



## DP-601: Implementing a Lakehouse with Microsoft Fabric

Microsoft - Data & AI

- **Nível:** Intermédio
  - **Duração:** 7h
- 

### Sobre o curso

**O presente curso é concebido de forma a fornecer aos participantes competências fundamentais em engenharia de dados no Microsoft Fabric, com um foco especial no conceito de Lakehouse.**

Neste curso explorar-se-á as poderosas capacidades do Apache Spark no que respeita o processamento de dados distribuídos e as técnicas essenciais para proceder à gestão eficiente de dados, incluindo confiabilidade e versionamento, ao trabalhar-se com tabelas do Delta Lake. Abordar-se-á também a ingestão e orquestração de dados utilizando Dataflows Gen2 e pipelines do Data Factory. O curso é composto por aulas teóricas e exercícios práticos que prepararão os participantes para trabalharem com lakehouses no Microsoft Fabric.

---

### Destinatários

O público-alvo principal deste curso são profissionais da área de dados (Analistas, Engenheiros e Cientistas de Dados) já familiarizados com a modelação, extração e analítica de dados. Este é também destinado a profissionais interessado em adquirir conhecimentos acerca da arquitetura Lakehouse e da plataforma Microsoft Fabric, e de como executar analítica end-to-end usando estas duas tecnologias.

---

### Pré-requisitos

You should be familiar with basic data concepts and terminology.

---

### Programa

- Introdução à analítica end-to-end usando o Microsoft Fabric
- Introdução aos Lakehouses no Microsoft Fabric
- Utilização do Apache Spark no Microsoft Fabric
- Trabalhar com tabelas Delta Lake no Microsoft Fabric
- Ingestão de dados com Dataflows Gen2 no Microsoft Fabric
- Orquestração de processos e movimentação de dados com o Microsoft Fabric
- Organização de um lakehouse do Fabric usando o design de arquitetura medallion

## **Introdução à analítica end-to-end usando o Microsoft Fabric**

Descubra como o Microsoft Fabric pode atender às necessidades de analítica da sua empresa numa única plataforma. Fique a saber mais sobre o funcionamento do Microsoft Fabric e identifique como o mesmo pode usado para o ajudar nas suas necessidades de analítica.

- Explorar a analítica end-to-end com o Microsoft Fabric
- Equipas de dados e Microsoft Fabric
- Ativar e usar o Microsoft Fabric

## **Introdução aos Lakehouses no Microsoft Fabric**

Os Lakehouses combinam a flexibilidade do armazenamento de data lakes com a analítica de data warehouses. O Microsoft Fabric oferece uma solução de lakehouse para procedimentos de analítica abrangentes numa única plataforma SaaS.

- Explorar o Microsoft Fabric Lakehouse
- Trabalhar com Lakehouses no Microsoft Fabric
- Exercício - Criar e ingerir dados com um Lakehouse do Microsoft Fabric

## **Utilização do Apache Spark no Microsoft Fabric**

O Apache Spark é uma tecnologia essencial para procedimentos de analítica de dados de larga escala. O Microsoft Fabric oferece apoio a clusters Spark e permite que se analise e processe dados num Lakehouse de grande escala.

- Preparar-se para usar o Apache Spark
- Executar código Spark
- Trabalhar com dados num dataframe do Spark
- Trabalhar com dados usando o Spark SQL
- Visualizar dados em um notebook do Spark
- Exercício - Analisar dados com o Apache Spark

## **Trabalhar com tabelas Delta Lake no Microsoft Fabric**

As tabelas em um lakehouse do Microsoft Fabric são baseadas no formato de armazenamento Delta Lake, comumente usado no Apache Spark. A utilização das capacidades aperfeiçoadas das tabelas delta permite criar soluções de analítica avançada.

- Compreender o Delta Lake
- Criar tabelas delta
- Trabalhar com tabelas delta no Spark
- Usar tabelas delta com dados em streaming
- Exercício - Usar tabelas delta no Apache Spark

## **Ingestão de dados com Dataflows Gen2 no Microsoft Fabric**

A ingestão de dados é crucial em analítica. O Data Factory do Microsoft Fabric oferece Dataflows (Gen2) que permitem a criação visual da ingestão e transformação de dados em múltiplos passos recorrendo ao Power

Query Online.

- Compreender Dataflows (Gen2) no Microsoft Fabric
- Explorar Dataflows (Gen2) no Microsoft Fabric
- Integrar Dataflows (Gen2) e Pipelines no Microsoft Fabric
- Exercício - Criar e usar um Dataflow (Gen2) no Microsoft Fabric

### **Orquestração de processos e movimentação de dados com o Microsoft Fabric**

O Microsoft Fabric inclui recursos do Data Factory, incluindo a capacidade de criar pipelines que orquestram tarefas de ingestão e transformação de dados.

- Compreender pipelines
- Usar a atividade Copiar Dados
- Usar templates de pipeline
- Executar e monitorizar pipelines
- Exercício - Ingerir dados com um pipeline

### **Organização de um lakehouse do Fabric usando o design de arquitetura medallion**

Explore o potencial do design de arquitetura medallion no Microsoft Fabric. Organize e transforme os seus dados através das camadas Bronze, Silver e Gold de um lakehouse de forma a obter procedimentos de analítica otimizados.

- Descrever a arquitetura medallion
- Implementar uma arquitetura medallion no Fabric
- Consultar e criar relatórios sobre dados num lakehouse do Fabric
- Considerações para gerir um lakehouse