

INSTAGRAM USER ANALYTICS

- AJAI ARUMUGAM

PROJECT DESCRIPTION

- ◆ User Analysis is the process of understanding how users interact with your product such as website. Mobile app or any software.
- ◆ User Analysis collect data on users to enhance the product for better user experience.
- ◆ Here in this project we are going to analyse Instagram user data and provide insights to the management team which could potentially influence the future development of the Instagram app

PROJECT APPROACH

Here I used SQL to execute queries and answer questions posed by management teams which is

A. Marketing Analysis

1. Loyal user reward
2. Inactive user engagement
3. Most liked photo
4. Most popular hastags
5. Best days to launch AD campaign

B. Investor Metrics

1. User engagement
2. Identify potential bot & fake accounts

SOFTWARE USED

MySQL Workbench 8.0 CE

INSIGHTS

A. MARKETING

1. Loyalty User Reward –
- Marketing team wants to reward top 5 most loyal users i.e. oldest users.
 - From the database here are the oldest users

The screenshot shows a database management interface. On the left, a 'Navigator' pane displays a schema named 'ig_clone' with a table 'comments' containing columns 'id', 'comment_text', 'user_id', 'photo_id', and 'created_at'. The 'Columns' section is expanded, showing these columns. Below the schema, there are sections for 'Administration' and 'Information'. The main area shows a SQL query in a text editor:

```
1 • USE ig_clone;
2 • SELECT * FROM users
3 • ORDER BY created_at
4 • LIMIT 5;
```

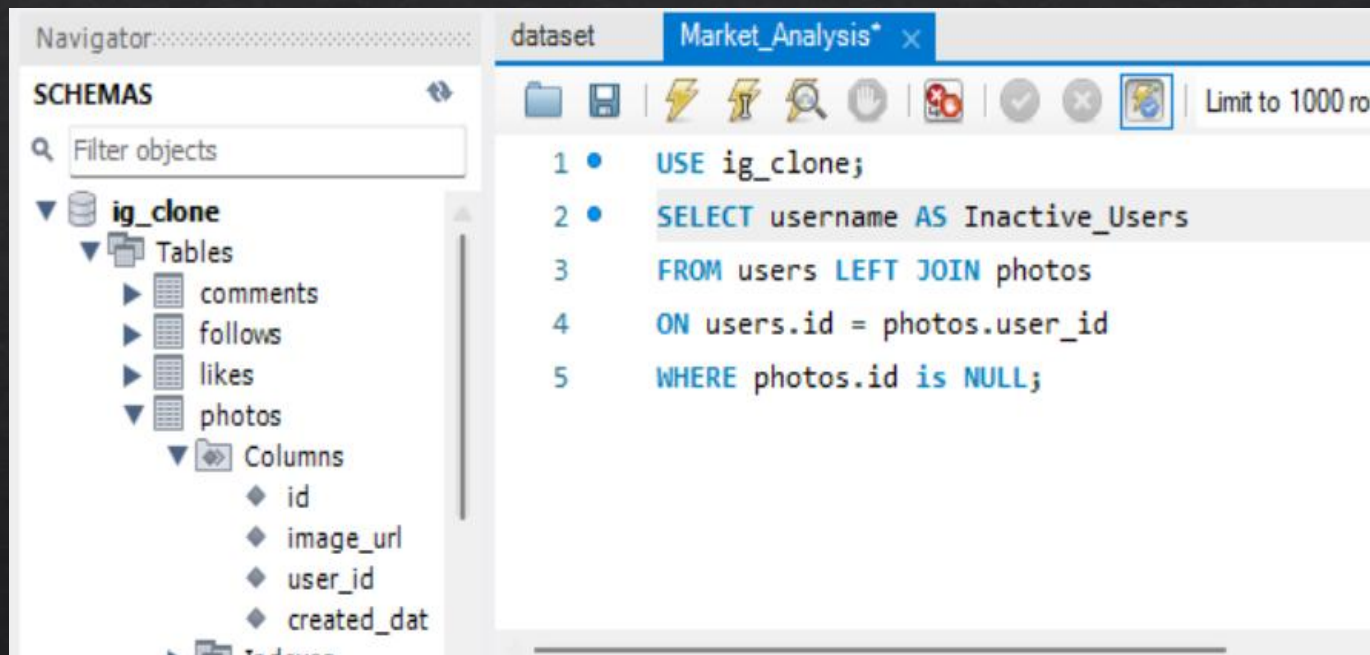
Below the query, a 'Result Grid' displays the top 5 oldest users:

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
NULL	NULL	NULL



Result Grid			
Filter Rows:			
	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
✱	NULL	NULL	NULL



2. Inactive user engagement -



- Marketing team wants to encourage inactive users to start posting by sending promotional emails.
- From the database here is the inactive users list



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

Result Grid				Filter
	Inactive_Users			
▶	Aniya_Hackett			
	Kasandra_Homenick			
	Jadyn81			
	Rocio33			
	Maxwell.Halvorson			
	Tierra.Trantow			
	Pearl7			
	Ollie_Ledner37			
	Mckenna17			

Result Grid				Filter Ro
	Inactive_Users			
	David.Osinski47			
	Morgan.Kassulke			
	Linnea59			
	Duane60			
	Julien_Schmidt			
	Mike.Auer39			
	Franco_Keebler64			

Result Grid				Filter
	Inactive_Users			
	Nia_Haag			
	Hulda.Macejkovic			
	Leslie67			
	Janelle.Nikolaus81			
	Darby_Herzog			
	Esther.Zulauf61			
	Bartholome.Bernhard			
	Jessyca_West			
	Esmeralda.Mraz57			
	Bethany20			

3. Contest winner declaration -

- Team wants to conduct a contest for finding the user with the most likes on a single photo wins.
- Here is the insight

The screenshot shows a database IDE interface. On the left, the 'Navigator' pane displays the 'ig_clone' schema with tables 'comments', 'follows', and 'likes'. The 'likes' table is selected, showing columns 'user_id', 'photo_id', and 'created_at'. The main editor pane shows a SQL query:

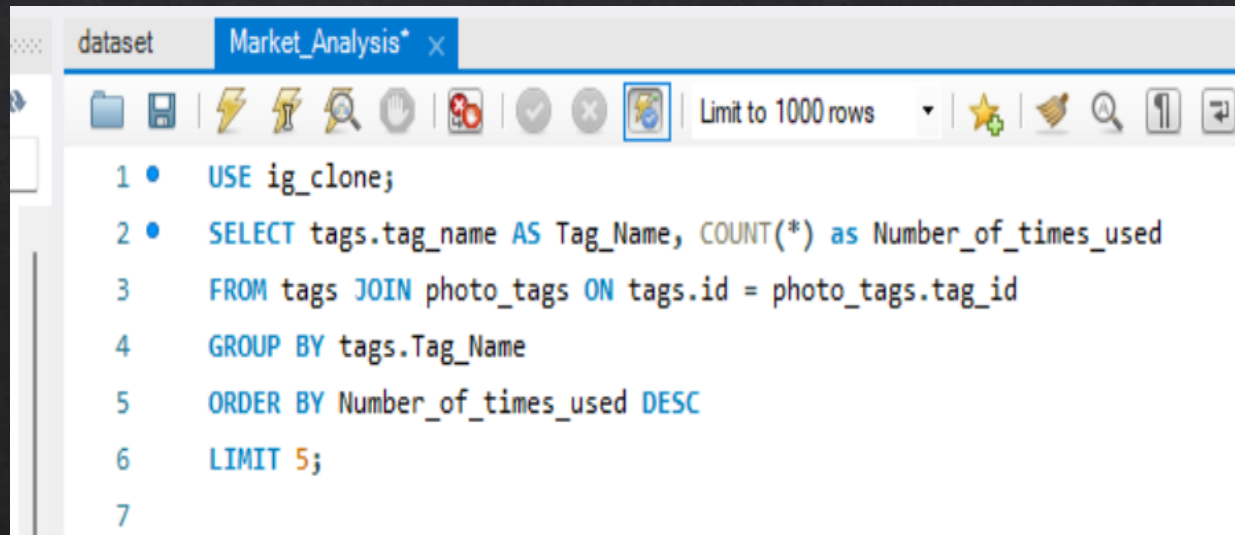
```
1 • USE ig_clone;
2 • SELECT users.id AS User_ID, users.Username, photos.id AS Photo_ID, photos.Image_URL,
3     COUNT(*) AS Total_Likes
4 FROM photos INNER JOIN likes
5 ON likes.Photo_ID= photos.id
6 INNER JOIN users ON photos.User_ID= users.id
7 GROUP BY photos.id
8 ORDER BY Total_Likes DESC
9 LIMIT 1;
```

The 'Result Grid' shows the results of the SQL query. The table has 6 columns: 'User_ID', 'Username', 'Photo_ID', 'Image_URL', and 'Total_Likes'. The first row shows the user with the most likes on a single photo.

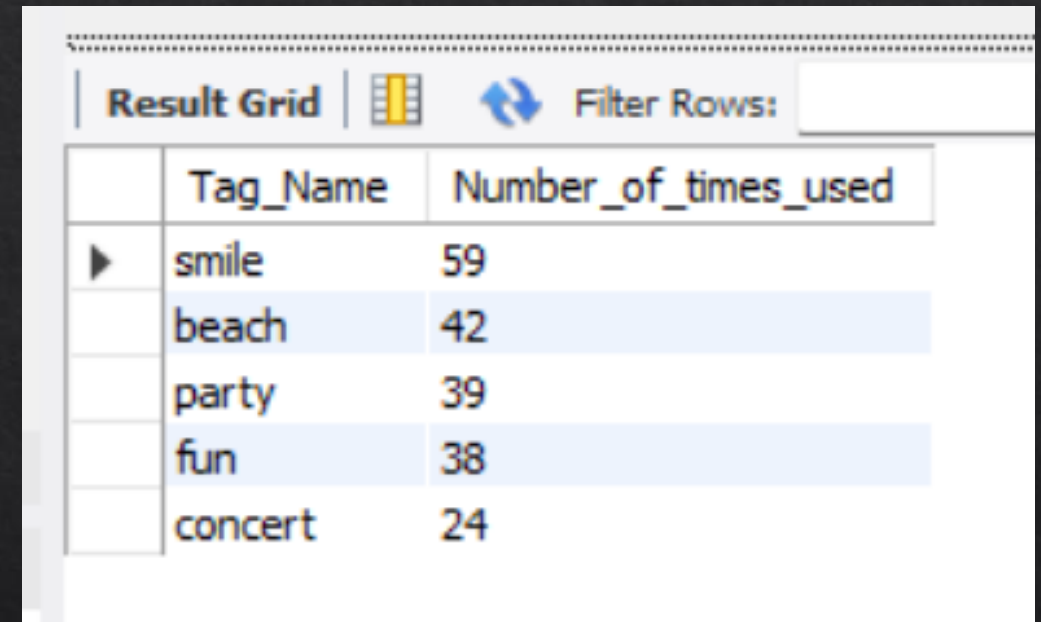
	User_ID	Username	Photo_ID	Image_URL	Total_Likes
▶	52	Zack_Kemmer93	145	https://jarret.name	48

4. Hashtag Research -

- Partner brand wants to know the most popular hashtags to use their posts to reach the most people
- Here is the insights from the given database



```
1 • USE ig_clone;
2 • SELECT tags.tag_name AS Tag_Name, COUNT(*) as Number_of_times_used
3   FROM tags JOIN photo_tags ON tags.id = photo_tags.tag_id
4  GROUP BY tags.Tag_Name
5  ORDER BY Number_of_times_used DESC
6  LIMIT 5;
7
```

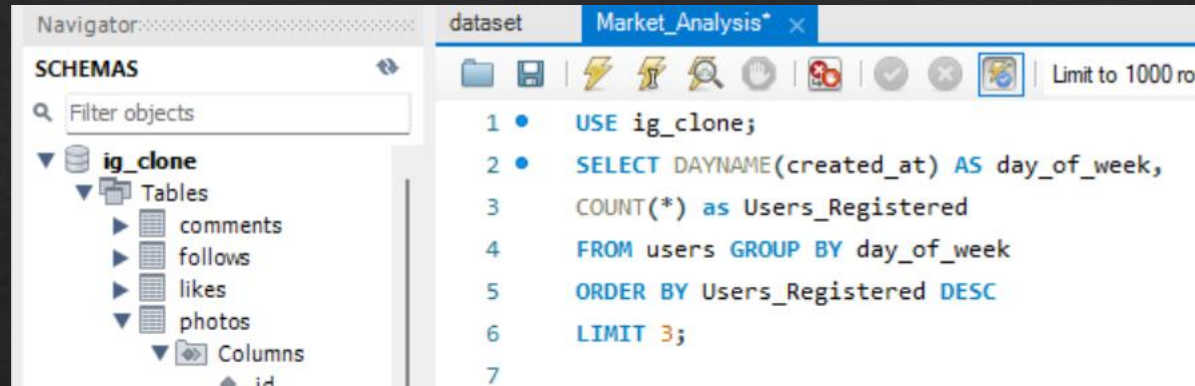


	Tag_Name	Number_of_times_used
▶	smile	59
	beach	42
	party	39
	fun	38
	concert	24

5. Ad Campaign Launch -

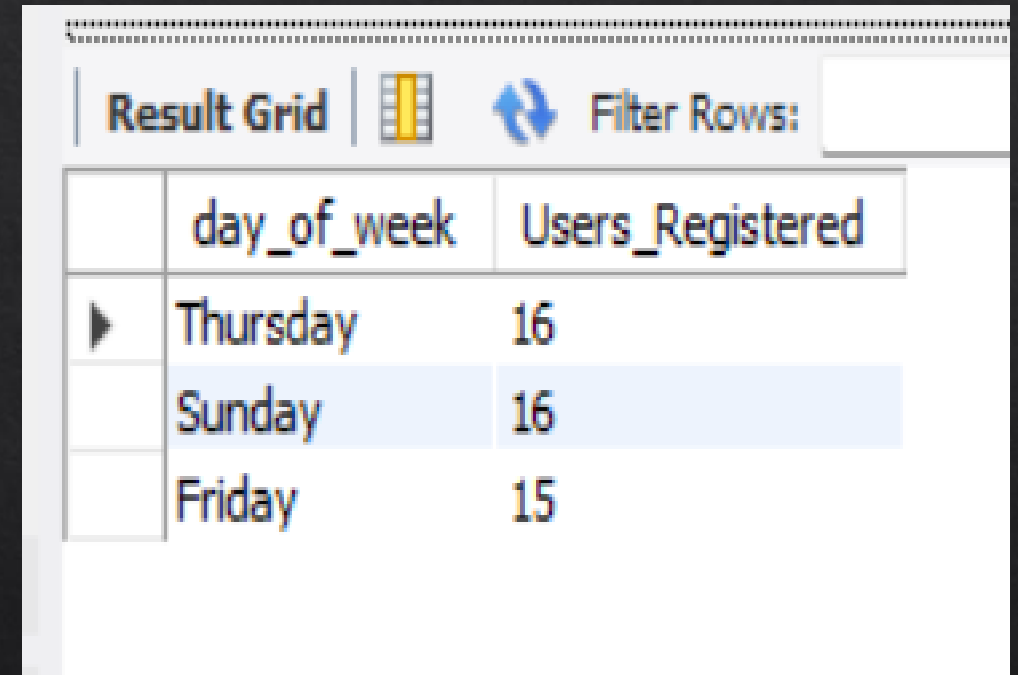
Team wants to know the best day of the week to launch ads.

Here is the result of top 3 day of the week where the most people register on Instagram.



The screenshot shows a database interface with a 'Navigator' pane on the left and a 'dataset' pane on the right. The 'Navigator' pane shows a schema named 'ig_clone' with tables 'comments', 'follows', 'likes', and 'photos'. The 'dataset' pane shows a SQL query being executed on a table named 'users'.

```
1 • USE ig_clone;
2 • SELECT DAYNAME(created_at) AS day_of_week,
3     COUNT(*) as Users_Registered
4     FROM users GROUP BY day_of_week
5     ORDER BY Users_Registered DESC
6     LIMIT 3;
7
```

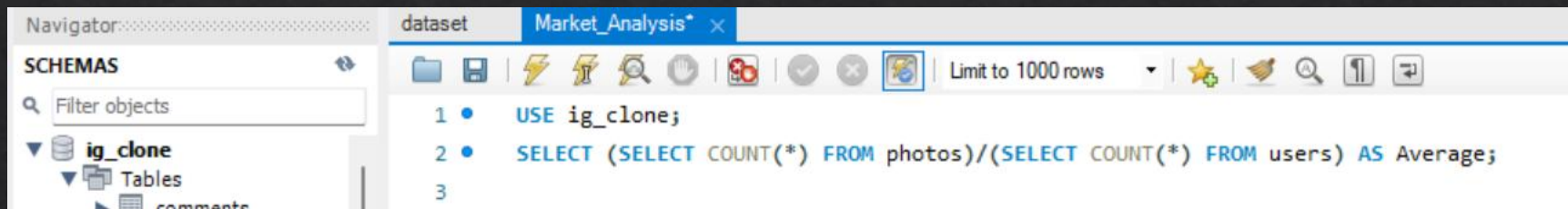


The screenshot shows a 'Result Grid' with a table of results. The table has two columns: 'day_of_week' and 'Users_Registered'. The results are sorted by 'Users_Registered' in descending order, showing Thursday, Sunday, and Friday as the top three days.

	day_of_week	Users_Registered
▶	Thursday	16
	Sunday	16
	Friday	15

B. INVESTOR METRICS

1. User engagement -
 - Investors want to know if the users are still active and posting on Instagram or if they are making fewer posts.
 - Here the insight of average number of posts per user on Instagram.



The screenshot shows a 'Result Grid' with a single row of data. The column header is 'Average' and the value is '2.5700'.

	Average
▶	2.5700

2. Bots and Fake accounts -

- Investors want to know if the platform is crowded with fake and dummy accounts.
- Here are the results of users who have liked every single photo on the site
- Because it is typically not possible for a normal user to do such things.

```
dataset Market_Analysis* x
1 • USE ig_clone;
2 • SELECT user_id, username, COUNT(*) AS total_likes_per_user
3   FROM users INNER JOIN likes ON users.id = likes.user_id
4   GROUP BY likes.user_id
5   HAVING total_likes_per_user = (SELECT COUNT(*) FROM photos);
6
```

	user_id	username	total_likes_per_user
▶	5	Aniya_Hackett	257
	14	Jadyn81	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

INSIGHTS GAINED

- Getting insights is the most important part to make informed decisions
- Insights gained are
 - The most loyal users of Instagram since 2016.
 - Instagram has many users who have never posted a single pic.
 - Most like photo on Instagram.
 - Most used hashtags.
 - Most users registered on Thursday and Sunday in Instagram, making it the best day to launch campaigns.
 - Per week Instagram users posts about 2 to 3 photos.
 - Found the evidence of the presence of bots and fake accounts.

RESULTS

- Through this project I have gained the understanding of how data analytics process works which is database creation and analysis and the importance of SQL in data analysis.
- We can provide insights through SQL which makes the life easier.

THANK YOU