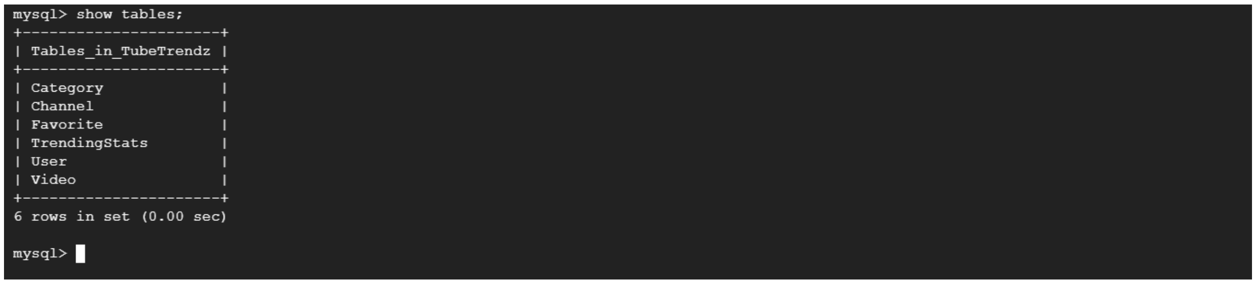
**Database Implementation:**

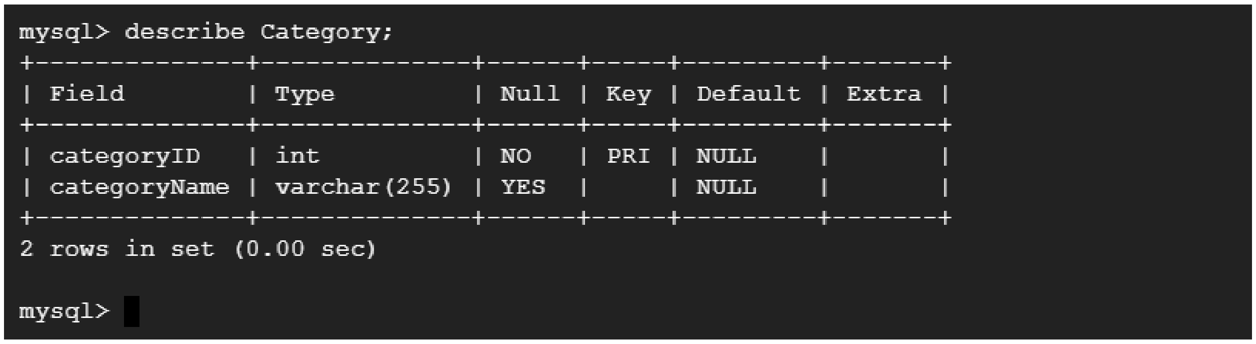
- Under the Doc folder in our git's main branch, **tubeTrendzSchema.sql** has the table creations for our TubeTrendz database schema within GCP

**TubeTrendz Database Schema:**



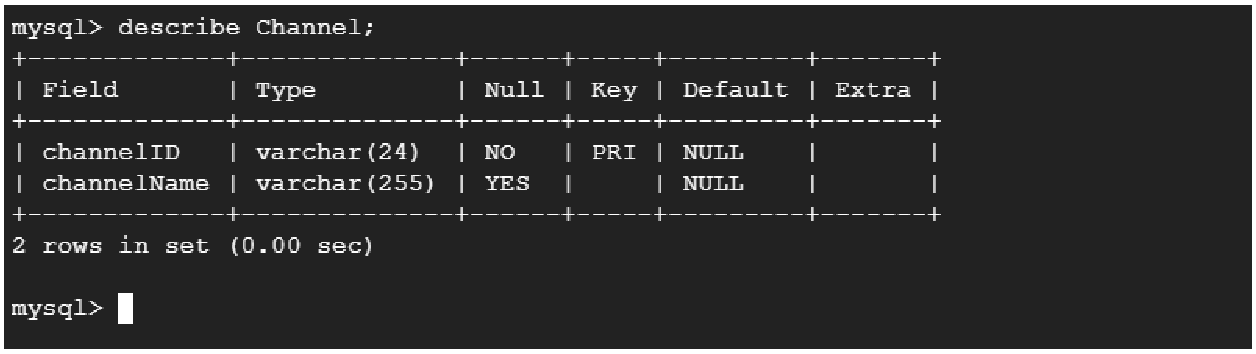
```
mysql> show tables;
+---------------------+
| Tables_in_TubeTrendz |
+---------------------+
| Category            |
| Channel             |
| Favorite            |
| TrendingStats       |
| User                |
| Video               |
+---------------------+
6 rows in set (0.00 sec)

mysql>
```

Category Table Schema:



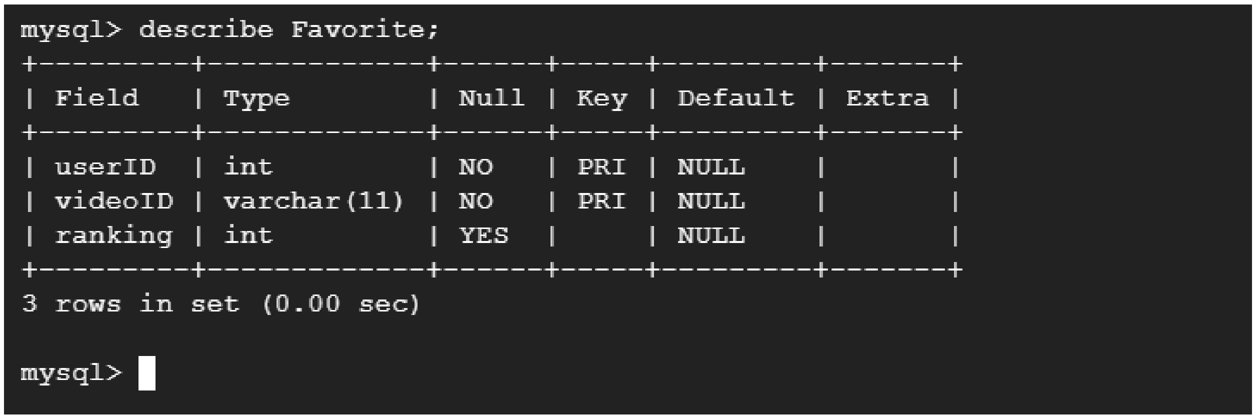
```
mysql> describe Category;
+--------------+--------------+------+-----+---------+-------+
| Field        | Type         | Null | Key | Default | Extra |
+--------------+--------------+------+-----+---------+-------+
| categoryID   | int          | NO   | PRI | NULL    |       |
| categoryName | varchar(255) | YES  |     | NULL    |       |
+--------------+--------------+------+-----+---------+-------+
2 rows in set (0.00 sec)

mysql>
```

Channel Table Schema:



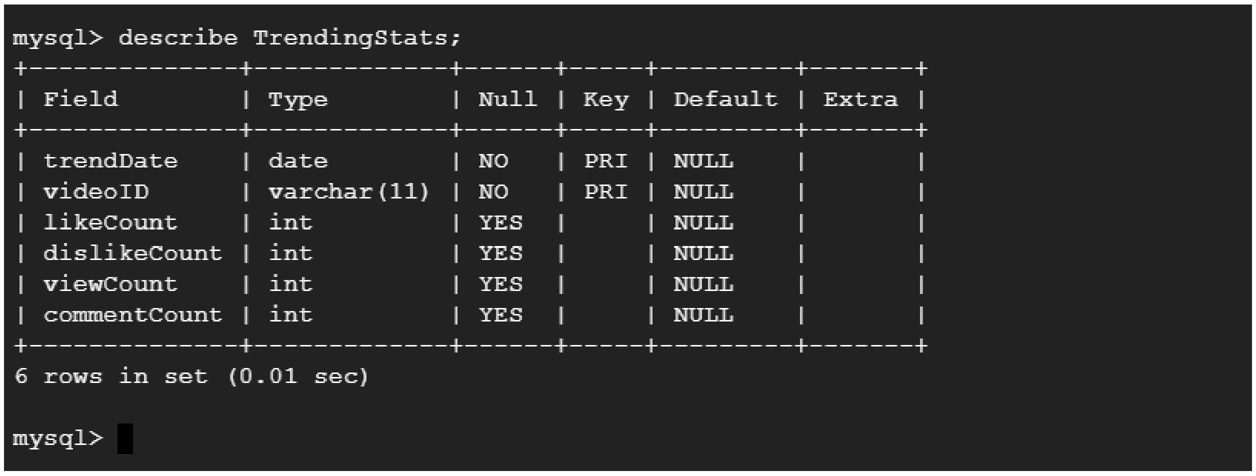
```
mysql> describe Channel;
+-------------+--------------+------+-----+---------+-------+
| Field       | Type         | Null | Key | Default | Extra |
+-------------+--------------+------+-----+---------+-------+
| channelID   | varchar(24)  | NO   | PRI | NULL    |       |
| channelName | varchar(255) | YES  |     | NULL    |       |
+-------------+--------------+------+-----+---------+-------+
2 rows in set (0.00 sec)

mysql>
```

Favorite Table Schema:



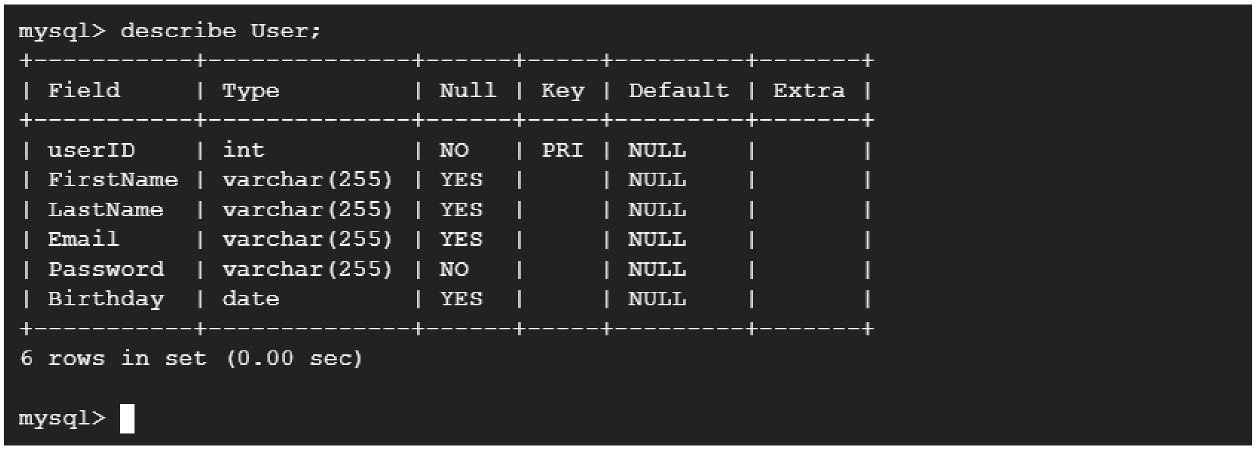
```
mysql> describe Favorite;
+---------+-------------+------+-----+---------+-------+
| Field   | Type        | Null | Key | Default | Extra |
+---------+-------------+------+-----+---------+-------+
| userID  | int         | NO   | PRI | NULL    |       |
| videoID | varchar(11) | NO   | PRI | NULL    |       |
| ranking | int         | YES  |     | NULL    |       |
+---------+-------------+------+-----+---------+-------+
3 rows in set (0.00 sec)

mysql> 
```

TrendingStats Table Schema:



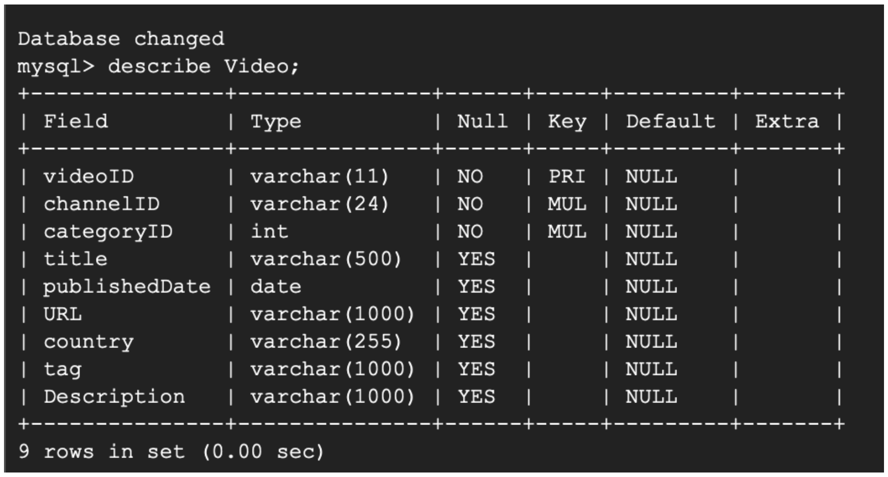
```
mysql> describe TrendingStats;
+--------------+-------------+------+-----+---------+-------+
| Field        | Type        | Null | Key | Default | Extra |
+--------------+-------------+------+-----+---------+-------+
| trendDate    | date        | NO   | PRI | NULL    |       |
| videoID      | varchar(11) | NO   | PRI | NULL    |       |
| likeCount    | int         | YES  |     | NULL    |       |
| dislikeCount | int         | YES  |     | NULL    |       |
| viewCount    | int         | YES  |     | NULL    |       |
| commentCount | int         | YES  |     | NULL    |       |
+--------------+-------------+------+-----+---------+-------+
6 rows in set (0.01 sec)

mysql> 
```

User Table Schema:



```
mysql> describe User;
+-----------+--------------+------+-----+---------+-------+
| Field     | Type         | Null | Key | Default | Extra |
+-----------+--------------+------+-----+---------+-------+
| userID    | int          | NO   | PRI | NULL    |       |
| FirstName | varchar(255) | YES  |     | NULL    |       |
| LastName  | varchar(255) | YES  |     | NULL    |       |
| Email     | varchar(255) | YES  |     | NULL    |       |
| Password  | varchar(255) | NO   |     | NULL    |       |
| Birthday  | date         | YES  |     | NULL    |       |
+-----------+--------------+------+-----+---------+-------+
6 rows in set (0.00 sec)

mysql> 
```

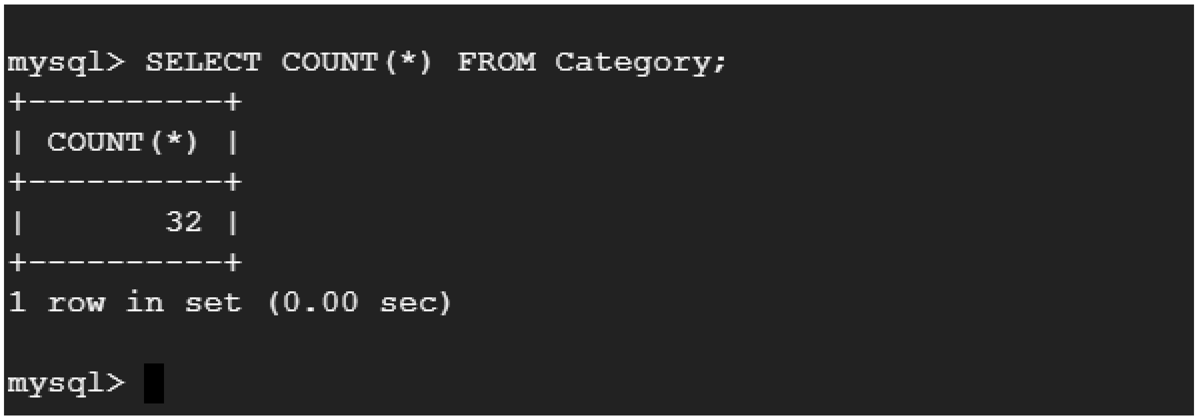
Video Table Schema:



```
Database changed
mysql> describe Video;
+---------------+---------------+------+-----+---------+-------+
| Field         | Type          | Null | Key | Default | Extra |
+---------------+---------------+------+-----+---------+-------+
| videoID       | varchar(11)   | NO   | PRI | NULL    |       |
| channelID     | varchar(24)   | NO   | MUL | NULL    |       |
| categoryID    | int           | NO   | MUL | NULL    |       |
| title         | varchar(500)  | YES  |     | NULL    |       |
| publishedDate | date          | YES  |     | NULL    |       |
| URL           | varchar(1000) | YES  |     | NULL    |       |
| country       | varchar(255)  | YES  |     | NULL    |       |
| tag           | varchar(1000) | YES  |     | NULL    |       |
| Description   | varchar(1000) | YES  |     | NULL    |       |
+---------------+---------------+------+-----+---------+-------+
9 rows in set (0.00 sec)
```

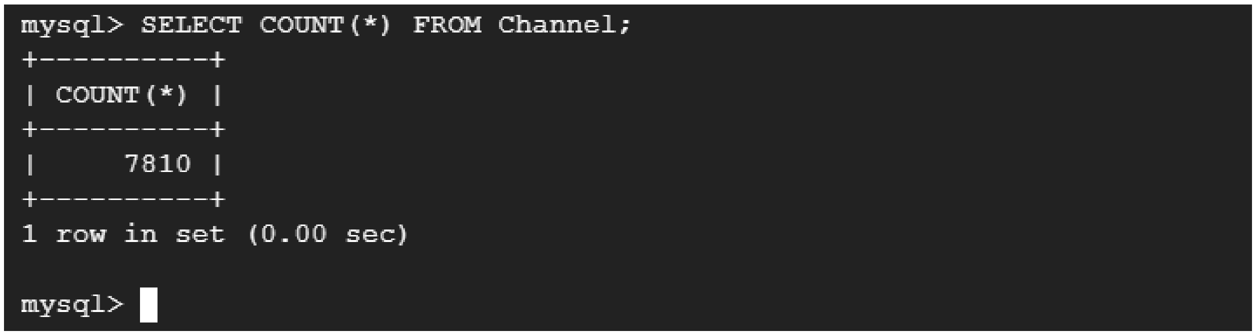
**Number of Tuples in Each Table:**

Category Table Tuple Count:



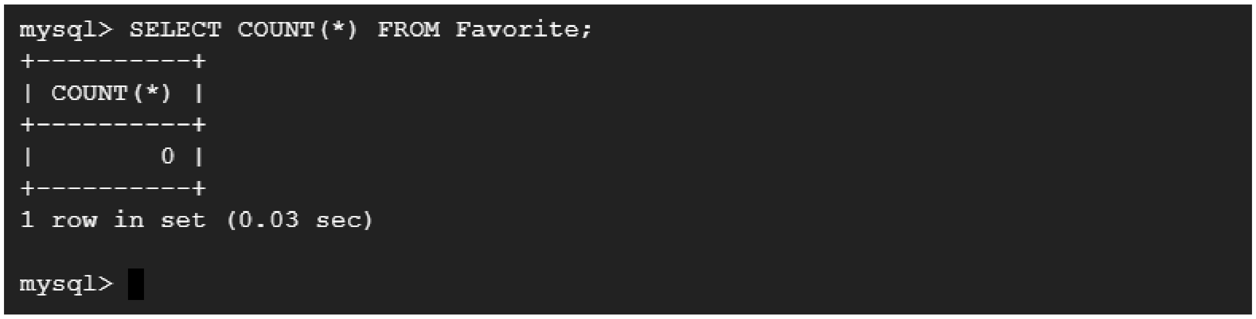
```
mysql> SELECT COUNT(*) FROM Category;
+----------+
| COUNT(*) |
+----------+
|       32 |
+----------+
1 row in set (0.00 sec)

mysql>
```

Channel Table Tuple Count:



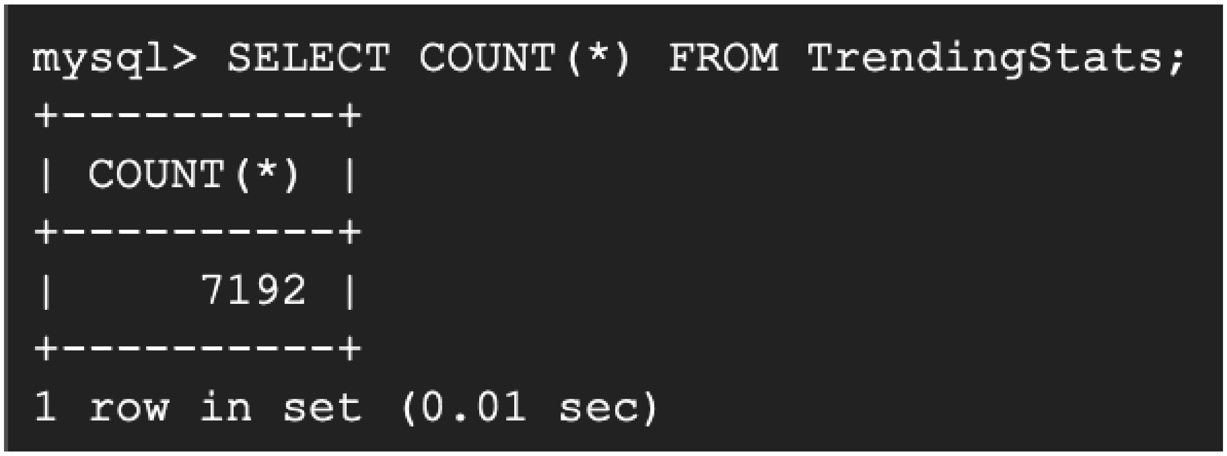
```
mysql> SELECT COUNT(*) FROM Channel;
+----------+
| COUNT(*) |
+----------+
|     7810 |
+----------+
1 row in set (0.00 sec)

mysql>
```

Favorite Table Tuple Count:



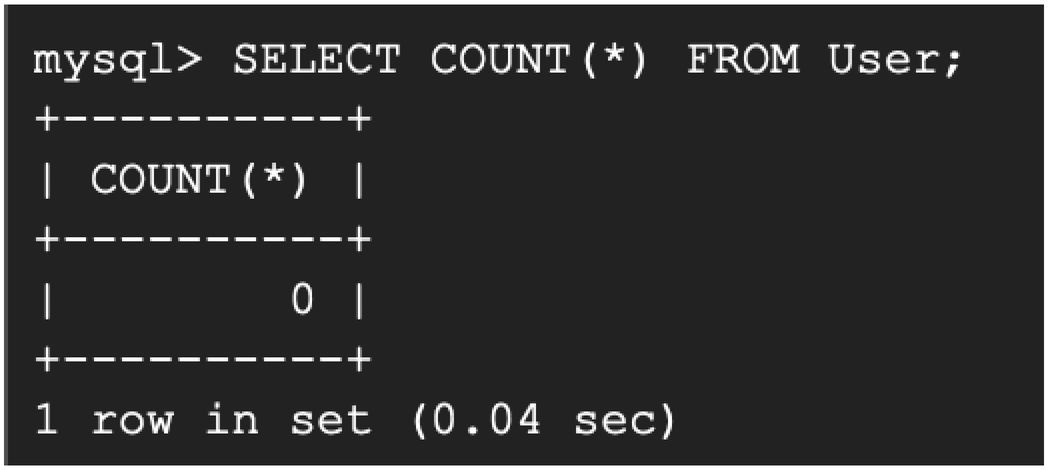
```
mysql> SELECT COUNT(*) FROM Favorite;
+----------+
| COUNT(*) |
+----------+
|        0 |
+----------+
1 row in set (0.03 sec)

mysql> 
```

TrendingStats Table Tuple Count:



```
mysql> SELECT COUNT(*) FROM TrendingStats;
+----------+
| COUNT(*) |
+----------+
|     7192 |
+----------+
1 row in set (0.01 sec)
```

User Table Tuple Count:



```
mysql> SELECT COUNT(*) FROM User;
+----------+
| COUNT(*) |
+----------+
|        0 |
+----------+
1 row in set (0.04 sec)
```

Video Table Tuple Count:

```
mysql> SELECT COUNT(*) FROM VIDEO;
ERROR 1146 (42S02): Table 'TubeTrendz.VIDEO' doesn't exist
mysql> SELECT COUNT(*) FROM Video;
+----------+
| COUNT(*) |
+----------+
|     1536 |
+----------+
1 row in set (0.22 sec)
```

Queries:

#1

Description: Our 1st SQL query returns the 15 most popular channels on Youtube based on the channel's total view count.

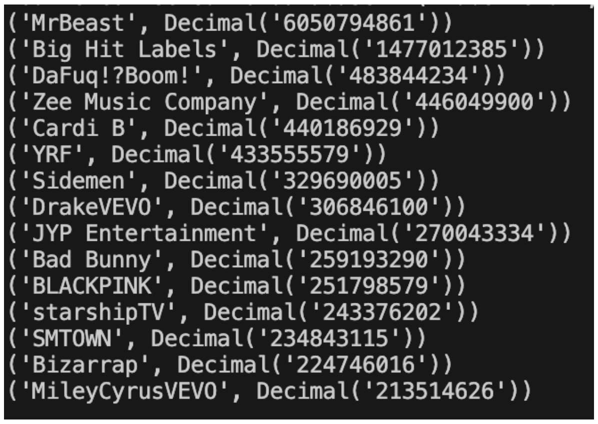
SQL Command:

SELECT Channel.channelName as channel, SUM(TrendingStats.viewCount) as views

FROM Video LEFT JOIN TrendingStats ON Video.videoID = TrendingStats.videoID

LEFT JOIN Channel on Channel.channelID = Video.channelID

GROUP BY Video.ChannelID ORDER BY views DESC LIMIT 15



```
('MrBeast', Decimal('6050794861'))
('Big Hit Labels', Decimal('1477012385'))
('DaFuq!?Boom!', Decimal('483844234'))
('Zee Music Company', Decimal('446049900'))
('Cardi B', Decimal('440186929'))
('YRF', Decimal('433555579'))
('Sidemen', Decimal('329690005'))
('DrakeVEVO', Decimal('306846100'))
('JYP Entertainment', Decimal('270043334'))
('Bad Bunny', Decimal('259193290'))
('BLACKPINK', Decimal('251798579'))
('starshipTV', Decimal('243376202'))
('SMTOWN', Decimal('234843115'))
('Bizarrap', Decimal('224746016'))
('MileyCyrusVEVO', Decimal('213514626'))
```

#2 - Only 11 rows fulfill this!

Description: Our 2nd SQL query returns the number of videos in each respective Category that has views greater than 1 million and likes greater than 100,000. As you can see, not all categories are included in our output. Consequently, these categories do not have videos that meet the demands of our video filters.

SQL Command:

SELECT c.categoryName, COUNT(*) AS NumVideos

FROM Video v NATURAL JOIN Category c

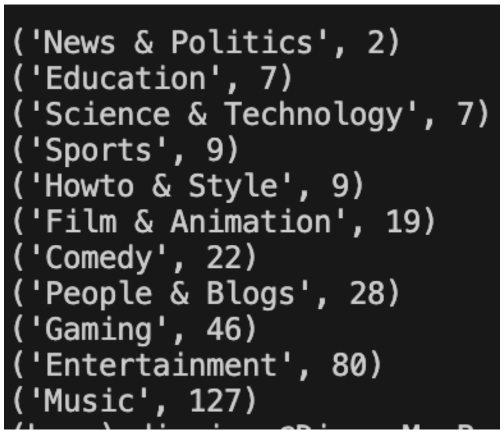
WHERE v.videoID  IN (SELECT videoID

FROM TrendingStats

WHERE viewCount > 1000000 and likeCount > 100000 )

GROUP BY c.categoryName

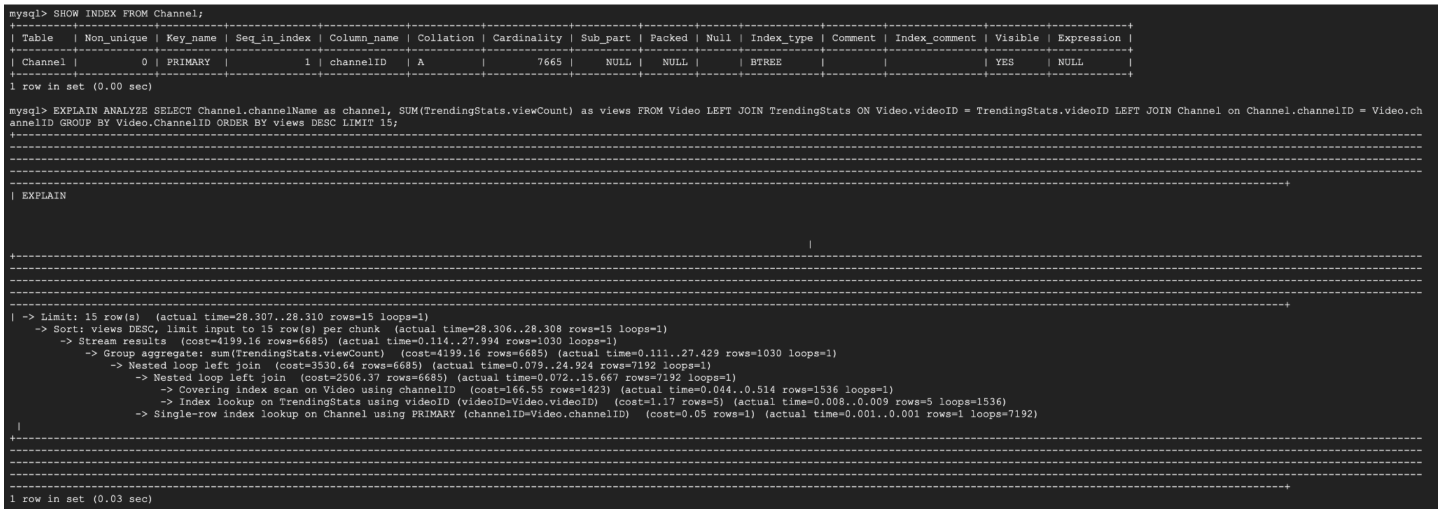
ORDER BY NumVideos ASC



```
('News & Politics', 2)
('Education', 7)
('Science & Technology', 7)
('Sports', 9)
('Howto & Style', 9)
('Film & Animation', 19)
('Comedy', 22)
('People & Blogs', 28)
('Gaming', 46)
('Entertainment', 80)
('Music', 127)
```

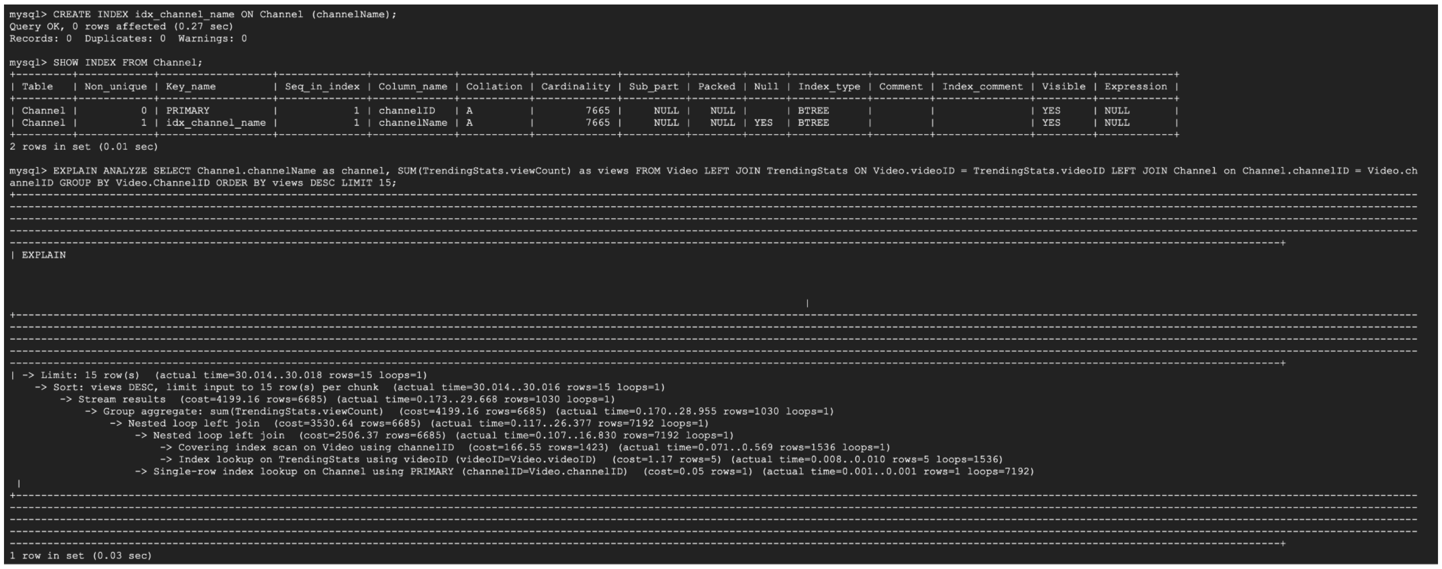
Indexing Analysis 1: Do not use an indexing scheme for this, performance is best without it since main accesses are mostly PKs.

Query #1



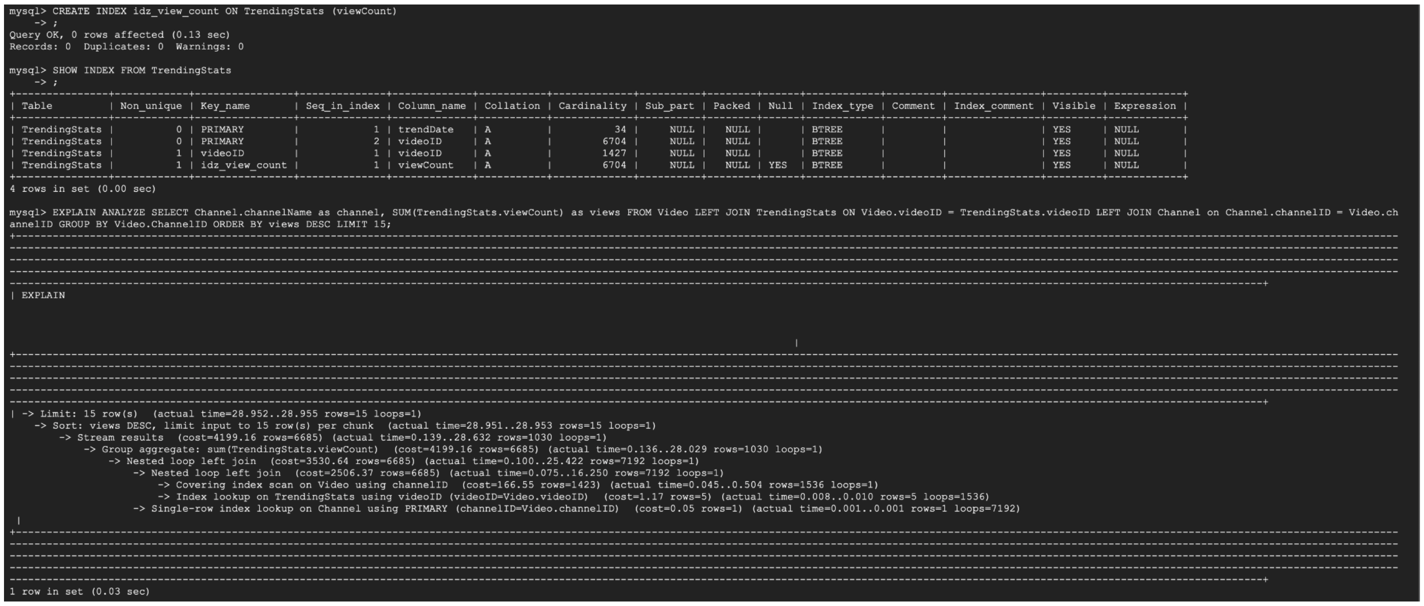
```
mysql> SHOW INDEX FROM Channel;
+---------+------------+----------+--------------+-------------+-----------+-------------+----------+--------+------+------------+---------+---------------+---------+------------+
| Table   | Non_unique | Key_name | Seq_in_index | Column_name | Collation | Cardinality | Sub_part | Packed | Null | Index_type | Comment | Index_comment | Visible | Expression |
+---------+------------+----------+--------------+-------------+-----------+-------------+----------+--------+------+------------+---------+---------------+---------+------------+
| Channel |          0 | PRIMARY  |            1 | channelID   | A         |        7665 |     NULL |   NULL |      | BTREE      |         |               | YES     | NULL       |
+---------+------------+----------+--------------+-------------+-----------+-------------+----------+--------+------+------------+---------+---------------+---------+------------+
1 row in set (0.00 sec)

mysql> EXPLAIN ANALYZE SELECT Channel.channelName as channel, SUM(TrendingStats.viewCount) as views FROM Video LEFT JOIN TrendingStats ON Video.videoID = TrendingStats.videoID LEFT JOIN Channel on Channel.channelID = Video.ch
annelID GROUP BY Video.ChannelID ORDER BY views DESC LIMIT 15;
+---------------------------------------------------------------------------------------------------------------------------------------
---------------------------------------------------------------------------------------------------------------------------------------
---------------------------------------------------------------------------------------------------------------------------------------
---------------------------------------------------------------------------------------------------------------------------------------
---------------------------------------------------------------------------------------------------------------------------------+
| EXPLAIN



                                                                                                                             |
+---------------------------------------------------------------------------------------------------------------------------------------
---------------------------------------------------------------------------------------------------------------------------------------
---------------------------------------------------------------------------------------------------------------------------------------
---------------------------------------------------------------------------------------------------------------------------------------
---------------------------------------------------------------------------------------------------------------------------------+
| -> Limit: 15 row(s)  (actual time=28.307..28.310 rows=15 loops=1)
    -> Sort: views DESC, limit input to 15 row(s) per chunk  (actual time=28.306..28.308 rows=15 loops=1)
        -> Stream results  (cost=4199.16 rows=6685) (actual time=0.114..27.994 rows=1030 loops=1)
            -> Group aggregate: sum(TrendingStats.viewCount)  (cost=4199.16 rows=6685) (actual time=0.111..27.429 rows=1030 loops=1)
                -> Nested loop left join  (cost=3530.64 rows=6685) (actual time=0.079..24.924 rows=7192 loops=1)
                    -> Nested loop left join  (cost=2506.37 rows=6685) (actual time=0.072..15.667 rows=7192 loops=1)
                        -> Covering index scan on Video using channelID  (cost=166.55 rows=1423) (actual time=0.044..0.514 rows=1536 loops=1)
                        -> Index lookup on TrendingStats using videoID (videoID=Video.videoID)  (cost=1.17 rows=5) (actual time=0.008..0.009 rows=5 loops=1536)
                    -> Single-row index lookup on Channel using PRIMARY (channelID=Video.channelID)  (cost=0.05 rows=1) (actual time=0.001..0.001 rows=1 loops=7192)
 |
+---------------------------------------------------------------------------------------------------------------------------------------
---------------------------------------------------------------------------------------------------------------------------------------
---------------------------------------------------------------------------------------------------------------------------------------
---------------------------------------------------------------------------------------------------------------------------------------
---------------------------------------------------------------------------------------------------------------------------------+
1 row in set (0.03 sec)
```

- Index on Channel.ChannelName



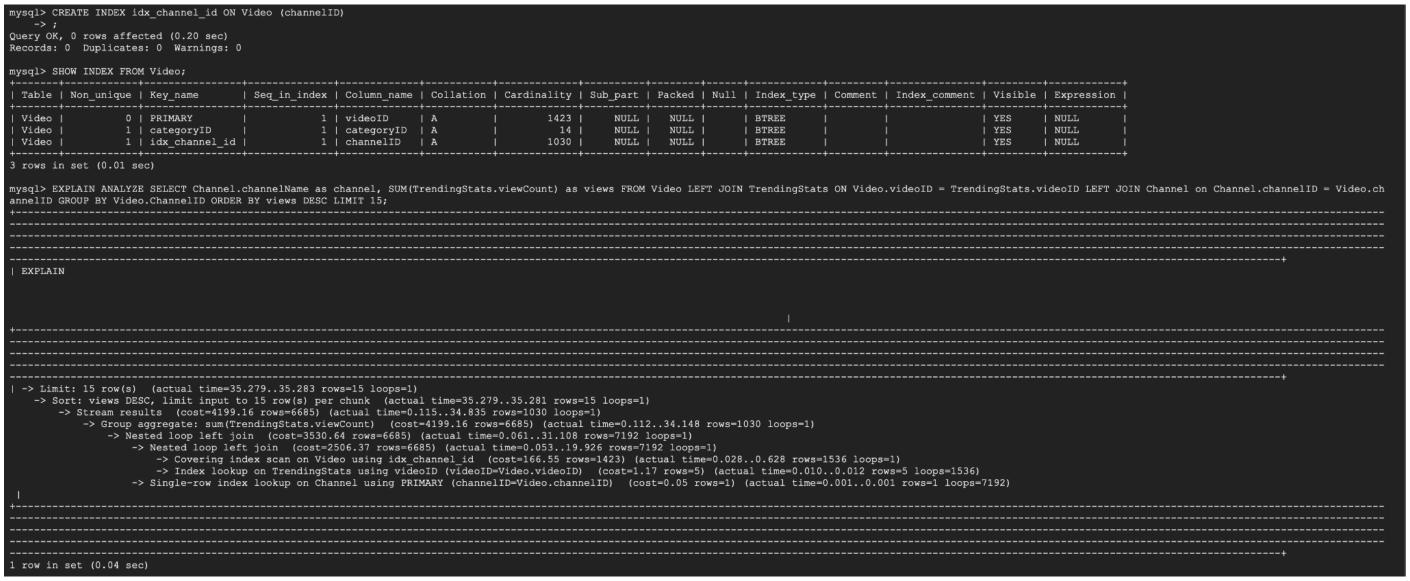
```
mysql> CREATE INDEX idx_channel_name ON Channel (channelName);
Query OK, 0 rows affected (0.27 sec)
Records: 0  Duplicates: 0  Warnings: 0

mysql> SHOW INDEX FROM Channel;
+---------+------------+------------------+--------------+-------------+-----------+-------------+----------+--------+------+------------+---------+---------------+---------+------------+
| Table   | Non_unique | Key_name         | Seq_in_index | Column_name | Collation | Cardinality | Sub_part | Packed | Null | Index_type | Comment | Index_comment | Visible | Expression |
+---------+------------+------------------+--------------+-------------+-----------+-------------+----------+--------+------+------------+---------+---------------+---------+------------+
| Channel |          0 | PRIMARY          |            1 | channelID   | A         |        7665 |     NULL |   NULL |      | BTREE      |         |               | YES     | NULL       |
| Channel |          1 | idx_channel_name |            1 | channelName | A         |        7665 |     NULL |   NULL | YES  | BTREE      |         |               | YES     | NULL       |
+---------+------------+------------------+--------------+-------------+-----------+-------------+----------+--------+------+------------+---------+---------------+---------+------------+
2 rows in set (0.01 sec)

mysql> EXPLAIN ANALYZE SELECT Channel.channelName as channel, SUM(TrendingStats.viewCount) as views FROM Video LEFT JOIN TrendingStats ON Video.videoID = TrendingStats.videoID LEFT JOIN Channel on Channel.channelID = Video.ch
annelID GROUP BY Video.ChannelID ORDER BY views DESC LIMIT 15;
+-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------+
| EXPLAIN



                                                                                                                          |
+-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------+
| -> Limit: 15 row(s)  (actual time=30.014..30.018 rows=15 loops=1)
    -> Sort: views DESC, limit input to 15 row(s) per chunk  (actual time=30.014..30.016 rows=15 loops=1)
        -> Stream results  (cost=4199.16 rows=6685) (actual time=0.173..29.668 rows=1030 loops=1)
            -> Group aggregate: sum(TrendingStats.viewCount)  (cost=4199.16 rows=6685) (actual time=0.170..28.955 rows=1030 loops=1)
                -> Nested loop left join  (cost=3530.64 rows=6685) (actual time=0.117..26.377 rows=7192 loops=1)
                    -> Nested loop left join  (cost=2506.37 rows=6685) (actual time=0.107..16.830 rows=7192 loops=1)
                        -> Covering index scan on Video using channelID  (cost=166.55 rows=1423) (actual time=0.071..0.569 rows=1536 loops=1)
                        -> Index lookup on TrendingStats using videoID (videoID=Video.videoID)  (cost=1.17 rows=5) (actual time=0.008..0.010 rows=5 loops=1536)
                    -> Single-row index lookup on Channel using PRIMARY (channelID=Video.channelID)  (cost=0.05 rows=1) (actual time=0.001..0.001 rows=1 loops=7192)
 |
+-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------+
1 row in set (0.03 sec)
```

- Index on TrendingStats.viewCount



```
mysql> CREATE INDEX idz_view_count ON TrendingStats (viewCount)
    -> ;
Query OK, 0 rows affected (0.13 sec)
Records: 0  Duplicates: 0  Warnings: 0

mysql> SHOW INDEX FROM TrendingStats
    -> ;
+---------------+------------+----------------+--------------+-------------+-----------+-------------+----------+--------+------+------------+---------+---------------+---------+------------+
| Table         | Non_unique | Key_name       | Seq_in_index | Column_name | Collation | Cardinality | Sub_part | Packed | Null | Index_type | Comment | Index_comment | Visible | Expression |
+---------------+------------+----------------+--------------+-------------+-----------+-------------+----------+--------+------+------------+---------+---------------+---------+------------+
| TrendingStats |          0 | PRIMARY        |            1 | trendDate   | A         |          34 |     NULL |   NULL |      | BTREE      |         |               | YES     | NULL       |
| TrendingStats |          0 | PRIMARY        |            2 | videoID     | A         |        6704 |     NULL |   NULL |      | BTREE      |         |               | YES     | NULL       |
| TrendingStats |          1 | videoID        |            1 | videoID     | A         |        1427 |     NULL |   NULL |      | BTREE      |         |               | YES     | NULL       |
| TrendingStats |          1 | idz_view_count |            1 | viewCount   | A         |        6704 |     NULL |   NULL | YES  | BTREE      |         |               | YES     | NULL       |
+---------------+------------+----------------+--------------+-------------+-----------+-------------+----------+--------+------+------------+---------+---------------+---------+------------+
4 rows in set (0.00 sec)

mysql> EXPLAIN ANALYZE SELECT Channel.channelName as channel, SUM(TrendingStats.viewCount) as views FROM Video LEFT JOIN TrendingStats ON Video.videoID = TrendingStats.videoID LEFT JOIN Channel on Channel.channelID = Video.ch
annelID GROUP BY Video.ChannelID ORDER BY views DESC LIMIT 15;
+-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------+
| EXPLAIN



                                                                                                                          |
+-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------+
| -> Limit: 15 row(s)  (actual time=28.952..28.955 rows=15 loops=1)
    -> Sort: views DESC, limit input to 15 row(s) per chunk  (actual time=28.951..28.953 rows=15 loops=1)
        -> Stream results  (cost=4199.16 rows=6685) (actual time=0.139..28.632 rows=1030 loops=1)
            -> Group aggregate: sum(TrendingStats.viewCount)  (cost=4199.16 rows=6685) (actual time=0.136..28.029 rows=1030 loops=1)
                -> Nested loop left join  (cost=3530.64 rows=6685) (actual time=0.100..25.422 rows=7192 loops=1)
                    -> Nested loop left join  (cost=2506.37 rows=6685) (actual time=0.075..16.250 rows=7192 loops=1)
                        -> Covering index scan on Video using channelID  (cost=166.55 rows=1423) (actual time=0.045..0.504 rows=1536 loops=1)
                        -> Index lookup on TrendingStats using videoID (videoID=Video.videoID)  (cost=1.17 rows=5) (actual time=0.008..0.010 rows=5 loops=1536)
                    -> Single-row index lookup on Channel using PRIMARY (channelID=Video.channelID)  (cost=0.05 rows=1) (actual time=0.001..0.001 rows=1 loops=7192)
 |
+-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------+
1 row in set (0.03 sec)
```

- Index on Video.ChannelID

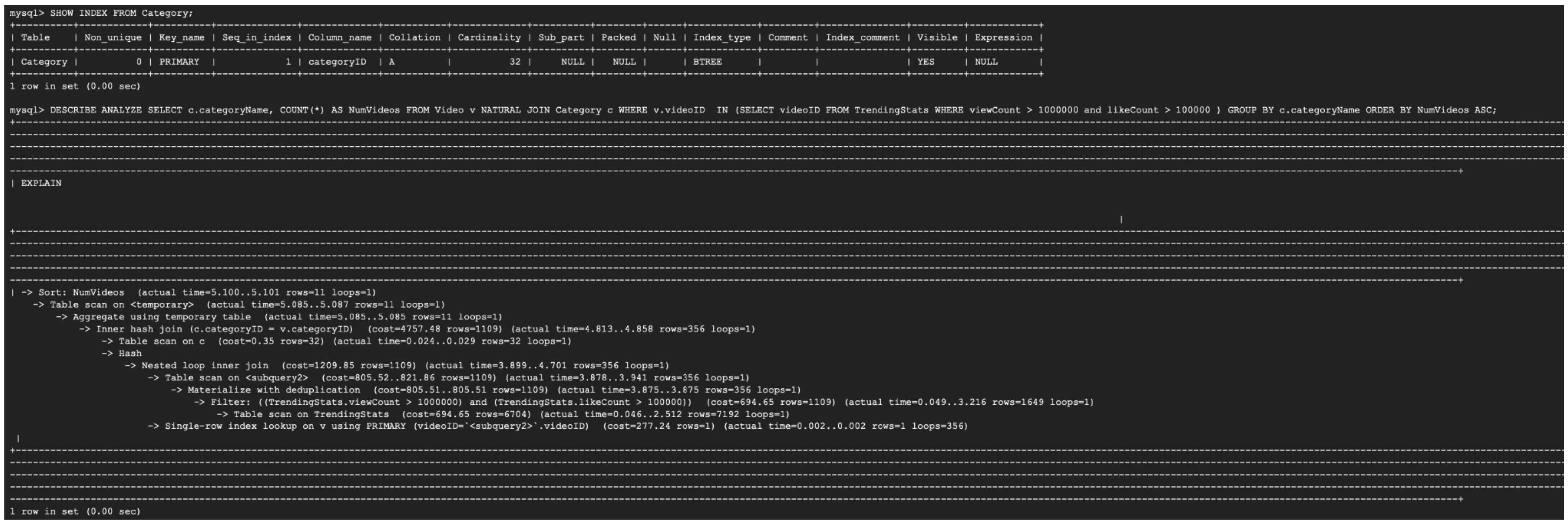


```
mysql> CREATE INDEX idx_channel_id ON Video (channelID)
    -> ;
Query OK, 0 rows affected (0.20 sec)
Records: 0  Duplicates: 0  Warnings: 0

mysql> SHOW INDEX FROM Video;
+-------+------------+----------------+--------------+-------------+-----------+-------------+----------+--------+------+------------+---------+---------------+---------+------------+
| Table | Non_unique | Key_name       | Seq_in_index | Column_name | Collation | Cardinality | Sub_part | Packed | Null | Index_type | Comment | Index_comment | Visible | Expression |
+-------+------------+----------------+--------------+-------------+-----------+-------------+----------+--------+------+------------+---------+---------------+---------+------------+
| Video |          0 | PRIMARY        |            1 | videoID     | A         |        1423 |     NULL |   NULL |      | BTREE      |         |               | YES     | NULL       |
| Video |          1 | categoryID     |            1 | categoryID  | A         |          14 |     NULL |   NULL |      | BTREE      |         |               | YES     | NULL       |
| Video |          1 | idx_channel_id |            1 | channelID   | A         |        1030 |     NULL |   NULL |      | BTREE      |         |               | YES     | NULL       |
+-------+------------+----------------+--------------+-------------+-----------+-------------+----------+--------+------+------------+---------+---------------+---------+------------+
3 rows in set (0.01 sec)

mysql> EXPLAIN ANALYZE SELECT Channel.channelName as channel, SUM(TrendingStats.viewCount) as views FROM Video LEFT JOIN TrendingStats ON Video.videoID = TrendingStats.videoID LEFT JOIN Channel on Channel.channelID = Video.ch
annelID GROUP BY Video.ChannelID ORDER BY views DESC LIMIT 15;
+----------------------------------------------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------------------------------------------------------------------+
| EXPLAIN



                                                                                                                |
+----------------------------------------------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------------------------------------------------------------------+
| -> Limit: 15 row(s)  (actual time=35.279..35.283 rows=15 loops=1)
    -> Sort: views DESC, limit input to 15 row(s) per chunk  (actual time=35.279..35.281 rows=15 loops=1)
        -> Stream results  (cost=4199.16 rows=6685) (actual time=0.115..34.835 rows=1030 loops=1)
            -> Group aggregate: sum(TrendingStats.viewCount)  (cost=4199.16 rows=6685) (actual time=0.112..34.148 rows=1030 loops=1)
                -> Nested loop left join  (cost=3530.64 rows=6685) (actual time=0.061..31.108 rows=7192 loops=1)
                    -> Nested loop left join  (cost=2506.37 rows=6685) (actual time=0.053..19.926 rows=7192 loops=1)
                        -> Covering index scan on Video using idx_channel_id  (cost=166.55 rows=1423) (actual time=0.028..0.628 rows=1536 loops=1)
                        -> Index lookup on TrendingStats using videoID (videoID=Video.videoID)  (cost=1.17 rows=5) (actual time=0.010..0.012 rows=5 loops=1536)
                    -> Single-row index lookup on Channel using PRIMARY (channelID=Video.channelID)  (cost=0.05 rows=1) (actual time=0.001..0.001 rows=1 loops=7192)
 |
+----------------------------------------------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------------------------------------------------------------------+
1 row in set (0.04 sec)
```

Indexing Analysis 2: We choose the likeCount index because the table scan cost is much smaller than the other options. This is, however, a tradeoff since the row by row time is shorter when we do not use an index.

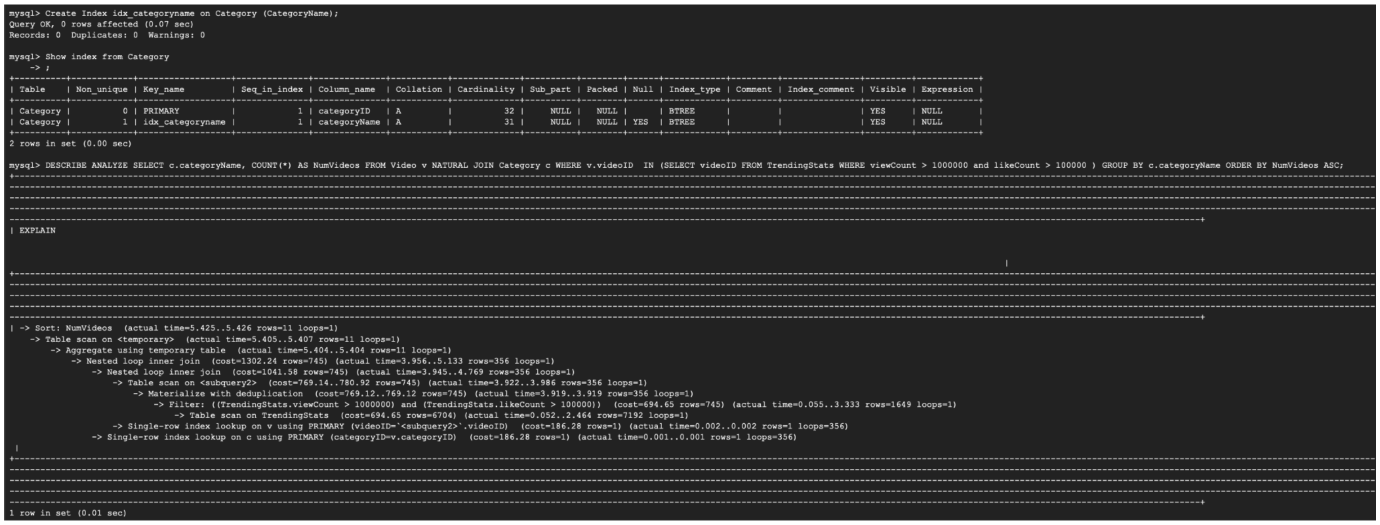
<u>Query #2</u>

SHOW INDEX FROM Category;



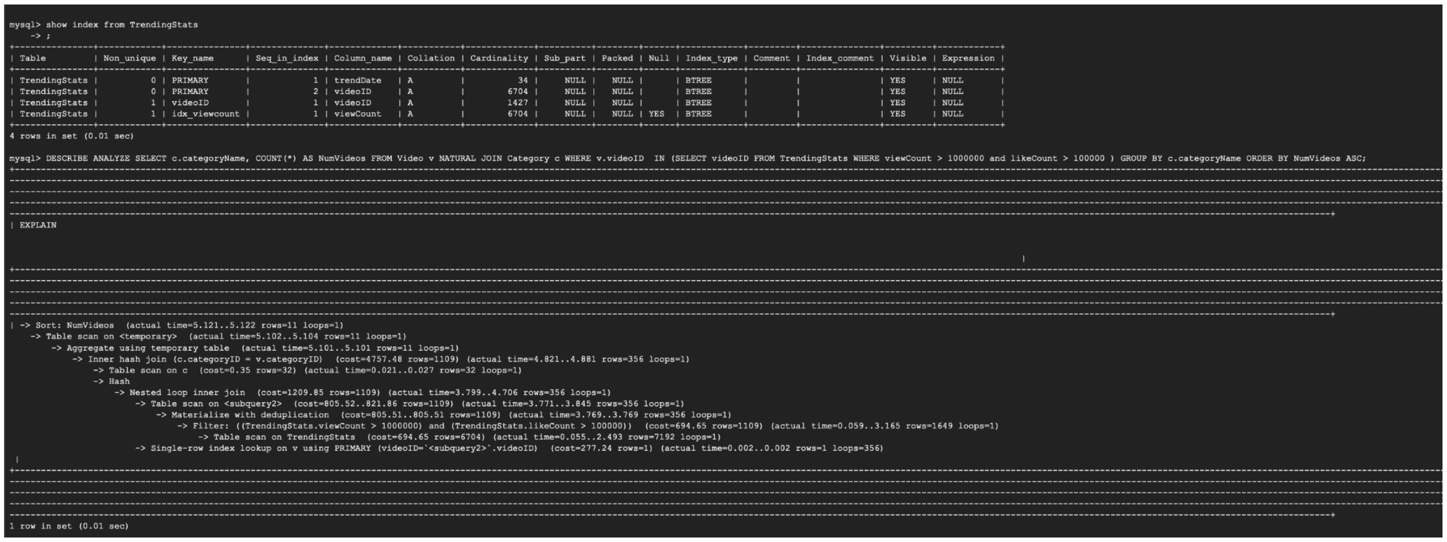
```
mysql> SHOW INDEX FROM Category;
+----------+------------+----------+--------------+-------------+-----------+-------------+----------+--------+------+------------+---------+---------------+---------+------------+
| Table    | Non_unique | Key_name | Seq_in_index | Column_name | Collation | Cardinality | Sub_part | Packed | Null | Index_type | Comment | Index_comment | Visible | Expression |
+----------+------------+----------+--------------+-------------+-----------+-------------+----------+--------+------+------------+---------+---------------+---------+------------+
| Category |          0 | PRIMARY  |            1 | categoryID  | A         |          32 |     NULL |   NULL |      | BTREE      |         |               | YES     | NULL       |
+----------+------------+----------+--------------+-------------+-----------+-------------+----------+--------+------+------------+---------+---------------+---------+------------+
1 row in set (0.00 sec)

mysql> DESCRIBE ANALYZE SELECT c.categoryName, COUNT(*) AS NumVideos FROM Video v NATURAL JOIN Category c WHERE v.videoID  IN (SELECT videoID FROM TrendingStats WHERE viewCount > 1000000 and likeCount > 100000 ) GROUP BY c.categoryName ORDER BY NumVideos ASC;
+----------------------------------------------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------------------------------------------------------------------+
| EXPLAIN



                                                                                                                |
+----------------------------------------------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------------------------------------------------------------------+
| -> Sort: NumVideos  (actual time=5.100..5.101 rows=11 loops=1)
    -> Table scan on <temporary>  (actual time=5.085..5.087 rows=11 loops=1)
        -> Aggregate using temporary table  (actual time=5.085..5.085 rows=11 loops=1)
            -> Inner hash join (c.categoryID = v.categoryID)  (cost=4757.48 rows=1109) (actual time=4.813..4.858 rows=356 loops=1)
                -> Table scan on c  (cost=0.35 rows=32) (actual time=0.024..0.029 rows=32 loops=1)
                -> Hash
                    -> Nested loop inner join  (cost=1209.85 rows=1109) (actual time=3.899..4.701 rows=356 loops=1)
                        -> Table scan on <subquery2>  (cost=805.52..821.86 rows=1109) (actual time=3.878..3.941 rows=356 loops=1)
                            -> Materialize with deduplication  (cost=805.51..805.51 rows=1109) (actual time=3.875..3.875 rows=356 loops=1)
                                -> Filter: ((TrendingStats.viewCount > 1000000) and (TrendingStats.likeCount > 100000))  (cost=694.65 rows=1109) (actual time=0.049..3.216 rows=1649 loops=1)
                                    -> Table scan on TrendingStats  (cost=694.65 rows=6704) (actual time=0.046..2.512 rows=7192 loops=1)
                        -> Single-row index lookup on v using PRIMARY (videoID=`<subquery2>`.videoID)  (cost=277.24 rows=1) (actual time=0.002..0.002 rows=1 loops=356)
 |
+----------------------------------------------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------
----------------------------------------------------------------------------------------------------------------------------------------------------+
1 row in set (0.00 sec)
```

## Create Index idx_categoryname on Category (CategoryName)



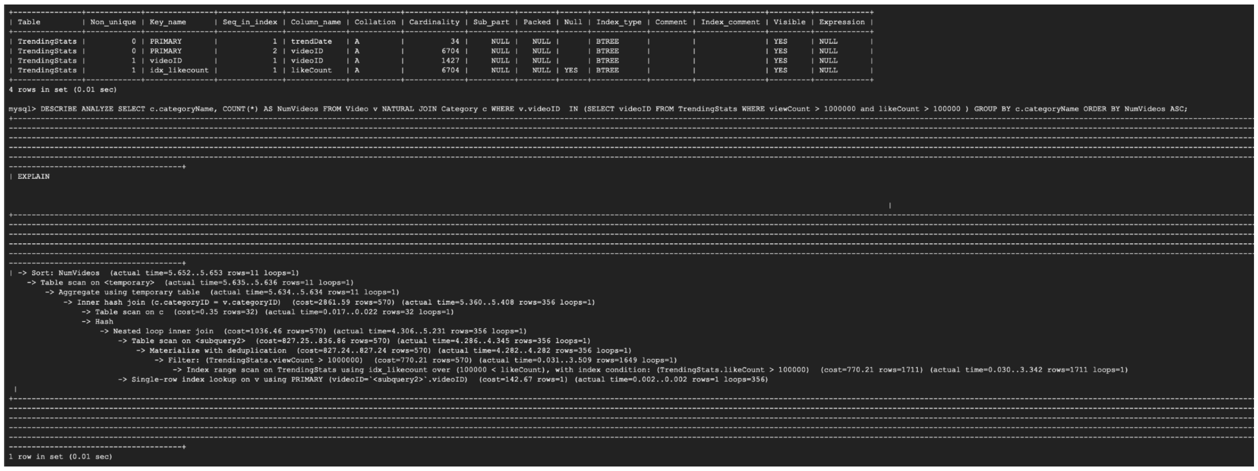
```
mysql> Create Index idx_categoryname on Category (CategoryName);
Query OK, 0 rows affected (0.07 sec)
Records: 0  Duplicates: 0  Warnings: 0

mysql> Show index from Category
    -> ;
+----------+------------+------------------+--------------+--------------+-----------+-------------+----------+--------+------+------------+---------+---------------+---------+------------+
| Table    | Non_unique | Key_name         | Seq_in_index | Column_name  | Collation | Cardinality | Sub_part | Packed | Null | Index_type | Comment | Index_comment | Visible | Expression |
+----------+------------+------------------+--------------+--------------+-----------+-------------+----------+--------+------+------------+---------+---------------+---------+------------+
| Category |          0 | PRIMARY          |            1 | categoryID   | A         |          32 |     NULL |   NULL |      | BTREE      |         |               | YES     | NULL       |
| Category |          1 | idx_categoryname |            1 | categoryName | A         |          31 |     NULL |   NULL | YES  | BTREE      |         |               | YES     | NULL       |
+----------+------------+------------------+--------------+--------------+-----------+-------------+----------+--------+------+------------+---------+---------------+---------+------------+
2 rows in set (0.00 sec)

mysql> DESCRIBE ANALYZE SELECT c.categoryName, COUNT(*) AS NumVideos FROM Video v NATURAL JOIN Category c WHERE v.videoID  IN (SELECT videoID FROM TrendingStats WHERE viewCount > 1000000 and likeCount > 100000 ) GROUP BY c.categoryName ORDER BY NumVideos ASC;
+--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------+
| EXPLAIN


                                                                                                                                                                                       |
+--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------+
| -> Sort: NumVideos  (actual time=5.425..5.426 rows=11 loops=1)
    -> Table scan on <temporary>  (actual time=5.405..5.407 rows=11 loops=1)
        -> Aggregate using temporary table  (actual time=5.404..5.404 rows=11 loops=1)
            -> Nested loop inner join  (cost=1302.24 rows=745) (actual time=3.956..5.133 rows=356 loops=1)
                -> Nested loop inner join  (cost=1041.58 rows=745) (actual time=3.945..4.769 rows=356 loops=1)
                    -> Table scan on <subquery2>  (cost=769.14..780.92 rows=745) (actual time=3.922..3.986 rows=356 loops=1)
                        -> Materialize with deduplication  (cost=769.12..769.12 rows=745) (actual time=3.919..3.919 rows=356 loops=1)
                            -> Filter: ((TrendingStats.viewCount > 1000000) and (TrendingStats.likeCount > 100000))  (cost=694.65 rows=745) (actual time=0.055..3.333 rows=1649 loops=1)
                                -> Table scan on TrendingStats  (cost=694.65 rows=6704) (actual time=0.052..2.464 rows=7192 loops=1)
                    -> Single-row index lookup on v using PRIMARY (videoID=`<subquery2>`.videoID)  (cost=186.28 rows=1) (actual time=0.002..0.002 rows=1 loops=356)
                -> Single-row index lookup on c using PRIMARY (categoryID=v.categoryID)  (cost=186.28 rows=1) (actual time=0.001..0.001 rows=1 loops=356)
 |
+--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------+
1 row in set (0.01 sec)
```

## Creating Index on ViewCount:



```
mysql> show index from TrendingStats
    -> ;
+---------------+------------+---------------+--------------+-------------+-----------+-------------+----------+--------+------+------------+---------+---------------+---------+------------+
| Table         | Non_unique | Key_name      | Seq_in_index | Column_name | Collation | Cardinality | Sub_part | Packed | Null | Index_type | Comment | Index_comment | Visible | Expression |
+---------------+------------+---------------+--------------+-------------+-----------+-------------+----------+--------+------+------------+---------+---------------+---------+------------+
| TrendingStats |          0 | PRIMARY       |            1 | trendDate   | A         |          34 |     NULL |   NULL |      | BTREE      |         |               | YES     | NULL       |
| TrendingStats |          0 | PRIMARY       |            2 | videoID     | A         |        6704 |     NULL |   NULL |      | BTREE      |         |               | YES     | NULL       |
| TrendingStats |          1 | videoID       |            1 | videoID     | A         |        1427 |     NULL |   NULL |      | BTREE      |         |               | YES     | NULL       |
| TrendingStats |          1 | idx_viewcount |            1 | viewCount   | A         |        6704 |     NULL |   NULL | YES  | BTREE      |         |               | YES     | NULL       |
+---------------+------------+---------------+--------------+-------------+-----------+-------------+----------+--------+------+------------+---------+---------------+---------+------------+
4 rows in set (0.01 sec)

mysql> DESCRIBE ANALYZE SELECT c.categoryName, COUNT(*) AS NumVideos FROM Video v NATURAL JOIN Category c WHERE v.videoID  IN (SELECT videoID FROM TrendingStats WHERE viewCount > 1000000 and likeCount > 100000 ) GROUP BY c.categoryName ORDER BY NumVideos ASC;
+--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------+
| EXPLAIN


                                                                                                                                                                                       |
+--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------+
| -> Sort: NumVideos  (actual time=5.121..5.122 rows=11 loops=1)
    -> Table scan on <temporary>  (actual time=5.102..5.104 rows=11 loops=1)
        -> Aggregate using temporary table  (actual time=5.101..5.101 rows=11 loops=1)
            -> Inner hash join (c.categoryID = v.categoryID)  (cost=4757.48 rows=1109) (actual time=4.821..4.881 rows=356 loops=1)
                -> Table scan on c  (cost=0.35 rows=32) (actual time=0.021..0.027 rows=32 loops=1)
                -> Hash
                    -> Nested loop inner join  (cost=1209.85 rows=1109) (actual time=3.799..4.706 rows=356 loops=1)
                        -> Table scan on <subquery2>  (cost=805.52..821.86 rows=1109) (actual time=3.771..3.845 rows=356 loops=1)
                            -> Materialize with deduplication  (cost=805.51..805.51 rows=1109) (actual time=3.769..3.769 rows=356 loops=1)
                                -> Filter: ((TrendingStats.viewCount > 1000000) and (TrendingStats.likeCount > 100000))  (cost=694.65 rows=1109) (actual time=0.059..3.165 rows=1649 loops=1)
                                    -> Table scan on TrendingStats  (cost=694.65 rows=6704) (actual time=0.055..2.493 rows=7192 loops=1)
                        -> Single-row index lookup on v using PRIMARY (videoID=`<subquery2>`.videoID)  (cost=277.24 rows=1) (actual time=0.002..0.002 rows=1 loops=356)
 |
+--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------+
1 row in set (0.01 sec)
```

## Create Index on LikeCount:



```
+---------------+------------+--------------+--------------+-------------+-----------+-------------+----------+--------+------+------------+---------+---------------+---------+------------+
| Table         | Non_unique | Key_name     | Seq_in_index | Column_name | Collation | Cardinality | Sub_part | Packed | Null | Index_type | Comment | Index_comment | Visible | Expression |
+---------------+------------+--------------+--------------+-------------+-----------+-------------+----------+--------+------+------------+---------+---------------+---------+------------+
| TrendingStats |          0 | PRIMARY      |            1 | trendDate   | A         |          34 |     NULL |   NULL |      | BTREE      |         |               | YES     | NULL       |
| TrendingStats |          0 | PRIMARY      |            2 | videoID     | A         |        6704 |     NULL |   NULL |      | BTREE      |         |               | YES     | NULL       |
| TrendingStats |          1 | videoID      |            1 | videoID     | A         |        1427 |     NULL |   NULL |      | BTREE      |         |               | YES     | NULL       |
| TrendingStats |          1 | idx_likecount |           1 | likeCount   | A         |        6704 |     NULL |   NULL | YES  | BTREE      |         |               | YES     | NULL       |
+---------------+------------+--------------+--------------+-------------+-----------+-------------+----------+--------+------+------------+---------+---------------+---------+------------+
4 rows in set (0.01 sec)

mysql> DESCRIBE ANALYZE SELECT c.categoryName, COUNT(*) AS NumVideos FROM Video v NATURAL JOIN Category c WHERE v.videoID  IN (SELECT videoID FROM TrendingStats WHERE viewCount > 1000000 and likeCount > 100000 ) GROUP BY c.categoryName ORDER BY NumVideos ASC;
+------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
-----------------------------------+
| EXPLAIN

                                                                                                                                                                                        |
+------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
-----------------------------------+
| -> Sort: NumVideos  (actual time=5.652..5.653 rows=11 loops=1)
    -> Table scan on <temporary>  (actual time=5.635..5.636 rows=11 loops=1)
        -> Aggregate using temporary table  (actual time=5.634..5.634 rows=11 loops=1)
            -> Inner hash join (c.categoryID = v.categoryID)  (cost=2861.59 rows=570) (actual time=5.360..5.408 rows=356 loops=1)
                -> Table scan on c  (cost=0.35 rows=32) (actual time=0.017..0.022 rows=32 loops=1)
                -> Hash
                    -> Nested loop inner join  (cost=1036.46 rows=570) (actual time=4.306..5.231 rows=356 loops=1)
                        -> Table scan on <subquery2>  (cost=827.25..836.86 rows=570) (actual time=4.286..4.345 rows=356 loops=1)
                            -> Materialize with deduplication  (cost=827.24..827.24 rows=570) (actual time=4.282..4.282 rows=356 loops=1)
                                -> Filter: (TrendingStats.viewCount > 1000000)  (cost=770.21 rows=570) (actual time=0.031..3.509 rows=1649 loops=1)
                                    -> Index range scan on TrendingStats using idx_likecount over (100000 < likeCount), with index condition: (TrendingStats.likeCount > 100000)  (cost=770.21 rows=1711) (actual time=0.030..3.342 rows=1711 loops=1)
                        -> Single-row index lookup on v using PRIMARY (videoID=`<subquery2>`.videoID)  (cost=142.67 rows=1) (actual time=0.002..0.002 rows=1 loops=356)
 |
+------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
-----------------------------------+
1 row in set (0.01 sec)
```