

Access for Data Cleaning

Reference: Courtesy of: R. Testa. (c) Field Museum of Natural History.
<https://mm.fieldmuseum.org/f5332017-e914-4619-838b-a6896895b7e7> (accessed on 16
June 2022)



Agenda

Getting Access

01

Getting Your Data

02

Views

03

Filters

Queries

Functions

Relating Your Data

04



Getting Access

Access Essentials

- What is Access?

- Access is the Microsoft Office database platform

- Why use Access?

- Available to most
- User interface doesn't require knowledge of computer languages (although SQL helps)

- When should Access be used?

- To visualize large quantities of data for cleaning
- To normalize and clean data
- To produce an end product for import into the final database platform

- Need help?

- Access has a built in help feature and online help:
<https://support.office.com/en-US/Access>
- There are plenty of user group online sites that can provide help as well. Google your question directly with the word "Access" and the version to get a list of help blogs.

Getting Access

- Access is available with the Microsoft Office suite
- Explanation of the differences between versions:
<https://www.stl-training.co.uk/versions/access-difference.php>
- Security settings for the program are available:
 - By clicking the **Office button** (2007)
 - By clicking the **Options button** on the **File menu** (2010)
- The Microsoft Trust Center allows you to control the privacy settings and determine where Access saves data.

What Can/Can't Access Do?

	Access
Formulas	<input checked="" type="checkbox"/>
Charts & Graphs	<input checked="" type="checkbox"/>
Auto-formats fields	
Allow creation of lookup lists	<input checked="" type="checkbox"/>
Allow character sets besides Latin1 (e.g., UTF-8)	<input checked="" type="checkbox"/>
Track changes after closing	
Maintain integrity between rows and columns	<input checked="" type="checkbox"/>
Enforce referential integrity between different sheets	<input checked="" type="checkbox"/>
Directly changes your original data	<input checked="" type="checkbox"/>
Assists in formula writing	Limited

What is Access?

Access is a database program (or platform).

Access:

- Uses Tables
- Allows permanent and temporary relationships between tables
- Enforces referential integrity in relational data
- Builds and saves queries of different types in a user interface or SQL
- And many more things (report development, forms, macros, links, etc.)



Access Help: <https://support.office.com/en-US/Access>

Security

Getting to the Trust Center

The image shows a screenshot of the Microsoft Access interface. On the left is a dark red sidebar with buttons for 'Print', 'Close', 'Account', 'Feedback', and 'Options'. The 'Options' button is highlighted with a pink box. An arrow points from this box to the 'Trust Center' option in the 'General' ribbon tab on the right, which is also highlighted with a pink box. A second arrow points from the 'Trust Center' box to the 'Trust Center Settings...' button in the 'Microsoft Access Trust Center' dialog box at the bottom. The dialog box contains the text: 'The Trust Center contains security and privacy settings. These settings help keep your computer secure. We recommend that you do not change these settings.'

Print

Close

Account

Feedback

Options

Search

Recent Pinned

Name

Database 1 Documents

General

Current Database

Datasheet

Object Designers

Proofing

Language

Client Settings

Customize Ribbon

Quick Access Toolbar

Add-ins

Trust Center

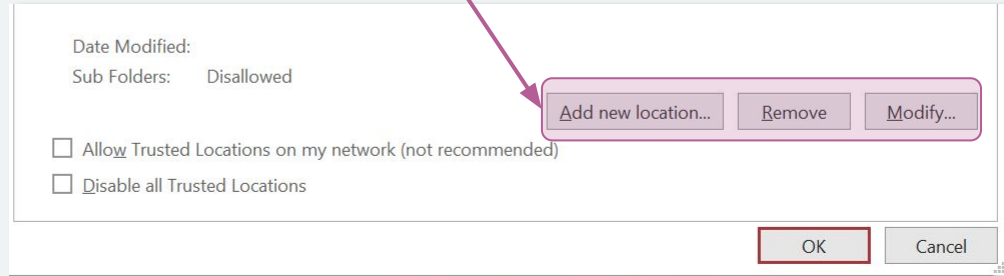
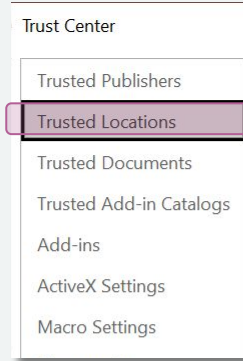
Trust Center Settings...

Microsoft Access Trust Center

The Trust Center contains security and privacy settings. These settings help keep your computer secure. We recommend that you do not change these settings.

Security

Setting a Trusted Location



Getting Your Data

Getting Your Data

Import and Link

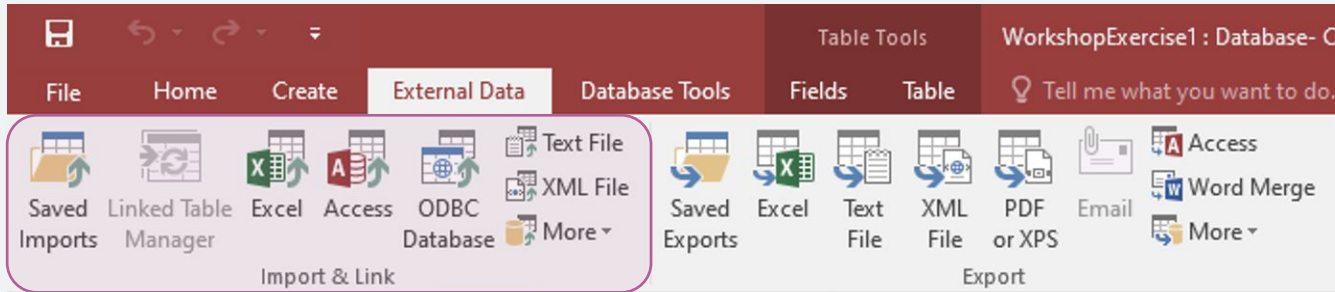
- Know where your data is
- Decide whether to import or link
 - **Import** (copies the data into Access)
 - **Link** (creates a link to the original data): Any changes made in Access are reflected in the data file. If the file is moved from the original path the link will break. (Linking data is found on the SharePoint list.)

Getting Your Data

Importing Data

In the **External Data** tab of the ribbon, select the option for the type of data being imported. For CSV files select **Text File**.

Follow the steps of the import wizard. The wizard will prompt the user with the option to save the import. Save the import only if you plan on re-importing the file in exactly the same way.



Views

Views

Filters

01

Queries

02

Functions

03

Views vs. Actions

Views

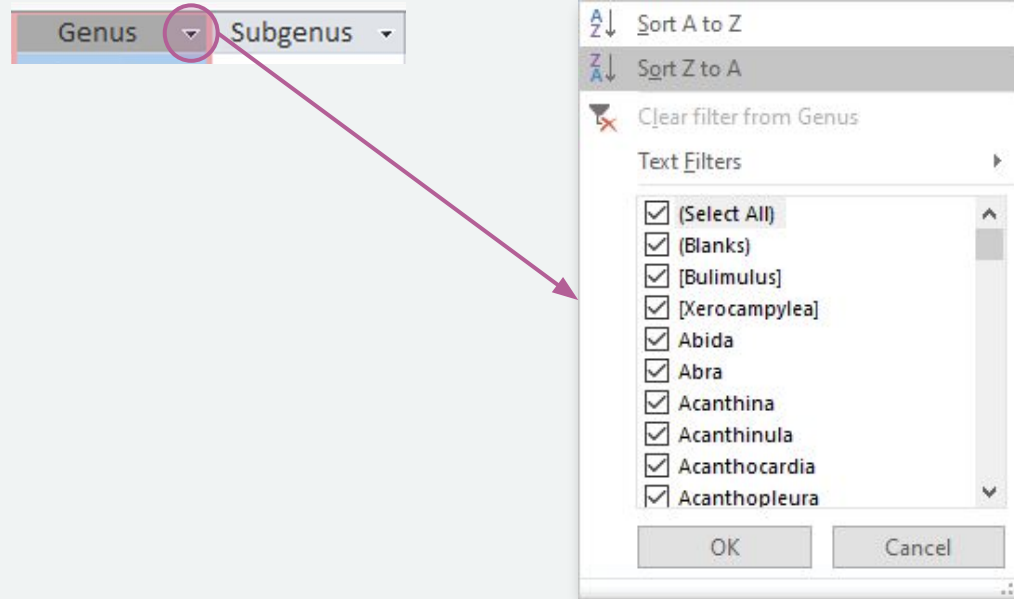
- A data view is a view or subset of data in a data table or virtual table.
 - You can create views that are filters/facets/preview pane/subsetting for content and/or fields.

Actions

- Actions are functions performed on a dataset that changes either the values themselves or the underlying structure.

Filters

Filters are available on every attribute in both tables and queries. To access them, click the drop-down icon to the right of the field name.



Creating Queries

On the Create tab, choose either **Query Wizard** (pre-canned queries, for example “Find duplicates”) or **Query Design** for almost every other kind of query.

The image shows a Microsoft Access window with the 'Create' tab selected. The ribbon contains 'Table Design', 'SharePoint Lists', 'Query Wizard', and 'Query Design'. A callout box labeled 'Design' points to the 'Query Design' option. A callout box labeled 'Wizard' points to the 'Query Wizard' option. A callout box labeled '1' points to the 'Query Type' menu in the 'Design' ribbon. A callout box labeled '2' points to the 'Query Type' menu in the 'Design' ribbon. A callout box labeled 'Click within this gray area to open the menu at left.' points to the gray area in the 'Query Type' menu.

1. The default query type is **Select**. Create a query and change the type by selecting it from the menu, or ...
2. ... by right-clicking anywhere inside the blank gray area within the new Query's tab and choosing it from the menu.

Types of Queries

To choose a query type right-click on **Query Design** and choose the **Query Type** option.

- a) **Select:** Just views data
- b) **Make Table:** Creates a new table from the selected data
- c) **Update:** Changes data according to the criteria designated by the user
- d) **Append:** Adds new rows to an existing table
- e) **Delete:** Deletes data from a table based on the criteria designated by the user
- f) **Crosstab:** Summarizes data and creates a new view; useful for data users plan to chart

Update Queries

Update queries change data in fields. Access allows users to define the change as simple text, another field, or a combination of the two.

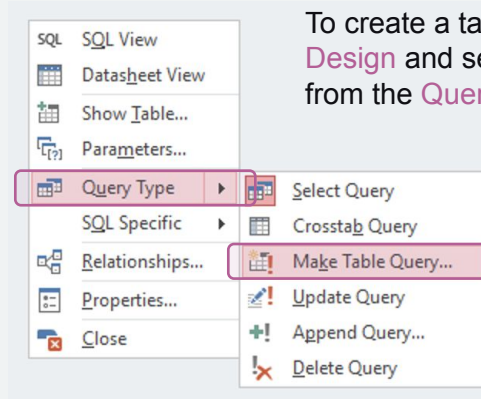
Field:	ID	ADP	Country	
Table:	WorkshopDataset	WorkshopDataset	WorkshopDataset	
Update To:			"Vanuatu"	
Criteria:			"New Hebrides"	
or:				

update To: Defines the value that will replace original data.

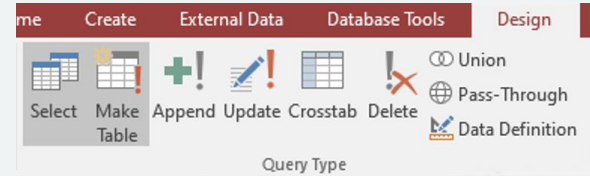
Criteria: Records must meet this criteria before the change happens.

Make Table Queries

Using tables to aid large data corrections.

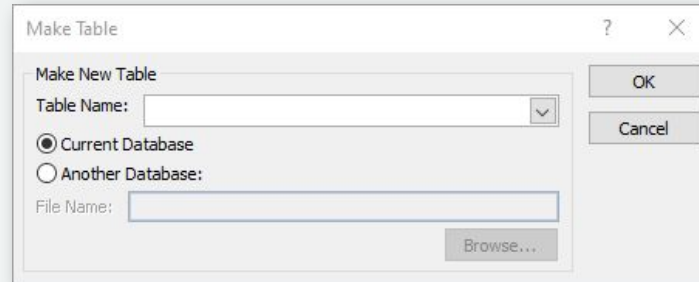


To create a table query, click on **Query Design** and select **Make Table Query...** from the **Query Type** menu; or ...



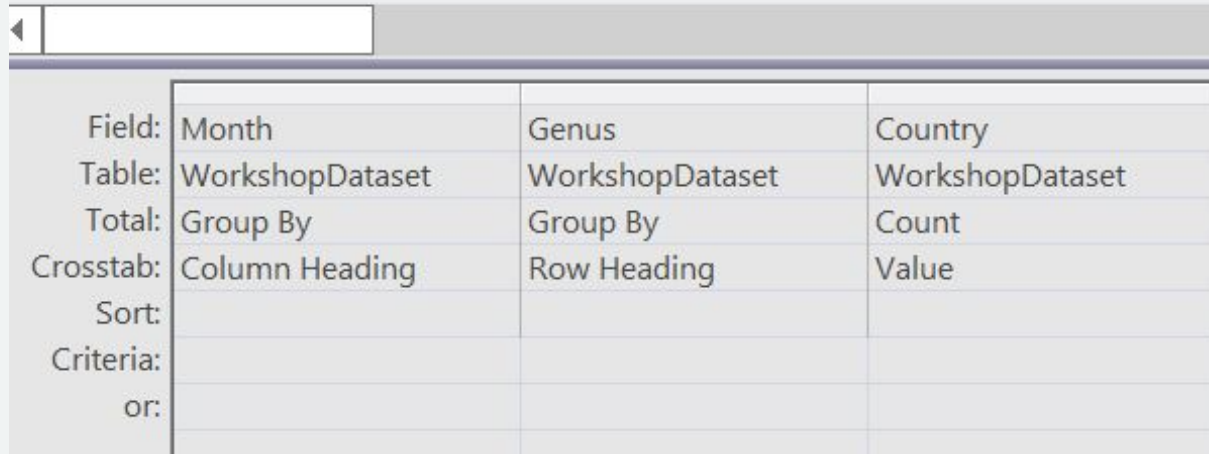
... create a query and change the type by clicking on **Make Table**.

Use the pop-up dialog to finish creating the table.



Crosstab Queries

Summarizes data and creates a new view; useful for data you plan to chart.



Field:	Month	Genus	Country
Table:	WorkshopDataset	WorkshopDataset	WorkshopDataset
Total:	Group By	Group By	Count
Crosstab:	Column Heading	Row Heading	Value
Sort:			
Criteria:			
or:			

To create a crosstab summary: select a field to serve as column headings and field to serve as row headings. Select a “value” field, the data which will be summarized in the new table.

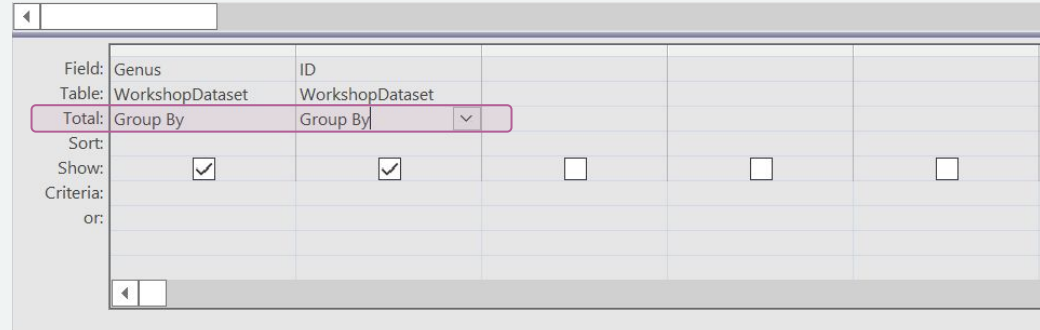
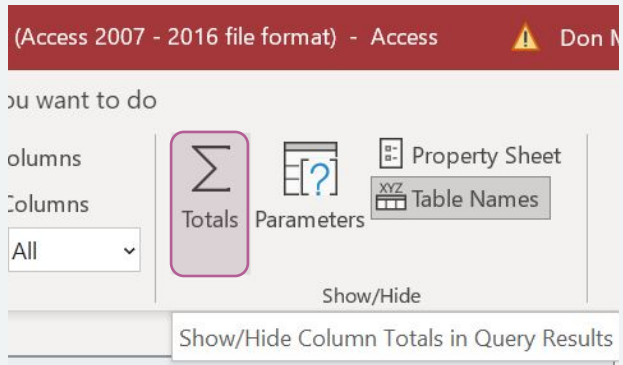


For more information, see:

<https://www.techrepublic.com/article/how-do-i-create-a-crosstab-query-in-microsoft-access/>

Summary Queries

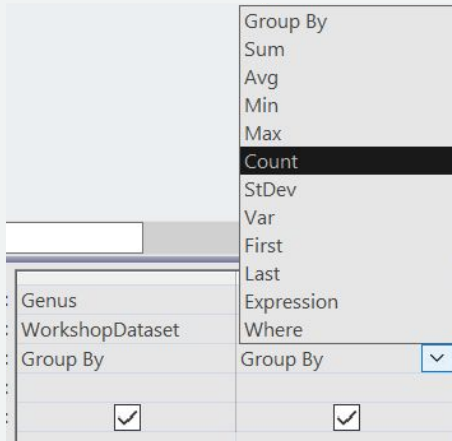
Although Access does not have a setting for “summary queries” as it does some other queries, the **Totals** Σ button can be used to add aggregating functions to a query.



Clicking the **Totals** icon adds a **Total:** option in the **Query Design** panel that can be used to select functions.

Summary Queries

Although Access does not have a setting for “summary queries” as it does some other queries, the Totals Σ button can be used to add aggregating functions to a query.



Using **Group By** and **Count** in this example yields a table that displays the number of records with each **Genus**

Genus	CountOfID
	1401
10	1
Acanthina	1
Acanthopleurin	1
Achatellinae	1
Achatellininae	1
Achatinellinae	84
Achatinelloidea	1
Achatinoidea	18
Acmaeinae	22
Adamsiellinae	8
Admetinae	1
Alcithoe	1
Alopiinae	59
Alycaeinae	7
Amblemidae	61
Amblemina	1
Ambleminae	78
Amnicolinae	41
Amphiboloidea	1
Amphidoxinae	23
Anadarinae	14
Ancillinae	3

Record: 1 of 481 No Filter

Using formulas to concatenate and rename fields

Use formulas in **Select** and **Make Table Query...** to complete actions.

- a) **Access reserved words and characters:** words or characters that have a specific meaning to the program. Learn more: [Learn about Access reserved words and symbols \(microsoft.com\)](https://learn.microsoft.com/en-us/sql/relational-databases/accessing-data-in-sql-server/accessing-data-in-sql-server?view=sql-server-15)
- b) **Split on character:**
 - [Right](#): “Returns a Variant (String) containing a specified number of characters from the right side of a string”
 - [Left](#): “Returns a Variant (String) containing a specified number of characters from the left side of a string.”
 - [InStr](#): “Returns a Variant (Long) specifying the position of the first occurrence of one string within another.”
 - [Len](#): “Returns a Long containing the number of characters in a string or the number of bytes required to store a variable.”
- c) **Concatenate:** Combines text values

Using formulas to concatenate and rename fields

Example **reserved** words and
characters:

- a) ALTER
 - b) BOOLEAN
 - c) COLUMN
 - d) DATABASE
 - e) EXISTS
 - f) FALSE
 - g) GROUP BY
- etc.

Using formulas to concatenate and rename fields

Use formulas in **Select** and **Make Table Query...** to complete actions.

IDYear: Right([UniqueSymbol], Len([UniqueSymbol]) - InStr([UniqueSymbol], " "))	Identifier: Left([UniqueSymbol], InStr([UniqueSymbol], " ") - 1)
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Access will copy into a new field called Genus (Genus: the Former identification) values starting from the left (**Left**) and ending at the point in the string (**InStr(...)**, "**|**") where there is a pipe, and then remove the pipe (-1).

In the second column Access will copy the values from Former Identification into a new field called species (species:) starting from the right (**Right**), starting at the end of the string (**Len**), starting after the first pipe (**-InStr**) and ("**|**").

Using formulas to concatenate and rename fields

Use formulas in **Select** and **Make Table Query...** to complete actions.

Split:

Field:	ID	Former identification	Genus: Left([Former	▼
Table:	WorkshopDataset	WorkshopDataset		
Sort:				
Show:	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Criteria:				
or:				

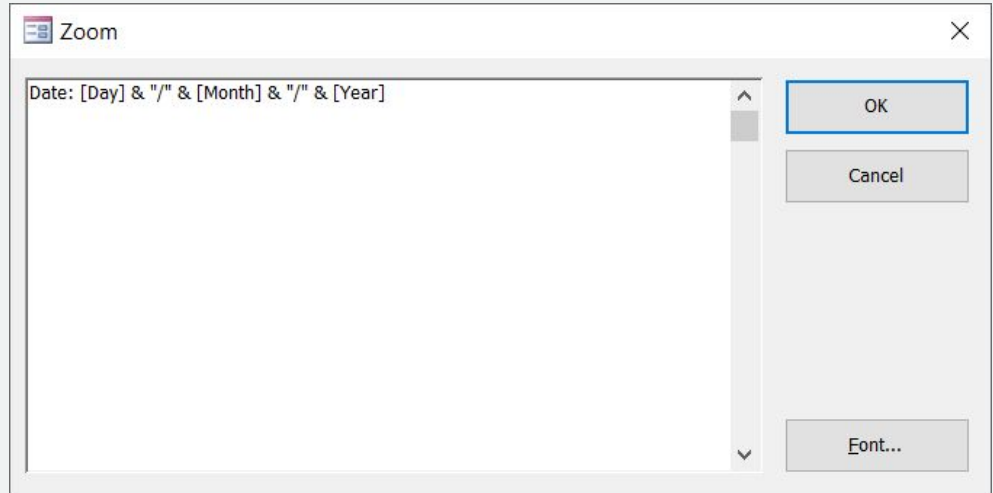
Concatenate:

ID	Year	Month	Day	Date: [Day]&"/"[Mor	▼
WorkshopDataset	WorkshopDataset	WorkshopDataset	WorkshopDataset		
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	

Using formulas to concatenate and rename fields

Use the keyboard shortcut **Shift+F2** to open a zoom pop-up window where the formulas are fully visible.

Concatenate:

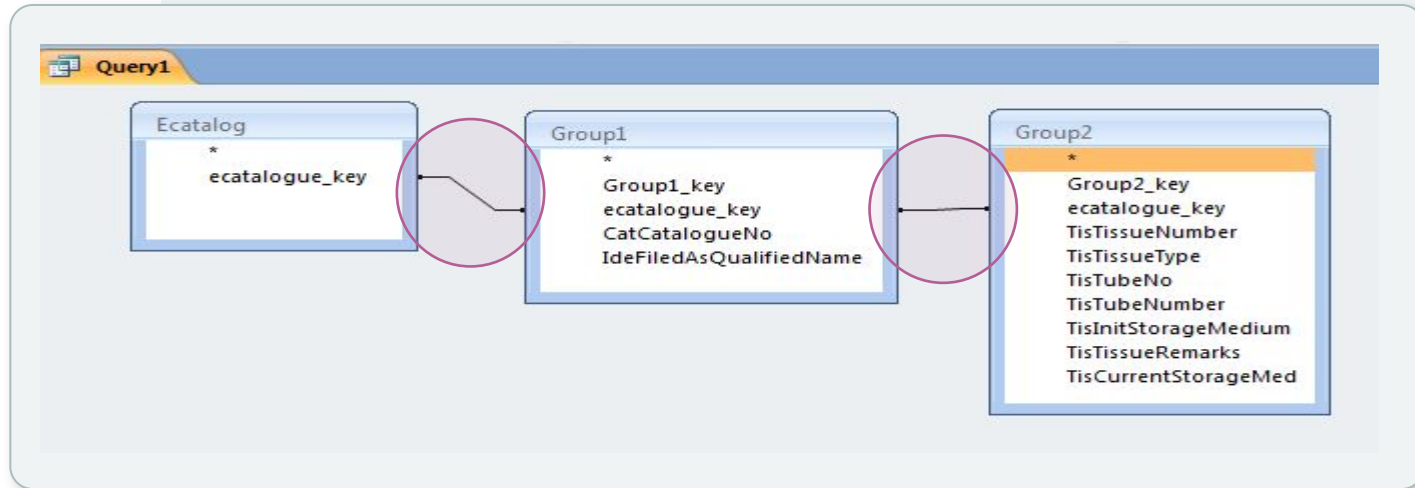


Relating Your Data

Relating Your Data

Use joins to create links between queries

- Create joins between tables in queries by dragging the join fields between tables.
 - The fields must be of the same type.
 - The default join is only where tables are equal.



Types of Joins

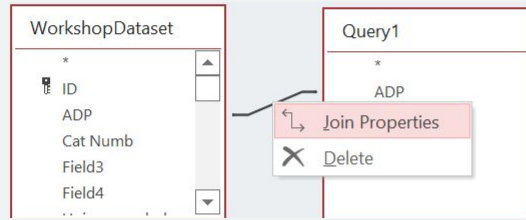
Right click on the join line to redefine the type of join.

- a) **Equal:** Only include rows where the joined fields from both tables are equal
- b) **Left:** Include all records from “First Table” and only those records from the “Second Table” where the joined fields are equal
- c) **Right:** Include all records from “Second table” and only those records from “First Table” where the joined fields are equal

Relating Your Data

Use joins to create links between queries

- To change the type of join between tables and queries, right click on the join line.
- Select the desired join type from the **join Properties** menu.



Join Properties

Left Table Name	Right Table Name
WorkshopDataset	Query1
Left Column Name	Right Column Name
ADP	ADP

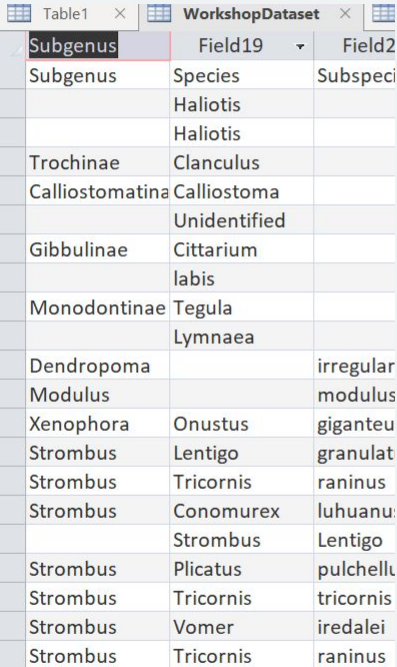
1: Only include rows where the joined fields from both tables are equal.

2: Include ALL records from 'WorkshopDataset' and only those records from 'Query1' where the joined fields are equal.

3: Include ALL records from 'Query1' and only those records from 'WorkshopDataset' where the joined fields are equal.

OK Cancel New

Renaming Fields



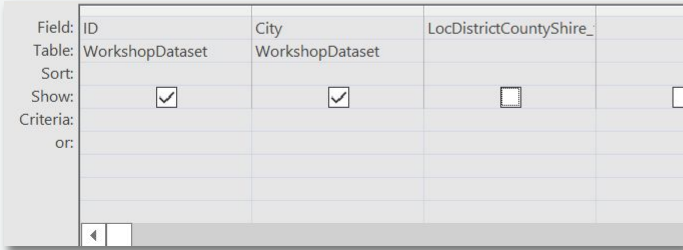
Subgenus	Field19	Field2
Subgenus	Species	Subspeci
	Haliotis	
	Haliotis	
Trochinae	Clanculus	
Calliostomatina	Calliostoma	
	Unidentified	
Gibbulinae	Cittarium	
	labis	
Monodontinae	Tegula	
	Lymnaea	
Dendropoma		irregular
Modulus		modulus
Xenophora	Onustus	giganteu
Strombus	Lentigo	granulat
Strombus	Tricornis	raninus
Strombus	Conomurex	luhuanu:
	Strombus	Lentigo
Strombus	Plicatus	pulchellu
Strombus	Tricornis	tricornis
Strombus	Vomer	iredalei
Strombus	Tricornis	raninus

Direct Design rename

← Double-click the field name to enter new text.

Using a query

LocDistrictCountyShire_tab:
[WorkshopDataset.City]

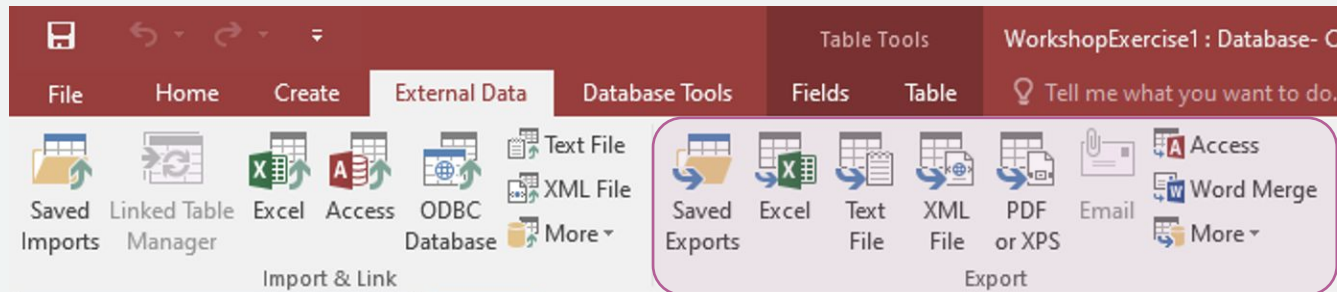


Field:	ID	City	LocDistrictCountyShire_
Table:	WorkshopDataset	WorkshopDataset	
Sort:			
Show:	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Criteria:			
or:			

Getting Your Data

Exporting Data

- Know what table/query you are exporting, where it is going and in what format you are exporting it too.
- Right click on the Access object and choose **Export**.
- Choose a format to export to (txt if you're exporting to csv).
- Follow the instructions of the wizard. If you're exporting to csv, be sure to change the extension either at the beginning or end of the export window.





Questions?

Ask now; we might have answers!

Presenter name

Title or credentials

- Bio
- Contact Information

Insert Presenter Image

Thank you!

Sharon Grant

sgrant@fieldmuseum.org

<https://www.fieldmuseum.org/>

Janeen Jones

jjones@fieldmuseum.org

<https://www.fieldmuseum.org/>

Kate Webbink

kwebbink@fieldmuseum.org

<https://www.fieldmuseum.org/>

Abigail McArthur-Self

amcarthur-self@fieldmuseum.org

<https://www.fieldmuseum.org/>

Alexis Ramirez

aramirez@fieldmuseum.org

<https://www.fieldmuseum.org/>