A) Marketing Analysis:

-- 1). Identifying the 5 oldest users on Instagram from the provided database (to reward them as a most loyal users). --

```
use ig_clone;
select username, created_at from users
order by created_at
limit 5 ;
```

➔ Output:

username	id	created_at
Darby_Herzog	80	06-05-2016 00:14
Emilio_Bernier52	67	06-05-2016 13:04
Elenor88	63	08-05-2016 01:30
Nicole71	95	09-05-2016 17:30
Jordyn.Jacobson2	38	14-05-2016 07:56

-- By sorting users according to the date on which they created account, we get the oldest users data and hence we can award them. --

-- 2). Identifying the users who have never posted a single photo on Instagram from provided database: --

```
select * from users, photos;
select username from users
left join photos
on users.id = photos.user_id
where photos.id is null
order by username;
```

➔ Output:

username
Aniya_Hackett
Bartholome.Bernhard
Bethany20
Darby_Herzog
David.Osinski47

Duane60
Esmeralda.Mraz57
Esther.Zulauf61
Franco_Keebler64
Hulda.Macejkovic
Jaclyn81
Janelle.Nikolaus81
Jessyca_West
Julien_Schmidt
Kasandra_Homenick
Leslie67
Linnea59
Maxwell.Halvorson
Mckenna17
Mike.Auer39
Morgan.Kassulke
Nia_Haag
Ollie_Ledner37
Pearl7
Rocio33
Tierra.Trantow

-- By sorting out the null value, we got the list of users who never posted a single photo. --

-- 3). Identifying the users with most likes on any single photo for the contest from the provided database: --

```
select * from photos,users,likes;
select username, photos.id, photos.image_url, count(*) as total_likes 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_likes desc
limit 1 ;
```

➔ Output:

username	id	image_url	total_likes
Zack_Kemmer93	145	https://jarret.name	48

-- First we sorted out the total number of likes then arranged them in a descending order and limited our order to one. This way we got the user with most likes on a single post hence we got our winner: 'Zack_Kemmer93'. --

-- 4). Identifying the top five most commonly used hashtags on Instagram: --

```
select * from photo_tags, tags;
select tag_name, count(photo_id) as hashtags from photo_tags
inner join tags on tags.id = photo_tags.tag_id
group by tag_name
order by hashtags desc;
```

➔ Output:

tag_name	hashtags
smile	59
beach	42
party	39
fun	38
concert	24
food	24
lol	24
hair	23
happy	22
beauty	20
dreamy	20
drunk	19
fashion	19
sunset	19
landscape	17
style	17
sunrise	17
photography	16
stunning	16
delicious	15
foodie	11

-- To reach the most people, one should use these 5 most popular tags: 'smile', 'beach', 'party', 'fun' and 'concert'. --

-- 5). Identifying the day of the week when most users registered on Instagram: --

```
select * from users;
select dayname(created_at) as created_day, count(*) as total_reg from users
group by created_day
order by total_reg desc;
```

➔ Output:

created_day	total_reg
Thursday	16
Sunday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12

-- After sorting out and analysis the registration, best day to launch the ad campaign is either 'Thursday' or 'Sunday'. --

B). Investor Metrics:

-- 1). Calculating the average number of posts per user on Instagram and finding total number of photos on Instagram divided by the total number of users: --

```
select count(*) as total_photos_on_ig from photos;
```


```
select count(*) as total_users_on_ig from users;
```


```
select (select count(*) from photos) / (select count(*) from users) as average;
```

➔ Output:

-- By performing the division, we got the average post per user on Instagram that will help analysing the active users –

-- 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 username, count(*) as fake_likes from users
inner join likes on users.id = likes.user_id
group by likes.user_id
having fake_likes = (select count(*) from photos);
```

➔ Output:

username	fake_likes
Aniya_Hackett	257
Jaclyn81	257
Rocio33	257
Maxwell.Halvorson	257
Ollie_Ledner37	257
Mckenna17	257
Duane60	257
Julien_Schmidt	257
Mike.Auer39	257
Nia_Haag	257
Leslie67	257
Janelle.Nikolaus81	257
Bethany20	257

-- Here we collected the data about same number of likes done by users which is unusual hence we got the list of bots. –

Project description:

The project is based on Instagram user analytics. In this project, we gathered and sorted and collected the information asked by using provided Instagram database.

This project will help us to understand and analyse how users behaves, interacts and follows to get a data driven insights and big picture of Instagram users which will help us to predict the areas of improvement and the changes required for the betterment of the app.

Approach:

I first understood the need, then use the proper syntax and formulas to execute the precise and accurate result. I moved step by step gaining information for required columns and tables in each question in a structured way, then I applied queries based on what we have taught and instructed. I got errors but it only helped me improving my coding skills.

Tech-stack used:

MySQL Workbench,

Version: 8.0.34 (MySQL Community Server – GPL),

Compiled for: Win64 (x86_64)

The purpose of using MySQL was to extract and execute the necessary queries on tables, from the provided database, to get the accurate result and derive a precise insights that has been asked from the team.

Insights:

While working on project, I got to learn the different methods, how to extract the data, how to handle it, how to work on tables and how to use database to get a desired results using proper queries in a structured way. The insights which asked in this project will help an analyst to improve the app features and balance variations in upcoming updates, marketing of the product and understand the changes which will come and required with time.

We were provided with two tasks which includes marketing analysis and investor metrics. In market analysis we needed to find the five oldest Instagram users, users who never posted a single photo, user with the most likes, top five commonly used hashtags and the day of a week where new users registered the most. For invertor metrics, we needed to find if users are still active and posting on Instagram or if they are making fewer posts and users (potential bots) who have liked every single photo on the site, as this is not typically possible for a normal user.

After getting the results I got to know the implementations that we can derive from these questions asked, the insights are very helpful to shape users behaviour.

Results:

While working on this project, my accomplishments are: I was able to performs tasks, derive insights from a given set of database, I can create-modify and extract the necessaries using MySQL queries in structured way.

THANK YOU