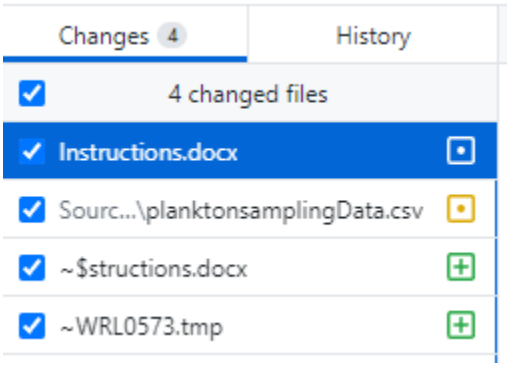
HSC Data Upload Procedure

Github Desktop

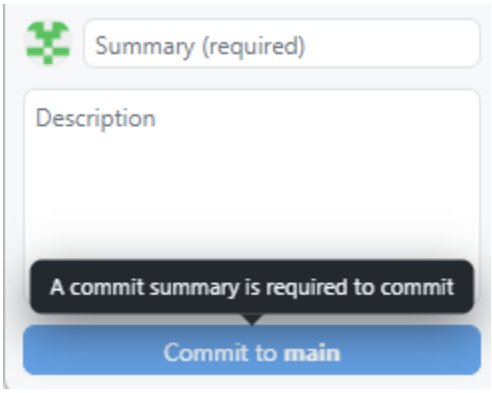
When any changes are made to any files in the github repository, these changes need to be uploaded via Github Desktop to also be reflected in Sharepoint and other areas. This includes making the survey plans, survey results, and updating the total data compendium. When making survey plans and results, the main path where data can be found or saved is:

```
C:\Users\herri\Documents\GitHub\HerringScience.github.io\Surveys\YEAR\SURVEY
```

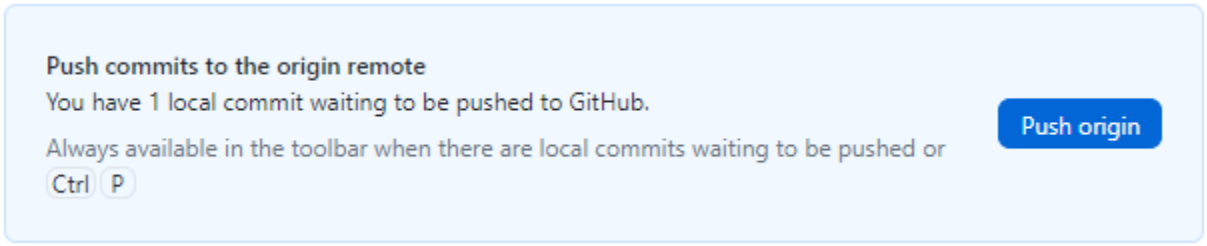
- Once all the steps are complete for either a survey plan or survey results, load Github desktop (either on the taskbar or search for it). 
- The first page should immediately show any and all changes that occurred to the files within the github folders (such as me writing these instructions):



- These changes need to be 'committed' (e.g. saved). The only other step is that this commit needs to have a summary title (e.g. "Survey Plans for SB5", "GB3 Survey Results", etc.). Afterwards, the 'Commit to main' button can be pressed.



- Finally, these saved changes can be uploaded by using the Push to Origin (e.g. upload to github online) button on the next screen.



- Any changes to .html files that are hosted by github pages (which are the links on Sharepoint) may take a few minutes to appear (e.g. the Total Data Compendium).

Creating a Survey Plan

Step 1) **Scots Bay ONLY:** Take screenshots of the tidal forecast for your ground:

Margaretsville station

The following screenshots should be taken:

1) A screenshot of the survey day (e.g. Sunday) and the proceeding day. **This needs to be saved exactly as "Daily.jpg"** and placed into the survey folder.

2022-08-07 (Sun)

Time AST	Height (m)	Height (ft)
00:04	1.3	4.4
06:17	9.2	30
12:30	1.6	5.1
18:43	9.7	31.8

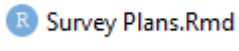
2022-08-08 (Mon)

Time AST	Height (m)	Height (ft)
01:07	1.4	4.5
07:20	9.1	29.8
13:32	1.6	5.4
19:45	9.7	31.8

2) A screenshot of the survey day and the proceeding day under hourly predictions. **This needs to be saved exactly as "Hourly.jpg"** and placed into the survey folder.

Hourly Predictions (m)

Event Date	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23
2023-01-10	7.4	8.6	9	8.3	7	5.2	3.5	2.4	2	2.5	3.8	5.6	7.5	8.8	9.3	8.8	7.5	5.6	3.7	2.3	1.6	1.8	2.9	4.7
2023-01-11	6.6	8.1	8.9	8.8	7.8	6.2	4.4	2.9	2.1	2.1	3	4.6	6.5	8.1	9.1	9.1	8.3	6.7	4.8	3	1.9	1.6	2.2	3.7

Step 2) Open the Survey Plans RMD Report in the Github folder 

Step 3) Change the options listed in the first code chunk to match the plan coordinated by the managers. Many of these options need to be formatted exactly as described.

Example German Bank:

```
## Global options
rm(list = ls())
#Change these options
surv = "GB"
surv2 = "German Bank"
surv.date = "2023-08-06 20:30:00"
surv.no = "1"
Allocation = "100" #set to "0" for a non-fishing survey
SIAllocation = "50" #Seal Island allocation for GB-only
Tagging = "Morning Star, Lady Janice II, and Lady Melissa"

#Set vessels below
vessels = 4
EVessel = NA #Set NA for GB or excluding box
NVessel = NA #Set NA for GB or excluding box
PlanktonVessel = "Lady Janice II"

#Main box only (e.g. if vessels = 9 with both SB boxes, sho
V1 = "Morning Star"
V2 = "Lady Melissa"
V3 = "Leroy and Barry"
V4 = PlanktonVessel
V5 = NA
V6 = NA
V7 = NA
V8 = NA
V9 = NA
```

Example notes: EVessel and NVessel are left “NA” as this only pertains to Scots Bay. As “vessels = 4”, there are four vessels named below under V1-V4. The PlanktonVessel (often Lady Janice II) should be the furthest east vessel (and last listed in V#) for German Bank, to line up with the CTD/tow box better.

Example Scots Bay (with both Boxes):

```
## Global options
rm(list = ls())
#Change these options
surv = "SB"
surv2 = "Scots Bay"
surv.date = "2023-08-06 20:30:00"
surv.no = "1"
Allocation = "100" #set to "0" for a non-fishing survey
SIAllocation = NA #Seal Island allocation for GB-only
Tagging = "Morning Star, Lady Janice II, and Lady Melissa"

#Set vessels below
vessels = 6
EVessel = "Fundy Monarch" #Set NA for GB or excluding box
NVessel = "Canada 100" #Set NA for GB or excluding box
PlanktonVessel = "Lady Janice II"

#Main box only (e.g. if vessels = 9 with both SB boxes, sh
V1 = "Morning Star"
V2 = PlanktonVessel
V3 = "Leroy and Barry"
V4 = "Lady Melissa"
V5 = NA
V6 = NA
V7 = NA
V8 = NA
V9 = NA
```

Example notes: In this case, 'vessels' is increased to 6 but only 4 are listed in the V1-V4 section as this is for the main box only. The other two vessels are listed as the EVessel (East Box) or NVessel (North Box). For Scots Bay, the main box V# vessels + EVessel + NVessel should equal the total 'vessels'. SIAllocation is ignored as this pertains only to Seal Island during German Bank surveys. Finally, the PlanktonVessel is often assigned V#2-3 as this will line up with the CTD and tow box.

Step 4) Once everything is in place, use the Knit button 

to create the survey plan document. If you run into any errors, try to source the problem and troubleshoot it. Make sure you have all required files saved (Hourly and Daily tides for Scots Bay) with their proper format and names.

Survey Plan Coding Dictionary

Creating the Survey Results

After a survey, the following data will have been collected:

Data	Data/Program(s) Needed	Where it goes
Plankton Tow Data	Excel	Fill out PlanktonData.csv in Survey Folder

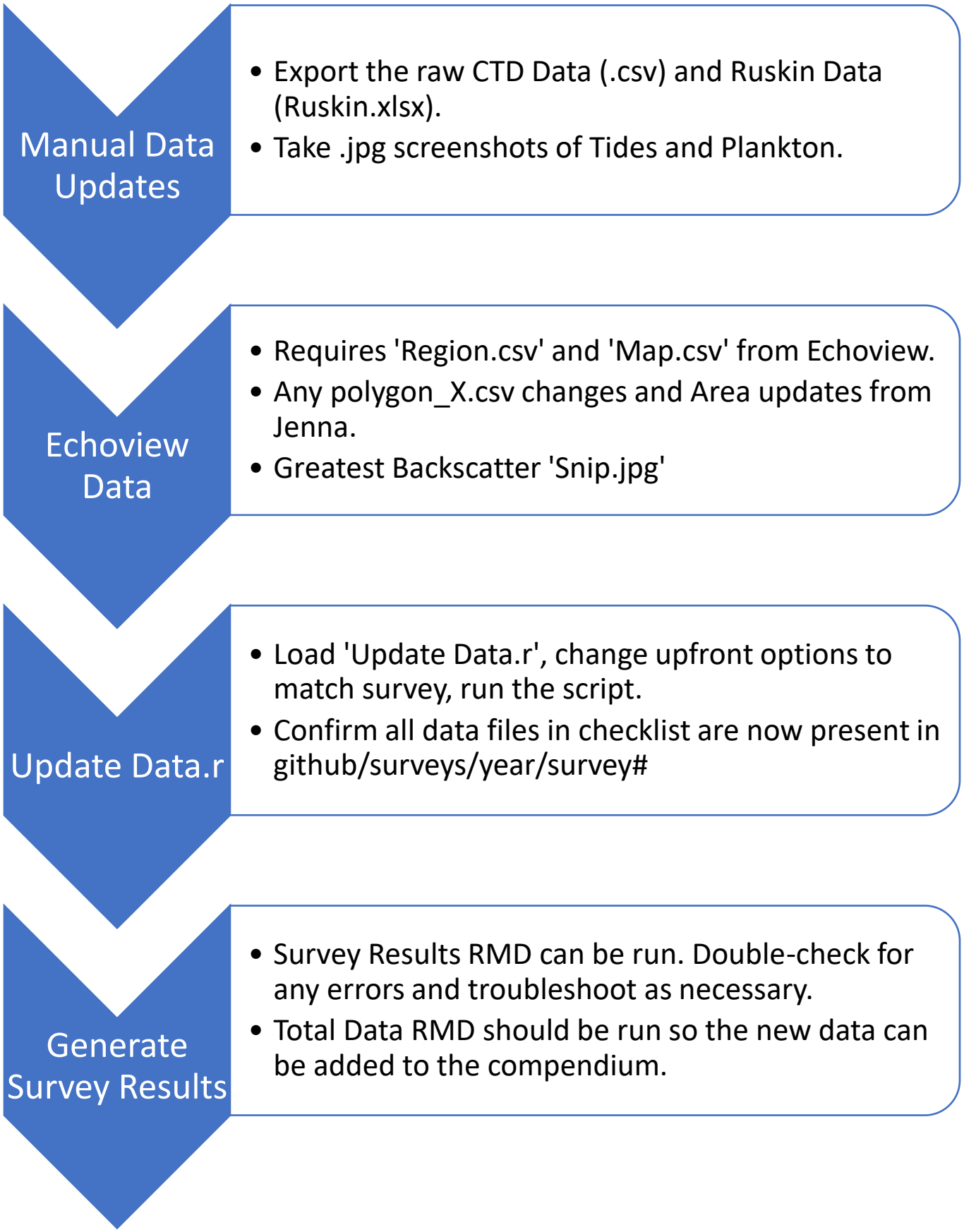
CTD Cast	YSI Castaway CTD	Export raw data .csv to Survey Folder
Depth Probe	Ruskin RBR	Export raw data .xlsx to Survey Folder as “Ruskin.xlsx”
Tagging Reports	Tagger Logs	HerringScience.github.io > HTML Markdown > Tagging Data.R
Map and Region .csv’s from Jenna’s Echoview	Echoview output	Github/Survey/Year as Map.csv and Region.csv
Survey Box Polygons from Jenna	polygon_SBEastern, polygon_SBNorthern, polygon_SB, and any GB changes	Github/Survey/Year as: polygon_SBEastern.csv polygon_SBNorthern.csv polygon_SB.csv
Greatest Backscatter Snip from Jenna	Save exactly as ‘Snip.jpg’	Github/Survey/Year/
Final Step: All data changes	Update Data.R script	HerringScience.github.io > HTML Markdown > Update Data.R

Full Survey Results Checklist

All the below files are made at some point in the results process, but this checklist can help if you find one is missing. **All files must be named exactly as below, and the same .jpg or .csv format!**

Daily.jpg	.jpg snip of the daily tides forecast for the survey day and the next day
Hourly.jpg	.jpg snip of the hourly tides forecast for the survey day and the next day
Distance.csv	Analysis of vessel transect distances made after Update RMD Data, if tableA and plan.csv exist
Map.csv	Output from Jenna’s Echoview work, should be manually saved into the survey folder
Performance Total.csv	Analysis of vessel performance made after Update Data, if tableA and plan.csv exist
survey plan.csv	
Plankton.jpg	Photograph of any jars of plankton from the tow(s), needs to be saved as .jpg
Region.csv	Output from Jenna’s Echoview work, should be manually saved into the survey folder
Snip.jpg	.jpg snip of the single-point with the greatest backscatter in Echoview, should be taken by Jenna and saved in the survey/year folder
Speed.csv	Analysis of vessel speed made after Update Data, if tableA and plan.csv exist
adhoc.csv	Should be saved directly to the survey/year folder if any vessels conducted an adhoc school survey after the scheduled survey
tableA.csv, tableB.csv, tableC.csv	Tables created in the Update Data script, if Map and Region exist. Also requires any “adhoc.csv” files to be present, if there was an adhoc survey conducted

Tow 1.jpg, Tow 2.jpg	Created in Update Data.r through ggplot and the raw Ruskin data
polygon_X.csv	.csv files of any polygon changes sent by Jenna. If they didn't change, need to copy the same set from the previous survey's folder, or the default boxes are in the "Box Coordinates" folder in the root github directory.
CTD Data: e.g., 10K100766_20230730_224751.csv	Raw .csv export of the CTD data via the CTD software. No name change necessary, but the CTD_ID (file name) needs to be recorded in the PlanktonData.csv sheet
Ruskin.xlsx	Raw .xlsx export of the Ruskin RBR Depth Probe data via their software.


Manual Data Updates
• Export the raw CTD Data (.csv) and Ruskin Data (Ruskin.xlsx).
• Take .jpg screenshots of Tides and Plankton.
Echoview Data
• Requires 'Region.csv' and 'Map.csv' from Echoview.
• Any polygon_X.csv changes and Area updates from Jenna.
• Greatest Backscatter 'Snip.jpg'
Update Data.r
• Load 'Update Data.r', change upfront options to match survey, run the script.
• Confirm all data files in checklist are now present in github/surveys/year/survey#
Generate Survey Results
• Survey Results RMD can be run. Double-check for any errors and troubleshoot as necessary.
• Total Data RMD should be run so the new data can be added to the compendium.

Plankton Tow Data

Each Github survey folder (Surveys/Year/Survey, e.g. Surveys/2023/SB5/) will have a blank "PlanktonData.csv" sheet that needs to be filled out as entirely as possible. There will be **one row per plankton tow completed**.

Set_Number	No_jars	Tow_No	Lon1	Lat1	Lon2	Lat2	Time1	Time2	TideDirection	Speed	Heading	Observers
WindDirection	WindSpeed	Swell	AirTemp	FlowReading1	FlowReading2	DiscDepthD	DiscDepthA	CTD_ID	CTD_Lat	CTD_Lon		

Plankton Data Columns

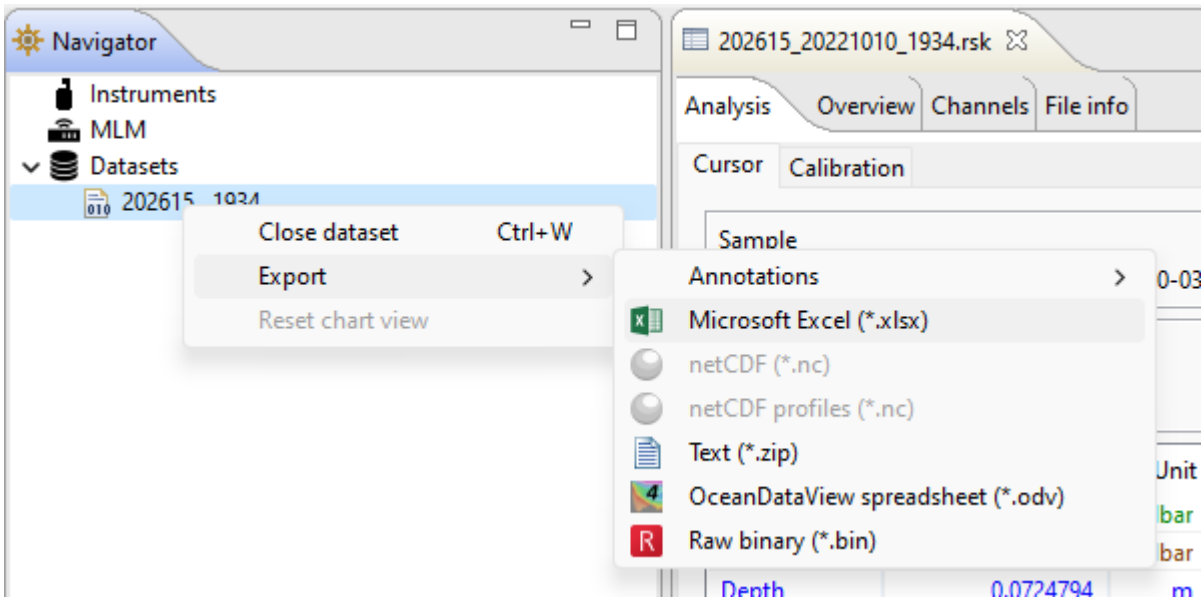
Ruskin Depth Probe Data

Data collected on the Ruskin RBR Depth Probes needs to be extracted using the Ruskin Software. This can generally be done on the boat and saved ahead of time, as you need to plug the probe into the software to turn it on/off during the plankton tows anyways.

Step 1) Load the Ruskin RBR software (can 'search' for "Ruskin" on windows). 

Step 2) Under the top left Navigator, right click the dataset you are currently using, choose Export, and select Microsoft Excel (*.xlsx) format. This should be saved in the Github folder for your current Year and Survey. **It needs to be named as "Ruskin.xlsx".** The exact path of this should be:

C:\Users\herri\Documents\GitHub\HerringScience.github.io\Surveys\YEAR\SURVEY\Ruskin.xlsx



Plankton Tow Jars

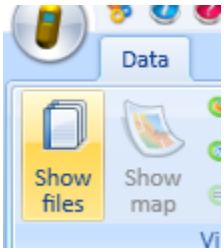
A photo of the plankton tow jars needs to be taken and saved as **"Plankton.jpg" in the survey folder.** It is best to take one photo of both jars, but if more detail is needed of the captured plankton then two separate photos can be taken and stitched together later (in Paint or other programs).

CTD Cast Data

For the CTD data, export the cast data as a .csv file using the YSI Castaway CTD software:

Step 1) Load the YSI Castaway CTD software (can 'search' for "CTD" on windows).

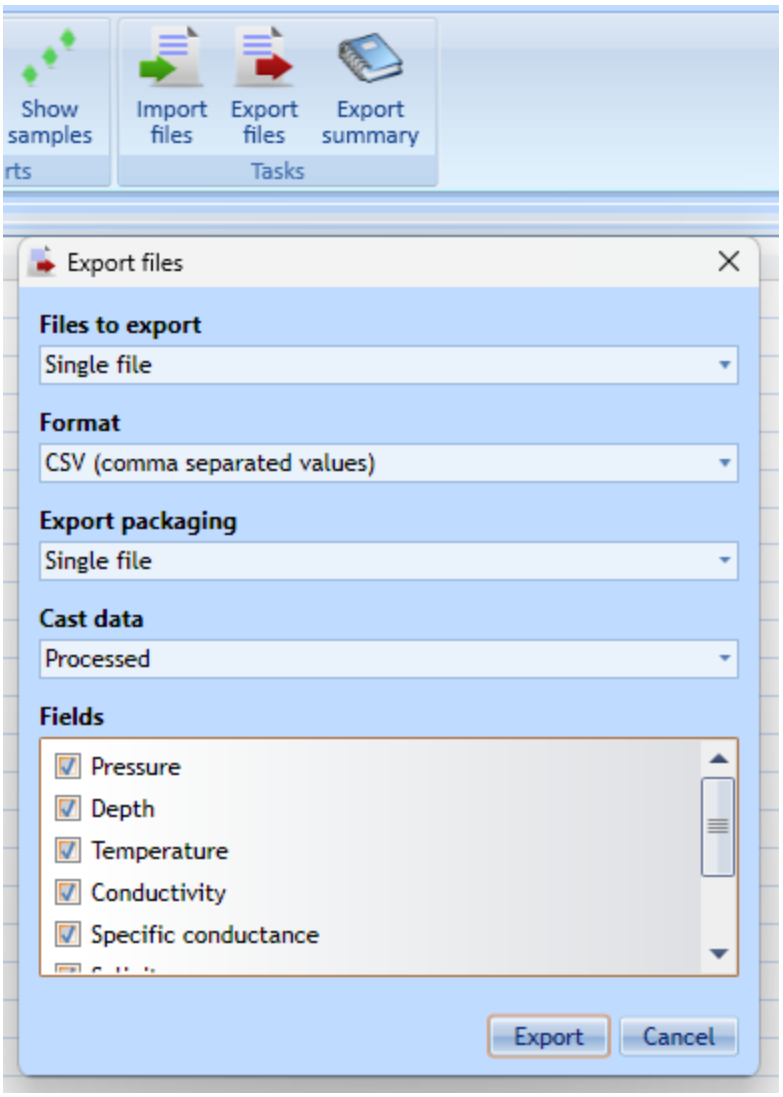
Step 2) Change from the default map page to "Show Files"



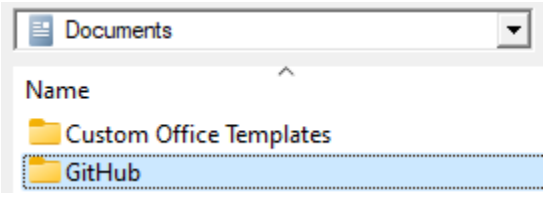
Step 3) Select the most recent cast (ordered chronologically with the newest on top), or find the cast you are interested in by date/time. Make sure that the cast has samples under "Number of samples" and that 'Sample type' is not invalid, which may indicate that you did a point sample cast instead of a regular cast (e.g. the bottom cast in this photo).

Device	Cast time (local)	File name	Number of samples	Location source	Sample type
10K100766	2022-10-23 18:32:53	10K100766_20221023_213253	562	None	Cast
10K100766	2022-10-10 19:22:34	10K100766_20221010_222234	551	GPS	Cast
10K100766	2022-09-25 19:32:22	10K100766_20220925_223222	457	GPS	Cast
10K100766	2022-09-25 19:31:20	10K100766_20220925_223120	0	GPS	Invalid

Step 4) With the cast selected, choose Export Files from the top right taskbar. Make sure format is set to .csv (should be defaulted) and leave all other checkboxes selected.



Step 5) It is probably best to export directly into the folder associated with your survey. The export defaults to Documents, which is where the Github folder already is.



You will need to navigate a bit further to find the exact survey folder.

The exact path you should follow will be: Documents/GitHub/Surveys/ and then pick your Year and Survey (SB# or GB#).

Tagging Reports

Data used to be manually added to TaggingEvents.csv but now can be added by using the Tagging Data.R script. Only add information to the first fields as marked and run the script, no other data needs to be added or modified normally. Lat/Lon is added as degree-min-sec (e.g. 44°16'23) but **written with only numbers and a space between (e.g. "44 16 23")**; the script will convert it to decimal degrees.

```
# Set all tagging log data here
Tag_Num = 123456:123460
Date = "2023-06-28"
Lon = "44 16 23"
Lat = "67 01 16"
Vessel = "Lady Janice" #As written: "Lady Melissa", "Sealife II", "Tasha Marie", "Lady Janice", "Morning Star"
Survey = NA #Survey number for Scots Bay or German Bank when tags were applied, otherwise "NA"
CTD = NA #Add the CTD id only if a TAGGER completed a cast (not the cast by the HSC tech)
```

If Vessel names or associated Taggers change, the below portion can be modified:

```
mutate(Tagger = ifelse(Vessel == "Lady Melissa", "Joseph Nickerson",
                ifelse(Vessel == "Sealife II", "Annik Doucette",
                ifelse(Vessel == "Tasha Marie", "Dale Fitzgerald",
                ifelse(Vessel == "Lady Janice", "Lee Surette",
                ifelse(Vessel == "Morning Star", "Nicholas D'entremont",
```

Map and Region Data Files

After each survey, Jenna processes the acoustic data in Echoview and outputs a "Map.csv" and "Region.csv" file. These need to go into the Github/Survey/Year file specific to the survey. Once these are in place, running the Update RMD Data script (details below) will process this data into tableA+B+C.csv files. Furthermore, **if there was an adhoc school survey conducted** the .csv for it needs to be saved as "adhoc.csv" in the same github/survey/year folder.

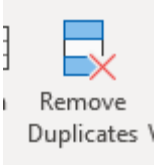
Final Step: Update Data.R

Update Data.R

Once all the above data is updated the Update RMD Data.R script needs to be run. The first options in the code need to be changed to reflect the current survey. 'ids' need to be set to the boat initials for only Scots Bay and only Main Box vessels. Area values need to be updated to match Jenna's values in tableC.csv., and then the script should be run in its entirety.

If for whatever reason you need to re-run this (e.g., you receive some tagger logs late and they need to be added or are troubleshooting errors**) there are three locations you need to manually go in and delete the entry for the current survey, that would have been entered from each time you've run the script.**

1) **Main Data/SSB Estimates.csv** and delete the entry for the current survey and any NA blanks.
2) **Source Data/CTD_Raw.csv** and delete any overlapping CTD entries from the same survey. You can also select the entire dataframe and press Data > Remove Duplicates
3) **Main Data/CTD Full.csv** and delete any overlapping CTD entries from the same survey. You can also select the entire dataframe and press Data > Remove Duplicates

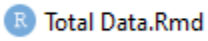



```
# IMPORTANT : SET GROUND, YEAR, AND SURVEY # HERE
surv="GB" #SB or GB
surv2="German Bank" #"German Bank" or "Scots Bay"
year="2023"
surv.no="1"
adhoc = "FALSE" #true or false if an adhoc survey
Sample = "N" #whether ("Y") or not ("N") they caug
Tow = "Y" #whether or not plankton tow(s) were cor

#(SB ONLY) Set main-box vessels
ids = c("FM", "LB", "LJ", "SL")

#Area and TS values
SB1= 661 #SB main area
SB2= 77 #SB north area
SB3= 115 #SB east area

GB1 = 805 #GB main area
GB2 = 267 #Seal Island area
GB3 = NA #Ad-hoc school survey area
```

Finally, the Total Data RMD Report can now be run from the main github.io folder: 

Once all these changes have been made, the steps for Github Desktop can be followed and after a few minutes check the Sharepoint version of the Data Compendium to make sure the data has been uploaded correctly.

Data Dictionary