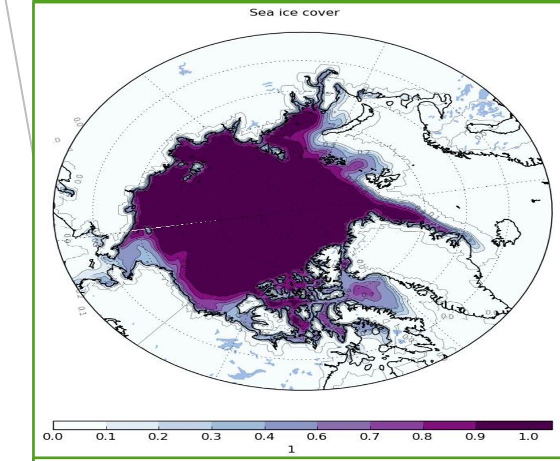
GSFC Internship: Eviz and Data Visualization

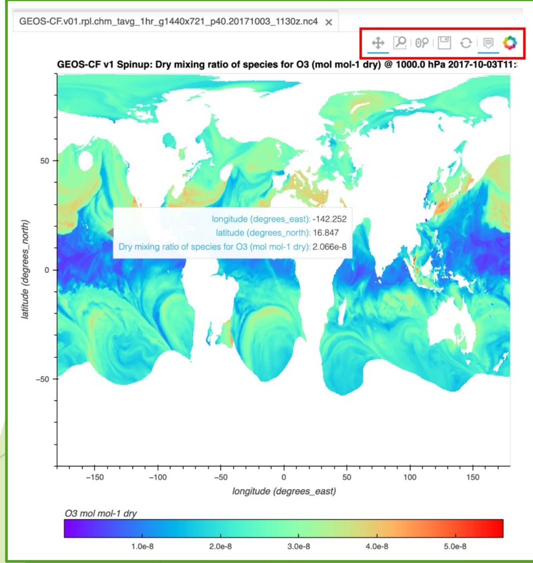
Deon Kouatchou

Supervisors: Vanessa Valenti and Carlos Cruz

Eviz Background Info

- It is known that current visualization methods can be limited by being static, non-interactive, and sometimes tedious and complex.
- Through the two tools of eViz and iViz, Eviz aims to improve visualization with its easy-to-use functionality, comparison and extensive mapping capabilities, and interactivity.
- The purpose is to make visualization more accessible and effective for scientists




GEOS-CF.v01.rpl.chm_tavg_1hr_g1440x721_p40.20171003_1130z.nc4

GEOS-CF v1 Spinup: Dry mixing ratio of species for O3 (mol mol-1 dry) @ 1000.0 hPa 2017-10-03T11:

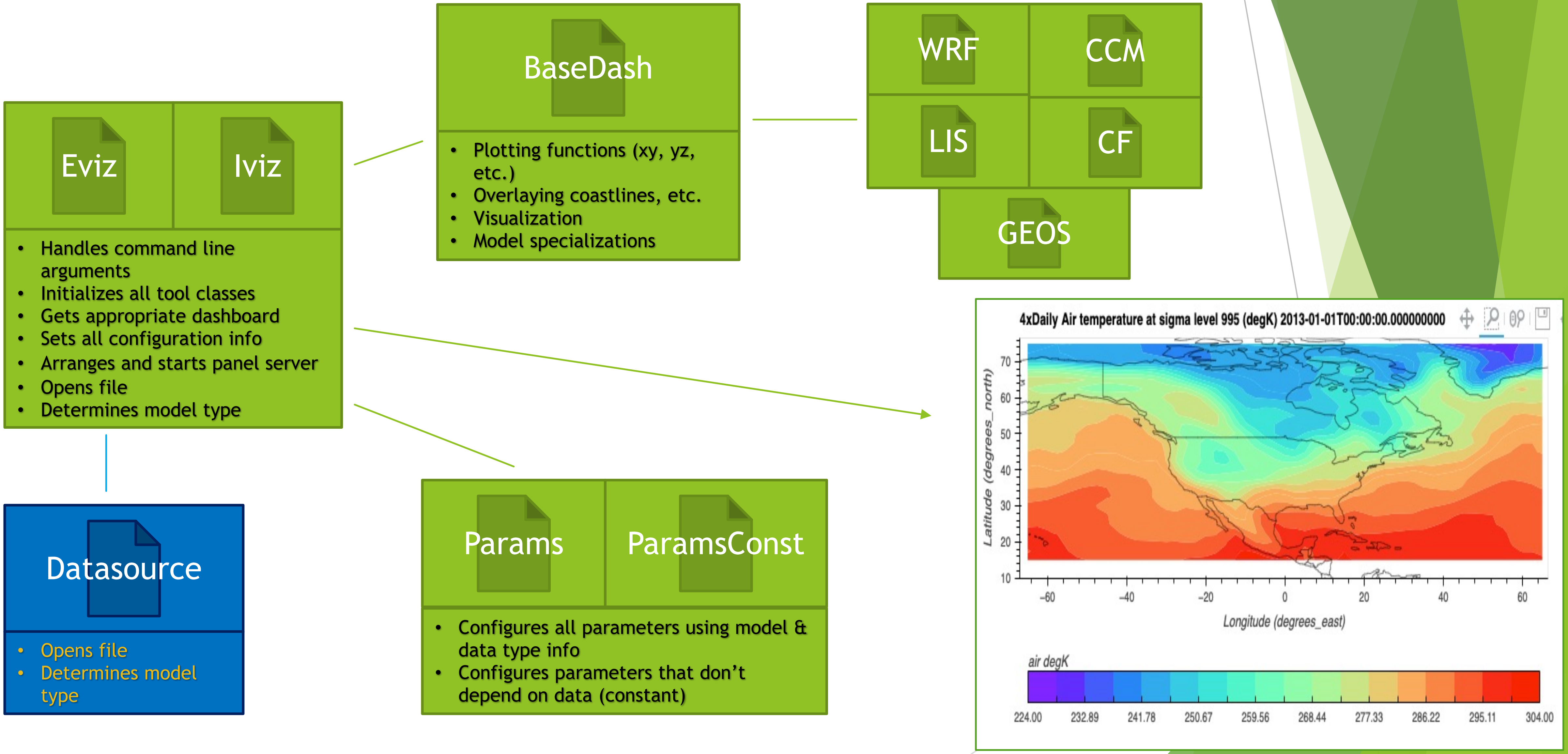
longitude (degrees_east): -142.252
latitude (degrees_north): 16.847
Dry mixing ratio of species for O3 (mol mol-1 dry): 2.066e-8

iViz

Introduction of the Problem

- As of the initial release, Eviz, while being very effective, has only been programmed to read earth systems data that are usually stored as NetCDF files.
- The authors have long considered expanding its capabilities by integrating data stored as other file types, especially satellite data
- The objective is to ultimately comparing earth systems model data with other types of data

Eviz General Architecture (Subset)



Eviz

Iviz

- Handles command line arguments
- Initializes all tool classes
- Gets appropriate dashboard
- Sets all configuration info
- Arranges and starts panel server
- Opens file
- Determines model type

BaseDash

- Plotting functions (xy, yz, etc.)
- Overlaying coastlines, etc.
- Visualization
- Model specializations

WRF

CCM

LIS

CF

GEOS

Datasource

- Opens file
- Determines model type

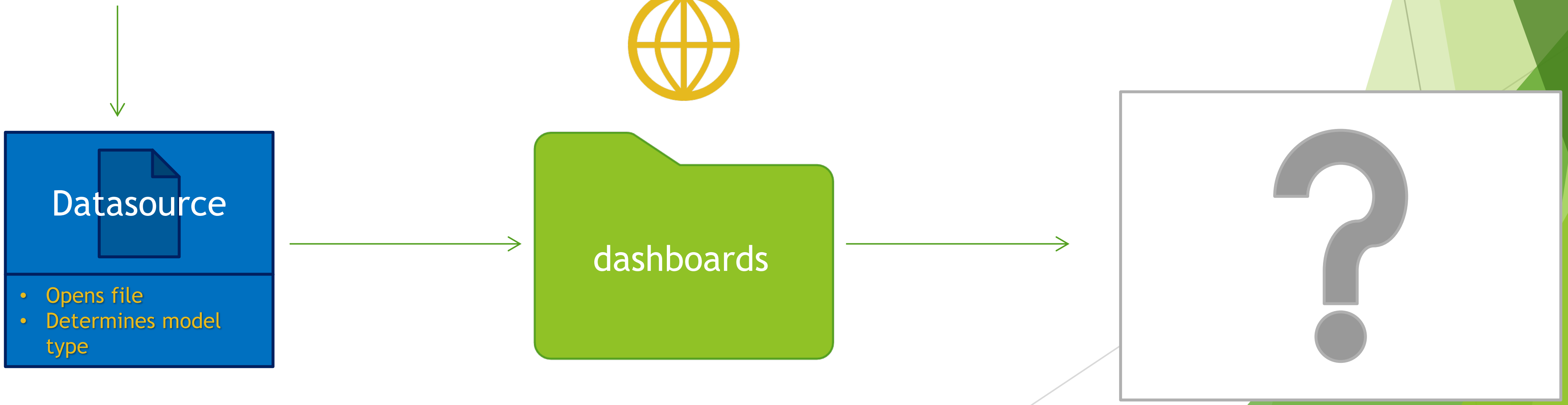
Params

ParamsConst

- Configures all parameters using model & data type info
- Configures parameters that don't depend on data (constant)

Overall Objective

- File types to open with class
  - NetCDF: climate/weather models (already implemented)
  - HDF: satellite data (objective)
  - TIF/TIFF: satellite data (future)
  - CSV: point and geospatial data (future)

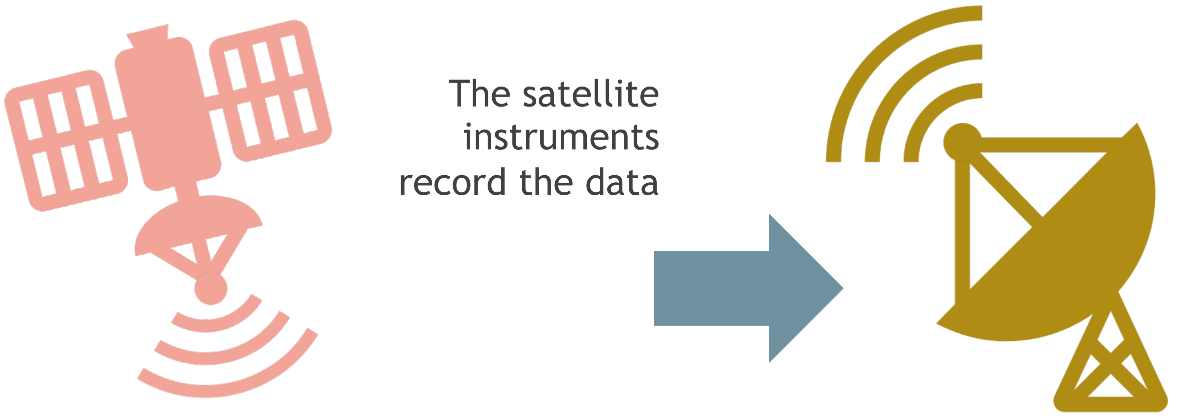


Datasource

- Opens file
- Determines model type

dashboards

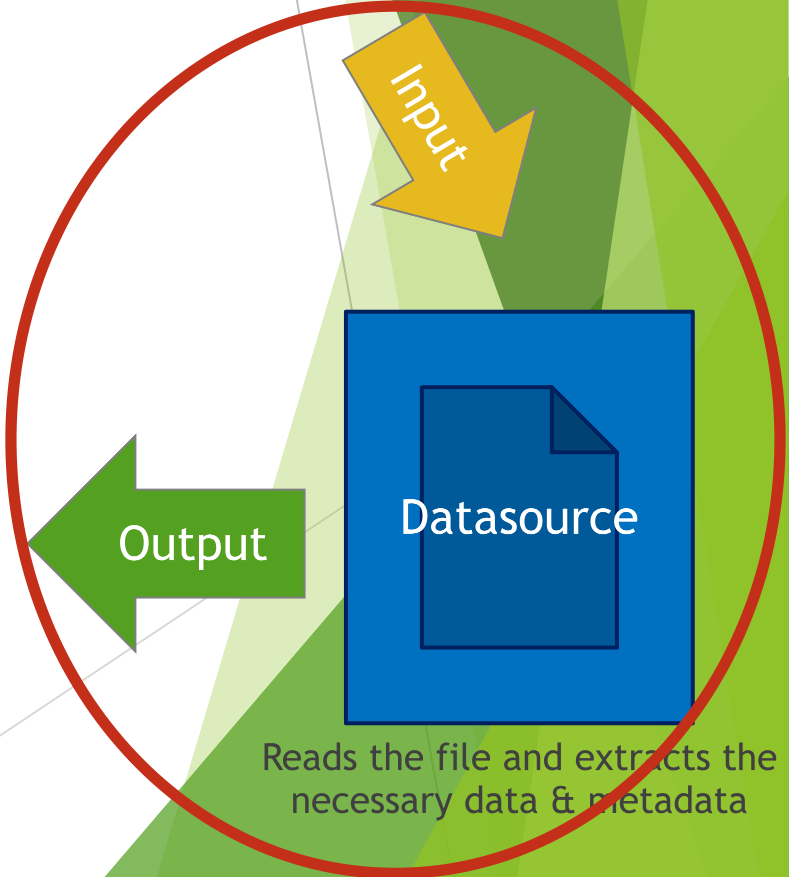
Task



The satellite instruments record the data

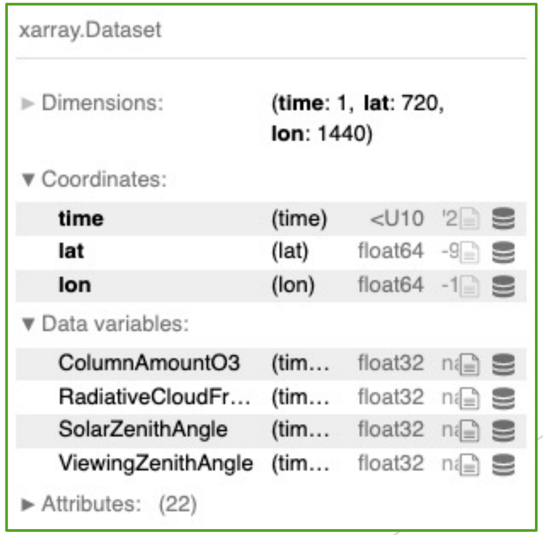
The ground station receives and streams the data

A data file from a database



Reads the file and extracts the necessary data & metadata

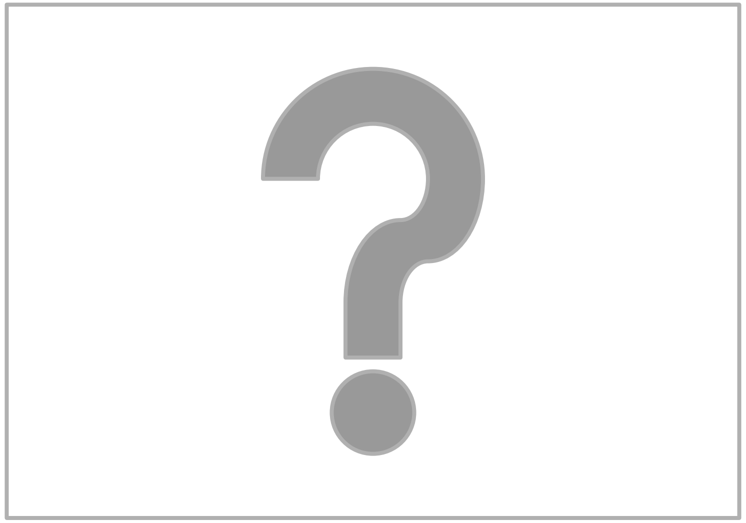
Resulting XArray object



```
xarray.Dataset

► Dimensions:           (time: 1, lat: 720,
                         lon: 1440)
▼ Coordinates:
   time                 (time)    <U10   '2
   lat                  (lat)     float64 -9
   lon                  (lon)     float64 -1
▼ Data variables:
   ColumnAmountO3       (tim...   float32 n
   RadiativeCloudFr...  (tim...   float32 n
   SolarZenithAngle     (tim...   float32 n
   ViewingZenithAngle   (tim...   float32 n
► Attributes:  (22)
```

Through iViz

Resulting Plot



Requirements

Image Source: python.org

Image Source: foundations.projectpythia.org

- ☑ Rudimentary Python and object-oriented programming skills
- □ Familiarity with Python application in data science and visualization
- □ Comprehension of Version Control and Git workflows
- □ A basic understanding of geographic data
- □ Knowing how eViz & iViz work
- □ Know what any of these file types even mean
- □ All these weird Python packages and modules

. . .

Image Source: github.com

Image Source: en.wikipedia.org

Image Source: holoviz.org

nage Source: numpy.org

Image Source: deployplace.com

Image Source: numfocus.org

Python Course Series for GSFC Interns

Image Source: commons.wikmedia.org

Covered Topics

- Git
- Input/Output on Text Files
- NumPy
- Matplotlib
- Pandas

Additional Learning

- Jupyter Notebook & Lab
- PyCharm
- Using Terminal
- XArray

Image Source: fullstackpython.org

Image Source: matplotlib.org

Image Source: pandas.pydata.org

☑Familiarity with Python application in data science
☑Comprehension of Git and Git workflows

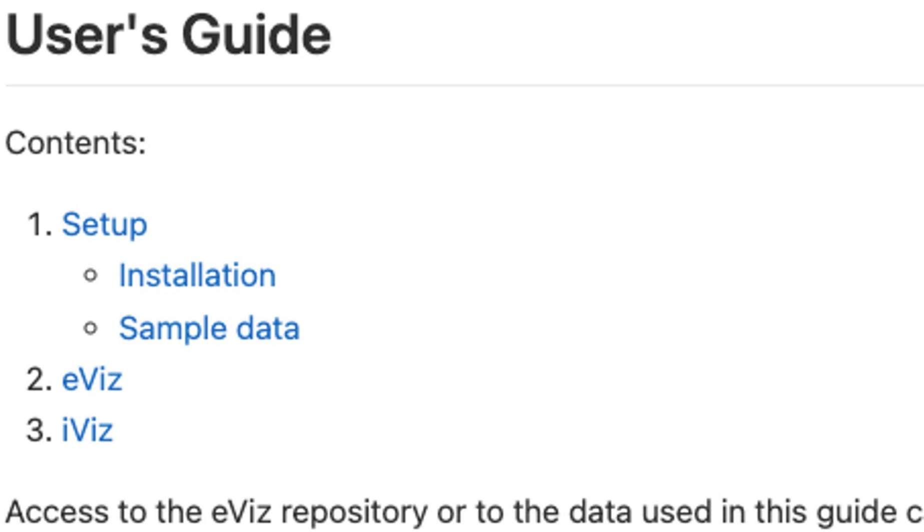
Image Source: support.apple.com

Eviz Familiarization

☑Familiarity with Python application in visualization
☑Understanding how eViz & iViz function
☑Knowledge about geographic data & file types

- Testing the User's Guide
- Holoviz Packages
- Understanding HDF Satellite Data
  - OMI (global) – H5py
  - Landsat (swath) - PyHDF

Image Source: hvplot.holoviz.org



User's Guide

Contents:

1. Setup
   - Installation
   - Sample data
2. eViz
3. iViz

Access to the eViz repository or to the data used in this guide can be obtained by contacting Carlos Cruz or Vanessa Valenti.

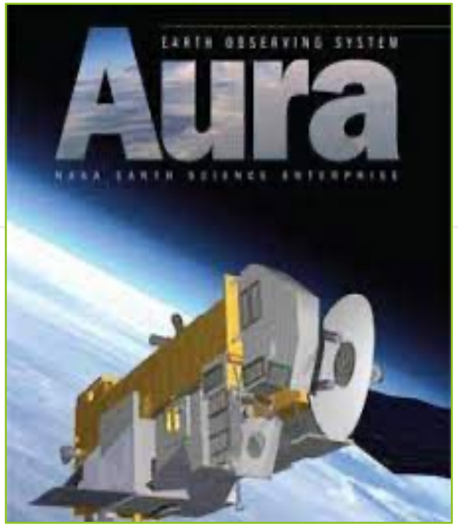
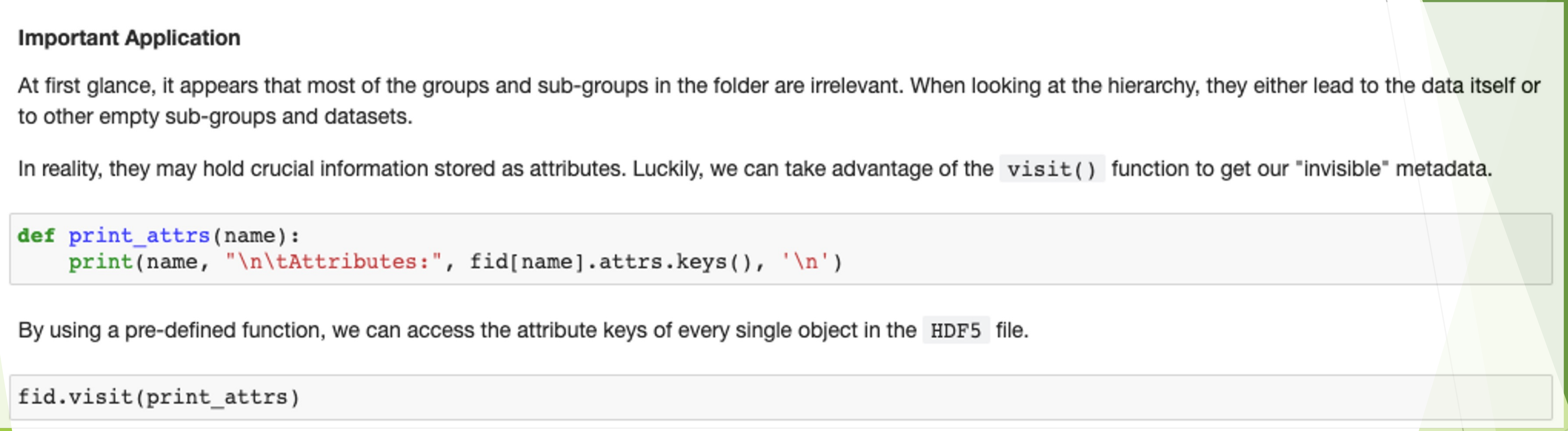


Image Source: aura.gsfc.nasa.gov

Image Source: usgs.gov

Jupyter Notebook Examples



Important Application

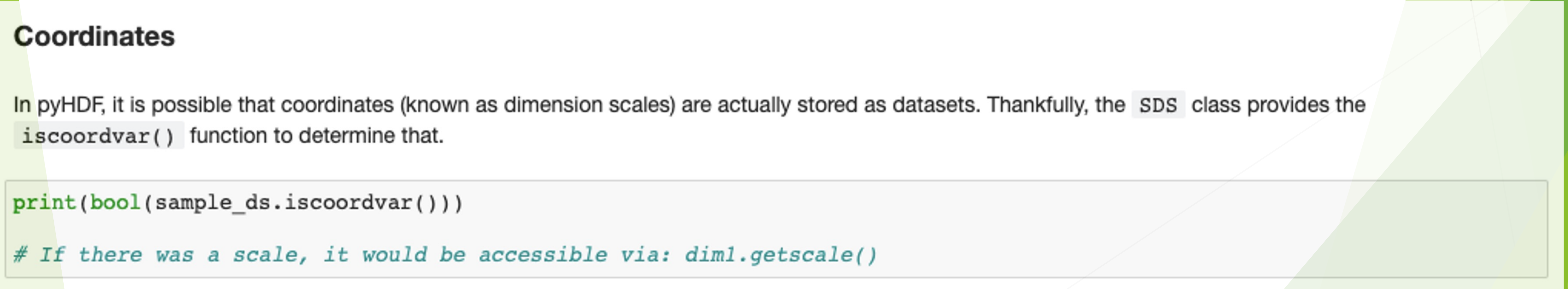
At first glance, it appears that most of the groups and sub-groups in the folder are irrelevant. When looking at the hierarchy, they either lead to the data itself or to other empty sub-groups and datasets.

In reality, they may hold crucial information stored as attributes. Luckily, we can take advantage of the `visit()` function to get our "invisible" metadata.

```
def print_attrs(name):
    print(name, "\n\tAttributes:", fid[name].attrs.keys(), '\n')
```

By using a pre-defined function, we can access the attribute keys of every single object in the `HDF5` file.

```
fid.visit(print_attrs)
```

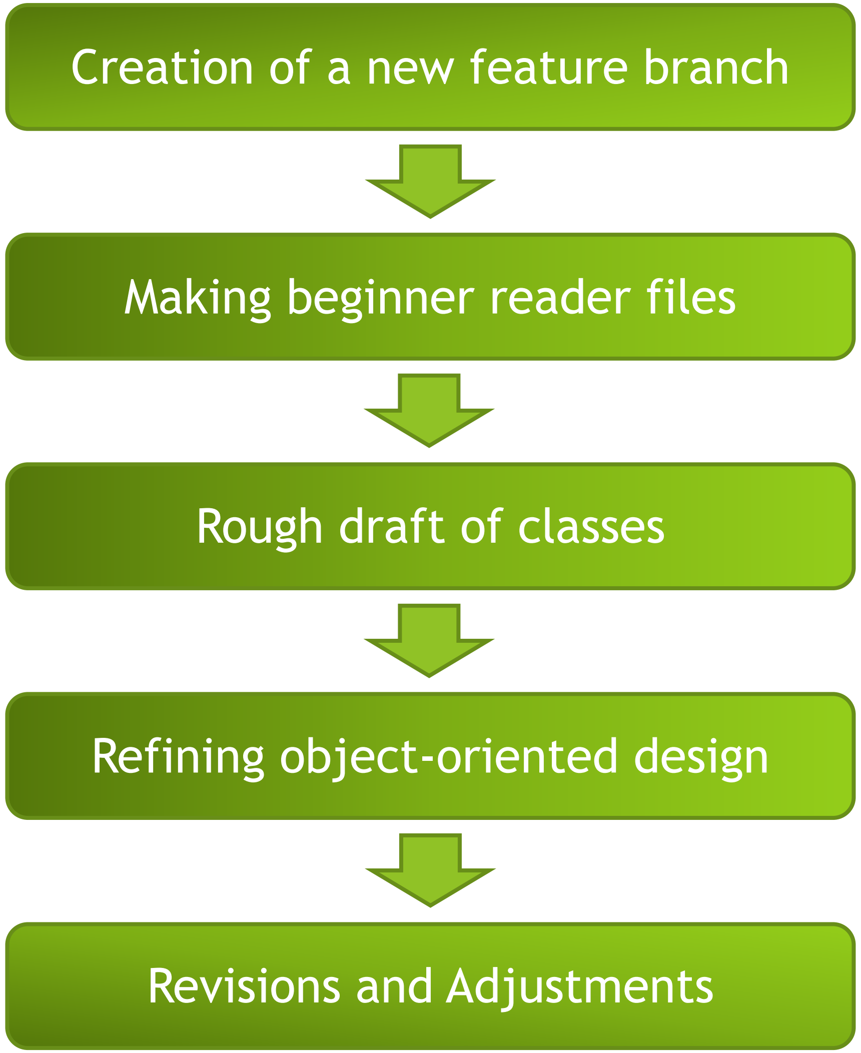


Coordinates

In pyHDF, it is possible that coordinates (known as dimension scales) are actually stored as datasets. Thankfully, the `SDS` class provides the `iscoordvar()` function to determine that.

```
print(bool(sample_ds.iscoordvar()))

# If there was a scale, it would be accessible via: dim1.getscale()
```

The Process I



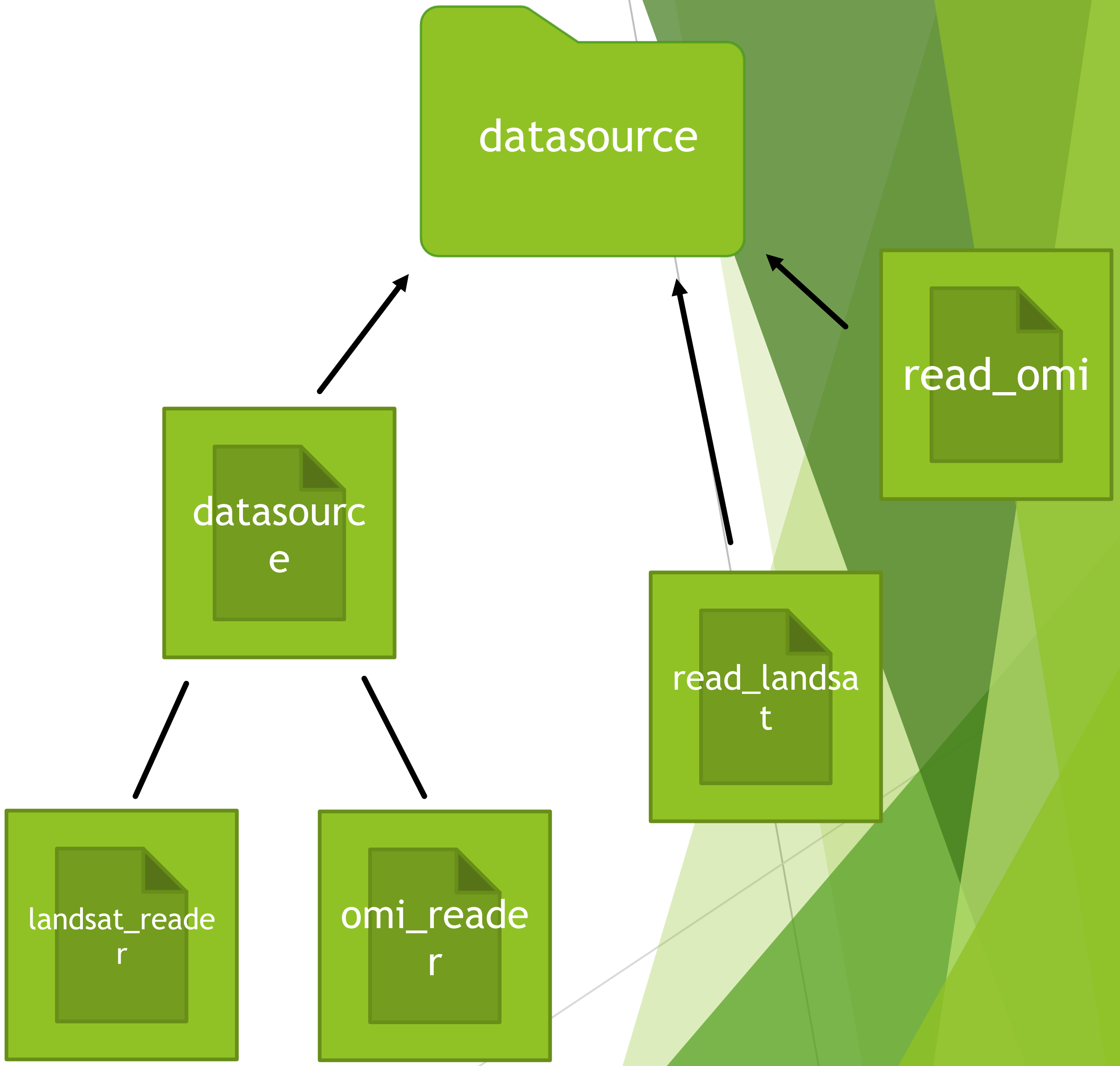
Creation of a new feature branch

Making beginner reader files

Rough draft of classes

Refining object-oriented design

Revisions and Adjustments



The Process II

Adjustments

- ➢ Using Python RegEx module
- ➢ Reading in multiple variables
- ➢ Exception Handling
- ➢ Adding time coordinate
- ➢ Python logging

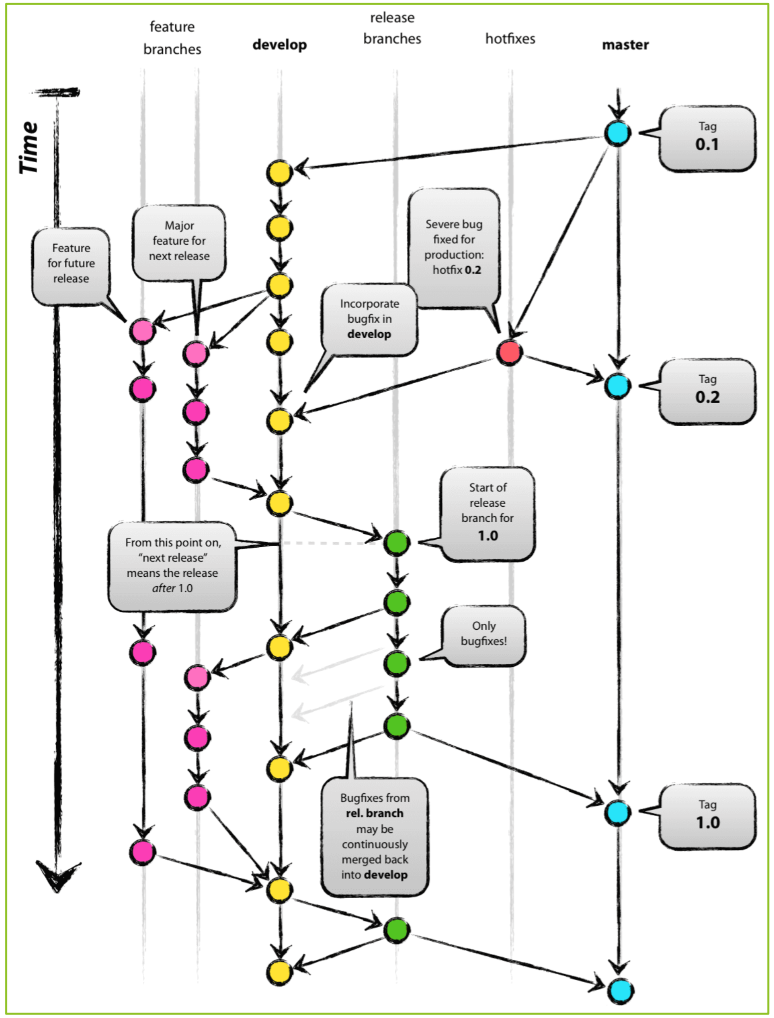
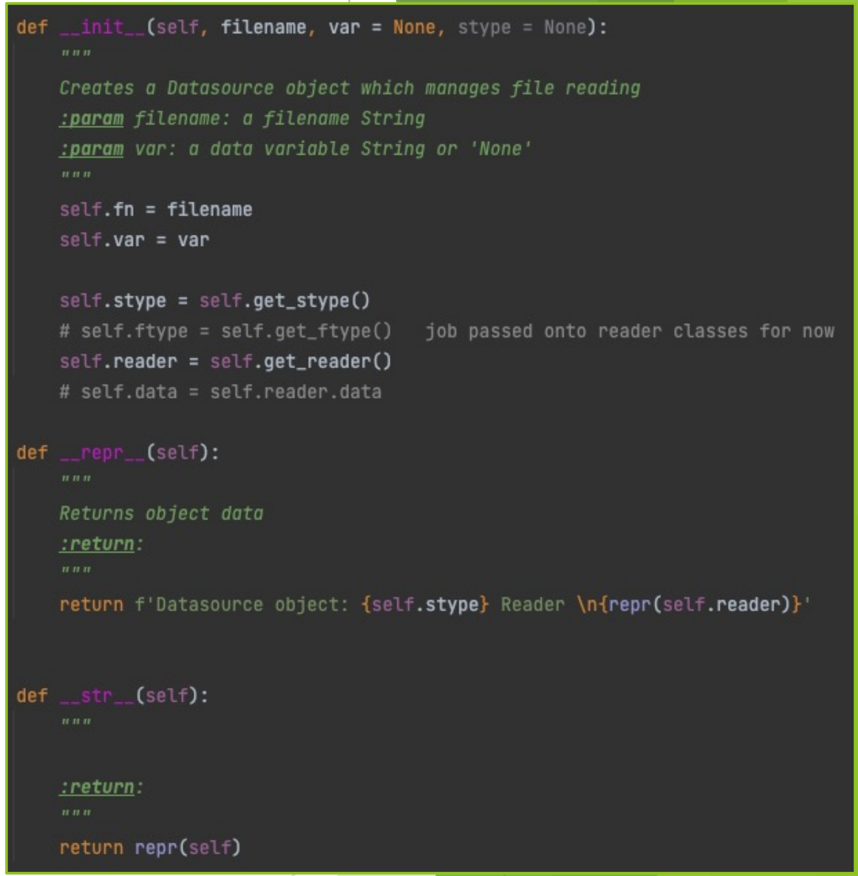

feature branches
develop
release branches
hotfixes
master
Time
Tag 0.1
Feature for future release
Major feature for next release
Severe bug fixed for production: hotfix 0.2
Incorporate bugfix in develop
Tag 0.2
Start of release branch for 1.0
From this point on, "next release" means the release after 1.0
Only bugfixes!
Bugfixes from rel. branch may be continuously merged back into develop
Tag 1.0

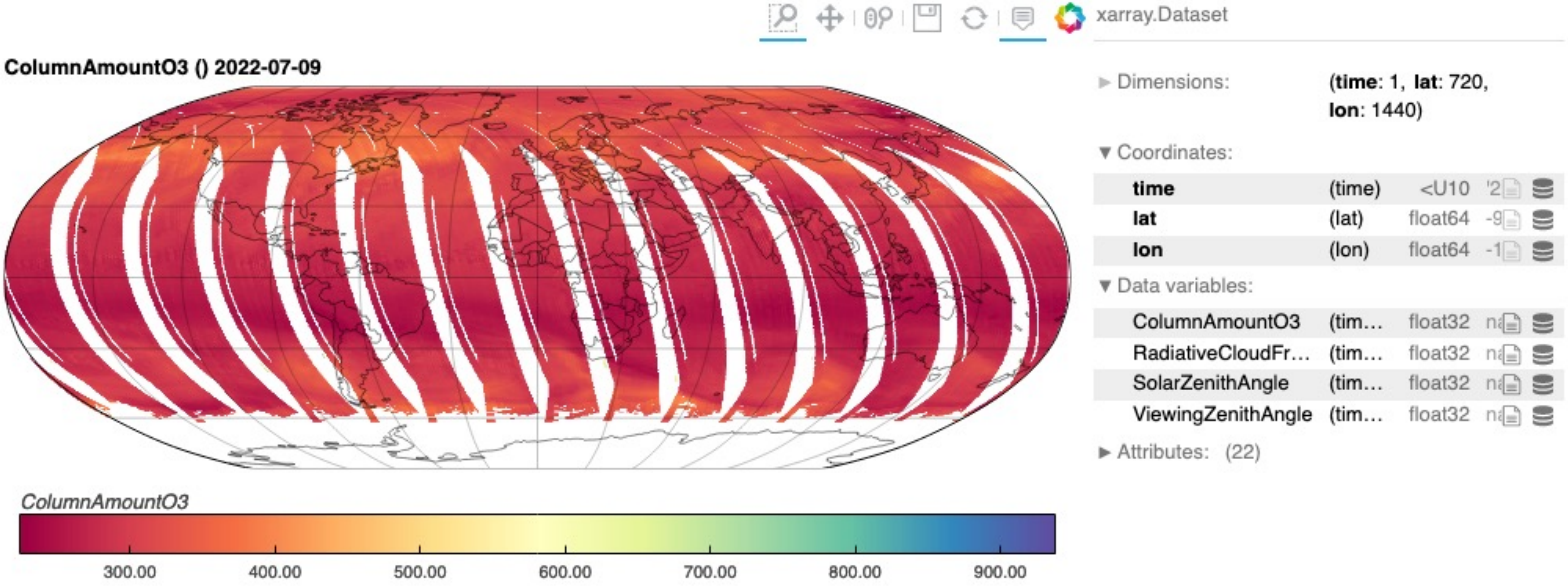
Image Source: nvie.com



```
def __init__(self, filename, var = None, stype = None):
    """
    Creates a Datasource object which manages file reading
    :param filename: a filename String
    :param var: a data variable String or 'None'
    """
    self.fn = filename
    self.var = var

    self.stype = self.get_stype()
    # self.ftype = self.get_ftype()   job passed onto reader classes for now
    self.reader = self.get_reader()
    # self.data = self.reader.data

def __repr__(self):
    """
    Returns object data
    :return:
    """
    return f'Datasource object: {self.stype} Reader \n{repr(self.reader)}'


def __str__(self):
    """

    :return:
    """
    return repr(self)
```

Results

Sample data from Ozone Monitoring Instrument (Level 3)



```
xarray.Dataset

▶ Dimensions:         (time: 1, lat: 720,
                      lon: 1440)
▼ Coordinates:
    time              (time)    <U10    '2
    lat               (lat)     float64 -9
    lon               (lon)     float64 -1
▼ Data variables:
    ColumnAmountO3    (tim...   float32 na
    RadiativeCloudFr… (tim...   float32 na
    SolarZenithAngle  (tim...   float32 na
    ViewingZenithAngle (tim...  float32 na
▶ Attributes:   (22)
```

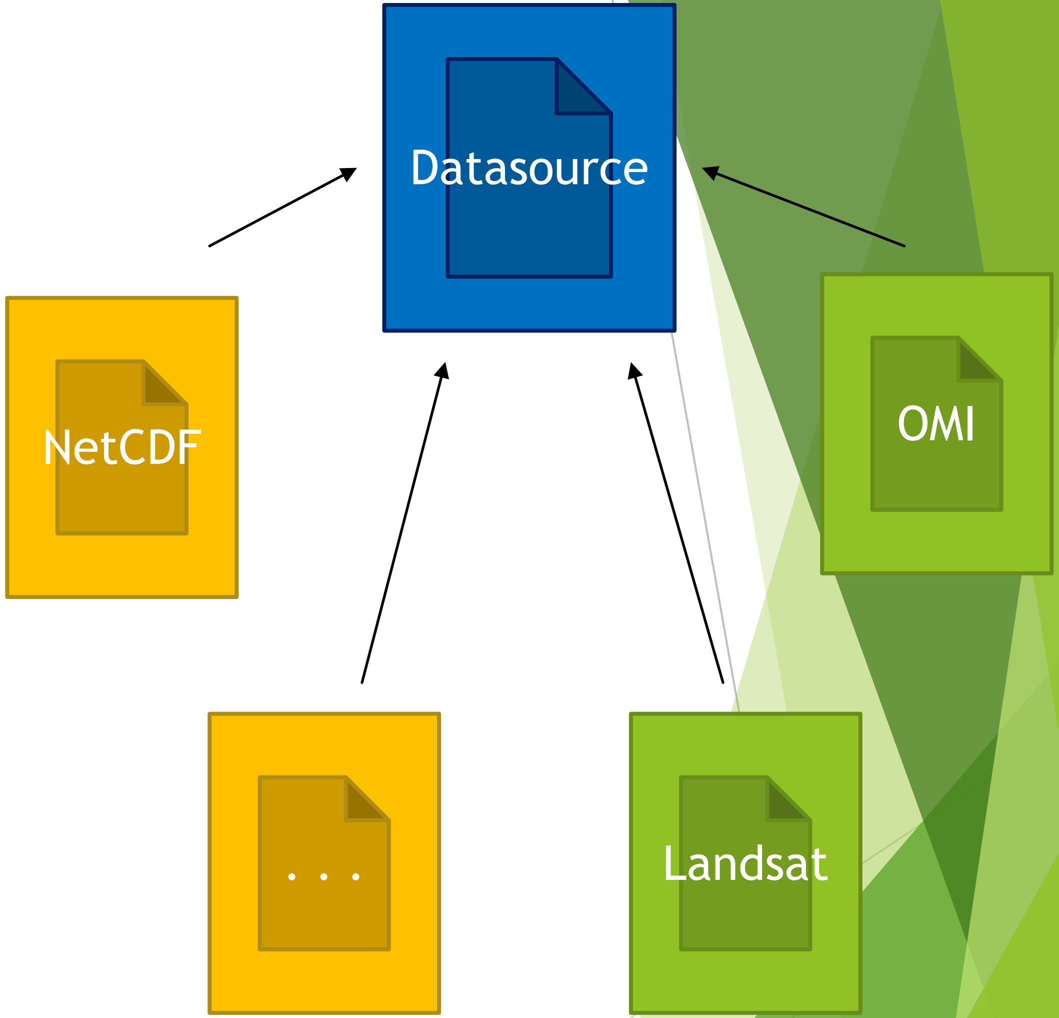
Impact

- Expanding upon Eviz’s efficacy by
  - Allowing scientists visualize some satellite data
  - Allowing scientists perform more extensive comparisons
  - Setting up a structure for future development

Image Source: theconversation.com

Future Ventures

- Expanding the Datasource class to implement:
    - More HDF-based satellite data such as from MODIS (Moderate Resolution Imaging Spectroradiometer)
    - NetCDF and models currently available to visualize (GEOS, CCM, WRF, etc)
    - CSV and TIFF file types
- Expand Datasource class to function with eViz



Thank You!

Any Questions?

NASA Goddard Space Flight Center
Advanced Software
Technology Group