INSY6112 Assignment 1

Mikhaeel Mahomed

ST10440073

Question1

I recommend a NoSQL database for the company to use for data storage due to its flexibility, scalability, and performance. It can handle dynamic content and user interactions in a seamless experience (geeksforgeeks, 2025).

A NoSQL database is any database that is not relational and stores data in a non-tabular format, allowing for flexibility and scalability (Google, 2025).

The platform generates large amounts of data. A NoSQL database uses horizontal scalability, which is adding more nodes to a database cluster, allowing the database to handle large amounts of data and high levels of traffic (geeksforgeeks, 2025).

A NoSQL database uses a flexible schema. Since the schema of the database is dynamic, it can house changing data structures and does not need schema alterations and migrations (geeksforgeeks, 2025).

NoSQL databases have fast read/write operations which allow a high and fast performance, which is necessary for a seamless user experience (geeksforgeeks, 2025).

There are four types of NoSQL databases:

Key-Value: This type of database has two columns: key and value, key containing strings and value containing JSON, strings, Blob, XML, etc (Katz, 2022). The reason for this design is to have a hash table with a unique key and a pointer to a data item. This design allows for the database to deal with large amounts of data, like the amount of data generated from the users on the platform (Katz, 2022).

Column-Oriented: As the name suggests, this database type stores data in columns instead of rows (Katz, 2022). This design is more efficient in reading and retrieving data, consuming less memory when analytics are run (geeksforgeeks, geeksforgeeks, 2025). This type of database is used to store large amounts of data.

Document-Oriented: This database type is similar to key-value databases; the data is stored in documents in a nested structure of keys and values (amazon, 2025). The documents are retrieved with keys, and the data in the document ranges from simple data types to arrays or only parts of a document (amazon, 2025). This database type improves query speed and flexibility by supporting nested documents and indexing (Katz, 2022).

Graph: In graph-based databases, each data element is stored as a node, and edges show the relationships between each node (Katz, 2022). In this type of database, no tables allow flexibility and scalability across different devices, which is essential for a social media platform (Katz, 2022). These databases are ahead in terms of performance regarding queries (amazon, 2025).

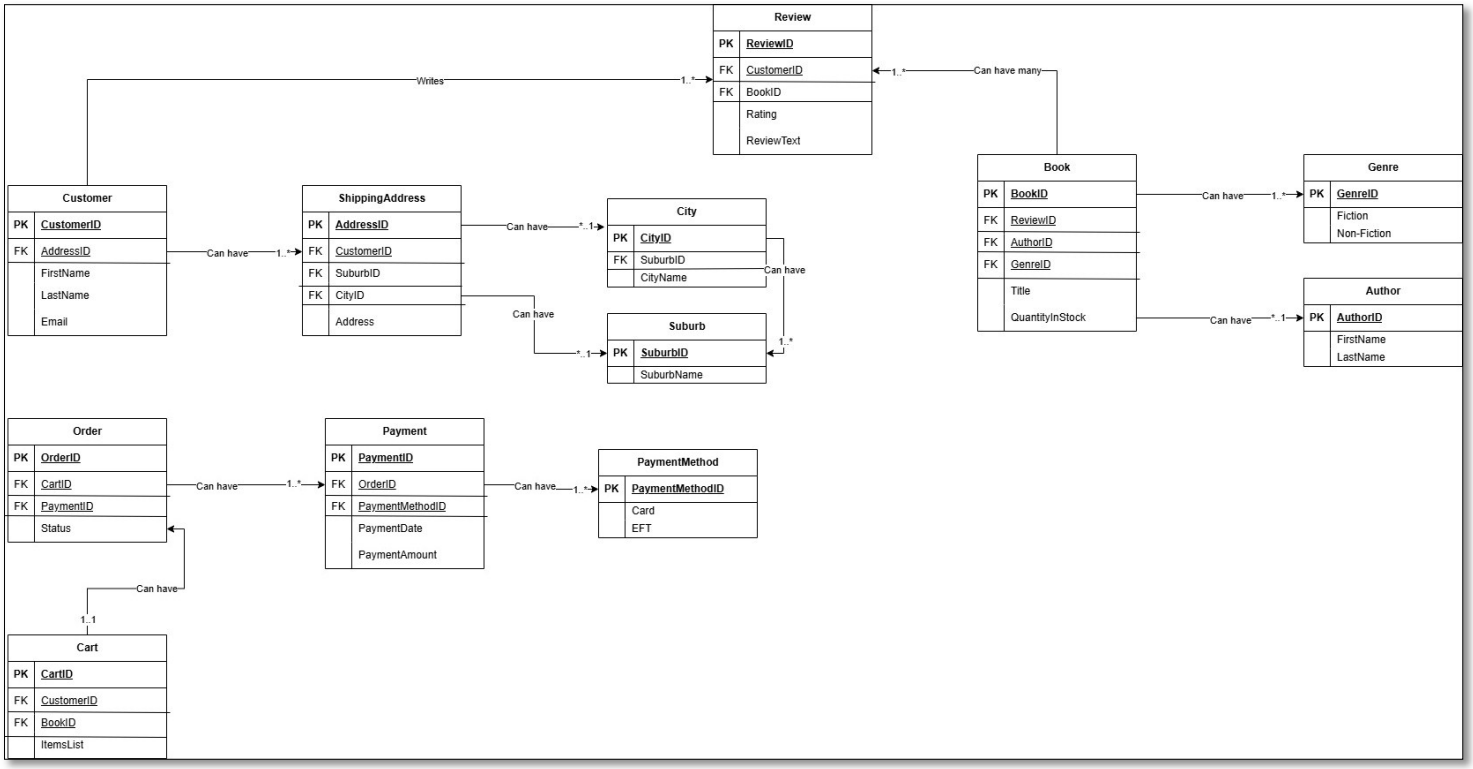
The three Vs of big data regarding a social media platform are:

Volume: This refers to the massive amount of data that is generated from users. Their posts, comments and interactions. This creates storage concerns since terabytes worth of data are being generated each day (BIG DATA LDN, 2018). The NoSQL database covers this concern due to its horizontal scalability, allowing it to handle large amounts of data and high traffic levels.

Velocity: This refers to the speed at which data is added to the database (Aggarwal, 2024). Since people around the world, on different devices and networks, are always online, interacting on social media at any given time, the database should be able to receive data at any time, requiring near real-time processing and analysis to stay relevant and responsive. NoSQL's fast read/write operations allow for a seamless and responsive experience.

Variety: This refers to the variety of data types that are stored in the database, including structured and unstructured data (brain, 2024). This data could come from many different interactions, such as status updates, videos, comments, etc (Aggarwal, 2024) . NoSQL databases have a flexible schema that allows the database to store changing data structures and many different types of data.

Question2



(draw.io)

References

Aggarwal, S. (2024, May 9). *qubole*. Retrieved from www.qubole.com: https://www.qubole.com/blog/social-media-marketing-best-practices-big-data

amazon. (2025). *aws.amazon*. Retrieved from docs.aws.amazon.com: https://docs.aws.amazon.com/whitepapers/latest/choosing-an-aws-nosql-database/types-of-nosql-databases.html

BIG DATA LDN. (2018, August 10). *bigdataldn*. Retrieved from www.bigdataldn.com: https://www.bigdataldn.com/en-gb/blog/data-engineering-platforms-architecture/big-data-the-3-vs-explained.html

brain, s. (2024, February 9). *ssp*. Retrieved from www.ssp.sh: https://www.ssp.sh/brain/three-vs-volume-velocity-and-variety/#:~:text=Big%20Data%20is%20a%20term,updates%20on%20social%20media%20platforms).

geeksforgeeks. (2025, March 12). *geeksforgeeks*. Retrieved from www.geeksforgeeks.org: https://www.geeksforgeeks.org/introduction-to-nosql/

geeksforgeeks. (2025, February 25). *geeksforgeeks*. Retrieved from www.geeksforgeeks.org: https://www.geeksforgeeks.org/types-of-nosql-databases/

Google. (2025). *cloud.google*. Retrieved from cloud.google.com: https://cloud.google.com/discover/what-is-nosql

Katz, E. (2022, June 22). *spectralops*. Retrieved from spectralops.io: https://spectralops.io/blog/nosql-databases/