

hive> show tables;

OK

temperature_data

temperature_data_vw

Time taken: 0.171 seconds, Fetched: 2 row(s)

----Creating table Olympics with mentioned columns and their respective data types as follows

hive> create table olympics(

> athlete_name string,

> athlete_age INT,

> country string,

> year INT,

> closing_date string,

> sport String,

> gold_medals int,

> silver_medals int,

> bronze_medals int,

> total_medals int

>)

> row format delimited fields terminated by '\t';

OK

Time taken: 0.497 seconds

-----Loading data set in to the table as follows---

hive> LOAD DATA LOCAL INPATH '/home/acadgild/Downloads/olympic_data.csv' into table olympics;

Loading data to table custom.olympics

OK

Time taken: 1.272 seconds

TASK 1:

3) Write a Hive Program to find the total number of medals each country won..

```
hive> select sum(total_medals),country from olympics group by country;
```

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 = acadgild_20180524003319_2c84c16d-c2c4-4a4e-9e80-bc9dcea3ede5

Total jobs = 1

Launching Job 1 out of 1

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_1527096493792_0008, Tracking URL =

http://localhost:8088/proxy/application_1527096493792_0008/

Kill Command = /home/acadgild/hadoop-2.7.2/bin/hadoop job -kill job_1527096493792_0008

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

2018-05-24 00:33:42,960 Stage-1 map = 0%, reduce = 0%

2018-05-24 00:34:08,306 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 3.49 sec

2018-05-24 00:34:36,395 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 5.64 sec

2018-05-24 00:34:41,431 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 7.07 sec

MapReduce Total cumulative CPU time: 7 seconds 70 msec

Ended Job = job_1527096493792_0008

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 7.07 sec HDFS Read: 528401 HDFS Write:

2742 SUCCESS

Total MapReduce CPU Time Spent: 7 seconds 70 msec

OK

2 Afghanistan

8 Algeria

141 Argentina

10 Armenia

609 Australia

91 Austria

25 Azerbaijan

24	Bahamas
1	Bahrain
1	Barbados
97	Belarus
18	Belgium
1	Botswana
221	Brazil
41	Bulgaria
20	Cameroon
370	Canada
22	Chile
530	China
20	Chinese Taipei
13	Colombia
2	Costa Rica
81	Croatia
188	Cuba
1	Cyprus
81	Czech Republic
89	Denmark
5	Dominican Republic
1	Ecuador
8	Egypt
1	Eritrea
18	Estonia
29	Ethiopia
118	Finland
318	France
1	Gabon
23	Georgia
629	Germany
322	Great Britain
59	Greece
1	Grenada
1	Guatemala
3	Hong Kong
145	Hungary
15	Iceland
11	India
22	Indonesia
24	Iran
9	Ireland

4	Israel
331	Italy
80	Jamaica
282	Japan
42	Kazakhstan
39	Kenya
2	Kuwait
3	Kyrgyzstan
17	Latvia
30	Lithuania
1	Macedonia
3	Malaysia
1	Mauritius
38	Mexico
5	Moldova
10	Mongolia
14	Montenegro
11	Morocco
1	Mozambique
318	Netherlands
52	New Zealand
39	Nigeria
21	North Korea
192	Norway
1	Panama
17	Paraguay
80	Poland
9	Portugal
2	Puerto Rico
3	Qatar
123	Romania
768	Russia
6	Saudi Arabia
31	Serbia
38	Serbia and Montenegro
7	Singapore
35	Slovakia
25	Slovenia
25	South Africa
308	South Korea
205	Spain
1	Sri Lanka

```
1      Sudan
181    Sweden
93     Switzerland
1      Syria
3      Tajikistan
18     Thailand
1      Togo
19     Trinidad and Tobago
4      Tunisia
28     Turkey
1      Uganda
143    Ukraine
1      United Arab Emirates
1312   United States
1      Uruguay
19     Uzbekistan
4      Venezuela
2      Vietnam
7      Zimbabwe
Time taken: 83.781 seconds, Fetched: 110 row(s)
```

2) Write a Hive program to find the number of medals won by each country in swimming.

```
hive> select sum(total_medals),country from olympics where sport="Swimming" GROUP BY country;
```

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 = acadgild_20180524003826_fc6c0158-4a90-4574-8618-0f4136bb9848

Total jobs = 1

Launching Job 1 out of 1

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_1527096493792_0009, Tracking URL =
http://localhost:8088/proxy/application_1527096493792_0009/

Kill Command = /home/acadgild/hadoop-2.7.2/bin/hadoop job -kill job_1527096493792_0009

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

2018-05-24 00:38:52,481 Stage-1 map = 0%, reduce = 0%

2018-05-24 00:39:11,948 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 4.48 sec

2018-05-24 00:39:30,684 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 8.16 sec

MapReduce Total cumulative CPU time: 8 seconds 160 msec

Ended Job = job_1527096493792_0009

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 8.16 sec HDFS Read: 529222 HDFS Write: 881 SUCCESS

Total MapReduce CPU Time Spent: 8 seconds 160 msec

OK

1 Argentina

163 Australia

3 Austria

2 Belarus

8 Brazil

5 Canada

35 China

2 Costa Rica

1 Croatia

1 Denmark

39 France

32	Germany
11	Great Britain
9	Hungary
16	Italy
43	Japan
1	Lithuania
46	Netherlands
2	Norway
3	Poland
6	Romania
20	Russia
1	Serbia
2	Slovakia
1	Slovenia
11	South Africa
4	South Korea
3	Spain
9	Sweden
1	Trinidad and Tobago
3	Tunisia
7	Ukraine
267	United States
7	Zimbabwe

Time taken: 65.477 seconds, Fetched: 34 row(s)

2) Write a Hive program to find the number of medals that India won year wise.

hive>select sum(total_medals), year from olympics where country="India" group by year;

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 = acadgild_20180524004155_1f0f16a7-b4a6-4845-b947-6a5f76f2b162

Total jobs = 1

Launching Job 1 out of 1

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_1527096493792_0010, Tracking URL =
http://localhost:8088/proxy/application_1527096493792_0010/

Kill Command = /home/acadgild/hadoop-2.7.2/bin/hadoop job -kill job_1527096493792_0010

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

2018-05-24 00:42:21,324 Stage-1 map = 0%, reduce = 0%

2018-05-24 00:42:41,642 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 4.35 sec

2018-05-24 00:42:57,720 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 7.64 sec

MapReduce Total cumulative CPU time: 7 seconds 640 msec

Ended Job = job_1527096493792_0010

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 7.64 sec HDFS Read: 529194 HDFS Write: 163 SUCCESS

Total MapReduce CPU Time Spent: 7 seconds 640 msec

OK

1 2000

1 2004

3 2008

6 2012

Time taken: 63.655 seconds, Fetched: 4 row(s)

4) Write a Hive program to find the number of gold medals each country won.

hive> select sum(gold_medals),country from olympics group by country;

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 = acadgild_20180524004405_dc3548f5-c3e6-4a15-aa94-f03d724163fe

Total jobs = 1

Launching Job 1 out of 1

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_1527096493792_0011, Tracking URL =
http://localhost:8088/proxy/application_1527096493792_0011/

Kill Command = /home/acadgild/hadoop-2.7.2/bin/hadoop job -kill job_1527096493792_0011

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

2018-05-24 00:44:28,557 Stage-1 map = 0%, reduce = 0%

2018-05-24 00:44:44,547 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.96 sec

2018-05-24 00:45:00,632 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.04 sec

MapReduce Total cumulative CPU time: 6 seconds 40 msec

Ended Job = job_1527096493792_0011

MapReduce Jobs Launched:

Stage-Stage-1: Map: 1 Reduce: 1 Cumulative CPU: 6.04 sec HDFS Read: 528396 HDFS Write: 2703 SUCCESS

Total MapReduce CPU Time Spent: 6 seconds 40 msec

OK

0	Afghanistan
2	Algeria
49	Argentina
0	Armenia
163	Australia
36	Austria
6	Azerbaijan
11	Bahamas
0	Bahrain
0	Barbados
17	Belarus
2	Belgium
0	Botswana
46	Brazil
8	Bulgaria
20	Cameroon
168	Canada
3	Chile

234	China
2	Chinese Taipei
2	Colombia
0	Costa Rica
35	Croatia
57	Cuba
0	Cyprus
14	Czech Republic
46	Denmark
3	Dominican Republic
0	Ecuador
1	Egypt
0	Eritrea
6	Estonia
13	Ethiopia
11	Finland
108	France
0	Gabon
6	Georgia
223	Germany
124	Great Britain
12	Greece
1	Grenada
0	Guatemala
0	Hong Kong

77	Hungary
0	Iceland
1	India
5	Indonesia
10	Iran
1	Ireland
1	Israel
86	Italy
24	Jamaica
57	Japan
13	Kazakhstan
11	Kenya
0	Kuwait
0	Kyrgyzstan
3	Latvia
5	Lithuania
0	Macedonia
0	Malaysia
0	Mauritius
19	Mexico
0	Moldova
2	Mongolia
0	Montenegro
2	Morocco
1	Mozambique

101	Netherlands
18	New Zealand
6	Nigeria
6	North Korea
97	Norway
1	Panama
0	Paraguay
20	Poland
1	Portugal
0	Puerto Rico
0	Qatar
57	Romania
234	Russia
0	Saudi Arabia
1	Serbia
11	Serbia and Montenegro
0	Singapore
10	Slovakia
5	Slovenia
10	South Africa
110	South Korea
19	Spain
0	Sri Lanka
0	Sudan
57	Sweden

21	Switzerland
0	Syria
0	Tajikistan
6	Thailand
0	Togo
1	Trinidad and Tobago
2	Tunisia
9	Turkey
1	Uganda
31	Ukraine
1	United Arab Emirates
552	United States
0	Uruguay
5	Uzbekistan
1	Venezuela
0	Vietnam
2	Zimbabwe

Time taken: 57.047 seconds, Fetched: 110 row(s)