

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

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

Table1 : customers

customer_id	name
1001	Nick
1002	Bob

Table 2: customers_orders

ID	customer_id	item	qty
1	1001	Item-1	10
2	1001	item-2	20
3	1001	item-3	30
4	1001	item-4	40

Write MySQL Queries for following:

[15]

1. Create customers & customers tables with data.

```
CREATE TABLE customers (  
customer_id TEXT,  
name TEXT  
);
```

```
INSERT INTO customers VALUES ('1001', 'Nick');  
INSERT INTO customers VALUES ('1002', 'Bob');
```

```
CREATE TABLE customers_orders (  
ID text,  
customer_id TEXT,  
item TEXT,  
qty text  
);  
INSERT INTO customers_orders VALUES ('1', '1001', 'Item-1', '10');  
INSERT INTO customers_orders VALUES ('2', '1001', 'Item-2', '20');  
INSERT INTO customers_orders VALUES ('3', '1001', 'Item-3', '30');  
INSERT INTO customers_orders VALUES ('4', '1001', 'Item-4', '40');
```

2. Perform INNER Join with two tables.

```
SELECT * FROM customers  
INNER JOIN customers_orders  
ON customers.customer_id = customers_orders.customer_id
```

Q2. Write simple IF conditional statement to variable contains positive value. [5]
For Example Variable: var1 = 99

```
var1 = 99
if var1>0:
    print("Number is Positive")
```

Q3. Demonstrate List Data structures in Python [10]

1. Create List with elements 0,1,2,3,4,5,6,7,8,9,10
`num_list = [0,1,2,3,4,5,6,7,8,9,10]`
2. Display all the list elements
`print(num_list)` or `num_list`
3. Display the number of elements present in List using len() function
`len(num_list)`
4. Add 20, 30 elements in to the existing list.
`num_list + [20,30]`

Or

```
new_ele = [20, 30]
num_list + new_ele
```

Q4. Demonstrate Tuple Data structures in Python

[5]

1. Create a Tuple with elements

10, 20, 30, 40, 50, 60, 70, 80, 90, 100

```
num_tuple = (10, 20, 30, 40, 50, 60, 70, 80, 90, 100)
```

2. Display First element of Tuple

```
num_tuple[0]
```

3. Display last element of Tuple

```
num_tuple[-1]
```

4. Display first 3 elements of tuple

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

5. Display last 3 elements of tuple

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

Q5. Demonstrate Pandas package to perform data analysis for IPL dataset. [15]

Step 1: Write below code to read data from Internet and IPL dataset will be loaded into data variable, use data variable name to perform below operations.

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

1. Display Shape of panda DataFrame
`data.shape`
2. Display all the columns names with its data types
`data.info()` or `data.dtypes`
3. Display quick summary of dataset
`data.describe()`
4. Display top 10 records
`data.head(10)`
5. Display last 10 records
`data.tail(10)`
6. Display all the values of TEAM column
`data['TEAM']`
7. Count frequency of TEAM values
`data['TEAM'].value_counts()`
8. Display unique TEAM values
`data['TEAM'].unique()`