High Distinction Level Project

Class Enrolment Data Storage in MySQL vs. MongoDB

COS20015 – Fundamentals of Data Management

Student Name: Cong Duc Danh (Dan) Ngo

Student ID: 104186810

Lab session: Wednesday 10:30 AM BA407

I.   Abstract

Databases are an essential part of the current era of data-driven decision making, hence there are many options for users, organisations, etc. to select for storing and accessing data. This report provides a comparison between MySQL and MongoDB, 2 popular Database Management Systems, based on a relational database that have been previously run on a MySQL data server.

II.   Introduction

MySQL is an RDBMS (Relational Database Management System) which organises data into structured, tabular formats. On the contrary, MongoDB, a NoSQL DBMS, employs a more adaptable approach by storing data as JSON documents. While both DBMSs excel in terms of performance and scalability, their strengths are particularly suited to different types of applications (AWS 2023). In this report, I will use the custom database called "Classes" that I have implemented in my Distinction Level Project (Ngo 2023), migrate it to a Mongo database, and compare the differences between the data storage in 2 DBMSs.

III.   Method & Results

In this section, I have conducted several methods aiming at comparing MySQL and MongoDB in terms of the following aspects:

1. Data Migration
- In order to insert data to a MySQL database, you will need to create a table with attributes, primary and foreign (optional) keys. The following figure shows the creation and data insertion for the "Unit" table:

```
15 ●    CREATE TABLE IF NOT EXISTS Unit
16   ⊖ (
17      Unit_ID CHAR(8) NOT NULL,
18      Convener_ID CHAR(7) NOT NULL,
19      Unit_name VARCHAR(50) NOT NULL,
20      PRIMARY KEY(Unit_ID),
21      FOREIGN KEY(Convener_ID) REFERENCES Convener(Convener_ID)
22    );
23
24 ●    INSERT INTO Unit(Unit_ID, Convener_ID, Unit_name)
25      VALUES
26                  ("COS20007", '1876543', "Object Oriented Programming"),
27                  ("SWE30009", '1876543', "Software Testing and Reliability"),
28                  ("TNE30023", '1245678', "Advanced Switching"),
29                  ("COS20015", '1245678', "Fundamentals of Data Management")
30      ;
```

Figure 1: Migrating Data in MySQL.

- In MongoDB, primary keys and foreign keys are not required, and each data record in an entity is created by a JSON object, and then import the objects to a MongoDB collection. The following figure shows the JSON file for the same Unit entity:

Figure 2: Migrating Data in MongoDB using JSON objects.

2. Data Storage
- In MySQL, data is stored in tabular format with strict rules based on the primary key, foreign key constraints and 1 row can only store 1 or 0 record for 1 attribute.

Figure 3: "Tutor" entity in MySQL

- In MongoDB, data is stored as a binary JSON, or BSON (MongoDB 2023) by inserting methods or importing JSON/CSV files. Data representation is MongoDB is not restricted to keys, attributes constraints or number of records stored in an attribute.
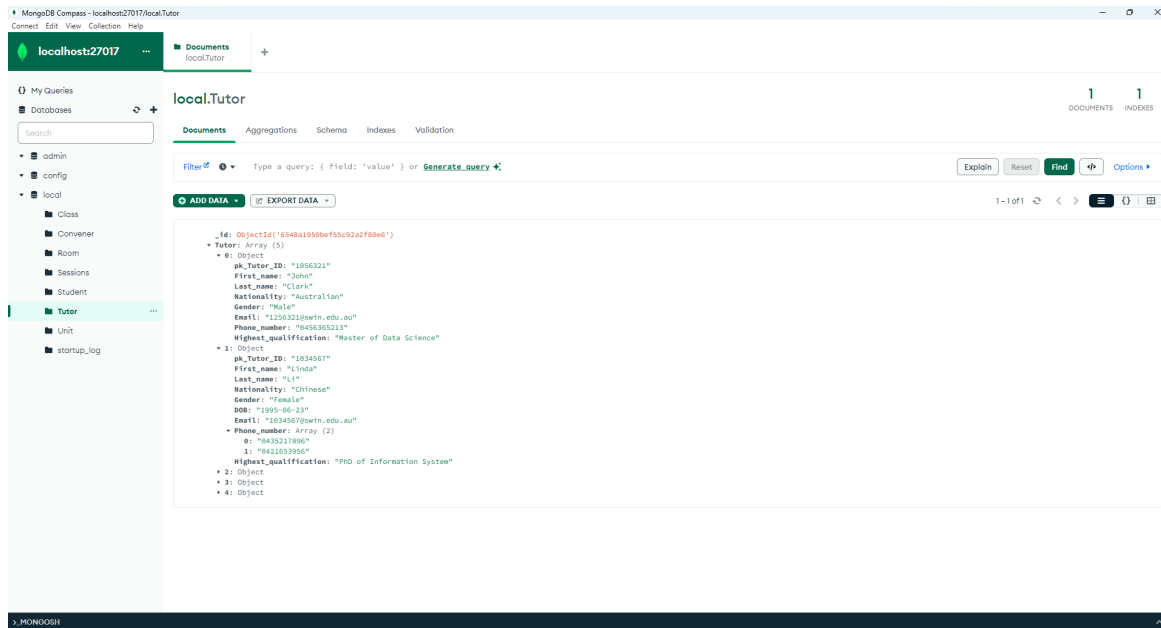

Figure 4: "Tutor" entity in MongoDB

- As we can see from Figures 3 and 4, while every row in the Tutor table must have 1 record for NOT NULL attributes and at most 1 record for NULLABLE attributes, objects in MongoDB are not restricted to those limitations. Tutor with ID 1056321 has no DOB attribute and Tutor 1034567 has 2 phone numbers.

3. Data Query Performance
- In a relational database like MySQL, queries are executed by SQL, which highly optimises the performance of structured, tabular data. MySQL employs techniques such as indexing or caching to enhance query performance.
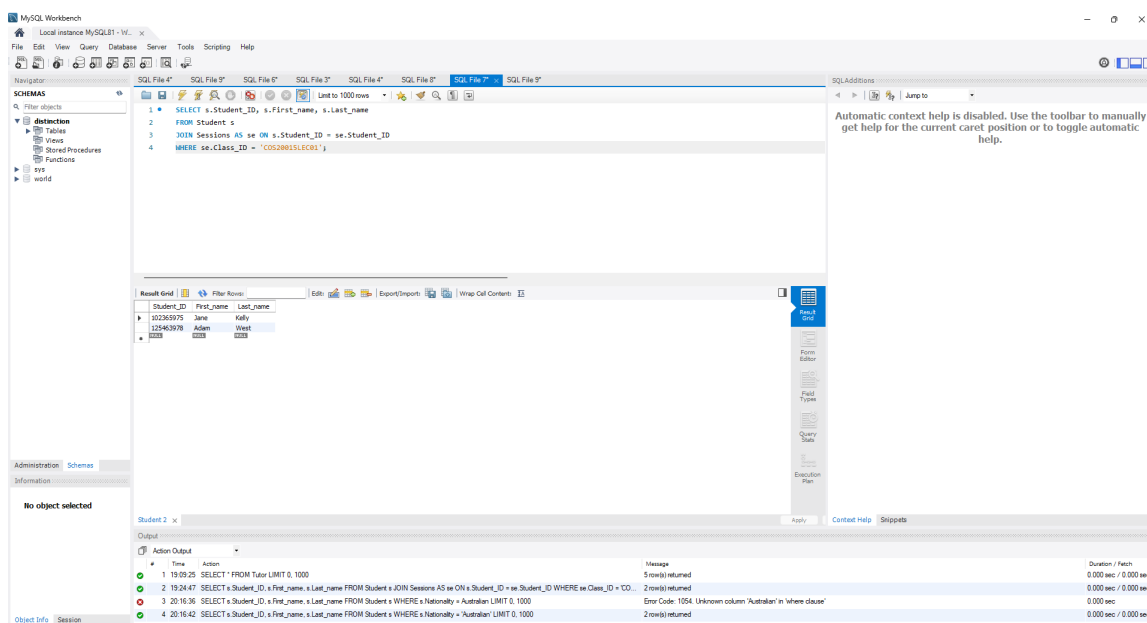

Figure 5: A SQL query example in MySQL database

- In MongoDB, search queries are performed by a query language that is designed to work with its flexible BSON documents. It is able to efficiently handle queries on unstructured or semi-structured data.
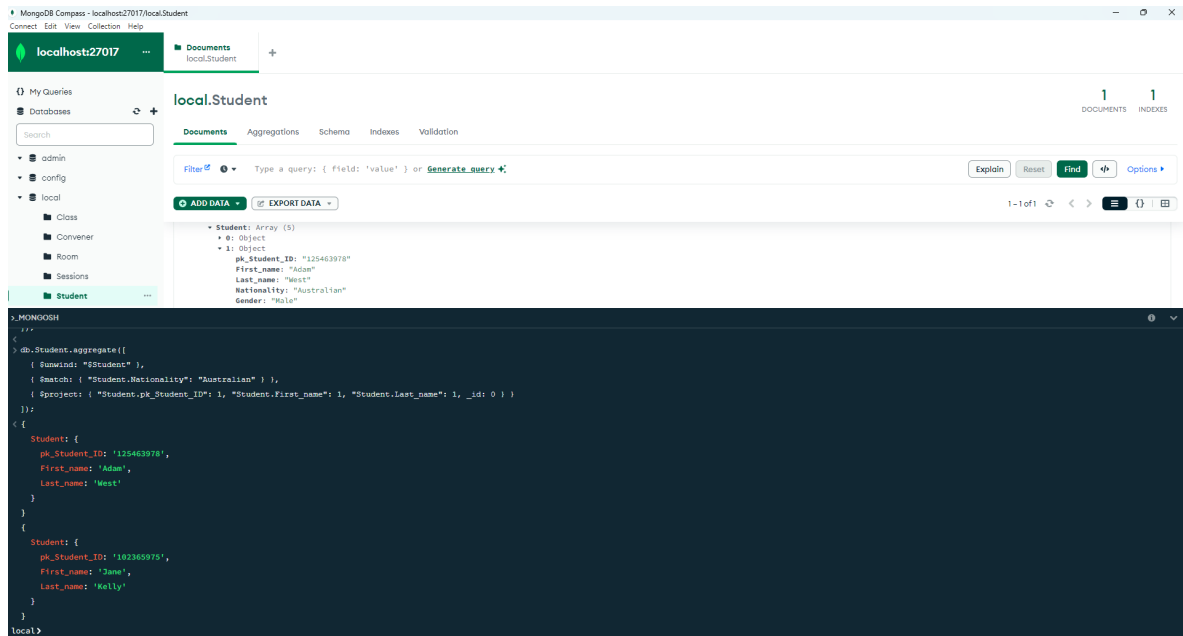

Figure 6: The same query in Figure 5 executed in MongoDB.

IV.   Discussion

In this section, I will discuss the differences between MySQL and MongoDB databases based on the research carried out in section III.

1. Data Migration
    - MySQL is a table-based system, the tables are designed for data query searching and is considered an SQL database. Therefore, data is searchable and accessible relative to another datapoint on set (IBM 2021). Creating entities and inserting data requires strict orders in several aspects such as: father-child table, primary key, foreign key, data integrity, and so on.
    - In contrast, MongoDB is a document-bases non-relational DBMS, also called as object-based system unlike MySQL. It is designed to go beyond the restrictions of the relational structure as an easier way to work with different types of data such as unstructured or semi-structured (IBM 2021). Creating database in MongoDB is more flexible than MySQL since you can insert tables and data in any order, and objects are not restricted to a certain of number of records.
2. Data storage
    - In MongoDB, each individual records are stored as documents, while in MySQL, each record is stored as rows or tuples in a structured table. However, the hierarchical structure of a record of MongoDB is more complex than MySQL. Documents with embedded objects in MongoDB belong to a particular collection, whereas a single table is used to store all records in MySQL (GeeksforGeeks 2018). Therefore, search queries in MongoDB are more difficult to conduct than MySQL in my personal opinion.

3. Data Query Performance
   - It is extremely difficult to compare the performance of 2 completely different DBMSs, since they have different approaches to search, sort and present data (MongoDB 2023).
   - For instance, MySQL is considered for executing high-performance JOIN operation across multiple tables that are well-indexed. On the other side, MongoDB offers support for a similar functionality through the $lookup operation, but it is generally less essential. This is due to the typical way in which MongoDB organises its documents following a hierarchical data model and concentrating most data within a single document. Consequently, this design minimises the necessity for joins across multiple documents (MongoDB 2023).

V.   Conclusion

In this report, I have briefly compared MySQL, a relational database, with MongoDB, a non-relational database, based on a database that has been designed in MySQL. To conclude, MongoDB is better for data storage in terms of flexibility, hence its security and data integrity aspects are might not be as good as MySQL, a structured SQL database. The preference between these 2 great DBMSs is different depending on the users, organisations in particular scenarios, the existence of them bring us unlimited potential for storing, sorting and presenting data for Software Development.

VI.   References

1. AWS 2023, What's the Difference Between MongoDB and MySQL?, Amazon Web Services, Inc. or its affiliates, viewed 23 February 2023, <https://aws.amazon.com/compare/the-difference-between-mongodb-vs-mysql>.
2. Ngo, D. 2023, Distinction Project Report. COS20015: Fundamentals of Data Management. Swinburne University of Technology. Unpublished assignment.
3. MongoDB 2023, MongoDB vs. MySQL Differences, MongoDB, Inc., viewed 5 June 2021, < https://www.mongodb.com/compare/mongodb-mysql>.
4. IBM 2021, MySQL vs. MongoDB: What's the Difference, IBM, viewed 3 November 2021, <https://www.ibm.com/blog/mysql-vs-mongodb/>.
5. GeeksforGeeks 2023, MongoDB vs MySQL, GeeksforGeeks, viewed 17 April 2018, <https://www.geeksforgeeks.org/mongodb-vs-mysql/>.