

第五章：高级数据工程

Delta Lake：版本控制与数据一致性

- Delta Lake提供事务性数据存储，确保数据一致性。

数据流处理与批处理

- 介绍实时数据流处理和批处理的工作原理。
 -

作业调度与自动化

- 实现任务的自动调度和依赖管理，提升数据处理效率。

高级数据工程设计模式与数据湖管理

- 设计模式包括数据清洗、转换，数据湖管理以确保数据高效存储。