INSTAGRAM USER ANALYTICS

PROJECT DESCRIPTION : This project is basically to analyse the raw data that is provided and run some queries based on questions asked by the management team to provide valuable insights. Various database management tools can be used for this project.

This project is about how the users engage and interact with Instagram. I will analyse these users in an attempt to derive business insights for marketing, product & development teams. These insights are then used by teams across the business to launch a new marketing campaign, decide on features to build for an app, track the success of the application by measuring user engagement and improve the experience altogether while helping the business grow.

APPROACH : Suppose, I am working with the product team of Instagram and the product manager has asked us to provide insights on the questions asked by the management team. I use SQL to derive different insights from the given dataset provided by the management team. Also, I will use various sorting and other queries for data extraction.

TECH-STACK USED : I am using MYSQL Workbench for this project because of its GUI. The left panel gives me access to all the tables in the datasets which makes it easy for all data analysis.

“I will mention the SQL queries along with the outputs in this file below. And also I will took snapshots of the SQL queries and will be attaching below the queries.”

INSIGHTS :

A.) MARKETING ANALYSIS :

1. LOYAL USER REWARD :

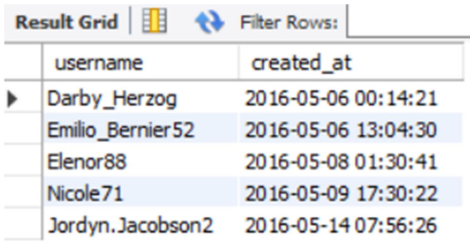
Identify the five oldest users on Instagram from the provided database.

QUERY :

SELECT username, created_at **FROM** users

ORDER BY created_at

LIMIT 5;



Result Grid | Filter Rows:

username	created_at
Darby_Herzog	2016-05-06 00:14:21
Emilio_Bernier52	2016-05-06 13:04:30
Elenor88	2016-05-08 01:30:41
Nicole71	2016-05-09 17:30:22
Jordyn.Jacobson2	2016-05-14 07:56:26

CONCLUSION : These users are 5 oldest users of Instagram.

2. INACTIVE USER ENGAGEMENT :

Identify users who have never posted a single photo on Instagram.

QUERY :

SELECT username **FROM** users **AS** u

LEFT JOIN photos **AS** p **ON** p.user_id=u.id

WHERE p.image_url **IS NULL**

ORDER BY u.username;

username
Aniya_Hackett
Bartholome.Bernhard
Bethany20
Darby_Herzog
David.Osinski47
Duane60
Esmeralda.Mraz57
Esther.Zulauf61
Franco_Keebler64
Hulda.Macejkovic
Jaclyn81
Janelle.Nikolaus81
Jessyca_West
Julien_Schmidt
Kasandra_Homenick
Leslie67
Linnea59
Maxwell.Halvorson
Mckenna17
Mike.Auer39
Morgan.Kassulke
Nia_Haag
Ollie_Ledner37
Pearl7
Rocio33
Tierra.Trantow

CONCLUSION : Above mentioned users didn't posted a single photo.

3. CONTEST WINNER DECLARATION :

Determine the winner of the contest and provide their details to the team.

QUERY :

```
SELECT likes.photo_id, users.username, COUNT(likes.user_id) AS no_of_likes
FROM likes
INNER JOIN photos ON likes.photo_id=photos.id
INNER JOIN users ON photos.user_id=users.id
GROUP BY likes.photo_id,users.username
ORDER BY no_of_likes DESC
LIMIT 1;
```

Result Grid | Filter Rows:

photo_id	username	no_of_likes
145	Zack_Kemmer93	48

CONCLUSION : Photo_ID 145 is the photo got most likes posted by username Zack_Kemmer93.

4. HASHTAG RESEARCH :

Identify and suggest the top five most commonly used hashtags on the platform.

QUERY :

```
SELECT t.tag_name, count(pt.photo_id) AS total_count
FROM photo_tags AS pt
INNER JOIN tags AS t ON t.id=pt.tag_id
GROUP BY t.tag_name
ORDER BY total_count DESC
LIMIT 5;
```

Result Grid | Filter Rows:

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

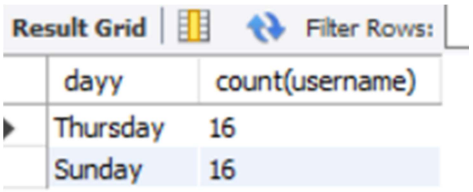
CONCLUSION : smile, beach, party, fun and concert are the most used hashtags on the platform.

5. AD CAMPAIGN LAUNCH :

Determine the day of the week when most users register on Instagram. Provide insights on when to schedule an ad campaign.

QUERY :

```
SELECT date_format((created_at),'%W') AS dayy, count(username)
FROM users
GROUP BY 1
ORDER BY 2 DESC
LIMIT 2;
```



dayy	count(username)
Thursday	16
Sunday	16

CONCLUSION : Most user registers on Thursday and Sunday.

B.) INVESTOR METRICS :

1. USER ENGAGEMENT :

Calculate the average number of posts per user on Instagram. Also, provide the total number of photos on Instagram divided by the total number of users.

QUERY :

```
WITH base AS(
SELECT u.id AS userid, COUNT(p.id) AS photoid FROM users AS u
LEFT JOIN photos AS p ON p.user_id =u.id
GROUP BY u.id)
SELECT SUM(photoid) AS total_photos, COUNT(userid) AS total_users,
SUM(photoid)/COUNT(userid) AS photo_per_user
FROM base;
```

Result Grid | Filter Rows:

total_photos	total_users	photo_per_user
257	100	2.5700

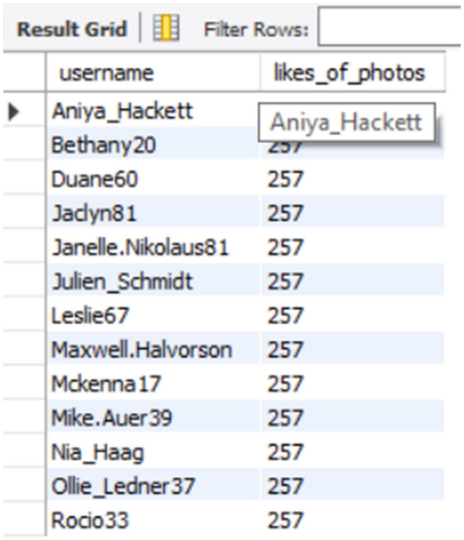
CONCLUSION : Average photos uploaded by user is 2.5700.

2. BOTS & FAKE ACCOUNT :

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

QUERY :

```
WITH base AS(
SELECT users.username, count(likes.photo_id) AS likes_of_photos FROM likes
INNER JOIN users ON users.id= likes.user_id
GROUP BY users.username)
SELECT username, likes_of_photos FROM base
WHERE likes_of_photos=(SELECT COUNT(*) FROM photos)
ORDER BY username;
```



Result Grid | Filter Rows:

username	likes_of_photos
Aniya_Hackett	257
Bethany20	257
Duane60	257
Jaclyn81	257
Janelle.Nikolaus81	257
Julien_Schmidt	257
Leslie67	257
Maxwell.Halvorson	257
Mckenna17	257
Mike.Auer39	257
Nia_Haag	257
Ollie_Ledner37	257
Rocio33	257

CONCLUSION : All the above mentioned users liked every single photos.

RESULT : Finally I would like to say learned quite a lot regarding JOINS and also loved that all the answers came correctly after many trials and errors. Learned new things also from this project and lastly the resources from trainity's learning materials helped a lot during this project.