# INSTAGRAM USER ANALYTICS

Project Report

# CONTENTS

# PROJECT DESCRIPTION

The project will provide insight into the questions asked by the management team. All the questions asked are answered here in this project report.

Through the project, we will find out

- when we should launch campaigns.

- If Instagram is performing well or becoming redundant like other social media platforms.

# APPROACH

Firstly, Learnt the basics required to complete the project i.e. SQL for this case. After learning, made a plan to complete the project i.e. devise a step-by-step plan, and divide the project accordingly. Now implemented the plan by working on the project. Monitored the results we were getting from the project. Hence finally completed the project

# TECH-STACK USED

The software used while doing the project is Oracle MySQL Workbench

I have used workbench because it is easy to execute queries in workbench.
And one can even import databases into it easily.

# INSIGHTS

# MARKETING TEAM

The marketing team wants to launch some campaigns; The following insights will help them with the task.

# REWARDING MOST LOYAL USERS

Id, username and date of creation of the account of the 5 most loyal Instagram users are found in the database.

They can be rewarded for being associated with Instagram for the longest time. It will attract other people to open their Instagram account.

```sql
select * from users order by created_at 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 |

# REMIND INACTIVE USERS TO START POSTING

We can send promotional emails to the 26 users mentioned here and ask them to post their first photo as they have yet to post any picture.

```
select * from users where not exists(select * from photos where users.id=user_id);
```

| id | username | created_at |
|---|---|---|
| 5 | Aniya_Hackett | 2016-12-07 01:04:39 |
| 7 | Kasandra_Homenick | 2016-12-12 06:50:08 |
| 14 | Jaclyn81 | 2017-02-06 23:29:16 |
| 21 | Rocio33 | 2017-01-23 11:51:15 |
| 24 | Maxwell.Halvorson | 2017-04-18 02:32:44 |
| 25 | Tierra.Trantow | 2016-10-03 12:49:21 |
| 34 | Pearl7 | 2016-07-08 21:42:01 |
| 36 | Ollie_Ledner37 | 2016-08-04 15:42:20 |
| 41 | Mckenna17 | 2016-07-17 17:25:45 |
| 45 | David.Osinski47 | 2017-02-05 21:23:37 |
| 49 | Morgan.Kassulke | 2016-10-30 12:42:31 |
| 53 | Linnea59 | 2017-02-07 07:49:34 |
| 54 | Duane60 | 2016-12-21 04:43:38 |
| 57 | Julien_Schmidt | 2017-02-02 23:12:48 |
| 66 | Mike.Auer39 | 2016-07-01 17:36:15 |
| 68 | Franco_Keebler64 | 2016-11-13 20:09:27 |
| 71 | Nia_Haag | 2016-05-14 15:38:50 |
| 74 | Hulda.Macejkovic | 2017-01-25 17:17:28 |
| 75 | Leslie67 | 2016-09-21 05:14:01 |
| 76 | Janelle.Nikolaus81 | 2016-07-21 09:26:09 |
| 80 | Darby_Herzog | 2016-05-06 00:14:21 |
| 81 | Esther.Zulauf61 | 2017-01-14 17:02:34 |
| 83 | Bartholome.Bernhard | 2016-11-06 02:31:23 |
| 89 | Jessyca_West | 2016-09-14 23:47:05 |
| 90 | Esmeralda.Mraz57 | 2017-03-03 11:52:27 |
| 91 | Bethany20 | 2016-06-03 23:31:53 |
| NULL | NULL | NULL |

# DECLARING CONTEST WINNER

The person with the highest number of likes on a single picture must be declared the constant winner.

It can be inferred from the query that a user with the username Zack_Kemmer93 has the highest number of likes and must be declared as the contest winner.

```sql
select u.id,u.username,l.photo_id,p.image_url,count(l.user_id) as number_of_likes,u.created_at,p.created_dat
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 l.photo_id
order by count(l.user_id) desc limit 1 ;
```

| id | username | photo_id | image_url | number_of_likes | created_at | created_dat |
|----|----------|----------|-----------|-----------------|------------|-------------|
| 52 | Zack_Kemmer93 | 145 | https://jarret.name | 48 | 2017-01-01 05:58:22 | 2023-01-30 19:24:12 |

# HASHTAG RESEARCHING

We need to find the hashtags used the maximum number of times on the platform.

Here are the five most commonly used hashtags

We can infer from the query that smile is the hashtag which is used most number of times, followed by beach, party, fun and concert

```sql
select tag_name
from photo_tags inner join tags on id=tag_id
group by tag_id
order by count(photo_id) desc limit 5;
```

| tag_name |
|----------|
| smile |
| beach |
| party |
| fun |
| concert |

# LAUNCH AD CAMPAIGN

The best day to launch an ad campaign would be the one on which the maximum number of customers register

Here 1 is Sunday, and 7 is Saturday.

As we can infer from the query result that most users register on Thursday or Sunday therefore add campaigns must be launched during these days.

```sql
select dayofweek(created_at) as days,count(username) as number_of_users_registered
from users
group by days
order by count(username) desc;
```

| days | number_of_users_registered |
|------|---------------------------|
| 5    | 16                        |
| 1    | 16                        |
| 6    | 15                        |
| 3    | 14                        |
| 2    | 14                        |
| 4    | 13                        |
| 7    | 12                        |

# INVESTOR METRICS

investors wanted to know if Instagram is performing well and is not becoming redundant. The following insights will help them assess the app.

# USER ENGAGEMENT

To check if the users are still active on Instagram, we calculated how much the average user posts on the app.

We also calculated the total number of users on Instagram and the total number of photos on Instagram.

o   An average user posts approx. 3 to 4 photos.

o   The total number of users on Instagram is 100

o   The total number of photos is 257

```
select avg(c) as average_users_posts from (select count(*) as c from photos group by user_id) as average;
select count(id) as number_of_photos from photos;
select count(id) as number_of_users from users;
```

| average_users_posts |
|---------------------|
| 3.4730              |

| number_of_photos |
|------------------|
| 257              |

| number_of_users |
|-----------------|
| 100             |

# BOTS & FAKE ACCOUNTS

Fake accounts would be the one which likes every photo on Instagram as any normal user would not do anything of this sort.

There were in total 13 accounts on the site which liked every photo present on the site

```
select * from users
where id in (select user_id
                from (select user_id,count(distinct photo_id) as cpid
                    from likes group by user_id) as no_of_likes where cpid=(select count(*) from photos));
```

| id | username | created_at |
|---|---|---|
| 5 | Aniya_Hackett | 2016-12-07 01:04:39 |
| 14 | Jaclyn81 | 2017-02-06 23:29:16 |
| 21 | Rocio33 | 2017-01-23 11:51:15 |
| 24 | Maxwell.Halvorson | 2017-04-18 02:32:44 |
| 36 | Ollie_Ledner37 | 2016-08-04 15:42:20 |
| 41 | Mckenna17 | 2016-07-17 17:25:45 |
| 54 | Duane60 | 2016-12-21 04:43:38 |
| 57 | Julien_Schmidt | 2017-02-02 23:12:48 |
| 66 | Mike.Auer39 | 2016-07-01 17:36:15 |
| 71 | Nia_Haag | 2016-05-14 15:38:50 |
| 75 | Leslie67 | 2016-09-21 05:14:01 |
| 76 | Janelle.Nikolaus81 | 2016-07-21 09:26:09 |
| 91 | Bethany20 | 2016-06-03 23:31:53 |
| NULL | NULL | NULL |

# RESULT

By doing this project, I have answered all the questions asked by the management team and investors' metrics.

It helped me understand the basics of data analytics and how it can help us answer big questions. It made me learn SQL and solve problems using it.

# THANK YOU

By Anjali Gupta