Project Description

This project is about user analysis of the Instagram platform using SQL. The purpose of this analysis is to provide insights that can help the marketing, product, and development teams make better decisions. The project involves analysing the database provided to find the 5 oldest users of the platform, users who have never posted a single photo, the winner of a contest, the top 5 most commonly used hashtags, and the best day to launch ADs. The project also aims to assess user engagement on Instagram and the prevalence of bots and fake accounts on the platform.

Approach

To perform this analysis, I used SQL to extract the required data from the provided database. I wrote SQL queries to find the 5 oldest users of the platform, users who have never posted a single photo, the winner of a contest, the top 5 most commonly used hashtags, and the best day to launch ADs. I also used SQL to assess user engagement on Instagram and the prevalence of bots and fake accounts on the platform.

Tech-Stack Used

I used MySQL Workbench v8.0.30.0 to interact with the database and execute SQL queries. MySQL is a free, cross-platform MySQL database manager that provides a user-friendly interface for managing MySQL databases.

Insights

Marketing

1. Top 5 oldest users

```
SELECT *
FROM ig_clone.users
ORDER BY created_at
 LIMIT 5;
```

Result:

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. Remind Inactive Users to Start Posting

```
SELECT
users.id,
username,
users.created_at
FROM ig_clone.users
LEFT JOIN ig_clone.photos
ON users.id = ig_clone.photos.user_id
WHERE ig_clone.photos.user_id IS NULL;
```

Result:

id	username	created_at
5	Aniya_Hackett	2016-12-07 01:04:39
7	Kasandra_Homenick	2016-12-12 06:50:08
14	Jaclyn81	2017-02-06 23:29:16
21	Rocio33	2017-01-23 11:51:15
24	Maxwell.Halvorson	2017-04-18 02:32:44
25	Tierra.Trantow	2016-10-03 12:49:21
34	Pearl7	2016-07-08 21:42:01
36	Ollie_Ledner37	2016-08-04 15:42:20
41	Mckenna17	2016-07-17 17:25:45
45	David.Osinski47	2017-02-05 21:23:37
49	Morgan.Kassulke	2016-10-30 12:42:31
53	Linnea59	2017-02-07 07:49:34
54	Duane60	2016-12-21 04:43:38
57	Julien_Schmidt	2017-02-02 23:12:48
66	Mike.Auer39	2016-07-01 17:36:15
68	Franco_Keebler64	2016-11-13 20:09:27
71	Nia_Haag	2016-05-14 15:38:50
74	Hulda.Macejkovic	2017-01-25 17:17:28
75	Leslie67	2016-09-21 05:14:01
76	Janelle.Nikolaus81	2016-07-21 09:26:09
80	Darby_Herzog	2016-05-06 00:14:21
81	Esther.Zulauf61	2017-01-14 17:02:34
83	Bartholome.Bernhard	2016-11-06 02:31:23
89	Jessyca_West	2016-09-14 23:47:05
90	Esmeralda.Mraz57	2017-03-03 11:52:27
91	Bethany20	2016-06-03 23:31:53

3. Contest Winner

```
SELECT
users.id AS user_id,
username,
photos.id AS photo_id,
photos.image_url,
COUNT(*) AS total_likes_count
FROM ig_clone.photos
JOIN ig_clone.likes
ON photos.id = likes.photo_id
JOIN ig_clone.users
ON users.id = photos.user_id
GROUP BY photos.id
ORDER BY total_likes_count DESC
LIMIT 1;
```

Result:

user_id	username	photo_id	image_url	total_likes_count
52	Zack_Kemmer93	145	https://jarret.name	48

4. Hashtag Researching

```
SELECT
ig_clone.tags.id AS tag_id,
ig_clone.tags.tag_name,
COUNT(*) as total
FROM ig_clone.tags
JOIN ig_clone.photo_tags
ON ig_clone.tags.id = ig_clone.photo_tags.tag_id
GROUP BY ig_clone.tags.id
ORDER BY total DESC
LIMIT 5;
```

Result:

tag_id	tag_name	total
21	smile	59
20	beach	42
17	party	39
13	fun	38
5	food	24

5. AD campaign

```
SELECT
DAYNAME(created_at) AS day_of_the_week,
COUNT(*) AS total_count
FROM ig_clone.users
GROUP by day_of_the_week
ORDER by total_count DESC;
```

Result:

day_of_the_week	total_count
Sunday	16
Thursday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12

B) Investor Metrics:

1. User Engagement

```
SELECT
ROUND(( SELECT COUNT(*) FROM ig_clone.photos ) / ( SELECT COUNT(*) FROM
ig_clone.users ), 2) AS avg_user_photo_post;
```

Result:

avg_user_photo_post
2.57

2. Bot Accounts

```
SELECT
ig_clone.users.id AS user_id,
ig_clone.users.username,
COUNT(*) AS total_user_likes
FROM ig_clone.users
JOIN ig_clone.likes
ON ig_clone.users.id = ig_clone.likes.user_id
GROUP BY users.id
HAVING total_user_likes = (
 SELECT COUNT(*) FROM ig_clone.photos );
```

Result:

user_id	username	total_user_likes
5	Aniya_Hackett	257
14	Jaclyn81	257
21	Rocio33	257
24	Maxwell.Halvorson	257
36	Ollie_Ledner37	257
41	Mckenna17	257
54	Duane60	257
57	Julien_Schmidt	257
66	Mike.Auer39	257
71	Nia_Haag	257
75	Leslie67	257
76	Janelle.Nikolaus81	257
91	Bethany20	257

Through this project, I gained insights into the following aspects of Instagram:

- Instagram has a large number of users who have never posted a single photo on the platform.
- The most loyal users of Instagram have been using the platform since 2016.
- The most commonly used hashtags on Instagram are related to party, fun, and beach.
- Sunday is the day when most users register on Instagram, making it the best day to launch ADs.
- On average, an Instagram user posts about 3 photos per week. However, there are a significant number of users who post much more frequently.
- There is evidence of the presence of bots and fake accounts on Instagram, as some accounts have liked every single photo on the site.

Result

Through this project, I was able to extract valuable insights from the Instagram database that can inform the decision-making process of various teams within the organization. By providing insights on user behaviour, engagement, and preferences, this analysis can help the marketing, product, and development teams make informed decisions that can help grow the business.