

# CNIT 372

Hunter King and Brian Kim

# Project Background

Data Type: Option A

Consisted of Youtube's trending video dataset that was collected until August 19th

Motivation: Wanted to analyze and see which videos has the most views in different conditions.

# Database

## Video

- VideoID (Primary Key) - int
- Title - varchar
- Description - text
- Genre - varchar
- UploadDate - date
- ChannelID (Foreign Key) - int

## Channel

- ChannelID (Primary Key) - int
- ChannelName - varchar
- SubscriberCount - int
- ChannelCategory - varchar

## Interaction

- InteractionID (Primary Key) - int
- VideoID (Foreign Key) - int
- ViewCount - int
- Likes - int
- Dislikes - int
- CommentCount - int
- TrendingRank - int

# Relationship

- Video - VideoID is the primary key, ChannelID is the foreign key referencing the table Channel.  
Foreign key ChannelID makes a one to many relationship between the tables Channel and Video.
- Channel - ChannelID is the primary key
- Interaction - InteractionID is the primary key, VideoID is the foreign key referencing the table Video.

# Problems

Asked 10 questions

Each had different conditions

Each had different context

# Solutions

What video has the highest like to view ratio?

**Answer=**

```
SELECT Title, ((LIKES / VIEWS)*100) AS LIKE_TO_VIEW_RATIO
```

```
FROM youtube
```

```
WHERE VIEWS > 0
```

```
ORDER BY LIKE_TO_VIEW_RATIO DESC
```

```
FETCH FIRST ROW ONLY;
```

TITLE	LIKE_TO_VIEW_RATIO
Bruno Mars - Finesse (Remix) [Feat. Cardi B] [Official Video]	29.046646

# Solutions

What video was most popular (had the most views) on November 15th?

**Answer=**

```
SELECT VIDEO_ID, TITLE, MAX(VIEWS) AS MAX_VIEWS
```

```
FROM youtube
```

```
WHERE TRENDING_DATE = '17.15.11' -- Use the correct date format as per your database
```

```
GROUP BY VIDEO_ID, TITLE
```

```
ORDER BY MAX_VIEWS DESC
```

```
FETCH FIRST ROW ONLY;
```

VIDEO_ID	TITLE	MAX_VIEWS
2Vv-BfVoq4g	Ed Sheeran - Perfect (Official Music Video)	39082222

# Solutions

What video from each genre has the highest views?

**Answer=**

```
WITH MaxViewsPerCategory AS (
```

```
    SELECT CATEGORY_ID, MAX(VIEWS) AS MAX_VIEWS
```

```
    FROM youtube
```

```
    GROUP BY CATEGORY_ID
```

```
)
```

```
SELECT t.VIDEO_ID, SUBSTR(t.TITLE, 1, 70) AS SHORT_TITLE, t.CATEGORY_ID, t.VIEWS
```

```
FROM youtube t
```

```
JOIN MaxViewsPerCategory mv ON t.CATEGORY_ID = mv.CATEGORY_ID AND t.VIEWS = mv.MAX_VIEWS;
```



VIDEO_ID	SHORT_TITLE	CATEGORY_ID	VIEWS
nWZMfPP34g8	BITTEN by a GIANT DESERT CENTIPEDE!	15	6187457
_9YmpuLDnwo	42 HOLY GRAIL HACKS THAT WILL SAVE YOU A FORTUNE	26	54155921
F1sCjmMhFmw	YouTube Rewind: The Shape of 2017   #YouTubeRewind	24	149376127
VRJmcxCrAOA	Real Life Trick Shots 2   Dude Perfect	17	29090799
oWjxSkJpxFU	Suicide: Be Here Tomorrow.	29	24286474
cgg5oc20nxk	Official Ram Trucks Super Bowl Commercial   Icelandic Vikings   We Will	2	25244097
BhIEIO0vaBE	To Our Daughter	22	56111957
3yvicqMrAHvQ	Turkish Airlines - 5 Senses with Dr. Oz	19	23932421
c8W-augg024	We Sent Garlic Bread to the Edge of Space, Then Ate It	27	7349435
uVcvhKTVZ8g	Clash Royale: CLAN WARS IS HERE!	20	16935442
4kDklB6-69Q	Poo Bear ft. Justin Bieber & Jay Electronica - Hard 2 Face Reality (Ly	23	29178096

VIDEO_ID	SHORT_TITLE	CATEGORY_ID	VIEWS
7_FJUSBFbJM	Game of Zones - Game of Zones - S5:E1: 'A Golden Summer'	43	1445949
Kc6CIwBVo5s	Hawaiian lava flows are faster than a turtle™	25	10277358
yDiXQl7grPQ	Do You Hear Yanny or Laurel? (SOLVED with SCIENCE)	28	42799458
ulNswX3If6U	Selena Gomez - Back To You (Lyric Video)	1	54863912
VYOjWnS4cMY	Childish Gambino - This Is America (Official Video)	10	225211923

15 rows selected

# Solutions

What video has the highest ratio of dislikes to views?

**Answer=**

```
SELECT VIDEO_ID,  
       substr(TITLE,1,50) as title,  
       (DISLIKES / NULLIF(VIEWS, 0)) AS DISLIKE_TO_VIEW_RATIO  
FROM youtube  
  
ORDER BY DISLIKE_TO_VIEW_RATIO DESC  
  
FETCH FIRST ROW ONLY;
```

VIDEO_ID	TITLE	DISLIKE_TO_VIEW_RATIO
LFhI6H6pRWg	PSA from Chairman of the FCC Ajit Pai	.191531483

# Solutions

What video has the most views by its date?

**Answer=**

WITH MonthlyTopVideos AS (

SELECT

EXTRACT(YEAR FROM TO\_DATE(TRENDING\_DATE, 'YY-DD-MM')) AS YEAR,

EXTRACT(MONTH FROM TO\_DATE(TRENDING\_DATE, 'YY-DD-MM')) AS MONTH,

VIDEO\_ID, TITLE, VIEWS, ROW\_NUMBER() OVER (

PARTITION BY EXTRACT(YEAR FROM TO\_DATE(TRENDING\_DATE, 'YY-DD-MM')),

EXTRACT(MONTH FROM TO\_DATE(TRENDING\_DATE, 'YY-DD-MM'))

ORDER BY VIEWS DESC) AS RN FROM youtube

)

SELECT YEAR, MONTH, VIDEO\_ID, TITLE,

VIEWS AS MAX\_VIEWS FROM MonthlyTopVideos

WHERE RN = 1 ORDER BY YEAR, MONTH;

YEAR	MONTH	VIDEO_ID	TITLE	MAX_VIEWS
2017	11	TyHvyGVs42U	Luis Fonsi, Demi Lovato - À%oçhame La Culpá	102012605
2017	12	FlsCjmMhFmw	YouTube Rewind: The Shape of 2017   #YouTubeRewind	149376127
2018	1	LsoLEjrDogU	Bruno Mars - Finesse (Remix) [Feat. Cardi B] [Official Video]	57951412
2018	2	BhIEIO0vaBE	To Our Daughter	56111957
2018	3	#NAME?	Maluma - El PrÃestamo (Official Video)	87264467
2018	4	ffxKSjUwKdU	Ariana Grande - No Tears Left To Cry	74523616
2018	5	VYOjWnS4cMY	Childish Gambino - This Is America (Official Video)	217750076
2018	6	VYOjWnS4cMY	Childish Gambino - This Is America (Official Video)	225211923

8 rows selected.

# Solutions

What video has the highest likes?

**Answer=** SELECT VIDEO\_ID, substr(TITLE,1,50), LIKES

FROM youtube

ORDER BY LIKES DESC

FETCH FIRST ROW ONLY;

VIDEO_ID	SUBSTR(TITLE,1,50)	LIKES
7C2z4GqqS5E	BTS (ë°@íf,,i+@ë...,ë<") 'FAKE LOVE' Official MV	5613827

# Solutions

What publish time has the most views?

**Answer=** SELECT

PUBLISH\_TIME,

SUM(VIEWS) AS TOTAL\_VIEWS

FROM

youtube

GROUP BY

PUBLISH\_TIME

ORDER BY

TOTAL\_VIEWS DESC

FETCH FIRST ROW ONLY;

PUBLISH_TIME	TOTAL_VIEWS
2018-05-06T04:00:07.000Z	3758488765

# Solutions

What is the video with the highest amount of dislikes and comments?

**Answer**=SELECT VIDEO\_ID, substr(TITLE,1,50) as title, DISLIKES, COMMENT\_count

FROM youtube

WHERE DISLIKES = (SELECT MAX(DISLIKES) FROM youtube)

AND COMMENT\_count = (SELECT MAX(COMMENT\_count) FROM youtube);

VIDEO_ID	TITLE	DISLIKES	COMMENT_COUNT
QwZT7T-TXI0	So Sorry.	1674420	1361580

# Solutions

What are the top five channels that have the highest views and likes?

**Answer**=SELECT substr(Channel\_TITLE,1,50) as Channel,

SUM(VIEWS) AS TOTAL\_VIEWS,

SUM(LIKES) AS TOTAL\_LIKES

FROM youtube

GROUP BY substr(Channel\_TITLE,1,50)

ORDER BY TOTAL\_VIEWS DESC, TOTAL\_LIKES DESC

FETCH FIRST 5 ROWS ONLY;

CHANNEL	TOTAL_VIEWS	TOTAL_LIKES
ChildishGambinoVEVO	3758488765	96700818
ibighit	2235906679	199247121
Dude Perfect	1870085178	60275557
Marvel Entertainment	1808998971	55873344
ArianaGrandeVevo	1576959172	52170970

# Solutions

What video has the highest views with comments disabled?

**Answer**=SELECT VIDEO\_ID, substr(TITLE,1,50) as title, VIEWS

FROM youtube

WHERE COMMENTS\_DISABLED like 'TRUE'

ORDER BY VIEWS DESC

FETCH FIRST ROW ONLY;

VIDEO_ID	TITLE	VIEWS
BhIEI00vaBE	To Our Daughter	56111957



# Conclusion

- Different condition affects the videos views
- There are correlations between the video and the video's viewers
- Not all correlations are explainable