

問 2 購買情報を管理する関係データベースの設計及び運用（データベース）（H24 秋・FE 午後問 2）

【解答】

- 〔設問 1〕 aーア, bーイ, cーウ (a, b は順不同)
- 〔設問 2〕 dーア
- 〔設問 3〕 eーア, fーカ, gーウ (e, f は順不同)
- 〔設問 4〕 hーイ

【解説】

関係データベースの設計と SQL 文に関する問題である。設問 1 は、データベースを統合する際の名寄せ作業についての問題で、説明文と図の内容から把握すれば解答できる。設問 3 も、求めたい情報が格納されている表の項目を把握すれば解答できる。

設問 2 と設問 4 の SQL 文は、出題頻度が高いグループ化の SQL 文である。文法を理解して問題を解くことで、実力を向上させたい。設問 2 は基本的な SQL 文であるが、設問 4 は応用力が要求される。

〔設問 1〕

名寄せの代表的な例は、金融機関に複数の口座をもっている一人の顧客のデータを一つにまとめる作業である。この問題では、同じ取引先が統合前の X 事業部と Y 事業部のデータベース両方に存在する場合、統合後のデータベースで一つにまとめる名寄せを実施する。

- ・空欄 a, b：名寄せを実施せず、同じ情報を表すデータが、統合されたデータベースに、複数格納された場合に発生する問題を答える。例えば、統合された取引先表に、同じ取引先が二つ格納された場合を考える。ここで、取引先の引越しによって所在地と電話番号を更新する際、二つのデータを両方とも更新するが、一つしか更新しないことが起こり得る。また、取引停止に伴い、取引先のデータを削除する場合も同じく、一つしか削除しないことが起こり得る。したがって、(ア) と (イ) が正解である。その他の選択肢は、名寄せを実施しないこととは関係しない。
- ・空欄 c：取引先の名寄せを実施する際、突合せを行う適切な項目を答える。突合せを行うためには、その取引先を一意に識別できる項目が望ましい。取引先表の項目の中では、表記法が統一された電話番号が適切である。図 2、図 3 の例から、名寄せの対象となる会社の電話番号が一致していることで分かる。したがって、(ウ) が正解である。(エ) の取引先コードは、桁数を含めコードの付番方法が異なるので、不適切である。また、(ア) の社名と (イ) の所在地は、同じ内容でも表記法が異なることがあるため、不適切である。例えば、図 2、図 3 の社名の情報商事株式会社と情報商事 (株) は、社名の表記法が異なるが同じ会社である。

〔設問 2〕

“エコ鉛筆黒”の平均購入単価を、取引先ごとに求めるグループ化の SQL 文である。問題の SQL 文は、設問文に「商品名“エコ鉛筆黒”について、取引先コードと平均購入単価を、金額が安い順に表示する」とあり、そのため、**GROUP BY** 句と集合関数の **SUM** を使って記述されている。**WHERE** 句では、グループ化前に抽出すべき条件を発注表、明細表と商品表を結合した上で適用するので、(ア) の記述 3 行の上 2 行は、発注表、明細表、商品表の結合条件を、3 行目は、“エコ鉛筆黒”の抽出条件を示している。したがって、(ア) が正解である。

参考までに、グループ化の SQL 文の構文は、次のようになる。

SELECT 句の指定した列名とは、**GROUP BY** 句で指定した列名という意味である。言い方を変えると、グループごとの値ということができる。もちろん、集合関数もグループごとの値である。

```
SELECT 指定した列名, 集合関数
FROM 表名
WHERE 結合条件, 抽出条件
GROUP BY 指定した列名
```

イ：副問合せによって、“エコ鉛筆黒”の定価の平均が抽出される。ここで、「商品名は一意に管理できている」という記述から、“エコ鉛筆黒”のデータは、商品表に 1 件しかないことが分かる。そのため、副問合せの結果は、商品表の“エコ鉛筆黒”の定価になる。これと明細表の購入額が等しいことが抽出条件となっており、不適切である。

ウ：副問合せによって、“エコ鉛筆黒”の定価の合計が抽出されるが、“エコ鉛筆黒”のデータは 1 件しかないため、商品表の“エコ鉛筆黒”の定価になる。(イ) と同じ結果となり、不適切である。

エ：副問合せによって、商品名が“エコ鉛筆黒”の行数 (件数) が抽出される。“エコ鉛筆黒”のデータは 1 件しかないため、1 になる。この件数の 1 と明細表の商品コードが等しいことが抽出条件となっており、不適切である。

<集合関数についての補足>

集合関数は、次の五つである。

SUM (項目名)：指定した項目の値の合計

AVG (項目名)：指定した項目の値の平均

MAX (項目名)：指定した項目の値の最大

MIN (項目名)：指定した項目の値の最小

COUNT (*) 又は **COUNT** (項目名)：データの行数 (件数)

<副問合せについての補足>

副問合せとは、**WHERE** 句中に記述した **SELECT** 文である。一方、元の **SELECT** 文を主問合せという。このように、**SELECT** 文が入れ子構造になっている。ここでは、副問合せの **SELECT** 文を先に単独処理した後、副問合せで抽出された値を使って、主問合せの **SELECT** 文を処理する。

〔設問 3〕

商品コードごとの平均割引率を求めるために必要な表と、それを取引先コードごとに集計したり、時期による違いを把握したりするために必要な表を答える設問である。

- ・空欄 e, f：商品コードごとの発注時の購入単価は、明細表の購入額を数量で割れば求められる。一方、商品コードごとの定価は商品表に格納されている。そして、商品コードごとの平均割引率は、発注時の購入単価の平均を定価で割れば求められる。したがって、(ア) と (カ) が正解である。なお、商品表と明細表は、両方の表に格納された主キーと外部キーの関係にある商品コードで結合する。
- ・空欄 g：取引先コードは、発注表に格納されている。また、時期を示す発注日と納品日も、発注表に格納されている。そのため、商品表と明細表を結合した表に対して、更に発注表を発注コードで結合すればよい。したがって、(ウ) が正解である。

〔設問 4〕

文具の発注実績が 10 回以上ある取引先と、その購入金額を求めるグループ化の SQL 文である。発注実績は、発注回数のことなので、集合関数の **COUNT** を指定し、発注データの行数 (件数) を求める。なお、**HAVING** 句では、グループ化した後の抽出条件をグループごとの値を使って指定するため、通常、集合関数を記述する。

ここで、**WHERE** 句の抽出条件として、品目の文具を指定していることに注目する。なぜなら、1 回の発注に対して、品目が文具である商品が、複数含まれることが考えられるからである。この場合、発注回数は 1 にする必要がある。そのため、発注コードが同じ行が複数存在するときは、**DISTINCT** を指定して、行の重複を排除する。したがって、(イ) が正解である。なお、**COUNT (*)** と **COUNT** (発注表.発注コード) は、発注コードが同じ行が複数存在しても重複を排除せず、どちらも同じ結果になる。

参考までに、この発注表の発注コードは **NULL** になることはないが、**COUNT** で指定した列が **NULL** になる場合は、**NULL** の列は行数に含まれないので、**COUNT** (列名) と **COUNT (*)** は行数が違うことがある。これが、表の全行数を知るには、通常 **COUNT (*)** を用いる理由である。