

NAME: ARIN KUMAR

DATE : 05/02/2023

PROJECT – 3

OPERATION ANALYTICS AND INVESTIGATING METRIC SPIKE

Project Description:

This project is about analyzing two datasets, job_data and users/events/email_events. The job_data dataset includes information about the jobs reviewed, including the unique identifier of the job, actor, event, language, time spent, organization, and date. The users/events/email_events dataset includes information about user activity, such as logins, messaging events, search events, and email events. The goal of this project is to answer various questions related to job review and user engagement using SQL.

Approach:

The approach taken in this project is to first create a database and tables based on the given structure and links. Then, use SQL to perform the analysis and answer the questions related to job review and user engagement.

Tech Stack Used:

The tech stack used in this project is SQL. The specific version of SQL used will depend on the database management system used.

Insights:

The insights gained from this project are related to the job review process and user engagement with a product. The results of the analysis provide information about the number of jobs reviewed per hour per day, the average throughput, the percentage share of each language, and the weekly user engagement, growth, retention, and email engagement.

CASE STUDY – 1

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

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

	date	hour	jobs_reviewed
▶	2020-11-25	0	1
	2020-11-26	0	1
	2020-11-27	0	1
	2020-11-28	0	2
	2020-11-29	0	1
	2020-11-30	0	2

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?

	job_id	avg_time_spent_7days
▶	11	104.0000
	20	45.0000
	21	15.0000
	22	25.0000
	23	56.0000
	23	39.0000
	23	32.6667
	25	11.0000

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?

	language	language_share
►	English	12.50000
	Arabic	12.50000
	Persian	37.50000
	Hindi	12.50000
	French	12.50000
	Italian	12.50000

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?

	job_id	actor_id	event	language	time_spent	org	ds	COUNT(*)
►	21	1001	skip	English	15	A	2020-11-30	1
	22	1006	transfer	Arabic	25	B	2020-11-30	1
	23	1003	decision	Persian	20	C	2020-11-29	1
	23	1005	transfer	Persian	22	D	2020-11-28	1
	25	1002	decision	Hindi	11	B	2020-11-28	1
	11	1007	decision	French	104	D	2020-11-27	1
	23	1004	skip	Persian	56	A	2020-11-26	1
	20	1003	transfer	Italian	45	C	2020-11-25	1

CASE STUDY 2

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?

	week_start	weekly_engagement
▶	2014-04-27	85
	2014-05-04	194
	2014-05-11	208
	2014-05-18	195
	2014-05-25	208
	2014-06-01	230
	2014-06-08	224
	2014-06-15	252
	2014-06-22	245
	2014-06-29	230
	2014-07-06	249
	2014-07-13	240
	2014-07-20	253
	2014-07-27	272
	2014-08-03	231
	2014-08-10	75
	2014-08-17	20
	2014-08-24	12

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

Your task: Calculate the user growth for product?

	week	user_count
	2013-12	92
	2013-13	86
	2013-14	96
	2013-15	93
	2013-16	100
	2013-17	102
	2013-18	105
	2013-19	108
	2013-20	104
	2013-21	113
	2013-22	32

C. **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?

	event_type	week	weekly_engaged_users
►	engagement	2014-05-01	41
	engagement	2014-05-02	34
	engagement	2014-05-03	11
	engagement	2014-05-04	10
	engagement	2014-05-05	26
	engagement	2014-05-06	37
	engagement	2014-05-07	41
	engagement	2014-05-08	40
	engagement	2014-05-09	40
	engagement	2014-05-10	10
	engagement	2014-05-11	8
	engagement	2014-05-12	30
	engagement	2014-05-13	44
	engagement	2014-05-14	38
	engagement	2014-05-15	48
	engagement	2014-05-16	42
	engagement	2014-05-17	10
	engagement	2014-05-18	8
	engagement	2014-05-19	35
	engagement	2014-05-20	46

	event_type	week	weekly_engaged_users
	engagement	2014-05-21	25
	engagement	2014-05-22	47
	engagement	2014-05-23	37
	engagement	2014-05-24	11
	engagement	2014-05-25	9
	engagement	2014-05-26	28
	engagement	2014-05-27	39
	engagement	2014-05-28	41
	engagement	2014-05-29	41
	engagement	2014-05-30	52
	engagement	2014-05-31	10
	engagement	2014-06-01	15
	engagement	2014-06-02	37
	engagement	2014-06-03	38
	engagement	2014-06-04	51
	engagement	2014-06-05	42
	engagement	2014-06-06	49
	engagement	2014-06-07	12
	engagement	2014-06-08	15
	engagement	2014-06-09	41

	event_type	week	weekly_engaged_users
	engagement	2014-06-10	34
	engagement	2014-06-11	47
	engagement	2014-06-12	45
	engagement	2014-06-13	42
	engagement	2014-06-14	14
	engagement	2014-06-15	14
	engagement	2014-06-16	46
	engagement	2014-06-17	55
	engagement	2014-06-18	56
	engagement	2014-06-19	40
	engagement	2014-06-20	47
	engagement	2014-06-21	13
	engagement	2014-06-22	10
	engagement	2014-06-23	51
	engagement	2014-06-24	28
	engagement	2014-06-25	48
	engagement	2014-06-26	47
	engagement	2014-06-27	54
	engagement	2014-06-28	14
	engagement	2014-06-29	11

	event_type	week	weekly_engaged_users
	engagement	2014-06-30	40
	engagement	2014-07-01	47
	engagement	2014-07-02	50
	engagement	2014-07-03	42
	engagement	2014-07-04	45
	engagement	2014-07-05	13
	engagement	2014-07-06	10
	engagement	2014-07-07	51
	engagement	2014-07-08	49
	engagement	2014-07-09	47
	engagement	2014-07-10	39
	engagement	2014-07-11	55
	engagement	2014-07-12	12
	engagement	2014-07-13	10
	engagement	2014-07-14	40
	engagement	2014-07-15	52
	engagement	2014-07-16	60
	engagement	2014-07-17	31
	engagement	2014-07-18	47
	engagement	2014-07-19	14

	event_type	week	weekly_engaged_users
	engagement	2014-07-20	12
	engagement	2014-07-21	42
	engagement	2014-07-22	44
	engagement	2014-07-23	53
	engagement	2014-07-24	50
	engagement	2014-07-25	48
	engagement	2014-07-26	15
	engagement	2014-07-27	20
	engagement	2014-07-28	49
	engagement	2014-07-29	41
	engagement	2014-07-30	58
	engagement	2014-07-31	48
	engagement	2014-08-01	53
	engagement	2014-08-02	18
	engagement	2014-08-03	15
	engagement	2014-08-04	34
	engagement	2014-08-05	53
	engagement	2014-08-06	36
	engagement	2014-08-07	53
	engagement	2014-08-08	41

	event_type	week	weekly_engaged_users
	signup_flow	2014-05-03	8
	signup_flow	2014-05-04	9
	signup_flow	2014-05-05	24
	signup_flow	2014-05-06	27
	signup_flow	2014-05-07	32
	signup_flow	2014-05-08	33
	signup_flow	2014-05-09	31
	signup_flow	2014-05-10	7
	signup_flow	2014-05-11	6
	signup_flow	2014-05-12	29
	signup_flow	2014-05-13	35
	signup_flow	2014-05-14	34
	signup_flow	2014-05-15	38
	signup_flow	2014-05-16	36
	signup_flow	2014-05-17	7
	signup_flow	2014-05-18	7
	signup_flow	2014-05-19	31
	signup_flow	2014-05-20	38
	signup_flow	2014-05-21	22
	signup_flow	2014-05-22	35

	event_type	week	weekly_engaged_users
	signup_flow	2014-05-23	34
	signup_flow	2014-05-24	9
	signup_flow	2014-05-25	8
	signup_flow	2014-05-26	24
	signup_flow	2014-05-27	32
	signup_flow	2014-05-28	37
	signup_flow	2014-05-29	33
	signup_flow	2014-05-30	39
	signup_flow	2014-05-31	10
	signup_flow	2014-06-01	11
	signup_flow	2014-06-02	33
	signup_flow	2014-06-03	29
	signup_flow	2014-06-04	44
	signup_flow	2014-06-05	32
	signup_flow	2014-06-06	39
	signup_flow	2014-06-07	8
	signup_flow	2014-06-08	12
	signup_flow	2014-06-09	34
	signup_flow	2014-06-10	28
	signup_flow	2014-06-11	37

	event_type	week	weekly_engaged_users
	signup_flow	2014-06-12	41
	signup_flow	2014-06-13	37
	signup_flow	2014-06-14	7
	signup_flow	2014-06-15	14
	signup_flow	2014-06-16	41
	signup_flow	2014-06-17	49
	signup_flow	2014-06-18	45
	signup_flow	2014-06-19	32
	signup_flow	2014-06-20	39
	signup_flow	2014-06-21	9
	signup_flow	2014-06-22	7
	signup_flow	2014-06-23	43
	signup_flow	2014-06-24	21
	signup_flow	2014-06-25	36
	signup_flow	2014-06-26	42
	signup_flow	2014-06-27	46
	signup_flow	2014-06-28	12
	signup_flow	2014-06-29	10
	signup_flow	2014-06-30	35
	signup_flow	2014-07-01	38

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

Your task: Calculate the email engagement metrics?

	user_type	emails_sent	emails_opened	emails_clicked
▶	1	1217	1717	1529
	2	1098	1701	1529
	3	1796	2509	2219

Result:

The result of this project is a report that can be presented to the leadership team. The report includes a brief description of the project, the approach taken, the tech stack used, the insights gained, and the results of the analysis. The results provide valuable information about the job review process and user engagement with a product, which can be used to make informed decisions and improvements.