Next Generation Higher National Unit Specification

Professional Practice in Data Science (SCQF level 8)

Unit code: J6C6 48

SCQF level: 8 (32 SCQF credit points)

Valid from: session 2024 to 25

Prototype unit specification for use in pilot delivery only (version 2.0) January 2024

This unit specification provides detailed information about this unit to ensure consistent and transparent assessment year on year.

This unit specification is for teachers and lecturers and contains all the mandatory information required to deliver and assess this unit.

Unit purpose

This specialist unit is a key part of the Higher National Diploma (HND) in Data Science. It integrates the knowledge and skills that learners on the HND programme achieve in learning data science concepts and the practical use of tools, including those for data acquisition and cleansing, exploration and analysis, and visualisation.

Learners are taught how to address and solve a data science problem in a professional manner, document their processes and communicate their findings to an audience. Most of their time is engaged in planning and managing the acquisition of a large dataset, and its investigation and analysis, while demonstrating the personal meta-skills that professional practice in data science requires.

Entry to this unit is at your centre's discretion, however, we recommend that learners have one or more of the following before they start:

♦ a wide range of data science and computing knowledge and skills
♦ digital skills such as the Digital Skills unit at SCQF level 7
♦ programming experience
♦ knowledge of project management concepts
♦ report writing skills

It is important that learners improve their professional practice and behaviours when developing their meta-skills.

On completing this unit, learners can progress to the PDA in Data Science at SCQF level 9 or to a degree programme in data science or computer science.

Unit outcomes

Learners who complete this unit can:

1   develop a solution by applying data science to a data problem
2   work collaboratively in a project team to plan and implement a solution to a data problem
3   apply data science methods and tools to transform and analyse data relating to a data problem
4   document and communicate the data science solution to a data problem and conclusions reached
5   exhibit professional practice and behaviours in the application of data science
6   develop their meta-skills in a vocational or academic context
7   develop sustainability knowledge and understanding, and skills, in a vocational context

Evidence requirements

Learners must provide both product evidence and performance evidence.

Product evidence

Evidence must include:

♦   an introduction to the data science problem and a description of the dataset selected and its source
♦   evidence of importing and cleaning data and data transformations
♦   a statement of the methodology used in the data analysis, including the programming language or other tools and techniques selected
♦   evidence from the implementation of the data analysis method
♦   a final project report containing conclusions and recommendations
♦   project management sections containing a project timeline, project diary and evidence of project progress meetings
♦   a reflection on the conduct of the project, the challenges met and overcome, and the ethical considerations of the team
♦   a self-assessment of meta-skills that includes the plan and activities to develop these skills, and reflection on improvement

This evidence must be a portfolio of work carried out over a period of time, under open-book conditions. You should arrange regular meetings to discuss their progress.

To initiate the project, you should ask each team of three, four or five learners to put together a short project outline describing what data they would like to analyse, the data source and the reasons for selection. Based on this, you should propose a data science project and set out requirements for each team.

The standard of evidence should be consistent with the SCQF level of this unit. You should use appropriate level descriptors when making judgements about the evidence.

You should distribute marks as follows:

♦ project report (85 per cent)
  — 30 per cent for data management and data analysis
  — 25 per cent for evidence and findings
  — 15 per cent for project management
  — 15 per cent for report writing and presentation

Performance evidence

Evidence is the team presentation of the investigation, the outcomes from data exploration and analysis, and its conclusions and recommendations. This should last for a minimum of 20 minutes, involve all team members, and be presented to a client audience. The team presentation of conclusions to client is worth 15 per cent of the marks.

Performance evidence should also demonstrate learners':

♦ appropriate use of digital platforms for collaboration
♦ integrity when dealing with data analysis and interpretation
♦ respect when collaborating with others
♦ communication in terms that are understandable to the client audience

The standard of evidence should be consistent with the SCQF level of this unit.

You should use appropriate level descriptors when making judgements about the evidence.

Meta-skills

Learners develop meta-skills in the course of doing all the units as part of this HND. In this unit, you also assess their meta-skills development as an outcome, following the evidence requirements set out below.

This meta-skills outcome is the same for the HNC (SCQF level 7) and the HND (SCQF level 8). Learners who progress from the HNC to the HND should continue to develop their meta-skills. They should gather evidence in line with the outcome requirements as they work through units and projects at HND level.

Evidence

Learners must gather evidence that shows they have:

♦ self-assessed their meta-skills baseline
♦ created a plan for their own meta-skills development
♦ carried out activities to develop and demonstrate their meta-skills
♦ used reflective practice to monitor and assess the meta-skills they have improved and developed

[Skills 4.0, a skills model to drive Scotland's future](), outlines three categories of meta-skills:

♦   self-management
♦   social intelligence
♦   innovation

Each of these comprises four meta-skills and a number of sub-skills.

There are many interrelationships and dependencies between these skills and, at SCQF level 7 and 8, learners should focus on holistic development relevant to their vocational or academic context.

See the HND Data Science Educator Guide for more information.

Outcome 7 (Learning for Sustainability)

Learners gather evidence that demonstrates they can:

♦   identify and describe sustainability in the context of the United Nations Sustainable Development Goals (UN SDGs)
♦   explain how one product or process relevant to networking and cloud infrastructure could be made more sustainable and help meet the aims of at least two selected UN SDGs

Knowledge and skills

The following table shows the knowledge and skills covered by the unit outcomes:

Knowledge	Skills

Learners should understand:
 	
♦ data sources, provenance and governance
♦ sources and conditions of use of open data
♦ the ETL process in the data pipeline
♦ approaches to data cleansing
♦ application of data transformation
♦ data warehousing and data formats
♦ data exploration and investigation
♦ data analysis and data insights
♦ programming concepts in a suitable language for data science (for example Python, R or Julia)
♦ data science libraries (for example SciKit, Pandas, Numpy, Plotly)
♦ software tools and workbenches for data management and analysis (such as Google Data Studio, IBM Watson, Weka)
♦ project management for a professional project
♦ ethics in the data science process
♦ meta-skills, specifically:
  — the categories of self-management, social intelligence and innovation, and associated meta-skills, as described in Skills 4.0
  — the importance of developing meta-skills, including employability, adaptability, and effectiveness
  — what meta-skills are most relevant to the learner's vocational context
	Learners can:
	
	♦ derive a project specification from client requirements
	♦ create a project plan
	♦ manage and monitor project progress
	♦ meet timescales and milestones
	♦ maintain ongoing record of activities and decisions
	♦ participate in team and client meetings
	♦ extract data from a source
	♦ cleanse and transform data
	♦ load data to a data warehouse
	♦ use data science tools and techniques to analyse data
	♦ plan a report template with a client and team
	♦ demonstrate programming skills in a high-level language
	♦ explore data with data science tools and techniques
	♦ present data in suitable formats, including data visualisations
	♦ draw conclusions from data analysis
	♦ use referencing in a report
	♦ create a report to a professional standard
	♦ document all data science processes, assumptions, models, algorithms and outcomes
	♦ plan a strategy for meta-skills development
	♦ implement and review plans for their meta-skills development
	♦ assess their meta-skills development

Knowledge	Skills
♦ approaches to developing meta-skills, in particular: — self-awareness: analysing preferences, strengths and weaknesses; meta-skills self-assessment — goal setting and action planning — reflective practice: principles of reflective practice; tools and approaches for effective reflective practice ♦ sustainability concepts ♦ the UN SDGs	Learners can demonstrate some or all of: ♦ performing data management effectively ♦ selecting algorithms to process data ♦ identifying and representing data structures ♦ working with data files in a programming language ♦ deploying a software tool or workbench to process data ♦ using statistical methods to describe data features ♦ using statistical methods to analyse data and test hypotheses ♦ using machine learning to build and test predictive models from data ♦ improve sustainability in a product or process

Meta-skills

Throughout this unit, learners develop meta-skills to enhance their employability in the data science sector.

The unit helps learners to develop the meta-skills of self-management, social intelligence and innovation. Learners should develop meta-skills naturally throughout the unit. You should encourage learners to develop a minimum of one area in each of the three categories, but they do not need to cover all suggested subsections. The following suggestions may help shape delivery and assessment, and vary depending on the chosen topics and assessment method.

Self-management

This meta-skill includes:

♦ focusing: sorting (sorting information into categories and understanding the relationships between information), filtering (filtering out non-essential information)

♦ integrity: self-awareness (being self-aware), ethics (raising questions of ethics), self-control (exercising self-control in reporting insights

♦ adapting: critical reflection (accepting new ideas and reflecting critically on them), self-learning (self-educating and responding constructively to change)

♦ initiative: risk taking (managing risks in processing data), self-belief (underpinned by self-belief and trust in their own judgement), responsibility (taking responsibility for actions)

Social intelligence

This meta-skill includes:

♦ communicating: listening (listening to and understanding the client brief), giving information (giving information to the team and using the outcomes to report to the client), storytelling (tell the story from the data)

♦ collaborating: relationship building, teamworking and collaboration, social perceptiveness, global and cross-cultural competence (cultural awareness)

♦ leading: inspiring others, influencing, developing others (sharing expertise and knowledge through coaching and exemplification)

Innovation

This meta-skill includes:

♦ curiosity: observation (noting significant aspects), questioning (questioning assumptions), information sourcing (exploration of data)

♦ creativity: idea generation, visualising (using visualisation in data storytelling), maker mentality (constructing solutions)

♦ sense-making: pattern recognition, analysis (systematic analysis of data to uncover relationships and an ability to see the big picture)

♦ critical thinking: deconstruction (deconstructing the data science problem), logical thinking (applying logical thinking), computational thinking (constructing a computational method for its solution)

Delivery of unit

The unit must be project-based and you should commence when learners have acquired a sufficiently wide range of vocational knowledge and skills in developing data science solutions.

The project must involve devising a solution to a real-world problem. The focus of the unit is not learning new technical knowledge and skills but applying knowledge and skills to a large-scale data analysis and interpretation activity.

You should deliver the unit as an ongoing project management unit, and provide learners with examples of data science problems, project management practices, and techniques for analysis and problem-solving. Throughout the unit you should arrange progress meetings with the team.

The time required varies depending on the previous experience of individual learners. Based on 160 hours delivery and assessment time, we suggest the following distribution:

Outcome 1 — develop a solution by applying data science to a data problem
(35 hours)

Outcome 2 — work collaboratively in a project team to plan and implement a solution to a data problem
(15 hours)

Outcome 3 — apply data science methods and tools to transform and analyse data relating to a data problem
(60 hours)

Outcome 4 — document and communicate the data science solution to a data problem and conclusions reached
(25 hours)

Outcome 5 — exhibit professional practice and behaviours in the application of data science
(10 hours)

Outcome 6 — develop their meta-skills in a vocational or academic context
(10 hours)

Outcome 7 — develop sustainability knowledge and understanding, and skills, in a vocational context
(5 hours)

Additional guidance

The guidance in this section is not mandatory.

Content and context for this unit

This unit involves a data science project. Learners can present their evidence as a single document in sections that match the bullets listed in the 'Evidence requirements' section. There should be scope for other presentation formats, such as an e-portfolio, showing the elements of the project and examples of progress. The e-portfolio might contain other media such as images, audio and video. If learners produce a single report, there should be a word count maximum of 2500 words.

Work with a team of learners to help them identify a dataset from open data that is of interest to them and formulate a client requirement based on this data. This forms the basis for the project objectives for the team.

In carrying out the data science investigation, learners must work in small groups (teams). In exceptional cases a learner may be allowed to work individually. In all cases, they must use project management methodologies throughout. The data science problem asks learners to engage with a client to discover the objectives of their investigation. You must perform the role of the client, and specify a project that draws on a data source selected by and agreed with learners. This could be an open data source provided by a local authority or a large public body. The project report must document the processes adopted, any data models and algorithms used, and the outcomes from the exploration and analysis of the data. The project must conclude with a presentation of the findings to an audience.

You should deliver this unit as an ongoing project management unit, and provide learners with examples of open data, project management practices and programming techniques. Throughout the unit, you should arrange regular progress meetings with the team.

Encourage learners to use open-source resources where possible. For example, they could use an open-source integrated development environment (IDE), such as Geany or Thonny, or RStudio or JupyterLab for data analysis. In these cases, you can support learners to install and configure software environments.

Assessment for grading

You should assess the product and performance evidence produced for the unit on a pass or fail basis, according to the standards set out in the assessment requirements. You can further consider the evidence as a component of grading in the HND Data Science, and apply the guidance on grading to the available evidence.

For guidance on grading, you should refer to the grading pack for HND Data Science.

Equality and inclusion

This unit is designed to be as fair and as accessible as possible with no unnecessary barriers to learning or assessment.

You should take into account the needs of individual learners when planning learning experiences, selecting assessment methods or considering alternative evidence.

Guidance on assessment arrangements for disabled learners and/or those with additional support needs is available on the [assessment arrangements web page](#).

Information for learners

Professional Practice in Data Science (SCQF level 7)

This information explains:

♦ what this unit is about

♦ what you should know or be able to do before you start

♦ what you need to do during this unit

♦ opportunities for further learning and employment

Unit information

In this unit you learn how to address and solve a data science problem in a professional manner, document your processes and communicate your findings to an audience. You produce a project report demonstrating your findings and your ability to work in a team to manage a project and conduct analysis of a dataset.

Before starting, we recommend that you have one or more of the following:

♦ a wide range of data science and computing knowledge and skills

♦ digital skills such as the Digital Skills unit at SCQF level 7

♦ programming experience

♦ knowledge of project management concepts

♦ report writing skills

Assessment is based on your project report and a team project presentation. A project proposal helps you to formulate and discuss your ideas with your lecturer before you start.

In the course of this unit you have the opportunity to practice data science skills such as:

♦ select and use software tools to manage the extraction of data from sources

♦ apply transformations and manage the dataset for analysis

♦ use a programming language to write data analysis programs

♦ source a suitable dataset and carry out analysis, then present the analysis

♦ document the project in a diary or timeline

♦ use data science-specific knowledge, as well as domain knowledge

♦ develop a mix of programming skills, project management skills, and skills in applied mathematics and statistics

♦ create reports and presentations

You develop a wide range of meta-skills and literacies while studying this unit. The meta-skills you develop cover self-management, social intelligence and innovation. You identify how data science methods have an impact on sustainability, and how they can be made more sustainable.

On completion of the unit, you may progress to qualification or degree programmes at SCQF level 9 in data science topics.

Administrative information

Published: January 2024 (version 2.0)

Superclass: AG

History of changes

Version	Description of change	Date
2.0	Change to grading model and addition of LfS.	January 2024

Note: please check [SQA's website](#) to ensure you are using the most up-to-date version of this document.