



AZURE DATA ENGINEER

with

MICROSOFT FABRIC

&

DATA ANALYTICS

*Master the essential tools to become an in-demand
full-stack Azure Data Professional*

www.tcdemy.com

Azure for the Data Engineer

- Cloud Computing basics
- Introduction to Microsoft Azure
- Applications of Azure
- Azure Services
- Understanding Data
- On-premises vs Cloud-based servers
- Azure Data Engineer roles & tasks
- Data Engineering processes
- Use cases for the Cloud

Data Engineering Services on Azure

- Azure Datalake Storage (ADLS) gen2
- Azure Blob Storage
- Azure SQL database
- Azure Data Factory
- Azure Databricks
- Azure Synapse Analytics
- Microsoft Fabric

Data Storages on Azure

- File stores
- Relational data store
- Non-Relational data stores
- Azure Blob storage
- NoSQL & Azure Cosmos DB
- Azure Blob Storage
- Azure Data Lake Storage
- Why Data Lake
- Data Lake architecture
- Azure SQL Database
- Azure database for MySQL
- Azure database for MariaDB
- Azure database for PostgreSQL
- Azure Synapse Analytics

Azure Data Factory (ADF)

- | | |
|--|--|
| <ul style="list-style-type: none">• Introduction to Azure Data Factory (ADF)• ETL vs ELT• On-prem vs Cloud ETL tools• Why ADF• ADF Components• Linked Services in ADF• Integration Runtime (IR):<ul style="list-style-type: none">• Azure• Self-hosted• SSIS• Pipelines and activities:<ul style="list-style-type: none">• Copy data• Control flow• Data Flow• Data Ingestion methods:<ul style="list-style-type: none">• Full load• Copy data tool• Batch load• Sequential load• Incremental load• CDC | <ul style="list-style-type: none">• Working with various Copy data scenarios• Working with Azure SQL database• Working with Control flow activities• Working with Data Flow Transformations• Working with Parameters in ADF• SCD Type 1 & 2 Pipelines• ADF Data Migration• Scheduling & Triggering Pipelines• Monitoring Pipelines• CI/CD in ADF• ADF Projects |
|--|--|

Python

- | | |
|---|---|
| <ul style="list-style-type: none">• Variables• Datatypes in Python• Operators in Python• Arithmetic• Logical• Comparison• Working with Strings• Indexing• Slicing• Data Structures in Python:<ul style="list-style-type: none">• List• Tuples• Sets• Dictionaries | <ul style="list-style-type: none">• Functions:<ul style="list-style-type: none">• Built-in• User defined• Conditional Flow Statements:<ul style="list-style-type: none">• If condition• While loop• For iteration• Range() function• Map• Filter• Reduce• Lambda expressions |
|---|---|

Apache Spark

- | | |
|--|---|
| <ul style="list-style-type: none"> • Introduction to Apache Spark • Hadoop vs Spark • Understanding Spark Architecture • Spark Context • Spark APIs • Spark Data structures: <ul style="list-style-type: none"> • RDDs • DataFrames • Datasets • Creating Spark RDDs • Creating Spark DataFrames • Spark Transformations • Spark Actions • Lazy Evaluation • Fault Tolerance | <ul style="list-style-type: none"> • Column manipulations: <ul style="list-style-type: none"> • WithColumn • WithColumnRenamed • Partitioning data using Partition by • Reading & Writing data of various formats • PySpark commands • Spark SQL • Group by & Aggregate Functions • Window Functions • Joins in PySpark • Date Functions • Conditional Functions • Caching • Schema evolution • DAG |
|--|---|

Data Engineering with Azure Databricks

- | | |
|--|--|
| <ul style="list-style-type: none"> • Introduction to Azure Databricks • Azure Databricks vs Databricks Community Edition • Databricks Runtime Engine • Databricks Components: <ul style="list-style-type: none"> • Workspace • Cluster • Catalog • DBFS • Tables • Hive Metastore • DBFS • Working with Files using DBUTILS • Creating Tables: <ul style="list-style-type: none"> • using UI • using Notebooks • Reading data from different files • Writing data to different formats • Visualization in Databricks Notebooks | <ul style="list-style-type: none"> • Integrating Azure Blob with Databricks • Integrating ADLS with Databricks • Mounting with Access Keys • Secret Scopes using Azure Key Vault • Service Principle • SAS Tokens • Delta Lake • Creating Delta Tables: <ul style="list-style-type: none"> • Managed tables • External tables • Lakehouse • Medallion Architecture: <ul style="list-style-type: none"> • Bronze layer • Silver layer • Gold layer • Scheduling Databricks Jobs using ADF Pipelines • Delta Live Tables (DLT) • Unity Catalog |
|--|--|

Microsoft Fabric Data Engineering

- | | |
|---|--|
| <ul style="list-style-type: none">• Introduction to Microsoft Fabric• Fabric Analytics platform• Fabric OneLake• Fabric Data Factory:<ul style="list-style-type: none">• Data Pipelines• Dataflows gen2• Transforming data• Loading data to Lakehouse• Copy job• Create Pipeline in Fabric• Transform data with Dataflow gen2• ADF to Fabric migration• Fabric Synapse Data Engineering• Fabric Notebooks• Analyze data with Apache Spark• Lakehouse• Delta tables | <ul style="list-style-type: none">• Fabric Synapse Datawarehouse• Creating a Warehouse in Fabric• Ways to Load data into Warehouse• Creating Tables in Warehouse• Loading Data using COPY INTO Command• Loading Data using Pipeline to Warehouse• Loading Data using DataFlow Gen2• Lakehouse vs Warehouse• Microsoft Power BI in Fabric• Access data using PBI Desktop• Direct Lake• Direct Query• Row-level security• Column-level security |
|---|--|

Hands-on Real-time ADE Scenarios

- Creating ETL/ELT Pipelines in Azure Data Factory
- Working with Control Flow activities in ADF Pipelines
- Working with Data Flow Transformations in ADF Pipelines
- Ingesting multiple File sources using ADF pipelines
- Ingesting multiple SQL tables using ADF Pipeline
- Implementing Batch load & Sequential load
- Incremental loading using last modify date & time
- Incremental loading and Change data capture
- SCD1 pipeline in ADF
- SCD2 pipeline in ADF
- Empty File scenarios
- Total File count in folder
- Validate Schema & Structure of files
- Copy data from Rest APIs
- Email Notifications
- Analyzing data with PySpark
- Analyzing data with Spark SQL

- Analyzing ADLS data from Databricks Notebooks
- Mounting & Service Principle
- Creating Delta Tables
- Lakehouse (Medallion) architecture
- Data Factory in Fabric
- Fabric Pipelines
- Fabric Dataflows gen2 to transform data
- Lakehouse in Fabric
- Warehouse in Fabric
- Visualizing data in Power BI
- Creating Dashboard in Power BI

DATA ANALYTICS

SQL

- SQL commands for Data Analysis (DDL & DML)
- Aggregate Functions & Grouping data
- SQL Joins & Set Operators
- Advanced Data Analysis: Subqueries
- Data Cleaning & Transformation
- Window & Logical Functions
- String & Date Functions
- Analytical Functions

Data Analytics with Power BI

- Introduction to Datawarehouse
- OLTP vs OLAP
- Star Schema vs Snowflake Schema
- Overview of Power BI
- Power BI Desktop
- Connecting Power BI with different data sources
- Power Query for Data Transformation
- Data Modeling
- DAX
- Reports & Visualization types
- Dashboards
- Publishing Reports
- Power BI Service

Real-time Projects

- ✓ ADF Data Migration Project (migrating on-prem SQL Server data to Azure)
- ✓ Incremental loading using Last Modified Date/Time
- ✓ Incremental loading using CDC
- ✓ SCD-2 Pipeline in ADF
- ✓ End-to-End Databricks Project (Medallion Architecture)
- ✓ Microsoft Fabric Project
- ✓ Power BI Project

Bonuses

- ADE Interview Questions and Answers
- Mock Interviews
- Placement Assistance
- Fabric Certification (DP700) Support
- Databricks certification support
- Resume/CV Preparation
- LinkedIn Profile Optimization