Higher National Unit Specification

General information

Unit title: Data Science Project (SCQF level 8)

Unit code: J4Y2 35

Superclass: CA

Publication date: December 2020

Source: Scottish Qualifications Authority

Version: 02

Unit purpose

This **specialist** unit is designed to show learners how to undertake a data science project, working in a team. The selection of the project is left to the team but must use real data to provide insights into real problems. It is presumed that learners already possess data analysis skills and wish to apply these to a real problem in a team context.

Learners will carry out an analysis of a large and complex dataset, in collaboration with other learners, to provide insights, using a range of collaboration, project management and data analysis software. The analysis may be descriptive or predictive and will involve the application of statistical methods to identify relationships within the dataset. Visualisation and communication tools will be used to present findings.

Learners will gain an appreciation of the data science process and the roles individuals play within that process. They will also use collaboration, project management and source control software to co-ordinate, control and manage their work.

On completion of this unit, learners may progress to appropriate units at SCQF level 9.

Outcomes

On successful completion of the unit the learner will be able to:

1 Collaborate with a team in a data science project.
2 Carry out an analysis of a large, complex dataset.
3 Communicate the results of an analysis.

Higher National Unit Specification: General information (cont)

Unit title: Data Science Project (SCQF level 8)

Credit points and level

1 Higher National Unit credit at Scottish Credit and Qualifications Framework (SCQF) level 8: (8 SCQF credit points at SCQF level 8)

Recommended entry to the unit

Learners must possess well developed analysis and communication skills before undertaking this unit since this unit seeks to integrate these skills in the context of a real project. Knowledge and experience of data analysis and visualisation software is assumed. This may be evidenced by possession of J4Y6 35 *Working with Data* at SCQF level 7 and J4Y7 35 *Communicating with Data* at SCQF level 8.

Core Skills

Achievement of this Unit gives automatic certification of the following Core Skills component:

Core Skill component	Critical Thinking at SCQF level 6 Working Co-operatively with Others at SCQF level 6
Complete Core Skill	Information and Communication Technology at SCQF level 6

There are also opportunities to develop aspects of Core Skills which are highlighted in the support notes of this unit specification.

Context for delivery

If this unit is delivered as part of a group award, it is recommended that it should be taught and assessed within the subject area of the group award to which it contributes.

This unit is part of the Professional Development Award (PDA) in Data Science at SCQF level 8. It may be delivered after the other mandatory units in this award (*Working with Data* and *Communicating with Data*).

This unit requires learners to work with others in a team. Learners are expected to already possess data analysis and communication skills. The focus of this unit is working with others to undertake an end-to-end project. The scale and scope of the project should be sufficient to provide every learner in the team with a valid role.

The project should involve real business problems and involve the use of real data.

Equality and inclusion

This unit specification has been designed to ensure that there are no unnecessary barriers to learning or assessment. The individual needs of learners should be taken into account when planning learning experiences, selecting assessment methods or considering alternative evidence.

Further advice can be found on our website **www.sqa.org.uk/assessmentarrangements**.

Higher National Unit Specification: Statement of standards

Unit title: Data Science Project (SCQF level 8)

Acceptable performance in this unit will be the satisfactory achievement of the standards set out in this part of the unit specification. All sections of the statement of standards are mandatory and cannot be altered without reference to SQA.

Where evidence for outcomes is assessed on a sample basis, the whole of the content listed in the knowledge and/or skills section must be taught and available for assessment. Learners should not know in advance the items on which they will be assessed and different items should be sampled on each assessment occasion.

Outcome 1

Collaborate with a team in a data science project.

Knowledge and/or skills

- Stages in a data science project
- Roles in a data science project team
- Collaboration tools
- Project management methods and software
- Source control systems
- Business understanding including gathering and synthesising requirements for the project
- Data requirements including data sources
- Ethical and legal considerations including data bias

Outcome 2

Carry out an analysis of a large, complex dataset.

Knowledge and/or skills

- Project planning and allocation of tasks and resources
- Quality control and quality assurance
- Data transformations including data cleaning
- Data architecture and data modelling
- Exploratory data analysis
- Statistical analyses to identify patterns, trends and relationships in the dataset
- Interpreting the results of statistical analysis

Higher National Unit Specification: Statement of standards (cont)

Unit title: Data Science Project (SCQF level 8)

Outcome 3

Communicate the results of an analysis.

Knowledge and/or skills

- Key performance indicators
- Selection of visualisations
- Dashboards and reports
- Presentation skills
- Tailoring communication to a specific audience
- Communicating the findings of the project

Evidence requirements for this unit

Learners will need to provide evidence to demonstrate their knowledge and/or skills across all outcomes. The evidence requirements for this unit will consist of two types of evidence:

1 Product evidence
2 Performance evidence

This evidence will be produced by a team, comprising two or more learners. Learners will undertake at least two roles during the project; all learners must contribute to the analysis and the presentation (see below). The roles must be significant and equitable.

The **product evidence** relates to outcome 1 and outcome 2. The product evidence will have two parts.

1 Description of the project, which references all of the knowledge and skills specified in Outcome 1 and some of the knowledge and skills specified in outcome 2.
2 Results of the analysis, which references the relevant knowledge and skills in outcome 2.

These two products must collectively address all of the knowledge and skills defined in outcome 1 and outcome 2. A single description and a single analysis will be produced by each team.

The description will be written and will be contextualised. For example, the roles within a project team will describe the roles within the learner's specific team; project management methods and software will describe the specific methods and software used by the learner's team.

The results of the analysis may be presented in a number of digital formats, including a dashboard and/or a report.

Higher National Unit Specification: Statement of standards (cont)

Unit title: Data Science Project (SCQF level 8)

The scope of the project and the scale of the dataset should reflect the level of this unit, and the group nature of the project. The dataset will consist of real data, which may be familiar to the learner. The data must be sourced by the learners. It will consist of a large dataset, created from at least two sources, which will require significant transformation prior to analysis.

The **performance evidence** will relate to outcome 3. It will take the form of a presentation of the results of their analysis. The presentation may be done in person or remotely. The presentation must provide evidence for each of the knowledge and skills statements in outcome 3. The performance evidence must be observed and recorded. Every learner in the team must make a contribution to the creation and delivery of the presentation.

Learners may be given guidance throughout the project; the details of the project should be left to learners, including the planning of the project and the allocation of work.

This evidence may be produced over an extended period of time, under loosely controlled conditions. When evidence is produced in loosely controlled conditions it must be authenticated. The *Guide to assessment* provides further advice on methods of authentication.

The SCQF level of this unit provides additional context on the nature of the required evidence and the associated standards. The level descriptors should be used (explicitly or implicitly) when making judgements about the evidence.

The *Guidelines on approaches to assessment* (see the support notes section of this specification) provides specific examples of instruments of assessment.

Higher National Unit Support Notes

Unit title: Data Science Project (SCQF level 8)

Unit support notes are offered as guidance and are not mandatory.

While the exact time allocated to this unit is at the discretion of the centre, the notional design length is 40 hours.

Guidance on the content and context for this unit

This unit is designed to permits learners to apply their existing knowledge and skills in data analysis to a realistic group project. The focus of the unit is group work and the application of their existing knowledge and skills to a specific business problem.

The following guidance does not seek to explain each knowledge/skills statement, which is left to the professionalism of the teacher. It seeks to clarify the Statement of Standards where it is potentially ambiguous. It also focuses on non-apparent teaching and learning issues that may be over-looked, or not emphasised, during unit delivery. As such, it is not representative of the relative importance of each knowledge/skill.

Teams should consist of at least two learners. To improve realism, it is recommended that the selection of team members is done by teachers (and not self-selected by learners). Where possible, larger teams, consisting of three or more members, should be formed.

There is not time within this unit for teaching or learning technical skills in data analysis. The time should be spent acquiring collaboration, team working and project management skills.

The project should relate to a real business problem and involve the analysis of real data. Datasets should be realistic in terms of their size and complexity. It is anticipated that the analysis will be descriptive, rather than predictive, although diagnostic analyses should be within the capabilities of learners.

Learners are required to play a number of roles during the project. For the purposes of generating evidence, every learner must be involved in the analysis and presentation of results. Teachers should approve the designated roles before commencement of the analysis to ensure an equitable distribution of work between learners. Teachers may wish to serve as team leader for each team or permit a learner to carry out this role. Some form of organisation will be required for decision making.

A key aspect of this unit is collaboration, team working and project management. A significant amount of time may be required for learners to acquire these skills. It is recommended that learners have the opportunity to use contemporary collaboration and project management tools. Learners are not required to use sophisticated project management software but should have the opportunity to use digital planning software.

Higher National Unit Support Notes (cont)

Unit title: Data Science Project (SCQF level 8)

Guidance on approaches to delivery of this unit

This is a project-based unit, which aims to apply learners' existing knowledge and skills in the context of a data science project.

A suggested distribution of time, across the outcomes, is:

Outcome 1: 10 hours
Outcome 2: 20 hours
Outcome 3: 10 hours

Before commencing teamwork, learners should be familiarised with the tools and techniques for team working, collaboration and project management (outcome 1). Platforms such as Microsoft Teams™ or Slack™ should be explored.

It is recommended that the composition of each team is determined by teachers (and not learners). A team leader should be identified. The team leader could be a teacher (serving in this role for each team) or a designated learner.

Each team should be permitted to choose its own project. A selection of projects could be offered or teams could be permitted to come up with their own projects. Teachers should approve projects to ensure that they have sufficient scope and complexity.

Roles and responsibilities, within each team, will have to be carefully identified and assigned to each learner. This should be moderated by teachers. The contribution of each learner to the overall project is an important aspect of assessment.

Once development (project) work commences, team should operate autonomously. However, teachers should maintain oversight of progress.

Guidance on approaches to assessment of this unit

Evidence can be generated using different types of assessment. The following are suggestions only. There may be other methods that would be more suitable to learners.

Centres are reminded that prior verification of centre-devised assessments would help to ensure that the national standard is being met. Where learners experience a range of assessment methods, this helps them to develop different skills that should be transferable to work or further and higher education.

The assessment for this unit will have two parts. The first part is project work. The second part is the presentation of findings.

Each team should develop a project brief that defines the scope and scale of the analysis. This should be moderated by teachers. Once a brief has been agreed, each team should develop a project report (the description defined in evidence requirements) and produce the results of the analysis, which is likely to take the form of a dashboard and/or a report. Both products (the description and the analysis) could be assessed using a checklist.

Higher National Unit Support Notes (cont)

Unit title: Data Science Project (SCQF level 8)

Each team should develop a project brief that defines the scope and scale of the analysis. This should be moderated by teachers. Once a brief has been agreed, each team should develop a project report (the description defined in Evidence Requirements) and produce the results of the analysis, which is likely to take the form of a dashboard and/or a report. Both products (the description and the analysis) could be assessed using a checklist.

The second part involves learners presenting their results. This must be observed and recorded. Every member of the team must make a contribution to the creation of the presentation and its delivery. An observation checklist could be used to assess the performance.

Peer assessment could be considered as a component of the overall assessment scheme. Although the evidence will be assessed holistically (for the whole team), there is scope for team members to assess one another's contribution (on a satisfactory/unsatisfactory basis). This would be moderated by the teacher.

Formative assessment should be used at various stages in this unit. For example, for some learners, participating in a technical presentation will be a new (and potentially daunting) experience, and may require practice and refinement after feedback.

Opportunities for e-assessment

E-assessment may be appropriate for some assessments in this unit. By e-assessment we mean assessment which is supported by Information and Communication Technology (ICT), such as e-testing or the use of e-portfolios or social software. Centres which wish to use e-assessment must ensure that the national standard is applied to all learner evidence and that conditions of assessment as specified in the evidence requirements are met, regardless of the mode of gathering evidence. The most up-to-date guidance on the use of e-assessment to support SQA's qualifications is available at **www.sqa.org.uk/e-assessment**.

Opportunities for developing Core and other essential skills

The unit provides opportunities to develop some of the following Core Skills:

- Information and Communication Technology (ICT) at SCQF level 6
- Problem Solving at SCQF level 6
- Numeracy at SCQF level 6
- Working with Others at SCQF level 6

Learners are expected to make use of software for performing statistical analysis, which may include generic software, such as Excel™ (with appropriate add-ins), and/or dedicated software, such as SPSS or R. They will address several components of the Core Skill in *Information and Communication Technology* (ICT) in so doing.

Team working is a vital part of this unit and there will be many opportunities, throughout the unit, to demonstrate the Core Skill of *Working with Others*.

Higher National Unit Support Notes (cont)

Unit title: Data Science Project (SCQF level 8)

The Core Skill of Information and Communication Technology at SCQF level 6 is embedded in this unit. When a learner achieves the unit, their Core Skills profile will also be updated to include this Core Skill.

The Critical Thinking component of Problem Solving at SCQF level 6 and Working Co-Operatively with Others component of Working with Others at SCQF level 6 are embedded in this unit. When a learner achieves these units, their Core Skills profile will also be updated to include these components

History of changes to unit

Version	Description of change	Date
02	Core Skills Components Critical Thinking and Working Co-Operatively with Others at SCQF level 6 embedded. Core Skill Information and Communication Technology at SCQF level 6 embedded	09/12/20

Unit template: June 2017

General information for learners

Unit title: Data Science Project (SCQF level 8)

This section will help you decide whether this is the unit for you by explaining what the unit is about, what you should know or be able to do before you start, what you will need to do during the unit and opportunities for further learning and employment.

This unit involves a group project where you will be solving a data science problem in a team. This unit should give you a chance to apply the skills you have developed in other units, gain practical experience of a large project, and teach you about tools for collaboration and project management.

You are expected to already possess data analysis skills before commencing this unit. You will work in a team to carry out a large-scale analysis (using real data) on a subject of your choice. Once you have carried out the analysis, you will present your findings to an audience.

The aim of the unit is to give you experience of carrying out a realistic data analysis project as part of a team. This unit includes the following.

- Stages in a data science project
- Roles and responsibilities in a data science project
- Collaboration tools
- Project management tools and techniques
- Source control systems
- Business requirements
- Data ethics
- Data transformation
- Statistical analysis
- Visualisations
- Dashboards and reports
- Presentation skills

The assessment will take the form of a project, which will involve you (along with other team members) in creating a project report, carrying out an analysis and then presenting your findings.

The unit will provide you with opportunities to develop aspects, at SCQF level 6, of the Core Skills in *Information and Communication Technology* (ICT), *Problem Solving*, *Numeracy* and *Working with Others*.

The Core Skill of Information and Communication Technology SCQF at level 6 is embedded in this unit. When a learner achieves the unit, their Core Skills profile will also be updated to include this Core Skill.

The Critical Thinking component of Problem Solving at SCQF level 6 and Working Co-operatively with Others component of Working with Others at SCQF level 6 are embedded in this unit. When a learner achieves these units, their Core Skills profile will also be updated to include these components.