

# ***01.O que é Apache Airflow?***

quinta-feira, 11 de janeiro de 2024

15:52

- **Projeto criado pela Airbnb em 2015 - Open source**
- **Desenvolvido em Python**
- **Orquestrador a pipeline de dados**
- **Coordena o processamento**
- **Em batch**
- **O processamento ocorre em:**
  - SO
  - Banco de dados
  - Spark
  - Elastic Search
  - Etc.
- **2 interfaces web - Web e CLI**
- **As Dags são desenvolvidas em códigos python**
- **Dentro das dags colocamos a pipeline**
- **Por que usar a Dag:**
  - Gerencia erros
  - Log/Auditoria
  - Monitorar/Alertas
  - Recuperação a partir de um Ponto
  - Dados Históricos
  - Alta disponibilidade
  - Distribuição de Carga