ncfe.

all you need to know.

NCFE Level 2 Certificate in Data Analytics (603/7482/2)

Qualification Specification

Contents

Section 1

About this qualification

Introduction

This qualification specification contains details of all the units and assessments required to complete this qualification.

To ensure that you are using the most up-to-date version of this qualification specification, please check the version number and date in the page footer against that of the qualification specification on QualHub.

If you advertise this qualification using a different or shortened name, you must ensure that students are aware that their final certificate will state the full regulated qualification title.

Reproduction by **approved** centres is permissible for internal use under the following conditions:

- you may copy and paste any material from this document; however, we do not accept any liability for any incomplete or inaccurate copying and subsequent use of this information
- the use of PDF versions of our support materials on QualHub will ensure that correct and up-to-date information is provided to students
- any photographs in this publication are either our exclusive property or used under licence from a third-party. They are protected under copyright law and cannot be reproduced, copied or manipulated in any form. This includes the use of any image or part of an image in individual or group projects and assessment materials. All images have a signed model release
- the resources and materials used in the delivery of this qualification must be age-appropriate and due consideration should be given to the wellbeing and safeguarding of students in line with your institute's safeguarding policy when developing or selecting delivery materials

Support handbook

This qualification specification must be used alongside the mandatory support handbook on the qualifications page on QualHub, which contains additional supporting information to help with the planning, delivery and assessment, such as:

- definition of total qualification time (TQT)
- quality assurance
- staffing requirements
- assessment
- qualification support
- diversity and equality

This qualification specification contains all of the qualification-specific information you will need that is not covered in the support handbook.

Qualification summary	
Qualification title	NCFE Level 2 Certificate in Data Analytics
Qualification number (QN)	(603/7482/2)
Aim reference	(60374822)
Total qualification time (TQT)	237
Guided learning hours (GLH)	215
Minimum age	16
Qualification purpose	This qualification is designed to equip students with the knowledge and skills required to enter junior positions within the digital workforce area as a Data Technician. For students wishing to progress their studies beyond Level 2, this qualification can be used as the technical skills option within a TQ transition programme for the T Level Technical Qualification in Digital Business Services (Level 3) (Delivered By NCFE) (603/6902/4).
Aims and objectives	This qualification aims to: • focus on the study of data in the digital sector • offer breadth and depth of study • provide opportunities to acquire a number of skills The objectives/themes of this qualification are to: • understand data and the relationship between different types of data • understand the principles of collecting, analysing and presenting data • understand the role of a data technician • understand data protection legislation and threats to ICT systems • source and prepare data for the purpose of data analysis • prepare, structure and blend data from multiple sources • interpret and present data to inform business decision making
Rules of combination	In order to achieve the Level 2 Certificate in Data Analytics, students must achieve: • a minimum of a pass grade in each of the two mandatory externally set internally marked assessments
Grading	Pass/merit/distinction
Assessment method	Externally set and internally marked assessments

Progression	Students who achieve this qualification could progress to: • T Level Technical Qualification in Digital Business Services (Level 3) (Delivered By NCFE) (603/6902/4) • Level 3 Data Technician Apprenticeship • Junior Data Analyst (or similar job role)
Regulation information	This is a regulated qualification. The regulated number for this qualification is 603/7482/2.
Funding	This qualification may be eligible for funding. For further guidance on funding, please contact your local funding provider.

Entry guidance

This qualification is designed to equip students with the knowledge and skills required to enter junior positions within the digital workforce area as a Data Technician.

Entry is at the discretion of the centre. However, students should be aged 16 and above to undertake this qualification.

There are no specific prior skills/knowledge a student must have for this qualification. However, students may find it helpful if they have already achieved a level 2 qualification.

Centres are responsible for ensuring that this qualification is appropriate for the age and ability of students. They need to make sure that students can fulfil the requirements of the learning outcomes and comply with the relevant literacy, numeracy and health and safety aspects of this qualification.

Students registered on this qualification should not undertake another qualification at the same level with the same or a similar title, as duplication of learning may affect funding eligibility.

Achieving this qualification

To be awarded this qualification, students must achieve a minimum of a pass grade in each of the two mandatory externally set internally marked assessments

Area of content

The following is a summary of the themes included within the area of content.

Area of content	Level	Theme
Area of content – Data Technician	2	• Theme 1: Data and the relationship between different types of data • Theme 2: The principles of collecting, analysing and presenting data • Theme 3: The role of a data technician • Theme 4: Data protection legislation and threats to ICT systems • Theme 5: Source and prepare data for the purpose of data analysis • Theme 6: Prepare, structure and blend data from multiple sources • Theme 7: Interpret and present data to inform business decision making

How the qualification is assessed

Assessment is the process of measuring a student's skill, knowledge and understanding against the standards set in a qualification.

This qualification has two assessments which are both externally-set by NCFE and internally-marked by centres. Both assessments contain extended tasks which assess knowledge and practical skills.

Students must be successful in **both** assessments to gain the Level 2 Certificate in Data Analytics.

The assessments have been designed to assess knowledge and skills through real-world tasks which reflect typical activities students will be required to carry out when working in junior data technician roles.

The assessments also contain tasks that students will undertake if they progress onto higher level Digital qualifications including the T Level Technical Qualification in Digital Business Services (Level 3) (Delivered By NCFE) (603/6902/4).

Students who are not successful can resubmit work within the registration period; however, a charge may apply.

Assessment and grading guidance

The following tables provide an overview of the themes covered in each assessment and the weighting applied to each assessment.

Assessment	Weighting	Content assessed
1. Knowledge-based assessment	TBC	• Theme 1: Data and the relationship between different types of data • Theme 3: The role of a data technician • Theme 4: Data protection legislation and threats to ICT systems
2. Practical assessment	TBC	• Theme 2: The principles of collecting, analysing and presenting data • Theme 5: Source and prepare data for the purpose of data analysis • Theme 6: Prepare, structure and blend data from multiple sources • Theme 7: Interpret and present data to inform business decision making

This qualification has **two** assessments. They will be sat in separate assessment windows, which will be scheduled at different points during the academic year.

Assessment 1 - Knowledge-based assessment

Assessment methods	Description
Extended written tasks assessing knowledge from themes 1, 3 and 4	• Externally set by NCFE • Tasks are internally marked by centres using levels-based mark schemes • Externally quality assured by NCFE • The completion time for the knowledge-based assessment is TBC hours • The assessment will take place during a window in TBC, specified by NCFE • The assessment must be completed under controlled conditions and supervised by the teacher • The assessment will target assessment objectives TBC • The assessment requires a computer with word-processing software and graphic design software capable of producing an infographic. Full details of the resources required for delivery are provided in section 4 of this qualification specification and in the sample assessment materials

Assessment 2 - Practical assessment

Assessment methods	Description
Practical tasks assessing knowledge and skills from themes 2, 5, 6 and 7	• Externally set by NCFE • Internally marked by centres using levels-based mark schemes • Externally quality assured by NCFE • The completion time for the practical assessment is TBC hours • The assessment will take place during a window in TBC, specified by NCFE • The assessment must be completed under controlled conditions and supervised by the teacher • The assessment will target assessment objectives TBC • The assessment requires a computer with TBC. Full details of the resources required for delivery are provided in section 4 of this qualification specification and in the sample assessment materials

Assessment objectives

The assessments are mapped against assessment objectives (AOs). These AOs provide a consistent framework for students, allowing students to show their knowledge, understanding and skills from across the full breadth and depth of the qualification.

The AOs that will be assessed against the content in this qualification are:

AO1	**Recall knowledge and show understanding** The emphasis here is for students to recall and communicate the fundamental elements of knowledge and understanding.
AO2	**Apply knowledge and understanding** The emphasis here is for students to apply their knowledge and understanding to real-world contexts and novel situations, including finding creative solutions.
AO3	**Analyse and evaluate knowledge and understanding** The emphasis here is for students to develop analytical thinking skills to make reasoned judgements and reach conclusions.
AO4	**Demonstrate and apply technical skills and processes** The emphasis here is for students to demonstrate the essential technical skills relevant to Data Analytics, by applying the appropriate processes, tools and techniques.

Assessment objective weightings

The table below shows the weightings for each of the AOs.

AOs	Assessment 1	Assessment 2	Overall weighting
AO1	TBC	TBC	TBC
AO2	TBC	TBC	TBC
AO3	TBC	TBC	TBC
AO4	TBC	TBC	TBC
Overall weighting of assessments	**TBC**	**TBC**	**100%**

The table below shows the weightings for the assessment tasks and the relevant AO(s) assessed in each task.

Assessment	Task	Task weighting	AO weighting
1	Part A	TBC	TBC
	Part B	TBC	TBC
2	Part A	TBC	TBC
	Part B	TBC	TBC
	Part C	TBC	TBC

Externally set and internally marked assessments

Each student is required to undertake two assessments.

For further information, centres should refer to the regulations for the conduct of external assessment and qualifications specific instructions for delivery documents available on the policies & documents page on QualHub.

Where qualifications have external assessment, centres must have entered students using the Portal to access the assessment.

Centres must enter students at least 10 working days in advance of the assessment window to avoid late entry fees.

If applicable, pre-release material will be made available by NCFE in advance of the assessment. All centres with entries will be notified.

The assessment material will be sent out in time for the start of the assessment. Assessment materials must be kept secure at all times.

Assessment windows

The assessments consists of an:

- assessment window (supervised). The centre arranges supervised periods of external assessment within a set window

For assessments sat in windows, the centre must enter students to the specified window. This will be either a set date and time assessment or a window in which the assessment will be completed.

For qualifications with 'entry on registration', the centre will choose the assessment window at the point of registering the student. The last date that we will accept student work for a specified assessment window is by that assessment window's cut-off date.

Please note: the 'cut-off date' is the last day that returned scripts will be accepted for the specified assessment window.

On completing their work at the end of the assessment window, students must sign the assessment declaration to authenticate the work produced as their own. Centres must ensure that all assessments are submitted for marking in accordance with the assessment windows.

Not yet achieved grade

A result that does not achieve a pass grade will be graded as a not yet achieved grade. Students may have the opportunity to resit.

Awarding the final grade

Each assessment is graded pass, merit or distinction. the grades for both assessments are aggregated to give an overall grade of pass, merit or distinction for the qualification.

Students must achieve a minimum of a pass in both assessments to achieve the overall qualification.

How grading works - link to grading calculator TBC

Worked example eg a table TBC

Descriptors

Descriptors have been written for the AO(s) assessed in each task. The descriptors are pitched at different levels in language and expectation and describe the student's performance at that band.

Assessors will use the descriptors to determine the banding decision for each AO in each task. This banding is based on the assessment of the evidence that the student submits for each task.

Overall grading descriptors

To achieve a pass, students will able to:

- TBC

To achieve a distinction, students will able to:

- TBC

Results - TBC

Resits - TBC

For further information on assessment, please refer to the user guide to the external quality assurance visit report.

Whilst NCFE does not anticipate any changes to our aggregation methods or any overall grade thresholds, there may be exceptional circumstances in which it is necessary to do so to secure the maintenance of standards over time. Therefore, overall grade thresholds published within this qualification specification may be subject to change.

Assessment strategy

Knowledge statements:

- assessors will need to be both occupationally knowledgeable and qualified to make assessment decisions
- internal quality assurers will need to be both occupationally knowledgeable and qualified to make quality assurance decisions

Competence/skills statements:

- assessors will need to be both occupationally competent and qualified to make assessment decisions
- internal quality assurers will need to be both occupationally knowledgeable and qualified to make quality assurance decisions

Section 2

Area of content

Area of content

This section provides details of the mandatory teaching content of this qualification.

The numbering is sequential throughout the themes in the area of content, from the first knowledge statement, following on through the skills statements. The 'K' and 'S' indicate whether the statement belongs to knowledge or skills.

The explanation of terms explains how the terms used are applied to this qualification. This document can be found in section 3.

For further information or guidance about this qualification, please contact our customer support team.

Data Technician

Theme 1 - Data and the relationship between different types of data
Knowledge - What you need to teach

The student must understand:

K1.1   what is meant by the term data:

- facts and statistics collected for reference and/or analysis (for example data on news channels/social media networks)

- real world examples of data that's recorded about individuals every day (for example card payments, mobile phone GPS, city size, car ownership, mobile phone ownership or personal data recorded on a daily basis etc).

K1.2   common sources of data:

- websites

- individuals

- databases

- online/offline

- Internet of Things (IoT) devices (for example smartphones).

K1.3   units used to measure the size of data:

- Byte (B)

- Kilobyte (KB)

- Megabyte (MB)

- Gigabyte (GB)

- Terabyte (TB)

K1.4   the difference between the following data structures:

- structured:
  - structured data is data that adheres to a pre-defined data model making it easier to analyse. Structured data usually comes in a tabular format with relationship between the different rows and columns

- un-structured:
  - unstructured data is information that either does not have a pre-defined data model or is not organised in a pre-defined manner

- semi-structured:
  - semi-structured data is information not contained within a relational database but is organised to an extent that makes it easier to analyse.

K1.5   the principles of big data:

- sometimes big data is known as complex data

- big data is concerned with analysing large amounts of data

Theme 1 - Data and the relationship between different types of data

- analysis of large data sets requires different techniques to traditional data analysis (for example, data mining)

K1.6   difference between qualitative and quantitative data:

- quantitative:
  - data expressing a certain quantity, amount or range:
    - integer
    - decimal
    - numerical
- qualitative:
  - non-numerical data related to characteristics and qualities rather than trends and statistics:
    - image data
    - text (free text)

K1.7   the purpose of different keys used in relational databases:

- primary key:
  - provides unique reference for each item in a database
  - ensures each column can be uniquely referenced
- foreign key:
  - allows the user to define relationships between tables
  - provides links between parent and child tables

K1.8   the advantages and disadvantages of using programming languages (eg SQL (Structured Query Language) and Excel for data storage, manipulation and retrieval:

- SQL:
  - advantages :
    - accessible for use when managing databases
    - efficient when handling structured data
    - allows users to access many records with a single command
  - disadvantages:
    - has a complex interface that may be not be user friendly for most users
    - some versions are costly which may limit some programmers from accessing it
    - due to hidden business rules, complete control is not given to the database
- Excel:
  - advantages:
    - ease of use for collecting and organising data
    - allows the user to create presentations where the data is analysed such as pie charts or tables for easy viewing and interpretation

Theme 1 - Data and the relationship between different types of data

- streamlines calculations (for example, if the spreadsheet is set up to calculate your gross profit, when any variable such as cost per unit, shipping costs, or

   sales discount are changed, the software automatically recalculates the new gross profit based on the new information

- spreadsheets can be shared across multiple users who need access to the same document

- disadvantages:

  - only the information that the user chooses for analysis is included in these presentations, and therefore, other pertinent information that may influence decision making might be excluded, unintentionally

  - if users input the wrong data, even in only one cell of the spreadsheet, all related calculations and cells will be affected and have incorrect data

  - spreadsheets are not that secure and therefore are at greater risk for data corruption or mismanagement of information

K1.9   what is meant by data transformation including the steps required to transform data:

- processing and preparing data in order to obtain information:

  - the steps of data transformation:

    - input (data)

    - process (transform)

    - storage

    - output (information)

K1.10   the characteristics that define the quality of data:

- accuracy and precision

- legitimacy and validity

- reliability and consistency

- timeliness and relevance

- completeness and comprehensiveness

- availability and accessibility

Theme 2 - The principles of collecting, analysing and presenting data

Knowledge - What you need to teach

The student must understand:

K2.1   the characteristics of data architecture:

- data architecture is composed of models, policies, rules or standards that governs:

  - data identified for collection

Theme 2 - The principles of collecting, analysing and presenting data

- o storage
- o arrangement
- o integration
- o applied (in data systems and in organisations)

K2.2   methods used for collecting data:

- methods of data collection:
  - o surveys
  - o interviews
  - o online tracking
  - o in-store traffic monitoring
  - o social media monitoring

K2.3   steps used to collected data for decision making:

- analyse the collected data
- assess the relevance of collected data
- apply the collected data to decision making

K2.4   the process of data analysis:

- process of inspecting, cleansing, transforming and modelling data with the goal of discovering useful information

K2.5   the primary reasons businesses or individuals might use data analysis:

- drive decision making:
  - o allows individuals to make informed decisions
  - o provide evidence for decision
- spot trends
- measure performance

K2.6   the different types of data analysis:

- diagnostic:
  - o specific situations
  - o problem solving
- descriptive:
  - o specific situations
  - o management of information

K2.7   the different types of data analytics:

- descriptive:
  - o investigate historical data (what was the result)

Theme 2 - The principles of collecting, analysing and presenting data

- diagnostic:
  - investigates the why of data (why did the result turn out in a certain way)
- predictive:
  - uses current and historic data to make predictions about future trends (what will be the result in the future)
- prescriptive:
  - investigates what can be done to influence future results (what can be done to alter results in the future)

K2.8   the characteristics of ethical data analysis:

- beneficial:
  - the value of data for the business/individual vs personal interests (for example the usefulness or merit that comes from solving the problem so it might be evaluated appropriately)
- progressive:
  - a culture of continuous improvement and data minimisation (for example what the business/customer learns from applying data should help deliver better and more valuable results and the least amount of data is necessary to meet the desired objective, with the understanding that minimising data usage promotes more sustainable and less risky analysis)
- sustainable:
  - insights identified with data is sustainable over time (for example big data insights, when placed into production, should provide value that is sustainable over a reasonable time frame)
- respectful:
  - ensure there has been transparency and inclusivity (for example any data analysis that affects individuals to whom the data pertains, businesses that originated the data, businesses that aggregate the data, and those that might regulate the data in different ways)
- fair:
  - potential impacts of data use is considered (for example ensuring no discrimination or offence against protected characteristics and considering if the data would strengthen or threaten customer relationships.

K2.9   difference between reports and dashboards:

- reports present a static picture of data
- dashboards provide a dynamic/configurable view of data

K2.10   strengths and limitations of different presentation techniques:

- graphs (for example line and bar graphs):
  - advantages:
    - can be created proportionally to the quantity it needs to represent

Theme 2 - The principles of collecting, analysing and presenting data

- displays multiple categories of data in one chart (for example, monthly expenses on food items categorised as fruits, vegetables, and packaged food)
- puts large sums of data into visual form for easy understanding
- offers easy calculations of data accuracy
- understood easily by different departments within a business and for media purposes
  - o disadvantages:
    - doesn't reveal exact values
    - multiple graphs are needed for time-lapse data
    - key assumptions, causes, effect, and patterns are not revealed
    - manipulated easily, causing false impressions or interpretations
- pie charts:
  - o advantages:
    - presents a simple and easy to understand picture
    - represents data visually as a fractional part of a whole, which can be an effective communication tool for an uninformed audience
    - enables the audience to see data comparison to make an immediate analysis or to understand information quickly.
    - pieces of data can be manipulated in the pie chart to emphasise points
  - o disadvantages:
    - if too many pieces of data are used, pie chart becomes less effective
    - a series of pie charts will be needed to compare multiple sets as this chart only represents one data set.
    - as the audience has to factor in angles and compare non-adjacent slices, it has its problems in comparing the data slices.
    - to make decisions based on visual impact rather than data analysis leads audience to draw inaccurate conclusions
- data tables:
  - o advantages:
    - can be useful to show comparisons between certain types of data, for example what services each department provides.
    - data is organised clearly in rows and columns
    - a large amount of data can be easily confined in a data table.
    - used for statistical analysis (for example, calculation of central tendency, dispersion)
  - o disadvantages:
    - less effective when presenting a trend over time (for example, how many patients entered an emergency department)

Theme 2 - The principles of collecting, analysing and presenting data

- ▪ the amount of data could impact the number of columns before the table width causes horizontal scrolling on smaller screens
- ▪ columns that are narrow to prevent horizontal scrolling will decrease readability of text in cells
- infographics:
  - o advantages:
    - ▪ simplifies complex concepts or ideas
    - ▪ increases the potential of knowledge retention
    - ▪ can be customised to align with a business brand image
  - o disadvantages:
    - ▪ they can be time consuming to create (for example, the layout needs to be designed so that it is easy to read, includes graphics that correspond with the data, and a crafted story throughout the infographic to make it a meaningful experience for the user)
    - ▪ due to limited space for in-depth information the data can be misinterpreted

K2.11   **factors that will influence the presentation of data:**

- requirements of the audience:
  - o internal (granular detail)
  - o external (clear and accessible)
- knowledge of audience
- audience familiarity with the dataset
- time constraints for presentation
- format of the data (for example face to face vs poster/infographic)

Theme 3 - The role of a data technician

Knowledge - What you need to teach

The student must understand:

K3.1   **different sectors that a data technician could work in:**

- retail/sales (for example Sales Analyst)
- health (for example Health Information Manager)
- finance (for example Financial Data Manager).
- workforce management (for example Performance Management)
- uniformed services (for example Emergency Response Coordinator)
- energy (for example Pricing Analyst)

Theme 3 - The role of a data technician

- marine (for example Maritime Data Analyst)

K3.2   key responsibilities of a data technician:

- data collecting
- maintain adherence with data protection requirements
- analyse data collected (for example, data analysis)
- design the structure of data (for example, data architecture)
- preparation and presentation of data
- data validation
- distribution and dissemination of data

K3.3   principal non-technical skills required for the data technician role:

- communication
- teamwork
- problem solving
- numeracy
- organisation
- reflective evaluation

K3.4   data handling skills that can be applied across sectors/organisations:

- database design and data architecture
- report structuring
- cleansing and preparing data

K3.5   the benefits a data technician can provide to an organisation:

- aids in the protection of data
- provides useful information through analysis by identifying:
  - customer/client/workforce needs
  - operational inefficiencies
  - data trends that may indicate fraudulent behaviour

K3.6   progression routes within business data management:

- Data Analyst Technician
- Data Analyst Advanced Technician:
  - Data Analyst
  - Data Architect
  - Information Systems (IS) Business Analyst
- Data Scientist

Theme 3 - The role of a data technician

- methods to support progression:
  - o academic and training:
    - Apprenticeships
    - short or long courses
    - Digital T-Levels
    - Level 4 Leadership and Management
  - o on the job development

K3.7 importance of continuous professional development:

- the definition of CPD:
  - o the improvement of personal and professional skills throughout a career
- importance of CPD:
  - o keeping up to date with new innovations
  - o best practice
  - o improving on skill set
  - o evidencing professional development

Theme 4 - Data protection legislation and threats to ICT systems

Knowledge - What you need to teach

The student must understand:

K4.1 what is meant by personal data:

- information that relates to an identified or identifiable individual
- types of personal data:
  - o name and surname
  - o home address
  - o email address (for example, name.surname@company.com)
  - o identification card number (for example, medical or banking)
  - o sensitive data:
    - racial or ethnic origin
    - political opinions
    - religion
    - membership of a trade union
    - health
    - sexual orientation

Theme 4 - Data protection legislation and threats to ICT systems

- criminal record

K4.2   the purpose of the General Data Protection Regulation (GDPR):

- protect individuals' fundamental rights and freedoms,
- protect individuals' right to protection of their personal data
- the right to one's private life within the European Convention on Human Rights (ECHR)

K4.3   key principles of the GDPR:

- lawfulness, fairness and transparency
- purpose limitation
- data minimisation
- accuracy
- storage limitation
- integrity and confidentiality

K4.4   the characteristics of a personal data breach:

- breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data transmitted, stored or otherwise processed
- types of data breaches:
  - o loss or theft of:
    - hard copy notes
    - USB drives
    - computers or mobile devices
  - o an unauthorised person gaining access to a laptop, email account or computer network
  - o sending an email with personal data to the wrong person

K4.5   responsibilities of employers and employees in the event of a personal data breach:

- employer responsibilities:
  - o setting up incident response
  - o investigating the breach
- employee responsibilities:
  - o report any identified data breaches
  - o provide follow up information to individual handling the data breach
- reporting procedures:
  - o follow business reporting procedures when contacting the appropriate person or team
  - o report a notifiable breach to the Information Commissioner's Officer (ICO) without undue delay, but no later than 72 hours after becoming aware of it

K4.6   consequences of data breaches:

Theme 4 - Data protection legislation and threats to ICT systems

- compensating affected customers

- investment into new security measures

- legal fees

- reputational damage (for example, loss of customer trust that their data will be protected)

- regulatory penalties that can be imposed for non-compliance with the GDPR

K4.7   the purpose of the DPA:

- protects the privacy and integrity of data held on individuals by businesses and other organisations. The act ensures that individuals (customers and employees) have access to their data and can correct it, if necessary.

K4.8   impact of the following elements of the DPA on the Data Technician role:

- general data processing

- law enforcement processing (for example tampering data, changing records)

- intelligence services processing (for example secured services)

- regulation and enforcement:
  - ICO Information Commissioner Office

K4.9   the differences between the DPA and the GDPR:

- the GDPR applies not only to EU organisations but also non-EU organisations in certain circumstances

- the DPA applies only to companies that control the processing of personal data (Controllers)

- the GDPR extends the law to those companies that process personal data on behalf of Controllers (Processors) (for example, if you buy a new television on the internet and give your contact details to the web-based store to enable delivery of that television, only the store would be liable for looking after your personal information under the DPA. Under the GDPR, the store and the delivery company could both be liable)

- the GDPR allows the regulator to fine non-compliant companies up to 4% of global turnover. under the DPA, the largest fine allowed is £500,000

- under the DPA the regulator recommends that organisations notify it if they experience a data breach. However, under the GDPR there is a requirement to notify the regulator and individuals' affected under certain circumstances

K4.10 common types of physical threats to ICT systems and data:

- physical access:
  - damage to computer systems hardware and infrastructure:
    - theft
    - material instability
    - improper storage environment (for example, temperature, humidity, light, dust)
    - overuse (for example, from physical contact media)

Theme 4 - Data protection legislation and threats to ICT systems

- natural disaster (for example, fire, flood, earthquake)
- infrastructure failure (for example, plumbing, electrical, climate control)
- inadequate hardware maintenance
- hardware malfunction
- vandalism

K4.11 common types of electronic threats to ICT systems and data:

- non-physical threats:
  - targets the software and data on the computer systems.
- denial of service (DoS):
  - a cyber-attack in which the perpetrator seeks to make a machine or network resource unavailable to its intended users by temporarily or indefinitely disrupting services of a host connected to the internet.
    - buffer overflow attack
    - ICMP flood (leverages misconfigured network devices by sending spoofed packets that ping every computer on the targeted network, instead of just one specific machine)
    - SYN flood (sends a request to connect to a server, but never completes the handshake)
- data theft or damage:
  - the act of stealing information stored on corporate databases, devices, and servers
- unauthorised use:
  - individuals gaining access to an organization's data, networks, endpoints, applications or devices, without permission. It is closely related to authentication – a process that verifies a user's identity when they access a system
- malware:
  - any software intentionally designed to cause damage to a computer, server, client, or computer network that include:
    - computer viruses
    - worms
    - trojan horses
    - ransomware
    - spyware
    - adware
    - rogue software
    - wiper
    - scareware
- social engineering:

Theme 4 - Data protection legislation and threats to ICT systems

- o the psychological manipulation of people into performing actions or divulging confidential information
- o types of social engineering include:
    - scamming
    - phishing
    - vishing
    - pretexting
    - baiting
    - quid pro quo
    - tailgating

K4.12 security vulnerabilities associated with remote access technologies:

- Remote Desktop Service (RDS) is exposed on the internet:
    - o if a password is guessed successfully, the resulting access could have substantial repercussions for an organisation and facilitate further attacks against trusted or connected infrastructure
- Man-in-the Middle Attacks (MiTM):
    - o lack of identity verification allows a malicious person, by deploying other criminal activities, to intercept all communications sent between a client and a Terminal Server
- encryption attacks:
    - o the medium setting may facilitate the use of weak encryption which could be decrypted in a reasonable time-frame and lead to the disclosure of sensitive information
- Denial of Service Network Level Authentication (NLA):
    - o Terminal Servers which support NLA but do not have it configured present a risk. NLA forces the client computer to present user credentials for authentication before the server will create a session for that user

Theme 5 - Source and prepare data for the purpose of data analysis

Knowledge and Skills - What you need to teach

The student must understand:

K5.1 what is meant by primary and secondary data:

- primary data:
    - o based on own criteria (for example, a time and money related activity)
    - o also known as raw data
    - o collected by researchers directly from main sources through interviews, surveys and experiments

Theme 5 - Source and prepare data for the purpose of data analysis

- o usually collected from the source (where the data originally originates from)
- secondary data:
  - o criteria is being dictated (for example, using someone else to support your model)
  - o data that is collected by someone other than the primary user
  - o common sources of secondary data for social science include censuses, information collected by government departments, organisational records and data that was originally collected for other research purposes

K5.2   internal sources of data:

- data retrieved from inside the organisation to make decisions for successful operations:
  - o sales:
    - ▪ data gathered from the sales department to determine revenue, profit, and the bottom line
  - o finance:
    - ▪ data in the form of cash flow reports, production reports, and a budget variance analysis
  - o marketing:
    - ▪ data gathered from website traffic statistics, phone reports and promotion codes
  - o human resources:
    - ▪ data gathered from rates of retention, annual leave, absenteeism, timesheets and expenses, workforce management

K5.3   external sources of data:

- data retrieved from outside the organisation which is the market, including customers and competitors:
  - o paid data from professional data providers (for example dedicated portals or software):
    - ▪ provided at a cost (for example, pay for use, subscriptions, freemium)
  - o social media data from user generated content (for example connection to official access points or web-crawling of social media platforms i.e Twitter, Facebook, Instagram):
    - ▪ freely available subject to copyrights
  - o questionnaires/surveys

K5.4   how to assess the reliability of external sources of data:

- social media data:
  - o data is unstructured and can be difficult to assess or verify
  - o data is mainly user-generated, or combines self-reported and behavioural traces
  - o users' online behaviour doesn't necessarily represent their offline beliefs
  - o organisations, public figures, bots and influencers populate social media and can bias datasets

Theme 5 - Source and prepare data for the purpose of data analysis

- o   continuous arrival of new data means a complete dataset cannot be available
- o   the ways we use social media and the popularity of platforms change rapidly
- paid data:
  - o   data is structured and is likely to be from a verifiable source
  - o   in most cases, the reports or publishers will have a listing of how the information is gathered
  - o   the providers are not likely to be the producers of the products and services they analyse therefore biases will likely be at a minimum

K5.5   what data validation checks are required ensure source data is accurate:

- peer assessed/reviewed (for example, from source such as author)
- corroboration:
  - o   comparing information from two separate sources and find similarities between them. When a second source provides the same or similar information to the first, the second source is considered to corroborate
- metadata checks:
  - o   metadata is data about data (for example, a document might include a collection of information like the author, file size, the date the document was created, and keywords to describe the document, metadata for an audiobook file might include the author's name, book title, and the year it was published)

K5.6   characteristics of fit for purpose data:

- contains accurate information
- relevant to user requirements
- consistent and reliable
- from legitimate and verifiable sources
- available and accessible

K5.7   considerations when collecting data from different sources:

- validity
- reliability
- repeatability
- cost
- time
- useability

K5.8   the importance ethical collection of data:

- minimises the chance of physical or emotional harm
- ensures participants are aware of the risks
- ensures the study remains unbiased

Theme 5 - Source and prepare data for the purpose of data analysis

- seeks to stop personal opinions interfering with the process

The student must be able to:

S5.1   establish the data collection methods to be used:

- interviews/focus groups
- questionnaires/surveys
- observations
- case studies
- documents and records
- physical measurements

S5.2   plan the process for collecting fit for purpose data:

- select audience
- select questions where appropriate
- decide on sample size
- decide what needs to be measured

S5.3   collect data using the selected method

S5.4   input collected data into a data set

S5.5   perform the following data cleansing measures:

- remove duplicates
- remove blanks
- remove repetitions
- length checks
- type checks
- sense checks

S5.6   select and store data in a secure file format:

- csv
- json
- doc
- pdf

Theme 6 - Prepare, structure and blend data from multiple sources

Knowledge and Skills - What you need to teach

The student must understand:

Theme 6 - Prepare, structure and blend data from multiple sources

K6.1 **the purpose of data blending:**

- combines multiple sources of useful data into a functioning data set
- identifies correlations between the different data sets without the time and expense of traditional data warehouse processes
- allows better business decisions based on the quantity of data from a single data set

K6.2 **steps required to prepare data for blending:**

- establish the need
- establish the source
- cleansing the data (normalising the data)
- merge the data
- identify the structure to blend the data
- identify the storage type
- establish the format required for the data to be accessible

K6.3 **methods used to blend data:**

- relationships:
  - can be used in most instances, including across tables with different levels of detail
  - are flexible and are adaptable to the structure of the analysis on a sheet by sheet basis
- joins:
  - combines tables by adding more columns of data across similar row structures.
  - can cause data loss or duplication if tables are at different levels of detail
  - joined data sources must be fixed before analysis can begin
- blends:
  - queries each data source independently
  - the results are aggregated to the appropriate level
  - the results are presented visually together in the view
  - can handle different levels of detail and works with published data sources

The student must be able to:

S6.1 **identify data requirements and sources:**

- establish the need
- establish the source

S6.2 **select suitable data for blending**

S6.3 **prepare data for blending:**

- validate data (for example, removing unwanted data)
- cleanse data (for example, normalising data)

Theme 6 - Prepare, structure and blend data from multiple sources

S6.4   select a tool for blending:

- open source (for example SQL/Python)
- Excel

S6.5   group the data into corresponding criteria

S6.6   blend data from multiple sources

S6.7   review the data:

- validation checks
- error checks

S6.8   select and store data in a secure file format

Theme 7 – Interpret and present data to inform business decision making

Knowledge and Skills - What you need to teach

The student must understand:

K7.1   how businesses use customer data:

- customer buying trends/patterns
- customer preferences (for example, colour of product)
- the impact of price changes on customer behaviour

K7.2   how businesses use competitor data:

- a business' own and competitors market share
- price of competitors products vs own product
- the quantity and characteristics of competitor products
- location where competitor products are sold
- perception/reputation of competitor products

K7.3   how businesses use internal data:

- to identify over or under-resourced areas
- processes that may be inefficient (for example, an out of sync production line)
- identify and/or assess employee performance

K7.4   advantages and disadvantages of using data from:

- single sources (for example, the individual or person in an interview, an external dataset):
  - advantage:
    - data provides a personal account which should be more accurate
  - disadvantage:

Theme 7 – Interpret and present data to inform business decision making

  - person may be an outlier and not representative of the wider sample
- multiple sources (for example, data for decisions to be made such as quality assurance, material tolerance):
  - advantage:
    - data provides a wider range of opinions from the overall sample
  - disadvantage:
    - data may lose granularity when focused on the high-level information

K7.5   what is meant by the term 'data driven business':

- refers to the utilisation of data to inform or enhance processes, decision making, and the revenue model of a business
- data-driven businesses systematically and methodically uses data to make the final business decisions instead of intuition or personal experience or preference

K7.6   what are the characteristics of a 'data driven business':

- business intelligence (BI):
  - strategies and technologies used by enterprises for the data analysis of business information.
  - combines business analytics, data mining, data visualisation, data tools and infrastructure, and best practices to help organisations make data-driven decisions.
  - prioritises flexible self-service analysis
  - governs data on trusted platforms
  - empowers business users
  - enables speed to insight
- competitive intelligence:
  - sometimes referred to as corporate intelligence
  - the ability to gather, analyse, and use information collected on competitors, customers, and other market factors that contribute to a business's competitive advantage
  - helps businesses understand their competitive environment and the opportunities and challenges it presents
  - businesses analyse the information to create effective and efficient business practices

The student must be able to:

S7.1   find patterns in single and multiple data sets:

- common issues such as prices, sales, process time
- trends (for example, a quantity that is generally increasing or decreasing)
- predictions (for example, analysing data to understand the rate at which the numbers are increasing)
- correlation (for example, the statistical relationship between two sets of numbers)

S7.2   interpret and provide recommendations based on:

Theme 7 – Interpret and present data to inform business decision making

- internal data
- external data

S7.3   select data appropriate to business needs

S7.4   present data using an appropriate method and format:

- graphs
- charts
- data tables
- written text
- verbal

Section 3
Explanation of terms

Explanation of terms

This table explains how the terms used at level 2 in the area of content are applied to this qualification (not all verbs are used in this qualification).

Apply	Link existing knowledge to new or different situations.
Assess	Consider information in order to make decisions.
Classify	Organise according to specific criteria.
Compare	Examine the subjects in detail looking at similarities and differences.
Define	State the meaning of a word or phrase.
Demonstrate	Show an understanding of the subject or how to apply skills in a practical situation.
Describe	Write about the subject giving detailed information.
Differentiate	Give the differences between 2 or more things.
Discuss	Write an account giving more than one view or opinion.
Distinguish	Show or recognise the difference between items/ideas/information.
Estimate	Give an approximate decision or opinion using previous knowledge.
Explain	Provide details about the subject with reasons showing how or why. Some responses could include examples.
Give (positive and negative points…)	Provide information showing the advantages and disadvantages of the subject.
Identify	List or name the main points. (Some description may also be necessary to gain higher marks when using compensatory marking.)
Illustrate	Give clear information using written examples, pictures or diagrams.
List	Make a list of key words, sentences or comments that focus on the subject.
Perform	Do something (take an action/follow an instruction) that the question or task asks or requires.
Plan	Think about and organise information in a logical way. This could be presented as written information, a diagram, an illustration or other suitable format.
Provide	Give relevant information about a subject.
Reflect	Students should look back on their actions, experiences or learning and think about how this could inform their future practice.
Select	Choose for a specific purpose.
Show	Supply sufficient evidence to demonstrate knowledge and understanding.

State	Give the main points clearly in sentences.
Use	Take or apply an item, resource or piece of information as asked in the question or task.

Section 4
Additional information

Additional information

Resource requirements

To assist in the delivery of this qualification, centres/students should have access to the following mandatory resources

- a digital device – either desktop, laptop or tablet
- a storage medium
- web browser software/applications
- generic spreadsheet software capable of performing data analysis functions (for example, Microsoft Excel, Google Sheets, Open Office)
- generic presentation software (for example, MS PowerPoint, Google Slides)
- a printer
- Internet connectivity/access
- suitable data sets.

There is no requirement to use any specific software/applications.

Support for centres

Qualification factsheet

This document outlines the key information of this qualification for the centre, student and employer.

Learning resources

We offer a wide range of learning resources and materials to support the delivery of our qualifications. Please check the qualifications page on QualHub for more information and to see what is available for this qualification.

Contact us

NCFE
Q6
Quorum Park
Benton Lane
Newcastle upon Tyne
NE12 8BT

Tel: 0191 239 8000*
Fax: 0191 239 8001
Email: customersupport@ncfe.org.uk
Websites: www.qualhub.co.uk
(www.ncfe.org.uk)

NCFE © Copyright 2021 All rights reserved worldwide.

DRAFT 1.0 May 2021
Information in this qualification specification is correct at the time of publishing but may be subject to change.

NCFE is a registered charity (Registered Charity No. 1034808) and a company limited by guarantee (Company No. 2896700).

CACHE; Council for Awards in Care, Health and Education; and NNEB are registered trademarks owned by NCFE.

All the material in this publication is protected by copyright.

* To continue to improve our levels of customer service, telephone calls may be recorded for training and quality purposes.