

**Task 1, Part 2: (15 marks)**

You are asked to design a Database solution for the data above. What would you use? How? Why did you make that choice? What are the advantages and disadvantages?

MongoDB is used when there is a big data to be stored, and the data above has huge records of 3431. In MongoDB records or data is defined as documents and each document has its own ID unlike a relational database which only allows one unique id per table. This makes MongoDB very attractive towards storing big data like above which has more than 3000 documents because it allows will make the database more structured and organised, this will save the user more time as it leaves the user low chance of getting confused when big data are being stored into the database. Furthermore, in the same collection of documents, the way the a document is stored can be different from a document in the same collection, this makes the MongoDB flexible than other databases like SQL which are restricted to the same schema and when it is a big data like above, the user will want the database to be flexible which in turn saves the user time.

In MongoDB, data can be stored in JSON format, which means that it will mean smaller size, more structural documents because different data types can be easily distinguished such number 11 and host name 11 Com string on the data above, this means in MongoDB data like number and number strings on the data above can be easily distinguished. This makes MongoDB stand out among other databases because the user does not have to explicitly define data type of every single document making the data input more faster and without a duplication as the MongoDB provides a unique ID for every document.

As for the scalability issue, MongoDB is perfect for the user as it has a process called sharding. Sharding is the process of storing documents across multiple machines and it MongoDB's approach to meeting the demands of data growth and this defines their quality of service. For the data above, sharding will work by horizontally portioning the database. This means that the rows of a document will be held somewhere else rather than splitting the column, each time there is a partition there will be a new shard(rows) which will be stored in a separate database or separation physical location. This will reduce the index size of the database of the big data above and this means there will be faster search which in turn improves the performance of the database. This also means that the data is not lost in-case there is a failure as the document is stored into separate databases or location and therefore will contain copies of data and this offers the user great quality of service. There are the reasons why MongoDB is the database solution of my choice.