NAME - ANANYA RAI

Presentation on :

FUNCTION & MODULE

DATA MANIPULATION IN PYTHON

Function & Module

FUNCTION IN PYTHON :-

Python function is a block of statements that return the specific task. The idea is to put some commonly or repeatedly done task together and make a function so that instead of writing the same code again and again for different inputs, we can do the function calls to reuse the code contained in it over and over again .

Some benefits of using Functions :

- Increase code Readability
- Increase code Reusability

Types of functions :-

There are mainly two types :-

- Built -in library Function : These are Standard Function in python that are available to use.
- User - defined Function : We can create our own function based on our requirements

Defining and calling a function with parameters :-

```
Def function_name(parameter : data_type)
return _type:
  “””Docstring”””
  #body of the function
Return expression
```

Function Examples :-

```
>>>def hello(name)
… results =”Hello World to”+name+” ”
… return results
…
>>> hello()
Traceback(most recent call last):
File “<stdin>”.line 1,in <module>
TypeError.hello() takes exactly 1 argument (0   given)
‘Hello World to Text’
>>>
```

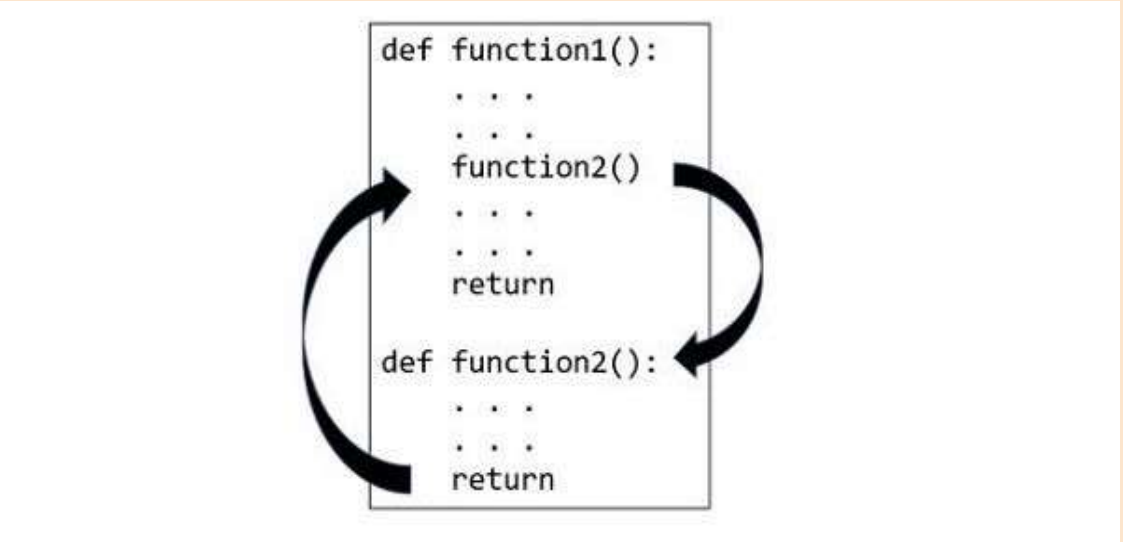
TASK : make script from this take name from command line

PRINT: result to screen

A function is a block of organized, reusable code that is used to perform a single, related action. Functions provide better modularity for your application and a high degree of code reusing.

A top-to-down approach towards building the processing logic involves defining blocks of independent reusable functions. A function may be invoked from any other function by passing required data (called parameters or arguments).

The called function returns its result back to the calling environment.



MODULE IN PYTHON :-

A python module is a file containing Python definitions and statements .

A module can define functions , classes, and variables .

A module can also include runnable code.

Grouping related code into a module makes the code easier to understand and use.

It also makes the code logically organized.

IMPORT MODULE

```
import example_module
```

A module is a file containing Python definitions and statements.

The file name is the module name with the suffix .py appended.

Within a module, the module's name (as a string) is available as the value of the global variable __name__.

For instance, use your favorite text editor to create a file called fibo.py in the current directory with the following contents.

Modules can import other modules. It is customary but not required to place all import statements at the beginning of a module (or script, for that matter). The imported module names, if placed at the top level of a module (outside any functions or classes), are added to the module's global namespace.

USING MODULE

- One of two import statements :

  1) import module name

  2)from module import function/constant

- If method 1:

  modulename.function(arguments)

- If method 2:

  1) function(arguments)- module name not needed

  2)beware of function name collision

- **Python itself allows for the creation of modules.**
- **Similar to the re (regular expression) module, a module can be primarily written in C programming language and then dynamically inserted at run-time.**
- **A built-in module, such as the itertools module, is inherently included in the interpreter**

A module is a file containing Python code, definitions of functions, statements, or classes. An example_module.py file is a module we will create and whose name is example_module.

Difference between Function & Module :-

The main difference between a module and a function is that a module is a collection of functions that are imported in multiple programs and can do various tasks.

A function is a small block of code and separates itself from the entire code and have a fixed functionality. This function can be used anywhere within the same program whereas modules can be used in multiple programs.

Conclusion :-

Modules and functions both have one main goal that is code reusability. Functions are used for small tasks whereas modules are used for larger tasks as it allows various classes and functions in it.

A module is used by importing it in another program where as a function is used by calling it.

DATA MANIPULATION IN PYTHON

Data Manipulation

Data manipulation with python is defined as a process in the python programming language that enables users in data organization in order to make reading or interpreting the insights from the data more structured and comprises of having better design.

It helps data analysts to make sense of complicated data sets and make them easier to understand. Another pro of using Python is its high readability.

The key feature of data manipulation is enabling faster business operations and also emphasize optimization in the process. Through proper manipulated data one can analyze trends, interpret insights from financial data, analyze consumer behaviour or pattern, etc

- Data Manipulation is one of the initial processes done in Data Analysis. It involves arranging or re arranging data points to make it easier for users/data analysts to perform necessary insights or business directives.
- Data Manipulation encompasses a broad range of tools and languages, which may include coding and non-coding techniques.
- It is not only used extensively by Data Analysts but also by business people and accountants to view the budget of a certain project.
- It also has its programming language, DML (Data Manipulation Language) which is used to alter data in databases. Let’s know what exactly Data manipulation is.

Data Manipulation Technique

1) Filtering - used for selecting a subset of the data based on a condition. This is done using boolean indexing or the query function.
2) Aggregation - used for computing summary statistics, such as mean, median, and standard deviation. This is done using the groupby function.
3) Sorting - used for sorting the data based on one or more columns. This is done using the sort_values function.
4) Joining - used for combining multiple data sources based on a common key. This is done using the merge function.
5) Reshaping - used for transforming the data from one shape to another. This is done using the pivot, melt, and stack/unstack functions.

Operations performing in Data Manipulation :-

- **Data Preprocessing**: Most of the raw data that is mined may contain errors, missing values and mislabeled data. This will hamper the final output if it is not dealt with in the initial stages.
- **Structuring data** (if it is unstructured): If there's any sort of data available in the database which can be structured into a table to query them effectively, we sort those data into tables for greater efficiency
- **Reduce the number of features:** As we know, data analysis is inherently computationally intensive. As a result, one of the reasons to perform data manipulation is to find out the optimum number of features needed for getting the result
- **Clean the data:** Delete unnecessary data points or outliers which may affect the final output. This is done to streamline the output.
- **Transforming data:** Some insights into data can be improved by transforming the data. This may involve transposing data, and arranging/rearranging them.

Uses of Data Manipulation

In today's world where every business has become competitive and undergoing digital transformation, the right data is paramount for all decision-making abilities.

Hence, to achieve our results easier and faster, we implement data manipulation.

There are many reasons why we need to manipulate our data. They are:

- *Increased Efficiency.*
- *Less Room for Error.*
- *Easier to Analyze data.*
- *Fewer chances for unexpected results.*

Conclusion

Due to unrestricted globalization, and near-digitization of all industries, there is a greater need for correct data for good business insights.

This calls for even more rigorous **Data Manipulation Techniques** in both the coding sphere and the **low code / no code** spheres.

Various programming languages and tools, such as Python with libraries like pandas, **R**, **SQL**, and Excel, are commonly used for data manipulation tasks. Data Manipulation may be hard if the data mined is unreliable.

Hence there are even more regulations on **data mining**, **Data Manipulation** and **Data Analysis**.


THANK
YOU