

Exercício Hive

```
docker exec -it hive-server bash
```

```
cd /input
```

```
cd home/input/dados
```

```
curl -O https://raw.githubusercontent.com/caiuafranca/dados_curso/main/cursos.csv
```

2. Mover os dados para o hdfs:

```
user/aluno/dados/curso
```

3. no banco de dados “treinamento”

4. Criar a Tabela Hive no BD treinamento

- Tabela externa: cursos_stg

Campos:

```
id_curso,id_unidade,codigo,nome,nivel,id_modalidade_educacao,id_municipio,
id_tipo_oferta_curso,id_area_curso,id_grau_academico,id_eixo_conhecimento,ativo
```

Propriedades

- Delimitadores: Campo ‘;’ | Linha ‘\n’
- Local ‘user/aluno/dados/curso’
- Sem Partição
- Tipo do arquivo: Texto

5. Visualizar a descrição da tabela cursos_stg

6. Verificar as 5 primeiras linhas da tabela cursos_stg

7. Contar linhas da tabela cursos_stg

8. Criar a Tabela Hive no BD treinamento

- Tabela interna: cursos

Campos:

id_curso,id_unidade,codigo,nome,nivel,id_modalidade_educacao,id_municipio,i
d_tipo_oferta_curso,id_area_curso,id_grau_academico,id_eixo_conhecimento,
ativo,dt_foto

OBS. Observar na origem o que é texto o que é numero

Propriedade

- Partição = dt_foto

8. Inserir dados da tabela cursos_stg para cursos

9. verificar partições da tabela

10. contar quantos registros foi inserido

11. contar a quantidade cursos ativo