Operational Analytics and Metric Spike Investigation Using SQL

Project Overview

This project focuses on analyzing operational data and identifying unusual activity patterns, known as metric spikes, using SQL. As the lead analyst, my task was to collaborate with various departments and derive actionable insights to support data-driven decisions. I utilized advanced SQL techniques such as joins, subqueries, rolling averages, and aggregations to perform comprehensive data analysis and uncover operational trends.

Analytical Workflow

1. Database Setup:

I began by setting up the MySQL database to accommodate the required datasets and queries for analysis.

2. Data Ingestion:

All relevant datasets were imported into the database for structured querying and manipulation.

3. Insight Generation:

Using SQL queries, I analyzed the datasets to address specific business questions and reveal meaningful insights.

Tech-Stack Used :

1. **MySQL**

Case Study 1: Job Data Analysis

```
SELECT DATE_FORMAT(ds, '%Y-%m-%d %H:00:00') AS hour,
 (sum(time_spent)/count(*)) as Time_spent_per_hour_per_job
FROM jobs
WHERE MONTH(ds) = 11 AND YEAR(ds) = 2020
```

hour	Time_spent_per_hour_per_job
30-11-2020 00:00	20
29-11-2020 00:00	20
28-11-2020 00:00	16.5
27-11-2020 00:00	104
26-11-2020 00:00	56
25-11-2020 00:00	45

Insight:

Review times fluctuated, with notable peaks (e.g., 104 seconds on 27-Nov-2020), indicating potential workload issues or process inefficiencies.

B. Throughput Analysis

SELECT SUM(time_spent) / TIMESTAMPDIFF(SECOND, MIN(ds), MAX(ds)) AS throughput

FROM jobs;

throughput
0.0007

Insight:

The dataset spans only 5 days, which limits the ability to calculate a reliable 7-day moving average. More data would be needed for smoother trends.

C. Language Distribution

```
select language, (count(*)/(select count(*) from jobs))*100 as perc
from jobs
group by language;
```

language	perc
English	12.5
Arabic	12.5
Persian	37.5
Hindi	12.5
French	12.5
Italian	12.5

Insight:
Persian dominated the content language with a 37.5% share. All other languages had equal representation (12.5% each), which might indicate regional focus or translation delays.

D. Duplicate Entries Check

```
select ds

from jobs

GROUP BY  ds

HAVING COUNT(*) > 1;
```

DS
30-11-2020 00:00
28-11-2020 00:00

Insight:
While no complete duplicate rows exist, some dates have multiple records—e.g., 30th and 28th Nov 2020—suggesting peak workload or recurring entries.

Case Study 2: Metric Spike Analysis

Weekly User Engagement Rate

```
SELECT

(SELECT COUNT(DISTINCT(user_id)) AS active_users

FROM users

WHERE state = 'active') / (SELECT COUNT(DISTINCT(user_id)) AS active_users

FROM users) * 100 as user_engagement_rate

FROM users
```

Engagement Rate = 49.2%

Insight:
Nearly half the users are actively engaged, highlighting a need to improve product stickiness or onboarding processes.

B. Device-Based User Growth

```
SELECT
new_users.device,
avg(new_users.new_users / first_month.total_users*100) AS growth_rate
FROM (
SELECT
device,
month(occurred_at) AS signup_month,
COUNT(DISTINCT user_id) AS new_users
FROM events
GROUP BY device, signup_month
) AS new_users
JOIN (
SELECT
device,
COUNT(DISTINCT user_id) AS total_users
FROM events
GROUP BY device
) AS first_month
ON new_users.device = first_month.device
group by new_users.device;
```

device	growth_rate
acer aspire desktop	36.49
acer aspire notebook	35.58
amazon fire phone	32.2
asus chromebook	34.89
dell inspiron desktop	35.56
dell inspiron notebook	36.38
hp pavilion desktop	35.3
htc one	33.37
ipad air	33.14
ipad mini	32.34
iphone 4s	34.7
iphone 5	35.11
iphone 5s	33.99
device	**growth_rate**
kindle fire	32.64
lenovo thinkpad	36.15
mac mini	39.41
macbook air	35.63
macbook pro	35.86
nexus 10	33
nexus 5	34.28
nexus 7	32.81
nokia lumia 635	36.25
samsumg galaxy tablet	31.85
samsung galaxy note	33.67
samsung galaxy s4	33.64
windows surface	32.35

Insight:
Mac Mini led the growth rate at 39.41%, followed by various popular devices. This information could guide device-specific product optimizations.

C. Weekly Retention Trends

```
SELECT
WEEK(occurred_at) AS week,
COUNT(DISTINCT user_id) AS signups,
COUNT(DISTINCT CASE WHEN event_type = 'engagement' THEN user_id END) AS
engagements,
COUNT(DISTINCT CASE WHEN event_type = 'engagement' THEN user_id END) /
COUNT(DISTINCT user_id) *100 AS retention_rate
FROM events
GROUP BY WEEK(occurred_at)
ORDER BY week;
```

week	signups	engagements	retention_rate
17	740	663	89.5946
18	1260	1068	84.7619
19	1287	1113	86.4802
20	1351	1154	85.4182
21	1299	1121	86.2972
22	1381	1186	85.8798
23	1446	1232	85.2006
24	1471	1275	86.6757
25	1459	1264	86.6347
26	1509	1302	86.2823
27	1573	1372	87.2219
28	1577	1365	86.5568
29	1607	1376	85.6254
30	1706	1467	85.9906
31	1514	1299	85.7992
32	1454	1225	84.2503
33	1438	1225	85.1878
34	1443	1204	83.4373
35	118	104	88.1356

Insight:
Retention rates remained consistently above 83%, peaking at 89.6% in Week 17. This reflects healthy user satisfaction and re-engagement.

D. Weekly Engagement Per Device

```
SELECT

WEEK(occurred_at) AS week,

COUNT(DISTINCT CASE WHEN event_type = 'engagement' THEN user_id END) /
COUNT(DISTINCT user_id)/(select count(distinct(device)) from events) * 100 AS
engagement_rate_per_device

FROM events

GROUP BY WEEK(occurred_at
```

week	engagement_rate_per_device
17	3.44594594
18	3.26007326
19	3.32616102
20	3.28531572
21	3.31912122
22	3.30306912
23	3.27694435
24	3.33368195
25	3.33210312
26	3.31855023
27	3.35468727
28	3.3291059
29	3.29328419
30	3.30733159
31	3.29996951
32	3.24039784
33	3.27645233
34	3.20912628
35	3.38983051

Insight:

Engagement hovered around 3.2–3.4%, showing consistent usage across different weeks and devices. A higher rate would indicate better device optimization.

E. Email Interaction Analysis

```
SELECT
engagements,
total_users,
engagements/total_users*100 AS engagement_rate
from (  SELECT
count(distinct(user_id)) AS total_users,
COUNT(DISTINCT
CASE
WHEN action = 'email_open' THEN user_id
WHEN action = 'email_clickthrough' THEN user_id END)  AS engagements
FROM email_events) as counts
```

engagements	total_users	engagement_rate
5927	6179	95.9217

Insight:
Email campaigns were highly successful in drawing user interaction—an excellent metric for marketing and CRM teams.