

```
hive> show databases;
OK
default
hiveproj
Time taken: 1.367 seconds, Fetched: 2 row(s)
hive> create database hiveassign;
OK
Time taken: 0.899 seconds
hive> use hiveassign;
OK
Time taken: 0.05 seconds
hive> create table sales_order
> (
>
>   ordernumber 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.
```

FAILED: ParseException line 24:20 cannot recognize input near 'fields terminated by','' in serde properties specification

```
hive> CREATE TABLE sales_orders (  
  >   ordernumber 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',';
```

OK

Time taken: 220.857 seconds

```
hive> load data local inpath'file:///home/databasess/sales_order_data.csv' into table sales_orders;
```

Loading data to table hiveassign.sales\_orders

OK

Time taken: 175.202 seconds

```
hive> CREATE TABLE sales_orders_orc(  
  >   ordernumber INT,
```

```
hive> CREATE TABLE sales_orders_orc(  
  >   ordernumber 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  
  > )  
  > stored as orc;
```

OK

Time taken: 0.7 seconds

```
hive> load data local inpath'file:///home/databases/sales_order_data.csv' into table sales_orders_orc;
```

```
hive> INSERT INTO TABLE sales_orders_orc
> SELECT
>   CAST(ordernumber AS INT),
>   CAST(quantityordered AS INT),
>   CAST(priceeach AS FLOAT),
>   CAST(orderlinenumber AS INT),
>   CAST(sales AS FLOAT),
>   status,
>   CAST(qtr_id AS INT),
>   CAST(month_id AS INT),
>   CAST(year_id AS INT),
>   productline,
>   CAST(msrp AS INT),
>   productcode,
>   phone,
>   city,
>   state,
>   postalcode,
>   country,
>   territory,
>   contactlastname,
>   contactfirstname,
>   dealsize
> FROM sales_orders;
```

WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.x.

Query ID = root\_20230802180403\_8dbc6f19-9edd-4bad-a16f-7084fcf58fa1

Total jobs = 1

Launching Job 1 out of 1

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

Job running in-process (local Hadoop)

2023-08-02 18:05:52,052 Stage-1 map = 0%, reduce = 0%

2023-08-02 18:06:04,627 Stage-1 map = 100%, reduce = 0%

Finished Job = job\_local1955442162\_0001

```
loading data to table niveassign.sales_orders_orc
MapReduce Jobs Launched:
Stage-Stage-1:  HDFS Read: 360233 HDFS Write: 398340 SUCCESS
Total MapReduce CPU Time Spent: 0 msec
```

```

JK
Time taken: 153.383 seconds
hive> select * from sales_orders limit 10;

```

[illegible]

```
Time taken: 0.288 seconds, Fetched: 10 row(s)
hive> select * from sales_orders_orc limit 10;
OK
```

[illegible]



10201	22	98.57	2	2168.54	Shipped	4	12	2003	Motorcycles	95	S10_1678	6505555787	San Francisco	CA	USA	NA	Murphy	Julie
Small																		

Time taken: 0.141 seconds, Fetched: 10 row(s)

hive> select year\_id, sum(sales) as total\_sales from sales\_orders where sum(sales) group by year\_id order by year\_id;

FAILED: SemanticException [Error 10128]: Line 1:66 Not yet supported place for UDAF 'sum'

hive> select year\_id, sum(sales) as total\_sales from sales\_orders group by year\_id order by year\_id;

WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases.

Query ID = root\_20230802182109\_9554cf17-035e-4086-bcdd-f3bcc4586bf7

Total jobs = 2

Launching Job 1 out of 2

Number of reduce tasks not specified. Estimated from input data size: 1

In order to change the average load for a reducer (in bytes):

set hive.exec.reducers.bytes.per.reducer=<number>

In order to limit the maximum number of reducers:

set hive.exec.reducers.max=<number>

In order to set a constant number of reducers:

set mapreduce.job.reduces=<number>

Job running in-process (local Hadoop)

2023-08-02 18:21:33,809 Stage-1 map = 0%, reduce = 0%

2023-08-02 18:22:06,397 Stage-1 map = 100%, reduce = 0%

2023-08-02 18:22:08,462 Stage-1 map = 100%, reduce = 100%

Ended Job = job\_local70862629\_0002

Launching Job 2 out of 2

Number of reduce tasks determined at compile time: 1

In order to change the average load for a reducer (in bytes):

set hive.exec.reducers.bytes.per.reducer=<number>

In order to limit the maximum number of reducers:

set hive.exec.reducers.max=<number>

In order to set a constant number of reducers:

set mapreduce.job.reduces=<number>

Job running in-process (local Hadoop)

2023-08-02 18:22:10,593 Stage-2 map = 100%, reduce = 100%

Ended Job = job\_local632341313\_0003

```
2023-08-02 18:22:10,593 Stage-2 map = 100%, reduce = 100%
Ended Job = job_local632341313_0003
MapReduce Jobs Launched:
Stage-Stage-1: HDFS Read: 1586856 HDFS Write: 796680 SUCCESS
Stage-Stage-2: HDFS Read: 1586856 HDFS Write: 796680 SUCCESS
Total MapReduce CPU Time Spent: 0 msec
OK
```

```
NULL      NULL
2003      3516979.547241211
2004      4724162.593383789
2005      1791486.7086791992
```

```
Time taken: 61.309 seconds, Fetched: 4 row(s)
```

```
hive> select productline, productcode, max(quantityordered) as max_product_order from sales_orders group by productline,productcode order by max_product_order desc;
```

```
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X
```

```
Query ID = root_20230802183013_a58f4115-0ac8-4f1d-9b90-cd66093e0895
```

```
Total jobs = 2
```

```
Launching Job 1 out of 2
```

```
Number of reduce tasks not specified. Estimated from input data size: 1
```

```
In order to change the average load for a reducer (in bytes):
```

```
    set hive.exec.reducers.bytes.per.reducer=<number>
```

```
In order to limit the maximum number of reducers:
```

```
    set hive.exec.reducers.max=<number>
```

```
In order to set a constant number of reducers:
```

```
    set mapreduce.job.reduces=<number>
```

```
Job running in-process (local Hadoop)
```

```
2023-08-02 18:30:27,746 Stage-1 map = 0%, reduce = 0%
```

```
2023-08-02 18:30:49,205 Stage-1 map = 100%, reduce = 100%
```

```
Ended Job = job_local471626233_0004
```

```
Launching Job 2 out of 2
```

```
Number of reduce tasks determined at compile time: 1
```

```
In order to change the average load for a reducer (in bytes):
```

```
    set hive.exec.reducers.bytes.per.reducer=<number>
```

```
In order to limit the maximum number of reducers:
```

```
    set hive.exec.reducers.max=<number>
```

```
In order to set a constant number of reducers:
```

[10000:10000](#) [🔗](#)

Logs    Inspect    **Terminal**    Files    Stats

```
MapReduce Jobs Launched:
Stage-Stage-1:  HDFS Read: 2307322 HDFS Write: 796680 SUCCESS
Stage-Stage-2:  HDFS Read: 2307322 HDFS Write: 796680 SUCCESS
Total MapReduce CPU Time Spent: 0 msec

OK
Classic Cars      S12_4675          97
Planes  S700_2466      85
Planes  S700_3167      77
Vintage Cars     S18_1749          76
Classic Cars     S24_3856          76
Classic Cars     S24_2766          76
Trucks and Buses S24_2300          70
Motorcycles     S10_1678          66
Motorcycles     S12_2823          66
Motorcycles     S10_4698          66
Vintage Cars     S18_4409          66
Classic Cars     S18_4933          66
Classic Cars     S18_3685          65
Classic Cars     S24_1628          64
Classic Cars     S10_4757          64
Planes  S24_3949          64
Vintage Cars     S32_4289          62
Vintage Cars     S24_1937          61
Classic Cars     S18_1889          61
Classic Cars     S18_1129          61
Classic Cars     S24_1046          60
Vintage Cars     S24_3151          60
Classic Cars     S18_3232          60
Motorcycles     S24_2000          59
Classic Cars     S24_2887          59
Classic Cars     S18_1589          59
Vintage Cars     S18_3856          58
Vintage Cars     S50_1341          56
Trucks and Buses S18_2319          56
Planes  S24_2841          55
```



		Logs	Inspect	Terminal	Files	Stats
nts	BETA	Classic Cars	S24_1444		55	
		Classic Cars	S18_3482		55	
EARLY ACCESS		Classic Cars	S18_2870		55	
		Trucks and Buses		S18_1097		54
er		Trucks and Buses		S12_4473		54
		Planes	S24_4278	52		
		Trains	S50_1514	51		
		Vintage Cars	S24_3816		51	
		Vintage Cars	S18_1342		51	
		Trains	S18_3259	50		
S		Ships	S72_3212	50		
		Ships	S700_3962	50		
		Ships	S700_3505	50		
		Ships	S700_2610	50		
		Ships	S700_1938	50		
		Ships	S700_1138	50		
		Ships	S24_2011	50		
		Planes	S72_1253	50		
		Planes	S700_4002	50		
		Planes	S700_2834	50		
		Trucks and Buses		S32_2509		50
		Planes	S18_2581	50		
		Motorcycles	S50_4713		50	
		Motorcycles	S32_4485		50	
		Motorcycles	S24_1578		50	
		Vintage Cars	S18_2248		50	
		Motorcycles	S10_2016		50	
		Vintage Cars	S18_2325		50	
		Classic Cars	S700_2824		50	
		Vintage Cars	S18_2795		50	
		Classic Cars	S12_1099		50	
		Vintage Cars	S18_2949		50	
		Classic Cars	S10_4962		50	
		Vintage Cars	S18_3320		50	
		Classic Cars	S18_4721		50	

	Logs	Inspect	Terminal	Files	Stats	<a href="#">OPEN IN SQL</a>
BETA	Vintage Cars	S24_2022	49			
	Classic Cars	S24_3191	49			
ACCESS	Classic Cars	S24_2840	49			
	Vintage Cars	S18_2957	49			
	Vintage Cars	S18_3136	49			
	Vintage Cars	S18_3140	49			
	Vintage Cars	S24_3969	49			
	Trucks and Buses	S50_1392	49			
	Planes	S24_1785	49			
	Classic Cars	S12_3990	49			
	Planes	S18_1662	49			
	Trucks and Buses	S12_1666	49			
	Classic Cars	S12_3148	49			
	Trucks and Buses	S32_3522	49			
	Vintage Cars	S18_1367	49			
	Classic Cars	S18_3278	49			
	Trains	S32_3207	49			
	Trucks and Buses	S32_1268	49			
	Ships	S18_3029	49			
	Classic Cars	S24_4620	49			
	Classic Cars	S24_4048	49			
	Classic Cars	S18_1984	49			
	Classic Cars	S24_2972	48			
	Classic Cars	S12_1108	48			
	Planes	S700_1691	48			
	Vintage Cars	S24_4258	48			
	Classic Cars	S12_3380	48			
	Motorcycles	S32_2206	48			
	Motorcycles	S18_3782	48			
	Vintage Cars	S24_3420	47			
	Motorcycles	S24_2360	47			
	PRODUCTLINE	PRODUCTCODE	NULL			
Time taken: 39.875 seconds, Fetched: 110 row(s)						
hive> SELECT year_id, quarter(DATE_FORMAT(CAST(CONCAT(year_id, '-', month_id, '-01') AS DATE), 'yyyy-MM-dd')) AS quarter, SUM(sales) AS total_sales FROM sales_orders GROUP BY year_id, quarter(DATE_FORMAT(CAST(CONCAT(year_id, '-', month_id, '-01') AS DATE), 'yyyy-MM-dd')) ORDER BY year_id, quarter;						

ESS

```
MOTORCYCLES    524_230W      4/  
PRODUCTLINE    PRODUCTCODE    NULL
```

```
Time taken: 39.875 seconds, Fetched: 110 row(s)
```

```
hive> SELECT year_id, quarter(DATE_FORMAT(CAST(CONCAT(year_id, '-', month_id, '-01') AS DATE), 'yyyy-MM-dd')) AS quarter, SUM(sales) AS total_sales FROM sales_orders GROUP BY ye  
quarter(DATE_FORMAT(CAST(CONCAT(year_id, '-', month_id, '-01') AS DATE), 'yyyy-MM-dd')) ORDER BY year_id, quarter;
```

```
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X rele
```

```
Query ID = root_20230802183256_bdae2f59-2620-482c-9953-d84446c85a75
```

```
Total jobs = 2
```

```
Launching Job 1 out of 2
```

```
Number of reduce tasks not specified. Estimated from input data size: 1
```

```
In order to change the average load for a reducer (in bytes):
```

```
    set hive.exec.reducers.bytes.per.reducer=<number>
```

```
In order to limit the maximum number of reducers:
```

```
    set hive.exec.reducers.max=<number>
```

```
In order to set a constant number of reducers:
```

```
    set mapreduce.job.reduces=<number>
```

```
Job running in-process (local Hadoop)
```

```
2023-08-02 18:32:59,554 Stage-1 map = 100%,  reduce = 100%
```

```
Ended Job = job_local222020777_0006
```

```
Launching Job 2 out of 2
```

```
Number of reduce tasks determined at compile time: 1
```

```
In order to change the average load for a reducer (in bytes):
```

```
    set hive.exec.reducers.bytes.per.reducer=<number>
```

```
In order to limit the maximum number of reducers:
```

```
    set hive.exec.reducers.max=<number>
```

```
In order to set a constant number of reducers:
```

```
    set mapreduce.job.reduces=<number>
```

```
Job running in-process (local Hadoop)
```

```
2023-08-02 18:33:01,038 Stage-2 map = 100%,  reduce = 100%
```

```
Ended Job = job_local1699879654_0007
```

```
MapReduce Jobs Launched:
```

```
Stage-Stage-1:  HDFS Read: 3027788 HDFS Write: 796680 SUCCESS
```

```
Stage-Stage-2:  HDFS Read: 3027788 HDFS Write: 796680 SUCCESS
```

2023-08-02 18:33:01,038 Stage-2 map = 100%, reduce = 100%

Ended Job = job\_local1699879654\_0007

MapReduce Jobs Launched:

Stage-Stage-1: HDFS Read: 3027788 HDFS Write: 796680 SUCCESS

Stage-Stage-2: HDFS Read: 3027788 HDFS Write: 796680 SUCCESS

Total MapReduce CPU Time Spent: 0 msec

OK

NULL	NULL	NULL
------	------	------

2003	1	445094.6897583008
------	---	-------------------

2003	2	562365.2218017578
------	---	-------------------

2003	3	649514.5415039062
------	---	-------------------

2003	4	1860005.094177246
------	---	-------------------

2004	1	833730.6786499023
------	---	-------------------

2004	2	766260.7305297852
------	---	-------------------

2004	3	1109396.2674560547
------	---	--------------------

2004	4	2014774.9167480469
------	---	--------------------

2005	1	1071992.3580932617
------	---	--------------------

2005	2	719494.3505859375
------	---	-------------------

Time taken: 4.086 seconds, Fetched: 11 row(s)

hive> set hive.cli.print.header=true;

hive> SELECT year\_id, quarter(DATE\_FORMAT(CAST(CONCAT(year\_id, '-', month\_id, '-01') AS DATE), 'yyyy-MM-dd')) AS quarter, SUM(sales) AS total\_sales FROM sales\_orders GROUP BY year\_id, quarter(DATE\_FORMAT(CAST(CONCAT(year\_id, '-', month\_id, '-01') AS DATE), 'yyyy-MM-dd')) ORDER BY year\_id, quarter;

WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases.

Query ID = root\_20230802183835\_e1a8aad2-17ed-41aa-b55f-28276c6cbb9b

Total jobs = 2

Launching Job 1 out of 2

Number of reduce tasks not specified. Estimated from input data size: 1

In order to change the average load for a reducer (in bytes):

set hive.exec.reducers.bytes.per.reducer=<number>

In order to limit the maximum number of reducers:

set hive.exec.reducers.max=<number>

In order to set a constant number of reducers:

set mapreduce.job.reduces=<number>



```

Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
    set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
    set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
    set mapreduce.job.reduces=<number>
Job running in-process (local Hadoop)
2023-08-02 18:38:40,865 Stage-2 map = 100%,  reduce = 100%
Ended Job = job_local1486923673_0009
MapReduce Jobs Launched:
Stage-Stage-1:  HDFS Read: 3748254 HDFS Write: 796680 SUCCESS
Stage-Stage-2:  HDFS Read: 3748254 HDFS Write: 796680 SUCCESS
Total MapReduce CPU Time Spent: 0 msec
OK
year_id quarter total_sales
NULL      NULL      NULL
2003      1        445094.6897583008
2003      2        562365.2218017578
2003      3        649514.5415039062
2003      4        1860005.094177246
2004      1        833730.6786499023
2004      2        766260.7305297852
2004      3        1109396.2674560547
2004      4        2014774.9167480469
2005      1        1071992.3580932617
2005      2        719494.3505859375
Time taken: 5.396 seconds, Fetched: 11 row(s)
hive>

```



```
hive> select city, quarter(DATE_FORMAT(CAST(CONCAT(year_id, '-', month_id, '-01') AS DATE), 'yyyy-MM-dd')) as quarter, sum(sales) as quarterly_sales from sales_order;
arter(DATE_FORMAT(CAST(CONCAT(year_id, '-', month_id, '-01') AS DATE), 'yyyy-MM-dd')) order by city, quarter;
WARNING: Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using H
Query ID = root_20230802191351_9b3dd3c2-d39f-4762-b27c-e4786ef9f639
Total jobs = 2
Launching Job 1 out of 2
Number of reduce tasks not specified. Estimated from input data size: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Job running in-process (local Hadoop)
2023-08-02 19:13:58,056 Stage-1 map = 100%,  reduce = 0%
2023-08-02 19:13:59,061 Stage-1 map = 100%,  reduce = 100%
Ended Job = job_local520744268_0001
Launching Job 2 out of 2
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Job running in-process (local Hadoop)
2023-08-02 19:14:01,040 Stage-2 map = 100%,  reduce = 100%
Ended Job = job_local1985756862_0002
MapReduce Jobs Launched:
```

ss	OK		
	Aaarhus 4	100595.5498046875	
	Allentown	2	6166.7998046875
	Allentown	3	71930.61041259766
	Allentown	4	44040.729736328125
	Barcelona	2	4219.2001953125
	Barcelona	4	74192.66003417969
	Bergamo 1	56181.320068359375	
	Bergamo 4	81774.40008544922	
	Bergen 3	16363.099975585938	
	Bergen 4	95277.17993164062	
	Boras 1	31606.72021484375	
	Boras 3	53941.68981933594	
	Boras 4	48710.92053222656	
	Boston 2	74994.240234375	
	Boston 3	15344.640014648438	
	Boston 4	63730.7802734375	
	Brickhaven	1	31474.7802734375
	Brickhaven	2	7277.35009765625
	Brickhaven	3	114974.53967285156
	Brickhaven	4	11528.52978515625
	Bridgewater	2	75778.99060058594
	Bridgewater	4	26115.800537109375
	Brisbane	1	16118.479858398438
	Brisbane	3	34100.030029296875
	Bruxelles	1	18800.089721679688
	Bruxelles	2	8411.949829101562
	Bruxelles	3	47760.479736328125
	Burbank 1	37850.07958984375	
	Burbank 4	8234.559936523438	
	Burlingame	1	13529.570190429688
	Burlingame	3	42031.83020019531
	Burlingame	4	65221.67004394531
	CITY NULL	NULL	
	Cambridge	1	21782.699951171875

					Logs	Inspect	Terminal	Files	Stats
S	BETA	RLY ACCESS	⋮		Reims	4	48895.59014892578		
					Salzburg	2	98104.24005126953		
					Salzburg	3	6693.2802734375		
					Salzburg	4	45001.10986328125		
					San Diego	1	87489.23010253906		
					San Francisco	1	72899.19995117188		
					San Francisco	4	151459.4805908203		
					San Jose	2	160010.27026367188		
					San Rafael	1	267315.2586669922		
					San Rafael	2	7261.75		
					San Rafael	3	216297.40063476562		
					San Rafael	4	163983.64880371094		
					Sevilla	4	54723.621154785156		
					Singapore	1	28395.18994140625		
					Singapore	2	92033.77014160156		
					Singapore	3	90250.07995605469		
					Singapore	4	77809.37023925781		
					South Brisbane	1	21730.029907226562		
					South Brisbane	3	10640.290161132812		
					South Brisbane	4	27098.800048828125		
					Stavern	1	54701.999755859375		
					Stavern	4	61897.19006347656		
					Strasbourg	2	80438.47985839844		
					Torino	3	94117.25988769531		
					Toulouse	1	15139.1201171875		
					Toulouse	3	17251.08056640625		
					Toulouse	4	38098.240234375		
					Tsawassen	2	31302.500244140625		
					Tsawassen	3	43332.349609375		
					Vancouver	4	75238.91955566406		
					Versailles	1	5759.419921875		
					Versailles	4	59074.90026855469		
					White Plains	4	85555.98962402344		
					Time taken: 9.324 seconds, Fetched: 183 row(s)				
					..				