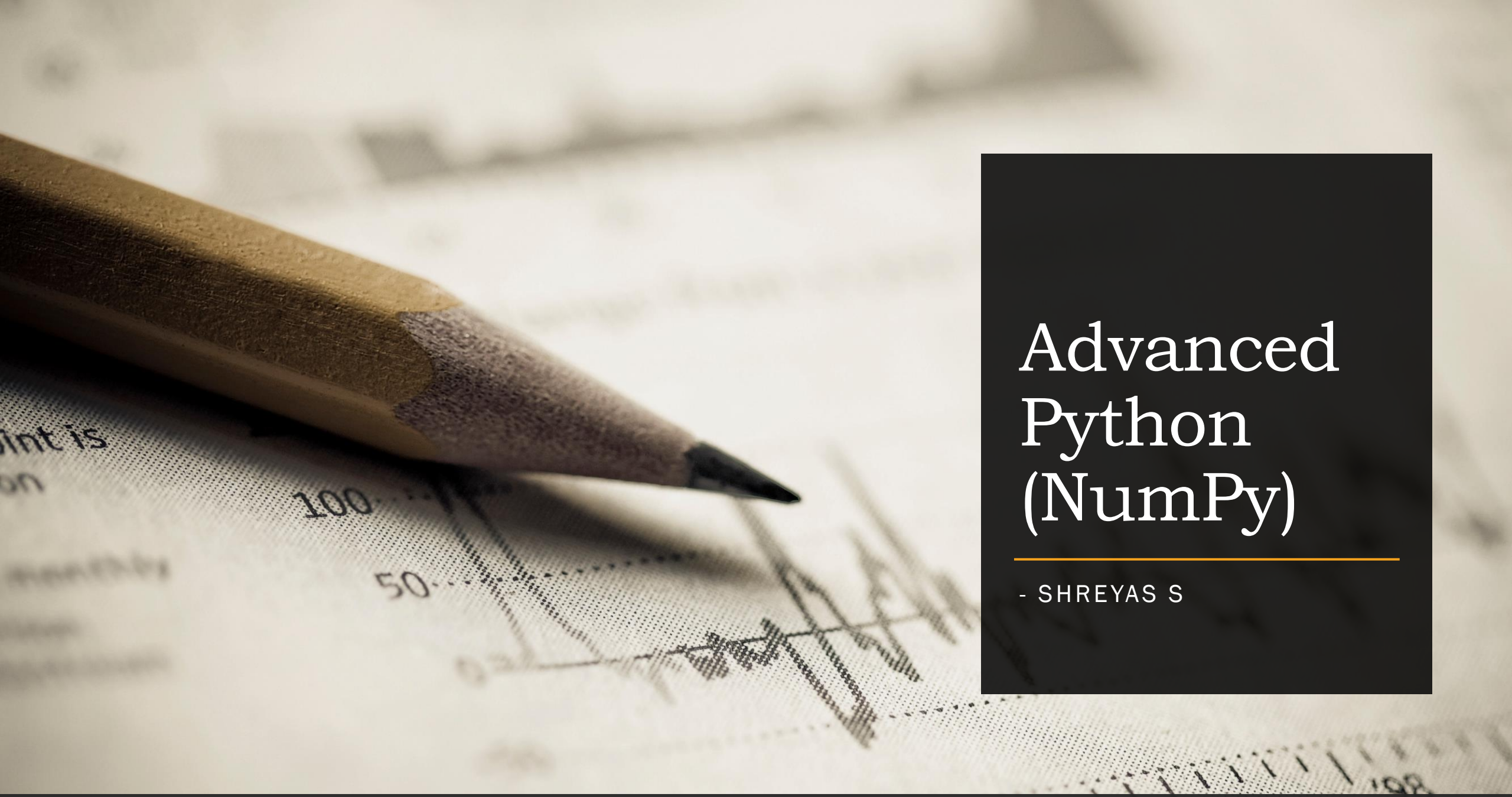

Advanced Python (NumPy)

- SHREYAS S

Agenda

- What is NumPy and why NumPy?
- NumPy Arrays
- NumPy Array Indexing and Array Slicing
- Data types in NumPy
- Shape and Reshape in numpy arrays
- Iterating over numpy arrays
- Array split and array join
- Array search and array sort
- Filtering numpy arrays
- Good to know functions in numpy

What is NumPy and why numpy?

- NumPy is a python library that is used for scientific computing.
- NumPy stands for Numerical Python.
- Numpy helps you to create multi-dimensional arrays.
- Numpy arrays have fixed size when they are created whereas python list size can grow dynamically. In numpy when you change the size or add elements then original array is deleted after creating new array.
- Unlike python list where you can have multiple values of multiple data types in single list, in numpy all elements are required to be of the same data type.
- All operations on numpy are performed faster than the python built-in sequences.
- In numpy as task in broken down to multiple sub-task and then parallelly executes all those tasks.
- Numpy internally integrates C++, C and Fortran codes in python which in turn reduces the execution time as these languages have less execution time when compared to python.

Numpy Arrays

- So now since we know that numpy are used to work with arrays next we would look deeper into how arrays are created and type of arrays.
- The array in numpy is stored in an object called ndarray and you can create ndarray object using array() function in numpy.
- Now there is a concept of dimensions in numpy arrays. These are nothing but the level of array depth.
- We can have n dimension in array but highly used are 0-D, 1-D, 2-D and 3-D arrays.
- Numpy also have ndim attribute which returns the number of dimension an array has.

Numpy Arrays Indexing and slicing

- Indexing in numpy is similar to list.
- Here the only difference is how we access elements in n-dimensional array.
- Similar to list you can also perform negative indexing on numpy arrays.
- You can even use the slicing functionality on numpy array similar to the way we do that in list or tuple.
- Always remember in slicing the result includes elements from start index upto the end index but not the end index. (similar to lists)

Data types in numpy array

- As we know python has some basic data types like int, float, str, Boolean,complex.
- But numpy support more number of data types other than the above.
- You can check all the list of supported data types in numpy in the below link: https://numpy.org/doc/stable/user/basics.types.html
- You can check the data type of numpy array using dtype property on created numpy array object.
- You can also mention the data type while creating the numpy array.
- In some cases while doing so it may give you error like you get in normal python coding like if you have an element as string eg: 'hello' and then while creating an array with this element if you mention data type as integer then this element can't be converted to integer so will get you an error.
- You can also convert the data type of an existing array using astype property on array object.

Shape and reshape in numpy

- Shape in numpy array represents the number of elements in each dimension of the array.
- Numpy arrays have an attribute shape which return the shape of each dimension in an array.
- The shape is returned as a tuple.
- Every integer in tuple represents the number of elements the corresponding dimension in the array has.
- Then we have a concept reshape which can be used to change the shape of array.
- By using reshape we can add/remove dimension or number of elements in it.
- While reshaping you can mention unknown dimension where you do not mention the exact number for one of the dimensions.
- You can also flatten the multi dimensional array to 1D array.

Iterating over numpy arrays

- Similar to list you can use for loop for iterating over numpy arrays.
- But to access each scalar values in numpy arrays of n dimension you need to use nested for loop.
- You can also use nditer which is a function by numpy to iterate easily over the array elements.
- Using nditer you can also change data type while iterating over it as well as you can use it to iterate with different step size.
- nditer has an argument called op_flags through which you can modify the existing array.
- You can enumerate your array while iterating using the ndenumerate function.

Array join and array split

- You can use numpy's built in functions like concatenate where we can pass n number of arrays that we want to join.
- Now we do have an option where we can join the arrays based on axis. If axis is given none then the arrays are concatenated by flattening them.
- You can also use stack to concatenate n number of arrays along a new axis.
- You can use hstack() for concatenating along rows and vstack() to concatenate along columns and dstack() to stack along height (same as depth).
- Splitting is doing the reverse of what we do in joining. It breaks an array into multiple arrays.
- We use array_split() function to split arrays where we pass the array that we want to split and also the number of splits.

Array join and array split

- But say you have a 1D array of 7 elements and you try to split the array into 4 parts then the end of the array would be adjusted accordingly.
- This same function can be used for 2D arrays where we split them into number of groups.
- You can also mention the axis along which you want the array to get splitted.
- Similar to hstack, vstack and dstack for splitting we have hsplit, vsplit and dsplit which have the opposite functionalities than what they have in concatenation. (here they split the arrays)

Array search and array sort

- In numpy you can use where function to search a particular value. This function returns the index of the value that you are trying to search in the array.
- The other method in search is the searchsorted which performs binary search on the array and return you the index of where the number that you specified should be inserted in the array to keep it in sorted format.
- Arrays can also be sorted using the sort function which sort the array in an ascending manner.
- In case you have strings in you arrays then they are sorted alphabetically.
- This function can be used to sort n dimensional arrays as well. Here the scalar values (elements) in the arrays are sorted.

Filtering Numpy arrays

- When you want to get some elements from the existing numpy then you can use the filtering functionality.
- For filtering the existing array you need to use an array with Boolean values and the their indexs are used to filter the array.
- There are multiple ways to do so. Lets look at them one by one.

Good to know function in numpy

- Creating numpy array of zeros
- Creating numpy arrays of ones
- Creating numpy arrays of random numbers
- Flattening numpy array of higher dimension
- Transpose in numpy arrays
- Max in numpy arrays (row wise and column wise)
- Sum of all elements in numpy arrays

Agenda for Next Session

- ❑ Pandas library
- ❑ Series and Dataframe
- ❑ Operation on series and dataframe

QUESTIONS?

THANK YOU