

DP - 601T00 : Implement a Lakehouse with Microsoft Fabric

Course Outline

Module: Introduction to End-to-End Analytics using Microsoft Fabric

1. Introduction to End-to-End Analytics

- **Describe end-to-end analytics in Microsoft Fabric:**
Understand how Microsoft Fabric provides a unified platform for ingesting, storing, analyzing, and visualizing data using integrated services.

2. Get Started with Lakehouses in Microsoft Fabric

Key Concepts:

- **Core Features & Capabilities:**
 - Combines data warehouse and data lake features
 - Supports open formats like Parquet and Delta
 - Built-in support for Spark and SQL queries
- **Create a Lakehouse:**
 - Use the Fabric UI to define a lakehouse with structured and unstructured data zones
- **Ingest Data:**
 - Load data into files and tables using built-in tools, pipelines, or notebooks
- **Query with SQL:**
 - Run T-SQL or Spark SQL queries on lakehouse tables

Lab: Create a Microsoft Fabric lakehouse

3. Use Apache Spark in Microsoft Fabric

Key Concepts:

- Configure Spark in a Fabric workspace
- Identify scenarios best suited for Spark notebooks and Spark jobs
- Use DataFrames to manipulate and transform data
- Query data using Spark SQL
- Visualize data within notebooks using display libraries

Lab: Analyze data with Apache Spark

4. Work with Delta Lake Tables

Key Concepts:

- Understand what Delta Lake and delta tables are in Fabric
- Create and manage delta tables using Spark
- Optimize delta tables using compaction and Z-ordering
- Transform and query delta tables using Spark
- Use Spark structured streaming to process data in real time

Lab: Use delta tables in Apache Spark

5. Ingest Data with Dataflows Gen2

Key Concepts:

- **Dataflow Capabilities:**

- No-code/low-code ETL tool
 - Built for data transformation and ingestion
- **Create Dataflow Solutions:**
 - Build reusable transformation pipelines with Power Query-like interface
- **Include in Pipelines:**
 - Integrate Dataflows into larger orchestrated processes

SpireTec Solutions