
Instagram User Data Analysis Report

---

1. Project Description

This project focuses on analyzing Instagram user interactions and engagement to extract meaningful insights. These insights will assist the marketing and product teams in making data-driven decisions, helping to enhance user engagement, improve platform features, and support business growth.

The analysis will address specific tasks requested by the management team, categorized into **Marketing Analysis** and **Investor Metrics.**

---

2. Approach

The project was carried out in the following steps:

1. **Database Setup**:
    - Imported the provided database into MySQL Workbench.
    - Verified table structures and relationships to understand the schema and data flow.
2. **Analysis Execution**:
    - Utilized SQL queries to extract the necessary information.
    - Optimized queries to ensure efficient data retrieval.
    - Documented queries and results for transparency and reproducibility.
3. **Report Compilation**:
    - Collected outputs and visualized data where applicable.
    - Organized findings into a concise and actionable format.

---

3. Tech Stack Used

- **Database Management Tool**: MySQL Workbench
    - **Version**: (Mention the version, e.g., 8.0.33)
    - Reason for choice: User-friendly interface, robust query handling, and support for complex SQL operations.

4. Insights

A) Marketing Analysis

1. **Loyal User Reward**

Sql Query used;

SELECT, username, created_at FROM users ORDER BY created_at ASC LIMIT 5;

Following is top 5 loyal user reward:-

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 7:56:26

2. **Inactive User Engagement**

Sql Query used;

SELECT id, username, created_at FROM users WHERE id NOT IN (SELECT DISTINCT user_id FROM photos);

Following are the inactive user engagement data:-

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
74	Hulda.Macejkovic	2017-01-25 17:17:28
75	Leslie67	2016-09-21 05:14:01
76	Janelle.Nikolaus81	2016-07-21 09:26:09
80	Darby_Herzog	2016-05-06 00:14:21
81	Esther.Zulauf61	2017-01-14 17:02:34
83	Bartholome.Bernhard	2016-11-06 02:31:23
89	Jessyca_West	2016-09-14 23:47:05
90	Esmeralda.Mraz57	2017-03-03 11:52:27
91	Bethany20	2016-06-03 23:31:53

3. Contest Winner Declaration

SQL Query used

SELECT user_id, photo_id, MAX(likes_count) AS max_likes FROM photos GROUP BY photo_id ORDER BY max_likes DESC LIMIT 1;

Photo_id	Maximum interactions
145	48

SQL Query;

SELECT user_id

FROM photos

WHERE id = 145;

===== Got id = 52

Again SQL Query;

select username from users

where id = 52;

===== Got name == **Zack_Kemmer93 is Identified as the contest winner with the highest likes on a single photo.**

4. **Hashtag Research**

SQL Query Used:

```
SELECT     tag_id,
   COUNT(tag_id) AS tag_count
FROM
   photo_tags
   GROUP BY
   tag_id
ORDER BY
   tag_count DESC
   Limit 5;
```

SQl Query to find tag names from tags table

```
SELECT tag_name
FROM tags
WHERE id IN (21, 20, 17, 13, 18);
```

Top 5 tags

Tag_id	Tag_count	Tag_name
21	59	smile
20	42	beach
17	39	Party
13	38	Fun
18	24	Concert

Top 5 tags are smile, beach, party, fun, concert.

5. **Ad Campaign Launch**

- **SQL Query Used**:

```
SELECT
   dayname(created_at) AS registration_date,
   COUNT(*) AS registration_count
FROM
   users
GROUP BY
   registration_date
ORDER BY
   registration_count DESC
LIMIT 1;
```

Registeration date	Regidteration count
Thursday	16

Insight: The most common day for user registration to schedule ad campaigns effectively in a week is Thursday.

B) Investor Metrics

1. **User Engagement**

SQL Query Used:

```
select
   AVG(photo_count) AS avg_posts_per_user,   -- Average number of posts per user
   COUNT(DISTINCT user_id) AS total_users,  -- Total users
   SUM(photo_count) AS total_photos         -- Total number of photos
```

```
FROM (
   -- Subquery: Count photos per user
   SELECT
      user_id,
      COUNT(*) AS photo_count
   FROM
      photos
   GROUP BY
      user_id
) AS photo_stats;
```

avg_posts_per_user	total_users	total_photos
3.4730	74	257

The average number of posts per user is 3.4730.

The total number of users is 74.

The total number of photos is 257.

2. **Bots & Fake Accounts**

SQL Query Used:

```
SELECT user_id, COUNT(*) AS total_likes
FROM likes
GROUP BY user_id
HAVING total_likes = (SELECT COUNT(*) FROM photos);
```

Following users (potential bots) who have liked every single photo on the site, as this is not typically possible for a normal user.

User_id	Total likes
5	257
14	257
21	257
24	257
36	257
41	257
54	257
57	257
66	257
71	257
75	257
76	257
91	257

5. Results

- Successfully identified user patterns, engagement levels, and potential anomalies.
- Insights provided actionable recommendations for marketing strategies, feature improvements, and operational adjustments.
- This project demonstrates how data-driven analysis can support strategic decisions and foster business growth.