

データベースシステム入門

7. 集計, 集約

リレーショナルデータベースシステム

コンピュータ



リレーショナル
データベース
管理システム



記憶
装置

リレーショナル
データベース

あわせて
リレーショナルデータベースシステム

	id	name	teacher_name	student_name	score	created_at	updated_at
1	1	Database	K	KK	85	2014-09-13 00:46:17	{null}
2	2	Database	K	AA	75	2014-09-13 00:48:26	{null}
3	3	Database	K	LL	90	2014-09-13 00:48:26	{null}
4	4	Programming	A	KK	85	2014-09-13 00:48:26	{null}
5	5	Programming	A	LL	75	2014-09-13 00:48:26	{null}

	id	year	month	day	customer_name	product_id	qty	created_at
1	1001	2014	10	26	kaneko	1	10	2014-09-14 14:28:38
2	1002	2014	10	26	miyamoto	2	2	2014-09-14 14:28:38
3	1003	2014	10	27	kaneko	3	8	2014-09-14 14:28:38
4	1004	2014	10	27	kaneko	3	8	2014-09-14 14:28:38

データの種類ごとに分かれた、
たくさんの**テーブル**が格納される

SQLをマスターするには

◆ SQL のキーワード

create table テーブル定義

select 射影など

from 扱うテーブルの指定

where 選択, 結合など

max 最大

min 最小

count 数え上げ

group by 集約 (**グループ**の基準の指定)

order by ソート (並べ替え)

distinct 重複除去 など

◆ SQL のデータ型

短いテキスト char

長いテキスト text

数値 integer, real

日付／時刻 datetime

YesやNo bit

など

今日の授業で学ぶこと

元データ

科目	受講者	得点
国語	A	85
国語	B	90
算数	A	90
算数	B	96
理科	A	95



SQLで

集計

A	3
B	2

行数

A	90
B	93

得点の平均

国語	2
算数	2
理科	1

行数

国語	87.5
算数	93
理科	95

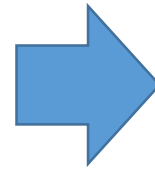
得点の平均

7-1 集計

集計の例

科目	受講者	得点
国語	A	85
国語	B	90
算数	A	90
算数	B	96
理科	A	95

元データ
(テーブル名は「成績」)



Aさんは 3科目
Bさんは 2科目受講した

A	3
B	2

集計の例

集計の例

科目	受講者	得点
国語	A	85
国語	B	90
算数	A	90
算数	B	96
理科	A	95

テーブル名は「成績」

SQLで、
集計も簡単にできます

```
select 受講者, count(*)  
from 成績  
group by 受講者;
```

A	3
B	2

受講者ごとの行数



問い合わせ
(クエリ)



結果

7-1-1 集計とは

- ◆ **集計**とは、**データの集まり**に対して、
何らかの計算を行うこと

科目	受講者	得点
国語	A	85
国語	B	90
算数	A	90
算数	B	96
理科	A	95

元データ



A	3
B	2



ノートページ

集計の方法のバリエーション

集計の方法は行数、合計、平均、最大、最小など

科目	受講者	得点
国語	A	85
国語	B	90
算数	A	90
算数	B	96
理科	A	95

元データ



A	270
B	186

得点の合計

A	3
B	2

行数

A	90
B	93

得点の平均

集計の例

7-1-2 集計の方法のバリエーション

◆集計の方法は

行数	count
合計	sum
平均	avg
最大	max
最小	min
など	



ノートページ

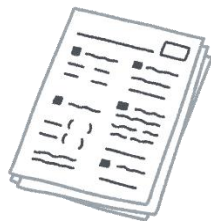
集計でのグループの基準



グループの基準が
受講者

科目	受講者	得点
国語	A	85
国語	B	90
算数	A	90
算数	B	96
理科	A	95

元データ



グループの基準が
科目

A	90
B	93

得点の平均

国語	87.5
算数	93
理科	95

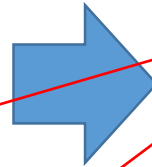
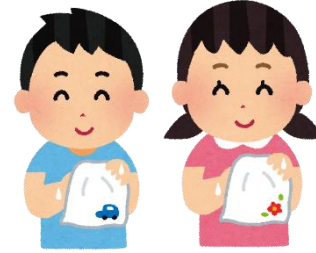
得点の平均

集計の例

グループの基準が受講者

科目	受講者	得点
国語	A	85
国語	B	90
算数	A	90
算数	B	96
理科	A	95

元データ



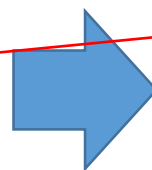
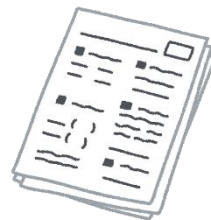
A	90
B	93

得点の平均

グループの基準が科目

科目	受講者	得点
国語	A	85
国語	B	90
算数	A	90
算数	B	96
理科	A	95

元データ



国語	87.5
算数	93
理科	95

得点の平均

7-1-3 集計とグループ化

- ◆ どのフィールドでグループ化するのか
= **グループの基準**



ノートページ

グループの基準が
受講者

A	90
B	93

得点の平均

グループの基準が
科目

国語	87.5
算数	93
理科	95

得点の平均

7-2 SQL で集計を行う

集計の例

科目	受講者	得点
国語	A	85
国語	B	90
算数	A	90
算数	B	96
理科	A	95

テーブル名は「成績」

SQLで、
集計も簡単にできます

```
select 受講者, count(*)  
from 成績  
group by 受講者;
```

A	3
B	2

受講者ごとの行数



問い合わせ
(クエリ)



結果

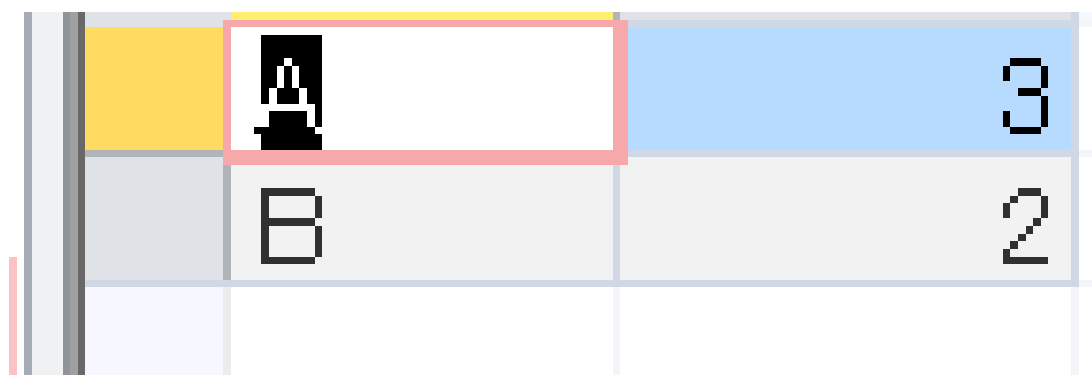
集計を行う SQL の例

```
select count(*) from 成績 group by 受講者;
```



3
2

```
select 受講者, count(*) from 成績 group by 受講者;
```



	3
B	2

集計結果だけでは分かりにくいので一工夫

集計を行う SQL の例

```
select 受講者, count(*) from 成績 group by 受講者;
```

A	3
B	2

```
select 科目, count(*) from 成績 group by 科目;
```

国語	2
算数	2
理科	1

グループの基準の違い

グループの基準の違い

A

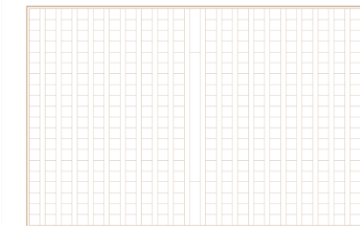


B

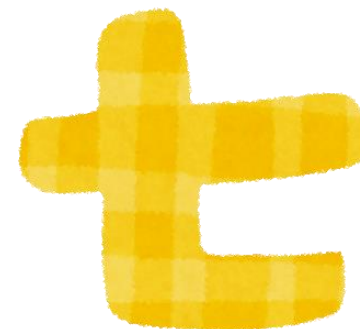


科目	受講者	得点
国語	A	85
国語	B	90
算数	A	90
算数	B	96
理科	A	95

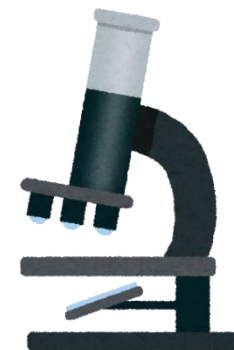
国語



算数



理科



グループの基準の違い

A

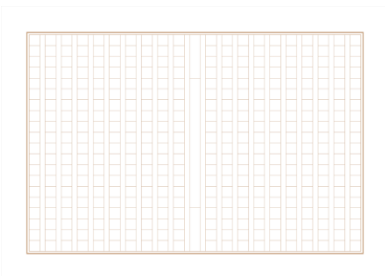


B

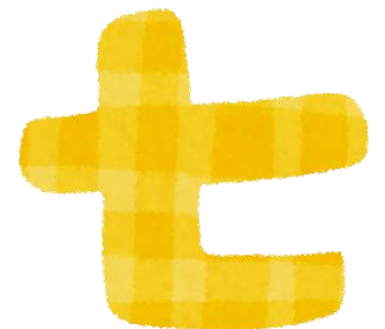


科目	受講者	得点
国語	A	85
国語	B	90
算数	A	90
算数	B	90
理科	A	95

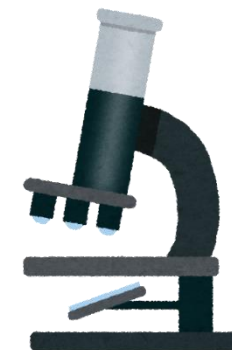
国語



算数



理科



グループの基準が**科目**

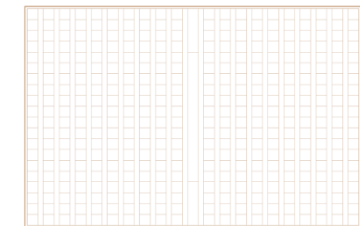
グループの基準の違い

A

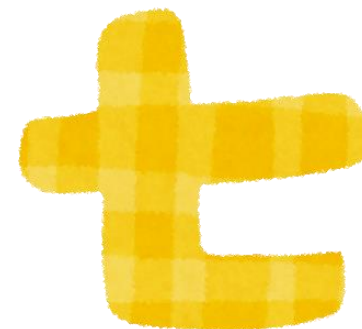


科目	受講者	得点
国語	A	85
国語	B	90
算数	A	90
算数	B	96
理科	A	95

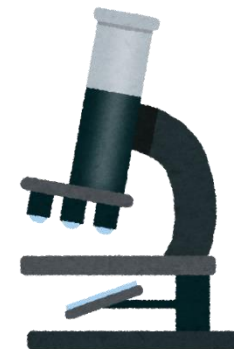
国語



算数



理科



B



グループの基準が**受講者**

集計を行うときの計画

科目	受講者	得点
国語	A	85
国語	B	90
算数	A	90
算数	B	96
理科	A	95

元データ
(テーブル名: **成績**)

合計、平均、最大、最小を
どのフィールドについて
求めるのかの指定

集計をどのように行うか 計画を立てる

グループの基準	☐	科目
	☐	受講者
	☐	得点
集計の方法	☐	行数
	☐	合計
	☐	平均
	☐	最大
	☐	最小
集計するフィールド	☐	科目
	☐	受講者
	☐	得点

集計を行う

科目	受講者	得点
国語	A	85
国語	B	90
算数	A	90
算数	B	96
理科	A	95

元データ

(テーブル名: 成績)

```
select 受講者, count(*)  
from 成績  
group by 受講者;
```

SQL

グループの基準

☐ 科目
☒ 受講者
☐ 得点

集計の方法

☒ 行数
☐ 合計
☐ 平均
☐ 最大
☐ 最小

集計するフィールド

☐ 科目
☐ 受講者
☐ 得点

A	3
B	2

実行結果

集計を行う

科目	受講者	得点
国語	A	85
国語	B	90
算数	A	90
算数	B	96
理科	A	95

元データ

(テーブル名：成績)

```
select 科目, count(*)  
from 成績  
group by 科目;
```

SQL

グループの基準



科目



受講者



得点

集計の方法



行数



合計



平均



最大



最小



最小

集計するフィールド



科目



受講者



得点

A	3
B	2

国語	2
算数	2
理科	1

実行結果

集計を行う

科目	受講者	得点
国語	A	85
国語	B	90
算数	A	90
算数	B	96
理科	A	95

元データ

(テーブル名：成績)

```
select 科目, sum(得点)
from 成績
group by 科目;
```

SQL

グループの基準

☒ 科目
☐ 受講者
☐ 得点

集計の方法

☐ 行数
☒ 合計
☐ 平均
☐ 最大
☐ 最小

集計するフィールド

☐ 科目
☐ 受講者
☒ 得点

国語	175
算数	186
理科	95

実行結果

集計を行う

科目	受講者	得点
国語	A	85
国語	B	90
算数	A	90
算数	B	96
理科	A	95

元データ

(テーブル名: 成績)

```
select 科目, avg(得点)
from 成績
group by 科目;
```

SQL

グループの基準

☒ 科目
☐ 受講者
☐ 得点

集計の方法

☐ 行数
☐ 合計
☒ 平均
☐ 最大
☐ 最小

集計するフィールド

☐ 科目
☐ 受講者
☒ 得点

国語	87.5
算数	93
理科	95

実行結果

集計を行う

科目	受講者	得点
国語	A	85
国語	B	90
算数	A	90
算数	B	96
理科	A	95

元データ

(テーブル名: 成績)

```
select 科目, max(得点)
from 成績
group by 科目;
```

SQL

グループの基準

☒ 科目
☐ 受講者
☐ 得点

集計の方法

☐ 行数
☐ 合計
☐ 平均
☒ 最大
☐ 最小

集計するフィールド

☐ 科目
☐ 受講者
☒ 得点

科目	Expr1 (
国語	90
算数	96
理科	95

実行結果

集計を行う

科目	受講者	得点
国語	A	85
国語	B	90
算数	A	90
算数	B	96
理科	A	95

元データ

(テーブル名：成績)

```
select 科目, min(得点)
from 成績
group by 科目;
```

SQL

グループの基準

- ☒ 科目
- ☐ 受講者
- ☐ 得点

集計の方法

- ☐ 行数
- ☐ 合計
- ☐ 平均
- ☐ 最大
- ☒ 最小

集計するフィールド

- ☐ 科目
- ☐ 受講者
- ☒ 得点

科目	Expr1 (
国語	85
算数	90
理科	95

実行結果

集計を行う SQL の例

```
select 受講者, count(*) from 成績 group by 受講者;
```

```
select 科目, count(*) from 成績 group by 科目;
```

```
select 科目, sum(得点) from 成績 group by 科目;
```

```
select 科目, avg(得点) from 成績 group by 科目;
```

```
select 科目, max(得点) from 成績 group by 科目;
```

```
select 科目, min(得点) from 成績 group by 科目;
```

集計の方法のバリエーション

7-3 集計を行う SQL の書き方

select ○, ○, ○

フィールド名、
count(*)
sum(集計するフィールド名)
avg(集計するフィールド名)
max(集計するフィールド名)
min(集計するフィールド名)

from <集計したいテーブル名>

group by <グループの基準>



ノートページ

データベース 対 Excel (表計算)

データベース

	路線コード番号	事業者コード番号	路線名称一般	路線名称一般カナ
1	1001	3	中央新幹線	チュウオウシンカンセン
2	1002	3	東海道新幹線	トウカイドウシンカンセン
3	1003	4	山陽新幹線	サンヨウシンカンセン
4	1004	2	東北新幹線	トウホクシンカンセン

エクセル

	A	B	C	D
1	路線コード番号	事業者コード番号	路線名称一般	路線名称一般カナ
2	1001	3	中央新幹線	チュウオウシンカンセン
3	1002	3	東海道新幹線	トウカイドウシンカンセン
4	1003	4	山陽新幹線	サンヨウシンカンセン
5	1004	2	東北新幹線	トウホクシンカンセン

- データ検索, データ共有, セキュリティ
- 集計
- 並べ替え (ソート)
- 「結合」の機能で,
複数のテーブルを1つにまとめる
- 他のプログラムとの連携は簡単
- SQLコマンド

- 表計算, グラフ
- 集計
- 並べ替え (ソート)
- 「参照」の機能で,
複数のワークシート間を参照
- 他のプログラムとの連携は簡単
- ビジュアルに操作

両方使うのも良い

Excel 2013 で集計を行う (1 / 4)

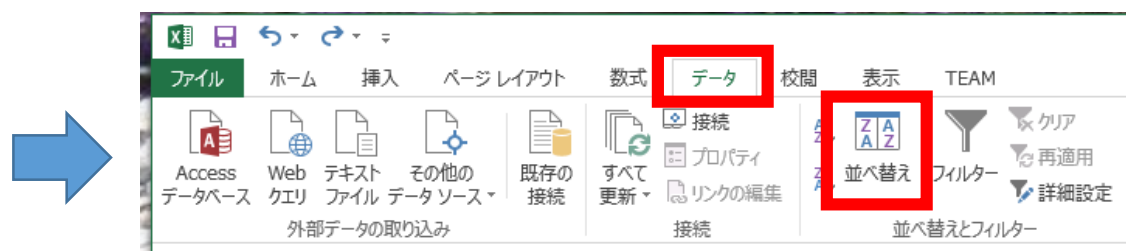
	A	B	C
1	科目	受講者	得点
2	国語	A	85
3	国語	B	90
4	算数	A	90
5	算数	B	96
6	理科	A	95
7			

元データ



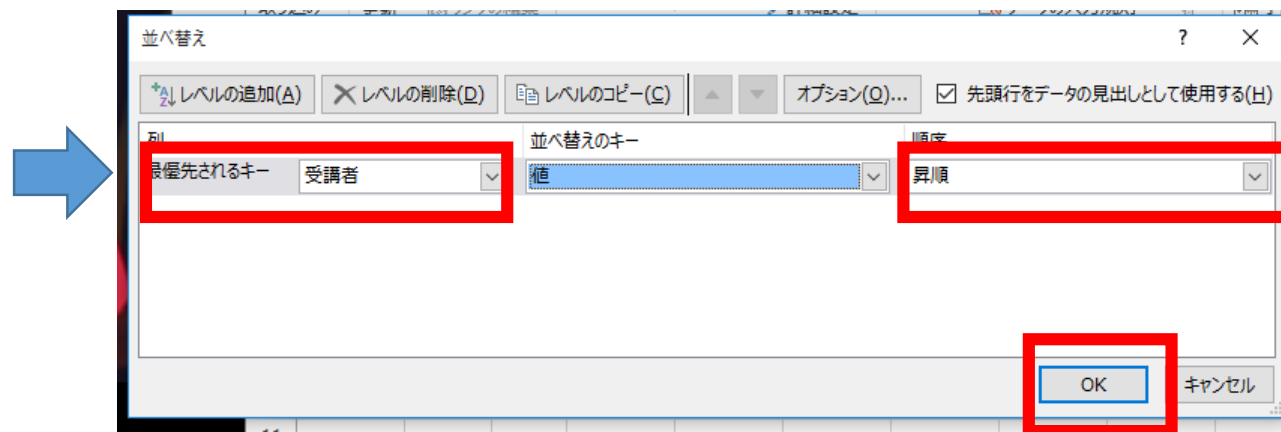
	A	B	C
1	科目	受講者	得点
2	国語	A	85
3	国語	B	90
4	算数	A	90
5	算数	B	96
6	理科	A	95
7			

- ① 左上をクリック. 全セルが選択される
(これは集計したい**範囲の選択**)

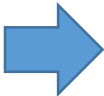


- ② リボンで「データ」→「並べ替え」

Excel 2013 で集計を行う (2 / 4)



③ 「最優先されるキー」には、グループの基準を指定。
「順序」には昇順を設定して「OK」



	A	B	C
1	科目	受講者	得点
2	国語	A	85
3	算数	A	90
4	理科	A	95
5	国語	B	90
6	算数	B	96

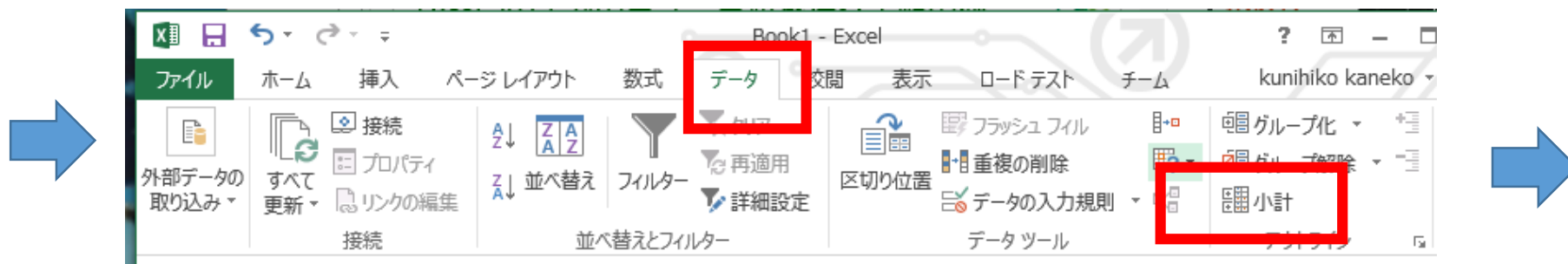
④ 並べ替え結果が得られる

Excel 2013 で集計を行う (3 / 4)



	A	B	C
1	科目	受講者	得点
2	国語	A	85
3	算数	A	90
4	理科	A	95
5	国語	B	90
6	算数	B	96
7			

⑤ 再び**左上**をクリック。
全セルが選択される
(これは集計したい
範囲の選択)



⑥ リボンで「データ」、「小計」

Excel 2013 で集計を行う (4 / 4)

集計の設定

グループの基準(A):
受講者

集計の方法(U):
データの個数

集計するフィールド(D):
☐ 科目
☐ 受講者
☒ 得点

☒ 現在の小計をすべて置き換える(C)
☐ グループごとに改ページを挿入する(P)
☒ 集計行をデータの下に挿入する(S)

すべて削除(R) OK キャンセル

受講者

行数

得点



	A	B	C
1	科目	受講者	得点
2	国語	A	85
3	算数	A	90
4	理科	A	95
5		Aデータの個数	3
6	国語	B	90
7	算数	B	96
8		Bデータの個数	2
9		総合計	5
10			

⑧ 集計できた

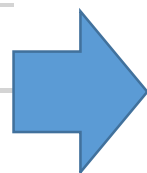
⑦ グループの基準、集計の方法、集計するフィールドを設定

7-3 基準が複数あるような集計

クロス集計表（ピボットテーブル）の例

名前		性別	申し込み
A		男性	済
B		男性	未
C		女性	済
D		女性	未
E		男性	済
F		男性	未

元データ



女性	済	1
女性	未	1
男性	済	2
男性	未	2

	男性	女性
済	2	1
未	2	1

クロス集計表

（ピボットテーブルともいう）の例

※ 形が違う 2 種類

SQL でのクロス集計表（ピボットテーブル）は

```
SELECT 性別, 申し込み, count(*)  
FROM 申し込み  
GROUP BY 性別, 申し込み;
```

SQL

女性	済	1
女性	未	1
男性	済	2
男性	未	2

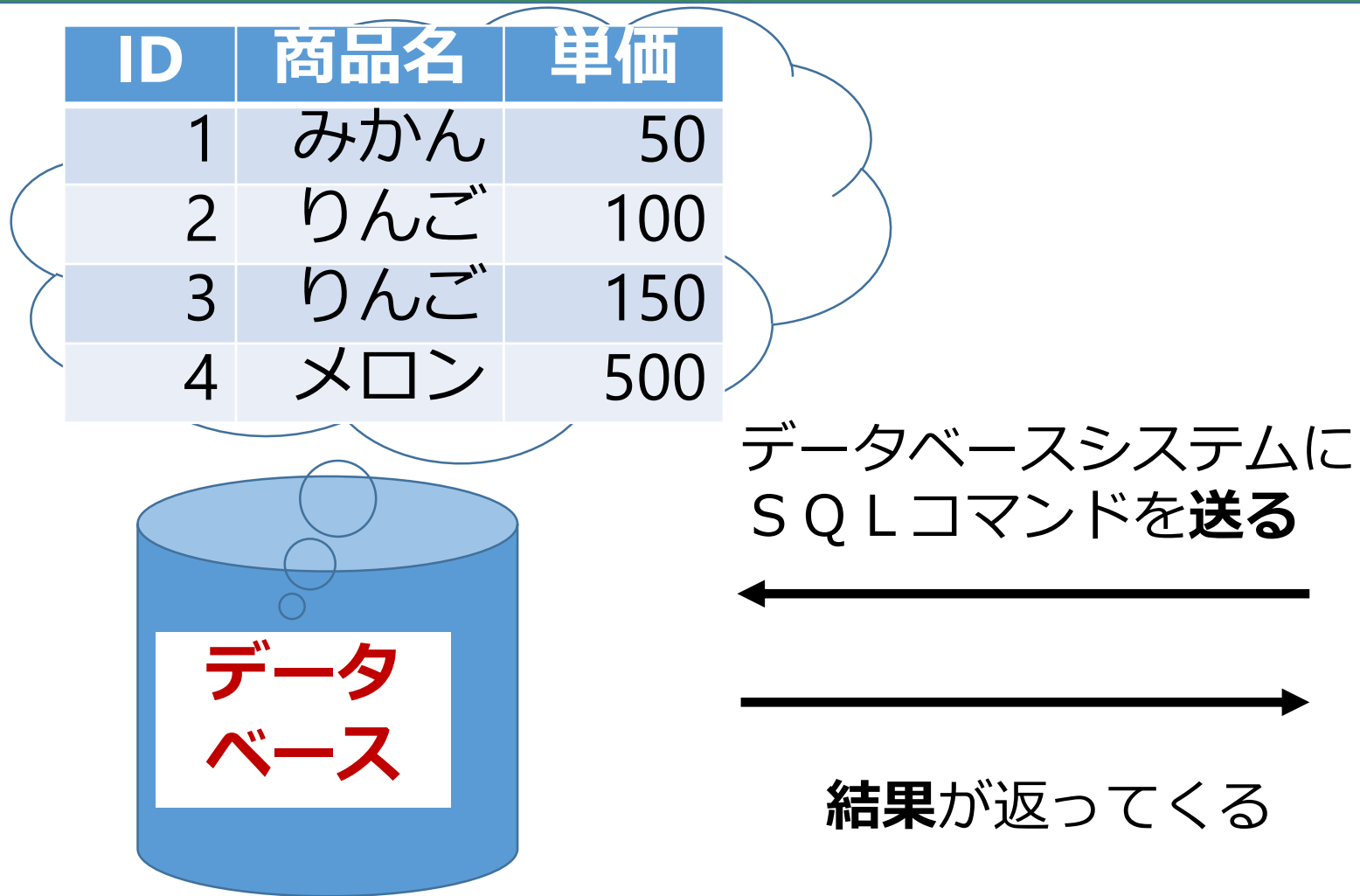
SQL の実行結果

女性	済	1
女性	未	1
男性	済	2
男性	未	2

SQL では、こちらの「形」
の**クロス集計表**（ピボットテーブル）
を簡単に作ることができる

7-5 SQL の復習と「分解」の紹介

SQLの使い方



データベース
利用者

SQLで、テーブルをそのまま表示

ID	商品名	単価
1	みかん	50
2	りんご	100
3	りんご	150
4	メロン	500

元のテーブルの
まま表示

1	<code>select * from 商品;</code>
2	
3	
4	

ID	商品名	単価
1	みかん	50
2	りんご	100
3	りんご	150
4	メロン	500



問い合わせ
(クエリ)



結果

フィールドの表示／非表示（射影）の例

ID	商品名	単価
1	みかん	50
2	りんご	100
3	りんご	150
4	メロン	500

フィールドの表示／非表示 （「射影」といいます）



問い合わせ
(クエリ)

 クエリ1
SELECT 商品名, 単価 FROM 商品;



結果

商品名	単価
みかん	50
りんご	100
りんご	150
メロン	500

レコードの絞り込み（選択）の例

ID	商品名	単価
1	みかん	50
2	りんご	100
3	りんご	150
4	メロン	500

全体で「射影＋選択」

レコードの絞り込み
(選択)



```
1 select 商品名, 単価 from 商品
2 where 単価 > 80;
3
```

商品名	単価
りんご	100
りんご	150
メロン	500

問い合わせ
(クエリ)

結果

※ 途中で改行してもいいし、改行しなくてもよい (SQL のルール)

結合の例

ID	商品名	単価
1	みかん	50
2	りんご	100
3	りんご	150
4	メロン	500

ID	名前	商品番号
1	X	3
2	Y	1

商品.ID	商品名	単価	購入.ID	名前	商品番号
1	みかん	50	1	X	3
1	みかん	50	2	Y	1
2	りんご	100	1	X	3
2	りんご	100	2	Y	1
3	りんご	150	1	X	3
3	りんご	150	2	Y	1
4	メロン	500	1	X	3
4	メロン	500	2	Y	1

```
select * from 商品, 購入;
```



問い合わせ
(クエリ)



結果

分解の例

講義名	担当教員	受講者
DB	K	CC
DB	K	AA
DB	K	BB
プロ	A	AA
プロ	A	DD

講義名	担当教員
DB	K
プロ	A

講義名	受講者
DB	CC
DB	AA
DB	BB
プロ	AA
プロ	DD

```
select DISTINCT 講義名 担当教員 from 授業;  
select DISTINCT 講義名 受講者 from 授業;
```

結果が2つ欲しい
とき、SQLも2つ



問い合わせ
(クエリ)



結果