Importing Data

Importing data into PISCES takes a little bit of setup. Prior to importing external data into PISCES, you need to configure the data in a format that the import tools will understand. Input filters are interpreters for datasets that configure the data to be compatible with PISCES. Data comes in many forms and formats, input filters are custom configurations to standardize common types of occurrence data to HUC12s.

Data Types

PISCES code does not automatically recognize the type of data in a dataset - you will need to specify which one it is when you configure it. A dataset that does not match these formats can still be imported, but will possibly need additional code to be functional.

Point, line or polygon data where:

1. all features in the file represent the occurrence of a single species
2. each feature has a field indicating a taxon
3. each feature has multiple fields where the field name indicates the taxon and the value of that field for each feature indicates a form of presence

Tables with a field for x coordinates and a field for y coordinates and:

1. each row has a field indicating a taxon
2. each feature has multiple fields where the field name indicates the taxon and the value of that field for each feature indicates a form of presence

Input Filters

Input filters are interpreters for datasets that configure the data to be compatible with PISCES. Data comes in many forms and formats, input filters are custom configurations to standardize common types of occurrence data to HUC12s. Input filters are classes of python code that handle a type of data. They are hierarchical by default due to python and extensible due to how PISCES is built. They are all based on a core set of code and have extensions that make them more useful to a particular type of source data or types of source data.

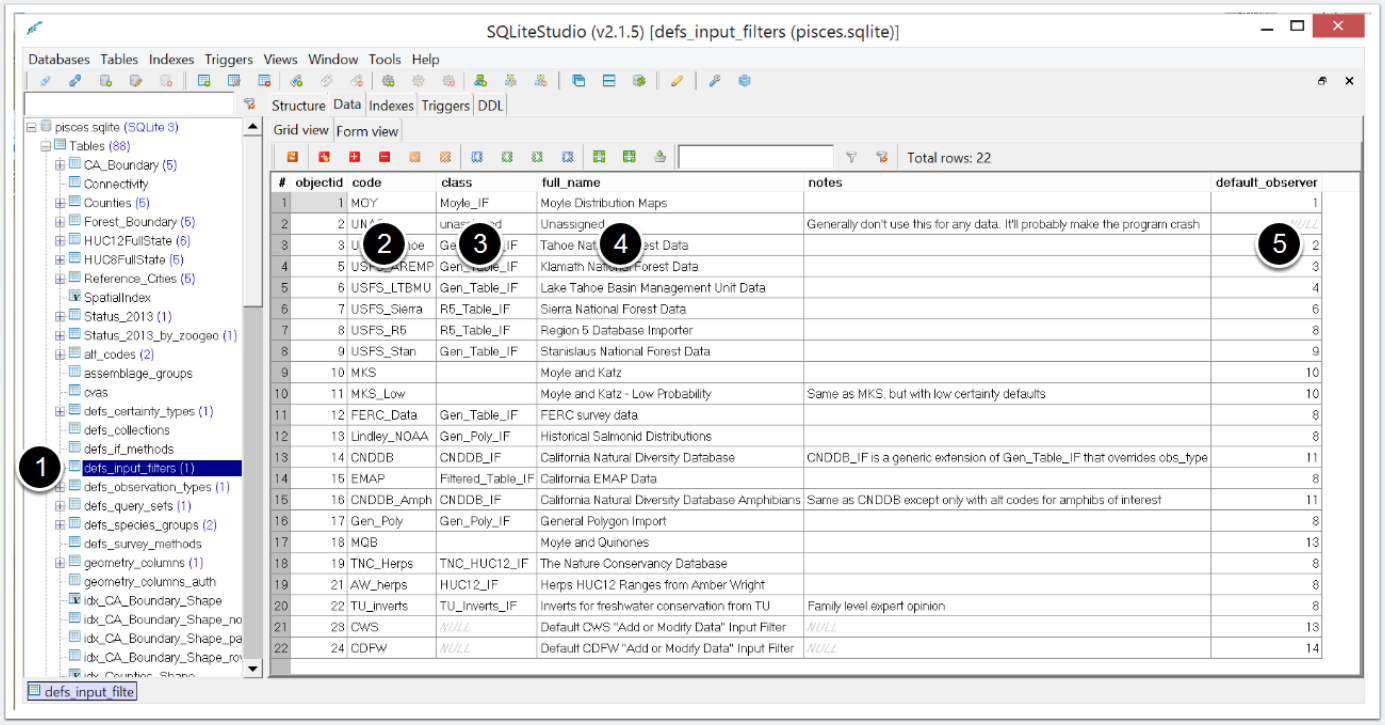
Input filters are defined in **defs_input_filters** and reference specific Python code classes that are configured to deal with different types of data.

Built-in input filters

Input filters are set in the PISCES database (pisces.sqlite). Please see the Database tutorial for more information about using SQLiteStudio to interact with the database.

To view the configured input_filters:

1. Open **defs_input_filters**
2. **Code -** unique shortname for the input filter.
3. **Class -** the Python class the input filter uses.
4. **Full_Name -** Description about the intended use for the input filter.
5. **Default_observer -** the default observer that is associated with the data. See the **Observers** table for full list.



#	objectid	code	class	full_name	notes	default_observer
1	1	MOY	Moyle_IF	Moyle Distribution Maps		1
2	2	UN…	una…ed	Unassigned	Generally don't use this for any data. It'll probably make the program crash	NULL
3	3	U…	Ge… IF	Tahoe Nat… …st Data		2
4	5	USFS_AREMP	Gen_Table_IF	Klamath National Forest Data		3
5	6	USFS_LTBMU	Gen_Table_IF	Lake Tahoe Basin Management Unit Data		4
6	7	USFS_Sierra	R5_Table_IF	Sierra National Forest Data		6
7	8	USFS_R5	R5_Table_IF	Region 5 Database Importer		8
8	9	USFS_Stan	Gen_Table_IF	Stanislaus National Forest Data		9
9	10	MKS		Moyle and Katz		10
10	11	MKS_Low		Moyle and Katz - Low Probability	Same as MKS, but with low certainty defaults	10
11	12	FERC_Data	Gen_Table_IF	FERC survey data		8
12	13	Lindley_NOAA	Gen_Poly_IF	Historical Salmonid Distributions		8
13	14	CNDDB	CNDDB_IF	California Natural Diversity Database	CNDDB_IF is a generic extension of Gen_Table_IF that overrides obs_type	11
14	15	EMAP	Filtered_Table_IF	California EMAP Data		8
15	16	CNDDB_Amph	CNDDB_IF	California Natural Diversity Database Amphibians	Same as CNDDB except only with alt codes for amphibs of interest	11
16	17	Gen_Poly	Gen_Poly_IF	General Polygon Import		8
17	18	MQB		Moyle and Quinones		13
18	19	TNC_Herps	TNC_HUC12_IF	The Nature Conservancy Database		8
19	21	AW_herps	HUC12_IF	Herps HUC12 Ranges from Amber Wright		8
20	22	TU_inverts	TU_Inverts_IF	Inverts for freshwater conservation from TU	Family level expert opinion	8
21	23	CWS	NULL	Default CWS "Add or Modify Data" Input Filter	NULL	13
22	24	CDFW	NULL	Default CDFW "Add or Modify Data" Input Filter	NULL	14

Common input filter methods

Input filter (IF) methods are python code classes that handle specific types of data. Input filter methods are python classes that handle a type of data. They are hierarchical by default due to python and extensible due to how PISCES is built. They are all based on a core set of code and have extensions that make them more useful to a particular type of source data or types of source data.

Input Filter	Description
Gen_Poly_IF	Generic Filter for polygons
Gen_Table_IF	Generic Filter for tables that have a species column and a column for x/y
HUC12_IF	Dataset that have a HUC12 Zone ID field
Filtered_Table_IF	Filters table using customized selection
Molye_IF	Moyle distribution maps importer
CNDDB_IF	Imports data from the California Natural Diversity Database
R5_Table_IF	Propriety format of USFS region 5 database

Adding input filters

Additional input filters can be expanded if needed but may require additional code to functional.

Species Codes

Most datasets have their own way of indicating what species a particular record documents. Some use scientific names, others common names, and others use various shorthands developed by those in the field.All taxa in PISCES have a unique species

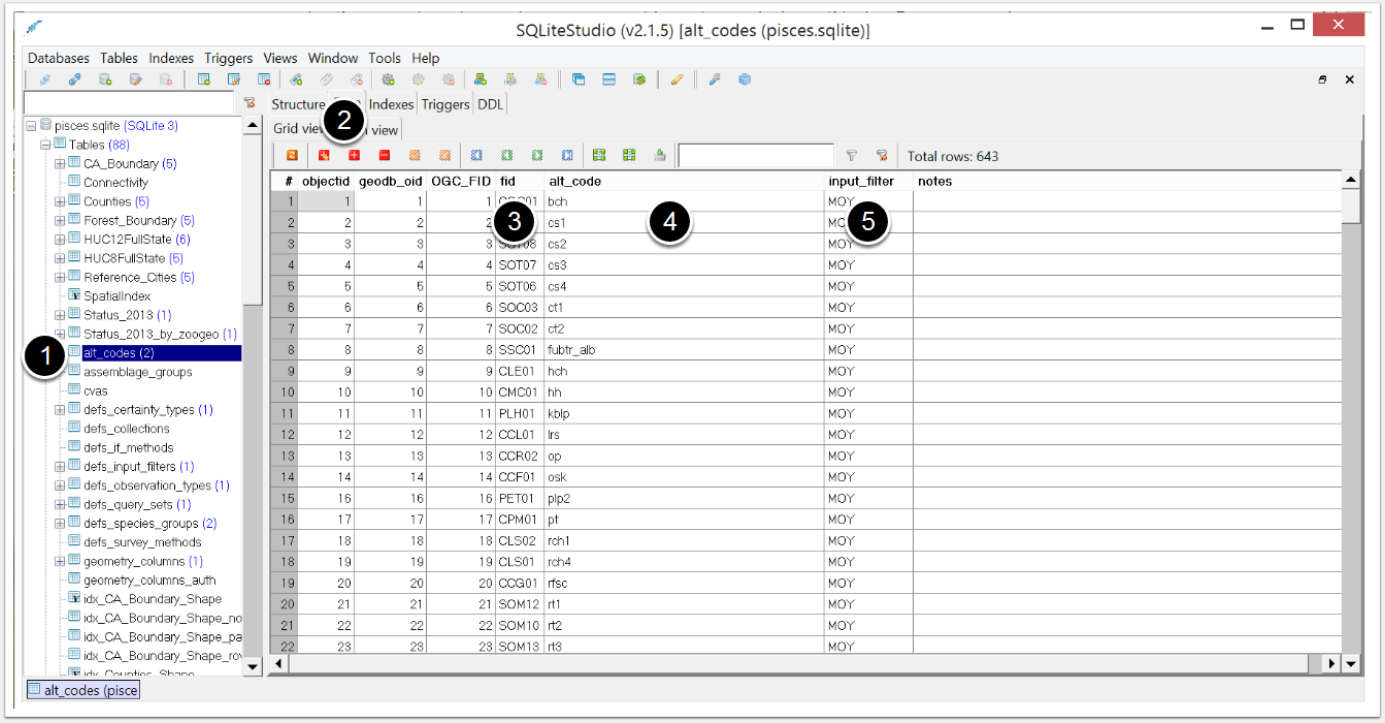
code for identification in the database. The species code used in PISCES is derived from first letter of the family, genus and scientific name plus two digits (ie ZZZ01) for each taxa. The species code is unique for each taxa in the database and all new data that is added to the database must be cross-referenced to the PISCES species code.

Alt_Codes

To import other datasets, PISCES needs to understand what species is specified by each of the taxa identifiers. This is handled via a lookup table called Alt_Codes. Each input filter needs to have the alt codes defined prior to import so that each record is assigned to the appropriate species in the database. Data bins can be used to import unknown or unresolved taxonomy but by default the records in the data bins are not used in any of the database queries or exports.

To manually add to Alt_Codes:

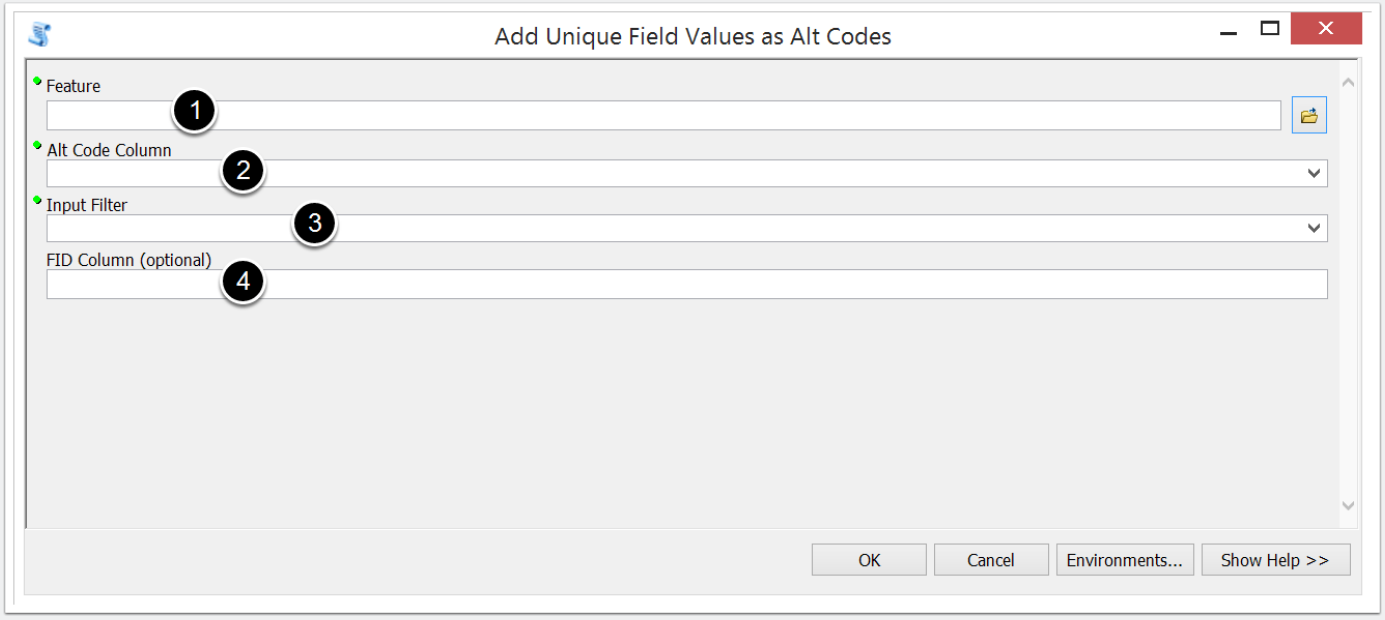
1. Open **Alt_Codes**
2. Click Add Custom Number of Rows (red plus button) - one row per identifier
3. **FID** - PISCES internal ID code for the species
4. **Alt_Code** - identifier in the data set to be imported, this can be a code, abbreviation, common name, or scientific name that uniquely identifies the species.
5. **Input_Filter** - the short name for the input_filter that the alt_codes will be associated with.



Add Unique Field Values as Alt Codes

The PISCES toolbox has a tool called **Add Unique Field Values** as Alt Codes that streamlines the process of adding species identifiers for new data sets. It extracts the unique values from a field in a table or feature class and creates a record for each one, by default without the corresponding species code. You will still need to edit the **Alt_Codes** table in pisces.sqlite and add the corresponding PISCES species code into the appropriate field.

1. Select the feature class or table that contain the values
2. Pick the attribute table column that contains the Alt Codes
3. Select the Input Filter that will be associated with these alt code matching
4. Optional: If the data set already contains PISCES species codes that is cross referenced to the alt_codes, select the column with the PISCES codes.
5. Open the **Alt_Codes** table and add the corresponding PISCES species codes to the unique values that were added.



Crosswalking Alt_Codes in Editing

It is possible to create the crosswalk reference between the alt_codes in a dataset and the PISCES species in an ArcMap editing session. This will allow you to use the **Add Unique Field Values as Alt_Codes** tool to load into the Alt_codes table using the optional parameter **FID Column.**

We can add a field to the same dataset with a column (FID column) that contains the PISCES species codes in order to make the references in ArcMap, which will then be loaded up into the alt_codes table.

1. Open attribute table for new data set
2. Add a new TEXT field
3. Start ArcMap Editing session
4. Fill in the new field with PISCES species codes (found in the Species table in the database) for all of the alt_codes
5. Save edits
6. Run **Add Unique Field Values as Alt_Codes** using the name of the new field as the option for the **FID Column** field

Import Tools

The PISCES toolbox has some tools used for configuring and importing data sets into the pisces.sqlite database.

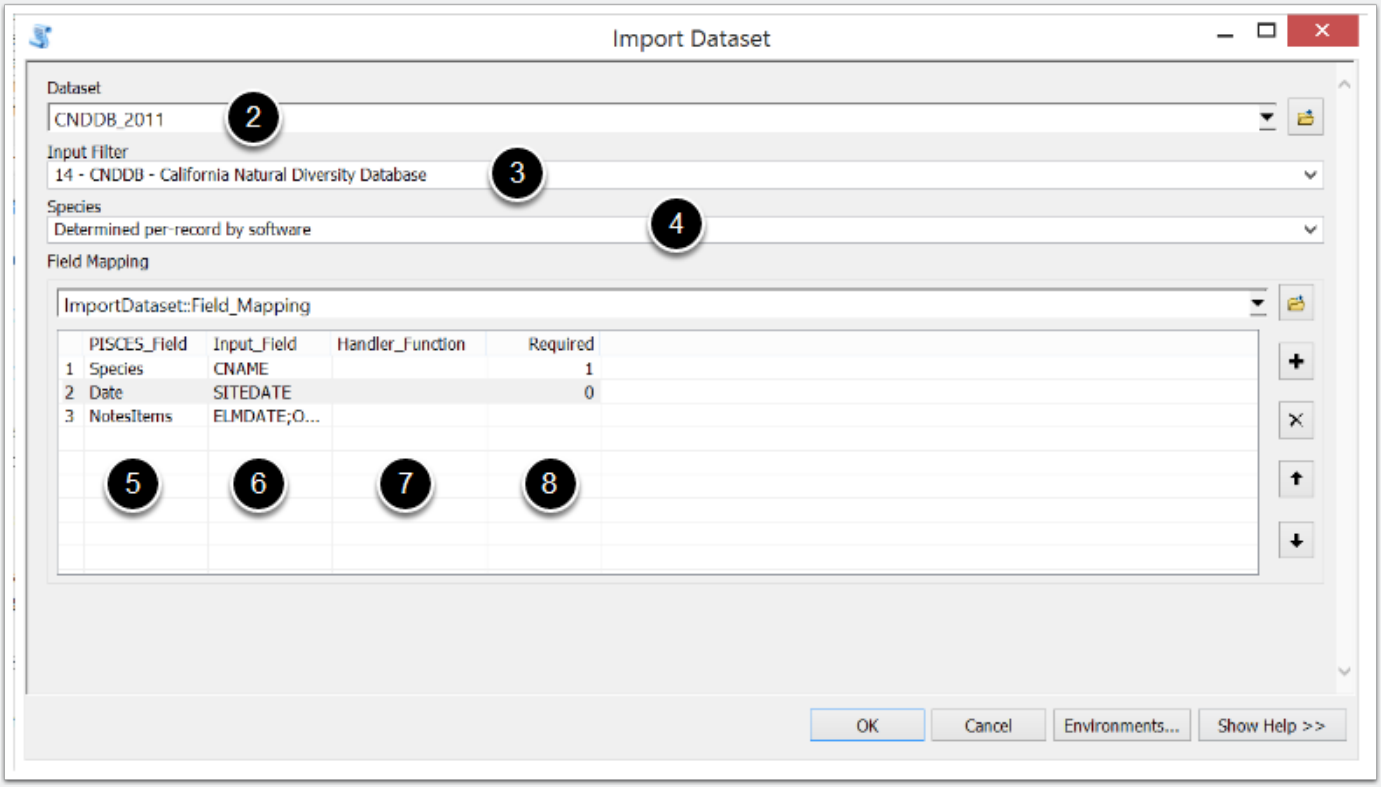
Importing Data

Import Data Tool

Prior to using the **Import Data Tool**, check that you have an **Input Filter** configured for the data that you want to import. It is also important that the **Alt_Codes** are already set up for your data if it contains multiple species.

1. Add feature or table with data to your working map document
2. Open the **Import Data Tool**
3. Select the feature or table name to be imported
4. Pick the Input Filter that is configured for that data type.
5. Select the species that the records are for. If it is multiple species, pick **Determined per-record by software** and make sure to map the species identification field
6. **Field Mapping** - user sets which fields in the import data set corresponds with the fields in the **Observations** table.
    - Species
    - Latitude
    - Longitude
    - Zone_ID
    - Date
    - Certainty
    - Observer
    - Observation Type
    - Survey Method
    - NotesItems (Multiple fields can be referenced to **NotesItems**. Each field name should be separated by a semicolon.)

7. **Handler_Function** - custom code that handles different types of data (such as mapping multiple presence types) for each field.
8. Is this field **required**? Records will not be imported if the field is missing values. 1 = required, 0 = not required



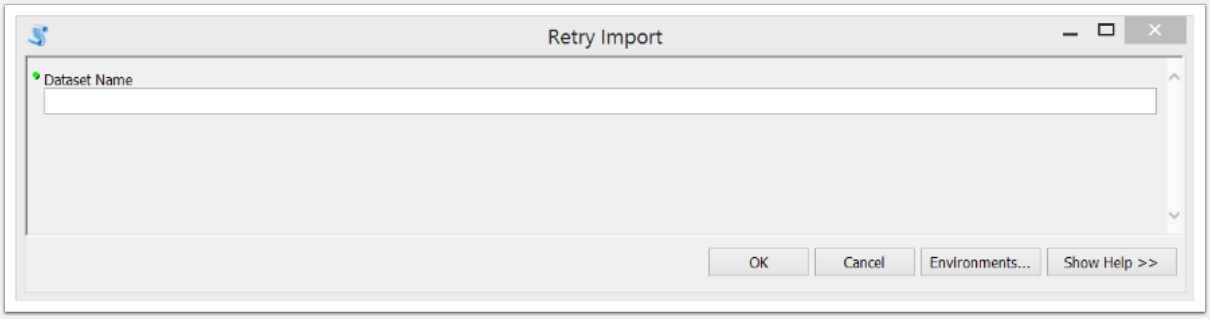
Retry Import Tool

If the import failed while using the **Import Data** tool, check to make sure that you have the **Input Filter**, **Alt Codes** and **Field Mappings** set up correctly. See the **Manually Configuring Data** section for instructions about how to edit the configuration manually in **New_Data.mdb**.

Imports occur in two stages - setting the metadata into the database to stage the import, then the actual import. If the metadata setup occurred successfully in the **Import Dataset Tool**, but the overall import failed, you can select the dataset to try again with here.

1. Open the Retry Import tool

2. Type the name of the **Dataset**



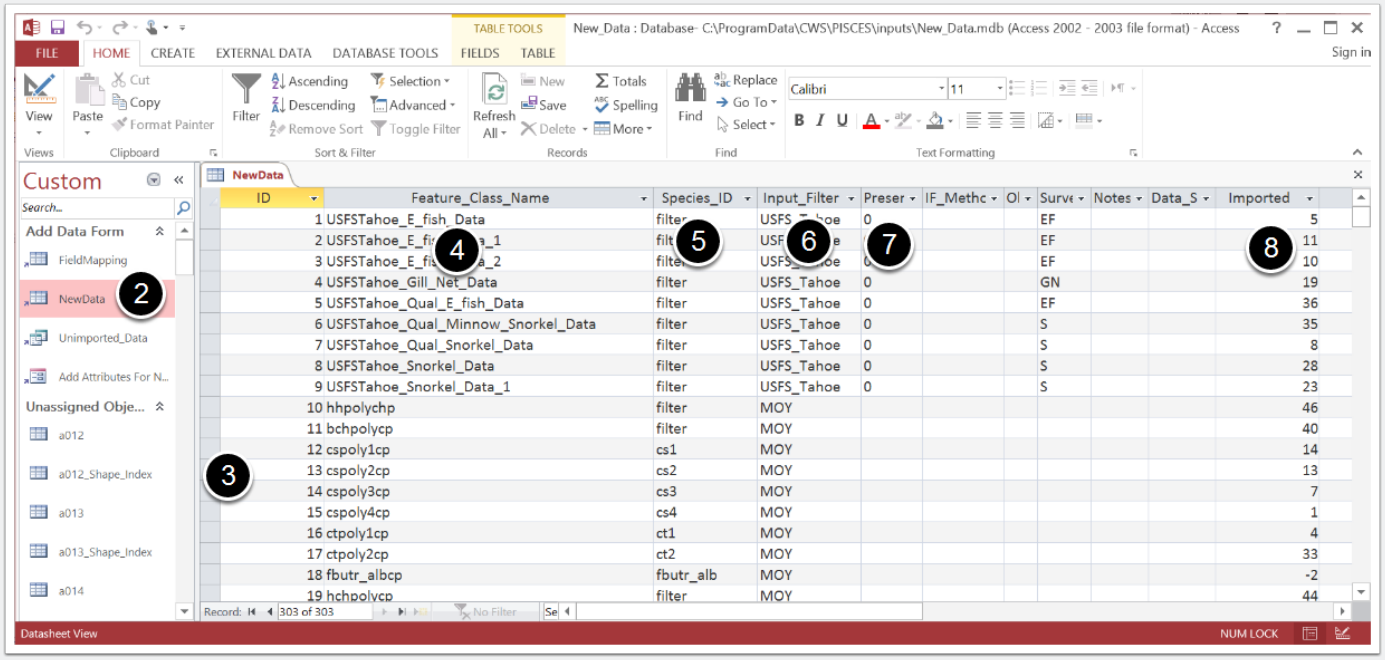
Manually Configuring Data

The configuration of new data sets takes place in **PISCES\inputs\New_Data.mdb** . Configuration of new data is required in order to tell the PISCES import tools what methods to use while adding the data to the pisces.sqlite database.

Prior to configuration, make sure that the data has a input_filter and alt_codes are already mapped.

Add feature to NewData.mdb

1. Copy the feature class or table to PISCES\inputs\New_Data.mdb.
2. Open the **NewData** table in Microsoft Access
3. Add row for new data set
4. **Feature_Class_Name** - name of the data feature or table just added to New_Data.mdb
5. **Species_ID** - a PISCES species FID or 'filter' if there are multiple species (identified by some attribute)
6. **Input_Filter** - the input filter that has been configured to use with this data set type
7. **Presence_Type** - default presence type (optional)
8. **Imported** - leave blank (field will automatically fill in with the count of the number of records successfully imported)



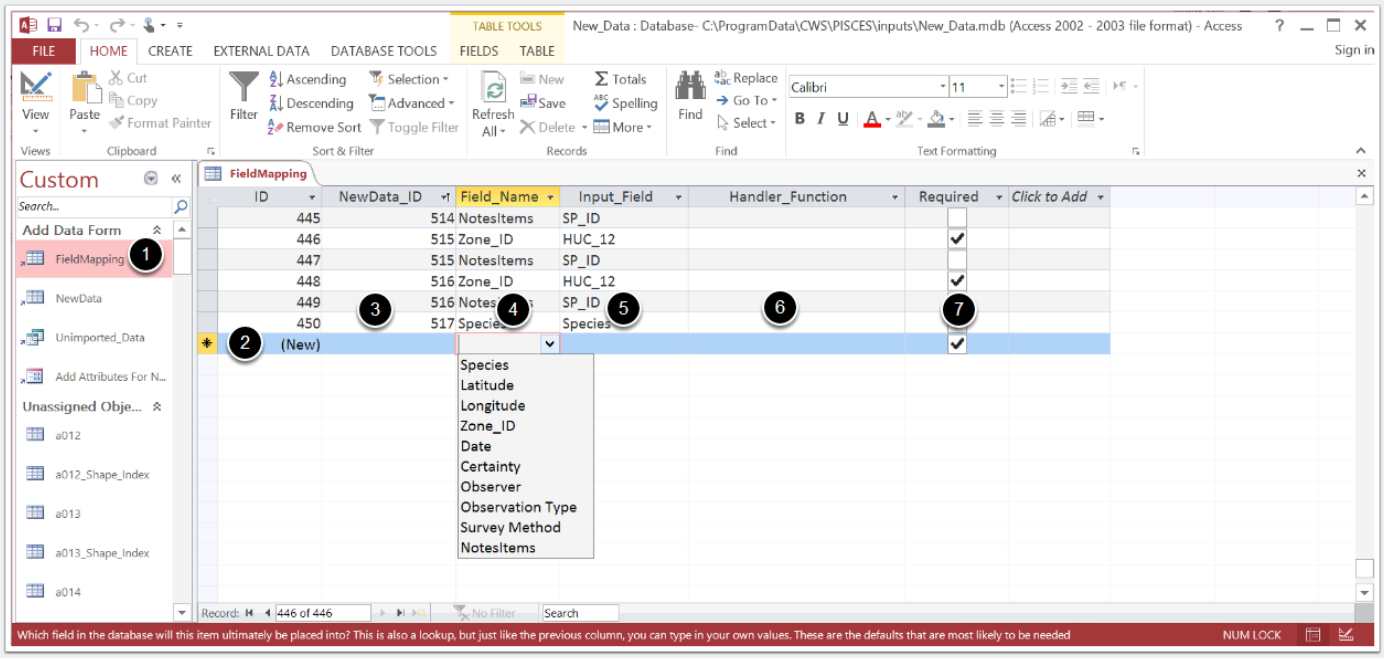
Set Field Mapping

Field mapping tell the importer what each of the fields in the new data set represent. These fields get matched with the fields that are in the **Observations** table in pisces.sqlite.

Importing Data

1. Open **FieldMapping** table in Microsoft Access
2. Add a new row for each field that needs to be mapped
3. Identify the integer ID of the feature in the NewData table
4. Select the **Field_Name** type: Species, Latitude, Longitude, Zone_ID, Date, Certainty, Observer, Observation Type, Survey Method, NotesItems
5. Write the **input_field** name for that type of data.
6. **Handler_Function** - custom code that handles different types of data (such as mapping multiple presence types)
7. Is this field **required**? Records will not be imported if this field is missing values.

Multiple fields can be referenced to the **NotesItems** field type. Each field name should be separated by a semicolon.



Command line import

The new data set can be imported using the Retry Import tool in the PISCES toolbox or through the command line.

1. Open command line at PISCES\scripts\PISCES

Importing Data

2. Run: python main.py import