NODE PROCESSING TIME

TEMP TABLE | CTE | AGGREGATION | PARTITION

TASK 1

For each node extract the earliest scan time for ENTRY and DELIVERY, but the latest for OUTBOUND

SQL | Saved to variable `df_1`

```
DROP TABLE IF EXISTS exp_table;
CREATE TEMP TABLE exp_table(
    node varchar(4)
    ,scan text
    ,event_date_utc timestamp
);

INSERT INTO exp_table
VALUES
 ('ABC1',103,'2023-01-10 10:00')
,('ABC1',201,'2023-01-10 14:00')
,('ABC1',202,'2023-01-10 18:00')
,('ABC1',254,'2023-01-10 20:00')
,('ABC1',301,'2023-01-11 13:00')
,('ABC1',302,'2023-01-11 15:00')
,('XYZ2',103,'2023-01-10 11:00')
,('XYZ2',201,'2023-01-10 13:00')
,('XYZ2',202,'2023-01-10 16:00')
,('XYZ2',254,'2023-01-10 19:00')
,('XYZ2',301,'2023-01-11 12:00')
,('XYZ2',302,'2023-01-11 13:00');


WITH base_table AS (
SELECT
node
,CASE
    WHEN scan IN ('103','201') THEN 'entry'
    WHEN scan IN ('202','254') THEN 'outbound'
    WHEN scan IN ('301','302') THEN 'delivery'
END AS scan_type
,MIN(event_date_utc) as min_date
,MAX(event_date_utc) as max_date
FROM exp_table
GROUP BY 1,2
)

SELECT
node
,scan_type
,CASE WHEN scan_type IN ('outbound') THEN max_date ELSE min_date END AS check_time
FROM base_table
ORDER BY node,check_time
```

	node object	scan_type object	check_time dateti...
0	ABC1	entry	2023-01-10 10:00:00
1	ABC1	outbound	2023-01-10 20:00:00
2	ABC1	delivery	2023-01-11 13:00:00
3	XYZ2	entry	2023-01-10 11:00:00
4	XYZ2	outbound	2023-01-10 19:00:00
5	XYZ2	delivery	2023-01-11 12:00:00

TASK 2

Which node had the fastest E2E process time?

SQL | Saved to variable `df_2`

```
DROP TABLE IF EXISTS exp_table;
CREATE TEMP TABLE exp_table(
    node varchar(4)
    ,scan text
    ,event_date_utc timestamp
);

INSERT INTO exp_table
VALUES
 ('ABC1',103,'2023-01-10 10:00')
,('ABC1',201,'2023-01-10 14:00')
,('ABC1',202,'2023-01-10 18:00')
,('ABC1',254,'2023-01-10 20:00')
,('ABC1',301,'2023-01-11 13:00')
,('ABC1',302,'2023-01-11 15:00')
,('XYZ2',103,'2023-01-10 11:00')
,('XYZ2',201,'2023-01-10 13:00')
,('XYZ2',202,'2023-01-10 16:00')
,('XYZ2',254,'2023-01-10 19:00')
,('XYZ2',301,'2023-01-11 12:00')
,('XYZ2',302,'2023-01-11 13:00');

WITH base_table AS (
SELECT
node
,CASE
    WHEN scan IN ('103','201') THEN 'entry'
    WHEN scan IN ('202','254') THEN 'outbound'
    WHEN scan IN ('301','302') THEN 'delivery'
END AS scan_type

,MIN(event_date_utc) as min_date
,MAX(event_date_utc) as max_date
FROM exp_table
GROUP BY 1,2
)

SELECT
node
,scan_type
,CASE WHEN scan_type IN ('outbound') THEN max_date ELSE min_date END AS check_time
,MAX(CASE WHEN scan_type IN ('outbound') THEN max_date ELSE min_date END) OVER (PARTITION BY node )
- MIN(CASE WHEN scan_type IN ('outbound') THEN max_date ELSE min_date END) OVER (PARTITION BY node )
 as e2e_time
FROM base_table
ORDER BY e2e_time,check_time
```

	node object	scan_type object	check_time dateti...	e2e_time timedelt...	
0	XYZ2	entry	2023-01-10 11:00:00	1 days 01:00:00	
1	XYZ2	outbound	2023-01-10 19:00:00	1 days 01:00:00	
2	XYZ2	delivery	2023-01-11 12:00:00	1 days 01:00:00	
3	ABC1	entry	2023-01-10 10:00:00	1 days 03:00:00	
4	ABC1	outbound	2023-01-10 20:00:00	1 days 03:00:00	
5	ABC1	delivery	2023-01-11 13:00:00	1 days 03:00:00	

TASK 2 *Alternative*

Which node had the fastest E2E process time?

SQL | Saved to variable `df_3`

```
DROP TABLE IF EXISTS exp_table;
CREATE TEMP TABLE exp_table(
    node varchar(4)
    ,scan text
    ,event_date_utc timestamp
);

INSERT INTO exp_table
VALUES
 ('ABC1',103,'2023-01-10 10:00')
,('ABC1',201,'2023-01-10 14:00')
,('ABC1',202,'2023-01-10 18:00')
,('ABC1',254,'2023-01-10 20:00')
,('ABC1',301,'2023-01-11 13:00')
,('ABC1',302,'2023-01-11 15:00')
,('XYZ2',103,'2023-01-10 11:00')
,('XYZ2',201,'2023-01-10 13:00')
,('XYZ2',202,'2023-01-10 16:00')
,('XYZ2',254,'2023-01-10 19:00')
,('XYZ2',301,'2023-01-11 12:00')
,('XYZ2',302,'2023-01-11 13:00');

WITH support_table AS (
SELECT
node
,CASE
    WHEN scan IN ('103','201') THEN 'entry'
    WHEN scan IN ('202','254') THEN 'outbound'
    WHEN scan IN ('301','302') THEN 'delivery'
END AS scan_type

,MIN(event_date_utc) as min_date
,MAX(event_date_utc) as max_date
FROM exp_table
GROUP BY 1,2
),

base_table AS (
SELECT
node
,scan_type
,CASE WHEN scan_type IN ('outbound') THEN max_date ELSE min_date END AS check_time
,ROW_NUMBER() OVER (PARTITION BY node ORDER BY CASE WHEN scan_type IN ('outbound') THEN max_date ELSE min_date END) as leg_numb
FROM support_table
)

SELECT
node
,MAX(CASE WHEN leg_number = 2 THEN check_time END) - MAX(CASE WHEN leg_number = 1 THEN check_time END) as leg_1
,MAX(CASE WHEN leg_number = 3 THEN check_time END) - MAX(CASE WHEN leg_number = 2 THEN check_time END) as leg_2
,MAX(CASE WHEN leg_number = 3 THEN check_time END) - MAX(CASE WHEN leg_number = 1 THEN check_time END) as e2e_time
FROM base_table
GROUP BY 1
ORDER BY e2e_time
```

	node object	leg_1 timedelta64[...	leg_2 timedelta64[...	e2e_time timedelt...	
0	XYZ2	0 days 08:00:00	0 days 17:00:00	1 days 01:00:00	
1	ABC1	0 days 10:00:00	0 days 17:00:00	1 days 03:00:00	