Python for DS

MCQs (1 mark)

Q1. In Spyder IDE, which of the following helps you clear all variables from the current workspace?

a) `Clear Console`

b) `Clear Variables`

c) `Clear All Files`

d) `Restart Kernel`

Answer: b) `Clear Variables`

Q2. Given the code:

```
nums = [1, 2, 3, 4, 5]
result = nums[1:4]
```

What does `result` contain?

a) `[2, 3, 4, 5]`

b) `[1, 2, 3]`

c) `[2, 3, 4]`

d) Error

Answer: c) `[2, 3, 4]`

Descriptive Questions

Q3.(5 MARK)

Explain the difference between a Python list and a NumPy `ndarray` with examples.

Answer:

A **Python list** is a built-in data structure that can hold elements of different data types (integers, floats, strings, etc.). It is flexible but not optimized for numerical computations. Operations on lists usually require loops, which are slower.

A **NumPy** `ndarray` is a specialized data structure for numerical computing. It stores data in a contiguous block of memory and supports **vectorized operations** (performing operations on the entire array without explicit loops), which makes it much faster for mathematical tasks.

Key differences:

1. **Data type:** Lists can hold mixed data types, while `ndarray` elements are usually of the same type.
2. **Performance:** NumPy arrays are optimized for speed and memory efficiency.
3. **Operations:** Mathematical operations on lists are element-wise only through loops, while arrays support direct element-wise vectorized operations.

Example:

```
import numpy as np

# Python list
lst = [1, 2, 3]
print(lst * 2)     # Output: [1, 2, 3, 1, 2, 3] (concatenation)

# NumPy array
arr = np.array([1, 2, 3])
print(arr * 2)     # Output: [2 4 6] (element-wise multiplication)
```

Q4. (7 Marks)

Q4. (7 Marks)

A Python script `sales.py` is written in an IDE (for example, VS Code, Jupyter Notebook, or Spyder) to load sales data stored in a file named `data.csv`. Describe how you would:

1. Set or check the working directory so the script can access `data.csv`.
2. Read the file into Python using an appropriate method.
3. Convert its "Quantity" column (initially stored as strings) to integers.
4. Print the total quantity sold.

Answer:

1. **Set Working Directory:**
   - In most IDEs, the default working directory is the folder where the script is saved.
   - To explicitly set it, use:

```
import os
os.chdir("C:/Users/YourName/Documents")  # adjust path accordingly
print(os.getcwd())  # to confirm
```

2. **Read the CSV File:**

```
import pandas as pd
df = pd.read_csv("data.csv")
```

3. **Convert the "Quantity" Column to Integers:**

```
df['Quantity'] = df['Quantity'].astype(int)
```

4. **Calculate and Print Total Quantity Sold:**

```
total_qty = df['Quantity'].sum()
print("Total Quantity Sold:", total_qty)
```

Explanation:

- **Step 1** ensures the IDE is pointing to the correct folder.
- **Step 2** loads the dataset into a Pandas DataFrame for easier analysis.
- **Step 3** fixes data type issues, ensuring arithmetic operations are possible.
- **Step 4** computes and prints the required business metric.