



Data Science With Python

DATA AND ARTIFICIAL INTELLIGENCE



Mathematical Computing with Python (NumPy)

Learning Objectives

By the end of this lesson, you will be able to:

- Explain NumPy and its importance
- O Discuss the basics of NumPy, including its fundamental objects
- Demonstrate how to create and print a NumPy array
- Analyze and perform basic operations in NumPy
- Utilize shape manipulation and copying methods
- Demonstrate how to execute linear algebraic functions
- Build basic programs using NumPy





Numpy

Quick Recap: Lists

Below are some of the properties of lists:

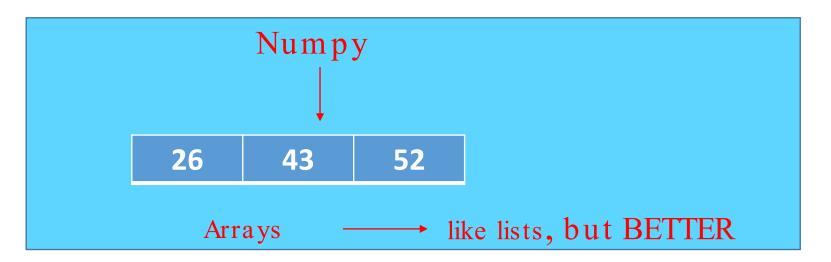
List

Limitations of Lists

You can change individual values in a list, but you cannot apply a mathematical operation over the entire list.

Why NumPy

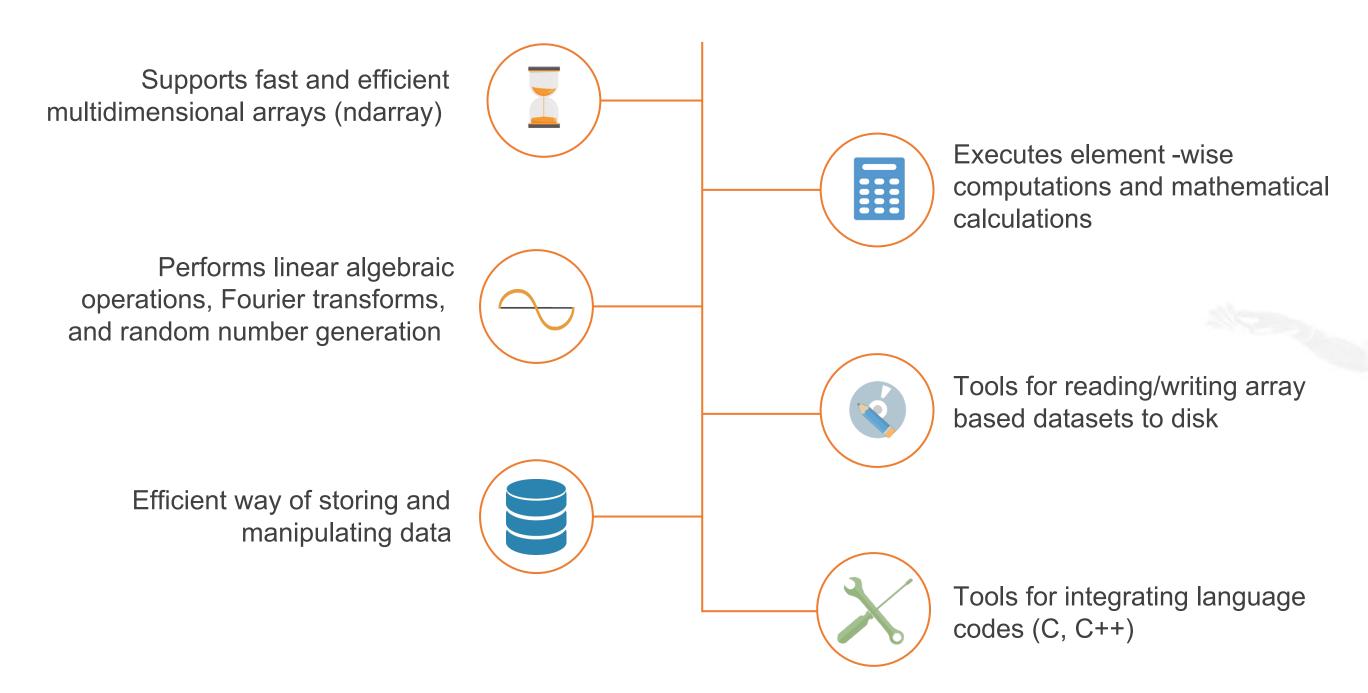
Numerical Python (NumPy) supports multidimensional arrays over which you can easily apply mathematical operations.



NumPy Overview

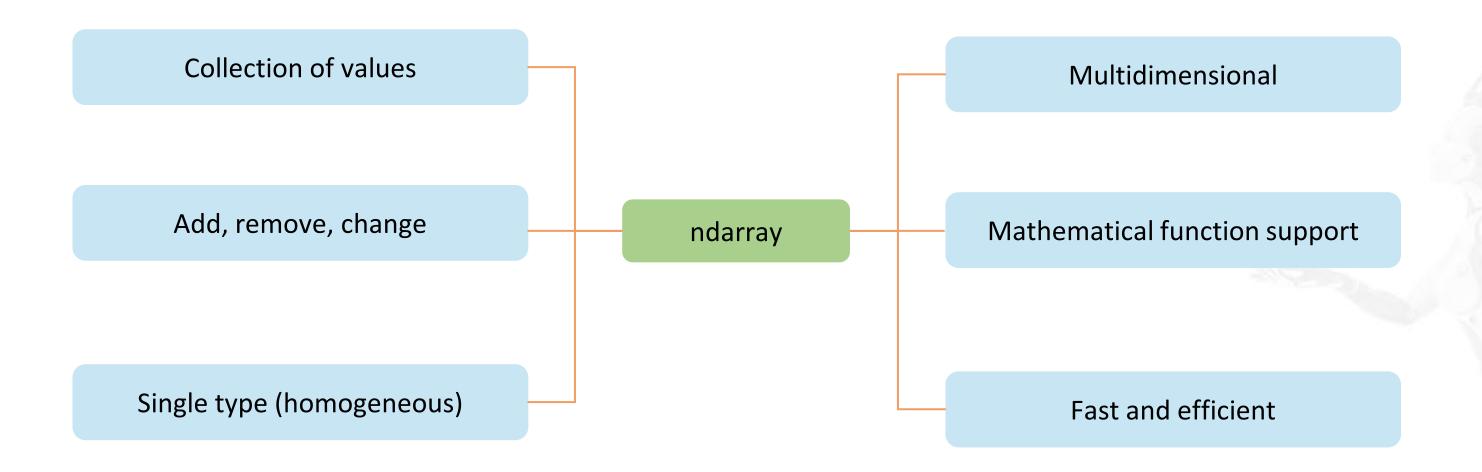
NumPy is the foundational package for mathematical computing in Python.

It has the following properties:



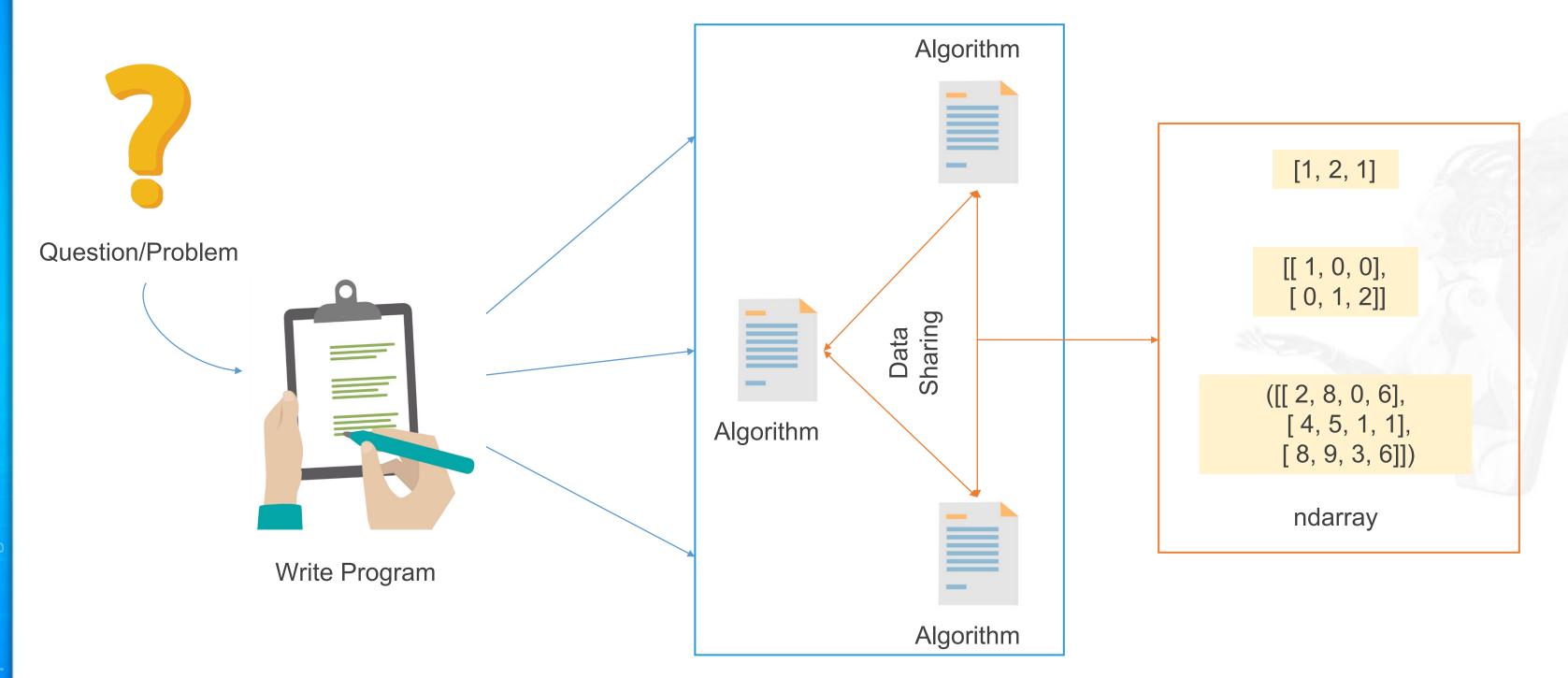
Properties of ndarray

An array in NumPy has the following properties:



Purpose of ndarray

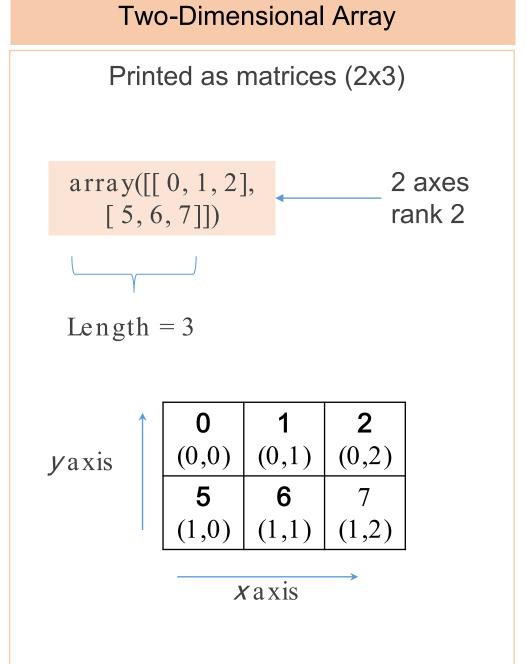
The ndarray in Python is used as the primary container to exchange data between algorithms.

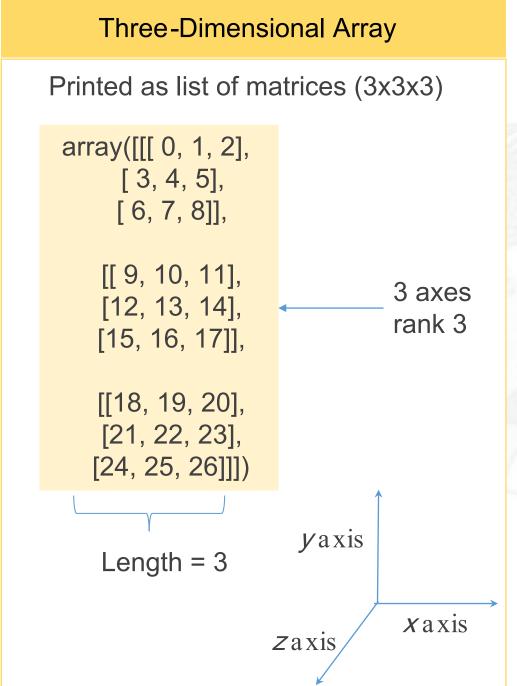


Types of Arrays

Arrays can be one -dimensional, two -dimensional, three -dimensional, or multi -dimensional.

One-Dimensional Array Printed as rows 1 axis array([5, 7,9]) rank 1 Length = 39 5 0 xaxis





Activity: Sequence it Right!

The code here is buggy. You must correct its sequence to debug it.

```
distance=[10,15,17,26]
time=[.30,.47,.55,1.20]
np_distance = np.array(distance)
np_time=np.array(time)
import numpy as np
speed=np_distance/np_time
speed
array([ 33.3333333, 31.91489362, 30.90909091, 21.66666667])
```

Activity: Sequence it Right!

```
1 distance=[10,15,17,26]
time=[.30,.47,.55,1.20]
```

2 import numpy as np

np_distance = np.array(distance)
np_time=np.array(time)

speed=np_distance/np_time

speed

speed

array([33.33333333, 31.91489362, 30.90909091, 21.66666667])

Creating and Printing an ndarray



Demonstrate how to create and print a ndarray.

Classes and Attributes of ndarray: .ndim

Numpy's array class is **ndarray**, also referred to as **numpy.ndarray**. The attributes of ndarray are:

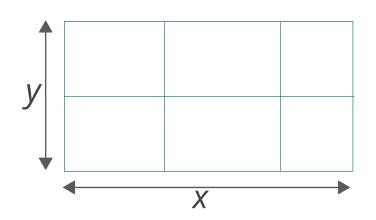
ndarray.ndim

ndarray.shape

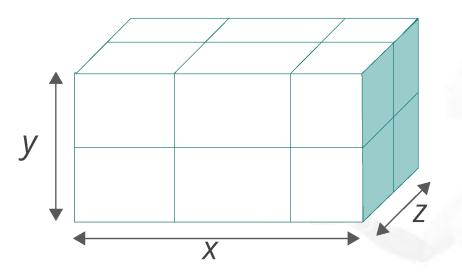
ndarray.size

ndarray.dtype

This refers to the number of axes (dimensions) of the array. It is also called the rank of the array.



Two axes or 2D array



Three axes or 3D array

Concept

Example



Classes and Attributes of ndarray: .ndim

The array np_city is one-dimensional, while the array np_city_with_state is twodimensional. ndarray.ndim In [108]: np_city = np.array(['NYC', 'LA', 'Miami', 'Houston']) In [109]: np_city.ndim ndarray.shape Out[109]: 1 In [110]: np_city_with_state = np.array([['NYC', 'LA', 'Miami', 'Houston'],['NY', 'CA', 'FL', 'TX']]) ndarray.size In [111]: np_city_with_state.ndim Out[111]: 2 ndarray.dtype Example Concept

Classes and Attributes of ndarray: .shape

Numpy's array class is ndarray, also referred to as numpy.ndarray. The attributes of ndarray are:

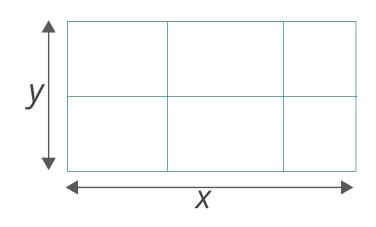
ndarray.ndim

ndarray.shape

ndarray.size

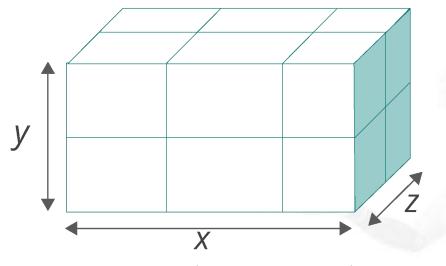
ndarray.dtype

This consists of a tuple of integers showing the size of the array in each dimension. The length of the **shape tuple** is the rank or ndim.



2 rows, 3 columns

Shape: (2, 3)



2 rows, 3 columns, 2 ranks

Shape: (2, 3, 2)

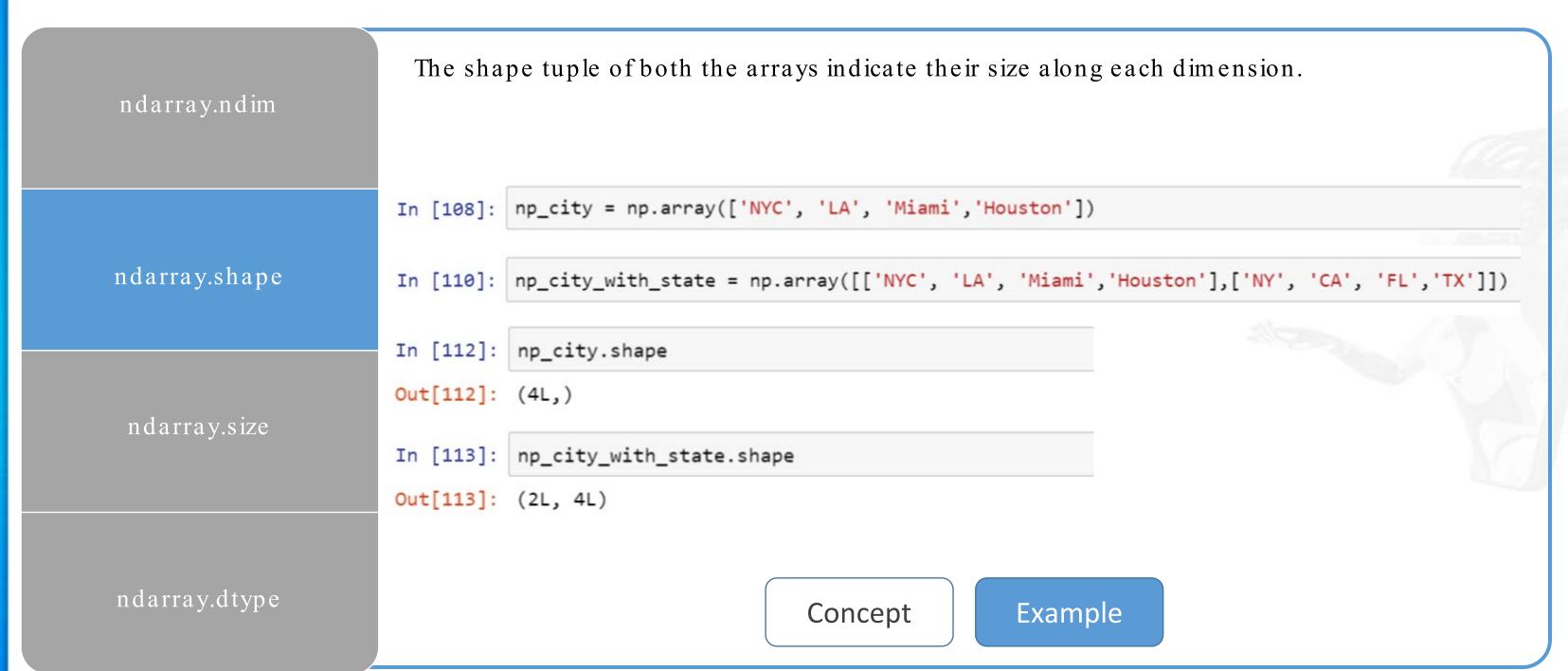
Concept

Example



Classes and Attributes of ndarray: .shape

Numpy's array class is ndarray, also referred to as numpy.ndarray. The attributes of ndarray are:



Classes and Attributes of ndarray: .size

Numpy's array class is ndarray, also referred to as numpy.ndarray. The attributes of ndarray are:

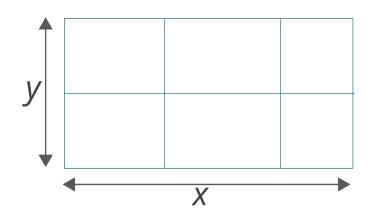
ndarray.ndim

ndarray.shape

ndarray.size

ndarray.dtype

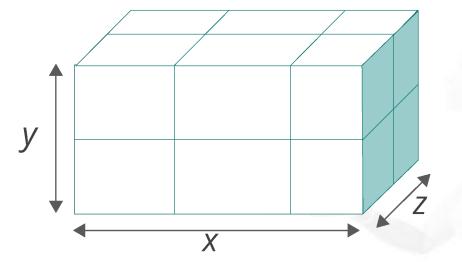
It gives the total number of elements in the array. It is equal to the product of the elements of the shape tuple.



Array contains 6 elements

Array
$$a = (2, 3)$$

Size = 6



Array contains 12 elements

Array
$$b = (2, 3, 2)$$

Size = 12

Concept

Example



Classes and Attributes of ndarray: .size

Numpy's array class is ndarray, also referred to as numpy.ndarray. The attributes of ndarray are:

Look at the examples to see how the shape tuples of the arrays are used to calculate their size. ndarray.ndim In [112]: np_city.shape Out[112]: (4L,) In [113]: np_city_with_state.shape ndarray.shape Out[113]: (2L, 4L) In [114]: np_city.size Out[114]: 4 ndarray.size In [115]: np_city_with_state.size Out[115]: 8 ndarray.dtype Example Concept



Classes and Attributes of ndarray: .dtype

Numpy's array class is ndarray, also referred to as numpy.ndarray. The attributes of ndarray are:

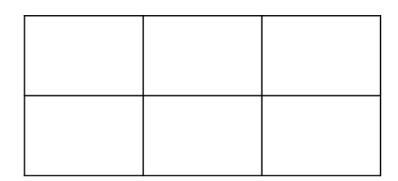
ndarray.ndim

ndarray.shape

ndarray.size

ndarray.dtype

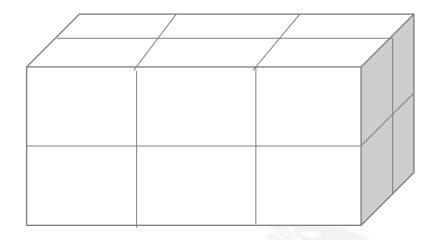
It's an object that describes the type of the elements in the array. It can be created or specified using Python.



Array contains integers

Array
$$a = [3, 7, 4]$$

[2, 1, 0]



Array contains floats

Array b =
$$[1.3, 5.2, 6.7]$$

 $[0.2, 8.1, 9.4]$

Concept

Example



Classes and Attributes of ndarray: .dtype

Numpy's array class is **ndarray**, also referred to as **numpy.ndarray**. The attributes of ndarray are:

ndarray.ndim

ndarray.shape

ndarray.size

ndarray.dtype

Both the arrays are of **string** data type (dtype) and the longest string is of length 7, which is Houston .

```
In [116]: np_city
Out[116]: array(['NYC', 'LA', 'Miami', 'Houston'],
                dtype=' S7')
In [117]: np_city_with_state
Out[117]: array([['NYC', 'LA', 'Miami', 'Houston'],
                 ['NY', 'CA', 'FL', 'TX']],
                dtype='|S7')
In [118]: np_city_with_state.dtype
Out[118]: dtype('S7')
```

Concept

Example



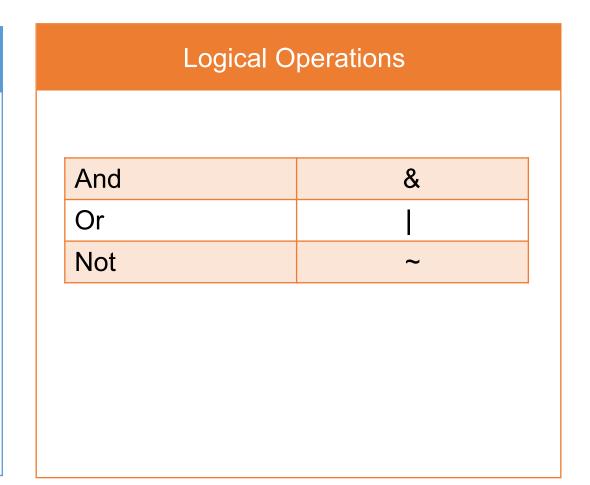
Operations



Basic Operations

Using the following operands, you can easily apply various mathematical, logical, and comparison operations on an array.

Mathematical Operations					
Addition	+				
Subtraction	-				
Multiplication	*				
Division	/				
Exponentiation	**				



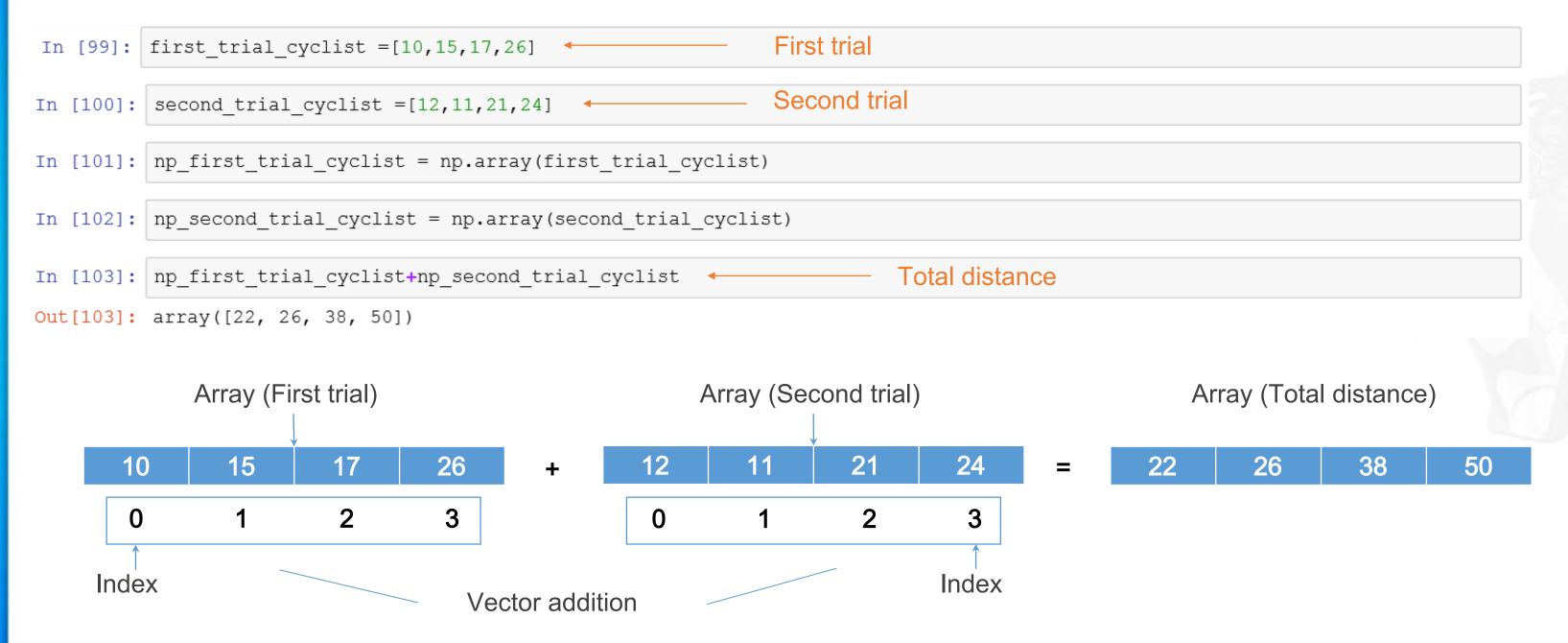
>
>=
<
<=
==
!=



Demonstrate how to apply some basic operations on an array.

Basic Operations: Example

NumPy uses the indices of the elements in each array to carry out basic operations. In this case, where we are looking at a dataset of four cyclists during two trials, vector addition of the arrays gives the required output.



Accessing Array Elements: Indexing

You can access an entire row of an array by referencing its axis index.

```
In [117]: cyclist_trials = np.array([[10,15,17,26],[12,11,21,24]]) 		 Create 2D array using cyclist trial data shown earlier

In [118]: first_trial =cyclist_trials[0] 		 First trial data

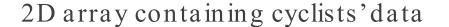
In [119]: first_trial

Out[119]: array([10, 15, 17, 26])

In [120]: second_trial = cyclist_trials[1] 		 Second trial data

In [121]: second_trial

Out[121]: array([12, 11, 21, 24])
```



10	15	17	26	←——	First trial (axis 0)
12	11	21	24	←—	Second trial (axis 1)



(1, 0)

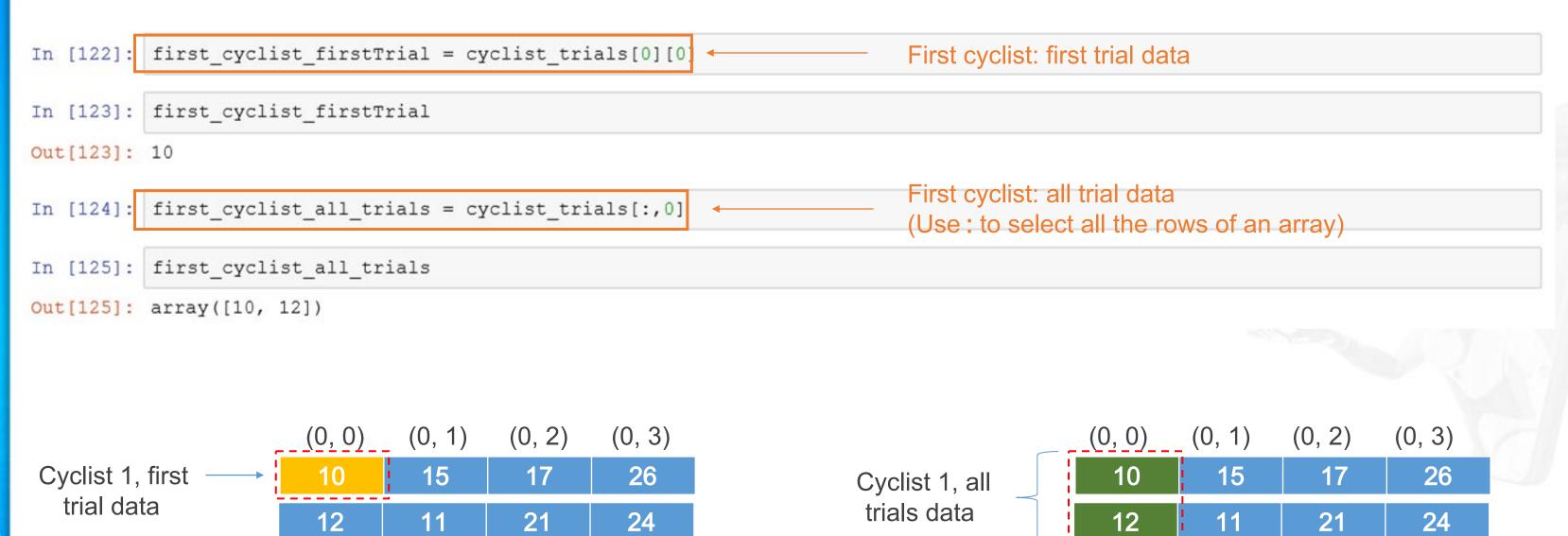
(1, 1)

(1, 2)

(1, 3)

Accessing Array Elements: Indexing

You can refer the indices of the elements in an array to access them. You can also select a particular index of more than one axis at a time.



(1, 1)

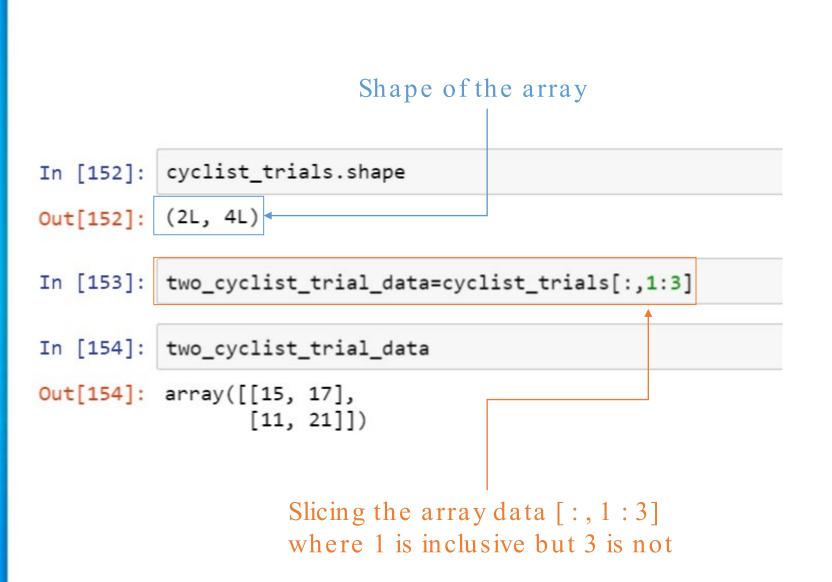
(1, 2)

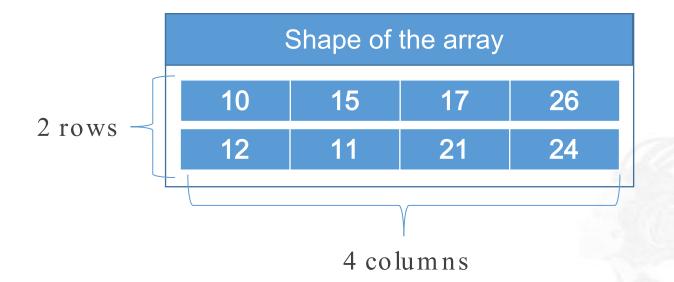
(1, 3)

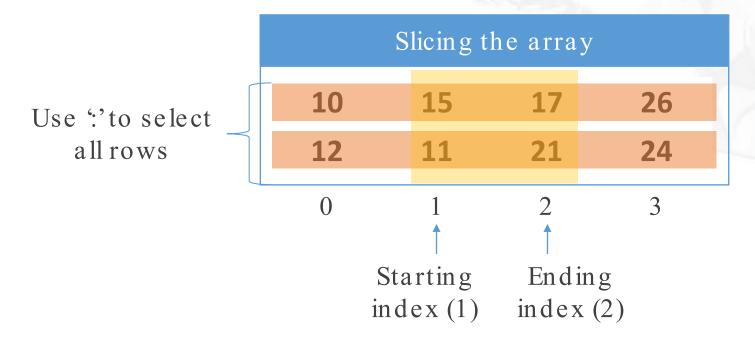
(1, 0)

Accessing Array Elements: Slicing

Use the slicing method to access a range of values within an array.







Activity: Slice It!

Select any two elements from the array to see how the statement required to slice the range changes.

Rules of the Game

- Choose the first element of the range. Then, choose the element that ends the range.
- See how the values in the statement change according to your choices.
- Refresh to try again.

5 8 10 21

example_array[1:3]

Select any two elements from the array.



Accessing Array Elements: Iteration

Use the iteration method to go through each data element present in the dataset.

```
In [117]: cyclist trials = np.array([[10,15,17,26],[12,11,21,24]])
In [153]: two_cyclist_trial_data=cyclist_trials[:,1:3]
In [154]: two_cyclist_trial_data
Out[154]: array([[15, 17],
                 [11, 21]])
                                                                                            Iterate with for loop
In [159]: for iterate_cyclist_trials_data in cyclist_trials:
                                                                                            through entire dataset
              print (iterate cyclist trials data)
          [10 15 17 26]
          [12 11 21 24]
                                                                                            Iterate with for loop through
In [160]: for iterate_two_cyclist_trial_data in two_cyclist_trial_data:
                                                                                            the two cyclist datasets
              print (iterate_two_cyclist_trial_data)
          [15 17]
          [11 21]
```

Indexing with Boolean Arrays

Boolean arrays are useful when you need to select a dataset according to set criteria.

True False



Test 2

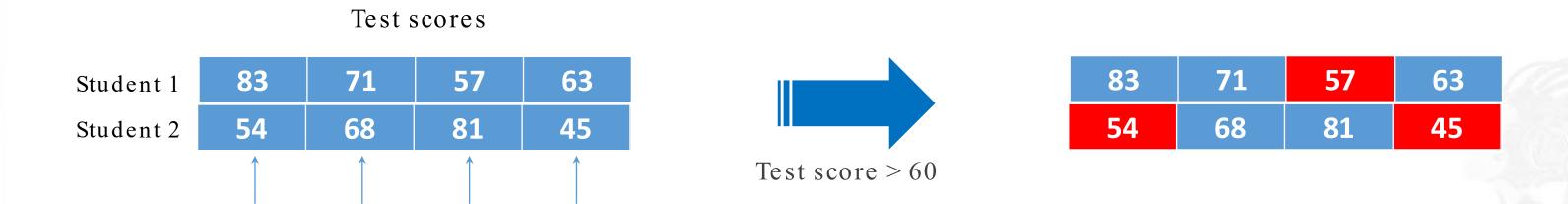
Test 3

Test 4

Test 1

Indexing with Boolean Arrays

Here, the original dataset contains test scores of two students. You can use a Boolean array to choose only the scores that are above a given value.



Simplilearn. All rights reserved

Copy and Views

When working with arrays, data is copied into new arrays only in some cases. Following are the three possible scenarios:



Copy and Views



In [296]: Boroughs_in_NYC

A view, also referred to as a shallow copy, creates a new array object.

View/Shallow Copy



Deep Copy

```
Out[296]: array(['Manhattan', 'Bronx', 'Brooklyn', 'Staten Island', 'Queens'],
                                                                                     Original dataset
               dtype='|S13')
In [298]: View of Borough in NYC = Boroughs in NYC.view()
In [299]: len(View_of_Borough_in_NYC)
Out[299]: 5
In [300]: View_of_Borough_in_NYC[4] ='Central Park' Change value in view object
In [301]: View of Borough in NYC
Out[301]: array(['Manhattan', 'Bronx', 'Brooklyn', 'Staten Island', 'Central Park'],
               dtype='|S13')
In [302]: Boroughs in NYC
Out[302]: array(['Manhattan', 'Bronx', 'Brooklyn', 'Staten Island', 'Central Park'],
                                                                                         Original dataset
               dtype='|S13')
                                                                                         changed
```

Copy and Views





View/Shallow Copy



Copy is also called **deep copy** because it entirely copies the original dataset. Any change in the copy will not affect the original dataset.

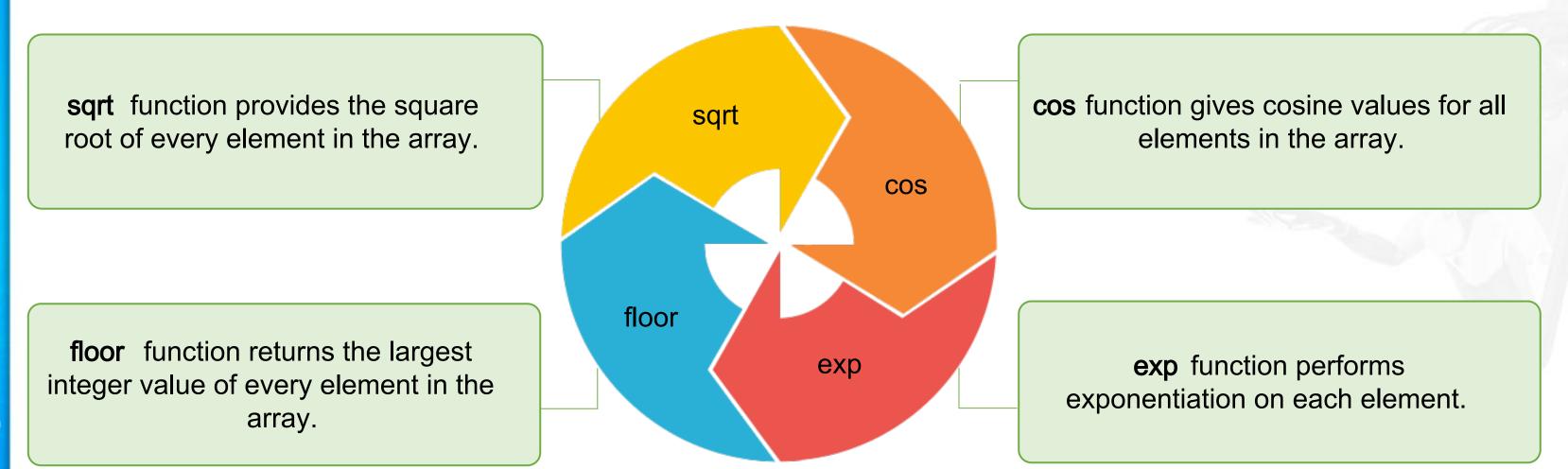
```
In [304]: Copy of NYC Borough = NYC Borough.copy()
                                                                    Shows copy and original object are different
In [305]: Copy of NYC Borough is NYC Borough
Out[305]: False
                                                                    Shows copy object data is not
In [306]: Copy of NYC Borough.base is NYC Borough
                                                                    owned by the original dataset
Out[306]: False
                                                                 In [307]: Copy_of NYC Borough[4]='Central Park'
In [308]: NYC Borough
Out[308]: array(['Manhattan', 'Bronx', 'Brooklyn', 'Staten Island', 'Queens'],

    Copy object changed

               dtype='|S13')
In [309]: Copy of NYC Borough
Out[309]: array(['Manhattan', 'Bronx', 'Brooklyn', 'Staten Island', 'Central Park'],
                                                                                         Original dataset
               dtype='|S13')
                                                                                           retained
```

Universal Functions (ufunc)

NumPy provides useful mathematical functions called Universal Functions. These functions operate element -wise on an array, producing another array as output. Some of these functions are:



Ufunc: Examples

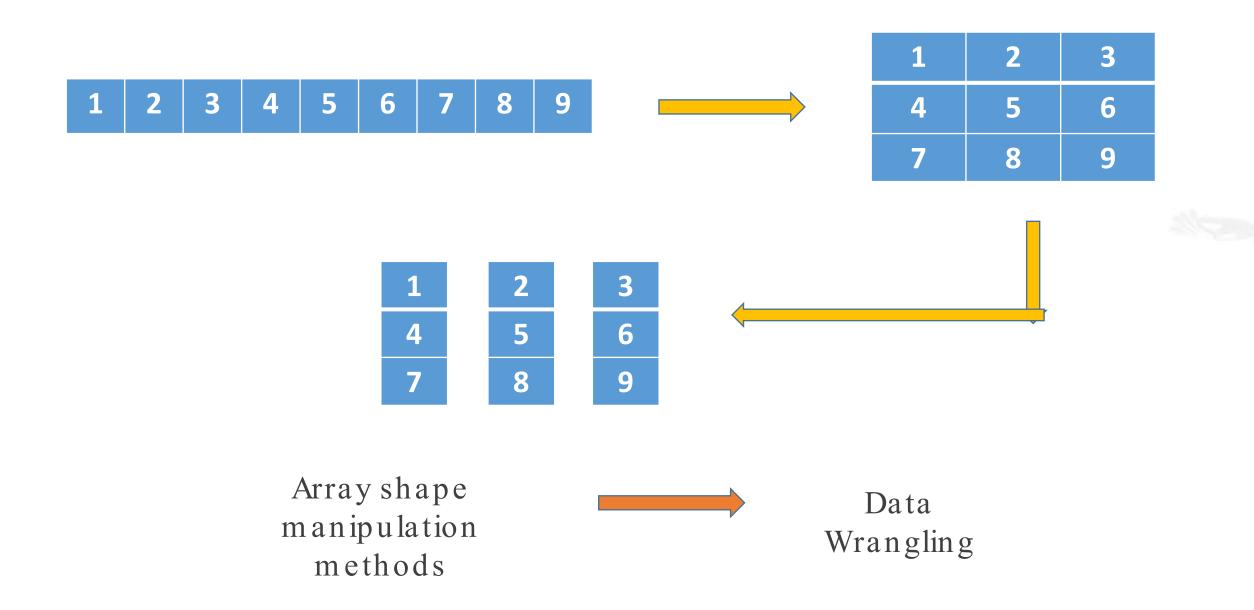
```
    Numbers for which square root will be calculated

In [186]: np_sqrt = np.sqrt([2,4,9,16])
In [187]: np_sqrt
                                                                             Square root values
Out[187]: array([ 1.41421356, 2.
                                      , 3.
                                                                             Import pi*
In [188]: from numpy import pi ←
          np.cos(0)
Out[188]: 1.0
In [189]: np.sin(pi/2) ←
                                                                             Trigonometric functions
Out[189]: 1.0
In [190]: np.cos(pi)
Out[190]: -1.0
In [191]: np.floor([1.5,1.6,2.7,3.3,1.1,-0.3,-1.4])
                                                                            Return the floor of the input element wise
Out[191]: array([ 1., 1., 2., 3., 1., -1., -2.])
                                                                             Exponential functions for complex
In [192]: np.exp([0,1,5])
                                                                             mathematical calculations
Out[192]: array([ 1. , 2.71828183, 148.4131591 ])
```

Shape Manipulation

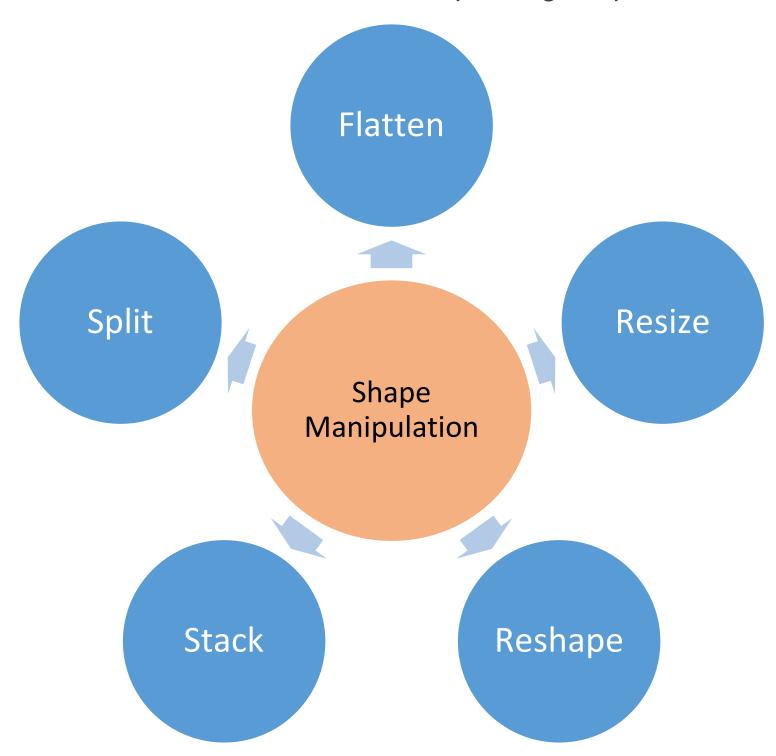
You can use certain functions to manipulate the shape of an array.

The shape of an array can be changed according to the requirement using the NumPy library functions.



Shape Manipulation

Some common methods for manipulating shapes are:



Shape Manipulation: Example

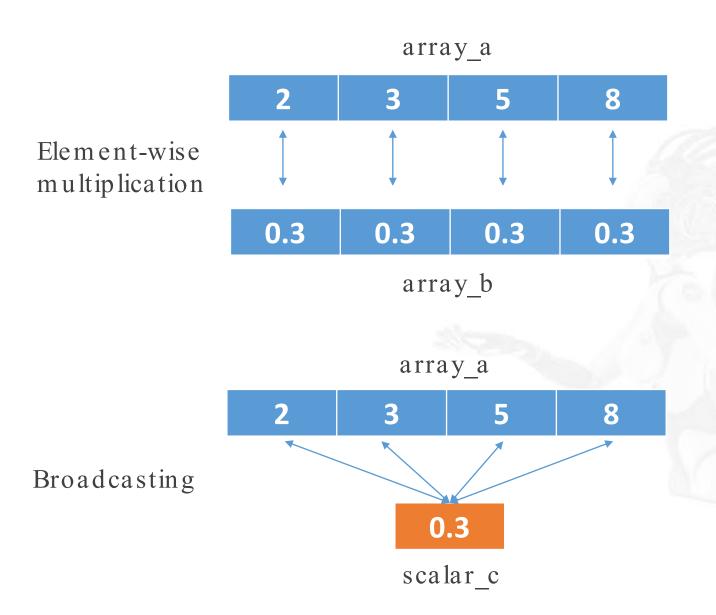
You can use certain functions to manipulate the shape of an array to do the following:

```
In [383]: new cyclist trials = np.array([[10,15,17,26,13,19],[12,11,21,24,14,23]])
                                                                    Flattens the dataset
In [384]: new cyclist trials.ravel()
Out[384]: array([10, 15, 17, 26, 13, 19, 12, 11, 21, 24, 14, 23])
                                                                    Changes or reshapes the dataset to 3 rows and 4 columns
In [385]: new cyclist trials.reshape(3,4)
Out[385]: array([[10, 15, 17, 26],
                [13, 19, 12, 11],
                 [21, 24, 14, 23]])
                                                                    Resizes again to 2 rows and 6 columns
In [386]: new cyclist trials.resize(2,6)
In [387]: new cyclist trials
Out[387]: array([[10, 15, 17, 26, 13, 19],
                 [12, 11, 21, 24, 14, 23]])
                                                                    Splits the array into two
In [388]: np.hsplit(new cyclist trials,2)
Out[388]: [array([[10, 15, 17],
                 [12, 11, 21]]), array([[26, 13, 19],
                  [24, 14, 23]])]
In [389]: new cyclist 1 = np.array([10,15,17,26,13,19])
In [390]: new cyclist 2 = np.array([12,11,21,24,14,23])
                                                                    Stacks the arrays together
In [391]: np.hstack((new cyclist 1,new cyclist 2))
Out[391]: array([10, 15, 17, 26, 13, 19, 12, 11, 21, 24, 14, 23])
```

Broadcasting

NumPy uses broadcasting to carry out arithmetic operations between arrays of different shapes. In this method, NumPy automatically broadcasts the smaller array over the larger array.

```
In [9]: import numpy as np
In [10]: #Create two arrays of the same shape
         array a = np.array([2, 3, 5, 8])
         array b = np.array([.3, .3, .3, .3])
In [11]: #Multiply arrays
         array a * array b
Out[11]: array([ 0.6, 0.9, 1.5, 2.4])
In [12]: #Create a variable with a scalar value
         scalar c = .3
In [13]: #Multiply 1D array with a scalar value
         array_a * scalar c
Out[13]: array([ 0.6, 0.9, 1.5, 2.4])
```



If the shape doesn't match with array_a, numpy doesn't have to create copies of scalar value.

Instead, broadcast scalar value over the entire array to find the product.



Broadcasting: Constraints

Though broadcasting can help carry out mathematical operations between different -shaped arrays, they are subject to certain constraints as listed below:

- When NumPy operates on two arrays, it compares their shapes element -wise. It finds these shapes compatible only if:
 - Their dimensions are the same or
 - One of them has a dimension of size 1.
- If these conditions are not met, a **ValueError** is thrown, indicating that the arrays have incompatible shapes.

Broadcasting: Example

Let's look at an example to see how broadcasting works to calculate the number of working hours of a worker per day in a certain week.

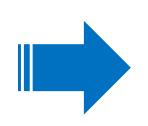
```
Week one earnings
In [246]: np_week_one =np.array([105, 135, 195, 120, 165])
           np_week_two =np.array([123, 156, 230, 200, 147])
                                                                               Week two earnings
In [247]: total_earning = np_week_one+np_week_two
                                                       Element-wise operation
In [248]:
           total earning
Out[248]: array([228, 291, 425, 320, 312])
                                                            Total earning for 2 weeks
In [249]: np_week_one_hrs = np_week_one / 15
                                                               Calculate week one hours
                                                               Hourly wage
In [250]:
           np_week_one_hrs
                                                      Number of working hours
Out[250]: array([ 7, 9, 13, 8, 11]) <
                                                      per day in week one
```

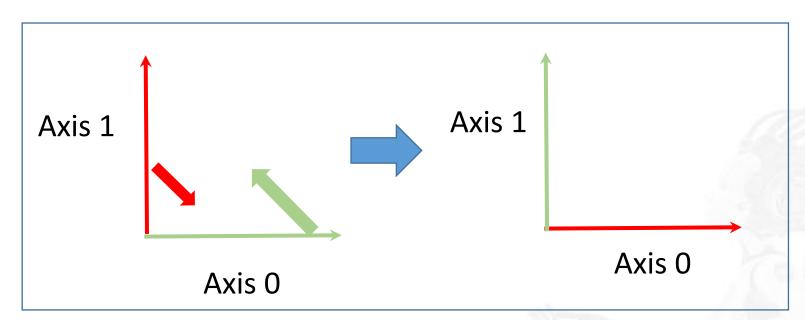
Linear Algebra: Transpose

NumPy can carry out linear algebraic functions as well. The **transpose()** function can help you interchange rows as columns, and vice -versa.

transpose()					
	83	71	57	63	
	54	68	81	45	
	54	00	01	45	

[63, 45]])





Linear Algebra: Inverse and Trace Functions

Using NumPy, you can also find the inverse of an array and add its diagonal data elements.

* Can be applied **only** on a square matrix



Evaluate the dataset containing the GDPs of different countries to:

- Find and print the name of the country with the highest GDP
- Find and print the name of the country with the lowest GDP
- Print out text and input values iteratively
- Print out the entire list of the countries with their GDPs
- Print the highest GDP value, lowest GDP value, mean GDP value, standardized GDP value, and the sum of all the GDPs



Evaluate the dataset of the Summer Olympics, London 2012 to:

- Find and print the name of the country that won maximum gold medals
- Find and print the countries who won more than 20 gold medals
- Print the medal tally
- Print each country name with the corresponding number of gold medals
- Print each country name with the total number of medals won

DATA AND ARTIFICIAL INTELLIGENCE



Knowledge Check



Which of the following arrays is valid?

- a. [1, 0.3, 8, 6.4]
- b. ["Lucy", 16, "Susan", 23, "Carrie", 37]
- c. [True, False, "False", True]
- d. [3.14j, 7.3j, 5.1j, 2j]



Which of the following arrays is valid?

- a. [1, 0.3, 8, 6.4]
- b. ["Lucy", 16, "Susan", 23, "Carrie", 37]
- c. [True, False, "False", True]
- d. [3.14j, 7.3j, 5.1j, 2j]



The correct answer is d

A NumPy ndarray can hold only a single data type, which makes it homogenous. NumPy supports integers, floats, Booleans, and even complex numbers. Of all the options provided, only the array containing complex numbers is homogenous. All the other options contain more than one data type.



2

Which function is most useful to convert a multidimensional array into a one dimensional array?

- a. ravel()
- b. reshape()
- c. resize() and reshape()
- d. All of the above





2

Which function is most useful to convert a multidimensional array into a one dimensional array?

- a. ravel()
- b. reshape()
- c. resize() and reshape()
- d. All of the above



The correct answer is a

The function ravel() is used to convert a multidimensional array into a one -dimensional array. Though reshape() also functions in a similar way, it creates a new array instead of transforming the input array.



The np.trace() method gives the sum of _____.

3

- a. the entire array
- b. the diagonal elements from left to right
- c. the diagonal elements from right to left
- d. consecutive rows of an array





3

The np.trace() method gives the sum of _____.

- a. the entire array
- b. the diagonal elements from left to right
- c. the diagonal elements from right to left
- d. consecutive rows of an array



The correct answer is **b**

The trace() function is used to find the sum of the diagonal elements in an array. It is carried out in an incremental order indices. Therefore, it can only add diagonal values from left to right and not vice versa.

of the

4

The function np.transpose() when applied on a one dimensional array gives _____.

- a. a reverse array
- b. an unchanged original array
- c. an inverse array
- d. all elements with zeroes





4

The function np.transpose() when applied on a one dimensional array gives _____.

- a. a reverse array
- b. an unchanged original array
- c. an inverse array
- d. all elements with zeroes



The correct answer is **b**

Transposing a one -dimensional array does not change it in any way. It returns an unchanged view of the original array.



Key Takeaways

You are now able to:

- Explain NumPy and its importance
- Discuss the basics of NumPy, including its fundamental objects
- Demonstrate how to create and print a NumPy array
- Analyze and perform basic operations in NumPy
- Utilize shape manipulation and copying methods
- Demonstrate how to execute linear algebraic functions
- Build basic programs using NumPy





Thank You

