

Operation Analytics and Investigating Metric Spike

SQL Project Report

Description

Operation Analytics and Investigating Metric Spike with basics and fundamentals of SQL (Structured Query Language).

Approach

Case Study 1 (Job Data)

We have created a database on Jobs having data tables **Table-1:** job_data

Case Study 2 (Investigating metric spike)

We have created a database on Instagram having data tables **Table-1:** users, **Table-2:** events, **Table-3:** email_events

Tech Stack Used

Online Platform - dumb matter (<https://dumbmatter.com/csv-sql-live/>)
Database - MySQL v5.7

Insights & Results

Case Study 1 (Job Data)

- Number of jobs reviewed

CSV SQL Live Loaded Tables Add New CSV

```
SELECT ds, ROUND(1.0*COUNT(job_id)*3600/SUM(time_spent),2) FROM job_data
WHERE ds LIKE '2020-11%'
GROUP BY ds
```

Any valid SQLite query is supported. Run Query

ds	ROUND(1.0*COUNT(job_id)*3600/SUM(time_spent),2)
2020-11-25	80
2020-11-26	64.29
2020-11-27	34.62
2020-11-28	218.18
2020-11-29	180
2020-11-30	180

- Throughput

CSV SQL Live Loaded Tables Add New CSV

```
SELECT ds, ROUND(1.0*COUNT(job_id)*3600/SUM(time_spent),2) AS throughput, COUNT(job_id) AS num_jobs, SUM(time_spent) AS totaltime FROM job_data
WHERE ds LIKE '2020-11%'
GROUP BY ds
```

Any valid SQLite query is supported. Run Query

ds	throughput	num_jobs	totaltime
2020-11-25	80	1	45
2020-11-26	64.29	1	56
2020-11-27	34.62	1	104
2020-11-28	218.18	2	33
2020-11-29	180	1	20
2020-11-30	180	2	40

- Percentage share of each language

CSV SQL Live Loaded Tables Add New CSV

```
SELECT Language, COUNT(job_id) AS numjobs, ROUND(100.0*COUNT(job_id)/COUNT(job_id),2) AS percjobs FROM job_data
WHERE ds LIKE '2020-11%'
GROUP BY ds
```

Any valid SQLite query is supported. Run Query

language	numjobs	percjobs
Italian	1	100
Persian	1	100
French	1	100
Hindi	2	100
Persian	1	100
Arabic	2	100

- Duplicate rows

CSV SQL Live Loaded Tables Add New CSV

```
SELECT ds, job_id, actor_id, COUNT(*) AS cnt FROM job_data
GROUP BY actor_id having count(*) = 1
```

Any valid SQLite query is supported. Run Query

ds	job_id	actor_id	cnt
2020-11-30	21	1001	1
2020-11-28	25	1002	1
2020-11-26	23	1004	1
2020-11-28	23	1005	1
2020-11-30	22	1006	1
2020-11-27	11	1007	1

Case Study 2 (Investigating metric spike)

- User Engagement

CSV SQL Live

Loaded Tables Add New CSV

```
SELECT DISTINCT(email_events.user_id),action FROM email_events
INNER JOIN users ON users.user_id = email_events.user_id
WHERE users.state='active'
AND action like 'email%'
```

Any valid SQLite query is supported.

Run Query

user_id	action
0.0	email_open
4.0	email_clickthrough
4.0	email_open
8.0	email_clickthrough
8.0	email_open
11.0	email_clickthrough
11.0	email_open
17.0	email_clickthrough

- User Growth in AUG'22

CSV SQL Live

Loaded Tables Add New CSV

```
SELECT user_id,created_at,state FROM users
WHERE state='active' AND created_at LIKE '2014-08%'
ORDER BY created_at ASC
```

Any valid SQLite query is supported.

Run Query

user_id	created_at	state
17079.0	2014-08-01 02:01:06	active
17109.0	2014-08-01 03:17:17	active
17142.0	2014-08-01 04:32:43	active
17138.0	2014-08-01 04:41:51	active
17127.0	2014-08-01 04:49:38	active
17089.0	2014-08-01 05:10:08	active
17093.0	2014-08-01 06:15:47	active
17084.0	2014-08-01 06:40:18	active

- Weekly Retention

← → ↺ dumbmatter.com/csv-sql-live/

CSV SQL Live

Loaded Tables Add New CSV

SELECT * FROM events
INNER JOIN users ON users.user_id=events.user_id
WHERE event_name='login' AND users.activated_at IS NOT NULL

Any valid SQLite query is supported.

Run Query

user_id	occurred_at	event_type	event_name	location	device	user_type	user_id	created_at	company_id	language	activated_at	state
<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="Search"/>
10522.0	2014-05-02	engagement	login	Japan	dell inspiron	3.0	10522.0	2014-04-04	1147.0	japanese	2014-04-04	active
10612.0	2014-05-01	engagement	login	Netherlands	iphone 5	1.0	10612.0	2014-04-07	11066.0	english	2014-04-07	active
10736.0	2014-05-09	engagement	login	Austria	iphone 4s	2.0	10736.0	2014-04-09	1532.0	german	2014-04-09	active
10965.0	2014-05-15	engagement	login	Finland	windows suri	3.0	10965.0	2014-04-14	5086.0	english	2014-04-14	active
11020.0	2014-05-08	engagement	login	Japan	macbook air	2.0	11020.0	2014-04-15	11365.0	japanese	2014-04-15	active
11037.0	2014-05-01	engagement	login	United Kingd	iphone 4s	1.0	11037.0	2014-04-16	471.0	french	2014-04-16	active
11040.0	2014-05-13	engagement	login	Japan	iphone 5s	2.0	11040.0	2014-04-16	18.0	japanese	2014-04-16	active
11133.0	2014-05-08	engagement	login	India	macbook prc	2.0	11133.0	2014-04-17	35.0	indian	2014-04-17	active

- Weekly Engagement

← → ↺ dumbmatter.com/csv-sql-live/

CSV SQL Live

Loaded Tables Add New CSV







SELECT * FROM events
INNER JOIN users ON users.user_id=events.user_id
WHERE users.activated_at IS NOT NULL
GROUP BY device

Any valid SQLite query is supported.

Run Query

user_id	occurred_at	event_type	event_name	location	device	user_type	user_id	created_at	company_id	language	activated_at	state
<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="Search"/>
18947.0	2014-08-29	signup_flow	create_user	Taiwan	acer aspire c		18947.0	2014-08-29	9132.0	chinese		pending
19056.0	2014-08-31	engagement	send_messa	Pakistan	acer aspire r	2.0	19056.0	2014-08-31	1234.0	english	2014-08-31	active
19052.0	2014-08-31	signup_flow	create_user	Venezuela	amazon fire i		19052.0	2014-08-31	13087.0	spanish		pending
19026.0	2014-08-30	signup_flow	create_user	United State	asus chrome		19026.0	2014-08-30	7148.0	english		pending
19049.0	2014-08-31	engagement	search_auto	Italy	dell inspiron	3.0	19049.0	2014-08-31	85.0	spanish	2014-08-31	active
19055.0	2014-08-31	engagement	like_messag	Germany	dell inspiron	3.0	19055.0	2014-08-31	9073.0	french	2014-08-31	active
19029.0	2014-08-30	signup_flow	create_user	United State	hp pavilion d		19029.0	2014-08-30	3071.0	english		pending
19039.0	2014-08-31	engagement	home_page	United State	htc one	3.0	19039.0	2014-08-31	5815.0	english	2014-08-31	active

- Email Engagement

← → ↻ dumbmatter.com/csv-sql-live/      

CSV SQL Live Loaded Tables Add New CSV

```
SELECT * FROM email_events  
WHERE action LIKE 'email%'  
GROUP BY user_id
```

Any valid [SQLite query](#) is supported. Run Query

user_id	occurred_at	action	user_type
<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="Search"/>
10037.0	2014-07-21 09:30:26	email_open	3.0
10038.0	2014-08-18 09:30:28	email_open	3.0
10045.0	2014-08-18 09:30:52	email_clickthrough	3.0
10046.0	2014-08-18 09:30:18	email_open	1.0
10047.0	2014-07-07 09:30:29	email_open	3.0
10065.0	2014-08-26 09:30:16	email_open	3.0
10067.0	2014-08-26 09:30:26	email_open	1.0
10070.0	2014-08-26 09:30:20	email_open	1.0

Operation Analytics and Investigating Metric Spike By – Sangam Mahajan