Implement data engineering solutions using Azure Databricks

Data Engineering mit Azure Databricks, Apache Spark und Delta Lake

Kurs ID: DP-750T00
Dauer: 4 Tage
ab € 2.490,00 zzgl. MwSt.

Mehr Details

In diesem Kurs lernen Sie, wie Sie Azure Databricks für moderne Data-Engineering- und Analytics-Szenarien einsetzen. Sie erfahren, wie Sie große Datenmengen mit Apache Spark verarbeiten, analysieren und für weiterführende Auswertungen in der Cloud aufbereiten. Sie beschäftigen sich mit zentralen Azure-Databricks-Konzepten, DataFrames, Spark-Clustern und Notebooks und lernen, wie Sie Daten effizient erfassen, untersuchen und visualisieren. Ein weiterer Schwerpunkt liegt auf Delta Lake, ACID-Transaktionen, Schema Enforcement sowie Datenversionierung und Time Travel. Darüber hinaus erhalten Sie einen Überblick über Data Governance mit Unity Catalog und Microsoft Purview und entwickeln ein fundiertes Verständnis dafür, wie sich skalierbare und zuverlässige Datenlösungen mit Azure Databricks umsetzen lassen.

Dieser Kurs ersetzt ab **30.09.2026** den eingestellten DP-3011: Implementing a Data Analytics Solution with Azure Databricks Kurs.

Lernziele

- End-to-End-Datenverarbeitung mit Azure Databricks
- Konfiguration und Verwaltung von Azure Databricks-Umgebungen
- Sichere Datenpipelines und Datengovernance mit Unity Catalog
- Bereitstellung und Optimierung von Datentechnik-Workloads
- Skalierbare Lakehouse-Architekturen für Unternehmensanforderungen

Kursinhalte

- **Einrichten und Konfigurieren einer Azure Databricks-Umgebung**
  - Erkunden von Azure Databricks
  - Grundlegendes zur Architektur von Azure Databricks
  - Grundlegendes zu Azure Databricks-Integrationen
  - Auswählen und Konfigurieren von Compute in Azure Databricks
  - Erstellen und Organisieren von Objekten im Unity-Katalog
- **Sichern und Steuern von Unity-Katalogobjekten in Azure Databricks**

    - Secure Unity Catalog-Objekte
    - Verwalten von Unity-Katalogobjekten
- **Vorbereiten und Verarbeiten von Daten mit Azure Databricks**
    - Entwerfen und Implementieren der Datenmodellierung mit Azure Databricks
    - Aufnehmen von Daten in den Unity-Katalog
    - Bereinigen, transformieren und laden von Daten in den Unity-Katalog
    - Implementieren und Verwalten von Datenqualitätseinschränkungen mit Azure Databricks
- **Bereitstellen und Verwalten von Datenpipelines und Workloads mit Azure Databricks**
    - Entwerfen und Implementieren von Datenpipelines mit Azure Databricks
    - Implementieren von Lakeflow-Aufträgen mit Azure Databricks
    - Implementieren von Entwicklungslebenszyklusprozessen in Azure Databricks
    - Überwachen, Beheben von Problemen und Optimieren von Workloads in Azure Databricks

Voraussetzungen

- Grundlegende Kenntnisse über Datenanalysekonzepte und Cloudspeicher
- Vertrautheit mit den Prinzipien der Datenorganisation
- Erfahrung im Arbeiten mit SQL sowie Python
- Gutes Verständnis für Azure Databricks-Arbeitsbereiche und Unity Catalog
- Vertrautheit mit Datenzugriffsmustern sowie Datentechnik- und Data-Warehouse-Konzepten
- Grundlegende Kenntnisse der Azure-Sicherheit einschließlich Microsoft Entra ID
- Vertrautheit mit den Grundlagen der Git-Versionssteuerung

Zielgruppe

- Data Engineers
- Data Analysts
- Cloud- und Azure-Spezialist:innen
- Entwickler:innen mit Datenfokus
- IT-Professionals im Data- und Analytics-Umfeld