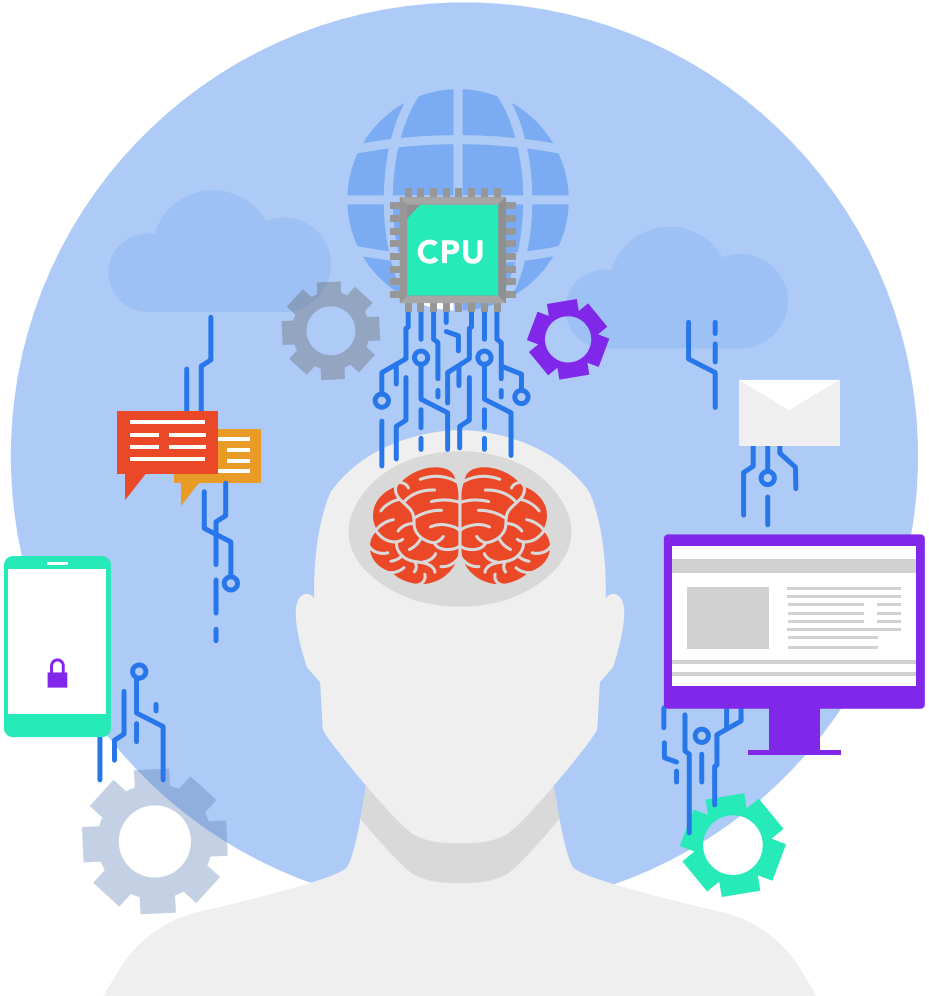


SQL Retail Analysis

Fauzia Y. Ayupuspita

Question 1 & Table Result

Question 1

Write an SQL statement to count the number of users per country

```sql
/*
    count the number of users per country
*/
SELECT country,
       COUNT(userid) number_of_user
FROM user_tab_Users_Sales
GROUP BY country
```

	country	number_of_user
1	ID	1,531
2	MY	1,528
3	PH	1,583
4	SG	1,562
5	TH	1,507
6	TW	1,539
7	VN	1,605

Question 2 & Table Result

Question 2

Write an SQL statement to count the number of orders per country

```sql
/*
    count the number of orders per country
*/
select u.country,
       count(distinct o.orderid) number_of_order
from user_tab_Users_Sales u
left join order_tab_Users_Sales o
on u.userid = o.userid
group by 1
order by 1 desc
```

country	number_of_order
VN	4,367
TW	4,403
TH	4,169
SG	4,167
PH	4,350
MY	4,173
ID	4,412

Question 3 & Table Result

Question 3

Write an SQL statement to find the first order date of each user

```sql
/*
    find the first order date of each user
*/
select userid,
       min(order_time) first_order_time
from order_tab_Users_Sales
group by 1
order by 2
```

	123 userid	first_order_time
1	10,565	01/01/2017
2	11,278	01/01/2017
3	11,367	01/01/2017
4	11,501	01/01/2017
5	11,526	01/01/2017
6	11,534	01/01/2017
7	11,566	01/01/2017
8	11,846	01/01/2017
9	12,311	01/01/2017
10	13,131	01/01/2017
11	13,599	01/01/2017
12	13,652	01/01/2017
13	13,747	01/01/2017
14	13,828	01/01/2017
15	14,462	01/01/2017
16	14,657	01/01/2017
17	14,799	01/01/2017
18	15,282	01/01/2017
19	15,608	01/01/2017
20	15,618	01/01/2017

Question 4 & Table Result

Question 4

Write an SQL statement to find the number of users who made their first order in each country, each day

```sql
/*
    find the number of users who made their first order
    in each country and each day
*/
select o.first_order_time,
       u.country,
       count(distinct o.userid) total_users
from (
    select userid, min(order_time) first_order_time
    from order_tab_Users_Sales
    group by 1
    ) o
left join user_tab_Users_Sales u
on o.userid = u.userid
group by 1,2
order by 1,2
```

	first_order_time	country	total_users
1	01/01/2017	ID	37
2	01/01/2017	MY	25
3	01/01/2017	PH	34
4	01/01/2017	SG	21
5	01/01/2017	TH	34
6	01/01/2017	TW	30
7	01/01/2017	VN	37
8	01/02/2017	ID	30
9	01/02/2017	MY	26
10	01/02/2017	PH	23
11	01/02/2017	SG	23
12	01/02/2017	TH	28
13	01/02/2017	TW	23
14	01/02/2017	VN	21
15	01/03/2017	ID	30
16	01/03/2017	MY	31
17	01/03/2017	PH	25
18	01/03/2017	SG	34
19	01/03/2017	TH	24
20	01/03/2017	TW	26

Question 5 & Table Result

Question 5

Write an SQL statement to find the first order GMV of each user. If there is a tie, use the order with the lower orderid

```sql
/*
    find the first order GMV of each user.
    If there is a tie, use the order with the lower orderid
*/
select userid,
       gmv
from (
    select userid, gmv,
    min(order_time) first_order_time,
    min(orderid) first_order_id
    from order_tab_Users_Sales
    group by 1,2
    ) o
```

	123 userid	ABC gmv
1	10,310	10,0
2	10,310	21,0
3	10,310	56,0
4	10,313	21,0
5	10,313	45,0
6	10,313	85,0
7	10,313	95,0
8	10,323	5,0
9	10,333	66,0
10	10,333	90,0
11	10,333	91,0
12	10,341	10,0
13	10,341	37,0
14	10,341	41,0
15	10,341	51,0
16	10,341	52,0
17	10,341	83,0
18	10,347	30,0
19	10,347	45,0
20	10,347	68,0

Question 6 & Table Result

Question 6

Write an SQL statement to find the total GMV per country

```
/*
    count total gmv per country
*/
select u.country,
count(o.gmv) total_gmv
from user_tab_Users_Sales u
left join order_tab_Users_Sales o
on u.userid = o.userid
group by 1
order by 2 desc
```

	country	total_gmv
1	ID	4,412
2	TW	4,403
3	VN	4,367
4	PH	4,350
5	MY	4,173
6	TH	4,169
7	SG	4,167

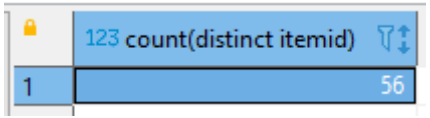
Question 7 & Table Result

Question 7

Find out what is wrong with the sample data

```sql
/*
    count the number of item have 1 price tag
*/
select count(distinct itemid)
from (
    select itemid
    from (
        select itemid, gmv
        from order_tab_Users_Sales
        group by 1,2
    ) o
    group by 1
    having count(gmv) > 1
) o

*/ There are 56 items have more than 1 price tag */
```



	123 count(distinct itemid)
1	56

Question 7 & Table Result

What is wrong with the sample data

There are many users buy before register

```sql
/*
    find the user buy before register
*/
select u.userid
from user_tab_Users_Sales u
left join order_tab_Users_Sales o
on u.userid = o.userid
where o.order_time < u.register_time

*/ There are many users buy before register */
```

	123 userid
1	10,313
2	10,313
3	10,313
4	10,323
5	10,333
6	10,341
7	10,341
8	10,353
9	10,360
10	10,366
11	10,376
12	10,384
13	10,389
14	10,389
15	10,389
16	10,391
17	10,409
18	10,409
19	10,409
20	10,409

Question 7 & Table Result

What is wrong with the sample data

There are many users buy before register

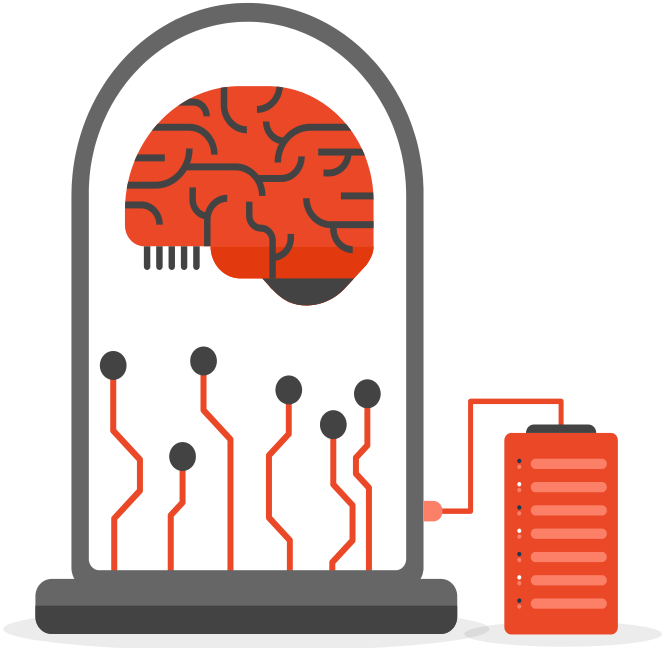
```sql
/*
    count the user buy before register
*/
select count(distinct o.userid)
from user_tab_Users_Sales u
right join order_tab_Users_Sales o
on u.userid = o.userid

*/ There are 10.181 users exsist in order tab but did not in user tab */
```

	123 count(distinct o.userid)
1	10,181

Conclution

- ID is the highest GMV given, because ID is the first country to open branches
- There are several problems we should fix, such:
  1. Many items have more than 1 price tag
  2. many users buy before register



THANK YOU