FN_Python SQL Azure AWS Questions

Test Summary
- No. of Sections: 4
- No. of Questions: 80
- Total Duration: 180 min

Section 1 - Python

Section Summary
- No. of Questions: 20
- Duration: 45 min

Additional Instructions:

None

Q1.   Consider the code given below.
What must be written in Line1 and Line 2 in the below code such that the function returns the maximum price for the given input?
Output
450

```
1   def find_price(product_price_list):
2     if len(product_price_list) == 1:
3        #Line1
4     else:
5        #Line2
6        if price > product_price_list[0]:
7           return price
8        else:
9           return product_price_list[0]
10  product_price_list = [100, 450, 270, 12, 58, 102, 375]
11  print (find_price(product_price_list))
12
```

    Line 1: price = product_price_list[0]
    Line 2: price = find_price(product_price_list[1:])

    Line 1: return product_price_list[0]
    Line 2: price = find_price(product_price_list[1:])

    Line 1: return product_price_list[0]
    Line 2: price = find_price(product_price_list)

    Line 1: return product_price_list[0]
    Line 2: return find_price(product_price_list[1:])

Q2.   What is the output of the code given below?

```
1   def demo(numl,num2):
2     try:
3        if(num==2):
4           nun3=numl/num2
5        else:
6           num3=num3/num2
7        print("Number divided")
8     except NameError:
9        print("Name not found")
10    except TypeError:
11       print("Invalid type")
12    except ZeroDivisionError:
13       print ("Division by zero")
14  try:
15    demo('2',0)
16    print("Done")
17  except:
18    print("Error occurred")
19  finally:
20    print("Program done")
```

Name not found
Done
Program done

Division by Zero
Error occurred
Program done

Invalid type
Done
Program Done

Name not found
Error occurred
Program done

Q3.   Consider the code given below.

Which of the following function calls will execute successfully?

```
1  def info(name, no_of_seats=100, *marks):
2      pass
3
```

info("John",350,21)

info("John")

info("John",100,21,46)

All the given function Calls

Q4.   What is the output of the code given below ?

```
1  values = ["823", "863"]
2  num = values[0][0:]
3  for row in range(0, len(values)):
4      for coloumn in range(0, len(values[row])):
5          if num > values[row][coloumn:]:
6              num = values[row][coloumn:]
7  print(num)
8
```

2

3

23

823

```
1    class NonDivisibilityException(Exception):
2        pass
3    class CheckStatus:
4        def is_div_by_two(self,list1):
5            try:
6                for value in list1:
7                    if value%2 !=0:
8                        raise NonDivisibilityException()
9                print("Divisible by 2")
10           except NonDivisibilityException:
11               print("Number Exception - inside")
12   try:
13       CheckStatus().is_div_by_two([2,13,22,3])
14   except NonDivisibilityException:
15       print("Number Exception - outside")
16   except Exception as e:
17       print("Some error occurred")
18   finally:
19       print("Inside finally")
20   print("Success")
```

Number Exception - inside
Number Exception - outside
Inside finally
Success

Number Exception - inside
Inside finally
Success

Number Exception - outside
Some error occurred

Number Exception - inside

Q6.   What is the output of the given code below?

```
1    def retrive_num( number, value):
2        if(number <(value//2)):
3            return 1
4        elif(number%2==0):
5            return retrive_num(number-1,value-1)
6        else:
7            return (value-1)*retrive_num(number-1,value-1)
8    print(retrive_num(7,4))
```

315

4

-3

10

Q7.   What is the output of the code given below ?

```
1    def function(input_list):
2        mid_pos=len(input_list)//2
3        low=0
4        high=len(input_list)-1
5        while(input_list[mid_pos]<input_list[low]):
6            low=low+1
7        if(low < high):
8            temp=input_list[low]
9            input_list[low]=input_list[high]
```

```
10        input_list[high]=temp
11     return input_list
12  list1 = [39,91,77,51,33,84]
13  sub_list = function(list1[:4])
14  print(sub_list)
15  |
```

[51, 91, 39, 77]

[51, 91, 77, 39]

[51, 91, 77, 84]

[39, 91, 77, 84]

Q8.   Evelyn is new to python programming. She has written the below code and found the **account_balance** is access outside the class.

Help her to choose the correct option to prevent the access of **account_balance** outside the Account class.

```
1  class Account:
2    def __init__(self,acc_name,acc_balance):
3       self.account_name= acc_name
4       self.accoutn_balance = acc_balance
5    |
```

---

self.account_balance = __acc_balance

---

private self.accoutn_balance = acc_balance

---

self.__account_balance = acc_balance

__self.account_balance = acc_balance

Q9.   What is the output of the code given below?

```
1  class NegativeMarksException(Exception):
2     pass
3  class Student:
4     counter = 0
5     def __init__(self, marks1, marks2):
6        self.__marks1 = marks1
7        self.__marks2 = marks2
8     def validate(self):
9        try:
10           if(self.__marks1 < 0):
11              Student.counter += 1
12              raise NegativeMarksException
13           if (self.__marks2 < 0):
14              Student.counter += 1
15              raise NegativeMarksException
16           print('Success')
17        finally:
18           Student.counter += 2
19        Student.counter += 1
20  try:
21     student1 = Student(4, -2)
22     student1.validate()
23     Student.counter += 1
24  except NegativeMarksException:
25     Student.counter += 1
26     print('Marks cannot be negative')
27  print(Student.counter)
```

> Marks cannot be negative
> 5

> Marks cannot be negative
> 4

> Marks cannot be negative
> 3

> Marks cannot be negative
> 2

Q10.   Which of the following two function calls will execute successfully?
        choose two correct options

```
1   def display_cat_details(color,name,pet_name=None):
2       print("Color:")
3       print(color)
4       print("Names")
5       print(name)
6       if pet_name!=None:
7           print(pet_name)
```

> display_cat_details("Brown","Nelly")

> display_cat_details("White")

> display_cat_details("Grey",pet_name="Grey")

> display_cat_details("Black","Roger","show")

Q11.   Predict the output for the following snippet.

```
1    value = [10, 20, 30, 40]
2    data = 10
3    try:
4        data = value[3]
5    except IndexError:
6        print('Index Error  ', end = '')
7    except:
8        print('Error  ', end = '')
9    finally:
10       print('Final block  ', end = '')
```

> Index Error

> Error

> Final block

> Error Final block

Q12.   Consider the below code snippet.

        Identify the most efficient test data set for testing the below code using 'Logic Coverage Technique'.

```
1    if previous_year_percentage>=75 and previous_year_percentage<=85:
2        scholarship=5000
```

```
3    elif previous_year_percentage>85 and previous_year_percentage<=95:
4       scholarship=8000
5    elif previous_year_percentage>95:
6       scholarship=10000
7    else:
8       scholarship=0
9
10
11   Option A - 79,87,91,99
12   Option B - 78,80,92,99
13   Option C - 74,77,90,100
14   Option D - 74,75,76,84,85,86,94,95,96,97
```

<table>
<tr><td>Option A</td></tr>
<tr><td>Option B</td></tr>
<tr><td>Option A & B</td></tr>
<tr><td>Option C & D</td></tr>
</table>

Q13. What is the output of the below python code?

```
1    class Employee:
2       __counter = 100
3
4       def __init__(self,name):
5          self.name = name
6          Employee.__counter+=1
7          self.id=Employee.__counter
8
9       @staticmethod
10      def get_counter():
11         return Employee.__counter
12
13
14   employee1= Employee("Tina")
15   employee2= Employee("Chris")
16   employee3= Employee("Robb")
17   print(Employee.get_counter()-100)
18
```

<table>
<tr><td>1</td></tr>
<tr><td>3</td></tr>
<tr><td>103</td></tr>
<tr><td>0</td></tr>
</table>

Q14. Consider the below code snippet.

How many static variables, local variables, instance variables respectively are present in the code snippet?

Notes:
While counting the variables do not consider **self** present at Line 1.
Line numbers are only for reference.

```
1    class Car:
2       __counter = 100
3       types = ['SUV', 'Hatchback', 'Coupe']
4       def __init__(self,model,doors): #Line1
5          self.model = model
6          self.doors = doors
7          self.color = None
8
9    car1 = Car("Ford", 4)
10   Car2 = Car("Porsche", 2)
11
```

Q15.   What is the content of dictionaries after the below code is executed?

```
1  def dict_items(dict1):
2    global dict2
3    #keys() returns the collection of keys in the dictionary
4    for key in dict1.keys():
5      dict2[key+1]=dict1[key]+key
6      dict1[key]=dict2[key+1]
7  dict2={}
8  dict_items({1:1,2:22,3:33,4:44,5:55})
9
```

dict1 - {1: 2, 2: 24, 3: 36, 4: 48, 5: 60}
dict2 - {2: 2, 3: 24, 4: 36, 5: 48, 6: 60}

dict1 - {1: 2, 2: 24, 3: 36, 4: 48, 5: 60}
dict2 - {1: 2, 2: 24, 3: 36, 4: 48, 5: 60}

dict1 - {1: 2, 2: 24, 3: 36, 4: 48, 5: 60}
dict2 - { }

dict1 - {1: 2, 2: 2, 3: 24, 4: 36, 5: 48, 6: 60}
dict2 - {2: 2, 3: 24, 4: 36, 5: 48, 6: 60}

Q16.   What will be the output of the below code?

```
1  class NonDivisibilityException(Exception):
2    pass
3  class CheckStatus:
4    def is_div_by_two(self,list1):
5      try:
6        for value in list1:
7          if value%2!=0:
8            raise NonDivisibilityException()
9        print("Divisible by 2")
10       except NonDivisibilityException:
11         print("Number Exception -Insdie")
12 try:
13   CheckStatus().is_div_by_two([2,13,22,3])
14 except NonDivisibilityException:
15   print("Number Exception - Outside")
16 except Exception as e:
17   print("Some error occurred")
18 finally:
19   print("Inside finally")
20 print("Success")
```

Number Exception - Inside
Number Exception - Outside
Inside finally
Success

> Number Exception - Inside
> Some error occurred
> Inside finally
> Success

> Divisible by 2
> Number Exception - Inside
> Divisible by 2
> Number Exception - Inside
> Inside finally
> Success

> Number Exception - Inside
> Inside finally
> Success

Q17.   What is the output of the below Python code?

```python
1   class ClassA:
2       def __init__(self,num):
3           self.__num=num
4       def get_num(self):
5           return  self.__num
6       def set_num(self, num):
7           self.__num = num
8       def method1(self,var):
9           return  self.__num+var
10  class ClassB:
11      def __init__(self,ref_a):
12          self.ref_a=ref_a
13          self.num1=10
14      def method2(self,val):
15          self.ref_a.set_num(self.num1+val)
16      def method3(self):
17          return self.ref_a.method1(3)+self.num1
18
19  obj_a=ClassA(5)
20
21  obj_b=ClassB(obj_a)
22  obj_b.method2(3)
23  print(obj_b.method3())
```

> 27

> 19

> 26

> 18

Q18.   What will be the output of the following Python code?

```python
1   x=10
2   y=8
3   assert x>y, 'X too small'
4
```

> Assertion Error

10 8

No output

108

Q19.   What will be the output of the following Python code?

```
1   def getMonth(m):
2     if m<1 or m>12:
3         raise ValueError("Invalid")
4     print(m)
5   getMonth(6)
6   |
```

ValueError

Invalid

6

ValueError("Invalid")

Q20.   Compare the following two Python codes shown below and state the output if the input entered in each case is -6?

```
1    ##########CODE 1
2    import math
3    num=int(input("Enter a number of whose factorial you want to find"))
4    print(math.factorial(num))
5
6
7
8    ##########CODE 2
9    num=int(input("Enter a number of whose factorial you want to find"))
10   print(math.factorial(num))
11   |
```

ValueError, NameError

AttributeError, ValueError

NameError, TypeError

TypeError, ValueError

Section 2 - SQL

Section Summary

- No. of Questions: 20
- Duration: 45 min

Additional Instructions:

None

Q1. You need to find out the employees which belong to the department of 'Jessica Butcher' and have salary greater than the salary of 'Jessica Butcher' who has an employee ID of 40 and department id of 100. Which of the following queries will work?

```
SELECT first_name, last_name
FROM employees
WHERE last_name = 'Butcher'
AND first_name = 'Jessica'
AND salary > 10000;
```

```
SELECT first_name, last_name
FROM employees
WHERE department = 100;
```

```
SELECT first_name, last_name
FROM employees
WHERE department = (SELECT department FROM employees
WHERE first_name = 'Jessica'
AND last_name = 'Butcher'
AND employee_id = 40)
AND salary > (SELECT salary FROM employees
WHERE first_name = 'Jessica'
AND last_name = 'Butcher'
AND employee_id = 40);
```

```
SELECT first_name, last_name
FROM employees
WHERE department = (SELECT department
FROM employees
WHERE first_name = 'Jessica'
AND last_name = 'Butcher'
AND department = 100);
```

Q2. What is true about using NOT IN when writing queries with sub-queries in them?

NOT IN ignores all the NULL values and gives only the NOT NULL values

NOT IN puts all the NULL values at the last and gives the NOT NULL to be displayed first

NOT IN should be not be used if a NULL value is expected in the result set

NOT IN is just a negation of the operator IN and can be changed without any caveat.

Q3. Examine the table structure as given.

Table Name: employees

```
EMPLOYEE_ID NOT NULL NUMBER(6)
FIRST_NAME VARCHAR2(20)
LAST_NAME NOT NULL VARCHAR2(25)
EMAIL NOT NULL VARCHAR2(25)
PHONE_NUMBER VARCHAR2(20)
HIRE_DATE NOT NULL DATE
JOB_ID NOT NULL VARCHAR2(10)
SALARY NUMBER(8,2)
COMMISSION_PCT NUMBER(2,2)
MANAGER_ID NUMBER(6)
DEPARTMENT_ID NUMBER(4)
```

This below query returns an error. What is the reason for error?

```
1  SELECT first_name, last_name
2  FROM employees
3  WHERE employee_id NOT IN
4  (SELECT manager_id, hire_date
5  FROM employees
6  WHERE manager_id is not null);
```

The NOT IN operator used is invalid

The WHERE clause in the sub-query is incorrectly written

The column in the sub-query SELECT clause should only be one when there's an inequality
used in the main query

The sub-query uses the same table as the main query

Q4.   What will be the outcome of the following query?
        Consider the given table structure:

        EMPLOYEE_ID NOT NULL NUMBER(6)
        FIRST_NAME VARCHAR2(20)
        LAST_NAME NOT NULL VARCHAR2(25)
        EMAIL NOT NULL VARCHAR2(25)
        PHONE_NUMBER VARCHAR2(20)
        HIRE_DATE NOT NULL DATE
        JOB_ID NOT NULL VARCHAR2(10)
        SALARY NUMBER(8,2)
        COMMISSION_PCT NUMBER(2,2)
        MANAGER_ID NUMBER(6)
        DEPARTMENT_ID NUMBER(4)

```
1   SELECT first_name, last_name, salary
2   FROM employees
3   WHERE salary ANY (SELECT salary FROM employees);
4   |
```

It executes successfully giving the desired results

It executes successfully but does not give the desired results

It throws an error

It executes successfully and gives two values for each row obtained in the
result set

Q5.   You need to find the 3rd maximum salary from the 'employees' table. Which of the following queries will give you the required
        results? Choose the most appropriate answer.

SELECT * FROM employees E WHERE salary = (SELECT count(distinct salary)
FROM employees S WHERE E.salary = S.salary);

SELECT * FROM employees E WHERE 1 = (SELECT count(distinct salary)
FROM employees S WHERE E.salary < S.salary)

 SELECT * FROM employees E WHERE 2 = (SELECT count(distinct salary)
FROM employees WHERE e.salary > S.salary)

```
SELECT * FROM employees E WHERE 3 = (SELECT count(distinct salary)
FROM employees S WHERE S.salary > E.salary);
```

Q6. Consider the two tables in a relational database with columns and rows as follows

Table : STUDENT

Roll No.	Name	Dept_id
1	Suresh	1
2	Mahesh	1
3	Ramesh	2
4	Paresh	3

Table : Department

Dept_id	Dept_Name
1	A
2	B
3	C

Roll_No is the Primary Key of Student_table
Dept_id is the Primary Key of Department table
Student.Dept _id is a Foreign Key refers Department.Dept_id
What will happen if we try to execute the following 2 SQL statement ?
(i) Update student Set Dept_id = NULL where Roll_No. = 1
(ii) Update Department Set Dept_id = NULL where Dept_id = 1

Both (i) and (ii) will fail

(i) will fail but (ii) will succeed

(i) will succeed but (ii) will fail

Both (i) and (ii) will succeed

Q7.

Workid	MobileNumber	JobTitle	Wagesperday
W01	9876543210	Carpenter	300
W02	9785461230	Electrician	500
W03	9632587410	Carpenter	400
W04	9630147852	Electrician	500
W05	9514786301	plumber	380
W06	7854601239	painter	460
W07	9894850636	Carpenter	300
W08	7708132548	painter	400

Workid is the primary key.
Following index is created on the table worker.
IX1 - job title, wages per day

Which of the following queries will use TABLE SCAN ?

[Choose TWO correct options]

SELECT workerid, iobtitle, mobilenumber FROM worker WHERE workerid ="W06";

SELECT mobilenumber FROM worker WHERE LOWER {job title} = "electrician";

SELECT mobilenumber FROM worker WHERE LOWER {job title} = "plumber";

SELECT workerid, jobtitle, mobilenumber FROM worker WHERE workerid ="W07";

Q8.

saleid	bookid	salelocation	price	studentid
SA1	B1	Boston	50	S3
SA2	B2	San Fransisco	35	S1
SA3	B3	Seattle	40	S6
SA4	B2	New york	35	S4
SA5	B1	Detroit	50	S2
SA6	B3	Boston	40	S5
SA7	B3	Seattle	40	S6

Table :booksale

What will be the output when the above query is executed?

1 SELECT bookid, COUNT(*) NUMBEROFSALES
2 FROM booksale WHERE price €40 GROUP BY bookid HAVING COUNT(*) IN
3 (SELECT COUNT(*) FROM booksale
4 WHERE salelocation LIKE e GROUP BY salelocation);

bookid	Noof sales
B2	2
B3	3

bookid	Noof sales
B1	2
B2	3

bookid	Noof sales
B1	3
B2	2

None

Q9.   Consider the table book and bookissue :
Table : book

bookid	bookname	authorname
B978	Home Alone	Lauren Jen
B232	Jack and Jill	Enid Blyton
B567	The Three Musketeers	Alexander Dumas
B234	End Game	Stan Lee

Table :bookissue

customerid	bookid	dateofissue
C101	B978	09-Feb-18
C102	B232	23-Jan-17
C103	B978	27-Nov-18
C104	B567	20-Jan-18
C105	B234	07-Dec-18
C106	B978	27-Nov-18

What will be the output if the below query is executed ?

1 SELECT SUBSTR(customerid, 2) customerid, bi.bookid
2 FROM bookissue bi INNER JOIN book b ON bi.bookid = b.bookid
3 WHERE dateofissue >= '21-Jan-2018' AND LENGTH(bookname) > 9;

```
CUSTOMERID BOOKID
101 B978
103 B978
106 B978
105 B234
```

```
CUSTOMERID BOOKID
101 B978
103 B978
106 B978
104 B567
```

```
CUSTOMERID BOOKID
101 B978
103 B978
106 B978
104 B567
105 B234
```

```
CUSTOMERID BOOKID
101 B978
103 B978
106 B978
```

Q10.   Consider the table employee and output table:

Table :employee

empid	empname	salary	dept	manager
111	Tim	75000	ENG	NULL
112	Bob	90000	Finance	NULL
113	Catie	25000	Finance	112
114	Harry	30000	Finance	112
115	Peter	40000	ENG	111

Expected output :

empid	empname	mgrid	mgrname
115	Peter	111	Tim
114	Harry	112	Bob
113	Catie	112	Bob

Which of the following query will fetch the above expected output ?

```
SELECT E.empid EMPID, E.empname EMPNAME, M.empid MGRID,
M.empname MGRNAME FROM employee E INNER JOIN employee M ON
E.empid = M.manager;
```

```
SELECT E.empid EMPID, E.empname EMPNAME, M.empid MGRID,
M.empname MGRNAME FROM employee E INNER JOIN employee M ON
M.empid = E.manager WHERE M.empid IS NULL;
```

```
SELECT E.empid EMPID, E.empname EMPNAME, M.empid MGRID,
M.empname MGRNAME FROM employee E INNER JOIN employee M ON
E.manager = M.empid;
```

```
SELECT E.empid EMPID, E.empname EMPNAME, M.empid MGRID,
M.empname MGRNAME FROM employee E INNER JOIN employee M ON
M.manager = E.empid WHERE M.manager IS NULL;
```

Q11.   Consider the table employeedetail given below :

empid	empname	empsalary	bonus	project
1	John	27000	400	Finance
2	Michael	20000	300	Telecom
3	John	40000	600	ENG
4	Mike	50000	700	ENG
5	Robert	25000	NULL	Telecom
6	Robert	20000	300	Telecom
7	Peter	30000	NULL	ENG
8	Carl	27000	NULL	Finance

Find the output when below query is executed ?
1   SELECT project, ROUND(AVG(empsalary)) AVGSAL, SUM(bonus) TOTALBONUS FROM employeedetails
2   GROUP BY project HAVING AVG(empsalary) > 20000 AND SUM(bonus) > 500;

```
PROJECTDEPT AVGSAL TOTALBONUS
Telecom 21667 600
Finance 40000 600
```

```
PROJECTDEPT AVGSAL TOTALBONUS
ENG 40000 1300
Finance 40000 600
```

```
PROJECTDEPT AVGSAL TOTALBONUS
Telecom 21667 600
Finance 40000 600
ENG 50000 700
```

```
PROJECTDEPT AVGSAL TOTALBONUS
Telecom 21667 600
ENG 40000 1300
```

Q12.   Consider the table account given below :

Custno	accountno	Ifsccode	location	Amount	accounttype
C301	10530002231	BANK0001053	Mysore	140000	Savings
C302	12340004235	BANK0001234	Delhi	225000	Current
C303	87900002367	BANK0008790	Bangalore	79000	Savings
C304	56630003454	BANK0005663	Bangalore	152000	Savings
C305	78390002343	BANK0007839	Delhi	73000	Current

Requirement:

Jack and John are given with the requirement to retrieve the account details and sort details according to location (alphabetically) and in decreasing order of the amount in each of that location.

Jack :
SELECT custno, accountno, ifsccode, location, amount, accounttype FROM account ORDER BY location ASC, amount DESC;

John :
SELECT custno, accountno, ifsccode, location, amount, accounttype FROM account ORDER BY 5 DESC, 4 ASC;

Whose query will be correct to satisfy the above requirements ?

> Only Jack

> Only John

> Both Jack and John

> Neither Jack nor John

Q13.   Consider given tables : customer, purchase respectively.

custid	custname	custtype	gender
C9012	Maria	Regular	F
C9022	William	Privileged	M
C9032	Robert	Regular	M
C9042	Jenny	Privileged	F
C9052	Eliza	Elite	F
C9062	Jane	Regular	F

billid	custid	billamount	billdate
7001	C9052	860	17-Oct-18
7201	C9042	962	11-Oct-18
7401	C9022	720	11-Sep-18
7601	C9032	1002	17-Oct-18
7801	C9012	515	22-Sep-18
7501	C9052	430	11-Sep-18

Which customer(s) details will be fetched when the above query is executed ?

```
1   SELECT custid, custname, gender
2   FROM customer WHERE custid IN (
3   SELECT custid FROM purchase WHERE billdate = (SELECT MIN(billdate) FROM purchase));
```

Eliza and Robert

Jenny and Maria

William and Eliza

Robert and Jenny

Q14.   Consider a MongoDB collection name Item given below :

db.item.insert([{_id: 1, description: "Sugar", price: 60, discount: 10},
{_id: 2, description: "Vinegar", price: 110, discount: 15},
{_id: 3, description: "Tea", price: 200, discount: 20},
{_id: 4, description: "Biscuits", price: 50, discount: 20},
{_id: 5, description: "Coffee", price: 75, discount: 5}]);

John executed the below mongoDB statements :

Which of the following statements will be True when the above statements are executed ?

1 db.item.update({$and:[{description: "Tea"},{discount: 20}]}, {$set: {discount:10}});
2 db.item.remove({discount:20});

Two items will have discount of 10

The item collection will have four documents

Two items will have price less than 70

Two items will have discount of 20

Q15.   Consider the table structure for the table student given below.

Column Name	DataType and size	Constraint
id	NUMBER(5)	NOT NULL
name	CHAR(10)	
gender	CHAR(10)	
dob	DATE	
age	NUMBER(2)	

Following ALTER statements are successfully executed on:

ALTER TABLE student ADD emailid VARCHAR2(20);
ALTER TABLE student MODIFY gender CHAR(1);
ALTER TABLE student DROP (age);

Which of the below INSERT statement will execute successfully on the modified student table?

INSERT INTO student VALUES(10001, Rony,'M','12-IAN-95', 'rony@gmail.com');

INSERT INTO student VALUES(10001,'Rony','Male' ,'12-JAN-95','rony@gmail.com');

INSERT INTO student VALUES(10001,'Rony','Male','12-JAN-95',23);

INSERT INTO student VALUES(10001,'Rony','Male','12-JAN-95');

Q16.   Consider the table account given below :

accountnum	customerid	balance	accttype	status
100001	1001	9999	Savings	Closed
100002	1002	2500	Current	Active
100003	1003	3600	Savings	Active
100004	1004	5800	Savings	Active

Which of the following SQL query will help John to fetch the account numbers of customers who have an active account with a balance given ?

SELECT accountnum FROM account WHERE status = 'Active' AND balance IN (SELECT AVG(balance) FROM account);

SELECT accountnum FROM account WHERE status = 'Active' AND balance < (SELECT AVG(balance) FROM account GROUP BY accountnum, balance);

SELECT accountnum FROM account WHERE status = 'Active' GROUP BY accountnum HAVING AVG(balance) < (SELECT AVG(balance) FROm account GROUP BY accountnum);

SELECT accountnum FROm account WHERE status = 'Active' AND balance < (SELECT AVG(balance) FROM account)

Q17.   Consider the following query.

SELECT SUBSTR(TO_DATE('01-02-2018','MM-DD-YYYY'),4,3)month FROM DUAL;

What will be the output when the above query is executed ?

MONTH
Jan

MONTH
01

MONTH
02

MONTH
Feb

Q18.   Following are the functional dependencies:

itemid -> descr, cost, itemtype

cost - discount

Custid -> custname

custid, itemid-> amount

Which of the following statement is TRUE?

item is in 2NF and customer is in 2NF

item is in 3NF and customer is in 1NF

item is in 2NF and customer is in 1N

item is in 1NF and customer is in 3NF

Q19.   Consider the table trainee given :

traineename	location	dateofjoining
Alex	Delhi	02-Mar-18
Jack	Mysore	21-May-18
John	Mumbai	02-May-18
Angel	Chennai	18-May-18
Joana	Hyderabad	21-May-18
Mathew	Hyderabad	17-Feb-18

SELECT traineename, location FROM trainee WHERE
TO_CHAR(dateofjoining, 'mm') = 5 ORDER BY 1 DESC,location;

Which trainee's details will be fetched in the first record when it is executed ?

Jack

Angel

Joana

John

Q20.   Consider the table instructor given :

instructorid	name	subject	university	salary
I201	Alex	java	Harward	70000
I202	Sam	ruby	Oxford	75000
I201	Alex	rdbms	Harward	60000
I203	Mitchel	networking	Cambridge	50000
I202	sam	rdbms	Harward	40000
I203	Mitchel	.net	Oxford	50000

SELECT university FROM Instructor WHERE salary = 40000
GROUP BY university HAVING COUNT(instructorid) >1;

How many rows will be fetched in the output after executing the above code ?

1

4

3

2

Section 3 - AWS

Q1. To deal with a surge in customer interactions from a specific area, a company's social media manager asks additional personnel on weekends from the rest of the organization, which is granted. QuickSight is required by the business in order to create a report that visualizes the patterns in weekend activity over the previous six months.

What format should the data be provided in?

A line graph plotting customer contacts vs. time, with a line for each region

A pie chart per region plotting customer contacts per day of week

A map of regions with a heatmap overlay to show the volume of customer contacts

A bar graph plotting region vs. volume of social media contacts

Q2. Amazon Elastic MapReduce (EMR) is used by an organization to perform a sequence of extract-transform-load (ETL) stages. Each step's output must be completely processed in future stages or it will be discarded.

Which of the following methods will most effectively fulfill this requirement?

Use the EMR File System (EMRFS) to store the outputs from each step as objects in Amazon Simple Storage Service (S3).

Use the s3n URI to store the data to be processed as objects in Amazon S3.

Define the ETL steps as separate AWS Data Pipeline activities.

Load the data to be processed into HDFS, and then write the final output to Amazon S3.

Q3. A business runs a worldwide operation that is serviced by a single AWS region. The business wishes to expand into a new market. Within 24 hours of the product transaction, the regulator for that nation expects the Data Architect to keep a record of financial transactions in that country. The production application is indifferent to latency. Additionally, the new country includes an AWS region.

How can this need be met in the most cost-effective manner possible?

Use CloudFormation to replicate the production application to the new region.

Use Amazon CloudFront to serve application content locally in the country; Amazon CloudFront logs will satisfy the requirement.

Continue to serve customers from the existing region while using Amazon Kinesis to stream transaction data to the regulator.

Use Amazon S3 cross-region replication to copy and persist production transaction logs to a bucket in the new countrys region.

Q4. Using Kinesis streams and Lambda, an administrator processes events in near real time. Due to a 5-munite time restriction, Lambda sometimes fails to process batches from one of the shards.

What is a potential solution to this issue?

Add more Lambda functions to improve concurrent batch processing.

Reduce the batch size that Lambda is reading from the stream.

Ignore and skip events that are older than 5 minutes and put them to Dead Letter Queue (DLQ).

Configure Lambda to read from fewer shards in parallel.

Q5. An company must store and handle sensitive data on Amazon S3 and Amazon EMR. At rest and in transit, data must be secured on Amazon S3 and Amazon EMR. The Data Analysis team interacts with this data through Thrift Server and HIVE. The company wishes to restrict access to certain databases and tables, allowing only the SELECT query to be used.

Which method will safeguard the data and restrict user access to the SELECT query that returns a subset of the data?

Configure Transparent Data Encryption on Amazon EMR. Create an Amazon EC2 instance and install Apache Ranger. Configure the authorization on the cluster to use Apache Ranger.

Configure data encryption at rest for EMR File System (EMRFS) on Amazon S3. Configure data encryption in transit for traffic between Amazon S3 and EMRFS. Configure storage and SQL base authorization on HiveServer2.

Use AWS KMS for encryption of data. Configure and attach multiple roles with different permissions based on the different user needs.

Configure Security Group on Amazon EMR. Create an Amazon VPC endpoint for Amazon S3. Configure HiveServer2 to use Kerberos authentication on the cluster.

Q6. The data warehouse of a company stores sales data for reporting reasons. Employees are not permitted to obtain consumers' credit card information under data governance rules.
How can these rules be followed while yet allowing a Data Scientist to aggregate transactions using the same credit card number?

Store a cryptographic hash of the credit card number.

Encrypt the credit card number with a symmetric encryption key, and give the key only to the authorized Data Scientist.

Mask the credit card numbers to only show the last four digits of the credit card number.

Encrypt the credit card number with an asymmetric encryption key and give the decryption key only to the authorized Data Scientist.

Q7. Daily depletion reports from the field are received by a major food distributor in the form of gzip archives or CSV files uploaded to Amazon S3. The files are between 500MB and 5GB in size. Each day, these files are processed by an EMR task.

Recently, it has been noted that file sizes fluctuate and EMR tasks take an excessive amount of time. With this little information, the distributor must adjust and optimize the data processing workflow in order to enhance the EMR job's performance.

Which suggestion is appropriate for an administrator to make?

Reduce the HDFS block size to increase the number of task processors.

Use bzip2 or Snappy rather than gzip for the archives.

Decompress the gzip archives and store the data as CSV files.

Use Avro rather than gzip for the archives.

Q8.   A real estate business uses Apache HBase on Amazon EMR to power a mission-critical application. A single master node is setup for Amazon EMR. The company's data is kept on a Hadoop Distributed File System in excess of 5 TB (HDFS). The organization is looking for a cost-effective way to increase the availability of its HBase data.

Which architectural design best fulfills the needs of the business?

Use Spot Instances for core and task nodes and a Reserved Instance for the EMR master node. Configure the EMR cluster with multiple master nodes. Schedule automated snapshots using Amazon EventBridge.

Store the data on an EMR File System (EMRFS) instead of HDFS. Enable EMRFS consistent view. Create an EMR HBase cluster with multiple master nodes. Point the HBase root directory to an Amazon S3 bucket.

Store the data on an EMR File System (EMRFS) instead of HDFS and enable EMRFS consistent view. Run two separate EMR clusters in two different Availability Zones. Point both clusters to the same HBase root directory in the same Amazon S3 bucket.

Store the data on an EMR File System (EMRFS) instead of HDFS and enable EMRFS consistent view. Create a primary EMR HBase cluster with multiple master nodes. Create a secondary EMR HBase read-replica cluster in a separate Availability Zone. Point both clusters to the same HBase root directory in the same Amazon S3 bucket.

Q9.   A human resources organization runs analytics queries on the company's data using a 10-node Amazon Redshift cluster. The Amazon Redshift cluster comprises two tables: one for products and one for transactions, both of which have a product sku field. The tables span more than 100 GB. Both tables are used in the majority of queries.

Which distribution pattern should the organization adopt to optimize query speed for the two tables?

An EVEN distribution style for both tables

A KEY distribution style for both tables

An ALL distribution style for the product table and an EVEN distribution style for the transactions table

An EVEN distribution style for the product table and an KEY distribution style for the transactions table

Q10.   A business has an application that reads records from a Kinesis data stream using the Amazon Kinesis Client Library (KCL).

The application saw a considerable rise in use after a successful marketing effort. As a consequence, a data analyst was forced to separate certain data shards. When the shards were divided, the program began intermittently issuing ExpiredIteratorExceptions.

What is the data analyst's role in resolving this?

Increase the number of threads that process the stream records.

Increase the provisioned read capacity units assigned to the streams
Amazon DynamoDB table.

Increase the provisioned write capacity units assigned to the streams
Amazon DynamoDB table.

Decrease the provisioned write capacity units assigned to the streams
Amazon DynamoDB table.

Q11. Every ten seconds, a streaming application reads data from Amazon Kinesis Data Streams and promptly writes it to an Amazon S3 bucket. Data is being read from hundreds of shards by the application. Due to a different need, the batch interval cannot be modified. Amazon Athena has access to the data. As time passes, users notice a deterioration in query performance.

Which step may aid in query performance optimization?

Merge the files in Amazon S3 to form larger files.

Increase the number of shards in Kinesis Data Streams.

Add more memory and CPU capacity to the streaming application.

Write the files to multiple S3 buckets.

Q12. A major government entity is utilizing Amazon Managed Streaming for Apache Kafka to gather events from multiple internal applications (Amazon MSK).

To keep data distinct, the business has setup a separate Kafka topic for each application. To ensure data security, the Kafka cluster is set to accept only TLS encrypted data and to encrypt data in transit.

A recent application upgrade revealed that one of the apps had been set improperly, resulting in data being written to another application's Kafka topic. As data from numerous apps surfaced on the same subject, this resulted in many failures in the analytics pipeline. Following this occurrence, the organization want to prohibit applicants from writing to a subject other than the one to which they are supposed to write.

Which option satisfies these criteria with the least effort?

Create a different Amazon EC2 security group for each application. Configure each security group to have access to a specific topic in the Amazon MSK cluster. Attach the security group to each application based on the topic that the applications should read and write to.

Install Kafka Connect on each application instance and configure each Kafka Connect instance to write to a specific topic only.

Use Kafka ACLs and configure read and write permissions for each topic. Use the distinguished name of the clients TLS certificates as the principal of the ACL.

Create a different Amazon EC2 security group for each application. Create an Amazon MSK cluster and Kafka topic for each application. Configure each security group to have access to the specific cluster.

Q13. A business requires the collection of streaming data from several sources and storage on the AWS Cloud. Although the dataset is well organized, analysts must execute multiple sophisticated SQL queries with consistent performance. Certain types of data are searched more often than others. The organization is looking for a cost-effective solution that satisfies its performance criteria.

Which solution satisfies these criteria?

Use Amazon Managed Streaming for Apache Kafka to ingest the data to save it to Amazon S3. Use Amazon Athena to perform SQL queries over the ingested data.

Use Amazon Managed Streaming for Apache Kafka to ingest the data to save it to Amazon Redshift. Enable Amazon Redshift workload management (WLM) to prioritize workloads.

Use Amazon Kinesis Data Firehose to ingest the data to save it to Amazon Redshift. Enable Amazon Redshift workload management (WLM) to prioritize workloads.

Use Amazon Kinesis Data Firehose to ingest the data to save it to Amazon S3. Load frequently queried data to Amazon Redshift using the COPY command. Use Amazon Redshift Spectrum for less frequently queried data.

Q14. A manufacturing business stores its data on Amazon S3. The organization intends to employ AWS Lake Formation to secure such data assets at the granular level. Apache Parquet is used to store the data. The corporation has assigned a consultant a deadline for constructing a data lake.

How should the consultant approach developing the MOST COST-EFFECTIVE solution that satisfies these requirements?

Run Lake Formation blueprints to move the data to Lake Formation. Once Lake Formation has the data, apply permissions on Lake Formation.

To create the data catalog, run an AWS Glue crawler on the existing Parquet data. Register the Amazon S3 path and then apply permissions through Lake Formation to provide granular-level security.

Install Apache Ranger on an Amazon EC2 instance and integrate with Amazon EMR. Using Ranger policies, create role-based access control for the existing data assets in Amazon S3.

Create multiple IAM roles for different users and groups. Assign IAM roles to different data assets in Amazon S3 to create table-based and column-based access controls.

Q15. Salesforce, MySQL, and Amazon S3 are all used by a marketing organization to store data. The organization wishes to use data from these three sites in order to provide mobile dashboards for its consumers. The organization is unclear how to develop the dashboards and need a solution that requires as little modification and code as feasible.

Which solution satisfies these criteria?

Use Amazon Athena federated queries to join the data sources. Use Amazon QuickSight to generate the mobile dashboards.

Use AWS Lake Formation to migrate the data sources into Amazon S3. Use Amazon QuickSight to generate the mobile dashboards.

Use Amazon Redshift federated queries to join the data sources. Use Amazon QuickSight to generate the mobile dashboards.

Use Amazon QuickSight to connect to the data sources and generate the mobile dashboards.

Q16. An insurance firm has raw data in JSON format that is transferred to an Amazon S3 bucket on an ad hoc basis through an Amazon Kinesis Data Firehose delivery stream. Every eight hours, an AWS Glue crawler is scheduled to update the schema of the tables contained in the S3 bucket's data catalog. Apache Spark SQL is used by data analysts to analyze data on Amazon EMR, which is configured using AWS Glue Data Catalog as the metastore. According to data analysts, they sometimes obtain outdated data. A data engineer must ensure that users have access to the most current data.

Which solution satisfies these criteria?

> Create an external schema based on the AWS Glue Data Catalog on the existing Amazon Redshift cluster to query new data in Amazon S3 with Amazon Redshift Spectrum.

> Use Amazon CloudWatch Events with the rate (1 hour) expression to execute the AWS Glue crawler every hour.

> Using the AWS CLI, modify the execution schedule of the AWS Glue crawler from 8 hours to 1 minute.

> Run the AWS Glue crawler from an AWS Lambda function triggered by an S3:ObjectCreated:* event notification on the S3 bucket.

Q17. A multinational corporation is divided into sub-organizations, each of which offers its goods and services in a separate country. Senior management of the corporation needs to swiftly determine which sub-organization is the best performance in each country. All sales data is saved in the Parquet format on Amazon S3.

Which technique will result in the images required by senior leadership with the least amount of effort?

> Use Amazon QuickSight with Amazon Athena as the data source. Use heat maps as the visual type.

> Use Amazon QuickSight with Amazon S3 as the data source. Use heat maps as the visual type.

> Use Amazon QuickSight with Amazon Athena as the data source. Use pivot tables as the visual type.

> Use Amazon QuickSight with Amazon S3 as the data source. Use pivot tables as the visual type.

Q18. A database professional is required to evaluate and improve a performance-related Amazon DynamoDB table. The database professional determines that the partition key is generating the hot partitions, and hence creates a new partition key. The database professional must apply the new partition key to all current and new data in an efficient manner.

How does one go about implementing this solution?

> Use Amazon EMR to export the data from the current DynamoDB table to Amazon S3. Then use Amazon EMR again to import the data from Amazon S3 into a new DynamoDB table with the new partition key.

> Use AWS DMS to copy the data from the current DynamoDB table to Amazon S3. Then import the DynamoDB table to create a new DynamoDB table with the new partition key.

> Use the AWS CLI to update the DynamoDB table and modify the partition key.

> Use the AWS CLI to back up the DynamoDB table. Then use the restore-table-from-backup command and modify the partition key.

Q19. A database professional maintains a fleet of Amazon RDS database instances that are configured to utilize the default database parameter group. A database expert must connect a custom parameter group with certain database instances.

When will the instances be allocated to this new parameter group once the database specialist performs this change?

Instantaneously after the change is made to the parameter group

In the next scheduled maintenance window of the DB instances

After the DB instances are manually rebooted

Within 24 hours after the change is made to the parameter group

Q20.   A business examines historical data and requires access to data stored in Amazon S3. Each day, new data is created as.csv files and saved on Amazon S3.

Amazon Athena is being used by the company's analysts to run SQL queries on a recent subset of the company's entire data. The volume of data fed into Amazon S3 has risen significantly over time, as has the query latency.

Which options should the business consider using to boost query performance? (Select two.)

Use MySQL Workbench on an Amazon EC2 instance, and connect to Athena by using a JDBC or ODBC connector. Run the query from MySQL Workbench instead of Athena directly.

Use Athena to extract the data and store it in Apache Parquet format on a daily basis. Query the extracted data.

Run a daily AWS Glue ETL job to convert the data files to Apache Parquet and to partition the converted files. Create a periodic AWS Glue crawler to automatically crawl the partitioned data on a daily basis.

Run a daily AWS Glue ETL job to compress the data files by using the .gzip format. Query the compressed data.

Run a daily AWS Glue ETL job to compress the data files by using the .lzo format. Query the compressed data.

Section 4 - Azure

Section Summary
- No. of Questions: 20
- Duration: 45 min

Additional Instructions:
None

Q1.   You are developing a Microsoft Azure web application. The application will be deployed to 20 web role instances. A minimum of 18 running instances is needed to meet scaling requirements.
You need to configure the application so that upgrades are performed as quickly as possible, but do not violate scaling requirements.
How many upgrade domains should you use?

1

2

5

10

Q2.   You are developing an application that reads and writes data from a SQL Server database.
You need to ensure transactional data integrity.
Which isolation level should you use?

Serializable

ReadCommitted

ReadUncommitted

Normal

Q3.   You have a website that is hosted on Azure. You connect to the site by using the URI http://www.contoso.com. You plan to publish a new version of the website.
You need to acquire the publishing profile for the website.
Which two actions will achieve the goal? Each correct answer presents a complete solution.

Run the following Windows PowerShell cmdlet:Get-AzurePublishSettingsFile

Run the following Windows PowerShell cmdlet:Get-AzureSubscription

Navigate to the following URI: https://www.contoso.com/download/publishprofile.aspx

Navigate to the following URI: https://windows.azure.com/download/publishprofile.aspx

Q4.   You manage a cloud service that has a web application named WebRole1. WebRole1 writes error messages to the Windows Event Log.
Users report receiving an error page with the following message: "Event 26 has occurred. Contact your system administrator."
You need to access the WebRole1 event log.
Which three actions should you perform? Each correct answer presents part of the solution.

Enable verbose monitoring.

Update the WebRole1 web.config file.

Update the cloud service definition file and the service configuration file.

Run the Set-AzureVMDiagnosticsExtensionPowerShell cmdlet.

Run the Enable-AzureWebsiteApplicationDiagnostic PowerShell cmdlet.

Create a storage account.

Q5.   You are migrating an existing solution to Azure. The solution includes a user interface tier and a database tier. The user interface tier runs on multiple virtual machines (VMs). The user interface tier has a website that uses Node.js. The user interface tier has a background process that uses Python. This background process runs as a scheduled job. The user interface tier is updated frequently. The database tier uses a self-hosted MySQL database.
The user interface tier requires up to 25 CPU cores. You must be able to revert the user interface tier to a previous version if updates to the website cause technical problems. The database requires up to 50 GB of memory. The database must run in a single VM.

You need to deploy the solution to Azure.
What should you do first?

> Deploy the entire solution to an Azure Web App. Use a web job that runs continuously to host the database.

> Configure Microsoft Visual Team Services to continuously deploy the user interface tier to the Azure Web App service. Deploy the production builds and the staging builds of the user interface tier to separate slots.

> Deploy the entire solution to an Azure Web App. Use a web job that runs continuously to host the user interface tier.

> Deploy the user interface tier to a VM. Use multiple availability sets to continuously deploy updates from Microsoft Visual Studio Online.

Q6.   You manage a cloud service that utilizes data encryption.
You need to ensure that the certificate used to encrypt data can be accessed by the cloud service application.
What should you do?

> Upload the certificate referenced in the application package.

> Deploy the certificate as part of the application package.

> Upload the certificate's public key referenced in the application package.

> Use RDP to install the certificate.

Q7.   You have an Apache Hive table that contains one billion rows.
You plan to use queries that will filter the data by using the WHERE clause. The values of the columns will be known only while the data loads into a Hive table.
You need to decrease the query runtime.
What should you configure?

> static partitioning

> bucket sampling

> parallel execution

> dynamic partitioning

Q8.   You are configuring the Hive views on an Azure HDInsight cluster that is configured to use Kerberos.
You plan to use the YARN logs to troubleshoot a query that runs against Apache Hadoop.
You need to view the method, the service, and the authenticated account used to run the query.
Which method call should you view in the YARN logs?

> HQL

> WebHDFS

> HDFS C* API

Ambari RESR API

Q9.   Your team is building a data engineering and data science development environment.
      The environment must support the following requirements:
      - support Python and Scala
      - compose data storage, movement, and processing services into automated data pipelines
      - the same tool should be used for the orchestration of both data engineering and data science
      - support workload isolation and interactive workloads
      - enable scaling across a cluster of machines
      You need to create the environment.
      What should you do?

      Build the environment in Apache Hive for HDInsight and use Azure Data
      Factory for orchestration.

      Build the environment in Azure Databricks and use Azure Data Factory for
      orchestration.

      Build the environment in Apache Spark for HDInsight and use Azure
      Container Instances for orchestration.

      Build the environment in Azure Databricks and use Azure Container Instances
      for orchestration.

Q10. You are developing a data science workspace that uses an Azure Machine Learning service.
      You need to select a compute target to deploy the workspace.
      What should you use?

      Azure Data Lake Analytics

      Azure Databricks

      Azure Container Service

      Apache Spark for HDInsight

Q11. You are creating a machine learning model. You have a dataset that contains null rows.
      You need to use the Clean Missing Data module in Azure Machine Learning Studio to identify and resolve the null and missing data
      in the dataset.
      Which parameter should you use?

      Replace with mean

      Remove entire column

      Remove entire row

      Hot Deck

      Custom substitution value

      Replace with mode

Q12.

You are a data engineer implementing a lambda architecture on Microsoft Azure. You use an open-source big data solution to collect, process, and maintain data.

The analytical data store performs poorly.

You must implement a solution that meets the following requirements:

- Provide data warehousing
- Reduce ongoing management activities
- Deliver SQL query responses in less than one second

You need to create an HDInsight cluster to meet the requirements.

Which type of cluster should you create?

Interactive Query

Apache Hadoop

Apache HBase

Apache Spark

Q13.   You are a data architect. The data engineering team needs to configure a synchronization of data between an on-premises Microsoft SQL Server database to Azure SQL Database.

Ad-hoc and reporting queries are being overutilized the on-premises production instance. The synchronization process must:

- Perform an initial data synchronization to Azure SQL Database with minimal downtime
- Perform bi-directional data synchronization after initial synchronization

You need to implement this synchronization solution.

Which synchronization method should you use?

transactional replication

Data Migration Assistant (DMA)

backup and restore

SQL Server Agent job

Azure SQL Data Sync

Q14.   You are designing the folder structure for an Azure Data Lake Storage Gen2 container.

Users will query data by using a variety of services including Azure Databricks and Azure Synapse Analytics serverless SQL pools. The data will be secured by subject area. Most queries will include data from the current year or current month.

Which folder structure should you recommend to support fast queries and simplified folder security?

/{SubjectArea}/{DataSource}/{DD}/{MM}/{YYYY}/{FileData}_{YYYY}_{MM}_{DD}.cs

/{DD}/{MM}/{YYYY}/{SubjectArea}/{DataSource}/{FileData}_{YYYY}_{MM}_{DD}.cs

/{YYYY}/{MM}/{DD}/{SubjectArea}/{DataSource}/{FileData}_{YYYY}_{MM}_{DD}.cs

```
/{SubjectArea}/{DataSource}/{YYYY}/{MM}/{DD}/{FileData}_{YYYY}_{MM}_{DD}.cs
```

Q15.   You have an Azure Data Lake Storage Gen2 container that contains 100 TB of data.

You need to ensure that the data in the container is available for read workloads in a secondary region if an outage occurs in the primary region. The solution must minimize costs.

Which type of data redundancy should you use?

geo-redundant storage (GRS)

read-access geo-redundant storage (RA-GRS)

zone-redundant storage (ZRS)

locally-redundant storage (LRS)

Q16.   You use YARN to manage the resources for a Spark Thrift Server running on a Linux-based Apache Spark cluster in Azure HDInsight.

You discover that the cluster does not fully utilize the resources. You want to increase resource allocation.

You need to increase the number of executors and the allocation of memory to the Spark Thrift Server driver.

Which two parameters should you modify? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

spark.dynamicAllocation.maxExecutors

spark.cores.max

spark.executor.memory

spark_thrift_cmd_opts

spark.executor.instances

Q17.   You have an Azure Stack integrated system that has 100 tenants.
You create a new offer that is Private.
You need to provide the offer to a tenant.
What should you do?

Delegate rights to the tenant user, and then instruct the tenant user to create
a new subscription.

Create a new subscription, and then assign the subscription to the tenant.

Run the New-AzsOffer cmdlet, and then specify the tenant user account.

Run the Set-AzsUserSubscription cmdlet, and then specify the subscription of
the tenant user.

Q18.   You have an Azure Stack integrated system.
A tenant requires assistance managing a subscription.

The tenant needs to create a custom RBAC role definition.

What should you instruct the tenant to do?

Establish a PowerShell session to the Azure Resource Manager (user) endpoint. Run the New-AzureRmPolicySetDefinition cmdlet and the New-AzureRoleTemplate cmdlet.

Establish a PowerShell session to the Azure Resource Manager (user) endpoint. Create a JSON file that contains the permission definitions. Run the New-AzureRmRoleDefinition cmdlet.

Establish a PowerShell session to the Azure Resource Manager (administrator) endpoint. Create an XML file that contains the permission definitions. Run the New-AzureRmRoleDefinition cmdlet

Establish a PowerShell session to the Azure Resource Manager (administrator) endpoint. Run the New-AzureRmPolicySetDefinition cmdlet and the New-AzureRoletemplate cmdlet.

Q19.　　You have an Azure Stack integrated system that uses Microsoft Azure Active Directory (Azure AD) for authentication.
You download and extract the Azure App Service files.
You need to configure the system to support the deployment of Node.js applications and Azure functions.
What should you do before you configure the resource provider?

Install certificates. Register a service principal. Deploy and configure a file server. Create a relying party trust.

Install certificates. Deploy and configure a file server. Deploy and configure a Microsoft SQL server. Create an Azure AD application.

Install certificates. Register a service principal. Deploy and configure a Microsoft SQL server. Create a relying party trust.

Register a service principal. Deploy and configure a file server. Implement and configure a MySQL resource provider. Configure the Azure functions.

Q20.　　You have an Azure Stack integrated system. You deploy a platform as a service (PaaS) service that uses a file server.
You need to ensure that tenant users can use PowerShell Desired State Configuration (DSC) to manage the file server from the Azure Stack portal.

What should you do?

From the Azure Stack administrator portal, add the DSC extension.

From a privileged endpoint, run the Import-DscResource cmdlet.

From a privileged endpoint, run the Start-DscConfiguration cmdlet.

From the Azure Stack user portal, add the DSC extension.

Section 1 - Python

Q1

   Line 1: return product_price_list[0]

   Line 2: price = find_price(product_price_list[1:])

   Solution

   No Solution

Q2

   Name not found

   Done

   Program done

   Solution

   No Solution

Q3

   All the given function Calls

   Solution

   No Solution

Q4

   23

   Solution

   No Solution

Q5

   Number Exception - inside

   Inside finally

   Success

   Solution

   No Solution

Q6

   315

   Solution

   No Solution

Q7

   [51, 91, 77, 39]

Q8

self.__account_balance = acc_balance

Q9

Marks cannot be negative

4

Q10

display_cat_details("Brown","Nelly")

display_cat_details("Black","Roger","show")

Q11

Final block

Q12

Option C & D

Q13

3

Q14

Static=2, Local=3, Instance=3

Q15

dict1 - {1: 2, 2: 24, 3: 36, 4: 48, 5: 60}

dict2 - {2: 2, 3: 24, 4: 36, 5: 48, 6: 60}

Solution

No Solution

Q16   Number Exception - Inside

Inside finally

Success

Solution

No Solution

Q17   26

Solution

No Solution

Q18   No output

Solution

No Solution

Q19   6

Solution

No Solution

Q20   ValueError, NameError

Solution

No Solution

Section 2 - SQL

Q1   SELECT first_name, last_name

FROM employees

WHERE department = (SELECT department FROM employees

WHERE first_name = 'Jessica'

AND last_name = 'Butcher'

AND employee_id = 40)

AND salary > (SELECT salary FROM employees

WHERE first_name = 'Jessica'

AND last_name = 'Butcher'

AND employee_id = 40);

Solution

No Solution

Q2   NOT IN should be not be used if a NULL value is expected in the result set

Solution

No Solution

Q3   The column in the sub-query SELECT clause should only be one when there's an inequality

used in the main query

Solution

The columns selected in the sub-query should be same as on the other side of

comparison operator. Any inequality of data type or number of columns would result in an

error.

Q4   It throws an error

Solution

Multi-row operators cannot be used in single-row sub-queries and vice versa.

Q5   SELECT * FROM employees E WHERE 3 = (SELECT count(distinct salary) FROM employees S WHERE S.salary > E.salary);

Solution

No Solution

Q6   (i) will succeed but (ii) will fail

Solution

First query will successfully execute and after execute this statement, attribute dept_id of student table with Roll_No = 1 becomes 'NULL'

which is foreign key from Department's Dept_id. Now, second statement will not execute because primary key cannot be NULL.

Q7   SELECT workerid, iobtitle, mobilenumber FROM worker WHERE workerid ="W06";

SELECT workerid, jobtitle, mobilenumber FROM worker WHERE workerid ="W07";

Solution

No Solution

Q8

Solution

No Solution

Q9

CUSTOMERID BOOKID

101 B978

103 B978

106 B978

105 B234

Solution

No Solution

Q10

SELECT E.empid EMPID, E.empname EMPNAME, M.empid MGRID, M.empname MGRNAME FROM employee E INNER JOIN employee M ON

E.manager = M.empid;

Solution

No Solution

Q11

PROJECTDEPT AVGSAL TOTALBONUS

Telecom 21667 600

ENG 40000 1300

Solution

No Solution

Q12

Only Jack

Solution

No Solution

Q13

Eliza and Robert

Solution

No Solution

Q14

Two items will have discount of 10

Solution

No Solution

Q15   INSERT INTO student VALUES(10001, Rony,'M','12-IAN-95', 'rony@gmail.com');

Solution

No Solution

Q16   SELECT accountnum FROm account WHERE status = 'Active' AND balance < (SELECT AVG(balance) FROM account)

Solution

No Solution

Q17   MONTH

Jan

Solution

No Solution

Q18   item is in 2NF and customer is in 2NF

Solution

No Solution

Q19   Jack

Solution

No Solution

Q20   1

Solution

No Solution

Section 3 - AWS

Q1   A map of regions with a heatmap overlay to show the volume of customer contacts

Solution

No Solution

Q2

Use the s3n URI to store the data to be processed as objects in Amazon S3.

Solution

No Solution

Q3

Use Amazon CloudFront to serve application content locally in the country; Amazon CloudFront logs will satisfy the requirement.

Solution

No Solution

Q4

Configure Lambda to read from fewer shards in parallel.

Solution

No Solution

Q5

Use AWS KMS for encryption of data. Configure and attach multiple roles with different permissions based on the different user needs.

Solution

No Solution

Q6

Mask the credit card numbers to only show the last four digits of the credit card number.

Solution

No Solution

Q7

Use bzip2 or Snappy rather than gzip for the archives.

Solution

No Solution

Q8

Store the data on an EMR File System (EMRFS) instead of HDFS and enable EMRFS consistent view. Run two separate EMR clusters in two

different Availability Zones. Point both clusters to the same HBase root directory in the same Amazon S3 bucket.

Solution

No Solution

Q9

A KEY distribution style for both tables

Solution

No Solution

Q10

Increase the provisioned write capacity units assigned to the streams Amazon DynamoDB table.

Solution

No Solution

Q11   Add more memory and CPU capacity to the streaming application.

Solution

No Solution

Q12   Install Kafka Connect on each application instance and configure each Kafka Connect instance to write to a specific topic only.

Solution

No Solution

Q13   Use Amazon Managed Streaming for Apache Kafka to ingest the data to save it to Amazon Redshift. Enable Amazon Redshift workload

management (WLM) to prioritize workloads.

Solution

No Solution

Q14   Install Apache Ranger on an Amazon EC2 instance and integrate with Amazon EMR. Using Ranger policies, create role-based access control

for the existing data assets in Amazon S3.

Solution

No Solution

Q15   Use Amazon Redshift federated queries to join the data sources. Use Amazon QuickSight to generate the mobile dashboards.

Solution

No Solution

Q16   Create an external schema based on the AWS Glue Data Catalog on the existing Amazon Redshift cluster to query new data in Amazon S3

with Amazon Redshift Spectrum.

Solution

No Solution

Q17   Use Amazon QuickSight with Amazon Athena as the data source. Use pivot tables as the visual type.

Solution

No Solution

Q18   Use the AWS CLI to back up the DynamoDB table. Then use the restore-table-from-backup command and modify the partition key.

Solution

No Solution

Q19   After the DB instances are manually rebooted

Solution

No Solution

Q20   Use Athena to extract the data and store it in Apache Parquet format on a daily basis. Query the extracted data.

Run a daily AWS Glue ETL job to convert the data files to Apache Parquet and to partition the converted files. Create a periodic AWS Glue

crawler to automatically crawl the partitioned data on a daily basis.

Solution

No Solution

Section 4 - Azure

Q1   5

Solution

No Solution

Q2   ReadUncommitted

Solution

No Solution

Q3   Run the following Windows PowerShell cmdlet:Get-AzurePublishSettingsFile

Navigate to the following URI: https://windows.azure.com/download/publishprofile.aspx

Solution

No Solution

Q4   Enable verbose monitoring.

Update the cloud service definition file and the service configuration file.

Create a storage account.

Solution

No Solution

Q5   Configure Microsoft Visual Team Services to continuously deploy the user interface tier to the Azure Web App service. Deploy the

production builds and the staging builds of the user interface tier to separate slots.

Solution

No Solution

Q6   Upload the certificate referenced in the application package.

Solution

No Solution

Q7   parallel execution

Solution

No Solution

Q8   Ambari RESR API

Solution

No Solution

Q9   Build the environment in Azure Databricks and use Azure Data Factory for orchestration.

Solution

No Solution

Q10   Azure Container Service

Solution

No Solution

Q11   Remove entire row

Solution

No Solution

Q12   Apache Spark

Solution

No Solution

Q13   Azure SQL Data Sync

Solution

No Solution

Q14

/{SubjectArea}/{DataSource}/{YYYY}/{MM}/{DD}/{FileData}_{YYYY}_{MM}_{DD}.csv

Solution

No Solution

Q15

read-access geo-redundant storage (RA-GRS)

Solution

No Solution

Q16

spark.dynamicAllocation.maxExecutors

spark.executor.memory

Solution

No Solution

Q17

Create a new subscription, and then assign the subscription to the tenant.

Solution

No Solution

Q18

Establish a PowerShell session to the Azure Resource Manager (user) endpoint. Create a JSON file that contains the permission definitions.

Run the New-AzureRmRoleDefinition cmdlet.

Solution

No Solution

Q19

Install certificates. Deploy and configure a file server. Deploy and configure a Microsoft SQL server. Create an Azure AD application.

Solution

No Solution

Q20

From the Azure Stack user portal, add the DSC extension.

Solution

No Solution