academy

DP-203: Data Engineering on Microsoft Azure [203T00-A] (EN)

Group Training

Training code	CGADP203CE
Spoken Language	English
Language Materials	English
Dayparts	8
Price	€2.400,00 excl. VAT No extra costs.

What is DP-203: Data Engineering on Microsoft Azure [203T00-A]

In this course, the student will learn about the data engineering patterns and practices as it pertains to working with batch and real-time analytical solutions using Azure data platform technologies. Students will begin by understanding the core compute and storage technologies that are used to build an analytical solution. They will then explore how to design an analytical serving layers and focus on data engineering considerations for working with source files. The students will learn how to interactively explore data stored in files in a data lake. They will learn the various ingestion techniques that can be used to load data using the Apache Spark capability found in Azure Synapse Analytics or Azure Databricks, or how to ingest using Azure Data Factory or Azure Synapse pipelines.

The students will also learn the various ways they can transform the data using the same technologies that is used to ingest data. The student will spend time on the course learning how to monitor and analyze the performance of analytical system so that they can optimize the performance of data loads, or queries that are issued against the systems. They will understand the importance of implementing security to ensure that the data is protected at rest or in transit. The student will then show how the data in an analytical system can be used to create dashboards, or build predictive models in Azure Synapse Analytics.

This course uses MOC (Microsoft Official Courseware) and will be given by an experienced MCT (Microsoft Certified Trainer).

See the below modules for more information:

Module 1: Explore compute and storage options for data engineering workloads

This module provides an overview of the Azure compute and storage technology options that are available to data engineers building analytical workloads. This module teaches ways to structure the data lake, and to optimize the files for exploration, streaming, and batch workloads. The student will learn how to organize the data lake into levels of data refinement as they transform files through batch and stream processing. Then

they will learn how to create indexes on their datasets, such as CSV, JSON, and Parquet files, and use them for potential query and workload acceleration.

Who should attend DP-203: Data Engineering on Microsoft Azure [203T00-A]

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 data analysts and data scientists who work with analytical solutions built on Microsoft Azure.

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

Objectives

After completing this course, you will be able to:
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 Synpase Analytics
- Perform Integrated Machine Learning Processes in Azure Synapse Analytics

If a third-party copyright applies to this course, you will find the copyright on
https://academy.capgemini.nl/en/topic/trademarks/

Capgemini Academy's general terms and conditions are applied to all products and services mentioned within this document. For the latest version please check https://academy.capgemini.com/. The rates of products and services mentioned in this document are subject to change. For the most recent rates, please also visit our website.

About Capgemini Academy

Capgemini Academy's professionals offer what people in IT need. Our professionals have a keen eye for motivation, talent and are aware of specific contexts and circumstances. They move people to move. Programmes and courses that originate from daily experience of our both didactical and substantively strong trainers, light a fire within the individual IT professionals.
Real life stories of our professionals' experience that tell how to solve problems and work with the people around it, do the rest.

An organization, like ours, helps people and their organizations day by day to get the best out of themselves and each other. We prepare them to defy tomorrow's challenges. We stimulate learning and curiosity. In order for individual IT professionals and their employers, to build better, longer and more intensive relationships. For mutual benefit.

Capgemini Academy. We transform IT professionals
academy.capgemini.nl

IN/3A-018.18