[Using expressions in your forms](https://malttanzania.surveycto.com/help.html" \l "Help_Designing-forms_core-concepts_expressions)

The easiest way to build expressions is to use the constraint-builder, the relevance-builder, and/or the calculation-builder, all of which are available both as actions in the *Your forms* section of your server's [Design](https://malttanzania.surveycto.com/main.html#Design) tab and, offline, as options in the *Tools... Form tools* menu of *SurveyCTO Sync*.

If you want to edit expressions yourself, you can refer to user responses and other fields as follows:

* For constraints, "." is used to refer to the user's proposed entry or selection for the current field (i.e., for the value you are testing to see if it's valid).
* Use "${fieldname}" to refer to a prior field's entry, selection, or calculated value.

And you can use all of the following operators in any expression:

|  |  |  |  |
| --- | --- | --- | --- |
| **Operation** | **Operator** | **Example** | **Example answer** |
| Addition | + | 1 + 1 | 2 |
| Subtraction | - | 3 - 2 | 1 |
| Multiplication | \* | 3 \* 2 | 6 |
| Division | div | 10 div 2 | 5 |
| Modulus | mod | 9 mod 2 | 1 |
| Equal | = | ${fieldname} = 3 | true or false |
| Not equal | != | ${fieldname} != 3 | true or false |
| Greater-than | > | ${fieldname} > 3 | true or false |
| >-or-equal | >= | ${fieldname} >= 3 | true or false |
| Less-than | < | ${fieldname} < 3 | true or false |
| <-or-equal | <= | ${fieldname} <= 3 | true or false |
| Or | or | ${fieldname} = 3 or ${fieldname} = 4 | true or false |
| And | and | ${fieldname} > 3 and ${fieldname} < 5 | true or false |
| Not | not() | not(${fieldname} > 3 and ${fieldname} < 5) | false or true |

And you can also call any of the following functions:

1. **string-length(field)**: Returns the length of the string field (as in the constraint "string-length(.) > 3 and string-length(.) < 10").
2. **count-selected(field)**: Returns the number of items selected in a *select\_multiple* field (as in the constraint "count-selected(.) = 3").
3. **selected(field, value)**: Returns true or false depending on whether the value indicated in the second parameter was selected in the *select\_one* or*select\_multiple* field indicated in the first parameter (as in the relevance "selected(${fieldname}, 'Male')"). Please note that the second parameter to the selected() function should always specify the internal value associated with the selection, as configured in the *choices* worksheet's *name* column. (You cannot specify the label configured in the *label* column.)
4. **selected-at(field, number)**: When the passed number is 0, returns the first selected item in a *select\_multiple* field; when the passed number is 1, returns the second selected item; etc. (as in the relevance "selected-at(${fieldname}, 0) = 'Shona'"). Note that the returned value will correspond with the internal value associated with the selection, as configured in the *choices* worksheet's *name* column.
5. **jr:choice-name(value, 'field')**: Returns the label for a *select\_one* or *select\_multiple* field choice (as in the calculate expression "jr:choice-name(${selectonefield}, '${selectonefield}')" to return the label for the currently-selected choice in the field named "selectonefield"). For *select\_multiple*fields, you can combine with selected-at() to get the label for individual selections (as in "jr:choice-name(selected-at(${selectmultfield}, 0), '${selectmultfield}')" to return the label for the first choice in the field named "selectmultfield").
6. **concat(fieldorstring, fieldorstring, ...)**: Concatenates fields (and/or strings) together (as in the calculate expression "concat(${firstname}, ' ', ${lastname})").
7. **duration()**: Returns the total amount of time spent, in seconds, filling or editing the current form submission. Call this function in a *calculate\_here* field to capture the user's number of seconds into a form when they first reach a particular point (expression for the *calculation* column: "once(duration())"); you can then subtract one captured duration from another to get the time spent in between. Or, use in a regular *calculate* field to capture the total duration spent on the form overall (expression for the *calculation* column: "duration()").
8. **count(repeatgroup)**: Returns the current number of times that a *repeat group* has repeated (i.e., the number of "instances" of the group, as in the calculate expression "count(${repeatgroupname})").
9. **sum(repeatedfield)**: For a field within a *repeat group*, calculates the sum of all values (as in the calculate expression "sum(${loan\_size})").
10. **join(string, repeatedfield)**: For a field within a *repeat group*, generates a string-separated list of values (as in the calculate expression "join(' ,', ${hh\_member\_name})" to generate a single comma-separated list from all entered names).
11. **min(repeatedfield)**: For a field within a *repeat group*, calculates the minimum of all values (as in the calculate expression "min(${hh\_member\_age})"). If more than one non-repeating field is passed to min(), then the minimum of the passed fields will be returned (as in "min(${father\_age}, ${mother\_age})").
12. **max(repeatedfield)**: For a field within a *repeat group*, calculates the maximum of all values (as in the calculate expression "max(${hh\_member\_age})"). If more than one non-repeating field is passed to max(), then the maximum of the passed fields will be returned (as in "max(${son\_age}, ${daughter\_age})").
13. **index()**: Called within a *repeat group*, returns the index number for the current group or instance (as in the calculate expression "index()", which will return 1 for the first instance of the repeat, 2 for the second, and so on). (We used to recommend the ODK position(..) function instead of index() – but that would fail if you called it from within a non-repeating group that was itself within a repeating group.)
14. **once(expression)**: For use in expressions in *calculate* or *calculate\_here* fields only, indicates that the enclosed expression should be calculated only once per form (as in the expression "once(random())" to draw a random number). If a calculated expression is not enclosed in once(), it will recalculate periodically, including each time that the form is edited and saved. In the case of *random()*, this would draw a new random number every time the form is edited and saved – which is not generally what you want.
15. **once(random())**: Returns a random number between 0 and 1 (as in the calculate expression "once(random())"). To introduce randomness into your forms, use once(random()) in calculated fields, then refer to those calculated fields in your relevance expressions. You should always call random() inside the once() function and you should never use random() directly in a relevance expression, because you don't want to generate a new random number every time the relevance is calculated: you want just one stable random number for each filled-out form. For an example, see the sample form in [*Randomization: Randomizing form elements*](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_samples).
16. **pulldata(source, colname, lookupcolname, lookupval)**: Pulls data from a dataset or .csv file (as in the calculate expression "pulldata('hhplotdata', 'plot1size', 'hhid\_key', ${hhid})", which will pull a value either from an attached .csv file named *hhplotdata.csv* or from an attached dataset with the unique ID *hhplotdata*; the value will come from the *plot1size* column of the pre-loaded data, and the *hhid* field will be used to identify the matching row in the pre-loaded data's *hhid\_key* column). See [*Preloading data into a form*](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_advanced-topics_preloading) for more information, or [*Pre-loading: Referencing pre-loaded .csv data*](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_samples_preloading) for a working example.
17. **indexed-repeat(repeatedfield, repeatgroup, index)**: References a field or group that is inside a prior *repeat group*. The first parameter specifies the prior field or group in which you are interested; the second specifies the prior repeat group within which the field or group of interest is located; and the third specifies the instance number, within the prior repeat group, to use. For example, the calculate expression "indexed-repeat(${name}, ${names}, 1)" will return the first name available when the "name" field is inside a prior repeat group named "names". From inside a later repeat group, the calculate expression "indexed-repeat(${name}, ${names}, index())" will pull the *x*th name from the prior repeat group, where *x* is the instance number of the current repeat group (e.g., if currently in the fourth instance of a repeat group, it will return the fourth name from the earlier repeat group). If you need to reference a field or group within multiple nested repeat groups, you can supply additional parameters to indicate the instance numbers to use for each level of nesting. For example, the calculate expression "indexed-repeat(${name}, ${families}, ${familynumber}, ${names}, ${membernumber})" will pull a particular family member's name when family member names are inside a repeat group that is itself inside a repeat group of families. For two examples, see [*Rosters: Two methods for repeated questions*](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_samples_rosters). Finally, note that when the passed-in instance number is invalid, an instance number of 1 will be automatically used instead (so the first instance will be returned for such cases).
18. **rank-index(index, repeatedfield)**: calculates the ordinal rank of the specified instance of a repeated field for use outside the repeat group (as in the calculate expression "rank(1, ${random\_draw})" to calculate the rank of the first instance, based on the value of its "random\_draw" field as compared with other instances' values). The rank of 1 is assigned to the instance with the highest value, the rank of 2 to the instance with the next-highest value, and so on. Instances with the same value are ordered arbitrarily (they are not given the same rank, so every instance will have a unique rank). If you pass an invalid index or an index to an instance with a non-numeric value, a rank of *999* will be returned. See below for a variation that can be used within the same repeat group.
19. **rank-value(fieldorvalue, fieldorlist)**: calculates the ordinal rank of a given value relative to a list of values (as in the calculate expression "rank(3, '4 2 1 9 3 7')" to calculate the rank of 3 in the given list, or as in "rank(${random\_draw}, ${list\_of\_draws})" if the value and the list are both stored in fields). The rank of 1 is assigned to the highest value in the list, the rank of 2 to the next-highest value, and so on. If you pass an empty list or a value that's not in the list, a rank of *999* will be returned. Use this variation to calculate the rank within the same repeat group as the field being ranked. In this use case, add a *calculate* field outside the repeat group to construct the full list of values using *join()* (as in "join(' ', ${random\_draw})" when the name of the repeated field is "random\_draw"); then, within the repeat group, find the current instance's rank with another *calculate* field (as in "rank(${random\_draw}, ${list\_of\_draws})" when the name of the joined list you added outside the repeat is "list\_of\_draws").
20. **de-duplicate(string, field)**: For a string-separated list of items, removes duplicates (as in the calculate expression "de-duplicate(' ', ${fieldname})" to remove duplicates from a space-separated list). For example, say you had a repeated multiple-choice field. Outside the repeat group itself, you might want to join all selections together into one calculated list, as in "join(' ', ${repeatedfield})" – but that list might then contain the same selections multiple times. The calculate expression "de-duplicate(' ', join(' ', ${repeatedfield}))" would join together all of the selections, and remove duplicates. You could then call *count-selected()*, *selected-at()*, etc. on the combined list in the calculated field.
21. **distance-between(geopointfield1, geopointfield2)**: Returns the distance, in meters, between two *geopoint* fields (as in the calculate expression "distance-between(${start\_gps}, ${end\_gps})"). (Keep in mind that the accuracy of the distance calculated will depend on the accuracy of the GPS readings, so try to be sure to get accurate GPS readings.)
22. **area(repeatedgeopointfield)**: Returns the area enclosed, in square-meters, within a series of repeated *geopoint* fields (as in the calculate expression "area(${gps\_reading})", called outside a *repeat group* that includes a field named "gps\_reading"). (Keep in mind that the accuracy of the area calculated will depend on the accuracy of the GPS readings, so try to be sure to get accurate GPS readings.)
23. **short-geopoint(geopointfield)**: Returns a string containing the GPS location with only the longitude and latitude, no altitude or accuracy (as in the calculate expression "short-geopoint(${location})"). You might use this if publishing data to outside systems like [Google Fusion Tables](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_publishing-forms_forms-to-google-ft), which might be confused by altitude or accuracy.
24. **substr(fieldorstring, startindex, endindex)**: Returns a substring starting at startindex and ending just before endindex (as in "substr(${phone}, 0, 3)" to get the first three digits of a phone number). Indexes start at 0 for the first character in the string.
25. **coalesce(field1, field2)**: Returns field1 if it isn't empty, otherwise returns field2 (as in the calculate expression "coalesce(${id}, ${id2})").
26. **round(field, digits)**: Rounds the numeric field value to the specified number of digits after the decimal place (as in the calculate expression "round(${interest\_rate}, 2)").
27. **regex(field, expression)**: Returns true or false depending on whether the field matches the [regular expression](http://en.wikipedia.org/wiki/Regular_expressions) specified (as in the constraint "regex(., '[A-Za-z0-9.\_%+-]+@[A-Za-z0-9.-]+\.[A-Za-z]{2,4}')" which checks for a valid-looking email address). Because regular expressions can be arbitrarily complex, this function allows for construction of advanced relevance and constraint expressions.
28. **if(expression, valueiftrue, valueiffalse)**: Returns one of two values, depending on whether an expression is true (as in the calculate expression "if(selected(${country}, 'South Africa') or selected(${country}, 'Zimbabwe'), 'SADC', 'Non-SADC')").
29. **pow(base, exponent)**: Returns the field, number, or expression in the first parameter raised to the power of the field, number, or expression in the second parameter (as in the calculate expression "pow(1+${annual\_interest\_rate}, ${years\_of\_interest})").
30. **format-number(field)**: Formats an *integer* or *decimal* field according to the user's locale settings (as in the calculate expression "format-number(${income})", which might format "120000" as "120,000").
31. **~~number(field)~~**~~: Converts field to a number (as in "number('34.8') = 34.8").~~
32. **~~int(field)~~**~~: Converts field to an integer (as in "int('39') = 39").~~
33. **~~string(field)~~**~~: Converts field to a string (as in "string(34.8) = '34.8'").~~
34. **~~date(string)~~**~~: Converts string into a date (as in the relevance "${fieldname} > date('2013-01-31')").~~
35. **~~date-time(string)~~**~~: Converts string into a date-time (as in the relevance "${fieldname} > date-time('2013-01-31T16:42:00')").~~
36. **format-date-time(field, format)**: Converts date and/or time into a string (as in the calculate expression "format-date-time(${fieldname}, '%Y-%b-%e %H:%M:%S')"). In the format string, *%Y* indicates four-digit year, *%y* two-digit year, *%m* two-digit month, *%n* one-or-two-digit month, *%b* three-letter month, *%d* two-digit day, *%e* one-or-two-digit day, *%H* two-digit hour, *%h* one-or-two-digit hour, *%M* two-digit minute, *%S* two-digit seconds, *%3* three-digit milliseconds, and *%a* three-letter day of week.
37. **today()**: Returns the current date (as in the calculate expression "format-date-time(today(), '%Y-%b-%e')").
38. **now()**: Returns the current date and time (as in the *calculate\_here* expression "once(format-date-time(now(), '%Y-%b-%e %H:%M:%S'))" for saving the date and time at which a particular point in your form is first reached).
39. **uuid()**: Calculates a unique random ID (as in the calculate expression "once(uuid())").
40. **version()**: Returns the version number of the current form (as in the calculate expression "version()").
41. **username()**: Returns the currently-configured username of the user filling in the form (as in the calculate expression "username()").
42. **linebreak()**: Returns a linebreak character (as in the calculate expression "concat(${field1}, linebreak(), ${field2}, linebreak(), ${field3})" for storing a list of three field values with linebreaks between them).
43. **hash(fieldorvalue, ...)**: Returns a hash value that represents the one or more parameters passed (as in the calculate expression "hash(${name})" or "hash(${name}, ${birthday})")). See [this page](https://en.wikipedia.org/wiki/Hash_function) for more details on hash values and their potential uses.

*Please note that literal strings within your expressions should always be enclosed in single-quotes, as in 'Male' or ' ' for a single space. If you use double-quotes (like "Male" or " "), your expressions will not work properly.*

[Updating an existing form](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_core-concepts_updating)

Even after your form has been deployed and users have started collecting data, you can safely make and deploy many changes. You just want to be careful to understand the limitations and implications discussed below.

To update an existing form:

1. **Update the form definition's *survey* and *choices* worksheets as desired.** In doing so, you will want to be careful about removing fields, renaming fields, or changing groups, as discussed below.
2. **Increase the version number on the *settings* worksheet.** If you started with a SurveyCTO form template or with [one of the sample forms](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_samples), then this is automatic: the version is set to a formula that automatically increments every minute. (Otherwise, you need to increment the version number yourself, taking care to keep the number of digits fixed: the version has to be a single whole number, and its number of digits has to stay the same from one version to the next. The convention is to use a ten-digit number that represents the current date and time, as in 1401130917 for 9:17AM on January 13, 2014. Alternatively, you can use 10-digit numbers like 0000000001 and 0000000024 or 3-digit numbers like 001 and 024.)
3. **Upload the new version of the form definition to your SurveyCTO server.** To do this, navigate to the [Design](https://malttanzania.surveycto.com/main.html#Design) tab, scroll down to the appropriate form in the *Your forms* list, and click the *Update* button. (As part of your upload, be sure to include any necessary media files referenced by your form, and/or check the box to *Keep old attachments*.)
4. **Deploy the new form to your users.** If you're collecting data via the web, this will happen automatically: whenever a web user begins a new submission, they will always use the latest version of your form. Your Android users who use *SurveyCTO Collect*, on the other hand, can click *Get Blank Form* to manually download the new version from your server. They can also configure Collect to auto-download new form versions whenever they are available, by going into Collect's *General settings* and checking *Auto download with Wi-Fi* and/or *Auto download with network* ("network" here refers to "cellular network"). When either of these options is enabled and the appropriate connection type is available, Collect will check hourly for form updates; available updates will then be downloaded automatically, but they won't be installed until the user confirms the installation by clicking on the*Install Form Updates* button that appears, or by confirming the update when about to fill out an older version of the form.

*SurveyCTO Collect* – on Android or on the web – will always present the latest installed version of your form when a user selects *Fill Blank Form*, but older versions of your form are kept to support users who had already begun submissions with those earlier versions. Whether a user was in the middle of editing an old version of the form when you updated it, or they came back much later to *Edit Saved Form* a form begun with an old version, existing submissions will not be affected by any of your changes. Only new submissions, begun *after* you updated your form, will benefit from the changes you made.

*We strongly encourage you to maintain back-ups of all deployed versions of your forms, and to keep detailed notes on the changes you make. One easy strategy is to add a "notes" worksheet to each of your survey forms, and to add detailed notes there as you make changes; SurveyCTO will simply ignore the extra worksheet, but it can be incredibly valuable to your team – particularly as they work to interpret collected data.*

Implications for your data exports

SurveyCTO keeps track of each form based on its unique form ID (on the *settings* worksheet), so two forms uploaded with the same form ID but different version numbers are, by definition, two versions of the same form, and two forms uploaded with different form IDs are, on the other hand, treated as two different forms. Ultimately, all data is stored, aggregated, and exported based on the form ID, so the data for different form versions that share a single form ID will mix together.

When exporting your data, the export will be based on the current (latest) version of the form available on your server. That version of the form is what will govern which fields and which groups of fields get exported. Data collected with earlier versions of the form may be missing some data (in the case of fields or groups that were added later) and may have some data for fields or groups that have since been deleted. In the former case, where fields are missing for a submission, those fields will be exported as blank or empty – unless you set the preferences option in *SurveyCTO Sync* to specify some other character or string to use for missing values. In the latter case, where data exists for fields or groups that no longer exist in the latest version of the form, that data will simply not be exported at all. (If you want to keep exporting data for fields that you would like to delete, don't actually delete those fields: instead, put the number *0* in their *relevance* column. That will keep the fields in the form so that they will be exported if present, but they will never show up to users filling out the latest version of your form.)

To help you in interpreting your data, every export will automatically include a *formdef\_version* column. For each row (i.e., for each submission), that column will contain the version number of the form definition used to fill out the form.

*You will want to take some care in how you alter the groups within which your form fields reside, because SurveyCTO internally tracks each field by a full name that includes all enclosing groups.* For example, a field named "age" within a group named "demographics" that is itself within a group named "module1" will be tracked as "/module1/demographics/age". If we were to collect some data for that field and then decide to shorten the name of the "demographics" group to just "demo", then we would subtly alter our form with respect to the "age" field: old form submissions would have a field named "/module1/demographics/age" and new submissions would start having a field named "/module1/demo/age" instead. Exports would, by default, include only the "/module1/demo/age" field; data collected for the earlier "/module1/demographics/age" field would *not* be included. Now, if we set the preference to "Ignore groups so fields with same name export together" in *SurveyCTO Sync*, then our exports *would* include merged data for both "age" fields. But that kind of merging slows down the export process, and it can't work when you change anything about the *repeat groups* that enclose a field. So your best bet is to avoid adding or removing groups, renaming groups, or moving fields across groups.

Changes to encryption settings

While you can change most things about an existing form, you cannot change the encryption settings. In other words, you can't change a form from unencrypted to encrypted, you can't change a form from encrypted to unencrypted, and you can't change the encryption key used to encrypt the form. To make any changes to a form's encryption settings, you will need to change the form's unique form ID on the *settings* worksheet so that the updated form will essentially be treated as a new form altogether. (This is because you simply cannot mix data for two form versions if those versions have different encryption settings. It would just never work.)

[Implementing skip patterns with "relevance"](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_core-concepts_relevance)

Enter an expression into the *relevance* column to control when a field or group displays to the user (i.e., to control when it is "relevant"). The field or group will only show up to users when the expression in the *relevance* column evaluates to true; otherwise, it will remain hidden.

Typically, whether a question displays will depend on the responses to prior questions. For example, you might ask for informed consent at the beginning of your survey. Then, for all remaining questions, you may want to include "${consent} = 1" in the *relevance* column. That way, all remaining questions are automatically skipped when consent is not given. Another approach is to enclose all consent-requiring questions within "begin group" and "end group" rows, then specify the relevance expression on the "begin group" row – so you only have to enter the expression once (see the [*Grouping and repeating questions*](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_core-concepts_groups)section above).

Of course, there may be multiple criteria for displaying a certain question. For example, the relevance expression for a question about the fourth household member might look like "${consent} = 1 and ${nhhmembers} > 3".

The easiest way to build a relevance expression is to use the relevance-builder on the server's [Design](https://malttanzania.surveycto.com/help.html#Design) tab: select *Tools*, then *Build relevance* in the *Your forms* section. To use the relevance-builder offline, simply run *SurveyCTO Sync* and select the relevance-builder from the *Tools* menu's *Form tools* sub-menu.

For details on hand-creating or hand-editing relevance expressions, see the [*Using expressions in your forms*](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_core-concepts_expressions) section below.

Please note that all relevance expressions are only evaluated once, when a screen is first displayed. Thus, if you use a "field-list group" to display multiple questions on one screen, one question cannot appear or disappear based on the answer to another question that appears above it on the same screen. To implement skip patterns, questions must appear on separate screens.

[Encrypting form data](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_additional-topics_encrypting)

If you are collecting sensitive, personally-identifiable data, then you will want to encrypt it so that your authorized team members are truly the only ones with access to that data. The model of encryption supported by SurveyCTO is the following:

1. As soon as a surveyor marks a filled-out form as "finalized," the form's contents will be encrypted using your public encryption key – except for those fields explicitly marked as *publishable* (those fields with a "yes" in their *publishable* column). Fields explicitly marked as publishable are left unencrypted so that they can be conveniently [published to cloud services](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_publishing-forms), [directly downloaded from the server](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_overview_accessing-data), and [analyzed directly in the cloud](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_overview_using-statwing).
2. From that point forward, the form can no longer be edited because not even the device on which the form was filled can decrypt it (the "device" in this case can be an Android device or the SurveyCTO server, in the case of web forms).
3. Whenever form data is transmitted via a 3g or other Internet network, it is encrypted in transit using SSL. This is true for all form data regardless of whether the form itself is configured to be encrypted – so, in a sense, encrypted forms are doubly-encrypted (once with your public key, and then again with SSL).
4. The SurveyCTO server stores the form data, but it remains encrypted and therefore unreadable by the server (and by anyone who might conceivably compromise that server, legally or otherwise). (Again, though, fields explicitly marked as *publishable* will remain readable by the server.)
5. When the form data is downloaded by an end-user using *SurveyCTO Sync*, it is again doubly-encrypted, with both your public key and the SSL protocol used for secure data transmission.
6. Sync stores a local backup of the form data in its local storage directory, but it remains encrypted and therefore unreadable.
7. The end-user who downloads the data presumably has your private encryption key, which is necessary to decrypt and read that data. However, for added security this user may choose not to store that private key on an Internet-connected computer. Thus, the user may first transport the data to a secure "cold-room" computer, and only then decrypt the data, export it for further processing, and begin the analysis.

The key thing is that the SurveyCTO model of encryption relies on *you* generating the encryption keys with which data is encrypted; this assures that you and you alone are able to access your data. So the first step is to generate your own public/private encryption key pair, which you can then use to encrypt all of your sensitive forms.

To create a new key pair, run *SurveyCTO Sync* and select *Create a public/private key pair...* from the *Tools* menu, then specify a name for the pair as well as a directory in which to save it (the name will simply be used for the names of the files that get saved). Sync will then save two files into the directory you specified: keyname\_public.pem, and keyname\_PRIVATEDONOTSHARE.pem.

*keyname\_public.pem is the public key used to****encrypt****data*. As the name suggests, it is public – so you don't need to worry about safeguarding it.

*keyname\_PRIVATEDONOTSHARE.pem is the private key used to****decrypt****data*. This private key you want to guard very closely. In fact, you may want to generate the key pair on a "cold room" computer that is extraordinarily secure (e.g., disconnected from the Internet) and never have other copies of the private key in less secure locations.

While you do want to be very careful with your private key, make sure that you do not lose it. The private key is your only way to decrypt data. If you lose it, you will also lose the ability to decrypt data. Thus, you will need some very secure way to back up or otherwise safeguard your private key against loss. If you have it only on a cold-room computer, for example, and that computer fails – how will you decrypt your data?

Once you have a key pair, you are ready to configure your forms for encryption. To start a new encrypted form, just click *Start new* in the SurveyCTO server's main [Design](https://malttanzania.surveycto.com/main.html#Design) tab, mark the checkbox to indicate that you want the form to be encrypted, and then upload your *public* key file when prompted (*not the private key: that you never upload to anybody*). The form template that you download will then include, on the *settings* worksheet, a copy of your public key. That will then be used to encrypt all finalized forms.

You will design and upload these encrypted forms just as you do non-encrypted forms, and surveyors will fill them out and submit them in the usual way. The only real difference will come in when you are ready to publish, download, or analyze your data.

When using *SurveyCTO Sync* to export your encrypted form's data to local .csv files, you will need to use your private key: use the *Browse* button to point Sync to where it resides on your local system. Without that private key, Sync cannot decrypt and export your data. (It can download the data and store it locally, but it cannot decrypt or export it.)

If you are keeping the private key only on a cold-room computer, follow these steps to download, decrypt, and export your data:

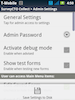
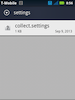
1. On an Internet-connected computer, run *SurveyCTO Sync* with *Server* as the data source and *External drive* as the data destination. When you download data, it will download and sync to an external hard drive (either a thumb drive or an external hard drive).
2. Eject the external drive, bring it to the cold-room computer, and plug it in.
3. Finally, run *SurveyCTO Sync* on the cold-room computer with *External drive* as the data source and *Local CSVs* as the data destination. Also tell Sync where to find the private key file using the *Browse* button.

If you want to publish some parts of an encrypted form [to a server dataset](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_advanced-datasets_datasets-intro) (e.g., to feed directly into other forms), publish some parts [to the cloud](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_publishing-forms) (e.g., to monitor key indicators via a Google Spreadsheets dashboard), or directly [download](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_overview_accessing-data) or [analyze](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_overview_using-statwing) safe subsets of your data in the cloud, you will need to explicitly mark some fields as *publishable* so that they will remain unencrypted (by putting a "yes" in their *publishable* column). Those fields will still be encrypted in transit, but they will be technically readable by the SurveyCTO server.

Copying device settings

Once you have the settings just the way you want them on one device, you may want to copy those settings to other devices so that you don't have to manually enter all of the settings on every device. You can do this by following these steps:

1. Save your desired settings to a file. Press your device's Menu button from the main Collect menu and choose *Admin Settings*. Enter your admin password (if any), then click your device's menu button again and select *Save Settings to Disk*. This will save all of your current settings to the*/SurveyCTO/settings/collect.settings* file on your device's SD card.

[](https://malttanzania.surveycto.com/images/surveycto/help/collect-savesettings.png) [](https://malttanzania.surveycto.com/images/surveycto/help/collect-savesettings2.png)

1. Copy that *collect.settings* file to your computer or to another location. (How you access the SD card and copy its files will depend on the device.)
2. Once you have installed *SurveyCTO Collect* onto another device and run it at least once, locate the /SurveyCTO directory on its SD card.
3. Copy the *collect.settings* file into the device's existing /SurveyCTO directory.
4. Run *SurveyCTO Collect* on the device and confirm that it has loaded all of your settings; when it finds the *collect.settings* file at start-up, Collect will show a small message to indicate that it has restored saved settings. If your settings are not restored when you run Collect, use your device's Task Manager to end the *SurveyCTO Collect* app, then run it again (or, alternatively, restart your device).

[](https://malttanzania.surveycto.com/images/surveycto/help/collect-settingsloaded.png)

You can repeat this process to copy the same configuration to all of your devices. That way, you can ensure that they are all configured in exactly the same way.

Copying both forms and settings

If you want to copy *everything* from one device to another – including saved forms – then you can copy the entire */SurveyCTO* directory tree from your model device, copy the *collect.settings* file from */SurveyCTO/settings* into */SurveyCTO*, then copy that entire directory tree onto another device's SD card. You can even copy the */SurveyCTO* tree onto a new device before *SurveyCTO Collect* has been installed onto it; in that case, you can install Collect and the first time it is run it will inherit all settings and data from the copied directory.

If you start copying */SurveyCTO* directories around, the one thing you need to be careful about is to never combine the contents of multiple devices'*/SurveyCTO* directories. So before pasting in a new one, always delete the old one if there is already one present.

#### [Loading multiple-choice options from pre-loaded data](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_advanced-topics_search-and-select)

See the help topic [just above](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_advanced-topics_preloading) for details on how to attach pre-loaded data to your survey form. Once your form has pre-loaded data attached, you can dynamically pull the choice lists for *[select\_one](https://malttanzania.surveycto.com/help.html" \l "Help_Designing-forms_core-concepts_field-types-select-one)* and *[select\_multiple](https://malttanzania.surveycto.com/help.html" \l "Help_Designing-forms_core-concepts_field-types-select-multiple)* fields from that data.

If your goal is to allow the user of your survey form to select the appropriate row for which you would like to pre-load data, your best bet is to download a form template from the [Design](https://malttanzania.surveycto.com/main.html#Design) tab (*Start new* in the *Your forms* section), specify the data that you wish to pre-load, and use the auto-generated template as a starting-point. You will be able to specify that the user should choose the data to pre-load in one, two, or even three steps (e.g., choosing district, then village, then household), and you will be able to select which pre-loaded data columns to use as the choice labels for each step.

Multiple-choice fields with dynamic choice lists follow the same general syntax as regular, static *select\_one* and *select\_multiple* fields: specify "select\_one listname" or "select\_multiple listname" in the *type* column of your *survey* sheet (where "listname" is the name of your choice list; e.g., select\_one country), specify any special appearance styles in the *appearance* column, and include one or more rows for your listname on the *choices* worksheet. When the choice list should be pulled from pre-loaded data, there are just three important differences in the syntax to be aware of:

1. In the *appearance* column of the *survey* worksheet, you should include a search() expression that specifies which rows of pre-loaded data to include in the choice list. If you want to specify another appearance style, that should be entered into the *appearance* column first, followed by a space, followed by the search() expression (as in "quick search()"). Syntax for the search() expression itself is included below.
2. On the *choices* worksheet, a row should indicate which pre-loaded data columns to use for the label and selected value. In the *list\_name* column, specify the name of your choice list as you normally would. In the *name* column, however, include the name of the column to use for uniquely identifying selected choices (this is normally an ID column of some sort). In the *label* column, include the name of the column to use for labeling the choices; if you wish to include multiple columns in the labels, just include a comma-separated list of all columns to include. At survey time, the *name*column will be dynamically populated based on the column name you put there, and the *label* column will be dynamically populated based on the column name(s) you put there.
3. Also on the *choices* worksheet, you can include a column name in the *image* column. If you do, the image filename to use will be pulled from the specified column. If you refer to image files in this way, you must always upload those image files as media file attachments when you upload your form to the SurveyCTO server (as you normally would).

For the search() expression itself, there are a series of options to indicate which rows of pre-loaded data to include in the choice list:

* "**search(source)**": This single-parameter search expression includes all distinct rows as choices; you just have to specify the data source, which is either the name of the attached .csv file or the unique ID of the attached dataset (e.g., "search('hhplotdata')" if either the *hhplotdata.csv* file was attached or a dataset with the ID *hhplotdata* was attached). All rows in the specified dataset or .csv file will be considered as choices, but only *distinct*rows – those with unique selection values – will be listed for the user. In other words, duplicates will be automatically filtered from the list shown to users.
* "**search(source, 'contains', columnsToSearch, searchText)**": This search expression includes all distinct rows that contain the specified text in the specified column(s) (e.g., "search('hhplotdata', 'contains', 'respondentname', ${nametofind})"). The third parameter specifies either a single column name to search, or a comma-separated list of column names to search. Rows with matches in any specified column will be included.
* "**search(source, 'startswith', columnsToSearch, searchText)**": This search expression includes all distinct rows that start with the specified text in the specified column(s) (e.g., "search('hhplotdata', 'startswith', 'respondentname', ${nameprefix})"). The third parameter specifies either a single column name to search, or a comma-separated list of column names to search. Rows with matches in any specified column will be included.
* "**search(source, 'endswith', columnsToSearch, searchText)**": This search expression includes all distinct rows that end with the specified text in the specified column(s) (e.g., "search('hhplotdata', 'endswith', 'respondentname', ${namesuffix})"). The third parameter specifies either a single column name to search, or a comma-separated list of column names to search. Rows with matches in any specified column will be included.
* "**search(source, 'matches', columnsToSearch, searchText)**": This search expression includes all distinct rows that exactly contain the specified text in the specified column(s) (e.g., "search('hhplotdata', 'matches', 'respondentname', ${nametofind})"). The third parameter specifies either a single column name to search, or a comma-separated list of column names to search. Rows with exact matches in any specified column will be included.
* "**search(source, searchType, columnsToSearch, searchText, columnToFilter, filterText)**": Finally, any of the four search types above can be further filtered to only include a subset of pre-loaded data. Simply add two extra parameters to any of the search types above, with the first extra parameter being the column name to filter and the second extra parameter being the exact value to filter. For whichever search is specified in the first four parameters, only rows exactly containing the sixth parameter value in the column named by the fifth parameter will be included (e.g., "search('hhplotdata', 'contains', 'respondentname', ${nametofind}, 'villageid', ${villageid})" to list all matching names within a particular village).

Rather than writing it out by hand, you can use the calculation-builder to construct your own search() expression. The builder is available as both an option on the server (from the [Design](https://malttanzania.surveycto.com/main.html#Design) tab, click *Tools* under *Your forms*) and as an option in *SurveyCTO Sync* (on the *Tools...Form tools* menu). Once you run the builder, choose *Pull pre-loaded data into a multiple-choice prompt* as the calculation type, then enter your details and click *Build*. You can just copy and paste the resulting expression into your form's *appearance* column.

For a working example, see [*Pre-loading: Searching and selecting from pre-loaded data*](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_samples_search-and-select).

Two additional notes on usage:

1. Choices will be ordered, by default, in the order that they appear in your dataset or .csv file. If you want to specify a different ordering, include a numeric column in your dataset or .csv file named *sortby*; if present, choices will be ordered numerically, according to this *sortby* column.
2. You can include one or more static choice options, in addition to the dynamic ones loaded from your pre-loaded data. Simply include static choices as you normally would, on the *choices* worksheet. These can appear before and/or after the row that indicates the columns to use for your dynamic choices. The only restriction is that the values you specify for your static choices in the *name* column must be numeric.

#### [Exporting, publishing, and using your data](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_overview_accessing-data)

SurveyCTO makes it easy to not only collect data, but also to export, publish, and start using that data. Everything you need to get started is located on your server console's [Export](https://malttanzania.surveycto.com/main.html#_blank) tab.

The quickest and easiest way to access and start using your data is to go to the *Your forms* section of the [Export](https://malttanzania.surveycto.com/main.html#_blank) tab. There, you will find two simple options available for each of your survey forms:

1. **Download**: directly download .csv data in the most common and convenient format. This option exports your data into a single wide-format .csv file that includes extra columns for both repeated data and multiple-choice questions. Media files included in your data (e.g., photos or audio recordings) are included in the .csv file as URLs that you can use to download those files.
2. **Analyze**: analyze your data directly using the powerful *Statwing* analysis platform, which is fully integrated with SurveyCTO. When data is passed to*Statwing* for analysis, it is formatted in the same wide-format .csv format described just above.

Exported form data always includes a fully-unique identifier for each and every filled-out form (also known as the submission's "key") in the *KEY* column; use this identifier to uniquely identify and track all form data. All exported data also includes the version number of the form definition used on the device itself, in the *formdef\_version* column; use this column to make sense of the data if your form has changed over time (for more on updating forms, see [*Updating an existing form*](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_core-concepts_updating)).

If you have [encrypted your form data](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_additional-topics_encrypting) with your own encryption keys, you can directly download or analyze only those form fields that were explicitly marked as *publishable* – but you can use *[SurveyCTO Sync](https://malttanzania.surveycto.com/help.html" \l "Help_Exporting-and-publishing-data_exporting-with-sync)* to download and export all of your data, encrypted or not, in a variety of formats.

While the above options for direct download and analysis are meant to help you get started working with your data, a richer, more full-featured range of additional options are also available to you. You can learn about them in the following help topics:

* [*Exporting data with SurveyCTO Sync*](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_exporting-with-sync)
* [*Publishing data to the cloud*](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_publishing-forms)
* [*Advanced publishing with datasets*](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_advanced-datasets)
* [*API access*](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_api-access)

For more about the format of exported data, see [*Understanding the format of exported data*](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_overview_data-format). For more about using *Statwing* to analyze your data, see [*Using Statwing to start visualizing, exploring, and analyzing data*](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_overview_using-statwing).

#### [Understanding the format of exported data](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_overview_data-format)

SurveyCTO defaults to exporting data in .csv files. These are formatted in a comma-separated-value format supported by just about any spreadsheet, database, or statistical analysis software. For the main .csv file exported for each of your forms, the first row (also called the "header row") contains field or column names, and each additional row contains data for a specific submission (also called a "filled-out form" or "instance"). Every file includes columns that correspond to the fields in your form, plus the following extra columns:

* **KEY**: a unique identifier for each submission, automatically assigned by SurveyCTO. You can use this key to uniquely identify and track each submission through your entire data pipeline.
* **instanceID**: a duplicate copy of the submission's key (can be ignored).
* **formdef\_version**: the version number of the form definition used to fill out the submission. Use this column to make sense of data when your form has changed over time (for more on updating forms, see [*Updating an existing form*](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_core-concepts_updating)).
* **SubmissionDate**: the date and time that the submission was submitted to the SurveyCTO server. This may well differ from the date and time at which the form was initially filled out.

Except for some special cases described below, each field in your survey will be exported as its own separate column. The name of each column (the value in the header row of exported .csv files) will simply be the name that you gave the field in your form definition – unless you are exporting data with*SurveyCTO Sync* and have selected the option to include group names in .csv column headers. In that case, column headers will include both field names and enclosing groups (e.g., *consented-agriculture-crops* for a field named *crops* that's inside groups named *consented* and *agriculture*).

Dates and times (like the *SubmissionDate* column discussed above) are always exported relative to the time zone of the exporting computer – *not* the time zone of the data-collection device. This is done as a convenience so that data collected in multiple time zones can be easily analyzed relative to a single fixed time zone. If you are exporting data with *[SurveyCTO Sync](https://malttanzania.surveycto.com/help.html" \l "Help_Exporting-and-publishing-data_exporting-with-sync)*, then the time zone used will be the one set in the exporting computer's region or locale settings; if exporting data directly from the server, the time zone used will always be UTC (for downloading .csv data directly from the server or publishing to the cloud).

GPS locations (captured with *[geopoint](https://malttanzania.surveycto.com/help.html" \l "Help_Designing-forms_core-concepts_field-types-geopoint)* fields) are exported as four columns rather than just one. For example, for a field named *gpsloc*, the four columns would be: *gpsloc-Latitude*, *gpsloc-Longitude*, *gpsloc-Altitude*, and *gpsloc-Accuracy*.

Multiple-choice fields that allow the user to choose more than one option (captured with *[select\_multiple](https://malttanzania.surveycto.com/help.html" \l "Help_Designing-forms_core-concepts_field-types-select-multiple)* fields) are exported differently depending on how you export your data. When downloading data directly from your server console's [Export](https://malttanzania.surveycto.com/main.html#_blank) tab, a 1/0 column is added for each choice; for example, a single field named *crops* that has choice values 1, 2, and 3 would be exported with *crops\_1*, *crops\_2*, and *crops\_3* columns, with submission values of 1 if selected or 0 if not selected. *SurveyCTO Sync* will export these 1/0 columns as well, if set to do so in the preferences; but it will also export a single column that includes a space-separated list of selected values (like "1 2 3" for a submission where all three crops were selected).

Fields inside [repeat groups](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_additional-topics_repeats) can have multiple values for a single submission, so they require the exported data to be structured a bit differently. Repeated data can be exported in either "long" or "wide" format:

* **Wide**: in wide format, additional columns are simply added to the primary .csv file in order to accommodate repeated data. For example, if a household roster included a repeated field named *age*, then the exported .csv file would include columns *age\_1*, *age\_2*, and so on; if the maximum number of household members in the data was 21, then there would be 21 *age* columns and 21 columns for every other field in the roster's repeat group. The server console's [Export](https://malttanzania.surveycto.com/main.html#_blank) tab defaults to exporting in this format, and there is a preferences option in *SurveyCTO Sync* to include wide-format exports as well.
* **Long**: in long format, a separate row is added for each repeat instance. Since the main .csv file is structured to have only one row per submission,*SurveyCTO Sync* exports each repeat group into a separate .csv file – so a form with repeat groups will export as a primary .csv file plus additional .csv files for each repeat group. Each row in a repeat group's .csv file will have its own unique *KEY* value, plus a *PARENT\_KEY* value that can be used to link to the primary submission or parent-group .csv file.

Exported .csv files are encoded in Unicode format so that they can support the widest possible range of characters and scripts. Unfortunately, Microsoft Excel defaults to importing .csv files as Latin text, so some accents or other scripts may be distorted; see [*Looking at data in Microsoft Excel*](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_exporting-with-sync_using-excel) for details on how to safely import your data into Excel.

SurveyCTO will also export data in other formats. See the help topics on *[SurveyCTO Sync](https://malttanzania.surveycto.com/help.html" \l "Help_Exporting-and-publishing-data_exporting-with-sync)* and [publishing data to the cloud](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_publishing-forms) for more details.

#### [Using Statwing to start visualizing, exploring, and analyzing data](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_overview_using-statwing)

You can use statistical software like Stata, SPSS, or SAS to visualize and analyze your data, or use other desktop software like Microsoft Excel. To get a quick start exploring, visualizing, and analyzing your data in the cloud, you can use *Statwing*.

*Statwing* is a powerful platform for data analysis, integrated directly into SurveyCTO, that allows you to very quickly and easily describe not only the individual fields in your form's data, but also the empirical relationships between those fields. Rather than having to tediously set up every graph, table, or regression, *Statwing* will do all of the hard work for you, deciding which graphs, tables, and regression analyses best represent your data.

To get started using *Statwing*, click the *Analyze* button in the *Your forms* section of your server console's [Export](https://malttanzania.surveycto.com/help.html#_blank) tab. You'll be able to choose which fields to send over to *Statwing*, so that you can keep any personally-identifiable or sensitive data safe on your SurveyCTO server. If you have [encrypted your form data](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_additional-topics_encrypting) with your own encryption keys, you will only be able to analyze those form fields that were explicitly marked as *publishable* – but you can also use[*SurveyCTO Sync*](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_exporting-with-sync) to download your data, decrypt it, and send a safe subset of it to *Statwing*. See [*Using Statwing*](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_exporting-with-sync_using-statwing) for more on using *Statwing* via *SurveyCTO Sync*.

Once you've loaded your data into *Statwing*, you'll find additional help available on-screen. Depending on your SurveyCTO subscription, use of *Statwing*may be completely free – or you may have to create an account and pay to keep using it beyond a free trial period. See [this help topic](https://malttanzania.surveycto.com/help.html#Help_getting-started_overview_your-subscription) for more about your SurveyCTO subscription.

#### [Installing and using SurveyCTO Sync](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_exporting-with-sync_using-sync)

*SurveyCTO Sync* is software you can use to download, process, and export data collected with SurveyCTO. Get started by downloading and installing the appropriate version for your computer's operating system:

* [SurveyCTO Sync for Windows](http://malttanzania.surveycto.com/client/SurveyCTO_Sync.exe)
* [SurveyCTO Sync for Mac (OSX)](http://malttanzania.surveycto.com/client/SurveyCTO_Sync.dmg)
* [SurveyCTO Sync for other platforms](http://malttanzania.surveycto.com/client/SurveyCTO_Sync.jar)

The first time you run Sync, any or all of the following may happen: (a) it may need to install Oracle's Java environment (on which it relies); (b) it may need to upgrade your Java environment in order to support strong data security (if so, it will walk you through the process); and (c) it will ask you where on your local hard drive to store downloaded form data.

[](https://malttanzania.surveycto.com/images/surveycto/help/client-jce-update.png)

(If you are running Windows and have trouble upgrading your Java encryption support, you can close Sync, right-click on its .exe file, and choose *Run as Administrator*. Upgrading generally works fine as long as your current login has sufficient permissions to access the Java system files. If your organization does not allow you permission to upgrade your own system, tell your IT department that you need them to install the "Java Cryptography Extension (JCE) Unlimited Strength Jurisdiction Policy Files.")

Once Sync is installed, you can use it to download, transport, and export data. Simply choose the source of the data (typically "Server", in which case you enter **malttanzania** as the server name, plus a valid username and password), a destination (typically "Local CSVs", in which case you select a local directory into which exported data should be saved), and then press *GO!*. Next, choose the forms for which you want to download, transport, and/or export data. Finally, press the second *GO!* button. Data will be transported and/or exported to the specified location.

[](https://malttanzania.surveycto.com/images/surveycto/help/client-demo1.png) [](https://malttanzania.surveycto.com/images/surveycto/help/client-demo2.png) [](https://malttanzania.surveycto.com/images/surveycto/help/client-demo3.png)

Form data is exported in .csv format by default, then Sync offers a range of options for exporting additional files in other formats as well. For more about the default .csv format, see [*Understanding the format of exported data*](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_overview_data-format). (You should review that help topic even if you plan to export in some other format, as the overall structure of the data is discussed there.)

Note that exported form data is cumulative in that it includes all data stored on the server and all data stored locally by Sync. However, this is not something upon which you should rely. Old data may eventually be purged from both the server and Sync, so exported .csv files should not be relied upon to always include all old data.

It is best to think of each export as containing all new data as well as some possibly-incomplete subset of old data. Therefore, your data-processing tasks should include:

1. Merging newly-exported data into a master dataset (probably ignoring all old data in the .csv exports).
2. Maintaining backups of both your master dataset as well as all raw .csv files exported by Sync. While it is redundant to back up both the raw .csv files and the master dataset, it is better safe than sorry. You should also rotate your backups so that some backups are maintained off-site, in case of theft, fire, or other office mishap.

You can configure Sync to execute processes that partially or completely automate both of these tasks. You can also export subsets of your data to Google Earth, to [www.Statwing.com](http://www.statwing.com/), or to other .csv files; mail-merge data with Microsoft Word templates; or track the status of your back-office-data processing. See the following help topics for details.

Within Sync's preferences, you will also find various settings to control how your data is stored and exported. On Windows, this will be an option under the*Tools* menu in the upper-left of the *SurveyCTO Sync* window; on OSX, it will be an option under the *SurveyCTO Sync* menu in the upper-left of your screen.

#### [Data export options](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_exporting-with-sync_export-options)

*SurveyCTO Sync* offers a wide range of options to control how data is processed and exported. You can find all of these options within the preferences (on Windows, find *Preferences* under the *Tools* menu in the upper-left of the Sync window; on OSX, find it under the *SurveyCTO Sync* menu in the upper-left of your screen). There, preference options are organized in a series of different tabs. Some of the less obvious options are discussed below.

Treatment of enclosing groups in exports

Internally, SurveyCTO stores fields according to a "full name" that includes the field's name as well as the names of all enclosing groups (as in "module1/field1"). One implication of this is that a field named "field1" in one group is treated as effectively different from another field named "field1" that is located in another group. In part as a reflection of this, Sync's default behavior is to include group names in the column headers of exported .csv files (as in "group1-field1" or even "outergroup-innergroup-field1").

Since most people almost always give unique names to all of the fields in their forms, regardless of which groups they're in, including the group names in the export headers is often unnecessary – and unwanted, because it makes the headers longer and more unwieldy. So, there are two options to exclude those group names from export headers:

1. *Exclude groups, but still use them internally*. This option simply drops the group names from the export headers, but it still considers fields in different groups to be different fields. So, for example, if you did have a field named "field1" in two different groups, you would end up with two "field1" columns in your exports (because Sync would be exporting them as different fields, even though they share the same name).
2. *Ignore groups so fields with same name export together (slowest)*. This option merges data for different fields (in different groups) that share the same name, so that there will never be more than one export column for a single field name. This option slows down the export process, but it can be a life-saver if you have edited your form over time and moved a field from one group to another: Sync will export old and new data for a field even if it has changed groups. (See [*Updating an existing form*](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_core-concepts_updating) for more.)

Note that, because Stata and mail-merge templates reference field names as they appear in your .csv headers, you will need to re-generate any [Stata .do file templates](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_exporting-with-sync_using-stata) or [Microsoft Word mail-merge templates](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_exporting-with-sync_using-word) after changing this option.

Exclude note fields from exports

Forms typically include a lot of [*note*](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_core-concepts_field-types-note) fields, which display something to the user but don't actually record any data. For completeness, Sync includes these fields in exports by default – but their columns are always blank. You can use this option to simply exclude those fields from your exports.

Field value to include for missing values

By default, "missing" values are exported as blanks. This includes all cases where a field was never displayed and so the user could not have given a response (the field might not have been *relevant*, or the field might not have existed in the version of the form filled out by the user). To distinguish these missing cases from cases where the field did display but the user simply left it blank, you can ask Sync to replace all missing values with some character or string. For example, if you wanted to follow the standard set by Stata, you could configure Sync to replace missing values with a single period (".").

Show test forms

By default, Sync won't show forms with titles that begin with "TEST - ", but you can override that by enabling this option. See [*Hiding forms during testing*](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_additional-topics_hiding-test-forms) for more.

Override the default value separator in .csv files

By default, SurveyCTO separates columns in .csv files using commas (after all, CSV stands for "comma-separated values"). Depending on your region, however, your system may prefer to use semi-colons (;'s) or some other character to separate columns in .csv files.

You can override this default and specify some other separator to use. For most cases, you will want to use either a comma (,) or a semi-colon (;).

Replace line-breaks in exported data

Depending on what software you use to open .csv files exported by Sync, line-breaks within field values might cause confusion. For example, Microsoft Excel can get confused and start a new row when a field includes a line-break. Therefore, you can have Sync auto-replace all line-breaks with some other character or string of your choosing. (A single space character is the most popular replacement for line-breaks.)

Export repeat groups in "wide" format

Sync defaults to exporting data from [repeat groups](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_additional-topics_repeats) into a series of separate long-format .csv files. However, if you wish, you can choose this option to also export a single wide-format .csv file that includes extra columns to accommodate repeated data (the long-format .csv files will still be exported as well). See[*Understanding the format of exported data*](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_overview_data-format) for more on this issue of long vs. wide export formats.

Export select\_multiple responses as series of 1/0 columns

By default, Sync will export each *[select\_multiple](https://malttanzania.surveycto.com/help.html" \l "Help_Designing-forms_core-concepts_field-types-select-multiple)* field as a single column that includes a space-separated list of selected choice values (e.g., "1 2 5" or "Boston NYC"). If you prefer a series of 1/0 "dummy" variables, Sync can also export an additional 1/0 column for each possible selection (though it will slow down the export process somewhat). See [*Understanding the format of exported data*](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_overview_data-format) for more on the structure of exported data.

#### [Looking at data in Microsoft Excel](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_exporting-with-sync_using-excel)

The default approach to exporting your data with *SurveyCTO Sync* is to export it to "Local CSVs", in which case one or more .csv files will be exported into the destination directory of your choosing. (The default behavior is to export all data in "long format", in which case data from [repeat groups](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_additional-topics_repeats) is exported into separate .csv files. You can also export a unified wide-format .csv file that includes all data, repeated or not. See [*Data export options*](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_exporting-with-sync_export-options) above for details.)

The first thing many people do is open their data in Microsoft Excel. You can certainly review, clean, or even analyze your data in Excel. However, please note the following:

1. If you double-click on a .csv file, most systems will default to opening the file in Excel. In this case, Excel will import the .csv file based on a series of assumptions, most of which will not alter your data. But it will assume that the characters in the file are Windows or Mac characters as opposed to Unicode characters, so it will mess up accents or foreign scripts.
2. The safer way to import a .csv file is to first open Excel and then choose either *File...Import* or *Data...Get External Data...Import Text File*. You will want to specify that it is a .csv file, that it is comma-delimited, and that its "file origin" or "encoding" is Unicode/UTF-8. You will then get the opportunity to preview the columns and override the cell formats used for each column. Finally, when it asks you where to put the data, there will be a *Properties...*button that you can click to see options for "refresh control": whether to re-import the data whenever you open the Excel workbook and whether to re-prompt for the source .csv file every time. Once you import this way, you can essentially re-import at any time by clicking *Refresh* on the *Data* tab.
3. Sadly, when you import data using Excel's import wizard, it gets confused when it sees line-breaks inside cells – which will be a problem for you if anybody ever presses Enter when entering a text response into one of your survey forms. Excel will end the row at the line-break and essentially break one row of data into multiple rows. To avoid this, there is an option in Sync's preferences to replace line-breaks in exported .csv files with some other character; the default is to replace all line-breaks with spaces.
4. Excel will sometimes convert things that aren't dates into dates. For example, say you have an ID number like "10-11-12": Excel will convert it into a date, encoding it internally (and saving it!) as the number of days between Jan. 1, 1980 and Oct. 11, 2012. Obviously, that is totally wrong. To prevent Excel from corrupting your data in this way, you need to take care when assigning the formatting of your columns on import: for ID numbers and other things that might be mistaken for dates, "General" is not a safe choice.
5. If you want to edit your data or make notes in new columns, you are better off configuring Sync to directly merge incoming data into an existing Excel workbook, rather than exporting to .csv format and then re-importing into Excel yourself. See [this help topic](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_exporting-with-sync_using-excel-bo) for more on merging directly into Excel.

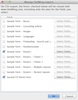
#### [Using Statwing](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_exporting-with-sync_using-statwing)

You can use statistical software like Stata, SPSS, or SAS to visualize and analyze your data, or use other desktop software like Microsoft Excel. To get quickly started exploring, visualizing, and analyzing your data in the cloud, you can use *Statwing*.

*Statwing* is a powerful platform for data visualization and analysis, integrated directly into SurveyCTO. Using it, you can very quickly and easily describe not only the individual fields in your form data, but also the empirical relationships between those fields. And rather than having to tediously set up every graph, table, or regression, *Statwing* will do all of the hard work for you, deciding which graphs, tables, and regression analyses best represent your data.

To use *Statwing* directly from the SurveyCTO server, click the *Analyze* button in the *Your forms* section of your server console's [Export](https://malttanzania.surveycto.com/help.html#_blank) tab. If you have[encrypted your form data](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_additional-topics_encrypting) with your own encryption keys, however, or would otherwise prefer to use *Statwing* via *SurveyCTO Sync*, you can do so easily.

Start by beginning the process of exporting data using Sync: press the first *GO!* button and select the forms that you would like to export (including any forms for which you would like to export data to *Statwing*). Before clicking the second *GO!* button, check the *Show export options after download* checkbox. Then, click the second *GO!* button and, when the export options window pops up: check the *Export subsets of some .csv data to www.Statwing.com for some forms* option, click *Manage Statwing exports...*, and select the forms for which you would like to export data.

[](https://malttanzania.surveycto.com/images/surveycto/help/client-statwing1.png) [](https://malttanzania.surveycto.com/images/surveycto/help/client-statwing2.png) [](https://malttanzania.surveycto.com/images/surveycto/help/client-statwing3.png)

For each form that you select, you will be asked to choose which subset of fields to share with *Statwing*. Typically, you will only want to share non-PII (non-personally identifiable information) fields, so that confidentiality and data security are maintained.

[](https://malttanzania.surveycto.com/images/surveycto/help/client-statwing4.png)

Once configured, your chosen subset of data will be automatically uploaded to *Statwing* whenever new .csv data is exported. A window will then pop up with hyperlinks to view each form's data in *Statwing*. Once in *Statwing*, you can find additional help available on-screen.

Depending on your SurveyCTO subscription, use of *Statwing* may be completely free – or you may have to create an account and pay to keep using it beyond a free trial period. See [this help topic](https://malttanzania.surveycto.com/help.html#Help_getting-started_overview_your-subscription) for more about your SurveyCTO subscription.

#### [Using Google Earth](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_exporting-with-sync_using-google-earth)

Whenever you export data to local .csv files, you can choose to also export a subset of that data to Google Earth ([click here](https://www.google.com/earth/#_blank) to learn more about Google Earth).

Start by beginning the *SurveyCTO Sync* data-export process the way you usually do: press the first *GO!* button and select the forms that you would like to export (including any forms that contain data that you would like to export to Google Earth). Before clicking the second *GO!* button, however, check the*Show export options after download* checkbox down at the bottom. Then, click the second *GO!* button, and when the export options window pops up, check the *Export Google Earth .kml files for some forms* option, click *Manage .kml exports...*, and select the forms for which you would like to view data in Google Earth. (Note that only forms containing at least one *geopoint* field will be listed since Google Earth requires GPS locations for all data points.)

[](https://malttanzania.surveycto.com/images/surveycto/help/client-google-earth1.png) [](https://malttanzania.surveycto.com/images/surveycto/help/client-google-earth2.png) [](https://malttanzania.surveycto.com/images/surveycto/help/client-google-earth3.png)

For each form that you select, you will be prompted to configure some export options. First, specify which field to use for the GPS location. Next, specify which field to use when labeling your data points in Google Earth. Finally, choose which subset of fields to include in Google Earth; the data for these fields will appear when you click on a data point.

[](https://malttanzania.surveycto.com/images/surveycto/help/client-google-earth4.png)

By default, all chosen data for a given form exports to a single .kml file. However, you can also organize your exported data into multiple .kml files by choosing the fields by which you want exports organized. For example, if you choose to organize first by district, then by block, and then by village, your .kml files will be named "Form title - district - block - village.kml", and each file will only include the appropriate subset of data. This makes it easy to, for example, monitor survey operations in different areas or for different teams.

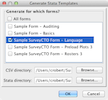
Once configured, your chosen subset of data will export to .kml file(s) whenever new .csv data is exported (to the same .csv export location).

#### [Using Stata](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_exporting-with-sync_using-stata)

*SurveyCTO Sync* can automatically generate Stata .do files that import, merge, and partially process your exported data. You can use these auto-generated .do files as they are, or you can use them as a starting-point for your own back-end processing code.

Start by beginning the *SurveyCTO Sync* data-export process the way you usually do: press the first *GO!* button and select the forms that you would like to export (including any forms for which you would like to generate Stata templates). Before clicking the second *GO!* button, check the *Show export options after download* checkbox down at the bottom. Then click the second *GO!* button, and when the export options window pops up, click the *Generate Stata templates...* button.

[](https://malttanzania.surveycto.com/images/surveycto/help/client-stata1.png) [](https://malttanzania.surveycto.com/images/surveycto/help/client-stata2.png)

[](https://malttanzania.surveycto.com/images/surveycto/help/client-stata3.png)

Next, you will be given a list of all survey forms that have been downloaded from your server. Check off those forms for which you would like to generate Stata templates. Also specify your .csv export location, as well as the location to which you would like to output your Stata files.

Once you press *OK* to continue, Sync will output your Stata templates. If any of your selected forms include multiple languages, Sync will prompt you to choose a language for each. Stata variables will then be labeled according to the languages you select.

For most forms, Sync will output two files: import\_FORMID.do in your chosen Stata directory, and FORMTITLE\_corrections.csv in your .csv export location (where FORMID and FORMTITLE are replaced by the form's ID and title, respectively). If your form has nested repeat groups (i.e., repeat groups inside other repeat groups), then additional .do files will also be output and called from the main import\_FORMID.do file.

Once you have exported form data into your chosen .csv export location, you can run any auto-generated template as-is. You can also copy auto-generated .do files into the *thenrun* subdirectory of your .csv export directory, to have them run automatically every time new data is exported (see [*How do I automatically execute processes when new data becomes available?*](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_exporting-with-sync_outside-processes)).

Note that Stata only allows variable names to be up to 32 characters long. If you have very long group and/or field names, you will run into Stata errors if the first 32 characters fail to uniquely identify a variable. If you run into this problem, you can choose not to export group names in your .csv column headers (see [*Data export options*](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_exporting-with-sync_export-options)), then re-generate your Stata templates and re-export your data; that will shorten many field names because they will no longer include any enclosing group names.

Even if you do not plan to customize them much, you should familiarize yourself with the Stata code contained in these templates. Most broadly, each template does the following:

1. Imports, labels, and formats all incoming data. This includes, for example, merging all repeat-group data into "wide" format. (Note, however, that it is unclear how to best merge repeat-group data when repeat groups are nested. Therefore, when repeat groups are nested, data for repeat groups are organized into separate .dta files, one for each repeat group. You can merge these data as you see fit, using the *key* and *parent\_key* variables.)
2. Merges with any previously-imported data, dropping any duplicates (such that previously-imported data is respected and not overwritten).
3. Applies data corrections, if any.
4. Saves the revised Stata dataset.

In more detail, each template:

1. Initializes Stata ("clear all", "set mem", etc.). Depending on your memory requirements and version of Stata, you may need to revise this code.
2. Initializes file locations in local macros. If you later want to change your .csv or Stata directories, you can update these macro definitions.
3. Lists any names of repeat groups in local macros, both as they would appear in .csv filenames and as they would appear in imported .csv file headers.
4. Lists any names of text, note, date, and date-time fields in local macros. The fields are listed with the names as they will come into Stata, from the exported .csv headers. Sync tries to make these lists as accurate as possible, based on the form definitions. However, it is possible that you might need to tweak them.
5. Imports the primary incoming .csv file exported by Sync.
6. Imports any secondary .csv files from repeat-group data. The template reshapes all repeat-group data into "wide" format (see Stata's help on the "reshape" command), then merges that data together with the primary data. The result is a single in-memory dataset with all of the form's data together in wide format (i.e., with one row per filled-out form). (But note that when repeat groups are nested, data for repeat groups are organized into separate .dta files, one for each repeat group. You can merge these data as you see fit, using the *key* and *parent\_key* variables.)
7. Drops any note fields, since they do not contain data.
8. Converts any date and date-time variables from text format into Stata's internal date/time format. That way, they sort and filter properly. Please note that the default code uses the clock() and date() functions to parse incoming dates, and it presumes that .csv dates are in month-day-year order. If your computer's regional settings instead format dates in day-month-year order, you will need to correct the code: search for "MDY" and replace all occurrences with "DMY".
9. Converts all text fields to text format. By default, this includes "calculate" fields – but you can destring them later if you want, or remove them from the*text\_fields* macro to not convert them at all.
10. Labels variables and *select\_one* values. Note that Stata can only label numeric values, so the template will only label *select\_one* values when all possible values are numeric. Note also that variable labels are truncated at 79 characters, but the template also adds a "note" to each field with the full text of the label (as found in the form definition).
11. Merges with any previously-imported data, dropping any duplicates. Because the Stata process is designed to run repeatedly – each time importing .csv files that likely contain both old and new data – the import process is designed to never overwrite existing data with incoming .csv data. That way, you can always update or extend existing data in the Stata file without fear of it being overwritten. The import process will only add new data to the existing data file.
12. Saves the updated data file.
13. Outputs the codebook and all variable notes to the Stata console.
14. Applies corrections, if any. If you wish, you can enter data corrections into the FORMTITLE\_corrections.csv file (where FORMTITLE is the title of the form).

Sync outputs an empty template into your .csv export directory at the same time that it writes the Stata template, and you can then edit this template and add a row for each correction you would like to make to the data. If you want to use Microsoft Excel to maintain the list of corrections, you should maintain a .xls or .xlsx file in which all cells are set as *Text* format; otherwise, Excel will do funny things like assume that door number "4/16" is April 16 and encode the value as a very long number instead of "4/16". If you maintain your corrections in .xls or .xlsx format, you can either "Save as" .csv format for the Stata template, or you can update the Stata template to import your .xls or .xlsx file directly.

If there are any corrections in the corrections .csv file, they will be applied in sequence, row by row. Each row should indicate the *key* of the row to correct, the name of the field to correct (as it appears in the header of .csv exports), the corrected value, and any notes you might wish to maintain in the corrections file (optional).

If there is any error applying a correction, an error message will be output to the Stata console, including the row number of the offending correction. For example, if you entered a correction with a value of "John" for a numeric field, or if you had a typo in a field name, there would be an error applying the correction.

Note that the Stata code for applying corrections is tricky. In essence, each correction is output as Stata code, into a temporary .do file, and the template executes the temporary .do file in order to apply the corrections. Thus, you have auto-generated Stata code outputting and running its own auto-generated Stata code.

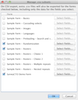
1. Saves the corrected data file.

The auto-generated Stata templates are mostly meant to get you started on your back-end data-processing. Feel free to revise, extend, and delete whichever parts you don't need or want.

#### [Sharing subsets of data](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_exporting-with-sync_sharing-data-subsets)

Whenever you use *SurveyCTO Sync* to export data to local .csv files, you can choose to also export a subset of that data to other .csv files. You might do this, for example, if you wanted to share a subset of non-PII (non-personally identifiable information) fields with another person, program, or website.

Start by beginning the *SurveyCTO Sync* data-export process the way you usually do: press the first *GO!* button and select the forms that you would like to export (including any forms for which you would like to export subsets of data). Before clicking the second *GO!* button, check the *Show export options after download* checkbox down at the bottom. Then, click the second *GO!* button, and when the export options window pops up: check the *Export subsets of .csv data for some forms* option, click *Manage .csv subsets...*, and select the forms for which you would like to export subsets of data.

[](https://malttanzania.surveycto.com/images/surveycto/help/client-subsets1.png) [](https://malttanzania.surveycto.com/images/surveycto/help/client-subsets2.png) [](https://malttanzania.surveycto.com/images/surveycto/help/client-subsets3.png)

For each form that you select, you will be prompted to choose which subset of fields you would like to export.

[](https://malttanzania.surveycto.com/images/surveycto/help/client-subsets4.png)

By default, all chosen data for a given form exports to a single .csv file. However, you can also organize your exported data into multiple .csv files. Simply choose the fields by which you want exports organized. For example, if you choose to organize first by district, then by block, and then by village, then your .csv files will be named like "Form title - SUBSET - district - block - village.csv", and each file will only include the appropriate subset of data. This makes it easy to, for example, monitor survey operations in different areas or for different teams.

Once configured, your chosen subset of data will export to .csv file(s) whenever new .csv data is exported (to the same .csv export location). Primary .csv export files will contain all data, and subset .csv export files (named with "SUBSET" in their file names) will contain only the fields you selected.

#### [Using Microsoft Word and mail merge to output and view data](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_exporting-with-sync_using-word)

By default, *SurveyCTO Sync* exports data as .csv or comma-separated-values format. This is ideal for opening the data in Excel, Stata, or any other program that treats data as columns and rows. Unfortunately, this format may not be ideal for users who wish to review the data. For example, you may have a data scrutiny process that partly involves people who manually look over incoming data (to catch mistakes that can't be easily caught by automated computer review). For human review, .csv format is not ideal.

The solution is to use Microsoft Word's "mail merge" feature to merge the computer-readable data with a more user-friendly template. SurveyCTO assists in two ways:

1. You can download a starter template for any existing form. On the [Design](https://malttanzania.surveycto.com/main.html#_blank) tab of your server console, select *Download* and then *Mail merge template*under any form in the *Your forms* section, then specify how you want to label your fields and [whether you have *SurveyCTO Sync* configured to include group names in .csv file headers](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_exporting-with-sync_export-options). You can then customize this template or use it as-is.

[](https://malttanzania.surveycto.com/images/surveycto/help/client-preferences-mailmerge.png)

1. Sync will auto-merge exported data for you. Go into Sync's preferences, click the *Export options* tab, and check the *Auto-run mail merge after export* checkbox. Then, copy the mail-merge template you downloaded from the server into the destination directory into which you export your .csv files, and name it *formid\_template.docx* (where formid is the unique ID of the form, like "sampleform"). Whenever a new .csv file is exported for that form, a *formid\_merged.docx* file will also be output; that file will contain the combination of the latest .csv data with the mail merge template.

Note that fields within [repeat groups](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_additional-topics_repeats) (i.e., fields enclosed by *begin repeat* and *end repeat* rows) are not included by default in the mail-merge output. This is because such fields are exported into separate .csv files, and the mail-merge process only works with a single, primary .csv file. To manually add repeat-group data to a mail merge template, you would need to do the following:

1. Enable the Sync preference to export repeated fields in "wide" format.
2. Open the template and manually merge it with the *(title)\_WIDE.csv* file now exported by Sync.
3. Add any of the repeated-field columns from the *(title)\_WIDE.csv* file into your template to merge the data from those columns.
4. Re-merge with the updated *(title)\_WIDE.csv* file whenever you want to update the mail-merge output with new data.

Also note that mail-merging large amounts of data will require large amounts of memory. If you receive an error about "Java heap space," then it means that you ran out of memory. By default, *SurveyCTO Sync* will use up to 1GB of memory; to use up to 4GB, you can download [SurveyCTO\_Sync.l4j.ini](https://malttanzania.surveycto.com/images/surveycto/help/SurveyCTO_Sync.l4j.ini) into the same directory from which you run *SurveyCTO Sync*, then close and re-run Sync. (The name of the .ini file must match the name of your *SurveyCTO Sync*application file. So, e.g., if your application is named "SurveyCTO Sync.app", change the underscore to a space so that the .ini file is named "SurveyCTO Sync.l4j.ini".)

#### [Using Microsoft Excel for back-office operations](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_exporting-with-sync_using-excel-bo)

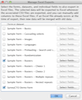
Your back-office work is likely far from over once you have exported your data, imported it into a data-analysis package, and/or mail-merged it into Word documents. Chances are, you have teams scrutinizing and correcting data, performing back-checks, or otherwise working to assure the quality of your data. Managing that process can be a challenge, particularly when new data is continually flowing in from the field.

To help make your job easier, SurveyCTO can automatically export form or dataset data to Excel documents, merging exported data with other back-office details that you record in that spreadsheet. Here's how it works:

1. Start by beginning the *SurveyCTO Sync* data-export process the way you usually do: press the first *GO!* button and select the forms and/or datasets that you would like to export (including any for which you would like to export to Excel). Before clicking the second *GO!* button, check the *Show export options after download* checkbox down at the bottom. Then, click the second *GO!* button, and when the export options window pops up, check the*Export and merge to Excel files for some forms and datasets* checkbox.

[](https://malttanzania.surveycto.com/images/surveycto/help/client-preferences-tracker0.png) [](https://malttanzania.surveycto.com/images/surveycto/help/client-preferences-tracker1.png)

1. Next, click the *Manage Excel exports...* button and check off any forms and/or datasets for which you would like Excel worksheets to be maintained. For each one that you check, click *Select fields...* to choose which fields you would like included in Excel. (For datasets, you will also need to specify a dataset field to use as the unique identifier. This is because each exported row must have a column that serves to uniquely identify it; for forms, this is automatic because each form submission has a unique key – but for datasets you have to specify an identifier column.)

[](https://malttanzania.surveycto.com/images/surveycto/help/client-preferences-tracker2.png) [](https://malttanzania.surveycto.com/images/surveycto/help/client-preferences-tracker3.png)

1. Once the Excel export has been enabled for a form or dataset, SurveyCTO will maintain a *TITLE\_status.xls* file for that form or dataset, located in the .csv output directory (where *TITLE* is the title of your form or dataset). Whenever new data is exported, SurveyCTO will add new rows to that spreadsheet. Each row will have columns filled in with the unique row identifier as well as the data associated with any fields you checked when configuring in Step 2 above.
2. You can edit the exported spreadsheet and add new columns to contain whatever additional, back-office details you wish.
3. Whenever you export new data, SurveyCTO will add rows to the spreadsheet and/or update data in the columns being exported; it will never delete rows and it will never change anything entered into any of the new columns that you added yourself. In those additional columns you can maintain whatever other details are useful for managing your back-office operations.

You have many options for how you manage your back-office operations, so you may not need these Excel exports. However, they are available if you need them.

#### [Automatically executing outside processes](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_exporting-with-sync_outside-processes)

After *SurveyCTO Sync* has downloaded and exported new data, you can have it automatically execute outside processes (e.g., a Stata .do file to process and review the new data).

To do this, go into Sync's preferences, click the *Export options* tab, and check off the *Auto-run other processors after export* checkbox.

[](https://malttanzania.surveycto.com/images/surveycto/help/client-preferences-autorun.png)

Then, within the destination directory into which you will export your .csv files, simply create a *thenrun* subdirectory (named "thenrun" exactly) that contains any .cmd, .do, or other executable files that you wish to run. Sync will automatically execute each of these files after exporting new .csv files to that destination. (When there are multiple files to execute, it will execute them in alphabetical order.)

If you will be analyzing your data in Stata, see [*How do I create templates for importing into Stata?*](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_exporting-with-sync_using-stata) for how to get started with auto-generated .do files.

#### [Running SurveyCTO Sync on a schedule](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_exporting-with-sync_scheduling-sync)

[](https://malttanzania.surveycto.com/images/surveycto/help/client-schedule.png)You might want to schedule *SurveyCTO Sync* to download and export new data automatically, in the background. To do so, run Sync and configure your options like normal, press *GO!*, and select which forms to download/export. Then, rather than pressing the second *GO!* button to execute once, choose *Schedule sync...* from the *Tools* menu. The download/export will happen right away, then again every *x* minutes, hours, or days, depending on the options you specify. You can leave Sync running in the background, then close it whenever you want it to stop.

#### [Working with cold-room computers](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_exporting-with-sync_cold-room-computers)

Typically, you use *SurveyCTO Sync* to sync from *Server* to *Local CSVs* in order to download the latest data from the server and export it to .csv format. However, there are other possible uses.

When dealing with highly sensitive [encrypted data](https://malttanzania.surveycto.com/help.html#Help_Designing-forms_additional-topics_encrypting), for example, you may want to decrypt and export data only on a non-Internet-connected "cold-room" computer. In this scenario, you would sync from *Server* to *External drive* on an Internet-connected computer, switch the external drive to your cold-room computer, then sync from *External drive* to *Local CSVs* on the cold-room computer. That way, you maintain your private key and decrypted data only on the cold-room computer, where they are safe.

Note that whenever you sync from *Server* or from *External drive*, all new data is always stored in the local Sync storage directory. This is true even when you select a destination other than *Local*. For example, a sync from *Server* to *Local CSVs* actually goes *Server* -> *Local* -> *Local CSVs*. At any time, then, you can choose to sync from your local storage (*Local*) to some destination, or sync from somewhere to your local storage. You might do this, for example, if you wanted to re-export .csv files without downloading any new data from the server.

And don't worry: your encrypted data is stored safely (as still encrypted) in Sync's local storage. It is only decrypted when you choose to export it to *Local CSVs*, using the private encryption key that you specify.

#### [Managing users and access](https://malttanzania.surveycto.com/help.html#Help_Monitoring-and-management_the-basics_managing-users)

If your user account permits you to administer users, you will see a *Your users* section at the bottom of each tab of your server console. There, you can add, edit, or delete user accounts, each of which can have a level of access appropriate for their role. In particular:

* For users who will fill out forms, you should create one or more accounts with "Data collection only" permission. Two things to keep in mind for these kinds of users: if users are filling out forms with a web browser, you can allow them to submit forms without logging in (in which case you don't need to create a user account for them); and if users are using the *SurveyCTO Collect* Android app to fill out forms, more than one device can be configured to share a single user account (so you don't need a separate account for each device).
* Users who will download and analyze your data will need "Data manager (collection and download)" permission.
* Users who will be adding, updating, or deleting forms or datasets on your server will need "Form and data manager (can administer forms and datasets)" permission.
* Your user account will need "Form, data, and user manager (can also administer users)" permission in order to manage user accounts. You can also create additional manager accounts that share this capability.

Note that the user named *manager* cannot be deleted and will always have full access. This is because the *manager* user account is the primary account through which your overall SurveyCTO subscription is managed.

You should be careful to assign secure passwords to every account, but be especially careful to use strong passwords for all of your manager accounts.

#### [Managing data](https://malttanzania.surveycto.com/help.html#Help_Monitoring-and-management_the-basics_managing-data)

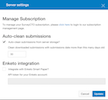
By default, none of the SurveyCTO components automatically delete any data. That means that old data piles up over time on devices, on the server, and on your computer. With old data being stored redundantly, there is little opportunity for data to be lost.

The accumulation of old data can cause the SurveyCTO system to slow down over time, however, as all data is constantly being stored, scanned, decrypted, and exported. It is therefore advisable to periodically clean out old data – once you are sure that it has been safely exported and backed up.

In *SurveyCTO Collect*, you can clean old form data by choosing *Delete Saved Form* from the main menu. Generally, you will only want to delete forms that are labeled as "Sent," as they are the forms have been safely transmitted to the server.

[](https://malttanzania.surveycto.com/images/surveycto/help/collect-deleted-saved.png)

On the server, you can manually eliminate old form data by choosing the *Delete* action for any form on the [Design](https://malttanzania.surveycto.com/main.html#Design) tab. This will delete the form, all of its data, and the history of all prior versions of the form. Another option is to choose *Purge form data* in the *Form submissions* section of the [Monitor](https://malttanzania.surveycto.com/main.html#_blank) tab to eliminate all submissions data for any listed form. This option will purge all submissions up to a date of your choosing (leaving the form itself intact, as well as the history of prior form versions). If you want, you can also configure the server to automatically clean data over a certain number of days old: as the "manager" user, simply click the gear icon in the top-right of your server console to configure this option.

[](https://malttanzania.surveycto.com/images/surveycto/help/server-preferences.png)

In the *SurveyCTO Sync* application on your computer, you can choose the *Preferences...* menu option to set the number of days to keep data in Sync's local storage. Data older than the number of days specified there will be cleaned from local storage and no longer included in subsequent exports.

[](https://malttanzania.surveycto.com/images/surveycto/help/client-preferences-autoclean.png)

We recommend that you set the server to maintain 30 days of data and Sync to maintain 45 days. You can choose any settings you wish, but note that you will not want to set Sync to store less data than the server. If you do, Sync will purge data, download it again from the server (since it will still be on the server), purge it again, and so on. This is because Sync always makes sure that it has all data from the server every time it syncs.

#### [Monitoring incoming data](https://malttanzania.surveycto.com/help.html#Help_Monitoring-and-management_the-basics_monitoring)

Data can be monitored as it comes in to your SurveyCTO server. Two tools, available on the [Monitor](https://malttanzania.surveycto.com/main.html#_blank) tab, allow you to view statistics on incoming submissions by time and by form:

1. **Submission statistics.** In the *Submission statistics* section of the [Monitor](https://malttanzania.surveycto.com/main.html#_blank) tab, you can plot form submissions over time for each of your forms. You can customize the time period displayed in the graph to show submissions over the last 7, 30, or 60 days, or, any individual month's submissions over the last year. This graph is downloadable to your local hard drive in either PNG or PDF format.
2. **Form submissions.** Also on the [Monitor](https://malttanzania.surveycto.com/main.html#_blank) tab, all of your forms will be listed under *Form submissions* with both the number of *Complete submissions*and date and time of the most recent submission reported below each. Once you have downloaded, exported, and backed up your data, you can also purge old data from here, so that it doesn't take up server storage or slow your data downloads.

For more sophisticated monitoring with dashboards or summary tables, you can publish subsets of your data – like key indicators – to Google Spreadsheets or Google Fusion Tables. From there, you can easily summarize the data and share it with your team. See [*Publishing data to the cloud*](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_publishing-forms) for a full discussion.

Also, it is not only important to track incoming submission statistics, but also to monitor the quality of your incoming data. See [*Monitoring the quality of incoming data*](https://malttanzania.surveycto.com/help.html#Help_Monitoring-and-management_managing-for-quality_monitoring-quality) for details on how you can configure automated quality checks.

#### [Managing users and access](https://malttanzania.surveycto.com/help.html#Help_Monitoring-and-management_the-basics_managing-users)

If your user account permits you to administer users, you will see a *Your users* section at the bottom of each tab of your server console. There, you can add, edit, or delete user accounts, each of which can have a level of access appropriate for their role. In particular:

* For users who will fill out forms, you should create one or more accounts with "Data collection only" permission. Two things to keep in mind for these kinds of users: if users are filling out forms with a web browser, you can allow them to submit forms without logging in (in which case you don't need to create a user account for them); and if users are using the *SurveyCTO Collect* Android app to fill out forms, more than one device can be configured to share a single user account (so you don't need a separate account for each device).
* Users who will download and analyze your data will need "Data manager (collection and download)" permission.
* Users who will be adding, updating, or deleting forms or datasets on your server will need "Form and data manager (can administer forms and datasets)" permission.
* Your user account will need "Form, data, and user manager (can also administer users)" permission in order to manage user accounts. You can also create additional manager accounts that share this capability.

Note that the user named *manager* cannot be deleted and will always have full access. This is because the *manager* user account is the primary account through which your overall SurveyCTO subscription is managed.

You should be careful to assign secure passwords to every account, but be especially careful to use strong passwords for all of your manager accounts.

#### [Managing data](https://malttanzania.surveycto.com/help.html#Help_Monitoring-and-management_the-basics_managing-data)

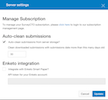
By default, none of the SurveyCTO components automatically delete any data. That means that old data piles up over time on devices, on the server, and on your computer. With old data being stored redundantly, there is little opportunity for data to be lost.

The accumulation of old data can cause the SurveyCTO system to slow down over time, however, as all data is constantly being stored, scanned, decrypted, and exported. It is therefore advisable to periodically clean out old data – once you are sure that it has been safely exported and backed up.

In *SurveyCTO Collect*, you can clean old form data by choosing *Delete Saved Form* from the main menu. Generally, you will only want to delete forms that are labeled as "Sent," as they are the forms have been safely transmitted to the server.

[](https://malttanzania.surveycto.com/images/surveycto/help/collect-deleted-saved.png)

On the server, you can manually eliminate old form data by choosing the *Delete* action for any form on the [Design](https://malttanzania.surveycto.com/main.html#Design) tab. This will delete the form, all of its data, and the history of all prior versions of the form. Another option is to choose *Purge form data* in the *Form submissions* section of the [Monitor](https://malttanzania.surveycto.com/main.html#_blank) tab to eliminate all submissions data for any listed form. This option will purge all submissions up to a date of your choosing (leaving the form itself intact, as well as the history of prior form versions). If you want, you can also configure the server to automatically clean data over a certain number of days old: as the "manager" user, simply click the gear icon in the top-right of your server console to configure this option.

[](https://malttanzania.surveycto.com/images/surveycto/help/server-preferences.png)

In the *SurveyCTO Sync* application on your computer, you can choose the *Preferences...* menu option to set the number of days to keep data in Sync's local storage. Data older than the number of days specified there will be cleaned from local storage and no longer included in subsequent exports.

[](https://malttanzania.surveycto.com/images/surveycto/help/client-preferences-autoclean.png)

We recommend that you set the server to maintain 30 days of data and Sync to maintain 45 days. You can choose any settings you wish, but note that you will not want to set Sync to store less data than the server. If you do, Sync will purge data, download it again from the server (since it will still be on the server), purge it again, and so on. This is because Sync always makes sure that it has all data from the server every time it syncs.

#### [Monitoring incoming data](https://malttanzania.surveycto.com/help.html#Help_Monitoring-and-management_the-basics_monitoring)

Data can be monitored as it comes in to your SurveyCTO server. Two tools, available on the [Monitor](https://malttanzania.surveycto.com/main.html#_blank) tab, allow you to view statistics on incoming submissions by time and by form:

1. **Submission statistics.** In the *Submission statistics* section of the [Monitor](https://malttanzania.surveycto.com/main.html#_blank) tab, you can plot form submissions over time for each of your forms. You can customize the time period displayed in the graph to show submissions over the last 7, 30, or 60 days, or, any individual month's submissions over the last year. This graph is downloadable to your local hard drive in either PNG or PDF format.
2. **Form submissions.** Also on the [Monitor](https://malttanzania.surveycto.com/main.html#_blank) tab, all of your forms will be listed under *Form submissions* with both the number of *Complete submissions*and date and time of the most recent submission reported below each. Once you have downloaded, exported, and backed up your data, you can also purge old data from here, so that it doesn't take up server storage or slow your data downloads.

For more sophisticated monitoring with dashboards or summary tables, you can publish subsets of your data – like key indicators – to Google Spreadsheets or Google Fusion Tables. From there, you can easily summarize the data and share it with your team. See [*Publishing data to the cloud*](https://malttanzania.surveycto.com/help.html#Help_Exporting-and-publishing-data_publishing-forms) for a full discussion.

Also, it is not only important to track incoming submission statistics, but also to monitor the quality of your incoming data. See [*Monitoring the quality of incoming data*](https://malttanzania.surveycto.com/help.html#Help_Monitoring-and-management_managing-for-quality_monitoring-quality) for details on how you can configure automated quality checks.