Source datasets

S: Order

o_order_id	o_customer_sk	date
1	11	2024-01-01
2	12	2024-01-02
3	11	2024-01-03
...	...	...

R_1 Lineitem

order_id	li_product_id	quantity	price
1	21	2	20
1	22	1	25
2	23	3	13
3	21	1	10
...	...		...

R_2 Order returns

order_id	product_id	or_return_quantity
1	21	1
...	...	...

Factorized Learning

Join

T: Materialized data

o_order_id	o_customer_sk	date	li_product_id	quantity	price	or_return_quantity
1	11	2024-01-01	21	2	20	0
1	11	2024-01-01	22	1	25	0
2	12	2024-01-02	23	3	13	0
3	11	2024-01-03	21	1	10	1
...	...	...	...	...	...	...

Learning over Materialized data