NumPy

For Beginners

Prepared for:

Boston Python User Group https://about.bostonpython.com/

Teach to Learn Data Science Study Group Series.

Glenn Lehman

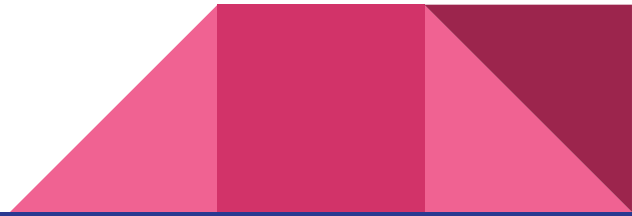
https://www.linkedin.com/in/glenndlehman/

Who is Glenn? A Few Weird Things to Know

- Working in programing since 1984
- Been to the north pole on a nuclear sub
- Familiar with SQL, VB, and Python
- Fun project - EZ Pass, Bill Payment Processing
- Spent a good amount of time cleaning data and designing data structures
- Finally in college to get a degree
- 4 kids x 4 grandkids

Agenda

- What is NumPy (overview and where does it fit in to Data Science)
- Installing NumPy (pip and conda)
- Difference between Python List and NumPy Array
- Creating Arrays (Vector and Matrix)
- Reshaping, Resizing, Indexing, and Slicing
- Simple Math
- Linear Regression Analysis (Start)
- Review Additional Resources
- Questions



What is NumPy

NumPy is the fundamental package for scientific computing in Python.

Provides a multidimensional array object, various derived objects (such as masked arrays and matrices), and an assortment of routines for fast operations on arrays

A growing plethora of scientific and mathematical Python-based packages are using NumPy arrays.

FAST - Vectorized code

Numpy in the World of Data Science

Nearly every scientist working in Python draws on the power of NumPy.

NumPy brings the computational power of languages like C and Fortran to Python, a language much easier to learn and use. With this power comes simplicity: a solution in NumPy is often clear and elegant.

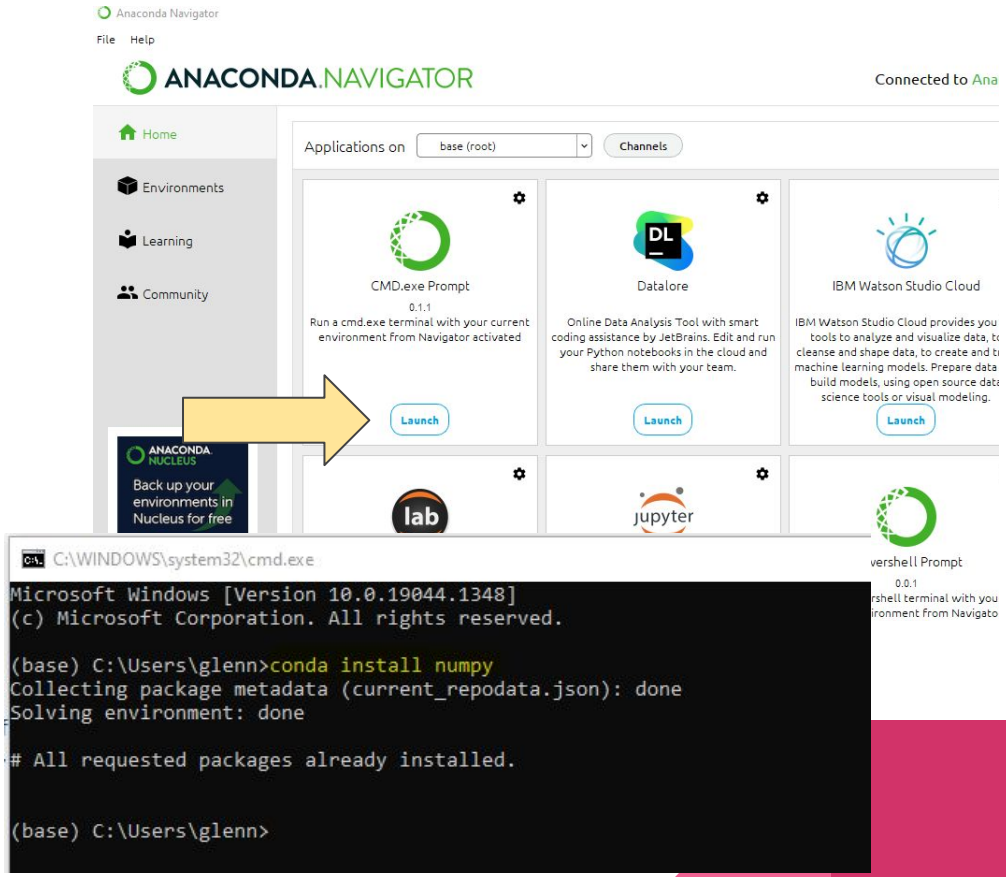
Quantum Computing	Statistical Computing	Signal Processing	Image Processing	Graphs and Networks	Astronomy Processes	Cognitive Psychology
QuTiP PyQuil Qiskit	Pandas statsmodels Xarray Seaborn	SciPy PyWavelets python-control	Scikit-image OpenCV Mahotas	NetworkX graph-tool igraph PyGSP	AstroPy SunPy SpacePy	PsychoPy

Bioinformatics	Bayesian Inference	Mathematical Analysis	Chemistry	Geoscience	Geographic Processing	Architecture & Engineering
BioPython Scikit-Bio PyEnsembl ETE	PyStan PyMC3 ArviZ emcee	SciPy SymPy cvxpy FEniCS	Cantera MDAnalysis RDKit	Pangeo Simpeg ObsPy Fatiando a Terra	Shapely GeoPandas Folium	COMPAS City Energy Analyst Sverchok

Numpy Install

1. Launch Anaconda Navigator
2. Launch CMD.exe Prompt
3. Create your virtual environment
4. Enter command:

conda install numpy



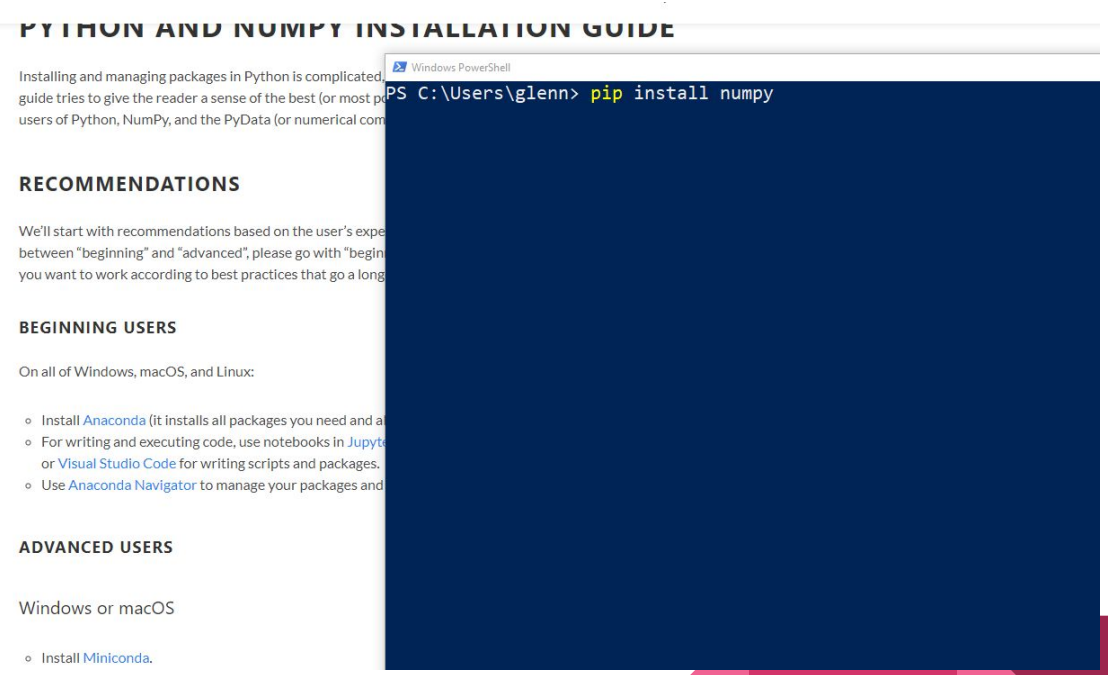
Numpy - Pip Install

From the windows powershell

pip install numpy

Installation for other operating systems and advance options are can be found at:

https://numpy.org/install/



Arrays - Vocabulary Lesson

ndarray	Shorthand for “N-dimensional array.” An N-dimensional array is simply an array with any number of dimensions
1-D	One-dimensional array
2-D	Two-dimensional array (And so on)
Vector	One-dimensional array
Matrix	Two-dimensional array
Tensor	3-D or Higher
Axis	A specific dimension in the array
Shape	Tuple of non negative integers specify the size of each dimension.

Attribute List:
https://numpy.org/doc/stable/reference/arrays.ndarray.html#arrays-ndarray

Python List vs Numpy Array

Python List

- Grow dynamically
- Multiple data types

Example

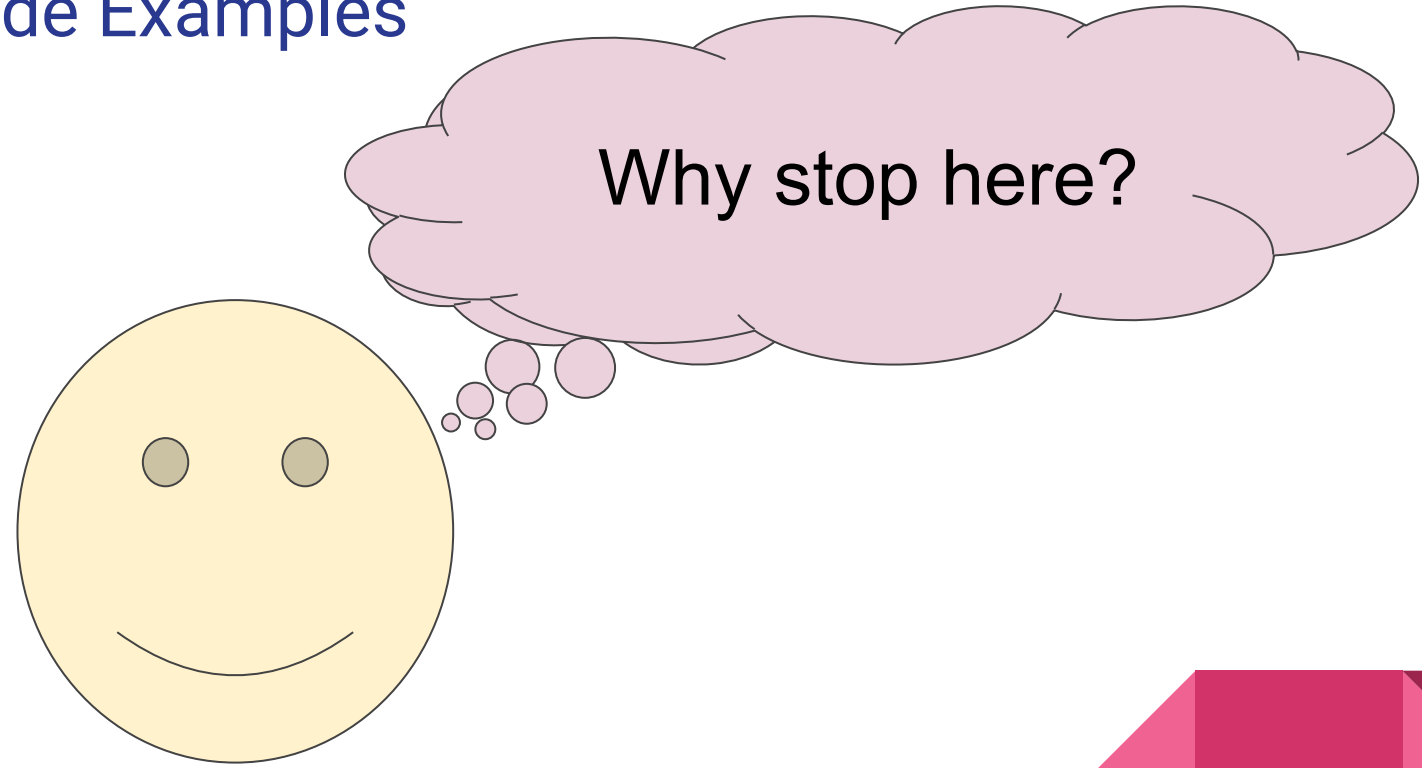
```
c = []
for i in range(len(a)):
    c.append(a[i]*b[i])
```

Numpy Array

- Fixed size at creation
- Same data type (same size in memory)*
- Facilitate advanced mathematical on large numbers of data (Less code execute more efficiently)

```
c = a * b
```

Real Code Examples



References and Resources

NumPy

NumPy: https://numpy.org/

NumPy API Reference: https://numpy.org/doc/stable/reference/index.html

NumPy Cheat Sheet: https://towardsdatascience.com/numpy-cheat-sheet-4e3858d0ff0e

NumPy Illustrated: The Visual Guide to NumPy: https://betterprogramming.pub/numpy-illustrated-the-visual-guide-to-numpy-3b1d4976de1d

Introduction to Numerical Computing with NumPy | SciPy 2019 Tutorial: https://www.youtube.com/watch?v=ZB7BZMhfPgk

Pure Python vs NumPy vs TensorFlow Performance Comparison: https://realpython.com/numpy-tensorflow-performance/

Real Python Tutorial: Your First Steps Into Data Science in Python https://realpython.com/numpy-tutorial/

Numpy Full Course: https://youtu.be/j31ah5Qa4QI

Statistics

StatQuest (Josh Starmer): https://statquest.org/

Stats 101 (Brandon Foltz): https://www.bcfoltz.com/blog/

Questions?