PROJECT_2

Prepared By

Vaishnavi Patel

Instagram User Analytics

Project Description

- The project "Instagram User Analytics" provides to work on one type of social media analysis with provided raw database.
- It includes giving insights to marketing team and investors based on the question asked by management team for better overall growth of social media platform.
- Team wants to know the current situation, involvement of users from past to present status of users and depending on the information they want to lunch campaigns, want to announce winner, and want all information show that improvement can be done and necessary steps can be taken to grow business.

Project Description

- The project report includes
  - ✓ Data retrieved from raw database depending upon query
  - ✓ Answer of the all questions solved by SQL (structured query language)
  - ✓ With SQL code and Result-set as output
  - ✓ Possible important information

Approach

- Approach to problem solving for better and effective completion project
  - Identification and understanding of problem and each task
  - What exactly team wants to know
  - Which type of data will be required and how many data will be needed
  - How can I solve it ?

SQL

- **Which tools** will be required (Ans: SQL language is used for analysis)
- Properly **written SQL code** for desired result
- **Collection of result-set** for particular task
- **Retrieving** information from result-set for each task
- Presenting result in **readable** and **presentable** manner

Approach

- This approach ensures that the project report is comprehensive, accurate, and easy to understand for all stakeholders involved.
- By using SQL, the report enables a structured and efficient analysis of the data, allowing for insights to be easily extracted and conclusions to be drawn.

- The inclusion of SQL code and result-sets as output ensures that the analysis can be reproduced and verified, providing transparency and credibility to the findings.
- Overall, the project report aims to provide a clear understanding of the data and its implications, enabling informed decision-making for the business.

TECH-STACK USED

- Software Details :
  - MySQL workbench version 8.0 CE
- Advantages of MySQL workbench :
  - It is open source
  - It can work on different operating systems
  - It works on web-database
  - Easy to learn and execute SQL code
  - We can easily debug the errors, see it and resolve it
  - Scalability is high
- Looking to advantages, it gives purpose to use this software.
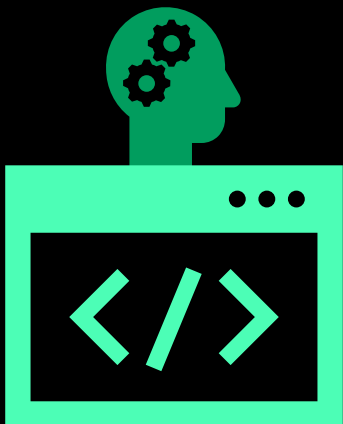
SQL code:

```
1    /* List of 5 Oldest Instagram Users*/
2 •  SELECT u.id, u.username, u.created_at
3    FROM users u
4    ORDER BY created_at ASC
5    LIMIT 5;
```

SQL result set:

Result Grid | Filter Rows:

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

A1. Rewarding Most Loyal Users:

People who have been using the platform for the longest time.

Task :

Find the **5 oldest users** of the Instagram from the database provided

SQL code:

```sql
/* Users who have never posted a photo */

SELECT u.username, p.user_id, u.id, p.image_url

FROM users u

LEFT JOIN photos p

ON u.id = p.user_id

WHERE image_url IS NULL
```

SQL result set:

username	user_id	id	image_url
Aniya_Hackett	NULL	5	NULL
Kasandra_Homenick	NULL	7	NULL
Jadyn81	NULL	14	NULL
Rocio33	NULL	21	NULL
Maxwell.Halvorson	NULL	24	NULL
Tierra.Trantow	NULL	25	NULL
Pearl7	NULL	34	NULL
Ollie_Ledner37	NULL	36	NULL
Mckenna17	NULL	41	NULL
David.Osinski47	NULL	45	NULL
Morgan.Kassulke	NULL	49	NULL
Linnea59	NULL	53	NULL
Duane60	NULL	54	NULL
Julien_Schmidt	NULL	57	NULL
Mike.Auer39	NULL	66	NULL
Franco_Keebler64	NULL	68	NULL
Nia_Haag	NULL	71	NULL

Result Grid | Filter Rows:

A2. Remind Inactive Users to Start Posting:

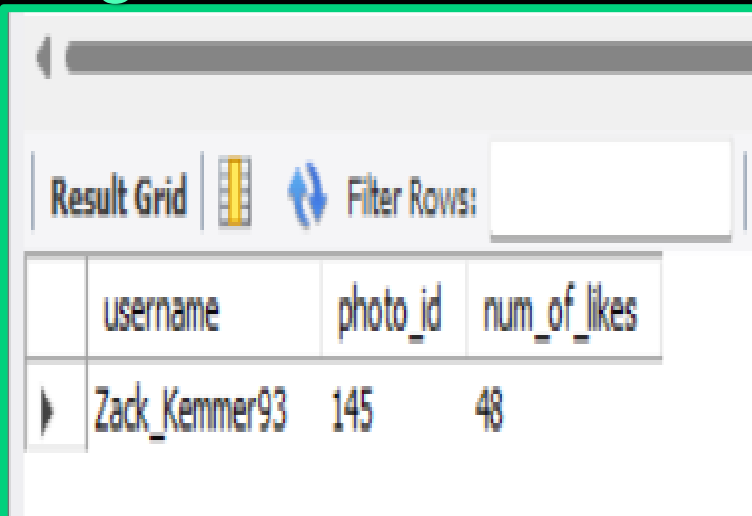By sending them promotional emails to post **their** 1st photo

Task:

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

SQL code:

```
/* Identify winner : The user who got most likes on single photo*/

SELECT u.username, l.photo_id, COUNT(l.user_id) as num_of_likes

FROM likes l

INNER JOIN photos p ON l.photo_id = p.id

INNER JOIN users u ON p.user_id = u.id

GROUP BY photo_id

ORDER BY num_of_likes DESC

LIMIT 1;
```

SQL result set:

username	photo_id	num_of_likes
Zack_Kemmer93	145	48

Result Grid | Filter Rows:

A3. 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.

Task :

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

SQL code:

```sql
/* List of Top 5 most commnly used hashtags on Instagram*/
SELECT t.tag_name, COUNT(pt.photo_id) as total_tag_count
FROM tags t
INNER JOIN photo_tags pt
ON t.id = pt.tag_id
GROUP BY tag_name
ORDER BY total_tag_count DESC
LIMIT 5;
```

SQL result set:

Result Grid		Filter Rows:

	tag_name	total_tag_count
▶	smile	59
	beach	42
	party	39
	fun	38
	concert	24

A4. Hashtag Researching:

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

Task :
Identify and suggest the top 5 most commonly used hashtags on the platform

SQL code:

```sql
/* To lunch ADs which day of the week would be the best*/
SELECT DAYNAME(created_at) as week_day,
       COUNT(id) as total_users
FROM users
GROUP BY week_day
ORDER BY total_users DESC;
```

SQL result set:

week_day	total_users
Thursday	16
Sunday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12

A5. Launch AD Campaign:

The team wants to know, which day would be the best day to launch ADs.

Task :

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

SQL code:

```
/* How many times an average user posts on instagram*/
SELECT distinct(COUNT(*)) FROM photos as total_photos;/*257*/
SELECT distinct(COUNT(*)) FROM users as total_users; /*100*/
/*User engagement can be obtained from calculating average of these*/
SELECT (
(SELECT distinct(COUNT(*)) FROM photos)
/
(SELECT distinct(COUNT(*)) FROM users)
) as avg_post_per_photos;/*2.57*/
```

SQL result set:

Result Grid	Filter Rows:
avg_post_per_photos
2.5700

B1. User Engagement:

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

Task :

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

SQL code:

```
/* List of users who have liked every single photo posted*/

SELECT u.id,u.username,COUNT(u.id) as total_like_by_user, l.user_id

FROM users u

JOIN likes l ON u.id = l.user_id

GROUP BY u.id

HAVING total_like_by_user = (SELECT COUNT(*) FROM photos); /*total_photos =257*/
```

SQL result set:

id	username	total_like_by_user	user_id
5	Aniya_Hackett	257	5
14	Jadyn81	257	14
21	Rocio33	257	21
24	Maxwell.Halvorson	257	24
36	Ollie_Ledner37	257	36
41	Mckenna17	257	41
54	Duane60	257	54

Result 50

B2. Bots & Fake Accounts:

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

Task :

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).

By running the SQL code I found below results.

It is noted step by step:

(A) Analysis of Marketing Data Research

A1. Rewarding Most Loyal Users

List of people who have been using the platform for the longest time : Darby_Herzog , Emilio_Bernier52, Elenor88, Nicole71, Jordyn.Jacobson2 are peoples are most five person who are using platform for the longest time.

A2. Remind Inactive Users to Start Posting

There are total of 26 users who have been inactive for so long time. So, we need to remind them by sending promotional emails to post their 1st photo.

A3. Declaring Contest Winner

The team started a contest and the user who gets the most likes on a single photo is a Zack_Kemmer93 and maximum likes are 48.

A4. Hashtag Researching

A partner brand wants to know, which hashtags to use in the post :
After running the SQL query smile, beach, party, fun and concert are most used hashtags and can be used.

A5.Launch AD Campaign

The team wants to know, which day would be the best day to launch Ads
SQL query tells that most users have created their account on Thursday and Sunday, then Friday , than Tuesday and so on. So, Thursday and Sunday can be a best day to lunch Ads.

B) Analysis of Data for Investor Metrics

B1. User Engagement

User engagement can be measured in variety of ways. By finding total number of post per users, photos likes by users, comments made by single user on any photo, counting any shared post or story will gives information on user engagement. Here, total number of photos is 257 and total users are 100. So, average post per photos is 2.57 which shows user engagement.

B2. Bots & Fake Accounts

No users will like each and every photo posted on Instagram. If we can count the users who have likes each and every photo we can say about Fake accounts. SQL code gives list of total 13 users who have like each photo on Instagram so can possibly are fake accounts.

RESULT

Understanding of different task and problems of marketing research, social media data analysis

Database creation, used DDL and DML commands, writing SQL query

Got problem solving skill sharpen and project report writing skills on PPT is improved

Got an idea about how social media analysis might are working on data, giving recommendation

"User Instagram Analytics" projects gave an idea about work of data analytics, how marketing and investors make decisions based on data analysis.