Tutorial: Python for Data Science on **ACES**

Richard Lawrence

2025/04/11

High Performance
Research Computing
DIVISION OF RESEARCH

Outline

- Overview of Python for Data Science
- Getting Started
- Arrays and Series
- Plots

Overview of Python for Data Science

Arrays

Numpy Arrays support common operations, such as arithmetic, on an element-by-element (or “vectorized”) basis.

Example:

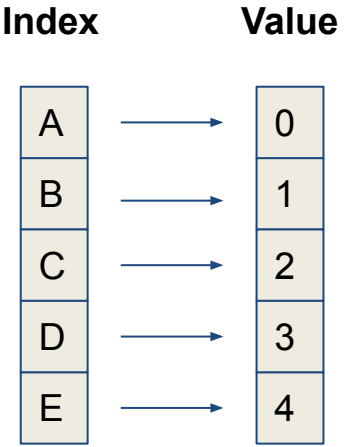
```
array C = array A + array B
```

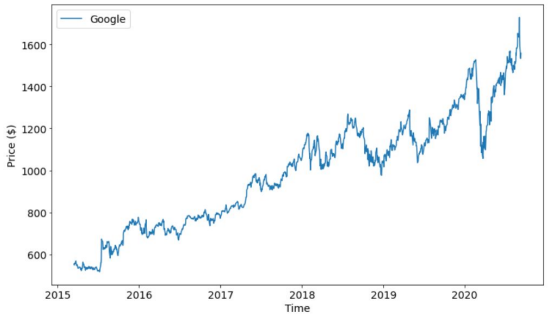
This adds the elements of A and B pair-wise (Instead of concatenating the elements as would happen with lists).

Pandas Series and DataFrames further expand on this.

Series

- One-dimensional labeled array
- Capable of holding any data type (integers, strings, floating point numbers, etc.)
- Example: time-series stock price data





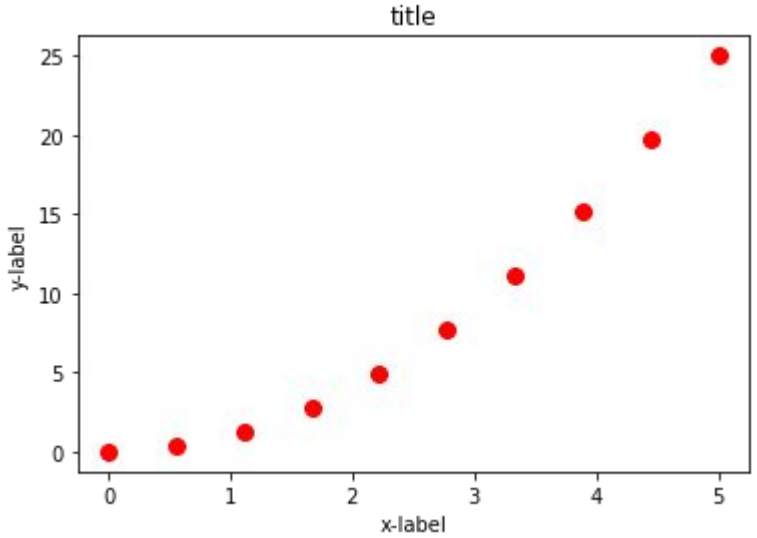
Anatomy of a Plot

Marker

- style
- size
- color

Figure

- title
- xlabel
- ylabel

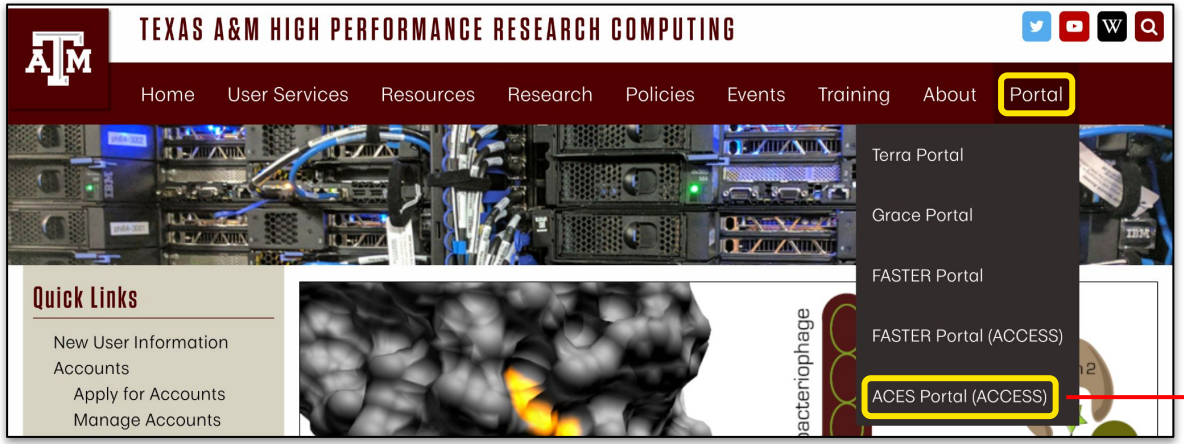


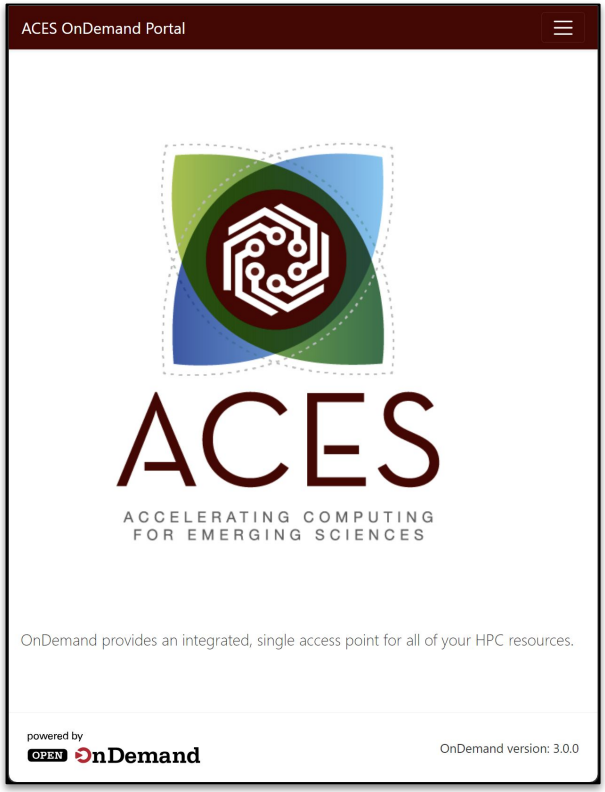
Learning Resources

- Slides on the course web page
  https://hprc.tamu.edu/training/aces_python4data.html
- HPRC's Knowledge Base
  https://hprc.tamu.edu/kb/Software/Python/
- HPRC on YouTube
  https://www.youtube.com/c/TexasAMHPRC
- ACCESS Links
  https://support.access-ci.org/ci-links

Getting Started

ACES Portal

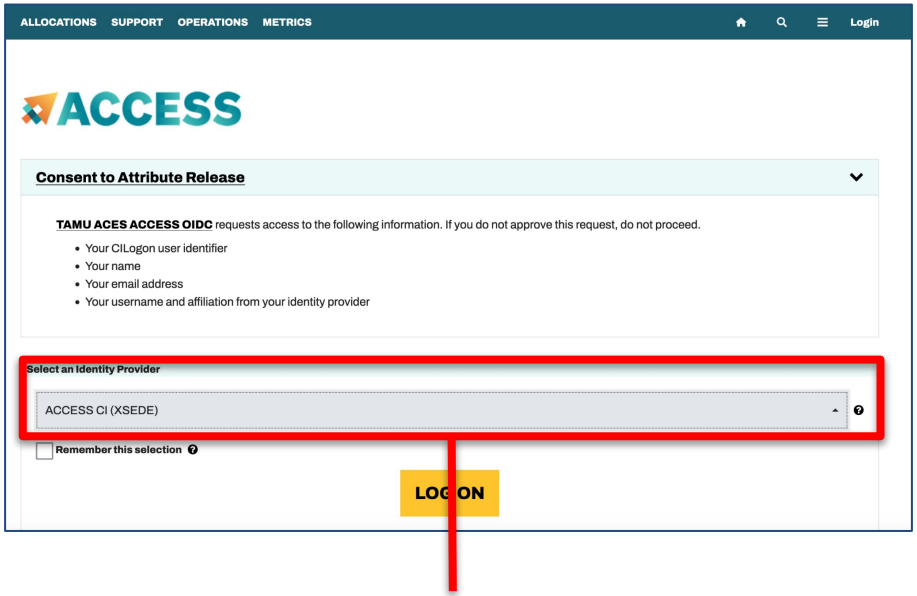




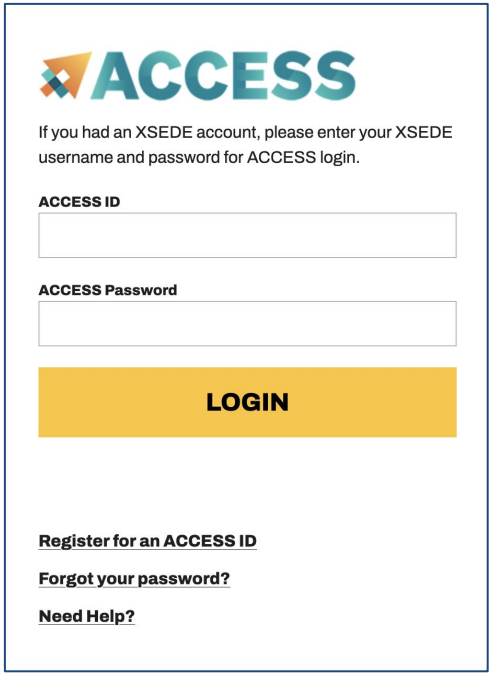
ACES Portal portal-aces.hprc.tamu.edu
is the web-based user interface for the ACES cluster

Open OnDemand (OOD) is an advanced web-based graphical interface framework for HPC users

Accessing ACES via the Portal (ACCESS)

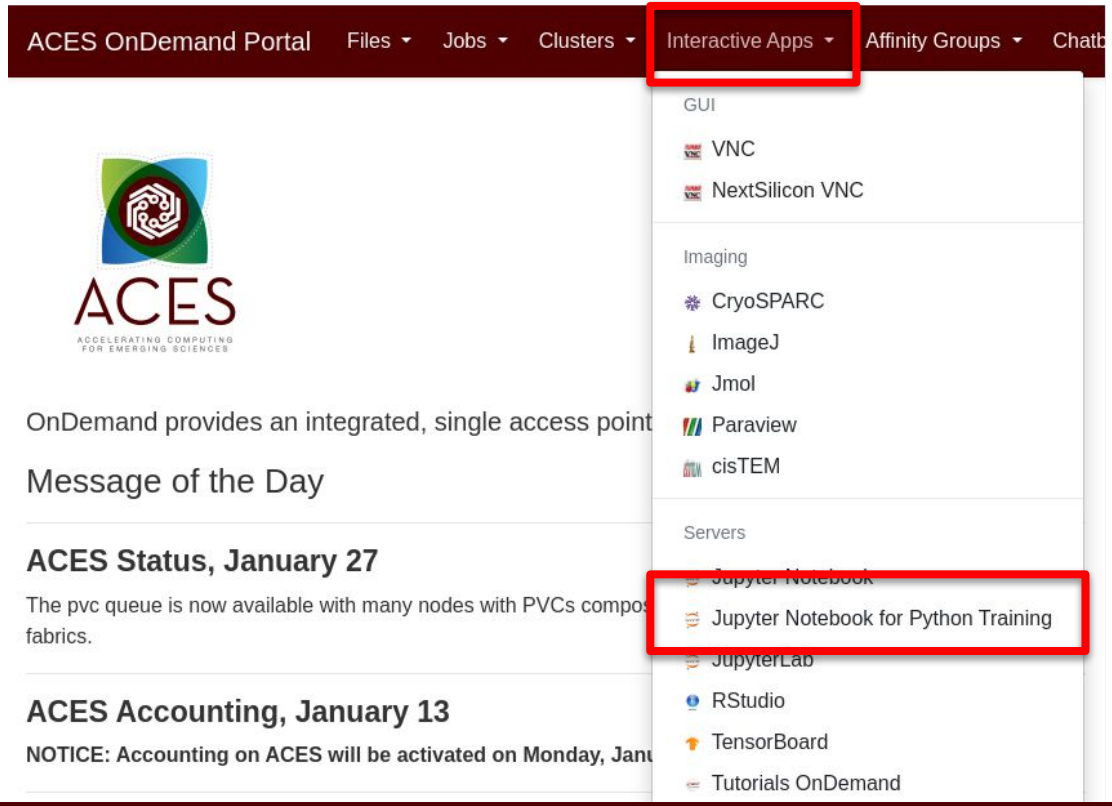


Select the Identity Provider appropriate for your account.

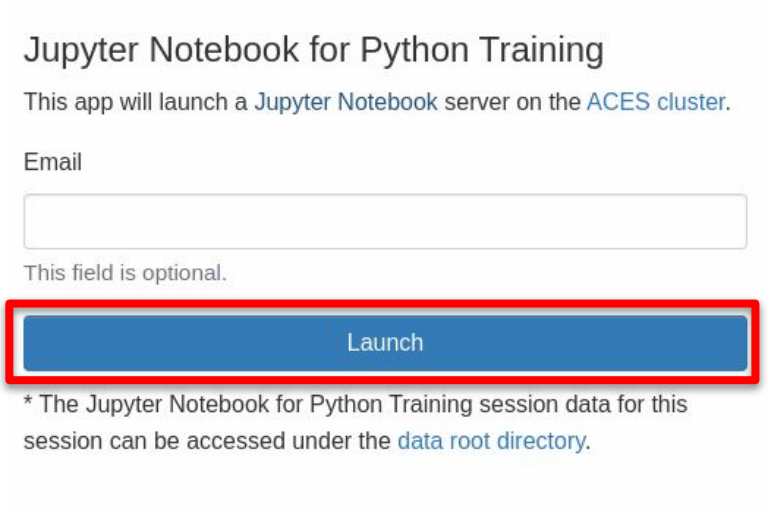


Log-in using your ACCESS or institutional credentials.

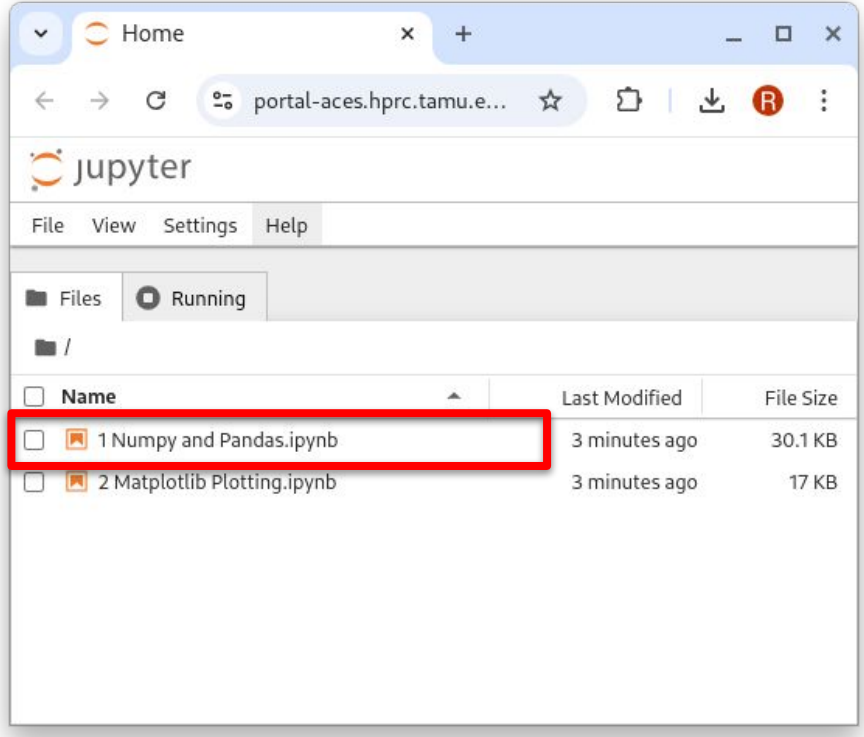
Jupyter Notebook for Python Training

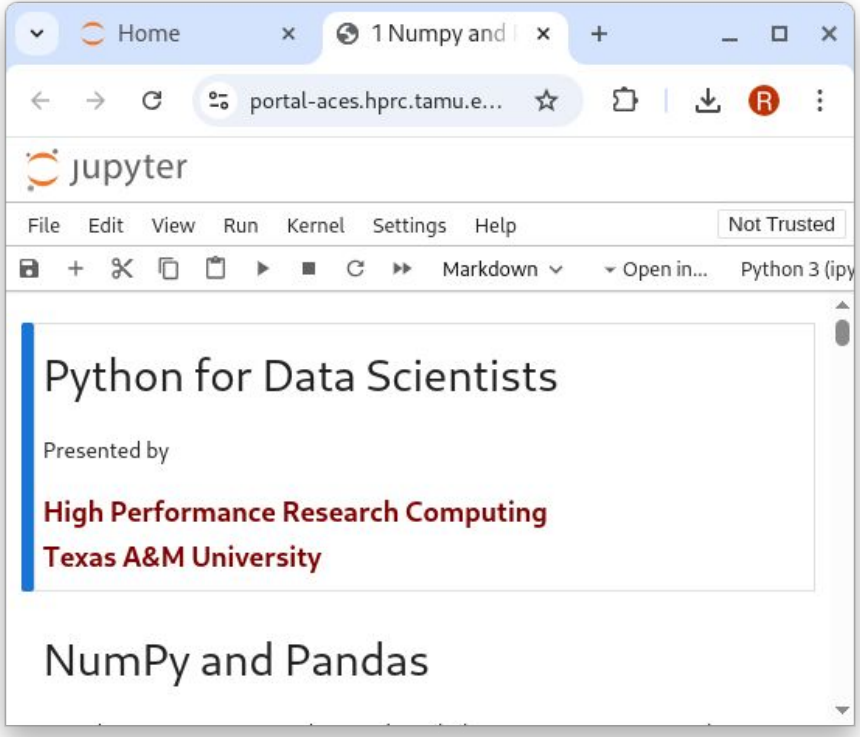


Jupyter Notebook for Python Training



Jupyter Notebook for Python Training





HIGH PERFORMANCE RESEARCH COMPUTING
TEXAS A&M UNIVERSITY

https://hprc.tamu.edu

HPRC Helpdesk:

help@hprc.tamu.edu
Phone: 979-845-0219

Take our short course survey!

HPRC Survey

https://u.tamu.edu/hprc_shortcourse_survey

Help us help you. Please include details in your request for support, such as, Cluster (ACES, FASTER, Grace, Launch), NetID (UserID), Job information (JobID(s), Location of your jobfile, input/output files, Application, Module(s) loaded, Error messages, etc), and Steps you have taken, so we can reproduce the problem.