

【グループを極める（その2）】

GROUP BY句に複数の列を指定する場合（その1）

前回の投稿（グループを極める（その1））において「GROUP BY句で指定していない列をSELECT句に指定することはできません」と記載しました。逆にいえば、GROUP BY句で指定すれば、SELECT句に指定することができます。しかし、GROUP BY句に指定すると、その列もグループ化されるため、むやみに指定するものではありません。

例えば、左の下表（受注状況）に対して、以下のSELECT文を実行した結果が、右の下表（実行結果その3）なのですが、受注日毎商品番号毎にソートされただけであり、先の投稿の表（実行結果その1）及び表（実行結果その2）と比較して、**グループ化の恩恵は見受けられません**。また、右の下表（実行結果その3）のデータは昇順になっていますが、これは暗黙的にソート（暗黙ソート）されていることが前提です。その前提がなければ、ORDER BY句を使う必要があります。

SELECT 受注日, 商品番号 FROM 受注状況 GROUP BY 受注日, 商品番号

表:受注状況

受注日	商品番号
4 / 1	200
4 / 1	300
4 / 1	400
4 / 1	200
4 / 2	100
4 / 2	300
4 / 2	500
4 / 2	400
4 / 2	300
4 / 2	100
4 / 3	400
4 / 3	500

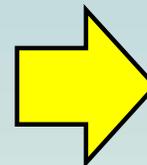
受注日 4 / 1 でグループ化
商品番号毎にグループ化
それぞれの数量が異なるため、
重複行はない

受注日 4 / 2 でグループ化
商品番号毎にグループ化
それぞれ数量が異なるため、
重複行はない

受注日 4 / 3 でグループ化
商品番号毎にグループ化
それぞれ数量が異なるため、
重複行はない

表:実行結果その3

受注日	商品番号
4 / 1	200
4 / 1	200
4 / 1	300
4 / 1	400
4 / 2	100
4 / 2	100
4 / 2	300
4 / 2	300
4 / 2	400
4 / 2	500
4 / 3	400
4 / 3	500



GROUP BY句に複数の列を指定する場合（その2）

前ページにおいて「グループ化の恩恵は見受けられません」と記載しましたが、その理由は、受注日毎商品番号毎にグループ化する目的が曖昧だからです。例えば、受注日毎商品番号毎の合計数量を得たいという目的があれば、以下のSELECT文を実行することで、左の下表（実行結果その3）は、右の下表（実行結果その4）のようになり、GROUP化の恩恵を得ることができます。

```
SELECT 受注日, 商品番号, SUM(数量) FROM 受注状況 GROUP BY 受注日, 商品番号
```

表:実行結果その3

受注日	商品番号	数量
4 / 1	200	3
4 / 1	200	4
4 / 1	300	2
4 / 1	400	6
4 / 2	100	2
4 / 2	100	7
4 / 2	300	5
4 / 2	300	8
4 / 2	400	12
4 / 2	500	10
4 / 3	400	9
4 / 3	500	10

商品番号200で数量をまとめることができる

商品番号100で数量をまとめることができる

商品番号300で数量をまとめることができる

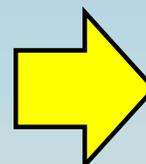


表:実行結果その4

受注日	商品番号	SUM(数量)
4 / 1	200	7
4 / 1	300	2
4 / 1	400	6
4 / 2	100	9
4 / 2	300	13
4 / 2	400	12
4 / 2	500	10
4 / 3	400	9
4 / 3	500	10

集約関数COUNTの注意点 (その1)

COUNTはグループ化することにより、ややこしくなります。例えば、以下の2つのSELECT文の相違点はDISTINCTだけですが、列 (商品番号) のCOUNT結果は、下表 (実行結果その5) 及び下表 (実行結果その6) のとおり、全く違ってきます。

```
SELECT 受注日, COUNT(商品番号), COUNT(*) FROM 受注状況 GROUP BY 受注日
SELECT 受注日, COUNT(DISTINCT 商品番号), COUNT(*) FROM 受注状況 GROUP BY 受注日
```

表:受注状況

受注日	商品番号	数量
4 / 1	200	3
4 / 1	300	2
4 / 1	400	6
4 / 1	200	4
4 / 2	100	2
4 / 2	300	8
4 / 2	500	10
4 / 2	400	12
4 / 2	300	5
4 / 2	100	7
4 / 3	400	9
4 / 3	500	10

受注日 4 / 1 でグループ化しており、COUNT(*)は、グループ内の行数なので、**4**
また、COUNT(商品番号)も、同様であり、重複行もカウントするため、**4**
ただし、COUNT(DISTINCT 商品番号)は、重複行をまとめるため、商品番号 2-0-0 は1行扱いとなり、**3**

上記と同様に、COUNT(*)は、グループ内の行数なので、**6**
また、COUNT(商品番号)も、重複行もカウントするため、**6**
ただし、COUNT(DISTINCT 商品番号)は、重複行をまとめるため、商品番号 1 0 0 及び 3 0 0 はそれぞれ1行扱いとなり、**4**

もともと、重複行はないため、全てのCOUNTは、2

表:実行結果その5

受注日	COUNT (商品番号)	COUNT (*)
4 / 1	4	4
4 / 2	6	6
4 / 3	2	2

表:実行結果その6

受注日	COUNT (DISTINCT 商品番号)	COUNT (*)
4 / 1	3	4
4 / 2	4	6
4 / 3	2	2

集約関数COUNTの注意点 (その2)

次に、受注日毎商品番号毎の行数を得たい場合には、以下のSELECT文を実行することで、左の下表（実行結果その7）のようになります。つまり、GROUP BY句に複数の列を指定した場合のCOUNT(*)は、最後に指定した列の行数となります。

ただし、GROUP BY句に複数の列を指定する場合、大カテゴリから小カテゴリになるように指定しないと、自動調整が働くわけではないので、変な結果となります。同様に、SELECT句で指定する列の順番も意識しないと、目的の結果を得ることはできません。抽象的な表現になりましたが、目的をもってSQL文を作成することが重要です。

```
SELECT 受注日, 商品番号, COUNT(*) FROM 受注状況 GROUP BY 受注日, 商品番号
```

表:受注状況

受注日	商品番号	数量
4 / 1	200	3
4 / 1	300	2
4 / 1	400	6
4 / 1	200	4
4 / 2	100	2
4 / 2	300	8
4 / 2	500	10
4 / 2	400	12
4 / 2	300	5
4 / 2	100	7
4 / 3	400	9
4 / 3	500	10

COUNT(*)は、受注日4 / 1でグループ化して、更に、商品番号毎にグループ化したグループ内の行数なので、200は、2
300と400は、それぞれ1

同様に、100と300は、それぞれ2
400と500は、それぞれ1

同様に、
400と500は、それぞれ1

表:実行結果その7

受注日	商品番号	COUNT(*)
4 / 1	200	2
4 / 1	300	1
4 / 1	400	1
4 / 2	100	2
4 / 2	300	2
4 / 2	400	1
4 / 2	500	1
4 / 3	400	1
4 / 3	500	1

集約関数COUNTの注意点 (その3)

これまでの投稿から、COUNT(*)は、グループ化されたグループ内の行数であり、グループ化された個数ではありません。よって、上表（受注状況）において、受注日でグループ化する場合、グループ化された個数は、4 / 1、4 / 2及び4 / 3の3つです。この3を得るためには、別途SELECT文「SELECT COUNT(DISTINCT 受注日) FROM 受注状況」により求める方が確実です。