Like, Tweet, Read

Exploratory Analyses of Social Media Data as an Indicator for Readership Behaviour in the Newspaper and Periodicals Industries

by

Tilman Queitsch

B.A., Ilmenau University of Technology, 2012

Project Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Publishing

in the
Publishing Program
Faculty of Communication, Art and Technology

SIMON FRASER UNIVERSITY

Fall 2014

Approval

Name: **Tilman Queitsch**

Degree: **Master of Publishing**

Title of Project: ***Like, Tweet, Read:***
Exploratory Analyses of Social Media Data as an Indicator for Readership Behaviour in the Newspaper and Periodicals Industries

Supervisory Committee:

John Maxwell, Ph.D.
Senior Supervisor
Director and Associate Professor,
Publishing Program

Juan Pablo Alperin, B.Math, M.A.
Supervisor
Instructor, Publishing Program

Alexandra Samuel, Ph.D.
Industry Supervisor
Vice President of Social Media
Vision Critical
Vancouver, British Columbia

Date Defended/Approved: December 4, 2014

Partial Copyright Licence

Ethics Statement

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

a. human research ethics approval from the Simon Fraser University Office of Research Ethics,

or

b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University;

or has conducted the research

c. as a co-investigator, collaborator or research assistant in a research project approved in advance,

or

d. as a member of a course approved in advance for minimal risk human research, by the Office of Research Ethics.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Burnaby, British Columbia, Canada

update Spring 2010

Abstract

Magazine and newspaper publishers benefit from readership studies conducted by large research organizations. They help publishing professionals keep track of readers' habits and their competitors' success. In most areas, surveys of readers and Internet users generate the findings that the publishing industry is interested in. In recent years, market research has developed a new approach combining such survey data with social media data. This approach offers new ways to analyze how social media audiences can be segmented, how readers choose between different media, how they use mobile devices, and how magazines or newspapers compare to their competitors.

Tackling each of these research scenarios, this report summarizes a series of analyses conducted at Vision Critical, a multinational market research technology company. By using basic functions in R, a freely available statistical programming language, the analyses show how this approach enriches results in a way that is useful for publishers.

Keywords: Social media; readership research; R; market research; statistics

Acknowledgements

This project greatly benefitted from the support I received from Alexandra Samuel during my work at Vision Critical, and the feedback from John Maxwell and Juan Pablo Alperin at SFU, who both inspired me to become way techier over the course of this project. On that note, I thank Noah Marconi, Director of Advanced Analytics at Vision Critical, for making me more curious about R.

I also highly appreciate the Mitacs Office's work as it helped me advance my career. My most sincere thanks go to my parents and my two sisters in Germany, and all of my friends. I would not be where I am today without all of you.

Table of Contents

List of Tables

List of Figures

List of Acronyms

Term	Initial components of the term
ARF	Angus Reid Forum
R	The R Project for Statistical Computing
SBA	Springboard America
VC	Vision Critical

Introduction

Industry surveys suggest that magazine and newspaper publishers still need to become more aware of how modern data ecosystems work, which possibilities data analysis offers, and how they can benefit from the insights it generates. A 2014 survey conducted by Cxense, a software firm, found that 78 percent of US news publishers do not know how many third parties are accessing their user data to drive their own revenue (Publishing Profitability Survey Shows Mix of Optimism, Naiveté 2012).

Research and data analysis can not only bring profitability, but also shape the strategies of companies, including publishing houses. Results published by marketing agencies and research organizations, such as the Pew Research Center, help publishers make strategic decisions in a variety of areas. For example, they can review current trends in usage of mobile devices to decide which ones to develop content for. In all of these areas, results can be complemented by combining them with another field that has become an everyday part of the publishing business for magazines and newspapers of all sizes—social media.

Even big papers struggle to understand social media users and to integrate them into their plans of action. Discussed by professionals across the industry, the *New York Times* innovation report that leaked in the spring of 2014 asked one essential question: “How can the Times become more digital while still maintaining a print presence, and what has to change?” Aside from the organizational concerns the report raises, it makes clear that understanding its social media followers and their behaviour would enable the *Times* to turn their 8 million fans on Facebook and more than 13 million followers on Twitter into a useful source of insight. This intersection between publishers’ research needs and social networks is exactly where market research can provide new possibilities through a combination of survey and social media data analysis. In recent years, this approach has been tested in various industries, yielding meaningful results: in

late 2013, the GfK Group connected the results of a social media analysis with survey data from its panels to show that users of YouTube, Facebook, and other platforms who had been exposed to a make-up campaign spent 20 percent more on cosmetic products than non-users (Waldheim 2013).

This report demonstrates how a combination of these two sources of insight – survey data and social media data – can provide value to publishers in the newspaper and periodicals industries. Summarizing a series of high-level analyses conducted at Vision Critical (VC), a multinational market research technology company, the report describes how these complimentary sources of insight can address known scenarios from readership and social media research. VC has published several reports and white papers on its projects combining social media with survey data, covering several fields such as the collaborative economy and charitable giving in North America (Owyang, Samuel and Grenville 2014). While all of these reports are based on studies that were specifically designed to address these areas, this report uses data from several separate surveys that focus on media usage. They were run as a part of VC's regular market panel surveys and not created with this particular project in mind. Hence this report does not describe one coherent study, but rather instances where a combination of survey and social media data can address research scenarios that publishers are interested in.

Looking at the research landscape that serves publishers, current solutions to known scenarios in the newspaper and periodicals industries do not integrate social media data to generate insights (see Part 1). To show how the offering can be complemented with a combination of survey responses and social media data, this report investigates four research scenarios (see Part 2). The first scenario analyzes how publishers can segment their social media audience based on data imported from Facebook. Scenarios two and three explore readers' media consumption and usage of mobile devices, respectively, highlighting patterns that show how social media usage affects these two kinds of readership behaviour. Lastly, the fourth scenario connects Facebook data with online readership in order to compare the popularity of two Canadian news websites. The results are discussed as a whole in Part 3.

Part 1.

The Research Landscape

To understand how social media data can complement the research for publishers, it is important to first review current findings in each of the relevant areas. By explaining methodologies and results for social media segmentation, choice of media, usage of mobile devices, and benchmarking, Part 1 provides the necessary topical background for each scenario.

1.1. Social media research: audience segmentation

As for segmenting one's audience on social media platforms, many market research agencies offer survey-based typologies of Facebook users. They argue that knowing about the different kinds of social media users helps publishers make informed decisions for their campaigns. Aimia, a company managing customer loyalty programs, has run several surveys to create a set of such social media personas. According to Aimia, so-called "no shows" make up a large part of the social media audience. They have low income and are barely active on social networks. On the opposite end of the scale are the "sparks", who are more active and deeply engaged social media enthusiasts, but make up only three percent of the adult US population (Rozen, Askalani and Senn 2012). Given that this is just one of innumerable approaches, there is definitely no perfect solution.

1.2. Readership research

1.2.1 Choice of media

For publishers it is essential to keep up with readership trends, such as which media different groups of readers prefer. One of the associations that publishes studies on this subject, the Pew Research Centre, saw social media on the rise as a news source in 2012. As opposed to print media and television, which both had lost the attention of large groups of media consumers, social media was continuously winning American readers. As the Centre's report notes, the percentage of people in the US who have read yesterday's news on a social media site had doubled from 9 to 19 percent between 2010 and 2012 alone (19 to 36 percent when looking at social media users only). The report also mentions that the amount of Americans who use a combination of traditional news sources, including print newspapers, increased to up to 38 percent since the last study. Hence, while social media has initially often been thought of as a channel driving traffic to publishers' own sites, it has also become a news source of its own (In Changing News Landscape, Even Television is Vulnerable 2012).

Studies like this one show media usage patterns of groups of readers, which affect newspaper and magazine publishers as much as they affect TV and radio broadcasters. They also remind publishers of the essential fact that readers differ by their media choices and that their usage patterns are always changing. The same can be said about the usage of mobile devices.

1.2.2 Usage of mobile devices

Readership patterns such as the one highlighted by the Pew Research Centre can also be influenced by usage of mobile devices, a research area investigated by a variety of organizations. Magazines Canada, for example, periodically publishes a Digital Magazine Fact Book covering this topic. The survey data for this public report comes from the Print Measurement Bureau, a Canadian non-profit organization that gathers statistics on print media readership. In the latest issue from 2013, the results show that Canadian ownership of e-reading devices – including tablets and e-readers – has

doubled from 9 to 21 percent between 2012 and 2013 (Digital Magazine Fact Book 2013). Compared to 2011, when only 7 percent of Canadians owned a tablet, these numbers show that mobile device ownership is increasing \ (Digital Magazines Fact Book 2011).

1.2.3. Benchmarking

Another topic of readership research relevant to publishers is how their publication compares to its competitors. Studies and metrics can tell them how many readers digital editions or websites attract, and allow them to distinguish top performers from magazines that have yet to become successful in digital markets. By doing so they can find out which publications were able to increase their readership, benchmark their own performance against them, and review their strategies to follow these more successful examples, which are often identified as such through industry studies. In 2013, GfK's Survey of the American Consumer showed not only that digital readership nearly doubled compared to the year before, but also that *ESPN The Magazine*, a sports title, was among the biggest winners in digital with about 1.1 million readers (Bazilian 2013).

1.2. The role of social media data for current research

In each of the above scenarios, social media data can enrich findings in several ways. In general, it allows publishers to see how the results differ among different groups or "segments" of Facebook and Twitter users.

The advantage of having data from Facebook available for audience segmentation is that it allows researchers to consider actual online behaviour when building their segments, without having to rely on self-reported usage of social networks, which may not reflect usage patterns accurately. To elaborate the importance of this difference, one might consider the accuracy of a user's guess of how often per month he or she posts an update to Facebook compared to the actual count as tracked by the platform itself. Hence, imported data widens the existing spectrum of possibilities to build segments that publishers can use for campaigns and other strategic decisions.

When it comes to readership research, analyses can show which kinds of media Facebook users with a lot of friends prefer and compare the results to respondents that have relatively few friends on this platform. As for the usage of mobile devices, analysts can explore if and how social media variables can predict how much time certain users spend on their smartphones and tablet PCs. Lastly, to complement publishers' ways of benchmarking their publication against competitors, social media data can clarify if certain successful titles attract more active or more influential social media users than their competition.

To explain some of the framework of the research, the following chapters will give an overview of the dataset, the methodology of the project, a brief description of the variables used in this report, and how they can be categorized.

Part 2.

Social Media Data Analysis

While the preceding chapters have explained the scenarios, Part 2 deals with the methodology of this project, including the structure of the data and the variables used for the analyses, as well as the actual analyses.

2.1. Research Overview and Methodology

2.1.1. Description of the dataset

As stated in the introduction, the data used for the following analyses is proprietary to VC. It was gathered in online surveys on two market panels that the company manages: the Angus Reid Forum (ARF), a Canadian online panel, and Springboard America (SBA), its US counterpart. The social media data used for this project was imported by VC after respondents had given the company permission to do so and self-reported in an online survey, depending on the variable in question (see Appendix A). The dataset used for this project contains 2,688 respondents who have provided Facebook usage data in one or both of the aforementioned ways and 1,231 respondents who have done the same for Twitter. As Figure 1 shows, there is an overlap of 871 people. All of these panelists have also taken a survey on their media usage.[1]

[1] All of the data points – survey responses, self-reported and imported social media data – stem from different surveys. They had to be merged into one dataset before the analyses. In this case, all files need to be merged using a unique identifier for respondent because it needs to be clear which piece of data belongs to which respondent so that the files can be combined correctly. R offers merge functions that can handle this case.

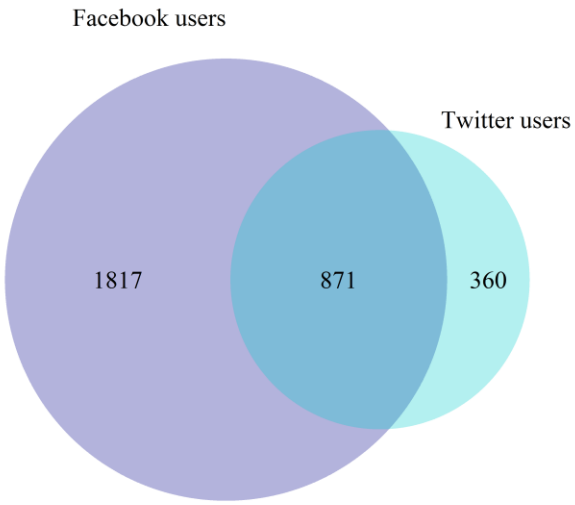


Figure 1 Dataset structure

2.1.2. Description of the variables used in this report

Looking at the social media variables imported from Facebook and Twitter, each of them can be understood as a different dimension of social media usage. While Facebook friends and Twitter followers say something about the influence of the respondent in question, the number of Facebook “likes” and Twitter friends can describe engagement with the content on the respective social network. The imported variables all contain behavioural data, i.e. data on the online behaviour of the respondents. The survey variables stem from studies conducted by VC on ARF and SBA, all of which are related to media consumption and the other scenarios outlined in Part 1. They can be considered different dimensions for measuring media usage and readership overall (see Appendix A for a complete list).

The first scenario, social media audience segmentation, will elaborate only social media variables to identify different groups of users. All other scenarios put them into relation with a readership variable in order to show how the two areas can correlate.

2.2. Analyses

2.2.1. Social media audience segmentation

For publishers, highly influential Facebook users might be a more attractive target group for campaigns than people with fewer connections. It is the definition of segments that can enable publishers to hone their strategies in this area. In this particular case, influential users can be identified through their number of Facebook friends. The number of influential Facebook users, for example, becomes clear when looking at the distribution of Facebook friends. This step also makes it easier to define concrete cut lines for more and less influential users.

The distribution of Facebook friends in Figure 2 indicates that the majority of Facebook users in this dataset have fewer than 250 connections on this platform.[2] Given the distribution, users with more than 500 or 750 friends can be considered as particularly influential.

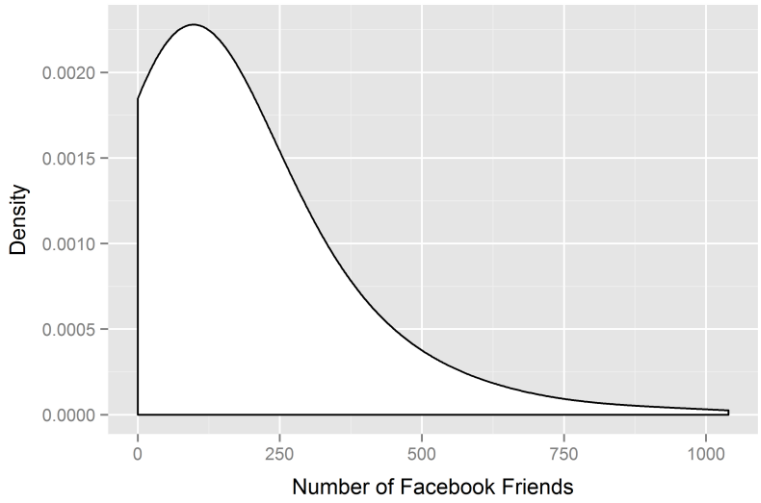


Figure 2 Kernel density plot for Facebook influence
Note. N=2129

[2] Like histograms, kernel density plots visualize the distribution of a variable, i.e. how often certain values occur in a dataset. For reasons of simplification, this plot only shows the area from 0 to about 1,000 Facebook friends (roughly four standard deviations of the plotted variable, the standard deviation being a unit for the dispersion of values from the average).

By adding visiting frequency to the chart, influence can be put into relation with another important variable (see Figure 3). The respondents can be divided up into three equally large groups according to how often they visit Facebook each week. The average number of Facebook friends in each of these groups, indicated by the dashed vertical lines, increases linearly from 106 (0 to 6 visits) to 176 (7 to 19 visits) to 240 (20 or more visits per week).

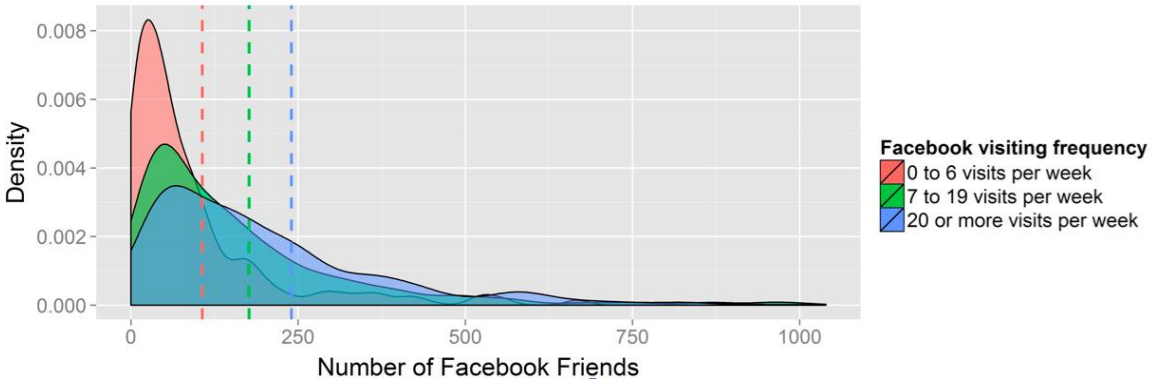


Figure 3 Kernel density plot for Facebook influence by visiting frequency
Note N=1575

When replacing visiting frequency with people's number of Facebook updates, similar differences occur. The means for the three groups in Figure 4 are 126, 191, and 235 Facebook friends—they increase just as steadily as in Figure 3. In other words, the more often respondents visit or post updates to Facebook, the more likely they are to have more friends.

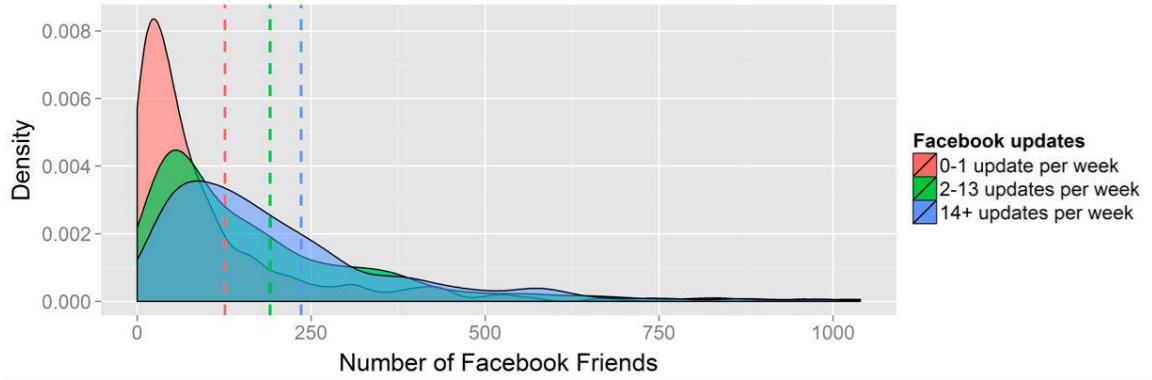


Figure 4 Kernel density plot for Facebook influence by updates
Note. N=2129

The number of Facebook friends might seem particularly interesting to publishers since it reflects influence and thereby people's potential to share campaign content with a lot of other users. Looking closer at the distribution for posting frequency can complement the picture of one's audience by determining how big "loud" and "quiet" certain groups of the audience are. For instance, publishers who wonder why their Facebook site does not get more comments from their followers might have to take into account that large parts of the audience on Facebook are, in fact, quiet (see Figure 5).[3]

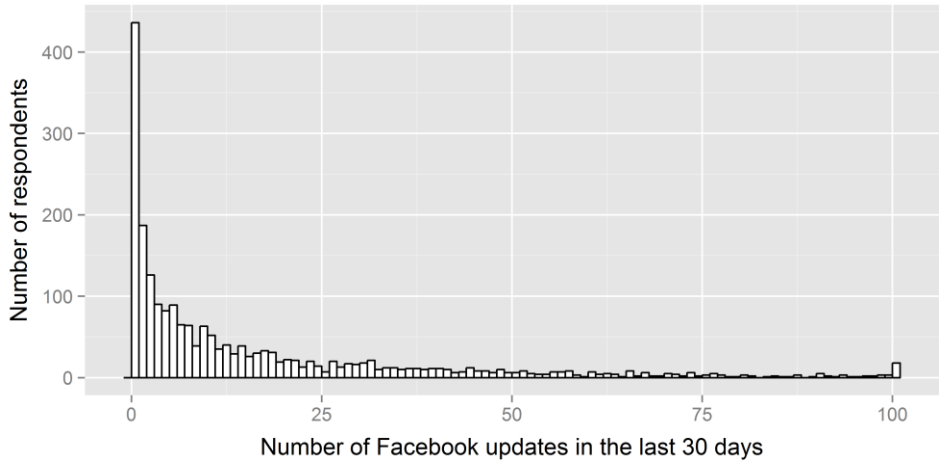


Figure 5 Histogram for Facebook updates
Note. N=2129

A big group of Facebook users has posted fewer than five updates in the past 30 days. This raises the question if those people are actually active on Facebook. Table 1, a crosstab of Facebook updates versus visiting frequency, can help describe this audience segment better.

[3] This histogram shows how many respondents have posted a certain number of Facebook updates in the last 30 days. Due to Facebook's restrictions on exporting data, the maximum value for this imported variable is set at 100 posts, so the respondents at the right end of this plot actually may have posted more updates.

Table 1 Facebook visiting frequency, segmented by number of updates

On average, how many times a week do you visit Facebook?	Number of Facebook updates posted in the last 30 days			Total
	0 to 1 FB update per week	2 to 13 FB updates per week	14 or more FB updates per week	
0 to 6 FB visits per week	333 56.3 %	76 9.8 %	44 5.8 %	453 21.3 %
7 to 19 FB visits per week	180 30.5 %	470 60.3 %	238 31.2 %	888 41.6 %
20 or more FB visits per week	78 13.2 %	233 29.9 %	482 63.1 %	793 37.2 %
Total	591 100.0 %	779 100.0 %	764 100.0 %	2134 100.0 %

$X^2=821.694 \cdot df=4 \cdot \Phi_c=0.439 \cdot p=0.000$

About 45 percent of respondents who posted up to one update in the last 30 days self-reportedly visit Facebook at least seven times a week, which is roughly equal to once a day.[4] In other words, even quiet Facebook users are actively visiting this social network. By finding out about the distribution for their individual follower base on social media, publishers can better adjust their strategies to their readers on social media overall, knowing that a large part of their audience might not be very vocal on this platform.

[4] The values at the bottom of this table are statistics that describe the extent of the relationship between the two variables. They refer to the entire table. Since it is not the purpose of this paper to explain the mathematics behind them, only a rough overview is given: in general, the chi-squared value X^2 indicates if the relationship is significant. In this case, 822 is a very high number and likely significant, which also depends on the degrees of freedom (df) and the significance level. The significance level is a threshold value of p that should not be exceeded by the p value of the analysis. Unless specified differently, if p is smaller than 0.05, the result can be considered as significant. Cramér's v, a value denoted as Φc, shows how significant and important the relationship actually is. It ranges from 0 to 1. In this case, 0.439 implies a very strong relationship (Crosstabulation with Nominal Variables n.d.).

This raises the question of why such a large part of the audience on Facebook is not posting updates and yet visiting the network. The answer might be simple: Facebook offers many different activities that do not involve posting, such as reading friends' updates. The passive consumption of content, which is not as measureable as liking and posting updates, always has been a part of the platform experience and presents a challenge for media strategists who want to drive measurable user engagement to their pages. As this scenario shows, data analysis can help publishers make this behaviour visible, allowing for more precise strategies, like testing which messaging or content resonates well with the quieter groups on social media.

2.2.2. Choice of media

This is the first research scenario in which social media data is put into relation with readership variables, such as newspaper readership. Before looking at the data, the question is which social media variable is most likely to predict or correlate the readership variable and can hence produce meaningful results in a contingency table. This variable will act as the column, which is also called a "banner" in market research, for the analysis, which is the base of respondents that researchers want to say something about. For this decision, researchers can look at the correlation between the variables they are interested in, i.e. each of the Facebook variables and print newspaper reading frequency.[5]

The amount of Facebook friends is most highly correlated with print newspaper reading frequency. As the results show, there are interesting differences in the reading habits of people who read newspapers daily and those who read them less often (see Table 2). There is a considerable change along the segments of the banner—in general, people who have a lot of friends on Facebook are less likely to belong to the group who

[5] This means that both variables are treated as numeric variables, even though they may be categorical (having values like "Agree" and "Disagree", for example), so that R can compute how a change in one variable affects the other. This procedure only makes sense when the scales of the categorical variables reflect an increasing or decreasing order, like "Agree," "Neither agree nor disagree," and "Disagree." It would not make sense to pass a categorical scale like "Agree," "Disagree," and "Don't know" to a correlation function since the third answer is not a meaningful successor of the second one. R offers a series of functions and tests to compute correlations and make decisions easier for analysts.

read print newspapers daily: from the lowest to the highest segment, the percentage of daily readers of print newspapers decreases from one third to about 19 percent, a change of roughly 13 percent.[6]

Table 2 Print newspaper reading frequency among Facebook users, segmented by influence

How often do you read print newspapers?	Five groups based on their number of friends on FB (segments based on quintiles)					Total
	Fewer than 39 FB friends	39 to 78 FB friends	79 to 149 FB friends	150 to 272 FB friends	273 or more FB friends	
Daily	131 32 %	139 32 %	105 24.4 %	81 19 %	81 18.9 %	537 25.2 %
Less often than daily	279 68 %	295 68 %	325 75.6 %	346 81 %	347 81.1 %	1592 74.9 %
Total	410 100.0 %	434 100.0 %	430 100.0 %	427 100.0 %	428 100.0 %	2129 100.0 %

$X^2=38.496 \cdot df=4 \cdot \Phi_c=0.134 \cdot p=0.000$

Looking at Table 2, it becomes clear how this particular dimension of Facebook usage can affect print newspaper readership. This pattern changes when using self-reported Facebook visiting frequency as a banner (see Table 3).

Interestingly, there is an inconsistent relationship between the two variables: going from left to right in the row for "Daily" in Table 3, the percentage of daily newspaper readers decreases for the segments for 0 to 14 visits per week and then increases again for people who visit between 15 and 29 times per week. The changes for the other reading frequency levels are similarly inconsistent (see Appendix B, Table 3a).

[6] This table only contains two levels for reading frequency: daily and less often than daily. This is to compare the two levels where the highest differences between the segments occur. The patterns are less discernable when looking at the other reading frequency levels ("Every few days," "Once a week or so," and less often; see the full tables in Appendix B).

Table 3 Print newspaper reading frequency among Facebook users, segmented by visiting frequency

How often do you read print newspapers?	On average, how many times a week do you visit Facebook? (segments based on quintiles)					Total
	0 to 5 visits per week	6 to 9 visits per week	10 to 14 visits per week	15 to 29 visits per week	30 or more visits per week	
Daily	126 30.5 %	111 25.3 %	84 22.2 %	128 27.1 %	98 22.7 %	547 25.6 %
Less often than daily	287 69.5 %	328 74.7 %	295 77.8 %	344 72.9 %	333 77.3 %	1587 74.3 %
Total	413 100.0 %	439 100.0 %	379 100.0 %	472 100.0 %	431 100.0 %	2134 100.0 %

$X^2=10.013 \cdot df=4 \cdot \Phi_c=0.068 \cdot p=0.040$

As for magazine readership, the results for both of these banners show less of a pattern: going from segment to segment, smaller differences occur between the percentages of daily readers. The pattern that was visible in Table 2 for print newspaper reading frequency does not reappear when looking at print magazines. The same goes for Facebook visiting frequency (see Tables 4 and 5).

Table 4 Print magazine reading frequency among Facebook users, segmented by influence

How often do you read print magazines?	Five groups based on their number of friends on FB (segments based on quintiles)					Total
	Fewer than 39 FB friends	39 to 78 FB friends	79 to 149 FB friends	150 to 272 FB friends	273 or more FB friends	
Daily	24 5.9 %	35 8.1 %	24 5.6 %	21 4.9 %	23 5.4 %	127 5.9 %
Less often than daily	386 94.1 %	399 91.9 %	406 94.4 %	406 95.1 %	405 94.6 %	2002 94 %
Total	410 100.0 %	434 100.0 %	430 100.0 %	427 100.0 %	428 100.0 %	2129 100.0 %

X^2=4.633 · df=4 · Φ_c=0.047 · p=0.327

Table 5 Print magazine reading frequency among Facebook users, segmented by FB visiting frequency

How often do you read print magazines?	On average, how many times a week do you visit Facebook? (segments based on quintiles)					Total
	0 to 5 visits per week	6 to 9 visits per week	10 to 14 visits per week	15 to 29 visits per week	30 or more visits per week	
Daily	23 5.6 %	25 5.7 %	22 5.8 %	27 5.7 %	35 8.1 %	132 6.2 %
Less often than daily	390 94.4 %	414 94.3 %	357 94.2 %	445 94.3 %	396 91.9 %	2002 93.9 %
Total	413 100.0 %	439 100.0 %	379 100.0 %	472 100.0 %	431 100.0 %	2134 100.0 %

X^2=3.505 · df=4 · Φ_c=0.041 · p=0.477

It has to be pointed out that there were not as many daily magazine readers as daily newspaper readers in this study. As for daily readers in this dataset, newspaper readers seem to be more strongly affected by Facebook usage in this comparison. Online news consumption habits follow different patterns. Tables 6 and 7 show that there is no discernable correlation for the imported variable for number of friends, but some correlation for Facebook visiting frequency.

Table 6 Online news reading/watching frequency among Facebook users, segmented by influence

How often do you read/watch online news?	Five groups based on their number of friends on FB (segments based on quintiles)					Total
	Fewer than 39 FB friends	39 to 78 FB friends	79 to 149 FB friends	150 to 272 FB friends	273 or more FB friends	
Daily	207 50.5 %	201 46.3 %	219 50.9 %	203 47.5 %	216 50.5 %	1046 49 %
Less often than daily	203 49.5 %	233 53.7 %	211 49.1 %	224 52.5 %	212 49.5 %	1083 50.8 %
Total	410 100.0 %	434 100.0 %	430 100.0 %	427 100.0 %	428 100.0 %	2129 100.0 %

$X^2=2.975 \cdot df=4 \cdot \Phi_c=0.037 \cdot p=0.562$

Table 7 Online news reading/watching frequency among Facebook users, segmented by FB visiting frequency

How often do you read/watch online news?	On average, how many times a week do you visit Facebook? (segments based on quintiles)					Total
	0 to 5 visits per week	6 to 9 visits per week	10 to 14 visits per week	15 to 29 visits per week	30 or more visits per week	
Daily	196 47.5 %	219 49.9 %	204 53.8 %	274 58.1 %	249 57.8 %	1142 53.6 %
Less often than daily	217 52.5 %	220 50.1 %	175 46.2 %	198 41.9 %	182 42.2 %	992 46.5 %
Total	413 100.0 %	439 100.0 %	379 100.0 %	472 100.0 %	431 100.0 %	2134 100.0 %

X^2=15.475 · df=4 · Φ_c=0.085 · p=0.004

To summarize, readers with a lot of friends on Facebook are less likely to read newspapers daily, but there is no similar effect for magazines. Periodicals, such as magazines, are published less frequently. The effort of reading a newspaper every day might make readers prefer other media. This seems possible, given that respondents visit Facebook more often the more friends they have.

Table 8 Facebook visiting frequency among FB users, segmented by number of friends

Facebook visiting frequency	Five groups based on their number of friends on FB (segments based on quintiles)					Total
	Fewer than 39 FB friends	39 to 78 FB friends	79 to 149 FB friends	150 to 272 FB friends	273 or more FB friends	
More than once a day	121 33.8 %	203 49.3 %	261 62.7 %	301 71.7 %	334 79.7 %	1220 60.3 %
About once a day	89 24.9 %	126 30.6 %	86 20.7 %	82 19.5 %	60 14.3 %	443 21.8 %
Several times a week	66 18.4 %	45 10.9 %	29 7 %	26 6.2 %	17 4.1 %	183 9 %
About once a week	38 10.6 %	24 5.8 %	20 4.8 %	10 2.4 %	6 1.4 %	98 4.9 %
Less often / Never	44 12.3 %	14 3.4 %	20 4.8 %	1 0.2 %	2 0.5 %	81 3.8 %
Total	358 100.0 %	412 100.0 %	416 100.0 %	420 100.0 %	419 100.0 %	2025 100.0 %

$X^2=298.751 \cdot df=16 \cdot \Phi_c=0.192 \cdot p=0.000$

Given these patterns, social media potentially replaces usage routines of other media, such as reading the newspaper every day. Magazines, being published less frequently, might require less of a commitment to consume. For this analysis, it has proven advantageous to have several Facebook variables available for banners since different dimensions of Facebook usage can correlate differently with reading habits. The same exploratory approach can be used for the other scenarios, such as the usage of mobile devices that people read content on, as the following chapter will elaborate.

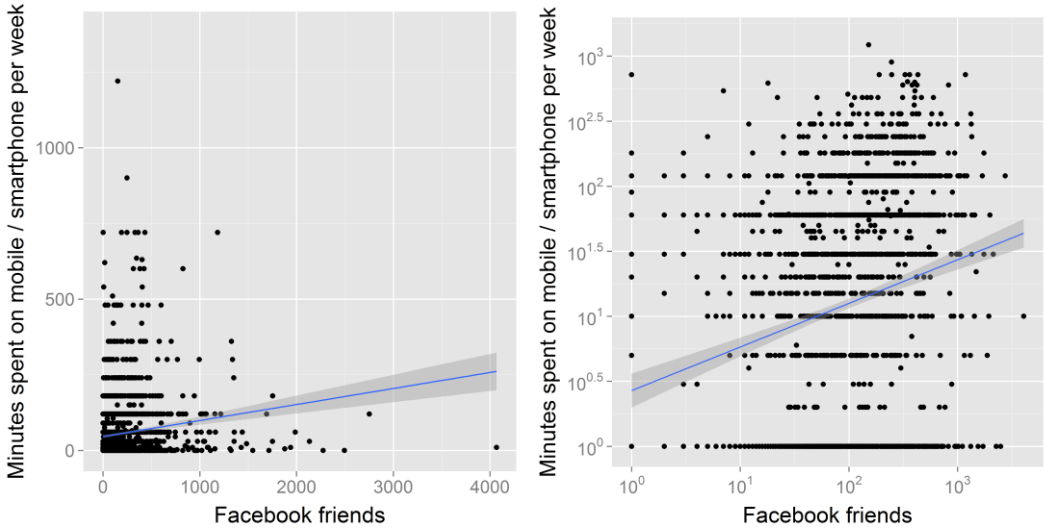
2.2.3. Usage of mobile devices

How much time people spend on their mobile devices can say something about the kind of content they expect from a magazine or newspaper on their tablet or smartphone: readers who are often browsing on these devices can be assumed to have different reading habits and demands toward content compared to those who barely use their mobile devices, for instance.

While contingency tables helped discuss the previous scenario, data visualizations can help identify correlation between continuous variables. The analysis is supposed to show how the time spent on a mobile device per week differs across social media variables in a meaningful way, i.e. if users who post a lot of updates spend more minutes per week on their tablets than those who are basically quiet on social networks.

Below are two scatter plots for the time respondents spend on their phones or smartphones each week and their number of Facebook friends. The regression lines show that there is a low positive correlation between the two variables. While the first plot has a linear scale in which the values for respondents with fewer than 500 Facebook friends are hard to discern, the second plot has a logarithmic scale, which “pulls” outliers closer to the other values in order to create a clearer visualization.[7]

[7] Plots with a logarithmic scales can be slightly harder to interpret compared to those with linear scales. In Figure 2, values on the x-axis increase by the same amounts from left to right (1,000, 2,000, 3,000), while the values in Figure 3 increase by higher and higher amounts (10, 100, 1,000). To many readers, an additive increase of the values from left to right seems more natural than the multiplicative increase used in Figure 3. Further explanations of how the data was transformed for Figure 3 can be found in the online repository (see appendix).

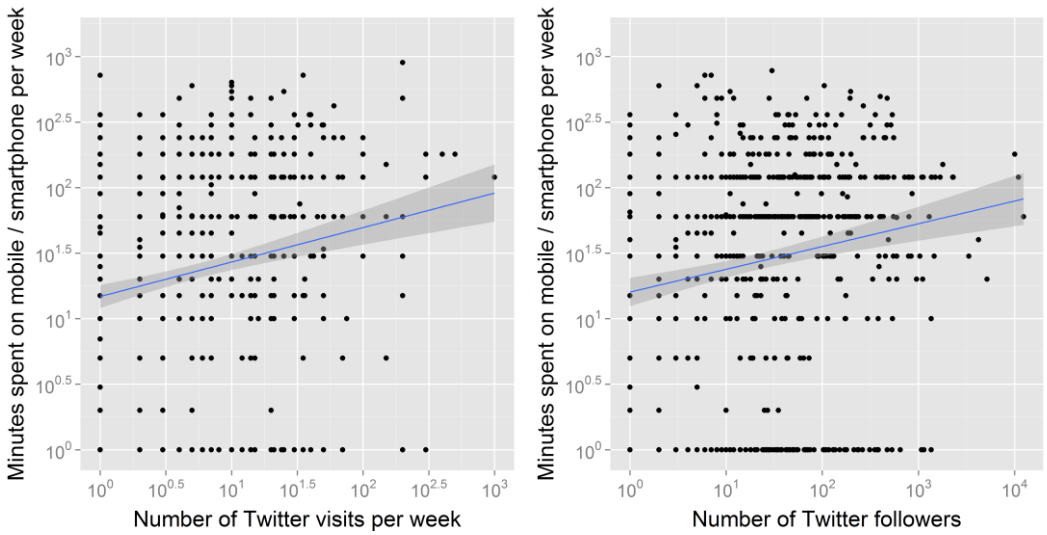


Figures 6-7 **Linear and logarithmic scatter plots of Facebook friends vs mobile/smartphone usage**

Note. N=2129

Even though a lot of respondents do not use smartphones or regular cell phones often, as indicated by the horizontal stroke at 100, the regression lines in both plots has a positive slope.[8] Accordingly, the number of Facebook friends can indicate how much time people spend on their phones and smartphones. Looking at the same plot for Twitter visiting frequency and iPad usage, the regression line shows a similar coherence between the two variables. A similar pattern shows up for number of Twitter followers.

[8] The regression line visualizes how one variable can predict the other. The grey area around the line is the confidence interval. It shows how, based on the data, R would predict the time spent on phones or smartphones for a given number of Facebook friends. Looking at Figure 2, for example, someone with 4,000 Facebook friends would spend about 250 minutes on their mobile device each week, give or take 50 minutes. Understandably, this interval is smaller where there are a lot of cases to be plotted, i.e. the confidence with which R can predict the values is higher.



Figures 7-9 **Logarithmic scatter plots of phone/smartphone usage vs Twitter visiting frequency and influence**

Note. N=850

The fact that the weekly duration of phone and smartphone usage is grouped together may have an effect on how meaningful the results are for publishers: non-smartphones do not necessarily allow users to access publishing content online. Still, mobile phone penetration in general can play a role for publishers that aim to reach their readers with mobile marketing campaigns that involve sending a text to enter a prize draw, for example.

Using the same method, one can look at the time respondents spend using iPad tablet computers and put it in relation to their Twitter visiting frequency.

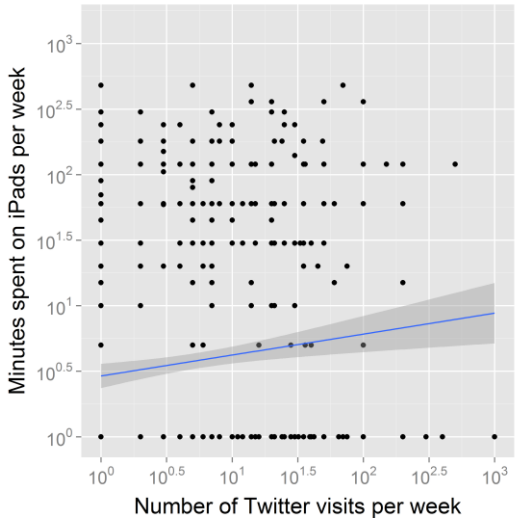


Figure 8 Logarithmic scatter plot of Twitter visiting frequency vs iPad usage
Note. N=850

As for time spent on tablets other than the iPad, on the other hand, the regression line is almost horizontal, indicating that there is little to no connection between the two variables.

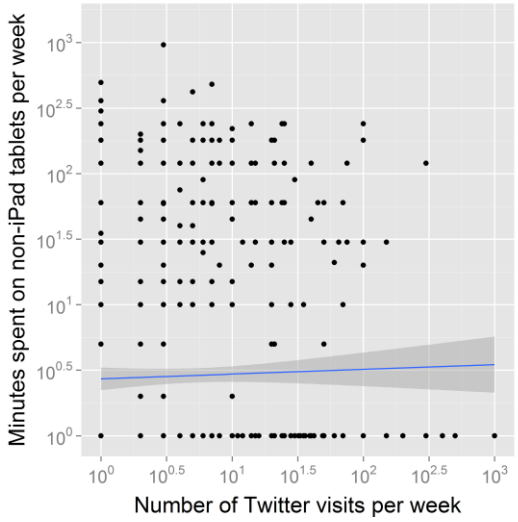


Figure 9 Logarithmic scatter plot of Twitter visiting frequency vs non-iPad tablet usage
Note. N=726

In this case it is worth to check the overall distribution of the variable again to see if there is even enough variation in the values to make for meaningful differences between any categories.

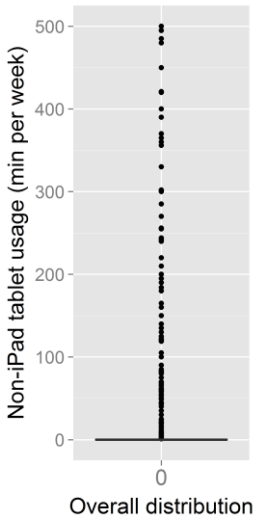


Figure 10 Box plot of non-iPad usage
Note. N=3048

The vast majority of respondents spends very little time on non-iPad tablets, which is why it is not surprising that no clear patterns show up in the scatter plots.[9]

In this scenario, the general duration of time spent on a phone, smartphone, or iPad was found to change along different categories for number of Facebook friends and to some degree by Twitter visiting frequency. Respondents may be using their mobile devices to stay in touch with their Facebook friends, the effort of which can increase and take more time more time as the number of friends increases. It is possible that differences between the two platforms come from the different level of engagement that

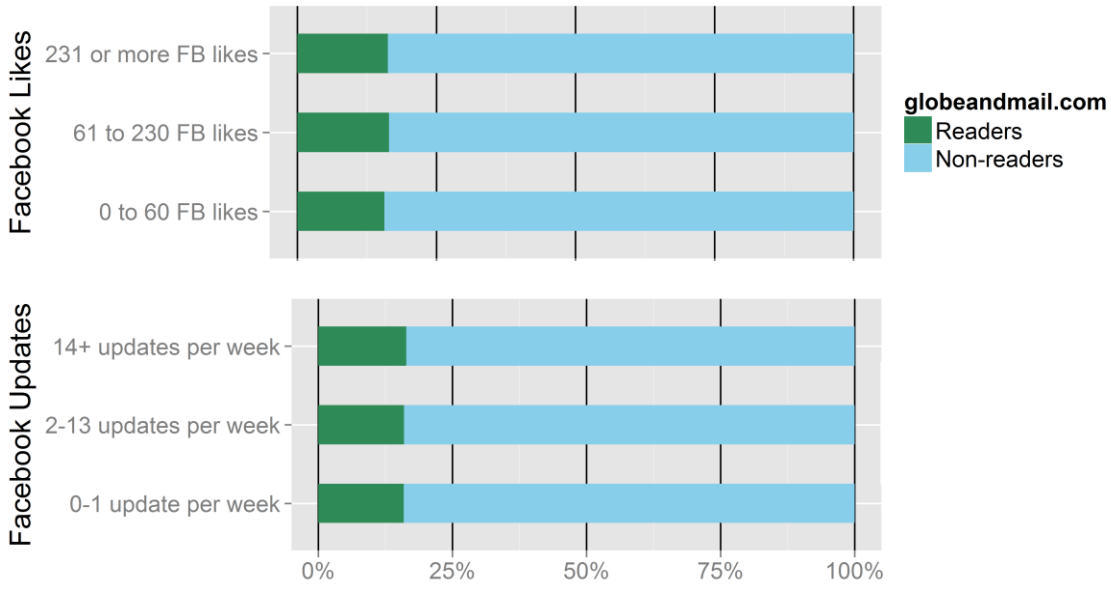
[9] This plot shows that the vast majority of respondents spend almost no time on non-iPad tablet computers each week. To elaborate, in a box plot, the lower and upper borders of the box are defined by the 25th and the 75th percentile. Percentiles split up the distribution according to their index, for example, 25 percent of all of the values are smaller than the 25th percentile. The thick horizontal stroke within the box is the median. Values above the 95th percentile are defined as outliers, i.e. values that are extremely high, and are plotted as dots. In Figure 8, there is no actual box and all of the above statistics are 0. The values greater than 0 are treated as outliers that form a vertical line above the box. They give the impression that there are numerous values bigger than 0, but in fact only 5 percent of the values plotted here are above 0.

they offer. Facebook might be used for more than just sharing an update with one's friends, but also playing one of the platform's games with them, while users' activities on Twitter may not go much further than tweeting an update to instantaneously reach all of their followers. These are just possible explanations for the differences shown in the charts, but they can help publishers see potential characteristics of social media users and to build their own personas of mobile device owners they want to reach on Twitter, for instance.

2.2.4. Benchmarking

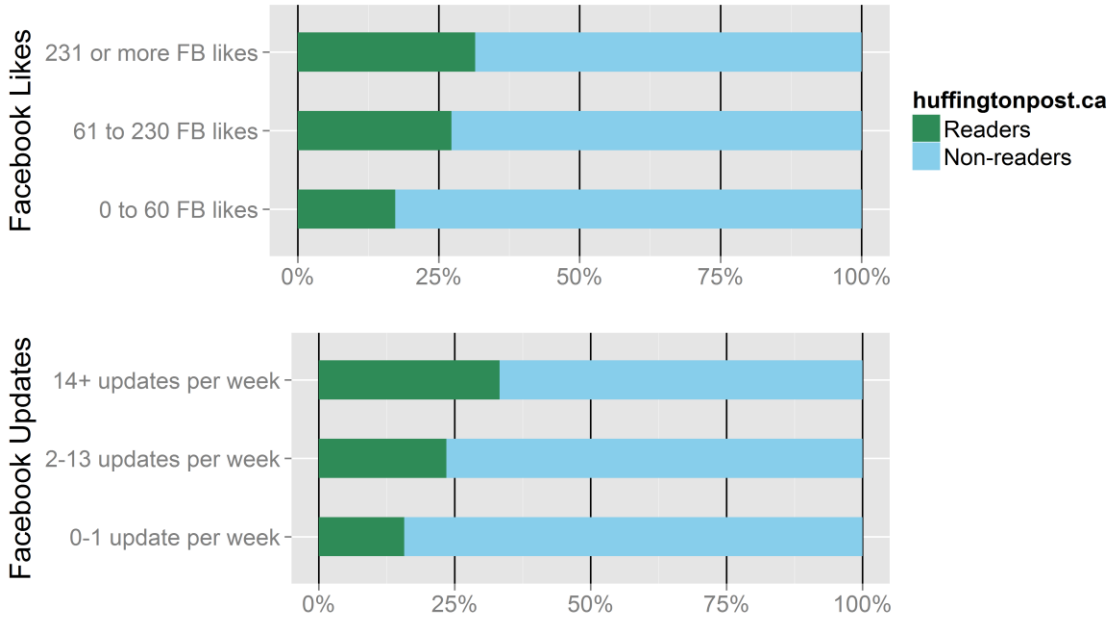
To be able to optimize their own digital strategies, it is useful for magazine and newspaper publishers to know which publications are more or less successful on social media than others. Measuring "success" in this context does not only include metrics like the number of "likes" on Facebook or how many unique visitors a certain magazine has, but also knowing among which kinds of social media users it is particularly known or popular. The social media variables for this scenario include the number of things, including pages, that users have liked on Facebook, and their number of status updates in the last 30 days before their data was imported.

The survey question that the readership variable for this scenario is based on asked respondents if they had visited several newspapers' websites in the past seven days. When comparing the website of *The Globe and Mail* with huffingtonpost.ca along the aforementioned Facebook variables, the following differences occur.



Figures 13-11 **Globeandmail.com readership vs Facebook engagement and activity**

Note. Total N=1757



Figures 15-12 **Huffingtonpost.com readership vs Facebook engagement and activity**

Note. Total N=1757

Looking at these charts, one could say that the online presence of the Huffington Post is more popular than The Globe and Mail's website among active Facebook users that either "like" a lot of content on Facebook or post more frequently than less active groups. These two examples are relatively large papers that have established more of an online readership than smaller newspapers. But the usage of social media also affects smaller papers' readership.

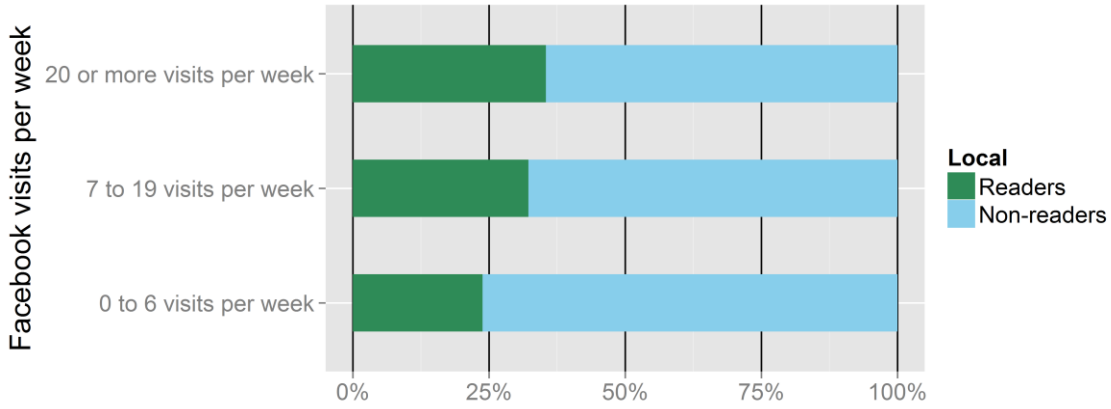


Figure 13 Readership of respondents' local newspaper's website vs Facebook engagement

Note. Total N=1816

The differences between the three groups in Figure 17 make sense when looking at possible ways for a paper to generate traffic through Facebook: many users might be lead to the site of a local newspaper by their friends who come from the same area and share articles of it on Facebook. As for The Globe and Mail and The Huffington Post, one could assume that people who like a lot of content on Facebook or visit the platform often are in general also more likely to come across a big newspaper's content and land on its website, but the question is still how some of them manage to be more successful than others in this coherence. It could be a certain popularity among social media users in general that drives sharing, liking, and commenting, or it might be the individual topics the publication is dealing with that make it differ from its competitors.

Part 3.

Discussion of the findings

This part of the report highlights trends and that appeared across several scenarios, discusses the limitation of this project, and gives suggestions for potential follow-up studies. Part 3 is followed by a summary of this project's implications, which underscores the advantages of data analysis for insight-driven publications.

1. Overall summary

Reviewing the analyses conducted and explained in this report, it becomes clear that correlations between social media data and readership behaviour can occur and be relevant to publishers. The fact that some of the findings did not hold up for Twitter implies that user behaviour differs in a meaningful way, a finding that in itself is of interest for publishers. Users of Facebook and Twitter should not be treated as one and the same since they probably use these networks for different reasons, to do different things. Their use of these platforms can also play a role in their overall media consumption, including magazines and newspapers.

In several analyses, using different social media variables lead to different results, which emphasizes the importance of testing how a variety of variables correlates with readership behaviour. Some patterns appeared for one, but did not reappear for other variables, like the one for daily newspaper readers among Facebook users (see Chapter 2.2.2.). This shows that the number of Facebook friends, visiting frequency, and the other social media variables can not be understood as variables measuring the same concept, which in this case would be online engagement.

How meaningful the findings are to publishers largely depends on the sample used for the analysis. Ideally, a large part of the sample should be made up of their magazine's followers on social media and also include potential readers. Other than this particular point, there are more limitations of the findings of this project that are worth discussing.

2. Limitations of the project and its findings

The results presented here are descriptive, they have little meaning for the larger populations of Facebook users whose attitudes and usage likely differs from the respondents who take part in VC's surveys. The dataset used for this project contained data gathered at different time points in different studies which raises methodological concerns, yet it does represent a considerable aggregation of social media and survey data. As such, it is appropriate to demonstrate the use cases and the value this kind of research can deliver to magazine and newspaper publishers as well as clients in other industries.

As it was not the aim of this project to make specific conclusions about the population of Canada, the United States of America, or the entirety of Facebook users, none of the datasets was weighted to represent any of these groups.[10] Hence, this project does not aim to deliver insights beyond the dataset involved, which is a combination of surveys conducted in 2013 and 2014.

Between the import of the social media data and the online surveys used for the analysis lay several months. One could argue that attitudes and behaviour do not change within months to an extent that would play a role for this project, but

[10] In general, weighting increases the extent to which the base of respondents can be considered representative of a certain population by assigning each respondent a multiplier so that, for instance, the proportions of inhabitants of certain Canadian provinces in the sample reflect the actual proportions as captured by the Canadian census. In terms of social media variables, like the number of Facebook friends, this is hard to achieve because the actual distribution of Facebook friends on the entire network is not available publicly.

methodologically it may have made a difference to collect both kinds of data at a single point in time.

3. The potential of social media research

The results presented here are examples of how the use of social media, broken up into the variables used for the analysis, can explain and deliver context for readership behaviour. Several benefits of this approach have been pointed out, such as the advantage of having several Facebook variables available for analysis, thereby taking into account how diverse users can be on each of these levels.

Although this new combination of data brings many possibilities, its value for publishers depends on the focus of the underlying research, which should specifically be designed for their industry. Larger studies measuring reading habits and behaviour with more diverse variables could advance both readership and social media research. These projects could include a survey on how and why people share a magazine's updates, and use segmentations based on social media data to identify meaningful differences.

For such a research project, one of the obvious challenges is to convince users of the valuable contribution they make by giving researchers access to their social media data. While it presents less of a problem to ask respondents in a survey how often they visit Facebook each week, it is more difficult to get their permission to import their number of status updates, for instance. It may simply depend on who is asking them for their data. A newspaper that they follow or like on Facebook anyway will be more likely to get access than an organization they have no connection to. Hence, such a project could be feasible if all other requirements, including having the necessary technology to gather all of the data, are met.

Implications

At the time of writing this report, the approach of combining data from surveys and social media is still relatively young, but within market research, one can expect it to keep gaining momentum. As for the development of this research approach toward an essential market research service, other industries than publishing are more likely to drive this development. While this paper has explained and demonstrated use cases for this new combination of data particularly for publishers, one important question for them is how they can gather this data using methods they have, especially in the case of smaller publishing houses. While this is undoubtedly determined by one's research budget and personnel, it can be advantageous to have bigger companies in other industries drive the progress of this new method. With larger research budgets than those in publishing, they can afford to run iterations of studies that help market research refine its methodology so that publishers can benefit from this progress at a later point. In this context, market research acts both as a forerunner and an innovator to develop services relevant to research clients in the publishing industry.

Besides making the results and use cases more palpable for the readers of this report, the other reason for including details of analytics and statistics is that there are enough opportunities available for publishers to hone their abilities of making sense of data, especially through freely available statistical software like R. The increase in popularity that R has experienced in recent years is remarkable. In 2013, its growth of capability outpaced the one of SAS, a widely used business analytics software (Muenchen 2013). Once publishers realize the potential in their businesses to refine their strategies with data analysis, they can make use of both the vast amount of available resources and the expertise of a large and growing user community. This can help them generate insights from surveys, online metrics, and social media data imported through application programming interface calls.

Opportunities to gather meaningful data are abundant in publishing. As industry experts have pointed out, every magazine or newspaper has chances to connect with its readers, be it through social media or other ways, and establish a connection that lets them share data with it. Emilie Harkin, marketing director at *Foreign Affairs*, emphasizes the role of information-sharing for her magazine. “Subscribers share information with us, and that data is valuable.” *Foreign Affairs* complements these sources with independent surveys from Erdos & Morgan to embed data analysis in its business. Harkin adds that, “Working alongside smart, organized analytics and data experts is really the best way to find order in the chaos.” For the whole team, skills in data mining and statistics play a considerable role, as she points out, “I am extremely fortunate to collaborate with colleagues who can see stories about our audiences in lines of raw data" (Peck 2014).

However, the benefits of data analysis for magazines or newspapers for their own ongoing research depends on the publication in question. But, by having more information on their readers, any publisher is able to describe his or her particular audiences better and to give advertisers more incentives to book ads or custom content with them. For instance, magazines with influential followers on social media that they have data on can calculate how big the potential readership of a promoted post can get if it is shared. This way they can offer advertisers more diversified and attractive ad packages.

Considering new approaches like the one explored in this report, the range of methods for research and data analysis available to publishers is increasing. Along with it, publishers have to become aware of how important this field is to them. Realizing how and where they can implement it in their own businesses will be another challenge and, in a time where all kinds of industries are becoming increasingly creative in how they can use data to their advantage, it also will be their own responsibility.

Bibliography

Online Sources

Bazilian, Emma. 2013. *Magazine Readership Growing, Survey Shows.* May 29. Accessed October 3, 2014. http://www.adweek.com/news/press/magazine-readership-growing-survey-shows-149863.

n.d. "Crosstabulation with Nominal Variables." *Computing in the Humanities and Social Sciences, Faculty of Arts & Science, University of Toronto.* Accessed October 3, 2014. http://groups.chass.utoronto.ca/pol242/Labs/LM-3A/LM-3A_content.htm.

2013. "Digital Magazine Fact Book." *Magazines Canada.* Accessed October 3, 2014. http://www.magazinescanada.ca/uploads/File/AdServices/FactBooks/2013/DigitalMagazineFactBook2013Final-Eng.pdf.

2011. "Digital Magazines Fact Book." *Magazines Canada.* Accessed October 3, 2014. http://www.magazinescanada.ca/uploads/File/AdServices/FactBooks/2011/DigitalEN.pdf.

2012. *In Changing News Landscape, Even Television is Vulnerable.* September 27. Accessed October 3, 2014. http://www.people-press.org/2012/09/27/in-changing-news-landscape-even-television-is-vulnerable/.

Muenchen, Robert A. 2013. *R's 2012 Growth in Capability Exceeds SAS' All Time Total.* March 19. Accessed October 3, 2014. http://r4stats.com/2013/03/19/r-2012-growth-exceeds-sas-all-time-total/.

Owyang, Jeremiah, Alexandra Samuel, and Andrew Grenville. 2014. " Sharing Is The New Buying." *Vision Critical.* Accessed October 3, 2014. http://www.visioncritical.com/sites/default/files/pdf/sharing-new-buying-collaborative-economy-report.pdf.

Peck, Gretchen A. 2014. *Data-Driven Publishing: Know Thy Audience.* February. Accessed October 3, 2014. http://www.pubexec.com/article/data-driven-publishing/1.

2012. *Publishing Profitability Survey Shows Mix of Optimism, Naiveté.* May. Accessed October 3, 2014. http://www.cxense.com/publishing-profitability-survey-shows-mix-optimism-naivete.

Rozen, Doug, Mona Askalani, and Tom Senn. 2012. "Staring at the sun: Identifying, understanding and influencing social media users." Accessed October 3, 2014. http://www.pamorama.net/wp-content/uploads/2012/06/Aimia-Social-Media-White-Paper-6-types-of-social-media-users.pdf.

Waldheim, Christian. 2013. *Establishing Social Media ROI through Social Media Analysis and Established Techniques.* September 19. Accessed October 3, 2014. http://blog.gfk.com/2013/09/establishing-social-media-roi-through-social-media-analysis-and-classical-research/.

Appendix A.

List of variables

Social media variable	Type	Mean	Median	Standard deviation	Used for
Number of FB friends	continuous	187.5	110.0	259.8	SMAS, COM, UMD, B
FB visits (per week)	continuous	22.4	14.0	56.6	SMAS, COM, B
FB updates in the last 30 days (before data was imported)	continuous	15.4	6.0	21.3	SMAS,
Number of FB “likes"	continuous	237.7	94.0	456.8	B
Twitter visits (per week)	continuous	16.5	5.0	50.0	UMD
Number of Twitter followers	continuous	173.1	20.0	802.0	UMD

Except for Facebook visiting frequency, all social media variables have been imported from the respective networks.

Readership behaviour variable	Type	Mean	Median	Standard deviation	Used for
Print newspapers reading frequency	categorical				COM
Print magazines reading frequency	categorical				COM
Online news reading/watching frequency	categorical				COM
Phone / smartphone usage (minutes per week)	continuous	41.0	10.0	83.8	UMD
iPad usage (minutes per week)	continuous	22.9	0.0	61.1	UMD
Non-iPad tablet usage (minutes per week)	continuous	19.7	0.0	60.8	UMD
Reading globeandmail.com (in the week before the survey)	categorical (yes/no)				B
Reading huffingtonpost.ca (in the week before the survey)	categorical (yes/no)				B
Reading one's local newspaper's website (in the week before the survey)	categorical (yes/no)				B

Note. SMAS = social media audience segmentation, COM = choice of media, UMD = usage of mobile devices, B = benchmarking

The R script files for all plots and research scenarios can be freely accessed at https://github.com/tiQu/LikeTweetRead/.

Appendix B.

Contingency tables for all levels of reading frequency

Table 2a **Print newspaper reading frequency among Facebook users, segmented by influence**

Print newspapers	Five groups based on their number of friends on FB (segments based on quintiles)					Total
	Fewer than 39 FB friends	39 to 78 FB friends	79 to 149 FB friends	150 to 272 FB friends	273 or more FB friends	
Daily	131 32 %	139 32 %	105 24.4 %	81 19 %	81 18.9 %	537 25.2 %
Every few days	69 16.8 %	80 18.4 %	73 17 %	65 15.2 %	61 14.3 %	348 16.4 %
Once a week or so	82 20 %	102 23.5 %	108 25.1 %	111 26 %	107 25 %	510 24 %
Once a month or so	30 7.3 %	33 7.6 %	34 7.9 %	53 12.4 %	52 12.1 %	202 9.5 %
Less than once a month	61 14.9 %	53 12.2 %	74 17.2 %	79 18.5 %	80 18.7 %	347 16.4 %
Never (do not ever use this type of media)	37 9 %	27 6.2 %	36 8.4 %	38 8.9 %	47 11 %	185 8.7 %
Total	410 100.0 %	434 100.0 %	430 100.0 %	427 100.0 %	428 100.0 %	2129 100.0 %

$X^2=61.017 \cdot df=20 \cdot \Phi_c=0.085 \cdot p=0.000$

Table 3a **Print newspaper reading frequency among Facebook users, segmented by visiting frequency**

Print newspapers	On average, how many times a week do you visit Facebook? (segments based on quintiles)					Total
	0 to 5 visits per week	6 to 9 visits per week	10 to 14 visits per week	15 to 29 visits per week	30 or more visits per week	
Daily	126 30.5 %	111 25.3 %	84 22.2 %	128 27.1 %	98 22.7 %	547 25.6 %
Every few days	71 17.2 %	74 16.9 %	58 15.3 %	69 14.6 %	74 17.2 %	346 16.2 %
Once a week or so	98 23.7 %	106 24.1 %	104 27.4 %	117 24.8 %	93 21.6 %	518 24.4 %
Once a month or so	27 6.5 %	39 8.9 %	36 9.5 %	50 10.6 %	54 12.5 %	206 9.6 %
Less than once a month	55 13.3 %	67 15.3 %	66 17.4 %	72 15.3 %	77 17.9 %	337 15.8 %
Never (do not ever use this type of media)	36 8.7 %	42 9.6 %	31 8.2 %	36 7.6 %	35 8.1 %	180 8.5 %
Total	413 100.0 %	439 100.0 %	379 100.0 %	472 100.0 %	431 100.0 %	2134 100.0 %

$X^2=25.219 \cdot df=20 \cdot \Phi_c=0.054 \cdot p=0.193$

Table 4a **Print magazine reading frequency among Facebook users, segmented by influence**

Print magazines	Five groups based on their number of friends on FB (segments based on quintiles)					Total
	Fewer than 39 FB friends	39 to 78 FB friends	79 to 149 FB friends	150 to 272 FB friends	273 or more FB friends	
Daily	24 5.9 %	35 8.1 %	24 5.6 %	21 4.9 %	23 5.4 %	127 5.9 %
Every few days	60 14.6 %	58 13.4 %	60 14 %	39 9.1 %	45 10.5 %	262 12.2 %
Once a week or so	101 24.6 %	93 21.4 %	80 18.6 %	74 17.3 %	70 16.4 %	418 19.7 %
Once a month or so	80 19.5 %	99 22.8 %	114 26.5 %	125 29.3 %	119 27.8 %	537 25.4 %
Less than once a month	98 23.9 %	115 26.5 %	116 27 %	122 28.6 %	115 26.9 %	566 26.5 %
Never (do not ever use this type of media)	47 11.5 %	34 7.8 %	36 8.4 %	46 10.8 %	56 13.1 %	219 10.3 %
Total	410 100.0 %	434 100.0 %	430 100.0 %	427 100.0 %	428 100.0 %	2129 100.0 %

$X^2=42.030 \cdot df=20 \cdot \Phi_c=0.070 \cdot p=0.003$

Table 5a **Print magazine reading frequency among Facebook users, segmented by FB visiting frequency**

Print magazines	On average, how many times a week do you visit Facebook? (segments based on quintiles)					Total
	0 to 5 visits per week	6 to 9 visits per week	10 to 14 visits per week	15 to 29 visits per week	30 or more visits per week	
Daily	23 5.6 %	25 5.7 %	22 5.8 %	27 5.7 %	35 8.1 %	132 6.2 %
Every few days	62 15 %	59 13.4 %	46 12.1 %	55 11.7 %	61 14.2 %	283 13.4 %
Once a week or so	78 18.9 %	105 23.9 %	77 20.3 %	88 18.6 %	74 17.2 %	422 19.8 %
Once a month or so	98 23.7 %	93 21.2 %	92 24.3 %	135 28.6 %	98 22.7 %	516 24.2 %
Less than once a month	106 25.7 %	111 25.3 %	107 28.2 %	136 28.8 %	114 26.5 %	574 26.9 %
Never (do not ever use this type of media)	46 11.1 %	46 10.5 %	35 9.2 %	31 6.6 %	49 11.4 %	207 9.8 %
Total	413 100.0 %	439 100.0 %	379 100.0 %	472 100.0 %	431 100.0 %	2134 100.0 %

$X^2=26.298 \cdot df=20 \cdot \Phi_c=0.056 \cdot p=0.156$

Table 6a **Online news reading/watching frequency among Facebook users, segmented by influence**

Online news	Five groups based on their number of friends on FB (segments based on quintiles)					Total
	Fewer than 39 FB friends	39 to 78 FB friends	79 to 149 FB friends	150 to 272 FB friends	273 or more FB friends	
Daily	207 50.5 %	201 46.3 %	219 50.9 %	203 47.5 %	216 50.5 %	1046 49 %
Every few days	81 19.8 %	82 18.9 %	82 19.1 %	95 22.2 %	99 23.1 %	439 20.8 %
Once a week or so	45 11 %	48 11.1 %	48 11.2 %	52 12.2 %	43 10 %	236 11.1 %
Once a month or so	21 5.1 %	28 6.5 %	23 5.3 %	23 5.4 %	25 5.8 %	120 5.7 %
Less than once a month	25 6.1 %	35 8.1 %	39 9.1 %	33 7.7 %	25 5.8 %	157 7.4 %
Never (do not ever use this type of media)	31 7.6 %	40 9.2 %	19 4.4 %	21 4.9 %	20 4.7 %	131 6.2 %
Total	410 100.0 %	434 100.0 %	430 100.0 %	427 100.0 %	428 100.0 %	2129 100.0 %

$X^2=23.303 \cdot df=20 \cdot \Phi_c=0.052 \cdot p=0.274$

Table 7a **Online news reading/watching frequency among Facebook users, segmented by FB visiting frequency**

Online news	On average, how many times a week do you visit Facebook? (segments based on quintiles)					Total
	0 to 5 visits per week	6 to 9 visits per week	10 to 14 visits per week	15 to 29 visits per week	30 or more visits per week	
Daily	196 47.5 %	219 49.9 %	204 53.8 %	274 58.1 %	249 57.8 %	1142 53.6 %
Every few days	80 19.4 %	79 18 %	76 20.1 %	96 20.3 %	98 22.7 %	429 20.1 %
Once a week or so	54 13.1 %	42 9.6 %	48 12.7 %	40 8.5 %	33 7.7 %	217 10.1 %
Once a month or so	24 5.8 %	35 8 %	13 3.4 %	16 3.4 %	16 3.7 %	104 4.7 %
Less than once a month	29 7 %	40 9.1 %	22 5.8 %	24 5.1 %	19 4.4 %	134 6.3 %
Never (do not ever use this type of media)	30 7.3 %	24 5.5 %	16 4.2 %	22 4.7 %	16 3.7 %	108 4.9 %
Total	413 100.0 %	439 100.0 %	379 100.0 %	472 100.0 %	431 100.0 %	2134 100.0 %

X^2=49.945 · df=20 · Φ_c=0.076 · p=0.000