

Introduction to Data Science (DS2001)

Course Instructor(s):

Ms. Saira Qamar

Section(s): (DS-A,B,C)

Sessional-I Exam

Total Time (Hrs): 1

Total Marks: 50

Total Questions: 3

Date: Sep 23, 2024

Roll No

Course Section

Student Signature

Do not write below this line.

Attempt all the questions.

[CLO 3: Comprehend the fundamental constructs of Python programming language]

Q1: Write the outputs of the following python codes in the respective output column. There are no syntax mistakes in the code. You can mention if there is any logical mistakes. **[10 marks]**

| | Code | Output |
|---|--|--------|
| 1 | <pre>s1 = "Hello" s2 = "World" result = (s1 + s2) * 2 print(result)</pre> | |
| 2 | <pre>l1 = [1, 2] l2 = [3, 4] result = l1 + l2 * 2 print(result)</pre> | |
| 3 | <pre>name = "Ali" age = 25 print(f"My name is {name} and I am {age+1} years old.")</pre> | |
| 4 | <pre>t = (10, 20, 30) t[1] = 40 print(t)</pre> | |

National University of Computer and Emerging Sciences
Islamabad Campus

| | | |
|----|---|--|
| 5 | <pre>matrix = [[1, 2, 3], [4, 5, 6], [7, 8, 9]] flat = [x for row in matrix for x in row if x % 2 != 0] print(flat)</pre> | |
| 6 | <pre>s = "Artificial Intelligence" result = s[-5:] print(result)</pre> | |
| 7 | <pre>t = (x * 2 for x in range(3)) print(tuple(t))</pre> | |
| 8 | <pre>result = list(map(lambda x: x * 2, [1, 2, 3, 4])) print(result)</pre> | |
| 9 | <pre>def fib(n): if n <= 1: return n else: return fib(n-1) + fib(n-2) print(fib(5))</pre> | |
| 10 | <pre>def func(x, y=5, z=10): return x + y * z print(func(2)) print(func(2, 3)) print(func(2, 3, 4))</pre> | |

[CLO 3: Comprehend the fundamental constructs of Python programming language]

Q2: Write a one-line python code in the given column to perform the following operations:

[20 marks]

| | Statement | One line code |
|---|--|----------------------|
| 1 | To create a list of squares from a list of numbers [1, 2, 3, 4]. | |
| 2 | To filter out odd numbers from the list [1, 2, 3, 4, 5, 6] and store the result in even_numbers. | |

National University of Computer and Emerging Sciences

Islamabad Campus

| | | |
|----|---|--|
| 3 | To convert a list of tuples [(1, 'a'), (2, 'b'), (3, 'c')] into a dictionary. | |
| 4 | To create a 4x4 identity matrix using NumPy. | |
| 5 | To create a NumPy array of 10 equally spaced numbers between 1 and 100. | |
| 6 | To create a line plot of x = [1, 2, 3, 4, 5] and y = [10, 20, 30, 40, 50] using matplotlib. | |
| 7 | To reshape a NumPy array arr = np.arange(12) into a 2D array with 3 rows and 4 columns. | |
| 8 | To drop the column 'A' from the pandas DataFrame df. | |
| 9 | To create a pandas DataFrame with columns ['A', 'B'] and data [[1, 2], [3, 4], [5, 6]]. | |
| 10 | To calculate the sum of each row in a pandas DataFrame df. | |

[CLO 3: Comprehend the fundamental constructs of Python programming language]

Q3: You are working as a data analyst for a sportswear company that tracks the sales of its products across three regions: North, East, and West. The company has provided sales data for the first three months of the year (January, February, and March) and wants you to analyze it to make strategic decisions for the next quarter. **[20 marks]**

Here is the sales data, where each product's sales are represented in three different regions:

```
sales_data = {
    'Shoes': {'North': (500, 520, 530), 'East': (300, 320, 310), 'West': (400, 450, 420)},
    'Shirts': {'North': (150, 170, 160), 'East': (200, 210, 190), 'West': (180, 175, 190)},
    'Jackets': {'North': (300, 320, 310), 'East': (250, 240, 230), 'West': (220, 210, 200)},
    'Caps': {'North': (100, 120, 110), 'East': (80, 90, 85), 'West': (75, 80, 78)}
}
```

Each value in the inner dictionary is a tuple representing sales for January, February, and March in the respective region.

National University of Computer and Emerging Sciences

Islamabad Campus

Tasks:

Data Organization: [3 marks]

- Convert the sales_data dictionary into a pandas DataFrame where each row represents a product, and columns represent monthly sales in the different regions ('North_Jan', 'North_Feb', 'North_Mar', 'East_Jan', etc.).

Total Sales Calculation: [3 marks]

- For each product, calculate the total sales for each region (North, East, West) over the three months and add them as new columns ('North_Total', 'East_Total', 'West_Total').

Average Sales with NumPy: [5 marks]

- Convert the sales data of each product in each region into a NumPy array and calculate the average sales of each product across all regions. Add a new column 'Average_Sales' to store the results. **(don't use the built-in average function)**

Visualizing Data: [4 marks]

- Create three line plots using matplotlib to visualize the sales trend for each product across the three months (January, February, and March). One line should represent each region (North, East, West). Use the same plot for each product and ensure proper labeling.
- The x-axis should represent the months (January, February, March).
- The y-axis should represent the sales figures.
- Each region's sales trend should be shown with a separate line.

Highest Sales in March: [2 marks]

- Identify the product with the highest sales in March across all regions.

Insights: [3 marks]

- Provide insights on the following:
 - Which product has the highest average sales across all regions?
 - Which region had the lowest overall total sales?