

Comparative Case Study: An Evaluation of Performance Computation Between SQL And NoSQL Database

Muhammad Zohaib Khan

Software and Data Engineer at SMBBIT Karachi, Pakistan
Email: zohaib_khan2017@yahoo.com

Fahim Uz Zaman

Software Engineer at Premier Systems Pvt. Ltd./Audi, Karachi, Pakistan
Email: faheemuzzaman@hotmail.com

Muhammad Adnan

Engineer at KPT Karachi, Pakistan
Email: m.adnan1225@gmail.com

Aisha Imroz

Software Engineer at Avanza Pvt. Ltd, Karachi, Pakistan
Email: aishaimroz4@gmail.com

Mahira Abdul Rauf

SAP Basis Engineer at TMC Pvt. Ltd, Karachi, Pakistan
Email: mahirab30@gmail.com

6th Zuhaib Phul

Web Engineer at Karachi Shipyard & Engineering Works Ltd, Karachi, Pakistan
Email: zuhaibphul53@gmail.com

Received: 15th October, 2022; Accepted: 13th February, 2023; Published: 28th February, 2023

Abstract: Data has always been the company's most valuable resource because it can be used for analysis, decision-making, and judgement. Hard data handling necessitates the use of complicated cache and accessibility concepts. The effectiveness of SQL and NoSQL database systems for producing scientific data is examined in this study. SQL databases and NoSQL databases are the most popular and structured types of database solutions. Another name for the SQL database is RDBMS (Relational Database Management System). Associations or tables are used to organize the data. A NoSQL database is a non-relational database management system. NoSQL databases, a new type of database system, were created to address this issue by providing an unstructured platform and scalability for large data applications. The term "NoSQL" refers to more than just SQL. Wide column stores, documents, graph databases, and key-value pairs are a few NoSQL database types that do not have the necessary standard structure. Additionally, in RDBMS, it might scale horizontally rather than vertically. To compare SQL and NoSQL databases, the data is organized in unstructured tables or relationships. Both of them are open source. The experiment assessed and supported database loading, response, and retrieval times for both SQL and NoSQL databases to discover if a database is smoother, more efficient, and performant.

Index Terms: SQL, NoSQL Database, MySQL, MongoDB.

1. INTRODUCTION

Currently, various companies generate data through a variety of apps. These apps also generate a large amount of disorganised data, making data management difficult. Companies must analyse and store data in specific structures since it is crucial to their operations. Relational databases can only handle structured data; hence, a NoSQL database management system is required because it can handle unstructured data. Slovakian railways contain a sizable collection of values that show the connections between cities and villages around the country. When this database was established,

relational databases existed only. A database is a repository of organised data or information stored in a computer device so that a user or computer programs may access it using a search language. Decision-making can be aided by the knowledge thus gathered. SRBD is the name of the computer programme used for data administration and querying. Database management systems are classified by computer programmers according to the DBMS they support. Relational databases, that were originally described, ruled technology in the late 1970s. They allow to include both rows and columns in a table collection. Bulk of them also enter and access information using SQL. Due to their unique query language, non-relational databases, such as NoSQL, have gained popularity in the past decade of the 2000s. [1]. Many databases are used today, mostly in the industry. They come in both SQL and NoSQL varieties. NoSQL is a database management system that is not relational, and an RDMS database defines the relationships between tables. In contrast to NoSQL databases, which have an unstructured schema, SQL databases have a predefined, permanent, or static pattern. Declarative programming (DML, DDL, and DCL) can be used effectively to gather declarations, while NoSQL database searches are simple to scale and do not require joining. A stable schema is not necessary with NoSQL. [2]. Many businesses use databases to manage accounts, websites, and a variety of other industries. A set of predefined data objects makes up a relational database. These items are grouped in table-like columns and rows. The columns of the table have headers, while the rows include the actual data that corresponds to the heading. A specific identification referred to as the main key can be used for every row in a table. In-stored procedures, join queries, view techniques, and some other upwardly adaptive techniques are used in relational database systems. Non-relational databases use a strategy designed to meet the essential needs of a provided information. For instance, data can be stored in non-relational databases used in NoSQL databases as records, charts, and simple keys and values sets. [3]. The first figure below illustrates the distinction between SQL and NoSQL databases.

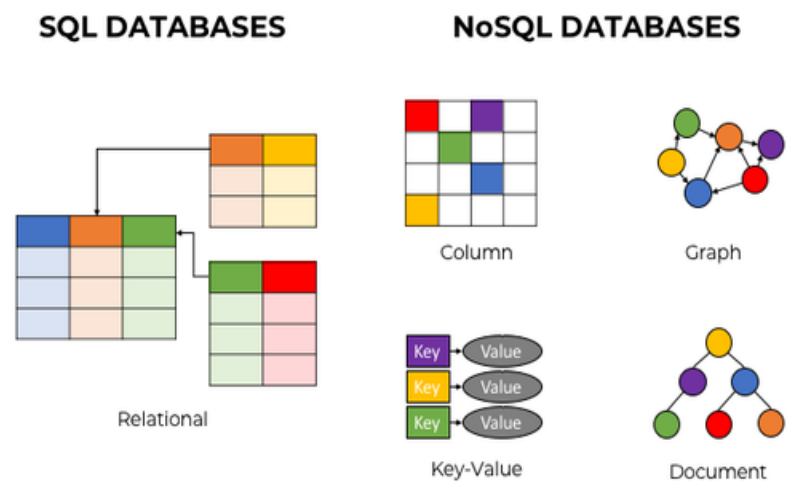


Fig.1. SQL versus NoSQL Databases

In this study, we apply these tools to investigate scientific results from SQL and NoSQL. In order to determine whether the database is speedier, more efficient, and more performant, the experiment evaluated and justified database loading, response, and retrieval times for both SQL and NoSQL databases.

2. BACKGROUND

SQL databases are normalized databases in which the data is divided up into a number of logical tables to prevent information duplication and redundancy. Alternative names for relational databases include SQL databases. In this situation, SQL databases perform joins, queries, updates, and other types of operations better than NoSQL databases.

SQL databases are designed for structured data, whereas Unstructured data can be graph-based, column-based, document-oriented, or any mix of the three. In this instance, the questioned data item is kept in a single location without being partitioned. As a result, importing or exporting a single data item appears to be significantly quicker when comparing NoSQL databases to SQL databases. Find out in this paper which is more dependable in terms of response time. [5,6].

2.1. SQL Database

Relational data is the most common type of data model. It is widely used to collect and extract information all over the world. All the qualities and abilities necessary for efficient data processing are contained in this straightforward paradigm. Figure 2 demonstrates how it saves data as a table. A relational data model uses a standard query language, also referred to as a structured query language. It handles relational databases, makes updates for them, and performs other tasks. It is used by several of the industry's top players, including Oracle, Microsoft, MySQL, and DB2. It will be straightforward to read and comprehend for customers. [7]. Figure 3 shows the many SQL components that are used for different activities.

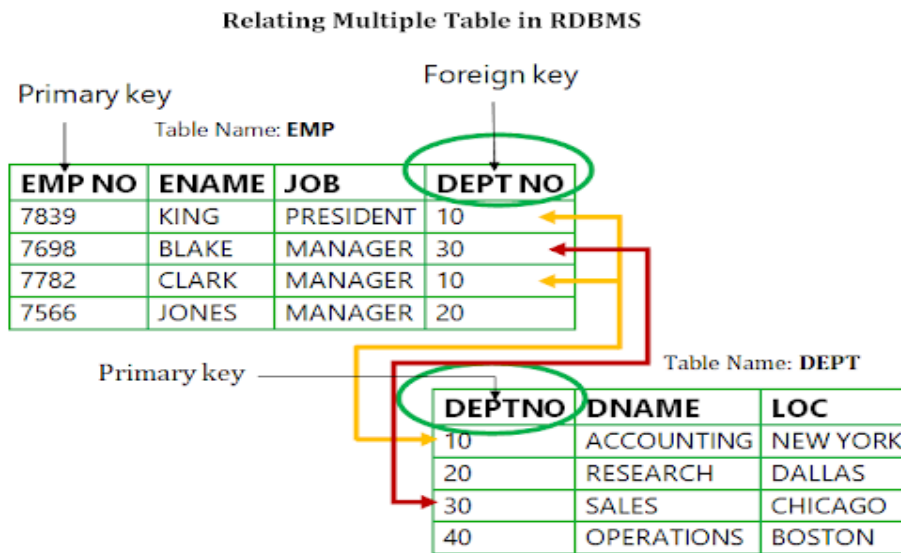


Fig.2. SQL versus NoSQL Databases

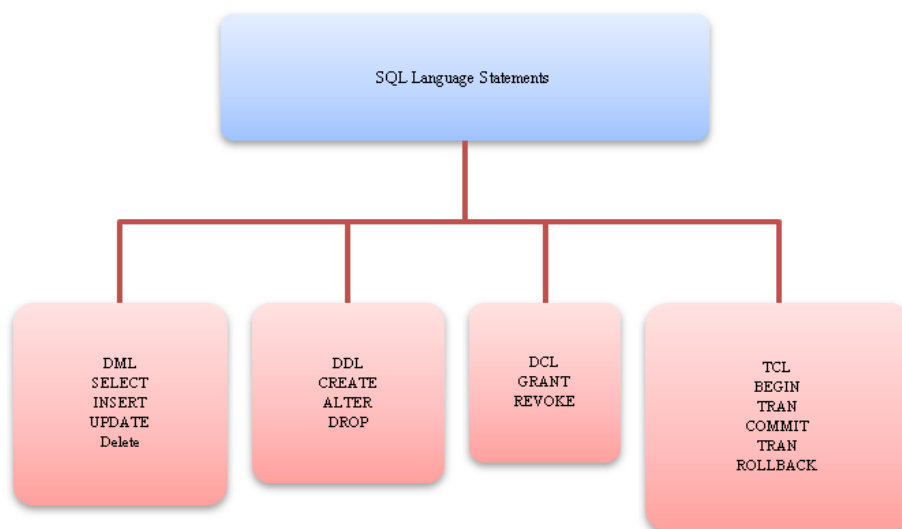


Fig.3. SQL versus NoSQL Databases

2.2. Document Data Model and MongoDB

The Document - oriented Model allows for a large amount of data to be stored within one document then nested into entries. Key-value stores, as the name implies, store keys and have a simple data format. Wide columns can store more features, data type variables, and active columns than row-based relational databases. The relational table document in MongoDB is shown in Figure 4. MongoDB, a NoSQL DBMS, is used for vast data storage. Documents, collections, and fields are all used. A document is the primary data type in MongoDB. Record sets and related tables, such as features, are included in the collections. [10].

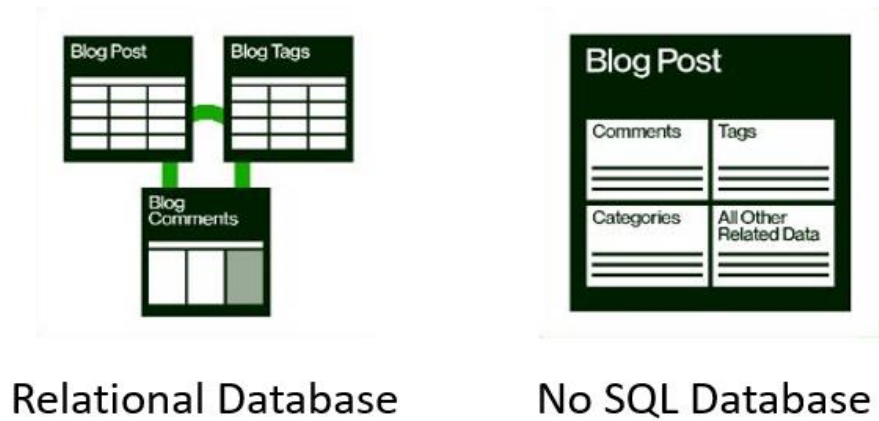


Fig.4. MongoDB Relational Documents two tables [11]

2.3. Comparisons SQL and NoSQL Database

The primary difference between the two concepts is the relational nature of the SQL database and the presence of foreign keys. In contrast, the NOSQL database is unchangeable and does not identify the relationships. various characteristics of both databases are shown in Table 1.

Table 1. Comparison between SQL and NoSQL both databases [12]

Property	SQL	NoSQL
The Method of Data Storage	Tabular Documents	Major Value
Data organization	Schema is predefined in SQL	Schema is dynamic in NoSQL.
Scalable	Vertically (Huge RAM, Strong Processor)	Horizontally (Extra Servers, Instances)
Language	Standard Query Language	Customized Query Language
Data Interaction	Relation Key	Embed Document
Safety	Isolated, Consistent, Transactions,	Non-existent

NoSQL may appear to offer more advantages than SQL, although this may not always be the case. Although NoSQL can read and write data, it has no security features. Terabytes of data are employed for big data applications as a result when they are anticipated. If we need to develop a thorough full-text search tool that considers relevant keywords, copying, or other rules of grammar into account,

NoSQL is an excellent solution. NoSQL also addresses the issue of uncertain database schematics. The drawback, on the other hand, is that we must deal with data consistency on the application side. NoSQL databases have the additional flaw of not supporting transaction processing. [13]. There are many more varieties of NoSQL than merely those that support full-text searches. Their uses and storage layouts differ. We are aware of the document graph in the NoSQL database, as displayed in the table. We utilize the key value database since we require a location to keep the information when utilizing the cache. Redis and Memcached are two of the most well-known. In the chart database, social networking sites appear first.[14]. Table 2 compares the features of SQL and NoSQL in a brief manner.

Table 2. Distinct characteristic between SQL vs NoSQL Databases [15]

Distinct Characteristic	SQL	NoSQL
Relational	It is relational	It is non-relational
Standard Query Language	It is a standard query language	It is not a standard query language
Document	It is a table-based database	It is a document-based database
Online Transaction Processing	It is not supported in SQL	It is fully Supported in NoSQL
Scalability	It is not scalable	It is highly scalable
Complexity of the Query	Complexity of the query is low	Complexity of the query is high
Distributed	Distributing computing is not used	This database uses distributed computing

3. ANALYSIS OF THE SQL AND NOSQL DATABASES PERFORMANCE

A high-level, open-source programming language called PHP is primarily employed in database and enterprise applications. By connecting the MySQL database and the MongoDB using the PHP programming language, merging two tables using the MySQL database's join query and the identical aggregation feature inner join solution in MongoDB, Utilizing the three different programs, the first program processes 10,000 records, the second program processes 20,000 records, and the third program processes 30,000 records, as shown in table 2. These programs are all operating on a computer the entire time. The computer's setup is described below.

- Second Generation, Intel(R) Core (TM) i5-2520M CPU @2.50 GHZ
- 4 GB RAM
- Operating system is 64-bit.
- Windows 10.
- 500 GB hard disc.

For empirical research, both SQL databases have been employed (MySQL and MongoDB). A database table and a technique combining inner join of SQL and aggregate function of NoSQL databases were used to compare the loading, response, and retrieval times of two databases.

4. RESULTS AND DISCUSSION

The study compares the performance of SQL and NoSQL databases based on file size and evaluates the performance characteristics in seconds using the parameters mentioned in Table3. By comparing performance, you may decide which one meets these criteria better. The evaluation of the inputs and outputs by these parameters, which are mentioned in Tables 3, 4, and 5, is shown in Figures 5, 6, and 7. A SQL database most likely gives far more effective performance

based on loading, response, and retrieval times. The record data input contains 10,000, 20,000, and 30,000 records, respectively. Figure 8 depicts the comparison graph between the records and outcomes from databases utilizing SQL and NoSQL.

Table 3. Result Ten Thousand Records SQL and NoSQL

Performance	SQL	NoSQL
File Size	4.6 mb	6.0 mb
Loading Time	28.12 sec	51.93 sec
Response Time	24.11 sec	16.11 sec
Retrieval Time	0.0034 sec	0.0136 sec

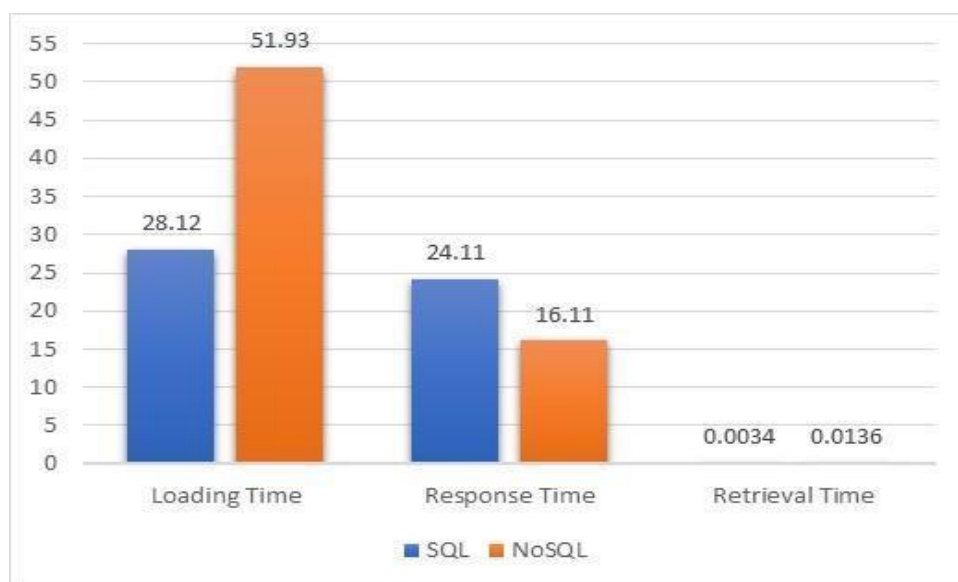


Fig.5. SQL and NoSQL Database Ten Thousand Records

Table 4. Result Twenty Thousand Records SQL and NoSQL

Performance	SQL	NoSQL
File Size	9.2 mb	12.0 mb
Loading Time	90 sec	174 sec
Response Time	24.26 sec	78 sec
Retrieval Time	0.073 sec	0.59 sec

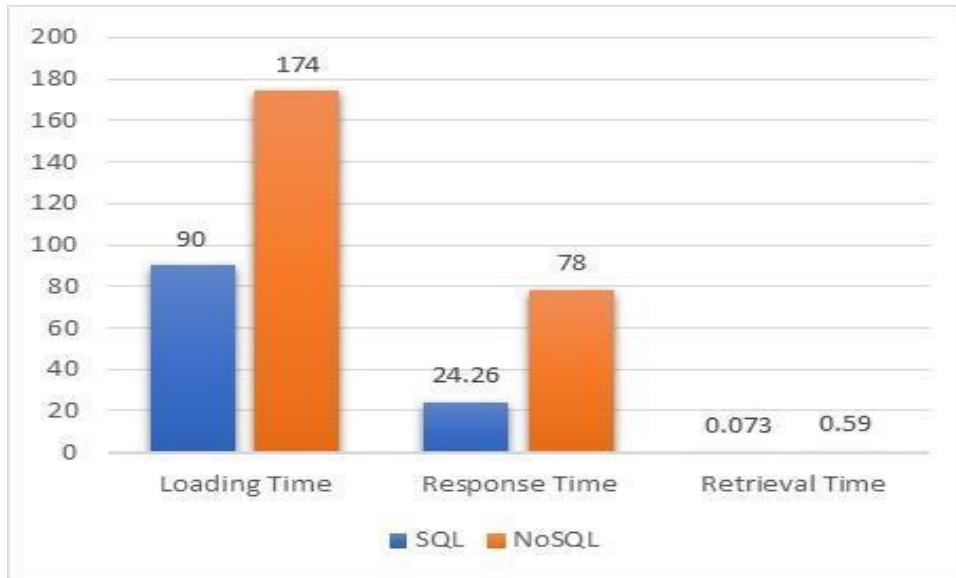


Fig.6. SQL and NoSQL Database Twenty Thousand Records

Table 5. Result Thirty Thousand Records SQL and NoSQL

Performance	SQL	NoSQL
File Size	13.8 mb	11.6 mb
Loading Time	252 sec	258 sec
Response Time	102 sec	138 sec
Retrieval Time	0.02866 sec	0.30031 sec

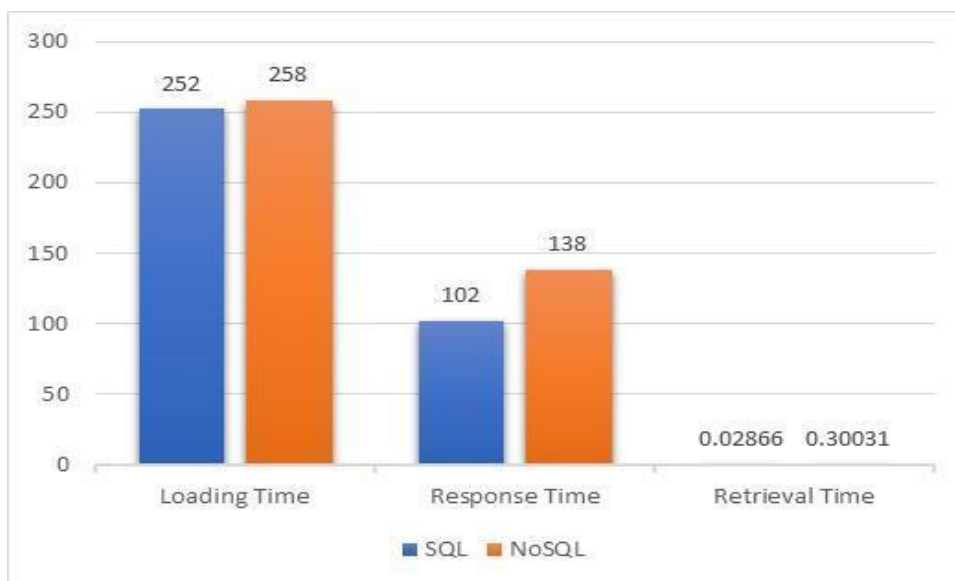


Fig.7. SQL and NoSQL Database Thirty Thousand Records

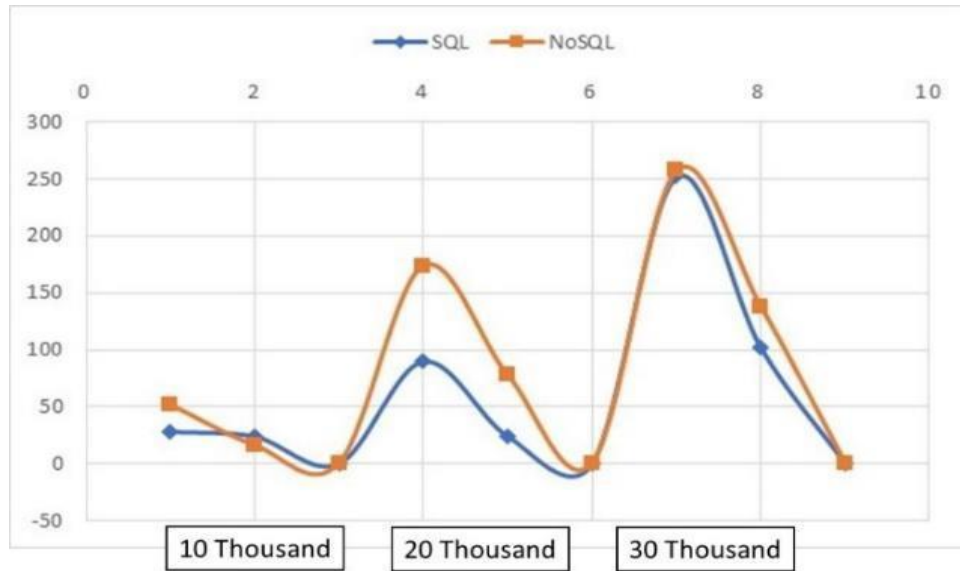


Fig.8. Comparison Graph SQL and NoSQL Database Records

5. CONCLUSION

In this study, the efficacy of SQL and NoSQL databases was investigated. It was feasible to monitor parameter loading times, response times, and retrieval times on both the databases throughout the analysis to differentiate their performance. This was made possible by the database connectivity of the PHP programming language. It has been discovered that the loading, response, and retrieval times for SQL (MySQL) are all considerably quicker than those for NoSQL (MongoDB). These findings show that in terms of loading, responding, and retrieval times, SQL databases are more efficient and swifter than NoSQL databases.

References

- [1] Roman Ceresnak1, Michal Kvet2, "Comparison of query performance in relational a non-relation databases", 13th International Scientific Conference on Sustainable, Modern and Safe Transport (TRANSCOM 2019), High Tatras, Novy Smokovec – Grand Hotel Bellevue, Slovak Republic, May 29-31, 2019
- [2] Nagarajan Madhavan1, 2019, <https://medium.com/@itfsMadhavan/sql-vs-nosql-databases-whats-the-difference-a05492b48d99>
- [3] Elad Israeli1, Eldad Farkash2, Aviad Harell3, Guy Boyangu4, Adi Azaria5, 2004, <https://www.sisense.com/glossary/relational-database/>
- [4] SQL and NoSQL Database, <https://qph.fs.quoracdn.net/main-qimg-548fdcb1dec7865303859def292f1f30>
- [5] Imtiaz Rashid: "Term Paper on Database Management System".
- [6] Don Chamberlin: "SQL", IBM Almaden Research Center, San Jose, CA.
- [7] Mohammad Muhtashim1, 2006, <https://www.tutorialspoint.com/sql/sql-overview.htm>
- [8] MySQL Relational Database Management System (RDMS), <https://i.ytimg.com/vi/cW-rgspRrX8/hqdefault.jpg>
- [9] SQL Database languages, <https://csharpcorner.azureedge.net/UploadFile/BlogImages/01032017000551AM/image002.png>

- [10] Dwight Merriman¹, Eliot Horowitz², Kevin Ryan³, 2007, <https://docs.mongodb.com/manual/core/data-modeling-introduction/#document-structure>
- [11] MongoDB Relational Database Management System (RDMS), <https://image.slidesharecdn.com/mongodbrdbmsmigrationwebinar-140202060406-phpapp02/95/migrating-from-relational-databases-to-mongodb-10-638.jpg?cb=1391321199>
- [12] Lourenço, J., Cabral, B., Carreiro, P., Vieira, M. and Bernardino, J. (2015). Choosing the right NoSql database for the job: a quality attribute evaluation. *Journal of Big Data*, 2(1).
- [13] Mehmood, N., Culmone, R. and Mostarda, L. (2017). Modeling temporal aspects of sensor data for MongoDB NoSql database. *Journal of Big Data*, 4(1).
- [14] Sanchez-de-Madariaga, R., Muñoz, A., Lozano-Rubí, R., Serrano-Balazote, P., Castro, A., Moreno, O. and Pascual, M. (2017). Examining database persistence of ISO/EN 13606 standardized electronic health record extracts: relational vs. NoSql approaches. *BMC Medical Informatics and Decision Making*, 17(1).
- [15] Sneha Binani¹, Ajinkya Gutti², Shivam Upadhyay³, "SQL vs. NoSQL vs. NewSQL- A Comparative Study", *Communications on Applied Electronics (CAE) – ISSN : 2394-4714 Foundation of Computer Science FCS, New York, USA Volume 6 – No.1, October 2016*.
- [16] Nidamanuri Amani¹, Yelchuri Rajesh², "Comparative Study of Open-Source NOSQL Document-Based Databases", 2020 *Information and Communication Technology for Intelligent Systems (ICTIS)*.
- [17] Zakariyaa Ait El Mouden¹, Abdeslam Jakimi², Moha Hajar³, Mohamed Boutahar⁴, "Graph Schema Storage in SQL Object-Relational Database and NoSQL Document-Oriented Database: A Comparative Study", 2019 *Innovation in Information Systems and Technologies to Support Learning Research*.
- [18] Wisal Khan¹, Waqas Ahmad², Bin Luo, Ejaz Ahmed³, "SQL Database with physical database tuning technique and NoSQL graph database comparisons", 2019 *IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*.
- [19] Fatima Yassine¹, Mamoun Adel Awad², "Migrating from SQL to NOSQL Database: Practices and Analysis", 2018 *International Conference on Innovations in Information Technology (IIT)*.
- [20] Sunita¹ Ghotiya², Juhi Mandal³, Saravanakumar Kandasamy¹, "Migration from relational to NoSQL database", 2017 *IOP Conference Series: Materials Science and Engineering*.
- [21] Jitender Kumar¹, Varsha Garg², "Security analysis of unstructured data in NOSQL MongoDB database", 2017 *International Conference on Computing and Communication Technologies for Smart Nation (IC3TSN)*.
- [22] Cornelia Györödi¹, Robert Györödi², George Pecherle³, Andrada Olah⁴, "A comparative study: MongoDB vs. MySQL", 2015 *13th International Conference on Engineering of Modern Electric Systems (EMES)*.
- [23] Matthew Aslett: "451 Research delivers market sizing estimates for NoSQL, NewSQL and MySQL ecosystem" May 22nd, 2012
- [24] Shiwei Yu: "ACID Properties in Distributed Databases", 20096
- [25] Ryan Betts: "SQL Time-Tested and still flourishing", VoltDB, Bedford.
- [26] Jenny Richards, Advantages and Disadvantages of NoSQL databases – what you should know, Hadoop360, September 24, 2015,
- [27] Venkat Gudivada, Dhana Rao and Vijay Raghavan: "NoSQL Systems for Big Data Management", June 2014, Conference: 2014 IEEE World Congress on Services, At Anchorage, Alaska.
- [28] A B M Moniruzzaman: "NewSQL: Towards Next Generation Scalable RDBMS for Online Transaction Processing (OLTP) for Big Data Management", Nov 2014.
- [29] Hoff, Todd: "Google Spanner's Most Surprising Revelation: NoSQL is Out and NewSQL is In". Retrieved 2012-10-07.
- [30] Stonebraker, Michael "NewSQL: An Alternative to NoSQL and Old SQL for New OLTP Apps". *Communications of the ACM Blog.*, 2012
- [31] Vinay Jain "Rise of NewSQL". *International Journal for Research in Emerging Science and Technology*.

- [32] Scholz, J.: Coping with Dynamic, Unstructured Data Sets – NoSQL: a Buzzword or a Savior? In: Schrenk, M., Popovich, V.V., Zeile, P. Proc. REAL CORP 2011 May, 1–9 (2011).
- [33] Stonebraker, M. et al.: C-store: a Column-Oriented DBMS. In: Proc. 31st VLDB Conference. pp. 553–564 (2005).
- [34] Zaki, A.K.: NoSQL Databases?: New millenium database for big data, big users, cloud computing and its security challenges. 403–409 (2014).
- [35] The Vertica R · Analytic Database Technical Overview White Paper (2010).
- [36] Abadi, D. et al.: Integrating Compression and Execution in Column-Oriented Database Systems. In: Proc. 2006 ACM SIGMOD Int. Conf. Manag. data – SIG-MOD’06. pp. 671 (2006).
- [37] Abadi, D.J. et al.: Column-Oriented Database Systems. In: Proc. VLDB ’09. pp. 1664–1665 (2009).
- [38] Abadi, D.J., Madden, S.R.: Column-Stores vs. Row-Stores?: How Different Are They Really? In: SIGMOD’08. pp. 967–980 (2008).
- [39] Angles, R., Gutierrez, C.: Survey of graph database models. ACM Comput. Surv. 40(1), 1–39 (2008).
- [40] Harizopoulos, S. et al.: Performance Tradeoffs in Read-Optimized Databases. In: Proc. VLDB Endowment. pp. 487–498 (2006).
- [41] Kyurkchiev, H., Kaloyanova, K.: Performance Study of Analytical Queries of Oracle and Vertica. In: Proc. 7th ISGT Int. Conf. pp. 127 – 139 (2013)
- [42] Lamb, A. et al.: The Vertica Analytic Database?: C-Store 7 Years Later. In: Proc. VLDB Endowment. pp. 1790–1801 (2012).
- [43] Mitreva, E., Kaloyanova, K.: NoSQL Solutions to Handle Big Data. In: Proc. Doctoral Conference in MIE. pp. 77–86 (2013).
- [44] Neil, B.O. et al.: The Star Schema Benchmark and Augmented Fact Table Indexing Outline of Talk. In: Nambiar, R., Poess, M. (eds.) Performance Evaluation and Benchmarking. LNCS, vol. 5895, pp. 237–252. Springer, Heidelberg (2009).
- [45] Padhy, R.P. et al.: RDBMS to NoSQL: Reviewing some next-generation non-relational database’s. Int. J. Adv. Eng. Sci. Technol. 11, 15–30 (2011).
- [46] Redmond, E., Wilson, J.R.: Seven Databases in Seven Weeks. The Pragmatic Book-shelf (2012).