Operation Analytics and Investigating Metric Spike

Tool: Advanced SQL

What is Operation Analysis?

Operation analysis, also known as operational analytics, is a powerful approach that involves measuring and analyzing real-time operations within a company to drive informed decision-making and optimize overall performance. Here are a few key points to understand about operation analysis:

- Real-Time Insights: Operational analytics provides the ability to monitor day-to-day operations in real time, allowing companies to gain immediate insights into their processes and performance.
- Proactive Decision-Making: By leveraging operational analytics, organizations can take proactive measures to improve customer satisfaction, enhance efficiency, and ultimately boost their bottom line.
- Predictive Capabilities: Through comprehensive analysis of end-to-end operations, operational analytics enables companies to predict the overall growth or decline of their business, providing valuable foresight for strategic planning.
- Cross-Functional Impact: The insights derived from operational analytics can be utilized by various departments such as operations, support, marketing, and more, fostering better collaboration and effective workflows across the organization.

In essence, operational analytics empowers businesses to harness the power of real-time data to drive continuous improvement, make data-backed decisions, and ultimately achieve greater operational efficiency and success.

Task:

Operational Analytics is a crucial process that involves analyzing a company's end-to-end operations. This analysis helps identify areas for improvement within the company. As a Data Analyst, you'll work closely with various teams, such as operations, support, and marketing, helping them derive valuable insights from the data they collect.

One of the key aspects of Operational Analytics is investigating metric spikes. This involves understanding and explaining sudden changes in key metrics, such as a dip in daily user engagement or a drop in sales.

I was provided with various datasets and tables, and my task was to derive insights from this data to answer questions posed by different departments within the company. My goal was to use my advanced SQL skills to analyze the data and

provide valuable insights that could help improve the company's operations and understand sudden changes in key metrics.

Case Study 1: Job Data Analysis

Table job_data has 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:

A. **Jobs Reviewed Over Time:**
   - Objective: Calculate the number of jobs reviewed per hour for each day in November 2020.

B. **Throughput Analysis:**
   - Objective: Calculate the 7-day rolling average of throughput (number of events per second).

C. **Language Share Analysis:**
   - Objective: Calculate the percentage share of each language in the last 30 days.

D. **Duplicate Rows Detection:**
   - Objective: Identify duplicate rows in the data.

Case Study 2: Investigating Metric Spike

I was working with the following 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:

A. **Weekly User Engagement:**
   - Objective: Measure the activeness of users on a weekly basis.

B. **User Growth Analysis:**
   - Objective: Analyze the growth of users over time for a product.

C. **Weekly Retention Analysis:**
   - Objective: Analyze the retention of users on a weekly basis after signing up for a product.

D. **Weekly Engagement Per Device:**

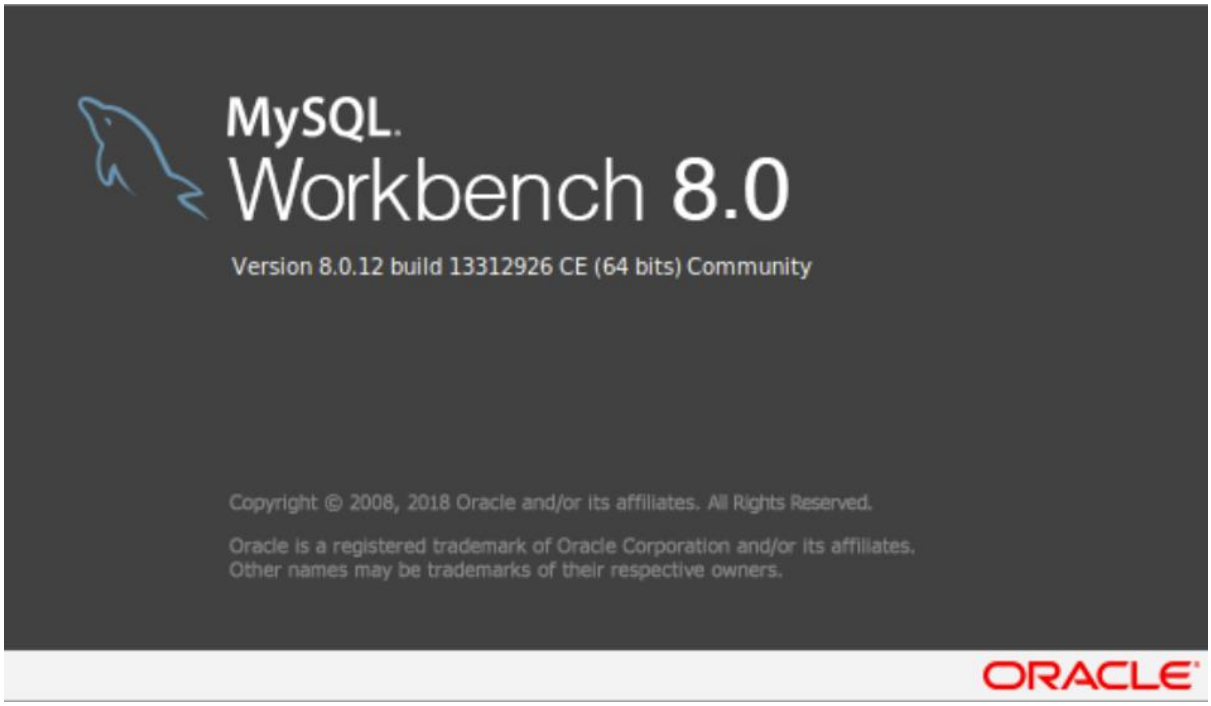
- Objective: Measure the activeness of users on a weekly basis per device.

E. **Email Engagement Analysis:**
   - Objective: Analyze how users are engaging with the email service.

PROJECT APPROACH

The project is executed using SQL Workbench, beginning with the creation of a database from the provided dataset. Subsequently, the data is loaded into SQL Workbench for analysis, aiming to extract valuable insights that can aid the operations, support, and marketing teams. The analysis seeks to address critical questions such as the reasons behind a decline in daily engagement and sales, emphasizing the need to investigate metric spikes regularly to provide timely answers to such inquiries.

TECH STACK USED



MySQL Workbench serves as a comprehensive visual editor that seamlessly integrates data modeling, SQL development, and database administration within a single interface. This tool enables users to intuitively design, generate, and oversee databases.

Widely recognized for its proficiency in handling structured data, MySQL Workbench stands as an open-source Relational Database Management System (RDBMS) developed by Oracle Corporation and Sun Microsystems. Leveraging Structured Query Language (SQL), it facilitates seamless interaction with databases.

Moreover, MySQL Workbench provides convenient database migration options, simplifying the process of transferring data to and from various RDBMS tables, including Microsoft SQL Server and Microsoft Access.

Insights:

Visit the following link to view the overall insights from the above tasks and cases

Operation Analytics and Investigating Metric Spike

RESULT

This project has been instrumental in helping me recognize the significance of operational analytics and how it can be leveraged as a secret weapon by companies. By analyzing metric spikes, businesses can gain valuable insights that enable them to make informed, data-backed decisions, optimize their strategies, and ultimately boost their return on investment.

However, the project was not without its challenges, particularly in dealing with the massive amount of data. To overcome this hurdle, I had to resort to using LOAD DATA statements and modify the datatype of certain columns to ensure smooth data importation.

In conclusion, operational analytics plays a crucial role in synchronizing real-time data and aggregating information from multiple sources to create a comprehensive, actionable solution that delivers analytical models in real-time. This approach ensures that operational routines and systems are utilized efficiently, leading to a significant positive impact on various areas and ultimately contributing to the greater good of society.