Project – 2

Instagram User Analytics

- **Project Description –**

In this age of social media, we are trying to understand the user's behaviour for our marketing strategies. To project better profits and engagement of users. We analysed users' registration patterns to launch our ad campaigns, identified bot and fake accounts by analysing inactive users, also identified which hashtags are popular to increase the engagement of users. We also took some initiative to announce a user who has the most likes on their post and rewarded some users who have been using the platform and are very active on consistent basis.

- **Approach –**

1) Data collection - Created a database of user's information
2) Identify User behaviour – Utilized SQL queries or data analysis process to identify fake accounts, most active users, loyal users, Popular hashtags, User engagement.
3) Reporting and Monitoring - Created a report based on data analysis

- **Tech-Stack Used –**

1) MySQL Workbench 8.0,(SQL Server Management Studio)
   MySQL workbench is a powerful and widely used tool for database design, database management and data analysis. It offers variety tools for data analysis, data management which align well with tasks of this project.
2) MS word to generate report

A) Marketing Analysis:

1. Loyal User Reward: Identify the five oldest users on Instagram

SQL Syntax –

```
SELECT * FROM users
ORDER BY created_at
LIMIT 5;
```

Output –

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

2 . Inactive User Engagement: Identify users who have never posted a single photo on Instagram.

SQL Syntax –

```
SELECT username
FROM users
LEFT JOIN photos ON users.id = photos.user_id
WHERE photos.id IS NULL;
```

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

3 . Contest Winner Declaration: Determine the winner of the contest and provide their details to the team.

SQL Syntax –

```
SELECT
  username,
  photos.id,
  photos.image_url,
  COUNT(*) AS total
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 DESC
LIMIT 1;
```

Output -

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

4. Hashtag Research: Identify and suggest the top five most commonly used hashtags on the platform.

SQL Syntax –

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

Output –

	tag_name	total
▶	smile	59
	beach	42
	party	39
	fun	38
	food	24

5. Ad Campaign Launch: Determine the day of the week when most users register on Instagram. Provide insights on when to schedule an ad campaign.

SQL Syntax –

```
select date_format(created_at, '%W') as Day_of_Week,
count(*) as total
from  users
Group By DAY_OF_week
Order by total desc
Limit 1;
```

Output –

	Day_of_Week	total
▶	Thursday	16

B) Investor Metrics:

1. **User Engagement:** Calculate the average number of posts per user on Instagram. Also, provide the total number of photos on Instagram divided by the total number of users.

SQL Syntax –

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

Output –

	avg
▶	2.5700

2. **Bots & Fake Accounts:** identify users (potential bots) who have liked every single photo on the site, as this is not typically possible for a normal user.

SQL Syntax –

```
select username, count(*) as Number_of_likes from users
INNER JOIN likes
on likes.user_id= users.id
Group By likes.user_id
Having Number_of_likes = (select Count(*) from photos);
```

Output –

	username	Number_of_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

Insights –

1) From derived results we found Loyal active users, fake and bot accounts which helped us to maintain the integrity of platform and investors perception.
2) It helps to understand how quality of data can manipulate the results of a task and increase awareness of quality control
3) This analysis can help to shift the focus to user-centric approach for better data driven decisions.

Report –

- The outcome of this Data analysis is a list of potential bot or fake accounts based on their liking patterns.
- This information is important for maintaining the integrity of engagement metrics and providing accurate data to the investors.
- This analysis helped to find loyal and active user to reward them to retain the active user metrics.
- Popular hashtags and trends help to increase the engagement of the users.
- Most active time and days help to launch ad campaigns to get the most profit.