

## Course duration

- 4 days

## Course Benefits

- Explore compute and storage options for data engineering workloads in Azure
- Design and Implement the serving layer
- Understand data engineering considerations
- Run interactive queries using serverless SQL pools
- Explore, transform, and load data into the Data Warehouse using Apache Spark
- Perform data Exploration and Transformation in Azure Databricks
- Ingest and load Data into the Data Warehouse
- Transform Data with Azure Data Factory or Azure Synapse Pipelines
- Integrate Data from Notebooks with Azure Data Factory or Azure Synapse Pipelines
- Optimize Query Performance with Dedicated SQL Pools in Azure Synapse
- Analyze and Optimize Data Warehouse Storage
- Support Hybrid Transactional Analytical Processing (HTAP) with Azure Synapse Link
- Perform end-to-end security with Azure Synapse Analytics
- Perform real-time Stream Processing with Stream Analytics
- Create a Stream Processing Solution with Event Hubs and Azure Databricks
- Build reports using Power BI integration with Azure Synapse Analytics
- Perform Integrated Machine Learning Processes in Azure Synapse Analytics

### Available Delivery Methods

#### Public Class

Public expert-led online training from the convenience of your home, office or anywhere with an internet connection. Guaranteed to run .

#### Private Class

Private classes are delivered for groups at your offices or a location of your choice.

### Microsoft Certified Partner

Webucator is a Microsoft Certified Partner for Learning Solutions (CPLS). This class uses official Microsoft courseware and will be delivered by a Microsoft Certified Trainer (MCT).

## Course Outline

1. Explore compute and storage options for data engineering workloads
  1. Introduction to Azure Synapse Analytics
  2. Describe Azure Databricks
  3. Introduction to Azure Data Lake storage
  4. Describe Delta Lake architecture
  5. Work with data streams by using Azure Stream Analytics
  6. Lab: Explore compute and storage options for data engineering workloads
    1. Combine streaming and batch processing with a single pipeline
    2. Organize the data lake into levels of file transformation
    3. Index data lake storage for query and workload acceleration
2. Design and implement the serving layer
  1. Design a multidimensional schema to optimize analytical workloads
  2. Code-free transformation at scale with Azure Data Factory
  3. Populate slowly changing dimensions in Azure Synapse Analytics pipelines
  4. Lab: Designing and Implementing the Serving Layer
    1. Design a star schema for analytical workloads
    2. Populate slowly changing dimensions with Azure Data Factory and mapping data flows
3. Data engineering considerations for source files
  1. Design a Modern Data Warehouse using Azure Synapse Analytics
  2. Secure a data warehouse in Azure Synapse Analytics
  3. Lab: Data engineering considerations
    1. Managing files in an Azure data lake
    2. Securing files stored in an Azure data lake
4. Run interactive queries using Azure Synapse Analytics serverless SQL pools
  1. Explore Azure Synapse serverless SQL pools capabilities
  2. Query data in the lake using Azure Synapse serverless SQL pools
  3. Create metadata objects in Azure Synapse serverless SQL pools
  4. Secure data and manage users in Azure Synapse serverless SQL pools
  5. Lab: Run interactive queries using serverless SQL pools
    1. Query Parquet data with serverless SQL pools
    2. Create external tables for Parquet and CSV files
    3. Create views with serverless SQL pools
    4. Secure access to data in a data lake when using serverless SQL pools
    5. Configure data lake security using Role-Based Access Control (RBAC) and Access Control List
5. Explore, transform, and load data into the Data Warehouse using Apache Spark
  1. Understand big data engineering with Apache Spark in Azure Synapse Analytics
  2. Ingest data with Apache Spark notebooks in Azure Synapse Analytics
  3. Transform data with DataFrames in Apache Spark Pools in Azure Synapse Analytics
  4. Integrate SQL and Apache Spark pools in Azure Synapse Analytics
  5. Lab: Explore, transform, and load data into the Data Warehouse using Apache Spark
    1. Perform Data Exploration in Synapse Studio

2. Ingest data with Spark notebooks in Azure Synapse Analytics
  3. Transform data with DataFrames in Spark pools in Azure Synapse Analytics
  4. Integrate SQL and Spark pools in Azure Synapse Analytics
6. Data exploration and transformation in Azure Databricks
  1. Describe Azure Databricks
  2. Read and write data in Azure Databricks
  3. Work with DataFrames in Azure Databricks
  4. Work with DataFrames advanced methods in Azure Databricks
  5. Lab: Data Exploration and Transformation in Azure Databricks
    1. Use DataFrames in Azure Databricks to explore and filter data
    2. Cache a DataFrame for faster subsequent queries
    3. Remove duplicate data
    4. Manipulate date/time values
    5. Remove and rename DataFrame columns
    6. Aggregate data stored in a DataFrame
7. Ingest and load data into the data warehouse
  1. Use data loading best practices in Azure Synapse Analytics
  2. Petabyte-scale ingestion with Azure Data Factory
  3. Lab: Ingest and load Data into the Data Warehouse
    1. Perform petabyte-scale ingestion with Azure Synapse Pipelines
    2. Import data with PolyBase and COPY using T-SQL
    3. Use data loading best practices in Azure Synapse Analytics
8. Transform data with Azure Data Factory or Azure Synapse Pipelines
  1. Data integration with Azure Data Factory or Azure Synapse Pipelines
  2. Code-free transformation at scale with Azure Data Factory or Azure Synapse Pipelines
  3. Lab: Transform Data with Azure Data Factory or Azure Synapse Pipelines
    1. Execute code-free transformations at scale with Azure Synapse Pipelines
    2. Create data pipeline to import poorly formatted CSV files
    3. Create Mapping Data Flows
9. Orchestrate data movement and transformation in Azure Synapse Pipelines
  1. Orchestrate data movement and transformation in Azure Data Factory
  2. Lab: Orchestrate data movement and transformation in Azure Synapse Pipelines
    1. Integrate Data from Notebooks with Azure Data Factory or Azure Synapse Pipelines
10. Optimize query performance with dedicated SQL pools in Azure Synapse
  1. Optimize data warehouse query performance in Azure Synapse Analytics
  2. Understand data warehouse developer features of Azure Synapse Analytics
  3. Lab: Optimize Query Performance with Dedicated SQL Pools in Azure Synapse
    1. Understand developer features of Azure Synapse Analytics
    2. Optimize data warehouse query performance in Azure Synapse Analytics
    3. Improve query performance
11. Analyze and Optimize Data Warehouse Storage
  1. Analyze and optimize data warehouse storage in Azure Synapse Analytics
  2. Lab: Analyze and Optimize Data Warehouse Storage
    1. Check for skewed data and space usage

2. Understand column store storage details
  3. Study the impact of materialized views
  4. Explore rules for minimally logged operations
12. Support Hybrid Transactional Analytical Processing (HTAP) with Azure Synapse Link
  1. Design hybrid transactional and analytical processing using Azure Synapse Analytics
  2. Configure Azure Synapse Link with Azure Cosmos DB
  3. Query Azure Cosmos DB with Apache Spark pools
  4. Query Azure Cosmos DB with serverless SQL pools
  5. Lab: Support Hybrid Transactional Analytical Processing (HTAP) with Azure Synapse Link
    1. Configure Azure Synapse Link with Azure Cosmos DB
    2. Query Azure Cosmos DB with Apache Spark for Synapse Analytics
    3. Query Azure Cosmos DB with serverless SQL pool for Azure Synapse Analytics
13. End-to-end security with Azure Synapse Analytics
  1. Secure a data warehouse in Azure Synapse Analytics
  2. Configure and manage secrets in Azure Key Vault
  3. Implement compliance controls for sensitive data
  4. Lab: End-to-end security with Azure Synapse Analytics
    1. Secure Azure Synapse Analytics supporting infrastructure
    2. Secure the Azure Synapse Analytics workspace and managed services
    3. Secure Azure Synapse Analytics workspace data
14. Real-time Stream Processing with Stream Analytics
  1. Enable reliable messaging for Big Data applications using Azure Event Hubs
  2. Work with data streams by using Azure Stream Analytics
  3. Ingest data streams with Azure Stream Analytics
  4. Lab: Real-time Stream Processing with Stream Analytics
    1. Use Stream Analytics to process real-time data from Event Hubs
    2. Use Stream Analytics windowing functions to build aggregates and output to Synapse Analytics
    3. Scale the Azure Stream Analytics job to increase throughput through partitioning
    4. Repartition the stream input to optimize parallelization
15. Create a Stream Processing Solution with Event Hubs and Azure Databricks
  1. Process streaming data with Azure Databricks structured streaming
  2. Lab: Create a Stream Processing Solution with Event Hubs and Azure Databricks
    1. Explore key features and uses of Structured Streaming
    2. Stream data from a file and write it out to a distributed file system
    3. Use sliding windows to aggregate over chunks of data rather than all data
    4. Apply watermarking to remove stale data
    5. Connect to Event Hubs read and write streams
16. Build reports using Power BI integration with Azure Synapse Analytics
  1. Create reports with Power BI using its integration with Azure Synapse Analytics
  2. Lab: Build reports using Power BI integration with Azure Synapse Analytics

1. Integrate an Azure Synapse workspace and Power BI
  2. Optimize integration with Power BI
  3. Improve query performance with materialized views and result-set caching
  4. Visualize data with SQL serverless and create a Power BI report
17. Perform Integrated Machine Learning Processes in Azure Synapse Analytics
1. Use the integrated machine learning process in Azure Synapse Analytics
  2. Lab: Perform Integrated Machine Learning Processes in Azure Synapse Analytics
    1. Create an Azure Machine Learning linked service
    2. Trigger an Auto ML experiment using data from a Spark table
    3. Enrich data using trained models
    4. Serve prediction results using Power BI

## Class Materials

Each student will receive a comprehensive set of materials, including course notes and all the class examples.

## Class Prerequisites

Experience in the following *is required* for this Azure class:

- Knowledge of cloud computing and core data concepts and professional experience with data solutions.

## Prerequisite Courses

Courses that can help you meet these prerequisites:

- [AZ-900T00 - Microsoft Azure Fundamentals](#)
- [DP-900T00 - Microsoft Azure Data Fundamentals](#)