COURSE CODE (CREDITS): 20B1WCI531   MAX. MARKS: 15

COURSE NAME: Foundation for Data Science and Visualization

COURSE INSTRUCTORS: Dr. Aman Sharma, Dr. Shubham Goel   MAX. TIME: 1 Hour

Note: (a) All questions are compulsory. (b) Marks are indicated against each question in square brackets. (c) The candidate is allowed to make Suitable numeric assumptions wherever required for solving problems.

Q1. What is the output of following R codes?   [4*4 Marks, CO-1, CO-2]

a)	b)
| ```r
x <- c(2, 4, 6)
y <- x[1] + x[length(x)]
z <- sum(x) - y
cat("x:", x, "y:", y, "z:", z)
``` | ```r
a <- 5
b <- 8
c <- 3

if (a > b) {
  if (b < c) {
    print("Inner")
  } else {
    print("Middle")
  }
} else {
  print("Outer")
}
``` |
| c) | d) |
| ```r
mat <- matrix(1:9, nrow = 3)
total_sum <- 0

for (i in seq_len(nrow(mat))) {
  for (j in seq_len(ncol(mat))) {
    total_sum <- total_sum + mat[i, j]
  }
}

cat("Total Sum:", total_sum)
``` | ```r
vector_a <- c(1, 2, 3)
vector_b <- c(2, 3, 4)

result <- sum((vector_a * vector_b) /
sqrt(vector_a + vector_b))
cat("Result:", result)
``` |

Q2. Explain the fundamental differences between computer science, data science, and real science. Highlight how the interdisciplinary nature of data science bridges the gaps between these fields, providing an example of how each contributes to the holistic understanding of complex real-world phenomena. **[2 Marks, CO-1]**

Q3. Write python code for the below two cases: [2 *2 Marks, CO-2]

a) Given a 2D NumPy array **matrix**, extract the elements from the bottom-right quarter of the array. For example, if **matrix** is:

matrix = np.array([[1, 2, 3, 4], [5, 6, 7, 8], [9, 10, 11, 12], [13, 14, 15, 16]])

b) Given two 1D NumPy arrays **a** and **b**, determine the positions where elements of **a** are greater than the corresponding elements of **b**. Reshape both arrays into a 2D format to facilitate element-wise comparison. For example:

a = np.array([4, 5, 2, 8])
b = np.array([3, 7, 1, 6])

Q4. Given a pandas DataFrame **df** containing information about students, extract the names of students who have scored above the mean in both Math and Science. Extract the names of students who scored above the mean in both Math and Science, i.e., **[John, Alex]**. For example, if **df** is:

Name Math Science   [3 Marks, CO-2]

0 John 85   92

1 Mary 78   88

2 Alex 92   95

3 Sara 65   75

Q5. Given a pandas DataFrame **data** containing temperature readings for different days, calculate the average temperature for each city. If a temperature reading is missing (NaN), replace it with the average temperature of that city. Display the cities in ascending order of their average temperatures. Calculate and display the cities with their average temperatures, i.e., **[A: 21.33, B: 28.33]**. For example:

data: [2 Marks, CO-1, CO-2]

City Temperature

0 A   20.0

1 B   25.0

2 A   NaN

3 B   30.0

4 A   22.0