Project-2 “Instagram User Analytics”

A.) **Project Description:** This project aims to extract useful insights from raw data/metadata, using various database management tools, and even visualize them to increase the platform’s efficiency.

B.) **Project Approach:** The project was executed using SQL, where queries were utilized to create a database from the provided raw data. Sorting and data Extracting queries were then implemented to obtain the required data/insights.

C.) **Tech Stack Used:** The Tech Stack used included MySQL Workbench v8.0.34, which was an excellent tool for querying the database, thanks to its ease of access, simple setup, and GUI, as well as its troubleshooting support.

Project Insights: (Raw Insights) –

A.) Marketing Analysis

1.) **Rewarding Most Loyal Users:** People who have been using the platform for the longest time.
Conclusion: These are the oldest user of Instagram.

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

CODE:

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

2.) **Remind Inactive Users to start posting:** By sending them promotional emails to post their 1st photo.
Conclusion: These users were inactive after their first post.

 # username

5 Aniya_Hackett
7 Kasandra_Homenick
14 Jaclyn81
21 Rocio33
24 Maxwell.Halvorson
25 Tierra.Trantow
34 Pearl7
36 Ollie_Ledner37
41 Mckenna17
45 David.Osinski47
49 Morgan.Kassulke
53 Linnea59

54 Duane60
57 Julien_Schmidt
66 Mike.Auer39
68 Franco_Keebler64
71 Nia_Haag
74 Hulda.Macejkovic
75 Leslie67
76 Janelle.Nikolaus81
80 Darby_Herzog
81 Esther.Zulauf61
83 Bartholome.Bernhard
89 Jessyca_West
90 Esmeralda.Mraz57
91 Bethany20

CODE:

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

3.) **Declaring Contest Winner:** The team started a contest and the user who gets the most likes on a single photo will win the contest now they wish to declare the winner.
Conclusion: He has the most likes in his one post.

Username	id	Image_url	total
Zack_Kemmer93	145	https://jarret.name	48

CODE:

```
SELECT username, photos.id, photos.image_url, COUNT(likes.user_id) 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;
```

4.) **Hashtag Researching:** A partner brand wants to know, which hashtags to use in the post to reach the most people on the platform.
Conclusion: These are some trending hashtags which a partner brand can use.

Tag_name	total
smile	59
beach	42
party	39
Fun	38
Concert	24

CODE:

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

5.) **Launch AD Campaign:** The team wants to know, which day would be the best day to launch Ads.
Conclusion: These days would be best for AD Campaign.

Day	Total
Thursday	16
Sunday	16

CODE:

```
Select
    DAYNAME(created_at) AS day, COUNT(*) as total
    FROM users
GROUP BY day
ORDER BY total DESC
LIMIT 2;
```

B.) Investor Metrices

1.) **User Engagement:** Are users still as active and post on Instagram or they are making fewer posts.
Conclusion: A users average post is more than 2.

AVG
2.5700

CODE:

```
SELECT
(SELECT COUNT(*) FROM photos) / (SELECT COUNT(*) FROM users) AS AVG;
```

2.) **Bots & Fake Accounts:** The Investors want to know if the platform is crowded with fake and dummy accountd.
Conclusion: These are some users who can be bots and fake account.

Username	num_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

CODE:

```
SELECT user_id, COUNT(*) AS num_likes
FROM likes
GROUP BY user_id
HAVING num_likes = (SELECT COUNT(*) FROM photos);
SELECT U.username, COUNT(*) as num_likes
FROM users u
JOIN likes I ON u.id = I.user_id
GROUP BY u.id
HAVING num_likes = (SELECT COUNT(*) FROM photos);
```

RESULT

Through this project, we have successfully answered the questions posed by the marketing team and investors at Instagram. We have identified the oldest users, inactive users, and the winner of a contest. We have also provided insights on the most commonly used hashtags and the best day to launch ad campaigns.

The analysis and insights provided in this project will help Instagram make data-driven decisions to reward loyal users, encourage inactive users, and ensure a high level of user engagement. Additionally, investors will gain a better understanding of user activity and the presence of potential bots on the platform.

The impact of this analysis is significant, as it empowers Instagram to enhance user experience, target marketing campaigns effectively, and maintain the trust of investors.

By following the project description and leveraging SQL skills, we have provided valuable insights that can influence the future development of Instagram, one of the world's most popular social media platforms.