

Instagram User Analytics

Name : Aniket Ashok Ubale

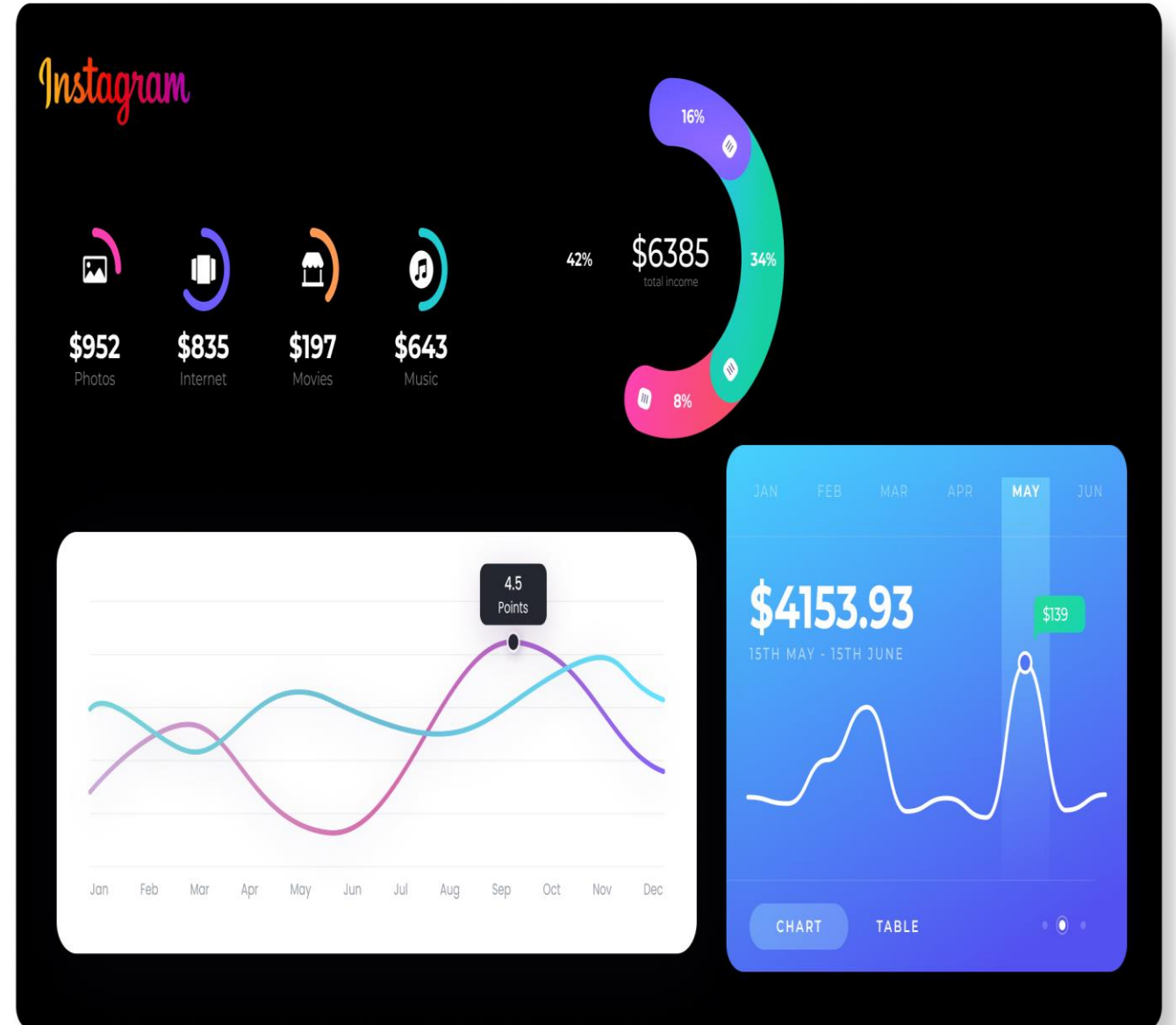
Class : B. Tech

Year : 3rd

Branch : Artificial intelligence and data science

Project : Instagram User Analytics

Date : 19-03-2023



Project Description

- Imagine you're a data analyst working with the product team at Instagram. Your role involves analyzing 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, you'll be using SQL and MySQL Workbench as your tool to analyze 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.
- Remember, the goal of this project is to use your SQL skills to extract meaningful insights from the data. Your findings could potentially influence the future development of one of the world's most popular social media platforms.

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.

How to do this Project?

1.Create a Database: Start by running the provided commands to create the necessary database for the project. The database file has been provided in the attachments.

2.Perform Analysis: Utilize SQL to perform the analysis and answer the questions mentioned in the project description. Use SQL queries to extract the required information from the database, ensuring the accuracy and efficiency of the queries.

3.Submit a Report: Prepare a report in PDF or PowerPoint format that will be presented to the leadership team. The report should include the following sections:

Note: Make sure you mention the SQL Queries along with the outputs in the report. You can take snapshots of the SQL Queries and the outputs. It is advised to do the project in MySQL Workbench.

Instagram user analytics

- Learning data analytics involves developing skills in data manipulation, analysis, and visualization.
- Data analytics help a business optimize its performance, perform more efficiently, maximize profit, or make more strategically-guided decisions.
- Data analytics is the science of analyzing raw data to make conclusions about that information.
- User analytics is the main important thing to every organization.

Project – Instagram user analytics

Description-

- The Instagram user analytics project is all about the analyzing a raw data and derive meaningful information from it.
- This project is about the user's that are on instagram and extracting the insights that are then used by teams across the business to launch a new marketing campaign, decide on features to build for an app, track the success of the app by measuring user engagement and improve the experience altogether while helping the business grow. It gives a brief about the behavior and the ratios of people of Instagram.
- The project gave a real time industry level experience to me.
- I handled all the things using data analytics skill and mysql tool.

Approach-

- To perform the required tasks and complete the project , I have used SQL queries using MySQL Command line client . According to the instructions provided to create the database and the corresponding tables , I feeded the data into MySQL and executed the appropriate queries to get the required insights
- The approach is simple as we are going to use the sql command to use the database as tables and the extract the insights from the database through various commands.

Tack Stack Used-

1. MySQL Workbench 8.0 CE (Command line client)
2. Microsoft Powerpoint (Office 365)

Insights-

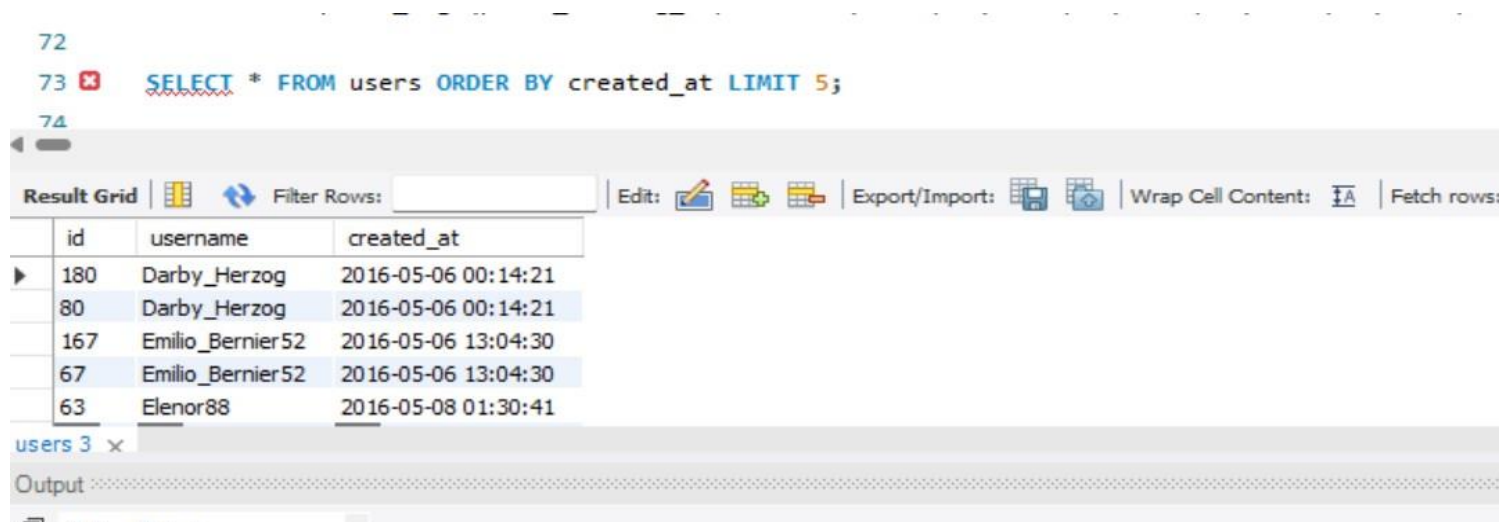
- I have learned how to handle datasets and how to analyze a dataset.
- I already had a brief hands-on experience with SQL in my Bachelor's degree curriculum, but it was only basic concepts which I had worked with. This Instagram user analytics project helped me to dig deeper into the world of SQL and helped me understand how complex queries work and how to cultivate business insights from given data. It enabled me to ask the right necessary questions and narrow down solutions to the given problems.

Rewarding Most Loyal Users: People who have been using the platform for the longest time.
Your Task: Find the 5 oldest users of the Instagram from the database provided

- Query used-

SELECT * FROM users ORDER BY created_at LIMIT 5;

➤ ORDER BY used for arrange a data in sequence mannar and LIMIT 5 is used fetching a upper 5 rows data.



The screenshot shows a database query interface. At the top, a SQL query is entered: `SELECT * FROM users ORDER BY created_at LIMIT 5;`. Below the query, a toolbar contains various icons for editing and exporting. The main area displays a table with the results of the query. The table has three columns: `id`, `username`, and `created_at`. The results are as follows:

	id	username	created_at
▶	180	Darby_Herzog	2016-05-06 00:14:21
	80	Darby_Herzog	2016-05-06 00:14:21
	167	Emilio_Bernier52	2016-05-06 13:04:30
	67	Emilio_Bernier52	2016-05-06 13:04:30
	63	Elenor88	2016-05-08 01:30:41

Below the table, there is a tab labeled `users 3` and an `Output` section.

Remind Inactive Users to Start Posting: By sending them promotional emails to post their 1st photo.

Your Task: Find the users who have never posted a single photo on Instagram

- Query used-

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

➤ LEFT JOIN is used for fetching matching all matching records as well as all user table data, I joined two tables based on id and user_id and IS NULL is used for fetched all null records.

```
80 • SELECT username  
81 FROM users  
82 LEFT JOIN photos  
83 ON users.id=photos.user_id  
84 WHERE photos.id IS NULL;  
85  
86
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: [IA](#)

username
Aniya_Hackett
Kassandra_Homenick
Jaclyn81
Rocio33
Maxwell.Halvorson
Tierra.Trantow
Pearl7
Ollie_Ledner37
Mckenna17
David.Osinski47

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.

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

- Query used-

```
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 photos.user_id  
= users.id GROUP BY photos.id ORDER BY total DESC LIMIT 1;
```

➤ Contest winner- zack_Kemmer93 id-145

```
85  
86 • SELECT username, photos.id, photos.image_url, COUNT(*) AS total  
87 FROM photos  
88 INNER JOIN likes ON likes.photo_id = photos.id  
89 INNER JOIN users ON photos.user_id = users.id  
90 GROUP BY photos.id  
91 ORDER BY total DESC  
92 LIMIT 1;
```

Result Grid	Filter Rows:	Export:	Wrap Cell Content:	Fetch rows:
username	id	image_url	total	
Zack_Kemmer93	145	https://jarret.name	48	

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

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

- Query used-

```
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;
```

➤ Top 5 hashtags are- smile, beach ,party ,fun ,lol

```
103 • SELECT tags.tag_name, COUNT(*) AS total
104 FROM photo_tags
105 JOIN tags ON photo_tags.tag_id = tags.id
106 GROUP BY tags.id
107 ORDER BY total DESC
108 LIMIT 5;
```

Result Grid	Filter Rows:	Export:	Wrap Cell Content:	Fetch rows:
tag_name	total			
smile	59			
beach	42			
party	39			
fun	38			
concert	24			

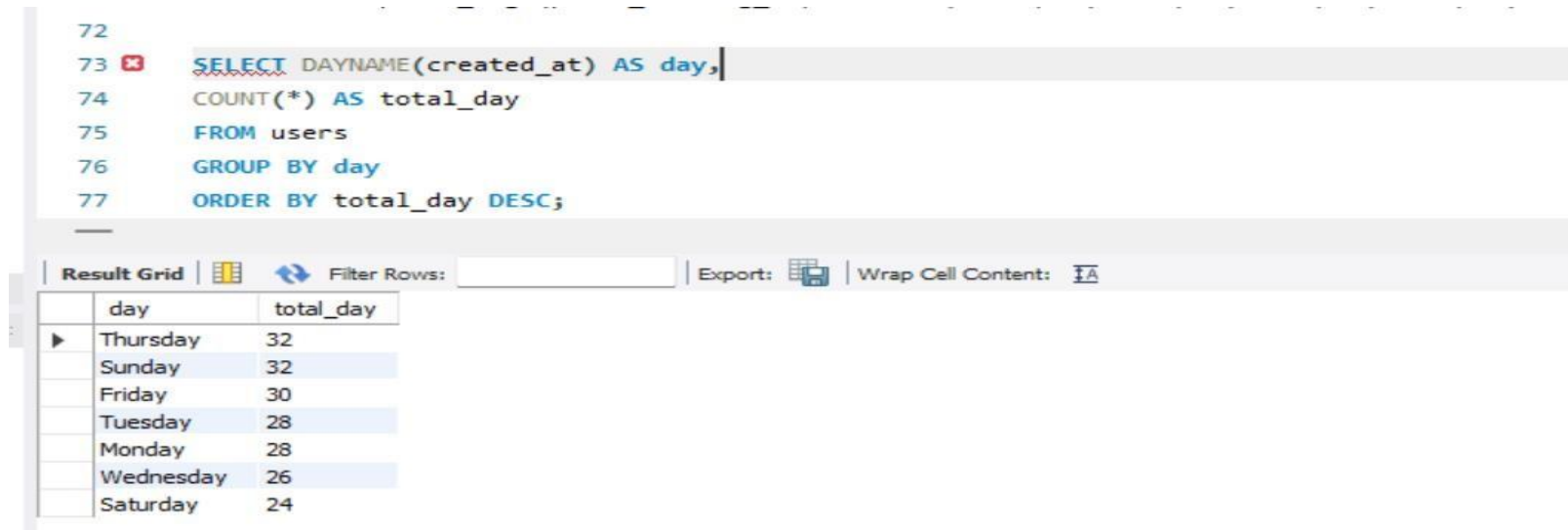
Launch AD Campaign: The team wants to know, which day would be the best day to launch ADs.

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

- Query used-

```
SELECT DAYNAME(created_at) AS day, COUNT(*) AS total_day FROM users GROUP BY day ORDER BY total_day DESC;
```

➤ COUNT used for counting the total days ,ORDER BY used for arrange a row ,DESC used for arrange a data in descending order



The screenshot shows a SQL query editor with the following query:

```
72  
73 SELECT DAYNAME(created_at) AS day,  
74 COUNT(*) AS total_day  
75 FROM users  
76 GROUP BY day  
77 ORDER BY total_day DESC;
```

Below the query editor, the results are displayed in a table with the following data:

	day	total_day
▶	Thursday	32
	Sunday	32
	Friday	30
	Tuesday	28
	Monday	28
	Wednesday	26
	Saturday	24

Investor Metrics:

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

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

- Query used –

`SELECT(SELECT COUNT(*) FROM photos)/(SELECT COUNT(*) FROM users);`

➤ Total photos-257

➤ Total users – 200

➤ Average user post – 1.2850

93

94 • `SELECT(SELECT COUNT(*) FROM photos)/(SELECT COUNT(*) FROM users);`

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	(SELECT COUNT(*) FROM photos)/(SELECT COUNT(*) FROM users)			
	1.2850			

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



Your 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)

- Query used-

```
SELECT username, COUNT(*) AS liked FROM users INNER JOIN likes ON
users.id = likes.user_id GROUP BY likes.user_id HAVING liked = (SELECT
COUNT(*) FROM photos);
```

➤ Total 13 users found

```
97 • SELECT username, COUNT(*) AS liked
98 FROM users INNER JOIN likes
99 ON users.id = likes.user_id
100 GROUP BY likes.user_id
101 HAVING liked = (SELECT COUNT(*) FROM photos);
102
```

Result Grid		
Filter Rows: <input type="text"/>		
Export:  Wrap Cell Content: 		
	username	liked
▶	Aniya_Hackett	257
	Jadyn81	257
	Rocio33	257
	Maxwell.Halvorson	257
	Ollie_Ledner37	257
	Mckenna17	257
	Duane60	257
	Julien_Schmidt	257
	Mike.Auer39	257
	Nia_Haag	257

Result-

- I've learned about numerous data analytics approaches and processes, including data pretreatment and exploratory data analysis.
- This understanding has aided my understanding of the sql language used in data analytics.
- How can user data be managed in real time in campaigns? that thought I had

Link for database creation: [link](#)

Thank you..!!