BDTM: Big Data Tools for Managers
2nd Internal Question Paper
Version-A

Q1. Create below table with data and demonstrate JOIN operation in SQL.
Table1 : student

usn	name
1SI22MBA01	John Nick
1SI22MBA02	Bob Smith

Table 2: student_marks

ID	usn	subject	marks
1	1SI22MBA01	sub-1	45
2	1SI22MBA01	sub-2	38
3	1SI22MBA01	sub-3	25
4	1SI22MBA01	sub-4	49

Write MySQL Queries for following:

1. Create student & student_marks tables with data. [10]

```
CREATE TABLE student (
usn text,
name text
);
INSERT INTO student VALUES ('1SI22MBA01','John Nick');
INSERT INTO student VALUES ('1SI22MBA02','Bob Smith');

CREATE TABLE student_marks (
ID TEXT,
usn text,
subject text,
marks text
);

INSERT INTO student_marks VALUES('1','1SI22MBA01','sub-1','45');
INSERT INTO student_marks VALUES('2','1SI22MBA01','sub-2','38');
INSERT INTO student_marks VALUES('3','1SI22MBA01','sub-3','25');
INSERT INTO student_marks VALUES('4','1SI22MBA01','sub-4','49');
```

2. Perform INNER Join with two tables. [5]

```
SELECT * FROM student
INNER JOIN student_marks
ON student.usn = student_marks.usn;
```

Q2. Write simple IF conditional statement to variable contains positive value. [5]
For Example Variable: var1 = 10

```
var1 = 10
if var1>0:
   print("Number is Positive")
```

Q3. Demonstrate List Data structures in Python

1. Create List with elements 10,20,30,40,50,60,70,80,90 [4]

```
num_list = [10,20,30,40,50,60,70,80,90]
```

2. Display all the list elements [2]

```
print(num_list)
```

3. Display the number of elements present in List using len() function [1]

```
print(len(num_list))
#or
len(num_list)
```

4. Add 100, 101 elements in to the existing list. [3]

```
new_ele = [100,101]
num_list + new_ele

or

num_list + [100,101]
```

Q4. Demonstrate Tuple Data structures in Python

1. Create a Tuple with elements 100,110,120,130,140,150,160,170,180,190, 200 [1]

   ```
   num_tuple = (100,110,120,130,140,150,160,170,180,190, 200)
   ```

2. Display First element of Tuple [1]

   ```
   num_tuple[0]
   ```

3. Display last element of Tuple [1]

   ```
   num_tuple[-1]
   ```

4. Display first 3 elements of tuple [1]

   ```
   num_tuple[0:3]
   ```

5. Display last 3 elements of tuple [1]

   ```
   num_tuple[-3:]
   ```

Q5. Demonstrate Pandas package to perform data analysis for IPL dataset.

```
import pandas as pd
data = pd.read_csv("https://bit.ly/3V0H3Ox")
```

1. Display Shape of panda DataFrame [1]

```
data.shape
```

2. Display all the columns names with its data types [2]

```
data.info()
```

3. Display quick summary of dataset [2]

```
data.describe()
```

4. Display top 10 records [2]

```
data.head(10)
```

5. Display last 10 records [2]

```
data.tail(10)
```

6. Display all the values of COUNTRY column [2]

```
data['COUNTRY']
```

7. Count frequency of COUNTRY values [2]

```
data['COUNTRY'].value_counts()
```

8. Display unique COUNTRY values [2]

```
data['COUNTRY'].unique()
```