



<http://sqoop.apache.org/>

Importación de datos de MySQL con apache Sqoop

Para preparar la máquina Cloudera:

```
$DEVSH/scripts/catchup.sh
```

Importación de tablas

```
sqoop help
```

```
sqoop list-tables --connect jdbc:mysql://localhost/loudacre \  
--username training --password training
```

```
sqoop help import
```

```
sqoop import --connect jdbc:mysql://localhost/loudacre \  
--username training --password training \  
--table accounts --target-dir /loudacre/accounts \  
--null-non-string '\\N'
```

Visualización de los datos importados

```
hdfs dfs -ls /loudacre/accounts
```

```
hdfs dfs -tail /loudacre/accounts/part-m-00000
```

```
hdfs dfs -tail /loudacre/accounts/part-m-00001
```

```
hdfs dfs -tail /loudacre/accounts/part-m-00002
```

```
hdfs dfs -tail /loudacre/accounts/part-m-00003
```

Importación a formato Parquet

```
sqoop import --connect jdbc:mysql://localhost/loudacre \  
--username training --password training \  
--table accounts --target-dir /loudacre/accounts_parquet \  
--as-parquetfile
```

Para trabajar con el formato parquet existen herramientas específicas como parquet-tools.

```
parquet-tools head \  
hdfs://localhost/loudacre/accounts_parquet/nombreArchivo.parquet
```