

**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 1 - Apache Spark Programming Model &gt; Lecture 1: Apache Spark Architecture and Programming Model &gt; The Structure Spectrum

Bookmark

## The Structure Spectrum

BERCS1052016-V000500



[Download video](#)[Download transcript](#)[.srt](#)

## Semi-Structured Data

(1 point possible)

Which of the following items are examples of semi-structured data:

☐ Excel spreadsheet ✓

☐ The contents of a user's Facebook posts

☒ A downloaded Comma-Separated Values (CSV) version of your bank statement ✓

☐ A collection of photos

☒ The metadata (time, date, location, exposure, etc.) associated with a collection of photos ✓

?

Note: Make sure you select all of the correct options—there may be more than one!

**EXPLANATION**

An Excel spreadsheet, CSV bank statement, and collection of photo metadata are all examples of semi-structured data. The contents of a user's Facebook posts and a collection of photos are examples of unstructured data.



*You have used 4 of 4 submissions*

## eXtensible Markup Language Documents

(1/1 point)

eXtensible Markup Language documents are an example of which type of data?

- ☐ Structured
- ☒ Semi-Structured ✓
- ☐ Unstructured
- ☐ Random

*You have used 1 of 3 submissions*

## Working with Semi-Structured Data

(1/1 point)

When working with semi-structured data, which statement is correct?

- ☐ The programmer must specify the data's schema
- ☐ The data's schema can be dynamically inferred by examining the data as it is loaded
- ☒ Either the programmer must specify the data's schema, or the schema can be dynamically inferred by examining the data as it is loaded ✓
- ☐ There is no schema when working with semi-structured data

*You have used 1 of 3 submissions*

## Working with Structured Data

(1/1 point)

When working with structured data, which statement is correct?

- ☒ The programmer must specify the data's schema ✓
- ☐ The data's schema can be dynamically inferred by examining the data as it is loaded

☐ Either the programmer must specify the data's schema, or the schema can be dynamically inferred by examining the data as it is loaded

☐ There is no schema when working with structured data

*You have used 2 of 3 submissions*

© ⓘ Ⓢ ⓘ 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.

POWERED BY  
OPENedX®

