

Comandos Hive

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

```
cd /input
mkdir dados
cd dados
curl -O https://github.com/caiuafranca/dados_curso/blob/main/alunos.csv
```

```
hdfs dfs -mkdir -p /user/aluno/dados/alunos
```

```
hdfs dfs -copyFromLocal alunos.csv /user/aluno/dados/alunos
```

```
beeline --help
beeline -u jdbc:hive2://localhost:10000
```

```
show databases;
```

```
create database treinamento;
```

```
=====
=====

CREATE EXTERNAL TABLE IF NOT EXISTS treinamento.alunos_stg(
  `id_discente` string,
  `nome` string,
  `ano_ingresso` string,
  `periodo_ingresso` string,
  `nivel` string,
  `id_forma_ingresso` string,
  `id_curso` string
)
COMMENT 'Tabela Externa de Alunos'
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ','
STORED AS TEXTFILE
LOCATION '/user/aluno/dados'
tblproperties ('skip.header.line.count'='1', 'store.charset'='UTF-8',
'retrieve.charset'='UTF-8');
```

```
=====
=====

create table treinamento.alunos (
  `id_discente` string,
  `nome` string,
  `ano_ingresso` int,
  `periodo_ingresso` int,
  `nivel` string,
  `id_forma_ingresso` string,
  `id_curso` string
)
PARTITIONED BY (DT_FOTO STRING)
ROW FORMAT SERDE 'org.apache.hadoop.hive.ql.io.orc.OrcSerde'
```

```
STORED AS INPUTFORMAT 'org.apache.hadoop.hive.ql.io.orc.OrcInputFormat'
OUTPUTFORMAT 'org.apache.hadoop.hive.ql.io.orc.OrcOutputFormat'
TBLPROPERTIES ( 'orc.compress'='SNAPPY');
```

```
SET hive.exec.dynamic.partition=true;
SET hive.exec.dynamic.partition.mode=nonstrict;
```

```
INSERT OVERWRITE TABLE
treinamento.alunos
```

```
PARTITION(DT_FOTO)
SELECT
id_discente,
nome,
ano_ingresso,
periodo_ingresso,
nivel,
id_forma_ingresso,
id_curso,
'21052022' as DT_FOTO
FROM treinamento.alunos_stg;
```

```
describe treinamento.alunos;
```

```
describe formatted treinamento.alunos;
```

```
msck repair table treinamento.alunos;
```