

Data Review Guidance

Table of Contents

- [Goals](#)
- [Data Release Rules](#)
- [Recommended Steps](#)
 - [Step 1 - Review Your Project Profile](#)
 - [Step 2 - Review Your Study Reaches](#)
 - [Step 3 - Review Your Physical and Community Samples](#)
 - [Using the Physical or Community Sample List to Determine Review Status](#)
 - [Using the Sample Data Entry Screen to Determine Review Status](#)
 - [Using the Data Release Status Tab to Determine Review Status](#)
 - [Step 4 - Review Your Taxonomic Data Records](#)
 - [Look for these issues in taxonomic data](#)
 - [Process for checking taxonomic data](#)
 - [STEP 1: Select the sample that you want to review](#)
 - [STEP 2: Set their review code for the taxonomic records](#)
- [Tips and Tricks](#)
 - [Use the Data Retrievals](#)

Goals

The goals of data review are to ensure that:

1. **ALL** of the data stored in BioData are as complete and accurate as possible
2. **ALL** of the data that **should** be released to the public **are** released to the public in a timely manner
3. **ALL** of the data that should **not** be released to the public (internal-use only and proprietary data) will **not** be released to the public

Data Release Rules

BioData places every sample in one of three categories

- **Public**
 - Anyone can view or retrieve the data
- **Project Staff**
 - Only Project Staff, National, and Water Science Center (WSC) Data Stewards can view or retrieve the data
 - Project Staff can see only data for projects to which they are assigned
- **Data Steward**
 - Only National and WSC Data Stewards can view or retrieve the data
 - WSC Data Stewards can see ALL data for ALL projects assigned to their science center, but have only Public rights for data of other WSCs

The BioData GUI uses icons for these categories

-  = Public
-  = Project Staff
-  = Data Steward

From a practical, operational perspective:

- If all of your samples **should** be released to the public then when you are done reviewing your data all of the dots in the sample list (physical or community samples tabs) need to be **green**.

Sample data is categorized "Public" when all of the following are true...

- Sample data passes all of the automated BioData data validation tests
- The [Project Profile](#) has been reviewed and accepted
- The project [Study Reach](#) for the sample has been reviewed and accepted
- The Sample collection information has been reviewed and accepted (on sample data entry screen)
- The Sample is flagged as "unrestricted" using the Analysis Status Code in the sample header (on sample data entry screen)
- Taxon id/count records have been input for community (invert, fish, algae) samples
- All of the taxon id/count records have been reviewed and accepted (in Review Taxa Records tab)

Sample data is categorized "Project Staff" when...

- any of the above items have **not** been reviewed and accepted
- the sample is flagged as "internal" or "proprietary" using the Analysis Status Code in the sample data entry screen information header.

Sample data is categorized "Data Steward" when...

- BioData's automated data-checking routines detect invalid or erroneous data

Recommended Steps

Step 1 - Review Your Project Profile

None of the samples for your project will be available to anyone but the project staff until the Project Profile has been reviewed and accepted. So get this out of the way before you worry about reviewing individual samples.

- fix any project profile validation errors flagged by system
- enter any missing data
- correct any typo's
- **review the project abstract - this is published as a description of the project.** Make sure it's good.
- When you are done set the Review Status Code to "Reviewed and Accepted".

To open the project profile click on the "Edit" link next to the Project drop-down box at the top of the page. Note that the "Edit" link is only displayed to Project Staff.

Project: BoiseR WQ (View | **Edit** | Go to Project List) Project Page Tips

BoiseR WQ - Boise River Water Quality...
Idaho Water Science Center - Dorene E MacCoy

Data Release Status | Study Reaches | Physical Samples | **Community Samples** | Manage Lab Orders | Review Taxa Records

Community Sample List Create New Community Sample Getting Started - Community Samples

Filter list by:

Click "Edit" to open the Project Profile screen

Community Sample List:

Page 1 of 1 Page Size: 15

Release Category	Actions	Valid	Review Status	SIDNO...	Community	Sample Type	Protoc...	Collecti... Date	Collecti... Time	Proj. S...	Site Nu...	Site Na...	Study ...	Site Visit Sampl...	Anal Stat
------------------	---------	-------	---------------	----------	-----------	-------------	-----------	------------------	------------------	------------	------------	------------	-----------	---------------------	-----------

[Project List](#) | [Work on this Project](#) | **Edit Project**
[Expand All](#) | [Collapse All](#)

(*) required field

Project Header

Creation Date: 2011-05-11
 Project ID: USGS_BioData-00000044

Project Name: * Boise River Water Quality
 Project Label: * BoiseR WQ [View Project Owner History](#)
 Science Center: * Idaho Water Science Center [Project Staff:](#)
 Project Owner: * Dorene E MacCoy
 Default Project Account:
 Review Status Code: * R -- Reviewed and Accepted
 Release Category: Public Release Category for the PROJECT PROFILE - not! for all of the data for the project data
 Project Abstract:

Between 1990 and 2005, Ada and Canyon County population increased nearly 60% to 509,320, and is expected to continue growing over the next ten years*. The lower Boise basin contains the most industrialized and urbanized areas in Idaho.

Within the last century, the lower Boise River has been transformed from a meandering, braided, gravel-bed river that supported large runs of salmon to a channelized, regulated, urban river that provides flood control and irrigation to over 1,200 mi2 of land.

Several lower Boise River tributaries are interconnected by a complex irrigation system of canals, laterals, and drains. Flow in the lower Boise between Lucky Peak Dam and the mouth is primarily controlled by reservoir regulation; irrigation withdrawals and return flows; and seepage of shallow groundwater. Three upper basin reservoirs have a combined storage capacity of approximately 1 million acre-ft. These reservoirs are

Step 2 - Review Your Study Reaches

If you've assigned your samples to study reaches sample data will be released only if the study reach location information has been reviewed and accepted.

Select the Study Reaches tab

- fix any validation errors flagged by system
- enter any missing data
- correct any typo's
- set Review Status Code to "Reviewed and Accepted"

If the project profile and study reach is good to go there should be a green icon next to it in the study reach list. If it's not, click on it to see what the problem is.

[Project Profile](#) | **[Study Reaches](#)** | [Physical Samples](#) | [Community Samples](#) | [Manage Lab Orders](#) | [Review Taxa Records](#) | [Data Release Status](#)

[Study Reach List](#) | [Create New Study Reach](#)

Filter list by:

Site: Review Status: Valid:

Study Reach List:

Release Category	Actions	Valid	Review Status	Site	Reach Name	Length	No. of Samples	Delineation Date
<input checked="" type="radio"/>	View Edit	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	12056500 - NF SKOKOMISH R BL STAIRCASE RPDS NR ...	12056500-B-PUGTTrn	285 m	10	2011-09-09
<input checked="" type="radio"/>	View Edit	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	12128000 - THORNTON CREEK NEAR SEATTLE, WA	12128000-A-PUGTTrn	80 m	4	2011-07-21

Step 3 - Review Your Physical and Community Samples

For each physical or community sample you need to:

- fix any validation errors flagged by system
- enter any missing data
- correct any typo's
- set Review Status Code to "Reviewed and Accepted"
- Set Analysis Status Code to "U - unrestricted" for samples that will be released to public. Samples set to "I - Internal" and "P - Proprietary" will be released only to project staff.

You can find the review status of your samples in either of two places:

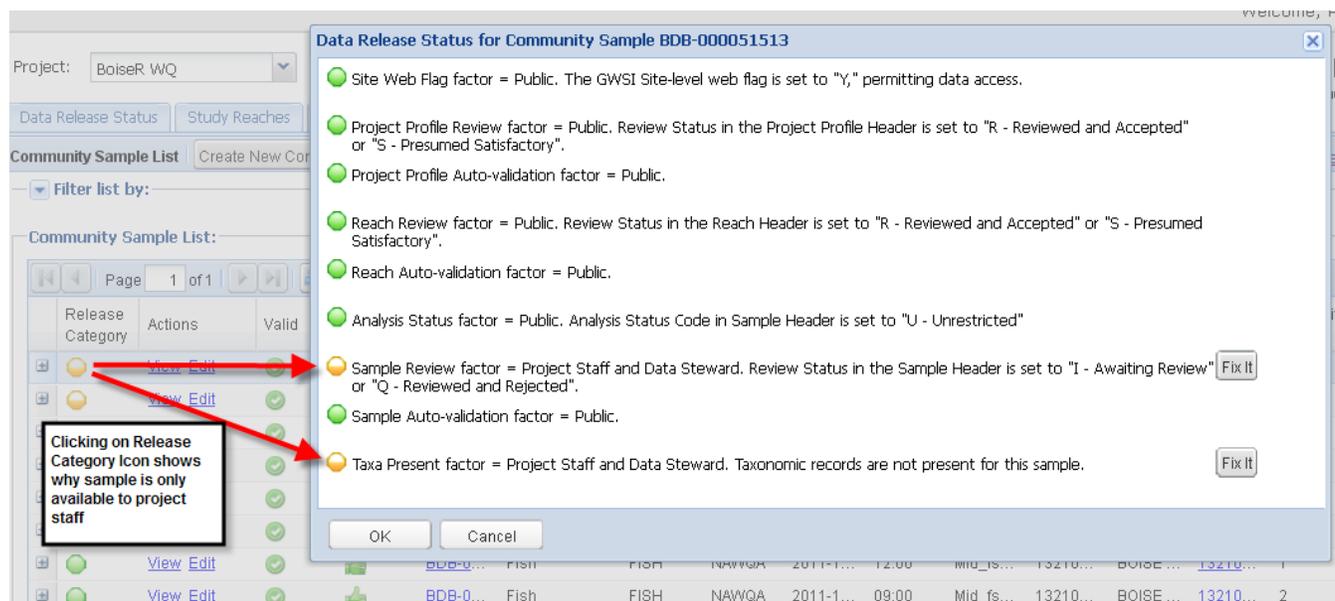
1. In the sample lists under the Physical or Community Samples tabs

2. In the Data Release Status tab
 - a. **NOTE:** option 1 may be your only choice for larger projects as this feature may be very slow or not work at all with large projects
3. By retrieving the sample inventory from the BioData Retrieval website.
 - a. **NOTE:** you need to log in to BioData and then go to the internal retrieval page to get non-public samples listed in the sample inventory

Using the Physical or Community Sample List to Determine Review Status

Select the Physical Samples or Community Samples tab. This displays a list of samples.

- The Review Status column indicates whether the sample has been reviewed and accepted.
- The Release Category column indicates whether the sample is being released to the public, to project staff, or only to data stewards. Click on the icon to display a popup window showing which factors are causing a sample to be withheld from the public.



Using the Sample Data Entry Screen to Determine Review Status

The sample data entry screens include fields for setting whether the sample *can* be released to the public (Analysis Status Code), and whether the sample information has been reviewed and accepted (Review Status Code).

The sample data entry screen also indicates the Release Category of the sample (Public, Project Staff, Data Steward). Clicking on the icon will display a popup that shows why a sample is not being released to the public (where you can click on the "Fix It" button to jump to the screen where the problem is).

[Data Release Status](#) | [Study Reaches](#) | [Physical Samples](#) | **[Community Samples](#)** | [Manage Lab Orders](#) | [Review Taxa Records](#)

[Community Sample List](#) | **Edit Community Sample** [Gettin](#)

SIDNO: BDB-000051513 **Site Visit Sample Number:** 1
Sample Type: NAWQA Invertebrate Targeted Habitat **Collection Date:** * 10/20/2011
Site: * 13203760 - BOISE RIVER AT ECKERT RD NR BOISE ID

Study Reach: 13203760-Eckert-BoiseR WQ **Start Time:** * 10:00
Project-assigned Sample Label: Eckert_invert2011 **Time Datum:** * MDT - Mountain Daylight Time
Analysis Status Code: U -- Unrestricted
Comments: Filamentous algae abundant

Sample Information Review Status
Review Status: I -- Awaiting Review
Release Category: Project Staff and Data Steward

Add From Staff | Add

Crew Leader	Name
<input checked="" type="checkbox"/>	Dorene MacCoy
<input type="checkbox"/>	Marshall Williams
<input type="checkbox"/>	Alex Etheridge
<input type="checkbox"/>	Rhonda Weakland

U - unrestricted means sample can be released to the public.

Click on icon to see reasons sample is not being released to public

Project: BoiseR WQ (View | Edit | Go)

[Data Release Status](#) | [Study Reaches](#) | [Physical Samples](#)

[Community Sample List](#) | **Edit Community Sample**

SIDNO: BDB-000051513
Sample Type: NAWQA Invertebrate Targeted Habitat
Site: * 13203760 - BOISE RIVER AT

Study Reach: 13203760-Eckert-BoiseR WQ
Project-assigned Sample Label: Eckert_invert2011
Analysis Status Code: U -- Unrestricted
Comments: Filamentous algae abundant

Sample Information Review Status
Review Status: I -- Awaiting Review
Release Category: Project Staff and Data Steward

Data Release Status for FieldSample

- Site Web Flag factor = Public. The GWSI Site-level web flag is set to "Y," permitting data access.
- Project Profile Review factor = Public. Review Status in the Project Profile Header is set to "R - Reviewed and Accepted" or "S - Presumed Satisfactory".
- Project Profile Auto-validation factor = Public.
- Reach Review factor = Public. Review Status in the Reach Header is set to "R - Reviewed and Accepted" or "S - Presumed Satisfactory".
- Reach Auto-validation factor = Public.
- Analysis Status factor = Public. Analysis Status Code in Sample Header is set to "U - Unrestricted"
- Sample Review factor = Project Staff and Data Steward. Review Status in the Sample Header is set to "I - Awaiting Review" or "Q - Reviewed and Rejected". Fix It
- Sample Auto-validation factor = Public.
- Taxa Present factor = Project Staff and Data Steward. Taxonomic records are not present for this sample. Fix It

Click on icon to open popup of release factors for the sample

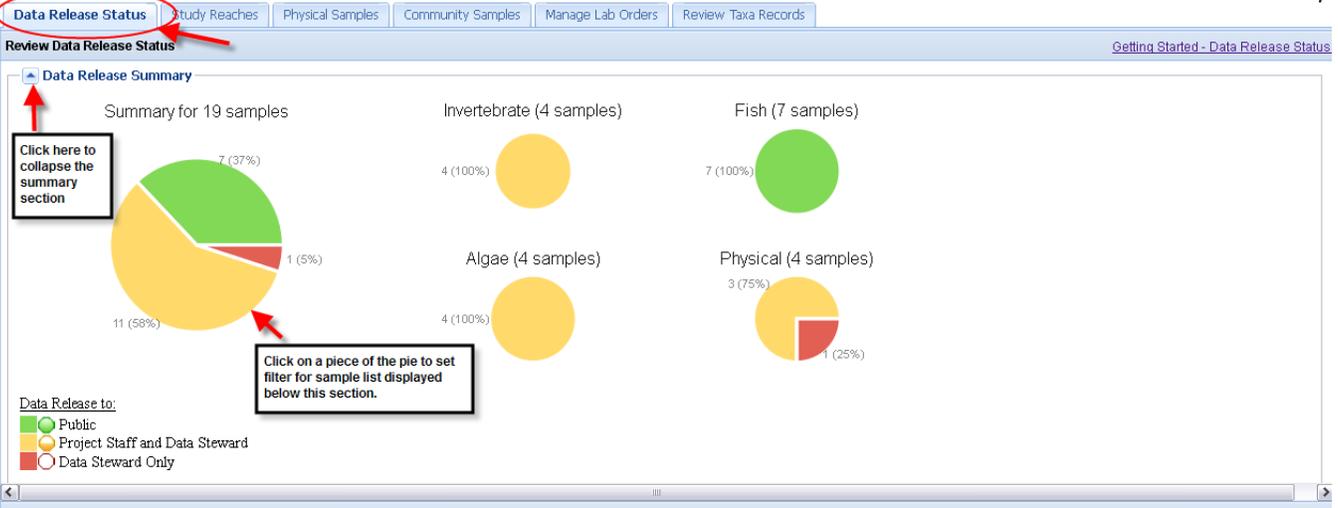
Click on the Fix It button to jump to screen where the problem is.

Using the Data Release Status Tab to Determine Review Status

Select the Data Release Status tab to determine which samples need work.

- NOTE:** this feature may be very slow or not work at all on larger projects

The summary section displays pie charts that summarize the overall release status of project samples. **i** Clicking on pie titles or slices will filter the sample list below.



The sample list in this tab indicates which samples will be released to the public, and what factors are preventing any particular sample from being released. It can be filtered to display a subset of all of the samples entered.

Filter list by:

Site: [] Sample Type: [] Data Category: []
Release Category: [] Collection Year: []

Sample List

Page: 1 of 2 | Page Size: 15 | Release Factor Icon Legend

Release Category	SIDNO	Collection Date	Data Category	Sample Type	Site W... Flag	Project Profile Review	Project Profile Valid	Reach Review	Reach Valid	Analysis Status	Sample Review	Sample Valid	Taxa Present	Taxa Review	Taxa Valid
Project Staff	BDB-0000515...			IRTH	●	●	●	●	●	●	●	●	●	-	-
Project Staff	BDB-0000515...	13203760	2011-10-20	Algae	ARTH	●	●	●	●	●	●	●	●	-	-
Public	BDB-0000515...	13203760	2011-10-20	Fish	FISH	●	●	●	●	●	●	●	●	●	●
Public	BDB-0000515...	13203760	2011-10-24	Fish	FISH	●	●	●	●	●	●	●	●	●	●
Public	BDB-0000515...	13210050	2011-10-25	Invertebrate	IRTH	●	●	●	●	●	●	●	●	-	-

Callouts: "Click on link for full description of columns and meaning of colored dots", "Because sample info has not been reviewed and accepted AND taxonomic ID/Count data have not been input yet.", "Sample is only visible to project staff", "Click on the SIDNO to open the sample for editing", "Sample is released to public", "Taxonomic ID/Count data HAVE been input AND have been reviewed and accepted"

Step 4 - Review Your Taxonomic Data Records

Look for these issues in taxonomic data

1) **For invertebrates and diatoms, confirm that minimum target counts were achieved (e.g. 300 for invertebrates or 600 for diatoms).** Tools in BioData help identify minimum target counts. If targets were not met, confirm that lab notes describe condition of the sample. For example, the entire invertebrate sample might have been sorted without reaching 300 organisms, or the 8 hour time limit specified in OFR 00-212 might have been reached. In another example, diatoms on slide might have been obscured by fine silts and clays. This information can be used by USGS biologist to confirm that sampling is successful, and if not, how to correct for future field collections.

2) **For invertebrates, confirm that organisms were identified to target levels** (e.g. OFR 00-212). Tools in BioData help identify minimum target counts. If targets were not met, confirm that lab notes document the shortfall. For example, a significant number of immature organisms will prevent identification to target level. This information can be used to consider shifting the sampling period to obtain mature individuals in future field collections.

3) **Determine if there are obvious differences between current and previous taxonomic data from the site.** For example, changes in the taxonomic level to which organisms are identified, appearance of new taxa, absence of formerly common taxa, or significant changes in relative abundance of taxonomic groups should alert the analyst to look further at the data, ask the lab about analyses, or contact one of the BioData taxonomic stewards for assistance.

4) **Consider comparing quantitative (e.g. IRTH) and qualitative (e.g. IQMH) results, if applicable.** Note obvious differences and determine if there are differences in collection procedures or personnel. This information can be used by USGS biologist to flag data, ask questions, alter procedures, or contact one of the BioData taxonomic stewards for assistance.

Process for checking taxonomic data

- 1) Select the sample that you want to review
- 2) Set the review code for the taxonomic records

STEP 1: Select the sample that you want to review

1. Apply filters to the display to show the samples that you want to select from.
 - a. Only one community will be displayed at a time
2. Consult the [Taxonomic Data Review Icons](#) page to understand what the icons mean
3. Click on the [Display](#) link to open that set of records

Idaho Water Science Center - Dorene E MacCoy

Data Release Status | Study Reaches | Physical Samples | Community Samples | Manage Lab Orders | **Review Taxa Records**

Find Taxonomic Result Set Click for help → [Getting Started - Taxonomic Result Review](#)

Filter list by:

Fish Invertebrate Algae Results:

Site: Reviewed: [Clear all filters](#)

Page 1 of 1 Page Size: 15 Displaying Found 1 - 7 c

Actions	Sidno	Review Complet...	Sample Type	Collection Date...	Site Number	Site Name
Display	BDB-000051515		FISH	2011-10-20	13203760	BOISE RIVER AT ECKERT RD NR BOISE ID
Display	BDB-000051518		FISH	2011-10-24	13203760	BOISE RIVER AT ECKERT RD NR BOISE ID
Display	BDB-000051523		FISH	2011-10-25	13210050	BOISE RIVER NR MIDDLETON ID
Display	BDB-000051524		FISH	2011-10-25	13210050	BOISE RIVER NR MIDDLETON ID
Display	BDB-000051528		FISH	2011-10-26	13213030	BOISE RIVER AT MOUTH NR PARMA ID
Display	BDB-000051529		FISH	2011-10-26	13213030	BOISE RIVER AT MOUTH NR PARMA ID

Filled in icon means all records have been reviewed

Click to display taxa count records

Click to open sample info data screen

STEP 2: Set their review code for the taxonomic records

- You can use the column sort and resizing features to examine your data and look for outliers
- The left hand summary panel can be hidden
- You need to set the review code for **every** record
- **Beware-** if your data set is large it may be spread across several pages (note "Page Size" and "Page x of y" below)
 - the 'select all' check box may not select all the records
 - you might want to consider using the "Set ALL Results to:" option at the bottom of the page.
- Return to the sample/lab order list and make sure the [Taxonomic Data Review Icon](#) is filled in appropriately
 - If it's not, you missed some records or made some errors.

Review Status definitions

- **Awaiting Review:** Default state given every result record when it is created.
- **Reviewed and Accepted:** Reviewer has reviewed and accepted the result.
- **Reviewed and Rejected:** Reviewer has rejected the record. One rejection in a result set means that none of the results in the result set will be published to BioData Retrieval.
- **Presumed Satisfactory:** Result is neither accepted or rejected, but will not inhibit the transfer of the result set to BioData Retrieval.

Fish Condition Factor - sort to identify typos

The Fish Condition Factor is provided to identify typos only, and is not intended for weight-length relationship analysis. Sort by the Fish Condition Factor to identify typos in fish total length and weight data entry. Pay special attention to the highest values (especially those over 10) and lowest values (those less than 0.6).

BioData uses the following calculation for condition factor:

The coefficient of condition has usually been represented and weighed in the metric system. The formula most

$$K = \frac{100,00}{L^3}$$

where: W = the weight of the fish in grams;

L = the standard length of the fish in millimeters

Source: Williams, J.E., 2000, Chapter 13, The Coefficient of Condition of Fish, in Schneider, James C., ed., Manual of fisheries survey methods II: with periodic updates: Michigan Department of Natural Resources, Fisheries Special Report 25, Ann Arbor.

Project Profile Study Reaches Physical Samples Community Samples Manage Lab Orders **Review Taxa Records** Data Release Status

Find Different Taxonomic Result Set **Review Results for FISH sample BDB-000050461 on 2010-09-20 at 05451770 - Iowa River at County Highway E49 near Tama, IA** [Getting Started - Taxonomic Result Review](#)

Summary Page 1 of 1 Page Size 200 [Show All Results \(Simple View\)](#) [Export List](#) Displaying Found 1 - 121 of 121

Taxonomic Result Summary
 121 Total records.
 Sum of abundance - 328
 17 distinct taxa

49 (15%) Cyprinidae specimens
 36 (11%) Catostomidae specimens
 2 (1%) Ictaluridae specimens
 0 (0%) Salmonidae specimens
 0 (0%) Cottidae specimens
 1 (0%) Centrarchidae specimens
 1 (0%) Percidae specimens

<input checked="" type="checkbox"/>	Revi... Stat...	Meth...	Pa...	Li...	Co...	Scie...	Publi...	Ab...	Total Leng...	Stan... Leng...	Weig...	Co... Fa...	DELT...	DE...	ER...	LE	TU...	ID at	Taxo...	Vouc...
<input checked="" type="checkbox"/>	R	Minn...	1	1	joh...	Etke...	100...	1	56		2.0	1.1	0	No	No	No	No	Field		
<input checked="" type="checkbox"/>	R	Minn...	1	2	nor...	Hyp...	100...	1	50		1.6	1.3	0	No	No	No	No	Field		
<input checked="" type="checkbox"/>	R	Minn...	1	3	sa...	Notr...	100...	1	52		1.0	0.7	0	No	No	No	No	Field		
<input checked="" type="checkbox"/>	R	Minn...	1	4	sa...	Notr...	100...	1	58		2.0	1	0	No	No	No	No	Field		
<input checked="" type="checkbox"/>	R	Minn...	1	5	sa...	Notr...	100...	1	53		1.0	0.7	0	No	No	No	No	Field		
<input checked="" type="checkbox"/>	R	Minn...	1	6	sa...	Notr...	100...	1	50		1.0	0.8	0	No	No	No	No	Field		
<input checked="" type="checkbox"/>	R	Minn...	1	7	sa...	Notr...	100...	1	48		1.0	0.9	0	No	No	No	No	Field		
<input checked="" type="checkbox"/>	R	Minn...	1	8	sa...	Notr...	100...	1	67		2.0	0.7	0	No	No	No	No	Field		
<input checked="" type="checkbox"/>	R	Minn...	1	9	sa...	Notr...	100...	1	55		1.0	0.6	0	No	No	No	No	Field		
<input checked="" type="checkbox"/>	R	Minn...	1	10	sa...	Notr...	100...	1	51		1.0	0.8	0	No	No	No	No	Field		
<input checked="" type="checkbox"/>	R	Minn...	1	11	bra...	Hyb...	100...	1	42		1.0	1.3	0	No	No	No	No	Field		
<input checked="" type="checkbox"/>	R	Boat ...	1	12	co...	Cyp...	100...	1	700		7500.0	2.2	1	No	No	Yes	No	Field		
<input checked="" type="checkbox"/>	R	Boat ...	1	13	co...	Cyp...	100...	1	562		4000.0	2.3	0	No	No	No	No	Field		
<input checked="" type="checkbox"/>	R	Boat ...	1	14	co...	Cyp...	100...	1	465		2500.0	2.5	0	No	No	No	No	Field		
<input checked="" type="checkbox"/>	R	Boat ...	1	15	co...	Cyp...	100...	1	660		6000.0	2.1	1	No	No	Yes	No	Field		
<input checked="" type="checkbox"/>	R	Boat ...	1	16	co...	Cyp...	100...	1	585		4000.0	2	0	No	No	No	No	Field		
<input checked="" type="checkbox"/>	R	Boat ...	1	17	co...	Cyp...	100...	1	455		2500.0	2.7	1	No	Yes	No	No	Field		
<input checked="" type="checkbox"/>	R	Boat ...	1	18	co...	Cyp...	100...	1	420		2000.0	2.7	0	No	No	No	No	Field		
<input checked="" type="checkbox"/>	R	Boat ...	1	19	co...	Cyp...	100...	1	450		2500.0	2.7	0	No	No	No	No	Field		

Set checked results to: - Awaiting Review R - Reviewed and Accepted Q - Reviewed and Rejected S - Presumed Satisfactory

Set ALL results to: I - Awaiting Review R - Reviewed and Accepted Q - Reviewed and Rejected S - Presumed Satisfactory

Tips and Tricks

Use the Data Retrievals