

Instagram User Analytics

PART A

PROJECT DESCRIPTION –

The objective of this project is to analyze raw data/metadata and generate useful insights. Database management tools can assist in extracting and visualizing insights, leading to improved platform efficiency.

APPROACH –

SQL was used to implement the project. SQL queries were used to create the database using the raw data provided. After creating the database, various sorting and extracting queries were used to get the required data/overview.

TECH STACK USED – MySQL Workbench , Excel,

INSIGHTS –

Instagram user analytics provides user insights into the age, names, trends, posts of user base. Such information used in identifying recent and current trends patterns as well as tailor marketing efforts to specific target audiences.

Q1. Find the 5 oldest users of the Instagram from the database provided.

Sorted the database according to ascending dates and selected the first 5 users from there.

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

	A	B	C
1	id	username	created_at
2	80	Darby_He	06-05-2016 00:14
3	67	Emilio_Be	06-05-2016 13:04
4	63	Elenor88	08-05-2016 01:30
5	95	Nicole71	09-05-2016 17:30
6	38	Jordyn.Jac	14-05-2016 07:56
7			

Q2. Find the users who have never posted a single photo on Instagram.

```
select id,username from users where id not in (select user_id from photos);
```

	A	B	C	D
1	id	username		
2		5 Aniya_Hackett		
3		7 Kasandra_Homenick		
4		14 Jaclyn81		
5		21 Rocio33		
6		24 Maxwell.Halvorson		
7		25 Tierra.Trantow		
8		34 Pearl7		
9		36 Ollie_Ledner37		
10		41 Mckenna17		
11		45 David.Osinski47		
12		49 Morgan.Kassulke		
13		53 Linnea59		
14		54 Duane60		
15		57 Julien_Schmidt		
16		66 Mike.Auer39		
17		68 Franco_Keebler64		
18		71 Nia_Haag		
19		74 Hulda.Macejkovic		
20		75 Leslie67		
21		76 Janelle.Nikolaus81		
22		80 Darby_Herzog		
23		81 Esther.Zulauf61		
24		83 Bartholome.Bernhard		
25		89 Jessyca_West		
26		90 Esmeralda.Mraz57		
27		91 Bethany20		

Q3. Identify the winner of the contest and provide their details to the team.

```
select photo_id, COUNT(*) AS freq from likes group by photo_id order by freq desc;
```

	A	B	C
1	photo_id	freq	
2	145	48	
3	127	43	
4	182	43	
5	123	42	
6	30	41	
7	52	41	
8	61	41	
9	147	41	
10	174	41	
11	192	41	
12	256	41	
13	13	40	
14	97	40	

The user having the photo_id = 145 has the maximum number of likes so he is the winner of the contest.

Q4. Identify and suggest the top 5 most used hashtags on the platform.

Found the frequency of the tag ids from photo tags and grouped them by tag id to order them by descending frequency

```
select tag_id, count(*) as freq from photo _ tags group by tag_id order by freq desc;
```

	A	B	
1	tag_id	freq	
2	21	59	
3	20	42	
4	17	39	
5	13	38	
6	5	24	
7	11	24	
8	18	24	
9	15	23	
10	12	22	
11	8	20	
12	10	20	
13	1	19	
14	16	19	
15	19	19	
16	3	17	
17	4	17	
18	14	17	
19	2	16	
20	9	16	
21	7	15	
22	6	11	

Q5. What day of the week do most users register on? Provide insights on when to schedule an ad campaign.

Used the “DAYNAME()” function to convert the sign up dates into days of the week and then found the frequency of the days and sorted accordingly to select the top days.

```
select dayname(created_at), AS freq from users group by dayname(created_at) order by freq desc
```

	A	B
1	dayname(freq	
2	Thursday	16
3	Sunday	16
4	Friday	15
5	Tuesday	14
6	Monday	14
7	Wednesday	13
8	Saturday	12

PART (B)

Q6. Provide how many times does average user posts on Instagram. Also, provide the total number of photos on Instagram/total number of users.

```
create table avg as select user_id, COUNT(*) AS freq from photos group by user_id order by freq desc;
```

```
select * from avg;
```

```
select avg(freq) from avg;
```

```
select count(*) from photos;
```

average user post on instagram

	A	B
1	avg(freq)	
2	3.473	

Total number of users

	A	B
1	count(*)	
2	257	
3		

Q7. 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).

	A	B	
1	user_id	freq	
2	71	257	
3	54	257	
4	91	257	
5	5	257	
6	57	257	
7	41	257	
8	75	257	
9	76	257	
10	36	257	
11	21	257	
12	66	257	
13	14	257	
14	24	257	

RESULT-

In this project I have got to work on real world Instagram dataset and learned how to extract and short the data which help and support in finding solutions , based on that the right decesions can be made.