

Introduction to Accessing DIVER Data Through Data Services

<https://www.diver.orr.noaa.gov/diver-data-services>

DIVER's ERDDAP data server provides users with a simple, consistent way to download subsets of scientific datasets in common file formats. Because ERDDAP users may view all exposed DIVER database fields, it is best if the user has prior knowledge of how the data are structured, including valid values of interest. In this guide we explain how to navigate the list of datasets, discuss data structure and valid values, and provide example queries with links to get started.

Regional Datasets

<https://www.diver.orr.noaa.gov/web/guest/data-service-regional-datasets>

Datasets are grouped by region, with a dataset for each available data category. Buttons at the top of the list of datasets allow users to navigate to a specific region. In addition to environmental data, each region has four supporting datasets:

- **Chemical Codes:** information about all chemicals available in DIVER.
- **Qualifier Codes:** definitions for all data qualifiers in a region.
- **Study Notes:** information about data source and available results.
- **Study References:** sources related to each study.

Navigating the Data Access Form

The screenshot shows the ERDDAP Data Access Form for the dataset 'Samples, Southeast'. The form includes a list of variables, two optional constraint columns, and a dropdown filter. Annotations explain the following features:

- Variables correspond to DIVER fields.** (Points to the variable list)
- Checked variables are included in export. Buttons enable checking and/or unchecking all variables.** (Points to the 'Check All' and 'Uncheck All' buttons)
- Operators:**
 - = < > <= >=
 - != (does not equal)
 - =~ (regular expression)
- Hover over question mark icons to access help information.** (Points to a question mark icon)
- Dropdown filters are available for certain fields. Menus *do not* update based on other constraints user has added.** (Points to a dropdown filter)

Available Export Formats

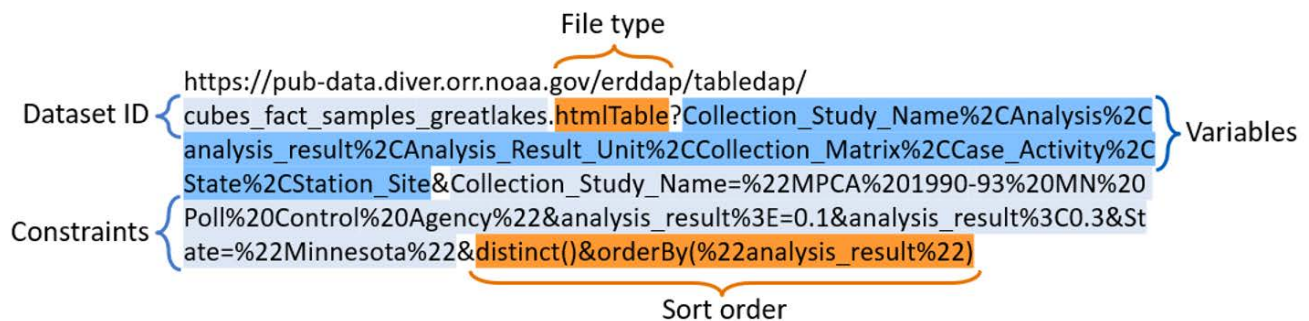
Users can choose to export data as any of roughly 40 [file types](#). A dropdown menu below the variables lists all available file types alongside a short description. Options include common file types (e.g., .csv, .json), metadata files (e.g., .iso19115, .fgdc), and extensions leading to other ERDDAP pages (e.g., .html, .htmlTable, .graph).

The screenshot shows a web interface for selecting an export format. At the top, there is a link for "(more info)". Below it is a dropdown menu currently set to ".htmlTable - View a UTF-8 .html web page with the data in a table. Times are ISO 8601 strings." Below the dropdown is a text input field labeled "Just generate the URL:" containing the URL "https://pub-datapre.diver.orr.noaa.gov/erddap/tabledap/fact_samples_greatlakes". There are links for "(Documentation / Bypass this form ?)". At the bottom is a "Submit" button with the text "(Please be patient. It may take a while to get the data.)".

URL Structure

Users may save queries by saving the unique URL (web address) generated in the data export process. URLs completely describe queries and may be created/edited directly without using the Data Access Form.

Tip: Spaces and special characters in URLs are converted to a [two-digit hexadecimal form](#). Each two-digit code is preceded by "%". In the URL below, **2C** represents a comma, **3E** represents ">", **22** represents quotation marks, **3C** represents "<", and **20** represents a space.



Query Examples

In the following examples, click the URL to see the ERDDAP web interface (note that the dataset will already be chosen). The instruction steps assume you start from the Data Access Form for a specific dataset. Modify to create your own queries.

Tip: It is possible to compose a query that combines field constraints to result in no data, or an error if valid values are not correctly included in the query. If queries are not valid the user will receive an error message. We recommend exploring data using the DIVER Explorer tool to understand the data structure and values as you compose your query.

Examples:

- 1) What are the valid values for Collection Method for samples in the Great Lakes?
- 2) What studies took place in Howards Bay between 2010 and 2018?
- 3) How many results are associated with each Station/Site for the studies in the "Great Lakes (Mussel Watch)" Case/Activity?
- 4) Where are sediment samples located in the Lower Duwamish River study group?
- 5) What PCB data are available for surface sediment in the Hudson River station group?

Example 1: What are the valid values for Collection Method for samples in the Great Lakes?

Step 1: From DIVER data services home page (<https://www.diver.orr.noaa.gov/diver-data-services>), navigate to "Regional Datasets"

Step 2: Under Great Lakes region, click "Data" link for "**Samples, Great Lakes**" dataset to navigate to Data Access Form

Step 3: Uncheck all variables

Step 4: Check **Collection_Method**

Step 5: Under Server-side Functions, check **distinct()** – This function deletes duplicate rows, allowing the user to access information at an overview level

Step 6: Set File type to **.htmlTable**

Step 7: Click **Submit**

Follow this URL to view/edit the Data Access Form for Example 1: [https://pub-data.diver.orr.noaa.gov/erddap/tabledap/cubes_fact_samples_greatlakes.html?Collection_Method&distinct\(\)](https://pub-data.diver.orr.noaa.gov/erddap/tabledap/cubes_fact_samples_greatlakes.html?Collection_Method&distinct())

Tip #1: Prior to running a query, make sure each field you want in the export is selected. Entering a constraint will not automatically include the field in the export; the field must be selected. Note, constraints will act on the data even if they are not included in the export.

Tip #2: Adding additional constraints, like Case/Activity, will usually make queries run faster by limiting how many data records are searched.

Example 2: What studies took place in Howards Bay between 2010 and 2018?

Step 1: From Regional Datasets page, click "Data" link for "Samples, Great Lakes" to access Data Access Form

Step 2: Uncheck all variables

Step 2: Check **Collection_Study_Name**

Step 3: Set **Collection_Study_Name** to =~ ".*Howard.*"

- Change operator to =~
- Type ".*Howard.*"

Step 4: Include a date range (2010 to 2018)

- Set **date_id** to =~"201[0-8].*"

Step 5: Select **distinct()**

Step 6: Choose a file type and click **Submit**

Follow this URL to view/edit the Data Access Form for Example 2: [https://pub-data.diver.orr.noaa.gov/erddap/tabledap/cubes_fact_samples_greatlakes.html?Collection_Study_Name&Collection_Study_Name=~%22.*Howard.*%22&date_id=~%22201%5B0-8%5D.*%22&distinct\(\)](https://pub-data.diver.orr.noaa.gov/erddap/tabledap/cubes_fact_samples_greatlakes.html?Collection_Study_Name&Collection_Study_Name=~%22.*Howard.*%22&date_id=~%22201%5B0-8%5D.*%22&distinct())

Tip #3: When using regular expressions, "." means "any character any number of times." To limit results to study names that begin with Howard, you would type "Howard.*"

Tip #4: The date format is yyyyymmdd, and the expression **date_id=~"201[0-8].*"** returns all records with "201" as the first three digits and a number 0-8 for the fourth digit, so all dates in 2010-2018 will be included.

Example 3: How many results are associated with each Station/Site for the studies in the "Great Lakes (Mussel Watch)" Case/Activity?

Step 1: Starting from Data Access Form for "Samples, Great Lakes", uncheck all variables

Step 2: Check **analysis_result**, **Collection_Matrix**, **Collection_Study_Name**, **Station_Site**

Step 3: Set **Case_Activity** to "Great Lakes (Mussel Watch)"

Step 4: Under Server-side Functions, select **orderByCount**

Step 5: Then select **Collection_Study_Name**, **Station_Site**, **Collection_Matrix** in the drop-down menus to the right

Step 6: Choose a file type and click **Submit**

Follow this URL to view/edit the Data Access Form for Example 3: [https://pub-data.diver.orr.noaa.gov/erddap/tabledap/cubes_fact_samples_greatlakes.html?Collection_Study_Name%2Canalysis_result%2CCollection_Matrix%2CStation_Site&Case_Activity=%22Great%20Lakes%20\(Mussel%20Watch\)%22&orderByCount\(%22Collection_Study_Name%2CStation_Site%2CCollection_Matrix%22\)](https://pub-data.diver.orr.noaa.gov/erddap/tabledap/cubes_fact_samples_greatlakes.html?Collection_Study_Name%2Canalysis_result%2CCollection_Matrix%2CStation_Site&Case_Activity=%22Great%20Lakes%20(Mussel%20Watch)%22&orderByCount(%22Collection_Study_Name%2CStation_Site%2CCollection_Matrix%22))

Tip #5: To order results by count when using orderByCount, add "&orderBy("Analysis_Result")" directly to the end of the URL.

Example 4: Where are sediment samples located in the Lower Duwamish River study group? (This example replicates a Station Locations guided query in DIVER Explorer)

Step 1: From Regional Datasets page, click "Data" link for "**Samples, Northwest**" to access Data Access Form.

Step 2: Uncheck all variables.

Step 3: Check Area_Description, Case_Activity, Collection_Study_Name, Data_Category, Data_Classification, date, DIVER_Dataset, File_Collection_ID, latitude, longitude, QM_Matrix, QM_Site_ID, QM_Station_ID, qm_study_id, Sharing_Status, Source_Type, Station_Group_List, Station_Site, Trip_ID, Workgroup

Step 4: In the drop-down menu for **Analysis_Category**, select "Contaminant Chemistry"

Step 5: Set **Analysis_Status** to ="Results Available"

- Change operator to =
- Type "Results Available"

Step 6: Set **Collection_Matrix** to "Sediment"

Step 7: Set **Station_Group_List** to =~".*Lower Duwamish River.*"

- Change operator to =~
- Type ". *Lower Duwamish River. *"

Step 8: Select **distinct()**

Step 9: Select **orderBy()**, then select **qm_study_id** and **QM_Station_ID**

Step 10: Choose a file type and click **Submit**

Tip #6: Quotation marks must be used around values for string (text) variables. Most DIVER variables are set as strings. The icon (🔗) next to each variable includes the variable type.

Tip #7: Unsure what values are available for Station Group? Load **Example 1** and replace Collection_Method with Station_Group_List.

Follow this URL to view/edit the Data Access Form for Example 4: [https://pub-data.diver.orr.noaa.gov/erddap/tabledap/cubes_fact_samples_northwest.html?QM_Site_ID%2Cqm_study_id%2CQM_Station_ID%2Clatitude%2Clongitude%2CCollection_Study_Name%2CQM_Matrix%2CSharing_Status%2CFile_Collection_ID%2CWorkgroup%2CSource_Type%2CCase_Activity%2CData_Classification%2CData_Category%2CDIVER_Dataset%2CTrip_ID%2Cdate%2CStation_Site%2CStation_Group_List%2CArea_Description&Analysis_Status=%22Results%20Available%22&Collection_Matrix=%22Sediment%22&Analysis_Category=%22Contaminant%20Chemistry%22&Station_Group_List=~%22.*Lower%20Duwamish%20River.*%22&distinct\(\)&orderBy\(%22qm_study_id%2CQM_Station_ID%22\)](https://pub-data.diver.orr.noaa.gov/erddap/tabledap/cubes_fact_samples_northwest.html?QM_Site_ID%2Cqm_study_id%2CQM_Station_ID%2Clatitude%2Clongitude%2CCollection_Study_Name%2CQM_Matrix%2CSharing_Status%2CFile_Collection_ID%2CWorkgroup%2CSource_Type%2CCase_Activity%2CData_Classification%2CData_Category%2CDIVER_Dataset%2CTrip_ID%2Cdate%2CStation_Site%2CStation_Group_List%2CArea_Description&Analysis_Status=%22Results%20Available%22&Collection_Matrix=%22Sediment%22&Analysis_Category=%22Contaminant%20Chemistry%22&Station_Group_List=~%22.*Lower%20Duwamish%20River.*%22&distinct()&orderBy(%22qm_study_id%2CQM_Station_ID%22))

Example 5: What PCB data are available for surface sediment in the Hudson River station group? (This example replicates a single contaminant Guided Query in DIVER Explorer)

Step 1: Navigate to Data Access Form for "**Samples, Northeast**", uncheck all variables

Step 2: Check Analysis, Analysis_Result_Unit, Area_Description, Case_Activity, Collection_Matrix, Collection_Study_Name, Collection_Workplan, Data_Category, Data_Classification, date, Depth_Category, DIVER_Dataset, File_Collection_ID, Lab_ID, Lab_Replicate, Lab_Result_Matrix_Detailed, latitude, latitude_end, longitude, longitude_end, QM_Sample_ID, QM_Site_ID, qm_study_id, Qualifier_Code, result_zero_nd, Sample_Depth_Unit, Sample_ID, Sample_Lower_Depth, Sample_Upper_Depth, Sharing_Status, Source_Type, Station_Group_List, Station_Site, Trip_ID, Workgroup

Step 3: Set **Analysis_Category** to "Contaminant Chemistry"

Step 4: Set **Analysis** to ="PCBs, total (QM calculated)"

Step 5: Set **Analysis_Status** to ="Results Available"

Step 6: Set **Depth_Category** to "Surface Sediment"

Step 7: Set **QM_Matrix** to ="Sediment"

Step 8: Set **Result_Type** to ="Result"

Step 9: Set **Station_Group_List** to =~".*Hudson.*"

Step 10: Select **orderBy**, then select **Case_Activity, Collection_Study_Name, QM_Sample_ID, Station_Site**

Step 11: Choose a file type and click **Submit**

Tip #8: Pay attention to quotation marks, case, and spacing. See Example 1 for guidance on getting lists of valid values.

Tip #9: Do you want your columns in a different order? You can rearrange the field order in the Variables section of the URL and send it directly to an export. (Note that when starting from the Data Access Form the export will be in the same order as the form.)

Follow this URL to view/edit the Data Access Form for Example 5: [https://pub-data.diver.orr.noaa.gov/erddap/taledap/cubes_fact_samples_northeast.htmlTable?QM_Site_ID%2Cqm_study_id%2Clatitude%2Clongitude%2CCollection_Study_Name%2CQM_Sample_ID%2CLab_Replicate%2CSample_Upper_Depth%2CSample_Lower_Depth%2CAnalysis%2CQualifier_Code%2CAnalysis_Result_Unit%2CLab_ID%2CSample_Depth_Unit%2Cresult_zero_nd%2CCollection_Matrix%2CSharing_Status%2CFile_Collection_ID%2CWorkgroup%2CCollection_Workplan%2CDepth_Category%2CSource_Type%2CCase_Activity%2CData_Classification%2CData_Category%2CDIVER_Dataset%2CTrip_ID%2CLab_Result_Matrix_Detailed%2CSample_ID%2Clatitude_end%2Clongitude_end%2Cdate%2CStation_Site%2CStation_Group_List%2CArea_Description&Analysis=%22PCBs%2C%20total%20\(QM%20calculated\)%22&QM_Matrix=%22Sediment%22&Analysis_Status=%22Results%20Available%22&Depth_Category=%22Surface%20Sediment%22&Result_Type=%22Result%22&Analysis_Category=%22Contaminant%20Chemistry%22&Station_Group_List=~%22.*Hudson.*%22&orderBy\(%22Case_Activity%2CCollection_Study_Name%2CQM_Sample_ID%2CStation_Site%22\)](https://pub-data.diver.orr.noaa.gov/erddap/taledap/cubes_fact_samples_northeast.htmlTable?QM_Site_ID%2Cqm_study_id%2Clatitude%2Clongitude%2CCollection_Study_Name%2CQM_Sample_ID%2CLab_Replicate%2CSample_Upper_Depth%2CSample_Lower_Depth%2CAnalysis%2CQualifier_Code%2CAnalysis_Result_Unit%2CLab_ID%2CSample_Depth_Unit%2Cresult_zero_nd%2CCollection_Matrix%2CSharing_Status%2CFile_Collection_ID%2CWorkgroup%2CCollection_Workplan%2CDepth_Category%2CSource_Type%2CCase_Activity%2CData_Classification%2CData_Category%2CDIVER_Dataset%2CTrip_ID%2CLab_Result_Matrix_Detailed%2CSample_ID%2Clatitude_end%2Clongitude_end%2Cdate%2CStation_Site%2CStation_Group_List%2CArea_Description&Analysis=%22PCBs%2C%20total%20(QM%20calculated)%22&QM_Matrix=%22Sediment%22&Analysis_Status=%22Results%20Available%22&Depth_Category=%22Surface%20Sediment%22&Result_Type=%22Result%22&Analysis_Category=%22Contaminant%20Chemistry%22&Station_Group_List=~%22.*Hudson.*%22&orderBy(%22Case_Activity%2CCollection_Study_Name%2CQM_Sample_ID%2CStation_Site%22))