

コンビニのレジには年齢・性別ボタンがあるらしい(2) -ExcelでDBっぽくデータを扱おう-

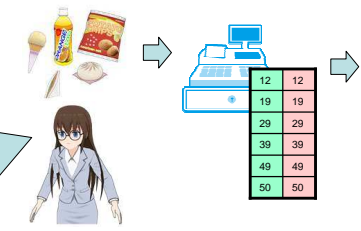
これがコンビニのレジのデータベースだって、何かすごいけど。

簡単なのか複雑なのか、わからないね。

さいきん、データベースという言葉をよく聞くと思います。今回の授業ではコンビニのデータベースをみていきましょう。

コンビニのレシートに入っている情報

コンビニのレシートに入っている情報をExcelの表にしてみました。データは例ですが、一枚のレシートでもいろいろな情報が入っていますね。念のためExcelの教材シート”A2元情報DB”で確認しておきましょう。



店舗番号	店舗住所	店舗電話番号	レシート番号	販売日	販売日時	レジ番号	担当者番号	商品番号	商品名	商品単価	商品購入数量	商品小計金額	商品小計金額	商品小計金額
1220301	市川市平田1丁目	047-322-0902	1001	5月19日	7:15	3	101	10101	アイス	130	1	130	130	2004
1220301	市川市平田1丁目	047-322-0902	1002	5月19日	7:15	1	102	10201	ジュース	238	1	238	238	4001
1220301	市川市平田1丁目	047-322-0902	1003	5月19日	8:00	3	101	10103	アイス	102	1	102	102	10101
1220301	市川市平田1丁目	047-322-0902	1004	5月20日	8:15	3	101	10202	ジュース	178	1	178	1023	
1220301	市川市平田1丁目	047-322-0902	1005	5月19日	7:22	1	101	10203	アイス	184	2	368	10201	
1220301	市川市平田1丁目	047-322-0902	1006	5月19日	7:45	1	101	99102	アイス	145	1	145		
1220301	市川市平田1丁目	047-322-0902	1007	5月19日	7:30	1	101	10203	アイス	114	1	114	10232	

店舗番号	商品名	商品単価	商品購入数量	商品小計金額	商品番号	商品名	商品単価	商品購入数量	商品小計金額	消費税	合計金額	合計金額	合計金額
41103	ポトパイ	118	1	118						374	29	403	403
20301	ジュース	102	1	102						547	43	590	600
41201	アイス	113	1	113						232	18	250	250
										407	32	439	500
										844	67	911	1000
										145	11	156	170
										298	23	321	321

シート: A2元情報DB

どうしてデータベースを使うの

- データベースの利点**
- データの検索や取り出しなどが簡単に速くできる。
 - 情報/データの重複が無く、また情報の誤りが少なくなる。
 - 情報/データの修正・変更などで楽である。
 - プログラムとデータを別々に考えることができる。
 - 複数の人で同時に使用できる。



データベースには一般的に上記のような特徴があります。少しずつ、その利点を確認していきましょう。そしてデータベースには、いろいろな方式がありますが、授業では現在広く使用されているリレーショナルデータベース(RDB: Relational DB)という方法を利用します。RDBは、右図のように、複数の表(テーブル)の情報を分けて、それと関係付けながらデータを管理、利用する方法です。例えば、店舗情報にFAX番号を追加したい場合は、店舗情報の表だけを修正すればいいことになります。また、新しい大きな区分の情報が追加された場合は、新しい表を追加することで対応できます。

元_店舗情報	元_販売情報
店舗名 店舗住所 店舗電話番号	レシート番号 販売日 販売日時 レジ番号 担当者番号 商品番号1 商品名1 商品単価1 商品購入数量1 商品小計金額1 商品番号2 商品名2 商品単価2 商品購入数量2 商品小計金額2 合計金額(税抜) 消費税 合計支払金額 残り金額
元_キャンペーン情報	
キャンペーン名 キャンペーン詳細	

大まかに分類して表(テーブル)に分けたデータ
シート:A3_単純分割 (これはRDBではなく、単なる三つの表)

主キーレコード

レコード:RDBの各テーブルの横の1行のデータをそれぞれ、レコードと読んでいます。
主キー:1つのテーブルの中の各レコードは必ず違うものを設定します(全く同じデータの複数のレコードは設定しない)。この時、各レコードを確実に区別するものを主キーと呼びます。一つの項目が主キーになる場合も、複数のキーが合わさって主キー(複合キー)になる場合があります。

RDBのルールや用語など見ていきましょう。レコードはExcelの1行のデータと同じと考えてください。また下記の図の店舗情報では店舗名でも各レコードの区別をすることができ、主キーとなることができ、通常主キーは検索や他のテーブルとのリンクに使われることが多く、処理の高速化などから、数値データをすることも多く、ここでは店舗番号のデータを追加しています。販売情報については、レシート番号は、毎日・店ごとに1番から始まることを想定して、そのため、店舗番号・レシート番号・販売日で、区別がつくようになります。



シート:NO キー有

店舗番号	店舗名	店舗住所	店舗電話番号
1220301	市川市平田1丁目	市川市平田1-18-14	047-322-6083
1220302	市川市大洲	市川市大洲3-10-13	047-379-5110
1220303	市川市八幡3丁目	市川市八幡3-3-2	047-324-1219

主キー

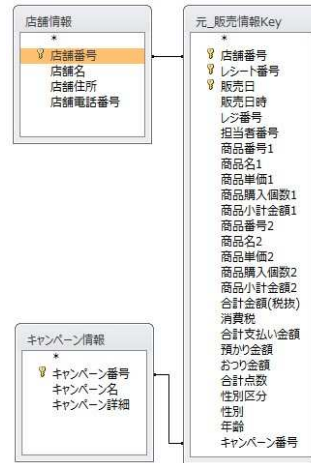
店舗番号	レシート番号	販売日	販売日時	レジ番号	担当者番号	合計金額(税抜)	消費税	合計支払
1220301	1001	5月19日	7:15	3	101	374	29	403
1220301	1002	5月19日	7:15	1	102	547	43	590
1220301	1003	5月19日	8:00	3	101	232	18	250
1220301	1004	5月20日	8:15	3	101	407	32	439
1220302	1001	5月19日	7:22	1	101	844	67	911
1220302	1002	5月19日	7:45	1	101	145	11	156
1220303	1002	5月19日	7:30	1	101	298	23	321

シート: NO キー有

主キー

テーブルの関連付け

では、いよいよバラバラだった表をくっつけましょう。基本的に二つのテーブルに同じ列のデータを入れて、この二つのテーブルを関連付けます(リンクします。参照します)
この例では、店舗情報と販売情報が店舗番号でリンクされます。この場合、ある店舗に複数のレシートがあることとなります。



店舗番号	店舗名	店舗住所	店舗電話番号
1220301	市川平田1丁目	市川市平田1-18-14	047-322-9093
1220302	市川大洲	市川市大洲9-10-13	047-379-5110
1220303	市川八幡3丁目	市川市八幡3-3-2	047-324-1219

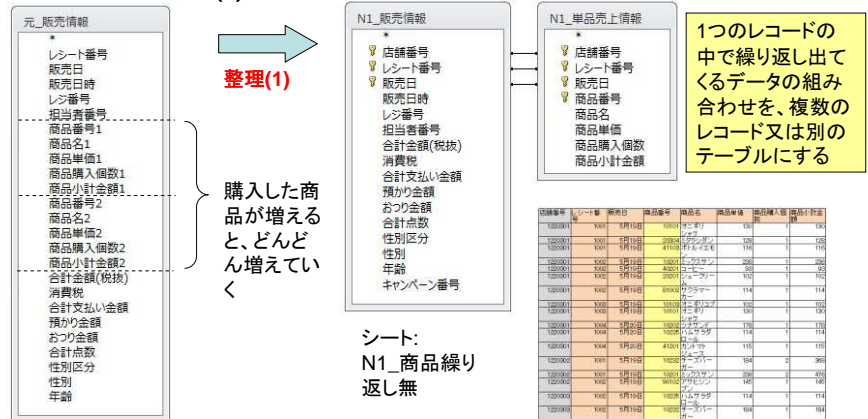
リンク: 参照

店舗番号	レシート番号	販売日	販売日時	レジ番号	担当者番号	合計金額(税抜)	消費税	合計支払い
1220301	1001	5月19日	7:15	3	101	374	29	
1220301	1002	5月19日	7:15	1	102	547	43	
1220301	1003	5月19日	8:00	3	101	232	18	
1220301	1004	5月20日	8:15	3	101	407	32	
1220302	1001	5月19日	7:22	1	101	844	67	
1220302	1002	5月19日	7:45	1	101	145	11	
1220303	1002	5月19日	7:30	1	101	298	23	

シート:
NO_キー有

5

テーブルを整理する(1)



シート:
N1_商品繰り返し無

RDBはできましたが、個々のテーブルを見ると、まだ改良の余地があります。ここでは販売情報を見直しましょう。このテーブルの一番の問題は、1つのレコードに商品の情報が複数個あります。そのため、この商品の領域をあらかじめいくつ用意しておけばいいかわかりません。100個同時に買う人がいるかもしれません。テーブルの整理で初めにやることは、このように同じ種類のデータが複数個レコードにある場合、それを分けてレコードを作ることです。ここでは、単品売上情報として別テーブルにして、1つの商品を1つのレコードにしています。



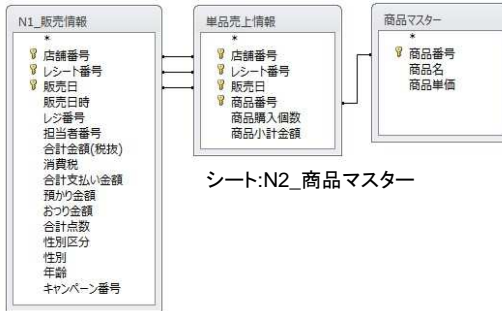
6

テーブルを、もっと整理する(2)

N1_単品売上情報

店舗番号	レシート番号	販売商品番号	商品名	商品単価	販売個数
1220301	1001	10101	オニギリシャケ	130	
1220301	1001	20304	ミックスサンド	128	
1220301	1001	41103	ロイヤルエビ	116	
1220301	1002	10201	ミックスサンド	238	
1220301	1002	40201	コービー	93	
1220301	1002	20201	シュークリーム	102	
1220301	1002	61002	カラマール	114	
1220301	1003	10103	オニギリ	102	
1220301	1003	10101	オニギリ	130	
1220301	1004	10202	ミックスサンド	178	
1220301	1004	10235	ラムサラ	114	
1220301	1004	41201	カスタードシュークリーム	115	
1220302	1001	10232	チーズシュークリーム	184	
1220302	1001	10201	ミックスサンド	238	
1220302	1002	90102	アサヒシソヨ	145	
1220303	1002	10235	ラムサラ	114	
1220303	1002	10232	チーズシュークリーム	184	

整理(2)



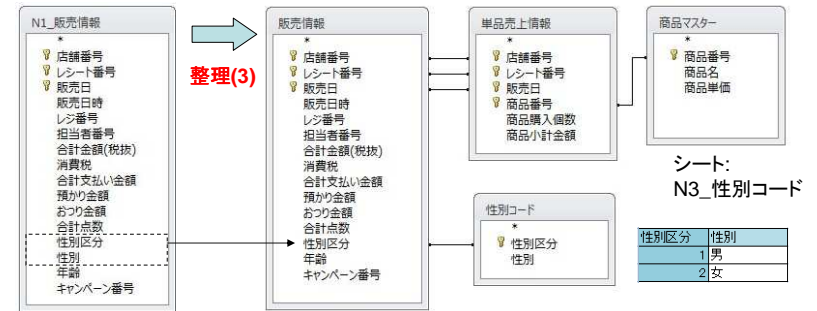
シート:N2_商品マスター

テーブルをもっと整理していきましょう。次の主キー(おもに複合主キー)に関連した情報に無駄がないか見ていきましょう。現在の単品売上情報には各レコードの中に商品名と商品単価が入っていますが、これは商品番号が決まるもので、テーブルのなかに同じ"オニギリシャケ"などのデータが複数入っています。また、これらは本来同じデータであるのに、別々のレコードに入っていて誤ってデータが設定されることも想定されます。そこで、ここでは、商品番号から決まる商品名/商品単価を別テーブルにします。



7

テーブルを、もっと、もっと整理する(3)



シート:
N3_性別コード

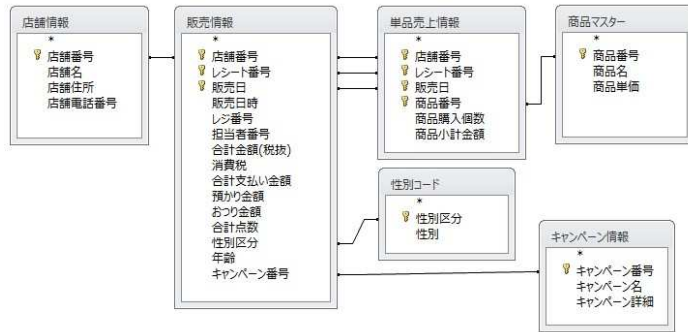
性別区分	性別
1	男
2	女

いよいよ、最後の整理です。テーブルの中に、まだ無駄が無いかわかりません。今回のテーブルでは性別については男女の文字だけでなく、1と2という数値で表しています(これは、検索や集計などの時に数値の方が処理しやすいという点からです)。この数値と男女の文字がダブっているため、性別コードのテーブルを作り、販売情報のテーブルには性別区分の数値だけを設定するようにしました。



8

完成したコンビニレシート用のRDB



店舗情報とキャンペーン情報のテーブルも元に戻して、これでコンビニレシート用にRDBの完成です。次のこれを使って、いろいろデータの処理をする方法が見ていきましょう。



コンビニレシート用のRDBの利用

コンビニのレシートのデータの使い方は次のようなものだけ

- 1.各店舗での利用
 - ・日別/時間帯別売り上げ
 - ・売れ筋商品
 - ・時間帯別売れ筋商品
 - ・客層/来客時間帯
 - ・仕入発注
 - ・返品計画
- 2.本社/本店での利用
 - ・売れ筋商品
 - ・客層/客層別売れ筋
 - ・店舗別売上
 - ・店舗別売れ筋商品
 - ・商品組み合わせ
 - ・時期別売れ筋商品
 - ・生産計画
- 3.各店舗と本社/本店とのやりとりでの利用
 - ・仕入情報
 - ・配送情報
 - ・返品情報
 - ・売れ筋商品(お勧め)

コンビニのレシートの情報って何に使うんだっけ



これからExcelを使ってコンビニレシートRDBを操作して、必要な情報をとりだしましょう。



操作: 射影=特定の列を取り出す

販売情報

- 店舗番号
- レシート番号
- 販売日
- 販売日時
- レジ番号
- 担当者番号
- 合計金額(税抜)
- 消費税
- 合計支払い金額
- 預かり金額
- おつり金額
- 合計点数
- 性別区分
- 年齢
- キャンペーン番号

射影

1#射影#販売情報

- 店舗番号
- レシート番号
- 販売日
- 合計支払い金額
- 合計点数

店舗番号	レシート番号	販売日	合計支払い金額	合計点数
1220901	1001	5月19日	408	3
1220901	1002	5月19日	590	4
1220901	1003	5月19日	250	2
1220901	1004	5月20日	439	3
1220902	1001	5月19日	911	2
1220902	1002	5月19日	156	1
1220903	1002	5月19日	321	2

シート:
1.射影.販売情報

射影はテーブルから特定の列だけを取り出す操作になります。

データベースに対する操作には、「選択」「射影」「結合」の基本的なものがあります。例えば、販売情報にはいろいろなデータが入っていますが、売上げだけに注目/計算したい場合は、支払い金額や合計点数だけを抽出してから処理すると、判りやすくなりますね。



Excelでの操作: 射影

D	E	F	G	H
販売日時	レジ番号	担当者番号	合計金額(税抜)	消費税
7:15	3	101	374	
7:15	1	102	547	
8:00	3	101	232	
8:15	3	101	407	
7:22	1	101	844	
7:45	1	101	145	
7:30	1	101	298	

列を削除する場合は、列を指定した後右クリックしてから削除を選択します。

Excelで射影の操作をすることは比較的簡単です。もとの表から不必要な列を削除するか、別のシートに必要な列をコピーするだけです。不要な列を削除する場合は、もとのシートを残すため、あらかじめシート全体を別のシートにコピーして、そのコピーしたシートから削除するようにしましょう。

シート全体をコピーする方法は
a. Ctrl + A で全体を選択してからコピー
b. 表の左上をクリックして全体を選択してからコピー



操作: **選択**=特定の条件に合うレコードを取り出す

店舗番号	レシート番号	販売日	合計支払い金額	合計点数
1220301	1001	5月19日	403	3
1220301	1002	5月19日	590	4
1220301	1003	5月19日	250	2
1220301	1004	5月20日	439	3
1220302	1001	5月19日	911	2
1220302	1002	5月19日	156	1
1220303	1002	5月19日	321	2



店舗番号	レシート番号	販売日	合計支払い金額	合計点数
1220301	1001	5月19日	403	3
1220301	1002	5月19日	590	4
1220301	1003	5月19日	250	2
1220301	1004	5月20日	439	3

選択

1220301を抽出
シート:
2.選択.販売情報S1



店舗番号	レシート番号	販売日	合計支払い金額	合計点数
1220301	1001	5月19日	403	3
1220301	1002	5月19日	590	4
1220302	1001	5月19日	911	2

選択

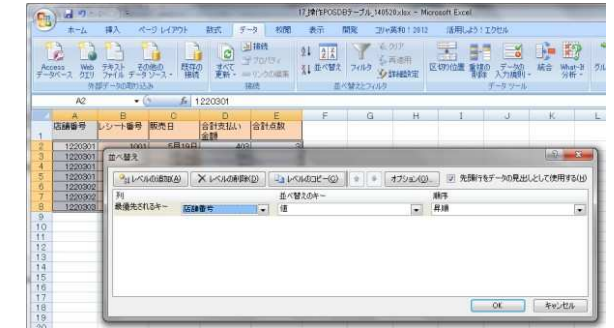
5月19日で
400円以上を抽出
シート:
2.選択.販売情報S2



例えば、特定の店舗の売り上げを計算する場合は、その店舗だけのレコードを取り出して計算する必要があります。このような時、**選択**の操作をすることになります。選択はある列の特定の条件に合うレコードだけを取り出すことができます。また、複数の列の条件を指定して取り出すこともできます。

Excelでの操作: **選択(1)** 行の削除又はコピー

選択は射影と同様にa) unnecessaryな行を削除する。b) necessaryな行をコピーすることで実現できます。



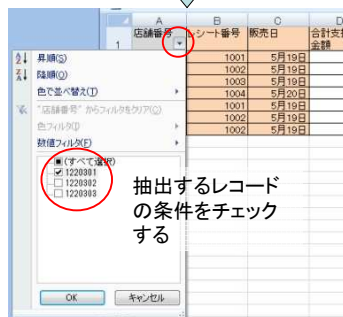
行の削除やコピーは、列に比べて操作するレコードの数が多くなります。そのため、上図のように「並び替え(ソート)」の機能を使って予めレコードをコピーする又は削除するものが連続して並ぶようにしておく、操作が楽になります。

Excelでの操作: **選択(2-a)** フィルタ機能を使用する



店舗番号	レシート番号	販売日	合計支払い金額	合計点数
1220301	1001	5月19日	403	3
1220301	1002	5月19日	590	4
1220301	1003	5月19日	250	2
1220301	1004	5月20日	439	3
1220302	1001	5月19日	911	2
1220302	1002	5月19日	156	1
1220303	1002	5月19日	321	2

条件を指定する矢印が出てくる



抽出するレコードの条件をチェックする

Excelで選択の操作をする、もう一つの方法はフィルタ機能を使うことです。フィルタ機能では、各行ごとに抽出する条件をチェックで指定することができます。慣れると、前のスライドの行を削除するなどより簡単ですね。



フィルタした結果

店舗番号	レシート番号	販売日	合計支払い金額	合計点数
1220301	1001	5月19日	403	3
1220301	1002	5月19日	590	4
1220301	1003	5月19日	250	2
1220301	1004	5月20日	439	3

指定した条件に合う行(レコード)だけが表示される: フィルタを解除すると、再度全部の行が表示される。

シート:
2.選択.販売情報S3

Excelでの操作: **選択(2-b)** フィルタ機能を使用した後で計算する

店舗番号	レシート番号	販売日	合計支払い金額
1220301	1001	5月19日	403
1220301	1002	5月19日	590
1220301	1004	5月20日	439
1220301	1003	5月19日	250
合計			3070
レコード数			7

コピーしてから計算する

店舗番号	レシート番号	販売日	合計支払い金額
1220301	1001	5月19日	403
1220301	1002	5月19日	590
1220301	1004	5月20日	439
1220301	1003	5月19日	250
合計			1692
レコード数			4

SUBTOTAL関数で計算する。

店舗番号	レシート番号	販売日	合計支払い金額
1220301	1001	5月19日	403
1220301	1002	5月19日	590
1220301	1004	5月20日	439
1220301	1003	5月19日	250
合計			3070
レコード数			7
=SUBTOTAL(9,D2:D8)			1692
=SUBTOTAL(2,D2:D8)			4
1 平均			
2 数値の個数			
3 データの個数			
9 合計			

SUBTOTAL(集計方法、範囲)

さあ、大変。フィルタで選択した後、SUM関数などを使って計算すると表示されていないレコードも計算対象になってしまいます。そんなとき
a) 別の場所・シートにコピーしてから計算する
コピーするとフィルタで選択したものだけが残ります。
b) SUBTOTAL関数を使用する
SUBTOTAL関数はフィルタで選択されたデータだけを対象に集計してくれます。



操作: 結合=二つのテーブルを合成してデータを取り出す

販売情報					店舗情報			
店舗番号	レシート番号	販売日	合計支払い金額	合計点数	店舗番号	店舗名	店舗住所	店舗電話番号
1220301	1001	5月19日	403	3	1220301	市川平田1丁目	市川市平田1-18-14	047-322-6083
1220301	1002	5月19日	590	4	1220302	市川大洲	市川市大洲3-10-13	047-379-5110
1220301	1003	5月19日	250	2	1220303	市川八幡3丁目	市川市八幡3-3-2	047-324-1219
1220301	1004	5月20日	439	3				
1220302	1001	5月19日	911	2				
1220302	1002	5月19日	156	1				
1220303	1002	5月19日	321	2				

店舗名販売情報					
店舗番号	店舗名	レシート番号	販売日	合計支払い金額	合計点数
1220301	市川平田1丁目	1001	5月19日	403	3
1220301	市川平田1丁目	1002	5月19日	590	4
1220301	市川平田1丁目	1003	5月19日	250	2
1220301	市川平田1丁目	1004	5月20日	439	3
1220302	市川大洲	1001	5月19日	911	2
1220302	市川大洲	1002	5月19日	156	1
1220303	市川八幡3丁目	1002	5月19日	321	2

結合は同一のデータ種類の列を持つ二つのテーブルを、その列の値を手がかりに合成して新しいデータを作成します。

シート: 選択.販売情報S3



DBの場合、テーブルを整理する必要がありましたが、実際レポートなど作成する時に、例えば店舗番号だけだと良くわからないことがあります。このような時実際の店舗名が入っていると判りやすいですね。結合の操作は、このような時、二つのテーブルを合成してデータを取り出すことができます。

Excelでの操作: 結合(a) 条件に一致したデータの取り出し

店舗名販売情報					店舗情報		
店舗番号	店舗名	レシート番号	販売日	合計支払い金額	店舗番号	店舗名	店舗住所
1220301	市川平田1丁目	1001	5月19日	403	1220301	市川平田1丁目	市川市平田1-18-14
1220301	市川平田1丁目	1002	5月19日	590	1220302	市川大洲	市川市大洲3-10-13
1220301	市川平田1丁目	1003	5月19日	250	1220303	市川八幡3丁目	市川市八幡3-3-2
1220301	市川平田1丁目	1004	5月20日	439			
1220302	市川大洲	1001	5月19日	911			
1220302	市川大洲	1002	5月19日	156			
1220303	市川八幡3丁目	1002	5月19日	321			

A	B
15	=VLOOKUP(A15,\$G\$15:\$H\$18,2,FALSE)
16	=VLOOKUP(A16,\$G\$15:\$H\$18,2,FALSE)
17	=VLOOKUP(A17,\$G\$15:\$H\$18,2,FALSE)

VLOOKUP関数は、範囲指定した表の一番左の値を比較して、一致したレコードの値を取り出すことができます。

演習1-a 販売情報テーブルをもとに、射影の操作をして下記の項目のテーブルを作成する。

店舗番号	レシート番号	販売日	合計支払い金額	性別区分	年齢
------	--------	-----	---------	------	----

演習1-b 単品売上情報から10101オニギリ シャケと10103オニギリコブを選択したテーブルを作成する。作成は並べ替えをしてから行を削除する。

演習1-c 単品売上情報と商品マスターを結合して下記の項目のテーブルを作成する。

店舗番号	レシート番号	販売日	商品番号	商品名	商品購入個数	商品小計金額
------	--------	-----	------	-----	--------	--------

演習2-a 単品売上情報から10101オニギリ シャケと10103オニギリコブを選択したテーブルを作成する。作成はフィルタを使用し、SUBTOTAL関数で小計を計算しなさい。

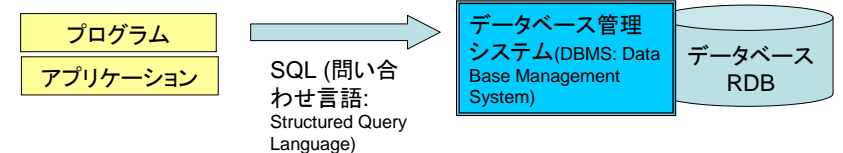
演習2-b 販売情報と展墓情報を結合・射影の操作をして下記の項目のテーブルを作成する。

店舗番号	店舗名	レシート番号	販売日	販売日時	合計支払い金額	性別区分	年齢
------	-----	--------	-----	------	---------	------	----

演習2-c 演習1-cのテーブルをもとに、行を削除して10101オニギリ シャケと10103オニギリコブを選択したテーブルを作成する。



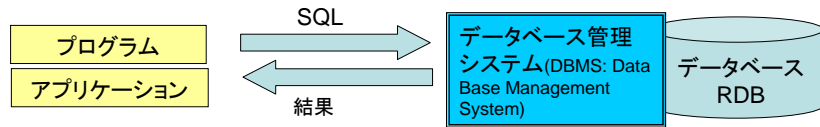
データベース管理システムとSQL



データベースはコンピュータ上のデータの集まりですが、それを管理しているデータベース管理システム(DBMS)があります。実際にプログラムでデータベースを使用する場合は、このDBMSを仲介してアクセスします。特にリレーショナルデータベースではSQLというデータベースを操作する問い合わせ言葉があり、この言葉を使ってデータベース管理システムに命令を出します。SQLはいろいろなシステムやメーカーで共通なので、異なるメーカーのデータベース管理システムを使っても、プログラムでのSQLは同じになります。



SQLとその結果



射影: SELECT 店舗番号, レシート番号, 販売日, 合計支払い金額, 合計点数 FROM 販売情報;

選択: SELECT 店舗番号, レシート番号, 販売日, 合計支払い金額, 合計点数 FROM 販売情報 WHERE 店舗番号 = 1220301;

結合: SELECT 販売情報.店舗番号, 店舗情報.店舗名, 販売情報.[レシート番号], 販売情報.販売日, 販売情報.合計支払い金額, 販売情報.合計点数 FROM 販売情報 INNER JOIN 店舗情報 ON 販売情報.店舗番号 = 店舗情報.店舗番号;

授業ではデータベースの操作をExcelの手作業で行いましたが、例で示した射影、選択、結合の操作をSQLで表現すると左のようになります。これをDBMSに要求するだけで、Excelで作ったようなデータが結果として得ることができます。SQLを知れば簡単にデータベースを操作できるようになります。

