Logging initialized using configuration in file:/etc/hive/conf.dist/hive-log4j.properties

WARNING: Hive CLI is deprecated and migration to Beeline is recommended.

hive> use automobiles;

OK

Time taken: 0.666 seconds

1. Total Number of Orders

hive> SELECT COUNT(DISTINCT ORDERNUMBER) AS total_orders FROM project_data;

Query ID = cloudera_20251027002121_103a2331-9251-48fe-9f28-0c785dddeed7

Total jobs = 1

Launching Job 1 out of 1

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>

Starting Job = job_1761547090470_0003, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1761547090470_0003/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1761547090470_0003

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

2025-10-27 00:21:35,443 Stage-1 map = 0%, reduce = 0%

2025-10-27 00:21:46,534 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.75 sec

2025-10-27 00:21:55,357 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 4.2 sec

MapReduce Total cumulative CPU time: 4 seconds 200 msec

Ended Job = job_1761547090470_0003

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 4.2 sec   HDFS Read: 212860 HDFS Write: 4 SUCCESS

Total MapReduce CPU Time Spent: 4 seconds 200 msec

OK

298

Time taken: 32.715 seconds, Fetched: 1 row(s)

2. Total Number of Products Sold

hive> SELECT SUM(QUANTITYORDERED) AS total_quantity FROM project_data;

Query ID = cloudera_20251027003131_b6fb658f-26f7-4480-9bdb-64b753e3694a

Total jobs = 1

Launching Job 1 out of 1

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>

Starting Job = job_1761547090470_0004, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1761547090470_0004/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1761547090470_0004

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

2025-10-27 00:31:53,820 Stage-1 map = 0%, reduce = 0%

2025-10-27 00:32:01,213 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.26 sec

2025-10-27 00:32:07,624 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 2.25 sec

MapReduce Total cumulative CPU time: 2 seconds 250 msec

Ended Job = job_1761547090470_0004

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 2.25 sec   HDFS Read: 212681 HDFS Write: 6 SUCCESS

Total MapReduce CPU Time Spent: 2 seconds 250 msec

OK

96428

Time taken: 22.985 seconds, Fetched: 1 row(s)

3. Total Revenue

hive> SELECT ROUND(SUM(SALES),2) AS total_revenue FROM project_data;

Query ID = cloudera_20251027003636_def43d91-b48a-453e-b050-4d03b755d162

Total jobs = 1

Launching Job 1 out of 1

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>

Starting Job = job_1761547090470_0005, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1761547090470_0005/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1761547090470_0005

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

2025-10-27 00:37:04,162 Stage-1 map = 0%, reduce = 0%

2025-10-27 00:37:11,750 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.16 sec

2025-10-27 00:37:20,279 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 2.35 sec

MapReduce Total cumulative CPU time: 2 seconds 350 msec

Ended Job = job_1761547090470_0005

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 2.35 sec   HDFS Read: 212980 HDFS Write: 11 SUCCESS

Total MapReduce CPU Time Spent: 2 seconds 350 msec

OK

9760221.71

Time taken: 25.805 seconds, Fetched: 1 row(s)

4. Top 5 Product Lines by Sales

hive> select productline, round(sum(sales),2) as total_sales from project_data group by productline order by total_sales desc limit 5;

Query ID = cloudera_20251027004444_a23b76b9-57e4-4dc6-bdf7-c8e43abec29d

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>

Starting Job = job_1761547090470_0006, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1761547090470_0006/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1761547090470_0006

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

2025-10-27 00:44:06,784 Stage-1 map = 0%, reduce = 0%

2025-10-27 00:44:13,057 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 0.93 sec

2025-10-27 00:44:20,394 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 2.04 sec

MapReduce Total cumulative CPU time: 2 seconds 40 msec

Ended Job = job_1761547090470_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>

Starting Job = job_1761547090470_0007, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1761547090470_0007/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1761547090470_0007

Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1

2025-10-27 00:44:27,807 Stage-2 map = 0%, reduce = 0%

2025-10-27 00:44:34,040 Stage-2 map = 100%, reduce = 0%, Cumulative CPU 0.66 sec

2025-10-27 00:44:41,469 Stage-2 map = 100%, reduce = 100%, Cumulative CPU 1.63 sec

MapReduce Total cumulative CPU time: 1 seconds 630 msec

Ended Job = job_1761547090470_0007

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 2.04 sec   HDFS Read: 212443 HDFS Write: 375 SUCCESS

Stage-Stage-2: Map: 1 Reduce: 1   Cumulative CPU: 1.63 sec   HDFS Read: 5289 HDFS Write: 116 SUCCESS

Total MapReduce CPU Time Spent: 3 seconds 670 msec

OK

Classic Cars   3842868.54

Vintage Cars   1806675.68

Trucks and Buses   1111559.19

Motorcycles   1103512.19

Planes   969323.42

Time taken: 43.52 seconds, Fetched: 5 row(s)

5. Top 5 Countries by Sales

hive> select country, round(sum(sales),2) as total_sales from project_data group by country order by total_sales desc limit 5;

Query ID = cloudera_20251027005252_c165fcf0-addb-45d1-8495-f5a96e8cff86

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>

Starting Job = job_1761547090470_0008, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1761547090470_0008/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1761547090470_0008

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

2025-10-27 00:52:26,427 Stage-1 map = 0%, reduce = 0%

2025-10-27 00:52:32,708 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 0.95 sec

2025-10-27 00:52:39,052 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 2.06 sec

MapReduce Total cumulative CPU time: 2 seconds 60 msec

Ended Job = job_1761547090470_0008

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>

Starting Job = job_1761547090470_0009, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1761547090470_0009/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1761547090470_0009

Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1

2025-10-27 00:52:47,204 Stage-2 map = 0%, reduce = 0%

2025-10-27 00:52:52,494 Stage-2 map = 100%, reduce = 0%, Cumulative CPU 0.71 sec

2025-10-27 00:52:59,908 Stage-2 map = 100%, reduce = 100%, Cumulative CPU 1.69 sec

MapReduce Total cumulative CPU time: 1 seconds 690 msec

Ended Job = job_1761547090470_0009

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 2.06 sec   HDFS Read: 212435 HDFS Write: 741 SUCCESS

Stage-Stage-2: Map: 1 Reduce: 1   Cumulative CPU: 1.69 sec   HDFS Read: 5647 HDFS Write: 82 SUCCESS

Total MapReduce CPU Time Spent: 3 seconds 750 msec

OK

USA   3355575.69

Spain   1215686.92

France   1110916.52

Australia   630623.1

UK   478880.46

Time taken: 42.645 seconds, Fetched: 5 row(s)

6. Monthly Sales Trend

hive> SELECT SUBSTR(ORDERDATE, 4, 7) AS month_year, ROUND(SUM(SALES),2) AS monthly_sales

   > FROM project_data

   > GROUP BY SUBSTR(ORDERDATE, 4, 7)

   > ORDER BY month_year;

Query ID = cloudera_20251027010000_22dc5b32-56ff-42c9-b9f0-7e0a6555f40c

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>

Starting Job = job_1761547090470_0010, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1761547090470_0010/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1761547090470_0010

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

2025-10-27 01:00:10,331 Stage-1 map = 0%, reduce = 0%

2025-10-27 01:00:16,705 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.24 sec

2025-10-27 01:00:23,103 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 2.35 sec

MapReduce Total cumulative CPU time: 2 seconds 350 msec

Ended Job = job_1761547090470_0010

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>

Starting Job = job_1761547090470_0011, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1761547090470_0011/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1761547090470_0011

Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1

2025-10-27 01:00:31,112 Stage-2 map = 0%, reduce = 0%

2025-10-27 01:00:36,401 Stage-2 map = 100%, reduce = 0%, Cumulative CPU 0.68 sec

2025-10-27 01:00:43,754 Stage-2 map = 100%, reduce = 100%, Cumulative CPU 1.63 sec

MapReduce Total cumulative CPU time: 1 seconds 630 msec

Ended Job = job_1761547090470_0011

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 2.35 sec   HDFS Read: 212866 HDFS Write: 1077 SUCCESS

Stage-Stage-2: Map: 1 Reduce: 1   Cumulative CPU: 1.63 sec   HDFS Read: 5872 HDFS Write: 531 SUCCESS

Total MapReduce CPU Time Spent: 3 seconds 980 msec

OK

01-2018	129753.6
01-2019	292688.1
01-2020	339543.42
02-2018	140836.19
02-2019	311419.53
02-2020	303982.56
03-2018	155809.32
03-2019	205733.73
03-2020	374262.76
04-2018	201609.55
04-2019	206148.12
04-2020	261633.29
05-2018	192673.11
05-2019	273438.39
05-2020	457861.06
06-2018	168082.56
06-2019	286674.22
07-2018	187731.88
07-2019	327144.09
08-2018	197809.3
08-2019	461501.27
09-2018	263973.36
09-2019	320750.91
10-2018	448452.95

10-2019	552924.25
11-2018	1029837.66
11-2019	1058699.29
12-2018	236444.58
12-2019	372802.66

ERDATE NULL

Time taken: 41.216 seconds, Fetched: 30 row(s)

7. Average Sale per Order

hive> SELECT ROUND(SUM(SALES)/COUNT(DISTINCT ORDERNUMBER),2) AS avg_sale_per_order

  > FROM project_data;

Query ID = cloudera_20251027010606_d94103f6-0f71-4fb9-b5ca-aa3489482d77

Total jobs = 1

Launching Job 1 out of 1

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>

Starting Job = job_1761547090470_0012, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1761547090470_0012/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1761547090470_0012

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

2025-10-27 01:06:30,640 Stage-1 map = 0%, reduce = 0%

2025-10-27 01:06:36,916 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 0.96 sec

2025-10-27 01:06:44,239 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 2.19 sec

MapReduce Total cumulative CPU time: 2 seconds 190 msec

Ended Job = job_1761547090470_0012

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 2.19 sec   HDFS Read: 214306 HDFS Write: 9 SUCCESS

Total MapReduce CPU Time Spent: 2 seconds 190 msec

OK

32752.42

Time taken: 22.031 seconds, Fetched: 1 row(s)

8. Deal Size Distribution

hive> SELECT DEALSIZE, COUNT(*) AS num_orders, ROUND(SUM(SALES),2) AS total_sales

   > FROM project_data

   > GROUP BY DEALSIZE

   > ORDER BY total_sales DESC;

Query ID = cloudera_20251027011111_94c8e703-c3bd-40d4-8e26-a62be7052355

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>

Starting Job = job_1761547090470_0013, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1761547090470_0013/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1761547090470_0013

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

2025-10-27 01:12:00,082 Stage-1 map = 0%, reduce = 0%

2025-10-27 01:12:05,445 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 0.94 sec

2025-10-27 01:12:13,816 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 2.06 sec

MapReduce Total cumulative CPU time: 2 seconds 60 msec

Ended Job = job_1761547090470_0013

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>

Starting Job = job_1761547090470_0014, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1761547090470_0014/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1761547090470_0014

Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1

2025-10-27 01:12:21,612 Stage-2 map = 0%, reduce = 0%

2025-10-27 01:12:24,214 Stage-2 map = 100%, reduce = 0%, Cumulative CPU 0.68 sec

2025-10-27 01:12:31,498 Stage-2 map = 100%, reduce = 100%, Cumulative CPU 1.67 sec

MapReduce Total cumulative CPU time: 1 seconds 670 msec

Ended Job = job_1761547090470_0014

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 2.06 sec   HDFS Read: 213012 HDFS Write: 225 SUCCESS

Stage-Stage-2: Map: 1 Reduce: 1   Cumulative CPU: 1.67 sec   HDFS Read: 5500 HDFS Write: 79 SUCCESS

Total MapReduce CPU Time Spent: 3 seconds 730 msec

OK

Medium   1349   5931231.47

Small   1246   2570033.84

Large   152   1258956.4

DEALSIZE   1   NULL

Time taken: 38.531 seconds, Fetched: 4 row(s)

9. Top 5 Best-Selling Products

hive> SELECT PRODUCTCODE, ROUND(SUM(SALES),2) AS total_sales

   > FROM project_data

   > GROUP BY PRODUCTCODE

   > ORDER BY total_sales DESC

   > LIMIT 5;

Query ID = cloudera_20251027011616_c265bc30-ad95-4ccd-982e-3849bb8a428f

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>

Starting Job = job_1761547090470_0015, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1761547090470_0015/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1761547090470_0015

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

2025-10-27 01:16:59,411 Stage-1 map = 0%, reduce = 0%

2025-10-27 01:17:05,715 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 0.92 sec

2025-10-27 01:17:13,031 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 2.06 sec

MapReduce Total cumulative CPU time: 2 seconds 60 msec

Ended Job = job_1761547090470_0015

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>

Starting Job = job_1761547090470_0016, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1761547090470_0016/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1761547090470_0016

Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1

2025-10-27 01:17:20,034 Stage-2 map = 0%, reduce = 0%

2025-10-27 01:17:25,327 Stage-2 map = 100%, reduce = 0%, Cumulative CPU 0.72 sec

2025-10-27 01:17:31,680 Stage-2 map = 100%, reduce = 100%, Cumulative CPU 1.68 sec

MapReduce Total cumulative CPU time: 1 seconds 680 msec

Ended Job = job_1761547090470_0016

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 2.06 sec   HDFS Read: 212443 HDFS Write: 3863 SUCCESS

Stage-Stage-2: Map: 1 Reduce: 1   Cumulative CPU: 1.68 sec   HDFS Read: 8777 HDFS Write: 95 SUCCESS

Total MapReduce CPU Time Spent: 3 seconds 740 msec

OK

S18_3232   284249.02

S10_1949   179815.23

S12_1108   168585.32

S10_4698   158202.48

S18_2238   154623.95

Time taken: 40.513 seconds, Fetched: 5 row(s)

10. Average Price per Product Line

hive> SELECT PRODUCTLINE, ROUND(AVG(PRICEEACH),2) AS avg_price

  > FROM project_data

  > GROUP BY PRODUCTLINE

> ORDER BY avg_price DESC;

Query ID = cloudera_20251027012222_e56e48cf-2b49-4ef3-a33b-79b599482f7d

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>

Starting Job = job_1761547090470_0017, Tracking URL =
http://quickstart.cloudera:8088/proxy/application_1761547090470_0017/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1761547090470_0017

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

2025-10-27 01:22:53,915 Stage-1 map = 0%, reduce = 0%

2025-10-27 01:23:00,307 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.03 sec

2025-10-27 01:23:07,758 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 2.21 sec

MapReduce Total cumulative CPU time: 2 seconds 210 msec

Ended Job = job_1761547090470_0017

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>

Starting Job = job_1761547090470_0018, Tracking URL =
http://quickstart.cloudera:8088/proxy/application_1761547090470_0018/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1761547090470_0018

Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1

2025-10-27 01:23:16,780 Stage-2 map = 0%, reduce = 0%

2025-10-27 01:23:23,196 Stage-2 map = 100%, reduce = 0%, Cumulative CPU 0.82 sec

2025-10-27 01:23:30,627 Stage-2 map = 100%, reduce = 100%, Cumulative CPU 2.08 sec

MapReduce Total cumulative CPU time: 2 seconds 80 msec

Ended Job = job_1761547090470_0018

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 2.21 sec   HDFS Read: 212702 HDFS Write: 375 SUCCESS

Stage-Stage-2: Map: 1 Reduce: 1   Cumulative CPU: 2.08 sec   HDFS Read: 5164 HDFS Write: 133 SUCCESS

Total MapReduce CPU Time Spent: 4 seconds 290 msec

OK

Classic Cars   115.2

Trucks and Buses   104.34

Motorcycles   99.77

Planes   90.52

Vintage Cars   90.01

Ships   88.17

Trains   84.11

PRODUCTLINE   NULL

Time taken: 44.137 seconds, Fetched: 8 row(s)

11. Identify Peak Selling Month

hive> SELECT SUBSTR(ORDERDATE, 4, 7) AS month_year, ROUND(SUM(SALES),2) AS total_sales

   > FROM project_data

   > GROUP BY SUBSTR(ORDERDATE, 4, 7)

   > ORDER BY total_sales DESC

   > LIMIT 1;

Query ID = cloudera_20251027013636_ebc92c5b-fbb9-468f-a02a-73f69b324912

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>

Starting Job = job_1761547090470_0019, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1761547090470_0019/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1761547090470_0019

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

2025-10-27 01:36:47,456 Stage-1 map = 0%, reduce = 0%

2025-10-27 01:36:53,956 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.39 sec

2025-10-27 01:37:03,573 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 2.77 sec

MapReduce Total cumulative CPU time: 2 seconds 770 msec

Ended Job = job_1761547090470_0019

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>

Starting Job = job_1761547090470_0020, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1761547090470_0020/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1761547090470_0020

Hadoop job information for Stage-2: number of mappers: 1; number of reducers: 1

2025-10-27 01:37:11,539 Stage-2 map = 0%, reduce = 0%

2025-10-27 01:37:17,920 Stage-2 map = 100%, reduce = 0%, Cumulative CPU 0.75 sec

2025-10-27 01:37:25,426 Stage-2 map = 100%, reduce = 100%, Cumulative CPU 1.98 sec

MapReduce Total cumulative CPU time: 1 seconds 980 msec

Ended Job = job_1761547090470_0020

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 2.77 sec   HDFS Read: 212865 HDFS Write: 1077 SUCCESS

Stage-Stage-2: Map: 1 Reduce: 1   Cumulative CPU: 1.98 sec   HDFS Read: 5989 HDFS Write: 19 SUCCESS

Total MapReduce CPU Time Spent: 4 seconds 750 msec

OK

11-2019   1058699.29

Time taken: 48.181 seconds, Fetched: 1 row(s)

12. Difference Between MSRP and Actual Price

hive> SELECT ROUND(AVG(MSRP - PRICEEACH),2) AS avg_discount

  > FROM project_data;

Query ID = cloudera_20251027014444_86c1ccda-2bfc-4d07-aca9-c5b6687fcd14

Total jobs = 1

Launching Job 1 out of 1

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>

Starting Job = job_1761547090470_0021, Tracking URL = http://quickstart.cloudera:8088/proxy/application_1761547090470_0021/

Kill Command = /usr/lib/hadoop/bin/hadoop job -kill job_1761547090470_0021

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

2025-10-27 01:45:06,694 Stage-1 map = 0%, reduce = 0%

2025-10-27 01:45:14,201 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1.56 sec

2025-10-27 01:45:21,673 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 2.88 sec

MapReduce Total cumulative CPU time: 2 seconds 880 msec

Ended Job = job_1761547090470_0021

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1   Cumulative CPU: 2.88 sec   HDFS Read: 214392 HDFS Write: 6 SUCCESS

Total MapReduce CPU Time Spent: 2 seconds 880 msec

OK

-0.41

Time taken: 26.828 seconds, Fetched: 1 row(s)

hive>