

Data Engineering Workshops

Get familiar with Azure Data stack and connect the dots within 5 days!

1st Day

Data concepts, Azure Cloud Basics, Azure Storage

- Data Platform Concepts Data Warehouse, Modern Data Warehouse, Data Mesh, Lakehouse
- Azure Platform and Azure Portal basics
- Setup your Azure environment
- Azure Data Lake Storage Gen2, concepts, security, work with AzCopy, Azure Storage Explorer, Azure Portal
- How to avoid to build Data Swamp on ADLS

2nd Day

Database lifecycle and data orchestration in Azure

- Introduction to Azure SQL Database: Deployment models, purchasing models, service tiers
- Work with Azure SQL Database: Azure SQL configuration, work with tools: SSMS, SSDT, ADS
- Development of Azure SQL Database: understand database development lifecycle with SSDT
- Introduction to Azure Data Factory, understand concepts and build first ADF pipeline

3rd Day

Generic ETL and Azure Synapse Analytics

- Advanced Azure Data Factory pipeline development
- Introduction to Azure Synapse Analytics. Understand the concepts and features. SQL Dedicated, SQL Serverless, Apache Spark Pool
- Azure Synapse Analytics Studio, how to develop code, Distributed and Replicated tables
- CETAS, CTAS, how to work with data on ADLS

4th Day

Azure Databricks

- Introduction to Azure Databricks, discuss the concepts and architecture. Lakehouse and Medallion architecture, Databricks SQL, Delta Live Tables, Unity Catalog
- Get familiar with Azure Databricks Workspace
- Integrate Azure Databricks with Azure Key Vault, Azure SQL Database, Azure Synapse Analytics, Azure Data Lake Storage
- Run parametrized notebooks with ADF

5th Day

DevOps and Do it yourself!

- Introduction to Azure DevOps, general DevOps concepts and Azure DevOps features (including CI/CD)
- Integrate Azure SQL Database, Azure Data Factory and Azure Databricks code with GIT repository
- Do it yourself – exercise performed alone which summarize everything what was learned during workshops