Instagram User Analytics

Project Description

This project is all about finding insights about Instagram and its users from the collected dataset. In this analytics report we track how users of Instagram engage and interact with Instagram in an attempt to derive business insights for marketing, product and development teams.
These insights are definitely useful for teams across the business to launch a new marketing campaign, decide on features to build for an app Instagram, track the success of the app by measuring user engagement and improve the experience altogether while helping the business grow.

Things I need to find –

1. Some oldest and loyal users of instagram.
2. Finding people who never posted anything.
3. Finding users who get most likes on a single posted photo.
4. Trending Hashtags.
5. Most crowded day in a week, when users are heavily engaged on Instagram.
6. Identifying bots and fake accounts.

Approach

In approaching a data analytics project using SQL, I will try to use the power of SQL to extract meaningful insights from the available dataset. I begin with knowing the project objectives and the questions I have to answer. I design appropriate SQL queries, aggregations and filtering conditions to get the required insights.Throughout the process, I ensure adherence to data security and privacy protocols to maintain the confidentiality of sensitive information.

Tech-Stack used

1. MySQL and Mysql workbench - to execute sql queries.

2. Google documents - to prepare a final report.

Insights

Users are using Instagram very oftenly but there are some id's who have never been as active. So we should send them promotional mail for posting their first photo. We get to know about the time where maximum people get engaged to Instagram and what they like to consume as a user.

Report

We get to conclude about the question we aim to solve in this analysis, we will get to the result one by one ,

Marketing Insights and results :-

Top 5 most loyal users:-

Sql command used :-

```
select id,username,created_at from users order by created_at asc limit 5;
```

These are the oldest users on Instagram

User_id	User_name	Instagram joining time
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.Jacobson 2	2016-05-14 07:56:26

Inactive users on Instagram:-

Sql command used :-

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

These users never post any photo till date

User Id	User name	User Id	User name
5	Aniya_Hackett	57	Julien_Schmidt
7	Kasandra_Homenick	66	Mike.Auer39
14	Jaclyn81	68	Franco_Keebler64
21	Rocio33	71	Nia_Haag
24	Maxwell.Halvorson	74	Hulda.Macejkovic
25	Tierra.Trantow	75	Leslie67
34	Pearl7	76	Janelle.Nikolaus81
36	Ollie_Ledner37	80	Darby_Herzog
41	Mckenna17	81	Esther.Zulauf61
45	David.Osinski47	83	Bartholome.Bernhard
49	Morgan.Kassulke	89	Jessyca_West
53	Linnea59	90	Esmeralda.Mraz57
54	Duane60	91	Bethany20

So we must mail them to step into the world of instagram and share their posts for better reach from the audience and create a growing environment .

Contest winner :-

Sql command used :-

```
SELECT photo_id, COUNT(photo_id) AS occurrence_count
FROM likes
GROUP BY photo_id
HAVING COUNT(photo_id) = (SELECT MAX(count_result) FROM (SELECT
COUNT(photo_id) AS count_result FROM likes GROUP BY photo_id) AS
subquery);
```

The team started a contest and the user get most likes on a single photo will win

And the winner is :-

User id - 52
User name - Zack_Kemmer93
Photo id - 145
Photo_url -'https://jarret.name'

5 most commonly used hashtag are :-

Sql command used :-

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

The top 5 most used hashtags are :-

Serial number	tag_name	Number of times used

1	smile	59
2	beach	42
3	party	39
4	fun	38
5	concert	24

Which day would be the best day to launch ADs :-

Sql command used

```
# for getting day

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

# for getting hour

SELECT DAYNAME(created_at) AS registration_day,
      HOUR(created_at) AS registration_hour,
      COUNT(*) AS registration_count
FROM users
GROUP BY registration_day, registration_hour
ORDER BY registration_count DESC;
```

These are the days and times where Instagram got most of the user registration.

Day	Time	Number of user register
Tuesday	12.00 pm	3
Sunday	05.00 pm	3

Friday	07.00 pm	3
Wednesday	05.00 pm	3
Friday	11.00 pm	3

Insights :

As we can see from the table, in these days the maximum number of user does register to Instagram,

But On the Friday we have the max count 3 registration at 7 pm and 3 more at 11 pm which give a total of 6 users registered in a single day,

So Friday would be a profitable day to launch ADs.

Investor metrics Insights and results :-

User Engagement :-

Sql command used

```
# total number photos post by users

select count(id) as total_number_of_photos from photos;


# total number of users

select count(id) as total_number_of_user from users;
```

The number of time average user post on Instagram

Result :-
Total_number_of_photos = 257
Total_number_of_user = 100

Average post an user make in Instagram =

total_number_of_photo / total_number_user

= 257 / 100 = 2.57

Bots and fake accounts :-

Sql command used

```
SELECT user_id
FROM likes
GROUP BY user_id
HAVING COUNT(DISTINCT photo_id) = (SELECT COUNT(DISTINCT photo_id) FROM
likes);
```

Someone who did like every single photo

Serial number	User_id	User_name
1	5	Aniya_Hackett
2	14	Jaclyn81
3	21	Rocio33
4	24	Maxwell.Halvorson
5	36	Ollie_Ledner37
6	41	Mckenna17
7	54	Duane60
8	57	Julien_Schmidt
9	66	Mike.Auer39
10	71	Nia_Haag
11	75	Leslie67
12	76	Janelle.Nikolaus81
13	91	Bethany20

Conclusion :-

The analysis brought useful insights from the data which, I believe, must be used in the upgrowth of the company and increasing user experience and sharing more entertaining content with creating a more safe and secure environment.