



Advanced Data Engineering with Databricks

Course ID #: 7000-808-ZZ-Z

Hours: 14

Course Content

Course Description:

In this course, you will build upon your existing knowledge of Apache Spark, Structured Streaming, and Delta Lake to unlock the full potential of the data lakehouse by utilizing the suite of tools provided by Databricks. This course places a heavy emphasis on designs favoring incremental data processing, enabling systems optimized to continuously ingest and analyze ever-growing data. By designing workloads that leverage built-in platform optimizations, data engineers can reduce the burden of code maintenance and on-call emergencies, and quickly adapt production code to new demands with minimal refactoring or downtime.

The topics in this course should be mastered prior to attempting the Databricks Certified Data Engineer Professional exam.

Course Objectives:

Upon successful completion of this course students will be able to:

- Design databases and pipelines optimized for the Databricks Lakehouse Platform.
- Implement efficient incremental data processing to validate and enrich data driving business decisions and applications.
- Leverage Databricks-native features for managing access to sensitive data and fulfilling right-to-be-forgotten requests.
- Manage error troubleshooting, code promotion, task orchestration, and production job monitoring using Databricks tools.

Register for this class by visiting us at:

www.tcworkshop.com or calling us at 800-639-3535



Advanced Data Engineering with Databricks

Course ID #: 7000-808-ZZ-Z

Hours: 14