INSTAGRAM USER ANALYTICS

SQL FUNDAMENTALS

PROJECT DESCRIPTION

This project helps to analyze raw data/metadata to create useful insights. Various database management tools can be extract useful insights and even visualize them. This enables a way to increase efficiency of a platform.

APPROACH

The approach to this project is kept very straight forward. In order to execute the project ,SQL is used . SQL queries is used to create a database using the raw data provided. Once the database was created, various sorting and data extracting queries is used to get the data/insights required

TECH STACK USED

MySQL Workbench v8.0.32.0 is used during project execution in order to query database. The ease of access and setup, troubleshooting support as well as the GUI made it a good tool for the project.

PROJECT INSIGHTS

A)MARKETTING-:

1- The 5 oldest users are-:

```
select * from users
order by created_at asc
limit 5;
```

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

2- The users who have never posted a single photo on Instagram-:

```
select
users.id,
users.username  ,
photos.user_id ,
photos.image_url
from users
left join photos
on users.id = photos.user_id
where photos.id is null;
```

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

3- The user who gets the most likes on a single photo-:

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

COUNT(*)	id	username
48	145	Zack_Kemmer93

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

```
SELECT tag_name, count(*) as total_tags
FROM photo_tags
JOIN tags
ON photo_tags.tag_id = tags.id
GROUP BY tags.id
ORDER BY total_tags DESC
LIMIT 5;
```

tag_name	total_tags
Smile	59
beach	42
party	39
fun	38
concert	24

5- The team wants to know, which day would be the best day to launch Ads-:

```
SELECT count(DISTINCT username),
dayname(created_at) FROM users
GROUP BY dayname(created_at)
ORDER BY count(DISTINCT username)DESC ;
```

COUNT(DISTINCT USERNAME)	DAYNAME created_at
16	Sunday
16	Thursday
15	Friday
14	Monday
14	Tuesday
13	Wednesday
12	Saturday

B) Investor Metrics:

1- Are users still as active and post on Instagram or they are making fewer posts-:

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

totalphotos
2.5700

2- The investors want to know if the platform is crowded with fake and dummy accounts-:

```
SELECT username,count(*) as num_likes
FROM users
INNER JOIN likes
ON users.id = likes.user_id
GROUP BY likes.user_id
HAVING num_likes = (select count(*) from photos) ;
```

username	num_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

RESULT

- This project is my first on SQL which helped me to sharpen my basic SQL skill.
- By this project I have achieved and gain knowledge how to clean the data with help of MySQL.
- And how to interact with database and how to customize the query to get the desired output.