

**BerkeleyX: CS105x Introduction to Apache Spark**

Bookmarks

- ▶ Week 1 - Apache Spark Programming Model
- ▼ Week 2 - The Structured Query Language and Spark SQL

Lecture 2: The Structured Query Language and Spark SQL

Quizzes



Lab 1A/1B - Learning Apache Spark (Due September 10, 2016 at 23:59 UTC)

Lab



- ▶ Week 3 - Analyzing Semi-Structured Data with Apache Spark

Week 2 - The Structured Query Language and Spark SQL > Lecture 2: The Structured Query Language and Spark SQL > The Big Data Problem

Bookmark

The Big Data Problem

BERCS1052016-V001700



▶ 0:00 / 1:38

▶ 1.0x



HD



Download video

Download transcript

.srt

Big Data Tools

(1/1 point)

Why is using traditional analysis tools for big data a poor choice?

- ☐ Storage is becoming less expensive.
- ☐ The time to read from a 1 TB drive is 3 hours.
- ☒ Big data does not fit on a single machine. ✓
- ☐ CPUs are getting faster and faster every year.

EXPLANATION

Big data is data that is many 10's of Terabytes to Petabytes in size, which means it does not fit on a single machine.



You have used 2 of 3 submissions

CC BY-NC-SA Some Rights Reserved



© edX Inc. All rights reserved except where noted. EdX, Open edX and the edX and Open EdX logos are registered trademarks or trademarks of edX Inc.

