# netCDF

# We are going to make an animated gif and not embarass a friend!

You need to take your time. Every cell is important today. If you rush you will get yourself in trouble.

```
In [2]:
         1 # The usual libraries
         2 import pandas as pd
         3 import numpy as np
         4 import matplotlib.pylab as plt
         5
         6 from scipy import stats
           from matplotlib.backends.backend pdf import PdfPages
         7
         8
           %matplotlib inline
         9
        10
        11 # The new ones
        12 import netCDF4
        13 import cartopy.crs as ccrs
        14 import cartopy.feature
        15 from cartopy.util import add_cyclic_point
        16 #import contextily as ctx
        17
        18 import datetime
```

You are going to read in a netcdf file. Plot sea surface temperatures. Animate them for the year and turn in an animated GIF.

The links to get the sea surface data don't always work. Be patient. WE DOWNLOADED THIS YESTERDAY.

- Start here <u>http://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.highres.html#detail</u> (<u>http://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.highres.html#detail</u>)
- · Every year this changes.
- Now I go into the box on the lower left.
- Scroll down to sea surface temperature man with 52. Click on it.
- then click a year and it will download. Do not do this year! It is not a full year because we aren't done yet!
- It is a 400mb file. It might be slow if you are on wifi
- Now make sure it is saved in your working directory and we can open it. url is the file name.
- It is just like the last packet!

# In [3]: 1 file=('sst.day.mean.2019.nc') 2 f=netCDF4.Dataset(file)

You saved the file into "f". So if we print f we will see part of it. But remember the f is not the whole file but more like a function that we can then call to get at the whole file. We will use "f." notation to learn about the file. You know the drill!

In [3]:	1 print (f)											
	<pre><class 'netcdf4netcdf4.dataset'=""></class></pre>											
	root group (NETCDF4_CLASSIC data model, file format HDF5):											
	Conventions: CF-1.5 title: NOAA/NCEI 1/4 Degree Daily Optimum Interpolation Sea Surface Temperatur											
	e (OISST) Analysis, Version 2.1											
	institution: NOAA/National Centers for Environmental Information											
	<pre>source: NOAA/NCEI https://www.ncei.noaa.gov/data/sea-surface-temperature-optim</pre>											
	<pre>um-interpolation/v2.1/access/avhrr/ (https://www.ncei.noaa.gov/data/sea-surface-te</pre>											
	<pre>mperature-optimum-interpolation/v2.1/access/avhrr/) Deferences https://www.acl.access/avhrr/)</pre>											
	<b>References:</b> https://www.psl.noaa.gov/data/gridded/data.noaa.oisst.v2.highres.h tml (https://www.psl.noaa.gov/data/gridded/data.noaa.oisst.v2.highres.html)											
	dataset_title: NOAA Daily Optimum Interpolation Sea Surface Temperature											
	version: Version 2.1											
	comment: Reynolds, et al.(2007) Daily High-Resolution-Blended Analyses for Sea											
	Surface Temperature (available at https://doi.org/10.1175/2007JCLI1824.1). (http											
	<pre>s://doi.org/10.1175/2007JCLI1824.1).) Banzon, et al.(2016) A long-term record of b landed catallite and in city can surface temperature for climate monitoring medal</pre>											
	lended satellite and in situ sea-surface temperature for climate monitoring, model ing and environmental studies (available at https://doi.org/10.5194/essd-8-165-201											
	6). (https://doi.org/10.5194/essd-8-165-2016).) Huang et al. (2020) Improvements o											
	f the Daily Optimum Interpolation Sea Surface Temperature (DOISST) Version v02r01,											
	submitted.Climatology is based on 1971-2000 OI.v2 SST. Satellite data: Pathfinder											
	AVHRR SST and Navy AVHRR SST. Ice data: NCEP Ice and GSFC Ice. dimensions(sizes): time(365), lat(720), lon(1440)											
	variables(dimensions): float64 time(time), float32 lat(lat), float32 lon(lon),											
	float32 sst(time, lat, lon)											
	groups:											
	<ipython-input-3-b30fe898003a>:1: DeprecationWarning: tostring() is deprecated. Us</ipython-input-3-b30fe898003a>											
	e tobytes() instead.											
	print (f)											

So what do we have? We have 365 time steps of lat and long with sea surface temperatures at every day. so lets try to pull it out. It is easy....

#### In [4]: 1 print (f.variables['time'])

```
<class 'netCDF4._netCDF4.Variable'>
float64 time(time)
    long_name: Time
    units: days since 1800-01-01 00:00:00
    delta_t: 0000-00-01 00:00:00
    avg_period: 0000-00-01 00:00:00
    axis: T
    actual_range: [79988. 80352.]
unlimited dimensions: time
current shape = (365,)
filling on, default _FillValue of 9.969209968386869e+36 used
<ipython-input-4-bcbaa96a41be>:1: DeprecationWarning: tostring() is deprecated. Us
e tobytes() instead.
print (f.variables['time'])
```

We are going to come back to the time values.

Now lets look at latitude.

In [5]: 1 print (f.variables['lat'])

```
<class 'netCDF4._netCDF4.Variable'>
float32 lat(lat)
    long_name: Latitude
    standard_name: latitude
    units: degrees_north
    actual_range: [-89.875 89.875]
    axis: Y
unlimited dimensions:
current shape = (720,)
filling on, default _FillValue of 9.969209968386869e+36 used
<ipython-input-5-bd934758bff6>:1: DeprecationWarning: tostring() is deprecated. Us
e tobytes() instead.
    print (f.variables['lat'])
```

There are 720 locations for latitude. Or 4 values every degree

Now lets look at longitude

In [6]: 1 print (f.variables['lon'])

```
<class 'netCDF4._netCDF4.Variable'>
float32 lon(lon)
long_name: Longitude
standard_name: longitude
units: degrees_east
actual_range: [1.25000e-01 3.59875e+02]
axis: X
unlimited dimensions:
current shape = (1440,)
filling on, default _FillValue of 9.969209968386869e+36 used
<ipython-input-6-3928fc27ac36>:1: DeprecationWarning: tostring() is deprecated. Us
e tobytes() instead.
print (f.variables['lon'])
```

longitude has 1440 values or twice as many as latitude as it goes around the whole globe

Now we can look at the actual sea surface temperature.

In

[7]:	<pre>1 print (f.variables['sst'])</pre>
	<class 'netcdf4netcdf4.variable'=""></class>
	float32 sst(time, lat, lon)
	long_name: Daily Sea Surface Temperature
	units: degC
	valid_range: [-3. 45.]
	missing_value: -9.96921e+36
	precision: 2.0
	dataset: NOAA High-resolution Blended Analysis
	var_desc: Sea Surface Temperature
	level_desc: Surface
	statistic: Mean
	parent_stat: Individual Observations
	actual_range: [-1.8 36.79]
	unlimited dimensions: time
	current shape = (365, 720, 1440)
	filling on, default _FillValue of 9.969209968386869e+36 used
	<ipython-input-7-1f712b81bbbd>:1: DeprecationWarning: tostring() is deprecated. Us</ipython-input-7-1f712b81bbbd>
	e tobytes() instead.
	<pre>print (f.variables['sst'])</pre>

The sst is really cool. Think of it as 365 days of data, each in an array of values in the shape of 720 by 1440. So for each day we will have a map of sea surface temperatures.

Can we actually look at some numbers?

In [8]:

## 1 f.variables['lat'][:]

riables['lat'][:]						
-86.875,	-86.625,	-86.375,	-86.125,	-85.875,	-85.625,	
-85.375,	-85.125,	-84.875,	-84.625,	-84.375,	-84.125,	
-83.875,	-83.625,	-83.375,	-83.125,	-82.875,	-82.625,	
-82.375,	-82.125,	-81.875,	-81.625,	-81.375,	-81.125,	
-80.875,	-80.625,	-80.375,	-80.125,	-79.875,	-79.625,	
-79 <b>.</b> 375,	-79.125,	-78.875,	-78.625,	-78.375,	-78.125,	
-77 <b>.</b> 875,	-77.625,	-77.375,	-77.125,	-76.875,	-76.625,	
-76.375,	-76.125,	-75.875,	-75.625,	-75.375,	-75.125,	
-74 <b>.</b> 875,	-74.625,	-74.375,	-74.125,	-73.875,	-73.625,	
-73.375,	-73.125,	-72.875,	-72.625,	-72.375,	-72.125,	
-71.875,	-71.625,	-71.375,	-71.125,	-70.875,	-70.625,	
-70.375,	-70.125,	-69.875,	-69.625,	-69.375,	-69.125,	
-68.875 <b>,</b>	-68.625,	-68.375,	-68.125,	-67.875,	-67.625,	
-67.375 <b>,</b>	-67.125,	-66.875,	-66.625,	-66.375,	-66.125,	
-65 <b>.</b> 875,	-65.625,	-65.375,	-65.125,	-64.875,	-64.625,	
-64.375 <b>,</b>	-64.125,	-63.875,	-63.625,	-63.375,	-63.125,	
-62 <b>.</b> 875,	-62.625,	-62.375,	-62.125,	-61.875,	-61.625,	
-61.375,	-61.125,	-60.875,	-60.625,	-60.375,	-60.125,	
-59 <b>.</b> 875,	-59.625,	-59.375,	-59.125,	-58.875,	-58.625,	
_5Q 375	_5Q 175	_57 Q75	_57 625	_57 275	_57 125	

So we can pull out numbers and look at them. They all come out in numpy arrays.

#### In [9]: 1 type(f.variables['lat'][:])

#### Out[9]: numpy.ma.core.MaskedArray

But we can pull them out and set them to a variable. We are not going to use Pandas and just leave everything in numpy arrays for today. This might take a minute because you are putting 400 Mb of data into memory. Your computer might get cranky and we might have to restart at some point.

In	[10]:	1	<pre>lon=f.variables['lon'][:]</pre>
		2	<pre>lat=f.variables['lat'][:]</pre>
		3	<pre>sst=f.variables['sst'][:]</pre>
		4	<pre>time=f.variables['time'][:]</pre>

You can look at the shape of each element. Remember 1 year has 365 days!

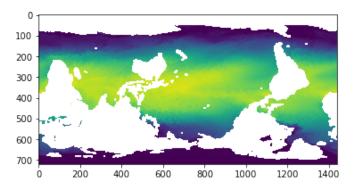
In [11]:	1	np.shape(time)									
Out[11]:	(365,)										
In [12]:	1	np.shape(lon)									
Out[12]:	(1440,)										
In [13]:	1	np.shape(lat)									
Out[13]:	(72	0,)									
In [14]:	1	np.shape(sst)									
Out[14]:	(36	5, 720, 1440)									

So what do we have? We have a lat/lon grid. Then each point has a sst. Then we have that repeated 365 times for each day of the year.

Let's make our easy plot and then we will make the nice map.

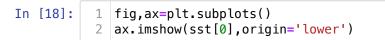
In [15]: 1 fig,ax=plt.subplots()
2 ax.imshow(sst[0])

Out[15]: <matplotlib.image.AxesImage at 0x7fda40784100>

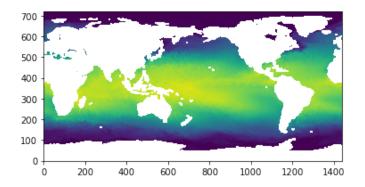


What just happened? This did not make a map? It just took the sst array and plotted it. Each array point has a latittude and longitude we could use to make a map. But we didn't do that. We just did a raw imshow which shows the array. It is a nice start. Read the help and see if you can flip the array?

#### In [16]: 1 ?plt.imshow



Out[18]: <matplotlib.image.AxesImage at 0x7fd9d4f80d90>



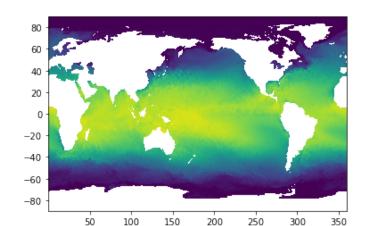
imshow is nice as a quick way to show an array but it is not a map. The next way to make a better map is poolormesh. It uses the lat and long so it is better but still not a map.

#### In [19]:

```
1 fig,ax=plt.subplots()
2 ax.pcolormesh(lon,lat,sst[0])
```

<ipython-input-19-ebfd4181c561>:2: MatplotlibDeprecationWarning: shading='flat' wh
en X and Y have the same dimensions as C is deprecated since 3.3. Either specify
the corners of the quadrilaterals with X and Y, or pass shading='auto', 'nearest'
or 'gouraud', or set rcParams['pcolor.shading']. This will become an error two mi
nor releases later.

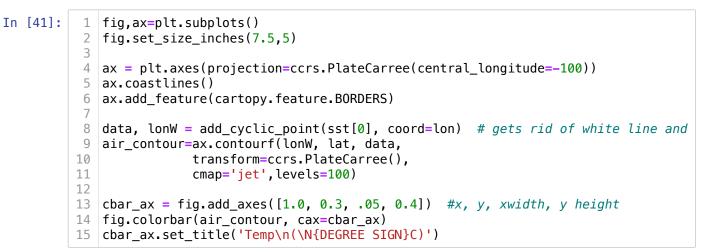
ax.pcolormesh(lon,lat,sst[0])



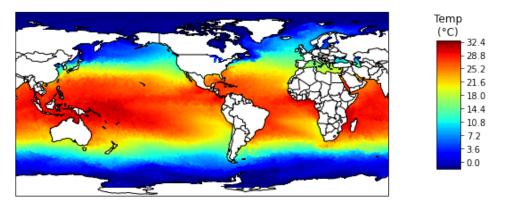
#### Out[19]: <matplotlib.collections.QuadMesh at 0x7fd9eca2b820>

Lets try to make that nicer

• Let's go back and grab our last map and paste it in here. You should be able to make it work

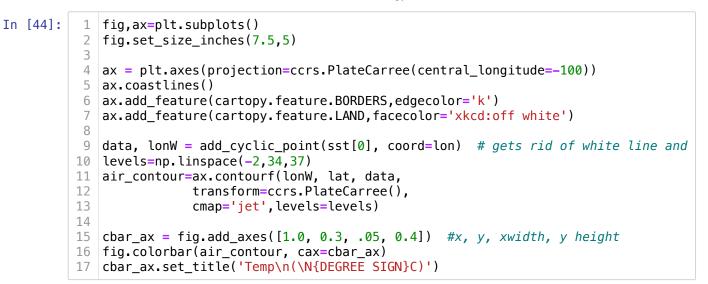


#### Out[41]: Text(0.5, 1.0, 'Temp\n(°C)')

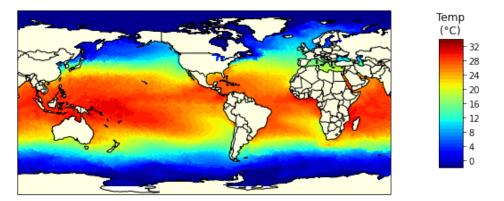


#### Let's make it look better

- I am going to make the land a slightly different color
- Then we need to set the levels better. The reason is different days have different ranges.
- We will use linspace to get levels and then pass the levels. You are setting the range. For contourf you need to be over the min and max.

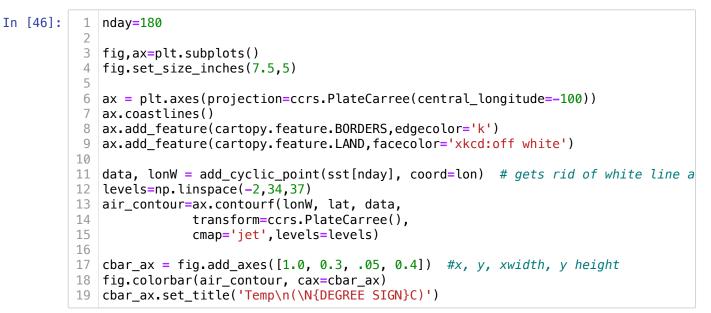


#### Out[44]: Text(0.5, 1.0, 'Temp\n(°C)')

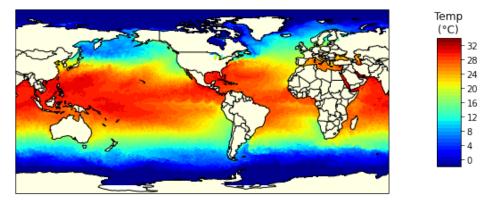


#### remember that we have 365 days of data.

- So lets set a parameterto day number that we can change
- If you know me you should know a for loop is coming and we are stepping our way there.
- Now make a variable called "nday" and set it equal to the day you want to plot. For example January 31 will be 30 because python starts at 0.
- Compare day 0 and day 180. Do the temperatures look different?



#### Out[46]: Text(0.5, 1.0, 'Temp\n(°C)')



#### That is a good looking map!!!!

I am starting to see a path forward for how to make a movie of the temperatures over the year.

- Python is not perfect for making movies.
- We are going to make a modern flip book.
- We will save an image for each frame and then we can use python or a web program to put them together. It is like turning a burst on your phone into a movie.
- We want to end with an animated GIF with python or a program. Things we will need to be able to do.

How do we do thi??

- Loop over all the data.
- Figure out the date for the data.
- plot the data.
- · save the data in a unique file for each day
- use python/website to put it all togehter.

So we need to learn each of these steps. I will start with figuring out the date for each slice.

#### Back to our crazy dates!

Still crazy! Time since 1800 in days! Let's deal with it like last time. Go cut and paste what we did.

- · Set your startdate
- · calculate your datedelta
- · Add them together to figure out your date
- · use ndays again

1 nday=180

· Change ndays and see if it

#### In [50]:

```
2 startdate=datetime.datetime(1800,1,1)
```

```
3 datedelta=datetime.timedelta(days=time[nday])
```

4 mapdate=startdate+datedelta

```
5 print ("The date is {:%B %d, %Y}".format(mapdate))
```

The date is June 30, 2019

#### We need a for loop

- I do not know the best way to for loop here.
- I see three options
- just give us a number to loop 20. e.g. 365. But would get an error on a leap year.
- We could do an enumerate loop.
- We could also just loop to length of a parameter.
- I am going to do np.arange(len(time))
- then print nday

In

[53] <b>:</b>	1 2	<pre>for nday in np.arange(len(time)):     print(nday)</pre>	
	0		
	1		
	2		
	3		
	4		
	5		
	6		
	7		
	8		
	9		
	10		
	11 12		
	12		
	14		
	15		
	16		
	17		
	18		
	10		

Now print each day in the foor loop. to save memory put the least amount of things needed in a for loop. You can do this!

In [54]:

1

The The The The The The The The The The	date date date date date date date date	on on on on on on on on on on on	day day day day day day day day day day	1 2 3 4 5 6 7 8 9 10 11 12 13 14	is is is is is is is is is is is is is i	January January January January January January January January January January January January January January	02, 03, 04, 05, 06, 07, 08, 09, 10, y 11, y 12, y 13, y 14, y 15,	2019 2019 2019 2019 2019 2019 2019 2019
The The	date date	on on	day day	11 12	is is	5 January 5 January	y 12, y 13,	, 2019 , 2019
The The The	date date date	on on on	day day day	15 16 17 18	is is is	5 Januar 5 Januar 5 Januar 5 Januar 5 Januar	y 16, y 17, y 18,	, 2019 , 2019 , 2019

#### this might not make sense.

If we are going to make 365 plots one for each day. We need to save each to a file. So we need to create a file name we can send to savfig.

- We will want our files nicely oraganized.
- We could use the dates to name the files. But we will lose their order. You could name the file yyyymmdd and that should keep them in order. I haven't tried yet.
- Instead I will do a number in the file name.
- But we will want a three digit number in the name. This will keep them all in order.
- it will look like sst\_000.png, sst\_001.png, sst\_002.png, .... sst\_364.png.

netcdf-v2-2020 - Jupyter Notebook

We can use format to pad the numbers. You remember this?
 <u>https://stackoverflow.com/questions/339007/nicest-way-to-pad-zeroes-to-string</u>

In [56]:

56] <b>:</b>	1 2 3 4 5 6	for r c n t	<pre>startdate=datetime.datetime(1800,1,1) for nday in np.arange(len(time)):     datedelta=datetime.timedelta(days=time[nday])     mapdate=startdate+datedelta     filename="sst_{:03d}.png".format(nday)     print ("The date on day {} is {:%B %d, %Y} with filename {}".format(nday,ma</pre>											
											filename			
				-			-	-			filename			
											filename			
											filename			
											filename			
				-			-	-			filename			
											filename			
											filename			
											filename			
				-			-	-			filename			
											filename			
											filename			
											filename			
				-			-	-			filename	_		
											filename			
											filename			
											filename			
											filename			
											filename			
	Tha	4-+-	~ ~	4~~	10		1		7010	· · ÷ + h	f:100000	· · · · · ·	10	

But I will want to put the files into a subfolder. So if you make a folder named sst/ we could add the prefix to the name. You need to make a folder on your computer. I am ging to repeat this. You named it sst/ but the folder doesn't exist until you make it and then it will put the files into it.

```
startdate=datetime.datetime(1800,1,1)
In [57]:
            1
             2
                for nday in np.arange(len(time)):
            3
                     datedelta=datetime.timedelta(days=time[nday])
                    mapdate=startdate+datedelta
            4
                     filename="sst/sst {:03d}.png" format(nday)
            5
                     print ("The date on day {} is {:%B %d, %Y} with filename {}".format(nday,ma
             6
           The date on day 0 is January 01, 2019 with filename sst/sst_000.png
           The date on day 1 is January 02, 2019 with filename sst/sst_001.png
The date on day 2 is January 03, 2019 with filename sst/sst_002.png
           The date on day 3 is January 04, 2019 with filename sst/sst_003.png
           The date on day 4 is January 05, 2019 with filename sst/sst_004.png
           The date on day 5 is January 06, 2019 with filename sst/sst_005.png
           The date on day 6 is January 07, 2019 with filename sst/sst_006.png
           The date on day 7 is January 08, 2019 with filename sst/sst_007.png
The date on day 8 is January 09, 2019 with filename sst/sst_008.png
The date on day 9 is January 10, 2019 with filename sst/sst_009.png
           The date on day 10 is January 11, 2019 with filename sst/sst_010.png
           The date on day 11 is January 12, 2019 with filename sst/sst_011.png
           The date on day 12 is January 13, 2019 with filename sst/sst_012.png
           The date on day 13 is January 14, 2019 with filename sst/sst_013.png
           The date on day 14 is January 15, 2019 with filename sst/sst_014.png
The date on day 15 is January 16, 2019 with filename sst/sst_015.png
           The date on day 16 is January 17, 2019 with filename sst/sst_016.png
           The date on day 17 is January 18, 2019 with filename sst/sst_017.png
           The date on day 18 is January 19, 2019 with filename sst/sst_018.png
           The date on day 10 is lanuary 20
                                                    2010 with filonome cot/cot 010 pro
```

Now we know how to

- Make a title name with the date.
- make a file name
- loop over the array

But before you begin you need to think a little.

Each timestep we are going to make a figure, make a map, and save it.

But making a figure and map are "expensive" and eat up cpu. And you need to this 365 times. So when you loop you want to set as many things before your loop as possible. Then you can use ax.cla() to clear the axes. If you don't clear the axis you keep putting each day over the last day and you computer will grind to a halt. So i would test things by running just a few days. You can do this by only calling the first so many parts of the array. Plus you can use kernal->interrupt to stop a run.

Don't forget a title with your name and date. You need your name in the title.

TRICK: the first time using the for loop do not go over the np.arange(len(time)). just do np.arange(10). This way if something goes wrong it just happens 10 times and not 365.

TRICK: you are making 365 files. make them small. use dpi=50. So you call to savefig will look like fig.savefig(filename,dpi=50,bbox\_inches='tight',facecolor='white') You need to add the facecolor='white' to make the titles look good.

The websites for making a gif are full or viruses. I would use python.

I found this stackoverflow example.

https://stackoverflow.com/questions/41228209/making-gif-from-images-using-imageio-in-python (https://stackoverflow.com/questions/41228209/making-gif-from-images-using-imageio-in-python)

Look at the answer and just update with your directories!

BUT THERE IS A BIG PROBLEM! If you look up os.listdir It says the order of files is arbitrary. On some computers is works and on some it doesn't and it jumbles the order.

We have two solutions.

- 1. Kai came up with a great one about reading the files in (See below)
- 2. I tryed to understand what went wrong and it looks like we can sort our files. So I googled this "python listdir in order". It then restated that listserv jumbles the order. But it says you can add the sorted function to the for loop. See below!

PS. This is slow. My computer takes 30 minutes and makes a lot of noise!

PPS. You cannot open your movie in preview. Use a web browser.



Brian's Map January 01, 2019 Brian's Map January 02, 2019

<ipython-input-69-1c80be8b3878>:16: MatplotlibDeprecationWarning: Adding an axes u sing the same arguments as a previous axes currently reuses the earlier instance. In a future version, a new instance will always be created and returned. Meanwhil e, this warning can be suppressed, and the future behavior ensured, by passing a u nique label to each axes instance.

ax = plt.axes(projection=ccrs.PlateCarree(central\_longitude=-100))

Brian's Map January 03, 2019 Brian's Map January 04, 2019 Brian's Map January 05, 2019 Brian's Map January 06, 2019 Brian's Map January 07, 2019 Brian's Map January 08, 2019 Brian's Map January 09, 2019 Brian's Map January 10, 2019 Brian's Map January 11, 2019 Brian's Map January 12, 2019

### Make the GIF. Choose a method.....

