

## Practice Questions: Data Filtration on YouTube Dataset

### 1. Filter Videos by Category

Hint: Use the ``WHERE`` or ``filter()`` clause on the ``category`` column in SQL or Python/Pandas.

### 2. Videos with More Than 1 Million Views

Hint: Check the ``views`` column for values greater than ``1,000,000``.

### 3. Top 5 Videos with Highest Likes

Hint: Sort the dataset by the ``likes`` column in descending order and select the first 5 rows.

### 4. Find Videos Uploaded After 2020

Hint: Use the ``upload_date`` column to filter dates greater than ``2020-01-01``.

### 5. Filter Videos by Specific Creator

Hint: Filter the dataset where the ``creator`` column matches "TechGuru."

### 6. Videos with Dislike Count Above 10% of Views

Hint: Calculate a derived column or condition: ``dislikes / views > 0.1``.

### 7. Videos with Tags Containing "Python"

Hint: Use a string search function (``LIKE`` in SQL or ``str.contains()`` in Python) on the ``tags`` column.

### 8. Short Videos Under 5 Minutes

Hint: Use the ``duration`` column, ensuring time format consistency (e.g., seconds or minutes).

### 9. Videos with Comments Disabled

Hint: Check for a boolean flag or a ``comments_disabled`` column value.

### 10. Most Popular Video in Each Category

Hint: Group the dataset by ``category`` and find the maximum value in the ``views`` column for each group.

### 11. Use ``loc`` to Extract All Rows for a Specific Category

Hint: Use `df.loc[df['category'] == 'Education']` in Pandas.

## 12. Select Specific Columns with `iloc`

Hint: Use column index positions with `iloc`, e.g., `df.iloc[:, [0, 2, 4]]` to select the 1st, 3rd, and 5th columns.

## 13. Use `loc` to Filter Rows with Views Greater Than 1 Million

Hint: Use `loc` with a condition, e.g., `df.loc[df['views'] > 1000000]`.

## 14. Retrieve a Subset of Rows and Columns with `iloc`

Hint: Select rows and columns by position, e.g., `df.iloc[0:10, 1:3]` for rows 0-9 and columns 2-3.

## 15. Use `loc` to Update a Specific Value

Hint: Access the desired cell using `loc`, e.g., `df.loc[5, 'likes'] = 10000`.