

Skill Area	Evidential Requirements	Types of evidence
<p>A. Data Privacy and Stewardship</p> <p>This skill relates to the security and protection of data, including design, creation, storage, distribution and associated risk.</p>	<p>1. Ensuring the protection of personal and sensitive data.</p>	<p>i. Assess risks and enact data protection policies and procedures.</p> <p>ii. Ensure safe and secure management of sensitive data, models and infrastructures.</p> <p>iii. Apply appropriate data controls, such as encryption, (pseudo)anonymisation, and synthetic data.</p> <p>iv. Risk management around environment and infrastructure.</p>
	<p>2. Managing sensitive data.</p>	<p>i. Act with integrity, giving due regard to legal and regulatory requirements.</p> <p>ii. Be aware of the actions that should be taken to respond to potential data loss in line with organisational, legal and regulatory procedures.</p>
	<p>3. Data stewardship and standards.</p>	<p>i. Incorporates the FAIR Guiding Principles for scientific data management and stewardship into practices, where appropriate and practicable.</p> <p>ii. Identify opportunities for efficient and creative reuse of data.</p> <p>iii. Understand the relationship between technical standards and regulation/governance, and their benefits for interoperability and knowledge sharing.</p>
<p>B. Definition, acquisition, engineering, architecture, storage and curation.</p> <p>This skill relates to the collection, manipulation and secure storage of data,</p>	<p>1. Data collection and management.</p>	<p>i. Source and access data appropriate for the problem.</p> <p>ii. Critically analyse the availability of appropriate data and resources to meet project requirements.</p> <p>iii. Critically evaluate and synthesise data.</p> <p>iv. Ensure data provenance processes are followed.</p> <p>v. Identify data characteristics (volume, velocity and variety).</p> <p>vi. Identify infrastructure requirements for data storage and analysis.</p>

the application of data management and analytical techniques.		vii. Show familiarity or experience with tabular and non-tabular data (e.g. unstructured and streaming data).
	2. Data engineering.	<ul style="list-style-type: none"> i. Source and access data appropriate for the problem. ii. Construct data sets, potentially drawing from multiple disparate sources using data linkage. iii. Perform data profiling and characterisation to understand the surface properties of the data. iv. Handle missing data, through principled inclusion/exclusion criteria and imputation methods. v. Take a systematic approach to data curation and the application of data quality controls. vi. Identify the most appropriate solutions (e.g. cloud vs on-premise) in response to business and project needs.
	3. Deployment.	<ul style="list-style-type: none"> i. Plan the deployment of data products with their end-users. ii. Develop monitoring and maintenance processes. iii. Deliver secure, stable and scalable data products to meet the needs of the organisation, e.g. Application Programming Interfaces (APIs), derivative datasets, dashboards, reports and do so according to modern software development best practices. iv. Design and deliver data products that meet appropriate accessibility standards for their users.
C. Problem definition and communication with stakeholders This skill is about engaging stakeholders,	1. Problem definition.	<ul style="list-style-type: none"> i. Identify and elicit project requirements. ii. Determine success criteria and frame these in the context of the business. iii. Clearly articulate the problem statement. iv. Identify and critically evaluate assumptions. v. Recognise and quantify biases and identify solutions to manage and mitigate these.

demonstrating the ability to clearly define a problem and agree on solutions.		vi. Assess risk. vii. Demonstrate sector/domain knowledge and knowledge of how data science can deliver value to these sectors/domains.
	2. Relationship management.	i. Communicate in an effective manner for diverse audiences, including technical colleagues, subject matter experts and leadership. ii. Effectively manage the expectations of diverse stakeholders with conflicting priorities to mediate equitable solutions. iii. Use relevant communication techniques (written, oral or visual), appropriate for the audience. iv. Build appropriate and effective business relationships. v. Show experience in human factors considerations with respect to data-driven solutions.
D. Problem solving, analysis, statistical modelling, visualisation. This skill relates to the identification and presentation of solutions using a range of methods, tools and techniques, demonstrating the ability to analyse a problem and define and present options.	1. Identifying and applying technical solutions and project management approaches.	i. Identify viable solutions based on requirements and data available. ii. Identify and provide guidance to technical and non-technical stakeholders on the most appropriate solution. iii. Apply appropriate technical and project management methodologies appropriate for the organisation and project.
	2. Data preparation and feature modelling.	i. Identify appropriate solutions, including statistical and machine learning approaches. ii. Identify and evaluate appropriate evaluation metrics, including computational performance and accuracy. iii. manipulate data with due regard for differences in characteristics. iv. Creation and evaluation of new data features.

	3. Data Analysis and Model building	<ul style="list-style-type: none"> i. Apply appropriate solutions, including statistical and machine learning approaches. Demonstrate competence in a modern programming language. ii. Use appropriate analysis platforms and tools. iii. Adopt a systematic approach to exploratory data analysis to embrace and manage ambiguity and uncertainty. iv. Critically analyse data and analytical results. v. Adopt appropriate methods to visualise data and communicate complex findings.
--	-------------------------------------	---

Section E is a cross-cutting consideration to be evidenced throughout.

<p>E. Evaluation and Reflection</p> <p>This skill is about reflecting on performance and outcomes, identifying development needs and applying important principles associated with ethics and sustainability.</p> <p>Note: when completing your evidence for this Skill Area you can refer to evidence provided in the Skill Areas A-D, together with ensuring that ethical evaluation is reflected throughout Skill Areas A-D.</p>	1. Project evaluation.	<ul style="list-style-type: none"> i. Ongoing monitoring of project performance and outcomes. ii. Identify and feed forward lessons learned. iii. Participate in and lead collaborative project evaluations, e.g. retrospectives.
	2. Ethical behaviour.	<ul style="list-style-type: none"> i. Identify and manage the risks of erroneous and biased data. ii. Act with integrity with respect to legal and regulatory requirements. iii. Uphold principles of ethical and safe use of data and AI technologies. iv. Implement data use procedures to ensure sensitive data is only used for its agreed purpose. v. Implement data retention strategies in line with regulatory and legal requirements.
	3. Sustainability and best Practices.	<ul style="list-style-type: none"> i. Evidence of incorporating the principles of open science and/or reproducible research within the organisation, and perhaps beyond. ii. Competence programmatic approaches to undertaking data science work.

		<ul style="list-style-type: none"> iii. Apply the scientific method in delivering solutions iv. Ensure high technical standards, in line with software development best practices; for example, software testing, version control, Continuous Integration and Continuous Delivery. v. Apply automation to promote reproducibility analyses.
	<p>4. Reflective practice and ongoing development.</p>	<ul style="list-style-type: none"> i. Learn from experience through self-assessment of one's own responses to practice situations. ii. Identify learning opportunities to maintain knowledge and skills in the relevant area of data science. iii. Take ownership of ongoing professional development. iv. Contribute to knowledge-sharing across their organisation and/or the wider community. v. Contribute to the management and empowerment of the broader team. vi. Engage with the latest developments across industry and academia and incorporate these into solutions.