

I read it on Reddit: Exploring the role of online communities in the 2016 US elections news cycle

Jon Roozenbeek¹ and Adrià Salvador Palau²

¹ Department of Slavonic Studies, Faculty of Modern and Medieval Languages, University of Cambridge, Cambridge, CB3 9DA, UK

jjr51@cam.ac.uk,

² Department of Engineering, Institute for Manufacturing, University of Cambridge, Cambridge, CB3 0FS, UK

Abstract. Reddit has developed into a significant platform for political discussion among Millennials. In this exploratory study, we examine subscription trends on three political sub-forums on Reddit during the 2016 US presidential elections: /The_Donald, /SandersForPresident, and /HillaryClinton. As a theoretical framework, we draw from work on online communities' group identity and cohesion. Concretely, we investigate how subscription dynamics relate to positive, negative and neutral news events occurring during the election cycle. We classify news events using a sentiment analysis of event-related news headlines. We observe that users who supported Sanders displayed no consolidation of support for Clinton after she won the Democratic Party's presidential nomination. Secondly, we show that negative news events affected Sanders and Clintons subscription trends negatively, while showing no effect for Donald Trump. This gives empirical credence to Trump's controversial claim that he could "stand in the middle of 5th Avenue and shoot somebody and not lose any voters". We offer a number of explanations for the observed phenomena: the nature of the content of the three subreddits, their cultural dynamics, and changing dynamics of partisanship. We posit that the 'death of expertise' expresses itself on Reddit as a switch in persuasion tactics from a policy-based to an emotions-based approach, and that group members' agreement on policy proved a weak marker for online communities' group identity and cohesion. We also claim that strong partisanship coupled with weak party affiliation among Millennials contributed to the low levels of Democratic support consolidation after Clinton won the nomination.

Keywords: Electoral politics; US elections 2016; social media; Reddit; Millennial generation

1 Introduction

Since Donald Trump’s electoral victory in the US presidential elections on November 8 of 2016, many have sought to explain Secretary of State Hillary Clinton’s loss to Donald Trump, as well as the insurgent candidacy of Senator Bernie Sanders during the Democratic primaries [7,39,13,11]. In 2016, the Millennial generation (people born between 1981 and 1997) was seen as a crucial demographic by Republican as well as Democratic party strategists [46], even more so than in the 2008 and 2012 elections [34,14]. In order to attract these Millennials, campaigns have actively engaged with young voters through social media activity such as Facebook, Twitter and Instagram [27,12].

This has prompted scholars of political communication and electoral campaigns to focus on the growing role of social media in US elections. Pancer and Maxwell [50] look at what kinds of political messages are shared on social media during the 2016 elections. Lee and Lim [37] investigate the differences in the ways Trump and Clinton communicated with their supporters on social media. Alashri et al. [2] look at Facebook dynamics, and specifically commentator sentiments with regards to candidates’ social media engagement. Hu et al. [29] employed a similar approach for user sentiments on Twitter. Effing et al. [24] show that politicians with higher social media engagement received more preference votes within their political parties. Groshek and Koc-Michalska [28] explore if social media use was linked to increased support for ‘populist’ candidates Bernie Sanders and Donald Trump. Their findings suggest that political social media use was linked to less support for Republican candidates, including Trump, but to more support for Sanders. In terms of the methodology used in these and similar studies, social network analytics and big data continue to provide insight into this interplay between the youngest voting generation, political campaigns, and online media [41,43].

Facebook and Twitter have received the bulk of attention in the above-mentioned studies. However, other social media outlets are increasingly asserting themselves as hubs of political engagement. One such platform is Reddit (www.Reddit.com), a social news aggregation and discussion site. The site is divided into sub-fora, commonly called ‘subreddits’. On these subreddits, users discuss topics that can range from music, film and politics to sex, practical advice and product reviews. During the 2016 presidential election cycle, several candidates actively used Reddit in their campaign strategy; for example, both Bernie Sanders and Donald Trump held ‘Ask me Anything’-sessions on the site, during which Reddit users could pose questions directly to them. As such, Reddit has become an important political platform where mostly young voters exchange ideas and events, raise funds for candidates, and work out strategies. Its role in the 2016 US presidential elections, however, has thus far remained unexplored by scholars.

In this study, we look at subreddit subscription trends for three subreddits dedicated to three 2016 presidential candidates: `/The_Donald` (Donald Trump), `/SandersForPresident` (Bernie Sanders), and `/HillaryClinton`. These three subreddits were by far the most popular out of all presidential candidates, with approximately 366.000, 215.000 and 40.000 subscribers, respectively, as of January 15, 2017. There are several reasons why we chose to focus on subscription dynamics and not content or traffic numbers. First of all, all three subreddits have a policy of allowing only subscribers to place comments. Furthermore, they only allow actual supporters to write posts, and ban or remove users who post content critical of the candidate. In addition, both `/The_Donald` and `/HillaryClinton` have hidden the downvote-buttons by means of a skin, which means that even subscribers would find it hard to vote down a particular thread. Therefore, especially for Trump’s and Clinton’s subreddits, we assume that most subscribers are actually supporters of the candidate, as there is no rationale to subscribe from a non-supporter point of view. We extend this assumption to `/SandersForPresident`. Traffic numbers, on the other hand, also include non-subscribing visitors, who may or may not support a candidate. Second, we do not quantitatively analyse the content of the posts, although we do cite illustrative examples in section 4. The reason for this is that the number of active Reddit users (defined as people who post three or more posts on a given subreddit) is disproportionately low compared to the number of visitors: approximately 6 in 10 Reddit users who posted at least one comment only posted *one* comment, and only a quarter of commenters posted three or more comments in total. This means that the level of active participation is quite low, but passive participation among subscribers (meaning reading but not commenting) can be very high [8]. Thus, subscription dynamics are the most effective way of mapping not only active, but also passive participation, while at the same time staying confined within the body of Reddit users that can be seen as actual supporters rather than visitors.

Our study is the first to marry Reddit subscription dynamics with news events occurring during electoral campaigns. To do so, we run a sentiment analysis algorithm to establish the average polarity of headlines on 3 major news outlets. This allows us to classify events as positive, neutral or negative for a particular candidate.

We place our study within the theoretical framework of online group identity and cohesion [51]. The difference in definition between these two concepts is that group cohesion is defined as how much group members like the other individuals in the group, whereas group identity refers to how much group members like the group’s purpose. This framework predicts that the ability of an online community to keep its members together and grow larger is dependent on fostering members’ interdependence, social interactions and interpersonal attraction. Furthermore, intergroup comparisons, social categorisation and setting common goals are key factors in determining the strength of an online community. In this paper, we consider the three subreddits as separate online communities. In terms of theoretical development, we study how communities interact with a relevant news cycle, and what factors determine any observed differences in the reception of positive and negative news.

Concretely, we ask the following research questions: First, to which extent does the positivity or negativity of news events affect Reddit subscription dynamics? Second, what differences can we discern among the three candidates’ subreddits? Third, how do we explain the observed differences? And fourth, how can these observations be placed within the context of the role of social media in electoral campaigns?

Answering these questions enables researchers as well as analysts working for electoral campaigns to investigate which types of news events resonate with social media users, and gain insight into how online political communities operate. Additionally, this opens up opportunities to build a more general framework on how news cycles and social media affect one another.

2 Methods

This paper’s data set consists of two main components: Reddit visitor and subscriber metrics, and a list of news events relevant during the 2016 US elections. In this section, we describe the methodology used in order to analyse the visitor and subscriber trends, and to classify the news events. Additionally, we discuss Reddit demographics as a proxy for broader political trends among Millennials.

2.1 Visitor trends

The primary data set consists of monthly visitor and daily subscriber numbers for /SandersForPresident, /The_Donald, and /HillaryClinton. It covers the period between the announcement of Bernie Sanders’ presidential campaign on April 30th, 2015, and November 9th, 2016, the day after the elections. This data was provided by a Reddit administrator team upon request by the authors, and complemented with data from www.redditmetrics.com.

We define two inter-related indicators; subscriber growth, or ΔS and accumulated subscribers, $S(T)$. Their relation can be simply stated as follows:

$$S(T) = \sum_{t=0}^{N(T)} \Delta S_t. \quad (1)$$

Where $t = 0$ corresponds to the date when the subreddit was created and $N(T)$ to the number of days until the time T . Note that ΔS_t can be both positive (subscriber increase) or negative (subscriber loss).

The reason for using subscriber metrics and not page views, comments or other indicators, is discussed in section 2.4.

The ideal indicator of the relation between subreddit subscription and subscriber support would be individual user subscription data. Unfortunately, such data is not available, since Reddit does not keep track of it. Therefore subscriber indicators remain the best available metrics for Reddit user support.

2.2 News event classification

In order to analyse the relation between news events occurring during the 2016 election campaign and subscriber variations on Reddit, we look at the association between positive, neutral and negative news events and subscriber trends. The list with these events was compiled using different timelines of the 2016 primary and general elections that are available online [64,5,26][36].

We selected those news events that were of direct political relevance to the election (such as primary elections, debates, party conventions, candidates’ speeches, et cetera) or particularly newsworthy otherwise (such as terrorist attacks and political scandals). We have allowed for a deviation from the day of the news event by one day; for example, a traffic bounce that occurs one day before or after a party

convention is considered to be associated with this convention. In case of unplanned or unpredictable events, the previous days are not considered.

Our method for determining if a news event is positive, negative or neutral relies primarily on a sentiment analysis of the headlines on these events on 3 major news outlets. For this, we used TextBlob, an open-source pre-trained Python library for processing natural language. We extracted the headlines for articles about each event in our data set from the New York Times, CNN and Wall Street Journal websites. These outlets were chosen based on their audiences' place on the ideological spectrum: according to a Pew Research survey from 2014, the New York Times' audience is predominantly somewhat liberal, CNN's is rather in the centre, and the Wall Street Journal's audience leans somewhat to the right [10].

By drawing from ideologically diverse sources, we sought to eliminate the possibility of a consistent ideological slant in the sentiment analysis sample as much as possible. In cases where there was more than one article per website about an event, we chose only one. We also disregarded any op-eds and editorials. Next, we ran Textblob's 'sentiment'-property over the three headlines to return an average event polarity between -1 and 1. We classified a polarity between -1 and -0.15 as a negative event; between -0.15 and .15 as a neutral event; and between 0.15 and 1 as a positive event.

Sentiment analysis, however, has its limitations. For example, Ted Cruz winning the Iowa primary is rather obviously a negative event for Donald Trump. Nonetheless, the algorithm returns a positive value, because the headlines in our data set emphasise Cruz' victory rather than Trump's loss. In order to smooth out such discrepancies, we therefore compared the sentiment analysis polarities to our independent classification. We individually classified each event as positive, negative or neutral to each candidate, and then compared our answers. In case a difference occurred, we discussed it and settled on a compromise.

After this, we compared the sentiment analysis classification to our own. In the majority of cases, the two overlapped. In cases such as the one mentioned above, where the algorithm is unambiguously wrong, we went with our classification. In all other cases, we gave the benefit of the doubt to the algorithm.

The full list of positively, negatively and neutrally classified news events can be found in the appendix (6).

In order to quantify the influence of news events on traffic and subscriber dynamics, we perform a linear fit with dummy nominal variables. We compare two equations, first a simple model for news events in general:

$$\Delta S = \beta_1 n + k. \quad (2)$$

Where ΔS is the total subscriber change, β_1 is the slope of the fit and k is the intercept. n is a dummy nominal variable, in other words, n is 1 when a relevant news event occurs and 0 when it does not. The standard error of the fit and the adjusted R^2 will give us an estimate of the predictive power of news events.

We compare this with a model that takes into account news categorised as positive and negative:

$$\Delta S = \beta_+ n_+ + \beta_- n_- + \gamma. \quad (3)$$

Where ΔS is the total subscriber change, β_+ is the slope related to positive news events and β_- the slope related to negative news events. n_+ and n_- are set to 0 except in the case when a positive or negative news event occurs in which they are correspondingly set to 1. γ is the intercept. The fits were all performed using the fitlm-function of the MATLAB mathematical computing software.

2.3 Reddit Demographics

We argue that Reddit's demographic make-up makes it a decent proxy for the Millennial population in the United States. There are three reports available on Reddit user demographics: a SurveyMonkey poll, a PingDom.com survey, and a Pew Research report. Their conclusions can be found in Fig. 4. This is a fairly limited data set, and it comes with a few caveats.

First, according to US census data, 61.6% of Americans consider themselves white, 17.6% non-white Hispanic, 13.3% black, 5.8% Asian or Pacific Islander, and 1.2% Native American. However, whereas Hispanics comprise 17.6% of the total US population [63], they form 24.4% of the population under 18 years of age [16]. The age pyramid for minorities in the US (particularly for Hispanics) is therefore skewed towards young people more so than for whites [19]. Since young people are more likely to use Reddit, it is possible that there is a certain amount of over-representation of Hispanics among US Reddit users. This, however, cannot be said for sure.

Furthermore, in terms of education levels, 9% of the Pew Research report's respondents without a high school diploma reported to use Reddit (n=99), more than those with a college degree (7%, n=79), some

college education (6%, n=517) and a high school degree (4%, n=473). The finding by Pew Research that non-high school graduates are the most likely to use Reddit is rather counter-intuitive (Fig. 4). Considering especially the small number of non-high school graduates in the sample, this should be seen as statistically suspicious.

Overall, the available surveys indicate that Reddit demographics are not fully representative of the US Millennial population as a whole. The website's user base somewhat skewed towards male, white, English-speaking users, although there are a few issues of statistical reliability and a lack of available data that make it difficult to state this definitively. In terms of ethnic origin, the only demographic within the US which can be safely assumed to be underrepresented in Reddit's user base is the black population; it is possible that whites and Asians are overrepresented on Reddit, but this cannot be said for sure. It is important to bear this imbalance in mind when using Reddit data to draw broader conclusions about Millennial political involvement.

Nonetheless, despite this somewhat skewed sample, Reddit's user base is certainly young and to a large extent politically engaged. Furthermore, considering the popularity of the platform and the growing relevance of online communities and social media in political campaigning [24,33,55], using Reddit could offer a new route to understanding Millennial online political participation.

2.4 Limitations

There are a number of limitations to our approach. First, we only look at one social media platform that is not used by a very large share of the US population. Therefore, drawing general conclusions about the 2016 elections should be done with care. Furthermore, as was mentioned above, Reddit as a platform is not fully representative of the US Millennial population. Therefore, the data set is not to be seen as a representative sample of Millennials in general. Nonetheless, the trends observed can be to some extent indicative of broader sentiments among members of this group, once the demographic differences are weighted. Furthermore, since social media plays an important role in modern election campaigns, the dynamics and trends visible within social media hubs are increasingly likely to reflect broader electoral dynamics, and as such are of value to researchers, campaign managers, and so on.

Second, our sentiment analysis does not measure audiences' perception of headlines or events. This is a common limitation for sentiment analysis [38]. Without looking at how individual Reddit users received and internalised a particular event and aggregating this data, it is difficult to draw firm conclusions about how news events affect group identity and candidates' levels of support. We nonetheless maintain that there is a measurable difference in the effects of positive and negative headlines that can inform substantial analysis.

Third, we limit ourselves mainly to subscription dynamics, which (as discussed above) we argue is the best available proxy for measuring true support for a candidate. However, there is a significant degree of active user overlap between /SandersForPresident and /HillaryClinton. We analysed this using the Reddit Subreddit Similarity and Algebra-tool developed by Trevor Martin at Short Tails [42]. Both subreddits have a subreddit similarity scoring (calculated as the cosine distance between positive pointwise mutual information vectors) of 0,74 (it is much lower for /The_Donald). This indicates that a significant number of active commenters were active on both threads. However, as we previously discussed, number of active users on Reddit is low compared to the total number of subscribers. Furthermore, it is important to acknowledge the difference in total number of subscribers between /SandersForPresident and /HillaryClinton. Nonetheless, we should be careful in our analysis to take this overlap into account.

3 Results

The accumulated subscriber data set against time can be divided into three different consecutive periods (see Fig. 1).

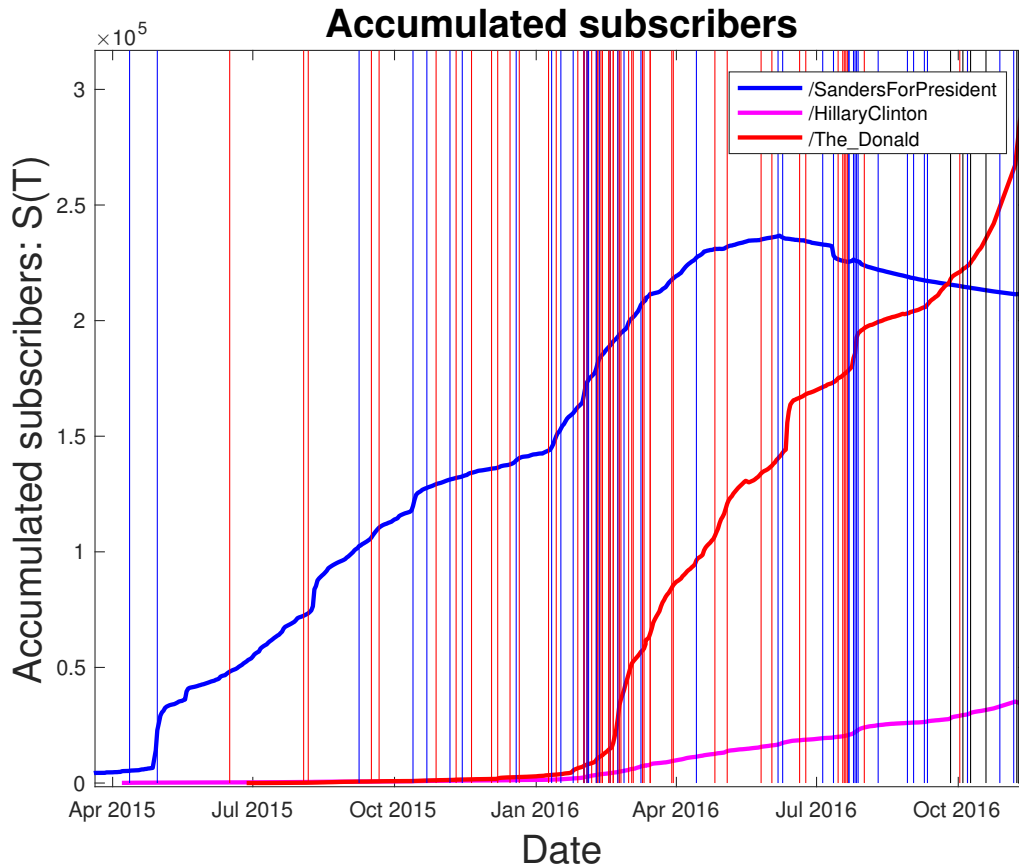


Fig. 1: Accumulated subscribers $S(T)$ per subreddit. The blue line shows the total number of subscribers for /SandersForPresident, the red line for /The_Donald and the pink line for /HillaryClinton. The vertical lines indicate relevant news events. Blue lines represent events mostly related to the Democratic candidates, and red lines to their Republican counterparts. Black lines are events that were classified as affecting both equally. The vertical axis is given in scientific notation.

The first period, between May 2015 and January 4th, 2016, is characterised by /HillaryClinton and /The_Donald being stagnant at low subscription rates and /SandersForPresident growing linearly. The second period, between January 2016 and 20th of February 2016, shows a strong growth of /The_Donald, coupled with /SandersForPresident maintaining an even higher growth rate than before, and /HillaryClinton growing at a much smaller pace. Finally, the third period, starting on the 20th of February 2016, consists of a decline in the growth of /SandersForPresident from its peak value of 236745 subscribers. In the same period /The_Donald overcomes /SandersForPresident between the 24th and 25th of September 2016. /HillaryClinton follows roughly the same ascending trend during the second and third period. Sanders was the most popular candidate among Reddit users up until February 2016, when the total amount of views on The Donald first exceeded those on SandersForPresident (the total amount of subscribers for The Donald only first exceeded those on SandersForPresident in September 2016).

Second, subscriber change can be divided into the same three consecutive periods (see Fig. 2).

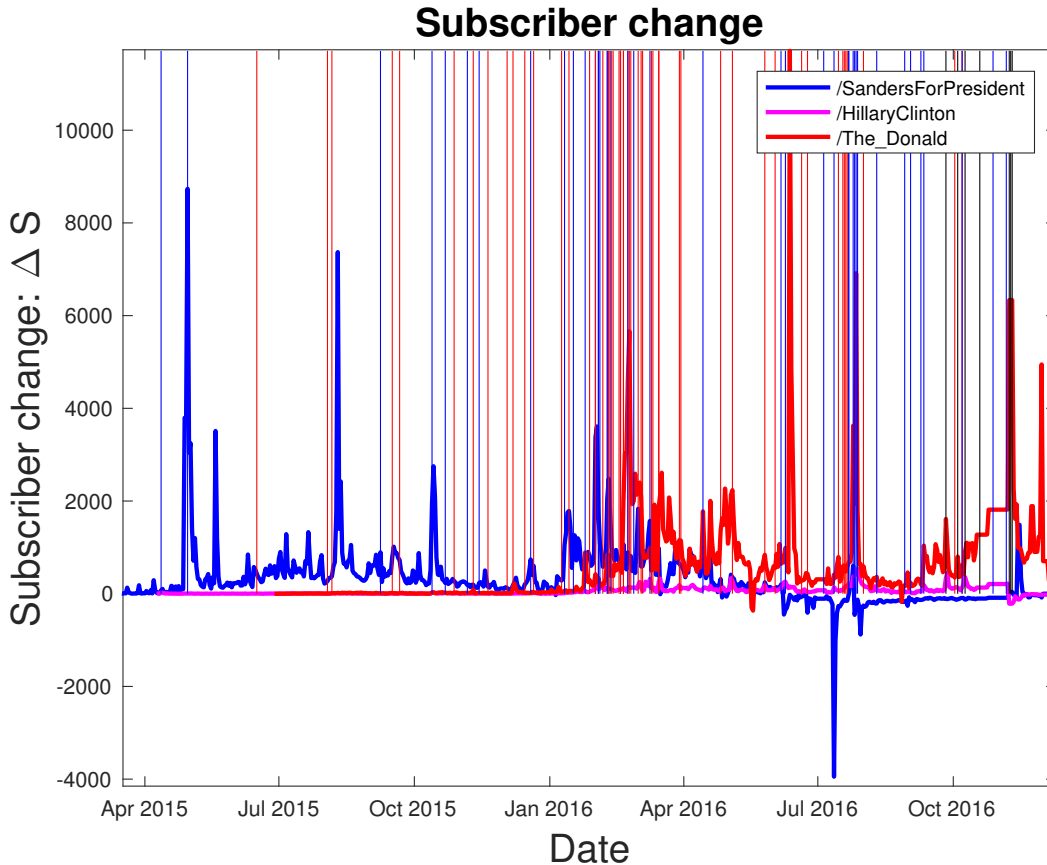


Fig. 2: Subscriber change, ΔS , per subreddit. The blue line shows the subscriber change for /Sander-sForPresident, the red line for /The_Donald and the pink line for /HillaryClinton. The vertical lines show relevant news events. Blue lines represent events mostly related to the Democratic candidates, and red lines to their Republican counterparts. Black lines are events that were classified as affecting both equally.

The average daily subscriber change within the first period was of 16.1 ± 22.6 for /The_Donald, 4.8 ± 4.6 for /HillaryClinton and 186.9 ± 562.8 for /SandersForPresident. It must be noted that the values of the standard deviations are high, indicating strong variability. For the second period, the average daily subscriber change increases for all subreddits: 297.8 ± 342 for /The_Donald, 64.9 ± 48.5 for /HillaryClinton and 1011.2 ± 727.5 for /SandersForPresident. In the third period, we see an abrupt decrease in the average growth of /SandersForPresident, which shows an average of 83.4 ± 447.5 , while the other two subreddits witness an increase. /HillaryClinton has an average subscriber growth of 100.3 ± 95.73 and /The_Donald has an average of 1035.9 ± 1212.1 . In reality, these standard deviations must not be treated under the typical Gaussian assumption, as the subscriber change rarely hits negative values and the variables do not seem to conform to any stable probability distribution. The standard deviation here stands simply as a measure of the sample variance, with no further assumption of any properties of the subscriber change variables. We use the unbiased estimator definition:

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N |\Delta S - \mu|^2}. \quad (4)$$

Where μ is the mean of the sample and N is the number of elements.

To further analyse the influence of new effects in subreddit subscriber change, we performed a linear fit as described in Eqs. (2) and (3). The results show that in all cases, eq. (2) predicted a statistically significant increment in subscriber growth. Concretely, in the case of /The_Donald, the extrapolated average subscriber growth with the presence of news events was 592 ± 90 , compared with 397 ± 53 for the case of non-existent relevant news events. Similarly, /HillaryClinton and /SandersForPresident showed

a clear difference between the presence and absence of news events, with 392 ± 44 compared to 127 ± 18 in the case of `/SandersForPresident`, and 60 ± 7 compared to 37 ± 4 for `/HillaryClinton`. Eq. (3), which considers separately positive and negative news events, explains the data significantly better than eq. (2) for the case of `/SandersForPresident` and `/HillaryClinton`, with respective adjusted R^2 of 0.24 and 0.16 compared to 0.07 and 0.11 from the one parameter linear model. In the case of `/The_Donald`, classifying news events as positive and negative did not change the adjusted R^2 , which was low at a constant 0.07. An alternative way to evaluate the mathematical soundness of this approach is to compare the fit parameters β_+ and β_- and their standard deviations. If the hypothesis of positive and negative influence of new events holds for a candidate, β_- is expected to be smaller than β_+ , as a negative event is expected to add less subscribers to the subscriber count. In Table. 1 we show how `/The_Donald` does not follow this behaviour as the other two subreddits do.

Table 1: Weight of positive and negative news events away from the mean

	<code>/SandersForPresident</code>	<code>/HillaryClinton</code>	<code>/The_Donald</code>
β_+	1474 ± 81	102 ± 9	694 ± 117
β_-	-456 ± 130	8 ± 15	572 ± 177

Fig. 3 shows the total views by subreddit per month, where the three time periods are again apparent.

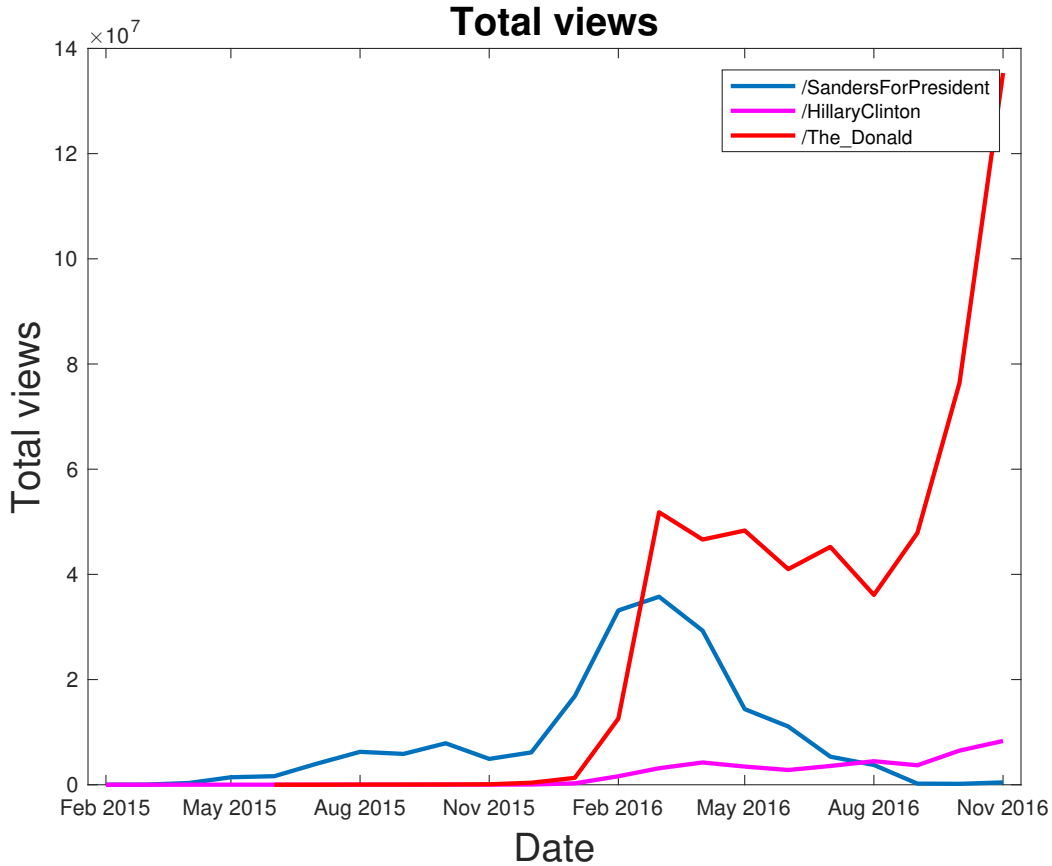


Fig. 3: Total monthly views per subreddit. The blue line shows the total views for `/SandersForPresident`, the red line for `/The_Donald` and the pink line for `/HillaryClinton`. News events are not shown because the data is given per month. The vertical axis is given in scientific notation.

4 Discussion

4.1 General trends

First, Fig. 1 shows that interest in the general election began to pick up steam around February of 2016. This is explained by the fact that the first primary elections, in Iowa and New Hampshire, were held on February 1st and February 9th, respectively. Bernie Sanders' subreddit, however, had been rather popular and growing continuously since April of 2015, when he declared his interest in running for president. /SandersForPresident remained the most popular subreddit of the three until September of 2016, two months after he lost the primary race and endorsed Hillary Clinton.

Second, Table 3 (see 6) shows that the news events described in 2 coincide with the fifteen days with the highest subscriber change in descending order. The most important news events for each candidate's subreddit are: the Orlando shooting for /The_Donald (June 12th, 2016), the third day of the Democratic National Convention for /HillaryClinton (July 28th, 2016), and the announcement of Bernie Sanders' candidacy for /SandersForPresident (April 30th, 2015). Only Bernie Sanders has a negative subscriber growth day in the list, July 7th, 2016, a date that coincides with Sanders' endorsement of Hillary Clinton before the Democratic National Convention. The events that showed a correlation with more than one candidate are marked in blue in the table. Peaks in Sanders' and Trump's subreddits are correlated with the aforementioned Wikileaks revelations of supposedly "rigged" primaries in the case of Sanders, and Trump's Reddit Ask Me Anything-session in the case of Trump. Trump's subreddit seems to be more affected by the Democratic convention than by the Republican one, perhaps explained by the fact that Trump was fairly dominant during the primaries, resulting relatively in low levels of interest in the Republican convention. The Democratic primaries were more of a tight contest between Clinton and Sanders, and speculation of a possible contested convention, as well as anger over the DNC email leak scandal, may have peaked Reddit users' interest [54,58].

Third, the data in Fig. 1 and 3 shows an unexpectedly large difference between /HillaryClinton and /The_Donald and /SandersForPresident. Both /The_Donald and /SandersForPresident were very popular compared to /HillaryClinton, particularly after the first primary elections were held in February of 2016. We know that Clinton's support among young voters was lower than Barack Obama's in both 2008 and 2012, with many Millennials preferring Sanders or third party candidates [35]. This lack of enthusiasm was thus reflected in Clinton's lower levels of online community engagement. However, her support among Millennials is not so low as to explain the full width of the subscription and visitor gaps visible in the data. In fact, according to exit polls, Hillary received 55% of the votes from citizens younger than 29, compared to 37% in the case of Donald Trump [35].

Fourth, the data does not show a clear transfer of traffic and subscriptions from /SandersForPresident to /HillaryClinton after Clinton obtained the Democratic Party's nomination, despite the fact that the total number of accumulated subscribers to /SandersForPresident went down after July 2016 (see Fig. 1). This indicates that Clinton did not manage to consolidate support among Democratic-leaning Reddit users after it became clear that she was going to be the nominee. This is not in line with polling data showing that Clinton enjoyed a typical amount of support from likely Democratic voters (of any age) [30,25]. For example, support for Clinton among Sanders supporters was higher in 2016 than support for Barack Obama was among Clinton supporters in 2008 (after Obama had won the nomination over Clinton; see [61]).

We propose three possible explanations for the observations mentioned above: skewed demographics, differences in subreddit cultures, and a decline in party affiliation paired with increased partisanship.

Skewed demographics The hypothesis of skewed demographics has been reviewed in section 2. The underrepresentation of minority groups on Reddit lends credence to the idea that Reddit would be more supportive of Trump than Clinton (compared to the average Millennial population), since Clinton drew more support from minority communities and women than Trump. Additionally, this hypothesis could help explain the lack of popularity of /HillaryClinton, since young women are underrepresented on Reddit. However, this sample skewing is not sufficient to explain the more than sevenfold difference between subscribers to Donald Trump's subreddit and subscribers to Hillary Clinton's. In fact, even if we only look at white male Millennials, Hillary Clinton would be expected to have roughly the same demographic support as Donald Trump [46]. Therefore, the skewed demographics hypothesis is not sufficient to explain the lack of support for Clinton on Reddit.

Differences in subreddit culture Internet fora are known to develop unique subcultures [47]. From its inception, /SandersForPresident was characterised by an energetic spirit, mixing ideas such as universal health care and raising taxes on the rich with an additional focus on social issues such as racial discrimination and LGBT rights [44]. The general tone of the discussion was optimistic; users created posts with personal stories of how they further contributed to the cause, and asked other users to match them in their efforts. This created a culture among /SandersForPresident subscribers where contributing to Sanders' political cause was encouraged [53,51].

Donald Trump's subreddit showed a markedly different culture. It is unclear if the forum was initially conceived ironically, or as a serious support group for Trump. Whichever is the case, the general tone was humorous and playful, with a penchant for criticising more establishment-friendly Republican and Democratic primary candidates. Many /The_Donald posts made fun of the attitude present in /SandersForPresident, for example: *I have just donated \$0 to Donald Trump's campaign, match me if you can!* [3].

On /The_Donald, users posted memes, jokes, video mashups and Trump quotes. These tropes grew in popularity with time. This may have attracted passive users that were not overt Trump supporters [27]. As we mentioned in 2.4, however, this is somewhat unlikely, since /The_Donald does not allow non-Trump supporters to post on the forum. Thus, /The_Donald's popularity is most likely best explained by the fact that it managed to become an attractive place for not only Trump supporters, but also the politically disaffected or disengaged, who were drawn in by the subreddit's humorous bent.

/HillaryClinton, which had fewer subscribers than the other two subreddits, generally struck a more policy-focused tone, with some attempts at replicating the mischievous nature of /The_Donald. An example of this is a post by the user name "Deceptiveideas" in June 2016 titled "Hillary engaging in the meme war" followed by a photo of an exasperated Hillary Clinton in front of the text "Q: *Saying [a judge] can't do his job because of his race - is that not the definition of racism? Trump: No.*" [20].

In general, the mood among /HillaryClinton users was mostly optimistic in nature, due in part to her good standing in many national election polls. As a social platform of support, however, /HillaryClinton had an insignificant amount of subscribers compared to other online support groups such as the "secret" pro-Clinton Facebook group PantsuitNation, which as of February 2017 had accumulated over 2.5 million users [18]. It is therefore likely that Clinton's supporters did not choose Reddit as their main avenue of expression, but instead preferred other online channels. In addition, /HillaryClinton did not see a spike in accumulated subscribers after /SandersForPresident's numbers began to drop after July 2016. This implies that /HillaryClinton did not absorb a significant number of Sanders supporters after Sanders dropped out of the race.

As the primary elections advanced, the sentiment in /The_Donald and /SandersForPresident subreddits evolved, with the former denying its candidates primary loss against Clinton and, following Wikileaks revelations that appeared to indicate that the Democratic National Committee (DNC) had been biased against Sanders' candidacy, a sense of betrayal towards Clinton and the DNC. Meanwhile, /The_Donald started a campaign to assimilate /SandersForPresident supporters. One post, published in March 2016, "The Reformed BernieBot's Guide on Assimilation", outlines the steps that Bernie supporters must follow in order to integrate into the r/The_Donald community [4]. Another post from May 2016 was titled "RIP BERNIE — CONVERT NOW" and featured an image portraying Donald Trump as the priest of the Disciples of Christ-church, blessing a coffin where an image of Bernie Sanders was photoshopped into [40]. The most upvoted comment on this particular thread was a direct quote by Donald Trump: *"Bernie Sanders is being treated very badly by the Democrats, the system is rigged against him"*.

Thus, the subreddit culture explanation for the observed trends states that users were more attracted to /The_Donald and /SandersForPresident due to their more humorous, ironic, or anti-establishment nature. The content of the posts in these subreddits was perhaps simply considered funnier or of better quality than in /HillaryClinton, and therefore was more successful at constructing a group identity and group cohesion [51]. The same goes for Clinton support groups outside of Reddit, such as the aforementioned PantSuitNation. We posit that /HillaryClinton remained relatively unpopular due in part to its excessive focus on 'boring' or serious topics. Content analysis should reveal more about this proposed negative relation between online political community popularity and the level of focus on policy.

Furthermore, the attempts undertaken by /The_Donald to assimilate Sanders supporters can be seen as a way to overcome /SandersForPresident's group identity by pointing to shared experiences and a common adversary [51]. This, in turn, affected group cohesion, as /SandersForPresident began to abandon the group. This is particularly noteworthy considering that Trump and Sanders have little in common politically, which is a strong indicator that online political communities do not see politics and policy as a primary driver for constructing and maintaining group identity. As such, these assimilation attempts are

in line with what Tom Nichols has dubbed the "death of expertise": the authority of experts, especially when it comes to political topics, is becoming less and less relevant. This development, he claims, occurred in tandem with the rise of the internet and the ubiquitous availability of media resources [48]. We put forward that policy and political ideology were not as relevant for group identity and cohesion as entertainment value. Nichols' 'death of expertise' thus expressed itself as a switch in persuasion tactics from a policy-based to an emotions-based approach.

Partisanship and declining party affiliation There are several factors pointing at the possibility that Millennials may have voted in a higher proportion for Donald Trump than was indicated by election polls. This, again, would not suffice to explain the sevenfold difference in subscribers between Trump's and Clinton's redds, but in conjunction with the aforementioned skewed demographics may have had a non-zero effect. This is in line with observations by Twenge et al. [62], who argue that the Millennial generation is more positively inclined towards conservative ideas than the Boomer and Gen X-generations were at the same age. This may have lowered the threshold to go from voting Democratic to voting Republican.

The relatively large number of Millennials who do not identify as Democrats forms part of the explanation for the popularity of /The_Donald. Furthermore, /SandersForPresident's popularity does not seem to have translated into loyalty to the Democratic Party or Hillary Clinton after Sanders' primary loss. One particular day in the data set is indicative of this. On July 12th, 2016, Sanders officially endorsed Clinton during a rally they both attended in Portsmouth, New Hampshire. On this day, /SandersForPresident lost 3947 subscribers, indicating a degree of disappointment among Sanders supporters for his decision to endorse Clinton. On the same day, /The_Donald gained 411 subscribers, and /HillaryClinton gained 136. Thus, it is safe to say that there was no transfer from /SandersForPresident to /HillaryClinton to speak of, and a slightly larger but still quite small transfer to /The_Donald.

It is therefore possible that a significant portion of /SandersForPresident subscribers did not, by and large, base their subreddit subscriptions on party identification or shared viewpoints on political issues [57,1,17]. This phenomenon can be explained by looking at trends of partisanship versus party affiliation. Julia Azari argues that the influence of party organisations and the interest groups that constitute parties (such as the NRA, the US Chamber of Commerce or environmental groups) has waned, both because voters do not have to listen to elites and because elites do not have to listen to each other [6]. This opens up avenues for non-traditional candidates such as Donald Trump and Bernie Sanders to do unexpectedly well in electoral cycles, in apparent violation of the idea that 'the Party decides' who is an acceptable (enough) candidate to win the party's primary elections. [22,15,60]. At the same time, partisanship is strong, as witnessed by increased congressional as well as societal polarisation and growing self-identification at the outer edges of the political spectrum. Thus, parties can no longer control whom they nominate, but voters and elites are prepared to support the eventual candidate. This is especially true for those who get the majority of their news from online sources; as [49] argue, consumers of internet news sources are likely to hold more extreme political views, while being interested in a wider range of political topics, than those who mostly watch TV or read newspapers. Reddit users can be seen to fit within this framework.

However, Azari's idea predicts more consolidation in favour of Clinton after she clinched the nomination than the data is able to show. Instead, we see almost no increase in subscriptions and page visits for Clinton's subreddit, despite a persistent decline in Sanders' after his loss. We therefore propose that it is important to take into account how negative news events affected subscription trends for both Clinton and Sanders. We explore this in the next section.

4.2 News event dynamics

The analysis of the relation between news events and Reddit activity shown in Sec. 3, adds weight to the hypothesis that subreddit subscriber count relates to specific news events rather than being mostly determined by common activity trends. Only one event shows up as significant for all candidates: the Democratic National Convention, which was expected to have an important effect on voter enthusiasm, for example by way of a convention bounce [56].

Generally speaking, most of the traffic and subscription spikes correspond to news events expected to be of interest during a political campaign (such as primary elections, inter- and intra-party debates, and party conventions), and 'Ask me Anything'-events in the case of Sanders and Trump. Three data points warrant special mention. First, 12 June 2016, the day of the Orlando mass shooting, was the day with the largest increase in subscribers on /The_Donald. Second, /SandersForPresident witnessed a

traffic spike on August 8, 2015. this was the day a Sanders rally was interrupted by Black Lives Matter activists. And finally, one data point appears to be absent: on October 8, 2016, a tape was leaked of Donald Trump making controversial comments about women. This scandal dominated the news cycle for a number of days, and one would expect a spike in /The_Donald’s traffic or subscriptions because of it. However, this did not occur: /The_Donald grew by a total of 1337 subscribers between October 8 and 9, and /HillaryClinton by a total of 566. This corresponds to a small local bump for both candidates. A content-analytical approach of Reddit posts on these days may reveal some insights about this ostensible paradox.

Finally, we will make a few broader observations about the relation between subscription dynamics and news events. As explained in 3, our linear fit model found in 3 shows that the impact of positive and negative news events on subscriber and traffic dynamics are significant for /HillaryClinton and /SandersForPresident, but not for /The_Donald (see also table 1). Broadly speaking, albeit with significant statistical caveats (explained in section 3), our observations support the idea that Donald Trump’s candidacy was not adversely affected by negative news events in the same way as other candidates. Whereas Hillary Clinton and especially Bernie Sanders experienced significant Reddit traffic and subscription dips during or briefly after some negative news events, this association does not exist for Donald Trump. Our observations are therefore in line with the idea that Trump’s candidacy was helped by his continuous domination of the news cycle and social media activity [32]. His presence online allowed him to run a low-cost campaign that relied on controversy to generate support [59]. This finding is also corroborated by Effing, Van Hillegersberg and Huibers [24], who argue that increased social media participation by politicians can increase preference voting. Additionally, there is the possibility that Trump’s supporters on /The_Donald simply did not care about news events perceived as negative to Trump’s campaign by what they saw as the “mainstream media” [9,45]. The perception that many media outlets were biased against Trump may have led to a sense that whatever report came out about his candidacy either did not matter, was either taken out of context or just as bad as other candidates’ campaign hiccups. /The_Donald developed an immunity against negative news that /HillaryClinton and /SandersForPresident did not possess. Further research into /The_Donald’s group dynamics and identity could shed light onto how this immunity emerged. We posit that looking at the content of posts on /The_Donald within the context of positive and negative news events could offer insight into how this immunity came to be.

5 Conclusion

In this paper, we have sought to understand social media dynamics in the 2016 US presidential election, particularly concerning Millennial voters, by measuring Reddit metrics and election-related news events. From the analysis of subscriber trends for Trump’s, Clinton’s and Sanders’ subreddits, no consolidation of support was observed for Hillary Clinton after her victory in the primaries. Additionally, Hillary Clinton had much less Reddit support than Donald Trump. We show how this deficit is unlikely to be caused by skewed demographics in Reddit, as the skewing is much smaller than the recorded difference in subscribers. We discuss three possible explanations of the observed phenomena: the difference of the quality of content produced on /HillaryClinton, /The_Donald, and /SandersForPresident; the efforts by /The_Donald subscribers to assimilate Sanders supporters not through policy but through emotional appeals; and high partisanship paired with low party identification in the United States. We suggest that a combination of these factors three explains candidates’ different levels of popularity on Reddit.

The correlation between news events during the electoral cycle and subscriber change on Reddit was also studied. We observe that positive and negative news events explains our observations better than just news events in general in the case of Sanders’ and Clinton’s subreddits. However, this was not the case for Donald Trump, whose subreddit seems to be affected positively by both positive and negative news events. In fact, one of the elections’ major scandals, Trump’s *Grab ’em by the pussy*-comments, did not seem to affect the number of /The_Donald subscribers in any significant way. Our findings thus provide support for Trump’s remark that he could “stand in the middle of Fifth Avenue and shoot somebody and not lose any voters” [21,24].

References

1. AELST, P. V., AND WALGRAVE, S. New media, new movements? The role of the internet in shaping the antiglobalization' movement. *Information, Communication and Society* 5, 4 (2002), 465–493.
2. ALASHRI, S., KANDALA, S. S., BAJAJ, V., RAVI, R., SMITH, K. L., AND DESOUZA, K. C. An analysis of sentiments on facebook during the 2016 U.S. presidential election. *Proceedings of the 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2016* (2016), 795–802.
3. ANONYMOUS. FOR EVERY UPVOTE I WILL DONATE EXACTLY \$0 TO DONALD TRUMP'S CAMPAIGN!!! BE HIGH ENERGY!!!! MAKE REDDIT GREAT AGAIN!!!! *Reddit.com February, 2* (2016).
4. ANONYMOUS. The Reformed BernieBot's Guide on Assimilation! *Reddit.com* (2016).
5. AOL.COM. Timeline: election. *AOL* (2016).
6. AZARI, J. R. Weak parties and strong partisanship are a bad combination. *Vox.com November, 3* (2016).
7. BALL, M. Why Hillary Clinton Lost. *The Atlantic Nov 15* (2016).
8. BARTHEL, M. How the 2016 presidential campaign is being discussed on Reddit. *Pew research center: Factank May, 26* (2016).
9. BATEMAN, O. L. Trump Supporters Were Hiding In Plain Sight In Online Comment Sections. *Vice.com November, 14* (nov 2016).
10. CENTER, P. R. Political Polarization & Media Habits: Ideological Placement of Each Sources Audience. *Journalism.org October, 20* (2014).
11. CHOMA, B. L., AND HANOCH, Y. Cognitive ability and authoritarianism: Understanding support for Trump and Clinton. *Personality and Individual Differences* 106 (2017), 287–291.
12. CHOU, L.-C., AND FU, C.-Y. The influence of Internet on politics: the impact of Facebook and the Internet penetration on elections in Taiwan. *Applied Economics Letters* 24, 7 (2017), 494–497.
13. CILLIZZA, C. One of Hillary Clinton's top aides nailed exactly why she lost. *The Washington Post November, 1* (2016).
14. CIRCLE. At Least 80 Electoral Votes Depended on Youth. *The Center For Information & Research On Civic Learning and Engagement November, 7* (2012).
15. COHEN, M., MCGRATH, M. C., ARONOW, P., AND ZALLER, J. Ideologically Extreme Candidates in U.S. Presidential Elections, 1948-2012. *The ANNALS of the American Academy of Political and Social Science* 667, 1 (2016), 126–142.
16. COLBY, S. L., AND ORTMAN, J. M. Projections of the Size and Composition of the U.S. Population: 2014 to 2060. *US Census: Population Estimates and Projections March* (2015), 10.
17. CONWAY, J. M. Anti-Globalization Movements. *The Wiley Blackwell Encyclopedia of Gender and Sexuality Studies* (2016).
18. CORREAL, A. Pantsuit Nation, a Secret' Facebook Hub, Celebrates Clinton. *The New York Times November, 8* (2016).
19. D. SANDEFUR, G., MATRIN, M., EGGERLING-BOECK, J., E. MANNON, S., AND M. MEIER, A. 3: An Overview of Racial and Ethnic Demographic Trends. In *America Becoming: Racial Trends and Their Consequences, Volume 1*. The National Academic Press, 2001, p. 48.
20. DECEPTIVEIDEAS. Hillary engaging in the meme war. *Reddit.com* (2016).
21. DIAMOND, J. Trump: I could 'shoot somebody and I wouldn't lose voters'. *CNN January, 24* (2016).
22. DOWDLE, A. The Party Decides: Presidential Nominations Before and After Reform by Marty Cohen, David Karol, Hanes Noel, and John Zaller. *Political Science Quarterly* 124, 3 (2009), 550–551.
23. DUGGAN, M., AND SMITH, A. 6% of Online Adults are reddit Users. *Pew Research Center July, 3* (2013).
24. EFFING, R., VAN HILLEGERSBERG, J., AND HUIBERS, T. Social media and political participation: Are Facebook, Twitter and YouTube democratizing our political systems? *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 6847 LNCS (2011), 25–35.
25. ENTEN, H. GOP Voters Are Rallying Behind Trump As If He Were Any Other Candidate. *fivethirtyeight June, 1* (2016).
26. GAMBINO, L., AND PANKHANIA, M. How we got here: a complete timeline of 2016's historic US election. *The Guardian November, 8* (2016).
27. GILMAN, H. R., AND STOKES, E. The Civic and Political Participation of Millennials. *New America* (2014), 57–60.
28. GROSHEK, J., AND KOC-MICHALSKA, K. Helping populism win? Social media use, filter bubbles, and support for populist presidential candidates in the 2016 US election campaign. *Information Communication and Society* 20, 9 (2017).
29. HU, G., KODALI, S., AND PADAMATI, A. Sentiment analysis of tweets on 2016 US presidential election candidates. In *29th International Conference on Computer Applications in Industry and Engineering, CAINE 2016* (2016), pp. 219–226.
30. HUFFPOSTPOLLSTER. Custom Chart: 2016 General Election: Trump vs. Clinton. *The Huffington Post* (2016).
31. HURRICANEXRIKS. Results Of The Reddit Demographics Survey 2016. *Reddit.com October, 11* (2016).

32. J. BOCZKOWSKI, P. Has Election 2016 been a turning point for the influence of the news media? *Nieman Labs November, 8* (2016).
33. JUNGHERR, A. Twitter use in election campaigns: A systematic literature review. *Journal of Information Technology and Politics 13*, 1 (2016), 72–91.
34. KEETER, S., TYSON, J., AND HOROWITZ, A. Young Voters in the 2008 Election. *Pew Research Center November, 1* (2008).
35. KHALID, A., AND ROSE, J. Millennials Just Didn't Love Hillary Clinton The Way They Loved Barack Obama. *NPR November, 1* (2016).
36. KURTZLEBEN, D. The Most 'Unprecedented' Election Ever? 65 Ways It Has Been. *NPR July, 3* (2016).
37. LEE, J., AND LIM, Y.-S. Gendered campaign tweets: The cases of Hillary Clinton and Donald Trump. *Public Relations Review 42*, 5 (2016), 849–855.
38. LIU, B. *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers, San Rafael, CA, 2012.
39. LONG, R. Why did Hillary Clinton lose? Simple. She ran a bad campaign. *Chicago Tribune November, 1* (2016).
40. LOREZMASTER. RIP BERNIE — CONVERT NOW. *Reddit.com June 7* (2016).
41. MALDONADO, M., AND SIERRA, V. Twitter predicting the 2012 US presidential election?: Lessons learned from an unconscious value co-creation platform. *Journal of Organizational and End User Computing 28*, 3 (2016), 10–30.
42. MARTIN, T. Interactive Map of Reddit and Subreddit Similarity Calculator. *Short Tails November, 28* (2016).
43. MIROWSKI, T., ROYCHOUDHURY, S., ZHOU, F., AND OBRADOVIC, Z. Predicting poll trends using twitter and multivariate time-series classification. *Social Informatics: 8th International Conference 10046 LNCS* (2016), 273–289.
44. MONTELLARO, Z. Bernie Sanders Is Dominating the Reddit Primary. *The Atlantic October, 5* (2015).
45. MORIN, R. Trump says social media was the key to victory. *Politico November, 12* (2016).
46. MOSENDZ, P. What This Election Taught Us About Millennial Voters. *Bloomberg November, 9* (2016).
47. MUGGLETON, D., AND WEINZIERL, R. *The Post-Subcultures Reader*, 1st ed. Berg Publishers, Oxford, 2003.
48. NICHOLS, T. *The Death of Expertise: The Campaign against Established Knowledge, and Why It Matters*. Oxford University Press, Oxford, 2017.
49. NIE, N. H., MILLER III, D. W., GOLDE, S., BUTLER, D. M., AND WINNEG, K. The World Wide Web and the U.S. Political News Market. *American Journal of Political Science 54*, 2 (2010), 428–439.
50. PANCER, E., AND POOLE, M. The popularity and virality of political social media: hashtags, mentions, and links predict likes and retweets of 2016 U.S. presidential nominees' tweets. *Social Influence 11*, 4 (2016), 259–270.
51. PANCONESI, G., AND GUIDA, M. *Handbook of Research on Collaborative Teaching Practice in Virtual Learning Environments*. IGI Global, Hershey, PA, 2017.
52. PINGDOM. Report: Social network demographics in 2012. *Pingdom.com August, 21* (2012).
53. POLLACK, L. How Reddit has boosted Bernie Sanders' US campaign. *Financial Times March, 29* (2016).
54. REILLY, K. Bernie Sanders: Democratic Convention Will Be Contested. *Time Magazine June, 4* (2016).
55. ROTHMUND, T., AND OTTO, L. The changing role of media use in political participation. *Journal of Media Psychology 28*, 3 (2016), 97–99.
56. SAAD, L. Average Convention "Bounce" Since 1964 Is Six Points. *Gallup News Service July, 26* (2000).
57. SAULL, R. Capitalism, crisis and the far-right in the neoliberal era. *Journal of International Relations and Development 18*, 1 (2015), 25–51.
58. SIDDIQUI, S., GAMBINO, L., AND ROBERTS, D. DNC apologizes to Bernie Sanders amid convention chaos in wake of email leak. *The Guardian July, 25* (2016).
59. SILVER, N. How Trump Hacked the Media. *Fivethirtyeight March, 30* (2016).
60. STEGER, W. P. Polls and Elections: Two Paradigms of Presidential Nominations. *Presidential Studies Quarterly 43*, 2 (2013), 377–387.
61. SUMNER, M. Sanders voters moving to Clinton much faster than Clinton voters moved to Obama. *Daily Kos June, 27* (2016).
62. TWENGE, J. M., HONEYCUTT, N., PRISLIN, R., AND SHERMAN, R. A. More Polarized but More Independent: Political Party Identification and Ideological Self-Categorization Among U.S. Adults, College Students, and Late Adolescents, 1970-2015. *Personality and Social Psychology Bulletin 42*, 10 (2016), 1364–1383.
63. UNITED STATES CENSUS BUREAU. United Status Census: Quick Facts. *Census.gov* (2017).
64. WIKIPEDIA. United States presidential election, 2016 timeline, 2016.

6 Appendix

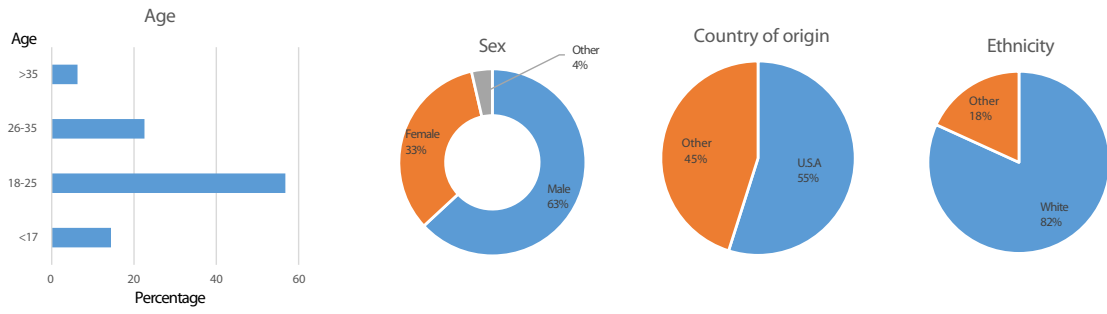
Table 2 contains the list of news events used in this paper and their classification. The relevant variables are set in columns, starting by labels which we describe as follows:

- **Year, Month, Day**: The date of the news event.
- **Label**: A label describing the news event(s) in the given day.
- **Party**: A numerical label used to draw the vertical lines in the figures. If equal to 0, represents a new event mostly relevant for the Republican party, if equal to 1, it represents a news event relevant for the Democratic party. If equal to 2 it represents a news event equally relevant for both.
- **PositiveH**: A numerical label used to classify news events positive to Hillary Clinton’s campaign as described in Sec. 2. If equal to 1, the event is considered as positive. If equal to 0 as neutral.
- **NegativeH**: A numerical label used to classify news events negative to Hillary Clinton’s campaign as described in Sec. 2. If equal to 1, the event is considered as negative. If equal to 0 as neutral.
- **PositiveT**: A numerical label used to classify news events positive to Donald Trump’s campaign as described in Sec. 2. If equal to 1, the event is considered as positive. If equal to 0 as neutral.
- **NegativeT**: A numerical label used to classify news events negative to Donald Trump’s campaign as described in Sec. 2. If equal to 1, the event is considered as negative. If equal to 0 as neutral.
- **PositiveS**: A numerical label used to classify news events positive to Bernie Sanders campaign as described in Sec. 2. If equal to 1, the event is considered as positive. If equal to 0 as neutral.
- **NegativeS**: A numerical label used to classify news events negative to Bernie Sanders campaign as described in Sec. 2. If equal to 1, the event is considered as negative. If equal to 0 as neutral.

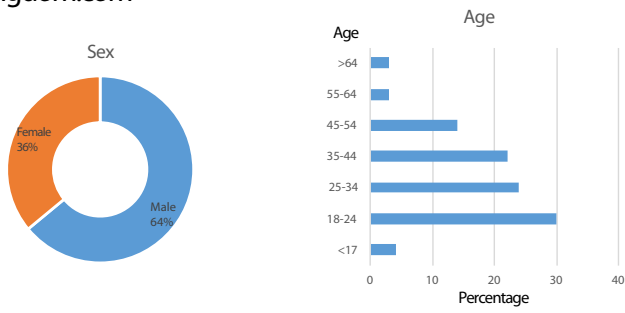
Table 2: News events considered for the analysis of subscriber and visitor trends.

Year	Month	Day	Label	Party	PositiveH	NegativeH	PositiveT	NegativeT	positiveS	negativeS	comparison	verdict	
2015	4	12	Hillary Clinton formally announces her candidacy for the presidential nomination of the Democratic Party	1	1	0	0	0	0	0	different	human	
2015	4	30	Bernie Sanders, of Vermont, formally announces his candidacy for the presidential nomination of the Democratic Party	1	0	0	0	0	1	0	different	human	
2015	6	16	Donald Trump, of New York, officially declares his candidacy for the presidential nomination of the Republican Party	0	0	0	1	0	0	0	same	same	
2015	8	3	First presidential forum, featuring 14 Republican candidates, was broadcast on C-SPAN	0	0	0	0	1	0	0	different	algorithm	
2015	8	6	First official presidential debate, featuring 10 Republican candidates, is held in Cleveland, Ohio	0	0	0	1	0	0	0	different	human	
2015	9	8	Hillary Clinton apologises for her emails	1	0	1	0	0	0	0	same	same	
2015	9	16	Second Republican debate is held in Simi Valley, California	0	0	0	1	0	0	0	same	same	
2015	10	13	First Democratic debate	1	1	0	0	0	0	1	0	same	same
2015	10	22	Hillary Clinton's hearing on Benghazi	1	0	1	1	0	0	0	different	algorithm	
2015	10	28	Third Republican debate is held in Boulder	0	0	0	0	0	0	0	different	algorithm	
2015	11	10	Fourth Republican debate is held in Milwaukee, Wisconsin	0	0	0	0	0	0	0	different	algorithm	
2015	11	14	Second Democratic debate is held in Des Moines, Iowa	1	1	0	0	0	1	0	same	same	
2015	12	3	The Republican Jewish Coalition Presidential Candidates	0	0	0	1	0	0	0	same	same	
2015	12	7	Trump calls for Muslim ban	0	0	0	0	0	1	0	different	algorithm	
2015	12	15	Fifth Republican debate is held in Las Vegas	0	0	0	1	0	0	0	same	same	
2015	12	19	Third Democratic debate is held in Manchester	1	1	0	0	0	1	0	same	same	
2016	1	9	The Republican Kemp Forum is held in Columbia	0	0	0	0	0	0	0	different	algorithm	
2016	1	11	Third Democratic forum is held in Des Moines	1	1	0	0	0	1	0	same	same	
2016	1	14	Sixth Republican debate is held in North Charleston	0	0	0	1	0	0	0	different	human	
2016	1	17	Fourth Democratic debate is held in Charleston, South Carolina	1	1	0	0	0	1	0	same	same	
2016	1	25	A Democratic forum, a Town Hall event, is held in Des Moines	0	0	0	0	0	1	0	same	same	
2016	1	28	Seventh Republican debate is held in Des Moines, Iowa. Trump does not participate	0	0	0	0	1	0	0	same	same	
2016	2	1	The Iowa Democratic caucus is won by Hillary Clinton	1	1	0	0	0	1	0	same	same	
2016	2	1	The Iowa Republican caucus is won by Ted Cruz	0	0	0	0	1	0	0	different	same	
2016	2	3	A Democratic Town Hall forum event is held in Derry, New Hampshire	1	1	0	0	0	1	0	same	same	
2016	2	4	Fifth Democratic debate is held in Durham	1	1	0	0	0	1	0	same	same	
2016	2	6	Eighth Republican debate is held in Manchester	0	0	0	0	0	0	0	different*	algorithm	
2016	2	9	The New Hampshire Republican primary is won by Donald Trump	1	1	0	0	0	1	0	same	same	
2016	2	9	The New Hampshire Democratic primary is won by Bernie Sanders	1	0	1	0	0	1	0	same	same	
2016	2	11	Sixth Democratic debate is held in Milwaukee	1	1	0	0	0	1	0	different*	human	
2016	2	13	Ninth Republican debate is held in Charleston, South Carolina	0	0	0	0	0	0	0	different	algorithm	
2016	2	17	CNN Republican town halls are held in Greenville	0	0	0	0	1	0	0	nodata	human	
2016	2	18	CNN Republican town halls are held in Greenville	0	0	0	0	1	0	0	nodata	human	
2016	2	18	Democratic Town Hall forum event is held in Las Vegas	1	1	0	0	0	1	0	nodata	human	
2016	2	18	Trump fights the Pope	0	0	0	0	0	0	0	different	algorithm	
2016	2	20	Trump wins South Carolina	0	0	0	1	0	0	0	same	same	
2016	2	23	Nevada Republican caucuses are won by Donald Trump	0	0	0	1	0	0	0	same	same	
2016	2	23	CNN Democratic town hall is held in Columbia	1	1	0	0	0	1	0	same	same	
2016	2	24	Republican town hall is held in Houston	0	0	0	1	0	0	0	nodata	human	
2016	2	25	10th Republican debate is held in Houston	0	0	0	0	0	0	0	different	algorithm	
2016	2	27	South Carolina Democratic primary is won by Hillary Clinton	1	1	0	0	0	0	1	same	same	
2016	3	1	Super Tuesday, big win for Trump	0	0	0	1	0	0	0	same	same	
2016	3	3	11th Republican debate. Rubio jokes about Trump's tiny penis	0	0	0	0	1	0	0	same	same	
2016	3	9	Eighth and final Democratic debate is held in Miami	1	1	0	0	0	1	0	same	same	
2016	3	10	Twelfth Republican debate is held in Miami, Florida	0	0	0	1	0	0	0	same	same	
2016	3	11	Corey Lewandowski (Trump's campaign manager) accused of assault	0	0	0	0	1	0	0	different	human	
2016	3	15	Trump wins Florida, Illinois, Missouri, North Carolina. Kasich wins Ohio	0	0	0	1	0	0	0	same	same	
2016	3	29	Republican town hall	0	0	0	0	0	0	1	same	same	
2016	3	30	Trump says women should be punished for having abortions	0	0	0	0	1	0	0	different	human	
2016	4	14	Ninth Democratic debate is held in Brooklyn	1	1	0	0	0	1	0	same	same	
2016	4	26	Trump wins Connecticut, Delaware, Maryland, Rhode Island, Pennsylvania	0	0	0	1	0	0	0	same	same	
2016	5	26	Donald Trump passes 1237 pledged delegates	0	0	0	1	0	0	0	different*	human	
2016	6	2	Paul Ryan says hell vote for Trump	0	0	0	1	0	0	0	different	human	
2016	6	6	Hillary Clinton passes 2383 pledged delegates	1	1	0	0	0	0	1	same	same	
2016	6	9	Obama endorses Clinton	0	1	0	0	0	0	1	same	same	
2016	6	20	Trump fires Corey Lewandowski	0	0	0	0	0	0	0	different	algorithm	
2016	6	24	Trump goes to Scotland the day after Brexit vote	0	0	0	0	0	0	0	different	algorithm	
2016	7	5	FBI director James Comey recommends not indicting Clinton	1	1	0	0	0	0	0	same	same	
2016	7	12	Bernie Sanders endorses Hillary Clinton	1	1	0	0	0	0	1	same	same	
2016	7	15	Republican presumptive nominee Donald Trump announces Indiana governor Mike Pence as his vice presidential running mate	0	0	0	1	0	0	0	different	human	
2016	7	18	Republican National Convention is held in Cleveland, Ohio. Donald Trump and Mike Pence are formally nominated for President	0	0	0	0	0	0	0	different	algorithm	
2016	7	19	Republican National Convention is held in Cleveland, Ohio. Donald Trump and Mike Pence are formally nominated for President	0	0	0	1	0	0	0	different	human	
2016	7	20	Republican National Convention is held in Cleveland, Ohio. Donald Trump and Mike Pence are formally nominated for President	0	0	0	1	0	0	0	same	same	
2016	7	21	Republican National Convention is held in Cleveland, Ohio. Donald Trump and Mike Pence are formally nominated for President	0	0	0	1	0	0	0	same	same	
2016	7	22	Democratic presumptive nominee Hillary Clinton announces Tim Kaine as her vicepresidential candidate	1	1	0	0	0	0	0	same	same	
2016	7	22	Wikileaks leaks 20,000 emails from the Democratic National Committee, revealing a systematic bias against Bernie Sanders.	1	0	1	0	0	0	1	0	same	same
2016	7	25	Democratic National Convention is held in Philadelphia	1	1	0	0	0	0	1	same	same	
2016	7	26	Democratic National Convention is held in Philadelphia, Pennsylvania.	1	1	0	0	0	0	1	same	same	
2016	7	27	Democratic National Convention is held in Philadelphia, Pennsylvania.	1	1	0	0	0	0	1	same	same	
2016	7	28	Democratic National Convention is held in Philadelphia, Pennsylvania.	1	1	0	0	0	0	1	same	same	
2016	8	1	Trump feuds with Khizr Khan, parent of dead soldier	0	1	0	0	1	0	0	same	same	
2016	8	10	Leaked emails show Clinton Foundations ties with State Department	1	0	1	1	0	0	0	different	human	
2016	8	29	New sexting scandal with Anthony Weiner & Huma Abedin	1	0	1	1	0	0	0	different	human	
2016	9	2	FBI releases more Clinton emails. Trump lashes onto them	1	0	1	0	0	0	0	same	same	
2016	9	9	Clinton calls half of Trump supporters deplorable	1	0	1	1	0	0	0	same	same	
2016	9	11	Clinton faints	1	0	1	1	0	0	0	same	same	
2016	9	26	First presidential general election debate was held at Hofstra University in Hempstead, New York	2	1	0	0	1	0	0	same	same	
2016	10	2	NYT report suggests Trump paid no federal income tax for 20 years	0	1	0	0	1	0	0	different	algorithm	
2016	10	4	Only vice presidential general election debate was held at Longwood University in Farmville, Virginia.	2	0	1	1	0	0	0	different*	human	
2016	10	7	Tapes are leaked out from Access Hollywood showing Donald Trump and Billy Bush bragging about sexual exploits in 2005	0	1	0	0	1	0	0	same	same	
2016	10	7	Hillary Clinton's transcripts of speeches to Wall Street bankers are leaked	1	1	0	1	0	0	0	different	algorithm	
2016	10	9	Second presidential general election debate was held at Washington University in St. Louis in St. Louis, Missouri.	2	0	1	0	1	0	0	different	algorithm	
2016	10	19	The third and final presidential debate between the two major candidates was held at the University of Nevada	2	1	0	1	0	0	0	same	same	
2016	10	28	James Comey announces that the FBI will be investigating newly discovered emails pertinent to Hillary Clinton	1	0	1	1	0	0	0	different	human	
2016	11	6	James Comey tells Congress there is no evidence in the recently discovered emails that Clinton should face charges	1	1	0	0	1	0	0	same	same	
2016	11	9	Election day	2	0	1	1	0	0	0	same	same	
2016	11	10	Day after the election	2	0	1	1	0	0	0	same	same	
2016	11	13	Donald Trump's national security and terrorism speech after Orlando massacre. He repeats his call to ban Muslim immigration.	0	1	0	0	0	0	0	different	algorithm	
2016	11	14	Donald Trump turns 70. Barack Obama, in a speech, denounces Trump's plan to stop immigration to the US from Muslim countries.	0	0	0	0	1	0	0	same	same	
2016	11	17	Trump claims there will be riots if he is not named the GOP nominee, in an interview with CNN.	0	1	0	0	1	0	0	same	same	
2016	11	3	Bernie Sanders wins Indiana primary. Ted Cruz drops out of the presidential race, effectively handing Trump the nomination.	1	0	1	1	0	1	0	same	same	
2016	11	16	Day after Clinton wins primaries in Florida, Ohio, Illinois and North Carolina.	1	1	0	0	0	0	1	same	same	
2016	11	4	Day after Indiana primary. News states the race is now effectively between Trump and Clinton.	1	1	0	1	0	0	1	same	same	
2016	11	30	Clinton takes 10-point lead over Trump in polls in first poll after DNC convention.	1	1	0	0	1	0	0	same	same	
2016	12	6	Day after Democratic debate (CNN).	1	1	0	0	0	0	1	same	same	
2015	10	8	Bernie Sanders admits he was insensitive to African-American issues and Black Lives Matter in an interview with Ebony Magazine.	1	1	0	0	0	0	1	same	same	
2015	10	8	Bernie Sanders rally in Portland is interrupted by Black Lives Matter activists.	1	1	0	0	0	0	1	different	algorithm	
2015	4	28	Bernie Sanders announces his candidacy for the Democratic Party nomination for President.	1	0	0	0	0	1	0	different	human	
2015	4	29	Day after Sanders announces his candidacy for president.	1	0	0	0	0	1	0	different	human	
2016	2	2	Day after Iowa Caucus, in which Bernie Sanders loses to Clinton by a tiny margin.	1	1	0	0	0	1	0	same	same	
2016	5	19	Day after Nevada Democratic National Convention chaos breaks out	1	1	0	0	0	0	1	different	human	
2016	2	1	Day of Iowa caucuses.	1	0	1	0	0	1	0	different	human	
2016	5	2	Day after Sanders appears on CBS Face the Nation. Day before Indiana primary.	1	0	0	0	0	0	0	different	algorithm	
2016	5	1	Day of Sander's appearance on CBS Face the Nation. May 1 st is also Labour Day.	1	0	0	0	0	0	0	different	algorithm	
2015	10	14	Day after first Democratic Party debate.	1	1	0	0	0	1	0	same	same	
2016	2	10	Day after Bernie Sanders wins New Hampshire primary.	1	0	1	0	0	1	0	same	same	
2015	8	12	Bernie Sanders is ahead of Clinton in new New Hampshire primary poll. First time Sanders is ahead of Clinton in any poll.	1	0	1	0	0	1	0	same	same	
2015	10	15	Two days after first Democratic debate.	1	1	0	0	0	1	0	nodata	human	
2016	6	12	Orlando Shooting	0	0	0	1	0	0	0	nodata	human	
2016	2	20	Clinton wins Nevada	1	1	0	0	0	0	1	nodata	human	
2016	3	12	Marco Rubio wins District of Columbia	0	0	0	0	1	0	0	same	same	
2016	4	26	Clinton wins 4, Sanders 1	1	1	0	0	0	0	1	same	same	

SurveyMonkey poll



Pingdom.com



Pew Research

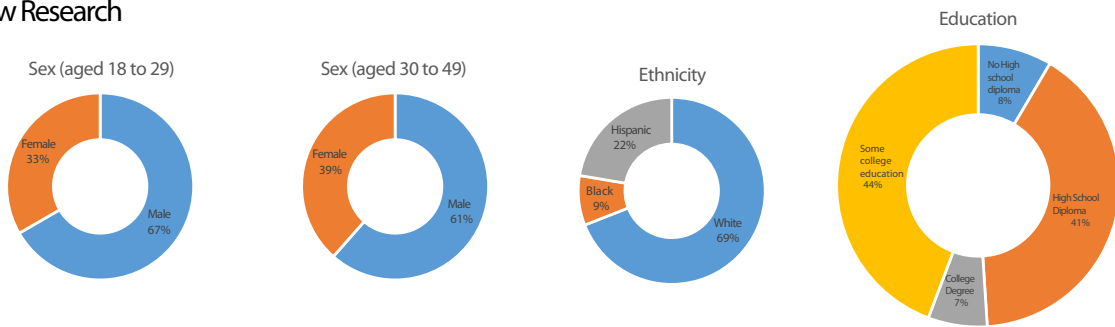


Fig. 4: Graphical representation of Reddit demographics according to different sources [31,52,23]

Table 3: News event occurrence versus subscriber change, per candidate

Donald Trump			Hillary Clinton			Bernie Sanders		
Date	Event	ΔS	Date	Event	ΔS	Date	Event	ΔS
2016-6-12	Orlando shooting	11712	2016-7-28	Dem. National Convention	534	2015-4-30	Sanders announces candidacy	8735
2016-7-27	Dem. Nat. Convention Trump's Reddit AMA	6914	2016-9-27	1 Pres. Debate. Aftermath	451	2015-8-10	BLM interrupts Sanders rally	7370
2016-11-10	Day after election	6335	2016-7-27	Dem. National Convention	441	2016-7-12	Sanders endorses Clinton	-3947
2016-11-9	Election day	6335	2016-7-29	Hillary accepts nomination	418	2015-4-28	Sanders Announces Candidacy	3795
2016-11-8	Day before election	6335	2016-7-26	Dem. National Convention	388	2015-4-29	Sanders Announces Candidacy +1	3772
2016-2-24	Nevada primary caucus	5664	2016-7-25	Dem. National Convention	372	2016-2-2	Day after IO Caucus	3609
2016-11-30	Trump illegal vote claim	4947	2016-10-9	Second presidential debate	358	2015-5-19	Sanders' Reddit AMA	3512
2016-6-13	Trump Speech on Orlando	4864	2016-9-26	First presidential debate	343	2016-2-1	Iowa caucus	3374
2016-2-23	CNN Dem. town hall	4618	2016-5-3	Sanders wins Indiana	335	2015-5-2	Two days after announcing candidacy	3249
2016-7-25	Dem. National Convention	3625	2016-3-16	Clinton won OH,IL,NC	324	2015-5-1	Sanders in CBS	3047
2016-2-22	South Car. Primary	3027	2016-3-9	Eighth and final Dem. debate	321	2015-10-14	After first Dem. debate	2751
2016-2-21	Trump in CBS	3023	2016-5-4	Hillary secures nomination	281	2016-2-10	Day after NH primary	2477
2016-6-14	Obama Speech vs Trump	2956	2016-7-30	Dem. Convention rebound (+10)	274	2015-8-12	NH poll: Sanders overtakes Clinton	2421
2016-3-17	Trump warns of riots	2611	2016-6-9	Obama endorses Clinton	260	2015-10-15	Two days after 1st Dem. deb	2094
2016-2-28	Trump in CBS	2589	2016-3-6	Aftermath Dem. Debate	260	2016-2-9	New Hampshire Primary	2048