

## *Hoja de Trabajo No. 6*

6.

```
[cloudera@quickstart ~]$ hdfs dfs -put /home/cloudera/Desktop/ejercicios/data/ratings_2012.txt analyst/dualcore/  
[cloudera@quickstart ~]$ hdfs dfs -put /home/cloudera/Desktop/ejercicios/data/ratings_2013.txt analyst/dualcore/  
[cloudera@quickstart ~]$
```

7.

The screenshot shows the Hue web interface. At the top, there is a navigation bar with the Hue logo, a 'Query' button, a 'Search data' input field, and links for 'Jobs', a refresh icon, and a user profile labeled 'cloudera'. Below the navigation bar, the breadcrumb path is '/ratings\_2012.txt'. The main section is titled 'SOURCE' and contains two input fields: 'Type' with a dropdown menu set to 'File', and 'Path' with the text '/user/cloudera/analyst/dualcore/ratings\_2012.txt'. At the bottom, there is a section titled 'FORMAT'.

9 y 12.

Load data in analyst.ratings

×

INFO : Loading data to table analyst.ratings from hdfs://quickstart.cloudera:8020/user/cloudera/analyst/dualcore/ratings\_2012.txt

INFO : Starting task [Stage-1:STATS] in serial mode

INFO : Table analyst.ratings stats: [numFiles=1, totalSize=27007]

INFO : Completed executing command(queryId=hive\_20200306162121\_67759ede-9438-41c8-933d-7f1548d6707c); Time taken: 0.98 seconds

INFO : OK

Load data in analyst.ratings

3 minutes ago

!

Creating table analyst.ratings

5 minutes ago

✓

Drop table ratings

10 minutes ago

!

Creating table analyst.ratings

Load data in analyst.ratings

×

INFO : Loading data to table analyst.ratings from hdfs://quickstart.cloudera:8020/user/cloudera/analyst/dualcore/ratings\_2013.txt

INFO : Starting task [Stage-1:STATS] in serial mode

INFO : Table analyst.ratings stats: [numFiles=2, totalSize=1267251]

INFO : Completed executing command(queryId=hive\_20200306162222\_85a31c50-9e66-47f4-8cf1-258bc86f6f02); Time taken: 0.536 seconds

INFO : OK

Load data in analyst.ratings

a minute ago

✓

Load data in analyst.ratings

3 minutes ago

!

Creating table analyst.ratings

6 minutes ago

✓

Drop table ratings

10.

	posted	cust_id	prod_id	rating	message
1	2012-05-21 12:52:48	1043182	1274362	5	This is truly fantastic!
2	2012-10-14 01:36:07	1242853	1273879	2	The product quality was OK
3	2012-10-14 02:41:50	1047430	1273799	2	Shoddy quality
4	2012-10-14 10:10:05	1087455	1274476	4	Quality was passable
5	2012-10-14 10:42:41	1170230	1273964	2	It was OK
6	2012-10-14 19:12:33	1063130	1274734	4	It was OK
7	2012-10-14 22:00:56	1031378	1274616	4	Quality was passable
8	2012-10-15 00:27:47	1203215	1273850	5	Awesome product
9	2012-10-15 01:14:26	1135616	1274218	4	Value of product was iust alriht

11 y 14. \*El resultado ya incluye ratings 2012 y 2013

	count(*)
1	21997

15.



```
0s default text
6 city string,
7 state string,
8 zipcode string,
9 job_title string,
10 email string,
11 active string,
12 salary int
13 )
14 location '/cloudera/analyst/dualcore/employee'

***

✓ Success.
```

16.

```
SELECT job_title, COUNT(*) AS num
FROM employees
GROUP BY job_title
```

17.

```
hive> sqoop import --connect jdbc:mysql://localhost/analyst_dualcore --username root --password cloudera --fields-terminated-by '\t' --warehouse-dir analyst\dualcore --table suppliers--hive-import--create-hive-table--hive-table analyst.suppliers
```

```
OK
Time taken: 4.76 seconds
Loading data to table analyst.suppliers
Table analyst.suppliers stats: [numFiles=4, totalSize=6060]
OK
Time taken: 0.544 seconds_
```

18.

Welcome to the Impala shell.

(Impala Shell v2.10.0-cdh5.13.0 (2511805) built on Wed Oct 4 10:55:37 PDT 2017)

After running a query, type SUMMARY to see a summary of where time was spent.

\*\*\*\*\*

[quickstart.cloudera:21000] > USE analyst;

Query: use analyst

[quickstart.cloudera:21000] > INVALIDATE METADATA suppliers;

Query: invalidate METADATA suppliers

Query submitted at: 2020-03-19 21:03:00 (Coordinator: http://quickstart.cloudera:25000)

Query progress can be monitored at: [http://quickstart.cloudera:25000/query\\_plan?query\\_id=2c4e66f3efd437bb:2b64b23600000000](http://quickstart.cloudera:25000/query_plan?query_id=2c4e66f3efd437bb:2b64b23600000000)

Fetches 0 row(s) in 0.19s

<  analyst

Tables

(6)  

 customers

 order\_details

 orders


 products

 ratings

 suppliers

count(\*)

19. `ALTER TABLE analyst.suppliers CHANGE company na`

 **suppliers**  
supp\_id (int)  
name (string)

20.

1	supp_id	int
2	name	string
3	contact	string
4	address	string
5	city	string
6	state	string
7	zipcode	string
8	phone	string

21.

`ALTER TABLE suppliers RENAME TO vendors`

 **vendors**

22.

```
| SELECT supp_id, name FROM vendors LIMIT 10;|
```

	supp_id	name
1	1000	Smithco Equipment Manufacturing
2	1001	Simian Tech Systems, Inc.
3	1002	Division Zero Corp.
4	1003	General Diode Co.
5	1004	JPD Companies
6	1005	Harvey Mettle Electrical Supply Co.
7	1006	The Connector Company