Импортировал таблицу work

**sqoop import --connect jdbc:postgresql://node3.novalocal/pg\_db --username exporter -P --table work --hive-import --hive-database carbon\_postgre\_import --hive-table work\_parquet --as-parquetfile**



Посмотрел размер

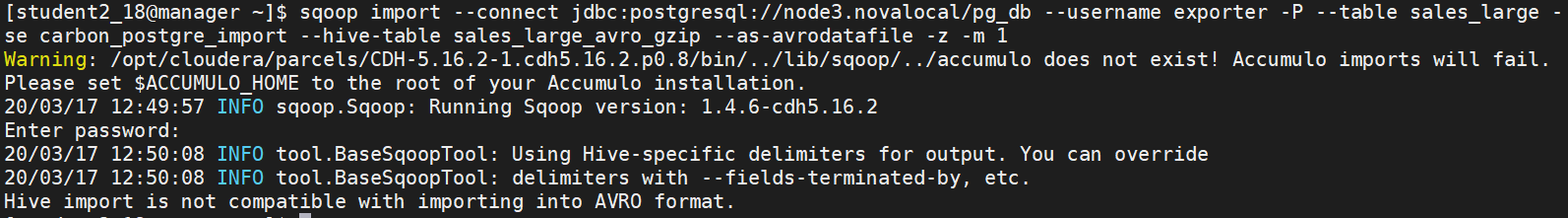


Сделал запросы на эту таблицу



Импортировал таблицу sales\_large, хотел импортировать в avro, не получилось, пришлось в паркет, но название таблицы забыл поменять, осталось sales\_large\_avro ☺

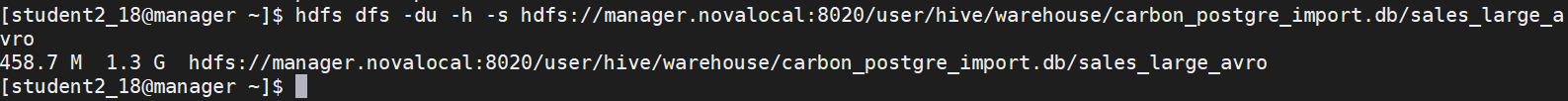
Здесь попытка импортировать в AVRO, написало, что несовместимый формат ☹



Для демонстрации компрессии импортировал sales\_large в формате паркет, но с компрессией gzip

**sqoop import --connect jdbc:postgresql://node3.novalocal/pg\_db --username exporter -P --table sales\_large --hive-import --hive-database carbon\_postgre\_import --hive-table sales\_large\_parquet\_gzip --as-parquetfile -z -m 1**

По умолчанию тип компрессии, что был выше - Snappy, размер этих файлов: 458 mb



При gzip компрессии: ……. А нет gzip компрессии, хоть я и указал ее вроде ключом –z. Соответственно, тип остался snappy

