



Operation Analytics AND Investigating Metric spike





Description

Operational Analytics is a crucial process that involves analyzing a company's end-to-end operations. This analysis helps identify areas for improvement within the company. The project focuses on conducting operational analytics for a company using data analysis techniques. I have been provided with various data sets and tables to derive insights and answer questions related to different operations. The goal is to analyze the data, identify areas of improvement, and predict the overall growth or decline of the company's performance.

APPROACH

- Created a database and tables to store the provided data sets.
- Utilized SQL queries to perform analysis on the tables and derive the required insights.

Answered specific questions related to the operations, such as the number of jobs reviewed, throughput, percentage share of each language, duplicate rows, user engagement, user growth, weekly retention, weekly engagement per device, and email engagement.

- Documented the findings and insights in a report format.



TECH-STACK USED





INSIGHTS

CASE STUDY 1

JOB DATA ANALYSIS

CASE STUDY 1 - Task A

Jobs Reviewed Over Time

```
SELECT ds AS dates,  
round((COUNT(job_id)/SUM(time_spent))*3600) AS  
'Jobs Reviewed per Hour per day'  
from job_data  
where ds between '2020-11-01' AND '2020-11-30'  
group by ds;
```

Objective: Calculate the number of jobs reviewed per hour for each day in November 2020.

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

CASE STUDY 1 - Task B

Throughput Analysis

Objective: Calculate the 7-day rolling average of throughput (number of events per second).

Using both metrics together can provide a comprehensive understanding of throughput performance, with the daily metric offering insights into short-term fluctuations and the rolling average revealing long-term trends.

```
SELECT ds AS date,  
AVG(COUNT(*)) OVER  
(ORDER BY ds ROWS  
BETWEEN 6  
PRECEDING AND  
CURRENT ROW) AS  
7_day_rolling_average  
FROM ritika.job_data  
GROUP BY ds  
ORDER BY ds;
```

```
SELECT ds,  
COUNT(DISTINCT event)  
AS event_per_day  
FROM ritika.job_data  
GROUP by ds;
```

	date	7_day_rolling_average
▶	25-11-2020	1.0000
	26-11-2020	1.0000
	27-11-2020	1.0000
	28-11-2020	1.2500
	29-11-2020	1.2000
	30-11-2020	1.3333

ds	event_per_day
25-11-2020	1
26-11-2020	1
27-11-2020	1
28-11-2020	2
29-11-2020	1
30-11-2020	2

CASE STUDY 1 - Task C

Language Share Analysis

```
SELECT language, COUNT(*) AS content_count,  
COUNT(*)*100.0/SUM(COUNT(*) OVER()) AS  
percentage_share FROM job_data GROUP BY language;
```

Objective: Calculate the percentage share of each language in the last 30 days

language	content_count	percentage_share
English	1	12.50000
Arabic	1	12.50000
Persian	3	37.50000
Hindi	1	12.50000
French	1	12.50000
Italian	1	12.50000

CASE STUDY 1 - Task D

Duplicate Rows Detection

```
SELECT *  
FROM ( SELECT *, ROW_NUMBER()OVER(PARTITION  
BY job_id) AS rownum  
FROM job_data )a WHERE rownum > 1 ;
```

Objective: Identify duplicate rows in the data.

ds	job_id	actor_id	event	language	time_spent	org	rownum
28-11-2020	23	1005	transfer	Persian	22	D	2
26-11-2020	23	1004	skip	Persian	56	A	3



INSIGHTS

CASE STUDY 2

INVESTIGATING METRIC SPIKE

CASE STUDY 2 • Task A

Weekly User Engagement

Objective: Measure the activeness of users on a weekly basis.

Week Number	Weekly Active Users
17	85
18	194
19	208
20	195
21	208
22	230
23	224
24	252
25	245
26	123
27	14
28	9
29	5
30	1
31	2
32	1

```
SELECT
extract(WEEK FROM occurred_at) AS "Week Number",
COUNT(DISTINCT user_id) AS "Weekly Active Users"
FROM events
GROUP BY 1;
```

CASE STUDY 2 • Task B

User Growth Analysis

months	users	Growth(%)
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

```
SELECT months, users,
round(((users/LAG(users,1)
over (ORDER BY months) -
1)*100),2)AS "Growth(%)"
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
)sub ;
```

Objective: Analyze the growth of users over time for a product.

CASE STUDY 2 - Task C

Weekly Retention Analysis

Objective: Analyze the retention of users on a weekly basis after signing up for a product.

```
SELECT COUNT(user_id)as users,
SUM(CASE WHEN retention_week = 1 THEN 1 ELSE 0 END ) AS week_1,
SUM(CASE WHEN retention_week = 2 THEN 1 ELSE 0 END ) AS week_2,
SUM(CASE WHEN retention_week = 3 THEN 1 ELSE 0 END ) AS week_3,
SUM(CASE WHEN retention_week = 4 THEN 1 ELSE 0 END ) AS week_4,
SUM(CASE WHEN retention_week = 5 THEN 1 ELSE 0 END ) AS week_5
FROM (
SELECT a.user_id,
a.sign_up_week,
b.engagement_week,
b.engagement_week - a.sign_up_week as retention_week
FROM ((select distinct user_id, extract(week from occurred_at) as
sign_up_week
from events
where event_type='signup_flow'and event_name='complete_signup' and
extract(week from occurred_at)=18)a
left JOIN
(select distinct user_id,extract(week from occurred_at) as
engagement_weekfrom events
where event_type='engagement')b
on a.user_id=b.user_id )
order by a.user_id )a
```

users	week_1	week_2	week_3	week_4	week_5
182	11	4	3	0	1

CASE STUDY 2 • Task D

Weekly Engagement Per Device

```
Select  
extract(year from occurred_at)as year,  
extract(week from occurred_at)as week,  
device,  
COUNT(distinct user_id) as num_users  
FROM events  
WHERE event_type='engagement'  
group by 1,2,3  
order BY1,2,3
```

Objective: Measure the activeness of users on a weekly basis per device.

CASE STUDY 2 - Task D Weekly Engagement Per Device

Objective: Measure the activeness of users on a weekly basis per device.(A part of result)

year	week	device	num_users
2014	17	samsung galaxy tablet	2
2014	17	samsung galaxy note	1
2014	17	samsung galaxy s4	7
2014	18	acer aspire desktop	4
2014	18	acer aspire notebook	4
2014	18	amazon fire phone	2
2014	18	asus chromebook	4
2014	18	dell inspiron desktop	3
2014	18	dell inspiron notebook	12
2014	18	hp pavilion desktop	6
2014	18	htc one	2
2014	18	ipad air	8
2014	18	ipad mini	7
2014	18	iphone 4s	4
2014	18	iphone 5	7
2014	18	iphone 5s	8
2014	18	kindle fire	5
2014	18	lenovo thinkpad	28
2014	18	mac mini	1
2014	18	macbook air	18

year	week	device	num_users
2014	17	acer aspire desktop	2
2014	17	acer aspire notebook	2
2014	17	amazon fire phone	1
2014	17	asus chromebook	3
2014	17	dell inspiron desktop	1
2014	17	dell inspiron notebook	4
2014	17	hp pavilion desktop	2
2014	17	htc one	2
2014	17	ipad air	1
2014	17	ipad mini	3
2014	17	iphone 4s	3
2014	17	iphone 5	11
2014	17	iphone 5s	5
2014	17	lenovo thinkpad	8
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

year	week	device	num_users
2014	18	macbook pro	41
2014	18	nexus 10	2
2014	18	nexus 5	9
2014	18	nexus 7	5
2014	18	nokia lumia 635	4
2014	18	samsung galaxy note	4
2014	18	samsung galaxy s4	12
2014	18	windows surface	1
2014	19	acer aspire notebook	8
2014	19	amazon fire phone	4
2014	19	asus chromebook	3
2014	19	dell inspiron desktop	1
2014	19	dell inspiron notebook	7
2014	19	hp pavilion desktop	5
2014	19	htc one	6
2014	19	ipad air	10
2014	19	ipad mini	1
2014	19	iphone 4s	7
2014	19	iphone 5	19
2014	19	iphone 5s	9

year	week	device	num_users
2014	19	kindle fire	4
2014	19	lenovo thinkpad	21
2014	19	mac mini	2
2014	19	macbook air	23
2014	19	macbook pro	37
2014	19	nexus 10	3
2014	19	nexus 5	15
2014	19	nexus 7	7
2014	19	nokia lumia 635	2
2014	19	samsung galaxy note	4
2014	19	samsung galaxy s4	14
2014	19	windows surface	3
2014	20	acer aspire desktop	2
2014	20	acer aspire notebook	4
2014	20	amazon fire phone	1
2014	20	asus chromebook	6
2014	20	dell inspiron desktop	11
2014	20	dell inspiron notebook	9
2014	20	hp pavilion desktop	2
2014	20	htc one	3

CASE STUDY 2 - Task E

Email Engagement Analysis

Objective: Analyze how users are engaging with the email service.

```
SELECT  action AS email_engagement_type,  
COUNT(DISTINCT user_id) AS engaged_users  
FROM    ritika.email_events  
GROUP BY  action ;
```

email_engagement_type	engaged_users
email_clickthrough	5277
email_open	5927
sent_reengagement_email	3653
sent_weekly_digest	4111

THANK YOU!

The project was instrumental in improving my SQL proficiency and provided practical insights into operational aspects, user engagement, and the product's growth trajectory. I successfully presented these findings to the leadership team, showcasing my ability to deliver valuable insights that can positively impact business strategies. Overall, this experience has boosted my confidence and independence in handling data analysis tasks in real-world scenarios.