PROJECT PYTHIA
Introduction: all about Pythia and Cookbooks

9 AUGUST 2023

Project Pythia

what, why, where is it headed?

NCAR
UCAR

Project motivation

PROJECT PYTHIA

1. Geoscience community moving to open source software and cloud computing for analysis to better support open, reproducible science
2. Python ecosystem, cloud computing are complex and dynamic environments
3. Geoscientists are not computer scientists
4. Little training material exists that is focused specifically on needs of geoscientists

… but which package should I use?

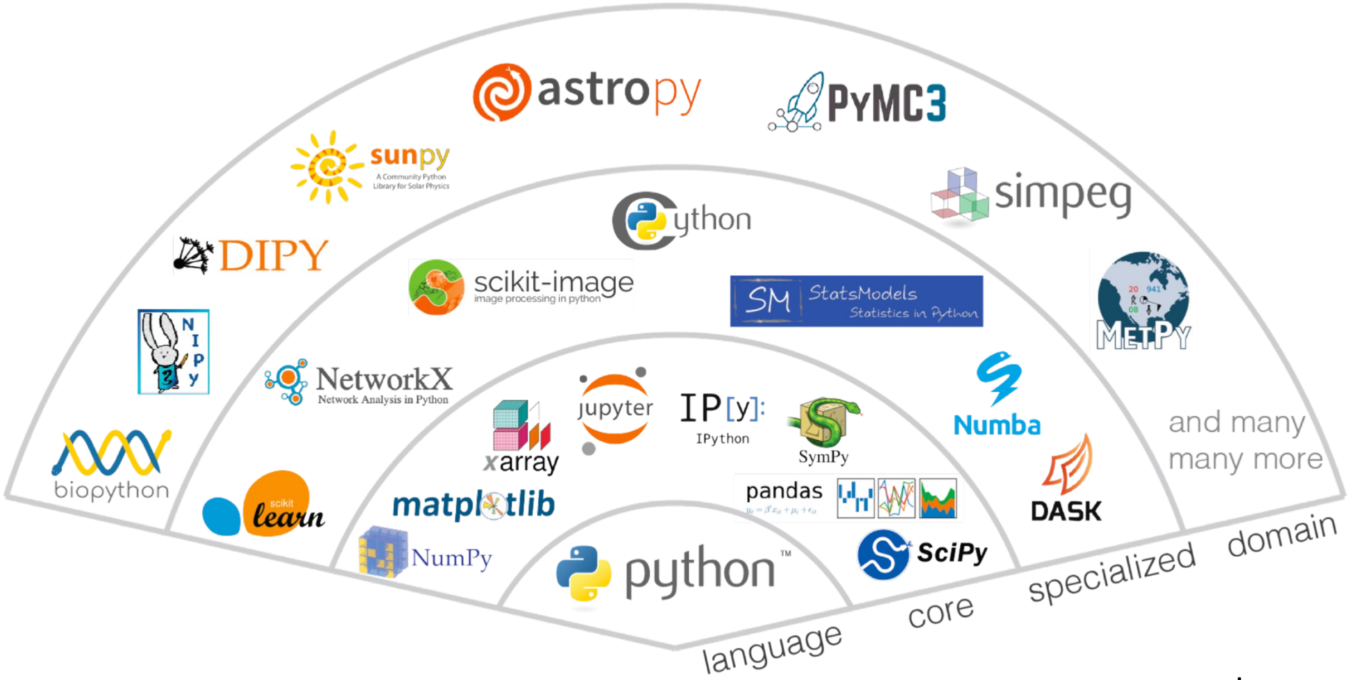
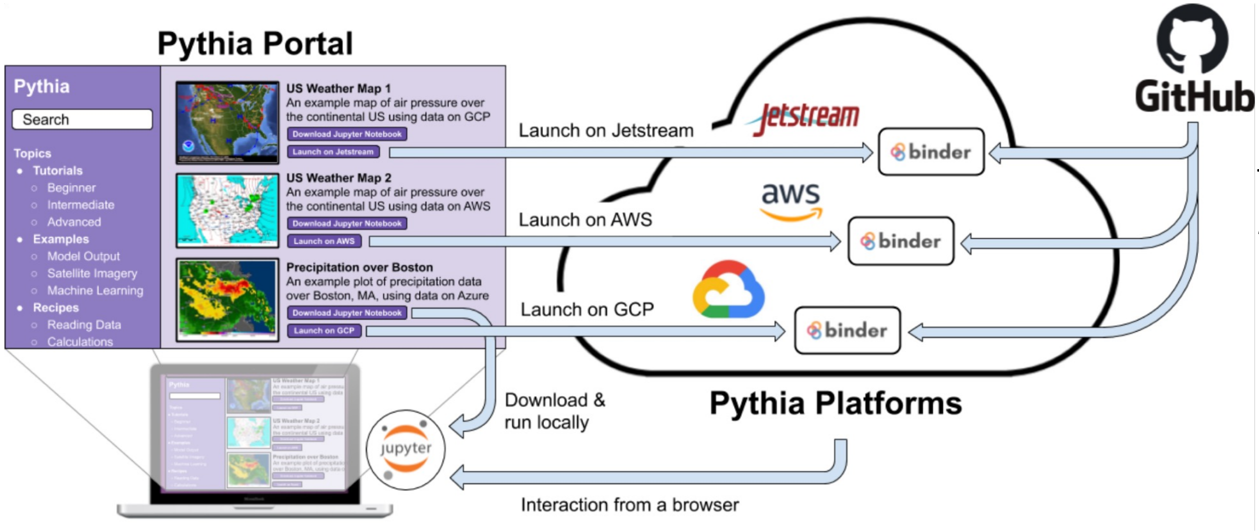


Image credit: VanderPlas, 2017

A Community Learning Resource for Geoscientists

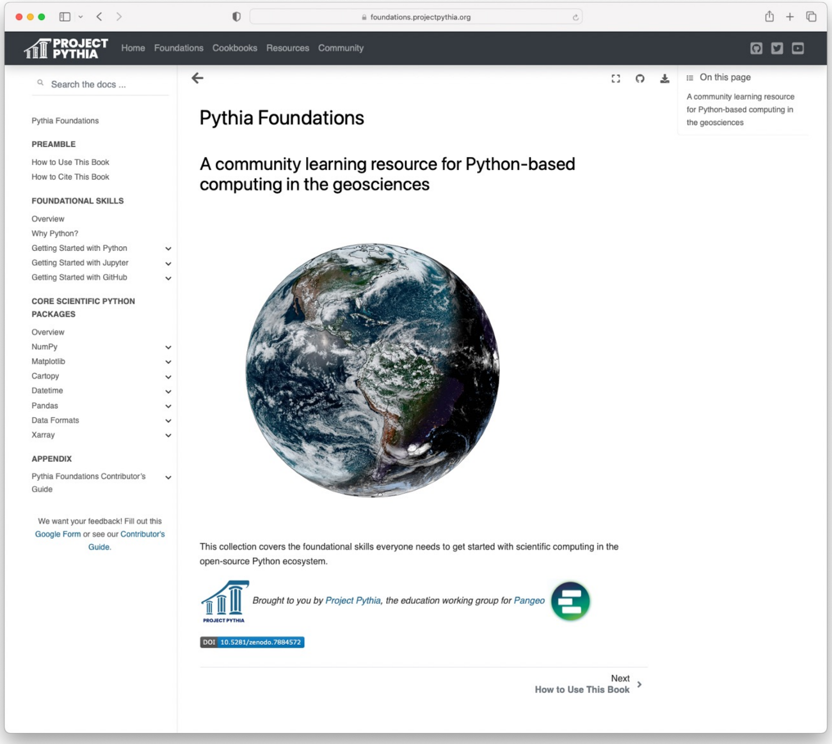
PROJECT PYTHIA



The goto resource for learning the *Scientific Python Ecosystem*

- Geoscience focused
- From beginner to the power user
- Tutorials, videos, examples, on-line courses, and sample data
- Community owned

The Foundations book



PROJECT PYTHIA

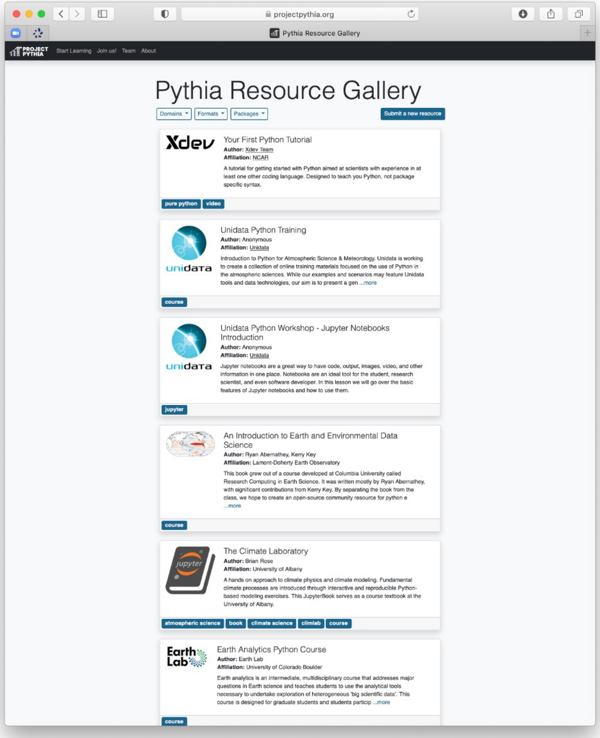
https://foundations.projectpythia.org/

*A comprehensive set of tutorials covering the **foundational skills** needed to get started with computing in the open-source Python ecosystem.*

*Serve as common references for more advanced and domain-specific **Cookbooks***

*All tutorials are **Binderized** for exploratory learning*

The Resource Gallery

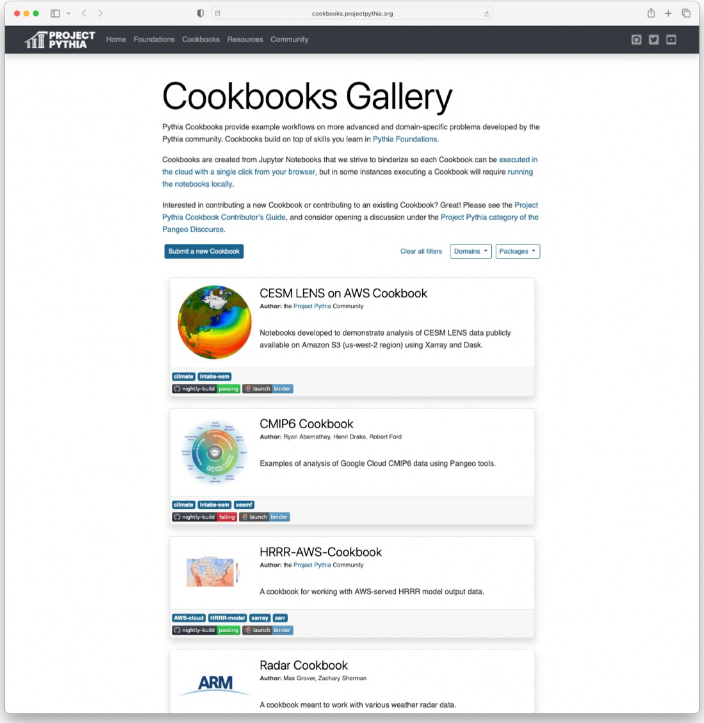


https://projectpythia.org/resource-gallery.html

A curated, searchable, and extensible gallery of links to external learning resources

An attempt to "beat Google" by gathering together relevant resources for the geosciences

Pythia Cookbooks



https://cookbooks.projectpythia.org

PROJECT PYTHIA

Cookbooks are community-contributed collections of advanced or domain-specific tutorials and example workflows

Essential features of Pythia Cookbooks:

- ➢ *explicitly build upon Foundations*
- ➢ *Demonstrate real workflows on publicly available data*
- ➢ *Backed by **automated testing infrastructure** to ensure that the example code "just works" and stays relevant*
- ➢ *Binderized for interactive learning*

Cookbooks are meant to serve as starting points for new geoscience analysis using the Python stack

Why Cookbooks?

What problems are Cookbooks trying to solve?

What problems are Cookbooks trying to solve?

Jupyter Notebooks are awesome, but…

PROJECT PYTHIA

- **Ambiguity**: Jupyter notebooks don't fully describe their own execution environment
- **Obsolescence**: most Notebooks found "in the wild" will not run and/or will not reproduce themselves
- **Collaboration**: Notebooks don't play very well with GitHub pull requests
- **Findable and Accessible**: using Notebooks to share knowledge about scientific workflows requires an audience!
- **Scalability**: tutorials that run in a limited sandbox don't offer clearest paths to doing new science on real data

What problems are Cookbooks trying to solve?

Jupyter Notebooks are awesome, but…

PROJECT PYTHIA

- **Ambiguity**: Jupyter notebooks don't fully describe their own execution environment

jupyter {book}

binder

A great tool for packaging Notebooks and conda environment descriptions into easy-to-navigate Web pages, with Binder links for execution

What problems are Cookbooks trying to solve?

Jupyter Notebooks are awesome, but…

PROJECT PYTHIA

- **Obsolescence**: most Notebooks found “in the wild” will not run and/or will not reproduce themselves

GitHub Actions

We need a CI service that can perform regular “health-checking” of notebook code!

What problems are Cookbooks trying to solve?

Jupyter Notebooks are awesome, but...

PROJECT PYTHIA

- **Collaboration**: Notebooks don’t play very well with GitHub pull requests

GitHub Actions

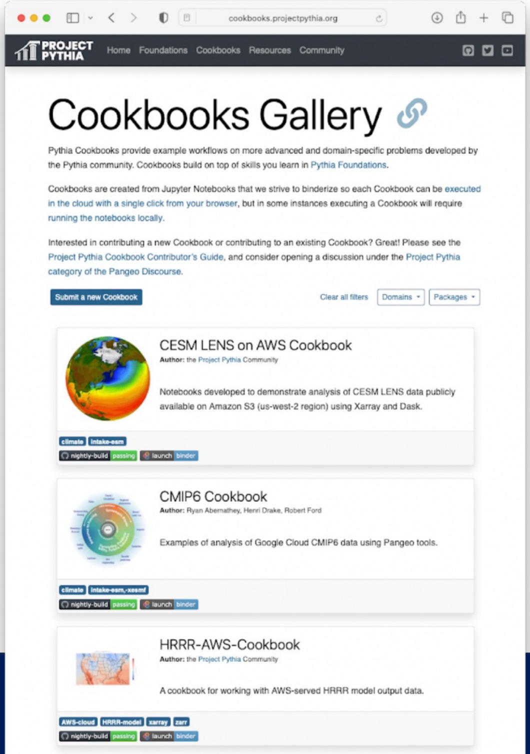
GitHub Pages

We need to execute notebooks and generate + deploy a preview of the rendered book to facilitate review and merge cycles

What problems are Cookbooks trying to solve?

Jupyter Notebooks are awesome, but…

PROJECT PYTHIA



- **Findable and Accessible**: using Notebooks to share knowledge about scientific workflows requires an audience!

  We should have a community repository for sharing workflows!
  And it should be organized and filterable

What problems are Cookbooks trying to solve?

Jupyter Notebooks are awesome, but…

PROJECT PYTHIA

We need to be able to route notebook execution to the appropriate compute resource for its content!

binder

- **Scalability**: tutorials that run in a limited sandbox don’t offer clearest paths to doing new science on real data

Pythia Cookbooks, a complete pipeline for reproducible, self-publishing notebooks

Got a cool workflow to demonstrate?

- Clone the Cookbook Template repo
- Commit some unexecuted notebooks
- Edit a short list of config files:
  - TOC for the book
  - URL for the appropriate BinderHub service
  - environment.yml for dependencies
- Switch on GitHub pages for your repo
- **That's it!** Your Cookbook should build and publish itself, with previews for all PRs

Want to host your Cookbook on the Pythia Gallery?

- Transfer the repo to the ProjectPythia org
- Open a PR on the cookbook-gallery with some simple filter tags (or just reach out!)

Radar Cookbook Improvements: Including the Rest of the Ecosystem

Search the docs ...

Radar Cookbook

FOUNDATIONS

Py-ART Basics

Py-ART Corrections

Py-ART Gridding

EXAMPLE WORKFLOWS

Looking at NEXRAD Data from Moore, Oklahoma

Plotting Data from a Field Campaign (TRACER)

Data Quality Check from the CACTI Field Campaign

Theme by Project Pythia.

ARM

Radar Cookbook

nightly-build failing

This Project Pythia Cookbook covers the basics of working with weather radar data in Python.

Motivation

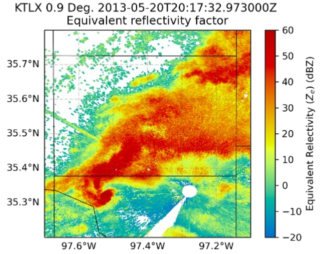
This cookbook provides the essential materials to learning how to work with weather radar data using Python.

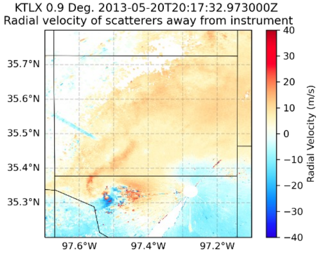
Most of the curriculum is focused around the Python ARM Toolkit, which is defined as:

“a Python module containing a collection of weather radar algorithms and utilities. Py-ART is used by the Atmospheric Radiation Measurement (ARM) user facility for working with data from a number of its precipitation and cloud radars, but has been designed so that it can be used by others in the radar and atmospheric communities to examine, processes, and analyze data from many types of weather radars.”

Once you go through this material, you will have the skills to read in radar data, apply data corrections, and visualize your data, building off of the core foundational Python material covered in the Foundations Book

radlib





Xradar

Zarr

Example workflow – 2D objective analysis CSAPR2 RHI sweep

Setup our Download Query

Before downloading our data, we need to make sure we have an ARM Data Account, and ARM Live token. Both of these can be found using this link:

- ARM Live Signup

Once you sign up, you will see your token. Copy and replace that where we have `arm_username` and `arm_password` below.

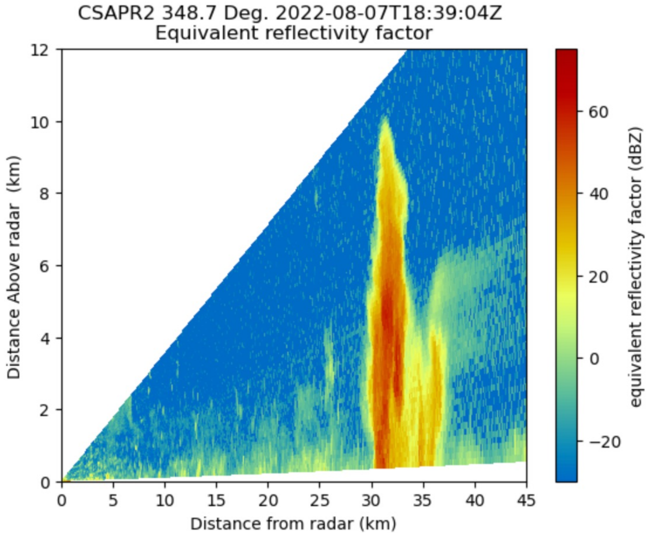
```
arm_username = os.getenv("ARM_USERNAME")
arm_password = os.getenv("ARM_PASSWORD")

datastream = "houcsapr2cfrS2.a1"

start_date = "2022-08-07T18:39:04"
end_date = "2022-08-07T18:39:05"
```

Plot one of the RHI scans

We read in the data corresponding to 7 August 2022 18:39:04 UTC, and plot a basic `RadarDisplay` which will automatically detect whether the plot is a vertical cross section (RHI or VPT), or a horizontal scan (PPI)

```
radar = pyart.io.read(radar_file)
display = pyart.graph.RadarDisplay(radar)
display.plot("reflectivity", 0)
plt.savefig(f"quicklooks/{Path(radar_file).stem}.png", dpi=200)
plt.xlim(0,45)
plt.ylim(0,12)
plt.show()
plt.close()
```

```
/srv/conda/envs/notebook/lib/python3.11/site-packages/numpy/core/fromnumeric.py:784: UserWa
  a.partition(kth, axis=axis, kind=kind, order=order)
```



Define a function to grid the RHI data from polar (antenna) coordinates to a two-dimensional Caretsian grid

We use numba to vectorize the `dist_func` function to calculate the distance of each range gate from the radar. This makes our code run faster than simply executing this function for each gate in a for loop.

Next, we use the `barnes` function from the fastbarnes Python package to interpolate the radar fields such as `equivalent reflectivity factor` (Z_H), `differential_reflectivity` (Z_{DR}), and `specific_differential_phase` (K_{DP}) to a uniform range-height Cartesian grid.

Geosci. Model Dev., 16, 1697–1711, 2023
https://doi.org/10.5194/gmd-16-1697-2023

Geoscientific
Model Development
Open Access

https://github.com/MeteoSwiss/fast-barnes-py

Fast approximate Barnes interpolation: illustrated by Python-Numba implementation fast-barnes-py v1.0

Bruno K. Zürcher

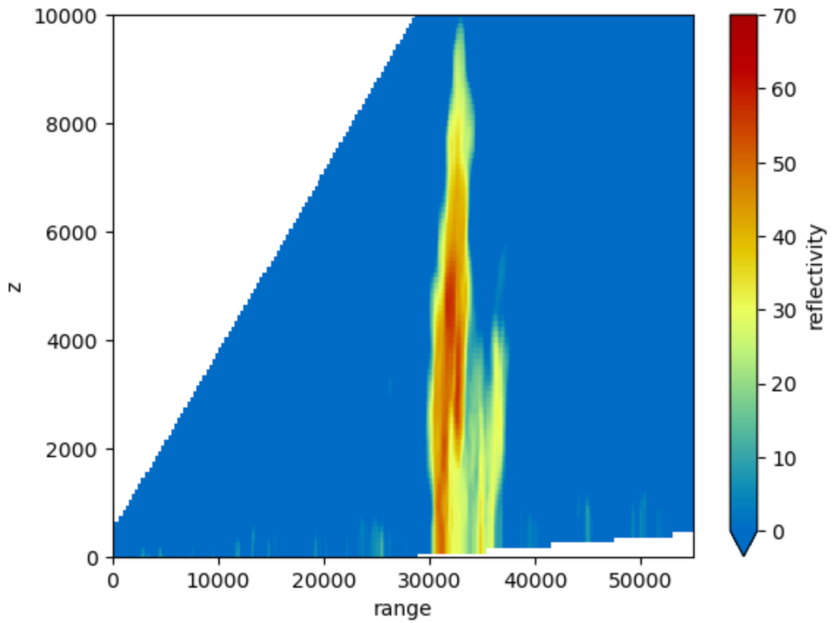
Federal Office of Meteorology and Climatology MeteoSwiss, Zurich, Switzerland

Correspondence: Bruno K. Zürcher (bruno.zuercher@meteoswiss.ch)

Received: 24 April 2022 – Discussion started: 19 July 2022
Revised: 14 January 2023 – Accepted: 16 January 2023 – Published: 27 March 2023

Grid and plot the RHI data

```
# Finally, plot the gridded reflectivity
fig,ax = plt.subplots()
grid_ds['reflectivity'].plot(vmin=0,vmax=70,cmap='pyart_HomeyerRainbow',ax=ax)
ax.set_xlim(0,55000)
ax.set_ylim(0,10000)
plt.show()
```



Echo top height calculation from NEXRAD PPI volume data:

An echo top is the radar indicated top of an area of precipitation. This notebook demonstrates how to calculate the echo top height (ETH) in a NEXRAD PPI volume scan to determine the maximum elevation angle at which a certain reflectivity threshold is exceeded.

This example uses the echo top height (ETH) calculation code written by Valentin Louf, available at this github repository.

---

https://projectpythia.org/radar-cookbook/notebooks/example-workflows/echo_top_height.html