

Youtube Analysis in AWS Athena

■ 1. Joining Tables

```
SELECT * FROM "AwsDataCatalog"."youtube-raw"."raw_statistics" as r
INNER JOIN "AwsDataCatalog"."youtube-cleaned"."cleaned_statistics_reference_data" as c
ON r.category_id = c.id LIMIT 10;
```

■ 2. Top 10 most viewed videos:

```
WITH ranked_videos AS (
  SELECT
    video_id,
    title,
    views,
    trending_date,
    channel_title,
    ROW_NUMBER() OVER (PARTITION BY title ORDER BY views DESC) AS rn
  FROM "AwsDataCatalog"."youtube-raw"."raw_statistics"
)
SELECT
  video_id,
  title,
  views,
  trending_date,
  channel_title
FROM ranked_videos
WHERE rn = 1
ORDER BY views DESC
LIMIT 10;
```

video_id	title	views	trending_date
_I_D_8Z4sJE	"Nicky Jam x J. Balvin - X (EQUIS) Video Oficial Prod. Afro Bros & Jeon"	424538912	18.07.04
kLpH1nSLJSs	"Bad Bunny - Amorfoda Video Oficial"	328860380	18.23.03
wfWkmURBNv8	"Ozuna x Romeo Santos - El Farsante Remix"	288811992	18.01.03
VYOjWnS4cMY	"Childish Gambino - This Is America (Official Video)"	259721696	18.11.06
xpVfcZ0ZcFM	"Drake - God's Plan"	258164991	18.25.03
ffxKSjUwKdU	"Ariana Grande - No Tears Left To Cry"	208876887	18.26.05
FlsCjmMhFmw	"YouTube Rewind: The Shape of 2017 #YouTubeRewind"	169884583	17.24.12
sGIm0-dQd8M	"Dura - Daddy Yankee (Video Oficial)"	167456025	18.14.02
2Vv-BfVoq4g	"Ed Sheeran - Perfect (Official Music Video)"	138578860	17.27.11
M4ZoCHID9GI	"The Weeknd - Call Out My Name (Official Video)"	138535053	18.19.05

■ 3. Top 10 most liked videos:

WITH ranked_videos AS (

SELECT

video_id,

title,

likes,

trending_date,

channel_title,

ROW_NUMBER() OVER (PARTITION BY title ORDER BY likes DESC) AS rn

FROM "AwsDataCatalog"."youtube-raw"."raw_statistics"

)

SELECT

video_id,

title,

likes,

trending_date,

channel_title

FROM ranked_videos

WHERE rn = 1

ORDER BY likes DESC

LIMIT 10;

video_id	title	likes	trending_date
zEf423kYfqk	"Becky G	200862743	18.26.05
TyHvyGVs42U	"Luis Fonsi	143408235	17.01.12
DkeiKbqa02g	"Calvin Harris	99282704	18.02.06
xTlNMmZKwpA	"Cardi B	94254507	18.14.06
WtE011iVx1Q	"Sebastián Yatra - Por Perro ft. Luis Figueroa	92838957	18.05.05
dfnCAmr569k	"Taylor Swift - End Game ft. Ed Sheeran	87043012	18.06.02
JQbjS0_ZfJ0	"Kendrick Lamar	69257303	18.15.03
lwz4P8HfGVM	"David Bisbal	67670181	18.21.04
31OMf1H5Rk0	"Matt Hunter	43761370	18.09.06
QgOXIEhHU1Y	"Diplo	40087764	18.14.06

■ 4. Top 10 most disliked videos:

WITH ranked_videos AS (

SELECT

video_id,

title,

dislikes,

trending_date,

channel_title,

ROW_NUMBER() OVER (PARTITION BY title ORDER BY dislikes DESC) AS rn

FROM "AwsDataCatalog"."youtube-raw"."raw_statistics"

)

SELECT

video_id,

title,

dislikes,

```

trending_date,
channel_title
FROM ranked_videos
WHERE rn = 1
ORDER BY dislikes DESC
LIMIT 10;

```

video_id ▾	title ▾	dislikes ▾	trending_date ▾
liczsmUSMA4	"J. Balvin	70001631	18.14.02
LPTlvQ1Zet0	"21 Savage	69907195	18.07.04
y0PMqR7IbAQ	"Ozuna	54067360	18.07.04
95Rzsn-zjr0	"Rae Sremmurd	46514582	18.18.04
9t9u_yPEidY	"Jennifer Lopez - Amor	31815783	17.25.11
p03TIGqEc8o	"Sofia Reyes - 1	31371692	18.23.03
8Y5-GI9hYcl	"Abraham Mateo	29075706	18.15.05
j-I0Uxbh1uY	"Heeriye Song Video - Race 3 Salman Khan	25952473	18.24.05
9v_rtaye2yY	"Migos	23575103	17.15.12
tjA7nAHOAww	"LSD - Audio (Official Video) ft. Sia	21986883	18.08.06

■ 5. Video with the highest engagement (likes + dislikes + comments):

```

WITH ranked_videos AS (
SELECT
video_id,
title,
likes + dislikes + comment_count AS engagement,
trending_date,
channel_title,
ROW_NUMBER() OVER (PARTITION BY title ORDER BY likes + dislikes + comment_count DESC) AS rn
FROM "AwsDataCatalog"."youtube-raw"."raw_statistics"
)

```

```

SELECT
  video_id,
  title,
  engagement,
  trending_date,
  channel_title
FROM ranked_videos
WHERE rn = 1
ORDER BY engagement DESC
LIMIT 10;

```

video_id ▼	title ▼	engagement ▼	trending_date ▼
zEf423kYfqk	"Becky G	202673730	18.26.05
TyHvyGVs42U	"Luis Fonsi	146232342	17.01.12
DkeiKbqa02g	"Calvin Harris	100332672	18.02.06
xTlNMmZKwpA	"Cardi B	96173734	18.14.06
WtE011iVx1Q	"Sebastián Yatra - Por Perro ft. Luis Figueroa	93380455	18.05.05
dfnCAmr569k	"Taylor Swift - End Game ft. Ed Sheeran	89322173	18.06.02
JQbjS0_ZfJ0	"Kendrick Lamar	70303146	18.15.03
lwz4P8HfGVM	"David Bisbal	68125113	18.21.04
31OMf1H5Rk0	"Matt Hunter	44615437	18.09.06
QgOXIEhHU1Y	"Diplo	40948704	18.14.06

■ 6. Total views per region:

WITH ranked_videos AS (

```

SELECT
  r.video_id,
  r.title,
  r.views,
  r.trending_date,

```

```

    r.channel_title,
    r.region,
    ROW_NUMBER() OVER (PARTITION BY r.title, r.region ORDER BY r.views DESC) AS rn
FROM "AwsDataCatalog"."youtube-raw"."raw_statistics" AS r
)
SELECT
    rv.region,
    SUM(rv.views) AS total_views
FROM ranked_videos AS rv
WHERE rv.rn = 1
GROUP BY rv.region
ORDER BY total_views DESC;

```

region	total_views
ca	18040783997
gb	13004397698
de	12608109057
us	10953065827
in	9379977506
fr	9377412246
mx	8320138621
ru	6024625462
kr	5292054396
jp	2773418256

■ 7. Top 5 channels by total video views:

```
SELECT
  channel_title,
  SUM(views) AS total_views
FROM
  "AwsDataCatalog"."youtube-raw"."raw_statistics"
GROUP BY
  channel_title
ORDER BY
  total_views DESC
LIMIT 5;
```

channel_title	total_views
"ChildishGambinoVEVO"	10652536179
"NickyJamTV"	9237321281
"Ozuna"	8462112733
"DrakeVEVO"	7468532897
"Marvel Entertainment"	7426837076

■ 8. Average views per video for each channel:

```
SELECT
  channel_title,
  AVG(views) AS average_views
FROM
  "AwsDataCatalog"."youtube-raw"."raw_statistics"
GROUP BY
  channel_title
ORDER BY
```

average_views DESC;

channel_title	average_views
"Bad Bunny"	1.6244876827906978E8
"ChildishGambinoVEVO"	1.3657097665384614E8
"NickyJamTV"	8.714454038679245E7
"Ozuna"	8.63480891122449E7
"DrakeVEVO"	8.584520571264368E7
"ArianaGrandeVevo"	6.6966598640449435E7
"YouTube Spotlight"	6.5737038421875E7
"RomeoSantosVEVO"	6.3169863428571425E7
"SebastianYatraVEVO"	6.042503061904762E7
YouTube Spotlight	5.696768784E7

■ 9. Number of videos with disabled comments

```
SELECT
  COUNT(*) AS comments_disabled_videos
FROM "AwsDataCatalog"."youtube-raw"."raw_statistics"
WHERE comments_disabled = TRUE;
```

comments_disabled_videos

7590

■ 10. Number of videos without ratings

```
SELECT
  COUNT(*) AS ratings_disabled_videos
FROM "AwsDataCatalog"."youtube-raw"."raw_statistics"
WHERE ratings_disabled = TRUE;
```


ratings_disabled_videos

6936