

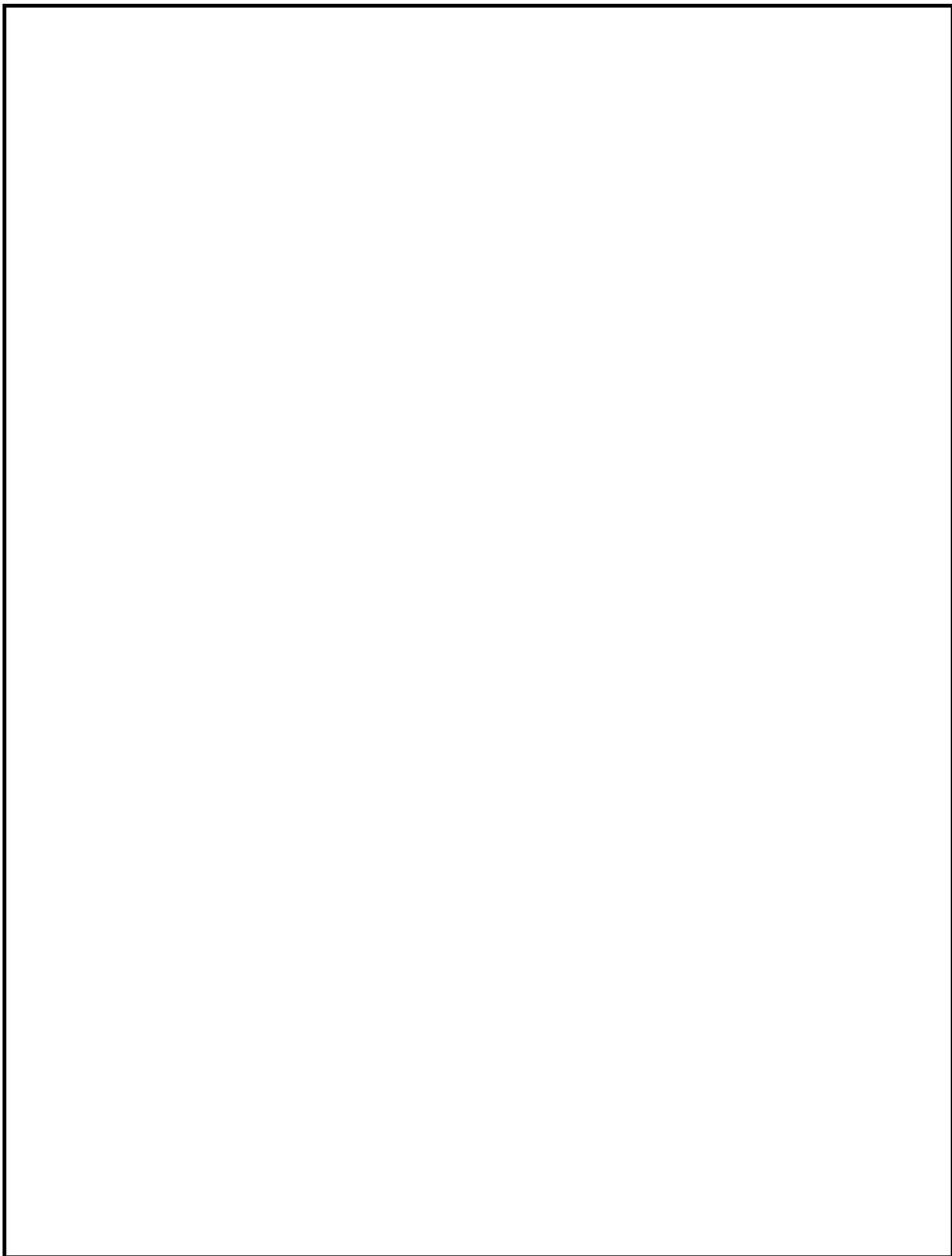
2023

PYTHON PROGRAMMING

BY
KIRAN AIGALI

Table of Contents

1. Import required libraries and read the provided dataset (youtube_dislike_dataset.csv) and retrieve top 5 and bottom 5 records.	1
2. Check the info of the data frame and write your inference on data types and shape of the dataset.	3
3. Check for the percentage of the missing values and drop or impute them.	4
4. Check whether the statistical summary of both numerical and categorical columns and write your inferences.	5
5. Convert datatype of column published at from object to pandas datetime.	6
6. Create a new column as 'published month' using the column published at (display the months only). 7	7
7. Replace the numbers in the column published month as names of the months that is 1 as 'Jan', 2 as 'Feb' and so on.	7
8. Find the number of videos published each month and arrange the months in a decreasing order based on the video count.	8
9. Find the count of unique video id, channel id, and channel title.	8
10. Find the top10 channel names having the highest number of videos in the dataset and the bottom10 having lowest number of videos.	9
11. Find the title of the video which has the maximum number of likes and the title of the video having minimum likes and write your inferences.	9
12. Find the title of the video which has the maximum number of dislikes and the title of the video having minimum dislikes and write your inferences.	10
13. Does the number of views have any effect on how many people disliked the video? Support your answer with a metric and a plot.	11
14. Display all the information about the videos that were published in January and mention the count of videos that were published in January.	12



1. Import required libraries and read the provided dataset (youtube_dislike_dataset.csv) and retrieve top 5 and bottom 5 records.

```
# Import Libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

```
# Read the Dataset
df = pd.read_csv('youtube_dislike_dataset.csv')
df
```

	video_id	title	channel_id	channel_title	published_at	view_count	likes	dislikes	comment_count	
0	--0bCF-iK2E	Jadon Sancho Magical Skills & Goals	UC6UL29enLNe4mqwTfAyeNuW	Bundesliga	2021-07-01 10:00:00	1048888	19515	226	1319	football alemn
1	--14w5SOEUu	Migos - Avalanche (Official Video)	UCGleIM2Dj3zza3xyV3pL3WQ	MigosVEVO	2021-06-10 16:00:00	15352638	359277	7479	18729	Migo Qu Music
2	--40TEbZ9Is	Supporting Actress in a Comedy: 73rd Emmys	UCIBKH8yZRcM4AsRjDVEdjMg	Television Academy	2021-09-20 01:03:32	925281	11212	401	831	
3	--4tfbSyYDE	JO1'YOUNG (JO1 ver.) PERFORMANCE VIDEO	UCsmXiDP8S40uBeJYxyvulmA	JO1	2021-03-03 10:00:17	2641597	39131	441	3745	PRODUC JO1 The!
4	--DKIzWVh-E	Why Retaining Walls Collapse	UCMOqf8ab-42UUQldVoKwjIQ	Practical Engineering	2021-12-07 13:00:00	715724	32887	367	1067	retaini Jer Dir
...	
37417	zsd4ydaFGR0	Lil Tjay - Calling My Phone (feat. 6LACK) [Off...	UCEB4a5o_6KfjxHwNMnmj54Q	Lil Tjay	2021-02-12 05:03:49	120408275	2180780	35871	81360	Lil Callin Call
37418	zziBybeSAtw	PELICANS at LAKERS FULL GAME HIGHLIGHTS Ja...	UCWJ2IWNubArHWmf3FIHbfcQ	NBA	2021-01-16 05:39:05	2841917	20759	1049	2624	NB game-(
37419	zzk09ESX7e0	[MV] (MAMAMOO) - Where Are We Now	UCuhAUMLzJxIP1W7mEk0_6IA	MAMAMOO	2021-06-02 09:00:10	13346678	720854	4426	90616	MAM/ WAW/ WAW W
37420	zzmQEb0Em5I	FELLIPE ESCUDERO- Master Podcast #12	UC8NjnNWMsRqq11NYvHAQb1g	Master Podcast	2020-10-20 20:59:30	252057	19198	1234	1471	master lord vint
37421	zzxPZwaA-8w	Gareth Bale brace secures dramatic comeback on...	UCEg25rdRZXg32iwai6N8I0w	Tottenham Hotspur	2021-05-23 21:00:31	2252090	34063	868	2004	Spur Hotspu

37422 rows × 12 columns

```
# Display top 5 records
```

```
df.head()
```

	video_id	title	channel_id	channel_title	published_at	view_count	likes	dislikes	comment_count	
0	--0bCF-iK2E	Jadon Sancho Magical Skills & Goals	UC6UL29enLNe4mqwTfAyeNuw	Bundesliga	2021-07-01 10:00:00	1048888	19515	226	1319	football soccer f alemn Bundes seaso
1	--14w5SOEU	Migos - Avalanche (Official Video)	UCGIeIM2Dj3zza3xyV3pL3WQ	MigosVEVO	2021-06-10 16:00:00	15352638	359277	7479	18729	Migos Avalor Quality Cor Music/Motown
2	--40TEbZ9Is	Supporting Actress in a Comedy: 73rd Emmys	UCIBKH8yZRcM4AsRjDVEdjMg	Television Academy	2021-09-20 01:03:32	925281	11212	401	831	
3	--4tfbSyYDE	JO1'YOUNG (JO1 ver.) PERFORMANCE VIDEO	UCsmXIDP8S40uBeJYxyuImA	JO1	2021-03-03 10:00:17	2641597	39131	441	3745	PRODUCE101JA JO1 TheSTAR S1
4	--DKKzWVh-E	Why Retaining Walls Collapse	UCMOqf8ab-42UUQIdVoKwjIQ	Practical Engineering	2021-12-07 13:00:00	715724	32887	367	1067	retaining wall Jersey high Direct Coni

```
# Display bottom 5 records
```

```
df.tail()
```

	video_id	title	channel_id	channel_title	published_at	view_count	likes	dislikes	comment_count	
37417	zsd4ydefGR0	Lil Tjay - Calling My Phone (feat. 6LACK) [Off...	UCEB4a5o_6KfjxHwNMnmj54Q	Lil Tjay	2021-02-12 05:03:49	120408275	2180780	35871	81360	Lil Tjay : Calling My Calling M
37418	zziBybeSAtw	PELICANS at LAKERS FULL GAME HIGHLIGHTS Ja...	UCWJ2IWNubArHWmf3FIHbfcQ	NBA	2021-01-16 05:39:05	2841917	20759	1049	2624	NBA G L Bas game-00220 L
37419	zzk09ESX7e0	[MV] (MAMAMOO) - Where Are We Now	UCuhAUMLzJxlP1W7mEk0_6IA	MAMAMOO	2021-06-02 09:00:10	13346678	720854	4426	90616	MAMAMOC WAW MAM WAW Whe We
37420	zzmQEb0Em5I	FELLIPE ESCUDERO- Master Podcast #12	UC8NjnNWMsRqq11NYvHAQb1g	Master Podcast	2020-10-20 20:59:30	252057	19198	1234	1471	masterpodca lord vinhete
37421	zzxPZwaA-8w	Gareth Bale brace secures dramatic comeback on...	UCEg25rdRZXg32iwai6N6I0w	Tottenham Hotspur	2021-05-23 21:00:31	2252090	34063	888	2004	Spurs Tott Hotspur Tott Leice

This code imports the panda's library, reads the dataset, and then displays the top5 and the bottom5 records from the dataset.

2. Check the info of the data frame and write your inference on data types and shape of the dataset.

```
# Display information about the dataset
```

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 37422 entries, 0 to 37421
Data columns (total 12 columns):
#   Column          Non-Null Count  Dtype
---  ---
0   video_id        37422 non-null  object
1   title           37422 non-null  object
2   channel_id      37422 non-null  object
3   channel_title   37422 non-null  object
4   published_at    37422 non-null  object
5   view_count      37422 non-null  int64
6   likes           37422 non-null  int64
7   dislikes        37422 non-null  int64
8   comment_count   37422 non-null  int64
9   tags            37422 non-null  object
10  description      37422 non-null  object
11  comments        37264 non-null  object
dtypes: int64(4), object(8)
memory usage: 3.4+ MB
```

```
df.dtypes
```

```
video_id      object
title         object
channel_id    object
channel_title object
published_at  object
view_count    int64
likes         int64
dislikes      int64
comment_count int64
tags          object
description   object
comments      object
dtype: object
```

```
df.shape
```

```
(37422, 12)
```

The shape of the dataset is (37422, 12), which it has 37,422 rows and 12 columns. This information confirms the number of records and attributes present in your dataset.

3. Check for the percentage of the missing values and drop or impute them.

```
# Percentage of missing values & drop or impute them
```

```
missing_percent = (df.isnull().sum() / len(df)) * 100  
print("Percentage of missing values:")  
print(missing_percent)
```

```
Percentage of missing values:  
video_id      0.000000  
title         0.000000  
channel_id    0.000000  
channel_title 0.000000  
published_at  0.000000  
view_count    0.000000  
likes         0.000000  
dislikes      0.000000  
comment_count 0.000000  
tags          0.000000  
description   0.000000  
comments      0.422212  
dtype: float64
```

```
# Drop or impute missing values
```

```
df['comments'].fillna('', inplace = True)  
threshold = 30  
df.dropna(axis = 1, thresh = len(df) * (1 - threshold / 100), inplace = True)  
df.dropna(axis = 0, inplace = True)  
print((df.isnull().sum() / len(df)) * 100)
```

```
video_id      0.0  
title         0.0  
channel_id    0.0  
channel_title 0.0  
published_at  0.0  
view_count    0.0  
likes         0.0  
dislikes      0.0  
comment_count 0.0  
tags          0.0  
description   0.0  
comments      0.0  
dtype: float64
```

In the code above, the missing values are calculated as a percentage for each column. Then we can decide whether to drop or impute them based on our analysis. To impute the missing values in the 'Comments' column with an empty string and to drop the columns and rows based on a threshold percentage.

4. Check whether the statistical summary of both numerical and categorical columns and write your inferences.

```
# Numeric summary
numerical_summary = df.describe()
print("Statistical summary of numerical columns:")
print(numerical_summary)
```

```
Statistical summary of numerical columns:
```

	view_count	likes	dislikes	comment_count
count	3.742200e+04	3.742200e+04	3.742200e+04	3.742200e+04
mean	5.697838e+06	1.668147e+05	4.989862e+03	9.924930e+03
std	2.426622e+07	5.375670e+05	3.070824e+04	1.171003e+05
min	2.036800e+04	0.000000e+00	0.000000e+00	0.000000e+00
25%	5.122970e+05	1.323350e+04	2.810000e+02	9.000000e+02
50%	1.319078e+06	4.233050e+04	7.960000e+02	2.328000e+03
75%	3.670231e+06	1.304698e+05	2.461750e+03	6.184000e+03
max	1.322797e+09	3.183768e+07	2.397733e+06	1.607103e+07

- View count, likes, dislikes, and comment count are numerical columns with non-negative values, as indicated by the minimum values being greater than or equal to 0.
- You can see the counts, mean, standard deviation, minimum, 25th percentile, median 50th percentile, and 75th percentile, and maximum values for these columns.

```
# Categorical summary
categorical_summary = df.describe(include = ['object']).T
print("\nStatistical summary of categorical columns:")
print(categorical_summary)
```

```
statistical summary of categorical columns:
```

	count	unique	top	freq
video_id	37422	37422	--0bCF-iK2E	1
title	37422	37113	www	21
channel_id	37422	10961	UCNAf1k0yIjyGu3k9BwAg3lg	533
channel_title	37422	10883	Sky Sports Football	533
published_at	37422	36772	2020-10-16 04:00:10	6
tags	37422	28799		3817
description	37422	35630		589
comments	37422	37265		158

- The statistical summary for categorical column includes the count, unique values, topmost frequent values, and the frequency of the top value.
- Columns like video id, title, channel id, channel title, published at, tags, description, comments have these categorical statistics.
- The categorical summary provides the information about the most common values and their frequency, but it doesn't give measures like mean and standard deviation as it does for numerical column.

5. Convert datatype of column published_at from object to pandas datetime.

```
# 5. Convert datatype of column published_at from object to pandas datetime.
```

```
df.columns
```

```
Index(['video_id', 'title', 'channel_id', 'channel_title', 'published_at',  
      'view_count', 'likes', 'dislikes', 'comment_count', 'tags',  
      'description', 'comments'],  
      dtype='object')
```

```
# Convert published_at columns to datetime
```

```
df['published_at'] = pd.to_datetime(df['published_at'])  
print(df['published_at'])
```

```
0      2021-07-01 10:00:00  
1      2021-06-10 16:00:00  
2      2021-09-20 01:03:32  
3      2021-03-03 10:00:17  
4      2021-12-07 13:00:00  
...  
37417   2021-02-12 05:03:49  
37418   2021-01-16 05:39:05  
37419   2021-06-02 09:00:10  
37420   2020-10-20 20:59:30  
37421   2021-05-23 21:00:31  
Name: published_at, Length: 37422, dtype: datetime64[ns]
```

```
# Check the data type
```

```
print(df.dtypes)
```

```
video_id      object  
title         object  
channel_id    object  
channel_title object  
published_at  datetime64[ns]  
view_count    int64  
likes         int64  
dislikes      int64  
comment_count int64  
tags          object  
description   object  
comments      object  
dtype: object
```

6. Create a new column as 'published month' using the column published at (display the months only).

```
# 6. Create new column published_month using published_at column (display only the months)
```

```
df['published_month'] = df['published_at'].dt.month
print(df['published_month'])
```

```
0      7
1      6
2      9
3      3
4     12
...
37417   2
37418   1
37419   6
37420  10
37421   5
Name: published_month, Length: 37422, dtype: int32
```

7. Replace the numbers in the column published month as names of the months that is 1 as 'Jan', 2 as 'Feb' and so on.

```
# 7. Replace the numbers in the column published_month as names of the months i.e., 1 as 'Jan', 2 as 'Feb' and so on.
```

```
# Replace numerical values with month names
```

```
month_mapping = {
    1: 'Jan', 2: 'Feb', 3: 'Mar', 4: 'Apr', 5: 'May', 6: 'Jun',
    7: 'July', 8: 'Aug', 9: 'Sep', 10: 'Oct', 11: 'Nov', 12: 'Dec'
}
```

```
df['published_month'] = df['published_month'].apply(lambda x: month_mapping.get(x))
print(df['published_month'])
```

```
-----
KeyError                                Traceback (most recent call last)
File C:\ProgramData\anaconda3\Lib\site-packages\pandas\core\indexes\base.py:3653, in Index.get_loc(self, key)
    3652 try:
-> 3653     return self._engine.get_loc(casted_key)
    3654 except KeyError as err:

File C:\ProgramData\anaconda3\Lib\site-packages\pandas\_libs\index.pyx:147, in pandas._libs.index.IndexEngine.get_loc()

File C:\ProgramData\anaconda3\Lib\site-packages\pandas\_libs\index.pyx:176, in pandas._libs.index.IndexEngine.get_loc()

File pandas\_libs\hashtable_class_helper.pxi:7080, in pandas._libs.hashtable.PyObjectHashTable.get_item()

File pandas\_libs\hashtable_class_helper.pxi:7088, in pandas._libs.hashtable.PyObjectHashTable.get_item()

KeyError: 'published_month'

The above exception was the direct cause of the following exception:

KeyError                                Traceback (most recent call last)
Cell In[82], line 1
----> 1 df['published_month'] = df['published_month'].apply(lambda x: month_mapping.get(x))
      2 print(df['published_month'])

File C:\ProgramData\anaconda3\Lib\site-packages\pandas\core\frame.py:3761, in DataFrame.__getitem__(self, key)
    3759 if self.columns.nlevels > 1:
    3760     return self._getitem_multilevel(key)
-> 3761 indexer = self.columns.get_loc(key)
    3762 if is_integer(indexer):
    3763     indexer = [indexer]

File C:\ProgramData\anaconda3\Lib\site-packages\pandas\core\indexes\base.py:3655, in Index.get_loc(self, key)
    3653     return self._engine.get_loc(casted_key)
    3654 except KeyError as err:
-> 3655     raise KeyError(key) from err
    3656 except TypeError:
    3657     # If we have a listlike key, _check_indexing_error will raise
    3658     # InvalidIndexError. Otherwise we fall through and re-raise
    3659     # the TypeError.
    3660     self._check_indexing_error(key)

KeyError: 'published_month'
```

8. Find the number of videos published each month and arrange the months in a decreasing order based on the video count.

```
# 8. Find the number of videos published each month and arrange the months in a decreasing order based on the video count.
```

```
video_count_per_month = df.groupby(['published_month'])['video_id'].count().sort_values(ascending = False)
print(video_count_per_month)
```

```
Series([], Name: video_id, dtype: int64)
```

9. Find the count of unique video id, channel id, and channel title.

```
# 9. Find the count of unique video_id, channel_id and channel_title.
```

```
unique_video_id_count = df['video_id'].nunique()
unique_channel_id_count = df['channel_id'].nunique()
unique_channel_title_count = df['channel_title'].nunique()

print("count of unique video_id:", unique_video_id_count)
print("count of unique channel_id:", unique_channel_id_count)
print("count of unique channel_title:", unique_channel_title_count)
```

```
count of unique video_id: 37422
count of unique channel_id: 10961
count of unique channel_title: 10883
```

10. Find the top10 channel names having the highest number of videos in the dataset and the bottom10 having lowest number of videos.

```
# 10. Find the top10 channel names having the highest number of videos in the dataset and the bottom10 having lowest number of videos.

# Top 10 channels

channel_video_count = df['channel_title'].value_counts()

top_10_channels = channel_video_count.head(10)
print("Top 10 channels with the highest number of videos")
print(top_10_channels)
```

```
Top 10 channels with the highest number of videos
channel_title
Sky Sports Football    533
The United Stand      301
BT Sport               246
NBA                   209
NFL                   162
WWE                   122
SSSniperWolf          99
SSundee               98
FORMULA 1             87
NHL                   86
Name: count, dtype: int64
```

```
# Bottom 10 channels

bottom_10_channels = channel_video_count.tail(10)
print("Bottom 10 channels with the highest number of videos")
print(bottom_10_channels)
```

```
Bottom 10 channels with the highest number of videos
channel_title
Eiaz                1
adidas              1
Universitetet i Agder  1
Team Tapia          1
TheVincentMottola    1
Blocktrainer        1
Movie Addicts       1
1 1 minute mom      1
FLORIN CERCEL       1
Master Podcast       1
Name: count, dtype: int64
```

11. Find the title of the video which has the maximum number of likes and the title of the video having minimum likes and write your inferences.

```
# 11. Find the title of the video which has the maximum number of likes and the title of the video having minimum likes and write your inferences.

max_likes_title = df.loc[df['likes'].idxmax(), 'title']
min_likes_title = df.loc[df['likes'].idxmin(), 'title']

print("Title of the video with the maximum number of likes:", max_likes_title)
print("Title of the video with the minimum number of likes:", min_likes_title)
```

```
Title of the video with the maximum number of likes: BTS () 'Dynamite' Official MV
Title of the video with the minimum number of likes: Kim Kardashian's Must-See Moments on "Saturday Night Live" | E! News
```

- The video with maximum number of likes is titled BTS () 'Dynamite' Official MV, and the video with minimum number of likes Kim Kardashian's Must-See Moments on "Saturday Night Live" | E! News
- It's important to consider other factors like view, comments and context of the videos to have a comprehensive understanding of their popularity or reception.

12. Find the title of the video which has the maximum number of dislikes and the title of the video having minimum dislikes and write your inferences.

```
# 12. Find the title of the video which has the maximum number of dislikes and the title of the video having minimum dislikes

max_dislikes_title = df.loc[df['dislikes'].idxmax(), 'title']
min_dislikes_title = df.loc[df['dislikes'].idxmin(), 'title']

print("Title of the video with the maximum number of dislikes:", max_dislikes_title)
print("Title of the video with the minimum number of dislikes:", min_dislikes_title)
```

Title of the video with the maximum number of dislikes: Cuties | Official Trailer | Netflix
 Title of the video with the minimum number of dislikes: Kim Kardashian's Must-See Moments on "Saturday Night Live" | E! News

- Video with the maximum number of dislikes is titled Cuties | Official Trailer | Netflix, While the video with the minimum number of dislikes is titled Kim Kardashian's Must-See Moments on "Saturday Night Live" | E! News
- As with likes and other metrics, its important to consider the context, content, audience perspective when interacting the dislike counts.

13. Does the number of views have any effect on how many people disliked the video? Support your answer with a metric and a plot.

```
# 13. Does the number of views have any effect on how many people disliked the video? Support your answer with a metric and a plot.

# Calculate the correlation between 'view_count' and 'dislikes'

correlation = df['view_count'].corr(df['dislikes'])

plt.figure(figsize = (10, 6))
sns.scatterplot(df = df, x = 'view_count', y = 'dislikes', alpha = 0.5)
plt.title(f"scatter plot of view count vs Dislikes\ncorrelation:{correlation: .2f}")
plt.xlabel('view count')
plt.ylabel('Dislikes')
plt.grid(True)
plt.show()

print("correlation between view count and dislikes:", correlation)
```

```
-----
ValueError                                Traceback (most recent call last)
Cell In[83], line 7
      4 correlation = df['view_count'].corr(df['dislikes'])
      6 plt.figure(figsize = (10, 6))
----> 7 sns.scatterplot(df = df, x = 'view_count', y = 'dislikes', alpha = 0.5)
      8 plt.title(f"scatter plot of view count vs Dislikes\ncorrelation:{correlation: .2f}")
      9 plt.xlabel('view count')

File C:\ProgramData\anaconda3\Lib\site-packages\seaborn\relational.py:742, in scatterplot(data, x, y, hue, size, style, palette, hue_order, hue_norm, sizes, size_order, size_norm, markers, style_order, legend, ax, **kwargs)
    732 def scatterplot(
    733     data=None, *,
    734     x=None, y=None, hue=None, size=None, style=None,
    (...)
    738     **kwargs
    739 ):
    741     variables = _ScatterPlotter.get_semantics(locals())
--> 742     p = _ScatterPlotter(data=data, variables=variables, legend=legend)
    744     p.map_hue(palette=palette, order=hue_order, norm=hue_norm)
    745     p.map_size(sizes=sizes, order=size_order, norm=size_norm)

File C:\ProgramData\anaconda3\Lib\site-packages\seaborn\relational.py:538, in _ScatterPlotter.__init__(self, data, variables, legend)
    529 def __init__(self, *, data=None, variables={}, legend=None):
    530
    531     # TODO this is messy, we want the mapping to be agnostic about
    532     # the kind of plot to draw, but for the time being we need to set
    533     # this information so the SizeMapping can use it
    534     self.default_size_range = (
    535         np.r_[.5, 2] * np.square(mpl.rcParams["lines.markersize"])
    536     )
--> 538     super().__init__(data=data, variables=variables)
    540     self.legend = legend

File C:\ProgramData\anaconda3\Lib\site-packages\seaborn\_oldcore.py:640, in VectorPlotter.__init__(self, data, variables)
    635 # var_ordered is relevant only for categorical axis variables, and may
    636 # be better handled by an internal axis information object that tracks
    637 # such information and is set up by the scale_* methods. The analogous
    638 # information for numeric axes would be information about log scaling
    639
    640
    641
    642
    643
    644
    645
    646
    647
    648
    649
    650
    651
    652
    653
    654
    655
    656
    657
    658
    659
    660
    661
    662
    663
    664
    665
    666
    667
    668
    669
    670
    671
    672
    673
    674
    675
    676
    677
    678
    679
    680
    681
    682
    683
    684
    685
    686
    687
    688
    689
    690
    691
    692
    693
    694
    695
    696
    697
    698
    699
    700
    701
    702
    703
    704
    705
    706
    707
    708
    709
    710
    711
    712
    713
    714
    715
    716
    717
    718
    719
    720
    721
    722
    723
    724
    725
    726
    727
    728
    729
    730
    731
    732
    733
    734
    735
    736
    737
    738
    739
    740
    741
    742
    743
    744
    745
    746
    747
    748
    749
    750
    751
    752
    753
    754
    755
    756
    757
    758
    759
    760
    761
    762
    763
    764
    765
    766
    767
    768
    769
    770
    771
    772
    773
    774
    775
    776
    777
    778
    779
    780
    781
    782
    783
    784
    785
    786
    787
    788
    789
    790
    791
    792
    793
    794
    795
    796
    797
    798
    799
    800
    801
    802
    803
    804
    805
    806
    807
    808
    809
    810
    811
    812
    813
    814
    815
    816
    817
    818
    819
    820
    821
    822
    823
    824
    825
    826
    827
    828
    829
    830
    831
    832
    833
    834
    835
    836
    837
    838
    839
    840
    841
    842
    843
    844
    845
    846
    847
    848
    849
    850
    851
    852
    853
    854
    855
    856
    857
    858
    859
    860
    861
    862
    863
    864
    865
    866
    867
    868
    869
    870
    871
    872
    873
    874
    875
    876
    877
    878
    879
    880
    881
    882
    883
    884
    885
    886
    887
    888
    889
    890
    891
    892
    893
    894
    895
    896
    897
    898
    899
    900
    901
    902
    903
    904
    905
    906
    907
    908
    909
    910
    911
    912
    913
    914
    915
    916
    917
    918
    919
    920
    921
    922
    923
    924
    925
    926
    927
    928
    929
    930
    931
    932
    933
    934
    935
    936
    937
    938
    939
    940
    941
    942
    943
    944
    945
    946
    947
    948
    949
    950
    951
    952
    953
    954
    955
    956
    957
    958
    959
    960
    961
    962
    963
    964
    965
    966
    967
    968
    969
    970
    971
    972
    973
    974
    975
    976
    977
    978
    979
    980
    981
    982
    983
    984
    985
    986
    987
    988
    989
    990
    991
    992
    993
    994
    995
    996
    997
    998
    999
    1000
    1001
    1002
    1003
    1004
    1005
    1006
    1007
    1008
    1009
    1010
    1011
    1012
    1013
    1014
    1015
    1016
    1017
    1018
    1019
    1020
    1021
    1022
    1023
    1024
    1025
    1026
    1027
    1028
    1029
    1030
    1031
    1032
    1033
    1034
    1035
    1036
    1037
    1038
    1039
    1040
    1041
    1042
    1043
    1044
    1045
    1046
    1047
    1048
    1049
    1050
    1051
    1052
    1053
    1054
    1055
    1056
    1057
    1058
    1059
    1060
    1061
    1062
    1063
    1064
    1065
    1066
    1067
    1068
    1069
    1070
    1071
    1072
    1073
    1074
    1075
    1076
    1077
    1078
    1079
    1080
    1081
    1082
    1083
    1084
    1085
    1086
    1087
    1088
    1089
    1090
    1091
    1092
    1093
    1094
    1095
    1096
    1097
    1098
    1099
    1100
    1101
    1102
    1103
    1104
    1105
    1106
    1107
    1108
    1109
    1110
    1111
    1112
    1113
    1114
    1115
    1116
    1117
    1118
    1119
    1120
    1121
    1122
    1123
    1124
    1125
    1126
    1127
    1128
    1129
    1130
    1131
    1132
    1133
    1134
    1135
    1136
    1137
    1138
    1139
    1140
    1141
    1142
    1143
    1144
    1145
    1146
    1147
    1148
    1149
    1150
    1151
    1152
    1153
    1154
    1155
    1156
    1157
    1158
    1159
    1160
    1161
    1162
    1163
    1164
    1165
    1166
    1167
    1168
    1169
    1170
    1171
    1172
    1173
    1174
    1175
    1176
    1177
    1178
    1179
    1180
    1181
    1182
    1183
    1184
    1185
    1186
    1187
    1188
    1189
    1190
    1191
    1192
    1193
    1194
    1195
    1196
    1197
    1198
    1199
    1200
    1201
    1202
    1203
    1204
    1205
    1206
    1207
    1208
    1209
    1210
    1211
    1212
    1213
    1214
    1215
    1216
    1217
    1218
    1219
    1220
    1221
    1222
    1223
    1224
    1225
    1226
    1227
    1228
    1229
    1230
    1231
    1232
    1233
    1234
    1235
    1236
    1237
    1238
    1239
    1240
    1241
    1242
    1243
    1244
    1245
    1246
    1247
    1248
    1249
    1250
    1251
    1252
    1253
    1254
    1255
    1256
    1257
    1258
    1259
    1260
    1261
    1262
    1263
    1264
    1265
    1266
    1267
    1268
    1269
    1270
    1271
    1272
    1273
    1274
    1275
    1276
    1277
    1278
    1279
    1280
    1281
    1282
    1283
    1284
    1285
    1286
    1287
    1288
    1289
    1290
    1291
    1292
    1293
    1294
    1295
    1296
    1297
    1298
    1299
    1300
    1301
    1302
    1303
    1304
    1305
    1306
    1307
    1308
    1309
    1310
    1311
    1312
    1313
    1314
    1315
    1316
    1317
    1318
    1319
    1320
    1321
    1322
    1323
    1324
    1325
    1326
    1327
    1328
    1329
    1330
    1331
    1332
    1333
    1334
    1335
    1336
    1337
    1338
    1339
    1340
    1341
    1342
    1343
    1344
    1345
    1346
    1347
    1348
    1349
    1350
    1351
    1352
    1353
    1354
    1355
    1356
    1357
    1358
    1359
    1360
    1361
    1362
    1363
    1364
    1365
    1366
    1367
    1368
    1369
    1370
    1371
    1372
    1373
    1374
    1375
    1376
    1377
    1378
    1379
    1380
    1381
    1382
    1383
    1384
    1385
    1386
    1387
    1388
    1389
    1390
    1391
    1392
    1393
    1394
    1395
    1396
    1397
    1398
    1399
    1400
    1401
    1402
    1403
    1404
    1405
    1406
    1407
    1408
    1409
    1410
    1411
    1412
    1413
    1414
    1415
    1416
    1417
    1418
    1419
    1420
    1421
    1422
    1423
    1424
    1425
    1426
    1427
    1428
    1429
    1430
    1431
    1432
    1433
    1434
    1435
    1436
    1437
    1438
    1439
    1440
    1441
    1442
    1443
    1444
    1445
    1446
    1447
    1448
    1449
    1450
    1451
    1452
    1453
    1454
    1455
    1456
    1457
    1458
    1459
    1460
    1461
    1462
    1463
    1464
    1465
    1466
    1467
    1468
    1469
    1470
    1471
    1472
    1473
    1474
    1475
    1476
    1477
    1478
    1479
    1480
    1481
    1482
    1483
    1484
    1485
    1486
    1487
    1488
    1489
    1490
    1491
    1492
    1493
    1494
    1495
    1496
    1497
    1498
    1499
    1500
    1501
    1502
    1503
    1504
    1505
    1506
    1507
    1508
    1509
    1510
    1511
    1512
    1513
    1514
    1515
    1516
    1517
    1518
    1519
    1520
    1521
    1522
    1523
    1524
    1525
    1526
    1527
    1528
    1529
    1530
    1531
    1532
    1533
    1534
    1535
    1536
    1537
    1538
    1539
    1540
    1541
    1542
    1543
    1544
    1545
    1546
    1547
    1548
    1549
    1550
    1551
    1552
    1553
    1554
    1555
    1556
    1557
    1558
    1559
    1560
    1561
    1562
    1563
    1564
    1565
    1566
    1567
    1568
    1569
    1570
    1571
    1572
    1573
    1574
    1575
    1576
    1577
    1578
    1579
    1580
    1581
    1582
    1583
    1584
    1585
    1586
    1587
    1588
    1589
    1590
    1591
    1592
    1593
    1594
    1595
    1596
    1597
    1598
    1599
    1600
    1601
    1602
    1603
    1604
    1605
    1606
    1607
    1608
    1609
    1610
    1611
    1612
    1613
    1614
    1615
    1616
    1617
    1618
    1619
    1620
    1621
    1622
    1623
    1624
    1625
    1626
    1627
    1628
    1629
    1630
    1631
    1632
    1633
    1634
    1635
    1636
    1637
    1638
    1639
    1640
    1641
    1642
    1643
    1644
    1645
    1646
    1647
    1648
    1649
    1650
    1651
    1652
    1653
    1654
    1655
    1656
    1657
    1658
    1659
    1660
    1661
    1662
    1663
    1664
    1665
    1666
    1667
    1668
    1669
    1670
    1671
    1672
    1673
    1674
    1675
    1676
    1677
    1678
    1679
    1680
    1681
    1682
    1683
    1684
    1685
    1686
    1687
    1688
    1689
    1690
    1691
    1692
    1693
    1694
    1695
    1696
    1697
    1698
    1699
    1700
    1701
    1702
    1703
    1704
    1705
    1706
    1707
    1708
    1709
    1710
    1711
    1712
    1713
    1714
    1715
    1716
    1717
    1718
    1719
    1720
    1721
    1722
    1723
    1724
    1725
    1726
    1727
    1728
    1729
    1730
    1731
    1732
    1733
    1734
    1735
    1736
    1737
    1738
    1739
    1740
    1741
    1742
    1743
    1744
    1745
    1746
    1747
    1748
    1749
    1750
    1751
    1752
    1753
    1754
    1755
    1756
    1757
    1758
    1759
    1760
    1761
    1762
    1763
    1764
    1765
    1766
    1767
    1768
    1769
    1770
    1771
    1772
    1773
    1774
    1775
    1776
    1777
    1778
    1779
    1780
    1781
    1782
    1783
    1784
    1785
    1786
    1787
    1788
    1789
    1790
    1791
    1792
    1793
    1794
    1795
    1796
    1797
    1798
    1799
    1800
    1801
    1802
    1803
    1804
    1805
    1806
    1807
    1808
    1809
    1810
    1811
    1812
    1813
    1814
    1815
    1816
    1817
    1818
    1819
    1820
    1821
    1822
    1823
    1824
    1825
    1826
    1827
    1828
    1829
    1830
    1831
    1832
    1833
    1834
    1835
    1836
    1837
    1838
    1839
    1840
    1841
    1842
    1843
    1844
    1845
    1846
    1847
    1848
    1849
    1850
    1851
    1852
    1853
    1854
    1855
    1856
    1857
    1858
    1859
    1860
    1861
    1862
    1863
    1864
    1865
    1866
    1867
    1868
    1869
    1870
    1871
    1872
    1873
    1874
    1875
    1876
    1877
    1878
    1879
    1880
    1881
    1882
    1883
    1884
    1885
    1886
    1887
    1888
    1889
    1890
    1891
    1892
    1893
    1894
    1895
    1896
    1897
    1898
    1899
    1900
    1901
    1902
    1903
    1904
    1905
    1906
    1907
    1908
    1909
    1910
    1911
    1912
    1913
    1914
    1915
    1916
    1917
    1918
    1919
    1920
    1921
    1922
    1923
    1924
    1925
    1926
    1927
    1928
    1929
    1930
    1931
    1932
    1933
    1934
    1935
    1936
    1937
    1938
    1939
    1940
    1941
    1942
    1943
    1944
    1945
    1946
    1947
    1948
    1949
    1950
    1951
    1952
    1953
    1954
    1955
    1956
    1957
    1958
    1959
    1960
    1961
    1962
    1963
    1964
    1965
    1966
    1967
    1968
    1969
    1970
    1971
    1972
    1973
    1974
    1975
    1976
    1977
    1978
    1979
    1980
    1981
    1982
    1983
    1984
    1985
    1986
    1987
    1988
    1989
    1990
    1991
    1992
    1993
    1994
    1995
    1996
    1997
    1998
    1999
    2000
    2001
    2002
    2003
    2004
    2005
    2006
    2007
    2008
    2009
    2010
    2011
    2012
    2013
    2014
    2015
    2016
    2017
    2018
    2019
    2020
    2021
    2022
    2023
    2024
    2025
    2026
    2027
    2028
    2029
    2030
    2031
    2032
    2033
    2034
    2035
    2036
    2037
    2038
    2039
    2040
    2041
    2042
    2043
    2044
    2045
    2046
    2047
    2048
    2049
    2050
    2051
    2052
    2053
    2054
    2055
    2056
    2057
    2058
    2059
    2060
    2061
    2062
    2063
    2064
    2065
    2066
    2067
    2068
    2069
    2070
    2071
    2072
    2073
    2074
    2075
    2076
    2077
    2078
    2079
    2080
    2081
    2082
    2083
    2084
    2085
    2086
    2087
    2088
    2089
    2090
    2091
    2092
    2093
    2094
    2095
    2096
    2097
    2098
    2099
    2100
    2101
    2102
    2103
    2104
    2105
    2106
    2107
    2108
    2109
    2110
    2111
    2112
    2113
    2114
    2115
    2116
    2117
    2118
    2119
    2120
    2121
    2122
    2123
    2124
    2125
    2126
    2127
    2128
    2129
    2130
    2131
    2132
    2133
    2134
    2135
    2136
    2137
    2138
    2139
    2140
    2141
    2142
    2143
    2144
    2145
    2146
    2147
    2148
    2149
    2150
    2151
    2152
    2153
    2154
    2155
    2156
    2157
    2158
    2159
    2160
    2161
    2162
    2163
    2164
    2165
    2166
    2167
    2168
    2169
    2170
    2171
    2172
    2173
    2174
    2175
    2176
    2177
    2178
    2179
    2180
    2181
    2182
    2183
    2184
    2185
    2186
    2187
    2188
    2189
    2190
    2191
    2192
    2193
    2194
    2195
    2196
    2197
    2198
    2199
    2200
    2201
    2202
    2203
    2204
    2205
    2206
    2207
    2208
    2209
    2210
    2211
    2212
    2213
    2214
    2215
    2216
    2217
    2218
    2219
    2220
    2221
    2222
    2223
    2224
    2225
    2226
    2227
    2228
    2229
    2230
    2231
    2232
    2233
    2234
    2235
    2236
    2237
    2238
    2239
    2240
    2241
    2242
    2243
    2244
    2245
    2246
    2247
    2248
    2249
    2250
    2251
    2252
    2253
    2254
    2255
    2256
    2257
    2258
    2259
    2260
    2261
    2262
    2263
    2264
    2265
    2266
    2267
    2268
    2269
    2270
    2271
    2272
    2273
    2274
    2275
    2276
    2277
    2278
    2279
    2280
    2281
    2282
    2283
    2284
    2285
    2286
    2287
    2288
    2289
    2290
    2291
    2292
    2293
    2294
    2295
    2296
    2297
    2298
    2299
    2300
    2301
    2302
    2303
    2304
    2305
    2306
    2307
    2308
    2309
    2310
    2311
    2312
    2313
    23
```

14. Display all the information about the videos that were published in January and mention the count of videos that were published in January.

```
# 14. Display all the information about the videos that were published in January, and mention the count of videos that were
```

```
videos = df[df['published_at'].dt.month == 1]
```

```
videos
```

	video_id	title	channel_id	channel_title	published_at	view_count	likes	dislikes	comment_count	tags
27	-2Gwm7QIBnE	Q&A With Naisha	UCYwNMbogQFzMcPSuy-pPWg	MianTwins	2021-01-21 00:05:47	872372	38826	239	621	
48	-4s#KSHSxZ	SURPRISING BRENT WITH HIS TIKTOK CRUSH!!	UCPpATKqmMV-CNRNWyaDUwiA	Alexa Rivera	2021-01-16 21:40:04	6504784	262477	5779	7907	
95	-AJD1Fc5rpQ	WE ARE HAVING A BABY! finding out i'm pregna...	UCVSTboAhpnuL6j-IDePWwQ	Tess Christine	2021-01-03 21:53:48	533084	38965	119	1650	
103	-AuJiwjsmWk	Do Ugly Foods Taste Worse? Taste Test	UCzpCo5n9hqIVC7HhPwoIKeg	Good Mythical MORE	2021-01-19 11:00:01	1057077	22526	531	773	gmmgr mythical morn rhetandlinkr
182	-JhqO2KWt5U	Schlatt gets fit	UCWZp4y1jqBuvLtiySs_ZBw	Big guy	2021-01-24 22:50:57	1724985	119431	325	1578	jschlattbig jschlatt highlig schlatt
...
37300	zmzFL5bG-jc	DEVINE MON PERSONNAGE AVANT AKINATOR ! (c'est ...	UCllr3byh8wmXgCPxTm8Ocw	Piwerre	2021-01-16 16:12:19	670357	54462	832	1249	Piwerre frere michou crou among us devi
37329	zpzjex7qwrA	Lampard Sacked Within Days Rorys Misery! Chel...	UCkD-ZOxd0a9FjIExDsHsbq	The Kick Off	2021-01-03 20:13:49	428646	12060	296	1505	Premier leag Chelsea chel: 1-3 Man CityC
37345	zqyB8mnBM	Lil Wayne - Ain't Got Time (Audio)	UCO9zJy7HWhIS3ojB4Lr7YqW	Lil Wayne	2021-01-21 05:00:10	2238244	58925	2365	5539	lil wayne we weezywednes wayne carter
37383	zwfu1-24T7Q	PRADA Cup Day1 Full Race Replay PRADA Cup...	UCo15ZYQ_XDRU9LI30OPbAg	America's Cup	2021-01-15 04:07:55	317382	2008	83	192	America's C Americas C AC36 AC Presente
37418	zzlBybeSAtw	PELICANS at LAKERS FULL GAME HIGHLIGHTS Ja...	UCWJ2lWNubArHWm3FIHbfoQ	NBA	2021-01-16 05:39:05	2841917	20759	1049	2624	NBA G Leag Basket game-0022000' Laker

2108 rows x 13 columns


```

print(videos)

27      video_id      title \
48      -26um7Qf8nE      Q&A With Naisha
95      -4sfXSHSxzA      SURPRISING BRENT WITH HIS TIKTOK CRUSH!!
103     -AJD1Fc5rpQ      WE ARE HAVING A BABY! | finding out i'm pregna...
182     -AuJiwjsmWk      Do Ugly Foods Taste Worse? Taste Test
182     -JhqO2KwR5U      Schlatt gets fit
...
37300   zmzFL5b6-jc      DEVINE MON PERSONNAGE AVANT AKINATOR ! (c'est ...
37329   zpzjex7qurA      Lampard Sacked Within Days Rorys Misery | Chel...
37345   zqyv-B6mn8M      Lil Wayne - Ain't Got Time (Audio)
37383   zwFu1-24T7Q      PRADA Cup Day 1 | Full Race Replay | PRADA Cup...
37418   zziBybeSAtw      PELICANS at LAKERS | FULL GAME HIGHLIGHTS | Ja...

27      channel_id      channel_title      published_at \
48      UCvNMb0gQFzMcP5uy-pPWg      MianTwins 2021-01-21 00:05:47
95      UCpPaTKqmV-CNRmWYaDUwIA      Alexa Rivera 2021-01-16 21:40:04
103     UCvS7boAhpnuL6j-tDePvWwQ      Tess Christine 2021-01-03 21:53:48
182     UCzPcC5n9hqVC7HHPwCtKEg      Good Mythical MORE 2021-01-19 11:00:01
182     UCWZp4yjq8BuvLtiyxSs_ZBw      Big guy 2021-01-24 22:50:57
...
37300   UCILr3byh6mXgcPx_Tm9Ocw      Pwierre 2021-01-16 16:12:19
37329   UCkD-ZOixI0a9FjIExDsHsbg      The Kick Off 2021-01-03 20:13:49
37345   UC09zJy7HwIS3ojB4Ln7YqW      Lil Wayne 2021-01-21 05:00:10
37383   UC015ZY0_XDRU9LI30OPtxAg      America's Cup 2021-01-15 04:07:55
37418   UCW21lWubArWmf3FIHbfcQ      NBA 2021-01-16 05:39:05

27      view_count      likes      dislikes      comment_count \
48      872372      38626      239      621
95      6504784      262477      5779      7907
103     533084      38965      119      1650
182     1057077      22526      531      773
182     1724965      119431      325      1578
...
37300   670357      54462      832      1249
37329   420646      12060      296      1505
37345   2238244      58925      2365      5539
37383   317382      2008      83      192
37418   2841917      20759      1049      2624

27      tags \
48
95
103     gmm good mythical morning rhettandlink rhett a...
182     jschlatt big guy jschlatt highlights schlatt j...
...
37300   Pwierre frere de michou crouton among us devin...
37329   Premier league Chelsea chelsea 1-3 Man City Ch...
37345   lil wayne weezy weezy wednesday wayne carter y...
37383   America's Cup Americas Cup AC36 AC75 Presented...
37418   NBA G League Basketball game-0022000187 Lakers...

27      description \
48      Hey Guys!!! this has been the most requested v...
95      He had no idea! Thank you guys so much for wat...
103     I am so happy to tell you that I am pregnant!!!!
182     Today, we're doing a blind taste test to deter...
182     #jschlatt #schlatt #bigguy #short
...
37300   Discord Pwierre : https://discord.gg/QBduPgAA...
37329   The Kick Off watched Manchester City destroy C...
37345   Official audio for Lil Wayne "Ain't Got Time",...
37383   The opening day of the PRADA Cup in Auckland, ...
37418   PELICANS at LAKERS | FULL GAME HIGHLIGHTS | Ja...

27      comments      published_month
48      I feel like Nate and Aishas personality match ...      None
95      Thank you guys for watching and don't forget t...      None
103     Okay I needed a moment to collect my thoughts ...      None
182     "there's nothing wrong with it being bent"\nI ...      None
182     Schlatt is single handedly wiping out all the ...      None

27      comments      published_month
48      I feel like Nate and Aishas personality match ...      None
95      Thank you guys for watching and don't forget t...      None
103     Okay I needed a moment to collect my thoughts ...      None
182     "there's nothing wrong with it being bent"\nI ...      None
182     Schlatt is single handedly wiping out all the ...      None
...
37300   Mdr michou quand c'est pas ses tournage il fou...      None
37329   True Its like a fighter who Georgie But I thou...      None
37345   RIP Juice Wrld, wrote that on two cups, pour o...      None
37383   Incredible how these boats evolve in a short t...      None
37418   Montreal Harrell is going crazy with the rebou...      None

[2108 rows x 13 columns]

```

```
df['published_at'] = pd.to_datetime(df['published_at'])  
count_january_videos = videos.shape[0]  
print("count of videos published in january:", count_january_videos)
```

count of videos published in january: 2108