Intro to Python

Fuzzy Rogers
Research Computing Administrator
Materials Research Laboratory (MRL)
Center for Scientific Computing (CSC)
fuz@mrl.ucsb.edu
MRL 2066B

Sharon Solis
Research Computing Consultant
Enterprise Technology Services (ETS)
Center for Scientific Computing (CSC)
swsolis@ucsb.edu
Elings Hall 3229

Paul Weakliem
CNSI Research Computing Support
California NanoSystems Institute (CNSI)
Center for Scientific Computing (CSC)
weakliem@cnsi.ucsb.edu
Elings Hall 3231

Other Research IT

- Ted cabeen (Life Sciences/Chemistry)
- Michael colee (eri)
- Steve Miley (bren)
- Jim Woods (MSI)
- LSIT (Letters & Sciences)
- Glen Schiferrel (Physics)

Pre-class Instructions:

Python

Python is a popular language for research computing, and great for general-purpose programming as well. Installing all of its research packages individually can be a bit difficult, so we recommend **Anaconda:** **https://www.anaconda.com/distribution**, an all-in-one installer.

Regardless of how you choose to install it, please make sure you install Python version 3.x (e.g., 3.7 is fine). Here are further instructions on how to install Anaconda on Windows, Mac or Linux:
http://carpentries.github.io/workshop-template/#python

Obtain lesson materials

1. Download Data Set:
   http://swcarpentry.github.io/python-novice-inflammation/data/python-novice-inflammation-data.zip
2. Download Python Code:
   http://swcarpentry.github.io/python-novice-inflammation/code/python-novice-inflammation-code.zip
3. Create a folder called swc-python on your Desktop.Move downloaded files into this newly created folder.
4. Unzip the files. You should now see two new folders called data and code in your swc-python directory on your Desktop.

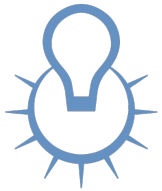
Post-It Notes

- **Red**: Help needed
- **Green**: Good to go

What is python™

- High level general purpose language, using more natural English language to emphasize code readability
- Created by Guido van Rossum in the Netherlands, first released in 1991
- Core philosophy, *The Zen of Python*:
  - Beautiful is better than ugly
  - Explicit is better than implicit
  - Simple is better than complex
  - Complex is better than complicated
  - Readability counts

Who uses Python?



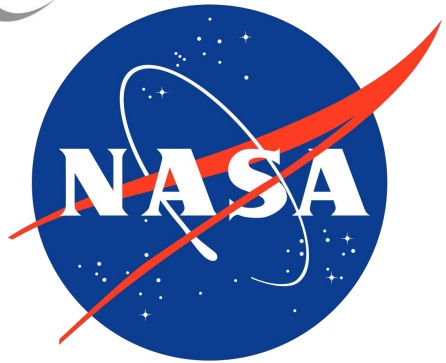
INDUSTRIAL
LIGHT & MAGIC

Google

Journyx

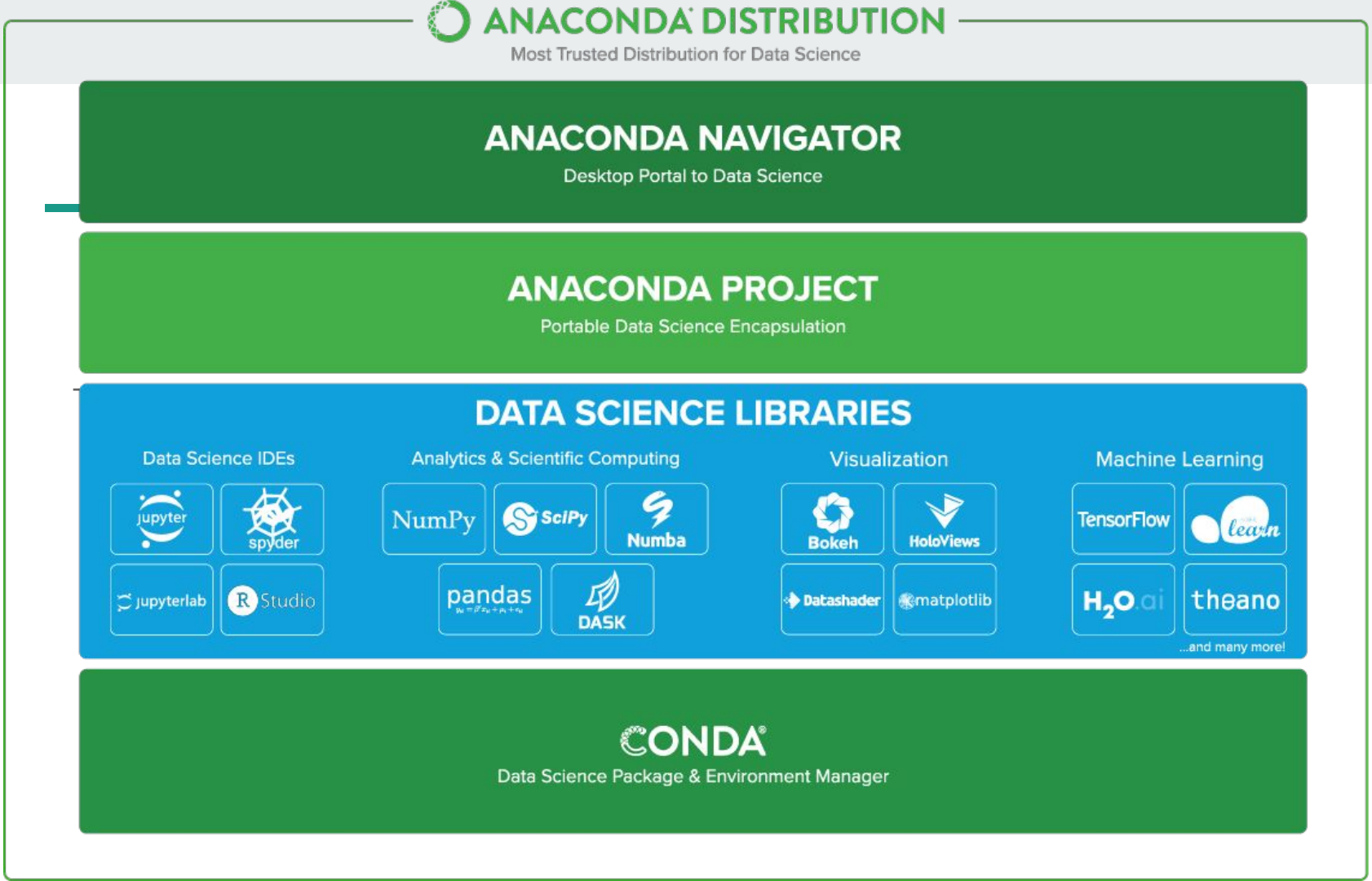
IBM

YouTube



Why Use Python?

- Open-source
- Easy to learn, good first language (great foundation for C, C++, Java)
- A large community of users, scientists, and developers
- Lots of libraries spanning many disciplines:
    - **Bioinformatics**: Biopython
    - **Numerical and Scientific computing:** Numpy, Scipy, Dask
    - **Statistics**: Scipy, Statmodels
    - **Visualization**: Matplotlib, Seaborn, Bokeh
    - **Data analysis**: Pandas
    - **Machine learning:** Scikit-learn
    - **Image processing**: OpenCV, Skimage
    - **Deep learning**: Tensorflow, Theano, Pytorch,….


ANACONDA DISTRIBUTION
Most Trusted Distribution for Data Science
ANACONDA NAVIGATOR
Desktop Portal to Data Science
ANACONDA PROJECT
Portable Data Science Encapsulation
DATA SCIENCE LIBRARIES
Data Science IDEs
jupyter
spyder
jupyterlab
RStudio
Analytics & Scientific Computing
NumPy
SciPy
Numba
pandas
DASK
Visualization
Bokeh
HoloViews
Datashader
matplotlib
Machine Learning
TensorFlow
learn
H2O.ai
theano
...and many more!
CONDA
Data Science Package & Environment Manager

Post-It Notes

- **Red**: Help needed
- **Green**: Good to go

Data Set

- Data Set, Code and Lesson Material available here: http://swcarpentry.github.io/python-novice-inflammation/

What is a Script?

- How to run code
- Save yourself work!
- Don’t need to type over and over again
- Move easily between machines

Using Python on a Cluster

- Use R not RStudio on the cluster
- Make sure your R code runs from start to end on your own machine
- Perform tests on your computer first
- A simple script (text file) can be used to submit to the queue:

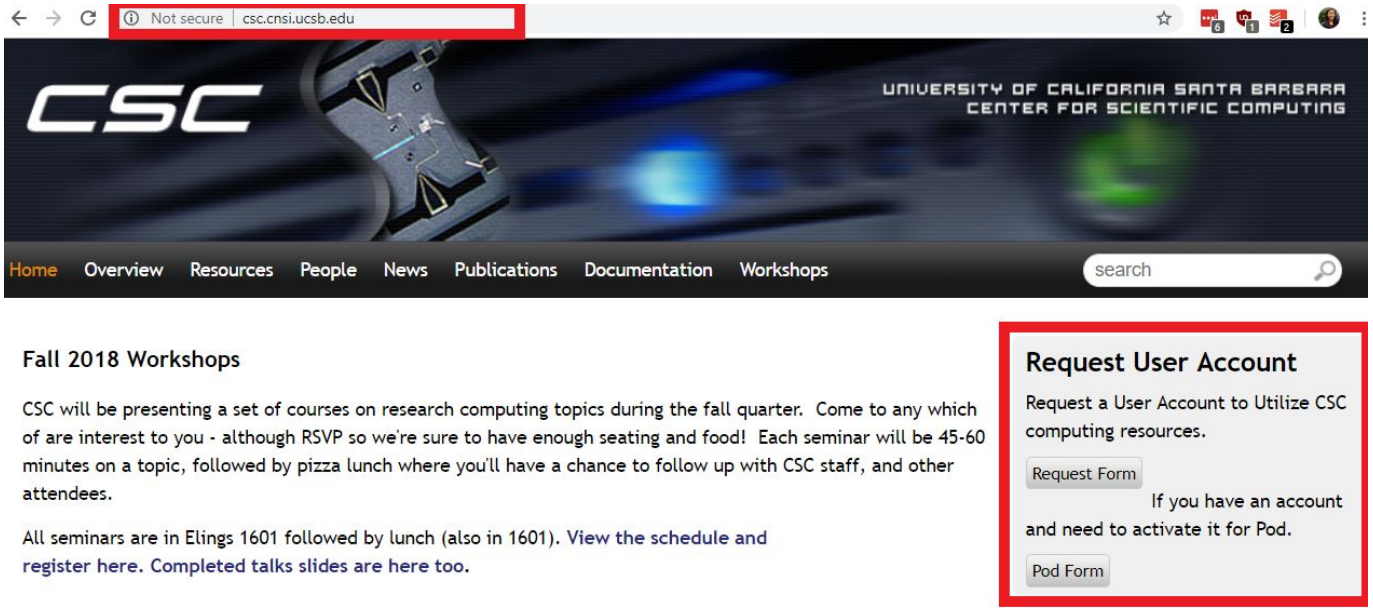
```
#!/bin/bash
#SBATCH -N 1 -n 1
cd $SLURM_SUBMIT_DIR
python myfile.py
```

Using Python on the Cluster

```
#!/bin/bash
#SBATCH -N 1 -n 1
cd $SLURM_SUBMIT_DIR
python myfile.py
```

#!/bin/bash : the shell you are using
#SBATCH -N 1 -n 1 : Asking for one node and one task per node
cd $SLURM_SUBMIT_DIR : change directory to the one where job is submitted from
Python myfile.py : Run your python myfile.py code

How to Request a User Account



How to Learn More Python

- Online Tutorials:
  - learnpython.org
  - Principles of Computing in Python from Carnegie Mellon's Open Learning Initiative
    - You can click the "Enter Course" button to take a look at the course without creating an account
    - If you want to save your work, you can create an account.
  - Lynda.com (available to UCSB employees, including student employees)
- One-on-One Consultation
  - Center for Scientific Computing (Elings Hall 3229)
  - Collaboratory
- Books
  - UCSB students, staff and faculty (i.e. anyone with a UCSBNetID, or access to computer on campus), you can get free access to a zillion great Computing related textbooks here http://proquest.safaribooksonline.com/

Contact Us

csc.cnsi.ucsb.edu

Sharon Solis
Research Computing Consultant
Enterprise Technology Services (ETS)
Center for Scientific Computing (CSC)
swsolis@ucsb.edu
Elings Hall 3229

Fuzzy Rogers
Research Computing Administrator
Materials Research Laboratory (MRL)
Center for Scientific Computing (CSC)
fuz@mrl.ucsb.edu
MRL 2066B

Paul Weakliem
CNSI Research Computing Support
California NanoSystems Institute (CNSI)
Center for Scientific Computing (CSC)
weakliem@cnsi.ucsb.edu
Elings Hall 3231