SQL queries

A/B test experiment to improve homepage.

--Can a user show up more than once in the activity table? Yes or no, and why?

-- Yes - they can make multiple purchases

```
select   uid,
  count(*)
from    activity
group by uid
order by count desc;
```

--**What type of join should we use to join the users table to the activity table?**

--LEFT JOIN

--**What SQL function can we use to fill in NULL values?**

--COALESCE()

--**What are the start and end dates of the experiment?**

-- Start: 2023-01-25. End: 2023-02-06

```
select min(join_dt) as min, max(join_dt) as max from groups
```

--**How many total users were in the experiment?**

--48943

```
select count(distinct uid) from groups;
```

--How many users were in the control and treatment groups?

```
select "group", count(*) from groups group by "group";
```

--**What was the conversion rate of all users?**

```
with cte_users_with_purchases as
(
  select
        u.id as uid,
        a.dt as dt,
        coalesce(a.spent, 0.0) as spent
        from
        users u
        left join
        activity a
        on
        u.id = a.uid
 ),
 cte_user_purchases_agg as
 (
   select
    uid,
    sum(spent) as total_spend,
    case when sum(spent) > 0 then 1 else 0 end as converted
```

```
 from cte_users_with_purchases
 group by uid
)
 select round(avg(converted)*100,2) from cte_user_purchases_agg


--What is the user conversion rate for the control and treatment groups?
with cte_users_with_purchases as
(
 select
        u.id as uid,
        a.dt as dt,
        coalesce(a.spent, 0.0) as spent,
        g.group as ab_group
        from
        users u
        left join
        activity a
        on
        u.id = a.uid
        left join
        groups g
 on
        u.id = g.uid
```

```
),
cte_user_purchases_agg as
(
  select
   uid,
        ab_group,
   sum(spent) as total_spend,
   case when sum(spent) > 0 then 1 else 0 end as converted
 from cte_users_with_purchases
 group by uid, ab_group
)
select
        ab_group,
 round(avg(converted)*100,2)
from
        cte_user_purchases_agg
group by
        ab_group
```

--What is the average amount spent per user for the control and treatment groups, including users who did not convert?

```
with cte_users_with_purchases as
(
  select
```

```
        u.id as uid,
        a.dt as dt,
        coalesce(a.spent, 0.0) as spent,
        g.group as ab_group
        from
        users u
        left join
        activity a
        on
        u.id = a.uid
        left join
        groups g
 on
        u.id = g.uid
),
cte_user_purchases_agg as
(
  select
   uid,
        ab_group,
   sum(spent) as total_spend,
   case when sum(spent) > 0 then 1 else 0 end as converted
 from cte_users_with_purchases
```

```
 group by uid, ab_group
)
 select
        ab_group,
 round(avg(converted)*100,2) as conversion_rate,
 round(avg(total_spend), 2) as avg_spend
from
        cte_user_purchases_agg
group by
        ab_group
```

--Why does it matter to include users who did not convert when calculating the average amount spent per user?

-- Because these users matter too for our business