

Travelers EDP Week 3 Review

1. Which of the following statements are false about Python? Select all that apply.
 - a) Python variable names may contain letters, digits, and underscores may not start with digits.
 - b) Python variable names are case sensitive.
 - c) In Python, you need to declare variable data types.
 - d) Jupyter is a browser-based python REPL.
 - e) None of the above
2. Which of the following statements are true in python environment? Select all that apply.
 - a) A Python module contains definitions of functions, classes, constants, and objects.
 - b) A Python package contains definitions of functions, classes, constants, and objects.
 - c) The Python sequence types are bool, int, float, and complex.
 - d) The Python sequence types are str, bytes, list, and tuple.
 - e) None of the above
3. What is lambda in Python?
 - a) A special class
 - b) Arrow function
 - c) A small anonymous function
 - d) A special package to perform mathematical operations
 - e) None of the above
4. What will be the output of the following python code?

```
i = 1
while True:
    if i%3 == 0:
        break
    print(i)

    i += 1
```

- a) 1 2 3
 - b) 1 2
 - c) **SyntaxError: invalid syntax**
 - d) No output
 - e) None of the above
5. Which of the following is used to define a block of code in Python?
 - a) Curly brackets
 - b) Brackets
 - c) **Indentation**
 - d) All the above
 - e) None of the above

6. What will be the output of the following Python code?

```
len([1,4, 8, 'awesome', 24, 'love'])
```

- a) 4
- b) 6
- c) SyntaxError: invalid syntax
- d) 32
- e) None of the above

7. What will be the output of the following Python code?

```
for i in [1, 2, 3, 4, 5][::-1]:  
    print (i)
```

- a) 0 1 2 3 4
- b) 1 2 3 4 5
- c) 5 4 3 2 1
- d) SyntaxError: invalid syntax
- e) None of the above

8. What will be the output of the following Python code?

```
message = "cool";  
x = 10;  
output = message + x;  
print(output)
```

- a) cool10
- b) TypeError
- c) No output
- d) coolcoolcoolcoolcoolcoolcoolcoolcoolcool
- e) None of the above

9. What is ETL in data engineering?

- a) Extract, Transfer, and Learn
- b) Extract, Transform, and Load
- c) Extract, Transform, and Lean
- d) Estimate, Transform, and Load
- e) None of the above

10. What are the 3 primary characteristics of Big Data?
- a) Volume, Velocity, Vertical Scaling
 - b) Volume, Variety, Velocity
 - c) Volume, Veracity, Vertical
 - d) Volume, Velocity, Vertical
 - e) None of the above
11. Which of the following statements are false about Data Engineering?
- a) Data processing includes data acquisition, data transfer, data transformation and data storage.
 - b) Apache Parquet is an open-source columnar storage data format.
 - c) Data Sharding is a technique to achieve vertical scalability.
 - d) Data Sharing is a technique to achieve horizontal scalability.
 - e) None of the above
12. Which of the following are part of Hadoop ecosystem?
- a) Apache Pig
 - b) Apache Hive
 - c) Apache Sqoop
 - d) Apache Flume
 - e) All the above
 - f) None of the above
13. Which of the following statements are true?
- a) The syntax `*args` in python function definitions is used to pass a variable number of arguments to a function.
 - b) The syntax `**kwargs` in python function definitions is used to pass a keyworded variable length argument list to a function.
 - c) DAG is an acronym for Database Aligned Graph in data engineering.
 - d) REPL is an acronym for Read Evaluate Print Lambda
 - e) None of the above
14. What are the desirable properties of any distributed system in data engineering?
- a) Consistency
 - b) High Availability
 - c) Tolerance to network partitions
 - d) All the above
 - e) None of the above

15. What will be the output of the following python code?

```
deposits = (10000, 20000, 50000, 100000)
interest = 0.04
print(list(map(lambda i: i * interest, deposits)))
```

- a) [10000.0, 20000.0, 50000.0, 100000.0]
- b) [10400.0, 20800.0, 52000.0, 10400.0]
- c) [400.0, 800.0, 2000.0, 4000.0]
- d) SyntaxError: invalid syntax
- e) None of the above

16. Which of the following is a package manager for Python?

- a) npm
- b) maven
- c) gradle
- d) pip
- e) none of the above

17. Which of the following functions is a built-in function in Python?

- a) seed()
- b) print()
- c) sqrt()
- d) factorial()
- e) none of the above

18. The Hadoop Distributed File System is a distributed, scalable, fault-tolerant, and portable file system. Is this True or False?

- a) True
- b) False
- c) Not Sure

19. Which of the following data related roles are common in data driven organizations?

- a) Data Analyst
- b) Data Scientist
- c) Data Engineer
- d) All the above
- e) None of the above

20. Each object in Python has a unique id. We can use the id() function to get the object's id. Is this True or False?

- a) True
- b) False
- c) Not sure