

Mais PythonOperator

- Permite adicionar qualquer funcionalidade Python ao Airflow:
 - Limpeza e tratamento de dados
 - Transformações, resumos
 - Extração de fontes diversas
 - Machine Learning
- Você pode usar qualquer módulo Python



Prática

- Vamos adicionar um path para uma pasta dados em docker-compose.yaml
- Colocamos o arquivo Churn.csv nesta pasta
 - Este arquivo está junto com o material do curso. Olhe na pasta desta seção!
- Vamos criar uma DAG que:
 - Lê este arquivo
 - Faz diversos tratamentos:
 - Preenche NAs com mediana ou com Moda
 - Exclui duplicados
 - Salva como novo arquivo

