INSTAGRAM USER ANALYTICS

1.1 PROJECT DESCRIPTION

The primary goal of the project is to analyze and strategize the user responses to various marketing tactics implemented by the team of Instagram. With users in humungous numbers, it is a tedious task to skim through every individual's record and activity. For the ease of the stated purpose, a database management tool that efficiently supports query management is used to store data systematically in tables and retrieve meaningful insights as per the required conditions through queries. These insights provide actionable recommendations for enhancing user engagement.

1.2 APPROACH

The approach involves storing the user data generated from Instagram application in the form of systematic tables related to each other in a MySQL database and using MySQL Workbench to convert the stored raw data into meaningful insights that can be used to form profitable strategies for the organization.

1.3 TECH STACK USED

MySQL is Relational Database Management tool used for storing the raw data generated from the social media platform. MySQL Workbench Version 8.0.36 provides a user-friendly GUI that makes query processing effortless. Designing and modeling databases as well as writing and SQL queries becomes extremely simple for users who prefer a visual approach to database management.

1.4 INSIGHTS

1.4.1 Marketing Analysis

⇨ **Loyal User Reward:**
- o The marketing team wants to reward the most loyal users. i.e., those who have been using the platform for the longest time.
- o **Code:**

```
97      #Q1. Loyal User Award
98  ●   SELECT
99          *
100     FROM
101         users
102     ORDER BY created_at
103     LIMIT 5;
```

- o **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
* NULL	NULL	NULL

- o **Inference:**
  - ▪ These are the top 5 users that have used the application since its creation. Rewarding them with goodies or small but interesting privileges will be an ethical leverage that will keep them and other fellow users attracted.
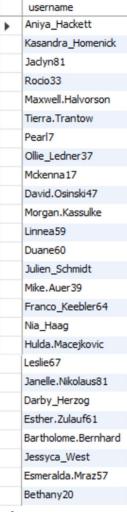
⇨ **Inactive User Engagement:**
- o The team wants to encourage inactive users to start posting by sending them promotional emails.
- o **Code:**

```
105     #Q2. Inactive User Engagement
106  ●  SELECT
107         username
108     FROM
109         users
110             LEFT JOIN
111         photos ON users.id = photos.user_id
112     WHERE
113         photos.id IS NULL;
```

- o **Output:**

username
Aniya_Hackett
Kasandra_Homenick
Jaclyn81
Rocio33
Maxwell.Halvorson
Tierra.Trantow
Pearl7
Ollie_Ledner37
Mckenna17
David.Osinski47
Morgan.Kassulke
Linnea59
Duane60
Julien_Schmidt
Mike.Auer39
Franco_Keebler64
Nia_Haag
Hulda.Macejkovic
Leslie67
Janelle.Nikolaus81
Darby_Herzog
Esther.Zulauf61
Bartholome.Bernhard
Jessyca_West
Esmeralda.Mraz57
Bethany20

- o **Inference:**
  - ▪ This list will assist the team to focus on encouraging the inactive users to get involved in the community by sending promotional mails and content based on their liking. Furthermore, this will also prove helpful in detecting trolls and spam accounts.

⇨ **Contest Winner Declaration:**
- o The team has organized a contest where the user with the most likes on a single photo wins.
- o **Code:**

```
115    #Q3 Contest Winner Declaration
116 •  SELECT
117        users.username,
118        photos.id,
119        photos.image_url,
120        COUNT(likes.user_id) AS total_likes
121    FROM
122        photos
123            INNER JOIN
124        likes ON likes.photo_id = photos.id
125            INNER JOIN
126        users ON photos.user_id = users.id
127    GROUP BY photo_id
128    ORDER BY total_likes DESC
129    LIMIT 1;
```

- o **Output:**

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

- o **Inference:**
  - ▪ Rewarding the user with the most likes on a single photo is an amazing promotional stunt that would not only help the winning creator increase their reach but also create a competitive atmosphere among the users of the platform and push them to create more content.

⇨ **Hashtag Research:**
- o A partner brand wants to know the most popular hashtags to use in their posts to reach the most people.
- o **Code:**

```
131      #Q4 Hashtag Research
132 ●    SELECT
133          tags.tag_name,
134          COUNT(tags.id) AS tags_use_count
135      FROM
136          tags
137              INNER JOIN
138          photo_tags ON tags.id = photo_tags.tag_id
139      GROUP BY id
140      ORDER BY tags_use_count DESC
141      LIMIT 5;
```

- o **Output:**

tag_name	tags_use_count
▶ smile	59
beach	42
party	39
fun	38
concert	24

- o **Inference:**
  - ▪ Once the currently most used tags are narrowed down, the brand can ask the creator they team up with to get maximum user indulgence and hence generate profitable revenue.

⇨ **Ad Campaign Launch:**
- o The team wants to know the best day of the week to launch ads.
- o **Code:**

```
143      #Q5 Ad Campaign Launch
144 ●    SELECT
145          DAYNAME(created_at) AS day, COUNT(*) AS total_acc_created
146      FROM
147          users
148      GROUP BY day
149      ORDER BY total_acc_created DESC
150      LIMIT 1;
```

- o **Output:**

	day	total_acc_created
▶	Thursday	16

- o **Inference:**
  - ▪ Ads generate revenue based on how many users engage with it. Launching the ads on the day when maximum accounts are being created, that is maximum user engagement is taking place will perfectly serve the purpose in question.

1.4.2 Investor Metrics

⇨ **User Engagement:**
- o Investors want to know if users are still active and posting on Instagram or if they are making fewer posts.
- o **Code:**

```sql
153     #Q1 User Engagement
154 ● ⊖ WITH base AS (
155         SELECT users.id AS usersid, COUNT(photos.id) AS photosid
156         FROM users
157         LEFT JOIN photos ON photos.user_id = users.id
158         GROUP BY users.id
159     )
160     SELECT
161         SUM(photosid) AS tot_photos,
162         COUNT(usersid) AS tot_users,
163         SUM(photosid) / COUNT(usersid) AS photos_per_user
164     FROM base;
```

- o **Output:**

	tot_photos	tot_users	photos_per_user
▶	257	100	2.5700

- o **Inference:**

  To detect the extent of activity by the users on Instagram, the team first calculates the total photos posted and the total users on the platform. The average of photos posted per user helps estimate the participation of the users. This can help the investors understand how they can make the platform more convincing and captivating for better engagement.

⇨ **Bots & Fake Accounts:**
  o Investors want to know if the platform is crowded with fake and dummy accounts.
  o **Code:**

```
167     #Q2 Bots and Fake Accounts
168  •⊖ WITH base AS (
169         SELECT
170             users.username,
171             COUNT(likes.photo_id) AS tot_likes
172         FROM
173             likes
174         INNER JOIN
175             users ON users.id = likes.user_id
176         GROUP BY
177             users.username
178     )
179     SELECT
180         username,
181         tot_likes
182     FROM
183         base
184     WHERE
185         tot_likes = (SELECT COUNT(*) FROM photos)
186     ORDER BY
187         username;
```

  o **Output:**

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

  o **Inference:**
  It is practically impossible for a genuine user to like every single post on the platform unless it is a bot. Based on this logic, the users who have liked every single photo on the social media platform is deemed fake or a bot. Actions like restriction or blocking of the account can be done to maintain authenticity of the community.

1.5 RESULTS

The user analysis generated helps in assessing various aspects across multiple dimensions. From a marketing team's point of view, analysis of parameters like active users, loyal users, most used hashtags, the days with maximum engagement and further rewarding the fitting users help publicize the platform. Having insights about the most used hashtags can provide exclusive strategic guidance to the marketing team. Launching an ad campaign on the best day of the week can help the partner company have a better reach. These factors prove to be hefty contributors to the revenue chart. From an investor's point of view, information about fake accounts, bots and even inactive users helps evaluate user engagement, irradicate degradation of the quality of content and aids in maintaining the safety and authenticity of the space. These results help form informed decisions that has resulted in turning Instagram into a globally celebrated social media platform.