MEASURING THE SUCCESS OF YOUR FACEBOOK FEED

MACKENZIE GARRETT AND ALLIE STEHNEY

INTRODUCTION

As discussed in class, the success of one's social feed, and how one measures success, depends on perspective. Social media platforms define success differently than organizations and organizations define paramaters of success differently than individuals. Moreover, the consumer of a social feed and the producer of a social feed measure success in different ways, which is what this project will focus on. This project involves self-tracking of our own social media feeds in order to determine the success of each feed.

For our content source, we chose to analyze our individual Facebook feeds. We are interested in determining the success of our own individual Facebook feeds, so we looked at a total of two (2) Facebook feeds -- each of our own. Since determining the success of an individual social feed depends on personal goals and information, we decided to each define what success would look like for our own Facebook feed. While the social media platforms themselves prefer to see several different actions happeni ng -- i.e., liking, commenting, sharing, an individual might not want the same thing out of their social feed. So, we each defined what success would look like in terms of our social feeds, seen later on in the project, and then created a calculation for a "success" score in order to maintain consistency across the project. This score takes into consideration the total number of Facebook friends we each have, the number of likes received per post, and the number of comments received per post.

We collected data via a python code that scraped the Facebook API and also manually put in data about our Facebook feeds into an excel spreadsheet. From there, we created data visualizations and infographics to represent all of the data we collected in an attempt to draw a conclusion about our findings. By creating these visualizations before conducting our analysis, we were able to see trends arise within the data sets which helped us to better come to a conclusion about the individual success of our feeds.

Mackenzie's Facebook Feed:

I have been using my Facebook since 2009, and my initial reason for using the site is similar to my reason for use now, which is to connect with my friends. However, a difference in my intention of use now is that I additionally desire to connect with my family, and since I am now away from home at college, I wish to stay updated with my friends from back home. Furthermore, I use my feed to share information that I think reflect updates in my life, for both those I see frequently and those I see rarely.

For this study, the success of my feed will be measured by:

-The number of people I interact with that I do not see often (i.e., old high school friends that live in different states)
-Number of likes on my personal posts
-Number of comments on my personal posts

Before this study, I believed that my feed was successful because I am on it daily and post frequently. I didn't believe that I interacted with all 750 of my Facebook friends, but I felt confident that i was interacting with a wide range of them. I am expecting to see results that align with this view. Additionally, I believe I post quite often, whether it be through status updates or shared posts. While I don't post everyday, I know I post multiple times per week.

Allie's Facebook Feed:

I became a Facebook user in 2009, back when I was a meager 14 year old. That age might sound shocking -- or at least to my parents who still hoarded fears of catfishing and internet scandals seven years ago. I initially joined so that I could stay updated with what my friends were doing and post statuses and pictures to connect with them. Fast forward seven years and I am still using my feed for the same purpose; however, as Mackenzie mentioned, I am also away from home right now, and as a result, have connected with a lot of my family members in the last year. My delayed connection with family members stemmed from two things: first, my initial desire as a young high schooler to not be associated with my family on social media, and second, my family's hesitancy to sign up for Facebook. Now, I log onto Facebeook to see updates from recent friends, old friends, and family members. For posting purposes, I tend to stick to posting only pictures, but 2009 me would beg to differ with all of the statuses she posted.

For this study, the success of my feed will be measured by:

-Average number of likes I receive per post
-Average number of comments I receive per post
-Number of people in my top likers and top commenters who live in a different city than me (so, I do not see in person on a regular basis)

Before conducting this study I thought that my Facebook feed was successful. I consistently uploaded pictures that my 1,074 Facebook friends could see. Moreover, I felt that I was interacting with a good portion of my set of Facebook friends. I expect to see results that show a broad range of interaction with different people.

Method

In order to collect data from each of our feeds, we wrote a code in Python to call the Facebook API.* To ensure that we got a reliable amount of data, we scraped data from our past 100 posts. We decided to cut off at the 100th post number rather than a specific date because the two of us do not post as frequently as one another. We wanted to have the same amount of data for each person, so measuring by the number of posts rather than by a specific date afforded us this opportunity. It is also important to note that we scraped our individual feeds -- so, not the newsfeed where posts from friends can be found -- but our individual feeds that contain content we generate. Scraping the newsfeed is more complicated since it involves receiving permission from all of the users in our feeds to look at their data.

Looking now at our code, we created it in order to analyze ten pieces of data:

Time span of posts
Find the dates of our most recent and our oldest post within the 100 post span; this information is helpful because it shows how often a user posts.

Total number of likes
Total amount of likes we each received on all of our 100 posts combined; needed in order to calculate success score.

Total number of comments
Total amount of comments we each received on all of our 100 posts combined; needed in order to calculate success score.

Number of unique likers
Total number of different people that have liked the last 100 posts of each of our feed. (Example: if Bob liked 5 of my past 10 posts and Tom liked 1, I would have a total of 2 unique likers for my last 10 posts).

Number of unique commenters
Total number of different people that have commented on the last 100 posts of each of our feed. (Example: if Bob commented on 5 of my past 10 posts and Tom commented on 1, I would have a total of 2 unique likers for my last 10 posts).

Top 3 Commenters
The three users who had the highest number of comments on our collective last 100 posts; needed in order to understand relationship between user and top commenters.

Top 3 Likers
The three users who liked the most of our past 100 posts; needed to understand the relationship between user and Facebook friends.

* the base code is from SI 106, but manipulations, changes, and additions were made to get the desired data for this project

```python
#Comm 404 Project
#Note: this code is based off of code from SI 106. We use the base code, along with the code we had to write to complete the set.
#Note: We also manipulated the code to search for different data.

import unittest
import requests
import json
from pprint import pprint

class Post():
    """object representing status update"""
    def __init__(self, post_dict={}):
        if 'message' in post_dict:
            self.message = "message"
        else:
            self.message = "none"
        if 'comments' in post_dict:
            self.comments =[]
            for x in post_dict['comments']['data']:
                self.comments.append(x)
        else:
            self.comments = []
        if 'likes' in post_dict:
            self.likes = []
            for x in post_dict['likes']['data']:
                self.likes.append(x)
        else:
            self.likes = []

access_token = "EAACEdEose0cBAJslxh0FakbXrISzn1wSCa8URYRiHfkQ3BMv6xX2aLRsoBtqpw7UcxZAzsD98HasSKzia3xgTjGgWZCOODwfUZBgSkAHncLckoFKkY92q7vrYzqbtJQSupSRtT3gf
baseurl = "https://graph.facebook.com/v2.3/me/feed"

# Building the Facebook parameters dictionary
url_params = {}
url_params["access_token"] = access_token
url_params["fields"] = "comments{comments{like_count,from,message,created_time},like_count,from,message,created_time},likes,message,created_time,from" # P
url_params["limit"] = 100

#request to the Facebook API
resp = requests.get("https://graph.facebook.com/v2.3/me/feed", params= url_params)
fb_data = json.loads(resp.text)

#Below is code for a list holding FB data that has Post instances
post_insts= []
for item in fb_data['data']:
    post_insts.append(Post(item))
#print pprint(fb_data)
#uncomment above to find data for first and last posts!

def second_elem(tup):
    return tup[1]

#code to collect top commenters and top likers
top = {}
for x in post_insts:
    for likes in x.likes:
        if likes["name"] in top:
            top[likes['name']]+=1
        else:
            top[likes['name']] = 1
likez = top.items()
#print likez

def second_elem(tup):
    return tup[1]

likez2 = sorted(likez, key = second_elem, reverse = True)
top_liker = likez2[:3]
#print top_liker

top_likers = []
for tup in top_liker:
    top_likers.append(tup[0])
print top_likers
print "above are the top 3 likers"

```

```
topc = {}
for x in post_insts:
    for com in x.comments:
        if com['from']['name']in topc:
            topc[com['from']['name']]+=1
        else:
            topc[com['from']['name']] =1


comz2 = sorted(topc.items(), key = second_elem, reverse = True)
top_commenter= comz2[:3]
top_commenters = []
for tup in top_commenter:
    top_commenters.append(tup[0])
print top_commenters
print "above are the top 3 commenters"

def unique_facebookers(posts):
    likz= {}
    for x in post_insts:
        for likes in x.likes:
            if likes["name"] in likz:
                likz[likes['name']]+=1
            else:
                likz[likes['name']] = 1
    likez = likz.items()
    print len(likez)
    print "above is number of unique likers"
    ents = {}
    for x in post_insts:
        for com in x.comments:
            if com['from']['name']in ents:
                ents[com['from']['name']]+=1
            else:
                ents[com['from']['name']] =1
    ments = ents.items()
    print len(ments)
    print "above is number of unique commenters"
    if len(ments) > len(likez):
        return "commenters"
    elif len(likez) > len(ments):
        return "likers"
    elif len(likez) == len(ments):
        return "equal"

print unique_facebookers(fb_data)
print "above is data for column for Posts having more comments or likes"

def facebookers(posts):
    total_likes = 0
    for x in post_insts:
        for likes in x.likes:
            total_likes += 1
    total_comments = 0
    for x in post_insts:
        for com in x.comments:
            total_comments += 1
    return total_likes, total_comments


print facebookers(fb_data)
print "above is total number of likes and then total number of comments"


# Creating the csv file
new_file_ref = open ('feed_data.csv', 'w')
new_file_ref.write('message, comment counts, like counts\n')
for x in post_insts:
    likes_count = 0
    for item in x.likes:
        likes_count +=1
    comm_count = 0
    for item in x.comments:
        comm_count += 1
    new_file_ref.write("{}, {}, {}\n".format(x.message, comm_count, likes_count))
new_file_ref.close()
```

Excel Spreadsheet Data Screenshot

We did not code for the number of Facebook friends that we each have, the average number of comments and likes per post, time span of each of our 100 posts, and relationship status and gender of top commenters and likers. This information was collected from our Facebook page (number of friends), from our personal knowledge (relationship status and gender), and from simple calculations (averages and time span). While we did not put it in our code, we thought that this information was still important to determine the success of our individual feeds. The number of Facebook friends, for example, is important especially in terms of comparing feeds because one person may receive more likes on their posts than the other, but strictly because they have more Facbeook friends than the other person; and, when looked at proportionally, the person who receives the smaller number of likes might have a more successful feed (if number of likes is part of their definition of success) if that person also has a smaller set of friends.

Calculating for time span is also important because one person could have 100 posts in one week on Facebook while another person might have to go back an entire year to find 100 posts. This variable shows how active a user is on Facebook, which can contribute to the success of the feed.

The relationship to top likers and commenters is also important because it highlights the span of a user's network. If the top liker for a user is a family member, rather than an old high school friend, the user may not be connecting out to the number of people that they thought they were. For example, it would prove interesting if a user's top likers and commenters were all family memebers because, usually, those can be the people that a user would see most often. So instead of using their feed to bridge connections with past friends and coworkers, the user is maintaining relationships with the people they are already closest to.

The below screenshot is a list of complied data from each of our CSV outputs. On the left, there are two columns for the comment counts per post; Mackenzie's data is column A (comment counts_M) and Allie's data is column B (comment counts_A); column D has Allie's likes per post data (like count_A) and column E has Mackenzie's likes per post data (like count_M).

	A	B	C	D	E
1	comment counts_M	comment counts_A		like counts_A	like counts_M
2	1	0		25	2
3	1	1		25	25
4	0	0		5	6
5	7	1		25	11
6	1	0		25	25
7	2	1		25	0
8	0	0		11	4
9	0	2		25	3
10	1	0		9	4
11	0	1		2	14
12	0	1		3	4
13	0	1		3	4
14	2	1		25	5
15	0	1		5	25
16	0	1		4	25
17	0	2		3	2
18	0	2		6	25
19	0	4		4	11
20	0	2		25	25
21	0	3		25	5
22	0	1		8	6
23	1	5		4	8
24	0	2		22	20
25	0	23		25	5
26	0	0		0	1
27	1	5		18	1
28	2	9		25	4

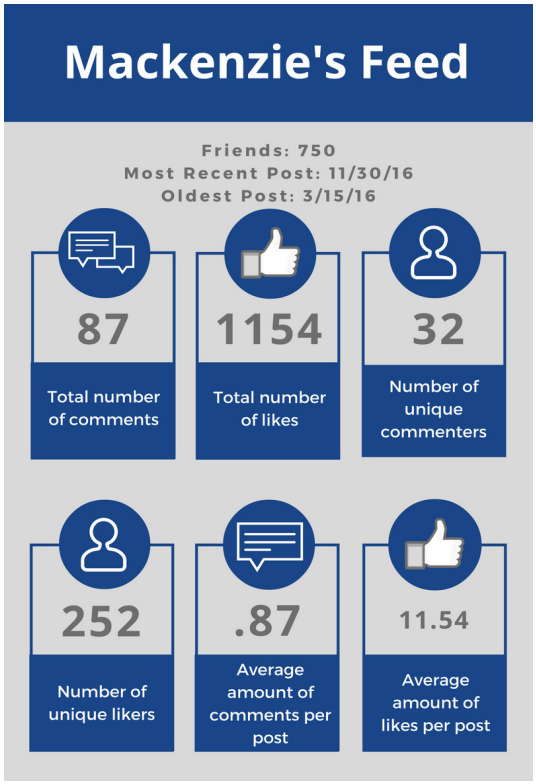
Excel Spreadsheet Data Screenshot

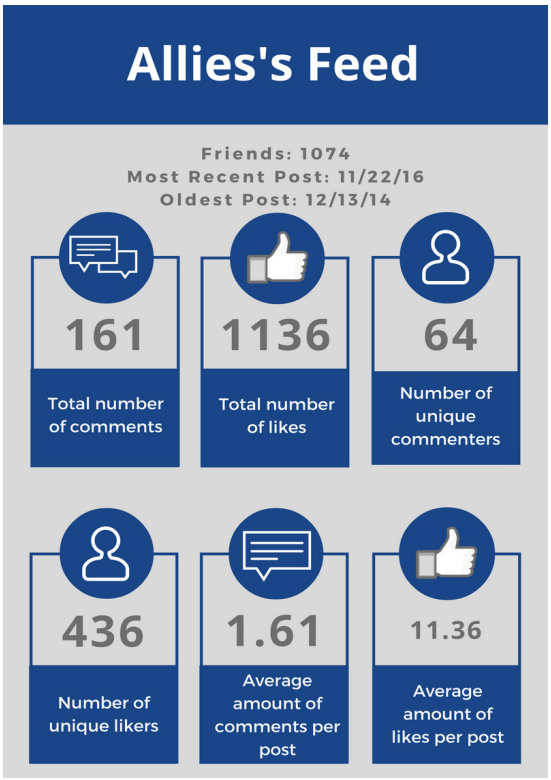
	A	B	C	D	E	F	G
1	message_M	comment counts_M	like counts_M		message_A	comment counts_A	like counts_A
2	message	3	20		message	0	25
3	message	1	2		none	1	25
4	message	7	11		message	0	5
5	none	1	25		none	1	25
6	message	2	0		message	0	25
7	none	0	4		none	1	25
8	message	0	3		message	0	11
9	message	1	4		message	0	0
10	message	0	14		none	0	9
11	message	0	4		none	1	2
12	message	0	4		none	1	3
13	message	2	5		none	1	3
14	none	0	25		none	1	25
15	none	0	25		none	1	5
16	message	0	2		none	1	4
17	none	0	25		message	2	3
18	none	0	11		none	2	6
19	none	0	25		none	4	4
20	message	0	5		message	0	0
21	none	0	6		none	3	25
22	none	1	8		none	1	8
23	message	0	20		none	5	4
24	message	0	5		none	2	22
25	none	0	1		message	23	25
26	message	1	1		message	0	0
27	message	2	4		message	5	18
28	message	3	10		none	0	0
29	message	1	25		none	1	2
30	none	1	13		none	3	25
31	none	0	25		none	3	25
32	none	2	22		none	1	3
33	none	0	1		none	0	8

After creating our original CSV file that told us the number of likes per post and the number of comments per post, we decided to rewrite some of our code to output a CSV file that tells us which of our posts include written messages. If the post included a written message (i.e., a status update) it printed 'message' otherwise it printed 'none'. Sharing a post or posting a picture for example, would print out 'none' since it is content that does not involve the user's own written message. If a message was attached to a share or to a picture, however, that post would print 'message'.

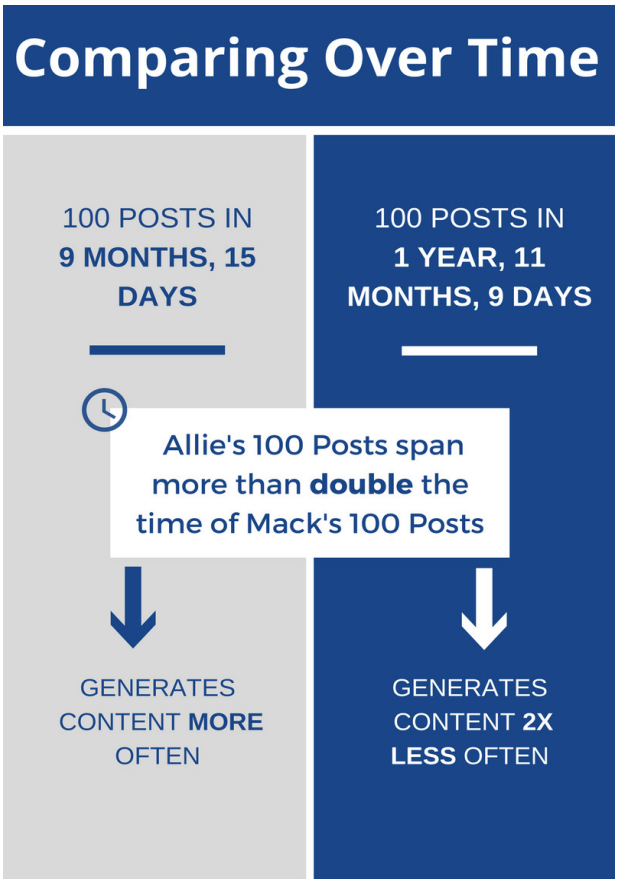
This set of data is particularly important because it provides some insight on the type of posts that each user is generating. Allie, for example, tends to post a lot of pictures, and mostly without captions, which is evident in the spreadsheet above. In this screenshot, Mackenzie's data is on the left (columns A-C) and Allie's data is on the right (E-G). The screenshots of the code on pages 4 and 5 reflect the change we made to output this CSV file.

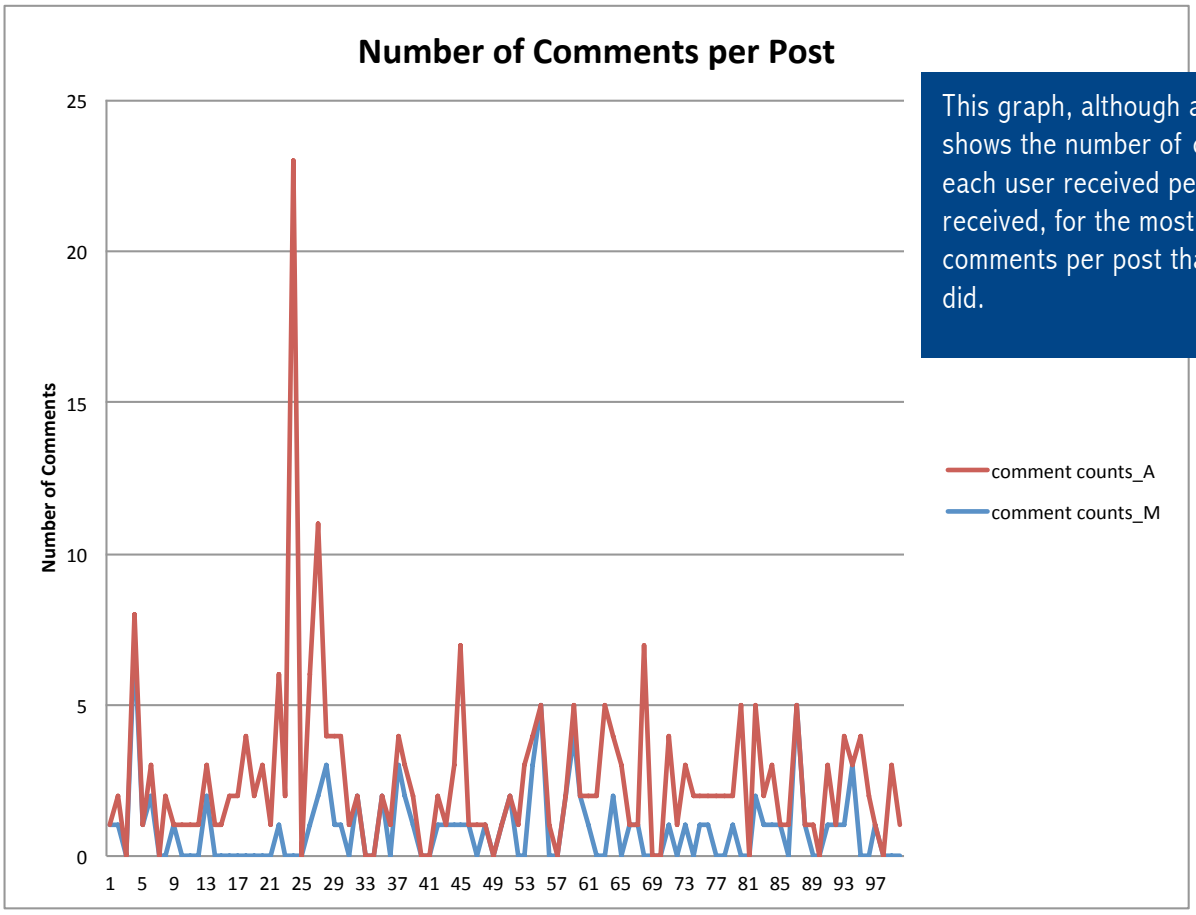
Infographics displaying the core data about each user's Facebook feed:



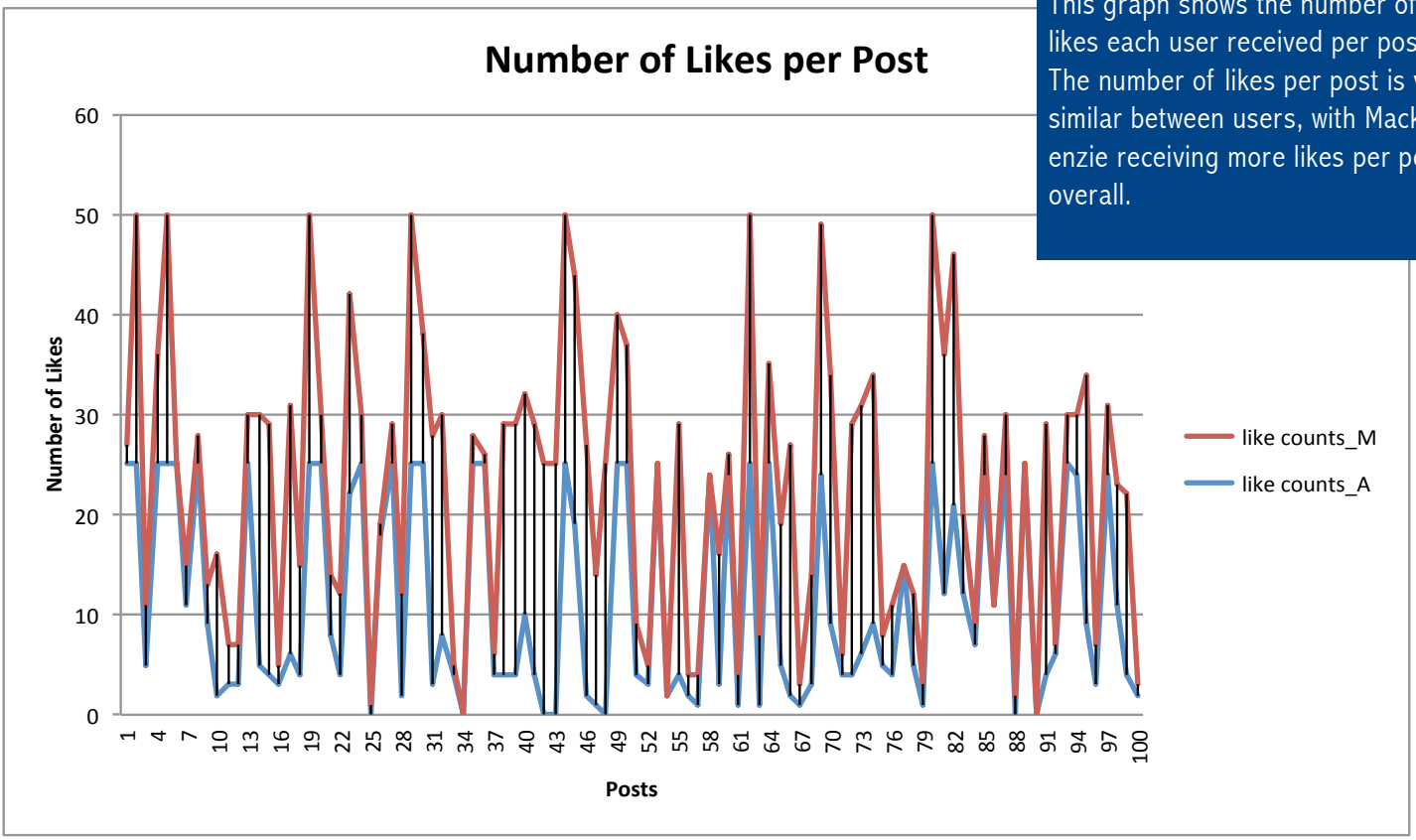


Infographic that highlights the time span for the most recent 100 posts for each user:



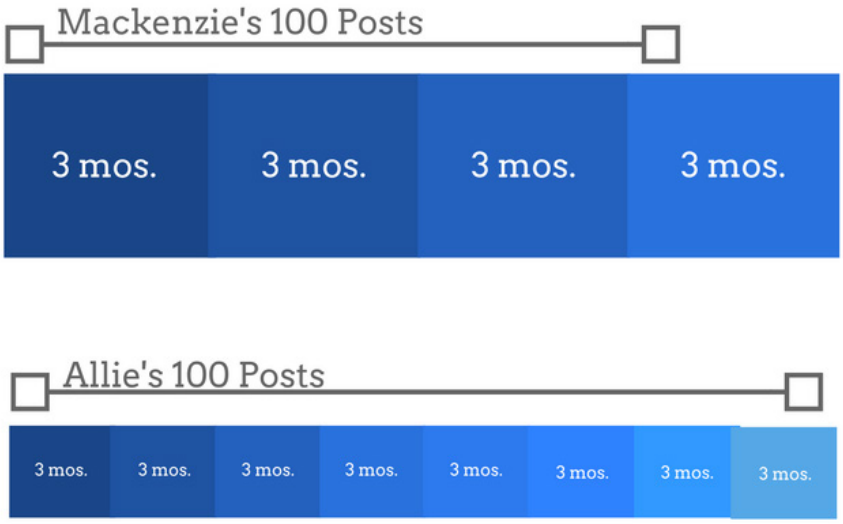


This graph, although a bit crowded, shows the number of comments each user received per post. Allie received, for the most part, more comments per post than Mackenzie did.



This graph shows the number of likes each user received per post. The number of likes per post is very similar between users, with Mackenzie receiving more likes per post overall.

Posts and Time



This visualization includes blocks of time (3 months each) to highlight the time span that the most recent 100 posts was for each user. As seen, Mackenzie's last 100 posts were found in the last 9 months, whereas Allie's go all the way back 1 year and 11 months.

All of these visualizations and infographics help to better organize our data and highlight trends that are not as obvious when looking at the initial data set. Andy Kirk's **Data Visualization: A Successful Design Process** underscores the importance of rearranging data from the initial vertical columns in order to see less obvious trends for our analysis. In an example in his paper he discusses the "sense of the physicality of the data and the potential influencing attributes that might shape [the] visualization architecture" (Kirk 72). Looking at our data in the excel spreadsheet, as well as the data from the code, it's hard to visualize the fact that Allie's most recent 100 post span almost double the time as Mackenzie's most recent 100 posts. And, this is particularly important because it is a factor in determining the success of a feed. A more active feed might be considered more successful by a user than a feed that does not generate as much original content.

Conclusion

Mackenzie's Feed: Unsuccessful

Comparing the parameters I defined for success and the analysis of my newsfeed, my feed proved to be unsuccessful. My top three likers are my roommates that I see everyday. Also, my top three commenters include a roommate, myself, and my a U of M classmate. While I knew prior to this study that I interact with these people often on Facebook, I was surprised to find out that they were the people I interacted with most. I predicted that I would have interacted with more people from back home, both family and friends, and these results tell me that my Facebook feed isn't as successful as I had hoped.

Additionally, my posts averaged 11.54 likes and .87 comments. While I don't think that this a bad average, I do think it is a little bit low considering that I have 750 Facebook friends. Plus, only 252 out 750 friends have liked at least one of my last 100 posts, and only 32 Facebook friends have commented at least once on one of my last 100 posts. Considering that I recently filtered out Facebook friends that I didn't know or only barely knew, I thought that I would have had higher numbers for both of these totals. Thus, I don't consider these scores to validate my feed as successful.

Overall, my last 100 posts span over a much shorter time frame than Allies, and this put into perspective that maybe I post too frequently. This could lead to people not finding a reason to constantly interact with my posts. While I don't believe that my feed is necessarily bad, I do believe that this study has pointed to some areas that I can change if I wish to make my feed more successful in my own terms. My results and analysis lead me to believe that I should perhaps post less often and also try to incorporate more updates about my life and wellbeing into my posts, leading to those I don't see often feeling compelled to interact.

Allie's Feed: Unsuccessful

Based on the parameters given to define success for Allie's Facebook feed, the analysis proved the feed to be unsucessful. The top three likers included two current rooommates and one good friend (who lives in the same city as me) and the top three commenters were comprised of myself, one of my current roommates, and one other good friend who also lives in the same city as me. These results are surprising since I initially thought that I interacted more with people living in a different area than me, but my Facebook feed has shown me otherwise. Important to note, however, is that I did have likes and comments from individuals living in different cities on my last 100 posts, but these individuals were not amongst the top people I interacted with. These results also beg the question of the purpose of my feed since I can interact face-to-face with my top likers and top commenters on a regular basis.

Moreover, the average number of likes I received per post was 11.36, which when looked at in comparison to my total number of friends, is not too high. Also, Mackenzie produced a similar average for the number of likes she received per post, but her total number of Facebook friends is ~300 less than mine; thus, I would argue that her feed is more successful than mine since herratio of number of likes per post to the total number of Facebook friends is higher.

My average number of comments per post is 1.61, which did not surprise me.

Overall, since my 100 posts spanned a time period of almost two years, I cannot blame my Facebook friends for not always commenting or liking my content since I do not generate content often. However, while there might not be proof of interaction with Facebook friends that live farther away, their presence in my total number of friends still shows some weak tie between myself and them. Thus, while my feed might not be successful based on the parameters I originally defined, I would argue that it is still relatively successful since content is being commented on and liked. Social media platforms define success based on the number of different actions that occur on a feed (likes, comments, shares, follows, etc.), so if I align myself with their perspective, my feed can be labeled relatively successful.

Conclusion

In terms of our project overall, we can draw two conclusions: first, that visualizing data did help us to uncover trends that were less obvious in the data set; and second, an individual's measurement of the success of their social feed does vary on what they are looking to get out of that social media platform. While both of these feeds wanted to see interaction with friends and acquaintances living farther away in order to be successful, someone else might define their feed as successful if they have a high number of interactions with colleagues who live in the same area as them (i.e., and provide answers to their questions when they post questions).

We chose to analyze two Facebook feeds together, rather than each of us looking at one, in order to compare different social media patterns amongst millenials and how they each define success of their feed. Moreover, we chose to look at our Facebook feeds because Facebook continues to dominate in the social media realm, according to a study put out by the American Press Institue. However, millenials do not necissarily go to Facebook nowadays to post their own content, but instead are searching for news or information from friends, family, and Facebook pages. Looking at the two feeds analyzed in this project, it is clear that one user follows this pattern closer since it takes her almost double the time of the other to reach 100 posts.

Studying the success of social feeds is relevant to today's digital culture. In an article in the Harvard Business Review, "What's Your Personal Social Media Strategy?", Soumitra Dutta persistently argues the importance of leaders within a company to have a voice on social media, Facebook and LinkedIn especially. Similarily to how Mackenzie and Allie defined success, Dutta urges company leaders to post public content and interact with "peers, employees, customers, and the broader public...in the same transparent and direct way they expect from everyone in their lives" (Dutta 2). Thus, a successful personal feed for a business executive involves reaching across connections, not just maintaining ties with individuals they interact with on a daily basis. While our feeds are examples of social feeds from millenials, not business executives, the same idea persists: a sucessful feed involves conversation, or at least some form of interaction, with everyone. Moreover, the more actions there are (likes, comments, shares, follows) the better.

This project set out to determine, via a self-tracking study and analysis of our own social media platforms, if an individual's Facebook feed is successful based on their predetermined definition of success. The data was then represented visually in order to get a better understanding of it. While both feeds proved to be unsuccessful based on the previous outlined parameters of success, the project as a whole succeeded in portraying the different ways the success of a social feed can be measured. Again, and as mentioned in class, a social media platform or an organization measures success differently than an individual user, but even amongst individuals the definition of success can vary.

Sources

Am, P. 0. (2015). How Millennials use and control social media. Retrieved December 02, 2016, from https://www.american-pressinstitute.org/publications/reports/survey-research/millennials-social-media/

Dutta, S. (2014). Managing Yourself: What's Your Personal Social Media Strategy? Retrieved December 01, 2016, from https://hbr.org/2010/11/managing-yourself-whats-your-personal-social-media-strategy

Kirk, Andy. Packt Publishingh (2012). A Successful Design Process. Retrieved November 08, 2016.

Canva.com (used to help create infographics and data visualizations)

Lampe, Cliff. Lecture Slides November 8.