

Instagram User Analytics

Data Analytics Project



Submitted By-
Anuj Pratap Singh

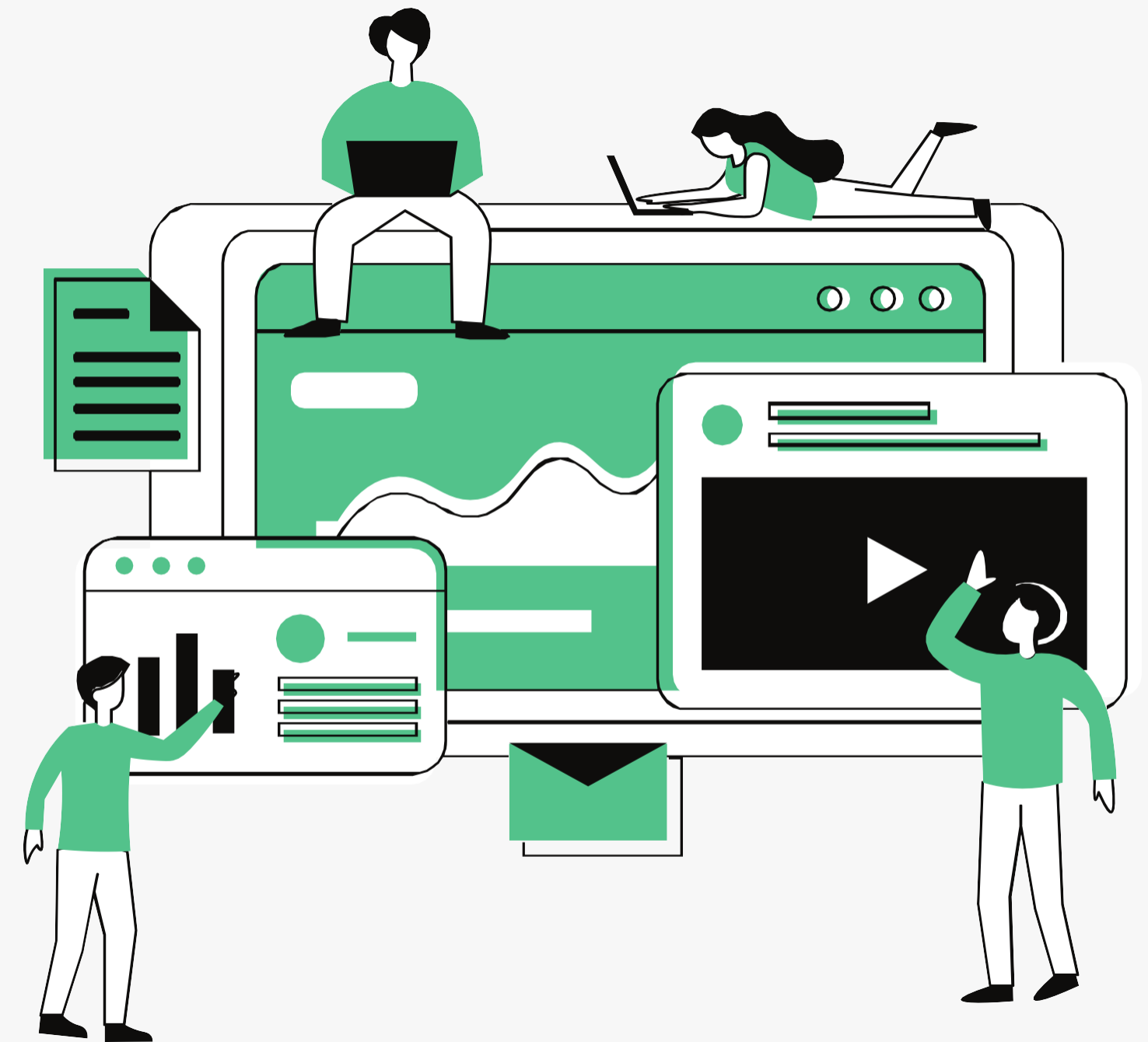


Table of
Contents

		Page
I	Project Description	3
II	Tech-Stack Used	4
III	Approach	5
IV	Insights	9
V	Result	18

I Project Description

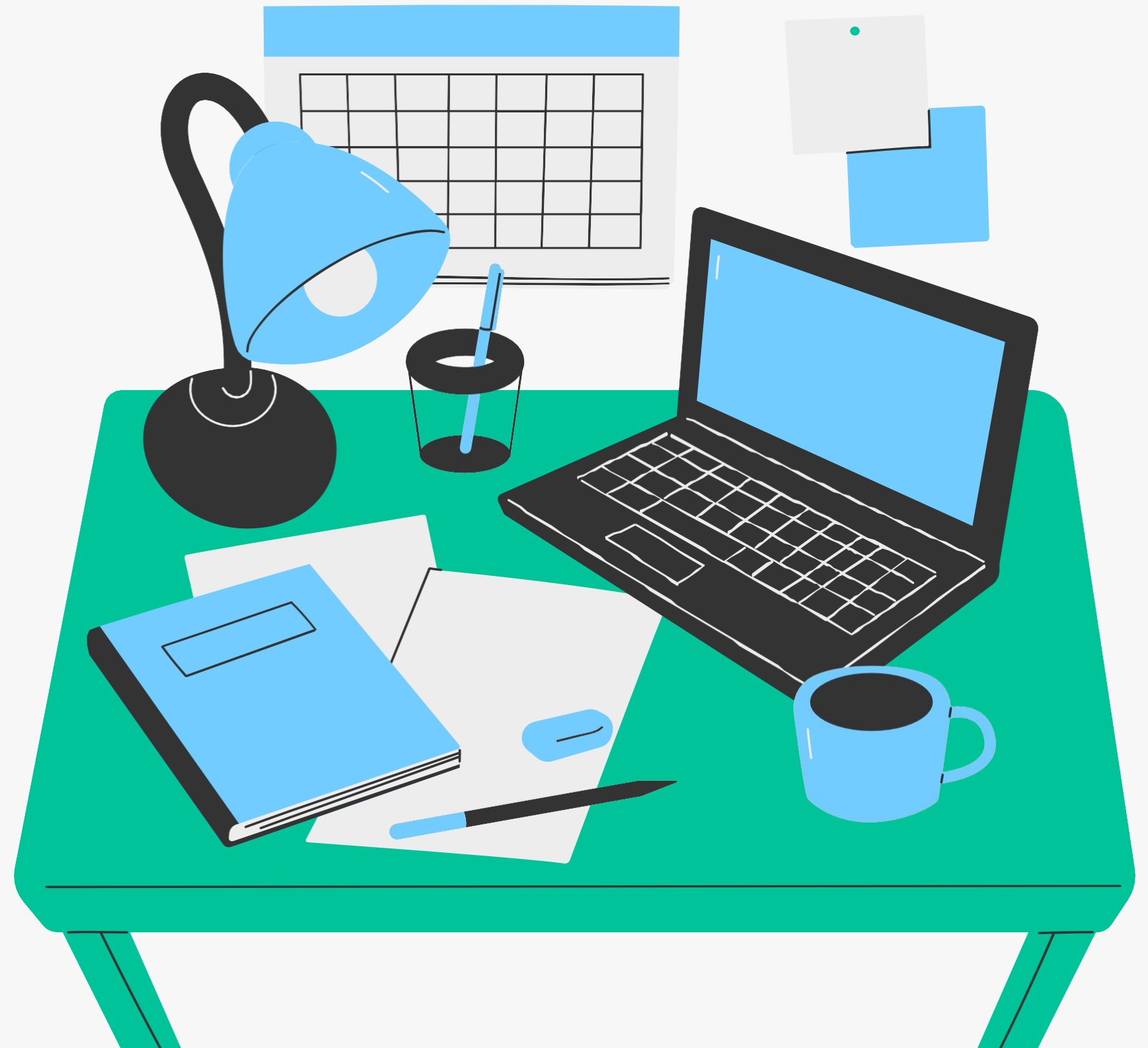
- This project focuses on finding insights into the leading questions that the Marketing team and the Investors of Instagram had.
- While working with the product team, I was given the task to perform an end-to-end user analysis to track how the users were interacting with the digital product with the scope to derive business insights for the marketing, product, and developing teams.
- The leading question covered areas like finding loyal users, inactive users, primary hashtags, and primary registering days, among others. They were asked with the intent of being used by the team to decide which new features to build, launch a new ad campaign, or just track the success of the product by measuring how the users are engaging with it.



II Tech-Stack Used

- MySQL Workbench 8.0 CE

I only used the MYSQL Workbench to clean the data, there were no duplicates and other unknown values, fortunately. I also created a relational schema and executed some SQL queries on the software to find insights about the scope of my study.



III Approach

Exploring the Datasets

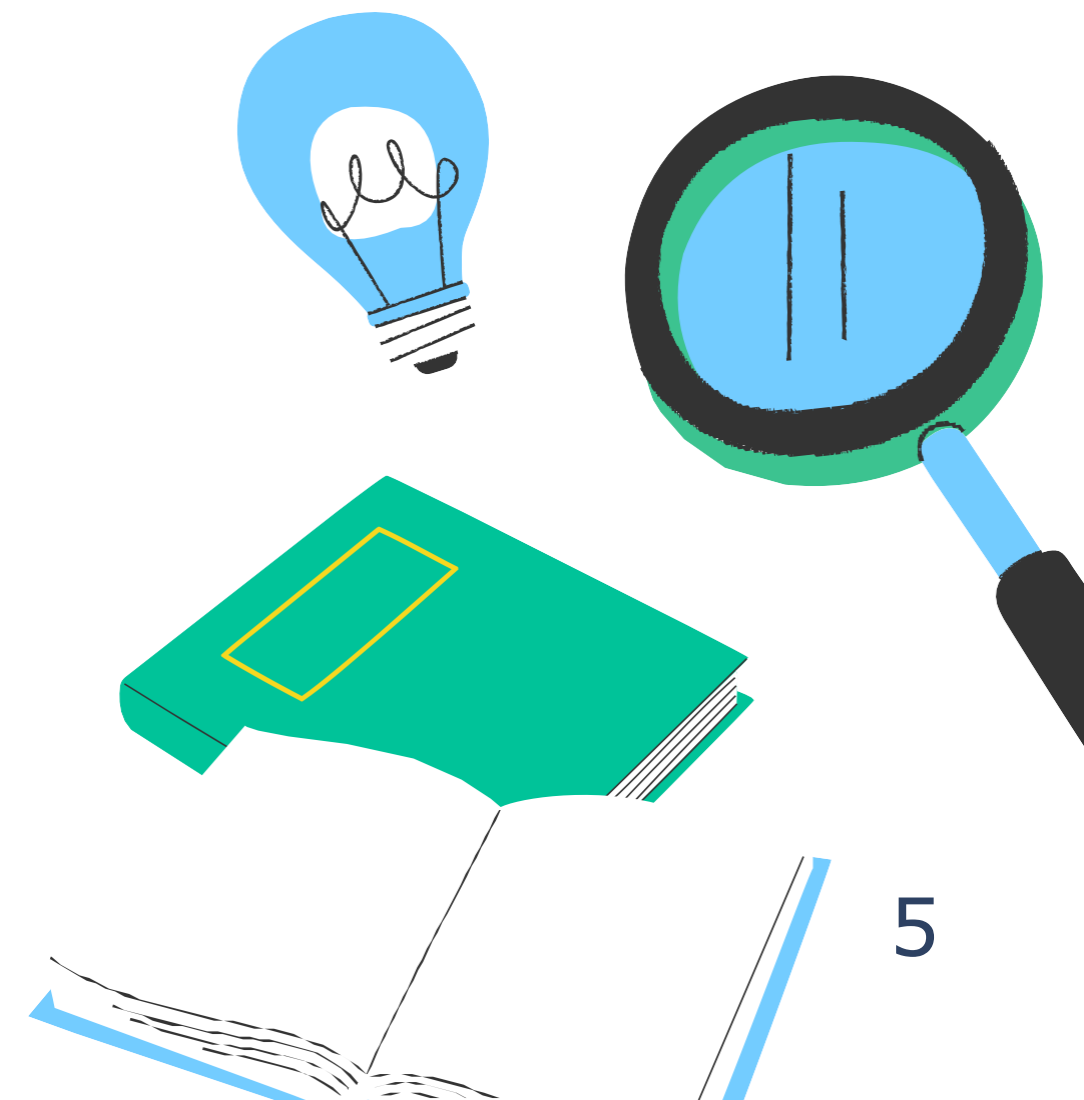
Initially, I explored all 7 tables that the database had and executed SQL queries to check all the rows and columns and remove duplicate or unknown values if there were any.

Analyzing the Datasets

Further, I analyzed the datasets and executed SQL queries to derive some crucial insights that were important for the strategy execs

Providing Recommendations

Finally, after deriving all the crucial insights, I provided some recommendations to help the management team identify white space opportunities and take further steps accordingly.



III.B Expanding all the Tables

Users Table

```
1 • |SELECT * FROM ig_clone.users;
```

	id	username	created_at
	1	Kenton_Kirlin	2017-02-16 18:22:11
	2	Andre_Purdy85	2017-04-02 17:11:21
	3	Harley_Lind18	2017-02-21 11:12:33
	4	Arely_Bogan63	2016-08-13 01:28:43
	5	Aniya_Hackett	2016-12-07 01:04:39
	6	Travon.Waters	2017-04-30 13:26:14
	7	Kassandra_Homenick	2016-12-12 06:50:08
	8	Tabitha_Schamberger11	2016-08-20 07:10:46

Photos Table

```
1 • |SELECT * FROM ig_clone.photos;
```

	id	image_url	user_id	created_at
	1	http://elijah.biz	1	2022-12-02 06:37:39
	2	https://shanon.org	1	2022-12-02 06:37:39
	3	http://vicky.biz	1	2022-12-02 06:37:39
	4	http://oleta.net	1	2022-12-02 06:37:39
	5	https://jennings.biz	1	2022-12-02 06:37:39
	6	https://quinn.biz	2	2022-12-02 06:37:39
	7	https://selina.name	2	2022-12-02 06:37:39
	8	http://melina.org	2	2022-12-02 06:37:39

III.C Expanding all the Tables

Photo Tags Table

1 •

SELECT * FROM ig_clone.photo_tags;

	photo_id	tag_id
	14	1
	21	1
	45	1
	75	1
	83	1
	85	1
	91	1

Likes Table

1 •

SELECT * FROM ig_clone.likes;

	user_id	photo_id	created_at
	2	1	2022-12-02 06:37:40
	2	4	2022-12-02 06:37:40
	2	8	2022-12-02 06:37:40
	2	9	2022-12-02 06:37:40
	2	10	2022-12-02 06:37:40
	2	11	2022-12-02 06:37:40
	2	12	2022-12-02 06:37:40

III.D Expanding all the Tables

Comments Table

1

SELECT * FROM ig_clone.comments;

id	comment_text	user_id	photo_id	created_at
1	unde at dolore	2	1	2022-12-02 06:37:40
2	quae ea ducimus	3	1	2022-12-02 06:37:40
3	alias a voluptatum	5	1	2022-12-02 06:37:40
4	facere suscipit sunt	14	1	2022-12-02 06:37:40
5	totam eligendi quaerat	17	1	2022-12-02 06:37:40
6	vitae quia aliquam	21	1	2022-12-02 06:37:40
7	exercitationem occaecati neque	24	1	2022-12-02 06:37:40
8	sint ad fugiat	31	1	2022-12-02 06:37:40
9	nesciunt aut nesciunt	36	1	2022-12-02 06:37:40
10	laudantium ut nostrum	41	1	2022-12-02 06:37:40

Tags Table

1

SELECT * FROM ig_clone.tags;

id	tag_name	created_at
1	sunset	2022-12-02 06:37:40
2	photography	2022-12-02 06:37:40
3	sunrise	2022-12-02 06:37:40
4	landscape	2022-12-02 06:37:40
5	food	2022-12-02 06:37:40
6	foodie	2022-12-02 06:37:40
7	delicious	2022-12-02 06:37:40
8	beauty	2022-12-02 06:37:40

**Note – I've not expanded 'Follows Table' here as it was nowhere used during this analysis

IV Insights

A Marketing: The Marketing team wants to launch some campaigns, and they need to gather some insights to successfully execute them.


1. Rewarding Most Loyal Users

Objective: Finding the 5 oldest users of instagram

Query

```
select username
from (SELECT username, created_at,
      dense_rank() over (order by created_at) as rnk
      FROM ig_clone.users) sub
where rnk <=5
```

Output



	username
▶	Darby_Herzog
	Emilio_Bernier52
	Elenor88
	Nicole71
	Jordyn.Jacobson2

Recommendation: These 5 users who have been using Instagram for years, could be rewarded by giving them free passes to some of the company sponsored events

IV Insights


2. Reminding Inactive Users to Start Posting

Objective: Finding users who have never posted a single photo

Query

```
1 • SELECT u.id, u.username
2 FROM ig_clone.users u
3 left join ig_clone.photos p
4 on u.id = p.user_id
5 where p.user_id is null
```

Output



id	username
5	Aniya_Hackett
7	Kassandra_Homenick
14	Jadlyn81
21	Rocio33
24	Maxwell.Halvorson

Recommendation: These 5 users have never posted a single photo on Instagram as of now. A few push notification on their phones would increase the chances of them posting a picture.

IV Insights


3. Declaring Contest Winner

Objective: Identify the photo which got liked the most

Query

```
select l.photo_id, p.user_id, count(l.photo_id) as total_likes
from ig_clone.likes l
join ig_clone.photos p
on l.photo_id = p.id
group by photo_id
order by total_likes desc
```

Output



photo_id	user_id	total_likes
145	52	48
127	46	43
182	65	43
123	44	42
30	10	41

Recommendation: The photo with id 145 which was posted by a User with id 52 got the maximum number of likes (i.e., 48 in total) and is the winner of the contest. The team can reward them by either posting t

IV Insights


4. Hashtag Researching

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

Query

```
1  SELECT p.tag_id, count(p.tag_id) as totalTagCount
2  FROM ig_clone.photo_tags p
3  JOIN ig_clone.tags t
4  ON p.tag_id = t.id
5  GROUP BY p.tag_id
6  ORDER BY totalTagCount desc
7  Limit 5
```

Output



tag_id	totalTagCount
21	59
20	42
17	39
13	38
18	24

Recommendation: These are the top 5 most commonly used hashtags and the Instagram can suggest their partner brand to add these on their new posts to reach out to most people who are using the platform

IV Insights


5. Launch Advertisement Campaign

Objective: Identify the days where most users register on

Query

```
SELECT dayname(created_at) as dayname , count(dayname(created_at)) as cnt
FROM ig_clone.users
group by dayname
order by cnt desc
```

Output



dayname	cnt
Thursday	16
Sunday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12

Recommendation: Thursdays and Sundays were the days when most users registered and these days would be the most appropriate days to launch Advertisements in order to reach the maximum audien

IV Insights

B Investor Metrics: The Investors want to know if Instagram is performing well and is not becoming redundant like Facebook, so they want to assess the app on some grounds


1.1 User Engagement

Objective: i. How many times does an average user post on Instagram

Query

```
1 • select avg(cnt) as average
2   from (
3       SELECT user_id, count(*) as cnt
4       FROM ig_clone.photos
5       group by user_id
6       order by cnt desc ) sub
```

Output



average
3.4730

Recommendation: On average, a user posts 3-4 photos in a year on Instagram which is not very sustainable and it clearly maps out that the user engagement is not appropriately high. This could be improved by adding new features, custom AD targetting, and organizing creative contests for the users

IV Insights


1.2 User Engagement

Objective: ii. Total number of photos on Instagram/Total number of users

Query

```
select count(distinct p.id) / count(distinct u.id) as perc  
from ig_clone.users u  
left join ig_clone.photos p  
on u.id = p.user_id
```

Output



	perc
	2.5700

Insights: There were a total of 257 photos and 100 users who were using Instagram. For every single user, on an average there are 2-3 photos posted on Instagram

IV Insights


2. Bots & Fake Accounts

Objective: Identify the users (bots) who have liked every single photo on Instagram (any normal user would not be able to do that)

Query

```
1 •  SELECT user_id, count(*) as likeCount
2    FROM ig_clone.likes
3   group by user_id
4  having count(*) = 257
5  order by likeCount desc
```

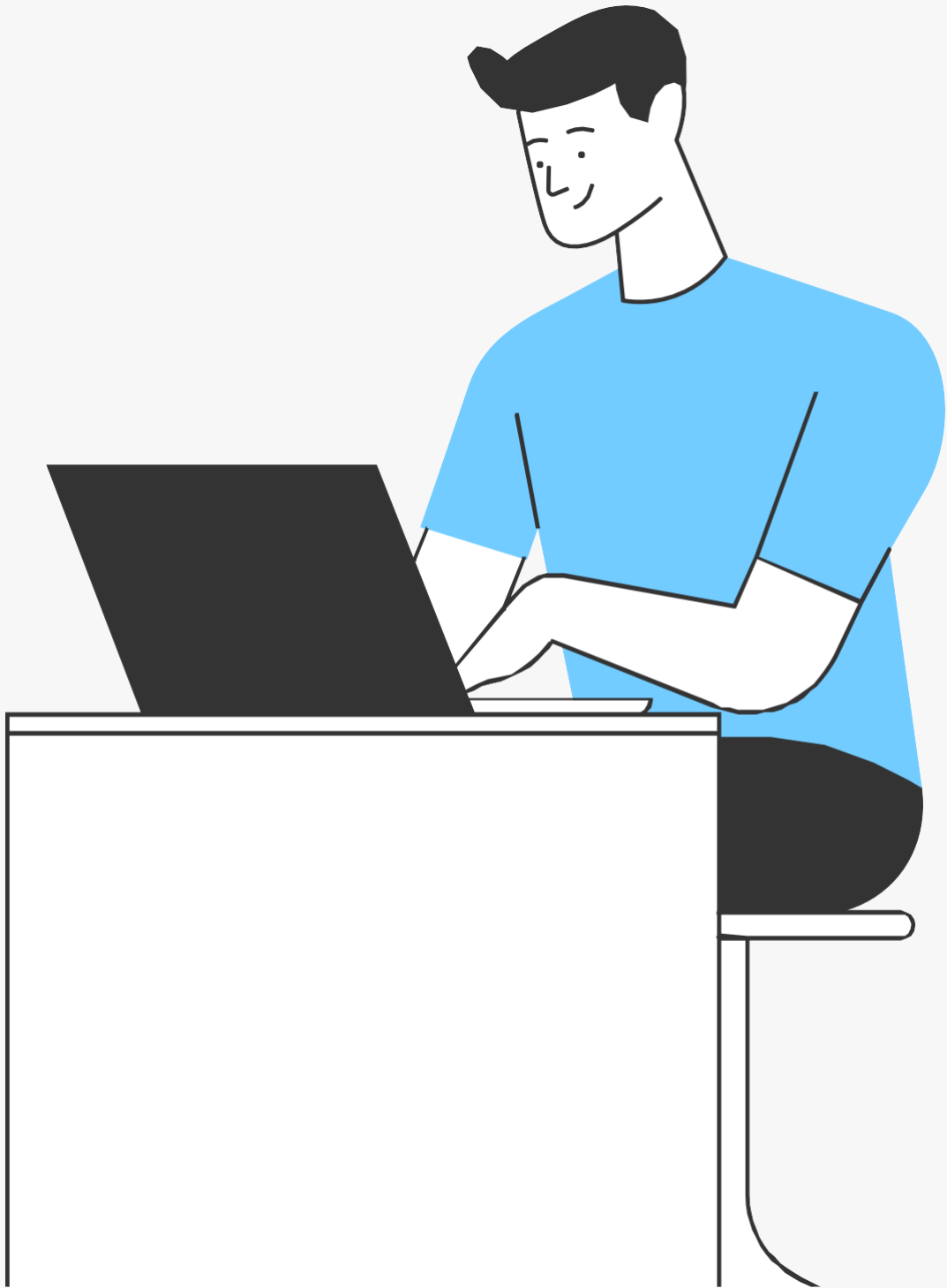
Output



user_id	likeCount
5	257
14	257
21	257
24	257
36	257
41	257
54	257
57	257
66	257
71	257
75	257
76	257
91	257

Recommendation: A total of 257 photos were posted by the users on Instagram. Since it's not possible for a human user to like every single photo posted, 13users (bots) were identified who liked every single photo posted on the platform. Using the user id of these bots, they can be removed from the platform to make Instagram a spam free community

IV Insights



Key Insights

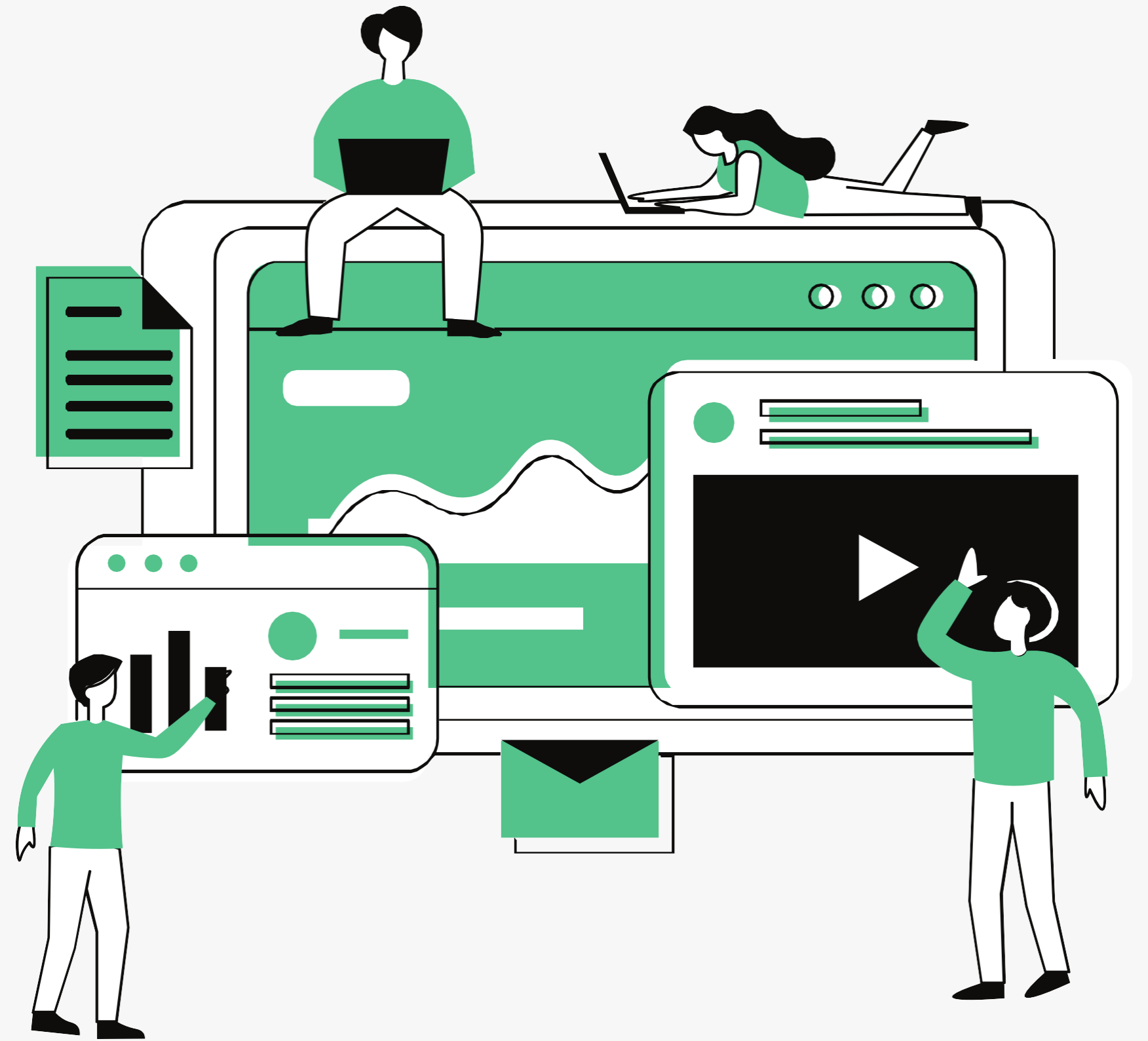
- The top 5 oldest customers who are considered to be the most loyal ones created their accounts in the month of May. Darby out of them was the first one to register
- There are 5 users who have never posted a single picture on Instagram as of yet. They could be sent some personalized messages to attract them to post
- There are 13 bots on the platform that could fill the app with spam and irrelevant content.
- It was found that Thursdays and Sundays were the days when most users registered.
- On an average, a user was posting 3-4 photos in a year.

V Result

- The key findings of the study would be the identification of the most loyal customers, inactive users, and primary hashtags which could be crucial for the strategy execs to take a data-driven decision.
- Thursday being the day when most people register on the platform was a little unusual according to me.
- I learned that a lot of metrics other than the monetary ones could have a huge impact on the business decision making process



Thank you!



<https://drive.google.com/drive/folders/1tk6jKHpyzyKUfn86bIXt3xBbRyn-oVZoI>