



データベースの試験問題を解いてみる

♡ 9



情報処理学会・学会誌「情報処理」

2022年2月7日 11:23





長瀬寛之（大阪電気通信大学）

情報Iでは「情報通信ネットワークとデータの利用」の中で、「データを蓄積，管理，提供する方法」としてのデータベースについての理解が求められています。また学習指導要領解説や教員研修用資料などで，データベースによるデータ管理の手法として関係データモデル，それを実現するリレーショナルデータベース（以下RDB）が取り上げられています。

さて，本稿ではデータベースを扱った問題として，第1回大学情報入試全国模擬試験 #002 第3問 を取り上げてみます。RDBがテーマですが，「データの蓄積や検索が容易となるような，適切なデータベース設計とはなにか」の理解が，問題を解くポイントになっています。

第3問 (必答問題)

次の文章を読んで、下の問い（問1～3）に答えよ。

ケーキ類を製造・販売する店で、材料やコストの管理のためのシステムを作ろうとしている。次に示すのはそのシステムで使うテーブル（表）の一部である。

テーブル名: 商品テーブル

商品番号	商品名	ピース数	販売価格
101	いちごケーキ	6	2400
102	アップルパイ	6	1800
103	チーズケーキ	8	1600
104	ロールケーキ	8	1500
105	ロールケーキ	4	800

テーブル名: 材料単価テーブル

材料番号	材料名	単位	単価
0001	薄力粉	グラム	0.2
0002	砂糖	グラム	0.1
0003	クリーム	グラム	1
0004	いちご	個	50
0005	いちご	個	30
0006	卵	個	40
0007	卵	個	20

最初の文から、ここでは「ケーキ類を製造・販売する店」が舞台で、データの活用目的は「材料やコスト管理」のようです。

また、すでに"商品テーブル"と"材料単価テーブル"が用意されていて、具体的なデータも示されています。ここから、商品テーブルは取り扱う商品の価格情報、材料単価テーブルはこの店で扱う材料の単価情報を管理していることが読み取れます。

▼ 目次

問1

問2

問3

問1

問1 上のテーブル設計では、商品ごとに「商品番号」、材料ごとに「材料番号」を割り当てて使っている。このような設計を行なう理由として適しているものを、次の解答群から二つ選べ。

解答群

- ア. 商品名や材料名が同じでも、異なる商品や材料を区別する必要があるため。
- イ. コンピュータでは文字どうしの比較が行えないため。
- ウ. 商品名や材料名が変更されても、同一のものであることが識別できるようにするため。
- エ. データベース管理システムでは整数値しか主キーにできないため。
- オ. 現在、何種類の商品や材料があるかを数えるのが容易になるから。

商品テーブルと材料単価テーブルにある「番号」は何のために入っているのか？という問題です。

データベースでは、各テーブルの一組のデータ=レコードに通し番号のような"重複しないID"を振り、それをレコードを一意に識別する**主キー**として扱う設計がよく見られます。この主キーがあることで、まず「異なるレコードの区別が行える」という利点があります。実際のデータでは、"同姓同名"のように、重複する情報を"同一"とみなしたい場合と、"異なる"レコードとして扱いたい場合があります。主キーを使うことで、この異なるレコードの区分けを明確にすることができます。今回の問題でも、同じ"いちご"に異なる材料番号を割り振ることで、データ上は明確に別の材料（恐らく別品種のいちご）として扱うことができます。

ほかに主キーを利用する利点は、情報の変更による影響を少なくできることで

す。今回の問題でも、ロールケーキの8ピースを10ピースに変更する、材料番号0005のいちごが50円に値上がりするなど、利用者側の都合で変更され得るデータが存在します。そういったデータをテーブル外から利用するときに、「主キーをもとに、付属データを参照する」という仕組みにしておくことで、データの変更が発生しても、それを参照・利用する側には「主キーから参照する」という仕組みを一切変更せず、データ変更を反映させることが可能となります。

さて、問1の選択肢を見ると、**ア.**と**ウ.**の説明が、通し番号を振る一般的な目的と一致しますので、これが正解として妥当と言えます。

ちなみに、**イ.**は明らかにあり得ない（文字同士の『等しい』『等しくない』の比較が可能なのは、文字検索の仕組みからも明らか）。**エ.**についてもそんな決まりはない（数字でも文字でも関係なく、主キーは"重複がないこと"が重要）。また**オ.**は、個数をカウントしたければレコード数を直接数えれば済む話。となると、消去法でも**ア.**と**ウ.**しか残りません。

問2

問2 特定商品の材料費の合計を求めるには新たなテーブルが必要となる。そのテーブル名を「材料テーブル」とするとき、「材料テーブル」に必要最低限含めるべき項目名を、次の解答群からすべて選べ。

解答群

ア. 商品番号	イ. 商品名	ウ. ピース数
エ. 販売価格	オ. 材料番号	カ. 材料名
キ. 単位	ク. 単価	ケ. 分量
コ. 合計金額		

新たに材料テーブルを設計するために、そこに必要な項目=フィールドを求めよ、という問題です。ここでは「特定商品の材料費を求める」ことが目的のテーブルなので、商品・材料単価のテーブルに今回「材料」テーブルで追加する項目によって、特定商品の材料費を求められるようにする必要があります。

ちなみにこの問題、「商品は、1つ以上の一定分量の『材料』の組合せで構成される」が暗黙の前提事項とされていますが、本来は問題文中で明示的に定義した方がよいと思います。

"商品"単位の情報とは商品テーブルに、"材料"の単価はすでに材料単価テーブルに情報がありますが、商品の材料費を求めるための「その商品では、何の材料をどれ

だけ使うか」の情報が不足しています。つまり、材料テーブルに「商品」と、その商品に紐づく「使う材料」と「その材料の分量」の情報があれば、後は材料の単価をかけることで材料費は算出できそうです。

そうすると、解答群の選択肢から選ぶものは自ずと絞られてきます。

まず商品テーブルと材料単価テーブルにある各項目は、主キーとなる「商品番号」「材料番号」があれば、付属情報として引っ張ってこれます。つまり材料テーブルには最低限「ア. 商品番号」と「オ. 材料番号」があればよく、商品番号に付随する「イ」「ウ」「エ」、材料番号に付随する「カ」「キ」「ク」は、いずれも材料テーブル内で保持する必要がありません。

残るは「ケ. 分量」と「コ. 合計金額」です。各材料の「分量」は商品を構成する要素で、かつほかのテーブルでは存在しないので、材料テーブルで管理すべき情報と言えます。対して合計金額はどうでしょう。一見管理すべきデータにも思えますが、ここで問題文中の「必要最低限」という条件が効いてきます。合計金額は、使用材料の「分量」×「単価」で求められるので、必ずしもデータとしてあらかじめ保持しておく必要がありません。

というわけで、材料テーブルに必要な最小限の項目は、「**ア. 商品番号**」「**オ. 材料番号**」「**ケ. 分量**」となります。

ちなみに、問題としては問われていませんが、この項目を使って具体的にどういうテーブルを作ればよいのか、問3の手順とも合わせながらぜひ考えてみましょう。

また、今回は「必要最小限」ということで除外しましたが、そもそも材料の合計金額は「商品」単位の情報になるので、あえて材料費の合計金額を保存するなら、材料テーブルでなく商品テーブルのフィールドとするのが妥当でしょう。

問3

問3 特定商品の商品番号を用いて、その商品の材料費の合計を求めるにはどうしたらいいか。次の(1)～(5)に最も適切なテーブル名または項目名を、下の解答群から選べ。

第1段階 特定商品の(1)を使って、テーブル(2)から、その商品に使うすべての材料の(3)と、それぞれの材料の分量を求める。

第2段階 テーブル(4)から、それぞれの(3)に対する(5)を求める。

第3段階 それぞれの材料の分量と(5)を掛け、それらを足し合わせて合計金額を求める。

解答群

ア. 商品番号	イ. 商品名	ウ. ピース数
エ. 販売価格	オ. 材料番号	カ. 材料名
キ. 単位	ク. 単価	ケ. 分量
コ. 合計金額	サ. 商品テーブル	シ. 材料単価テーブル
ス. 材料テーブル		

最後に、商品番号からその商品の材料費を求める過程について、説明文の穴埋めを行って完成させる問題です。問2で材料テーブルに最低限必要なフィールドが何かが分かったので、あとは各テーブルのデータを適切な順番で組み合わせれば、目的の「商品の材料費」が求められるはずです。

第1段階では、特定の商品番号から、その商品に使われる「それぞれの材料の分量」を取り出したいことが分かります。商品と材料が関連付けられているのは「材

料テーブル」．そして材料テーブルには商品と材料の「番号」と、各材料の「分量」のみ格納されています．そこで、特定商品の「商品番号」をキーとして、材料レコード内の商品番号を含む全レコードを取り出し、その商品番号に紐づく「材料番号」と分量を取り出します．

(1) ア.商品番号 (2) ス.材料テーブル (3) オ.材料番号

第2段階は文のほとんどが空欄なので、その後の第3段階と合わせて考えます．第3段階では、各材料の分量と(5)を掛けて、それらを足し合わせて材料費を求める、とあります．ということは、材料費を求めるのに必要な情報は、材料の「分量」以外ではその材料の「単価」となります．

つまり第2段階は、使用材料の「単価」を求めるプロセスだと分かります．そして、単価情報が入っているのは「材料単価テーブル」．商品に使う材料は第1段階で材料番号を取り出しているのです、それを主キーとする各材料の単価を、材料単価テーブルから取り出せばよいことが分かります．

(4) セ：材料単価テーブル (5) ク：単価

そもそもデータベースは「どういう目的でデータを蓄積・活用するのか」によって、必要なデータやその管理方法が大きく変わってきます．そのため、問題の前提として「ここでは何を目指して、何のデータを、どんな目的で活用しようとしているか」が必ず示されているはずなので、そこをあらかじめ理解しておくことが必要です．それなしに直感で必要なデータの当たりをつける解き方では、安定して正確

な答えは得られません。

また今回の問題は、RDBの「主キー」や「正規化」や「関係演算」の概念を含んでいますが、実は直接その知識やスキルを問うていない点も重要です。ただテクニックを習得するのではなく、その技術と理論がどのように「適切なデータの蓄積・検索」を実現しているかを、実例も踏まえて理解しておくことが大事です。

(2022年1月24日受付)

(2022年2月7日note公開)

■長瀧寛之（正会員）

大阪電気通信大学 特定業務特任准教授。コンピュータと教育研究会幹事。主にコンピュータ活用教育、情報教育の研究に従事。データベース学習システム sAccess を開発したりもしている。

情報処理学会ジュニア会員へのお誘い

小中高校生、高専生本科～専攻科1年、大学学部1～3年生の皆さんは、情報処理学会に**無料で入会**できます。会員になると**有料記事の閲覧**、**情報処理を学べるさまざまなイベント**に**お得に参加できる等のメリットがあります**。ぜひ、入会をご検討ください。入会は[こちら](#)から！