NATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY

FACULTY OF APPLIED SCIENCES

DEPARTMENT OF INFORMATICS AND ANALYTICS

PROGRAMMING FOR DATA SCIENCE

SIDS 5102

Main Examination Paper

First Semester 2024

This examination paper consists of 4 pages

Time allowed: 3 hours

Total Marks: 100

Special Requirements: None

Examiner's Name: Mrs B Ndlovu

External Examiner's Name: Professor L. Sakala

INSTRUCTIONS TO THE CANDIDATES

1. Answer any four (4) questions
2. Each question carries 25 marks
3. All programming questions to be answered in Java/Python/R

MARK ALLOCATION

QUESTION	MARKS
1.	25
2.	25
3.	25
4.	25
5.	25

QUESTION ONE

Explain the following terms:

   i.   Parallel Processing   **[5]**

  ii.   Client Server Architecture   **[5]**

 iii.   Yet Another Resource Manager (YARN)   **[5]**

 iv.   Scaling   **[5]**

  v.   Unstructured data   **[5]**

QUESTION TWO

(a) Discuss Flynn's taxonomy of computer architectures and how it classifies different types of computer processors based on their instruction stream and data stream characteristics.

(b)   Using a programming language of your choice implement Flynn's Taxonomy **[25]**

QUESTION THREE

List and explain 5 (five) use cases of big data in Hadoop   **[25]**

QUESTION FOUR

Write a program that presents a simulated annealing solution for a maximisation problem. The objective function should be to:

(a) find the maximum number of 1s in a specified array. **[5]**

(b) distribute the problem to 4 cores. **[10]**

(c) indicate which core was the quickest to arrive at a solution and which was the longest **[10]**

QUESTION FIVE

With the aid of a clearly labelled diagram, describe the Multi-tier Architecture

[25]

THE END