NoSQL DATABASE PROJECT BY MD.SAIFUL ISLAM

MongoDB / NoSQL :

from pymongo import MongoClient

# Connect with MongoDB
client = MongoClient('mongodb://localhost:27017')

'''
I have tried to create a database of the top 10 news channels worldwide.
The information here is completely authentic and reliable.
I initiated my project on June 8th, 2024.
'''

db = client.YouTube_Database

# Create the Collections
channels_collection = db.channels
videos_collection = db.videos
playlists_collection = db.playlists
comments_collection = db.comments

Insert data into channels Collection
'''
 I have selected these ten news channels based on google research.
'''
channels = [
  {'channel_id': 1, 'channel_name': 'ABC News', 'country': 'USA', 'subscriber': 16700000,
   'videos_count': 89952, 'total_views': 14766436670,
'estimated_monthly_earnings_in_$': '26100-417000'},
  {'channel_id': 2, 'channel_name': 'Fox News', 'country': 'USA', 'subscriber': 11300000,
   'videos_count': 105817, 'total_views': 16870380190,
'estimated_monthly_earnings_in_$': '43300-692700'},
  {'channel_id': 3, 'channel_name': 'CNN', 'country': 'USA', 'subscriber': 16400000,
   'videos_count': 165130, 'total_views': 15922958540,
'estimated_monthly_earnings_in_$': '29700-474400'},
  {'channel_id': 4, 'channel_name': 'NBC News', 'country': 'USA', 'subscriber': 10000000,
   'videos_count': 65233, 'total_views': 7349074254,
'estimated_monthly_earnings_in_$': '28200-450600'},

```python
  {'channel_id': 5, 'channel_name': 'WION', 'country': 'INDIA', 'subscriber': 8790000,
   'videos_count': 155807, 'total_views': 4977235701,
'estimated_monthly_earnings_in_$': '15400-246800'},
  {'channel_id': 6, 'channel_name': 'Sky News', 'country': 'UK', 'subscriber': 7760000,
   'videos_count': 48412, 'total_views': 4397479437,
'estimated_monthly_earnings_in_$': '5000-79500'},
  {'channel_id': 7, 'channel_name': 'CBS News', 'country': 'USA', 'subscriber':
5850000,
   'videos_count': 123232, 'total_views': 5053137602,
'estimated_monthly_earnings_in_$': '10900-174500'},
  {'channel_id': 8, 'channel_name': 'BBC News', 'country': 'UK', 'subscriber':
16100000,
   'videos_count': 21837, 'total_views': 5344390233,
'estimated_monthly_earnings_in_$': '13600-217300'},
  {'channel_id': 9, 'channel_name': 'Al Jazeera English', 'country': 'QATAR',
'subscriber': 13600000,
   'videos_count': 115166, 'total_views': 4564028127,
'estimated_monthly_earnings_in_$': '21700-347000'},
  {'channel_id': 10, 'channel_name': 'TheYoungTurks', 'country': 'USA', 'subscriber':
5910000,
   'videos_count': 63197, 'total_views': 6736003984,
'estimated_monthly_earnings_in_$': '8200-131400'}
]

channels_collection.insert_many(channels)

# Insert data into videos Collection
# Focused on analyzing the newest videos from each channel.
videos = [
  {'video_id': 1, 'title': 'U.S. weapons used in Israeli airstrikes, plus U.S. military
officials repair damaged pier for aid', 'channel_id': 1, 'views': 6200, 'like': 45},
  {'video_id': 2, 'title': 'Trump: My revenge will be success', 'channel_id': 2, 'views':
380000, 'like': 11000},
  {'video_id': 3, 'title': 'Boeing spacecraft carrying two NASA astronauts lifts off in
historic launch', 'channel_id': 3, 'views': 205137, 'like': 2900},
  {'video_id': 4, 'title': 'Back-to-back Florida shark attacks startle start to summer
season', 'channel_id': 4, 'views': 24127, 'like': 155},
  {'video_id': 5, 'title': 'Russia,s border regions of Kursk & Belgorod under attack',
'channel_id': 5, 'views': 20000, 'like': 314},
  {'video_id': 6, 'title': 'PMQs: Rishi Sunak and Keir Starmer focus on Infected Blood
Inquiry', 'channel_id': 6, 'views': 5900, 'like': 59},
  {'video_id': 7, 'title': 'Israel conducts deadly strike on U.N. school in Gaza, says
Hamas militants operated there', 'channel_id': 7, 'views': 21000, 'like': 168},
```

```python
    {'video_id': 8, 'title': 'Will flying ever be sustainable?', 'channel_id': 8, 'views': 77000,
'like': 889},
    {'video_id': 9, 'title': 'Iran candidacy registration', 'channel_id': 9, 'views': 6400, 'like':
91},
    {'video_id': 10, 'title': 'Men In Remote Tribe Got Addicted To THIS After Getting
Internet', 'channel_id': 10, 'views': 44000, 'like': 1600}
]

videos_collection.insert_many(videos)

# Insert data into playlists Collection
# A particular playlist was chosen for the videos I analyzed.
playlists = [
    {'playlist_id': 1, 'channel_id': 1, 'title': 'Israel-Hamas War'},
    {'playlist_id': 2, 'channel_id': 2, 'title': 'Former President Donald Trump'},
    {'playlist_id': 3, 'channel_id': 3, 'title': 'Science & Technology'},
    {'playlist_id': 4, 'channel_id': 4, 'title': 'Top News'},
    {'playlist_id': 5, 'channel_id': 5, 'title': 'Global News 2024'},
    {'playlist_id': 6, 'channel_id': 6, 'title': 'Infected blood scandal'},
    {'playlist_id': 7, 'channel_id': 7, 'title': 'Israel-Hamas War'},
    {'playlist_id': 8, 'channel_id': 8, 'title': 'BBC ideas'},
    {'playlist_id': 9, 'channel_id': 9, 'title': 'Iran's president killed in helicopter crash'},
    {'playlist_id': 10, 'channel_id': 10, 'title': 'TheYoungTurks'}
]

playlists_collection.insert_many(playlists)

# Insert comments into comments Collection
# I selected comments that received the most likes.
comments = [
    {'comment_id': 1, 'video_id': 1, 'author': '@user-qu8lt3hf7e', 'comment_text':
'Remove Israel prime Minister'},
    {'comment_id': 2, 'video_id': 2, 'author': '@autotechandspecs', 'comment_text': 'You
can't have woke military'},
    {'comment_id': 3, 'video_id': 3, 'author': '@Pfromm007', 'comment_text': 'Stepping
into a Boeing spacecraft Proof that astronauts are the bravest human beings'},
    {'comment_id': 4, 'video_id': 4, 'author': '@justmejenny7986', 'comment_text': '4
separate attacks in a week or two. California, Texas and 2 in Florida. And summer
hasn,t even begun'},
    {'comment_id': 5, 'video_id': 5, 'author': '@user-sz9ue6xr5g', 'comment_text':
'Ukraine took the war to Russia,s ground.Strength to Ukraine'},
    {'comment_id': 6, 'video_id': 6, 'author': '@bryanwood8050', 'comment_text': 'Well,
wasn,t that easy - constructive conversations: no shouting, screaming, slagging off,
name calling or heckling. Brilliant.'},
```

{'comment_id': 7, 'video_id': 7, 'author': '@skycloud4802', 'comment_text': 'If they keep up with this, at this rate, there will be no Palestines left. Is that the intention?'},
    {'comment_id': 8, 'video_id': 8, 'author': '@user-qu8lt3hf7e', 'comment_text': 'Remember, it,s YOUR fault flying is not sustainable... definitely not all the super rich taking their private jets out for a 15 minute flight'},
    {'comment_id': 9, 'video_id': 9, 'author': '@mostafar8514', 'comment_text': 'as an iranian we hoped to vote raesi second round.'},
    {'comment_id': 10, 'video_id': 10, 'author': '@user-nf8id6di9b', 'comment_text': 'They probably should have put some restrictions'}
]

comments_collection.insert_many(comments)

Python Program:

```python
from pymongo import MongoClient

# Connect to the MongoDB server
client = MongoClient('mongodb://localhost:27017')

# Access the database
db = client.YouTube_Database

# Define the collections
collections = {
    'channels': db.channels,
    'videos': db.videos,
    'playlists': db.playlists,
    'comments': db.comments
}

# Process each collection to count the documents and print representative samples.
for name, collection in collections.items():
    document_count = collection.count_documents({})
    print(f'Count of documents in the {name} collection: {document_count}')

    print(f'Sample documents from the {name} collection:')

    # Fetch and print a sample of 3 documents from the collection
    sample_documents = collection.find().limit(3)
    for document in sample_documents:
        print(document)
```

```
    print('\n')

pipeline = [
    {
        '$lookup': {
            'from': 'videos',
            'localField': 'channel_id',
            'foreignField': 'channel_id',
            'as': 'videos'
        }
    },
    {'$limit': 3}
]

# Execute the aggregation pipeline on the channels collection
joined_result = db.channels.aggregate(pipeline)

print('Joined data (channels and videos):')

# Print each document from the joined result
for doc in joined_result:
    print(doc)
```

Result Output:

```
Count of documents in the channels collection: 10
Sample documents from the channels collection:
{'_id': ObjectId('666c1ab79cd2b8576bde92ab'), 'channel_id': 1, 'channel_name': 'ABC News', 'country': 'USA', 'subscriber': 16700000, 'videos_c
{'_id': ObjectId('666c1ab79cd2b8576bde92ac'), 'channel_id': 2, 'channel_name': 'Fox News', 'country': 'USA', 'subscriber': 11300000, 'videos_c
{'_id': ObjectId('666c1ab79cd2b8576bde92ad'), 'channel_id': 3, 'channel_name': 'CNN', 'country': 'USA', 'subscriber': 16400000, 'videos_count'

Count of documents in the videos collection: 10
Sample documents from the videos collection:
{'_id': ObjectId('666c1ab79cd2b8576bde92b5'), 'video_id': 1, 'title': 'U.S. weapons used in Israeli airstrikes, plus U.S. military officials r
{'_id': ObjectId('666c1ab79cd2b8576bde92b6'), 'video_id': 2, 'title': 'Trump: My revenge will be success', 'channel_id': 2, 'views': 380000, '
{'_id': ObjectId('666c1ab79cd2b8576bde92b7'), 'video_id': 3, 'title': 'Boeing spacecraft carrying two NASA astronauts lifts off in historic la

Count of documents in the playlists collection: 10
Sample documents from the playlists collection:
{'_id': ObjectId('666c1ab79cd2b8576bde92bf'), 'playlist_id': 1, 'channel_id': 1, 'title': 'Israel-Hamas War'}
{'_id': ObjectId('666c1ab79cd2b8576bde92c0'), 'playlist_id': 2, 'channel_id': 2, 'title': 'Former President Donald Trump'}
{'_id': ObjectId('666c1ab79cd2b8576bde92c1'), 'playlist_id': 3, 'channel_id': 3, 'title': 'Science & Technology'}

Count of documents in the comments collection: 10
```

```
Joined data (channels and videos):
{'_id': ObjectId('666c1ab79cd2b8576bde92ab'), 'channel_id': 1, 'channel_name': 'ABC News', 'country': 'USA', 'subscriber': 16700000, 'videos_co
{'_id': ObjectId('666c1ab79cd2b8576bde92ac'), 'channel_id': 2, 'channel_name': 'Fox News', 'country': 'USA', 'subscriber': 11300000, 'videos_co
{'_id': ObjectId('666c1ab79cd2b8576bde92ad'), 'channel_id': 3, 'channel_name': 'CNN', 'country': 'USA', 'subscriber': 16400000, 'videos_count':

Process finished with exit code 0
```