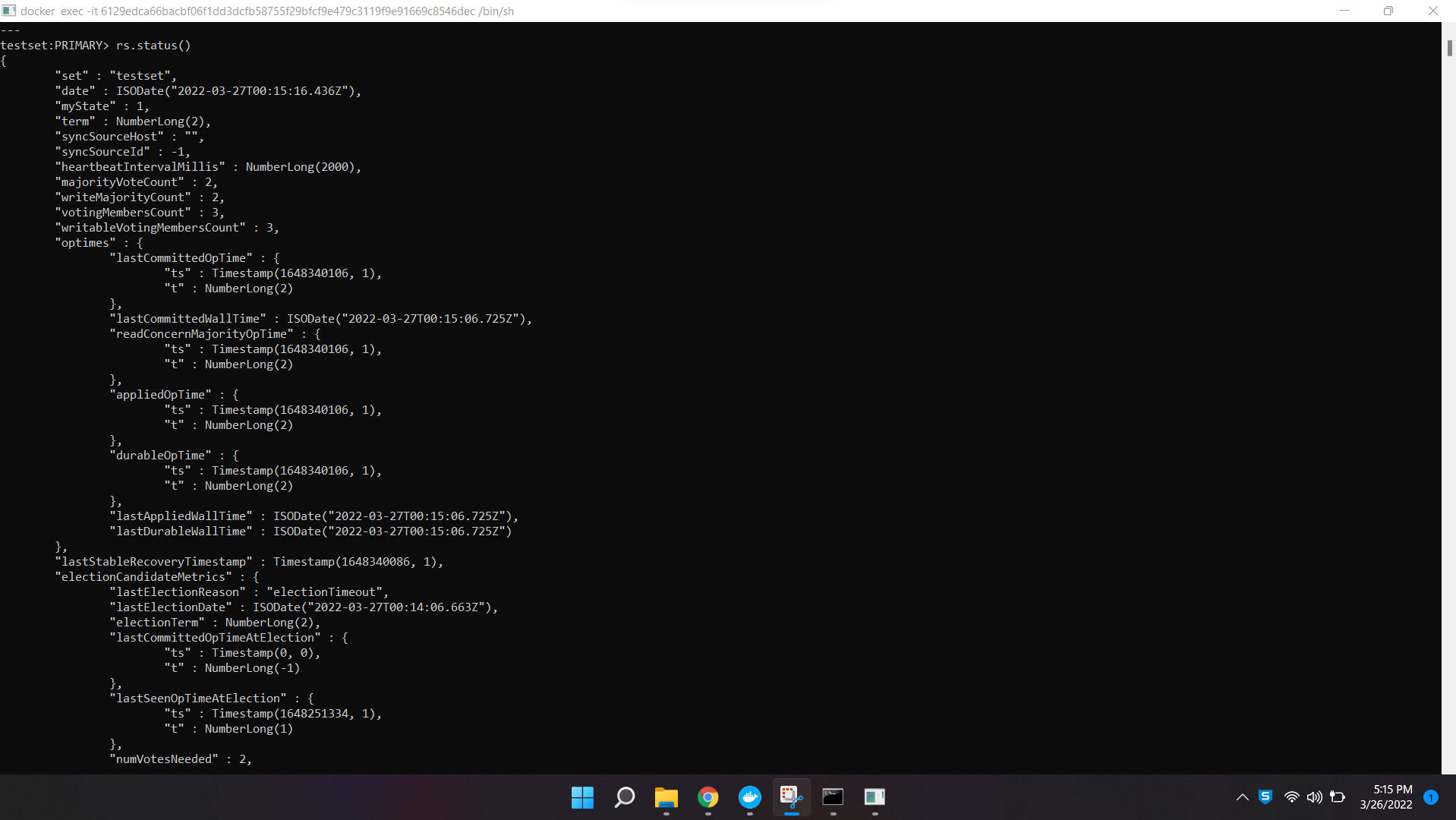
#### **Spring 2022 CS157C: NoSQL Database Systems**

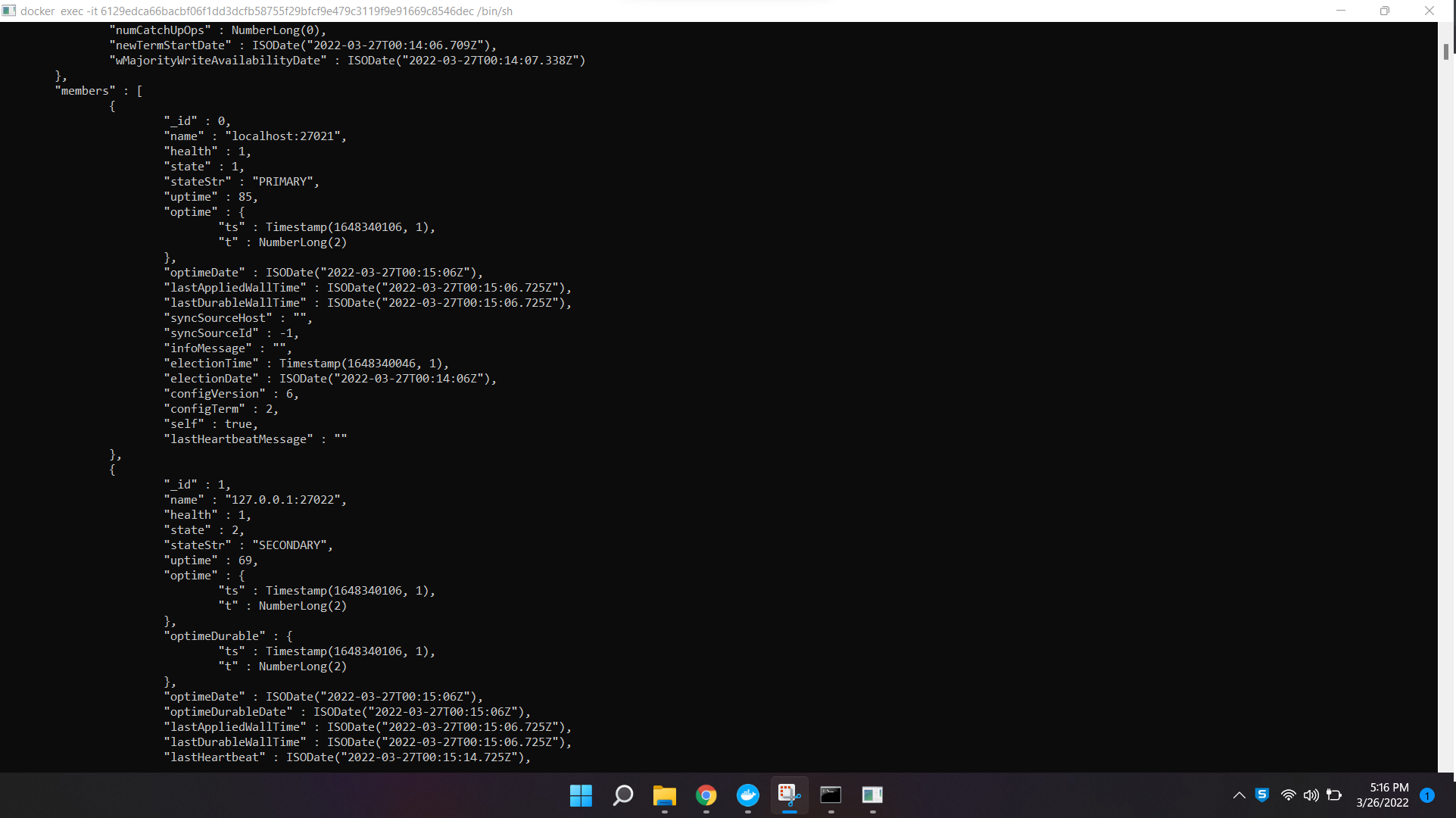
Programming Assignment 3

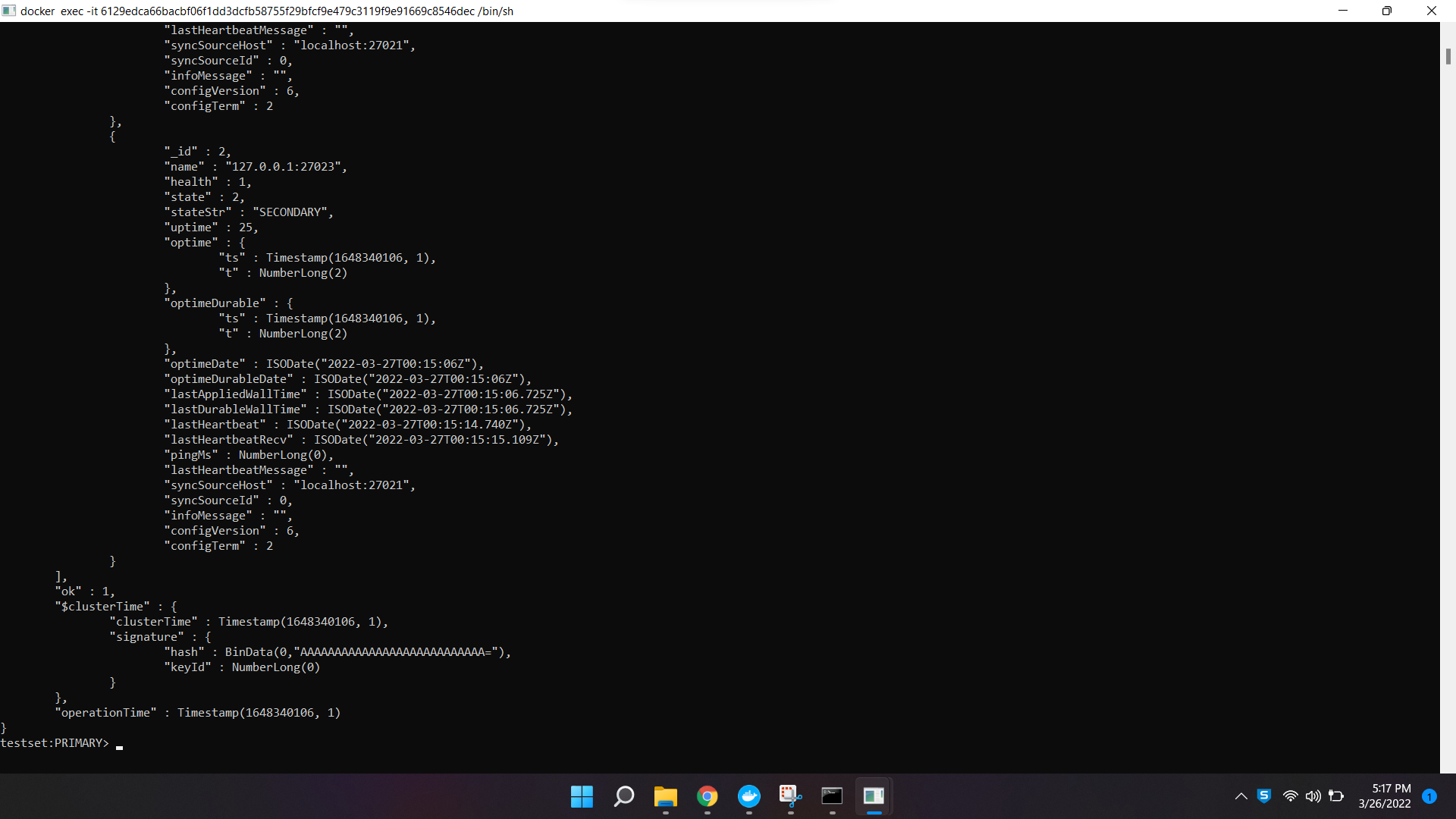
Task 1:

Screen1:

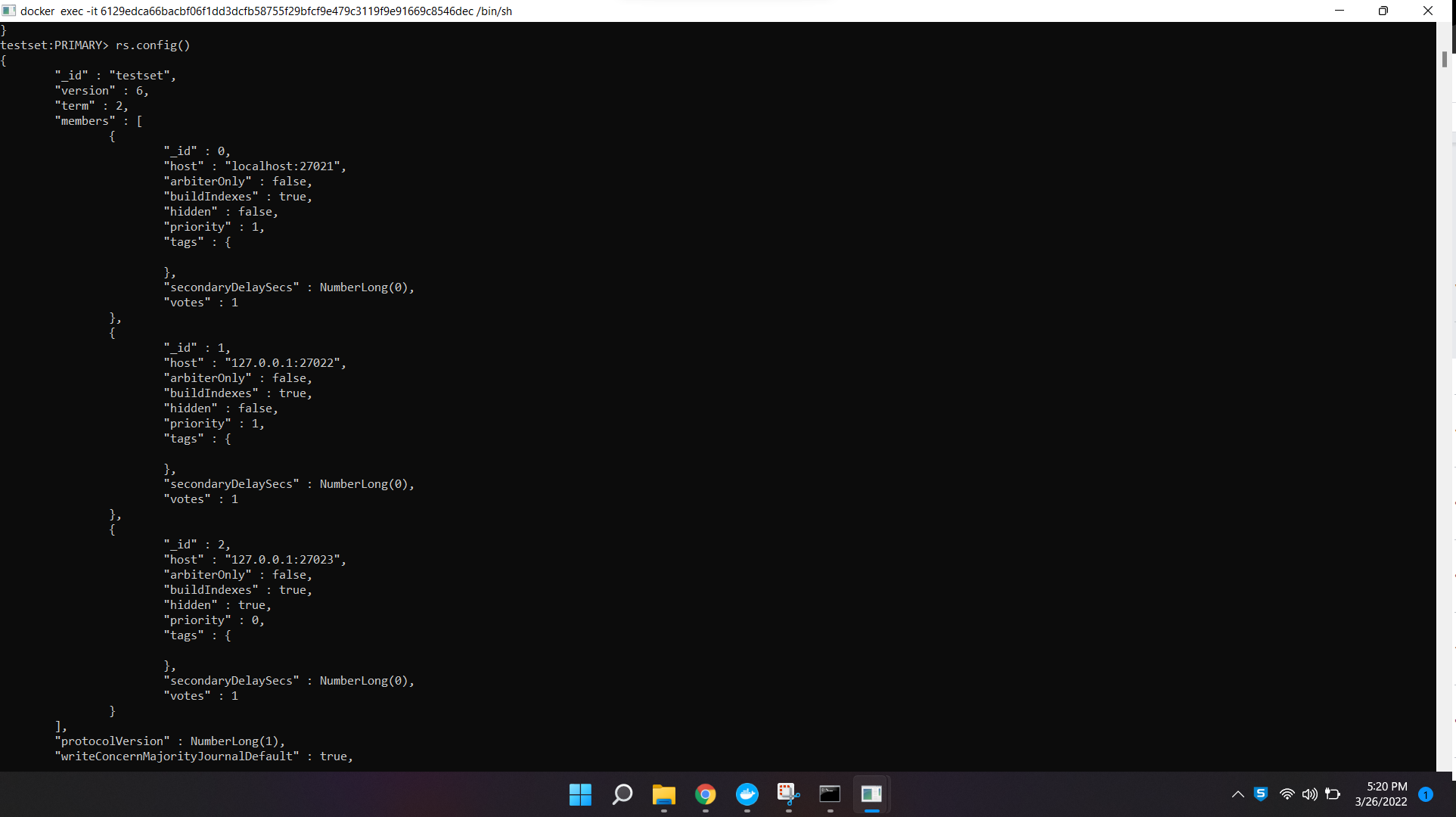
rs.status()

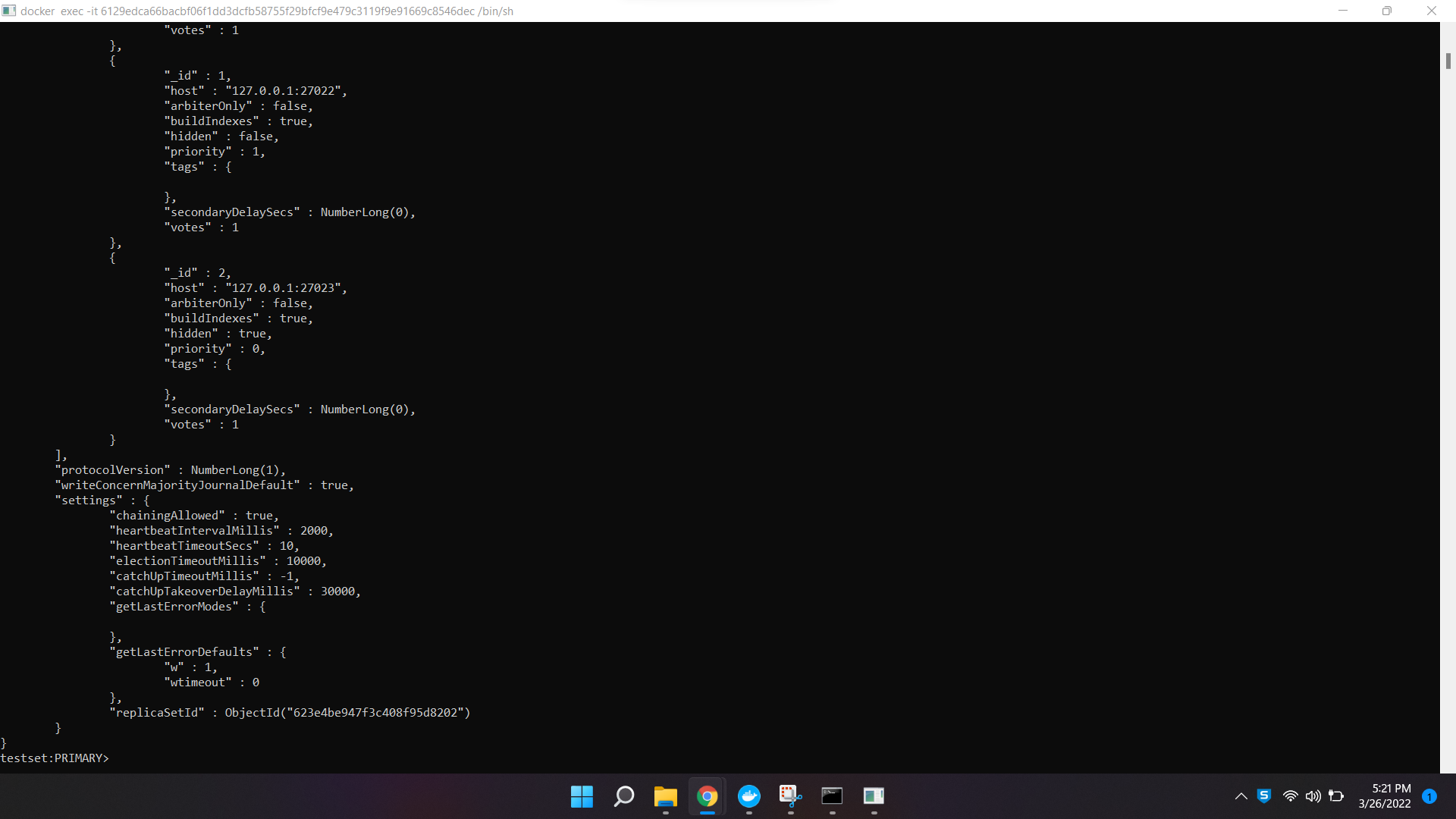




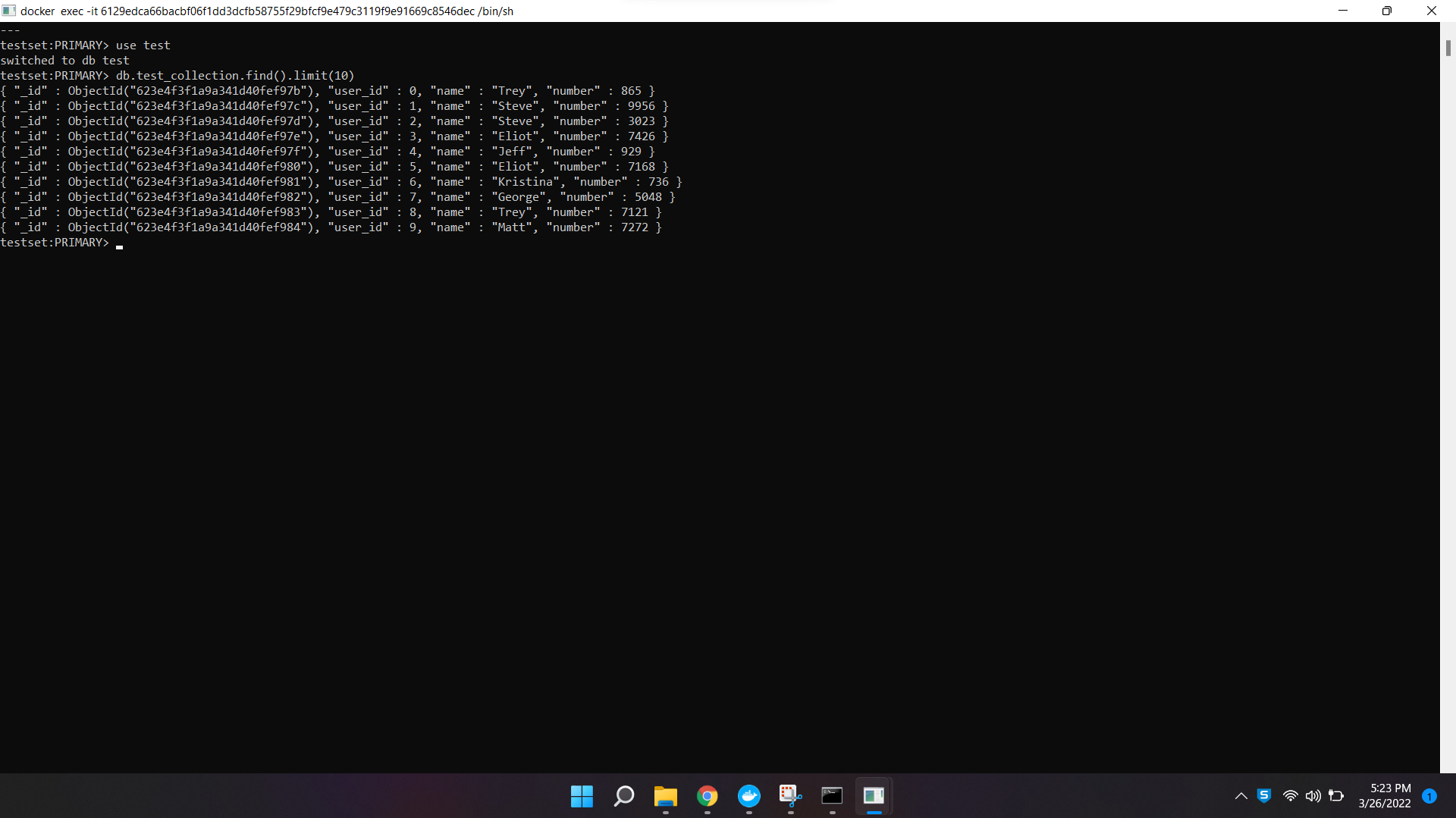


rs.config()

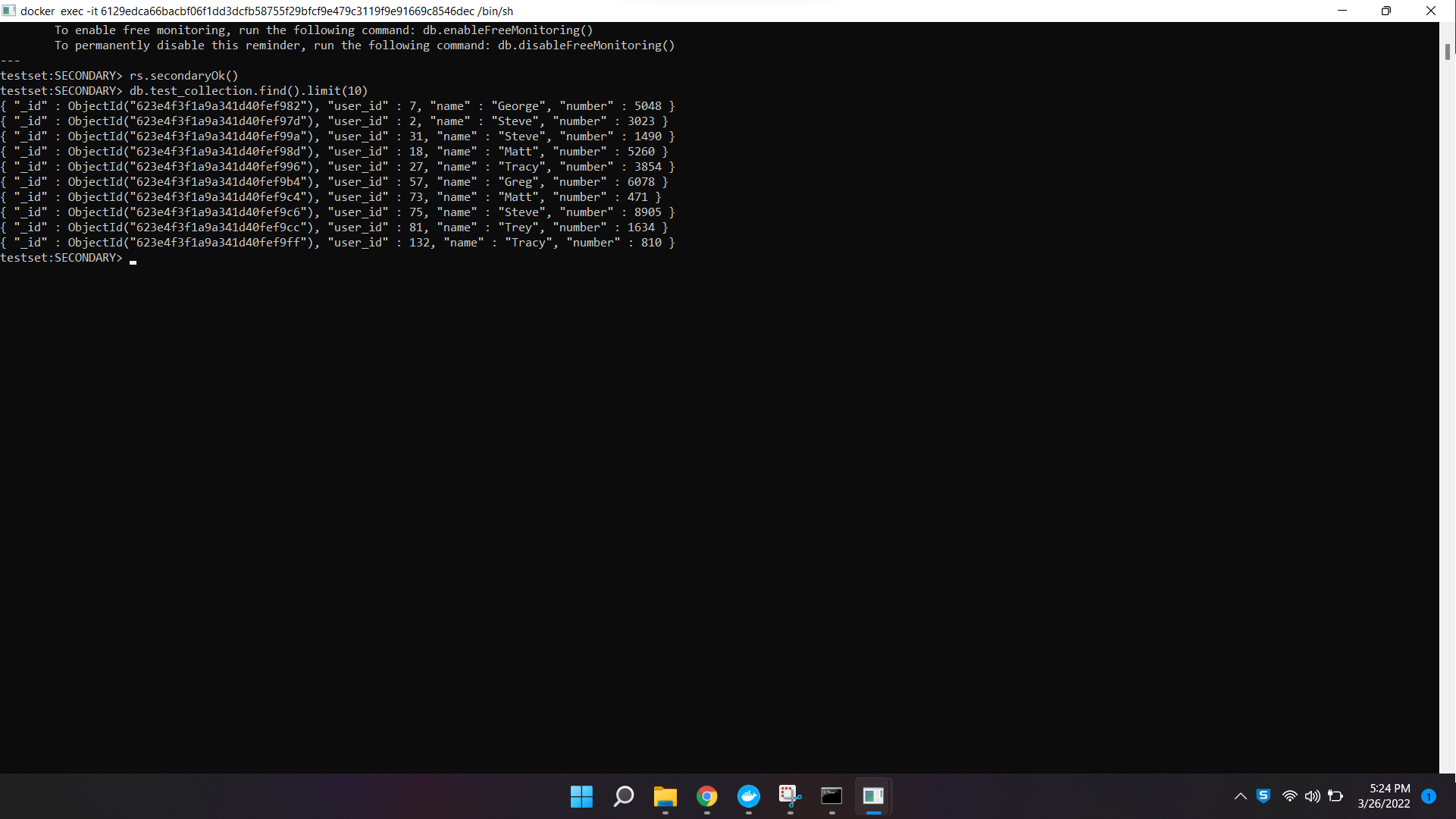




Screen 2:



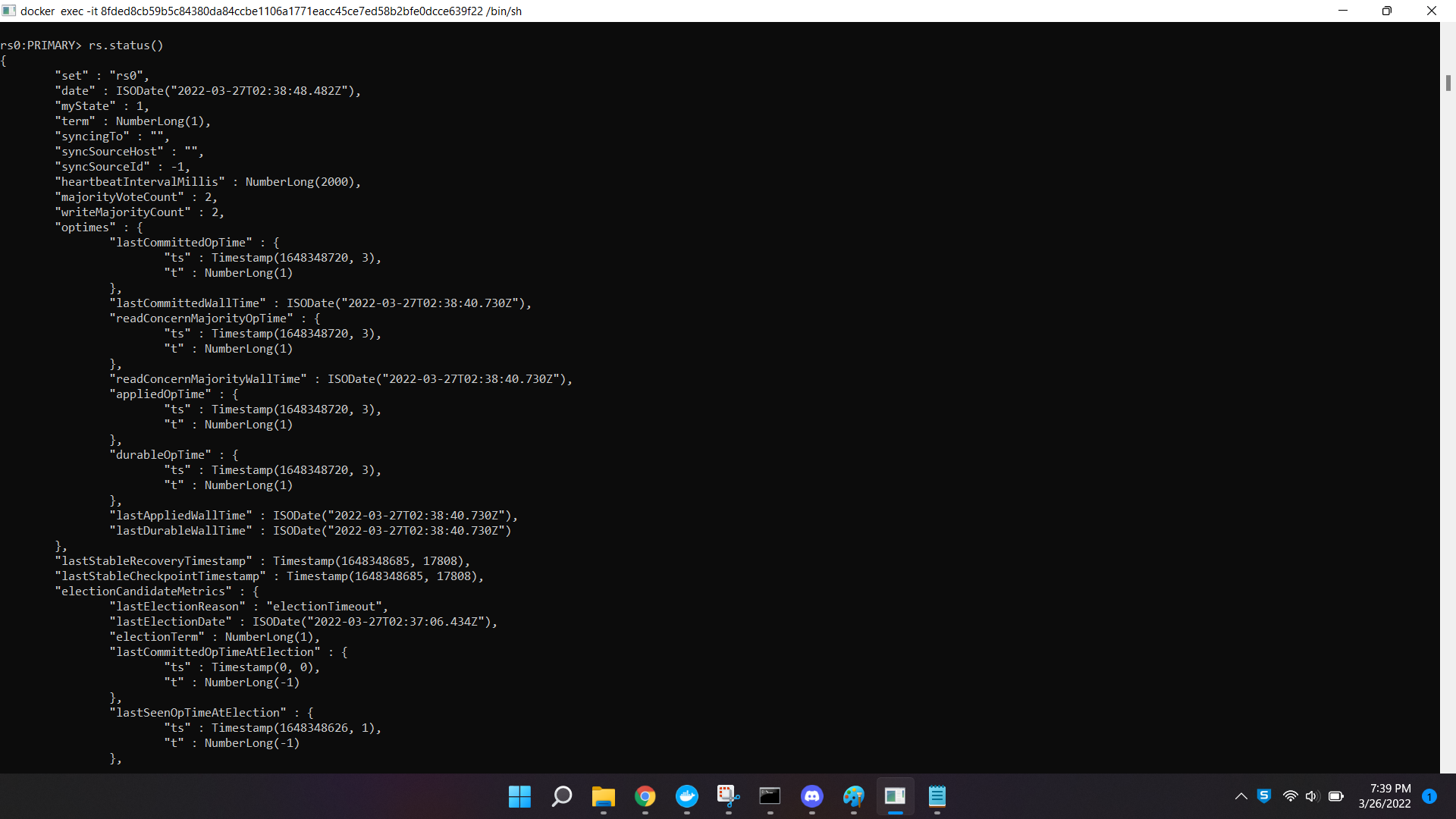
Screen 3:

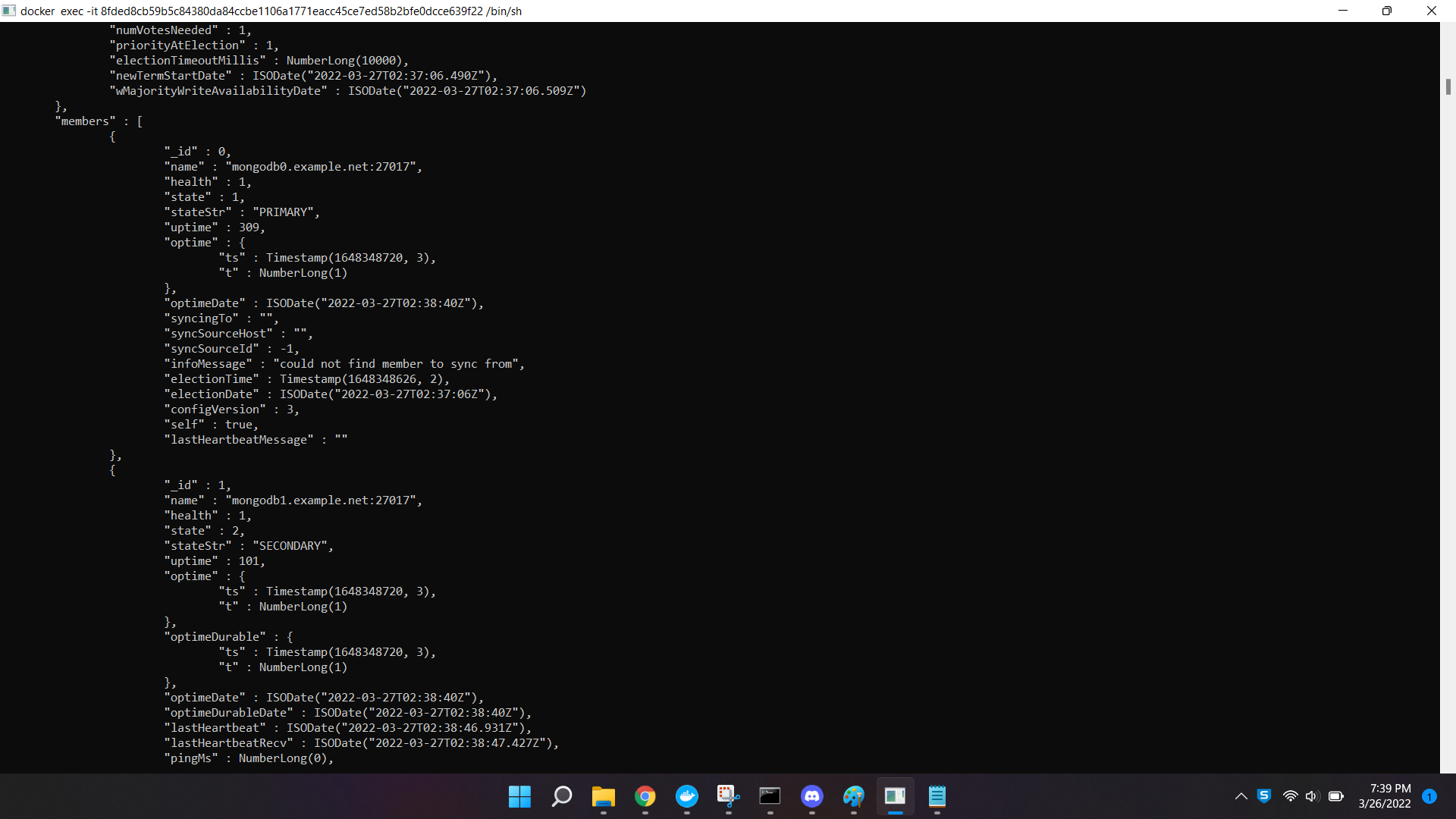


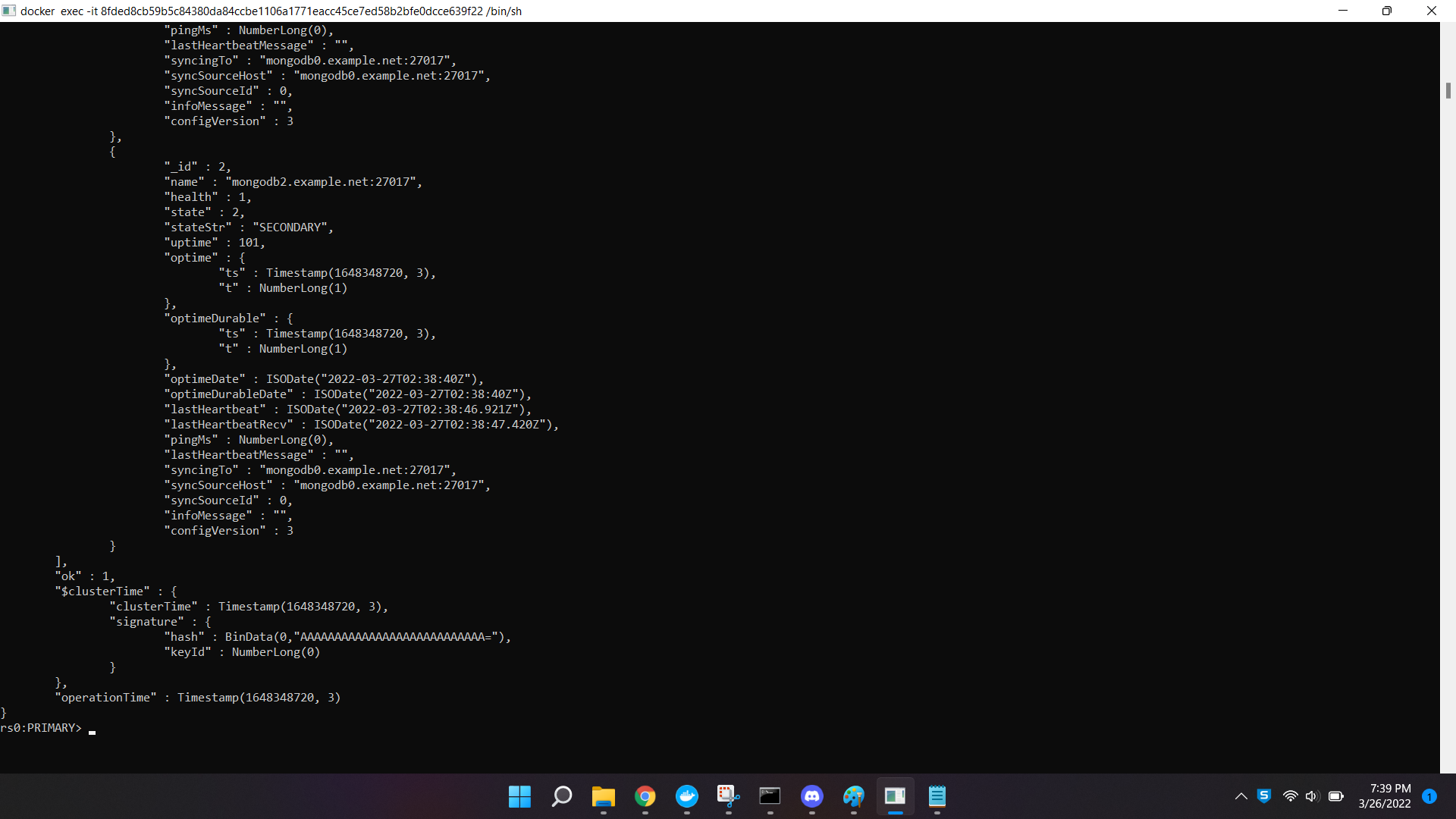
Task 2

Screen 4:

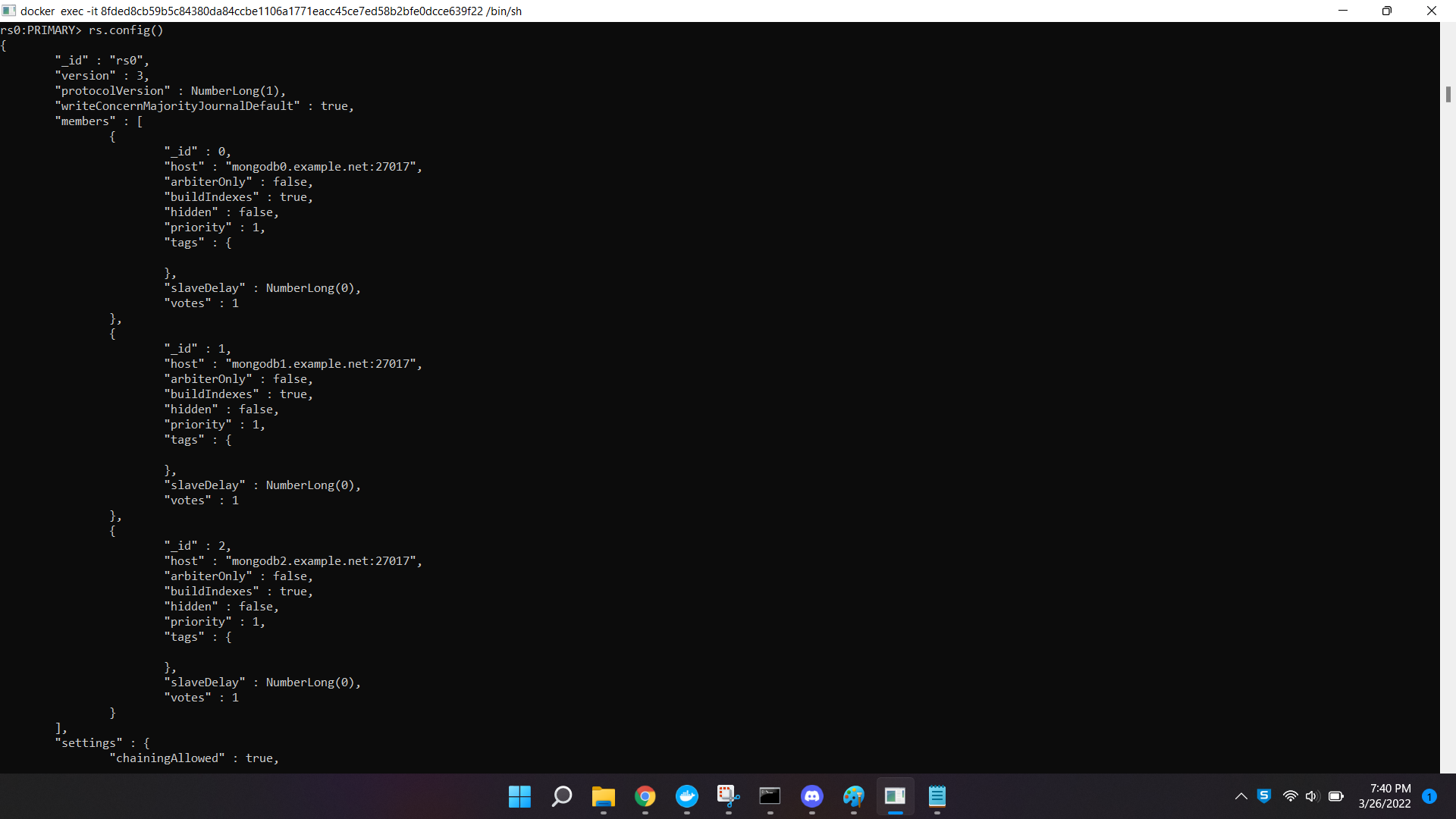
rs.status():

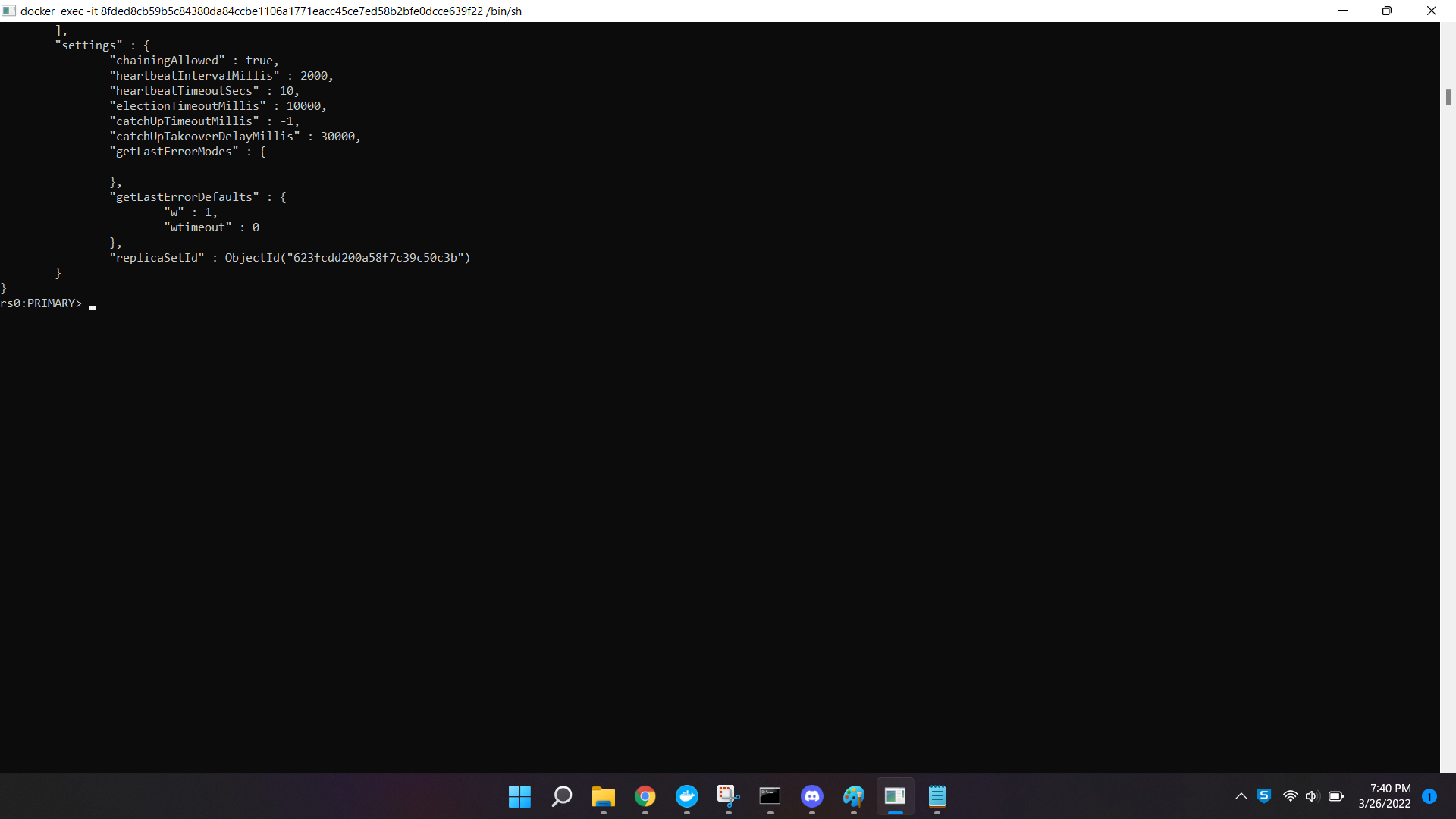




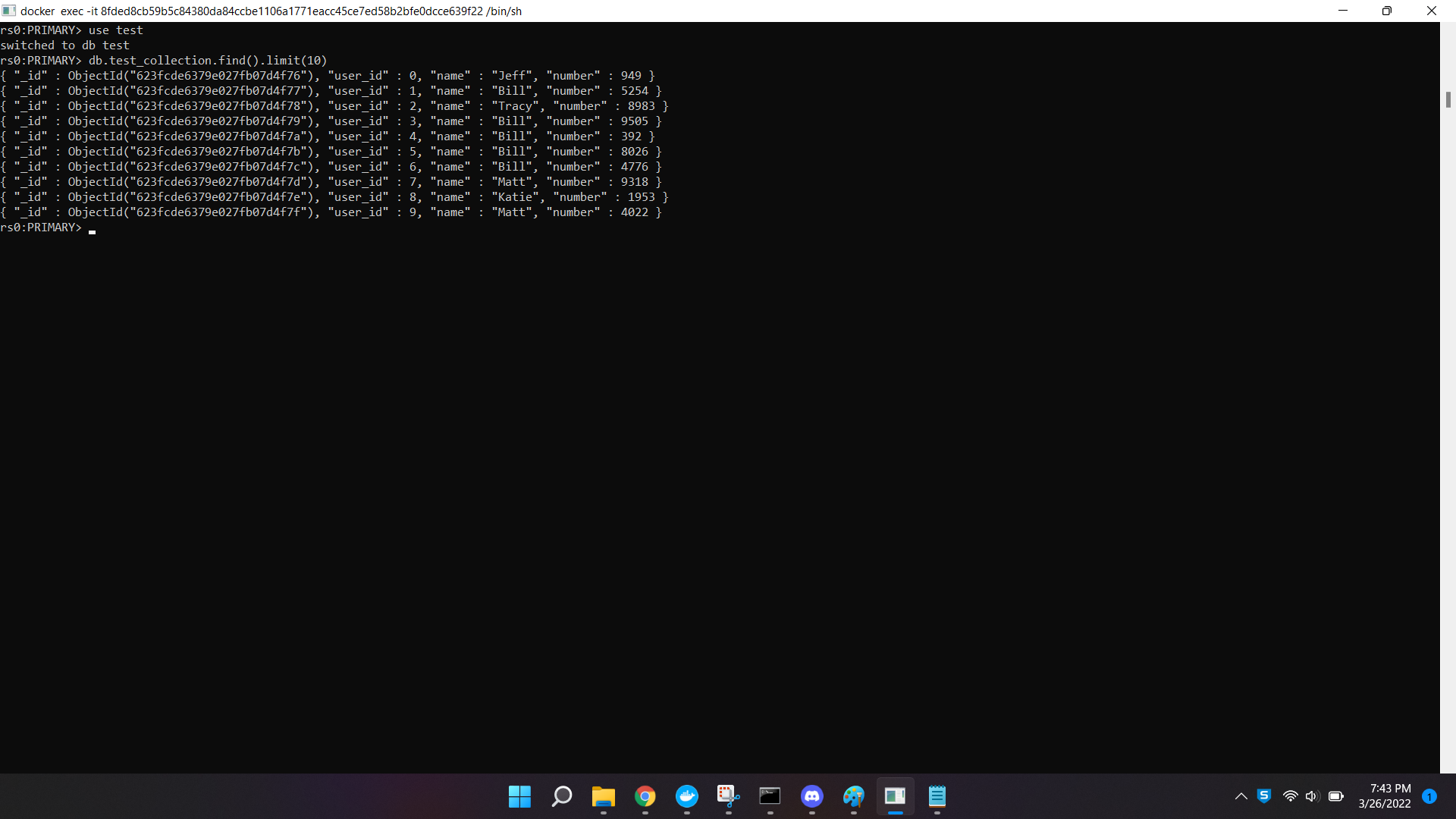


rs.config():

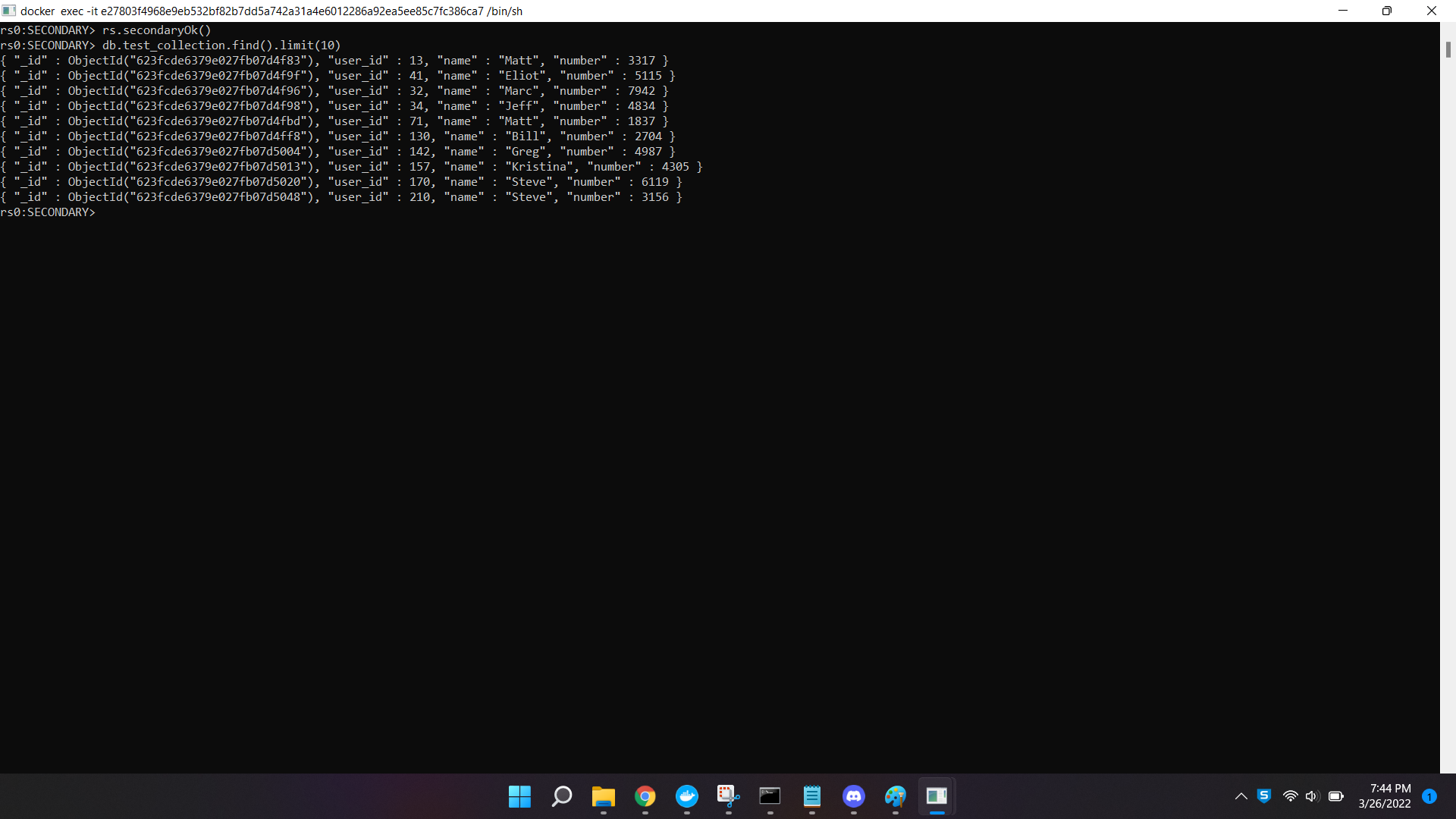




Screen 5



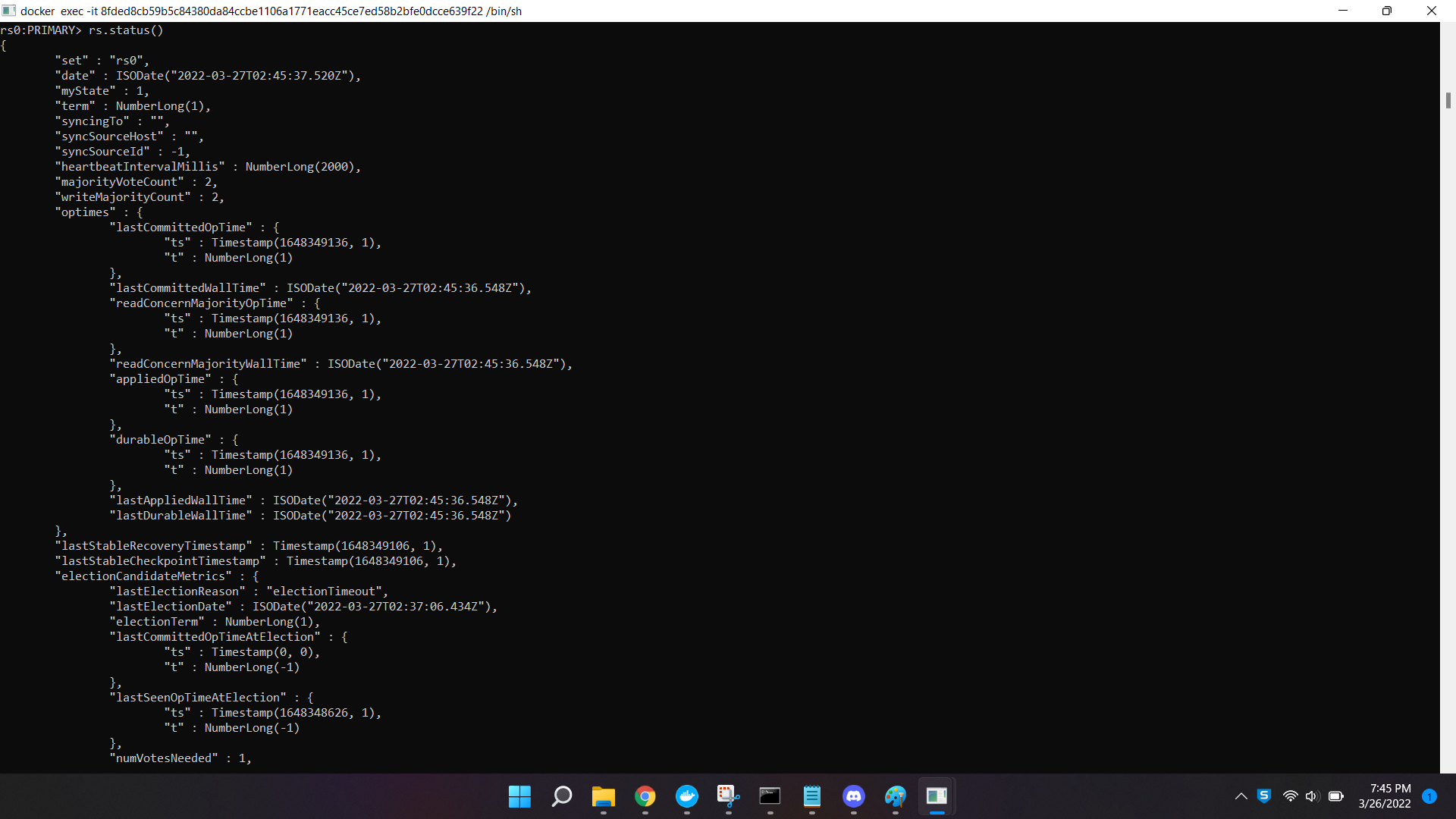
Screen 6:

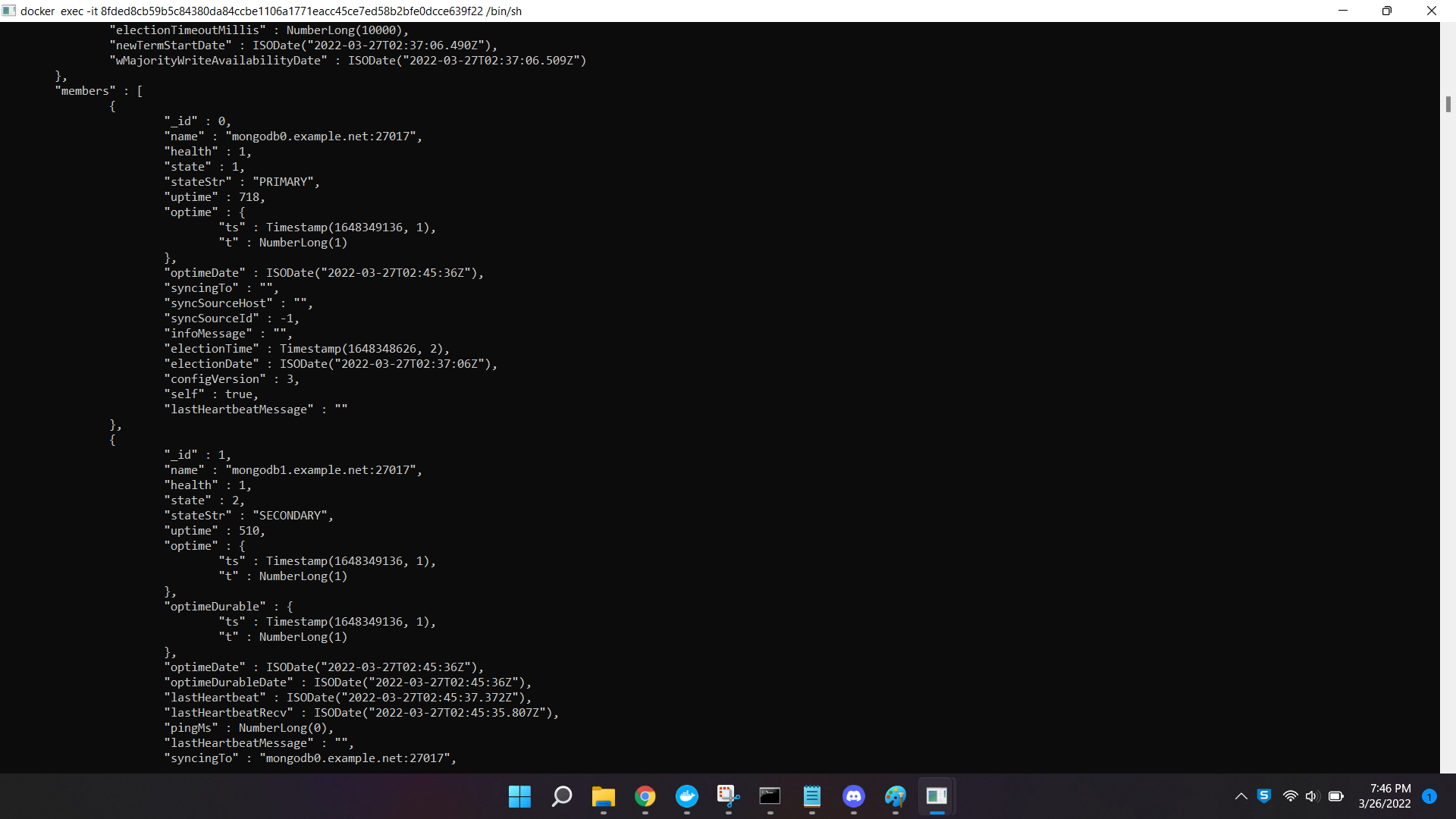


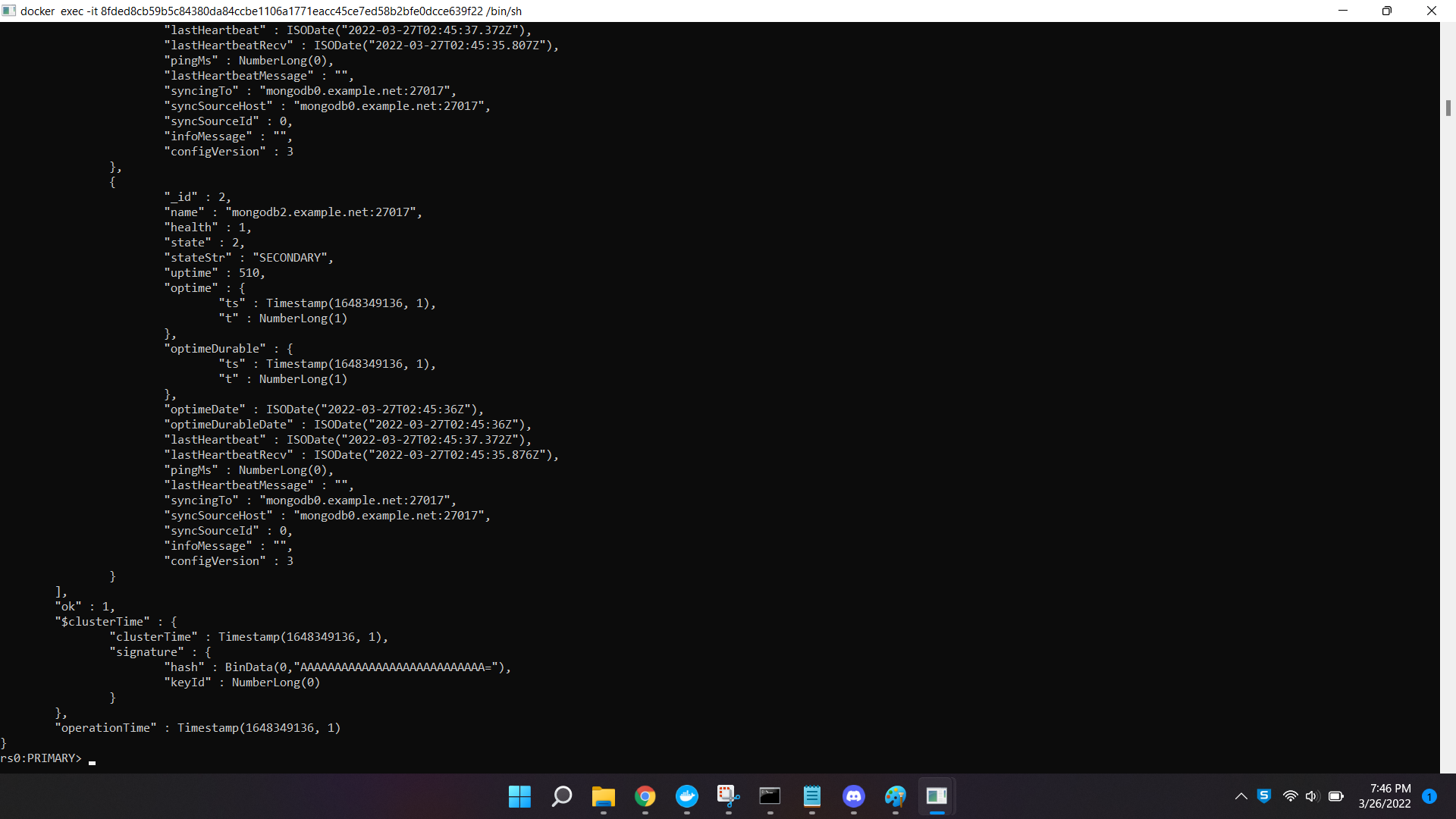
Screen 7:

rs.status() before stepping down the primary

The hostname of the primary: **"mongodb0.example.net:27017”**



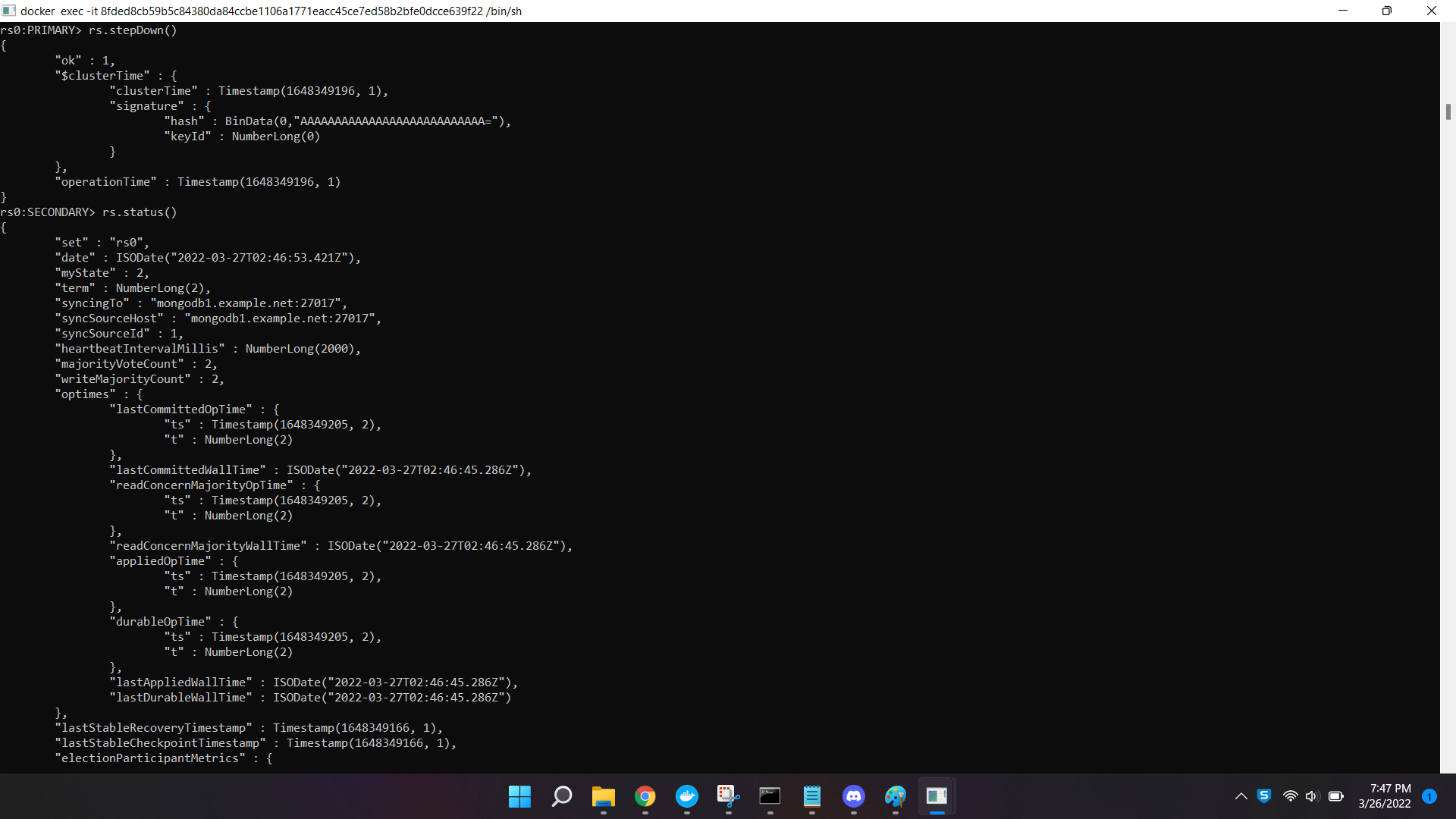


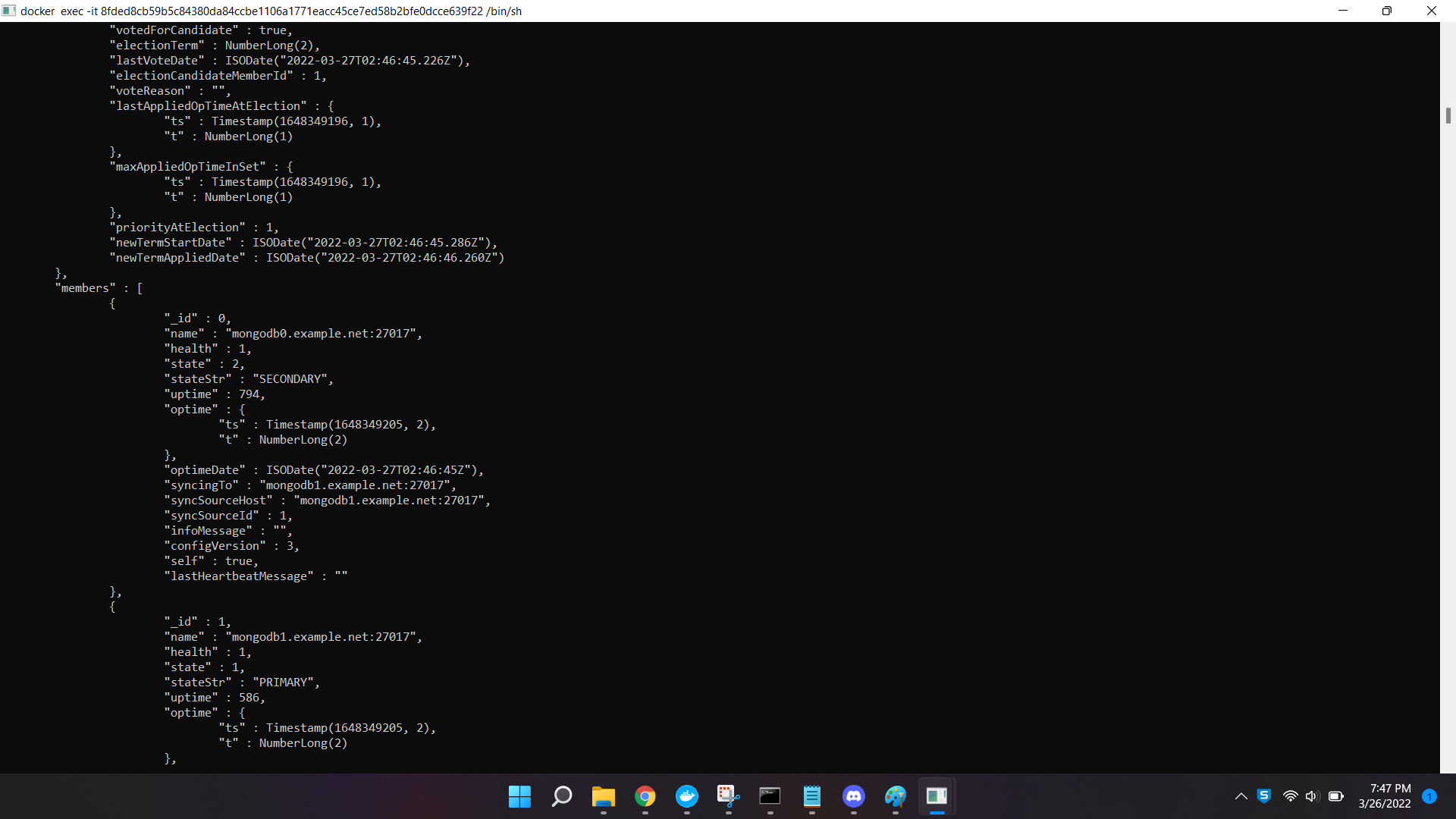


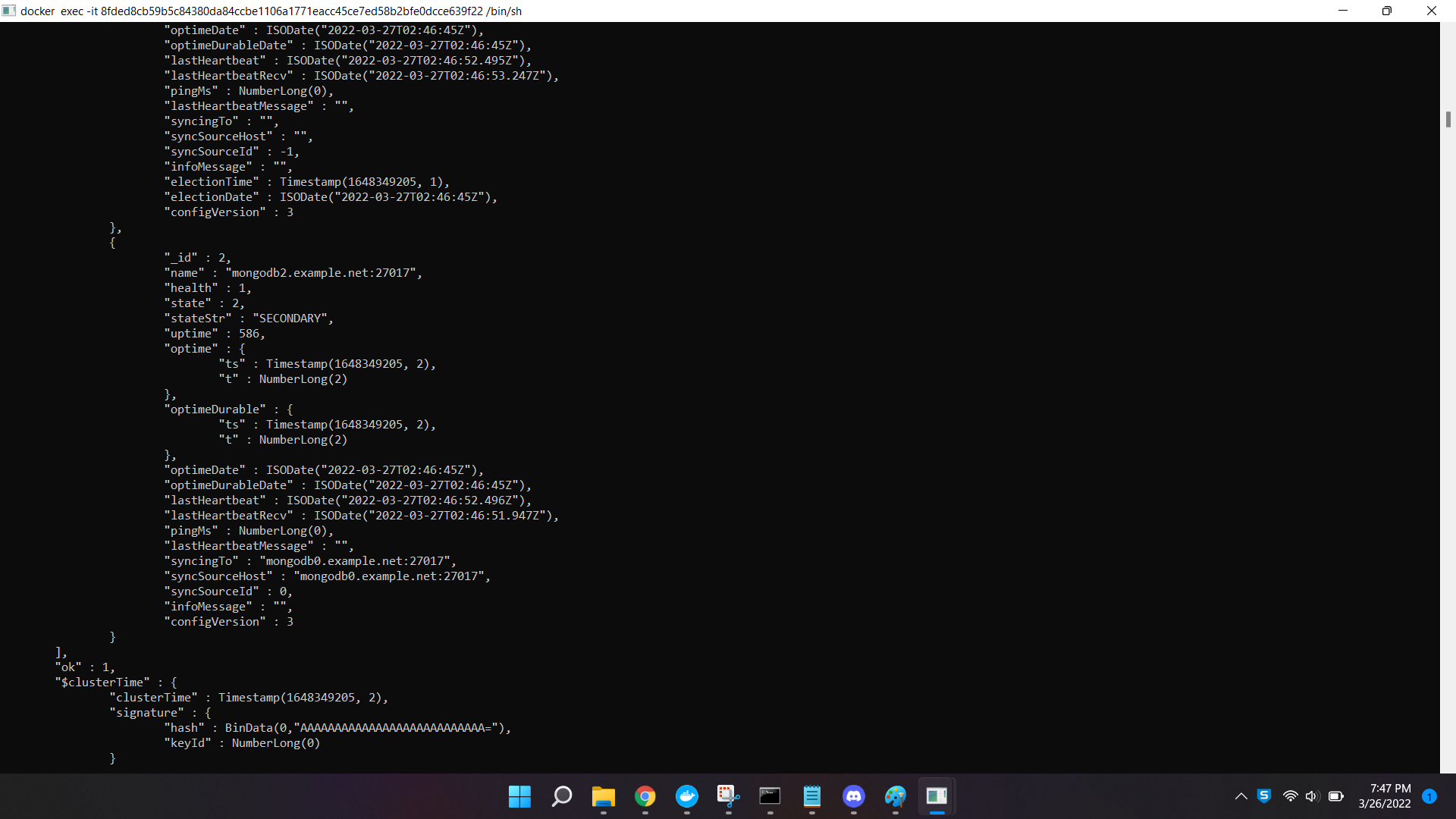
Screen 8

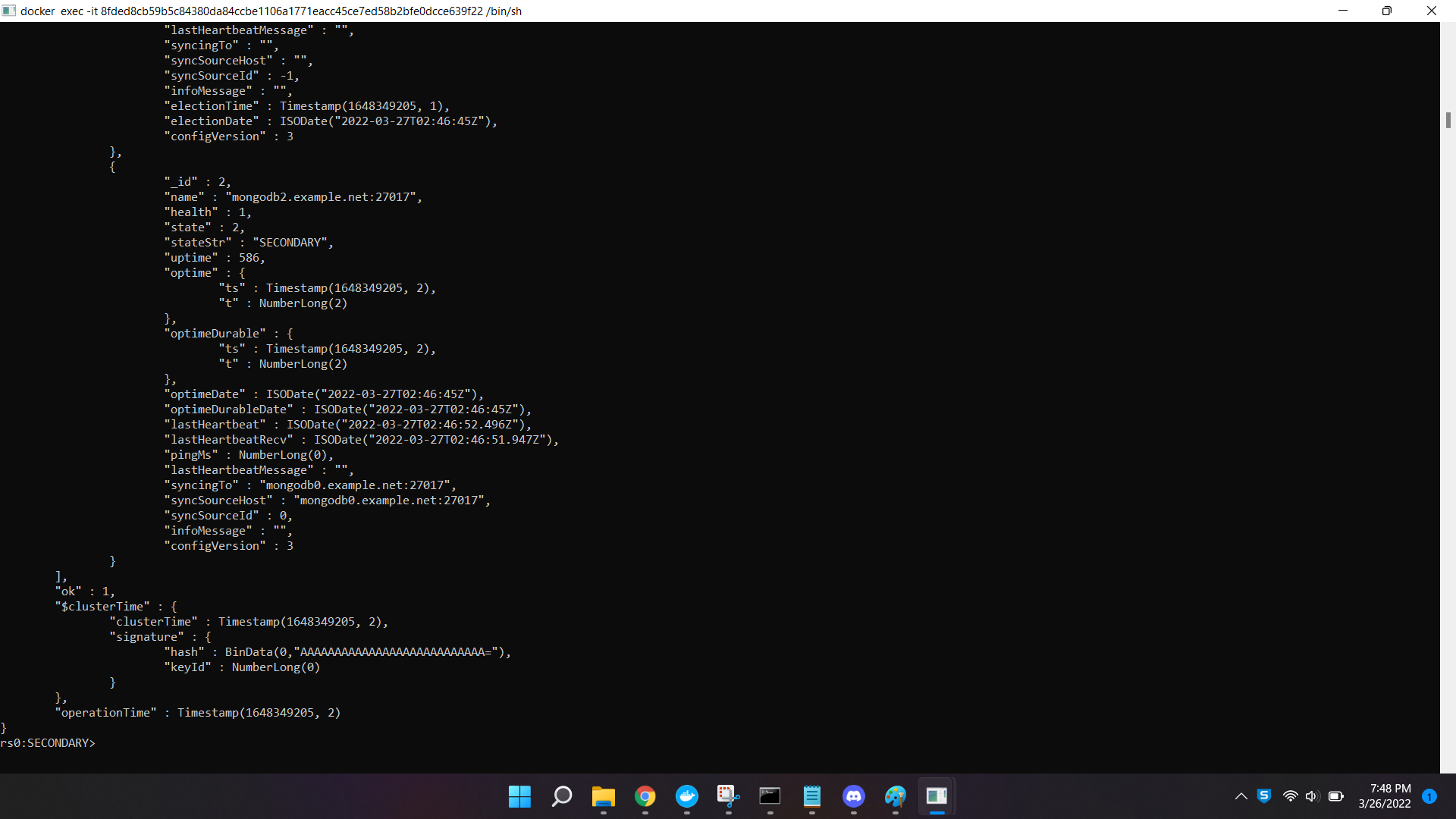
rs.status() before stepping down the primary

The hostname of the primary: **"mongodb1.example.net:27017”**

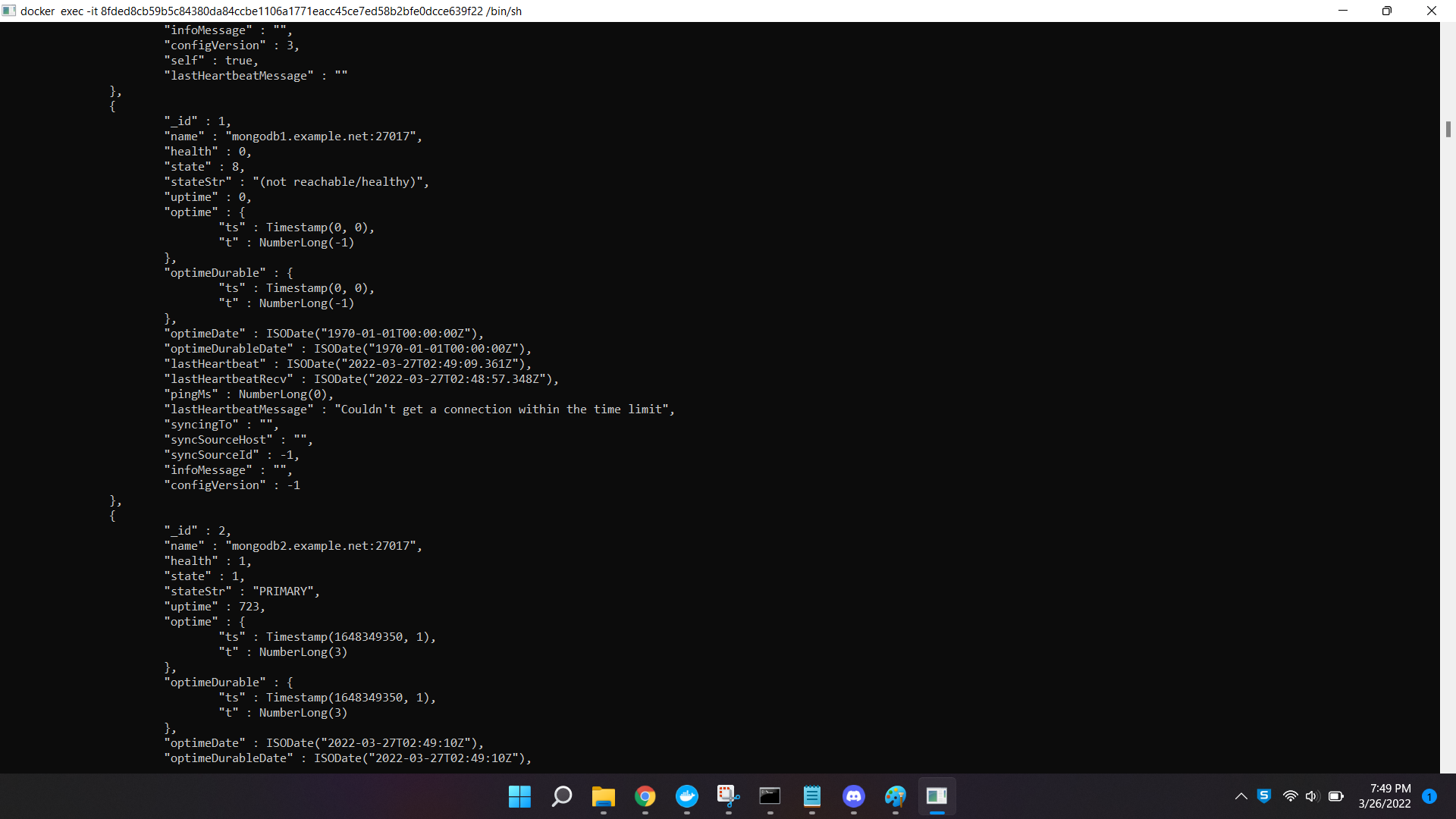




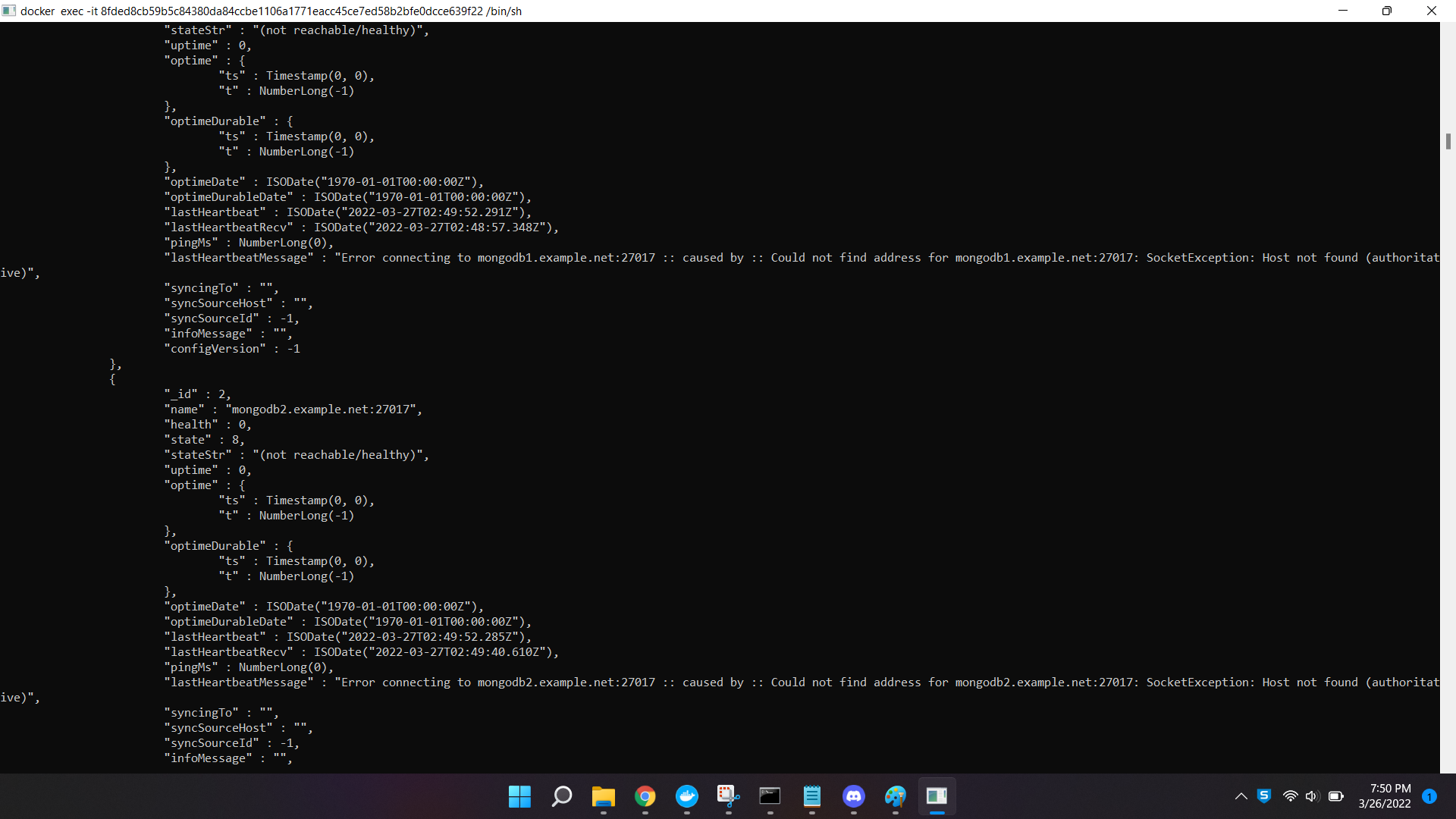




* When the primary container which is <https://mongodb1.example.net/> is stopped it can be observed that the stateStr of the respective host is changed to “**not reachable/healthy**”. This is because the primary has been shut down completely and we can also observe that the lastHeartbeatMessage is "**Couldn't get a connection within the time limit**”. Another replica is automatically elected as the new primary which in this case is “**mongodb2.example.net:27017**”. Screenshot attached below for reference.



* Now we stop the new primary container <https://mongodb2.example.net/> and we observe that stateStr of both **mongodb1.example.net** and **mongodb2.example.net** is changed to “**not reachable/healthy”**. lastHeartbeatMessage for both of them is "**Error connecting::Caused by::Could not find address for mongodb1.example.net:27017: SocketException: Host not found (authoritative)**". This is probably because both the primary hosts are down now and there is no primary host. The only other replica which is still existing i.e, **mongodb0.example.net:27017** is still in the SECONDARY state. Screenshot attached below for reference.



Task 3:

**VI-3:**

Documents are distributed in each shard as below:

**Shard a** contains 57.14% of data, 57.14% docs in cluster and avg obj size on shard: 43B

**Shard b** contains 23.33% of data, 23.33% docs in cluster and avg obj size on shard: 43B

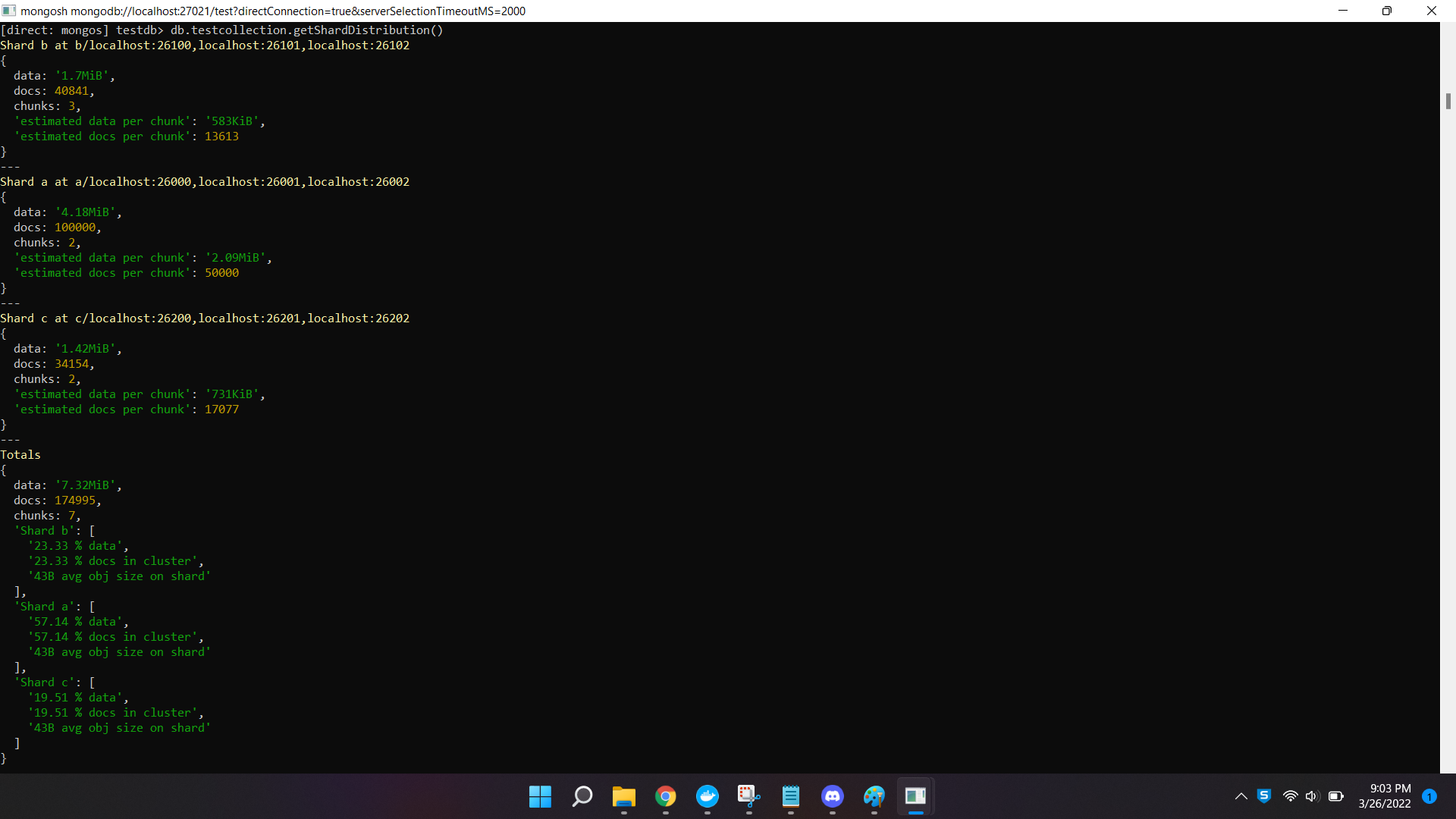
**Shard a** contains 19.51% of data, 19.51% docs in cluster and avg obj size on shard: 43B

**Shard a** contains **highest** number of documents i.e, 100000

**Shard b** contains 2nd highest number of documents i.e, 40841

**Shard c** contains **least** number of documents i.e, 34154

Screenshot of the same attached below for reference



**VI-4:**

After inserting additional documents

Total number of docs: 252456

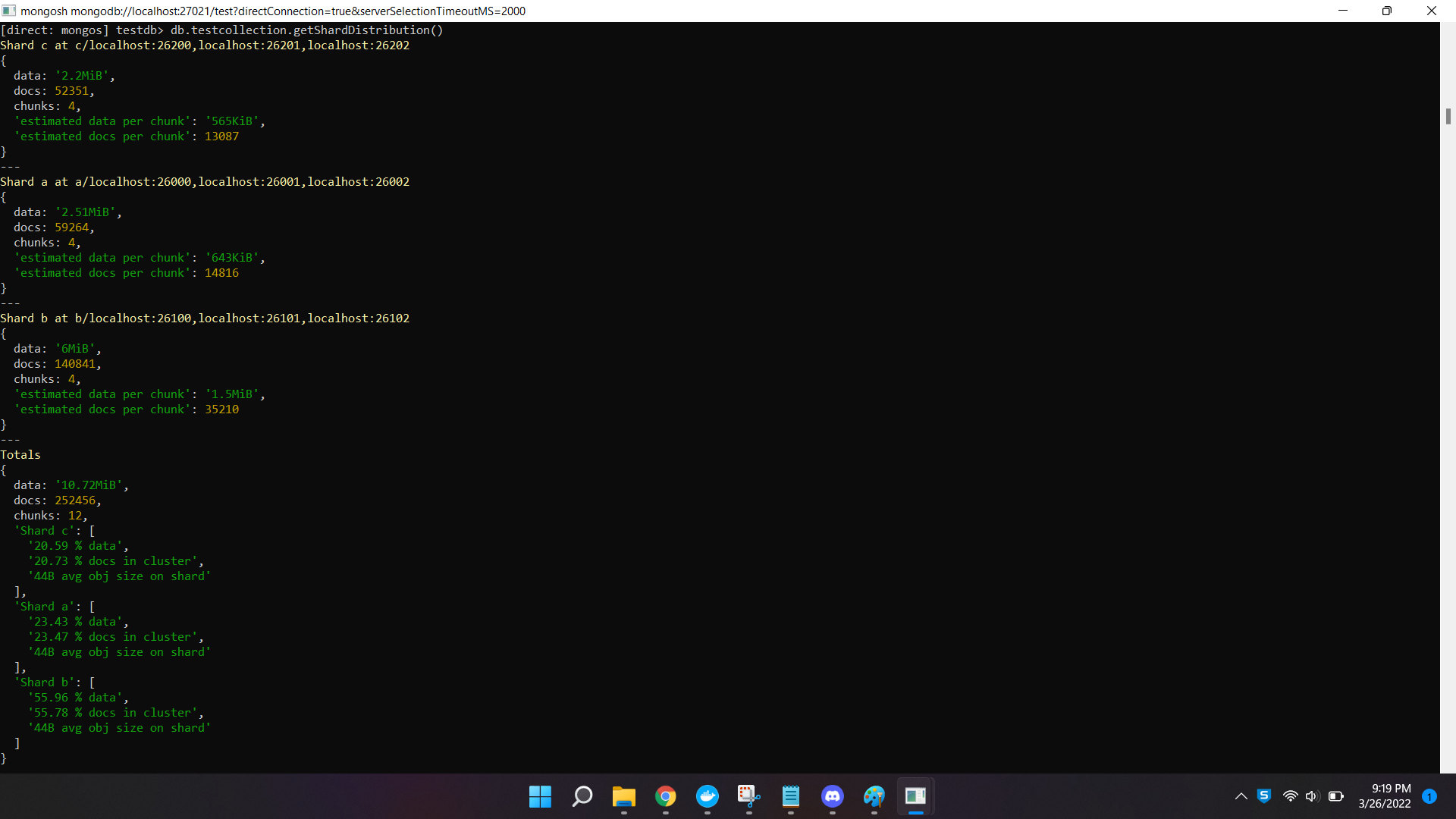
Now, it can be observed that the shard b has the highest number of documents. Below is the data distribution across the shards

**Shard b** contains **highest** number of documents i.e, 140841, '55.96 % data', '55.78 % docs in cluster'

**Shard a** contains 2nd highest number of documents i.e, 59264, '23.43 % data', '23.47 % docs in cluster'

**Shard c** contains **least** number of documents i.e, 52351, '20.59 % data', ‘20.73 % docs in cluster'

Screenshot of the same attached below for reference



VI-5:

sh.status()

