Querying Databases for Data Extraction

Part 1: Office Roleplay Dialogue

Scenario: A Data Analyst, Kenji, is discussing SQL **queries** and **data extraction** techniques with his colleague, Maria.

Maria: Hey Kenji, I need to pull customer transaction data for the last six months. What's the best way to do that?

Kenji: You can write a **query** to extract the relevant records. Do you need details from multiple tables?

Maria: Yes, I need purchase history and customer profiles. Should I use a **JOIN** to combine them?

Kenji: Exactly! A **JOIN** will link both tables using a common key, like the customer ID.

Maria: That makes sense. What if I only need specific details from a subcategory within the dataset?

Kenji: Then you can use a **subquery** inside your main **query** to filter the data before retrieving it.

Maria: Got it. Also, I need to summarize the total purchase amounts per customer. How do I do that?

Kenji: You can use an **aggregate** function like SUM() to calculate total sales for each customer.

Maria: That's helpful! What if I need to export the results for reporting?

Kenji: Once the **data extraction** is done, you can export it as a CSV file or connect it to a dashboard tool.

Maria: Perfect! I'll write the **query** and test it now. Thanks for your help!

Kenji: No problem! Let me know if you run into any issues.

Part 2: Comprehension Questions

1. What does Kenji suggest Maria use to retrieve data from multiple tables?
(A) Delete
(B) Backup
(C) Format
(D) JOIN

2. What is a subquery used for?
(A) Creating a new database
(B) Filtering data within a larger query
(C) Replacing customer information
(D) Updating all records

3. Which function does Kenji recommend for summarizing purchase amounts?
(A) Reset function
(B) Format function
(C) Aggregate function (SUM)
(D) Delete function

4. How can Maria export the extracted data?
(A) As a CSV file or connect it to a dashboard
(B) By printing it on paper

(C) By manually writing each value in a report
(D) By using a calculator

Part 3: Key Vocabulary Definitions in Japanese

1. **Query (クエリ)** – データベースに対して情報を検索、取得するための命令文。
2. **Join (結合)** – 2 つ以上のテーブルを関連付けてデータを組み合わせる SQL 操作。
3. **Subquery (サブクエリ)** – 別のクエリ内に埋め込まれたクエリで、データのフィルタリングに使用される。
4. **Aggregate (集計関数)** – SUM() や AVG() などの関数を使ってデータの集計を行う手法。
5. **Data Extraction (データ抽出)** – 必要なデータをデータベースから取得し、分析やレポートに活用するプロセス。

Part 4: Questions & Correct Answers

1. **What does Kenji suggest Maria use to retrieve data from multiple tables?**
   ✅ (D) JOIN
2. **What is a subquery used for?**
   ✅ (B) Filtering data within a larger query

3. **Which function does Kenji recommend for summarizing purchase amounts?**

   ✅ (C) Aggregate function (SUM)

4. **How can Maria export the extracted data?**

   ✅ (A) As a CSV file or connect it to a dashboard