Instagram User Analytics
SQL Fundamentals

Project Description
In this project user analysis will be done understand the engagement of user with Instagram. With this analysis, insights will be derived for marketing, product and development teams which will help Instagram to launch a new marketing campaign, decide on features to build for an app, track the success of the app by measuring user engagement and improve the experience altogether while helping the business grow.

Approach
Using existing database of Instagram users, SQL queries will be run to user analysis

Tech-Stack Used: MySQL Tutorial
- MySQL is a widely used relational database management system.
- It is free and open-source.
- It is ideal for both small and large application.
- It is very easy to write query in MySQL
- MySQL uses a standard form of the well-known SQL data language.
- MySQL works very quickly and works well even with large data sets.

Insights

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.

   Given list contain the names of top 5 oldest users of the Instagram:

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

   They are the most loyal user of Instagram whom will reward.

   Query used:

```
SELECT  * FROM ig_clone.users
order by created_at
limit 5;
```

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

   List of inactive Instagram user who haven't posted even once on Instagram are as follow

id	username	NoOfImgPosted
5	Aniya_Hackett	0
7	Kasandra_Homenick	0
14	Jaclyn81	0
21	Rocio33	0
24	Maxwell.Halvorson	0
25	Tierra.Trantow	0
34	Pearl7	0
36	Ollie_Ledner37	0
41	Mckenna17	0
45	David.Osinski47	0
49	Morgan.Kassulke	0
53	Linnea59	0
54	Duane60	0
57	Julien_Schmidt	0
66	Mike.Auer39	0
68	Franco_Keebler64	0
71	Nia_Haag	0
74	Hulda.Macejkovic	0
75	Leslie67	0
76	Janelle.Nikolaus81	0
80	Darby_Herzog	0
81	Esther.Zulauf61	0
83	Bartholome.Bernhard	0
89	Jessyca_West	0
90	Esmeralda.Mraz57	0
91	Bethany20	0

   Query used:

```
SELECT users.id, users.username, count(image_url) as NoOfImgPosted
 FROM ig_clone.users left join ig_clone.photos
 on users.id = photos.user_id
 group by id
 order by NoOfImgPosted, id;
```

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.

   Instagram user with username Zack_Kemmer93, user id 52 is the winner of contest with the 48 likes on photo with photo id 145

username	user_id	photo_id	likes
Zack_Kemmer93	52	145	48

Query used:

```
SELECT users.username, photos.user_id, likes.photo_id, count(likes.user_id) as likes
FROM
ig_clone.users inner join ig_clone.photos
on users.id = photos.user_id
inner join ig_clone.likes
on photos.id = likes.photo_id
group by likes.photo_id
order by likes desc
limit 1;
```

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

   List of top 5 most commonly used Instagram hashtag as follow:

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

Query used:

```
select tags.tag_name, count(photo_tags.photo_id) as UsedTimes
from  ig_clone.tags inner join ig_clone.photo_tags
on tags.id = photo_tags.tag_id
group by photo_tags.tag_id
order by UsedTimes desc
limit 5;
```

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

   Thursday and Sunday are the day on which most of user use Instagram so Thursday are the best days to launch AD Campaign

weekdays	NoOfDays
Thursday	16
Sunday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12

Query used:

```
SELECT
weekdays, count(weekdays) AS NoOfDays
FROM
(
    SELECT MAX(DAYNAME(CREATED_AT)) AS weekdays
    FROM USERS
    GROUP BY CREATED_AT
) t
GROUP BY weekdays
ORDER BY NoOfDays DESC;
```

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.

- Total no of photo on Instagram: 257
- Total no users: 100
- Therefore on average a user post a photo 2.57 time on Instagram

AvgPost
2.5700

Query used:

```
SELECT
AVG(ImgPosts1) as AvgPost
FROM
(
    SELECT users.id, users.username,
           count(image_url) as ImgPosts, COUNT(image_url) AS ImgPosts1
    FROM ig_clone.users left join ig_clone.photos
    ON users.id = photos.user_id
    group by users.id
) t  ;
```

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

- There 257 photos on Instagram in total

List of the account with have liked each photos are as follow:

	BotAccount	user_id	Liked
▶	Rocio33	21	257
	Nia_Haag	71	257
	Aniya_Hackett	5	257
	Mike.Auer39	66	257
	Mckenna17	41	257
	Jadyn81	14	257
	Julien_Schmidt	57	257
	Maxwell.Halvorson	24	257
	Janelle.Nikolaus81	76	257
	Leslie67	75	257
	Duane60	54	257
	Bethany20	91	257
	Ollie_Ledner37	36	257

So we can say that these are the fake accounts which we found on Instagram.

Query used:

1.

```
SELECT count(distinct photo_id) FROM ig_clone.likes;
```

2.

```
SELECT users.username as BotAccount, user_id, count(photo_id) as Liked
FROM  ig_clone.users inner join ig_clone.likes
ON users.id = likes.user_id
group by user_id
order by Liked desc;
```

Result:

- In this project I have gain practical hands on knowledge about database and how to solve SQL queries.
- This will further help me to perform data analysis in real world scenarios.