
Instagram User
Analytics

Project Description

- In this project scenario, my task is to analyze the Instagram users based on the questions asked by the management team and the product manager to provide valuable business insights for marketing, product & development teams which helps in developing a product.

- And these insights will be effective for the teams across the business to launch a new marketing campaign and it will help them decide what features will satisfy the customers to include in an app. Such that success rate of the app can be tracked easily by measuring user engagement so as to improve the user experience.

- As a Data Analyst, I'm going to create a database named "ig_clone" given which contains of many tables and query the required details for the product manager.

Project Description

In this project, I'm expected to provide a detailed report by covering the following tasks:

Marketing:

- Task 1 – To find the oldest users of the Instagram.
- Task 2 – To find the users who have never posted a single photo on Instagram.
- Task 3 – To find the user who have got the most likes on a single photo.
- Task 4 – To suggest top 5 commonly used hashtags.
- Task 5 – To suggest an optimal day to launch the AD campaign.

Investor Metrics:

- Task 6 – To provide how many times does average user posts on Instagram and the total number of photos / users.
- Task 7 – To provide data on users who have liked every single photo (termed as bots).

Project Approach

- For this project, SQL was used necessarily to create a database as provided in the dashboard. With the help of the created database, data extraction and manipulations queries were performed to obtain the required insights by the product manager

Tech Stack Used

MySQL Workbench v8.0.30.0 was used in this project to query the database.

Purpose:
The ease of access to use and query the database.

Insights

A) Marketing

1. Rewarding Most Loyal Users: Top 5 People who have been using the platform the longest time

Query:

```
SELECT *
FROM 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

Conclusion:

Users 80, 67, 63, 95, 38 are the top 5 people who have been using Instagram for the longest time.

2. Remind Inactive Users to Start Posting: Users who have never posted a single photo on Instagram.

Query:

```
SELECT ID, USERNAME
FROM USERS
WHERE (USERNAME NOT IN
(SELECT USERNAME FROM USERS
JOIN PHOTOS
ON USERS.ID = PHOTOS.USER_ID))
```

Result - >

ID	USERNAME
5	Aniya_Hackett
7	Kasandra_Homenick
14	Jaclyn81
21	Rocio33
24	Maxwell.Halvorson
25	Tierra.Trantow
34	Pearl7
36	Ollie_Ledner37
41	Mckenna17
45	David.Osinski47
49	Morgan.Kassulke
53	Linnea59
54	Duane60
57	Julien_Schmidt
66	Mike.Auer39
68	Franco_Keebler64
71	Nia_Haag
74	Hulda.Macejkovic
75	Leslie67
76	Janelle.Nikolaus81
80	Darby_Herzog
81	Esther.Zulauf61
83	Bartholome.Bernhard
89	Jessyca_West
90	Esmeralda.Mraz57
91	Bethany20

Conclusion:

It seems that about 26% (26 out of 100) of users have never posted a single photo on Instagram.

3. Declaring Contest Winner: Identify the winner who got the most likes on a single photo on Instagram.

Query:

```
SELECT P.USER_ID, U.USERNAME, PHOTO_ID, COUNT(L.USER_ID) AS NUM_OF_LIKES
FROM PHOTOS AS P
JOIN LIKES AS L
ON L.PHOTO_ID = P.ID
JOIN USERS AS U
ON P.USER_ID = U.ID
GROUP BY PHOTO_ID
ORDER BY COUNT(USER_ID) DESC
LIMIT 1;
```

Result:

USER_ID	USERNAME	PHOTO_ID	NUM_OF_LIKES
52	Zack_Kemmer93	145	48

Conclusion:

The Winner of the contest is Zack_Kemmer93 (ID: 52) with 48 no of likes on photo ID: 145 !!!

4. Hashtag Researching: Identify top 5 commonly used hashtags.

Query:

```
SELECT TAG_ID, TAG_NAME AS HASHTAG, COUNT(PHOTO_ID) AS COUNTS
FROM TAGS AS T
JOIN PHOTO_TAGS AS P
ON T.ID = P.TAG_ID
GROUP BY TAG_ID
ORDER BY COUNTS DESC
LIMIT 5;
```

Result:

TAG_ID	HASHTAG	COUNTS
21	smile	59
20	beach	42
17	party	39
13	fun	38
18	concert	24

Conclusion:

The above listed hashtags were used more commonly in many posts by more people.
Hence these tags can be used for wider reach of audience.

5. Launch AD campaign: Find which day of the week do most users register on.

Query:

```
-- From User Interactions such as Likes, Follows, Posts, Comments.
-- Here I've queried the day for posts were created
SELECT DAYNAME(DATE(CREATED_DAT)) AS DAY
FROM PHOTOS
LIMIT 1;

-- Query for the common day where most users were created
SELECT DAYNAME(DATE(CREATED_AT)) AS DAY
FROM USERS
GROUP BY DAYNAME(DATE(CREATED_AT))
ORDER BY COUNT(DAYNAME(DATE(CREATED_AT))) DESC
LIMIT 1;
```

Result:

DAY
Thursday

Conclusion:

From the common user interactions such as Likes, Comments, Posts were often created on Thursdays. Most users were also created on Thursdays. Thus, it is optimal to schedule an Ads campaign on Thursdays.

B) Investor Metrics

1. User Engagement: Provide how many times does an average user posts on Instagram. Also, provide the total number of photos on Instagram/total number of users.

Queries:

```
-- Number of posts by an average user
SELECT ROUND(AVG(C)) AS NO_OF_POSTS_OF_AN_AVERAGE_USER
FROM (SELECT COUNT(ID) AS C FROM PHOTOS GROUP BY USER_ID) AS S;

--Total number of photos / total number of users.
SELECT COUNT(P.ID) AS NO_OF_PHOTOS_AND_USERS
FROM PHOTOS AS P
UNION
SELECT COUNT(U.ID)FROM USERS AS U
```

Results:

NO_OF_POSTS_OF_AN_AVERAGE_USER
3

NO_OF_PHOTOS_AND_USERS
257
100

Conclusion:

It is found that an average user posts about 3 times on Instagram.
And the total number of photos & users on Instagram has been queried from the database.

2. Bots & Fake Accounts: Provide data on users who have liked every single photo.

Query:

```
SELECT USERS.USERNAME AS BOTS, LIKES.USER_ID AS ID
FROM PHOTOS
JOIN LIKES
ON PHOTOS.ID = LIKES.PHOTO_ID
JOIN USERS
ON USERS.ID = LIKES.USER_ID
GROUP BY LIKES.USER_ID
HAVING COUNT(PHOTOS.ID) = (SELECT COUNT(PHOTOS.ID) FROM PHOTOS)
```

Result:

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

Conclusion:

From the results, the above listed users are likely to be bots or fake accounts, because they've liked every single photo.

Result

- In this project, I've learned how to query and manipulate the data according to the questions asked.
- I've also gained the knowledge about how to create a database and insert tables and rows into it.
- This project has helped me in such a way that it feels little bit easier to work using MySQL because of lot of thinking, understanding, developing a solution, coding and practice.