

Instagram User Analytics

Project Description:

This project helps to analyse the raw data/metadata to create useful insights. Various database management tools can be used to extract useful insights and even visualise them. This enables a way to increase efficiency of a platform.

Project Approach:

In order to execute the project, SQL 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 data/insights required.

Tech Stack Used:

MySQL Workbench v8.0.32.0 was used during project execution in order to query the database. The ease of access and setup, troubleshooting support as well as the GUI made it a good tool for the project.

Insights:

A) Marketing:

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

QUERY:

/* 5 oldest users of the Instagram from the database provided */

SELECT * FROM users; /* Reference */

	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
*	HULL	HULL	HULL

SELECT * FROM users ORDER BY created_at ASC LIMIT 5;

CONCLUSION:

Users 80, 67, 63, 95, 38 are the 5 oldest users on the platform.

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

QUERY:

/ Users who have never posted a single photo on Instagram */*

SELECT * FROM photos; /* Reference */

SELECT * FROM users; /* Reference */

**SELECT users.id AS UserID, photos.id AS PostID,
users.username FROM users LEFT JOIN photos ON users.id =
photos.user_id WHERE photos.id IS NULL ORDER BY
username;**

CONCLUSION:

There are 26 users mentioned above who have never posted a single photo on Instagram.

	UserID	PostID	username
▶	5	NULL	Aniya_Hackett
	83	NULL	Bartholome.Bernhard
	91	NULL	Bethany20
	80	NULL	Darby_Herzog
	45	NULL	David.Osinski47
	54	NULL	Duane60
	90	NULL	Esmeralda.Mraz57
	81	NULL	Esther.Zulauf61
	68	NULL	Franco_Keebler64
	74	NULL	Hulda.Macejkovic
	14	NULL	Jacyln81
	76	NULL	Janelle.Nikolaus81
	89	NULL	Jessyca_West
	57	NULL	Julien_Schmidt
	7	NULL	Kasandra_Homenick
	75	NULL	Leslie67
	53	NULL	Linnea59
	24	NULL	Maxwell.Halvorson
	41	NULL	Mckenna17
	66	NULL	Mike.Auer39
	49	NULL	Morgan.Kassulke
	71	NULL	Nia_Haag
	36	NULL	Ollie_Ledner37
	34	NULL	Pearl7
	21	NULL	Rocio33
	25	NULL	Tierra.Trantow

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.

QUERY:

/ The winner of the contest the most likes on a single photo */*

SELECT * FROM users; /* Reference */

SELECT * FROM photos; /* Reference */

SELECT * FROM likes; /* Reference */

**SELECT photos.user_id , likes.photo_id AS PostID, COUNT(likes.photo_id) AS
TimesLiked, users.username FROM likes LEFT JOIN photos ON likes.photo_id =
photos.id INNER JOIN users ON photos.user_id = users.id GROUP BY photo_id ORDER
BY TimesLiked DESC LIMIT 1;**

	user_id	PostID	TimesLiked	username
▶	52	145	48	Zack_Kemmer93

CONCLUSION:

The winner of the contest ‘the most likes on a single photo’ is user 52 with a total of 48 likes on post 145.

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

QUERY:

/* Identify and suggest the top 5 most commonly used hashtags on the platform */

SELECT * FROM tags; /* Reference */

SELECT * FROM photo_tags; /* Reference */

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

	tag_id	tag_name	TimesUsed
▶	21	smile	59
	20	beach	42
	17	party	39
	13	fun	38
	18	concert	24

CONCLUSION:

The tags smile, beach, party, fun, concert are 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.

QUERY:

/* What day of the week do most users register on? Provide insights on when to schedule an ad campaign */

SELECT * FROM users; /* Reference */

**SELECT DAYOFWEEK(created_at) AS WeekDays,
COUNT(DAYOFWEEK(created_at)) AS UsersRegistered
FROM users GROUP BY DAYOFWEEK(created_at) ORDER
BY UsersRegistered DESC;**

/* Note: 1=Sunday, 2=Monday, 3=Tuesday, 4=Wednesday,
5=Thursday, 6=Friday, 7=Saturday. */

	WeekDays	UsersRegistered
▶	5	16
	1	16
	6	15
	3	14
	2	14
	4	13
	7	12

CONCLUSION:

Most users register on Sundays and Thursdays.

Note: 1=Sunday, 2=Monday, 3=Tuesday, 4=Wednesday, 5=Thursday, 6=Friday, 7=Saturday.

B) Investor Metrics:

1. **User Engagement:** Are users still as active and post on Instagram or they are making fewer posts

QUERY:

/* Provide how many times an average user posts on Instagram. Also, provide the total number of photos on Instagram/total number of users */

SELECT * FROM photos; /* Reference */

SELECT COUNT(id) / MAX(user_id) AS AvgPost, MAX(id) AS TotalPosts, MAX(user_id) AS TotalUsers FROM photos;

	AvgPost	TotalPosts	TotalUsers
▶	2.5700	257	100

CONCLUSION:

A user posts 2.57 posts on an average.

There are 257 photos in total on Instagram.

There are 100 users in total on Instagram.

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

QUERY:

/* 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) */

SELECT * FROM likes; /* Reference */

SELECT likes.user_id, users.username, COUNT(user_id) AS LikeCount FROM likes INNER JOIN users ON likes.user_id = users.id GROUP BY user_id HAVING COUNT(user_id)=MAX(photo_id);

	user_id	username	LikeCount
►	5	Aniya_Hackett	257
	14	Jadlyn81	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

CONCLUSION:

There are 13 bots in total on the platform who have liked all the photos.

Result:

While making the project I learnt how the raw data can be used for our benefit and the actual execution of collecting insights from raw data. In the process of making this project I learnt SQL. I learnt how to create a database and query the database for extracting data.