##### **Operation Analytics and Investigating Metric Spike**

**Project Description**

The project "Operation Analytics and Investigating Metric Spike" involves analyzing operational data to investigate spikes in specific metrics. This analysis aims to understand the causes and implications of sudden increases in metrics within a system or process.

**Approach**

* *Prepare & Analyze*

The dataset is initially prepared and analyzed through the utilization of SQL functions such as SELECT, WHERE, GROUP BY, JOIN, among others, to extract pertinent insights from the data. These functions serve to filter, aggregate, and combine data, facilitating the extraction of meaningful information from the dataset for analysis.

* *Identify KPIs*

After the completion of data cleaning and transformation, the subsequent stage involves the identification of Key Performance Indicators (KPIs) pertinent to the business. These KPIs encompass metrics such as throughput, user engagement, and user retention, among others, which serve as crucial indicators of business performance and effectiveness.

* *Visualize Data*

The concluding phase involves leveraging Tableau to craft interactive charts and dashboards. Tableau serves as a pivotal tool for developing dynamic visualizations, aiding in the discernment of trends and patterns within the data.

**Tech-Stack Used**

* Data Collection: Custom scripts, APIs, or data extraction tools depending on the source of data.
* Data Preprocessing: Python and MySQL Workbench

**Insights**

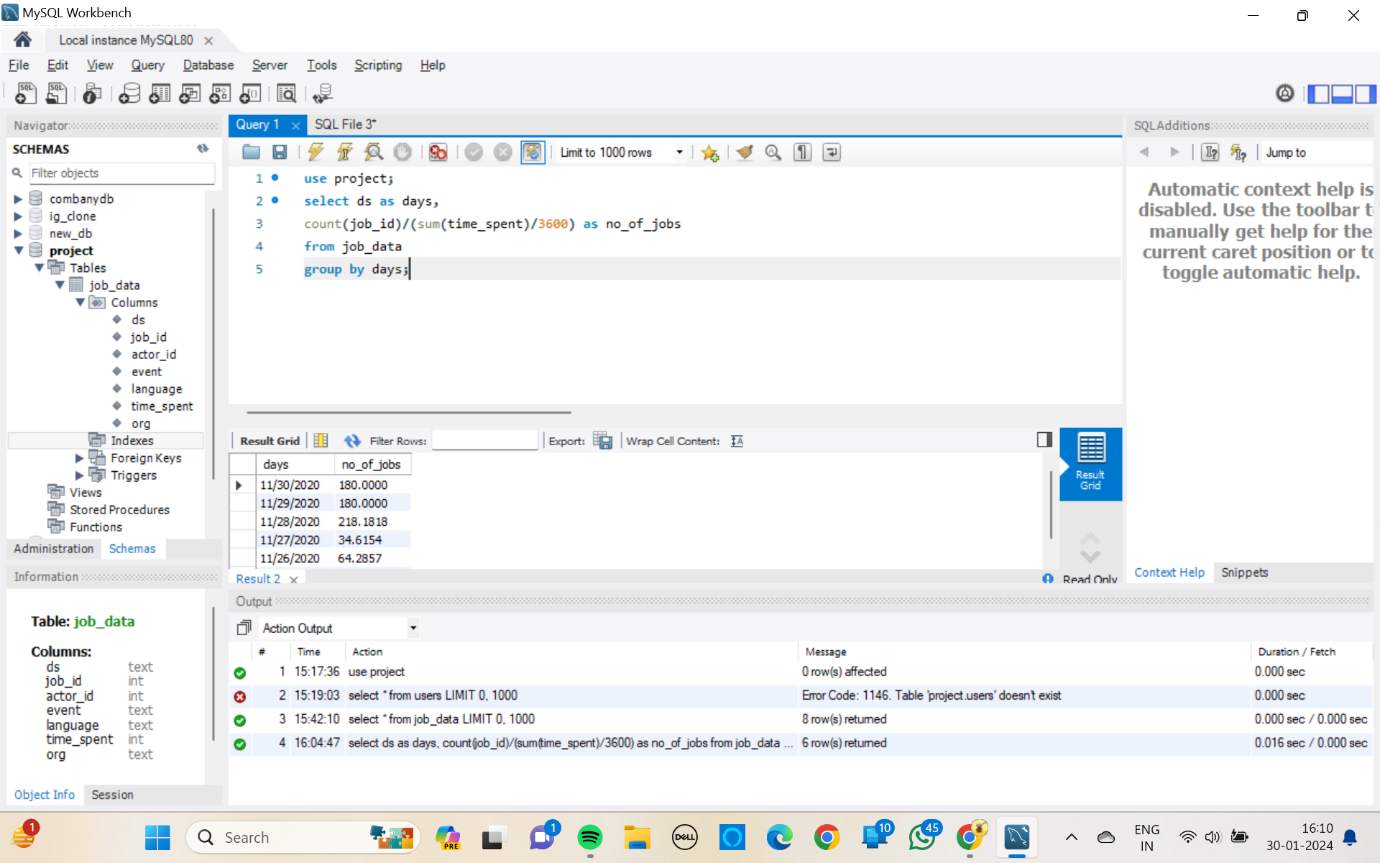
**Case Study 1: Job Data Analysis**

**You will be working with a table named job\_data with the following columns:**

* **job\_id:**Unique identifier of jobs
* **actor\_id:**Unique identifier of actor
* **event:**The type of event (decision/skip/transfer).
* **language:**The Language of the content
* **time\_spent:**Time spent to review the job in seconds.
* **org:**The Organization of the actor
* **ds:**The date in the format yyyy/mm/dd (stored as text).

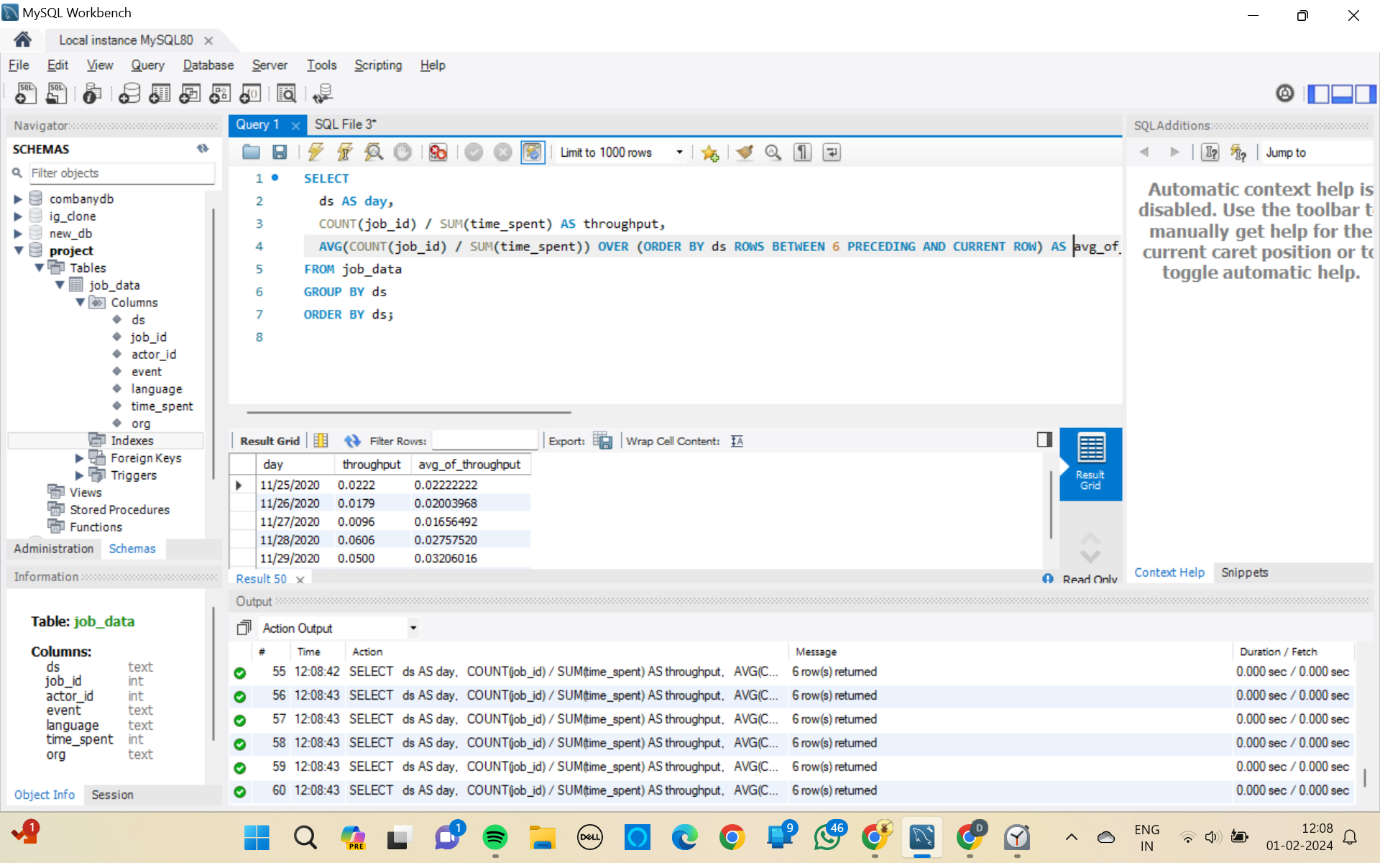
**Tasks:**

1. **Jobs Reviewed Over Time:**
   * Objective: Calculate the number of jobs reviewed per hour for each day in November 2020.
   * Your Task: Write an SQL query to calculate the number of jobs reviewed per hour for each day in November 2020.



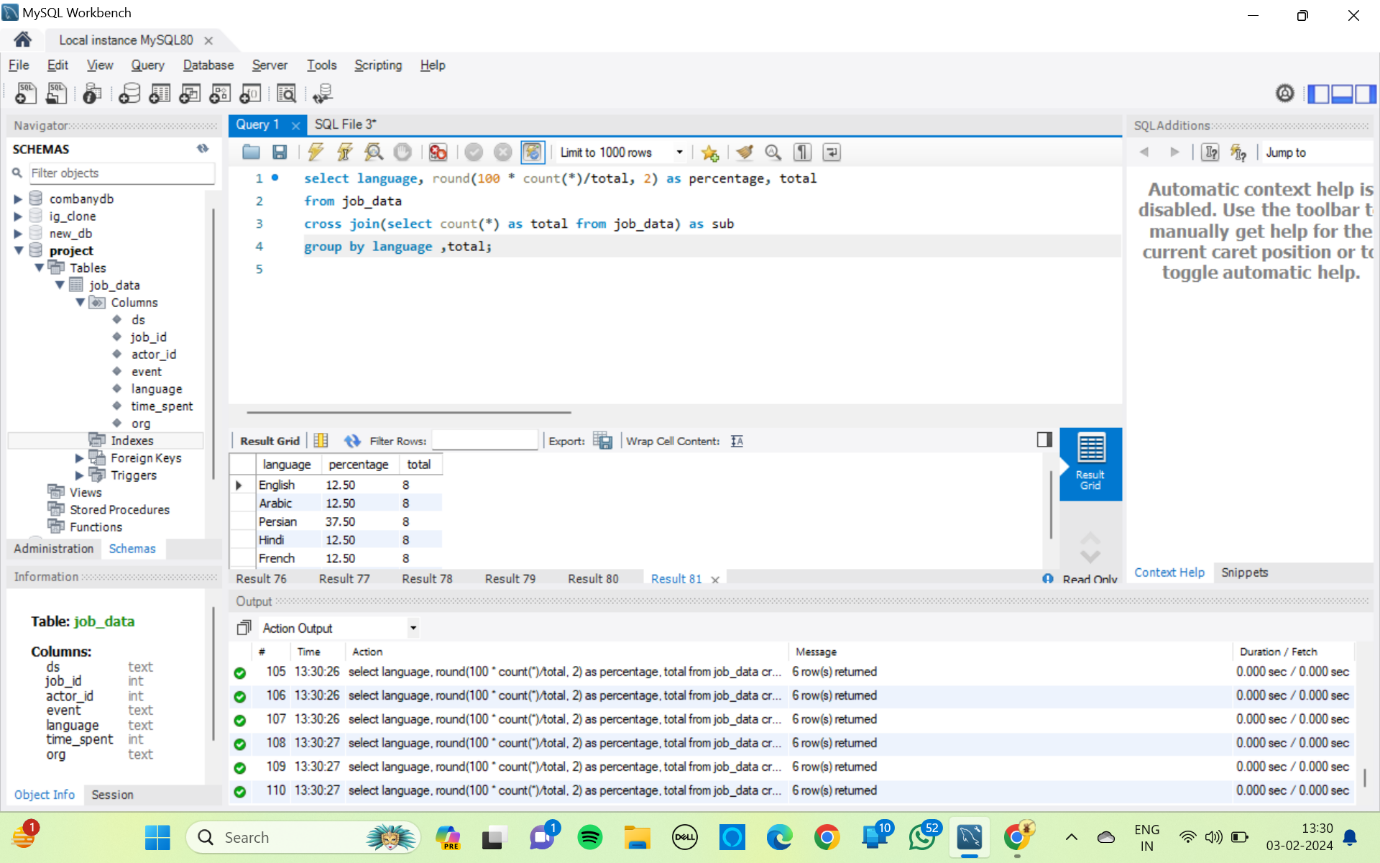
* Highest number of jobs were reviewed on 28 November 2020 with 218 jobs per hour.
* Lowest number of jobs were reviewed on 27 November with 34 jobs per hour.

1. **Throughput Analysis:**
   * Objective: Calculate the 7-day rolling average of throughput (number of events per second).
   * Your Task: Write an SQL query to calculate the 7-day rolling average of throughput. Additionally, explain whether you prefer using the daily metric or the 7-day rolling average for throughput, and why.



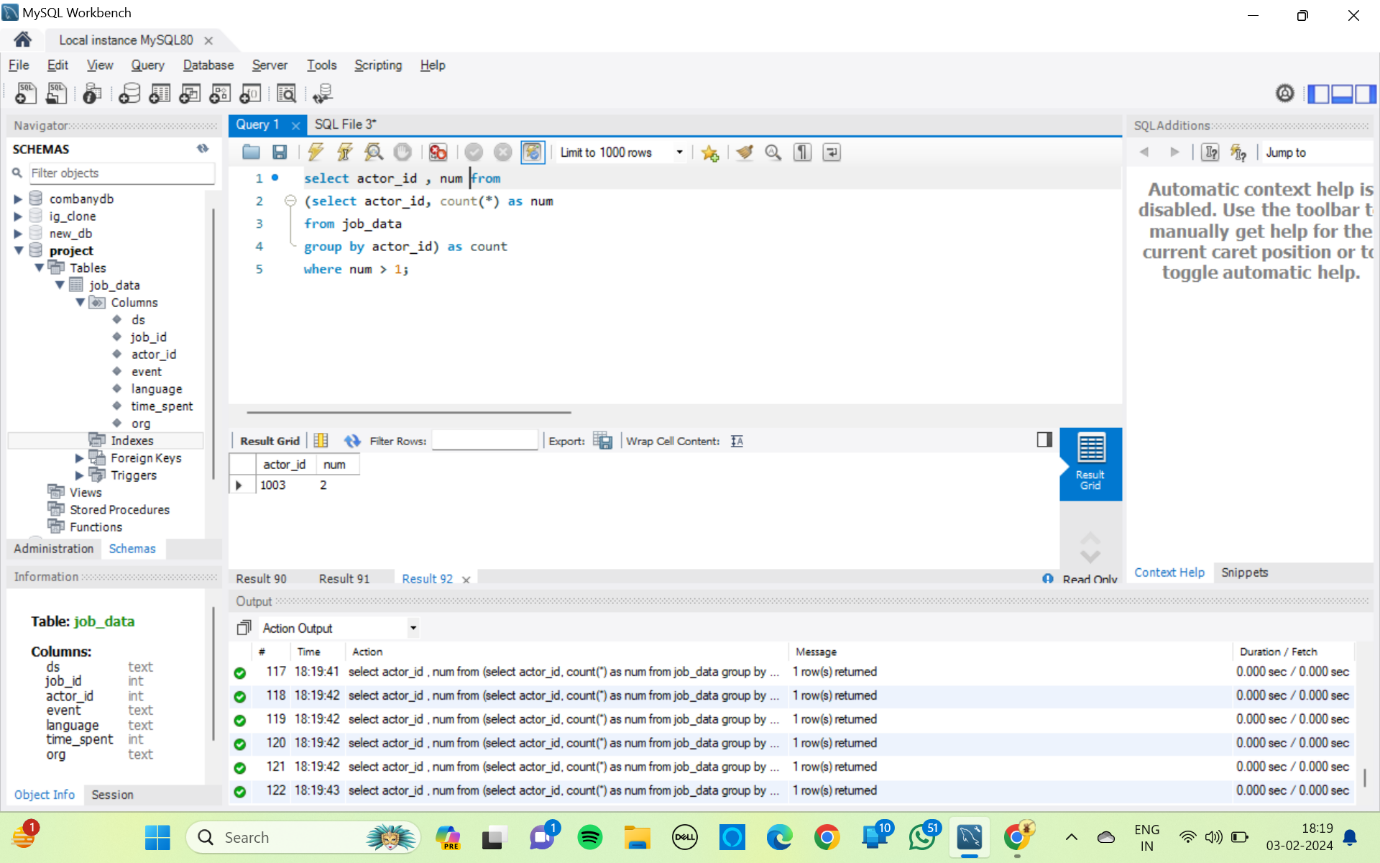
* The rolling average is consistent and helps understand the trend over time without much fluctuations compared to the daily metric.

1. **Language Share Analysis:**
   * Objective: Calculate the percentage share of each language in the last 30 days.
   * Your Task: Write an SQL query to calculate the percentage share of each language over the last 30 days.



* Persian is the most used language with percentage share of 37.5 %
* The rest of the languages, Arabic, Hindi, Italian, French and English has the same percentage share of 12.5%

1. **Duplicate Rows Detection:**
   * Objective: Identify duplicate rows in the data.
   * Your Task: Write an SQL query to display duplicate rows from the job\_data table.



* No duplicate rows were present in the data

**CASE STUDY 1: Analysis**

* Strengths – Productivity : 126 jobs reviewed per hour in a month
* Weaknesses : Relatively low throughput.
* Opportunities : Reward employees for performance to retain the productivity, Introduce programs to improve use of other languages.

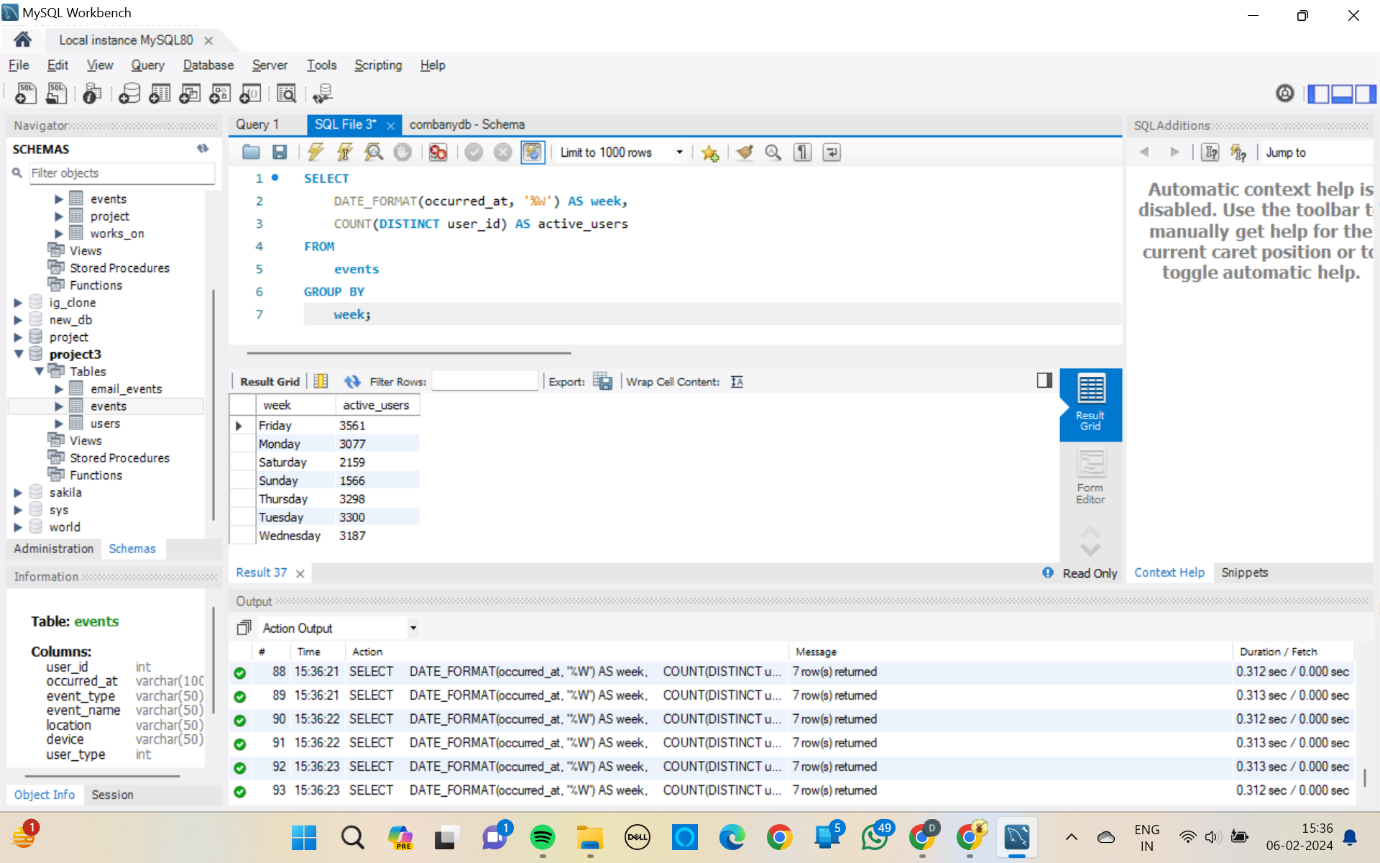
**Case Study 2: Investigating Metric Spike**

**You will be working with three tables:**

* **users**: Contains one row per user, with descriptive information about that user’s account.
* **events**: Contains one row per event, where an event is an action that a user has taken (e.g., login, messaging, search).
* **email\_events**: Contains events specific to the sending of emails.

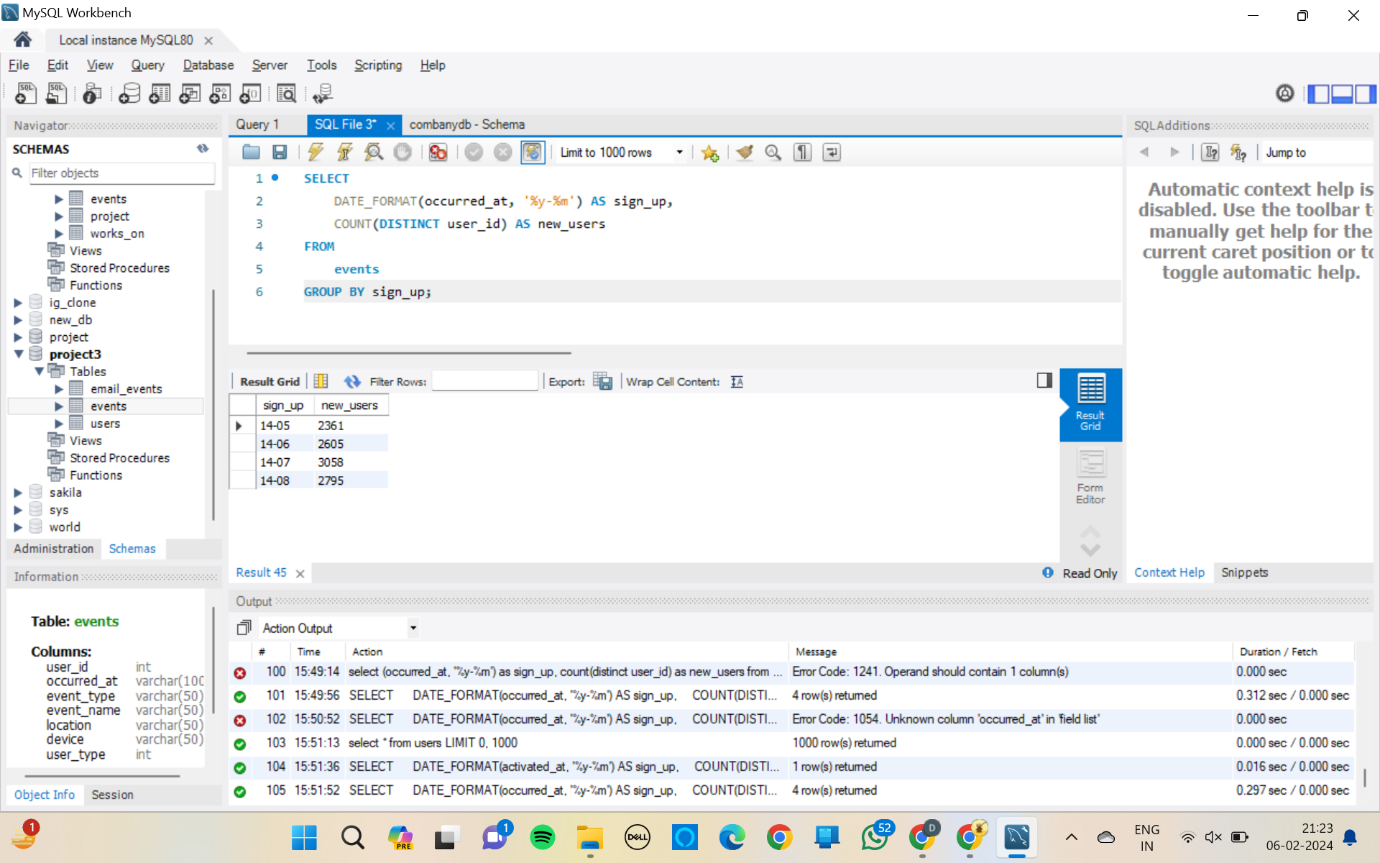
**Tasks:**

1. **Weekly User Engagement:**
   * Objective: Measure the activeness of users on a weekly basis.
   * Your Task: Write an SQL query to calculate the weekly user engagement.

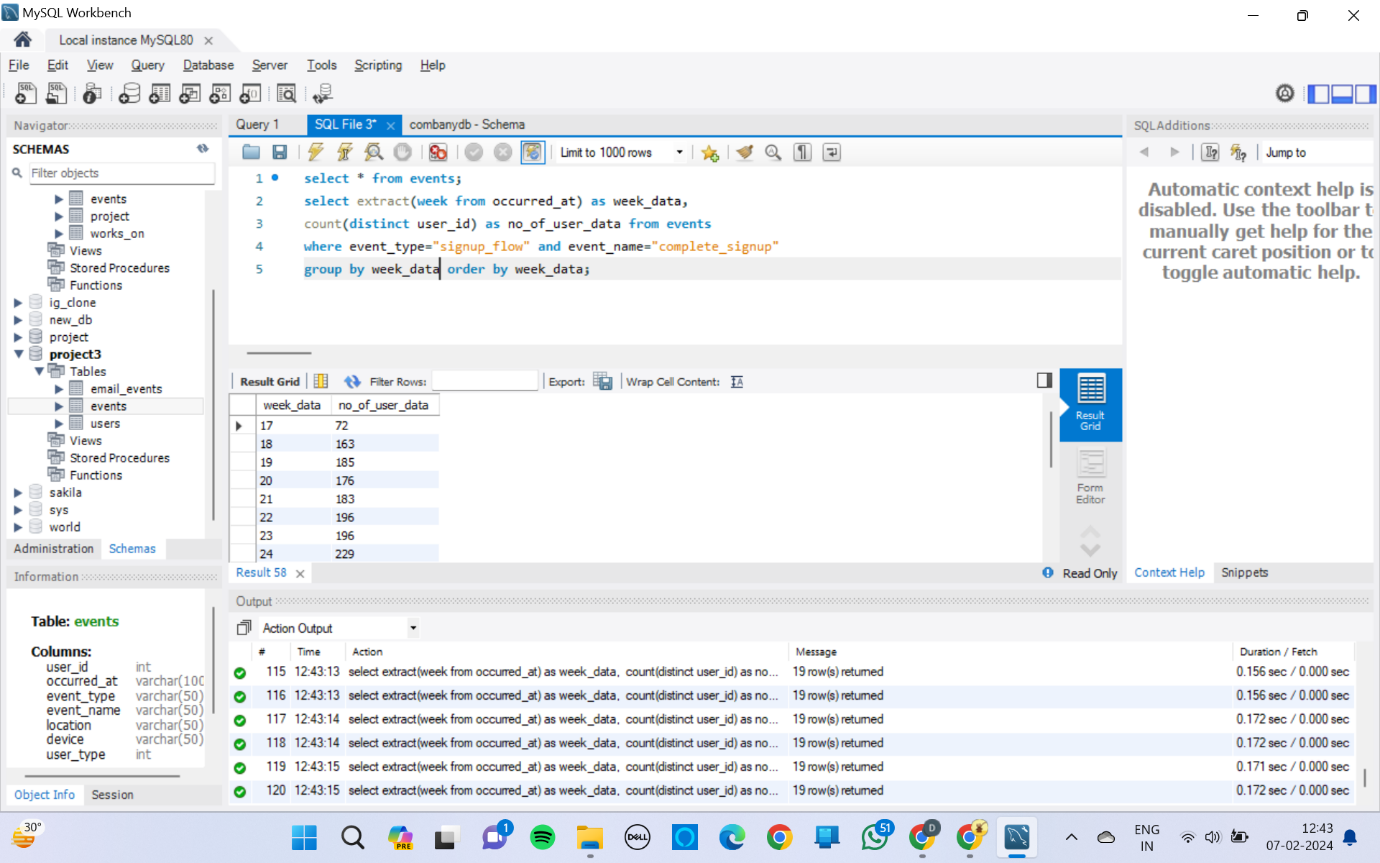


* The highest engagement was seen on Tuesday, with 3300 users.
* The lowest engagement was seen on Sunday, with 1566 users.

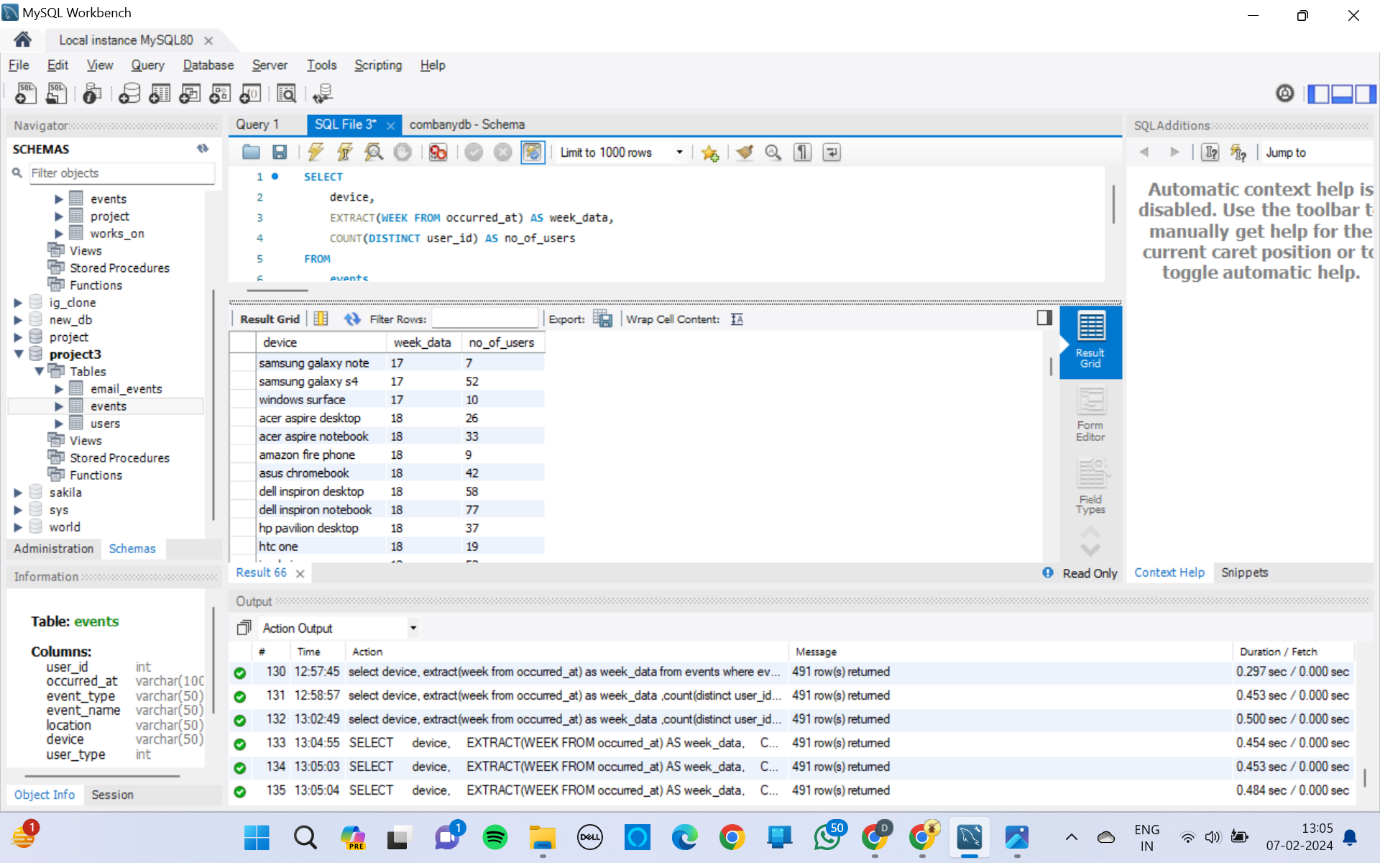
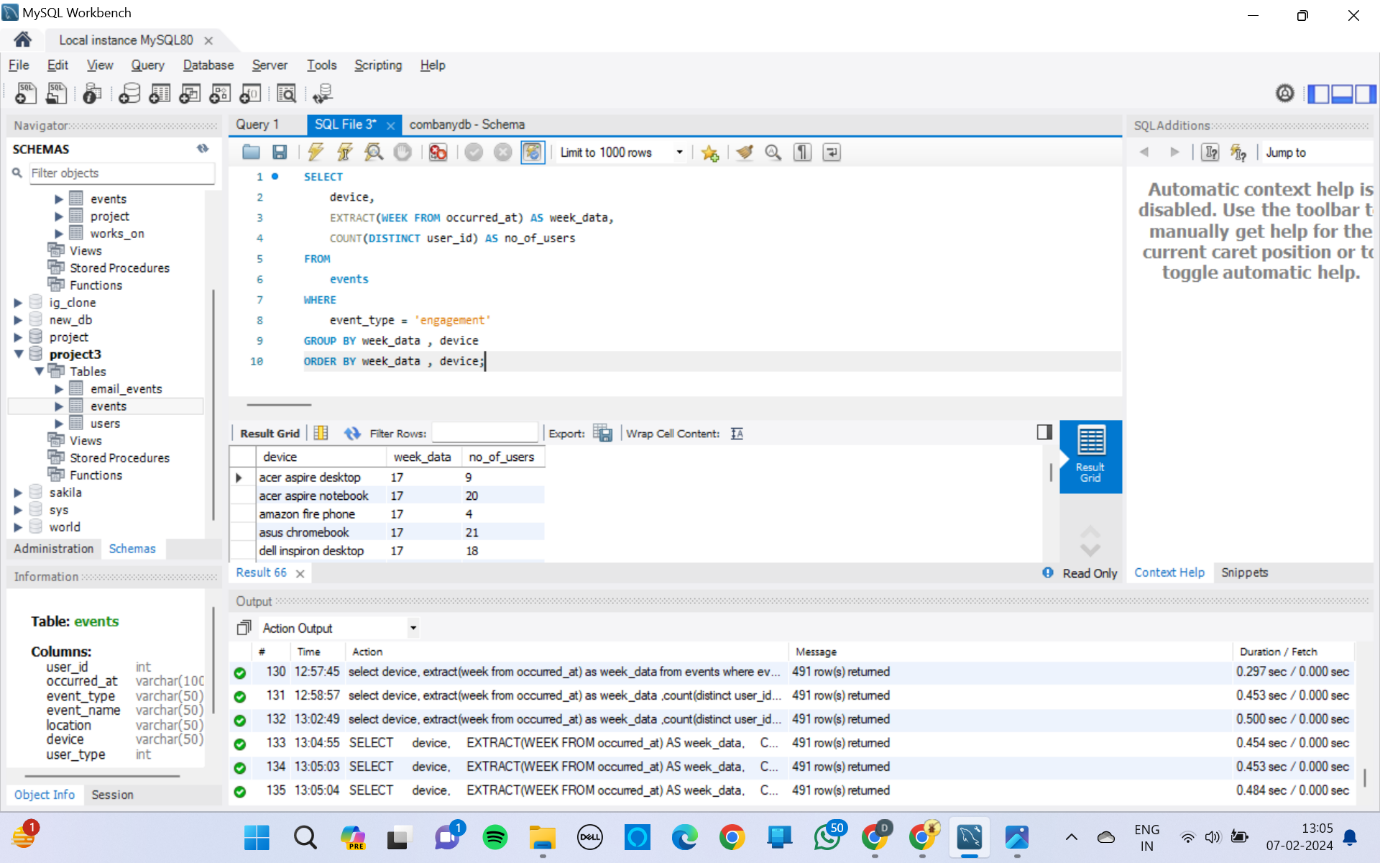
1. **User Growth Analysis:**
   * Objective: Analyze the growth of users over time for a product.
   * Your Task: Write an SQL query to calculate the user growth for the product.



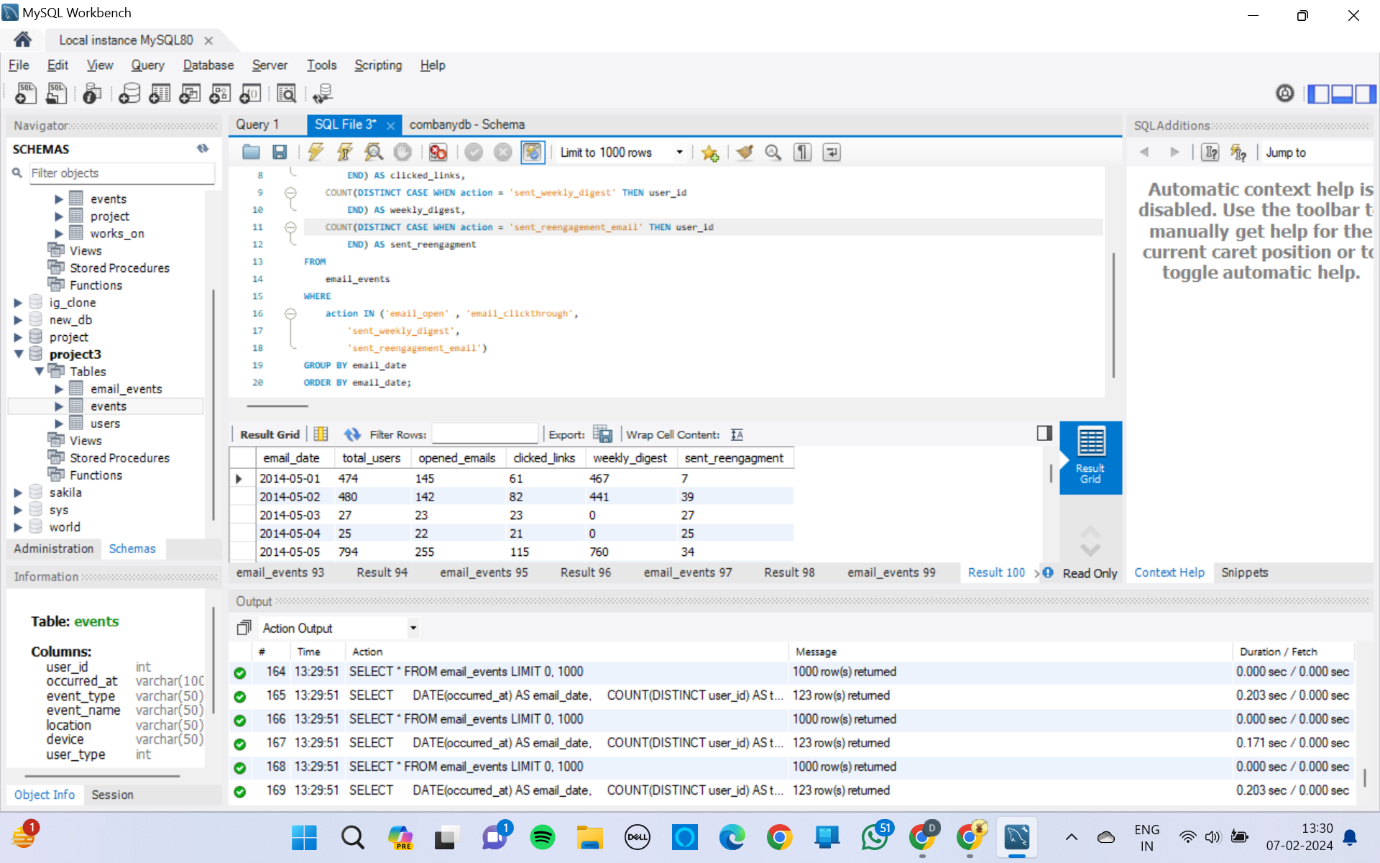
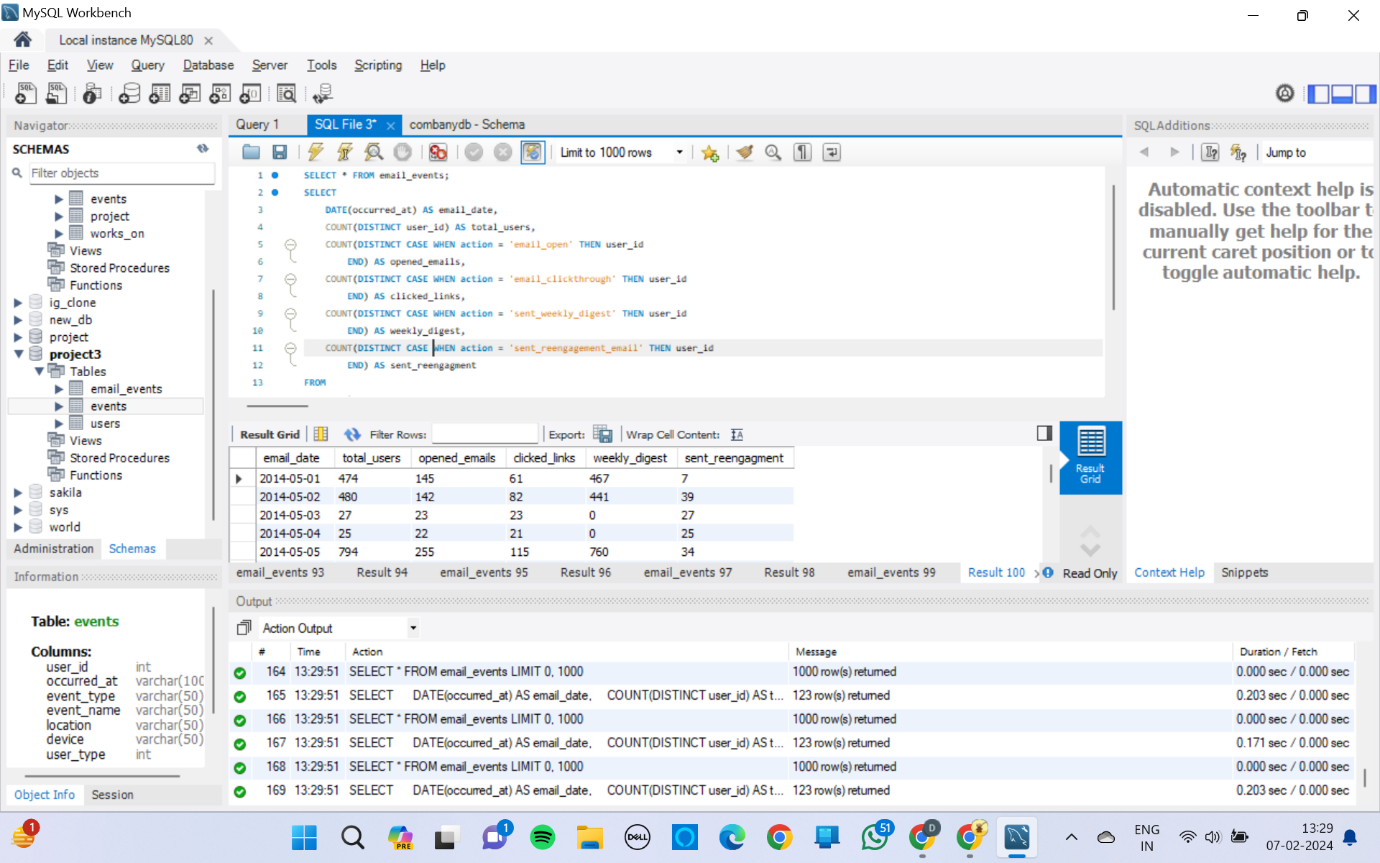
1. **Weekly Retention Analysis:**
   * Objective: Analyze the retention of users on a weekly basis after signing up for a product.
   * Your Task: Write an SQL query to calculate the weekly retention of users based on their sign-up cohort.



1. **Weekly Engagement Per Device:**
   * Objective: Measure the activeness of users on a weekly basis per device.
   * Your Task: Write an SQL query to calculate the weekly engagement per device.



1. **Email Engagement Analysis:**
   * Objective: Analyze how users are engaging with the email service.
   * Your Task: Write an SQL query to calculate the email engagement metrics.



**Result:** Participating in this project presented an invaluable learning experience, enabling me to gain practical skills and enhance my proficiency in SQL and Tableau. Furthermore, the project facilitated my comprehension and acquisition of new business concepts, including Key Performance Indicators (KPIs) and various metrics. I extend my gratitude to the team for their unwavering support and assistance throughout the project. Their guidance and provision of resources greatly facilitated my learning process and contributed to the success of the project. I am genuinely appreciative of the opportunity to be a part of this endeavour.