

📌 Milestone 1 | YouTube Trending Videos

INTRODUCTION: Welcome to your first Milestone. Milestones are a great opportunity for you to practice your skills, both in using SQL, but also in interpreting the information that comes out of the queries you write.

In this Milestone, you will practice the fundamentals of making queries into a SQL database, including using keywords to specify how much data is retrieved and whether or not it is sorted. We will focus on a real-world data set capturing popular YouTube videos (the same one Dr. Alvarez used in his applied lesson). You will pull out information about these videos' views, likes, dislikes, and comments and use the query outputs to make some observations about what separates the top videos from the rest.

HOW IT WORKS: Follow the prompts in the questions below to investigate your data. Post your answers in the provided boxes: the **yellow boxes** for the queries you write, and **blue boxes** for text-based answers. When you're done, export your document as a pdf file and submit it on the Milestone page – see instructions for creating a PDF at the end of the Milestone.

RESOURCES: If you need hints on the Milestone or are feeling stuck, there are multiple ways of getting help. Attend Drop-In Hours to work on these problems with your peers, or reach out to the HelpHub if you have questions. Good luck!

PROMPT: You've been hired by a YouTube content creator to analyze trends on YouTube. Your employer is interested in the patterns of views, likes, dislikes, and comments earned by YouTube videos that make the top trending list.

SQL App: [Here's that link](#) to our specialized SQL app, where you'll write your SQL queries and interact with the data.

– Data Set **Description**

The YouTube Trending Videos (**youtube.trending**) consists of 6351 videos that were listed in the Trending Videos in the United States, recorded between November 2017 and June 2018. There are 16 columns in the dataset; we'll be working with the following columns in this skill builder: **title**, **channel_title**, **views**, **likes**, and **dislikes**.

– **Task 1:** Top User Engagement

To start, you've been asked to look at the videos with the highest levels of user engagement, in terms of likes, dislikes, and comments.

- A.** Write a query that returns these columns: title, channel_title, views, likes, dislikes, and comment_count. Run your query then copy the query into the box below

```
SELECT title, channel_title, views, likes, dislikes,  
comment_count  
FROM youtube.trending
```

- B.** Add **ORDER BY** to find the video that has the highest number of likes. What is the name of that video? Post your query into the yellow box, and the name of the most-liked video in the blue box.

```
SELECT title, channel_title, views, likes, dislikes,  
comment_count  
FROM youtube.trending  
ORDER BY likes DESC  
LIMIT 1;
```



BTS (방탄소년단) 'FAKE LOVE' Official MV

- C. What is the name of the video with the highest number of dislikes? (a query goes in the yellow box, a text answer in the blue box.)

```
SELECT title, channel_title, views, likes, dislikes,  
comment_count  
FROM youtube.trending  
ORDER BY dislikes DESC  
LIMIT 1;
```

```
So Sorry.  
  
Logan Paul Vlogs
```

- D. What is the name of the video with the highest number of comments?

```
SELECT title, channel_title, views, likes, dislikes,  
comment_count  
FROM youtube.trending  
ORDER BY comment_count DESC  
LIMIT 1;
```

```
So Sorry.  
  
Logan Paul Vlogs
```

– Task 2: Comments Counts Large and Small

Your employer wants to go further into the patterns of user engagement via comments on top trending videos.

- A.** Continuing from the queries of the previous task, modify the query to return only the top 10 videos in terms of comment count.

```
SELECT title, channel_title, views, likes, dislikes,  
comment_count  
FROM youtube.trending  
ORDER BY comment_count DESC  
LIMIT 10;
```

- B.** How many comments are on the 10th-most-commented video? What is the ratio of this comment count to the top commented video (from Task 1D)? (The ratio is obtained by dividing the first number by the second. This should be done with a calculator outside of SQL using what was returned from the part A query, and not with a new SQL query.)

```
371864/1361580= .273
```

- C.** Let's dig deeper down the rankings. What is the number of comments on the 100th-ranked video? Use the `OFFSET` keyword to skip past the top 99 videos so that the first row returned will be the 100th rank. (In other words, don't just change the `LIMIT` to 100 and check the last row returned.)

```
SELECT title, channel_title, views, likes, dislikes,  
comment_count  
FROM youtube.trending  
ORDER BY comment_count DESC  
LIMIT 1 OFFSET 99;
```

53665

- D. Okay, let's take one more step down the rankings. What is the number of comments on the 1000th-ranked video?

```
SELECT title, channel_title, views, likes, dislikes,  
comment_count  
FROM youtube.trending  
ORDER BY comment_count DESC  
LIMIT 1 OFFSET 999;
```

7155

– Level Up

- A. Let's reflect on the data we just looked at. In each step from part B through D, we looked at the 10th, 100th, and 1000th most-commented videos, a 10-fold increase in rank number. How different are the videos from one another at the top rankings compared to those in the middle rankings in terms of comment count? Write a sentence or two to summarize what your

takeaways are. (Feel free to run extra queries on your own if it will help build your understanding or intuition of the trends in the data!)

It seems that people comment more on controversial videos. We see this with multiple Logan Paul videos and Childish Gambino "This is America." Music videos have a big trend of comments throughout the dataset.

– Submission

Great work completing your first Milestone! To submit your completed Milestone, you will need to download / export this document as a PDF and then upload it to the Milestone submission page. You can find the option to download as a PDF from the File menu in the upper-left corner of the Google Doc interface.