders needs to understand the requirements of the program (e.g., identification of indicators in an iterative process spanning several weeks, technical preparation and training of system users, regular meetings to reinforce concepts and skills) and be willing to provide the personnel, time, and resources required to successfully implement and sustain the program.

5. Will your institution provide sufficient staff with the necessary skills and time to support the development and implementation and use of the Dean’s Dashboard?

   > It is important to find individuals with the right characteristics to serve as users to increase the success of the Dean’s Dashboard initiative.

6. Will your institution allow time for personnel to attend training?

   > While the Dean’s Dashboard will be designed by participating institutions, some direct staff training will be required. The full training element of the program includes up to XX hours of classroom learning; however, your institution may decide to train only one segment (e.g., a specific indicator or set of related indicators) that may require 1 hour. To preclude disruptions and

scheduling problems and to maximize learning, participants should be excused from all duties while attending class.

3.3. Best practices in planning deployment/project management

3.3.1. Stakeholder Engagement

A dashboard development committee can be composed of the following stakeholders:

1. Director of the School of Medicine (Chair)
2. Faculty Administrator
3. Coordinator of the Undergraduate training programs
4. Coordinator of the Registrar (Residency) training program
5. Faculty examination officer
6. Representative of the Nurse training program
7. Representative of the Allied Health training programs
8. Representative of the Academic Planning unit
9. Representative of the Institutional Finance Department
10. Representative of the Human Resource Department.

3.3.1.1. Key Questions

The Stakeholder Leadership Group should consider the following questions:

1. What indicators that needed to be tracked on regular basis – should they just be educational (student performance) indicators or should issues related to program administration be included.
2. At what level should the dashboard operate – should it be at departmental level, school level, or at Faculty level?
3. Who would run the project – sourcing the data and handling the data
4. Was there sufficient expertise in IT within the Faculty to develop the Dashboard?
5. Did the Faculty have the infrastructure to house the dashboard – the computer server and the necessary software?
6. How much would it cost the Faculty per annum to run this project?

3.3.2. project management

3.3.3. defining scope

The dashboard development committee decides on which indicators for each of the identified key performance areas would be tracked, and how often these indicators would be monitored.

3.4. Getting data:

3.4.1. What is on the dashboard?

3.4.2. Where does the data come from?

3.4.3. who has the data

3.4.4. what parts of the institution have the data

Chapter 4. Key Indicators

4.1. Overview

The educational indicators to be tracked should be at a macro-level – dealing with program performance rather than of individual student/staff performance. The following key performance areas were identified to tracked on the Dean’s dashboard:

1. Student selection, admission and numbers
2. Student progression and achievement
3. Student development support and guidance
4. Teaching and learning
5. Research
6. Human Resources
7. Budgeting and fiscal management
8. Alumni (graduate tracking)

4.2. Example Key Indicators

4.2.1. Student selection and admissions

1. Number of applicants by gender, race and province
2. Number of students admitted by gender, province, economic class, and rurality
3. Number of students self-fee paying, on scholarship, on loans, and unable to pay fees.

4.2.2. Student progression and achievements

1. Numbers of students graduating from the programs each year
2. Student throughput rates
3. Student graduation rates
4. Student attrition rates

4.2.3. Human Resources

1. Established posts and vacancy rates for academics and support staff
2. Academic staff by rank, highest academic qualification, age, gender race, and nationality

4.2.4. Teaching and Learning

1. Established posts and vacancy rates for academics and support staff
2. Academic staff by rank, highest academic qualification, age, gender race, and nationality

4.2.5. Teaching and Learning

1. Workload in in FTEs (Full Time Equivalents) and in hours per staff member per program
2. Utilization levels of lecture theatres, class laboratories, tutorial rooms, and community hospitals

4.2.6. Research

1. Number of research applications submitted to IRB and numbers approved and rejected
2. Number of grants application and numbers of successful and unsuccessful applications
3. Number of papers published in peer reviewed journals – proportion where faculty was the lead author
4. Numbers of Masters and Doctorate students graduating

4.2.7. Finance

Chapter 5. Step-by-Step Tutorial

5.1. Step-by-step walk-through (from Jehu)

for one use case/indicator with screenshots

Part Volume II: End User Guide. Volume II: End User Guide

Chapter 6. Introduction to the End Users' Guide

6.1. Welcome!

This guide is for you if you are going to be a primary user of the Dean's Dashboard. Your responsibilities include one or more of the following:

- Entering data into the Dean's Dashboard
- Ensuring data is updated routinely
- Designing reports and Dashboard
- Performing analysis on the data
- Making decisions with the data

Chapter 7. What is DHIS 2?

After reading this chapter you will be able to understand:

- What is DHIS 2 and what purpose it serves with respect to health information systems (HIS)?
- What are the major technological considerations when it comes to deploying DHIS 2, and what are the options are for extending DHIS 2 with new modules?
- What is the difference between patient based and aggregate data?
- What are some of the benefits and challenges with using Free and Open Source Software (FOSS) for HIS?

7.1. DHIS 2 Background

DHIS 2 is a tool for collection, validation, analysis, and presentation of aggregate statistical data, tailored (but not limited) to integrated health information management activities. It is a generic tool rather than a pre-configured database application, with an open meta-data model and a flexible user interface that allows the user to design the contents of a specific information system without the need for programming. DHIS2 and upwards is a modular web-based software package built with free and open source Java frameworks.

DHIS2 is open source software released under the BSD license and can be used at no cost. It runs on any platform with a Java Runtime Environment (JRE 6 or higher) installed.

DHIS 2 is developed by the Health Information Systems Programme (HISP) as an open and globally distributed process with developers currently in India, Vietnam, Tanzania, Ireland, and Norway. The development is coordinated by the University of Oslo with core support from NORAD.

As of October 2012, the DHIS 2 software is used in more than 30 countries in Africa, Asia, and Latin America, and countries that have adopted DHIS 2 as their nation-wide HIS software include Kenya, Tanzania, Uganda, Rwanda, Ghana, Liberia, and Bangladesh. A rapidly increasing number of countries and organisations are starting up new deployments.

The documentation provided herewith, will attempt to provide a comprehensive overview of the application. Given the abstract nature of the application, this manual will not serve as a complete step-by-step guide of how to use the application in each and every circumstance, but rather will seek to provide illustrations and examples of how DHIS2 can be implemented in a variety of situations through generalized examples.

Before implementing DHIS 2 in a new setting, we highly recommend reading the DHIS 2 Implementation Guide (a separate manual from this one), also available at the main DHIS2 website.

7.2. Key features and purpose of DHIS 2

The key features and purpose of DHIS 2 can be summarised as follows:

- Provide a comprehensive HIS solution based on data warehousing principles and a modular structure which can easily be customised to the different needs of the health systems - and supports the idea of an integrated HIS at all levels of the health hierarchy.
- Customisation and local adaptation through the user interface. No programming required to start using DHIS 2 in a new setting (country, region, district etc.).
- Provide data entry tools which can either be in the form of standard lists or tables, or can be customised to replicate paper forms.
- Provide different kinds of tools for data validation and improvement of data quality.
- Provide easy to use - one-click reports with charts and tables for selected indicators or summary reports using the design of the data collection tools. Integration with popular external report design tools like iReport and BIRT allows super-users to flexibly add more custom reports accessible to all users.
- Flexible and dynamic (on-the-fly) data analysis in the Data Visualizer and the GIS modules.

- A user-specific dashboard for quick access to the relevant monitoring and evaluation tools including indicator charts and links to favourite reports, maps and other key resources in the system.
- Easy to use user-interfaces for metadata management e.g. for adding/editing datasets or health facilities. No programming needed to set up the system in a new setting.
- Functionality to design and modify calculated indicator formulas.
- User management module for passwords, security, and fine-grained access control (user roles).
- Messages can be sent to system users for feedback and notifications. Messages can also be delivered to email and SMS.
- Users can share and discuss their data in charts and reports using Interpretations, enabling an active information-driven user community.
- Functionalities of export-import of data and metadata, supporting synchronisation of offline installations as well as interoperability with other applications.
- Integration with other software systems – using the DHIS 2 Web-API and the Integration Engine.
- Further modules can be developed and integrated as per user needs, either as part of the DHIS 2 portal user interface or a more loosely-coupled external application interacting through the DHIS 2 Web-API.

In summary, DHIS2 provides a comprehensive HIS solution for the reporting and analysis needs of health information users at any level.

7.3. Use of DHIS 2 in HIS: data collection, processing, interpretation, and analysis.

The wider context of HIS can be comprehensively described through the information cycle presented in Figure 1.1 below. The information cycle pictorially depicts the different components, stages and processes through which the data is collected, checked for quality, processed, analysed and used.

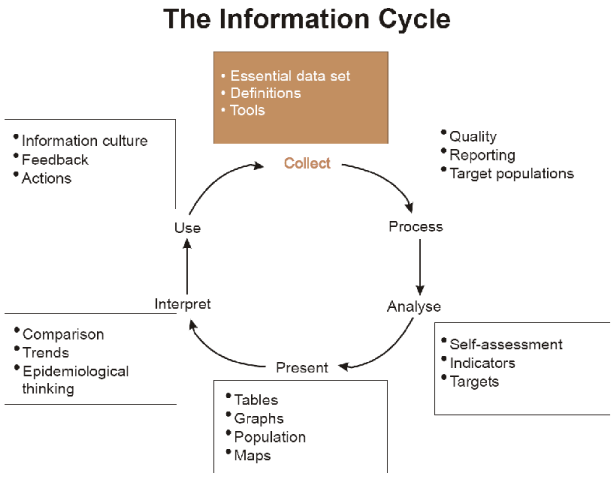


Figure 7.1. The health information cycle

DHIS 2 supports the different facets of the information cycle including:

- Collecting data.
- Running quality checks.
- Data access at multiple levels.
- Reporting.
- Making graphs and maps and other forms of analysis.
- Enabling comparison across time (for example, previous months) and space (for example, across facilities and districts).

- See trends (displaying data in time series to see their min and max levels).

As a first step, DHIS 2 serves as a data collection, recording and compilation tool, and all data (be it in numbers or text form) can be entered into it. Data entry can be done in lists of data elements or in customised user defined forms which can be developed to mimic paper based forms in order to ease the process of data entry.

As a next step, DHIS 2 can be used to increase data quality. First, at the point of data entry, a check can be made to see if data falls within acceptable range levels of minimum and maximum values for any particular data element. Such checking, for example, can help to identify typing errors at the time of data entry. Further, user can define various validation rules, and DHIS 2 can run the data through the validation rules to identify violations. These types of checks help to ensure that data entered into the system is of good quality from the start, and can be improved by the people who are most familiar with it.

When data has been entered and verified, DHIS 2 can help to make different kinds of reports. The first kind are the routine reports that can be predefined, so that all those reports that need to be routine generated can be done on a click of a button. Further, DHIS 2 can help in the generation of analytical reports through comparisons of for example indicators across facilities or over time. Graphs, maps, reports and health profiles are among the outputs that DHIS 2 can produce, and these should routinely be produced, analysed, and acted upon by health managers.

7.4. Technical background

7.4.1. DHIS as a platform

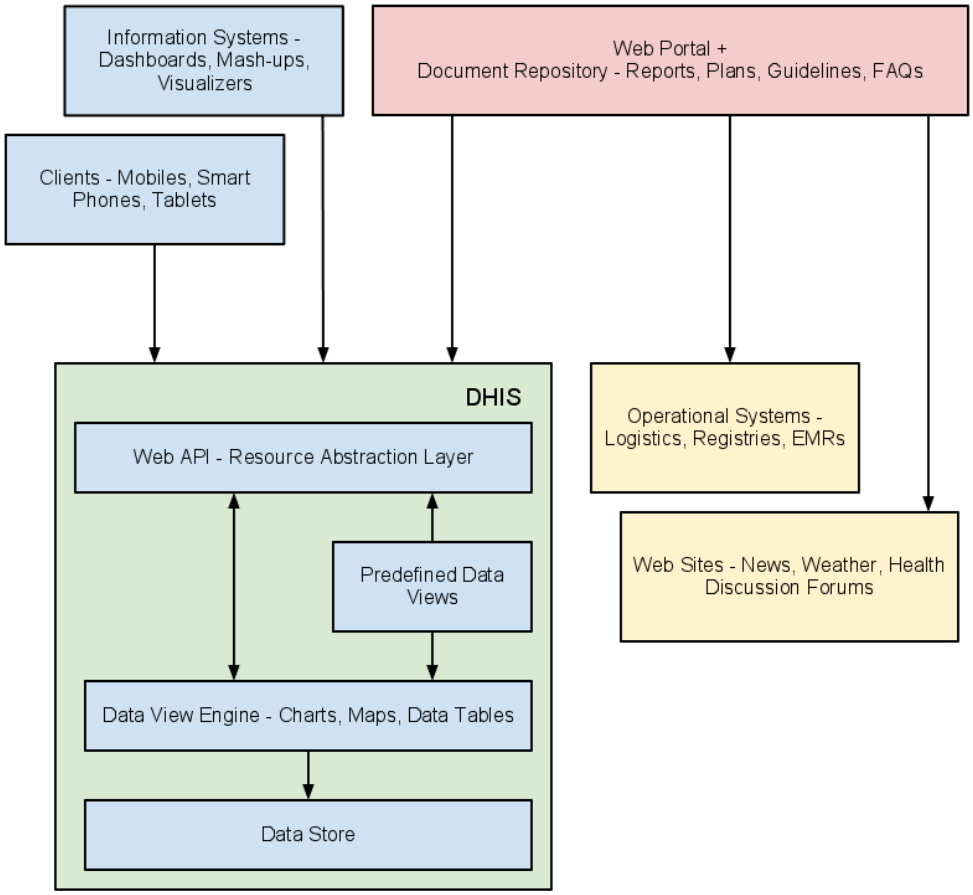
DHIS can be perceived as a platform on several levels. First, the application database is designed ground-up with flexibility in mind. Data structures such as data elements, organisation units, forms and user roles can be defined completely freely through the application user interface. This makes it possible for the system to be adapted to a multitude of locale contexts and use-cases. We have seen that DHIS supports most major requirements for routine data capture and analysis emerging in country implementations. It also makes it possible for DHIS to serve as management system for domains such as logistics, labs and finance.

Second, due to the modular design of DHIS it can be extended with additional software modules. These software modules can live side by side with the core modules of DHIS and can be integrated into the DHIS portal and menu system. This is a powerful feature as it makes it possible to extend the system with extra functionality when needed, typically for country specific requirements as earlier pointed out.

The downside of the software module extensibility is that it puts several constraints on the development process. The developers creating the extra functionality are limited to the DHIS technology in terms of programming language and software frameworks, in addition to the constraints put on the design of modules by the DHIS portal solution. Also, these modules must be included in the DHIS software when the software is built and deployed on the web server, not dynamically during run-time.

In order to overcome these limitations and achieve a looser coupling between the DHIS service layer and additional software artifacts, the DHIS development team decided to create a Web API. This Web API complies with the rules of the REST architectural style. This implies that:

- The Web API provides a navigable and machine-readable interface to the complete DHIS data model. For instance, one can access the full list of data elements, then navigate using the provided hyperlink to a particular data element of interest, then navigate using the provided hyperlink to the list of forms which this data element is part of. E.g. clients will only do state transitions using the hyperlinks which are dynamically embedded in the responses.
- Data is accessed through a uniform interface (URLs) using a well-known protocol. There are no fancy transport formats or protocols involved - just the well-tested, well-understood HTTP protocol which is the main building block of the Web today. This implies that third-party developers can develop software using the DHIS data model and data without knowing the DHIS specific technology or complying with the DHIS design constraints.
- All data including meta-data, reports, maps and charts, known as resources in REST terminology, can be retrieved in most of the popular representation formats of the Web of today, such as HTML, XML, JSON, PDF and PNG. These formats are widely supported in applications and programming languages and gives third-party developers a wide range of implementation options.



7.4.2. Understanding platform independence

All computers have an Operating System (OS) to manage it and the programs running it. The operating system serves as the middle layer between the software application, such as DHIS 2, and the hardware, such as the CPU and RAM. DHIS 2 runs on the Java Virtual Machine, and can therefore run on any operating system which supports Java. Platform independence implies that the software application can run on ANY OS - Windows, Linux, Macintosh etc. DHIS 2 is platform independent, and is extremely useful in the context of public health, where multiple operating systems may be in use.

Furthermore, DHIS 2 is also platform independent when it comes to the Database Management System (DBMS). DHIS 2 uses the Hibernate database abstraction framework and is compatible with any DBMS supported by Hibernate, such as PostgreSQL, MySQL, H2, MS SQL Server, Oracle and many more. PostgreSQL is the recommended DBMS for DHIS 2.

Lastly, and perhaps most importantly, since DHIS2 is a browser-based application, the only real requirement to interact with the system is with a web browser. DHIS2 supports most web browsers, although currently either Google Chrome, Mozilla Firefox or Opera are recommended.

7.4.3. Deployment strategies - online vs offline

DHIS 2 is a network enabled application and can be accessed over the Internet, a local intranet and as a locally installed system. The deployment alternatives for DHIS 2 are in this chapter defined as i) offline deployment ii) online deployment and iii) hybrid deployment. The meaning and differences will be discussed in the following sections.

7.4.3.1. Offline Deployment

An offline deployment implies that multiple standalone offline instances are installed for end users, typically at the district level. The system is maintained primarily by the end users/district health officers who enters data and generate reports from the system running on their local server. The system will also typically be maintained by a national super-

user team who pay regular visits to the district deployments. Data is moved upwards in the hierarchy by the end users producing data exchange files which are sent electronically by email or physically by mail or personal travel. (Note that the brief Internet connectivity required for sending emails does not qualify for being defined as online). This style of deployment has the obvious benefit that it works when appropriate Internet connectivity is not available. On the other side there are significant challenges with this style which are described in the following section.

- **Hardware:** Running stand-alone systems requires advanced hardware in terms of servers and reliable power supply to be installed, usually at district level, all over the country. This requires appropriate funding for procurement and plan for long-term maintenance.
- **Software platform:** Local installs implies a significant need for maintenance. From experience, the biggest challenge is viruses and other malware which tend to infect local installations in the long-run. The main reason is that end users utilize memory sticks for transporting data exchange files and documents between private computers, other workstations and the system running the application. Keeping anti-virus software and operating system patches up to date in an offline environment are challenging and bad practices in terms of security are often adopted by end users. The preferred way to overcome this issue is to run a dedicated server for the application where no memory sticks are allowed and use an Linux based operating system which is not as prone for virus infections as MS Windows.
- **Software application:** Being able to distribute new functionality and bug-fixes to the health information software to users are essential for maintenance and improvement of the system. Relying on the end users to perform software upgrades requires extensive training and a high level of competence on their side as upgrading software applications might a technically challenging task. Relying on a national super-user team to maintain the software implies a lot of traveling.
- **Database maintenance:** A prerequisite for an efficient system is that all users enter data with a standardized meta-data set (data elements, forms etc). As with the previous point about software upgrades, distribution of changes to the meta-data set to numerous offline installations requires end user competence if the updates are sent electronically or a well-organized super-user team. Failure to keep the meta-data set synchronized will lead to loss of ability to move data from the districts and/or an inconsistent national database since the data entered for instance at the district level will not be compatible with the data at the national level.

7.4.3.2. Online deployment

An online deployment implies that a single instance of the application is set up on a server connected to the Internet. All users (clients) connect to the online central server over the Internet using a web browser. This style of deployment currently benefits from the huge investments in and expansions of mobile networks in developing countries. This makes it possible to access online servers in even the most rural areas using mobile Internet modems (also referred to as *dongles*).

This online deployment style has huge positive implications for the implementation process and application maintenance compared to the traditional offline standalone style:

- **Hardware:** Hardware requirements on the end-user side are limited to a reasonably modern computer/laptop and Internet connectivity through a fixed line or a mobile modem. There is no need for a specialized server for each user, any Internet enabled computer will be sufficient. A server will be required for online deployments, but since there is only one (or several) servers which need to be procured and maintained, this is significantly simpler (and cheaper) than maintaining many separate servers is disparate locations.
- **Software platform:** The end users only need a web browser to connect to the online server. All popular operating systems today are shipped with a web browser and there is no special requirement on what type or version. This means that if severe problems such as virus infections or software corruption occur one can always resort to re-formatting and installing the computer operating system or obtain a new computer/laptop. The user can continue with data entry where it was left and no data will be lost.
- **Software application:** The central server deployment style means that the application can be upgraded and maintained in a centralized fashion. When new versions of the applications are released with new features and bug-fixes it can be deployed to the single online server. All changes will then be reflected on the client side the next time end users connect over the Internet. This obviously has a huge positive impact for the process of improving the system as new features can be distributed to users immediately, all users will be accessing the same application version, and bugs and issues can be sorted out and deployed on-the-fly.
- **Database maintenance:** Similar to the previous point, changes to the meta-data can be done on the online server in a centralized fashion and will automatically propagate to all clients next time they connect to the server. This

effectively removes the vast issues related to maintaining an upgraded and standardized meta-data set related to the traditional offline deployment style. It is extremely convenient for instance during the initial database development phase and during the annual database revision processes as end users will be accessing a consistent and standardized database even when changes occur frequently.

This approach might be problematic in cases where Internet connectivity is volatile or missing in long periods of time. DHIS2 however has certain features which requires Internet connectivity to be available only part of the time for the system to work properly, such as offline data entry and the MyDatamart tool presented in a separate chapter in this guide, which cater to information flow in situations when Internet connectivity may be challenging.

7.4.3.3. Hybrid deployment

From the discussion so far one realizes that the online deployment style is favourable over the offline style but requires decent Internet connectivity where it will be used. It is important to notice that the mentioned styles can co-exist in a common deployment. It is perfectly feasible to have online as well as offline deployments within a single country. The general rule would be that districts and facilities should access the system online over the Internet where sufficient Internet connectivity exist, and offline systems should be deployed to districts where this is not the case.

Defining decent Internet connectivity precisely is hard but as a rule of thumb the download speed should be minimum 10 Kbyte/second and accessibility should be minimum 70% of the time.

In this regard mobile Internet modems which can be connected to a computer or laptop and access the mobile network is an extremely capable and feasible solution. Mobile Internet coverage is increasing rapidly all over the world, often provide excellent connectivity at low prices and is a great alternative to local networks and poorly maintained fixed Internet lines. Getting in contact with national mobile network companies regarding post-paid subscriptions and potential large-order benefits can be a wort-while effort. The network coverage for each network operator in the relevant country should be investigated when deciding which deployment approach to opt for as it might differ and cover different parts of the country.

7.4.3.4. Server hosting

The online deployment approach raises the question of where and how to host the server which will run the DHIS 2 application. Typically there are several options:

1. Internal hosting within the Ministry of Health
2. Hosting within a government data centre
3. Hosting through an external hosting company

The main reason for choosing the first option is often political motivation for having “physical ownership” of the database. This is perceived as important by many in order to “own” and control the data. There is also a wish to build local capacity for server administration related to sustainability of the project. This is often a donor-driven initiatives as it is perceived as a concrete and helpful mission.

Regarding the second option, some places a government data centre is constructed with a view to promoting and improving the use and accessibility of public data. Another reason is that a proliferation of internal server environments is very resource demanding and it is more effective to establish centralized infrastructure and capacity.

Regarding external hosting there is lately a move towards outsourcing the operation and administration of computer resources to an external provider, where those resources are accessed over the network, popularly referred to as “cloud computing” or “software as a service”. Those resources are typically accessed over the Internet using a web browser.

The primary goal for an online server deployment is provide long-term stable and high-performance accessibility to the intended services. When deciding which option to choose for server environment there are many aspects to consider:

1. Human capacity for server administration and operation. There must be human resources with general skills in server administration and in the specific technologies used for the application providing the services. Examples of such technologies are web servers and database management platforms.
2. Reliable solutions for automated backups, including local off-server and remote backup.
3. Stable connectivity and high network bandwidth for traffic to and from the server.

4. Stable power supply including a backup solution.
5. Secure environment for the physical server regarding issues such as access, theft and fire.
6. Presence of a disaster recovery plan. This plan must contain a realistic strategy for making sure that the service will be only suffering short down-times in the events of hardware failures, network downtime and more.
7. Feasible, powerful and robust hardware.

All of these aspects must be covered in order to create an appropriate hosting environment. The hardware requirement is deliberately put last since there is a clear tendency to give it too much attention.

Looking back at the three main hosting options, experience from implementation missions in developing countries suggests that all of the hosting aspects are rarely present in option one and two at a feasible level. Reaching an acceptable level in all these aspects is challenging in terms of both human resources and money, especially when compared to the cost of option three. It has the benefit that is accommodates the mentioned political aspects and building local capacity for server administration, on the other hand can this be provided for in alternative ways.

Option three - external hosting - has the benefit that it supports all of the mentioned hosting aspects at a very affordable price. Several hosting providers - of virtual servers or software as a service - offer reliable services for running most kinds of applications. Example of such providers are Linode and Amazon Web Services. Administration of such servers happens over a network connection, which most often anyway is the case with local server administration. The physical location of the server in this case becomes irrelevant as that such providers offer services in most parts of the world. This solution is increasingly becoming the standard solution for hosting of application services. The aspect of building local capacity for server administration is compatible with this option since a local ICT team can be tasked with maintaining the externally hosted server, but with not being burdened with worrying about power supply and bandwidth constraints which usually exist outside of major data centres.

An approach for combining the benefits of external hosting with the need for local hosting and physical ownership is to use an external hosting provider for the primary transactional system, while mirroring this server to a locally hosted non-critical server which is used for read-only purposes such as data analysis and accessed over the intranet.

7.5. Difference between Aggregated and Patient data in a HIS

Patient data is data relating to a single patient, such as his/her diagnosis, name, age, earlier medical history etc. This data is typically based on a single patient-health care worker interaction. For instance, when a patient visits a health care clinic, a variety of details may be recorded, such as the patient's temperature, their weight, and various blood tests. Should this patient be diagnosed as having "Vitamin B 12 deficiency anaemia, unspecified" corresponding to ICD-10 code D51.9, this particular interaction might eventually get recorded as an instance of "Anaemia" in an aggregate based system. Patient based data is important when you want to track longitudinally the progress of a patient over time. For example, if we want to track how a patient is adhering to and responding to the process of TB treatment (typically taking place over 6-9 months), we would need patient based data.

Aggregated data is the consolidation of data relating to multiple patients, and therefore cannot be traced back to a specific patient. They are merely counts, such as incidences of Malaria, TB, or other diseases. Typically, the routine data that a health facility deals with is this kind of aggregated statistics, and is used for the generation of routine reports and indicators, and most importantly, strategic planning within the health system. Aggregate data cannot provide the type of detailed information which patient level data can, but is crucial for planning and guidance of the performance of health systems.

In between the two you have case-based data, or anonymous "patient" data. A lot of details can be collected about a specific health event without necessarily having to identify the patient it involved. Inpatient or outpatient visits, a new case of cholera, a maternal death etc. are common use-cases where one would like to collect a lot more detail that just adding to the total count of cases, or visits. This data is often collected in line-listing type of forms, or in more detailed audit forms. In is different from aggregate data in the sense that it contains many details about a specific event, whereas the aggregate data would count how many events of a certain type, e.g. how many outpatient visits with principal diagnosis "Malaria", or how many maternal deaths where the deceased did not attend ANC, or how many cholera outbreaks for children under 5 years. In DHIS 2 this data is collected through programs of the type single event without registration.

Patient data is highly confidential and therefore must be protected so that no one other than doctors can get it. When in paper, it must be properly stored in a secure place. For computers, patient data needs secure systems with passwords, restrained access and audit logs.

Security concerns for aggregated data are not as crucial as for patient data, as it is usually impossible to identify a particular person to a aggregate statistic . However, data can still be misused and misinterpreted by others, and should not be distributed without adequate data dissemination policies in place.

7.6. Free and Open Source Software (FOSS): benefits and challenges

Software carries the instructions that tell a computer how to operate. The human authored and human readable form of those instructions is called source code. Before the computer can actually execute the instructions, the source code must be translated into a machine readable (binary) format, called the object code. All distributed software includes the object code, but FOSS makes the source code available as well.

Proprietary software owners license their copyrighted object code to a user, which allows the user to run the program. FOSS programs, on the other hand, license both the object and the source code, permitting the user to run, modify and possibly redistribute the programs. With access to the source code, the users have the freedom to run the program for any purpose, redistribute, probe, adapt, learn from, customise the software to suit their needs, and release improvements to the public for the good of the community. Hence, some FOSS is also known as free software, where “free” refers, first and foremost, to the above freedoms rather than in the monetary sense of the word.

Within the public health sector, FOSS can potentially have a range of benefits, including:

- Lower costs as it does not involve paying for prohibitive license costs.
- Given the information needs for the health sector are constantly changing and evolving, there is a need for the user to have the freedom to make the changes as per the user requirements. This is often limited in proprietary systems.
- Access to source code to enable integration and interoperability. In the health sector interoperability between different software applications is becoming increasingly important, meaning enabling two or more systems to communicate metadata and data. This work is a lot easier, and sometimes dependent on the source code being available to the developers that create the integration. This availability is often not possible in the case of proprietary software. And when it is, it comes at a high cost and contractual obligations.
- FOSS applications like DHIS 2 typically are supported by a global network of developers, and thus have access to cutting edge research and development knowledge.

Chapter 8. Getting started with DHIS 2

8.1. Getting started with DHIS 2

After reading this chapter you will be able to understand:

- Start DHIS 2 from the desktop
- How to log-in from the desktop
- Create new users and user roles
- What steps are needed to design a DHIS 2 database for your organisation

8.1.1. Prerequisites

You must be sure that you have a current version of the Java Runtime installed on your machine. Depending on your operating system, there are different ways of installing Java. The reader is referred to this website for detailed information on getting Java installed.

8.1.2. Starting the DHIS 2 Live package

The DHIS 2 Live package is the easiest way to get started with DHIS2. DHIS2 Live is appropriate for a stand-alone installation and demos. Simply download the application from here. Once the file is downloaded, you can simply double-click the downloaded file, and get started using DHIS 2.

8.1.2.1. Starting up with a blank database

The live package comes with a demo database just like what you see on the online demo (which is based on the national Sierra Leone HMIS), and if you want to start with a blank system/database and build up your own system then you need to do the following:

1) Stop DHIS2 live if it is already running. Right click on the tray icon and select Exit. The tray icon is the green symbol on the bottom right of your screen (on Windows) which should say' DHIS 2 Server running' if you hold your mouse pointer over it.

2) Open the folder where the DHIS 2 live package is installed and locate the folder called "conf".

3) In conf/ open the file called 'hibernate.properties' in a text editor (notepad or similar) and do the following modification: locate the string 'jdbc:h2:./database/dhis2' and replace the 'dhis2' part with any name that you want to give to your database (e.g. dhis2_test).

4) Save and close the hibernate.properties file.

5) Start DHIS 2 Live by double-clicking on the file dhis2-live.exe in the DHIS 2 Live installation folder or by using a desktop shortcut or menu link that you might have set up.

6) Wait for the browser window to open and the login screen to show, and then log in with username: admin and password: district

7) Now you will see a completely empty DHIS 2 system and you should start by adding your users, organisational hierarchy, data elements, and datasets etc. Please refer to the other sections of the user manual for instructions on how to do this.

8.1.3. Working directly with the H2 database

DHIS 2 Live uses an embedded H2 database. This has several advantages - there is no need to install a separate database engine such as PostgreSQL or MySQL, and backup can be made by just copying the file. The whole database exists

in memory, which means high performance. The disadvantage is need for RAM. It is also not suitable for multi-user server installations.

In general, it is recommended to work with the database through the DHIS2 user interface, but in some situations one may need to manipulate the data directly. If one downloads H2 separately, it comes with a web interface. It can also be manipulated using OpenOffice.org, using the following procedure. This assumes that dhis2-live is located in the user's Linux home directory (represented by ~). Substitute the absolute path to your dhis2-live installation.

- Start OpenOffice Word Processor and select Tools - Options, then Java - Class Path ... and click on Add Archive...
- Select the following file (version may differ): ~/dhis2-live/webapps/dhis/WEB-INF/lib/h2-1.1.119.jar
- Close OpenOffice completely and then open OpenOffice.org Database. Select connect to an existing database - JDBC
- Datasource URL is h2:~/dhis2-live/database/dhis2;AUTO_SERVER=TRUE, and JDBC driver class is org.h2.Driver
- User name is sa, password not needed. Finally, select a name and folder for the .odb file.

More tips

8.1.4. Downloading and installing the server version

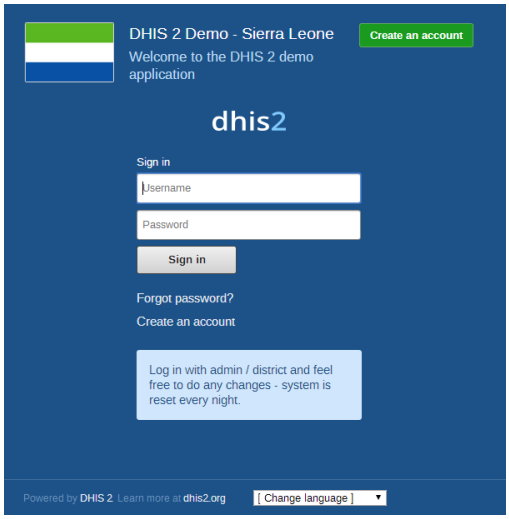
The latest stable server version can be downloaded from this website. For detailed information on how to install it please refer to the installation chapter in the implementation manual.

8.2. Logging on to DHIS 2

Regardless of whether you have installed the server version of the desktop Live version, you will use a web-browser to log on to the application. DHIS2 should be compatible with most modern web-browsers, although you will need to ensure that Java Script is enabled.

To log on to the application just enter http://localhost:8080/dhis if you are using the DHIS2 live package, or replace localhost with the name or IP address of the server where the server version is installed.

Once you have started DHIS2, either on-line or off-line, the displayed screen will prompt you to enter your registered username and password. After entering the required information click on log-in button to log into the application. The default user name and password are 'admin' and 'district'. They should be changed immediately upon logging on the first time.



You can select the language which you wish to display DHIS2 in from the "Change language" dialog box at the bottom of the screen. Not all languages may be available.

Should you have forgotten your password, you can click on the "Forgot password?" link. You must have informed DHIS2 of your email address and the server must be properly configured to send emails.

If you want to create your own account (and the server administrator allows this), simply click "Create an account" and follow the directions provided.

Once you have logged into DHIS2, refer to the specific sections in this manual for the different functionality which is available.

8.3. Logging out of DHIS 2

Just click on the Profile->Log in the top-right corner of the DHIS2 menu.

8.4. Quick intro to designing a DHIS 2 database

The DHIS 2 application comes with a set of tools for data collection, validation, reporting and analysis, but the contents of the database, e.g. what to collect, who should collect it and on what format will depend on the context of use. This metadata need to be populated into the application before it can be used, and this can be done through the user interface and requires no programming or in-depth technical skills of the software. We call this initial process database design or customisation.

This section will provide a very quick and brief introduction to DHIS 2 database design and mainly explain the various steps needed to prepare a new DHIS 2 system for use. How to do each step is explained in other chapters, and best practices on design choices will be explained in an implementers manual (expected during first half of 2011). Here are the steps to follow:

1. Set up an organisational hierarchy

2. Define data elements

3. Define data sets and data entry forms

4. Define validation rules

5. Define indicators

6. Define report tables and design reports

7. Set up the GIS module

8. Design charts and customise the dashboard

8.4.1. The organisational hierarchy

The organisational hierarchy defines the organisation using the DHIS 2, the health facilities, administrative areas and other geographical areas used in data collection and data analysis. This dimension to the data is defined as a hierarchy with one root unit (e.g. Ministry of Health) and any number of levels and nodes below. Each node in this hierarchy is called an organisational unit in DHIS 2.

The design of this hierarchy will determine the geographical units of analysis available to the users as data is collected and aggregated in this structure. There can only be one organisational hierarchy at the same time so its structure needs careful consideration. Additional hierarchies (e.g. parallel administrative groupings such as "Facility ownership") can be modelled using organisational groups and group sets, however the organisational hierarchy is the main vehicle for data aggregation on the geographical dimension. Typically national organisational hierarchies in public health have 4-6 levels, but any number of levels is supported. The hierarchy is built up of parent-child relations, e.g. a Country or MoH unit (the root) might have e.g. 8 parent units (provinces), and each province again (at level 2) might have 10-15 districts as their children. Normally the health facilities will be located at the lowest level, but they can also be located

at higher levels, e.g. national or provincial hospitals, so skewed organisational trees are supported (e.g. a leaf node can be positioned at level 2 while most other leaf nodes are at level 5).

Typically there is a geographical hierarchy defined by the health system. e.g. where the administrative offices are located (e.g. MoH, province, district), but often there are other administrative boundaries in the country that might or might not be added, depending on how its boundaries will improve data analysis. When designing the hierarchy the number of children for any organisational unit may indicate the usefulness of the structure, e.g. having one or more 1-1 relationships between two levels is not very useful as the values will be the same for the child and the parent level. On the other extreme a very high number of children in the middle of the hierarchy (e.g. 50 districts in a province) might call for an extra level to be added in between to increase the usefulness of data analysis. The lowest level, the health facilities will often have a large number of children (10-60), but for other levels higher up in the hierarchy approx. 5-20 children is recommended. Too few or too many children might indicate that a level should be removed or added.

Note that it is quite easy to make changes to the upper levels of the hierarchy at a later stage, the only problem is changing organisational units that collect data (the leaf nodes), e.g. splitting or merging health facilities. Aggregation up the hierarchy is done based on the current hierarchy at any time and will always reflect the most recent changes to the organisational structure. Refer to the chapter on Organisation Units to learn how to create organisational units and to build up the hierarchy.

8.4.2. Data Elements

The Data Element is perhaps the most important building block of a DHIS 2 database. It represents the "WHAT" dimension, it explains what is being collected or analysed. In some contexts this is referred to an indicator, but in DHIS 2 we call this unit of collection and analysis a *data element*. The data element often represents a count of something, and its name describes what is being counted, e.g. "BCG doses given" or "Malaria cases". When data is collected, validated, analysed, reported or presented it is the data elements or expressions built upon data elements that describes the WHAT of the data. As such the data elements become important for all aspects of the system and they decide not only how data is collected, but more importantly how the data values are represented in the database, which again decides how data can be analysed and presented.

It is possible to add more details to this "WHAT" dimension through the disaggregation dimension called data element categories. Some common categories are Age and Gender, but any category can be added by the user and linked to specific data elements. The combination of a data element's name and its assigned category defines the smallest unit of collection and analysis available in the system, and hence describes the raw data in the database. Aggregations can be done when zooming out of this dimension, but no further drill-down is possible, so designing data elements and categories define the detail of the analysis available to the system (on the WHAT dimension). Changes to data elements and categories at a later stage in the process might be complicated as these will change the meaning of the data values already captured in the database (if any). So this step is one of the more decisive and careful steps in the database design process.

One best practice when designing data elements is to think of data elements as a unit of data analysis and not just as a field in the data collection form. Each data element lives on its own in the database, completely detached from the collection form, and reports and other outputs are based on data elements and expressions/formulas composed of data elements and not the data collection forms. So the data analysis needs should drive the process, and not the look an feel of the data collection forms. A simple rule of thumb is that the name of the data element must be able to stand on its own and describe the data value also outside the context of its collection form. E.g. a data element name like "Total referrals" makes sense when looking at it in either the "RCH" form or the "OPD" form, but on its own it does not uniquely describe the phenomena (who are being referred?), and should in stead be called "Total referrals from Maternity" or "Total referrals from OPD". Two different data elements with different meanings, although the field on the paper form might only say "Total referrals" since the user of the form will always know where these referrals come from. In a database or a repository of data elements this context is no longer valid and therefore the names of the data elements become so important in describing the data.

Common properties of data elements can be modelled through what is called data element groups. The groups are completely flexible in the sense that they are defined by the user, both their names and their memberships. Groups are useful both for browsing and presenting related data, but can also be used to aggregate data elements together. Groups are loosely coupled to data elements and not tied directly to the data values which means they can be modified and added at any point in time without interfering with the raw data.

8.4.3. Datasets and data entry forms

All data entry in DHIS 2 is organised through the use of Datasets. A Dataset is a collection of data elements grouped together for data collection, and in the case of distributed installs they also define chunks of data for export and import between instances of DHIS 2 (e.g. from a district office local installation to a national server). Datasets are not linked directly to the data values, only through their data elements and frequencies, and as such a dataset can be modified, deleted or added at any point in time without affecting the raw data already captured in the system, but such changes will of course affect how new data will be collected.

A dataset has a period type which controls the data collection frequency, which can be daily, weekly, monthly, quarterly, six-monthly, or yearly. Both which data elements to include in the dataset and the period type is defined by the user, together with a name, short name, and code.

In order to use a dataset to collect data for a specific orgunit you must assign the orgunit to the dataset, and this mechanism controls which orgunits that can use which datasets, and at the same time defines the target values for data completeness (e.g. how many health facilities in a district expected to submit RCH data every month).

A data element can belong to multiple datasets, but this requires careful thinking as it may lead to overlapping and inconstant data being collected if e.g. the datasets are given different frequencies and are used by the same orgunits.

8.4.3.1. Data entry forms

Once you have assigned a dataset to an orgunit that dataset will be made available in Data Entry (under Services) for the orgunits you have assigned it to and for the valid periods according to the dataset's period type. A default data entry form will then be shown, which is simply a list of the data elements belonging to the dataset together with a column for inputting the values. If your dataset contains data elements with categories such as age groups or gender, then additional columns will be automatically generated in the default form based on the categories. In addition to the default list-based data entry form there are two more alternatives, the section-based form and the custom form.

8.4.3.1.1. Section forms

Section forms allow for a bit more flexibility when it comes to using tabular forms and are quick and simple to design. Often your data entry form will need multiple tables with subheadings, and sometimes you need to disable (grey out) a few fields in the table (e.g. some categories do not apply to all data elements), both of these functions are supported in section forms. After defining a dataset you can define it's sections with subsets of dataelements, a heading and possible grey fields i the section's table. The order of sections in a dataset can also be defined. In Data Entry you can now start using the Section form (should appear automatically when sections are available for the selected dataset). You can switch between default and section forms in the top right corner of the data entry screen. Most tabular data entry forms should be possible to do with sections forms, and the more you can utilise the section forms (or default forms) the easier it is for you. If these two types of forms are not meeting your requirements then the third option is the completely flexible, although more time-consuming, custom data entry forms.

8.4.3.1.2. Custom Forms

When the form you want to design is too complicated for the default or section forms then your last option is to use a custom form. This takes more time, but gives you full flexibility in term of the design. In DHIS 2 there is a built in HTML editor (FcK Editor) for the form designer and you can either design the form in the UI or paste in your html directly (using the Source window in the editor. In the custom form you can insert static text or data fields (linked to data elements + category) in any position on the form and you have complete freedom to design the layout of the form. Once a custom form has been added to a dataset it will be available in data entry and used automatically. You can switch back to default and section (if exists) forms in the top right corner of the data entry screen.

8.4.4. Validation rules

Once you have set up the data entry part of the system and started to collect data then there is time to define data quality checks that help to improve the quality of the data being collected. You can add as many validation rules as you like and these are composed of left and right side expressions that again are composed of data elements, with an operator between the two sides. Typical rules are comparing subtotals to totals of something. E.g. if you have two data

elements "HIV tests taken" and "HIV test result positive" then you know that in the same form (for the same period and organisational unit) the total number of tests must always be equal or higher than the number of positive tests. These rules should be absolute rules meaning that they are mathematically correct and not just assumptions or "most of the time correct". The rules can be run in data entry, after filling each form, or as a more batch like process on multiple forms at the same time, e.g. for all facilities for the previous reporting month. The results of the tests will list all violations and the detailed values for each side of the expression where the violation occurred to make it easy to go back to data entry and correct the values.

8.4.5. Indicators

Indicators represent perhaps the most powerful data analysis feature of the DHIS 2. While data elements represent the raw data (counts) being collected the indicators represent formulas providing coverage rates, incidence rates, ratios and other formula-based units of analysis. An indicator is made up of a factor (e.g. 1, 100, 100, 100 000), a numerator and a denominator, the two latter are both expressions based on one or more data elements. E.g. the indicator "BCG coverage <1 year" is defined a formula with a factor 100, a numerator ("BCG doses given to children under 1 year") and a denominator ("Target population under 1 year"). The indicator "DPT1 to DPT3 drop out rate" is a formula of 100 % x ("DPT1 doses given"- "DPT3 doses given") / ("DPT1 doses given").

Most report modules in DHIS 2 support both data elements and indicators and you can also combine these in custom reports, but the important difference and strength of indicators versus raw data (data element's data values) is the ability to compare data across different geographical areas (e.g. highly populated vs rural areas) as the target population can be used in the denominator.

Indicators can be added, modified and deleted at any point in time without interfering with the data values in the database.

8.4.6. Report tables and reports

Standard reports in DHIS 2 is a very flexible way of presenting the data that has been collected. Data can be aggregated by any organisational unit or orgunit level, by data element, by indicators, as well as over time (e.g. monthly, quarterly, yearly). The report tables are custom data sources for the standard reports and can be flexibly defined in the user interface and later accessed in external report designers such as iReport or BIRT. These report designs can then be set up as easily accessible one-click reports with parameters so that the users can run the same reports e.g. every month when new data is entered, and also be relevant to users at all levels as the organisational unit can be selected at the time of running the report.

8.4.7. GIS

In the integrated GIS module you can easily display your data on maps, both on polygons (areas) and as points (health facilities), and either as data elements or indicators. By providing the coordinates of your organisational units to the system you can quickly get up to speed with this module. See the GIS section for details on how to get started.

8.4.8. Charts and dashboard

On of the easiest way to display your indicator data is through charts. An easy to use chart dialogue will guide you through the creation of various types of charts with data on indicators, organisational units and periods of your choice. These charts can easily be added to one of the four chart sections on your dashboard and there be made easily available right after log in. Make sure to set the dashboard module as the start module in user settings.

Chapter 9. Data entry

9.1. Data entry with DHIS 2

To open the data entry window click on the services tab displayed in the main menu. A drop down menu will appear listing the services provided by DHIS 2. To use the data entry module, simply search for "Data Entry" in the "Apps" menu.

The data entry module is where data is manually registered in the DHIS 2 database. Data is registered for an organisation unit (also known as an orgunit) , a period, and a set of data elements at a time.

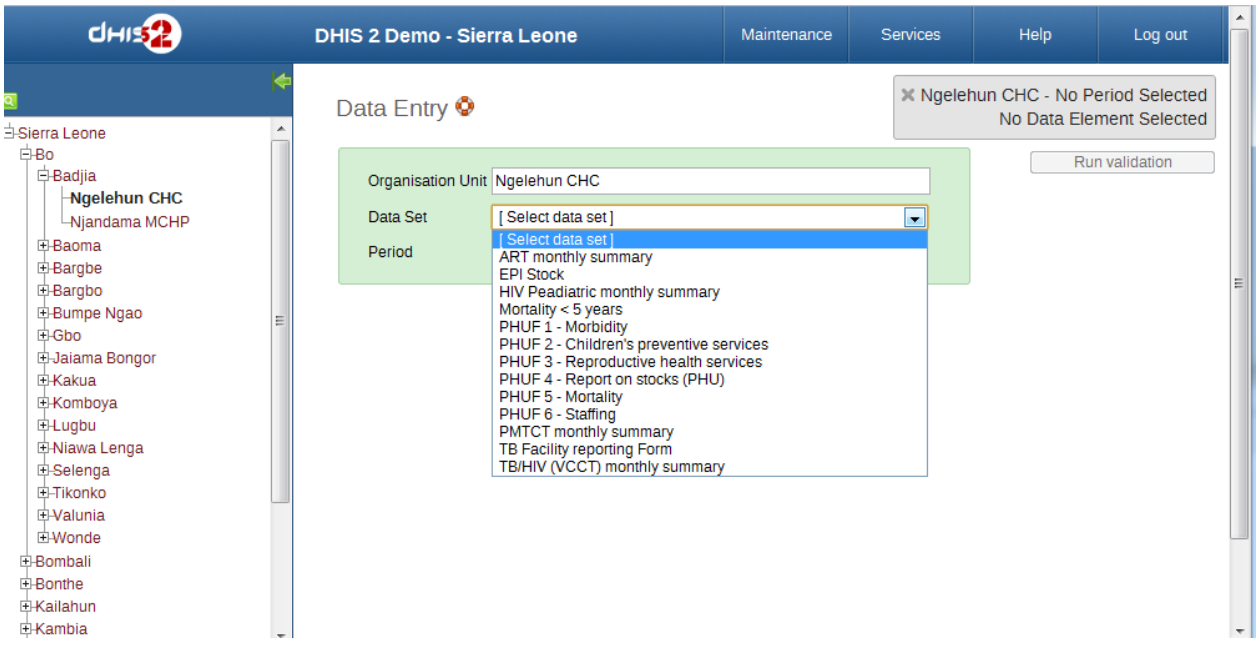
In the context of the Deans Dashboard, you would normally be entering data for a specific unit, such as a faculty or school.

9.1.1. Selecting the data entry form

To start entering data, the first step is to open the correct form by following these steps:

1. Locate the orgunit you want to register data for in the tree menu to the left. You can expand a branch of the organisation unit tree by clicking on the + symbol to the left of the name of the orgunit you are interested in. You can close a branch of thr orgunit tree by clicking on the - symbol to the left of the orgunit you are interested in. A quick way to find an orgunit is to use the search box just above the tree (the green symbol). Once you have located the orgunit, click on the name. The name of the orgunit should now be highlighted in orange when selected.
2. Once you have selected an orgunit from the tree, a possible list of data sets will be displayed in the drop -down box called "Datasets". Select a data set from the drop-down list of data sets.
3. Next, select a period to register data for. The available periods are controlled by the period type of the data set (reporting frequency). You can jump a year back or forward by using the arrows above the period. For some data sets, you may not be able to enter data in the future.
4. Certain datasets may have an attribute. If the dataset has one, it will be displayed below the period drop-down box after you select the period you are entering data for. Select the appropriate attribute.

By now you should see the data entry form. From a form design perspective, there are three types for forms: default forms, section forms and custom forms. If a custom form exists, it will be displayed, followed in order of precedence by a section form, and finally a default form.



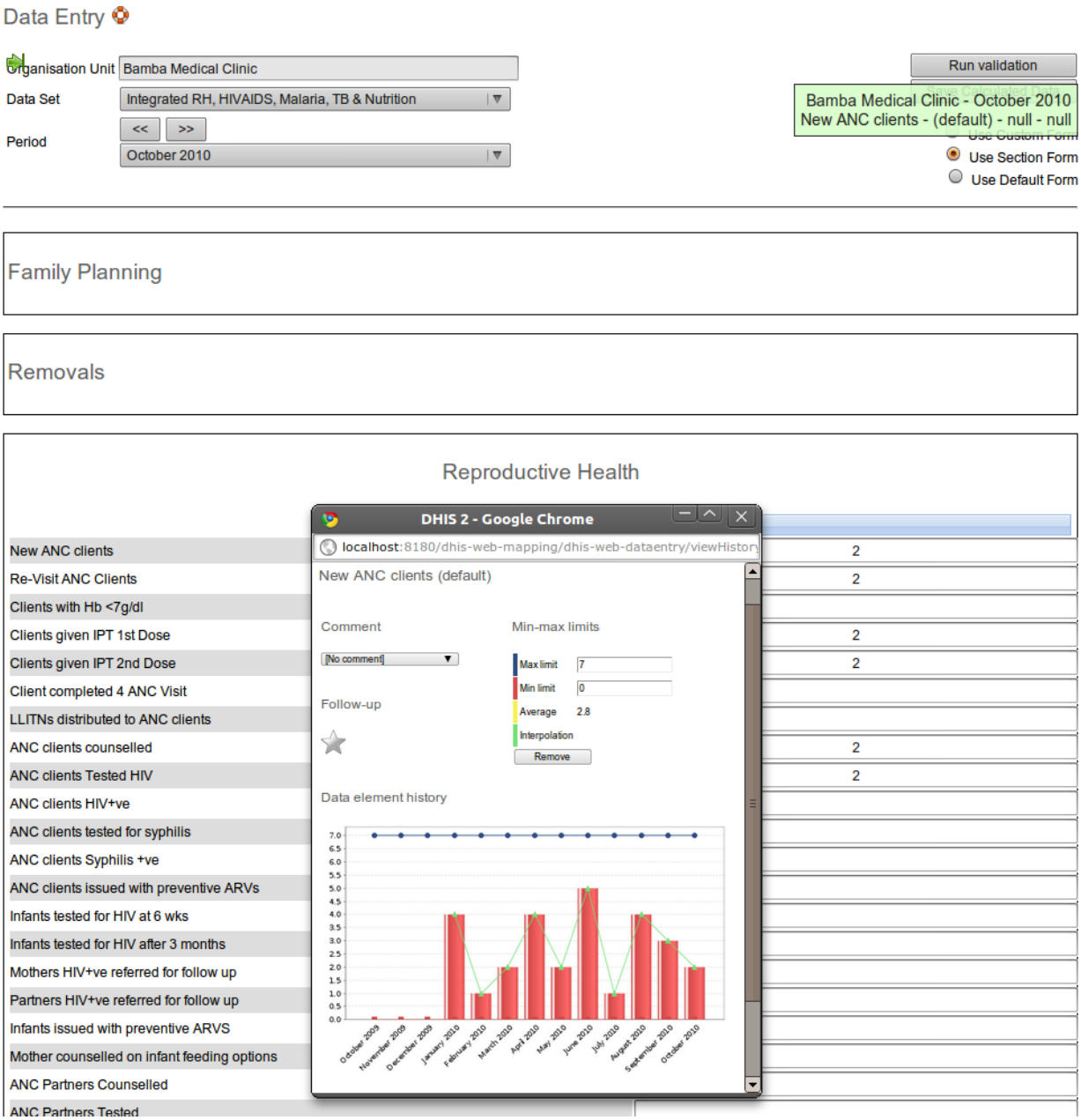
9.1.2. Entering data

Start entering data by clicking inside the first field and type in the value. Move to the next field using the Tab button. Shift+Tab will take you back one step. You can also use the "up" and "down" arrow keys to navigate between the form cells. The values are saved immediately and do not require to be saved at a later stage. A green field indicates that the value has been saved in the system (on the server).

Input validation: If you type in an invalid value, e.g. a character in a field that only accepts numeric values you will get a pop-up that explains the problem and the field will be coloured yellow (not saved) until you have corrected the value. If you have defined a min/max range for the field (data element+organisation unit combination) a pop-up message will notify you when the value is out of range, and the value will remain unsaved until you have changed the value (or updated the range and then re-entered the value).

Disabled fields: If a field is disabled (grey) it means that the field can and should not be filled. The cursor will automatically jump to the next open field.

Data history: By double-clicking on any input field in the form a data history window opens showing the last 12 values registered for the current field (organisation unit+data element+categoryoptioncombo) in a bar chart. This window also shows the min and max range and allows for adjusting the range for the specific organisation unit and data element combination.



Follow Up: In the data history window there is also a feature to tag or star a value. E.g. a suspicious value that needs further investigation can be kept in the system, but marked for Follow-Up. In the Data Quality module you can run a Follow-Up analysis and view all values marked for Follow-Up, and then later edit the values if proved incorrect.

9.1.3. Editing and deleting data

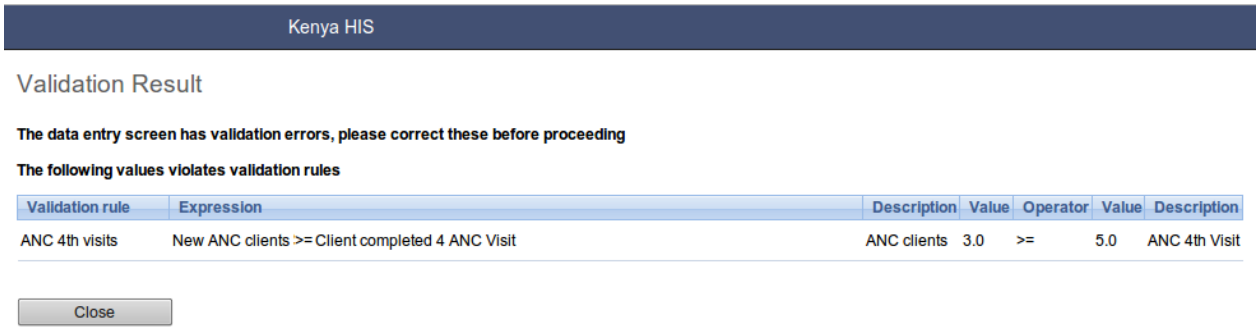
If you wish to enter data which has already been entered, simply replace the data entry value with the update values.

If you want to delete a data value completely, you should select the value of interest, and press "Delete" on your keyboard. If you enter a zero and the data element has been configured to not store zeros, the previous data value (i.e. the one you wish to modify) will not be overwritten with the new value. Therefore, it is better practice to delete the data value completely (waiting for the cell to turn green) and then to enter the new value.

9.1.4. Validating data in the form

When all the available values for the form has been filled in you can run a validation check on the data in the form. Click on the "Run Validation" button in the top right corner. All validation rules which involves data elements in the current form (dataset) will be run against the new data. Upon completion you will be presented with a list of violations or a simply a message that says "The data entry screen successfully passed validation". See the Data Quality chapter for information on how to define such validation rules.

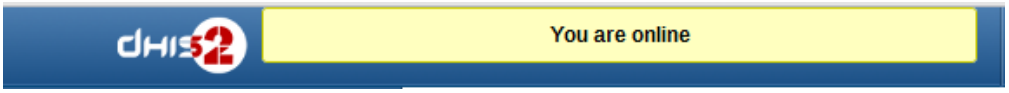
When you have corrected any erroneous values and are done with the form the recommended practice is to click on the Complete button below the form to register the form as complete. This information is used when generating completeness reports for district, county, province or the national level.



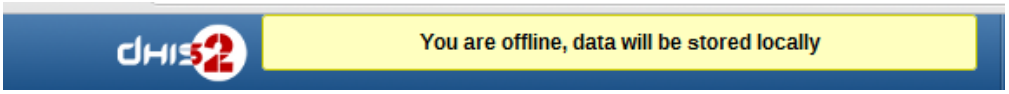
9.1.5. Off-line data entry

The data entry module will function even if during data entry the Internet connectivity is not stable. In order to utilize this functionality, you must login to the server while Internet connectivity is present, but if during data entry, the Internet link between your computer and the server becomes unstable, data can still be entered into the data entry form, saved to your local computer, and then pushed to the server once the Internet connectivity has been restored. Data can be entered and stored locally while being off-line and uploaded to the central server when on-line. This means that the on-line deployment strategy will be more viable in areas with unstable Internet connectivity. The total bandwidth usage is greatly reduced since forms no longer are retrieved from the server for each rendering.

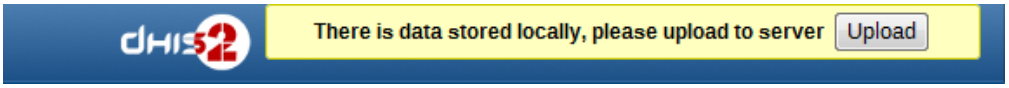
When the server is able to be reached through the Internet connection, a message is displayed at the top of the data entry screen below.



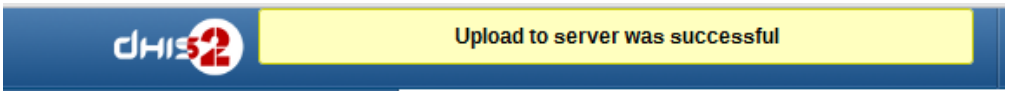
If the Internet connection should disconnect for some reason during the data entry process, this will be detected by the application, and you will be informed that your data will be stored locally.



Data entry can proceed as normal. Once you have entered all of the necessary data, and the application detects that the server is back on-line, you will be informed that you have data which needs to be synchronized with the server.



Once the data has successfully synchronized with the server, you will receive a confirmation message that the data has been successfully uploaded to the server.



9.1.6. Multi-organisation unit data entry

In some scenarios it is beneficial to enter data for multiple organisation units in the same data entry form, for instance if there are few data elements in the form and a huge number of organisation units in the hierarchy. In that case you can enable multi-organisation unit data entry by going to "System settings" and tick the "Enable multi organisation unit forms" setting. Then, in data entry, select the organisation unit immediately above the organisation unit you want to enter for in the hierarchy. Note that this only work for the "section" based forms. You should now see the data elements appearing as columns and the organisation units appearing as rows in the form. Note that the data entry forms should still be assigned to the facilities that you actually enter data for, i.e. the organisation units now appearing in the form.

Chapter 10. Using Data Quality functionality

The data quality module provides means to improve the accuracy and reliability of the data in the system. This can be done through validation rules and various statistical checks. All the functionality described below can be accessed from the left side menu in the Services->Data Quality module.

10.1. Overview of data quality checks

Ensuring data quality is a key concern in building an effective HMIS. Data quality has different dimensions including:

- *Correctness:* Data should be within the normal range for data collected at that facility. There should be no gross discrepancies when compared with data from related data elements.
- *Completeness:* Data for all data elements for all health facilities should have been submitted.
- *Consistency:* Data should be consistent with data entered during earlier months and years while allowing for changes with reorganization, increased work load, etc. and consistent with other similar facilities.
- *Timeliness:* All data from all reporting orgunits should be submitted at the appointed time.

10.2. Data quality checks

Data quality checking can be done through various means, including:

1. At point of data entry, the software can check the data entered to see if it falls within the min-max ranges of that data element (based on all previous data registered).
2. Defining various validation rules, which can be run once the user has finished data entry. The user can also check the entered data for a particular period and Organization Unit(s) against the validation rules, and display the violations for these validation rules.
3. Analysis of data sets, i.e. examining gaps in data.
4. Data triangulation which is comparing the same data or indicator from different sources.

10.3. Running Validation Rule Analysis

You can access Validation Rule Analysis from the Services->Data Quality menu.

A validation rule is based on an expression which defines a relationship between a number of data elements. The expression has a left side and a right side and an operator which defines whether the former must be less than, equal to or greater than the latter. The expression forms a condition which should assert that certain logical criteria are met. For instance, a validation rule could assert that the total number of vaccines given to infants is less than or equal to the total number of infants.

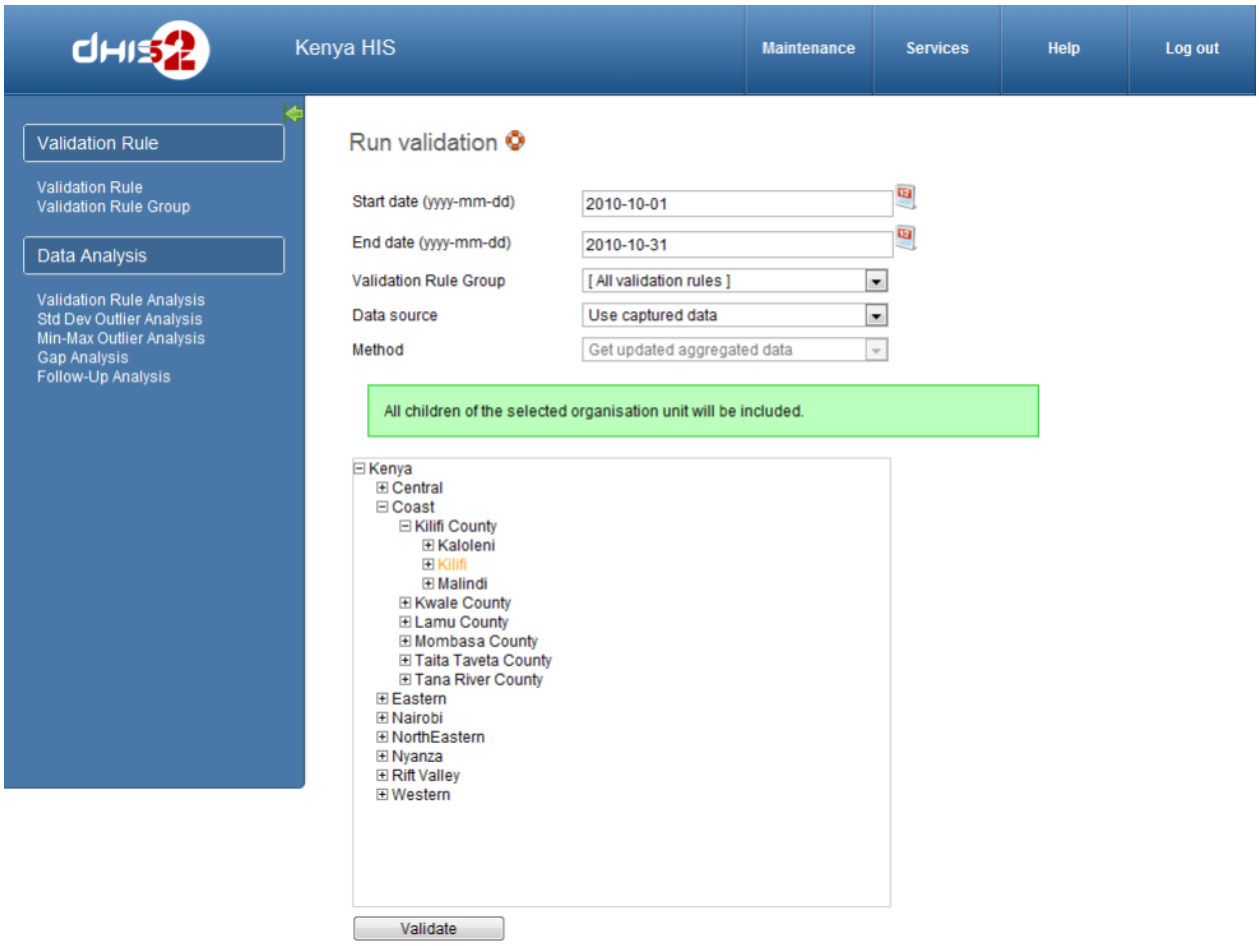
The validation rule analysis function will test validation rules against the data registered in the system. Validation violations will be reported in cases where the condition defined through the validation rule expression is not met, i.e. the condition is false.

Selecting what data to validate:

First, enter a start date and an end date for which data should be included in the analysis. The date picker widget may be used to select dates.

Second, choose between including all validation rules or all validation rules from a single group.

Third, choose between including the selected organisation unit only or the selected organisation unit with all children in the analysis. Fourth, select the organisation unit. Finally, click *validate*.



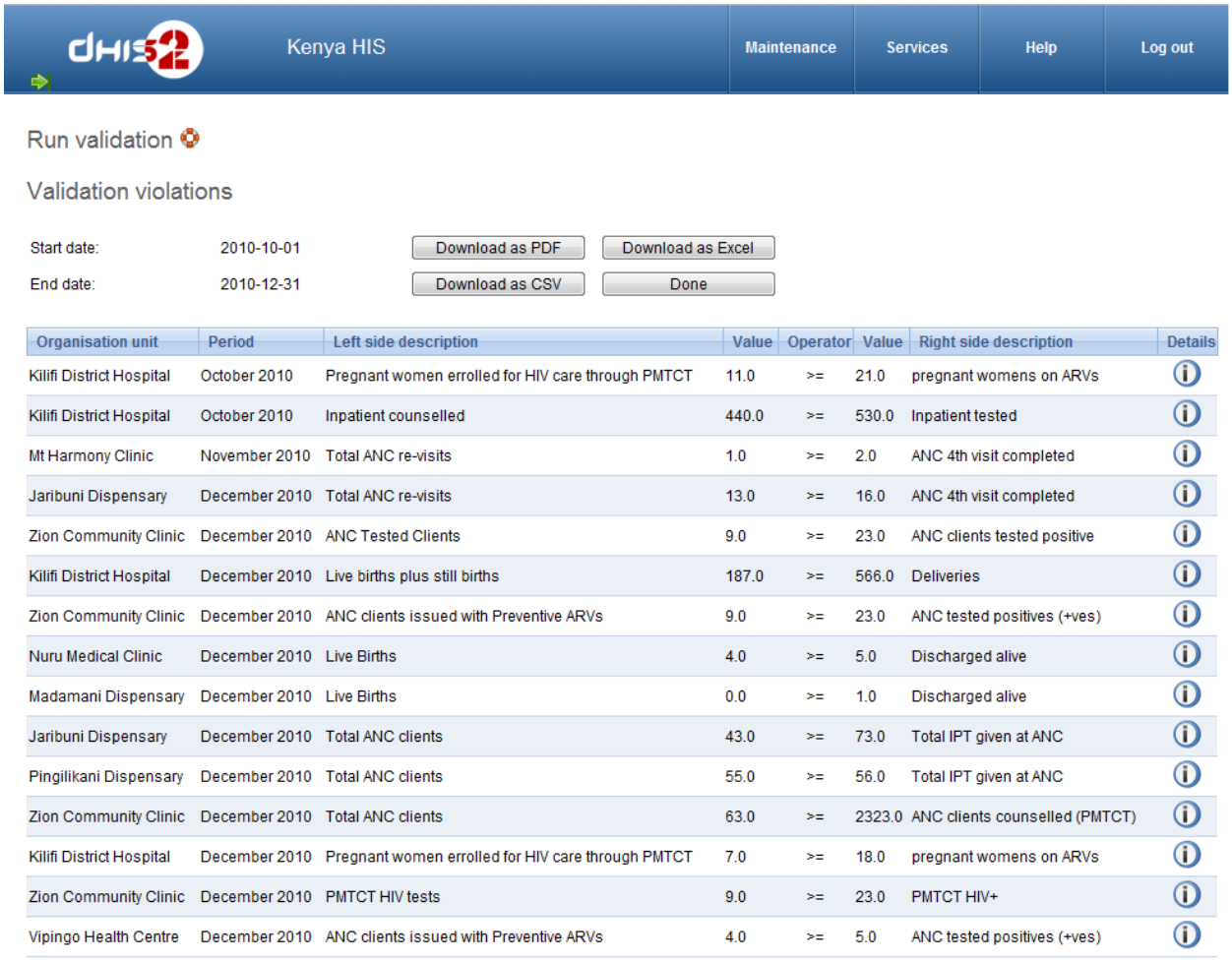
Validation results:

The analysis process will run for a while depending on the amount of data that is being analysed. If there were no violations of the validation rules a message saying *validation passed successfully* is displayed.

If validation violations were found, they will be presented in a list. The organisation unit, period, left side description and value, operator, and right side value and description for each validation violation are displayed.

The show details icon can be clicked in order to get more information about a validation violation. This will open a popup screen that provides information about the data elements included in the validation rules and their corresponding data values. This information can be used in order to fix incorrect data.

The validation violations can be exported to a PDF document by clicking on the Download as PDF button, and to a Microsoft Excel workbook by clicking on the Download as Excel button.



Organisation unit	Period	Left side description	Value	Operator	Value	Right side description	Details
Kilifi District Hospital	October 2010	Pregnant women erolled for HIV care through PMTCT	11.0	>=	21.0	pregnant womens on ARVs	
Kilifi District Hospital	October 2010	Inpatient counselled	440.0	>=	530.0	Inpatient tested	
Mt Harmony Clinic	November 2010	Total ANC re-visits	1.0	>=	2.0	ANC 4th visit completed	
Jaribuni Dispensary	December 2010	Total ANC re-visits	13.0	>=	16.0	ANC 4th visit completed	
Zion Community Clinic	December 2010	ANC Tested Clients	9.0	>=	23.0	ANC clients tested positive	
Kilifi District Hospital	December 2010	Live births plus still births	187.0	>=	566.0	Deliveries	
Zion Community Clinic	December 2010	ANC clients issued with Preventive ARVs	9.0	>=	23.0	ANC tested positives (+ves)	
Nuru Medical Clinic	December 2010	Live Births	4.0	>=	5.0	Discharged alive	
Madamani Dispensary	December 2010	Live Births	0.0	>=	1.0	Discharged alive	
Jaribuni Dispensary	December 2010	Total ANC clients	43.0	>=	73.0	Total IPT given at ANC	
Pingilikani Dispensary	December 2010	Total ANC clients	55.0	>=	56.0	Total IPT given at ANC	
Zion Community Clinic	December 2010	Total ANC clients	63.0	>=	2323.0	ANC clients counselled (PMTCT)	
Kilifi District Hospital	December 2010	Pregnant women erolled for HIV care through PMTCT	7.0	>=	18.0	pregnant womens on ARVs	
Zion Community Clinic	December 2010	PMTCT HIV tests	9.0	>=	23.0	PMTCT HIV+	
Vipingo Health Centre	December 2010	ANC clients issued with Preventive ARVs	4.0	>=	5.0	ANC tested positives (+ves)	

10.4. Std Dev Outlier Analysis

You can access Outlier analysis from the Services->Data Quality menu.

The standard deviation based outlier analysis provides a mechanism for revealing values that are numerically distant from the rest of the data. Outliers can occur by chance, but they often indicate a measurement error or a heavy-tailed distribution (leading to very high numbers). In the former case one wishes to discard them while in the latter case one should be cautious in using tools or interpretations that assume a normal distribution. The analysis is based on the standard normal distribution.

Select what data to analyse:

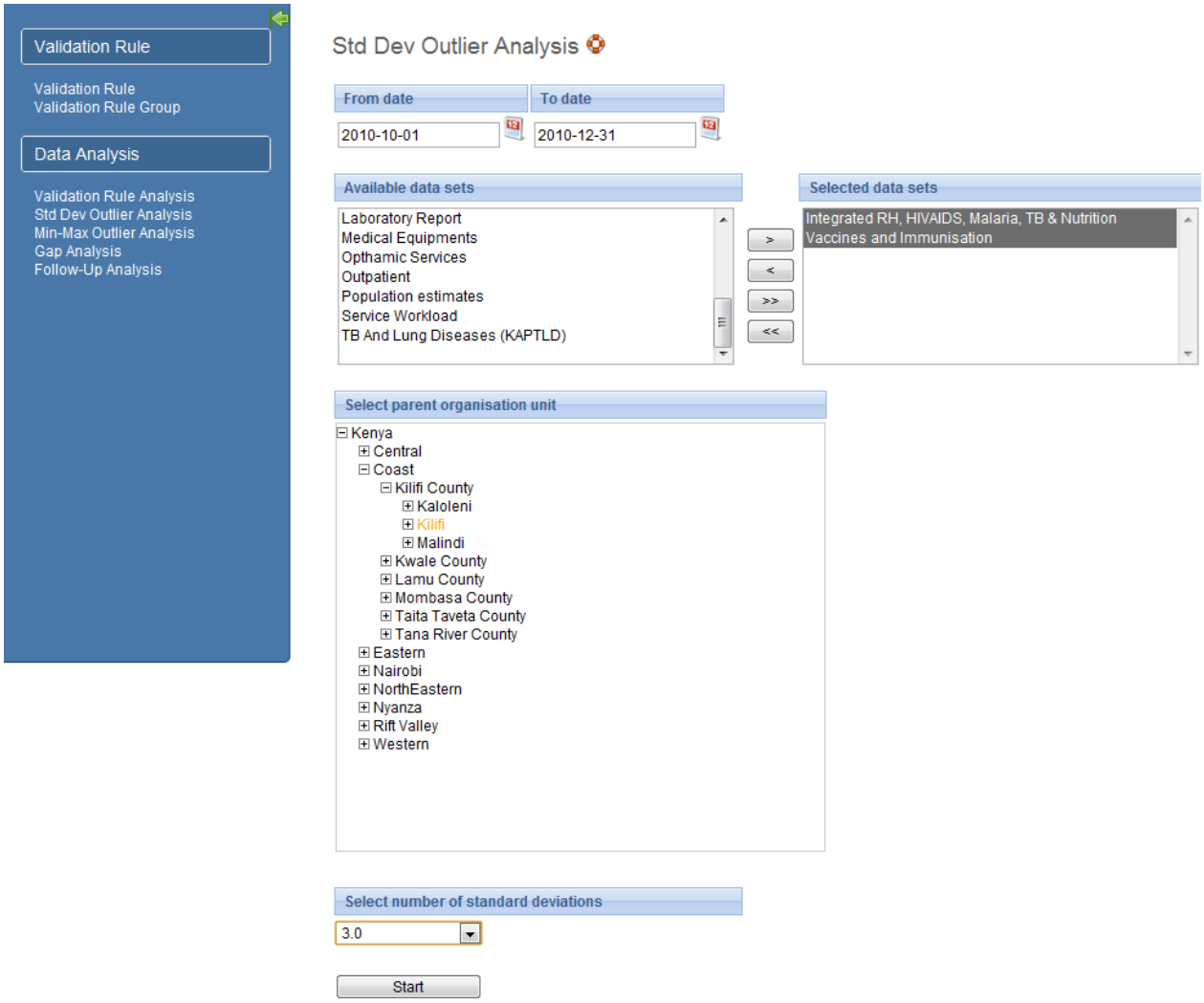
First, select the from and to date for the data to include in the analysis.

Second, select the data set from which to pick data elements from.

Third, select all or some of the data elements in the data set by double-clicking or marking them and clicking the add/remove buttons.

Fourth, select the parent organisation unit to use. All children of the organisation unit will be included.

Fifth, select the number of standard deviations. This refers to the number of standard deviations the data is allowed to deviate from the mean before it is classified as an outlier.



Analysis result:

The potential outlier values discovered will be presented in a list after the analysis process is finished. The data element, organisation unit, period, minimum value, actual value, and maximum value will be displayed for each outlier. The minimum and maximum values refer to the border values derived from the number of standard deviations selected for the analysis.

Each outlier value can be modified directly in the analysis result page. The value can be modified by clicking inside the corresponding field in the value column, entering a value and then navigate away from that field either by clicking tab or anywhere outside the field. The system will provide an alert if the value is still outside the defined minimum and maximum values, but the value will saved in any case. The field will have a red background color if the value is outside the range, and a green if inside.

Each outlier value can be marked for further follow-up by clicking the star icon.

10.5. Min-Max Outlier Analysis

The min-max value based outlier analysis provides a mechanism for revealing values that are outside the pre-defined minimum and maximum values. Minimum and maximum values can be custom defined or automatically defined by the system in the data administration module. See the section about Std dev outlier analysis for further details on usage.

10.6. Gap Analysis

The gap analysis provides a mechanism for revealing gaps in the data. A gap exists for a specific data element and organisation unit. A gap is a period with preceding and succeeding periods which have registered data values, but without registered data values itself. Such a gap might indicate a data capture error or omission and could be further investigated. See the section about Std dev outlier analysis for further details on usage.

10.7. Follow-Up Analysis

The follow-up analysis function will list all data values which are marked for follow-up. A data value can be marked for follow-up in the data entry module and in the other validation analysis variants in this module.

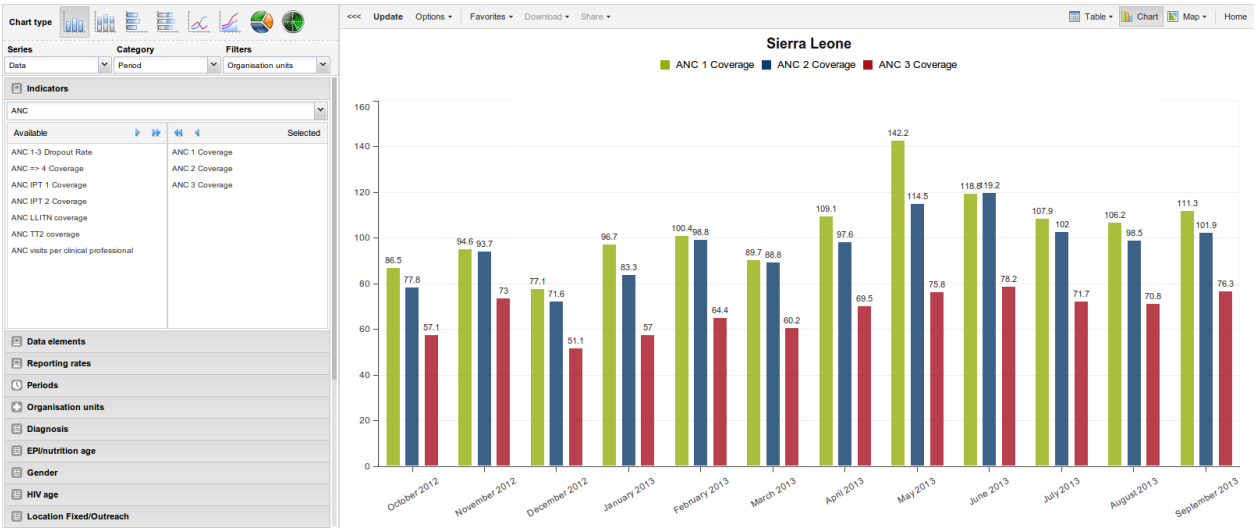
Chapter 11. Using Data Visualizer

11.1. Data Visualizer overview

The data visualizer module enables users to easily create dynamic data analysis and visualizations through charts and data tables. You can freely select content (like indicators, periods and organisation units) for your analysis. This module can be accessed by going to "Services - Data Visualizer" in the main menu. The image below shows the viewport of the module. For a quick start:

1. Look under the "Indicator" heading and select an indicator group from the list of groups.
2. Look under "Available indicators" and select a few indicators from the list by double-clicking on them.
3. Click "Update" in the top bar and see the chart unfold.

The data visualizer is designed firstly to be easy-to-use - you can simply select the indicators, data elements, periods and organisation units you want to include and click "Update" to get your visualization. Secondly it is designed to be fast and work well over poor Internet connections - charts are generated in the web browser and very little data is transferred over the Internet.



11.2. Selecting chart type

The visualizer module provides seven different chart types, each with different characteristics. You can select the type of your chart by clicking on one of the icons in top left bar titled "Chart type".

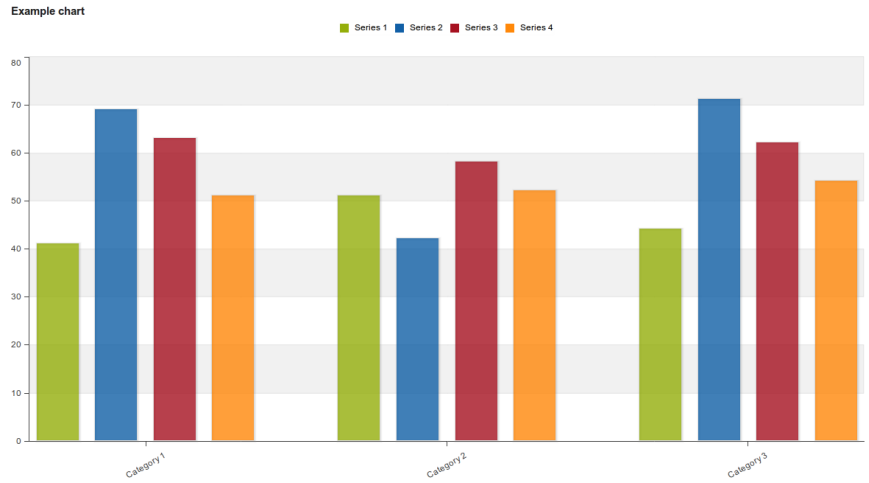
1. Column chart: Chart which displays information as vertical rectangular columns with lengths proportional to the values they represent. Useful e.g. for comparing performance of different districts.
2. Stacked column chart: Chart with vertical rectangular columns where bars representing multiple categories are stacked on top of each other. Useful e.g. for displaying trends or sums of related data elements.
3. Bar chart: Same as column chart, only with horizontal bars.
4. Stacked bar chart: Same as stacked column chart, only with horizontal bars.
5. Line chart: Graph which displays information as a series of points connected by straight lines. Also referred to as time series. Useful e.g. to visualize trends in indicator data over multiple time periods.
6. Area chart: Chart which is based on line chart, with the space between the axis and the line filled with colors and the lines stacked on top of each other. Useful for comparing the trends of related indicators.
7. Pie chart: Circular chart divided into sectors (or slices). Useful e.g. to visualize the proportion of data for individual data elements compared to the total sum of all data elements in the chart.
8. Radar chart: Displaying multivariate data on axes starting from the same point. Also known as spider chart.

11.3. Selecting series, category and filter

This section lets you define which dimension of the data you want to appear as series, category and filter. This asks for a closer explanation. Dimension in this regard refers to the elements which describe the data values in the system. We have three main dimensions in the system:

1. Data: Includes data elements, indicators and datasets (reporting rates), describing the phenomena or event of the data.
2. Periods: Describes when the event took place.
3. Organisation units: Describes where the event took place.

The visualization module lets you use these dimensions completely flexible in terms of appearing as series, categories and filter. Understanding these concepts is most easily done by looking at the screenshot from the opening page below:



More formally this can be described as following:

1. Series: A series is a set of continuous, related elements (e.g. periods or data elements) which you want to visualize in order to emphasize trends or relations in its data.
2. Categories: A category is a set of elements (e.g. indicators or organisation units) for which you want to compare its data.
3. Filter: Since most charts are two-dimensional, a filter must be used on the third dimension in order to use only a single element for the chart to become meaningful.

11.4. Selecting indicators and data elements

The visualizer module can display any number of indicators and data elements in a chart and data table. Both indicators and data elements can be selected and appear together in the same chart. You can select indicators by clicking at the "Indicators" header and selecting an indicator group from the list of groups below it. This will make the indicators in the selected group appear in the list under "Available indicators" to the left. From that list you can double click on any indicator in order to select it, this will move it to the list under "Selected indicators". Alternatively you can mark one or more indicators and click the single-arrow button. To select all indicators you simply click on the double-arrow button. To deselect indicators you can do correspondingly in the "Selected indicators" list.

To select data elements click on the "Data elements" header. The same principle for selecting and deselecting applies as for indicators.

11.5. Selecting reporting rates

The visualizer can display reporting rates in a chart, by itself or together with indicators and data elements. Reporting rates can be selected by clicking at the "Reporting rates" header. Reporting rates are defined by data sets. It can be selected by double-clicking in the list of available data sets to the left.

11.6. Selecting fixed and relative periods

Click on the "Periods" header. For fixed periods, select a period type from the combo box. You can select any number of fixed periods from any period type. Below the fixed period you will find the relative period checkboxes and you may select as many as you like. The names should be fairly self-descriptive and they are relative to the current date, meaning that if the current month is March and you select "Last month", the month of February will be included in the chart. You are also free to combine fixed periods and relative periods in the same chart. Overlapping periods will be filtered so that they only appear once.

11.7. Selecting organisation units

You can select which organisation units to include in the chart by clicking the "Organisation units" header. This section features three ways of selecting organisation units. The default mode is called "Organisation units" and lets you select the organisation units you want to appear in the chart from the tree. This mode also features three checkboxes. Checking "User org unit" will disable the organisation unit tree and give you the organisation unit that is related to the current/logged in user instead. This is also useful for administrators as they can create a meaningful "system" favorite with this option checked and all users will find their respective organisation unit when they open it. The the same concept goes for "Org unit children" and "Org unit grand children". The second mode is called "Select levels". Here you can select all organisation units at one or more levels. However, at the same time you also have the option to select parent organisation units in the tree, which makes it easy to select e.g. all facilities inside one or more districts. The same thing goes for the third mode called "Select groups". Here you can select all organisation units inside one or more groups and parent organisation units at the same time.

11.8. Selecting organisation unit group sets and data element group sets

All dimension tabs listed below "Organisation units" are organisation unit group sets and data element group sets. You are free to add groups from any of these group sets to your chart. Remember to add the group set in either the series, category or filters combobox.

11.9. Selecting chart options

You can set various chart options by clicking on the "Options" button the chart toolbar.

- Show values: Shows the values above the series in the chart.
- Hide empty category items: Hides category items with no data from the chart.
- Show trend lines: The trend line will visualize how your data evolves over time - e.g. is performance improving or deteriorating. Makes sense when periods are selected as category.
- Target line value/title: Displays a horizontal line at the given domain value. Useful e.g. when you want to compare your performance to the current target.
- Base line value/title: Displays a horizontal line at the given domain value. Useful e.g. when you want to visualize how your performance has evolved since the beginning of a process.
- Range axis max/min: Defines the maximum and minium value which will be visible on the range axis.
- Range axis tick steps: Defines the number of ticks which will be visible on the range axis.
- Range axis decimals: Defines the number of decimals which will be used for range axis values.
- Range axis title: Displays a label next to the range axis (also referred to as the Y axis). Can give context information to the chart, e.g. the unit of measure being used.
- Domain axis title: Displays a label below the domain axis (also referred to as the X axis). Can give context information to the chart, e.g. the type of periods being listed.
- Hide chart legend: Hides the legend and leaves more room for the chart itself.
- Hide chart title: Hides the title of your chart.

- Chart title: Type any title you like and it will appear above the chart.

11.10. Displaying a chart

You can display a chart based on your selections simply by clicking the "Update" button on the top centre menu. This requires that you have selected one or more elements from all of the three dimensions - data (indicators, data elements, reporting rates), periods (relative, fixed) and organisation units (units or groups). Note that "Months this year" from the period dimension and the root organisation unit are selected by default.

Notice that you can hide and show individual data series in the chart by clicking directly on the series label in the chart - they appear either at the top or to the right of the chart.

If you want to give the chart more space on your screen you can click on the triple left-arrow button on the top centre menu. This will collapse the left side menu. You can get this menu back by clicking on the same button again.

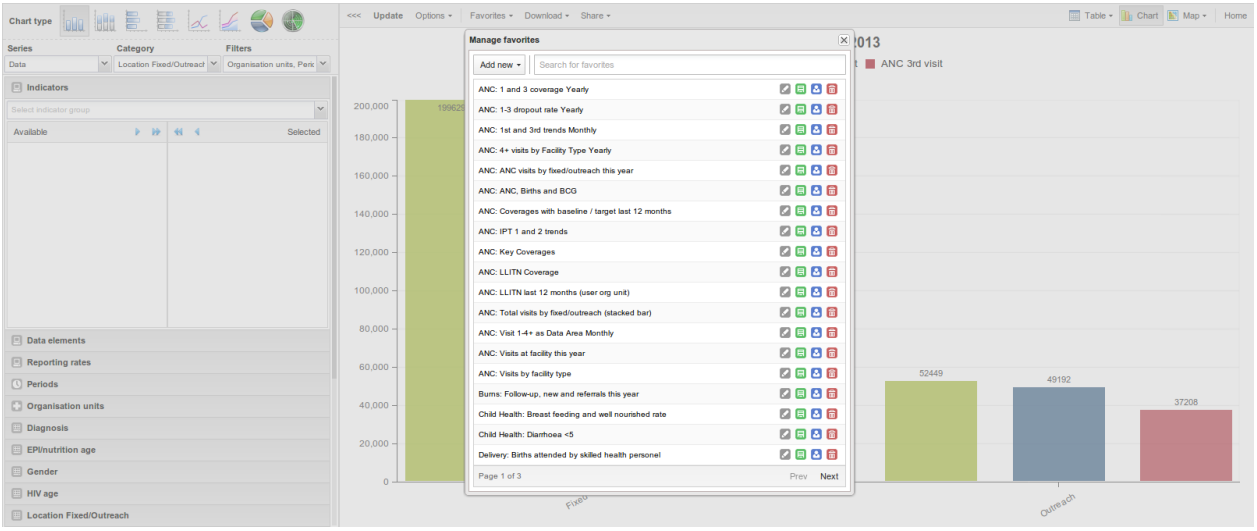
11.11. Downloading chart as image or PDF

After you have rendered a chart you can download it to your local computer as and image or pdf by clicking on "Download" on the top centre menu. The file will be automatically downloaded to your computer - for instance can you now embed the image file into a text document as part of a report. You can also download the data source behind the chart in json, xml, csv or Microsoft Excel format.

11.12. Saving chart as favorite

Once you have rendered a chart you can save it as a favorite to able to access it easily at a later point. Click on the "Favorites" button on the top menu to open up the favorites window. Click "Add new" and in the name field enter the desired name for your chart. The chart will be visible only to yourself. For every favorite in the list you have four options to the right. You can rename the chart (grey button), overwrite the chart (green button), modify the sharing settings of the chart (blue button) or delete the chart (red button).

These favorite charts can later be included on your personal dashboard. After saving you can navigate to the dashboard module, click on the "Insert" link over the chart areas and select your preferred chart.



11.13. Sharing interpretations

For certain analysis-related resources in DHIS, like pivot tables, charts and maps, one can share a data interpretation. An interpretation is simply a link to the relevant resource together with a text expressing some insight about the data.

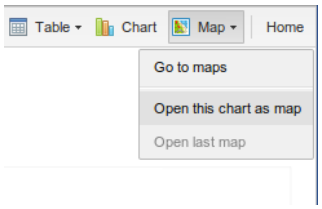
If you want to share a chart interpretation you need to first save the chart you want to share as a favorite. Then, without making any changes to the chart, click the "Share" button the toolbar. A window will open up and this is where you write your interpretation. When you are done, click share button in the bottom right corner of the window. The window will close automatically and if the interpretation was shared successfully you will find a notification on the bottom toolbar.

11.14. Embed charts in any web page

Certain analysis-related resources in DHIS, like pivot tables, charts and maps, can be embeded in any web page by using a plugin. If you have created a chart in the Data Visualizer you will get the plugin configuration for this chart by clicking the "Share" button the toolbar and then "Embed as plugin". You will find more information about the plugins in the web api chapter.

11.15. Analysis integration

The analysis apps in DHIS 2 are completely integrated, so you can easily switch between pivot table, chart and map visualization of your data. When you have made a chart you can click e.g. "Map" in the top right corner and then select "Open this table as map".



11.16. Exiting the data visualizer module

If you want to exit the module and go back to the DHIS start page you can click on the "Home" button to the right side of the top centre menu.

Chapter 12. Dashboards

Dashboards are intended to provide quick access to individual users to the data which has been stored in DHIS2. Dashboards consist of several sections, some of which provide links to reports or mapview which have already been defined. Other sections of the dashboard allow users to add charts which have been defined and made available through the charting module.

12.1. Setting up the dashboard

The dashboard is divided into two main sections. The right-side pane (denoted as A in the screenshot below) can be used to contain links to reports, documents (static reports), report tables, map views, and an RSS Health feed. The left-side zone (denoted as B in the screen shot) can be used to contain up to six separate charts which have been previously created in the charting reporting module.

In this screen shot, the dashboard has already been populated with a number of reports and mapviews. Simply clicking on one of the blue links will bring you automatically to the report or map view. Clicking on one of the charts will display a larger chart, which you can save as an image, and include in a report or other document.

You can redefine the structure of the dashboard by clearing the each of the windows by clicking the "Clear" link. By clicking "Insert" again, you can then select a new chart to appear in the window.

All reports, documents, report tables and charts can be added to the list of available options by clicking on the "Add to dashboard" icon  in each of the respective modules. Please refer to the sections to each of the sections in this manual for more detailed instructions. Once you have added the object to the list of available objects, you can"Insert" it on the dashboard.

Note

Dashboard are configured for each individual user.

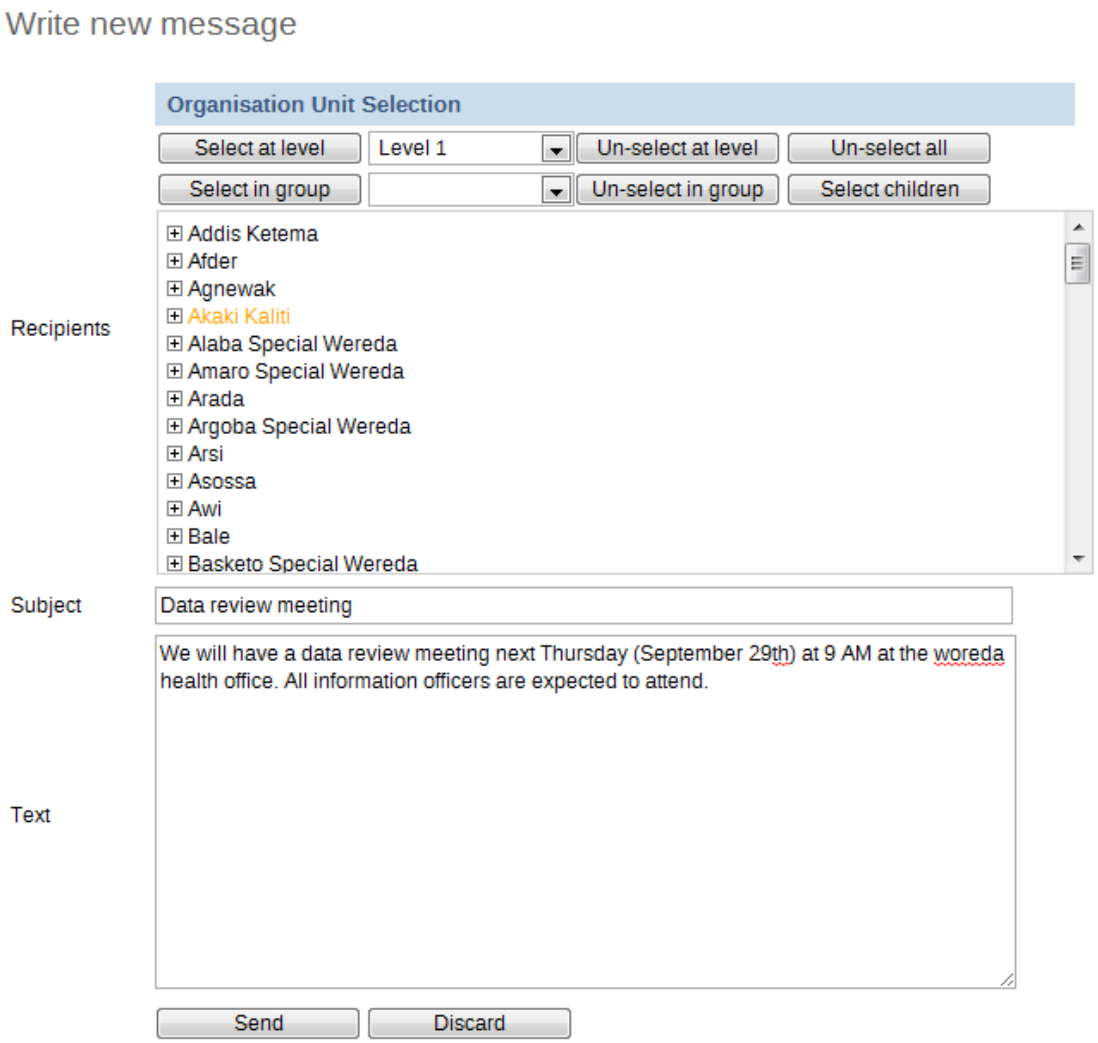
12.2. Messages and feedback

DHIS2 has certain functions to facilitate communication between different users and user groups. This type of communication is important to facilitate feedback regarding data quality, timeliness of submissions, or to simply answer a question which a particular user may have.

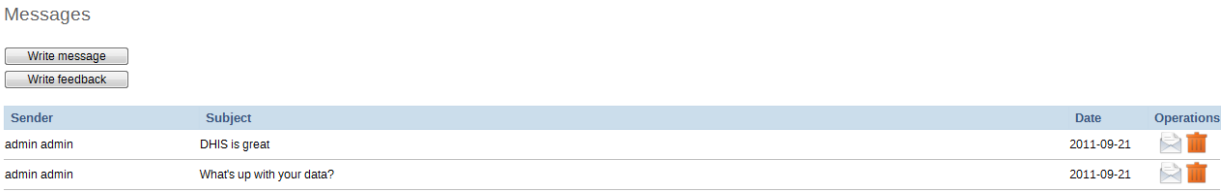
Feedback messages are sent to a particular group of users and can be sent by all users who have access to the dashboard module. In order to enable the receipt of feedback messages sent from the dashboard, you must set the system setting "Feedback recipients" which is available from the Maintenance->System settings dialog. Be sure to define a user group (e.g. "Feedback recipients") with all of the users who should receive feedback messages. Refer to the section in this manual on "User groups" for more information of how to do this. Once the "Feedback recipients" user group has been defined, each time a feedback message is sent, it will appear as a message in each of the "Feedback recipients" message queue within DHIS2. Note that messages will not be sent to users email addresses, but will only appear within the DHIS2 application.

To send a new feedback message, simply select "Write feedback" from the dashboard. Provide a subject and text in the respective text boxes. The message will appear in all of the specified users message queue.

Messages can be sent to specific groups of users who have been assigned to particular organisation units. To write a new message, simply click "Messages" from the dashboard screen and then press the "Write message" button.. Select an organisation unit (or group of organisation units) from the "Recipients" organisational unit tree. Provide a Subject and Text. To send the message, press the "Send" button. You can discard the message by pressing the "Discard" button as seen in the screenshot below.



To read messages which have been sent to you, select "Messages" from the "Dashboard" . You messages will be displayed as a list. Click the desired message to read all of the messages in this particular conversation.


Messages
Write message
Write feedback
Sender
Subject
Date
Operations
admin admin
DHIS is great
2011-09-21
admin admin
What's up with your data?
2011-09-21

Part Volume III: Administrator Guide. Volume III: Administrator Guide

Chapter 13. Introduction to the Administrator's Guide

13.1. Welcome!

This guide is for you if you are going to be responsible for maintaining the Dean's Dashboard. Your responsibilities include one or more of the following:

- Backing up data
- Restarting and upgrading the server
- Creating user accounts
- Linking data with other systems

Chapter 14. Organisation units

In this section you will learn how to:

- Create a new organisation unit and build up the organisation unit hierarchy
- Create organisation unit groups, group sets, and assigning organisation units to them
- How to make changes to the organisational unit hierarchy

14.1. The organisational hierarchy

The organisational hierarchy defines the organisation structure of the DHIS2 instance, such as how the health facilities, administrative areas and other geographical areas are arranged with respect to each other. It is essentially the "where" dimension of DHIS2, similar to how periods represent the "when" or time dimension. DHIS2 is structured so that the organisational unit hierarchy is a geographical hierarchy, and the GIS module depends on this. Non-geographical hierarchies are discouraged, and would better to be represented through the use of organisational unit groups. This dimension to the data is defined as a hierarchy with one root unit (e.g. Ministry of Health or a country) and any number of levels and nodes below. Each node in this hierarchy is called an organisational unit in DHIS2.

The design of this hierarchy will determine the geographical units of analysis available to the users as data is collected and aggregated in this structure. There can only be one organisational hierarchy at the same time so its structure needs careful consideration.

Additional hierarchies (e.g. parallel administrative boundaries to the health care sector) can be modeled using organisational groups and group sets, but the organisational hierarchy is the main vehicle for data aggregation on the geographical dimension. Typically national organisational hierarchies in public health have 4-6 levels, but any number of levels is supported.

The hierarchy is built up of parent-child relations. For instance a country might have eight provinces, and each province again might have a number of districts as their children. Normally the health facilities (from which data is typically collected) will be located at the lowest level, but they can also be located at higher levels, e.g. national or provincial hospitals, so skewed organisational trees are supported (e.g. a leaf node can be positioned at level 2 while most other leaf nodes are at level 5).

Note that it is quite easy to make changes to the upper levels of the hierarchy at a later stage, the only problem is changing organisational units that collect data (the leaf nodes), e.g. splitting or merging health facilities. Aggregation up the hierarchy is done based on the current hierarchy at any time and will always reflect the most recent changes to the organisational structure.

Important

Because the most recent information which is contained in the organisational unit hierarchy is always used for aggregation, it is important to keep in mind that changes to it (such as the division of districts into two separate districts) will not be respected over time. As an example, District A may be sub-divided into District B and District C. This is a process which often happens for political reasons. Facilities which belong to District A would need to be reassigned to District B and C. However, any historical data, which was entered before the split actually occurred, would still be registered as belonging to District B and C and not the defunct District A. This temporal representation of the organisational hierarchy across time will be lost.

14.2. Organisation unit maintenance

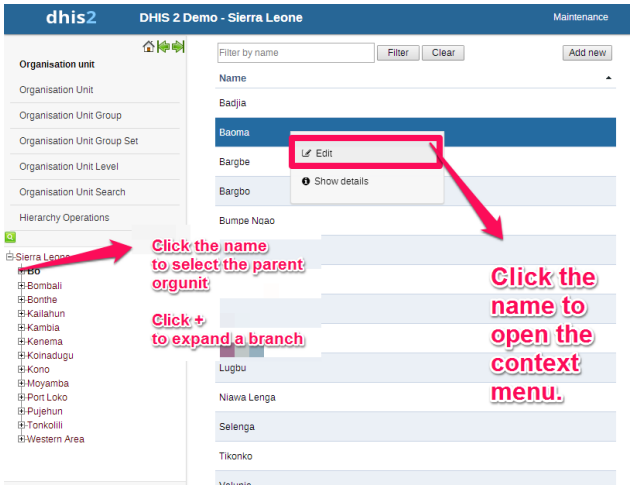
14.2.1. Organisation units

This is where you can create organisation units (from now on referred to as orgunits) and build up the orgunit hierarchy. Orgunits are added one by one as either root unit or a child of a selected unit. The left side menu represents the current

organisational hierarchy and if you select a unit there you will see its children listed in the main list of orgunits in the middle of the screen. When an orgunit is selected in the left side menu you can also add new child units to it. To locate an orgunit in the hierarchy you can either navigate through the tree by expanding the branches (click on the + symbol), or search for it by opening the search field (click the green symbol above the root of the hierarchy). In search you can either search for the orgunit name or its code, both will only show exact matches (case-insensitive). To add a new orgunit first select its parent and then click on the Add new button in the top right corner of the list of orgunits. To add a new root orgunit make sure no orgunit is selected in the menu and click on "Add new". The details of adding a new orgunit are explained in Section 14.2.1.1, “Editing organisation units”.

14.2.1.1. Editing organisation units

To edit the properties of an existing orgunit first select its parent (if any) in the left side menu, then locate the orgunit in the listed orgunits, and click on the name of the orgunit that you want to modify. A context menu will appear, and you should select "Edit". Refer to the screen-shot below to see how it works.



The following properties can be defined in the Edit (or Create new) window:

- Name: Define the precise name of the orgunit in this field. Each orgunit must have a unique name.
- Short name: Typically, an abbreviation of the full name. This attribute is often used in reports to display the name of the orgunit, where there is limited space available.
- Code: In many countries, orgunits are assigned a code. This code can be entered in this field.
- Description: A description can be a longer piece of text which can be used to describe the organisationunit.
- Opening date: Used to control which orgunits that where existing at a point in time, e.g. when analysing historical data. This attribute is required. The default date for opening of organisation units is 1900-01-01, but can be set to any date (even dates which occur in the future).
- Registers data: This property is used to identify which orgunits that can register data or not. Sometimes administrative orgunits at higher levels in the hierarchy are not supposed to register any data. This can help control the data entry process as only orgunits with this property set to Yes will be available for data entry.
- Comment: Any additional information that you would like to add can be put here.
- Coordinates: This field is used to create the maps in the GIS module. Paste in the coordinates of the orgunit in this field, either a polygon (for orgunits that represent an administrative boundary) or a point (for health facilities). Without this information the GIS module will not work. It might be more efficient to import these coordinates later as a batch job for all orgunits using the import module. See the GIS chapter for more details.
- URL: You can use this field to insert a URL link to an external web site that has additional information about this specific orgunit.
- Contact information: A contact person, address, email, and phone number can be entered in these fields. This information can be vital for facilitating follow-up.
- Datasets: Datasets can be assigned to organisational units here. See the chapter on "Data sets" for more detailed information on assigning datasets to organisational units.
- Organisation unit groups: Assignments to organisational units group sets can be assigned through the individual drop-down boxes which appear for each group set.

In addition to all of the options listed above, if you have added any attributes to organisation units, your custom attributes may also appear there. Please refer to the section on "Attributes" for more information about how attributes can be used.

14.2.2. Organisation unit group sets

Group sets can be understood as a flexible tool to add more categorisation to orgunits. Any number of group sets can be added, but as a default start all databases will have the two group sets "Type" and "Ownership". Using these group sets will simplify how reporting is done, and facilitate analysis through the use of tools such as Excel PivotTables.

While a group set like "Type" describes a measure dimension, the actual categories are represented by the groups, and the categorisation of an orgunit through the orgunit's group memberships. This can be understood as a parallel hierarchy of orgunits with the group set as the root ("Type"), the groups at level 2 (e.g. "Clinic", "Hospital", "Dispensary"), and the actual orgunits at level 3. The group set can as such provide additional information and dimensionality to the data analysis as data is easily filtered, organised, or aggregated by groups within a group set.

For this aggregation to work without any duplication in the data some rules are necessary. A group set is always exclusive, which means that an orgunit cannot be member of more than one group in a group set. Therefore, when creating a new organisational unit, you will only be allowed to select a single organisational group membership for each group set. Furthermore it is possible to define whether a group set is compulsory or not, which will affect the completeness of the data when analysing data using group sets. Compulsory means that ALL orgunits must be member of a group in that group set.

We recommend that you approach the orgunit grouping in the following sequence (and one group set at a time):

1. Define a new group set, such as "Location".
2. Add new groups (such as "Urban", "Rural" and "Peri-urban"). Once all groups have been defined, return to the organisational unit group set and assign each of the desired groups to the group set.
3. Go back to each group, one by one, go to edit mode and assign the orgunits that should be member of the group. Should you follow this route, you can place multiple organisation units at a time in a group. However, you must be careful not to place the same organisational units in two groups which itself is a member of an organisation unit group set. This will result in a data integrity violation. If you have organisation unit groups which are not exclusive, they should not be members of a group.
4. A better way to ensure that you do not mistakenly assign an organisation unit to multiple members of a group set is you can use the edit feature of each organisational unit to assign memberships to each group set. You will only be able to assign a single organisation unit at a time however.

It is important to keep in mind when using the "Organisational unit group" set function, that unless great care is taken, organisational units can be assigned to multiple groups of a group set. This can be checked through the "Data Integrity" module, which will report which organisational units are not members of a compulsory organisational unit group set, and which organisational units have been assigned to more than one member of a group set.

14.2.2.1. Editing organisation unit group sets

Click on the name of the organisation unit group set you wish to modify, followed by "Edit" from the context menu which will appear. The following properties can be defined in the Edit (or Create new) window:

- Name: Provide a precise name for the group set.
- Description: Describe the phenomena the group set is measuring/capturing.
- Compulsory: Indicate whether ALL orgunits need to be member of a group in this group set or not.
- Available groups/Selected groups: Here you assign groups to your group set by using the arrow buttons to move highlighted groups between the two lists (/selected). If no groups appear in the list then you must go to orgunit groups and create new groups there first. Note that assigning groups that will violate the exclusive rule on group sets is not possible, e.g. adding a group that already has assigned an orgunit that again is already member of a group that has already been selected by this group set, will not be possible since one orgunit will end up with two group memberships in the same group set. To avoid such situations we recommend first adding groups to group sets, and then orgunits to groups.

14.2.3. Organisation unit groups

This function will allow you to add new and manage existing organisation groups and their memberships. It can be accessed by choosing Maintenance->Organisation units->Organisation Unit group from the main menu. To add a new orgunit group click on the "Add new" button in the top right corner of the list of groups.

14.2.3.1. Editing organisation unit groups

Click on name of the orgunit group that you want to modify and then select "Edit" from the context menu which will appear. The following properties can be defined in the Edit (or Create new) window:

- Name: Provide a precise,unique and descriptive name for the orgunit group.
- Short name: This name should be less than 25 characters, and will be used in certain places in DHIS2 when the number of characters needs to be restricted due to space constraints.
- Symbol: Select a symbol which will be used to display the organisation unit (points only) when the layer is displayed in the GIS.
- Organisation unit tree selection: This is where you assign orgunits to the group. The tree supports multiple selection so select all the orgunits that you want to add (the selected ones appear with orange color) and click on "Save". Click on "Cancel" to undo your changes and return to the list of orgunit groups. Use the "Select at level" button and dropdown if you want to select all orgunits at a specific level in the hierarchy (e.g. all districts).
- Datasets: If you assign a dataset to an organisational unit group, all organisation units which are currently assigned to the dataset will be also present in this organisation unit group.

14.2.4. Organisation unit level

Here you specify a contextual name for each level in the hierarchy, e.g. "Country", "Province", "District", "Health Facility", and these names will be used all over the application where levels are referred to. This page will take some time to load if the orgunit hierarchy is very big.

14.2.5. Hierarchy operations

Here you can move orgunits around in the hierarchy by changing the parent of a selected orgunit. This process is done in three steps:

1. Select the orgunit you want to move (in the hierarchy in the left side menu) and click "Confirm" under the "Select an organisation unit to move" label.

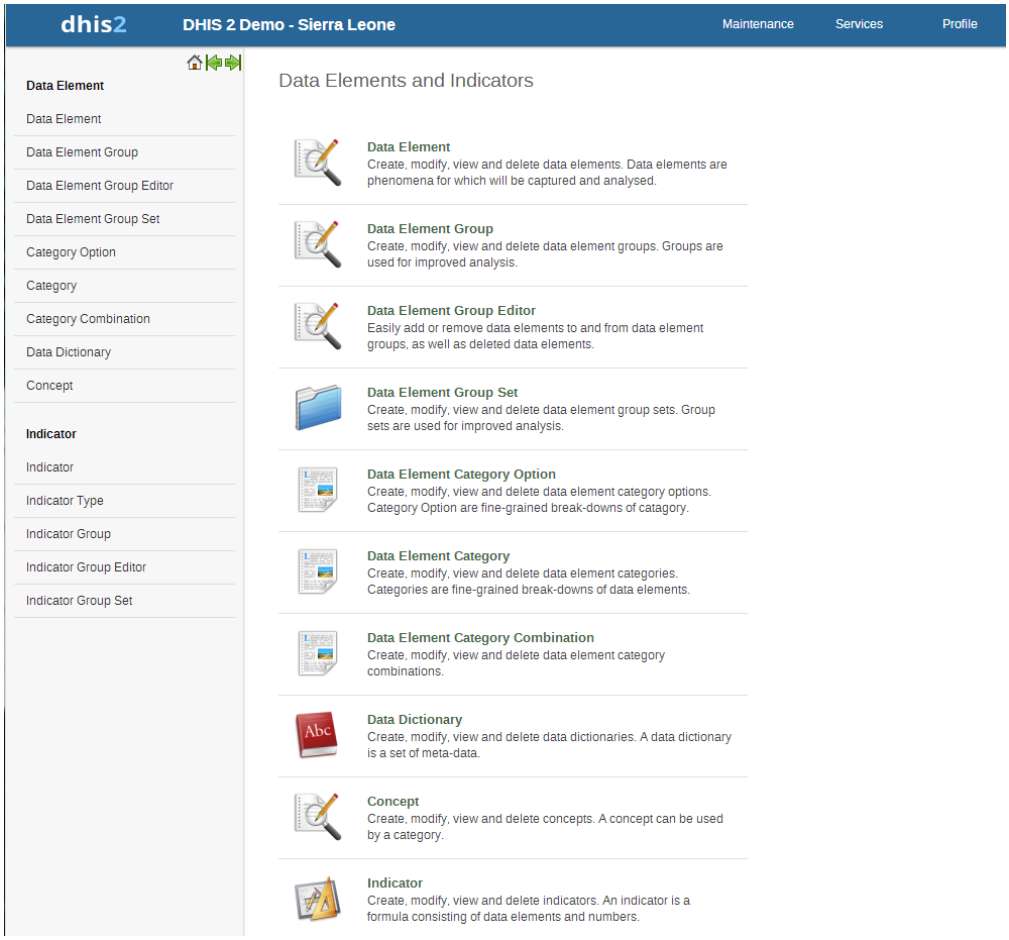
2. Select the new parent orgunit (again by using the hierarchy in the left side menu). If no parent is selected then the orgunit will be moved up to root level (top of the hierarchy). Click on the "Confirm" button under the "Select the new parent organisation unit for the one to move" label.

3. Click on the "Move" button to apply your changes to the hierarchy.

Your changes will be immediately reflected in the left side menu hierarchy. At any time in the process (before hitting the Move button) you can click on the "Reset" button to deselect orgunit to move and the new parent.

Chapter 15. Data elements

When the ‘Data Elements and Indicators’ options is chosen from the main Maintenance menu, the following screen appears:



From the left side menu or by clicking on the sections listed in the central area you can access the various sections on data elements and indicators.

15.1. Data element maintenance

Each of the options for maintenance of data elements will be described in the following section.

- Data element

  Create, modify, view and delete data elements.

- Data element group

  Create, modify, view and delete data element groups.

- Data element group editor

  Easily add or remove data elements to and from data element groups.

- Data element group set editor

  Create, modify, view and delete data elements group sets.

- Data element category options, categories and category combinations

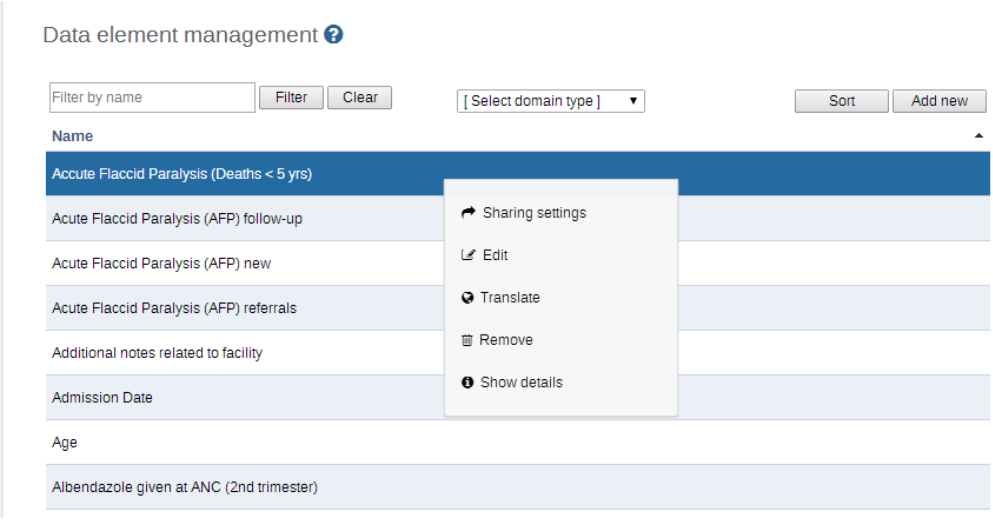
  Create, modify, view and delete data element categories.

15.1.1. Data elements

Data elements form the basis of DHIS2. Data elements define what is actually recorded in system, e.g. number of immunisations or number of cases of malaria. The actual creation and definition of the data elements themselves are far beyond the scope of this manual to describe, but it is assumed that an administrator will be provided with a list of standardised data elements for inclusion into the DHIS2 system.

To access the data element maintenance module, choose Maintenance -> Data elements and Indicators -> Data element.

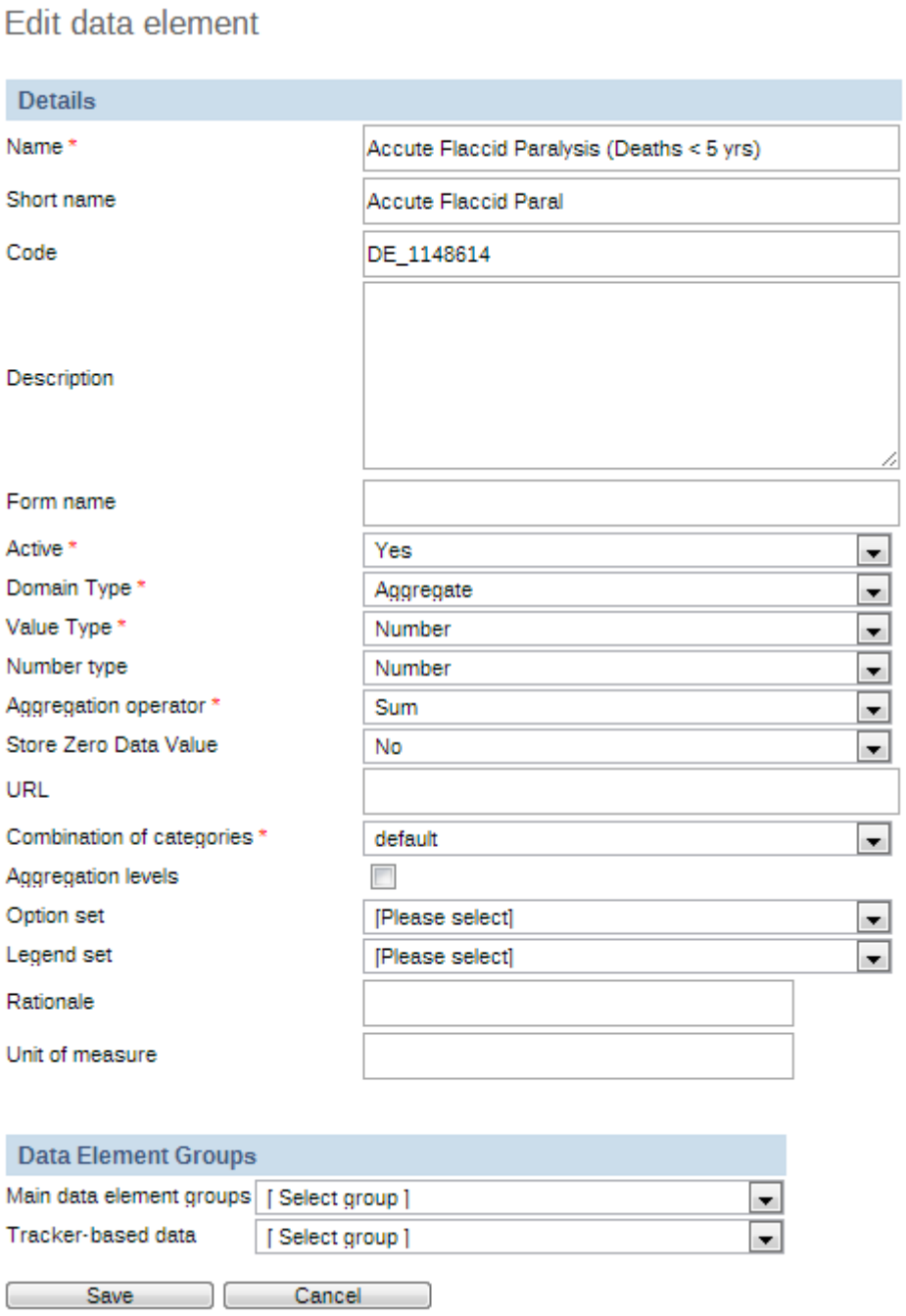
The 'Filter by name' will allow you to filter a range of data elements if you know either the full name of the data element, or just a part of it. Type the name into the search field and any matching data elements are displayed below. The 'Sort' button can be used to sort the data elements into alphabetical order.



To add a new data element, click the 'Add new' button. There are various options available from this page that allow the user to modify data elements already present in the database. Each of the options are described below in the "Editing data elements".

15.1.1.1. Editing data elements

To edit an existing data element, click the name of data element you wish to modify, and then select "Edit" from the context menu which will appear.



- **Name:** Define the precise name of the data element in this field. Each data element must have a unique name.
- **Short name:** Typically, an abbreviation of the full data element name. This attribute is often used in reports to display the name of the data element, where there is limited space available.
- **Code:** In many countries, data elements are assigned a code. This code can be entered in this field.
- **Description:** Allows a full textual description of the data element to be entered. The user should be as precise as possible, and include full information on how the data element is measured and what its meaning is.
- **Active:** Defines whether a given data element is active or not. Data elements marked as inactive, will not be displayed in the data entry screens.
- **Domain type:** Defines whether a data element is an aggregate or patient type of data element.
- **Value type:** Defines the type of data this data element will be used to record. Currently there are five options:
  - Number: Numeric values.
  - Text: Textual values.
  - Yes/No: Boolean values, will render as drop-down lists in data entry.
  - Yes only: True values, will render as check-boxes in data entry.

  - Date: Dates, will render as calendar widget in data entry.
- **Number type:** DHIS 2 supports several different number types. During data entry, users will be restricted to enter the defined number types only. Each of the available options are described below.
  - Number: This number type supports any real value with a single decimal point, an optional negative sign, and no thousands separators.
  - Integer: Any whole number (positive and negative), including zero.
  - Positive integer: Any whole number greater than (but not including) zero.
  - Negative integer: Any whole number less than (but not including) zero.
  - Positive or zero integer: Any whole number greater than or equal to zero.
  - Unit interval: Continuous number between 0 and 1.
  - Percentage: Whole number inclusive between 0 and 100.
- **Text type:** The detailed type relevant to text value type.
  - Text: Free text, rendered as standard input field.
  - Long text: Free text, rendered as text area in data entry.
- **Aggregation operator:** Defines the default aggregation operation that will be used on this data element. Most data elements should have the "SUM" option set. This includes all data elements which should be added together. Other data elements, such as staffing levels, should be set to use the "AVERAGE" operator, when values along the time dimension should not be added together, but rather averaged. The complete list of aggregation operators:
  - Sum: Sum of data values in the period and organisation unit dimension
  - Average: Average of data values in the period dimension, sum in the organisation unit dimensions.
  - Count: Count of data values.
  - Std dev: Standard deviation (population-based) of data values.
  - Variance: Variance (population-based) of data values.
  - Min: Minimum of data values.
  - Max: Maximum of data values.
- **Store Zero Data Value**: By default, DHIS2 will not store zeros which are entered in the data entry module. If zeros need to be saved for a particular reason, this option can be set to "Yes".
- **URL:** A URL having an in-depth description of the data element can be entered in the ‘URL’ field. This could be for instance, a link to a metadata repository or registry that contains detailed technical information about the definition and measurement of the data element.
- **Combination of categories:** Defines which category combination the data element should have.
- **Aggregation levels:** The Aggregation Levels option allows the data element to be aggregated at one or more levels. When the user clicks on the Aggregation levels option, a drop down menu appears which displays available aggregation levels. The desired aggregation level is then selected by clicking the ‘Add Selected’ button. By default, the aggregation will start at the lowest assigned organisation unit. If e.g. Chiefdom is selected below it means that Chiefdom, District, and National aggregates will use Chiefdom (the highest aggregation level available) as the data source, and PHU data will not be included. PHU data will still be available for the PHU level, but not included in aggregations to the levels above. If District and Chiefdom are both selected then the District and National level aggregates will use District data as their source, Chiefdom will use Chiefdom, and PHU will use PHU. Read more about aggregation levels in the Reporting chapter i the section on data sources for reporting.
- **Option set for data values:** Option sets are predefined lists of options which can be used in data entry.
- **Option set for comments**: Option sets for comments are predefined list of options which can be used to specify standardized comments for data values in data entry.
- **Legend set**: Legend sets can be used in the GIS module to display certain data elements with certain icons. Refer to the GIS module documentation for more information on legend sets.
- **Attributes**: Data element attributes (if they have been defined) can be defined. In this example, "Rationale" and "Unit of measure" are both data element attributes.
- **Data element group sets:** If data element group sets have been defined, each will appear in the "Data element groups" section. Select each data element group from the list of group sets provided.

After making all the required changes, click ‘Save’. The ‘Cancel’ button aborts all changes made.

15.1.1.2. Deleting a data element

In order to delete a data element, click the name of the data element you wish to delete, and then select "Remove" from the context menu. Note that this operation is only possible if there is no data attached to the data element itself. The user will be prompted to ensure that the data element should be deleted.

15.1.1.3. Displaying data element details

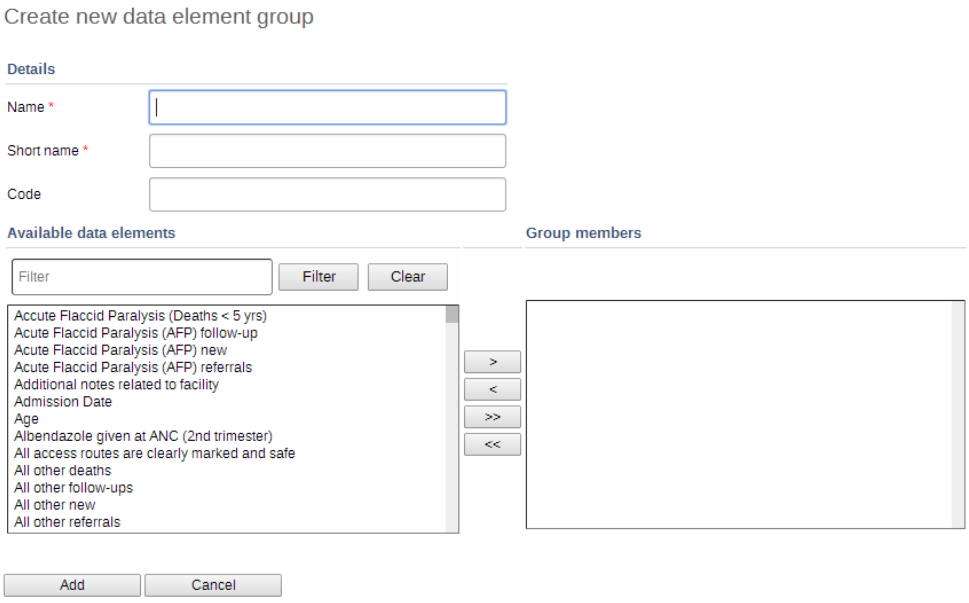
This operation displays an in-line panel in the browser which displays all metadata about a given data element. Click the name of the data element and then select "Show details" from the context menu.

15.1.2. Data element groups

Data element groups provide a mechanism for classifying related data elements into a common theme. For instance, two data elements "Measles immunisation" and "BCG Immunisation" might be grouped together into a data element group "Childhood immunisation". To access the data element group maintenance page, click Maintenance -> Data elements and Indicators -> Data Element Group.

Similar to the "Data element" maintenance page, data elements groups can be searched with by entering a search string in the "Filter by name" field.

To add a new data element group, click the "Add new" button and the following screen will be displayed:



Fill in the "Name" field and then select all data elements that should belong to the group from the left panel. Click the "Move selected"  button to add the selected data elements to the data element group. Click the "Remove selected" button to remove all data elements from the group that have been selected in the right panel. Finally, click the "Add" button to save changes, or the "Cancel" button to discard any changes.

15.1.3. Data element group editor

The data element group editor provides advanced functionality to the administrator to allow multiple data elements to be added or removed from a group. It is also possible to create new data element groups, rename existing groups, and delete groups entirely. To access the data element group editor, go to "Maintenance -> Data elements and Indicators -> Data Element Group Editor". The following screen will appear.

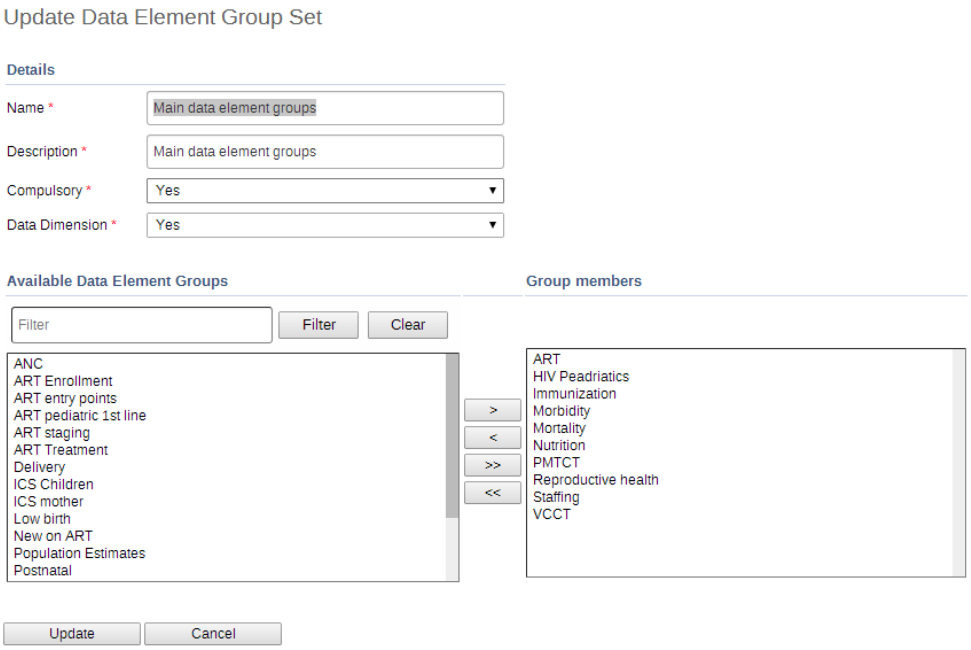
Data element groups area listed alphabetically in the leftmost panel. By clicking on a data element group, the current members of that group (data elements) are listed in the centre panel. Available data elements that can be added to the

data element group appear are listed alphabetically in the rightmost panel. To remove an existing data element from the group, click the name of the data element in the centre panel, and then press the "Move right" > button. To add data elements to the group, select them from the leftmost panel, and click the "Move left" < button. Press the "Update data element group member" button to save your changes.

15.1.4. Data element group sets

Data element group sets allow multiple data element groups to be categorised into a set. Data element group sets are used during analysis and reporting to combine similar data element groups into a common theme. To access the data element group set maintenance module, choose "Maintenance -> Data elements and Indicators -> Data Element Group Set". Similar to the other data element maintenance modules, new data element group sets can be added by pressing the "Add new button". Other operations include Edit, Translate, Delete and Information, similar to data elements and data element groups as described in the previous sections.

Existing data element group set members can be edited by clicking the name and selecting "Edit" from the context menu of the desired data element group set as seen below.



Available data element groups are displayed in the left panel. They can be moved into the selected data element group set by pressing the "Move right" > button. Data element groups that are currently members of the data element group set are displayed in the right hand panel. They can be removed from the data element group set by clicking the desired data element group and pressing the "Move left" < button. The ordering of the data element groups can be set with the "Move Up" and "Move Down" arrows. This ordering will be used in the datamart and reports to order the data element groups. Press the "Update" button to save any changes and the "Cancel" button to discard all changes.

15.1.5. Categories

Categories can be used to disaggregate data elements into individual atomic components. They can also be used to assign attributes to all data recorded in a specific dataset, such as "Implementing partner" and "Funding agency."

Data element categories are typically a concept, such as Gender, Age or Disease Status. Data elements such as "Number of cases of confirmed malaria" are often broken into smaller component parts to determine, for instance, the number

of confirmed malaria cases of particular age groups. As an example, three data element categories: Under 1, 1-5 and Over 5 could be created. They could be assigned as categories to the data element, which would then create in the data entry screens, three separate fields for this data element namely:

- Number of confirmed malaria cases (Under 1)
- Number of confirmed malaria cases (1-5)
- Number of confirmed malaria cases (Over 5)

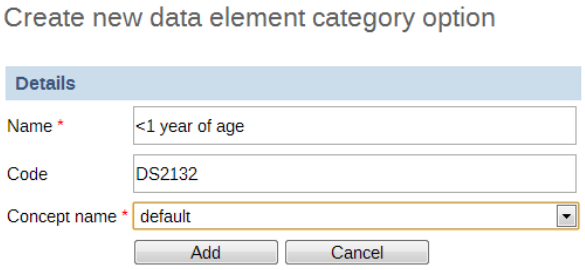
Effective use of data element categories greatly simplifies the process of setting up the DHIS2 system, as the data element categories can be reused to disaggregate many different data elements. Otherwise, each of the data elements listed above, would need to be created separately. Judicious use of data element categories will greatly simplify the DHIS2 implementation, and allow for subsequent advanced analysis.

Where possible, category options should be recycled. For instance, there might be two categories which might share a particular category option (e.g. <1 year of age). When creating the categories, this category option could be reused. This is important if particular category options (or category option combinations) need to be analyzed together.

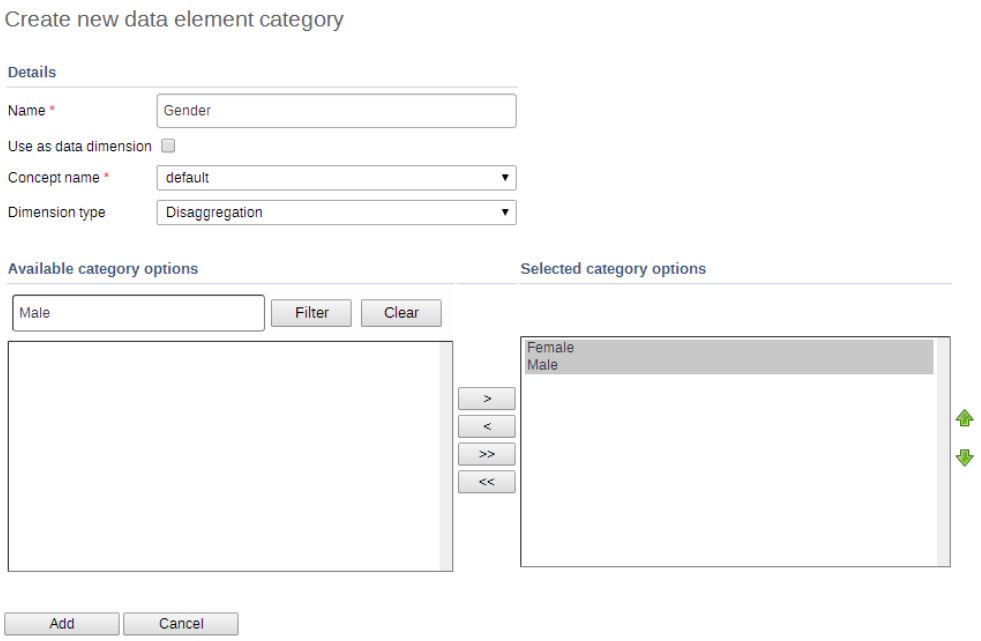
1. All possible category options should be defined.
2. Categories should be composed of multiple category options created in Step 1.
3. Category combinations should be composed of either one, or multiple categories.
4. Data elements should be created and assigned a particular category combination.

A category option consists of a name, along with an optional code, and a concept.

Categories can be added by accessing the "Data Element Category Option" dialog. (Maintenance -> Data Elements and Indicators->Data Element Category Options" as seen below. The category option must consist of a name, an optional code, and a concept.



Once all category options have been defined for a particular category, the category can be defined with the "Data Element Category" dialog.(Maintenance -> Data Elements and Indicators->Data Element Category). Enter the name of the data element category and press "Save". Once you have defined the name of the category, you can assign category options to it by selecting the category which you have just defined, clicking its name, and then selecting "Edit" from the context menu as seen below.



Type the name of the new data element category in the "Name" field in the "Details" region and assign a concept to the category. Category options can be added by moving category options from the left-side pane (Available category options) into the right-side pane(Selected category options). Category options can be reordered using the "Move Up" and "Move Down" buttons. . Once all data element categories options have been added to the data element category, press the "Add" button to save all changes or the "Cancel" button to discard any changes.

A category can have two types "Disaggregation" or "Attribute". For disaggregation of data elements, you should select "Disaggregation". A dimension type of "Attribute" will allow the category to be used to assign a combination of categories to data recorded through a dataset.

If the "Use as data dimension" box is ticked, the category will be available to the analytics as another dimension, in addition to the standard dimensions of "Period" and "Organisation unit".

15.1.6. Category combinations

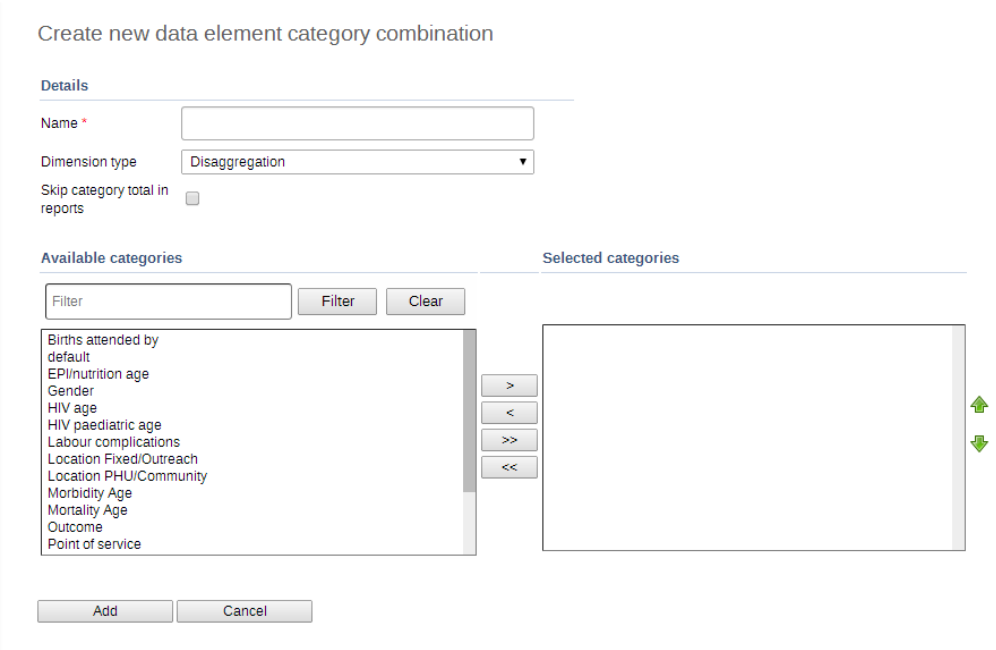
Category combinations allow multiple categories to be combined into a related set. As an example, a data element "Number of new HIV infections" might be disaggregated according to the following categories.

- Age: "Under 5", "5-15", "15-24", "24 and above"
- Gender: Male, Female

In this example, there would be two levels of disaggregation, consisting of two separate data element categories, each consisting of several data element category options. In most HMIS systems, different data elements are disaggregated according to a common set of categories. By combining these different categories into a category combination and assigning these combinations to data elements, the appropriate disaggregation levels can be applied efficiently and quickly to a large number of data elements.

To access the category combination maintenance module, select "Maintenance->Data element and indicators->Data element category combinations" from the main DHIS2 menu. As with the other maintains modules, you can filter the listed category combinations by entering the name (or portion thereof) of the category combination. Other operations such as "Edit", "Delete" and "Information" should be familiar to the reader.

To add a new category combination, click the "Add new" button. The following dialogue will be displayed.



Type the name of the category combination in the "Name" field, and then select the desired categories from the left panel. Press the "Move right" button to add the selected categories to the category combination. Press "Move left" to remove any categories that should not be part of the category combination.

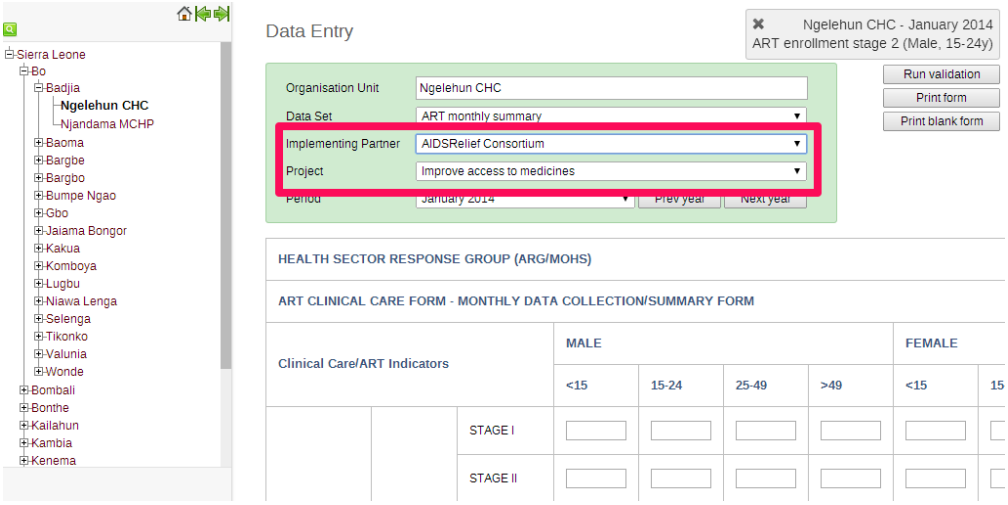
Categories can only be added to a category combination at this step. Categories can be removed from category combinations later by editing the category combination, however, it is not allowed to add additional categories once the combination has been created. Ensure that the category combination and its respective categories is final before you create the category combination and assign it to a data element.

15.1.7. Using categories as attributes

In some deployments of DHIS2, information on attributes such as "Implementing partner" and "Project" are important to record as an attribute of each data value. To provide an example, let us suppose that an NGO is providing ART services in a given facility. They would need to report each month on the "ART monthly summary", which would be reported monthly and contain a number of data elements. The NGO and project could potentially change over time. In order to be able to effectively attribute data to a given NGO and project at any point in time, this information would need to be recorded along with each data value at the time of data entry.

When categories and category combinations have a type of "Attribute", they can serve as a means of applying a common set of attributes (e.g. Implementing partner and project) to a related set of data values contained in a dataset. Let us suppose we create two categories called "Implementing partner" and "Projects". Each of these categories would be created with a dimension type of "Attribute." Finally, a category combination called "Implementing partners and projects" would be created with these two categories and assigned a dimension type of "Attribute". Finally, we can use this dimension to apply an attribute to the dataset "ART monthly summary", by choosing the "Implementing partners and projects" as the "Combination of categories".

The following screen shot of an example data entry screen illustrates all of these concepts.



Note that when data is entered, an "Implementing partner" and "Project" can be selected. Each data value recorded in this data entry screen, would be assigned a specific combination of these categories as an attribute. These attributes (when specified as a dimension) can be used in the analysis modules similar to other dimensions, such as the period and orgunit. In summary, when category combinations are used as attribute, they effectively serve as another dimension (similar to "Period" and "Organisationion unit") which can be used for analysis.

15.1.8. Category option groups and group sets

Category options can be groups and classified using category option groups. Category option groups can be created from "Category option group" > "Add new", and contains a set of category options.

Category option groups can be included in category option group sets. Category option group sets can be created from "Category option group set" > "Add new", and contains a set of category option groups.

The main purpose of the these entities is to be able to add more dimensionality to your captured data for analysis in apps such as pivot table and data visualizer. An example of where this becomes useful: In a system, data is collected by "projects", where projects is modeled as category options. The system is required to do data analysis according to which donor supports the project. In this case, a category option group set called "Donor" can be created. Each donor can be created as a category option group, where each category option / project is put in the appropriate group. In data analysis apps, the "Donor" group set will appear as a data dimension, while each donor appear as dimension items, ready for inclusion in reports.

15.1.9. Translations of data elements and other objects

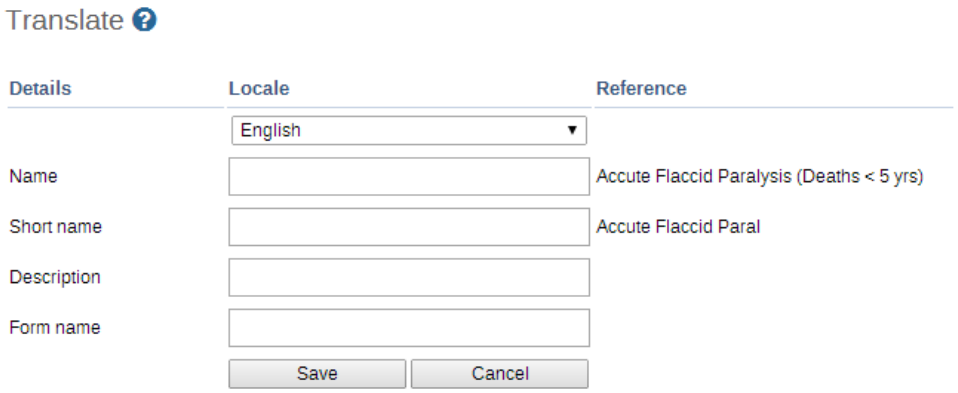
DHIS 2 provides functionality for translations of database content like data elements, data element groups, indicators, indicator groups, validation rules and more. These elements can be translated to any number of locales. A locale represents a specific geographical, political, or cultural region.

To add a translation click the Translate icon next to the element you would like to translate. Start by selecting the desired locale from the Locale select box. In the Translate screen, select your locale and enter values for the available element properties. The reference property values are shown on the right. These values are the values which have been entered in the regular add or update user interface for the current object.

Translations can be enabled by selecting the desired locale under Database Language under User General Settings in the Settings module.

15.1.9.1. Data element translation

DHIS2 provides functionality to translate existing data elements into other languages. Click the name of the data element you wish to translate, and select "Translate" from the context menu which will appear.. The following dialogue will be displayed.



The reference language is displayed in the upper right portion of the dialogue. Choose a locale to translate the data element into by selecting an option from the locale drop-down menu. Specify the name, short-name, description and form name in the target language. Press "Save" to save your changes.

Chapter 16. Data sets and data entry forms

16.1. Data sets

All data entry in DHIS2 is organised through the use of data sets. You can add and edit data sets in Maintenance->Data sets. A data set is a collection of data elements grouped together for data collection and data export between instances of DHIS2 (e.g. from a district office local installation to a national server).

A data set has a data collection frequency which can be set through the period type property. The frequency can be daily, weekly, monthly, quarterly, six-monthly, or yearly. Which data elements to include in the data set and the frequency are set in the Add/Edit Data set window. In order to use a data set to collect data for a specific orgunit you must assign the orgunit to the data set, and this mechanism controls which org units that can use which data sets.

Data sets also are assigned to specific organisation units which will be allowed to enter data for all data elements in that data set. You can assign org units to a data set in the Data set management by clicking on the blue folder icon, the first icon under Operations, next to the data set you would like to modify. Alternatively you can manage orgunit assignments for all data sets together in the Data set Assignment Editor (available in the right-side menu).

A data set has several properties that must be entered when creating a new one. Name, short name, code and description should be used to identify and describe the data set. The other properties deserve an explanation:

- Expiry days: Controls for how long it should be possible to enter data in data entry for this data set. Expiry days refer to the number of days after the end date of the selected data entry period where the data entry form should be open for entry. After the number of days has expired, the data set will be locked for further entry. You can set manual exceptions to this using the lock exception functionality in data administration module.

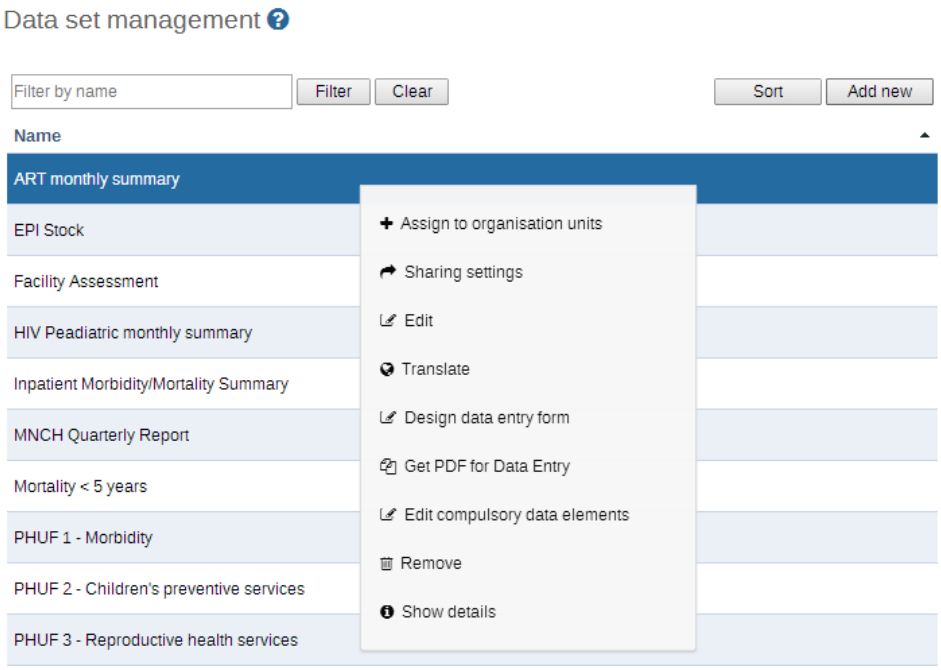
  Note

  If the number of expiry days is set to zero, this will allow data entry into all possible historical time periods.
- Complete notification recipients: Sets which users should receive a message with a notification about this data set being marked as complete in data entry. In this list you can select a user group, and all members in this group will receive a notification. The message will be delivered through the DHIS messaging system.
- Approve data: Define whether data for this data set should be Approved. (See the Data approval chapter.)
- Skip aggregation: Define whether data for this data set should be skipped during data mart generation. You should leave this on no, which is the default behavior, in most situations. Can be useful if you have limited server resources and are setting up new experimental data sets.
- Allow future periods: Defines whether it should be possible to enter data for future periods for this data set in data entry. The default behavior is to allow data entry only for periods which have passed, i.e. the end date is after today's date. If set to yes you can enter data for future periods, which is useful e.g. for population, target and planning data.'
- All fields for data elements required: Defines whether it is mandatory to fill all values for a data element in data entry if one or more values have been filled. This means that if the user enters one data value for a data element in an entry field (i.e. for a category option combination), then she must enter data for all fields belonging to that data element (i.e. all category option combinations).
- Complete allowed only if validation passes: Controls whether it should be possible to mark a data entry form as complete only if the validation of that form is successful. Default behavior is yes. If set to no, then a user cannot mark the form complete if validation fails.
- Skip Offline: Controls whether this data entry form should be downloaded and saved in the user's web browser. Normally you should leave this on no, which is the default behavior. If you have forms which are rarely used and are very big you can consider setting it to yes to speed up initial loading of the data entry module.

Your data set will then be ready to be used in Services->Data Entry for the org units that you have assigned and for periods according to your selected frequency (period type).

16.1.1. Data set management

The data set management function allows you to create new data sets and manage existing ones. The dialog can be reached by choosing Maintenance->Data sets->Data set. A sample dialog is displayed below.



Click on the name of the dataset you would like to alter, and a context menu will be displayed. Each of the available functions are described below.

- Sort: This controls the custom sort order. Depending on the systems settings, users will see the data sets ordered in the specific order which you provide.
- Add new: Adds a new data set. When pressing this button, you can create a new data set. You need to provide a name, short name and frequency. The "Code" attribute is optional. Data elements can be added to the "Selected data element" list by selecting them individually. and pressing the > button. Indicators can also be added to data sets and will be available to be placed in custom data entry forms when they need to be shown along with data elements on the same data entry form. Press "Save" to add the new data set.
- Assign organisation units to data sets: This function will allow you to assign individual organisational units to a data set. Only organisational units which have been assigned to a data set will be allowed to enter data into the data set.
- Sharing settings: Assign different rights to the dataset based on user groups and user roles.
- Edit data set: This will allow you to edit existing data sets, for instance when you need to add or remove data elements and indicators to a given data set.
- Translate: Allows you to translate the name of a data set to a different language.
- Design data entry form. Refer to the section on custom data entry forms for detailed information of how to use this function.
- Get PDF for Data Entry: Download a PDF file which can be used for offline data entry.
- Edit compulsory data elements: This dialog will allow you to add or remove data elements which will be marked as compulsory during data entry.
- Delete: Completely removes a data set from the system.

Warning

Any data set which is deleted from the system is irrevocably lost. All data entry forms, and section forms which may have been developed will also be removed. Ensure that you have made a backup of your database before deleting any data set in case you need to restore it at some point in time.

- Show details: Display some informative information about the data set, including the number of data elements, the frequency, and which data entry form has been assigned to the data set.

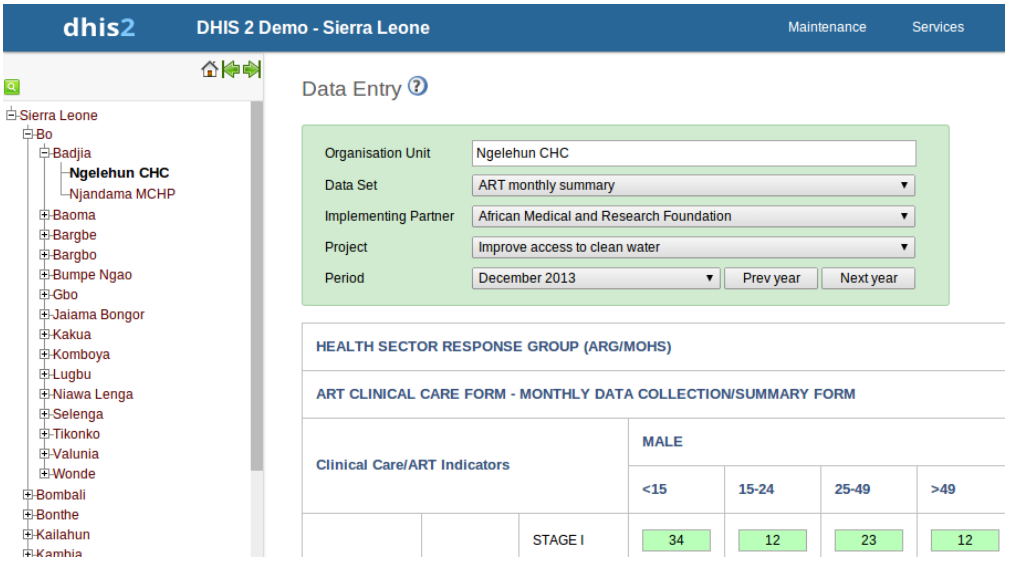
16.1.2. Data set categories

Before reading this section it is recommended to familiarize oneself with the sections on categories in the data element chapter. Whereas data element categories can be used for capturing disaggregations of data elements, data set categories are used to capture information which is common to an entire form.

To set up categories for data set, start by creating category options, categories and category combinations like described in the data element chapter. Make sure that you set the type of categories and category combinations to "Attribute". To assign a category combination to a data set, you can select it while creating or updating the data set from the "Combination of categories" drop-down box.

When a data set is linked to a category combination, those categories will be displayed as drop-down boxes in the data entry module. Data captured in the form will then be linked to the selected category options from those drop-down boxes.

An scenario for when data set categories are useful is when you need to capture a data entry form for a implementer partner organisation and a project. In that case, start by creating category options and categories for all partner organisations and projects, before linking these in a new category combination. Then, link the category combination to the data set (form) for which you need to capture this information. When opening this data set in data entry module, the partner organisation and project categories will automatically be rendered as drop-down boxes, allowing you to select a specific implementing partner organisation and project before continuing to do data entry.



16.2. Data Entry Forms

Once you have assigned a data set to an orgunit that data set will be made available in Data Entry for the orgunits you have assigned it to. A default data entry form will then be shown, which is simply a list of the data elements belonging to the data set together with a column for inputting the values. If your data set contains data elements with a non-default categorycombination, such as age groups or gender then additional columns will be automatically generated in the default form based on the different options/dimensions.

If you use more than one dataelement category combination you will get multiple columns in the data entry form with different column headings for the options. In addition to the default list-based data entry form there are two more alternatives, the section-based form and the custom form.

16.2.1. Section forms

Section forms allow for a bit more flexibility when it comes to using tabular forms and are quick and simple to design. Often your data entry form will need multiple tables with subheadings, and sometimes you need to disable (grey out) a few fields in the table, both of these functions are supported in section forms. This function can be access by choosing Maintenance->Data set Section.

16.2.2. Adding a new section form

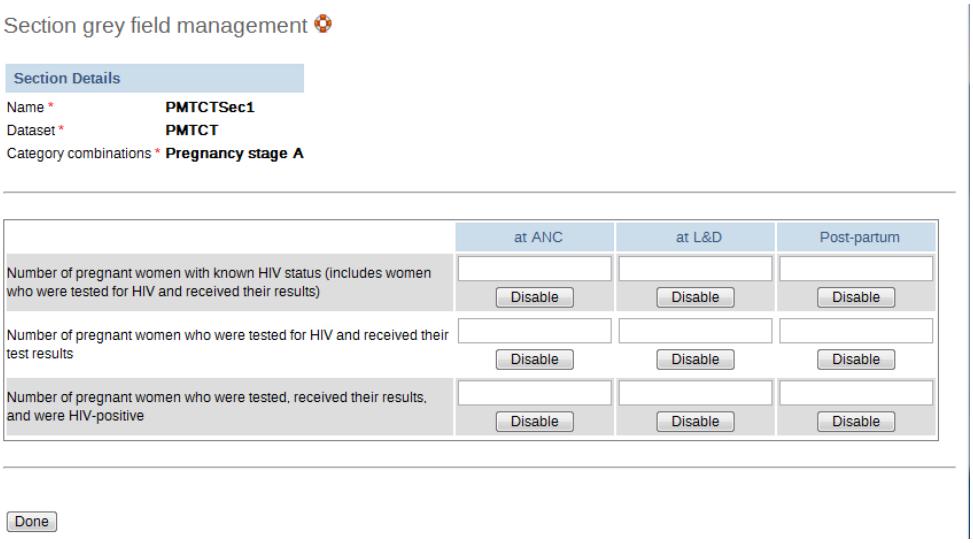
Section forms are separated automatically by data element category combinations, which produce a spreadsheet like data entry form for each section.

When designing a section form the procedure is as follows:

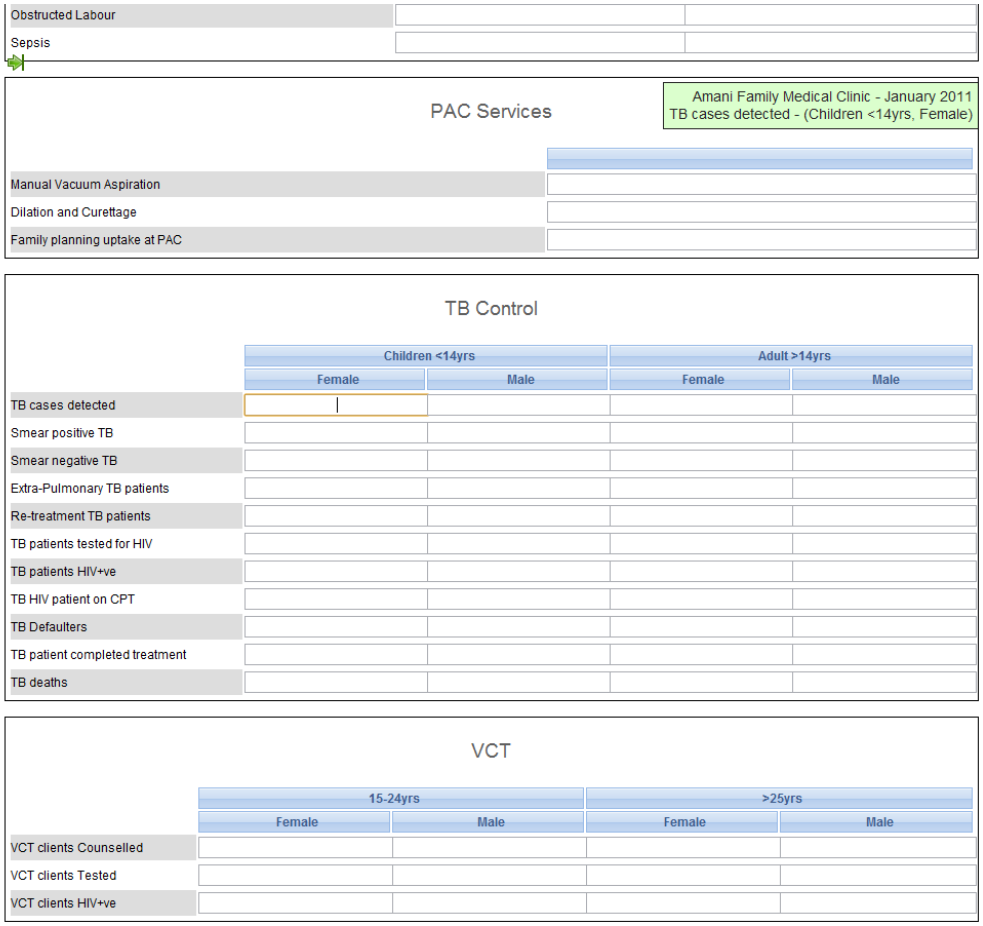
1. Set up your data set as described in Section 16.1, “Data sets”
2. Open the Data Set Section window (from right side menu under Data sets) and add your sections one by one. To add a new section to a section form, first choose the data set from the "Select data set" combo box. Then choose the specific category combo and press "Add new". You can now add data elements from the "Available data element" list on the left to the "Selected data elements" list on the right. Data elements can be sorted within the section with the use of the "Move up" and "Move down" buttons. Be sure to press "Save" once you have finished.

   Note

   You can only use one data element category combination per section.
3. You may need to control how the data element sections are displayed on the final form. In Data set Section management, select the data set from the "Data set" drop-down box, then leave [All] in the "Select Category Combo" drop-down. Click on "Sort section" to sort the order of appearance of your sections in the data entry form.
4. In Data Entry you can now start using the Section form (should appear automatically when sections are available for the selected data set). Data sets which have section forms will automatically display the section form.
5. Certain data elements may need to be disabled for data entry. Clicking on the "Section grey field management" menu item will allow you to disable specific data element category options as seen below. Pressing the "Disable" button will prevent data from being entered into this specific data element/category option during data entry. Be sure to press "Done" to save your changes.



A sample section form is displayed in the next figure. Notice how each data element category has been separated into a separate section, and a data entry table has been automatically generated by the system. Use of section forms in combination with data element categories can drastically reduce the amount of time which is required to create data entry forms for data sets.



16.2.3. Custom Forms

When the form you want to design is to complicated for the default or section forms then your last option is to use a custom form. This takes more time, but gives you full flexibility in term of the design. DHIS2 uses a built-in HTML editor (FcK Editor) for the form designer and you can either design the form in the UI or paste in your HTML directly (using the Source window in the editor). A complete reference for use of the editor can be found here.

One of the big advantages of custom forms, is that they can be created to mimic existing paper aggregation forms. This makes data entry much easier for users, and should reduce the number of data elements which are incorrectly entered, as they are more easily identifiable when entering data from a paper form.

Once a custom form has been added to a data set it will be available in data entry and used automatically.

Note

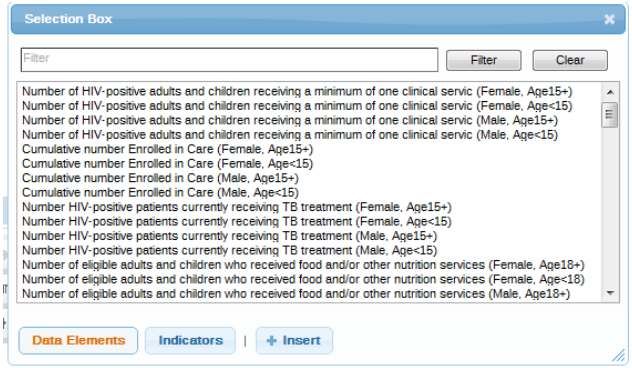
Custom forms are preferentially displayed over section forms. If a data set has both a section form and a custom form, the custom form will be displayed during data entry. Users will not be able to select which method they wish to input data, so be sure that your custom form contains all data elements which may be required.

To add a custom form design to a data set then first locate your data set in the Data set Management window and click on the Design data entry form icon under Operations (the fifth icon), see the mouse-over text to be sure.

First provide a Name for the form. There are a few important buttons in the Editor that you must pay special attention to. The blue monitor icon is the full screen mode on/off button, which can be very useful. The there is a Source button that shows the HTML code for your form.

If you already have the HTML for your form then you should start by pasting it in here. Click on Source again to go back to preview/non-HTML mode. Then there is an icon in the top right corner with a + sign on it, this will open a list of available data elements to add to your form, the Data Element Selector window.

All the input fields need to have a link to a data element or indicator. To add new data elements to the form, double-click them from the data element/indicator box as shown below. You can also select a data element/indicator and press the "+Insert" button. You can switch between either data elements or indicators by pressing the respective buttons.



You can to intermediary saving by clicking on the Save button, and this will not close the window. It is recommended to save often to ensure you do not loose your work.

When you are done or want to test your form in data entry click on <Save and close>.

16.2.4. Scripting in custom forms

In custom data entry form you can use javascript to create dynamic behavior and customziations. The DHIS 2 data entry module provides a range of events which you can register for and use to perform actions at certain times. The events are registered on the document element. The jQuery event object and the data set identifier are always the first two arguments provided to the callback functions. The table below provides an overview of the events and when they are triggered.

Table 16.1. Data entry events

Key	Description	Arguments
dis2.de.event.formLoaded	Triggered after the data entry form is rendered, but before data values are set in entry fields.	Event | Data set ID
dhis2.de.event.dataValuesLoaded	Triggered after data values are set in entry fields.	Event | Data set ID
dhis2.de.event.formReady	Triggered when the data entry form is completely rendered and loaded with all elements.	Event | Data set ID
dhis2.de.event.dataValueSaved	Triggered when a data value is saved successfully.	Event | Data set ID | Data value object
dhis2.de.event.completed	Triggered when a data set is successfully marked as complete.	Event | Data set ID | Complete registration object
dhis2.de.event.uncompleted	Triggered when a data set is successfully marked as incomplete.	Event | Data set ID
dhis2.de.event.validationSuccess	Triggered when validation is done and there were no violations.	Event | Data set ID
dhis2.de.event.validationError	Triggered when validation is done and there were one or more violations.	Event | Data set ID
dhis2.ou.event.orgUnitSelected	Triggered when one or more organisation units are selected in the org unit web tree.	Event | Org unit IDs | Org unit names | Sub org unit IDs

To register for such an event:

```
<script type="text/javascript">

dhis2.util.on( 'dhis2.de.event.formReady', function( event, ds ) {
  console.log( 'The form with id: ' + ds + ' is loaded!' );
} );

dhis2.util.on( 'dhis2.de.event.dataValueSaved', function( event, ds, dv ) {
  console.log( 'Data value: ' + dv.value + ' was saved with data element: ' + dv.de );
} );

dhis2.util.on( 'dhis2.de.event.completed', function( event, ds, cr ) {
  console.log( 'Form was completed for org unit: ' + cr.ou );
} );

</script>
```

Be careful to only use "namespaced" events like the ones in the example above and not general ones like "click" as the dhis2.util.on method will unregister the event first.

If your function only applies to certain data sets you can use the supplied data set identifier and shortcut your function for unwanted data sets like this:

```
dhis2.de.on( 'dhis2.de.event.validationSuccess', function( event, ds ) {
  if ( $.inArray( ds, ['utXOiGbEj14', 'Re7qzHEThSC'] ) == -1 ) {
    return false;
  }
  console.log( 'Form with id: ' + ds + ' validated successfully!' );
} );
```

The identifiers of the input fields in the data entry form is on the format described below. This format can be used to select the input fields in your script and perform actions on them:

```
<dataelementid>-<optioncomboid>-val
```

Since the data set identifier is provided for all events a feasible alternative is to utilize the "files" Web API resource and keep your callback functions in a single file, where you let the javascript code take action based on which data set is currently loaded.

16.2.5. Data set assignment editor

The data set assignment editor is a tool for adding and removing many data sets to organisation units in batch style. Start by selecting an organisation unit from the selection tree. In the area below the tree a grid will be displayed showing all data sets as columns and the child organisation units as rows.

From the grid you can now assign or unassign data sets simply by clicking on of the corresponding icons in the grid. If you want to assign or unassign an organisation unit to all data sets you can check or uncheck the checkbox next to the organisation unit. Your changes will automatically be saved.

Chapter 17. User management

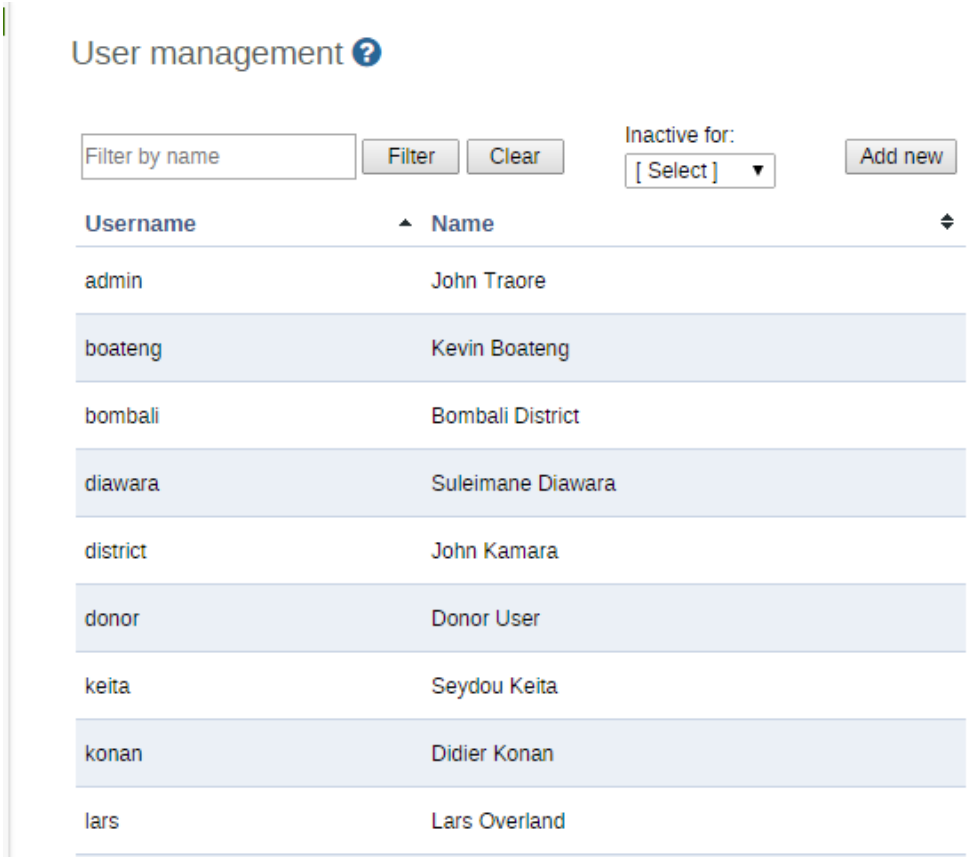
DHIS2 allows for multiple users to access the system simultaneously, each with a defined set of permissions. These permissions can be finely tuned so that certain users can only enter data, while others may generate reports. Multiple user roles can be created, each with their own set of permissions, and then assigned to users which grant them certain privileges within the system. This chapter describes how to manage users and user roles.

17.1. Creating new users and roles

This section will describe how to add new users and manage existing users to the DHIS2 application. You can create as many user names as you need. Each user can be assigned certain privileges, and can be assigned to certain organisation units for which they will be enabled to enter data on behalf of. To access the user module, choose *Users* from the Apps menu and then click "User" from the menu items on the left-hand pane.

17.1.1. User maintenance

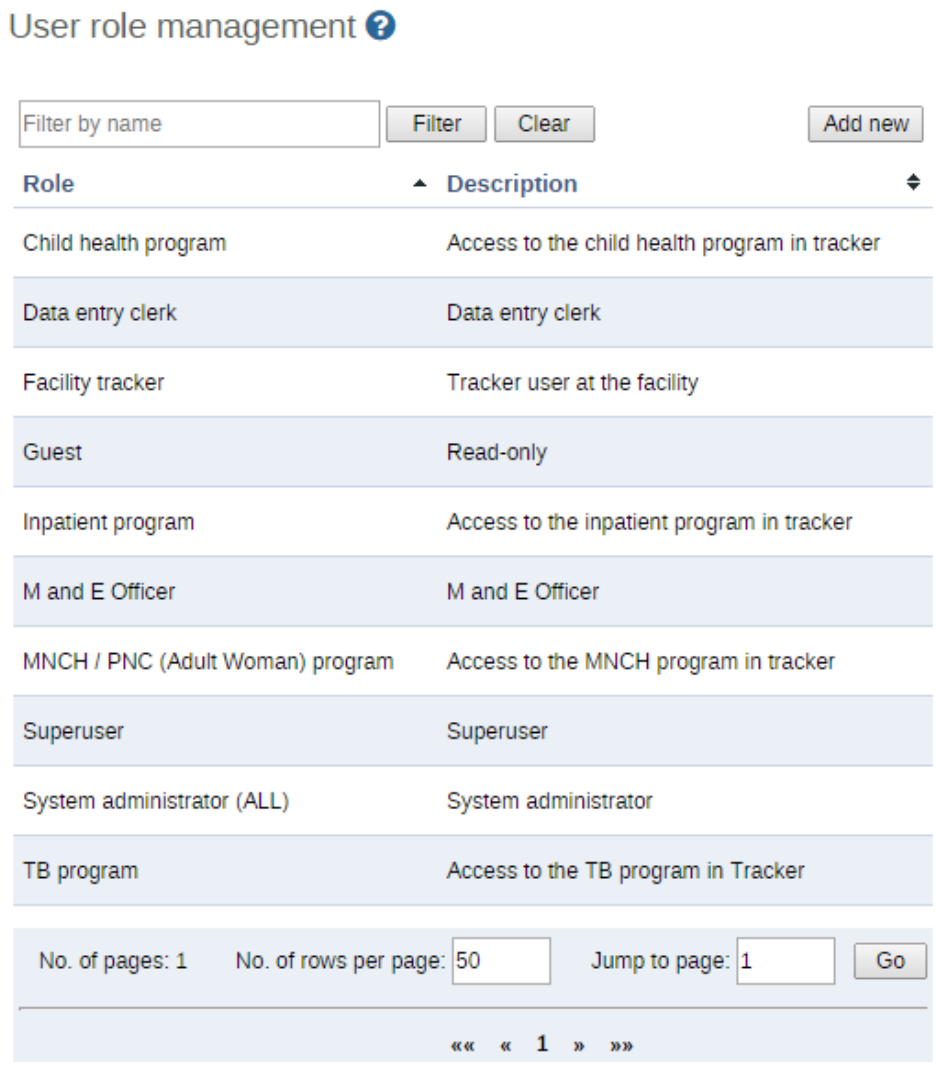
User names already registered will appear as a list as seen in the screen shot below.



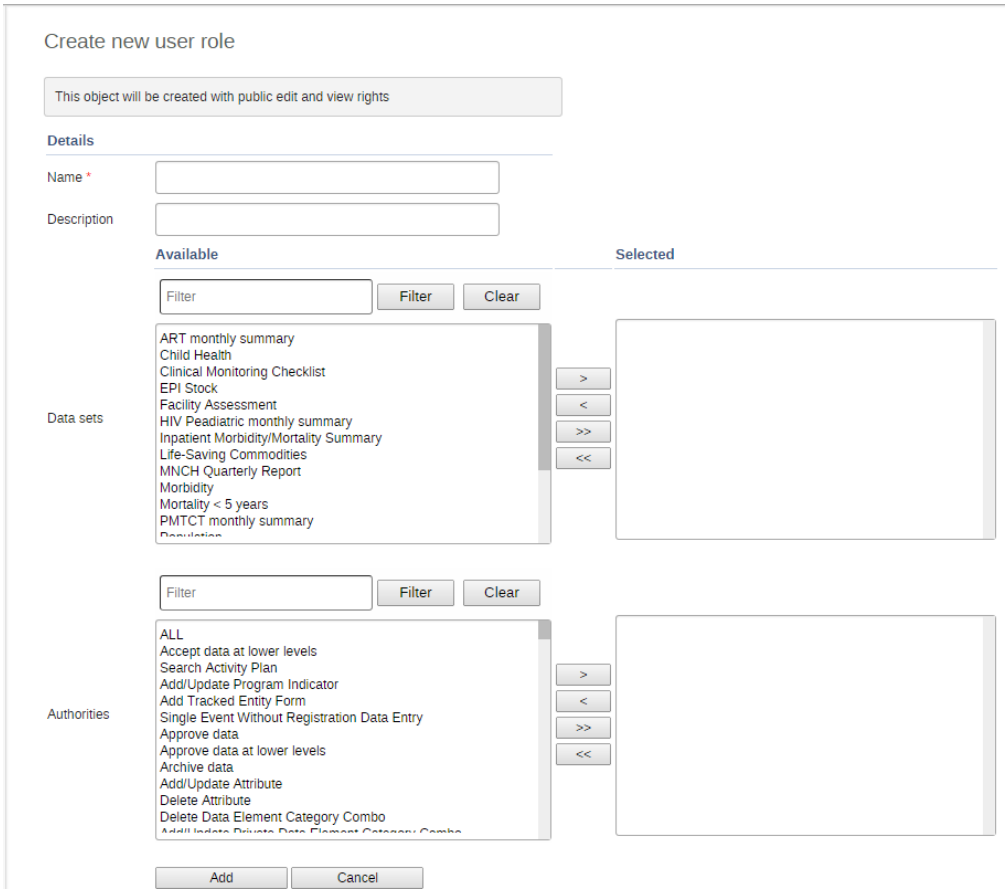
You can search for specific user names in the user list by entering the name in the *Filter by name* field as shown above.

17.1.2. User role management

As part of creating a user name you are required to define the user role. Do so by clicking on *User Role* on the left side of the displayed screen. This will lead you to the User role management page where you can click on Add new to create a new role.



The following screen will open and here in the first text box you need to give a *Name* of the Role such as Super User, Admin User, etc. The second text box called *Description* gives more information about the type of User Role that is being created for e.g. State Admin User, District Data Entry.



Next you will specify the particular data set(s) that are to be made available to the particular role. You will also need to specify the type of *Authority* to be given to the particular user. For each of the three options namely Datasets, Reports and Authorities user can select multiple options from the scroll down menu provided against each field. A user can choose multiple options either by moving them one-by-one.

In order for particular users to be able to enter data, you must add them to both a dataset as well as an organisational unit level. You can also select multiple datasets individually by pressing the Ctrl key on the keyboard and clicking on individual datasets.

Finally when you have entered the required fields click on Save which is located on the lower part of the displayed screen. The desired user role and related authorisation will be saved to the database, and can then be assigned to a particular user.

17.1.3. User management

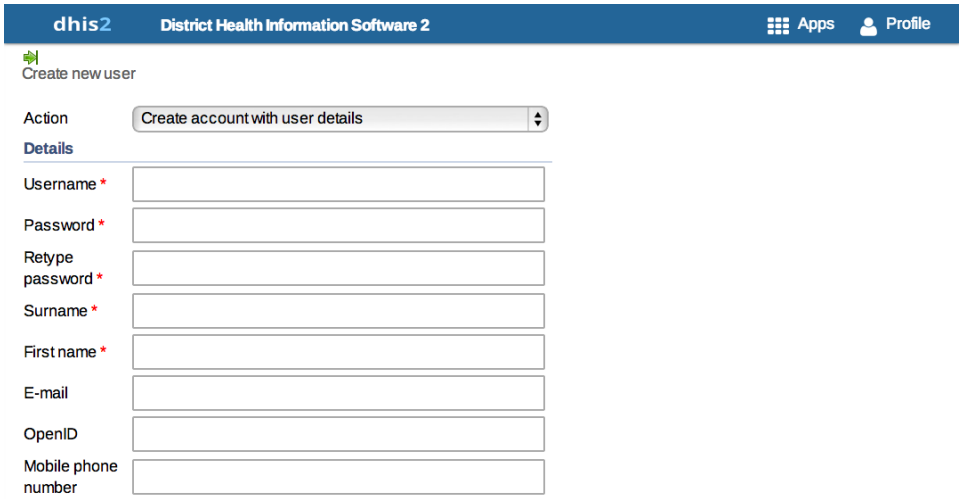
Under particular user role there can be more than one user. To manage users, click on *User* on the left side of the screen. This will lead you to the *User management* page.

To add a new user, follow these steps:

1. Click on the Add New button.
2. Choose whether you want to fill in all the personal user information now, or invite the user by email to complete the rest of the user information:
   - *Create account with user details* - Choose this if you would like to enter all the details of the new user such as name, password, etc.

     If you choose this action, then enter the following information: username, password, surname, first name, E-mail, OpenID account (if any) and mobile phone number (if any).

     After you finish adding the user, the account will be ready for them to use with the username and password that you supply.

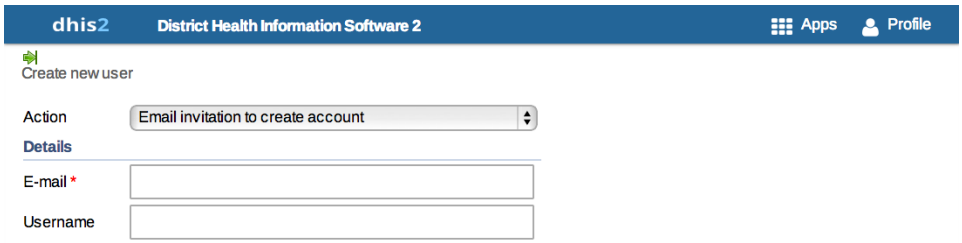


- *Email invitation to create account* - Choose this if you would like to send by email an invitation for the user to return to the system and finish setting up their user account. The user will then return to the system and fill in most of their personal information. The account that the user finishes setting up will be limited according to how you configure it below.

  Note that you may not select this option to create an account with "critical" system authorities such as All, Scheduling Administration, Perform maintenance tasks, Merge organisation units, Eliminate duplicate data elements, Sql View Management, Change system settings, and List, Add or Delete user roles.

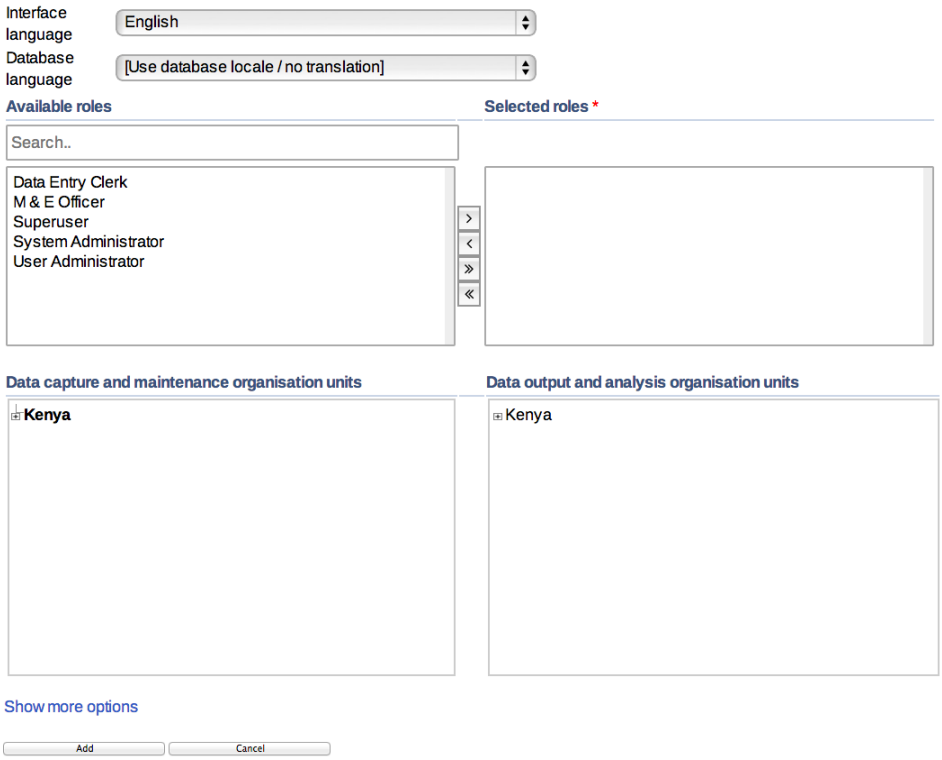
  If you choose this action, then enter the email address to which the invitation should be sent. If you want to, you may also enter the user name that the account will have. If you leave the username empty, then the user may choose their own username when they respond to the invitation (as long as it is not taken already for another user.)

  After you finish adding the new user, two emails will be sent to the address you provided. One contains a unique web link by which the user can return to the system and activate their account by entering the rest of their user information. The other email contains a unique code that they must enter into the system in order to complete the registration, after following the link in the first email. The user must finish setting up the account within three months, or the invitation becomes invalid.

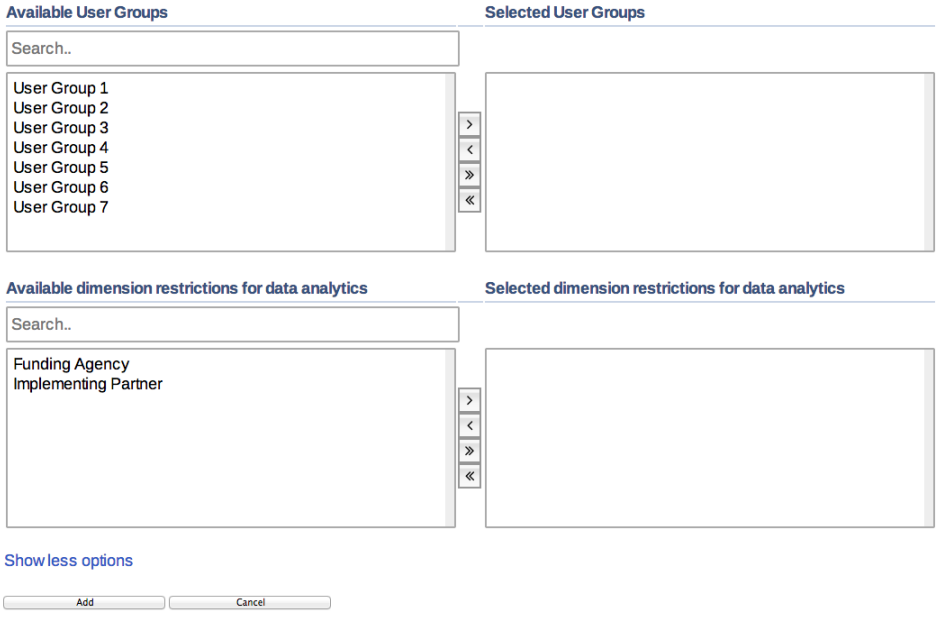


3. Select the *Interface language* for the user. You may choose a language into which fixed elements of the DHIS2 user interface have been translated.
4. Select the *Database language* for the user. You may choose a language into which implementation-supplied items have been translated in the database, for example data element names, organisation unit level names, etc.
5. Users must be assigned to at least one data capture and maintenance organisation unit. Users will have access to all children of the organisation units which have been assigned to them. For instance, if a user has been assigned to a district which has several facilities contained in the district, the user would have access to the district's data, as well as all of the facilities contained within the district. The data approval organisation units control for which organisation units the user can do data entry.
6. Users can be assigned to any number of data view organisation units. This controls which organisation units the user can view aggregated data for in analysis modules. Giving access to an organisation unit implicitly gives access to all organisation unit below it in the organisation unit hierarchy. Note that data view organisation units are optional. If you do not specify any, the user will have access to the full organisation unit hierarchy for viewing aggregated data.

   In several places in the analysis modules one can select "user organisation unit" for the organisation unit dimension. This mechanism will first attempt to use the data view organisation units linked to the current user. If not found, it will use the data capture organisation units.



7. (Click on *Show more options*.) You may optionally assign users to user groups on this page.
8. (Click on *Show more options*.) You may optionally restrict the values this user sees in data analytics by selecting dimensions that will restrict the users view. For example, let's say you have defined *Implementing Partner* as a category option group set, and you have shared with this user only one or more specific implementing partners (category option groups.) If you want to insure that the user does not see totals in analytics that include values from other groups, assign *Implementing Partner* to this user. This insures that any data visible to the user through DHIS2 analytics will be filtered to select only the Implementing Partner category option group(s) which are visible to the user.



9. Click on the Add button to complete adding the new user.

The recently created new user can be seen in main *User management* screen

You can edit (like password, surname, etc.) and delete the details of new/old users by selecting corresponding User's Edit and Remove menu options.

17.1.4. User by organisation unit

The *User by organisation unit* function allows you see which users have been assigned to a particular organisation unit. Simply select the organisation unit from the tree on the left, and a list of users which have been assigned to this particular organisation unit will be displayed

17.1.5. OpenID Support

DHIS 2 supports the OpenID standard, which allows third party login using a OpenID provider, please see http://openid.net for more information. To create a custom OpenID URL for a username you can visit this URL and log in with your OpenID provider: http://openid-provider.appspot.com.

To enable support for this in DHIS 2, two steps must be done:

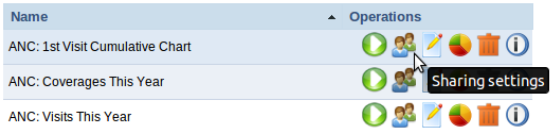
- **Set your OpenID provider**: This can be done inside system settings, under "Access". Here you can set both the OpenID provider, and also the label to display on the login page to login with this provider (defaults to Login with OpenID).
- **Set the OpenID identifier on the user**: For every user that should be able to login with his openid identifier, you will need to set this on the user itself. This can be done in user management, under the email field, there is noe a field called *OpenID* which can be used to fill in the OpenID identifier.

Chapter 18. Sharing

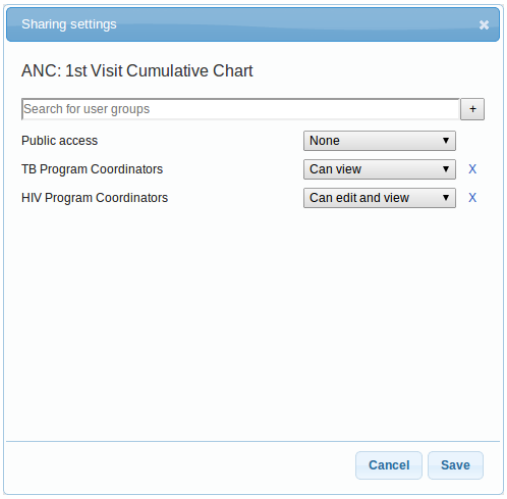
This chapter discusses the sharing of entities feature in DHIS 2.

18.1. Sharing of objects

A large part of the objects in DHIS 2, like reports, charts, maps and indicators, can be shared. Sharing means making an object, like a report, available for reading or modification to a group of users or to everyone. For instance for reports, the sharing dialog can be opened by clicking on the "Sharing settings" button next to each report in the list.



Clicking on the "Sharing settings" button will open the sharing dialog. The sharing dialog is where sharing is controlled.



You can share your report with everyone or with a number of user groups. Next to "Public access" you can choose your public access option: "None", "Can view" or "Can edit and view". Edit also implies deleting the report. To share with a group, simply start typing the name of the group and the "Search for user groups" input field and select your desired group. Click on the "+" icon next to the input field to share with that group. For each group you can set an access option, similar to public access.

Sharing with a user group implies that all users in that group will get access to the shared object. To create a user group you can go to the dashboard module and click on "Groups" on the front page. This will lead you to the list of groups where you can click "Add new" in the top right corner. Creating user groups is open for everyone from the dashboard module.

18.2. Sharing and access control

The objects which support sharing are indicator, indicator group, indicator group set, data dictionary, data set, program, standard report, resource, report table, chart, map and user group. Out of those objects, report table, chart, map and user group are open for everyone to create privately. Private means that the objects are available only to yourself or potentially to a number of user groups if you choose to share the object. These objects are referred to as "open" objects and can be created by all users. The remaining objects require that your user account has the authority to create them. These objects are referred to as "non-open" objects.

A user can be granted the authority to create publicly accessible objects or privately accessible objects. In order to create a publicly accessible object (available for viewing or editing by anyone) your user account must have the authority to do so. As an example, to create a publicly accessible chart, your user must have the "Create public chart" authority

granted. The authority to create private objects applies only to non-open objects. For example, to allow a user to create indicators which will only be accessible to that user and not to everyone, the user can be issued with the "Create private indicator" authority.

Sharing a non-open object with another person and let her edit the object requires that the person's user account has the authority for updating that type of objects granted. For instance, if you want to let another person edit your indicator, that person's user account must have the "Update indicator" authority granted. This does not apply for open objects.

When you create a new object it will automatically become viewable for everyone if your user account has the authority to create public objects. As an example, if you create a standard report and you have the "Create public standard report" authority granted, the report will become viewable for everyone. If you do not have that authority granted the report will be viewable only to yourself. After you have created an object, you may navigate to the "Sharing settings" dialog and set your desired access control level.

If you need a user account which is able to view absolutely all objects you can create a user role with the "ALL" authority and assign a user to that role. If you need to switch between a "complete" view of objects and a "personal" view of objects it is recommended to create two user accounts, one assigned with the "ALL" authority and one without.

18.3. Sharing applied

The sharing functionality is useful in several scenarios. One use-case is setting up a DHIS instance for a global organisation with operations in multiple countries. Typically the organisation has a set of global data sets, indicators and reports which should apply to all countries, while all countries will have the need for country-specific data sets, indicators and reports. In this scenario the following approach could work:

- Set up one user group for global personnel.
- Set up a user group for personnel in each country.
- Create global data sets and reports, make them viewable for everyone and editable for the global user group only.
- Create country-specific data sets and reports, make them viewable and editable for the country user group and the global user group only.

This way, the global indicators and reports could be viewed and analysed by everyone, but maintained by the global user group only. The country-specific data sets, indicators and reports could be viewed and maintained by the country and global personnel, without being visible or impacting the system for other countries in the organisation.

A similar approach could work for a scenario with a donor, multiple funding agencies and implementing partners in a country, where user groups could be set up for each of those entities. That way each implementing partner could create and share their reports within their organisation without affecting or allowing access to others. Reports could also be shared with supervisors and funding agencies at the end of reporting periods.

Another use-case is a country department of health with multiple health programs. Typically there is a need for having general reports and charts for the department while allowing the health programs to develop specific reports and charts for internal use. This can be achieved by creating user groups for each health program. Later, when developing reports and charts, these can be made viewable and editable to the program user group only. This way the reports will not be visible to other programs and users. This is beneficial because the reports are kept internal to the program and because the visible list of reports of other users are kept shorter and more relevant.

Chapter 19. Setting up Data Quality functionality

The data quality module provides means to improve the quality of the data in the system. This can be done through validation rules and various statistical checks.

19.1. Overview of data quality check

Ensuring data quality is a key concern in building an effective HMIS. Data quality has different dimensions including:

- *Correctness:* Data should be within the normal range for data collected at that facility. There should be no gross discrepancies when compared with data from related data elements.
- *Completeness:* Data for all data elements for all health facilities/blocks/Taluka/districts should have been submitted.
- *Consistency:* Data should be consistent with data entered during earlier months and years while allowing for changes with reorganization, increased work load, etc. and consistent with other similar facilities.
- *Timeliness:* All data from all health facilities/blocks/Taluka/districts should be submitted at the appointed time.

19.2. Data quality checks

Data quality checking can be done through various means, including:

1. At point of data entry, the software can check the data entered to see if it falls within the min-max ranges of that data element over the last six months or as defined by the user.
2. Defining various validation rules, which can be run once the user has finished data entry. The user can also check the entered data for a particular period and Organization Unit(s) against the validation rules, and display the violations for these validation rules.
3. Analysis of data sets, i.e. examining gaps in data.
4. Data triangulation which is comparing the same data or indicator from different sources.

19.3. Data quality check at the point of data entry

Data quality can be checked at the point of data entry through setting the minimum and maximum value range for each element manually or generating the min-max values using the DHIS 2 if there is historical data available for that data element.

19.3.1. Setting the minimum and maximum value range manually

If you are using the default entry screen click on the element for which you want to set the min-max value. A pop-up window will appear in which you can enter the values. On subsequent data entry if the value entered does not fall within the set min-max range the text box will change colour to red. The user will also get a pop-up as shown below. This change in colour is a prompt to check the data entered and make necessary correction. On the data entry screen the users also have the option to add a comment on how the discrepant figure might be explained (if required). This you can do by using the drop down menu of the ‘comment’ box. In case you are using the custom data entry screen which is displayed when you deselect the ‘default data entry form’ option on the top right corner of the screen. In this case the minimum and maximum values can be added by double-clicking on the data entry box instead of the data element.

19.3.2. Generated min-max values

It is possible to generate the min-max value, element-wise, using the DHIS2. In such case you merely need to click on the ‘Generate min-max’ button near the upper right corner. In case of default data entry screen the min and max

values, when generated, will appear on the left and right side of the data entry box. In case you deselect the default data entry form the generated values will appear on the top right end of the screen.

19.4. Validation Rule

This module provides management of validation rules. A validation rule is based on an expression which defines a relationship between a number of data elements. The expression has a left side and a right side and an operator which defines whether the former must be less than, equal to or greater than the latter. The expression forms a condition which should assert that certain logical criteria are met. For instance, a validation rule could assert that the total number of vaccines given to infants is less than or equal to the total number of infants.

To add a validation rule, just follow these steps:

- Click on the *Add new* button
- Provide a descriptive *Name* for the validation rule. The name must be unique among the validation rules.
- Provide a *Description* for the validation rule.
- Select the *Importance* you wish to assign to the validation rule: *high*, *medium* or *low*.
- Select the *Rule type*. For validating correct data, choose *Validation*. For monitoring data according to a rule, choose *Surveillance* and follow the extra instructions below.
- Select the *Period type* for the data being validated.
- Select an *Operator*. The operator options are *equal*, *not equal*, *greater than*, *greater than or equal*, *less than*, *less than or equal to* and *compulsory pair*.
- Define the *left side* and *right side* of the validation rule expression. First, provide a description for the expression. Second, build the expression with the expression builder. The expression is mathematical and contain data elements as well as integers and mathematical operators. Data elements can be included by double-clicking one in the available data elements list to the right. Alternatively one can select a data element and click the insert button. Mathematical operators can be included by clicking the corresponding button under the expression builder area.
- Save each expression by clicking *Save*, then save the validation rule by clicking *Save*.

When editing an expression, check the setting *skip for missing values* to require that all data element values in the expression must be present. If any values are missing, the validation rule will be skipped. If this setting is unchecked it means that the expression will be evaluated even if some of the values are missing.

The compulsory pair operator allows you to require that data values must be entered for a form for both left and right sides of the expression, or for neither side. In other words, you can require that if one field in a form is filled, then one or more other fields must also be filled.

To edit a validation rule, click the *edit*icon next to the relevant validation rule in the list. Then follow the same procedures as above.

To delete a validation rule, click the *delete*icon next to the relevant validation rule in the list.

To view validation rule details, click the *view details*icon next to the relevant validation rule in the list.

19.5. Surveillance Rule

As well as checking for correct data, you can use a validation rules to find unexpected data values when compared with data from previous time periods. This kind of validation rules are called *surveillance* rules.

To add a surveillance rule, follow the steps above for validation rules, choosing a rule type of *Surveillance*. This adds the following validation rule options:

- Choose an *Organisation unit level* for this surveillance rule. If the data you wish to monitor is not entered at this level, it will be aggregated from lower level organisation units in the organisation unit hierarchy.
- Enter a *Sequential sample count*. This is the number of time periods immediately preceding against which you wish to compare the data. For example, if the period type is *Weekly* and this count is 2, the current data will be compared with past data averaged over each of the 2 previous weeks.

- Enter an *Annual sample count*. This is the number of preceding years over which you will compare the data, from periods at the same time of preceding years. For example if the period type is Weekly and the count is 3, data for a week starting on the first of September will be compared with data averaged over the week containing September 1 in each of the past 3 years.

  The sequential sample count *or* annual sample count must be at least 1.

  You may use the sequential and annual sample counts together. For example, say the period type is *Weekly*, the sequential count is 2, and the annual count is 3. The data will be compared with the average of the following time periods: the 2 weeks immediately preceding, and for each of the previous 3 years the data at the same time of year, plus the preceding 2 weeks, plus the following 2 weeks. This makes a total of 17 past periods: 2 immediate past periods plus 5 periods for each of the 3 preceding years. Be aware that when you have many past periods like this, evaluating the surveillance rule may take significant time and system resources.

- Enter a number of *High outliers* of past data that you wish to exclude from comparison. This is useful if some past periods may have had unusually high data values, and you wish to compare against the average of all but the highest past period values. This gives the number of highest past period values to exclude before the past period values are averaged and compared with the current value.

- Enter a number of *Low outliers* of past data that you wish to exclude from comparison. This is like High outliers except that it gives the number of lowest past period values to exclude before the past period values are averaged and compared with the current value.

  You may use the high and low outliers in combination, but the sum of high outliers and low outliers must be less than the total number of past samples as determined by the sequential and annual sample counts. In the example above where there are 17 past periods, the sum of high and low outliers must be 16 or less.

  If data is not found for all the desired past periods, the high and low outliers will be reduced in proportion to the number of periods for which data is found. For example, say we are looking for 17 past periods with high outliers set to 4 and low outliers set to 2. If data is found for only 9 of these past periods, only 2 high outliers will be discarded, and only 1 low outlier will be discarded.

When a surveillance rule is evaluated, the *left side* of the equation is evaluated for the current period, and the *right side* is evaluated for each of the past periods. The right side values for past periods are averaged, and the average is compared to the left side value according to the operator. Any high or low outliers are removed before the right-side average is taken.

A surveillance rule may contain data elements for periods that are longer than the period given for the rule. This is useful, for example, if you want to divide a data element value by the population count, and the population count is entered annually. The data for the longer period type (e.g. population count) must be entered for a period that overlaps with the start of the period being evaluated. For example, if the period being evaluated is the week starting January 6, 2014, the yearly population count must be entered for the year containing January 6 2014. If the data element is from a longer period type than the rule period type, it must have an aggregation operator of average, not sum. In other words, only data elements that don't sum through time (like population count) can be used from longer periods than the rule period.

For surveillance rules, the *skip for missing* option is given an additional meaning when data is being collected and aggregated from lower levels of the organisation unit hierarchy. When this option is selected and the data is present for some descendants at a lower level, but not all descendants at that level, the rule is skipped.

19.6. Validation Rule Group

A validation rule group provides a mechanism for classifying related validation rules. Another advantage of using validation rule groups is that it can later be run separately, instead of running all validation rules.

You can also use a validation rule group to configure how users are notified of alerts from scheduled validation runs. To do this, you should identify a set of validation rules you want to evaluate regularly, and a group of users who should be notified of any exceptions to these rules. Then:

- Be sure that one or more user groups are defined containing all the users you wish to notify.
- Define a validation rule group for a set of validation rules. In the section User groups to alert, select one or more user groups to be notified.

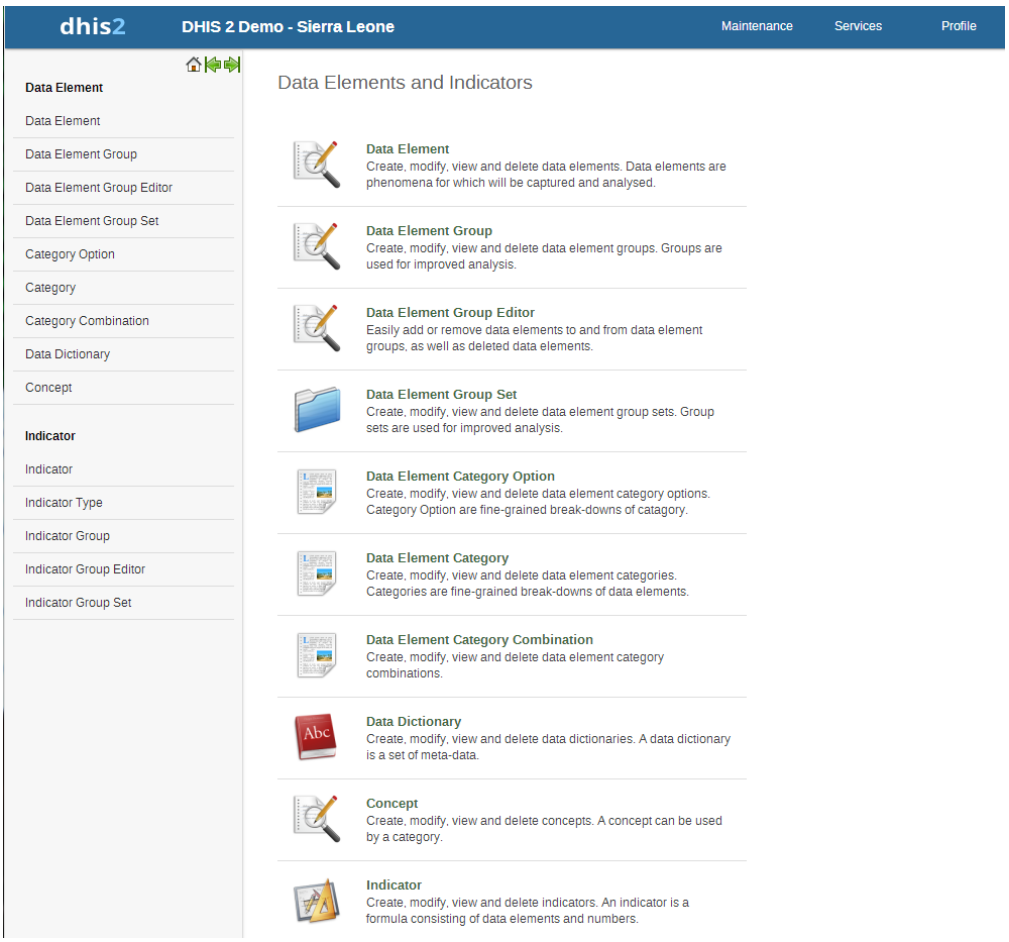
By repeating these two steps, you can build any set of relations between validation rules and users to fit your needs.

When you create or edit a validation rule group, there is an option called *Only organisation unit related users are alerted*. If this is set to *Yes*, then each user in the group(s) will be alerted only for validation exceptions for an organisation unit to which the user has been assigned through user management -- or for any lower-level organisation unit under that organisation unit. If this is set to *No*, then each user in the assigned user group(s) will be alerted for all validation exceptions in the group, regardless of organisation unit.

To enable routine scheduling of data validation runs, choose Data Administration from the Maintenance menu. Then click on Scheduling. If scheduling is active, click on the Stop button, then select the Data monitoring strategy of All daily. Finally enable scheduling by clicking on the Start button.

Chapter 20. Indicator maintenance

Indicators can be maintained by choosing "Maintenance->Data elements and indicators" from the main menu. The following screen will appear.



From the left side menu or by scrolling down the central area you can access the various sections on Indicators;

Indicator, Indicator Type, Indicator Group, Indicator Group Editor, and Indicator Group Set.

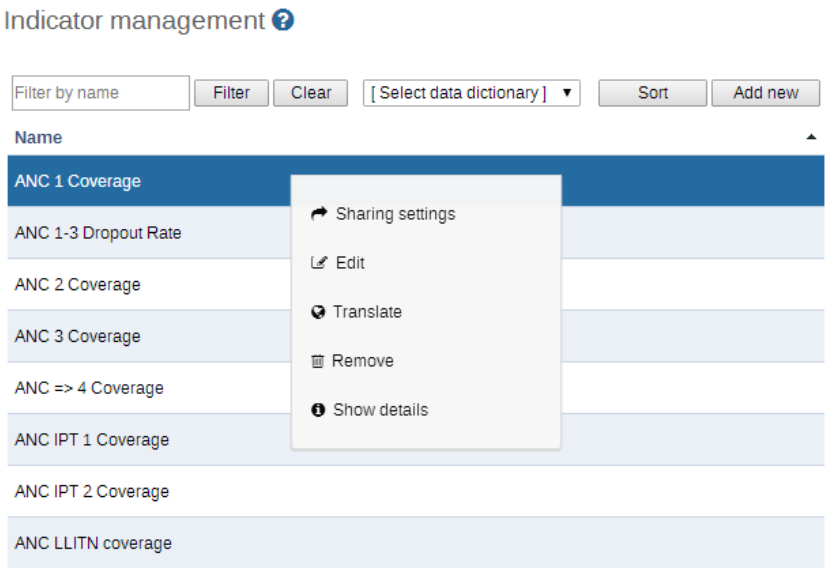
20.1. Indicator maintenance

Indicator maintenance functions essentially the same as each of the respective sections in the previous section on data elements. The basic operations will be described in this section, but the reader should refer to the corresponding sections above for detailed instructions.

20.1.1. Indicators

Indicators are composed of multiple data elements, and typically consist of a numerator and denominator. Calculated totals do not have a denominator. Indicators are never entered directly in DHIS2 through data entry, but are derived from combinations of data elements and factors. Indicators are used to calculate coverage rates, incidence and other values are a result of data element values that have been entered into the system.

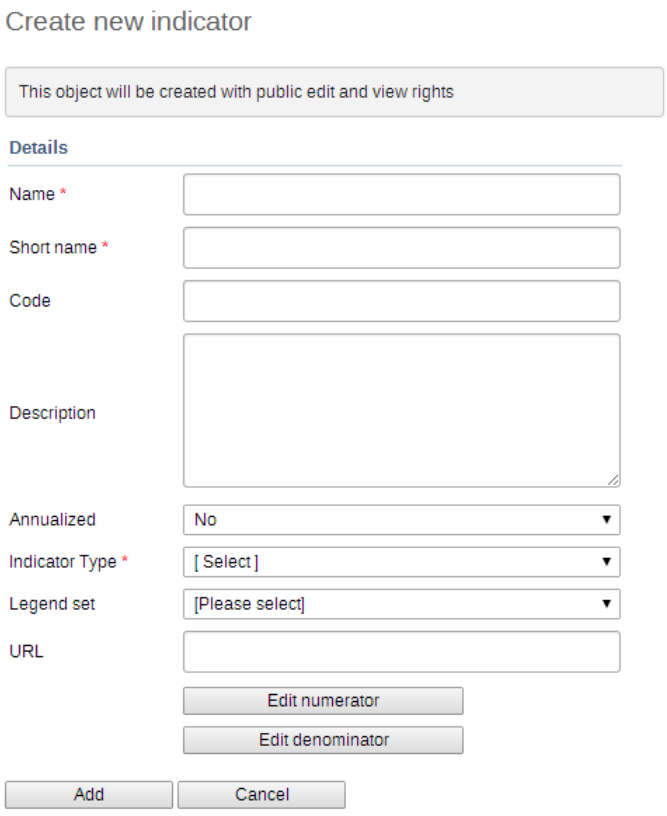
To access the Indicator maintains page, press Maintenance -> Data Element and Indicators and then pressing "Indicator" from the left pane or the main dialog. Similar to data elements, you can add, delete, modify and view extra information about the indicators in the system.



Indicators can be filtered by entering the name or a part of the indicator name in the "Filter by name" field. Similar to data elements, indicators can be added by pressing the "Add new" button. Other operations available from the context menu (available by clicking on the indicator name) are as follows.

- Sharing settings: Assign sharing privileges to user groups.
- Edit: Edit an existing indicator
- Translate: Translate an existing indicator to a different language.
- Remove: Delete an existing indicator.
- Show details: Get detailed information about this indicator.

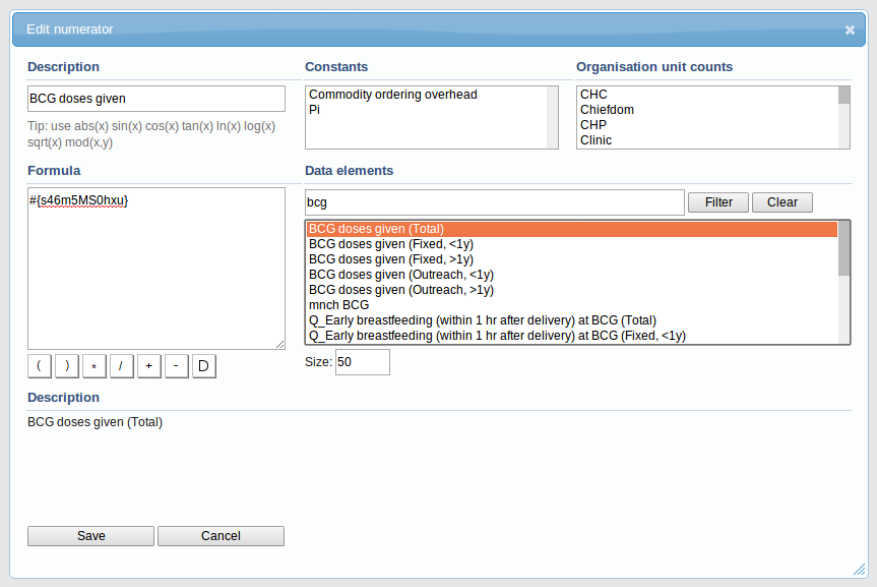
To add a new indicator, click the "Add new" button. The following screen is displayed.



Each of the fields marked with an asterisk are compulsory. A description of each field is provided below.

- Name: The full name of the indicator, such as "Incidence of confirmed malaria cases per 1000 population"
- Short name: An abbreviated name of the indicator such as "Inc conf. malaria per 1000 pop". The short name must be less than or equal to 25 characters, including spaces.
- Code: In many countries, indicators are often assigned a particular code. This code can be entered here.
- Description: A brief, informative description of the indicator and how it is calculated can be entered here.
- Annualized: Determines whether or not an annualization factor is applied during the calculation of the indicator. Typically, annualized indicator's numerator are multiplied by a factor of 12, and the denominator is for instance a yearly population figure. This allows for monthly coverage values to be calculated with yearly population figures.
- Indicator type: This field will determine a factor that will automatically be applied during the calculation of the indicator. Possible choices are determined by the Indicator Types (described below). For instance, a "Percent" indicator will automatically be multiplied by a factor of 100 when exported to the data mart, so that it will display as a percentage.
- Legend set: Define a legend set for this indicator. Refer to the section on "GIS" for more information about legend sets.
- URL: Can be used as a link to an indicator registry, where a full metadata description of the indicator can be made available.

To define the numerator and denominator, simply press the respective button, and the following dialogue will be displayed.



Essentially, an indicator is a formula that can be composed of multiple data elements, constants, organisation unit group counts and mathematical operators. In order to define a new indicator proceed with the following steps.

1. A description of the numerator/denominator must be provided in the "Description field". This should provide a clear description of the numerator/denominator. This field is required.
2. Define the formula of the indicator by selecting the data elements that should compose the numerator from the "Data elements" field. Simply select the data element, and double click it. It will now appears as part of the formula. You formula must be mathematically valid, including the proper use of parentheses when necessary. You can double click on each of the mathematical operator buttons below the indicator formula definition to add them to your formula.
3. Click the Save button to save all changes to the numerator. Click cancel to discard any changes that you have made.
4. Follow the same procedure in order to define the denominator.

The components of an indicator formula are listed below.

- Data elements: Will be substituted by the data value captured for the data element.
- Constants: Constants are numerical values which will remain the same for all indicator calculations. This is useful in order to have a single place to change values that might change over time.
- Days: A special operator "D" is available in formulas. This operator will always provide the number of days for a given indicator calculation. As an example, if you wish to calculate the "Percentage of time vaccine refrigerator was

non-functional", you could define the numerator as ("D-"Number of days vaccine refrigerator was available"")/"D". For example, if the fridge was available 25 days in June, the indicator would be calculated as (30-25/25)*100=17%. If you were however to calculate the total for Quarter 1, the number of days ("D") would be equal to 31+28+31=90. Thus, the "D" parameter will always be the number of days in the period of interest.

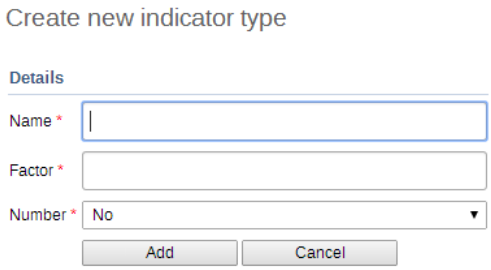
- Org unit group counts: Org unit groups can be utilized in formulas and will be substituted by the number of organisation units in the group. During aggregation, the org units in the group will be intersected with the part of the org unit hierarchy being requested. As an example, this lets you use the number of public facilities in a specific district in indicators. This is useful e.g. when creating facility infrastructure surveys and reports.
- Constants: Indicators may make use of constants. Constants are static values which will be applied uniformly to the indicator formula. Constants are applied AFTER data element values have been aggregated.

20.1.2. Indicator types

Indicator types simply define a factor that will be applied during aggregation. Indicator values that are calculate during a data mart export or report table generation process will appear properly formatted, and will therefore not require an additional multiplier (e.g. 100 in the case of percents) for the values to appear correctly formatted.

The indicator type maintenance panel has all of the same functions (Add new, Edit, Translate, Delete, and Information) as the Indicator maintenance section.

There are only two fields that need to be filled-in to create an indicator type, Name and Factor, as seen below. Name refers to the Indicator type (e.g. Per cent, Per thousand, Per ten thousand, etc). The factor is the numeric factor that will be multiplied to the indicator formula during the calculation of the indicator.



Note

As of version 2.4 of DHIS2, the "Calculated data element" object has been deprecated. Instead, a calculated data element can be created by creating an indicator type with a factor of "1" and by setting the "Number" option to "Yes". The effect of setting the "Number" option to "Yes" will be that the indicator will effectively not have a denominator. You will therefore only be able to define a numerator, which will serve as the formula of the calculated data element.

20.1.3. Indicator groups

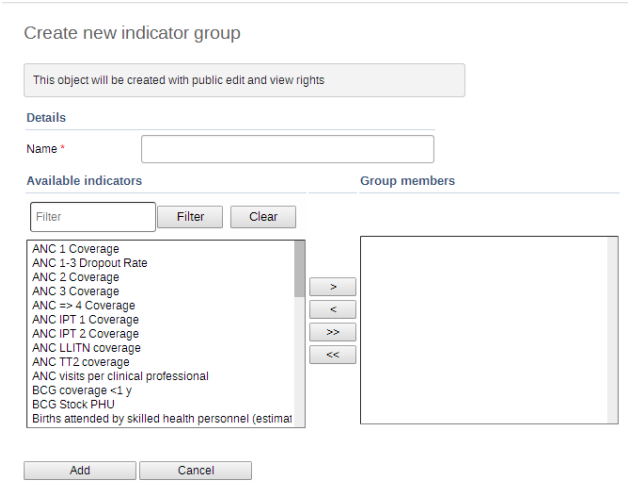
Indicator groups function essentially the same as data element groups. Multiple indicators can be assigned to a group for easy filtering and analysis. To assign indicators to groups, simple press Maintenance->Data elements and indicators->Indicator groups. See the section on Data element groups for detailed instructions of how to use this module.

20.1.4. Indicator group editor

The indicator group editor module functions essentially the same as the data element group editor module, except on indicators. You can easily rearrange the groups that indicators belong to with this module. To access it, choose To assign indicators to groups, simple press Maintenance->Data elements and indicators->Indicator group editor from the main menu. See the section on Data element group editor for further instructions.

20.1.5. Indicator group sets

Similar to data element group sets, indicator group sets serve to create combined groups of similar indicators. For instance, you might have a group of indicators called "Malaria" and "Leishmaniasis". Both of these groups could be combined into a group set called "Vector-borne diseases". Indicator groups sets are used during analysis of data to combine similar themes of indicators. To access this module choose Maintenance->Data elements and indicators->Indicator group sets from the main menu and then press "Add new". The following dialogue will appear.



Supply a name for the indicator group set, and then move the desired members from the "Available Indicator Groups" to the "Group members". Click "Add" to save your changes and "Cancel" to discard any changes.

Note

Similar to data element and indicator group sets, typically, indicator groups should be exclusive, meaning that one indicator group should not belong to multiple indicator group sets.

Chapter 21. Using reporting functionality

21.1. Reporting functionality in DHIS 2

The reporting module in DHIS 2 provides a range of reporting alternatives, and this section will explain how to use them to view and analyse data. Another section explains how to configure and set up the various reporting tools.

Standard reports: Standard reports are built on report tables, but are more advanced in its design allowing for more cosmetics and styles. These reports can also combine multiple tables and charts in the same report and be made available as one-click reports that are very easy to use. These reports can be downloaded as PDF files which makes them ideal for printing as well as sharing offline.

Dataset reports: Dataset reports are simply a printer friendly way to look at the data entry forms with either raw or aggregated data (over time or place). The design used in data entry will be used also in the data set reports. This will work only for data sets that has a custom data entry form set up.

Dashboard: The fastest way to view your data. The dashboard can display up to four updated charts as well as shortcuts to your favourite reports, report tables, and map views. Each user can configure a personal dashboard.

Data Visualizer: Do flexible visualizations of your data as charts and data tables. Any number of indicators and data elements can be included. Several chart types are available, such as column, stacked column, line, area and pie charts. The charts can be saved in order to be easily retrieved later and can also be put on your personal dashboard. Charts can be downloaded as image and PDF files to your local computer.

Report tables: These are very configurable table outputs of your data, either showing raw or aggregated data, as well as indicator data. These tables are used as either a data source for more advanced reports, for export to external systems, or as a crude report itself, and are exportable to PDF, excel, CSV and jasper design files. These tables represent a very dynamic, flexible and quick way to look at the data. Report tables can be set up with parameters to make them reusable over time and place.

Orgunit distribution reports: These reports are generated off the orgunit group set information and can show what types (and how many of each type) of health facilities that are located in a given area (any level in the hierarchy). These reports are automatically generated and display the information in both tables and charts, and downloads in PDF, excel, and CSV are available.

Reporting rate summary: These reports provide a nice overview of how many facilities that have submitted their data for a given dataset and period. Here you can get both the counts and the percentages showing the reporting rate for all or single data sets.

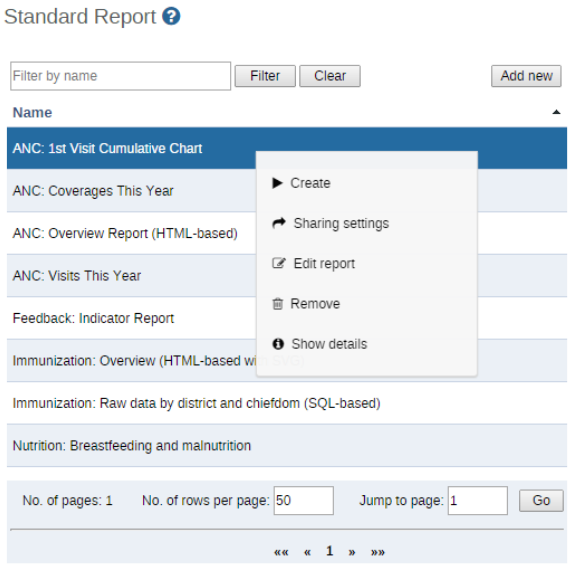
Excel pivot tables: Excel pivot tables represents a very powerful way to analyse your data and DHIS 2 links directly to the pivot tables so that all the data will be available and updated in your Excel file. This can be a very useful tool for users that prefer working with the data offline. To update your local pivot tables you need the myDatamart tool which connects to the online server and downloads the latest data. This update will typically take place once a month when new data is available, but do not require a constant internet connection like the other reporting tools (if you are connecting to an online DHIS 2 server).

Web-based pivot tables: The built in pivot table tool is a simple web-based tool to display indicator data by orgunit and period in a typical pivot table view and allows for some basic pivoting manipulations of the tables. It is a quick and easy way to look at many indicator values at the same time (by orgunit and/or period), but does not have the same functionality as the offline Excel pivot tables.

GIS: Present and analyse your data using thematic maps. You can view both data elements and indicators and given that you have coordinates for all your orgunits you can drill down the hierarchy and view maps for all levels from country polygons to facility points. See the separate chapter on GIS for more details. All the map information is built into DHIS 2 and all you need to do is to register coordinates for your organisation units and the maps will be available.

21.2. Using standard reports

You access the available reports from the Services drop-down menu, by selecting Reports. In the report menu in the left bar, click Standard Report. A list of all pre-defined reports will appear in the main window.

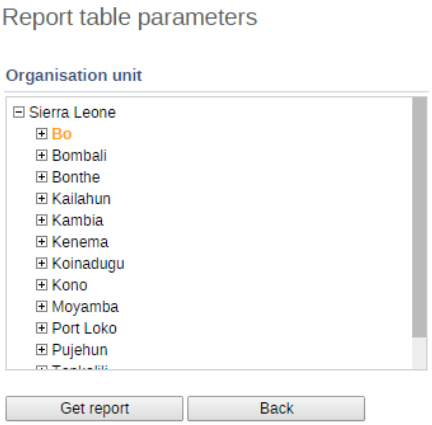


You run/view a report by clicking on the name of the report and then selecting "Create" from the contextual menu. If there are any pre-define paramaters, you will see a report parameter window where you must fill in the values needed for orgunit and/or reporting month, depending on what has been defined in the underlying report table(s). Click on "Get Report" when you are ready. The report will either appear directly in your browser or be available as a PDF file for download, depending on your browser settings for handling PDF files. You can save the file and keep it locally on your computer for later use.

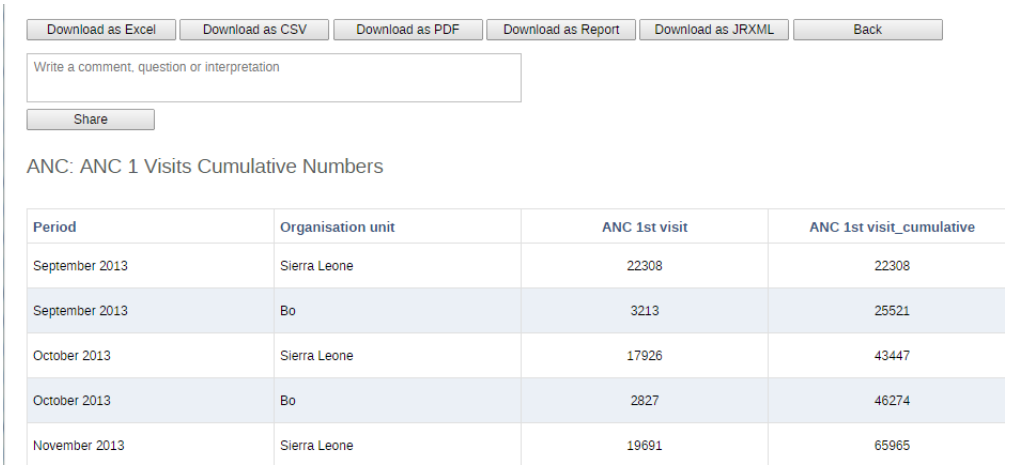
21.3. Using report tables

Report tables are a simple-to-use tool for creating tabular analysis. To run a report table first navigate to the list of available report tables in Services->Reports->Report Tables and then the name of the report table you wish to use. Select "Create" from the contextual menu. If the report table has any pre-defined paramaters, you will need to select them in the next screen. Finally, press "Get report" to view the report table.

Report parameters: Most report tables have parameters, which means that you can filter which orgunits and/or periods you want in the report. This makes the reports much more reusable. When you run the report table a Report parameter window will open and ask the user to input values for the selected parameters. The possible parameters are Reporting Month and Organisation Unit, and either one of these or both will show in the window. After selecting the values click on the Get Report button.



Export/view options: When the report table is ready it will be displayed in a HTML view. The report table can be exported to PDF (for better printing and easier saving), Excel, CSV, and also to a standard report format (Jasper) with a nicer table and a chart shown in PDF, or as a Jasper design file (JRXML) for further improvements and changes to the report design before uploading it as a standard report (see the Creating standard reports section in the Developers Guide for more detail information).



You can also share a comment or intepretation about this report table from the report table view, by simply writing a comment in the box and pressing "Share".

21.4. Using dataset reports

Dataset reports are printer friendly views of the data entry screen filled with either raw or aggregated data. These are only available for data sets that have custom data entry forms and not for default or section forms.

You can access data set reports from the Report menu under Services.

A Criteria window will appear where you fill in the details for your report:

Dataset: The data set you want to display.

Reporting period: The actual period you want data for. This can be aggregated as well as raw periods. This means that you can ask for a quarterly or annual report even though the data set is collected monthly. A data set's period type (collection frequency) is defined in data set maintenance. First select the period type (Monthly, Quarterly, Yearly etc.) in the drop down next to Prev and Next buttons, and then select one of the available periods from the dropdown list below. Use Prev and Next to jump one year back or forward.

Use data for selected unit only: Use this option if you want a report for an orgunit that has children, but only want the data collected directly for this unit and not the data collected by its children. If you want a typical aggregated report for an orgunit you do not want to tick this option.

Reporting Organisation unit: Here you select the orgunit you want the report for. This can be at any level in the hierarchy as the data will be aggregated up to this level automatically (if you do not tick the option above).

When you are done filling in the report criteria you click on "Generate". The report will appear in html view in a printer-friendly format. Use print and save as functions in the browser to print or save (as html) the report.

21.5. Using resources

The resource tool allows you to upload both files from your local computer to the DHIS server and to add links to other resources on the Internet through URLs. If you want to share the direct link to the DHIS resource you can right click on the "view resource" button and copy the link address.

The create a resource click on the "Add new" button. Enter a name for the resource, then choose between uploading a file or external URL. If you chose file upload click "Choose file" and select your file your local computer. If you chose URL enter the link to the resource on the Internet. Then click "Save".

21.6. Using reporting rate summary

Access the reporting rate summary from the Services->Reports menu. Reporting rate summaries will show how many datasets (forms) that have been submitted by organisation unit and period. There are two methods available to calculate reporting rates (completeness):

- Based on complete data set registrations. A complete data set registration refers to a user marking a data entry form as complete, typically by clicking the complete button in the data entry screen, hereby indicating to the system that she considers the form to be complete. This is i.e. a subjective approach to calculating completeness.
- Based on compulsory data element: You can define any number of data elements in a data set to be compulsory. This implies that data values must be captured for all data elements which have been marked as compulsory in order for the data set to be considered complete. This is i.e. an objective approach to calculating completeness.

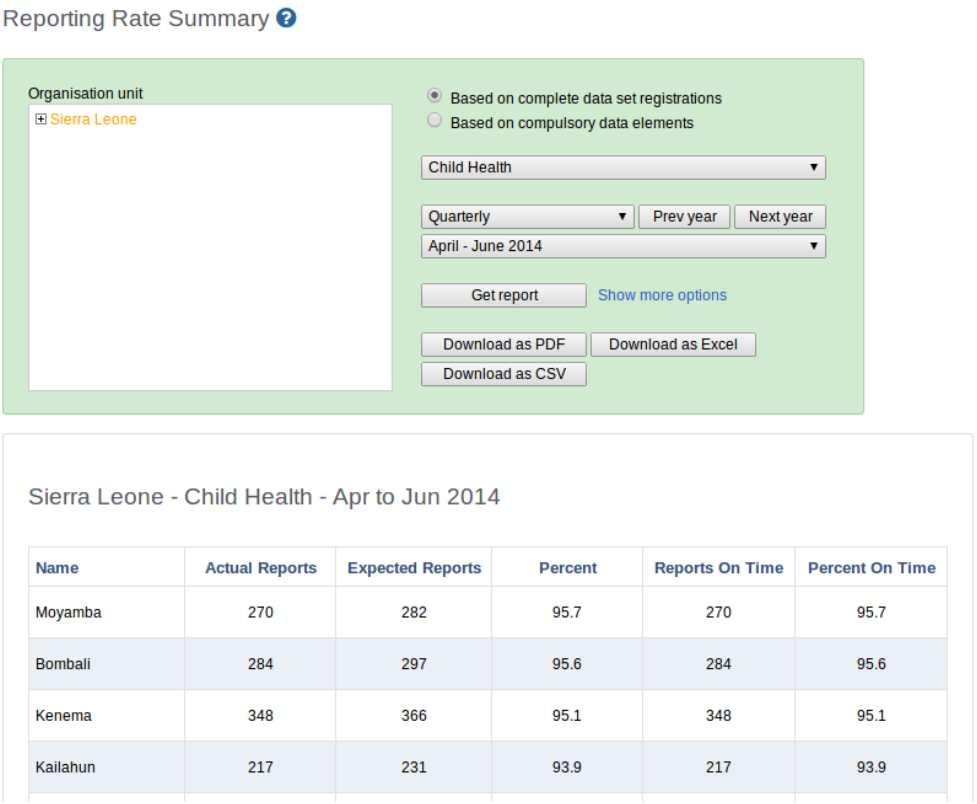
The reporting rate summary will for each row show a range of measures:

- Actual reports: Indicates the number of data entry complete registrations for the relevant data set.
- Expected reports: Indicates how many data entry complete registrations are expected. This number is based on the number of organisation units the relevant data set has been assigned to (enabled for data entry).
- Percent: The percentage of reports registered as complete based on the number expected.
- Reports on time: Same as actual reports, only reports registered as complete within the maximum number of days after the end of the reporting period. This number of days after reporting period can be defined per data set in the data set management.
- Percent on time: Same as percentage, only reports registered as complete on time used as numerator.

To run the report you can follow these steps:

- Select an orgunit from the tree.
- Select one of the completeness methods to use to calcuate the reporting rates.

  Select all or one data set. All will give you a report with all data sets for the selected organisation unit. A single data set will give you a report with completeness for all children of the selected organisation unit.
- Select a period type and a period from the list of available periods for that period type. Move back/forward one year by using the prev/next buttons.
- The report will then be rendered. Change any of the parameters above and the report will be updated automatically.



21.7. Using organisation unit distribution reports

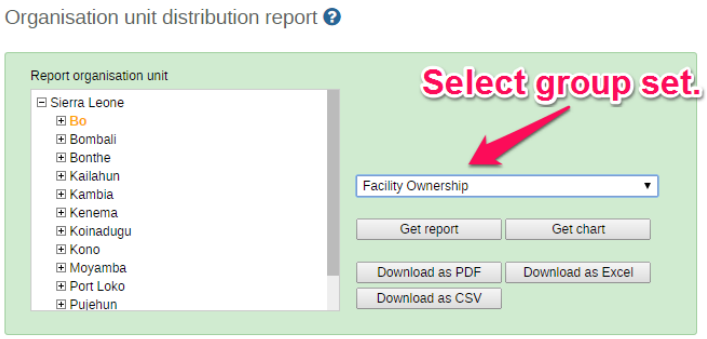
You can access the Orgunit Distribution reports from the left side menu in the Services->Reports module.

Orgunit distribution reports are reports that show how the orgunits are distributed on various properties like type and ownership, and by geographical areas.

The result can be presented in a table-based report or in a chart.

Running a report:

To run a report first select an orgunit in the upper left side orgunit tree. The report will be based on orgunits located under the selected orgunit. The select the orgunit group set that you want to use, typically these are Type, Ownership, Rural/Urban, but can be any user-defined orgunit group set. The you can click on either Get Report to get the table-based presentation or Get chart to get the same result in a chart. You can also download other format such as PDF, Excel and CSV.



Chapter 22. Using Pivot Table

22.1. Pivot Table overview

The pivot table module enables users to create pivot tables, using all available data *dimensions* in DHIS 2. A pivot table is a dynamic tool for data analysis which lets you quickly summarize and arrange data according to its dimensions. Examples of data dimensions in DHIS 2 are data elements (explaining what the data means), periods (representing the time aspect) and the organisational hierarchy (representing the geographical location of the data). From these dimensions you can freely select dimension *items* to include in the pivot table.

A pivot table can arrange data dimensions on *columns*, *rows*, and as *filters*. When you place a data dimension on columns, the pivot table will display one column per dimension item. If you place multiple data dimensions on columns, the pivot table will display one column for all combinations of the items in the selected dimensions. When you place a data dimension on rows, the pivot table will display one row per dimension item in a similar fashion. The dimensions you select as filters will not be included in the pivot table, but will aggregate and filter the table data based on the selected filter items.

The workflow for creating a simple pivot table is:

1. Select dimension items in the left menu, for instance a few data elements.
2. Click "Layout" on the top menu and arrange the data dimensions as columns, rows, and filters. You can leave the selection as it is if desired.
3. Click "Update".

Based on the demo database, a pivot table approximately as below will be displayed. Notice how indicators are listed on columns and periods as rows.

	BCG doses given	Penta1 doses given	Fully Immunized child	Measles doses given	Penta3 doses given	Total
October 2012	50 024	16 253	14 065	15 763	14 106	110 211
November 2012	17 308	17 816	14 812	16 679	16 034	82 649
December 2012	13 432	13 503	10 836	11 798	10 812	60 381
January 2013	20 136	18 738	14 493.2	16 389	14 679	84 435.2
February 2013	20 563	19 539	15 684	18 208	15 899	89 893
March 2013	19 414	19 062	16 163	17 563	15 600	87 802
April 2013	20 295	18 693	15 480	17 422	15 790	87 680
May 2013	22 443	20 870	17 743	19 363	17 191	97 610
June 2013	23 027	21 023	15 714	17 806	16 601	94 171
July 2013	21 309	20 745	15 705	17 063	16 622	91 444
August 2013	20 334	21 366	17 384	19 144	18 247	96 475
September 2013	20 975	20 778	17 793	19 610	18 062	97 218
Total	269 260	228 386	185 872.2	206 808	189 643	1 079 969.2

22.2. Selecting dimension items

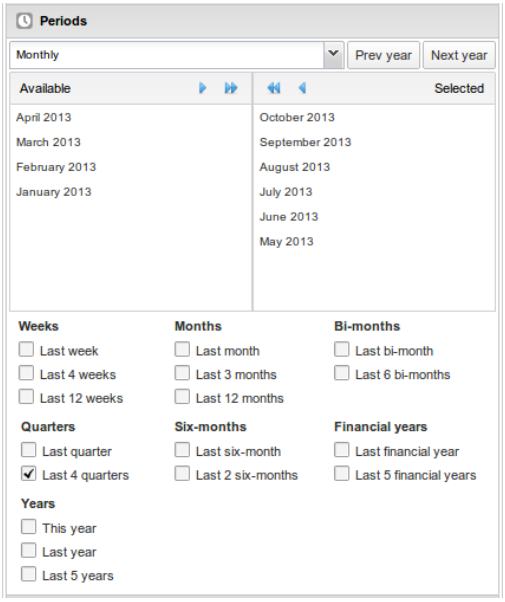
The left menu will list sections for all available data dimensions. From each section you can select any number of dimension items. As an example, you can open the section for data elements and select any number of data elements from the available list. You can select an item by marking it and clicking on the arrow in the section header or simply double-clicking on the item. Before you can use a data dimension in your pivot table you must at least select one dimension item. If you arrange a dimension as columns or rows but do not select any dimension items, the dimension will be ignored.

For the indicator and data element dimensions you must first select one or all groups from the group list. You can then select data elements from the available items list.

For the period dimension you can choose between using fixed periods or relative periods. An example of a fixed period is "January 2012". To select fixed periods start by selecting a period type from the period type list. You can then select periods from the list of available periods. Relative periods are periods relative to the current date. Examples of relative periods are "Last month", "Last 12 months", "Last 5 years". Relative periods can be selected by ticking the checkboxes

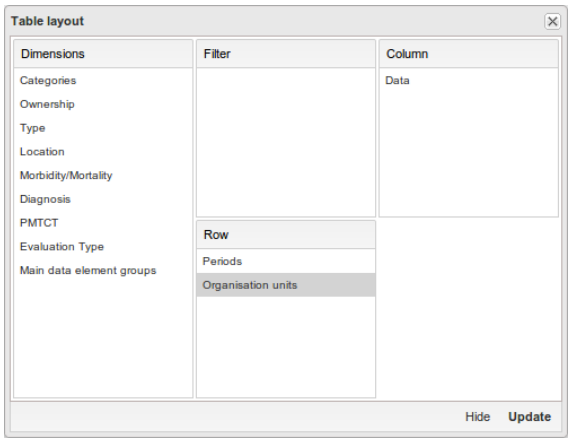
next to each period. The main advantage of using relative periods is that when you save a pivot table favorite, it will stay updated with the latest data as time goes by without the need for constantly updating it.

For the organisation unit dimension you can select any number of organisation units from the hierarchy. To select all organisation units below a specific parent organisation unit, right click and click "Select all children". To manually select multiple organisation units, click and hold the Ctrl button while clicking on organisation units. You can tick "User organisation unit", "User organisation unit children" or "User organisation unit grand children" in order to dynamically insert the organisation unit or units associated with your user account. This is useful when you save a pivot table favorite and want to share it with other users, as the organisation units linked with the other user's account will be used when viewing the favorite.

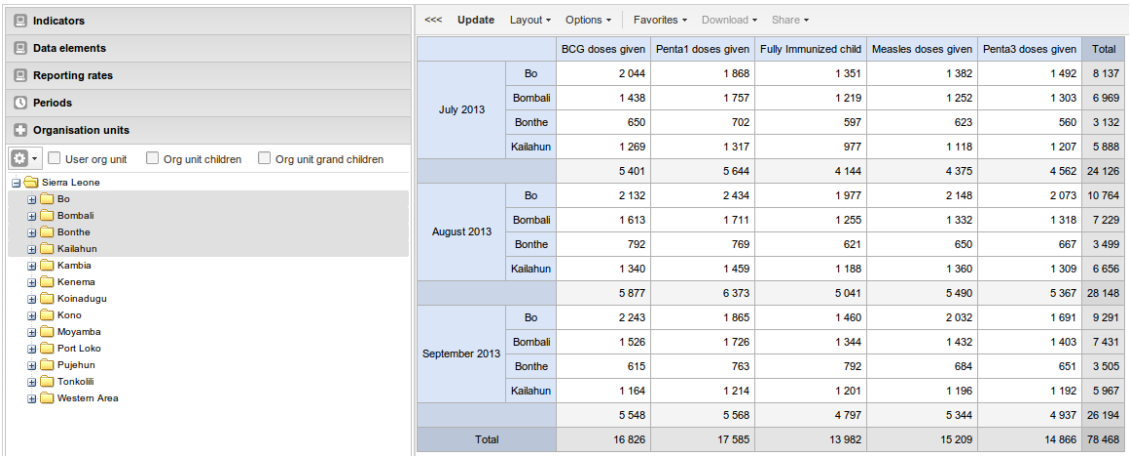


22.3. Arranging the table layout

After selecting data dimensions it is time to arrange your pivot table. Click "Layout" in the top menu to open the layout screen. In this screen you can position your data dimensions as table columns, rows or filters by clicking and dragging the dimensions from the dimensions list to the respective column, row and filter lists. You can set any number of dimensions in any of the lists. For instance, you can click on "Organisation units" and drag it to the row list in order to position the organisation unit dimension as table rows. Note that indicators, data elements and data set reporting rates are part of the common "Data" dimension and will be displayed together in the pivot table. For instance, after selecting indicators and data elements in the left menu, you can drag "Data" from the available dimensions list to the row dimension list in order to arrange them as rows in the pivot table.



After you have set up your pivot table you can click "Update" to render your pivot table, or click "Hide" to hide the layout screen without any changes taking effect. Since we in our example have selected both the period and organisation unit dimension as rows, the pivot table will generate all combinations of the items in these dimensions and produce a table like this:



		BCG doses given	Penta1 doses given	Fully Immunized child	Measles doses given	Penta3 doses given	Total
July 2013	Bo	2 044	1 868	1 351	1 382	1 492	8 137
	Bombali	1 438	1 757	1 219	1 252	1 303	6 969
	Bonthe	650	702	597	623	560	3 132
	Kailahun	1 269	1 317	977	1 118	1 207	5 888
		5 401	5 644	4 144	4 375	4 562	24 126
August 2013	Bo	2 132	2 434	1 977	2 148	2 073	10 764
	Bombali	1 613	1 711	1 255	1 332	1 318	7 229
	Bonthe	792	769	621	650	667	3 499
	Kailahun	1 340	1 459	1 188	1 360	1 309	6 656
		5 877	6 373	5 041	5 490	5 367	28 148
September 2013	Bo	2 243	1 865	1 460	2 032	1 691	9 291
	Bombali	1 526	1 726	1 344	1 432	1 403	7 431
	Bonthe	615	763	792	684	651	3 505
	Kailahun	1 164	1 214	1 201	1 196	1 192	5 967
		5 548	5 568	4 797	5 344	4 937	26 194
Total		16 826	17 585	13 982	15 209	14 866	78 468

22.4. Using table options

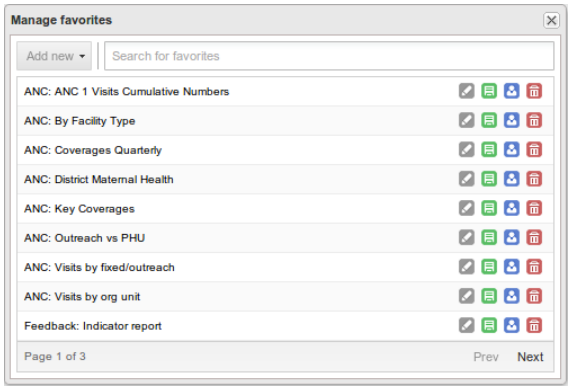
Several table options are available when working with a pivot table. Open the options screen by clicking on "Options" in the top menu. The following options are available:

- Show totals: Display total values in the table for each row and column, as well as a grand total for all values in the table.
- Show sub-totals: Display subtotals in the table for each dimension. In the screenshot above, notice how subtotals are generated for each of the periods in the period dimension. Note that subtotals will be hidden for columns or rows if there is only one selected dimension, as the values in that case are equal to the subtotals.
- Hide empty rows: Hides empty rows from the table, which is useful when looking at large tables where a big part of the dimension items do not have data in order to keep the table more readable.
- Show hierarchy: Shows the name of all ancestors for organisation units, e.g. "Sierra Leone / Bombali / Tamabaka / Sanya CHP" for Sanya CHP. The organisation units are then sorted alphabetically which will order the organisation units perfectly according to the hierarchy.
- Display density: Controls the size of the cells in the table. Can be set to "comfortable", "normal" and "compact". The "compact" option is handy in order to fit large tables into the browser screen.
- Font size: Controls the size of the table text font. Can be set to "large", "normal" and "small".
- Digit group separator: Controls which character to separate groups of digits or "thousands". Can be set to "comma", "space" and "none".
- Legend set: Shows a color indicator next to the values. Currently the GIS legend sets are being used.

22.5. Creating a favorite

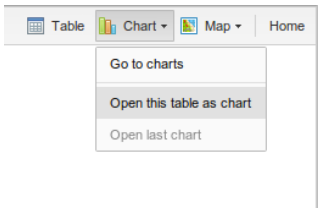
When you have set up a pivot table it is convenient to save it as a favorite. To do so, click "Favorites" on the top menu, click "Add new", give the favorite a descriptive name and click "Create". You can search for favorites through the search input field at the top. To load an existing favorite, simply click the name of the favorite in the list.

To rename a favorite, click the grey "Rename" icon next to the favorite in the list, change the name and click "Update". To overwrite an existing favorite with the current pivot table, click the green "Overwrite" icon. To share a favorite with everyone or a user group, click the blue "Share" icon. To delete a favorite, click the red "Delete" icon.

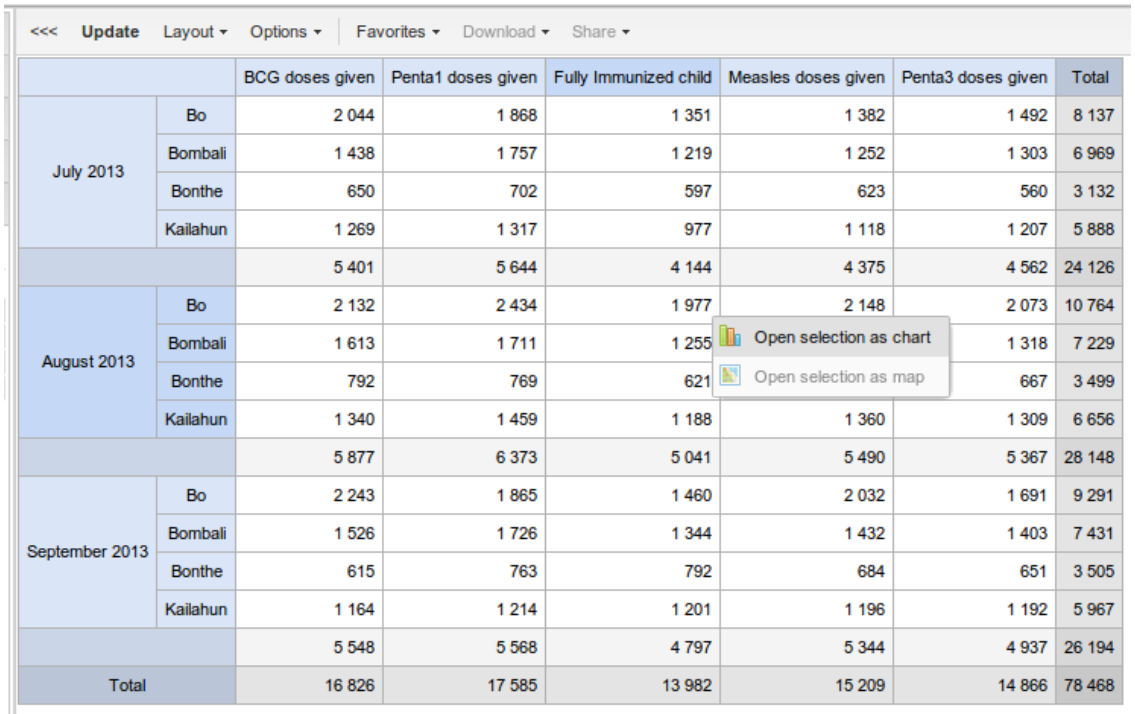


22.6. Analysis integration

The analysis apps in DHIS 2 are completely integrated, so you can easily switch between pivot table, chart and map visualization of your data. When you have made a pivot table you can click e.g. "Chart" in the top right corner and then select "Open this table as chart".



If you just want to visualize a small part of your pivot table as a chart, you can click directly on a value in the table instead. A menu will appear. If you mouse hover the "Open selection as chart" option you can see that some of the dimension headers in the table are highlighted, indicating what data will be visualized as a chart.



		BCG doses given	Penta1 doses given	Fully Immunized child	Measles doses given	Penta3 doses given	Total
July 2013	Bo	2 044	1 868	1 351	1 382	1 492	8 137
	Bombali	1 438	1 757	1 219	1 252	1 303	6 969
	Bonthe	650	702	597	623	560	3 132
	Kailahun	1 269	1 317	977	1 118	1 207	5 888
		5 401	5 644	4 144	4 375	4 562	24 126
August 2013	Bo	2 132	2 434	1 977	2 148	2 073	10 764
	Bombali	1 613	1 711	1 255		1 318	7 229
	Bonthe	792	769	621		667	3 499
	Kailahun	1 340	1 459	1 188	1 360	1 309	6 656
		5 877	6 373	5 041	5 490	5 367	28 148
September 2013	Bo	2 243	1 865	1 460	2 032	1 691	9 291
	Bombali	1 526	1 726	1 344	1 432	1 403	7 431
	Bonthe	615	763	792	684	651	3 505
	Kailahun	1 164	1 214	1 201	1 196	1 192	5 967
		5 548	5 568	4 797	5 344	4 937	26 194
Total		16 826	17 585	13 982	15 209	14 866	78 468

22.7. Downloading data

You can download the data in the current pivot table by clicking on "Download" in the top menu. The data can be downloaded in MS Excel and CSV format. The data table will have one column per dimension and contain names of the dimension items. You can easily create a pivot table in Microsoft Excel from the downloaded Excel file by clicking

on "pivot table" in the top panel, then clicking on "create pivot table", then marking the data range in the spreadsheet before clicking "OK".

Data can also be downloaded in JSON and XML format. The data format is specified in the Web API chapter under the "Analytics" section. The data document will use identifiers of the dimension items and will be opened in a new browser window in order to reveal the URL of the request to the Web API in the address bar. This will be useful for developers of apps and other client modules based on the DHIS 2 Web API.

22.8. Sharing interpretations

For certain analysis-related resources in DHIS, like pivot tables, charts and maps, one can share a data interpretation. An interpretation is simply a link to the relevant resource together with a text expressing some insight about the data. If you want to share a pivot table interpretation you need to first save the table you want to share as a favorite. Then, without making any changes to the table, click the "Share" button the toolbar. A window will open up and this is where you write your interpretation. When you are done, click share button in the bottom right corner of the window. The window will close automatically and if the interpretation was shared successfully you will find a notification on the bottom toolbar.

22.9. Embed tables in any web page

Certain analysis-related resources in DHIS, like pivot tables, charts and maps, can be embeded in any web page by using a plugin. If you have created a table in the Pivot Table app you will get the plugin configuration for this table by clicking the "Share" button the toolbar and then "Embed as plugin". You will find more information about the plugins in the web api chapter.

22.10. Constraints

When selecting and arranging dimensions there are a few constraints that apply. All of these constraints are validated and the pivot table module will provide feedback if any constraint is violated.

- At least one dimension must be selected on columns or rows.
- At least one period must be included in the pivot table.
- Data element group sets and reporting rates cannot appear in the same pivot table.
- A table cannot contain more than 10 000 cells for performance reasons.

Chapter 23. Data approval

DHIS 2 has an optional feature that allows authorized users to approve data that has been entered. It allows data to be reviewed and approved at selected levels in the organisation unit hierarchy, so the approval follows the structure of the hierarchy from lower levels to higher levels.

Data is approved for a combination of (a) data set, (b) period and (c) organisation unit. Data may be approved for the organisation unit for which it is entered, as well as for higher-level organisation units to which the data is aggregated. As part of system settings, you can choose the organisation unit level(s) at which data is approved. It can be approved at higher levels only after it has been approved for all that organisation unit's descendants at lower levels for the same data set and period.

After a data set, period and organisation unit combination has been approved, the form will be locked and any further data entry or modification will be prohibited, unless it is later un-approved.

For example, the following diagram illustrates that data has already been approved for organisation units C and D, for a given data set and period. It may now be approved for organisation unit B for the same data set and period. But it is not ready to be approved for organization unit A. Before it can be approved for organisation unit A, it must be approved for B, and for any other children of organisation unit A, for that data set and period.

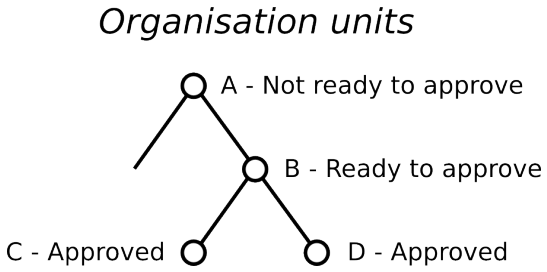


Figure 23.1. Approving at organisation units

23.1. Approving and accepting

DHIS2 supports two different types of approval processes: either a one-step process where the data is approved at each level, or a two-step process where data is first approved and then accepted at each level. This is illustrated in the following diagram:

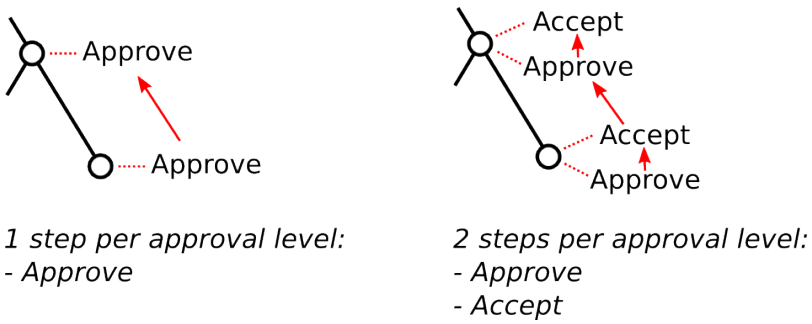


Figure 23.2. Approving and accepting

In the one-step process, data is approved at one level, and then approved at the next higher level. Until it is approved at the next higher level, it may be unapproved at the first level. (For example, if the data was approved my mistake, this allows the approver to undo their mistake.) Once the data is approved at the next higher level, it may not be unapproved at the lower level unless it is first unapproved at the higher level.

In the two-step process, data is approved at one level, and then the approval is accepted at the same level. This acceptance is done by a user who is authorized to approve data at the next higher level. Once the data is accepted, it may not be changed or unapproved unless it is first *unaccepted*.

The two-step process is not required by DHIS2. It is an optional step for a user reviewing data at the next higher level. It has the benefit of locking the acceptance from the level below, so reviewer does not have to worry that the data could be changing from below while it is being reviewed. It can also be used by the higher-level user to keep track of which lower-level data has already been reviewed.

23.2. Authorities for approving data

To approve data, you must be assigned a role containing one of these authorities:

- **Approve data** - You may approve data for the organisation unit(s) to which you are assigned. Note that this authority does not allow you to approve data for lower-levels below the organisation unit(s) to which you are assigned. This is useful to separate the users authorized to approve at one level from the users authorized to approve at levels below.
- **Approve data at lower levels** - Allows you to approve data for all lower levels below the organisation units assigned to you. This is useful if, for example, you are a district-level user whose role includes approving the data for all the facilities within that district, but not for the district itself. If you are assigned this as well as the *Approve data* authority, you may approve data at the level of the organisation unit(s) to which you have been assigned, and for any level below.
- **Accept data at lower levels** - Allows you to accept data for the level just below the organisation unit(s) assigned to you. This authority can be given to the same users as approve data. Or it may be given to different users, if you want to have some users who accept data from the level below, and a different set of users who approve data to go up to the next level above.

23.3. Configure approving data

When you add or edit a data set, you specify whether or not data entered for that data set will be available for approval. This is done with the *Approve data* option when adding or editing a data set.

You can configure the organisation unit levels at which you want to approve data in these data sets. This is done in the *Settings* section, under *System Approval Settings*. Click the Add new button on this page and select the organisation unit level at which you want approvals. It will be added to the list of approval settings. You may configure the system for approving data at every organisation unit level, or only at selected organisation unit levels.

Note that when you add a new approval level, you may optionally choose a Category option group set. This feature is discussed later in this chapter.

On the *System Approval Settings* page, you may select the option *Hide unapproved data in analytics* to hide unapproved data in reports, pivot table, data visualizer and GIS. If this option is checked, unapproved data will be hidden from users assigned to higher-level organisation units than where approval is needed. Users who are assigned to organisation units where data is ready for approval can still view the data, as can users assigned to higher-level organisation units if they have the *Approve data at lower levels* authority. If this option is not checked, then all data is shown whether approved or not.

23.4. Approving data

To approve data, go to *Reports* and choose *Data Approval*. When this report shows data that is configured for approval, it shows the approval status of the data in the report. The approval status will be one of the following:

- **Waiting for lower level org units to approve** - This data is not yet ready to be approved, because it first needs to be approved for all the child organisation units to this organisation unit, for the same data set and period.
- **Ready for approval** - This data may now be approved by an authorized user.
- **Approved** - This data has already been approved.
- **Approved and accepted** - This data has already been approved, and also accepted.

If the data you are viewing is in an approval state that can be acted upon, and if you have sufficient authority, one or more of the following actions will be available to you on the *Data Approval* form:

- **Approve** - Approve data that has not yet been approved, or that was formerly approved and has been unapproved.
- **Unapprove** - Return to an unapproved state data that has been approved or accepted.
- **Accept** - Accept data that has been approved.
- **Unaccept** - Return to an unaccepted (but still approved) state data that has been accepted.

In order to unapprove data for a given organisation unit, you must have the authority to approve data for that organisation unit or to approve data for a higher-level organisation unit to which that data is aggregated. The reason for this is as follows: If you are reviewing data for approval at a higher organisation unit level, you should consider whether the data at lower organisation units are reasonable. If all lower-level data looks good, you can approve the data at the higher level. If some lower-level data looks suspect, you can unapprove the data at the lower level. This allows the data to be reviewed again at the lower level, corrected if necessary, and re-approved up through the organisation unit levels according to the hierarchy.

23.5. Approving by category option group set

When defining an approval level, you specify the organisation unit level at which data will be approved. You may also optionally specify a category option group set. This is useful if you are using category option groups to define additional dimensions of your data, and you want approvals to be based on these dimensions. The following examples illustrate how this can be done within a single category option group set, and by using multiple category option group sets.

23.5.1. Approving by one category option group set

For example, suppose you define a category option group set to represent NGOs who serve as healthcare partners at one or more organisation units. Each category option group within this set represents a different partner. The category option group for Partner 1 may group together category options (such as funding account codes) that are used by that partner as a dimension of the data. So data entered by Partner 1 is attributed to a category option in Partner 1's category option group. Whereas data entered by partner 2 is attributed to a category option in Partner 2's category option group:

Table 23.1. Example Category Option Groups

Category option group set	Category option group	Category options
Partner	Partner 1	Account 1A, Account 1B
Partner	Partner 2	Account 2A, Account 2B

Each partner could enter data for their accounts independently of the other, for the same or different data sets, at the same or different facilities. So for example, data can be entered and/or aggregated at the following levels for each partner, independently of each other:

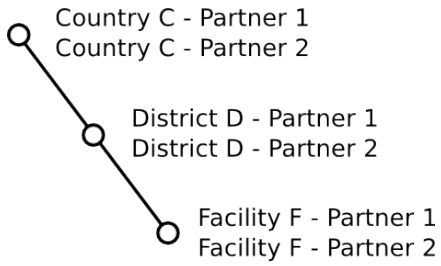


Figure 23.3. Example category option groups

Tip

You can use the sharing feature on category options and category option groups to insure that a user can enter data (and/or see data) only for certain category options and groups. If you don't want users to see data that is aggregated beyond of their assigned category options and/or category option groups, you can assign *Selected dimension restrictions for data analysis*, when adding or updating a user.

You can optionally define approval levels for partner data within any or all of these organisation unit levels. For example, you could define any or all of the following approval levels:

Table 23.2. Example Category Option Group Set approval levels

Approval level	Organisation unit level	Category option group set
1	Country	Partner
2	District	Partner
3	Facility	Partner

23.5.2. Approving by multiple category option group sets

You can also define approval levels for different category option group sets. To continue the example, suppose that you have various agencies that manage the funding to the different partners. For example, Agency A funds accounts 1A and 2A, while Agency B funds accounts 1B and 2B. You could set up category option groups for Agency A, and Agency B, and make them both part of a category option group set called Agency. So you would have:

Table 23.3. Example Multiple Category Option Group Sets

Category option group set	Category option group	Category options
Partner	Partner 1	Account 1A, Account 1B
Partner	Partner 2	Account 2A, Account 2B
Agency	Agency A	Account 1A, Account 2A
Agency	Agency B	Account 1B, Account 2B

Now suppose that at the country level, you want each partner to approve the data entered by that partner. Once this approval is done, you want each agency to then approve the data from accounts that are managed by that agency. Finally, you want to approve data at the country level across all agencies. You could do this by defining the following approval levels:

Table 23.4. Example Multiple Category Option Group Set approval levels

Approval level	Organisation unit level	Category option group set
1	Country	
2	Country	Agency
3	Country	Partner

Note that multiple approval levels can be defined for the same organisation unit level. In our example, Partner 1 would approve country-wide data at approval level 3 from category options Account 1A and Account 1B. Next, Agency A would approve country-wide data at approval level 2 from category options Account 1A (after approval by Partner 1) and Account 2A (after approval by Partner 2.) Finally, after approval from all agencies, country-wide data can be approved at approval level 1 across all category options. Note that approval level 1 does not specify a category option group set, meaning that it is for approving data across all category options.

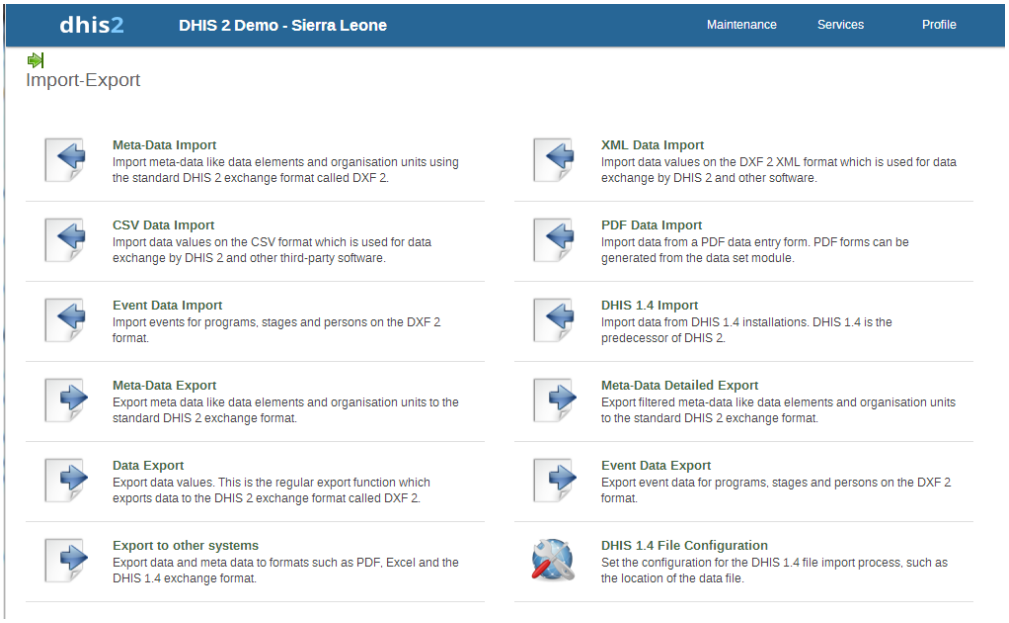
This example is meant to be illustrative only. You may define as many category option groups as you need, and as many approval levels as you need at the same organisation unit level for different category option group sets.

If you have multiple approval levels for different category option group sets at the same organisation unit level, you may change the approval ordering in the *Settings* section, under *System Approval Settings*. Just click on the approval level you wish to move, and select *Move up* or *Move down*. If you have an approval level with no category option groups set, it must be the highest approval level for that organisation unit level.

Chapter 24. Import and export

In a primary health system, the HMIS typically involves a distributed application, where the same application is running in different geographical locations (PHCs,CHCs, hospitals, districts, and state). Most of these physical locations do not have Internet connectivity, and hence they work off-line. At some point (normally at the district level), the data needs to be synchronised in order to have a consolidated database for the a particular geographical region. For this, it is important to be able to export data from one location (which is working offline, say at the health facility level) to another one say at the district level where the data would need to be imported. This feature of exporting and importing is thus a crucial function of a HMIS. This feature also helps us overcome the dependency on Internet to some degree, as data updates can be transferred via USB key where there is no connectivity, or through email where there is limited Internet connectivity. DHIS 2 provides robust export-import functionality to fulfill these needs.

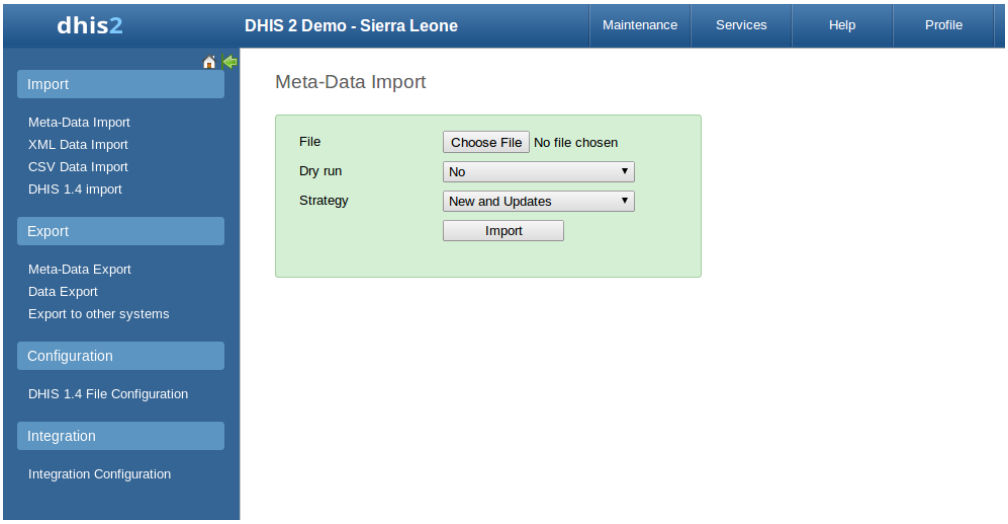
To access the main Import-Export module, choose Services->Import-Export. A number of services are available, all of which will be described in detail in respective sections below.



24.1. Meta-data import

24.1.1. XML Meta-data import

Meta-data in the XML format known as DXF2 can be easily imported by accessing Services->Import-Export->Meta-Data import. Select the XML file by pressing "Select" and choose the file from your local file system. Both raw XML as well as a ZIP archive of the XML are supported.



There are two separate options for importing data.

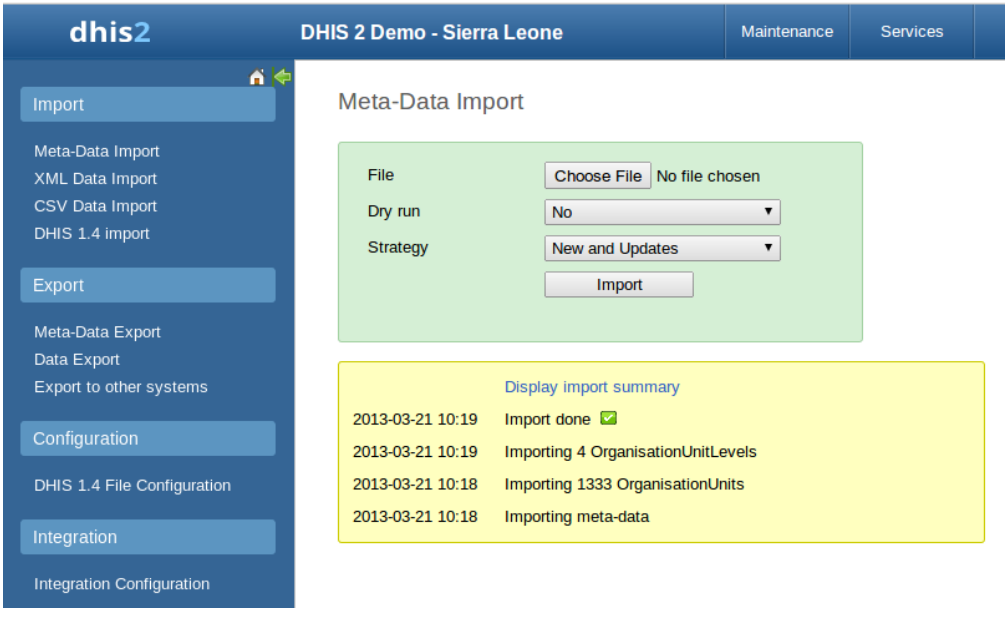
- Dry run: This is similar to the old preview option, this will do a dry run import, and give you information about any errors.
- Strategy: There are three options here, "New and Updates", "New only", and "Update only". New and updates tells the importer to expect both new meta-data, and updates to existing ones. New only will only accept fresh meta-data, usually you would use this on a empty database. Updates only will only allow the meta-data to match meta-data that is already there (same identifiers, new name etc).

Note

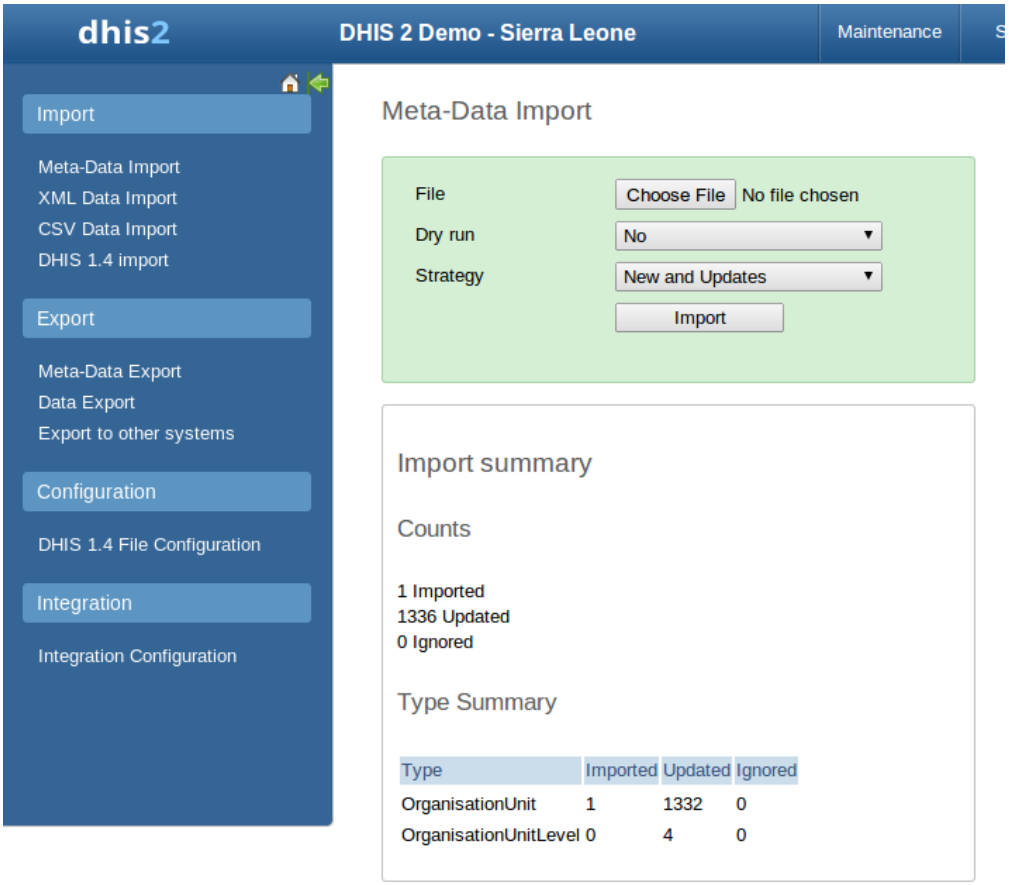
It is highly recommend always using the Dry run option when importing data to make sure you keep control over any changes to your meta-data and databases being out of synch on data elements or orgunit names

24.1.1.1. Dry run before importing

Before doing the import into your database, it is highly recommended that you run the import with the dry run option set to true first. This will enabled you to have a look at how many new, updates, and ignored meta-data there will be. After you have selected your file, set dry run to true, you can now click the import button. After doing this, you will be greeted by this window.

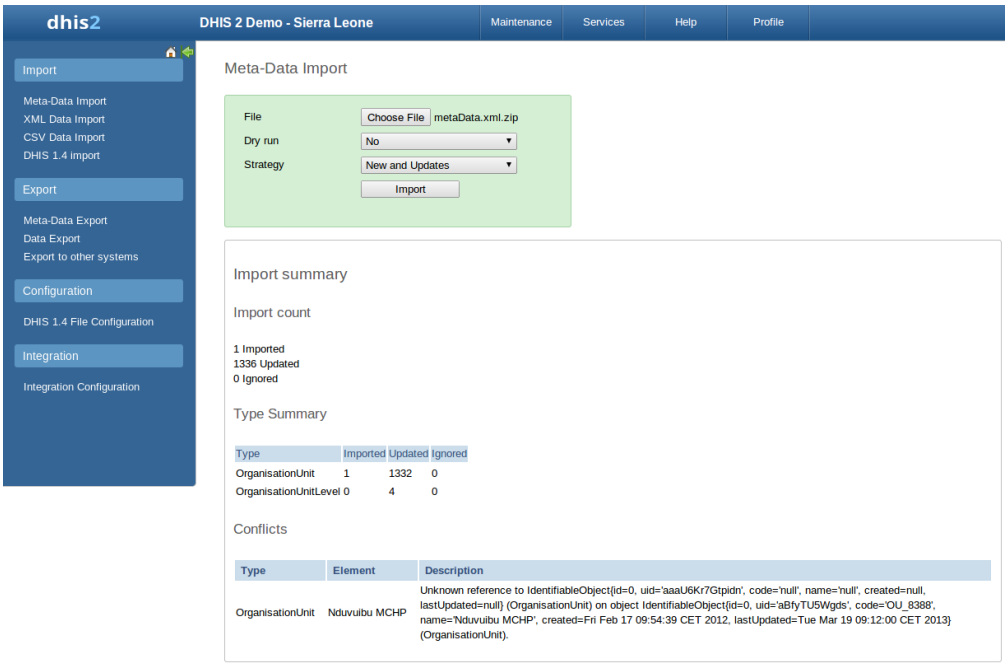


Here you can see a short summary of what was contained in your import file. To see further details, please click on the "display import summary" link.



Here you can see that the import dry run was successful and the import contained **1** new organisation unit, **1332** updated, and **4** organisation unit levels.

Below you can see another example.



Here you can see that the organisation unit Nduvuibu MCHP had a unknown reference to an object with ID "aaaU6Kr7Gtpidn", which means that an object with this ID was not present in your imported file, and it could not be found in the existing database. Its important to note that even if a reference could not be found, the object will still be imported, but you might have to fix this directly in DHIS 2 later, if the reference should have pointed to something else.

24.1.1.2. Matching identifiers in DXF2

The DXF2 format currently support matching for two identifiers, the internal DHIS2 identifier (known as a UID), and also using an external identifier called called a "code". When the importer is trying to search for references (like the one above), it will first go to the UID field, and then to the code field. This allows you to import from legacy systems without having a UID for every meta-data object. I.e. if you are importing facility data from a legacy system, you can leave out the ID field completely (DHIS 2 will fill this in for you) and the put the legacy systems own identifiers in the code field, this identifier is required to be unique. This not only works for organisation units, but for all kinds of meta-data, allowing for easy import from other systems.

24.1.2. Importing CSV meta-data

DHIS 2 supports import of meta-data in the CSV format. Columns which are not required can be omitted in the CSV file, but the order will be affected. If you would like to specify columns which appear late in the order but not specify columns which appear early in the order you can include empty columns ("") for them.

To import CV meta-data go to import-export module and select CSV Meta-Data Import form the left side menu. You must select the object type which your CSV file contains. You can only upload one type of objects at the time. Upload your file and click update. The following object types are supported:

- Data elements
- Data element groups
- Category options
- Category option groups
- Organisation units
- Organisation unit groups
- Validation rules
- Option sets

The formats for the currently supported object types for CSV import are listed below.

Table 24.1. Data Element CSV Format

Column	Required	Value (default first)	Description
Name	Yes		Name. Max 230 char. Unique.
UID	No	UID	Stable identifier. Max 11 char. Will be generated by system if not specified.
Code	No		Stable code. Max 50 char.
Short name	No	50 first char of name	Will fall back to first 50 characters of name if unspecified. Max 50 char. Unique.
Description	No		Free text description.
Form name	No		Max 230 char.
Domain type	No	aggregate | tracker	Domain type for data element, can be aggregate or tracker. Max 16 char.
Value type	No	int | string | bool | trueOnly | date | unitInterval	Value type. Max 16 char.
Number type	No	int | posInt | negInt | number | zeroPositiveInt	Only relevant if type is int. Max 16 char.
Text type	No	text | longText	Only relevant if type is string. Max 16 char.
Aggregation operator	No	sum | average | count | stddev | variance	Operator indicating how to aggregate data in the time dimension. Max 16 char.

Column	Required	Value (default first)	Description
Category combination UID	No	UID	Stable identifier. Max 11 char. Will default to default category combination if not specified.
Url	No		URL to data element resource. Max 255 char.
Zero is significant	No	false | true	Indicates whether zero values will be stored for this data element.

Table 24.2. Organisation Unit CSV Format

Column	Required	Value (default first)	Description
Name	Yes		Name. Max 230 characters. Unique.
UID	No	UID	Stable identifier. Max 11 char. Will be generated by system if not specified.
Code	No		Stable code. Max 50 char.
Parent UID	No	UID	Stable identifier. Max 11 char.
Short name	No	50 first char of name	Will fall back to first 50 characters of name if unspecified. Max 50 characters. Unique.
Description	No		Free text description.
UUID	No		UUID. Max 36 char.
Opening date	No	1970-01-01	Opening date of organisation unit in YYYY-MM-DD format.
Closed date	No		Closed date of organisation unit in YYYY-MM-DD format, skip if currently open.
Comment	No		Free text comment for organisation unit.
Feature type	No		Can be Point, Polygon, MultiPolygon. Max 50 char.
Coordinates	No		Coordinates used for geospatial analysis in Geo JSON format.
URL	No		URL to organisation unit resource. Max 255 char.
Contact person	No		Contact person for organisation unit. Max 255 char.
Address	No		Address for organisation unit. Max 255 char.
Email	No		Email for organisation unit. Max 150 char.
Phone number	No		Phone number for organisation unit. Max 150 char.

Table 24.3. Validation Rule CSV Format

Column	Required	Value (default first)	Description
Name	Yes		Name. Max 230 characters. Unique.
UID	No	UID	Stable identifier. Max 11 char. Will be generated by system if not specified.
Code	No		Stable code. Max 50
Description	No		Free text description.
Instruction	No		Free text instruction.
Importance	No	medium | high | low	
Rule type	No	validation | surveillance	

Column	Required	Value (default first)	Description
Operator	No	equal_to | not_equal_to | greater_than | greater_than_or_equal_to | less_than | less_than_or_equal_to | compulsory_pair	
Period type	No	Monthly | Daily | Weekly | Quarterly | SixMontly | Yearly	
Left side expression	Yes		Mathematical formula based on data element and option combo UIDs.
Left side expression description	Yes		Free text.
Left side null if blank	No	false | true	Boolean.
Right side expression	Yes		Mathematical formula based on data element and option combo UIDs.
Right side expression description	Yes		Free text.
Right side null if blank	No	false | true	Boolean.

Table 24.4. Option Set CSV Format

Column	Required	Value (default first)	Description
Name	Yes		Name. Max 230 characters. Unique. The option set values should be repeated for each option.
UID	No	UID	Stable identifier. Max 11 char. Will be generated by system if not specified.
Code	No		Stable code. Max 50 char.
Option	Yes		Option. Free text. The option set values should be repeated for each option.

Table 24.5. Data Element Group, Category Option, Category Option Group, Organisation Unit Group CSV Format

Column	Required	Value (default first)	Description
Name	Yes		Name. Max 230 characters. Unique.
UID	No	UID	Stable identifier. Max 11 char. Will be generated by system if not specified.
Code	No		Stable code. Max 50 char.

An example of a CSV file for data elements can be seen below. The first row will always be ignored. Notice how you can skip columns and rely on default values or simply leave columns blank:

```
name,uid,code,shortname,description,formname,domaintype,type,numbertype,texttype,aggregationopera
"Women participated in skill development training",,"D0001","Women participated
 development training"
```

```
"Women participated in community organizations",,"D0002","Women participated community
 organizations"
```

A minimal example for importing organisation units with a parent unit looks like this:

```
name,uid,code,parent
"West province",,"WESTP","ImspTQPwCqd"
"East province",,"EASTP","ImspTQPwCqd"
```

The format for option sets is special. One record represents an option, and the three first values representing the option set should be repeated for each option (record):

```
name,uid,code,option
"Color",,,"Blue"
"Color",,,"Green"
"Gender",,,"Female"
"Gender",,,"Male"
```

24.1.3. GML data import

The GML import function can be used to import data prepared in the Geography Markup Language (GML). GML can be used to update the coordinates (both polygons and points). Once you have prepared your GML file as detailed in the chapter on "Importing coordinates", you can load the file with this function.

It is currently not possible to import an organisation unit hierarchy with GML. Therefore, you should create the organisation unit hierarchy separately, and then use GML to update the coordinates once the hierarchy has been created.

Note

Currently, it is only possible to import GML data by matching on the name of the organisation unit itself. Therefore, if you have organisation units with the same name in your organisation unit hierarchy (e.g. two clinics with exactly the same name), you will need to distinguish them before importing GML. Otherwise, both organisation units may be updated with the same set of coordinates.

24.2. Importing data

The import option allows different instance of DHIS 2 to receive standardised sets of data in the absence of a networked system. The functionality can also be used to import data produced by another system (perhaps on a regular basis) or to import legacy data which has been transformed into a format which DHIS2 can understand. Typically, a data set is exported from one DHIS2 instance (e.g. a district level system) to another system (e.g. a provincial level system). DHIS 2 is also capable of importing data directly from a DHIS 1.4 Access database. Each of these options will be discussed in the following sections.

24.2.1. XML data import

To import data in XML format, simply select Services->Import-export->XML data import. Similar options to the XML meta-data import facility are available. Refer to the section in the Web API on XML data import for details on the XML format which is used by DHIS2 for aggregate data import.

24.2.2. Importing CSV data

DHIS2 supports import of data in the CSV (Comma Separated Values) format. This can be used to import exchange file produced by DHIS2 itself. It also comes in handy when you want to import data from a third-party system as CSV is widely supported in applications and is easy to produce manually.

To import a CSV data exchange file navigate to the *CSV Data Import* item in the left-side menu. Upload the exchange file and click *Import*.

The following section describes the CSV format used in DHIS2. The first row is assumed to be a header row and will be ignored during import.

Table 24.6. CSV format of DHIS 2

Column	Required	Description
Data element	Yes	Refers to ID by default, can also be name and code based on selected id scheme
Period	Yes	In ISO format
Org unit	Yes	Refers to ID by default, can also be name and code based on selected id scheme
Category option combo	No	Refers to ID
Attribute option combo	No	Refers to ID (from version 2.16)
Value	No	Data value
Stored by	No	Refers to username of user who entered the value
Last updated	No	Date in ISO format
Comment	No	Free text comment
Follow up	No	true or false

The following is an example CSV file which can be imported into DHIS 2. It can be imported both as plain text file or as compressed ZIP file archive.

```
"dataelement","period","orgunit","categoryoptioncombo","attributeoptioncombo","value","storedby",
"DUSpd8Jq3M7","201202","gP6hn503KUX","Prlt0C1RF0s",,"7","bombali","2010-04-17",,"false"
"DUSpd8Jq3M7","201202","gP6hn503KUX","V6L425pT3A0",,"10","bombali","2010-04-17",,"false"
"DUSpd8Jq3M7","201202","OjTS752GbZE","V6L425pT3A0",,"9","bombali","2010-04-06",,"false"
```

24.2.3. Event data import

Event data can also be easily imported with the "Event data import" function. Refer to the section in the WebAPI for more information on the format used for importing events. Otherwise, the functionality is the same as the XML data import used for aggregate data.

24.2.4. Importing PDF data

DHIS2 supports import of data in the PDF format. This can be used to import data produced by off-line PDF data entry forms. Please refer to the section "Data set management" for details on how to produce a PDF form which can be used for off-line data entry.

To import a PDF data file, navigate to the *PDF Data Import* item in the left-side menu. Upload the completed PDF file and click *Import*.

After the import process is finished, you can follow the link to the import summary, which will inform you about the outcome of the import process in terms of number of records imported, updated and ignored and potential conflicts.

24.2.5. Importing data from DHIS 1.4

There are two ways to import data from a DHIS 1.4 database; 1) through the 1.4 XML-based export files, or 2) directly from the DHIS 1.4 data file (.mdb). Both are accessible from the DHIS 1.4 Import menu under Import in the Import-Export module.

It is critical that all data integrity violations which are present in the DHIS 1.4 database be fully resolved before attempting an import into DHIS2. You can check the data integrity of DHIS 1.4 through the CORE Module->Advanced->Data integrity checks. A report will be generated of all data integrity violations which should be resolved prior to importing into DHIS2.

Warning

When data is imported from DHIS 1.4, both the meta-data as well as data are imported. You should therefore be exceedingly careful that the meta-data present in DHIS 1.4 is compatible with your DHIS2 data, otherwise during a DHIS 1.4 data import, the meta-data in the DHIS2 system will be potentially overwritten by the i

24.2.5.1. DHIS 1.4 File (database) Import

This method is recommend when doing large imports from 1.4, and especially when importing into a new DHIS2 database.

DHIS 1.4 File Configuration

Before you can start the 1.4 file import you need to provide a few details about the 1.4 database:

Datafile(#): Here you put the full path to the DHIS 1.4 data file you want to import from, e.g. C:\DHIS14\DHIS_#LR_LIBERIA.mdb.

Username: Leave blank (unless you have set up extra security measures on the file)

Password: Leave blank (unless you have set up extra security measures on the file)

Levels: Provide the number of levels in the orgunit hierarchy in your 1.4 database, e.g. 5.

Click "Save" and you will return to the DHIS 1.4 File Import window.

Import Type:

As with other imports you have the options to Import (directly), Preview, or Analyse the import. We recommend using the Analyse option first to check that the 1.4 database is OK and ready to be imported.

When importing a large database into a new blank DHIS 2 database we recommend using the Import option to save time.

For smaller incremental imports the Preview is OK.

Last Updated:

If you want the full import, all the data in the 1.4 database you leave this field blank.

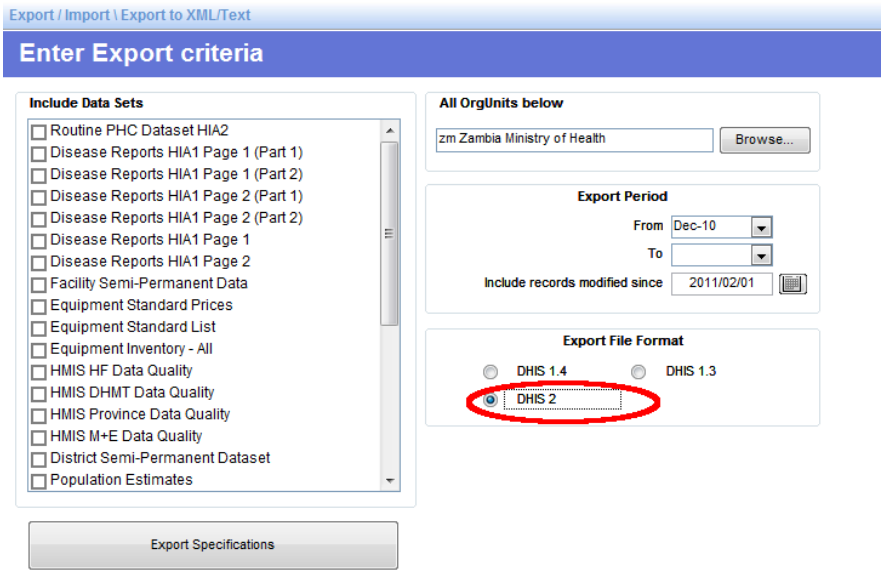
If you only want to do an incremental import into an already existing DHIS 2 database you can use this field to filter which data values to import. Only values added or edited after the date you specify will be imported. This filter makes use of the LastUpdated column in the RoutineData table in the DHIS 1.4 data file.

Import process:

When you are done selecting Method, and LastUpdated you can begin the import by clicking on the Import button. This process might take a long time, depending on how many data values you are importing. On a reasonable spec. computer the import takes about 2 million records per 30 minutes.

24.2.5.2. DHIS 1.4 XML Import

Import though XML data from DHIS 1.4 is also possible using the standard DHIS 1.4 export format. Just be sure that the DHIS 1.4 export format has been set to "DHIS 2" as illustrated in the screen shot below. After the data has been exported by DHIS 1.4, you can import the data by choosing "Services->Import-Export->DHIS 1.4 Import->DHIS 1.4 XML Import" and proceeding via the procedure outline in the previous section.



24.2.5.3. Limitations to DHIS 1.4 imports

Although it is possible to import and export data between instances of DHIS 1.4 and DHIS 2, there are significant limitations. Currently, the import of some metadata is not supported from DHIS 1.4 to DHIS 2. This includes:

- Organisational unit alternate names
- Compulsory data element pairings
- Custom data entry forms
- Dataset data entry levels

It is also important that the aggregation operator defined in DHIS 1.4 be set to the correct value. Some data, such as population, should have their aggregation operator set to "Average" in DHIS2, as this controls how the aggregation of data is handled over time (but not within the organisational unit hierarchy).

24.3. Exporting data and meta-data

DHIS2 makes it possible to export various types of data in multiple data formats. You can export your data, also referred to as measures or facts; and your meta-data, which refers to the information describing your data. Data can be exported both in raw, non-aggregated format and in aggregated format. It is also possible to export a combination of data and meta-data in case you have special requirements. This chapter covers mainly how to export data and meta-data through the user interface of the import-export DHIS 2 module. Data can also be exported programmatically through the DHIS 2 Web API, and it is useful to be aware of the various options available:

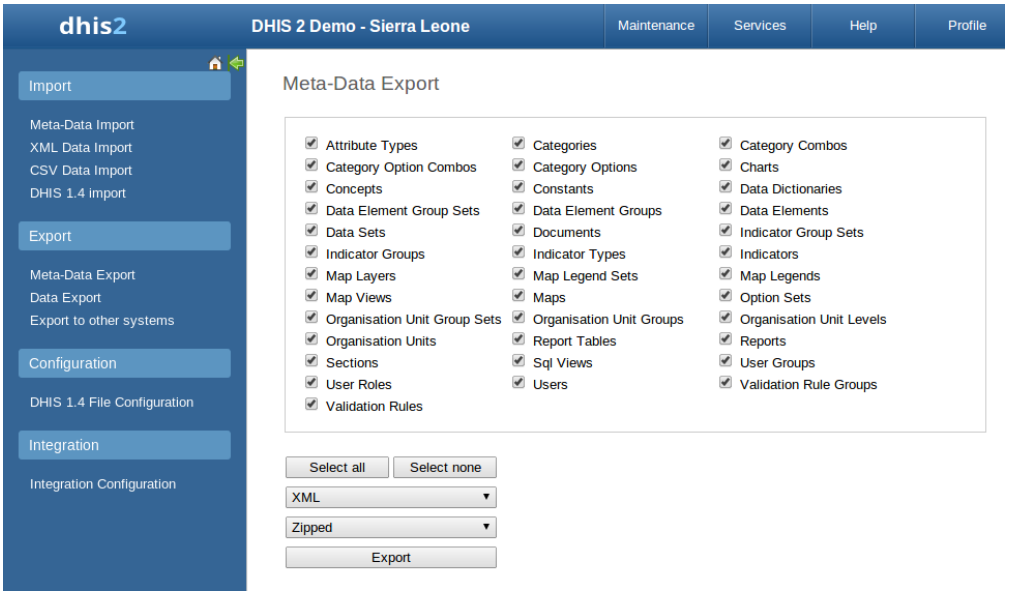
- Export raw, non-aggregated data through the user interface: Covered in this chapter.
- Export aggregated data programmatically through the Web API: Please refer to the Web API chapter, section on Analytics.
- Export a combination of data and meta-data: Please refer to the SQL view sections in the Data administration chapter and Web API chapter.
- Export meta-data through the user interface: Covered in this chapter.
- Export meta-data programmatically through the Web API: Please refer to the Web API chapter, section on meta data.

Data can be exported on various formats, including DXF 2 (the DHIS 2 meta-data and data exchange format based on XML), CSV, PDF, MS Excel and the DHIS 1.4 XML format.

Another aspect of data export is the type of DHIS 2 deployment. In the case of online deployment, all data is saved into a single database only. In an offline deployment, each instance will store data in a separate database in their local system. In an offline deployment, after the data entry is finished, data will have to be manually sent to the next level in the organizational hierarchy. In an on-line application this is not required, as all data is captured over an Internet connection and stored in a central location.

24.3.1. Metadata export

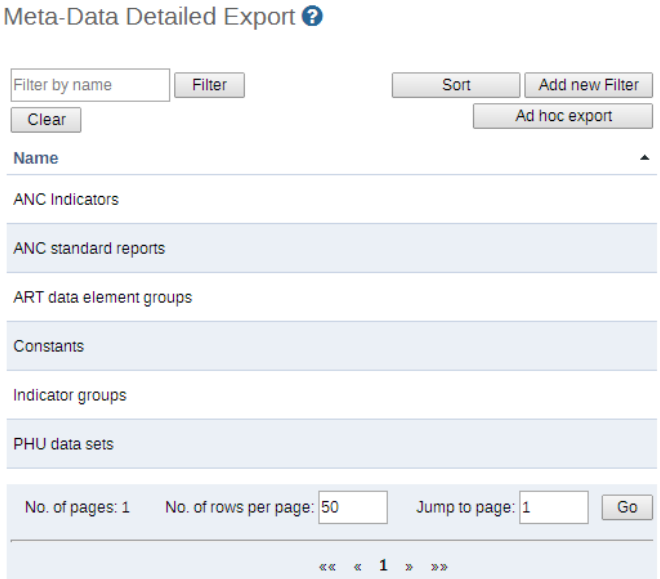
Meta-data is "data about data". In the context of DHIS 2, meta-data consists of definitions of data elements, indicators, the structure and names contained in the organizational hierarchy, and other options. Click on the "Meta-data export" link from the main "Data export" screen in order to access this. Just select the features, format, and the compression that you want and click "Export". This metadata file can then be transmitted just like a data file, except it will contain information on the definitions of the various features, as opposed to the values of the data themselves.



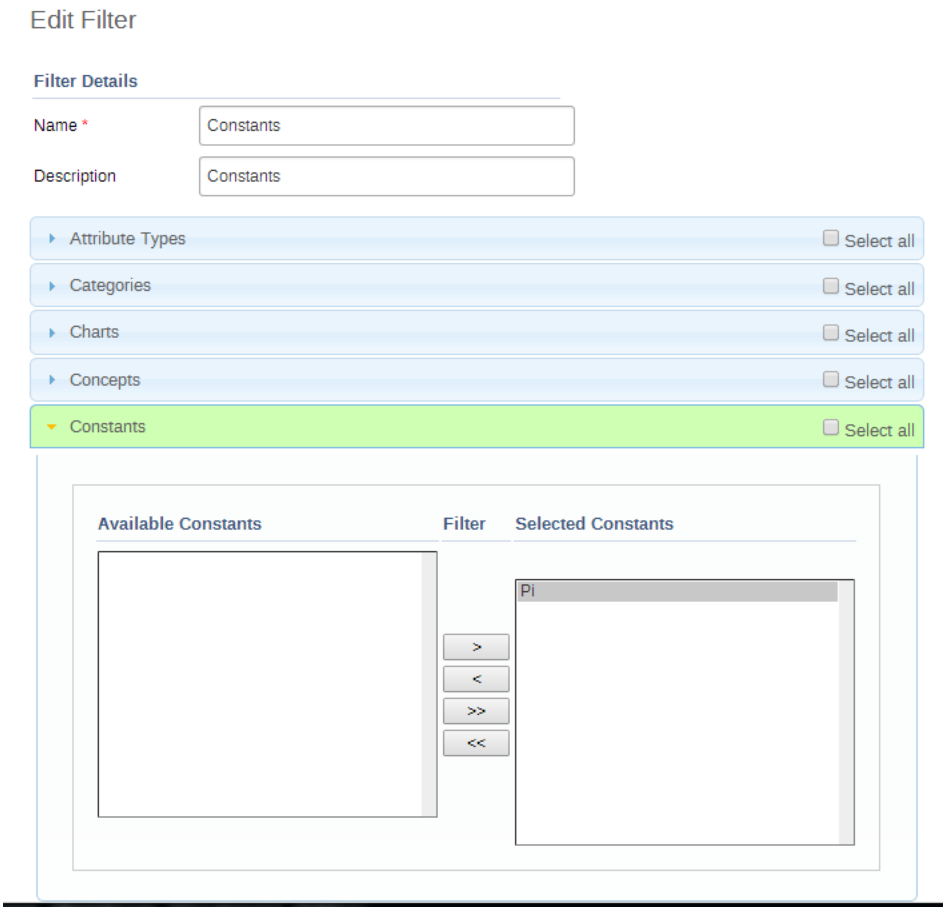
Simply choose the objects which you would like to export, and click "Export".

24.3.2. Meta-data detailed export

In certain implementations, it may be useful to create groups of metadata objects which should always be exported together. To access this feature, click " Meta-data detailed export" from the left side menu of the "Import-export" dialog.

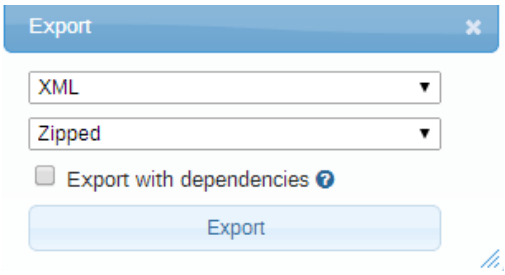


A list of available metadata export filters will be shown. Click on the name of the filter you would like to edit and then select "Edit" from the context menu. An example of the dialog is shown below.



In this example, we have created a metadata export filter, to export all "Constants" from the system. If any meta-data objects have been selected to be part of the filter, they will be highlighted in green. Click on the name of the objects you would like to export, and add them to the right side to be selected to be part of the export. Once you have composed the filter as needed, click "Save" (when editing) or "Add" (when adding a new metadata filter).

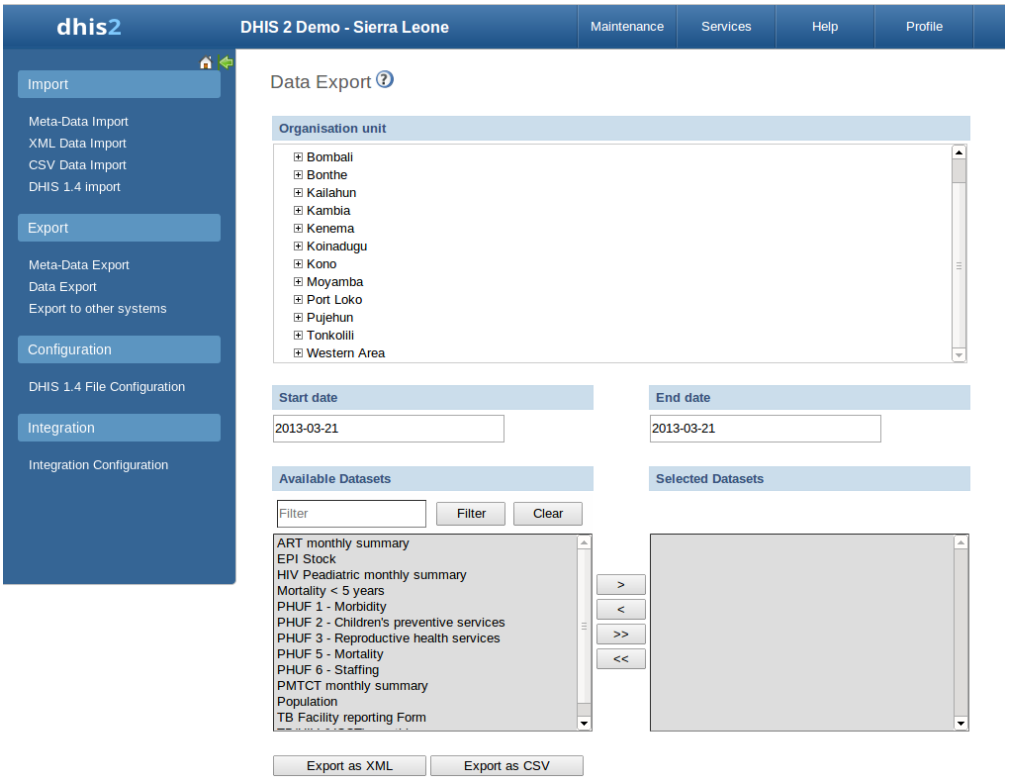
To export the filter, click on the name from the main "Meta-data detailed export" menu, and choose "Export" from the context menu. The following dialog will appear.



You can choose to export the data as XML or JSON format. You can also specify whether or not the result should be zipped or unzipped. Lastly, if you click check "Export with dependencies", all dependent meta-data objects which the ones you have selected depend upon, will also be exported.

24.3.3. Data export

To export raw data from DHIS2, choose "Services->Import-export->Data export". Select the organisation unit(s), the start and end date, and dataset or data sets for which data export should be selected. Click "Export as XML" to export data as XML and "Export as CSV" to export data in CSV format.



A pop-up save option will appear on the displayed screen (see picture below) prompting the saving of the exported data. You may save the export folder on your desktop or any other folder by selecting the ‘Save to Disk’ option from the pop-up prompt.

24.3.4. Event data export

Event (or tracker) data can be exported by accessing the "Event data export" function by choosing "Services->Import-export->Event data export" from the main menu. Choose the organisation unit of interest, along with the Programs and Program stages and start and end dates. The "Inclusion" option has three options.

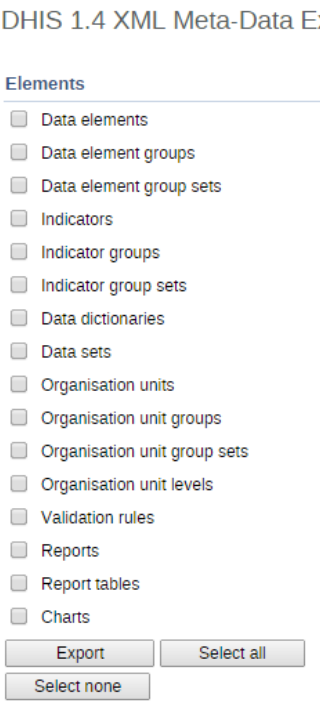
- Selected organisation unit: Export event data only for the selected organisation unit
- Include children of organisation unit: Export event data for the children of the organisation unit as well as the selected organisation unit itself.
- Include descendants of organisation unit: Export event data for the descendants of the organisation unit as well as the selected organisation unit itself.

Event data can be exported either as XML or JSON, and either compressed or uncompressed in a ZIP file. Once you have set all options, press "Export" . The export file will now by downloaded to your local computer.

24.3.5. Exporting data to other systems

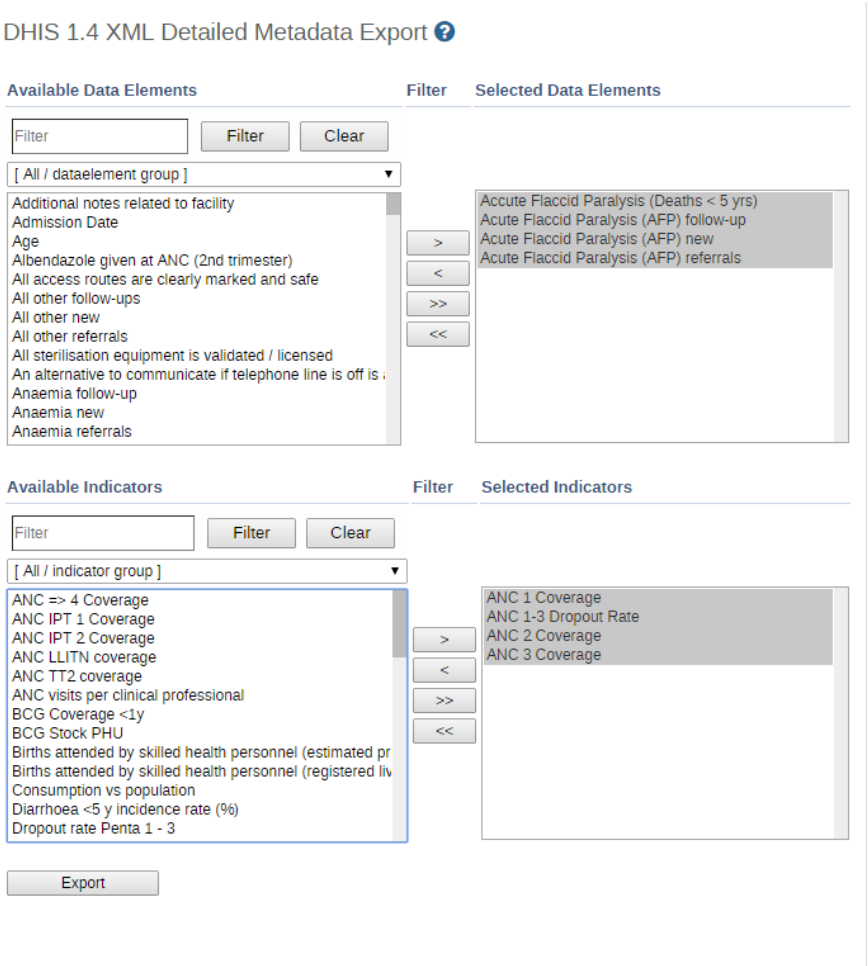
24.3.5.1. DHIS 1.4 Meta-data export

The DHIS 1.4 Meta-data export functionality provides the same functionality as the standard DHIS2 meta-data export, except that the resulting file can be used to transmit meta-data information to DHIS 1.4 systems.



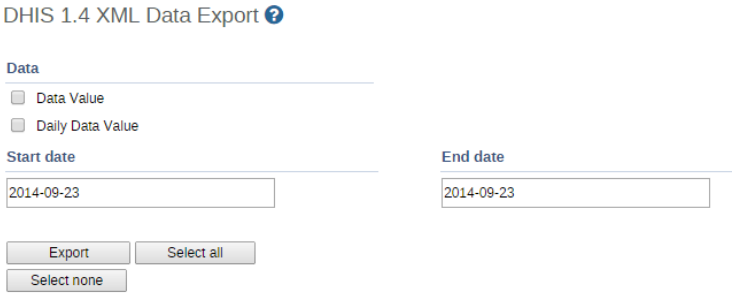
24.3.5.2. DHIS 1.4 Detailed Metadata Export

The DHIS 1.4 Metadata export functionality provides the same functionality as the detailed DHIS 2 metadata export, except that the resulting file can be used to transmit meta-data information to DHIS 1.4 systems. Simply select the data elements and indicators that you want and click "Export" to begin the export process.



24.3.5.3. DHIS 1.4 Data export

This service allows you to export data to a format which can easily be imported into a compatible DHIS 1.4 database.



24.3.5.4. XLS metadata export

Meta-data can be exported directly to the XLS format with this function. Simply click all of the available object which you wish to export and click "Export".

Chapter 25. Data Administration

The data administration module provides a range of functions to ensure that the data stored in the DHIS2 database is integral and that the database performance is optimised. These functions should be executed on a regular basis by a data administrator to ensure that the quality of the data stored is optimal.

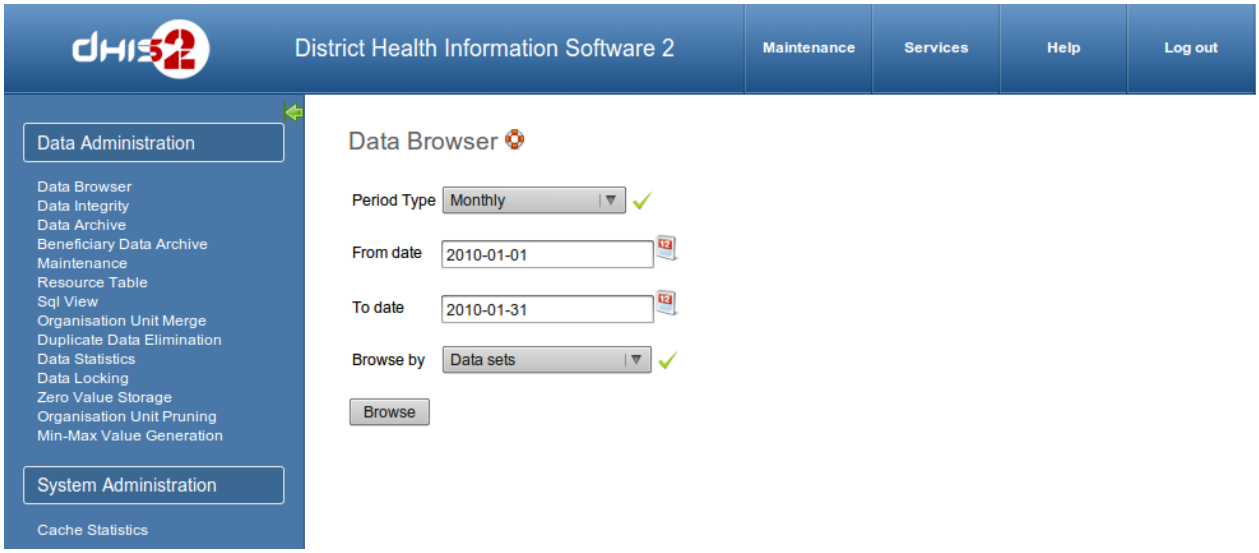
25.1. Data browser

The data browser maintenance and analysis module which allows the user to produce a summary of the data contained in the DHIS2 database. The summary view provides a count of data elements which have been entered at the selected organisation unit as well as its descendants. Raw data for all data elements for a range of time periods and a given organisational unit can be browsed and exported to Excel, CSV, or PDF formats. There are four modes of the data browser, which determine how the data is summarized

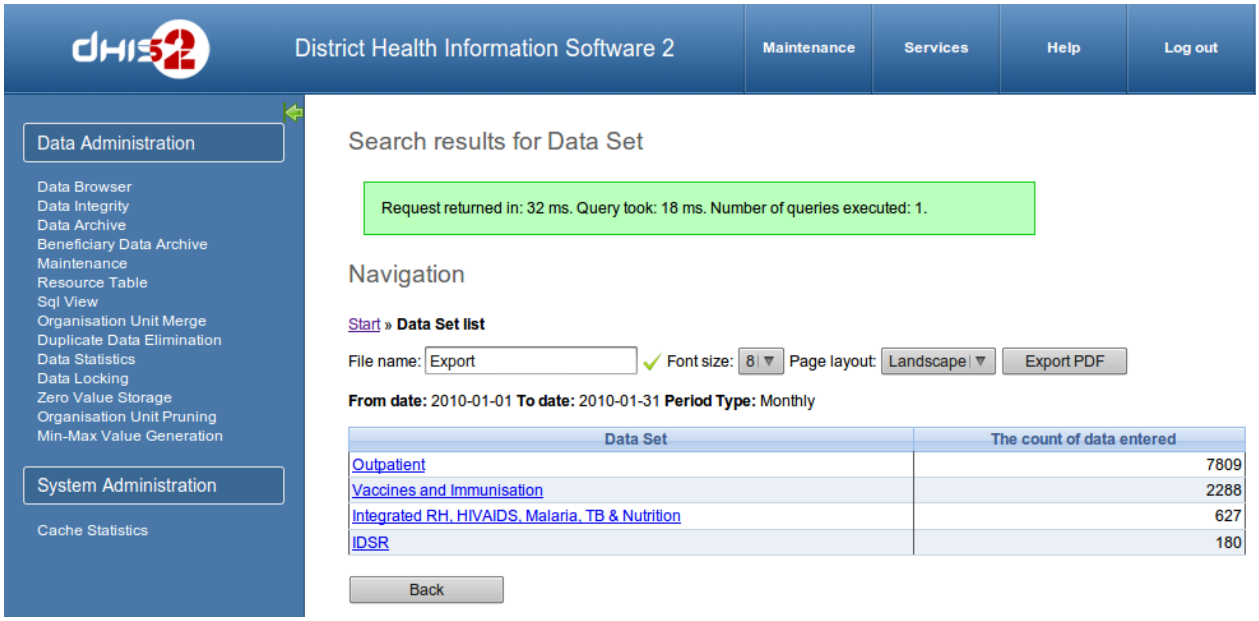
- Data sets
- Data element groups
- Organisational unit groups
- Organisational units

Each of these options can be accessed by selecting the desired option from "Browse by" drop-down menu.

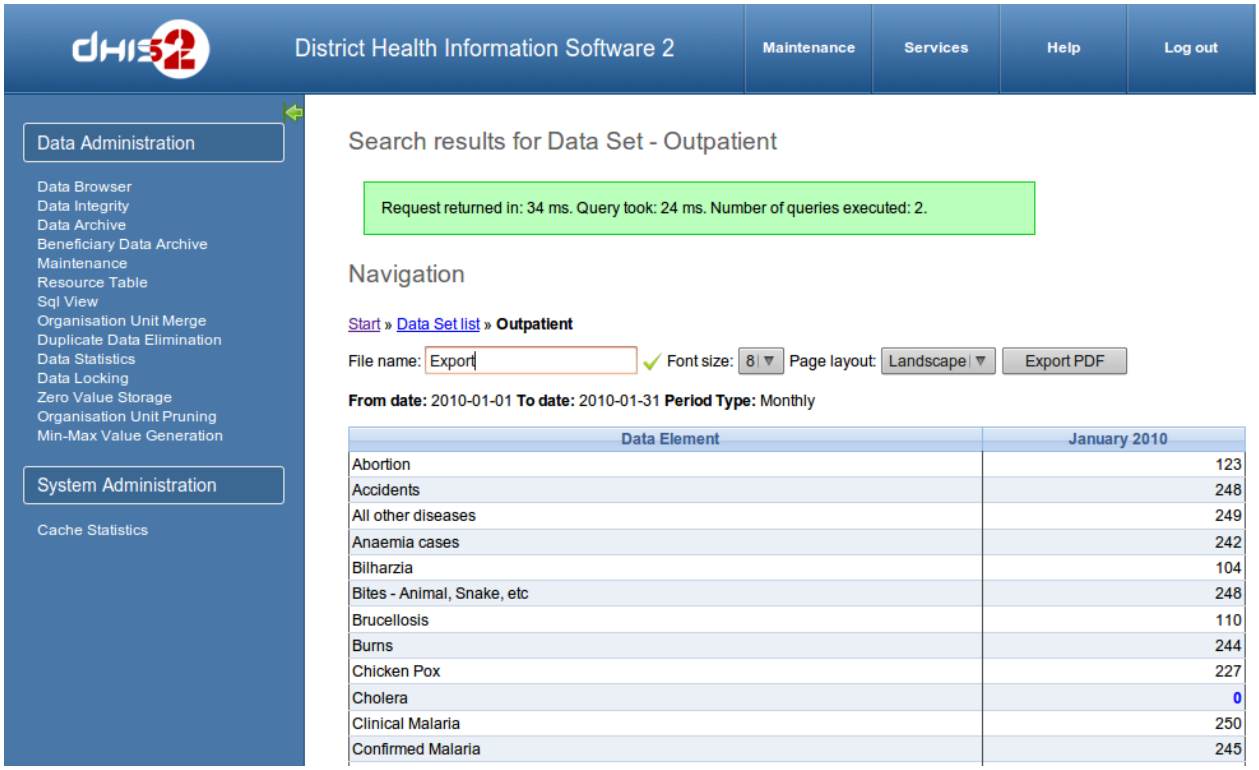
In order to produce a summary of submitted data for a given period and grouped by data sets, the user should follow this procedure. Begin by selecting a given periodicity type (e.g. Weekly, monthly, yearly, etc) and then a "From date" and "To date". (e.g. January 2009 to March 2009). Select the type of summary to be produced (e.g. Dataset) from the "Browse by" drop-down menu. Click the "Browse" button to view the summary.



A summary of the number of data element values that have been submitted over the user selected time period is shown below.

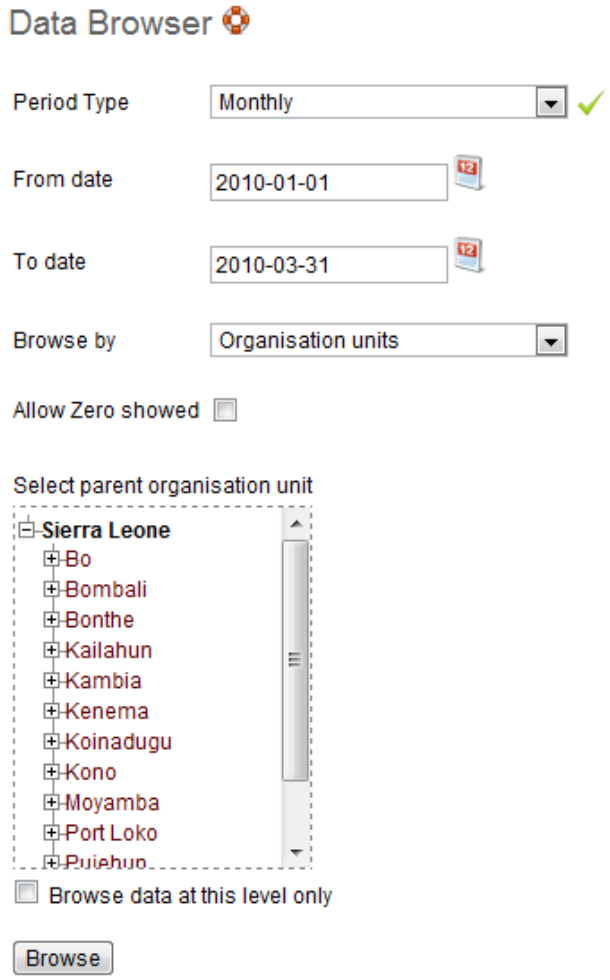


By clicking on the name of the individual dataset, a more detailed summary of each data element can be obtained as shown below. A cross-tab table summarising each time period will be shown.



The functionality of the grouping by Datasets, Data element groups, and Organisational groups is essentially the same.

The functionality of grouping by organisation units will be discussed below. Begin by selecting "Organisation units" from the "Browse by" drop-down menu. The organisational hierarchy present in the database will now be displayed. Organisational units can be expanded by clicking on the plus symbol in the organisational tree view.



By clicking on an organisational unit, and the clicking the "Browse" button, a summary of submitted data elements present in the database is returned for all immediate children of the selected organisational as shown below:

Organisation units - Bo

From date: 2010-01-01 To date: 2010-03-31, Period Type: Monthly

Download as Excel | Download as CSV | Download as PDF | Back

Start » Sierra Leone » **Bo** (Hit(s))

Organisation Unit	January 2010	February 2010	March 2010
Badjia	836	451	460
Baoma	2673	2691	2730
Bargbe	1257	1175	1316
Bargbo	1128	1214	974
Bumpe Ngao	1672	1587	1792
Gbo	253	234	229
Jaiama Bongor	1313	1283	1267
Kakua	3547	3393	3435
Komboya	707	792	675
Lugbu	1625	1491	1824
Niawa Lenga	477	784	770
Selenga	359	344	337
Tikonko	1273	1020	1144
Valunia	987	1045	1286
Wonde	418	476	511

By clicking on one of the organisational units, two drill down modes are presented to the user.

- Summary drill down- Drill down to the selected organisational units children to see the count of data elements.
- View raw data at this level: View the actual raw data at the selected organisational unit. A typical view of the raw data export can be seen below.

Organisation units - Ngelehun CHC

From date: 2010-01-01 To date: 2010-03-31, Period Type: Monthly

Download as Excel | Download as CSV | Download as PDF | Back

Start » Sierra Leone » Bo » Badjia » **Ngelehun CHC** (Raw data)

Data Element	January 2010	February 2010	March 2010
Acute Flaccid Paralysis (AFP) follow-up (0-11m)	1		
Acute Flaccid Paralysis (AFP) follow-up (12-59m)			
Acute Flaccid Paralysis (AFP) follow-up (15y+)			
Acute Flaccid Paralysis (AFP) follow-up (5-14y)			
Acute Flaccid Paralysis (AFP) new (0-11m)	11		
Acute Flaccid Paralysis (AFP) new (12-59m)			
Acute Flaccid Paralysis (AFP) new (15y+)			
Acute Flaccid Paralysis (AFP) new (5-14y)			
Acute Flaccid Paralysis (AFP) referrals (0-4y)	1		
Acute Flaccid Paralysis (AFP) referrals (5y+)			
Albendazole given at ANC (2nd trimester) (Fixed)	5	8	10
Albendazole given at ANC (2nd trimester) (Outreach)			
All other deaths (At PHU, 0-28d.)	1		

Data can be exported into Excel, CSV and as a PDF report by clicking the appropriate button.

25.2. Data integrity

DHIS2 can perform a wide range of data integrity checks on the data contained in the database. Identifying and correcting data integrity issues is extremely important for ensuring that the data used for analysis purposes is valid.

Each of the data integrity checks that are performed by the system will be described, along with general procedures that can be performed to resolve these issues.

25.2.1. Data elements without data set

Each data element must be assigned to a data set. Values for data elements will not be able to be entered into the system if a data element is not assigned to a data set. Choose Maintenance->Datasets->Edit from the main menu and then add the "orphaned" data element to the appropriate data set.

25.2.2. Data elements without groups

Some Data Elements have been allocated to several Data Element Groups. This is currently not allowed, because it will result in duplication of linked data records in the PivotSource recordsets that `feed` the pivot tables. Go to Maintenance -> Data Element Groups to review each Data Element identified and remove the incorrect Group allocations.

25.2.3. Data elements violating exclusive group sets

Some data elements have been allocated to several data element groups that are members of the same data element group set. All group sets in DHIS are defined as exclusive, which means that a data element can *only* be allocated to *one* data element group within that group set. Go to Maintenance -> Data elements and indicators ->Data element groups to review each data element identified in the integrity check. Either remove the data element from all groups except the one that it should be allocated to, or see if one of the groups should be placed in a different group set.

25.2.4. Data elements assigned to data sets with different period types

Data Elements should not be assigned to two separate data sets whose period types differ. The recommended approach would be to create two separate data elements (for instance a monthly and yearly data element) and assign these to respective datasets.

25.2.5. Data sets not assigned to organisation units

All data sets should be assigned to at least one organisation unit.

25.2.6. Indicators with identical formulas

Although this rule will not affect data quality, it generally does not make sense to have two indicators with the exact same definition. Review the identified indicators and their formulas and delete or modify any indicator that appears to be the duplicate.

25.2.7. Indicators without groups

All Data Elements and Indicators must be assigned to at least one group, so these Indicators need to be allocated to their correct Data Element and Indicator Group. Go to Maintenance -> Indicator Groups, and allocate each of the `Orphaned` Indicators to its correct group.

25.2.8. Invalid indicator numerators

Violations of this rule may be caused by an incorrect reference to a deleted or modified data element. Review the indicator and make corrections to the numerator definition.

25.2.9. Invalid indicator denominators

Violations of this rule may be caused by an incorrect reference to a deleted or modified data element. Review the indicator and make corrections to the denominator definition.

25.2.10. Indicators violating exclusive group sets

Some indicators have been allocated to several indicator groups that are members of the same indicator group set. All group sets in DHIS are defined as exclusive, which means that an indicator can *only* be allocated to *one* indicator group within that group set. Go to Maintenance -> Data elements and indicators ->Indicator groups to review each indicator identified in the integrity check. Either remove the indicator from all groups except the one that it should be allocated to, or see if one of the groups should be placed in a different group set.

25.2.11. Organisation units with cyclic references

Organisation units cannot be both parent and children of each other, directly nor indirectly.

25.2.12. Orphaned organisation units

All organisation units must exist within the organisation unit hierarchy. Go to Organisation->Hierarchy Operations and move the offending organisation unit into the proper position in the hierarchy.

25.2.13. Organisation units without groups

All organisation units *must* be allocated to at least *one* group. The problem might either be that you have not defined any `compulsory` OrgUnit Group Set at all, or that there are violations of the `compulsory` rule for some OrgUnits . NOTE: If you have defined no `compulsory` OrgUnit Group Sets, then you must first define them by going to Maintenance -> Organisation units->Organisation unit group sets and define at least one `compulsory` Group Set (the group set `OrgUnitType` are nearly universally relevant). If you have the relevant group sets, go to Maintenance -> OrgUnit Groups to review each OrgUnit identified and add the relevant Group allocation.

25.2.14. Organisation units violating compulsory group sets

These organisation units have not been assigned to the any organisation unit group within one of the *compulsory* organisation unit group sets. When a group set is defined as `compulsory`, it means that an organisation unit must be allocated to at least one organisation unit group within that group set. For instance, all organisation units must belong to one of the groups in the `organisation unitType` group set. It might belong to the `Hospital` or the `Clinic` or any other `type` group - but it must belong to exactly one of them. Go to Maintenance -> organisation unit->Organisation unit groups to review each organisation unit identified in the integrity check. Allocate all organisation units to exactly one group.

25.2.15. Organisation units violating exclusive group sets

Some organisation units have been allocated to several organisation unit groups that are members of the same organisation unit group set. All group sets in DHIS are defined as exclusive, which means that an organisation unit can *only* be allocated to *one* organisation unit group within that Group Set. For instance, one organisation unit cannot normally belong to the both the 'Hospital' and 'Clinic' groups , but rather to only to one of them. Go to Maintenance -> organisation unit->Organisation unit groups to review each organisation unit identified in the integrity check. Remove the organisation unit from all groups except the one that it should be allocated to.

25.2.16. Organisation unit groups without group sets

The organisation unit groups listed here have not been allocated to a group set. Go to Maintenance->Organisation unit->Organisation unit group sets and allocate the Organisation unit group to the appropriate group set.

25.2.17. Validation rules without groups

All validation rules must be assigned to a group. Go to **Services->Data quality->Validation rule group** and assign the offending validation rule to a group.

25.2.18. Invalid validation rule left side expressions

An error exists in the left-side validation rule definition. Go to **Services->Data quality->Validation rule** and click the "Edit" icon on the offending rule. Press "Edit left side" and make the corrections that are required.

25.2.19. Invalid validation rule right side expressions

An error exists in the left-side validation rule definition. Go to **Services->Data quality->Validation rule** and click the "Edit" icon on the offending rule. Press "Edit right side" and make the corrections that are required.

25.3. Maintenance

The data maintenance module has five options, each described below.

- Clear analytics tables

  Completely deletes the analytics tables, used to generate aggregate data for the pivot tables, GIS and reports.

- Clear data mart (aggregated indicator and data value values)

  The data mart is where DHIS 2 stores aggregated data produced during the export to data mart process. This function clears the database table which contains aggregated indicator and data element values.

- Rebuild data mart index

  Rebuilds the database indices on the aggregated data generated during a data mart process.

- Clear zero values

  This function removes zero data values from the database. Values registered for data elements with aggregation operator *average* is not removed, as such values will be significant when aggregating the data, contrary to values registered for data elements with aggregation operator *sum*. Reducing the number of data values will improve system performance.

- Clear dataset completeness

  This function removes aggregated dataset completeness values. This data is produced and used by report tables.

- Prune periods

  This function removes all periods which have no registered data values. Reducing the number of periods will improve system performance.

- Update category option combinations

  Rebuilds the category option combinations. This may be required after altering the category options which belong to a given category.

25.4. Resource tables

Resource tables are supporting tables that are used during analysis of data. One would typically join the contents of these tables with the data value table when doing queries from third-party applications like Microsoft Excel. They are also used extensively by the analysis modules of DHIS2. Regeneration of the resource tables should only be done once

all data integrity issues are resolved. The resource tables are also generated automatically, every time the analytics process is run by the system.

- Organisation unit structure (_orgunitstructure)

  This table should be regenerated any time there have been any changes made to the organisational unit hierarchy. This table provides information about the organisation unit hierarchy. It has one row for each organisation unit, one column for each organisation unit level and the organisation unit identifiers for all parents in the lineage as values.

- Data element group set structure (_dataelementgroupsetstructure)

  This table provides information about which data elements are members of which data element group sets. The table has one row for each data element, one column for each data element group set and the names of the data element group as values.

- Indicator group set structure (_indicatorgroupsetstructure)

  This table provides information about which indicators are members of which indicator group sets. The table has one row for each indicator, one column for each indicator group set and the names of the indicator group as values.

- Organisation unit group set structure (_organisationunitgroupsetstructure)

  This table provides information about which organisation units are members of which organisation unit group sets. The table has one row for each organisation unit, one column for each organisation unit group set and the names of the organisation unit groups as values.

- Category structure (_categorystructure)

  This table provides information about which data elements are members of which categories. The table has one row for each data element, one column for each category and the names of the category options as values.

- Data element category option combo name (_categoryoptioncomboname)

  This table should be regenerated any time there have been changes made to the category combination names. It contains readable names for the various combinations of categories.

- Data element structure (_dataelementstructure)

  This table provides information about all data elements and which period type (frequency) they capture data at. The period type is determined through the data set membership and hence relies on data elements to be member of data sets with similar period types to have a defined behavior.

- Period structure (_dataperiodstructure)

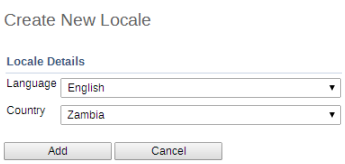
  This table provides information about all periods and which period type they are associated with. For each period type with lower frequency than itself, it contains information about which period if will fall within.

- Data element category option combinations (_dataelementcategoryoptioncombo)

  This table provides a mapping between data elements and all possible category option combinations.

25.5. Locale Management

It is posible to create custom locales in DHIS2. In addition to the locales available through the system, you might want to add a custom locale such as "English" and "Zambia" to the system. This would allow you to translate metadata objects to local languages, or to account for slight variants between countries which use a common metadata definition.



The locale is composed of a language along with a country. Select the desired values and press "Add". This custom locale will now be available as one of the translation locales in the system.

25.6. SQL View

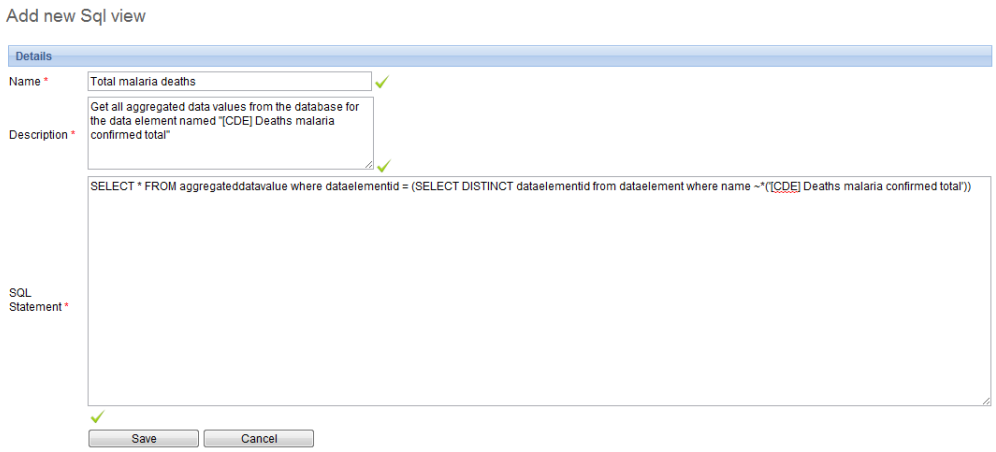
The SQL View functionality of DHIS2 will store the SQL view definition internally, and then materialize the view when requested.

Database administrators must be careful about creating database views directly in the DHIS 2 database. For instance, when the resource tables are generated, all of them will first be dropped and then re-created. If any SQL views depend on these tables, an integrity violation exception will be thrown and the process will be aborted.

The SQL views are dropped in reverse alphabetical order based on their names in DHIS 2, and created in regular alphabetical order. This allows you to have dependencies between SQL views, given that views only depend on other views which come earlier in the alphabetical order. For instance, "ViewB" can safely depend on "ViewA". Otherwise, having views depending on other view result in an integrity violation error.

25.6.1. Creating a new SQL view

To create a new SQL view, choose Maintenance->SQL view and click the "Add new" button.



The "Name" attribute of the SQL view will be used to determine the name of the table that DHIS2 will create when the view is materialized by the user. The "Description" attribute allows one to provide some descriptive text about what the SQL view actually does. Finally, the "SQL statement" should contain the SQL view definition. Only SQL "SELECT" statements are allowed and certain sensitive tables (i.e. user information) are not accessible Press "Save" to store the SQL view definition.

25.6.2. SQL View management

In order to utilize the SQL views, simply press the "Execute query" button  from the "SQL View management page. Once the process is completed, you will be informed that a table has been created. The name of the table will be provided, and is composed from the "Description" attribute provided in the SQL view definition. Once the view has been materialized, click on the "View" button  .

25.7. Organisation unit merge

This function is useful when two organisation units need to be merged, e.g. it is decided that one facility will be shut down and its services will be provided by a nearby facility.

Start by selecting the organisation unit to eliminate from the tree and click *confirm*. Then select the organisation unit to keep and click *confirm* again. Finally, verify the selection and click *merge*.

In the situation where data exist for the organisation unit to eliminate and not for the one to keep, the data will be moved to the one to keep. When data exists for both organisation units, the data will be summarized and moved to the one to keep. When data exists only for the one to keep, no action is taken. The organisation unit to eliminate will eventually be deleted.

25.8. Duplicate data elimination

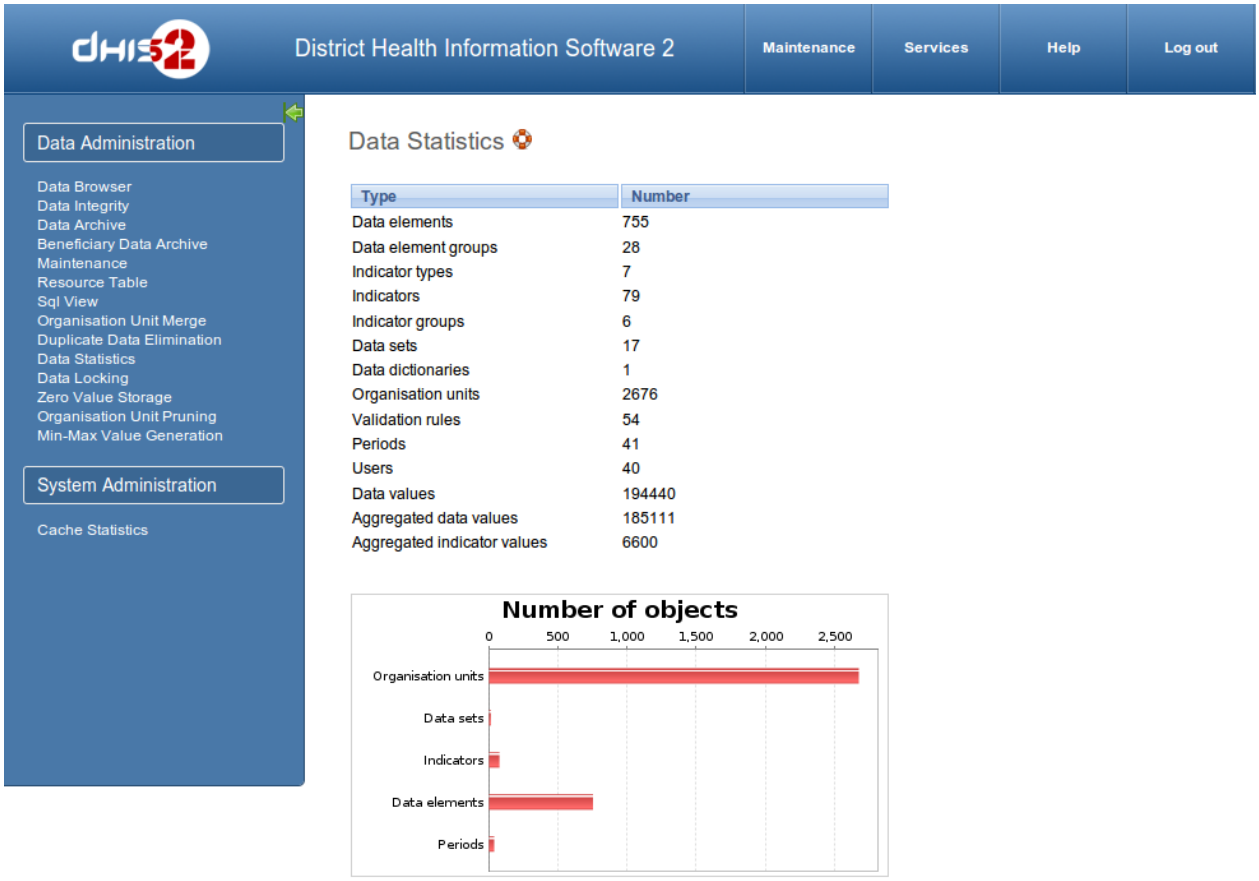
This function is useful when data has been entered mistakenly for two data elements which represents the same phenomena.

Start by selecting the data element to eliminate from the list and click confirm. Then select the data element to keep and click confirm again. Finally, verify the selection and click merge.

In the situation where data exists for the data element to eliminate and not for the one to keep, the data will be moved to the one to keep. When data exists for both data elements, the data which was updated last will be used. When data exists only for the one to keep, no action will be taken. The data element to eliminate will eventually be deleted, except when it is a multidimensional data element and has other data registered.

25.9. Data statistics

The data statistics module provides an overview of the number of objects stored in the DHIS2 database.

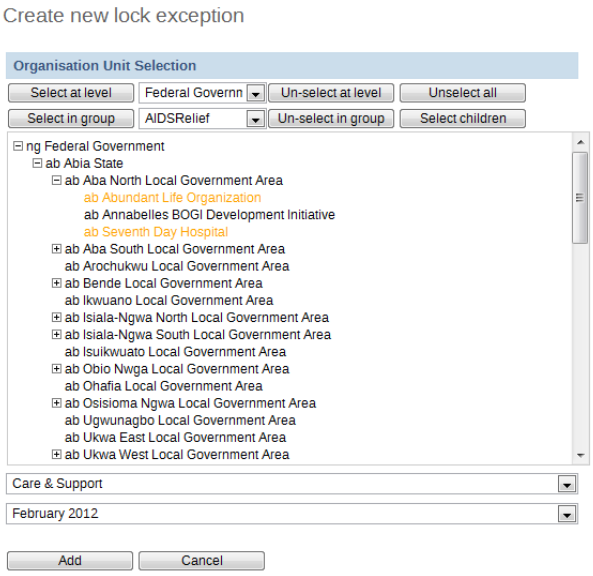


The total number of each type of object is presented in a table, as well as a graph.

25.10. Lock exceptions

Lock exceptions provide fine-grained control over exemption from a locked data set. After the expiry of the data set, data entry will be denied by default, unless an exception has been granted through the Lock exception interface. To

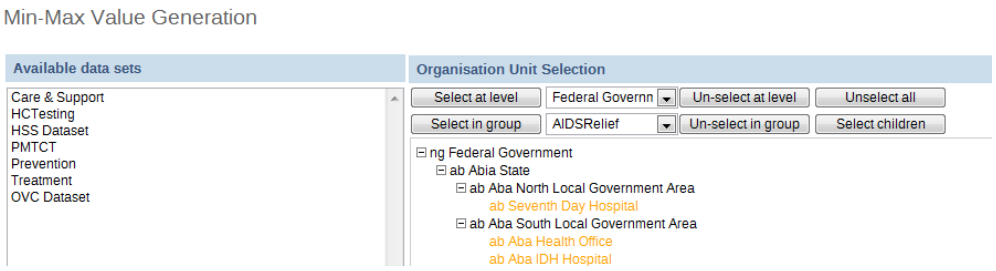
enable a lock exception, select the desired organization units, data sets, and time period and press "Add". By granting a lock exception, data entry will be enabled even after the expiry period of the data set has passed.



In the example above, a data lock exception would be created for "ab Abundant Life Organization" and "ab Seventh Day Hospital" for the "Care and Support" dataset for "February 2012".

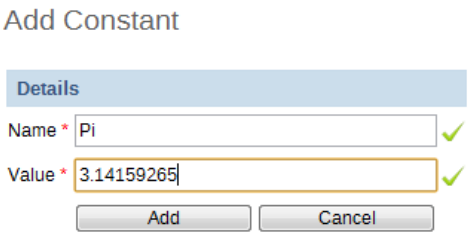
25.11. Min-Max Value Generation

This administrative function can be used to generate min-max values, which are used as part of the data quality and validation process for specific organization units and data sets. Simply select the dataset from the left hand frame, and then select the required orgunits to generate the min-max values for from the organisational units selector on the right. Press the "Generate" button to generate or regenerate all min-max values. Press "Remove" to remove all min-max values which are currently stored in the database.



25.12. Constant

Constants are static values which can be made available to users for use in data elements and indicators. Some indicators, such as "Couple year protection rate" depend on constants which usually do not change over time. Simply press "Add" and provide a name in the "Name" field and define it's value in the "Value" field. Press "Add" . The constant will now be available to users for use in their expressions.



25.13. Option sets

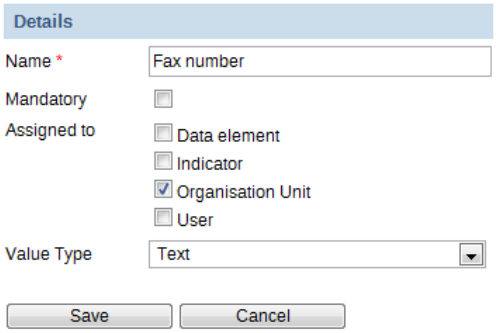
Option sets can be associated with data elements in the add / update data element interface for name-based data elements. You can define any kind of options, for instance an option set called "Delivery type" where "Normal", "Breach", "Caesarian" and "Assisted" would be the options. This option set can later be associated with any number of data elements. When doing data entry in name-based records module those elements will then appear in the form as drop-down lists with auto-completion support.

25.14. Cache Statistics

This option is for system administrators only to use. The cache statistics shows the status of the application level cache. The application level cache refers to the objects and query results that the application is caching in order to speed up performance. If the database has been modified directly the application cache needs to be cleared for it to take effect.

25.15. Attributes

Dynamic attributes can be used to add additional information to certain objects (namely data elements, indicators, organisation units and users). In addition to the standard attributes each of these objects have, it may be required in certain installations to have additional attributes, such as a fax number which is associated with an organisation unit. To add a new dynamic attribute to an object, select "Maintenance->Data administration" from the main menu, then "Attribute" from the left side panel, and press the "Add new" button.



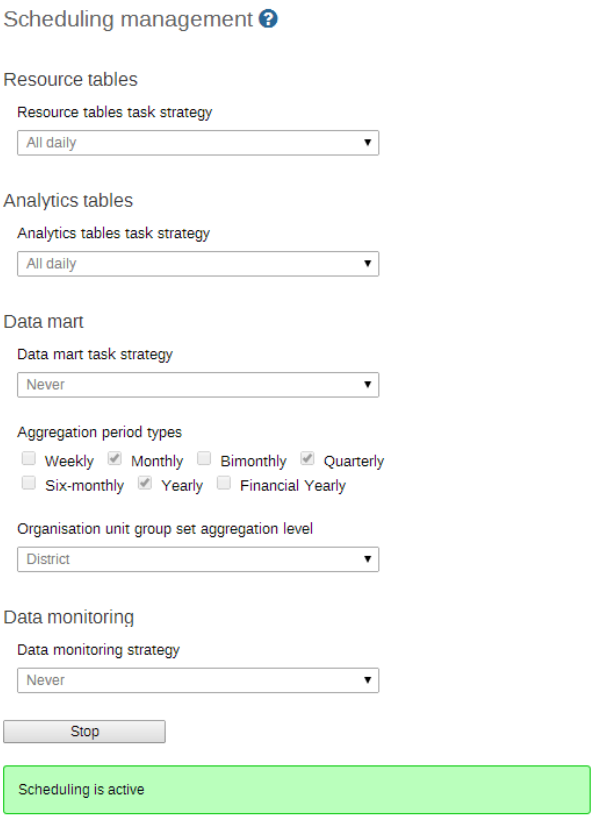
To create a new attribute, assign it a name. Each attribute should have a unique name. Check the tick-box "Mandatory" if the object should always have the dynamic attribute. Next, select which object (or objects) the attribute should be assigned to. Lastly, select the value type. You can choose from "Text", "Yes/No", "Date", "Number", "Integer", "Positive integer" and "Negative integer". If the value supplied for the attribute does not match the value type, an error will result. Finally, click "Save" to save the attribute.

The dynamic attribute will now be present in the object which you assigned it to in the respective "Edit" screen of each the object.

25.16. Scheduling

The analytics, resource tables and data mart can be automatically scheduled to run on regular intervals. Simply select the aggregation period types, organisation unit group set aggregation level, and strategy to configure how the scheduled job should run. If you are using surveillance rules, you can choose to run them "All daily" by selecting this option.

Pressing "Start" will enable the scheduled job to run at a pre-determined time (always at midnight based on the server time).



25.17. Data synchronization

DHIS 2 provides a feature for synchronizing data being captured on the local instance with a another, remote instance of DHIS 2. This can be useful e.g. when you have deployed multiple stand-alone instances of DHIS 2 which are required to submit data values to a central DHIS 2 instance.

These are the steps to enable data synchronziation:

- Go to Settings > Synchronization, enter the remote server URL, username and password and click Save. You can test your settings by clicking on the "Test settings" link.
- Go to Data administration > Scheduling. Under Data synchronization set strategy to Enabled, and click Start.

Some aspects of the data synchronization feature to be aware of:

- The local DHIS 2 instance will store the password of the user account on the remote instance encrypted in the local database. The remote account is used for authencation when transferring data. For security purposes make sure you set the "enryption.password" configuration parameter in hibernate.properties to a strong password.
- Deploying the remote server on SSL/HTTPS is strongly recommended as the username and password are sent in clear text using basic authentication and could be intercepted by an attacker.
- The data synchronization uses the UID property of data elements, category option combos and organisation units to match the meta-data. Hence the synchronization is dependent on these three meta-data objects being harmonized on the local and remote instance in order to work appropriately.
- The very first time DHIS 2 attempts to synchronize data the system will include data entered during the last three days. For the subsequent attempts the system will store the time of the last successful data synchronization and only include data saved or edited since that time. A synchronization job is considered successful only if data was submitted, authenticated and saved successfully on the remote server,.
- The system will attempt a synchronization every minute. If the local server does not have a working Internet connection at the time, the synchronization will be silenly aborted and re-attempted after a minute.
- You can see the time of last successful synchronization with remote server in the scheduling screen next to the "Last success" label.

Chapter 26. Settings

The settings module provides a set of application configuration options. There are two main groups of settings: the *system settings* apply to the whole system and all its users while the *user settings* apply to the environment of the currently logged in user. The system settings can be accessed from the maintenance menu, settings module. The user settings can be accessed under the profile menu, settings page.

26.1. System settings

The system settings section provides general configuration options and options specifically for appearance and email.

26.1.1. General settings

- Cache strategy: Decides for how long reports and responses related to analysis should be cached. If you are using the scheduled, nightly data mart tasks it makes sense to put this on "Cache until 6 AM tomorrow". This is because we know that data in reports change at that time, and you can safely cache data up to the moment when the data mart is updated. If you are loading data continuously into the datamart you should set it to "No cache". If you load data very infrequently into data mart you should consider setting it to "Cache for two weeks".
- Infrastructural data elements: This setting defines a data element group where the member data elements should describe data about the infrastructure of organisation units. Examples of such infrastructural data elements could be population, doctors, beds, internet connectivity and climate. This infrastructural data can currently be viewed in the GIS module in the facility information sheet.
- Infrastructural period type: Sets the frequency for which the data elements in the infrastructural data elements group are captured. This will typically be yearly. When viewing the infrastructural data you will be able to select the time period of the data source.
- Feedback recipients: This setting defines a user group where the members will receive all messages being sent through the function for writing feedback in the dashboard module. This will typically be members of the super user team who are able to support and answer questions coming from end-users.
- Maximum offline organisation unit levels: This setting defines how many levels in the organisation unit hierarchy will be available offline in the organisation unit tree widget. Under normal circumstances you can leaves this on the lowest level, which is default behavior. Setting it to a higher level might be useful in order to reduce initial load time in cases where you have extremely many organisation units, typically more than 30 000.
- Data analysis std dev factor: Sets the number of standard deviations for use in the outlier analysis performed on the captured data in the data entry module. The default value is 2; a high value will catch less outlier values than a low value.
- Days after period end to qualify for timely data submission: Sets the number of days after the end of a period in which a data entry form must be marked as complete in order to be considered timely. This affects the "reporting rate" tool in the reporting module which lists forms marked as complete as well as marked as complete in time. The default value is 15.
- Phone number area code: The area code for the area in which your deployment is located. Used for sending and receiving SMS.
- Google Analytics (Universal Analytics) Key: Set your Google UA key here to provide analytics for your DHIS 2 instance. Most places are covered, but it will not be provided for custom apps. Read more about Google Analytics at http://google.com/analytics .
- Enable multi-organisation unit forms: Enable support for entering data forms for multiple organisation units at the same time, in data entry, click on the parent organisation unit for the children that you want to enter data for, and the dataset list will include datasets that are assigned to the children of that parent.
- Omit indicator values with zero numerator value in data mart: Defines whether aggregated indicator values with zero as the numerator value should be written to the indicator data mart table. Having such values written is required for instance when connecting Excel pivot tables to the data mart as Excel will need the numerator data to correctly aggregate up in the organisation unit hierarchy. If third-party tools like Excel are not used with the application this

will reduce the total number of values written to the data mart (which again will improve performance) and could safely be set to omit.

- Put analytics in maintenance mode: Puts the analytics engine / Web API resource in maintenance mode, implying that 503 Service Unavailable will be returned for all requests. This is useful when you need to perform maintenance on the server like rebuilding indexes while the server is running in production, in order to reduce load and more efficiently carry out the maintenanec.

26.1.2. System appearance settings

- Application title: Sets the application title on the top menu.
- Application introduction: Sets an introduction of the system. Will be visible on the login page.
- Application notification: Sets a notification which should be displayed to users. Will be visible on the front page.
- Style: Sets the style / look-and-feel of the system. The corresponding user style setting overrides this.
- Flag: Sets the flag which is displayed in the left menu of the dashboard module.
- Start page: Sets page / module which the user will be redirected to after logging in. The dashboard module is the recommended start module.

26.1.3. Email settings

- Host name: Refers to the host name of the SMTP server. For instance when using Google SMTP services this should be smtp.gmail.com.
- Port: The port to connect to the SMTP server.
- User name: The user name of the user account with the SMTP server. For instance mail@dhis2.org.
- Password: The password of the user account with the SMTP server.
- TLS: Refers to whether the SMPT server requires TLS for connections.

26.1.4. Access settings

- Self registration account user role: Defines which user role should be given to self-registered user accounts. To enable self-registration of users, select any user role from the list. To disable it, select "Do not allow self registration". When enabled, a link to the self-registration form will be displayed on the login page.
- Do not require recaptcha for self registration: Whether or not to use reCAPTCHA for user registration.
- Self registration account organisation unit: Defines which organisation unit should be associated with self-registered users. Any organisation unit must be selected in order to enable self registration.
- Enable user account recovery: Defines whether users are allowed to restore the password of their account if they forgotten it. When enabled, a link to the account recovery form will be displayed on the front page. User account recovery requires that you have configured email settings (SMTP).
- Enable user account invite: Defines user account invites can be sent. Account invites let you invite new users to create their own accounts by sending an email invitation.
- Allow users to grant own user roles: Defines whether users should be allowed to grant the user roles they are granted themselves to others.
- Users must belong to a group controlled by the user manager: This allows user groups to play a role in user management. When checked, user A can manage user B only if user B belongs to a user group to which user A has read/write access. When user A creates a new user, she or he must assign the new user to such a user group.
- Require user account password change: Require that users change their password every 3,6,12 months. Please note that for 2.14 release, they will have to login through the desktop to change passwords.
- OpenID provider: Defines the OpenId provider.
- OpenID provider label: Defines the label to display for the specified OpenID provider.

26.1.5. Calendar settings

- Calendar: Defines which calendar system should be used throughout the system.

- Date format: Defines which date format should be used throughout the system.

26.1.6. Synchronization settings

- Remote server URL: The URL of the remote server running DHIS 2 to upload data values to. Use of SSL/HTTPS is recommended since username and password is sent with the request (using basic authentication). The system will attempt to synchronize data once every minute. Note that you must enable data synch from Data administration > Scheduling.
- Remote server username: The username of the DHIS 2 user account on the remote server to use for data synchronization.
- Remote server password: The password of the DHIS 2 user account on the remote server. The password will be stored encrypted.

Chapter 27. Data dimensions in DHIS2

27.1. The core building blocks describing the data

A data value in DHIS2 (core module for aggregated data) is at minimum described by three dimensions, 1) data element, 2) organisation unit, and 3) period, which also forms the core building blocks of the data model. E.g. if you want to know how many children that were immunised for measles in Gerehun CHC in December 2009, the three dimensions to that value are the Data Element "Measles doses given", the Organisation Unit "Gerehun CHC", and the Period "Dec-09". All data values have at least these three dimensions describing What, Where and When.

Table 27.1.

Organisation Unit	Data Element	Period	Value
Gerehun CHC	Measles doses given	Dec-09	22
Tugbebu CHP	Measles doses given	Dec-09	18

27.2. The data element dimension

27.2.1. Data element categories

The data element above called "Measles doses given", although capturing the core essence of what the data is can be further elaborated and disaggregated into what is called data element categories. The system administrators of DHIS are free to define any data element category dimensions to data elements using the user interface for this in the data element maintenance section, but here are some examples of possible categories to use.

Given the example of Measles vaccination, if you want to know whether these vaccines were given at the facility (fixed) or out in the community as part of the outreach services then you could add a dimension called e.g. "Place of service" with the two possible options "Fixed" and "Outreach". Then all data collected on measles immunisation would have to be disaggregated along these to options. In addition to this you might be interested in knowing how many of these children who were under 1 year or above 1 year of age. If so you can add an Age dimension to the data element with the two possible options "<1 y" and ">1 y". This implies further detail on the data collection process. You can also apply both categories Place of service and Age and combine these into a data element category combination e.g. called "EPI disaggregation" and thereby you would be able to look at 4 different more detailed values in stead of only 1 as in the example above for the data element "Measles doses given": 1) "Fixed and <1 y, 2) Fixed and >1 y, 3) Outreach and <1 y, and 4) Outreach and >1 y. This adds complexity to how data is collected by the health facilities, but at the same time opens up for new possibilities of detailed data analysis of Measles immunisation.

Table 27.2. Example of detailed storage of data values when using data element categories "Place of Service" and "Age" (simplified for readability compared to the actual database table)

Organisation Unit	Data Element	Place of service	Age	Period	Value
Gerehun CHC	Measles doses given	Fixed	<1 y	Dec-09	12
Gerehun CHC	Measles doses given	Outreach	<1 y	Dec-09	4
Gerehun CHC	Measles doses given	Fixed	>1 y	Dec-09	4
Gerehun CHC	Measles doses given	Outreach	>1 y	Dec-09	2

Organisation Unit	Data Element	Place of service	Age	Period	Value
Tugbebu CHP	Measles doses given	Fixed	<1 y	Dec-09	10
Tugbebu CHP	Measles doses given	Outreach	<1 y	Dec-09	4
Tugbebu CHP	Measles doses given	Fixed	>1 y	Dec-09	3
Tugbebu CHP	Measles doses given	Outreach	>1 y	Dec-09	1

27.2.2. Data element group sets

While the data element categories and their options described above dictated the level of detail (disaggregation) at the point of data collection and how data values get stored in the database, the data element group sets and groups can be used to add more information to data elements after data collection. E.g. if looking at a lot of data elements at the same time in a report you would want to group these based on some criteria, e.g. if looking at all the data captured in a form for immunisation and nutrition you might want to separate or group data elements along a programme dimension (called group set) where "Immunisation" (or EPI) and "Nutrition" would be the two groups. Expanding the report to include data from other programs or larger themes of health data would mean more groups to such a group set dimension, like "Malaria", "Reproductive Health", "Stocks". For this example you would create a data element group set called "Programme" (or whatever name you find appropriate), and to represent the different programmes in this dimension you would define data elements groups called "EPI", "Nutrition", "Malaria", "Reproductive health" and so on, and add all these groups to the "Programme" group set. To link or tag the data element "Measles doses given" to such a dimension you must (in our example) add it to the "EPI" group. Which groups you add "Measles doses given" to does not affect how health facilities collect the data, but adds more possibilities to your data analysis. So for the group set dimensions there are three levels; the group set (e.g. "Programme"), the group (e.g. "EPI"), and the data element (e.g. "Measles doses given").

Indicators can be grouped into indicator groups and further into indicator group sets (dimensions) in exactly the same way as data elements.

Table 27.3.

Organisation Unit	Data Element	Programme	Period	Value
Gerehun CHC	Measles doses given	EPI	Dec-09	22
Gerehun CHC	Vitamin A given	Nutrition	Dec-09	16
Tugbebu CHP	Measles doses given	EPI	Dec-09	18
Tugbebu CHP	Vitamin A given	Nutrition	Dec-09	12
Gerehun CHC	Malaria new cases	Malaria	Dec-09	32
Tugbebu CHP	Malaria new cases	Malaria	Dec-09	23

27.3. The organisation unit dimension

Organisation units in DHIS2 can be either any type of health facility like Community Health Centres or referral hospitals, or an administrative unit like "MoHS Sierra Leone", "Bo District" or "Baoma Chiefdom". Orgunits are represented in a default hierarchy following the health system at large, and are therefore assigned an organisational level. E.g. Sierra Leone has 4 levels; National, District, Chiefdom, and PHU, and all orgunits are linked to one of these levels. An orgunit hierarchy in DHIS can have any number of levels. Normally data is collected at the lowest level, at the PHUs, and then data values are linked to individual PHUs. When designing reports at higher levels with data aggregated by chiefdom or district, the DHIS will use the hierarchy structure to sum up all the health facilities' data for any given unit at any level. The orgunit level capturing the data always represents the lowest level of detail

that is possible to use in data analysis, and the organisational levels define the available levels of aggregation along a geographical dimension.

27.3.1. Organisation unit group sets and groups

While PHU is the lowest geographical level for disaggregation in DHIS2 there are ways to flexibly group organisation units into any number of dimensions by using the group sets and groups functionality. E.g. if all PHUs are given an official type like CHC, CHP, MCHP etc. it is possible to create an orgunit group set called "Orgunit type" and add groups with the names of the types mentioned above. Then you can link your orgunits to their corresponding groups (orgunit types). Other common orgunit dimensions are Rural/Urban (Rural, Urbal, Peri-urban), and Ownership (Public, Private. NGO etc.). When analysing data from the PHU level it then becomes possible to aggregate data by these dimensions, e.g. look at Measles immunisation in BO district by the type of PHU in stead of the PHUs themselves.

27.3.1.1. Alternative orgunit hierarchies - advanced use of group sets and groups

A more advanced use of orgunit group sets and groups is to create alternative hierarchies e.g. use administrative borders from other ministries. In Sierra Leone that could mean an alternative hierarchy of 1:MoHS, 2:Districts, and 3: Local councils, instead of the 4 level hierarchy with chiefdoms and PHUs. E.g. if all PHUs where linked to a specific local council it would be possible to look at data aggregated by local council instead of chiefdom. Then you would first need to create a group set called "Local council" and then create one orgunit group for every local council, and finally link all PHUs to their corresponding local council group.

Table 27.4.

District	OrgUnit Type	Data Element	Period	Value
Bo	CHC	Measles doses given	Dec-09	121
Bo	CHP	Measles doses given	Dec-09	98
Bo	MCHP	Measles doses given	Dec-09	87
Bombali	CHC	Measles doses given	Dec-09	110
Bombali	CHP	Measles doses given	Dec-09	67
Bombali	MCHP	Measles doses given	Dec-09	59

27.3.2. Best practice on the use of group sets and groups

Groups that are not members of group sets are more or less useless for analysis (should we allow this at all?). With all the groups you get by e.g. including all different diagnoses as groups the full list of groups does not makes sense in any pivot table, and you easily get duplicates in your pivot tables since data elements or indicators are members of multiple groups. This can be controlled by the use of group sets. So the recommodation is to assign all your groups to a group set, and then you pull the group sets you need into your reports and analysis.

It is also recommended to have one group set that is used for the major organising of all data in e.g. a pivot table, e.g. use health programs or other larger themes for data elements or indicators that together cover all the data. That will provide a nicely organised overview of all your raw data or indicator data.

The resource table gives all groupsets as columns with groups as rows, and 1 DE per row. This means that all groupsets are joined in when e.g. creating a view to a pivot table. Resource tables have to be generated from the Maintenance->Data Administration->Resource Table window, and more information on this is available in a later chapter called Data Administration in DHIS2 or from the inline help text in that window.

27.4. The time (period) dimension

The period dimension becomes an important factor when analysing data over time e.g. when looking at cumulative data, when creating quarterly or annual aggregated reports, or when doing analysis that combines data with different characteristics like monthly routine data, annual census/population data or six-monthly staff data.

27.4.1. Period Types

In DHIS2 the periods are organised according to a set of fixed period types: 1) Daily, 2) Weekly, 3) Monthly, 4) Quarterly, 5) Six-monthly, 6) Yearly, 7) Two-yearly, and 8) a special period type called Relative.

As a rule of thumb all organisation units should collect the same data using the same frequency or periodicity, so first of all the periods play an important role in standardising data collection across the country. A data entry form therefore needs to know its period type to make sure data is always collected according to the correct and same periodicity across the country.

It is possible however to collect the same data elements using different period types by assigning the same data elements to multiple data sets with different period types, however then it becomes crucial to make sure no orgunit is collecting data using both data sets/period types as that would create overlap and duplication of data values. If set up correctly the aggregation service in DHIS2 (through datamart and report tables) will aggregate the data together, e.g. the monthly data from one part of the country with quarterly data from another part of the country into a national quarterly report. For simplicity and to avoid data duplication we strongly advice to use the same period type for all orgunits for the same data elements when possible.

27.4.2. Relative periods

When creating reports within the DHIS (report tables, standard reports, charts) it is possible to make use of the relative periods functionality. The simplest scenario is when you want to design a monthly report that can be reused every month without having to make changes to the report template to accommodate for the changes in period. The relative period called reporting_month allows for this, and the user can at the time of report generation through a report parameter select the month to use in the report.

A slightly more advanced use case is when you want to make a monthly summary report for immunisation and want to look at the data from the current (reporting) month together with a cumulative value for the year so far. The relative period called "So far this year" provides such a cumulative value relative to the reporting month selecting when running the report. Other relative periods are the last 3,6, 9 or 12 months periods which are cumulative values calculated back from the selected reporting month. If you want to create a report with data aggregated by quarters (the ones that have passed so far in the year) you can select "Individual quarters this year". Other relative periods are described under the reporting table section of the manual. Common for all the relative periods is that they are relative to a selected reporting month. Even quarterly or annual reports need to know their reporting month to derive the year, the quarter and so on. The reporting month then becomes one of the report parameters the users have to select when running a report based on relative periods.

Table 27.5.

Organisation Unit	Data Element	Reporting month	So far this year	Reporting month name
Gerehun CHC	Measles doses given	15	167	Oct-09
Tugbebu CHP	Measles doses given	17	155	Oct-09

27.4.3. Aggregation of periods

While data needs to be collected on a given frequency to standardise data collection and management, this does not put limitations on the period types that can be used in data analysis and reports. Just like data gets aggregated up the organisational hierarchy, data is also aggregated according to a period hierarchy, so you can create quarterly and annual reports based on data that is being collected on a Monthly basis. The defined period type for a data entry form (data set) defines the lowest level of period detail possible in a report.

27.4.3.1. Sum and average aggregation along the period dimension

When aggregating data on the period dimension there are two options for how the calculation is done; 1) sum and 2) average, which is specified per data element in the DHIS2 through the use of the 'aggregation operator' attribute in the Add/Edit Data Elements window.

Most of the data collected on a routinely basis should be aggregated by summing up the months or weeks, e.g. you create a quarterly report on Measles immunisation by summing up the three monthly values for "Measles doses given".

Other types of data that are more permanently valid over time like "Number of staff in the PHU" or an annual population estimate of "Population under 1 year" need to be aggregated differently. These values are static for all months as long as there are valid data. E.g. the estimated population under 1 calculated from the census data is the same for all months in a year, or the number of nurses working in a PHU is the same for every month in the 6 months period the number is reported for.

This difference e.g. becomes important when calculating an annual value for the indicator morbidity service burden for a PHU. The monthly headcounts are summed up for the 12 months to get the annual headcount while the number of staff for the PHU is calculated as the average of the two 6-monthly values reported through the 6-monthly staff report. So in this example the data element "OPD headcount" would have the aggregation operator "sum" and the data element "Number of staff" would have it set to "average".

Another important feature of average data elements is the validity period concept. Average data values are standing values for any period type within the borders of the period they are registered for. E.g. an annual population estimate following the calendar year will have the same value for any period that falls within that year no matter what the period type. E.g. if the population under 1 for a Tugbebu CHP is 250 for the year of 2009 that means that the value will be 250 for Jan-09, for Q3-09, for Week 12 of 2009 and for any period within 2009. This has implications for how e.g. coverage indicators are calculated as the full annual population will be used as denominator value even when doing monthly reports. If you want to look at an estimated annual coverage value for a given month then you will have the option of setting the indicator to "Annualised" which means that a monthly coverage value will be multiplied by 12, a quarterly value by 4 etc. The annualised indicator feature can therefore be used to mimic the use of monthly population estimates.

27.5. Data collection vs. data analysis

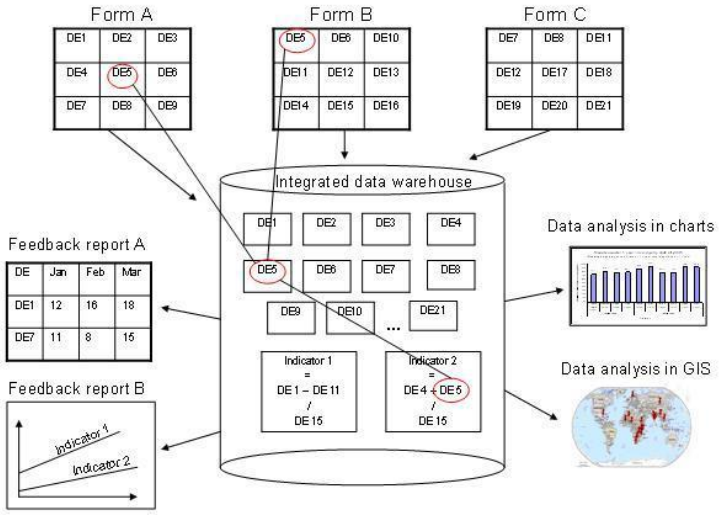
27.5.1. Data collection and storage

The datasets determine what raw data that is available in the system as they describe how data is collected. Through the data sets we define the building blocks of the data to be captured and stored in the DHIS. For each data dimension we decide what level of detail the data should be collected on; 1) the data element (e.g. diagnosis, vaccine, or any event taking place) and its categories (e.g. age and gender), 2) the period/frequency dimension, and 3) the orgunit dimension. For any report or data analysis you can never retrieve more detailed data than what is defined in the datasets so the design of the datasets and their corresponding data entry forms (the data collection tools) dictate what kind of data analysis that will be possible.

27.5.2. Input != Output

It is important to understand that the data entry forms or datasets themselves are not linked to the data (values) and that the meaning of data (the What) is only described by the data element (and its categories). This makes it perfectly safe to modify datasets and forms without altering the data (as long as the data elements stay the same). This loose coupling between forms and data makes DHIS flexible when it comes to designing and changing new forms and in providing exactly the form the users want.

Another benefit of only linking data to data elements and not to forms is the flexibility of creating indicators and validation rules based on data elements, and also in providing any kind of output report (in pivot tables, charts, maps etc) that can combine data individually or across forms, e.g. to correlate data from different health programs. Due to this flexibility of enabling integration of data from various programs (forms) and sources (routine and semi permanent (population, staff, equipment)) a DHIS database is used as an integrated data repository for many or all parts of the aggregated data in a larger HIS. The figure below illustrates this flexibility.



27.6. Some more examples

The table below combines data element the two group sets Diagnosis (all the diseases) and Morbidity/Mortality (New cases, Follow-ups, Referrals, Deaths) with the data element category PHU/Community. Deaths are captured in a separate form with other dimensions (e.g. the PHU/Community) than morbidity.

Diagnosis	New cases	Follow-up	Referrals	Deaths		Total
				PHU	*Community*	
Malaria	19	9			1	29
Anaemia	1	1	4		23	29
All Others	1	1	6	5	5	18
Diarrhoea With Blood (Dysentery)	13				1	14
Yellow Fever	3	3		7		13
Wounds/Trauma				11		11
Tetanus					11	11
Lassa Fever				11		11
Measles		1	5	2		8
Etc. etc.						

This output table combines the two data element categories HIV_Age and Gender with the data element group set ART Group. The group enables subtotals for staging and entry points summing up the data elements in that group. Subtotals for either age groups and gender would be other possible columns to easily include here.

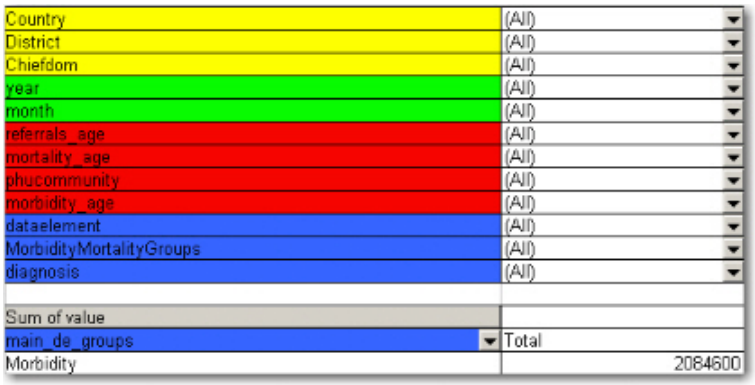
ART enrollment		Female				Male				Tot
ART Group	Data Element	<15y	15-24y	25-49y	>49 y	<15	15-24	25-49	>49	
ART staging	ART enrollment stage 1	1	8	3	11	1	1	2	1	28
	ART enrollment stage 2	1	2	15	2	2	1	3	1	27
	ART enrollment stage 3	11	4	8	1	1	2	12	2	41
	ART enrollment stage 4	11	1	1	1	1	1	11	1	28
ART staging total		24	15	27	15	5	5	28	5	124
ART entry points	ART entry point: Old patients	1	10	10	1	1	1	15	2	41
	ART entry point: Other	1	11	1	1	1	1	5	11	32
	ART entry point: Transfer in	11	1	3	1	1	11	4	1	33
	ART entry point: Walk in	1	3	15	3	2	1	7	3	35
	ART entry point: PMTCT	1	1	1	1	1	1	1	1	8
	ART entry point: transfer out	1	1	1	1	11	1	1	1	18
ART entry points total		16	27	31	8	17	16	33	19	167

27.7. How this works in pivot tables

When doing data analysis in Excel pivot tables or any other OLAP based tool the dimensions become extremely powerful in providing many different views into the data. Each data element category or group set become a pivot field, and the options or groups become values within each of these fields. In fact categories and groupsets are treated exactly the same way in pivot tables, and so are orgunits, periods, and data elements. All these become dimensions to the data value that can be used to rearrange, pivot, filter, and to drill down into the data. Here we will show some examples of how the data dimensions are used in pivot tables.

Using the example of morbidity and mortality data, a pivot table can show how the dimensions can be used to view data for different aggregation levels.

The completely aggregated number is viewed when none of the pivot fields are arranged in the table area, as column or row fields, but are listed above the table itself as page field (filter).



Here we have selected to look at the Morbidity total. The various data elements on morbidity have been ordered into the main_de_groups Morbidity (we will get back to Mortality later). The fields above the table itself are all set to "All", meaning that the totals in the table will contain data from all Countries, Districts, Chiefdom, ou_type, year, months, the various categories as listed in the red fields, and all data elements in the Morbidity group.

As we have seen, this is not a very useful representation, as Morbidity is organized into new cases, follow-ups, referrals, and then again in age groups. Also, we do not see the various diagnoses. The first step is to include the diagnoses field (which is a group set), which is done by dragging the "diagnosis" field down to be a row field, as shown in the figure below, and to add the group set called "morbiditymortality" in the column field to display new cases, follow-up, and referrals.

Contrast this figure above to the one below.

Country	(All)				
District	(All)				
Chiefdom	(All)				
year	(All)				
month	(All)				
referrals_age	(All)				
mortality_age	(All)				
phucommunity	(All)				
morbidity_age	(All)				
diagnosis	(All)				
Sum of value		MorbidityMortalityGroups			
main_de_groups	dataelement	New cases	Follow-up	Referrals	Grand Total
Morbidity	Acute Flaccid Paralysis (AFP) follow-up		12		12
	Acute Flaccid Paralysis (AFP) new	269			269
	Acute Flaccid Paralysis (AFP) referrals			1	1
	All other follow-ups		3		3
	All other new	628441			628441
	All other referrals			2	2
	Anaemia follow-up		4		4
	Anaemia new	14			14
	Anaemia referrals			2	2
	ARI treated with antibiotics (pneumonia) follow-up		4		4
	ARI treated with antibiotics (pneumonia) new	4			4
	ARI treated with antibiotics (pneumonia) referrals			2	2
	ARI treated without antibiotics (cough) follow-up		14		14
	ARI treated without antibiotics (cough) new	4			4

They both show the same data (some of the rows have been cut in the screenshot due to image size), albeit in a different way.

- The "dataelement" field, used in the bottom figure, displays each diagnosis as three elements; one follow-up, one new, and one referrals. This is the way the data elements have been defined in DHIS, as this makes sense for aggregation. You would not like to aggregate follow-ups and new, thus these have not been made as categories, the whole point of is to ease aggregation and disaggregation.
- The "diagnosis" group set has instead been made to lump these three (follow-up, new, referrals) together, which can then be split with another group set, namely the one called "morbiditymortality". This allows us to organize the data as in the first of the two figures, where we have the single diagnosis per row, and the groups new, follow-up, referrals as rows.

The idea of using group sets is that you can combine, in any set, different data elements. Thus, if we add the mortality data (by checking it from the drop-down menu of the main_de_groups field, and moving this field out of the table) we can see also the deaths, since the mortality data elements have been included as a "death" group in the "morbiditymortality" group set. The result is shown below.

Country	(All)			
District	(All)			
Chiefdom	(All)			
year	(All)			
month	(All)			
referrals_age	(All)			
mortality_age	(All)			
phucommunity	(All)			
morbidity_age	(All)			
main_de_groups	(Multiple Items)			
dataelement	(All)			
Sum of value	MorbidityMortalityGroups			
diagnosis	New cases	Follow-up	Referrals	Deaths
Acute Flaccid Paralysis (AFP)	269	12	1	
All Others	628441	3	2	1553
Anaemia	14	4	2	30
ARI Treated With Antibiotics (Pneumonia)	352274	4	2	353
ARI Treated Without Antibiotics (Cough)	4	14		
Burns	4	4	2	
Cholera	319			
Clinical Malnutrition	93091	34	2	176
Diarrhoea With Blood (Dysentery)	29341	4	12	79
Diarrhoea With Severe Dehydration	102812	4	2	208
Diarrhoea Without Severe Dehydration	4	4		
Eye Infection	4	4	2	

The result is a much more user-friendly pivot table. Now, another figure shows the relationship between the group sets and elements (these are fake data values).

This small detail of the pivot table show how the actual data elements link to the group sets:

- The four data elements, as defined in DHIS, are Measles death, Measles follow-up, Measles new, and Measles referrals
- They all belong to the group set "diagnosis", where they have been lumped together in the group Measles
- The group set "morbiditymortality" contains the groups New cases, Follow-up, Referrals, and Deaths.
- Only the data element Measles deaths has data related to the group Deaths, thus this is where the data value (20) is shown, at the upper right corner. The same for Measles new; the value (224) is shown at the intersection of the data element Measles new and the group New cases (in the group set morbiditymortality)
- All the intersections where the data element does not link with the groups in morbiditymortality are left blank. Thus in this case we would get a nice table if we excluded the dataelement from the table, and just had diagnosis and the group set morbiditymortality, as in the figure shown earlier

Now lets see how the data element categories can be used. In the data entry form for Morbidity the new cases and follow-ups use one age category, the referral data another,, and the mortality data a third age breakup, so these are available as three individual age group fields in the pivot tables called morbidity_age, referrals_age and mortality_age. It doesn't make sense to use these while looking at these data together (as in the examples above), but e.g. if we only want to look at the only the new cases we can put the MobidityMortalityGroups field back up as a page field and there select the New cases group as a filter. Then we can drag the Morbidity_age field down to the column area and we get the following view:

Country	(All)			
District	(All)			
Chiefdom	(All)			
year	(All)			
month	(All)			
referrals_age	(All)			
mortality_age	(All)			
phucommunity	(All)			
main_de_groups	(All)			
dataelement	(All)			
MorbidityMortalityGroups	New cases			
Sum of value	morbidity_age			
diagnosis	0-11m	12-59m	5-14y	15y+
Acute Flaccid Paralysis (AFP)	1	76	29	163
All Others	1	129705	53752	444983
Anaemia	1	1	11	1
ARI Treated With Antibiotics (Pneumonia)	1	204517	45752	102004
ARI Treated Without Antibiotics (Cough)	1	1	1	1
Burns	1	1	1	1
Cholera		88	104	127
Clinical Malnutrition	1	71259	11970	9861
Diarrhoea With Blood (Dysentery)	1	6868	4334	18138
Diarrhoea With Severe Dehydration	1	56345	15375	31091

The following table illustrates the benefits of reusing data element categories across datasets and categorycombinations. The VCCT, ART and PMTCT data are collected in three different datasets, the first two with both gender and age breakdown, and the PMTCT only age (gender is given). All three share the same age groups and therefore it is possible to view data elements from all these three datasets in the same table and use the age dimension. In the previous example with morbidity and mortality data this was not possible since new cases, referrals and deaths all have different age groups.

Country	(All)					
District	(All)					
Chiefdom	(All)					
year	(All)					
month	(All)					
anc_age	(All)					
pmtct	(All)					
Gender	(All)					
Sum of value		hiv_age				
main_de_groups	dataelement	<15y	15-24y	25-49y	>49y	Grand Total
ART	ART enrollment stage 1		7	3		10
	ART enrollment stage 2	1	2	7	1	11
	ART enrollment stage 3		6	23	2	31
	ART new clients started on ARV		6	24	4	34
	ART treatment stopped due to death			5		5
	ART treatment stopped due to loss to follow-up			5	1	6
	ART enrollment stage 4			1		1
	ART clients with new adverse clinical event			1	2	3
	ART entry point: No walk in	1	3	13	5	22
	ART No started Opportunist Infection prophylaxis		8	16	3	27
	ART entry ponit: No diagnostic testing		3	11		14
	ART entry point: No transfer in			5		5
	ART entry point: No old patients	6	9	27	1	43
	ART entry point: No other			4		4
	ART No clients who stopped TRT due to TRT failure			2		2
PMTCT	PMTCT HIV positive w. received ART for own health	1	136	118	1	256
	PMTCT HIV positive women received ZDV at ANC	1	199	193		393
	PMTCT male partners tested for HIV	30	1292	1277	25	2624
	PMTCT male partners tested HIV positive	1	74	82	1	158
	PMTCT No HIV women who delivered in the health facility (Normal)		57	86		143
	PMTCT reported mother death		4	5		9
	PMTCT women counselled for infant feeding	246	20791	15992	143	37172
	PMTCT women HIV positve	3	339	275		617
	PMTCT women HIV1	2	302	242	2	548
	PMTCT women HIV1 and HIV2		26	29		55
	PMTCT women HIV2		13	9	2	24
	PMTCT women lost to contact	4	71	67	6	148
	PMTCT women received complete ARV for PMTCT		77	93		170
	PMTCT women received NVP during labour		55	75		130
	PMTCT women received post test counselling and result	197	14901	10647	86	25831
	PMTCT women received ZDV & 3CT after delivery		96	118		214
	PMTCT women seen for 1st ANC	262	15838	11376	173	27649
	PMTCT women tested for HIV	228	15000	10813	78	26119
VCCT	VCCT No positive test HIV1 and HIV2	3	11	24	4	42
	VCCT No positive test HIV1 only	18	127	333	35	513
	VCCT No positive test HIV2 only		4	9		13
	VCCT No receiving positive test results	36	374	710	88	1208
	VCCT No receiving Pre-test counselling	810	3878	5009	683	10380

In the table below PMTCT data has been removed from the table and the gender category added to the column area so that you can analyse the data for VCCT and ART by age and gender. An optional subtotal for gender has also been added, as well as a grand total for all age and gender.

Country	(All)											
District	(All)											
Chiefdom	(All)											
year	(All)											
month	(All)											
anc_age	(All)											
pmtct	(All)											
Sum of va		Gender	hiv_age									
		Female				Female Total	Male				Male Total	Grand Total
main_d	dataelement	<15y	15-24y	25-49y	>49y		<15y	15-24y	25-49y	>49y		
ART	ART enrollment stage 1			7	2	9			1		1	10
	ART enrollment stage 2			2	5	1	8	1		2		3	11
	ART enrollment stage 3			4	10	1	15		2	13	1	16	31
	ART new clients started on ARV			4	13	2	19		2	11	2	15	34
	ART treatment stopped due to death				2		2			3		3	5
	ART treatment stopped due to loss to follow-up				4	1	5			1		1	6
	ART enrollment stage 4									1		1	1
	ART clients with new adverse clinical event				1	2	3						3
	ART entry point: No walk in			2	5	3	10	1	1	8	2	12	22
	ART No started Opportunist Infection prophylaxis			7	9	3	19		1	7		8	27
	ART entry ponit: No diagnostic testing			2	7		9		1	4		5	14
	ART entry point: No transfer in				2		2			3		3	5
	ART entry point: No old patients	6	9	9		24				18	1	19	43
	ART entry point: No other									4		4	4
	ART No clients who stopped TRT due to TRT failure				1	1				1		1	2
VCCT	VCCT No positive test HIV1 and HIV2	2	8	13	2	25	1	3	11	2	17	42
	VCCT No positive test HIV1 only	7	88	187	22	304	11	39	146	13	209	513
	VCCT No positive test HIV2 only		3	4		7		1	5		6	13
	VCCT No receiving positive test results	15	218	367	49	649	21	156	343	39	559	1208
	VCCT No receiving Pre-test counselling	452	2342	2737	352	5883	368	1536	2272	331	4497	10380
	VCCT No reiciving result & post-test counselling	401	1901	2221	267	4790	317	1236	1891	267	3711	8501
	VCCT No Test	449	2262	2697	303	5601	357	1469	2193	296	4315	9916

27.8. From paper for to multidimensional datasets - lessons learned

Typically the design of a DHIS 2 dataset is based on some requirements from a paper form that is already in use. The logic of paper forms are not the same as the data element and data set model of DHIS, e.g. often a field in a tabular paper form is described both by column headings and text on each row, and sometimes also with some introductory table heading that provides more context. In the database this is captured in one atomic data element with no reference to a position in a visual table format, so it is important to make sure the data element with the optional data element categories capture the full meaning of each individual field in the paper form.

Another important thing to have in mind while designing datasets is that the dataset and the corresponding data entry form (which is a dataset with layout) is a data collection tool and not a report or analysis tool. There are other far more sophisticated tools for data output and reporting in DHIS than the data entry forms. Paper forms are often designed with both data collection and reporting in mind and therefore you might see things such as cumulative values (in addition to the monthly values), repetition of annual data (the same population data reported every month) or even indicator values such as coverage rates in the same form as the monthly raw data. When you store the raw data in DHIS every month

and have all the processing power you need within the computerised tool there is no need (in fact it would be stupid and most likely cause inconsistency) to register manually calculated values such as the ones mentioned above. You only want to capture the raw data in your datasets/forms and leave the calculations to the computer, and presentation of such values to the reporting tools in DHIS.

27.8.1. From tables to category combinations - designing multidimensional data sets

As we have seen in the examples above the data element categories and their options are helpful in representing tabular data, when adding dimensions to a field in a paper form. We have also seen how the data element is the central dimension and that the data element categories are used to provide further details or disaggregation to the data element. As we will see in the example below there are often more than one way to represent a paper form in DHIS, and it can be difficult to know which dimension to represent with a data element name and which to represent as categories, or even as groups as we have seen above. Here are some general lessons learned from working with data element and category combinations:

- Design your dimensions with data use in mind, not data collection. This means that disaggregation of data values at collection time should be easily aggregated up along the various dimensions, as in adding up to a meaningful total.
- Reuse dimensions as much as possible as it increased the ability to compare disaggregated data (e.g. age groups, fixed/outreach, gender). Not necessary to share all dimensions, it helps to share only one as well (much better than none).
- Disaggregation dimensions should add up to a total
- Different levels of dimensions; 1) disaggregation and 2) grouping. Disaggregation dimensions dictate how you collect and how detailed you store your data, so plan these carefully. The group dimension is more flexible and can be changed and added to even after data collection (think of it as tagging).
- Think integrated data repository and not forms or programs when designing the metadata model and revising forms. Use the same disaggregation for the same or similar data across forms. Reuse definitions so that the database can integrate even though the forms might be duplicating.

STEP BY STEP APPROACH TO DESIGNING DATASETS

1. Identify the different tables (or sub datasets) in the paper form that share the same dimensions

2. For each table identify the dimensions that describe the data fields

3. Identify the key dimension, the one that makes most sense to look at in isolation (when the others are collapsed, summed up). This is your data element dimension, the starting point and core of your multidimensional model (sub dataset). The data element dimension can be a merger of two or more dimensions if that makes more sense for data analysis. The key is to identify which total that makes most sense to look at alone when the other dimensions are collapsed.

4. For all other/additional dimensions identify their options, and come up with explanatory names for dimensions and their options.

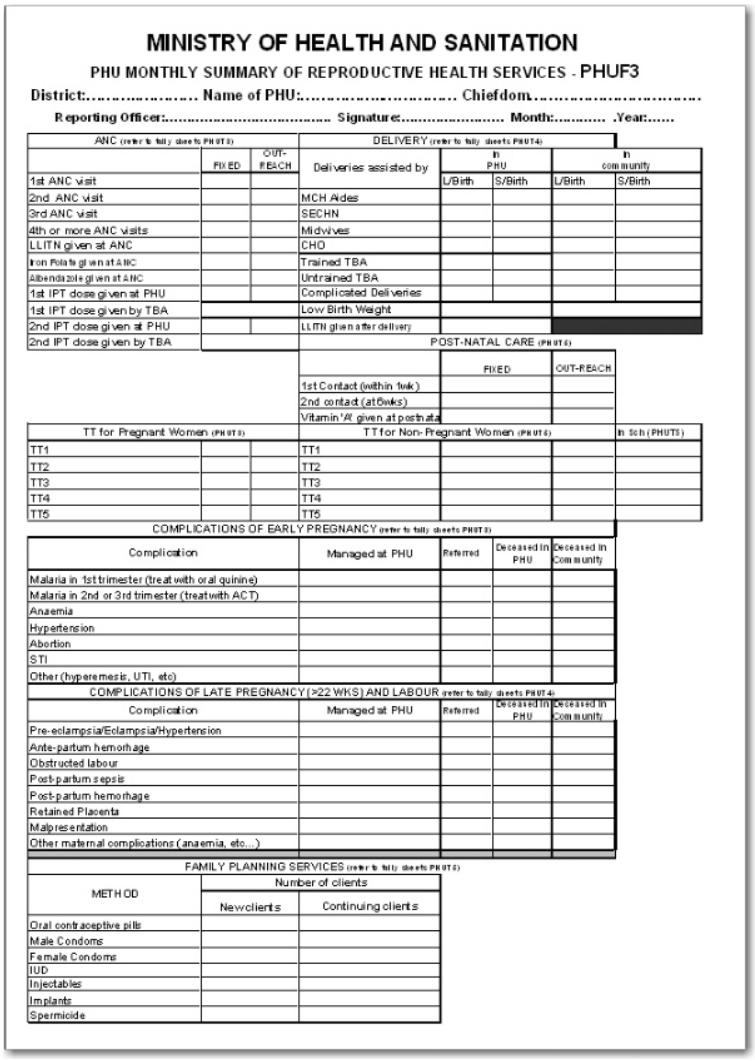
5. Each of these additional dimensions will be a data element category and their options will be category options.

6. Combine all categories for each sub dataset into one category combination and assign this to all the data elements in your table (or sub dataset if you like).

7. When you are done with all the tables (sub datasets), create a new dataset and add all the data elements you have identified (in the whole paper form) to that dataset.

8. Your dataset will then consist of a set of data elements that are linked to one or more category combinations.

In order to better explain the approach and the possibilities we present an example paper form and will walk through it step by step and design data elements, categories, category options and category combinations.



MINISTRY OF HEALTH AND SANITATION

PHU MONTHLY SUMMARY OF REPRODUCTIVE HEALTH SERVICES - PHUF3

District:..... Name of PHU:..... Chiefdom.....

Reporting Officer:..... Signature:..... Month:..... .Year:.....

ANC (refer to tally sheets PHUT3)

	FIXED	OUT-REACH
1st ANC visit		
2nd ANC visit		
3rd ANC visit		
4th or more ANC visits		
LLITN given at ANC		
Iron Folate given at ANC		
Albendazole given at ANC		
1st IPT dose given at PHU		
1st IPT dose given by TBA		
2nd IPT dose given at PHU		
2nd IPT dose given by TBA		

DELIVERY (refer to tally sheets PHUT4)

Deliveries assisted by	In PHU		In community	
	L/Birth	S/Birth	L/Birth	S/Birth
MCH Aides				
SECHN				
Midwives				
CHO				
Trained TBA				
Untrained TBA				
Complicated Deliveries				
Low Birth Weight				
LLITN given after delivery				

POST-NATAL CARE (PHUT4)

	FIXED	OUT-REACH
1st Contact (within 1wk)		
2nd contact (at 6wks)		
Vitamin 'A' given at postnata		

TT for Pregnant Women (PHUT3)			TT for Non-Pregnant Women (PHUT4)			In Sch (PHUT5)
TT1			TT1			
TT2			TT2			
TT3			TT3			
TT4			TT4			
TT5			TT5			

COMPLICATIONS OF EARLY PREGNANCY (refer to tally sheets PHUT3)

Complication	Managed at PHU	Referred	Deceased in PHU	Deceased in Community
Malaria in 1st trimester (treat with oral quinine)				
Malaria in 2nd or 3rd trimester (treat with ACT)				
Anaemia				
Hypertension				
Abortion				
STI				
Other (hyperemesis, UTI, etc)				

COMPLICATIONS OF LATE PREGNANCY (>22 WKS) AND LABOUR (refer to tally sheets PHUT4)

Complication	Managed at PHU	Referred	Deceased in PHU	Deceased in Community
Pre-eclampsia/Eclampsia/Hypertension				
Ante-partum hemorhage				
Obstructed labour				
Post-partum sepsis				
Post-partum hemorhage				
Retained Placenta				
Malpresentation				
Other maternal complications (anaemia, etc...)				

FAMILY PLANNING SERVICES (refer to tally sheets PHUT4)

METHOD	Number of clients	
	Newclients	Continuing clients
Oral contraceptive pills		
Male Condoms		
Female Condoms		
IUD		
Injectables		
Implants		
Spermicide		

This form has many tables and each of them potentially represent a data element category combination (from now on referred to as a catcombo). As such there is no restriction on a dataset to only have one set of dimensions or catcombo, it can have m For any and as we see above this is necessary as the dimensions are very different from table to table. We will walk through this table by table and discuss how to represent it in the DHIS.

ANC table. This table in the top left corner is one the simpler ones in this form. It has two dimensions, the first column with the ANC activity or service (1st visit, IPT 2nd dose etc) and the 2nd and 3rd column which represent the place where the service was given with the two options fixed and outreach. Since the ANC service is the key phenomena to analyse here and often there is a need for looking at e.g. total of ANC 1st visits no matter where (fixed+outreach) it makes a lot of sense to use this dimension as the data element dimension. So all items on the first column from 1st ANC visit to 2nd IPT dose given by TBA are represented as individual data elements. The place dimension is represented as a data element category (from now on referred to as category) with the name "fixed/outreach" with the two data element category options (from now on catoptions) "fixed" and "outreach". There is no other dimension here so we add a new catcombo with the name "Fixed/Outreach" with one category "Fixed/Outreach". Strictly speaking there is another dimension in this table, and that is the at PHU or by TBA dimension which is repeated for the two doses of IPT, but since none of the other ANC services listed have this dimension it does not seem like a good idea to separate out two data elements from this table and give them another catcombo with both fixed/outreach and at PHU/by TBA. reusing the same catcombo for all the ANC services makes more sense since it will be easier to look at these together in reports etc. and also the fact that there is not much to loose by repeating the at PHU or by TBA information as part of the data element name when it is only for four data elements in a table of totally 11 data elements.

DELIVERY table. This table is more tricky as it has a lot of information and you can see that not all the rows have the same columns (some columns are merged and a one field is grayed out/disabled.). If we start by looking at the first column "Deliveries assisted by" that seems to be one dimension, but only down to the "Untrained TBA" row, as the remaining three rows are not related to who assisted the delivery at all. Another dimension is the place of delivery, either In PHU or in Community as stated on the top column headings. These deliveries are further split into the outcome

of the delivery, whether it is a live or still birth, which seems to be another dimension. So if we disregard the three bottom rows for a moment there seems to be 3 dimensions here, 1) assisted by, 2) place of delivery, and 3) delivery outcome. The key decision to make is what to use as the data element, the main dimension, the total that you will most often use and want easily available in reports and data analysis. We ended up using the outcome dimension as total live births is a very commonly used value in many indicators (maternal mortality ratio, births attended by skilled health personnel etc.). In this case the Assisted By dimension could also have been used without any problem, but the added value of easily getting the total live births information was the decisive point for us. This means that from this table (or sub-table of row 1 to 6) there are only two data elements; "Live births" and "Still births". Then there are two more dimensions, the "PHU/Community" with its two options and a "Births attended by" with options ("MCH Aides", "SECHN", "Midwives", "CHO", "Trained TBA", "Untrained TBA"). These two categories make up the catcombo "Births" which is assigned to the two data elements "Live births" and "Still births". Considering the final three rows of the delivery table we can see that "Complicated Deliveries" does not have the assisted by dimension, but has the place and the outcome. "Low birth weight" also does not have the assisted by dimension and not the outcome either. The LLITN given after delivery does not have any additional dimension at all. Since not any of the three rows can share catcombo with any other row we decided to represent these fields as so called flat data elements, meaning data elements with no categories at all, and simply adding the additional information from the column headings to the data element name, and therefore ended up with the following data elements with the default (same as none) catcombo; "Complicated deliveries in PHU live birth", "Complicated deliveries in PHU still births", "Complicated deliveries in community live birth", "Complicated deliveries in community still births", "Low birth weight in PHU", "Low birth weight in community", and "LLITN given after delivery".

POST-NATAL CARE table This table is simple and we used the same approach as for the ANC table. 3 data elements listed in the first column and then link these to the catcombo called "fixed/outreach". Reusing the same category fixed/ outreach for these data elements enables analysis on fixed/outreach together with ANC data and other data using the same category.

TT table This is a bit more tricky. We decided to use "TT1", "TT2" ... "TT5" as data elements which makes it easy to get the total of each one of these. There is fixed/outreach dimension here, but there is also the In school place that is only applied to the Non-Pregnant, or more correctly to any of the two as the school immunisation is done whether the girls are pregnant or not. We consulted the program people behind the form and found out that it would be OK to register all school TT immunisations as non-pregnant, which simplifies the model a bit since we can reuse the "TT1" to "TT5" data elements. So we ended up with a new category called "TT place" with the three options (Fixed, Outreach, In School), and another category called "Pregnant/Non-pregnant" with two options. The new catcombo "TT" is then a combination of these two and applied to the 5 TT data elements. Since we agreed to put all In Schools immunisations under Non-pregnant in means that the combination of options (Pregnant+In School) will never be used in any data entry form, and hence become a possible optioncombo, which is OK. As long as the form is custom designed then you can choose which combinations of options to use or not, and therefore it is not a problem to have such passive or unused catoptions. Having school as one option in the TT place category simplifies the model and therefore we thought it was worth it. The alternative would be to create 5 more data elements for "TT1 in school" ... "TT5 in school", but then it would be a bit confusing to add these together with the "TT1" ..."TT5" plus TT catcombo. Having school as a place in the TT place category makes it a lot easier to get the total of TT1.. TT5 vaccines given, which are the most important numbers and most often used values for data analysis.

Complications of early and late pregnancy and labour tables We treat these two tables as one, and will explain why. These two tables are a bit confusing and not the best design. The major data coming out of these tables are the pregnancy complications and the maternal deaths. These are the major things for data analysis. And then there is further detail on the cause of the complication or death (the first column in both tables), as well as a place of death (in PHU or community), and a outcome of the complication (when its not a death) that can be either Managed at PHU or Referred. We decided to create two data elements for these two tables; "Pregnancy complications", and "Maternal Deaths", and two category combinations, one for each of the data elements. For the Pregnancy Complications data element there are two additional dimensions, the cause of the complication (the combined list of the first column in the two tables) and the outcome (managed at PHU or Referred), so these are the categories and options that make up that category combination. For the "Maternal deaths" data element the same category with the different causes are used and then another category for the place of death (in PHU or In community). This way the two data elements can share one category and it will be easy to derive the total number of pregnancy complications and maternal deaths. While the list of complications on the paper form is divided into two (early and late/labour) you can see that e.g. the malaria in 2nd and 3rd trimester are listed under early, but in fact are for a later phase of the pregnancy. There is no clear divide between early and late complications in the form, and therefore we gave up trying to make this distinction in the database.

Family Planning Services table This table has 2 dimensions, the family planning method (contraceptive) and whether the client is new or continuing. We ended up with one data element only "Family planning clients" and then added two categories "FP method" with all the contraceptives as options, and another category "FP client type" with new or continuing as options. This way it will be easy to get the total number of family planning clients which is the major value to look at in data analysis, and from there you can easily get the details on method or how many new clients there are.

Chapter 28. User profiles

With profiles, users can alter the appearance of their DHIS2, change the interface language, and enabled SMS and email notifications. Users can also enter information into their "Profile".

28.1. User settings

To access the user settings menu, select "Profile->Settings". You can select the interface language and the database language. The interface language refers to the translation of the user interface of the DHIS2 software. The database language refers to the translated content of the metadata contained in the system, for instance, data elements and indicators. Different interface styles can be set to alter the apperarnce of the system.

User can choose to receive their messages via email and/or SMS by ticking the appropriate check boxes.

28.2. User profile

Users can choose to enter more data, such as their email and phone number, job title, gender, birthday, etc into their profile. Just select "Profile->Profile" from the main menu.

DHIS2 Glossary

A

Aggregation	In the context of DHIS2, aggregation refers to how data elements are combined within a particular hierarchical relationship. As an example, all the academic departments in a particular school would contribute to the total value for the particular school in question. Different aggregation operators are supported within DHIS2, such as SUM, AVERAGE, and COUNT.
Aggregate data	In the context of DHIS2, aggregate data refers to either data elements or indicators that have been derived from other hierarchical data sources. For instance, aggregate departmental data would result from the aggregate totals of all students that have attended courses that department for a particular training program. Aggregate district data would result from the aggregate totals of all academic departments in that particular training program.
Application programming interface	An application programming interface is a specification of how different software components should interact with each other. The DHIS2 API (or WebAPI) can be used to interface DHIS2 with other software, to build reports or custom data entry forms.

B

Bimonthly	Refers to a two-month period, such as January 1st to February 28th.

C

Category	Categories are groups of category options. They are used in combinations to disaggregate data elements. Categories are typically a single type of concept, such as "Age" , "Gender","Race","Geographic Location", "Employment status", "Socio-economic status".
Category combinations	Category combinations are used to disaggregate data elements. As an example, the data element "Number of confirmed cases of malaria" could be disaggregated subdivided into categories: "Age" and "Gender". In turn each of these categories, would consist of several category options, such as "Male" and "Female" for the gender category, “black", “white"and “mixed race"for race category; Urban" and “rural" for geographical location category;" income or expenditure quintiles for socio-economic status category; “Diploma", “Bachelors", “Masters", “PhD"for level of study category; “lecturer", “senior lecturer", associate professor", “professor"for academic rank category; “permanent"and “temporary" for employment status category; “medicine", “nursing", “pharmacy", “"nutrition", e.t.c. for academic program category.Category combinations may consist of one or several categories.
Category combination options	Category combination options are dynamically composed of all of the different combinations of category options which compose a category combination. As an example, two categories "Gender" and "Academic Rank", might have options such as "Male"/"Female" and "Professor"/"Associate Professor"/"Senior Lecturer"/"Lecturer". The category combination options would then consist of: (Male/Professor) (Male/Associate Professor) (Male/Senior Lecturer) (Male/Lecturer) (Female/Professor)

	(Female/Associate Professor) (Female/Senior Lecturer) (Female/Lecturer)
Category option	Category options are atomic elements that are grouped into categories.
Comma separated values	Comma separated values are series of tabular data stored in a plain-text format. They are commonly used with DHIS2 to export and import data values.

D

Data dictionary	A collection of data elements and indicators, which can be exchanged with other DHIS systems. Typically used to define a set of data elements and indicators when setting up the DHIS system.
Data exchange format	In the context of DHIS2, the "data exchange format" refers to a XML schema that enables the transportation of data and metadata between disconnected DHIS instances, as well as between different applications that support the DXF schema.
Datamart	A set of database tables in DHIS2 that contains processed data elements and indicator values that is generated based on aggregation rules and calculated data element and indicator formular. Datamart tables are used for analysis and report production. Typically, users should not work directly with unaggregated data values, but rather with values that have resulted from a datamart export for analysis.
Data element	A data element is the fundamental building block of DHIS2. It is an atomic unit of data with well-defined meaning. Essentially it is a data value that has been actually observed or recorded which is further characterized by a number of dimensions. Examples are: applications (head counts) per educational program in the institution; enrollment (head counts) per educational program in the institution; course repeat rate; year repeat rate; attrition rate; retention rate; success rate; graduation rate; throughput rate; certification ratehead count of academic staff; headcount of administrative staff; academic staff qualification; institutional unit income; institutional unit expenditure; cost of training a graduate; source of financial support to students. Data elements are always linked to a period as well as an organizational unit. They optionally may be linked to other dimensions.
Data element group	Data element groups are used to categorize multiple data elements according to a common theme, such as .Student selection, admission and numbers; Student progression and achievement; Student development support and guidance; Teaching and Learning; Research; Human Resources; Budgeting and fiscal management; Alumni tracking. Typically, they are used during reporting and analysis to allow related data elements to be analyzed together.
Data element group sets	Data element groups are used to categorize multiple data element groups into a common theme.
Dimension	A dimension is used to categorize data elements during analysis. Dimensions provide a mechanism to group and filter data based on common characteristics. Typically, related data elements may be aggregated or filtered during analysis with the use of dimensions. Dimensions may be a member of a hierarchy. For instance the "Period" dimension may be broken down into "Day->Month->Quarter->Year".
DXF	See Data exchange format.

E

Education Management Information System (EMIS)	Typically, an electronic database system that is used to record aggregated data on academic performance, financial performance, infrastructure, human resource data and

other information used to evaluate the performance of higher educational institutions. Some of the commonly used educational data management systems include the enterprise resource planning (ERP) systems, the learning management systems (LMS), and the Integrated Tertiary Software (ITS). UNESCO has developed an open-source system called Open EMIS that can be easily modified and adapted to the needs of secondary and tertiary educational institutions (https://www.openemis.org/products/school)

I

Indicator

The divisor of an indicator. Can be composed of multiple data elements with the use of an indicator formula.

$$indicator = \frac{numerator}{denominator} \quad (1)$$

This is obviously a very generalized example. The numerator and indicator themselves can be composed of various data elements, factors, and the four basic operands (addition, multiplication, division and subtraction).

N

Numerator

The dividend of a indicator. Can be composed of multiple data elements and factors with the use of indicator formulas.

O

Organisational unit

Organizational units in educational context is used to specify an administrative unit, such as an academic department within a Faculty. The organisational unit dimension specifies essentially "where" a particular data value occurs.

Organisational unit level

Refers to a level within an organizational hierarchy. Typically, higher educational institutions are administered at different levels, such as 1)college 2) Faculty 3) School 4) Academic Department. In the context of DHIS2,academic departments typically are the lowest orgunit level. Data is aggregated upwards from the lowest orgunit level to the highest.

P

Period

A period is a specific time interval which consists of a start date and end date. For instance "January 2011" would refer to the time interval of January 1st 2011-January 31st 2011.

U

Unique identifier

A unique identifier (UID) is a semi-random series of letters and numbers used by DHIS2 to identify specific resources. UIDs begin with a letter, and are followed by exactly 10 letters or digits.

Bibliography

[AlSaid2010] Said Salah Eldin Al Said. *The health information system in Sudan*. The University of Oslo. 2010. http://urn.nb.no/URN:NBN:no-27062.

[Berg2007] Eivind Anders Berg. *The challenges of implementing a health information system in Vietnam*. The University of Oslo. 2007. http://urn.nb.no/URN:NBN:no-15021.

[BraaHedeberg2002] Jørn Braa and Calle Hedberg. "The Struggle for District-Based Health Information Systems in South Africa". *Information Society*. 18. 113-127. 2002. http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=6705438&site=ehost-live.

[BraaNetworksAction2004] Eric; Braa Jørn; Monteiro and Sundeep Sahay. "Networks of Action: Sustainable Health Information Systems Across Developing Countries". *MIS Quarterly*. 28. 3. 2004. http://aisel.aisnet.org/misq/vol28/iss3/3/.

[Brucker2007] Øyvind F Brucker. *Internationalization and localization - A case study from HISP*. The University of Oslo. 2007. http://urn.nb.no/URN:NBN:no-15774.

[Damitew2005] Hirut Gebrekidan Damitew and Netsanet Haile Gebreyesus. *Sustainability and optimal use of Health Information Systems*. The University of Oslo. 2005. http://urn.nb.no/URN:NBN:no-11506.

[Jacucci06exploringtensions] Ved Anfinsen Edoardo Jacucci Cover Inger S. "EXPLORING TENSIONS IN INFORMATION SYSTEMS STANDARDIZATION Two Case Studies from Healthcare in Norway and South Africa". 2006. http://folk.uio.no/edoardo/MatNatAvh_Jacucci_rettet.pdf.

[Gjendem2008] Anders Gjendem. *Recruitment, training, communication and Open Source*. The University of Oslo. 2008. http://urn.nb.no/URN:NBN:no-19821.

[Gjerull2006] Nils Fredrik Gjerull. *Open Source Software Development in Developing Countries*. The University of Oslo. 2006. http://urn.nb.no/URN:NBN:no-13117.

[Heldre2006] Thor Helge Heldre. *Study of a Health Information System pilot project in Tanzania*. The University of Oslo. 2006. http://urn.nb.no/URN:NBN:no-12362.

[Jacobsen2006] Petter Jacobsen. *Design and development of a global reporting solution for DHIS*. The University of Oslo. 2006. http://urn.nb.no/URN:NBN:no-12659.

[BraaStandards2007] Arthur Heywood Woishet Mohammed Vincent Shaw Jørn Braa Ole Hanseth. "DEVELOPING HEALTH INFORMATION SYSTEMS IN DEVELOPING COUNTRIES: THE FLEXIBLE STANDARDS STRATEGY". *MIS Q*. 31. 1. 2007. http://heim.ifi.uio.no/~vshaw/Files/Published%20Papers%20included%20in%20Kappa/4_Braa_Flexible%20standards.pdf.

[BraaSahayPowerToUsers] Sundeep Sahay Jørn Braa. "Integrated Health Information Architecture - Power to the Users". Matrix Publishers. 384. 2012.

[Lewis2005] John Lewis. *Design and development of spatial GIS application for primary healthcare sector*. The University of Oslo. 2005. http://urn.nb.no/URN:NBN:no-11504.

[Mangset2005] Lars Mangset. *DHIS-2 - A Globally Distributed Development Process*. The University of Oslo. 2005. http://urn.nb.no/URN:NBN:no-10640.

[Ngoma2007] Caroline Ngoma. *Cultivation Strategies in the Implementation of Health Management Information System in Zanzibar*. The University of Oslo. 2007. http://urn.nb.no/URN:NBN:no-16911.

[Nguyen2007] Thanh Ngoc Nguyen. *OSS For Health Care in Developing Countries*. The University of Oslo. 2007. http://urn.nb.no/URN:NBN:no-17859.

[Saeb2009] E.K. Golly-Kobrissa R.T. Titlestad O. Braa J. Saeb J. Kossi. "Integrating health information systems in Sierra Leone". 379 - 391. 2009.

[ShawComplexityInspried2009] Vincent Shaw. "A complexity inspired approach to co-evolutionary hospital management information systems development". 2009. http://folk.uio.no/vshaw/Files/VShaw%20Kappa%20Final%20Version/2_V_Shaw%20Intro%20Chapter_no%20annex.pdf.

[Staring_Titlestad_2008] Knut Staring and O H Titlestad. "Development as a Free Software: Extending Commons Based Peer Production to the South". *ICIS 2008 Proceedings*. 50. 2008. http://aisel.aisnet.org/icis2008/50.

[Store2007] Margrethe Store. *Explore the challenges of providing documentation in open source projects*. The University of Oslo. 2007. http://urn.nb.no/URN:NBN:no-15782.

[Storset2010] Leif Arne Storset. *Integration of Health Management Information Systems*. The University of Oslo. 2010. http://urn.nb.no/URN:NBN:no-25666.

[ShawScaling2007] Jorn Braa Vincent Shaw Shegaw Anagaw Mengiste. "Scaling of Health Information Systems in Nigeria and Ethiopia- Considering the Options". 2007. http://heim.ifi.uio.no/~vshaw/Files/Published%20Papers%20included%20in%20Kappa/6_Shaw_IFP9.4%20Scaling%20of%20HIS_Considering%20the%20Options.pdf.

[Vo2009] Kim Anh Thi Vo. *Challenges of Health Information Systems Programs in Developing Countries: SUCCESS and FAILURE*. The University of Oslo. 2009. http://urn.nb.no/URN:NBN:no-23652.

[Overland2010] Jan Henrik Øverland. *An Open Source Approach to Improving GIS Implementations in Developing Countries*. The University of Oslo. 2010. http://urn.nb.no/URN:NBN:no-24751.

[Overland2006] Lars Helge Øverland. *Global Software Development and Local Capacity Building*. University of Oslo. 2006. http://urn.nb.no/URN:NBN:no-13609.

Index