Instagram User Analytics

- **K Vikram Sundar**

Project Description :

The project is based on Instagram User analytics where the goal is to analyse the given dataset and come up with insights so that the marketing team can make data-driven decisions.

Approach :

We import the given dataset into MySQL Workbench and analyze the data by writing SQL queries based on the requirement from the marketing team and provide the results to the team.

Tools/Technologies Used

In this project , we use MySQL Workbench as the primary tool to gather relevant insights from the given dataset.
We use SQL queries to communicate with the database which contains the data from the dataset to retrieve only the data which we require based on the requirement from the marketing team.

Insights

Learnt how to use SQL queries by writing concise and optimized queries and also gained knowledge on several functions being used in MySQL to gather data effectively. The inferences made from the data are present in the below page.

Result

The results as per the requirements from the marketing team are present in the below page.

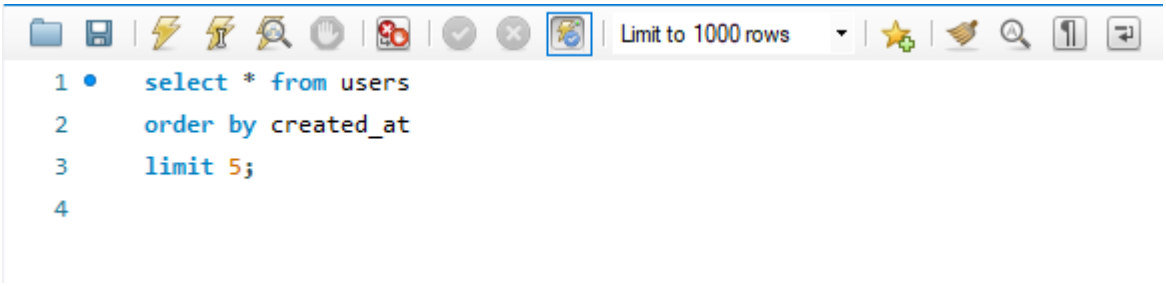
The same will aid the team to make better decisions on how to proceed further with their tasks.

Overall Learning and Progress

This project really helped me to understand the working of databases and how to communicate with the database to gather required data and also learn about different functions present in MySQL which helps to further filter data as per the requirement.

1. **Loyal User Reward**

Query:



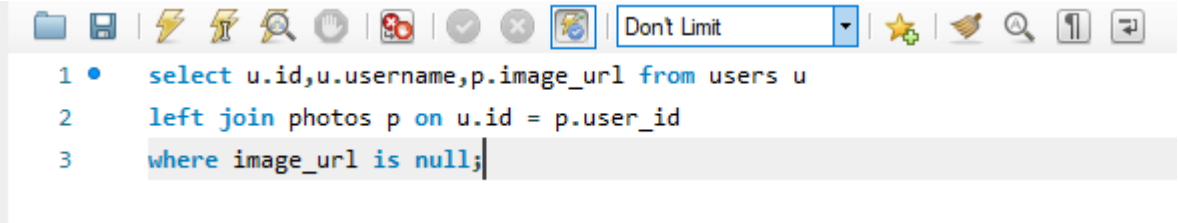
Output :

	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

Outcome : Retrieved the five oldest users on the platform and the oldest user is Darby Herzog whose profile was created on 06-05-2016 at 12:14 AM.

2. **Inactive User Engagement**

Query:



Output :

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

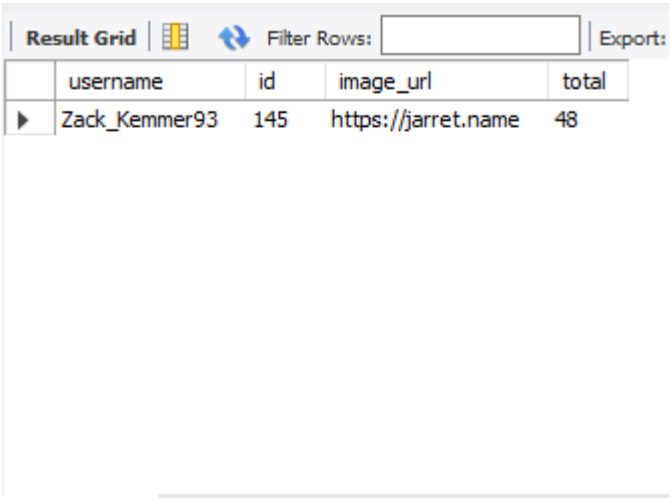
Outcome : There are 26 profiles who never posted any photo on Instagram till date.

3. Contest Winner Declaration

Query :

```
SELECT username , photos.id  , photos.image_url , COUNT(*) as total FROM photos
INNER JOIN likes ON likes.photo_id = photos.id
INNER JOIN users on users.id = photos.user_id
group by id
order by total desc
limit 1;
```

Output :



Result Grid | Filter Rows: | Export:

username	id	image_url	total
Zack_Kemmer93	145	https://jarret.name	48

Outcome : The user with the most likes on a single photo is Zack_Kemmer93. Hence he is the winner of the contest.

4. **Hashtag Research**

Query :

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

Output :

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

Outcome : The top 5 most commonly used hashtags are smile , beach , party , fun and concert with smile being the most used hashtag with count of 59.

5. **Ad Campaign Launch**

Query :

```
select dayname(created_at) as day , count(*) as total from users
group by day
order by total desc
limit 2;
```

Output :

day	total
Thursday	16
Sunday	16

Outcome : The best days to launch the ad campaign is either on Thursday or Sunday. But the best option is to host the campaign on Sunday since it's a holiday and the engagement of users will increase drastically.

Investor Metrics

1. **User Engagement**

Query :

```
select (select count(*) from photos) / (select count(*) from users) as average;
```

Output :

	average
▶	2.5700

Outcome : A user posts an average of 2.57 posts on Instagram based on the given data.

2. **Bots and Fake Accounts**

Query :

```
select u.username , count(*) as total_likes from users u
inner join likes l on u.id = l.user_id
group by l.user_id
having total_likes = (select count(*) from photos);
```

Output :

username	total_likes
Aniya_Hackett	257
Jaclyn81	257
Rocio33	257
Maxwell.Halvorson	257

Ollie_Ledner37	257
Mckenna17	257
Duane60	257
Julien_Schmidt	257
Mike.Auer39	257
Nia_Haag	257
Leslie67	257
Janelle.Nikolaus81	257
Bethany20	257

Outcome : There are 13 users who are potentially identified to be bots as they have liked every single photo on the platform.