



BUỔI 6 **BẢNG TỔNG HỢP DỮ LIỆU (PIVOT TABLE)** **& CHỦ ĐỀ XẾP HẠNG (RANKING TOPIC)**

1

Phương pháp PIVOT TABLE

Phần 1: Chọn các cột cần hiển thị

SELECT

Customer_name,
"Office Supplies",
"Furniture",
"Technology"

FROM

(
SELECT Customer_name,
product_category,
profit
FROM orders
) AS PivotData

PIVOT

(
SUM(profit)
FOR product_category IN ("Office Supplies", "Furniture", "Technology")
) AS PivotTable

Input

Customer_name	Product_category	Profit
Muhammed MacIntyre	Office Supplies	2.382399999999996
Ruben Dartt	Office Supplies	-0.7296
Liz Pelletier	Furniture	777.0084
Liz Pelletier	Furniture	233.9584

Customer_name	Office Supplies	Furniture	Technology
Tamara Chand	614.4988	1026.5506	11.015
Brian Moss	1759.367	2884.2064	2142.5348
Justin MacKendrick	-25.0334	7194.4288	7249.0196
Michael Chen	1439.2566	-38.2226	7833.9255

Out put

Phần 2: Data cần pivot

Phần 3: Chuyển đổi dòng thành cột

Cú pháp như sau:

SELECT

<non-pivoted column>, ---> Cột không xoay chuyển

[first pivoted column] **AS** <column name>, ---> Giá trị đầu tiên được xoay chuyển thành cột

[second pivoted column] **AS** <column name>, ---> Giá trị thứ hai được xoay chuyển thành cột

[last pivoted column] **AS** <column name> ---> Giá trị cuối được xoay chuyển thành cột

FROM

(<SELECT query that produces the data>) **AS** PivotData

PIVOT

(<Hàm tổng hợp>

FOR [<Cột chứa các giá trị sẽ trở thành tiêu đề cột>

IN ([first pivoted column], [second pivoted column], ... [last pivoted column])) **AS** PivotTable

;

Chủ đề XẾP HẠNG (RANKING TOPIC)

Sử dụng hàm **ROW_NUMBER()**, **RANK()** và **DENSE_RANK()**

Phân biệt **ROW_NUMBER**, **RANK** và **DENSE_RANK**

Đầu tiên, bạn nhìn ví dụ sau:

```
SELECT Customer_name, order_id, Order_Date,
ROW_NUMBER() OVER(ORDER BY Order_Date ASC) AS [ROW_NUMBER],
RANK() OVER(ORDER BY Order_Date ASC) AS [RANK],
DENSE_RANK() OVER(ORDER BY Order_Date ASC) AS [DENSE_RANK]
FROM Orders
WHERE Customer_name = 'Tamara Chand';
```

Results

Messages

	Customer_name	order_id	Order_Date	ROW_NUMBER	RANK	DENSE_RANK
1	Tamara Chand	57415	2009-03-24 00:00:00.000	1	1	1
2	Tamara Chand	6592	2009-12-01 00:00:00.000	2	2	2
3	Tamara Chand	6592	2009-12-01 00:00:00.000	3	2	2
4	Tamara Chand	640	2010-01-22 00:00:00.000	4	4	3
5	Tamara Chand	640	2010-01-22 00:00:00.000	5	4	3
6	Tamara Chand	44960	2011-05-03 00:00:00.000	6	6	4
7	Tamara Chand	44960	2011-05-03 00:00:00.000	7	6	4
8	Tamara Chand	44960	2011-05-03 00:00:00.000	8	6	4

Chủ đề XẾP HẠNG (RANKING TOPIC)

<div> <div>Results</div> <div>Messages</div> </div>						
	Customer_name	order_id	Order_Date	ROW_NUMBER	RANK	DENSE_RANK
1	Tamara Chand	57415	2009-03-24 00:00:00.000	1	1	1
2	Tamara Chand	6592	2009-12-01 00:00:00.000	2	2	2
3	Tamara Chand	6592	2009-12-01 00:00:00.000	3	2	2
4	Tamara Chand	640	2010-01-22 00:00:00.000	4	4	3
5	Tamara Chand	640	2010-01-22 00:00:00.000	5	4	3
6	Tamara Chand	44960	2011-05-03 00:00:00.000	6	6	4
7	Tamara Chand	44960	2011-05-03 00:00:00.000	7	6	4
8	Tamara Chand	44960	2011-05-03 00:00:00.000	8	6	4

Nhận xét:

Cả 3 hàm đều Order by **Order_Date**

Hàm ROW_NUMBER dùng để xếp hạng kết quả một cách tuần tự bắt đầu từ 1 và không quan tâm đến các giá trị giống nhau.

Tại dòng thứ 2 và dòng 3, giá trị Order_Date đều giống nhau. RANK và DENSE_RANK đều trả về giá trị là 2.

Tại dòng thứ 4, hàm RANK trực tiếp bắt kịp giá trị của ROW_NUMBER bằng cách bỏ đi giá trị 3. Còn DENSE_RANK thì trả về giá trị tiếp theo dựa trên dòng trước đó.

Tương tự với các dòng dưới

Chủ đề XẾP HẠNG (RANKING TOPIC)

<div> <div>Results</div> <div>Messages</div> </div>						
	Customer_name	order_id	Order_Date	ROW_NUMBER	RANK	DENSE_RANK
1	Tamara Chand	57415	2009-03-24 00:00:00.000	1	1	1
2	Tamara Chand	6592	2009-12-01 00:00:00.000	2	2	2
3	Tamara Chand	6592	2009-12-01 00:00:00.000	3	2	2
4	Tamara Chand	640	2010-01-22 00:00:00.000	4	4	3
5	Tamara Chand	640	2010-01-22 00:00:00.000	5	4	3
6	Tamara Chand	44960	2011-05-03 00:00:00.000	6	6	4
7	Tamara Chand	44960	2011-05-03 00:00:00.000	7	6	4
8	Tamara Chand	44960	2011-05-03 00:00:00.000	8	6	4

Nhận xét:

Cả 3 hàm đều Order by **Order_Date**

Hàm ROW_NUMBER dùng để xếp hạng kết quả một cách tuần tự bắt đầu từ 1 và không quan tâm đến các giá trị giống nhau.

Tại dòng thứ 2 và dòng 3, giá trị Order_Date đều giống nhau. RANK và DENSE_RANK đều trả về giá trị là 2.

Tại dòng thứ 4, hàm RANK trực tiếp bắt kịp giá trị của ROW_NUMBER bằng cách bỏ đi giá trị 3. Còn DENSE_RANK thì trả về giá trị tiếp theo dựa trên dòng trước đó.

Tương tự với các dòng dưới

Chủ đề XẾP HẠNG (RANKING TOPIC)

HÀM ROW_NUMBER()

```
ROW_NUMBER ( ) OVER ( [ PARTITION BY Columns] ORDER BY Column ASC|DESC ) AS Row_num
```

- Lấy danh sách các đơn hàng (order_id) dựa vào ngày giao hàng (shipping_date) và cho biết đơn hàng đó là đơn hàng thứ mấy được ship đi.
- Kết quả hiển thị cần các cột: order_id, order_date, row_num (thứ tự ship hàng)

```
Select order_id, shipping_date,  
ROW_NUMBER ( ) OVER ( ORDER BY shipping_date ASC )  
AS Row_num  
from orders
```

	order_id	shipping_date	Row_num
1	28774	2009-01-02 00:00:00.000	1
2	37537	2009-01-02 00:00:00.000	2
3	44069	2009-01-02 00:00:00.000	3
4	13729	2009-01-03 00:00:00.000	4
5	9285	2009-01-04 00:00:00.000	5
6	9762	2009-01-04 00:00:00.000	6
7	7909	2009-01-04 00:00:00.000	7
8	44069	2009-01-04 00:00:00.000	8

Chủ đề XẾP HẠNG (RANKING TOPIC)

HÀM DENSE_RANK()

```
DENSE_RANK ( ) OVER ( [ PARTITION BY Columns ] ORDER BY Columns ASC|DESC ) AS dense_rank_col
```

- Xếp hạng giá trị lợi nhuận tăng dần của mỗi đơn hàng ở mỗi tỉnh
- Kết quả hiển thị gồm các cột: province, order_id, profit, dense_rank_col (Cột xếp hạng)

```
SELECT  
province, order_id, profit,  
DENSE_RANK() OVER ( PARTITION BY province ORDER BY profit ASC )  
AS dense_rank_col  
from orders
```

	province	order_id	profit	dense_rank_col
1	Alberta	55490	-38.2592	1
2	Alberta	9221	-37.0075	2
3	Alberta	54949	-33.9382	3
4	Alberta	52870	-21.0175	4
5	Alberta	774	-20.8208	5
6	Alberta	46533	-12.24	6
7	Alberta	38817	-2.9268	7
8	Alberta	27844	-2.632	8

HÀM RANK()

```
RANK ( ) OVER ( [ PARTITION BY Columns ] ORDER BY Columns ASC|DESC )
```

- Xếp hạng thứ bậc giá trị lợi nhuận tăng dần của mỗi đơn hàng ở mỗi tỉnh
- Kết quả hiển thị gồm các cột: province, order_id, profit, rank_col (Cột xếp hạng)

```
SELECT  
province, order_id, profit,  
RANK() OVER ( PARTITION BY province ORDER BY profit ASC )  
AS rank_col  
from orders
```

	province	order_id	profit	dense_rank_col
1	Alberta	55490	-38.2592	1
2	Alberta	9221	-37.0075	2
3	Alberta	54949	-33.9382	3
4	Alberta	52870	-21.0175	4
5	Alberta	774	-20.8208	5
6	Alberta	46533	-12.24	6
7	Alberta	38817	-2.9268	7
8	Alberta	27844	-2.632	8

Bài 1:

Lấy ra danh sách các đơn hàng (order_id) có lợi nhuận (profit) xếp hạng cao nhất (xếp hạng thứ nhất) của mỗi tỉnh

Gợi ý: Sử dụng dense_rank() để xếp hạng

Kết quả như ảnh:

	Results	Messages			
	province	order_id	profit	dense_rank_col	
1	Alberta	29187	16074.0755	1	
2	British Columbia	41895	16215.4636	1	
3	Manitoba	48800	16495.4628	1	
4	New Brunswick	29766	53764.8586	1	
5	Newfoundland	14727	3349.4	1	
6	Northwest Territories	52321	15193.4925	1	
7	Nova Scotia	2208	13259.2886	1	
8	Nunavut	1761	6813.3475	1	

Bài 2:

Lấy danh sách top 3 loại sản phẩm có tổng lợi nhuận cao nhất ở mỗi tỉnh

Gợi ý: Sử dụng row_number() để xếp hạng

Kết quả như ảnh:

	Results	Messages			
	province	product_category	total_profit	row_num	
1	Alberta	Technology	409896.6965	1	
2	Alberta	Furniture	171780.9896	2	
3	Alberta	Office Supplies	151299.7148	3	
4	British Columbia	Technology	357901.3043	1	
5	British Columbia	Furniture	205175.4446	2	
6	British Columbia	Office Supplies	199583.7439	3	
7	Manitoba	Technology	245513.2356	1	
8	Manitoba	Furniture	168960.2275	2	

RANK()

Xếp hạng theo thứ bậc tăng dần, các giá trị giống nhau có cùng thứ bậc, giá trị tiếp theo có thể nhảy n bậc nếu thứ bậc trước có n giá trị giống nhau

Value	Rank
1	1
2	2
2	2 Cùng giá trị, cùng bậc
3	4 Nhảy bậc
4	5

DENSE_RANK()

Xếp hạng theo thứ tự tăng dần, các giá trị giống nhau có thứ hạng bằng nhau.

Value	Dense_Rank
1	1
2	2
2	2 Cùng giá trị, cùng xếp hạng
3	3
4	4

ROW_NUMBER()

Xếp hạng theo thứ tự tăng dần không quan tâm đến các giá trị giống nhau

Value	Row_num
1	1
2	2
2	3
3	4
4	5

Bài tập về nhà:

Bài 1: Sử dụng phương pháp pivotTable

Hãy tính tổng giá trị (total_value) của mỗi tỉnh theo từng loại sản phẩm (product_category).
Kết quả được hiển thị như bảng:

	Results	Messages			
	province	Office Supplies	Furniture	Technology	
1	Nunavut	36891.075	41513.2716	34855.6708	
2	Prince Edward Island	91595.6212	173344.4506	144993.3991	
3	Saskatchewan	366594.1441	408644.8557	691354.6055	
4	Quebec	335460.0555	589755.483	557147.5132	
5	Newfoundland	22603.0039	53369.8747	26979.1774	
6	Ontario	882544.2788	1076605.9753	1063209.2815	
7	Manitoba	317919.543	487825.2497	528632.0013	
8	Northwest Territories	206605.2503	272101.3145	305231.6746	
9	Nova Scotia	170143.7353	330919.4921	308559.1083	
10	Alberta	305031.862	514189.6279	872452.519600001	
11	British Columbia	455750.8455	622719.3729	787608.3419	

Bài tập về nhà:

Bài 2: Sử dụng row_number ()
Lấy danh sách top 3 sản phẩm (product_name) có tổng lợi nhuận (total_profit) nhỏ nhất của mỗi loại sản phẩm (product_category).

Results

Messages

	product_category	product_name	total_profit	row_num
1	Furniture	DAX Wood Document Frame	52.795	1
2	Furniture	Eldon® Executive Woodline II Desk Accessories, Ma...	58.7349	2
3	Furniture	DAX Charcoal/Nickel-Tone Document Frame, 5 x 7	63.66	3
4	Office Supplies	Hoover® Commercial Lightweight Upright Vacuum	-175.4	1
5	Office Supplies	Hoover Portapower™ Portable Vacuum	-151.1776	2
6	Office Supplies	Bravo II™ Megaboss® 12-Amp Hard Body Upright, Re...	-134.37	3
7	Technology	Sony IBM Color Diskettes, 25/Pack	3.0494	1
8	Technology	Imation 3.5 IBM Diskettes, 10/Box	73.7886	2
9	Technology	DS/HD IBM Formatted Diskettes, 10/Pack - Staples	81.6726	3

Bài tập về nhà:

Bài 3: Sử dụng `dense_rank()`
Xác định những sản phẩm (`product_name`) có tổng lợi nhuận xếp hạng 3 (từ trên xuống) của mỗi tỉnh

<div><div><div></div></div>Results<div><div></div></div>Messages</div>				
	province	product_name	total_profit	denseRank
1	Alberta	Hewlett-Packard Business Color Inkjet 3000 [N, DTN]...	22616.0539	3
2	British Columbia	Canon imageCLASS 2200 Advanced Copier	16215.4636	3
3	Manitoba	Riverside Palais Royal Lawyers Bookcase, Royale Ch...	17891.7018	3
4	New Brunswick	Polycom ViewStation™ Adapter H323 Videoconferenc...	12776.942	3
5	Newfoundland	Bretford CR8500 Series Meeting Room Furniture	2864.585	3
6	Northwest Territories	SAFCO PlanMaster Heigh-Adjustable Drafting Table B...	20836.5165	3
7	Nova Scotia	Canon PC1060 Personal Laser Copier	11847.3404	3
8	Nunavut	Global Troy™ Executive Leather Low-Back Tilter	4633.114	3
9	Ontario	Bretford CR8500 Series Meeting Room Furniture	23108.8634	3
10	Prince Edward Island	Safco Contoured Stacking Chairs	6412.31	3
11	Quebec	Hewlett Packard LaserJet 3310 Copier	21838.6552	3

Bài tập về nhà:

Bài 4: Sử dụng rank ()

Xác định những sản phẩm (product_name) có tổng lợi nhuận xếp vị trí thứ 3 (từ dưới lên) của mỗi tỉnh

<div><div><div></div></div>Results<div><div></div></div>Messages</div>				
	province	product_name	total_profit	Rankk
1	Alberta	Hoover Portapower™ Portable Vacuum	-20.8208	3
2	British Columbia	Eldon Wave Desk Accessories	-2.0949	3
3	Manitoba	Array® Parchment Paper, Assorted Colors	-3.142	3
4	New Brunswick	Tennsco Commercial Shelving	-21.5756	3
5	Newfoundland	Xerox 1984	-0.7048000000000001	3
6	Northwest Territories	Imation 3.5 IBM Diskettes, 10/Box	1.162	3
7	Nova Scotia	Tennsco Lockers, Gray	-6.614800000000003	3
8	Nunavut	Eldon Base for stackable storage shelf, platinum	2.3823999999999996	3
9	Ontario	Rogers Handheld Barrel Pencil Sharpener	-1.3628	3
10	Prince Edward Island	Recycled Eldon Regeneration Jumbo File	-1.4636	3
11	Quebec	Hoover Portapower™ Portable Vacuum	-4.737599999999999	3