Youtube Analysis

Creating and Loading data into Tables

Create Youtube Table and load youtubedata.txt file into table

Query:

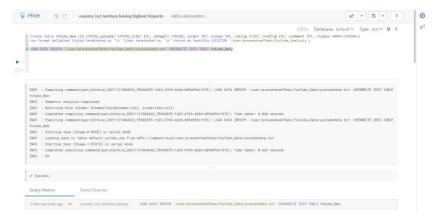
Create table youtubedata (id STRING,uploader STRING,inter INT, category STRING, length INT, nviews INT, rating FLOAT, nrating INT, ncomment INT, rvideos ARRAY<STRING>) row format delimited fields terminated by '\t' lines terminated by '\n' stored as textfile LOCATION /user/praveenkum79edu/YouTube_Analysis;

Hive ScreenSHot:



Loading data:

LOAD DATA INPATH '/user/praveenkum79edu/YouTube_Data/youtubedata.txt' OVERWRITE INTO TABLE Yotube_New;



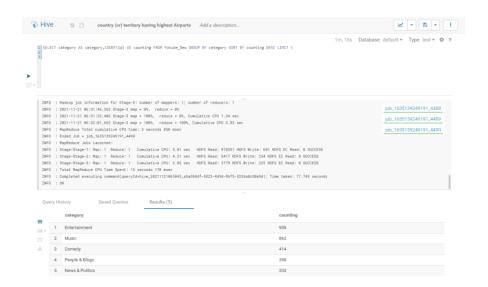
Problem Statements:

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

Solution:

SELECT category AS category, COUNT(id) AS counting FROM Yotube_New GROUP BY category SORT BY counting DESC LIMIT 5

Hive Screenshot:

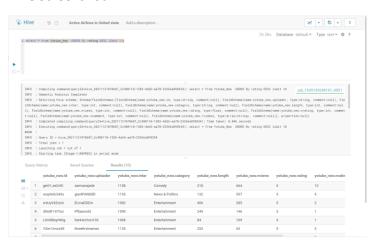


Find out the top 10 rated videos

Solution:

select * from Yotube_New ORDER By rating DESC limit 10;

Hive Screenshot:



Find out the most viewed videos.

Solution:

select id, category, rating, nviews from Yotube_New order by nviews desc limit 1;

Hive Screenshot:

