PROJECT – III

Operation Analytics and Investigating Metric Spike

PROJECT DESCRIPTION

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 like- Why is there a **dip** in daily engagement? Why have sales taken a dip? Etc. Questions like these must be answered daily and for that it's very important to investigate metric spike.

You (I) are (am) required to provide a detailed report for the two operations by mentioning the answers for the related questions that are asked in Case Study-1 and Case Study-2.

The Project first requires us to create a **database** and import it any SQL platform. And then it requires us to give answers to the questions that they have asked.

I would be using MYSQL in the further report.

I know this project will help me clear my doubts regarding **SQL programming** and **databases**. Having strong concepts, initially, about any topic is very important for further strengthening of the concepts and through this project I will learn to create a database and then import in MYSQL platform and then further learn to code in the form of making **queries.**

APPROACH

I will be following the steps as directed for completing this project . Learning at every step will be my goal. I have watched all the videos pertaining to this project . I have understood the concepts at every step, and I will be applying that knowledge here to finish this project with great respect for the mentors. I have practiced few questions at my level to test my knowledge before jumping into the project.

Hopefully with all the knowledge that I have gained from the videos and my practice would help me in solving the questions and finish the project.

TECH – STACK USED

FOLLOWING ARE THE TECH STACKS USED BY ME FOR COMPLETEING THS PROJECT

- MySQL Workbench 8.0 CE
- Excel

INSIGHTS: CASE STUDY – 1 (Job Data)

Below is the structure of the table with the definition of each column that we have to work on:

Table-1: job_data

job_id: unique identifier of jobs

actor_id: unique identifier of actor

event: decision/skip/transfer

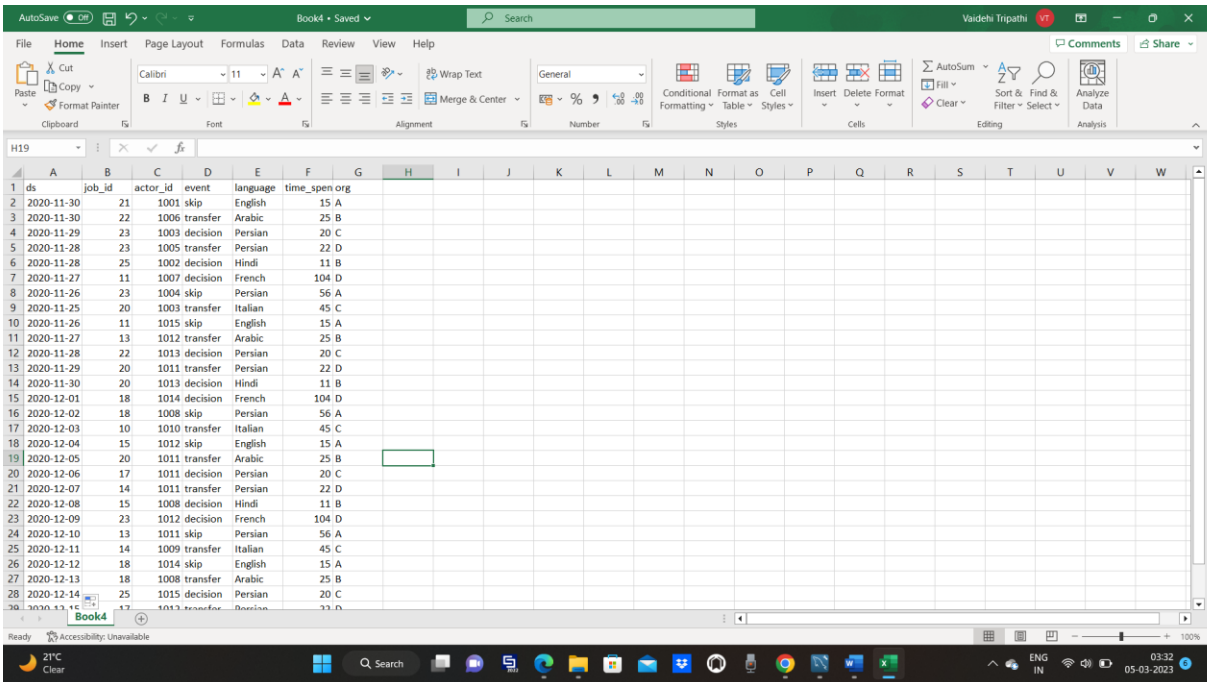
language: language of the content

time_spent: time spent to review the job in seconds

org: organization of the actor

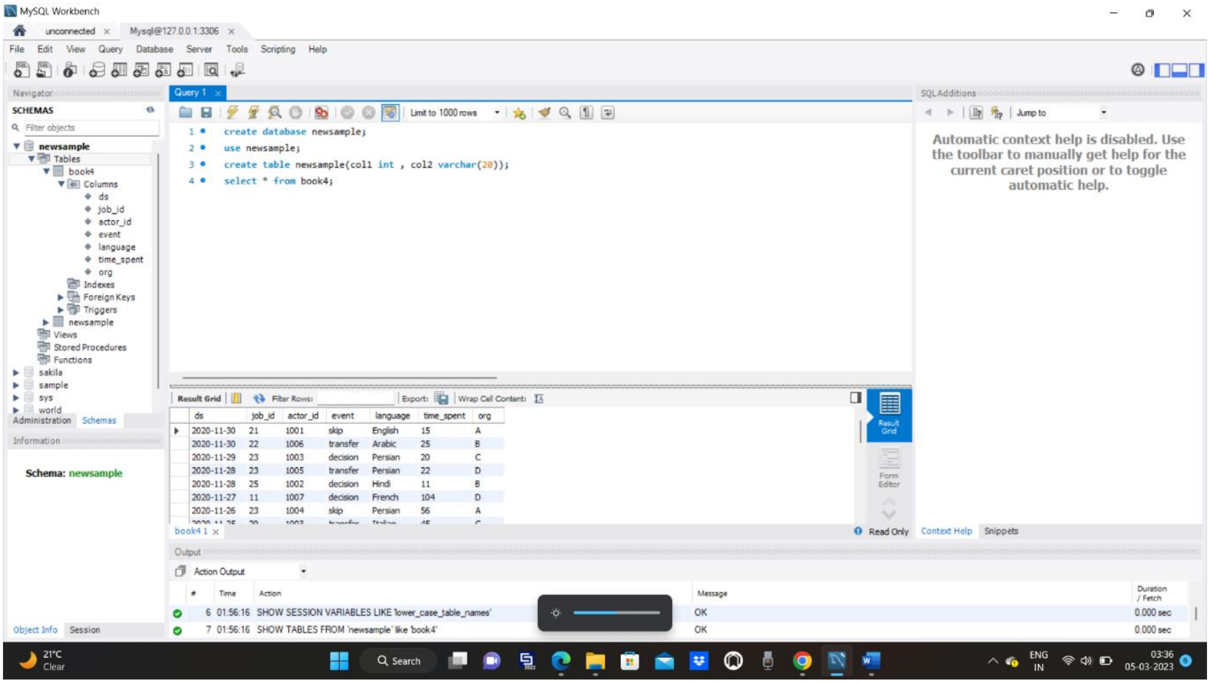
ds: date in the yyyy/mm/dd format. It is stored in the form of text, and we use presto to run. no need for date function

NOW FIRST I HAVE FORMED THE DATASET FOR 30 DAYS USING EXCEL



	A	B	C	D	E	F	G
1	ds	job_id	actor_id	event	language	time_spen	org
2	2020-11-30	21	1001	skip	English	15	A
3	2020-11-30	22	1006	transfer	Arabic	25	B
4	2020-11-29	23	1003	decision	Persian	20	C
5	2020-11-28	23	1005	transfer	Persian	22	D
6	2020-11-28	25	1002	decision	Hindi	11	B
7	2020-11-27	11	1007	decision	French	104	D
8	2020-11-26	23	1004	skip	Persian	56	A
9	2020-11-25	20	1003	transfer	Italian	45	C
10	2020-11-26	11	1015	skip	English	15	A
11	2020-11-27	13	1012	transfer	Arabic	25	B
12	2020-11-28	22	1013	decision	Persian	20	C
13	2020-11-29	20	1011	transfer	Persian	22	D
14	2020-11-30	20	1013	decision	Hindi	11	B
15	2020-12-01	18	1014	decision	French	104	D
16	2020-12-02	18	1008	skip	Persian	56	A
17	2020-12-03	10	1010	transfer	Italian	45	C
18	2020-12-04	15	1012	skip	English	15	A
19	2020-12-05	20	1011	transfer	Arabic	25	B
20	2020-12-06	17	1011	decision	Persian	20	C
21	2020-12-07	14	1011	transfer	Persian	22	D
22	2020-12-08	15	1008	decision	Hindi	11	B
23	2020-12-09	23	1012	decision	French	104	D
24	2020-12-10	13	1011	skip	Persian	56	A
25	2020-12-11	14	1009	transfer	Italian	45	C
26	2020-12-12	18	1014	skip	English	15	A
27	2020-12-13	18	1008	transfer	Arabic	25	B
28	2020-12-14	25	1015	decision	Persian	20	C
29	2020-12-15	17	1012	transfer	Persian	22	D

THEN I HAVE IMPORTED THIS .CSV FILE INTO MYSQL PLATFORM



```
create database newsample;
use newsample;
create table newsample(col1 int , col2 varchar(20));
select * from book4;
```

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Schema: newsample

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

```
#  Time      Action                                            Message  Duration / Fetch
6  01:56:16  SHOW SESSION VARIABLES LIKE 'lower_case_table_names'  OK       0.000 sec
7  01:56:16  SHOW TABLES FROM 'newsample' like 'book4'             OK       0.000 sec
```

NOW THE TASKS !

(A) Number of jobs reviewed: Amount of jobs reviewed over time.

Your task: Calculate   jobs reviewed per hour per day for November 2020?

ANSWER:

select ds,

extract(month from ds) as monthnum,

extract(day from ds) as daydate,

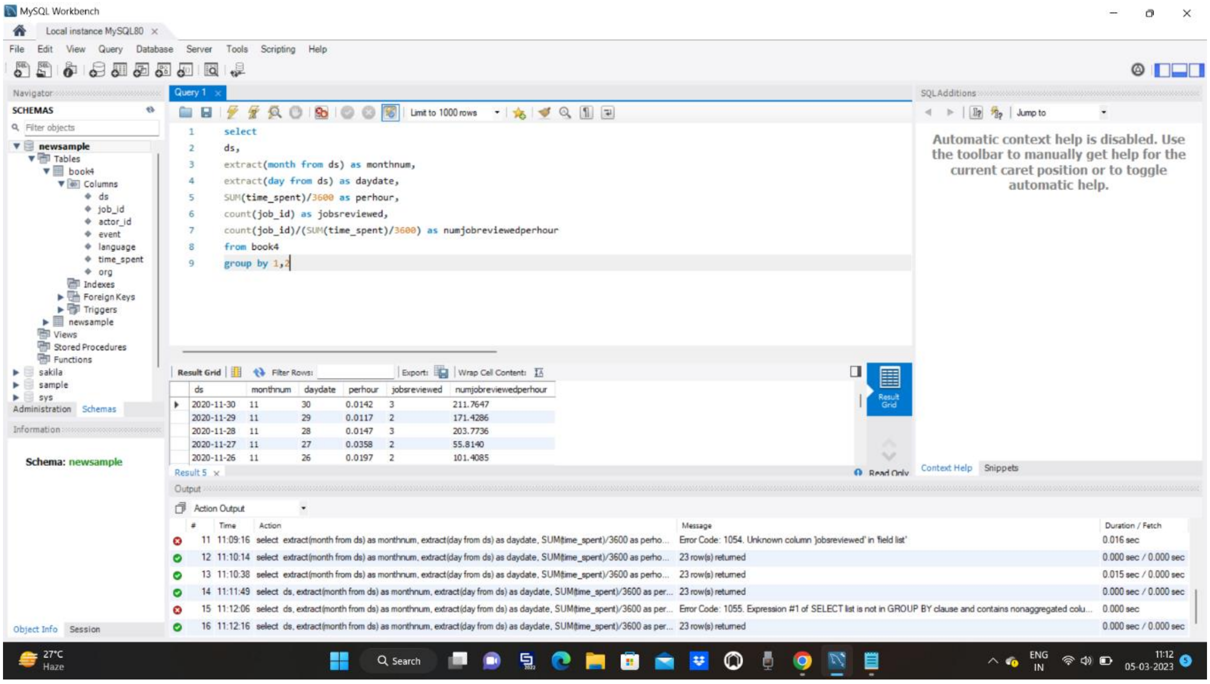
SUM(time_spent)/3600 as perhour,

count(job_id) as jobsreviewed,

count(job_id)/(SUM(time_spent)/3600) as numjobreviewedperhour

from book4

group by 1,2



```
select
ds,
extract(month from ds) as monthnum,
extract(day from ds) as daydate,
SUM(time_spent)/3600 as perhour,
count(job_id) as jobsreviewed,
count(job_id)/(SUM(time_spent)/3600) as numjobreviewedperhour
from book4
group by 1,2
```

ds	monthnum	daydate	perhour	jobsreviewed	numjobreviewedperhour
2020-11-30	11	30	0.0142	3	211.7647
2020-11-29	11	29	0.0117	2	171.4286
2020-11-28	11	28	0.0147	3	203.7736
2020-11-27	11	27	0.0358	2	55.8140
2020-11-26	11	26	0.0197	2	101.4085

#	Time	Action	Message	Duration / Fetch
11	11:09:16	select extract(month from ds) as monthnum, extract(day from ds) as daydate, SUM(time_spent)/3600 as perho...	Error Code: 1054. Unknown column 'jobsreviewed' in 'field list'	0.016 sec
12	11:10:14	select extract(month from ds) as monthnum, extract(day from ds) as daydate, SUM(time_spent)/3600 as perho...	23 row(s) returned	0.000 sec / 0.000 sec
13	11:10:38	select extract(month from ds) as monthnum, extract(day from ds) as daydate, SUM(time_spent)/3600 as perho...	23 row(s) returned	0.015 sec / 0.000 sec
14	11:11:49	select ds, extract(month from ds) as monthnum, extract(day from ds) as daydate, SUM(time_spent)/3600 as per...	23 row(s) returned	0.000 sec / 0.000 sec
15	11:12:06	select ds, extract(month from ds) as monthnum, extract(day from ds) as daydate, SUM(time_spent)/3600 as per...	Error Code: 1055. Expression #1 of SELECT list is not in GROUP BY clause and contains nonaggregated colu...	0.000 sec
16	11:12:16	select ds, extract(month from ds) as monthnum, extract(day from ds) as daydate, SUM(time_spent)/3600 as per...	23 row(s) returned	0.000 sec / 0.000 sec

INSIGHT: By using group by and different functions like sum and count, I was able to find out the answer to the question asked. Firstly I extracted day and month from the given dates and then I used sum, count, group by functions.

(B)Throughput: It is the no. of events happening per second. Your task: Let's say the above metric is called throughput. Calculate 7 day rolling average of throughput? For throughput, do you prefer daily metric or 7-day rolling and why?

ANSWER:

select a.*,

avg(throughput_per_day) over(order by ds desc rows between current row and 7 following) as seven_days_rolling_avg_throughput

from

(select

ds,

COUNT(event) AS num_events_per_day,

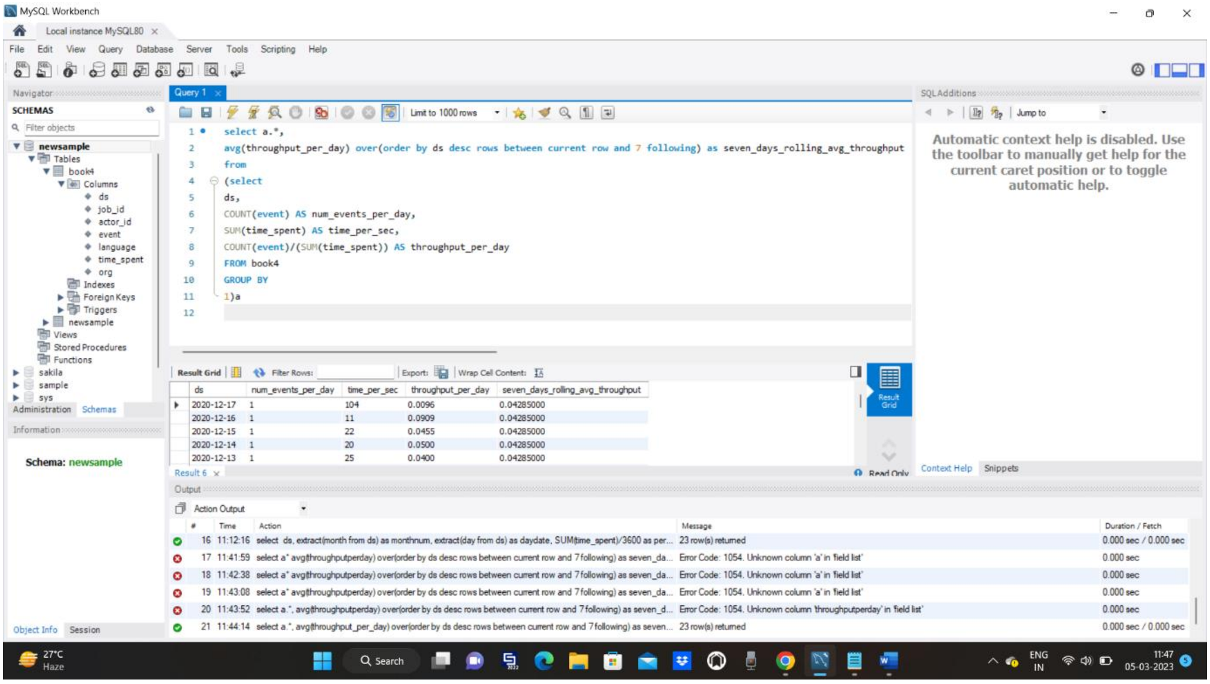
SUM(time_spent) AS time_per_sec,

COUNT(event)/(SUM(time_spent)) AS throughput_per_day

FROM book4

GROUP BY

1)a



```
select a.*,
avg(throughput_per_day) over(order by ds desc rows between current row and 7 following) as seven_days_rolling_avg_throughput
from
(select
ds,
COUNT(event) AS num_events_per_day,
SUM(time_spent) AS time_per_sec,
COUNT(event)/(SUM(time_spent)) AS throughput_per_day
FROM book4
GROUP BY
1)a
```

ds	num_events_per_day	time_per_sec	throughput_per_day	seven_days_rolling_avg_throughput
2020-12-17	1	104	0.0096	0.04285000
2020-12-16	1	11	0.0909	0.04285000
2020-12-15	1	22	0.0455	0.04285000
2020-12-14	1	20	0.0500	0.04285000
2020-12-13	1	25	0.0400	0.04285000

```
#   Time      Action                                                                                          Message                                                     Duration / Fetch
16  11:12:16  select  ds, extract(month from ds) as monthnum, extract(day from ds) as daydate, SUM(time_spent)/3600 as per...  23 row(s) returned                                          0.000 sec / 0.000 sec
17  11:41:59  select a* avg(throughputperday) over(order by ds desc rows between current row and 7following) as seven_da...  Error Code: 1054. Unknown column 'a' in 'field list'        0.000 sec
18  11:42:38  select a* avg(throughputperday) over(order by ds desc rows between current row and 7following) as seven_da...  Error Code: 1054. Unknown column 'a' in 'field list'        0.000 sec
19  11:43:08  select a* avg(throughputperday) over(order by ds desc rows between current row and 7following) as seven_da...  Error Code: 1054. Unknown column 'a' in 'field list'        0.000 sec
20  11:43:52  select a.*, avg(throughputperday) over(order by ds desc rows between current row and 7following) as seven_d...  Error Code: 1054. Unknown column 'throughputperday' in 'field list'  0.000 sec
21  11:44:14  select a.*, avg(throughput_per_day) over(order by ds desc rows between current row and 7following) as seven...  23 row(s) returned                                          0.000 sec / 0.000 sec
```

MySQL Workbench
Local instance MySQL80
File Edit View Query Database Server Tools Scripting Help
Query 1
Limit to 1000 rows
SCHEMAS
newsample
book4
Schema: newsample
Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.
Result 6
Action Output
27°C Haze
11:47 05-03-2023

INSIGHT: Here I have made use of the concept of subquery and in that subquery, I have calculated throughput per day by using sum, count, group by functions. Then in the outer query I calculated the average for 7 rolling days.

(C)Percentage share of each language: Share of each language for different contents.
Your task: Calculate the percentage share of each language in the last 30 days?

ANSWER:

SELECT

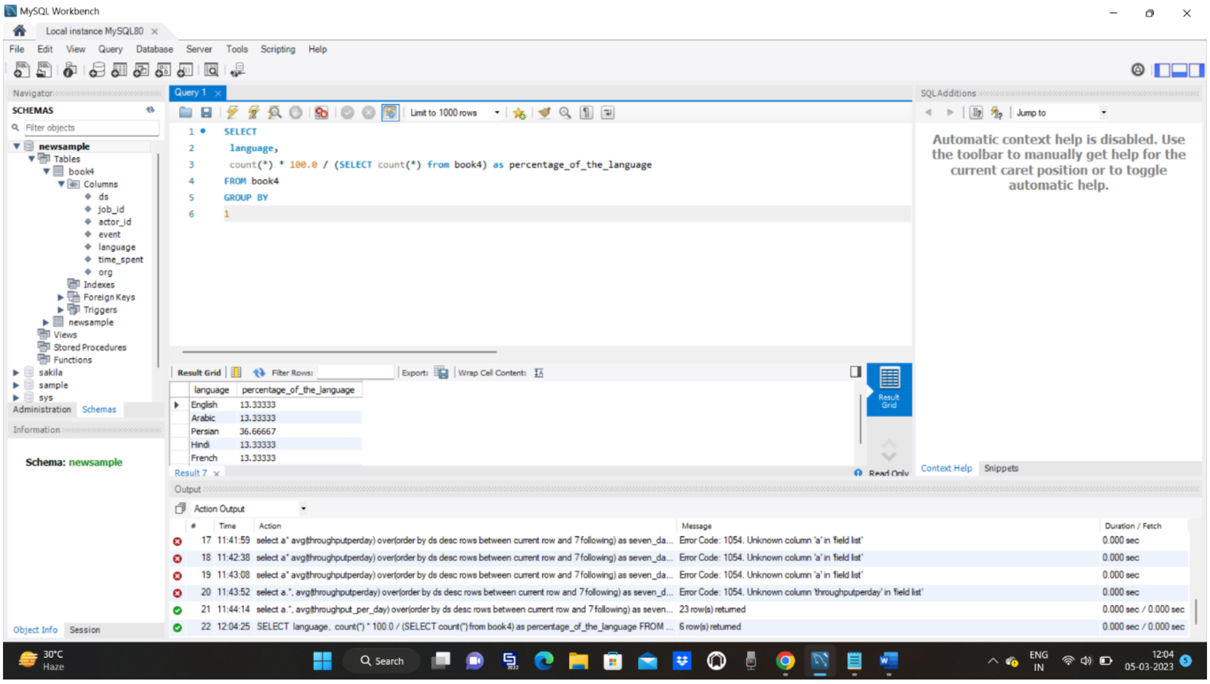
 language,

 count(*) * 100.0 / (SELECT count(*) from book4) as percentage_of_the_language

FROM book4

GROUP BY

1



```
SELECT
 language,
 count(*) * 100.0 / (SELECT count(*) from book4) as percentage_of_the_language
FROM book4
GROUP BY
1
```

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

language	percentage_of_the_language
English	13.33333
Arabic	13.33333
Persian	36.66667
Hindi	13.33333
French	13.33333

```
#   Time      Action                                                                                              Message                                                     Duration / Fetch
17  11:41:59  select a* avg(throughputperday) over(order by ds desc rows between current row and 7following) as seven_da...  Error Code: 1054. Unknown column 'a' in 'field list'        0.000 sec
18  11:42:38  select a* avg(throughputperday) over(order by ds desc rows between current row and 7following) as seven_da...  Error Code: 1054. Unknown column 'a' in 'field list'        0.000 sec
19  11:43:08  select a* avg(throughputperday) over(order by ds desc rows between current row and 7following) as seven_da...  Error Code: 1054. Unknown column 'a' in 'field list'        0.000 sec
20  11:43:52  select a.*, avg(throughputperday) over(order by ds desc rows between current row and 7following) as seven_d...  Error Code: 1054. Unknown column 'throughputperday' in 'field list'  0.000 sec
21  11:44:14  select a.*, avg(throughput_per_day) over(order by ds desc rows between current row and 7following) as seven...  23 row(s) returned                                          0.000 sec / 0.000 sec
22  12:04:25  SELECT language, count(*) * 100.0 / (SELECT count(*) from book4) as percentage_of_the_language FROM ...      6 row(s) returned                                           0.000 sec / 0.000 sec
```

INSIGHT: Here I have first grouped the languages using group by function, then I have count the total number of times the particular language is used divided by total number of languages.

(D)Duplicate rows: Rows that have the same value present in them. **Your task:** Let’s say you see some duplicate rows in the data. How will you display duplicates from the table?

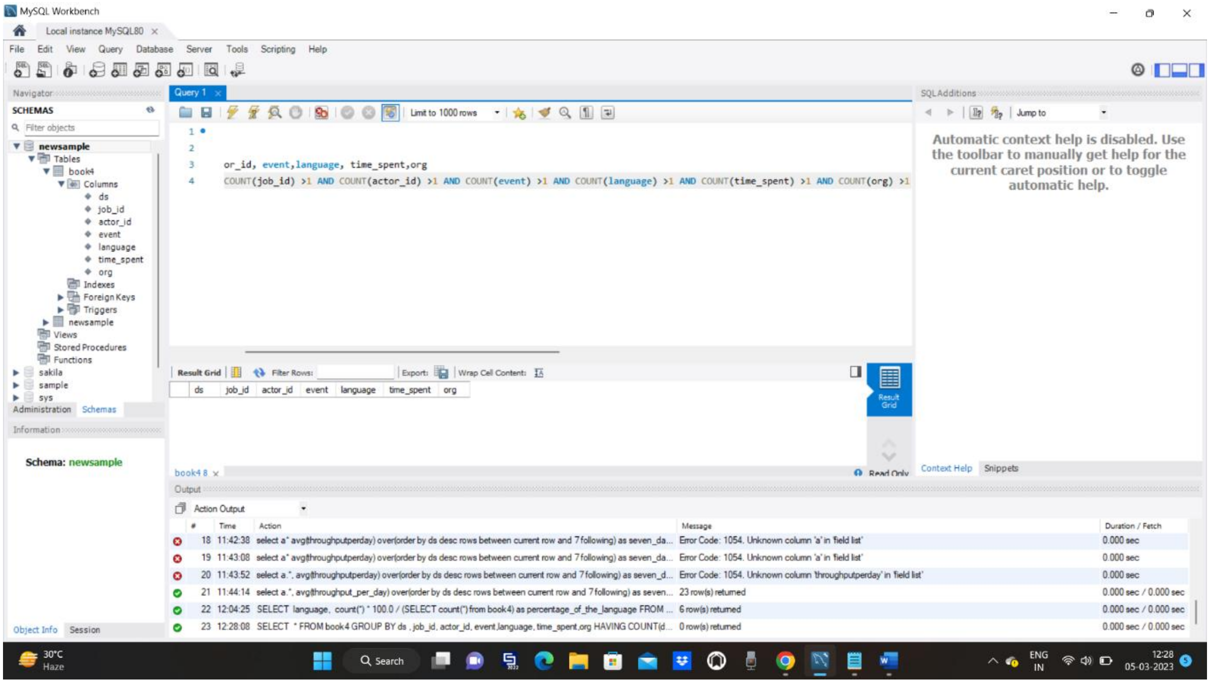
ANSWER:

SELECT *

FROM book4

GROUP BY ds , job_id, actor_id, event,language, time_spent,org

HAVING COUNT(ds) >1 AND COUNT(job_id) >1 AND COUNT(actor_id) >1 AND COUNT(event) >1 AND COUNT(language) >1 AND COUNT(time_spent) >1 AND COUNT(org) >1



```
or_id, event,language, time_spent,org
COUNT(job_id) >1 AND COUNT(actor_id) >1 AND COUNT(event) >1 AND COUNT(language) >1 AND COUNT(time_spent) >1 AND COUNT(org) >1
```

ds	job_id	actor_id	event	language	time_spent	org

```
#   Time      Action                                                                                              Message                                                      Duration / Fetch
18  11:42:38  select a* avg(throughputperday) over(order by ds desc rows between current row and 7following) as seven_da...  Error Code: 1054. Unknown column 'a' in 'field list'           0.000 sec
19  11:43:08  select a* avg(throughputperday) over(order by ds desc rows between current row and 7following) as seven_da...  Error Code: 1054. Unknown column 'a' in 'field list'           0.000 sec
20  11:43:52  select a.*, avg(throughputperday) over(order by ds desc rows between current row and 7following) as seven_d...  Error Code: 1054. Unknown column 'throughputperday' in 'field list'  0.000 sec
21  11:44:14  select a.*, avg(throughput_per_day) over(order by ds desc rows between current row and 7following) as seven...  23 row(s) returned                                            0.000 sec / 0.000 sec
22  12:04:25  SELECT language, count(*) * 100.0 / (SELECT count(*) from book4) as percentage_of_the_language FROM ...         6 row(s) returned                                             0.000 sec / 0.000 sec
23  12:28:08  SELECT * FROM book4 GROUP BY ds , job_id, actor_id, event,language, time_spent,org HAVING COUNT(d...           0 row(s) returned                                             0.000 sec / 0.000 sec
```

INSIGHT: Here I have used group by to group all the same dates , events, languages, time_spent,actor_id,job_id etc. Rows having all the data same are to be shown, since my dataset does not contains any such row therefore it is showing nothing inside the dataset , which is fine.

CASE STUDY – 2 (Investigating metric spike)

The structure of the table with the definition of each column that you must work on is present in the project image

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, 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.

Use the dataset then answer the questions that follows:

TABLES:

https://drive.google.com/drive/folders/1bB-uqONISA6wiI1hw1LzISpe0-kHg0Nx

NOW THE TASKS!

(A) User Engagement: To measure the activeness of a user. Measuring if the user finds quality in a product/service.
Your task: Calculate the weekly user engagement?

ANSWER:

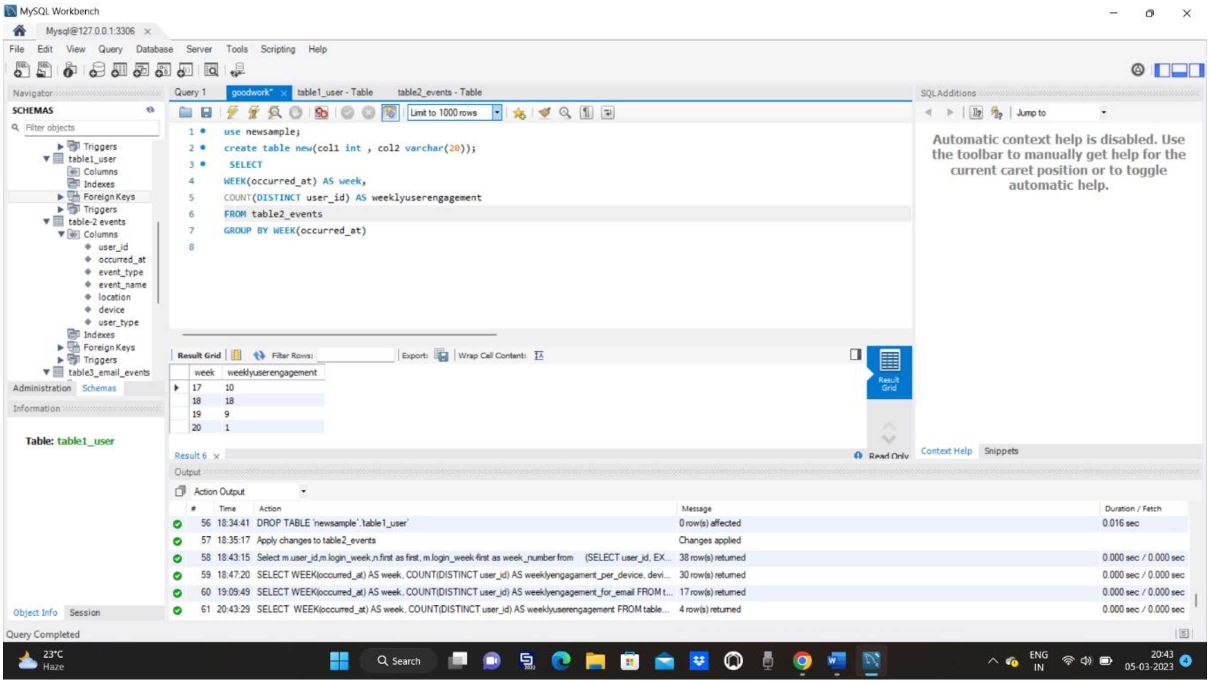
```
SELECT

WEEK(occurred_at) AS week,

COUNT(DISTINCT user_id) AS weeklyuserengagement

FROM table2_events

GROUP BY WEEK(occurred_at)
```



```
use newsample;
create table new(col1 int , col2 varchar(20));
 SELECT
WEEK(occurred_at) AS week,
COUNT(DISTINCT user_id) AS weeklyuserengagement
FROM table2_events
GROUP BY WEEK(occurred_at)
```

week	weeklyuserengagement
17	10
18	18
19	9
20	1

```
#   Time      Action                                                                                              Message             Duration / Fetch
56  18:34:41  DROP TABLE `newsample`.`table1_user`                                                                0 row(s) affected   0.016 sec
57  18:35:17  Apply changes to table2_events                                                                      Changes applied
58  18:43:15  Select m.user_id,m.login_week,n.first as first, m.login_week-first as week_number from    (SELECT user_id, EX...   38 row(s) returned   0.000 sec / 0.000 sec
59  18:47:20  SELECT WEEK(occurred_at) AS week, COUNT(DISTINCT user_id) AS weeklyengagament_per_device, devi...   30 row(s) returned   0.000 sec / 0.000 sec
60  19:09:49  SELECT WEEK(occurred_at) AS week, COUNT(DISTINCT user_id) AS weeklyengagement_for_email FROM t...   17 row(s) returned   0.000 sec / 0.000 sec
61  20:43:29  SELECT  WEEK(occurred_at) AS week, COUNT(DISTINCT user_id) AS weeklyuserengagement FROM table...     4 row(s) returned    0.000 sec / 0.000 sec
```

_(B)_User Growth: Amount of users growing over time for a product.

Your task: Calculate the user growth for product?

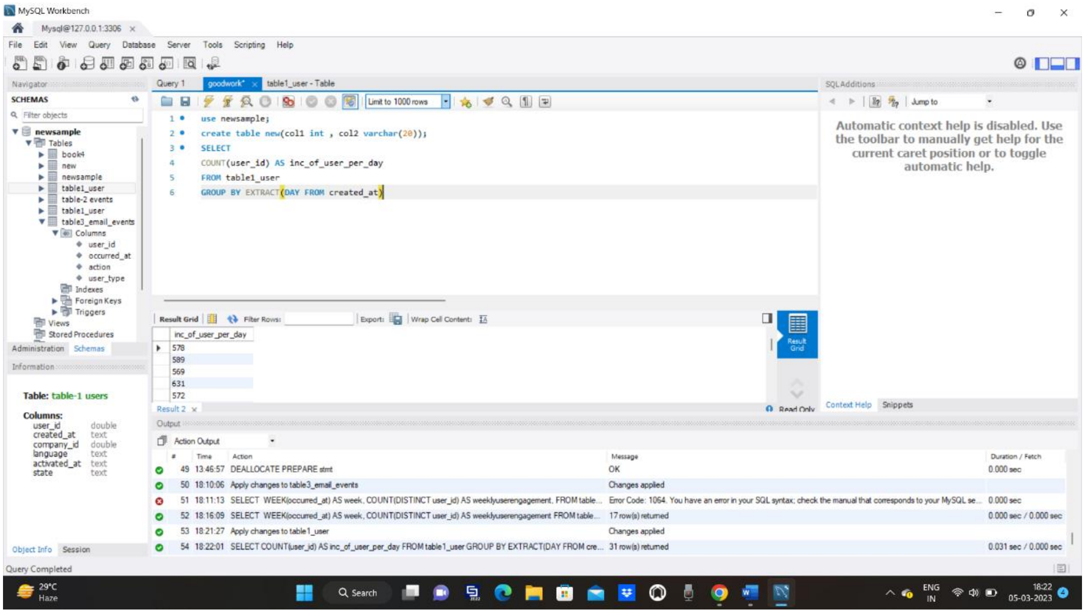
ANSWER:

SELECT

COUNT(user_id) AS inc_of_user_per_day

FROM table1_user

GROUP BY EXTRACT(DAY FROM created_at)



```
use newsample;
create table new(col1 int , col2 varchar(20));
SELECT
COUNT(user_id) AS inc_of_user_per_day
FROM table1_user
GROUP BY EXTRACT(DAY FROM created_at)
```

Table: table-1 users

Columns:

```
user_id       double
created_at    text
company_id    double
language      text
activated_at  text
state         text
```

inc_of_user_per_day
578
589
569
631
572

```
#   Time      Action                                                                                  Message                                                                                  Duration / Fetch
49  13:46:57  DEALLOCATE PREPARE stmt                                                                 OK                                                                                       0.000 sec
50  18:10:06  Apply changes to table3_email_events                                                    Changes applied
51  18:11:13  SELECT WEEK(occured_at) AS week, COUNT(DISTINCT user_id) AS weeklyuserengagement, FROM table...  Error Code: 1064. You have an error in your SQL syntax; check the manual that corresponds to your MySQL se...  0.000 sec
52  18:16:09  SELECT WEEK(occured_at) AS week, COUNT(DISTINCT user_id) AS weeklyuserengagement FROM table...   17 row(s) returned                                                                      0.000 sec / 0.000 sec
53  18:21:27  Apply changes to table1_user                                                            Changes applied
54  18:22:01  SELECT COUNT(user_id) AS inc_of_user_per_day FROM table1_user GROUP BY EXTRACT(DAY FROM cre...  31 row(s) returned                                                                      0.031 sec / 0.000 sec
```

Query Completed

THIS RESULT SHOWS THE NUMBER OF USES THAT JOINED THE PRODUCT ON DAILY BASIS.

(C)Weekly Retention: Users getting retained weekly after signing-up for a product.

Your task: Calculate the weekly retention of users-sign up cohort?

ANSWER:

Select m.user_id,m.login_week,n.first as first,

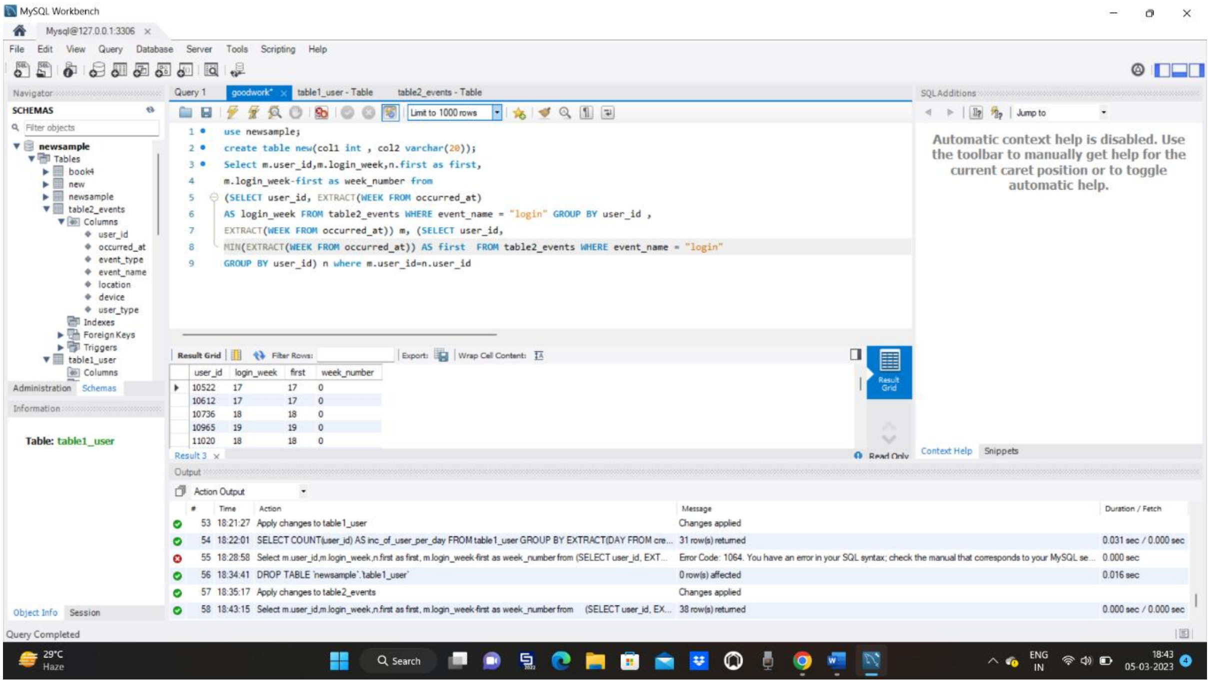
m.login_week-first as week_number from

(SELECT user_id, EXTRACT(WEEK FROM occurred_at)

AS login_week FROM table2_events WHERE event_name = "login" GROUP BY user_id , EXTRACT(WEEK FROM occurred_at)) m, (SELECT user_id,

MIN(EXTRACT(WEEK FROM occurred_at)) AS first FROM table2_events WHERE event_name = "login"

GROUP BY user_id) n where m.user_id=n.user_id



```
use newsample;
create table new(col1 int , col2 varchar(20));
Select m.user_id,m.login_week,n.first as first,
m.login_week-first as week_number from
(SELECT user_id, EXTRACT(WEEK FROM occurred_at)
AS login_week FROM table2_events WHERE event_name = "login" GROUP BY user_id ,
EXTRACT(WEEK FROM occurred_at)) m, (SELECT user_id,
MIN(EXTRACT(WEEK FROM occurred_at)) AS first  FROM table2_events WHERE event_name = "login"
GROUP BY user_id) n where m.user_id=n.user_id
```

user_id	login_week	first	week_number
10522	17	17	0
10612	17	17	0
10736	18	18	0
10965	19	19	0
11020	18	18	0

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

#	Time	Action	Message	Duration / Fetch
53	18:21:27	Apply changes to table1_user	Changes applied	
54	18:22:01	SELECT COUNT(user_id) AS inc_of_user_per_day FROM table1_user GROUP BY EXTRACT(DAY FROM cre...	31 row(s) returned	0.031 sec / 0.000 sec
55	18:28:58	Select m.user_id,m.login_week,n.first as first, m.login_week-first as week_number from (SELECT user_id, EXT...	Error Code: 1064. You have an error in your SQL syntax; check the manual that corresponds to your MySQL se...	0.000 sec
56	18:34:41	DROP TABLE `newsample`.`table1_user`	0 row(s) affected	0.016 sec
57	18:35:17	Apply changes to table2_events	Changes applied	
58	18:43:15	Select m.user_id,m.login_week,n.first as first, m.login_week-first as week_number from (SELECT user_id, EX...	38 row(s) returned	0.000 sec / 0.000 sec

The results show excellent user retention and that most of the users returned to login in the same week with a very exceptions that returned the next week.

(D)Weekly Engagement: To measure the activeness of a user. Measuring if the user finds quality in a product/service weekly. **Your task:** Calculate the weekly engagement per device?

ANSWER:

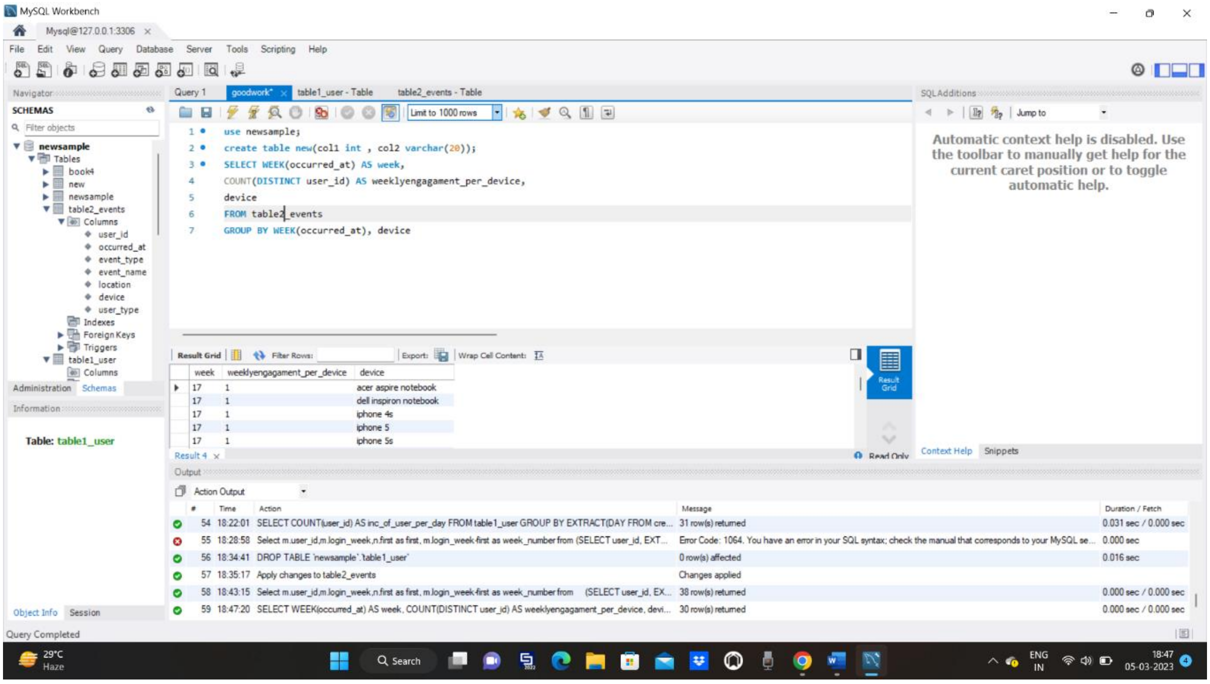
SELECT WEEK(occurred_at) AS week,

COUNT(DISTINCT user_id) AS weeklyengagament_per_device,

device

FROM table2_events

GROUP BY WEEK(occurred_at), device



```
use newsample;
create table new(col1 int , col2 varchar(20));
SELECT WEEK(occurred_at) AS week,
COUNT(DISTINCT user_id) AS weeklyengagament_per_device,
device
FROM table2_events
GROUP BY WEEK(occurred_at), device
```

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

week	weeklyengagament_per_device	device
17	1	acer aspire notebook
17	1	dell inspiron notebook
17	1	iphone 4s
17	1	iphone 5
17	1	iphone 5s

Output

Action Output

#	Time	Action	Message	Duration / Fetch
54	18:22:01	SELECT COUNT(user_id) AS inc_of_user_per_day FROM table1_user GROUP BY EXTRACT(DAY FROM cre...	31 row(s) returned	0.031 sec / 0.000 sec
55	18:28:58	Select m.user_id,m.login_week,n.first as first, m.login_week-first as week_number from (SELECT user_id, EXT...	Error Code: 1064. You have an error in your SQL syntax; check the manual that corresponds to your MySQL se...	0.000 sec
56	18:34:41	DROP TABLE 'newsample'.'table1_user'	0 row(s) affected	0.016 sec
57	18:35:17	Apply changes to table2_events	Changes applied	
58	18:43:15	Select m.user_id,m.login_week,n.first as first, m.login_week-first as week_number from (SELECT user_id, EX...	38 row(s) returned	0.000 sec / 0.000 sec
59	18:47:20	SELECT WEEK(occurred_at) AS week, COUNT(DISTINCT user_id) AS weeklyengagament_per_device, devi...	30 row(s) returned	0.000 sec / 0.000 sec

(E)Email Engagement: Users engaging with the email service.

Your task: Calculate the email engagement metrics?

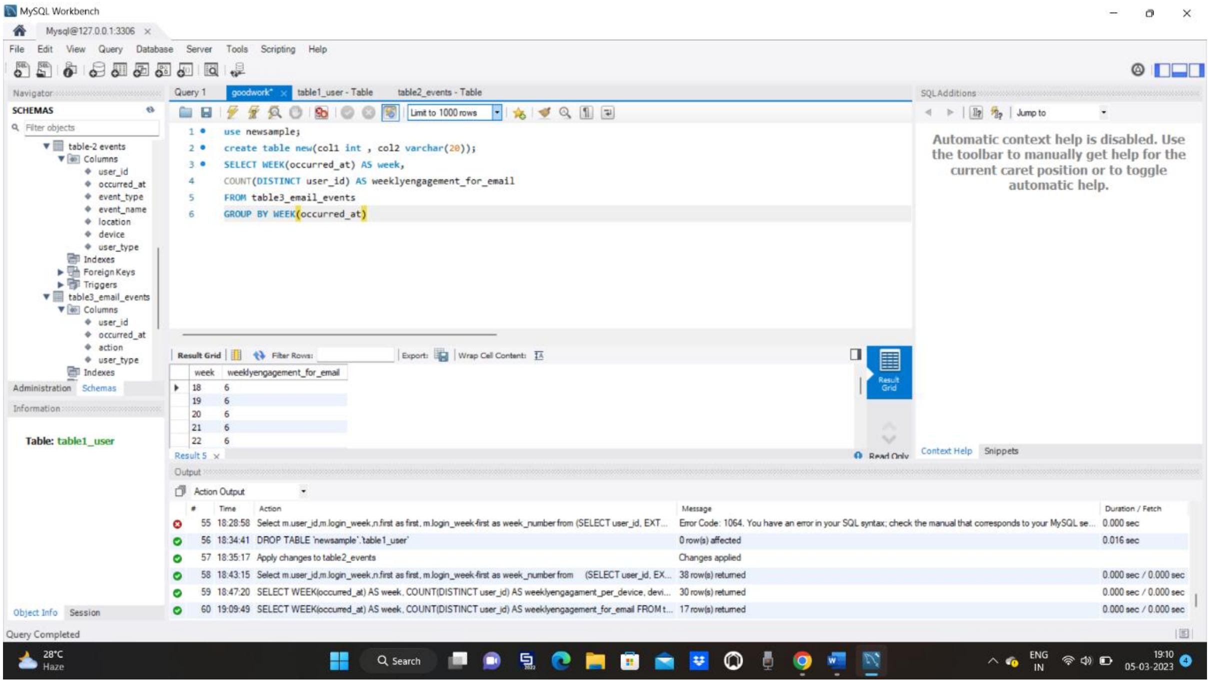
ANSWER:

SELECT WEEK(occurred_at) AS week,

COUNT(DISTINCT user_id) AS weeklyengagement_for_email

FROM table3_email_events

GROUP BY WEEK(occurred_at)



```
use newsample;
create table new(col1 int , col2 varchar(20));
SELECT WEEK(occurred_at) AS week,
COUNT(DISTINCT user_id) AS weeklyengagement_for_email
FROM table3_email_events
GROUP BY WEEK(occurred_at)
```

week	weeklyengagement_for_email
18	6
19	6
20	6
21	6
22	6

```
#   Time      Action                                                                                      Message                                                                                    Duration / Fetch
55  18:28:58  Select m.user_id,m.login_week,n.first as first, m.login_week-first as week_number from (SELECT user_id, EXT...  Error Code: 1064. You have an error in your SQL syntax; check the manual that corresponds to your MySQL se...  0.000 sec
56  18:34:41  DROP TABLE `newsample`.`table1_user`                                                        0 row(s) affected                                                                          0.016 sec
57  18:35:17  Apply changes to table2_events                                                              Changes applied
58  18:43:15  Select m.user_id,m.login_week,n.first as first, m.login_week-first as week_number from    (SELECT user_id, EX...  38 row(s) returned                                                                0.000 sec / 0.000 sec
59  18:47:20  SELECT WEEK(occurred_at) AS week, COUNT(DISTINCT user_id) AS weeklyengagament_per_device, devi...  30 row(s) returned                                                        0.000 sec / 0.000 sec
60  19:09:49  SELECT WEEK(occurred_at) AS week, COUNT(DISTINCT user_id) AS weeklyengagement_for_email FROM t...  17 row(s) returned                                                          0.000 sec / 0.000 sec
```

calculating email service retention:

Select m.user_id,m.login_week,n.first as first,

m.login_week-first as week_number from

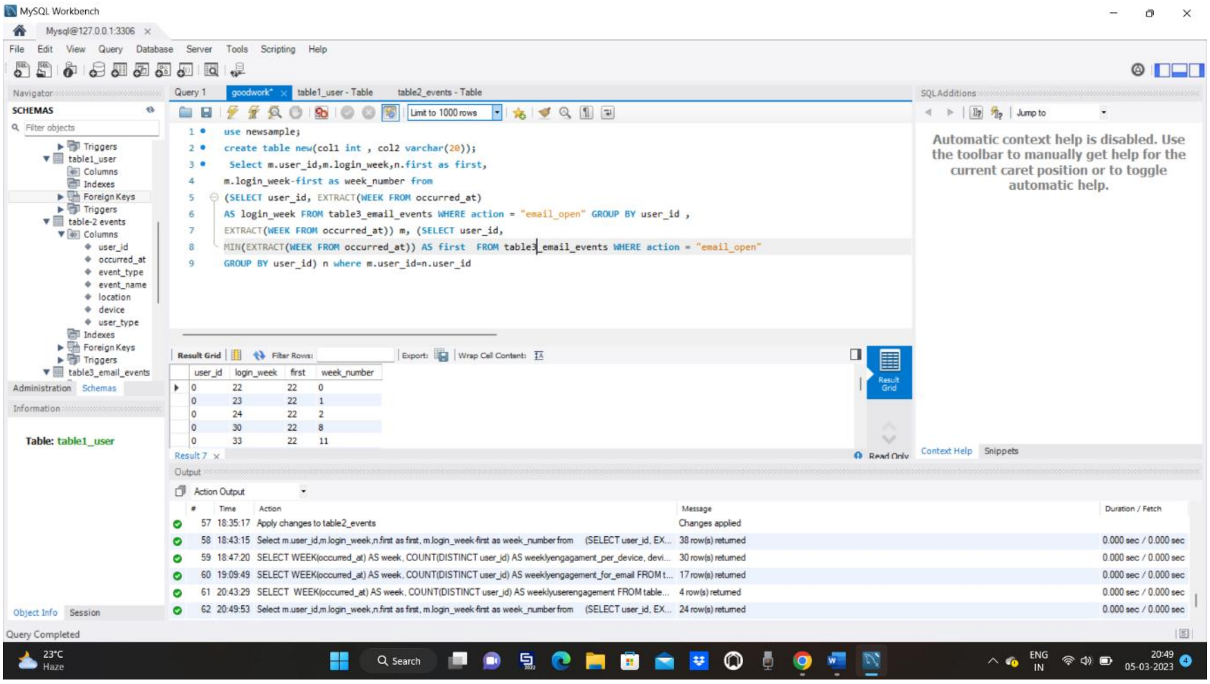
(SELECT user_id, EXTRACT(WEEK FROM occurred_at)

AS login_week FROM table3_email_events WHERE action = "email_open" GROUP BY user_id ,

EXTRACT(WEEK FROM occurred_at)) m, (SELECT user_id,

MIN(EXTRACT(WEEK FROM occurred_at)) AS first FROM table3_email_events WHERE action = "email_open"

GROUP BY user_id) n where m.user_id=n.user_id



```
use newsample;
create table new(col1 int , col2 varchar(20));
 Select m.user_id,m.login_week,n.first as first,
m.login_week-first as week_number from
(SELECT user_id, EXTRACT(WEEK FROM occurred_at)
AS login_week FROM table3_email_events WHERE action = "email_open" GROUP BY user_id ,
EXTRACT(WEEK FROM occurred_at)) m, (SELECT user_id,
MIN(EXTRACT(WEEK FROM occurred_at)) AS first  FROM table3_email_events WHERE action = "email_open"
GROUP BY user_id) n where m.user_id=n.user_id
```

user_id	login_week	first	week_number
0	22	22	0
0	23	22	1
0	24	22	2
0	30	22	8
0	33	22	11

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

```
#   Time      Action                                                                                          Message              Duration / Fetch
57  18:35:17  Apply changes to table2_events                                                                  Changes applied
58  18:43:15  Select m.user_id,m.login_week,n.first as first, m.login_week-first as week_number from    (SELECT user_id, EX...  38 row(s) returned   0.000 sec / 0.000 sec
59  18:47:20  SELECT WEEK(occurred_at) AS week, COUNT(DISTINCT user_id) AS weeklyengagament_per_device, devi...  30 row(s) returned   0.000 sec / 0.000 sec
60  19:09:49  SELECT WEEK(occurred_at) AS week, COUNT(DISTINCT user_id) AS weeklyengagement_for_email FROM t...  17 row(s) returned   0.000 sec / 0.000 sec
61  20:43:29  SELECT  WEEK(occurred_at) AS week, COUNT(DISTINCT user_id) AS weeklyuserengagement FROM table...   4 row(s) returned    0.000 sec / 0.000 sec
62  20:49:53  Select m.user_id,m.login_week,n.first as first, m.login_week-first as week_number from    (SELECT user_id, EX...  24 row(s) returned   0.000 sec / 0.000 sec
```

Query Completed

RESULT:

This project gave me deep knowledge about the topics that I learned from the mentors and while doing this project I had to search few things and concepts which again helped me in gaining more knowledge. I am happy to complete this project at my level.

THANK YOU