Databricks Certified Data Engineer Associate Course Overview

Karthick Selvam

Last updated: May 10, 2025

Exam Topics

The certification exam covers five broad categories:

- **Databricks Lakehouse Platform**
- **ETL with Apache Spark**
- **Incremental Data Processing**
- **Production Pipelines**
- **Data Governance**

For detailed subtopics, refer to the Official Exam Guide.

Course Structure

The course follows the exam topics, with some adjustments due to overlapping features.

1. Cloud Platform Setup (Microsoft Azure)

- Databricks is available on AWS, Azure, and Google Cloud.
- **Azure is used in this course** due to its first-party integration with Databricks.
- A guide on setting up an Azure subscription will be provided.
- *Note:* The exam does **not** test cloud platform knowledge—only Databricks.
- **Skip this section** if you already have an Azure subscription.
- *Community Edition is not sufficient* for exam preparation.

2. Introduction to Databricks Lakehouse Platform

- Overview of Databricks Lakehouse architecture.
- Key components of the Databricks workspace.

3. Unity Catalog (Metadata & Governance)

- Introduction to Unity Catalog for managing Lakehouse assets.
- Covers metadata handling and later explores data governance.

4. Apache Spark for Data Processing

- Querying and transforming data in the Lakehouse.
- Hands-on exercises with Spark.

5. Spark Structured Streaming

- Incremental data processing with Spark.

6. Delta Lake

- Features of Delta Lake for Lakehouse architecture.
- Includes incremental processing and other capabilities.

7. Delta Live Tables (ETL Framework)

- Simplifying ETL pipelines with Databricks.

8. Databricks Jobs

- Building production-grade data pipelines.

9. Databricks SQL

- Role of data engineers in SQL workflows.
- *Note:* Some SQL-related questions may appear in other exam sections.

10. Data Governance with Unity Catalog

- Security, access control, and compliance features.

11. Full-Length Practice Exam

- Simulates the actual certification test.

Study Recommendations

- Follow sections **in order** for the best learning experience.
- Complete **all hands-on exercises** before moving forward.
- If already proficient, **at least complete Section 5 (Unity Catalog)** to set up metastore and storage for later exercises.