Instagram User Analytics

Project Description

The project aims to analyze popular social media app Instagram to help the marketing and investor teams make informed business decisions.by studying Instagram the marketing team can discover trends such as oldest and inactive users on Instagram, users with the most likes on a single photo, popular hashtags and best day of the week to launch ads. For investors, the project provides insights on user engagement and identifies potential fake accounts, helping them understand Instagram's value and growth potential. This analyze offers valuable insights to guide marketing strategies and investment decisions.

Approach

Firstly, I created a database called ig_clone and established seven tables: comments, follows, likes, photo_tags, photos, tags and users within the database, related to Instagram using sql queries. After setting up the table, I executed sql queries to provide answers to the marketing and investors teams questions, helping them make important business decisions.

Tech-Stack Used

I used My SQL Workbench version 8.0.40 to execute SQL query for analyzing Instagram. Because it has a user-friendly interface that makes managing databases easy and efficient.

Project Insights

Below The insights gained by marketing and investors team by analysing Instagram data.

A) Marketing team Analysis

1.loyal user reward:

Identify the five oldest users on Instagram

Query

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

Result

These are the 5 oldest users on Instagram

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**

identify users who have never posted a single photo on Instagram

Query

```
select users.id, users. username from users left join photos
on users.id = photos.user_id where photos.id is null ;
```

Result

These are the users who have never posted a single photo on Instagram.

id	username
5	Aniya_Hackett
7	Kasandra_Homenick
14	Jadyn81
21	Rocio33
24	Maxwell.Halvorson
25	Tierra.Trantow
34	Pearl7
36	Ollie_Ledner37
41	Mckenna17
45	David.Osinski47
49	Morgan.Kassulke
53	Linnea59
54	Duane60
57	Julien_Schmidt
66	Mike.Auer39
68	Franco_Keebler64
71	Nia_Haag
74	Hulda.Macejkovic
75	Leslie67
76	Janelle.Nikolaus81
80	Darby_Herzog
81	Esther.Zulauf61
83	Bartholome.Bernhard
89	Jessyca_West
90	Esmeralda.Mraz57
91	Bethany20

3.contest winner declaration

Identify the user who has most likes for single photo and declare them as winner in the contest that is organised by the marketing team

Query

```
Select users.username, likes.photo_id , photos.image_url,
count(likes.photo_id) as total from
photos inner join likes on likes.photo_id = photos.id
inner join users on photos.user_id = users.id
group by photos.id order by total desc limit 1;
```

Result

The details of the contest winner is given below:

username	photo_id	image_url	total
Zack_Kemmer93	145	https://jarret.name	48

4. Hashtag research

Identify the top 5 commonly used hashtags

Query

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

Result

This are the top 5 most commonly used hashtags on Instagram

tag_name	total
smile	59
beach	42
party	39
fun	38
concert	24

5.Ad campaign launch

Determine the day of the week when most users register on Instagram

Query

```
select dayname (created_at) as day , count(*) as total_users
group by day order by total_users desc;
```

Result

Given below is the list of users registered on each day

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

Based on this list Thursday and Sunday are the days where most users register on Instagram. So, the marketing team can launch ads on either of these days.

B) Investor Metrics

1. user engagement

calculate the average number of posts per user on Instagram. also provide total no of photos on Instagram divided by the total number of users

Query

```
select
((select count(*) from photos) / (select count(*) from users))
 as  average_ photos_per_user
((select count(image_url) from photos) / (select count (distinct user_id) from photos)) as average_no_of posts_per_users ;
```

Result

average_photos_per_users	average_no_of_posts_per_users
2.5700	3.4730

The average number of posts per user = 3.4730

Total number of photos divided by total number of users = 2.5700

2. Bots and fake accounts

Identify users who have liked every single photo on the site

Query

```
select  likes.user_id, users.username from likes
inner join users on users.id = likes.user_id
group by users.id
having count(likes.photo_id) = (select count(id) from photos);
```

Result

This are the list of users who have liked every single photo on the site.

user_id	username
5	Aniya_Hackett
14	Jaclyn81
21	Rocio33
24	Maxwell.Halvorson
36	Ollie_Ledner37
41	Mckenna17
54	Duane60
57	Julien_Schmidt
66	Mike.Auer39
71	Nia_Haag
75	Leslie67
76	Janelle.Nikolaus81
91	Bethany20

All of the above are fake account s. Because it is not possible for normal users to like every single photo on the site.

Conclusion

In this project I gained experience analysing data using SQL to extract insights that can support business decisions. I identified inactive users, the oldest users, the user with the most likes on a single photo, popular hashtags, the best day of the week for user engagement and potential fake accounts. All of this is was achieved with simple sql queries, providing valuable information for the marketing and investor teams to make informed choices.