SurveyCTO

Help - Monitoring and management - The basics

Below, please find topics covering the basics of overall process management and monitoring incoming data.

Managing your server

If you have administrative or user-management access to this SurveyCTO server, then the Configure tab is where you go to configure server settings or manage user accounts. If you're the administrator for an enterprise installation of SurveyCTO, the Configure tab is also where you go to manage your enterprise configuration (including, e.g., custom user roles).

Individual server settings are discussed elsewhere in the documentation. For details on managing user accounts, see *Managing users and access*, just below.

Managing users and access

If your user account permits you to administer users, you will see a *Your users* section at the bottom of the Configure tab of your server console. There, you can add, edit, or delete user accounts. Each user account should have an access level appropriate to their role, and all SurveyCTO servers include at least the following roles:

- **Data collection only.** Users who will only fill out forms and submit data should be given "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).
- **Data manager (collection and download).** Users who will download and analyze your data will need "Data manager (collection and download)" permission.
- **Form and data manager (can administer forms and datasets).** 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.
- **Form, data, and user manager (can also administer users).** Users who need to be able to add, edit, or delete other user accounts will need "Form, data, and user manager (can also administer users)" permission.

- **Administrator (full access, plus subscription management).** Finally, users who need to be able to fully administer the account, including billing and other subscription options, will need "Administrator (full access, plus subscription management)" permission.

If you're using SurveyCTO as part of an enterprise configuration, shared with other teams, then user roles will likely be defined differently. In that case, the users listed in the *Your users* section will be those with access less than or equal to your own (e.g., everybody in your program team or country office).

All user accounts are identified by email addresses – except "Data collection only" users, who can be identified by either an email address or a non-email username. Sometimes, non-email usernames are easier to create, give out, configure on devices, etc.

If you create a user account with an email address, that user will be automatically sent a confirmation email with instructions on where and how to login. When you create such an account, you can also decide whether to invite the user to set their own password or choose a password for them; if you choose a password, you can opt to include or not include that password in the email confirmation.

If you choose a password, be careful to choose something secure. And be especially careful to use strong passwords for all accounts with administrator or user-management privileges.

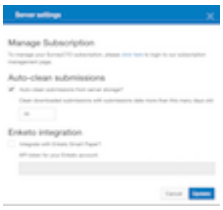
Managing data

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.

On the server, you can manually eliminate old form data by choosing the *Delete* action for any form on the 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 and dataset data* section of the Monitor 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 a user with full administrator access, simply go to the Configure tab, scroll down to *Server settings*, and choose *Storage* to configure this option.



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.

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

You'll want to monitor data as it comes in to your SurveyCTO server. Options available on the Monitor tab of your server console include:

1. **Submission statistics.** In the *Submission statistics* section of the Monitor 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 tab, all of your forms will be listed under *Form submissions and dataset data* 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.
3. **Monitor form data.** Also under *Form submissions and dataset data*, you can click the "Monitor form data" action for any form to jump into SurveyCTO's built-in *Data Explorer*. There, you can configure and save a monitoring workbook, review aggregate data as well as individual submissions, and catch potential data-quality issues right away. See *Using the Data Explorer to monitor incoming data* for more.
4. **Review workflow.** Also under *Form submissions and dataset data*, you can click the "Review workflow" action for any form to enable a review and correction workflow. Once this workflow has been enabled, you will have the opportunity to review and correct new submissions before they are released for publishing or export. See *Reviewing and correcting incoming data* for more.
5. **Automated quality checks.** Further down on the Monitor tab, the *Automated quality checks* section is where you can go to configure automated checks to supplement and complement your manual

monitoring efforts. See *Using quality checks to monitor the quality of incoming data* for details.

For more sophisticated monitoring with dashboards or summary tables, you can publish subsets of your data – like key indicators – to Google Sheets or Google Fusion Tables. From there, you can easily summarize the data and share it with your team. See *Publishing data to the cloud* for a full discussion.

Using the Data Explorer to monitor incoming data

You can use statistical software like Stata, SPSS, or R to visualize and analyze your data, or other software like Microsoft Excel or Google Sheets (both of which have free statistical-analysis add-ons). But to quickly and easily monitor your incoming data, you can use SurveyCTO's built-in *Data Explorer*.

Using the *Data Explorer*, you can easily summarize data submitted for individual fields, summarize the empirical relationships between fields, and drill down to browse individual submissions. With it, you can start learning from your data right away.

There are two ways into the *Data Explorer*:

1. Click the "Monitor form data" action for any form in the *Form submissions and dataset data* section of the Monitor tab.
2. Click the "Explore" action for any form in the *Your data* section of the Export tab.

The *Data Explorer* is mostly the same whichever path you take, but you can save and maintain separate workbooks for monitoring vs. exploration/analysis. Also, when you enter via the Monitor tab, quality-check results are summarized with your data (see the *Monitoring* section below).

However you enter the *Data Explorer*, you'll first need to choose which submissions – and maybe which fields – to load. Because all of the submissions you choose will be downloaded and loaded into your web browser's memory, you might need to choose only recent data or a random subset of data if your dataset is very large. Some computers with fast Internet connections, a lot of memory, and modern browsers will be able to load many thousands of submissions, but others could become too slow or even crash when loading too much data. If you're loading data for an encrypted form, you'll also need to choose whether to load all data (which requires the private key) or just fields that have been flagged as *publishable*.

Loading the *Data Explorer* might take a while, particularly the first time you open it; for large datasets, you might want to go grab a tea or coffee while it loads. But don't worry: downloaded data will be stored in your local browser's cache so that you won't need to download it again the next time you go in.

And once you're all loaded up, you can use the *Data Explorer* fully offline: you only need to connect again if you want to view or download an attached file (like a photo or audio recording), view or explore maps, or save your workbook. Saving your workbook is handy so that, when you come back later, all of your summaries will be there (ordered, organized, and configured just the way you left them – but also with any new data that's come in).

Exploring your data

When you first load the *Data Explorer*, you'll start with a mostly-blank canvas. At the top, you'll see some summary information about your form, including the submissions and fields you chose to load. Below that, you have the opportunity to add field summaries, relationship summaries, and groups:

1. **Field summaries.** These are summaries of your data, one field at a time. To quickly add field summaries for every field in your form, just click to add field summaries, choose *Select all*, and save

(but if you have hundreds or thousands of fields in your form, you might want to be more selective). Multiple-choice data will be summarized with horizontal bar or pie charts, numeric data with histograms (counting frequencies in an adjustable number of equal-sized "bins"), GPS data with maps, date/time data with vertical bar charts (with aggregation to days, weeks, or months), and text data with a simple list ordered by frequency. When multiple summary views are possible, you can click the "eye" icon to change a summary from one view to another (e.g., from categorical to numeric). All field summaries include a count ("N") of observations as well as a count of missing observations ("N missing"); min, max, median, mean, and standard deviation ("SD") are also included for all numeric data.

2. **Relationship summaries.** These are summaries of the bivariate relationship between two fields. For convenience, you can add more than one relationship at a time: just select your primary field first, then all of the other fields to which you'd like to relate that primary field (e.g., choose enumerator ID as primary and then a series of other fields, in order to see the relationship between enumerator and those other fields). Relationships will be summarized as either a scatterplot (for two numeric fields), a trend view (date/time + numeric), a map view (*geopoint* + anything), or a table (anything + anything); when multiple views are possible, you can always click the "eye" icon to switch between views.

   For scatterplots, correlation coefficient, R-squared, and OLS beta estimates are given, and you can show or hide an OLS line of best fit. For map views, each location pin is shaded or colored based on the value of the other field, and that other field's value is shown on hover. For trend views, a line chart is shown, and you can aggregate data at the day, week, or month level. For tables, you can choose between crosstab or summary-statistic views, depending on whether one or both fields is numeric; in a crosstab view, numeric data is grouped into an adjustable number of equal-sized "bins" just like histograms, but empty rows or columns are automatically hidden so that the crosstab doesn't get too big.

3. **Groups.** If you wish, you can create a group and then add summaries to that group (or drag and drop existing summaries into it). You can then collapse and expand groups, to show or hide the summaries within them. If you have a workbook with a lot of summaries, you might find groups to be helpful for organization and navigation.

Importantly, you can "drill down" to view individual submissions all through the *Data Explorer*: just click on any histogram bar, crosstab cell, scatterplot point, map marker, "N" count, etc. to view submission details. Your data will be nicely formatted (and even printable!) based on the latest version of your form. When viewing a submission, you can click to open or download any submission attachments, such as photos or audio recordings, or click to show question administration timing from an attached text audit – though this will require an active Internet connection. You can also click to copy a direct hyperlink to the submission details, in order to re-visit it later or share with a colleague (valid login and sufficient access rights will be required, of course). Finally, if you have the unique key (or UUID) for a submission, you can also jump straight to the submission-detail view from the Monitor tab's *Look up by key* action.

If your form has been translated into multiple languages, there will be a language selector in the top-right of the *Data Explorer*. Use that to choose the language in which to show field and option labels throughout the summaries and submissions.

You can feel free to add summaries and groups, drag and drop them into your preferred order, and then save your workbook so that it will be the same next time you come back – though every time you come back the data itself will update as appropriate. If you ever want to start completely over, just choose *Reset* in the upper-right.

Monitoring

A key to collecting high-quality data is keeping a close eye on data as it comes in. Use a combination of automated quality checks and manual inspection to catch potential problems quickly, so that they can be dealt with quickly (e.g., while teams are still available in the field). Configure your *Data Explorer* workbook and return to it frequently, as new data is collected.

If you enter the *Data Explorer* via the Monitor tab, then automated quality checks and quality-check results will be noted with your data. For example, say that you have a quality check configured to warn if there are submissions with *duration* less than 600 (i.e., less than 10 minutes). If you show a field summary for the *duration* field, then that quality check will be listed, along with any associated warnings from the current quality-check report. You'll even be able to easily click to view individual submissions that triggered warnings, and the details for those submissions will also note the related warnings. Finally, a summary of all configured checks and warnings will be available in the upper-right of the *Data Explorer*, in the CHECKS section of the page header.

Filtering and excluding data

Once you start using the *Data Explorer*, you're likely to want to filter or exclude some data. There are three options available to you:

1. **Exclude entire submissions.** You might have some submissions that were test or invalid submissions, which you want to exclude from summaries globally, for your entire workbook. Just click to view a submission, then click *View options* and *Exclude submission from summaries*. Back in the main window with your summaries, that submission will then be excluded. A yellow bar at the top of the screen will remind you of any exclusions; you can click the count of excluded submissions to view and un-exclude submissions individually, or you can click the icon in the yellow bar to quickly un-exclude all submissions.
2. **Exclude certain values.** Some outliers might make your field summaries look unpleasant, or they might throw off your summary statistics. For example, a -999 entered as "don't know" for income can really throw off the graph, the mean, and the standard deviation. Just click on a graph bar or data value, then select *Exclude from this summary* to drop the outlier(s) from the summary. Only the current summary view will be affected (not the workbook overall), and an orange bar at the top of the summary will remind you of the exclusion; click the icon next to the "Some values excluded" note in this bar to view or clear exclusions.
3. **Filter data.** As you explore your data, you may want to filter your workbook to focus on particular subsets. For example, you might want to see how the data looks when you consider only young women. To do this, you could click the "Female" bar in your field summary for *gender*, choose *Add to global data filter*, then click one or more bars in the *age* field's histogram to also add those to the global filter. This would filter all other summaries in the workbook, to only include submissions with the gender and age(s) you specified. Like with submission exclusions, a yellow bar at the top of the screen will remind you of any global filters, and you can click the filter icon in that bar to view or clear them. Using global filters, you can zero in on particular subsets of data and get a quick view of how that data looks.

Reviewing and correcting data

If the review and correction workflow is enabled for a form, then the *Data Explorer* is where you review, comment on, correct, approve, and reject submissions. Whenever you view the details for an individual submission, you will see the data as it exists after any corrections, plus any comments that have been made; options at the top of the screen will allow you to comment, correct, approve, or reject. Note that you will be able to review submissions and make changes while offline, but to apply those changes you will need to press the option to save them – which will require a working connection to the server.

Calculating and aggregating data

Sometimes you'll be interested in looking at the sum of multiple income measures, the average age of household members, or some other aggregation or calculation. Since the *Data Explorer* can only look at individual fields or individual relationships between two fields, you'll need to include important aggregations or calculations as fields in your form so that the results are readily available in your data. To do this, include one or more calculate fields to perform the appropriate aggregations or calculations. Once you do, you'll be able to add automated quality checks on those new fields, and they'll also be conveniently included in your exported data.

Advanced mode

In both the *Form submissions and dataset data* section of the Monitor tab and the *Your data* section of the Export tab, there is an *Advanced mode* button that allows you to enable a set of more powerful *Data Explorer* tools. Meant for more expert users, advanced mode allows you to:

- **Configure multiple workbooks.** By default, each form has one workbook on the Monitor tab and one on the Export tab. In advanced mode, you can configure as many additional workbooks as you like, tailoring each to a particular view, team, or workflow.
- **Attach datasets to workbooks.** Advanced mode also allows you to attach server datasets to workbooks, so that you can supplement incoming form data with earlier listing data, QC results from outside systems, and more.
- **Download and upload workbook definitions.** Finally, advanced mode includes *Download* and *Upload* buttons that allow you to export and import workbook definitions. These definitions are Excel spreadsheets, similar to form definitions and edited in a similar way; instead of defining fields in a form, however, these workbook definitions define summaries in a *Data Explorer* workbook.

See the help topic on advanced-mode usage for more details.

Data security

For the highest possible data security, we always recommend that you encrypt data using your own encryption keys. If you do, only you and your team will be able to read the data, because only you and your team will have access to the private encryption key necessary to decrypt it.

The *Data Explorer* can decrypt your encrypted data, but only if you allow your web browser to temporarily use your private encryption key. Here's how it works: your data is downloaded and stored in your web browser's local cache so that it doesn't have to be re-downloaded every time, but it's stored safely in its encrypted form; every time the *Data Explorer* loads, it will need you to select your private encryption key so that it can decrypt your data; your encryption key will not be saved in the browser cache and it will not be transmitted anywhere (it will only be used in memory); your decrypted data will likewise be decrypted in memory but it will not be saved anywhere, nor will it be transmitted anywhere; once you close the *Data Explorer* tab, the memory is freed and the decrypted data is no longer accessible.

This approach to data security has a few costs. For one, you need to select your private encryption key every time you load encrypted data (since we won't store it in your browser's local storage). For another, loading takes longer because we decrypt data every time you load (since we also won't store decrypted data in your browser's local storage). But the benefits are worth the costs: your data stays truly "for your eyes only," and your respondents' confidentiality is protected. We don't think that you'll find a more secure data-collection platform anywhere.

Hyperlinking into the *Data Explorer*

You may want to hyperlink directly from outside systems into SurveyCTO's *Data Explorer*. For example, if you have a dashboard to monitor top-level indicators, you may want to allow users to drill down into individual data points in the *Data Explorer*. To hyperlink directly into the submission-details view for a particular submission, just use a hyperlink like the following:

https://[servername].surveycto.com/view/submission.html?uuid=[KEY]

The one thing to be aware of is that the unique ID (the KEY) for the submission has a colon in it, and for a colon to appear in a URL it should be changed to *%3A* instead. If you were using Microsoft Excel and you had your SurveyCTO server name in cell A1 and your submission's KEY in cell B2, this formula would generate the proper hyperlink:

="https://" & A1 & ".surveycto.com/view/submission.html?uuid=" & SUBSTITUTE(B1,":","%3A")

But then, if you are using Excel – and you are on Windows – then you should know about another complication: Microsoft Office products on Windows are terrible about hyperlinking to secure pages like this one (pages that require a login); they'll freeze for a little while, and then they'll fail to open the page at all. Luckily, we've constructed a work-around. Say that you have your server name in cell A1, your submission KEY in A2, and the formula above (which generates the URL) in A3. In that case, this formula will generate a hyperlink that *does* work in the Windows version of Excel:

="https://" & A1 & ".surveycto.com/officelink.html?url=" & SUBSTITUTE(SUBSTITUTE(SUBSTITUTE(SUBSTITUTE(SUBSTITUTE(SUBSTITUTE(SUBSTITUTE(SUBSTITUTE(SUBSTITUTE(C1, "%", "%25"), " ", "%20"), "?", "%3F"), "&", "%26"), "=", "%3D"), "{", "%7B"), "}", "%7D"), "[", "%5B"), "]", "%5D")

That "officelink" URL sends users through an unsecured landing page that has a *Continue* button to login and continue. It's a bit of a hassle, but often worth it to get working hyperlinks directly from back-office Excel sheets.

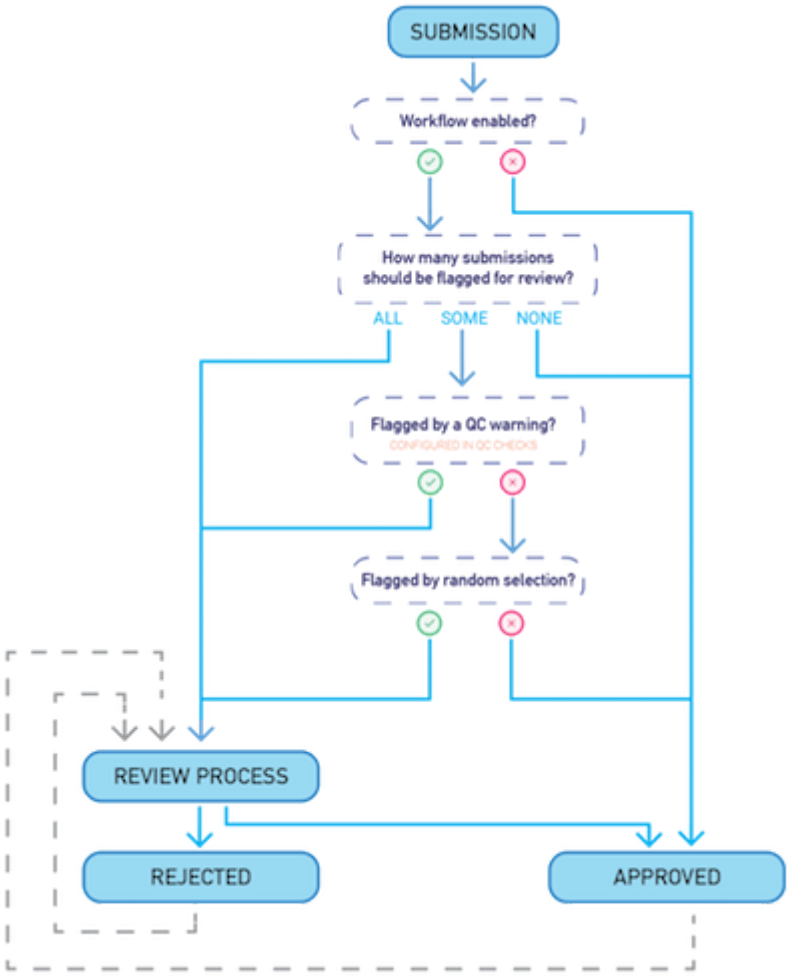
Reviewing and correcting incoming data

The data you collect with SurveyCTO comes in from web forms or mobile devices, it passes through the SurveyCTO server, and then it continues out to other systems via data exports, real-time publishing, or our various API's. You can think of the data as passing through a pipeline, transiting through the SurveyCTO server as part of its journey. For each form, there are two ways that you can configure the SurveyCTO part of the pipeline:

1. **Pass data immediately through (the default).** While you can and always should monitor the quality of your incoming data, SurveyCTO defaults to essentially auto-approving every incoming submission as soon as it comes in, releasing it for immediate publishing or export. Once a submission has been finalized and submitted to the SurveyCTO server, it is passed to you unchanged, however it came in. If you find problems in the data, you correct those problems further along in the pipeline, outside of SurveyCTO (e.g., in Stata).
2. **Review and correct incoming data, before releasing it downstream.** Alternatively, you can enable the review and correction workflow for a form, in which case new submissions can be initially held awaiting review. During review, you can examine submissions, comment on them, and make corrections to the data; when ready, you either approve or reject each submission. Only approved submissions are then released downstream via data exports, real-time publishing, and our various API's.

Review and correction workflow: enabling and configuring

You can use the *Review workflow* button to enable or disable the review and correction workflow for any form in the *Form submissions and dataset data* section of the Monitor tab. When the workflow is disabled (the default), submissions pass through the SurveyCTO server unchanged. When enabled, you choose which submissions to flag and hold for review, before releasing them for publishing or export:



- **ALL.** This is the default and almost certainly the setting you should use for new forms. All incoming submissions will be held awaiting review, only released to downstream systems once they have been approved for export or publishing. When testing or piloting a new form, it's often best to look closely at all incoming data, and it gives you a chance to get used to the review and correction process.
- **SOME.** When you don't have the resources to closely examine every incoming submission, this is often the best choice for deployed forms. Here, you would hold only some submissions for close inspection, based on quality checks and/or random selection. (Exact options discussed below.)
- **NONE.** If your QC process requires that you use downstream systems or processes to flag submissions for review, you might default to holding no incoming submissions for review. (See our separate help topic for more on advanced workflow configurations like this.)

Whichever submissions *aren't* held for review will be auto-approved for export and publishing as soon as they come in. If you choose to hold *SOME* submissions for review, then you get to decide which submissions to hold; you can select any of the following options:

- **Flag incoming submissions based on results of quality checks.** If you configure automated quality checks for your form, you can use the results of those checks to flag which submissions warrant additional review. (Choose the "only critical" option if you want to restrict consideration to only those quality checks configured as *critical*.)
    - **Flag any submission with a submission-specific QC warning.** If you select this option and an incoming submission triggers warnings for *Value is too low*, *Value is too high*, or *Value is an outlier* checks, the submission will be held for review. (Note that outlier tests for incoming submissions will be based on the interquartile range calculated the last time the full set of quality checks were run. This is typically the prior night when checks are configured to run nightly.)

    - **Flag any submission that is part of a group that triggered a QC warning during the last full evaluation.** If you select this option and an incoming submission is part of a group that triggered a *Group mean is different* or *Group distribution is different* warning in the prior run (generally, the prior night when all checks are run nightly), the submission will be held for review. So for example, if you have a quality check to monitor the gender distribution by enumerator, and a particular enumerator triggered a warning due to their distribution being different, selecting this option would result in all submissions from that enumerator being held for review.
    - **Flag any submission that would further contribute to a field-specific QC warning raised during the last full evaluation.** If you select this option and an incoming submission has a field value that would worsen an existing *Value is too frequent*, *Mean is too low*, or *Mean is too high* warning, that submission will be flagged for review. For example, if the last full quality-check evaluation triggered a warning for a specific field having too many -888 (don't know) responses and an incoming submission also had a -888 response in that field, the submission would be held for review. Likewise, if a field's mean is already triggering warnings because it is above or below a certain threshold – and an incoming submission has a value also above or below that threshold – then the submission would be held for review.
- **Flag a random percentage of submissions.** Even if you use quality checks to flag submissions for review, it's always good practice to review some percentage completely at random. After all, your quality checks can't catch everything, and in fact random review can help you to identify where and how to strengthen your automated checks. So here, you can specify what percentage of submissions to flag and hold at random. This percentage only applies to those submissions not already flagged based on quality checks (so, e.g., if 20% of incoming submissions get flagged for quality checks and you configure 10% to be flagged at random, only 10% of the 80% that aren't already flagged would be flagged at random).

Note that you can use *Value is too low* or *Value is too high* warnings to trigger review based on arbitrary criteria. For example, say that you wanted to review all submissions for women over 75 years old. You could add a *calculate* field into your form, with a calculation expression like "if(${gender}='F' and ${age}>75, 1, 0)", and then add a *Value is too high* quality check on that field, to warn whenever it is greater than 0. You could then configure your workflow to flag incoming submissions for review whenever they have submission-specific QC warnings, which would include any submissions for women over 75 years old.

In addition to choosing which submissions to hold for review, you can also configure a range of other workflow options. For example, you can require that users add comments to explain every correction they make to the data, enable or disable a bulk "approve all" option, and, for encrypted forms, choose whether to encrypt comments or leave them open (so that team members without the private key can see them). You can also allow un-approving or un-rejecting submissions, so that your team can reconsider any submission at any time; before you enable un-approving and un-rejecting, though, you should read about the potential consequences for downstream data systems.

Review and correction workflow: actual review, correction, and classification

When there are submissions awaiting review, you will be alerted on the Monitor tab, and you will find options to review and approve submissions in the *Form submissions and dataset data* section of that tab. There, you will be able to approve waiting submissions in bulk, review them one-by-one, or explore them in aggregate.

You review submissions in the *Data Explorer*, where you can safely review even encrypted data. For each submission you review, your tasks are as follows:

1. **Review.** Fundamentally, you want to take a careful look at each submission, to assess its quality. Factors to consider include the logical consistency between different responses, corroborating photos

and GPS locations, audio audits, "speed limit" violations, the time spent on individual questions, and more. See the help topic on *Collecting high-quality data* for a broad discussion of considerations and tools at your disposal.

2. **Classify.** You'll need to classify the initial quality of the submission, how it was on receipt. Your options are: *GOOD* (no problems found), *OKAY* (one or more minor problems), *POOR* (serious and/or many problems), or *FAKE* (fake or fraudulent responses). Classification is subjective, so just do your best to choose the most reasonable classification for each submission.
3. **Comment and correct.** As you discover problems, you can make corrections to the data. You can also add comments, either to the submission overall, or to individual fields or corrections. When it comes to making corrections, please note the following:
    - SurveyCTO allows you to edit the raw data for most fields, and there is no validation to constrain the changes you make – so you have to be careful not to introduce mistakes or invalid data.
    - If you encrypt your data, don't worry: even your corrections will be encrypted so that they are only visible to those who have your private key.
    - The only correction you can make to a GPS field is to delete/clear a value that you believe to be incorrect; you cannot edit such a field to change one GPS position to another.
    - The only correction you can make to a file field – like a photo or audio recording – is to delete/clear values that you believe to be incorrect; you cannot edit such fields to alter the attached files in any way.
    - You cannot change the number of responses within a repeat group; so if there are responses for eight household members in a household roster repeat group, for example, then you can change and/or clear the individual responses, but you cannot add or delete entire repeat instances (there will remain eight, as per the original submission).
4. **Approve or reject.** Ultimately, you will need to approve or reject each submission, at which point it exits the review and correction process. Approved submissions are immediately released for export and publishing. (If you configure your workflow to allow un-approving and un-rejecting, this determination is not final. But you should read about the potential consequences for downstream data systems if you enable these options.)

Review and correction workflow: exporting data

When the review and correction workflow is enabled for a form, you still export your data as you normally do. However, please note the following:

1. When you download data from the *Your data* section of the Export tab, the default is to include only approved submissions – but you can also download rejected or awaiting-review submissions if you select the options to include those as well. If you do include submissions that haven't yet been approved, the .csv export file will include an extra *review_status* column so that you will know the status of each submission.
2. SurveyCTO Sync v2.40+ will only export approved submissions. If you have an older version of Sync on your computer (pre-2.40), it will export all submissions regardless of status. (Older versions of Sync will also fail to apply any corrections to exported data. If you're using the review and corrections workflow, you should be sure to use the latest version of Sync available at the top of your server console's Export tab.)

3. Your exported .csv data will include *review_comments* and *review_corrections* columns with a history of comments and corrections (and changes in approval status will be automatically included as comments).
4. Your exported .csv data will include a *review_quality* column with the quality classification recorded for each submission (if any).
5. Your exported .csv data will include all corrections made to the data – unless you are using an older version of Sync or Sync v2.40+ with the *Apply corrections* setting manually turned off.
6. If you download data from the *Your data* section of the Export tab and corrections have been made to that data, then the download page will include a *yourformid_correction_log.csv* file that contains a complete history of corrections made to the data; the columns in that correction log are as follows: *KEY*, *field*, *old value*, *new value*, *user*, *date/time*, *comment*.

To learn more about the format of exported data, see *Understanding the format of exported data*.

Review and correction workflow: publishing data

When the review and correction workflow is enabled for a form that publishes to the cloud or publishes to a server dataset, only approved submissions will publish. This includes submissions that are auto-approved on receipt (unless your workflow is configured to hold *all* submissions for review), and those submissions that are held for review will publish only if and when they are approved. Also, all corrections will always be applied to published data. (If you allow un-approving, submissions can publish more than once. See the help topic on advanced workflows to learn more.)

Review and correction workflow: quality checks

When you configure automated quality checks for a form that uses the review and correction workflow, you have a decision to make: which submissions should be included when the full set of automated quality checks are evaluated?

By default, all submissions are included – including those that have not yet been reviewed and approved. After all, if you're checking for outliers (for example), then it might be useful to know that a new submission you're reviewing is an outlier; the quality-check reports might be one key input into your review process, as you decide which submissions to approve, what corrections to make, etc.

In the *Automated quality checks* section of the Monitor tab, you can click *Options* for any form to change which submissions should be included in the checks. You might, for example, choose to exclude rejected submissions, so that they don't throw off the statistical distributions used to assess differences, outliers, and means.