

Level 5 - 20 months + EPA

Data Engineer

There is nothing standard about the new apprenticeship Standards!

In 2024 the modern digital apprenticeship range expanded yet further, to provide apprenticeships for even more roles within digital departments, and to provide better vocational training options for employers. Apprenticeships are an amazing solution for new career starts, professional upskilling or changes in career direction.

Programme Overview:

A Data Engineer could be found in any organisation that works with large data sets.

The broad purpose of the role is to build systems that collect, manage, and convert data into usable information (often called data products) for data science, artificial intelligence, analysis and business intelligence.

With the right data products established, a business has all the data they require on tap, all the time, to support decisions, optimise performance and enhance governance.

At the start of the programme, we teach apprentices how to create a data product using SQL and Python so they can provide value to their business early. Apprentices are also taught how to work safely with data and how to collaborate effectively with stakeholders.

Apprentices will learn how to apply data quality and validation; software development principles and automation to the data product they created at the start of the programme.

We consider data architecture and evaluate the advantages and disadvantages of different organisations' architectures, highlighting what a modern data stack looks like.

Apprentices will also be equipped with the necessary skills to successfully cost, design, deliver and maintain a data product. Skills include how to successfully manage projects, technology and services.

Our programme teaches the most up-to-date tools, technologies and ways of working. We also equip apprentices with their own strategies to identify new relevant developments in order to continuously guide improvements for their organisation.

Who is it for?

Relevant job titles include (but are not limited to):

- Data Engineer
- Big Data Engineer
- Data Engineering Specialist
- ETL/ELT Engineer

Entry Requirements:

Entry requirements exist for all funded Further Education programmes. These ensure the value, gain and success of the programme. LearnTech will conduct the processes with employers and prospective apprentices to determine correct funding eligibility.

Job role eligibility (known as Competency Role Map):

The job role must contain opportunity for an apprentice to practice the content set out in the Level 5 Data Engineer apprenticeship Standard in order to achieve vocational competency.

Apprentices must have the opportunity to practice the knowledge taught in training sessions in order to convert new knowledge in to sustainable skills applied in the workplace. Where required, we can work with employers to ensure the job role meets the essential criteria.

If you have any questions about this apprenticeship Standard or would like to book an information session, please contact our team: info@learn-tech.co.uk or **+44(0)330 380 0249**

Initial assessment of existing knowledge and skills:

A candidate must stand to gain significant knowledge and skills from the programme. If the apprenticeship is too advanced, or if they already know a significant amount of the knowledge and skills the apprenticeship would provide, then they may not be eligible for the funding.

LearnTech will review existing qualifications, knowledge and skills to determine if the prospective apprentice will benefit from the proposed apprenticeship such that it meets the funding criteria. In most instances this is straightforward, however in some instances funding is adjusted in order to fund only the parts of an apprenticeship that would be relevant. LearnTech provides the assessment for these scenarios.

The Level 5 Data Engineer is highly technical, so whilst employers can select their own entry criteria, they should include; at least 5 GCSEs at grade 4 or above, including English and maths. English and maths can be included in the apprenticeship, if required.

In many cases, this type of apprenticeship can demand a higher capability of English and maths than is taught at GCSE or A-Level. For example, advanced report writing, budgeting, complex structured explanations and/or advanced formulae and statistics. LearnTech will provide both functional and advanced English and maths diagnostics and teaching to ensure each apprentice is fully supported in these areas.

Live, Virtual Delivery Model:

Our employer-designed delivery model has been developed over time using extensive employer feedback. The core aims are highly effective training for apprentices and as much flexibility as possible for employers.

The only weekly fixed element of training is a 1.5-2 hour training session, with the remaining learning being agreed between the employer and apprentice around workload and responsibilities.

The live classrooms are held via Microsoft Teams. This software provides a full suite of educational tools, including everything you would find in a conventional classroom and more e.g. live open interactions, private breakout rooms, notes, questions, and interactive illustration boards. These sessions are supported with practical labs, together with numerous other tools, selected specifically for their effectiveness in the virtual environment.

Whilst maximum flexibility is offered, if an employer prefers a day-release arrangement, it can be agreed that an entire day is set aside for the training.

All other requirements, including monthly one-to-one mentoring sessions, are arranged between the apprentice and their dedicated skills mentor. Quarterly progress reviews are arranged conveniently between the apprentice, their manager and their mentor.

Programme Duration:

This apprenticeship is delivered over 20 months for full-time employees. For part-time employees, the term is extended depending on contracted hours.

End Point Assessment (EPA):

An allowance of three months is required for the EPA process, which begins once all the training is completed.

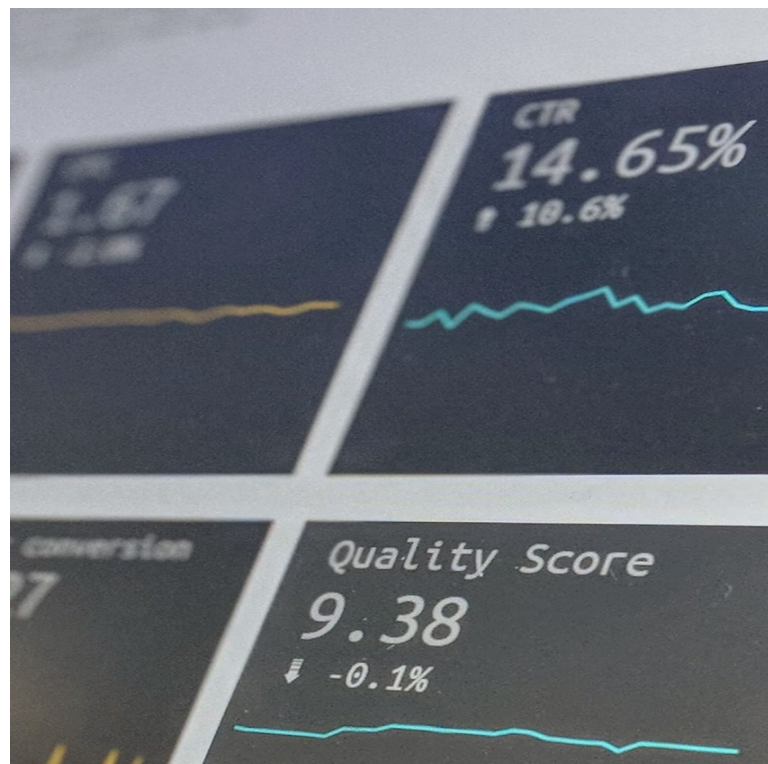
Successful apprentices will obtain a Pass, Merit or Distinction grade. The grade is determined by the EPA process. Like almost all examinations, a mock will take place before the final assessment.

Once all components of the apprenticeship have been achieved, including the mock, a final review is conducted to ensure everything has been covered, this is called gateway. Then the apprentice will undergo their EPA.

Whilst other qualifications that can be embedded in an apprenticeship, the most important and valuable goal is completing the apprenticeship Standard, which is certified by the Institute of Apprenticeships.

The EPA for this programme consists of:

1. Portfolio of workplace evidence and professional discussion
2. Project with report, presentation and questioning



If you have any questions about this apprenticeship Standard or would like to book an information session, please contact our team: info@learn-tech.co.uk or +44(0)330 380 0249

Programme Structure:

Apprentices are taught principles, techniques and technologies. The education incorporates knowledge, skills and behaviours, as well as self-management and an objective led approach.

Technical Competencies:

- Select and use appropriate tools to process data, in any format, such as structured and unstructured data and in any mode of delivery, such as streaming or batching. Adapt to legacy systems as required
- Ensure resilience is built into data products incorporating business continuity and disaster recovery plans, and document change management to limit service outages
- Analyse requirements, research scope and options and present recommendations for solutions to stakeholders
- Support the implementation of prototype or proof-of-concept data products within a production environment
- Maintain data solutions as continually evolving products, to service the organisation, user or client requirements. Collaborate with technical support teams and stakeholders from implementation to management
- Work within compliance and contribute towards data governance, organisational policies, standards and guidelines for data engineering
- Monitor the data system to meet service requirements to enable solutions such as data analysis, dashboards, data products, pipelines, and storage solutions
- Follow and develop processes to monitor and optimise the performance of the availability, management and performance of data products
- Follow and develop methodologies for moving data from one system to another for storage and further handling
- Follow and develop data normalisation principles and the advantages they achieve in databases for data protection, redundancy and inconsistent dependency
- Follow and develop frameworks for data quality, covering dimensions such as accuracy, completeness, consistency, timeliness and accessibility
- Mitigate inherent risks of data, such as incomplete data, ethical data sources and how to ensure data quality
- Utilise software development principles for data products, including debugging, version control, and testing
- Utilise principles of sustainable data products and organisational responsibilities for environmental social governance
- Deploy various approaches for new data pipelines and processes automation
- Build data products that comply with regulatory requirements
- Follow concepts of data governance, including regulatory requirements, privacy, security, and quality control, including legislation and its application to the safe use of data
- Follow data and information security standards, ethical practices, policies and procedures relevant to data management activities, such as data lineage and metadata management
- Cost and build systems whilst ensuring that organisational strategies for sustainable, net zero technologies are considered
- Understands implications of financial, strategic and compliance regarding to security, scalability, compliance and cost of local, remote or distributed solutions
- Considers on-demand Cloud computing platforms in a public or private environment, such as Amazon AWS, Google Cloud, Hadoop, IBM Cloud, Salesforce and Microsoft Azure
- Understands data warehousing principles, including techniques such as star schemas, data lakes and data marts
- Understands principles of data, including open and public data, administrative data, and research data, including the value of external data sources that can be used to enrich internal data.
- Understands how businesses use direct data acquisition to support or augment business operations
- Can combine disparate data sources to deliver value
- Uses streaming, batching and on-demand services to move data from one location to another
- Differentiates structured, semi-structured and unstructured data
- Uses a variety data engineering tools and applications
- Follows policies and strategies to ensure business continuity for operations, particularly in relation to data provision
- Support and respond to incidents through the application of technology and service management best practice, including configuration, change and incident management
- Undertakes analysis and root cause investigation
- Evaluates prototypes and takes them to implementation within a production environment
- Follows the lifecycle of implementing data solutions in a business, from scoping, through prototyping, development, production and continuous improvement
- Utilises data development frameworks and approved organisational architectures
- Utilises principles of descriptive, predictive and prescriptive analytics
- Follows continuous improvement methodologies, including how to: capture good practice and lessons learned
- Follows strategies for keeping up-to-date with new ways of working and technological developments in data science, data engineering and Artificial Intelligence
- Uses methods and techniques to best communicate messages to meet the needs of the audience

If you have any questions about this apprenticeship Standard or would like to book an information session, please contact our team: info@learn-tech.co.uk or **+44(0)330 380 0249**

Programme Structure:

Skills:

- Collate, evaluate and refine user requirements to design the data product
- Collate, evaluate and refine business requirements including cost, resourcing and accessibility to design the data product
- Design data products to serve multiple needs which incorporate scalability, efficiency, and security
- Automate data pipelines such as batch, real-time, on-demand and other processes using programming languages and data integration platforms with graphical user interfaces
- Produce and maintain technical documentation explaining the data product, that meets organisational, technical and non-technical user requirements, retaining critical information
- Systematically clean, validate, and describe data at all stages of extract, transform, load (ETL) and extract, load, transform (ELT)
- Work with different types of data stores, such as SQL, NoSQL and distributed file system
- Identify and troubleshoot issues with data processing pipelines
- Query and manipulate data using tools and programming, such as SQL and Python
- Manage database access and implement automated validation checks
- Communicate downtime and issues with database access to stakeholders to mitigate the operational impact of unforeseen issues
- Evaluate opportunities to extract value from existing data products through further development, considering costs, environmental impact and potential operational benefits
- Maintain a working knowledge of data use cases within organisations
- Use data systems securely to meet requirements and in-line with organisational procedures and legislation
- Identify new tools and technologies and recommend potential opportunities for use in own department or organisation
- Optimise data ingestion processes by making use of appropriate data ingestion frameworks such as batch, streaming and on-demand
- Develop algorithms and processes to extract structured data from unstructured sources
- Apply and advocate for software development best practice when working with other data professionals throughout the business. Contribute to standards and ways of working that support software development principles
- Develop simple forecasts and monitoring tools to anticipate or respond immediately to outages and incidents
- Identify and escalate risks with suggested mitigation/resolutions, as appropriate
- Investigate and respond to incidents, identifying the root cause and resolution with internal and external stakeholders
- Identify and remediate technical debt, assess for updates and obsolescence as part of continuous improvement
- Develop, maintain collaborative relationships using adaptive business methodology with stakeholders, such as business users, data scientists, data analysts and business intelligence teams
- Present, communicate, and disseminate messages about the data product, tailoring the message and medium to the needs of the audience
- Evaluate the strengths and weaknesses of prototype data products and how these integrate within an organisation's overarching data infrastructure
- Assess and identify gaps in existing tools and Identify data quality metrics and track them to ensure the quality, accuracy and reliability of the data product
- Selects and apply sustainable solutions to contribute to net zero and environmental strategies across the various stages of product and service delivery
- Horizon scanning to identify new technologies that offer increased performance of data products
- Implement personal strategies to keep up to date with new technology and ways of working



If you have any questions about this apprenticeship Standard or would like to book an information session, please contact our team: info@learn-tech.co.uk or +44(0)330 380 0249

Level 5 - 20 months + EPA Data Engineer



Behavioural Development Embedded:

- Maintain knowledge of technological developments, trends and innovation from industry sources
- Acts proactively and takes ownership, adapting positively to changing work priorities and ensuring deadlines are met
- Works collaboratively with stakeholders and colleagues, developing strong working relationships to achieve common goals. Support an inclusive culture and treat technical and non-technical colleagues and stakeholders with the same respect
- Quality focus that promotes continuous improvement utilising peer review techniques, innovation and creativity to the data system development process to improve processes and address business challenges
- Takes personal responsibility towards net zero and prioritises environmental sustainability outcomes in how they carry out the duties of their role
- Use initiative and innovation to problem solve and trouble shoot, providing creative solutions. Keeps abreast of developments in emerging, contemporary and advanced technologies to optimise sustainable data products and services

The designated mentor will support the employer and apprentice throughout the programme as a single point of contact for questions and queries. This includes additional support for portfolio and project preparation, along with any advice and guidance needed.

Progression:

Following on from this apprenticeship there are a variety of Level 6 and 7 apprenticeships that could be progressed on to, or alternatively a variety of professional qualifications that can be completed.

Next steps:

To configure an ideal apprenticeship we will meet with you virtually to discuss your requirements, present the options and collaborate to determine the best apprenticeships to meet your needs. We will provide ongoing support including:

- Recruitment of apprentices
- Quality assured Information Advice and Guidance
- Updates and information on legislation and funding
- Support and guidance for apprentice and employer throughout the apprenticeship
- Access to a comprehensive suite of resources and support material
- Industry specialist qualified trainers and mentors



If you have any questions about this apprenticeship Standard or would like to book an information session, please contact our team: info@learn-tech.co.uk or +44(0)330 380 0249