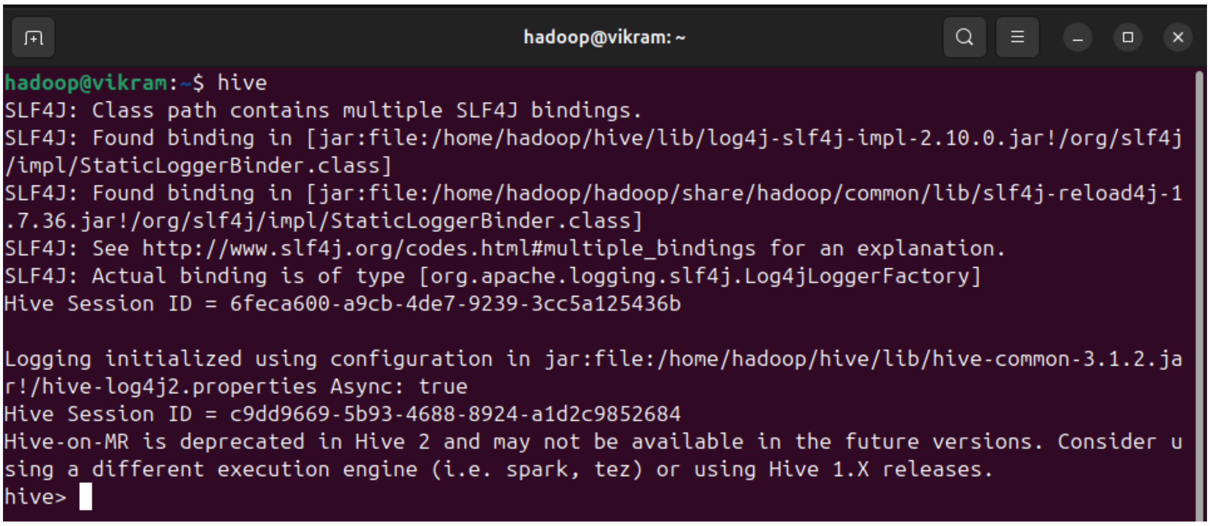
Exp. No : 5a

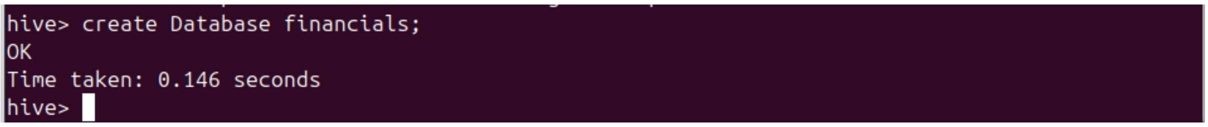
Designing and testing various schema models to optimize data storage and retrieval using Hive.

1. Start hive



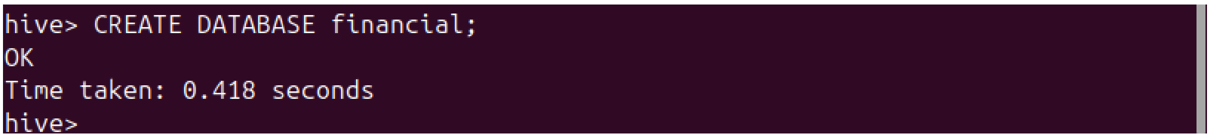
```
hadoop@vikram:~$ hive
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/hadoop/hive/lib/log4j-slf4j-impl-2.10.0.jar!/org/slf4j
/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/hadoop/hadoop/share/hadoop/common/lib/slf4j-reload4j-1
.7.36.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
Hive Session ID = 6feca600-a9cb-4de7-9239-3cc5a125436b

Logging initialized using configuration in jar:file:/home/hadoop/hive/lib/hive-common-3.1.2.ja
r!/hive-log4j2.properties Async: true
Hive Session ID = c9dd9669-5b93-4688-8924-a1d2c9852684
Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider u
sing a different execution engine (i.e. spark, tez) or using Hive 1.X releases.
hive>
```

2. CREATE DATABASE financials in hive



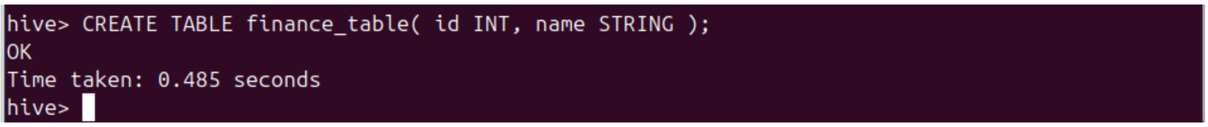
```
hive> create Database financials;
OK
Time taken: 0.146 seconds
hive>
```

3. Use financials database in hive



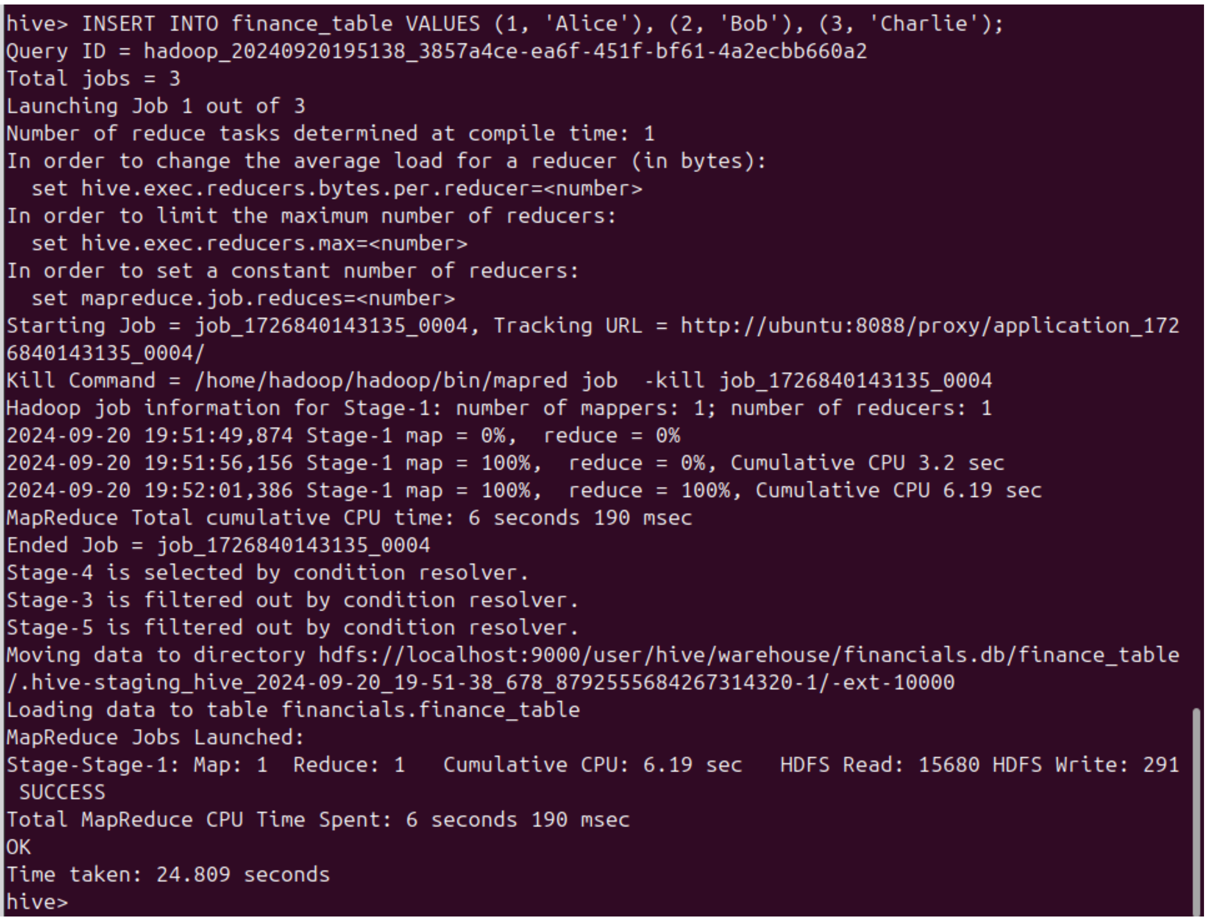
```
hive> CREATE DATABASE financial;
OK
Time taken: 0.418 seconds
hive>
```

4. Create Finance_table table in hive



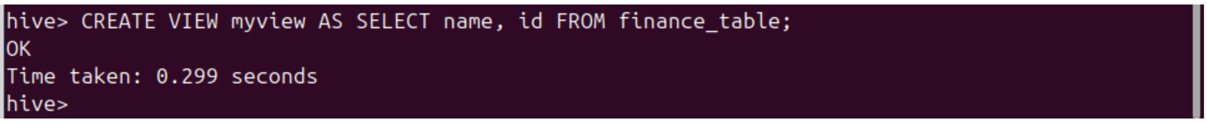
```
hive> CREATE TABLE finance_table( id INT, name STRING );
OK
Time taken: 0.485 seconds
hive>
```

5. Insert records in finance_table table



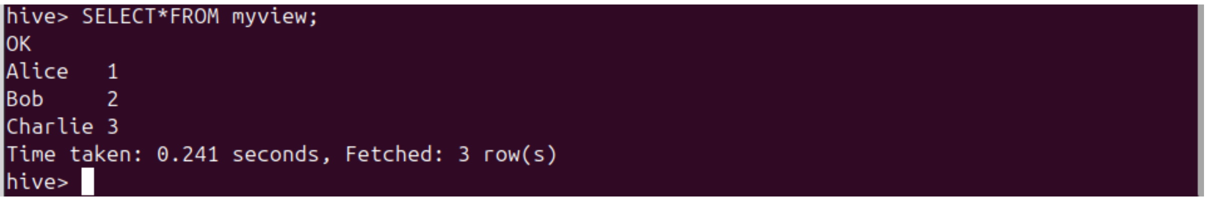
```
hive> INSERT INTO finance_table VALUES (1, 'Alice'), (2, 'Bob'), (3, 'Charlie');
Query ID = hadoop_20240920195138_3857a4ce-ea6f-451f-bf61-4a2ecbb660a2
Total jobs = 3
Launching Job 1 out of 3
Number of reduce tasks determined at compile time: 1
In order to change the average load for a reducer (in bytes):
  set hive.exec.reducers.bytes.per.reducer=<number>
In order to limit the maximum number of reducers:
  set hive.exec.reducers.max=<number>
In order to set a constant number of reducers:
  set mapreduce.job.reduces=<number>
Starting Job = job_1726840143135_0004, Tracking URL = http://ubuntu:8088/proxy/application_172
6840143135_0004/
Kill Command = /home/hadoop/hadoop/bin/mapred job  -kill job_1726840143135_0004
Hadoop job information for Stage-1: number of mappers: 1; number of reducers: 1
2024-09-20 19:51:49,874 Stage-1 map = 0%,  reduce = 0%
2024-09-20 19:51:56,156 Stage-1 map = 100%,  reduce = 0%, Cumulative CPU 3.2 sec
2024-09-20 19:52:01,386 Stage-1 map = 100%,  reduce = 100%, Cumulative CPU 6.19 sec
MapReduce Total cumulative CPU time: 6 seconds 190 msec
Ended Job = job_1726840143135_0004
Stage-4 is selected by condition resolver.
Stage-3 is filtered out by condition resolver.
Stage-5 is filtered out by condition resolver.
Moving data to directory hdfs://localhost:9000/user/hive/warehouse/financials.db/finance_table
/.hive-staging_hive_2024-09-20_19-51-38_678_8792555684267314320-1/-ext-10000
Loading data to table financials.finance_table
MapReduce Jobs Launched:
Stage-Stage-1: Map: 1  Reduce: 1   Cumulative CPU: 6.19 sec   HDFS Read: 15680 HDFS Write: 291
 SUCCESS
Total MapReduce CPU Time Spent: 6 seconds 190 msec
OK
Time taken: 24.809 seconds
hive>
```

6. Creating new VIEW named myview for finance_table



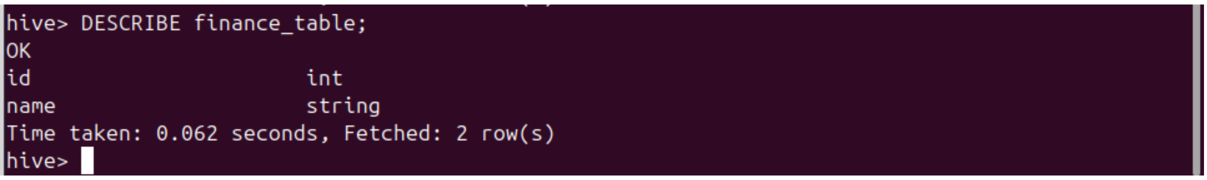
```
hive> CREATE VIEW myview AS SELECT name, id FROM finance_table;
OK
Time taken: 0.299 seconds
hive>
```

7. Display myview.



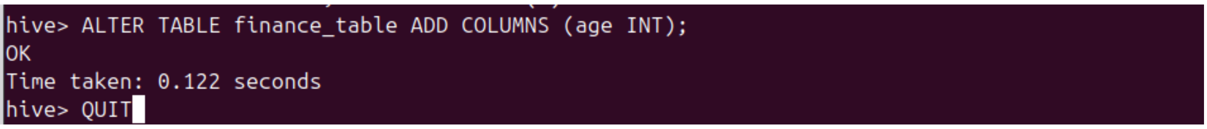
```
hive> SELECT*FROM myview;
OK
Alice   1
Bob     2
Charlie 3
Time taken: 0.241 seconds, Fetched: 3 row(s)
hive>
```

8. Describing finance_table structure.



```
hive> DESCRIBE finance_table;
OK
id                      int
name                    string
Time taken: 0.062 seconds, Fetched: 2 row(s)
hive>
```

9. Add new age column to Finance_table



```
hive> ALTER TABLE finance_table ADD COLUMNS (age INT);
OK
Time taken: 0.122 seconds
hive> QUIT
```