

training code: DP-203 / ENG DL 4d / EN

Data Engineering on Microsoft Azure

Authorized Data Engineering on Microsoft Azure training DP-203
Distance Learning course The training is addressed to:

- Administrator
- IT specialist
- Data engineer
- Database administrator
- Analyst



Purpose of the training

- Azure Synapse Analytics
- Azure Databricks
- Azure Data Lake Storage
- Azure Stream Analytics

Main audience of the course are data specialists, data architects and business analysis specialists who would like to know more about data engineering and creating analytical solutions with data platforms on Microsoft Azure. Another target audience are data analysts and scientists dealing with data who work with analytical solutions based on Microsoft Azure



Benefits of completing the training

At DP-203 course the student will be acquainted with patterns and practices of data engineering concerning analytical solutions in batch mode and real time with student Azure data platform technologies.

Students:

- will begin with understanding fundamental computation and mass storage technologies, which are used to create analytical solution.
- then will examine how to design analytical service layers and focus on data engineering issues while

working with source files.

- will learn how to interactiely explore data stored in files in data lake.
- will learn different acquisition techniques which can be used to load data with Apache Spark function available in Azure Synapse Analytics or Azure Databricks, or how to acquire data with Azure Data Factory lor Azure Synapse streams.
- will also learn various methods of data transfromation with the same technologies used to acquire data.

Student will devote his or her time to attend the course, by learning how to monitor and analyze the performance of naalytical system in order to be able to optimize performance of loading data or queries, sent to systems. They will understand how important it is to deploy securities in order to provide security of data at rest or during transfer. Next the student will show how data in analytical system will be used to create navigation desktops or develop prediction models in Azure Synapse Analytics.



Examination method

On-line exam. Record at: <https://home.pearsonvue.com/Clients/Microsoft.aspx>



Exam description

After the DP-203 course, you can take Microsoft certification exams:an Authorized Test Center,online being monitored by an offsite proctor. Details on the website:<https://learn.microsoft.com/en-us/certifications/exams/dp-203/>



Expected Listener Preparation

- Knowledge of cloud processing
- Knowledge of basic data concepts

Previous trainings: AZ-900, DP-900

To make your work more convenient and training more effective, we suggest using an additional screen. Lack of such screen does not exclude participation in the training; however, it significantly influences working comfort during classes.

You can find information and requirements of participation in Distance Learning trainings at:

<https://www.altkomakademia.pl/distance-learning/#FAQ>



Training Language

- **Training:** English
- **Materials:** English

Training Includes

- electronic handbook available at:
<https://learn.microsoft.com/pl-pl/training/>
- access to Altkom Akademia stu

Duration

4 days / 28 hours

Training agenda

1. Get to know computing and mass storage solutions for overloads related to data engineering

- Introduction to Azure Synapse Analytics
- Describe Azure Databricks
- Introduction to Azure Data Lake Storage
- Describe the architecture of Delta Lake
- Work with data streams with Azure Stream Analytics

2: Design and implement service layer

- Design multi-dimensional scheme to optimize analytical overloads
 - Non-code transformation on a huge scale thanks to Azure Data Factory
 - Complete slowly changing dimensions in Azure Synapse Analytics streams
- 3: Data engineering considerations for source files
- Design a Modern Data Warehouse using Azure Synapse Analytics
 - Secure a data warehouse in Azure Synapse Analytics
- 4: Launch interactive queries with serverless SQL Azure Synapse Analytics pools
- Be acquainted with serverless SQL pools in Azure Synapse
 - Performing data related queries in the lake with SQL Azure Synapse serverless pools
 - Develop metadata objects in serverless SQL Azure Synapse pools
 - Secure data and manage users in serverless SQL Azure Synapse pools
- 5: Exploring and transforming data in Azure Databricks
- Describe Azure Databricks
 - Reading and recording data in Azure Databricks
 - Work with DataFrames in Azure Databricks
 - Work with advanced DataFrames methods in Azure Databricks
- 6: Downloading and loading data to data warehouse
- Use the best solutions related to data loading in Azure Synapse Analytics
 - Acquiring within the scale of petabytes with Azure Data Factory
- 7: Transform data with Azure Data Factory or Azure Synapse Pipelines
- Data integration with Azure Data Factory or Azure Synapse Pipelines
 - Non-code transformation on a huge scale thanks to Azure Data Factory or Azure Synapse Pipelines
- 8: Transfer orchestration and data transformation in Azure Synapse Pipelines
- Organize transfer and data transformation in Azure Data Factory
- 9: Query performance optimization thanks to dedicated Azure Synapse pools
- Optimize data warehouse wydajność zapytań magazynu danych w usłudze Azure Synapse Analytics
 - Get acquainted with functions of data warehouse developers in Azure Synapse Analytics
- 10: Analysing and optimising data storage in data warehouse
- 11: Hybrid Transactional/Analytical Processing (HTAP) service with Azure Synapse link
- Project Hybrid Transactional/Analytical Processing with Azure Synapse Analytics
 - Skonfiguruj łącze Azure Synapse za pomocą usługi Azure Cosmos DB
 - Send queries to Azure Cosmos DB with Spark pools
 - Send queries to Azure Cosmos DB with serverless SQL pools
 - Analyze and optimize data warehouses in Azure Synapse Analytics
- 12: Complex securities thanks to Azure Synapse Analytics
- Secure data warehouses in Azure Synapse Analytics
 - Configure secret entries and manage them in Azure Key Vault
 - Implement sensitive data compatibility control
- 13: Stream processing in real-time with Stream Analytics service
- Activate reliable message service for Big Data applications with Azure Event Hubs
 - Working with data streams with Azure Stream Analytics

- Process data streams with Azure Stream Analytics

14: Creating stream processing with Event Hubs and Azure Databricks

- Process stream transfer data with structural stream Azure Databricks transfer

15: Creating reports using Power BI integration with Azure Synapse Analytics

- Create reports with Power BI service, using its integration with Azure Synapse Analytics

16: Performing integrated machine learning processes in Azure Synapse Analytics

- Use integrated machine learning process in Azure Synapse Analytics