**Advanced Databases**

**MongoDB CA**

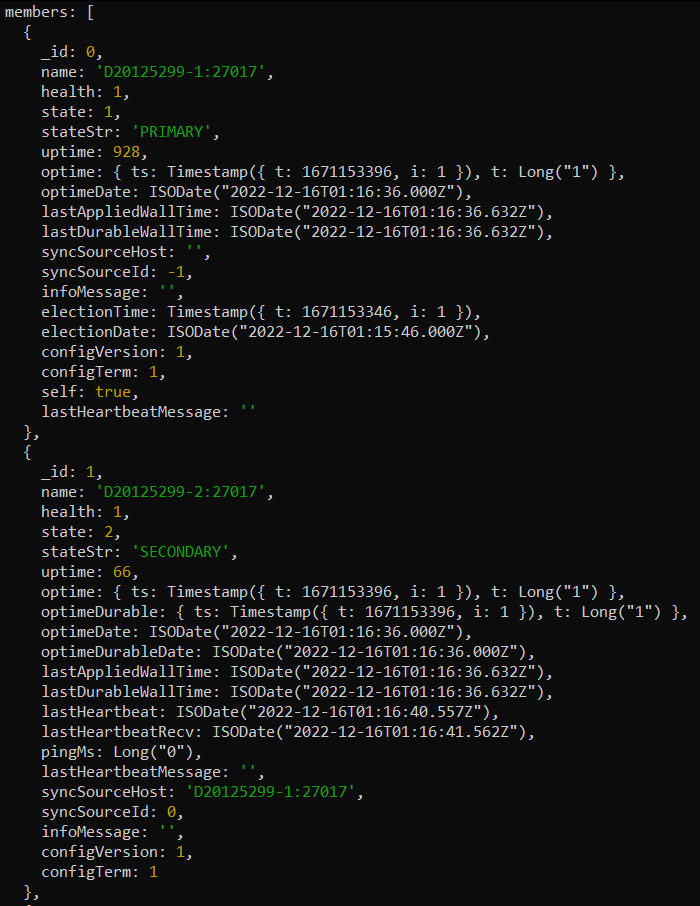
**Student Number:** d20125299

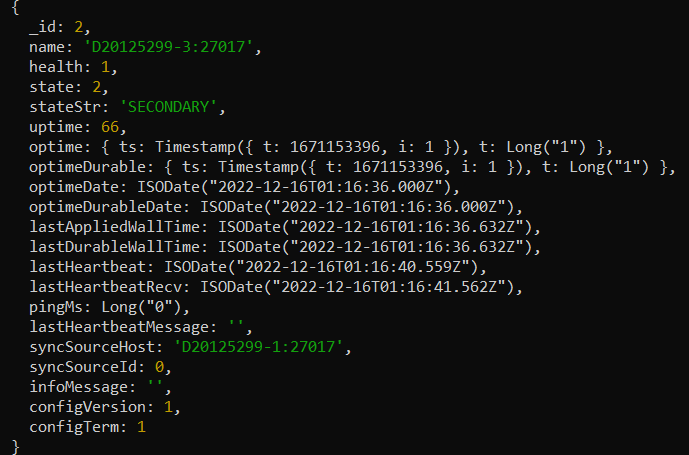
**Student Name:** Luke Hallinan

**Programme Code:** TU856

1. Setting up the cluster and replication

A cluster of 3 nodes was used. These are shown below.

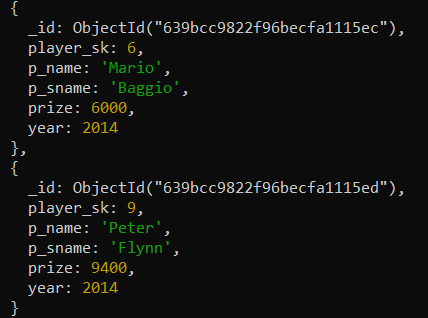


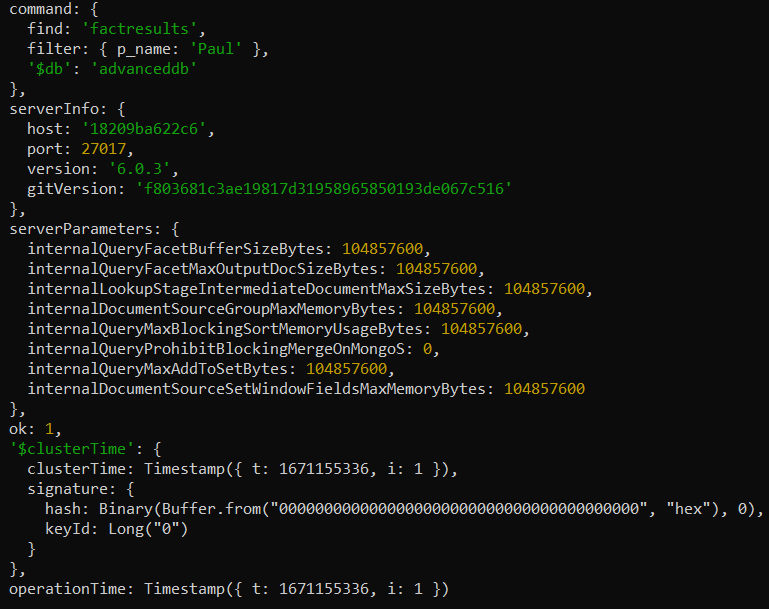
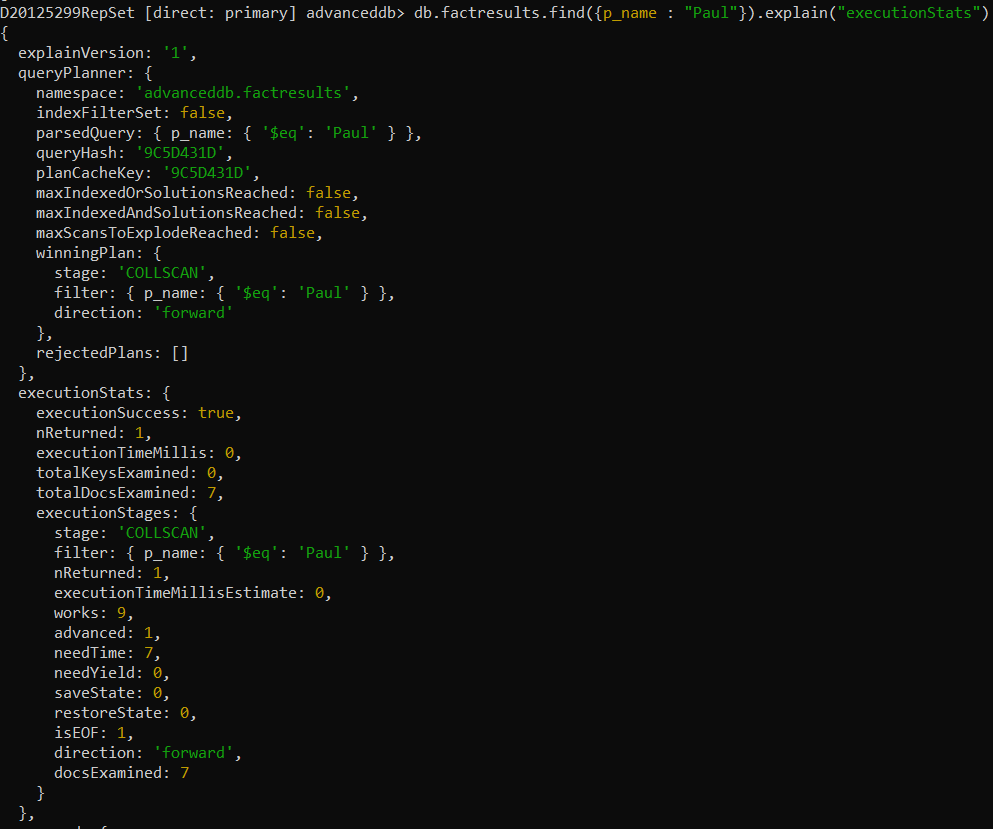


The first is set as the primary with the other two being secondary. All are in a single cluster. Both shild nodes are linked to the first on syncSourceHost: 'D20125299-1:27017'.

1. Porting the data to Mongo

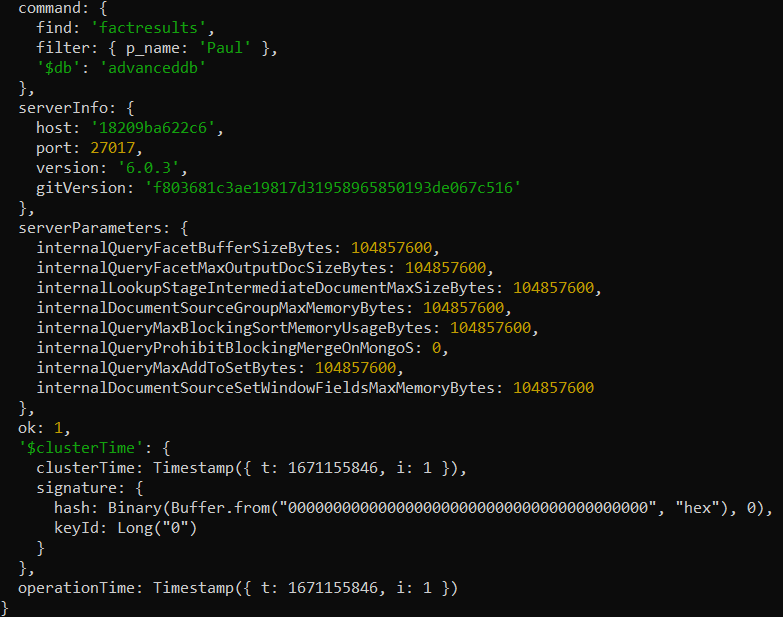




Working with the Golf collection in MongoDB:

This query will find anyone with the name paul in the data. it gives the number returned being 1, the total documents examines which is 7 and the timestamp of the operation. It does a collection scan with a filter on the whole document.

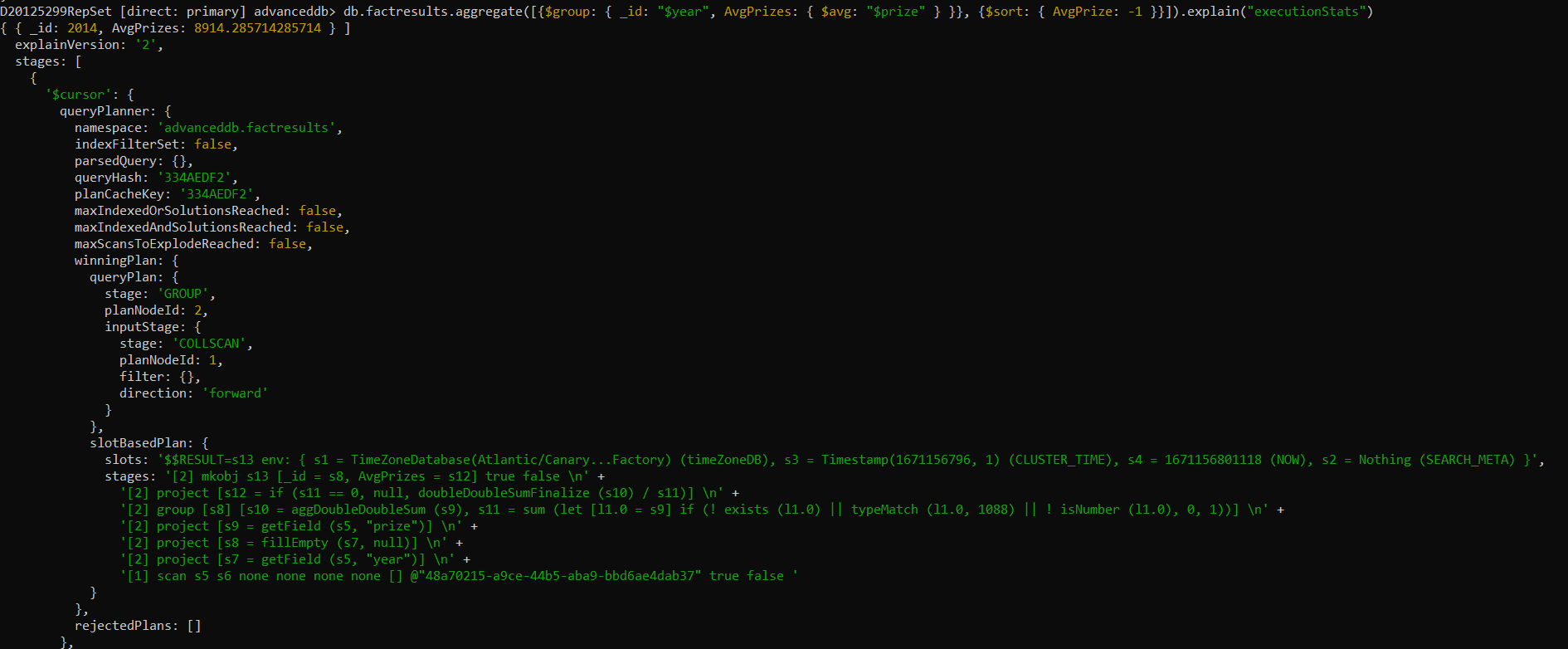
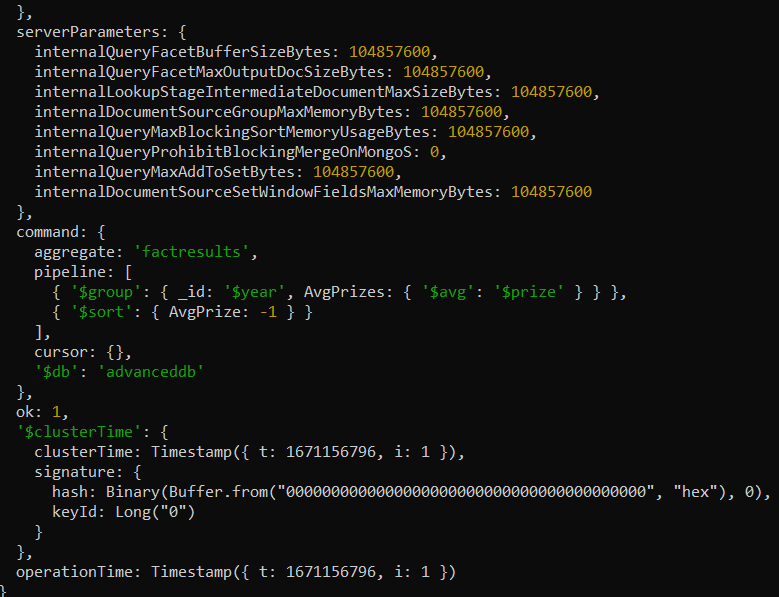
* 1. Adding a secondary index to golf data on a text field.

Added index to p\_name field. This shows that an index was used by IXSCAN as well as the index called p\_name\_1. Also returned is the information from the non indexed such as number of values found and the time stamp. It has total keys examined which is 1.

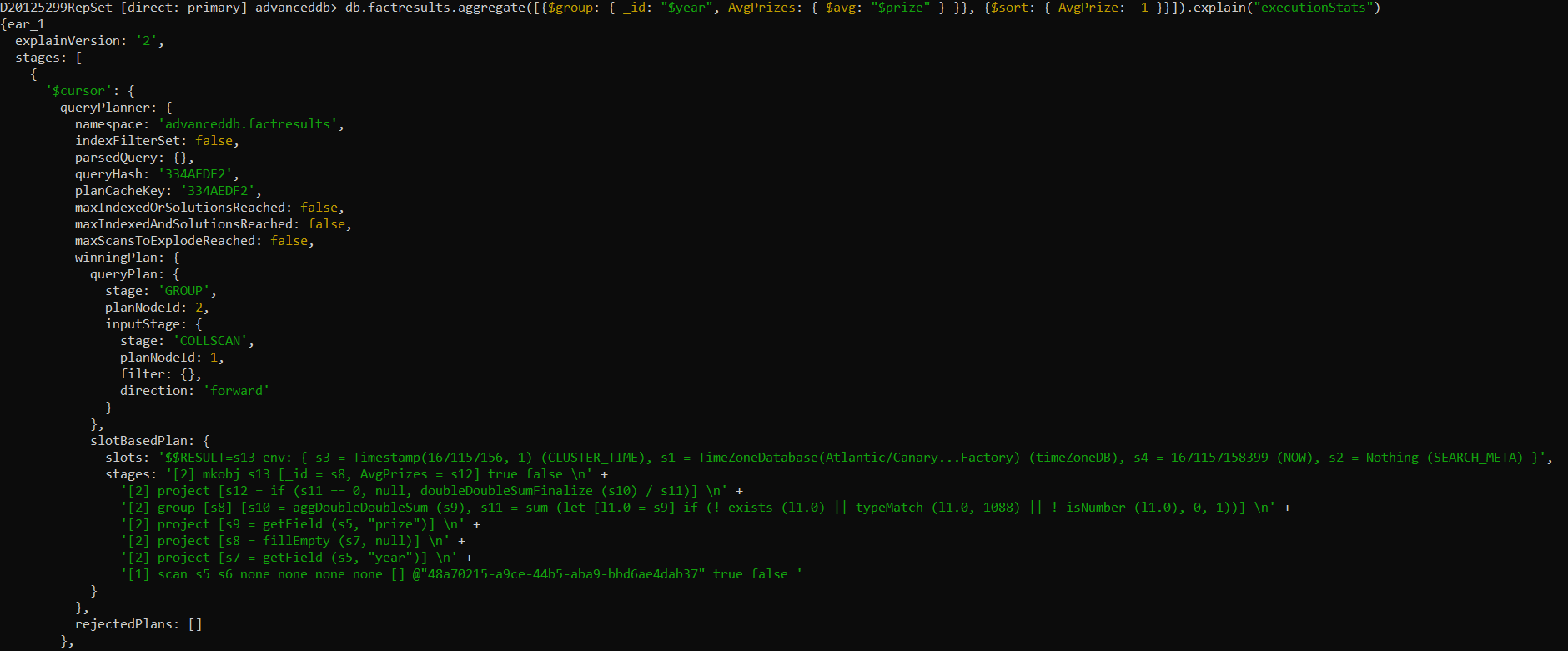
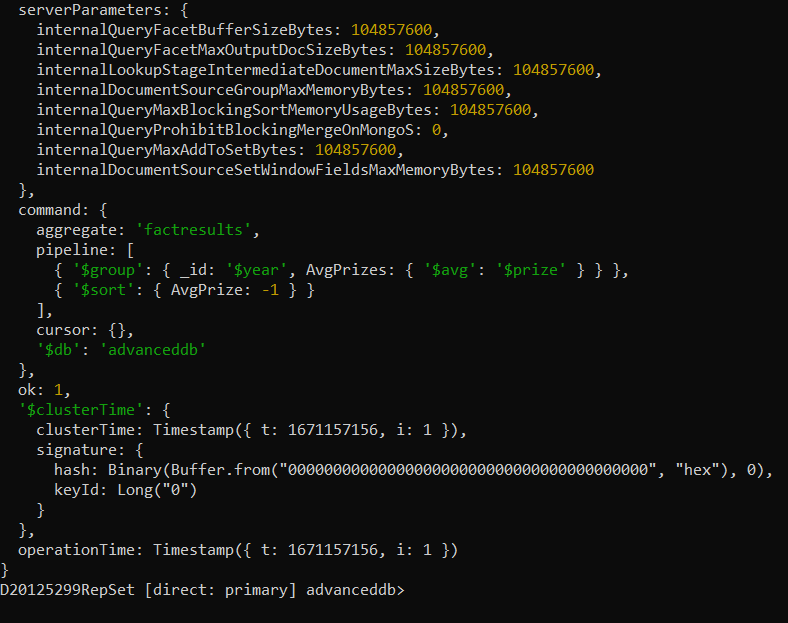
1. Working with aggregation in MongoDB:
   1. Create an aggregation pipeline



This shows the method used for the aggregation pipeline. It uses a collection scan and a group. The slots based plan allows the aggregation to work using valuesrather than documents.

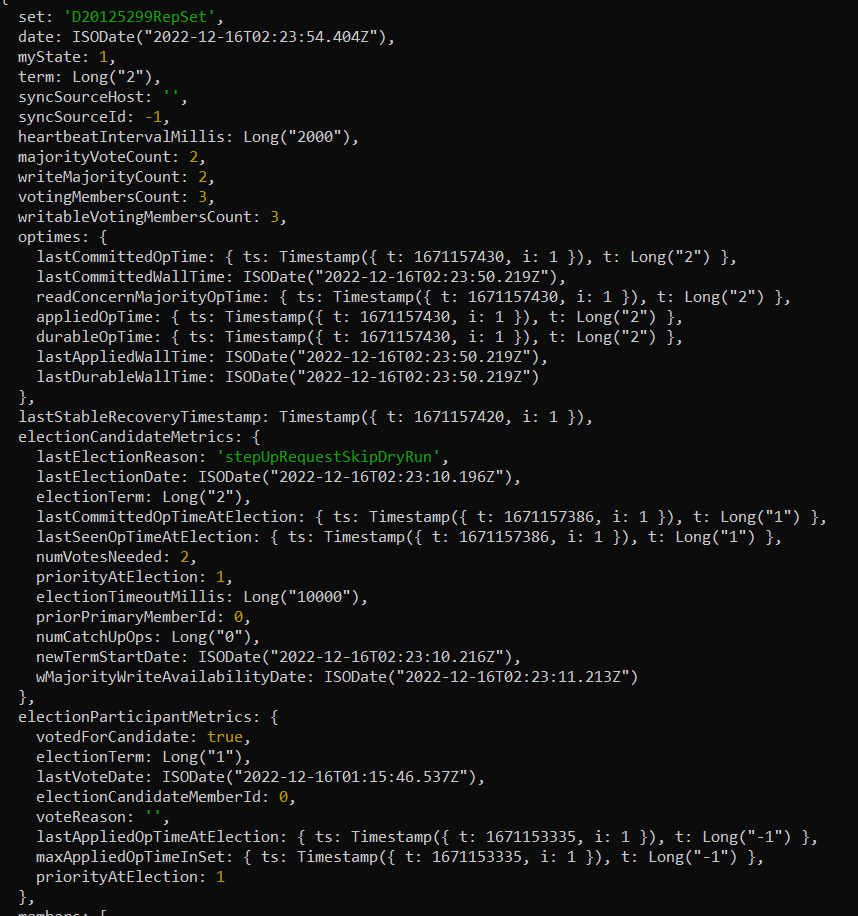
* 1. Add relevant indexes and reorder your stages.

an index was added to the year column. In the case of this dataset there was not a lot of change but with larger more varied ones this can change the method and speed greatly.

<<use explain, capture the output and include it here, comment on what is happening>>

1. Replication working

As can be seen above the first instance was stopped. When this happen the second instance was autimatically reassigned to PRIMARY while the first instance is now (not reachable/healthy). This is a great way to protect agains data loss and interuption by have an automatic switch if one fails.