Operation Analytics and Investigating Metric Spike

BY PRERNA

DESCRIPTION

- The analysis of an organization's whole end-to-end operations is commonly referred to as operation analytics. The organization can then employ this to identify the areas where it needs to improve. I collaborate closely with the operations, support, marketing, and other teams to gain insights from the information they acquire.

- This type of analysis is utilized to anticipate the general rise or fall of an organization's future because it counts as one of its biggest and most significant elements. It indicates greater automation, more understanding among teams from various departments, and more efficient workflows.

- Analyzing metric spikes also serves as an essential component of operational analytics. As a Data Analyst, you must additionally be able to grasp or enable other teams to understand queries such as, "The reason why is there a drop in daily engagement?" The reason why have sales dropped? Etc. This kind of question must be answered regularly, which necessitates the investigation of metric spikes.

- I'm a data analyst lead for an organization like Microsoft, and my responsibilities include analyzing various data sets and tables to draw conclusions and respond to inquiries from various departments.

APPROACH

The investigation of metric spikes is a significant component of Operation Analytics. Understanding and resolving queries about variations in metrics such as daily sales and customer engagement is crucial for a Data Analyst. This entails regular investigation and analysis of metric spikes to discover the factors behind these changes.

SQL may be a great tool for performing Operation Analytics and investigating metric spikes. SQL enables data analysts to efficiently query and analyze big databases. Data may be retrieved, processed, and aggregated by executing SQL queries to acquire insights into many operational areas of the organization. This involves finding patterns, and trends in data that can give significant insights for making choices and implementing changes.

TECH-STACK USED

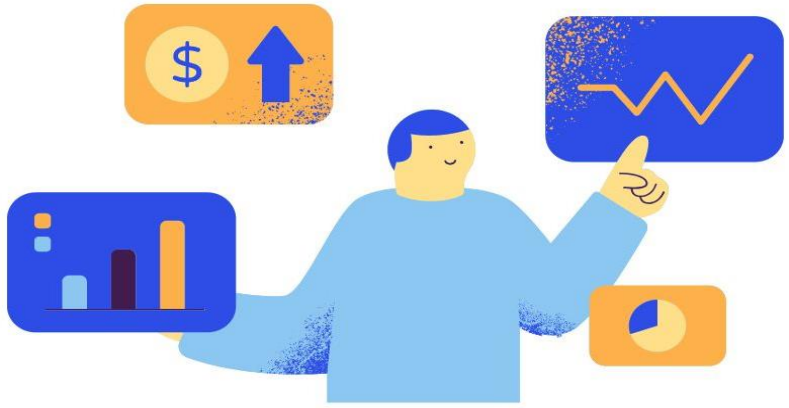
MySQL Workbench 8.0 CE enables you to build, create, and look through your database schemas, interact with database components, and insert data. It also enables you to create and execute SQL queries to interact with stored data.

OPERATION ANALYTICS

- To carry out this operation, the necessary information from various sources is taken and transformed into a format that can be analyzed.
- Specific analyses should be performed to address operational difficulties and objectives.
- SQL queries may be used to collect and analyze data and pinpoint particular metrics.

INVESTIGATING THE METRIC SPIKES

- By examining the relevant information, consider potential causes or events that may have influenced the rise.
- Determining patterns or trends in the results should be done through comparative analysis of data.
- Share findings with the various stakeholders, illustrating or outlining recommendations based on what has been observed.



Case Study 1 (Job Data)

- **Table-1: job_data**
  - **job_id:** unique identifier of jobs
  - **actor_id:** unique identifier of actor
  - **event:** decision/skip/transfer
  - **language:** language of the content
  - **time_spent:** time spent to review the job in seconds
  - **org:** organization of the actor
  - **ds:** date in the yyyy/mm/dd format. It is stored in the form of text and we use presto to run. no need for date function

INSIGHTS

Case Study 1 (Job Data)

Number of jobs reviewed

```
select count(distinct job_id)/(30*24) as jobs_per_day
from job_data
where ds >= '2020-11-01' and ds<= '2020-11-30';
```

jobs_per_day
0.0083

According to the data, it seems that throughout November only a small percentage of jobs were assessed every hour of the day. the precise figure is less than 0.01 jobs.

Throughput

Weekly Throughput

```
select
num_of_events/total_time_spent as week_throughput
from(select count(event) as num_of_events,
sum(time_spent) as total_time_spent
from job_data) j
```

week_throughput
0.0268

The weekly Throughput is 0.0268
The highest daily Throughput is 0.0606

Daily Throughput

```
select ds,
num_of_events/total_time_spent as daily_throughput
from(select ds,
count(event) as num_of_events,
sum(time_spent) as total_time_spent
from job_data
group by ds) j
```

ds	daily_throughput
30-11-2020	0.05
29-11-2020	0.05
28-11-2020	0.0606
27-11-2020	0.0096
26-11-2020	0.0179
25-11-2020	0.0222

Percentage share of each language

```
select language,count(language),
COUNT(*) * 100.0 / SUM(COUNT(*)) over() as percentage
from job_data
group by language
order by language
```

language	count(language)	percentage
Arabic	1	12.5
English	1	12.5
French	1	12.5
Hindi	1	12.5
Italian	1	12.5
Persian	3	37.5

According to the data, Persian appears to rank as the most spoken language with a percentage of 37.5 amongst the various other languages listed.

Duplicate Rows

```
select ds,job_id,actor_id,event,language,time_spent,org,count(*)
from job_data
group by ds,job_id,actor_id,event,language,time_spent,org
having count(*)>1
```

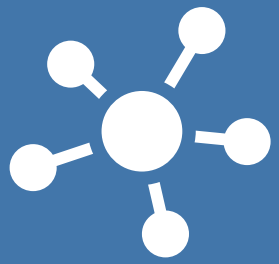
ds	job_id	actor_id	event	language	time_spent	org	count(*)

According to the desired outcomes, it is possible to conclude that there were no duplicates found. The lack of duplicate records indicates accuracy and dependability, enabling the process of decision-making.

OVERVIEW

Invest in the continual development of knowledge and abilities to increase your job competitiveness in the workplace.

To speed up processes, pinpoint areas that might benefit from automated technology advances.



Survey or get suggestions from your intended audience to better comprehend their cultural preferences and needs.

Continuous development promotes constant data precision as well as dependability while enabling the uncovering of opportunities for improvement.

Case Study 2 (Investigating metric spike)

- **Table-1: users**

This table includes one row per user, with descriptive information about that user's account.

- **Table-2: events**

This table includes one row per event, where an event is an action that a user has taken. These events include login events, messaging events, search events, events logged as users progress through a signup funnel, and events around received emails.

- **Table-3: email_events**

This table contains events specific to the sending of emails. It is similar in structure to the events table above.

User Engagement

```sql
Select
EXTRACT(WEEK from occurred_at) as weeks,
count(distinct user_id) as weekly_active_users
from events
where event_type = 'engagement'
and event_name = 'login'
group by weeks
order by weeks;
```

weeks	weekly_active_users
17	85
18	93
19	14
20	6
21	2
22	1
23	2

There were around 85 active users in week 17, which indicates a respectable level of involvement. With 93 active users in week 18, this figure grew slightly, indicating a healthy trend in user engagement. Just 14 active users were present in week 19, a considerable decrease.

Over the past several weeks, the drop in active users raises concerns regarding user retention and engagement.

A more thorough study and research are required to find the underlying issues and create plans to boost user engagement and boost the number of active users.

User Growth

```
select year,week,users,sum(users) over(order by year,week) as user_growth
from
(
SELECT extract(year from created_at) as year,extract(week from activated_at) as
week,count(distinct user_id) as users
FROM users
where state = 'active'
group by year,week
order by year,week
) a
```

year	week	users	usergrowth
2013	17	49	680
2013	18	44	724
2013	19	57	781
2013	20	39	820
2013	21	49	869
2013	22	54	923
2013	23	50	973
2013	24	45	1018
2013	25	57	1075
2013	26	56	1131
2014	20	176	6297
2014	21	183	6480
2014	22	196	6676
2014	23	196	6872
2014	24	229	7101
2014	25	207	7308
2014	26	201	7509
2014	27	222	7731
2014	28	215	7946
2014	29	221	8167
2014	30	238	8405
2014	31	193	8598
2014	32	245	8843
2014	33	261	9104
2014	34	259	9363
2014	35	18	9381

activated-users
9381

In 2013, user growth begins with 49 users in week 17 and gradually climbs over the following weeks. Starting with week 20, 2014, there is a considerable increase in user numbers. The data concludes in week 35, with an enormous decrease to 18 users.
The data ends at week 35 when there is a considerable reduction to 18 users.

Weekly Retention

```
SELECT extract(YEAR FROM OCCURRED_AT)AS YEAR,
extract(WEEK FROM OCCURRED_AT)AS WEEK,
DEVICE,
count(DISTINCT user_id) as USERS
from events
WHERE EVENT_TYPE ='ENGAGEMENT'
GROUP BY year,week,device
ORDER BY year,week,device
```

YEAR	WEEK	DEVICE	USERS
2014	17	mac mini	1
2014	17	macbook air	4
2014	17	macbook pro	13
2014	17	nexus 5	4
2014	17	nexus 7	4
2014	17	nokia lumia 635	2
2014	17	samsumg galaxy tablet	2
2014	17	samsung galaxy note	1
2014	17	samsung galaxy s4	7
2014	18	acer aspire notebook	1
2014	18	ipad air	3
2014	18	ipad mini	4
2014	18	iphone 4s	1
2014	18	iphone 5	3
2014	18	iphone 5s	1
2014	18	kindle fire	1
2014	18	lenovo thinkpad	14
2014	18	mac mini	1
2014	18	macbook air	12
2014	18	macbook pro	26
2014	18	nexus 10	1
2014	18	nexus 5	5

On week 18 of 2014, most users signed up using a MacBook Pro device.

Weekly Engagement

week	device	engagement_count
18	macbook	26
18	lenovo	14
17	macbook	13
18	macbook	12
17	iphone 5	11
17	lenovo	8
17	samsung	7
18	samsung	7
17	iphone	5
18	nexus 5	5

```
select
extract(week from occurred_at) as week,
device,
count(distinct user_id) as users
from events
group by week,device
order by users desc;
```

The "Macbook Pro" has the highest amount of engagement with 26 users. It displays the devices with the highest rates of engagement throughout particular weeks, showcasing the devices that grabbed users' attention and resulted in more interactions.

Email Engagement

year	month	action	num_of_emails
2014	8	sent_weekly_digest	16480
2014	7	sent_weekly_digest	15902
2014	6	sent_weekly_digest	13155
2014	5	sent_weekly_digest	11730
2014	8	email_open	5011
2014	7	email_open	4781
2014	6	email_open	3869
2014	5	email_open	3656
2014	7	email_clickthrough	1962
2014	6	email_clickthrough	1565
2014	5	email_clickthrough	1519
2014	8	email_clickthrough	1123
2014	5	sent_reengagement	142

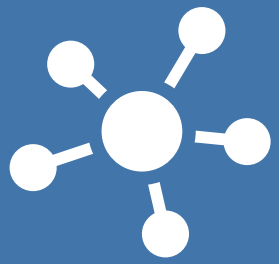
```
select
extract(year from occurred_at) as year,
extract(month from occurred_at) as month,
action,
count(action) as num_of_emails
from email_events
group by year,month,action
order by num_of_emails desc;
```

In August, users received the highest amount of weekly digest emails.
It displays the number of emails sent, Analyzes trends over some time, evaluates the efficiency of various email formats, and measures the impact of these efforts.

OVERVIEW

Enhance users experiences by providing additional features or giving promotional opportunities that will keep people actively engaged.

To speed up processes, pinpoint areas that might benefit from automated technology advances.



Organizations may improve their retention and build long-term relationships by studying user behavior, meeting their requirements, and offering an outstanding user experience.

Greater engagement and boosting email effectiveness necessitates the use of a variety of methods geared to your audience and product. Better outcomes need constant evaluation, experimenting, and a particular emphasis on giving benefits to consumers.

CONCLUSION

- Operation Analytics is an important strategy that entails analyzing the operations of an organization. It tries to find areas for development and improvement by thoroughly reviewing many factors like processes, support, promotional activities, and so on. Data analysts may extract significant conclusions from the data collected.

- The analysis of metric spikes is a key component of Operation Analytics. Being a Data Analyst, I need to understand and resolve queries about variations in key trends. Analysts may carry out deep evaluations and inquiries to find the root causes leading to these spikes by utilizing the power of SQL.

- In conclusion, combining the use of SQL and Operation Analytics enables organizations to generate data-driven decisions, enhance their operational workflows, and boost development. Organizations can come up with targeted initiatives to deal with areas of issue and meet their operational goals by utilizing the insights provided by metric spike analysis.


Thank
you