**Q1. What is the difference between list and tuples in Python?**

|  |  |
| --- | --- |
| **LIST vs TUPLES** | |
| **LIST** | **TUPLES** |
| Lists are mutable i.e they can be edited. | Tuples are immutable (tuples are lists which can’t be edited). |
| Lists are slower than tuples. | Tuples are faster than list. |
| Syntax: list\_1 = [10, ‘Chelsea’, 20] | Syntax: tup\_1 = (10, ‘Chelsea’ , 20) |

Q2. What is pass in Python?

The pass keyword represents a null operation in Python. It is generally used for the purpose of filling up empty blocks of code which may execute during runtime but has yet to be written. Without the pass statement in the following code, we may run into some errors during code execution.

Q3. What is \_\_init\_\_?

\_\_init\_\_ is a contructor method in Python and is automatically called to allocate memory when a new object/instance is created. All classes have a \_\_init\_\_ method associated with them. It helps in distinguishing methods and attributes of a class from local variables.

Q4. What is docstring in Python?

Documentation string or docstring is a multiline string used to document a specific code segment. The docstring should describe what the function or method does.

Q5. Explain how can you make a Python Script executable on Unix?

Script file must begin with #!/usr/bin/env python

Q6. What is lambda in Python? Why is it used?

Lambda is an anonymous function in Python, that can accept any number of arguments, but can only have a single expression. It is generally used in situations requiring an anonymous function for a short time period.

Q7. What is pickling and unpickling?

Python library offers a feature - serialization out of the box. Serializing an object refers to transforming it into a format that can be stored, so as to be able to deserialize it, later on, to obtain the original object. Here, the pickle module comes into play.

Pickling: Pickling is the name of the serialization process in Python. Any object in Python can be serialized into a byte stream and dumped as a file in the memory. The process of pickling is compact but pickle objects can be compressed further. Moreover, pickle keeps track of the objects it has serialized and the serialization is portable across versions. The function used for the above process is pickle.dump() .

Unpickling: Unpickling is the complete inverse of pickling. It deserializes the byte stream to recreate the objects stored in the file and loads the object to memory. The function used for the above process is pickle.load()

Q8. What are generators in Python?

Generators are functions that return an iterable collection of items, one at a time, in a set manner. Generators, in general, are used to create iterators with a different approach. They employ the use of yield keyword rather than return to return a generator object.

Q9. What is the difference between .py and .pyc files?

.py files contain the source code of a program. Whereas, .pyc file contains the bytecode of your program. We get bytecode aer compilation of .py file (source code). .pyc files are not created for all the files that you run. It is only created for the files that you import.

Q10. Explain how to delete a file in Python?

Use command os.remove(file\_name)

Q11. What are negative indexes and why are they used?

Negative indexes are the indexes from the end of the list or tuple or string. Arr[-1] means the last element of array Arr[].

Q12. Define pandas dataframe.

A dataframe is a 2D mutable and tabular structure for representing data labelled with axes - rows and columns.

Q13. How will you identify and deal with missing values in a dataframe?

We can identify if a dataframe has missing values by using the isnull() and isna() methods. missing\_data\_count=df.isnull().sum()

We can handle missing values by either replacing the values in the column with 0 as follows: df[‘column\_name’].fillna(0)

Or by replacing it with the mean value of the column df[‘column\_name’] = df[‘column\_name’].fillna((df[‘column\_name’].mean()))

Q14. What do you understand by reindexing in pandas?

Reindexing is the process of conforming a dataframe to a new index with optional filling logic. If the values are missing in the previous index, then NaN/NA is placed in the location. A new object is returned unless a new index is produced that is equivalent to the current one. The copy value is set to False. This is also used for changing the index of rows and columns in the dataframe.

Q15. Define GIL.

GIL stands for Global Interpreter Lock. This is a mutex used for limiting access to python objects and aids in effective thread synchronization by avoiding deadlocks. GIL helps in achieving multitasking (and not parallel computing).

Q16. Differentiate between deep and shallow copies.

Shallow copy does the task of creating new objects storing references of original elements. This does not undergo recursion to create copies of nested objects. It just copies the reference details of nested objects.

Deep copy creates an independent and new copy of an object and even copies all the nested objects of the original element recursively