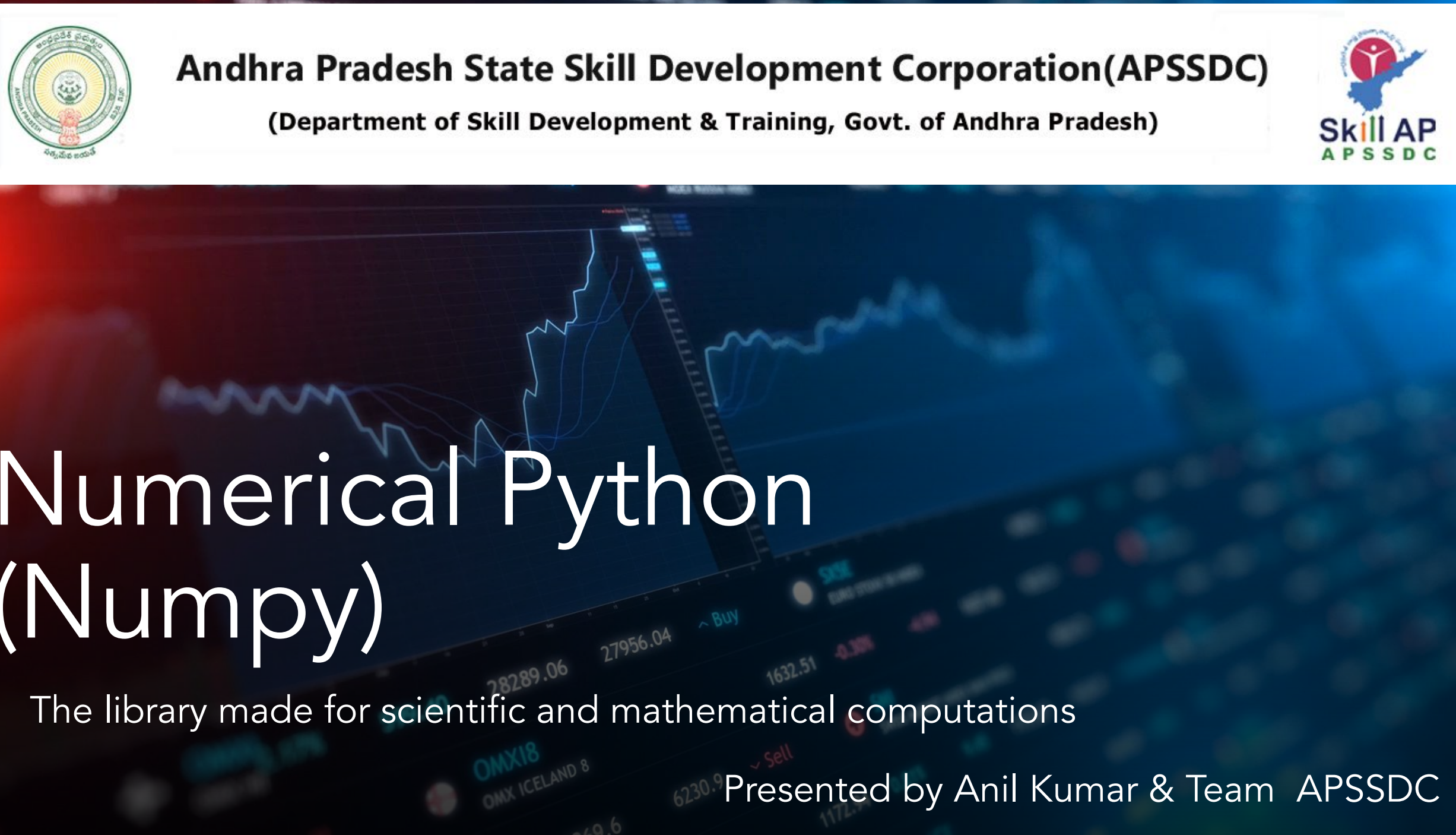


Numerical Python (Numpy)

The library made for scientific and mathematical computations

Presented by Anil Kumar & Team APSSDC

If we look into any dataset

- What do we see in the dataset ?
- What is the type of data present in the dataset?
- How can we perform operations on such a dataset?
- Need of libraries like Pandas, Numpy, matplotlib, scipy

What is Numpy?

- Numerical Python, popularly known as Numpy has been designed to carry out mathematical computations at a faster and easier rate.
- Further this library enriches the programming language Python by providing powerful data structures like multi dimensional arrays beyond matrices and linear arrays.
- Besides that, Numpy provides a large library of high level mathematical functions to operate on these structures.

How to install Numpy

- In command line

  pip install numpy

- Anaconda distribution

  conda install numpy

Python Objects vs Numpy

python objects

1. high-level number objects: integers, floating point

2. containers: lists (costless insertion and append), dictionaries (fast lookup)

Numpy provides

1. extension package to Python for multi-dimensional arrays

2. closer to hardware (efficiency)

3. designed for scientific computation (convenience)

4. Also known as array oriented computing

Why Numpy when we have “Lists” ?

Python has inbuilt data structure “List” which is also technically an array which allows different data types.

The answer to this question comes in following three aspects

1. Size – Numpy data structures take less space
2. Performance – They are inherently faster than lists.
3. Functionality – Scipy and Numpy have optimized functions.
4. Vectorization of the operations

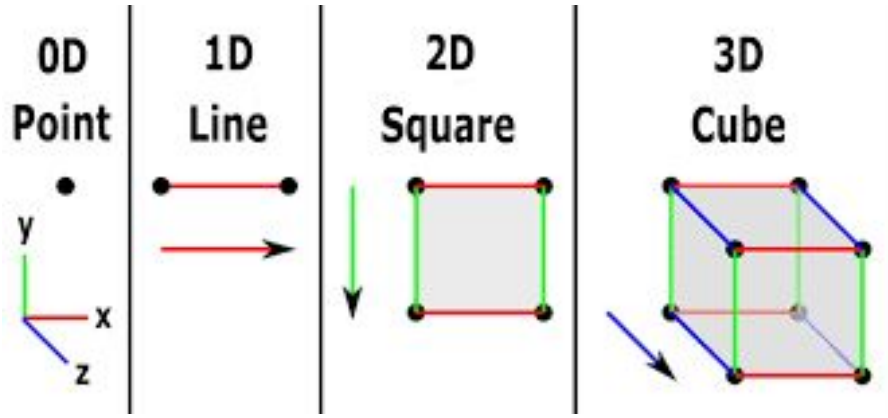
Now let’s create some numpy arrays and play around with them!!

Nd-array object

- Ndarray is multidimensional object which can contain only single data type objects.
- It can be a string type or numeric or integer data type.
- If we mixture of strings and numbers are used, all are converted to strings.

Attributes of ND array object

- Dimension – It tells us the number of dimensions of the nd array object. Number of dimensions can range from 1 to 100s and 1000s
- Shape – It gives the shape of the nd array object. That is the length of each dimension.



Attributes of ND array object

- Size - Total number of elements in numpy array
- Dtype – It tells about the type of data being stored in the object.
- Strides – How many steps to be taken to move to next row!!

Some miscellaneous numpy arrays

- Np.zeros()
- Np.arange()
- Np.linspace()
- Np.full()

Resize, reshape , flatten and ravel

- Resize adds zeros if you want to create size larger than current one.

Note – resize() does not work on view of nd array but original one

- Reshape – reshapes array to any size and dimension
- Flatter and ravel – They help in flattening multidimensional array into a single array.

Array indexing and slicing operations

They are straight forward ways to manipulate data into numpy arrays. Let's see how to do it for simple toy numpy arrays

Lets manipulate numpy arrays!

Trigonometric operations

- np.sin(), np.cos(), np.exp(), np.sqrt()

Comparison of numpy objects

- np.equal(ar1, ar2), np.not_equal(ar1, ar2)

Logical operations

- np.logical_or(), np.logical_and(), np.logical_not()

Broadcasting of numpy arrays

- To put it in a more practical context, you often have an array that's somewhat larger and another one that's slightly smaller. Ideally, you want to use the smaller array multiple times to perform an operation (such as a sum, multiplication, etc.) on the larger array.
- 1.First off, to make sure that the broadcasting is successful, the *dimensions of your arrays need to be compatible*.
- 2.Two dimensions are also compatible when *one of them is 1*

Matrix operations

- Numpy provides a range of functions to carry out various matrix operations

- Addition
- Matrix dot product
- Matrix element wise multiplication
- Matrix multiplication

Let's create arrays with random numbers

The source of randomness which we inject into machine learning projects is called Pseudo randomness.

1. np.random.rand()
2. np.random.randint()
3. np.random.shuffle()

Concatenate, append and stack numpy arrays

- Often, we might want to join different numpy arrays in different ways like column wise, row wise.
- Numpy offers a range of functions to do the same namely

np.append()

np.concatenate()

np.vstack()

np.hstack()

Let’s visualize numpy arrays

- The numeric data in numpy arrays can be better interpreted with visualizations from matplotlib.
- np.histogram()
- np.meshgrid()

Let’s save the nd array object to a file now!

- After all the processing and manipulation of the data, the most important step that comes is to save data into a file.
- Numpy provides savetxt(<file name>,<nd_array_object>, <delimiter>) function for the same.

Numpy quick start links

- Original documentation

https://numpy.org/devdocs/user/quickstart.html

- Cheat sheet

https://www.datacamp.com/community/blog/python-numpy-cheat-sheet

Thanks for taking part in today's session

Skill AP
APSSDC

NumPy