# Operation Analytics and Investigating Metric Spike

**By Sathish**

**Data Analyst**

**Date: 06/01/2024**

**Project Description:** The objective of this project is to analyse job data to extract meaningful insights and identify areas of improvement. Investigating and explaining sudden changes in key metrics to better understand the company's end-to-end operations.

**Approach:** The project was executed in SQL (Structured Query Language) by creating a database from the given dataset. Necessary data wrangling techniques were undertaken in order to understand and extract meaningful insights from the data.

**Tech-Stack Used:** All the SQL queries were executed in My SQL Workbench version 8.0.35.0. I prefer My SQL workbench due to its ease to use interface and powerful SQL editor.

**Insights:**

**1) Case Study 1: Job Data Analysis:**

**A. <u>Jobs Reviewed Over Time:</u>**

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

**Conclusions:** Following is the list of jobs reviewed per hour for each day in November 2020. 28/11/2020 saw the highest number of job reviews per hour and 27/11/2020 saw the lowest reviews.

| ds | num_of_job_reviewed_per_day_per_hour |
|---|---|
| 25-11-2020 | 80 |
| 26-11-2020 | 64.2857 |
| 27-11-2020 | 34.6154 |
| 28-11-2020 | 218.1818 |
| 29-11-2020 | 180 |
| 30-11-2020 | 180 |

**MySQL Query:**

```
SELECT ds, (COUNT(job_id)*3600/SUM(time_spent)) AS
        num_of_job_reviewed_per_day_per_hour
        FROM job_data_new
        GROUP BY ds
        ORDER BY ds;
```

## B. Throughput Analysis:

**Task:** To calculate the 7-day rolling average of throughput (number of events per second).

**Conclusions:** Following table shows the 7-day rolling average of throughput.

| ds | throughput_7d | throughtput_daily |
|---|---|---|
| 25-11-2020 | 0.02 | 0.02 |
| 26-11-2020 | 0.02 | 0.02 |
| 27-11-2020 | 0.01 | 0.01 |
| 28-11-2020 | 0.02 | 0.02 |
| 29-11-2020 | 0.02 | 0.02 |
| 30-11-2020 | 0.03 | 0.03 |

**MySQL Query: SELECT** ds,
**ROUND**(**SUM**(**COUNT**(job_id)) **OVER** (**ORDER BY** ds rows between 6 preceding and current
row) / **SUM**(**SUM**(time_spent)) **OVER** (**ORDER BY** ds rows between 6 preceding and current row),2) **AS**
throughput_7d,
**ROUND**(**SUM**(**COUNT**(job_id)) **OVER** (**ORDER BY** ds) / **SUM**(**SUM**(time_spent)) **OVER** (**ORDER BY** ds),2) **AS**
throughtput_daily
**FROM** job_data_new
**GROUP BY** ds
**ORDER BY** ds;

## C. Language Share Analysis:

**Task:** Calculate the percentage share of each language in the last 30 days.

**Conclusions:** Following table shows the percentage share of each language in the last 30 days.

Persian language has the highest usage during the last 30 days.

| language | count | Percent_language_share |
|---|---|---|
| Arabic | 1 | 12.5 |
| English | 1 | 12.5 |
| French | 1 | 12.5 |
| Hindi | 1 | 12.5 |
| Italian | 1 | 12.5 |
| Persian | 3 | 37.5 |

**MySQL Query: SELECT** language, **COUNT**(language) **AS** count,

**COUNT**(language)/(**SELECT COUNT**(*) **FROM** job_data_new)*100 **AS** Percent_language_share
**FROM** job_data_new
**GROUP BY** language;

### D. Duplicate Rows Detection:

**Task:** To identify duplicate rows in the data.

**Conclusions:** There are no duplicate rows in the provided dataset.

| ds | job_id | actor_id | event | language | org | time_spent |
|---|---|---|---|---|---|---|
| | | | | | | |

**MySQL Query: SELECT** * **FROM** job_data_new
**GROUP** by ds, job_id, actor_id, event, language, time_spent, org
**HAVING COUNT**(*)>1;

## 2) Case Study 2: Investigating Metric Spike:

### A. Weekly User Engagement:

**Task:** To measure the activeness of users on a weekly basis.

**Conclusions:** Following table shows the activeness of users on a weekly basis. 30$^{th}$ week of year 2014 saw the highest number of active users and 17$^{th}$ week of year 2014 saw the lowest number of active users.

| year | week | Count |
|---|---|---|
| 2014 | 17 | 663 |
| 2014 | 18 | 1068 |
| 2014 | 19 | 1113 |
| 2014 | 20 | 1154 |
| 2014 | 21 | 1121 |
| 2014 | 22 | 1186 |
| 2014 | 23 | 1232 |
| 2014 | 24 | 1275 |
| 2014 | 25 | 1264 |
| 2014 | 26 | 1302 |
| 2014 | 27 | 1372 |
| 2014 | 28 | 1365 |
| 2014 | 29 | 1376 |
| 2014 | 30 | 1467 |
| 2014 | 31 | 1299 |
| 2014 | 32 | 1225 |
| 2014 | 33 | 1225 |
| 2014 | 34 | 1204 |
| 2014 | 35 | 104 |

**MySQL Query: SELECT** extract(year **FROM** occurred_at) **AS** year,
**EXTRACT**(week from occurred_at) **AS** week, **COUNT**(distinct user_id) **AS** Count
**FROM** events_new
**WHERE** event_type = 'engagement'
**GROUP BY** week, year;

## B. User Growth Analysis:

**Task:** To analyze the growth of users over time for a product.

**Conclusions:** Following table shows the growth of users over time for a product. July 2014 saw the highest number of growth with 124 new users.

| month_year | count_of_users | growth |
|---|---|---|
| Jan-13 | 160 | |
| Feb-13 | 160 | 0 |
| Mar-13 | 150 | -10 |
| Apr-13 | 181 | 31 |
| May-13 | 214 | 33 |
| Jun-13 | 213 | -1 |
| Jul-13 | 284 | 71 |
| Aug-13 | 316 | 32 |
| Sep-13 | 330 | 14 |
| Oct-13 | 390 | 60 |
| Nov-13 | 399 | 9 |
| Dec-13 | 486 | 87 |
| Jan-14 | 552 | 66 |
| Feb-14 | 525 | -27 |
| Mar-14 | 615 | 90 |
| Apr-14 | 726 | 111 |
| May-14 | 779 | 53 |
| Jun-14 | 873 | 94 |
| Jul-14 | 997 | 124 |
| Aug-14 | 1031 | 34 |

**MySQL Query: SELECT CONCAT**(extract(month from created_at),'-',**EXTRACT**(year from created_at)) **AS** month_year,
**COUNT**(user_id) **AS** count_of_users,
**COUNT**(user_id) - **LAG**(**COUNT**(user_id),1) **OVER** () **AS** growth
**FROM** users_new
**GROUP BY** month_year;

## C. Weekly Retention Analysis:

**Task:** To analyze the retention of users on a weekly basis after signing up for a product.

**Conclusions:** Following table shows the retention of users on a weekly basis after signing up for a product.

| sign_up_week | count_of_users |
|---|---|
| 17 | 72 |
| 18 | 163 |
| 19 | 185 |
| 20 | 176 |
| 21 | 183 |
| 22 | 196 |
| 23 | 196 |
| 24 | 229 |
| 25 | 207 |
| 26 | 201 |
| 27 | 222 |
| 28 | 215 |
| 29 | 221 |
| 30 | 238 |
| 31 | 193 |
| 32 | 245 |
| 33 | 261 |
| 34 | 259 |
| 35 | 18 |

**MySQL Query: SELECT EXTRACT**(week from occurred_at) **AS** sign_up_week,**COUNT**(distinct user_id) **AS** count_of_users
**FROM** events_new
**WHERE** event_type = 'signup_flow'
**AND** event_name = 'complete_signup'
**GROUP BY** sign_up_week;

## D. Weekly Engagement Per Device:

**Task:** To Measure the activeness of users on a weekly basis per device.

**Conclusions:** Following table shows the activeness of users on a weekly basis per device.

(**Note:** The result of the query has been broke down to several lines to make it more readable)

| year | week_number | acer aspire desktop | acer aspire notebook | amazon fire phone | asus chromebook | dell inspiron desktop | dell inspiron notebook |
|---|---|---|---|---|---|---|---|
| 2014 | 17 | 67 | 206 | 83 | 251 | 187 | 503 |
| 2014 | 18 | 295 | 363 | 177 | 523 | 683 | 953 |
| 2014 | 19 | 242 | 406 | 141 | 268 | 444 | 1193 |
| 2014 | 20 | 226 | 483 | 104 | 461 | 504 | 975 |
| 2014 | 21 | 328 | 462 | 29 | 550 | 568 | 955 |
| 2014 | 22 | 255 | 431 | 46 | 631 | 737 | 1011 |
| 2014 | 23 | 240 | 449 | 181 | 671 | 494 | 1034 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 2014 | 24 | 289 | 509 | 147 | 471 | 713 | 1082 |
| 2014 | 25 | 263 | 605 | 131 | 432 | 644 | 1209 |
| 2014 | 26 | 313 | 328 | 137 | 632 | 672 | 1077 |
| 2014 | 27 | 296 | 582 | 109 | 510 | 541 | 992 |
| 2014 | 28 | 303 | 541 | 51 | 511 | 744 | 1183 |
| 2014 | 29 | 212 | 592 | 92 | 529 | 560 | 1179 |
| 2014 | 30 | 403 | 646 | 170 | 565 | 685 | 1488 |
| 2014 | 31 | 383 | 554 | 157 | 636 | 465 | 1242 |
| 2014 | 32 | 354 | 580 | 161 | 611 | 519 | 1229 |
| 2014 | 33 | 363 | 498 | 114 | 581 | 408 | 1011 |
| 2014 | 34 | 281 | 561 | 106 | 557 | 465 | 1088 |
| 2014 | 35 | 7 | 28 | 0 | 38 | 4 | 66 |

| year | week_number | hp pavilion desktop | htc one | ipad air | ipad mini | iphone 4s | iphone 5 |
|---|---|---|---|---|---|---|---|
| 2014 | 17 | 0 | 190 | 330 | 205 | 217 | 706 |
| 2014 | 18 | 0 | 174 | 520 | 309 | 448 | 1328 |
| 2014 | 19 | 0 | 270 | 595 | 381 | 546 | 1190 |
| 2014 | 20 | 0 | 365 | 611 | 264 | 608 | 1302 |
| 2014 | 21 | 0 | 260 | 428 | 280 | 522 | 1567 |
| 2014 | 22 | 0 | 323 | 591 | 342 | 477 | 1394 |
| 2014 | 23 | 0 | 234 | 404 | 258 | 459 | 1748 |
| 2014 | 24 | 0 | 136 | 619 | 313 | 609 | 1475 |
| 2014 | 25 | 0 | 285 | 640 | 234 | 442 | 1645 |
| 2014 | 26 | 0 | 225 | 580 | 464 | 503 | 1602 |
| 2014 | 27 | 0 | 225 | 482 | 344 | 720 | 1867 |
| 2014 | 28 | 0 | 318 | 547 | 364 | 783 | 1654 |
| 2014 | 29 | 0 | 227 | 620 | 364 | 605 | 1593 |
| 2014 | 30 | 0 | 306 | 656 | 373 | 740 | 1535 |
| 2014 | 31 | 0 | 117 | 573 | 240 | 631 | 1491 |
| 2014 | 32 | 0 | 146 | 463 | 257 | 318 | 1301 |
| 2014 | 33 | 0 | 138 | 376 | 260 | 349 | 1203 |
| 2014 | 34 | 0 | 271 | 306 | 252 | 472 | 966 |
| 2014 | 35 | 0 | 18 | 0 | 21 | 57 | 9 |

| year | week_number | iphone 5s | kindle fire | lenovo thinkpad | mac mini | macbook air | macbook pro |
|---|---|---|---|---|---|---|---|
| 2014 | 17 | 473 | 57 | 793 | 59 | 490 | 1516 |
| 2014 | 18 | 778 | 265 | 1732 | 159 | 1604 | 3301 |
| 2014 | 19 | 964 | 225 | 2143 | 255 | 1331 | 3159 |
| 2014 | 20 | 1024 | 242 | 2203 | 272 | 1443 | 3097 |
| 2014 | 21 | 804 | 355 | 1893 | 281 | 1269 | 3044 |
| 2014 | 22 | 920 | 231 | 1803 | 276 | 1714 | 3046 |
| 2014 | 23 | 859 | 352 | 1882 | 170 | 1431 | 3123 |
| 2014 | 24 | 959 | 239 | 1806 | 298 | 1716 | 3028 |
| 2014 | 25 | 963 | 209 | 2096 | 230 | 1269 | 2932 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 2014 | 26 | 1026 | 254 | 2214 | 140 | 1489 | 3309 |
| 2014 | 27 | 952 | 268 | 2233 | 169 | 1652 | 3548 |
| 2014 | 28 | 1039 | 342 | 2564 | 340 | 1671 | 3461 |
| 2014 | 29 | 932 | 293 | 2438 | 342 | 1691 | 3155 |
| 2014 | 30 | 1164 | 234 | 2584 | 316 | 1731 | 3578 |
| 2014 | 31 | 663 | 136 | 2114 | 232 | 1574 | 3608 |
| 2014 | 32 | 775 | 67 | 1898 | 120 | 1265 | 3320 |
| 2014 | 33 | 712 | 120 | 2156 | 331 | 1622 | 3182 |
| 2014 | 34 | 715 | 119 | 1908 | 395 | 1456 | 3141 |
| 2014 | 35 | 22 | 32 | 123 | 25 | 64 | 122 |

| year | week_number | nexus 10 | nexus 5 | nexus 7 | nokia lumia 635 | samsumg galaxy tablet | samsung galaxy note |
|---|---|---|---|---|---|---|---|
| 2014 | 17 | 145 | 382 | 177 | 128 | 70 | 116 |
| 2014 | 18 | 370 | 938 | 252 | 341 | 79 | 139 |
| 2014 | 19 | 232 | 944 | 334 | 215 | 66 | 117 |
| 2014 | 20 | 217 | 1278 | 372 | 151 | 78 | 160 |
| 2014 | 21 | 258 | 1020 | 211 | 190 | 63 | 196 |
| 2014 | 22 | 311 | 1231 | 433 | 308 | 70 | 252 |
| 2014 | 23 | 533 | 997 | 303 | 273 | 98 | 183 |
| 2014 | 24 | 369 | 1151 | 418 | 489 | 101 | 245 |
| 2014 | 25 | 258 | 935 | 557 | 446 | 159 | 134 |
| 2014 | 26 | 261 | 912 | 428 | 456 | 139 | 97 |
| 2014 | 27 | 320 | 1015 | 366 | 325 | 146 | 131 |
| 2014 | 28 | 261 | 946 | 342 | 405 | 102 | 118 |
| 2014 | 29 | 272 | 793 | 392 | 461 | 120 | 141 |
| 2014 | 30 | 282 | 983 | 629 | 333 | 137 | 156 |
| 2014 | 31 | 274 | 684 | 407 | 298 | 80 | 112 |
| 2014 | 32 | 318 | 678 | 235 | 357 | 77 | 116 |
| 2014 | 33 | 163 | 641 | 255 | 215 | 82 | 127 |
| 2014 | 34 | 208 | 770 | 328 | 151 | 119 | 96 |
| 2014 | 35 | 15 | 34 | 17 | 7 | 0 | 6 |

| year | week_number | samsung galaxy s4 | windows surface |
|---|---|---|---|
| 2014 | 17 | 449 | 87 |
| 2014 | 18 | 1130 | 107 |
| 2014 | 19 | 1024 | 163 |
| 2014 | 20 | 1001 | 194 |
| 2014 | 21 | 960 | 183 |
| 2014 | 22 | 1020 | 189 |
| 2014 | 23 | 1094 | 155 |
| 2014 | 24 | 940 | 209 |
| 2014 | 25 | 1102 | 234 |
| 2014 | 26 | 1114 | 177 |

| | | | |
|---|---|---|---|
| 2014 | 27 | 1154 | 345 |
| 2014 | 28 | 1260 | 304 |
| 2014 | 29 | 1462 | 260 |
| 2014 | 30 | 1203 | 170 |
| 2014 | 31 | 1068 | 216 |
| 2014 | 32 | 812 | 86 |
| 2014 | 33 | 729 | 130 |
| 2014 | 34 | 885 | 160 |
| 2014 | 35 | 29 | 30 |

**MySQL Query:**

```
SELECT EXTRACT(year from occurred_at) AS year, EXTRACT(week from occurred_at) AS week_number,
COUNT(CASE WHEN device = 'acer aspire desktop' THEN user_id ELSE null END) AS 'acer aspire desktop',
COUNT(CASE WHEN device = 'acer aspire notebook' THEN user_id ELSE null END) AS 'acer aspire notebook',
COUNT(CASE WHEN device = 'amazon fire phone' THEN user_id ELSE null END) AS 'amazon fire phone',
COUNT(CASE WHEN device = 'asus chromebook' THEN user_id ELSE null END) AS 'asus chromebook',
COUNT(CASE WHEN device = 'dell inspiron desktop' THEN user_id ELSE null END) AS 'dell inspiron desktop',
COUNT(CASE WHEN device = 'dell inspiron notebook' THEN user_id ELSE null END) AS 'dell inspiron notebook',
COUNT(CASE WHEN device = 'hp pavilion desktopp' THEN user_id ELSE null END) AS 'hp pavilion desktop',
COUNT(CASE WHEN device = 'htc one' THEN user_id ELSE null END) AS 'htc one',
COUNT(CASE WHEN device = 'ipad air' THEN user_id ELSE null END) AS 'ipad air',
COUNT(CASE WHEN device = 'ipad mini' THEN user_id ELSE null END) AS 'ipad mini',
COUNT(CASE WHEN device = 'iphone 4s' THEN user_id ELSE null END) AS 'iphone 4s',
COUNT(CASE WHEN device = 'iphone 5' THEN user_id ELSE null END) AS 'iphone 5',
COUNT(CASE WHEN device = 'iphone 5s' THEN user_id ELSE null END) AS 'iphone 5s',
COUNT(CASE WHEN device = 'kindle fire' THEN user_id ELSE null END) AS 'kindle fire',
COUNT(CASE WHEN device = 'lenovo thinkpad' THEN user_id ELSE null END) AS 'lenovo thinkpad',
COUNT(CASE WHEN device = 'mac mini' THEN user_id ELSE null END) AS 'mac mini',
COUNT(CASE WHEN device = 'macbook air' THEN user_id ELSE null END) AS 'macbook air',
COUNT(CASE WHEN device = 'macbook pro' THEN user_id ELSE null END) AS 'macbook pro',
COUNT(CASE WHEN device = 'nexus 10' THEN user_id ELSE null END) AS 'nexus 10',
COUNT(CASE WHEN device = 'nexus 5' THEN user_id ELSE null END) AS 'nexus 5',
COUNT(CASE WHEN device = 'nexus 7' THEN user_id ELSE null END) AS 'nexus 7',
```

```
COUNT(CASE WHEN device = 'nokia lumia 635' THEN user_id ELSE null END) AS 'nokia lumia 635',
COUNT(CASE WHEN device = 'samsumg galaxy tablet' THEN user_id ELSE null END) AS 'samsumg galaxy tablet',
COUNT(CASE WHEN device = 'samsung galaxy note' THEN user_id ELSE null END) AS 'samsung galaxy note',
COUNT(CASE WHEN device = 'samsung galaxy s4' THEN user_id ELSE null END) AS 'samsung galaxy s4',
COUNT(CASE WHEN device = 'windows surface' THEN user_id ELSE null END) AS 'windows surface'
FROM events_new
WHERE event_type = 'engagement'
GROUP BY year, week_number
ORDER BY year, week_number;
```

## E. Email Engagement Analysis:

**Task:** To analyze how users are engaging with the email service.

**Conclusions:** Following table shows how users are engaging with the email service.

| year | week_number | sent_weekly_digest | email_open | email_clickthrough | sent_reengagement_email |
|---|---|---|---|---|---|
| 2014 | 18 | 2602 | 912 | 430 | 157 |
| 2014 | 19 | 2665 | 972 | 477 | 173 |
| 2014 | 20 | 2733 | 1004 | 507 | 191 |
| 2014 | 21 | 2822 | 1014 | 443 | 164 |
| 2014 | 22 | 2911 | 987 | 488 | 192 |
| 2014 | 23 | 3003 | 1075 | 538 | 197 |
| 2014 | 24 | 3105 | 1155 | 554 | 226 |
| 2014 | 25 | 3207 | 1096 | 530 | 196 |
| 2014 | 26 | 3302 | 1165 | 556 | 219 |
| 2014 | 27 | 3399 | 1228 | 621 | 213 |
| 2014 | 28 | 3499 | 1250 | 599 | 213 |
| 2014 | 29 | 3592 | 1219 | 590 | 213 |
| 2014 | 30 | 3706 | 1383 | 630 | 231 |
| 2014 | 31 | 3793 | 1351 | 445 | 222 |
| 2014 | 32 | 3897 | 1337 | 418 | 200 |
| 2014 | 33 | 4012 | 1432 | 490 | 264 |
| 2014 | 34 | 4111 | 1528 | 490 | 261 |
| 2014 | 17 | 908 | 310 | 166 | 73 |
| 2014 | 35 | 0 | 41 | 38 | 48 |

**MySQL Query:**

```
SELECT EXTRACT(year from occurred_at) AS year, EXTRACT(week from occurred_at)
AS week_number,
COUNT(case WHEN action = 'sent_weekly_digest' THEN user_id ELSE null END) AS
'sent_weekly_digest',
COUNT(CASE WHEN action = 'email_open' THEN user_id ELSE null END) AS
'email_open',
COUNT(CASE WHEN action = 'email_clickthrough' THEN user_id ELSE null END) AS
'email_clickthrough',
COUNT(CASE WHEN action = 'sent_reengagement_email' THEN user_id ELSE null END)
AS 'sent_reengagement_email'
FROM email_events_new
GROUP BY year, week_number;
```

**Result:**

Insights derived from the above analysis helps the management team to better understand company's end-to-end operations and showcase the sudden changes in the key metrics. Also, it identify the areas of improvement and guides the management team in taking necessary actions.

# Thank You