

第四章：Spark核心操作

Spark基本概念回顾

- 回顾RDD、DataFrames、Spark SQL等核心概念。

使用DataFrames和Spark SQL进行 数据处理

- 支持数据选择、过滤、转换和SQL查询。

聚合、过滤与分组操作

- 执行数据聚合、过滤和分组分析。

大数据分析的常用模式和性能优化

- 介绍流数据、批处理和性能优化方法。