General Assembly

STUDY GUIDE

NUMPY AND SCIPY

Key Terms and Definitions

» **Import:** The import statement loads a library into a Python script for use throughout.

» **Library:** A bundle of Python modules, which are the individual pieces of code that contain definitions and statements. Essentially, a library brings in a set of classes (with associated methods and attributes) and functions that are used for a common purpose.

» **Anaconda:** A Python package manager.

» **pip install:** A manual alternative to package management used in the command line.

» **NumPy (pronounced num-pi):** A fundamental package for scientific computing with Python, including powerful vector and matrix creation, array math operations, and a rich set of advanced math operations and random number-generation capabilities.

- **np.random:** A module within NumPy used to generate random numbers and variables.

» **Random Values:** Encountered when it is impossible to predict future values based on previous ones. A random variable contains random values.

- **Discrete** random variables have outcomes that are "countable."
- **Continuous** random variables have outcomes that are not countable. They are still numeric but could be an infinite range of values.

» **SciPy:** Contains more fully featured versions of the linear algebra modules than NumPy, as well as many other numerical algorithms. SciPy is easy to use, computationally optimized, and able to help us do heavy lifting. If you are doing scientific computing with Python, we recommend you install both NumPy and SciPy.

» **Array:** An ordered set of objects with one or more dimensions. A one-dimensional array is a list (think of a column of information), a two-dimensional array would be a list of lists (think of rows within a table, each of which has information in each column), and so on. These get harder to visualize as we expand into higher dimensions.

» **Nested Lists:** The "Pythonic" version of an array. NumPy contains array objects, while Python itself presents these simply as lists within lists.

» **Vectors:** A sequence of numbers arranged as a one-dimensional array. We can use linear algebra to perform additions and other operations with vectors and matrices.

- **Matrices:** Two-dimensional arrays.
- **Multi-Dimensional Arrays:** Arrays with three or more dimensions.
- **scikit-learn:** A library built on NumPy, SciPy, and Matplotlib that seeks to make machine learning and data science models accessible and intuitive.

Guiding Questions

1. Would coin flips constitute a discrete or continuous random value?
2. Why are arrays so critical for performing mathematical operations in code?
3. What is the benefit of having several different math-based libraries for Python?

Additional Resources

1. GA NumPy/SciPy Demo Video
2. HackerRank Challenges
    - Challenge 1
    - Challenge 2
    - Challenge 3
3. DataCamp:
    - Intro to Python for Data Science. Specifically, see Section 4, "NumPy."