

## **Data Scientist Integrated Degree Apprenticeship – Level 6**

### **Occupational Profile**

Data science is a broad and fast-moving multi-disciplinary field spanning maths and statistics, software engineering and communications. Data Scientists blend experience and knowledge from a wide range of fields and organisations, and continuously seek to expand their range of technical skills. They can work in any sector, public or private, and will often work as part of a multidisciplinary team with Data Architects, Data Engineers and Analysts.

Data Scientists find information in diverse datasets to address complex problems and improve organisational processes. They explore and visualise data of all kinds, presenting the 'story' of the data in a meaningful way to a range of technical and non-technical audiences. They support strategic and operational decision making through sourcing, accessing, and manipulating data and engineering data processes. They use the insights gathered about the data they have analysed to enable change.

Data Scientists are dynamic and adaptable, addressing varied problems with varied techniques. They actively explore innovative ways to use existing and new statistical, algorithmic and machine learning tools and techniques, to find significant and valuable patterns in data and transform this into information for their organisation. They seek out new sources of data, and understand how to combine datasets to increase their value. They build credible statistical models from the data, use good coding practices, and maintain data, tools and processes for the operationalisation of workloads.

Data Scientists have an impact at a strategic and operational level by building and maintaining strong relationships with key stakeholders, subject matter experts and colleagues at all levels. They engage with the wider data science community to share ideas, techniques and experiences.

### **Typical Roles**

Data Scientist, Informatics

### **Entry Requirements**

Individual employers will set the selection criteria, but this is likely to include 3 A levels, including 1 STEM subject, a level 3 apprenticeship or other relevant qualifications. Other relevant or prior experience may also be considered by the employer.

**Knowledge.** A Data Scientist must understand:

1. The context and history of data science and the data science community and its relationship to computer science, statistics and software engineering.
2. How data science operates within the context of data governance, data security, ethics, and communications. How data science can be applied to improve an organisation's processes, operations, and outputs.

3. How data can be used systematically in an organisation, including
  - 3.1. Data processing and storage, including on-premise and cloud technologies.
  - 3.2. Data systems, including traditional data warehousing and modern approaches to data systems.
  - 3.3. Data driven decision making and the good use of evidence and analytics in making choices and decisions.
  
4. How to design, implement and optimise algorithms, as prototypes, and at production scale using:
  - 4.1. Traditional statistical and mathematical methods
  - 4.2. Advanced modern analytics and machine learning practices
  - 4.3. Simulations, optimisation, computer vision and Natural Language Processing (NLP).
  - 4.4. An awareness of the computing and organisational resource constraints and trade-offs involved in selecting models, algorithms and tools.
  - 4.5. Appropriate development standards, including good coding practice, testing, & source control.
  
5. The data landscape: knowing how to critically analyse, interpret and evaluate complex information through the understanding of:
  - 5.1. Database technologies: including relational, NoSQL and graph databases.
  - 5.2. Sources of data including but not exclusive to files, operational systems, data warehouses, and services, open data, government data, news and social media.
  - 5.3. Data formats, structures and data delivery methods.
  - 5.4. Common patterns in real-world data

**Skills.** A Data Scientist is able to:

1. Collaborate with colleagues to gather requirements, understand data sources, and determine appropriate analytical techniques and tools. Determine the value of analysis to the organisation and how to contribute with impact.
  
2. Identify, clarify and structure problems their organisation faces, and reformulate them into data science problems. Map solutions to problems and make decisions in context by seeking feedback from stakeholders.
  
3. Perform data processing and use preparation tools and techniques to access, explore, profile, pipeline, transform and store data, and apply governance (quality control, security, privacy) to data. Scale a system up/out and data exploration.
  
4. Apply statistical analyses, machine learning, optimisation, and simulations, using the appropriate techniques for the problem. Use statistical testing to validate results. Use analytical results and models to inform and improve organisational outcomes.
  
5. Apply scientific methodology through experiment design, measurement, hypothesis testing and delivery of results.
  
6. Implement solutions, using appropriate architectures and design patterns. Cloud vs. on-premise deployment. Assess value for money through return on investment (ROI) calculations.

7. Present, communicate and disseminate data science products effectively and with high impact through usage of creative storytelling and tailoring the message for the audience. Use the most appropriate medium, such as technical writing, reporting and visualisation to tell compelling and actionable stories relevant for organisation goals.

8. Use a range of programming languages and tools for data science work. Identifies and uses the appropriate language or tool for data manipulation, analysis, visualisation, system integration. Use test driven development and Agile methods. Use appropriate data structures and algorithms for the problem. Work in accordance with common software development standards, including security, accessibility, code quality and version control..

9. Develop and maintain collaborative relationships, using methods of organisational empathy (human, organisation and technical) and building relationships through active listening and trust development.

10. Use project delivery techniques and tools appropriate to their data science project and organisation. Plan, organise and manage resources to successfully run a small data science project, achieve organisational goals and enable effective change.

**Behaviours.** A Data Scientist demonstrates:

1. A curious and questioning approach. Shows curiosity and a desire to explore new questions, opportunities, data, and techniques, being relentless and creative in their approach to solutions.
2. Empathy and positive engagement to enable working and collaborating in teams. Champions and highlights ethics and diversity in data work.
3. Pragmatism in the face of real-world scenarios and responsiveness to organisational timescales.
4. Thinking at a strategic level: considering problems in the context of the organisation and organisational goals.
5. A scientific/hypothesis-driven approach to work.

**Duration:** Typically 36 months

**Qualifications:** BSc in Data Science

**English and maths:** Apprentices without level 2 English and maths will need to achieve this level prior to taking the end-point assessment'.

**Level:** This is a level 6 apprenticeship

**Review Date:** After three years