**ABOUT THE PROJECT**

|  |  |
| --- | --- |
| HACK THE FEED SOCIAL MEDIA INSIGHTS, REPORT, Findings AND Recommendations.  Dicey tech hackathon | TABLE OF CONTENTS  •About the project •Aim and objectives of the project •Insights and reports generation  By: Oluwatobiloba Odelola Solomon  08/10/2023 |

Playhouse Communication is one of Nigeria's leading digital marketing agencies. They combine design and media planning with cutting-edge tech solutions to reimagine what marketing is all about. Their client roster is a mix of global juggernauts and nimble SMEs, each redefining their sectors.   
They rolled out the ultimate arena for innovation in data and set the stage for up and coming data scientists and analysts to showcase their skills, win huge cash prizes, and boost their careers. The "Hack the Feed" hackathon is a showdown where data analytics meets creative prowess.

Note: I used Microsoft excel, python for data analysis, machine learning modeling and power bi for this project till completion.

**AIM AND OBJECTIVE**

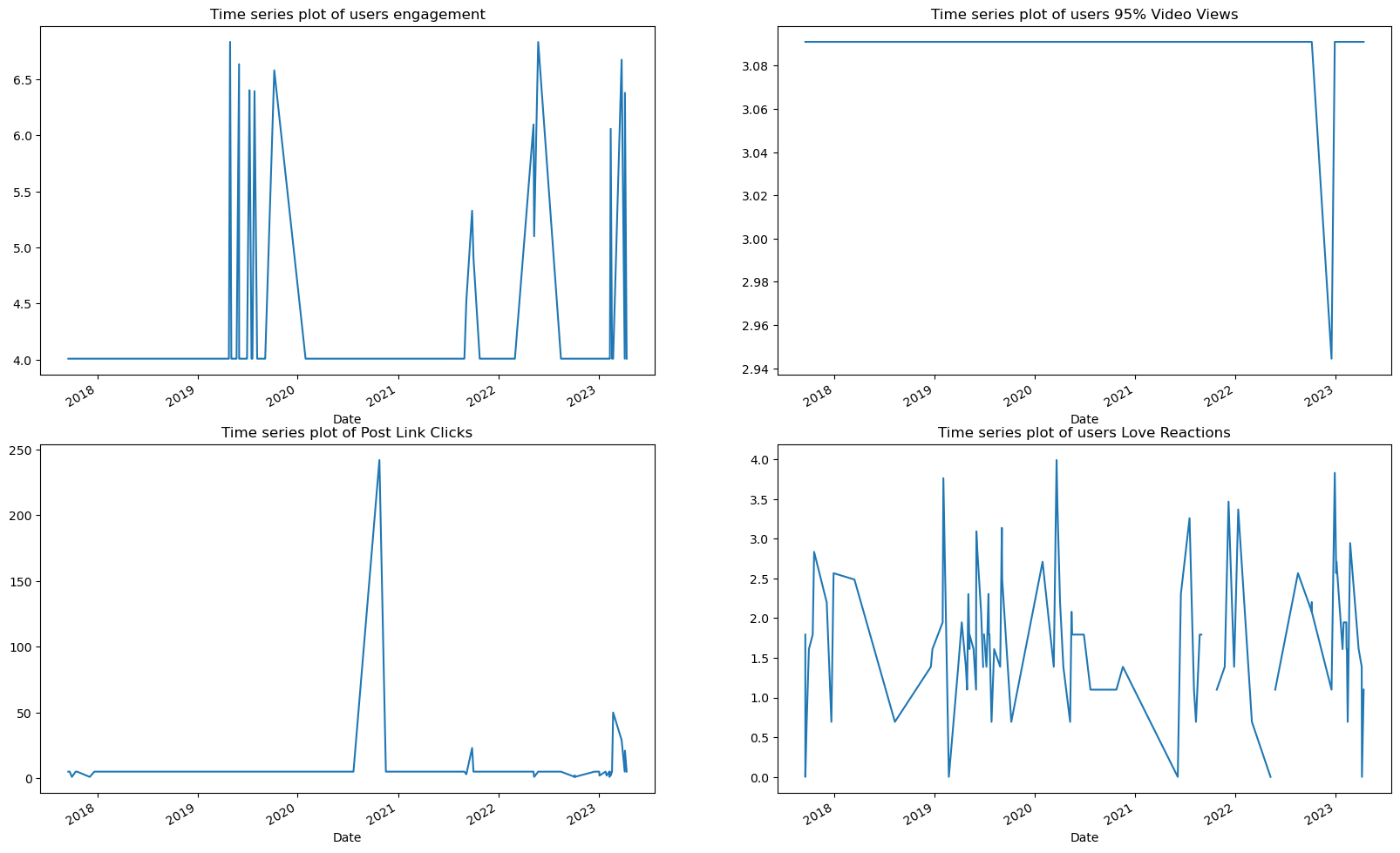
This project aims to:

* Perform data wrangling in order to decode the insights in the social media data provided and transform it into game-changing insights.
* Create a simple and engaging visualization of your results & analysis.
* Propose actionable recommendations based on the insights.
* Create a comprehensive and reproducible report detailing findings.

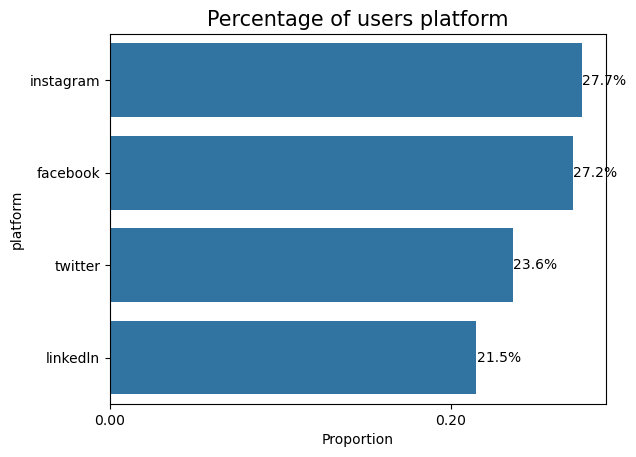
**Insights and Reports based on findings**.

In doing this, we will be answering leading questions that will help us look deeper into our data with the aid of visualizations.

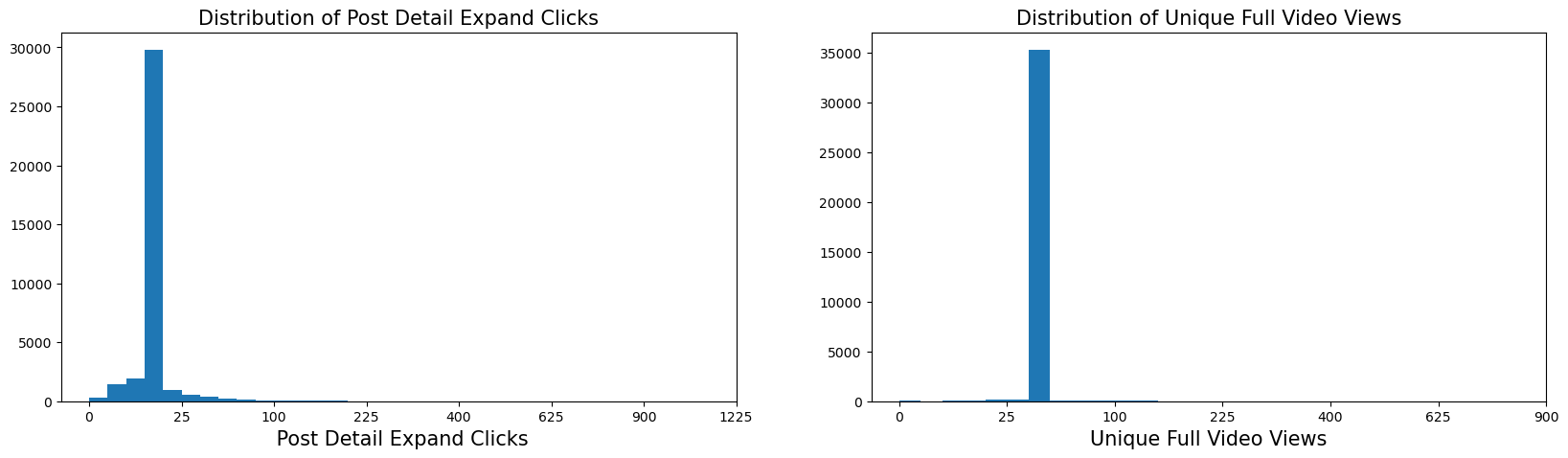
* Question 1 ( What is the trend in the engagements, potential reach, love reactions and the 95% video views of users?)

****

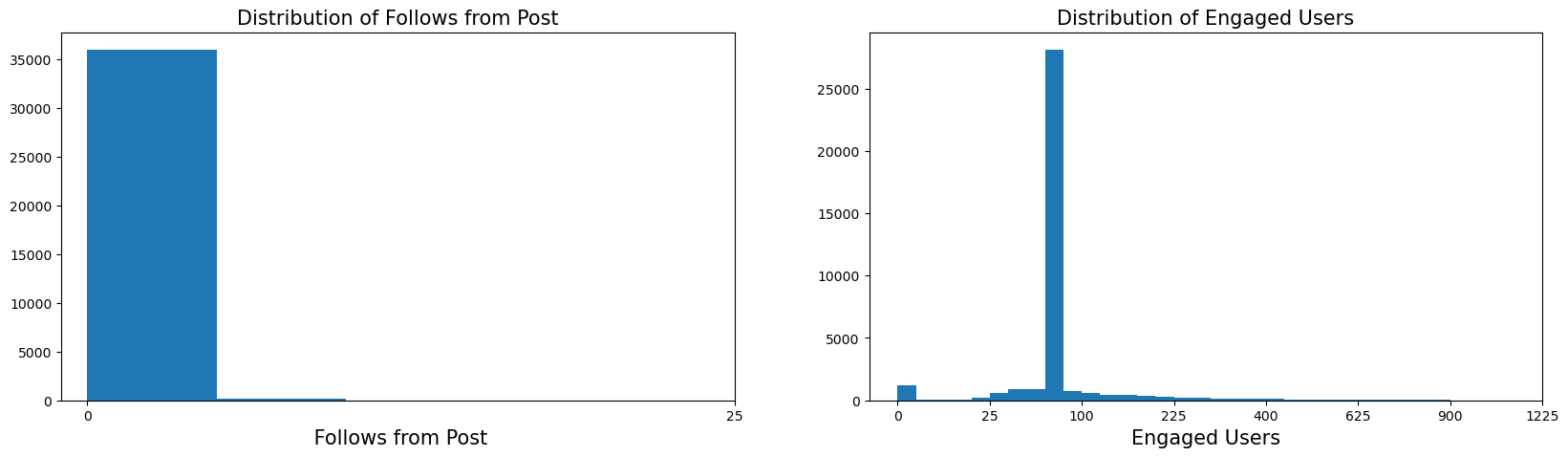
* **result/ findings :** From the charts above, it is seen that there was no particular trend in time from 2018 till 2023 between users engagement, which shows that the way users engage with posts does not follow a particular order(trend). There were some points where users had a constant engagement of 4 from 2018 till 2020, 2020 to 2022 and also at some points, there was a very drastic rise but irregular.
* For the users who watched about 95% of the videos there was a constant number of that particular occurrence above 3.08 from 2018 to 2023.
* Even for the time series plot of users love reactions, there is a slightly good trend at some points but the reactions from the users was not consistent.
* Also between 2022 and 2023, there was a period where users almost gave no love reactions, I suggest the company look at that particular period to really know what was wrong with probably the posts within that particular time frame
* Question2(What Network of the users has the Highest Percentage?)?)

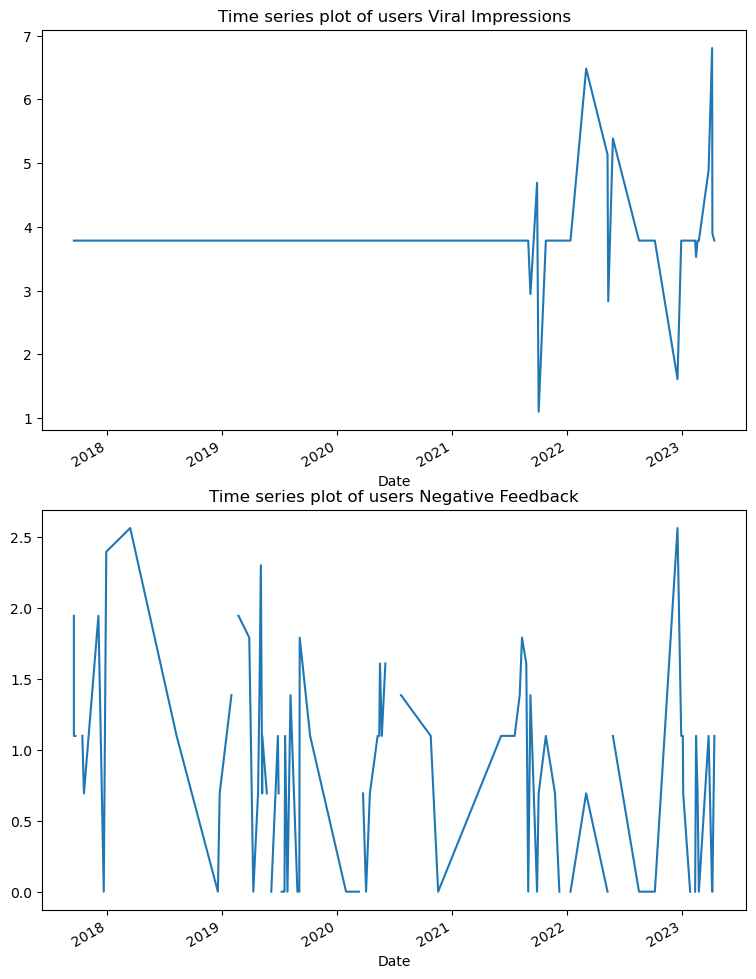


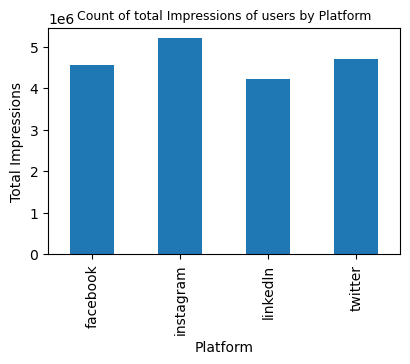
* **result/ findings :** it is obvious that Instagram network had the highest percentage of the number of users, followed by Facebook, twitter and then linkdln. This could be due to a lot reasons which I drill deep into in my python notebook using visuals. It is reasonable to say that Instagram has been having more attention of the users followed by facebook, so it will be wise if the company focuses its resources more on the two network platforms for more profits.
* Question 3 (What is the Distribution Post Detail Expand Clicks and Unique Full Video Views of users?)



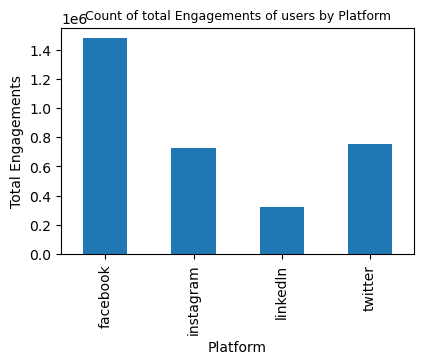
* **result/ findings : t**he distribution above shows that both variables are not normally distributed, and I saw that most of the variables were not normally distributed when I displayed the summary statistics of each numerical variables. Hence it shows that there are possible outliers in this model which I removed before building the machine learning models.
* Question 4 (What is the Distribution Follows from Post and Engaged Users of users?)

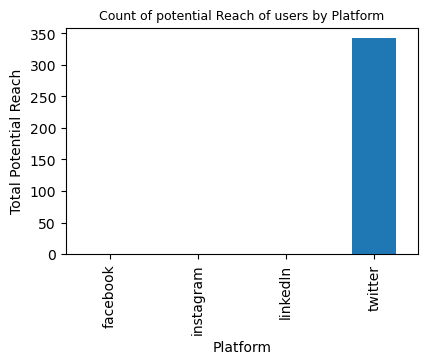


* **result/ findings : t**he distribution above shows that both variables are not normally distributed, and I saw that most of the variables were not normally distributed when I displayed the summary statistics of each numerical variables. Hence it shows that there are possible outliers in this model which I removed before building the machine learning models
* Question 5( What is the trend in the users Viral Impressions and love reactions and Negative Feedback ?) 
* **Results / Findings:** the time series plot of viral impressions shows that there was a constant or uniform or we say not trend between year 2018 and 2022, until which it now started now started receiving some impressions from users.
* Question6 (Which platform has the highest number of user engagement and love reactions?)

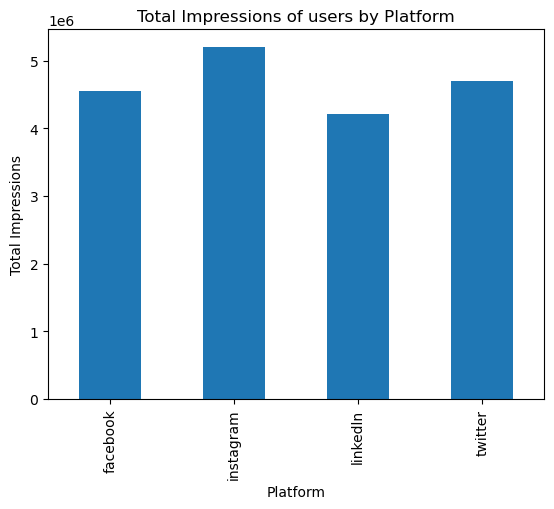


Obviously, Instagram had the highest count, so if the company could focus more of its contents on Instagram and Facebook, the probability that more income will be generated is very high.

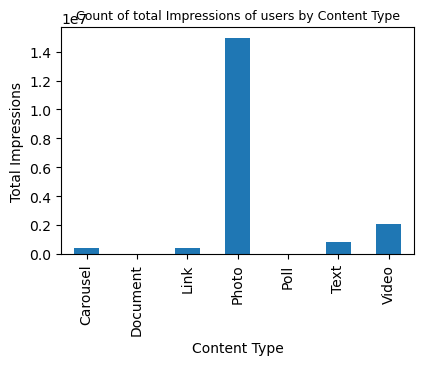


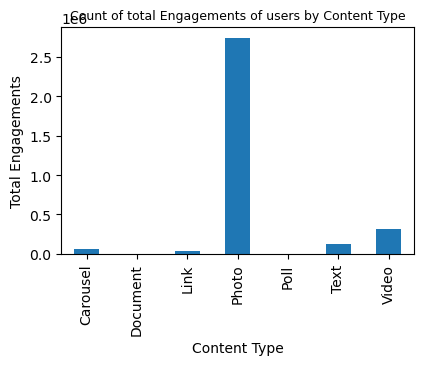


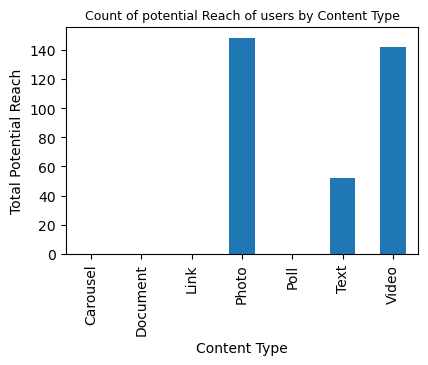
* **Results/findings:** funny and obviously enough, the first plot here shows that users engage more with Facebook than Instagram so it is still with the context of what I explained earlier that more posts should be directed to Instagram and Facebook. Linkdln had the lowest user engagements, from experience,linkedln is more professional and there a limited number of people who have the high confidence of being there especially if they do not have a skill. Hence it is meaningful to see why it had the lowest engagement from users. For twitter it also had as good number of user engagements in comparison with Instagram.
* **question 7 (Which platform has the highest number of impressions ?)**

****

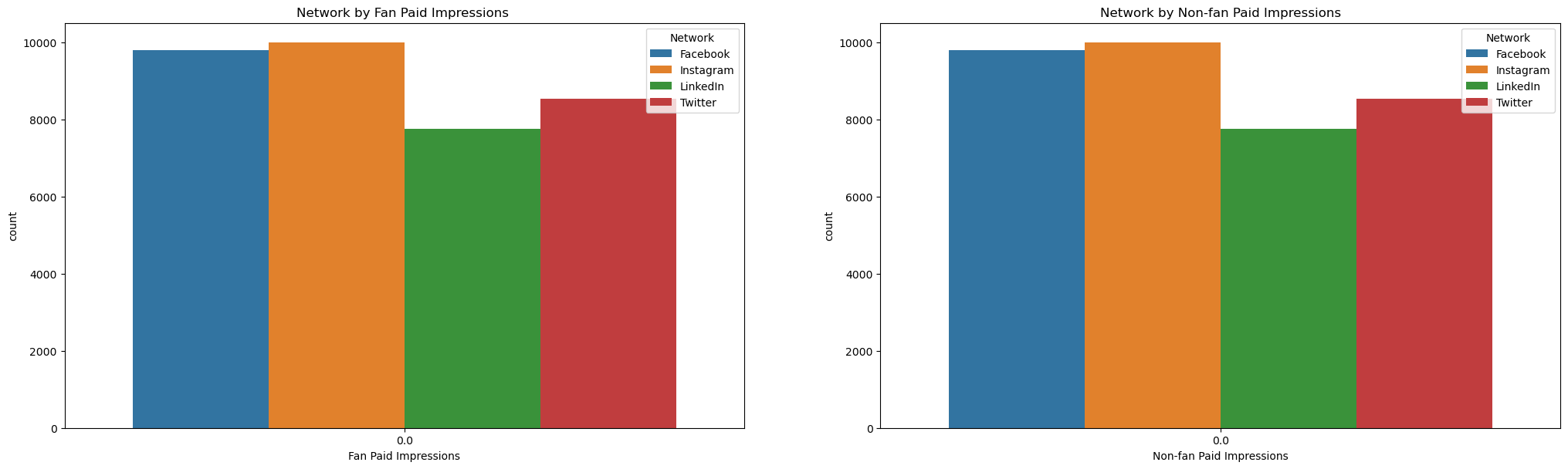
* **Results/findings:** it is still clear that Instagram, Facebook and twitter had the highest impressions, hence those networks are the best to focus on.
* **Question 8 (Which of the type of contents has the highest average Impressions, Engagements and Reach?)**

****

****

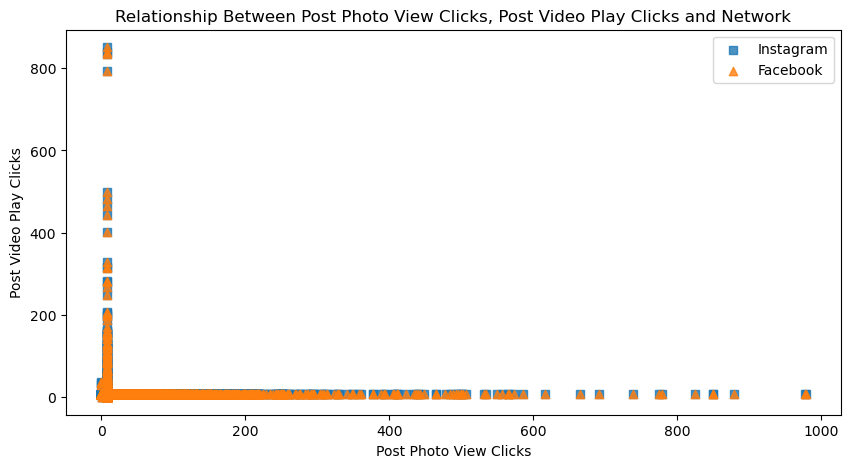
****

* **Results/findings:** Looking at the three plots, it follows consistently that posts that come in form of Photo had more engagements, potential reach and impressions than every other posts, so it means users are more interested in posts than other type of posts, hence, the company should post more of photos.
* **Question 9 ( Are there differences in the types of paid impressions based on Network?)**

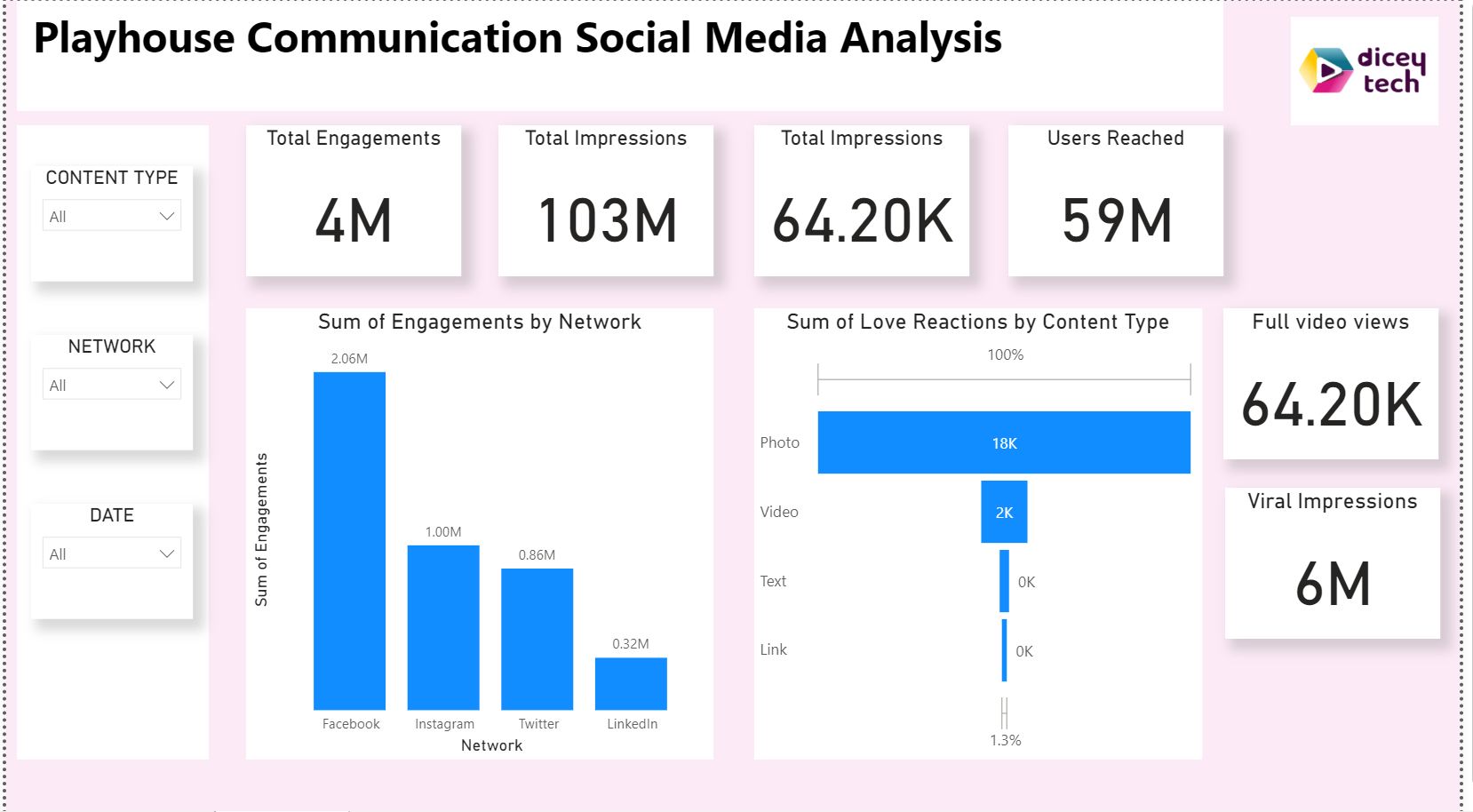
****

* **Results/findings:** obviously there is no difference in the result after paying to get fan impressions and even when money was not paid, so there is no point paying to get users impressions henceforth.

**Question 11 (What is the relationship between Post Photo View Clicks,Post Video Play Clicks and Network?)**

****

* **Results/findings:** there is no meaningful relationship between relationship between Post Photo View Clicks,Post Video Play Clicks and Network as it is neither a positive nor negative correlation. Hence, they overlap .

****

**Recommendations and summary**:

From the different plots and key findings, it is clear that facebook, Instagram and twitter had the higest user engagements, impressions, reach and even views, so it is advisable that the company should post most of her content on these platforms. Also users engage more with photos than other type of contents, so the company should post more of photo contents, the reasons users reason with photo more is very reasonable because some of the videos might be large and some users may not have enough electronic space or even data to watch for long.

**Limitations**:

More exploration would have been done if the data had information about users greographical loactions so that I can be able to see if there are geographical issues that could possibly limit users engagements with certain posts and network. Another limitation is that the data having the income details of the users will also aid the analysis so that I could see if it was due to money that users do not want to really view video posts , because it is obvious that in some very poor countries, users might be watching their pockets and then get scared of using too much data since video consumes more data than photos.