

9. Library - NumPy

Jin Wang

MSIS, Rutgers

- 9. Library - NumPy
 - 1. Introduction
 - 2. Create a basic array
 - 3. Indexing and slicing
 - 4. Operations - 1: Basic
 - 5. Operations - 2: Broadcasting
 - 6. Conditionals used in NumPy
 - 7. Statistical calculations
- Practice questions
 - References

1. Introduction

Install Numpy

Windows

```
bash    pip install numpy
```

Macbook

```
bash    pip3 install numpy
```

What is NumPy?

- NumPy stands for *Numerical Python*.
- NumPy is a Python library used for working with arrays.
- It also has functions for working in domain of linear algebra, fourier transform, and matrices.
- It is an open source project and you can use it freely.

Why Use NumPy?

- In Python we have lists that serve the purpose of arrays, but they are slow to process.
- *NumPy aims to provide an array object that is up to 50x faster than traditional Python lists.*
- The array object in NumPy is called `ndarray`, it provides a lot of supporting functions that make working with `ndarray` very easy.
- Arrays are very frequently used in data science, where speed and resources are very important.

2. Create a basic array

```
>>> import numpy as np
>>> np.array([1, 2, 3])
array([1, 2, 3])
```

3. Indexing and slicing

4. Operations - 1: Basic

5. Operations - 2: Broadcasting

6. Conditionals used in NumPy

7. Statistical calculations

- min
- max
- sum
- standard deviation - `std`
- percentiles (Q_1 , Q_2 (median), Q_3 ,) - `np.percentile(data, perc)`

Practice questions

1. Counting. Given a list, try to get the number of positive numbers.

2. Calculate the area of a circle, given a list of values of radius.

3. Given the following two lists, one includes the student names, another one includes the ages of the students. Try to get the students whose ages are greater than 20.

```
names = ['Jack', 'Mark', 'Mary', 'Jenny', 'April', 'Jin']  
ages = [23, 18, 21, 19, 22, 19]
```

4. Given the following grades about homework, project, exam_1 and exam_2, and the weights for the three assignments are 20%, 20%, 30%, and 30%, try to calculate the percentage of A (≥ 90), B (< 90 and ≥ 80), C (< 80 and ≥ 70), D (< 70 and ≥ 60), and F (< 60).

```
homework = [81, 83, 89, 98, 70, 71, 72, 91, 80, 61, 50]
project = [90, 92, 85, 82, 84, 86, 83, 79, 70, 81, 60]
exam_1 = [83, 70, 78, 90, 82, 88, 68, 59, 59, 75, 77]
exam_2 = [82, 73, 60, 65, 95, 88, 68, 59, 62, 75, 50]
```


References

- https://numpy.org/doc/stable/user/absolute_beginners.html
- https://www.w3schools.com/python/numpy/numpy_intro.asp

End