

YouTube Trending Videos

SkillBuilder

This SkillBuilder uses data about trending videos on YouTube in the USA, from late 2017 to mid-2018. This data can be found in the table **YouTubeVideos** in the SkillBuilder workbook.

Each video appears at most once in the data table; information such as views and likes are based on the last day that it appears in the trending ranking. Note as well on the Questions sheet two additional items: a summary table and a chart. Both of these are based off of the YouTubeVideos data. Consider these a preview of some of the data summaries you'll learn how to create through this course!

Part 1 - Understanding the Data

Before we do analysis, let's take some time to understand the data we're working with. Be sure to make use of the data dictionary to get more information about the data columns in the dataset.

1. How many columns take on numeric values?
2. Are there any numeric columns that should NOT be treated like standard numbers? (In other words, the numeric values aren't meaningful, and are just arbitrary.) If so, which ones are these?
3. How many columns take on text values?
4. Are there any text columns that have some intrinsic ordering that lets them be interpreted as more than just text? (For example, text that describes clothing sizes has an inherent ordering from Extra Small, Small, Medium, Large, and Extra Large, rather than just being sortable alphabetically.) If so, which ones are these?
5. How many columns take on boolean values?

Part 2 - Using Functions

Now it's time to look into the actual data. Use functions to answer the following questions:

6. How many videos are there in the dataset?
7. What is the total number of likes across all trending videos in the dataset?
8. What is the average number of views on a trending video in this dataset?

Part 3 - Sorting and Filtering

Use sorting and filtering directly on the YouTubeVideos table to answer the following questions

9. How many trending videos were published before 2010?
 - Use a sort or filter on the publish_time column (Column C).
10. What is the name of the video with the most views that is NOT from the Music or Entertainment categories?
 - Video categories can be found in the category_name column (Column G)
 - Don't forget to clear any previous filters so that you are making a judgment on all of the data.
11. How many videos in the Gaming category have more than 10 million views?
12. How many videos in the Nonprofits & Activism category have more than the average number of views (from question 8)?

Part 4 - Thinking about the Data

This final question doesn't require any further analysis of the data. Instead, you should think about some of the things you saw in Part 3 of the SkillBuilder, along with some additional supplemental material provided to you in the Questions sheet. There is one summary table that shows the number of videos and average number of views for each video category, and one histogram that shows the distribution of how many views each video got. If you don't know how to read these right now, that's okay! Consider this a preview of things you'll learn as you go through the course. Just do your best to think about the data to answer the following question.

13. Based on what you've observed about view counts in trending videos, would you think of a video with 2 million views as being a particularly exceptional video? Why or why not?