

Implementing a Data Analytics Solution with Azure Synapse Analytics

Duration: 1 Day Course Code: M-DP3012

Overview:

This is a single day Instructor Lead Course designed to give the learners instruction on the SQL dedicated and serverless Spark pools and providing instruction of data wrangling and the ELT process using Synapse Pipelines which is very similar to those familiar with Azure Data Factory (ADF) to move data into the Synapse dedicated pool database.

Target Audience:

This course is destinated to administrators and data specialists looking to implement a Data Analytics Solution with Azure Synapse Analytics.

Objectives:

- During this course you will learn to:
 - Process large data sets with serverless Spark pools.
 - Utilize Spark SQL & DataFrames for data exploration & transformation.
 - Implement Delta Lake for reliable data storage & version control.
 - Work with streaming data using Synapse SQL Streaming.
 - Integrate machine learning models with Spark MLlib & other frameworks.
 - Preprocess & prepare data for machine learning tasks.
 - Train & evaluate machine learning models within Synapse Analytics.
 - Deploy & manage machine learning models in production.
 - Understand the business value of data analytics & big data projects.
 - Learn best practices for building & deploying data solutions with Synapse Analytics.
 - Prepare for data engineer, analyst, & architect roles using Azure Synapse Analytics.
- Master Azure Synapse Analytics architecture and key concepts.
- Build data pipelines with Synapse Pipelines.
- Leverage dedicated SQL pools & serverless Spark pools for data warehousing & big data analysis.
- Develop data models and perform SQL queries for analysis.
- Analyze data with Spark and Delta Lake.
- Visualize & report data using Power BI.
- Monitor & optimize data pipelines for performance.
- Design & build data warehouse models (star/snowflake schemas).
- Load data efficiently into dedicated SQL pools.
- Perform complex queries on large data sets.
- Manage & secure Synapse Analytics data warehouses.

Prerequisites:

-

Content:

MODULE 1: Introduction to Azure Synapse Analytics

- Identify the business problems that Azure Synapse Analytics addresses.
- Describe core capabilities of Azure Synapse Analytics.
- Determine when to use Azure Synapse Analytics.

MODULE 2: Use Azure Synapse serverless SQL pool to query files in a data lake

- Identify capabilities and use cases for serverless SQL pools in Azure Synapse Analytics
- Query CSV, JSON, and Parquet files using a serverless SQL pool
- Create external database objects in a serverless SQL pool

MODULE 3: Analyze data with Apache Spark in Azure Synapse Analytics

- Identify core features and capabilities of Apache Spark.
- Configure a Spark pool in Azure Synapse Analytics.
- Run code to load, analyze, and visualize data in a Spark notebook.

MODULE 4: Use Delta Lake in Azure Synapse Analytics

- Describe core features and capabilities of Delta Lake.
- Create and use Delta Lake tables in a Synapse Analytics Spark pool.
- Create Spark catalog tables for Delta Lake data.
- Use Delta Lake tables for streaming data.
- Query Delta Lake tables from a Synapse Analytics SQL pool.

MODULE 5: Analyze data in a relational data warehouse

- Design a schema for a relational data warehouse.
- Create fact, dimension, and staging tables.
- Use SQL to load data into data warehouse tables.
- Use SQL to query relational data warehouse tables.

MODULE 6: Build a data pipeline in Azure Synapse Analytics

- Describe core concepts for Azure Synapse Analytics pipelines.
- Create a pipeline in Azure Synapse Studio.
- Implement a data flow activity in a pipeline.

Initiate and monitor pipeline runs.

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