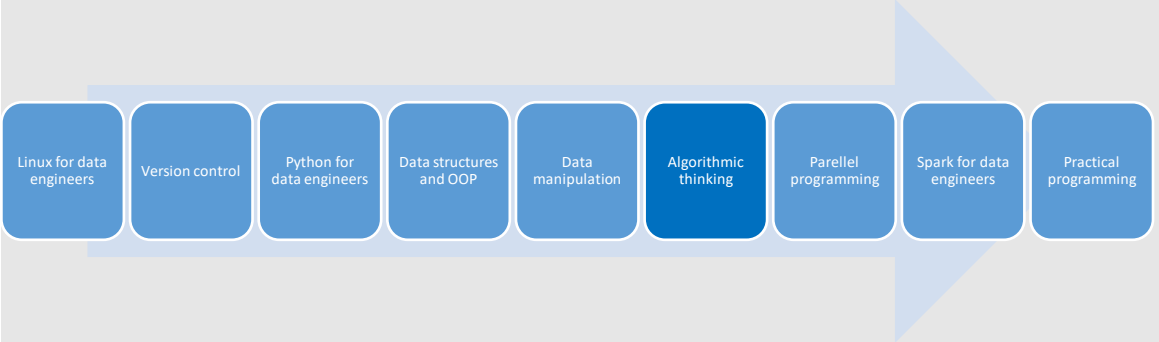
Topic 6 – Algorithmic thinking

This document is the handbook for Topic 6 – **Algorithmic thinking** – within Module 3 – **Programming and Scripting Essentials**.

The purpose of this document is to guide your learning throughout this topic and help you to maximise the value you get from the materials provided by the BPP School of Technology.

Context

This handbook is for one of 9 topics for this Module.



Every topic contributes towards the ultimate learning objectives for the Module, which you will be assessed on at the end of the term.

Module Learning Outcomes

On successfully completing this module, you will be able to:

- **Employ** software development tools and techniques for designing, deploying and maintaining secure data products and pipelines, including debugging, version control and testing.
- **Construct** algorithms that correctly and efficiently handle data at scale whilst mitigating risks.
- **Demonstrate** the knowledge of the steps needed to prepare the code for production.

Module Assessment

The Level 5 Data Engineer EPA has two assessment methods, each with its own mapping of KSBs. The Assessment plan and assessment guidance documents

above list the criteria and KSBs that are assessed. The criteria group the KSBs and describe what the apprentice needs to do to achieve a pass or distinction for that assessment method.

Both assessment methods need to be passed by the candidate:

(1) Project with report

The learner will complete a project and write a report of 3500 words. Project brief submitted at gateway:

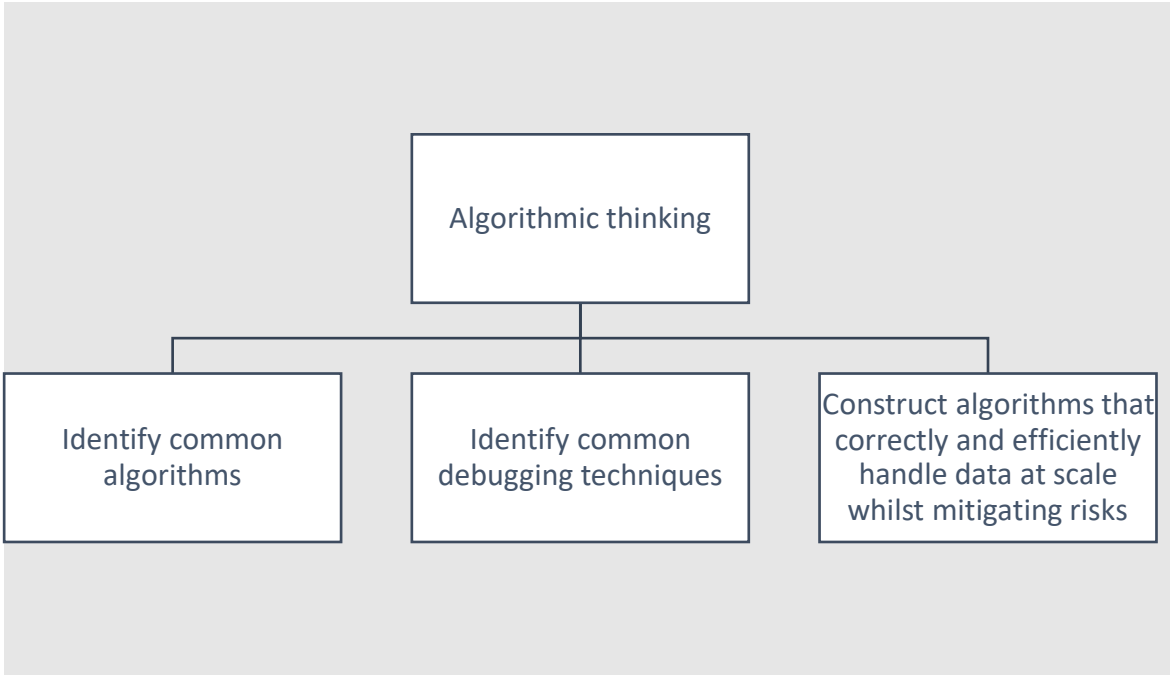
- Learners will have 10 weeks to complete the project and submit the report to the EPAO
- Learners also need to prepare and give a presentation to an independent assessor on their project
- The presentation with questions will last at least 50 minutes. The independent assessor will ask at least 6 questions about the project and presentation
- The project has to have real business application and benefit. Candidates are expected to showcase the use of appropriate standards for sustainability, privacy and security, thoroughly document their data pipeline designs, explain the choice of relevant tooling and demonstrate operational awareness of deployment, access control, risks, and how other stakeholders may be impacted positively and negatively

(2) Professional discussion underpinned by a portfolio of evidence

- Learners will have a professional discussion with an independent assessor. It will last 80 minutes
- They will be asked at least 10 questions about Data Engineering
- The portfolio of evidence will be used to help answer the questions
- We expect the candidates to demonstrate examples of working with data teams on data projects and data products, showcase ideas for future-proofing data, be clear on applying problem-solving skills, show regulatory awareness, and sensitivity towards data quality, data governance and areas for continuous improvement, both personal and organisational

Topic Learning Outcomes

As a step towards build your skills towards the final module assessment, the learning objectives for this topic are:



Introduction

Algorithmic thinking is a fundamental skill that will empower you to design efficient, scalable, and robust solutions for complex data processing challenges. This critical ability will set you apart in the field of data engineering, where the ability to conceptualise and implement effective algorithms is paramount.

In the realm of big data, traditional approaches to data processing often fall short when faced with massive datasets, real-time processing requirements, and complex data relationships. As a data engineer, you'll be tasked with creating data pipelines and systems that can handle these demands efficiently and accurately. Without strong algorithmic thinking skills, you may struggle to optimise performance, manage resource utilisation, and ensure data integrity across large-scale systems. The challenges you'll face require not just technical knowledge, but also the ability to think critically about how to structure and process data in the most effective way possible.

To illustrate the real-world application of algorithmic thinking, consider a leading streaming service like Netflix. They process enormous amounts of user data to provide personalised content recommendations, manage video streaming quality, and predict viewing trends. Their recommendation algorithm alone analyses billions of records, considering factors like viewing history, ratings, time of day, and device

type. This system must work in near real-time, scale to millions of users, and adapt to constantly changing content and user behaviours. The complexity of this task underscores the importance of sophisticated algorithmic thinking in modern data engineering.

By the end of this topic, you will have developed a deep understanding of algorithmic complexity and efficiency, a comprehensive grasp of common algorithmic patterns and their applications in data engineering, insight into optimisation techniques for data-intensive algorithms, and an understanding of how to analyse and improve algorithm performance. These knowledge areas form the foundation of effective algorithmic thinking in data engineering.

Structure

Topics for this programme follow a Prepare-Collaborate-Apply structure:

Prepare

This is the stage where you build the knowledge to underpin your learning. This might involve completing interactive e-learning packages, watching videos, or working through reading materials.

It is essential that you make the most of the learning materials provided before attending webinars, as this will allow you to test your knowledge and stretch you understanding further.

The e-learning for this topic covers building a data-driven culture, fundamentals of data types and sizes, standards and best practices in data engineering, working with different data sources and types through the data management lifecycle, and utilisation of visualisation tools.

Collaborate

This is where you will receive guidance from our expert tutors and coaches to shape and refine your understanding through in-depth explanation, discussion, testing and

carrying out more advanced practical and realistic tasks. This also helps to develop valuable team-working skills.

The webinar for this topic will focus on building a data-driven culture, data fundamentals like types/sizes, key standards and best practices in data engineering, a practical lab, and summarised the core learning concepts.

Apply

You now apply the knowledge you have developed to real-world tasks.

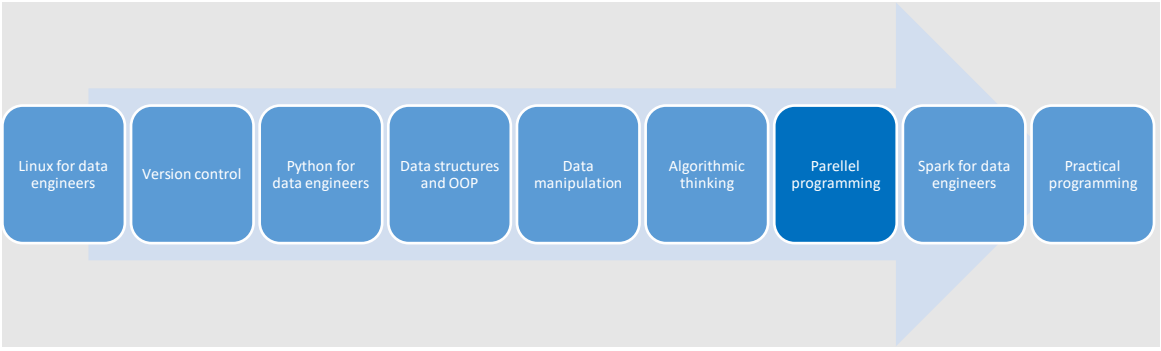
Off-the-job learning tasks

This stage is all about ensuring you truly grasp and retain what you've learned. Through completion of off-the-job (OTJ) revision tasks and tests, you'll get plenty of practice applying your knowledge. Plan to dedicate 6-8 hours each week to guided study and portfolio work, with sessions typically on the same day each week.

Task 1 brief: Your tutor will provide further guidance.

Link

This handbook is for one of 9 topics for this Module.



The sequence of topics in this module is carefully designed so that your knowledge and skills will develop as you progress.

The next topic is **Parallel programming**.