Instagram

USER ANALYTICS

CONTENT

PROJECT DESCRIPTION

The Instagram User Analytics project aims to analyze and derive valuable insights from user data on the Instagram platform. The project focuses on understanding user behavior, engagement patterns, and overall platform usage. By analyzing various data points and metrics, the project aims to provide meaningful information that can help individuals or organizations gain a better understanding of their Instagram user base and optimize their strategies accordingly.

APPROACH

For marketing purposes, I identified the 5 oldest users and users who have never posted a photo. This allowed the marketing team to reward loyal users and encourage inactive users to start posting. I also determined the contest winner based on the most likes on a single photo, providing their details.

For this Instagram User Analytics project, I adopted a data-driven approach to extract and analyze relevant information from the Instagram database using MySQL. By leveraging SQL queries, I was able to retrieve the necessary data to address the marketing and investor metrics objectives.

In this project, I used SQL queries to retrieve accurate data and extract valuable insights. By combining database querying, data analysis, and clear reporting, I provided insights into user behavior, engagement patterns, and platform performance on Instagram. This enables informed decision-making and strategic planning.

In terms of investor metrics, I assessed user engagement by calculating the average number of posts per user. I also examined the presence of bots and fake accounts by identifying users exhibiting unusual behavior.

TECH-STACK USED

For my Instagram User Analytics project, I utilized MySQL for data extraction and analysis. By leveraging MySQL's capabilities, I efficiently retrieved and queried the data from the database. Employing various analysis techniques, I gained valuable insights into user behavior, campaign effectiveness, and platform performance. This enabled data-driven decision-making and optimization of marketing strategies. Overall, the combination of MySQL and data analysis techniques proved effective in understanding and improving user engagement on Instagram.

INSIGHTS - MARKETING

```
mysql> SELECT id, username, created_at
    -> FROM users
    -> ORDER BY created_at ASC
    -> LIMIT 5;
+-----+-----------------+---------------------+
| id  | username        | created_at          |
+-----+-----------------+---------------------+
| 180 | Darby_Herzog    | 2016-05-06 00:14:21 |
|  80 | Darby_Herzog    | 2016-05-06 00:14:21 |
| 280 | Darby_Herzog    | 2016-05-06 00:14:21 |
|  67 | Emilio_Bernier52 | 2016-05-06 13:04:30 |
| 267 | Emilio_Bernier52 | 2016-05-06 13:04:30 |
+-----+-----------------+---------------------+
5 rows in set (0.00 sec)
```

Rewarding Most Loyal Users: By identifying the 5 oldest users on Instagram, we can acknowledge their long-term commitment to the platform. This insight allows us to prioritize these users for special rewards, exclusive offers, or loyalty programs, fostering their continued engagement and strengthening their loyalty even further.

- Multiple entries exist for the user "Darby_Herzog" with IDs 180, 80, and 280 and the user "Emilio_Bernier52" with IDs 67 and 267, indicating possible data inconsistencies or multiple associated accounts.

INSIGHTS - MARKETING

```
mysql> SELECT id, username
    -> FROM users
    -> WHERE id NOT IN (SELECT DISTINCT user_id FROM photos);
+-----+---------------------+
| id  | username            |
+-----+---------------------+
|   5 | Aniya_Hackett       |
```

⋮ ⋮ ⋮

```
| 299 | Alek_Watsica        |
| 300 | Javonte83           |
+-----+---------------------+
226 rows in set (0.00 sec)
```

Remind Inactive Users to Start Posting: Identifying users who have never posted a single photo provides an opportunity to re-engage them with the platform. By sending personalized promotional emails, we can encourage these inactive users to share their first photo. This strategy aims to increase user activity, generate new content, and revitalize the overall user experience. There are 226 users in total who have never posted a photo.

INSIGHTS - MARKETING

```
mysql> SELECT u.id, u.username, p.id AS photo_id, COUNT(*) AS likes_count
    -> FROM users u
    -> JOIN photos p ON u.id = p.user_id
    -> JOIN likes l ON p.id = l.photo_id
    -> GROUP BY u.id, u.username, p.id
    -> HAVING likes_count = (SELECT MAX(likes_count)
    ->                 FROM (SELECT COUNT(*) AS likes_count
    ->                 FROM likes GROUP BY photo_id) AS t);
+----+--------------+----------+-------------+
| id | username     | photo_id | likes_count |
+----+--------------+----------+-------------+
| 52 | Zack_Kemmer93 |      145 |          48 |
+----+--------------+----------+-------------+
1 row in set (0.02 sec)
```

Declaring Contest Winner: Identifying the user with the most likes on a single photo allows us to declare the contest winner accurately. This insight helps us recognize and reward the user who has generated the highest level of engagement, boosting user participation in future contests and driving overall user interaction on the platform.

The contest winner is "Zack_Kemmer93" with photo ID 145, which received the highest number of likes (48) among all the contest photos.

INSIGHTS - MARKETING

```
mysql> SELECT
    ->   t.Ranking,
    ->   t.tag_name,
    ->   t.tag_count
    -> FROM (
    ->   SELECT
    ->     @rank := IF(@prev_count = t.tag_count, @rank, @rank + 1) AS Ranking,
    ->     t.tag_name,
    ->     @prev_count := t.tag_count AS tag_count
    ->   FROM (
    ->     SELECT tag_name, COUNT(*) AS tag_count
    ->     FROM tags t JOIN photo_tags pt ON t.id = pt.tag_id
    ->     GROUP BY tag_name
    ->     ORDER BY tag_count DESC
    ->   ) t
    ->   CROSS JOIN (SELECT @rank := 0, @prev_count := NULL) r
    -> ) t
    -> WHERE
    ->   t.Ranking <= 5;
+---------+----------+-----------+
| Ranking | tag_name | tag_count |
+---------+----------+-----------+
| 1       | smile    |        59 |
| 2       | beach    |        42 |
| 3       | party    |        39 |
| 4       | fun      |        38 |
| 5       | concert  |        24 |
| 5       | food     |        24 |
| 5       | lol      |        24 |
+---------+----------+-----------+
7 rows in set, 4 warnings (0.00 sec)
```

Hashtag Researching: Identifying the top 5 most commonly used hashtags on the platform helps us understand the preferences and interests of the user base. This knowledge allows us to align our content strategy and marketing campaigns with popular trends, improving the discoverability and reach of our posts, and effectively targeting a larger audience. By using these popular hashtags in their posts, the partner brand can increase their reach and visibility on the platform, potentially attracting more users and engagement.

INSIGHTS - MARKETING

```
mysql> SELECT DAYNAME(created_at) AS registration_day, COUNT(*) AS registration_count
    -> FROM users
    -> GROUP BY registration_day
    -> HAVING registration_count = (
    ->     SELECT MAX(registration_count)
    ->     FROM (
    ->         SELECT DAYNAME(created_at) AS registration_day, COUNT(*) AS registration_count
    ->         FROM users
    ->         GROUP BY registration_day
    ->     ) subquery
    -> );
+------------------+--------------------+
| registration_day | registration_count |
+------------------+--------------------+
| Thursday         |                 48 |
| Sunday           |                 48 |
+------------------+--------------------+
2 rows in set (0.00 sec)
```

Launch AD Campaign: Analyzing the day of the week when most users register provides insights for scheduling ad campaigns. By launching ads on days when user registrations are highest, we can maximize the visibility and impact of our advertisements, ensuring optimal exposure and potential reach to the target audience. When scheduling an ad campaign, it would be advisable to consider launching it on either Thursday or Sunday to maximize the visibility and potential reach to new users.

INSIGHTS - INVESTOR METRICS

```
mysql> SELECT
    ->     (SELECT COUNT(*) FROM photos) AS total_photos,
    ->     (SELECT COUNT(*) FROM users) AS total_users,
    ->     (SELECT COUNT(*) FROM photos) / (SELECT COUNT(*) FROM users) AS photos_per_user,
    ->     (SELECT AVG(post_count) FROM (SELECT user_id, COUNT(*)
    ->             AS post_count FROM photos GROUP BY user_id)
    ->          AS user_posts) AS average_posts_per_user;
+--------------+-------------+----------------+----------------------+
| total_photos | total_users | photos_per_user | average_posts_per_user |
+--------------+-------------+----------------+----------------------+
|          771 |         300 |         2.5700 |                10.4189 |
+--------------+-------------+----------------+----------------------+
1 row in set (0.01 sec)
```

User Engagement: Understanding user engagement is crucial. The average number of posts per user offers insights into activity levels. If this average declines over time, it may signal reduced engagement. To re-engage users, strategies like introducing new features, enhancing the user interface, or launching interactive campaigns can be implemented. On Instagram, users post an average of 10 times, but with a total of only 771 photos, activity levels vary. Some highly active users skew the average, while others post less frequently. These insights aid in evaluating user engagement and understanding posting habits.

INSIGHTS - INVESTOR METRICS

```
mysql> SELECT user_id
    -> FROM likes
    -> GROUP BY user_id
    -> HAVING COUNT(DISTINCT photo_id) = (SELECT COUNT(*) FROM photos);
Empty set (0.01 sec)
```

Bots & Fake Accounts: Executing a query to identify users who have liked every photo on the platform yielded an empty set, indicating no apparent instances of bots or fake accounts engaging in systematic liking. However, it's important to note that this query alone cannot guarantee the absence of such accounts. Other indicators, user behavior analysis, and advanced detection algorithms may be necessary to comprehensively address the presence of bots or fake accounts on the platform. Further investigation is advised to determine the extent of their presence.

RESULT

During the process of making this project, I have achieved several significant milestones. Firstly, I gained hands-on experience in working with a real-world dataset and using SQL queries to extract meaningful information. This allowed me to develop strong data analysis and querying skills. Additionally, I learned how to interpret the insights derived from the data, such as user engagement patterns, popular hashtags, and identifying potential fake accounts. This project has helped me strengthen my analytical thinking, problem-solving, and presentation skills. It has provided me with practical knowledge in the field of data analysis and has boosted my confidence in handling similar projects in the future.

THANK YOU!

PRESENTED BY: MIHIR PATEL