## Project Overview

**Project Description:** As the Lead Data Analyst, my role is to utilize advanced SQL skills for analyzing operations data. I will investigate metric spikes and sudden changes in key metrics, providing valuable insights to various departments. Through data-driven decision-making, I will optimize operational efficiency and drive continuous improvement, helping our company achieve its business goals.

### Approach

My approach towards the project involved a systematic and data-driven methodology to execute the analysis effectively. I started by thoroughly understanding the objectives of the project, which involved operational analytics and investigating metric spikes. I then proceeded with the following steps:

1. **Data Understanding:** Reviewed provided datasets and tables to understand data structure and content.
2. **Query Formulation:** Utilized advanced SQL skills to create specific queries for data extraction.
3. **Data Analysis:** Conducted in-depth analysis to identify patterns and sudden changes in key metrics.
4. **Insight Generation:** Derived valuable insights and root causes behind metric spikes and fluctuations.
5. **Communication:** Presented findings through clear reports to respective departments.
6. **Continuous Improvement:** Emphasized data-driven decision-making and encouraged a culture of ongoing improvement.

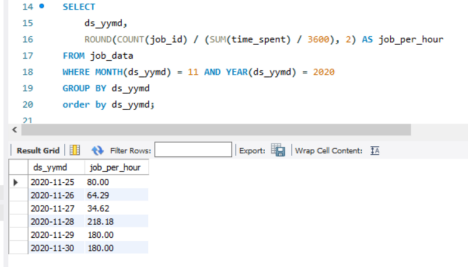
### Tech-Stack Used

* **Microsoft Excel:** Employed for data exploration, simple data cleaning, and basic data visualization.
* **MySQL 8:** Used for storing and managing the datasets, providing efficient data storage and retrieval.

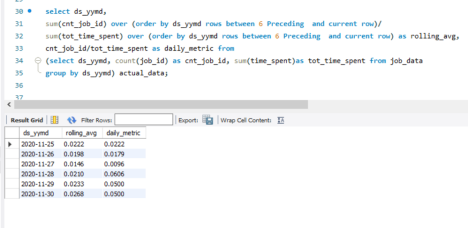
By utilizing MySQL 8 and Microsoft Excel, the project achieved efficient data analysis and visualization capabilities, empowering the Data Analyst to derive valuable insights and make informed decisions to improve operational efficiency.

### Insights

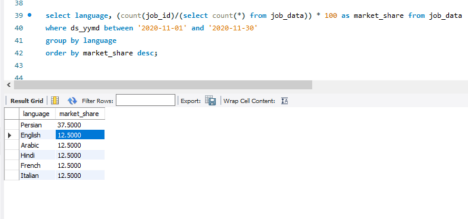
1. **Jobs Reviewed Over Time:** The following query returns data on how many jobs were reviewed per hour.



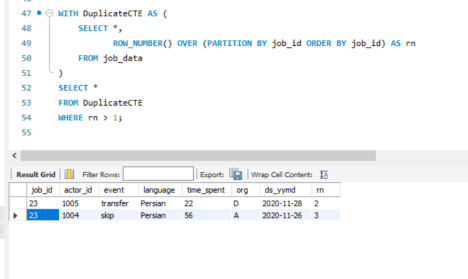
1. **Throughput Analysis:** Both throughput and daily metrics provide valuable insights, focusing on different aspects of the data.



1. **Language Share Analysis:** In the last 30 days, 37.5% of jobs reviewed were in Persian language.

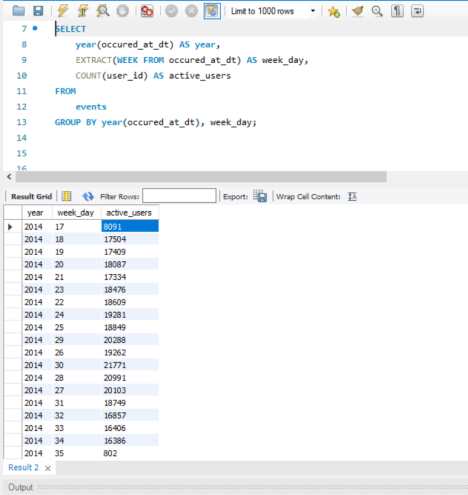


1. **Duplicate Rows Detection:** If job ID is unique, it should not be assigned to a new job.

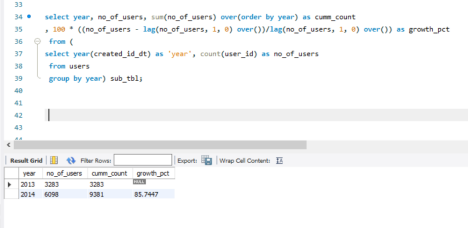


### User Engagement Analysis

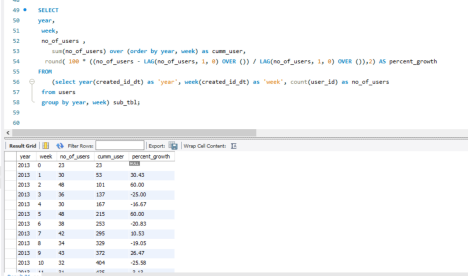
1. **Weekly User Engagement:** This result shows weekly user engagement on our platform.



1. **User Growth Analysis:**
   * Our first query shows year growth in the customer base, increasing by 85.8%.



* + The second query shows percent growth on a weekly basis.



1. **Weekly Retention Analysis:** 4-week retention rate of users on a weekly basis.

Code:

WITH cohort\_week AS (

SELECT

user\_id,

WEEK(created\_id\_dt) AS joining\_week,

YEAR(created\_id\_dt) AS joining\_year,

created\_id\_dt

FROM users

ORDER BY 1

),

user\_activities AS (

SELECT

e.user\_id AS id,

occured\_at\_dt,

created\_id\_dt,

DATEDIFF(occured\_at\_dt, created\_id\_dt),

CEIL(DATEDIFF(occured\_at\_dt, created\_id\_dt) / 7) AS active\_diff\_week, joining\_week,

joining\_year

FROM cohort\_week c

LEFT OUTER JOIN events e ON e.user\_id = c.user\_id

WHERE event\_type <> 'signup\_flow'

GROUP BY id, active\_diff\_week

),

cohort\_size AS (

SELECT

joining\_year,

joining\_week,

COUNT(user\_id) AS num\_users

FROM cohort\_week

GROUP BY joining\_year, joining\_week

ORDER BY joining\_year, joining\_week

),

retention\_table AS (

SELECT

C.joining\_year,

C.joining\_week,

A.active\_diff\_week,

COUNT(id) AS num\_users

FROM cohort\_week C

LEFT JOIN user\_activities A ON A.id = C.user\_id

GROUP BY joining\_year, joining\_week, active\_diff\_week

)

SELECT

joining\_year,

joining\_week, MAX(total\_users) AS total\_users,

round(coalesce(MAX(CASE WHEN active\_diff\_week = 0 THEN percentage END),0), 2) AS week\_0,

round(coalesce(MAX(CASE WHEN active\_diff\_week = 1 THEN percentage END), 0), 2) AS week\_1,

round(coalesce(MAX(CASE WHEN active\_diff\_week = 2 THEN percentage END), 0), 2) AS week\_2,

round(coalesce(MAX(CASE WHEN active\_diff\_week = 3 THEN percentage END), 0), 2) AS week\_3

FROM (

SELECT

R.joining\_year,

R.joining\_week,

S.num\_users AS total\_users,

R.active\_diff\_week,

R.num\_users \* 100 / S.num\_users AS percentage

FROM cohort\_size S

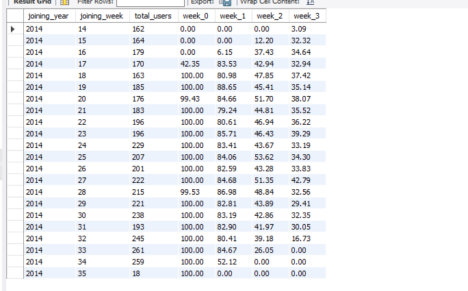
LEFT JOIN retention\_table R ON R.joining\_year = S.joining\_year AND R.joining\_week = S.joining\_week

WHERE R.active\_diff\_week < 4

) AS pivot\_tbl

GROUP BY joining\_year, joining\_week

ORDER BY joining\_year, joining\_week;



1. **Weekly Engagement Per Device:** Analysis of user engagement based on device type.

**Code:**

SELECT

weekly,

SUM(CASE WHEN device\_type = 'Tablet' THEN device\_count ELSE 0 END) AS Tablet, SUM(CASE WHEN device\_type = 'Notebook' THEN device\_count ELSE 0 END) AS Notebook,

SUM(CASE WHEN device\_type = 'Mobile' THEN device\_count ELSE 0 END) AS Mobile, SUM(CASE WHEN device\_type = 'Other' THEN device\_count ELSE 0 END) AS Other, sum(device\_count) as Total

FROM(

select week(occured\_at\_dt) as weekly, CASE

WHEN device IN ('ipad mini', 'nexus 7', 'samsung galaxy tablet') THEN 'Tablet' WHEN device IN ('dell inspiron notebook', 'macbook air', 'macbook pro', 'acer aspire notebook', 'asus chromebook', 'mac mini', 'hp pavilion desktop', 'acer aspire desktop') THEN 'Notebook'

WHEN device IN ('iphone 5', 'iphone 4s', 'iphone 5s', 'nexus 5', 'samsung galaxy s4', 'htc one', 'amazon fire phone', 'nokia lumia 635') THEN 'Mobile'

WHEN device IN ('windows surface', 'kindle fire', 'nexus 10', 'samsung galaxy note') THEN 'Tablet'

ELSE 'Other'

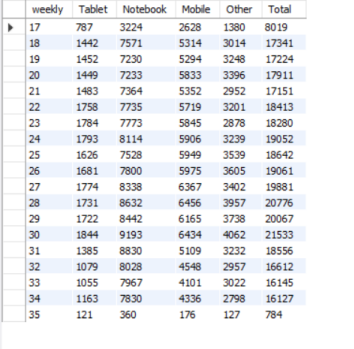
END AS device\_type , count(user\_id) as device\_count from events

where event\_name <> 'complete\_signup'

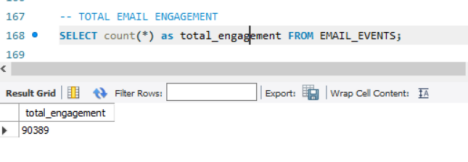
group by weekly, device\_type) pivot\_tbl

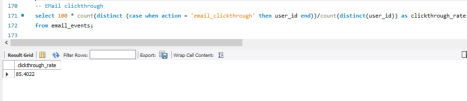
group by weekly

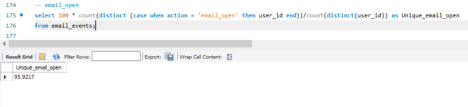
order by weekly;



1. **Email Engagement:** Analysis of email engagement metrics.







Result: With this analysis, we identified various pros and cons for our company, understanding our performance in reviewing jobs daily and ensuring growth year by year.