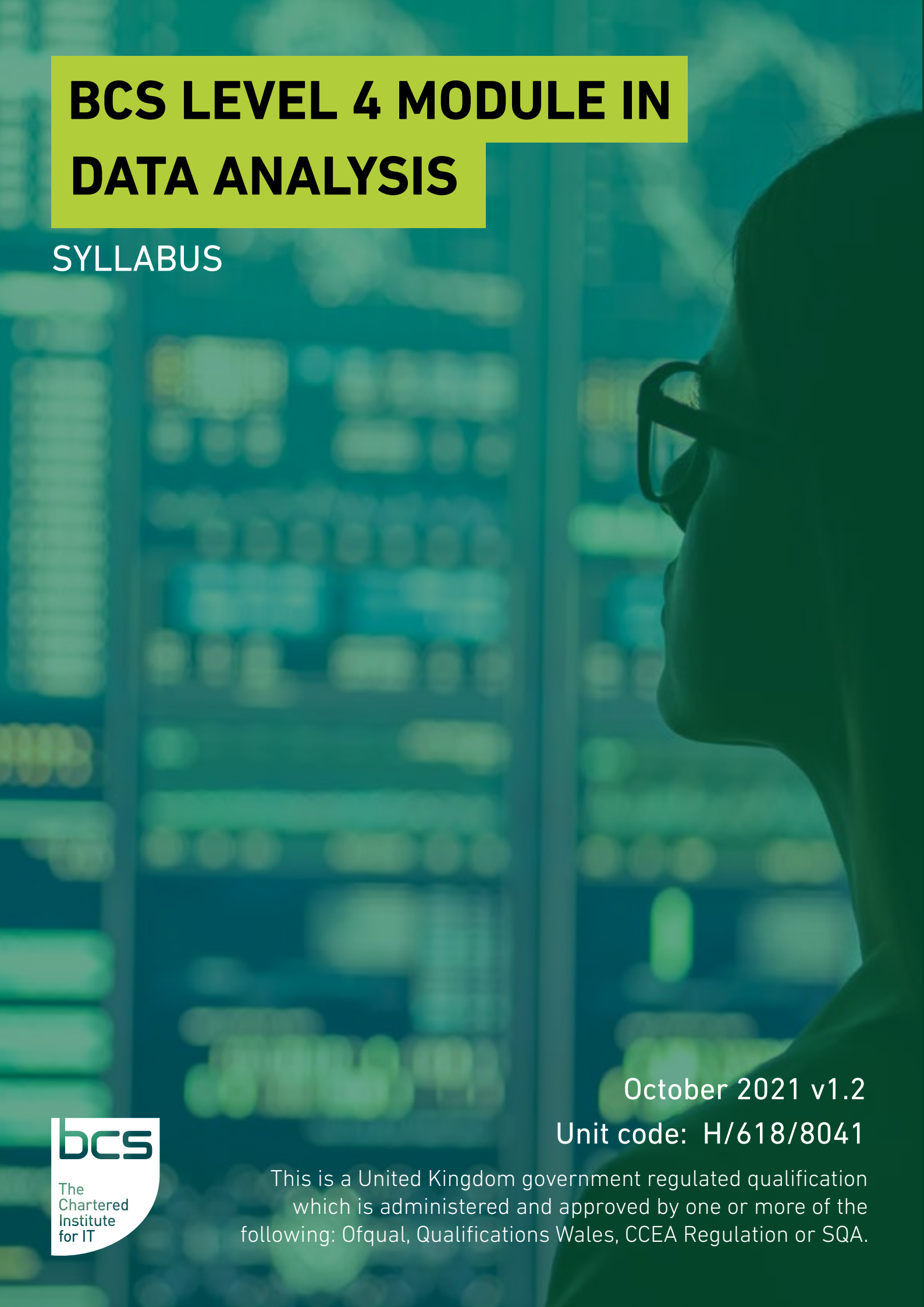
# BCS LEVEL 4 MODULE IN DATA ANALYSIS

SYLLABUS



October 2021 v1.2

Unit code: H/618/8041

This is a United Kingdom government regulated qualification which is administered and approved by one or more of the following: Ofqual, Qualifications Wales, CCEA Regulation or SQA.

bcs The Chartered Institute for IT

# CONTENTS




bcs
The Chartered Institute for IT

# Introduction

As data becomes an important currency in the world, enabling businesses to gain greater insight into their market and allowing them to better plan for the future, it is imperative that all organisations understand how best to use the data that is available to them. This requires having individuals within the business who are able to identify the requirements for using data, understanding which sources of data are best to use, who are able to use the systems and the tools available to integrate data from multiple sources, model it and prepare it for analysis, presenting the information it provides to key stakeholders to enable them to make better informed business decisions.

This Level 4 module covers the key concepts, skills and tools required by anyone working within a Data Analysis role to be able to successfully undertake each of the tasks required for the integration, preparation and analysis of data.

# Qualification Suitability and Overview

There are no mandatory requirements for candidates to be able to undertake this certificate qualification, although candidates will need a good standard of written English and Maths. Centres must ensure that learners have the potential and opportunity to gain the qualification successfully.

This qualification is suitable for candidates who are looking to progress their career within a data role. It can be taken as a standalone qualification, or in combination with other units and modules as part of a wider programme, such as an Apprenticeship.

This is an occupationally focused qualification which will:

- test a learner's applied knowledge, skills and behaviours to a range of scenarios.
- demonstrate a practical understanding of key concepts across the topic areas.
- enable a learner to progress in their career.

Candidates can study for this certificate by attending a training course provided by a BCS accredited Training Provider or through self-study

| Total Qualification Time | Guided Learning Hours | Independent Learning | Assessement Time |
|---|---|---|---|
| 320 hours | 227 hours | 92 hours | 1.5 hours |

# Trainer Criteria

It is recommended that to effectively deliver this certification, trainers should possess:

- 10 days training experience or have a train the trainer qualification
- A minimum of 3 years practical experience in a Data related role.

# SFIA Levels

This module provides candidates with the level of knowledge highlighted within the table, enabling candidates to develop the skills to operate successfully at the levels of responsibility indicated.

| Level | Levels of Knowledge | Levels of Skill and Responsibility (SFIA) |
|---|---|---|
| K7 | | Set strategy, inspire and mobilise |
| K6 | Evaluate | Initiate and influence |
| K5 | Synthesise | Ensure and advise |
| K4 | **Analyse** | **Enable** |
| K3 | **Apply** | **Apply** |
| K2 | **Understand** | **Assist** |
| K1 | **Remember** | **Follow** |

## SFIA Plus

This syllabus has been linked to the SFIA knowledge skills and behaviours required at level 4 for an individual working in a Data role.

## KSB01 Analytical Thinking:

Acquiring a proper understanding of a problem or situation by breaking it down systematically into its component parts and identifying the relationships between these parts. Selecting the appropriate method/tool to resolve the problem and reflecting critically on the result, so that what is learnt is identified and assimilated.

## KSB02 Conceptual Thinking:

Acquiring understanding and insights regarding the underlying issues in complex problems or situations through the development of abstract representations, the identification of patterns and the analysis of hypotheses.

## KSC09 Information Modelling Tools:

Proficient in using tools (manual or automated) to record the structure, relationships and use of information within an organisation. Examples, but not limited to class diagram and relational data model.

## KSC06 Database Software:

Familiar with software which enables the user to capture, create, populate and manipulate data structures and where appropriate unstructured data. Examples, but not limited to: MongoDB and NoSQL.

## KSC51 Big Data, Data Analytics and AI:

Aware of the discipline associated with data sets so large and/or complex that traditional data processing applications are inadequate. The data files may include structured, unstructured and/or semi-structured data, such as unstructured text, audio, video, etc. Challenges include analysis, capture, curation, search, sharing, storage, transfer, manipulation, analysis, visualization and information privacy.

## KSCA5 Data Handling

Proficient in the ability to harvest, clean, curate, manage, process and manipulate data in a variety of formats.

### KSD04 Information Elicitation Techniques:

Proficient in the selection and application of information elicitation methods, tools and techniques which are appropriate to the information required and the sources available. Examples, but not limited to: focus groups and surveys/questionnaires.

### KSD85 Stakeholder Engagement/ Analysis:

Proficient in establishing relationships, analysing perspectives and managing stakeholders from a variety of backgrounds and disciplines. Adapting stakeholder engagement style to meet the needs of different audiences. The identification of key business stakeholders and an assessment of their level of power and interests, and their perspectives to inform the way(s) in which they should be considered and managed.

Further detail regarding the SFIA Levels can be found at www.bcs.org/levels.

# Learning Outcomes

Upon completion of the certificate candidates will be able to demonstrate a practical understanding of:

- how to classify different types of data and the stages of the data lifecycle.
- how structured and unstructured data can complement each other to derive rich insight through data analysis.
- Identify the context for data analysis and how to gather customer requirements for data analysis.
- how to ensure good quality data whilst complying with Data Protection regulation.
- the principles of data architecture, data modelling and database design.
- how to integrate data from multiple sources and prepare it for analysis.
- how to undertake each of the stages of the Data Analysis Lifecycle.
- how to interpret the data, visualise it, and document and communicate the information to key stakeholders.

# Syllabus

## 1. Classifying data (4%) (K3)

### Learners will be able to:

**1.1** Describe how different forms of data can be applied to complex business situations.

#### Indicative content

a. Data, information and knowledge.
b. Working with multiple formats and sources of data.
c. Benefits and limitations of different data.

#### Guidance

Learners will be able to understand the requirements for working with complex data sets in multiple formats, processing raw data into information, in order to enable a business to gain knowledge and insight for informing business decisions.

---

**1.2** Explain the range of different types of data and the implications for allowable use, data quality, privacy concerns and availability.

#### Indicative content

a. Open, public and proprietary data.
b. Types of data e.g. research, operational.

#### Guidance

Learners will be able to work with different types of data, understanding the requirements for its use and selecting the most suitable data for the purposes of the business requirement.

---

**1.3** Demonstrate how to classify data, understanding its use within the business situation.

## Indicative content

a. Structured and unstructured data.
b. Quantitative and qualitative data.
c. Categorical (discrete) and continuous data.
d. Binomial and NOIR (Nominal Ordinal Interval Ratio) data.

## Guidance

Learners will be able to identify the types of data required to provide the necessary information for a particular business situation. This includes working with multiple types of data, using sources of both structured and unstructured data.

---

**1.4** Analyse and interpret the flow of an information system's data, understanding the business requirements at each stage of the lifecycle.

## Indicative content

a. The Data Lifecycle.

## Guidance

Learners will be able to explain how data is used at each of the stage of the data lifecycle, and the systems and processes that are required to manage the flow of data at each stage within the context of a business.

## 2. Data Structures (14%) (K3)

### Learners will be able to:

**2.1** Identify different data structures, explaining how they are used to form logical groupings.

| Indicative content | Guidance |
|---|---|
| a. Lists.<br>b. Arrays.<br>c. Records.<br>d. Trees.<br>e. Tables.<br>f. Graphs. | Learners will be able to work with different and often complex data structures, understanding their use in identifying, accessing, and manipulating data attributes by forming logical groupings of attributes. The list shown includes general terms for different groupings of data. Learners should be aware that when using specific languages, e.g. R or Python, that other terms might be used to represent these groupings. |

**2.2** Explain common sources of structured data.

| Indicative content | Guidance |
|---|---|
| a. Data files organised sequentially or organised serially.<br>b. Tables stored within a database management system.<br>c. Extensible Markup Language. | Learners will be able recognise different sources of structured data, understanding their use in a business context as well as their limitations. |

**2.3** Explain how structured data can be processed by data analysis tools.

| Indicative content | Guidance |
|---|---|
| a. Importing and linking to data. | Learners will be able use data analysis tools to process structured data files e.g. text files, csv, tables. They should be able to explain the complexities of analysing unstructured data. |

**2.4** Identify various formats of unstructured data.

## Indicative content

a. Examples include:
   - Word processor and PowerPoint files.
   - Audio.
   - Video.
   - Sensor and log data.
   - Social media feeds.
   - Paper-based documents.

## Guidance

Learners will be able to work with different types of unstructured data that is generated through often complex and multiple sources, such as through the use of devices and applications (e.g. cloud technologies, smartphones, mobile apps, social media, or through non-digital data collection methods).

---

**2.5** Illustrate how structured and unstructured data can complement each other to derive rich insight.

## Indicative content

a. Enhance analysis of the other (structured or unstructured text data).
b. Combined into a common model.
c. Big Data analytics.

## Guidance

Learners will be able to discuss the need for and the benefits of analysing different types of data to enhance a business's awareness and understanding of its internal and external environment. They should examine the concept of Big Data and the need to manage data from multiple sources that sit outside of a structured format. They may be encouraged to explore business analytics tools such as Power BI and how they are commonly being used by businesses for data visualisation to share, report, and make better use of data. They may wish to also consider the uses of Machine Learning within the analysis of structured and unstructured data – although they should not expect to be tested on this within the assessment.

---

## 3. The Principles of Data Analysis (12%) (K4)

### Learners will be able to:

**3.1** Apply relevant domain (industry/organisation) knowledge to enable effective data analysis.

| Indicative content | Guidance |
|---|---|
| a. Understanding the business context<br>• Business aims and goals<br>• Organisational structure, culture and operations<br>b. Professional standards; skills, knowledge, behaviours. | Learners will be able to interpret and evaluate the needs of a business organisation, and the context in which data analysis is required in order to apply solutions to business-critical systems. They will recognise the practical, theoretical and technical knowledge required for their role in order to problem solve and deliver effective business solutions. They will understand the value of continuous professional development towards ensuring their knowledge and skills remain current and up-to-date. |

**3.2** Implement different types of data analysis to solve specific business problems.

| Indicative content | Guidance |
|---|---|
| a. Decision analytics.<br>b. Descriptive analytics.<br>c. Predictive analytics.<br>d. Prescriptive analytics. | Learners will be able to undertake different forms of data analysis, identifying which approach will deliver the most effective results for the business based on its requirements. They will be able to illustrate the value of data and the ways in which it can enable a business to make better business decisions. |

**3.3** Explain the need to comply with Data Protection regulation.

| Indicative content | Guidance |
|---|---|
| a. Rights and obligations.<br>b. Enforcement agencies.<br>c. Regulatory and legal penalties.<br>d. Organisational policies and procedures. | Learners should understand the laws they are bound by in relation to data (collecting, storing, using, sharing) with consideration towards the principles of GDPR regulation. Learners may be encouraged to evaluate the policies and procedures in place within their own context that govern the operational approaches towards data management and use, and what the implications may be for them and the business if these procedures are not followed. |

**3.4** Apply data protection principles to manage any privacy issues that may occur during data analysis activities.

## Indicative content

a. Discuss the types, formats and activities that are protected:
   - Personally Identifiable Information.
   - Sensitive data.

b. Roles and responsibilities in using personal data:
   - Data controllers.
   - Data subject.
   - Data processor.

c. Individual rights and the right to be forgotten.

## Guidance

Learners will be able to identify protected information e.g. personal information, with an applied understanding of the rules and rights associated with how it may be collected, stored and used by an organisation; and therefore the implications of this in relation to data analysis activities.

---

**3.5** Analyse customer requirements and recognise the best way to obtain the relevant information.

## Indicative content

a. Classifying different types of requirements:
   - General requirements, such as business policies and standards.
   - Technical requirements.

b. Explain the difference between validation and verification.

## Guidance

Learners will be able to demonstrate how they classify different types of customer requirements, and how they validate and verify these requirements prior to undertaking any analysis of data.

---

**3.6** Apply the requirements elicitation process.

## Indicative content

a. Explicit vs. tacit knowledge.

b. Different elicitation techniques. For example, observe, recount, enact.

## Guidance

Learners will be able to select an appropriate elicitation technique in order to gather the requirements from their stakeholders. The ability to categorise and elicit both tacit and explicit knowledge is integral to the requirements elicitation process. Elicitation is concerned with purposefully extracting requirements from stakeholders, a process which requires different skills and techniques to simply gathering information. Learners should consider different techniques towards requirements elicitation to ensure they are able to achieve a full set of requirements.

---

**3.7** Interpret various data models used in the requirements gathering process.

## Indicative content

a. Logical, physical, and conceptual data models.

## Guidance

Learners will be able to select an appropriate data model, understanding when each type of data model might be used, for what purpose, and who they might be used with. They should be able to use data modelling as a means to determine or verify customer requirements.

---

**3.8** Explain the importance and necessity of good quality data.

## Indicative content

a. Benefits of having good data.
b. Data quality processes, practices.
c. Legal and regulatory compliance.
d. Commercial and intellectual property.
e. Confidentiality, integrity, and availability.
f. Improved business decision making.

## Guidance

Learners will be able to evaluating the factors that influence the quality of a business's data, advising others on how to ensure good quality data and why it is important.

---

**3.9** Demonstrate how to identify common sources of errors and how to avoid and/or resolve them.

## Indicative content

a. Sources:
   - Entry/Transcription (type of errors).
   - Data integration or migration activities.
b. Avoid or resolve through:
   - Process (a process for data quality and checking).
   - Identification (formal process or standardised means to identify types of errors).
   - Usage (using the right data for the purposes of the task).
   - Validity (checking the source of the data is valid).
   - Structure (ensuring data is in the right format and location e.g. so that the right tables are used to enable associations to be made between data).

## Guidance

Learners will be able to identify errors within data, and the sources of each error. They should be able to take suitable actions in order to resolve each error, suggesting how further errors can be avoided. They should consider the policies and processes in place within their own context that are designed to reduce data errors.

---

**3.10** Explain how minor data errors can cause major issues for data analysis.

## Indicative content

a. Cost.
b. Accuracy.
c. Inconsistency.
d. Cleanliness.

## Guidance

Learners will be able to advise on the implications associated with even small errors in data, explaining how their role in data analysis can enable a positive outcome towards each of these implications.

---

**3.11** Explain how to take account of data quality in preparing data for analysis to improve accuracy, quality and usefulness.

## Indicative content

a. Data profiling.
b. Data quality dimensions:
    - Completeness.
    - Uniqueness.
    - Timeliness.
    - Validity.
    - Accuracy.
    - Consistency.
c. What happens when an error or issue is found?:
    - Accept.
    - Reject.
    - Correct error.
    - Create default value.

## Guidance

Learners will be able to undertake data profiling in order to determine what good data looks like, and then how to identify and address any errors or issues with data that do not conform to the data quality dimensions as listed.

## 4. Data Architecture (35%) (K4)

### Learners will be able to:

**4.1** Explain how an organisation's data architecture defines how data is stored, managed, used and integrated within an organisation and its database systems.

| Indicative content | Guidance |
|---|---|
| a. Rules.<br>b. Policies.<br>c. Standards.<br>d. Models. | Data Architecture typically comprises of Rules, Policies, Standards, and Models that govern how data is collected and stored by a business. Learners will be able to contextualise Data Architecture in relation to their own environment through the evaluation of examples. |

**4.2** Explain the nature of the Data Architecture functions.

| Indicative content | Guidance |
|---|---|
| a. Data migration.<br>b. Data modelling.<br>c. Data integration.<br>d. Data warehousing.<br>e. Database design. | Learners will have a practical understanding of each of these functions and how they contribute towards data management. Learners may be directed to consider the Data Science Lifecycle if they wish to gain a deeper understanding of these functions and how they can support the implementation of an AI solution and the use of Machine Learning, for example. |

**4.3** Explain the nature and challenges of data volumes being processed through integration activities and how a programming approach can improve this.

## Indicative content

a. Big data:
   - Unstructured data.
   - Structured data.

b. Technical requirements for managing large data sets:
   - The location of data and challenge of restrictions due to the computer architecture (software and the system).

c. Data migration.
d. Master data management.
e. Integration design:

- Rules and requirements.
- Objectives and deliverables.
- Support models and SLAs.

f. Data integration tools (SQL):
   - Future scalability.
   - Implementation.
   - Support costs.

g. Data synchronisation:
   - Data ownership.
   - Frequency of updates.
   - Format.
   - Security.
   - Data quality.
   - Performance.
   - Maintenance.

## Guidance

Learners will be able to explain the requirements for working with large volumes of data, the need for a greater level of data governance when working with large datasets, and how an organisation may work to achieve 'one single source of truth' when approaching data management.

---

**4.4** Apply data modelling techniques within database design, producing data models from conceptual, logical and physical perspectives.

## Indicative content

a. Conceptual.
b. Logical.
c. Physical.
d. ERDs (Entry relationship diagrams).
e. Data dictionary.

## Guidance

Learners will be able to use data modelling in order to design and build a database based on the customer requirements gathered during the elicitation process.

---

**4.5** Recognise the most common forms of database.

## Indicative content

a. Relational.
b. Hierarchical.
c. Network.
d. Object-oriented.
e. Multi-dimensional (data cubes and hypercubes).
f. NoSQL.

## Guidance

Learners will have a practical understanding of each of these database types, in which context they would be used and for which types of data. For example, a relational database is commonly used for structured data whereas a NoSQL is commonly used for unstructured data. Learners will be able to propose the use of a particular database type, understanding the benefits and limitations of each type.

---

**4.6** Demonstrate how a logical data model can be transformed into a physical database design.

## Indicative content

a. Normalisation:
   - Redundancy free.
   - Unambiguous.
   - Flexible/extensible.
b. De-normalisation:
   - Introduction of derivable data (cumulative values, flags /status values).
   - Splitting logical data structures.
   - Combining logical data structures.
   - Introducing potentially redundant relationships.

## Guidance

Learners will be able to demonstrate normalisation to 3rd normalised form. They should be able to discuss de-normalisation, however there is no requirement for them to be able to carry it out at this level due to its complexity.

---

**4.7** Demonstrate how data can be queried within a database through the use of SQL queries.

## Indicative content

a. Single queries (SELECT, UPDATE, INSERT and DELETE)
   - Select and select* statements.
   - From, Where, AND, OR.
   - Use of wildcards and ordering.

b. Querying multiple tables of different information:
   - Joins.
     - Inner and outer.
     - Right and left.
     - Full.
     - Union.
     - Joins with duplicate values.
     - Joining on multiple fields.
     - Select into.
     - Subqueries.
   - Selecting the first/last of occurrences.

c. Aggregation.
   - Avg.
   - Count.
   - Max.
   - Min.
   - Group by.
   - Round.

d. Expressions and Functions:
   - CASE.
   - DATETIME.
   - Compound.
   - Cast.
   - Convert.
   - ISNULL.

e. Explicit and Implicit data conversion.

## Guidance

Learners should be able to create and run SQL queries on one or more tables.

---

**4.8** Explain the importance of database maintenance.

## Indicative content

a. Performance.
b. Data compaction.
c. Defragmentation.
d. Integrity check.
e. Log file maintenance.
f. Data warehousing.

## Guidance

Learners will have a practical understanding of database maintenance as an essential activity designed to keep a database running smoothly. They will be able to identify the cause of issues associated with a lack of maintenance such as slow performance or a loss of functionality, understanding how these can be resolved through regular maintenance tasks.

## 5. Data Preparation and Integration (35%) (K4)

### Learners will be able to:

**5.1** Demonstrate how data from multiple sources can be integrated to provide a unified view of the data.

#### Indicative content

a. Identifying suitable data sources.
b. Filtering data to ensure only relevant data is combined to underpin business objectives.
c. Data integration techniques:
   - Common user interface.
d. Dashboard.
e. Scorecard.
f. Dynamic.
   - Virtual integration.
g. Communication channels.
h. Data transfer.
   - Physical data integration.
i. ETL (extract - transform - load).
j. ELT (extract - load - transform)

#### Guidance

Learners will be able to create a basic dashboard for data. They should understand the terms ETL and ELT but not demonstrate or apply them.

---

**5.2** Explain how data manipulation is achieved and the purpose and outputs of data integration activities.

#### Indicative content

a. Functional requirements.
b. Non-functional requirements:
   - Speed.
   - Time available.
c. Information structure and rules;
   - Policies.
   - Practices.
d. Rationale for using and integrating data from multiple sources.
e. Importance of data in a business context.

#### Guidance

Learners will be able to consider why data is used and what the use of Big Data can enable a business to achieve.

---

**5.3** Analyse and compare the capabilities of statistical programming languages and software analysis tools.

## Indicative content

a. Speed, cost, function.
b. Capabilities and functions of statistical programming languages:
   - R.
   - Python.
c. Capabilities and functions of software tools e.g.:
   - Excel.
   - PowerBI.
   - Tableau.

## Guidance

Learners will be able to compare the features of different languages and tools used within Data Analysis (including those used within their own context), and how they can be used for manipulating, processing and cleaning data.

---

**5.4** Demonstrate how statistical programming languages are used in preparing data for analysis and within analysis projects.

## Indicative content

a. Preparation techniques:
   - Searching and sorting.
   - Grouping.
   - Filtering.
   - Modelling.
b. Data cleaning to remove a range of data issues:
   - Types of errors.
c. Missing data.
d. Inconsistencies.
e. Redundancy:
   - Invalid values.
   - Data that is out of range.
   - Outliers.
f. Processing and analysing:
   - Mean, median, mode and range.
   - Probability.
   - Bias.
   - Statistical significance.
   - Linear regression.
   - Scatter plots and correlation.
   - AND / OR probability.
   - Stem and leaf plots (frequency and distribution).
   - Box and whisker plots.
g. Methods for presenting results:
   - Tables.
   - Charts.
   - Graphs.

## Guidance

Learners will be able to use their chosen software to undertake initial data analysis, using different sets of data.

# 6. The Data Analysis Lifecycle (15%) (K2)

## Learners will be able to:

**6.1** Implement the typical routine steps of data analysis.

### Indicative content

a. Data Analysis lifecycle:
   - Problem hypothesis.
   - Identify what to measure.
   - Collect data.
   - Summarise data (inc visualisation).
   - Cleanse data.
   - Model data.
   - Analyse data.
   - Interpret results.
   - Document and communicate results (including visualisation).

### Guidance

Learners will be able to undertake each of the typical steps undertaken as part of the Data Analysis Lifecycle, reviewing the effectiveness and appropriateness of the methods and actions undertaken, and the results achieved.

---

**6.2** Explain how routine data analysis includes creating a problem hypothesis and identifying what to measure.

### Indicative content

a. Creating a problem hypothesis:
   - Understanding the importance of null and alternative hypotheses.
   - Understanding the subject area for analysis.
   - Finding similar previous analysis and exploring existing definition, assumptions, and reconciliation requirements.

b. Identifying what to measure:
   - Selecting the data sources.
   - Selecting aggregation and/or summarisation level.

### Guidance

Learners will be able to identify the problem that needs to be solved, and to determine what data will be needed and which sources of data might be used.

---

**6.3** Explain that routine data analysis includes clarification and confirmation of the requirement and identification of the right data and location.

## Indicative content

a. Collecting data:
   - Understand the size, nature and content of the data.
   - Identification of the data security and accessibility.
   - Complete data extraction.
   - Complete data transfer.

b. Summarising the datasets through visualisation.

c. Data loading cleansing data through:
   - Filtering.
   - Interpolation.
   - Transformation.
   - Masking.
   - Blending.

## Guidance

Once they have established the data required to solve the problem, learners will employ suitable methods for collecting data and summarising it through visualisation (in order to make an assessment as to whether the data is suitable for their purposes) before taking time to load and clean it.

---

**6.4** Explain that routine data analysis includes modelling data.

## Indicative content

a. Data preparation.
b. Provide assumptions made in preparation.
c. Identifying and selecting an appropriate model.
d. Model definition.
e. Train the model.
f. Testing and Verification of predictive models .
g. Troubleshooting.
h. Validation testing.
i. Strategy for improving model performance.

## Guidance

Learners should be able to identify which data modelling technique to use for the purposes of the task they are undertaking.

---

**6.5** Identify testing requirements to ensure that unified data sets are correct, complete and up-to-date.

## Indicative content

a. Testing dataset is complete and up-to-date.
b. Update model.
c. Further training.

## Guidance

Learners will understand that developing a data model can be an iterative process, where a model may be improved each time it is updated, trained and tested to ensure it is delivering the intended results. Learners will appreciate that a data model will need to be maintained after it has been deployed to ensure it continues to deliver the results needed for the customer based on their requirements.

**6.6** Explain how routine data analysis is used for analysing data, as well as for interpreting, documenting, and communicating results.

### Indicative content

a. Analysing data:
   - Reconcile and compare with other sources.

b. Interpreting results:
   - Understand the relationship between variables.
   - Show and compare the results in terms of real-world objects.

c. Visualising data:
   - Understand which type of visual data is suitable for the customer.
   - Types of charts (such as line graph, column and bar charts, pie chart, scatter plot, histogram, radar / spider chart, waterfall chart).
   - Geospatial distributions such as heat maps, bubble maps.
   - Time series such as time plot, Gantt chart.
   - Unstructured data such as Word Cloud.

d. Documenting and communicating results:
   - List the models and assumptions.
   - Understand your customer and stakeholders needs and communication style.

### Guidance

Learners will be encouraged to think about the needs of their customer and how they may best communicate the results of their analysis, with consideration towards which information will be most relevant/of interest to their audience and which information may be redacted. They should consider ways of presenting the results that provide the best visual description of the information being presented to ensure understanding is achieved.

# Examination Format

This module is assessed through completion of an invigilated online exam which candidates will only be able to access at the date and time they are registered to attend.

| | |
|---|---|
| **Type** | Digital exam which includes:<br>• a knowledge-based assessment consisting of 20 questions.<br>• a set of scenario-driven situational judgement assessments consisting of 4 scenarios each with a set of 5 questions, 20 in total. |
| **Duration** | 90 minutes |
| **Supervised** | Yes |
| **Open Book** | No (no materials can be taken into the examination room) |
| **Passmark** | 26/40 (65%) |
| **Delivery** | Digital format only |

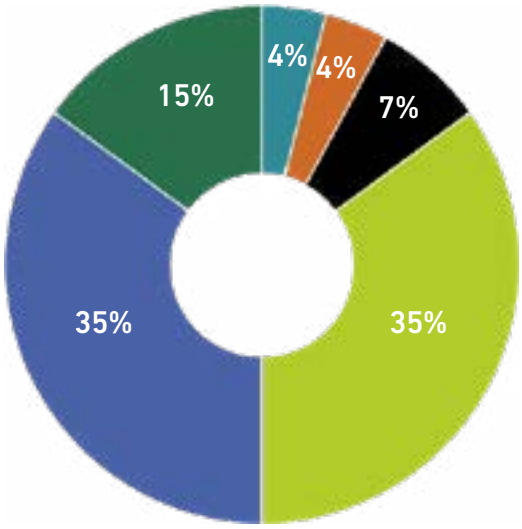
Adjustments and/or additional time can be requested in line with the BCS reasonable adjustments policy for candidates with a disability, or other special considerations including English as a second language.

# Question Weighting

Each major subject heading in this syllabus is assigned a percentage weighting. The purpose of this is:

1. Guidance on the proportion of content allocated to each topic area of an accredited course.
2. Guidance on the proportion of questions in the exam.

| Syllabus Area | Question type | |
|---|---|---|
| **1.** Classifying Data | Multiple choice and scenario-based multiple response | **4%** |
| **2.** Data Structures | | **4%** |
| **3.** The Principles of Data Analysis | | **7%** |
| **4.** Data Architecture | | **35%** |
| **5.** Data Preparation and Integration | | **35%** |
| **6.** The Data Analysis Lifecycle | | **15%** |
| **Total** | | **100%** |




**Syllabus Weighting**

4% 4% 7% 35% 35% 15%

# Recommended Reading

The following titles are suggested reading for anyone undertaking this award. Candidates should be encouraged to explore other available sources.

| | |
|---|---|
| **Title:** | Data Analyst |
| **Author:** | Harish Gulati, Charles Joseph, Rune Rasmussen, Clare Stanier, Obi Umegbolu Edited by Rune Rasmussen |
| **Publisher:** | BCS |
| **Publication Date:** | March 2019 |
| **ISBN:** | 9781780174327 |

| | |
|---|---|
| **Title:** | Data Governance: Governing data for sustainable business |
| **Author:** | Alison Holt, Benoit Aubert, David Sutton, Frédéric Gelissen, Alisdair McKenzie, Geoff Clarke, Rose Pan, Ming Li, Rohan Light, Beenish Saeed, Nathalie de Marcellis-Warin, Abdelaziz Khadraoui Edited by Alison Holt |
| **Publisher:** | BCS |
| **Publication Date:** | September 2021 |
| **ISBN:** | 9781780173757 |

| | |
|---|---|
| **Title:** | Managing Data Quality: A Practical Guide |
| **Author:** | Tin Ming, Julian Schwarzenbach |
| **Publisher:** | BCS |
| **Publication Date:** | May 2020 |
| **ISBN:** | 9781780174594 |

# Using BCS Books

Accredited Training Organisations may include excerpts from BCS books in the course materials. If you wish to use excerpts from the books you will need a license from BCS. To request a license, please contact the Head of Publishing at BCS outlining the material you wish to copy and the use to which it will be put.

# Document Change History

Any changes made to the syllabus shall be clearly documented with a change history log.  This shall include the latest version number, date of the amendment and changes made. The purpose is to identify quickly what changes have been made.

| Version Number | Changes Made |
|---|---|
| **Version 1.0** | Document creation. |

# CONTACT

For further information please contact:

**BCS**
The Chartered Institute for IT
3 Newbridge Square
Swindon
SN1 1BY

**T** +44 (0)1793 417 445

www.bcs.org



The Chartered Institute for IT