Azure Data engineer Training Course Outline

What is Cloud computing and Azure Data factory

ADF Pipeline

What is pipeline and why we using.

- Copy activity (Loading data from different files, tables and creation of files and tables)
- Explain importance of Storages and Database in ADF projects, as pre project wise scenarios
- Handling the files, and tables using Variables and parameters.
- ADF control flows utilization and other activities
- Get meta data, for each, filter, switch, copy, delete, script,if else activity etc.
- Explain the importance of control flow and other activities as per the project wise end to end explanations.

Data flow Transformation-Code less transformation

- Exists Transformation,Union Transformation,Join Transformation, Aggregate,Surrogate Key,Select,Lookup,Derived Column,Pivot,Unpivot,Rank,Window and Filter
- Above transformation are explained as per the project wise scenarios

SQL(ADD ons)

- Basic of SQL
- Explaining why we used SQL and required SQL covering helps to understand writing scripting in ADF.

Python(ADD ons)

- Basic of Python
- Explaining why we used Python and required concepts covering helps to understand writing scripting in ADF.

Data Brick-Pyspark

- What is Databrick and why we using it
- What is Pyspark and its components.
- Transformations using
- Mount to ADLS or Blob from Azure Data-bricks using PySpark
- Pyspark ingestion methods, Pyspark transformations and Aggregation,
- SQL utilization in Pyspark.
- Key difference between spark and pyspark.

Azure Synapse Analytics Data warehouse

- Create a Synapse workspace
- Analyze using server-less SQL pool
- Analyze using a Data Explorer pool
- Analyze using a server-less Spark pool
- Analyze using a dedicated SQL pool
- Analyze data in a storage account
- Pipleline and Dataflow key highlights
- Apache Spark in Synapse.

Microsoft Fabric Training Course Outline

Module 1: Introduction to End-to-End Analytics Using Microsoft Fabric

- Introduction
- Explore End-to-End Analytics with Microsoft Fabric
- Data Teams and Microsoft Fabric
- Enable and Use Microsoft Fabric

Module 2: Get Started with Lakehouses in Microsoft Fabric

- Introduction
- Explore the Microsoft Fabric Lakehouse
- Work with Microsoft Fabric Lakehouses

Exercise - Create and Ingest Data with a Microsoft Fabric Lakehouse

Module 3: Use Apache Spark in Microsoft Fabric

- Introduction
- Prepare to Use Apache Spark
- Run Spark Code
- Work with Data in a Spark Dataframe
- Work with Data Using Spark SQL
- Visualise Data in a Spark Notebook
- Exercise - Analyse Data with Apache Spark

Module 4: Work with Delta Lake Tables in Microsoft Fabric

- Introduction
- Understand Delta Lake
- Create Delta Tables
- Work with Delta Tables in Spark

Module 5: Use Data Factory Pipelines in Microsoft Fabric

- Introduction
- Understand Pipelines
- Use the Copy Data Activity
- Use Pipeline Templates
- Run and Monitor Pipelines
- Exercise - Ingest Data with a Pipeline

Module 6: Ingest Data with Dataflows Gen2 in Microsoft Fabric

- Introduction
- Understand Dataflows (Gen2) in Microsoft Fabric
- Explore Dataflows (Gen2) in Microsoft Fabric
- Integrate Dataflows (Gen2) and Pipelines in Microsoft Fabric
- Exercise - Create and Use a Dataflow (Gen2) in Microsoft Fabric

Module 7: Get Started with Data Warehouses in Microsoft Fabric

- Introduction
- Understand Data Warehouse Fundamentals
- Understand Data Warehouses in Fabric

Key-highlights

- As per student profile customize the contents.
- Covering Basic to advanced topics
- Real-time projects wise scenarios on training
- Resume making & Mock interview sessions
- Assignment assigning to students and followup
- one to one session for students queries
- Mentor guidance till getting a job.
- Real-time projects & hands-on labs
- Microsoft Certification guidance