

Lesson 3.3: Accessing Data

# DISTRIBUTED COMPUTING WITH SPARK SQL

Accessing Data



**Conor Murphy**  
Data Scientist  
Databricks

**UC DAVIS**  
Continuing and Professional Education

Slide 2: Welcome Back!



## Welcome Back!

Connecting to two common data sources:

BLOB

JDBC

Slide 3: BLOB and JDBC

## **BLOB and JDBC**

BLOB is a common way of storing large amounts of data

JDBC is an application programming interface (API) for Java environments

Predicate push down functionality

Slide 4: Learning Objectives



## Learning Objectives

Read data from a BLOB store

Read data in both serial and parallel from  
a JDBC connection