

【解答】

〔設問1〕 エ、カ

〔設問2〕 aーア

〔設問3〕 bーア

〔設問4〕 イ

【解説】

関係データベースにおける SQL 文とインデックスの設定に関する問題である。題材は所蔵する書籍とその貸出しをデータベースで管理するというものであり、図書館での書籍の貸出しやレンタルビデオ店でのビデオ／DVD の貸出しなどで比較的イメージしやすい問題になっている。

設問1は、従来のテーブル構造では検索することができない情報を選択する問題で、与えられたテーブルの説明を一つずつ読み進めていけば解答できる。

設問2と設問3は SQL 文の穴埋め問題となっている。特に設問3は出題頻度が高いグループ化による集合関数の SQL 文である。設問2、設問3ともに与えられたテーブルの構成と説明をよく理解して問題に取り組みば解答できる。

設問4は、レコード件数・抽出条件・結合条件から最も効果が期待できるインデックスの設定を検討する問題となっている。

学習に当たっては、問題文に沿ってデータベースにテーブルを定義し、データを登録した状態で、実際に SQL を実行し結果を確認してほしい。

〔設問1〕

従来の書籍表（図1）及び社員表（問題文にテーブル構造は示されていない）を使って検索できない情報を選択する問題である。社員表は、具体的なテーブル構造が示されていないが、〔書籍表に関する説明〕から、少なくとも図Aのような構造であると推測できる。

社員番号	氏名	連絡先
110028	佐藤太郎	東京都台東区〇〇〇1-23
...

図A 社員表のデータ格納例

また、書籍表は所蔵する書籍と現在の貸出状況をまとめて管理しており、貸出時と返却時に貸出日、返却予定日、社員番号を更新していることが分かる。

解答群を一つずつ確認し SQL を考えてみると、

ア：現在、貸出中でない書籍の一覧
貸出中でない：貸出日 IS NULL という選択条件で検索することができる。
SELECT 書籍表.書籍名 FROM 書籍表 WHERE 書籍表.貸出日 IS NULL

イ：現在、貸出中の書籍の一覧
貸出中：貸出日 IS NOT NULL という選択条件で検索することができる。
SELECT 書籍表.書籍名 FROM 書籍表 WHERE 書籍表.貸出日 IS NOT NULL

ウ：現在、書籍を貸出中の利用者の連絡先
イの条件に加え、書籍表と社員表を社員番号で結合することで検索することができる。
SELECT 社員表.連絡先 FROM 書籍表, 社員表
WHERE 書籍表.貸出日 IS NOT NULL AND 書籍表.社員番号 = 社員表.社員番号
又は、貸出中以外の書籍は、書籍表の社員番号も null なので、
SELECT 社員表.連絡先 FROM 書籍表, 社員表
WHERE 書籍表.社員番号 = 社員表.社員番号

エ：書籍ごとの累積貸出回数
貸出しの履歴が記録されていないので検索することができない。

オ：返却予定日を過ぎても貸出中の書籍の一覧
イの条件に加え、返却予定日が現在の日付よりも前という条件で検索することができる。
SELECT 書籍表.書籍名 FROM 書籍表
WHERE 書籍表.貸出日 IS NOT NULL AND 書籍表.返却予定日 < CURRENT_DATE
又は、貸出中以外の書籍は、書籍表の返却予定日も null なので、
SELECT 書籍表.書籍名 FROM 書籍表
WHERE 書籍表.返却予定日 < CURRENT_DATE

カ：利用者ごとの貸出履歴
エと同様で、貸出しの履歴が記録されていないので検索することができない。
以上のことから、(エ)と(カ)が検索できない情報となる。

〔設問2〕

設問1の要望に対応するために、従来の書籍表を、所蔵する書籍の情報と貸出履歴の情報の二つにテーブルを分割した構造で再構築している。問題文の SQL 文は「運用開始後に延滞したことがある利用者の社員番号と、延滞した書籍名を、社員番号の昇順に表示」するものである。

空欄となっているのは WHERE 句の部分で、結合条件と抽出条件で適切なものを選択する。図2の書籍表と貸出表の結合キーはそれぞれのテーブルの書籍番号であり、
書籍表.書籍番号 = 貸出表.書籍番号
の条件はどの解答群にも記述があるため、結合条件では解答の絞り込みはできない。抽出条件は「運用開始後に延滞したことがある利用者」であるが、これを満たすには二つのパターンを考慮する必要がある。

- ① まだ返却されていない書籍の延滞
返却日が null で返却予定日を過ぎている（現在の日付よりも前）つまり、
貸出表.返却日 IS NULL AND 貸出表.返却予定日 < CURRENT_DATE
- ② 既に返却されている書籍の延滞
返却予定日より後に返却された。つまり、
貸出表.返却日 > 貸出表.返却予定日
- ①と②はいずれかを満たせばよいので、これらを組み合わせた（ア）が正解である。

イ：①の条件と合致しているが、貸出日から返却予定日の間に返却されたもの（延滞していないもの）を合わせて抽出する条件となっているので、誤りである。

ウ：返却日と返却予定日が異なるものを抽出しており、返却予定日より前に返却したもの（延滞していないもの）を含む抽出条件となっているので、誤りである。

エ：返却日が入っているものを抽出しており、延滞の有無にかかわらず返却されたものを抽出する条件となっているので、誤りである。

〔設問3〕

書籍番号ごとに貸出回数を集計し、貸出回数の降順に表示する SQL 文である。WHERE 句や GROUP BY 句、ORDER BY 句は既に記載があり、SELECT 句の貸出回数の部分が空欄となっている。

貸出表は貸出しごとに1レコード挿入しており、貸出回数は貸出表のレコード数を数えればよいので、（ア）の COUNT(*)が正解である。

イ：MAX(書籍表.書籍番号)はグループの中で最も大きい書籍番号を表しており、誤りである。

ウ：SUM(貸出表.貸出番号)はグループの貸出番号の合計を表しており、誤りである。

エ：貸出表.貸出番号は貸出回数を表していないため、誤りである。

オ：書籍表.書籍番号は SELECT 句の一つ目の項目と同じで書籍番号を表しており、誤りである。

〔設問4〕

レスポンス低下に伴うインデックスの検討に関する問題である。一般的にインデックスはたくさんのレコードの中から少数のレコードを抽出する場合やデータのソートをする場合に用いられ、結合条件や抽出条件、ソートの項目に対して設定する。

設問では、書籍名を入力して貸出中か否かを表示する処理でレスポンス低下が顕在化しており、レコード件数がそれぞれ書籍表 865 件、貸出表 10,382 件である点を考慮しなければならない。

書籍名の入力については問題文に明記されていないが、書籍名の一部、例えば「技術」という言葉が入る書籍を検索することが考えられるので、抽出条件は、
書籍表.書籍名 LIKE '%技術%'
のように部分一致のあいまい検索を想定し、貸出中か否かを判断するために、書籍表と貸出表を結合する必要があるため、結合条件は、
書籍表.書籍番号 = 貸出表.書籍番号
となる。これらの条件に登場する項目に対しインデックスの設定の有効性を考えると、
書籍表.書籍番号：既に書籍表の主キーとなっているため、インデックスの設定の必要はない。
貸出表.書籍番号：貸出表のレコード件数も多く、インデックスの設定が有効である。

書籍表.書籍名：LIKE を使った部分一致のあいまい検索ではインデックスが有効ではない。

書籍表.書籍名 LIKE '技術%'
のように前方一致のあいまい検索を想定した場合や、
書籍表.書籍名 = '情報処理技術の歴史'
のように完全一致の検索を想定した場合は、インデックスは有効に機能する。

ここで、貸出表.書籍番号と書籍表.書籍名のどちらにインデックスを設定するのが有効であるのかを考える。インデックスを利用した場合の検索処理では2分探索やハッシュ探索が考えられるが、まずは2分探索で考えてみる。N 件のデータに対する2分探索での、平均探索回数は $\log_2 N - 1$ 回となる。一方、インデックスを利用しない場合は、線形探索になるため、平均探索回数は $N \div 2$ 回となる（ただし、探索以外の計算量を除く）。

書籍表.書籍名にインデックスを設定し書籍名の検索に有効に機能するとした場合、

書籍表の平均探索回数は $\log_2 865 - 1 \div 9$ 回
貸出表の平均探索回数は $10,382 \div 2$ 回
貸出表.書籍番号にインデックスを設定した場合、
書籍表の平均探索回数は $865 \div 2$ 回
貸出表の平均探索回数は $\log_2 10,382 - 1 \div 12$ 回
となり、明らかに貸出表.書籍番号にインデックスを設定した方が検索の計算量が少なくなる。また、ハッシュ探索で考えた場合、書籍名のあいまい検索ができないほか、インデックスを適用した表の探索は、原則1回になるので、この場合も貸出表にインデックスを設定した方が有利である。

したがって、(イ)が正解である。

ア：書籍表.書籍名は LIKE を使ったあいまい検索となるため、有効ではない。

ウ：貸出表.社員番号は抽出条件や結合条件に利用されないため、有効ではない。

エ：貸出表.貸出日は抽出条件や結合条件に利用されないため、有効ではない。

オ：貸出表.返却日は貸出中か否かを判断する際に使用するが、貸出表.返却日に null となるレコードが存在するか否かで判断することになる。null 値にはインデックスが作成されないなので、有効ではない。