Crossing the Chasm:
The Political Value of Facebook and Implications for Modern Congressional Campaigns

By

Kevin Guiang

A Senior Honors Thesis Submitted to the Department of Political Science,
University of California, San Diego
April 1, 2013

I dedicate this work to Seth Hill, Gerry Mackie, and James Fowler

Thank you for guiding me to achieve the joy of discovery - a learning process I will forever hold dear.

It is a mania shared by philosophers of all ages to deny what exists and to explain what does not exist. – Rousseau

Contents

Tables & Figures

**organized by chapter*

Chapter 1 – The Digital Landscape

Introduction

Many questions have arisen as a result of the political uprisings in the Middle East to the advent of unrestricted information dissemination in East Asia and Russia that have encouraged academics to study the role of modern social communication with respect to contemporary political participation. Academics assert that regardless of the role of Twitter in the Arab Spring, uprising would have occurred despite the accessibility of modern communications. While conversing with Professor Samuel Popkin, he argued - assuming the role of devil's advocate - that even in the absence of social media, protestors would still have successfully lined the walks of the Lincoln memorial in advocacy of their civil rights in the 1960's. It is easy to perceive this technology to assume the role of revolutionary beyond the ages, but it is truth that society has discovered means by which to communicate and congregate –on massive scales- in the absence of it. By what perspective, then, should we look at social media with respect to not only social relations, but also political participation in the United States? Is it a medium by which almost any candidate, any journalist, any political dissenter can access? Is it – in theory – the most disruptive social equalizer this, or perhaps any generation has the privilege of experiencing? Is it then, a combative medium by which large portions of society have the capacity to metaphorically usurp the media monopoly of the corporate kings? We understand that it is a disruptive technology, but exactly to what extent, and by when, will it assimilate itself with contemporary society?

Research Overview

I seek to contribute to the academic discourse on modern communications and Political Science by examining the political value of Facebook. I chose Facebook primarily because of its vast user base and integration with American society – 1 in 7 people are on Facebook in the entire world, while more than half of Americans access the site on a daily basis[1]. I also chose Facebook on the relative transparency of its methods of organization, primarily on the news feed, known as EdgeRank.

The concept behind my study derives from findings of research conducted by notable Political Scientists including Alan Gerber and Donald Green, Charles Atkin, and James Fowler. I primarily touch upon [1] social pressure and political participation [2] the beliefs that mere exposure to TV ads does not necessarily correlate with audience engagement, and that [3] social actions conducted by individuals within a specific network incline other individuals within that network to move to similar action – riding upon the assumption that social nodes organize themselves in accordance with similar interests.

Understanding these core concepts, I seek to contribute to the academic discourse on Political Science by discovering a relative measure of the political value of Facebook. Because a large proportion of Americans have taken interest in the site, it is possible to question to what extent its use has benefitted politicians due to its low barrier to access. It is important to differentiate between the importance of Facebook for Congressional campaigns from those running for the seat of the President because Congressional candidates, by default, have more limited appeal than candidates for President. With a

[1] Facebook Newsroom – Key Facts http://newsroom.fb.com/Key-Facts

limited constituency, candidates must engage their voter bases with appeals specific-to-specific demographic groups.

Therefore, I seek to answer several questions, including how well Congressional candidates are engaging their constituents on Facebook, what appeals - and what types of posts - receive the most engagement on Facebook, and during what periods of time do candidates yield the highest levels of engagement, if any. I also seek to discover a relationship between levels of user engagement on Facebook with respect to [1] frequency of posts [2] the number of "likes" associated with a page, and [3] district level demographics, including but not limited to constituent education levels and income.

I developed a research model by which I selected 70 Congressional districts on the basis of [1] the margin of victory [2] identification of "tossup" by the Cook Political Report, Rothenberg Political Report, Roll Call, and Open Secrets and [3] campaign spending, selecting districts by which the victor spent less than his/her opponent. From here, I accessed the Facebook pages of the top 2 candidates from each district, logging the number of likes, comments, shares, and date of each post from October 1st until ~ November 7-November 10, the time at which the candidate either expressed gratitude in winning the election, or expressed his/her concession. The final N value of my original data set resulted in the analysis of 3,533 different posts.

I established a system by which I would categorize each post in accordance with 15 different classifications, ranging from types of text posts, photo posts, and video posts. I also logged the number of likes the page achieved, the primary audience that accessed the page, retrieved from Facebook insights, as well as other district level and candidate level variables accessible on the basis of online mediums.

From here, I was able to develop a formula expressing the engagement efficacy of each post and candidate. On both the basis of post type and candidate, I accumulated the sum of the engagement values for likes, comments and shares.

I organized the 15 posts in to subsets, dividing them by appeals to candidate intimacy, civic duty, the national election, the local election, or general campaign information.

Within my dataset, I was able to find [1] what individual type of Facebook post receives the greatest levels of engagement on each basis of likes, comments, and shares [2] what modes of appeals on Facebook result in the highest levels of engagement, on the basis of likes, comments, and shares [3] the relationship between post frequency over time and user engagement [4] variables contributing to the highest levels of engagement efficacy.

My findings, in short, reveal that candidates have not yet demonstrated substantial levels of proficiency associated with Facebook user engagement. On the basis of content appeals on Facebook, appeals to civic duty and candidate intimacy, on average, are better able to engage users on Facebook than do content appeals related to the national politics, local politics, or general information about the campaign. While candidates post more on Facebook on the days leading up to the election, late campaign advertising does not, on average, yield greater levels of engagement on Facebook. Additionally, it is also possible to posit that it is easier to engage users on the basis of likes than it is on the basis of comments and shares. I can assert, with moderate confidence, that districts with lower levels of income and lower levels of educational attainment are more likely to engage with Facebook pages, candidates are able to engage their users despite increasing audience size, and candidates that post more often on Facebook are less likely to engage the users on their Facebook page.

The Significance of Digital Communication in 2013 – Examining the Digital Landscape

The Nielsen U.S. Consumer Usage Report

According to Nielsen, 212M of 278M people that access the web are active online. The percentage of households that only have access to broadcast TV has diminished, from 16% in 2003 to 9% in 2012[1]. Social media users have topped 160 million mark, while smartphone users lay at 85 million. A majority of time spent on the computer is dedicated to social networks and blogs – at 20.1% of time, while email dominates 7.1% of a users time[1]. On smartphones, users dedicate about 14.1% of time to text messaging, 10.2% of time to social networks, and 5.3% of their time on email and instant messaging, an aggregate of 29.6% of time *connecting with other people*[1].

When considering access from the perspective of the economic divide, it is clear that media is not out of reach to those in the lower income strata. Higher levels of income generally only correspond to access to more devices and platforms – economically strapped consumers generally consume more levels of media, despite a smaller range of devices to access it from[2].

Observing the Nielsen Global Trust in Advertising Report when asked *to what extent respondents trust various forms of advertising and brand messaging platforms*, *when looking for information about the products you want and need, to what extent are*

---

[1] Nielsen – State of the Media: U.S. Consumer Usage Report 2013

[2] Nielsen – The Economic Divide: How Consumer Behavior Differs Across the Economic Spectrum, 2012

the following advertising most relevant to you? 90% of respondents ascertained that recommendations from people they know are highly/somewhat relevant, followed by 75% responding that it was on the basis of consumer opinions posted online[1]. When asked *to what extent do you trust the following forms of advertising/recommendation,* the answers were very similar to the question posted above: 90% trust completely or trust somewhat recommendations from people they know, while 68% trust completely or trust somewhat consumer opinions posted online. Confidence diminished from ads on TV, magazines, billboards, newspapers, radio, TV programs and product placements, ads served in search engine results, online video ads, ads on social networks, and online banner ads, diminishing from 46% respectively down to 28%. Therefore, it is possible to infer that people derive their actions from evaluations of trust, and that greater levels of trust exist between nodes in close social circles rather than in those participating in loosely structured networks. Thus, individuals are more likely to act based off the recommendation by a peer, than by an acquaintance, than by a social outsider, and so on and so forth.

Non – Disruptive Advertising

The digital age has a fueled a new social, marketing, and political landscape. Access to political commentary is available while a deluge of facts, and rhetoric can be contested within the span of minutes. According to Google's Charles Scrase, modern political campaigns have "rel[ied] on disruptive messaging — catching people's attention while they're doing something else. What we think is so powerful about the web is that we're reaching people at moments of decision points, when they're thinking about an

[1] Nielsen – Global Trust in Advertising and Brand Messages, 2012

issue and wanting to learn more"[1]. According to the Pew Research Center, over a tenth of Americans "dual-screened" the 2012 Presidential debates, dual- screening defined by the act of either looking up facts related to the debate or sharing information related to the debate with friends and family over social networks[2]. Additionally, social media has allowed for a more nuanced form of political discourse available to a wider array of people.

It is apparent then, that perhaps the most beneficial aspect of social media in political campaigns is that user engagement has been made possible for the masses, and that more people are engaging with political content. Relevant studies have underscored the notion that the capacity to influence political efficacy is contingent upon degrees of personal connection. Understanding that political influence proliferates more effectively through the links between strong ties within a social network, the best means by which to further the discussion relevant to the most effective political campaign appeals is related to a campaign's success relative to how successfully they can engage specific voting groups and demographics.

Social Networks as Enablers

I contend that social media has impacted American politics by defining political credibility as acquired through social support and engagement rather than by mere name recognition and exposure. Recent literature on the impact of social media reveals that online advertisements leave at most fleeting impressions on voter behavior[3], while most social networks consist of weak-tie relationships, those connections that yield little

---

[1] Mashable – Google to Political Campaigns: Here's How You Win, 2012
[2] Pew Research Center – *One-in-Ten 'Dual Screened' the Presidential Debate*, 2012
[3] Broockman, Green – Do Online Advertisements Increase Political Candidates' Name Recognition or Favorability? 2013

influence between ties[1]. Other studies have concluded that the most effective voter mobilization efforts have been those that appeal to social pressure and the propensity for feeling a sense of shame in not fulfilling one's personal civic duty[2].

Reporting results from a randomized controlled trial of political mobilization messages delivered to 61 million Facebook users during the 2010 US Congressional elections, the results show that strong ties are instrumental in spreading both online and real world behavior in human social networks[3]. While most political campaigns have existed under the guise of passivity, mostly aiming to achieve name recognition, these studies underscore that a contingent factor influencing successful online political campaigns are those that engage voters most efficiently.

Large Scale Data Analytics and Sentiment Analysis

With this knowledge, it is apparent that societal sentiment can easily be sparked by the opinion of a single actor. Because individual nodes in separate networks rely on links between their nodes to develop social and political opinions, it is possible then to assume that aggregate actions will move at a faster pace then ever. It is important to understand the structure and development of online social networks before moving forward with this discussion. Observing the social composition of online social networks, it has shown that, contrary to hypothesis suggested by Wellman, Salaff, Dimitrova, Garton, Gulia, and Haythornthwaite in 1996-that the advent of social media will encourage users to align their social networks around shared interests rather than by

---

[1] Christakis, Fowler, *Connected, 2009*

[2] Gerber, Green, Larimer – Social Pressure and Voter Turnout: Evidence from a Large-Scale Experiment, 2008

[3] Bond et al. A 61-Million-Person Experiment in Social Influence and Political Mobilization

geographic proximity[1], social media has instead been used as a tool by which users translate offline networks into the online world.[2] Furthermore, studies of network structures reveal that networks with higher degrees of social reinforcement and overlaying clusters influence behavior to a farther extent and greater speed than corresponding random networks with little social reinforcement.[3]

Individuals are more connected than ever, and it is important to consider social media not by virtue of *who* can access it, but to what extent that others can see *what* other people are doing. This social and political transparency contributes to greater deals of social pressure, and with deliberate innovation in terms of innovation technology, it is more possible than ever to observe other users interactions on a voyeuristic basis. It is not uncommon for a friend of a friend to note that he/she has checked in at location y, that a friend is a fan of Stephen Colbert, or that he/she has attended a GOTV drive. In this context, the landscape of political action is changing.

Thesis Question

It is possible to understand that the greatest influences in social mobilization exist between the interactions shared by individuals in strong tie networks. Because social networks on the internet have been more representative of networks in the real world (rather than being aligned by interests), as noted above, the speed at which individuals within specific networks are alerted of what their friends are doing occur at a much faster rate.

---

1 Salaff et al. – Computer networks as social networks: Virtual community, computer-supported cooperative work and telework, 1996
2 Ellison , Steinfield, Lamp -The Benefits of Facebook "Friends:" Social Capital and College Student's Use of Online Social Network Sites, 2007
3 Centola – The Spread of Behavior in an Online Social Network Experiment, 2010

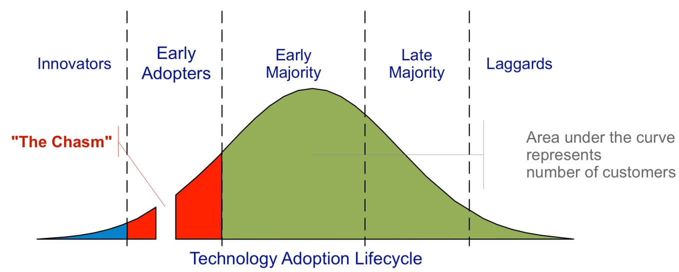
Because we are influenced by those we are close to, it is important to understand how Facebook, a tool that is becoming the increasingly preferred medium (if not already) of communication exists in relation to human behavior. We will discuss later the quantitative research methods that have successfully analyzed the average reach of a Facebook post, how strong ties influence others to political action, and how content is generate on a Facebook page. This research helps us form the basis for our study, including how well Congressional candidates are engaging their constituents on Facebook, what appeals - and what types of posts - receive the most engagement on Facebook, and during what periods of time do candidates yield the highest levels of engagement, if any. I also seek to discover a relationship between levels of user engagement on Facebook with respect to [1] frequency of posts [2] the number of "likes" associated with a page, and [3] district level demographics, including but not limited to constituent education levels and income.

By discovering the answers to these questions, it is possible to further understand how social interactions on Facebook influence action in the real world. It is important to evaluate the engagement level of a post because it is a greater indicator as to the extent of the reach of that interaction, as we will discuss later.

In reference to the technology adoption lifecycle, we have not yet overcome the "chasm" as to the integration of modern media in the political and social agendas of Congressional candidates. While an average district harbors 720,000 inhabitants, most candidate Facebook pages exhibit a reach that extends to 3,000 people or less, a figure that is less than half a percentage of the population (0.004). Even when we figure that on average, only 16% of eligible individuals in a district do choose to vote, (120,000), average demonstrated reach is only 2.5% of the population. Although we will see later that this 2.5% figure generally has the capacity to extend as far as 15% - 20%, this is still far from reaching the sought after early/late majority of technology adoption.

We understand that social media is becoming more integrated in to the American lifestyle, so it is advantageous to realize how we can leverage it to our advantage leading up to the 2014, 2016, and even 2018 elections.

[Figure 1 – The Technology Adoption Lifecycle]



Thesis Structure

Through the course of my thesis, I will elaborate upon the theories of classical political science that have motivated my research, the "science" of Facebook, including the technology that drives its content organization and relevant studies related to the field. I will follow these two fundamental levels of understanding and relate them to my research design and findings, and I will close with remarks regarding the relevance of my findings

Chapter 2 – Contemporary Political Science

Green and Gerber – Social Pressure and Political Participation

It is important to understand the role of social pressure when considering the primary motivations an individual takes in to account when choosing to vote. In the 2008 study *Social Pressure and Voter Turnout: Evidence from a Large-Scale Experiment,* Gerber and Green discovered that the reason why millions of citizens nonetheless vote "is that they are willing to pay the slight costs in time and effort to avoid the feelings of shame associated with not voting, or, conversely, to enjoy the satisfaction of voting."[1] Furthermore, when considering the Riker and Ordeshook model – that an individual's perceived benefit or reward from voting in a given election is determined by the probability of their vote mattering, multiplied by the benefit of voting (determined by individual candidate preference), in relation to the time/effort spent voting in addition to the psychological and civic benefit of voting (the sense of goodwill feeling by fulfilling citizen duty) – it is possible to assess that a primary motivation in a citizen's participation in the Democratic process is determined by how *others* perceive them.

$$R = pB - C + D$$

As Gerber and Green note, "voting is widely reared as a citizen duty, and citizens worry that others will think less of them if they fail to participate in elections." Social pressure is a greater motivation to vote. It is, as Gerber and Green note, "sometimes pointed out that people whose friends and coworkers voter at high rates are themselves more likely to vote."[1]

---

1 Gerber, Green - *Social Pressure and Voter Turnout: Evidence from a Large-Scale Experiment, 2008*

Berelson et al. Selective Media Consumption Theory

Observing other factors effecting political behavior, a wide literature suggests that the media's role is minimal. In 1954, Berelson, Lazarsfield, and McPhee found that sociological contexts are the main influence on voting decisions. They find that family is the main source of an individual's partisanship. Considering that most of an individual's decisions take place within a social context, they note that individuals (at least politically) tend to surround themselves with a homogeneous social circle of friends and coworkers because individuals tend to prefer other people that reinforce existing bias.[1] From the home, to the workplace, to any other social setting, individuals mutually reinforce their political views. From here, the mass media only works to reinforce existing bias because individuals generally choose to consume media that generally reaffirms their personal bias. Therefore, Berelson et al conclude that the media and campaigns have "supposedly minimal effects on voting behavior" as a consequence of selective media consumption.[1]

Atkin – Selective Exposure Principle

This view is reinforced by Atkin's 1973 study in *Quality Versus Quantity in Televised Political Ads.* Atkin notes that, during the 1970's, "a conventional wisdom concerning the effectiveness of political advertising"[2] subscribed to several main tenets:

- The brief spot ads reach a much larger proportion of the electorate than longer programs
- The greater the frequency of a candidate's ads, the greater the level of exposure and attention among voters
- Frequency of presentation is more important than quality of presentation

---

[1] Berelson, Lazarsfield, Mcphee – Voting: A Study of Opinion Formation in a Presidential Campaign, 1954

[2] Atkin, Bowen, Nayman, Sheinkopf, Quality Versus Quantity in Televised Political Ads, 1973

- The candidate's personality, image, and symbolic appeal take precedence over specific issue positions
- Getting the candidates name across is only a few steps removed from having his ballot lever pulled

At the conclusion of his study, Atkin discovered that the frequency of TV spots has a direct impact on exposure, but has little effects on the audience's attention levels. He underscored that different variables determine attention and information gain, ranging from messaging content to audience characteristics - audience characteristics often defined by partisanship. These factors together work to influence voting decisions or produce shifts in the strength of voting intentions. Atkin found, however, that individuals are more likely to engage with messages that support their attitudinal predispositions. Similar to Berelson's *Minimal Effects Hypothesis*, Atkin attributes the audience's attitudinal predisposition as a consequence of the *Selective Exposure Principle.*[1]

Broockman – The Minimal Effects of Digital Advertising

As the cost of running a campaign – and unseating an incumbent – have consistently increased, the advent of the Internet and social media encouraged political science to reconsider the consequence of political persuasion in the modern age. While it is understood that decreasing levels of intimacy correlate directly with decreasing levels of turnout (Gerber, Green – The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout)[2], it is valuable to examine the effects of online advertising on political persuasion, primarily because it is cheap relative to the aforementioned mediums. In a study conducted by David Broockman in 2012 assessing the effects of online

---

[1] Atkin, Bowen, Nayman, Sheinkopf, Quality Versus Quantity in Televised Political Ads, 1973

[2] Gerber, Green – The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment

advertisements on political persuasion, he discovered that "even frequent exposure to advertising messages may be insufficient to produce attitude change." Observing two groups, one frequently exposed to candidate ads on Facebook, and an alternative group not exposed, the former group was no more likely to recall remembering the candidate in question when questioned at a subsequent time.[1]

Fowler – The Role of Social Networks in Political Participation

In his article, *Turnout in a Small World*, James Fowler contends that

> *"Turnout is highly correlated between friends, family, and co-workers even when controlling for socioeconomic status and selection effects…if people choose whether or not to vote in part based on the turnout decision of their friends and acquaintances, then a single person may affect not only her acquaintances, but her acquaintances' acquaintances, her acquaintances acquaintances' acquaintances, and so on throughout the population… even a small conditional correlation between acquaintances can cause a chain reactions that leads to large aggregate changes in turnout."*[2]

He dubs this idea the *turnout cascade*. In a related, sociological study Easley describes that the methods by which new practices are adopted within a specific population depend largely upon the fact that people influence one another. Individuals have a fundamental inclination to behave relative to how others are behaving and are intrinsically motivated to conform to a crowd. This exists on the basis that, as individuals, we understand that the motivations behind our actions are driven by some mode of private information. When we see that others – or a large body of people, are acting in discordance with our actions, we believe that they have access to private information – that we have not yet been

[1] Broockman, Green – Do Online Advertisements Increase Political Candidates' Name Recognition or Favorability? Evidence from Randomized Field Experiments, 2013
[2] Fowler – Turnout in a Small World, 2005

privileged to understand, motivating them to act in a specific way. Therefore, we're motivated to do as they do, fearing that they are operating upon knowledge that we do not understand. And in this way, our strong ties affect the way we live, act, learn, and play.[1]

The Riker and Ordeshook Model

$$R = pB - C + D$$

R = the reward gained from voting in a given election (R, then, is a proxy for the probability that the voter will turn out)
p = probability of vote "mattering"
B = "utility" benefit of voting–differential benefit of one candidate winning over the other
C = costs of voting (time/effort spent)
D = citizen duty, goodwill feeling, psychological and civic benefit of voting (this term is not included in Downs's original model)

An overarching theme exists between the findings of the aforementioned scholars underscore that individuals are inclined to political participation by influences from their intimate social networks and that these social networks primarily reaffirm existing individual bias. The Riker and Ordeshook model demonstrates that the probability an individual will turn out to vote is primarily dependent upon the costs of voting and the psychological satisfaction of voting by virtue of fulfilling civic duty. Social media affects the model in two ways, first by reducing the cost of retrieving information, assuming that because individuals are already on Facebook most of the time, being able to access information they do not access on a regular basis on a medium that they do access on a regular basis drive down the costs of retrieving this information. Secondly, as I will more fully demonstrate in the chapter below, citizen interactions are more transparent on Facebook, so it amplifies the feeling of civic duty gained/lost from existing on a social

[1] Easley, Kleinberg – Networks, Crowds, and Markets, 2010

network, in that it is easier to share with people that one has voted, or it is easier for an individual to realize that they are a minority of individuals that have not voted. Because individual actions are more apparent, it is more possible to question our actions relevant to those of others – and to fulfill our perpetual need for group approval.

Chapter 3 – The Science of Facebook

Relevant Facebook Features for this Discussion – A quick synopsis of Facebook for non-Facebook users

News Feed

> News Feed—the center column of the home page— is a constantly updating list of stories from people and Pages that users follow on Facebook. News feed stories includes status updates, photos, videos, links, app activity and likes.
> This is the main page one access' Facebook from

Timeline

> An individual Facebook user's profile that highlights photos and posts created by the user as well as a user's interests, such as "likes" and links to other media that they choose to share.
> This is the primary means by which to display one's online persona, and a means by which other people can look at other people's profiles.

Ticker

> An overview of what is happening on Facebook including status updates, friendships, photos, videos, links, app activity, likes and comments as they occur in real time. This is seen on the top right bar of a user's Facebook page.
> This is seen on the top right bar of a user's Facebook page.

How Facebook Populates the News Feed – The EdgeRank Algorithm[1]

The news feed algorithm is known as EdgeRank, and decides which stories appear in each user's news feed. When a user's friend posts a status update, comments on another status update, tags a photo, joins a fan page, or RSVP's to an event, it generates a value that determines whether or not this interaction will appear on their friend's news feed.

---

[1] Information accessed from edgerank.net

[Figure 3 – The EdgeRank Algorithm]

$$\sum_{\text{edges } e} u_e w_e d_e$$

u_e ~ affinity score between viewing user and edge creator

w_e ~ weight for this edge type (create, connect, like, tag, ect.)

d_e ~ time decay factor based on how long ago the edge was created

***Affinity Score*[1]**

The affinity score observes the strength of each link between different nodes, factoring in 1) the strength of the action, 2) how strong the links between nodes exist, 3) how long ago the user participated in the action 4) how many mutual friends are shared between nodes.

The measure of an affinity score does not only measure an individual's actions, but also his/her friends' actions, and their friends' actions. If a user comments on a fan page, subsequent posts are more likely to appear on his/her news feed than if a friend decides to comment on that page, or if a friend's friend decides to.

To further elaborate, the EdgeRank value does not measure all friends' action equally. If a user, for instance, has a significant number of mutual friends and frequently comments on a specific person's status update and writes on their friends wall regularly, the affinity score between user A and user B rises significantly, meaning that content published by user B is more likely to appear on

---

[1] Information accessed from edgerank.net

user A’s news feed, rather than content posted by user C, whom user A infrequently interacts with.

The measure of the affinity score exists on a unilateral, rather than bilateral, basis. If user A frequently engages in content by user B, while user B hardly responds to the content or does not reciprocate the activity by similarly posting on user A’s wall, content posted by user A is no more likely to appear in user B’s news feed.

Edge Weight[1]

Every action a Facebook user engages in creates a value, and each of these values carries a different weight. A user is more likely to see another user’s action if he/she decides to comment on a page than like a page.

The strength of an action is a reflection of the effort required for the action – more effort from the user demonstrates more interest in the content. For instance, commenting on a post is worth more than liking it, which is worth more than clicking on it, which is worth more than passively scrolling by it (an action that carries no weight).

Time Decay [2]

The EdgeRank value is perpetually updated. When a user logs on to Facebook or refreshes his/her respective homepage, the newsfeed is populated by content that reflects high EdgeRank values at that moment in time.

[1] Information accessed from edgerank.net

Essentially, the time decay value of the EdgeRank formula is a measure of how much time has occurred since posting, or 1/x, x being time elapsed since initial posting. In other words, "as a story gets older, it loses points because it's 'old news.'"[1]

The Digital Footprint – Facebook Privacy and the Transparency of Our Digital Actions[2]

It is possible to see status updates and other interactions from users who are not directly connected with one another

If user A is friends with user B, and user B comments on user C's status update (user A and C are not friends), it is still possible for user A to observe this interaction by virtue of user A being friends with user B. However, the likelihood of user A observing this activity is dependent upon the affinity score between user A and B.

Privacy settings are not universal; they are specific to an individual user's specific timeline

A Facebook user is only in control of the privacy on his/her own timeline. While it is possible for a user to limit the audience on a specific post, ranging from public to custom, if (for example) user A posts to user B's wall, the proliferation of this interaction is contingent upon user B's privacy settings. While user A may exercise the highest levels of privacy on his own wall, if user B chooses to make his/content public, it is possible for these interactions to be observed by anyone. This is significant because if user A decides to comment on user B's status, any content regarding that interaction is

[1] Information accessed from edgerank.net

[2] Information accessed from Facebook Privacy Settings FAQ

visible to anyone who is friends with A or B, despite the fact that user A believes that he/she is exercising utmost control over the proliferation of his/actions.

Quantifying Audience Size – Bronson et al.

Observing 220,000 Facebook users post over the course of one month, the Stanford University HCI Group Computer Science Department in conjunction with the Facebook Data Science Group found that Facebook users manager to reach roughly 35% of their friends with each post, ultimately reaching 61% of their friends (with regards to exposure to that post) over the course of one month.[1] This helps us determine to what extent our Facebook interactions proliferate over a range over a specific number of user's friends.

Effects of Strong Ties on Social Networks – Bond et al.

During a randomized control trial of 61 million Facebook users during the 2010 Congressional elections, Robert Bond, Christopher Fariss, James Fowler et al. underscored the value of strong ties in spreading online and real world behavior in human social networks (Bond, Fariss et al., A 61-million-person experiment in social influence and political mobilization, 2012). Individuals with stronger ties (determined by the number of tagged photos one individual shared with another) had a greater propensity to influence his/her peers to action than those that did not constitute strong ties.[2,3]

---

[1] Bernstein et al – Quantifying the Invisible Audience in Social Networks, 2013

[2] Bond, Fariss et al., A 61-million-person experiment in social influence and political mobilization, 2012)

[3] Although "strong" ties were able to move their peers to action more so than "weak" ties, this only occurred on a 1-2% basis – a relatively small value. The group contends, however, that a 1-2% increase in political participation over the range of such a large sample (61 million), results in a large magnitude effect.

Chapter 3 Conclusion

Together, the technological behind Facebook's organization methods in conjunction with the quantitative experiments evaluating the network's effects on organizational behavior suggest that our social connectivity in the digital world relays important consequences relative to our actions in the real world.

Understanding that individuals generally have a need to conform, that the content we're exposed to on Facebook is generally sorted by those that we have a greater "affinity" with only further amplifies our social group organization bias – in other words, an individual's predispositions are only reaffirmed by Facebook's organizational methods. Individuals are less likely to be exposed to information that exists in contrary to their beliefs; by constantly being surrounded by individuals that hold the same beliefs that they do, individuals would in theory re-think his/her actions if those actions do not conform with those of the group.

Chapter 4 – Research Design

The Case for Research

Because an individual's attitudinal predispositions are mostly reinforced by their online social communities on Facebook, the strength of appeals posed by different candidates pose important implications for user engagement on Facebook. A candidate's capacity to engage his/her audience increases the probability that his/her message will proliferate over a larger group of people. The EdgeRank algorithm underscores that post popularity decays with time and that posts that rely mostly on time to appear on the news feed quickly disappear. However, posts that harness higher levels of engagement are more likely to appear on news feeds for longer periods of time. It is important, then, to understand what interactions on Facebook yield in the greatest levels of user interaction.

Purpose of Research

My research design is constructed with the intention of addressing three questions:

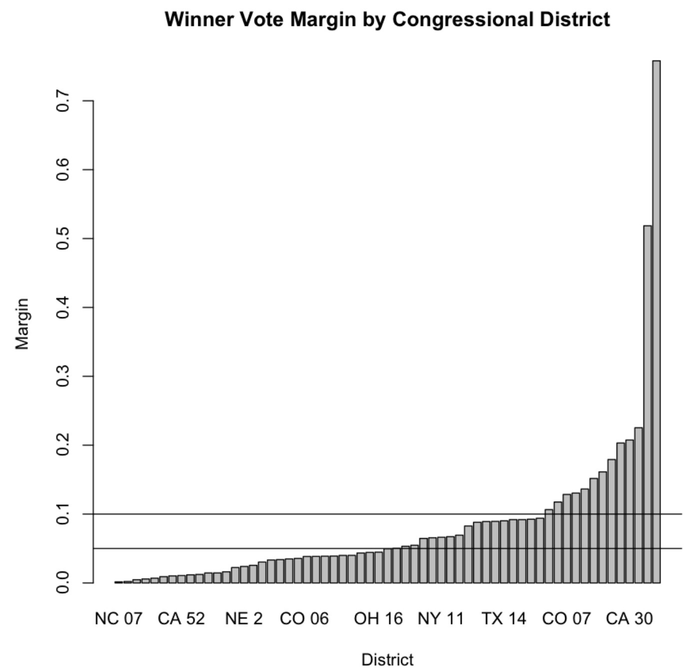
1. What specific types of posts and, more broadly, what types of appeals yield the greatest level of engagement on Facebook?
2. Is there a relationship between the time leading up to the election and user engagement?
3. Regarding external variables, what other factors affect user engagement on Facebook?

Design Overview - Selection of Variables

61 Congressional Districts were chosen on the basis of variations in levels of competition. Of the 61 districts, 32 districts had margins of victory below 5%, 16 districts

had margins of victory below 10%, while the remaining 10 districts exceed 10% but resulted in elections during which victorious candidates spent less than their opponents.

[Figure 4.1 Winner Vote Margin by Congressional District]



District demographics resulted in a standard distribution of income and educational attainment (measured by % of households that have achieved bachelor's degrees).

Design Overview – Data Collection on Facebook

Of the 61 Congressional Districts chosen, I accessed the Facebook pages of the top 2 candidates running in every district for a total of 122 Facebook pages. Upon accessing these pages, 10 candidates did not have Facebook pages or had deleted them following the election, all of which lost in 2012. All winners maintained a social media presence.

Of the pages observed, I logged every post occurring between October 1^{st} to November 7^{th} – November 10^{th}, the date depending upon when the candidate would write a statement either accepting congratulations for a victory or conceding a loss. Upon access, I logged [1] the total number of fans that had "liked" the page [2] the primary audience the page had interacted with, retrieved from Facebook insights, and [3] the post type, determined by factors I will elaborate below, and number of likes, comments, and shares per post. At the conclusion of the data retrieval process, I manually logged 3,553 posts over a range of 112 Facebook pages.

Design Error

Upon reviewing my data, I believed the addition of a time series analysis would be beneficial to my study. I developed a code system by which I logged October 1^{st} with the value of 1, October 2^{nd} as 2, and so on and so forth. Upon reaching posts logged in November, I would categorize November 1^{st} as n+1, or 32, until reaching posts from November 7^{th} – November 10^{th}.

As I re-accessed these pages, I observed that during the original logging process, I had systematically skipped certain Facebook posts during access. Therefore, during my original log of 3,553 posts, I only collected about 40% of the posts posted by candidates. I proceeded to log data for a second time, re-logging 3,440 posts across a measure of 53 Facebook pages, instead of 3,553 posts over the course of 112 Facebook pages.

In deciding which data set to utilize over the course of my experimentation, I sought to discover whether or not the sample dataset derived from the same distribution. I proceeded to run the Kolmogorov-Smirnov test to determine whether or not this was the case. I compared the engagement value of likes from the first set, retrieved from 122 pages that omitted posts, with the engagement value of likes from the second, more accurate set observing 53 pages. I retrieved a p-value of 0.00818, revealing that the two samples came from different distributions.

Therefore, I chose to use my first dataset to observe post engagement on an aggregate level because of the greater differentiation in sample size observed. Because I was looking at engagement values of posts across 112 districts, observing the median value over the max value, it would not skew the results of the data (as opposed to observing just 53 pages).

For the next two experiments to be conducted, I utilized the second data set that consisted of a more accurate measure of candidate-user. I utilized this dataset to discover the relationship between post frequency and post engagement. In the following experiment, I evaluated candidates on what I dubbed to be their "engagement efficacy," a measure of their average post interaction with their max post interaction. I will utilize this dataset to discover what factors contribute to greater levels of engagement efficacy, ranging from post frequency, candidate incumbency, and other district level variables (income, educational attainment).

Research Design 1: User Engagement and Modes of Appeals

Purpose

The purpose of my first research design is to discover what individual type of Facebook post yields the greatest user engagement on Facebook, ranging from likes, comments, and shares. From this set of 15 factor variables, I grouped them in broad categories defined by appeals to candidate intimacy, civic duty, national politics, local politics, and the general election, to examine specifically what mode of content yields the greatest engagement on Facebook.

Methodology

For the purpose of this design, I compiled a list of 10 categorical variables and crossed referenced these variables with 5 different college students asking each individual: *do you believe that this is an exhaustive list of post types a Congressional candidate would post on Facebook?* I surveyed 3 males and 2 females of which 3 were Political Science Majors, 1 a Computer Science Major, and the last being a Nano Engineer. After reclassifying, modifying and adding different factor categories, I came up with the 15 different variables. Inputting information in to my first data set, I included the factor category type, likes, comments, and shares on a per post basis from October 1st – November 7th-10th for 112 different Congressional candidate pages on Facebook.

From the 15 factor variables I had developed, I grouped them on the basis of type of appeal for the second part of my first research design. I grouped the 15 categorical variables in to 5 different subgroups emphasizing different appeals, the first being an appeal to a candidate's personal life followed by appeals to civic duty, the national election, the local election, or general information about the election.

Factor Variables

Text Posts

- Competitive, positive tone
    - "The race is very close! We can do it!"
    - "There are only 5 days left to register. Our campaign depends on your support. Please register now!"
- Competitive, negative tone
    - "Candidate X has no integrity and will push our country farther and farther in debt and is the absolute last leader our country needs"
    - "I'm appalled by the slander my opponent has published about me. This is not true"
- General information about the campaign
    - "I will be at Rubio's tonight to speak about my campaign"
    - Facebook events regarding campaign events
- Information about candidate receiving an endorsement
    - "I'm grateful to receive the endorsement of the NRA"
- Information linking to external media
    - Links to newspaper articles, candidate website, or other media related to the candidate/events
- Patriotic – National
    - "My heart goes out to the victims of Hurricane Sandy"
    - "National spending is out of control, we need to stand together and rise up against President Barack Obama"
    - "Congress is performing at an all time low, we must
- Patriotic – Local
    - "Congratulations to Boy Scout Troop 658 for achieving a gold star"
    - "Thank you to the local Lion's Club for hosting me at your community dinner tonight"
- Personal Life
    - "Happy birthday to my dearest husband"
    - "From the bottom of my heart, I thank you for voting for me"
- Call to action
    - "Like this post if you agree with my statement!"
    - "Joe Biden and Paul Ryan put on a great show at tonight's debate. I believe that Biden came out victorious. What do you think?"

Photo

- Local event
    - Photos of campaign canvassing, sign waving, local appearances
- Personal life
    - Photo of candidate with family/other intimate setting

- Call to action
    - "Share this photo if you voted for Candidate A"
- Popular media
    - Media posted by external sources uploaded/shared by candidate

Video

- General campaign advertisement
- Public service announcement

Post Appeal Types

- Candidate Intimacy
    - Photo, personal life
    - Text, personal life
    - Video, general campaign advertisement
- Civic Duty
    - Photo, call to action
    - Text, call to action
    - Text, competitive positive
    - Text, competitive negative
- National Politics
    - Photo, popular media
    - Text, patriotic national
- Local Politics
    - Photo, local media
    - Text, patriotic local
- General Campaign Information
    - Text, link to other web
    - Text, general info
    - Text, endorsement
    - Video, public service announcement

Calculating Post Engagement Values

I logged the number of times each post occurred, and from this value, I calculated the median number of likes, comments, and shares each type of post each. I utilized the median instead of the mean because of the great deals of variance that occurred amongst each page (a specific candidate could potentially log 6,000 likes upon a single post, substantially higher than the max value for most other Facebook pages).

With this information, I determined to what extent each type of post yielded how many user interactions on the basis of likes, comments, and shares. For the second part of my experiment, I logged the max engagement value each post achieved, calculated by the highest number of likes, comments, or shares each post achieved over the course of the 40 days.

I proceeded to calculate the median/max of each post type to discover what engagement efficacy each post generally yielded. I dubbed this term engagement efficacy because the max count reveals how many people are willing to interact with the page, while the median reveals the average engagement. By dividing the two values, I was able to examine which type of post yielded the greatest engagement over potential engagement.

I calculated the values for each 15 variables, and proceeded to take the average of these values across the groups I had established earlier to reveal what appeal yields the greatest engagement efficacy.

[Figure 4.2 Engagement Value Formula]

Engagement Value = Average Post Interaction / Potential Post Interaction

**Average post interaction defined by median, potential post interaction defined by max*

Research Design 2 - Time Series Analysis, User Engagement Up to and After the Election

Purpose

I seek to discover whether or not late campaign advertising yields greater engagement values and more exposure leading up to and on the day of the election on Facebook.

Methodology

For the second design, I replicated the process I practiced for my first research design, this time logging the date of each post. This data set was more complete in that it recorded the number of posts and date of each post, while it encompasses a smaller range of candidates.

I assembled the data to illustrate the frequency of posts up to and leading to the election utilizing bar plots. In succeeding graphs, I observed the median value of likes, comments, and shares per day from October 1st until November 10th to illustrate candidate engagement leading up to and after the election.

Research Design 3 – Candidate Engagement Efficacy & Relationship with External Factors

Purpose

My third research design has two intentions. The first is to develop a design that is a more accurate measure of calculating user interactions and engagement efficacy on Facebook. The second is to discover what factors contribute to various levels of engagement efficacy on the basis of three sets of attributes: candidate social media use, electoral outcomes, and district level attributes.

Calculating Engagement Efficacy.

Common measures of post engagement on Facebook include a measure of the sum of likes, comments, and shares a post receives on a given day divided by the total number of fans on a given day.

[Figure 4.3 Common Post Engagement Formula]

Post Engagement = (Likes + Comments + Shares) *on day* / Total fans *on day*

This model is flawed for two reasons. The first is that it gives the same weight to likes, comments, and shares. This is erroneous in that, in accordance with EdgeRank measures, shares are more valuable than comments than are likes because of the amount of effort it takes to engage in each activity. It is inaccurate to weight likes and shares because a share will transplant information from news feed A to news feed B, exposing the post to an entirely new user base. Alternatively, as mentioned above, if a user "likes" a post, there are many variables contributing to whether or not another user will see that like and be exposed to the original content.

The second reason this method is flawed is because observing the number of fans on a specific day also takes in to account passive users on a specific day. This is a problem because simply scrolling down a page is the lowest level of engagement a user can engage with on Facebook.

For the purpose of my experiment, I utilized a modified formula, which I believe to be a more accurate measure of user engagement. Similar to the methodology I posed above, I divided the average value of likes, comments, and shares by the respective potential value for each type of post. I went further by taking in to account the weight of each variable to account for their value on our basis of evaluation, value being measured by whether or not user engagement will lead to exposure by that user's friends.

[Figure 4.4 Formula for Engagement Efficacy]

Engagement Efficacy =

(((Average Likes / Potential Likes) / 20) +

((Average Comments / Potential Comments) * 4) +

((Average Shares / Potential Shares) * 6)))

For 53 candidates, I calculated the median likes, comments, and shares, respectively by the corresponding "maximum" for each value. The resultant values illustrated that, on average, candidates were able to achieve a proportion of 0.517 engagement related to their likes – or more shortly, if a candidate demonstrated that he/she was capable of receiving 200 likes, he/she on average achieved 100 likes per post. For comments, I recorded a value of 0.037, and for shares, I recorded a value of 0.031.

The engagement value for likes demonstrated a great deal of variance in comparison to the values recorded for comments and shares, so I arbitrarily divided the value by 20 to put the value on parity closer to comments and shares to avoid skews.

[Figure 4.5 Unmodified Engagement Efficacy Value]

$$0.025 + 0.037 + 0.031 = 0.093$$

This figure reveals that on a median basis, candidates achieve 9% engagement on the potential engagement for each post. However, similar to the problems associated with the earlier problem, it omits the weight of each interaction. I proceeded to multiply each value by how much more “valuable” each post was, paralleling this modification with that exercised by EdgeRank, to achieve a more representative measure of engagement efficacy.

[Figure 4.6 Weighted Engagement Efficacy Value]

Likes + Comments (4) + Shares (6) = Engagement Efficacy Value

0.025 + 0.148 + 0.186 = 0.359

As the second formula reveals, on average, candidates were able to achieve 36 engagement efficacy per post. This is a more accurate measure of their engagement, taking in to account the significance of each interaction. I applied this methodology for the 53 candidates observed to evaluate them on the basis of their capacity to engage visitors on their Facebook page.

With this information, I regressed this value with district level attributes, election outcomes, and candidate social media use attributes to illustrate what factors influence engagement efficacy and what factors are effected by different levels of engagement efficacy.

Evaluating Statistical Significance

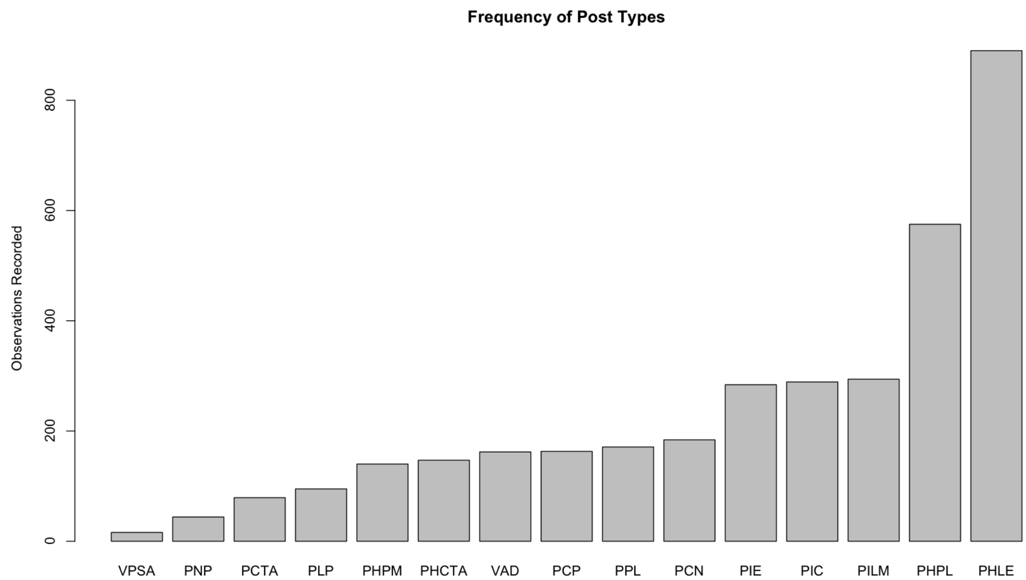
It is easy to assume that strong correlation values may imply relationships between variables, but as statisticians understand, correlation does not always imply causality. I did take this in to account over the course of my research, and while it may be interesting to analyze the relationships between the different variables in my dataset, they serve, at best, observational purposes. While I plotted relationships between variables, I ran three tests to consider a quantitative measure of statistical significance.

The first test I ran observed the correlation between variables, going further than simply looking at graphs. I tested these correlations for significance using the Spearman method, considering that large variances in engagement efficacy did occur, and that my values for engagement efficacy did not come for a normal population. I also ran robust analysis of variance for non-normal data, utilizing the Kruskal Wallis test. This test enabled me to observe whether or not two variables existed independent of one another. For each test (Spearman and Kruskal), p-values < 0.05 denoted statistical significance, something I have also noted in the tables below.

Chapter 5 – Results

Research Design 1 Part A: User Engagement on a Per Post Basis

[Figure 5.1 Frequency of Post Types]

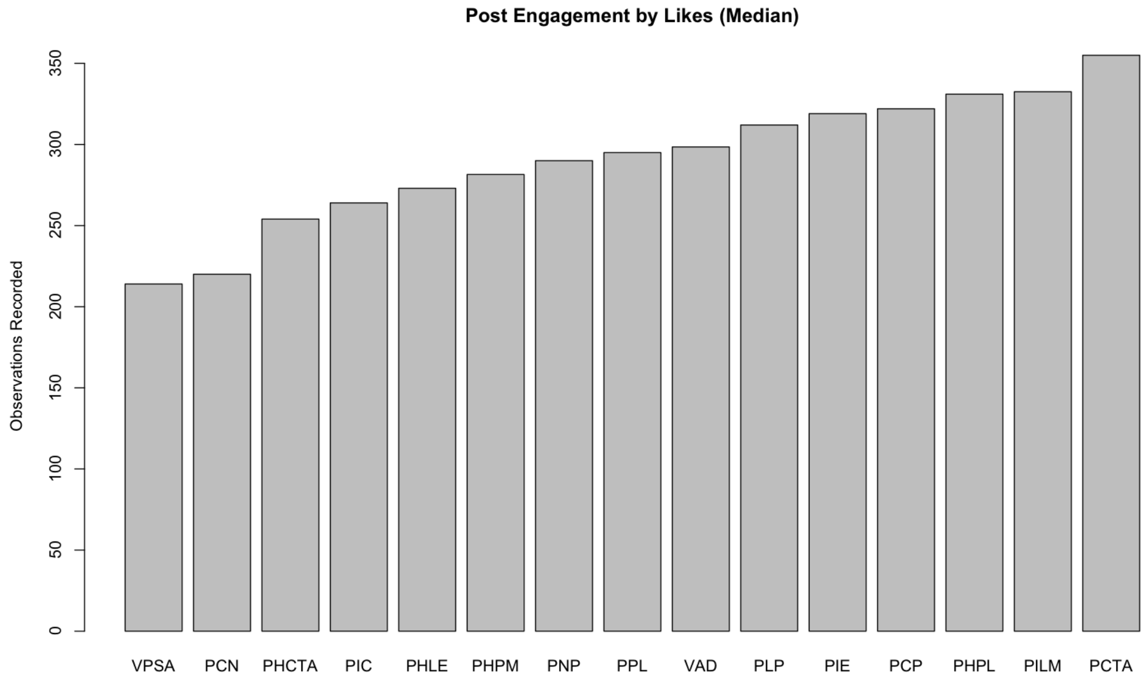


[Table 5.1 Frequency of Post Types]

Post Type	Frequency
Photo posted by candidates of local events	890
Photo posted by candidate of their personal life	575
Informational texts posts linking to external media	294
Informational text posts of information related to the campaign	289
Informational text posts about an endorsement	284
Text posts w/competitive message with w/negative tone	184
Text post about the candidate's personal life	171
Text posts w/competitive message with w/positive tone	163
Video of a general campaign ad	162
Photo post call to action	147
Photo containing popular media	140
Text post related to local politics	95
Text post call to action	79
Text post related to national politics	44
Video of public service announcement	16
Total	**3533**

Upon conducting the Shapiro-Wilk normality test to observe whether or not the sample size is normally distributed, it failed to achieve the p value necessary to indicate that the sample is normally distributed - achieving a p-value of 0.0004853. For the purpose of this design, it is important to note that we are dealing with a non-parametric sample in that we are not necessarily looking to achieve high levels of external validity, but rather, are trying to observe trends over a specific sample size to discover patterns within our distribution of mined data. Therefore, it is appropriate to observe the trends I will illustrate in the following tables and figures.

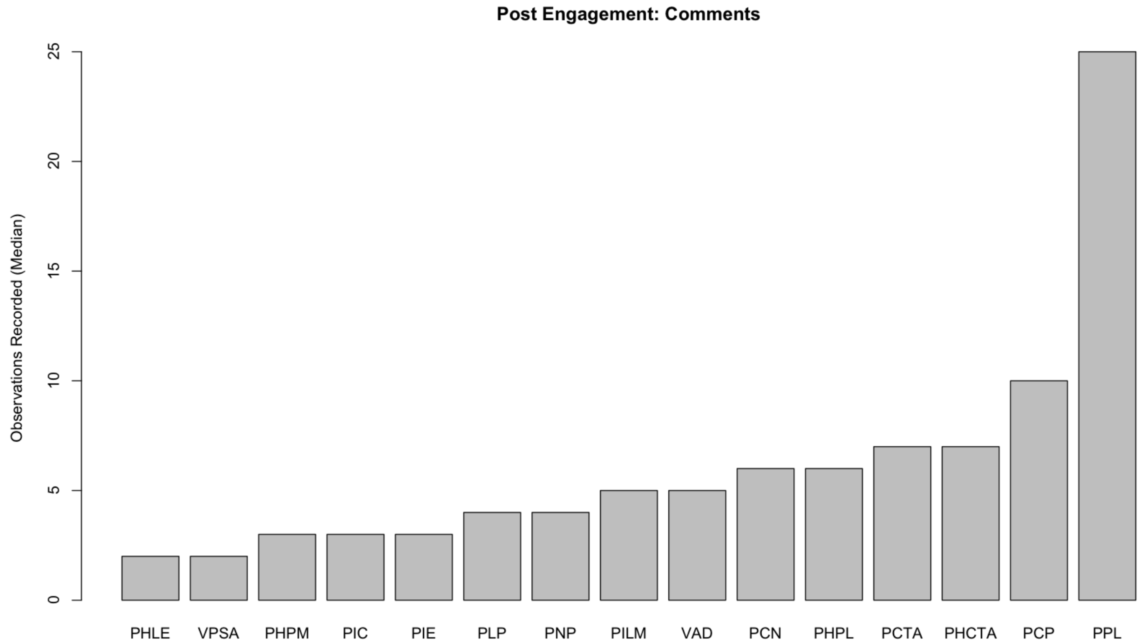
[Figure 5.2 Average Post Engagement by "Likes"]



[Table 5.2 Average Post Engagement by "Likes"]

Post Type	Median Engagement
Text post call to action	355
Text post of information linking to external media	333
Photo of the candidate's personal life	331
Text post, competitive, positive tone	322
Text post declaring candidate endorsement	319
Text post related to local politics	312
Video of a general campaign ad	299
Text posts of a candidate's personal life	295
Text posts related to national politics	290
Photo of popular media	282
Photo of a local event	273
Text post of information about the campaign	264
Photo call to action	254
Text post, competitive, negative tone	220
Video of public service announcement	214

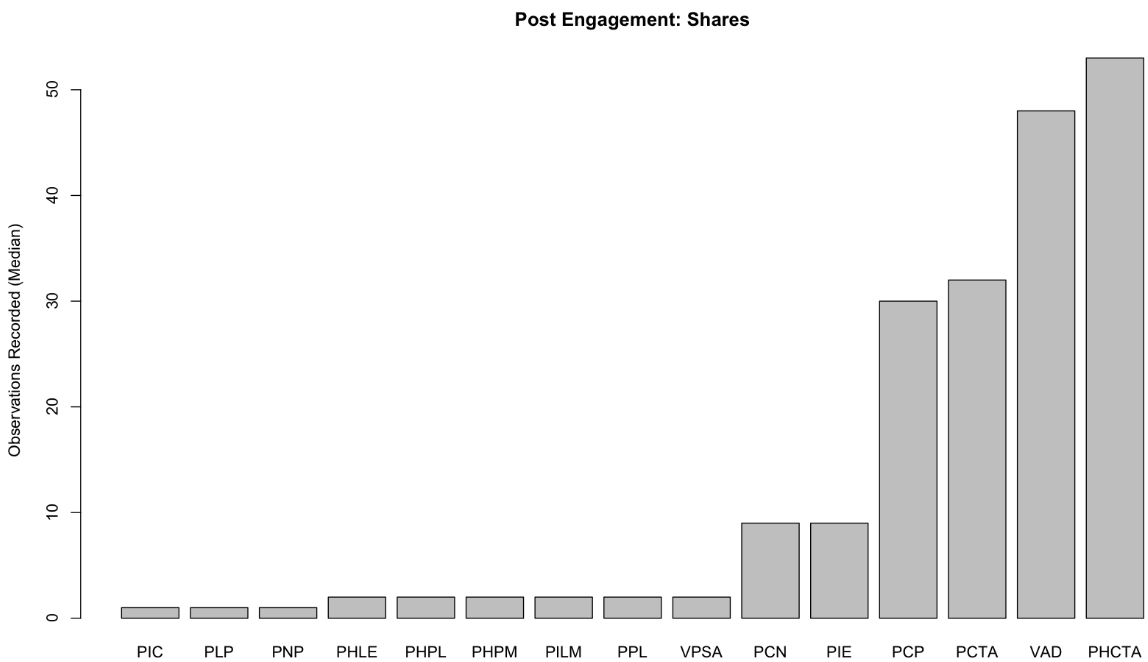
[Figure 5.3 Average Post Engagement by Comments]



[Table 5.3 Average Post Engagement by Comments]

Post Type	Media Engagement
Text post about the candidate's personal life	25
Text post, competitive, positive tone	10
Photo, call to action	7
Text post, call to action	7
Photo, candidate's personal life	6
Text post, competitive, negative tone	6
Video, general campaign ad	5
Text post, link to external media	5
Text post, national campaign	4
Text post, local campaign	4
Text post, declaring endorsement	3
Text post, general information about campaign	3
Photo, popular media (national campaign)	3
Video, public service announcement	2
Photo, local event	2

[Figure 5.4 Average Post Engagement by Shares]



[Table 5.4 Average Post Engagement by Shares]

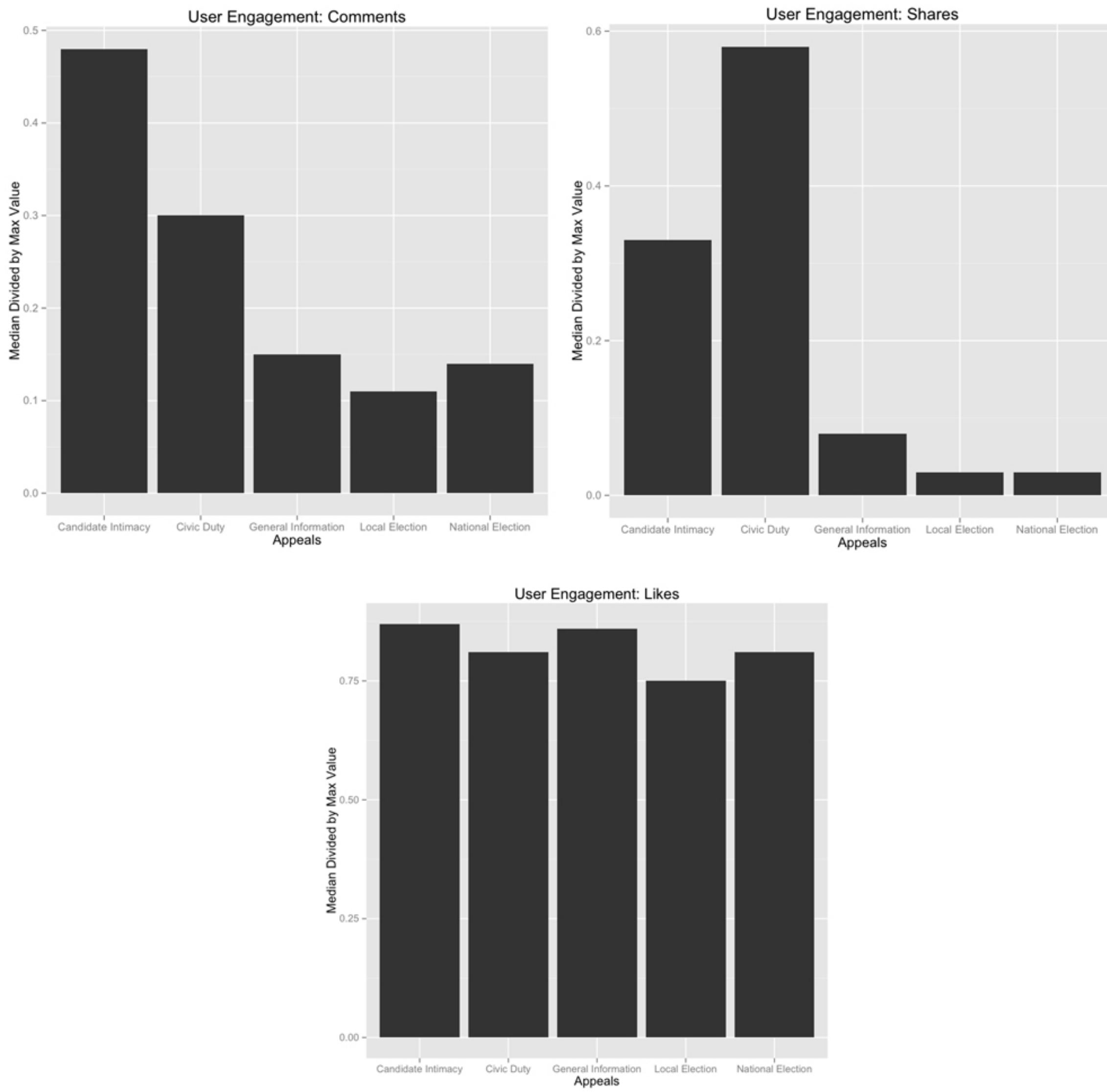
Post Type	Media Engagement
Photo call to action	53
Video of a general campaign ad	48
Text post call to action	32
Text post, competitive, positive tone	30
Text post declaring candidate endorsement	9
Text post, competitive, negative tone	9
Video public service announcement	2
Text post of candidate's personal life	2
Text post of information linking to external media	2
Photo of popular media	2
Photo of candidate's personal life	2
Photo of a local event	2
Text post regarding national politics	1
Text post regarding local politics	1
Text post of information related to the campaign	1

From the data above, it is apparent that post frequency does not necessarily lead to greater levels of user engagement, even though there are more opportunities to do so. Furthermore, it is generally easier for a candidate to achieve engagement by likes that it is by comments or shares. This demonstrates that the principal components to analyze relevant to the study of engagement on Facebook involved how to achieve greater engagement by having users comment on posts or sharing specific posts.

It is interesting to note that posts with competitive appeals generally received more engagement when they were written with a positive, rather than a negative tone. I will elaborate upon overarching trends in engagement in Part B of this experiment.

Research Design 1 Part B – User Engagement by Post Content and Appeal

[Figure 5.5 Engagement Value by Appeal – Comments, Shares, Likes]



The principal components of analysis are readily evident by observing the graphs in Figure 4.4. Appeals to candidate intimacy and civic duty net higher values of user engagement for comments than do the other three appeals. For shares, appeals to civic duty far surpass the other modes of appeals associated with shares for Facebook posts, while appeals to candidate intimacy also do fairly well.

This information helps us understand the main reasons users access candidate Facebook pages. We can infer that they do so to receive general information about the campaign, or to simply learn more about the candidate in question. While they may do so for all modes of appeals, for the sake of our study, we are concerned with whether or not a user's Facebook interactions will be seen by others. Therefore, we can assume that likes are of lesser statistical significance than are comments and shares, and we will focus on comments and shares for now.

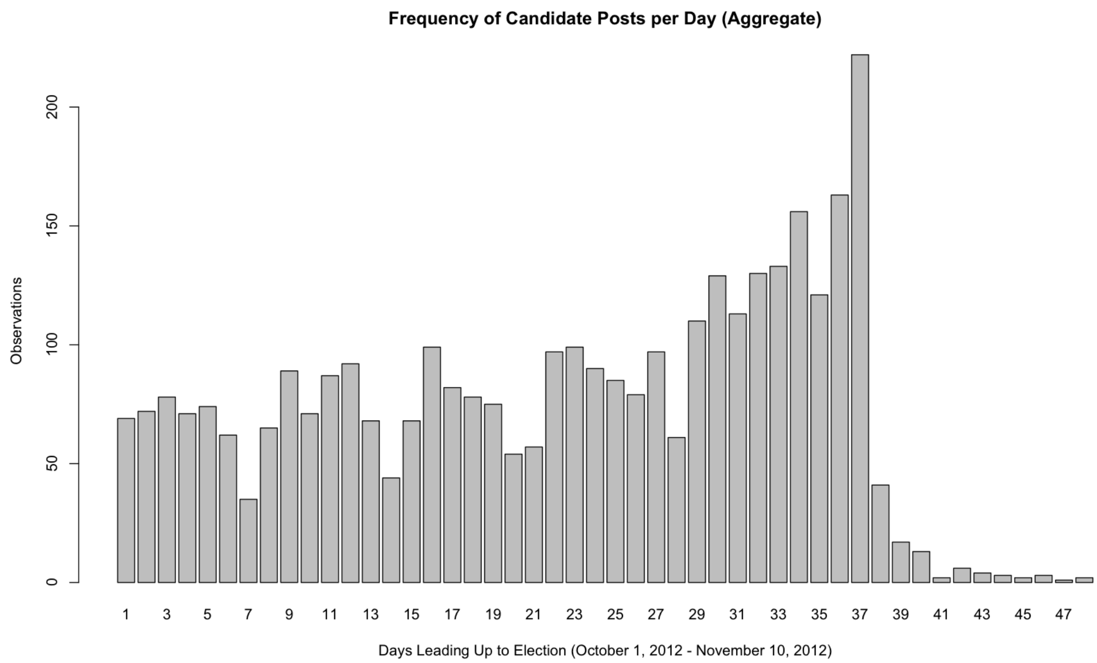
Perhaps engagement efficacies by comments are more important that those achieved by shares by virtue of polarization on online social networks. We understand that most people befriend people that are like them, and that online networks are reflective of physical, real world networks. Thus, sharing a post will most likely expose it to people who are similar to us, and this is successful in moving a specific group to vote, as noted by James Fowler in his work *Turnout In A Small World*.

Referring to both the Riker and Ordeshook model and studies conducted by Gerber and Green, appeals to civic duty are successful in motivating political action both in terms of psychological satisfaction as well as being able to avoid shame by other people.

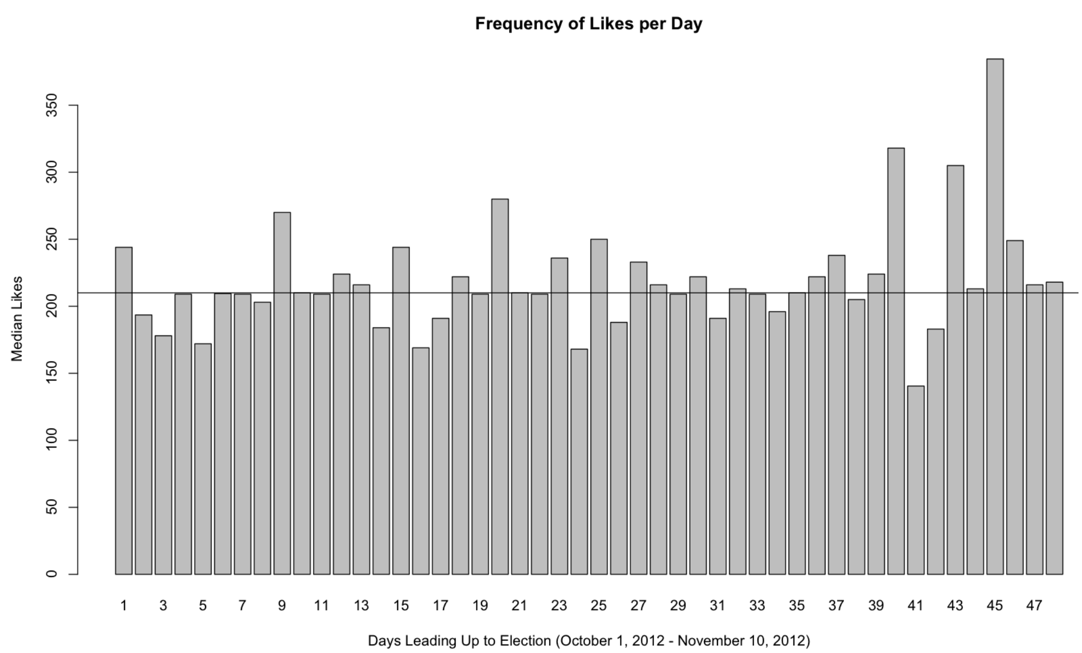
When individual users "share" pictures underscoring the importance of voting in a campaign, it has the capacity to result in a turnout cascade.

Research Design 2 - Time Series Analysis, User Engagement Up to and After the Election

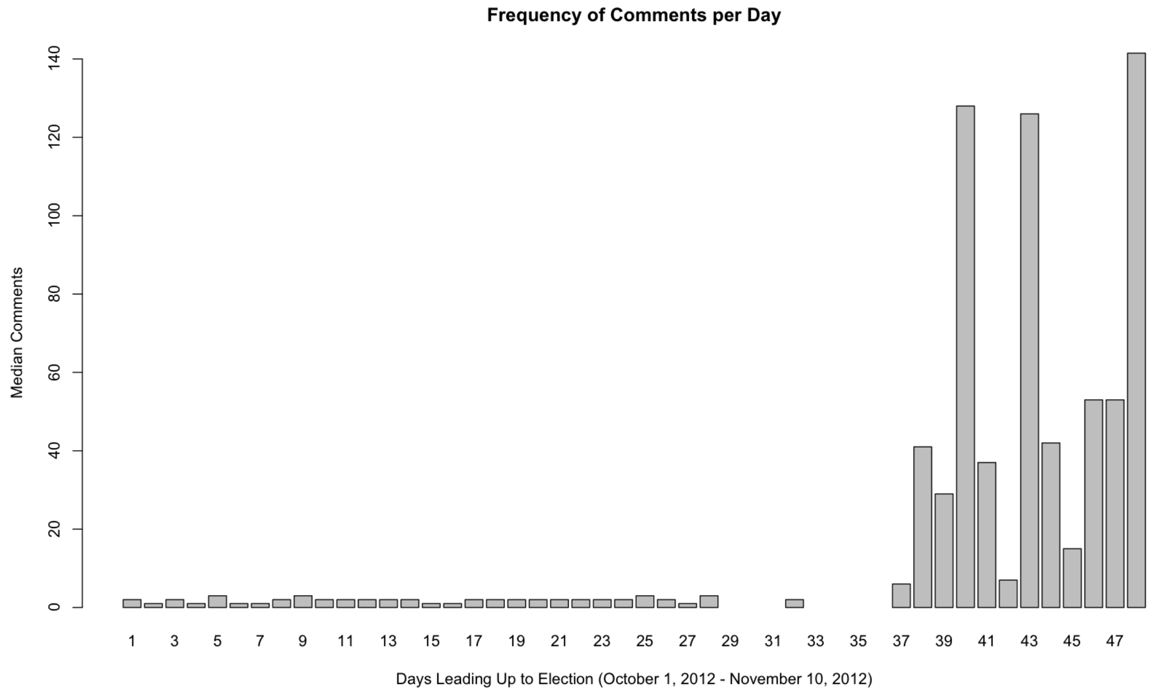
[Figure 5.6 Frequency of Candidate Posts per Day Leading Up to the Election]



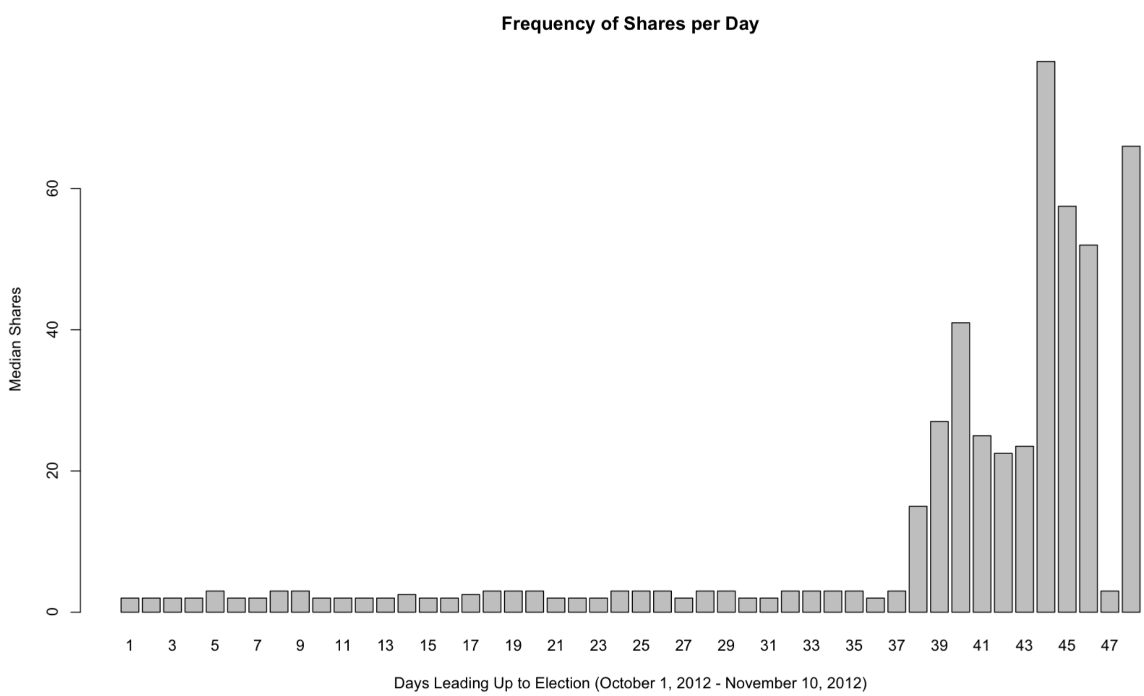
[Figure 5.7 Average User Likes per Day Leading Up to the Election]



[Figure 5.8 Average User Comments per Day Leading Up to the Election]



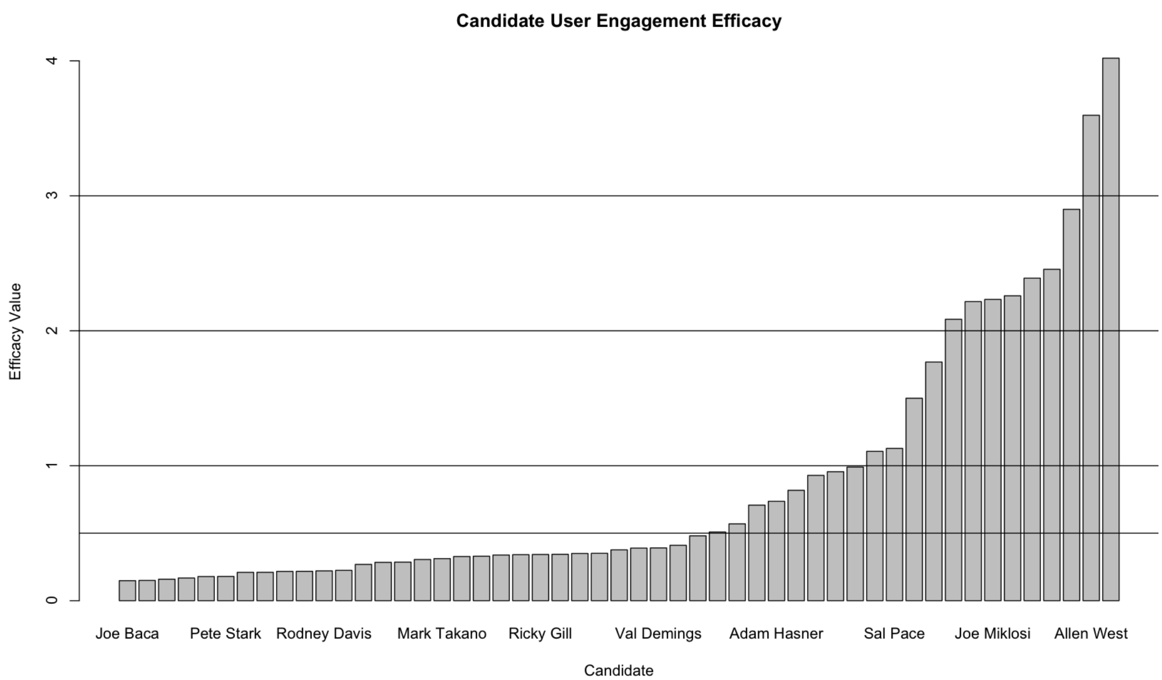
[Figure 5.9 Average User Shares per Day Leading Up to the Election]



An apparent trend exists in that candidates tended to post with greater frequency leading up to the election. The results reveal 2 observations. The first suggests that engagement does not improve with frequency (we will further elaborate upon this with our bivariate analysis in the next design). The second is that candidates do have the potential to reach audiences on Facebook, but they are doing so during times that are not most beneficial to the campaign. Most page interactions occur *after* the election has occurred. This suggests that a generous proportion of individuals do frequent a page; however, they are not being encouraged strongly enough to engage with it.

Research Design 3 – Candidate Engagement Efficacy & Relationship with External Factors

[Figure 5.10 Candidate User Engagement Efficacy]



We can observe that, for the most part, a majority of candidates do not really known how to engage their fan base on Facebook. This reinforces the notion that Facebook has not yet crossed the "chasm" of the technology adoption life cycle. Despite greater social media usage, political participation via social mediums has not proliferated as quickly as social media in general.

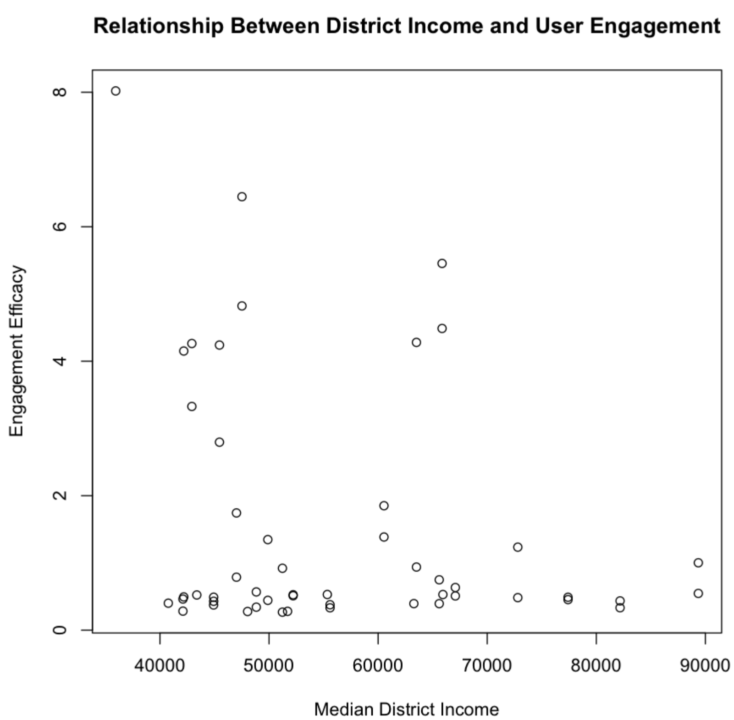
Our sample size is relatively small, so it will not be of much value to create predictive models on an individual candidate basis. Instead, I've regressed several variables with candidate engagement efficacy, running correlations, correlation significance tests, linear regression models, and Kruskal-Wallis tests instead of the typical analysis of variance because of the non-normal distribution of my values. It is important to note that correlation does not necessarily imply causality, but at least the results of our experiments will help us evaluate specific trends that currently occur with respect to social media.

[Table 5.5 Relationships Between Engagement Efficacy & External Variables]

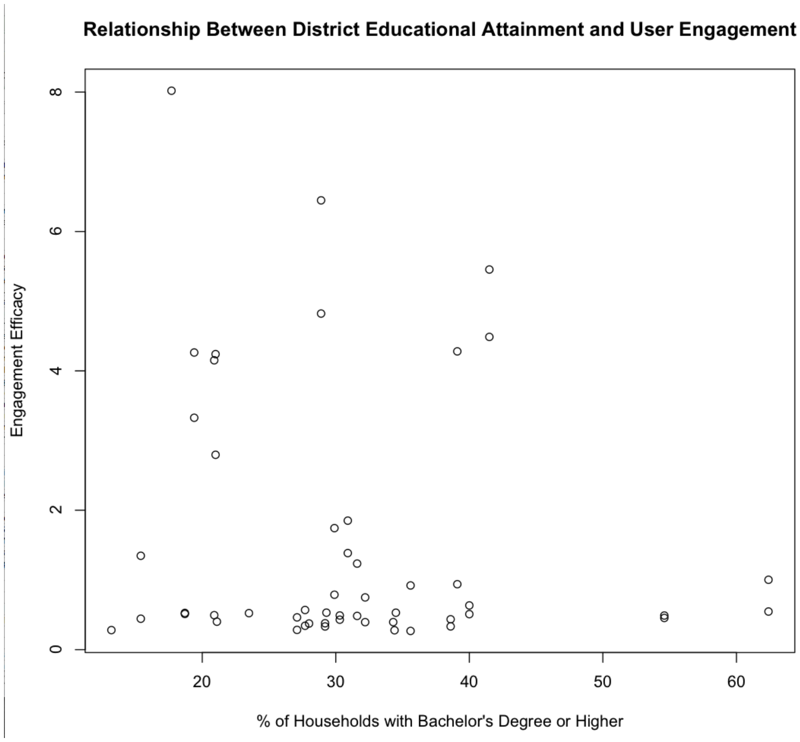
Relationship between engagement and:	Correlation	Spearman	Kruskal Wallis
District Income	-0.2489934	0.6309	0.0661
Educational Attainment	-0.1608589	0.812	0.0661
Candidate Likes	0.5000205	0.06121	**0.4734**
Post Frequency	-0.1417445	0.0614	**0.03783**

** For Spearman correlation significance test, $p < 0.05$ suggests statistical significance*
*** For Kruskal Wallis Robust ANOVA, $p < 0.05$ suggests differentiation between variables*

[Figure 5.11 Relationship Between Engagement and District Income]

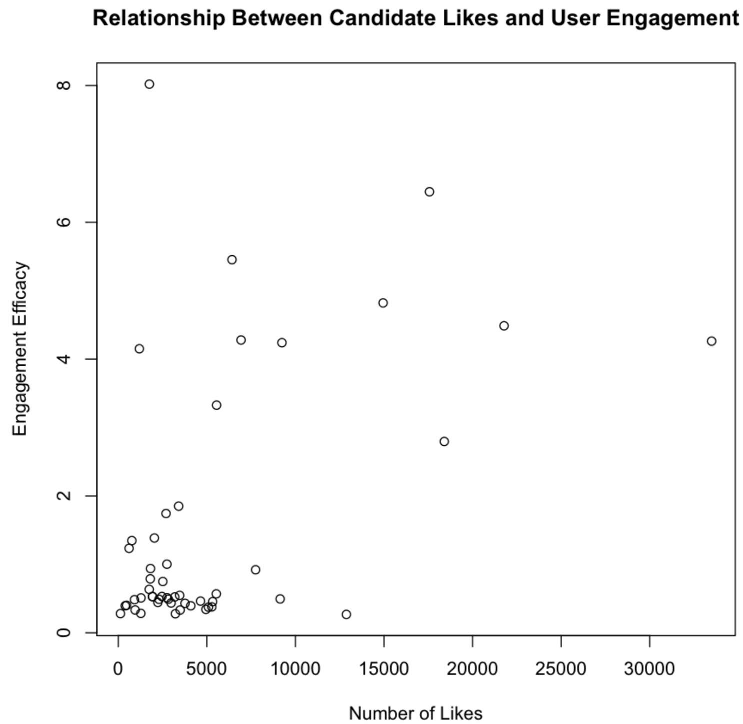


[Figure 5.12 Relationship Between Engagement and District Educational Attainment]

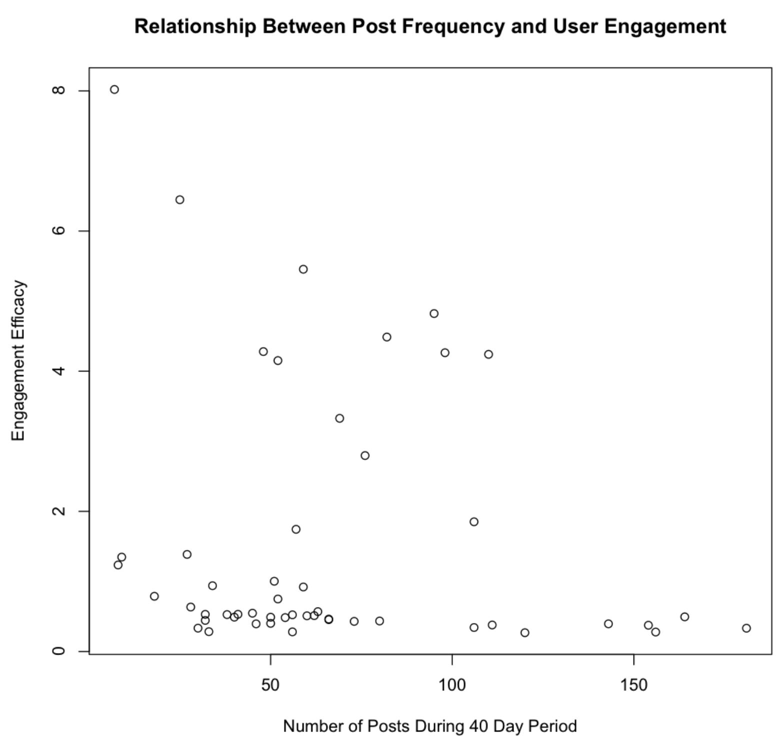


The most striking aspects of the data observed in Figure 4.9 and Figure 4.10 are the clusters that occur as the values on the X-Axis increase. Although the tests for the significance of correlation and independence of variables do not necessarily meet the required threshold, an observable pattern does occur. It seems that districts with lower incomes and districts with lower levels of educational attainment generally participate more with candidate Facebook campaigns. Correlation does not imply causation, however, but it is an interesting trend. Nielsen reported in 2012 that households with lower incomes generally consume more social media, so perhaps this is true too for their social media usage in accessing political content.

[Figure 5.13 Relationship Between Engagement and Number of Candidate Likes]



[Figure 5.14 Relationship Between Engagement and Number of Candidate Likes]

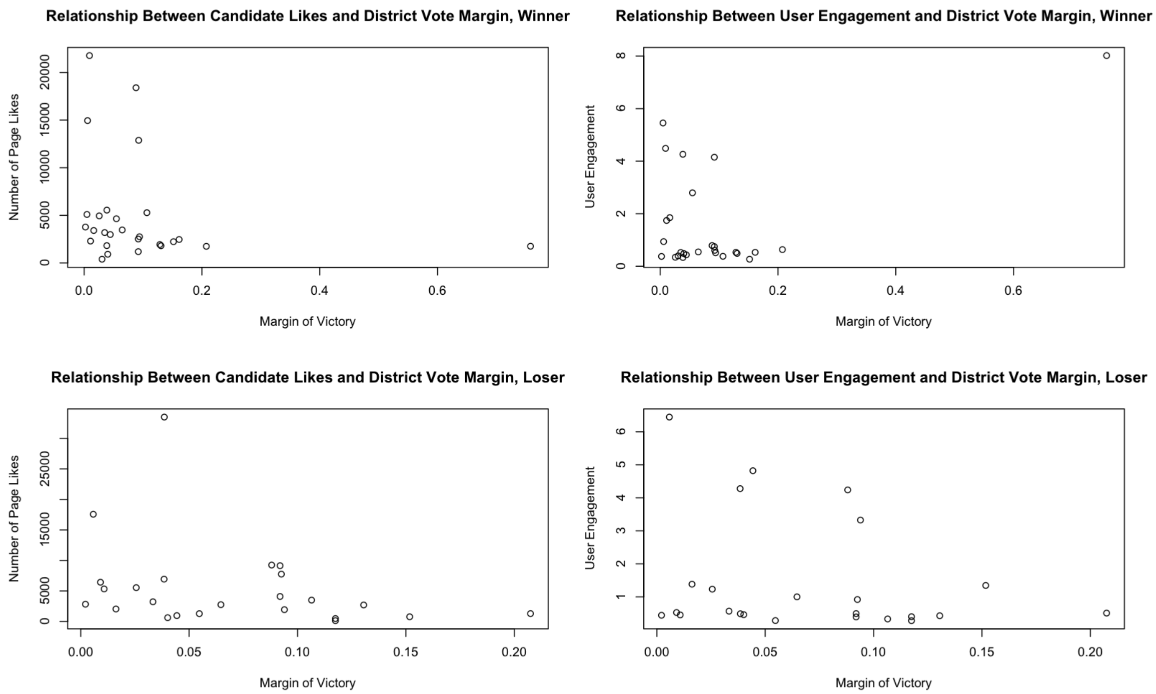


Regarding Figures 4.11 and Figures 4.12, partial correlations can be observed between the relationships evaluating engagement, total likes, and post frequency, respectively.

Focusing on the cluster in the bottom left quadrant of Figure 4.11, a minor correlation can be observed that suggests that candidate engagement improves with the number of likes a candidate accumulates on his or her page. It is possible to conclude (with minor confidence) that candidates are not achieving superficial likes, but have the ability, and the appeal, to keep users engaged after their first mode of interaction. Perhaps an individual "liking" a page suggests that he/she already had the propensity to seek engagement; it would be interesting to study the relationship between the acquisition of likes and the frequency by which [1] individuals sought out the page to "like" it, voluntarily, or [2] what percentage of candidates were on the fringe regarding "liking" a page, and decided to "like" a page as a result of a specific appeal.

In Figure 4.12, a (y = - x) relationship across the span of the x - axis. A similar relationship occurs in the bottom cluster. It is possible to infer that the quality of posts is better than the quantity – that it is better to invest in "Edge" value rather than focusing primarily on quantity to ensure that posts get exposure. Perhaps we can infer that regarding the EdgeRank algorithm, (uwd), the affinity score (u), is a better metric to focus on, representing quality – engagement of likes, comments, shares, versus the (d), time variable. The time variable decreases at the rate of 1/x, so it may be better to focus on the former to ensure maximum visibility of posts – a focus in this experiment.

[Figure 5.15 Relationship Between Engagement Efficacy, Likes and Vote Margin (4x4)]



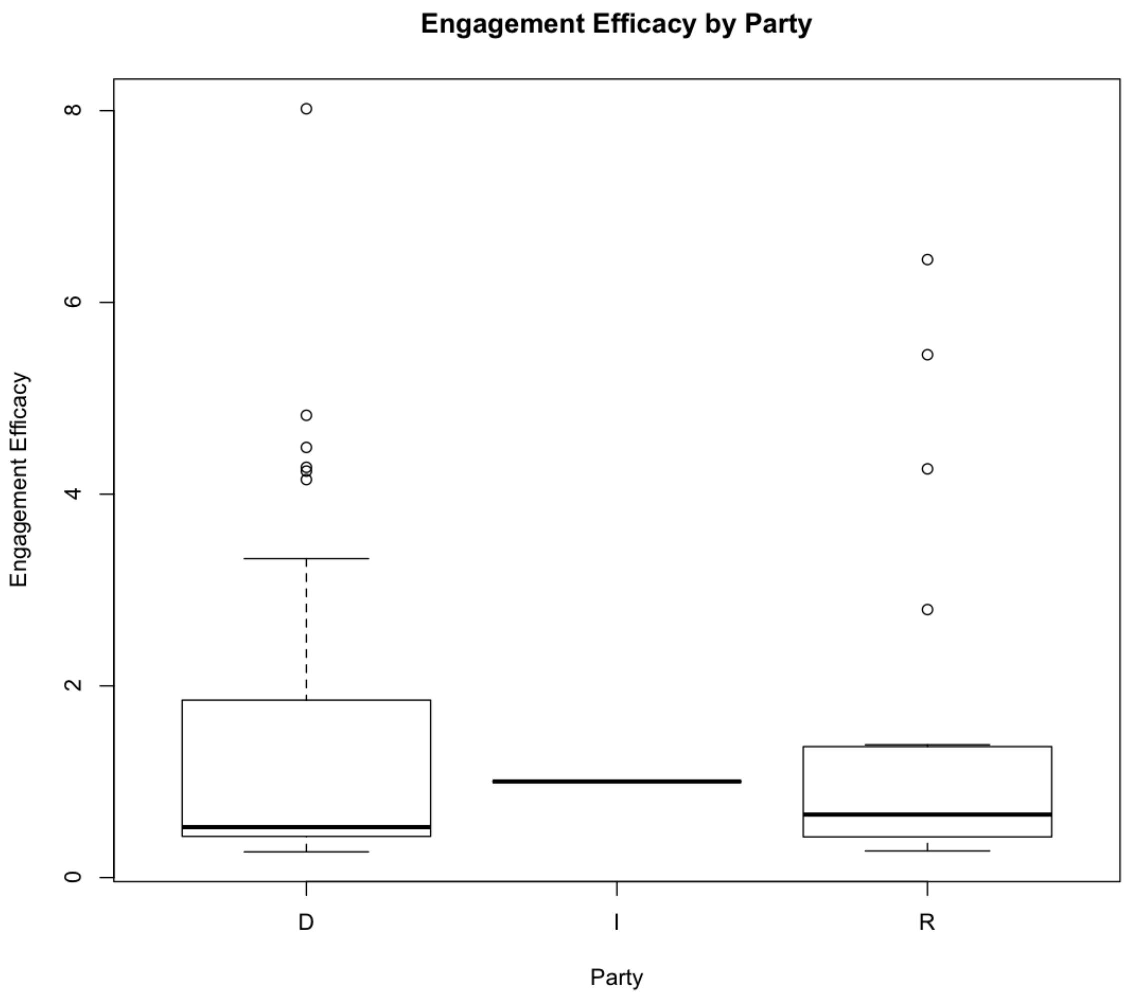
[Table 5.6 Relationship Between Engagement Efficacy, Likes and Vote Margin]

Relationship between margin of victory and:	**Correlation**	**Spearman**	**Kruskal Wallis**
Winner Engagement Efficacy	0.5023605	0.5945	**0.4631**
Winner Total Page Likes	-0.2002625	**0.03393**	**0.4631**
Loser Engagement Efficacy	-0.2333564	**0.1752**	**0.4108**
Loser Total Page Likes	-0.2979978	**0.03083**	**0.4608**

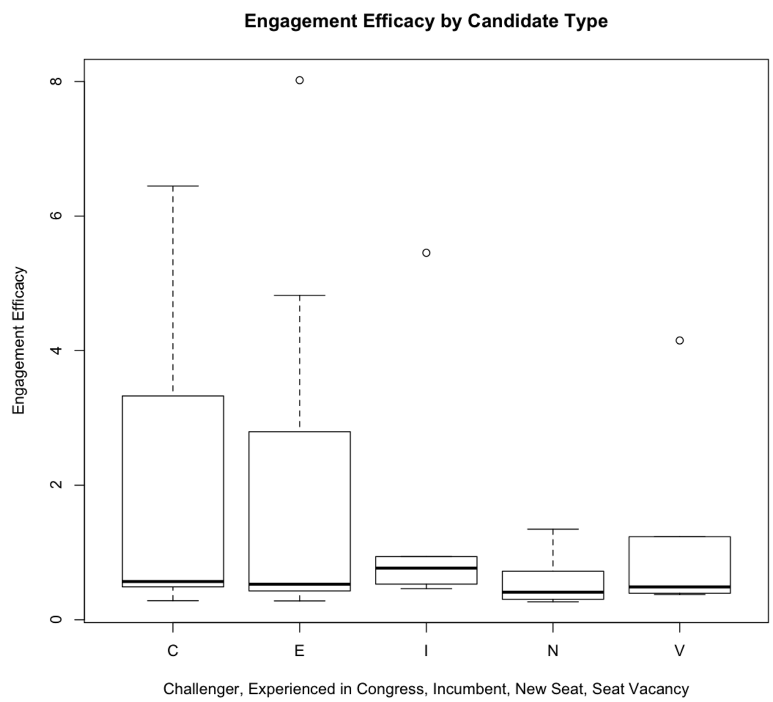
** For Spearman correlation significance test, $p < 0.05$ suggests statistical significance*
*** For Kruskal Wallis Robust ANOVA, $p < 0.05$ suggests differentiation between variables*

Although noise within Figures 14.3 and 14.4 do exist, it is apparent that there is a correlation between the strength of a candidate's engagement efficacy and the margin of victory during the 2012 Congressional Elections. As noted above, in the discussion of Figure 4.11, there is a relationship between total candidate page likes and user engagement. Examining these graphs, however, the trends appears to be that there is a more pronounced correlation when observing the relationship between vote margin on the basis of engagement efficacy rather than total page likes.

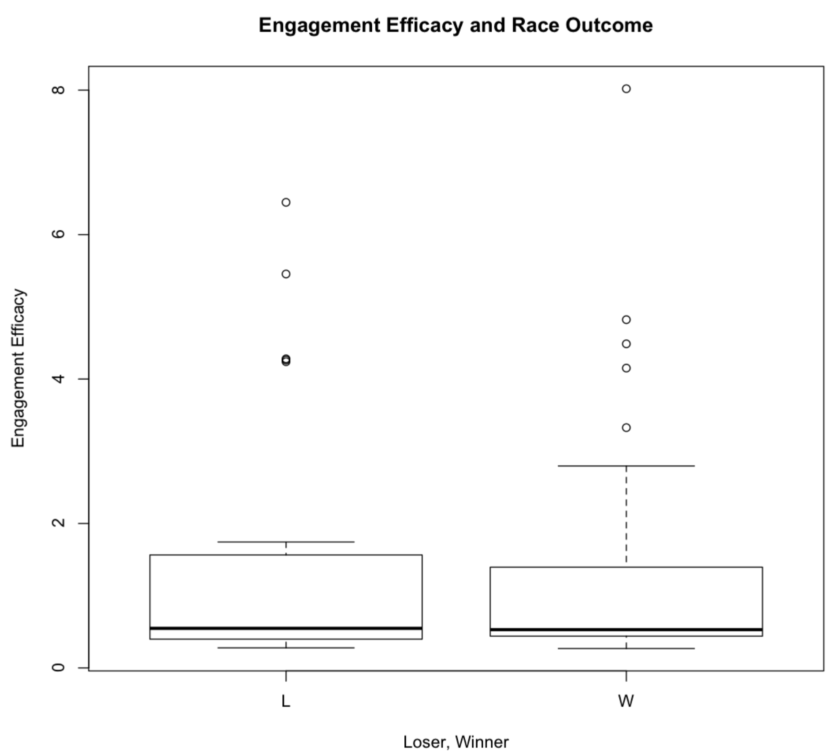
[Figure 5.16 Engagement Efficacy by Party]



[Figure 5.17 Engagement Efficacy by Candidate Type]



[Figure 5.18 Engagement Efficacy, Winners and Losers]



Figures 4.14, 4.15, and 4.16 display, respectively, each party's engagement efficacy, each type of candidate's engagement efficacy [challenger : redistricted incumbent (experienced in congress) : incumbent : new seat : vacant seat], and losers and winners. Democrats have a greater propensity than incumbents or republicans to engage their audiences, indicated by the larger spread of its upper and lower quartiles, with similar results for challengers and redistricted incumbents with respect to engagement in comparison to incumbents, those competing for new seats, and those competing for newly vacant seats. Although I would like to establish generalizations across these observations, that sample size is too small to do so (a total number of 53 pages with smaller subgroups per factor variable).

Chapter 6 – Conclusion

While candidates have not yet demonstrated substantial levels of proficiency associated with Facebook user engagement, it is possible to generalize some observations on the basis of the research involved with this study. On the basis of content appeals on Facebook, appeals to civic duty and candidate intimacy, on average, are better able to engage users on Facebook than do content appeals related to the national politics, local politics, or general information about the campaign. While candidates post more on Facebook on the days leading up to the election, late campaign advertising does not, on average, yield greater levels of engagement on Facebook. From this analysis, it is also possible to posit that it is easier to engage users on the basis of likes than it is on the basis of comments and shares. I can assert, with moderate confidence, that districts with lower levels of income and lower levels of educational attainment are more likely to engage with Facebook pages, candidates are able to engage their users despite increasing audience size, and candidates that post more often on Facebook are less likely to engage the users on their Facebook page. Additionally, comparing vote share with engagement efficacy is more accurate than comparing vote share with number of page likes, although these variables are somewhat correlated.

I have made these discoveries in the absence of a strong computer science background. An n – sample of 3,500 posts is relatively small for online mediums that have the power to process large amounts of information utilizing great computing power. Utilizing scripting language to optimize the data mining process, it would be possible to generalize these results with greater confidence as a by-product of a larger sample study size.

Nonetheless, the results I have uncovered contribute to the discourse on political science in that users don't necessarily understand the social media digital landscape. With a greater emphasis on engagement over mere frequency and attention to repetitive exposure, candidates will better be able to make valuable connections with their constituents. Although candidates now do not utilize Facebook to its fullest potential, I predict that over the course of the next two years, bodies of literature will contribute to the effectiveness of different social media tactics, and over the course of the next two to three elections, candidates will be better able to engage their audiences on Facebook.

References

Atkin, Bowen, Nayman, Sheinkopf, Quality Versus Quantity in Televised Political Ads, 1973

Berelson, Lazarsfield, Mcphee – Voting: A Study of Opinion Formation in a Presidential Campaign, 1954

Bernstein et al – Quantifying the Invisible Audience in Social Networks, 2013

Bond et al. A 61-Million-Person Experiment in Social Influence and Political Mobilization

Broockman, Green – Do Online Advertisements Increase Political Candidates' Name Recognition or Favorability? 2013

Centola – The Spread of Behavior in an Online Social Network Experiment, 2010

Christakis, Fowler, *Connected, 2009*

Easley, Kleinberg – Networks, Crowds, and Markets, 2010

Ellison , Steinfield, Lamp -The Benefits of Facebook "Friends:" Social Capital and College Student's Use of Online Social Network Sites, 2007

Fowler – Turnout in a Small World, 2005

Gerber, Green - *Social Pressure and Voter Turnout: Evidence from a Large-Scale Experiment, 2008*

Gerber, Green – The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment, 2008

Gerber, Green, Larimer – Social Pressure and Voter Turnout: Evidence from a Large-Scale Experiment, 2000

Pew Research Center – *One-in-Ten 'Dual Screened' the Presidential Debate*, 2012

Salaff et al. – Computer networks as social networks: Virtual community, computer-supported cooperative work and telework, 1996

Appendices

Original Dataset Workbook

DISTRICT	INCOME	EDU	RACE.TYPE	T.VS	T.EXP	T.REA	T.ENG	WINNER	W.AGE	W.EDU	W.PARTY2	W.TYPE	W.VS	W.EXP	W.REA	W.ENG	W.M.AGE	LOSER	L.AGE	L.PARTY	L.EDU	L.TYPE	L.VS	L.EXP	L.REA	L.ENG	L.M.AGE
AZ 1	43377	23.5	OS	226930	3844636	3169	3308	A. Kirkpatric	62	JD Private	D	55-64	117422	2355342	3155	3308		J. Paton	41	R	MA Public		109508	1489294	14	0	NA
AZ 2	44921	30.3	IV	275392	4128660	6504	5926	R. Barber	67	BA Public	D	25-44	137993	2690570	3670	2501		M. McSally	47	R	BA Private		137399	1438090	2834	3425	45-54
AZ 9	48033	34.4	ND	209145	3289318	26304	36945	K. Sinema	36	JD Public	D	25-44	108056	2118461	23091	33756		V. Parker	53	R	JD Private		101089	1170857	3213	3189	45-54
CA 07	60537	30.9	CV	235628	6248285	5461	6078	A. Bera	47	MD Public	D	25-34	119726	3531711	3411	5863		D. Lungren	66	R	JD Private		115902	2716574	2050	215	45-54
CA 09	52209	18.7	IV	188389	5691934	4696	3343	J. McNerney	61	BA Public	D	25-34	103044	2744514	2732	1628		R. Gill	25	R	JD Public		85345	2947420	1964	1715	25-34
CA 15	82179	38.6	CV	221566	2196715	3841	3220	E. Swalwell	32	JD Public	D	25-34	115694	799576	2908	3014		P. Stark	81	D	MA Public		105872	1397139	933	206	25-34
CA 26	72804	31.6	OS	232443	4483916	1443	1686	J. Brownley	60	BA Private	D	25-34; 55-64	120881	2106886	837	1553		T. Strickland	43	R	BA Private		111562	2377030	606	133	25-44
CA 30	67079	40	I V I	194407	11944794	3945	2103	B. Sherman	58	JD Private	D	18-24	117374	6172614	1957	1285		H. Berman	71	D	BA Public		77033	5772180	1988	818	25-34
CA 33	89354	62.4	IV	267023	10645394	6177	4914	H. Waxman	73	JD Public	D	55-64	142142	2663179	3421	3522		B. Bloomfiel	66	I	MA Private		124881	7982215	2756	1392	13-24
CA 35	51699	13.2	CV	129866	1456885	576	1173	G. McLeod	71	CC	D	25-34	72562	294428	455	847		J. Baca	66	D	BA Public		57304	1162457	121	326	35-44
CA 36	42922	19.4	CV	190350	4356585	33612	39119	R. Ruiz	40	MD Private	D	45-64	98843	1933069	546	15260		M. Bono Ma	51	R	BA Private		91507	2423516	33066	23859	55-64
CA 41	49887	15.4	OS	155589	2788415	2934	2708	M. Takano	52	BA Private	D	18-34	89599	1433468	2159	2137		J. Tavaglione	65	R	BA Private		65990	1354947	775	571	45-64
CA 47	55590	29.2	OS	201589	2589871	8729	3366	A. Lowenthal	71	PhD Private	D	45-54	111529	1171131	5264	1923		G. Delong	52	R	MA Private		90060	1418740	3465	1443	45-54
CA 52	77409	54.6	CV	246832	7125007	7603	4183	S. Peters	54	JD Private	D	45-54	124746	4352737	2267	1888		B. Bilbray	62	R	CC		122086	2772270	5336	2295	45-54
CO 03	47012	29.9	IV	317987	4178209	4057	2649	S. Tipton	56	BA Public	R	45-64	179745	2225171	1364	238		S. Pace	36	D	MA Public		138242	1953038	2693	2411	25-44
CO 06	63513	39.1	IV	290026	5533114	8787	8338	M. Coffman	57	BA Public	R	45-64	150587	3838137	1818	1462		J. Miklosi	76	D	MA Public		139439	1694977	6969	6876	35-54
CO 07	55341	29.3	IV	304601	7820466	1884	1612	E. Perlmutte	59	JD Public	D	25-34	171874	2988391	1884	1612		J. Coors	67	R	MA Public		132727	4832075	0	0	NA
CT 5	63275	34.3	OS	277401	4802514	2282	3050	E. Esty	53	JD Private	D	25-34	142912	3228060	928	467		A. Roraback	53	R	JD Public		134489	1574454	1354	2583	
FL 02	42107	27.1	OS	333441	2443203	5925	2876	S. Southerlan	47	BA Public	R	45-54	175842	1722317	4648	2465		A. Lawson	64	D	MA Public		157599	720886	1277	411	35-44
FL 10	48832	27.7	IV	321519	3436640	10482	9041	D. Webster	63	BA Public	R	45-54	164873	1498872	4961	4649		V. Demings	55	D	BA Public		156646	1937768	5521	4392	35-44
FL 18	47516	28.9	CV	330539	22951644	31263	38565	P. Murphy	29	BA Public	D	45-64	166223	4480428	14814	13316		A. West	52	R	MA Public		164316	18471216	16449	25249	45-54
FL 20	35941	17.7	IV	243726	1025445	1696	18	A. Hastings	76	JD Public	D	35-44	214245	558062	1696	18		R. Terry	53	I	BA Private		29481	467383	0	0	NA
FL 22	51227	35.6	OS	312845	6779342	20583	11564	L. Frankel	64	JD Private	D	45-64	170899	3410565	12805	10029		A. Hasner	43	R	JD Public		141946	3368777	7778	1535	55-64
H I01	65602	32.2	IV	210187	2086078	6598	1826	C. Hanabusa	61	JD Public	D	45-64	114756	1162296	2477	1133		C. Djou	42	R	JD Private		95431	923782	4121	693	35-54
IA 04	45454	21	IV	369154	7171713	27716	18919	S. King	63	BA Public	R	55-64	200831	3815765	18436	8807		C. Vilsack	62	D	BA Private		168323	3355948	9280	10112	25-34; 55-64
IL 10	65864	41.5	CV	259523	7560689	28392	17662	B. Schneider	51	MA Private	D	55-64	130941	3029605	21858	13949		R. Dold	43	R	JD Public		128582	4531084	6534	3713	18-24; 25-44
IL 11	65938	34.5	CV	240870	6767995	2472	2194	B. Foster	57	JD Private	D	25-34	139860	3532806	2436	2194		J. Biggert	75	R	JD Private		101010	3235189	36	0	NA
IL 12	42181	20.9	OS	283203	2513897	10355	4404	W. Enyart	63	JD Public	D	25-34	154621	1167686	1185	1526		J. Plummer	31	R	BA Public		128582	1346211	9170	2878	25-34
IL 13	44915	28	OS	271905	2705614	11585	13115	R. Davis	43	BA Private	R	35-54	136596	1383024	5061	5447		D. Gill		D	MD Public		135309	1322590	6524	7668	45-64
IL 17	41194	16.7	CV	278199	4694631	8530	5660	C. Bustos	51	MA Public	D	25-34; 45-54	148229	2187283	2211	2743		B. Schilling	49	R	CC		129970	2507348	6319	2917	45-54
IN 2	44494	20.1	OS	263783	3146042	9000	10579	J. Walorski	49	BA Private	R	35-54	133806	1878019	5171	1790		B. Mullen		D	BA Private		129977	1268023	3829	8789	25-34
KY 06	43399	29.3	CV	294657	4744720	3596	3536	A. Barr	39	JD Public	R	25-44	153221	2211676	3596	3536		B. Chandler	53	D	JD Public		141436	2533044	0	0	NA
MA 06	76130	40.3	IV	355556	4651798	8152	7858	J. Tierney	61	JD Private	D	25-34	179603	2299261	2869	1885		R. Tisel	50	R	BA Private		175953	2352537	5283	5973	45-54
MI 01	40765	21.1	IV	331369	3616116	4332	3987	D. Benishek	60	MD Public	R	45-54	166833	2128930	2516	1548		G. McDowell	60	D	BA Public		164536	1487186	1816	2439	25-34
MI 03	48010	27.2	IV	318646	3270291	38192	15394	J. Amash	32	JD Public	R	25-34	173529	1193611	31362	14483		S. Pestka	61	D	JD Public		145117	2076680	6830	911	13-17
MI 11	69397	44.1	OS	340695	1159801	1769	1168	K. Bentivolio	61	MD Private	R	25-34	181796	482219	1769	1168		S. Taj	66	D	MD Public		158899	677582	0	0	NA
MN 06	65461	28.1	IV	356054	25004512	482900	444889	M. Bachman	56	JD Private	R	45-54	180131	22725128	470754	423531		J. Graves	59	D	BA Public		175923	2279384	12146	21358	45-64
MN 08	46692	20.9	CV	353861	3579980	9421	9996	R. Nolan	69	MA Public	D	55-64	192748	1194314	2422	3024		C. Cravaack	53	R	MA Public		161113	2385666	6999	6972	45-54
NC 07	41935	21.9	IV	334647	3725191	5090	2961	M. McIntyre	56	JD Public	D	25-34; 45-54	167590	2253834	3713	1650		D. Rouzer	41	R	BA Public		167057	1471357	1377	1311	35-54
NC08	38549	16.2	CV	294117	2751739	0	0	R. Hudson	41	BA Public	R	NA	159226	1443064	0	0		L. Kissell	62	D	BA Private		134891	1308675	0	0	NA
ND01	51704	26.3	OS	304301	2310930	2762	2748	K. Cramer	52	MA Private	R	18-34	172905	1291733	1288	744		P. Gulleson	48	D	MA Public		131396	1019197	1474	2004	25-34
NE 2	55114	35.7	IV	248284	2618417	6804	6024	L. Terry	51	JD Private	R	45-54	127119	2033246	5056	1791		J. Ewing	52	D	MA Public		121165	585171	1748	4233	35-54
NH 01	63587	33	CV	329838	3544376	7584	6722	C. Shea-Port	60	MA Public	D	25-34	171356	1696703	3233	4212		F. Guinta	42	R	JD Private		158482	1847673	4351	2510	35-44
NH 02	61832	33.8	CV	320812	5246062	3757	3068	A. Kuster	56	JD Private	D	25-34	168954	3173921	3757	3068		C. Bass	61	R	BA Private		151858	2072141	0	0	NA
NY 01	83114	33.6	IV	254003	5316186	945	524	T. Bishop	62	MA Private	D	45-64	132525	3031036	934	524		R. Altschuler	42	R	MA Private		121478	2285150	11	0	NA
NY 11	62054	30.3	IV	176503	3121996	10932	5178	M. Grimm	43	JD Private	R	35-54	94102	2204598	10410	4737		M. Murphy	41	D	BA Private		82401	917398	522	441	25-34
NY 18	71399	33.7	CV	252373	5500080	12329	10228	S. Maloney	46	JD Public	D	25-34; 45-54	130462	2246008	1761	3128		N. Hayworth	53	R	MD Private		121911	3254072	10568	7100	25-34
NY 19	53769	25.5	IV	259318	3480574	2027	3610	C. Gibson	48	PhD Private	R	35-54	138657	2153561	864	2010		J. Schreibma	41	D	JD Private		120661	1327013	1163	1600	45-64
NY 21	48759	21.7	IV	226687	3909788	400	714	B. Owens	64	JD Private	D	18-34	115883	1939276	368	714		M. Doheny	41	R	JD Private		110804	1970512	32	0	NA
NY 23	44518	23.3	IV	243574	2835265	9977	15931	T. Reed	41	JD Private	R	35-44	126519	2006454	4072	3779		N. Shinagaw	28	D	MA Private		117055	828811	5905	12152	25-34
NY 24	51274	28.4	CV	249547	4103086	4813	4036	D. Maffei	44	MA Private	D	25-44	130969	2047007	1642	1916		A. Buerkle	61	R	JD Private		118578	2056079	3171	2120	45-54
NY 27	55340	26.9	CV	296580	5915626	8175	4838	C. Collins	62	MA Public	R	45-54	150446	1312830	3756	1593		K. Hochul	54	D	JD Private		146134	4602796	4419	3245	25-34 : 45-54
OH 16	56251	30.5	I V I	346775	5837872	5585	3739	J. Renacci	54	BA Public	R	45-54	181137	3284373	4094	1836		B. Sutton	49	D	JD Public		165638	2553499	1491	1903	25-34
OR04	40197	24.1	IV	347257	2633851	143196	11419	P.DeFazio	65	MA Public	D	55-64	208906	1401802	14679	10145		A. Robinson	70	R	PhD Public		138351	1232049	128517	1274	55-64
PA 12	53080	29.2	CV	334737	4598639	1808	3614	K. Rothfus	50	JD Private	R	45-54	173340	1993455	1404	2031		M. Critz	51	D	BA Public		161397	2605184	404	1583	25-34
RI 02	57448	30.7	IV	198408	2130309	997	499	J. Langevin	48	MA Private	D	25-34	121556	1131248	735	462		M. Riley	55	R	MA Private		76852	999061	262	37	45-54
TN 09	36142	22.8	IV	247921	2388984	4173	4101	S. Cohen	63	JD Public	D	25-34	188245	820327	3050	3756		G. Finn	68	R	MD Public		59676	1568657	1123	345	35-54
TX 14	50178	22.4	OS	240201	2464745	8974	12829	R. Weber	59	BA Public	R	45-54	130937	1289706	3354	4956		N. Lampson	68	D	MA Public		109264	1175039	5620	7873	45-64
TX 23	48232	20.7	CV	183732	4302073	3983	5807	P. Gallego	51	JD Public	D	35-54	96477	1767938	1699	3801		Q. Canseco	63	R	JD Private		87255	2534135	2284	2006	45-54
UT 04	57124	27	IV	213904	4753749	47329	87628	J. Matheson	52	MA Public	D	35-54	108275	2383305	5354	12407		M. Love	37	R	BA Public		105629	2370444	41975	75221	35-54
WA 06	51982	27.2	OS	303079	3533563	2115	2875	D. Kilmer	39	PhD Private	D	35-44	178701	1706202	2115	2875		B. Driscoll	50	R	MA Private		124378	1827361	0	0	NA

Modified Dataset Workbook, After Accounting for Error

CANDIDATE	ENGAGEMEN	VOTE.SHARE	MARGIN	AUDIENCE	PARTY	D.INCOME	D.EDU	C.EDU
Adam Hasne	0.92086877	141946	0.09254743	55-64	R	51227	35.6	JD Private
Al Lawson	0.28325143	157599	0.05471133	35-44	D	42107	27.1	BA Public
Alan Lowentl	0.37849117	111529	0.10649887	45-54	D	55590	29.2	PhD Private
Alcee Hastin	8.01997041	214245	0.7580808	35-44	D	35941	17.7	JD Public
Allen West	6.44679183	164316	0.00576937	45-54	R	47516	28.9	BA Public
Ami Bera	1.85116621	119726	0.01622897	25-34	D	60537	30.9	MD Public
Ann Kirkpatri	0.52357508	117422	0.03487419	55-64	D	43377	23.5	JD Private
Bill Bloomfiel	1.00243835	124881	0.06464237	13-24	I	89354	62.4	JD Public
Bill Enyart	4.151638	154621	0.09194465	25-34	D	42181	20.9	JD Public
Bill Foster	0.52999754	139860	0.16129032	25-34	D	65938	34.5	JD Private
Brad Schneid	4.48731962	130941	0.00908975	55-64	D	65864	41.5	MA Private
Brad Sherma	0.6343924	117374	0.20750796	18-24	D	67079	40	JD Private
Brian Bilbray	0.45444187	122086	0.01077656	45-54	R	77409	54.6	JD Private
Charles Djou	0.3945379	95431	0.09194194	35-54	R	65602	32.2	JD Public
Christie Vilsa	4.24023327	168323	0.08806081	25-34; 55-64	D	45454	21	BA Public
Colleen Hana	0.74929081	114756	0.09194194	45-64	D	65602	32.2	JD Public
Dan Lungren	1.38549556	115902	0.01622897	45-54	R	60537	30.9	MD Public
Daniel Webs	0.34242989	164873	0.02558791	45-54	R	48832	27.7	BA Public
Ed Perlmutte	0.53063201	171874	0.12851895	25-34	D	55341	29.3	JD Public
Elizabeth Est	0.39490645	142912	0.03036399	25-34	D	63275	34.3	JD Private
Eric Swalwell	0.43548762	115694	0.04432991	25-34	D	82179	38.6	JD Public
Gary DeLong	0.33256976	90060	0.10649887	45-54	R	55590	29.2	PhD Private
Gloria McLec	0.4014175	164536	0.11749034	25-34	D	40765	21.1	MD Public
Henry Waxm	0.54668949	142142	0.06464237	55-64	D	89354	62.4	JD Public
Howard Berr	0.50941032	77033	0.20750796	25-34	D	67079	40	JD Private
Jason Plumm	0.49464266	128582	0.09194465	25-34	R	42181	20.9	JD Public
Jerry McNerr	0.51193925	103044	0.09394922	25-34	D	52209	18.7	BA Public
Joe Baca	0.28032103	57304	0.11749034	35-44	D	51699	13.2	CC
Joe Miklosi	4.27978001	139439	0.03843793	35-54	D	63513	39.1	BA Public
John Tavaglic	1.34691328	65990	0.15173952	45-64	R	49887	15.4	BA Private
Julia Brownle	0.48394109	120881	0.04009155	25-34; 55-64	D	72804	31.6	BA Private
Lois Frankel	0.26831151	170899	0.09254743	45-64	D	51227	35.6	JD Private
Mark Takano	0.4437789	89599	0.15173952	18-34	D	49887	15.4	BA Private
Martha McSa	0.49029458	137399	0.00215693	45-54	R	44921	30.3	BA Public
Mary Bono N	4.26351183	91507	0.03853953	55-64	R	42922	19.4	MD Private
Mike Coffma	0.93888847	150587	0.03843793	45-64	R	63513	39.1	BA Public
Patrick Murp	4.8222463	166223	0.00576937	45-64	D	47516	28.9	BA Public
Pete Stark	0.33336798	105872	0.04432991	25-34	D	82179	38.6	JD Public
Raul Ruiz	3.32726473	98843	0.03853953	45-64	D	42922	19.4	MD Private
Ricky Gill	0.5268464	85345	0.09394922	25-34	R	52209	18.7	BA Public
Robert Dold	5.45445037	128582	0.00908975	18-24; 25-44	R	65864	41.5	MA Private
Rodney Davis	0.3748199	136596	0.00473327	35-54	R	44915	28	BA Private
Ron Barber	0.42989451	137993	0.00215693	25-44	D	44921	30.3	BA Public
Sal Pace	1.74362975	138242	0.13051791	25-44	D	47012	29.9	BA Public
Scott Peters	0.48963864	124746	0.01077656	45-54	D	77409	54.6	JD Private
Scott Tipton	0.78801969	179745	0.13051791	45-64	R	47012	29.9	BA Public
Steve King	2.79597862	200831	0.08806081	55-64	R	45454	21	BA Public
Steve Southe	0.46281558	175842	0.05471133	45-54	R	42107	27.1	BA Public
Tony Strickla	1.23475385	111562	0.04009155	25-44	R	72804	31.6	BA Private
Val Demings	0.56899221	156646	0.02558791	35-44	D	48832	27.7	BA Public
Vernon Park	0.27799057	101089	0.03331182	45-54	R	48033	34.4	JD Public

Sample Data Mining Set before Additions w/Date (1/150 pages)

District	Candidate	Post Type2	Likes	Comments	Shares
AZ 1	Ann Kirkpatri	PLP	55	9	3
AZ 1	Ann Kirkpatri	PLP	45	5	0
AZ 1	Ann Kirkpatri	PLP	56	3	0
AZ 1	Ann Kirkpatri	PLP	59	10	3
AZ 1	Ann Kirkpatri	PLP	45	2	4
AZ 1	Ann Kirkpatri	PHLE	57	2	4
AZ 1	Ann Kirkpatri	PIE	140	12	14
AZ 1	Ann Kirkpatri	PIC	40	3	1
AZ 1	Ann Kirkpatri	PIE	101	19	6
AZ 1	Ann Kirkpatri	PIE	48	3	3
AZ 1	Ann Kirkpatri	PLP	53	2	0
AZ 1	Ann Kirkpatri	PIE	43	4	4
AZ 1	Ann Kirkpatri	PHPL	80	19	5
AZ 1	Ann Kirkpatri	PHLE	70	3	1
AZ 1	Ann Kirkpatri	PHPL	56	4	1
AZ 1	Ann Kirkpatri	PNP	48	1	0
AZ 1	Ann Kirkpatri	PNP	50	6	1
AZ 1	Ann Kirkpatri	PHPL	79	9	3
AZ 1	Ann Kirkpatri	PIC	4	0	0
AZ 1	Ann Kirkpatri	PPL	52	5	3
AZ 1	Ann Kirkpatri	PHLE	113	10	1
AZ 1	Ann Kirkpatri	PLP	126	8	0
AZ 1	Ann Kirkpatri	PHLE	93	7	4
AZ 1	Ann Kirkpatri	PHLE	113	8	11
AZ 1	Ann Kirkpatri	PIC	238	46	4
AZ 1	Ann Kirkpatri	PIC	50	14	0
AZ 1	Ann Kirkpatri	PCP	660	122	40
AZ 1	Ann Kirkpatri	PPL	254	25	3
AZ 1	Jonathan Pat	NA	0	0	0
AZ 2	Ron Barber	PCP	69	10	3
AZ 2	Ron Barber	PIE	117	6	2
AZ 2	Ron Barber	PHPL	71	6	3
AZ 2	Ron Barber	PIE	101	8	13
AZ 2	Ron Barber	PIE	144	5	2
AZ 2	Ron Barber	PIE	76	12	1
AZ 2	Ron Barber	PIC	50	2	0
AZ 2	Ron Barber	PIC	9	1	0
AZ 2	Ron Barber	PHPL	155	17	2
AZ 2	Ron Barber	PHLE	76	9	0
AZ 2	Ron Barber	PIE	74	3	1
AZ 2	Ron Barber	PPL	94	7	3