**Guidelines for data entry into RStudio**

**For all data types:**

* **Always perform a “Pull” every time RStudio is opened**
* Directory to save data files should be:

GIT/ISUmonarch/data-raw/“data”/“year”/”month”/”day”/”observer”

* *Example:* GIT/ISUmonarch/data-raw/nectar/2016/08/15/CoryH
* Observers are on data sheets. If that is blank, use “Recorder” or “Team Lead”
* Type names with first name and last initial. Ex: CoryH
* If multiple observers are listed, type them all in alphabetical order. No spaces.
* Create new folders for dates and observers as you go.
* When saving a new data entry as a text file, name it as *“site\_transect\_round.csv”*
* *Example: Uthe 1 round 1:* uth1\_tuth1a\_1.csv

*Uthe 2 round 3:* uth2\_tuth2a\_3.csv

* Scanned data sheets should be named *“site\_transect\_round.pdf”*
* Do not capitalize anything while entering data values. (Some headers may have capitalized words, leave them as-is)
* Perform a ‘Commit’ after every data file is created/saved/changed/updated
* For newly added data, Commit message should be:

Add YYYY/MM/DD site\_transect\_round

* Every time a data file is altered (file title will turn red) it will need to be saved and then updated with a ‘commit’
* For updating/fixing typos in already-saved data files, Commit message should be:

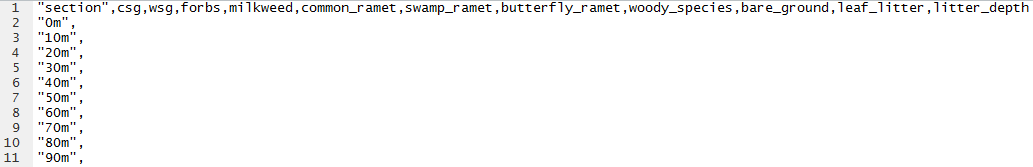
Update YYYY/MM/DD site\_transect\_round

* Then press ‘enter’ twice to move down two lines and explain in detail exactly what you did to the file.
* **ALWAYS double or triple check your entry** before saving and performing a commit. It gets annoying really quickly to have to go back and make small fixes to multiple data entries.
* Perform a ‘Push’ every time you walk away from the computer or take a break

**Nectar data**

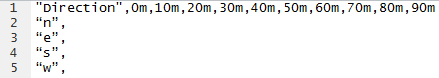
*  Copy and paste the header from a previously entered text file csv (from previous year):
* Pay close attention to transect lengths from site to site. Change accordingly for shorter transects.
* Enter species totals for each 10m interval, separated by commas. No spaces.
* *Example:*
* Start a new row/line for every different nectar species.
* If a site has no nectar present, simply type the header, fill in zeros for milkweed ramets, and leave the rest blank.
* Always use common names from the “Nectar Plant ID Guide Spreadsheet” located in the Survey folder in CyBox.
* Common names of nectar species should be in quotations and *should use spaces.*
* Do not capitalize nectar species names. Do not put dashes/hyphens/apostrophes in nectar species names.
* For the milkweed strip sites (this only applies to arm1 and nkn1) the row header is “milkweed strip: common milkweed ramet” for the common milkweed counted in the milkweed strip
* \*\*teams were not able to differentiate between Pennsylvania and swamp smartweed. Enter whatever is written down. **Will use an r-script to combine them into one species name.**
* \*\*2017 uth1, uth2: ‘b’ means over tile ‘uth1b’ and ‘uth2b’. No data for uth1a or uth2a R1. Uth1a and uth2a are the true plots. ‘b’ will only be in 2017 data for nectar, daubenmire, and robel

**Daubenmire** (Robel data files will be separate files from Daubenmire data)

* Copy and paste the header from a previously entered text file csv (from previous year):
* For each 10m point on the transect, enter the value of each data type separated by commas (corresponding to the header). No spaces.
* On transects that are shorter than 100m, remove the 10m sections that weren’t surveyed.
* \*\*2017 uth1, uth2: ‘b’ means over tile ‘uth1b’ and ‘uth2b’. No data for uth1a or uth2a R1. Uth1a and uth2a are the true plots. ‘b’ will only be in 2017 data for nectar, daubenmire, and robel

**Robel**

* These will be placed in the same directory as the Daubenmire files and will be named the exact same per site, but with a “2” added after the .csv extension: *site\_transect\_round.csv2*
* Copy and paste the header from a previously entered text file csv (from previous year):



* Enter the values of each cardinal direction for every 10m point on the transect, separated by commas. No spaces.
* \*\*2017 uth1, uth2: ‘b’ means over tile ‘uth1b’ and ‘uth2b’. No data for uth1a or uth2a R1. Uth1a and uth2a are the true plots. ‘b’ will only be in 2017 data for nectar, daubenmire, and robel

**Bees**

* Copy and paste the header from a previously entered text file csv (from previous year):

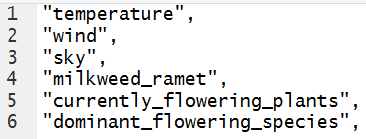


* Pay close attention to transect lengths from site to site. Change them accordingly for shorter transects.
* If no pollinating species were found at a site, simply type the header and leave the rest blank.
* Actual names of nectar and pollinator species should be in quotations and *should use spaces.*
* Do not capitalize nectar or bee species names. Do not put dashes/hyphens/apostrophes in nectar or bee species names.
* There are three options for bee names: “honey bee” “bumble bee” “solitary bee”
* Start time should be written in *xx:xxPM* format. Do not include a zero in the hour slot if it isn’t 10, 11, or 12.

*Examples:* 9:36AM *or*  12:28PM

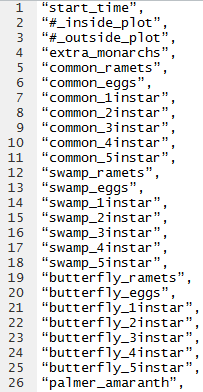
**Environment**

* Copy and paste the header from a previously entered text file csv (from previous year):



* Do not type % for any of the ‘currently flowering plants’ data
* Do not type the > for >0-5%. Simply type 0-5
* Dominant flowering species
  + Always use common names from the “Nectar Plant ID Guide Spreadsheet” located in the Survey folder in CyBox
    - \*\*rule difference\*\* all **common name spaces should be replaced with underscores.** Example: wild\_parsnip
      * This is different from all other data type entries of nectar plant common names. This is because in the ‘Environment’ data set the common names are outside of quotes and therefore, must remain linked in order to maintain them together
  + If no dominant species present:
    - “dominant\_flowering\_species”,
* If multiple dominant species present: (semi colon and space)
  + “dominant\_flowering\_species”,wild\_parsnip; canada\_thistle

**Monarch**

*  Copy and paste the header from a previously entered text file csv (from previous year):
* Start time should be written in *xx:xxPM* format. Do not include a zero in the hour slot if it isn’t 10, 11, or 12.
* *Examples:* 9:36AM *or*  12:28PM
* Simply enter the number value from the data sheet that corresponds to the information on each line.
* Fis1 8-15-17. No monarch survey: rain. Entered all monarch data as ‘NA’