

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

Course Details

Course Code: DP-3012

Duration: 1 day

Notes:

- This course syllabus should be used to determine whether the course is appropriate for the students, based on their current skills and technical training needs.
- Course content, prices, and availability are subject to change without notice.
- Terms and Conditions apply

Elements of this syllabus are subject to change.

About this course

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.

Audience Profile

The Audience should have familiarity with notebooks that use different languages and a Spark engine, such as Databricks, Jupyter Notebooks, Zeppelin notebooks and more. They should also have some experience with SQL, Python, and Azure tools, such as Data Factory.

Academy IT Pty Ltd

Level 4, 45 Grenfell Street
ADELAIDE 5000

Email: sales@academyit.com.au

Web: www.academyit.com.au

Phone: 08 7324 9800

Brian: 0400 112 083

Introduction to Azure Synapse Analytics

Learn about the features and capabilities of Azure Synapse Analytics - a cloud-based platform for big data processing and analysis.

Learning objectives

In this module, you'll learn how to:

- 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

With Azure Synapse serverless SQL pool, you can leverage your SQL skills to explore and analyze data in files, without the need to load the data into a relational database.

Learning objectives

After the completion of this module, you will be able to:

- 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

Apache Spark is a core technology for large-scale data analytics. Learn how to use Spark in Azure Synapse Analytics to analyze and visualize data in a data lake.

Learning objectives

After completing this module, you will be able to:

- 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

Delta Lake is an open source relational storage area for Spark that you can use to implement a data lakehouse architecture in Azure Synapse Analytics.

Learning objectives

In this module, you'll learn how to:

- 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

Relational data warehouses are a core element of most enterprise Business Intelligence (BI) solutions, and are used as the basis for data models, reports, and analysis.

Learning objectives

In this module, you'll learn how to:

- 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.