

# Operation Analytics and Investigating Metric Spike

## Advanced SQL

### Description:

Operation Analytics 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.

Being one of the most important parts of a company, this kind of analysis is further used to predict the overall growth or decline of a company's fortune. It means better automation, better understanding between cross-functional teams, and more effective workflows.

Investigating metric spike is also an important part of operation analytics as being a Data Analyst you must be able to understand or make other teams understand questions like- Why is there a dip in daily engagement? Why have sales taken a dip? Etc. Questions like these must be answered daily and for that its very important to investigate metric spike.

You are working for a company like Microsoft designated as Data Analyst Lead and is provided with different data sets, tables from which you must derive certain insights out of it and answer the questions asked by different departments.

**You are required to provide a detailed report for the below two operations mentioning the answers for the related questions:**

### Case Study 1 (Job Data)

**Below is the structure of the table with the definition of each column that you must work on:**

- **Table-1:** job\_data
  - **job\_id:** unique identifier of jobs
  - **actor\_id:** unique identifier of actor
  - **event:** decision/skip/transfer
  - **language:** language of the content
  - **time\_spent:** time spent to review the job in seconds
  - **org:** organization of the actor
  - **ds:** date in the yyyy/mm/dd format. It is stored in the form of text and we use presto to run. no need for date function

Use the dataset attached in the Dataset section below the project images then answer the questions that follows

- A. **Number of jobs reviewed:** Amount of jobs reviewed over time.  
**Your task:** Calculate the number of jobs reviewed per hour per day for November 2020?
- B. **Throughput:** It is the no. of events happening per second.  
**Your task:** Let's say the above metric is called throughput. Calculate 7 day rolling average of throughput? For throughput, do you prefer daily metric or 7-day rolling and why?
- C. **Percentage share of each language:** Share of each language for different contents.  
**Your task:** Calculate the percentage share of each language in the last 30 days?
- D. **Duplicate rows:** Rows that have the same value present in them.  
**Your task:** Let's say you see some duplicate rows in the data. How will you display duplicates from the table?

### Case Study 2 (Investigating metric spike)

The structure of the table with the definition of each column that you must work on is present in the project image

- **Table-1:** users  
This table includes one row per user, with descriptive information about that user's account.
- **Table-2:** events  
This table includes one row per event, where an event is an action that a user has taken. These events include login events, messaging events, search events, events logged as users progress through a signup funnel, events around received emails.
- **Table-3:** email\_events  
This table contains events specific to the sending of emails. It is similar in structure to the events table above.

Use the dataset attached in the Dataset section below the project images then answer the questions that follows

- A. **User Engagement:** To measure the activeness of a user. Measuring if the user finds quality in a product/service.  
**Your task:** Calculate the weekly user engagement?
- B. **User Growth:** Amount of users growing over time for a product.  
**Your task:** Calculate the user growth for product?
- C. **Weekly Retention:** Users getting retained weekly after signing-up for a product.  
**Your task:** Calculate the weekly retention of users-sign up cohort?
- D. **Weekly Engagement:** To measure the activeness of a user. Measuring if the user finds quality in a product/service weekly.  
**Your task:** Calculate the weekly engagement per device?
- E. **Email Engagement:** Users engaging with the email service.  
**Your task:** Calculate the email engagement metrics?