Implement data engineering solutions using Azure Databricks (DP-750)

Duration: 4 Days **Course Code: M-DP750** **Delivery Method: Virtual Learning**

Overview:

Design and deliver scalable data pipelines using Azure Databricks and the lakehouse architecture.
This course moves from foundational setup to production deployment, covering environment configuration and enterprise-grade governance. Learn to build robust ingestion pipelines, implement security with Unity Catalog, and deploy optimized workloads. By the end, you will have the practical skills to implement, secure, and maintain scalable lakehouse solutions that meet rigorous enterprise requirements.

Virtual Learning

This interactive training can be taken from any location, your office or home and is delivered by a trainer. This training does not have any delegates in the class with the instructor, since all delegates are virtually connected. Virtual delegates do not travel to this course, Global Knowledge will send you all the information needed before the start of the course and you can test the logins.

Target Audience:

The target audience is data engineers who have fundamental knowledge of data analytics concepts, a basic understanding of cloud storage, and familiarity with data organization principles. They should be comfortable working with SQL and have experience using Python, including notebooks, for data engineering tasks. Learners are expected to have a good understanding of Azure Databricks workspaces and Unity Catalog, along with familiarity with data access patterns and core data engineering and data warehouse concepts. In addition, they should have foundational knowledge of Azure security, including Microsoft Entra ID, and be familiar with Git version control fundamentals.

Objectives:

- **By the end of this course, learners will be able to:**
- Configure and manage Azure Databricks environments, including workspaces, clusters, and compute resources.
- Ingest, transform, and model data using SQL and Python to support analytics and downstream workloads.
- Build, deploy, and maintain scalable data pipelines using Azure Databricks and lakehouse patterns.
- Apply data governance and security best practices using Unity Catalog for access control, governance, and data quality.
- Integrate Azure Databricks with core Azure services, including Azure Storage, Azure Data Factory, Azure Monitor, Microsoft Entra ID, and Azure Key Vault.
- Troubleshoot, monitor, and optimize Databricks workloads to ensure reliable, production-ready data engineering solutions.

Prerequisites:

Basic understanding of data and analytics concepts is recommended.

Testing and Certification

None

Follow-on-Courses:

None

Content:

Module 1: Set up and configure an Azure Databricks environment

- Explore Azure Databricks
- Understand Azure Databricks architecture
- Understand Azure Databricks Integrations
- Select and Configure Compute in Azure Databricks
- Create and organize objects in Unity Catalog

Module 2: Secure and govern Unity Catalog objects in Azure Databricks

- Secure Unity Catalog objects
- Govern Unity Catalog objects

Module 3: Prepare and process data with Azure Databricks

- Design and implement data modeling with Azure Databricks
- Ingest data into Unity Catalog
- Cleanse, transform, and load data into Unity Catalog
- Implement and manage data quality constraints with Azure Databricks

Module 4: Deploy and maintain data pipelines and workloads with Azure Databricks

- Design and implement data pipelines with Azure Databricks
- Implement Lakeflow Jobs with Azure Databricks
- Implement development lifecycle processes in Azure Databricks
- Monitor, troubleshoot and optimize workloads in Azure Databricks

Further Information:

For More information, or to book your course, please call us on 00 971 4 446 4987

training@globalknowledge.ae

www.globalknowledge.com/en-ae/

Global Knowledge, Dubai Knowledge Village, Block 2A,First Floor, Office F68, Dubai, UAE