

DP-203T00: Data Engineering on Microsoft Azure

Course Description

In this course, the student will learn how to implement and manage data engineering workloads on Microsoft Azure, using Azure services such as Azure Synapse Analytics, Azure Data Lake Storage Gen2, Azure Stream Analytics, Azure Databricks, and others. The course focuses on common data engineering tasks such as orchestrating data transfer and transformation pipelines, working with data files in a data lake, creating and loading relational data warehouses, capturing and aggregating streams of real-time data, and tracking data assets and lineage.

Course Duration

4 days

Prerequisites

Successful students start this course with knowledge of cloud computing and core data concepts and professional experience with data solutions.

Specifically completing:

- AZ-900 - Azure Fundamentals
- DP-900 - Microsoft Azure Data Fundamentals

Course Outline

- Introduction to data engineering on Azure
- Introduction to Azure Data Lake Storage Gen2
- Introduction to Azure Synapse Analytics
- Use Azure Synapse serverless SQL pool to query files in a data lake
- Use Azure Synapse serverless SQL pools to transform data in a data lake
- Create a lake database in Azure Synapse Analytics
- Analyze data with Apache Spark in Azure Synapse Analytics
- Transform data with Spark in Azure Synapse Analytics
- Use Delta Lake in Azure Synapse Analytics
- Analyze data in a relational data warehouse
- Load data into a relational data warehouse
- Build a data pipeline in Azure Synapse Analytics
- Use Spark Notebooks in an Azure Synapse Pipeline
- Plan hybrid transactional and analytical processing using Azure Synapse Analytics
- Implement Azure Synapse Link with Azure Cosmos DB
- Implement Azure Synapse Link for SQL
- Get started with Azure Stream Analytics
- Ingest streaming data using Azure Stream Analytics and Azure Synapse Analytics
- Visualize real-time data with Azure Stream Analytics and Power BI
- Introduction to Microsoft Purview
- Integrate Microsoft Purview and Azure Synapse Analytics
- Explore Azure Databricks



- Use Apache Spark in Azure Databricks
- Run Azure Databricks Notebooks with Azure Data Factory



Who Should Attend

The primary audience for this course is data professionals, data architects, and business intelligence professionals who want to learn about data engineering and building analytical solutions using data platform technologies that exist on Microsoft Azure.

The secondary audience for this course includes data analysts and data scientists who work with analytical solutions built on Microsoft Azure