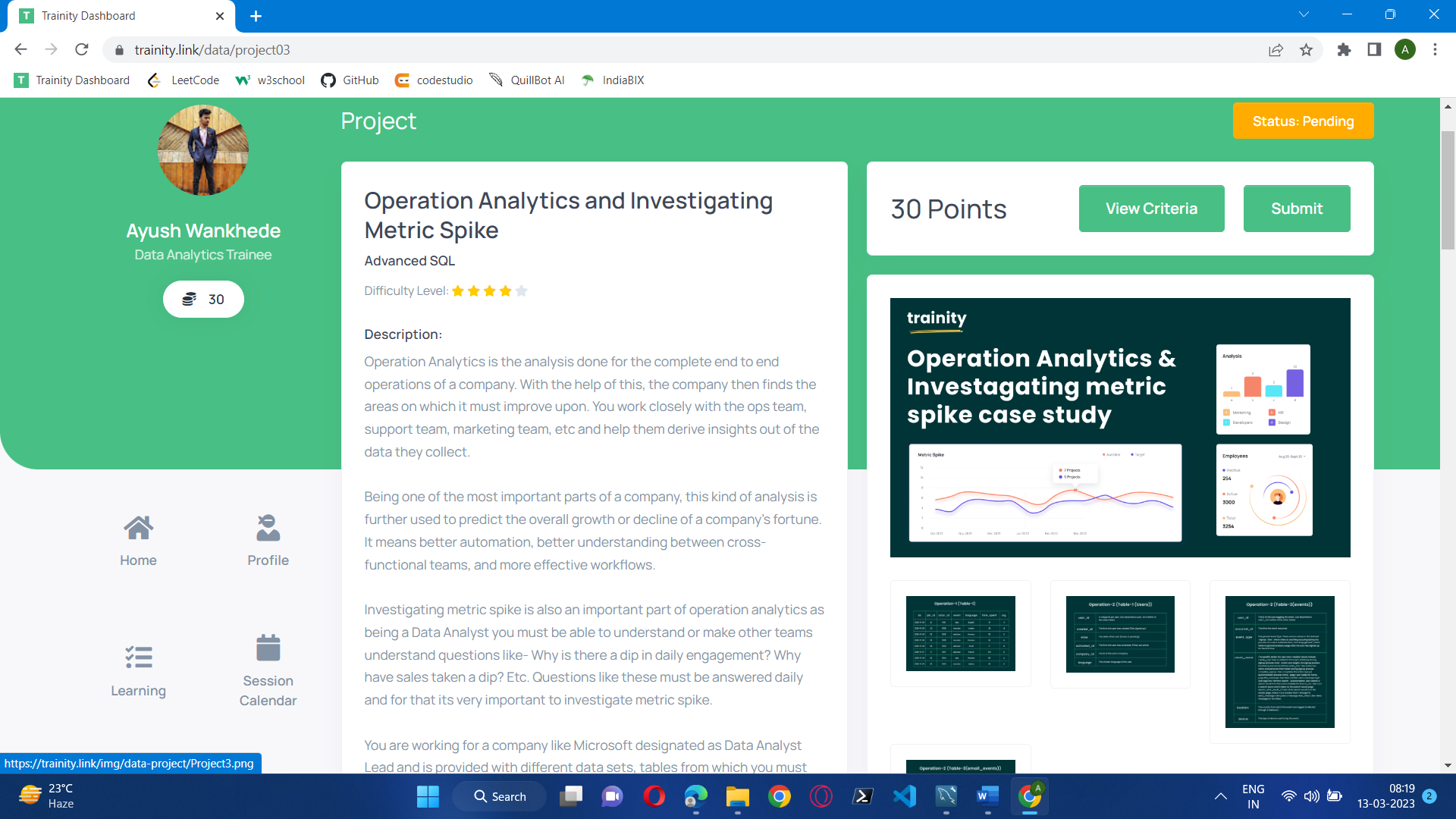
**Operation Analytics and Investigating Metric Spike**

**Advanced SQL**



**Project Description:**

The goal of this project is to analyse the data provided by various departments and respond to their inquiries. So that team, support team, and marketing team fortune may use this information. Better automation and cross-effective workflow knowledge are the results.

In case study 1 there is job\_data table while in case study 2 there are users, events Investigating Metric Spike .

**Approach:**

SQL Workbench is being used to create this project. I need to first establish a database using the dataset file that the company sent me. Next analysis and locate the information that will the assist the understand questions like - Why is there a fall in daily engagement? Why Daily answers to questions like these are required, thus it's crucial to look into metric spike.

**Tech-Stack Used:**

A visual editor called MySQL Workbench combines database management, SQL development, and data modelling into one user interface. You may create, administer, and design databases visually with it.

While working with structured data, MySQL Workbench is frequently used. It is an open Relational Database Management System (RDBMS) created by Oracle Corporation and Sun Microsystems that communicates with databases using SQL.

**Insights:**

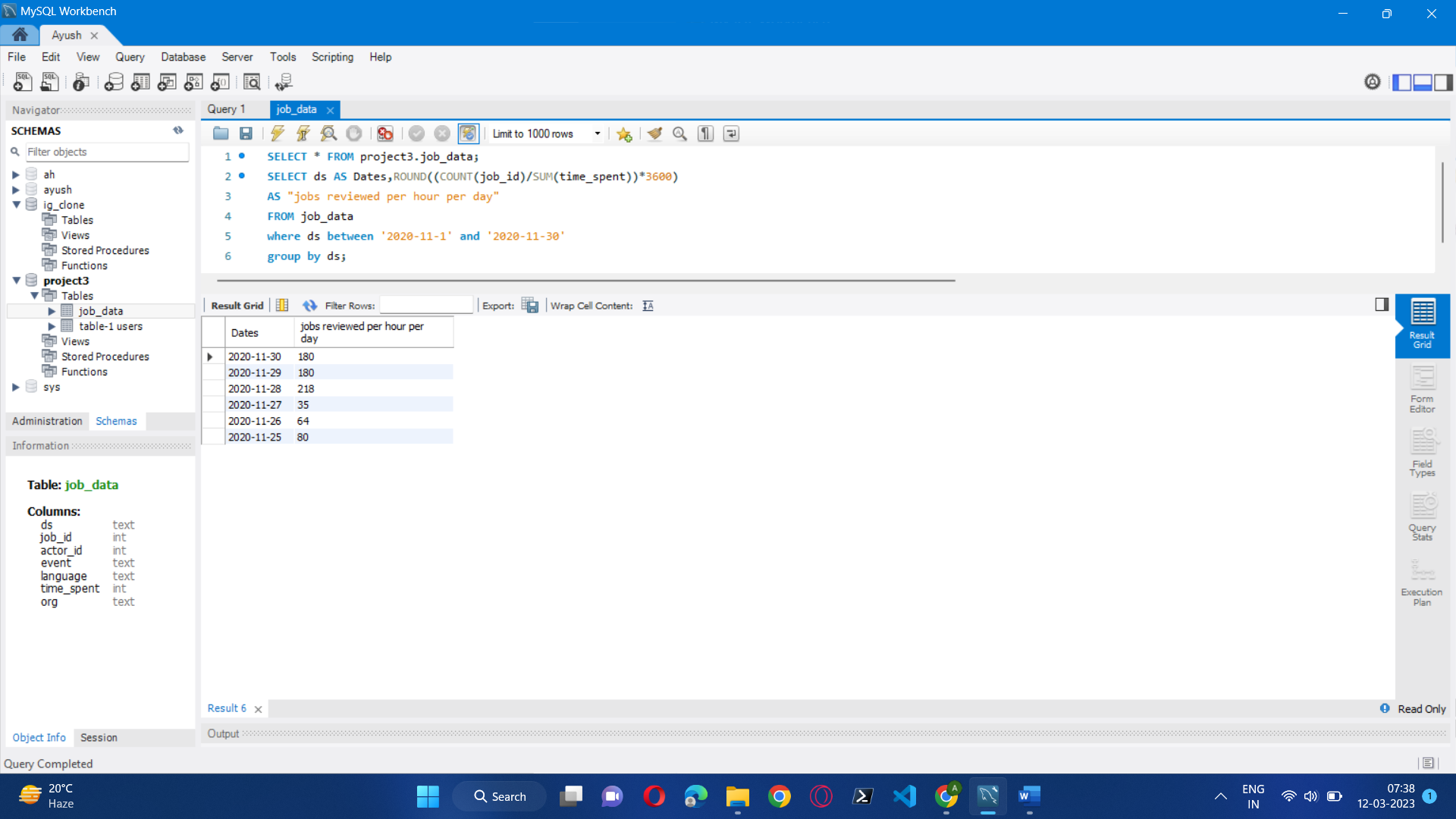
**Case Study 1 (Job Data)**

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

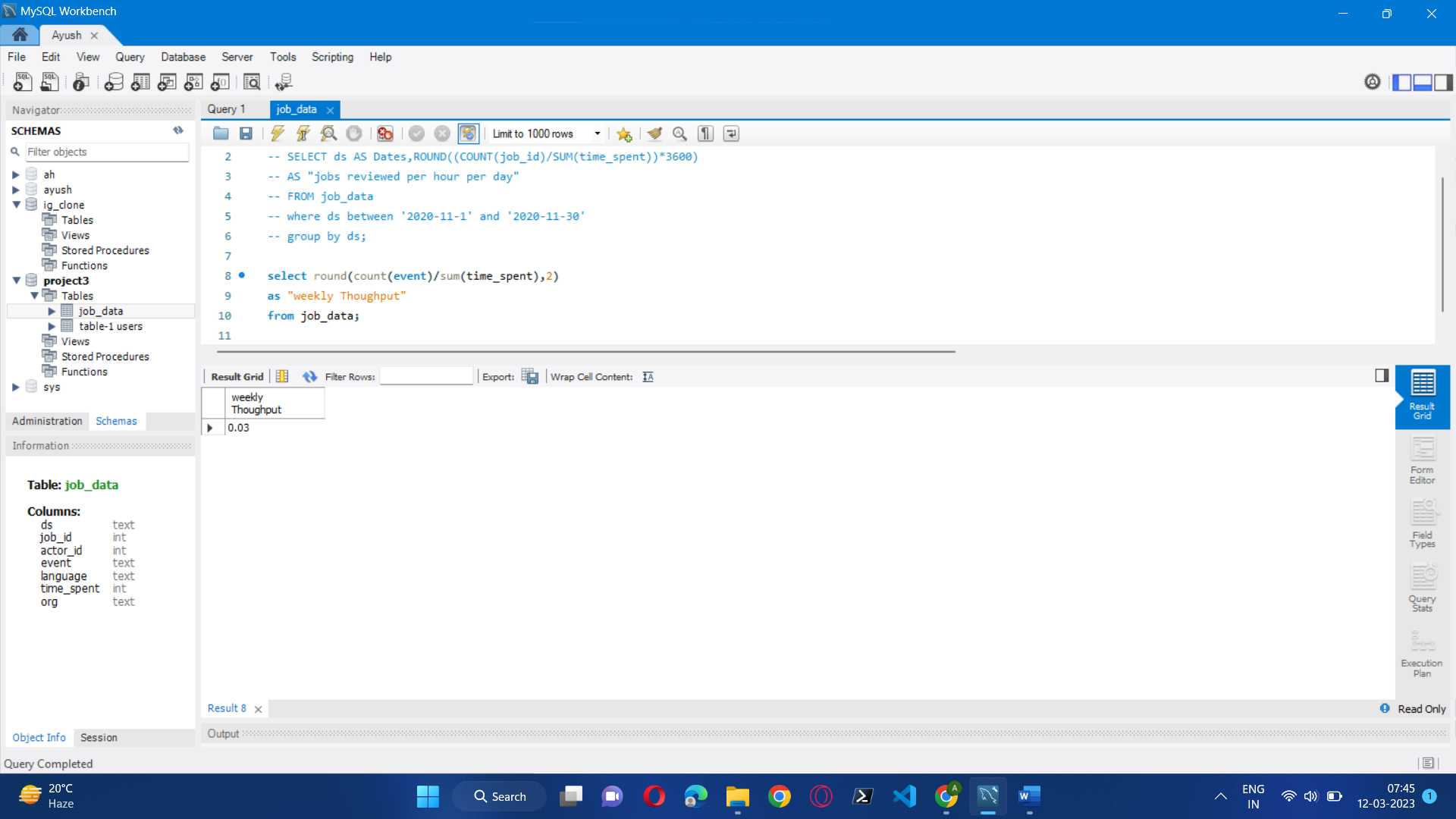
* **Table-1:**job\_data  
  + **job\_id:**unique identifier of jobs
  + **actor\_id:**unique identifier of actor
  + **event:**decision/skip/transfer
  + **language:**language of the content
  + **time\_spent:**time spent to review the job in seconds
  + **org:**organization of the actor
  + **ds:**date 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

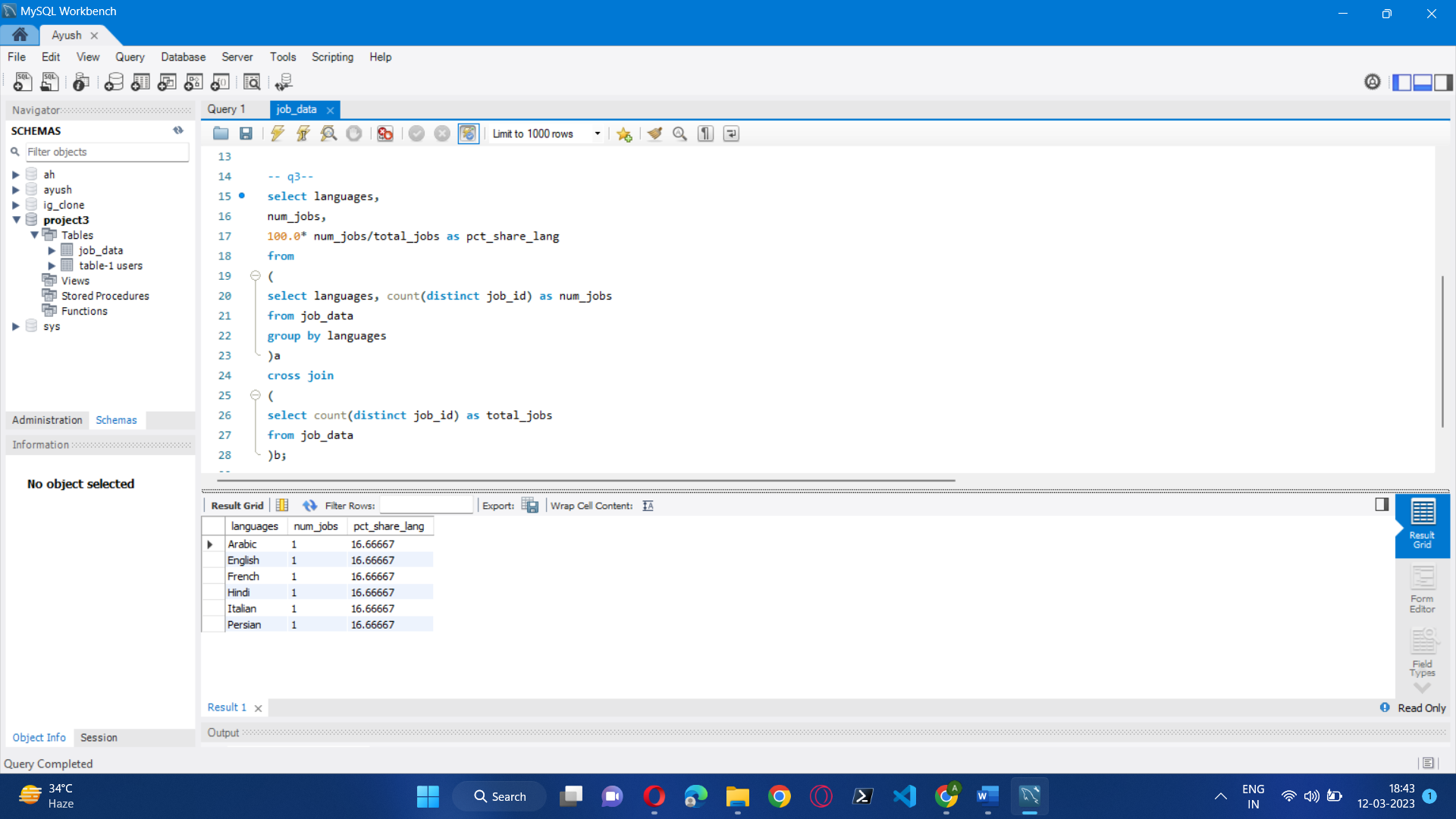
1. **Number of jobs reviewed:**Amount of jobs reviewed over time.  
   **Your task:** Calculate the number of jobs reviewed per hour per day for November 2020?



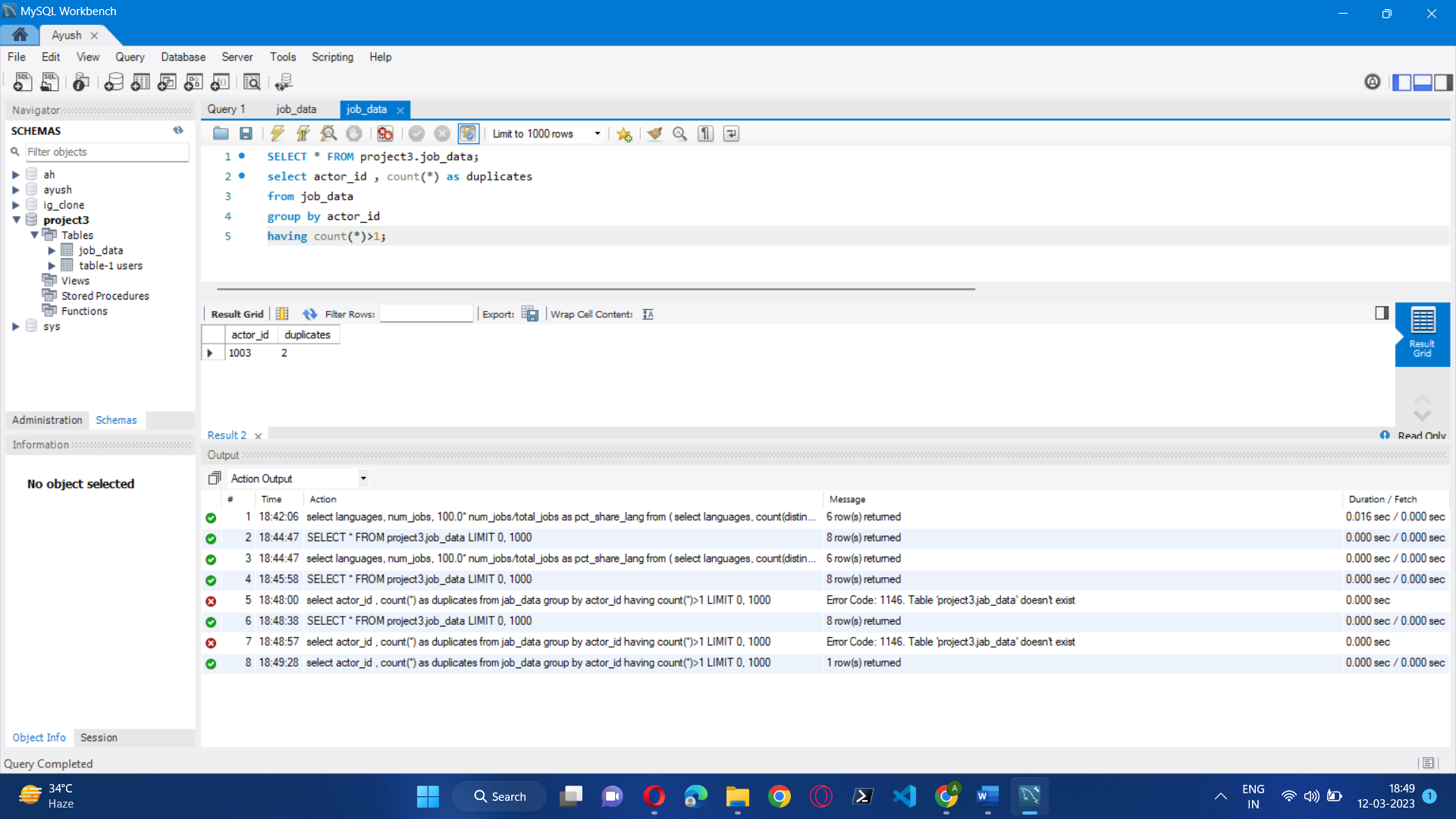
1. **Throughput:**It is the no. of events happening per second.  
   **Your task:** Let’s say the above metric is called throughput. Calculate 7 day rolling average of throughput? For throughput, do you prefer daily metric or 7-day rolling and why?



1. **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?



1. **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?



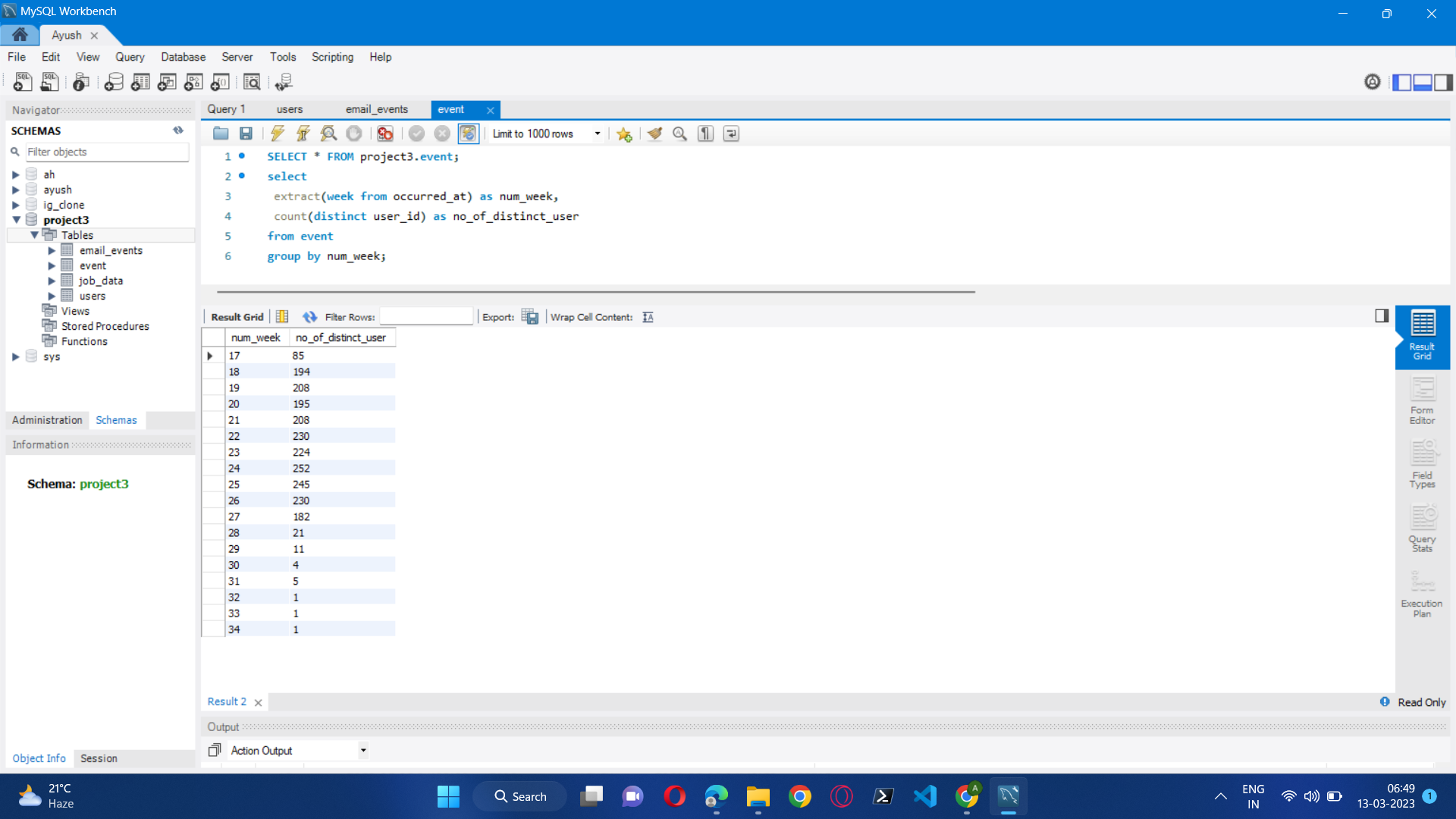
**Case Study 2 (Investigating metric spike)**

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

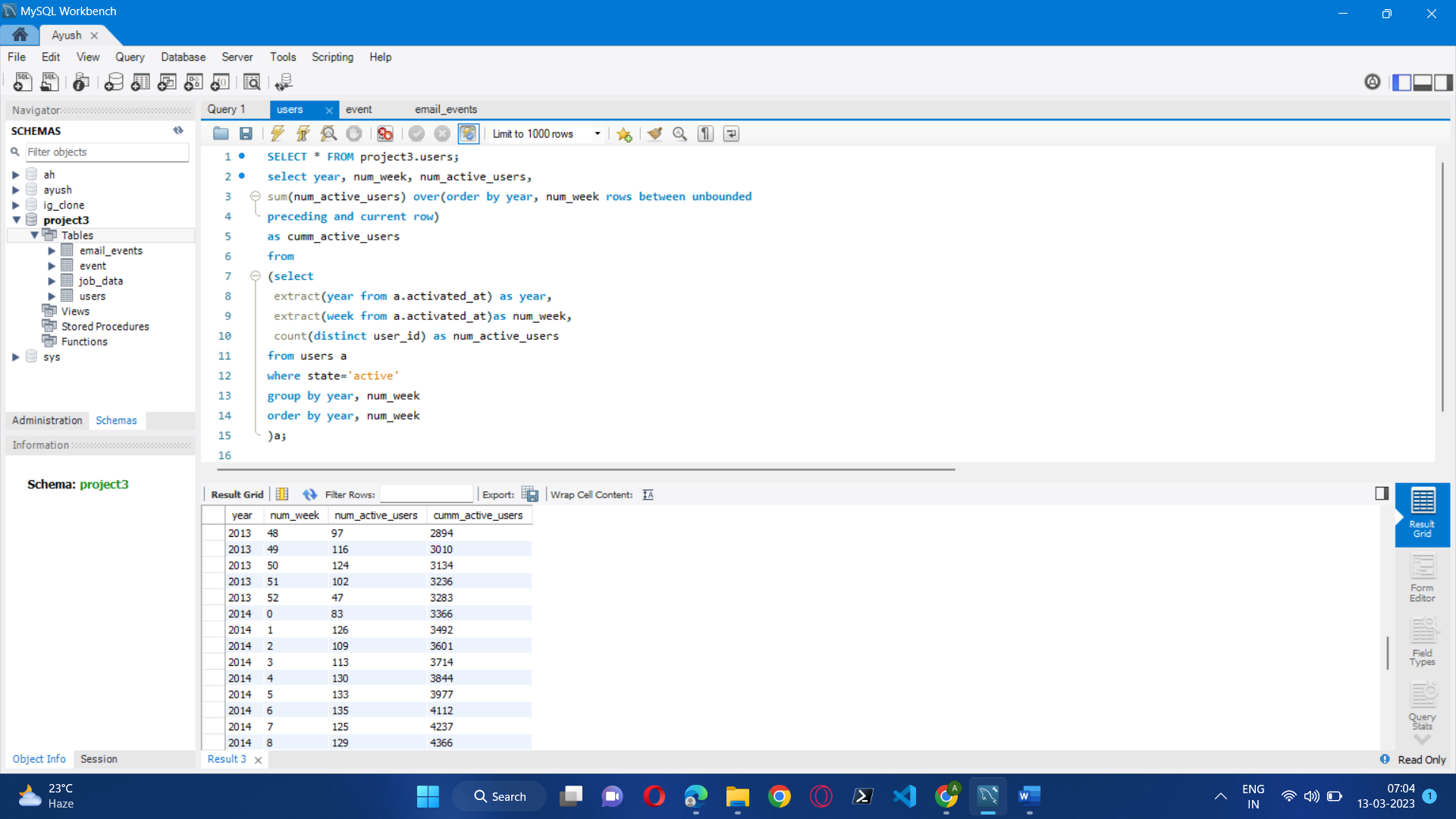
* **Table-1:**users  
  This table includes one row per user, with descriptive information about that user’s account.
* **Table-2:**events  
  This table includes one row per event, where an event is an action that a user has taken. These events include login events, messaging events, search events, 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

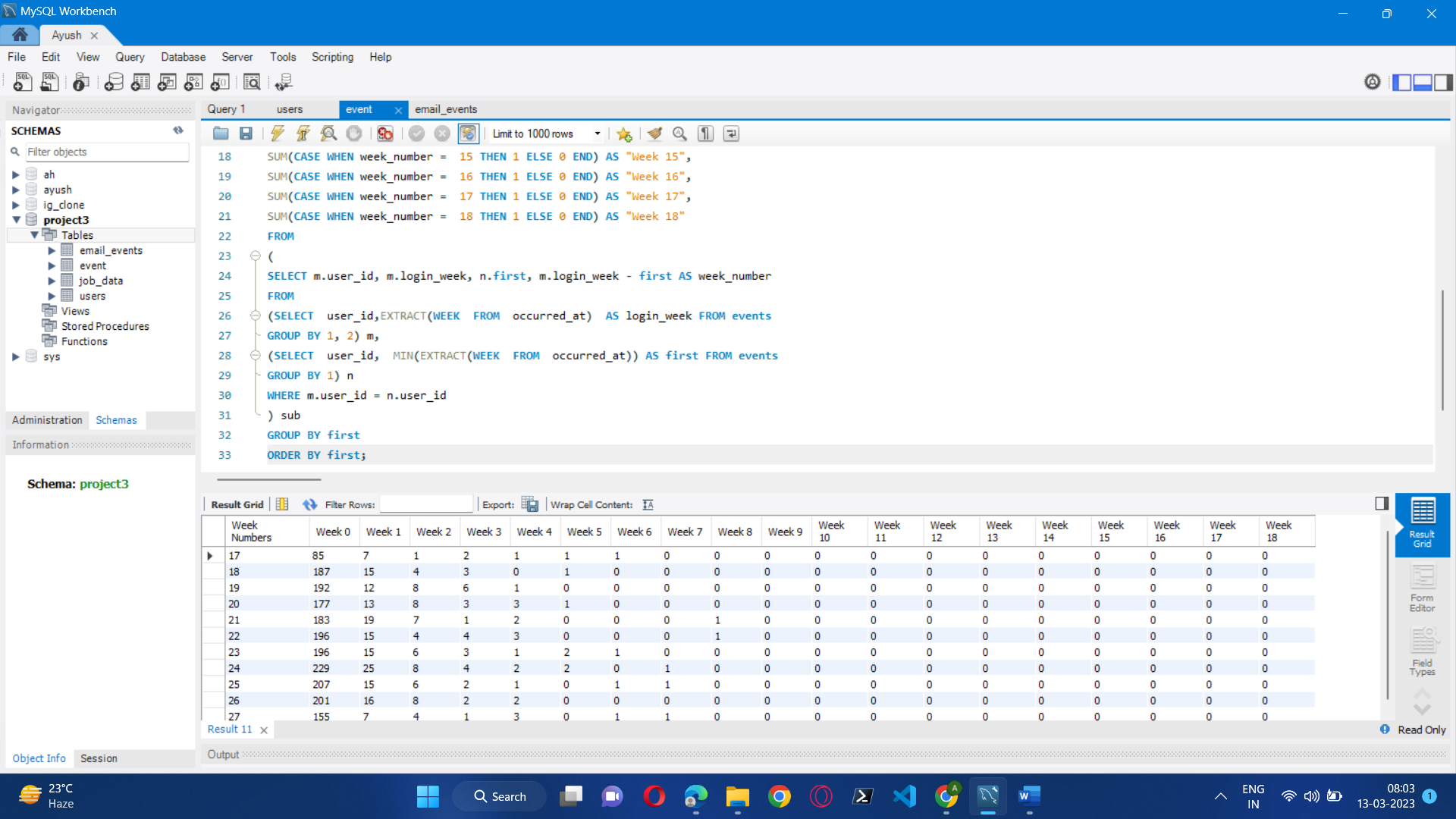
1. **User Engagement:**To measure the activeness of a user. Measuring if the user finds quality in a product/service.  
   **Your task:** Calculate the weekly user engagement?



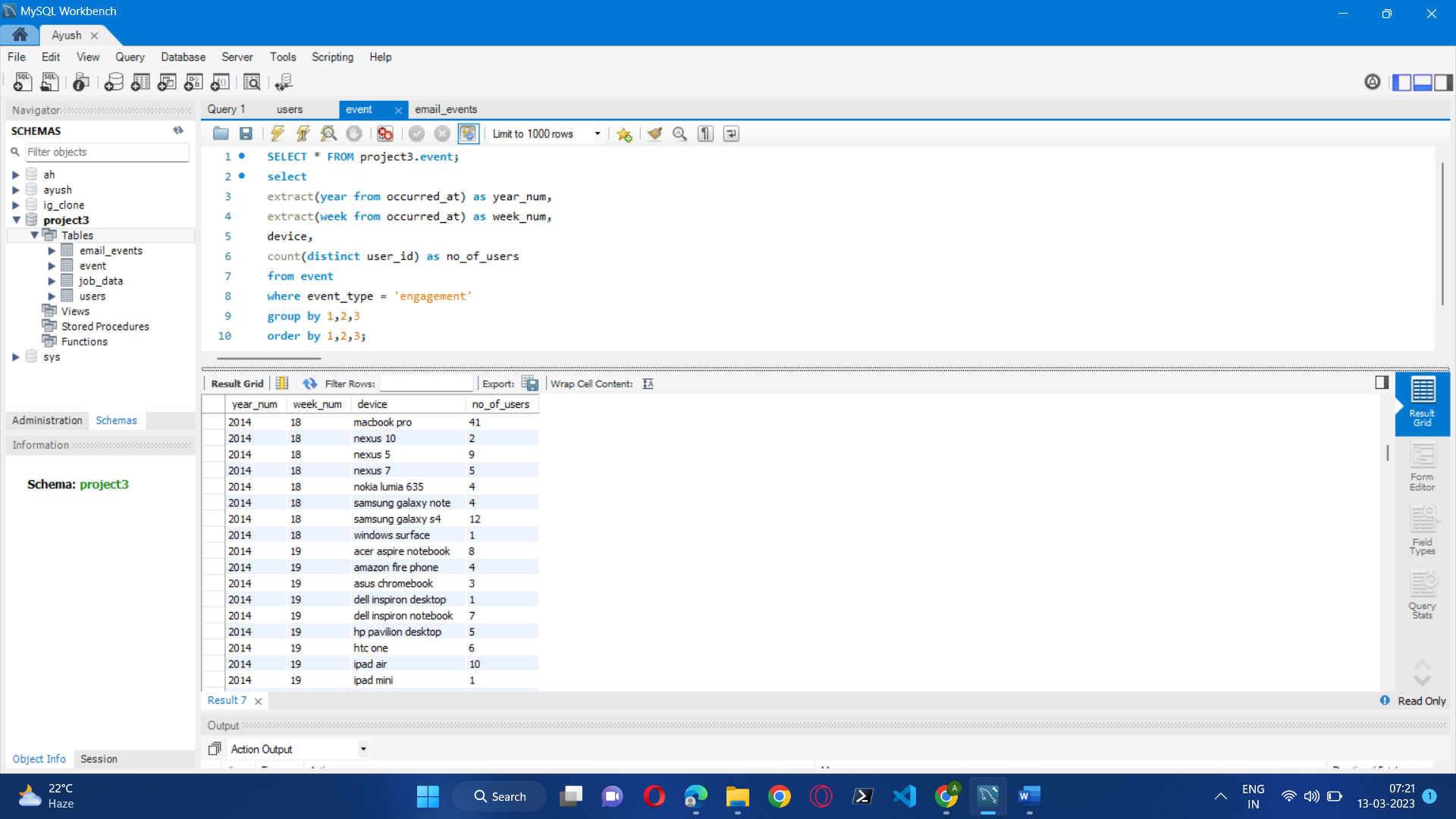
1. **User Growth:**Amount of users growing over time for a product.  
   **Your task:** Calculate the user growth for product?



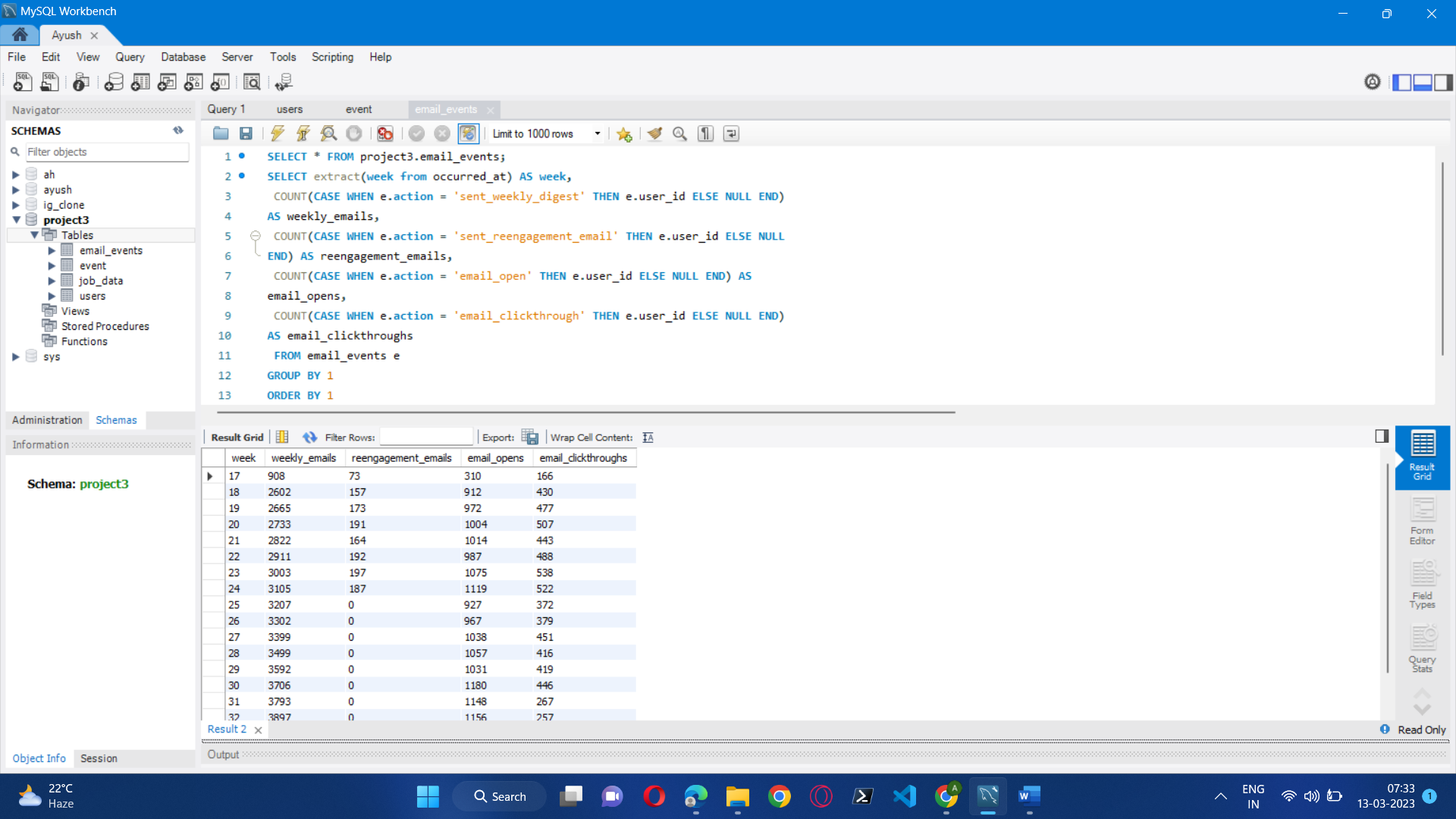
1. **Weekly Retention:**Users getting retained weekly after signing-up for a product.  
   **Your task:** Calculate the weekly retention of users-sign up cohort?



1. **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?



1. **Email Engagement:**Users engaging with the email service.  
   **Your task:** Calculate the email engagement metrics?



**Result:**

I learned how to use advanced SQL concepts, such as Windows Functions, in this project. I was aware of how the real-world industry functions. I was able to learn my SQL concepts as a result.

I gained the ability to ask the appropriate questions in light of the situation. Given the information and

questions, which columns to take into account, and how to get the priceless insights that support corporate growth. I discovered how the business identifies several business-related areas for improvement. I learned about looking into metric spikes