



INSTAGRAM USERS ANALYSIS

Praveen Joshi
Praveenjoshi899@gmail.com



Instagram User Analysis

Project Description

This Purpose of this project is to understand the behaviour of Instagram users and Based on that derive valuable insights for various decision-making.

Approach

The main objective was to understand how they interact with our product g(Instagram). To do this, I analysed data from the database provided. By understanding our users' needs, we can create a product that is tailored to their needs and make sure that it meets all their expectations.

Tech-Stack Used

I used Mysql workbench 8.0.33 for this project.

Conclusion

All the conclusions are provided at the end of the document.



Instagram User Analysis

THE PROBLEM

A) Marketing: The marketing team wants to launch some campaigns, and they need your help with the following

1. Rewarding Most Loyal Users: People who have been using the platform for the longest time.

Your Task: Find the 5 oldest users of Instagram from the database provided

2. Remind Inactive Users to Start Posting: By sending them promotional emails to post their 1st photo.

Your Task: Find the users who have never posted a single photo on Instagram

3. Declaring Contest Winner: The team started a contest and the user who gets the most likes on a single photo will win the contest now they wish to declare the winner.

Your Task: Identify the winner of the contest and provide their details to the team

4. Hashtag Researching: A partner brand wants to know, which hashtags to use in the post to reach the most people on the platform.

Your Task: Identify and suggest the top 5 most commonly used hashtags on the platform

5. Launch AD Campaign: The team wants to know, which day would be the best day to launch ADs.

Your Task: What day of the week do most users register on? Provide insights on when to schedule an ad campaign

B) Investor Metrics: Our investors want to know if Instagram is performing well and is not becoming redundant like Facebook, they want to assess the app on the following grounds

1. User Engagement: Are users still as active and post on Instagram or they are making fewer posts



Instagram User Analysis

Your Task: Provide how many times does average user posts on Instagram. Also, provide the total number of photos on Instagram/total number of users

2. Bots & Fake Accounts: The investors want to know if the platform is crowded with fake and dummy accounts

Your Task: Provide data on users (bots) who have liked every single photo on the site (since any normal user would not be able to do this).



Instagram User Analysis

A) MARKETING

FINDINGS - 1

To find the most loyal , the top 5 oldest users of Instagram:-

1. We will use the user table data by selecting the username and created_at columns.
2. Then using the order by function we will order the desired output by sorting with the created column in ascending order.
3. Then using the limit function, the output will be displayed for the top 5 oldest Instagram users.

Input

```
SELECT * FROM users ORDER BY created_at ASC LIMIT 5;
```

Output

username	created_at
Darby_Herzog	06-05-2016 00:14
Emilio_Bernier52	06-05-2016 13:04
Elenor88	08-05-2016 01:30
Nicole71	09-05-2016 17:30
Jordyn.Jacobson2	14-05-2016 07:56

*Above Data table was exported to Excel from MYSQL....



Instagram User Analysis

FINDINGS -2

To Find the most inactive users. The users who have never posted any photo on Instagram:

1. We will first select the username column from the users table.
2. Then we will left join photos table on the users table, on users.id = photos.user_id because, both the users.id and photos.user_id have common contents in them.
3. Then we will find rows from the users table where the photos.id ISNULL

Input

```
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;
```

Output

username	user_id
Aniya_Hackett	5
Kasandra_Homenick	7
Jaclyn81	14
Rocio33	21
Maxwell.Halvorson	24
Tierra. Trantow	25
Pearl7	34
Ollie_Ledner37	36
Mckenna17	41



Instagram User Analysis

David.Osinski47	45
Morgan. Kassulke	49
Linnea59	53
Duane60	54
Julien_Schmidt	57

Mike.Auer39	66
Franco_Keebler64	68
Nia_Haag	71
Hulda.Macejkovic	74
Leslie67	75
Janelle.Nikolaus81	76
Darby_Herzog	80
Esther.Zulauf61	81
Bartholome.Bernhard	83
Jessyca_West	89
Esmeralda.Mraz57	90
Bethany20	91

Mike.Auer39	66
Franco_Keebler64	68
Nia_Haag	71
Hulda.Macejkovic	74
Leslie67	75
Janelle.Nikolaus81	76
Darby_Herzog	80
Esther.Zulauf61	81
Bartholome.Bernhard	83
Jessyca_West	89
Esmeralda.Mraz57	90
Bethany20	91



Instagram User Analysis

FINDINGS - 3

To find the most likes on a single photo(will be :

- FIRST WE WILL SELECT THE USERS.USERNAME, PHOTOS.ID, PHOTOS.IMAGE_URL AND COUNT(*) AS TOTAL
- THEN, WE WILL INNER JOIN THE THREE TABLES WIZ : PHOTOS, LIKES AND USERS, ON LIKES.PHOTO_ID = PHOTOS.ID AND PHOTOS.USER_ID = USERS.ID
- THEN, BY USING GROUP BY FUNCTION WE WILL GROUP THE OUTPUT ON THE BASIS OF PHOTOS.ID
- THEN, USING ORDER BY FUNCTION WE WILL SORTING THE DATA ON THE BASIS OF THE TOTAL IN DESCENDING ORDER
- THEN, TO FIND THE MOST LIKED PHOTO WE WILL USING LIMIT FUNCTION TO VIEW ONLY THE TOP LIKED PHOTO'S INFORMATION.

Input

```
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;
```

Output

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

*Above Data table was exported to Excel from MYSQL....



Instagram User Analysis

FINDINGS – 4

TO FIND THE TOP 5 MOST COMMONLY USED HASHTAGS ON INSTAGRAM:

1. WE NEED TO SELECT THE TAG_NAME COLUMN FROM THE TAG TABLE AND THE COUNT(*) AS TOTAL FUNCTION SO AS TO COUNT THE NUMBER OF TAGS USED INDIVIDUALLY.
2. THEN, WE NEED TO JOIN TAGS TABLE AND PHOTO_TAGS TABLE, ON TAGS.ID = PHOTO_TAGS.TAG_ID CAUSE THEY CONTAIN THE SAME CONTENT IN THEM I.E. TAG_ID.
3. THEN USING THE GROUP BY FUNCTION WE NEED TO GROUP THE DESIRED OUTPUT ON THE BASIS OF TAGS.TAG_NAME.
4. THEN USING THE ORDER BY FUNCTION WE NEED TO SORT THE OUTPUT ON THE BASIS OF TOTAL(TOTAL NUMBER OF TAGS PER TAG_NAME) IN DESCENDING ORDER.
5. THEN, TO FIND THE TOP 5 MOST USED TAG NAMES WE WILL USE THE LIMIT 5 FUNCTION.

Inputs

```
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;
```

Output:-

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



Instagram User Analysis

FINDINGS - 5

To find the day of the week on which most users register on Instagram and which day would be the best day to launch ADs.:

1. FIRST WE DEFINE THE COLUMNS OF THE DESIRED OUTPUT TABLE USING SELECT
DAYNAME(CREATED_AT) AS DAY_OF_WEEK AND COUNT(*) AS
TOTAL_NUMBER_OF_USERS_REGISTERED FROM THE USERS TABLE.
2. THEN USING THE GROUP BY FUNCTION WE GROUP THE OUTPUT TABLE ON THE BASIS OF
DAY_OF_WEEK
3. THEN USING THE ORDER BY FUNCTION WE ORDER/SORT THE OUTPUT TABLE ON THE BASIS OF
TOTAL_NUMBER_OF_USERS_REGISTERED IN DESCENDING ORDER

Inputs

```
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;
```

Output

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



Instagram User Analysis

B) Investor Metrics:

1. USER ENGAGEMENT

TO FIND THE HOW MANY TIMES DOES AVERAGE POSTS ON INSTAGRAM:

1. FIRST, WE NEED TO FIND FIRST THE COUNT NUMBER OF PHOTOS(POSTS) THAT ARE PRESENT IN THE PHOTOS.ID COLUMN OF THE PHOTOS TABLE I.E. COUNT(*) FROM PHOTOS.
2. SIMILARLY, WE NEED TO FIND THE NUMBER OF USERS THAT ARE PRESENT IN THE USERS.ID COLUMN OF THE USERS TABLE I.E. COUNT(*) FROM USERS.
3. NEXT, WE NEED TO DIVIDE BOTH THE VALUES I.E. COUNT(*) FROM PHOTOS/COUNT(*) FROM USERS AND HENCE WE WOULD GET THE TOTAL NUMBER OF PHOTOS / TOTAL NUMBER OF USERS.
4. TO FIND HOW MANY TIMES THE USERS POSTS ON INSTAGRAM WE NEED TO FIND THE TOTAL OCCURRENCES OF EACH USER_ID IN PHOTOS TABLE.

Input

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

Output

total_photos_divide_total_photos
2.57



Instagram User Analysis



Instagram User Analysis

user_id	user_post_count
1	5
2	4
3	4
4	3
6	5
8	4
9	4
10	3
11	5
12	4
13	5
15	4
16	4
17	3
18	1
19	2
20	1
22	1
23	12
26	5
27	1
28	4
29	8
30	2
31	1
32	4
33	5
35	2
37	1
38	2
39	1
40	1
42	3
43	5

44	4
46	4
47	5
48	1
50	3
51	5
52	5
55	1
56	1
58	8
59	10
60	2
61	1
62	2
63	4
64	5
65	5
67	3
69	1
70	1
72	5
73	1
77	6
78	5
79	1
82	2
84	2
85	2
86	9
87	4
88	11

92	3
93	2
94	1
95	2
96	3
97	2
98	1
99	3
100	2

2. Bot/Fake Accounts

TO FIND THE BOTS AND FAKE ACCOUNTS :-

1. FIRST, WE SELECT THE USER_ID COLUMN FROM THE PHOTOS TABLE.
2. THEN WE SELECT THE USERNAME COLUMN FROM THE USERS TABLE.
3. THEN, WE SELECT THE COUNT(*) FUNCTION TO COUNT TOTAL NUMBER OF LIKES FROM THE LIKES TABLE.
4. THEN WE INNER JOIN USERS AND LIKES TABLE ON THE BASIS OF USERS.ID AND LIKES.USER_ID, USING THE ON FUNCTION/CLAUSE.
5. THEN BY USING THE GROUP BY FUNCTION WE GROUP THE DESIRED OUTPUT TABLE ON THE BASIS OF LIKES.USER_ID.
6. THEN, WE SEARCH FOR THE VALUES FROM THE COUNT(*) FROM PHOTOS HAVING EQUAL VALUES WITH THE TOTAL_LIKES_PER_USER.

Inputs

```
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
);
```

Output

user_id	username	total_likes_per_user
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



INSTAGRAM USER ANALYSIS

ANALYSIS

AFTER PERFORMING THE ANALYSIS I HAVE THE



**FOLLOWING POINTS:- THE MOST LOYAL USERS THE
TOP 5 OLDEST USERS ARE:**

username	created_at
Darby_Herzog	06-05-2016 00:14
Emilio_Bernier52	06-05-2016 13:04
Elenor88	08-05-2016 01:30
Nicole71	09-05-2016 17:30
Jordyn.Jacobson2	14-05-2016 07:56



OUT OF THE 100 TOTAL USERS, THERE ARE 26 USERS WHO ARE INACTIVE
AND THEY HAVE NEVER POSTED ANY KIND OF STUFF ON INSTAGRAM MAY IT
BE ANY PHOTO, VIDEO OR ANY TYPE OF TEXT. SO, THE MARKETING TEAM OF
INSTAGRAM NEEDS TO REMIND SUCH INACTIVE USERS.



SO, THE USER NAMED ZACK_KEMMER93 WITH USER_ID 52 IS THE WINNER OF
THE CONTEST BECAUSE HIS PHOTO WITH PHOTO_ID 145 HAS THE HIGHEST
NUMBER OF LIKES I.E. 48.



THE TOP 5 MOST COMMONLY USED #HASHTAGS ALONG WITH THE TOTAL
COUNT ARE SMILE(59), BEACH(42), PARTY(39), FUN(38) AND CONCERT(24)



MOST OF THE USERS REGISTERED ON THURSDAY AND SUNDAY I.E. 16 AND
HENCE IT WOULD PROVE BENEFICIAL TO START AD CAMPAIGN ON THESE TWO
DAYS

◆ SO, THERE ARE IN TOTAL 257 ROWS I.E. 257 PHOTOS IN THE PHOTOS TABLE AND 100 ROWS I.E. 100 IDS IN THE USERS TABLE WHICH MAKE THE DESIRED OUTPUT TO BE $257/100 = 2.57$ (AVG. USERS POSTS ON INSTAGRAM)

◆ OUT OF THE TOTAL USER ID'S, THERE ARE 13 SUCH USER ID'S WHO HAVE LIKED EACH AND EVERY POST ON INSTAGRAM (WHICH IS NOT PRACTICALLY POSSIBLE) AND SO SUCH USER ID'S ARE CONSIDERED AS BOTS AND FAKE ACCOUNTS.

ANALYSIS

USING THE 5 WHYS APPROACH I AM FINDING THE ROOT CAUSE OF THE FOLLOWING:-

- **WHY DID THE MARKETING TEAM WANT TO KNOW THE MOST INACTIVE USERS?**

SO, THEY CAN REACH OUT TO THOSE USERS VIA MAIL AND ASK THEM WHAT'S KEEPING THEM AWAY FROM USING INSTAGRAM.

- **WHY DID THE MARKETING TEAM WANT TO KNOW THE TOP 5 #HASHTAGS USED?**

MAYBE THE TECH TEAM WANTED TO ADD SOME FILTER FEATURES FOR PHOTOS AND VIDEOS POSTED USING THE TOP 5 MENTIONED #HASHTAGS

- **WHY DID THE MARKETING TEAM WANT TO KNOW ON WHICH DAY OF THE WEEK THE PLATFORM HAD THE MOST NEW USERS REGISTERED?**

SO, THAT THEY CAN RUN MORE ADS OF VARIOUS BRANDS DURING SUCH DAYS AND ALSO GET PROFIT FROM IT

- ◆
- **WHY DID THE INVESTORS WANT TO KNOW ABOUT THE AVERAGE NUMBER OF POSTS PER USER ON INSTAGRAM?**

IT IS A FACT THAT EVERY BRAND OR SOCIAL PLATFORM IS DETERMINED BY THE USER ENGAGEMENT ON SUCH PLATFORMS, ALSO INVESTORS WANTED TO KNOW WHETHER THE PLATFORM HAS THE RIGHT AND AUTHENTICATED USER BASE. IT ALSO HELPS THE TECH TEAM DETERMINE HOW TO HANDLE SUCH TRAFFIC ON THE PLATFORM WITH

THE LATEST TECH WITHOUT DISRUPTING THE SMOOTH AND EFFICIENT FUNCTIONING OF THE PLATFORM

- **WHY DID THE INVESTORS WANT TO KNOW THE COUNT OF BOTS AND FAKE ACCOUNTS IF ANY?**

SO THAT THE INVESTORS ARE ASSURED THAT THEY ARE INVESTING IN AN ASSET AND NOT A FUTURE LIABILITY

CONCLUSION

In conclusion, I would like to conclude that such Analysis to find insights from their customer data which in turn helps the firms to find customers who will be an asset to the firm in the future and not some Liability not only on Instagram but many other social media and commercial firms uses similar analytics to keep track of the user behaviour and activities. we identified trends in user behaviour that could be used to improve user experience, such as providing helpful feedback loops or streamlining processes for specific tasks. Overall, our user analysis project has provided us with valuable insights that can help us create better products in the future.

◆