

BCS THE CHARTERED INSTITUTE FOR IT

BCS HIGHER EDUCATION QUALIFICATIONS
BCS Level 5 Diploma in IT

BIG DATA MANAGEMENT

Thursday 16th April 2026 – Afternoon

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

Time: TWO hours

Answer any Section A questions you attempt in Answer Book A
Answer any Section B questions you attempt in Answer Book B

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 Section A questions in Answer Book A

A1.

- a) Describe **five** of the main challenges for organisations in implementing a big data strategy and explain the ways in which these challenges could be addressed.
(10 marks)
- b)
- i. Describe what is meant by the term data quality.
(4 marks)
 - ii. Explain why it is important for organisations to achieve a high level of data quality.
(4 marks)
- c) Define personal data as described in the General Data Protection Regulation (GDPR) and provide **four** examples of personal data.
(7 marks)

A2.

- a) Explain **five** advantages of in-house and **five** advantages of outsourced solutions for big data.
(10 marks)
- b) Explain the main risks in implementing an in-house big data solution and describe the steps that can be taken as mitigation.
(10 marks)
- c) Describe the main financial cost components in the implementation of a cloud-based solution for big data.
(5 marks)

A3.

- a) Describe each of the three main types of data: structured, unstructured and streamed. For **each** of the three types, provide **one** example of its use.
(15 marks)
- b) Explain what is meant by the term Kafka data modelling and describe what the challenges are in using it.
(10 marks)

Section B
Answer Section B questions in Answer Book B

B4.

- a) Describe noSQL, including its key features.
(8 marks)
- b) Describe the advantages of noSQL over Relational Database Management Systems (RDBMS).
(8 marks)
- c) Assess the use of BASE versus ACID approaches when dealing with transaction processing and concurrency control, comparing the two.
(9 marks)

B5.

- a) Describe cloud storage methods and onsite storage methods.
(5 marks)
- b) What are the main advantages and disadvantages for an organisation in choosing to implement cloud storage? Include **five** examples of **each**.
(10 marks)
- c) What are the **five** primary impacts of the Hadoop Distributed File System (HDFS) on big data processing?
(5 marks)
- d) What are **five** of the main risks of cloud-based technology?
(5 marks)

B6.

- a) Name the **three** main types of machine learning algorithms and briefly describe how **each** is used.
(6 marks)
- b) Explain what is meant by logistic regression and provide **two** examples of how this could be used in practice.
(6 marks)
- c) Describe how AI is used with big data.
(5 marks)
- d) Provide **four** examples of how AI is used with big data.
(8 marks)

END OF EXAMINATION