

BCS Level 4 Diploma in Data Analysis Concepts Specimen Paper  
Answer Key and Rationale – QAN 603/0823/0

Question	Answer	Explanation / Rationale	Syllabus Sections
1	A	Proprietary data is internally generated data or documents that contain technical or other types of information controlled by a firm to safeguard its competitive edge. It may be protected under copyright or trade secret laws.	1.2
2	A	Operational data includes internal control and operational environment information such as data on the company's workforce, direct competitors, creditors, suppliers and information on customers.	1.2
3	B	Open data is information that can be freely used, reused and redistributed by anyone with no existing local, national or international legal restrictions on access or usage.	1.2
4	B	This is a representation of the chronological order of the ETL (Extract, Transform and Load) process.	10.2
5	A	Predictive tools do not provide the actual outcome and other business knowledge will not provide key relevance and therefore accuracy.	11.1
6	A	Migration does not feature in the data lifecycle.	2.2
7	A	Extensible Markup Language (XML) is a common source of structured data. Recognising the way XML looks as it structures data allows a data analyst to understand how that data can be queried, transported, exported and more.	3.2
8	D	Structured data refers to any data that resides in a fixed field within a record or file. This includes data contained in relational databases and spreadsheets. Unstructured data formats are often layered within other formats. For example, plain text could be stored inside a document contained within a zip file.	3.2
9	D	Structured and unstructured data are both used extensively in data analysis.	3.5
10	D	Requirements analysis takes account of the possibly conflicting requirements of the various stakeholders, analysing, documenting, validating and managing software or system requirements.	4.2
11	A	Requirements elicitation is the practice of collecting the requirements of a system from users, customers and other stakeholders	4.3
12	C	Data validation is the process of ensuring that a program operates on clean, correct and useful data.	4.2
13	D	The quality of data is not subjective to the location in which it is stored.	5.1

Question	Answer	Explanation / Rationale	Syllabus Sections
14	C	Inaccurate data will prevent reliable analysis and insight. Correcting this will then take additional time to remedy. Typically, poor data quality will not have a positive effect on a business.	5.2
15	A	Obscuring or removing personal data will protect the individual by preventing accurate, genuine data from being available.	5.1
16	D	A high level of data quality will ensure compliance with legislative direction as well as providing insight into key organisational statistics, allowing informed business planning.	5.4
17	A	A hypothesis test is a statistical test that is used to determine whether there is enough evidence in a sample of data to infer that a certain condition is true for the entire population.	6.1
18	A	The (null) hypothesis is that there is no significant difference between specified populations, any observed difference being due to sampling or experimental error.	6.2
19	A	Interpolation is a method of constructing new data points within the range of a discrete set of known data points.	6.3
20	D	A pie chart is an accurate graphical representation of the categorical distribution of data.	6.4