

DIME Field Coordinator Workshop

Track 2: SurveyCTO (advanced)

June 18, 2018
Washington, DC



Agenda for advanced track

1. **Conceptual introduction:** Lessons on quality
2. **Catching up:** Recent SurveyCTO updates
 - a. Likert scale
 - b. Review and correction workflow
 - c. Data Explorer (DE) improvements
3. **Tooling up:** Advanced features and Q&A
 - a. Advanced programming
 - b. Case management
4. **Coming soon:** Machine learning training data



Conceptual introduction: Lessons on data quality



Today, thousands of users in more than 150 countries depend on SurveyCTO for their data collection.





DIME, IPA, and J-PAL

Collecting quality data in some
of the world's toughest places.

Lessons from our users

Consistent lessons have emerged:

- Enumerator jobs are hard
 - Rain, snow, extreme heat
 - Travel, finding respondents
 - When to ask questions
 - How to ask questions
- The temptations are many
 - Faking an interview
 - Interviewing the wrong respondent
 - Auto-answering
 - Incentive to answer one way
 - Just know the answer
- Local firms have weaker incentives for quality

Must design with these lessons in mind.



Designing for quality data

What you can do in survey design:

- Skip irrelevant questions
- Constrain responses
- Design for monitoring
 - Field-level supervision
 - Office-level monitoring
 - Audio audits, speed limits
 - Automated quality checks (HFC's)
 - Publishing and dashboards
 - Watching the data from day one
 - Tracking corrections, quality





Catching up: Recent updates

Photo: Laura Elizabeth Pohl / EGPAF

Likert appearance

SurveyCTO

likert appearance

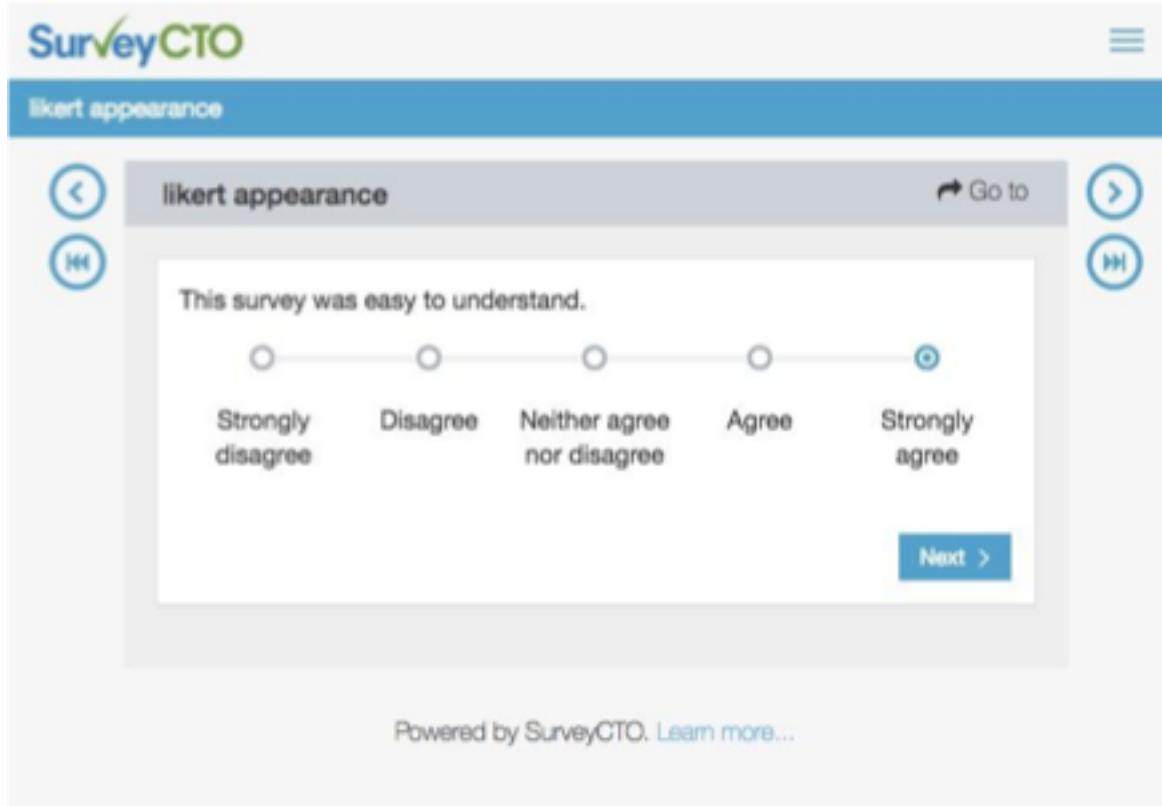
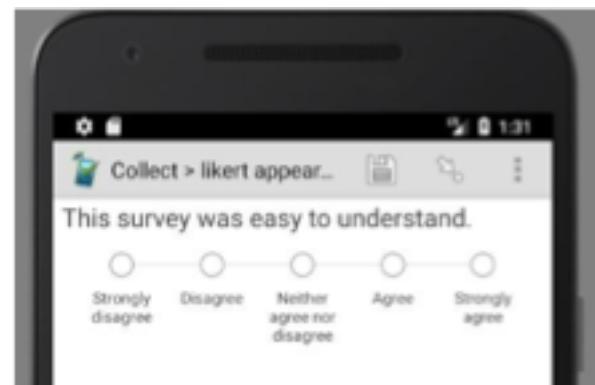
likert appearance Go to

This survey was easy to understand.

Strongly disagree Disagree Neither agree nor disagree Agree Strongly agree

Next >

Powered by SurveyCTO. [Learn more...](#)

The screenshot shows a mobile application interface for SurveyCTO. At the top, the SurveyCTO logo is visible. Below it, the title "likert appearance" is displayed. A navigation bar includes arrows for back and forward, a menu icon, and a "Go to" button. The main content area contains a question "This survey was easy to understand." followed by a five-point Likert scale with radio buttons. The options are "Strongly disagree", "Disagree", "Neither agree nor disagree", "Agree", and "Strongly agree". A blue "Next >" button is at the bottom right. At the very bottom of the screen, there is a footer note: "Powered by SurveyCTO. [Learn more...](#)".

Likert appearance - image

SurveyCTO

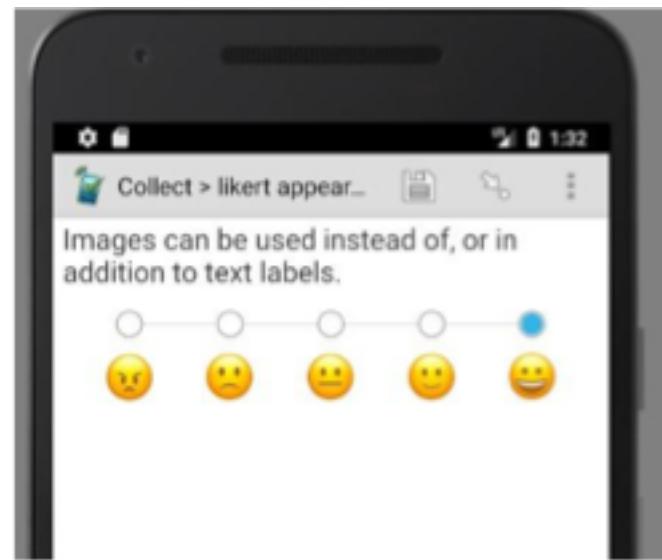
likert appearance

likert appearance Go to

Images can be used instead of, or in addition to text labels.

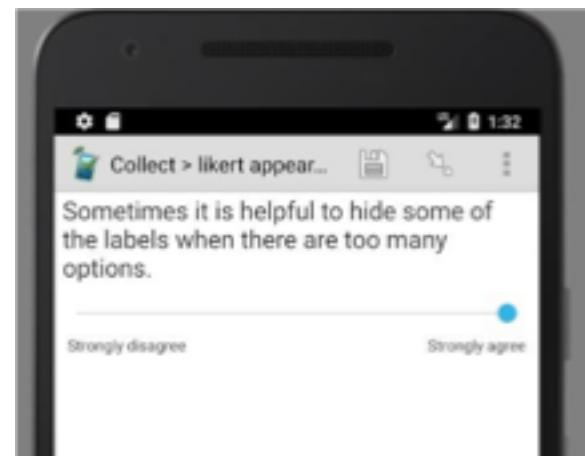
Next >

The SurveyCTO mobile application interface displays a Likert scale with five yellow circular emoji faces arranged horizontally. From left to right, the emoji faces represent increasing levels of agreement: a very angry face, a sad face, a neutral face, a smiling face, and a very happy face. Above the emoji faces is a horizontal line with five open circles, where the fourth circle from the left is filled blue, corresponding to the selected emoji. Navigation buttons for back, forward, and menu are visible on the left, and a 'Go to' button is on the right. A message at the top states: "Images can be used instead of, or in addition to text labels." A "Next >" button is at the bottom right. The footer says "Powered by SurveyCTO. Learn more..."



Likert appearance - slider

The screenshot shows a SurveyCTO mobile form titled "likert appearance". The main section contains a question: "Sometimes it is helpful to hide some of the labels when there are too many options." Below the question is a horizontal slider with two labels: "Strongly disagree" on the left and "Strongly agree" on the right. A blue dot indicates the current position of the slider. At the bottom of the slide area is a blue "Next >" button. On the left side of the screen are navigation icons: a back arrow, a forward arrow, and a refresh symbol. On the right side are icons for more options and a menu. The footer of the screen says "Powered by SurveyCTO. Learn more...".



Hyperlinking from Excel

Example cell A1:

```
presentation
```

Example cell B1:

```
https://presentation.surveycto.com/view/submission.html?uuid=uuid%3A06d355e9-7c34-4e2a-87cb-25f0840151a3
```

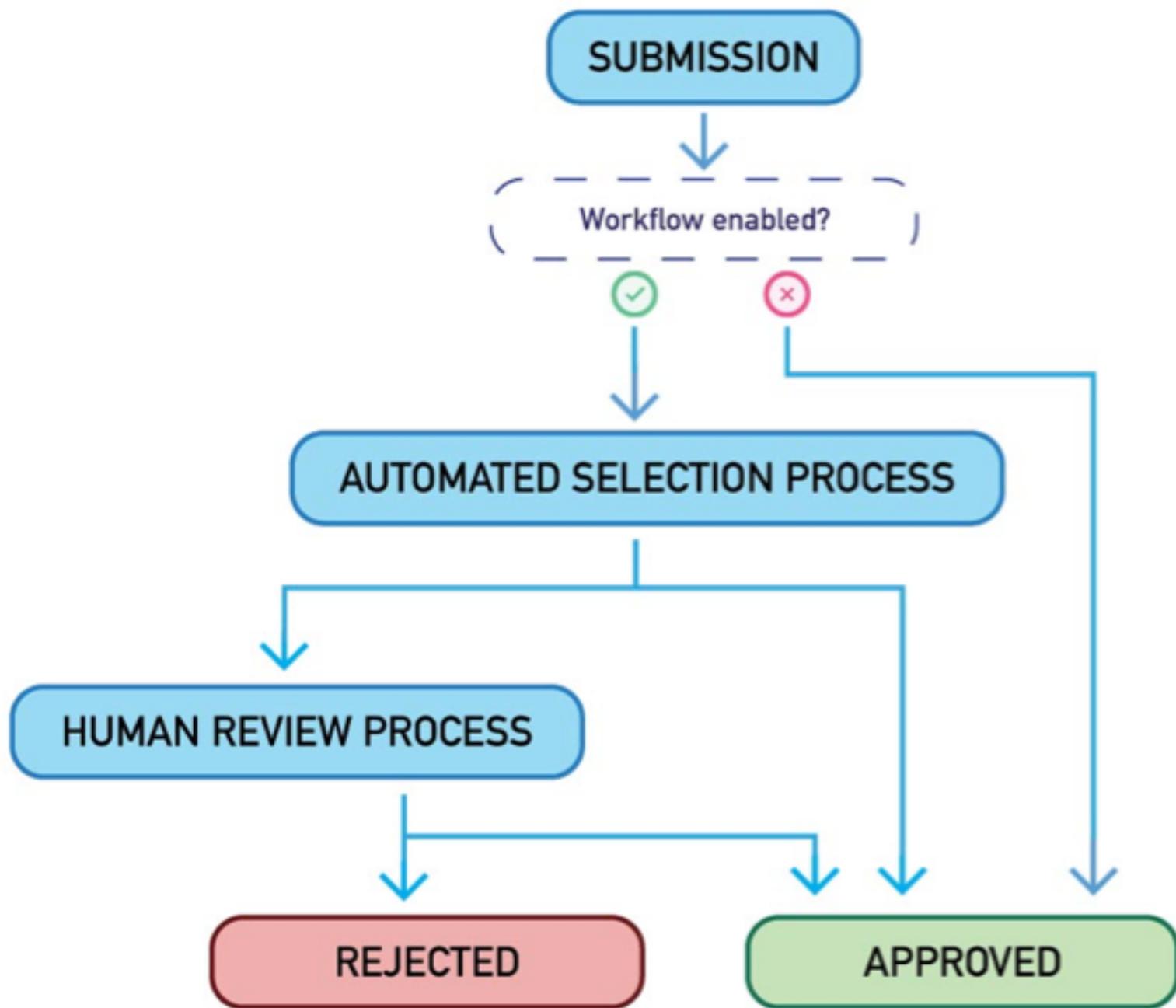
Example cell C1:

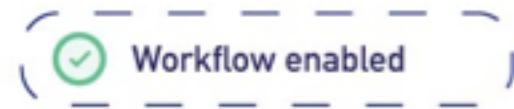
```
="https://" & A1 &
".surveycto.com/officelink.html?url=" &
SUBSTITUTE(SUBSTITUTE(SUBSTITUTE(
SUBSTITUTE(SUBSTITUTE(SUBSTITUTE(
SUBSTITUTE(SUBSTITUTE(SUBSTITUTE(B1, "%",
"%25"), " ", "%20"), "?", "%3F"), "&",
"%26"), "=" , "%3D"), "{", "%7B"), "}" ,
"%7D"), "[", "%5B"), "]", "%5D")
```



Review and correction workflow







Workflow enabled

SUBMISSION



(How many submissions
should be flagged for review?)

ALL SOME NONE



(Flagged by a QC warning?)

CONFIGURED IN QC CHECKS



(Flagged by random selection?)



HUMAN REVIEW PROCESS



Review and correction - workflow options

all field types

Form ID: all_field_types, Complete submissions: 5 (latest Mar. 7, 2018 at 5:43:48PM; approved: 4)

Review workflow settings

The review and correction workflow allows you to review each submission before it is released for publishing or export, making corrections, rejecting, or approving as appropriate. If this workflow is enabled, all submissions must be approved before they are published or exported. [Learn more...](#)

Enable review and correction workflow for this form? YES

Submissions to flag and hold for review
Set which submissions will be flagged for review. These submissions will require review before being published.

None Some All
All incoming submissions will be flagged for review.

Comments required for corrections
Require the user to enter a comment before they are allowed to make a correction Comments are not required

Allow un-approve and un-reject
Allow submissions which have already been either approved or rejected to be un-approved or un-rejected and placed back into the review process for further commenting or correcting. Un-approve and un-reject are disabled

Allow approve all
 Approve all is enabled

Review and correction - selective flagging

Submissions to flag and hold for review

Set which submissions will be flagged for review. These submissions will require review before being published.

None Some All

Some incoming submissions will be flagged for review.

Flag incoming submissions based on results of quality checks

Requires at least one quality check to be set up. Learn how to set up quality checks [here...](#)

All Only critical

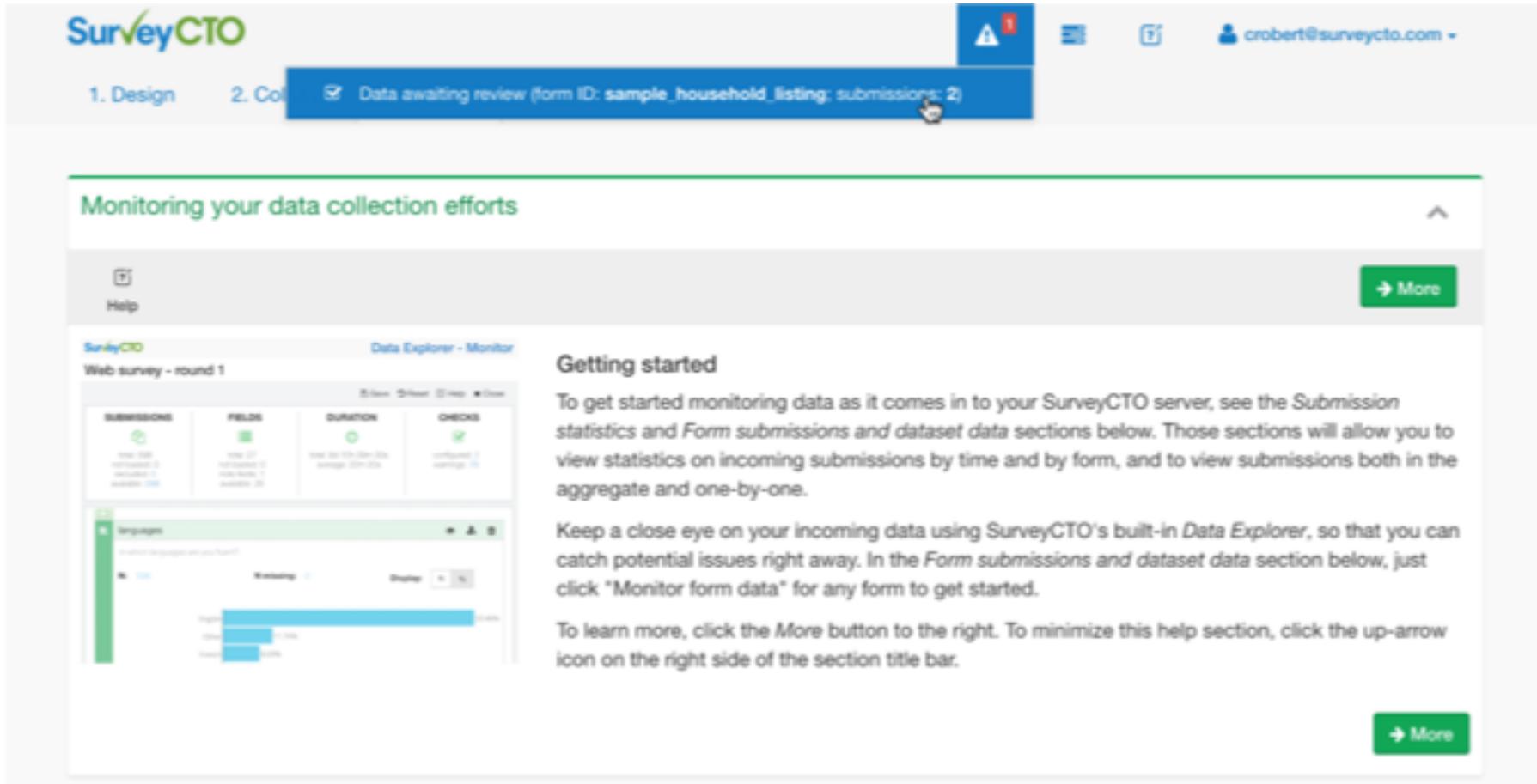
Remember to set quality checks to run nightly to update group-specific statistics.

- Flag any submission with a submission-specific QC warning (e.g., "value is too high")
- Flag any submission that is part of a *group* that triggered a QC warning during the last full evaluation (e.g., "group mean is different")
- Flag any submission that would further contribute to a field-specific QC warning raised during the last full evaluation (e.g., "value is too frequent")

Flag a random percentage of submissions

Select percent at random

Review and correction - using - 1



The screenshot shows the SurveyCTO Data Explorer - Monitor interface. At the top, there's a navigation bar with tabs for 'Design' and 'Collection'. A blue banner indicates 'Data awaiting review (form ID: sample_household_listing; submissions: 2)'. Below the banner, the main area is titled 'Monitoring your data collection efforts'. It features a 'Help' button and a 'More' button. On the left, there's a sidebar with a 'SurveyCTO' logo and 'Web survey - round 1'. The main content area has four sections: 'SUBMISSIONS' (2 total), 'FIELDS' (1 total), 'DURATION' (Total: 00:00:00, Average: 00:00:00), and 'CHECKS' (1 confirmed, 0 pending). Below these sections is a 'Languages' section with a chart showing submission counts for English, French, and Spanish.

Getting started

To get started monitoring data as it comes in to your SurveyCTO server, see the Submission statistics and Form submissions and dataset data sections below. Those sections will allow you to view statistics on incoming submissions by time and by form, and to view submissions both in the aggregate and one-by-one.

Keep a close eye on your incoming data using SurveyCTO's built-in *Data Explorer*, so that you can catch potential issues right away. In the *Form submissions* and *dataset data* section below, just click "Monitor form data" for any form to get started.

To learn more, click the *More* button to the right. To minimize this help section, click the up-arrow icon on the right side of the section title bar.

More

Review and correction - using - 2

The screenshot displays the SurveyCTO software interface, specifically the 'Monitor' section. At the top, there's a navigation bar with tabs: '1. Design', '2. Collect', '3. Monitor', and '4. Export'. The '3. Monitor' tab is selected, indicated by a dropdown arrow. To the right of the tabs, there are icons for a red alert, a blue square, a yellow square, and an envelope, followed by the email address 'crobert@surveycto.com'. Below the navigation bar, a header bar says 'Form submissions and dataset data'. The main content area shows two datasets:

- Sample household listing:** Form ID: sample_household_listing. Complete submissions: 2 (latest Oct. 13, 2017 at 10:10:55AM; approved: 0). A message box indicates 'New data awaiting review and approval (submissions: 2). (Unapproved submissions are not released for publishing or export.)' with three buttons: 'view now', 'Monitor now', and 'Approve all'. Action buttons for 'Monitor form data', 'Review workflow', 'Purge form data', and 'Look up by key' are also present.
- Case management example:** Dataset ID: cases, Records: 19. A 'Cases' section is shown with a 'Purge data' button. Action buttons for 'Monitor form data', 'Review workflow', 'Purge form data', and 'Look up by key' are also present.

Review panel

household_survey_town_1 | Household Survey - Town 1 X

Submission from Thu Apr 05 2018 08:30:23 GMT-0400 (EDT) #8 of 8

Options |

Flagged and held for review (not yet released) ?

This submission has been flagged for further review. Please follow the steps below to complete the review process.

1. Classify submission **2. Comment and correct** **3. Approve or reject**

Good Comment... Approve
 Okay How do I make a correction? Reject
 Poor How do I comment on a field?
 Fake

starttime
Thu Apr 05 2018 08:28:43 GMT-0400 (EDT)

endtime
Thu Apr 05 2018 08:30:22 GMT-0400 (EDT)

deviceid
(web)

subscriberid

Review panel - in progress

household_survey_town_1 | Household Survey - Town 1 X

Submission from Thu Apr 05 2018 08:30:23 GMT-0400 (EDT) #8 of 8

Options

Flagged and held for review (not yet released)

This submission has been flagged for further review. Please follow the steps below to complete the review process.

1. Classify submission

Good
 Okay
 Poor
 Fake

2. Comment and correct

How do I make a correction?
How do I comment on a field?

3. Approve or reject

Approve
 Reject

cking@surveycto... a minute ago
[5252] changed to [52].
age

cking@surveycto... a minute ago
I spoke to Nemesh; he typed the age twice in a row.
age

John Doe
address - What is your address? 50 Main St
age - How old are you? 52
gender - What is your gender? Male

Review panel - complete

household_survey_town_1 | Household Survey - Town 1

Submission from Thu Apr 05 2018 08:30:23 GMT-0400 (EDT) #8 of 8

Options

✓ Reviewed (Approved) Classified as okay [Click here to un-approve](#)

address - What is your address?
50 Main St

age - How old are you?
52

gender - What is your gender?
Male

num_in_hh - How many people live in your household?
5

Other details

KEY - Unique submission ID (instanceID)
uuid:50d8edb4-4f0e-4ece-aa57-5927f5aa806f

formdef_version - Form version used
1804041246

cking@surveycto... a minute ago
[5252] changed to [52].
age

cking@surveycto... a minute ago
I spoke to Nemesh; he typed the age twice in a row.
age

cking@surveycto... a few seconds ago
[Submission approved.
Classified as OKAY.]

Submission details - corrections and comments

district - ENUMERATOR: What district are you in?

Guruve

Guruve



Add a comment to describe why you are making this correction...

district - ENUMERATOR: What district are you in?

Mazowe

cking@surveycto.com

a few seconds

I checked the GPS location field and this survey was actually done in Mazowe district.

cking@surveycto.com

a few seconds ago

[Guruve] changed to [Mazowe].

Global comments with links to fields

hh_listing_2 | Sample household listing form with sample data >

Submission from Sun Mar 18 2018 21:49:19 GMT-0400 (EDT) #1 of 1

Options



ENUMERATOR: Which enumeration team are you on?

Southern Team

enumerator - ENUMERATOR: Which enumerator are you?

Malcolm Tonderai

province - ENUMERATOR: What province are you in?

Mashonaland Central

district - ENUMERATOR: What district are you in?

Mazowe

gps_location - ENUMERATOR: Please record your location.

42.5244492 -70.8932004 0.0 25.0 [View](#)

consent - ENUMERATOR: Does an adult in this household give informed consent?

Yes

cking@surveycto...

a few seconds ago

[Guruve] changed to
[Mazowe].

district

cking@surveyct...

a few seconds

I checked the GPS location
field and this survey was
actually done in Mazowe
district.

district

cking@surveyct...

a few seconds ago

[Submission approved.]

▲ Questions if have consent

respondent_name - What is your name?

Chuck

Field comments

enumerator_team - Welcome to the sample household listing form, meant mostly to provide some sample data for exploration and export.

ENUMERATOR: Which enumeration team are you on?

Southern Team

Comment

Cancel

district - ENUMERATOR: What district are you in?

Mazowe



Field comments (and correction history)

Questions to ask if consent has been given

respondent_name - What is your name?

John Doe

address - What is your address?

50 Main St

age - How old are you?

52



cking@surveycto.com

11 minutes ago

I spoke to Nemesh; he typed the age twice in a row.



cking@surveycto.com

11 minutes ago

[5252] changed to [52].

Review and correction - export - corrected data

	G	H	I	J	K	L	M	N	O	P	C
1	enumerator	province	district	gps_location-Latitude	gps_location-Longitude	gps_location-Altitude	gps_location-Accuracy	consent	respondent_name	address	num_i
2	Precious_Rutendo	Manicaland	Mutare	-19.01314962	32.58230732	0	0	1	Tadiwa James	172 Faradza Close	
3											
4											
5											
6											
7											
8											
9											
10											
11											
12											
13											
14											
15											
16											
17											
18											
19											
...											

Review and correction - export - new columns

	T	U	V	W	X
1	instanceID	formdef_version	KEY	review_comments	review_corrections
2	uuid:945b6981-5a34-4be4-a6c5-4729d572c122	1710131403	uuid:945b6981-5a34-4be4-a6c5-4729d572c122	crobert@surveycto.com [Oct 13, 2017 2:34:50 PM]: Believe that s was missing from last name. crobert@surveycto.com [Oct 13, 2017 2:35:13 PM]: [Submission approved.]	crobert@surveycto.com [Oct 13,
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
^					

Review and correction - export - correction log

Sample household listing

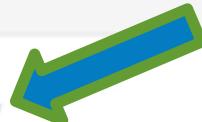
[☰ Explore](#) [Help](#) [✖ Close](#)

Please find your export files below. Click on each to save it to your local computer.

Note: The export format and time zone for date and time fields have been set according to your local web browser (English, United States).

Exported data

Sample household listing_WDE.csv
sample_household_listing_correction_log.csv



OFF [Show attachments? \(2\)](#)

Other download options:

[SurveyCTO Sync](#)

[Printable version](#)

[Mail merge template](#)

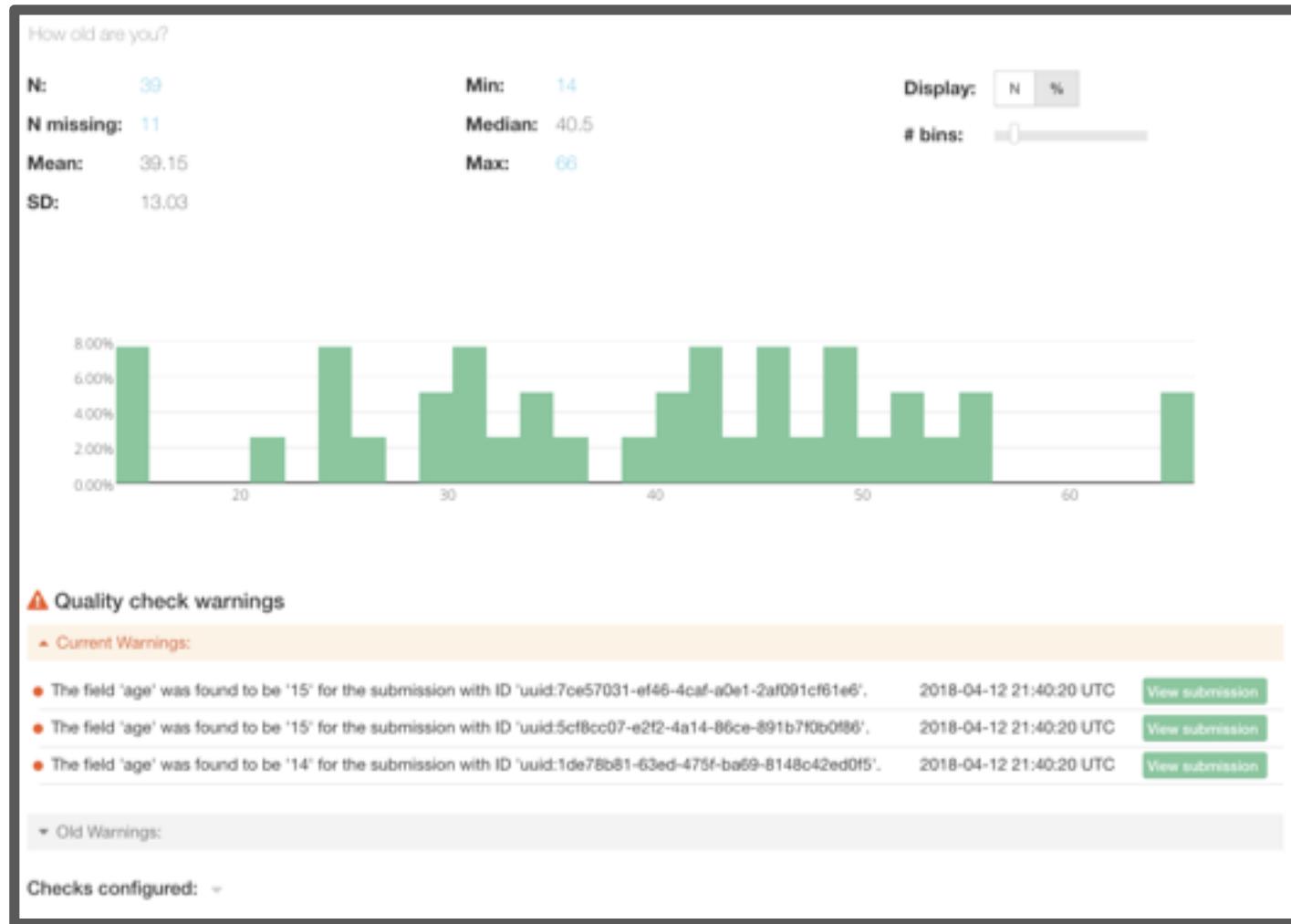
[Stata .do template](#)

The most powerful and flexible way to download, process, and export your data is to use SurveyCTO Sync, software that runs on your local computer. [Click here](#) to learn more about SurveyCTO Sync, and get started by downloading and installing the appropriate version, based on your computer's operating system: [download for Windows](#), [download for Mac \(OSX\)](#), or [download for other platforms](#).

Review and correction - export - correction log

A	B	C	D	E	F	G	H	I
KEY	field	old value	new value	user	date/time			
2	uuid:945b6981-5a34-4be4-a6c5-4729d572c122	on_consent[1]/respondent_name	Tadiwa Jame	Tadiwa James	crobert@surveycto.com	10/13/17 14:34		
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								
24								
25								
26								
27								
28								
29								
30								
31								

QC warnings in Data Explorer



QC warnings in submission details

Submission from Wed Apr 04 2018 14:53:52 GMT-0400 (EDT) #1 of 3

Options | [Edit](#) | [Delete](#) | [Share](#)

✓ Reviewed (Rejected) Classified as poor [Click here to un-reject](#)

▲ Quality checks warnings Current: 1, Old: 0

▲ Current warnings

- Value is too low: last reported on 2018-04-12 21:40:20 UTC age
The field 'age' was found to be '15'.

enumerator - Which enumerator are you?
Billy

consent - ENUMERATOR: Does an adult in this household give informed consent?
Yes

▲ Questions to ask if consent has been given

respondent_name - What is your name?
John Doe

address - What is your address?
45 Main St

age - How old are you?
15 A B

QC warnings in submission details (at field)

▲ Questions to ask if consent has been given

respondent_name - What is your name?

John Doe

address - What is your address?

80 Main St

age - How old are you?

47

gender - What is your gender?

Male

num_in_hh - How many people live in your household?

12

● Value is an outlier: The value '12' is an outlier in the field 'num_in_hh' (Q1: 3.0; Q3: 4.5).



Attaching a dataset to a workbook

The screenshot shows the SurveyCTO software interface. At the top, there is a navigation bar with tabs: 1. Design, 2. Collect, 3. Monitor (selected), 4. Export, and Configure. To the right of the navigation bar is a user profile icon and the email address `meletis@surveycto.com`. Below the navigation bar, the main area displays the title "Likert dyn tests" and the form ID "likert_dyn_tests". It also shows "Complete submissions: 0". On the right side of this section are four buttons: Review workflow, Purge form data, Look up by key, and Advanced mode.

The main content area is titled "Monitor (default)" and shows "Workbook ID: monitor, Summaries: 0". Below this, there is a section titled "Attach datasets to this workbook". This section includes a "Help" button and a "My datasets" list which contains an item named "somedataset: SomeDataset". A note below the list states: "We will merge your dataset with your form data, linking using an ID column in your dataset that can be matched with an ID field in your form." There is a link "Click here... to learn more." and a diagram illustrating the data merging process. The diagram shows a grid of colored squares (blue, green, purple) connected by arrows to a central search icon, representing how the dataset is linked to the form data.

At the bottom of the "Attach datasets to this workbook" dialog, there are two input fields: "Form ID field" set to "KEY" and "Dataset ID Column" set to "a". There are also "Cancel" and "Save" buttons at the bottom right.

Data Explorer workbook definitions

The screenshot shows a Microsoft Excel spreadsheet titled "test_wkbk-1803190713". The table has columns labeled A through F. Column A contains the type of field, column B contains the field name, column C contains the field_2 name, column D contains the label, column E contains the exclusion, and column F contains the options. The bins column is empty. The rows represent various field types and their corresponding names and labels.

	A	B	C	D	E	F	bins
1	type	field	field_2	label	exclusion	options	
2	temporal	starttime					
3	temporal	endtime					
4	text	deviceid					
5	text	subscriberid					
6	text	simid					
7	text	devicephonenum					
8	text	username					
9	numeric	duration					
10	text	caseid					
11	numeric	int1					
12	numeric	int2					
13	text	KEY					
14	categorical	formdef_version					
15	temporal	SubmissionDate					
16	categorical	review_status					
17							
18							
19							
20							
21							

Uploading a new DE workbook definition

The screenshot shows the SurveyCTO interface for managing a 'Crosstab data' form. At the top, there are several icons: a blue square with a white minus sign, a magnifying glass, a circular arrow, a search icon, and an 'Advanced mode' icon. Below these are buttons for 'Review workflow', 'Purge form data', 'Look up by key', and another 'Advanced mode' button. A message indicates 'Form ID: crosstab_data, Complete submissions: 18 (latest Mar. 8, 2018 at 7:54:18PM)'. In the center, there are two green buttons: 'Start new workbook' and 'Upload workbook definition'. A modal dialog box titled 'Upload workbook definition' is open. It has a help icon and a question mark icon. It includes tabs for 'Upload from computer' (selected) and 'Upload from Google Drive'. A text field prompts 'Please choose a workbook definition file to upload (.xlsx):' with a 'Select file...' button. A preview area shows the file 'Monitor__default_-1803190711.xlsx'. At the bottom are 'Cancel' and 'Upload' buttons.

Crosstab data

Form ID: crosstab_data, Complete submissions: 18 (latest Mar. 8, 2018 at 7:54:18PM)

Start new workbook

Upload workbook definition

Upload workbook definition

Help ?

Upload from computer Upload from Google Drive

Please choose a workbook definition file to upload (.xlsx):

Select file...

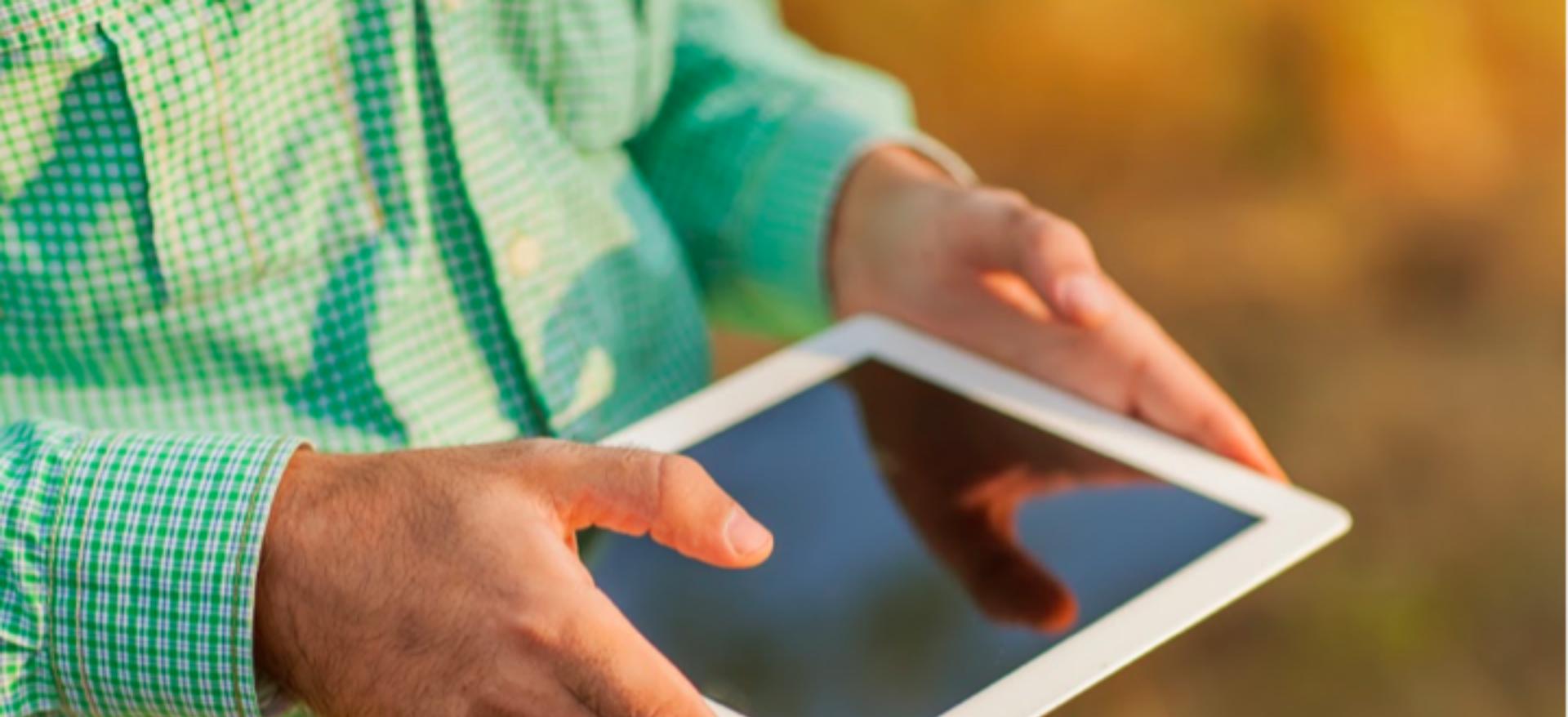
Monitor__default_-1803190711.xlsx

Cancel Upload

Our recommended QC workflow

1. Enable review and correction workflow, randomly selecting some small subset for close review
2. Run HFC's in Stata as per usual
3. Upload HFC results into a server dataset, with the submission KEY in one column
4. Attach that dataset to the *monitor* Data Explorer workbook for the form
5. Hyperlink into the Data Explorer to investigate potential problems
6. If you need to make corrections to the data, un-approve, correct, then re-approve





Tooling up: Advanced features

Advanced features: agenda

1. SurveyCTO's datasets and preloading features
2. Case management
3. SurveyCTO integrations with external platforms
4. SurveyCTO's offline sync capability
5. Non-linear form navigation
6. Advanced programming example: random draws without duplicates
7. Other Q&A



Datasets



Datasets and preloading

SurveyCTO A 1 ≡ ? faizan.diwan@gmail.com ▾

1. Design ▾ 2. Collect 3. Monitor 4. Export Configure

How to design your survey forms ▾

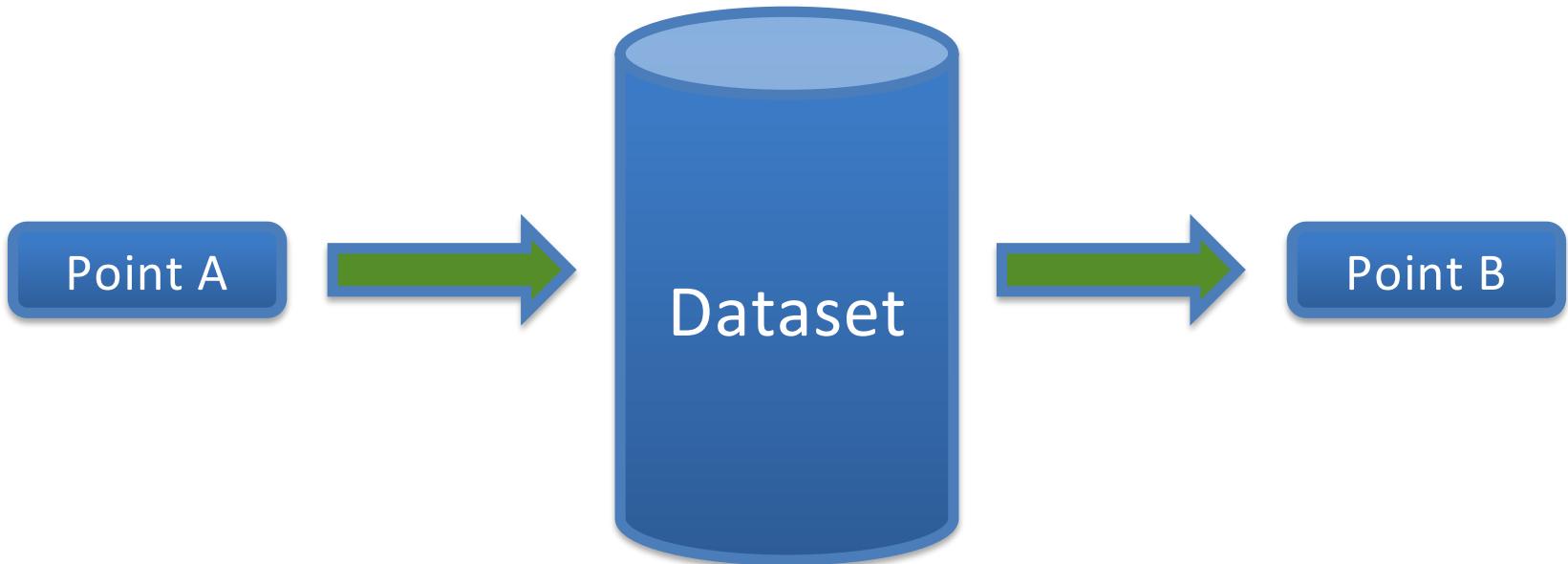
Your forms and datasets ^ ?

Tools Refresh Search Organize Help

 Sample household listing form with sample data	Preview	Edit	Upload	Tools	Encryption	Download	Duplicate	Delete
Form ID: hh_listing_2, Deployed version: 1710070852								
 Sample dataset	Attach	Publish into	Edit	Upload	Download	Rename	Purge data	Delete
Dataset ID: sample_dataset, Version: 0, Records: 0								

 Click here to start new form, upload form definition, add group, or add server dataset

Datasets and preloading



Datasets and preloading

Imagine a spreadsheet:

ID	Field2	Field3	Field4
Survey1			
Survey2			
Survey3			
Survey4			

All you see is this:

SurveyCTO

1. Design ▾ 2. Collect 3. Monitor 4. Export Configure

faizan.diwan@gmail.com ▾

How to design your survey forms ▾

Your forms and datasets ▾ ?

Tools Refresh Search Organize Help

	Sample household listing form with sample data	Preview	Edit	Upload	Tools	Encryption	Download	Duplicate	Delete
	Form ID: hh_listing_2, Deployed version: 1710070852								
	Sample dataset	Attach	Publish into	Edit	Upload	Download	Rename	Purge data	Delete
	Dataset ID: sample_dataset, Version: 0, Records: 0								

Click here to start new form, upload form definition, add group, or add server dataset



Click “Add server dataset”:

The screenshot shows the SurveyCTO web application interface. At the top, there is a navigation bar with tabs: 1. Design (selected), 2. Collect, 3. Monitor, 4. Export, and Configure. To the right of the tabs are icons for notifications (with 1 notification), settings, and user information (faizan.diwan@gmail.com). Below the navigation bar, there is a section titled "How to design your survey forms" with a dropdown arrow icon. Underneath this, there is a section titled "Your forms and datasets" with a collapse/expand arrow icon and a help icon. A toolbar below this section includes icons for Tools, Refresh, Search, Organize, and Help. The main content area displays a form titled "Sample household listing form with sample data". To the right of the form are buttons for Preview, Edit, Upload, Tools, Encryption, Download, Duplicate, and Delete. Below the form title, it says "Form ID: hh_listing_2, Deployed version: 1710070852". At the bottom of the page, there are several buttons: "Start new form", "Add group", "Upload form definition", and "Add server dataset". The "Add server dataset" button is highlighted with a red rectangular box.

1. Design

2. Collect

3. Monitor

4. Export

Configure

Your forms and datasets

Create a new dataset

?



Help

New dataset for data

New dataset for cases

New dataset from definition

Dataset title:

The dataset title here...

Dataset ID:

The dataset id...

Optionally, upload a file with initial dataset contents (.csv file). You can upload it directly from your computer or use Google Drive:

Upload from computer

Upload from Google Drive

Please choose a dataset file to upload (.csv, optional):

Select file...

Cancel

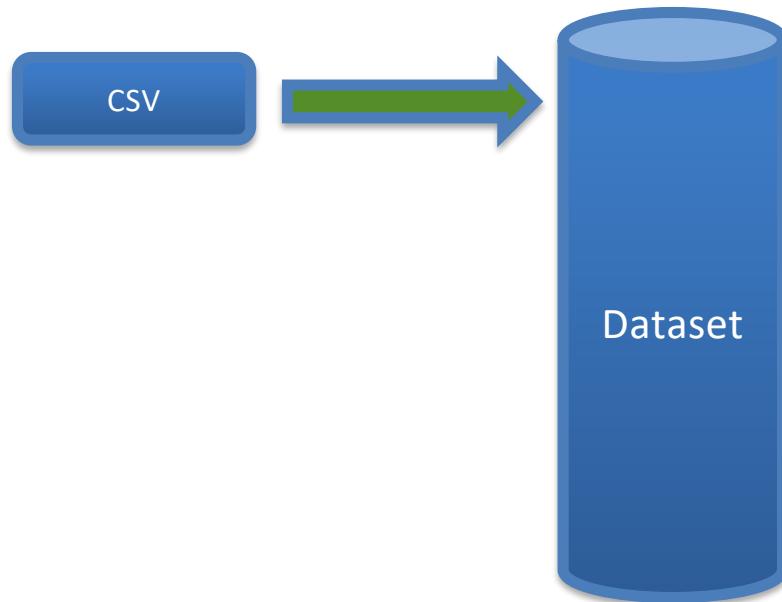
Create dataset

Datasets and preloading

Initially, the dataset is completely empty:

Datasets and preloading

You can get data into a dataset in several ways:



Upload a .csv when you create a dataset:

The screenshot shows the SurveyCTO interface for creating a new dataset. At the top, there are tabs for '1. Design', '2. Collect', '3. Monitor', '4. Export', and 'Configure'. The 'Configure' tab is selected. On the right, it shows 'Your forms and datasets' with a user icon and email 'faizan.diwan@gmail.com'. Below the tabs, the main title is 'Create a new dataset' with a help icon and question mark.

Three options are available for dataset creation: 'New dataset for data' (selected), 'New dataset for cases', and 'New dataset from definition'.

Form fields include:

- Dataset title:** A text input field with placeholder 'The dataset title here...'
- Dataset ID:** A text input field with placeholder 'The dataset id...'

Below these fields, a note says: 'Optionally, upload a file with initial dataset contents (.csv file). You can upload it directly from your computer or use Google Drive.'

Two upload options are shown:

- 'Upload from computer'
- 'Upload from Google Drive'

A red box highlights the 'Select file...' button in the 'Upload from computer' section, which is associated with the CSV upload area. The text in this area says: 'Please choose a dataset file to upload (.csv, optional)'.

At the bottom right are two buttons: 'Cancel' and 'Create dataset'.

1. Design

2. Collect

3. Monitor

4. Export

Configure

Your forms and datasets



Sample dataset



Attach

Publish into

Edit



Upload



Download



Rename



Purge data



Delete

Dataset ID: sample_dataset, Version: 0, Records: 0

Upload dataset



Help

Upload from computer

Upload from Google Drive

File with new dataset contents (.csv file):

Select file...

- Append this data to existing dataset contents
- Merge this data with existing dataset contents
- Replace existing dataset contents with this data

Cancel

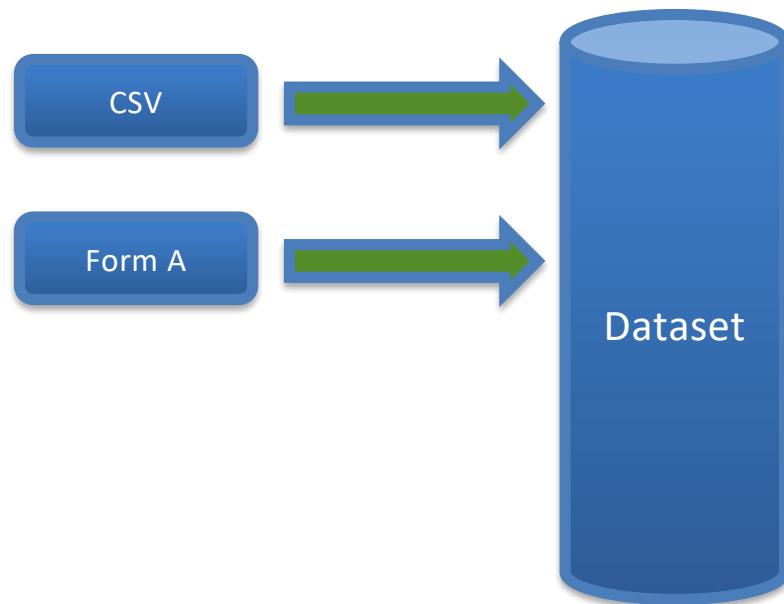
Upload



Click here to start new form, upload form definition, add group, or add server dataset

Datasets and preloading

You can get data into a dataset in several ways:



Stream data from a form into a dataset:



faizan.diwan@gmail.com

1. Design

2. Collect

3. Monitor

4. Export

Configure

Your forms and datasets

Published to datasets: Sample dataset (sample_dataset)



Sample dataset



Publish into



Dataset ID: sample_dataset, Version: 0, Records: 0

Forms that publish to dataset with ID 'sample_dataset'

Sample household listing form with sample data
ID: hh_listing_2

Edit

Delete

+ Add a form

Close

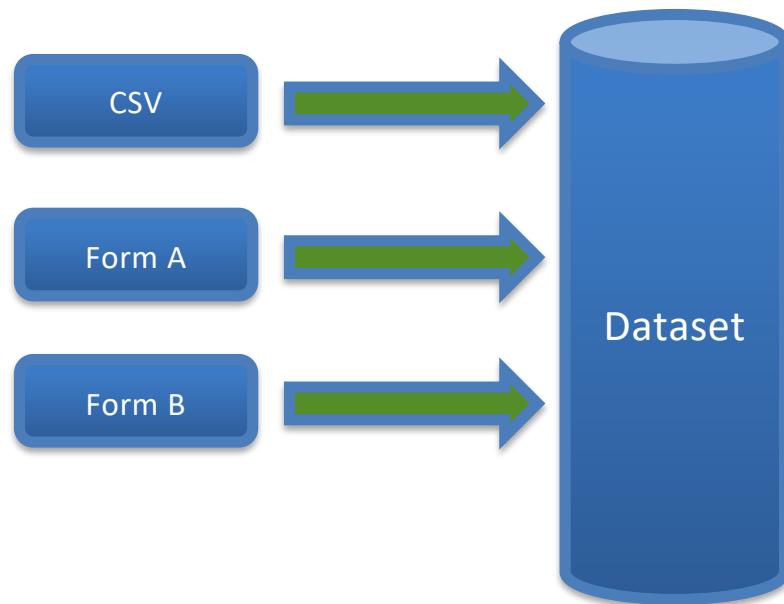
Published from: Sample household listing form with sample data (hh_listing_2)



← Click here to start new form, upload form definition, add group, or add server dataset

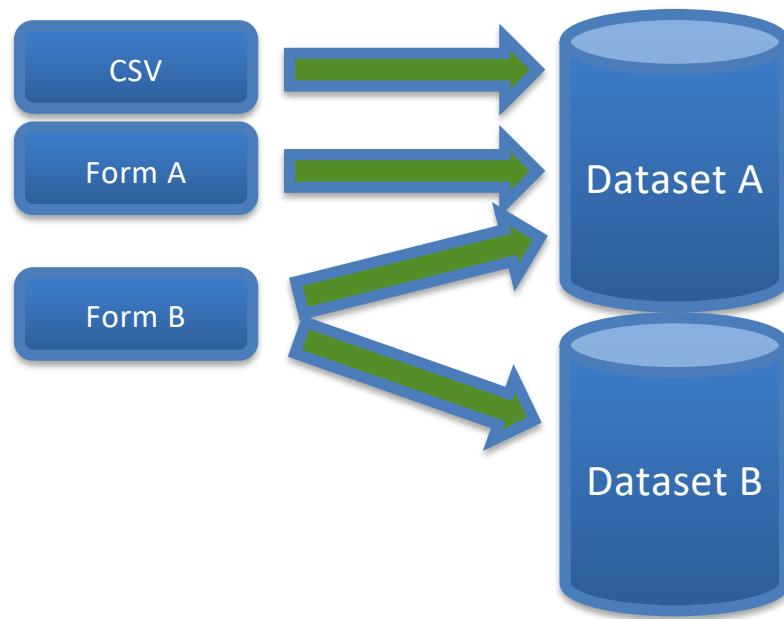
Datasets and preloading

You can get data into a dataset in several ways:



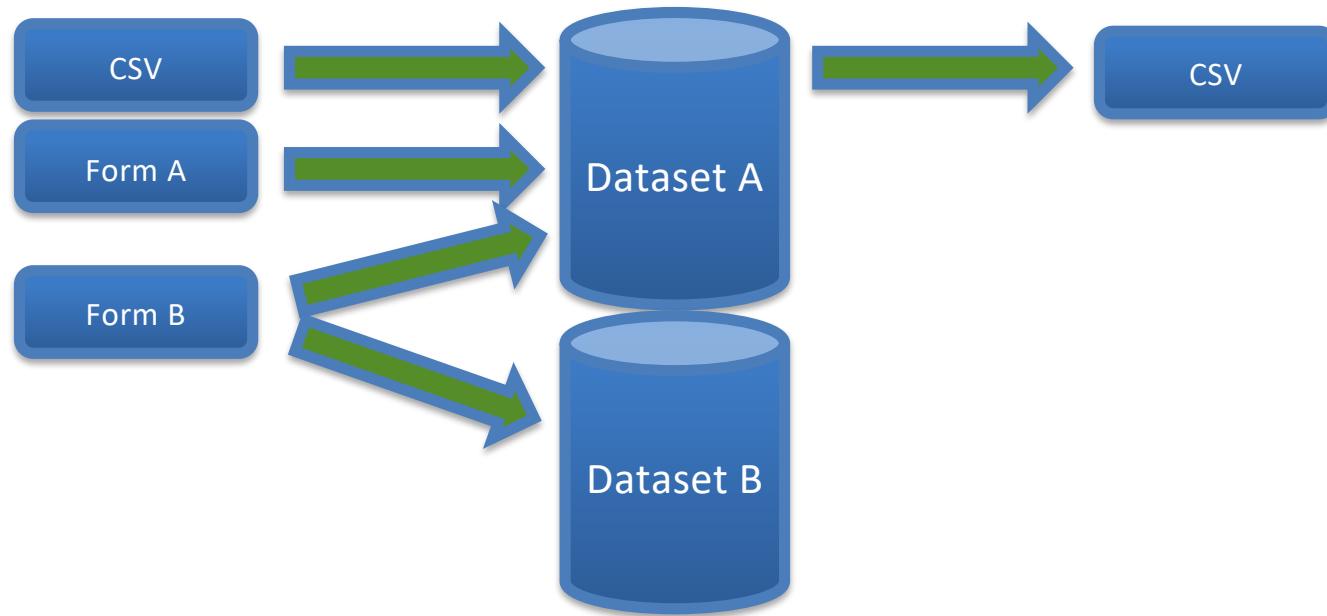
Datasets and preloading

You can get data into a dataset in several ways:



Datasets and preloading

You can then get data out of a dataset in several ways:



Download data as a CSV:

SurveyCTO

1. Design ▾ 2. Collect 3. Monitor 4. Export Configure

Your forms and datasets

Your forms and datasets ?

Tools Refresh Search Organize Help

Sample household listing form with sample data
Form ID: hh_listing_2, Deployed version: 1710070852
Published to datasets: Sample dataset (sample_dataset)

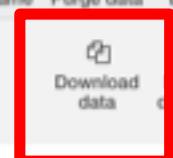
Sample dataset
Dataset ID: sample_dataset, Version: 0, Records: 0
Published from: Sample household listing form with sample data (hh_listing_2)

Click here to start new form, upload form definition, add group, or add server dataset

Preview Edit Upload Tools Encryption Download Duplicate Delete

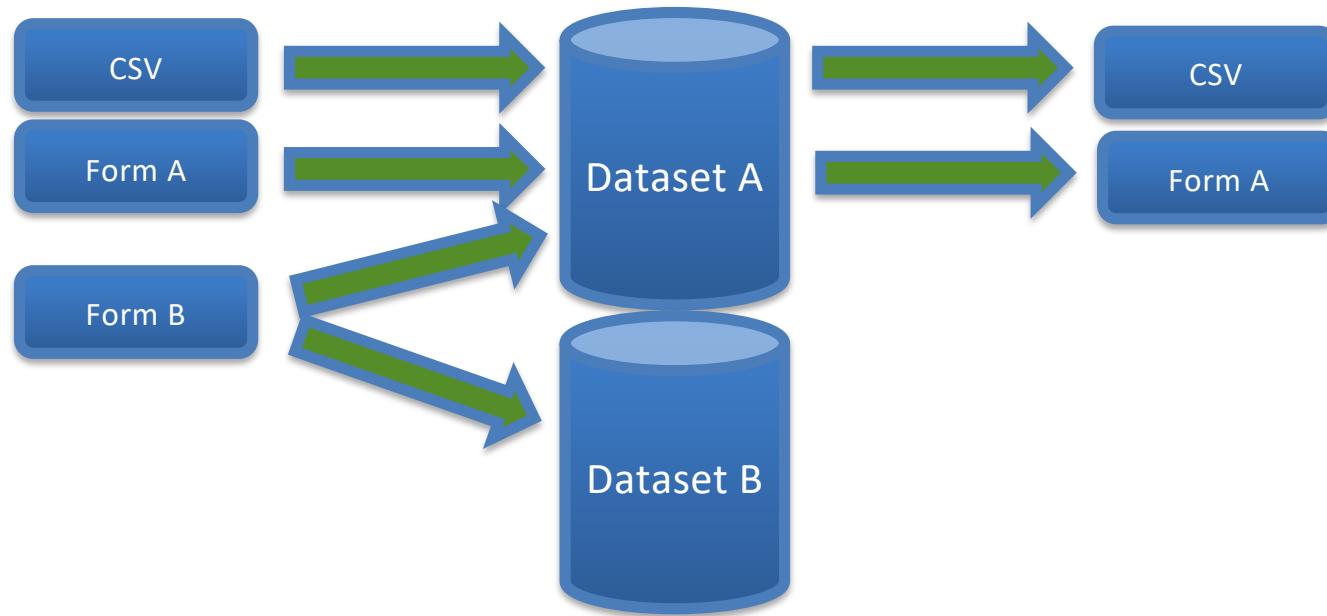
Attach Publish into Edit Upload Download Rename Purge data Delete

Download data Dataset definition



Datasets and preloading

You can then get data out of a dataset in several ways:



Preload data from a dataset into a form:

SurveyCTO

1. Design 2. Collect 3. Monitor 4. Export Configure faizan.diwan@gmail.com

Your forms and datasets

Tools Refresh Search Organize Help

Sample household listing form with sample data

Form ID: hh_listing_2, Deployed version: 1710070852
Published to datasets: Sample dataset (sample_dataset)

Preview Edit Upload Tools Encryption Download Duplicate Delete

Sample dataset

Dataset ID: sample_dataset, Version: 0, Records: 0

Attach Publish into Edit Upload Download Rename Purge data Delete

Forms to which this dataset is attached as pre-loaded .csv data:

hh_listing_2: Sample household listing form with sample data

Close

Published from: Sample household listing form with sample data (hh_listing_2)

Click here to start new form, upload form definition, add group, or add server dataset

Datasets and preloading

Preload data from a dataset into a form:
[Also requires pulldata() in the form]

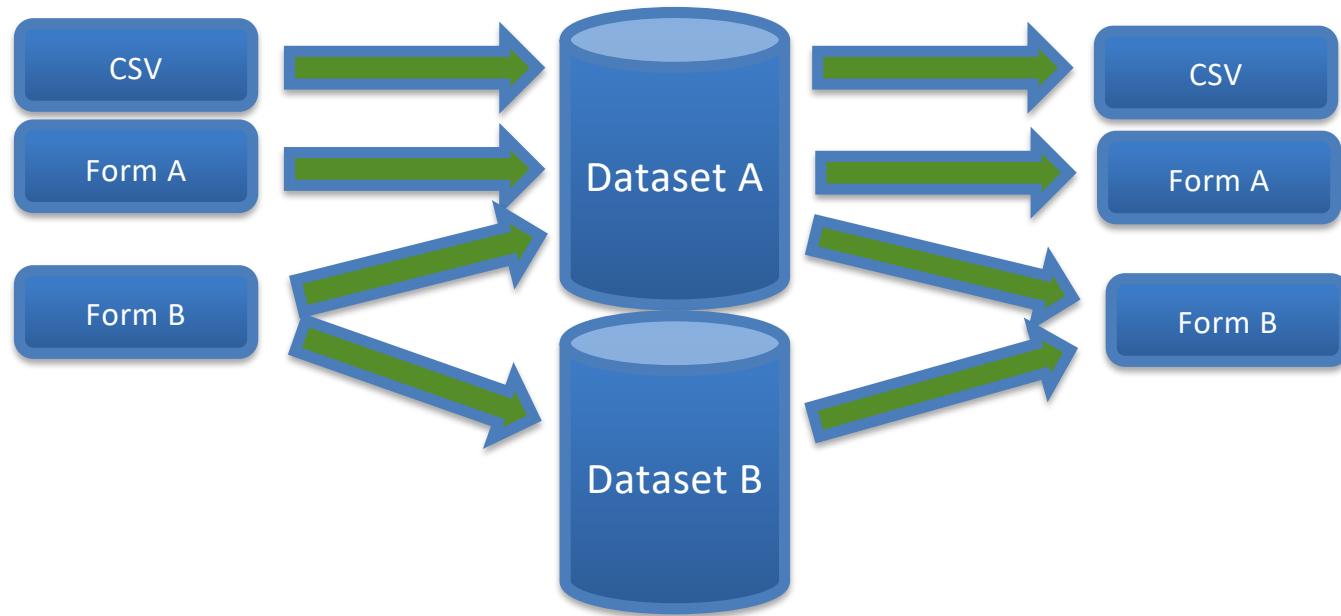
type	name	label	calculation
calculate	nplots		pulldata('datasetid', 'columntopull', 'uniqueidcolumn', \${uniqueid})
calculate	plot1description		pulldata('datasetid', 'columntopull2', 'uniqueidcolumn', \${uniqueid})

Find full sample preloading forms at:

Help > Designing forms > Advanced topics > Pre-loading data into a form

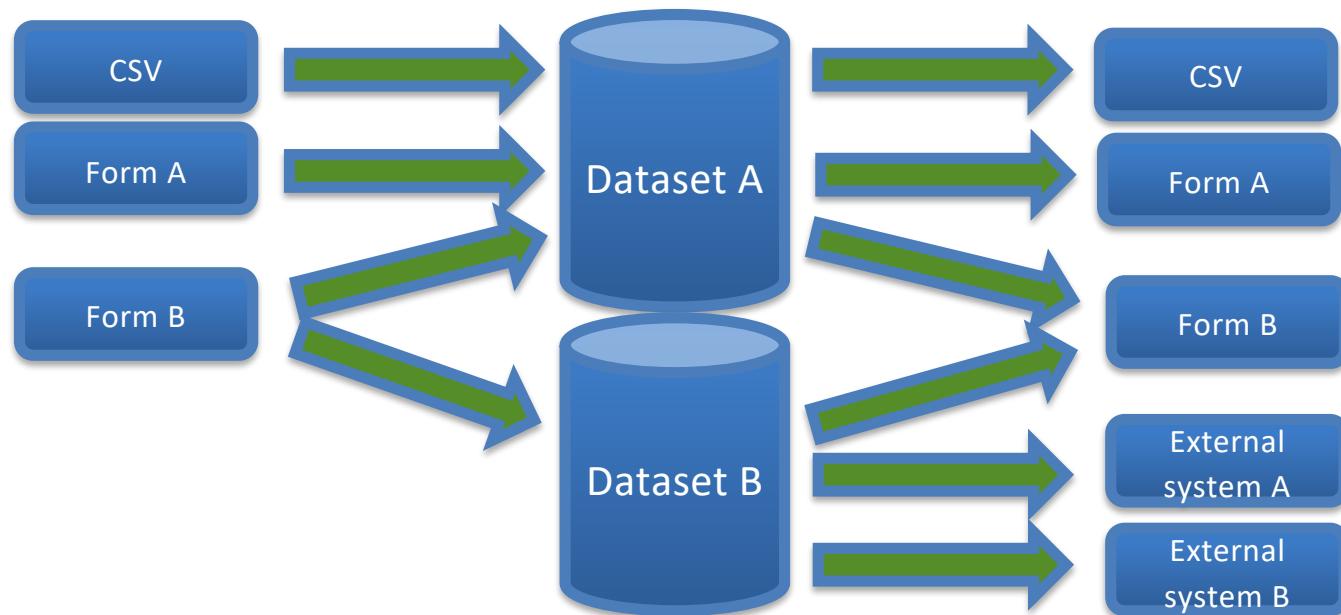
Datasets and preloading

You can then get data out of a dataset in several ways:



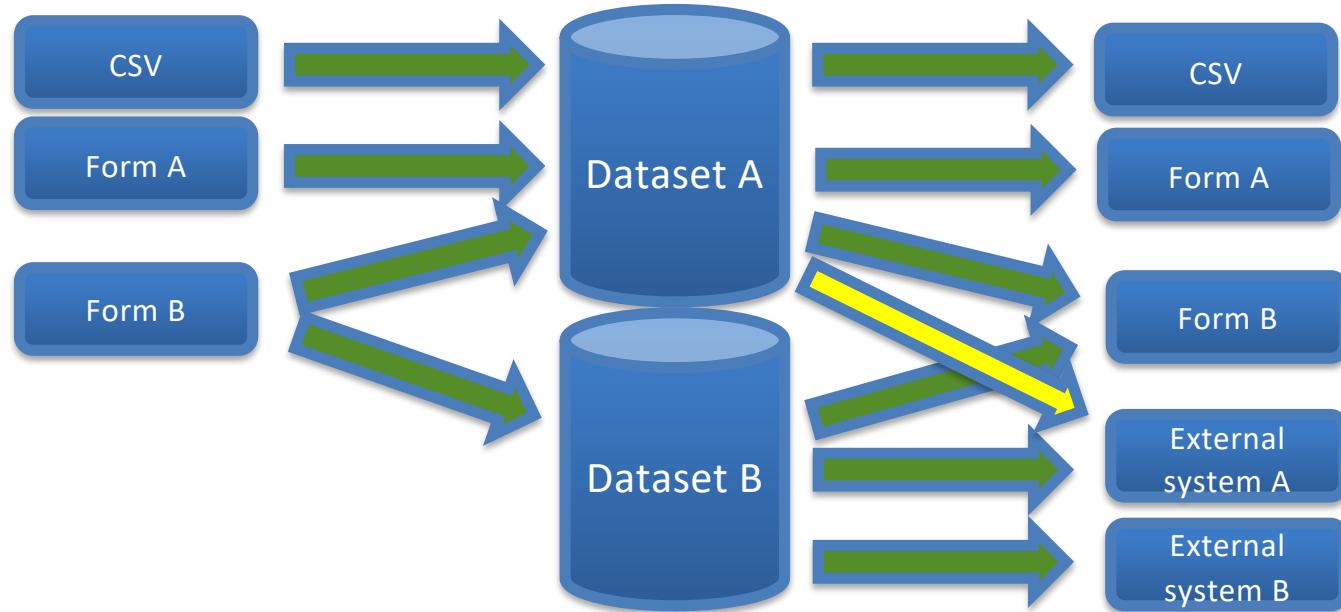
Datasets and preloading

You can then get data out of a dataset in several ways:



Datasets and preloading

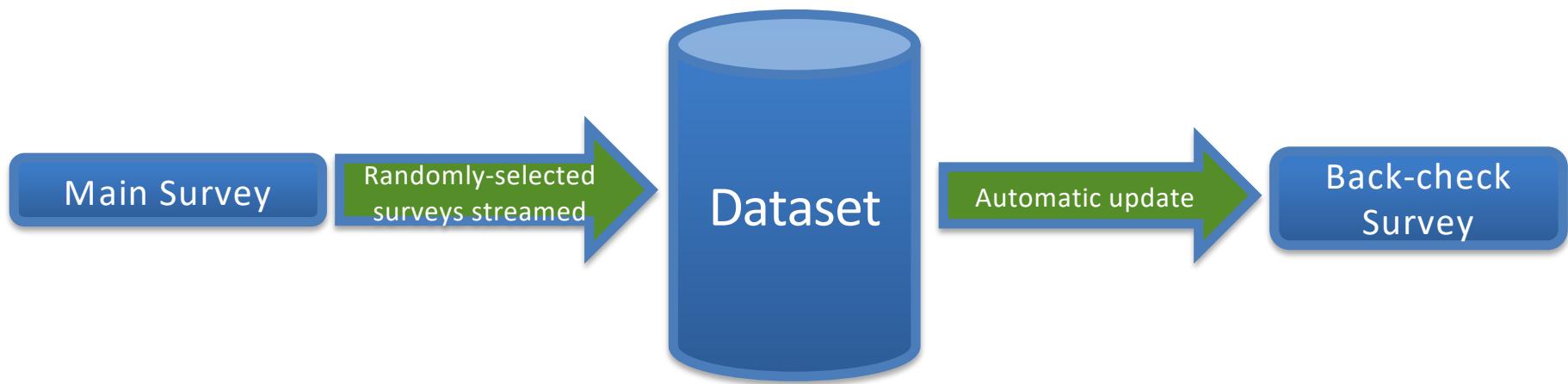
You can then get data out of a dataset in several ways:



You can have any combination of all of
these links at the same time!

Datasets and preloading

This allows you to create all sorts of automated workflows:

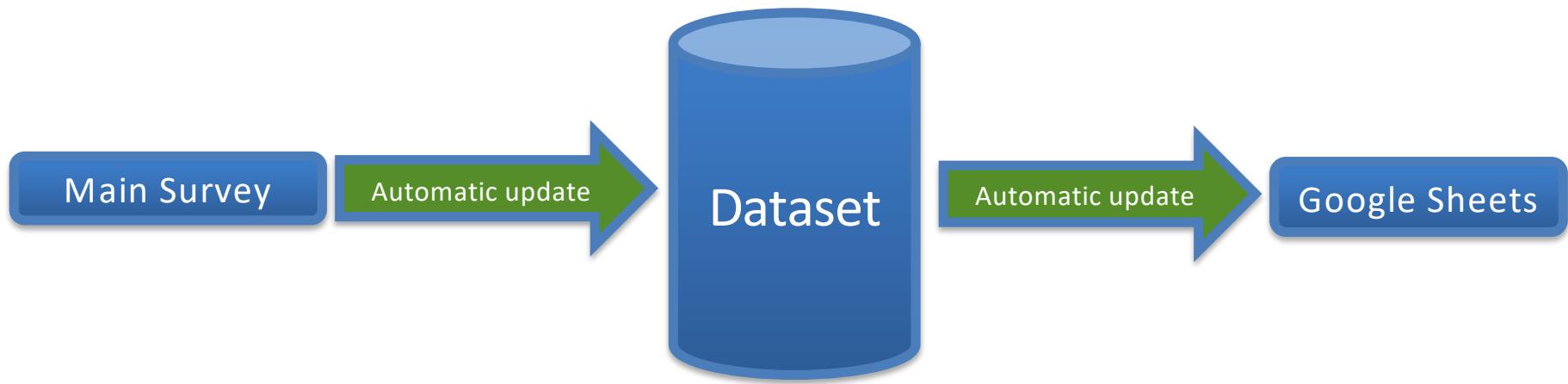


Automated back check workflow.

*Full sample: Help > Designing forms > Sample Forms
> Dataset back-checks*

Datasets and preloading

This allows you to create all sorts of automated workflows:

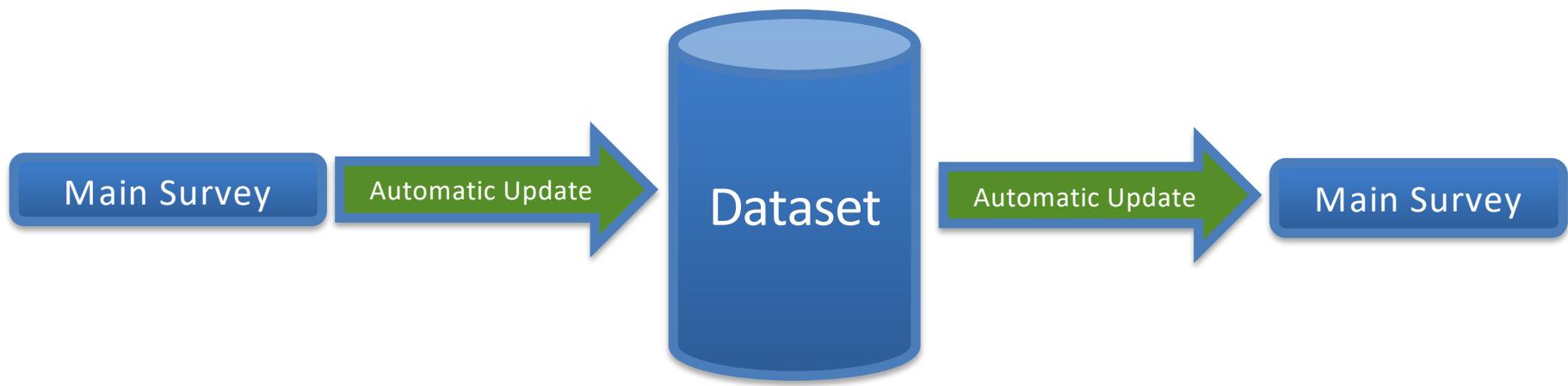


Monitor survey progress in real-time

Full sample: blog.surveycito.com > “Real-time project monitoring and visualization”

Datasets and preloading

This allows you to create all sorts of automated workflows:



Auto-updated panel survey

Similar sample: Help > Designing forms > Sample forms > Dataset basics

You can also run quality checks on data in a dataset:

SurveyCTO

1. Design 2. Collect 3. Monitor 4. Export Configure

Form submissions and dataset data

Sample dataset

Dataset ID: sample_dataset, Records: 0

Purge data

Automated quality checks

Refresh Search Help

?

Sample household listing form with sample data

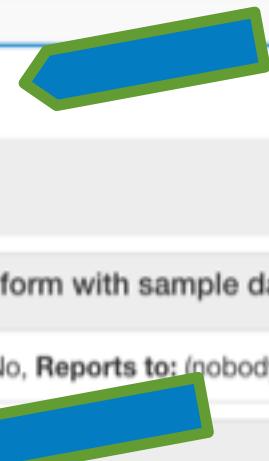
Options Checks Run now Report

Quality checks: 0, Nightly? No, Reports to: (nobody)

Sample dataset

Options Checks Run now Report

Quality checks: 0, Nightly? No, Reports to: (nobody)



Datasets and preloading

Things to keep in mind

- The one type of field that you cannot publish from a form into a dataset is an encrypted field. An encrypted field would be any field in an encrypted form that has not explicitly been marked as “publishable.”
- While rows can be automatically written to datasets, they cannot be automatically deleted/removed from a dataset. They can only be overwritten by merging on a unique ID.

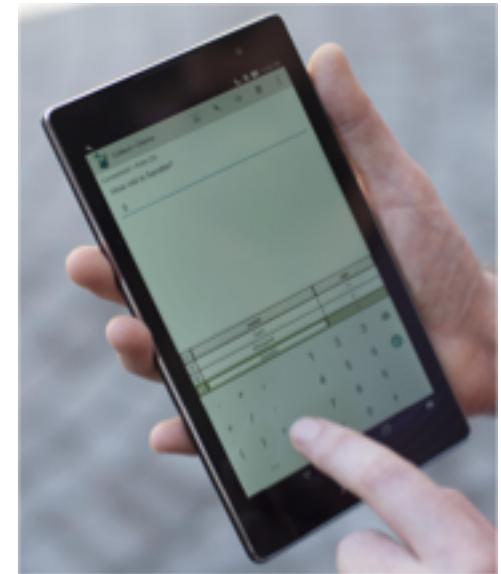
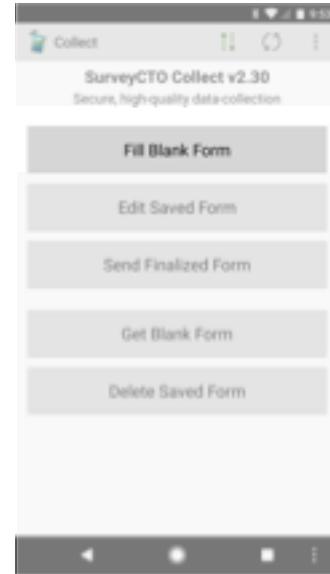


Case management



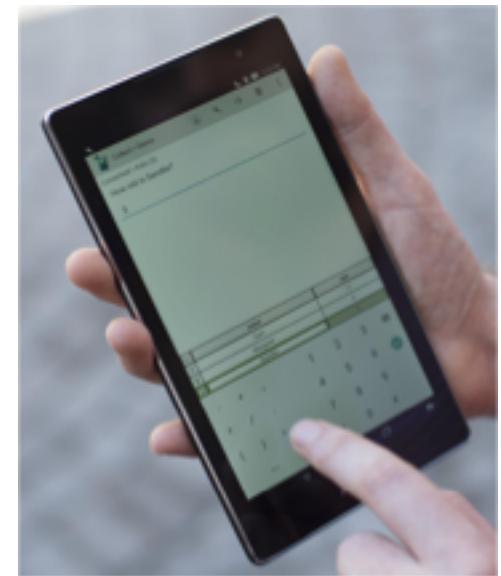
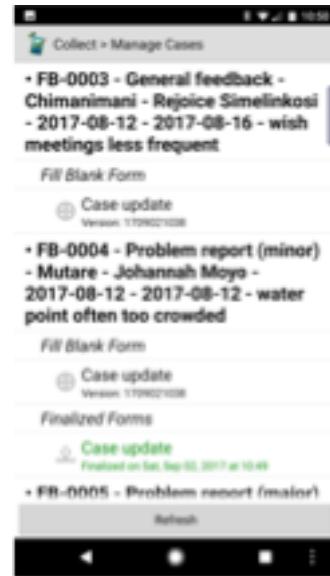
Case management vs. mobile data collection

Mobile data collection:



Case management vs. mobile data collection

Mobile case management:



Mobile case management - behind the scenes

Case list as dynamic dataset:

	A	B	C	D	E	F
1	id	label	formids	users	roles	sortby
		FB-0002 - General feedback - Buhera - Precious Rutendo - 2017-08-12 - 2017-08-16 - happy with community meeting				
2	FB-0002	FB-0003 - General feedback - Chimanimani - Rejoice Simelinkosi - 2017-08-12 - 2017-08-16 - wish meetings less frequent	case_update			
3	FB-0003	FB-0004 - Problem report (minor) - Mutare - Johannah Moyo - 2017-08-12 - 2017-08-12 - water point often too crowded	case_update			
4	FB-0004	FB-0004 - Problem report (minor) - Mutare - Johannah Moyo - 2017-08-12 - 2017-08-12 - water point often too crowded	case_update			



Recent improvement (late 2017)

Shifting from a simple label to describe cases:

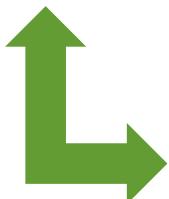
	A	B	C	D	E	F
1	id	label	formids	users	roles	sortby
		FB-0002 - General feedback - Buhera - Precious Rutendo - 2017-08-12 - 2017-08-16 - happy with community meeting				
2	FB-0002	FB-0003 - General feedback - Chimanimani - Rejoice Simelinkosi - 2017-08-12 - 2017-08-16 - wish meetings less frequent	case_update			
3	FB-0003	FB-0004 - Problem report (minor) - Mutare - Johannah Moyo - 2017-08-12 - 2017-08-12 - water point often too crowded	case_update			
4	FB-0004	FB-0005 - Problem report (major) -	case_update			



How can the technology evolve to help?

To a richer table of information:

	A	B	C	D	E	F	G	H	I	J	K	L
1	id	label	formids	users	roles	sortby	type	name	description	location	opened	updated
		FB-0002 - General feedback - Buhera - Precious Rutendo - 2017-08-12 - 2017-08-16 - happy with community meeting										
2	FB-0002		case_update				General feedback	Precious Rutendo	happy with community meeting	Buhera	2017-08-12	2017-08-16
		FB-0003 - General feedback - Chimanimani - Rejoice Simelinkosi - 2017-08-12 - 2017-08-16 - wish meetings less frequent										
3	FB-0003		case_update				General feedback	Rejoice Simelinkosi	wish meetings less frequent	Chimanimani	2017-08-12	2017-08-16



Collect > Manage Cases

ID	location	name	type	updated
FB-0002	Buhera	Precious Rutendo	General feedback	2017-08-16
FB-0003	Chimanimani	Rejoice Simelinkosi	General feedback	2017-08-16
FB-0004	Mutare	Johannah Moyo	Problem report (minor)	2017-08-16
FB-0005	Chipinge	Mercy Wadzanai	Problem report (major)	2017-08-16
FB-0006	Makoni	Tafara Zheve	General feedback	2017-08-16
FB-0007	Mutasa	Honest Tamboaga	Problem report (major)	2017-08-16
FB-0008	Nyanga	Tendai Muwinimi	Problem report (urgent)	2017-08-16
FB-0009	Bindura	Prosper Nyamuda	General feedback	2017-08-16

Refresh Filter

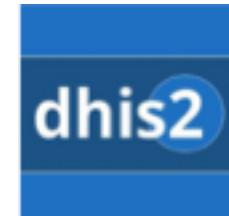


Integrations

Photo: Laura Elizabeth Pohl / EGPAF

We live in a wonderful world

Technology has proliferated and we have lots of great, specialized tools at our disposal:



We live in a fragmented world

Especially in the context of MERL, we often want the same data going through the pipeline of tools at our disposal.

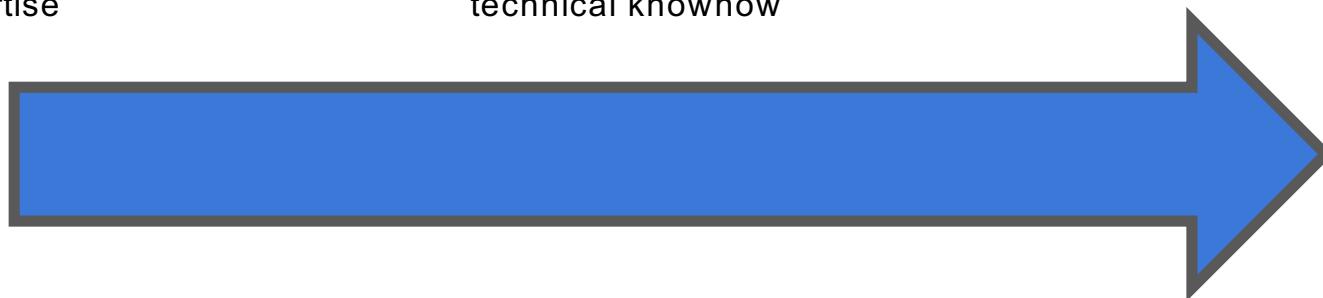


Integrating tools has become increasingly easy

Integrations very hard, expensive, require technical expertise

Integrations become easier, but still require considerable technical knowhow

Almost anyone can link tools together in a few clicks



Custom integrations

Integrations not possible

Better interoperability between tools

APIs

Easy-to-use tools that act as “bridges” - Sheets, Zapier, OpenFN, IFTTT

Webhooks

More APIs

This opens up lots of possibilities

Quickly link different best-in-class, scalable, affordable, off-the-shelf software to create your own “super-software” system, rather than trying to build a single expensive, custom, multi-purpose software platform.



Zapier

SurveyCTO



zapier*



Airtable

SMS

salesforce

Microsoft Azure

Power BI

ArcGIS®

Gmail

MySQL

Zapier

The screenshot shows the SurveyCTO web application interface. At the top, there is a navigation bar with tabs: 1. Design, 2. Collect, 3. Monitor, 4. Export (selected), and Advanced: publishing form and dataset data to the cloud. The user's email, crobert@surveycto.com, is also visible in the top right corner.

The main area displays a project titled "Project E". Within this project, a section titled "Sample Data - Integrations" is shown. A specific form, "farm_plot_listing", is selected. The publishing settings for this form are displayed, indicating it is currently publishing to Zapier.

A modal window titled "Form publishing options" is open, listing the publishing destinations:

- GOOGLE SPREADSHEETS (Inactive) with a "+ Add connection" button.
- GOOGLE FUSION TABLES (Inactive) with a "+ Add connection" button.
- ZAPIER - Account API key... (farm_plot_listing) with "Edit" and "Delete" buttons.

At the bottom of the modal, there is a "+ Add Zapier connection" button.

Zapier

Edit Zapier connection: "farm_plot_listing"

Add a new Zapier connection for this form, and choose which fields you want to share with Zapier (you can add multiple connections if you want to share different sets of fields). After you save, we'll give you an API key to use in configuring the Zapier side...

Please provide a name for this connection:

farm_plot_listing

Please choose the fields that you want to publish via this connection:



Select all

 **SubmissionDate:** Date/time submission received on server

 **starttime**

 **endtime**

 **username**

 **duration**

 **enumerator_team:** Welcome to the sample household listing form, mean...

 **enumerator:** ENUMERATOR: Which enumerator are you?

Google Sheets

SurveyCTO



Google
Sheets



infogr.am

C A R T O



ArcGIS®

Google Sheets

GOOGLE SPREADSHEETS
Inactive

+ Add connection

New Google Spreadsheets connection

Where do you want to publish this form to?

test spreadsheet

Field mapping:

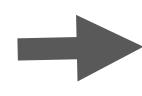
			+ Add	+ Add all	
<input type="checkbox"/>	SubmissionDate	→	<input type="checkbox"/>	SubmissionDate	
<input type="checkbox"/>	starttime	→	<input type="checkbox"/>	starttime	
<input type="checkbox"/>	endtime	→	<input type="checkbox"/>	endtime	
<input type="checkbox"/>	username	→	<input type="checkbox"/>	username	
<input type="checkbox"/>	duration	→	<input type="checkbox"/>	duration	

OpenFN

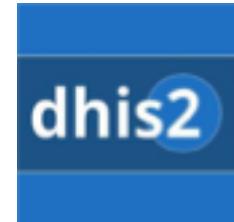
SurveyCTO



Open Fn



C A R T O



Caveats

- Different tools have different quirks! So some trial and error will always be involved. But you'll get there.
- Organize and structure data appropriately on the form-side as much as possible, to minimize work later.
- There are some trade-offs between real-time integrations and security of sensitive data.

Organize/structure data beforehand

Add calculate field

Save now Cancel

Shortcuts:

Required options Relevance

Add calculation - Step 2: Calculation parameters

Help Back Save

String concatenation

Concatenate this field:	first_name
And this text:	
And this field (optional):	last_name
And this field (optional):	

Cancel Back Save

Data security

Publishing data to Google Spreadsheet "HH Listing_TC111"

Field mapping:

Action	From Field	To Field	Action
+ Add	SubmissionDate	SubmissionDate	✖ Delete
+ Add all	starttime	starttime	✖ Delete
	endtime	endtime	✖ Delete
	username	username	✖ Delete
	duration	duration	✖ Delete
	enumerator_team	enumerator_team	✖ Delete
	enumerator	enumerator	✖ Delete
	province	province	✖ Delete

Try it yourself

1. Sign up for a free [Zapier](#) account
2. Review the integrations with Google Sheets
3. Pick an integration to test with your sample data*
4. Think about ways you could use integrations to support your own work
5. Check out the [free resources for using Google Sheets](#) from Zapier

* Note: This may take a little more time and patience than you would expect.

More ideas

1. Use [Google Sheets add-ons](#) when working with your data
2. Try [Google Data Studio \(beta\)](#) for reporting



Offline sync



Offline sync

SurveyCTO's normal workflow:



Offline sync

SurveyCTO's offline workflow:



Offline sync

A travel router like this will work well:



The HooToo Wireless Router by TripMate Nano (~\$20)

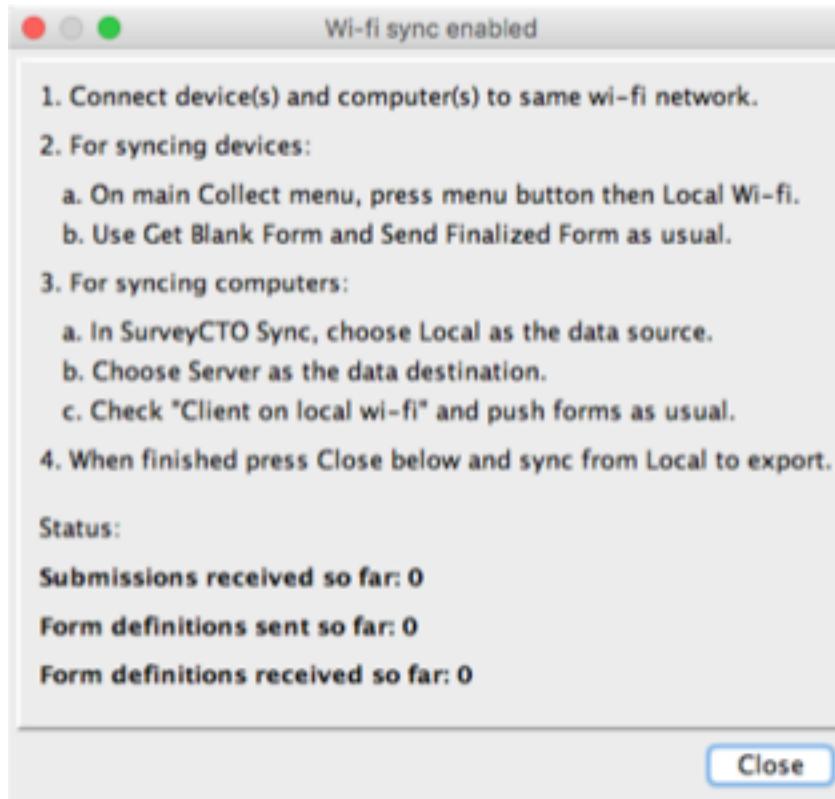
Offline sync

Use SurveyCTO Sync as a local server:



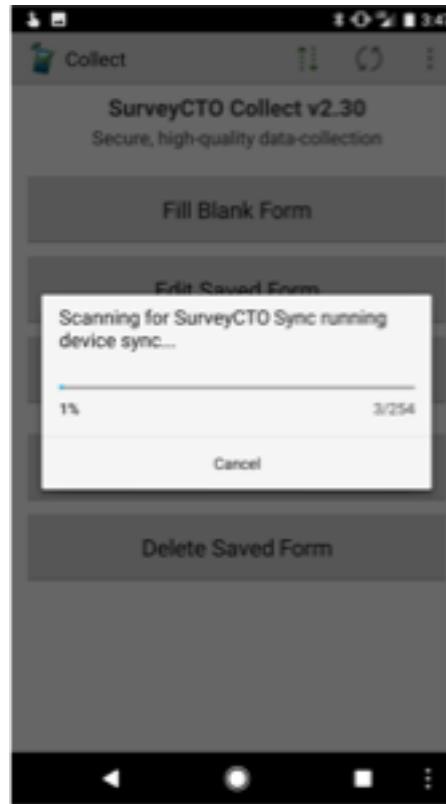
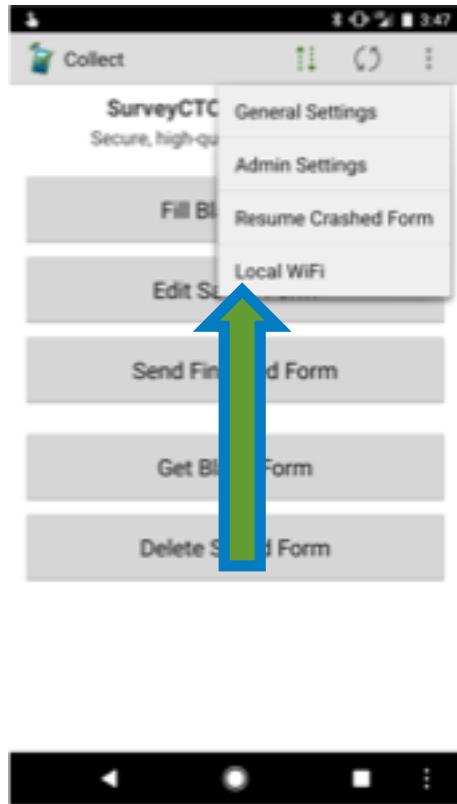
Offline sync

Use SurveyCTO Sync as a local server:



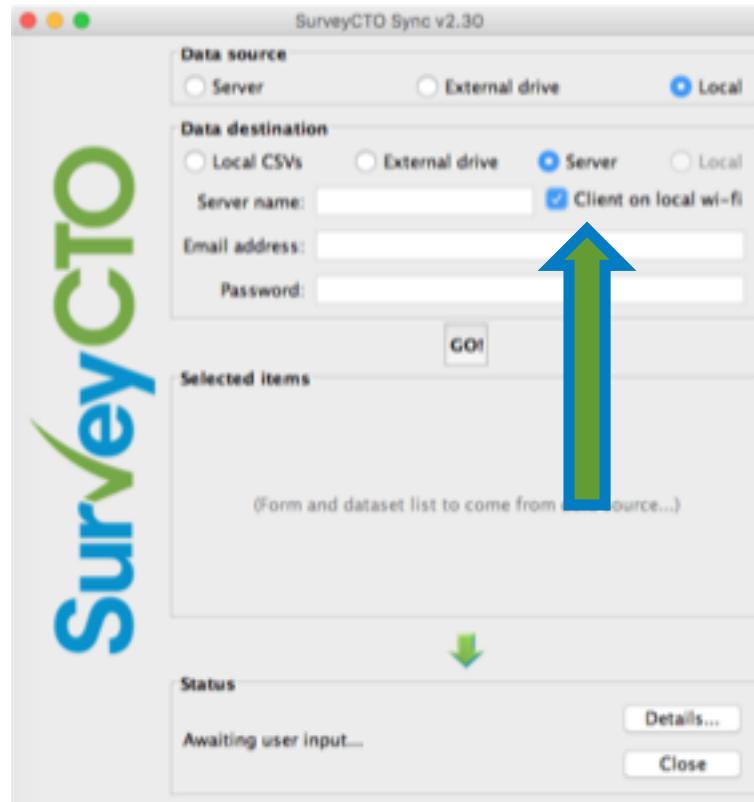
Offline sync

Connect, then Get Blank / Send Finalized:



Offline sync

Can even push to server, or between laptops:





Non-linear form navigation



Non-linear form navigation

- SurveyCTO Collect supports hyperlinking
- If hyperlinking is activated in Collect, any references to field or group names in labels will appear as hyperlinks on the device screen
- This allows users to create “navigation menus” via note fields



Advanced programming techniques



Example randomization form

- `once(random())` is often not enough: how to make random draws without duplicates?
- Sample: <https://tinyurl.com/advancedrandom>
 - Clever “auto-expanding” repeat group (`dynamic repeat_count`)
 - Strategic use of `join()`, `de-duplicate()`, and `selected-at()`



Next steps: Machine learning training

On the horizon: Machine learning roadmap

We believe that creative application of machine-learning technologies can meaningfully improve quality-control efforts – in the near-term.

1. But this requires humans to (accurately!) classify data quality.
 - a. Which requires humans to review data.
 - i. Which requires the tool to make review easy enough.
 - b. Which requires humans to make accurate judgments.
 - i. Which requires the tool to make accurate judgments easy enough.
2. It also requires meta-data that (a) correlates well with quality, and (b) is fundamentally safe (non-PII, non-revealing).
 - a. Ideas: summary statistics from light, sound, movement, timing.
3. And, of course, it requires clever use of machine-learning technologies.

Classify data now to train ML systems later

This submission has been flagged.

2¹ Please follow the steps below to complete the review process.

1. Classify submission

Correct
Action?
Has the user filled in all required fields?
Has the user answered all questions?

3. Approve or reject

Approve
 Reject



Feedback, discussion, and Q&A



Thank you!

Dr. Christopher Robert
Founder and CEO, Dobility, Inc.
crobert@surveycto.com

