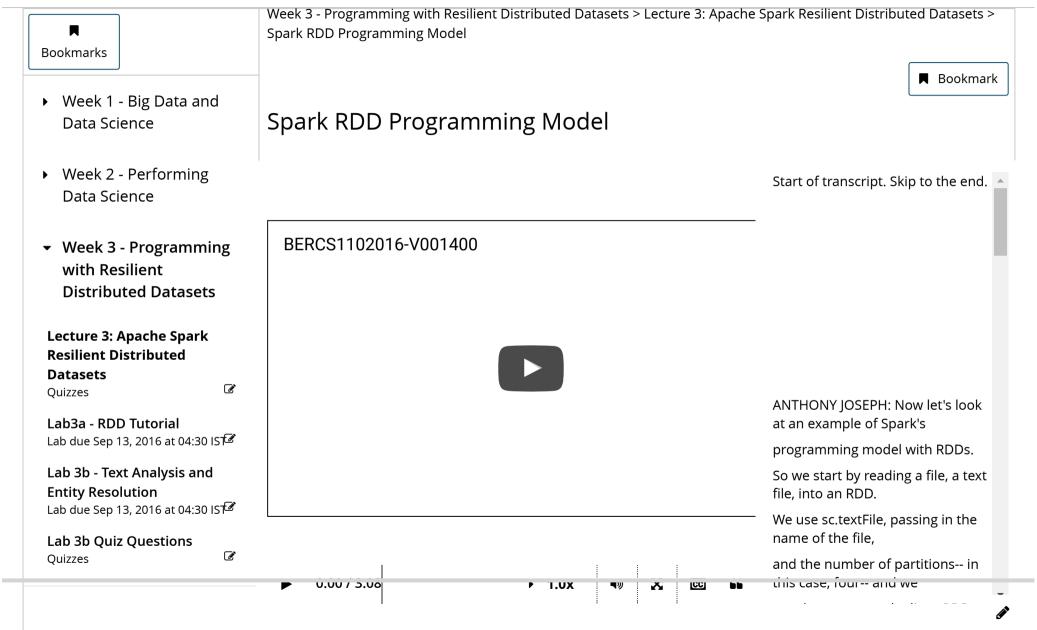


BerkeleyX: CS110x Big Data Analysis with Apache Spark



Download video Download transcript .srt

Spark RDD Program Lifecycle

(1/1 point)

Which of the following are part of a Spark RDD program's lifecycle?

- RDDs that are reused may be cached
- Transformations cause parallel computation to be immediately executed
- Actions cause parallel computation to be immediately executed
- Transformations lazily create new RDDs
- Actions create recipes for peforming parallel computation on datasets



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

EXPLANATION

Transformations specify how to perform parallel computation in a lazily evaluated manner. Action cause the transformations to be executed. If you plan to reuse an RDD, you should cache it.

⊚ (1) (\$) (2) 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.















