Preparing database

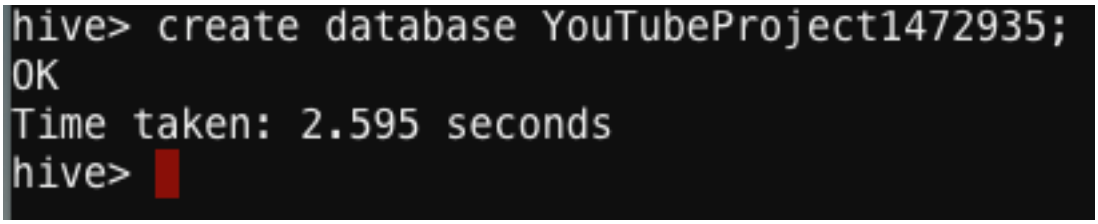
Step 1:

Create Database

Command:

create database YouTubeProject1472935;

Screenshot:



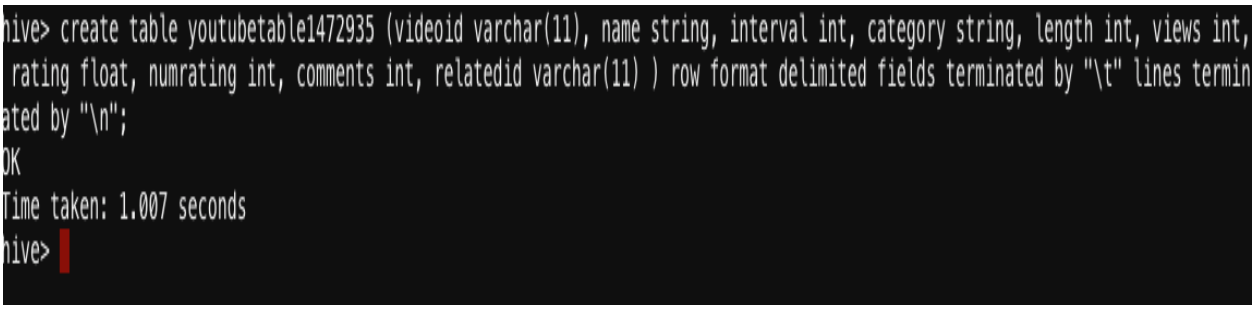
Step 2:

Create Table with the specified fields

Command:

create table youtubetable1472935 (videoid varchar(11), name string, interval int, category string, length int, views int, rating float, numrating int, comments int, relatedid varchar(11)) row format delimited fields terminated by "\t" lines terminated by "\n";

Screenshot:



Step 3:

Load data into table

Command:

load data local inpath " /mnt/home/edureka_1472935/YouTube_Analysis_Project /youtubedata.txt" into table youtubetable1472935;

Screenshot:

```
hive> load data local inpath "/mnt/home/edureka_1472935/YouTube_Analysis_Project/youtubedata.txt" into table youtubetable147293
5;
Loading data to table default.youtubetable1472935
Table default.youtubetable1472935 stats: [numFiles=1, totalSize=969389]
OK
Time taken: 0.896 seconds
hive>
```

Solutions:

A. Find out the top 5 categories with maximum number of videos uploaded.

Command:
Select category, count(*) A from youtubetable1472935 group by category order by A desc limit 5;

Screenshot:

```
hive> Select category, count(*) A from youtubetable1472935 group by category order by A desc limit 5;
Query ID = edureka_1472935_20210425184949_de2d5aa0-b589-4405-a799-d27e8d0e986e
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>
```

Result:

```
Total MapReduce CPU Time Spent: 8 seconds 80 msec
OK
Entertainment   908
Music   862
Comedy  414
People & Blogs  398
News & Politics 333
Time taken: 73.526 seconds, Fetched: 5 row(s)
hive>
```

B. Find out the top 10 rated videos.

Command:
Select videoid, rating from youtubetable1472935 order by rating desc limit 10;

Screenshot:

```
hive> Select videoid, rating from youtubetable1472935 order by rating desc limit 10;
Query ID = edureka_1472935_20210425185252_f41c6ae6-a17d-49fc-adc3-cf5826a90d3a
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>
```

Result:

```
Total MapReduce CPU Time Spent: 5 seconds 410 msec
OK
geUY_esOrt0     5.0
osqnkdc349s     5.0
wzUyV42Izz4     5.0
dh6dF1XY3uI     5.0
L3mR8syHNIg     5.0
YZev1imoxX8     5.0
gP0jnBrVEpI     5.0
Y_hAXbgxXp8     5.0
lD3GuCoJxT4     5.0
3TYqkBJ9YRk     5.0
Time taken: 30.765 seconds, Fetched: 10 row(s)
hive>
```

C. Find out the most viewed videos.

Command:

Select videoid, views from youtubetable1472935 order by views desc;

Screenshot:

```
hive> Select videoid, views from youtubetable1472935 order by views desc;
Query ID = edureka_1472935_20210425185959_6bd79cbb-1800-4dd6-98dc-2270e6f7f653
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:
```

Result:

```
Total MapReduce CPU Time Spent: 5 seconds 360 msec
OK
12Z3J1uzd0Q     65341925
4DC4Rb9quKk     33754615
LU8DDYz68kM     27721690
kHmvkRoEowc     18235463
Md6rURKhZmA     18141492
EwTZ2xpQwpA     16841569
A2f3cuUXXRs     13038204
rZBA0SKmQy8     11007201
irp8CNj9qBI     10172172
ZCYaw5tGYAs     8944331
YCj-RyKCmHQ     8892789
Pxs5EM3xlus     8361412
eBGIQ7ZuuiU     8119804
rDB9zwlXrB8     7990144
Qit3ALTelOo     7939352
O61Do03ZCjw     7553998
veYbj2zypaM     7155928
iikKzQwgBJc     7043921
pFlcqWQVVuU     6618744
GfPJeDssBOM     6101232
Z_gKOCb4QBA     6024441
-hXKRilPNpc     5994026
EBM854BTGL0     5766247
HPPj6viIBmU     5565385
Ugrlzm7fySE     5360384
leEsz9ci5XE     5322084
yH8b5ruc_-E     5239520
Oy-nqyfR_w4     5205518
```