



Data Pipelines and Workflow Orchestration

About

By the end of this lesson, you will be able to:

- **Analyze** the fundamental concepts of data pipelines and workflow orchestration.
- **Design** scalable, reliable, and maintainable data pipelines.
- **Implement** an ETL pipeline using a pre-built dataset (NYC Taxi Data).
- **Orchestrate** workflows and manage dependencies effectively.
- **Optimize** and troubleshoot data pipelines for performance and efficiency.
- **Apply** best practices for building and managing data pipelines.

Content

Lesson	Est. Delivery Time	Skills
Setup	2 min	Set up the development environment.
Introduction	5 min	Overview of data pipelines and workflow orchestration.
Hands-On: Running a Data Pipeline	40 min	Implementing an ETL pipeline using the NYC Taxi dataset.
Optimization & Troubleshooting	30 min	Performance tuning, debugging, and error handling.
Recap & Preparing for Apache Airflow	10 min	Key takeaways and next steps for building pipelines from scratch.
Total Content	~90 min	Hands-on, practical application-focused learning.