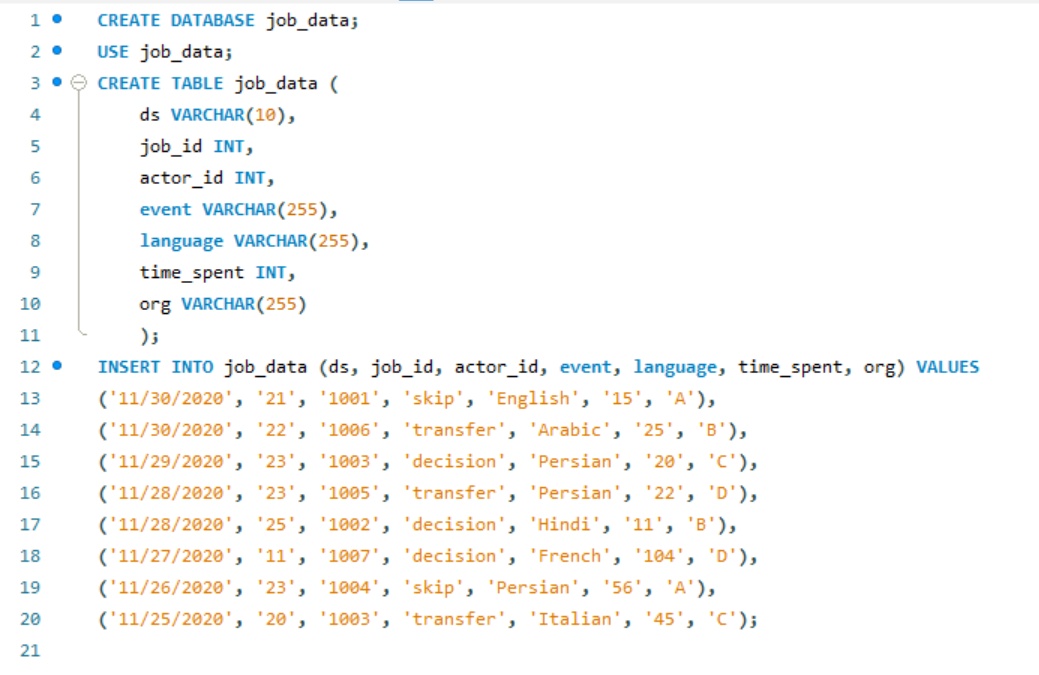
Operation Analytics and Investigating Metric Spike

**Overview:**

Operational analytics is critical for increasing corporate efficiency by evaluating a company's whole operations. The focus of this research is on using data analytics approaches to extract relevant insights from job-related data. The key goals are to discover opportunities for improvement, recognize patterns in job-related metrics, and provide actionable recommendations to other departments, such as operations, support, and marketing.

The research is based on a comprehensive dataset known as "job\_data," which contains key information such as unique identities for jobs and actors, event kinds (decision/skip/transfer), content language, time spent on job evaluations, organizational characteristics, and date stamps. The project's duties include using advanced SQL skills to answer particular questions and identify relevant patterns from the data.

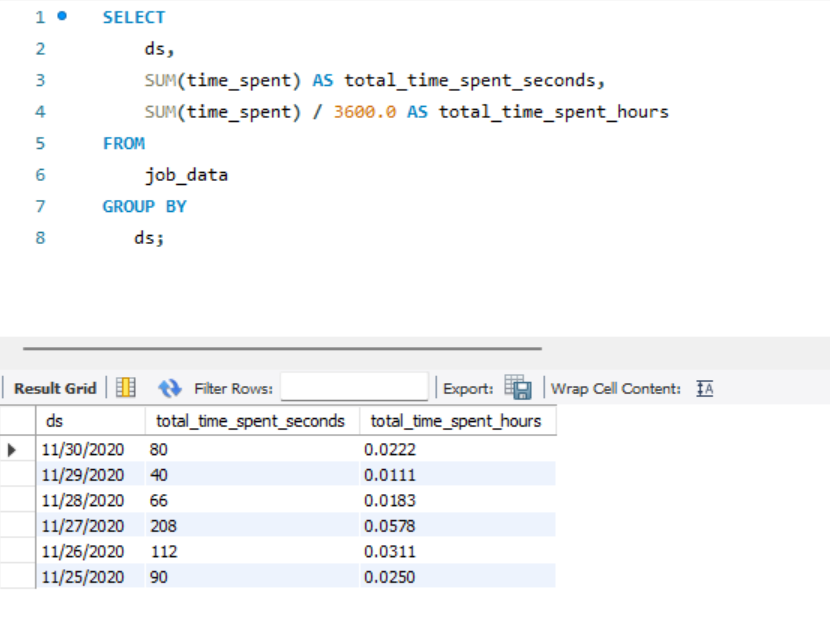


**Approach:**

1. **Jobs Reviewed Over Time:**

Objective: Determine the number of jobs examined each hour for each day in November 2020.

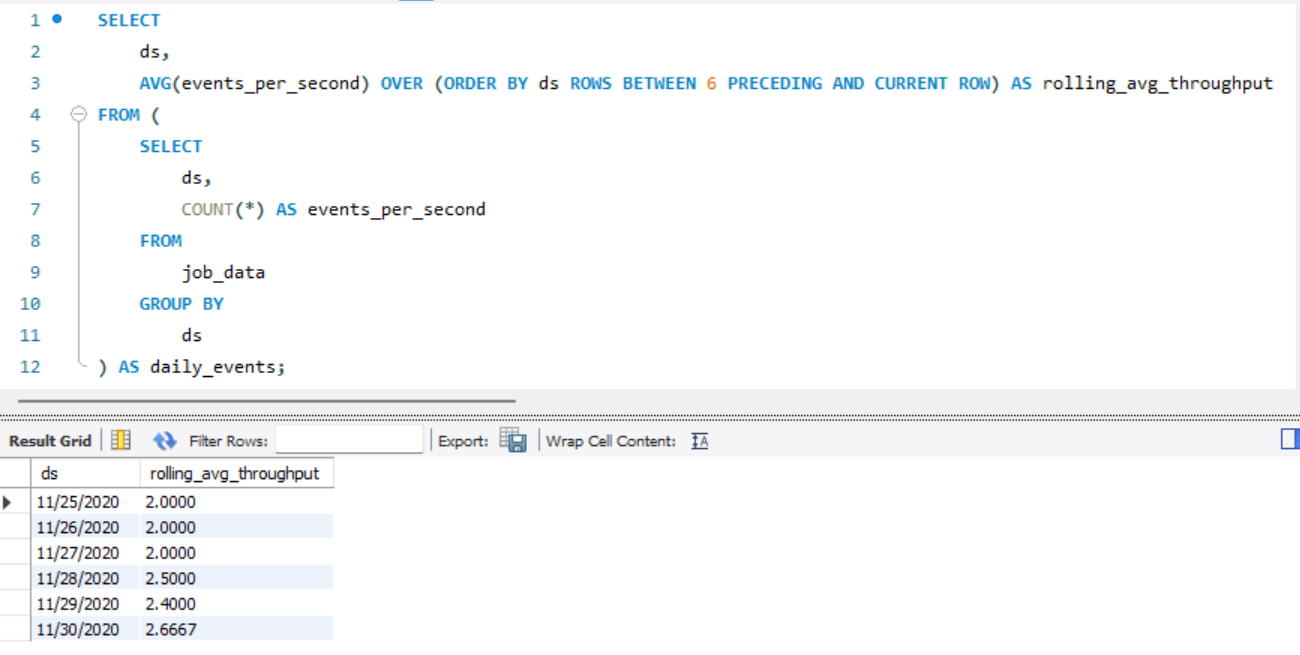
Execution: Used SQL queries to gather and summarize data by date and hour, resulting in a more intricate knowledge of job review patterns during the provided timeframe.



1. **Throughput Analysis:**

Objective: Determine the 7-day rolling average of throughput (events per second).

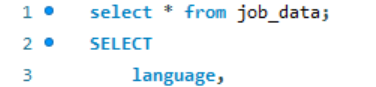
Execution: Used SQL window techniques to smooth out short-term swings in event frequency, resulting in a more stable depiction of trend over time.

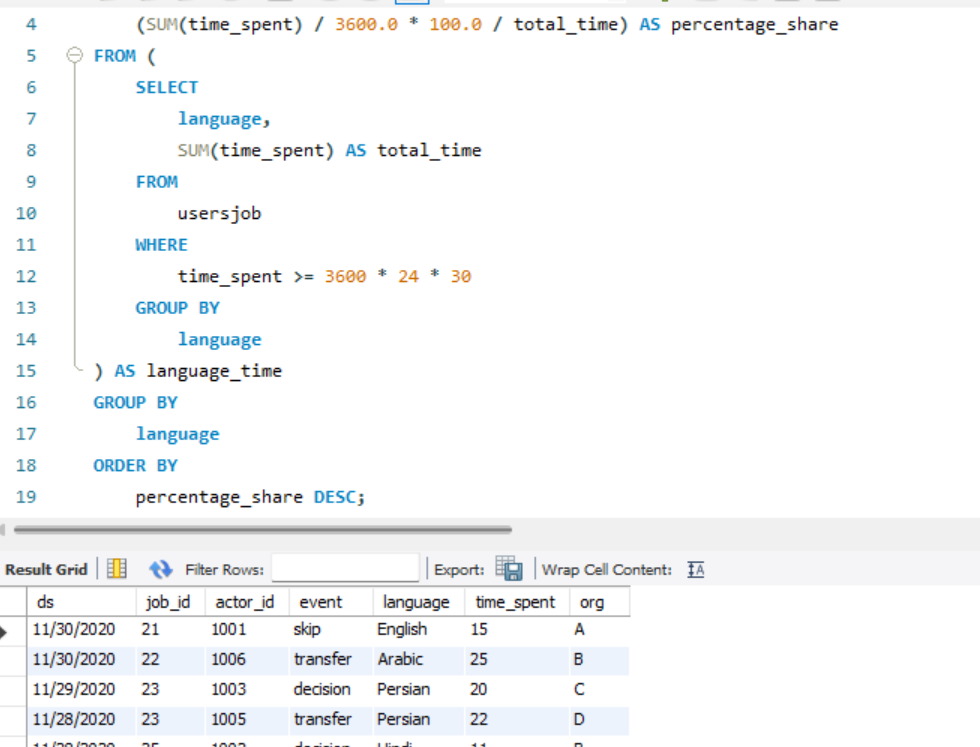


1. **Language Share Analysis:**

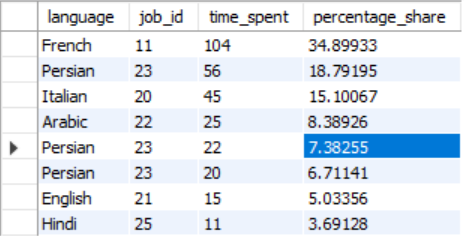
Objective: Calculate the percentage share of each language in the last 30 days.

Execution: Used SQL queries to aggregate language data and establish the distribution of languages, providing insights into language trends.





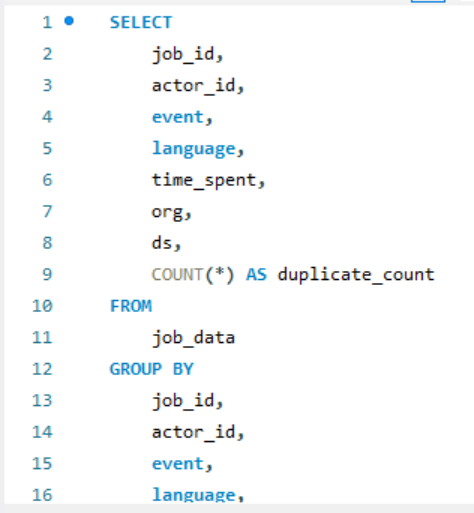
Output:

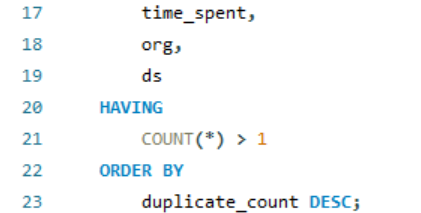


1. **Duplicate Rows Detection:**

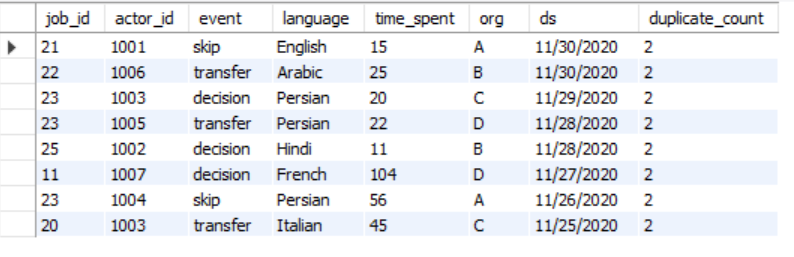
Objective: Identify duplicate rows in the job\_data table.

Execution: Formulated SQL queries to detect and display duplicate rows based on key columns, contributing to data cleanliness and integrity.

****

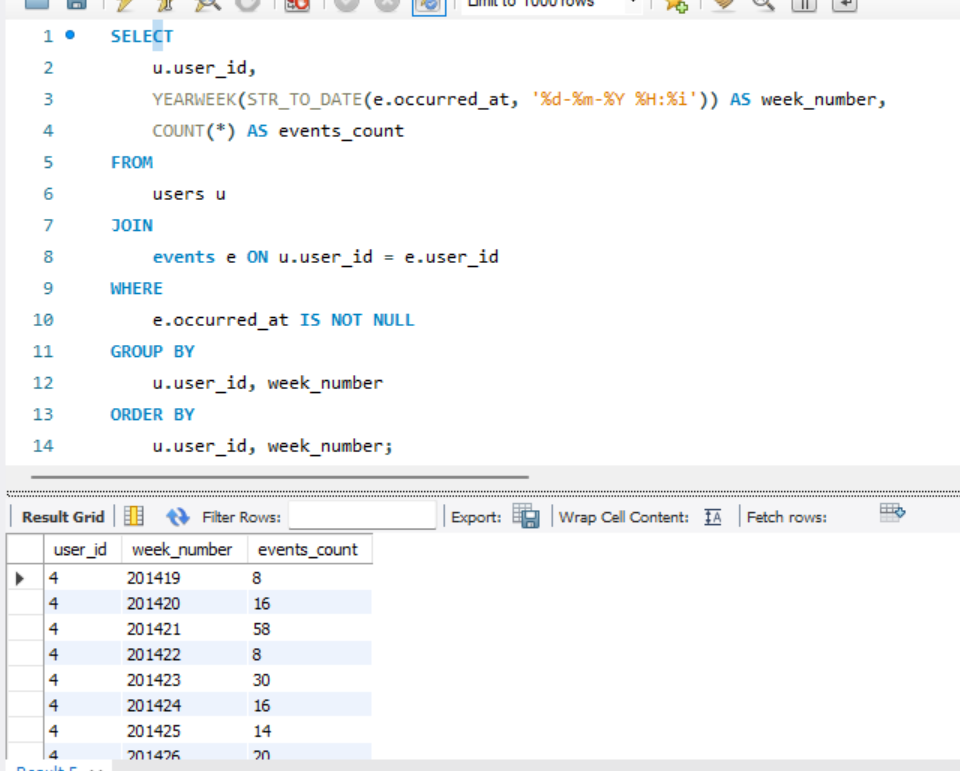


**Output:**



1. **Weekly User Engagement (Case Study 2):**

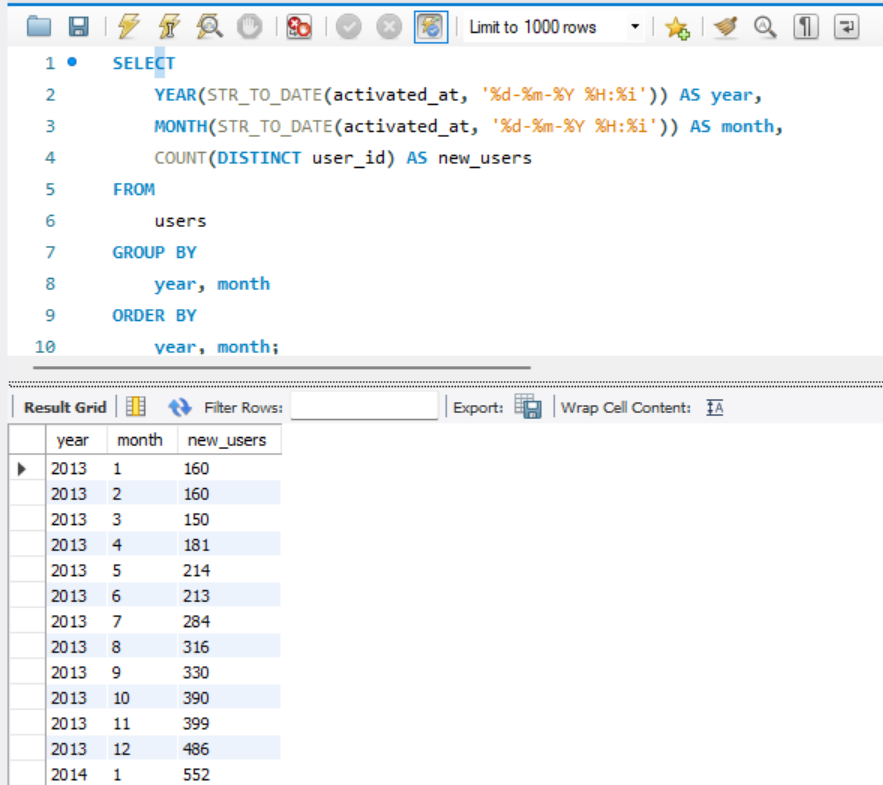
Calculated the weekly user engagement by counting events per user for each week. Presented the results for measuring user activity.



1. **User Growth Analysis:**

Analyzed the growth of users over time for a product.

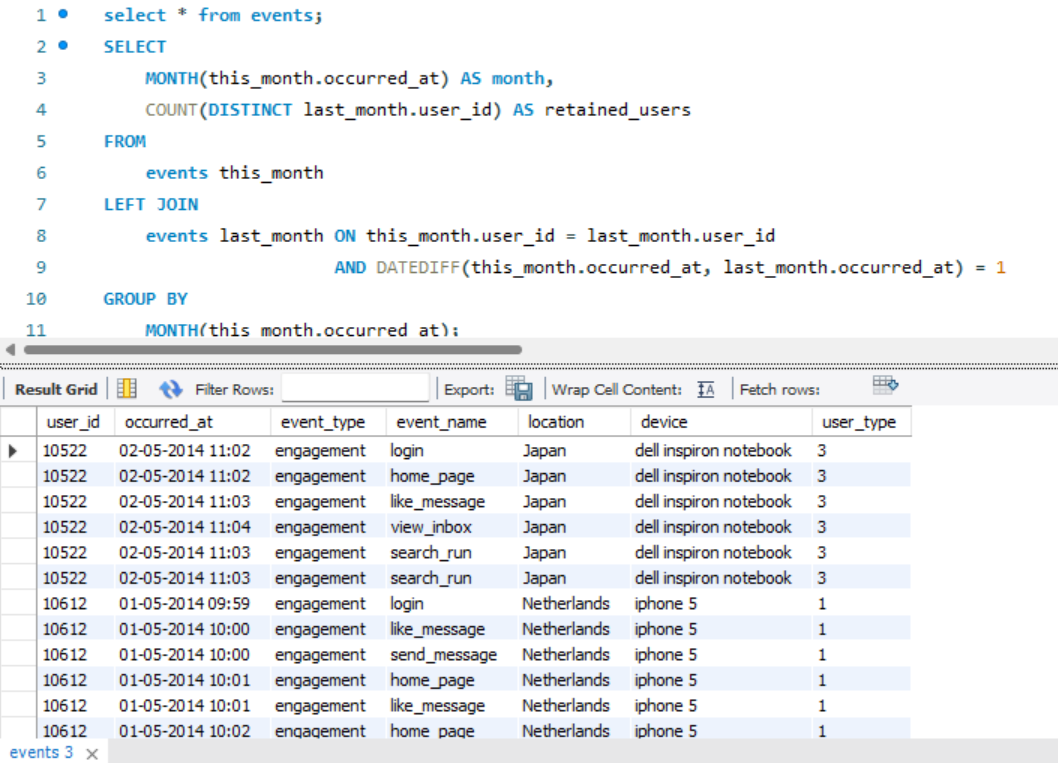
Extracted month-year information from user creation dates to understand user growth patterns.



1. **Weekly Retention Analysis:**

Explored the retention of users on a weekly basis after signing up.

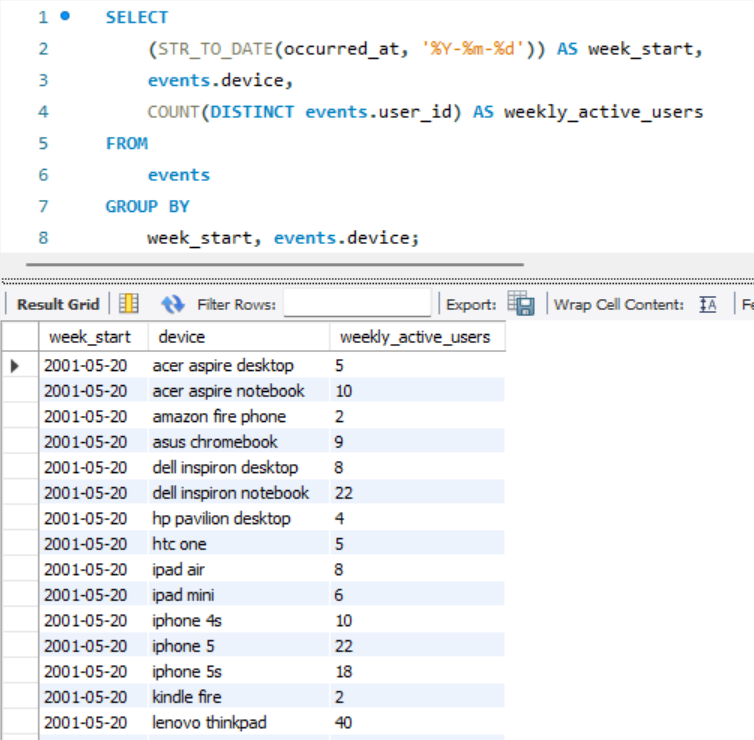
Determined retained users based on cohorts formed by their signup weeks.



1. **Weekly Engagement Per Device:**

Measured user activeness on a weekly basis per device.

Grouped users by activation week and device to analyze engagement patterns.



**Tech-Stack Used:**

* Database Tool: MySQL Workbench
* Query Language: SQL

**Insights:**

1. Identified peak hours for job reviews in November 2020.
2. Recognized the importance of a rolling average in understanding throughput trends.
3. Discovered language distribution patterns over the last 30 days.
4. Detected and visualized duplicate entries within job data.
5. Analyzed user engagement, growth, and retention trends over time.
6. Evaluated device-wise engagement and user interaction with email events.

**Results:**

* Generated actionable insights for optimizing job review processes.
* Provided a comprehensive understanding of user behavior and growth.
* Identified potential areas for operational improvement.
* The project successfully utilized SQL queries to gain actionable insights into various operational aspects.
* The provided analyses can serve as a foundation for data-driven decision-making, ultimately enhancing the company's overall operational efficiency.