**Python Programming for Data Science & AI**

**Week 1**  
  
It covers simple coding with numbers, types, indexing, slicing, stride, expressions, string operations and such.

**Week 2**

**Lists & Tuples**

These are called compound data types and are one of the key types of data structures in Python.

**Tuples** are an ordered sequence.

Tuples are expressed as comma separated elements within **parentheses**.

These are values inside the parentheses. And Tuples are immutable.

**Lists** are also a popular data structure in Python. Lists are also an ordered sequence.

A list is represented with **square brackets**.

In many aspects, lists are like tuples. One key difference is lists are mutable.

Lists can contain strings, floats, integers. We can nest other lists. We also nest tuples and other data structures.

The same indexing conventions apply for nesting like tuples, each element of a list can be accessed via an index.

**.split() function is used to convert the string into list.**

**Week 3**

**Range() Function:**

The range function outputs and ordered sequence as a list I. If the input is a positive integer, the output is a sequence. The sequence contains the same number of elements as the input but starts at zero.

**For loop:**

**While loop:** While loops are similar to for loops but instead of executing a statement a set number of times a while loop will only run if a condition is met.

*The main difference between a while loop and a for loop in Python is how they control the flow of execution and handle iterations.*

Example:

*Animals = ["lion", "giraffe", "gorilla", "parrots", "crocodile","deer", "swan"]*

*New = []*

*i=0*

*while i<len(Animals):*

*j=Animals[i]*

*if(len(j)==7):*

*New.append(j)*

*i=i+1*

*print(New)*

**Functions:**

A function is a block of code which only runs when it is called.

You can pass data, known as parameters, into a function.

A function can return data as a result. A function can have multiple params.

**Local variable:** A variable that is declared inside a function is called a local variable. The parameter only exists within the function (i.e. the point where the function starts and stops).

**Global Variable:** A variable that is declared outside a function definition is a global variable, and its value is accessible and modifiable throughout the program. We will discuss more about global variables at the end of the lab.

A function combines many instructions into a single line of code. Once a function is defined, it can be used repeatedly.

The scope of a variable is the part of that program where that variable is accessible. Variables that are declared outside of all function definitions.

Functions can be incredibly powerful and versatile. They can accept (and return) data types, objects and even other functions as arguments.

**Exception Handling:**

**Exception:** An exception is an error that occurs during the execution of code. This error causes the code to raise an exception and if not prepared to handle it will halt the execution of the code.

A **try except** will allow you to execute code that might raise an exception and in the case of any exception or a specific one we can handle or catch the exception and execute specific code. This will allow us to continue the execution of our program even if there is an exception.

Python tries to execute the code in the try block. In this case if there is any exception raised by the code in the try block, it will be caught and the code block in the except block will be executed. After that, the code that comes after the try except will be executed.

A specific try except allows you to catch certain exceptions and also execute certain code depending on the exception. This is useful if you do not want to deal with some exceptions and the execution should halt. It can also help you find errors in your code that you might not be aware of. Furthermore, it can help you differentiate responses to different exceptions. In this case, the code after the try except might not run depending on the error.

**Object:** In Python, each is an object. Every object has the following:

* a type,
* internal representation,
* a set of functions called methods to interact with the data.

An object is an instance of a particular type.

**Methods:** A class or type's methods are functions that every instance of that class or type provides. It's how you interact with the object. We have been using methods all this time, for example, on lists. Sorting is an example of a method that interacts with the data in the object.

**What is Text analysis?**

Text analysis, also known as text mining or text analytics, refers to the process of extracting meaningful information and insights from textual data.

**Pandas Library:**

Pandas is a popular library for data analysis built on top of the Python programming language. Pandas generally provide two data structures for manipulating data, They are:

* DataFrame
* Series

A **DataFrame** is a two-dimensional data structure, i.e., data is aligned in a tabular fashion in rows and columns.

* A Pandas DataFrame will be created by loading the datasets from existing storage.
* Storage can be SQL Database, CSV file, Excel file, etc.
* It can also be created from lists, dictionaries, and from a list of dictionaries.

**Series** represents a one-dimensional array of indexed data. It has two main components :

1. An array of actual data.
2. An associated array of indexes or data labels.

The index is used to access individual data values. You can also get a column of a dataframe as a **Series**. You can think of a Pandas series as a 1-D dataframe.

loc() is a label-based data selecting method which means that we have to pass the name of the row or column that we want to select. This method includes the last element of the range passed in it.

Simple syntax for your understanding:

* loc[row\_label, column\_label]

iloc() is an indexed-based selecting method which means that we have to pass an integer index in the method to select a specific row/column. This method does not include the last element of the range passed in it.

Simple syntax for your understanding:

* iloc[row\_index, column\_index]

**NumPy Library:**

**1D Array (One Dimensional - Array)**

A **numpy array** or ND array is similar to a list. It's usually fixed in size and each element is of the same data type such as int, float, string etc.

**NumPy** is a Python library used for working with arrays, linear algebra, fourier transform, and matrices.

A numpy array is similar to a list.

NumPy stands for Numerical Python and it is an open source project.

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.

**The attribute shape is a tuple of integers indicating the size of the array in each dimension:**

**Linspace:**

A useful function for plotting mathematical functions is linspace. Linspace returns evenly spaced numbers over a specified interval.

numpy.linspace(start, stop, num = int value)

start : start of interval range

stop : end of interval range

num : Number of samples to generate.

**2D Array (Two Dimensional - Array)**

It is basically metrices (array with 2d is matrices)

**API (Application Programming Interface)**

You use the API to communicate with other software via inputs and outputs. Just like a function, you don’t have to know how the API works, just its inputs and outputs.

Maa:

Maa, aapka karj to utaar nahi skta hu,

Bin aapke ek kissa ban gaya hu!!

Maine Jannat to nahi dekhi, Apko dekha hai,

Ye soch ke, aankhon me aansu liye baitha rahta hun!!

An API lets two pieces of software talk to each other. Just like a function, you don't have to know how the API works, only its inputs and outputs. An essential type of API is a REST API that allows you to access resources via the internet. In this lab, we will review the Pandas Library in the context of an API, we will also review a basic REST API.

**REST APIs, Webscraping, and Working with Files**

**URL = Uniform Resource Locator**

**HTTP = Hypertext Transfer Protocol**

Url has three parts:

* **Scheme:** http://
* **Internet Address:** [www.ibm.com](http://www.ibm.com) or [www.google.com](http://www.google.com)
* **Route:** Location on the webserver, such as /images/IDSNlogo.png

**Web scraping**

**Beautiful Soup Objects¶**

Beautiful Soup is a Python library for pulling data out of HTML and XML files, we will focus on HTML files. This is accomplished by representing the HTML as a set of objects with methods used to parse the HTML. We can navigate the HTML as a tree, and/or filter out what we are looking for.

**Working with Different File formats:**

Such as xlsx, csv, json, xml etc.