DROP TABLE IF EXISTS intermediate,

cust\_loyalty,

customer\_session;

--temp table for calculate intermediate

CREATE TEMP TABLE intermediate AS

--use WITH to creat a temp table within another temp table

WITH temptable as (SELECT customer\_id, count(customer\_id) as order\_times

FROM public.order

GROUP BY customer\_id)

SELECT COALESCE(t1.order\_times, 0) AS orders,

t2.customer\_id,

t2.sex,

t2.start\_year,

t2.last\_order\_date,

(CASE WHEN t2.last\_order\_date < date'2018-01-01' OR t2.last\_order\_date IS NULL THEN '2018-01-01' ELSE t2.last\_order\_date END) - date'2021-01-01' AS intermediate --intermediate do not have meaning just for calculating customer\_loyalty later

FROM customer t2

LEFT JOIN temptable t1 ON t1.customer\_id = t2.customer\_id;

-- This temp table for calculating customer\_loyalty

CREATE TEMP TABLE cust\_loyalty AS

SELECT \*,

50\*orders + intermediate AS customer\_loyalty

FROM intermediate

ORDER BY customer\_loyalty;

--This temp table for calculation session info and conversation rate.

CREATE TEMP TABLE customer\_session AS

SELECT t2.session\_id,

t1.customer\_id,

COUNT(CASE WHEN t1.article\_event\_type='40' THEN 1 ELSE NULL END) AS Sales,

COUNT(CASE WHEN t1.article\_event\_type='10' THEN 1 ELSE NULL END) AS ProductViews,

CAST(100.0 \* COUNT(CASE WHEN article\_event\_type='40' THEN 1 ELSE NULL END) / NULLIF(COUNT(CASE

WHEN article\_event\_type='10' THEN 1 ELSE NULL END),0) AS DECIMAL(6,2)) AS Conversion\_rate,

t2.session\_date

FROM PUBLIC.session t2

LEFT JOIN articleevents t1 ON t2.session\_id = t1.session\_id

GROUP BY t2.session\_id,

t1.customer\_id,

t2.session\_date;

---this is the final table, combined session info with customer info

SELECT c2.customer\_id,

c1.session\_id,

c1.sales,

c1.productviews,

c1.conversion\_rate,

c1.session\_date,

c2.sex,

COALESCE(c2.orders, 0) AS orders,

c2.start\_year,

COALESCE(c2.last\_order\_date, date'2018-01-01') AS last\_order\_date,

c2.intermediate,

c2.customer\_loyalty

FROM customer\_session c1

LEFT JOIN cust\_loyalty c2 ON c1.customer\_id = c2.customer\_id

ORDER BY customer\_loyalty ASC

**--Sanity check**

**SELECT \***

**FROM PUBLIC.SESSION -- It has 126375 rows;**

**--Check activity of customer ID "653FDF87C3848373C7CEC6EBA7FE270C"**

**SELECT \***

**FROM articleevents**

**WHERE session\_id = '653FDF87C3848373C7CEC6EBA7FE270C'**

--Test for missing data customers(replacing the last line:

--Type 1:

WHERE conversion\_rate IS null AND sex IS null AND productviews=0

--Type 2:

WHERE conversion\_rate is null AND sex IS NOT null ---6095records

WHERE conversion\_rate is null AND sex IS NOT null AND sales!=0 --1748 records

WHERE conversion\_rate is null AND sex IS NOT null AND sales=0 ---4347 records

--Type 3:

WHERE conversion\_rate IS not null AND sex IS null