Instagram
User Analytics

Project Description:

The project focuses on User Analytics for Instagram where we track how users engage and interact with our digital product (software or mobile application) in an attempt to derive business insights for marketing, product & development teams.

To handle this project, a database is provided by Instagram will be used, containing user details, their post data and other engagement information. The SQL will be applied to query and manipulate the database to extract the insights.

These insights are then used by teams across the business to launch a new marketing campaign, decide on features to build for an app, track the success of the app by measuring user engagement and improve the experience altogether while helping the business grow.

The project is aimed to find answers to the questions given below:

A. **Marketing**: The marketing team wants to launch some campaigns, and they need your help with the following

1. *Rewarding Most Loyal Users*: People who have been using the platform for the longest time.

   Task: Find the 5 oldest users of the Instagram from the database provided

2. *Remind Inactive Users to Start Posting*: By sending them promotional emails to post their 1st photo.

   Task: Find the users who have never posted a single photo on Instagram

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.
   Task: Identify the winner of the contest and provide their details to the team

4. *Researching:* A partner brand wants to know, which hashtags to use in the post to reach the most people on the platform.

   Task: Identify and suggest the top 5 most commonly used hashtags on the platform

5. *Launch AD Campaign*: The team wants to know, which day would be the best day to launch ADs.
   Task: What day of the week do most users register on? Provide insights on when to schedule an ad campaign

B. **Investor Metrics**: Our investors want to know if Instagram is performing well and is not becoming redundant like Facebook, they want to assess the app on the following grounds

   1. *User Engagement*: Are users still as active and post on Instagram or they are making fewer posts
      Task: Provide how many times does average user posts on Instagram. Also, provide the total number of photos on Instagram/total number of users

   2. *Bots & Fake Accounts*: The investors want to know if the platform is crowded with fake and dummy accounts
      Task: 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).

Project Approach:

In order to execute the project, SQL and Tableau was used. SQL queries were used to create a database using the raw data provided. Once the database was created, various sorting and data extracting queries were used to get the results and necessary data required. In order to create graphical representation of the results and to understand the result better Tableau is used to retrieve the database by connecting to MySQL.

Tech Stack Used:

MySQL Workbench v8.0.29.0 community edition is used for project execution in order to query the database and gather the required results.

Tableau Desktop v2019.4.1 professional edition is used to create visual representation of results for creation of graphs and other charts to understand the result better and make better decisions.

Results and Insights:

```
96    /* Finding 5 oldest users of the Instagram for rewarding most loyal users */
97
98    select username, created_at from users order by created_at asc limit 5;
99
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

username	created_at
Darby_Herzog	2016-05-06 00:14:21
Emilio_Bernier52	2016-05-06 13:04:30
Elenor88	2016-05-08 01:30:41
Nicole71	2016-05-09 17:30:22
Jordyn.Jacobson2	2016-05-14 07:56:26

```
100  /* Identifying inactive users who have not posted a single photo on Instagram
101     and sending them promotional mails for posting first photo */
102
103  select username from users left join photos on users.id = photos.user_id where photos.image_url is null;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

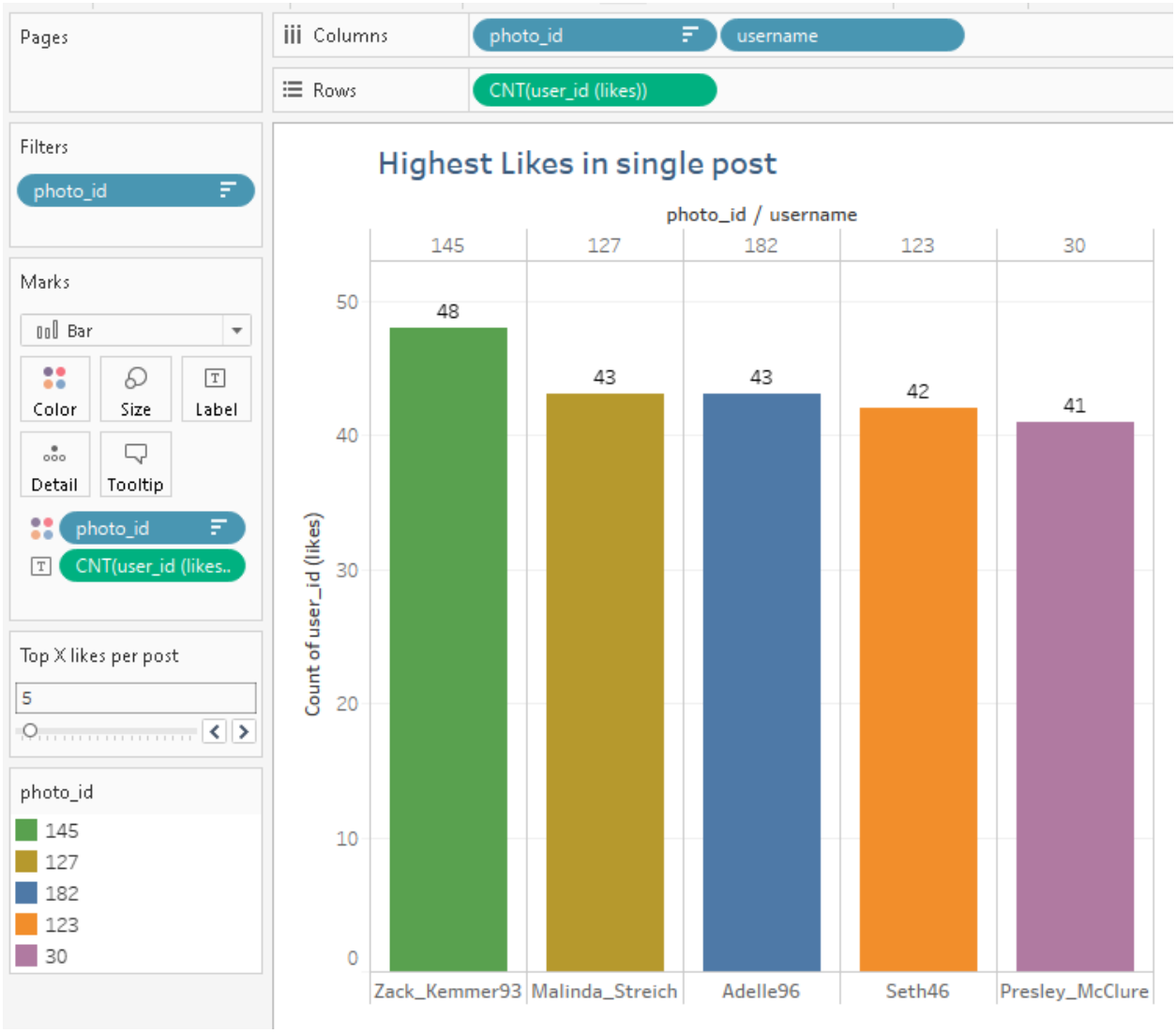
username
Aniya_Hackett
Kasandra_Homenick
Jaclyn81
Rocio33
Maxwell.Halvorson
Tierra.Trantow
Pearl7
Ollie_Ledner37
Mckenna17
David.Osinski47
Morgan.Kassulke
Linnea59
Duane60
Julien_Schmidt
Mike.Auer39
Franco_Keebler64
Nia_Haag
Hulda.Macejkovic
Leslie67
Janelle.Nikolaus81
Darby_Herzog
Esther.Zulauf61
Bartholome.Bernhard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

```sql
/* Identifying the winner of the contest with most likes on a single photo
   and provide their details to the Instagram team */

select username, photos.id, photos.image_url, count(likes.user_id) as highest_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 highest_likes desc limit 1;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

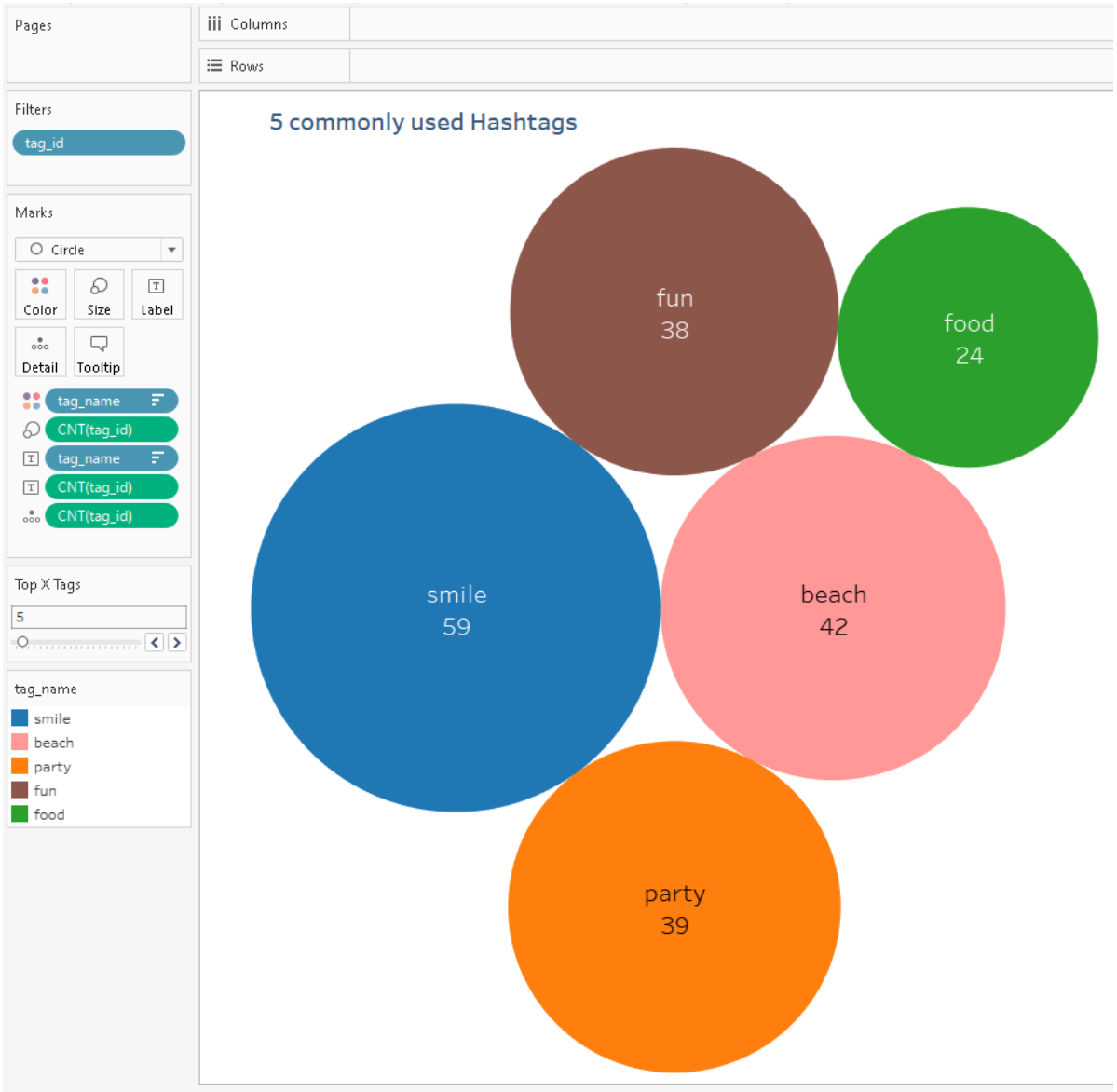
username	id	image_url	highest_likes
Zack_Kemmer93	145	https://jarret.name	48



```
112  /* identifying and suggesting the top 5 most commonly used hashtags to the partner brands from the platform
113     to reach most people for advertisements/promo campaign that users use while posting photos/videos */
114
115  select tag_name, count(tag_id) as num_tags from tags
116  inner join photo_tags on tags.id = photo_tags.tag_id
117  group by tag_name
118  order by num_tags
119  desc limit 5;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

tag_name	num_tags
smile	59
beach	42
party	39
fun	38
concert	24

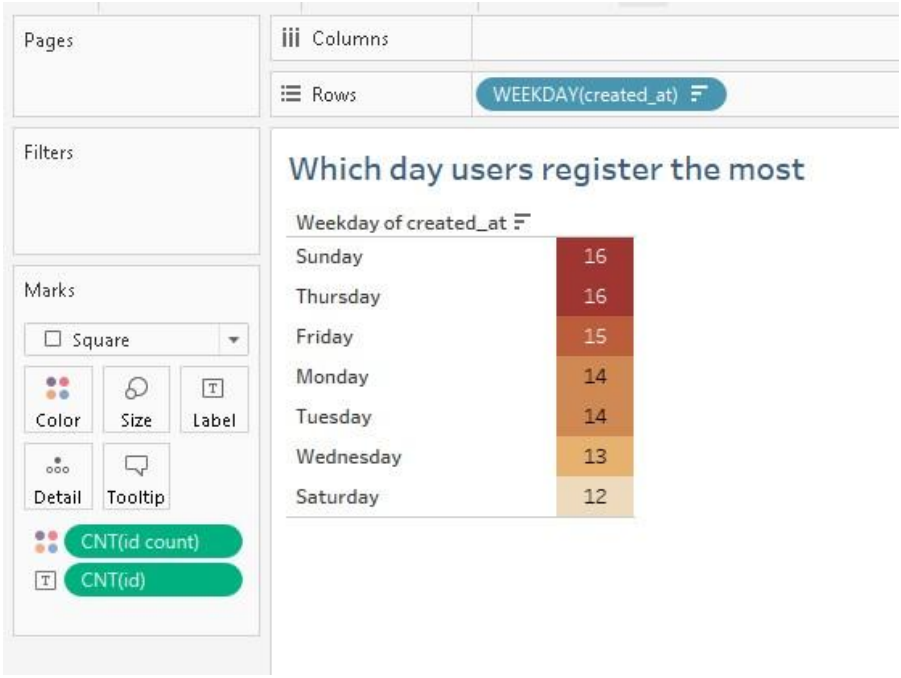

Pages
Columns
Rows
Filters
tag_id
Marks
Circle
Color
Size
Label
Detail
Tooltip
tag_name
CNT(tag_id)
tag_name
CNT(tag_id)
CNT(tag_id)
Top X Tags
5
tag_name
smile
beach
party
fun
food
5 commonly used Hashtags
fun
38
food
24
smile
59
beach
42
party
39

```sql
/*  The team wants to know, which day would be the best day to launch ADs.
    What day of the week do most users register to schedule an ad campaign */

select DAYNAME(created_at) as day,
COUNT(*) as Total_registrations from users
group by day
order by Total_registrations desc;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

day	Total_registrations
Thursday	16
Sunday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12



```sql
/* how many times on average does the active user posts on Instagram
    (excluding inactive users with no posts) */

select
  sum(photos) / count(users) as average_posts_per_active_user
from (
  select users.id as users, count(photos.id) as photos
  from users
  left join photos on users.id = photos.user_id
  group by users
) as a
where photos > 0;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

average_posts_per_active_user
3.4730

```sql
/*  total number of photos on Instagram/total number of users */

select
  (select count(*) from users) as total_users,
  (select count(*) from photos) as total_photos,
  (select count(*) from photos) / (select count(*) from users) as average_photos_per_user;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

total_users	total_photos	average_photos_per_user
100	257	2.5700

```sql
/* Finding users acting as bots who have liked every single photo on the site
   helping investors to classify or filter out dummy accounts*/

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

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

id	username	total_likes_by_user
5	Aniya_Hackett	257
14	Jaclyn81	257
21	Rocio33	257
24	Maxwell.Halvorson	257
36	Ollie_Ledner37	257
41	Mckenna17	257
54	Duane60	257
57	Julien_Schmidt	257
66	Mike.Auer39	257
71	Nia_Haag	257
75	Leslie67	257
76	Janelle.Nikolaus81	257
91	Bethany20	257