

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

Course Outline

1) Implement a Data Analytics Solution with Azure Synapse Analytics

- 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
- 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
- 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
- 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
- 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
- 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

2) Build data analytics solutions using Azure Synapse serverless SQL pools

- Use Azure Synapse serverless SQL pool to query files in a data lake
 - Identify capabilities and use cases for serverless SQL pools
 - Query CSV, JSON, and Parquet files using a serverless SQL pool
 - Create external database objects in a serverless SQL pool
- Use Azure Synapse serverless SQL pools to transform data in a data lake
 - Use a CREATE EXTERNAL TABLE AS SELECT (CETAS) statement to transform data
 - Encapsulate a CETAS statement in a stored procedure
 - Include a data transformation stored procedure in a pipeline

- Create a lake database in Azure Synapse Analytics
 - Understand lake database concepts and components
 - Describe database templates in Azure Synapse Analytics
 - Create a lake database
- Secure data and manage users in Azure Synapse serverless SQL pools
 - Choose an authentication method in serverless SQL pools
 - Manage users in serverless SQL pools
 - Manage user permissions in serverless SQL pools

3) Perform data engineering with Azure Synapse Apache Spark Pools

- 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
- Transform data with Spark in Azure Synapse Analytics
 - Use Apache Spark to modify and save dataframes
 - Partition data files for improved performance and scalability
 - Transform data with SQL
- Use Delta Lake in Azure Synapse Analytics
 - Describe core features and capabilities of Delta Lake
 - Create and use Delta Lake tables in a 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

4) Work with Data Warehouses using Azure Synapse Analytics

- 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
- Load data into a relational data warehouse
 - Load staging tables in a data warehouse
 - Load dimension tables in a data warehouse
 - Load time dimensions in a data warehouse
 - Load slowly changing dimensions in a data warehouse
 - Load fact tables in a data warehouse
 - Perform post-load optimizations in a data warehouse
- Manage and monitor data warehouse activities in Azure Synapse Analytics
 - Scale compute resources in Azure Synapse Analytics
 - Pause compute in Azure Synapse Analytics
 - Manage workloads in Azure Synapse Analytics
 - Use Azure Advisor to review recommendations
 - Use Dynamic Management Views to identify and troubleshoot query performance
- Secure a data warehouse in Azure Synapse Analytics
 - Understand network security options for Azure Synapse Analytics

- Configure Conditional Access
- Configure Authentication
- Manage authorization through column and row level security
- Manage sensitive data with Dynamic Data Masking
- Implement encryption in Azure Synapse Analytics

5) Transfer and transform data with Azure Synapse Analytics pipelines

- 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
- Use Spark Notebooks in an Azure Synapse Pipeline
 - Describe notebook and pipeline integration
 - Use a Synapse notebook activity in a pipeline
 - Use parameters with a notebook activity