Indian Statistical Institute
BSDS Semester-I 2026-2027
Subject: Introduction to Computation
Class Test 2 (23 November, 2026)
Maximum marks: 20 Duration: 1.5 hrs.

For the following code snippets, assume that the necessary packages have been imported in the usual way, and initialised as required.

1. (12 marks)

   (a) Let `A` be a NumPy ndarray. Given the attributes `A.itemsize` and `A.shape`, write Python **statements** to compute the attributes `A.ndim`, `A.size` and `A.nbytes` in terms of the given attributes and/or attributes calculated earlier. [1 + 2 + 1 = 4]

   ```
   A.ndim
   ```

   ```
   A.size
   ```

   ```
   A.nbytes
   ```

   (b) Consider the following code.

   ```
   A = np.linspace(5, 19, 8).reshape(2, 4)
   B = np.linspace(20, 41, 8).reshape(4, 2)
   C = np.concat([A[0:2,1:3], B[1:3,0:2]], axis=1)
   D = A.sum(axis=0)
   E = C + D
   ```

   Write the elements of the five ndarrays `A`–`E` in the form of matrices. Show your calculations. [1.5 + 1 + 2.5 + 1 + 2 = 8]

2. (10 marks) Consider the following code.

```
A = np.array([ i for i in range(71,95) ]).reshape(4, 6)
pdA = pd.DataFrame(A, index=["Gupi", "Bagha", "Haru", "Bhombol"], columns=list("123456"))
```

Assume that the strings comprising the index of `pdA` are **full names** of students. Thus, each row represents a student. Also assume that each column represents a course, and each cell represents the corresponding student's marks in the corresponding course.

Keep in mind that in real life, the number of students will generally be much larger than 4. The code that you write for this question should not be specific to the "4 students, 6 courses" scenario given in the example.

(a) The teacher for **course 1** forgot to include marks for the homework component of the course in the total, and submits these marks separately. These marks are stored in a Python list `hw` (with 4 elements for the above example). Write code to update `pdA` with the correct scores for **course 1**; [1.5]

(b) Write a Python statement to calculate the mean score for each course. [1]

(c) Write Python statements to calculate each student's aggregate and overall percentage. [1.5]

(d) Write code to construct a list containing the names of all students in **alphabetically sorted** order. [1]

(e) Use list comprehension to generate a list of roll numbers from the above list. Roll numbers have to be of the form `BSDS26XX`, where `XX` is a 2 digit number (assume that the number of students is less than 100). These should be assigned to the students in alphabetic order starting with `BSDS2601`, `BSDS2602`, …. [2]

(f) Construct another DataFrame `pdB` with a single column named `"RollNo"` and an index given by the list constructed in part (d). The entries in the single column should be the roll numbers generated in part (e). [1.5]

(g) Show how you would correctly combine the information in DataFrames `pdA` and `pdB`. [1.5]