Instagram User Analytics

Project by Trainity

Shahnawaz Akhtar
shahnawazakhtar2@gmail.com

Instagram User Analytics

Description:

This is a project to do an extensive research on user behavioral patterns within Instagram. Through user interactions and behaviors on the platform we need to serve useful information that can drive new features from product team. These insights will essentially help to enhance and further the quality of User Experience by increasing satisfaction with Instagram.

This project mainly focuses on two
Major aspects Marketing and Investors Metrics.
Based on the user engagement and the data collected, the insights need to be carried out and presented to product team. The project will answer to such key questions as:

- Rewarding Loyal Users
- Remind Inactive Users to Start Posting
- Declaring Contest Winner
- Hashtag Researching
- Launch AD Campaign
- User Engagement
- Bots & Fake Accounts

Tech-Stack Used

MySQL Workbench 8.0 CE

Purpose – This tool is used o create the database and store records. It is also used to carry out the required analysis by writing SQL queries.

Microsoft® Excel® 2019

Purpose - This tool is used to create graphical representation of the results and to understand the result set better.


iNSiGHT

A) Marketing Analysis:

Task-1- Loyal User Reward: Identify the five oldest users on Instagram from the provided database.

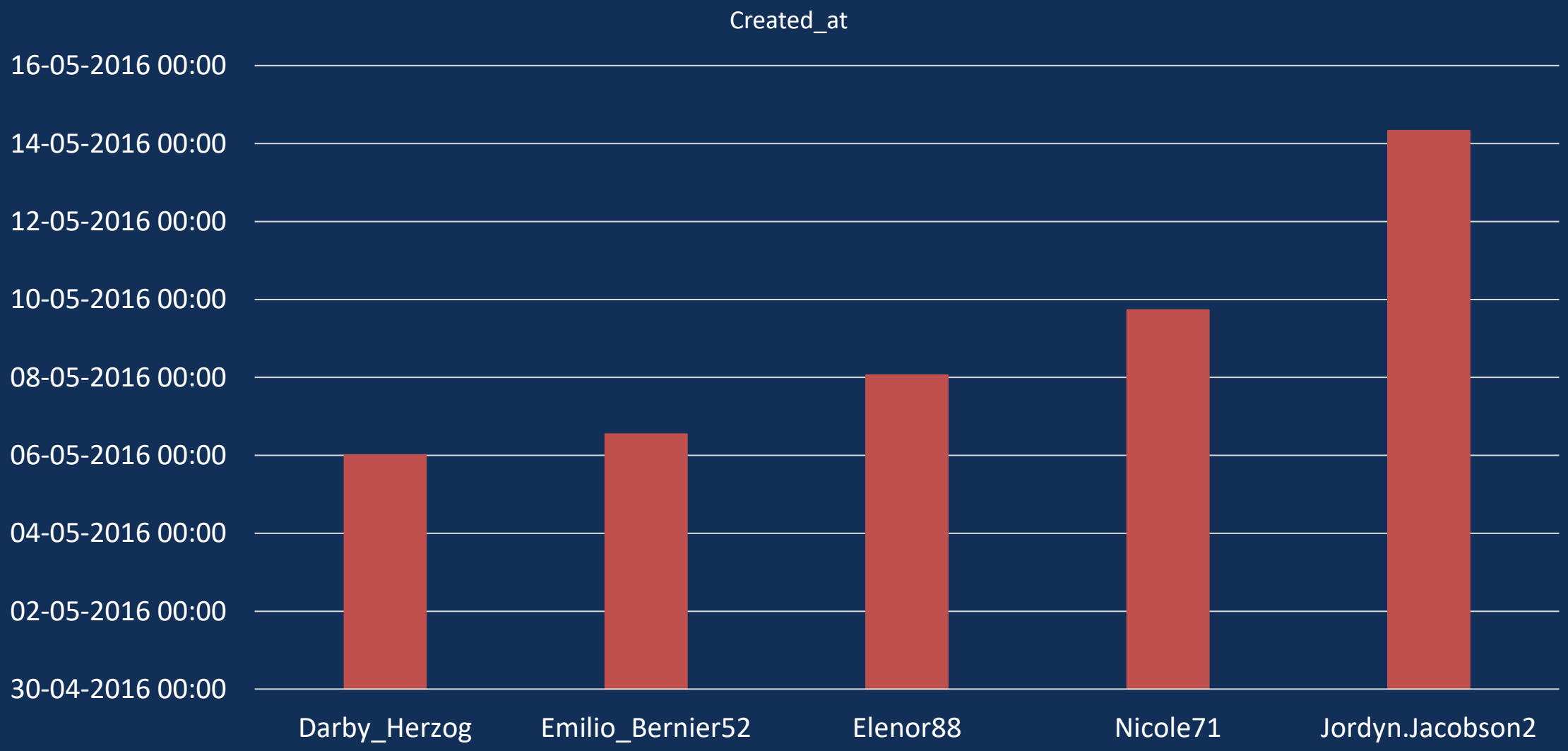
Solution Query-

```
SELECT username, created_at
FROM users ORDER BY created_at
LIMIT 5;
```

Output-

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

Task-1: Visualization



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

Solution Query-

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

Output- Inactive Users Never Posted a Photo

id	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

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

Solution Query-

```
with mostlikedphoto as (select  photo_id, count(user_id) as total_likes
from likes
group by photo_id
order by count(user_id)
desc limit 1)
select u.username, u.id, p.id as photo_id, mostlikedphoto.total_likes
from mostlikedphoto
join photos p on mostlikedphoto.photo_id = p.id
join users u on p.user_id = u.id;
```

Output-

Username	Id	Photo_id	Total_likes
Zack_Kemmer93	52	145	48

- According to data table user Zack_kemmer93 has posted a photo having photo Id 145 which has a total of 48 likes highest amongst all the users.

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

Solution Query-

```
with top_tags as (
  select tag_id from photo_tags
   group by tag_id
   order by count(tag_id)
   desc limit 5)
select t.tag_name from top_tags
join tags t on top_tags.tag_id = t.id;
```

Output-

tag_name
smile
beach
party
fun
concert

- According to the Data the there are five recommended hashtags to use in the posts to reach most people on platform

Task-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.

Solution Query-

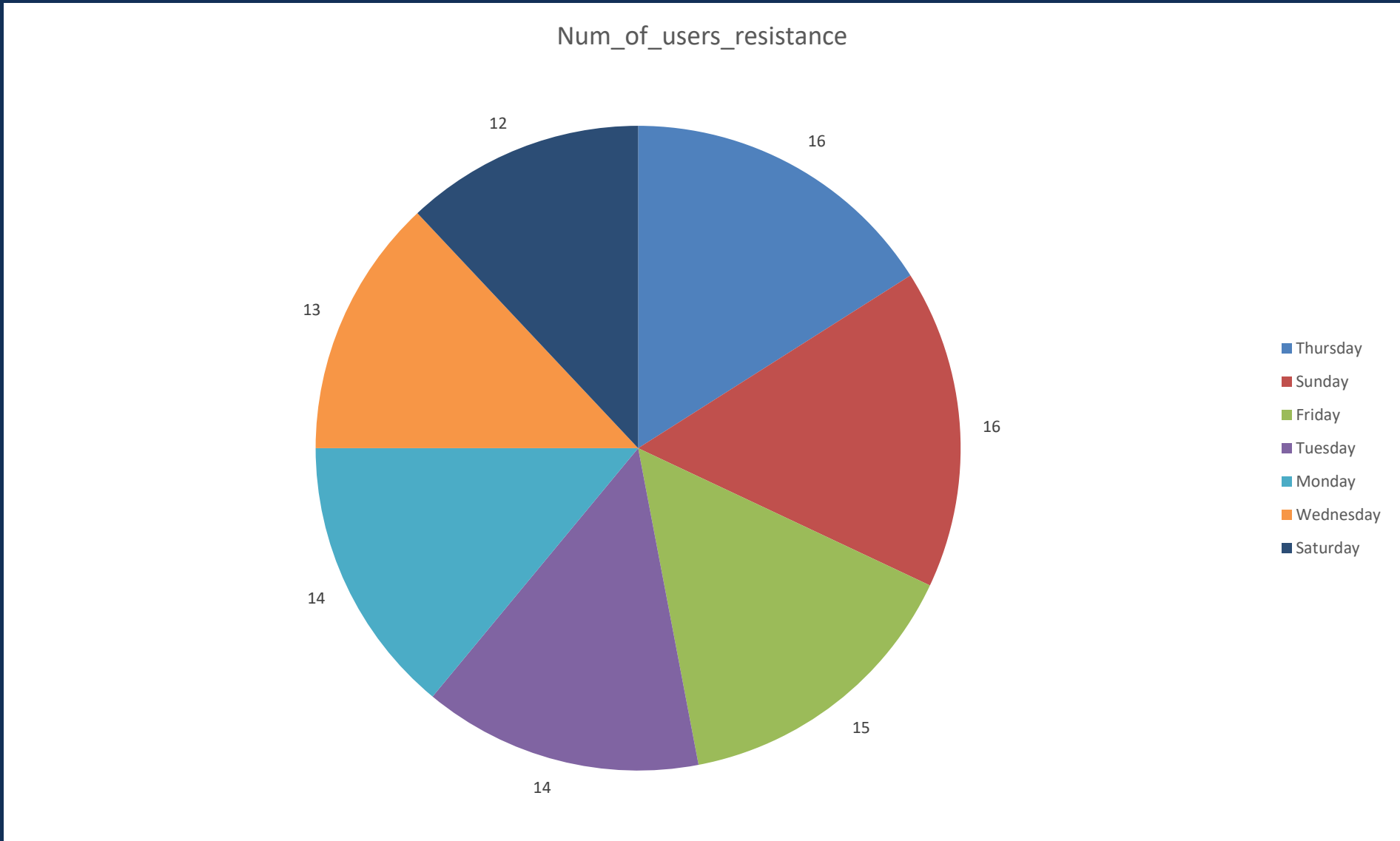
```
 select dayname(created_at) as days_of_week
count(*) as num_of_users_resisters
from users
group by dayname(created_at)
order by num_of_users_resisters desc;
```

Output-

Days_of_week	Num_of_users_resistance
Thursday	16
Sunday	16
Friday	15
Tuesday	14
Monday	14
Wednesday	13
Saturday	12

- According to the data users have registered most on Sundays and Thursdays
- The Recommended days to launch Ads are Thursday and Sunday

Task-5: Visualization



B) Investor Metrics:

Task-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.

Solution Query:

- ✓ Calculate the average number of posts per user on Instagram post count by user-

```
select user_id, count(*)  as posts_count
from photos
group by user_id
order by posts_count desc;
```

- ✓ total number of photos on Instagram divided by the total number of users-

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

Task-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.

Solution Query:

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

- **According to the data there are some fake/bot accounts as there who have liked every single photo on the site, as this is not typically possible for a normal user.**

Output:

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

Conclusion

- The project addresses the queries and delivers the necessary insights that are pertinent to the product team's needs.
- This project has deepened my appreciation for the practical application of data in business contexts, and has significantly improved my proficiency in SQL and Excel.

Thank You !