



*Fundamentos de Engenharia de Dados*

# Fundamentos de Engenharia de Dados

Plataforma Databricks

# Fundamentos de Engenharia de Dados

Curso Online, Gratuito, Em Português e com Certificado de Conclusão

[www.datascienceacademy.com.br](http://www.datascienceacademy.com.br)



Databricks é uma plataforma de análise e inteligência artificial baseada em nuvem que pode ser usada em Engenharia de Dados para ajudar a criar, gerenciar e analisar grandes conjuntos de dados.

Aqui estão algumas maneiras de usar a plataforma Databricks em Engenharia de Dados:

**Processamento de dados em larga escala:** A plataforma Databricks é construída sobre o Apache Spark, um framework de processamento de dados em larga escala. Isso permite que a plataforma seja usada para processar grandes conjuntos de dados de maneira eficiente, escalável e distribuída.

**Implantação de soluções de dados:** A plataforma Databricks pode ser usada para implantar soluções de dados em nuvem. A plataforma permite que os usuários criem ambientes de desenvolvimento e teste de maneira fácil e rápida, além de implantar soluções de dados em produção em um ambiente de nuvem seguro e escalável.

**Análise de dados:** A plataforma Databricks oferece recursos de análise de dados avançados, permitindo que os usuários executem análises complexas em grandes conjuntos de dados. A plataforma também oferece recursos de visualização de dados, permitindo que os usuários criem gráficos e dashboards para visualizar os resultados das análises.

**Aprendizado de máquina:** A plataforma Databricks inclui recursos de aprendizado de máquina (Machine Learning), permitindo que os usuários treinem modelos de aprendizado de máquina em grandes conjuntos de dados. A plataforma oferece recursos de automação de fluxo

de trabalho, permitindo que os usuários criem fluxos de trabalho de aprendizado de máquina complexos e escaláveis.

**Colaboração:** A plataforma Databricks é projetada para facilitar a colaboração entre equipes de Engenharia de Dados e Cientistas de Dados. A plataforma permite que os usuários compartilhem código, dados e notebooks, além de colaborar em tempo real em projetos de dados.

Databricks é uma plataforma poderosa para Engenharia de Dados, oferecendo recursos de processamento de dados em larga escala, implantação de soluções de dados em nuvem, análise de dados, aprendizado de máquina e colaboração. Ao usar a plataforma Databricks, as empresas podem processar grandes conjuntos de dados de maneira eficiente, criar soluções de dados escaláveis e seguras, executar análises avançadas e treinar modelos de aprendizado de máquina. Visite o site oficial da plataforma:

<https://www.databricks.com>

A plataforma Databricks faz parte de diversos projetos da Formação Engenheiro de Dados e Arquiteto de Dados aqui na DSA.