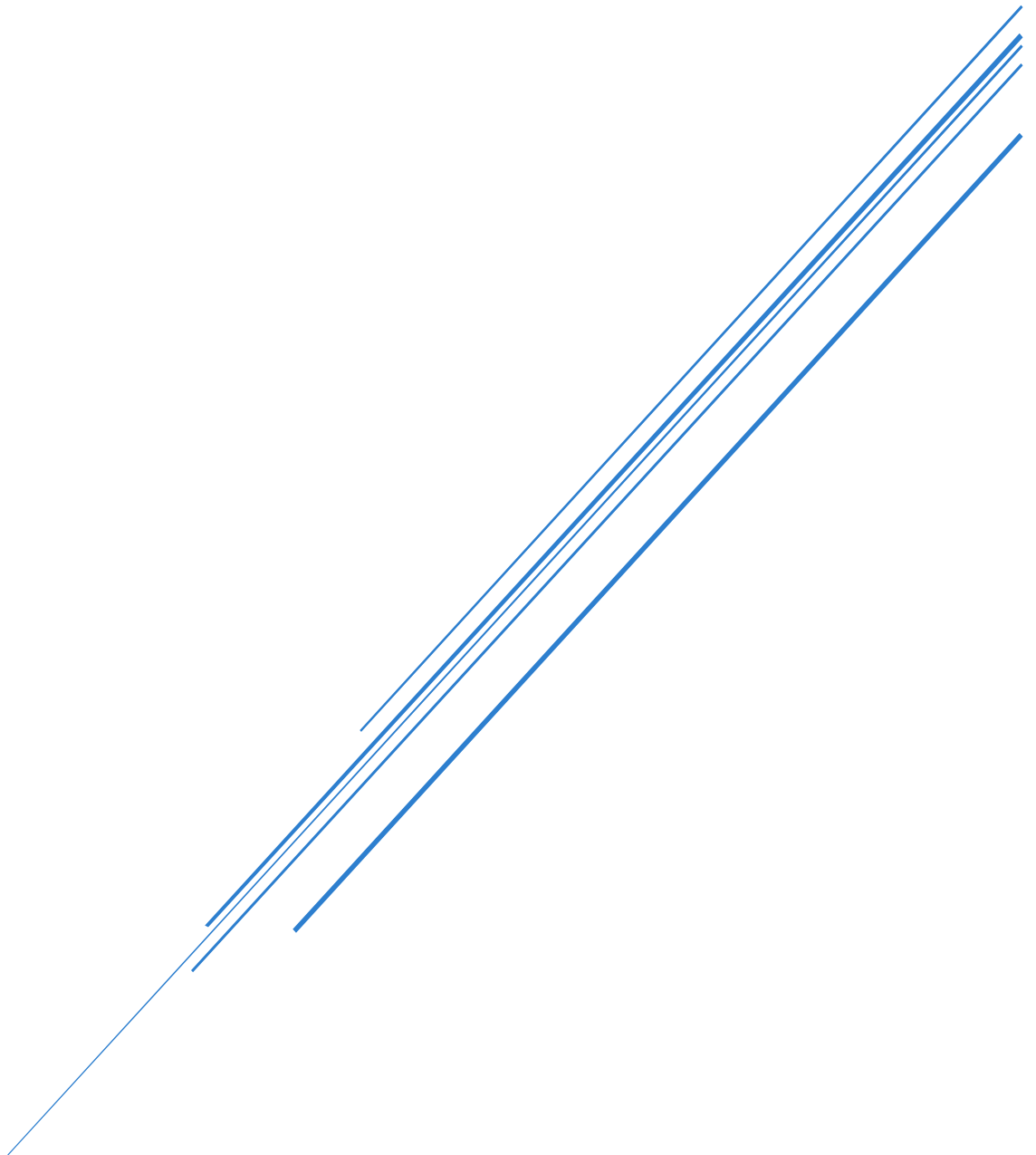


INSY6112 ASSIGNMENT

ST10451846



28/03/2025

CONTENTS PAGE

Question 1	page 2 – 4
Question 2	page 5
References	page 6 – 7

Question 1

Recommended data storage plan for a social media platform

The recommended database for the social media platform should be NoSQL (Not only SQL). It is a database which is designed in such a way to be able to control large amounts of data that is both structured and semi-structured (GeeksforGeeks, 2025). It has more adaptable structures that would make storing different types of data like photos and videos become easily to do and it is enabled to potentially expand, which makes this kind of database more useful when it comes to modern applications e.g. Instagram (Amazon,2025).

Motivation

It is recommended because this type of database can expand like add more servers to the system as the number of users would grow which can be helpful in both the handling and storing of large amounts of data (Amazon,2025). In this case, since a social media platform uses all different kinds of data e.g. photos and videos, NoSQL can successfully manage all the different data in its database (GeeksforGeeks, 2025). Due to the social media platform having live actions implemented in it e.g. commenting on a post, this type of data needs to be handled as soon as possible which NoSQL can do due to it having a good performance quality and is able to control large amounts of internet traffic and data compared to traditional databases (GoogleCloud, 2025). Also, NoSQL doesn't have a fixed database structure which means it is able to adjust if there are changes in the data e.g. increase in the number of users without having to make major adjustments to its database structure (GeeksforGeeks, 2025).

Kinds of data that would be stored in the database:

User Data

It is usually data that has the user's personal information e.g. name and phone number, that the user would use to either create an account or login into it and it would help the platform to suggest content to the user based on their likes (Amazon,2025). By having this kind of data, it helps to authenticate the user so that the correct account would be loaded, for example if a user decides to login into their account, the system needs to be able to load the data which is obtained from the user to identify that the correct user is login in (GeeksforGeeks, 2025).

Content Data

This type of data is used when users make their own content like their own posts on their social media page and when they would share information with their followers like a story update (GoogleCloud, 2025). For example, if a picture would be posted on the user's page, that data needs to be kept and accessed immediately which would allow the picture to be shown on their media page (InterSystems,2025).

Interaction Data

It is data that shows how the users would be able to interact with the social media platform and with other users like their followers (Amazon,2025). This would help the business by knowing what changes could be made on the platform in order to improve user experience (InterSystems,2025). Example, when an image is posted, the system needs to update the likes the photo gets and needs to notify the user that posted the image about each time the image is liked (GeeksforGeeks, 2025).

Relationship Data

This data would show how the users are connected to one another on the platform and also the interaction between them like tagging in a post or a mention in a story (Amazon,2025). For example, if a user has to tag another user in a post, the system needs to update recent activities and needs to notify the user which was tagged (InterSystems,2025).

Recommended Databases:

Document based databases

It is a type of database that saves data usually in a document structure like JSON (MongoDB, 2025). This helps to decrease the translation of the usage of data in the applications (GoogleCloud, 2025). By doing this, it helps to reduce the converting of data in the applications due to the document structure which the data is kept which leads to the application being able to understand the data easily (GeeksforGeeks, 2025). Also, in this kind of database there are certain data parts which can be used to access data that is being looked for by just searching for it by using the special index that it comes with, this leads to searching become successful (InterSystems,2025). In this scenario, this kind of database would be useful when it comes to storing both the user's profiles and posts since both have structured and semi structured data like photos (MongoDB, 2025).

Key-Value Stores

It is one of the simplest structures in a NoSQL database. In this database, the data is kept in a "Key-value" format, which means that a key like a user's name which would be joined to a certain value like a number (InterSystems,2025). The key can be used to either keep the value or get the value (GeeksforGeeks, 2025). This type of database is mostly used for user profiles or shopping carts (HazelCast, 2025). In the social media platform, this type of database can store and receive data that has a basic structure form e.g. username. This helps the users account to be able to load quicker and their feed which is customised from their likes (InterSystems,2025).

Column Database

This type of database stores data in columns compared to storing it in rows (TechTarget, 2025). This makes access to data efficient because not a lot of memory would be needed to use for data that is not required (InterSystems,2025). It was created in such a way in which data would be easy to find and read as well as it would become faster and more successful (GeeksforGeeks, 2025). It is commonly

used due to being able to store large volumes of data. In this kind of platform, this database would be useful when it comes to keeping and accessing large bundles of data that are statistical e.g. likes, comments (TechTarget, 2025).

Graph Database

It is database that is created to manage data that have complicated relationships. So, in this database the data is kept as entities as well as edges (GeeksforGeeks, 2025). By having this, it makes this database more acceptable and is also able to manage real-world situations e.g. comments on posts (InterSystems,2025). In this scenario, graph database is useful when it comes to showing how users keep in touch with one another e.g. like mentioning a user or liking a post, this also helps the platform to be able to suggest content to the user, which could be similar to what they have liked in the past (MongoDB, 2025).

Three Big V's of Data:

Volume

According to Rajput (2024), since it's a social media platform, huge amounts of data would be created daily e.g. posts, comments (GoogleCloud, 2025). As the platform would start to grow in the future, the number of users would increase which would thus lead to the volume of data increasing as well e.g. if a post goes viral, the system will need to be able to store the data e.g. comments successfully in the database and it should be easy to access that data as well (GeeksforGeeks, 2025).

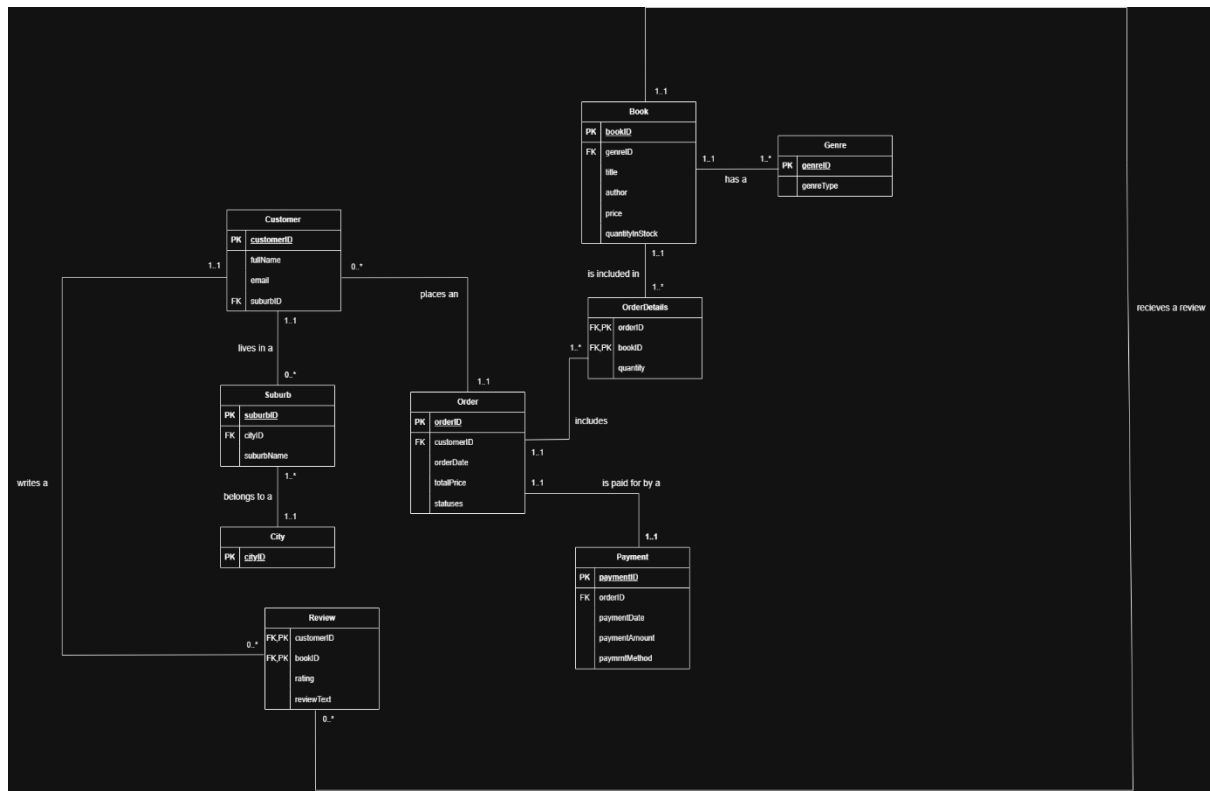
Velocity

According to Rajput (2024), since a lot of data would be created, the speed at which it is created at and being updated as well on the system usually depends on how the users communicate with the app (GoogleCloud, 2025). In this case, the social media platform needs to be able to manage the amount of data that would be delivered as soon as possible to keep up with the updates in the streams and also being able to send out notifications to the users for that specific stream without any disruptions e.g. when a user mentions one of their followers on the story, the user who is mentioned needs to receive a notification instantly(GeeksforGeeks, 2025).

Variety

In this scenario, variety means all different data like GIFS, images that would be created which the social media platform has to manage with (GeeksforGeeks, 2025). According to DuckWorth (2025), the NoSQL database needs to successfully be able to both manage and keep structured, and semi structured without any disruptions. In this scenario, both an image post and comment must show at the same time without any problems occurring e.g. can't show the image properly (GoogleCloud, 2025).

Question 2



In the ERD, the suburb and city are their own entities due to avoiding the reoccurrence of information and also because one suburb belongs to a city, so by doing this it allows for the control of the data to be much simpler. Genre is its own entity as well since it would make arranging books in each category much easier since the genre details won't have to be inserted each time a book is purchased, which also leads to arranging and looking for a book much simpler.

REFERENCES

- Amazon, 2025. *What Is a Graph Database*. [online]. Available at: <<https://aws.amazon.com/nosql/graph/>> [Accessed 21 March 2025].
- Amazon, 2025. *What are NoSQL databases*. [online]. Available at: <<https://aws.amazon.com/nosql/>> [Accessed 18 March 2025].
- DuckWorth, D . 2023. The 3 V's of Big Data: Velocity Remains A Challenge for Many. *HazelCast*, [blog], 4 January 2023. Available at: <<https://hazelcast.com/blog/velocity-remains-a-challenge/>> [Accessed 27 March 2025].
- GeeksforGeeks, 2025. *Difference between Relational database and NoSQL*. [online]. Available at: <<https://www.geeksforgeeks.org/difference-between-relational-database-and-nosql/>> [Accessed 18 March 2025].
- GeeksforGeeks, 2025. *Introduction to NoSQL*. [online]. Available at: <<https://www.geeksforgeeks.org/introduction-to-nosql/>> [Accessed 18 March 2025].
- GeeksforGeeks, 2025. *What is a Columnar Database*. [online]. Available at: <<https://www.geeksforgeeks.org/what-is-a-columnar-database/>> [Accessed 19 March 2025].
- GeeksforGeeks, 2025. *Introduction to Graph Database on NoSQL*. [online]. Available at: <<https://www.geeksforgeeks.org/introduction-to-graph-database-on-nosql/>> [Accessed 21 March 2025].
- GeeksforGeeks, 2025. *Types of NoSQL Databases*. [online]. Available at: <<https://www.geeksforgeeks.org/types-of-nosql-databases/>> [Accessed 19 March 2025].
- GoogleCloud, 2025. *What is a NoSQL database*. [online]. Available at: <<https://cloud.google.com/discover/what-is-nosq>> [Accessed 19 March 2025].
- GoogleCloud, 2025. *What is Big Data*. [online]. Available at: <<https://cloud.google.com/learn/what-is-big-data>> [Accessed 21 March 2025].
- HazelCast, 2025. *What is a key-value store*. [online]. Available at: <<https://hazelcast.com/foundations/data-and-middleware-technologies/key-value-store/>> [Accessed 27 March 2025].
- InterSystems, 2025. *NoSQL Databases Explained: Advantages, Types, and Use Cases*. [online]. Available at: <<https://www.intersystems.com/za/resources/nosql-databases-explained-advantages-types-and-use-cases/>> [Accessed 18 March 2025].
- MongoDB, 2025. *What is a Document Database*. [online]. Available at: <<https://www.mongodb.com/resources/basics/databases/document-databases>> [Accessed 27 March 2025].
- MongoDB, 2025. *What is NoSQL*. [online]. Available at: <<https://www.mongodb.com/resources/basics/databases/nosql-explained>> [Accessed 27 March 2025].

Rajput, M. 2024. Three V's of Big Data. *LinkedIn*, [blog], 22 January 2024. Available at: < <https://www.linkedin.com/pulse/three-vs-big-data-monika-rajpud-jjfqc/> > [Accessed 21 March 2025].

TechTarget, 2025. *Columnar database*. [online]. Available at: < <https://www.techtarget.com/searchdatamanagement/definition/columnar-database> > [21 March 2025].