

Chapter 12. Data Sets and Forms

This chapter discusses data sets and forms, what types of forms are available and describes best practises for the process of moving from paper based to electronic forms.

12.1. What is a data set?

All data entry in DHIS 2 is organised through the use of data sets. A data set 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). Data sets are not linked directly to the data values, only through their data elements and frequencies, and as such a data set 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 data set has a period type which controls the data collection frequency, which can be daily, weekly, monthly, quarterly, six-monthly, or yearly. Both the data elements to include in the data set and the period type is defined by the user, together with a name, short name, and code. If calculated fields are needed in the collection form (and not only in the reports), then indicators can be assigned to the data set as well, but these can only be used in custom forms (see further down).

In order to use a data set to collect data for a specific organisation unit the user must assign the organisation unit to the data set. This mechanism controls which organisation units that can use which data sets, and at the same time defines the target values for data completeness (e.g. how many health facilities in a district are expected to submit the RCH data set every month).

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

12.2. What is a data entry form?

Once you have assigned a data set to an organisation unit that data set will be made available in Data Entry (under Services) for the organisation units you have assigned it to and for the valid periods according to the data set's period type. 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 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.

12.2.1. Types of data entry forms

DHIS 2 currently features three different types of forms which are described in the following.

12.2.1.1. Default forms

A default data entry form 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 category combination, 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 category combination in a data set you will get one table per category combination in the default form, with different column headings for the options.

12.2.1.2. 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 data set you can define its sections with subsets of data elements, a heading and possible grey fields in the section's table. The order of sections in a data set 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 data set). Most tabular data entry forms should be possible to do with sections forms. Utilizing the section or default forms makes life easy as there is no need to maintain a fixed form design which includes references to data elements. 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.

12.2.1.3. 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 terms of the design. In DHIS 2 there is a built in HTML editor (CK Editor) in the form designer which allows you to either design the form in the GUI 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 option combination) 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 data set it will be available in data entry and used automatically.

When using a custom form it is possible to use calculated fields to display e.g. running totals or other calculations based on the data captured in the form. This can e.g. be useful when dealing with stock or logistics forms that need item balance, items needed for next period etc. In order to do so, the user must first define the calculated expressions as indicators and then assign these indicators to the data set in question. In the custom form designer the user can then assign indicators to the form the same way data elements are assigned. The limitation to the calculated expression is that all the data elements use in the expression must be available in the same data set since the calculations are done on the fly inside the form, and are not based on data values already stored in the database.

12.3. From paper to electronic form - Lessons learned

When introducing an electronic health information system the system being replaced is often paper based reporting. The process of migrating to electronic data capture and analysis has some challenges. The following sections suggest best practises on how to overcome these.

12.3.1. Identify self-contained data elements

Typically the design of a DHIS 2 data set 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 for 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.

12.3.2. Leave calculations and repetitions to the computer - capture raw data only

Another important thing to have in mind while designing data sets is that the data set and the corresponding data entry form (which is a data set 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 2 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 the DHIS 2 database every month and have all the processing power you need within the computerised tool, there is no need (in fact it would be wrong 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 data sets/forms and leave the calculations to the computer, and presentation of such values to the reporting tools in DHIS. Through the functionality of data set

reports all tabular section forms will automatically get extra columns at the far right providing subtotal and total values for each row (data element).

