

BCS Higher Education Qualification

Diploma

October 2022

EXAMINERS' REPORT TEMPLATE

Big Data Management

Questions Report:

A1	Syllabus area: 2
Phil Barker	<p>a) This question on costs of big data was well answered by most who correctly mentioned amongst other factors maintenance and staffing as main factors.</p> <p>b) This question asked candidates to consider the main issues for organisations with respect to characteristics of data. Most answers correctly identified data gathering being an issue in relation to the different data collection standards amongst others and data storage being an issue in particular for costs, security and infrastructure. Many candidates showed some uncertainty in the main concerns for data quality and data resilience. Some answers concerning data quality were the same as for data storage, rather than an emphasis on the lack of business processes and procedures in addition to factors such as poor staff training and the difficulty that might stem from dealing with a wide variety of data formats. The sub question on data resilience to growth was not generally well answered. Answers that acknowledged capacity planning, disaster recovery procedures and adequate record retention and disposal policy and similar would have gained fuller marks.</p> <p>c) This question on challenges of managing personal data was correctly answered with reference to adherence to data protection standards and maintaining standards of safety from acquisition of personal data from third parties.</p> <p>d) The importance of external compliance regulations was well recognised by many candidates with a few referring to the role of external compliance providing benchmarks for an organisation to work to and the benefit to ensuring a standard to judge third party suppliers by. In general answers were less complete in this question compared to c) above.</p>
A2	Syllabus area: 5.1 & 5.3
Phil Barker	<p>a) Many candidates correctly identified differences between cloud and onsite storage systems and gave an illustration of both</p> <p>b) This question proved difficult for some candidates. Many considered a range of advantages and disadvantages for cloud storage and gave good answers. Other answers gave comparison between cloud storage and onsite storage and gained fewer marks. In some cases, the answers centred around cloud storage as a service and cloud storage as a product. Making this distinction proved difficult to explore advantages/disadvantages</p> <p>c) This question on the impact of Hadoop on big data processing was best approached from emphasis on the use of cheap commodity servers able to offer affordable distributed resources and scalability opportunities. Many candidates gave answers based on that emphasis. Some candidates did not directly address the question and gave descriptions of the underlying mechanisms of Hadoop technology in isolation from the question context.</p> <p>d) Many candidates answered this question correctly by reference to the costs and security issues around using third party cloud-based providers. The possibility of costs rising with increased storage requirements and the legal implications when having sensitive data in a third party repository were well explored in good answers.</p>

A3	Syllabus area: 6
Phil Barker	<p>This question was answered by only a small minority of candidates. The question had five sub sections. The first two subsections a) and b) dealt with the Gartner data maturity model and elements of data analysis. The Gartner stages of maturity was well referenced by some candidates and gained full marks. The main elements of data analysis such as text, inferential, predictive and other types was in general not answered well by candidates. Many omitted this part of the question. Of the remaining three sub sections:</p> <p>c) Asked for some examples of AI in analysis of data. Answers were sought that referred to business intelligence using machine intelligence, or pattern detection in data on consumer behaviour and similar applications. Some candidates gave good accounts based on these example types . In many cases the question was avoided.</p> <p>d) and e) These sub sections asked for examples of neural networks and description of Bloom filters. There were no real attempts made to answer these questions.</p>
B4	Syllabus area: 1.1 & 7.3
Thomas Hayes	This question was well answered in the main, with the exception of B4(d) where most candidates scored lower than the pass mark.
B5	Syllabus area: 1.3 & 3.5
Thomas Hayes	B5(a) was well answered and candidates had a good grasp of the subject. (b) posed difficulties for most candidates and only one third scored more than the pass mark.
B6	Syllabus area: 4.3
Thomas Hayes	B6(a) This was well answered and only one out of six candidates who answered this question failed to meet the pass mark. (b) This posed a difficulty for all but one out of 5 candidates, who was the only one to achieve the pass mark.