

Final Project Data Extraction File for Social Media Analytics

Task 1 - Let's gather our own data (e.g. minimum 1000 posts or records) related to a hot topic in Somali politics or sociology from a Social Media Outlet by using its API. Put into one CSV file for all the collected Comments that are in Somali language.

Fetching Data using Youtube's API

In [1]:

```
1 from googleapiclient.discovery import build
2 from googleapiclient.errors import HttpError
3 import pandas as pd
4 # Replace 'YOUR_API_KEY_HERE' with your actual API Key
5 # Replace 'VIDEO_ID' with the actual video ID
6 API_KEY = 'AIzaSyBpxUg3oEgIA8Pv9E0hf5uCqH5FiRGeFBY'
7 api_service_name = "youtube"
8 api_version = "v3"
```

First Time Data Fetching

In [29]:

```
1 # actual video link 'https://www.youtube.com/watch?v=CDF61GYYLps'
2
3 youtube = build('youtube', 'v3', developerKey=API_KEY)
4
5 def get_video_comments(video_id):
6     comments = []
7     response = youtube.commentThreads().list(
8         part='snippet,replies',
9         videoId=video_id,
10        maxResults=2150, # Adjust based on your needs
11        textFormat='plainText'
12    ).execute()
13
14    while response:
15        for item in response['items']:
16            comment = item['snippet']['topLevelComment']['snippet']
17            comments.append({
18                'author': comment['authorDisplayName'],
19                'date': comment['publishedAt'],
20                'Comment': comment['textDisplay'],
21                'like_count': comment['likeCount'],
22                'reply_count': item['snippet']['totalReplyCount'],
23                'comment_length': len(comment['textDisplay'])
24            })
25
26        # If there are replies, fetch them as well
27        if item['snippet']['totalReplyCount'] > 0:
28            for reply in item['replies']['comments']:
29                reply_comment = reply['snippet']
30                comments.append({
31                    'author': reply_comment['authorDisplayName'],
32                    'date': reply_comment['publishedAt'],
33                    'Comment': reply_comment['textDisplay'],
34                    'like_count': reply_comment['likeCount'],
35                    'reply_count': 0, # Replies do not have further
36                    'comment_length': len(reply_comment['textDisplay'])
37                })
38
39        # Check if there are more pages of comments
40        if 'nextPageToken' in response:
41            response = youtube.commentThreads().list(
42                part='snippet,replies',
43                videoId=video_id,
44                pageToken=response['nextPageToken'],
45                maxResults=100, # Adjust based on your needs
46                textFormat='plainText'
47            ).execute()
48        else:
49            break
50
51    return comments
52
53 # Example usage
54 video_id = 'ItNAou5B7UQ'
55 comments = get_video_comments(video_id)
56 for comment in comments:
57     print(comment)
58
```

```
{'author': '@maryamadirie8722', 'date': '2024-04-19T09:43:45Z', 'Comment': 'Cadaab ka rabi bilaash uma san mujrimiin Rabiyaw umada kaqabo', 'like_count': 0, 'reply_count': 0, 'comment_length': 61}
{'author': '@hagiohaji7937', 'date': '2024-04-19T03:35:24Z', 'Comment': 'Beenaale gacan ku dhiigle tuug, hala dilo.', 'like_count': 0, 'reply_count': 0, 'comment_length': 42}
{'author': '@fadumawarsamemumin4232', 'date': '2024-04-18T12:17:27Z', 'Comment': 'All Shaba a le lagu Haye Polis dada cun ah Qamriga iyo qaramark Cadi utahay illahow adig nomaqan', 'like_count': 0, 'reply_count': 0, 'comment_length': 96}
{'author': '@user-we3ty2ct8f', 'date': '2024-04-17T22:02:09Z', 'Comment': 'Hadu I yaqaano ha ii abtiriyo kulahaa kk ma hoyadaaba dhashay marku kuu abtirinaayo qof kastoo ku yaqanaa ma inu kuu abtirsha', 'like_count': 0, 'reply_count': 0, 'comment_length': 125}
{'author': '@abdiali4991', 'date': '2024-04-17T08:07:17Z', 'Comment': 'Sarkaal kan bistolda loo dibtay waa ina asagana talabo lagadho hubka dawal ada ayu suga kuu haysanaya shaki ayaa jira', 'like_count': 0, 'reply_count': 0, 'comment_length': 117}
{'author': '@abdiali4991', 'date': '2024-04-17T08:05:05Z', 'Comment': 'Ni
```

In [30]:

```
1 # Convert the List of comments to a pandas DataFrame
2 df_comments = pd.DataFrame(comments)
3
4 # Save the DataFrame to a CSV file
5 csv_file_name = "kiiska_qoslaaye1.csv"
6 df_comments.to_csv(csv_file_name, index=False, encoding='utf-8-sig')
7 print(f"Comments have been saved to {csv_file_name}")
```

Comments have been saved to kiiska_qoslaaye1.csv

```
In [32]: 1 df1 = pd.read_csv('kiiska_qoslaaye1.csv')
2 df1.head()
```

Out[32]:

	author	date	Comment	like_count	reply_count	comment_len
0	@maryamadirie8722	2024-04-19T09:43:45Z	Cadaab ka rabi bilaash uma san mujrimiin Rabiy...	0	0	
1	@hagiohaji7937	2024-04-19T03:35:24Z	Beenaale gacan ku dhiigle tuug, hala dilo.	0	0	
2	@fadumawarsamemumin4232	2024-04-18T12:17:27Z	All Shaba a le lagu Haye Polis dada cun ah Qam...	0	0	
3	@user-we3ty2ct8f	2024-04-17T22:02:09Z	Hadu I yaqaano ha ii abtiryo kulahaa kk ma ho...	0	0	
4	@abdiali4991	2024-04-17T08:07:17Z	Sarkaal kan bistoolda loo dibtay waa ina asaga...	0	0	



```
In [33]: 1 records = len(df1)
2 records
```

Out[33]: 79

Second Time Data Fetching

In [34]:

```
1 # actual video link 'https://www.youtube.com/watch?v=JOC6hFkWlmy'
2
3 youtube = build('youtube', 'v3', developerKey=API_KEY)
4
5 def get_video_comments(video_id):
6     comments = []
7     response = youtube.commentThreads().list(
8         part='snippet,replies',
9         videoId=video_id,
10        maxResults=2150, # Adjust based on your needs
11        textFormat='plainText'
12    ).execute()
13
14    while response:
15        for item in response['items']:
16            comment = item['snippet']['topLevelComment']['snippet']
17            comments.append({
18                'author': comment['authorDisplayName'],
19                'date': comment['publishedAt'],
20                'Comment': comment['textDisplay'],
21                'like_count': comment['likeCount'],
22                'reply_count': item['snippet']['totalReplyCount'],
23                'comment_length': len(comment['textDisplay'])
24            })
25
26        # If there are replies, fetch them as well
27        if item['snippet']['totalReplyCount'] > 0:
28            for reply in item['replies']['comments']:
29                reply_comment = reply['snippet']
30                comments.append({
31                    'author': reply_comment['authorDisplayName'],
32                    'date': reply_comment['publishedAt'],
33                    'Comment': reply_comment['textDisplay'],
34                    'like_count': reply_comment['likeCount'],
35                    'reply_count': 0, # Replies do not have further
36                    'comment_length': len(reply_comment['textDisplay'])
37                })
38
39        # Check if there are more pages of comments
40        if 'nextPageToken' in response:
41            response = youtube.commentThreads().list(
42                part='snippet,replies',
43                videoId=video_id,
44                pageToken=response['nextPageToken'],
45                maxResults=100, # Adjust based on your needs
46                textFormat='plainText'
47            ).execute()
48        else:
49            break
50
51    return comments
52
53 # Example usage
54 video_id = 'JOC6hFkWlmy'
55 comments = get_video_comments(video_id)
56 for comment in comments:
57     print(comment)
58
```

```
{'author': '@adarosmanahmed5635', 'date': '2024-03-04T22:12:59Z', 'Comment': 'Naga daaya', 'like_count': 1, 'reply_count': 0, 'comment_length': 10}
{'author': '@user-qe8em2eo7k', 'date': '2024-03-02T12:32:45Z', 'Comment': 'Delaa😊', 'like_count': 1, 'reply_count': 0, 'comment_length': 6}
{'author': '@user-uv8zn3yc9n', 'date': '2024-03-01T15:52:44Z', 'Comment': '😂', 'like_count': 0, 'reply_count': 0, 'comment_length': 1}
{'author': '@Deeqasalaad-zt6pp', 'date': '2024-03-01T11:16:32Z', 'Comment': 'Illahay hakudilo balaqabayoho', 'like_count': 0, 'reply_count': 0, 'comment_length': 29}
{'author': '@naimaabduallah7994', 'date': '2024-03-01T08:36:59Z', 'Comment': 'What was the reason he killed her?', 'like_count': 0, 'reply_count': 0, 'comment_length': 34}
{'author': '@user-cn9zg7ps1c', 'date': '2024-02-27T10:54:19Z', 'Comment': 'Kakgu nikani♥♥♥♥♥', 'like_count': 0, 'reply_count': 0, 'comment_length': 18}
{'author': '@user-fj4lz8bf4u', 'date': '2024-02-27T07:58:21Z', 'Comment': 'Cadalad majirto wlhi mahkamada somalia', 'like_count': 0, 'reply_count': 0, 'comment_length': 38}
```

In [35]:

```
1 # Convert the List of comments to a pandas DataFrame
2 df_comments = pd.DataFrame(comments)
3
4 # Save the DataFrame to a CSV file
5 csv_file_name = "kiiska_qoslaaye2.csv"
6 df_comments.to_csv(csv_file_name, index=False, encoding='utf-8-sig')
7 print(f"Comments have been saved to {csv_file_name}")
```

Comments have been saved to kiiska_qoslaaye2.csv

In [41]:

```
1 df2 = pd.read_csv('kiiska_qoslaaye2.csv')
2 df2.head()
```

Out[41]:

	author	date	Comment	like_count	reply_count	comment_length
0	@adarosmanahmed5635	2024-03-04T22:12:59Z	Naga daaya	1	0	10
1	@user-qe8em2eo7k	2024-03-02T12:32:45Z	Delaa😊	1	0	6
2	@user-uv8zn3yc9n	2024-03-01T15:52:44Z	😂	0	0	1
3	@Deeqasalaad-zt6pp	2024-03-01T11:16:32Z	Illahay hakudilo balaqabayoho	0	0	29
4	@naimaabduallah7994	2024-03-01T08:36:59Z	What was the reason he killed her?	0	0	34

In [42]:

```
1 len(df2)
```

Out[42]:

220

Third Time Data Fetching

In [38]:

```
1 # actual video link 'https://www.youtube.com/watch?v=v0plj00s_W4'
2
3 youtube = build('youtube', 'v3', developerKey=API_KEY)
4
5 def get_video_comments(video_id):
6     comments = []
7     response = youtube.commentThreads().list(
8         part='snippet,replies',
9         videoId=video_id,
10        maxResults=2150, # Adjust based on your needs
11        textFormat='plainText'
12    ).execute()
13
14    while response:
15        for item in response['items']:
16            comment = item['snippet']['topLevelComment']['snippet']
17            comments.append({
18                'author': comment['authorDisplayName'],
19                'date': comment['publishedAt'],
20                'Comment': comment['textDisplay'],
21                'like_count': comment['likeCount'],
22                'reply_count': item['snippet']['totalReplyCount'],
23                'comment_length': len(comment['textDisplay'])
24            })
25
26        # If there are replies, fetch them as well
27        if item['snippet']['totalReplyCount'] > 0:
28            for reply in item['replies']['comments']:
29                reply_comment = reply['snippet']
30                comments.append({
31                    'author': reply_comment['authorDisplayName'],
32                    'date': reply_comment['publishedAt'],
33                    'Comment': reply_comment['textDisplay'],
34                    'like_count': reply_comment['likeCount'],
35                    'reply_count': 0, # Replies do not have further
36                    'comment_length': len(reply_comment['textDisplay'])
37                })
38
39        # Check if there are more pages of comments
40        if 'nextPageToken' in response:
41            response = youtube.commentThreads().list(
42                part='snippet,replies',
43                videoId=video_id,
44                pageToken=response['nextPageToken'],
45                maxResults=100, # Adjust based on your needs
46                textFormat='plainText'
47            ).execute()
48        else:
49            break
50
51    return comments
52
53 # Example usage
54 video_id = '15ywlmz-zN9Y'
55 comments = get_video_comments(video_id)
56 for comment in comments:
57     print(comment)
58
```

```

{'author': '@Gmaaad', 'date': '2024-04-19T13:02:01Z', 'Comment': 'Ceeleks dad kale ayaa lagu dilay ee mala baaro waxaa laga yaaba dad kale inee ku jiraan', 'like_count': 0, 'reply_count': 0, 'comment_length': 87}
{'author': '@xuseenibrahim4070', 'date': '2024-04-19T02:43:21Z', 'Comment': 'Australia AU', 'like_count': 0, 'reply_count': 0, 'comment_length': 12}
{'author': '@faziaahmed9537', 'date': '2024-04-18T17:42:42Z', 'Comment': 'Allah ayaa idinkashifay khaatiliinyahay 😞', 'like_count': 0, 'reply_count': 0, 'comment_length': 40}
{'author': '@faziaahmed9537', 'date': '2024-04-18T17:21:24Z', 'Comment': 'Kan afka weeyn muxuu ku qoslaa? Axmaqa xun 😞', 'like_count': 0, 'reply_count': 0, 'comment_length': 43}
{'author': '@arooryo7463', 'date': '2024-04-18T13:22:53Z', 'Comment': 'War bixinadaan Nisa maba aamino ani. Waxey camera da hor istaajiyaa qof si deey usoo jukeynaayeen xataa "dameeraan ahay" qiranaayo', 'like_count': 0, 'reply_count': 0, 'comment_length': 130}
{'author': '@user-ql8hg7cu5e', 'date': '2024-04-18T10:28:06Z', 'Comment': 'Walle dowlad cadi mahan\nXogti odan wala kashifay', 'like_count': 0, 'reply_count': 0, 'comment_length': 48}

```

In [39]:

```

1 # Convert the List of comments to a pandas DataFrame
2 df_comments = pd.DataFrame(comments)
3
4 # Save the DataFrame to a CSV file
5 csv_file_name = "kiiska_qoslaaye_3.csv"
6 df_comments.to_csv(csv_file_name, index=False, encoding='utf-8-sig')
7 print(f"Comments have been saved to {csv_file_name}")

```

Comments have been saved to kiiska_qoslaaye_3.csv

In [43]:

```

1 df3 = pd.read_csv('kiiska_qoslaaye_3.csv')
2 df3.head()

```

Out[43]:

	author	date	Comment	like_count	reply_count	comment_length
0	@Gmaaad	2024-04-19T13:02:01Z	Ceeleks dad kale ayaa lagu dilay ee mala baaro...	0	0	87
1	@xuseenibrahim4070	2024-04-19T02:43:21Z	Australia AU	0	0	12
2	@faziaahmed9537	2024-04-18T17:42:42Z	Allah ayaa idinkashifay khaatiliinyahay 😞	0	0	40
3	@faziaahmed9537	2024-04-18T17:21:24Z	Kan afka weeyn muxuu ku qoslaa? Axmaqa xun 😞	0	0	43
4	@arooryo7463	2024-04-18T13:22:53Z	War bixinadaan Nisa maba aamino ani. Waxey cam...	0	0	130

In [44]: 1 len(df3)

Out[44]: 1100

Fourth Time Data Fetching

In [45]:

```
1 # actual video link 'https://www.youtube.com/watch?v=9XuFdUVSaY0'
2
3 youtube = build('youtube', 'v3', developerKey=API_KEY)
4
5 def get_video_comments(video_id):
6     comments = []
7     response = youtube.commentThreads().list(
8         part='snippet,replies',
9         videoId=video_id,
10        maxResults=2150, # Adjust based on your needs
11        textFormat='plainText'
12    ).execute()
13
14    while response:
15        for item in response['items']:
16            comment = item['snippet']['topLevelComment']['snippet']
17            comments.append({
18                'author': comment['authorDisplayName'],
19                'date': comment['publishedAt'],
20                'Comment': comment['textDisplay'],
21                'like_count': comment['likeCount'],
22                'reply_count': item['snippet']['totalReplyCount'],
23                'comment_length': len(comment['textDisplay'])
24            })
25
26        # If there are replies, fetch them as well
27        if item['snippet']['totalReplyCount'] > 0:
28            for reply in item['replies']['comments']:
29                reply_comment = reply['snippet']
30                comments.append({
31                    'author': reply_comment['authorDisplayName'],
32                    'date': reply_comment['publishedAt'],
33                    'Comment': reply_comment['textDisplay'],
34                    'like_count': reply_comment['likeCount'],
35                    'reply_count': 0, # Replies do not have further
36                    'comment_length': len(reply_comment['textDisplay'])
37                })
38
39        # Check if there are more pages of comments
40        if 'nextPageToken' in response:
41            response = youtube.commentThreads().list(
42                part='snippet,replies',
43                videoId=video_id,
44                pageToken=response['nextPageToken'],
45                maxResults=100, # Adjust based on your needs
46                textFormat='plainText'
47            ).execute()
48        else:
49            break
50
51    return comments
52
53 # Example usage
54 video_id = 'sNXmmowt6Ds'
55 comments = get_video_comments(video_id)
56 for comment in comments:
57     print(comment)
58
```

```

{'author': '@fadumonour9815', 'date': '2024-04-19T07:13:16Z', 'Comment': 'Eeda waxaa iskaleh Dawladda Somalia, oo dad aan u qalma abniga u qortey oo qabiil sal ku leh. Waa qaran Dumay oo qabiil iyo mooryaan ka taliso. A abahay 1960 buu police noqdey. Dhacdaan aduunka lagama maqal oo nin police ah ku kaco. Dawladda maxay ka qanbneysaa Dadka ay police ka u qorte y ama u qoran doonto .😢😢😢😢😢', 'like_count': 1, 'reply_count': 0, 'comment_length': 318}
{'author': '@fadumonour9815', 'date': '2024-04-19T06:58:53Z', 'Comment': '😢😢😢', 'like_count': 1, 'reply_count': 0, 'comment_length': 3}
{'author': '@muuseaweys', 'date': '2024-04-18T18:36:41Z', 'Comment': 'Allow dowlaadeena noo sug oo noo xooji amni iyo cadaalad noogu xooji', 'like_count': 0, 'reply_count': 0, 'comment_length': 67}
{'author': '@muuseaweys', 'date': '2024-04-18T18:35:16Z', 'Comment': 'ماشاء الله تبارك الله', 'like_count': 0, 'reply_count': 0, 'comment_length': 20}
{'author': '@saidkhalif7456', 'date': '2024-04-18T17:01:41Z', 'Comment': 'Xaqq weye nimankas in la laayo', 'like_count': 0, 'reply_count': 0, 'comment_length': 31}
{'author': '@user-ql8hg7cu5e', 'date': '2024-04-18T13:32:49Z', 'Comment': 'Mashallah ciidanka SID a cadi ma ahan', 'like_count': 0, 'reply_count': 0, 'comment_length': 31}

```

In [46]:

```

1 # Convert the List of comments to a pandas DataFrame
2 df_comments = pd.DataFrame(comments)
3
4 # Save the DataFrame to a CSV file
5 csv_file_name = "kiiska_qoslaaye4.csv"
6 df_comments.to_csv(csv_file_name, index=False, encoding='utf-8-sig')
7 print(f"Comments have been saved to {csv_file_name}")

```

Comments have been saved to kiiska_qoslaaye4.csv

In [47]:

```

1 df4 = pd.read_csv('kiiska_qoslaaye4.csv')
2 df4.head()

```

Out[47]:

	author	date	Comment	like_count	reply_count	comment_length
0	@fadumonour9815	2024-04-19T07:13:16Z	Eeda waxaa iskaleh Dawladda Somalia, oo dad aa...	1	0	318
1	@fadumonour9815	2024-04-19T06:58:53Z	😢😢😢	1	0	3
2	@muuseaweys	2024-04-18T18:36:41Z	Allow dowlaadeena noo sug oo noo xooji amni iyo...	0	0	67
3	@muuseaweys	2024-04-18T18:35:16Z	ماشاء الله تبارك الله	0	0	20
4	@saidkhalif7456	2024-04-18T17:01:41Z	Xaqq weye nimankas in la laayo	0	0	31

In [48]:

```

1 len(df4)

```

Out[48]:

517

Fifth Time Data Fetching

In [49]:

```
1 # actual video link'https://www.youtube.com/watch?v=RKWWs9_BHto&t=1s'
2
3 youtube = build('youtube', 'v3', developerKey=API_KEY)
4
5 def get_video_comments(video_id):
6     comments = []
7     response = youtube.commentThreads().list(
8         part='snippet,replies',
9         videoId=video_id,
10        maxResults=2150, # Adjust based on your needs
11        textFormat='plainText'
12    ).execute()
13
14    while response:
15        for item in response['items']:
16            comment = item['snippet']['topLevelComment']['snippet']
17            comments.append({
18                'author': comment['authorDisplayName'],
19                'date': comment['publishedAt'],
20                'Comment': comment['textDisplay'],
21                'like_count': comment['likeCount'],
22                'reply_count': item['snippet']['totalReplyCount'],
23                'comment_length': len(comment['textDisplay'])
24            })
25
26        # If there are replies, fetch them as well
27        if item['snippet']['totalReplyCount'] > 0:
28            for reply in item['replies']['comments']:
29                reply_comment = reply['snippet']
30                comments.append({
31                    'author': reply_comment['authorDisplayName'],
32                    'date': reply_comment['publishedAt'],
33                    'Comment': reply_comment['textDisplay'],
34                    'like_count': reply_comment['likeCount'],
35                    'reply_count': 0, # Replies do not have further
36                    'comment_length': len(reply_comment['textDisplay'])
37                })
38
39        # Check if there are more pages of comments
40        if 'nextPageToken' in response:
41            response = youtube.commentThreads().list(
42                part='snippet,replies',
43                videoId=video_id,
44                pageToken=response['nextPageToken'],
45                maxResults=100, # Adjust based on your needs
46                textFormat='plainText'
47            ).execute()
48        else:
49            break
50
51    return comments
52
53 # Example usage
54 video_id = 'UzNp89G8hqE'
55 comments = get_video_comments(video_id)
56 for comment in comments:
57     print(comment)
58
```

```
{'author': '@user-yo8ms8di3m', 'date': '2024-04-18T18:59:31Z', 'Comment': 'Caadiyaa uga jawabe meeshaan dood mataalo kulahaa 🎉🎉', 'like_count': 0, 'reply_count': 0, 'comment_length': 52}
{'author': '@Sarjipw', 'date': '2024-04-18T06:03:33Z', 'Comment': "Waxay dileen qof aan waxba galabsan, wayna qoslayaan cabsi la'aan. Aduunka kama jiro qof dil lagu xukumay oo faraxsan mise waxay ogyahiin in ladaynayo", 'like_count': 0, 'reply_count': 0, 'comment_length': 150}
{'author': '@user-zj2bc1dg4f', 'date': '2024-04-17T23:57:47Z', 'Comment': 'Qaliyara ninba wasodil yidhi bal kawaran', 'like_count': 0, 'reply_count': 0, 'comment_length': 40}
{'author': '@saidaahmedmohamed4996', 'date': '2024-04-16T09:29:35Z', 'Comment': 'Kuli ha ladilo tiirkha hala geeyo', 'like_count': 0, 'reply_count': 0, 'comment_length': 33}
{'author': '@dii-1055', 'date': '2024-04-16T05:04:32Z', 'Comment': 'Waxaa s aa kuu police ah ilaaha noo maqan', 'like_count': 0, 'reply_count': 0, 'comment_length': 40}
{'author': '@Nimcoabdimuuse', 'date': '2024-04-15T22:45:47Z', 'Comment': 'Mansha Allah allow Ada awodleh siyey marxunka xaqisa', 'like_count': 0, 'reply_count': 0, 'comment_length': 52}
```

In [50]:

```
1 # Convert the List of comments to a pandas DataFrame
2 df_comments = pd.DataFrame(comments)
3
4 # Save the DataFrame to a CSV file
5 csv_file_name = "kiiska_qoslaaye5.csv"
6 df_comments.to_csv(csv_file_name, index=False, encoding='utf-8-sig')
7 print(f"Comments have been saved to {csv_file_name}")
```

Comments have been saved to kiiska_qoslaaye5.csv

```
In [51]: 1 df5 = pd.read_csv('kiiska_qoslaaye5.csv')
2 df5.head()
```

Out[51]:

	author	date	Comment	like_count	reply_count	comment_leng
0	@user-yo8ms8di3m	2024-04-18T18:59:31Z	Caadiyaa uga jawabe meeshaan dood mataalo kula...	0	0	
1	@Sarjipw	2024-04-18T06:03:33Z	Waxay dileen qof aan waxba galabsan, wayna qos...	0	0	1
2	@user-zj2bc1dg4f	2024-04-17T23:57:47Z	Qaliyara ninba wasodil yidhi bal kawaran	0	0	.
3	@saidaahmedmohamed4996	2024-04-16T09:29:35Z	Kuli ha ladilo tiirka hala geeyo	0	0	.
4	@dii-1055	2024-04-16T05:04:32Z	Waxaas aa kuu police ah ilaaha noo maqan	0	0	.



```
In [52]: 1 len(df5)
```

Out[52]: 172

Sixth Time Data Fetching

In [53]:

```
1 # actual video link 'https://www.youtube.com/watch?v=MCMxqdu6A8s&t=3s'
2
3 youtube = build('youtube', 'v3', developerKey=API_KEY)
4
5 def get_video_comments(video_id):
6     comments = []
7     response = youtube.commentThreads().list(
8         part='snippet,replies',
9         videoId=video_id,
10        maxResults=2150, # Adjust based on your needs
11        textFormat='plainText'
12    ).execute()
13
14    while response:
15        for item in response['items']:
16            comment = item['snippet']['topLevelComment']['snippet']
17            comments.append({
18                'author': comment['authorDisplayName'],
19                'date': comment['publishedAt'],
20                'Comment': comment['textDisplay'],
21                'like_count': comment['likeCount'],
22                'reply_count': item['snippet']['totalReplyCount'],
23                'comment_length': len(comment['textDisplay'])
24            })
25
26        # If there are replies, fetch them as well
27        if item['snippet']['totalReplyCount'] > 0:
28            for reply in item['replies']['comments']:
29                reply_comment = reply['snippet']
30                comments.append({
31                    'author': reply_comment['authorDisplayName'],
32                    'date': reply_comment['publishedAt'],
33                    'Comment': reply_comment['textDisplay'],
34                    'like_count': reply_comment['likeCount'],
35                    'reply_count': 0, # Replies do not have further
36                    'comment_length': len(reply_comment['textDisplay'])
37                })
38
39        # Check if there are more pages of comments
40        if 'nextPageToken' in response:
41            response = youtube.commentThreads().list(
42                part='snippet,replies',
43                videoId=video_id,
44                pageToken=response['nextPageToken'],
45                maxResults=100, # Adjust based on your needs
46                textFormat='plainText'
47            ).execute()
48        else:
49            break
50
51    return comments
52
53 # Example usage
54 video_id = 'HD6k3uWzF4Y'
55 comments = get_video_comments(video_id)
56 for comment in comments:
57     print(comment)
58
```

```

{'author': '@HidaayoCali', 'date': '2024-04-18T20:42:09Z', 'Comment': ':)', 'like_count': 0, 'reply_count': 0, 'comment_length': 1}
{'author': '@blessedme8532', 'date': '2024-04-18T12:27:57Z', 'Comment': 'Cadaawe Wuxuu u hadlaa Qof sheekadaa waxbo ka ogeen', 'like_count': 0, 'reply_count': 0, 'comment_length': 51}
{'author': '@user-fs7bp5is6j', 'date': '2024-04-18T10:07:26Z', 'Comment': 'Alaahu akbar:((', 'like_count': 1, 'reply_count': 0, 'comment_length': 14}
{'author': '@habibosaid2872', 'date': '2024-04-17T21:26:11Z', 'Comment': 'xamar bahalo galeen bey noqoteey wixii dadka dhacey oo diley baa ciidamo lagadhigey mooryaan mafiyoo ah bacidanka kadhex shaqeysaneysa oo dadkii ho rey ubabi iyey masakiinta Xamar kunool een danbiga gelin allaha usahlo.', 'like_count': 0, 'reply_count': 0, 'comment_length': 216}
{'author': '@asiyabare995', 'date': '2024-04-17T19:19:34Z', 'Comment': 'Nimankaan xukun adag bay ubaahanyihiin', 'like_count': 0, 'reply_count': 0, 'comment_length': 38}
{'author': '@farixaabdirahman7979', 'date': '2024-04-17T03:29:50Z', 'Comment': 'Danbiga gelin allaha usahlo.'}

```

In [54]:

```

1 # Convert the List of comments to a pandas DataFrame
2 df_comments = pd.DataFrame(comments)
3
4 # Save the DataFrame to a CSV file
5 csv_file_name = "kiiska_qoslaaye6.csv"
6 df_comments.to_csv(csv_file_name, index=False, encoding='utf-8-sig')
7 print(f"Comments have been saved to {csv_file_name}")

```

Comments have been saved to kiiska_qoslaaye6.csv

In [55]:

```

1 df6 = pd.read_csv('kiiska_qoslaaye6.csv')
2 df6.head()

```

Out[55]:

	author	date	Comment	like_count	reply_count	comment_length
0	@HidaayoCali	2024-04-18T20:42:09Z	😊	0	0	1
1	@blessedme8532	2024-04-18T12:27:57Z	Cadaawe Wuxuu u hadlaa Qof sheekadaa waxbo ka ...	0	0	51
2	@user-fs7bp5is6j	2024-04-18T10:07:26Z	Alaahu akbar:(,	1	0	14
3	@habibosaid2872	2024-04-17T21:26:11Z	xamar bahalo galeen bey noqoteey wixii dadka dh...	0	0	216
4	@asiyabare995	2024-04-17T19:19:34Z	Nimankaan xukun adag bay ubaahanyihiin	0	0	38

In [25]:

```
1 len(df6)
```

Out[25]: 315

Seventh Time Data Fetching

In [56]:

```
1 # actual video link 'https://www.youtube.com/watch?v=FnWk-HTdSmg&t=3s'
2
3 youtube = build('youtube', 'v3', developerKey=API_KEY)
4
5 def get_video_comments(video_id):
6     comments = []
7     response = youtube.commentThreads().list(
8         part='snippet,replies',
9         videoId=video_id,
10        maxResults=2150, # Adjust based on your needs
11        textFormat='plainText'
12    ).execute()
13
14    while response:
15        for item in response['items']:
16            comment = item['snippet']['topLevelComment']['snippet']
17            comments.append({
18                'author': comment['authorDisplayName'],
19                'date': comment['publishedAt'],
20                'Comment': comment['textDisplay'],
21                'like_count': comment['likeCount'],
22                'reply_count': item['snippet']['totalReplyCount'],
23                'comment_length': len(comment['textDisplay'])
24            })
25
26        # If there are replies, fetch them as well
27        if item['snippet']['totalReplyCount'] > 0:
28            for reply in item['replies']['comments']:
29                reply_comment = reply['snippet']
30                comments.append({
31                    'author': reply_comment['authorDisplayName'],
32                    'date': reply_comment['publishedAt'],
33                    'Comment': reply_comment['textDisplay'],
34                    'like_count': reply_comment['likeCount'],
35                    'reply_count': 0, # Replies do not have further
36                    'comment_length': len(reply_comment['textDisplay'])
37                })
38
39        # Check if there are more pages of comments
40        if 'nextPageToken' in response:
41            response = youtube.commentThreads().list(
42                part='snippet,replies',
43                videoId=video_id,
44                pageToken=response['nextPageToken'],
45                maxResults=100, # Adjust based on your needs
46                textFormat='plainText'
47            ).execute()
48        else:
49            break
50
51    return comments
52
53 # Example usage
54 video_id = 'mGbYpCK3IhE'
55 comments = get_video_comments(video_id)
56 for comment in comments:
57     print(comment)
58
```

```

{'author': '@djdahabdjdahab1443', 'date': '2024-04-14T22:30:07Z', 'Comment': 'No way ceelka waa in ay qoda ama furtta Dowlada somaliya kadib lafah a wax lagabarikra DNA qofwalba dadkisa lagu ogankara', 'like_count': 0, 'reply_count': 0, 'comment_length': 120}
{'author': '@BUNTUQ_channel', 'date': '2024-04-14T16:11:04Z', 'Comment': 'Ikraan unbaad ku heesaan meeqaa ladilaaa nagala kala daaya intaad isleed ahahy daawasho hel😊😊', 'like_count': 0, 'reply_count': 0, 'comment_length': 92}
{'author': '@shiine111', 'date': '2024-04-14T15:49:04Z', 'Comment': 'Markii dad isku wada qoys ah gacanta uu u galoo amniga gobolka waxaan weeye wa xaa kadhsha waa in wax wayn laga badalaa qaab dhismeedka haayadaha amniga waana in lahelaa dad aan isku hayb ahayn', 'like_count': 0, 'reply_count': 0, 'comment_length': 191}
{'author': '@musemahamud9806', 'date': '2024-04-14T12:26:27Z', 'Comment': ' ', 'like_count': 0, 'reply_count': 0, 'comment_length': 1}
{'author': '@caashocabdshiremaxamed4204', 'date': '2024-04-14T08:29:31Z', 'Comment': '😢', 'like_count': 0, 'reply_count': 0, 'comment_length': 1}
{'author': '@Minnesota416', 'date': '2024-04-14T01:24:00Z', 'Comment': 'A'}

```

In [57]:

```

1 # Convert the list of comments to a pandas DataFrame
2 df_comments = pd.DataFrame(comments)
3
4 # Save the DataFrame to a CSV file
5 csv_file_name = "kiiska_qoslaaye7.csv"
6 df_comments.to_csv(csv_file_name, index=False, encoding='utf-8-sig')
7 print(f"Comments have been saved to {csv_file_name}")

```

Comments have been saved to kiiska_qoslaaye7.csv

In [58]:

```

1 df7 = pd.read_csv('kiiska_qoslaaye7.csv')
2 df7.head()

```

Out[58]:

	author	date	Comment	like_count	reply_count	comment_length
0	@djdahabdjdahab1443	2024-04-14T22:30:07Z	No way ceelka waa in ay qoda ama furtta Dowlad...	0	0	120
1	@BUNTUQ_channel	2024-04-14T16:11:04Z	Ikraan unbaad ku heesaan meeqaa ladilaaa nagal...	0	0	92
2	@shiine111	2024-04-14T15:49:04Z	Markii dad isku wada qoys ah gacanta uu u galoo...	0	0	191
3	@musemahamud9806	2024-04-14T12:26:27Z		0	0	1
4	@caashocabdshiremaxamed4204	2024-04-14T08:29:31Z	😢	0	0	1

```
In [59]: 1 len(df6)
```

```
Out[59]: 315
```

```
In [60]: 1 import pandas as pd  
2  
3 # Assuming das_1 and das_2 are the two dataset variables  
4  
5 # Combine the datasets into one variable  
6 combined_data = pd.concat([df1, df2, df3, df4, df5, df6, df7])  
7  
8 # Save the combined data to a CSV file  
9 combined_data.to_csv('combined_data.csv', index=False)
```

```
In [ ]: 1
```

```
In [ ]: 1
```

```
In [ ]: 1
```