PROJECT 2

INSTAGRAM USER ANALYTICS

By SANTASISH BHATTACHARYA

Table of Contents

- Project Description
- Approach
- Tech Stack Used
- Insights
- Results

Project Description: Instagram User Analytics

This project focuses on conducting a thorough user analysis of Instagram, the popular social media platform. The primary aim is to provide valuable insights and data-driven recommendations to the management team. By examining user behavior and engagement patterns, the project will support decision-making processes related to marketing campaigns, investor metrics, user experience enhancements, and overall business growth.

The project will address various key areas, including:

- **Rewarding Most Loyal Users:** Recognizing and appreciating long-time users who have been with Instagram since the beginning. This will help foster loyalty and improve user satisfaction.
- **Reminding Inactive Users to Start Posting:** Identifying users who have never posted a single photo on Instagram. This will enable targeted promotional efforts to encourage them to engage and contribute content to the platform.
- **Declaring Contest Winner:** Determining the winner of a contest based on the user who receives the most likes on a single photo. This will ensure a fair and engaging contest experience for participants.
- **Hashtag Researching:** Identifying the most used hashtags on the platform. This information will assist partner brands in reaching a wider audience and increasing engagement with their posts.
- **Launch AD Campaign:** Analyzing user registration patterns to determine the optimal day of the week to launch ad campaigns. This will help maximize the reach and impact of the campaigns.

Furthermore, the project will provide insights on investor metrics, including user engagement and the presence of fake or suspicious accounts:

- **User Engagement:** Assessing the level of user engagement by determining the average number of posts per user on Instagram. This metric will help gauge the overall activity and participation of users on the platform.
- **Bots & Fake Accounts:** Identifying users (bots) who have liked every single photo on the platform, indicating potential fake or suspicious accounts. This analysis will help ensure the authenticity and integrity of the platform.

The project's goal is to contribute to the success and growth of Instagram by providing actionable insights that drive effective marketing strategies, enhance user engagement, and maintain a trusted platform for users and investors alike.

Approach to Project Execution

Our project aimed to achieve several objectives related to user engagement, rewarding loyalty, email marketing, contest winner selection, hashtag research, ad campaign scheduling, investor metrics, and identifying fake accounts. Here is an **overview** of our approach and execution for each objective:

1. **Rewarding Most Loyal Users:**
    - Identified the top five users who registered earliest by using the DATEDIFF() function on the 'created_at' column in the users table.
    - Sorted the usernames based on the length of time they have been on the platform.
2. **Reminding Inactive Users to Start Posting:**
    - Determined the number of users who registered but never posted a photo.
    - Utilized a LEFT JOIN between the users and photos tables to identify users without photo records.
    - Filtered the results to obtain the count of users who have not posted.
3. **Declaring Contest Winner:**
    - Identified the photo with the most likes using a subquery.
    - Retrieved the user details and additional post information based on the winning photo's photo_id.
4. **Hashtag Researching:**
    - Leveraged data from the 'tags' and 'photo-tags' tables to count hashtags used on the platform.
    - Ordered the results by tag count in descending order and limited to the top five hashtags.
5. **Launch Ad Campaign:**
    - Analyzed the registration count based on individual days of the week using the DAYNAME() function.
    - Analyzed the registration count based on time baskets using the CONCAT() and FLOOR() functions.

Investor Metrics:

6. **User Engagement:**
    - Provided the total number of photos on the platform divided by the number of users.

- Calculated the number of posts by individual users compared to the average number of posts on the platform.

7. **Bots & Fake Accounts:**
   - Detected users (bots) who have liked every single photo on the platform.
   - Utilized a scalar subquery to compare the count of likes by users with the total number of photos in the 'likes' table.

This approach allowed us to effectively address the project objectives and provide valuable insights to our stakeholders.

Tech Stack Utilized:

To accomplish the objectives outlined in this report, the tech stack incorporates **MySQL Workbench 8.0 CE** as a crucial tool. This powerful database management system enables the creation and management of databases to store the relevant data. A dedicated database has been created and utilized for this project '**ig_clone**'.

The core technology employed for data analysis is SQL (Structured Query Language). Through the utilization of SQL queries, the data stored in the database will be efficiently retrieved, manipulated, and analyzed. These queries will extract the necessary information to generate comprehensive reports tailored to the needs of the leadership team.

By leveraging the capabilities of **MySQL Workbench 8.0 CE** and employing **SQL queries**, the tech stack facilitates the seamless execution of the analysis process.

Insights:

Marketing Team

Rewarding Most Loyal Users

- The usernames 'Darby_Herzog' and 'Emilio_Bernier52' were the earliest registered users on the platform.
- Despite being early users, 'Darby_Herzog' and 'Nia_Haag' have not posted any photos, suggesting low engagement.
- 'Nia_Haag' is suspected to be a bot since it has liked all the platform's photos.
- The position of the most loyal user can be awarded to 'Jordyn.Jacobson2', who registered on the same day as 'Nia_Haag' and ranks next in terms of earliest users.

Username of the 5 Oldest users
Darby_Herzog
Emilio_Bernier52
Elenor88
Nicole71
Nia_Haag (bot) / Jordyn.Jacobson2

Note: The usernames have been sorted according to the highest number of days joined.

Reminding Inactive Users to Start Posting:

- Approximately 26% of users on the platform are inactive, having created profiles but not posted any photos.
- Out of the 26 inactive users, two of them are also among the top five most loyal users, indicating very low engagement.

- The list of inactive users who need to be reminded includes usernames: -

id	username
80	Darby_Herzog
71	Nia_Haag
91	Bethany20
66	Mike.Auer39
34	Pearl7
41	Mckenna17
76	Janelle.Nikolaus81
36	Ollie_Ledner37
89	Jessyca_West
75	Leslie67
25	Tierra.Trantow
49	Morgan.Kassulke
83	Bartholome.Bernhard
68	Franco_Keebler64
5	Aniya_Hackett
7	Kasandra_Homenick
54	Duane60
81	Esther.Zulauf61
21	Rocio33
74	Hulda.Macejkovic
57	Julien_Schmidt
45	David.Osinski47
14	Jaclyn81
53	Linnea59
90	Esmeralda.Mraz57
24	Maxwell.Halvorson

These are the **26** users who have a profile but have not posted yet!

Declaring Contest Winner

Winner: Zack_Kemmer93

Zack_Kemmer93 (user_id=52) is the **winner of the contest** for posting the photo (photo_id=145, image_url= https://jarret.name) on 2023-06-13.Zack_Kemmer93 has joined the platform on 2017-01-01 and posted 5 photos since then.

His winning photo has recieved **48** likes which is the most by any single post on the platform. The hashtags Zack used in this post are: -

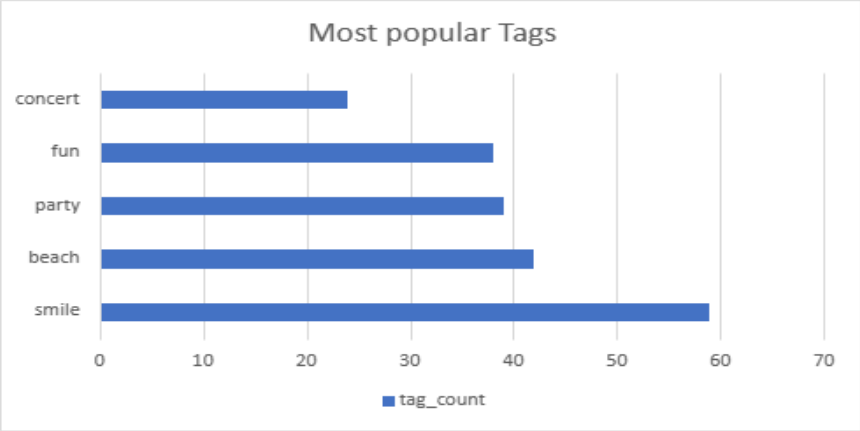
- fun
- party
- concert
- drunk
- Smile.

Hashtag Researching

1.The most used hashtag on the platform **#smile**

2. The best hashtags to use to reach the most people on the platform (ordered according to their popularity)

- Smile
- Beach
- Party
- Fun
- Concert



Launch Ad Campaign

- Our users have registered maximum on **Thursdays** (16%) and **Sundays** (16%), followed by Fridays (15%).

Day wise analysis

Day_of_week	Registration_count
Thursday	16
Sunday	16
Tuesday	14
Saturday	12
Wednesday	13
Monday	14
Friday	15

Thursdays and Sundays have the maximum registrations followed by Friday.

Hour-wise/Time Basket wise analysis

Time Baskets	Registration_count
16:00 - 19:59	23
8:00 - 11:59	11
0:00 - 3:59	17
12:00 - 15:59	16
4:00 - 7:59	11
20:00 - 23:59	22

- **45%** users have registered between 8 hrs of a day! (**4pm to midnight**)
- **62%** users have registered between **noon** (12:00) and **midnight** (0:00).

Thursday post noon can be a good time to begin the ad campaign as the window will continue on to Friday and then Sunday.

Investor Metrics

User Engagement:

- Total number of photos in Instagram: 257
- Total number of users: 100
- No. of active users who have posted: 74
- **Total number of photos on Instagram/total number of users**: 2.57
- User engagement shows that **58.11%** of active users (users who have posted at least once on the platform) posts more than the average number of posts per user, while **41.89%** active users post less than the average.

Users_posting	count	avg_post_per_user	percentage
more_than_average	34	2.57	58.11
less_than_average	40	2.57	41.89

Note: In this analysis, 'percentage' is calculated based on (Total photos)/ (Active users).

Bots & Fake Accounts:

- **13%** of the accounts on the platform are bots or dummy accounts that have liked all posts.
- Among the bots, approximately **69.23%** registered in 2017 and the remaining **30.76%** registered in 2016.
- No bots have registered on the platform since 2017.

Number_of_bots_registered	year
9	2016
4	2017

These insights provide valuable information regarding user behavior, engagement levels, hashtag usage, optimal ad campaign timing, and the presence of bots. They serve as a foundation for making informed decisions and formulating effective strategies to enhance user engagement, attract a wider audience, and ensure the authenticity of the platform.

Results

Throughout the project, we conducted a comprehensive analysis of various aspects of our platform, gaining valuable insights that can inform strategic decision-making.
By rewarding the most loyal users, we identified early registrants who had not yet posted and suspected bots, allowing us to **recognize and reward genuine user engagement**. Reminding inactive users to start posting presents an **opportunity to re-engage** them and increase overall platform activity.
We declared a contest winner based on the photo with the highest number of likes, providing **recognition and encouragement** for users to participate in such events. Hashtag researching helped us understand popular tags on the platform, enabling us to optimize **content visibility** and reach a wider audience. Analyzing user registration counts by day and time allowed us to **identify** optimal days and time periods for launching ad campaigns, maximizing user registrations.
Investor metrics, including **user engagement** and the **presence of bots**, provided valuable insights into the overall health and authenticity of our platform. We now have a better understanding of **user behavior**, allowing us to make informed decisions to enhance engagement and attract more users.

This project has been instrumental in providing actionable insights that can guide our marketing strategies, user engagement initiatives, and overall platform growth. It has helped us understand user preferences, optimize content reach, and identify potential areas for improvement. With these insights, we can create a more engaging and vibrant platform that resonates with our target audience.

In conclusion, the project has enabled us to **reward loyal users, encourage inactive users, identify contest winners, optimize hashtag usage, determine optimal ad campaign timings, assess user engagement, and address the presence of bots**. These findings empower us to make data-driven decisions that drive user engagement, improve user experience, and foster the growth of our platform.