Microsoft Fabric Workshop

Course Description:

This 4-day comprehensive training program is designed to provide participants with an in-depth understanding of Microsoft Fabric, a cutting-edge unified data platform that seamlessly integrates data engineering, warehousing, real-time analytics, and data science. The training begins with an introduction to Microsoft Fabric's architecture, tools, and data integration processes, before moving into advanced topics like Apache Spark, Delta Lakes, Power BI reporting, and real-time data science solutions.

The course emphasizes practical, hands-on learning to ensure participants not only grasp the theoretical concepts but also gain experience through labs. Participants will work on end-to-end data pipelines, data warehousing setups, machine learning model training, and real-time analytics in Microsoft Fabric. By the end of the course, they will have mastered the tools and techniques to implement robust data solutions that drive innovation and operational efficiency in their organizations.

Target Audience:

This course is designed for professionals who are either working with or planning to adopt Microsoft Fabric to handle their data engineering, data warehousing, analytics, and data science needs. The following groups will benefit the most from this course:

- **Data Engineers** looking to leverage the unified power of Microsoft Fabric for scalable data integration and processing using Spark and Delta Lakes.
- **Data Analysts** who want to build dynamic reports, integrate with Power BI, and streamline data-driven insights.
- **Data Scientists** seeking to implement machine learning models and real-time analytics in a fully integrated environment.
- **BI Developers and Data Warehouse Specialists** who want to design, monitor, and optimize data warehouses and ensure seamless data ingestion and querying.
- **IT Administrators** responsible for managing Microsoft Fabric environments, ensuring security, and overseeing workspace management.

Prerequisites:

- **Basic understanding of data engineering concepts** such as data pipelines, data lakes, and data warehouses.
- **Familiarity with Microsoft Azure or cloud platforms** is beneficial but not mandatory.

- **Working knowledge of Apache Spark, SQL, or Python** is recommended, as the course involves hands-on labs with data engineering and processing tasks.
- **Experience with Power BI or other data visualization tools** will help participants in the modules focusing on reporting and business intelligence.

For participants without prior exposure to these concepts, introductory resources will be recommended before the training to ensure a smooth learning experience.

Content Coverage :

Day 1:

Module 1: Introduction to Microsoft Fabric

- Overview of Microsoft Fabric
- Microsoft Fabric Terminology
- Copilot in Microsoft Fabric
- Self-help with the Fabric contextual Help pane
- Searching for content
- Microsoft Fabric settings
- Working with Workspaces
- Discovering data in OneLake and data hub
- Promoting or certifying items
- Administering Microsoft Fabric

Module 2: Getting started with Data Integration and Engineering Experiences using Microsoft Fabric

- Overview of Data Engineering in Microsoft Fabric
- Lakehouse Overview
- Ways to create a Lakehouse in Microsoft Fabric
- Getting data into Fabric Lakehouse
- Data Factory in Microsoft Fabric
- Working with Data Pipelines in Microsoft Fabric
- Ingesting Data with Dataflows Gen2

- Introduction to Spark compute in Microsoft Fabric
- Spark Administration settings

Day 2:

Module 3: Data Engineering using Apache Spark, Delta Lakes, and Notebooks

- Apache Spark job definition
- Apache Spark monitoring in Microsoft Fabric
- Delta Lake table optimization and V-Order
- How to use Microsoft Fabric notebooks
- Synapse Visual Studio Code extension
- Managing a Workspace with Git
- Decision Guide for copy activity, dataflow, or Spark
- Lakehouse end-to-end scenario and architecture

Hands-on Labs:

- Creating a Lakehouse, ingesting sample data, and building a report
- Analyzing data with Apache Spark
- Using Delta tables in Apache Spark
- Creating and using a Dataflow (Gen2) in Microsoft Fabric

Day 3:

Module 4: Data Warehousing, Power BI Reporting in Microsoft Fabric

- Introduction to Data Warehousing
- Creating a Data Warehouse in Microsoft Fabric
- Data Warehouse connectivity
- Decision guide: Data Warehouse or Lakehouse
- Better together: The Lakehouse and Warehouse
- Data ingestion options
- Security and Performance
- Querying and monitoring the Data Warehouse

- Creating models and building Power BI reports
- Building reports at blazing speed with Direct Lake mode

Hands-on Labs:

- Building a Data Warehouse in Microsoft Fabric

Day 4:

Module 5: Implementing Real-Time Analytics and Data Science in Microsoft Fabric

- Introduction to Data Science in Microsoft Fabric
- Preparing Data
- Training machine learning models
- Tracking models, experiments, and Autologging
- Model Scoring
- Overview of Real-Time Analytics in Microsoft Fabric
- Event Streams in Microsoft Fabric
- Creating a KQL Database
- Getting data to KQL Database from Data Factory
- Querying data from a KQL Queryset
- Visualizing your KQL Database data
- Driving alerts and actions from your data with Data Activator

Hands-on Labs:

- Training and tracking a model in Microsoft Fabric
- Implementing a Data Science scenario in Microsoft Fabric
- Developing, evaluating, and scoring a forecasting model for superstore sales
- Performing Real-Time Analytics using Microsoft Fabric