



databricks

Apache Spark[™] Programming with Databricks

Code:	ASPWD
Length:	2 days
URL:	View Online

In this course, you will explore the fundamentals of Apache Spark and Delta Lake on Databricks. You will learn the architectural components of Spark, the DataFrame and Structured Streaming APIs, and how Delta Lake can improve your data pipelines. Lastly, you will execute streaming queries to process streaming data and understand the advantages of using Delta Lake.

• This course will prepare you to take the Databricks Certified Associate Developer for Apache Spark exam.

Skills Gained

- Define Spark's architectural components
- Describe how DataFrames are transformed, executed, and optimized in Spark
- Apply the DataFrame API to explore, preprocess, join, and ingest data in Spark
- Apply the Structured Streaming API to perform analysis on streaming data
- Use Delta Lake to improve the quality and performance of data pipelines

Prerequisites

- Completion of Introduction to Python for Data Science & Data Engineering, OR familiarity with Python and basic programming concepts, including data types, lists, dictionaries, variables, functions, loops, conditional statements, exception handling, accessing classes, and using third-party libraries
- Basic knowledge of SQL, including writing queries using SELECT, WHERE, GROUP BY, ORDER BY, LIMIT, and JOIN

Course Details

Course Outline

Day 1

- Spark overview
- Databricks platform overview
- Spark SQL

- DataFrame reader, writer, transformation, and aggregation
- Datetimes
- Complex types

Day 2

- User-defined functions (UDFs) and vectorized UDFs
- Spark internals
- Query optimization
- Partitioning
- Streaming API
- Delta Lake

Schedule (as of 3)

Date	Location	
Apr 29, 2024 – May 2, 2024	Virtual	<u>Enroll</u>
May 13, 2024 - May 16, 2024	Virtual	<u>Enroll</u>
May 28, 2024 - May 31, 2024	Virtual	<u>Enroll</u>
Jun 10, 2024 - Jun 13, 2024	Virtual	<u>Enroll</u>
Jun 18, 2024 - Jun 21, 2024	Virtual	Enroll

Download Whitepaper: Accelerate Your Modernization Efforts with a Cloud-Native Strategy Get Your Free Copy Now

ExitCertified® Corporation and iMVP® are registered trademarks of ExitCertified ULC and ExitCertified Corporation, respectively Copyright © 2024 ExitCertified ULC & ExitCertified Corporation. All Rights Reserved.

Generated 4