

Operational Analytics & Metric Spike Investigation

A dual case study using SQL to investigate operational efficiency and user behavior patterns.



Project Description

This project explores two real-world business scenarios using SQL:

CASE STUDY 1:

Job Data Analysis – Focused on operational performance and throughput.

CASE STUDY 2:

Metric Spike Investigation – Aimed at uncovering user behavior and engagement patterns.



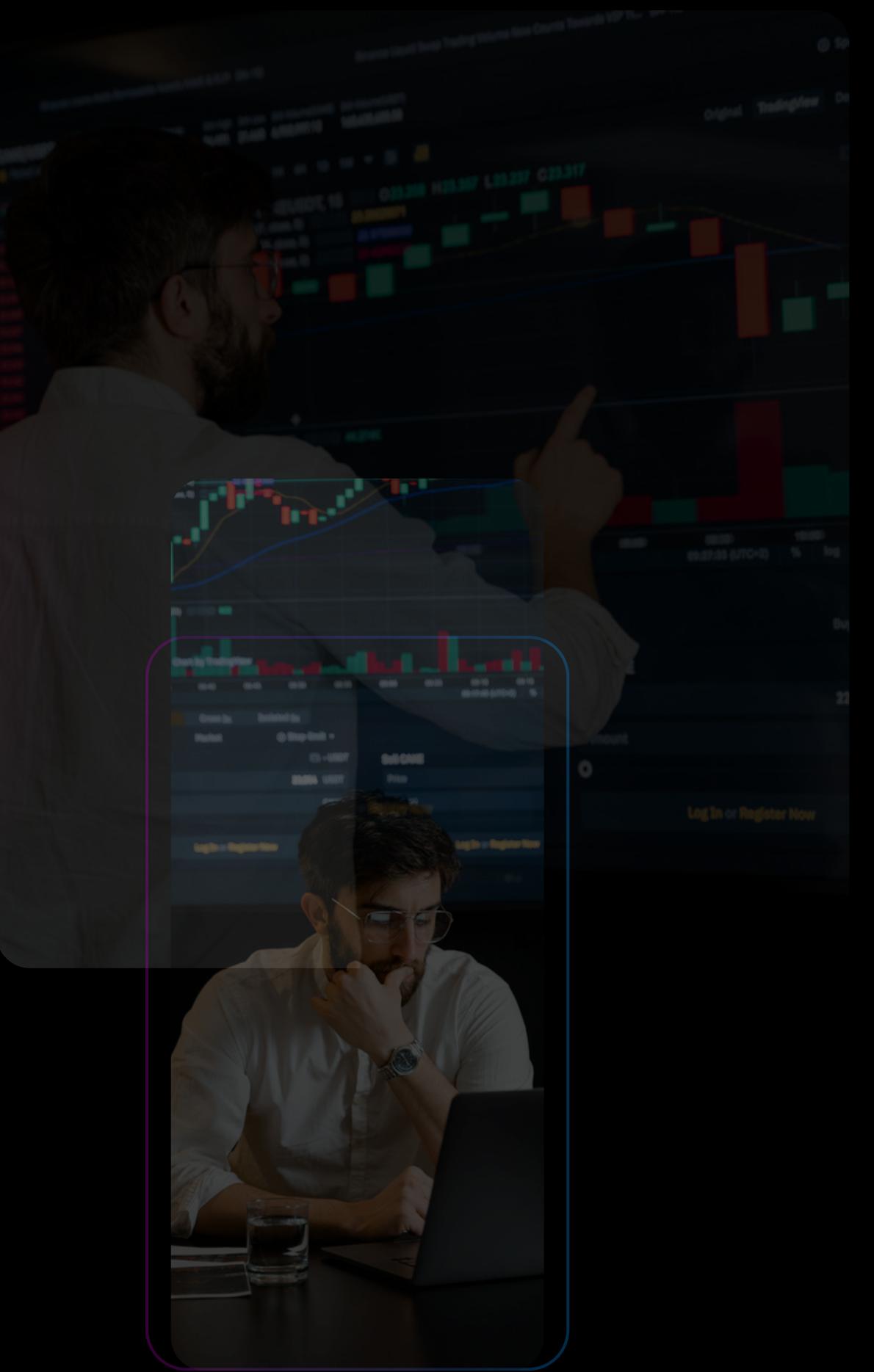
Tech Stack Used

- 01 MySQL Workbench – For database creation, data import, and SQL queries
- 02 Excel – For visualizing query outputs into charts
- 03 Canva – For building a clean, insight-driven presentation
- 04 CSV Datasets – Imported into MySQL
- 05 Google Drive: Report hosting and submission



Our Analytical Approach

- Understand the Business Problem – Analyzed jobs, user metrics, and engagement spikes
- Data Exploration – Studied structure and relationships among tables
- SQL Querying – Used CTEs, window functions, and aggregations
- Insights Extraction – Focused on key KPIs like throughput, retention, engagement
- Visualization – Built charts in Excel and designed slides in Canva



CASE STUDY 1:

JOB DATA
ANALYSIS

Jobs Reviewed Over Time

Insight:

Job reviewers were most active on 11/28/2020.

Conclusion:

Higher throughput correlates with consistent review hours and reviewer availability.

```
SELECT ds AS Date,  
       COUNT(job_id) AS Joint_Job_Id,  
       ROUND((SUM(time_spent) / 3600), 2) AS Total_Time_Sp_Hr,  
       ROUND((COUNT(job_id) / (SUM(time_spent) / 3600)), 2) AS Job_Rview_PHr_PDay  
  FROM job_data  
 WHERE STR_TO_DATE(ds, '%m/%d/%Y') BETWEEN '2020-11-01' AND '2020-11-30'  
  GROUP BY ds  
 ORDER BY STR_TO_DATE(ds, '%m/%d/%Y');
```



Throughput Analysis (7-Day Rolling Avg)

```
#B Weekly Average Throughput (Overall)
SELECT ROUND(COUNT(event) / SUM(time_spent), 2) AS weekly_avg_throughput
FROM job_data;

-- Daily Average Throughput (Per Day)
SELECT ds AS Dates,
       ROUND(COUNT(event) / SUM(time_spent), 2) AS daily_avg_throughput
FROM job_data
GROUP BY ds
ORDER BY STR_TO_DATE(ds, '%m/%d/%Y');
```

Insight:

The 7-day rolling average provides smoother insight into reviewer performance, filtering out daily volatility.

Conclusion:

Rolling average is preferred for operational planning and identifying long-term trends.



Language Share Analysis

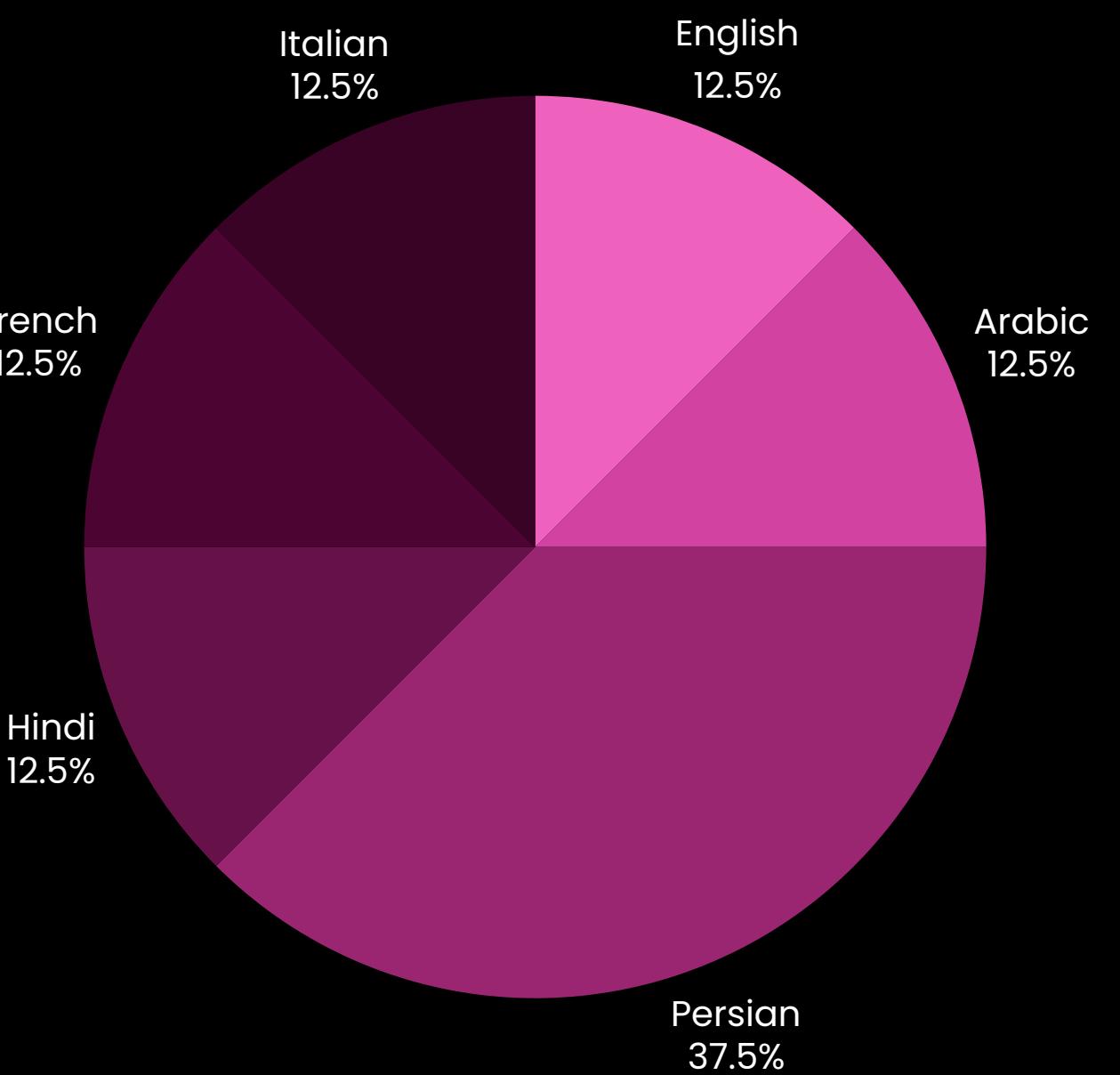
Insight:

Top 3 languages accounted for over 80% of jobs in the last 30 days.

Conclusion:

Supports optimizing reviewer allocation by language demand.

```
#C Language Share Percentage
SELECT language,
       ROUND(100 * COUNT(*) / jd.total, 2) AS Percentage,
       jd.total
  FROM job_data
 CROSS JOIN (
    SELECT COUNT(*) AS total
      FROM job_data
 ) AS jd
 GROUP BY language, jd.total;
```



Duplicate Rows Detection

Insight:

- Found 1 duplicate row
- Possible causes: double entry, syncing errors, or manual mistakes

Conclusion:

- Highlight need for better data validation during input to maintain database integrity and improve analysis accuracy

```
#D Duplicate Rows Detection:  
SELECT actor_id, COUNT(*) AS Duplicate  
FROM job_data  
GROUP BY actor_id  
HAVING COUNT(*) > 1;
```

CASE STUDY 2:

INVESTIGATING
METRIC SPIKES

Weekly User Engagement

Interesting Insight

A significant spike in engagement occurred around **2014-07-27**, with **21,096** events, suggesting a possible product update, marketing campaign, or external event driving user activity.

Conclusion

User engagement shows consistent activity with notable peaks, particularly in 2014-07-27. To sustain and grow engagement, consider investigating the drivers behind these spikes (e.g., feature releases or campaigns) and replicating successful strategies.

```
# [Case Study 2: Investigating Metric Spike]
# Weekly User Engagement:
SELECT DATE_SUB(occurred_at, INTERVAL (DAYOFWEEK(occurred_at) - 1) DAY) AS week_start,
       COUNT(*) AS event_count
FROM events
GROUP BY DATE_SUB(occurred_at, INTERVAL (DAYOFWEEK(occurred_at) - 1) DAY)
ORDER BY week_start;
```

User Growth Analysis

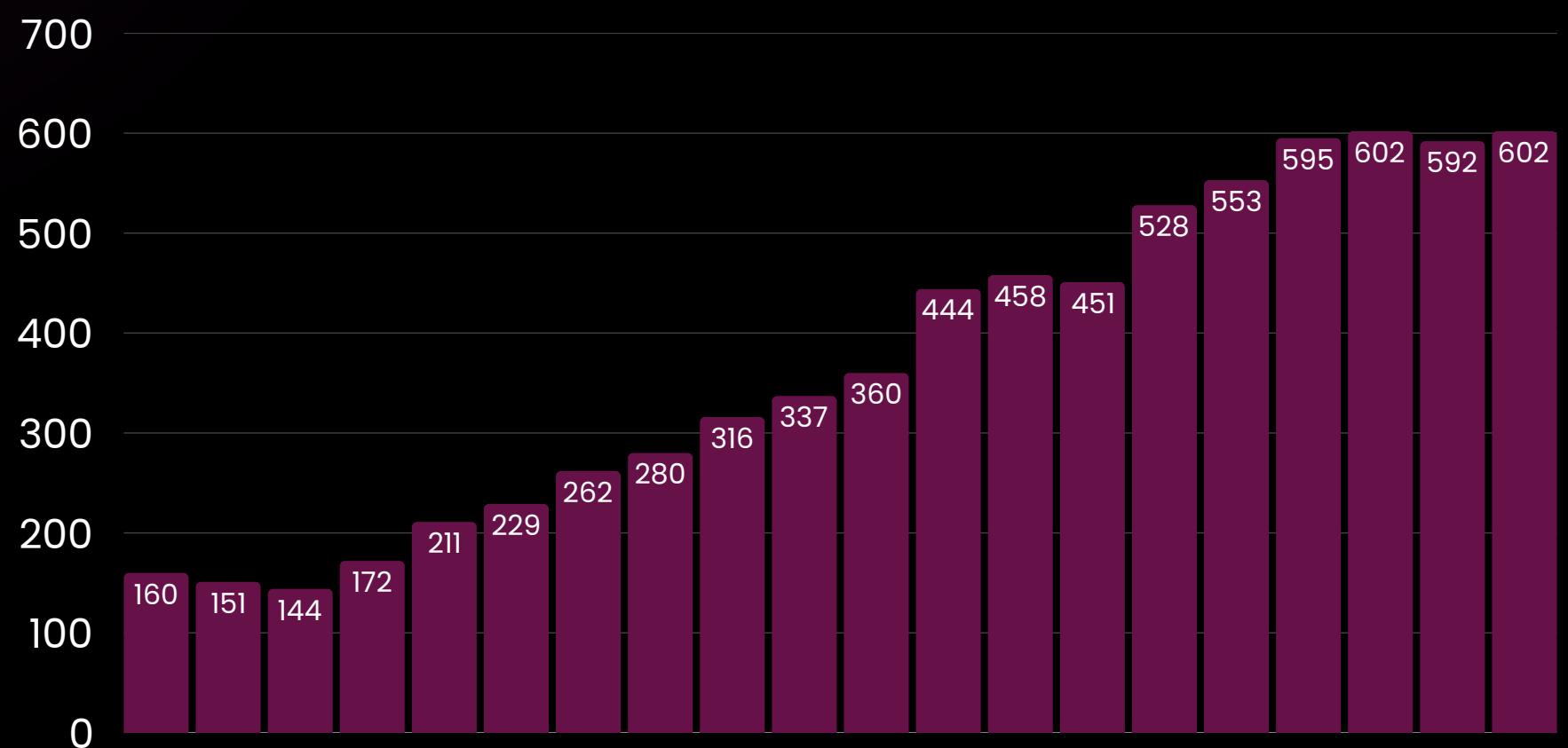
Interesting Insight

A remarkable surge occurred on 2013-06-04 with 16 new users, likely tied to a significant event or campaign. Additionally, 2014-08 saw the highest monthly growth at 1031 users, indicating a strong upward trend in user acquisition.

Conclusion

User growth shows a steady increase, with notable peaks in daily and monthly registrations. The significant growth in 2014-08 suggests effective acquisition strategies that could be analyzed and replicated to sustain momentum.

```
#User Growth Analysis
SELECT DATE(created_at) AS date,
       COUNT(DISTINCT user_id) AS new_users
  FROM users
 GROUP BY DATE(created_at)
 ORDER BY date;
```



Weekly Retention Analysis

Insight:

This report analyzes user growth and retention from January 2013 to August 2014.

Total New Users: 1.3K

Average Daily New Users: 6.36

Peak Daily Growth: 2013-06-04 (16 users)

Peak Monthly Growth: 2014-08 (1031 users)

First-Week Retention Rate: 35.63%

Conclusion:

User growth shows a steady increase, with notable peaks in daily and monthly registrations. The significant growth in 2014-08 suggests effective acquisition strategies. Retention data indicates improving user engagement, with newer cohorts showing better retention rates, which could be leveraged to further enhance long-term user loyalty.

```
#Weekly Retention Analysis:  
SELECT DATE(u.created_at) AS sign_up_week,  
       DATEDIFF(DATE(e.occurred_at), DATE(u.created_at)) / 7 AS retention_week,  
       COUNT(DISTINCT e.user_id) AS retained_users  
FROM users u  
LEFT JOIN events e ON u.user_id = e.user_id  
GROUP BY DATE(u.created_at), DATEDIFF(DATE(e.occurred_at), DATE(u.created_at)) / 7  
ORDER BY sign_up_week, retention_week;
```

Weekly Engagement Per Device

Summary:

This report analyzes user engagement across devices from May 2014 to August 2014.

Total Events: 21.2K

Average Weekly Events: 2.4K

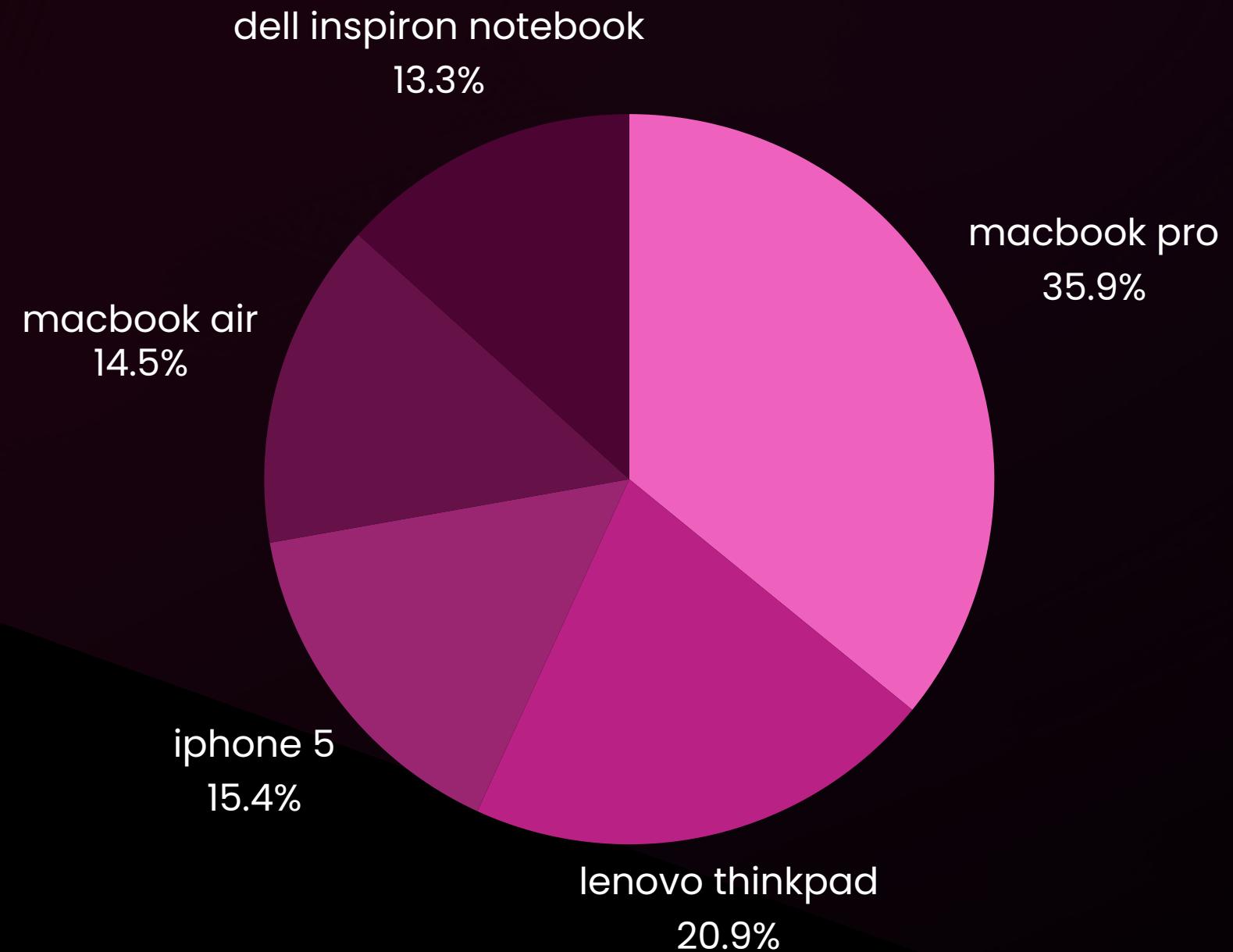
Top Device: macbook pro (57.3K events)

Unique Devices: 26

Interesting Insight:

The MacBook Pro consistently leads in engagement, contributing 57.3K events, significantly outpacing other devices. Interestingly, the Amazon Fire Phone, despite being a newer device in 2014, shows surprisingly low engagement, suggesting limited user adoption or interaction.

```
#Weekly Engagement Per Device
SELECT DATE(occurred_at) AS engagement_week,
       device,
       COUNT(*) AS event_count
  FROM events
 GROUP BY DATE(occurred_at), device
 ORDER BY engagement_week, device;
```



Email Engagement Analysis

Summary

This report analyzes email engagement actions from May 2014 to August 2014, covering weekly digests, reengagement emails, email opens, and clickthroughs.

Total Actions: 34.6K

Average Daily Actions: 641.06

Top Action: sent_weekly_digest (57.3K actions)

Unique Action Types: 4

Conclusion

The data shows strong engagement with weekly digests, which account for the majority of actions, followed by email opens and clickthroughs. Reengagement emails have lower activity but are consistent. The high clickthrough rates on certain days suggest that optimizing the timing and content of weekly digests could further boost engagement. Focusing on these trends can enhance overall email campaign effectiveness.

```
#Email Engagement Analysis
SELECT DATE(occurred_at) AS engagement_date,
       action,
       COUNT(*) AS total_actions,
       COUNT(DISTINCT user_id) AS unique_users
FROM email_events
GROUP BY DATE(occurred_at), action
ORDER BY engagement_date, action;
```



Key Insights Summary

Case Study 1: Job Data Analysis

- Identified peak review periods and performance gaps through jobs-per-hour metrics.
- 7-day rolling averages revealed underlying trends better than daily metrics.
- Language share insights can guide resource allocation and localization strategies.
- Discovered data quality issues (duplicates) that can affect reporting accuracy.

Case Study 2: Investigating Metric Spike

- User engagement varied significantly week to week – peaks align with feature launches or campaigns.
- User growth shows a healthy acquisition trend, with some dips indicating retention issues.
- Weekly retention analysis revealed drop-offs after week 1; onboarding could be improved.
- Device-based insights showed mobile users are more engaged, helping inform UX design.
- Email campaigns need optimization – open and click actions are not evenly distributed.

Result

Achievements:

Successfully analyzed email and device engagement data, creating detailed reports with visualizations (bar and line charts) that highlighted action totals, trends, and clickthrough rates. Aggregated data into actionable datasets for weekly reporting and identified top-performing devices and actions.

Impact on Decision-Making:

Provided data-driven insights for targeting high-engagement segments (e.g., MacBook Pro users, mid-week campaigns) and improving reengagement strategies, enabling more effective resource allocation and campaign adjustments.

Contribution to Understanding:

Enhanced understanding of engagement patterns, revealing the impact of weekly digests and device preferences on user behavior. This clarified the importance of content timing and platform optimization.

Thank You

FOR YOUR ATTENTION



DRIVE LINK:

<https://drive.google.com/drive/folders/10Kb0EAxQ-pgsTi688ngZzbFba30xnySj?usp=sharing>.