

# **CS 5513-001: ADVANCED DATABASE MANAGEMENT SYSTEM**

## **HOMEWORK 3**

### **Problem 1: Detailed Comparison of ACID in Oracle 23ai vs. MongoDB**

#### **1. Atomicity**

##### **Oracle 23ai:**

- **Definition & Guarantee:** Oracle defines a transaction as a logical, atomic unit of work composed of one or more SQL statements. All operations within this transaction are either fully completed or entirely undone.
- **Rollback Mechanism:** If the system fails after executing part of the transaction (e.g., 20 out of 100 row updates), Oracle automatically rolls back all partial changes, ensuring no intermediate state is visible or committed.
- **Transaction Identification:** Each transaction is uniquely identified by a Transaction ID, allowing Oracle to manage concurrent access and recovery accurately.

“All tasks of a transaction are performed or none of them are. There are no partial transactions. For example, if a transaction starts updating 100 rows, but the system fails after 20 updates, then the database rolls back the changes to these 20 rows.” – **Oracle 23ai Documentation**

##### **MongoDB:**

- **Default Behavior:** MongoDB ensures atomicity within a single document by default. Even if the document includes multiple nested fields or embedded subdocuments, the operation is atomic.
- **Multi-Document Transactions:** Since version 4.0 (replica sets) and 4.2 (sharded clusters), MongoDB allows multi-document, multi-collection atomic transactions.
- **Failure Handling:** If a system crash or error occurs during a multi-document transaction, all changes are rolled back to ensure atomicity is preserved.

“In MongoDB, a write operation is atomic on the level of a single document... MongoDB supports distributed transactions...” – **MongoDB ACID Docs**

**Comparison:** Oracle Database 23ai enforces atomicity by ensuring that all operations within a transaction are treated as an indivisible unit: either all changes are committed to the database or none are, with automatic rollback mechanisms handling partial failures. If, for example, a transaction updates 20 out of 100 rows before a failure, Oracle rolls back those 20 changes entirely. In contrast, MongoDB guarantees atomicity at the single-document level by default, ensuring that any changes made to one document—no matter how nested—are all-or-nothing. However, in modern

versions (4.0+), MongoDB supports multi-document atomic transactions on both replica sets and sharded clusters, enabling distributed transaction support similar to relational systems.

## 2. Consistency

### Oracle 23ai:

- **Definition:** A transaction must transition the database from one consistent state to another. It cannot violate referential integrity, domain constraints, check constraints, etc.
- **Constraint Enforcement:** Oracle ensures consistency by validating all data against schema constraints before committing a transaction.
- **Use Case:** For example, a bank transfer debits one account and credits another. If any part fails, no partial update occurs—preserving consistency.

“The transaction takes the database from one consistent state to another... a failure must not cause the database to credit only one account...” – **Oracle 23ai Documentation**

### MongoDB:

- **Consistency in Flexible Schemas:** MongoDB offers flexible schemas (schema-less), so consistency enforcement can vary. Developers can denormalize or duplicate data, and may need to handle manual consistency between collections.
- **Transactional Consistency:** In distributed transactions, MongoDB guarantees consistency across multiple documents and collections using write concerns and read snapshots.
- **Schema Implications:** With normalized schemas, MongoDB guarantees consistency using atomic writes. But with denormalized models, consistency depends on how/when duplicated fields are updated.

“Consistency guarantees that changes made within a transaction are populated across the database system... If data consistency is going to be negatively impacted... the entire transaction will fail.” – **MongoDB ACID Docs**

**Comparison:** When it comes to consistency, Oracle transitions the database from one valid state to another by strictly enforcing constraints such as referential integrity, check constraints, and domain rules. It ensures that no transaction can leave the database in an invalid state. MongoDB, while offering transactional consistency guarantees, provides schema flexibility—developers can denormalize or duplicate data to optimize for performance. In these cases, consistency enforcement partially becomes the developer’s responsibility. However, for transactions involving multiple documents or collections, MongoDB will automatically abort any transaction that would result in an inconsistent state.

### 3. Isolation:

#### Oracle 23ai:

- **Definition:** Isolation ensures that operations in one transaction are not visible to other concurrent transactions until committed.
- **Isolation Levels:** Oracle supports:
  - Read Committed (default) – avoids dirty reads
  - Serializable – strictest; simulates serial transaction execution
  - Read-Only – for queries without modifications
- **Concurrency Control:** Oracle uses multiversion concurrency control (MVCC) to isolate transaction reads from uncommitted writes.

“The effect of a transaction is not visible to other transactions until the transaction is committed... it appears to users as if transactions are executing serially.” – **Oracle 23ai Documentation**

#### MongoDB:

- **Snapshot Isolation:** Each transaction operates on a snapshot of data as it existed at the start. Reads are guaranteed to see a consistent view, and writes are only visible post-commit.
- **Conflict Resolution:** MongoDB aborts transactions that conflict with other concurrent writes.
  - Read & Write Concerns: Developers can configure:
    - Read Concern: Local, Majority, or Snapshot
    - Write Concern: Acknowledgement by majority of replica nodes

“MongoDB employs snapshot isolation... Transactions can read data from the snapshot of data committed at the time the transaction starts.”

“Any conflicting updates will cause the transaction to abort.” – **MongoDB ACID Docs**

**Comparison:** In terms of isolation, Oracle 23ai uses multiversion concurrency control (MVCC) and supports traditional isolation levels such as Read Committed and Serializable, giving developers fine-grained control over concurrent access and visibility of uncommitted data. Oracle ensures that a transaction’s intermediate state is never visible to other users, simulating serial execution. MongoDB implements snapshot isolation, where each transaction works with a consistent snapshot of the database taken at the beginning of the transaction. Reads see only committed data as of the transaction’s start time, and if any conflicting write occurs during the transaction, MongoDB aborts the operation. Developers can further configure

isolation strength using read concerns (e.g., snapshot) and write concerns (e.g., majority).

## 4. Durability

### Oracle 23ai:

- **Durability Mechanism:** Oracle writes transaction logs (redo logs) to disk before commit acknowledgment.
- **Crash Recovery:** In the event of a crash, Oracle uses online redo logs and archive logs to recover committed transactions.
- **Commit Guarantees:** Once the COMMIT command returns success, the changes are guaranteed to persist even in hardware failures.

“Changes made by committed transactions are permanent. After a transaction completes, the database ensures through its recovery mechanisms that changes from the transaction are not lost.” – **Oracle 23ai Documentation**

### MongoDB:

- **Oplog and Journaling:** MongoDB maintains a persistent operation log (Oplog) and journal files that record writes before committing them.
- **Write Concern:** Using majority write concern, MongoDB ensures data is written to the majority of replica nodes before confirming the commit.
- **Failure Recovery:** If the system crashes, MongoDB replays operations from the Oplog to ensure all committed writes are restored.

“MongoDB creates an OpLog... If there is an unforeseen event... the OpLog can be used when the system starts again to replay any writes...” – **MongoDB ACID Docs**

**Comparison:** Finally, for durability, Oracle ensures that once a transaction is committed, it is permanent. It writes all changes to redo logs on disk before acknowledging the commit, allowing complete recovery even in the case of system crashes or power failures. MongoDB uses a combination of journaling and an operation log (Oplog) to guarantee durability. Every write is recorded in the journal and replicated across nodes (if configured with majority write concern) before the transaction is considered committed. In the event of a crash, MongoDB replays operations from the Oplog to restore the last consistent state. In both systems, durability is tightly integrated into their recovery mechanisms, but while Oracle provides this behavior as a default, MongoDB’s guarantees depend on explicitly configured write concerns.

## Problem 2: MongoDB Implementation

### Task (a): Dataset Creation (articleData.txt)

```
Last login: Sun Apr 13 20:07:34 on ttys000
(base) ravisribhashyam@Sasanks-Macbook-2 ~ % scp ~/Desktop/articleData.txt srib0000@gpel9.cs.nor.ou.edu:~
srib0000@gpel9.cs.nor.ou.edu's password:
[articleData.txt          100%   34KB 427.7KB/s  00:00
(base) ravisribhashyam@Sasanks-Macbook-2 ~ % ssh srib0000@gpel9.cs.nor.ou.edu
[srib0000@gpel9.cs.nor.ou.edu's password:
[Welcome to Ubuntu 24.04.1 LTS (GNU/Linux 6.8.0-52-generic x86_64)

 * Documentation: https://help.ubuntu.com
 * Management:   https://landscape.canonical.com
 * Support:      https://ubuntu.com/pro

1 device has a firmware upgrade available.
Run `fwupdmgr get-upgrades` for more information.

Expanded Security Maintenance for Applications is not enabled.

264 updates can be applied immediately.
To see these additional updates run: apt list --upgradable

15 additional security updates can be applied with ESM Apps.
Learn more about enabling ESM Apps service at https://ubuntu.com/esm

*** System restart required ***

1 device has a firmware upgrade available.
Run `fwupdmgr get-upgrades` for more information.

Last login: Sun Apr 13 20:08:21 2025 from 10.199.192.103
srib0000@gpel9:~$ ls -l ~/articleData.txt
[-rw----- 1 srib0000 srib0000 34874 Apr 13 20:29 /home/srib0000/articleData.txt
```

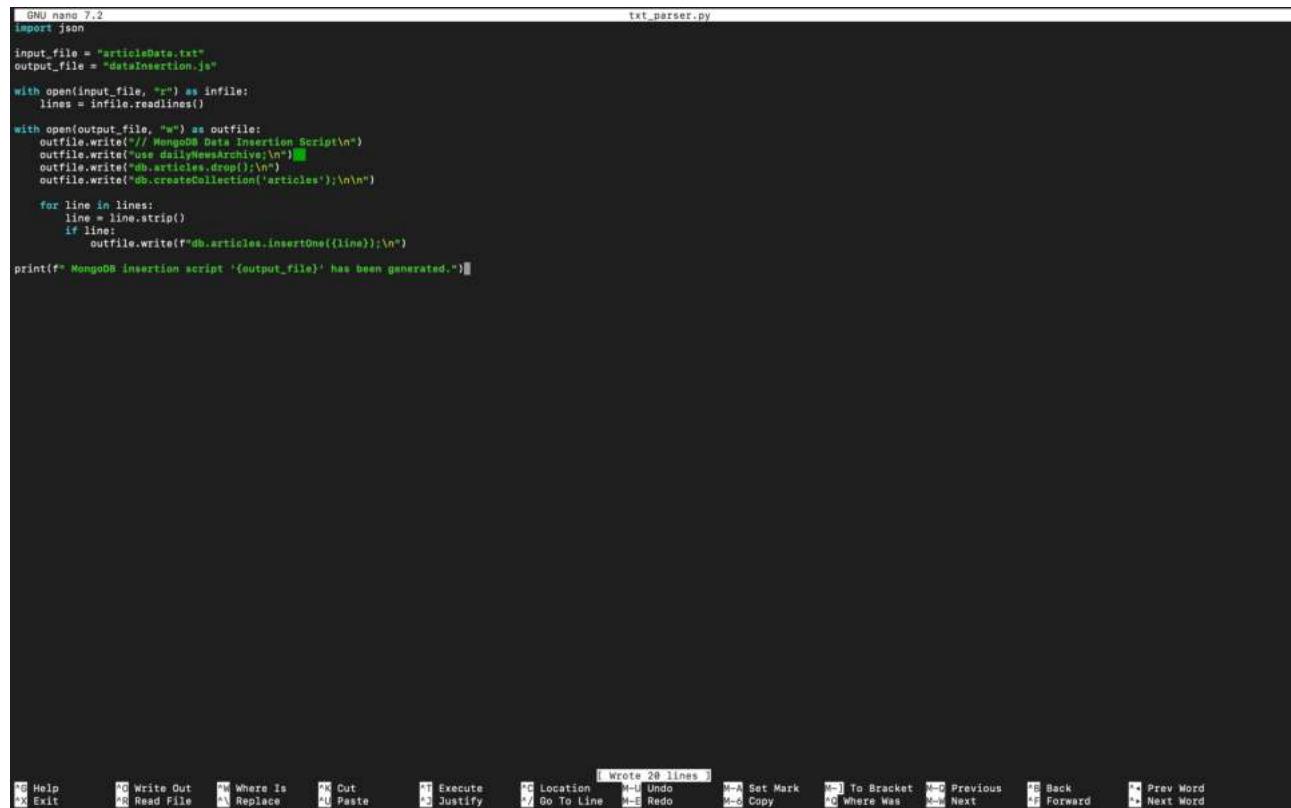
### Printing the first 5 entries of the dataset:

```
and comment"; "score": 4})]
srib0000@gpel9:~$ head -n 5 articleData.txt
{"article_type": "news article", "article_section": "Technology", "article_id": 1, "post_on_date": {"day": 2, "month": 1, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 1", "similar_stories": ["https://thedaily.com/politics/election-story-1.html", "https://thedaily.com/sports/election-story-1.html"], "num_times_read": 18964, "article_text": "This article discusses the latest developments in election.", "comments": [{"comment_id": 1, "article_id": 1, "user_id": "user1@mail.com", "posted_on_date": {"day": 2, "month": 1, "year": 2023, "time": "00:00:00"}, "comment_text": "Well-written and concise.", "score": -7/3}]}
{"article_type": "feature article", "article_section": "Technology", "article_id": 2, "post_on_date": {"day": 3, "month": 1, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 2", "similar_stories": ["https://thedaily.com/politics/education-story-2.html", "https://thedaily.com/sports/education-story-2.html"], "num_times_read": 6363, "article_text": "This article discusses the latest developments in education.", "comments": [{"comment_id": 2, "article_id": 2, "user_id": "user2@mail.com", "posted_on_date": {"day": 3, "month": 1, "year": 2023, "time": "00:00:00"}, "comment_text": "Needs better research.", "score": -25}]}
{"article_type": "feature article", "article_section": "Sports", "article_id": 3, "post_on_date": {"day": 4, "month": 1, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 3", "similar_stories": ["https://thedaily.com/politics/budget-story-3.html", "https://thedaily.com/sports/budget-story-3.html"], "num_times_read": 9484, "article_text": "This article discusses the latest developments in budget.", "comments": [{"comment_id": 3, "article_id": 3, "user_id": "user3@mail.com", "posted_on_date": {"day": 4, "month": 1, "year": 2023, "time": "00:00:00"}, "comment_text": "Not convinced with the facts.", "score": 4}]}
{"article_type": "news article", "article_section": "Technology", "article_id": 4, "post_on_date": {"day": 5, "month": 1, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 4", "similar_stories": ["https://thedaily.com/politics/education-story-4.html", "https://thedaily.com/sports/education-story-4.html"], "num_times_read": 6187, "article_text": "This article discusses the latest developments in education.", "comments": [{"comment_id": 4, "article_id": 4, "user_id": "user4@mail.com", "posted_on_date": {"day": 5, "month": 1, "year": 2023, "time": "00:00:00"}, "comment_text": "Well-written and concise.", "score": -3}]}
{"article_type": "news article", "article_section": "Entertainment", "article_id": 5, "post_on_date": {"day": 6, "month": 1, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 5", "similar_stories": ["https://thedaily.com/politics/film-story-5.html", "https://thedaily.com/sports/film-story-5.html"], "num_times_read": 4977, "article_text": "This article discusses the latest developments in film.", "comments": [{"comment_id": 5, "article_id": 5, "user_id": "user5@mail.com", "posted_on_date": {"day": 6, "month": 1, "year": 2023, "time": "00:00:00"}, "comment_text": "I disagree with the viewpoint.", "score": -75}]}]
```

## Task (b): MongoDB Insertion

```
1 import json
2
3 input_file = "articleData.txt"
4 output_file = "dataInsertion.js"
5
6 with open(input_file, "r") as infile:
7     lines = infile.readlines()
8
9 with open(output_file, "w") as outfile:
10    outfile.write("// MongoDB Data Insertion Script\n")
11    outfile.write("use dailyNewsArchive;\n")
12    outfile.write("db.articles.drop();\n")
13    outfile.write("db.createCollection('articles');\n\n")
14
15    for line in lines:
16        line = line.strip()
17        if line:
18            outfile.write(f"db.articles.insertOne({line});\n")
19
20 print(f" MongoDB insertion script '{output_file}' has been generated.")
```

```
GNU nano 7.2                                     txt_parser.py
import json
input_file = "articleData.txt"
output_file = "dataInsertion.js"
with open(input_file, "r") as infile:
    lines = infile.readlines()
with open(output_file, "w") as outfile:
    outfile.write("// MongoDB Data Insertion Script\n")
    outfile.write("use dailyNewsArchive;\n")
    outfile.write("db.articles.drop();\n")
    outfile.write("db.createCollection('articles');\n\n")
    for line in lines:
        line = line.strip()
        if line:
            outfile.write(f"db.articles.insertOne({line});\n")
print(f" MongoDB insertion script '{output_file}' has been generated.")
```



## Task (c): Populating MongoDB with Insertions

```
zihb400@opel19: ~ $ python3 txt_parser.py
MongoDB insertion script 'dataInsertion.js' has been generated.
zihb400@opel19: ~ $ mongoosh --port 24157 < dataInsertion.js
Current Mongosh Log ID: 67fc77235b70bea4836b14ea
Connecting to: mongosh://127.0.0.1:24157/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+2.4.2
Using Mongosh: 2.4.2
Using Mongosh: 2.4.2
mongosh 2.8.0 is available for download: https://www.mongodb.com/try/download/shell
For mongosh info see: https://www.mongodb.com/docs/mongodb-shell

The server generated these startup warnings when booting
2025-04-13T18:53:07.873-05:00: Using the XFS filesystem is strongly recommended with the WiredTiger storage engine. See http://dochub.mongodb.org/core/prodnotes-filesystem
2025-04-13T18:53:08.266-05:00: Access control is not enabled for the database. Read and write access to data and configuration is unrestricted
2025-04-13T18:53:08.266-05:00: Soft rlimits for open file descriptors too low
2025-04-13T18:53:08.266-05:00: For customers running the current memory allocator, we suggest changing the contents of the following sysfsFile
2025-04-13T18:53:08.266-05:00: For customers running the current memory allocator, we suggest changing the contents of the following sysfsFile
2025-04-13T18:53:08.266-05:00: We suggest setting the contents of sysfsFile to 0.
2025-04-13T18:53:08.266-05:00: Your system has glibc support for rseq built in, which is not yet supported by tcmalloc-google and has critical performance implications. Please set the environment variable GLIBC_TUNABLES=glibc_pthread_rseq=0
2025-04-13T18:53:08.266-05:00: We suggest setting swappiness to 0 or 1, as swapping can cause performance problems.

test> // MongoDB Data Insertion Script
test> use dailyNewsArchive;
switched to db dailyNewsArchive
dailyNewsArchive> db.articles.drop();
true
dailyNewsArchive> db.createCollection('articles');
{ ok: 1 }
dailyNewsArchive>
```

```
dailyNewsArchive> db.articles.insertOne({"article_type": "news article", "article_section": "Technology", "article_id": 1, "post_on_date": {"day": 2, "month": 1, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 1", "similar_stories": [{"https://thedsaily.com/politics/election-story-1.html", "https://thedsaily.com/sports/selection-story-1.html"}, "num_times_read": 18964, "article_text": "This article discusses the latest developments in election.", "comments": [{"comment_id": 1, "article_id": 1, "user_id": "user1@mail.com", "posted_on_date": {"day": 2, "month": 1, "year": 2023, "time": "00:00:00"}, "comment_text": "Well-written and concise.", "score": -77}]});
{
  acknowledged: true,
  insertedId: ObjectID("67fc77245b70bea4836b14eb")
}
dailyNewsArchive> db.articles.insertOne({"article_type": "feature article", "article_section": "Technology", "article_id": 2, "post_on_date": {"day": 3, "month": 1, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 2", "similar_stories": [{"https://thedsaily.com/politics/education-story-2.html", "https://thedsaily.com/sports/education-story-2.html"}, "num_times_read": 6363, "article_text": "This article discusses the latest developments in education.", "comments": [{"comment_id": 2, "article_id": 2, "user_id": "user2@mail.com", "posted_on_date": {"day": 3, "month": 1, "year": 2023, "time": "00:00:00"}, "comment_text": "Needs better research.", "score": -20}]});
{
  acknowledged: true,
  insertedId: ObjectID("67fc77245b70bea4836b14ec")
}
dailyNewsArchive> db.articles.insertOne({"article_type": "feature article", "article_section": "Sports", "article_id": 3, "post_on_date": {"day": 4, "month": 1, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 3", "similar_stories": [{"https://thedsaily.com/politics/budget-story-3.html", "https://thedsaily.com/sports/budget-story-3.html"}, "num_times_read": 9684, "article_text": "This article discusses the latest developments in budget.", "comments": [{"comment_id": 3, "article_id": 3, "user_id": "user3@mail.com", "posted_on_date": {"day": 4, "month": 1, "year": 2023, "time": "00:00:00"}, "comment_text": "Not convinced with the facts.", "score": 4}]});
{
  acknowledged: true,
  insertedId: ObjectID("67fc77245b70bea4836b14ed")
}
dailyNewsArchive> db.articles.insertOne({"article_type": "news article", "article_section": "Technology", "article_id": 4, "post_on_date": {"day": 5, "month": 1, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 4", "similar_stories": [{"https://thedsaily.com/politics/education-story-4.html", "https://thedsaily.com/sports/education-story-4.html"}, "num_times_read": 6187, "article_text": "This article discusses the latest developments in education.", "comments": [{"comment_id": 4, "article_id": 4, "user_id": "user4@mail.com", "posted_on_date": {"day": 5, "month": 1, "year": 2023, "time": "00:00:00"}, "comment_text": "Well-written and concise.", "score": -3}]}];
{
  acknowledged: true,
  insertedId: ObjectID("67fc77245b70bea4836b14ef")
}
dailyNewsArchive> db.articles.insertOne({"article_type": "news article", "article_section": "Entertainment", "article_id": 5, "post_on_date": {"day": 6, "month": 1, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 5", "similar_stories": [{"https://thedsaily.com/politics/film-story-5.html", "https://thedsaily.com/sports/film-story-5.html"}, "num_times_read": 4977, "article_text": "This article discusses the latest developments in film.", "comments": [{"comment_id": 5, "article_id": 5, "user_id": "user5@mail.com", "posted_on_date": {"day": 6, "month": 1, "year": 2023, "time": "00:00:00"}, "comment_text": "I disagree with the viewpoint.", "score": -10}]}];
{
  acknowledged: true,
  insertedId: ObjectID("67fc77245b70bea4836b14ef")
}
dailyNewsArchive> db.articles.insertOne({"article_type": "feature article", "article_section": "Politics", "article_id": 6, "post_on_date": {"day": 7, "month": 1, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 6", "similar_stories": [{"https://thedsaily.com/politics/AI-story-6.html", "https://thedsaily.com/sports/AI-story-6.html"}, "num_times_read": 18876, "article_text": "This article discusses the latest developments in AI.", "comments": [{"comment_id": 6, "article_id": 6, "user_id": "user6@mail.com", "posted_on_date": {"day": 7, "month": 1, "year": 2023, "time": "00:00:00"}, "comment_text": "Needs better research.", "score": -43}]}];
{
  acknowledged: true,
  insertedId: ObjectID("67fc77255b70bea4836b1410")
}
dailyNewsArchive> db.articles.insertOne({"article_type": "feature article", "article_section": "Technology", "article_id": 7, "post_on_date": {"day": 8, "month": 1, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 7", "similar_stories": [{"https://thedsaily.com/politics/education-story-7.html", "https://thedsaily.com/sports/education-story-7.html"}, "num_times_read": 14548, "article_text": "This article discusses the latest developments in education.", "comments": [{"comment_id": 7, "article_id": 7, "user_id": "user7@mail.com", "posted_on_date": {"day": 8, "month": 1, "year": 2023, "time": "00:00:00"}, "comment_text": "Not convinced with the facts.", "score": 19}]}];
{
  acknowledged: true,
  insertedId: ObjectID("67fc77255b70bea4836b1411")
}
dailyNewsArchive> db.articles.insertOne({"article_type": "news article", "article_section": "Technology", "article_id": 8, "post_on_date": {"day": 9, "month": 1, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 8", "similar_stories": [{"https://thedsaily.com/politics/cricket-story-8.html", "https://thedsaily.com/sports/cricket-story-8.html"}, "num_times_read": 9871, "article_text": "This article discusses the latest developments in cricket.", "comments": [{"comment_id": 8, "article_id": 8, "user_id": "user8@mail.com", "posted_on_date": {"day": 9, "month": 1, "year": 2023, "time": "00:00:00"}, "comment_text": "Very informative piece.", "score": 97}]}];
{
  acknowledged: true,
  insertedId: ObjectID("67fc77255b70bea4836b1412")
```

```

dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Sports", "article_id": 9, "post_on_date": {"day": 10, "month": 1, "year": 2023}, "time": "00:00:00"}, "reporter_name": "Reporter 9", "similar_stories": [{"url": "https://thedaily.com/politics/election-story-9.html", "https://thedaily.com/sports/election-story-9.html"}, {"num_times_read": 8243, "article_text": "This article discusses the latest developments in election.", "comments": [{"comment_id": 9, "article_id": 9, "user_id": "user9@mail.com", "posted_on_date": {"day": 10, "month": 1, "year": 2023}, "time": "00:00:00"}, {"comment_text": "Keep up the good work!", "score": 60}]}];
{
  acknowledged: true,
  insertedId: Objectid('67fc7725b76bea4836b1413')
}
dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Sports", "article_id": 10, "post_on_date": {"day": 11, "month": 1, "year": 2023}, "time": "00:00:00"}, "reporter_name": "Reporter 10", "similar_stories": [{"url": "https://thedaily.com/politics/file-story-10.html", "https://thedaily.com/sports/file-story-10.html"}, {"num_times_read": 11859, "article_text": "This article discusses the latest developments in film.", "comments": [{"comment_id": 10, "article_id": 10, "user_id": "user10@mail.com", "posted_on_date": {"day": 11, "month": 1, "year": 2023}, "time": "00:00:00"}, {"comment_text": "Keep up the good work!", "score": -60}]}];
{
  acknowledged: true,
  insertedId: Objectid('67fc7725b76bea4836b1414')
}
dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Health", "article_id": 11, "post_on_date": {"day": 12, "month": 1, "year": 2023}, "time": "00:00:00"}, "reporter_name": "Reporter 11", "similar_stories": [{"url": "https://thedaily.com/politics/election-story-11.html", "https://thedaily.com/sports/election-story-11.html"}, {"num_times_read": 1941, "article_text": "This article discusses the latest developments in election.", "comments": [{"comment_id": 11, "article_id": 11, "user_id": "user11@mail.com", "posted_on_date": {"day": 12, "month": 1, "year": 2023}, "time": "00:00:00"}, {"comment_text": "Very informative piece.", "score": 67}]}];
{
  acknowledged: true,
  insertedId: Objectid('67fc7725b76bea4836b1415')
}
dailyNewsArchive> db.articles.insertOne({article_type: "feature article", "article_section": "Health", "article_id": 12, "post_on_date": {"day": 13, "month": 1, "year": 2023}, "time": "00:00:00"}, "reporter_name": "Reporter 12", "similar_stories": [{"url": "https://thedaily.com/politics/space-story-12.html", "https://thedaily.com/sports/space-story-12.html"}, {"num_times_read": 1710, "article_text": "This article discusses the latest developments in space.", "comments": [{"comment_id": 12, "article_id": 12, "user_id": "user12@mail.com", "posted_on_date": {"day": 13, "month": 1, "year": 2023}, "time": "00:00:00"}, {"comment_text": "Bold take on current affairs.", "score": 69}]}];
{
  acknowledged: true,
  insertedId: Objectid('67fc7725b76bea4836b1416')
}
dailyNewsArchive> db.articles.insertOne({article_type: "feature article", "article_section": "Health", "article_id": 13, "post_on_date": {"day": 14, "month": 1, "year": 2023}, "time": "00:00:00"}, "reporter_name": "Reporter 13", "similar_stories": [{"url": "https://thedaily.com/politics/education-story-13.html", "https://thedaily.com/sports/education-story-13.html"}, {"num_times_read": 7520, "article_text": "This article discusses the latest developments in education.", "comments": [{"comment_id": 13, "article_id": 13, "user_id": "user13@mail.com", "posted_on_date": {"day": 14, "month": 1, "year": 2023}, "time": "00:00:00"}, {"comment_text": "Very informative piece.", "score": -65}]}];
{
  acknowledged: true,
  insertedId: Objectid('67fc7725b76bea4836b1417')
}
dailyNewsArchive> db.articles.insertOne({article_type: "feature article", "article_section": "Technology", "article_id": 14, "post_on_date": {"day": 15, "month": 1, "year": 2023}, "time": "00:00:00"}, "reporter_name": "Reporter 14", "similar_stories": [{"url": "https://thedaily.com/politics/election-story-14.html", "https://thedaily.com/sports/election-story-14.html"}, {"num_times_read": 7389, "article_text": "This article discusses the latest developments in election.", "comments": [{"comment_id": 14, "article_id": 14, "user_id": "user14@mail.com", "posted_on_date": {"day": 15, "month": 1, "year": 2023}, "time": "00:00:00"}, {"comment_text": "Bold take on current affairs.", "score": -123}]}];
{
  acknowledged: true,
  insertedId: Objectid('67fc7725b76bea4836b1418')
}
dailyNewsArchive> db.articles.insertOne({article_type: "feature article", "article_section": "Sports", "article_id": 15, "post_on_date": {"day": 16, "month": 1, "year": 2023}, "time": "00:00:00"}, "reporter_name": "Reporter 15", "similar_stories": [{"url": "https://thedaily.com/politics/budget-story-15.html", "https://thedaily.com/sports/budget-story-15.html"}, {"num_times_read": 2085, "article_text": "This article discusses the latest developments in budget.", "comments": [{"comment_id": 15, "article_id": 15, "user_id": "user15@mail.com", "posted_on_date": {"day": 16, "month": 1, "year": 2023}, "time": "00:00:00"}, {"comment_text": "Needs better research.", "score": 85}]}];
{
  acknowledged: true,
  insertedId: Objectid('67fc7725b76bea4836b1419')
}
dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Health", "article_id": 16, "post_on_date": {"day": 17, "month": 1, "year": 2023}, "time": "00:00:00"}, "reporter_name": "Reporter 16", "similar_stories": [{"url": "https://thedaily.com/politics/budget-story-16.html", "https://thedaily.com/sports/budget-story-16.html"}, {"num_times_read": 1692, "article_text": "This article discusses the latest developments in budget.", "comments": [{"comment_id": 16, "article_id": 16, "user_id": "user16@mail.com", "posted_on_date": {"day": 17, "month": 1, "year": 2023}, "time": "00:00:00"}, {"comment_text": "I disagree with the viewpoint.", "score": 99}]}];
{
  acknowledged: true,
  insertedId: Objectid('67fc7725b76bea4836b141a')
}
dailyNewsArchive> db.articles.insertOne({article_type: "feature article", "article_section": "Politics", "article_id": 17, "post_on_date": {"day": 18, "month": 1, "year": 2023}, "time": "00:00:00"}, "reporter_name": "Reporter 17", "similar_stories": [{"url": "https://thedaily.com/politics/space-story-17.html", "https://thedaily.com/sports/space-story-17.html"}, {"num_times_read": 16364, "article_text": "This article discusses the latest developments in space.", "comments": [{"comment_id": 17, "article_id": 17, "user_id": "user17@mail.com", "posted_on_date": {"day": 18, "month": 1, "year": 2023}, "time": "00:00:00"}, {"comment_text": "Amazing insight!", "score": 40}]}];
{
  acknowledged: true,
  insertedId: Objectid('67fc7725b76bea4836b141b')
}
dailyNewsArchive> db.articles.insertOne({article_type: "feature article", "article_section": "Health", "article_id": 18, "post_on_date": {"day": 19, "month": 1, "year": 2023}, "time": "00:00:00"}, "reporter_name": "Reporter 18", "similar_stories": [{"url": "https://thedaily.com/politics/cricket-story-18.html", "https://thedaily.com/sports/cricket-story-18.html"}, {"num_times_read": 16461, "article_text": "This article discusses the latest developments in cricket.", "comments": [{"comment_id": 18, "article_id": 18, "user_id": "user18@mail.com", "posted_on_date": {"day": 19, "month": 1, "year": 2023}, "time": "00:00:00"}, {"comment_text": "I disagree with the viewpoint.", "score": 71}]}];
{
  acknowledged: true,
  insertedId: Objectid('67fc7725b76bea4836b141c')
}
dailyNewsArchive> db.articles.insertOne({article_type: "feature article", "article_section": "Technology", "article_id": 19, "post_on_date": {"day": 20, "month": 1, "year": 2023}, "time": "00:00:00"}, "reporter_name": "Reporter 19", "similar_stories": [{"url": "https://thedaily.com/politics/file-story-19.html", "https://thedaily.com/sports/file-story-19.html"}, {"num_times_read": 3099, "article_text": "This article discusses the latest developments in file.", "comments": [{"comment_id": 19, "article_id": 19, "user_id": "user19@mail.com", "posted_on_date": {"day": 20, "month": 1, "year": 2023}, "time": "00:00:00"}, {"comment_text": "Looking forward to more.", "score": -6}]}];
{
  acknowledged: true,
  insertedId: Objectid('67fc7725b76bea4836b141d')
}
dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Health", "article_id": 20, "post_on_date": {"day": 21, "month": 1, "year": 2023}, "time": "00:00:00"}, "reporter_name": "Reporter 20", "similar_stories": [{"url": "https://thedaily.com/politics/covid-story-20.html", "https://thedaily.com/sports/covid-story-20.html"}, {"num_times_read": 5920, "article_text": "This article discusses the latest developments in covid.", "comments": [{"comment_id": 20, "article_id": 20, "user_id": "user20@mail.com", "posted_on_date": {"day": 21, "month": 1, "year": 2023}, "time": "00:00:00"}, {"comment_text": "Not convinced with the facts.", "score": 79}]}];
{
  acknowledged: true,
  insertedId: Objectid('67fc7725b76bea4836b141e')
}
dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Technology", "article_id": 21, "post_on_date": {"day": 22, "month": 1, "year": 2023}, "time": "00:00:00"}, "reporter_name": "Reporter 21", "similar_stories": [{"url": "https://thedaily.com/politics/startup-story-21.html", "https://thedaily.com/sports/startup-story-21.html"}, {"num_times_read": 5115, "article_text": "This article discusses the latest developments in startup.", "comments": [{"comment_id": 21, "article_id": 21, "user_id": "user21@mail.com", "posted_on_date": {"day": 22, "month": 1, "year": 2023}, "time": "00:00:00"}, {"comment_text": "Fantastic read!", "score": -33}]}];
{
  acknowledged: true,
  insertedId: Objectid('67fc7725b76bea4836b141f')
}
dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Politics", "article_id": 22, "post_on_date": {"day": 23, "month": 1, "year": 2023}, "time": "00:00:00"}, "reporter_name": "Reporter 22", "similar_stories": [{"url": "https://thedaily.com/politics/election-story-22.html", "https://thedaily.com/sports/election-story-22.html"}, {"num_times_read": 3196, "article_text": "This article discusses the latest developments in election.", "comments": [{"comment_id": 22, "article_id": 22, "user_id": "user22@mail.com", "posted_on_date": {"day": 23, "month": 1, "year": 2023}, "time": "00:00:00"}, {"comment_text": "Not convinced with the facts.", "score": 72}]}];
{
  acknowledged: true,
  insertedId: Objectid('67fc7725b76bea4836b1420')
}
dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Technology", "article_id": 23, "post_on_date": {"day": 24, "month": 1, "year": 2023}, "time": "00:00:00"}, "reporter_name": "Reporter 23", "similar_stories": [{"url": "https://thedaily.com/politics/ai-story-23.html", "https://thedaily.com/sports/ai-story-23.html"}, {"num_times_read": 7968, "article_text": "This article discusses the latest developments in AI.", "comments": [{"comment_id": 23, "article_id": 23, "user_id": "user23@mail.com", "posted_on_date": {"day": 24, "month": 1, "year": 2023}, "time": "00:00:00"}, {"comment_text": "Needs better research.", "score": -68}]}];
{
  acknowledged: true,
  insertedId: Objectid('67fc7725b76bea4836b1421')
}
dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Sports", "article_id": 24, "post_on_date": {"day": 25, "month": 1, "year": 2023}, "time": "00:00:00"}, "reporter_name": "Reporter 24", "similar_stories": [{"url": "https://thedaily.com/politics/file-story-24.html", "https://thedaily.com/sports/file-story-24.html"}, {"num_times_read": 4046, "article_text": "This article discusses the latest developments in film.", "comments": [{"comment_id": 24, "article_id": 24, "user_id": "user24@mail.com", "posted_on_date": {"day": 25, "month": 1, "year": 2023}, "time": "00:00:00"}, {"comment_text": "Fantastic read!", "score": -40}]}];
{
  acknowledged: true,
  insertedId: Objectid('67fc7725b76bea4836b1422')
}

```

```
dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Entertainment", "article_id": 25, "post_on_date": {"day": 26, "month": 4, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 25", "similar_stories": [{"https://thedaily.com/politics/election-story-26.html", "https://thedaily.com/sports/election-story-26.html"}, {"num_times_read": 9833, "article_text": "This article discusses the latest developments in election.", "comments": [{"comment_id": 26, "article_id": 25, "user_id": "user29@mail.com", "posted_on_date": {"day": 26, "month": 1, "year": 2023, "time": "00:00:00"}, {"comment_text": "Hold take on current affairs.", "score": 86}]}], "acknowledged": true, "insertedId": ObjectId('67fc7725b76bea4836b1429')}
```

```
dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Politics", "article_id": 26, "post_on_date": {"day": 27, "month": 1, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 26", "similar_stories": [{"https://thedaily.com/politics/climate-story-26.html", "https://thedaily.com/sports/climate-story-26.html"}, {"num_times_read": 2711, "article_text": "This article discusses the latest developments in climate.", "comments": [{"comment_id": 26, "article_id": 26, "user_id": "user36@mail.com", "posted_on_date": {"day": 27, "month": 1, "year": 2023, "time": "00:00:00"}, {"comment_text": "I disagree with the viewpoint.", "score": 88}]}], "acknowledged": true, "insertedId": ObjectId('67fc7725b76bea4836b1429')}
```

```
dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Sports", "article_id": 27, "post_on_date": {"day": 28, "month": 1, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 27", "similar_stories": [{"https://thedaily.com/politics/education-story-27.html", "https://thedaily.com/sports/education-story-27.html"}, {"num_times_read": 251, "article_text": "This article discusses the latest developments in education.", "comments": [{"comment_id": 27, "article_id": 27, "user_id": "user27@mail.com", "posted_on_date": {"day": 28, "month": 1, "year": 2023, "time": "00:00:00"}, {"comment_text": "Keep up the good work!", "score": -21}]}], "acknowledged": true, "insertedId": ObjectId('67fc7725b76bea4836b1429')}
```

```
dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Entertainment", "article_id": 28, "post_on_date": {"day": 29, "month": 1, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 28", "similar_stories": [{"https://thedaily.com/politics/file-story-28.html", "https://thedaily.com/sports/file-story-28.html"}, {"num_times_read": 7659, "article_text": "This article discusses the latest developments in file.", "comments": [{"comment_id": 28, "article_id": 28, "user_id": "user28@mail.com", "posted_on_date": {"day": 29, "month": 1, "year": 2023, "time": "00:00:00"}, {"comment_text": "Keep up the good work!", "score": -32}]}], "acknowledged": true, "insertedId": ObjectId('67fc7725b76bea4836b1429')}
```

```
dailyNewsArchive> db.articles.insertOne({article_type: "feature article", "article_section": "Sports", "article_id": 29, "post_on_date": {"day": 30, "month": 1, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 29", "similar_stories": [{"https://thedaily.com/politics/climate-story-29.html", "https://thedaily.com/sports/climate-story-29.html"}, {"num_times_read": 1110, "article_text": "This article discusses the latest developments in climate.", "comments": [{"comment_id": 29, "article_id": 29, "user_id": "user29@mail.com", "posted_on_date": {"day": 30, "month": 1, "year": 2023, "time": "00:00:00"}, {"comment_text": "Needs better research.", "score": 78}]}], "acknowledged": true, "insertedId": ObjectId('67fc7725b76bea4836b1429')}
```

```
dailyNewsArchive> db.articles.insertOne({article_type: "feature article", "article_section": "Entertainment", "article_id": 30, "post_on_date": {"day": 31, "month": 1, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 30", "similar_stories": [{"https://thedaily.com/politics/file-story-30.html", "https://thedaily.com/sports/file-story-30.html"}, {"num_times_read": 16109, "article_text": "This article discusses the latest developments in file.", "comments": [{"comment_id": 30, "article_id": 30, "user_id": "user30@mail.com", "posted_on_date": {"day": 31, "month": 1, "year": 2023, "time": "00:00:00"}, {"comment_text": "Well-written and concise.", "score": 72}]}], "acknowledged": true, "insertedId": ObjectId('67fc7725b76bea4836b1429')}
```

```
dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Sports", "article_id": 31, "post_on_date": {"day": 1, "month": 2, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 31", "similar_stories": [{"https://thedaily.com/politics/film-story-31.html", "https://thedaily.com/sports/film-story-31.html"}, {"num_times_read": 12499, "article_text": "This article discusses the latest developments in film.", "comments": [{"comment_id": 31, "article_id": 31, "user_id": "user31@mail.com", "posted_on_date": {"day": 1, "month": 2, "year": 2023, "time": "00:00:00"}, {"comment_text": "Not convinced with the facts.", "score": -30}]}], "acknowledged": true, "insertedId": ObjectId('67fc7725b76bea4836b1429')}
```

```
dailyNewsArchive> db.articles.insertOne({article_type: "feature article", "article_section": "Technology", "article_id": 32, "post_on_date": {"day": 2, "month": 2, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 32", "similar_stories": [{"https://thedaily.com/politics/ai-story-32.html", "https://thedaily.com/sports/ai-story-32.html"}, {"num_times_read": 6239, "article_text": "This article discusses the latest developments in AI.", "comments": [{"comment_id": 32, "article_id": 32, "user_id": "user32@mail.com", "posted_on_date": {"day": 2, "month": 2, "year": 2023, "time": "00:00:00"}, {"comment_text": "Keep up the good work!", "score": -90}]}], "acknowledged": true, "insertedId": ObjectId('67fc7725b76bea4836b1429')}
```

```
dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Entertainment", "article_id": 33, "post_on_date": {"day": 3, "month": 2, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 33", "similar_stories": [{"https://thedaily.com/politics/election-story-33.html", "https://thedaily.com/sports/election-story-33.html"}, {"num_times_read": 6798, "article_text": "This article discusses the latest developments in election.", "comments": [{"comment_id": 33, "article_id": 33, "user_id": "user33@mail.com", "posted_on_date": {"day": 3, "month": 2, "year": 2023, "time": "00:00:00"}, {"comment_text": "Looking forward to more.", "score": 83}]}], "acknowledged": true, "insertedId": ObjectId('67fc7725b76bea4836b1429')}
```

```
dailyNewsArchive> db.articles.insertOne({article_type: "feature article", "article_section": "Health", "article_id": 34, "post_on_date": {"day": 4, "month": 2, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 34", "similar_stories": [{"https://thedaily.com/politics/budget-story-34.html", "https://thedaily.com/sports/budget-story-34.html"}, {"num_times_read": 5189, "article_text": "This article discusses the latest developments in budget.", "comments": [{"comment_id": 34, "article_id": 34, "user_id": "user34@mail.com", "posted_on_date": {"day": 4, "month": 2, "year": 2023, "time": "00:00:00"}, {"comment_text": "I disagree with the viewpoint.", "score": -3}]}], "acknowledged": true, "insertedId": ObjectId('67fc7725b76bea4836b1429')}
```

```
dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Space", "article_id": 35, "post_on_date": {"day": 5, "month": 2, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 35", "similar_stories": [{"https://thedaily.com/politics/space-story-35.html", "https://thedaily.com/sports/space-story-35.html"}, {"num_times_read": 14888, "article_text": "This article discusses the latest developments in space.", "comments": [{"comment_id": 35, "article_id": 35, "user_id": "user35@mail.com", "posted_on_date": {"day": 5, "month": 2, "year": 2023, "time": "00:00:00"}, {"comment_text": "Bold take on current affairs.", "score": 99}]}], "acknowledged": true, "insertedId": ObjectId('67fc7725b76bea4836b1429')}
```

```
dailyNewsArchive> db.articles.insertOne({article_type: "feature article", "article_section": "Entertainment", "article_id": 36, "post_on_date": {"day": 6, "month": 2, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 36", "similar_stories": [{"https://thedaily.com/politics/election-story-36.html", "https://thedaily.com/sports/election-story-36.html"}, {"num_times_read": 31286, "article_text": "This article discusses the latest developments in election.", "comments": [{"comment_id": 36, "article_id": 36, "user_id": "user36@mail.com", "posted_on_date": {"day": 6, "month": 2, "year": 2023, "time": "00:00:00"}, {"comment_text": "Bold take on current affairs.", "score": -67}]}], "acknowledged": true, "insertedId": ObjectId('67fc7725b76bea4836b1429')}
```

```
dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Sports", "article_id": 37, "post_on_date": {"day": 7, "month": 2, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 37", "similar_stories": [{"https://thedaily.com/politics/ai-story-37.html", "https://thedaily.com/sports/ai-story-37.html"}, {"num_times_read": 16879, "article_text": "This article discusses the latest developments in AI.", "comments": [{"comment_id": 37, "article_id": 37, "user_id": "user37@mail.com", "posted_on_date": {"day": 7, "month": 2, "year": 2023, "time": "00:00:00"}, {"comment_text": "Needs better research.", "score": -30}]}], "acknowledged": true, "insertedId": ObjectId('67fc7725b76bea4836b1429')}
```

```
dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Entertainment", "article_id": 38, "post_on_date": {"day": 8, "month": 2, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 38", "similar_stories": [{"https://thedaily.com/politics/budget-story-38.html", "https://thedaily.com/sports/budget-story-38.html"}, {"num_times_read": 1259, "article_text": "This article discusses the latest developments in budget.", "comments": [{"comment_id": 38, "article_id": 38, "user_id": "user38@mail.com", "posted_on_date": {"day": 8, "month": 2, "year": 2023, "time": "00:00:00"}, {"comment_text": "Amazing insight!", "score": -14}]}], "acknowledged": true, "insertedId": ObjectId('67fc7725b76bea4836b1429')}
```

```
dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Entertainment", "article_id": 39, "post_on_date": {"day": 9, "month": 2, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 39", "similar_stories": [{"https://thedaily.com/politics/climate-story-39.html", "https://thedaily.com/sports/climate-story-39.html"}, {"num_times_read": 16636, "article_text": "This article discusses the latest developments in climate.", "comments": [{"comment_id": 39, "article_id": 39, "user_id": "user39@mail.com", "posted_on_date": {"day": 9, "month": 2, "year": 2023, "time": "00:00:00"}, {"comment_text": "I disagree with the viewpoint.", "score": 40}]}], "acknowledged": true, "insertedId": ObjectId('67fc7725b76bea4836b1429')}
```

```
dailyNewsArchive> db.articles.insertOne({article_type: "feature article", "article_section": "Health", "article_id": 40, "post_on_date": {"day": 10, "month": 2, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter 40", "similar_stories": [{"https://thedaily.com/politics/cricket-story-40.html", "https://thedaily.com/sports/cricket-story-40.html"}, {"num_times_read": 468, "article_text": "This article discusses the latest developments in cricket.", "comments": [{"comment_id": 40, "article_id": 40, "user_id": "user40@mail.com", "posted_on_date": {"day": 10, "month": 2, "year": 2023, "time": "00:00:00"}, {"comment_text": "Not convinced with the facts.", "score": -50}]}], "acknowledged": true, "insertedId": ObjectId('67fc7725b76bea4836b1429')}
```

```

dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Technology", "article_id": 41, "post_on_date": {"day": 11, "month": 2, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter A1", "similar_stories": ["https://thedaily.com/politics/file-story-41.html", "https://thedaily.com/sports/file-story-41.html"], "num_times_read": 4038, "article_text": "This article discusses the latest developments in film.", "comments": [{"comment_id": 41, "article_id": 41, "user_id": "user49@mail.com", "posted_on_date": {"day": 11, "month": 2, "year": 2023, "time": "00:00:00"}, {"comment_text": "Well-written and concise.", "score": 38}]}];
{
  acknowledged: true,
  insertedId: ObjectId('67fc772b0b76bea4836b1433')
}
dailyNewsArchive> db.articles.insertOne({article_type: "feature article", "article_section": "Sports", "article_id": 42, "post_on_date": {"day": 12, "month": 2, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter A2", "similar_stories": ["https://thedaily.com/politics/education-story-42.html", "https://thedaily.com/sports/education-story-42.html"], "num_times_read": 16386, "article_text": "This article discusses the latest developments in education.", "comments": [{"comment_id": 42, "article_id": 42, "user_id": "user49@mail.com", "posted_on_date": {"day": 12, "month": 2, "year": 2023, "time": "00:00:00"}, {"comment_text": "Well-written and concise.", "score": 73}]}];
{
  acknowledged: true,
  insertedId: ObjectId('67fc772b0b76bea4836b1434')
}
dailyNewsArchive> db.articles.insertOne({article_type: "feature article", "article_section": "Politics", "article_id": 43, "post_on_date": {"day": 13, "month": 2, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter A3", "similar_stories": ["https://thedaily.com/politics/ai-story-43.html", "https://thedaily.com/sports/ai-story-43.html"], "num_times_read": 15403, "article_text": "This article discusses the latest developments in AI.", "comments": [{"comment_id": 43, "article_id": 43, "user_id": "user49@mail.com", "posted_on_date": {"day": 13, "month": 2, "year": 2023, "time": "00:00:00"}, {"comment_text": "Looking forward to more.", "score": 93}]}];
{
  acknowledged: true,
  insertedId: ObjectId('67fc772b0b76bea4836b1435')
}
dailyNewsArchive> db.articles.insertOne({article_type: "feature article", "article_section": "Entertainment", "article_id": 44, "post_on_date": {"day": 14, "month": 2, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter A4", "similar_stories": ["https://thedaily.com/politics/election-story-44.html", "https://thedaily.com/sports/election-story-44.html"], "num_times_read": 19983, "article_text": "This article discusses the latest developments in election.", "comments": [{"comment_id": 44, "article_id": 44, "user_id": "user49@mail.com", "posted_on_date": {"day": 14, "month": 2, "year": 2023, "time": "00:00:00"}, {"comment_text": "Amazing insights!", "score": -45}]}];
{
  acknowledged: true,
  insertedId: ObjectId('67fc772b0b76bea4836b1436')
}
dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Technology", "article_id": 45, "post_on_date": {"day": 15, "month": 2, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter A5", "similar_stories": ["https://thedaily.com/politics/file-story-45.html", "https://thedaily.com/sports/file-story-45.html"], "num_times_read": 6405, "article_text": "This article discusses the latest developments in file.", "comments": [{"comment_id": 45, "article_id": 45, "user_id": "user49@mail.com", "posted_on_date": {"day": 15, "month": 2, "year": 2023, "time": "00:00:00"}, {"comment_text": "Looking forward to more.", "score": 26}]}];
{
  acknowledged: true,
  insertedId: ObjectId('67fc772b0b76bea4836b1437')
}
dailyNewsArchive> db.articles.insertOne({article_type: "feature article", "article_section": "Politics", "article_id": 46, "post_on_date": {"day": 16, "month": 2, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter A6", "similar_stories": ["https://thedaily.com/politics/covid-story-46.html", "https://thedaily.com/sports/covid-story-46.html"], "num_times_read": 19132, "article_text": "This article discusses the latest developments in covid.", "comments": [{"comment_id": 46, "article_id": 46, "user_id": "user49@mail.com", "posted_on_date": {"day": 16, "month": 2, "year": 2023, "time": "00:00:00"}, {"comment_text": "Needs better research.", "score": 29}]}];
{
  acknowledged: true,
  insertedId: ObjectId('67fc772b0b76bea4836b1438')
}
dailyNewsArchive> db.articles.insertOne({article_type: "feature article", "article_section": "Entertainment", "article_id": 47, "post_on_date": {"day": 17, "month": 2, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter A7", "similar_stories": ["https://thedaily.com/politics/climate-story-47.html", "https://thedaily.com/sports/climate-story-47.html"], "num_times_read": 13623, "article_text": "This article discusses the latest developments in climate.", "comments": [{"comment_id": 47, "article_id": 47, "user_id": "user49@mail.com", "posted_on_date": {"day": 17, "month": 2, "year": 2023, "time": "00:00:00"}, {"comment_text": "Fantastic read!", "score": 19}]}];
{
  acknowledged: true,
  insertedId: ObjectId('67fc772b0b76bea4836b1439')
}
dailyNewsArchive> db.articles.insertOne({article_type: "feature article", "article_section": "Sports", "article_id": 48, "post_on_date": {"day": 18, "month": 2, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter A8", "similar_stories": ["https://thedaily.com/politics/space-story-48.html", "https://thedaily.com/sports/space-story-48.html"], "num_times_read": 17121, "article_text": "This article discusses the latest developments in space.", "comments": [{"comment_id": 48, "article_id": 48, "user_id": "user49@mail.com", "posted_on_date": {"day": 18, "month": 2, "year": 2023, "time": "00:00:00"}, {"comment_text": "Fantastic read!", "score": -17}]}];
{
  acknowledged: true,
  insertedId: ObjectId('67fc772b0b76bea4836b143a')
}

```

```

dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Entertainment", "article_id": 49, "post_on_date": {"day": 19, "month": 2, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter A9", "similar_stories": ["https://thedaily.com/politics/cricket-story-49.html", "https://thedaily.com/sports/cricket-story-49.html"], "num_times_read": 7956, "article_text": "This article discusses the latest developments in cricket.", "comments": [{"comment_id": 49, "article_id": 49, "user_id": "user49@mail.com", "posted_on_date": {"day": 19, "month": 2, "year": 2023, "time": "00:00:00"}, {"comment_text": "Not convinced with the facts.", "score": -61}]}];
{
  acknowledged: true,
  insertedId: ObjectId('67fc772b0b76bea4836b143b')
}
dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Politics", "article_id": 50, "post_on_date": {"day": 20, "month": 2, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter A10", "similar_stories": ["https://thedaily.com/politics/education-story-50.html", "https://thedaily.com/sports/education-story-50.html"], "num_times_read": 1380, "article_text": "This article discusses the latest developments in education.", "comments": [{"comment_id": 50, "article_id": 50, "user_id": "user50@mail.com", "posted_on_date": {"day": 20, "month": 2, "year": 2023, "time": "00:00:00"}, {"comment_text": "Fantastic read!", "score": -77}]}];
{
  acknowledged: true,
  insertedId: ObjectId('67fc772b0b76bea4836b143c')
}
dailyNewsArchive> db.articles.insertOne({article_type: "feature article", "article_section": "Sports", "article_id": 51, "post_on_date": {"day": 21, "month": 2, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter A11", "similar_stories": ["https://thedaily.com/politics/covid-story-51.html", "https://thedaily.com/sports/covid-story-51.html"], "num_times_read": 18628, "article_text": "This article discusses the latest developments in covid.", "comments": [{"comment_id": 51, "article_id": 51, "user_id": "user51@mail.com", "posted_on_date": {"day": 21, "month": 2, "year": 2023, "time": "00:00:00"}, {"comment_text": "Not convinced with the facts.", "score": 19}]}];
{
  acknowledged: true,
  insertedId: ObjectId('67fc772b0b76bea4836b143d')
}
dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Health", "article_id": 52, "post_on_date": {"day": 22, "month": 2, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter A12", "similar_stories": ["https://thedaily.com/politics/budget-story-52.html", "https://thedaily.com/sports/budget-story-52.html"], "num_times_read": 2739, "article_text": "This article discusses the latest developments in budget.", "comments": [{"comment_id": 52, "article_id": 52, "user_id": "user52@mail.com", "posted_on_date": {"day": 22, "month": 2, "year": 2023, "time": "00:00:00"}, {"comment_text": "Not convinced with the facts.", "score": -23}]}];
{
  acknowledged: true,
  insertedId: ObjectId('67fc772b0b76bea4836b143e')
}
dailyNewsArchive> db.articles.insertOne({article_type: "feature article", "article_section": "Politics", "article_id": 53, "post_on_date": {"day": 23, "month": 2, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter A13", "similar_stories": ["https://thedaily.com/politics/cricket-story-53.html", "https://thedaily.com/sports/cricket-story-53.html"], "num_times_read": 6215, "article_text": "This article discusses the latest developments in cricket.", "comments": [{"comment_id": 53, "article_id": 53, "user_id": "user53@mail.com", "posted_on_date": {"day": 23, "month": 2, "year": 2023, "time": "00:00:00"}, {"comment_text": "Needs better research.", "score": -93}]}];
{
  acknowledged: true,
  insertedId: ObjectId('67fc772b0b76bea4836b143f')
}
dailyNewsArchive> db.articles.insertOne({article_type: "news article", "article_section": "Health", "article_id": 54, "post_on_date": {"day": 24, "month": 2, "year": 2023, "time": "00:00:00"}, "reporter_name": "Reporter A14", "similar_stories": ["https://thedaily.com/politics/ai-story-54.html", "https://thedaily.com/sports/ai-story-54.html"], "num_times_read": 8205, "article_text": "This article uses the latest developments in AI.", "comments": [{"comment_id": 54, "article_id": 54, "user_id": "user54@mail.com", "posted_on_date": {"day": 24, "month": 2, "year": 2023, "time": "00:00:00"}, {"comment_text": "Fantastic read!", "score": 13}]}];
{
  acknowledged: true,
  insertedId: ObjectId('67fc772b0b76bea4836b1440')
}

```

```

test> use dailyNewsArchive
switched to db dailyNewsArchive
dailyNewsArchive> db.articles.countDocuments()
55
dailyNewsArchive>

```

## Task (d) Index Creation:

**Query 1:** Insert a (news or feature) article into the archive

// No Index Needed

**Justification:** Insertion operations do not benefit from indexes unless there are constraints (like uniqueness), which is not the case here. Adding indexes would slow down insertion due to index maintenance overhead. Since articles have unique IDs but no uniqueness constraint enforced here, no index is needed specifically for insertion.

**Query 2:** Retrieve the latest news

```
article_id_1
dailyNewsArchive> db.articles.createIndex({
...   "post_on_date.year": -1,
...   "post_on_date.month": -1,
...   "post_on_date.day": -1,
...   "post_on_date.time": -1
... })
post_on_date.year_-1_post_on_date.month_-1_post_on_date.day_-1_post_on_date.time_-1
```

**Justification:** To efficiently retrieve the most recent article, a compound index on the full date and time fields in descending order is necessary. This allows MongoDB to quickly sort documents by publication date and return the latest one without scanning the entire collection.

**Query 3:** Retrieve the top 10 most read news in the past month

```
dailyNewsArchive> db.articles.createIndex({
...   "post_on_date.year": -1,
...   "post_on_date.month": -1,
...   "post_on_date.day": -1,
...   "num_times_read": -1
... })
post_on_date.year_-1_post_on_date.month_-1_post_on_date.day_-1_num_times_read_-1
```

**Justification:** This query requires filtering by date (to restrict to the past month) and sorting by num\_times\_read to get the most-read articles. A compound index starting with date fields followed by num\_times\_read supports both the filter and sort operations efficiently.

#### Query 4: Add a comment to an article

```
db.articles.createIndex({ article_id: 1 });
```

**Justification:** To update or add comments to a specific article, the article\_id is used to locate the target document. Indexing this field ensures quick access to the correct article without scanning all documents.

#### Query 5: Add a story to the ‘similar stories’ section for articles containing a given word and published within the past 2 years

```
db.articles.createIndex({ article_text: "text" });
db.articles.createIndex({
  "post_on_date.year": -1,
  "post_on_date.month": -1,
  "post_on_date.day": -1
});
```

##### Justification:

**Text index (article\_text):** Enables efficient full-text search for keywords across all articles. This is essential to match articles containing the given word or phrase.

**Date index:** Filters articles published in the past 2 years. A descending index on year, month, and day improves performance for time-based range queries, especially when combined with the text search.

## Task (e) Frequently Used Queries:

Query 1: Insert a (news or feature) article into the archive

```
dailyNewsArchive> db.articles.insertOne({
...   article_type: "news article",
...   article_section: "Technology",
...   article_id: 999,
...   post_on_date: { day: 12, month: 4, year: 2025, time: "12:00:00" },
...   reporter_name: "MS Dhoni",
...   similar_stories: [],
...   num_times_read: 0,
...   article_text: "A brand new tech breakthrough in quantum computing.",
...   comments: []
... })
{
  acknowledged: true,
  insertedId: ObjectId('67fc7f6395be239aba6b140b')
}
dailyNewsArchive> db.articles.find({ article_id: 999 }).pretty()
[
  {
    _id: ObjectId('67fc7f6395be239aba6b140b'),
    article_type: 'news article',
    article_section: 'Technology',
    article_id: 999,
    post_on_date: { day: 12, month: 4, year: 2025, time: '12:00:00' },
    reporter_name: 'MS Dhoni',
    similar_stories: [],
    num_times_read: 0,
    article_text: 'A brand new tech breakthrough in quantum computing.',
    comments: []
  }
]
dailyNewsArchive> █
```

Query 2: Retrieve the latest news

```
dailyNewsArchive> db.articles.find({ article_type: "news article" })
...   .sort({ "post_on_date.year": -1, "post_on_date.month": -1, "post_on_date.day": -1, "post_on_date.time": -1 })
...   .limit(1)
[...   .pretty()
[
  {
    _id: ObjectId('67fc7f6395be239aba6b140b'),
    article_type: 'news article',
    article_section: 'Technology',
    article_id: 999,
    post_on_date: { day: 12, month: 4, year: 2025, time: '12:00:00' },
    reporter_name: 'MS Dhoni',
    similar_stories: [],
    num_times_read: 0,
    article_text: 'A brand new tech breakthrough in quantum computing.',
    comments: []
  }
]
dailyNewsArchive> █
```

### Query 3: Retrieve the top 10 most read news in the past month

```
dailyNewsArchive> db.articles.find({
...   article_type: "news article",
...   $or: [
...     { "post_on_date.year": 2023, "post_on_date.month": 2 },
...     { "post_on_date.year": 2023, "post_on_date.month": 1, "post_on_date.day": { $gte: 12 } }
...   ]
... })
... .sort({ num_times_read: -1 })
... .limit(10)
... .pretty()
```

```
[
  {
    _id: ObjectId('67fc772b5b76bea4836b142f'),
    article_type: 'news article',
    article_section: 'Entertainment',
    article_id: 37,
    post_on_date: { day: 7, month: 2, year: 2023, time: '00:00:00' },
    reporter_name: 'Reporter 37',
    similar_stories: [
      'https://thedaily.com/politics/AI-story-37.html',
      'https://thedaily.com/sports/AI-story-37.html'
    ],
    num_times_read: 18879,
    article_text: 'This article discusses the latest developments in AI.',
    comments: [
      {
        comment_id: 37,
        article_id: 37,
        user_id: 'user37@mail.com',
        posted_on_date: { day: 7, month: 2, year: 2023, time: '00:00:00' },
        comment_text: 'Needs better research.',
        score: -30
      }
    ]
  },
  {
    _id: ObjectId('67fc77265b76bea4836b141a'),
    article_type: 'news article',
    article_section: 'Health',
    article_id: 16,
    post_on_date: { day: 17, month: 1, year: 2023, time: '00:00:00' },
    reporter_name: 'Reporter 16',
    similar_stories: [
      'https://thedaily.com/politics/budget-story-16.html',
      'https://thedaily.com/sports/budget-story-16.html'
    ],
    num_times_read: 16892,
    article_text: 'This article discusses the latest developments in budget.',
    comments: [
      {
        comment_id: 16,
        article_id: 16,
        user_id: 'user16@mail.com',
        posted_on_date: { day: 17, month: 1, year: 2023, time: '00:00:00' },
        comment_text: 'I disagree with the viewpoint.',
        score: 99
      }
    ]
  }
],
```

```
{
  _id: ObjectId('67fc772b5b76bea4836b1431'),
  article_type: 'news article',
  article_section: 'Entertainment',
  article_id: 39,
  post_on_date: { day: 9, month: 2, year: 2023, time: '00:00:00' },
  reporter_name: 'Reporter 39',
  similar_stories: [
    'https://thedaily.com/politics/climate-story-39.html',
    'https://thedaily.com/sports/climate-story-39.html'
  ],
  num_times_read: 15636,
  article_text: 'This article discusses the latest developments in climate.',
  comments: [
    {
      comment_id: 39,
      article_id: 39,
      user_id: 'user39@mail.com',
      posted_on_date: { day: 9, month: 2, year: 2023, time: '00:00:00' },
      comment_text: 'I disagree with the viewpoint.',
      score: 46
    }
  ]
},
{
  _id: ObjectId('67fc77295b76bea4836b1429'),
  article_type: 'news article',
  article_section: 'Sports',
  article_id: 31,
  post_on_date: { day: 1, month: 2, year: 2023, time: '00:00:00' },
  reporter_name: 'Reporter 31',
  similar_stories: [
    'https://thedaily.com/politics/film-story-31.html',
    'https://thedaily.com/sports/film-story-31.html'
  ],
  num_times_read: 12499,
  article_text: 'This article discusses the latest developments in film.',
  comments: [
    {
      comment_id: 31,
      article_id: 31,
      user_id: 'user31@mail.com',
      posted_on_date: { day: 1, month: 2, year: 2023, time: '00:00:00' },
      comment_text: 'Not convinced with the facts.',
      score: -30
    }
  ]
}
```

```
{
  _id: ObjectId('67fc772a5b76bea4836b142e'),
  article_type: 'news article',
  article_section: 'Sports',
  article_id: 36,
  post_on_date: { day: 6, month: 2, year: 2023, time: '00:00:00' },
  reporter_name: 'Reporter 36',
  similar_stories: [
    'https://thedaily.com/politics/election-story-36.html',
    'https://thedaily.com/sports/election-story-36.html'
  ],
  num_times_read: 11286,
  article_text: 'This article discusses the latest developments in election.',
  comments: [
    {
      comment_id: 36,
      article_id: 36,
      user_id: 'user36@mail.com',
      posted_on_date: { day: 6, month: 2, year: 2023, time: '00:00:00' },
      comment_text: 'Bold take on current affairs.',
      score: -57
    }
  ]
},
{
  _id: ObjectId('67fc77285b76bea4836b1423'),
  article_type: 'news article',
  article_section: 'Entertainment',
  article_id: 25,
  post_on_date: { day: 26, month: 1, year: 2023, time: '00:00:00' },
  reporter_name: 'Reporter 25',
  similar_stories: [
    'https://thedaily.com/politics/election-story-25.html',
    'https://thedaily.com/sports/election-story-25.html'
  ],
  num_times_read: 9833,
  article_text: 'This article discusses the latest developments in election.',
  comments: [
    {
      comment_id: 25,
      article_id: 25,
      user_id: 'user25@mail.com',
      posted_on_date: { day: 26, month: 1, year: 2023, time: '00:00:00' },
      comment_text: 'Bold take on current affairs.',
      score: 86
    }
  ]
},
```

```
_id: ObjectId('67fc77285b76bea4836b1421'),
article_type: 'news article',
article_section: 'Technology',
article_id: 23,
post_on_date: { day: 24, month: 1, year: 2023, time: '00:00:00' },
reporter_name: 'Reporter 23',
similar_stories: [
  'https://thedaily.com/politics/AI-story-23.html',
  'https://thedaily.com/sports/AI-story-23.html'
],
num_times_read: 7960,
article_text: 'This article discusses the latest developments in AI.',
comments: [
  {
    comment_id: 23,
    article_id: 23,
    user_id: 'user23@mail.com',
    posted_on_date: { day: 24, month: 1, year: 2023, time: '00:00:00' },
    comment_text: 'Needs better research.',
    score: -68
  }
],
},
{
  _id: ObjectId('67fc772d5b76bea4836b143b'),
  article_type: 'news article',
  article_section: 'Entertainment',
  article_id: 49,
  post_on_date: { day: 19, month: 2, year: 2023, time: '00:00:00' },
  reporter_name: 'Reporter 49',
  similar_stories: [
    'https://thedaily.com/politics/cricket-story-49.html',
    'https://thedaily.com/sports/cricket-story-49.html'
  ],
  num_times_read: 7955,
  article_text: 'This article discusses the latest developments in cricket.',
  comments: [
    {
      comment_id: 49,
      article_id: 49,
      user_id: 'user49@mail.com',
      posted_on_date: { day: 19, month: 2, year: 2023, time: '00:00:00' },
      comment_text: 'Not convinced with the facts.',
      score: -51
    }
  ]
},
```

```
_id: ObjectId('67fc77295b76bea4836b1426'),
article_type: 'news article',
article_section: 'Entertainment',
article_id: 28,
post_on_date: { day: 29, month: 1, year: 2023, time: '00:00:00' },
reporter_name: 'Reporter 28',
similar_stories: [
  'https://thedaily.com/politics/film-story-28.html',
  'https://thedaily.com/sports/film-story-28.html'
],
num_times_read: 7559,
article_text: 'This article discusses the latest developments in film.',
comments: [
  {
    comment_id: 28,
    article_id: 28,
    user_id: 'user28@mail.com',
    posted_on_date: { day: 29, month: 1, year: 2023, time: '00:00:00' },
    comment_text: 'Keep up the good work!',
    score: -32
  }
],
{
  _id: ObjectId('67fc772c5b76bea4836b1437'),
  article_type: 'news article',
  article_section: 'Technology',
  article_id: 45,
  post_on_date: { day: 15, month: 2, year: 2023, time: '00:00:00' },
  reporter_name: 'Reporter 45',
  similar_stories: [
    'https://thedaily.com/politics/film-story-45.html',
    'https://thedaily.com/sports/film-story-45.html'
  ],
  num_times_read: 6605,
  article_text: 'This article discusses the latest developments in film.',
  comments: [
    {
      comment_id: 45,
      article_id: 45,
      user_id: 'user45@mail.com',
      posted_on_date: { day: 15, month: 2, year: 2023, time: '00:00:00' },
      comment_text: 'Looking forward to more.',
      score: 25
    }
  ]
}
```

#### Query 4: Add a comment to an article

```
dailyNewsArchive> db.articles.updateOne(  
...   { article_id: 999 },  
...   { $push: {  
...     comments: {  
...       comment_id: 123,  
...       article_id: 999,  
...       user_id: "newuser@mail.com",  
...       posted_on_date: { day: 12, month: 4, year: 2025, time: "13:00:00" },  
...       comment_text: "This is a new comment for testing.",  
...       score: 10  
...     }  
...   }  
... )  
{  
  acknowledged: true,  
  insertedId: null,  
  matchedCount: 1,  
  modifiedCount: 1,  
  upsertedCount: 0  
}
```

```
dailyNewsArchive> db.articles.find({ article_id: 999 }, { comments: 1, _id: 0 }).pretty()  
[  
  {  
    comments: [  
      {  
        comment_id: 123,  
        article_id: 999,  
        user_id: 'newuser@mail.com',  
        posted_on_date: { day: 12, month: 4, year: 2025, time: '13:00:00' },  
        comment_text: 'This is a new comment for testing.',  
        score: 10  
      }  
    ]  
  }  
]  
dailyNewsArchive>
```

**Query 5:** Add a story to the ‘similar stories’ section for articles containing a given word and published within the past 2 years.

```
dailyNewsArchive> db.articles.updateMany(  
...   {  
...     $text: { $search: "quantum" },  
...     $or: [  
...       { "post_on_date.year": { $gt: 2023 } },  
...       { "post_on_date.year": 2023, "post_on_date.month": { $gt: 4 } },  
...       { "post_on_date.year": 2023, "post_on_date.month": 4, "post_on_date.day": { $gte: 12 } }  
...     ]  
...   },  
...   {  
...     $push:  
...       {  
...         similar_stories: "https://thedaily.com/science/quantum-research-breakthrough.html"  
...       }  
...   }  
... )  
{  
  acknowledged: true,  
  insertedId: null,  
  matchedCount: 1,  
  modifiedCount: 1,  
  upsertedCount: 0  
}
```

```
dailyNewsArchive> db.articles.find(  
...   { article_id: 999 },  
...   { article_text: 1, similar_stories: 1, _id: 0 }  
... ).pretty()  
[  
  {  
    similar_stories: [  
      'https://thedaily.com/science/quantum-research-breakthrough.html'  
    ],  
    article_text: 'A brand new tech breakthrough in quantum computing.'  
  }  
]  
dailyNewsArchive> █
```

## Task (f): Schema Update for Reporter Details

```
dailyNewsArchive> db.articles.insertOne({
...   article_type: "news article",
...   article_section: "Health",
...   article_id: 1001,
...   post_on_date: { day: 12, month: 4, year: 2025, time: "14:00:00" },
...   reporters: [
...     { name: "MS Dhoni", location: "Chennai" },
...     { name: "Suresh Raina", location: "Ghaziabad" }
...   ],
...   similar_stories: [],
...   num_times_read: 1500,
...   article_text: "This article discusses healthcare AI in India.",
...   comments: []
... })
{
  acknowledged: true,
  insertedId: ObjectId('67fc87b295be239aba6b140c')
}
```

```
dailyNewsArchive> db.articles.find({ article_id: 1001 }).pretty()
[
  {
    _id: ObjectId('67fc87b295be239aba6b140c'),
    article_type: 'news article',
    article_section: 'Health',
    article_id: 1001,
    post_on_date: { day: 12, month: 4, year: 2025, time: '14:00:00' },
    reporters: [
      { name: 'MS Dhoni', location: 'Chennai' },
      { name: 'Suresh Raina', location: 'Ghaziabad' }
    ],
    similar_stories: [],
    num_times_read: 1500,
    article_text: 'This article discusses healthcare AI in India.',
    comments: []
  }
]
```

```
dailyNewsArchive> db.articles.insertOne({
...   article_type: "feature article",
...   article_section: "Technology",
...   article_id: 1002,
...   post_on_date: { day: 12, month: 4, year: 2025, time: "15:00:00" },
...   reporters: [
...     { name: "Virat Kohli", location: "Delhi" },
...     { name: "Rohit Sharma", location: "Mumbai" },
...     { name: "KL Rahul", location: "Bangalore" },
...     { name: "MS Dhoni", location: "Chennai" }
...   ],
...   similar_stories: [],
...   num_times_read: 2600,
...   article_text: "This article explores emerging cloud platforms.",
...   comments: []
... })
{
  acknowledged: true,
  insertedId: ObjectId('67fc87f295be239aba6b140d')
}
```

```
, dailyNewsArchive> db.articles.find({ article_id: 1002 }).pretty()
[
  {
    _id: ObjectId('67fc87f295be239aba6b140d'),
    article_type: 'feature article',
    article_section: 'Technology',
    article_id: 1002,
    post_on_date: { day: 12, month: 4, year: 2025, time: '15:00:00' },
    reporters: [
      { name: 'Virat Kohli', location: 'Delhi' },
      { name: 'Rohit Sharma', location: 'Mumbai' },
      { name: 'KL Rahul', location: 'Bangalore' },
      { name: 'MS Dhoni', location: 'Chennai' }
    ],
    similar_stories: [],
    num_times_read: 2600,
    article_text: 'This article explores emerging cloud platforms.',
    comments: []
  }
]
```

## Task (g): Replication Setup

The replication scheme is implemented as a MongoDB replica set named "theDailyReplica" with three nodes, each hosted on a different GPEL machine:

- Primary Node: gpel9.cs.nor.ou.edu:24157
- Secondary Node: gpel10.cs.nor.ou.edu:24157
- Secondary Node: gpel11.cs.nor.ou.edu:24157

The replica set was initialized using the following configuration:

```
Last login: Sun Apr 13 21:56:07 on ttys001
(base) ravisribhashyan@Sasank-Macbook-2 ~ % ssh srib0000@gpel10.cs.nor.ou.edu
srib0000@gpel10.cs.nor.ou.edu's password:
Welcome to Ubuntu 24.04.1 LTS (GNU/Linux 6.11.0-19-generic x86_64)

 * Documentation: https://help.ubuntu.com
 * Management: https://landscape.canonical.com
 * Support: https://ubuntu.com/pro

1 device has a firmware upgrade available.
Run 'fwupdmgr get-upgrades' for more information.

Expanded Security Maintenance for Applications is not enabled.

259 updates can be applied immediately.
To see these additional updates run: apt list --upgradable

16 additional security updates can be applied with ESM Apps.
Learn more about enabling ESM Apps service at https://ubuntu.com/esm

*** System restart required ***

1 device has a firmware upgrade available.
Run 'fwupdmgr get-upgrades' for more information.

Last login: Sun Apr 13 20:01:11 2025 from 10.199.192.103
srib0000@gpel10:~ % mkdir -p ~/mongodb-data
srib0000@gpel10:~ % mongod --replSet theDailyReplica --port 24157 --dbpath ~/mongodb-data --bind_ip_all --logpath ~/mongo.log

ps -ef | grep mongod
*Caribou0000@gpel10:~ % ps -ef | grep mongod
mongodb 2464 1 0 Mar26 ? 00:19:22 /usr/bin/mongod --config /etc/mongod.conf
dnsmsnq 3978 3795 0 Mar26 ? 00:15:16 mongod --repSet overleaf --bind_ip_all
ared0000 465095 463969 1 21:24 pts/3 00:01:20 mongod --dbpath /home/ared0000/mongodata --port 22526 --repSet rs0
adda0003 506661 1 1 21:52 ? 00:01:14 mongod --repSet dailyReplicaSet --dbpath /home/adda0003/mongo-repl/data --port 29372 --bind_ip_all --fork --logpath /home/adda0003/mongo-repl/mongod.log
rep00000 553097 1 1 22:21 ? 00:00:47 mongod --fork --dbpath /home/rep0000/data/gpel10/ --port 22986 --logpath mongod.log --repSet my0ReplicaSet --bind_ip_all
ared0000 568404 467902 0 22:29 pts/4 00:00:04 mongosh mongod://12
mail0019 565854 463974 1 22:46 ? 00:00:33 mongod --fork --dbpath /home/mail0019/data/gpel10/ --port 23245 --logpath mongod.log --repSet my0ReplicaSet --bind_ip_all
adda0004 637937 1 2 23:01 ? 00:00:12 mongod --fork --dbpath /home/adda0004/data/gpel10/ --port 28294 --logpath /home/adda0004/data/gpel10/mongod.log --repSet gpelRep1Set --bind_ip_all
bodd0008 3929866 3587966 32 23:22 pts/5 00:00:46 mongod --repSet rs0 --port 20008 --dbpath /home/bodd0008/adms/hw3/db --bind_ip localhost,gpel10.cs.nor.ou.edu
taus0001 3938931 1 1 13:59 ? 00:07:29 mongod --fork --port 31005 --repSet ktreplset --logpath /home/taus0001/mongo.log --dbpath /home/taus0001/data/mongo --bind_ip_all
srib0000@gpel10:~ % mongod --repSet theDailyReplica --port 24157 --dbpath ~/mongodb-data --bind_ip_all --fork --logpath ~/mongo.log
about to fork child process, waiting until server is ready for connections.
forked process: 658088
child process started successfully, parent exiting
srib0000@gpel10:~ %
```

```
Last login: Sun Apr 13 23:17:22 on ttys003
(base) ravisribhashyan@Sasank-Macbook-2 ~ % ssh srib0000@gpel11.cs.nor.ou.edu
srib0000@gpel11.cs.nor.ou.edu's password:
Welcome to Ubuntu 24.04.1 LTS (GNU/Linux 6.8.0-49-generic x86_64)

 * Documentation: https://help.ubuntu.com
 * Management: https://landscape.canonical.com
 * Support: https://ubuntu.com/pro

1 device has a firmware upgrade available.
Run 'fwupdmgr get-upgrades' for more information.

Expanded Security Maintenance for Applications is not enabled.

263 updates can be applied immediately.
To see these additional updates run: apt list --upgradable

16 additional security updates can be applied with ESM Apps.
Learn more about enabling ESM Apps service at https://ubuntu.com/esm

*** System restart required ***

1 device has a firmware upgrade available.
Run 'fwupdmgr get-upgrades' for more information.

Last login: Sun Apr 13 20:05:15 2025 from 10.199.192.103
srib0000@gpel11:~ % mkdir -p ~/mongodb-data
srib0000@gpel11:~ % mongod --replSet theDailyReplica --port 24157 --dbpath ~/mongodb-data --bind_ip_all --logpath ~/mongo.log
*Caribou0000@gpel11:~ % ps -ef | grep mongod
taus0001 3564470 1 1 13:59 ? 00:09:38 mongod --fork --port 31005 --repSet ktreplset --logpath /home/taus0001/mongo.log --dbpath /home/taus0001/data/mongo --bind_ip_all
rep00000 3585878 1 1 22:21 ? 00:01:08 mongod --fork --dbpath /home/rep0000/data/gpel11/ --port 22989 --logpath mongod.log --repSet my0ReplicaSet --bind_ip_all
mail0019 3590400 1 1 22:46 ? 00:00:44 mongod --fork --dbpath /home/mail0019/data/gpel11/ --port 23246 --logpath mongod.log --repSet my0ReplicaSet --bind_ip_all
adda0004 35904483 1 2 23:17 ? 00:00:07 mongod --fork --dbpath /home/adda0004/data/gpel11/ --port 28294 --logpath /home/adda0004/data/gpel11/mongod.log --repSet gpelRep1Set --bind_ip_all
bodd0008 3592866 3587966 32 23:22 pts/5 00:00:46 mongod --repSet rs0 --port 20008 --dbpath /home/bodd0008/adms/hw3/db --bind_ip_all
srib00000 3593587 3592342 0 23:24 pts/7 00:00:00 grep --color=auto mongod
srib0000@gpel11:~ % mongod --repSet theDailyReplica --port 24157 --dbpath ~/mongodb-data --bind_ip_all --fork --logpath ~/mongo.log
about to fork child process, waiting until server is ready for connections.
forked process: 3593783
child process started successfully, parent exiting
srib0000@gpel11:~ %
```

```

$ mongosh --port 24157
Current Mongosh Log ID: 67fc51c71321398db6b140s
Connecting to: mongodb://127.0.0.1:24157/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+2.4.2
Using MongoDB: 8.0.0
Using Mongosh: 2.4.2
mongosh 2.8.0 is available for download: https://www.mongodb.com/try/download/shell
For mongosh info see: https://www.mongodb.com/docs/mongodb-shell/

The server generated these startup warnings when booting
2025-04-13T23:39:02.814-05:00: Using the XFS filesystem is strongly recommended with the WiredTiger storage engine. See http://dochub.mongodb.org/core/prodnotes-filesystem
2025-04-13T23:39:02.826-05:00: Access control is not enabled for the database. Read and write access to data and configuration is unrestricted
2025-04-13T23:39:02.826-05:00: Soft rlimits for open file descriptors too low
2025-04-13T23:39:02.826-05:00: For customers running the current memory allocator, we suggest changing the contents of the following sysfsFile
2025-04-13T23:39:02.826-05:00: For customers running the current memory allocator, we suggest changing the contents of the following sysfsFile
2025-04-13T23:39:02.826-05:00: We suggest setting the contents of sysfsFile to 0.
2025-04-13T23:39:02.826-05:00: Your system has glibc support for rseq built in, which is not yet supported by tcmalloc-google and has critical performance implications. Please set the environment variable MONGODB_PTHREAD_RSEQ=0
2025-04-13T23:39:02.826-05:00: We suggest setting swappiness to 0 or 1, as swapping can cause performance problems.

test> rs.initiate({
...   _id: 'theDailyReplica',
...   members: [
...     { _id: 0, host: 'gpel9.cs.nor.ou.edu:24157' },
...     { _id: 1, host: 'gpel10.cs.nor.ou.edu:24157' },
...     { _id: 2, host: 'gpel11.cs.nor.ou.edu:24157' }
...   ]
... })
{
  ok: 1,
  $clusterTime: {
    clusterTime: timestamp({ t: 1744605654, i: 1 }),
    signature: {
      hash: Binary.createFromBase64('AAAAAAAAAAAAAAAAAAAAAA='), b64: true
      keyId: Long('0')
    }
  },
  operationTime: timestamp({ t: 1744605654, i: 1 })
}
theDailyReplica [direct: other] test> rs.status()
{
  set: 'theDailyReplica',
  date: ISODate('2025-04-14T04:41:25.331Z'),
  myState: 1,
  term: 1,
  syncSourceHost: '',
  syncSourceId: -1,
  heartbeatIntervalMillis: Long('2000'),
  majorityVoteCount: 2,
  writeMajorityCount: 2,
  votingMembersCount: 3,
  writeableVotingMembersCount: 3,
  options: {
    lastCommittedOpTime: { ts: Timestamp({ t: 1744605654, i: 1 }), t: Long('1') },
    lastCommittedWalltime: ISODate('2025-04-14T04:41:24.783Z'),
    readConcernMajorityOpTime: { ts: Timestamp({ t: 1744605654, i: 1 }), t: Long('1') },
    appliedOpTime: { ts: Timestamp({ t: 1744605654, i: 1 }), t: Long('1') },
    durableOpTime: { ts: Timestamp({ t: 1744605654, i: 1 }), t: Long('1') },
    writtenOpTime: { ts: Timestamp({ t: 1744605654, i: 1 }), t: Long('1') },
    lastAppliedWalltime: ISODate('2025-04-14T04:41:24.783Z'),
    lastDurableWalltime: ISODate('2025-04-14T04:41:24.783Z')
  },
  electionCandidateMetrics: {
    lastElectionReason: 'electionTimeout',
    lastElectionDate: ISODate('2025-04-14T04:41:04.701Z'),
    electionTerm: Long('1'),
    lastCommittedOpTimeAtElection: { ts: Timestamp({ t: 1744605654, i: 1 }), t: Long('-1') },
    lastSeenWrittenOpTimeAtElection: { ts: Timestamp({ t: 1744605654, i: 1 }), t: Long('-1') },
    lastSeenOpTimeAtElection: { ts: Timestamp({ t: 1744605654, i: 1 }), t: Long('-1') },
    numVotesNeeded: 2,
    priorityAtElection: 1,
    electionTimeoutMillis: Long('10000'),
    numCatchUpOps: Long('0'),
    newTermStartDate: ISODate('2025-04-14T04:41:04.736Z'),
    wMajorityWriteAvailabilityDate: ISODate('2025-04-14T04:41:05.221Z')
  },
  members: [
    {
      _id: 0,
      name: 'gpel9.cs.nor.ou.edu:24157',
      health: 1,
      state: 1,
      stateStr: 'PRIMARY',
      uptime: 143,
      optime: { ts: Timestamp({ t: 1744605684, i: 1 }), t: Long('1') },
      optimeDate: ISODate('2025-04-14T04:41:24.000Z'),
      optimeWritten: { ts: Timestamp({ t: 1744605684, i: 1 }), t: Long('1') },
      optimeWrittenDate: ISODate('2025-04-14T04:41:24.000Z'),
      lastAppliedWallTime: ISODate('2025-04-14T04:41:24.753Z'),
      lastDurableWallTime: ISODate('2025-04-14T04:41:24.753Z'),
      lastWrittenWallTime: ISODate('2025-04-14T04:41:24.753Z'),
      syncSourceHost: '',
      syncSourceId: -1,
      infoMessage: 'Could not find member to sync from',
      electionTime: timestamp({ t: 1744605664, i: 1 }),
      electionDate: ISODate('2025-04-14T04:41:04.000Z'),
      configVersion: 1,
      configTerm: 1,
      self: true,
      lastHeartbeatMessage: ''
    },
    {
      _id: 1,
      name: 'gpel10.cs.nor.ou.edu:24157',
      health: 1,
      state: 2,
      stateStr: 'SECONDARY',
      uptime: 30,
      optime: { ts: Timestamp({ t: 1744605664, i: 16 }), t: Long('1') },
      optimeDurable: { ts: Timestamp({ t: 1744605664, i: 16 }), t: Long('1') },
      optimeWritten: { ts: Timestamp({ t: 1744605664, i: 16 }), t: Long('1') },
      optimeDate: ISODate('2025-04-14T04:41:04.000Z'),
      optimeDurableDate: ISODate('2025-04-14T04:41:04.000Z'),
      optimeWrittenDate: ISODate('2025-04-14T04:41:04.000Z'),
      lastAppliedWallTime: ISODate('2025-04-14T04:41:24.753Z'),
      lastDurableWallTime: ISODate('2025-04-14T04:41:24.753Z'),
      lastWrittenWallTime: ISODate('2025-04-14T04:41:24.753Z'),
      lastHeartbeat: ISODate('2025-04-14T04:41:24.721Z'),
      lastHeartbeatRecv: ISODate('2025-04-14T04:41:25.223Z'),
      pingMs: Long('0'),
      lastHeartbeatMessage: '',
      syncSourceHost: 'gpel9.cs.nor.ou.edu:24157',
      syncSourceId: 0,
      lastHeartbeat: ISODate('2025-04-14T04:41:24.721Z')
    }
  ]
}

```

```

lastStableRecoveryTimestamp: timestamp({ t: 1744605654, i: 1 }),
electionCandidateMetrics: {
  lastElectionReason: 'electionTimeout',
  lastElectionDate: ISODate('2025-04-14T04:41:04.701Z'),
  electionTerm: Long('1'),
  lastCommittedOpTimeAtElection: { ts: timestamp({ t: 1744605654, i: 1 }), t: Long('-1') },
  lastSeenWrittenOpTimeAtElection: { ts: timestamp({ t: 1744605654, i: 1 }), t: Long('-1') },
  lastSeenOpTimeAtElection: { ts: timestamp({ t: 1744605654, i: 1 }), t: Long('-1') },
  numVotesNeeded: 2,
  priorityAtElection: 1,
  electionTimeoutMillis: Long('10000'),
  numCatchUpOps: Long('0'),
  newTermStartDate: ISODate('2025-04-14T04:41:04.736Z'),
  wMajorityWriteAvailabilityDate: ISODate('2025-04-14T04:41:05.221Z')
},
members: [
  {
    _id: 0,
    name: 'gpel9.cs.nor.ou.edu:24157',
    health: 1,
    state: 1,
    stateStr: 'PRIMARY',
    uptime: 143,
    optime: { ts: timestamp({ t: 1744605684, i: 1 }), t: Long('1') },
    optimeDate: ISODate('2025-04-14T04:41:24.000Z'),
    optimeWritten: { ts: timestamp({ t: 1744605684, i: 1 }), t: Long('1') },
    optimeWrittenDate: ISODate('2025-04-14T04:41:24.000Z'),
    lastAppliedWallTime: ISODate('2025-04-14T04:41:24.753Z'),
    lastDurableWallTime: ISODate('2025-04-14T04:41:24.753Z'),
    lastWrittenWallTime: ISODate('2025-04-14T04:41:24.753Z'),
    syncSourceHost: '',
    syncSourceId: -1,
    infoMessage: 'Could not find member to sync from',
    electionTime: timestamp({ t: 1744605664, i: 1 }),
    electionDate: ISODate('2025-04-14T04:41:04.000Z'),
    configVersion: 1,
    configTerm: 1,
    self: true,
    lastHeartbeatMessage: ''
  },
  {
    _id: 1,
    name: 'gpel10.cs.nor.ou.edu:24157',
    health: 1,
    state: 2,
    stateStr: 'SECONDARY',
    uptime: 30,
    optime: { ts: timestamp({ t: 1744605664, i: 16 }), t: Long('1') },
    optimeDurable: { ts: timestamp({ t: 1744605664, i: 16 }), t: Long('1') },
    optimeWritten: { ts: timestamp({ t: 1744605664, i: 16 }), t: Long('1') },
    optimeDate: ISODate('2025-04-14T04:41:04.000Z'),
    optimeDurableDate: ISODate('2025-04-14T04:41:04.000Z'),
    optimeWrittenDate: ISODate('2025-04-14T04:41:04.000Z'),
    lastAppliedWallTime: ISODate('2025-04-14T04:41:24.753Z'),
    lastDurableWallTime: ISODate('2025-04-14T04:41:24.753Z'),
    lastWrittenWallTime: ISODate('2025-04-14T04:41:24.753Z'),
    lastHeartbeat: ISODate('2025-04-14T04:41:24.721Z'),
    lastHeartbeatRecv: ISODate('2025-04-14T04:41:25.223Z'),
    pingMs: Long('0'),
    lastHeartbeatMessage: '',
    syncSourceHost: 'gpel9.cs.nor.ou.edu:24157',
    syncSourceId: 0,
    lastHeartbeat: ISODate('2025-04-14T04:41:24.721Z')
  }
]

```

```

    },
    {
      _id: 2,
      name: 'gpel11.cs.nor.ou.edu:24157',
      health: 1,
      state: 2,
      stateStr: 'SECONDARY',
      uptime: 30,
      optime: { ts: Timestamp({ t: 1744605664, i: 16 }), t: Long('1') },
      optimeDurable: { ts: Timestamp({ t: 1744605664, i: 16 }), t: Long('1') },
      optimeWritten: { ts: Timestamp({ t: 1744605664, i: 16 }), t: Long('1') },
      optimeDate: ISODate('2025-04-14T04:41:04.000Z'),
      optimeDurableDate: ISODate('2025-04-14T04:41:04.000Z'),
      optimeWrittenDate: ISODate('2025-04-14T04:41:04.000Z'),
      lastAppliedWallTime: ISODate('2025-04-14T04:41:24.753Z'),
      lastDurableWallTime: ISODate('2025-04-14T04:41:24.753Z'),
      lastWrittenWallTime: ISODate('2025-04-14T04:41:24.753Z'),
      lastHeartbeat: ISODate('2025-04-14T04:41:24.723Z'),
      lastHeartbeatRecv: ISODate('2025-04-14T04:41:23.722Z'),
      pingMs: Long('0'),
      lastHeartbeatMessage: '',
      syncSourceHost: 'gpel9.cs.nor.ou.edu:24157',
      syncSourceId: 0,
      infoMessage: '',
      configVersion: 1,
      configTerm: 1
    }
  ],
  ok: 1,
  '$clusterTime': {
    clusterTime: Timestamp({ t: 1744605684, i: 1 }),
    signature: {
      hash: Binary.createFromBase64('AAAAAAAAAAAAAAAAAAAAAAA='),
      keyId: Long('0')
    }
  },
  operationTime: Timestamp({ t: 1744605684, i: 1 })
}
theDailyReplica [direct: primary] test> █

```

## Justification for Design:

### Fault Tolerance and High Availability

- MongoDB replica sets ensure that even if one node fails (e.g., due to server crash, network issues), the system can elect a new primary and continue to process read/write operations.
- Using three machines guarantees automatic failover with quorum-based elections — as soon as the majority (2 out of 3) are alive, the set remains operational.
- For example, if the primary node on gpel9 fails, gpel10 or gpel11 will be elected as the new primary automatically.

```

members: [
  {
    _id: 0,
    name: 'gpel9.cs.nor.ou.edu:24157',
    health: 0,
    state: 0,
    stateStr: '(not reachable/healthy)',
    uptime: 0,
    optime: { ts: Timestamp({ t: 0, i: 0 }), t: Long('-1') },
    optimeDurable: { ts: Timestamp({ t: 0, i: 0 }), t: Long('-1') },
    optimeWritten: { ts: Timestamp({ t: 0, i: 0 }), t: Long('-1') },
    optimeDate: ISODate('1970-01-01T00:00:00.000Z'),
    optimeDurableDate: ISODate('1970-01-01T00:00:00.000Z'),
    optimeWrittenDate: ISODate('1970-01-01T00:00:00.000Z'),
    lastAppliedWallTime: ISODate('2025-04-14T04:51:14.777Z'),
    lastDurableWallTime: ISODate('2025-04-14T04:51:14.777Z'),
    lastWrittenWallTime: ISODate('2025-04-14T04:51:14.777Z'),
    lastHeartbeat: ISODate('2025-04-14T04:51:58.039Z'),
    lastHeartbeatRecv: ISODate('2025-04-14T04:51:21.005Z'),
    pingMs: Long('0'),
    lastHeartbeatMessage: 'Error connecting to gpel9.cs.nor.ou.edu:24157 (10.194.8.51:24157) :: caused by :: onInvoke :: caused by :: Connection refused',
    syncSourceHost: '',
    syncSourceId: -1,
    infoMessage: '',
    configVersion: 1,
    configTerm: 1
  },
  {
    _id: 1,
    name: 'gpel10.cs.nor.ou.edu:24157',
    health: 1,
    state: 1,
    stateStr: 'PRIMARY',
    uptime: 1686,
    optime: { ts: Timestamp({ t: 1744606310, i: 1 }), t: Long('2') },
    optimeDurable: { ts: Timestamp({ t: 1744606310, i: 1 }), t: Long('2') },
    optimeWritten: { ts: Timestamp({ t: 1744606310, i: 1 }), t: Long('2') },
    optimeWrittenDate: ISODate('2025-04-14T04:51:50.000Z'),
    lastAppliedWallTime: ISODate('2025-04-14T04:51:50.024Z'),
    lastDurableWallTime: ISODate('2025-04-14T04:51:50.024Z'),
    lastWrittenWallTime: ISODate('2025-04-14T04:51:50.024Z'),
    syncSourceHost: '',
    syncSourceId: -1,
    infoMessage: 'Could not find member to sync from',
    electionTime: Timestamp({ t: 1744606290, i: 1 }),
    electionDate: ISODate('2025-04-14T04:51:30.000Z'),
    configVersion: 1,
    configTerm: 2,
    self: true,
    lastHeartbeatMessage: ''
  }
],
{
  _id: 2,
  name: 'gpel11.cs.nor.ou.edu:24157',
  health: 1,
  state: 2,
  stateStr: 'SECONDARY',
  uptime: 664,
  optime: { ts: Timestamp({ t: 1744606310, i: 1 }), t: Long('2') },
  optimeDurable: { ts: Timestamp({ t: 1744606310, i: 1 }), t: Long('2') },
  optimeWritten: { ts: Timestamp({ t: 1744606310, i: 1 }), t: Long('2') },
  optimeDate: ISODate('2025-04-14T04:51:50.000Z'),
  optimeDurableDate: ISODate('2025-04-14T04:51:50.000Z'),
  optimeWrittenDate: ISODate('2025-04-14T04:51:50.000Z'),
  lastAppliedWallTime: ISODate('2025-04-14T04:51:50.024Z'),
  lastDurableWallTime: ISODate('2025-04-14T04:51:50.024Z'),
  lastWrittenWallTime: ISODate('2025-04-14T04:51:50.024Z'),
  lastHeartbeat: ISODate('2025-04-14T04:51:58.027Z'),
  lastHeartbeatRecv: ISODate('2025-04-14T04:51:58.539Z'),
  pingMs: Long('0'),
  lastHeartbeatMessage: '',
  syncSourceHost: 'gpel10.cs.nor.ou.edu:24157',
  syncSourceId: 1,
  infoMessage: '',
  configVersion: 1,
  configTerm: 2
},
],
ok: 1,
'$clusterTime': {
  clusterTime: Timestamp({ t: 1744606310, i: 1 }),
  signature: {
    hash: Binary.createFromBase64('AAAAAAAAAAAAAAAAAAAAAAA='),
    keyId: Long('0')
  }
},
operationTime: Timestamp({ t: 1744606310, i: 1 })
}

```

## Read Scalability

- Secondary nodes (gpel10, gpel11) can serve read-only workloads, such as analytics queries (e.g., retrieving top 10 articles or latest news).
- This reduces the load on the primary node, which handles writes (like article insertions and comments).

## Data Redundancy

- Data is automatically replicated to all members. This protects against data loss if one server fails or data corruption occurs.
- Backups and maintenance operations can be performed on secondary nodes without affecting the primary.

## Geographical or Logical Separation

- Even though the machines are in the same data center (OU GPEL), distributing them logically across nodes improves resilience.
- In real-world extensions, you could spread them across zones or racks to protect against rack-level outages.

## Supports Use Case Requirements

Each of the frequent queries defined in the project is read-heavy, including:

- Retrieving latest articles
- Viewing most-read content
- Running full-text search

These are ideal for offloading to secondary nodes, while writes (new articles, comments) go to the primary.

## Task (h): Primary Node Failure Simulation

### (i) Killing the primary node gpel9:

```
srib0000@gpel9:~$ ps -ef | grep "mongod --port 24157"
srib0000 2196614 2175284  0 23:49 pts/20    00:00:00 grep --color=auto mongod --port 24157
srib0000@gpel9:~$ kill -9 2182492
srib0000@gpel9:~$
```

**(ii) Electing on the secondary nodes gpel10 or gpel11 as primary:**

```
theDailyReplica [direct: primary] test> rs.status()
{
  set: 'theDailyReplica',
  date: ISODate('2025-04-14T04:51:59.794Z'),
  myState: 1,
  term: Long('2'),
  syncSourceHost: '',
  syncSourceId: -1,
  heartbeatIntervalMillis: Long('2000'),
  majorityVoteCount: 2,
  writeMajorityCount: 2,
  votingMembersCount: 3,
  writableVotingMembersCount: 3,
  optimes: {
    lastCommittedOpTime: { ts: Timestamp({ t: 1744606310, i: 1 }), t: Long('2') },
    lastCommittedWallTime: ISODate('2025-04-14T04:51:50.024Z'),
    readConcernMajorityOpTime: { ts: Timestamp({ t: 1744606310, i: 1 }), t: Long('2') },
    appliedOpTime: { ts: Timestamp({ t: 1744606310, i: 1 }), t: Long('2') },
    durableOpTime: { ts: Timestamp({ t: 1744606310, i: 1 }), t: Long('2') },
    writtenOpTime: { ts: Timestamp({ t: 1744606310, i: 1 }), t: Long('2') },
    lastAppliedWallTime: ISODate('2025-04-14T04:51:50.024Z'),
    lastDurableWallTime: ISODate('2025-04-14T04:51:50.024Z'),
    lastWrittenWallTime: ISODate('2025-04-14T04:51:50.024Z')
  },
  lastStableRecoveryTimestamp: Timestamp({ t: 1744606264, i: 1 }),
  electionCandidateMetrics: {
    lastElectionReason: 'electionTimeout',
    lastElectionDate: ISODate('2025-04-14T04:51:30.017Z'),
    electionTerm: Long('2'),
    lastCommittedOpTimeAtElection: { ts: Timestamp({ t: 1744606274, i: 1 }), t: Long('1') },
    lastSeenWrittenOpTimeAtElection: { ts: Timestamp({ t: 1744606274, i: 1 }), t: Long('1') },
    lastSeenOpTimeAtElection: { ts: Timestamp({ t: 1744606274, i: 1 }), t: Long('1') },
    numVotesNeeded: 2,
    priorityAtElection: 1,
    electionTimeoutMillis: Long('10000'),
    numCatchUpOps: Long('0'),
    newTermStartDate: ISODate('2025-04-14T04:51:30.024Z'),
    wMajorityWriteAvailabilityDate: ISODate('2025-04-14T04:51:30.124Z')
  },
  electionParticipantMetrics: {
    votedForCandidate: true,
    electionTerm: Long('1'),
    lastVoteDate: ISODate('2025-04-14T04:41:04.702Z'),
    electionCandidateMemberId: 0,
    voteReason: '',
    lastWrittenOpTimeAtElection: { ts: Timestamp({ t: 1744605654, i: 1 }), t: Long(-1) },
    maxWrittenOpTimeInSet: { ts: Timestamp({ t: 1744605654, i: 1 }), t: Long(-1) },
    lastAppliedOpTimeAtElection: { ts: Timestamp({ t: 1744605654, i: 1 }), t: Long(-1) },
    maxAppliedOpTimeInSet: { ts: Timestamp({ t: 1744605654, i: 1 }), t: Long(-1) },
    priorityAtElection: 1
  }
},
```

(iii) Re running queries from task (e) with new primary node gpel10:

**Query 1:** Insert a (news or feature) article into the archive

```
theDailyReplica [direct: primary] test> use dailyNewsArchive
switched to db dailyNewsArchive
theDailyReplica [direct: primary] dailyNewsArchive> db.articles.insertOne({
...   article_type: "news article",
...   article_section: "Science",
...   article_id: 56,
...   post_on_date: { day: 1, month: 3, year: 2023, time: "10:00:00" },
...   reporter_name: "Reporter 56",
...   similar_stories: [],
...   num_times_read: 500,
...   article_text: "This article discusses the latest developments in science.",
...   comments: []
... })
{
  acknowledged: true,
  insertedId: ObjectId('67fc972c5c1241b3f56b140b')
}
theDailyReplica [direct: primary] dailyNewsArchive> █
```

```
theDailyReplica [direct: primary] dailyNewsArchive> db.articles.find({ article_id: 56 }).pretty()
[
  {
    _id: ObjectId('67fc972c5c1241b3f56b140b'),
    article_type: 'news article',
    article_section: 'Science',
    article_id: 56,
    post_on_date: { day: 1, month: 3, year: 2023, time: '10:00:00' },
    reporter_name: 'Reporter 56',
    similar_stories: [],
    num_times_read: 500,
    article_text: 'This article discusses the latest developments in science.',
    comments: []
  }
]
theDailyReplica [direct: primary] dailyNewsArchive> █
```

## Query 2: Retrieve the latest news

```
theDailyReplica [direct: primary] dailyNewsArchive> db.articles.find({ article_type: "news article" }).sort({ "post_on_date.year": -1, "post_on_date.month": -1, "post_on_date.day": -1 }).limit(1).pretty()
[
  {
    _id: ObjectId('67fc87b295be239abs6b140c'),
    article_type: 'news article',
    article_section: 'Health',
    article_id: 1001,
    post_on_date: { day: 12, month: 4, year: 2023, time: '14:00:00' },
    reporters: [
      { name: 'MS Dhoni', location: 'Chennai' },
      { name: 'Suresh Raina', location: 'Ghaziabad' }
    ],
    similar_stories: [],
    num_times_read: 1500,
    article_text: 'This article discusses healthcare AI in India..',
    comments: []
  }
]
theDailyReplica [direct: primary] dailyNewsArchive>
```

## Query 3: Retrieve the top 10 most read news in the past month

```
theDailyReplica [direct: primary] dailyNewsArchive> db.articles.find({
...   "post_on_date.year": 2023,
...   "post_on_date.month": { $gte: 2 }
... }).sort({ num_times_read: -1 }).limit(10).pretty()
[
  {
    _id: ObjectId('67fc772c5b76bea4836b1436'),
    article_type: 'feature article',
    article_section: 'Entertainment',
    article_id: 44,
    post_on_date: { day: 14, month: 2, year: 2023, time: '00:00:00' },
    reporter_name: 'Reporter 44',
    similar_stories: [
      'https://thedaily.com/politics/election-story-44.html',
      'https://thedaily.com/sports/election-story-44.html'
    ],
    num_times_read: 19903,
    article_text: 'This article discusses the latest developments in election.',
    comments: [
      {
        comment_id: 44,
        article_id: 44,
        user_id: 'user44@mail.com',
        posted_on_date: { day: 14, month: 2, year: 2023, time: '00:00:00' },
        comment_text: 'Amazing insight!',
        score: -79
      }
    ]
  },
  {
    _id: ObjectId('67fc772c5b76bea4836b1438'),
    article_type: 'feature article',
    article_section: 'Politics',
    article_id: 46,
    post_on_date: { day: 16, month: 2, year: 2023, time: '00:00:00' },
    reporter_name: 'Reporter 46',
    similar_stories: [
      'https://thedaily.com/politics/covid-story-46.html',
      'https://thedaily.com/sports/covid-story-46.html'
    ],
    num_times_read: 19132,
    article_text: 'This article discusses the latest developments in covid.',
    comments: [
      {
        comment_id: 46,
        article_id: 46,
        user_id: 'user46@mail.com',
        posted_on_date: { day: 16, month: 2, year: 2023, time: '00:00:00' },
        comment_text: 'Needs better research.',
        score: 29
      }
    ]
  }
]
```

```
{
  "_id": ObjectId("67fc772b5b76bea4836b142f"),
  "article_type": "news article",
  "article_section": "Entertainment",
  "article_id": 37,
  "post_on_date": { "day": 7, "month": 2, "year": 2023, "time": "00:00:00" },
  "reporter_name": "Reporter 37",
  "similar_stories": [
    "https://thedaily.com/politics/AI-story-37.html",
    "https://thedaily.com/sports/AI-story-37.html"
  ],
  "num_times_read": 18879,
  "article_text": "This article discusses the latest developments in AI.",
  "comments": [
    {
      "comment_id": 37,
      "article_id": 37,
      "user_id": "user37@mail.com",
      "posted_on_date": { "day": 7, "month": 2, "year": 2023, "time": "00:00:00" },
      "comment_text": "Needs better research.",
      "score": -30
    }
  ]
},
{
  "_id": ObjectId("67fc772d5b76bea4836b143d"),
  "article_type": "feature article",
  "article_section": "Sports",
  "article_id": 51,
  "post_on_date": { "day": 21, "month": 2, "year": 2023, "time": "00:00:00" },
  "reporter_name": "Reporter 51",
  "similar_stories": [
    "https://thedaily.com/politics/covid-story-51.html",
    "https://thedaily.com/sports/covid-story-51.html"
  ],
  "num_times_read": 18628,
  "article_text": "This article discusses the latest developments in covid.",
  "comments": [
    {
      "comment_id": 51,
      "article_id": 51,
      "user_id": "user51@mail.com",
      "posted_on_date": { "day": 21, "month": 2, "year": 2023, "time": "00:00:00" },
      "comment_text": "Not convinced with the facts.",
      "score": 19
    }
  ]
},
```

```
_id: ObjectId('67fc772d5b76bea4836b143a'),
article_type: 'feature article',
article_section: 'Sports',
article_id: 48,
post_on_date: { day: 18, month: 2, year: 2023, time: '00:00:00' },
reporter_name: 'Reporter 48',
similar_stories: [
  'https://thedaily.com/politics/space-story-48.html',
  'https://thedaily.com/sports/space-story-48.html'
],
num_times_read: 17121,
article_text: 'This article discusses the latest developments in space.',
comments: [
  {
    comment_id: 48,
    article_id: 48,
    user_id: 'user48@mail.com',
    posted_on_date: { day: 18, month: 2, year: 2023, time: '00:00:00' },
    comment_text: 'Fantastic read!',
    score: -17
  }
],
},
{
  _id: ObjectId('67fc772b5b76bea4836b1431'),
  article_type: 'news article',
  article_section: 'Entertainment',
  article_id: 39,
  post_on_date: { day: 9, month: 2, year: 2023, time: '00:00:00' },
  reporter_name: 'Reporter 39',
  similar_stories: [
    'https://thedaily.com/politics/climate-story-39.html',
    'https://thedaily.com/sports/climate-story-39.html'
  ],
  num_times_read: 15636,
  article_text: 'This article discusses the latest developments in climate.',
  comments: [
    {
      comment_id: 39,
      article_id: 39,
      user_id: 'user39@mail.com',
      posted_on_date: { day: 9, month: 2, year: 2023, time: '00:00:00' },
      comment_text: 'I disagree with the viewpoint.',
      score: 46
    }
  ]
},
```

```
_id: ObjectId('67fc772c5b76bea4836b1435'),
article_type: 'feature article',
article_section: 'Politics',
article_id: 43,
post_on_date: { day: 13, month: 2, year: 2023, time: '00:00:00' },
reporter_name: 'Reporter 43',
similar_stories: [
  'https://thedaily.com/politics/AI-story-43.html',
  'https://thedaily.com/sports/AI-story-43.html'
],
num_times_read: 15403,
article_text: 'This article discusses the latest developments in AI.',
comments: [
  {
    comment_id: 43,
    article_id: 43,
    user_id: 'user43@mail.com',
    posted_on_date: { day: 13, month: 2, year: 2023, time: '00:00:00' },
    comment_text: 'Looking forward to more.',
    score: -9
  }
]
},
{
  _id: ObjectId('67fc772c5b76bea4836b1434'),
  article_type: 'feature article',
  article_section: 'Sports',
  article_id: 42,
  post_on_date: { day: 12, month: 2, year: 2023, time: '00:00:00' },
  reporter_name: 'Reporter 42',
  similar_stories: [
    'https://thedaily.com/politics/education-story-42.html',
    'https://thedaily.com/sports/education-story-42.html'
  ],
  num_times_read: 15180,
  article_text: 'This article discusses the latest developments in education.',
  comments: [
    {
      comment_id: 42,
      article_id: 42,
      user_id: 'user42@mail.com',
      posted_on_date: { day: 12, month: 2, year: 2023, time: '00:00:00' },
      comment_text: 'Well-written and concise.',
      score: -73
    }
  ]
},
```

```
_id: ObjectId('67fc772a5b76bea4836b142d'),
article_type: 'feature article',
article_section: 'Entertainment',
article_id: 35,
post_on_date: { day: 5, month: 2, year: 2023, time: '00:00:00' },
reporter_name: 'Reporter 35',
similar_stories: [
  'https://thedaily.com/politics/space-story-35.html',
  'https://thedaily.com/sports/space-story-35.html'
],
num_times_read: 14880,
article_text: 'This article discusses the latest developments in space.',
comments: [
  {
    comment_id: 35,
    article_id: 35,
    user_id: 'user35@mail.com',
    posted_on_date: { day: 5, month: 2, year: 2023, time: '00:00:00' },
    comment_text: 'Bold take on current affairs.',
    score: 99
  }
]
},
{
  _id: ObjectId('67fc772e5b76bea4836b1441'),
  article_type: 'feature article',
  article_section: 'Health',
  article_id: 55,
  post_on_date: { day: 25, month: 2, year: 2023, time: '00:00:00' },
  reporter_name: 'Reporter 55',
  similar_stories: [
    'https://thedaily.com/politics/budget-story-55.html',
    'https://thedaily.com/sports/budget-story-55.html'
  ],
  num_times_read: 14479,
  article_text: 'This article discusses the latest developments in budget.',
  comments: [
    {
      comment_id: 55,
      article_id: 55,
      user_id: 'user55@mail.com',
      posted_on_date: { day: 25, month: 2, year: 2023, time: '00:00:00' },
      comment_text: 'Well-written and concise.',
      score: 41
    }
  ]
}
]
theDailyReplica [direct: primary] dailyNewsArchive> █
```

## Query 4: Add a comment to an article

```
'theDailyReplica [direct: primary] dailyNewsArchive> db.articles.updateOne(  
...   { article_id: 56 },  
...   {  
...     $push: {  
...       comments: {  
...         comment_id: 56,  
...         article_id: 56,  
...         user_id: "newuser@mail.com",  
...         posted_on_date: { day: 2, month: 3, year: 2023, time: "12:00:00" },  
...         comment_text: "Interesting article!",  
...         score: 10  
...       }  
...     }  
...   }  
[... )  
{  
  acknowledged: true,  
  insertedId: null,  
  matchedCount: 1,  
  modifiedCount: 1,  
  upsertedCount: 0  
}
```

```
'theDailyReplica [direct: primary] dailyNewsArchive> db.articles.find({ article_id: 56 }, { comments: 1 }).pretty()  
[  
  {  
    _id: ObjectId('67fc972c5c1241b3f56b140b'),  
    comments: [  
      {  
        comment_id: 56,  
        article_id: 56,  
        user_id: 'newuser@mail.com',  
        posted_on_date: { day: 2, month: 3, year: 2023, time: '12:00:00' },  
        comment_text: 'Interesting article!',  
        score: 10  
      }  
    ]  
  }  
]  
theDailyReplica [direct: primary] dailyNewsArchive> █
```

**Query 5:** Add a story to the ‘similar stories’ section for articles containing a given word and published within the past 2 years

```
theDailyReplica [direct: primary] dailyNewsArchive> db.articles.updateMany(  
...   {  
...     article_text: { $regex: /cricket/i },  
...     $or: [  
...       { "post_on_date.year": { $gt: 2023 } },  
...       {  
...         "post_on_date.year": 2023,  
...         "post_on_date.month": { $gte: 3 }  
...       }  
...     ]  
...   },  
...   {  
...     $addToSet: {  
...       similar_stories: "https://thedaily.com/sports/cricket-analysis-2025.html"  
...     }  
...   }  
[... ]  
{  
  acknowledged: true,  
  insertedId: null,  
  matchedCount: 0,  
  modifiedCount: 0,  
  upsertedCount: 0  
}
```

```
theDailyReplica [direct: primary] dailyNewsArchive> db.articles.find({  
...   article_text: /cricket/i,  
...   $or: [  
...     { "post_on_date.year": { $gt: 2023 } },  
...     {  
...       "post_on_date.year": 2023,  
...       "post_on_date.month": { $gte: 3 }  
...     }  
...   ]  
[... ]).pretty()  
  
theDailyReplica [direct: primary] dailyNewsArchive> █
```

Re running queries from task (e) with new primary node gpel11:

```
this incident has been reported to the administrator.
srib0000@gpel10:~$ lsof -nPiTCP:24157 | grep LISTEN
mongod 658600 srib0000 10u IPv4 142466113 0t0 TCP *:24157 (LISTEN)
srib0000@gpel10:~$ kill -9 658600
srib0000@gpel10:~$
```

```
srib0000@gpel11:~$ mongosh --port 24157
Current Mongosh Log ID: 67fdca0dfeeb879a246b148
Connecting to: mongos://127.0.0.1:24157/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+2.4.2
Using MongoDB: 8.0.6
Using Mongosh: 2.4.2
mongosh 2.4.6 is available for download: https://www.mongodb.com/try/download/shell
For mongosh info see: https://www.mongodb.com/docs/mongodb-shell/
-----
The server generated these startup warnings when booting
2025-04-13T23:25:28.362-05:00: Using the XFS filesystem is strongly recommended with the WiredTiger storage engine. See http://dochub.mongodb.org/core/prodnotes-filesystem
2025-04-13T23:25:28.362-05:00: Access control is not enabled for the database. Read and write access to data and configuration is unrestricted
2025-04-13T23:25:28.362-05:00: Soft limits for open file descriptors too low
2025-04-13T23:25:28.362-05:00: For customers running the current memory allocator, we suggest changing the contents of the following sysfsFile
2025-04-13T23:25:28.362-05:00: For customers running the current memory allocator, we suggest changing the contents of the following sysfsFile
2025-04-13T23:25:28.362-05:00: We suggest setting the contents of sysfsFile to 0
2025-04-13T23:25:28.362-05:00: Your system has glibc support for rseq built in, which is not yet supported by tcmalloc-google and has critical performance implications. Please set the environment variable GLIBC_TUNABLES=libc.rthread.rseq=0
2025-04-13T23:25:28.362-05:00: We suggest setting swappiness to 0 or 1, as swapping can cause performance problems.
-----
theDailyReplica [direct: secondary] test> rs.status()
{
  set: 'theDailyReplica',
  date: ISODate('2025-04-15T02:52:12.000Z'),
  myState: 2,
  term: Long('2'),
  syncSourceHost: '',
  syncSourceId: -1,
  heartbeatIntervalMillis: Long('2000'),
  majorityVoteCount: 2,
  writeMajorityCount: 2,
  votingMembersCount: 3,
  writeConcerningMembersCount: 3,
  options: {
    lastCommittedOpTime: { ts: Timestamp({ t: 1744685532, i: 1 }), t: Long('2') },
    lastCommittedWallTime: ISODate('2025-04-15T02:52:12.000Z'),
    readConcernMajorityOpTime: { ts: Timestamp({ t: 1744685532, i: 1 }), t: Long('2') },
    appliedOpTime: { ts: Timestamp({ t: 1744685532, i: 1 }), t: Long('2') },
    writtenOpTime: { ts: Timestamp({ t: 1744685532, i: 1 }), t: Long('2') },
    lastAppliedWallTime: ISODate('2025-04-15T02:52:12.000Z'),
    lastDurableWallTime: ISODate('2025-04-15T02:52:12.000Z'),
    lastWrittenWallTime: ISODate('2025-04-15T02:52:12.000Z')
  },
  lastStableRecoveryTimestamp: Timestamp({ t: 1744685532, i: 1 }),
  electionParticipantMetrics: {
    votedForCandidate: true,
    electionTerm: Long('2'),
    lastVoteDate: ISODate('2025-04-14T04:51:30.018Z'),
    electionCandidateMemberId: 1,
    voteReason: ''
  },
  lastWrittenOpTimeInElection: { ts: Timestamp({ t: 1744686274, i: 1 }), t: Long('1') },
  maxTimestampInSet: { ts: Timestamp({ t: 1744686274, i: 1 }), t: Long('1') },
  lastAppliedOpTimeInElection: { ts: Timestamp({ t: 1744686274, i: 1 }), t: Long('1') },
  maxAppliedOpTimeInSet: { ts: Timestamp({ t: 1744686274, i: 1 }), t: Long('1') },
  priorityAtElection: 1,
  newTermStartDate: ISODate('2025-04-14T04:51:30.024Z'),
  newTermAppliedDate: ISODate('2025-04-14T04:51:30.028Z')
},
members: [
  {
    _id: 0,
    name: 'gpel10.cs.mor.ou.edu:24157',
    health: 0,
    state: 8,
    stateStr: '(not reachable/healthy)',
    uptime: 0,
    optime: { ts: Timestamp({ t: 0, i: 0 }), t: Long('1') },
    optimeDurability: { ts: Timestamp({ t: 0, i: 0 }), t: Long('1') },
    optimeWritten: { ts: Timestamp({ t: 0, i: 0 }), t: Long('1') },
    optimeDate: ISODate('1970-01-01T00:00:00.000Z'),
    optimeLastAppliedDate: ISODate('1970-01-01T00:00:00.000Z'),
    optimeWrittenDate: ISODate('1970-01-01T00:00:00.000Z'),
    lastAppliedWallTime: ISODate('2025-04-14T04:51:14.777Z'),
    lastDurableWallTime: ISODate('2025-04-14T04:51:14.777Z'),
    lastWrittenWallTime: ISODate('2025-04-14T04:51:14.777Z'),
    lastHeartbeat: ISODate('2025-04-15T02:54:06.570Z'),
    lastHeartbeatRecv: ISODate('2025-04-14T04:51:21.005Z'),
    pingMs: Long('0'),
    lastHeartbeatMessage: 'Error connecting to gpel10.cs.mor.ou.edu:24157 (10.194.8.51:24157) :: caused by :: onInvoke :: caused by :: Connection refused',
    syncSourceHost: '',
    syncSourceId: -1,
    infoMessage: '',
    configVersion: 1,
    configTerm: 1
  },
  {
    _id: 1,
    name: 'gpel10.cs.mor.ou.edu:24157',
    health: 0,
    state: 8,
    stateStr: '(not reachable/healthy)',
    uptime: 0,
    optime: { ts: Timestamp({ t: 0, i: 0 }), t: Long('1') },
    optimeDurability: { ts: Timestamp({ t: 0, i: 0 }), t: Long('1') },
    optimeWritten: { ts: Timestamp({ t: 0, i: 0 }), t: Long('1') },
    optimeDate: ISODate('1970-01-01T00:00:00.000Z'),
    optimeLastAppliedDate: ISODate('1970-01-01T00:00:00.000Z'),
    optimeWrittenDate: ISODate('1970-01-01T00:00:00.000Z'),
    lastAppliedWallTime: ISODate('2025-04-15T02:52:12.000Z'),
    lastDurableWallTime: ISODate('2025-04-15T02:52:12.000Z'),
    lastWrittenWallTime: ISODate('2025-04-15T02:52:12.000Z'),
    lastHeartbeat: ISODate('2025-04-15T02:54:06.473Z'),
    lastHeartbeatRecv: ISODate('2025-04-15T02:52:15.033Z'),
    pingMs: Long('0'),
    lastHeartbeatMessage: 'Error connecting to gpel10.cs.mor.ou.edu:24157 (10.194.8.52:24157) :: caused by :: onInvoke :: caused by :: Connection refused',
    syncSourceHost: '',
    syncSourceId: -1,
    infoMessage: '',
    configVersion: 1,
    configTerm: 2
  },
  {
    _id: 2,
    name: 'gpel11.cs.mor.ou.edu:24157',
    health: 1,
    state: 2,
    stateStr: 'SECONDARY',
    uptime: 88924,
    optime: { ts: Timestamp({ t: 1744685532, i: 1 }), t: Long('2') },
    optimeDate: ISODate('2025-04-15T02:52:12.000Z'),
    optimeLastAppliedDate: ISODate('2025-04-15T02:52:12.000Z'),
    optimeWrittenDate: ISODate('2025-04-15T02:52:12.000Z'),
    lastAppliedWallTime: ISODate('2025-04-15T02:52:12.000Z')
  }
]
```

```
_id: 0,
name: 'gpel10.cs.mor.ou.edu:24157',
health: 0,
state: 8,
stateStr: '(not reachable/healthy)',
uptime: 0,
optime: { ts: Timestamp({ t: 0, i: 0 }), t: Long('1') },
optimeDurability: { ts: Timestamp({ t: 0, i: 0 }), t: Long('1') },
optimeWritten: { ts: Timestamp({ t: 0, i: 0 }), t: Long('1') },
optimeDate: ISODate('1970-01-01T00:00:00.000Z'),
optimeLastAppliedDate: ISODate('1970-01-01T00:00:00.000Z'),
optimeWrittenDate: ISODate('1970-01-01T00:00:00.000Z'),
lastAppliedWallTime: ISODate('2025-04-14T04:51:14.777Z'),
lastDurableWallTime: ISODate('2025-04-14T04:51:14.777Z'),
lastWrittenWallTime: ISODate('2025-04-14T04:51:14.777Z'),
lastHeartbeat: ISODate('2025-04-15T02:54:06.570Z'),
lastHeartbeatRecv: ISODate('2025-04-14T04:51:21.005Z'),
pingMs: Long('0'),
lastHeartbeatMessage: 'Error connecting to gpel10.cs.mor.ou.edu:24157 (10.194.8.51:24157) :: caused by :: onInvoke :: caused by :: Connection refused',
syncSourceHost: '',
syncSourceId: -1,
infoMessage: '',
configVersion: 1,
configTerm: 1
},
members: [
  {
    _id: 1,
    name: 'gpel10.cs.mor.ou.edu:24157',
    health: 0,
    state: 8,
    stateStr: '(not reachable/healthy)',
    uptime: 0,
    optime: { ts: Timestamp({ t: 0, i: 0 }), t: Long('1') },
    optimeDurability: { ts: Timestamp({ t: 0, i: 0 }), t: Long('1') },
    optimeWritten: { ts: Timestamp({ t: 0, i: 0 }), t: Long('1') },
    optimeDate: ISODate('1970-01-01T00:00:00.000Z'),
    optimeLastAppliedDate: ISODate('1970-01-01T00:00:00.000Z'),
    optimeWrittenDate: ISODate('1970-01-01T00:00:00.000Z'),
    lastAppliedWallTime: ISODate('2025-04-15T02:52:12.000Z'),
    lastDurableWallTime: ISODate('2025-04-15T02:52:12.000Z'),
    lastWrittenWallTime: ISODate('2025-04-15T02:52:12.000Z'),
    lastHeartbeat: ISODate('2025-04-15T02:54:06.473Z'),
    lastHeartbeatRecv: ISODate('2025-04-15T02:52:15.033Z'),
    pingMs: Long('0'),
    lastHeartbeatMessage: 'Error connecting to gpel10.cs.mor.ou.edu:24157 (10.194.8.51:24157) :: caused by :: onInvoke :: caused by :: Connection refused',
    syncSourceHost: '',
    syncSourceId: -1,
    infoMessage: '',
    configVersion: 1,
    configTerm: 2
  },
  {
    _id: 2,
    name: 'gpel11.cs.mor.ou.edu:24157',
    health: 1,
    state: 2,
    stateStr: 'SECONDARY',
    uptime: 88924,
    optime: { ts: Timestamp({ t: 1744685532, i: 1 }), t: Long('2') },
    optimeDate: ISODate('2025-04-15T02:52:12.000Z'),
    optimeLastAppliedDate: ISODate('2025-04-15T02:52:12.000Z'),
    optimeWrittenDate: ISODate('2025-04-15T02:52:12.000Z'),
    lastAppliedWallTime: ISODate('2025-04-15T02:52:12.000Z')
  }
]
```

```

sr1b0000@gpel19:~$ ps -ef | grep "mongod --port 24157"
sr1b0000 3651431 3650226 0 22:00 pts/12 00:00:00 grep --color=auto mongod --port 24157
sr1b0000@gpel19:~$ mongod --repSet theDailyReplica --port 24157 --dbpath ~/mongodb-data --bind_ip_all --fork --logpath ~/mongo.log
about to fork child process, waiting until server is ready for connections.
forked process: 3652247
child process started successfully, parent exiting
sr1b0000@gpel19:~$ mongo
Current Mongosh Log ID: 67fdcc44dbdb3c3806b2148s
Connecting to: mongodb://127.0.0.1:24157/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+2.4.2
Using MongoDB: 8.0.6
Using Mongosh: 2.4.2
mongosh 2.8.0 is available for download: https://www.mongodb.com/try/download/shell
For mongosh info see: https://www.mongodb.com/docs/mongodb-shell/
-----
The server generated these startup warnings when booting
2025-04-14T22:02:15.433-05:00: Using the XFS filesystem is strongly recommended with the WiredTiger storage engine. See http://dochub.mongodb.org/core/prodnotes-filesystem
2025-04-14T22:02:15.946-05:00: Access control is not enabled for the database. Read and write access to data and configuration is unrestricted
2025-04-14T22:02:15.946-05:00: Soft rlimits for open file descriptors too low
2025-04-14T22:02:15.946-05:00: For customers running the current memory allocator, we suggest changing the contents of the following sysfsFile
2025-04-14T22:02:15.946-05:00: For customers running the current memory allocator, we suggest changing the contents of the following sysfsFile
2025-04-14T22:02:15.946-05:00: We suggest setting the contents of sysfsFile to 0.
2025-04-14T22:02:15.946-05:00: Your system has glibc support for rseq built in, which is not yet supported by tcmalloc-google and has critical performance implications. Please set the environment variable GLIBC_TUNABLES=glibc.pthread.rseq=0
2025-04-14T22:02:15.946-05:00: We suggest setting swappiness to 0 or 1, as swapping can cause performance problems.
-----
theDailyReplica [direct: secondary] test> rs.status()
{
  set: 'theDailyReplica',
  date: ISODate('2025-04-15T03:02:37.648Z'),
  myState: 2,
  term: Long('3'),
  syncSourceHost: 'gpel11.cs.norou.edu:24157',
  syncSourceId: 2,
  heartbeatIntervalMillis: Long('2000'),
  maxTermVoteCount: 3,
  writeMajorityCount: 3,
  votingMembersCount: 3,
  writableVotingMembersCount: 3,
  optimes: {
    lastCommittedOptime: { ts: Timestamp({ t: 1744686156, i: 1 }), t: Long('3') },
    lastCommittedWalltime: ISODate('2025-04-15T03:02:36.148Z'),
    readConcernMajorityOptime: { ts: Timestamp({ t: 1744686156, i: 1 }), t: Long('3') },
    appliedOptime: { ts: Timestamp({ t: 1744686156, i: 1 }), t: Long('3') },
    durableOptime: { ts: Timestamp({ t: 1744686156, i: 1 }), t: Long('3') },
    writeMajorityWalltime: ISODate('2025-04-15T03:02:36.148Z'),
    lastAppliedWalltime: ISODate('2025-04-15T03:02:36.148Z'),
    lastDurableWalltime: ISODate('2025-04-15T03:02:36.148Z'),
    lastWrittenWalltime: ISODate('2025-04-15T03:02:36.148Z')
  },
  lastStableRecoveryTimestamp: Timestamp({ t: 1744686244, i: 1 }),
  electionParticipantMetrics: {
    votedForCandidate: true,
    electionTerm: Long('3'),
    lastAppliedWalltime: ISODate('2025-04-15T03:02:16.124Z'),
    electionCandidateMemberId: 2,
    voteReason: ''
  },
  lastWrittenOptimaAtElection: { ts: Timestamp({ t: 1744686274, i: 1 }), t: Long('1') },
  maxWrittenOptimeInSet: { ts: Timestamp({ t: 1744685532, i: 1 }), t: Long('2') },
  lastAppliedOptimeAtElection: { ts: Timestamp({ t: 1744686274, i: 1 }), t: Long('1') },
  maxAppliedOptimeInSet: { ts: Timestamp({ t: 1744685532, i: 1 }), t: Long('2') },
  members: [
    {
      _id: 0,
      name: 'gpel19.cs.norou.edu:24157',
      health: 5,
      state: 2,
      stateStr: 'SECONDARY',
      uptime: 22,
      optime: { ts: Timestamp({ t: 1744686156, i: 1 }), t: Long('3') },
      optimeDate: ISODate('2025-04-15T03:02:36.000Z'),
      optimeWritten: { ts: Timestamp({ t: 1744686156, i: 1 }), t: Long('3') },
      optimeWrittenDate: ISODate('2025-04-15T03:02:36.000Z'),
      lastAppliedWalltime: ISODate('2025-04-15T03:02:36.148Z'),
      lastDurableWalltime: ISODate('2025-04-15T03:02:36.148Z'),
      lastWrittenWalltime: ISODate('2025-04-15T03:02:36.148Z'),
      syncSourceHost: 'gpel11.cs.norou.edu:24157',
      syncSourceId: 2,
      infoMessage: '',
      configVersion: 1,
      configTerm: 3,
      self: true,
      lastHeartbeatMessage: ''
    },
    {
      _id: 1,
      name: 'gpel10.cs.norou.edu:24157',
      health: 0,
      state: 8,
      stateStr: '(not reachable/healthy)',
      uptime: 0,
      optime: { ts: Timestamp({ t: 0, i: 0 }), t: Long('-1') },
      optimeDurable: { ts: Timestamp({ t: 0, i: 0 }), t: Long('-1') },
      optimeWritten: { ts: Timestamp({ t: 0, i: 0 }), t: Long('-1') },
      optimeWrittenDate: ISODate('1970-01-01T00:00:00.000Z'),
      optimeDurableDate: ISODate('1970-01-01T00:00:00.000Z'),
      optimeWrittenDate: ISODate('1970-01-01T00:00:00.000Z'),
      lastAppliedWalltime: ISODate('1970-01-01T00:00:00.000Z'),
      lastDurableWalltime: ISODate('1970-01-01T00:00:00.000Z'),
      lastWrittenWalltime: ISODate('1970-01-01T00:00:00.000Z'),
      lastHeartbeat: ISODate('2025-04-15T03:02:36.711Z'),
      lastHeartbeatRecv: ISODate('1970-01-01T00:00:00.000Z'),
      pingMs: Long('0'),
      lastHeartbeatMessage: 'Error connecting to gpel10.cs.norou.edu:24157 (19.194.8.82:24157) :: caused by :: envoke :: caused by :: Connection refused',
      syncSourceHost: '',
      syncSourceId: 1,
      infoMessage: '',
      configVersion: -1,
      configTerm: -1,
      self: false
    },
    {
      _id: 2,
      name: 'gpel11.cs.norou.edu:24157',
      health: 1,
      state: 3,
      stateStr: 'PRIMARY',
      uptime: 21,
      optime: { ts: Timestamp({ t: 1744686156, i: 1 }), t: Long('3') },
      optimeDurable: { ts: Timestamp({ t: 1744686156, i: 1 }), t: Long('3') },
      optimeWritten: { ts: Timestamp({ t: 1744686156, i: 1 }), t: Long('3') }
    }
  ]
}

```

```

priorityAtElection: 1,
newTermStartDate: ISODate('2025-04-15T03:02:16.146Z'),
newTermAppliedDate: ISODate('2025-04-15T03:02:16.705Z'),
},
members: [
  {
    _id: 0,
    name: 'gpel19.cs.norou.edu:24157',
    health: 5,
    state: 2,
    stateStr: 'SECONDARY',
    uptime: 22,
    optime: { ts: Timestamp({ t: 1744686156, i: 1 }), t: Long('3') },
    optimeDate: ISODate('2025-04-15T03:02:36.000Z'),
    optimeWritten: { ts: Timestamp({ t: 1744686156, i: 1 }), t: Long('3') },
    optimeWrittenDate: ISODate('2025-04-15T03:02:36.000Z'),
    lastAppliedWalltime: ISODate('2025-04-15T03:02:36.148Z'),
    lastDurableWalltime: ISODate('2025-04-15T03:02:36.148Z'),
    lastWrittenWalltime: ISODate('2025-04-15T03:02:36.148Z'),
    syncSourceHost: 'gpel11.cs.norou.edu:24157',
    syncSourceId: 2,
    infoMessage: '',
    configVersion: 1,
    configTerm: 3,
    self: true,
    lastHeartbeatMessage: ''
  },
  {
    _id: 1,
    name: 'gpel10.cs.norou.edu:24157',
    health: 0,
    state: 8,
    stateStr: '(not reachable/healthy)',
    uptime: 0,
    optime: { ts: Timestamp({ t: 0, i: 0 }), t: Long('-1') },
    optimeDurable: { ts: Timestamp({ t: 0, i: 0 }), t: Long('-1') },
    optimeWritten: { ts: Timestamp({ t: 0, i: 0 }), t: Long('-1') },
    optimeWrittenDate: ISODate('1970-01-01T00:00:00.000Z'),
    optimeDurableDate: ISODate('1970-01-01T00:00:00.000Z'),
    optimeWrittenDate: ISODate('1970-01-01T00:00:00.000Z'),
    lastAppliedWalltime: ISODate('1970-01-01T00:00:00.000Z'),
    lastDurableWalltime: ISODate('1970-01-01T00:00:00.000Z'),
    lastWrittenWalltime: ISODate('1970-01-01T00:00:00.000Z'),
    lastHeartbeat: ISODate('2025-04-15T03:02:36.711Z'),
    lastHeartbeatRecv: ISODate('1970-01-01T00:00:00.000Z'),
    pingMs: Long('0'),
    lastHeartbeatMessage: 'Error connecting to gpel10.cs.norou.edu:24157 (19.194.8.82:24157) :: caused by :: envoke :: caused by :: Connection refused',
    syncSourceHost: '',
    syncSourceId: 1,
    infoMessage: '',
    configVersion: -1,
    configTerm: -1,
    self: false
  },
  {
    _id: 2,
    name: 'gpel11.cs.norou.edu:24157',
    health: 1,
    state: 3,
    stateStr: 'PRIMARY',
    uptime: 21,
    optime: { ts: Timestamp({ t: 1744686156, i: 1 }), t: Long('3') },
    optimeDurable: { ts: Timestamp({ t: 1744686156, i: 1 }), t: Long('3') },
    optimeWritten: { ts: Timestamp({ t: 1744686156, i: 1 }), t: Long('3') }
  }
]

```

```

        optimeDate: ISODate('2025-04-15T03:02:36.000Z'),
        optimeDurableDate: ISODate('2025-04-15T03:02:36.000Z'),
        optimeWrittenDate: ISODate('2025-04-15T03:02:36.000Z'),
        lastAppliedWallTime: ISODate('2025-04-15T03:02:36.148Z'),
        lastDurableWallTime: ISODate('2025-04-15T03:02:36.148Z'),
        lastWrittenWallTime: ISODate('2025-04-15T03:02:36.148Z'),
        lastHeartbeat: ISODate('2025-04-15T03:02:37.159Z'),
        lastHeartbeatRecv: ISODate('2025-04-15T03:02:36.157Z'),
        pingMs: Long('0'),
        lastHeartbeatMessage: '',
        syncSourceHost: '',
        syncSourceId: -1,
        infoMessage: '',
        electionTime: Timestamp({ t: 1744686136, i: 1 }),
        electionDate: ISODate('2025-04-15T03:02:16.000Z'),
        configVersion: 1,
        configTerm: 3
    }
],
ok: 1,
'$clusterTime': {
    clusterTime: Timestamp({ t: 1744686156, i: 1 }),
    signature: {
        hash: Binary.createFromBase64('AAAAAAAAAAAAAAAAAAAAAAA='),
        keyId: Long('0')
    }
},
operationTime: Timestamp({ t: 1744686156, i: 1 })
}
theDailyReplica [direct: secondary] test> █

```

### Query 1: Insert a (news or feature) article into the archive

```

theDailyReplica [direct: primary] test> use dailyNewsArchive
switched to db dailyNewsArchive
theDailyReplica [direct: primary] dailyNewsArchive> db.articles.insertOne({
...     article_type: "feature article",
...     article_section: "Health",
...     article_id: 57,
...     post_on_date: { day: 2, month: 3, year: 2023, time: "00:00:00" },
...     reporter_name: "Suresh Raina",
...     similar_stories: ["https://thedaily.com/health/story-57.html"],
...     num_times_read: 5643,
...     article_text: "A new report reveals major healthcare trends for 2023.",
...     comments: []
... })
{
    acknowledged: true,
    insertedId: ObjectId('67fdced967fb77e11b6b140b')
}

```

```
theDailyReplica [direct: primary] dailyNewsArchive> db.articles.find({ article_id: 57 }).pretty()
[
  {
    _id: ObjectId('67fdced967fb77e11b6b140b'),
    article_type: 'feature article',
    article_section: 'Health',
    article_id: 57,
    post_on_date: { day: 2, month: 3, year: 2023, time: '00:00:00' },
    reporter_name: 'Suresh Raina',
    similar_stories: [ 'https://thedaily.com/health/story-57.html' ],
    num_times_read: 5643,
    article_text: 'A new report reveals major healthcare trends for 2023.',
    comments: []
  }
]
theDailyReplica [direct: primary] dailyNewsArchive> █
```

## Query 2: Retrieve the latest news

```
theDailyReplica [direct: primary] dailyNewsArchive> db.articles.find().sort({
...   "post_on_date.year": -1,
...   "post_on_date.month": -1,
...   "post_on_date.day": -1,
...   "post_on_date.time": -1
[... ]).limit(1).pretty()
[
  {
    _id: ObjectId('67fc87f295be239aba6b140d'),
    article_type: 'feature article',
    article_section: 'Technology',
    article_id: 1002,
    post_on_date: { day: 12, month: 4, year: 2025, time: '15:00:00' },
    reporters: [
      { name: 'Virat Kohli', location: 'Delhi' },
      { name: 'Rohit Sharma', location: 'Mumbai' },
      { name: 'KL Rahul', location: 'Bangalore' },
      { name: 'MS Dhoni', location: 'Chennai' }
    ],
    similar_stories: [],
    num_times_read: 2600,
    article_text: 'This article explores emerging cloud platforms.',
    comments: []
  }
]
theDailyReplica [direct: primary] dailyNewsArchive> █
```

### Query 3: Retrieve the top 10 most read news in the past month

```
theDailyReplica [direct: primary] dailyNewsArchive> db.articles.find({
...   "post_on_date.year": 2023,
...   "post_on_date.month": 1
... }).sort({ num_times_read: -1 }).limit(10).pretty()
```

```
[
  {
    _id: ObjectId('67fc77245b76bea4836b140b'),
    article_type: 'news article',
    article_section: 'Technology',
    article_id: 1,
    post_on_date: { day: 2, month: 1, year: 2023, time: '00:00:00' },
    reporter_name: 'Reporter 1',
    similar_stories: [
      'https://thedaily.com/politics/election-story-1.html',
      'https://thedaily.com/sports/election-story-1.html'
    ],
    num_times_read: 18964,
    article_text: 'This article discusses the latest developments in election.',
    comments: [
      {
        comment_id: 1,
        article_id: 1,
        user_id: 'user1@mail.com',
        posted_on_date: { day: 2, month: 1, year: 2023, time: '00:00:00' },
        comment_text: 'Well-written and concise.',
        score: -77
      }
    ]
  },
  {
    _id: ObjectId('67fc77265b76bea4836b141a'),
    article_type: 'news article',
    article_section: 'Health',
    article_id: 16,
    post_on_date: { day: 17, month: 1, year: 2023, time: '00:00:00' },
    reporter_name: 'Reporter 16',
    similar_stories: [
      'https://thedaily.com/politics/budget-story-16.html',
      'https://thedaily.com/sports/budget-story-16.html'
    ],
    num_times_read: 16892,
    article_text: 'This article discusses the latest developments in budget.',
    comments: [
      {
        comment_id: 16,
        article_id: 16,
        user_id: 'user16@mail.com',
        posted_on_date: { day: 17, month: 1, year: 2023, time: '00:00:00' },
        comment_text: 'I disagree with the viewpoint.',
        score: 99
      }
    ]
  }
],
```

```
{
  _id: ObjectId('67fc77275b76bea4836b141b'),
  article_type: 'feature article',
  article_section: 'Politics',
  article_id: 17,
  post_on_date: { day: 18, month: 1, year: 2023, time: '00:00:00' },
  reporter_name: 'Reporter 17',
  similar_stories: [
    'https://thedaily.com/politics/space-story-17.html',
    'https://thedaily.com/sports/space-story-17.html'
  ],
  num_times_read: 16364,
  article_text: 'This article discusses the latest developments in space.',
  comments: [
    {
      comment_id: 17,
      article_id: 17,
      user_id: 'user17@mail.com',
      posted_on_date: { day: 18, month: 1, year: 2023, time: '00:00:00' },
      comment_text: 'Amazing insight!',
      score: 40
    }
  ]
},
{
  _id: ObjectId('67fc77295b76bea4836b1428'),
  article_type: 'feature article',
  article_section: 'Entertainment',
  article_id: 30,
  post_on_date: { day: 31, month: 1, year: 2023, time: '00:00:00' },
  reporter_name: 'Reporter 30',
  similar_stories: [
    'https://thedaily.com/politics/AI-story-30.html',
    'https://thedaily.com/sports/AI-story-30.html'
  ],
  num_times_read: 16109,
  article_text: 'This article discusses the latest developments in AI.',
  comments: [
    {
      comment_id: 30,
      article_id: 30,
      user_id: 'user30@mail.com',
      posted_on_date: { day: 31, month: 1, year: 2023, time: '00:00:00' },
      comment_text: 'Well-written and concise.',
      score: 72
    }
  ]
},
```

```
{
  _id: ObjectId('67fc77255b76bea4836b1411'),
  article_type: 'feature article',
  article_section: 'Technology',
  article_id: 7,
  post_on_date: { day: 8, month: 1, year: 2023, time: '00:00:00' },
  reporter_name: 'Reporter 7',
  similar_stories: [
    'https://thedaily.com/politics/education-story-7.html',
    'https://thedaily.com/sports/education-story-7.html'
  ],
  num_times_read: 14545,
  article_text: 'This article discusses the latest developments in education.',
  comments: [
    {
      comment_id: 7,
      article_id: 7,
      user_id: 'user7@mail.com',
      posted_on_date: { day: 8, month: 1, year: 2023, time: '00:00:00' },
      comment_text: 'Not convinced with the facts.',
      score: 19
    }
  ]
},
{
  _id: ObjectId('67fc77255b76bea4836b1414'),
  article_type: 'news article',
  article_section: 'Sports',
  article_id: 10,
  post_on_date: { day: 11, month: 1, year: 2023, time: '00:00:00' },
  reporter_name: 'Reporter 10',
  similar_stories: [
    'https://thedaily.com/politics/film-story-10.html',
    'https://thedaily.com/sports/film-story-10.html'
  ],
  num_times_read: 11859,
  article_text: 'This article discusses the latest developments in film.',
  comments: [
    {
      comment_id: 10,
      article_id: 10,
      user_id: 'user10@mail.com',
      posted_on_date: { day: 11, month: 1, year: 2023, time: '00:00:00' },
      comment_text: 'Keep up the good work!',
      score: -5
    }
  ]
},
```

```
{
  _id: ObjectId('67fc77255b76bea4836b1410'),
  article_type: 'feature article',
  article_section: 'Politics',
  article_id: 6,
  post_on_date: { day: 7, month: 1, year: 2023, time: '00:00:00' },
  reporter_name: 'Reporter 6',
  similar_stories: [
    'https://thedaily.com/politics/AI-story-6.html',
    'https://thedaily.com/sports/AI-story-6.html'
  ],
  num_times_read: 10076,
  article_text: 'This article discusses the latest developments in AI.',
  comments: [
    {
      comment_id: 6,
      article_id: 6,
      user_id: 'user6@mail.com',
      posted_on_date: { day: 7, month: 1, year: 2023, time: '00:00:00' },
      comment_text: 'Needs better research.',
      score: -43
    }
  ]
},
{
  _id: ObjectId('67fc77255b76bea4836b1412'),
  article_type: 'news article',
  article_section: 'Technology',
  article_id: 8,
  post_on_date: { day: 9, month: 1, year: 2023, time: '00:00:00' },
  reporter_name: 'Reporter 8',
  similar_stories: [
    'https://thedaily.com/politics/cricket-story-8.html',
    'https://thedaily.com/sports/cricket-story-8.html'
  ],
  num_times_read: 9871,
  article_text: 'This article discusses the latest developments in cricket.',
  comments: [
    {
      comment_id: 8,
      article_id: 8,
      user_id: 'user8@mail.com',
      posted_on_date: { day: 9, month: 1, year: 2023, time: '00:00:00' },
      comment_text: 'Very informative piece.',
      score: 97
    }
  ]
},
```

```
{
  _id: ObjectId('67fc77285b76bea4836b1423'),
  article_type: 'news article',
  article_section: 'Entertainment',
  article_id: 25,
  post_on_date: { day: 26, month: 1, year: 2023, time: '00:00:00' },
  reporter_name: 'Reporter 25',
  similar_stories: [
    'https://thedaily.com/politics/election-story-25.html',
    'https://thedaily.com/sports/election-story-25.html'
  ],
  num_times_read: 9833,
  article_text: 'This article discusses the latest developments in election.',
  comments: [
    {
      comment_id: 25,
      article_id: 25,
      user_id: 'user25@mail.com',
      posted_on_date: { day: 26, month: 1, year: 2023, time: '00:00:00' },
      comment_text: 'Bold take on current affairs.',
      score: 86
    }
  ]
},
{
  _id: ObjectId('67fc77245b76bea4836b140d'),
  article_type: 'feature article',
  article_section: 'Sports',
  article_id: 3,
  post_on_date: { day: 4, month: 1, year: 2023, time: '00:00:00' },
  reporter_name: 'Reporter 3',
  similar_stories: [
    'https://thedaily.com/politics/budget-story-3.html',
    'https://thedaily.com/sports/budget-story-3.html'
  ],
  num_times_read: 9484,
  article_text: 'This article discusses the latest developments in budget.',
  comments: [
    {
      comment_id: 3,
      article_id: 3,
      user_id: 'user3@mail.com',
      posted_on_date: { day: 4, month: 1, year: 2023, time: '00:00:00' },
      comment_text: 'Not convinced with the facts.',
      score: 4
    }
  ]
}
]
theDailyReplica [direct: primary] dailyNewsArchive> █
```

## Query 4: Add a comment to an article

```
theDailyReplica [direct: primary] dailyNewsArchive> db.articles.updateOne(  
...   { article_id: 57 },  
...   { $push: {  
...     comments: {  
...       comment_id: 102,  
...       article_id: 57,  
...       user_id: "user102@mail.com",  
...       posted_on_date: { day: 2, month: 3, year: 2023, time: "12:00:00" },  
...       comment_text: "Very informative piece!",  
...       score: 15  
...     }  
...   }  
... }  
{  
  acknowledged: true,  
  insertedId: null,  
  matchedCount: 1,  
  modifiedCount: 1,  
  upsertedCount: 0  
}
```

```
theDailyReplica [direct: primary] dailyNewsArchive> db.articles.find({ article_id: 57 }, { comments: 1, _id: 0 }).pretty()  
[  
  {  
    comments: [  
      {  
        comment_id: 102,  
        article_id: 57,  
        user_id: 'user102@mail.com',  
        posted_on_date: { day: 2, month: 3, year: 2023, time: '12:00:00' },  
        comment_text: 'Very informative piece!',  
        score: 15  
      }  
    ]  
  }  
]  
theDailyReplica [direct: primary] dailyNewsArchive> █
```

**Query 5:** Add a story to the ‘similar stories’ section for articles containing a given word and published within the past 2 years

```
theDailyReplica [direct: primary] dailyNewsArchive> db.articles.updateMany(  
...   {  
...     article_text: { $regex: /health/i },  
...     $or: [  
...       { "post_on_date.year": 2023 },  
...       { "post_on_date.year": 2024 },  
...       { "post_on_date.year": 2025 }  
...     ]  
...   },  
...   {  
...     $addToSet: {  
...       similar_stories: "https://thedaily.com/health/global-health-initiative.html"  
...     }  
...   }  
... )  
{  
  acknowledged: true,  
  insertedId: null,  
  matchedCount: 2,  
  modifiedCount: 2,  
  upsertedCount: 0  
}  
  
theDailyReplica [direct: primary] dailyNewsArchive> db.articles.find({  
...   similar_stories: "https://thedaily.com/health/global-health-initiative.html"  
... }).pretty()  
[  
  {  
    _id: ObjectId('67fc87b295be239aba6b140c'),  
    article_type: 'news article',  
    article_section: 'Health',  
    article_id: 1001,  
    post_on_date: { day: 12, month: 4, year: 2025, time: '14:00:00' },  
    reporters: [  
      { name: 'MS Dhoni', location: 'Chennai' },  
      { name: 'Suresh Raina', location: 'Ghaziabad' }  
    ],  
    similar_stories: [ 'https://thedaily.com/health/global-health-initiative.html' ],  
    num_times_read: 1500,  
    article_text: 'This article discusses healthcare AI in India.',  
    comments: []  
  },  
  {  
    _id: ObjectId('67fdced967fb77e11b6b140b'),  
    article_type: 'feature article',  
    article_section: 'Health',  
    article_id: 57,  
    post_on_date: { day: 2, month: 3, year: 2023, time: '00:00:00' },  
    reporter_name: 'Suresh Raina',  
    similar_stories: [  
      'https://thedaily.com/health/story-57.html',  
      'https://thedaily.com/health/global-health-initiative.html'  
    ],  
    num_times_read: 5643,  
    article_text: 'A new report reveals major healthcare trends for 2023.',  
    comments: [  
      {  
        comment_id: 102,  
        article_id: 57,  
        user_id: 'user102@mail.com',  
        posted_on_date: { day: 2, month: 3, year: 2023, time: '12:00:00' },  
        comment_text: 'Very informative piece!',  
        score: 15  
      }  
    ]  
  }  
]  
theDailyReplica [direct: primary] dailyNewsArchive>
```

## **REFERENCES:**

- 1. Oracle 23ai Official Documentation:** Oracle. (2024). Introduction to Transactions. Oracle 23ai Database Concepts.
- 2. MongoDB Official Documentation Citation:** MongoDB Inc. (n.d.). ACID Transactions in MongoDB. MongoDB Documentation.