 **🧪 Lab**  | SQL YouTube Trending Videos

**INTRODUCTION:** Welcome to your first lab in this module. Labs 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 lab, 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 you saw back in Module 2: Excel!). 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.

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.

**—** 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 lab: **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.

1. 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.

(paste your query below 👇)

|  |
| --- |
| select  title,  channel\_title,  views,  likes,  dislikes,  comment\_count  from youtube\_trending |

1. 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.

(paste your query below 👇)

|  |
| --- |
| select  title,  channel\_title,  views,  likes,  dislikes,  comment\_count  from youtube\_trending  order by 4 DESC |

(write your **answer** below 👇)

|  |
| --- |
| BTS (방탄소년단) 'FAKE LOVE' Official MV |

1. Modify the ORDER BY line in your query to find the video that has the highest number of dislikes. What is the name of that video? (As before, a query goes in the yellow box, a text answer in the blue box.)

(paste your query below 👇)

|  |
| --- |
| Select   title,  channel\_title,  views,  likes,  dislikes,  comment\_count  from youtube\_trending  order by 5 DESC |

(write your **answer** below 👇)

|  |
| --- |
| So Sorry. |

1. Modify the ORDER BY line in your query once more to find the name of the video with the highest number of comments? What is the video?

(paste your query below 👇)

|  |
| --- |
| select  title,  channel\_title,  views,  likes,  dislikes,  comment\_count  from youtube\_trending  order by 6 DESC |

(write your **answer** below 👇)

|  |
| --- |
| So Sorry. |

**— 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.

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

(paste your query below 👇)

|  |
| --- |
| select  title,  channel\_title,  views,  likes,  dislikes,  comment\_count  from youtube\_trending  order by 6 DESC  limit 10 |

1. 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.)

(write your **answer** below 👇)

|  |
| --- |
| 371864/1361580 = 0.273 |

1. 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.)

(paste your query below 👇)

|  |
| --- |
| select  channel\_title,  views,  likes,  dislikes,  comment\_count  from youtube\_trending  order by 5 DESC  LIMIT 1 offset 99 |

(write your **answer** below 👇)

|  |
| --- |
| 53665 |

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

(paste your query below 👇)

|  |
| --- |
| select  channel\_title,  views,  likes,  dislikes,  comment\_count  from youtube\_trending  order by 5 DESC  LIMIT 1 offset 999 |

(write your **answer** below 👇)

|  |
| --- |
| 7155 |

**— LevelUp**

1. 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!)

(write your **answer** below 👇)

|  |
| --- |
| With a decrease in comment counts we can see that there is also a decrease in the number of like that video received. |