Q1. Can you use this data to generate monthly cohorts using SQL? A monthly cohort table would have these results:

|  |  |  |  |
| --- | --- | --- | --- |
| Cohort Month | Month Number | Cohort Size | Retained users |
|  |  |  |  |
|  |  |  |  |

Where

* *Cohort month:* is a month for which we are looking at retention
* *Month number:*  would be the number of months elapsed the since the first purchase
* *Cohort size:* is the number of users who came in a particular cohort month
* *Retained users* is the number of the users who have ordered in the *month\_number*

*This box can’t be generated as we don’t have repeat user details in data set and no date column which indicates first purchase date. Only Cohort month , cohort size can be calculated, for rest repeat purchase details in terms or order or date details required.*

Dummy SQL will be likewise below:

Assuming there is repeat purchase by same customer id and no first purchase date column present.

Logic –

1. Rank order by order date patriation by customer id
2. Extract month, year from order date
3. Count customerid where rank is greater than 1 group by month,year,

select

distinct cohort, count(distinct cid)Cohort\_Size , count(cust\_type) Retained\_users #case when cust\_type = 'retained' then count(cust\_type) else 0 end as

,case when cust\_type= 'first' then 0 else month\_number end as month\_number

from

(select \*,

case when rank> 1 then 'retained' when rank = 1 then 'first' else 'others' end as cust\_type,

LEAD(cohort) OVER (PARTITION by cid order by cid,cohort)

from

(SELECT \*, EXTRACT(month FROM date) cohort , date\_diff (date ,'2022-01-01',month) as month\_number,RANK() OVER ( PARTITION BY cid ORDER BY date ) AS rank

FROM `mbaa20006-ayushjha.Testing\_learning.tb`))

group by 1,cust\_type,month\_number

having cust\_type = 'retained'

order by 1,2,3

Q2. If we plot Retained users against month number for each cohort month, what is the shape of the curve you would expect?

If we plot the graph it would be Declining Curve as number of users gets reduced in any business in retentions over the time period. Decline may differ as slow or faster depending upon the business.