

# Instagram User Analytics

-Saksham Gupta

## Project Description

This project presents solution to the needs and concerns of marketing and investors' team. It focuses on user analytics where it is tracked how users engage and interact with the digital product in an attempt to derive business insights.

The marketing team's wish to launch some campaigns around the following: Rewarding most loyal users, remind inactive users to start posting, declaring contest winner, hashtag researching and launch AD campaign.

While, the investors want to know if Instagram is performing well, they want to assess the app on the following grounds:- User Engagement and Bots & Fake Accounts: The investors want to know if the platform is crowded with fake and dummy accounts.

## Approach

The approach used in this project is pretty straightforward. Accessing the database, applying SQL knowledge to extract the information, thus answering the questions for the marketing team and the investors.

## Tech-Stack used

-MySQL installer community 8.0.30.0.msi was used for this project it contains:- -MySQL Workbench 8.0 CE

## Insights

Got hands-on experience by applying my knowledge gained through this program into this project. It gave me a chance to put my theoretical knowledge to the test and practically performing a task boosted my confidence thus motivating me to explore more into the vast field of data analytics.

## Solution to Question 1

The 5 oldest users of Instagram are DARBY HERZOG, EMILO\_BERNIER52, ELENOR88, NICOLE71, JORDYN.JACOBSON2 based on the date and time they created their account.

ID	Username	Created_at
1	Darby_Herzog	5/6/2016 12:14:21 AM
2	Emilio_Bernier52	5/6/2016 1:04:30 PM
3	Elenor88	5/8/2016 1:30:41 AM
4	Nicole71	5/9/2016 5:30:22 PM
5	Jordyn.Jacobson2	5/14/2016 7:56:26 AM

The screenshot shows the MySQL Workbench interface. The SQL editor contains the following query:

```
1 use ig_clone;
2 select * from users
3 order by created_at
4 limit 5
```

The Results Grid displays the following data:

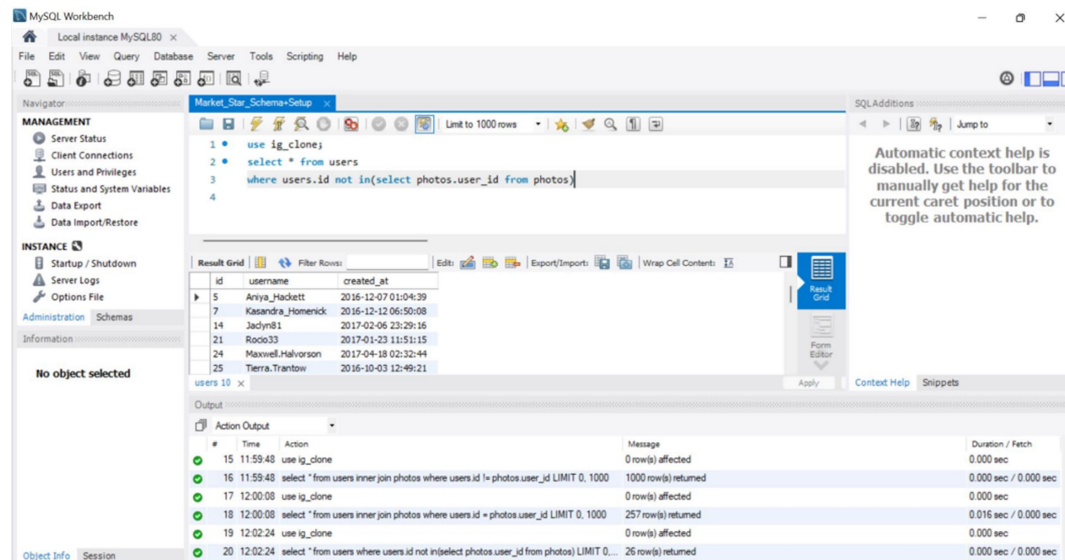
id	username	created_at
80	Darby_Herzog	2016-05-06 00:14:21
67	Emilio_Bernier52	2016-05-06 13:04:30
63	Elenor88	2016-05-08 01:30:41
95	Nicole71	2016-05-09 17:30:22
38	Jordyn.Jacobson2	2016-05-14 07:56:26

The Output tab shows the execution log:

#	Time	Action	Message	Duration / Fetch
9	14:00:33	use ig_clone	0 row(s) affected	0.000 sec
10	14:00:33	select * from users LIMIT 0, 1000	100 row(s) returned	0.016 sec / 0.000 sec
11	14:01:02	use ig_clone	0 row(s) affected	0.000 sec
12	14:01:02	select * from users order by created_at LIMIT 0, 1000	100 row(s) returned	0.000 sec / 0.000 sec
13	14:01:14	use ig_clone	0 row(s) affected	0.000 sec
14	14:01:14	select * from users order by created_at limit 5	5 row(s) returned	0.000 sec / 0.000 sec

## Solution to Question 2

The users who have never posted on Instagram are as follows:



username	
Aniya_Hackett	
Kassandra_Homenick	
Jaclyn81	
Rocio33	
Maxwell.Halvorson	
Tierra.Trantow	
Pearl7	
Ollie_Ledner37	
Mckenna17	
David.Osinski47	Leslie67
Morgan.Kassulke	Janelle.Nikolaus81
Linnea59	Darby_Herzog
Duane60	Esther.Zulauf61
Julien_Schmidt	Bartholome.Bernhard
Mike.Auer39	Jessyca_West
Franco_Keebler64	Esmeralda.Mraz57
Nia_Haag	Bethany20
Hulda.Macejkovic	

## Solution to Question 3

The winner of the contest and details:

The screenshot shows the MySQL Workbench interface with a query editor and a result grid. The query is as follows:

```
4 where likes.user_id = users.id
5 group by photo_id
6 order by Number_of_likes
7 desc
8 limit 5
```

The result grid displays the top 5 users by the number of likes on their photos:

username	photo_id	Number_of_likes
Harley_Lind18	145	48
Andre_Purdy85	127	43
Harley_Lind18	182	43
Aniya_Hackett	123	42
Andre_Purdy85	30	41

The bottom panel shows the Action Output with the following messages:

#	Time	Action	Message	Duration / Fetch
25	14:12:57	use ig_clone	0 row(s) affected	0.000 sec
26	14:12:57	select user_id, photo_id, count(user_id) as Number_of_likes from likes inner join pho...	Error Code: 1052. Column 'user_id' in field list is ambiguous	0.000 sec
27	14:13:16	use ig_clone	0 row(s) affected	0.000 sec
28	14:13:16	select user_id, photo_id, count(user_id) as Number_of_likes from likes group by pho...	257 row(s) returned	0.000 sec / 0.000 sec
29	14:14:27	use ig_clone	0 row(s) affected	0.000 sec
30	14:14:27	select username, photo_id, count(user_id) as Number_of_likes from likes inner join u...	257 row(s) returned	0.031 sec / 0.000 sec

## Solution to Question 4

The top 5 most commonly used hashtags on the platform are: #smile, #beach, #party, #fun and #concert

The screenshot shows the MySQL Workbench interface with a query editor and a result grid. The query is as follows:

```
4 where tags.id = photo_tags.tag_id
5 group by tag_name
6 order by Count_of_Tags
7 desc
8 limit 5
```

The result grid displays the top 5 most commonly used hashtags:

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

The bottom panel shows the Action Output with the following messages:

#	Time	Action	Message	Duration / Fetch
53	12:53:17	use ig_clone	0 row(s) affected	0.000 sec
54	12:53:17	select tag_name, count(photo_id) as Count_of_Tags from photo_tags inner join tag...	Error Code: 1064. You have an error in your SQL syntax; check the manual that cor...	0.000 sec
55	12:53:23	use ig_clone	0 row(s) affected	0.000 sec
56	12:53:23	select tag_name, count(photo_id) as Count_of_Tags from photo_tags inner join tag...	21 row(s) returned	0.016 sec / 0.000 sec
57	12:53:45	use ig_clone	0 row(s) affected	0.000 sec
58	12:53:45	select tag_name, count(photo_id) as Count_of_Tags from photo_tags inner join tag...	5 row(s) returned	0.000 sec / 0.000 sec

## Solution to Question 5

Total number of users on Instagram, total number of posts on Instagram and number of times an average user posts on Instagram

numberOfUsers	numberOfPosts	numberOfTimesAverageUserPosts
283	257	0.9081

The screenshot shows the MySQL Workbench interface. The main window displays a query in the SQL editor:

```
1 • use ig_clone;
2 • select user_id, count(image_url) as Total_Number_of_Post from photos
3 • group by user_id
```

The query results are shown in the Result Grid:

user_id	Total_Number_of_Post
1	5
2	4
3	4
4	3
6	5
8	4

The Output pane shows the execution log:

#	Time	Action	Message	Duration / Fetch
85	13:02:35	use ig_clone	0 row(s) affected	0.000 sec
86	13:02:35	select user_id, count(image_url), avg(Total_Number_of_Post) as Total_Number_of_...	Error Code: 1054. Unknown column 'Total_Number_of_Post' in field list	0.000 sec
87	13:03:17	use ig_clone	0 row(s) affected	0.000 sec
88	13:03:17	select user_id, count(image_url), sum(Total_Number_of_Post)/count(user_id) as Tot...	Error Code: 1054. Unknown column 'Total_Number_of_Post' in field list	0.015 sec
89	13:03:33	use ig_clone	0 row(s) affected	0.000 sec
90	13:03:33	select user_id, count(image_url) as Total_Number_of_Post from photos group by use...	74 row(s) returned	0.000 sec / 0.000 sec

## Solution to Question 6

Users (bots) who have liked every single photo on the platform

The screenshot shows the MySQL Workbench interface. The SQL editor contains the following query:

```
1 use ig_clone;
2 select user_id, username, count(photo_id) as Number_of_Post_Liked from likes
3 inner join users
4 where users.id = likes.user_id
5 group by user_id
6 having Number_of_Post_Liked = 767
```

The Results grid shows the following data:

user_id	username	Number_of_Post_Liked
5	Aniya_Hackett	257
14	Jaclyn81	257
21	Rocio33	257
24	Maxwell.Halvors on	257
36	Ollie_Ledner37	257
41	Mckenna17	257

The Output tab shows the execution log:

#	Time	Action	Message	Duration / Fetch
107	13:10:09	use ig_clone	0 row(s) affected	0.000 sec
108	13:10:09	select user_id, count(photo_id) as Number_of_Post_Liked from likes group by user_...	13 row(s) returned	0.016 sec / 0.000 sec
109	13:11:21	use ig_clone	0 row(s) affected	0.000 sec
110	13:11:21	select user_id, username, count(photo_id) as Number_of_Post_Liked from likes inn...	0 row(s) returned	0.672 sec / 0.000 sec
111	13:11:49	use ig_clone	0 row(s) affected	0.000 sec
112	13:11:49	select user_id, username, count(photo_id) as Number_of_Post_Liked from likes inn...	13 row(s) returned	0.032 sec / 0.000 sec

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

### Solution to Question 7

Best day to launch an AD

dayOfTheWeek	numberOfAccountsRegistered
5	16
1	16

### Result

The project turned out to be a success as all the problems were analysed and the expected solutions were extracted from the database.