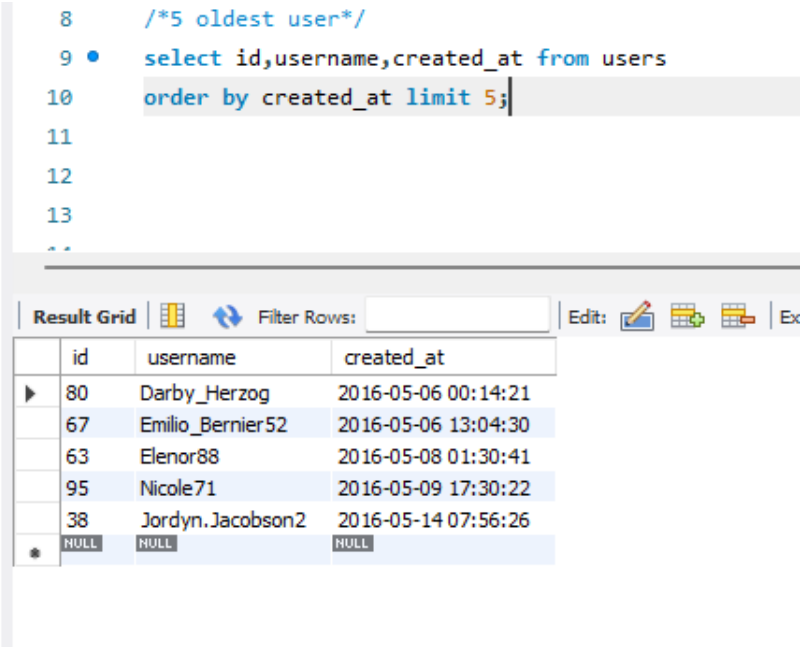
Instagram User Analytics

1. **Project Description:** This project focuses on analysing user interactions and engagement with Instagram, deriving meaningful insights using SQL quires. The primary goal is to understand the concepts of database and management tools and analyse average posts, identify bots, and more.
2. **Approach:** To tackle the project, I setup a MySQL environment using MySQL Workbench.
    - Database: I first understood the structure of the tables and their relationships, for efficient SQL queries.
    - Query Execution: Queries were utilized to create a database from the provided dataset. SQL queries were implemented for sorting and extracting data to obtain valuable insights, such as Loyal users, Inactive users, most commonly used hashtags etc.
3. **Tech-Stack:** I utilized MySQL Workbench v8.0.36 for its robust query building capabilities and visual database design. For extraction of data SQL (Structured Query Language) is utilized for its efficient handling in relational database and performing complex queries or retrieval of data.
4. **Insights:** Through this project, several key insights were gained,
    1. **Loyal user:** Here, I identified 5 oldest users on Instagram i.e. the users that created accounts or using platform for the longest time. Knowing these users helps Instagram maintain a strong relationship with its earliest adopters.



2. **Inactive User Engagement:** Users who have registered but never posted may need re-engagement strategies. I identified users who have never posted a single photo, and that data is retrieved using subquery and selecting distinct id's.

```
17    /*Inactive user*/
18 •  SELECT id, username, created_at
19    FROM users
20    WHERE id NOT IN (SELECT DISTINCT user_id FROM photos);
```

Result Grid | Filter Rows: | Edit: | Export/Import:

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

3. **Contest Winner Declaration:** Determine user with most likes. Recognizing the user with most likes on a post is achieved using Joins. This insight helps in creating similar engagement-driven contests in the future.

```
19    /*Contest Winner Declaration*/
20 •  select l.photo_id,u.username,count(l.user_id) totallikes
21    from likes l join photos p on l.photo_id=p.id
22    inner join users u on p.user_id=u.id
23    group by l.photo_id,u.username
24    order by totallikes desc
25    limit 3;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

photo_id	username	totallikes
145	Zack_Kemmer93	48
182	Adelle96	43
127	Malinda_Streich	43

4. **Hashtag Research:** Identify top 5 most used hashtags. Retrieval of data is done using aggregate function and joins. Understanding the most popular hashtags helps brands and influencers optimize their content for maximum reach.

```
28      /* 5 Most used hashtags*/
29 •    SELECT t.tag_name, COUNT(p.photo_id) as total_count
30      FROM tags t
31      JOIN photo_tags p ON t.id = p.tag_id
32      GROUP BY t.tag_name
33      ORDER BY total_count DESC
34      LIMIT 5;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

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

5. **Ad Campaign Launch:** Determine the day of the week when most users register, DAYNAME() function is used to retrieve the day of the week. Knowing the peak days can help in launching ads that might capture the attention of the new users.

```
37      /* Most user registertion*/
38 •    SELECT DAYNAME(created_at) AS day_of_week, COUNT(*) AS registrations
39      FROM users
40      GROUP BY day_of_week
41      ORDER BY registrations DESC;
```

esult Grid | Filter Rows: | Export: | Wrap Cell Content:

day_of_week	registrations
Thursday	16
Sunday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12

B. Investor Metrics

1. **User Engagement:** Calculate the average number of posts and total number of photos, using the nested subquery placement. It shows the user engagement or user activity.

```
/* B Invester metrics
Calculate average number of posts per user*/
SELECT
(SELECT COUNT(*) FROM photos) / (SELECT COUNT(*) FROM users) AS 'photos/userratio',
(SELECT COUNT(*) FROM photos) / (SELECT COUNT(*) FROM users) AS 'avg_posts/user'
FROM users
limit 1;
```

ult Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

photos/userratio	avg_posts/user
2.5700	2.5700

2. **Bot & Fake accounts:** Identify users who have liked every single photo on the site. Detecting these accounts for crucial for maintaining the integrity of user engagement metrics. This insight helps Instagram in taking measures to remove or monitor the accounts for bots or fake accounts.

```
/*bots fake accounts*/
SELECT u.*, COUNT(l.user_id) AS total_likes
FROM users u
JOIN likes l ON u.id = l.user_id
GROUP BY u.id
HAVING total_likes = (SELECT COUNT(*) FROM photos);
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

id	username	created_at	total_likes
5	Aniya_Hackett	2016-12-07 01:04:39	257
14	Jaclyn81	2017-02-06 23:29:16	257
21	Rocio33	2017-01-23 11:51:15	257
24	Maxwell.Halvorson	2017-04-18 02:32:44	257
36	Ollie_Ledner37	2016-08-04 15:42:20	257
41	Mckenna17	2016-07-17 17:25:45	257
54	Duane60	2016-12-21 04:43:38	257
57	Julien_Schmidt	2017-02-02 23:12:48	257
66	Mike.Auer39	2016-07-01 17:36:15	257
71	Nia_Haag	2016-05-14 15:38:50	257
75	Leslie67	2016-09-21 05:14:01	257
76	Janelle.Nikolaus81	2016-07-21 09:26:09	257
91	Bethany20	2016-06-03 23:31:53	257

Result: This project successfully achieved multiple objectives, significantly benefiting both the organisation and personal skill development. The project required writing complex SQL queries including subqueries, joins, and aggregates which significantly enhanced my SQL skills. Tackling various analytical tasks and solving problems related to user engagement, content popularity, and potential bot activity developed by problem solving skills. The insights from this project helps the platform to continuously improve for better user experience.