



## DP-750: Implement data engineering solutions using Azure Databricks

Microsoft - Data & AI

- **Nível:**
  - **Duração:** 28h
- 

### Sobre o curso

**Domina a engenharia de dados end-to-end com Azure Databricks e Unity Catalog, desde a configuração do ambiente até à implementação em produção.**

Este curso percorre todo o processo, desde a configuração inicial até à implementação em produção, abrangendo a configuração do ambiente e a governação de nível empresarial. Aprende a criar pipelines de ingestão robustos, implementar segurança com Unity Catalog e disponibilizar workloads otimizados.

No final, terás competências práticas para implementar, proteger e manter soluções lakehouse escaláveis que respondem a requisitos empresariais exigentes.

---

### Destinatários

- Data Engineers
- 

### Objetivos

- Configurar e preparar um ambiente Azure Databricks para workloads de engenharia de dados.
  - Proteger e governar objetos Unity Catalog no Azure Databricks.
  - Preparar e processar dados com Azure Databricks.
  - Implementar e manter pipelines de dados e workloads com Azure Databricks.
  - Desenhar modelos de dados, ingerir dados, transformar datasets e aplicar controlos de qualidade de dados.
  - Implementar práticas de ciclo de vida de desenvolvimento, monitorização, troubleshooting e otimização de workloads no Azure Databricks.
- 

### Condições

**Para particulares**

- 10% do valor total pago no ato da inscrição; restante valor até 7 dias antes do início do curso.
- Formandos não residentes em Portugal: pagamento de 50% no ato da inscrição; restante valor até 7 dias antes do início do curso.
- Possibilidade de pagamento faseado.
- Possibilidade de beneficiar do Cheque Formação+Digital até 750€ (conforme elegibilidade).
- Isenção de IVA para particulares.

### **Para empresas**

- Empresas nacionais: pagamento a 30 dias, contra fatura (acresce IVA à taxa legal em vigor).
- Empresas da UE e fora da UE: valores isentos de IVA e pagamento a pronto.

---

## Pré-requisitos

- Conhecimentos fundamentais de conceitos de data analytics.
- Conhecimentos básicos de cloud storage.
- Familiaridade com princípios de organização de dados.
- À-vontade a trabalhar com SQL.
- Experiência na utilização de Python, incluindo notebooks, para tarefas de engenharia de dados.
- Boa compreensão de workspaces Azure Databricks e de Unity Catalog.
- Familiaridade com padrões de acesso a dados e conceitos fundamentais de engenharia de dados e data warehouse.
- Conhecimentos base de segurança no Azure, incluindo Microsoft Entra ID.
- Familiaridade com fundamentos de controlo de versões com Git.

---

## Programa

- Configurar e preparar um ambiente Azure Databricks
- Proteger e governar objetos do Unity Catalog no Azure Databricks
- Preparar e processar dados com Azure Databricks
- Implementar e manter pipelines de dados e workloads com Azure Databricks

### **Configurar e preparar um ambiente Azure Databricks**

Este módulo introduz a plataforma Azure Databricks e o seu papel na engenharia de dados moderna. Foca-se na arquitetura do serviço, integrações, opções de compute e organização de objetos em Unity Catalog, para compreenderes como preparar um workspace para workloads escaláveis de analytics, engenharia de dados e IA no Azure.

- Explorar o Azure Databricks
- Compreender a arquitetura do Azure Databricks
- Compreender as integrações do Azure Databricks
- Selecionar e configurar compute no Azure Databricks
- Criar e organizar objetos no Unity Catalog

## **Proteger e governar objetos do Unity Catalog no Azure Databricks**

Este módulo aborda as capacidades de segurança e governação do Unity Catalog no Azure Databricks. Trabalha controlo de acessos, permissões granulares, gestão de credenciais, data lineage, audit logging e práticas de partilha segura, para protegeres e gerires ativos de dados em ambientes empresariais.

- Proteger objetos do Unity Catalog
- Governar objetos do Unity Catalog

## **Preparar e processar dados com Azure Databricks**

Este módulo foca-se no workflow central de engenharia de dados em Azure Databricks. Abrange modelação de dados, ingestão para Unity Catalog, limpeza, transformação, carregamento e implementação de restrições de qualidade de dados, para construíres pipelines fiáveis e dados prontos para analytics numa arquitetura lakehouse.

- Desenhar e implementar modelação de dados com Azure Databricks
- Ingerir dados para o Unity Catalog
- Limpar, transformar e carregar dados no Unity Catalog
- Implementar e gerir restrições de qualidade de dados com Azure Databricks

## **Implementar e manter pipelines de dados e workloads com Azure Databricks**

Este módulo aborda a implementação e operação contínua de pipelines e workloads em Azure Databricks. Inclui desenho de pipelines, Lakeflow Jobs, processos de ciclo de vida de desenvolvimento e monitorização de workloads, para gerires orquestração, fiabilidade, troubleshooting e otimização de desempenho em produção.

- Desenhar e implementar pipelines de dados com Azure Databricks
- Implementar Lakeflow Jobs com Azure Databricks
- Implementar processos de ciclo de vida de desenvolvimento no Azure Databricks
- Monitorizar, fazer troubleshooting e otimizar workloads no Azure Databricks