

	Training – Practice Guide 5	
	Trainee:	Date:

OBJETIVOS

- Crear una tabla en Impala con los datos ingestados de la base de datos

GUIA

1. Copiar la ruta del primer archivo parquet en directorio HDFS

```
hdfs dfs -ls /user/<tu_nombre>/orden_trabajo
```

2. Acceder la consola de Hue o impala-shell (consola línea de comando)

3. Crear una base de datos en Impala con tu nombre

```
create database <tu_nombre>;
```

4. Crear la tabla externa

```
create external table < nombre>.orden like parquet '<ruta del parquet>' stored as parquet location 'orden_trabajo';
```

Ejemplo ruta parquet: /user/<tu_nombre>/orden_trabajo/6b672ea3-45fa-4fca-b4ad-47617245b76e.parquet

5. Describir los campos de la tabla

```
describe <tu_nombre>.orden;
```

6. Ejecutar queries libres con los datos utilizados.

Ejemplo: select * from <tu_nombre>.orden;

Ejemplo: select count(*) from <tu_nombre>taller.orden;

Ejemplo: select tipo_orden_nombre, count(*) from table <tu_nombre>.orden group by tipo_orden_nombre;

REFERENCIA

https://www.cloudera.com/documentation/enterprise/5-8-x/topics/impala_create_table.html