

SQL - GROUP BY & AGGREGATION

Diagram illustrating the aggregation process:

O-ID	Product	Pc.	Amt
1		stationery	10
2		vegetable	5
3	(+)	fruits	8
4		milk products	7
5		vegetables	20
6		fruits	15

Aggregation steps:

- Step 1: Sum of row 3 (fruits) = 8
- Step 2: Sum of rows 3 and 6 (fruits) = 8 + 15 = 23

Final result: $\underline{\underline{23}}$

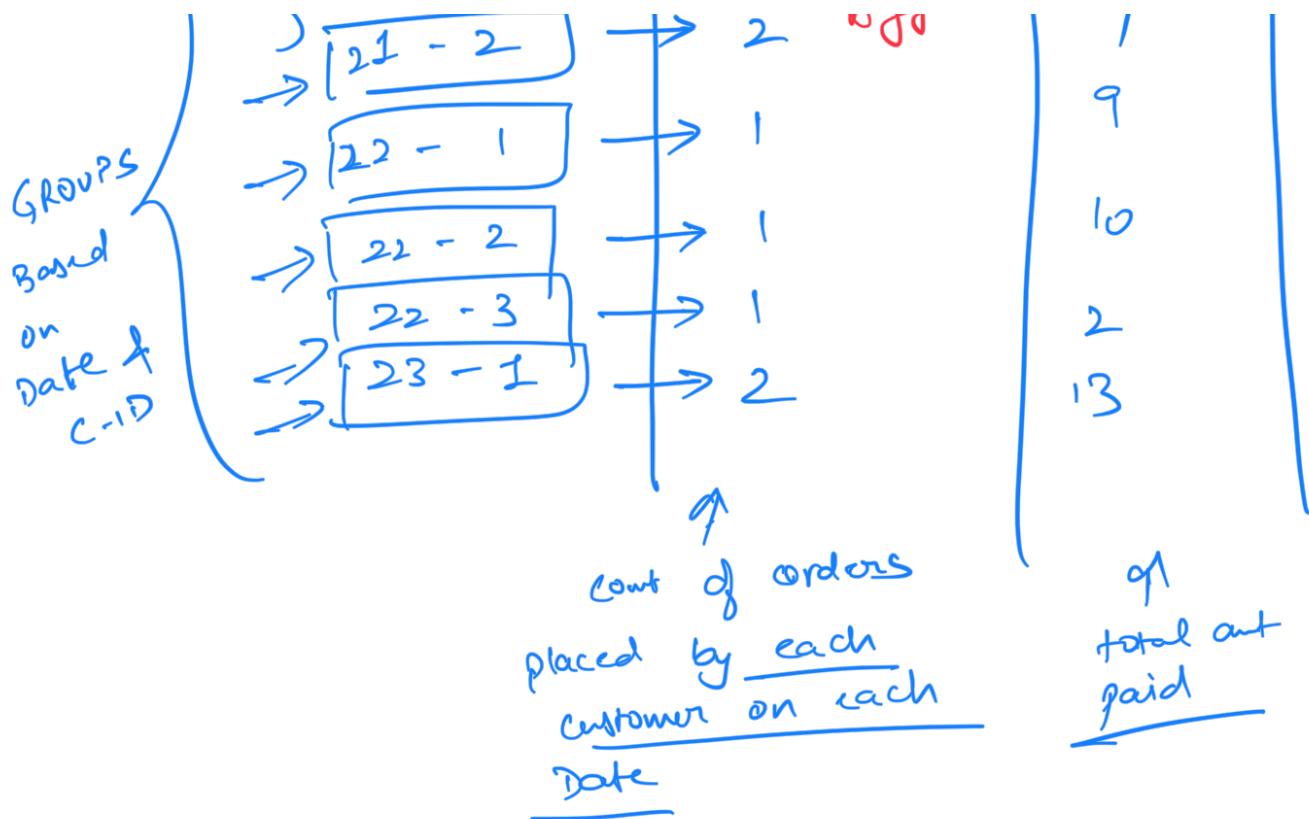
Customer-purchases

Date	C-ID	Amt
21	1	5
21	1	2
21	1	8
21	2	4
21	2	3
21	2	9
22	1	10
22	2	2
22	3	8
23	1	5

GROUP BY market-date, cust-id

Aggregation results:

- Count of Orders: 3
- Total Amt: $\frac{15}{\sum \text{Amt})}$



Syntax:

```

SELECT quantity
      , SUM(amount) AS total-amount
FROM   tbl-name
WHERE 
GROUP BY
HAVING
  
```

Aggregation functions

COUNT(*) → total no. of purchases

.....

SUM(C-1D)

Date	C-1D	Qty
21	1 ↗ 1 ↘	2
21	1 ↗ 1 ↘	3
21	1 ↗ 1 ↘	4
22	2 ↗ 2 ↘	2 + 3 ↗ 4 → 9
22	2 ↗ 2 ↘	3
22	2 ↗ 2 ↘	3
22	2 ↗ 2 ↘	3

→ [21] — 3

GROUP BY market_date

Date	C-1D
21	1 ↗ 1 ↘
22	2 ↗ 2 ↘

GROUP BY market_date,
customer_id

Date	C-1D
21	1 ↗ 1 ↘
22	1 ↗ 1 ↘
22	2 ↗ 2 ↘

Product id

Date	C-1D	Product id
21	1 ↗ 1 ↘	A ✓
21	1 ↗ 1 ↘	C
21	1 ↗ 1 ↘	A ✓

1	1
22	1
22	3
22	3
22	3
24	3

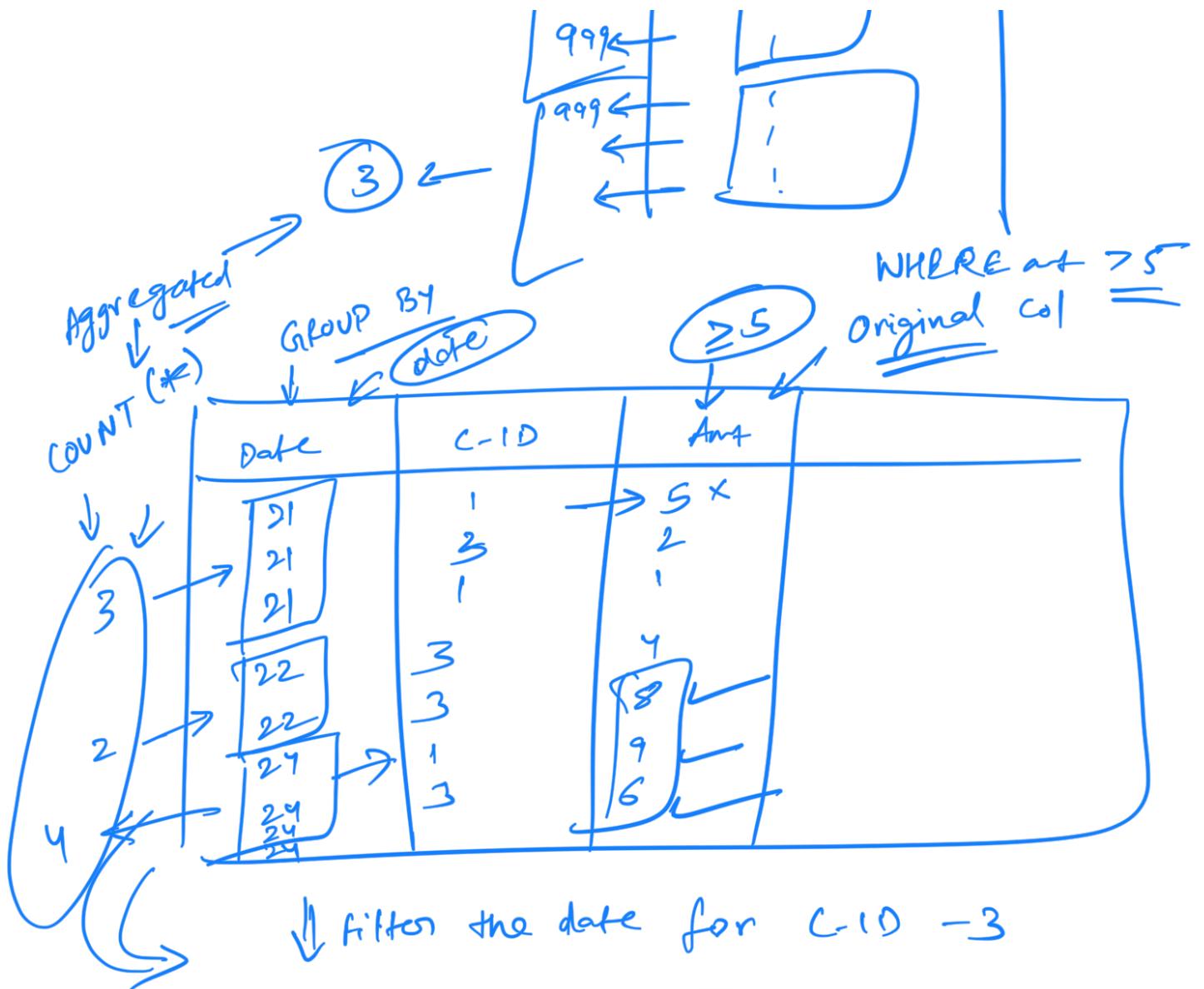
$\Rightarrow \text{COUNT}(\text{DISTINCT product-id})$

$\Rightarrow \text{GROUP BY market-date-, c-id}$

COUNT (*)	COUNT (?)	COUNT (col-name)	COUNT(DISTINCT col-name)
Counts all the values → including NULLs → commonly used → less confusing	← → includes NULL → same ↑	Counts duplicate but IGNORES NULLS ↑	Neither duplicates NOR NULLs ↑

X COUNT (?) → First column

COUNT('abc')	Dot
999 999 999	⋮



Date	C-ID	Qty	Cost	Amt
21	3	1	4	4
22	3	2	2	4
22	3	3	6	18
29	3	2	5	10

