Assignment

HBase – Hive Connectivity

Create Hbase Table:

Create 'sales_hbase', 'sale_details', 'product_details', 'contact_details'

Create Hive table:

hive> create external table if not exists sales_hive

> (

> key int

> ,

> quantityordered int,

> priceeach float,

> orderlinenumber int,

> sales float,

> status string,

> qtr_id int,

> month_id int,

> year_id int,

> productline string,

> msrp int,

> productcode string,

> phone int,

> city string,

> state string,

> postalcode string,

> country string,

> territory string,

> contactlastname string,

> contactfirstname string,

> dealsize string

```
> )

> row format delimited

>  fields terminated by ','

> stored by 'org.apache.hadoop.hive.hbase.HBaseStorageHandler'

> with serdeproperties

> ("hbase.columns.mapping"=":key,sale_details:quantityordered, sale_details:priceeach,
sale_details:orderlinenumber, sale_details:sales, sale_details:status, sale_details:qtr_id,
sale_details:month_id,
sale_details:year_id,product_details:productline,product_details:msrp,product_details:productco
de,contact_details:phone,contact_details:city,contact_details:state,contact_details:postalcode,co
ntact_details:country,contact_details:territory,contact_details:contactlastname,contact_details:co
ntactfirstname, sale_details:dealsize")

> tblproperties("hbase.table.name"="sales_hbase");
```

OK

Time taken: 3.163 seconds

Create Hive stage table/ dummy table:

```
hive> create table sales_order_csv

> (

> key int,

> quantityordered int,

> priceeach float,

> orderlinenumber int,

> sales float,

> status string,

> qtr_id int,

> month_id int,

> year_id int,

> productline string,

> msrp int,

> productcode string,

> phone string,

> city string,

> state string,
```

> **postalcode string,**

> **country string,**

> **territory string,**

> **contactlastname string,**

> **contactfirstname string,**

> **dealsize string**

> **)**

> **row format delimited**

> **fields terminated by ','**

> **tblproperties("skip.header.line.count"="1");**

OK

Time taken: 6.163 seconds

sales_order_data.csv file at '/home/cloudera' location.

Loading the data into a stage table:

hive> load data local inpath 'file:///home/cloudera/sales_order_data.csv' into table sales_order_csv;

Loading data to table default.sales_order_csv

Table default.sales_order_csv stats: [numFiles=1, totalSize=360233]

OK

Time taken: 1.798 seconds

Insert data from stage table to main table:

hive> insert into table sales_hive select * from sales_order_csv;

Query ID = cloudera_20221031053333_4a5bf535-b7a4-4fcd-978b-c3f26787a60a

Total jobs = 1

Launching Job 1 out of 1

Number of reduce tasks is set to 0 since there's no reduce operator

Starting Job = job_1667216548526_0001, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1667216548526_0001/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1667216548526_0001

Hadoop job information for Stage-0: number of mappers: 1; number of reducers: 0

2022-10-31 05:33:39,640 Stage-0 map = 0%, reduce = 0%

2022-10-31 05:34:03,806 Stage-0 map = 100%, reduce = 0%, Cumulative CPU 11.7 sec

MapReduce Total cumulative CPU time: 11 seconds 700 msec

Ended Job = job_1667216548526_0001

MapReduce Jobs Launched:

Stage-Stage-0: Map: 1   Cumulative CPU: 11.7 sec   HDFS Read: 375309 HDFS Write: 0 SUCCESS

Total MapReduce CPU Time Spent: 11 seconds 700 msec

OK

Time taken: 57.599 seconds

Fetch some records from hive main table:

hive> select * from sales_hive limit 5;

OK

10100	49	34.47	1	1689.03	Shipped	1	1	2003	Vintage Cars		41
	S24_3969	NULLNashua		NH	62005	USA	NA	Young	Valarie	Small	
10101	46	53.76	2	2472.96	Shipped	1	1	2003	Vintage Cars		44
	S24_2022	NULLFrankfurt			60528	Germany	EMEA	Keitel	Roland	Small	
10102	41	50.14	1	2055.74	Shipped	1	1	2003	Vintage Cars		53
	S18_1367	2125551500		NYC	NY	10022	USA	NA	Frick	Michael	Small
10103	42	100.0	6	4460.82	Shipped	1	1	2003	Classic Cars		101
	S700_2824		NULLStavern		4110	Norway	EMEA	Bergulfsen	Jonas	Medium	
10104	32	53.31	2	1705.92	Shipped	1	1	2003	Trains	58	S50_1514 NULL
	Madrid		28034	Spain	EMEA	Freyre	Diego	Small			

Time taken: 0.931 seconds, Fetched: 5 row(s)

Scan hbase table:

hbase(main):009:0> scan 'sales_hbase', LIMIT => 5

ROW	COLUMN+CELL
10100	column=contact_details:city, timestamp=1667219642768, value=Nashua
10100	column=contact_details:contactfirstname, timestamp=1667219642768, value=Valarie
10100	column=contact_details:contactlastname, timestamp=1667219642768, value=Young
10100	column=contact_details:country, timestamp=1667219642768, value=USA
10100	column=contact_details:postalcode, timestamp=1667219642768, value=62005
10100	column=contact_details:state, timestamp=1667219642768, value=NH
10100	column=contact_details:territory, timestamp=1667219642768, value=NA
10100	column=product_details:msrp, timestamp=1667219642768, value=41
10100	column=product_details:productcode, timestamp=1667219642768, value=S24_3969
10100	column=product_details:productline, timestamp=1667219642768, value=Vintage Cars
10100	column=sale_details:dealsize, timestamp=1667219642768, value=Small
10100	column=sale_details:month_id, timestamp=1667219642768, value=1
10100	column=sale_details:orderlinenumber, timestamp=1667219642768, value=1
10100	column=sale_details:priceeach, timestamp=1667219642768, value=34.47
10100	column=sale_details:qtr_id, timestamp=1667219642768, value=1
10100	column=sale_details:quantityordered, timestamp=1667219642768, value=49
10100	column=sale_details:sales, timestamp=1667219642768, value=1689.03

10100	column=sale_details:status, timestamp=1667219642768, value=Shipped
10100	column=sale_details:year_id, timestamp=1667219642768, value=2003
10101	column=contact_details:city, timestamp=1667219641975, value=Frankfurt
10101	column=contact_details:contactfirstname, timestamp=1667219641975, value=Roland
10101	column=contact_details:contactlastname, timestamp=1667219641975, value=Keitel
10101	column=contact_details:country, timestamp=1667219641975, value=Germany
10101	column=contact_details:postalcode, timestamp=1667219641975, value=60528
10101	column=contact_details:state, timestamp=1667219641975, value=
10101	column=contact_details:territory, timestamp=1667219641975, value=EMEA
10101	column=product_details:msrp, timestamp=1667219641975, value=44
10101	column=product_details:productcode, timestamp=1667219641975, value=S24_2022
10101	column=product_details:productline, timestamp=1667219641975, value=Vintage Cars
10101	column=sale_details:dealsize, timestamp=1667219641975, value=Small
10101	column=sale_details:month_id, timestamp=1667219641975, value=1
10101	column=sale_details:orderlinenumber, timestamp=1667219641975, value=2
10101	column=sale_details:priceeach, timestamp=1667219641975, value=53.76
10101	column=sale_details:qtr_id, timestamp=1667219641975, value=1
10101	column=sale_details:quantityordered, timestamp=1667219641975, value=46
10101	column=sale_details:sales, timestamp=1667219641975, value=2472.96
10101	column=sale_details:status, timestamp=1667219641975, value=Shipped
10101	column=sale_details:year_id, timestamp=1667219641975, value=2003
10102	column=contact_details:city, timestamp=1667219635675, value=NYC
10102	column=contact_details:contactfirstname, timestamp=1667219635675, value=Michael
10102	column=contact_details:contactlastname, timestamp=1667219635675, value=Frick
10102	column=contact_details:country, timestamp=1667219635675, value=USA
10102	column=contact_details:phone, timestamp=1667219635675, value=2125551500
10102	column=contact_details:postalcode, timestamp=1667219635675, value=10022
10102	column=contact_details:state, timestamp=1667219635675, value=NY
10102	column=contact_details:territory, timestamp=1667219635675, value=NA
10102	column=product_details:msrp, timestamp=1667219635675, value=53
10102	column=product_details:productcode, timestamp=1667219635675, value=S18_1367
10102	column=product_details:productline, timestamp=1667219635675, value=Vintage Cars
10102	column=sale_details:dealsize, timestamp=1667219635675, value=Small
10102	column=sale_details:month_id, timestamp=1667219635675, value=1
10102	column=sale_details:orderlinenumber, timestamp=1667219635675, value=1
10102	column=sale_details:priceeach, timestamp=1667219635675, value=50.14
10102	column=sale_details:qtr_id, timestamp=1667219635675, value=1
10102	column=sale_details:quantityordered, timestamp=1667219635675, value=41

10102	column=sale_details:sales, timestamp=1667219635675, value=2055.74
10102	column=sale_details:status, timestamp=1667219635675, value=Shipped
10102	column=sale_details:year_id, timestamp=1667219635675, value=2003
10103	column=contact_details:city, timestamp=1667219642768, value=Stavern
10103	column=contact_details:contactfirstname, timestamp=1667219642768, value=Jonas
10103	column=contact_details:contactlastname, timestamp=1667219642768, value=Bergulfsen
10103	column=contact_details:country, timestamp=1667219642768, value=Norway
10103	column=contact_details:postalcode, timestamp=1667219642768, value=4110
10103	column=contact_details:state, timestamp=1667219642768, value=
10103	column=contact_details:territory, timestamp=1667219642768, value=EMEA
10103	column=product_details:msrp, timestamp=1667219642768, value=101
10103	column=product_details:productcode, timestamp=1667219642768, value=S700_2824
10103	column=product_details:productline, timestamp=1667219642768, value=Classic Cars
10103	column=sale_details:dealsize, timestamp=1667219642768, value=Medium
10103	column=sale_details:month_id, timestamp=1667219642768, value=1
10103	column=sale_details:orderlinenumber, timestamp=1667219642768, value=6
10103	column=sale_details:priceeach, timestamp=1667219642768, value=100.0
10103	column=sale_details:qtr_id, timestamp=1667219642768, value=1
10103	column=sale_details:quantityordered, timestamp=1667219642768, value=42
10103	column=sale_details:sales, timestamp=1667219642768, value=4460.82
10103	column=sale_details:status, timestamp=1667219642768, value=Shipped
10103	column=sale_details:year_id, timestamp=1667219642768, value=2003
10104	column=contact_details:city, timestamp=1667219642768, value=Madrid
10104	column=contact_details:contactfirstname, timestamp=1667219642768, value=Diego
10104	column=contact_details:contactlastname, timestamp=1667219642768, value=Freyre
10104	column=contact_details:country, timestamp=1667219642768, value=Spain
10104	column=contact_details:postalcode, timestamp=1667219642768, value=28034
10104	column=contact_details:state, timestamp=1667219642768, value=
10104	column=contact_details:territory, timestamp=1667219642768, value=EMEA
10104	column=product_details:msrp, timestamp=1667219642768, value=58
10104	column=product_details:productcode, timestamp=1667219642768, value=S50_1514
10104	column=product_details:productline, timestamp=1667219642768, value=Trains
10104	column=sale_details:dealsize, timestamp=1667219642768, value=Small
10104	column=sale_details:month_id, timestamp=1667219642768, value=1
10104	column=sale_details:orderlinenumber, timestamp=1667219642768, value=2
10104	column=sale_details:priceeach, timestamp=1667219642768, value=53.31
10104	column=sale_details:qtr_id, timestamp=1667219642768, value=1
10104	column=sale_details:quantityordered, timestamp=1667219642768, value=32

10104	column=sale_details:sales, timestamp=1667219642768, value=1705.92
10104	column=sale_details:status, timestamp=1667219642768, value=Shipped
10104	column=sale_details:year_id, timestamp=1667219642768, value=2003

5 row(s) in 0.1330 seconds