Loading files (Colab, Google Drive), loading data (readlines, numpy), and an intro to plotting (matplotlib)

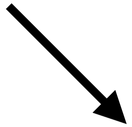
ESS 116 | Fall 2024

Prof. Henri Drake, Prof. Jane Baldwin, and Prof. Michael Pritchard

(Modified from Ethan Campbell and Katy Christensen's materials for UW's Ocean 215)

Importing matplotlib

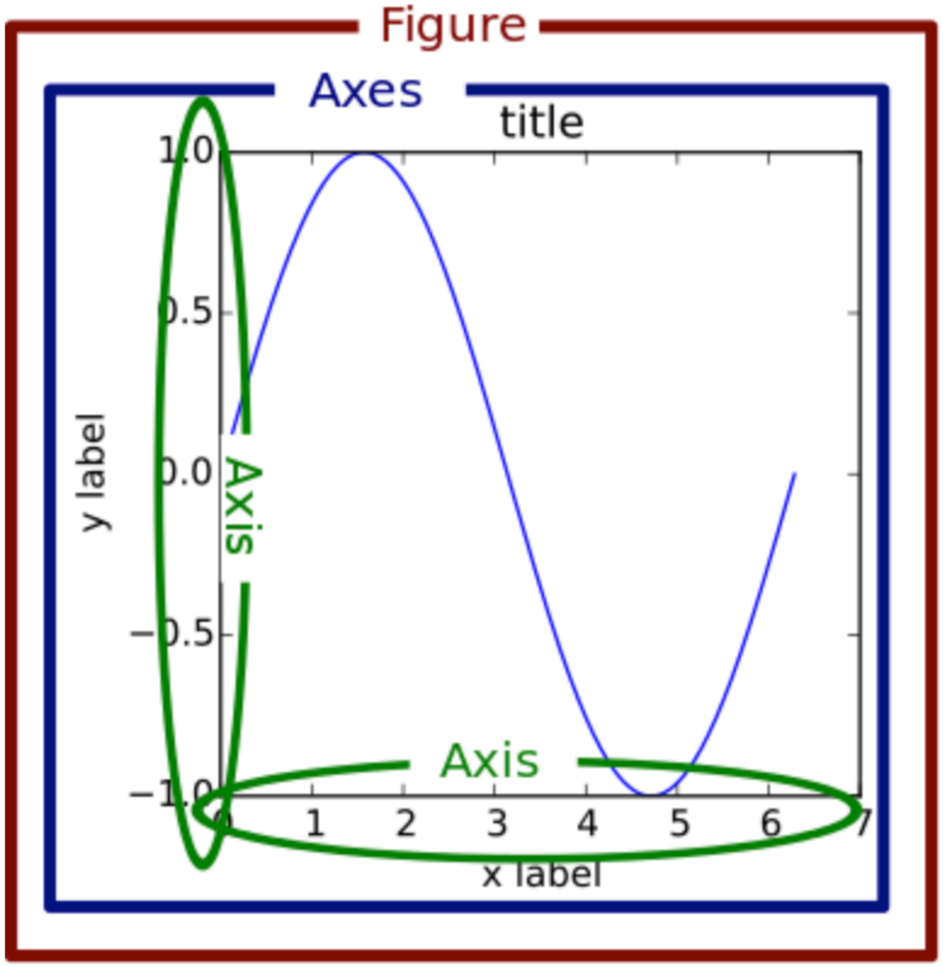
This is a shortcut;
you can choose any name
but `plt` is most common



```
import matplotlib.pyplot as plt
```

This part is
technically optional

Matplotlib objects



https://realpython.com/python-matplotlib-guide/

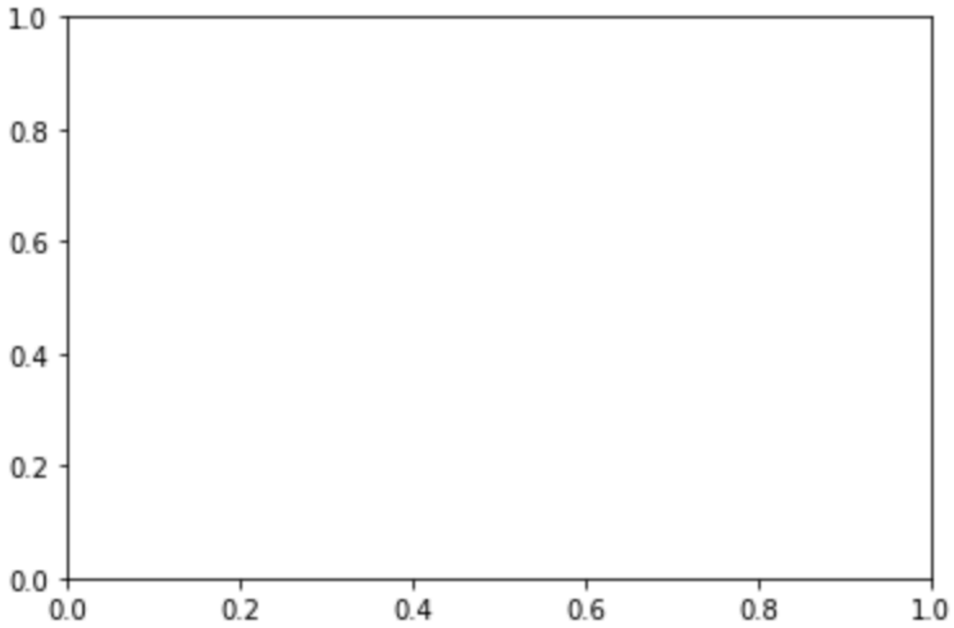
Main matplotlib objects:

1) Figure: this is outer container for plotting
2) Axes: this is an individual graph
3) Axis (and smaller…): these are the small formatting to refine your plot

Creating figures

Creating a figure with a blank axes object:

```
1 import matplotlib.pyplot as plt
2 fig,ax = plt.subplots()
```



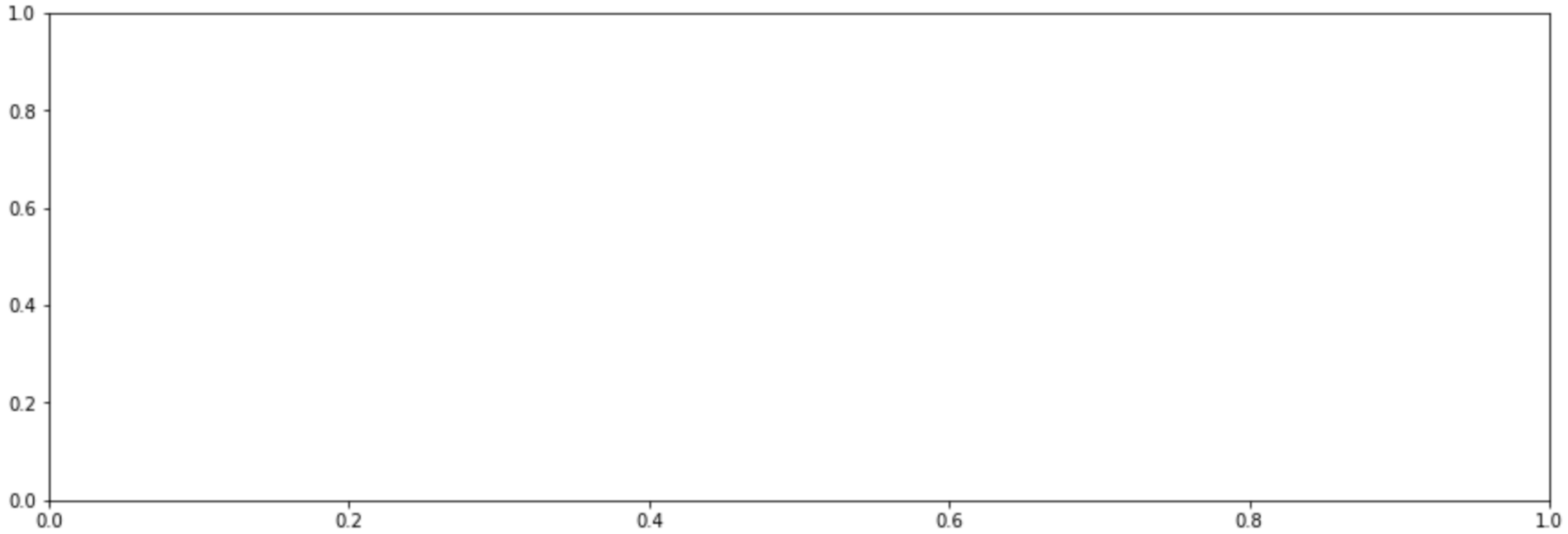
These become the variable names for the figure and axes objects, respectively.

Creating figures

Creating a figure with a blank axes object of custom size:

(width, height) in inches

```
1 import matplotlib.pyplot as plt
2 fig,ax = plt.subplots(figsize=(15,5))
```

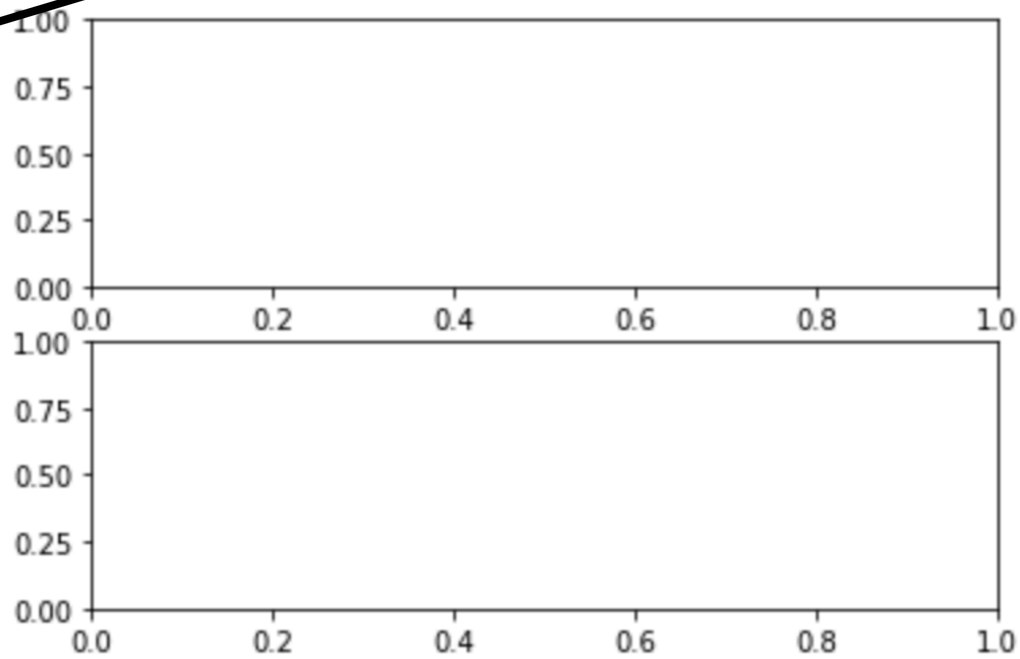


Creating figures

Creating a figure with multiple axes objects:

```
1 import matplotlib.pyplot as plt
2 fig,(ax0,ax1) = plt.subplots(nrows=2, ncols=1)
```

This is so that each axes has a variable name



ax0

ax1

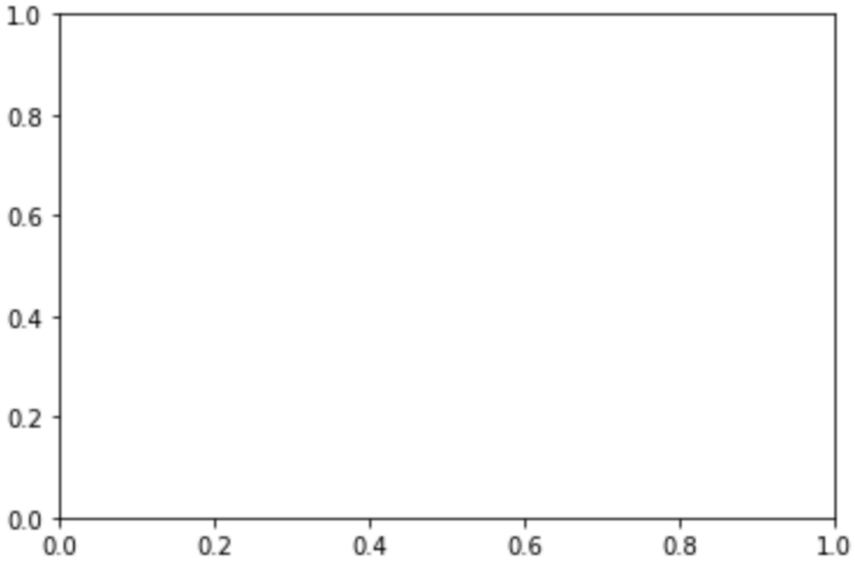
Simple line plot

Our data:

```
1 import numpy as np
2 filepath = 'drive/My Drive/Data_folder/Seattle_tides_predicted_20201001_20201024.txt'
3
4 data = np.genfromtxt(filepath,skip_header=14,dtype=float,usecols=3,delimiter=None)
5
6 time = np.linspace(0,len(data)/10,len(data)) # 6 min freq. so len(data)/10 = # of hours
7
```

Start by creating a figure with an empty axes object:

```
1 import matplotlib.pyplot as plt
2 fig,ax = plt.subplots()
3
```



Simple line plot

Our data:

```
1 import numpy as np
2 filepath = 'drive/My Drive/Data_folder/Seattle_tides_predicted_20201001_20201024.txt'
3
4 data = np.genfromtxt(filepath,skip_header=14,dtype=float,usecols=3,delimiter=None)
5
6 time = np.linspace(0,len(data)/10,len(data)) # 6 min freq. so len(data)/10 = # of hours
7
```

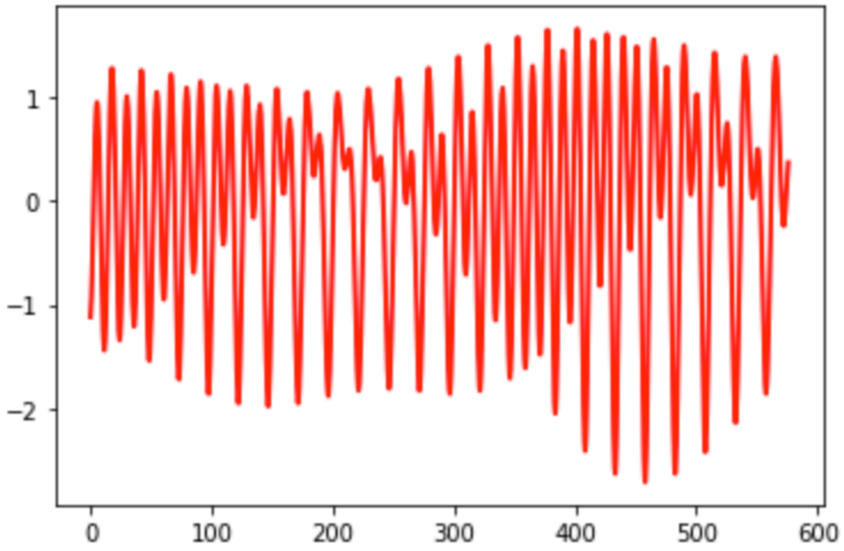
Plot our data on the axis object:

```
1 import matplotlib.pyplot as plt
2 fig,ax = plt.subplots()
3
4 ax.plot(time, data, c='r',linestyle='-', linewidth=2, marker=None)
```

x-axis, y-axis

(c=color)

These are optional arguments, but they make the figure more appealing.



Simple line plot

Our data:

```
import numpy as np
filepath = 'drive/My Drive/Data_folder/Seattle_tides_predicted_20201001_20201024.txt'

data = np.genfromtxt(filepath,skip_header=14,dtype=float,usecols=3,delimiter=None)

time = np.linspace(0,len(data)/10,len(data)) # 6 min freq. so len(data)/10 = # of hours

```

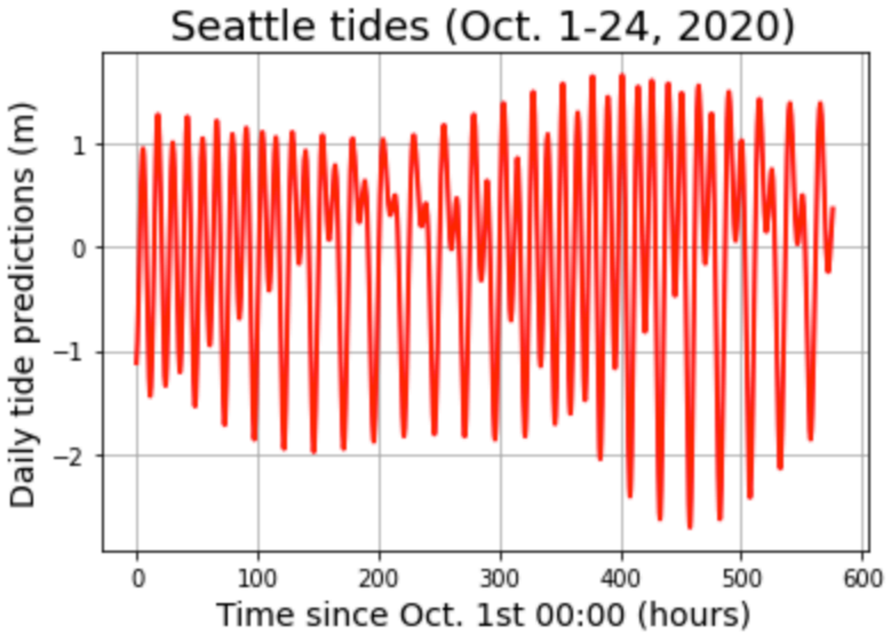
Create a title, labels, and figure formatting:

```
import matplotlib.pyplot as plt
fig,ax = plt.subplots()

ax.plot(time, data, c='r',linestyle='-', linewidth=2, marker=None)


ax.grid()
ax.set_title('Seattle tides (Oct. 1-24, 2020)', fontsize=18)
ax.set_xlabel('Time since Oct. 1st 00:00 (hours)', fontsize=14)
ax.set_ylabel('Daily tide predictions (m)', fontsize=14)

```



Scatter plot

These are changeable, and won't affect where the dots are on the plot.

```
plt.scatter( x, y, size= , color= , alpha=)
```

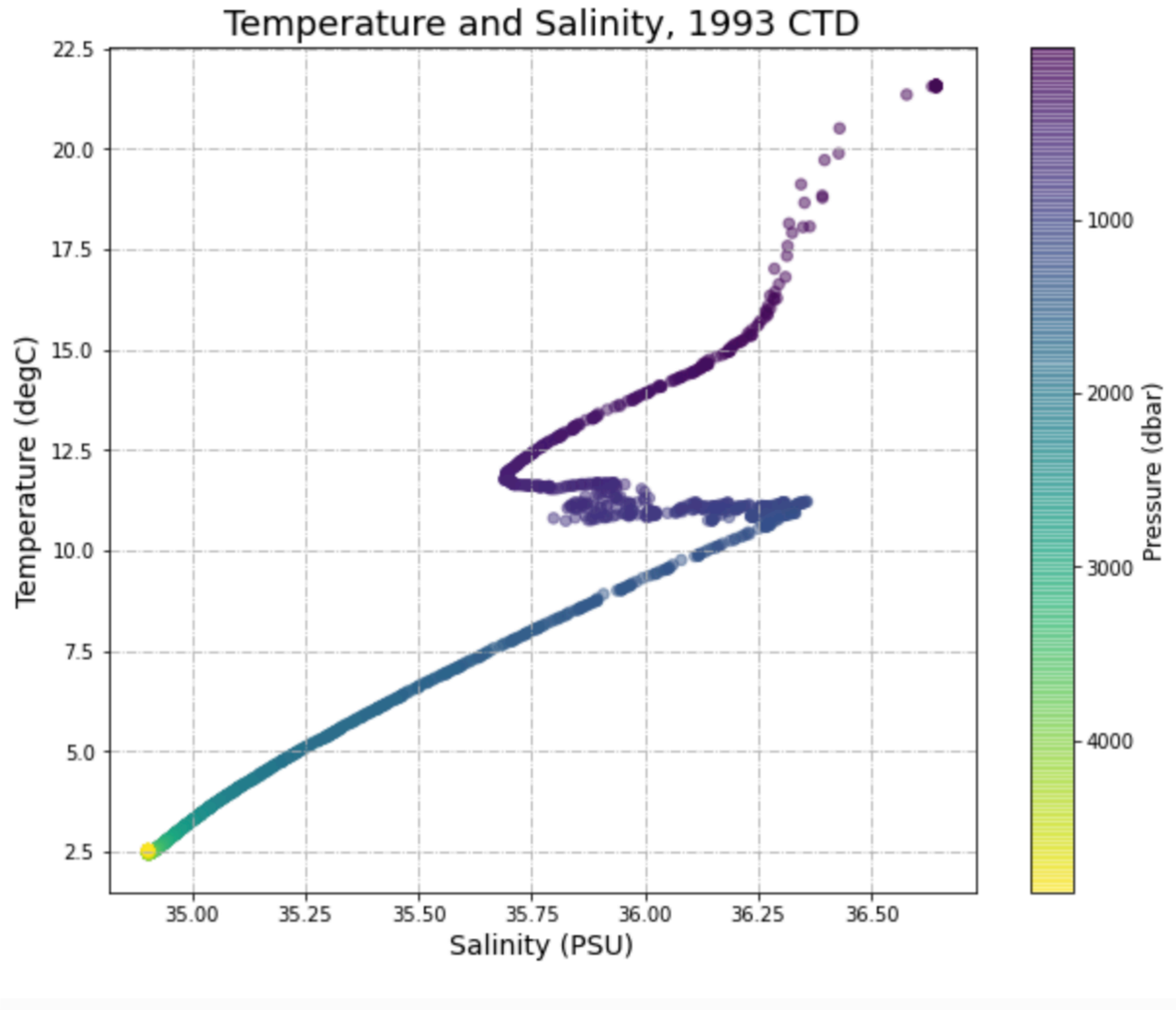
Salinity

Temperature

Constant

Pressure

Transparency



Scatter plot

Example data: CTD data from 1993 WOCE


a03_00011_1993CTD_data.csv

```
1 import matplotlib.pyplot as plt
2 import numpy as np
3 from google.colab import drive
4
5 drive.mount('/content/drive')
```

```
1 filepath = 'drive/My Drive/Data_folder/a03_00011_1993CTD_data.csv'
2
3 file_obj = open(filepath, 'r')
4
5 for index in range(90):
6   line = file_obj.readline()
7   print(line)
8
9 file_obj.close()
10
```

```
#Software Version: CTD_Exchange_Encode_v1.0g (Diggs)
#SUMFILE_NAME:      a03su.txt
#SUMFILE_MOD_DATE: Tue Feb 17 08:35:04 2004
#CTDFILE_NAME:      CT40D011.WCT
#CTDFILE_MOD_DATE: Tue Feb 17 07:55:02 2004
#DEPTH_TYPE       : COR
#EVENT_CODE       : BO
NUMBER_HEADERS = 10
EXPOCODE = 90CT40_1
SECT = A03
STNNBR =    11
CASTNO =     1
DATE = 19930925
TIME =  0312
LATITUDE =   36.2247
LONGITUDE =  -10.4520
DEPTH =  4842
CTDPRS,CTDPRS_FLAG_W,CTDTMP,CTDTMP_FLAG_W,CTDSAL,CTDSAL
DBAR,,ITS-90,,PSS-78,,UMOL/KG,,
```

```
 1.0,2,  21.5315,2,  36.6439,2,  -999.0,9
 3.0,2,  21.5861,2,  36.6421,2,  -999.0,9
 5.0,2,  21.5689,2,  36.6417,2,  -999.0,9
 7.0,2,  21.5673,2,  36.6426,2,  -999.0,9
 9.0,2,  21.5698,2,  36.6438,2,  -999.0,9
11.0,2,  21.5761,2,  36.6436,2,  -999.0,9
13.0,2,  21.5770,2,  36.6439,2,  -999.0,9
15.0,2,  21.5773,2,  36.6443,2,  -999.0,9
17.0,2,  21.5771,2,  36.6438,2,  -999.0,9
19.0,2,  21.5771,2,  36.6436,2,  -999.0,9
21.0,2,  21.5771,2,  36.6441,2,  -999.0,9
23.0,2,  21.5776,2,  36.6436,2,  -999.0,9
25.0,2,  21.5790,2,  36.6439,2,  -999.0,9
27.0,2,  21.5793,2,  36.6435,2,  -999.0,9
29.0,2,  21.5793,2,  36.6434,2,  -999.0,9
31.0,2,  21.5784,2,  36.6432,2,  -999.0,9
33.0,2,  21.5759,2,  36.6428,2,  -999.0,9
...
4877.0,2,   2.5475,2,  34.9021,2,  -999.0,9
END_DATA
```

Scatter plot

Loading data:

```
filepath = 'drive/My Drive/Data_folder/a03_00011_1993CTD_data.csv'

# Load the data
data = np.genfromtxt(filepath,skip_header=20,skip_footer=1,delimiter=',',usecols=(0,2,4))

# Separate out the columns into individual variables
P = data[:,0]
T = data[:,1]
S = data[:,2]

```

Plotting:

```
filepath = 'drive/My Drive/Data_folder/a03_00011_1993CTD_data.csv'

# Load the data
data = np.genfromtxt(filepath,skip_header=20,skip_footer=1,delimiter=',',usecols=(0,2,4))

# Separate out the columns into individual variables
P = data[:,0]
T = data[:,1]
S = data[:,2]

# Create the figure and scatter the data
fig,ax = plt.subplots(figsize=(10,8))
scpl = ax.scatter(S, T, s=30, c=P, alpha=0.5)

# Format the figure
ax.set_title('Temperature and Salinity, 1993 CTD', fontsize=18)
ax.set_ylabel('Temperature (degC)', fontsize=14)
ax.set_xlabel('Salinity (PSU)', fontsize=14)
ax.grid(linestyle='-.')
c = fig.colorbar(scpl,ax=ax)
c.set_label('Pressure (dbar)',fontsize=12)
```

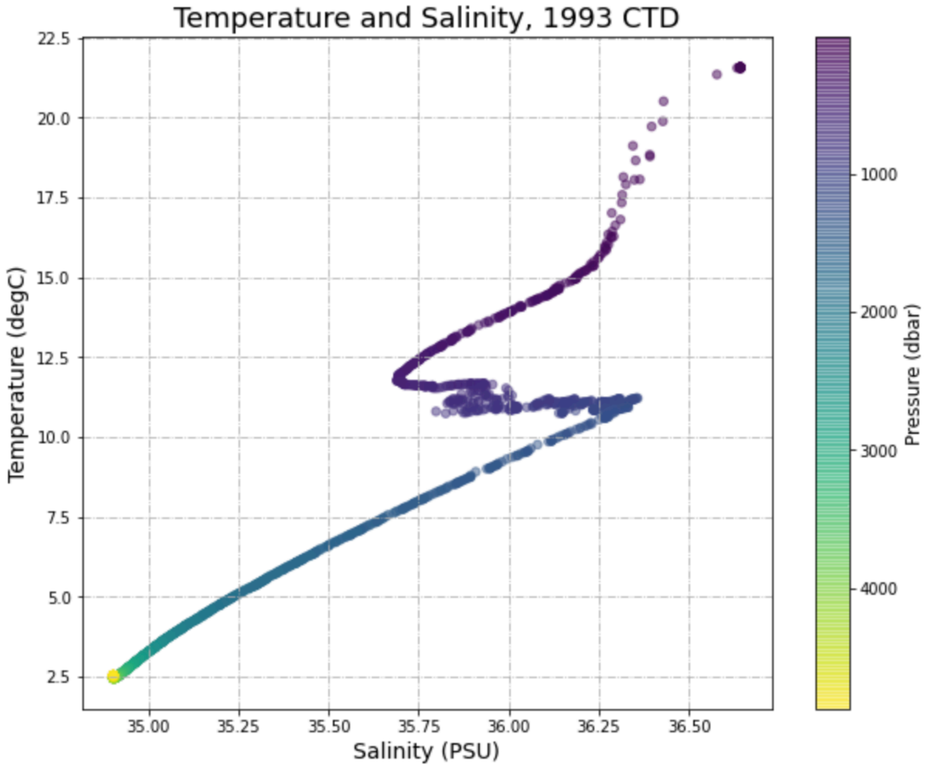
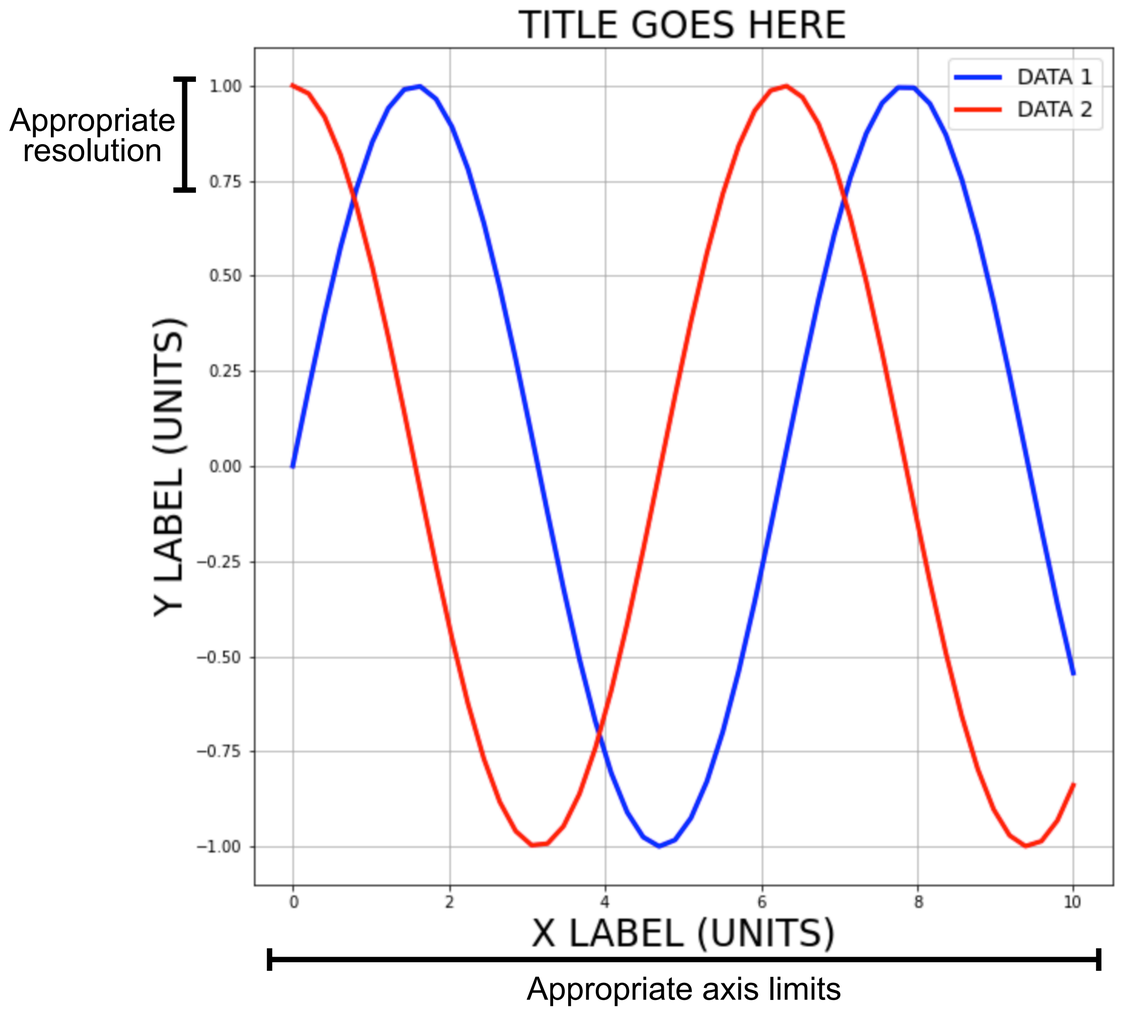
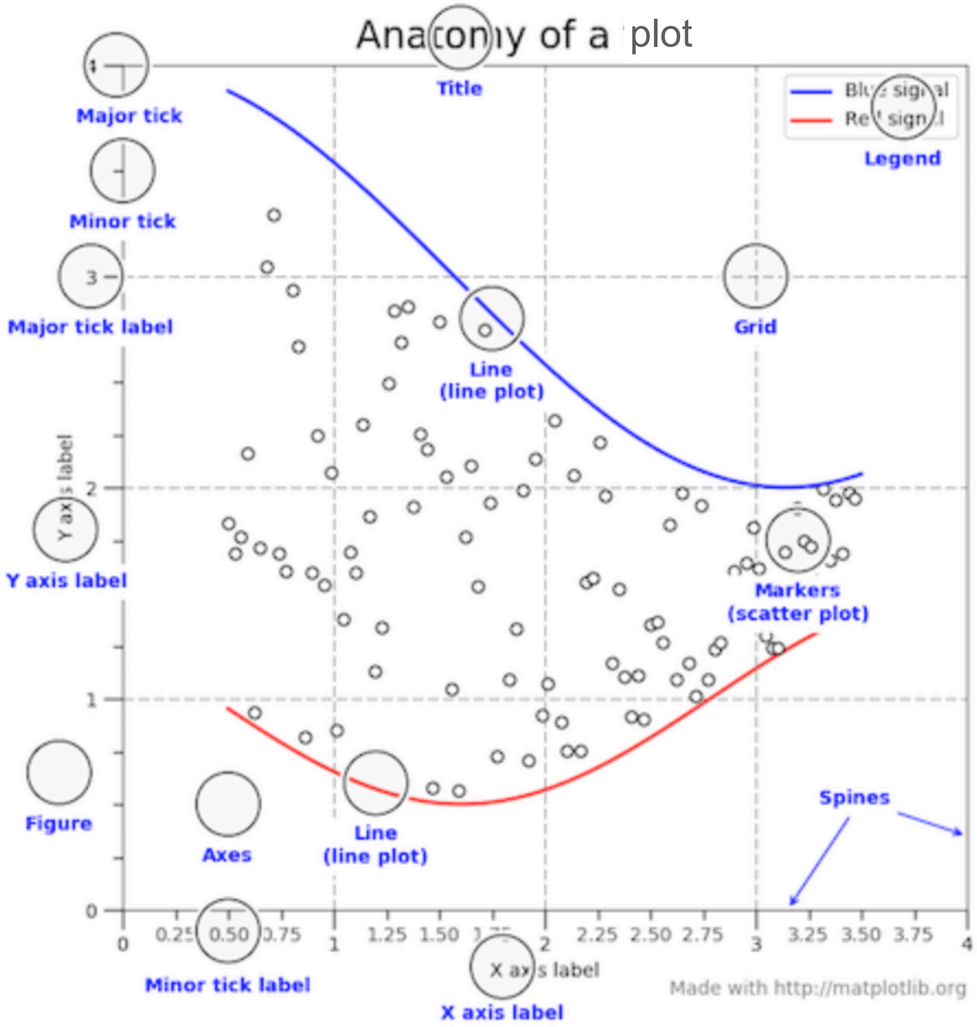


Figure requirements for this course

1) Title
2) Axis labels (with units, when possible)
3) Appropriate axis limits (e.g. max/min)
4) Appropriate tick resolution
5) Legend for different datasets, when applicable
6) Large enough fontsizes



Everything is customizable when plotting



https://realpython.com/python-matplotlib-guide/

You can change anything in a plot if you know how.

You can usually find how to do something by searching the documentation or searching the internet.

Official matplotlib documentation:

https://matplotlib.org/3.3.2/index.html

Resources

Official matplotlib documentation:

https://matplotlib.org/3.3.2/index.html