INFORMATION SYSTEMS

ASSIGNMENT 1
INSY6112

NAME: ELISHUA EMMANUEL NAIDOO

STUDENT NUMBER: ST10438817

LECTURER: TALIA DHANAMPAL

QUESTION 1

Based on the scenario outlined, which likely involves large volumes of data, unstructured or semi-structured data, high-velocity data streams, and the need for horizontal scalability, a NoSQL database is the recommended choice.

DEFINITION OF NoSQL DATABASE

NoSQL databases store non-relational data operating at exceptional performance levels, ***(Introduction to nosql, 2024***). NoSQL databases store large amounts of various data in an unstructured and semi-structured format, ***(Introduction to nosql, 2024***). Thus, making it flexible and easy to interact with. NoSQL uses key-value pairs to store unstructured and semi-structured data, ***(Introduction to nosql, 2024***). It can increase and decrease in size with ease. There are many categories or sub sections of NoSQL databases. NoSQL databases are commonly used for storing data from social media apps, ***(Introduction to nosql, 2024***), thus making it the perfect database for the above scenario.

MOTIVATION FOR NoSQL DATABASE

High Volume Data: Working with a social media platform automatically implies that there will be a large number of users combined with an ever-growing number of users thus the volume of data will grow exponentially as a result. So, our database needs to be scalable to accommodate for such factors. NoSQL databases are known for their scalability, one of their major characteristics and they can distribute data across multiple servers effortlessly, ***(Introduction to nosql, 2024***). NoSQL databases are prepared to handle high volume data as quickly and efficiently as possible, ***(Introduction to nosql, 2024***).

Complex Data Structures: A social media platform contains a variety of different types of data and provides a variety of ways to interact with the data. NoSQL databases can handle a variety of data allowing for flexible data models, ***(Introduction to nosql, 2024***). NoSQL databases can store data ranging from: text, images, videos, GIFs, likes, shares and comments, which are all typically found on a social media platform, ***(Introduction to nosql, 2024***). It can also cater for the introduction of new data types, due to its flexibility, ***(Introduction to nosql, 2024***).

Real-Time Analytics: Social media platforms must provide trending topics and updating the users with the latest news. NoSQL databases are designed for high performance levels dealing with large amounts of various data, ***(Introduction to nosql, 2024***). It is designed to read and write large amounts of data, ***(Introduction to nosql, 2024***). Thus, making it suitable for real-time applications like a social media platform which needs instantaneous updates of data to provide to users.

DIFFERENT TYPES OF DATA STORED

Demographic Data: This refers to user personal and account information about the user, ***(Binnewies, 2022)***. This includes profile information such as name, date of birth, profile picture, and bio, ***(Binnewies, 2022)***. As well as account settings such as privacy settings (public or private account), notification settings, and restricting who can interact with your content, ***(Binnewies, 2022)***. Authentication data to restrict who gain access to your social media account, ***(Binnewies, 2022)***.

Psychographic Data: This refers to the user's personality and lifestyle based on the content created and shared, ***(Binnewies, 2022)***. It allows users to express themselves in a unique way, embracing their unique individuality. Other users can also gather information about the type of person they are following or interacting with, ***(Binnewies, 2022)***. It ranges from posts, captions, comments, images, videos, and live streams, ***(Binnewies, 2022)***.

Location Data: This refers to the user's actual real-time location and location at which certain pictures were shot (geotagging), ***(Binnewies, 2022)***. This all provides information about user activity, where and what exactly the user enjoys doing, ***(Binnewies, 2022)***.

Product/Advertisement Data: This refers to a product's social media platform to promote their product or a user who is sponsored to promote certain products for a company, ***(Binnewies, 2022)***. Social media advertisement helps a company to promote their product to influence people to buy, thus increasing profitability, ***(Binnewies, 2022)***. It helps companies deliver personalised advertisements, ***(Binnewies, 2022)***.

TYPES OF NoSQL DATABASES

NoSQL-Document Store: It is a non-relational database that stores data in a document format, typically using JSON, BSON and XML documents, (***Types of databases***). Documents can be accessed, stored and retrieved from a network in the form of data objects and the data must be assembled and disassembled when the data is transferred, (***Types of databases***). NoSQL Document Databases offers flexibility to

improve the structure of documents to satisfy the needs of an application, as a result speeds up development, **(*Types of databases*)**.

NoSQL-Key Value Store: It is a non-relational database that stores data as key value pairs which consist of two columns: an attribute name and a value, **(*Types of databases*)**. Each key is unique and is allocated a value, then the key-value pairs are stored on separate records, ***(Key-value data model in nosql, 2022)***. A NoSQL-Key Value Database reflects an array which is controlled by a DBMS, ***(Key-value data model in nosql, 2022)***.

NoSQL-Column Store: It is a non-relational database that stores data in rows and columns, it reads the data row by row as it is organised in columns, **(*Types of databases*)**. Columns are of the same data type making it easier and faster to read but writing to the database is extremely challenging, **(*Types of databases*)**. It is great for performing analytics, it can run analytics of a number of columns, for example: providing the aggregate value of a column, **(*Types of databases*)**.

NoSQL-Graph Store: It is a non-relational database that stores data in graph structure, focusing on the relationship between elements, **(*Types of databases*)**. A node contains an element, and a relationship is created between the different connected elements, **(*Types of databases*)**. The connections between elements in a graph database can be captured and searched in order to defeat the purpose of joining tables in SQL, **(*Types of databases*)**.

3 V's OF BIG DATA (VOLUME, VARIETY AND VELOCITY)

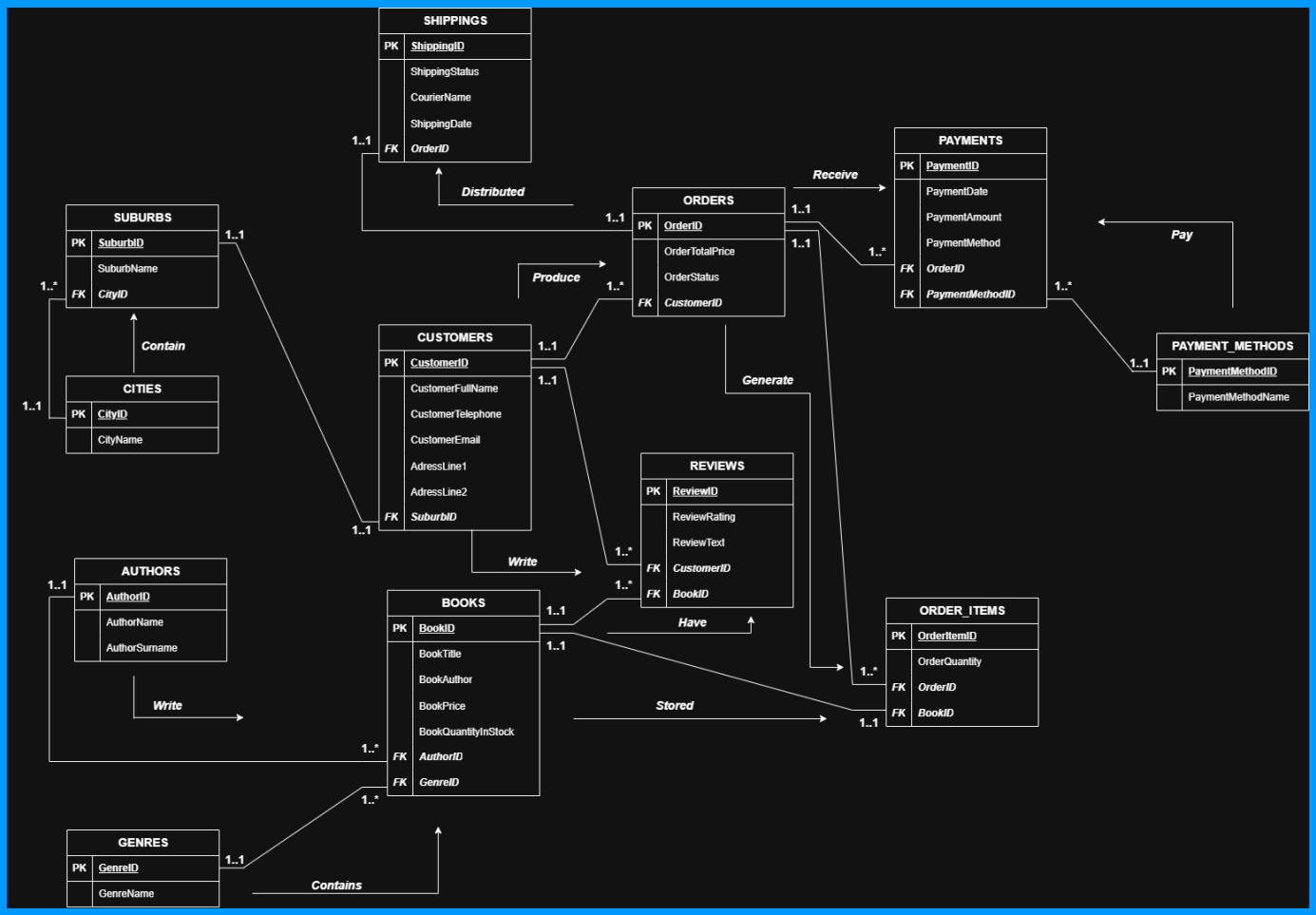
Volume: Volume refers to the quantity or capacity of data that currently exists, keeping in mind that this volume can increase in size, ***(Robinson & Gillis, 2023)***. In terms of a social media platform, there are millions of users interacting on the platform in a variety of ways such as posts, comments, likes, shares, images, and videos, which all contributes to a large volume of data, ***(Sellas, 2023)***. Social media platforms always have an ever-growing user base, so the amount of data is going to increase exponentially, ***(Sellas, 2023)***. As a result, the database needs to be able to increase in volume effortless. NoSQL databases are designed to deal with high volumes of data as they can scale horizontally, ***(Introduction to nosql, 2024***).

Variety: Variety refers to the structured, unstructured and semi-structured data as well as the various data types, ***(Robinson & Gillis, 2023)***. Texts, tweets, pictures, and videos fall under structured data, ***(Understanding the 3 vs of big data - volume, velocity and variety, 2025)***. Social media platforms have a wide range of data ranging from:

Demographic Data, Psychographic Data, Behavioural Data, Product Data, Referral Data, Location Data, and Intention Data, ***(Binnewies, 2022)***, thus we can also conclude there is a wide range of various data types. NoSQL-Document Store and NoSQL-Key Value Store are suited for dealing with this variety because of their flexibility and ability store a variety of data, ***(Introduction to nosql, 2024***).

Velocity: Velocity refers to how fast can the data be generated, ***(Robinson & Gillis, 2023)***. Based on our social media platform, millions of users update, interact with, and create: text posts, photos, videos, comments, likes, and shares, ***(Understanding the 3 vs of big data - volume, velocity and variety, 2025)***. The social media platform needs to be able to process high volume data quickly, and effectively to provide real-time analytics such as trending topics and hot posts, ***(Robinson & Gillis, 2023)***. NoSQL databases are equipped for high-speed data processes, enabling the platform to meet real-time analytics, ***(Introduction to nosql, 2024***).

QUESTION 2


SHIPPINGS
PK ShippingID
ShippingStatus
CourierName
ShippingDate
FK OrderID
Distributed
PAYMENTS
PK PaymentID
PaymentDate
PaymentAmount
PaymentMethod
FK OrderID
FK PaymentMethodID
Receive
Pay
SUBURBS
PK SuburbID
SuburbName
FK CityID
Contain
ORDERS
PK OrderID
OrderTotalPrice
OrderStatus
FK CustomerID
Produce
PAYMENT_METHODS
PK PaymentMethodID
PaymentMethodName
CITIES
PK CityID
CityName
CUSTOMERS
PK CustomerID
CustomerFullName
CustomerTelephone
CustomerEmail
AdressLine1
AdressLine2
FK SuburbID
Generate
REVIEWS
PK ReviewID
ReviewRating
ReviewText
FK CustomerID
FK BookID
Write
AUTHORS
PK AuthorID
AuthorName
AuthorSurname
Write
BOOKS
PK BookID
BookTitle
BookAuthor
BookPrice
BookQuantityInStock
FK AuthorID
FK GenreID
Have
Stored
ORDER_ITEMS
PK OrderItemID
OrderQuantity
FK OrderID
FK BookID
GENRES
PK GenreID
GenreName
Contains
1..1
1..*

REFRENCE LIST

1. (No date) Types of databases MongoDB. Available at: https://www.mongodb.com/resources/basics/databases/types#:~:text=NoSQL%20databases%20are%20different%20from,oriented%20databases%2C%20and%20graph%20databases. (Accessed: 25 February 2025).

2. Binnewies, S. (2022) Types of data. FutureLearn. Available at: https://www.futurelearn.com/info/courses/applied-big-data-analytics/0/steps/52407 (Accessed: 26 February 2025).

3. Introduction to nosql (2024) GeeksforGeeks. Available at: https://www.geeksforgeeks.org/introduction-to-nosql/ (Accessed: 25 February 2025).

4. Key-value data model in nosql (2022) GeeksforGeeks. Available at: https://www.geeksforgeeks.org/key-value-data-model-in-nosql/ (Accessed: 25 February 2025).

5. Robinson, S. and Gillis, A.S. (2023) What are the 5 V's of big data? Definition techtarget, Search Data Management. Available at: https://www.techtarget.com/searchdatamanagement/definition/5-Vs-of-big-data (Accessed: 26 February 2025).

6. Sellas, B.B. (2023) Volume of Social Media Interactions: A deep dive*. LinkedIn. Available at: https://www.linkedin.com/pulse/volume-social-media-interactions-deep-dive-brooke-b-sellas/ (Accessed: 26 February 2025).

7. Understanding the 3 vs of big data - volume, velocity and variety (2025) Coforge. Available at: https://www.coforge.com/what-we-know/blog/understanding-the-3-vs-of-big-data-volume-velocity-and-variety (Accessed: 26 February 2025).