

File Structure for USGS_DRO database

Note: If you are getting data from Sam Droege, you are getting one of two file types. Either an Excel spreadsheet with data already extracted for you, or the entire data set which is compressed into a file called **USGS_DRO_flat.gz**. The file **USGS_DRO_flat.gz** is how this file is usually shipped when we send the entire database. If you get this file it is compressed and needs to be uncompressed with 7-zip, gzip, or something similar ...after unzipping it becomes a very large txt file which can be imported into a database program for manipulation and searching. Data can also be extracted from BISON and GBIF and we assume the file structure below is reflected in their data also.

Be sure to check for spelling errors...we do global checks only every year or two.

Column Headings for USGS Bee Flat File (or an Excel Sheet if we send a subset)

ID. - Unique 6 digit specimen number with the database identifier (USGS_DRO) in front

name - Scientific name; ('Destroyed' = label destroyed without being used; 'NONBEE' = An insect that is not a bee, but was given a label; Blank = not yet identified, can indicate that specimen is waiting for identification, is lost, or was missed during data entry)

Sex - 'f' = female; 's' = male; 'u' = unknown

DeterminedBy - Who did the identification of the specimen

DeterminedWhen - Roughly when the identification was done

WhoScanned - Who scanned or entered the data

DateScanned - Date data were scanned

SpeciesNotes - Any notes about the species such as changes in identification

DateEntered - Date entered into the new database (done automatically)

COLLECTION.db - Collection event number. Up to 5 digit code preceded by database identifier (USGS_DRO) of the date, time, place of the collection event

gmt - Greenwich Mean Time of I think the time in which the Collection Event Number was created

Latitude - Latitude in decimal degrees

Longitude - Longitude in decimal degrees

accuracy - Degree of association of specimens to the Latitude/Longitude (1 = Specimens from the level of a country/state; 2 = specimens found in area of a county; 3 = specimens found in area of a park or refuge; 4 and 5 = Specimens located very close to coordinates)

elevation - Elevation of location (rarely used)

country - Country where collection occurred

state - State where collection occurred

county - County where collection occurred

city - Nearest city or the geographic unit (i.e. Park or refuge) where collected

site - A site name, number, or designation within a city/park

position - similar to "site" but used less often

time1 - Date/time collection started or traps put out (format is yyyyymmddmmss)

time2 - Date/time collection ended or traps picked up

days - Number of trapping days (used inconsistently)

who - The collector

email - Almost always Sam Droege's OLD email address

how0 - Technique used to capture bees

how1 - Number of traps that remained full of trap fluid when traps picked up

how2 - Bowl/ trap size

how3 - Trap color (rarely used since mostly we use several colors)

how4 - Trap liquid/soap

habitat - Rarely used

field_note - 'field_note' and 'note' used interchangeably to take notes about collection event

note - 'field_note' and 'note' used interchangeably to take notes about collection event