

# Instagram User Analytics

## Project Description:

In this project analysis on engagement and interaction of users with the digital product is performed to get insights from the given data to help the business to grow.

I am using mysql to perform the required analysis on the given data.

Here I am going to answer following questions of marketing team and investors:

- 1) Most Loyal Users
- 2) Inactive Users
- 3) Contest Winner
- 4) Most used Hashtags
- 5) Best day to launch ADs
- 6) Users Engagement
- 7) Bots and fake Accounts

**Approach:** I am going to write queries to fetch data from the database to get insights and to answer the questions and that will help the business to grow.

**Tech Stack Used:** MySQL Workbench

## Questions asked & Insights:

- 1) **Rewarding Most Loyal Users:** Find the 5 oldest users of Instagram from the database provided.

**Approach:**

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

**Solution:** These are the five oldest users of instagram from the provided database

	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) **Remind Inactive Users to Start Posting:** Find the users who have never posted a single photo on Instagram.

**Approach:**

```
SELECT * FROM users
WHERE id NOT IN (SELECT user_id FROM photos)
```

**Solution:** The image of the table given below provide the details of users who have never posted a single photo on Instagram.

	id	username	created_at				
▶	5	Aniya_Hackett	2016-12-07 01:04:39		66	Mike.Auer39	2016-07-01 17:36:15
	7	Kassandra_Homenick	2016-12-12 06:50:08		68	Franco_Keebler64	2016-11-13 20:09:27
	14	Jadyn81	2017-02-06 23:29:16		71	Nia_Haag	2016-05-14 15:38:50
	21	Rocio33	2017-01-23 11:51:15		74	Hulda.Macejkovic	2017-01-25 17:17:28
	24	Maxwell.Halvorson	2017-04-18 02:32:44		75	Leslie67	2016-09-21 05:14:01
	25	Tierra.Trantow	2016-10-03 12:49:21		76	Janelle.Nikolaus81	2016-07-21 09:26:09
	34	Pearl7	2016-07-08 21:42:01		80	Darby_Herzog	2016-05-06 00:14:21
	36	Ollie_Ledner37	2016-08-04 15:42:20		81	Esther.Zulauf61	2017-01-14 17:02:34
	41	Mckenna17	2016-07-17 17:25:45		83	Bartholome.Bernhard	2016-11-06 02:31:23
	45	David.Osinski47	2017-02-05 21:23:37		89	Jessyca_West	2016-09-14 23:47:05
	49	Morgan.Kassulke	2016-10-30 12:42:31		90	Esmeralda.Mraz57	2017-03-03 11:52:27
	53	Linnea59	2017-02-07 07:49:34		91	Bethany20	2016-06-03 23:31:53
	54	Duane60	2016-12-21 04:43:38				
	57	Julien_Schmidt	2017-02-02 23:12:48				

- 3) **Declaring Contest Winner:** Identify the winner of the contest and provide their details to the team.

**Approach:** To find winner

```
SELECT * FROM (SELECT COUNT(user_id) Total_Likes, photo_id FROM likes
GROUP BY photo_id ) win ORDER BY Total_Likes DESC LIMIT 1
```

To find details of winner

```
SELECT * FROM users
WHERE id IN (SELECT user_id FROM photos
WHERE id = 145)
```

**Solution:** The first query helps us to find the photo id of the photo which got the maximum likes i.e, a photo with photo id 145 got 48 likes which is maximum. So the user who posted that photo is the winner of the contest.

Then the second query is to fetch the details of the user who wins the contest.

	Total_Likes	photo_id		id	username	created_at
▶	48	145	▶	52	Zack_Kemmer93	2017-01-01 05:58:22
			*	NULL	NULL	NULL

- 4) **Hashtag Researching:** Identify and suggest the top 5 most commonly used hashtags on the platform

**Approach:**

```
SELECT * FROM (SELECT COUNT(photo_id) Times_Used, tag_id FROM photo_tags  
GROUP BY tag_id ) MOST ORDER BY Times_Used DESC LIMIT 5
```

**Solution:** Below the given table shows the id of top 5 most used tags with the number of times it is used.

	Times_Used	tag_id
▶	59	21
	42	20
	39	17
	38	13
	24	18

- 5) **Launch AD Campaign:** What day of the week do most users register on? Provide insights on when to schedule an ad campaign

**Approach:**

```
SELECT COUNT(id) number_of_registration, day FROM (SELECT id,  
DATE_FORMAT(cast(created_at AS date), '%a') day, cast(created_at AS Time) time from users) temp  
GROUP BY day ORDER BY number_of_registration DESC LIMIT 5
```

**Solution:** The table below shows the day when the maximum registration done.

So to launch or schedule an ad campaign the best days of week are Thursday and Sunday.

	number_of_registration	day
▶	16	Thu
	16	Sun
	15	Fri
	14	Tue
	14	Mon

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

**Approach:** To find total number of users and total posts(i.e, photos)

```
SELECT SUM(posts) total_posts, COUNT(user_id) total_users
FROM (SELECT COUNT(id) posts, user_id FROM photos
GROUP BY user_id) sums1
```

To find how many times an average user posts

```
SELECT posts, COUNT(user_id) users FROM
(SELECT COUNT(id) posts, user_id FROM photos
GROUP BY user_id ORDER BY posts) table1
GROUP BY posts ORDER BY users DESC
```

To find the total number of photos on Instagram/total number of users.

```
SELECT total_posts/total_users FROM
(SELECT SUM(posts) total_posts, COUNT(user_id) total_users FROM
(SELECT COUNT(id) posts, user_id FROM photos
GROUP BY user_id) sums1)final
```

**Solution:** The first table gives the information that how many users and posts are there. There are 257 posts and 74 users on the platform.

The second table gives the details of how many users posted how many times. From this table we got that an average user posts somewhere between 1 to 5 photos.

The third table gives information of the total number of photos on Instagram/total number of users which is approx 3.

	total_posts	total_users
▶	257	74

	total_posts/total_users
▶	3.4730

	posts	users
▶	1	18
	5	14
	2	13
	4	13
	3	9
		13

- 7) **Bots & Fake Accounts:** 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).

**Approach:**

```
SELECT * FROM users
WHERE id IN (SELECT user_id FROM (SELECT * FROM (SELECT user_id, COUNT(photo_id) total FROM likes
GROUP BY user_id ORDER BY total DESC) tab1
WHERE total = (SELECT COUNT(DISTINCT(photo_id)) FROM likes))tab2)
```

**Solution:** Below table gives details about the users who seems to be bots because they've liked every single photo on the site.

	id	username	created_at
▶	5	Aniya_Hackett	2016-12-07 01:04:39
	14	Jadyn81	2017-02-06 23:29:16
	21	Rocio33	2017-01-23 11:51:15
	24	Maxwell.Halvorson	2017-04-18 02:32:44
	36	Ollie_Ledner37	2016-08-04 15:42:20
	41	Mckenna17	2016-07-17 17:25:45
	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
	71	Nia_Haag	2016-05-14 15:38:50
	75	Leslie67	2016-09-21 05:14:01
	76	Janelle.Nikolaus81	2016-07-21 09:26:09
	91	Bethany20	2016-06-03 23:31:53
	HULL	HULL	HULL

**ACCORDING TO THE DATABASE PROVIDED:**

- The site has a total 74 users and a total 257 posts.
- Darby\_herzog is the oldest user of the site. He registered himself on 06-05-2016.
- There are 26 total users who have never posted on the site.
- The photo with id 145 got the maximum likes. Zack\_kemmer93 posted this so he is the winner of the contest.
- Tag with the tag id 21 is the most used tag on the site(59 times).
- Most of the registration on the site is done on Thursday and sunday. So these two days of the week are best to schedule ad campaigns.
- Average user posts are somewhere between 1 to 5 posts.
- 13 accounts may be fake or bots because they've liked every post on the site.

**Result:**

I've learned concepts of mysql database and implemented it in this project. I learned a lot by researching on queries I've used to get the results and insights. It helped me to complete the project and build a strong foundation of mysql databases. It may also help me in my future goal of becoming a Data Analyst.