Q1

- **NoSQL**, often spelled "No-SQL" or simply "non-SQL," refers to a design or set of approaches that allows storage and querying of data in non-relational database structures. Other than storing data found within a relational database management system, NoSQL stores it differently than an RDBMS. A choice of whether to use an RDB or otherwise is generally dependent on the context and application. Unlike relational databases which store data in tables, NoSQL databases store data within a single data structure, such as a JSON document. Because non-relational databases need no complex schema to fit the data into, the scaling for large and mostly unstructured data sets is pretty straightforward and fast (IBM, 2022).
- **The motive**
- Scalability:
    - Social media platforms require rapid and unforeseen growth in user base and data volume. Considering that NoSQL databases are horizontally scalable, they can easily distribute data and operations among several servers (nodes). This provides an easy avenue for scaling out to huge datasets and user traffic while avoiding performance degradation (IBM, 2022).
- Flexibility with Unstructured and Semi-structured Data:
    - Social media content such as posts, comments, or likes, may come in varying formats and structures, from user profiles to multimedia. NoSQL databases provide schema flexibility, allowing for diverse data types to be stored and managed without a strictly defined format. This flexibility is necessary for social media data, which can be ever-changing (IBM, 2022).
- High Performance and Low Latency:
    - Social media platforms require real-time interactivity, fast data retrieval systems for news feeds, notifications, and personalized content. NoSQL databases are best optimized for faster write and read operations, hence allowing the lightening-fast access to large data sets that keep users happy, even during periods of peak demand (IBM, 2022).
- **Types of data**
- User Data
  About people and entities on the platform:

- Usernames, names, and contact info (GeeksforGeeks, 2024).
- Demographics, interests, and bios (GeeksforGeeks, 2024).
- Login credentials, profile pictures (GeeksforGeeks, 2024).
- Content Data
  Actual content shared by users:
  - Looking at the text: posts and comments with images, videos, links (GeeksforGeeks, 2024).
  - Hashtags and mentions inside posts (GeeksforGeeks, 2024).
- Engagement Data
  Information about user activities related to content and, or other users:
  - Likes and reactions, sharing, retweeting (GeeksforGeeks, 2024).
  - Commenting, replying, and direct messaging (GeeksforGeeks, 2024).
- Network Data
  Data concerning the ties amongst the users:
  - Following, following, and friend-side connections (GeeksforGeeks, 2024).
  - Group membership (GeeksforGeeks, 2024).
- Usage and Behavioral Data
  Statistics and patterns of user activities:
  - Time on platform and feature use (GeeksforGeeks, 2025).
  - Device info and traffic sources (GeeksforGeeks, 2025).
  - User flow and site interactions (GeeksforGeeks, 2025).
- Transactional and Personal Data
  Data concerning commerce and characteristics of the user:
  - Purchase history and customer loyalty information (GeeksforGeeks, 2025).
  - Sensitive personal data on characteristics, political views, purchasing habits (GeeksforGeeks, 2025).
- Qualitative Data
  Text and interaction rendered by users implying sentiment and feedback:
  - Customer services information and feedback (GeeksforGeeks, 2025).
  - General user (GeeksforGeeks, 2025).

- **Type of No-SQL database:**

- Graph Databases
    - What they are: They represent the data as nodes and edges, zeroing in on relationships between data points (Cuello, 2025).
    - Social Media Use Case: -- Great for social features, such as friend lists, recommendations, or seeing how users are connected to one another (as in Instagram or TikTok) (Cuello, 2025).
    - Examples: Neo4j, Amazon Neptune (Cuello, 2025).

- Document Databases
    - What they are: They're a match for semi-structured data as they store information in supple, JSON-like documents (Cuello, 2025).
    - Use case for Social Media: Great for user profiles, posts, and other content with a dynamic structure and fields (like in a CMS or an online shop) (Cuello, 2025).
    - Examples: MongoDB, Couchbase (Cuello, 2025).

- Key-Value Stores
    - What they do: Basic databases that store data in the form of unique keys paired with values, allowing for fast read/write access to the data (MongoDB, 2025).
    - Social media Use case: This is the best use case for very high speed, simple data access (eg use preferences, real time feeds or session info) (MongoDB, 2025).
    - Examples: Redis, DynamoDB (MongoDB, 2025).
- Column-Oriented Databases (Wide-Column Stores)
    - What they are: Data is kept in rows but set up in columns or column families, making it possible to store and retrieve large amounts of data with relative ease (MongoDB, 2025).
    - Social media use case: Great for ingesting large volumes of social data (posts, messages, etc.) and big data analytics (MongoDB, 2025).
    - Examples: Apache Cassandra, employed by platforms such as Instagram (MongoDB, 2025).
- **The Three Vs of Big Data in the Scenario:**
    1- Volume
        - The system handles massive amounts of data generated daily by millions of users (OpenAI, 2025).
        - This includes posts, images, and videos, and even VR content, likes, comments, and analytics (OpenAI, 2025).

- It needs a horizontally scalable NoSQL solution to store and manage this amount efficiently (OpenAI, 2025).

2- Velocity
- The information is generated in real time as the users engage on the platform (OpenAI, 2025).
- Likes, shares, comments, and live streams need real-time processing to keep feeds current (OpenAI, 2025).
- NoSQL databases enable high-speed data ingestion and real-time analysis (OpenAI, 2025).

3- Variety
- The platform is holding structured, semi-structured, and unstructured data in text, image, video, GIF, and VR streaming formats (OpenAI, 2025).
- NoSQL databases are ideal for handling this mixed collection of data formats without rigid schema restrictions (OpenAI, 2025).

Q2

Customer	
customerID	PK
name	
email	
addressLine1	
addressLine1	
suburb	
city	

Order	
orderID	PK
customerID	FK
orderStatus	
totalPrice	

Payment	
paymentID	PK
orderID	FK
paymentdate	
paymentAmount	
paymentMethod	

place 1..*

associated 1..1

have 1..1

ordered 1

appear 1..*

written 1

write

contain 1..*

written 1

have 1..*

Book	
bookID	PK
author	
title	
genre	
price	
quantityInStock	

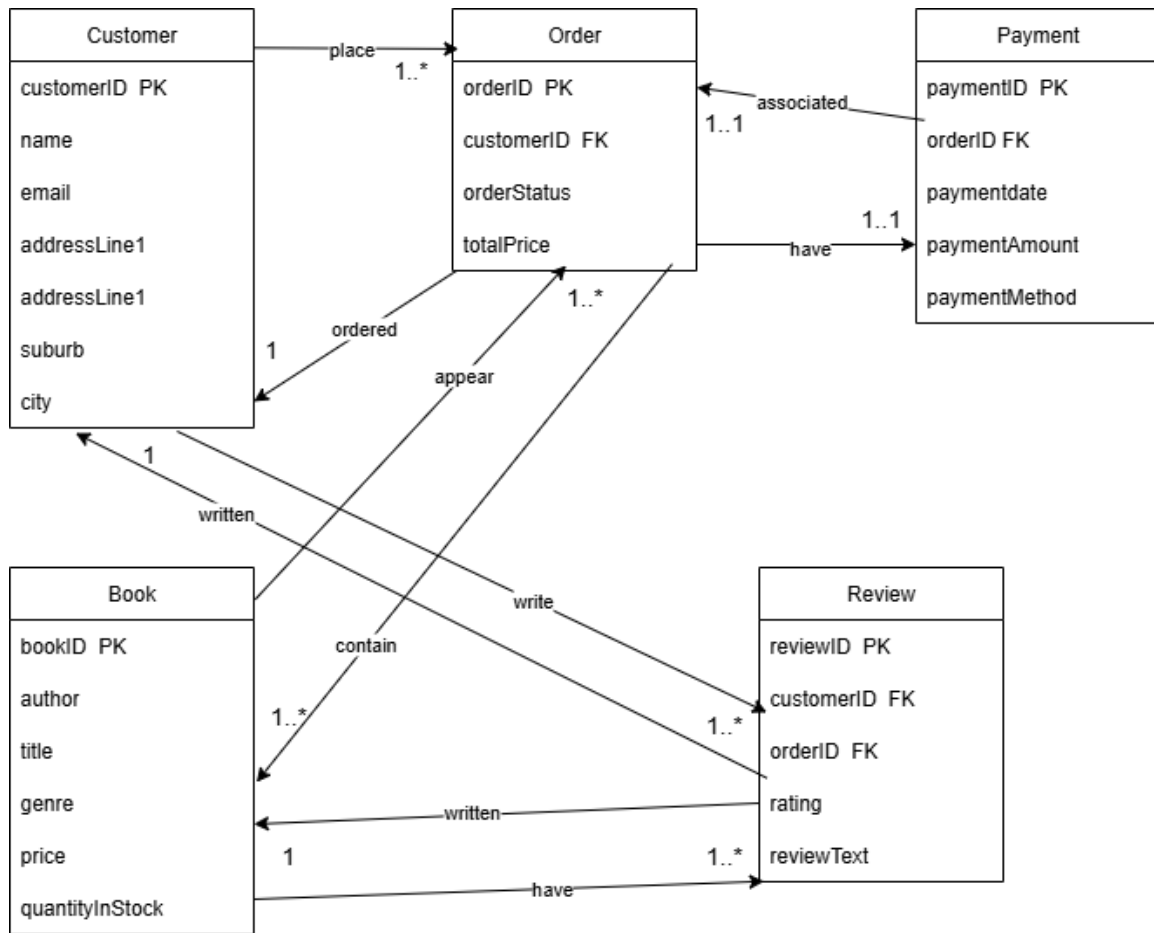| Review	
reviewID	PK
customerID	FK
orderID	FK
rating	
reviewText	

Figure 1 draw.io

References

Cuello, C., 2025. *Understanding The Different Types Of Databases & When To Use Them*. [Online]
Available at: https://rivery.io/data-learning-center/database-types-guide/#:~:text=Instead%20of%20tables%2C%20NoSQL%20databases,commerce%2C%20and%20big%20data%20analytics.
[Accessed 26 08 2025].

GeeksforGeeks, 2024. *How to Design Database for Social Media Platform*. [Online]
Available at: https://www.geeksforgeeks.org/sql/how-to-design-database-for-social-media-platform/
[Accessed 25 08 2025].

GeeksforGeeks, 2025. *How to Use SQL for Social Media Data Analysis.*. [Online]
Available at: https://www.geeksforgeeks.org/sql/how-to-use-sql-for-social-media-data-analysis/
[Accessed 26 08 2025].

IBM, 2022. *What is a NoSQL database?*. [Online]
Available at: https://www.ibm.com/think/topics/nosql-databases#:~:text=NoSQL%2C%20also%20referred%20to%20as,applications%20in%20nearly%20every%20industry.
[Accessed 26 08 2025].

MongoDB, 2025. *What is NoSQL?*. [Online]
Available at: https://www.mongodb.com/resources/basics/databases/nosql-explained#:~:text=NoSQL%20databases%20come%20in%20a,data%20and%20high%20user%20loads.
[Accessed 26 08 2025].

OpenAI. 2025. Chat-GPT (version 5). [Large Language model]. Available at:
https://chatgpt.com/c/68aca7f8-6430-832e-a814-112d5e30af57 [Accessed: 28 August 2025