BCS THE CHARTERED INSTITUTE FOR IT

BCS HIGHER EDUCATION QUALIFICATIONS
BCS Level 5 Diploma in IT

BIG DATA MANAGEMENT

Wednesday 4th October 2023 - Morning

Answer any FOUR questions out of SIX. All questions carry equal marks.

Time: TWO hours

Answer any <u>Section A</u> questions you attempt in <u>Answer Book A</u>
Answer any <u>Section B</u> questions you attempt in <u>Answer Book B</u>

The marks given in brackets are **indicative** of the weight given to each part of the question.

Calculators are **NOT** allowed in this examination.

Section A
Answer any Section A questions in Answer Book A

A1.

a) Outline some of the main challenges that an organisation should consider when planning to embark on a new Big Data implementation.

(5 marks)

b) Describe **TWO** services that are included in the provision of Big Data management.

(6 marks)

c) Give **FIVE** examples of personal data as defined by the General Data Protection Regulation (GDPR).

(5 marks)

d) Describe unstructured data and why it is a problem in managing personal data for data protection compliance.

(5 marks)

e) What are considered the main objectives of governments when introducing data privacy regulations?

(4 marks)

A2.

a) Compare and contrast the main features of cloud storage methods and onsite storage methods. Evaluate the advantages and disadvantages of **each**.

(10 marks)

b) What are the main advantages and disadvantages for an organisation in adopting Cloud storage?

(5 marks)

c) Describe the attractions and effect of using distributed file systems in the operation of IT systems in organisations.

(5 marks)

d) Discuss the main risks to an organisation in the adoption of cloud-based technology.

(5 marks)

A3.

a) Describe Gartner's Ascendancy model.

(7 marks)

b) Describe basic machine learning algorithms and provide **THREE** examples.

(7 marks)

c) Explain what is meant by Artificial Intelligence (AI) and provide **THREE** examples of its use to analyse data.

(7 marks)

d) Describe the main characteristics of Neural Networks and give **THREE** examples of their use .

(4 marks)

[Turn Over]

Section B
Answer any Section B questions in Answer Book B

B4.

a)

    i.   Describe and give **TWO** examples of machine generated big data.

(3 marks)

    ii.   Describe and give **TWO** examples of human generated big data.

(3 marks)

b) Explain the following three types of data formats commonly found in big data processing tasks:

    i.   Structured data;

(3 marks)

    ii.   Unstructured data;

(3 marks)

    iii.   Semi-structured data.

(3 marks)

c) Describe **FOUR** characteristics of streamed data that distinguish it from other more traditional types of data.

(4 marks)

d) Compare the characteristics of Event Stream Processing with Complex Event Stream Processing.

(6 marks)

B5.

a) Explain what the Apache Kafka platform is and describe its principal features.

(9 marks)

b) Explain, and give an example of, why you might want to use Kafka in a big data application.

(7 marks)

c) Give **THREE** examples of how the Hadoop Hierarchical Data File Store (HDFS) is particularly suited to big data applications.

(9 marks)

B6.

a) Explain the basic features of a Not-only SQL (NoSQL) database.

(4 marks)

b) Describe the following three types of databases.

    i.   Document database;

(3 marks)

    ii.   Column-oriented database;

(3 marks)

    iii.   Graph database.

(3 marks)

c) Explain why the MongoDB document-oriented database can maintain consistency and partition tolerance but not availability.

(10 marks)

d) Describe how the R language and statistical environment is useful for the analysis of big data.

(2 marks)

END OF EXAMINATION