

# Instalação do Airflow

---

Vamos partir para a instalação do Airflow!

Ele assemelha-se a uma biblioteca ou um aplicativo, por isso, às vezes, sua instalação pode ser um pouco complexa.

No terminal, acessaremos a pasta que criamos anteriormente, "airflow" (cd airflow) e ativaremos o ambiente virtual (source venv/bin/activate) para uso de bibliotecas específicas para esse projeto.

```
cd airflow
source venv/bin/activate
```

Criação do ambiente virtual:

```
python3.9 -m venv venv
```

Ativação do ambiente virtual:

```
source venv/bin/activate
```

Agora, vamos instalar o Airflow. Utilizaremos o pacote pip, do Python, para instalá-lo e passaremos a versão desejada, 2.3.2. Em seguida, passaremos o arquivo de restrição, que especifica todas as versões das bibliotecas que o Airflow precisa para funcionar corretamente, então adicionaremos --constraint e a url que direciona para este arquivo.

```
pip install 'apache-airflow==2.3.2' --constraint
"https://raw.githubusercontent.com/apache/airflow/constraints-
2.3.2/constraints-3.9.txt"
```

Observe que na url constam a versão do Airflow que estamos instalando e a versão do Python que estamos utilizando.

Após o carregamento da instalação do Airflow, podemos iniciá-lo localmente, mas, antes, precisamos exportar a variável de ambiente Airflow Home. A exportaremos com o diretório no qual instalamos o Airflow, ou seja, o caminho da localização. Vale ressaltar que é necessário exportá-la sempre que formos executar o Airflow.

```
export AIRFLOW_HOME=~/.Documents/airflow
```

Agora, podemos executar o comando responsável por iniciar o banco de dados padrão do Airflow, criar um novo usuário e senha para que possamos acessá-lo, e executar os dois principais serviços do Airflow: o scheduler e o web services.

```
airflow standalone
```

Podemos observar, no log de execução, que há um trecho indicando que o Airflow está pronto ("Airflow is ready") seguido de um usuário ("admin") e uma senha, que consta em "password". Caso você não consiga identificar a senha, é possível encontrá-la na pasta de execução do Airflow, no arquivo chamado "standalone\_admin\_password.txt" ou no "docker-compose.yaml" na tag *AIRFLOW\_WWW\_USER*.

No navegador, acessaremos "localhost:8080", e logaremos com o usuário e senha que nos foram fornecidos para acessar a interface do Airflow.

## Comando airflow standalone

Quando queremos executar o Airflow, nós temos a opção de executar cada um de seus componentes de forma separada, inicializando o banco de dados do Airflow, executando o scheduler e o webserver. Tudo de forma manual. No entanto, também temos a opção de utilizar o comando:

```
airflow standalone
```

Esse comando já faz tudo de forma mais automática, ou seja, inicia o banco de dados padrão do airflow, cria um novo usuário e inicia os serviços principais (o webserver e o scheduler).

Para finalizar a execução desses processos localmente, basta acessar o terminal onde o Airflow está sendo executado e pressionar Ctrl + C. É importante finalizar o processo antes de fechar o terminal para não termos problemas quando executarmos o Airflow novamente.

Para mais informações, você pode acessar a documentação do Airflow: Rodando o [Airflow localmente](#).

## Interface do Airflow

A interface de usuário do Airflow é um recurso que apresenta informações sobre os DAGs e DAG Runs. Ela é uma ferramenta muito útil para entender, monitorar e solucionar problemas de nossos pipelines.

Nos vídeos anteriores, nós já tivemos uma visão geral de alguns dos recursos e visualizações mais úteis na IU do Airflow. Existem algumas outras visualizações de DAG que não abordaremos em profundidade aqui, mas é bom conhecermos. São elas:

Task Duration: apresenta um gráfico de linhas com a duração de cada tarefa ao longo do tempo;  
Task Tries: apresenta um gráfico de linhas com o número de tentativas para cada tarefa em um DAG Run ao longo do tempo;  
Landing Times: apresenta um gráfico de linhas com a hora e o dia em que cada tarefa foi iniciada ao longo do tempo;

Gantt: apresenta um gráfico de Gantt com a duração de cada tarefa para um DAG Run específico.

Para mais informações sobre a interface do Airflow, acesse a documentação: [UI/Screenshots](#).