**Operation Analytics and Investigating Metric Spike**

**Using Advanced SQL**

**Project Description**

It's important to analyze a company's whole operation as part of the process known as operational analytics. The company can improve in the areas that our study identifies. Working directly with numerous departments, including operations support, and marketing, you will perform the duties of a data analyst while assisting them in gaining insightful information from the data they gather.

Examining metric spikes is one of the major components of operational analytics. For example, a reduction in daily user engagement or spikes in sales are examples of crucial indicators that must be understood and explained when they occur suddenly. Knowing how to look into these metric spikes is vital since as a data analyst, you'll be required to respond to these queries frequently.

For instance, as a Data Analyst Lead for a firm like **META**, I receive various data sets, tables from various social media platforms that they run, and I must draw conclusions from them and respond to inquiries from various departments using various data analysis operational tools.

In this project I have dealt with two case studies which includes some parameters, they are

Case Study – 01 (Job Data Analysis)

* Jobs reviewed over time.
* Throughput Analysis.
* Language share Analysis.
* Duplicate rows detection.

Case Study – 02 (Investigating Metric Spike)

* Weekly User Engagement.
* User Growth Analysis.
* Weekly Retention Analysis.
* Weekly Engagement per device.
* Email Engagement Analysis.

I'll be given a variety of datasets and tables to work with, and my job will be to use SQL queries to obtain those data tables in MySQL Workbench and then analyze them for insights that will help me to respond to inquiries from various departments within the organization.

**Understanding and Proceeding**

First, I took my time to comprehend the data and table that were provided. I answered the questions about what columns are available in the datasets provided to create the table in the database and how to retrieve big data from the datasets provided to MySQL Workbench using the video instructions given by one of the Trainity members, as well as what factors to take into account when reviewing the data. I utilize SQL to extract several insights from the dataset that the management team provided. Using the team's provided structure and links, I first created the database "umeshdb" and then the tables. Then, we conducted analysis to produce insightful data for the business.

**Tech-Stack Used**

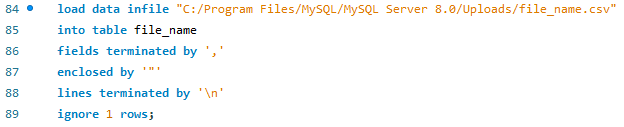
**MySQL Workbench (Version 8.0.34 CE):** MySQL Workbench offers SQL creation, data modeling, and a number of administration tools for configuration. Additionally, a graphical user interface is available for structured database interaction. MySQL makes it simple and cost-free to build databases and conduct analyses to respond to the issues posed in the description.

**Microsoft Excel:** It is used to remove blank cells in the provided datasets.

**Microsoft Word:** It is to create a report (PDF) that will be delivered to the leadership team.

**Insights & Execution**

Big data retrieval to the data table created in a database in MySQL Workbench can be done using the these commands



To reverse or modify the date structure provided in the datasets from dd-mm-yyyy to yyyy-mm-dd along with time constraints was done using this command



This shows the path that where the MySQL server be stored in local system



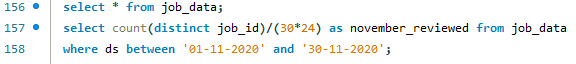


**Case Study – 01**

Here I **will be working with a table named job\_data**

1. **Jobs reviewed over time**: Here my task is to calculate the number of jobs reviewed per hour for each day in November 2020 using SQL queries

Code:



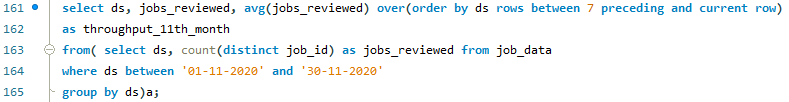
Result:



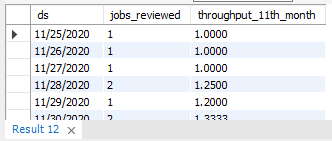
The number of jobs reviewed per hour per day for November 2020 is 83%.

1. **Throughput Analysis: Here I need to perform with the** calculation of 7-day rolling average of throughput.

**Code:**



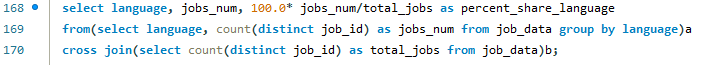
Result:



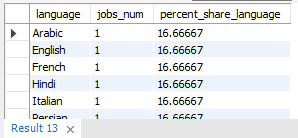
I was given the 7-day rolling average of throughput since it provides the average for each day from day 1 to day 7, unlike daily metrics, which only provide the average for the current day.

1. **Language Share Analysis: My task is to perform the** calculation of percentage share of each language over the last 30 days.

**Code:**



Result:



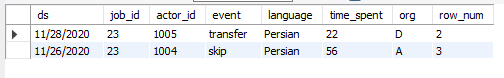
The percentage share of all languages is same.

1. **Duplicate Rows Detection: My job is to find out the duplicate rows in the Table.**

**Code:**



Result:



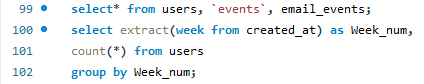
From the table job\_data, we get two duplicate rows. But if we examine the overall columns, each row is distinct.

**Case Study – 02**

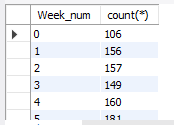
Here I’ll be working with 3 tables named **“users”, “events”** and **“email\_events”**

1. **Weekly User Engagement: Here I need to calculate the** weekly user engagement.

**Code:**



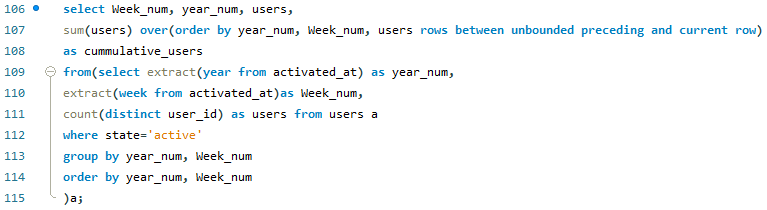
Result:



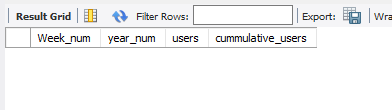
From week 18 to week 34, the weekly user engagement rose, and it then began to fall after that. This indicates that some customers have not found the product to be very high quality in recent weeks.

1. **User Growth Analysis: Here I was calculated the user growth for a particular product**

**Code:**

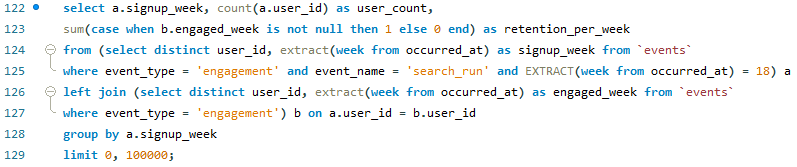


Result:



1. **Weekly Retention Analysis: I need to perform the** Calculation of weekly retention of users-sign up cohort.

**Code:**



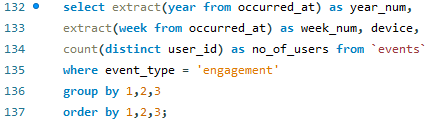
Result:



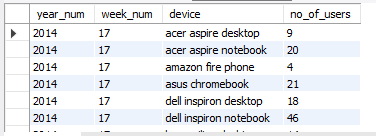
There are in total 966 active users from 18th week.

1. **Weekly Engagement per Device: Here I need to** write an SQL query to calculate the weekly engagement per device.

**Code:**



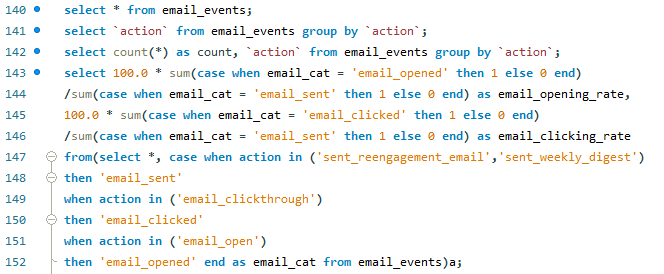
Result:



MacBook and iPhone users have the highest overall weekly involvement per device utilized.

1. **Email Engagement Analysis: Here I need to perform email engagement metrics using SQL queries.**

**Code:**



Result:



33.58% of emails are opened, while 14.78% are clicked. It's encouraging for the company to grow that consumers are actively using the email service.

**Conclusion:** I learned how to use sophisticated SQL concepts, such as Windows Functions, in this project. I was aware of how the real-world economy 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. How to extract the relevant insights that will help the business grow from the given data and queries, including which columns to take into account. I discovered how the business identifies several business-related areas for improvement. I learned about looking at metric spikes, such as why there is a peak and why there is a decline.