

20776A : Performing Big Data Engineering on Microsoft Cloud Services

Course Description

This five-day instructor-led course teaches data engineers how to design, build, and maintain big data solutions on Microsoft Azure. Participants learn to process streaming and batch data, orchestrate workflows, manage large datasets, and integrate advanced analytics using Python and R.

Learning Objectives

By the end of this course, students will be able to:

- Design big data architectures using Azure services
- Process real-time event streams with Azure Stream Analytics
- Implement custom logic and integrate machine learning into streaming jobs
- Store, secure, and query large datasets with Azure Data Lake and U-SQL
- Orchestrate data movement and transformation pipelines with Azure Data Factory
- Load, query, and optimize performance in Azure SQL Data Warehouse
- Incorporate Python and R scripts into Azure workflows for advanced analytics

Duration: 5 days (40 Hours)

Prerequisites

- Basic knowledge of cloud computing concepts
- Familiarity with relational databases and SQL
- Experience with programming (preferably in Python, R, or C#)
- Understanding of fundamental data analysis concepts

Content Coverage

Module 1: Architectures for Big Data Engineering with Azure

- Big data architecture design
- Processing big data with Azure services

Module 2: Processing Event Streams using Azure Stream Analytics

- Introduction to Stream Analytics
- Configuring jobs

Module 3: Performing Custom Processing in Azure Stream Analytics

- Implementing custom functions
- Integrating ML models

Module 4: Working with Azure Data Lake

- Introduction to Data Lake Store
- Securing and managing data
- Processing data with U-SQL

Module 5: Processing Big Data using Azure Data Lake Analytics

- Introduction to Data Lake Analytics
- Running U-SQL queries
- Implementing custom operations
- Monitoring performance

Module 6: Managing Data with Azure SQL Data Warehouse

- Introduction to SQL DW
- Loading and querying datasets
- Performance optimization

Module 7: Performing Analytics with Azure SQL Data Warehouse

- Using SQL DW for analytics
- Query optimization and reporting

Module 8: Integrating Python and R with Azure Services

- Using Python and R for analytics
- Incorporating scripts into workflows

Labs Covered

- Processing Event Streams using Azure Stream Analytics
- Performing Custom Processing in Azure Stream Analytics
- Managing Big Data in Azure Data Lake Store
- Processing Big Data using Azure Data Lake Analytics
- Implementing Custom Operations and Monitoring Performance in Azure Data Lake Analytics
- Implementing Azure SQL Data Warehouse
- Performing Analytics with Azure SQL Data Warehouse