OPERATION ANALYTICS AND INVESTIGATING METRIC SPIKE

Shruti Jha

PROJECT INSIGHTS

Operation Analytics is the analysis done for the complete end to end operations of a company. With the help of this, the company then finds the areas on which it must improve upon. You work closely with the ops team, support team, marketing team, etc and help them derive insights out of the data they collect.

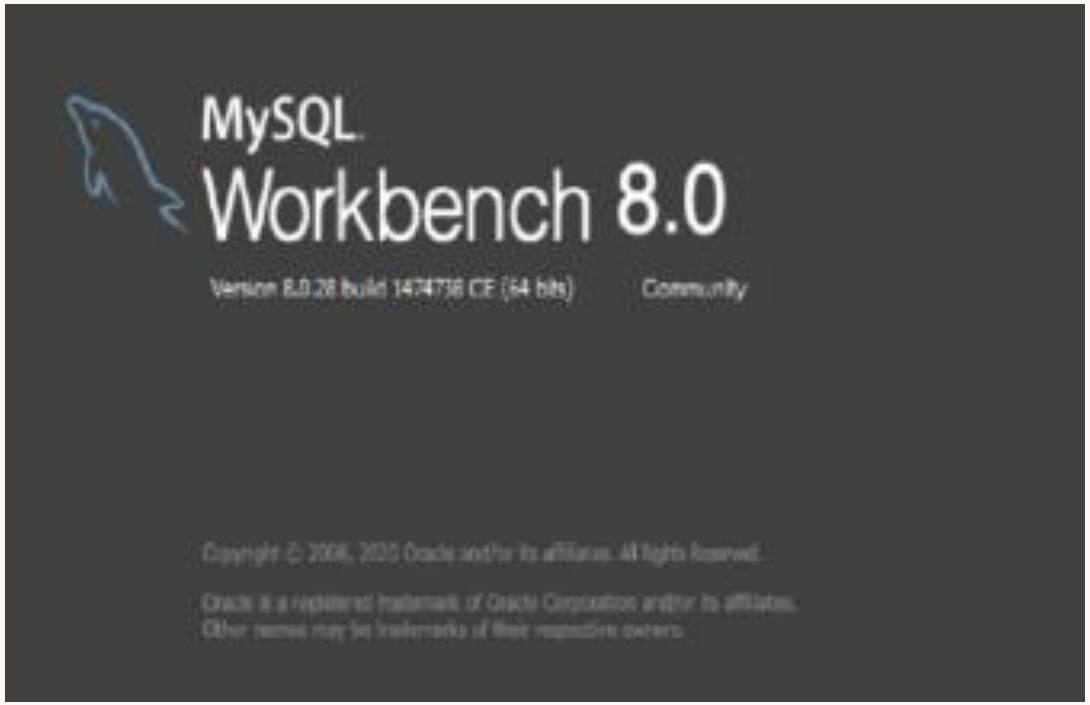
Being one of the most important parts of a company, this kind of analysis is further used to predict the overall growth or decline of a company's fortune. It means better automation, better understanding between cross-functional teams, and more effective workflows.

Investigating metric spike is also an important part of operation analytics as being a Data Analyst you must be able to understand or make other teams understand questions.

SOFTWARE USED

In order to perform the required analysis, I used SQL where I worked with MYSQL Workbench Version 8.0.3

I used the data provided from the organization about the employees, inclusive of their details and worked with them to find the necessary insights for better automation, better understanding between cross-functional teams, and more effective workflows. metrics.



CASE STUDY 1 (JOB DATA)

Using this piece of code in SQL, we can initialize our database for employees and find out the required details for analysis.

```
CREATE DATABASE operation_analysis;
USE operation_analysis;
CREATE TABLE job_data
(
    ds DATE,
    job_id INT NOT NULL,
    actor_id INT NOT NULL,
    event VARCHAR(15) NOT NULL,
    language VARCHAR(15) NOT NULL,
    time_spent INT NOT NULL,
    org CHAR(2)
);
INSERT INTO job_data (ds, job_id, actor_id, event, language, time_spent, org)
VALUES ('2020-11-30', 21, 1001, 'skip', 'English', 15, 'A'),
    ('2020-11-30', 22, 1006, 'transfer', 'Arabic', 25, 'B'),
    ('2020-11-29', 23, 1003, 'decision', 'Persian', 20, 'C'),
    ('2020-11-28', 23, 1005,'transfer', 'Persian', 22, 'D'),
    ('2020-11-28', 25, 1002, 'decision', 'Hindi', 11, 'B'),
    ('2020-11-27', 11, 1007, 'decision', 'French', 104, 'D'),
    ('2020-11-26', 23, 1004, 'skip', 'Persian', 56, 'A'),
    ('2020-11-25', 20, 1003, 'transfer', 'Italian', 45, 'C');
```

Operation-1 (Table-1)

ds	job_id	actor_id	event	language	time_spent	org
2020-11-30	21	1001	skip	English	15	A
2020-11-30	22	1006	transfer	Arabic	25	B
2020-11-29	23	1003	decision	Persian	20	C
2020-11-28	23	1005	transfer	Persian	22	D
2020-11-28	25	1002	decision	Hindi	11	B
2020-11-27	11	1007	decision	French	104	D
2020-11-26	23	1004	skip	Persian	56	A
2020-11-25	20	1003	transfer	Italian	45	C

CASE STUDY 1 (JOB DATA)

- **/* Number of jobs reviewed: Amount of jobs reviewed over time.**
- **Calculate the number of jobs reviewed per hour per day for November 2020 */**
- select ds as "Dates",ROUND((COUNT(job_id)/SUM(time_spent))*3600) as "Jobs reviewed per hour/day" from job_data where ds between '2020-11-01' and '2020-11-30' group by ds;

Dates	Jobs reviewed per hour/day
2020-11-28	218
2020-11-30	180
2020-11-29	180
2020-11-25	80
2020-11-26	64
2020-11-27	35

CASE STUDY 1 (JOB DATA)

- **/* Number of events happening per second.**
- **Calculate 7 day rolling average of throughput */**

•select ds as dates, round (count(event)/sum(time_spent),2) as "Daily throughput" from job_data group by ds order by ds;

select round(count(event)/sum(time_spent), 2) as "7 day rolling average throughput" from job_data;

dates	Daily 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

7 day rolling average throughput
0.03

CASE STUDY 1 (JOB DATA)

- **/* Percentage share of each language: Share of each language for different contents.**
- **Calculate the percentage share of each language in the last 30 days */**
- select language as Language, count(*)
select language as Languages, round(100 * count(*)/total, 2) as Percentage from job_data
- cross join (select count(*) as total from job_data) sub group by language;

Languages	Percentage
English	12.50
Arabic	12.50
Persian	37.50
Hindi	12.50
French	12.50
Italian	12.50

CASE STUDY 1 (JOB DATA)

- **/* Duplicate rows: Rows that have the same value present in them**
- **Display duplicates from the table */**
- select *, count(*) as duplicates from job_data group by actor_id having count(*) > 1;

ds	job_id	actor_id	event	language	time_spent	org	duplicates
2020-11-29	23	1003	decision	Persian	20	C	2

Operation-2 (Table-1 (Users))

user_id	A unique ID per user. Can be joined to user_id in either of the other tables.
created_at	The time the user was created (first signed up)
state	The state of the user (active or pending)
activated_at	The time the user was activated, if they are active
company_id	The ID of the user's company
language	The chosen language of the user

Operation-2 (Table-2(events))

user_id	The ID of the user logging the event. Can be joined to user_id in either of the other tables.
occurred_at	The time the event occurred.
event_type	The general event type. There are two values in this dataset: "signup_flow", which refers to anything occuring during the process of a user's authentication, and "engagement", which refers to general product usage after the user has signed up for the first time
event_name	The specific action the user took. Possible values include: create_user: User is added to Yammer's database during signup process enter_email: User begins the signup process by entering her email address enter_info: User enters her name and personal information during signup process complete_signup: User completes the entire signup/ authentication process home_page: User loads the home page like_message: User likes another user's message login: User logs into Yammer search_autocomplete: User selects a search result from the autocomplete list search_run: User runs a search query and is taken to the search results page search_click_result_X: User clicks search result X on the results page, where X is a number from 1 through 10. send_message: User posts a message view_inbox: User views messages in her inbox
location:	The country from which the event was logged (collected through IP address).
device:	The type of device used to log the event.

Operation-2 (Table-3(email_events))

user_id	The ID of the user to whom the event relates. Can be joined to user_id in either of the other tables.
occurred_at	The time the event occurred.
action	The name of the event that occurred. "sent_weekly_digest" means that the user was delivered a digest email showing relevant conversations from the previous day. "email_open" means that the user opened the email. "email_clickthrough" means that the user clicked a link in the email.

CASE STUDY 2 (INVESTIGATING METRIC SPIKE)

/*__User Engagement:__ To measure the activeness of a user. Measuring if the user finds quality in a product/service.
Calculate the weekly user engagement?*/

```
select extract(week from occurred_at) as
Week_number, count(distinct user_id) as
Distinct_users from events where
event_type="engagement" group by 1;
```

Week number	Distinct users
17	663
18	1068
19	1113
20	1154
21	1121
22	1186
23	1232
24	1275
25	1264
26	1302
27	1372
28	1365
29	1376
30	1467
31	1299
32	1225
33	1225
34	1204
35	104

CASE STUDY 2 (INVESTIGATING METRIC SPIKE)

User Growth: Amount of users growing over time for a product.

Calculate the user growth for product?

```
select months,users,round(((users/lag(users,1)over(order by months)-1)*100), 2) as "Growth in %" from( select extract(month from created_at) as months, count(activated_at) as users from users where activated_at not in("")group by 1 order by 1 );
```

Months	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
7	1281	17.96
8	1347	5.15
9	330	-75.50
10	390	18.18
11	399	2.31
12	486	21.80

CASE STUDY 2 (INVESTIGATING METRIC SPIKE)

Weekly Retention: Users getting retained weekly after signing-up for a product.
Calculate the weekly retention of users-sign up cohort

select months,users,round(((users/lag(users,1)over(order by months)-1)*100), 2) as "Growth in %" from(select extract(month from created_at) as months, count(activated_at) as users from users where activated_at not in("")group by 1 order by 1);

Week Numbers	Week 0	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11	Week 12	Week 13	Week 14	Week 15	Week 16	Week 17	Week 18
17	740	472	324	251	205	187	167	146	145	145	136	131	132	143	116	91	82	77	5
18	788	362	261	203	168	147	144	127	113	122	106	118	127	110	97	85	67	4	0
19	601	284	173	153	114	95	91	81	95	82	68	65	63	42	51	49	2	0	0
20	555	223	165	121	91	72	63	67	63	65	67	41	40	33	40	0	0	0	0
21	495	187	131	91	74	63	75	72	58	48	45	39	35	28	2	0	0	0	0
22	521	224	150	107	87	73	63	60	55	48	41	39	31	1	0	0	0	0	0
23	542	219	138	101	90	79	69	61	54	47	35	30	0	0	0	0	0	0	0
24	535	205	143	102	81	63	65	61	38	39	29	0	0	0	0	0	0	0	0
25	500	218	139	101	75	63	50	46	38	35	2	0	0	0	0	0	0	0	0
26	495	181	114	83	73	55	47	43	29	0	0	0	0	0	0	0	0	0	0
27	493	199	121	106	68	53	40	36	1	0	0	0	0	0	0	0	0	0	0
28	486	194	114	69	46	30	28	3	0	0	0	0	0	0	0	0	0	0	0
29	501	186	102	65	47	40	1	0	0	0	0	0	0	0	0	0	0	0	0
30	533	202	121	78	53	3	0	0	0	0	0	0	0	0	0	0	0	0	0
31	430	145	76	57	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32	496	188	94	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33	499	202	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	518	44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35	32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

CASE STUDY 2 (INVESTIGATING METRIC SPIKE)

Weekly Engagement: To measure the activeness of a user. Measuring if the user finds quality in a product/service weekly.
Calculate the weekly engagement per device

```
select extract(week from occured_at) as "Week numbers",device
from events where event_type = 'engagement'
group by 1 order by 1;
```

Week Numbers	Dell Inspiron Notebook	iPhone 5	iPhone 4S	Windows Surface	Macbook Air	iPhone 5S	Macbook Pro	Kindle Fire	iPad Mini	Nexus 7	Nexus 5	Samsung Galaxy S4	Lenovo Thinkpad	Samsung Galaxy Tablet
17	46	65	21	10	54	42	143	6	19	18	40	52	86	8
18	77	113	46	10	121	73	252	27	30	30	73	82	153	11
19	83	115	44	16	112	79	266	21	36	41	87	91	178	6
20	84	125	55	21	119	79	256	23	32	32	103	93	173	9
21	80	137	45	17	110	74	247	30	23	29	91	84	167	6
22	92	125	45	15	145	71	251	21	34	45	96	105	176	10
23	103	152	53	14	124	79	266	25	33	36	88	99	176	14
24	99	142	53	22	152	79	255	25	39	49	87	101	165	11
25	105	137	40	22	121	78	275	24	30	51	89	99	197	12
26	89	152	50	21	134	94	269	26	43	46	87	112	192	12
27	89	163	67	33	142	83	302	25	35	40	84	116	202	15
28	103	151	61	33	148	93	295	31	35	39	85	122	220	9
29	113	144	60	28	148	90	295	37	34	45	77	123	209	13
30	127	152	65	19	159	103	322	25	35	62	84	103	206	9
31	113	135	56	19	147	71	321	14	27	38	69	100	207	8
32	104	119	34	10	125	67	307	12	30	25	67	82	179	6
33	110	110	35	15	133	65	312	14	28	30	70	80	191	12
34	105	101	50	18	136	70	292	13	25	33	70	90	193	14
35	9	2	6	3	10	3	17	3	2	2	4	6	16	0

CASE STUDY 2 (INVESTIGATING METRIC SPIKE)

Email Engagement: Users engaging with the email service. Calculate the email engagement metrics

```
select ((weekly_digest/total*100),2) AS "Weekly Digest Rate",((email_opens/total*100),2) AS "Email Open
Rate",((email_clickthroughs/total*100),2) AS "Email Clickthrough Rate",((reengagement_emails/total*100),2) as
"Reengagement Email Rate
from (select extract (Week from  occurred_at) AS Week,
count(case when action = 'sent_weekly_digest' then  user_id else null end) as weekly_digest,
count(case when action = 'email_open' then  user_id else null end) as email_opens,
count(case when action = 'email_clickthrough' then  user_id else null end) as email_clickthroughs,
count(case when action = 'sent_reengagement_email' then user_id else null end) as reengagement_emails
count(user_id) as total
from mail_events group by 1)
group by 1 order by 1;
```

SUMMARY

By analyzing the data, we can derive various results that can be of use to the organization to improve the existing product for users based on their activities via databases and queries

This project helps me to understand the importance of operation analytics. Through this project I am able to understand how the companies use metric spike as a secret weapon. With an informed and proactive approach, they can leverage insights to make data-backed decisions that optimize their strategy.

THANK YOU

Operational Analytics has the capability to aggregate data from multiple data sources into a cumulative, organized, actionable solution capable of delivering analytical models in real-time to create individual customer profiles and a holistic view of operations for a company.

Prepared by: Shruti Jha