**README: Figshare**

*Created by Kaija Gahm on 17 August 2020*

*Updated 28 June 2021*

**Contents of this repository**

* MFEdb\_20210423.db *the database, an SQLite .db file. This is version 4.5.4 of the database and it was created on 4/23/2021.*
* Metadata Files.zip *when unzipped: a folder, containing a whole bunch of documents describing lab and field procedures used to collect and process data. The subfolder “General or Secondary Metadata” contains metadata documents relating to the database as a whole, or metadata documents referenced in other metadata documents. Each of the metadata documents in the main “Metadata Files” folder is referenced in the METADATA database table, and corresponds to one or more metadataID’s. These metadataID’s, in turn, can be found in the `metadataID` columns of the database tables, as a means of tying individual observations back to their corresponding procedures.*
* Experiments.zip *when unzipped: a folder, containing subfolders for experiments conducted, usually in the lab. The data from these experiments may not fit the normal database table structures, so it has been saved separately, outside of the database. Sometimes experiments were smaller lab-based projects; sometimes they were pre-sampling experiments to test for e.g. between-sample variability or error rate.*
* dbSchemaKeys.20210405.pdf *a pdf database schema, illustrating the relationships between the database tables.*
* README.docx *this document.*
* DbTableDescriptions.pdf *brief ‘abstracts’ for each database table, including primary/foreign keys.*
* columnNames\_allDatabaseTables\_20210423.csv  *a csv version of UNITS, though note that this csv is slightly more up to date than the UNITS table itself.*

**Basic Structure**

The MFE database has a hierarchical structure (see the schema). The highest-level table is LAKES, followed by SITES, \*\_SAMPLES (SAMPLES, FISH\_SAMPLES, etc.), and then the tables containing data/measurements. Each data table is linked to the PROJECTS table, which describes the project for which the samples were collected. Both the SAMPLES table and the data tables are linked to the METADATA table. The ISOTOPE\_SAMPLES tables (red outlines on the schema) are on the same level as the other data tables and are linked to the SAMPLES table. The ISOTOPE\_RESULTS table is linked only to each of the ISOTOPE\_SAMPLES tables. The ISOTOPE\_BATCHES table links to the ISOTOPE\_RESULTS table.

Fish data (blue outlines on the schema) and creel survey data (yellow outlines on the schema) are somewhat separated. There is a FISH\_SAMPLES table, which contains info about the fish sampling event (lake, date, time, gear, etc.)––this is the equivalent of the SAMPLES table, but specifically for fish. The other fish tables are linked to FISH\_SAMPLES (with the exception of FISH\_YOY, which is kind of off doing its own thing––this table will be more consistently linked to the others in a future database update.) Similarly, the CREEL\_SAMPLES table gives details for each unique creel sampling/surveying event, with the other CREEL tables linked to it. Creel surveys were performed as part of the Fishscapes project, to understand “the social-ecological dynamics of recreational fishery landscapes.”

**Other Tables**

You should also be aware of some of the non-primary tables, since they’re important to navigating through the database.

* OTU: gives operational taxonomic units (otu's) for organisms whose data show up in the database. An operational taxonomic unit, in this case, is the most specific taxonomic level to which the organism could be identified.
* METADATA contains metadata information about sampling schemes, experimental procedures, etc., as well as pointing to file names when there is a separate metadata file. Metadata files are available in the “Metadata Files” folder. Each metadata file corresponds with a unique metadataID in METADATA. *NOTE: while the metadataID is sometimes the same as the name of the corresponding metadata document, this is not always the case. If you’re trying to find the metadata document that corresponds to a given metadataID, make sure you look up the correct fileName in METADATA.*
* UNITS lists column/variable names for all database tables—it is the “data dictionary” for the database. The file “columnNames\_allDatabaseTables.csv” is the same as UNITS, just in CSV format for slightly easier viewing.
* UPDATE\_METADATA describes updates made to various database tables, including the date, updater initials, and details about what was changed and why. Each database table contains an “updateID” column, corresponding to the data in UPDATE\_METADATA. *Note: for brevity, the updateID column is not included in the schema.*
* PROJECTS gives a general overview of projects for which the database data were collected, and the names of the people in charge.

**Database Schema**

The database schema shows the primary and foreign keys of each table, with arrows and colors indicating their relationships. For narrative descriptions of each table, see “dbTableDescriptions.pdf”, which can be found in “Metadata Files/General or Secondary Metadata/”.

Italicized variable names in “dbSchemaKeys.20201012.pdf” are primary keys. If there are two italicized variables in one table, then they form a composite primary key. Non-key variables are not shown; they can be viewed in UNITS.