

AIM

-> Data Analysis of my friend youtube channel "Telugu Tech Store".

-> Getting to know Youtube API and how to obtain video data.

-> Analyzing video data and verify different common "myths" about what makes a video do well on Youtube, for example

1.Best Performing Videos based on the views.

2.Worst Performing Videos based on the views.

3.View Distribution per Video.

4.Views influencing Likes and Comments.

5.Does the Tags influence the views Count.

6.Longest Video and shortest video

7.Upload Schedule based on the Weekdays

```
In [4]: ## My Youtube API key  
api_key = "AIzaSyAIGKHoLd6p7Rkb3PQOx1S2Mwwp4mpm4yE"
```

```
In [5]: channel_id = ['UCASo6BRw8pmTTUEW-Vf8pnw',]
```

```
In [9]: from googleapiclient.discovery import build  
import pandas as pd  
from IPython.display import JSON  
import warnings  
warnings.filterwarnings('ignore')
```

```
In [10]: api_service_name = "youtube"  
api_version = "v3"  
  
# Get credentials and create an API client  
youtube = build(  
    api_service_name, api_version, developerKey=api_key)
```

In [11]: *## Get Playlist ID*

```
def get_channel_stats(youtube, channel_id):
    all_data=[]

    request= youtube.channels().list(
        part="snippet,contentDetails,statistics",
        id=', '.join(channel_id)
    )
    response= request.execute()

    for item in response['items']:
        data= {'channelName':item['snippet']['title'],
              'subscribers': item['statistics']['subscriberCount'],
              'views':item['statistics']['viewCount'],
              'totalvideos': item['statistics']['videoCount'],
              'playlistId': item['contentDetails']['relatedPlaylists']['uploads']}

        all_data.append(data)

    return(pd.DataFrame(all_data))
```

In [12]: channel_stats= get_channel_stats(youtube, channel_id)

In [13]:

channel_stats

Out[13]:

	channelName	subscribers	views	totalvideos	playlistId
0	Telugu Techstore	17100	1424853	325	UUASo6BRw8pmTTUEW-Vf8pnw

```
In [15]: ## Get Video IDs
playlist_id = "UUASo6BRw8pmTTUEW-Vf8pnw"

def get_video_ids(youtube,playlist_id):

    video_ids = []
    request = youtube.playlistItems().list(
        part="snippet,contentDetails",
        playlistId=playlist_id,
        maxResults = 50
    )
    response = request.execute()

    for item in response['items']:
        video_ids.append(item['contentDetails']['videoId'])

    next_page_token = response.get('nextPageToken')
    more_pages = True

    while more_pages:
        if next_page_token is None:
            more_pages = False
        else:
            request = youtube.playlistItems().list(
                part='contentDetails',
                playlistId = playlist_id,
                maxResults = 50,
                pageToken = next_page_token)
            response = request.execute()

            for i in range(len(response['items'])):
                video_ids.append(response['items'][i]['contentDetails']['videoId'])

            next_page_token = response.get('nextPageToken')

    return video_ids
```

```
In [16]: video_ids = get_video_ids(youtube,playlist_id)
```

```
In [17]: len(video_ids)
```

```
Out[17]: 325
```

```
In [18]: request = youtube.videos().list(
    part="snippet,contentDetails,statistics",
    id=video_ids[:5]
)
response = request.execute()

JSON(response)
```

```
Out[18]: <IPython.core.display.JSON object>
```

In [24]: `##Get Video Details`

```
def get_video_details(youtube, video_ids):
    all_videos_info= []

    for i in range(0,len(video_ids),50):
        request=youtube.videos().list(part='snippet,contentDetails,Statistics',id=

        response= request.execute()

        for video in response['items']:
            stats_to_keep= { 'snippet':['channelTitle','title','description','tags',
                                'statistics':['viewCount','likeCount','favouriteCount'],'
                                'contentDetails':['duration','definition','caption']}

            video_info={}
            video_info['video_id']=video['id']

            for k in stats_to_keep.keys():
                for v in stats_to_keep[k]:
                    try:
                        video_info[v]=video[k][v]
                    except:
                        video_info[v]=None

            all_videos_info.append(video_info)

    return pd.DataFrame(all_videos_info)
```

In [25]: `video_df = get_video_details(youtube,video_ids)`In [26]: `video_df`

3	OZYPTt6Cuew	Techstore	lock tor incognito tabs...	lock tor incognito tabs...	incognito mode telu...	02T13:58:40Z
4	gBfcJt81p8o	Telugu Techstore	How to change paytm upi pin in telugu by Pra...	How to change paytm upi pin in telugu by Pra...	[how to change upi pin in phonepe, how to chan...	2023-03- 29T12:54:40Z
...
320	WQMK8nwjiXk	Telugu Techstore	how to create vijay devarakonda whatsapp video...	Hai friends,\n I have explai...	[kinemaster editing, video call with celebrity...	2020-04- 24T15:43:13Z
321	GKllgDm71Ss	Telugu Techstore	make trending #tiktok video in telugu l	How to create trending tiktok video	[tik tok videos creation, how to	2020-04- 20T13:45:32Z

In [19]: *## get comments of the videos*

```
def get_comments(youtube, video_ids):
    all_comments = []

    for video_id in video_ids:
        try:
            request = youtube.commentThreads().list(
                part="snippet,replies",
                videoId=video_id
            )
            response = request.execute()

            comments_in_video = [comment['snippet']['topLevelComment']['snippet']
                                for comment in response['comments']]
            comments_in_video_info = {'video_id': video_id, 'comments': comments_in_video}

            all_comments.append(comments_in_video_info)

        except:
            # When error occurs - most likely because comments are disabled on
            print('Could not get comments for video ' + video_id)

    return pd.DataFrame(all_comments)
```

In [20]: comments_df= get_comments(youtube, video_ids)
comments_df

Out[20]:

	video_id	comments
0	IIMIH2TVJOY	[Gsea food app real or fake cheppava bro .]
1	w6JfcT-suFI	[Gsea food app real or fake cheppava bro]
2	J1IG_1NmMuM	[]
3	O2YPTt6c0ew	[Bro incognito endhuku bro]
4	gBfcJt81p8o	[]
...
320	WQMK8nwjiXk	[Nice bro, Nice, super ee, Super]
321	GklIgDm71Ss	[]
322	0C0msNH_UIg	[]
323	Gsu2EuC2pOQ	[keka, Nice ra]
324	BP3IRPxwavE	[Fist video 😊, Anna Kine master app link pettav...

325 rows × 2 columns

Data Preprocessing

```
In [17]: video_df.isnull().sum()
```

```
Out[17]: video_id          0
channelTitle        0
title               0
description          0
tags                9
publishedAt         0
viewCount           0
likeCount           1
favouriteCount      325
commentCount        0
duration            0
definition          0
caption             0
dtype: int64
```

```
In [18]: video_df.dtypes
```

```
Out[18]: video_id          object
channelTitle        object
title               object
description          object
tags                object
publishedAt         object
viewCount           object
likeCount           object
favouriteCount      object
commentCount        object
duration            object
definition          object
caption             object
dtype: object
```

```
In [19]: video_df.columns
```

```
Out[19]: Index(['video_id', 'channelTitle', 'title', 'description', 'tags',
               'publishedAt', 'viewCount', 'likeCount', 'favouriteCount',
               'commentCount', 'duration', 'definition', 'caption'],
              dtype='object')
```

```
In [28]: video_df.columns=video_df.columns.str.lower()
```

```
In [29]: video_df.columns
```

```
Out[29]: Index(['video_id', 'channeltitle', 'title', 'description', 'tags',
               'publishedat', 'viewcount', 'likecount', 'favouritecount',
               'commentcount', 'duration', 'definition', 'caption'],
              dtype='object')
```

```
In [30]: #change data type of 'viewcount','likecount','commentcount'
numeric_cols= ['viewcount','likecount','commentcount']
video_df[numeric_cols]= video_df[numeric_cols].apply(pd.to_numeric,errors='coer
```

```
In [23]: video_df.dtypes
```

```
Out[23]: video_id          object
channeltitle      object
title             object
description        object
tags              object
publishedat       object
viewcount         float64
likecount         float64
favouritecount    object
commentcount      float64
duration          object
definition        object
caption           object
dtype: object
```

```
In [24]: ##Published date into date
from dateutil import parser
video_df['publishedat'] = video_df['publishedat'].apply(lambda x: parser.parse
video_df['publishDayName'] = video_df['publishedat'].apply(lambda x: x.strftime
```

In [34]: video_df.head()

Out[34]:

	video_id	channeltitle	title	description	tags	publishedat	viewcount	like
0	IIMIH2TVJOY	Telugu Techstore	Is fast charging harmful to your mobile batter...	Is fast charging harmful to your mobile batter...	[is fast charging good or bad in telugu, mobil...	2023-04-21 13:00:25+00:00	42.0	
1	w6JfcT-suFI	Telugu Techstore	These 5 things will damage your mobile battery...	These 5 things will damage your mobile battery...	[these things will damage your mobile battery ...	2023-04-16 12:54:13+00:00	85.0	
2	J1IG_1NmMuM	Telugu Techstore	How to get refund for a wrong upi transaction ...	How to get refund for a wrong upi transaction ...	[wrong upi transaction in telugu, wrong upi tr...	2023-04-06 12:34:43+00:00	229.0	
3	O2YPTt6c0ew	Telugu Techstore	How to set fingerprint lock for incognito tabs...	How to set fingerprint lock for incognito tabs...	[incognito mode in chrome, incognito mode telu...	2023-04-02 13:58:40+00:00	77.0	
4	gBfcJt81p8o	Telugu Techstore	How to change paytm upi pin in telugu by Pra...	How to change paytm upi pin in telugu by Pra...	[how to change upi pin in phonepe, how to chan...	2023-03-29 12:54:40+00:00	40.0	

In [1]: !pip install isodate

Looking in indexes: <https://pypi.org/simple>, (<https://pypi.org/simple>,) <http://us-python.pkg.dev/colab-wheels/public/simple/> (<https://us-python.pkg.dev/colab-wheels/public/simple/>)

Collecting isodate

Downloading isodate-0.6.1-py2.py3-none-any.whl (41 kB)

41.7/41.7 kB 4.8 MB/s eta 0:00:00

Requirement already satisfied: six in /usr/local/lib/python3.9/dist-packages (from isodate) (1.16.0)

Installing collected packages: isodate

Successfully installed isodate-0.6.1

In [27]: *#Duration into Duration in seconds*
 from isodate import parse_duration

```
video_df['durationSecs'] = video_df['duration'].apply(lambda x: parse_duration(x))
video_df['durationSecs'] = video_df['durationSecs'].astype('timedelta64[s]')
```



```
In [28]: video_df[['duration', 'durationSecs']]
```

```
Out[28]:
```

	duration	durationSecs
0	PT4M54S	294.0
1	PT5M50S	350.0
2	PT5M45S	345.0
3	PT3M29S	209.0
4	PT2M26S	146.0
...
320	PT9M18S	558.0
321	PT8M22S	502.0
322	PT6M19S	379.0
323	PT4M12S	252.0
324	PT12M35S	755.0

325 rows × 2 columns

```
In [29]: #Adding Tag count
video_df['tagcount'] = video_df['tags'].apply(lambda x: 0 if x is None else len
```

```
In [30]: video_df[['tags', 'tagcount']]
```

```
Out[30]:
```

	tags	tagcount
0	[is fast charging good or bad in telugu, mobil...	9
1	[these things will damage your mobile battery ...	10
2	[wrong upi transaction in telugu, wrong upi tr...	9
3	[incognito mode in chrome, incognito mode telu...	11
4	[how to change upi pin in phonepe, how to chan...	12
...
320	[kinemaster editing, video call with celebrity...	11
321	[tik tok videos creation, how to make videos o...	13
322	[kinemaster editing, whatsapp status video edi...	11
323	[kinemaster tutorials telugu, fire effect vide...	11
324	[kinemaster editing, kinemaster lyrical video ...	12

325 rows × 2 columns

In [32]: video_df

Out[32]:

	video_id	channeltitle	title	description	tags	publishedat	viewcount
0	IIMIH2TVJOY	Telugu Techstore	Is fast charging harmful to your mobile batter...	Is fast charging harmful to your mobile batter...	[is fast charging good or bad in telugu, mobil...	2023-04-21 13:00:25+00:00	42.0
1	w6JfcT-suFI	Telugu Techstore	These 5 things will damage your mobile battery...	These 5 things will damage your mobile battery...	[these things will damage your mobile battery ...	2023-04-16 12:54:13+00:00	85.0
2	J1IG_1NmMuM	Telugu Techstore	How to get refund for a wrong upi transaction ...	How to get refund for a wrong upi transaction ...	[wrong upi transaction in telugu, wrong upi tr...	2023-04-06 12:34:43+00:00	229.0
3	O2YPTt6c0ew	Telugu Techstore	How to set fingerprint lock for incognito tabs...	How to set fingerprint lock for incognito tabs...	[incognito mode in chrome, incognito mode telu...	2023-04-02 13:58:40+00:00	77.0
4	gBfcJt81p8o	Telugu Techstore	How to change paytm upi pin in telugu by Pra...	How to change paytm upi pin in telugu by Pra...	[how to change upi pin in phonepe, how to chan...	2023-03-29 12:54:40+00:00	40.0
...
320	WQMK8nwjiXk	Telugu Techstore	how to create vijay devarakonda whatsapp video...	Hai friends,\n I have explai...	[kinemaster editing, video call with celebrity...	2020-04-24 15:43:13+00:00	606.0
321	GklIgDm71Ss	Telugu Techstore	make trending #tiktok video in telugu How to...	How to create trending tiktok video in telugu....	[tik tok videos creation, how to make videos o...	2020-04-20 13:45:32+00:00	201.0
322	0C0msNH_UIg	Telugu Techstore	make love whatsapp status video in telugu 2020...	How to make awesome whatsapp status video in t...	[kinemaster editing, whatsapp status video edi...	2020-04-16 07:48:44+00:00	255.0
323	Gsu2EuC2pOQ	Telugu Techstore	make fire effect whatsapp status video in telugu	How to create fire effect whatsapp status vide...	[kinemaster tutorials telugu, fire effect vide...	2020-04-11 07:49:59+00:00	407.0
324	BP3IRPxwavE	Telugu Techstore	create lyrical video in telugu 2020 make lyr...	How to create lyrical video in kinemaster in t...	[kinemaster editing, kinemaster lyrical video ...	2020-04-07 16:40:45+00:00	541.0

325 rows × 16 columns

In []:

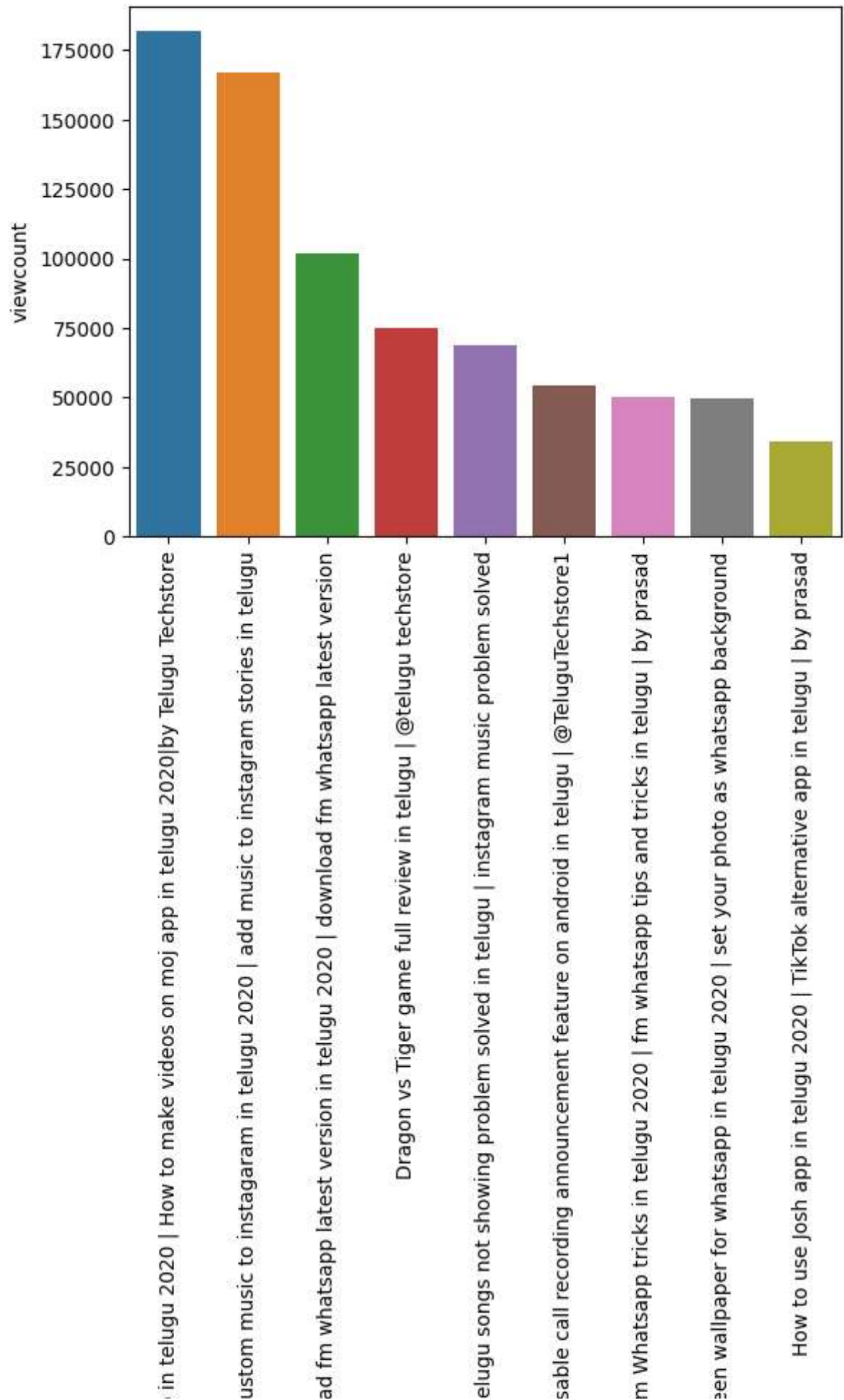


Exploratory Data Analysis (EDA)

```
In [101]: ## import all Necessary Libraries  
import seaborn as sns  
import matplotlib.pyplot as plt
```

Best Performing Videos

```
In [39]: ax=sns.barplot(x="title",y="viewcount",data=video_df.sort_values('viewcount',a
plot = ax.set_xticklabels(ax.get_xticklabels(), rotation=90)
plt.show()
```

How to use moj app

How to add c

How to downlo

instagram story t

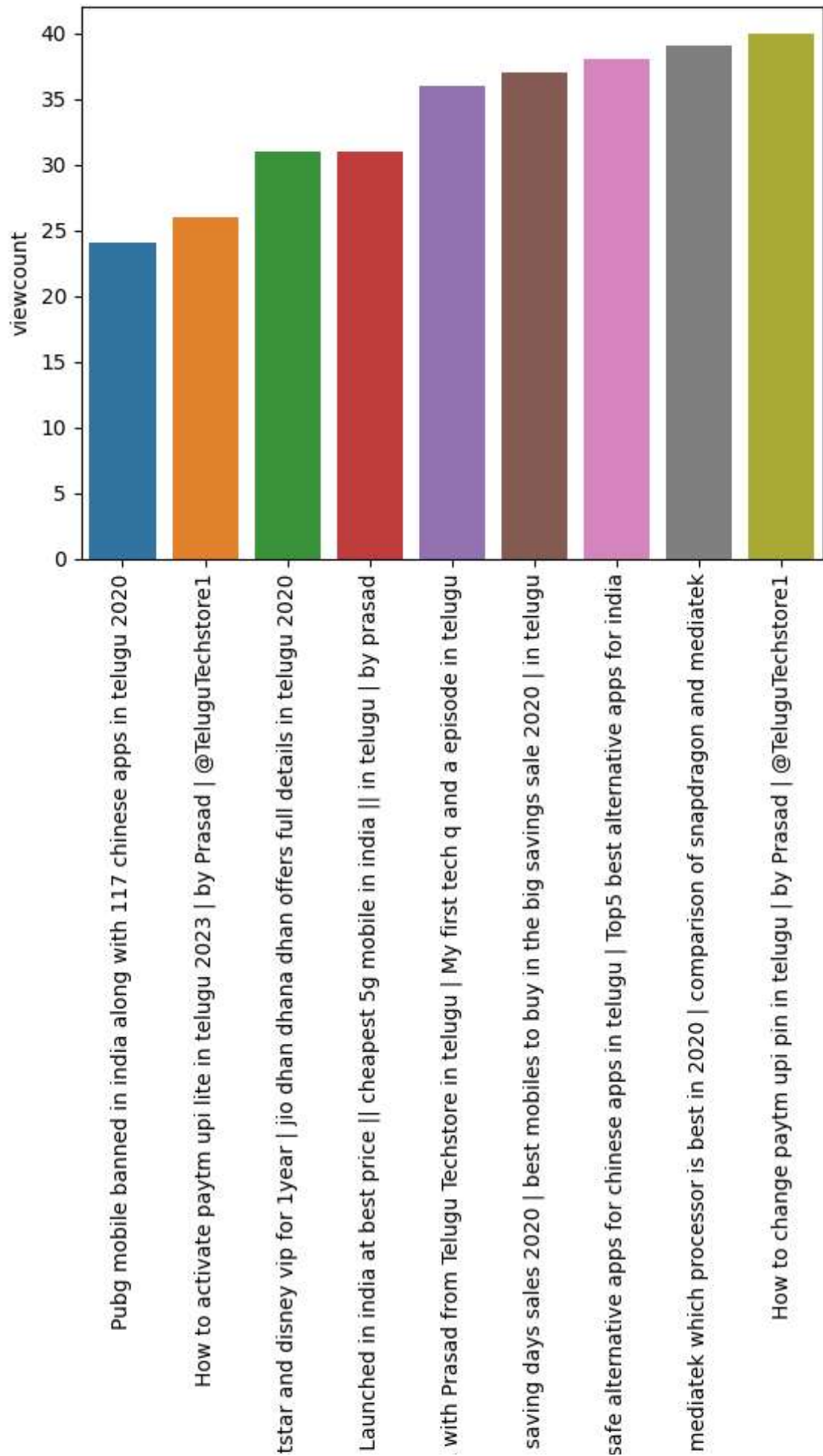
How to dii

Top 15 Fi

How to set home scre

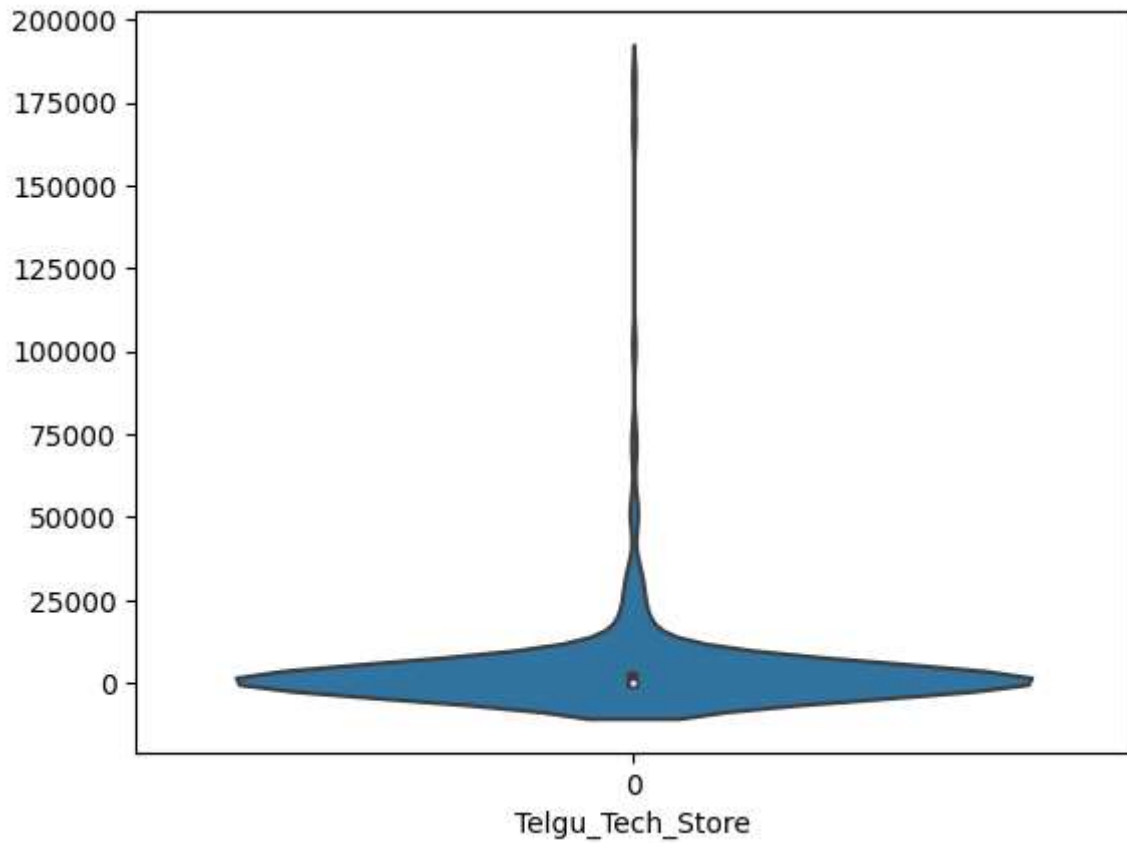
title


```
In [40]: ax=sns.barplot(x="title",y="viewcount",data=video_df.sort_values('viewcount',a
plot = ax.set_xticklabels(ax.get_xticklabels(), rotation=90)
plt.show()
```

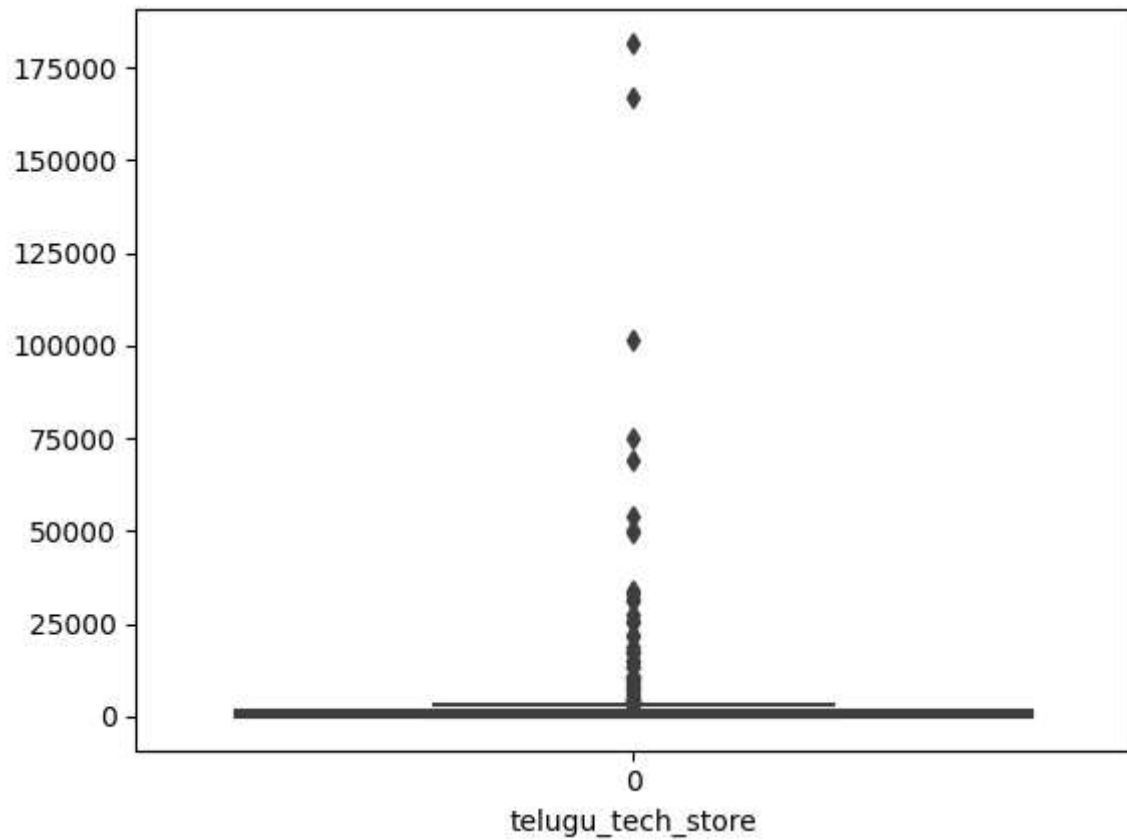



```
In [51]: ax=sns.violinplot(video_df['viewcount'],  
ax.set_xlabel("Telgu_Tech_Store"))
```

```
Out[51]: Text(0.5, 0, 'Telgu_Tech_Store')
```

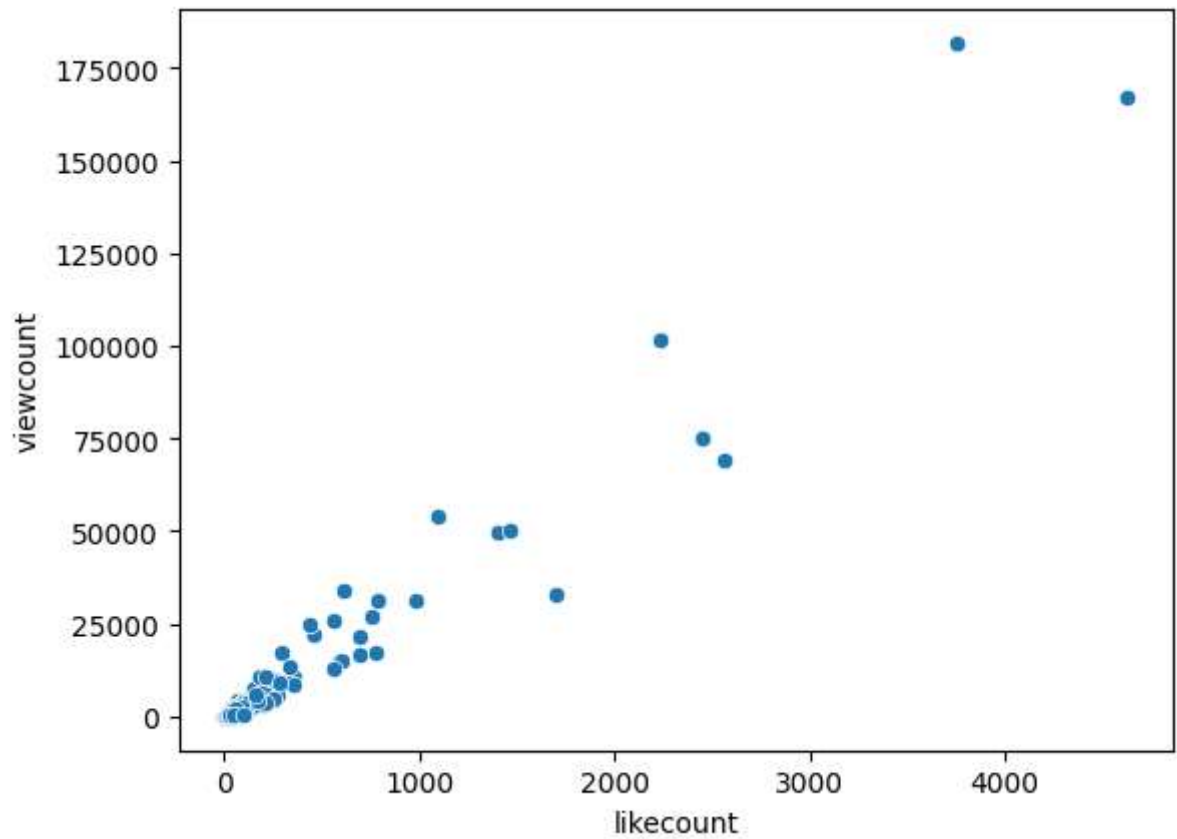


```
In [61]: ax=sns.boxplot(video_df['viewcount'])  
ax.set_xlabel("telugu_tech_store")  
plt.show()
```



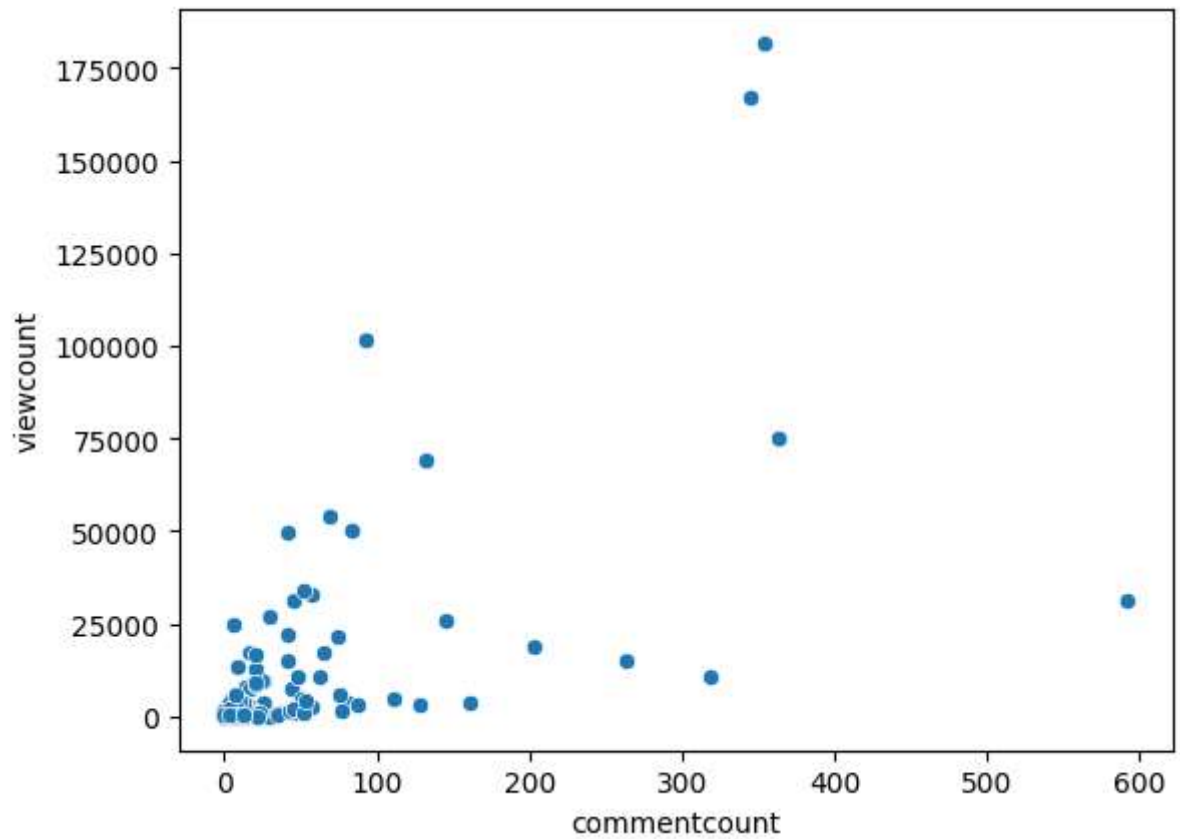
Views vs Likes

```
In [65]: sns.scatterplot(x="likecount",y="viewcount",data=video_df)  
plt.show()
```



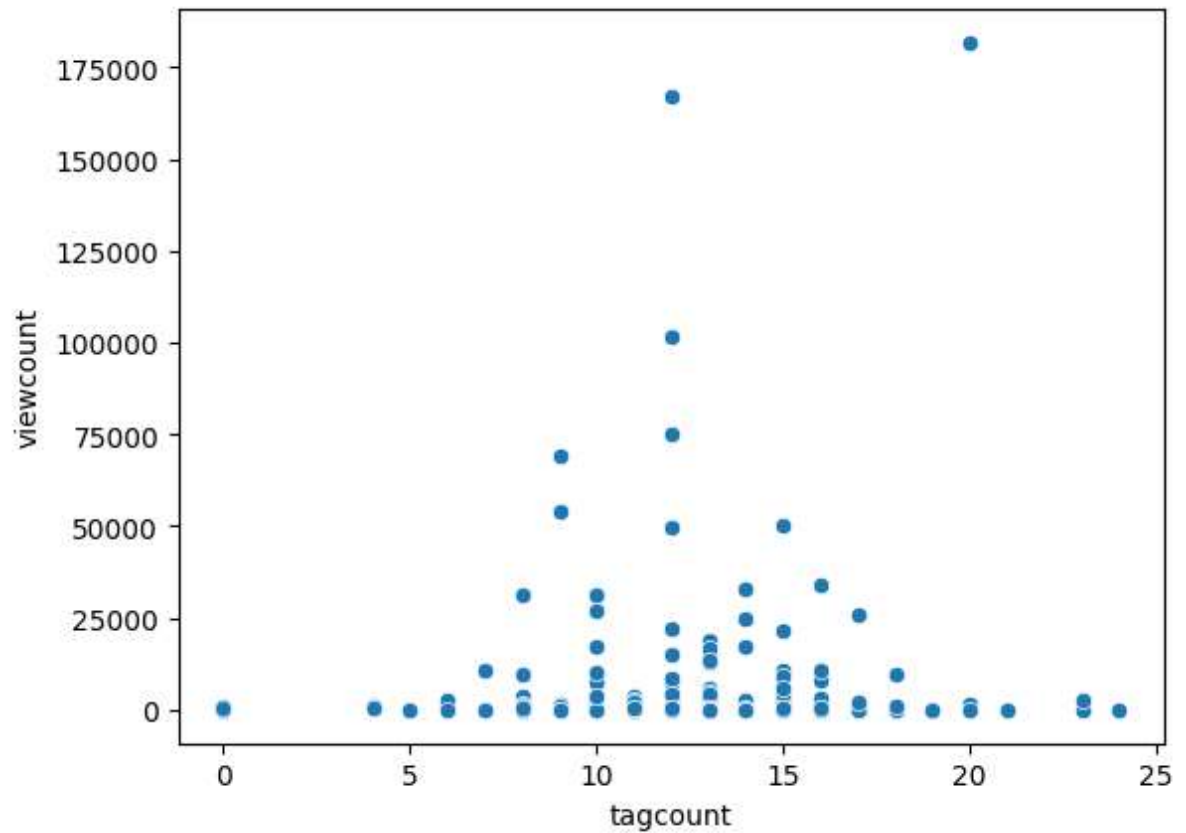
Views vs Comments

```
In [67]: sns.scatterplot(x="commentcount",y="viewcount",data=video_df)  
plt.show()
```



Views vs No of Tags

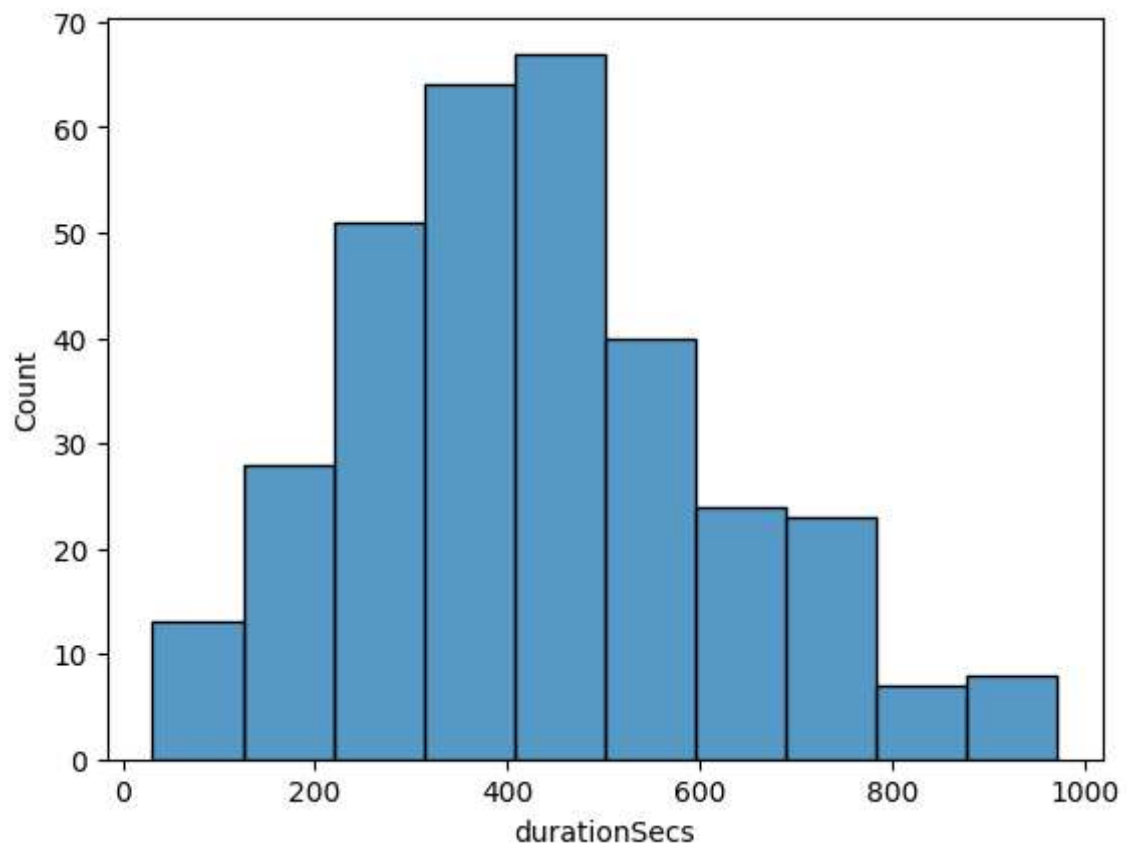
```
In [102]: sns.scatterplot(x="tagcount",y="viewcount",data=video_df)  
plt.show()
```



Video Duration


```
In [69]: sns.histplot(x='durationSecs',data=video_df,bins=10)
```

```
Out[69]: <Axes: xlabel='durationSecs', ylabel='Count'>
```

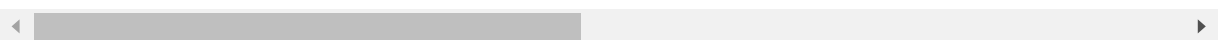


Longest Video

```
In [71]: video_df[video_df['durationSecs']==video_df['durationSecs'].max()]
```

```
Out[71]:
```

	video_id	channeltitle	title	description	tags	publishedat	viewcount	likeco
274	GMqlchDjlsE	Telugu Techstore	How to make lyrical videos on alight motion in...	How to make lyrical videos on alight motion in...	[how to create/make trending lyrical videos in...	2020-07-28 08:29:16+00:00	115.0	1



Shortest Video

```
In [73]: video_df[video_df['durationSecs']==video_df['durationSecs'].min()]
```

```
Out[73]:
```

	video_id	channeltitle	title	description	tags	publishedat	viewcount	li
318	zFnQwnfmENQ	Telugu Techstore	Trending phone charging connected to hand gree...	Trending phone charging connected to hand gree...	[green screen video for kinemaster, trending g...	2020-05-01 13:04:49+00:00	6190.0	

Upload Schedule

```
In [84]: df=video_df.groupby('publishDayName')['video_id'].count().reset_index().sort_v
```

```
In [91]: df
```

```
Out[91]:
```

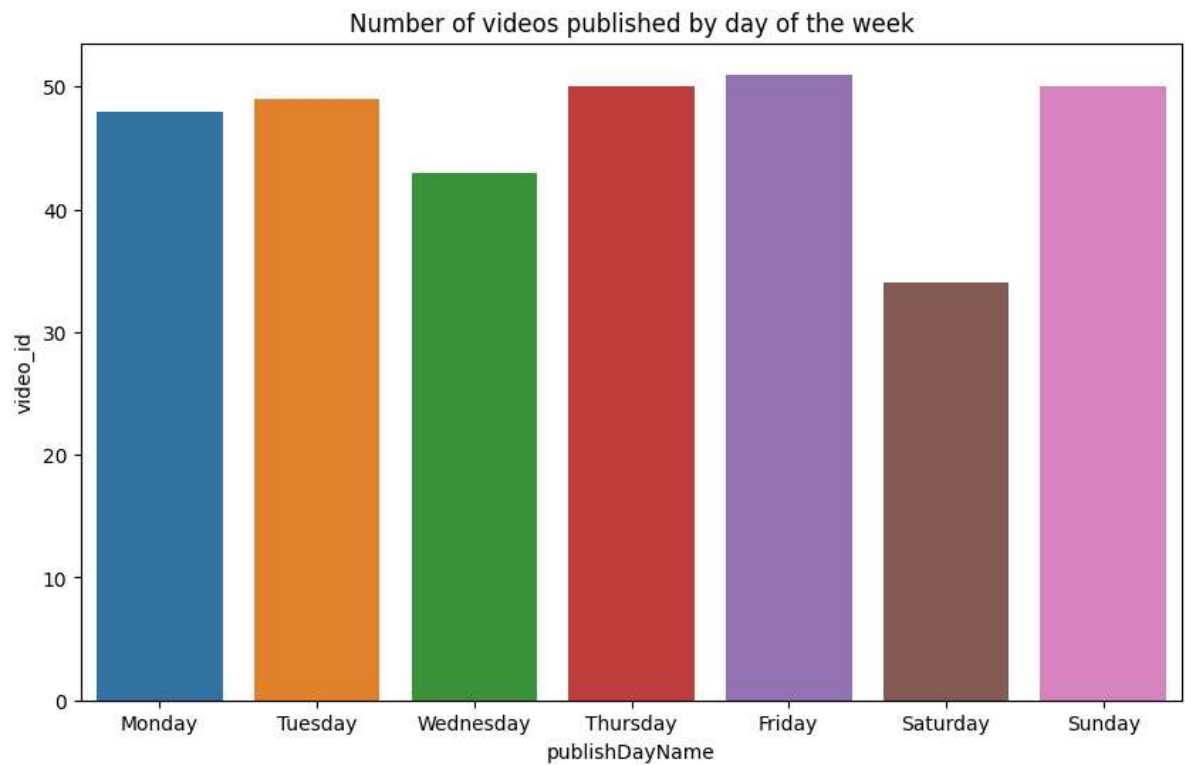
	publishDayName	video_id
0	Friday	51
1	Monday	48
2	Saturday	34
3	Sunday	50
4	Thursday	50
5	Tuesday	49
6	Wednesday	43

```
In [92]: cats=['Monday','Tuesday','Wednesday','Thursday','Friday','Saturday','Sunday']
df['publishDayName']=pd.Categorical(df['publishDayName'],categories=cats, orde
df=df.sort_values('publishDayName')
df
```

```
Out[92]:
```

	publishDayName	video_id
1	Monday	48
5	Tuesday	49
6	Wednesday	43
4	Thursday	50
0	Friday	51
2	Saturday	34
3	Sunday	50

```
In [103]: plt.figure(figsize=(10,6))  
sns.barplot(x="publishDayName",y="video_id",data=df)  
plt.title('Number of videos published by day of the week')  
plt.show()
```



```
In [ ]:
```