Lesson 2.5: Broadcast Joins

DISTRIBUTED COMPUTING WITH SPARK SQL

Broadcast Joins



Brooke Wenig Machine Learning Practice Lead Databricks



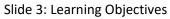


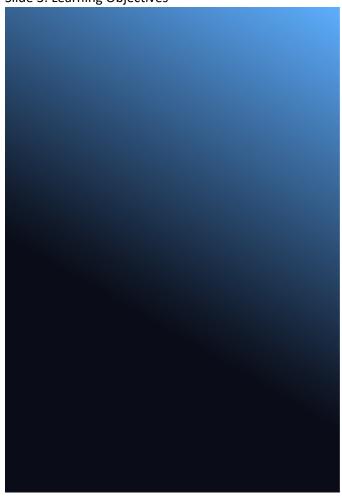


Welcome Back!

Previous courses showed how to join two tables

Learn how to broadcast joins using Spark SQL





Learning Objectives

Identify candidate datasets for broadcast joins

Understand how this helps to decrease data shuffle and query time