Examples in this section rely on the table stored in Hive during the previous module on Hive.

Check that the data are available

Start PySpark, check if this is working by running a simple query on the customer table:

```
sqlCtx.sql("SELECT customer_fname, customer_lname FROM customers").limit(2).collect()
```

The result should be:

```
[Row(customer_fname=u'Richard', customer_lname=u'Hernandez'),
 Row(customer_fname=u'Mary', customer_lname=u'Barrett')]
```

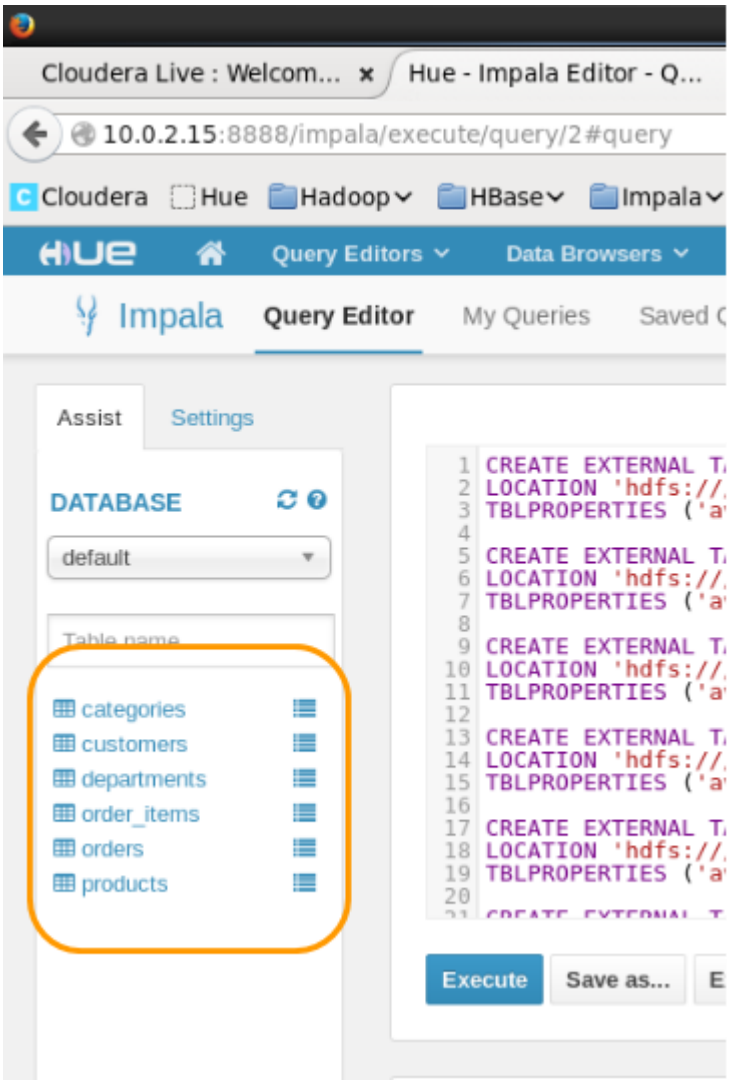
Insert data from the Cloudera Quickstart

ONLY if the previous step fails, you need to follow again the instructions on the Cloudera Quickstart to import data into Hive.

1. open browser inside the Cloudera VM
2. Click on the Getting started icon on the bookmarks bar, last item on the right
3. Click on "Start Tutorial"
4. Execute the commands in Tutorial 1 and Tutorial 2

Finally verify that the tables have been created in Hue

1. open browser inside the VM
2. connect to the address quickstart.cloudera:8888 (or look for the link to Hue)
3. Check that the tables show up on the left panel (hit the refresh button if they don't), see screenshot:


Cloudera Live : Welcom...
Hue - Impala Editor - Q...
10.0.2.15:8888/impala/execute/query/2#query
Cloudera
Hue
Hadoop
HBase
Impala
Query Editors
Data Browsers
Impala
Query Editor
My Queries
Assist
Settings
DATABASE
default
Table name
categories
customers
departments
order_items
orders
products
CREATE EXTERNAL T
LOCATION 'hdfs://
TBLPROPERTIES ('a
Execute
Save as...