

Level 5 Data Engineer.

This apprenticeship transforms fragmented, unreliable data into a powerful business asset. Learners build automated data pipelines, enhance data quality, and integrate systems across cloud, distributed, and legacy platforms, creating scalable, future-proof data foundations. The result: faster reporting, fewer errors, lower costs, and smoother operations. Your organisation gains data specialists who unlock new opportunities, remove bottlenecks, and deliver clean, trusted insight that drives smarter decisions and sustained competitive advantage.

Who's it for?

This programme is ideal for analytical, technically minded individuals who work with large datasets and want to evolve into skilled data engineers. It's perfect for those who enjoy solving complex problems, building data systems, and ensuring data is accurate, accessible, and ready for data scientists and analysts to use.

Business impacts

- ✓ Maximise the value of business data by creating efficient data pipelines and models
- ✓ Respond quicker than competitors and with more agility with in-house prototypes and data solutions
- ✓ Improve decision-making with accurate, timely data that strengthens strategic planning
- ✓ Reduce operational costs with optimised data processes and infrastructures
- ✓ Build scalable data operations with cloud computing integration
- ✓ Confidently manage data securely and in compliance with regulations

Built for performance

- ✓ **AI embedded across every programme.** Equipping your workforce to develop job-ready skills that boost productivity and drive business impact.
- ✓ **Measurable business impact, not just learning outcomes.** Learners complete real workplace projects, supported by Coaches who help translate skills into demonstrable ROI.
- ✓ **Hands-on learning through virtual labs.** Guided virtual labs allow learners to safely experiment with real tools and technologies, mirroring real-world technical environments.
- ✓ **Relevant, tailored learning.** Data-sets and case studies can be adapted to ensure learning remains relevant to the individual and the sector or business area they work in.
- ✓ **Verified digital credentials employers can trust.** Digital badges and vendor-aligned certifications provide secure, shareable evidence of skills and progression.
- ✓ **Personalised continuous development.** Learners have the 'digital edge' – masterclasses, podcasts, guest speakers, in-person meet-ups, on-demand e-learning tech modules – all available according to their need and interests.

Programme overview.

Apprenticeship standard: Data Engineer

Cost: £19,000

Duration: 19 months

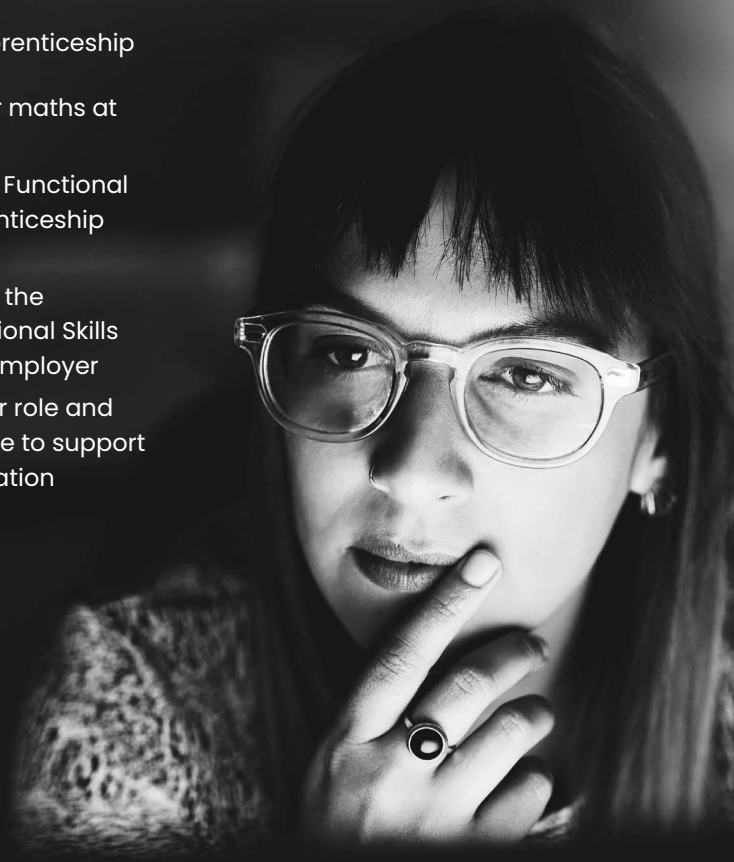
Entry requirements

As a minimum learners will need to have:

- A minimum of 720 hours of technology-related work experience (IT, software or engineering), or
- An A-Level in computer science or equivalent and have been in a technical role for three months or longer, or
- A Level 3 or 4 data/computing/engineering apprenticeship

For learners that do not have GCSE English and/or maths at grades 9 to 4 (A* to C):

- Learners aged 16–18 years must study and pass Functional Skills English and/or maths as part of the apprenticeship programme
- Learners aged 19 or above on the day they start the programme do not need to study or pass Functional Skills English and/or maths, unless required by their employer
- Learners must work with big data as part of their role and have access to the necessary tools and software to support data engineering activities across their organisation



Prepare for the challenges of tomorrow

Get ahead of evolving workforce skills needs

Every BPP apprentice has access to our exclusive Emerging Skills programme.

Comprised of four bespoke courses, the programme combines expertise from BPP, Microsoft and xUnlocked to give learners essential knowledge and skills in the rapidly emerging areas of AI, cyber security and sustainability.

- ✓ Available to all learners at no extra cost
- ✓ Accessible anytime, anywhere via our virtual learning platform
- ✓ Self-paced learning to fit into any busy schedule



Programme contents

Generative AI Fundamentals (Four modules)

Developed by BPP's expert data scientists, this course offers an introduction to working with generative AI effectively, safely and ethically.

Introduction to Sustainability (Six modules)

Developed in partnership with sustainability experts, xUnlocked, this course builds fundamental knowledge on sustainability and sustainable working practices.

Microsoft AI and Security Essentials (Seven modules)

This course begins with a core AI Essentials pathway, followed by a choice between AI Fundamentals or Security. Microsoft digital badges are awarded throughout, with the option to earn a recognised Microsoft Certification upon completion.

Cyber Security (Three modules)

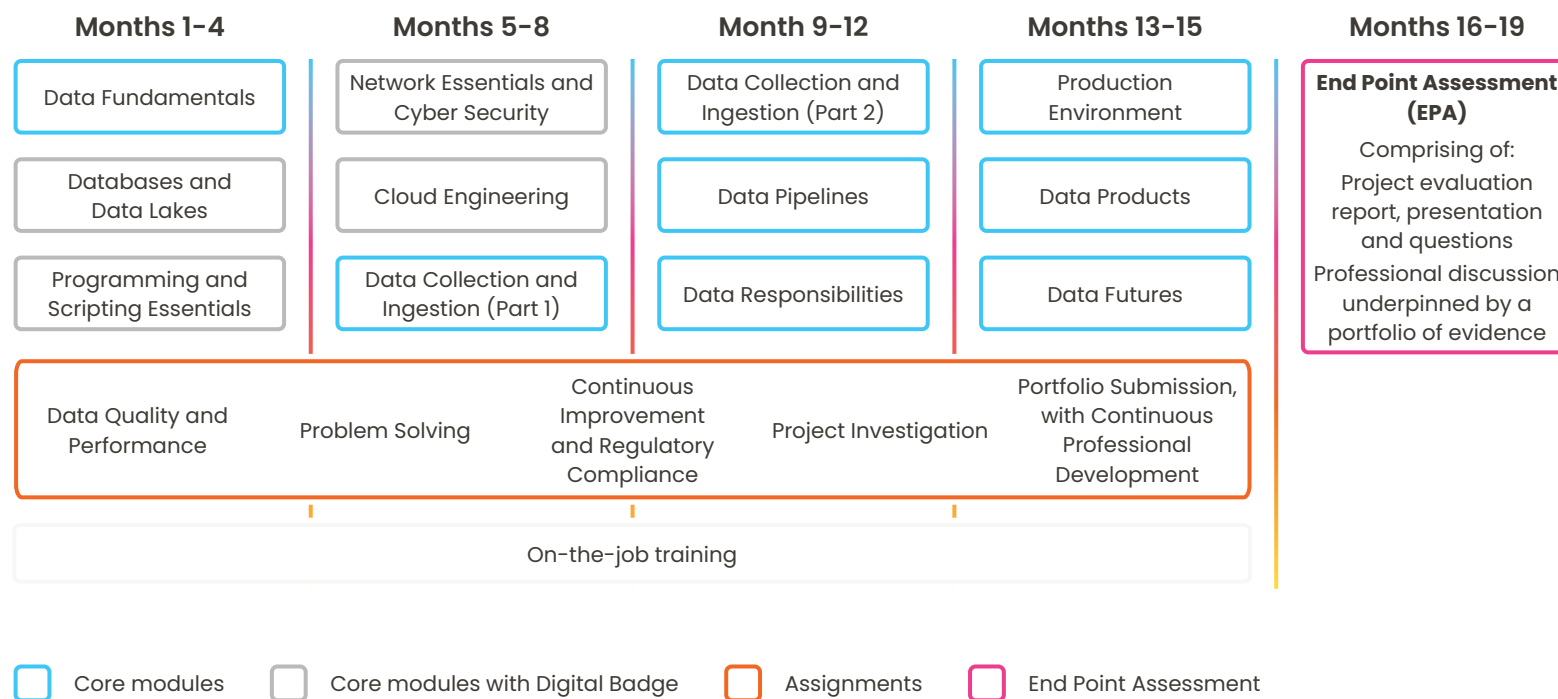
Designed by BPP's cyber experts, this course provides a clear and practical introduction to the importance of cyber security, the most common attack techniques everyone should be aware of, and the fast-changing digital threat landscape.

Level 5 Data Engineer.

Study mode

Online weekly: flexible learning that fits busy schedules, with two to three hours of online live lectures and seminars.

19-month programme (inc. EPA)



Apprenticeship standard

Data Engineer

Delivered by

BPP

Time commitment*

- 15 months** on programme
- 189 days** in online live training sessions comprised of:
 - 57 x three-hour webinars
 - 3 x six-hour hackathons
- 3-4 hours** guided self-study, per module, via our virtual learning platform
- 1 hour** performance coaching session, every month
- 1 hour** per week, or a half-day per month, dedicated to workplace shadowing (*recommended*)
- 1 hour** progress review, every eight weeks
- 1 hour** EPA preparation session
- 4 months** in End Point Assessment
- Off-the-job training** (*average eight hours a week*) Apprentices have dedicated learning time within their normal working hours to develop new skills directly linked to their apprenticeship. Training, mentoring, online learning and project work all count toward this time, and sit alongside their normal duties, not outside them.

*On average.

Programme modules.

Data Fundamentals

Learners discover how data drives business performance. They see how teams use information, how value is created, and what makes data trustworthy. They build confidence working with stakeholders and understanding the systems that keep data flowing, so they can support decisions and contribute to meaningful improvements across the organisation.

Databases and Data Lakes

Learners build the foundations that keep your organisation's data organised, accurate and accessible. They learn how to structure information, optimise performance and maintain secure, reliable storage. This helps ensure your teams and systems can work with clean data that supports better analysis and business insight.

Programming and Scripting Essentials

Learners use coding to streamline how data moves through your organisation. They turn manual steps into repeatable, reliable processes and keep data flows running smoothly. Their scripting, testing and debugging skills help reduce errors and support consistent, well-engineered data operations.

Network Essentials and Cyber Security

Learners grow their understanding of how data travels through networks and how to protect it along the way. They recognise risks, apply security principles and work within organisational standards to keep systems resilient. Their skills help maintain safe, uninterrupted access to the information your business relies on.

Cloud Engineering

Learners create cloud-based data services that are secure, scalable and ready for real-world use. They design and deploy systems that fit modern business needs, using flexible architectures that adapt as your organisation grows. Their work supports faster delivery and more dependable data products.

Data Collection and Ingestion (Part 1)

Learners bring data into your organisation in a controlled, high-quality way. They gather requirements, connect to the right sources and ensure incoming data is structured and dependable. Their approach supports smoother analytics and better decision-making from the start.

Data Collection and Ingestion (Part 2)

Learners handle fast, continuous flows of data using modern streaming tools. They build real-time pipelines that stay reliable under pressure, helping your organisation act quickly when situations change. Their skills keep event-driven systems stable, secure and responsive.

Data Pipelines

Learners build the end-to-end data routes that keep information moving cleanly through your organisation. They test, monitor and improve these flows to ensure data arrives accurate and on time, supporting dependable reporting, automation and analysis.

Data Responsibilities

Learners develop the judgement needed to handle data safely and responsibly. They follow governance, privacy and security expectations while working confidently with stakeholders. Their approach helps your organisation stay compliant and maintain trust in how data is managed and shared.

Programme modules.

Production Environment

Learners gain confidence working with systems that run in real business environments. They use DevOps and MLOps practices to release updates safely, test how components work together, and keep services healthy through good monitoring and logging. They stress-test systems and document what they build, helping your organisation run stable, well-supported data products.

Data Products

Learners build the skills to create and improve data products your teams will actually use. They work with AI and machine learning tools, understand business needs, and fix issues as they arise. They also visualise data clearly and track whether a product is delivering value, helping your organisation make better, faster decisions.

Data Futures

Learners develop the ability to judge when new technologies are worth adopting and how they fit into real business needs. They explore data ethics, sustainable tech choices, and emerging tools through horizon scanning. They learn to assess risks, spot opportunities and keep their skills current, helping your organisation innovate with confidence.

EPA: Project, Project Evaluation Report, and Presentation

Project

Learners complete a work-based project that tackles a real business problem or improvement opportunity, demonstrating the skills and knowledge gained throughout the programme.

Evaluation Report

Learners produce a 3,500-word report covering the project's scope, research, approach and findings.

Presentation

Learners present their project, summarising the scope, actions and outcomes.

Transformative *technologies*.

Designed to make real-world impact, our on-demand, self-paced asynchronous modules build awareness of transformative technologies, with coverage of artificial intelligence and an evolving roadmap to explore digital twins, edge computing and other emerging technologies.

- ✓ Available to all learners at no extra cost
- ✓ Bespoke to BPP, developed by our expert data scientists
- ✓ Self-guided online learning to fit into any busy schedule
- ✓ Accessible anytime, anywhere via our virtual learning platform



Knowledge and skills gained

Focusing on practical application of technical and non-technical AI skills, the modules explore AI's capacity to optimise structured interactions, align governance frameworks, and deploy scalable solutions, with a significant focus on ethical considerations and operational efficiency.

- ✓ Ability to design and implement complex prompts for diverse use cases
- ✓ Understanding and application of HITL techniques and fact-checking principles
- ✓ Ability to conduct prompt A/B testing
- ✓ Understanding of AI governance frameworks and compliance requirements
- ✓ Ability to adapt AI messaging for different stakeholders
- ✓ Ability to assess AI ecosystems
- ✓ Understanding of API-driven generative AI (GenAI) benefits



Tools

Large language models (LLMs) (for example Copilot or Google Gemini)
No-code AI automation platforms (no prior coding knowledge required)

Optional masterclasses

 Live online sessions available every month

Example topics include:

- Data Leadership
- Ethical Hacking and Cyber Security
- Discovering and Analysing Market Trends
- Sustainable Technology and Green Computing
- Responsible AI
- Setting AI Strategy
- Emerging Landscapes – AI
- Quantum Computing Fundamentals

