

Instagram User Analytics

Project Description: The aim of this project is to pull out actionable insights from raw data by utilizing different database management tools and to visualize these insights, ultimately improving the platform's efficiency.

Approach: SQL was used to execute the project, where queries helped build a database from the raw data. These queries then facilitated the sorting and extraction of data to uncover the necessary insights.

Tech-Stack Used: MySQL WorkBench v8.0.40 was utilized as part of the tech stack, proving to be an excellent tool for database querying due to its easy access, simple configuration, intuitive graphical user interface, and effective troubleshooting support.

Insights:

A) Marketing Analysis:

1. **Loyal User Reward:** The marketing team wants to reward the most loyal users, i.e., those who have been using the platform for the longest time. Your Task: Identify the five oldest users on Instagram from the provided database.

Code:

```
select * from users  
  
order by created_at limit 5;
```

Output:

	id	username	created_at
▶	80	Darby_Herzog	2016-05-06 00:14:21
	67	Emilio_Bernier52	2016-05-06 13:04:30
	63	Elenor88	2016-05-08 01:30:41
	95	Nicole71	2016-05-09 17:30:22
	38	Jordyn.Jacobson2	2016-05-14 07:56:26

2. Inactive User Engagement: The team wants to encourage inactive users to start posting by sending them promotional emails.
Your Task: Identify users who have never posted a single photo on Instagram.

Code:

```
select username, image_url from photos as p
right join users as u
on p.user_id = u.id
where p.image_url is null;
```

Output:

	username	image_url
▶	Aniya_Hackett	NULL
	Kasandra_Homenick	NULL
	Jadlyn81	NULL
	Rocio33	NULL
	Maxwell.Halvorson	NULL
	Tierra.Trantow	NULL
	Pearl7	NULL
	Ollie_Ledner37	NULL
	Mckenna17	NULL
	David.Osinski47	NULL
	Morgan.Kassulke	NULL
	Linnea59	NULL
	Duane60	NULL
	Julien_Schmidt	NULL
	Mike.Auer39	NULL
	Franco_Keebler64	NULL
	Nia_Haag	NULL
	Hulda.Macejkovic	NULL
	Leslie67	NULL
	Janelle.Nikolaus81	NULL
	Darby_Herzog	NULL
	Esther.Zulauf61	NULL
	Bartholome.Bernhard	NULL
	Jessyca_West	NULL
	Esmeralda.Mraz57	NULL
	Bethany20	NULL

3. Contest Winner Declaration: The team has organized a contest where the user with the most likes on a single photo wins.

Your Task: Determine the winner of the contest and provide their details to the team.

Code:

```
select likes.photo_id, users.username, count(likes.user_id) as no_of_likes
from likes inner join photos on likes.photo_id = photos.id
inner join users on photos.user_id = users.id
group by likes.photo_id, users.username
order by no_of_likes desc limit 3;
```

Output:

	photo_id	username	no_of_likes
►	145	Zack_Kemmer93	48
	182	Adelle96	43
	127	Malinda_Streich	43

Winner is Zack_Kemmer93 with 48 likes on photo_id 145.

4. Hashtag Research: A partner brand wants to know the most popular hashtags to use in their posts to reach the most people.

Your Task: Identify and suggest the top five most commonly used hashtags on the platform.

Code:

```
select tag_name, count(photo_id) as no_of_hashtags
from photo_tags inner join tags
on tags.id = photo_tags.tag_id
group by tags.tag_name order by no_of_hashtags desc limit 5;
```

Output:

	tag_name	no_of_hashtags
▶	smile	59
	beach	42
	party	39
	fun	38
	concert	24

5. Ad Campaign Launch: The team wants to know the best day of the week to launch ads.

Your Task: Determine the day of the week when most users register on Instagram. Provide insights on when to schedule an ad campaign.

Code:

```
select count(username) as registers , date_format((created_at), '%W') as 'day'
from users group by 2
order by 1 desc;
```

Output:

	registers	day
▶	16	Thursday
	16	Sunday
	15	Friday
	14	Tuesday
	14	Monday
	13	Wednesday
	12	Saturday

According to the data, scheduling an ad campaign on Sunday and Thursday would be beneficial.

B) Investor Metrics:

1. **User Engagement:** Investors want to know if users are still active and posting on Instagram or if they are making fewer posts.

Your Task: Calculate the average number of posts per user on Instagram. Also, provide the total number of photos on Instagram divided by the total number of users.

Code:

with abcd as

```
(select users.id as u_id, count(photos.id) as p_id
```

```
from users left join photos
```

```
on photos.user_id = users.id
```

```
group by users.id)
```

```
select sum(p_id) as total_photos, count(u_id) as total_users,
```

```
sum(p_id) / count(u_id) as posts_per_user
```

```
from abcd;
```

Output:

	total_photos	total_users	posts_per_user
▶	257	100	2.5700

2. Bots & Fake Accounts: Investors want to know if the platform is crowded with fake and dummy accounts.

Your Task: Identify users (potential bots) who have liked every single photo on the site, as this is not typically possible for a normal user.

Code:

with abc as

(select users.username, count(likes.photo_id) as no_of_likes

from likes inner join users on users.id = likes.user_id

group by username)

select username, no_of_likes from abc where no_of_likes = (select count(*)
from photos)

order by username;

Output:

	username	no_of_likes
▶	Aniya_Hackett	257
	Bethany20	257
	Duane60	257
	Jadyn81	257
	Janelle.Nikolaus81	257
	Julien_Schmidt	257
	Leslie67	257
	Maxwell.Halvorson	257
	Mckenna17	257
	Mike.Auer39	257
	Nia_Haag	257
	Ollie_Ledner37	257
	Rocio33	257

Results: The project allowed me to transform raw Instagram user data into actionable insights. I successfully extracted, sorted, and visualized data to understand user engagement patterns. This experience honed my SQL skills and reinforced the critical role of data analytics in improving platform efficiency and user experience. The insights derived from the analysis highlighted key areas for potential growth and optimization, making the project both an educational and impactful experience.