

Atividade Spark transformando dados

- iniciar o cluster

-> docker-compose up -d

Habilitar a porta para ter acesso o jupyter notebook

1 – Com o dataframe df_stage criado, criar as colunas do calendário

Ex(<https://acervolima.com/adicionar-nova-coluna-com-valor-padrao-no-dataframe-pyspark/>)

- Ano
- Mês
- Dia
- Trimestre