

Operation Analytic and Investigating Metric Spike

Description:

Operation Analytic is the analysis done for the complete end to end operations of a company. With the help of this, the company then finds the areas on which it must improve upon. You work closely with the ops team, support team, marketing team, etc and help them derive insights out of the data they collect.

Technology used: I have used MySQL for executing my Queries as I am already using it for a long time.

Approach: firstly I have created the database in MySQL through which I am going to perform some operations on operation analytic and investigating metric spike and understanding each problem I performed the queries to get insights.

Task 1: Calculate the number of jobs reviewed per hour per day for November 2020?

```
Output: SELECT ds, COUNT(*)/24 AS no_of_job
FROM task_3.`task 3 trainity`
WHERE ds BETWEEN '01/11/2020' AND '30/11/2020'
GROUP BY ds
;
```

Task 2: Calculate 7 day rolling average of throughput? For throughput, do you prefer daily metric or 7-day rolling and why?

```
SELECT ds, COUNT(*) AS events_per_second,
AVG(COUNT(*)) OVER (ORDER BY ds ROWS BETWEEN 6 PRECEDING AND CURRENT ROW) AS rolling_avg
FROM task_3.`task 3 trainity`
GROUP BY ds
ORDER BY ds;
```

Task 3: Calculate the percentage share of each language in the last 30 days?

```
SELECT language, COUNT(*) AS num_jobs_reviewed,
```

```
COUNT(*)*100 / SUM(COUNT(*)) OVER() AS pct_share  
  
FROM task_3.`task 3 trainity`  
  
WHERE ds >= DATE_TRUNC('day', NOW()) - INTERVAL '30 day'  
  
GROUP BY language  
  
ORDER BY pct_share DESC;
```

Task 4: Let's say you see some duplicate rows in the data. How will you display duplicates from the table?

```
SELECT *  
  
FROM task_3.`task 3 trainity`  
  
WHERE (job_id, actor_id, event, language, time_spent, org, ds) IN (  
  
    SELECT job_id, actor_id, event, language, time_spent, org, ds  
  
    FROM task_3.`task 3 trainity`  
  
    GROUP BY job_id, actor_id, event, language, time_spent, org, ds  
  
    HAVING COUNT(*) > 1  
  
);
```

Case Study 2 (Investigating metric spike)

Task 1: Calculate the weekly user engagement?

```
SELECT  
  
date_trunc('week', timestamp) as week_start,  
  
COUNT(DISTINCT user_id) as weekly_active_users  
  
FROM  
  
task_3.events
```

GROUP BY
week_start

Task 2: Calculate the user growth for product?

```
SELECT  
  
date_trunc('week', signup_date) as week_start,  
  
COUNT(*) as new_users  
  
FROM  
  
task_3.users  
  
GROUP BY  
  
week_start
```

Task 3: Calculate the weekly retention of users-sign up cohort?

```
SELECT  
  
date_trunc('week', e.timestamp) as week_start,  
  
COUNT(DISTINCT e.user_id) as weekly_retained_users,  
  
(date_trunc('week', e.timestamp) - date_trunc('week', u.signup_date)) / 7 as weeks_since_signup  
  
FROM  
  
task_3.events e  
  
JOIN users u ON e.user_id = u.user_id  
  
GROUP BY  
  
week_start,  
  
weeks_since_signup
```

Task 4: Calculate the weekly engagement per device?

```
SELECT

    date_trunc('week', timestamp) as week_start,

    device,

    COUNT(DISTINCT user_id) as weekly_active_users

FROM

    task_3.events

GROUP BY

    week_start,

    device
```

Task 5: Calculate the email engagement metrics?

```
SELECT

    COUNT(*) as email_events,

    COUNT(DISTINCT user_id) as unique_email_users

FROM

    task_3. email_events
```

Result: I understand how to perform and understand the real time problems for useful insights .this task helped me a lot to develop my logic as well as skills.