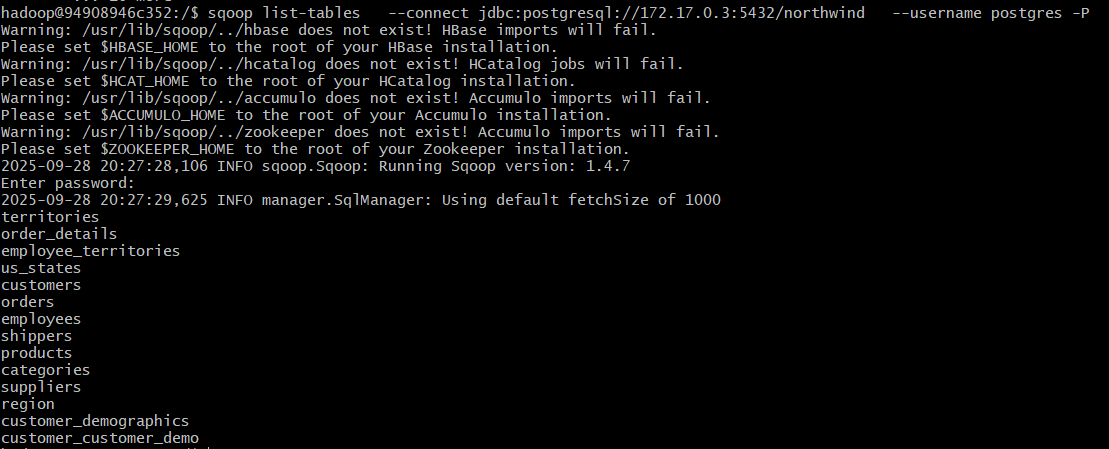
**SQOOP**

1) Mostrar las tablas de la base de datos northwind

sqoop list-tables \

--connect jdbc:postgresql://172.17.0.3:5432/northwind \

--username postgres -P



2) Mostrar los clientes de Argentina

sqoop eval \

--connect jdbc:postgresql://172.17.0.3:5432/northwind \

--username postgres \

--P \

--query "SELECT \* FROM customers WHERE country = 'Argentina'"



3) Importar un archivo .parquet que contenga toda la tabla orders. Luego ingestar el archivo a HDFS (carpeta /sqoop/ingest)

sqoop import \

--connect jdbc:postgresql://172.17.0.3:5432/northwind \

--username postgres \

--table orders \

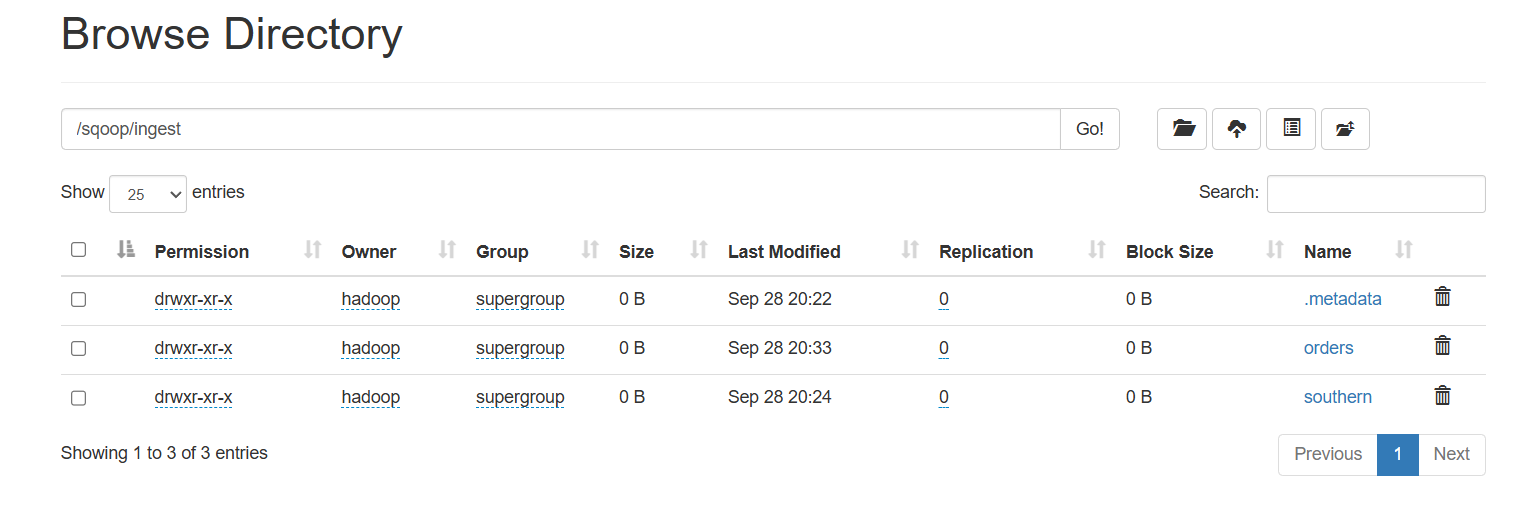
-m 1 \

--P \

--target-dir /sqoop/ingest/orders \

--as-parquetfile \

--delete-target-dir



4) Importar un archivo .parquet que contenga solo los productos con mas 20 unidades en stock, de la tabla Products . Luego ingestar el archivo a HDFS (carpeta ingest)

sqoop import \

--connect jdbc:postgresql://172.17.0.3:5432/northwind \

--username postgres \

--table products \

-m 1 \

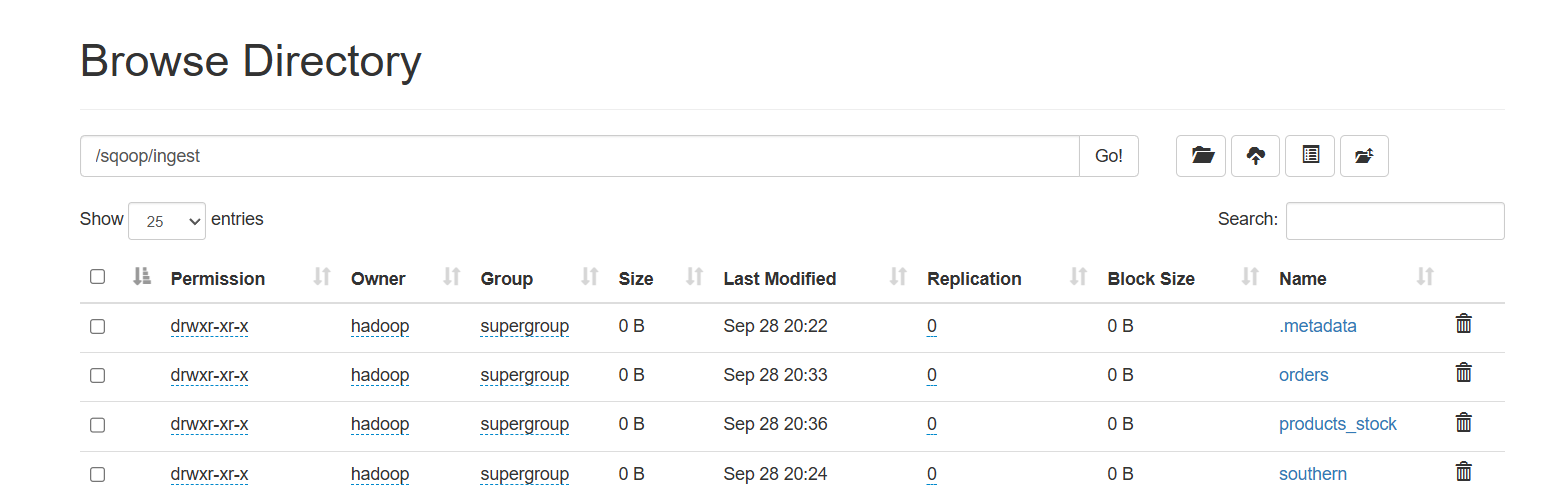
--P \

--target-dir /sqoop/ingest/products\_stock \

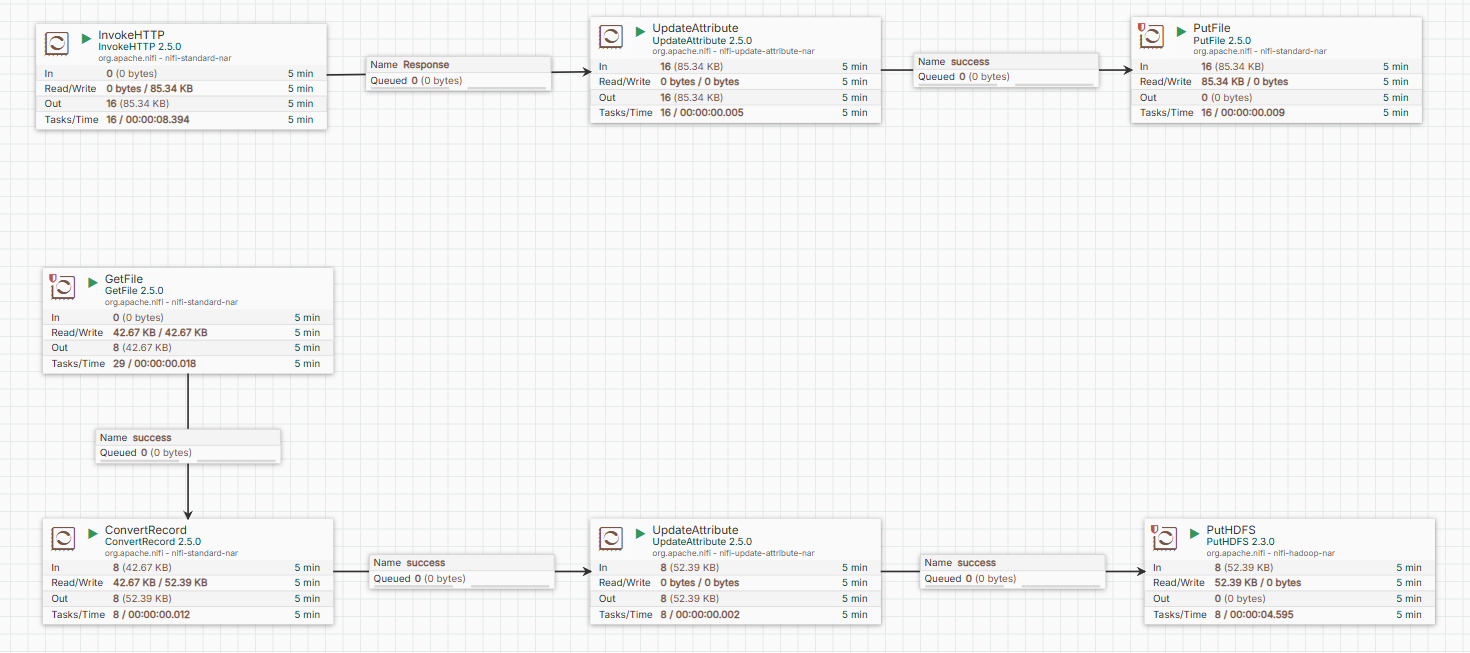
--as-parquetfile \

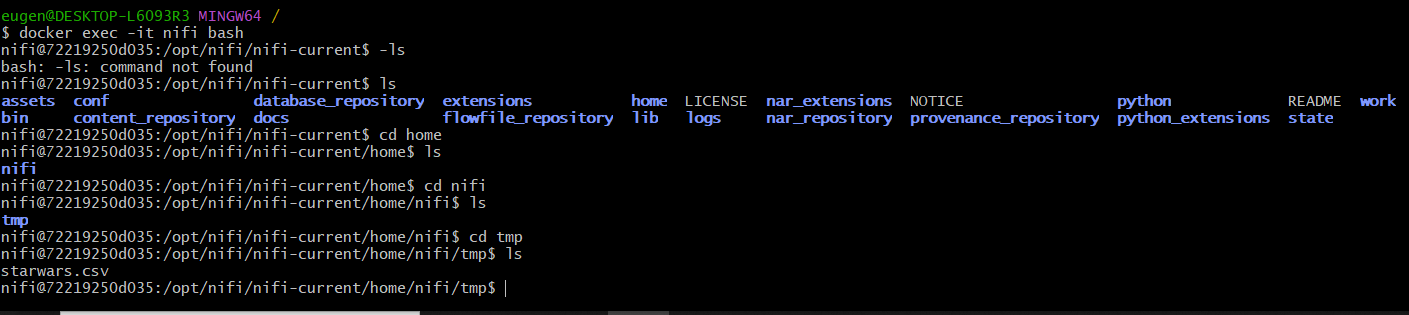
--where "units\_in\_stock > 20" \

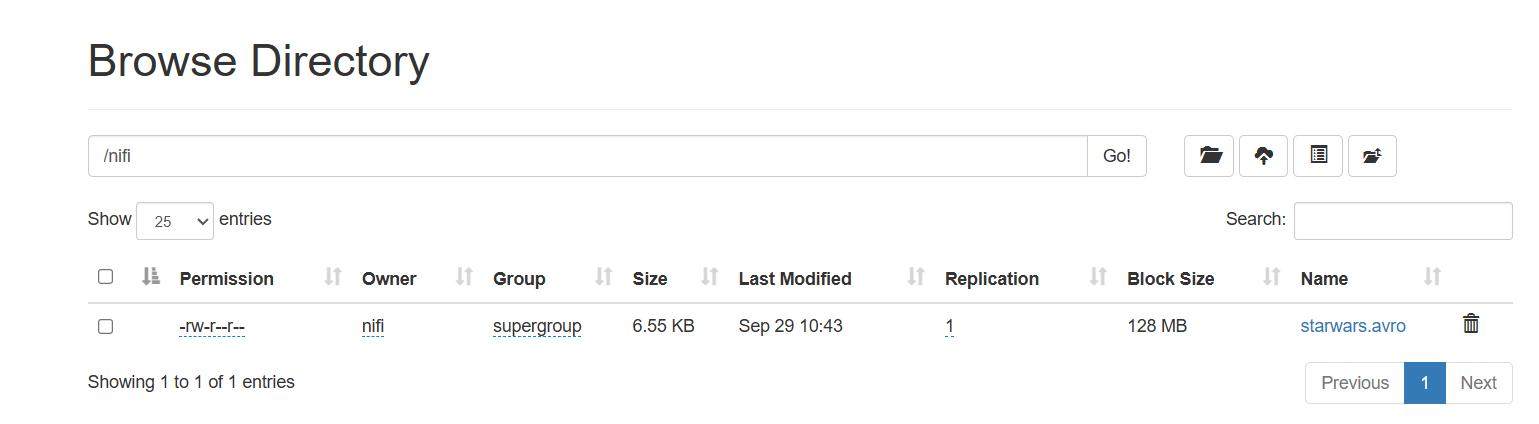
--delete-target-dir

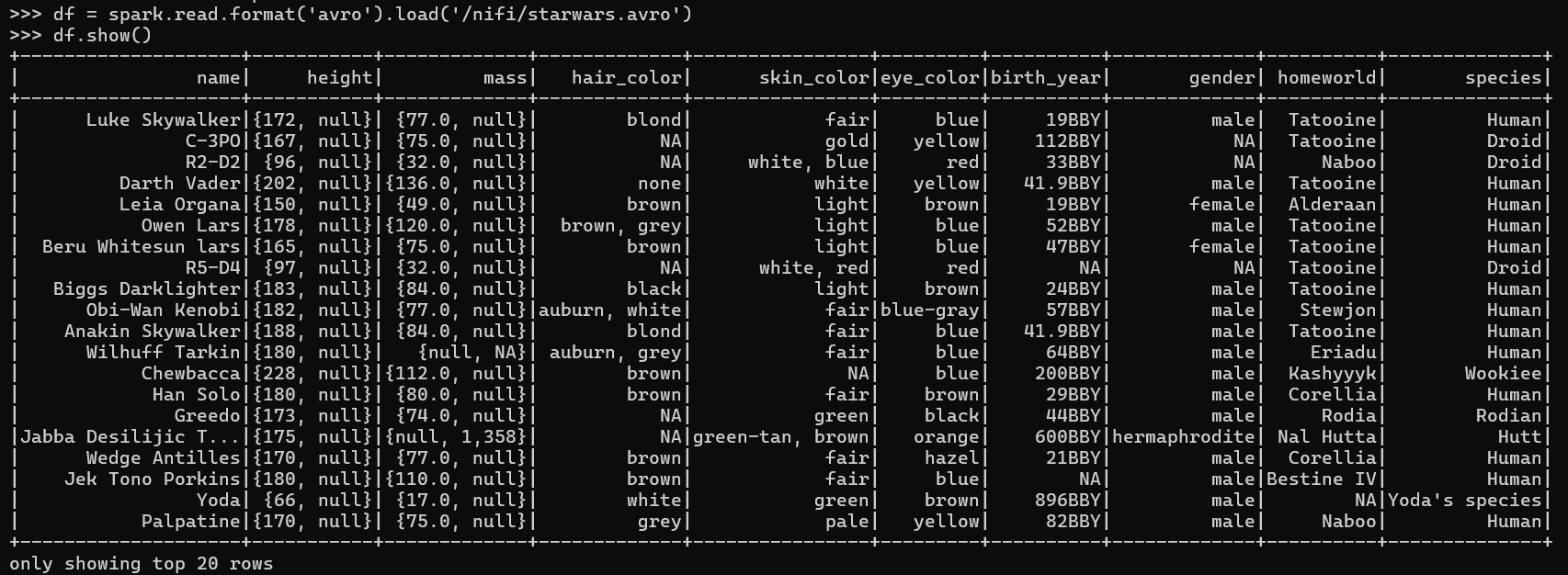


**NIFI**









\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Tuve este problema, lo documento por si vuelve a suceder:

## **🔹 Problema inicial**

NiFi no podía escribir en HDFS y devolvía:

File ... could only be written to 0 of the 1 minReplication nodes.

There are 0 datanode(s) running

👉 Esto significa que el **NameNode estaba vivo**, pero **ningún DataNode** estaba conectado → no había dónde guardar los bloques.

# **📝 Checklist: “0 datanode(s) running” en HDFS**

### **1️⃣ Verificar estado de procesos**

jps

* ¿Ves NameNode, SecondaryNameNode y DataNode?
* Si **no aparece DataNode** → problema confirmado.

### **2️⃣ Verificar estado del cluster**

hdfs dfsadmin -report

* Si dice Live datanodes (0) → ningún DataNode está activo.

### **3️⃣ Revisar logs del DataNode**

cd ~/hadoop/logs

tail -n 50 hadoop-hadoop-datanode-\*.log

* Buscá **ERROR** o **FATAL**.  
  Caso típico: Incompatible clusterIDs ... namenode clusterID = X ; datanode clusterID = Y

### **4️⃣ Solución: resetear directorios**

**Parar HDFS** stop-dfs.sh

**Borrar datos del DataNode** rm -rf /home/hadoop/hadoopdata/hdfs/datanode

mkdir -p /home/hadoop/hadoopdata/hdfs/datanode

chown -R hadoop:hadoop /home/hadoop/hadoopdata

### **5️⃣ Reiniciar y verificar**

start-dfs.sh

jps

hdfs dfsadmin -report

* Ahora debería aparecer DataNode en jps
* Y Live datanodes (1) en el reporte. ✅

**🚀 Con esto:**

* NiFi ya puede escribir con PutHDFS en /nifi.
* Evitás el error *“could only be written to 0 of the 1 minReplication nodes”*.