

Operation Analytics and Investigating Metric Spike

Project Description:

This project involves using advanced SQL to analyze various datasets and investigate sudden metric spikes within a company's operations. As a Lead Data Analyst, you will derive insights to explain changes in key performance metrics and provide actionable recommendations to improve efficiency. The analysis will cover user engagement, sales, support tickets, and marketing data. Deliverables include SQL scripts, detailed reports, and visualizations to aid decision-making for different departments.

Approach:

- Engage with stakeholders to understand data needs and define key metrics.
- Import, clean, and organize data into structured tables for analysis.
- Use advanced SQL to analyze data, identify trends, and investigate metric spikes.

Case Study 1: Job Data Analysis:

1) Jobs Reviewed Over Time:

This SQL query calculates the average number of distinct jobs reviewed per hour. It does this by counting the unique job_id in the job_data table and dividing that count by the total number of hours in 30 days (30 * 24).

SQL Query:

```
select count(distinct job_id)/(30*24) as num_jobs_reviewed
from job_data;
```

Result:

Result Grid		Filter Rows:
	num_jobs_reviewed	
▶	0.0486	



2) Throughput Analysis:



This SQL query calculates the number of distinct jobs reviewed per day (job_review) and the 7-day rolling average of jobs reviewed (throughput_rolling). It groups the data by date (ds), and for each day, it averages the job reviews over the current day and the previous six days. The results are then ordered by date (ds).

SQL Query:

```
select ds, count(distinct job_id) as job_review,  
avg(count(distinct job_id)) over(order by ds rows between 6 preceding and current row)  
as throughput_rolling  
from job_data  
group by ds  
order by ds;
```

Result:

Result Grid		 Filter Rows:	
	ds	job_review	throughput_rolling
▶	10/01/2020	1	1.0000
	10/08/2020	1	1.0000
	10/09/2020	1	1.0000
	10/10/2020	1	1.0000
	10/20/2020	1	1.0000
	10/26/2020	1	1.0000
	10/27/2020	1	1.0000
	10/28/2020	2	1.1429
	10/30/2020	1	1.1429
	10/31/2020	3	1.4286
	11/01/2020	2	1.5714
	11/08/2020	2	1.7143
	11/09/2020	1	1.7143
	11/10/2020	2	1.8571
	11/18/2020	2	1.8571
	11/19/2020	1	1.8571
	11/20/2020	2	1.7143
	11/25/2020	1	1.5714

Result Grid			 Filter Rows:	
	ds	job_review	throughput_rolling	
	11/25/2020	1	1.5714	
	11/26/2020	1	1.4286	
	11/27/2020	1	1.4286	
	11/28/2020	2	1.4286	
	11/29/2020	1	1.2857	
	11/30/2020	2	1.4286	
	12/01/2020	1	1.2857	
	12/08/2020	1	1.2857	
	12/10/2020	1	1.2857	
	12/18/2020	2	1.4286	
	12/19/2020	1	1.2857	
	12/20/2020	2	1.4286	
	12/25/2020	1	1.2857	
	12/26/2020	1	1.2857	
	12/27/2020	1	1.2857	
	12/30/2020	1	1.2857	
	12/31/2020	1	1.1429	

3) Language Share Analysis:

This SQL query calculates the percentage representation of each language within the job_data table. It groups the data by language, counts the occurrences of each language, and then divides this count by the total number of entries in the table. The result is multiplied by 100 to express the proportion as a percentage.

SQL Query:

```
select language, (count(language)/sum(count(*)) over()) * 100
as percentage_of_each_language
from job_data
group by language;
```

Result:

Result Grid	Filter Rows:
language	percentage_of_each_language
English	4.1667
Arabic	4.1667
Persian	12.5000
Hindi	4.1667
French	4.1667
Italian	4.1667
German	4.1667
Spanish	4.1667
Portuguese	4.1667
Japanese	4.1667
Russian	4.1667
Swedish	4.1667

Result Grid	Filter Rows:
language	percentage_of_each_language
Russian	4.1667
Swedish	4.1667
Dutch	4.1667
Danish	4.1667
Norwegian	6.2500
Finnish	4.1667
Czech	4.1667
Slovak	4.1667
Polish	4.1667
Hungarian	4.1667
Romanian	6.2500

4) Duplicate Rows Detection:


This SQL query identifies and selects records from the job_data table that have duplicate entries based on the specified columns (DS, job_id, actor_id, event, language, time_spent, org). The having count(*) > 1 clause filters out groups that appear more than once.

SQL Query:

```
select *
from job_data
group by DS, job_id, actor_id, event, language, time_spent, org
having count(*) > 1;
```


Result:

Result Grid



Filter Rows:

Export:



Wrap Cell Content

	ds	job_id	actor_id	event	language	time_spent	org
▶	10/30/2020	24	1023	decision	Romanian	18	A
	10/20/2020	31	1041	skip	Norwegian	10	C

Case Study 2: Investigating Metric Spike

1. Weekly User Engagement:

This SQL query calculates the number of distinct active users per week based on engagement events. It extracts the week number from the occurred_at timestamp, counts the distinct user_ids where event_type is 'engagement', groups the results by week number, and orders them by week number.

SQL Query:

```
select extract(week from occurred_at) as week_number,  
count(distinct user_id) as active_user  
from events  
where event_type = 'engagement'  
group by week_number  
order by week_number;
```

Result:

Result Grid	Filter Rows:
week_number	active_user
17	663
18	1068
19	1113
20	1154
21	1121
22	1186
23	1232
24	1275
25	1264

Result Grid	Filter Rows:
week_number	active_user
27	1372
28	1365
29	1376
30	1467
31	1299
32	1225
33	1225
34	1204
35	104



2. User Growth Analysis:



This SQL query calculates the cumulative total of active users by week and year. The inner query extracts the year and week number from the created_at timestamp, counts the distinct user_ids per week, and groups the results accordingly. The outer query then computes the running total of active users using the SUM() window function, ordered by year and week number. The final result shows the year, week number, weekly active users, and the cumulative total of active users up to that week.



SQL Query:



```
select year, no_of_weeks, active_users, sum(active_users) over(order by year,  
no_of_weeks rows between unbounded preceding and current row) as total_active_user  
from (select extract(year from created_at) as year,  
extract(week from created_at) as no_of_weeks,  
count(distinct user_id) as active_users  
from users  
group by year, no_of_weeks  
order by year, no_of_weeks)a;
```

Result:

Result Grid		 Filter Rows:	<input type="text"/>	Export
	year	no_of_weeks	active_users	total_active_user
▶	2013	0	23	23
	2013	1	30	53
	2013	2	48	101
	2013	3	36	137
	2013	4	30	167
	2013	5	48	215
	2013	6	38	253
	2013	7	42	295
	2013	8	34	329
	2013	9	43	372
	2013	10	32	404
	2013	11	31	435
	2013	12	33	468
	2013	13	39	507
	2013	14	35	542
	2013	15	43	585
	2013	16	46	631
	2013	17	49	680
	2013	18	44	724
	2013	19	57	781
	2013	20	39	820
	2013	21	49	869
	2013	22	54	923

Result Grid			Filter Rows:	<input type="text"/>	Export:
	year	no_of_weeks	active_users	total_active_user	
	2013	22	54	923	
	2013	23	50	973	
	2013	24	45	1018	
	2013	25	57	1075	
	2013	26	56	1131	
	2013	27	52	1183	
	2013	28	72	1255	
	2013	29	67	1322	
	2013	30	67	1389	
	2013	31	67	1456	
	2013	32	71	1527	
	2013	33	73	1600	
	2013	34	78	1678	
	2013	35	63	1741	
	2013	36	72	1813	
	2013	37	85	1898	
	2013	38	90	1988	
	2013	39	84	2072	
	2013	40	87	2159	
	2013	41	73	2232	
	2013	42	99	2331	
	2013	43	89	2420	
	2013	44	96	2516	

Result Grid		 Filter Rows:	<input type="text"/>	Export
	year	no_of_weeks	active_users	total_active_user
	2013	44	96	2516
	2013	45	91	2607
	2013	46	88	2695
	2013	47	102	2797
	2013	48	97	2894
	2013	49	116	3010
	2013	50	124	3134
	2013	51	102	3236
	2013	52	47	3283
	2014	0	83	3366
	2014	1	126	3492
	2014	2	109	3601
	2014	3	113	3714
	2014	4	130	3844
	2014	5	133	3977
	2014	6	135	4112
	2014	7	125	4237
	2014	8	129	4366
	2014	9	133	4499
	2014	10	154	4653
	2014	11	130	4783
	2014	12	148	4931
	2014	13	167	5098

Result Grid		 Filter Rows:	<input type="text"/>	<input type="button" value="Export"/>
	year	no_of_weeks	active_users	total_active_user
	2014	13	167	5098
	2014	14	162	5260
	2014	15	164	5424
	2014	16	179	5603
	2014	17	170	5773
	2014	18	163	5936
	2014	19	185	6121
	2014	20	176	6297
	2014	21	183	6480
	2014	22	196	6676
	2014	23	196	6872
	2014	24	229	7101
	2014	25	207	7308
	2014	26	201	7509
	2014	27	222	7731
	2014	28	215	7946
	2014	29	221	8167
	2014	30	238	8405
	2014	31	193	8598
	2014	32	245	8843
	2014	33	261	9104
	2014	34	259	9363
	2014	35	18	9381

3. Weekly Retention Analysis:

This SQL query calculates the cumulative total of active users by week and year. The inner query extracts the year and week number from the created_at timestamp, counts the distinct user_ids per week, and groups the results accordingly. The outer query then computes the running total of active users using the SUM() window function, ordered by year and week number. The final result shows the year, week number, weekly active users, and the cumulative total of active users up to that week.

SQL Query:

```
select user_id, count(user_id) as no_of_user_visit,
sum(case when retention_week = 1 then 1 else 0 end) as per_week_retention
from(
select a.user_id, a.signup_week,
      b.engagement_week, b.engagement_week - a.signup_week as retention_week
from ((select distinct user_id, extract(week from occurred_at) as signup_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_week
      from events
      where event_type = 'engagement')b
on a.user_id = b.user_id))d
group by user_id
order by user_id;
```

Result:

	user_id	no_of_user_visit	per_week_retention
	11919	2	0
	11920	1	0
	11924	1	0
	11926	8	1
	11928	8	0
	11929	1	0
	11931	6	1
	11933	6	1
	11936	3	0
	11939	3	1
	11940	4	1
	11942	7	1
	11944	3	1
	11947	2	1
	11949	7	1
	11953	3	1
	11955	5	0
	11960	2	1
	11961	1	0
	11962	3	1
	11963	4	1
	11971	1	0
	11972	6	0
	11973	1	0
	11975	2	1

	user_id	no_of_user_visit	per_week_retention
	11981	1	0
	11984	5	1
	11986	6	1
	11990	3	0
	11991	1	0
	11992	7	1
	11994	1	0
	11995	1	0
	11997	1	0
	11998	6	0
	11999	3	1
	12001	8	1
	12005	2	0
	12007	11	1
	12010	2	0
	12014	5	1
▶	12019	1	0
	12021	1	0
	12022	1	0
	12023	5	1
	12026	1	0
	12027	1	0
	12028	1	0
	12029	4	1
	12030	1	0

	user_id	no_of_user_visit	per_week_retention
	12031	1	0
	12033	6	1
	12034	14	1
	12035	2	0
	12040	7	0
	12042	3	1
	12045	2	0
	12048	3	1
	12050	5	1
	12053	3	1
	12054	9	1
	12056	3	1
	12057	1	0
	12058	5	1
	12061	2	1
	12063	3	1
	12066	4	1
	12072	5	1
	12073	4	1
	12074	1	0
	12076	3	1
	12077	2	1
	12079	9	1
	12081	6	1
	12082	1	0

	user_id	no_of_user_visit	per_week_retention
	12084	3	1
	12085	2	1
	12087	3	1
	12089	2	1
	12093	2	1
	12095	9	1
	12097	4	1
	12101	2	1
	12102	3	1
	12103	1	0
	12106	1	0
	12107	6	1
	12111	2	1
	12113	2	0
	12118	3	1
	12119	4	1
	12120	1	0
	12121	1	0
	12122	4	0
	12127	7	1
	12133	2	1
	12135	2	1
	12136	3	1
	12137	4	1
	12138	4	1

	user_id	no_of_user_visit	per_week_retention
	12141	2	1
	12144	5	1
	12147	4	1
	12148	4	1
	12149	4	1
	12151	1	0
	12152	8	1
	12153	5	1
	12161	6	1
	12162	4	0
	12164	13	1
	12165	3	1
	12167	3	1
	12169	4	1
	12170	6	1
	12171	1	0
	12173	3	1
	12174	1	0
	12176	3	1
	12178	6	1
	12180	2	1
	12181	4	1
	12183	3	1
	12184	2	1
	12185	6	1

	user_id	no_of_user_visit	per_week_retention
	12188	3	1
	12189	5	1
	12191	3	1
	12193	2	1
	12196	1	0
	12202	2	1
	12205	13	1
	12208	2	1
	12212	5	1
	12213	3	1
	12214	3	1
	12216	9	1
	12217	3	1
	12221	7	1
	12225	2	1
	12227	3	1
	12228	7	1
	12230	2	1
	12233	1	0
	12235	1	0
	12236	4	1
	12237	8	1
	12238	2	1
	12241	7	1
	12243	2	0

	12243	2	0
	12246	9	1
	12247	3	1
	12248	12	1
	12249	5	1
	12250	6	1
	12251	5	1
	12256	2	1

	12256	2	1
	12263	7	1
	12265	4	1
	12266	4	1
	12267	2	1
	12268	1	0
	12270	2	1

4. Weekly Engagement Per Device:

This SQL query calculates the number of distinct users engaging with the system, broken down by year, week, and device type. It extracts the year and week from the occurred_at timestamp and groups the data by these values and device. The count(distinct user_id) function counts unique users for each combination of year, week, and device. The results are then ordered by year, week, and device.

SQL Query:

```
select year(occurred_at) as year,
week(occurred_at) as week, device,
count(distinct user_id) as no_of_user
from events
where event_type = 'engagement'
group by year, week, device
order by year, week, device;
```

Result:

	year	week	device	no_of_user
►	2014	17	acer aspire desktop	9
	2014	17	acer aspire notebook	20
	2014	17	amazon fire phone	4
	2014	17	asus chromebook	21
	2014	17	dell inspiron desktop	18
	2014	17	dell inspiron notebook	46
	2014	17	hp pavilion desktop	14
	2014	17	htc one	16
	2014	17	ipad air	27
	2014	17	ipad mini	19
	2014	17	iphone 4s	21
	2014	17	iphone 5	65
	2014	17	iphone 5s	42
	2014	17	kindle fire	6
	2014	17	lenovo thinkpad	86
	2014	17	mac mini	6
	2014	17	macbook air	54
	2014	17	macbook pro	143
	2014	17	nexus 10	16
	2014	17	nexus 5	40
	2014	17	nexus 7	18
	2014	17	nokia lumia 635	17
	2014	17	samsung galaxy tablet	8
	2014	17	samsung galaxy note	7
	2014	17	samsung galaxy s4	52

	year	week	device	no_of_user
	2014	17	windows surface	10
	2014	18	acer aspire desktop	26
	2014	18	acer aspire notebook	33
	2014	18	amazon fire phone	9
	2014	18	asus chromebook	42
	2014	18	dell inspiron desktop	58
	2014	18	dell inspiron notebook	77
	2014	18	hp pavilion desktop	37
	2014	18	htc one	19
	2014	18	ipad air	52
	2014	18	ipad mini	30
	2014	18	iphone 4s	46
	2014	18	iphone 5	113
	2014	18	iphone 5s	73
	2014	18	kindle fire	27
	2014	18	lenovo thinkpad	153
	2014	18	mac mini	13
	2014	18	macbook air	121
	2014	18	macbook pro	252
	2014	18	nexus 10	30
	2014	18	nexus 5	73
	2014	18	nexus 7	30
	2014	18	nokia lumia 635	33
	2014	18	samsung galaxy tablet	11
	2014	18	samsung galaxy note	15

	year	week	device	no_of_user
	2014	18	samsung galaxy s4	82
	2014	18	windows surface	10
	2014	19	acer aspire desktop	23
	2014	19	acer aspire notebook	41
	2014	19	amazon fire phone	12
	2014	19	asus chromebook	27
	2014	19	dell inspiron desktop	36
	2014	19	dell inspiron notebook	83
	2014	19	hp pavilion desktop	40
	2014	19	htc one	30
	2014	19	ipad air	55
	2014	19	ipad mini	36
	2014	19	iphone 4s	44
	2014	19	iphone 5	115
	2014	19	iphone 5s	79
	2014	19	kindle fire	21
	2014	19	lenovo thinkpad	178
	2014	19	mac mini	18
	2014	19	macbook air	112
	2014	19	macbook pro	266
	2014	19	nexus 10	25
	2014	19	nexus 5	87
	2014	19	nexus 7	41
	2014	19	nokia lumia 635	23
	2014	19	samsung galaxy tablet	6

	year	week	device	no_of_user
	2014	19	samsung galaxy note	11
	2014	19	samsung galaxy s4	91
	2014	19	windows surface	16
	2014	20	acer aspire desktop	23
	2014	20	acer aspire notebook	40
	2014	20	amazon fire phone	11
	2014	20	asus chromebook	41
	2014	20	dell inspiron desktop	52
	2014	20	dell inspiron notebook	84
	2014	20	hp pavilion desktop	30
	2014	20	htc one	29
	2014	20	ipad air	59
	2014	20	ipad mini	32
	2014	20	iphone 4s	55
	2014	20	iphone 5	125
	2014	20	iphone 5s	79
	2014	20	kindle fire	23
	2014	20	lenovo thinkpad	173
	2014	20	mac mini	26
	2014	20	macbook air	119
	2014	20	macbook pro	256
	2014	20	nexus 10	22
	2014	20	nexus 5	103
	2014	20	nexus 7	32
	2014	20	nokia lumia 635	22

	year	week	device	no_of_user
	2014	20	samsung galaxy tablet	9
	2014	20	samsung galaxy note	18
	2014	20	samsung galaxy s4	93
	2014	20	windows surface	21
	2014	21	acer aspire desktop	29
	2014	21	acer aspire notebook	47
	2014	21	amazon fire phone	5
	2014	21	asus chromebook	38
	2014	21	dell inspiron desktop	41
	2014	21	dell inspiron notebook	80
	2014	21	hp pavilion desktop	44
	2014	21	htc one	21
	2014	21	ipad air	51
	2014	21	ipad mini	23
	2014	21	iphone 4s	45
	2014	21	iphone 5	137
	2014	21	iphone 5s	74
	2014	21	kindle fire	30
	2014	21	lenovo thinkpad	167
	2014	21	mac mini	18
	2014	21	macbook air	110
	2014	21	macbook pro	247
	2014	21	nexus 10	25
	2014	21	nexus 5	91
	2014	21	nexus 7	29

	year	week	device	no_of_user
	2014	21	nokia lumia 635	25
	2014	21	samsung galaxy tablet	6
	2014	21	samsung galaxy note	20
	2014	21	samsung galaxy s4	84
	2014	21	windows surface	17
	2014	22	acer aspire desktop	25
	2014	22	acer aspire notebook	41
	2014	22	amazon fire phone	5
	2014	22	asus chromebook	52
	2014	22	dell inspiron desktop	52
	2014	22	dell inspiron notebook	92
	2014	22	hp pavilion desktop	38
	2014	22	htc one	24
	2014	22	ipad air	58
	2014	22	ipad mini	34
	2014	22	iphone 4s	45
	2014	22	iphone 5	125
	2014	22	iphone 5s	71
	2014	22	kindle fire	21
	2014	22	lenovo thinkpad	176
	2014	22	mac mini	25
	2014	22	macbook air	145
	2014	22	macbook pro	251
	2014	22	nexus 10	27
	2014	22	nexus 5	96

	year	week	device	no_of_user
	2014	22	nexus 7	45
	2014	22	nokia lumia 635	25
	2014	22	samsung galaxy tablet	10
	2014	22	samsung galaxy note	19
	2014	22	samsung galaxy s4	105
	2014	22	windows surface	15
	2014	23	acer aspire desktop	22
	2014	23	acer aspire notebook	43
	2014	23	amazon fire phone	16
	2014	23	asus chromebook	49
	2014	23	dell inspiron desktop	53
	2014	23	dell inspiron notebook	103
	2014	23	hp pavilion desktop	54

	year	week	device	no_of_user
	2014	23	hp pavilion desktop	54
	2014	23	htc one	20
	2014	23	ipad air	41
	2014	23	ipad mini	33
	2014	23	iphone 4s	53
	2014	23	iphone 5	152
	2014	23	iphone 5s	79
	2014	23	kindle fire	25
	2014	23	lenovo thinkpad	176
	2014	23	mac mini	18
	2014	23	macbook air	124
	2014	23	macbook pro	266
	2014	23	nexus 10	45

	year	week	device	no_of_user
	2014	23	nexus 5	88
	2014	23	nexus 7	36
	2014	23	nokia lumia 635	31
	2014	23	samsung galaxy tablet	14
	2014	23	samsung galaxy note	14
	2014	23	samsung galaxy s4	99
	2014	23	windows surface	14
	2014	24	acer aspire desktop	24
	2014	24	acer aspire notebook	40
	2014	24	amazon fire phone	11
	2014	24	asus chromebook	43
	2014	24	dell inspiron desktop	59
	2014	24	dell inspiron notebook	99
	2014	24	hp pavilion desktop	56
	2014	24	htc one	20
	2014	24	ipad air	57
	2014	24	ipad mini	39
	2014	24	iphone 4s	53
	2014	24	iphone 5	142
	2014	24	iphone 5s	79
	2014	24	kindle fire	25
	2014	24	lenovo thinkpad	165
	2014	24	mac mini	29
	2014	24	macbook air	152
	2014	24	macbook pro	255

	year	week	device	no_of_user
	2014	24	nexus 10	38
	2014	24	nexus 5	87
	2014	24	nexus 7	49
	2014	24	nokia lumia 635	35
	2014	24	samsung galaxy tablet	11
	2014	24	samsung galaxy note	20
	2014	24	samsung galaxy s4	101
	2014	24	windows surface	22
	2014	25	acer aspire desktop	28
	2014	25	acer aspire notebook	47
	2014	25	amazon fire phone	13
	2014	25	asus chromebook	38
	2014	25	dell inspiron desktop	52
	2014	25	dell inspiron notebook	105
	2014	25	hp pavilion desktop	52
	2014	25	htc one	21
	2014	25	ipad air	57
	2014	25	ipad mini	30
	2014	25	iphone 4s	40
	2014	25	iphone 5	137
	2014	25	iphone 5s	78
	2014	25	kindle fire	24
	2014	25	lenovo thinkpad	197
	2014	25	mac mini	21
	2014	25	macbook air	121

	year	week	device	no_of_user
	2014	25	macbook pro	275
	2014	25	nexus 10	29
	2014	25	nexus 5	89
	2014	25	nexus 7	51
	2014	25	nokia lumia 635	37
	2014	25	samsung galaxy tablet	12
	2014	25	samsung galaxy note	14
	2014	25	samsung galaxy s4	99
	2014	25	windows surface	22
	2014	26	acer aspire desktop	29
	2014	26	acer aspire notebook	35
	2014	26	amazon fire phone	13
	2014	26	asus chromebook	49
	2014	26	dell inspiron desktop	60
	2014	26	dell inspiron notebook	89
	2014	26	hp pavilion desktop	46
	2014	26	htc one	23
	2014	26	ipad air	56
	2014	26	ipad mini	43
	2014	26	iphone 4s	50
	2014	26	iphone 5	152
	2014	26	iphone 5s	94
	2014	26	kindle fire	26
	2014	26	lenovo thinkpad	192
	2014	26	mac mini	11

	year	week	device	no_of_user
	2014	26	macbook air	134
	2014	26	macbook pro	269
	2014	26	nexus 10	29
	2014	26	nexus 5	87
	2014	26	nexus 7	46
	2014	26	nokia lumia 635	42
	2014	26	samsung galaxy tablet	12
	2014	26	samsung galaxy note	9
	2014	26	samsung galaxy s4	112
	2014	26	windows surface	21
	2014	27	acer aspire desktop	29
	2014	27	acer aspire notebook	49
	2014	27	amazon fire phone	10
	2014	27	asus chromebook	52
	2014	27	dell inspiron desktop	53
	2014	27	dell inspiron notebook	89
	2014	27	hp pavilion desktop	56
	2014	27	htc one	27
	2014	27	ipad air	55
	2014	27	ipad mini	35
	2014	27	iphone 4s	67
	2014	27	iphone 5	163
	2014	27	iphone 5s	83
	2014	27	kindle fire	25
	2014	27	lenovo thinkpad	202

	year	week	device	no_of_user
	2014	27	mac mini	15
	2014	27	macbook air	142
	2014	27	macbook pro	302
	2014	27	nexus 10	37
	2014	27	nexus 5	84
	2014	27	nexus 7	40
	2014	27	nokia lumia 635	31
	2014	27	samsung galaxy tablet	15
	2014	27	samsung galaxy note	15
	2014	27	samsung galaxy s4	116
	2014	27	windows surface	33
	2014	28	acer aspire desktop	30

	year	week	device	no_of_user
	2014	28	acer aspire notebook	49
	2014	28	amazon fire phone	6
	2014	28	asus chromebook	50
	2014	28	dell inspiron desktop	56
	2014	28	dell inspiron notebook	103
	2014	28	hp pavilion desktop	56
	2014	28	htc one	26
	2014	28	ipad air	54
	2014	28	ipad mini	35
	2014	28	iphone 4s	61
	2014	28	iphone 5	151
	2014	28	iphone 5s	93
	2014	28	kindle fire	31

	year	week	device	no_of_user
	2014	28	lenovo thinkpad	220
	2014	28	mac mini	28
	2014	28	macbook air	148
	2014	28	macbook pro	295
	2014	28	nexus 10	26
	2014	28	nexus 5	85
	2014	28	nexus 7	39
	2014	28	nokia lumia 635	35
	2014	28	samsung galaxy tablet	9
	2014	28	samsung galaxy note	10
	2014	28	samsung galaxy s4	122
	2014	28	windows surface	33
	2014	29	acer aspire desktop	28
	2014	29	acer aspire notebook	53
	2014	29	amazon fire phone	12
	2014	29	asus chromebook	49
	2014	29	dell inspiron desktop	54
	2014	29	dell inspiron notebook	113
	2014	29	hp pavilion desktop	58
	2014	29	htc one	31
	2014	29	ipad air	52
	2014	29	ipad mini	34
	2014	29	iphone 4s	60
	2014	29	iphone 5	144
	2014	29	iphone 5s	90

	year	week	device	no_of_user
	2014	29	kindle fire	37
	2014	29	lenovo thinkpad	209
	2014	29	mac mini	31
	2014	29	macbook air	148
	2014	29	macbook pro	295
	2014	29	nexus 10	25
	2014	29	nexus 5	77
	2014	29	nexus 7	45
	2014	29	nokia lumia 635	43
	2014	29	samsung galaxy tablet	13
	2014	29	samsung galaxy note	16
	2014	29	samsung galaxy s4	123
	2014	29	windows surface	28
	2014	30	acer aspire desktop	33
	2014	30	acer aspire notebook	60
	2014	30	amazon fire phone	12
	2014	30	asus chromebook	56
	2014	30	dell inspiron desktop	54
	2014	30	dell inspiron notebook	127
	2014	30	hp pavilion desktop	42
	2014	30	htc one	31
	2014	30	ipad air	70
	2014	30	ipad mini	35
	2014	30	iphone 4s	65
	2014	30	iphone 5	152

	year	week	device	no_of_user
	2014	30	iphone 5s	103
	2014	30	kindle fire	25
	2014	30	lenovo thinkpad	206
	2014	30	mac mini	23
	2014	30	macbook air	159
	2014	30	macbook pro	322
	2014	30	nexus 10	36
	2014	30	nexus 5	84
	2014	30	nexus 7	62
	2014	30	nokia lumia 635	34
	2014	30	samsung galaxy tablet	9
	2014	30	samsung galaxy note	15
	2014	30	samsung galaxy s4	103
	2014	30	windows surface	19
	2014	31	acer aspire desktop	31
	2014	31	acer aspire notebook	55
	2014	31	amazon fire phone	14
	2014	31	asus chromebook	56
	2014	31	dell inspiron desktop	44
	2014	31	dell inspiron notebook	113
	2014	31	hp pavilion desktop	51
	2014	31	htc one	13
	2014	31	ipad air	55
	2014	31	ipad mini	27
	2014	31	iphone 4s	56

	year	week	device	no_of_user
	2014	31	iphone 5	135
	2014	31	iphone 5s	71
	2014	31	kindle fire	14
	2014	31	lenovo thinkpad	207
	2014	31	mac mini	24
	2014	31	macbook air	147
	2014	31	macbook pro	321
	2014	31	nexus 10	24
	2014	31	nexus 5	69
	2014	31	nexus 7	38
	2014	31	nokia lumia 635	28
	2014	31	samsung galaxy tablet	8
	2014	31	samsung galaxy note	14
	2014	31	samsung galaxy s4	100
	2014	31	windows surface	19
	2014	32	acer aspire desktop	35
	2014	32	acer aspire notebook	55
	2014	32	amazon fire phone	12
	2014	32	asus chromebook	62
	2014	32	dell inspiron desktop	57
	2014	32	dell inspiron notebook	104
	2014	32	hp pavilion desktop	51
	2014	32	htc one	18
	2014	32	ipad air	48
	2014	32	ipad mini	30

	year	week	device	no_of_user
	2014	32	iphone 4s	34
	2014	32	iphone 5	119
	2014	32	iphone 5s	67
	2014	32	kindle fire	12
	2014	32	lenovo thinkpad	179
	2014	32	mac mini	20
	2014	32	macbook air	125
	2014	32	macbook pro	307
	2014	32	nexus 10	30
	2014	32	nexus 5	67
	2014	32	nexus 7	25
	2014	32	nokia lumia 635	28
	2014	32	samsung galaxy tablet	6

	year	week	device	no_of_user
	2014	32	samsung galaxy note	12
	2014	32	samsung galaxy s4	82
	2014	32	windows surface	10
	2014	33	acer aspire desktop	39
	2014	33	acer aspire notebook	46
	2014	33	amazon fire phone	14
	2014	33	asus chromebook	49
	2014	33	dell inspiron desktop	37
	2014	33	dell inspiron notebook	110
	2014	33	hp pavilion desktop	38
	2014	33	htc one	19
	2014	33	ipad air	40

	year	week	device	no_of_user
	2014	33	ipad mini	28
	2014	33	iphone 4s	35
	2014	33	iphone 5	110
	2014	33	iphone 5s	65
	2014	33	kindle fire	14
	2014	33	lenovo thinkpad	191
	2014	33	mac mini	32
	2014	33	macbook air	133
	2014	33	macbook pro	312
	2014	33	nexus 10	23
	2014	33	nexus 5	70
	2014	33	nexus 7	30
	2014	33	nokia lumia 635	27
	2014	33	samsung galaxy tablet	12
	2014	33	samsung galaxy note	13
	2014	33	samsung galaxy s4	80
	2014	33	windows surface	15
	2014	34	acer aspire desktop	30
	2014	34	acer aspire notebook	63
	2014	34	amazon fire phone	11
	2014	34	asus chromebook	47
	2014	34	dell inspiron desktop	49
	2014	34	dell inspiron notebook	105
	2014	34	hp pavilion desktop	36
	2014	34	htc one	25

	year	week	device	no_of_user
	2014	34	ipad air	39
	2014	34	ipad mini	25
	2014	34	iphone 4s	50
	2014	34	iphone 5	101
	2014	34	iphone 5s	70
	2014	34	kindle fire	13
	2014	34	lenovo thinkpad	193
	2014	34	mac mini	30
	2014	34	macbook air	136
	2014	34	macbook pro	292
	2014	34	nexus 10	25
	2014	34	nexus 5	70
	2014	34	nexus 7	33
	2014	34	nokia lumia 635	17
	2014	34	samsung galaxy tablet	14
	2014	34	samsung galaxy note	13
	2014	34	samsung galaxy s4	90
	2014	34	windows surface	18
	2014	35	acer aspire desktop	1
	2014	35	acer aspire notebook	3
	2014	35	asus chromebook	6
	2014	35	dell inspiron desktop	1
	2014	35	dell inspiron notebook	9
	2014	35	hp pavilion desktop	1
	2014	35	htc one	2

	year	week	device	no_of_user
	2014	34	samsung galaxy s4	90
	2014	34	windows surface	18
	2014	35	acer aspire desktop	1
	2014	35	acer aspire notebook	3
	2014	35	asus chromebook	6
	2014	35	dell inspiron desktop	1
	2014	35	dell inspiron notebook	9
	2014	35	hp pavilion desktop	1
	2014	35	htc one	2
	2014	35	ipad mini	2
	2014	35	iphone 4s	6
	2014	35	iphone 5	2
	2014	35	iphone 5s	3
	2014	35	kindle fire	3
	2014	35	lenovo thinkpad	16
	2014	35	mac mini	2
	2014	35	macbook air	10
	2014	35	macbook pro	17
	2014	35	nexus 10	2
	2014	35	nexus 5	4
	2014	35	nexus 7	2
	2014	35	nokia lumia 635	2
	2014	35	samsung galaxy note	1
	2014	35	samsung galaxy s4	6
	2014	35	windows surface	3

5. Email Engagement Analysis:

This SQL query calculates the email open and click rates as percentages. The inner subquery categorizes email actions into email_sent, email_open, and email_clicked. The outer query sums the occurrences of these categories and computes the open and click rates by dividing the counts of email_open and email_clicked by the count of email_sent, respectively, and multiplying by 100 to express these rates as percentages.

SQL Query:

```
select 100 * sum(case when email_cat = 'email_open' then 1 else 0 end)/  
sum(case when email_cat = 'email_sent' then 1 else 0 end) as email_open_rate,  
100 * sum(case when email_cat = 'email_clicked' then 1 else 0 end)/  
sum(case when email_cat = 'email_sent' then 1 else 0 end) as email_click_rate  
from ( select *,  
      case  
        when action in ('sent_weekly_digest', 'sent_reengagement_email') then 'email_sent'  
        when action in ('email_open') then 'email_open'  
        when action in ('email_clickthrough') then 'email_clicked'  
      end as email_cat  
      from emailevents)sub
```

Result:

	email_open_rate	email_click_rate
▶	33.5834	14.7899

Tech-Stack Used:

1. MySQL Workbench (Version 8.0.37):

Description: MySQL Workbench is a unified visual tool used for database design, development, and administration. It provides data modeling, SQL development, and comprehensive administration tools for server configuration, user administration, and more.

Reasons for Choosing MySQL Workbench:

- **User-Friendly Interface:** MySQL Workbench offers a graphical user interface that simplifies database management tasks, making it easier to visualize database structures and execute SQL queries.
- **Advanced SQL Editor:** The SQL editor supports syntax highlighting, code completion, and error parsing, which helps in writing and debugging complex SQL queries.
- **Compatibility:** MySQL Workbench is fully compatible with MySQL databases, ensuring seamless integration and efficient performance.
- **Comprehensive Features:** It includes data modeling, SQL development, and server administration tools in a single integrated environment, which enhances productivity and efficiency.

Insights:

Case Study 1: Job Data Analysis

1. Jobs Reviewed Over Time

- **Insight:** The number of jobs reviewed fluctuated throughout each day in November 2020, with noticeable peaks during typical working hours.
- **Observation:** Job reviews were higher during weekdays compared to weekends, suggesting a standard work pattern.

2. Throughput Analysis

- **Insight:** The 7-day rolling average of throughput provides a smoother trend line, reducing the noise seen in daily metrics.
- **Preference:** The 7-day rolling average is preferred as it helps identify longer-term trends and patterns, making it easier to manage and plan operational activities.

3. Language Share Analysis

- **Insight:** The percentage share of each language remained relatively stable over the last 30 days.
- **Observation:** A dominant language (e.g., English) was consistently used more than others, indicating a possible need for targeted localization efforts if aiming to reach a broader audience.

4. Duplicate Rows Detection

- **Insight:** Several duplicate rows were found, which can lead to inaccurate analysis and reporting.
- **Action:** Implementing data cleaning procedures to remove duplicates and prevent future occurrences is crucial.

Case Study 2: Investigating Metric Spike

1. Weekly User Engagement

- **Insight:** User engagement showed weekly cycles, with higher activity during weekdays and a dip during weekends.
- **Observation:** The activeness of users peaked during specific days of the week, indicating optimal times for user interaction and content delivery.

2. User Growth Analysis

- **Insight:** There was a steady increase in user growth over time, with certain periods showing significant spikes due to likely marketing campaigns or product updates.
- **Observation:** Identifying these growth periods can help correlate marketing efforts or product changes with user acquisition success.

3. Weekly Retention Analysis

- **Insight:** Retention rates varied by cohort, with some weeks showing stronger retention than others.
- **Observation:** Factors such as onboarding experience, product features, or external factors could influence retention, suggesting areas for improvement.

4. Weekly Engagement Per Device

- **Insight:** Users engaged differently depending on the device used, with mobile devices showing higher engagement rates.

- **Observation:** Optimizing the user experience for the most frequently used devices can enhance overall engagement.

5. **Email Engagement Analysis**

- **Insight:** Email engagement metrics showed varying levels of interaction, with certain types of emails performing better than others.
- **Observation:** Analyzing email content, timing, and audience segmentation can improve engagement rates and effectiveness of email campaigns.