Project Description

The project is Instagram Analytics and is mainly about gaining insights from the Instagram data. The data is available in text format and we need to perform all the operations from data loading to finally gaining insights.

Approach

As data was available in plane text, first task was to load data into any SQL software to query the data to gain insights so I used DB-Fiddle to load the data and perform querying on it.

Tech-Stack Used

I used Db-Fiddle software to load and query the data as Mysql workbench was giving errors while installing on my laptop.

Insights

I will share the insights gained by question asked in the assignment.

1) Find the 5 oldest users of the Instagram from the database provided

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

Writing above query I gained information about 5 oldest users on platform.

2) Find the users who have never posted a single photo on Instagram

```
select u.username,p.image_url from users u left join photos p on u.id=p.user_id
where p.user_id is null;
```

Writing above query I gained information about users who have never posted.

3) Identify the winner of the contest and provide their details to the team

```
select u.username,p.id,photo_likes.like_count from photos p inner join
    (select photo_id, COUNT(*)as like_count from likes group by photo_id) as photo_likes
    inner join users u on u.id=p.user_id  order by like_count desc limit 1;
```

Writing above query I gained information about winner of contest and their details.

4) Identify and suggest the top 5 most commonly used hashtags on the platform

```
select tag_name,count(*) as tag_count from photo_tags inner join tags on
photo_tags.tag_id=tags.id group by tag_id  order by tag_count desc limit 5;
```

Writing above query I gained information about top 5 mostly used hashtags.

5) What day of the week do most users register on? Provide insights on when to schedule an ad campaign.

```
select dayofmonth(created_at) as days,count(*) as day_count from users
group by days
order by day_count desc
limit 1;
```

Gained insights for on which day to launch campaign by querying the data.

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

```
6a) select count(image_url)/ count(distinct(user_id)) as avg_user_posts from photos
group by user_id;
6b) select count(image_url)/ count(distinct(user_id)) as
total_photos_by_total_users_ratio from photos;
```

Writing above query I gained information about how many times average user posts on Instagram also ratio of total no of photos by total no of users.

7) 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).

```
select user_id , count(photo_id) from likes group by user_id order by 2 desc limit 13 ;
```

I gained insights about bots and identified the all bots.

Result

I achieved all the results required by querying the data.

Drive Link

https://drive.google.com/drive/folders/1A6xVY5S91lffNFiaicI88yGDvN2Y7Rlu?usp=sharing