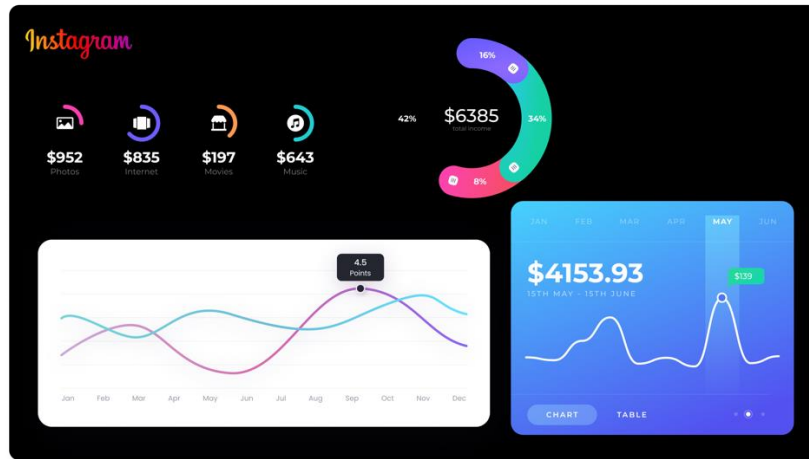


Project 02

Instagram User Analytics



Project Description:

Imagine you're a data analyst working with the product team at Instagram. Your role involves analysing user interactions and engagement with the Instagram app to provide valuable insights that can help the business grow.

User analysis involves tracking how users engage with a digital product, such as a software application or a mobile app. The insights derived from this analysis can be used by various teams within the business. For example, the marketing team might use these insights to launch a new campaign, the product team might use them to decide on new features to build, and the development team might use them to improve the overall user experience.

In this project, I have leveraged SQL and MySQL Workbench as tool to analyse Instagram user data and answer questions posed by the management team. Your insights will help the product manager and the rest of the team make informed decisions about the future direction of the Instagram app.

SQL Tasks:

A) Marketing Analysis:

1. **Loyal User Reward:** The marketing team wants to reward the most loyal users, i.e., those who have been using the platform for the longest time.

Your Task: Identify the five oldest users on Instagram from the provided database.

2. **Inactive User Engagement:** The team wants to encourage inactive users to start posting by sending them promotional emails.

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

3. **Contest Winner Declaration:** The team has organized a contest where the user with the most likes on a single photo wins.

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

4. **Hashtag Research:** A partner brand wants to know the most popular hashtags to use in their posts to reach the most people.

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

5. **Ad Campaign Launch:** The team wants to know the best day of the week to launch ads.

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

B) Investor Metrics:

1. **User Engagement:** Investors want to know if users are still active and posting on Instagram or if they are making fewer posts.

Your Task: 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.

2. **Bots & Fake Accounts:** Investors want to know if the platform is crowded with fake and dummy accounts.

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

Approach

For this project, I have used My SQL Knowledge to connect with SQL Server using My SQL Workbench and extracted the required data from the given database using the Join function, subqueries, Aggregation, where Clause, Group by, Distinct and other functions are required.

keeping the Primary key and foreign key in consideration provided all the reports asked by the marketing department and metrics from the Investor's POV

Tech-Stack Used

MySQL Workbench 8.0 CE

A) Marketing Analysis

Loyal User Reward: The marketing team wants to reward the most loyal users, i.e., those who have been using the platform for the longest time.

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

```
SELECT * FROM users
ORDER BY created_at DESC
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	
▶	NULL	NULL	NULL	

Inactive User Engagement: The team wants to encourage inactive users to start posting by sending them promotional emails.

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

```
SELECT u.id, u.username, u.created_at
FROM users u
LEFT JOIN photos p ON u.id = p.user_id
WHERE p.id IS NULL;
```

	id	username	created_at	
▶	5	Aniya_Hackett	2016-12-07 01:04:39	
▶	83	Bartholome.Bernh...	2016-11-06 02:31:23	
▶	91	Bethany20	2016-06-03 23:31:53	
▶	80	Darby_Herzog	2016-05-06 00:14:21	
▶	45	David.Osinski47	2017-02-05 21:23:37	
▶	54	Duane60	2016-12-21 04:43:38	
▶	90	Esmeralda.Mraz57	2017-03-03 11:52:27	
▶	81	Esther.Zulauf61	2017-01-14 17:02:34	
▶	68	Franco_Keebler64	2016-11-13 20:09:27	
▶	74	Hulda.Macejkovic	2017-01-25 17:17:28	
▶	14	Jaclyn81	2017-02-06 23:29:16	
▶	76	Janelle.Nikolaus81	2016-07-21 09:26:09	
▶	89	Jessyca_West	2016-09-14 23:47:05	
▶	57	Julien_Schmidt	2017-02-02 23:12:48	
▶	7	Kasandra_Homen...	2016-12-12 06:50:08	
▶	75	Leslie67	2016-09-21 05:14:01	
▶	53	Linnea59	2017-02-07 07:49:34	
▶	24	Maxwell.Halvorson	2017-04-18 02:32:44	
▶	41	Mckenna17	2016-07-17 17:25:45	
▶	66	Mike_Auer39	2016-07-01 17:36:15	
▶	49	Morgan.Kassulke	2016-10-30 12:42:31	
▶	71	Nia_Haag	2016-05-14 15:38:50	
▶	36	Ollie_Ledner37	2016-08-04 15:42:20	
▶	34	Pearl7	2016-07-08 21:42:01	
▶	21	Rocio33	2017-01-23 11:51:15	
▶	25	Tierra.Trantow	2016-10-03 12:49:21	

Contest Winner Declaration: The team has organized a contest where the user with the most likes on a single photo wins.

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

```
SELECT u.id, u.username, p.id AS photo_id, COUNT(l.user_id) AS likes_count
FROM users u
JOIN photos p ON u.id = p.user_id
LEFT JOIN likes l ON p.id = l.photo_id
GROUP BY u.id, u.username, p.id
ORDER BY likes_count DESC
LIMIT 1;
```

	id	username	photo_id	likes_count
▶	52	Zack_Kemmer93	145	48

Hashtag Research: A partner brand wants to know the most popular hashtags to use in their posts to reach the most people.

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

```
SELECT tag_name, COUNT(*) AS hashtag_count
FROM photo_tags
JOIN tags ON photo_tags.tag_id = tags.id
GROUP BY tag_name
ORDER BY hashtag_count DESC
LIMIT 5;
```

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

Ad Campaign Launch: Determining the day of the week when most users register on Instagram to schedule an ad campaign.

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

```
SELECT DAYNAME(created_at) AS day_of_week, COUNT(*) AS user_count
FROM users
GROUP BY day_of_week
ORDER BY user_count DESC
```

day_of_week	user_count
Thursday	16
Sunday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12

B) Investor Metrics

User Engagement: Investors want to know if users are still active and posting on Instagram or if they are making fewer posts

- 1. 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.

Average Number of Posts per User

```
SELECT COUNT(*) / COUNT(DISTINCT user_id) AS average_posts_per_user
FROM photos;
```

average_posts_per_user
3.4730

Total Number of Photos Divided by the Total Number of Users

```
SELECT COUNT(*) AS total_photos, COUNT(DISTINCT user_id) AS total_users
FROM photos;
```

	total_photos	total_users	
▶	257	74	

Bots & Fake Accounts: Investors want to know if the platform is crowded with fake and dummy accounts.

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

```
SELECT * FROM users, likes;
WITH base AS(
SELECT u.username, count(l.photo_id) AS likess
FROM likes AS l
INNER JOIN users AS u
ON u.id=l.user_id
GROUP BY u.username
)
```

```
SELECT username, likess FROM base WHERE likess=(SELECT count(*) FROM photos) ORDER
BY username;
```

	username	likess	
▶	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	

Insights

Marketing Analysis

Rewarding the Most Loyal Users:

Finding the 5 oldest users who have been engaging on the platform for the longest time will facilitate marketing team to introduce loyalty programs and provide reward points

Reminding Inactive Users to Increase Engagement

Identifying users who have never posted a single photo on Instagram will let marketing team to encourage inactive users to start posting by introducing promotional campaigns

Declaring Winner of the Contest

Identifying maximum number of likes for the Photo and respective information of the user. So it will help the Marketing team to declare the Winner of that Contest.

Searching Hashtags

Identifying the top 5 most used hashtags by users will help partner brands to know the most popular hashtags used in their posts in order to reach the most people.

Launching Ad Campaign

Identifying days of the week for the Marketing team to gauge the best day for launching ad Thursday and Sunday turn out to be the days when most number of users register on the platform

Investor Metrics

Users Engagement

Calculating the average number of posts per user on Instagram. So It will help the Investors to know if users are still active and posting on Instagram or if they are making fewer posts.

Total Number of Photos on Instagram = 257

Total Number of Users = 100

Average = $257/100 = 2.57$

Since the number of active users is 74

Average number of posts per user on Instagram = $257/74 = 3.47$

Average user post 3-4 times

Bot and Fake Accounts

Identifying users (potential bots) who have liked every single photo on the platform as this is highly unlikely for a user. This information will make sure investors gauge if the platform has fake and dummy accounts. We have identified a potential impact for the investors.

Result

This project has helped me delve into data analysis using SQL fundamentals and advanced concepts. It has been a fulfilling exercise for gathering useful insights from the given data