Next Generation Higher National Unit Specification

Mathematics for Data (SCQF level 8)

Unit code: J6CH 48

SCQF level: 8 (16 SCQF credit points)

Valid from: session 2023–24

Prototype unit specification for use in pilot delivery only (version 1.0) June 2023

This unit specification provides detailed information about the 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 the unit.

This edition: June 2023 (version 1.0)

Unit purpose

This unit focuses on the mathematical knowledge and skills that underpin data science, and is designed for learners studying machine learning-related courses. We recommend that learners have passed Higher Mathematics at SCQF level 6 or have equivalent knowledge and skills in calculus, algebra, vectors and functions before starting this unit.

Learners develop the mathematical skills that are essential to fully understand how modern machine learning algorithms operate. The unit covers two areas of mathematics: calculus and convex optimisation. It specifically covers:

- preliminaries: sets, real numbers, intervals, Cartesian coordinates, graphs, equations of straight lines
- properties of mathematical functions: graph of a function, surjection, injection, bijection, even and odd functions, convexity
- limits and differentiation: continuity of a function, difference quotient, product and chain rules, derivative of the most relevant linear and non-linear functions
- linear algebra: vectors and matrices, operations with vectors and matrices, determinants, matrix inverse, linear system of equations, eigenvectors and eigenvalues
- convex optimisation: multivariate functions, global optimum, gradient descent methods

Within calculus, the unit focuses on non-linear functions relevant to data science, such as square roots, logarithms, exponentials, trigonometric, and hyperbolic functions, in particular Tanh.

On completing this unit, learners have the necessary mathematical skills to progress in data science and machine learning-related Higher National units, such as:

- Machine Learning at SCQF level 8 or 9
- Artificial Intelligence at SCQF level 8
- Big Data at SCQF level 8

Unit outcomes

Learners who complete this unit can:

1 analyse mathematical functions to determine continuity and discontinuous points with the use of limits
2 apply differentiation rules to linear and non-linear functions
3 solve linear algebra problems
4 determine global optimum of mathematical functions using the concepts of convex optimisation

Evidence requirements

Learners must provide knowledge evidence.

It must be produced for each knowledge and skill, and collectively demonstrate that they can meet the outcomes. You can sample evidence if using testing. The test must be carried out under timed and controlled conditions without reference to materials, but formula sheets and calculators are allowed. The sample must include aspects of every outcome.

The ‘Additional guidance’ section provides further information on assessment standards and suggestions on how you should assess learners.

Authentication is required when the evidence is produced in lightly controlled conditions.

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.

Knowledge and skills

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

Knowledge	Skills
Learners should understand:	Learners can demonstrate skills in:
♦ mathematical preliminaries: — sets (particularly numerical sets) — real numbers — intervals — two-dimensional Cartesian coordinates — straight lines	♦ mathematical preliminaries: — set operations: union, intersection, difference — plot graphs in a cartesian coordinate system — determine the parameters of a straight line (slope and intercept)
♦ mathematical functions and main properties: — two-dimensional function graphs — surjection — injection — bijection — even and odd functions — convex functions	♦ mathematical functions: — determine function correspondence (surjection, injection, bijection) — determine symmetric relations in a function: even and odd
♦ limits and differentiation: — concepts of function continuity — difference quotient — methods of differentiation (product and chain rules) — stationary points — derivative test	♦ calculation of limits and derivatives of functions: — analyse discontinuity points in a function — application of the product rule — application of the chain rule — analysis of stationary points — determine if a function is convex with the use of derivative test
♦ derivatives of well-known linear and non-linear functions: — square root — logarithms — exponential — trigonometric functions — hyperbolic functions — second-order derivatives	

Knowledge	**Skills**
Learners should understand: ♦ linear algebra: — matrix and vectors — square matrices — identity matrix — matrix determinant — matrix inverse — eigenvectors and eigenvalues ♦ convex optimisation: — concept of function convexity — multivariate functions — partial derivatives — convex optimisation — gradient descent methods — Lagrange multipliers	Learners can demonstrate skills in: ♦ matrix operations: — compute scalar product — perform matrix multiplication — determine matrix determinant — determine matrix inverse using the analytic solution — solve linear equations of the form Ax+b=0 — determine eigenvectors and eigenvalues — use eigen decomposition to calculate the matrix inverse ♦ convex optimisation: — compute partial derivatives — apply gradient descent methods to determine global optimum — use Lagrange multipliers to enforce constrained optimisation

Meta-skills

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

Self-management

This meta-skill includes:

- focusing: sorting
- initiative: independent thinking

Social intelligence

This meta-skill includes:

- communicating: receiving information, giving information,

Innovation

This meta-skill includes:

- curiosity: observation, information sourcing, problem recognition
- creativity: visualising
- sense-making: synthesis, analysis
- critical thinking: deconstruction, logical thinking

Literacies

Throughout this unit, learners have opportunities to develop their literacy skills.

Numeracy

This is implicit in all areas of the unit and heavily used in some. Learners enhance their numeracy skills while studying the topics included in this unit.

Delivery of unit

This unit provides the mathematical background relevant to machine learning-related units. We recommend that you deliver this before the Big Data at SCQF level 8 and the Artificial Intelligence at SCQF level 8 units. It provides the knowledge and skills that are necessary for the Probability and Statistics for Data Science at SCQF level 8 unit.

We recommend that you deliver this unit as follows:

- lectures (60 per cent)
- tutorials (40 per cent)

You should design lectures to deliver the knowledge and use tutorials to deliver and develop skills, where learners can then solve mathematical exercises with your support.

We suggest the following distribution of time:

- Lectures (48 hours):
  - mathematical preliminaries: sets, real numbers, graphs, cartesian coordinate system, straight lines: 4 hours
  - properties of mathematical functions: non-linear functions (square root, logarithms, exponential, trigonometric, and hyperbolic functions): 6 hours
  - limits and differentiation: 10 hours
  - linear algebra: 16 hours
  - convex optimisation: 12 hours
- Tutorials (32 hours) with exercises on:
  - mathematical preliminaries and properties of mathematical function: 8 hours
  - limits and differentiation:10 hours
  - linear algebra: 7 hours
  - convex optimisation: 7 hours

Additional guidance

The guidance in this section is not mandatory.

It is important that this unit focuses on delivering the knowledge and skills to prepare learners for a deeper understanding of modern machine learning algorithms.

You should emphasise the importance of solving problems involving non-linear functions. Learners may already have knowledge of derivatives of simpler functions (quotients and/or polynomials) and this unit should encourage them to apply their existing knowledge and skills to solve more complex problems. This is important because some non-linear functions, such as square root, Tanh, or log, are widely used in data science, especially in statistics and machine learning. You should introduce and formalise such non-linear functions, highlighting their main characteristics and properties.

You can deliver parts of tutorials alongside computing examples, solving calculus problems using numerical approaches. For example, derivatives of continuous functions could be numerically solved in a computer program using the finite differences method, linear algebra or convex optimisation. In this case, learners can work on more complex problems (such as matrices of size 100×100). Python or MATLAB programming languages are well suited for this numerical approach to calculus. The decision to include computer programs to solve numerical problems in tutorial time is at your centre's discretion.

In general, you should state main theorems, but you can omit proofs. We also recommend that, after having completed the lectures on differentiation, you should introduce the basic concepts of integration. We recommend that learners practise linear algebra with a matrix of size up to 4×4. Convex optimisation, introduced in the last part of this unit, should include Lagrange multipliers and the reasons they are important in convex optimisation. This is an essential part in machine learning, supporting the needs of regular users when training neural networks. We recommend that you introduce convex optimisation using linear algebra and show learners how derivatives of matrices and vectors can be computed. This helps you to introduce partial derivatives of a multivariate function expressed in a vector form. Gradients are at the core of how neural networks are optimised, by way of gradient descent methods included in this unit.

Approaches to assessment

You can use a test to assess knowledge. This test can last 2 hours and should take place under closed-book conditions; however, you can provide learners with a formula sheet. As stated in the 'Evidence requirements' section, you can sample a subset of knowledge and skills, provided that you assess all outcomes. We recommend that the questions are independent of each other — this avoids a situation where a mistake in one question can negatively impact the whole assessment.

A sampled assessment could include the following example questions:

- Given some functions (for example, two), learners should determine domain, critical points (and whether the functions converge or diverge to such points with the use of limits), stationary points (with the use of first and second derivatives), and their classification (for example, maximum, minimum, or inflection points).
- Given some matrices (for example, four of different shapes, such as 2×2, 3×3, or 4×4), learners should perform some basic matrix operations (such as multiplication and sums), determine the determinant and find the inverse (if applicable). For 2×2 matrices, you could task learners with determining eigenvalues and eigenvectors (it is not advisable to use bigger matrices for the calculation of eigenvalues and eigenvectors because it can take a lot of time-consuming manual computations, jeopardising the whole assessment).
- Learners must calculate the derivative or gradient of objective functions.
- Learners must formulise a constrained optimisation problem using Lagrange multiplies.

The assessment should include some theoretical questions covering the knowledge offered in this unit. For example, you could ask learners to:

- provide some definitions, such as the definition of bijective function, derivative, stationary points, or limit
- define the determinant and/or the inverse of a matrix (and when this can be computed)
- define a convex function

These are examples of sampled questions that cover the expected learning outcomes. We recommend a threshold score of 50 per cent.

An alternative approach would be for learners to produce a portfolio of mathematics, maintained over a period of time. If you use this approach, it is vital that the portfolio provides evidence for every knowledge and skill (sampling is not acceptable), and that the material is the learner's own work. Learners must understand the material contained in their portfolios and authentication is essential. The Guide to Assessment provides further advice on methods of authentication.

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

Mathematics for Data (SCQF level 8)

This information explains:

- what the unit is about
- what you should know or be able to do before you start
- what you need to do during the unit
- opportunities for further learning and employment

Unit information

Data science is a discipline that has gained enormous success in the past few years. All major companies perform data analysis to improve the commercialisation of their products and services. However, data science is not a new field of study. It embeds knowledge and skills from several other disciplines, such as probability and statistics, information theory and artificial intelligence. All these disciplines are formalised with the language of maths.

This unit prepares you to study other data science-related units, as they all speak the same language. We recommend that you have passed Higher Mathematics at SCQF level 6 or equivalent knowledge and skills in calculus, algebra, vectors and functions before you start this unit.

You are introduced to two important topics of maths, calculus and optimisation, that are essential to advance towards the next steps in data science. You start with preliminary topics introducing you to sets, numbers, and cartesian coordinate systems. After that, you learn what a mathematical function is and how you can analyse it, with the use of two widespread and important mathematical tools: limits and derivatives. You learn how to deal with linear systems of equations using the array representations: vectors and matrix. This gives you the solid knowledge base to advance to the last part of the unit, covering convex optimisation. This topic is essential, because most machine learning algorithms use convex optimisation to discover new patterns in data.

The unit covers how you can operate with functions and determine their main properties. You analyse a function by calculating its limits and derivatives at specific points. You also learn how to operate with matrices and vectors to solve linear systems of equations, and how to use gradient descent methods to optimise convex functions.

All the knowledge and skills you acquire can be assessed with a question paper, covering both theoretical and practical questions. This should be completed over a limited period of time (for example 2 hours).

Through studying this unit, you enhance a range of skills, such as logical and independent thinking, which is developed through solving calculus and linear algebra problems. Problem recognition and deconstruction are other skills that you improve through solving complex problems by dividing them into smaller and more manageable problems. Analysis and synthesis are skills enhanced in this unit, as you analyse mathematical functions and summarise their properties using the calculus knowledge you learn. Other skills developed

are visualisation, information sourcing, and the ability to receive and give information through sorting. These all contribute towards improving your numeracy literacy.

You can also enhance your meta-skills in self-management, social intelligence and innovation.

On completing this unit, you will have the necessary mathematical skills to progress in data science and machine learning-related Higher National units, such as:

- Machine Learning at SCQF level 8 or 9
- Artificial Intelligence at SCQF level 8
- Big Data at SCQF level 8

Administrative information

Published: June 2023 (version 1.0)

Superclass: RB

History of changes

Version	Description of change	Date

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