Data Analysis Tutorial

Follow this step-by-step guide to begin analyzing your adopted float's data.

In this tutorial, we'll use data from a single profile.

Similar steps can be applied to all of the float's data, or to data collected from multiple floats.

Step 1.
Follow the data access tutorial to download your float's data

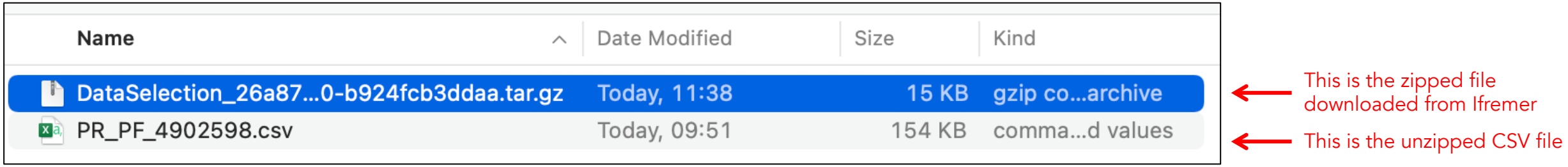
Watch the Data Access tutorial on the Adopt-a-Float "Data Explorer" page to download float data

And check out the the "Introduction to Biogeochemical Properties" lesson.

Step 2.
Download and open data from a single float profile

Follow step 5 in the data access tutorial to download data from a single profile.

Click on the e-mail link to download the data. The data will be packaged in a zipped file (.tar.gz). Unzip the file, and then open the CSV file.



In this tutorial, we will use Microsoft Excel to open and analyze the data. You may also open the data in Google Sheets, for example.

NOTE to teachers: it may take several minutes or hours for the data link to become available. Consider downloading the data ahead of class.

Step 3.
Begin exploring the data in M/S Excel

The data will look something like this:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	PLATFORM_CODE	DATE (YYYY-MM-DDTHH:MI:SSZ)	DATE_QC	LATITUDE (degree_north)	LONGITUDE (degree_east)	POSITION_QC	PRES (decibar)	PRES_QC	PSAL (psu)	PSAL_QC	TEMP (degree_Celsius)	TEMP_QC	DOX2 (micromole/kg)	DOX2_QC	CPHL (milligram/m3)
2	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	-0.2	1	35.094	1	16.996	1			
3	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	0.5	1	35.093	1	17.015	1			
4	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	1.6	1	35.092	1	17.016	1			
5	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	2.6	1	35.093	1	17.016	1			
6	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	3.5	1	35.093	1	17.017	1			
7	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	4.5	1	35.094	1	17.016	1			
8	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	5.5	1	35.093	1	17.017	1			
9	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	6.4	1	35.094	1	17.017	1			
10	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	7.5	1	35.093	1	17.017	1			
11	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	8.6	1	35.093	1	17.018	1			
12	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	9.5	1	35.093	1	17.018	1			
13	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	10.6	1	35.093	1	17.018	1			
14	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	11.5	1	35.093	1	17.019	1			
15	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	12.5	1	35.094	1	17.019	1			
16	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	13.5	1	35.093	1	17.018	1			
17	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	14.4	1	35.094	1	17.018	1			
18	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	15.5	1	35.094	1	17.02	1			
19	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	16.5	1	35.095	1	17.02	1			
20	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	17.5	1	35.094	1	17.02	1			
21	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	18.4	1	35.094	1	17.02	1			
22	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	19.5	1	35.094	1	17.02	1			
23	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	20.4	1	35.094	1	17.02	1			
24	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	21.5	1	35.094	1	17.02	1			
25	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	22.5	1	35.093	1	17.019	1			
26	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	23.4	1	35.093	1	17.019	1			
27	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	24.5	1	35.094	1	17.02	1			
28	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	25.4	1	35.096	1	17.02	1			
29	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	26.4	1	35.096	1	17.019	1			
30	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	27.4	1	35.097	1	17.02	1			
31	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	28.4	1	35.096	1	17.02	1			
32	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	29.5	1	35.095	1	17.02	1			
33	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	30.5	1	35.095	1	17.02	1			

PR_PF_4902598

Step 3 (continued).
Begin exploring the data in M/S Excel

Each column contains a different piece of information. For now, the most relevant columns are:

PRES (decibar), PSAL (psu), TEMP (degree_Celsius), DOX2 (micromole/kg), CPHL (milligram/m3)

The words before the brackets refer to the property, and the values in the brackets are the units of measurement.

PRES = pressure, or depth, measured in decibars. 1 decibar = 1 meter. Low values are near to the ocean surface, large values are deep in the ocean.

PSAL = salinity, or saltiness, measured in practical salinity units (PSU)

TEMP = temperature, measured in degrees Celsius

DOX2 = dissolved oxygen, measured in micromoles per kilogram of water

CPHL = chlorophyll, measured in milligrams per kilogram of water

Step 4. Make a profile plot

A profile represents a series of measurements of an ocean property from the surface (PRES = 0) to the maximum water depth (PRES = 2000 m for BGC-Argo).

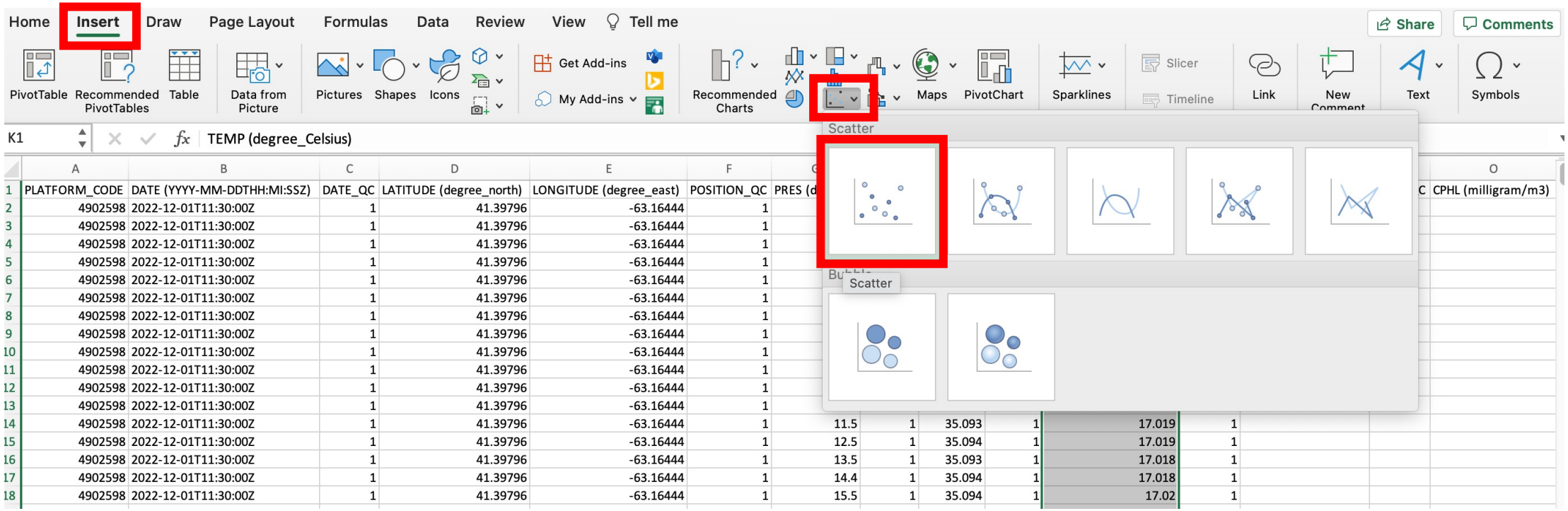
To make a profile plot in Excel, select the data column of the property that you wish to view. Let's start with temperature.

K1 | TEMP (degree_Celsius)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	PLATFORM_CODE	DATE (YYYY-MM-DDTHH:MI:SSZ)	DATE_QC	LATITUDE (degree_north)	LONGITUDE (degree_east)	POSITION_QC	PRES (decibar)	PRES_QC	PSAL (psu)	PSAL_QC	TEMP (degree_Celsius)	TEMP_QC	DOX2 (micromole/kg)	DOX2_QC	CPHL (milligram/m3)
2	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	-0.2	1	35.094	1	16.996	1			
3	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	0.5	1	35.093	1	17.015	1			
4	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	1.6	1	35.092	1	17.016	1			
5	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	2.6	1	35.093	1	17.016	1			
6	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	3.5	1	35.093	1	17.017	1			
7	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	4.5	1	35.094	1	17.016	1			
8	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	5.5	1	35.093	1	17.017	1			
9	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	6.4	1	35.094	1	17.017	1			
10	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	7.5	1	35.093	1	17.017	1			
11	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	8.6	1	35.093	1	17.018	1			
12	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	9.5	1	35.093	1	17.018	1			
13	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	10.6	1	35.093	1	17.018	1			
14	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	11.5	1	35.093	1	17.019	1			

Step 4 (continued). Make a profile plot

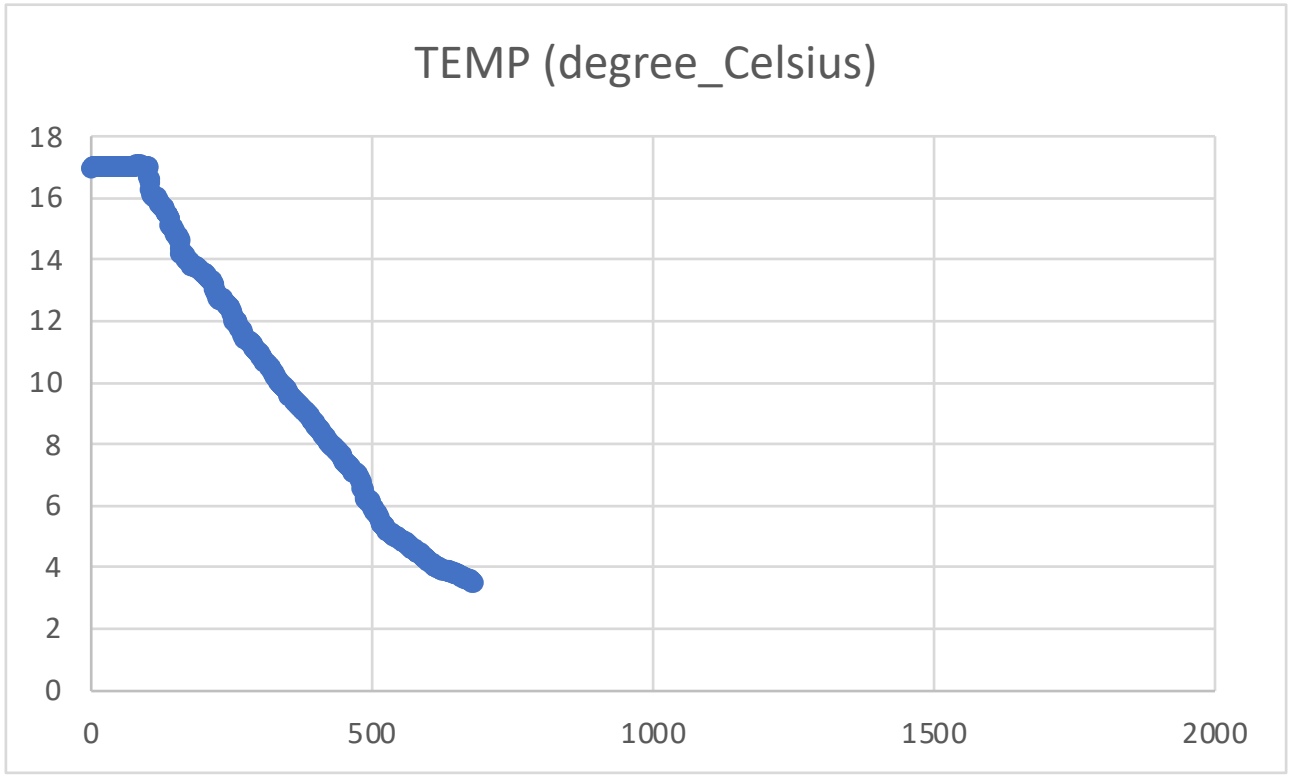
Next, click on “Insert” (located at the top of the Excel menu bar), then “XY Scatter”. Select the first plot option.



	A	B	C	D	E	F	G	H	I	J	K	L	O
1	PLATFORM_CODE	DATE (YYYY-MM-DDTHH:MI:SSZ)	DATE_QC	LATITUDE (degree_north)	LONGITUDE (degree_east)	POSITION_QC	PRES (d						C CPHL (milligram/m3)
2	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1							
3	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1							
4	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1							
5	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1							
6	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1							
7	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1							
8	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1							
9	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1							
10	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1							
11	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1							
12	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1							
13	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1							
14	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	11.5	1	35.093	1	17.019	1	
15	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	12.5	1	35.094	1	17.019	1	
16	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	13.5	1	35.093	1	17.018	1	
17	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	14.4	1	35.094	1	17.018	1	
18	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	15.5	1	35.094	1	17.02	1	

Step 4 (continued). Make a profile plot

A chart, similar to this one, will appear.



But, it's not very useful at this point...

Step 4 (continued). Make a profile plot

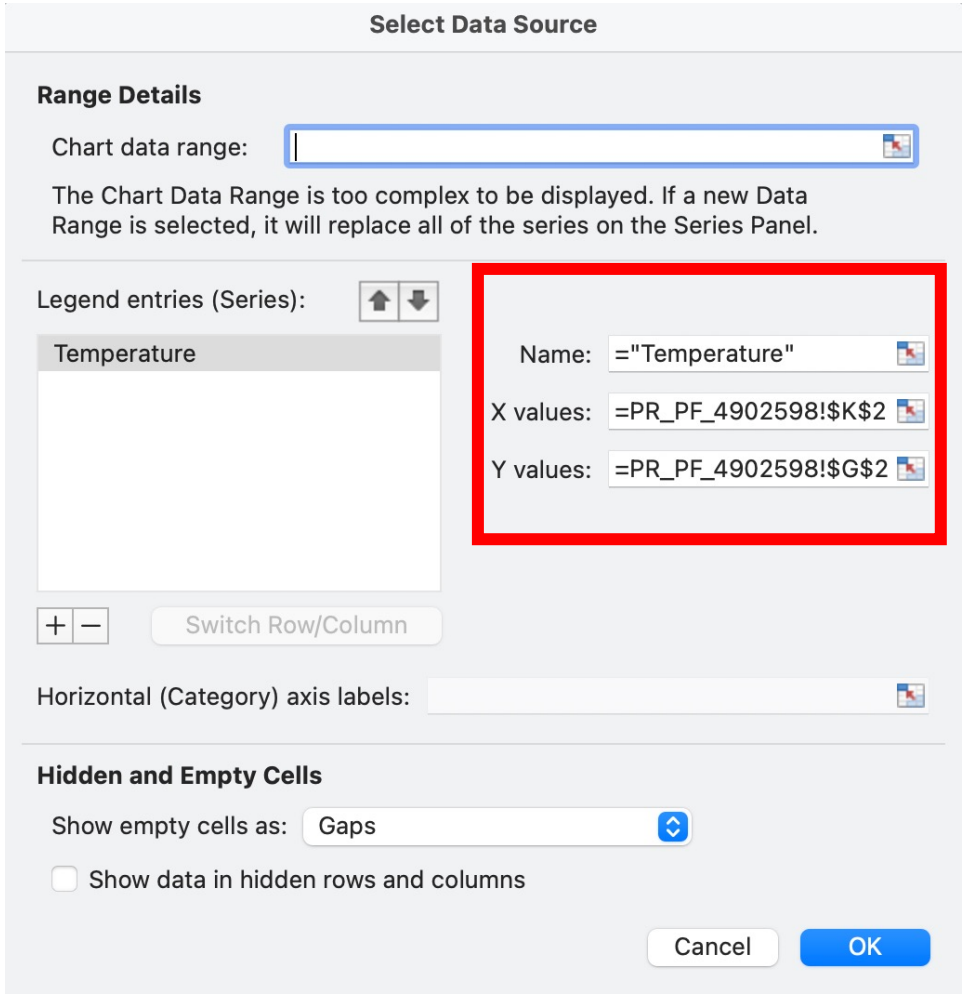
To modify the plot, right-click on the plot and click on “Select Data…”

Set the name to “Temperature”.

Change the X values to the temperature column data.

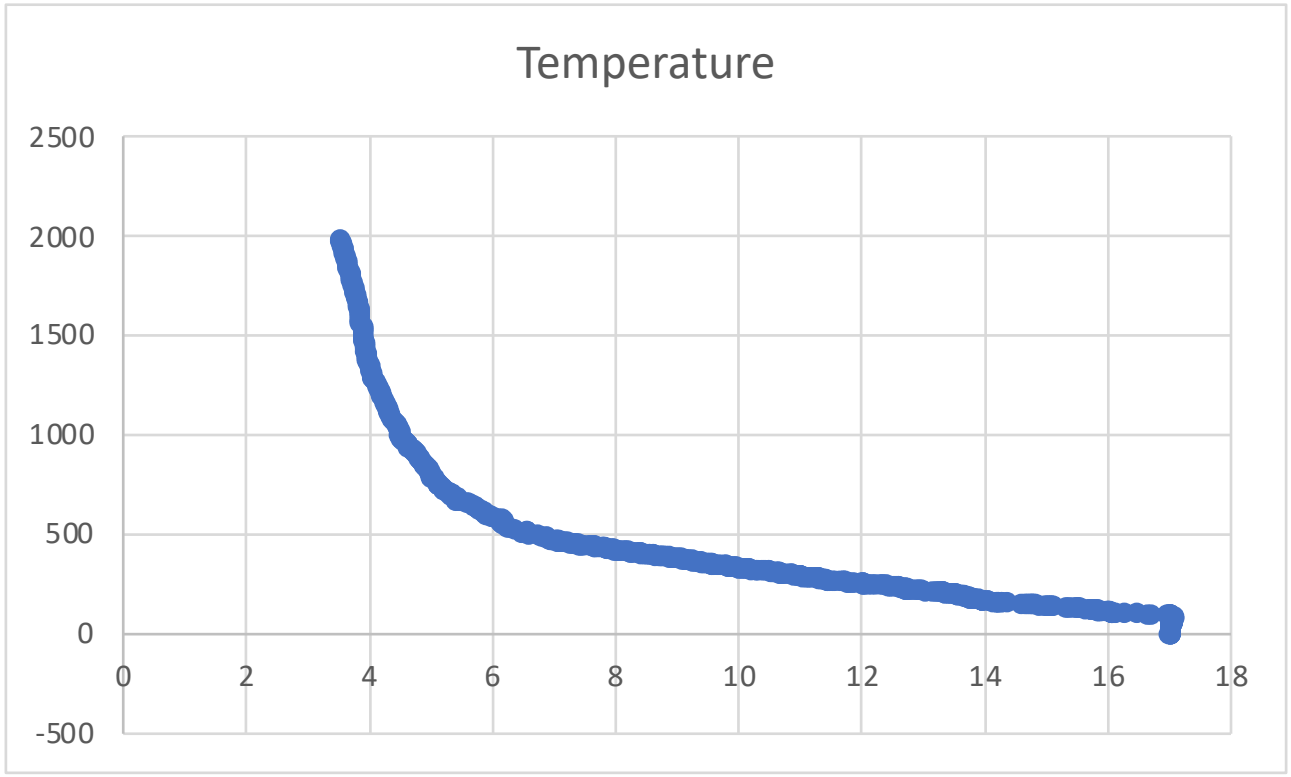
Change the Y values to the pressure column data.

Click “OK”



Step 4 (continued). Make a profile plot

The plot is looking better!



But it's not very intuitive, and it's not clear what the axis values represent.

Step 5. Make the plot intuitive

In oceanography, profiles are typically displayed with the measured property of interest (e.g., temperature) on the horizontal (X) axis, and depth or pressure on the vertical (Y) axis. That's why, in step 4, we set the plot's X values to the temperature column, and the Y values to the pressure column.

Oceanographers display data this way because it's intuitive. Imagine looking at a slice of the ocean from the side. As you look up-and-down that slice, it's like you're moving from shallow to deep depths.

So, the plot we currently have (on the previous page) is not very intuitive, since high depth values are at the top, and low depth values are at the bottom.

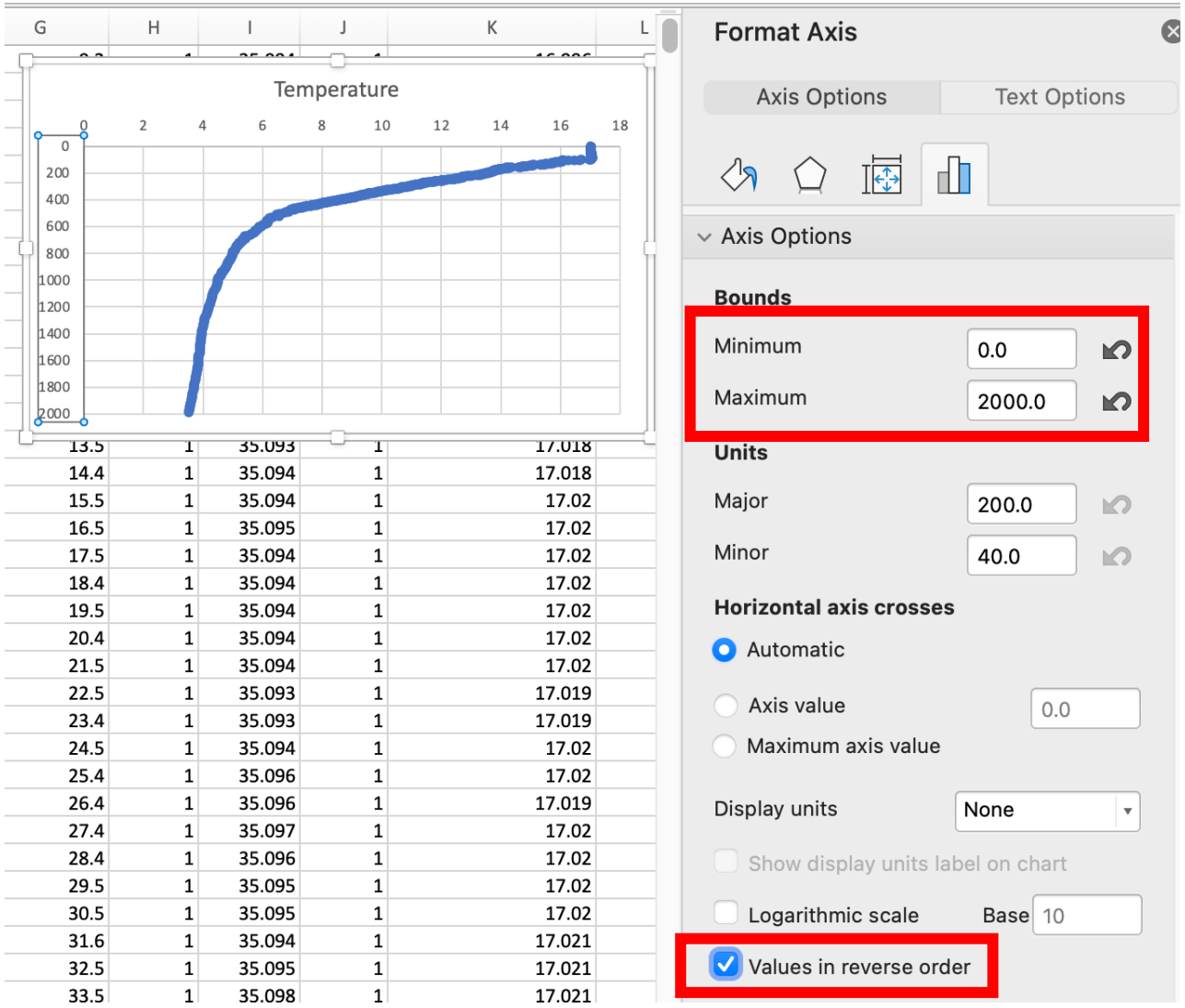
Step 5 (continued). Make the plot intuitive

To make the plot more intuitive, right-click on Y-axis numbers, and select “Format Axis”

Tick the box next to “Values in reverse order” to display shallow values at the top, and deep values at the bottom.

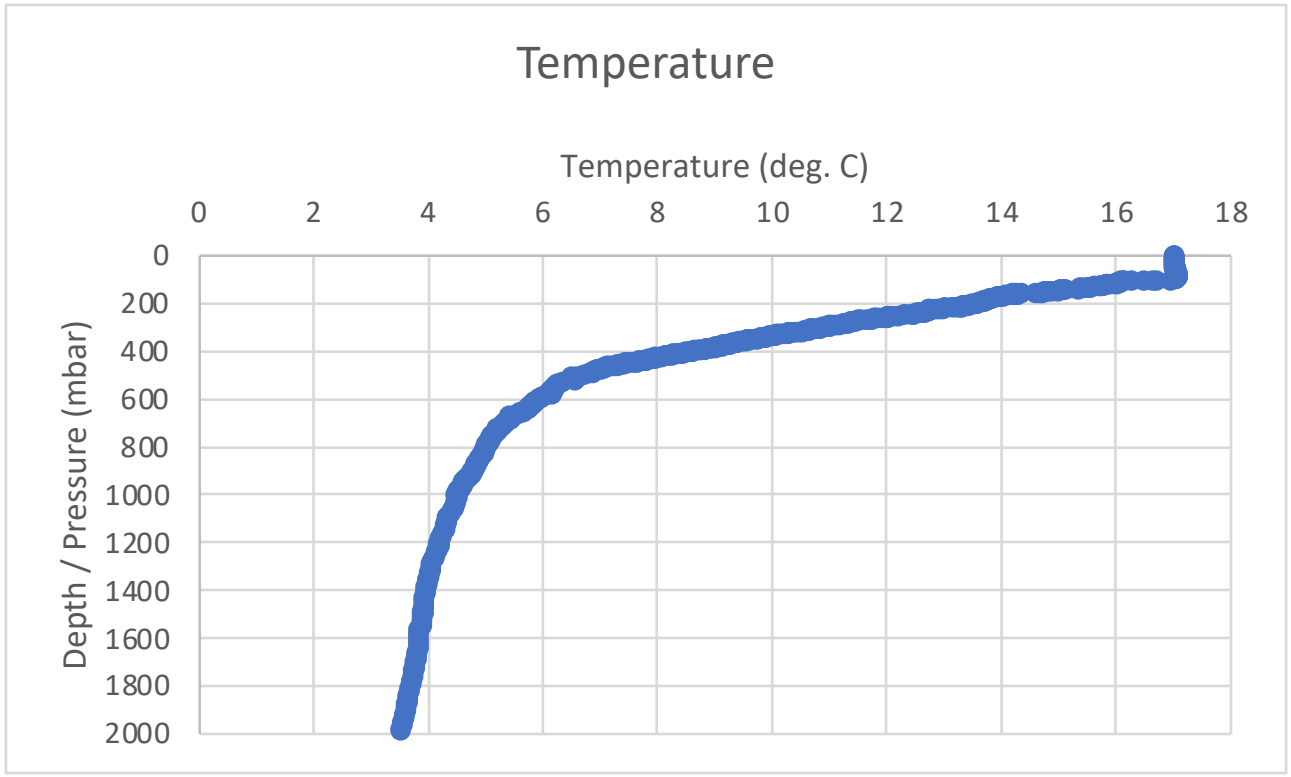
Then, set the minimum and maximum bounds to 0 and 2000 m, respectively.

Now the plot looks more intuitive!



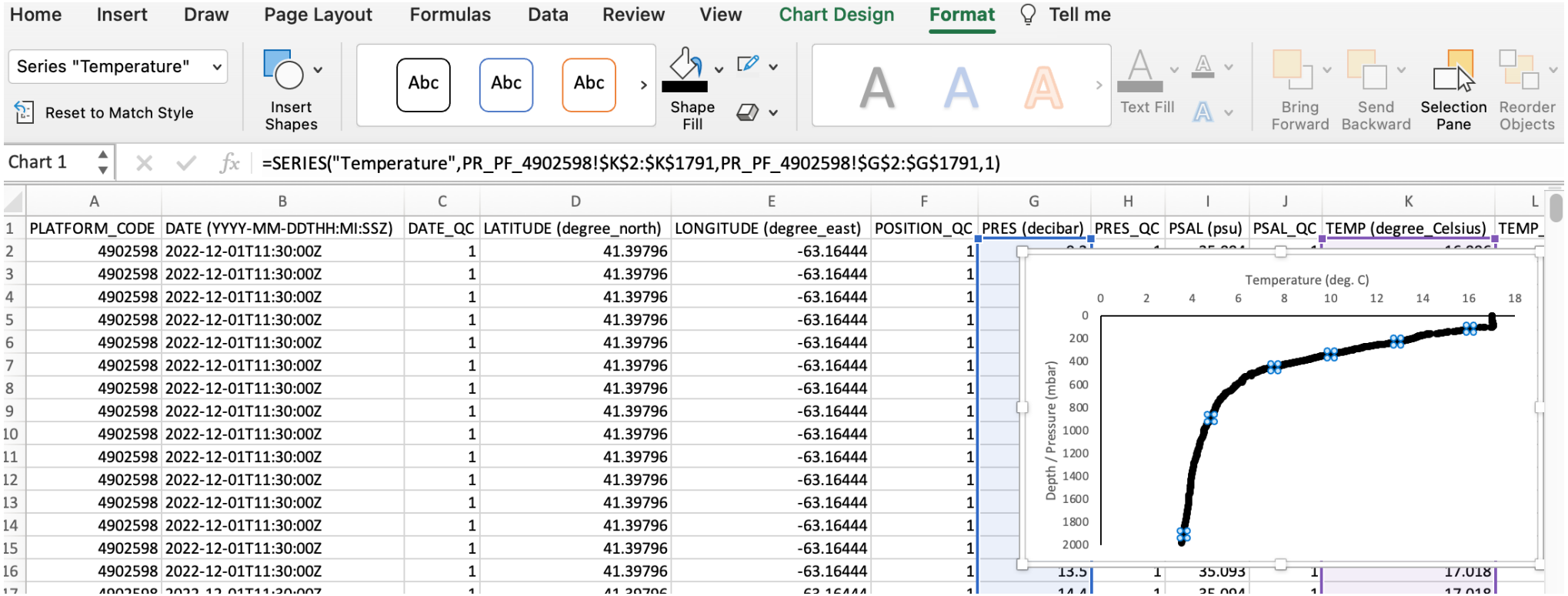
Step 6. Clean up the plot

Let's clean up the plot by adding X- and Y-axis labels. Click on the chart, click "Chart Design" (located at the top of the Excel menu bar), then select "Add Chart Element" to add both horizontal and vertical axis titles. Give the axes appropriate labels and remember to include the units!



Step 6 (continued). Clean up the plot

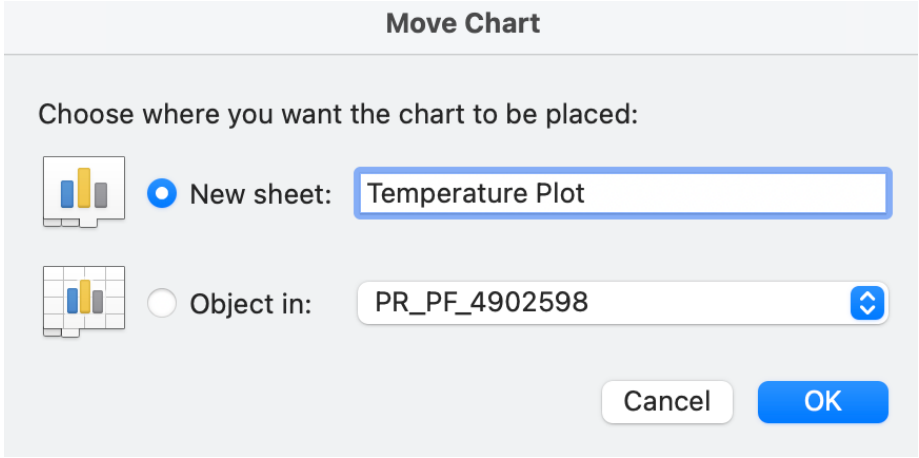
Finally, adjust the plot appearance by clicking on different elements of the chart, and changing various options under the "Format" panel.



	A	B	C	D	E	F	G	H	I	J	K	L
1	PLATFORM_CODE	DATE (YYYY-MM-DDTHH:MI:SSZ)	DATE_QC	LATITUDE (degree_north)	LONGITUDE (degree_east)	POSITION_QC	PRES (decibar)	PRES_QC	PSAL (psu)	PSAL_QC	TEMP (degree_Celsius)	TEMP_
2	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1						
3	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1						
4	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1						
5	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1						
6	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1						
7	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1						
8	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1						
9	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1						
10	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1						
11	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1						
12	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1						
13	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1						
14	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1						
15	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1						
16	4902598	2022-12-01T11:30:00Z	1	41.39796	-63.16444	1	13.5	1	35.093	1	17.018	

Step 7. Move the chart to a new tab

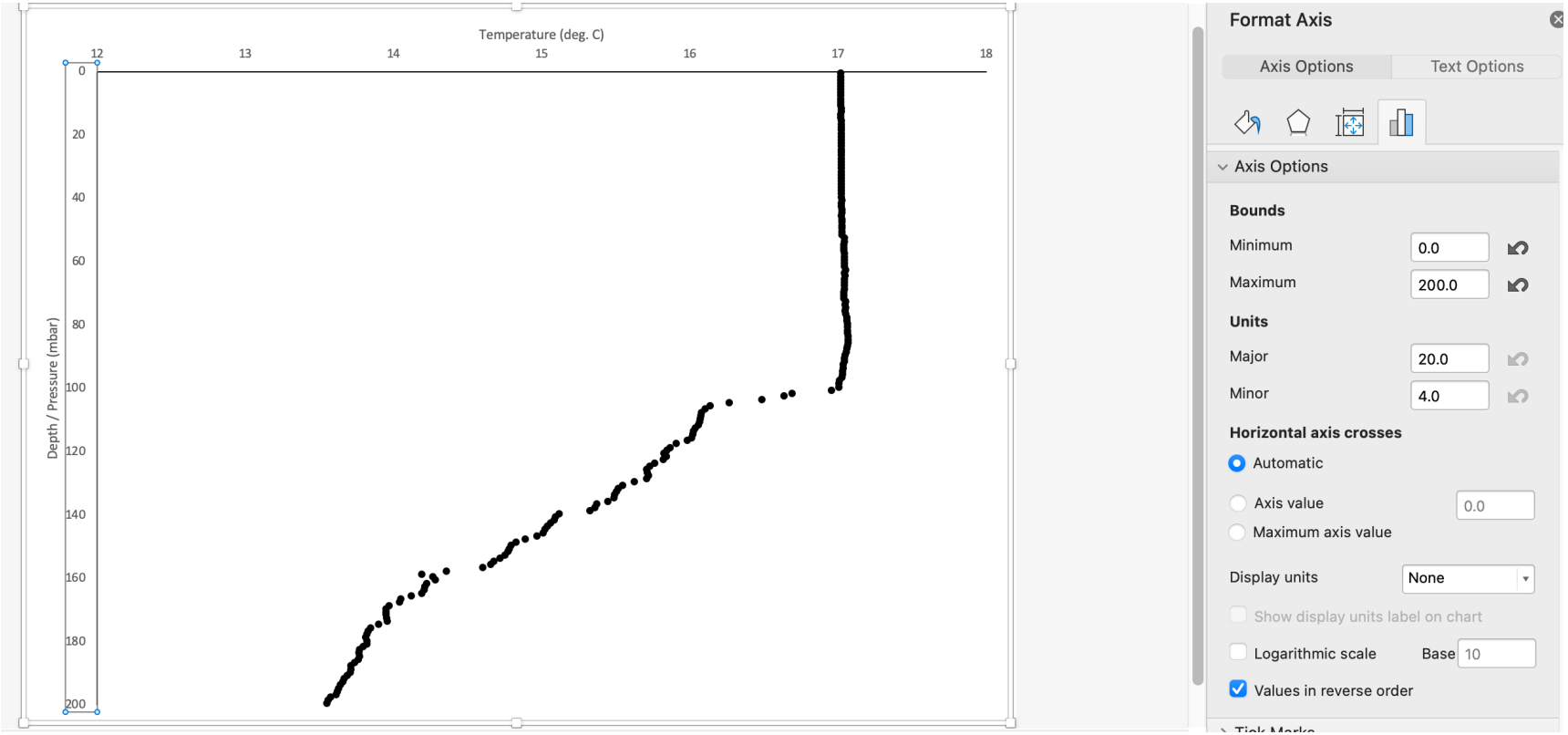
Move the chart to a new tab in the Excel workbook by right clicking on the chart, and selecting "Move Chart." Select "New Sheet", and give the sheet an intuitive name, like "Temperature plot"



A new tab will appear at the bottom of the sheet.

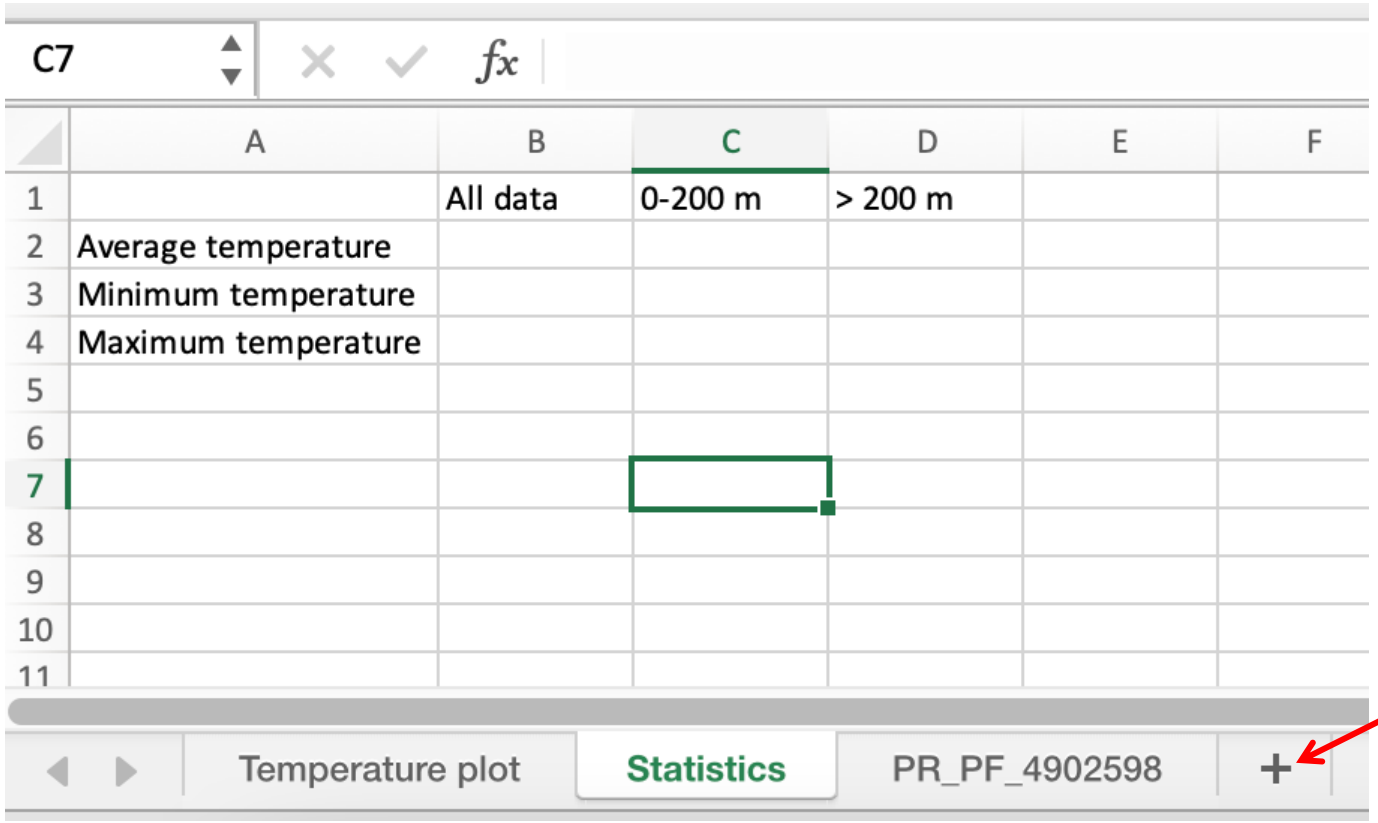
Step 8. Explore the data in the plot

Now we have a plot that is intuitive, and clean! You can explore the data in a number of ways. One is by adjusting the Y-axis range in the "Format Axis" panel. You can, for example, change the values here to display data from the upper 200 m only.



Step 9. Explore the data statistics

You can also explore the data using basic statistics. Let's start by creating a new tab in the worksheet, called "Statistics", with the following contents:



Step 9 (continued). Explore the data statistics

Now, we can determine what the average, minimum and maximum temperature values are within different parts of the profile. The following functions will be useful:

	All data	0-200 m	> 200 m
Average temperature	=AVERAGE(PR_PF_4902598!K:K)	=AVERAGEIF(PR_PF_4902598!G:G,"<=200",PR_PF_4902598!K:K)	=AVERAGEIF(PR_PF_4902598!G:G,">200",PR_PF_4902598!K:K)
Minimum temperature	=MIN(PR_PF_4902598!K:K)	=MINIFS(PR_PF_4902598!K:K,PR_PF_4902598!G:G,"<=200")	=MINIFS(PR_PF_4902598!K:K,PR_PF_4902598!G:G,">200")
Maximum temperature	=MAX(PR_PF_4902598!K:K)	=MAXIFS(PR_PF_4902598!K:K,PR_PF_4902598!G:G,"<=200")	=MAXIFS(PR_PF_4902598!K:K,PR_PF_4902598!G:G,">200")

Calculate the average, minimum and maximum values from the temperature data column (column K on the PR_PF_4902598 sheet)

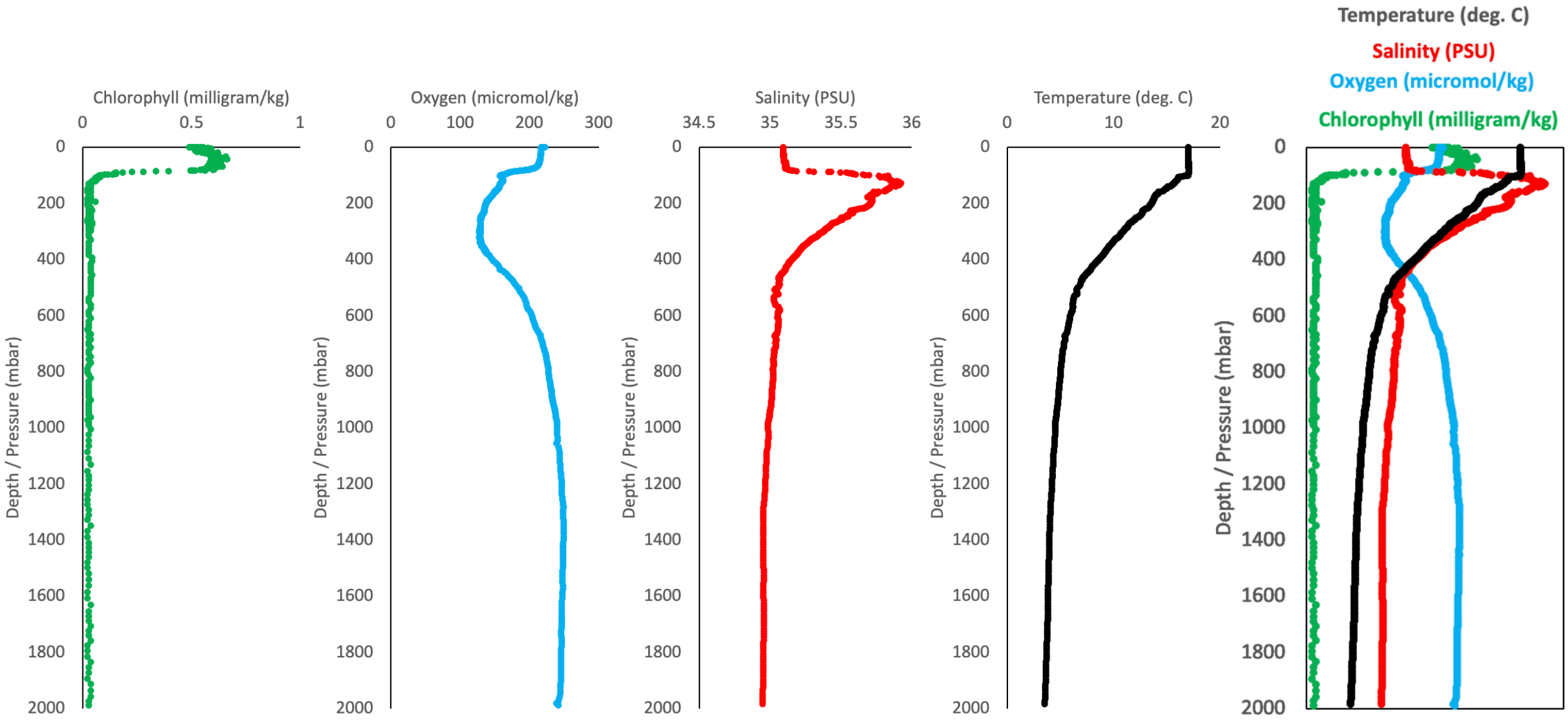
Calculate the average, minimum and maximum values when the pressure (depth) is less than or equal to 200. The "<=200" condition checks to see when values in column G (pressure) are below 200

Calculate the average, minimum and maximum values when the pressure (depth) is greater than 200.

The answer to those equations is as follows:

	All data	0-200 m	> 200 m
Average temperature	10.19	15.94	7.77
Minimum temperature	3.52	13.55	3.52
Maximum temperature	17.07	17.07	13.54

Step 10. Repeat steps 4-9 for other properties!



Step 11. Save your progress

Up until now, we have been working in a CSV document. However, CSV files are not compatible with charts and multiple sheets.

To make sure that you don't lose your progress, save your worksheet as an .xlsx document. Click "File", "Save As". Change the file format to Excel Workboox (.xlsx), and save the file. Consider giving the file an intuitive name, to identify the profile number, date, and float name.

You can also save the profile plots by right clicking on a plot, and selecting "Save As Picture…"

Step 12. Try it for yourself!

Try these exercises for yourself, using your own float data, and repeat the steps using all of the float's data (follow step 6 in the Data Access Tutorial to download all of the float's data), or data from multiple floats (step 7).

Step 13. Try using alternative software tools.

For more advanced users, check out these software tools (www.go-bgc.org/getting-started-with-go-bgc-data – scroll down to "BGC Float Toolbox") designed to help you explore your data using *Python*, *R*, or Matlab computer languages.