

Operation Analytics and Investigating Metric Spike

-by Sandeep Tomar



Contents



- **Introduction**
- **Case Study 1 (Operation Analytics)**
- **Case Study 2 (Investigating metric spike)**
- **Approach**
- **Tech-Stack Used**
- **Insights and Result**

Introduction

In this project, I worked as a Data Analyst Lead for a company like Microsoft, where I was provided with different data sets and tables to derive insights out of them.

The main goal of the project was to perform operation analytics and investigate metric spikes to help different departments of the company make data-driven decisions.

I used SQL to analyze the data and answer the questions asked by different departments.



Case Study 1

(Operation Analytics)



- **01. Number of jobs reviewed.**
Calculate the number of jobs reviewed per hour per day for November 2020?
- **02. Throughput.**
Calculate 7 day rolling average of throughput?
- **03. Percentage share of each language.**
Calculate the percentage share of each language in the last 30 days?
- **04. Duplicate rows.**
How will you display duplicates from the table?



01. Number of jobs reviewed.

date	job_reviewed_per_hour_per_day
2020-11-25	80
2020-11-26	64
2020-11-27	35
2020-11-28	218
2020-11-29	180
2020-11-30	180



02. The 7-day rolling average of throughput.

The weekly throughput is 0.03.



Date	Rolling_Avg_Throughput
2020-11-25	0.02
2020-11-26	0.02
2020-11-27	0.01
2020-11-28	0.06
2020-11-29	0.05
2020-11-30	0.05



03. Percentage share of each language.

Language	Percentage_Share
Persian	37.5
Italian	12.5
Hindi	12.5
French	12.5
English	12.5
Arabic	12.5



04. How will you display duplicates from the table?

Ans.

```
SELECT job_id, COUNT(*) AS num_duplicates  
FROM job_data  
GROUP BY job_id  
HAVING COUNT(*) > 1;
```

job_id	num_duplicates
23	3



Case Study 2

(Investigating metric spike)



- **01. User Engagement.**
Calculate the weekly user engagement?
- **02. User Growth.**
Calculate the user growth for product?
- **03. Weekly Retention.**
Calculate the weekly retention of users-sign up cohort?
- **04. Weekly Engagement.**
Calculate the weekly engagement per device?
- **05. Email Engagement.**
Calculate the email engagement metrics?





01. User Engagement.

Ans.

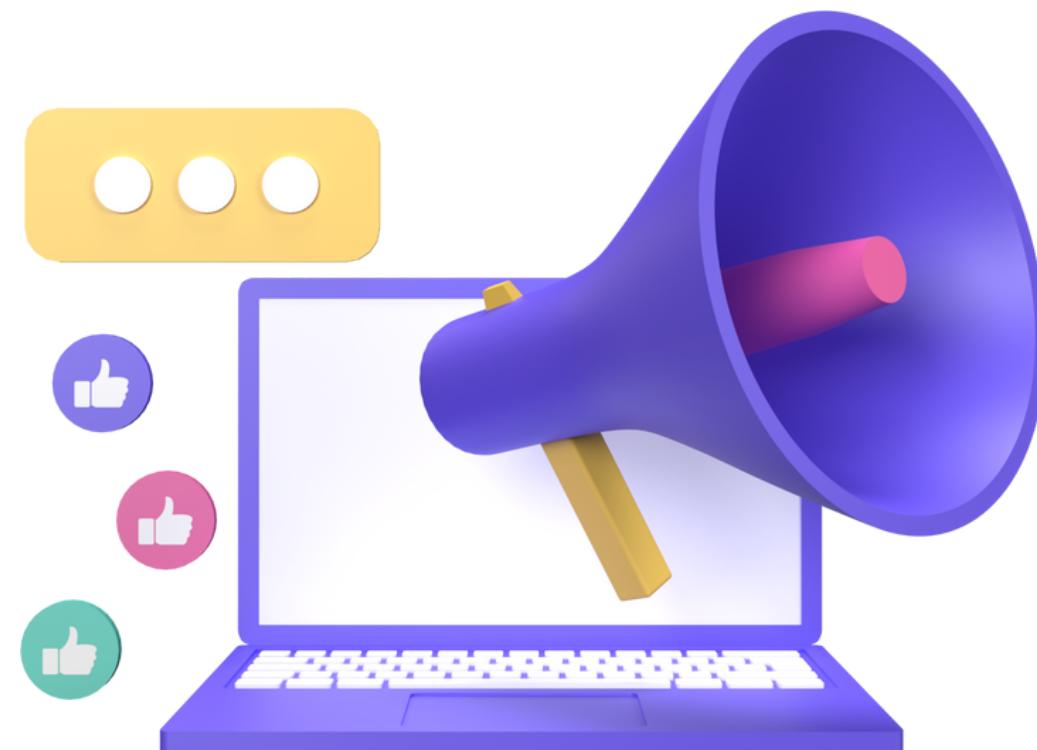
SELECT

```
extract(week from occurred_at) AS week,  
COUNT(DISTINCT user_id) AS weekly_active_users
```

FROM events

Where event_type = "engagement"

GROUP BY week;



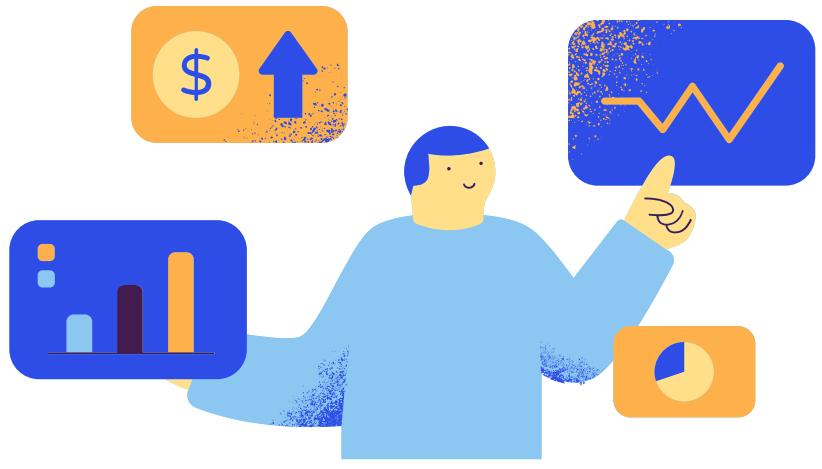
week	weekly_active_users
null	6142

02. User Growth.



month	users	growth in %
1	712	null
2	685	-3.79
3	765	11.68
4	907	18.56
5	993	9.48
6	1086	9.37





month	users	growth in %
7	1281	17.96
8	1347	5.15
9	330	-75.50
10	390	18.18
11	399	2.31
12	486	21.80

03. Weekly Retention



Ans.

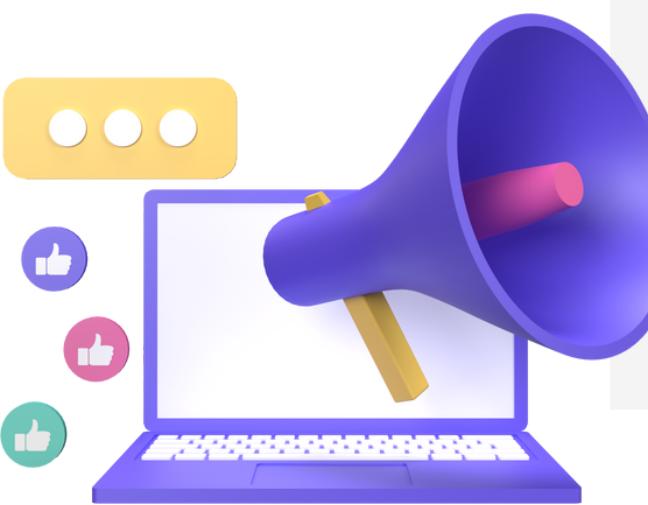
```
SELECT  
    extract(week from occurred_at) AS week,  
    COUNT(DISTINCT user_id) AS users_sign_up_cohort  
FROM events  
Where event_type = "signup_flow"  
GROUP BY week;
```

week	users_sign_up_cohort
null	7298

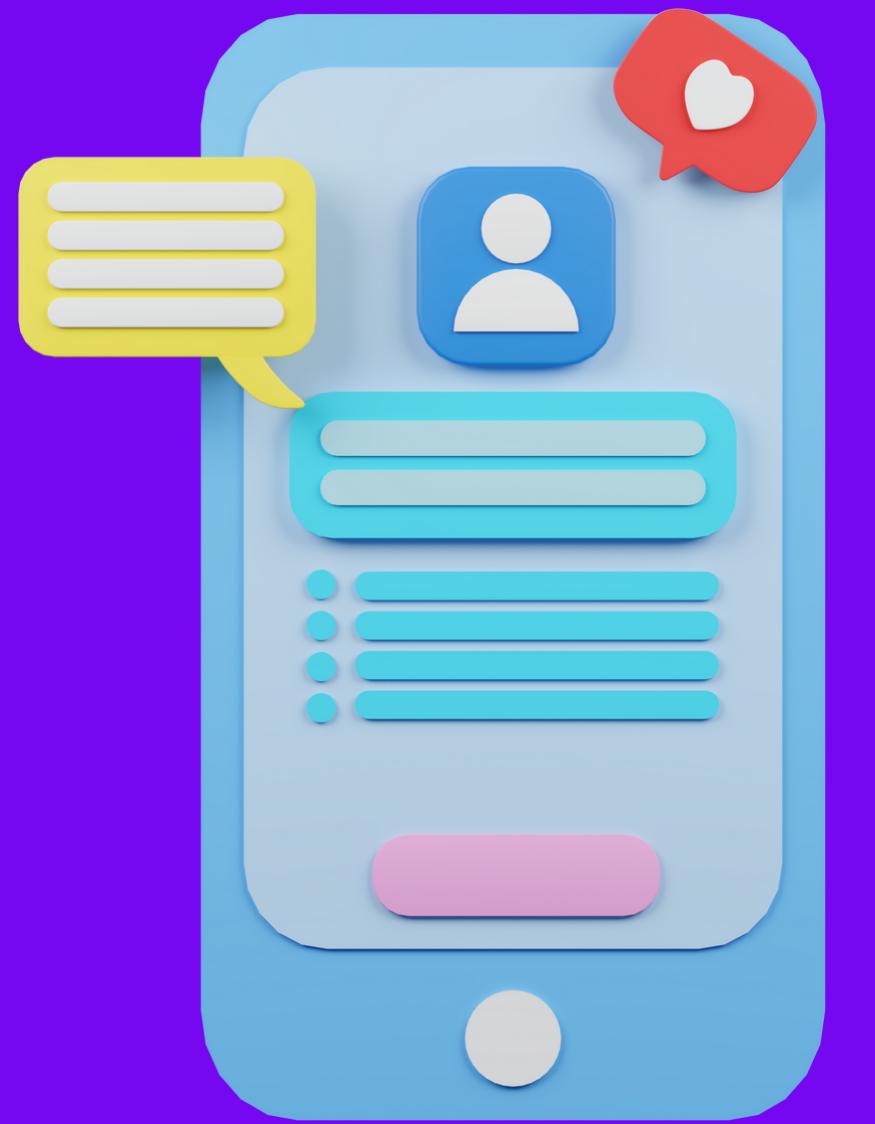
04. Weekly Engagement Per Device



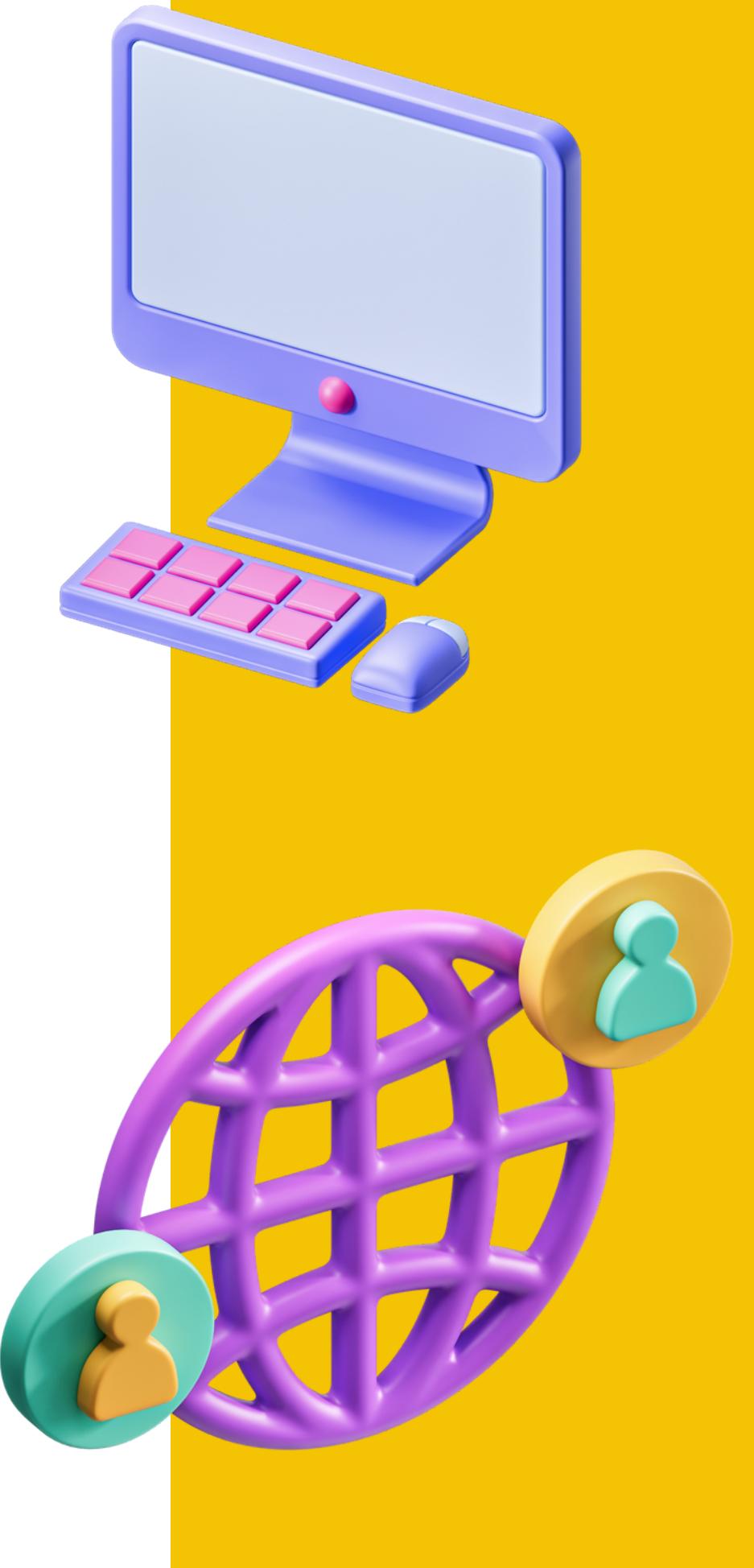
week	device	engagement_per_device
null	acer aspire desktop	198
null	acer aspire notebook	338
null	amazon fire phone	89
null	asus chromebook	355
null	dell inspiron desktop	360
null	dell inspiron notebook	677

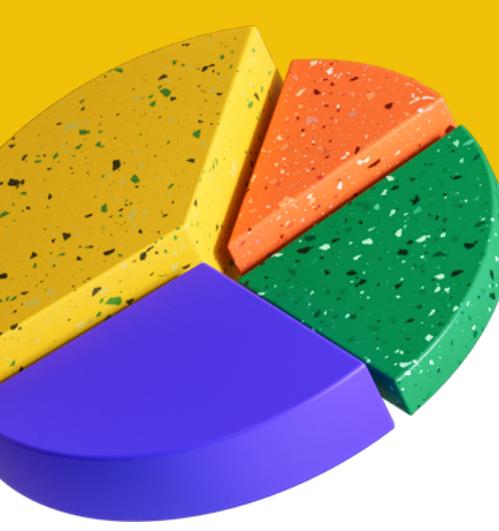


week	device	engagement_per_device
null	hp pavilion desktop	339
null	htc one	196
null	ipad air	478
null	ipad mini	292
null	iphone 4s	409
null	iphone 5	1025
null	iphone 5s	626




week	device	engagement_per_device
null	kindle fire	205
null	lenovo thinkpad	1309
null	mac mini	150
null	macbook air	950
null	macbook pro	1952
null	nexus 10	273
null	nexus 5	621





week	device	engagement_per_device
null	nexus 7	355
null	nokia lumia 635	211
null	samsung galaxy tablet	107
null	samsung galaxy note	119
null	samsung galaxy s4	803
null	windows surface	182



05. Email Engagement Metrics.



`weekly_digest_rate`

63.36

`email_opened_rate`

22.63

`email_clickthrough_rate`

9.97

`reengagement_email_rate`

4.04



Approach

I started by learning about the structure of the tables and the definitions of each column. Then I created a database and tables following the framework and links given.

I applied SQL to analyze and respond to the case study questions.

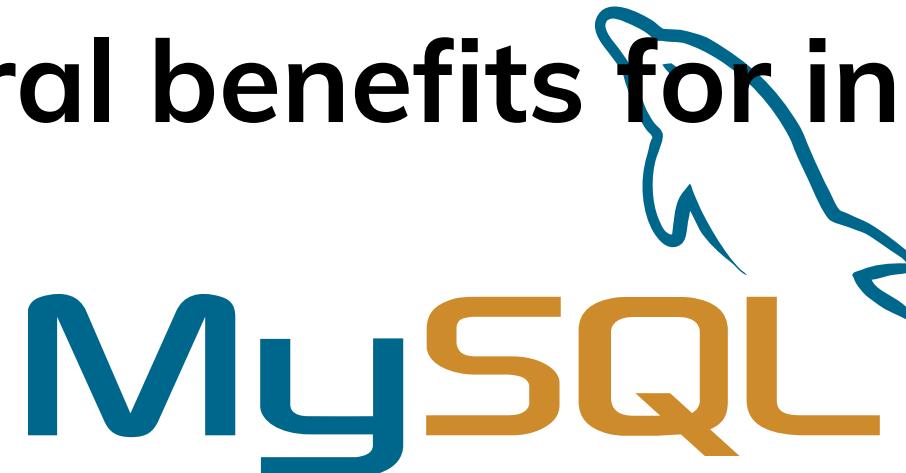
I calculated the required metrics and generated insights from the data using various SQL methods and techniques.



Tech-Stack Used



In this project, I used SQL to perform data analysis on the database. And the software used for this project is MySQL Workbench version 8.0.27 because MySQL is a popular relational database management system that offers several benefits for individuals.

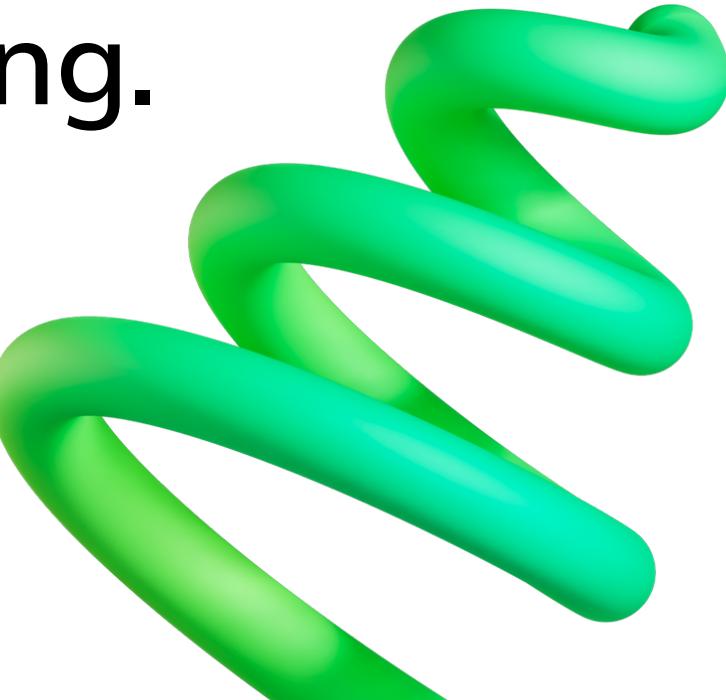


Insights and Result



Through this project, I was able to deliver useful insights and analytics to various departments inside the company.

The company may enhance all aspects of its operations, forecast growth or decline, and have more efficient workflows by adopting data-driven decision-making.





Thank You