Q1.

For each customer, product, month and state combination, compute (1) the customer's average sale of this product for the given month and state, (2) the customer’s average sale for the given month and state, but for all other products (3) the customer’s average sale for the given product and state, but for all other months and (4) the average sale of the product and the month but for all other states.

=>

with T1 as

(

select cust, prod, month, state, round(avg(quant),0) as Cust\_Avg

from sales

group by cust, prod,month, state

),

otherProd as

(

select T1.cust, T1.prod, T1.state,T1.month, T1.Cust\_Avg, round(avg(s.quant),0) as Other\_Prod\_Avg

from T1,sales s

where T1.cust = s.cust and T1.state = s.state and T1.month=s.month and T1.prod != s.prod

group by T1.cust, T1.prod,T1.month, T1.state, T1.Cust\_Avg

),

otherMonth as

(

select T1.cust, T1.prod,T1.month, T1.state, T1.Cust\_Avg, round(avg(s.quant),0) as Other\_Month\_Avg

from T1,sales s

where T1.cust = s.cust and T1.prod = s.prod and T1.state = s.state and T1.month!=s.month

group by T1.cust, T1.prod, T1.state, T1.cust\_avg,T1.month

),

otherstate as

(

select T1.cust, T1.prod,T1.month, T1.state, T1.cust\_avg, round(avg(s.quant),0) as other\_state\_avg

from T1,sales s

where T1.cust = s.cust and T1.prod = s.prod and T1.state != s.state and T1.month!=s.month

group by T1.cust, T1.prod, T1.state, T1.Cust\_Avg,t1.month

)

select T1.cust as customer, T1.prod as product,T1.month,T1.state, t1.Cust\_Avg, opa.Other\_Prod\_Avg,oma.Other\_Month\_Avg, osa.Other\_State\_Avg

from T1

full outer join otherProd opa on T1.cust=opa.cust and T1.prod=opa.prod and T1.state=opa.state and T1.month=opa.month

full outer join otherMonth oma on T1.cust=oma.cust and T1.prod=oma.prod and T1.state=oma.state and T1.month=oma.month

full outer join otherState osa on T1.cust=osa.cust and T1.prod=osa.prod and T1.state=osa.state and T1.month=osa.month

order by T1.cust,T1.prod,T1.month,T1.state

Q2. For customer, product and state, show the average sales before and after each quarter (e.g., for Q2, show average sales of Q1 and Q3. For “before” Q1 and “after” Q4, display . The “YEAR” attribute is not considered for this query – for example, both Q1 of 2017 and Q1 of 2018 are considered Q1 regardless of the year.

=>

with T1 as

(

select cust, prod, state, 1 as q1, Round(avg(quant)) as average

from sales

where month in (1,2,3)

group by cust, prod, state

),

T2 as

(

select cust, prod, state, 2 as q1, Round(avg(quant)) as average

from sales

where month in(4,5,6)

group by cust, prod, state

),

T3 as

(

select cust, prod, state, 3 as q1, Round(avg(quant)) as average

from sales

where month in (7,8,9)

group by cust, prod, state

),

T4 as

(

select cust, prod, state, 4 as q1, Round(avg(quant)) as average

from sales

where month in (10,11,12)

group by cust, prod, state

),

T5 as

(

select \* from T1 union select \* from T2 union select \* from T3 union select \* from T4

order by cust, prod, state, q1

),

T6 as

(

select \* from T1 union select \* from T2 union select \* from T3 union select \* from T4

order by cust, prod, state, q1

)

select t.cust, t.prod, t.state, t.q1, (select T6.average from T6 where T6.q1=t.q1-1 and T6.cust=t.cust

and T6.prod=t.prod and T6.state=t.state) as before\_avg,

(select T6.average from T6 where T6.q1=t.q1+1 and T6.cust=t.cust

and T6.prod=t.prod and T6.state=t.state) as after\_avg from T5 t

Q3. For each product, find the median sales quantity (assume an odd number of sales for simplicity of presentation). (NOTE – “median” is defined as “denoting or relating to a value or quantity lying at the midpoint of a frequency distribution of observed values or quantities, such that there is an equal probability of falling above or below it.” E.g., Median value of the list {13, 23, 12, 16, 15, 9, 29} is 15.

=>

with T1 as

(

select s.prod, s.quant,(select count(\*) from sales where s.quant>=quant and s.prod=prod) as count

from sales s

order by s.prod,count

),

T2 as

(

select T1.prod, count(T1.prod) as total\_count, count(T1.prod)/2 + 1 as count\_half

from T1

group by T1.prod order by prod

),

T3 as

(

select T1.prod, T1.quant from T1, T2 where T1.prod=T2.prod and T2.count\_half<=T1.count

order by T1.prod, T1.quant

)

select T3.prod as "PRODUCT", min(T3.QUANT) as "MEDIAN QUANT"

from T3

group by prod

Q4. For customer and product, find the month by which time, 75% of the sales quantities have been purchased. Again, for this query, the “YEAR” attribute is not considered. Another way to view this query is to pretend all 10,000 rows of sales data are from the same year.

=>

with base as

(

select cust, prod, month, sum(quant)

from sales

group by cust, prod, month

order by cust, prod, month

),

total as

(

select cust, prod, sum(sum)

from base

group by cust, prod

),

cumulative\_count as

(

select b1.cust, b1.prod, b1.month, b1.sum, sum(b2.sum) as cumulative\_sum

from base as b1 join base as b2

on b2.month <= b1.month and b1.cust = b2.cust and b1.prod = b2.prod

group by b1.cust, b1.prod, b1.month, b1.sum

order by cust, prod, month

),

pre\_final as

(

select r.cust, r.prod, r.month

from cumulative\_count as r, total as t

where r.cust = t.cust and r.prod = t.prod and r.cumulative\_sum >= (3.0/4.0 \* t.sum)

order by cust, prod, month

)

select cust "CUSTOMER", prod "PRODUCT", min(month) "75% PURCHASED BY MONTH"

from pre\_final

group by cust, prod

order by cust, prod