Question 1

1. **Recommendation on whether to use a Relational or NoSQL database**

**1.1 Definition of the recommended database type**

A NoSQL database would be recommended based on the scenario of a huge social media network managing massive amounts of varied quickly changing data (text posts, comments, videos, photos, and user interactions).

A non-relational data store with flexible schemas that can handle large data volumes is called a NoSQL database. As opposed to relational databases, which enforce a set table structure, NoSQL systems are capable of managing and storing unstructured or semi-structured data in a number of different formats.

As stated in (The Independent Institute of Education [IIE], 2025: p.11-22). “Databases are appropriate for web-scale applications due to their high-performance reads and writes, horizontal scalability, and schema flexibility”

**1.2 Detailed Motivation**

**High data volume and velocity**

The volume of data rises exponentially as millions of users create posts, comments, and other activities. By adding more common servers, a NoSQL database can expand horizontally to accommodate growing workloads. Compared to vertically scaling a single large relational database server, this method is frequently more expensive.

**Adaptable schema conditions**

Text, photos, videos, GIFs, reactions (likes, shares), and real-time analytics are all examples of social media data. Development can be slowed down and made more difficult by enforcing a strict schema, as in relational databases. Because of NoSQL's schema flexibility, different data types can be stored without requiring frequent schema upgrades.

**Low latency and Real-time analytics**

NoSQL databases are built for fast reads and writes, which are essential for providing end users with real-time feeds, alerts, and analytics. A smooth user experience requires low latency. Relational databases, on the other hand, may find it difficult to handle complicated joins and extensive data updates in real-time settings.

**1.3 Discussion of types of data stored**

* **User interactions and social media feeds:**

NoSQL is effective at managing a variety of quickly evolving information forms, including text, photos, links, and reactions.

* **High-Throughput applications:**

Programs that need to read and write data millions of times at once, including news feeds, live chat, or real-time analytics dashboards.

* **Content management:**

Is the process of storing unstructured or semi-structured data, such as media metadata, user-generated content, and personal information.

(The Independent Institute of Education [IIE], 2025: p.11-22)"Insufficient engagement may result from delays in updating these feeds". Therefore, for the best user experience, a NoSQL database with low latency and fast throughput is essential.

**1.4 Four types of recommended databases**

There are four primary types of NoSQL databases found in modern large-scale systems.

**1 Key-Value shops:**

* Definition: Data should be stored as a set of key-value pairs, each of which should be unique.
* It is ideal for storing user sessions, user profiles, or quickly available information (like Redis or Amazon DynamoDB).

**2 Document stores**

* Definition: Allow arrays and nested fields while storing semi-structured data in documents, usually in the form of JSON or BSON.
* Outstanding for applications and content management systems that need flexible schemas (e.g., CouchDB, MongoDB).

**3 Family stores in columns**

* Definition: Data is stored in tables, but for scalability, columns are spread across several servers; each "column family" can be modified to fit query patterns.
* Excellent for wide-column queries and large-scale analytics (e.g., Apache Cassandra, HBase).

**4 Graph databases:**

* Definition: These databases enable complex relationship queries by storing entities (nodes) and relationships (edges) in a graph structure.
* recommendation engines, social networks, and fraud detection (e.g., Neo4j).

**1.5 The three V’s of big data**

**Volume:**

Describes the enormous volume of data produced by millions of users who share, like, comment, and submit content. Petabytes of data must be efficiently managed and stored by the organization.

**Velocity:**

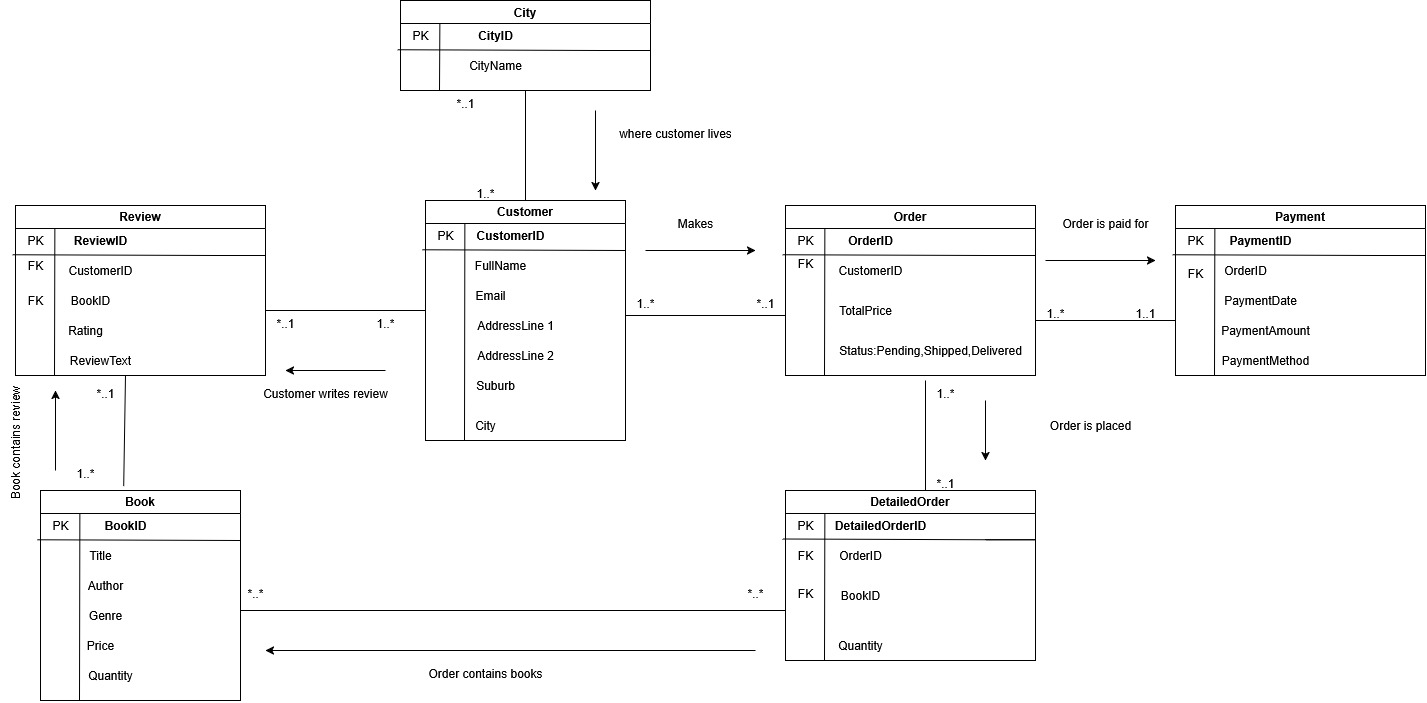
The rate at which information is produced, delivered, and retrieved (for example, real-time feeds and alerts). To prevent delay issues, the platform must receive and serve data at a high throughput.

**Variety:**

Data can be in the form of text, pictures, videos, unstructured user-generated information, and structured metadata. This variety is best handled by NoSQL databases.

In conclusion a NoSQL database system is best suited for the volume, velocity, and variety of data seen in modern social media environments. The platform will stay responsive and scalable thanks to this decision, which tackles the fundamental issues of large traffic, real-time analytics, and schema flexibility. The business can adapt its storage to various use cases by choosing the appropriate NoSQL database type (key-value, document, column-family, or graph), finally using big data (the three Vs) for insightful information and improved user experience.

Question 2



**REFERENCING**

Cattell, R. (2011) ‘Scalable SQL and NoSQL Data Stores’, ACM SIGMOD Record, 39(4), pp. 12–27.

Gandomi, A. and Haider, M. (2015) ‘Beyond the hype: Big data concepts, methods, and analytics’, International Journal of Information Management, 35(2), pp. 137–144.

Han, J., E, H., Le, G. and Du, J. (2011) ‘Survey on NoSQL Database’, International Conference on Pervasive Computing and Applications, 1, pp. 363–366.

Stonebraker, M. (2012) ‘What Goes Around Comes Around’, Communications of the ACM, 55(7), pp. 5–5.

The Independent Institute of Education (IIE) (2025) INYS6112MM: Learning Unit 1: Introduction to Database Concepts, [Module Manual], pp. 11–28.

The Independent Institute of Education (IIE) (2025) INYS6112MM: Learning Unit 2: Designing Relational Databases, [Module Manual], pp. 33–61.