DP-3012: Implementing a Data Analytics Solution with Azure Synapse Analytics

Course Duration: 8 Hours **Course code: DP-3012**

1. Course Overview

DP-3012 walks you through the core capabilities of Azure Synapse Analytics and how to build a comprehensive, end-to-end data analytics solution using it. You'll learn to ingest, explore, prepare, transform, and serve data at scale. The course emphasizes hands-on implementation, covering both on-demand serverless and provisioned (dedicated SQL pool) methods. By course end, you'll be able to design, implement, monitor, and optimize data pipelines and analytics workloads using Synapse's integrated ecosystem.

2. What you'll learn?

- Understand Azure Synapse Analytics architecture — including workspaces, SQL pools, Apache Spark pools, and pipelines.
- Ingest and orchestrate data using Azure Data Factory capabilities within Synapse.
- Query data using serverless SQL pools and provisioned SQL pools (dedicated).
- Leverage Spark notebooks for data exploration, transformation, and machine learning.
- Design ETL/ELT pipelines for structured, semi-structured, and unstructured data.
- Secure data access using authentication, role-based access control, and data classification.
- Optimize performance through indexing, distribution, caching, and partition strategies.
- Implement monitoring and alerting, using built-in Synapse monitoring tools and integration with Azure Monitor.

Phone: +91-999-911-1686 **Mail Us:** info@ssdntech.com

- Deploy and manage resources using ARM templates, PowerShell, or CLI.

3. Target Audience

This course is ideal for:

- Data Engineers aiming to architect and implement scalable, cloud-based analytics solutions.
- BI Developers and Data Analysts looking to leverage Synapse for data prep, transformation, and ad-hoc querying.
- Solution Architects who need to design end-to-end analytics pipelines, integrating various Azure data services.
- DevOps Engineers focused on deploying and automating Synapse workloads.

4. Pre-Requisites

- A strong foundational understanding of the following is expected:
- Basic knowledge of Azure fundamentals, including core services (e.g. storage, networking).
- Experience with SQL querying and data modelling concepts.
- Familiarity with Python or Spark for processing workloads (recommended but not mandatory).
- Exposure to ETL workflows or orchestration tools (Azure Data Factory or similar).
- Understanding of data warehousing concepts (fact/dimension tables, partitioning, indexing).

5. Course content

Module 1: Introduction to Azure Synapse Analytics

• Overview of Azure Synapse Analytics architecture

• Synapse workspace components: SQL pools, Spark pools, Pipelines

- Serverless SQL pool vs Dedicated SQL pool
- Data storage integration with ADLS and Blob Storage
- Understanding Synapse Studio interface
- Knowledge check

Module 2: Data Ingestion and Orchestration

- Connecting to structured, semi-structured, and unstructured data sources
- Ingesting data using Synapse Pipelines and Copy activity
- Integrating Synapse with Azure Data Factory capabilities
- Scheduling and triggering data workflows
- Hands-on: Building an ingestion pipeline from Azure Blob Storage
- Knowledge check

Module 3: Querying Data with SQL Pools

- Serverless SQL pool basics and query execution
- Provisioning and configuring Dedicated SQL pools
- Loading data into dedicated pools using PolyBase and COPY statement
- Query optimization techniques
- Using external tables and views
- Hands-on: Running queries across multiple data sources
- Knowledge check

Module 4: Data Processing with Apache Spark in Synapse

- Introduction to Apache Spark pools in Synapse
- Creating and managing Spark notebooks

• Data exploration and transformation with PySpark/Scala
• Integrating Spark with machine learning workflows
• Hands-on: Cleaning and transforming raw data using Spark
• Knowledge check

Module 5: Security and Governance

• Authentication methods in Synapse (AAD, key-based)
• Role-Based Access Control (RBAC) and workspace permissions
• Data masking and column-level security
• Data encryption and classification in Synapse
• Auditing and compliance features
• Knowledge check

Module 6: Performance Optimization

• Distribution methods (hash, round-robin, replicated)
• Partitioning strategies for large datasets
• Materialized views and result set caching
• Monitoring performance and workload management
• Hands-on: Tuning queries for faster execution
• Knowledge check

Module 7: Monitoring, Troubleshooting, and Operations

• Built-in Synapse monitoring tools
• Integrating Synapse with Azure Monitor and Log Analytics
• Setting up alerts and diagnostics

- Troubleshooting common performance and connectivity issues
- Cost monitoring and optimization strategies
- Knowledge check

Module 8: Deployment and Automation

- Using ARM templates for Synapse deployment
- Automating tasks with Azure CLI and PowerShell
- Implementing CI/CD pipelines for Synapse
- Hands-on: Deploying an end-to-end solution via automation
- Knowledge check

Module 9: Capstone Project

- Designing an end-to-end data analytics solution
- Ingesting and transforming multi-format data
- Loading data into dedicated SQL pool
- Implementing security, optimization, and monitoring
- Final project presentation

Module 10: Course Summary

- Debrief and recap of key concepts
- Real-world best practices for Synapse implementation
- What's next? (certification paths, advanced learning)