**\*\*\*Active User Retention [Facebook SQL Interview Question]**

**https://datalemur.com/questions/user-retention**

;with cte as (

SELECT user\_id,COUNT(DISTINCT date\_part('month',event\_date)) as monthly\_active\_users

,MAX(date\_part('month',event\_date)) as mth

FROM user\_actions

where event\_date >='06/01/2022' and event\_date<='07/31/2022'

group by user\_id)

select mth,count(\*) from cte

where monthly\_active\_users>=2

group by mth

**\*\*\*Y-on-Y Growth Rate [Wayfair SQL Interview Question]**

**https://datalemur.com/questions/yoy-growth-rate**

with cte as (

SELECT

date\_part('year',transaction\_date) as year,

product\_id,spend as curr\_year\_spend

,lag(spend)over(partition by product\_id order by date\_part('year',transaction\_date))

as prev\_year\_spend

FROM user\_transactions)

select \* ,

ROUND((curr\_year\_spend-prev\_year\_spend)\*100.0/prev\_year\_spend,2) as yoy\_rate

from cte

**\*\*\*Maximize Prime Item Inventory [Amazon SQL Interview Question]**

**https://datalemur.com/questions/prime-warehouse-storage**

with cte as (

SELECT item\_type

,count(\*) as batch\_size,

SUM(square\_footage) as area\_per\_item\_type

FROM inventory

group by item\_type)

,prime\_item as (

select item\_type,

area\_per\_item\_type,

floor(500000/area\_per\_item\_type) as prime\_item\_batch\_count,

floor(500000/area\_per\_item\_type)\*batch\_size as prime\_item\_count

from cte where item\_type='prime\_eligible')

,non\_prime as (

select item\_type,

area\_per\_item\_type,

floor((500000-(select area\_per\_item\_type \*prime\_item\_batch\_count from prime\_item))/area\_per\_item\_type) as non\_prime\_item\_batch\_count,

floor((500000-(select area\_per\_item\_type \*prime\_item\_batch\_count from prime\_item))/area\_per\_item\_type)\*batch\_size as non\_prime\_item\_count

from cte where item\_type='not\_prime'

)

select item\_type,non\_prime\_item\_count as item\_count from (

select \* from non\_prime

union all

select \* from prime\_item)a

order by item\_count desc

**\*\*\*Median Google Search Frequency [Google SQL Interview Question]**

**https://datalemur.com/questions/median-search-freq**

with cte as (

SELECT searches

FROM search\_frequency

GROUP BY

searches,GENERATE\_SERIES(1, num\_users))

select ROUND(PERCENTILE\_CONT(.50) WITHIN GROUP (order by searches ):: decimal,1)

as median

from cte

**\*\*\*Advertiser Status [Facebook SQL Interview Question]**

**https://datalemur.com/questions/updated-status**

with cte as (

SELECT COALESCE(a.user\_id,dp.user\_id) as user\_id,a.status

,dp.paid as payment

from

advertiser a

full outer join daily\_pay dp on a.user\_id=dp.user\_id)

,cte2 as (

select \*,

case when payment is null then 'Not Paid' else 'Paid' end as paymentt

from cte )

select user\_id,new\_status from (

select \*,

case when paymentt ='Not Paid' then 'CHURN'

when paymentt='Paid' and status='CHURN' then 'RESURRECT'

when status is null then 'NEW'

else 'EXISTING'

end as new\_status

from cte2)a

order by a.user\_id

**\*\*\*Consecutive Filing Years [Intuit SQL Interview Question]**

**https://datalemur.com/questions/consecutive-filing-years**

with cte as (

select \* from filed\_taxes

where user\_id in (

select user\_id

from filed\_taxes

group by user\_id having count(1)>=3

)

),final as (

select \*,

lag(product) over (partition by user\_id order by filing\_date) as prev\_prod,

lead(product) over (partition by user\_id order by filing\_date) as next\_prod

from cte )

select DISTINCT user\_id from (

select \*,

Case when left(product,8)=left(prev\_prod,8) and left(product,8)=left(next\_prod,8) then 1 else 0 end as flag

from final)a where a.flag=1

**\*\*\*Marketing Touch Streak [Snowflake SQL Interview Question]**

**https://datalemur.com/questions/marketing-touch-streak**

;with cte as (

select \*

from marketing\_touches

where contact\_id in

(select contact\_id from marketing\_touches group by contact\_id HAVING

count(1)>=3)

),final as (

select \*, date\_trunc('week',event\_date) as start\_of\_week

from cte

where contact\_id IN

(select DISTINCT contact\_id from cte where event\_type='trial\_request'))

,final2 as (

select \*,

lead(start\_of\_week,1)over(PARTITION BY contact\_id order by start\_of\_week) as next\_week,

lag(start\_of\_week,1)over(PARTITION BY contact\_id order by start\_of\_week) as prev\_week

from final)

select cc.email as email

from final2 ff

inner join crm\_contacts cc on ff.contact\_id=cc.contact\_id

where

DATE\_PART('week', start\_of\_week::date)-DATE\_PART('week', prev\_week::date)=1

and

DATE\_PART('week', next\_week::date)-DATE\_PART('week', start\_of\_week::date)=1

**\*\*\*3-Topping Pizzas [McKinsey SQL Interview Question]**

**https://datalemur.com/questions/pizzas-topping-cost**

;with master as (

SELECT \*

FROM pizza\_toppings)

select

m1.topping\_name || ',' || m2.topping\_name||','|| m3.topping\_name as pizza,

m1.ingredient\_cost+m2.ingredient\_cost+m3.ingredient\_cost as total\_cost

from master m1 inner join master m2 on m1.topping\_name < m2.topping\_name

inner join master m3 on m2.topping\_name < m3.topping\_name

order by total\_cost desc,pizza

**\*\*\*Compressed Median [Alibaba SQL Interview Question]**

**https://datalemur.com/questions/alibaba-compressed-median**

WITH summary AS (

SELECT item\_count

FROM items\_per\_order

GROUP BY

item\_count,

GENERATE\_SERIES(1, order\_occurrences)

)

SELECT

ROUND(

PERCENTILE\_CONT(0.50) WITHIN GROUP (

ORDER BY item\_count)::DECIMAL, 1) AS median

FROM summary;

**\*\*\*Average Vacant Days [Airbnb SQL Interview Question]**

**https://datalemur.com/questions/average-vacant-days**

;with cte as (

select l.listing\_id as lid,b.\*,

case when checkin\_date < '1/1/2021 00:00:00' then '1/1/2021 00:00:00'

else checkin\_date end as correct\_checkin,

case when checkout\_date > '12/31/2021 00:00:00' then '12/31/2021 00:00:00'

else checkout\_date end as correct\_checkout

from listings l

LEFT JOIN bookings b on l.listing\_id=b.listing\_id

where l.is\_active=1)

,final as (

select lid, cast((correct\_checkout-correct\_checkin)as int) as number\_of\_days

from cte)

select ROUND(SUM(total\_vacant\_days)\*1.0/COUNT(DISTINCT lid) ) as avg\_vacant\_days

from (

select lid,365-COALESCE(SUM(number\_of\_days),0) as total\_vacant\_days

from final

group by lid

)a

**\*\*\*Patient Support Analysis (Part 3) [UnitedHealth SQL Interview Question]**

**https://datalemur.com/questions/patient-call-history**

WITH call\_history AS (

SELECT

policy\_holder\_id,

call\_date AS current\_call, -- Remove this column

LAG(call\_date) OVER (

PARTITION BY policy\_holder\_id ORDER BY call\_date) AS previous\_call, -- Remove this column

ROUND(EXTRACT(EPOCH FROM call\_date

- LAG(call\_date) OVER (

PARTITION BY policy\_holder\_id ORDER BY call\_date)

)/(24\*60\*60),2) AS time\_diff\_days

FROM callers

)

SELECT COUNT(DISTINCT policy\_holder\_id) AS policy\_holder\_count

FROM call\_history

WHERE time\_diff\_days <= 7;

**\*\*\*Patient Support Analysis (Part 4) [UnitedHealth SQL Interview Question]**

**https://datalemur.com/questions/long-calls-growth**

;with cte as (

select EXTRACT(year from call\_date) as yr,

EXTRACT(month from call\_date) as mnth,

count(case\_id) as call\_duration\_secs,

lag(count(case\_id))OVER

(order by EXTRACT(year from call\_date),EXTRACT(month from call\_date)) as prev\_month\_duration

from callers

where call\_duration\_secs >=300

group by yr,mnth)

select yr,mnth as mth,

ROUND((call\_duration\_secs-prev\_month\_duration)\*100.0/prev\_month\_duration,1) as long\_calls\_growth\_pct

from cte

**\*\*\*Same Week Purchases [Etsy SQL Interview Question]**

**https://datalemur.com/questions/same-week-purchases**

;with cte as (

select s.user\_id,s.signup\_date,a.purchase\_date

from signups s

left join (

select \*

,row\_number()over(PARTITION BY user\_id order by purchase\_date) as rn

from user\_purchases )a on s.user\_id=a.user\_id and a.rn=1)

select ROUND(SUM(number\_of\_days)\*100.0/count(1),2) as single\_purchase\_pct

from (

select \*,

case when

EXTRACT(EPOCH FROM (purchase\_date - signup\_date))/(24\*60\*60) <=7 then 1 else 0 END

as number\_of\_days

from cte )a

**\*\*\*Follow-Up Airpod Percentage [Apple SQL Interview Question]**

**https://datalemur.com/questions/follow-up-airpod-percentage**

;with cte as (

select \*,

lag(product\_name)over(PARTITION BY customer\_id order by transaction\_id)

as prev\_item\_bought,

case

when product\_name='AirPods' and

lag(product\_name)over(PARTITION BY customer\_id order by transaction\_id)='iPhone'

then 1 else 0 end as flag

from transactions )

select ROUND(SUM(flag)\*100.0/count(DISTINCT customer\_id)) as round

from cte