*THIS WORK IS UNPUBLISHED AND HAS NOT BEEN PEER REVIEWED*

Modeling Political Misinformation and Bias in Online News Media Using Lexical Analyses and Historical Content Extraction

Jesse Waite, *jessewaite86@gmail.com*

Washington State University (unaffiliated with this work)

*Measure what is measurable, and make measurable what is not so.*

*Galileo Galilei*

Abstract

This work presents quantitative methods of analyzing misinformation, influence campaigns, and political bias using lexical models and historical content extraction. Computational methods for quantifying political misinformation and bias over time using large scale archival data are surprisingly absent from the literature, despite their utility. Likewise, no previous work has incorporated historical web content extraction methods. In this work, methods for quantifying political bias and misinformation are defined and applied to 2016 United States presidential election news and social media content generated by several foreign and domestic news organizations. A long term partisan analysis is also performed for news content disseminated by Google News (US), a news aggregator, from June 2011 through June 2017. The results demonstrate that strategic misinformation patterns can be ascertained using temporal sentiment analyses, and subsequently that political bias is both evident and increasing among leading news media platforms. The results are replicated for a second known dataset from [Faris et al., 2017], which refute its authors claims per 2016 presidential election objectivity by demonstrating substantial strategic bias. The results demonstrate that partisan anti-Trump and anti-conservative biases were prevalent during the election among leading American news organizations, and further, obeyed strategic misinformation patterns. This work demonstrates that historical content extraction methods and quantitative language models effectively identify strategic information patterns that qualitative approaches are insufficient to reproduce reliably. The results presented in this paper are only one application of a class of methods for describing political bias and misinformation using quantitative analyses of online content. The robustness of these approaches invites similar computational analyses of misinformation, influence campaigns, bias, and overall information dynamics with applications in political science, network science, health science, and security.

Introduction

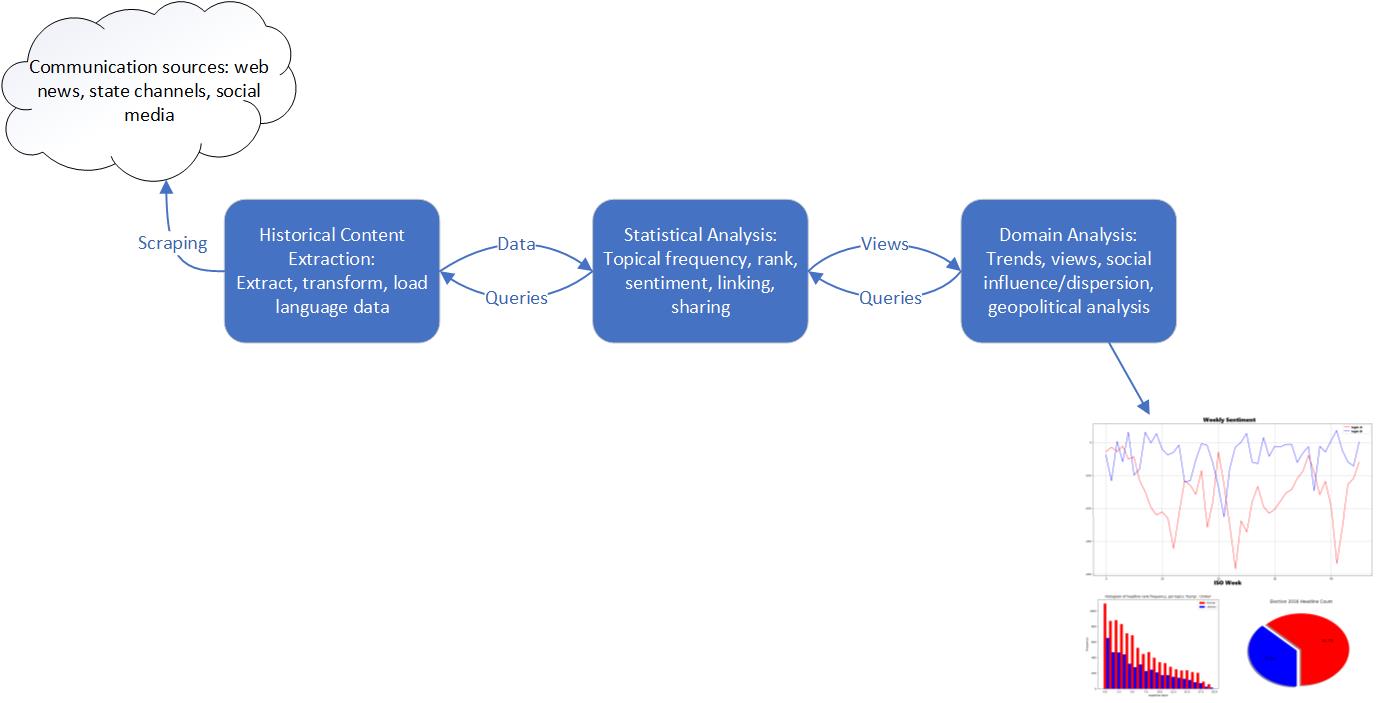
Rapid political and technological change, increasingly sophisticated misinformation operations (Lin et al, 2017), abundant examples of major media misreporting and partisan bias, polarized electorates, and decreasing public institutional trust call for robust and reproducible methods of analyzing strategic communication patterns. The information monopoly of content-oriented social and news media platforms, with their unwieldy potential to both empower and undermine democratic discourse, further animates the need to research their information dynamics using reproducible quantitative methods. Fortunately, the growth of these platforms has also generated plentiful content by which to historically analyze political communication trends.

Tightly coupled with these developments are the social and cognitive biases that have accompanied their growth. Sustained bias commits social groups to many in-group reasoning patterns and cognitive biases: confirmation bias, attentional bias, the availability heuristic, and many others examined in (Baron, 2009). Additional structural properties, such as network homophily, systematically reinforce information homogeneity whereby online audiences are attended by services and community structures more likely to provide information and sources correlated with their prior biases and preferences (Bessi, 2016). The web provides abundant examples of confirmation bias, allowing users to select information sources favoring their prior beliefs. In complementary fashion, marketing algorithms present them with homophilic content to maximize site interaction and to propagate ideas and products. This model is a contemporary extension of (Gentzlow et al, 2005), which provided a game theoretic model in which sources slant their content toward the prior beliefs of their customers to improve their reputation for quality. Strategic bias patterns are thus a form of misinformation by exploiting social flaws which supplant objective discourse, eliding information that could contradict prevailing narratives and belief.

Thus, media bias is closely wedded to cognitive and social biases and assumes several consistent and measurable forms. Coverage volume and propagation bias reinforce the availability heuristic, focusing attentional emphasis on certain topics by disproportionately maximizing their exposure and suppressing contrary topics and views, making it easier to process and contextualize over-reported topics versus their less prominent counterparts. Tonal coverage bias is identified by differences in the embedded language sentiment between topics of objectively equal merit. This bias feeds a moral license effect as audiences’ prior beliefs are consistently reaffirmed by the mistreatment or negative representation of scapegoat groups, decreasing audience introspection. Likewise, as social pressures toward reputation increase, organizations have fewer incentives to report objective information which might contradict audiences’ views, and greater incentive to avoid the inference of negative behavior, placing these institutions in a position of licensing a bias toward “correctness” to avoid negative inferences associated with reporting controversial information (Princeton, 2003). Fuller treatment of models of media bias can be found in (Gentzkow, 2015), and likewise cognitive biases in (Baron, 2000).

The connections between media bias and social cognition highlight digital society’s susceptibility to misinformation, and hence its ease of exploitation via strategically oriented information distributions. By examining these distributions, empirical models of bias and misinformation can be generated. This work’s accomplishment is to demonstrate that bias, political influence campaigns, and malicious disinformation activity affect these temporal distributions in reproducible ways, and thus can be detected using large-scale textual analyses of extracted historical content. This is perhaps obvious, since strategically disseminating polarizing information entails skewing a topic’s frequency as well as the embedded distribution of negative/positive and subjective/objective language in its context. An organization cannot promote specific perceptions on a topic without coincidingly positive/negative language, nor without the content-dissemination biases, given by the linking and citation patterns between information sources, social media pages, user shares, and other engagement statistics. By evaluating these distributions over time, reproducible characteristics of bias and misinformation can be ascertained across topics, time, regions, and languages.

Historical Content Analysis Architecture



This diagram details the basic architecture of historical content analysis. Content extraction can be done in-situ, or persisted in a database. Statistical analysis embodies low-level analyses suiting domain objectives. Domain analysis refers to the high-level objectives guided by research interests or theoretical objectives. Historical content analysis accommodates a wide range of research objectives, and perhaps more interestingly, supports unsupervised pattern discovery without the steering assumptions of prior research objectives.

Data Collection

Large scale distributional language analyses are accommodated by historical web archives such as Common Crawl (Elbaz), historical protocols such as MEMENTO (Van de Sompel et al, 2009), and tools such as MEMEX (DARPA, 2014). Such resources enable the derivation of comprehensive content datasets under a chosen data model such as daily headlines, web page snapshots, or citation structures, and thus content sources with significant online presence can be extracted and analyzed. The data can also be augmented with corresponding social media metadata such as share counts, link structure, comments, and other audience response data using available social media programmatic interfaces (APIs). Accurate analyses require data models accommodating the requirements of the analytical objectives, and moreover, lots of data. Previous work by Leetaru suggested the strategic capability of textual analysis of similar massive text archives for forecasting major regional conflagrations in the Balkans, Egypt, and Tunisia [LEETARU, 2011].

In this work, front page site snapshots were gathered for several organizations, including Google News (U.S.), Russia Today (RT), Sputnik News, Fox News, Politico, CNN, BBC, NPR, and the Washington Post for the 2016 U.S. presidential election. Other prominent news organizations are available on request. This set comprises several top-rated mainstream American organizations; a top-rated online news aggregator, Google News; beltway sources, Politico, Washington Post, and NPR; as well as the American audience branches of three foreign news organizations, BBC, Sputnik, and RT, the latter two of which are formally accused of strategic misinformation activity by American intelligence agencies (ODNI, 2017). Additionally, to analyze long term partisan trends, extended snapshots were extracted for Google News from June 2011 through June 2017.

Deterministic HTML extractors were written to extract all visible headlines from historical snapshots of each organization’s primary news page, which for simplicity is assumed to model the target distribution of information disseminated by that organization. Headlines consisted of each visible headline and textual synopsis, link, and relative ranking on the page when the snapshot was taken. Each headline was also augmented with social metadata including share counts, textual description, and other attributes via its URL using the Facebook Graph API (v2.11), a programmatic interface for extracting metadata of social media content. As such, each organizational dataset consisted of a historical collection of front page snapshots (several per day on average), each containing the ordered headlines posted by a source, and with each headline augmented by Facebook metadata. Practitioners should note that sharing statistics (counts, impressions, and so forth) are dynamic and depend on when such figures were queried; the social content metadata in this work was queried in fall 2017, well after such counts would have stabilized for content prior to summer 2017.

This data model provides the comprehensive temporal description of the distribution of information disseminated by news organizations and Google News, a popular multi-source news content aggregator. The benefit of including an aggregator is that they often approximate the expected propagation and popularity of headlines and sources. The data permits queries for two powerful analyses not present in the literature: 1) aggregate historical analyses foregrounding the overall bias of organizations 2) analysis of temporal characteristics by which organizations vary the distribution of information they disseminate over time. The latter provides direct insight into the strategic communication patterns of media organizations beyond mere aggregate metrics, and without the need for subjective bias assessment by human analysts, which cannot be reproduced or automated for large scale analyses.

These are the primary objective of this work, with an emphasis on sentiment analysis, although the data model supports myriad research objectives. The combination of language content with Facebook metadata permits estimating the overall tenor and bias of content for an organization or a complete social platform, the behavior of these distributions over time, as well as audience response to certain language features like polarizing language. It also allows estimating characteristics of link homophily, audience homophily via sharing distributions, and possibly even malicious coordination by online actors, as examined by (Faris et al., 2017). This rich combination of structural and language data contrasts with the scarcity of quantitative research of the tonal distributions of information within digital platforms, how they behave over time, recurring organizational bias patterns, analysis of misinformation actors, and other behavioral and economic impacts.

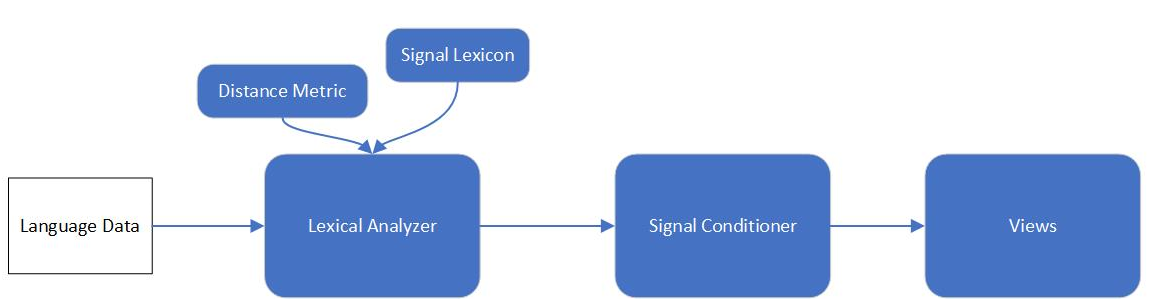
Proposed Metrics

The most relevant measures of political bias are topical coverage frequency and embedded sentiment. However, demonstrating disparate coverage volume for equivalent topics is insufficient without also analyzing its tenor. For instance, a political candidate may receive more coverage than their opponent, but the coverage may be more positive. Therefore, the most immediate way to analyze bias and political activity by news organizations is to separately analyze coverage volume and sentiment, with the latter desirably controlling for disparities in coverage volume. These may then be combined in a variety of ways to give a statistical expectation of bias.

Analyzing coverage volume is a simple matter of applying a topic-term frequency metric, aggregating these frequencies over binned time periods. Metrics for propagation and audience exposure likelihood could be devised but are left to future work. Although coverage volume foregrounds topical frequency, sentiment analysis provides deeper insight into the embedded tonal biases of content purportedly arising from an objective, random process of journalistic discovery. This work incorporates sentiment analysis into long-range content analysis to make quantitative statements about the bias of news organizations, and further, to discover temporal patterns indicative of strategic patterns.

Sentiment analysis research focuses primarily on two approaches: statistical models and learning models using derived representations to determine textual polarity, such as the VADER tool (Hutto et al, 2014). The former better suit regression tasks but require more data to estimate patterns, whereas the latter attempt to learn the syntactical structure of signaling language, typically for classification tasks, such as in (Recasens et al, 2013). Financial researchers have contributed significantly to the advancement of statistical sentiment modeling by demonstrating its effectiveness for deriving financial value functions (Zubair, et al 2015) (Mishne et al, 2006). A recent survey of sentiment analysis is provided by (Medhat et al, 2014).

In this work, the statistical model is justified by access to large quantities of language data, consisting of thousands of news articles and millions of words, with a focus on lexicon-based sentiment analysis. This accommodates the target objective of modelling the sentiment distribution of query topics (e.g. candidates, issues, political parties), where sentiment distributions are evaluated via the associative frequency of topic terms and signaling language. Estimates for topical sentiment are given by analyzing the mathematical distance between topical and signal terms. Using these estimates, both aggregate and temporal estimates of bias can be derived. This approach lends the following basic architecture for lexicon-based signals analysis.



This model generalizes lexical analysis, since sentiment is only one possible lexical query. Others include subjectivity and similar orientation-indicator lexica, such as the various lexica of the Harvard IV-4 General Inquirer (Stone et al, 1966). The signal conditioning component incorporates averaging or other signal processing algorithms to derive a smooth signal for the chosen query inputs, metrics, and data.

The lexical analyzer takes both a distance metric and a signal lexicon as input, for greater flexibility. A pointwise measure of associative strength between topic terms and signal lexicon is then:

For sentiment analysis, the signal lexicon divides into positive and negative terms, and net sentiment is calculated as:

See (Liu, 2012) for similar lexicon-based sentiment expressions of semantic orientation. Distance metrics may be rule based, distributional, or co-occurrence based. This work utilizes term co-occurrence metrics which are reviewed by (Medhat et al, 2014) as well as (Manning and Schutze, 1999). The output depends on the input lexicon, which may be pre-built or generated algorithmically from domain data. Pre-built lexica preserve a common baseline, whereas the latter can improve domain precision. For example, the term “scandal” indicates negative sentiment in a political domain yet may not be included in a generic pre-built sentiment lexicon predominantly consisting of adjectives. Lexicon generation can flexibly identify such signaling terms for a specific domain. These terms change often, as different topics become artifices of social sentiment circulating online. Another advantage is that lexicon generation accommodates languages for which pre-built signal lexica are unavailable.

Distance Metrics

There are many term-based distance metrics, most ultimately based on the co-occurrence frequency of signal and query terms within some context. Given lots of language data to estimate these frequencies, semantic associativity can be measured by summing the similarity between subsets of topical and signal words. Perhaps the most powerful is cosine similarity, since it encompasses the complete associative vector space of word co-occurrences. Let **F** be the symmetric, triangular term co-occurrence matrix for all terms in some bag of language , for which the row represents the co-occurrence frequency between word and all other words . The cosine similarity of and is given by:

Cosine similarity detects both similarity (closer to 1.0) and dissimilarity (closer to -1.0). Its semantic grasp lies in hyponymy, the detection of indirect similarity via the normalized inner-product of their word co-occurrence vectors (Lyons, 1995). Distance metrics based on immediate co-occurrence frequencies cannot measure similarity between words which never co-occur in the data, such as “red” and “dalmation”. By contrast, cosine similarity captures their semantic relationship via their shared co-occurrences with all other words, hence they might have slightly positive cosine similarity via their mutual association with “fire” and “station”. However, with quality comes greater computation cost, since computing cosine similarity values for a co-occurrence frequency matrix **F** requires calculating inner products between every row and column of **F**.

Other common word similarity metrics can be taken directly from Jurafsky or from generic statistical distance methods: pointwise mutual information (PMI) (Fano, 1961), positive-PMI (PPMI) (Church et al, 1989), Dice, Jaccard, Chi-squared, and the t-test (Manning, 1999). PPMI was used in this work because it provides the best tradeoff between computational speed and signal quality, and it is a fast and high-quality approximation to cosine similarity. PPMI is simply the positive range of PMI:

PPMI possesses the benefit of weighting by term probabilities, mitigating differences in topical coverage volume. However, topical coverage volume differences can also be controlled by comparing only article sets of equal size for the queried topics. For example, given topics *t1* and *t2*, and their respective articles and for some period, select articles from and , where .

Misinformation and Bias Evaluation

As described, the datasets consisted of multiple daily snapshots of several leading news organizations. For each of these, coverage volume and sentiment per two partisan topic term sets: and were analyzed. These terms were sufficiently unique to prevent false positive hits, and sufficiently inclusive to include most coverage of the leading candidates for U.S. president in 2016. Likewise, to test broader partisan biases beyond the presidential race, a second partisan topic set was also analyzed: and . To identify tonal bias unique to each topic set, the query results for sentiment analysis were cross-filtered of opposing topics, hence any headline containing both “clinton” and “trump” was omitted due to ambiguity. Since multiple snapshots may contain the same headline, only unique daily headlines were retained for analysis, determined via URL. Coverage volume analysis was not filtered in these ways, since its intent is to track these characteristics. A range of additional features were also analyzed, such as headline duration and rank, but are beyond the scope of this work. The algorithmic outline for two topics is outlined below.

**Algorithm 1: Lexicon-Based Temporal Sentiment Analysis**

**Input** *t1, t2*: Topic term sets (e.g., )

*metric:* Lexical distance metric or method

*extractor:* Historical snapshot/content extractor for a target organization

*grouping:*  A group-by clause for temporal bins (“week”, “month”, et cetera)

*begin:* Begin date for query

*end:* End date for query

**Output**  *plot:* A plot of net sentiment

1. h1 = getTopicalHeadlines(extractor, t1, begin, end) //historical snapshot querying
2. h2 = getTopicalHeadlines(extractor, t2, begin, end)
3. h1 = topicFilter(h1, t2) //topical cross-filtering
4. h2 = topicFilter(h2, t1)
5. headlines = h1 h2
6. headlines = deduplicate(headlines) //duplicate-headline filtering
7. bins = binArticlesBySpan(headlines, bin=“week”)
8. s1Weekly = list()
9. s2Weekly = list()
10. for bin in bins:
11. s1 = netSentiment(t1, S, bin, metric)
12. s1Weekly.add(s1)
13. s2 = netSentiment(t2, S, bin, metric)
14. s2Weekly.add(s2)
15. return plot(s1, s2)

This outline captures the essential elements of temporal sentiment analysis and can be generalized to more than two topic sets. Bin values can be plotted raw, or trend conditioned using additional signal processing, such as basic smoothing techniques or Kalman filtering. In this work, sentiment was binned by ISO week and averaged over two weeks, but raw plots are given in the appendix. Aggregate sentiment for a topic is calculated by summing the values over some time span.

If the distance metric is weighted by term frequency, then the plots are akin to a rate of negative/positive coverage, eliminating potential bias resulting from disproportional topical coverage volume. Multiplying these rates by the topical frequency at that time-slice thus provides a rudimentary estimate of expected coverage sentiment, combining both the polarity and its estimated propagation. This provides a much closer estimate of the expected-value of sentiment. This is done by modifying the lines above, where the operator *‘#’* represents the frequency of a topic within a bin of headlines. In fact, ‘#’ could be any function returning a measure of dissemination, such as returning social media share counts, and thereby providing a measure of audience response to the sentiment distribution disseminated by an organization.

**Algorithm 1.1: A modification to 1.0 for expected sentiment**

1. for bin in bins:
2. s1 = netSentiment(t1, S, bin, metric) \* #(t1, bin)
3. s1Weekly.add(s1)
4. s2 = netSentiment(t2, S, bin, metric) \* #(t2, bin)
5. s2Weekly.add(s2)

The sentiment lexicon used in this work was from (Hu et al, 2004), filtered of topical terms (e.g., as “trump” is a positive adjective), providing a common baseline. This is a lexicon derived from social media consisting of nearly 7,000 definitive sentiment signaling terms, 2,005 positive terms and 4,783 negative terms. Such compact lexica generate better signal patterns than comprehensive or multi-valued sentiment lexica due to their unambiguity.

To evaluate partisan bias over a period of years, an additional query was performed to analyze long-term bias patterns. The only organization for which such extensive data was collected and analyzed was Google News, for coverage between June 2011 and December 2016, and only partisan queries were analyzed, *t3* and *t4*. The articles pertaining to these topics were analyzed for coverage volume, sentiment via algorithm 1.0. Link histograms were also gathered, to evaluate news-item sources and their change over time for this aggregator and topic sets *t3* and *t4*. Full visual results are listed in the appendices.

**2016 Presidential Election: Coverage of Presidential Candidates**

|  |  |
| --- | --- |
| BBC | CNN |
| Fox News | Google News |
| NBC News | NPR |
| Politico | Russia Today (RT) |
| Sputnik News | Washington Post |

2016 coverage volume statistics show overwhelmingly more coverage of Donald Trump than Hillary Clinton, immediately highlighting the attentional bias of the media overall, and likely contributed to the widespread perception that Clinton was very likely to win. Sum volumes are shown in the plot legends.

|  |  |
| --- | --- |
| BBC | CNN |
| Fox News | Google News |
| NBC News | NPR |
| Politico | Russia Today (RT) |
| Sputnik News | Washington Post |

The sentiment analysis results from algorithm 1.0 reveal that coverage of Trump was overwhelmingly more negative than coverage of Clinton. The reliability of this method is given by observing that a candidate’s large positive and negative spikes precisely correspond to specific news events and the tenor of the associated organizational narratives. These spikes are very important, as they represent news narratives that become artifices for disseminating tonally oriented content which influences perception of a topic. The latter ISO-week [CITE ISO WEEK] negative spikes for Trump reflect the release of damaging audio recordings, while the positive spikes for Clinton around ISO-week [CITE ISO WEEK] reflect tonally promotional coverage of the Democratic national convention.

Despite claims of conservative leaning bias, Fox News showed the least bias of any organization, and was tonally neutral and insignificantly anti-Trump. Similarly, while RT and Sputnik are accused of engaging in strategic reporting bias in favor of Donald Trump (ODNI, 2017), this claim is contradicted by real data. No similar quantitative evidence was presented by the ODNI, whose report offered only soft intelligence. Both Russian state-run media outlets exhibited nearly neutral aggregate bias, nor was there a consistent bias in favor of Trump, nor against Clinton, directly contradicting a variety of claims made in (ODNI, 2017). In short, the data do not support that there was a cohesive pro-Trump or anti-Clinton information strategy by these organizations. This contrasts with mainstream organizations like BBC, The Washington Post, and CNN, NPR, and Politico, which bore remarkably deterministic anti-conservative and pro-Clinton bias, and reach far larger mainstream American audiences. In fact, if the concern is one of foreign state media interference, then UK-funded BBC demonstrated the most reliably biased and partisan oriented foreign news organization of the three, much more so than Sputnik and RT.

A key observation is that the two plots demonstrate a high degree of auto-correlation emerging in the final 12 weeks of the election, whereby coverage trended negative for one candidate and trended positive for the other, and vice versa. Notably, the pattern emerges after election primaries completed, in the general head-to-head phase of the election. The patterns imply that immediately, media coverage became oriented by a left-versus-right, anti-conservative and anti-Trump strategic logic in the general election phase. This pattern was reproduced by multiple organizations and is very worthy of further analysis using distributional approaches. The patterns demonstrate the effectiveness of using temporal sentiment metrics to bear out adversarial patterns of misinformation and influence.

In that regard, one can remark that strategic reporting patterns by RT and Sputnik are apparent in the final days of the election, with Clinton news trending more negatively and Trump coverage trending more positively. However, these signatures are born out far more strongly by most other western organizations, for the inverted anti-Trump/pro-Clinton strategic pattern. These results imply that if the distribution of information of state-run RT and Sputnik demonstrate directed and strategic misinformation activity, as alleged by (ODNI, 2017), then far stronger and more deterministic behavior is evidenced by leading Western media organizations with far greater U.S. propagation, most notably during the most sensitive final weeks of the 2016 election. That is, if the patterns demonstrated by RT and Sputnik represent alleged strategic misinformation, then far worse patterns are strongly evident among major Western news organizations in terms of homogenous aggregate bias, as well as strategic misinformation patterns over time. A second interpretation is that the RT/Sputnik patterns should not be characterized by their sustained partisan patterns, which largely cancel over time. This interpretation, a subject for future work, would characterize misinformation less by sustained long-term patterns of candidate preferences than by volatile short-term preference toward discord. But in both Russian and Western cases, the sentiment spikes demonstrate how organizations select stories and their narratives as artifices of sustained political influence.

2016 Presidential Election: Partisan Coverage

[visual]

The same patterns emerge for partisan topics, and . Once again, coverage volume bias is very evident in the disproportionate coverage of right-leaning topics.

[visual]

Foremost, aggregate bias for 2016 was strongly negative toward Republican and conservative topics, whereas headlines about Democrats tended to be more positive. Neither party enjoyed aggregate positive coverage, but Democrat topics enjoyed several periods of net positive coverage, which should be regarded with alarm, since critical and objective language should tend to be either neutral or negative. Likewise, a high degree of autocorrelation emerges once again, indicating a dependence between the rate-of-change of coverage sentiment for these topics, with more negative coverage of Republicans coinciding with more positive coverage of Democrats.

Google News: Long Term Trend Analysis

[long term visual] [subtract both plots]

Google News partisan analysis of topic sets *t3* and *t4* show consistently more negative coverage of coverage of conservative topics and more positive coverage of left-leaning topics. The results also show that bias has increased within the last three years and continues to worsen. This can be shown by plotting the distance between the two plots, as shown.

[histogram]

A headline link histogram of 2016 news about topics *t1* and *t2* indicates a highly non-uniform link distribution, with overwhelming linking toward strongly leftwing news organizations. The further left leaning the organization based on the previous sentiment results, the greater its featured frequency in Google News, per reporting on the 2016 general election presidential candidates. Since Google News purports to aggregate headlines algorithmically, this highly uneven distribution suggests structural network features of information bias, across platforms and organizations.

Reproduction

To further demonstrate the reliability and capability of the previous results, the same evaluation methods were applied to the dataset used by the work “Partisanship, Propaganda, and Disinformation: Online Media and the 2016 U.S. Presidential Election” by (Faris et al., 2017). The authors generously provided a replication dataset of web-based election news gathered from a wide variety of sources and around 2.1 million news items, upon which they based many of their conclusions. The importance of this dataset is that unlike the organizational extractions from the previous sections, it approximates the complete distribution of 2016 web election news content. For textual analysis purposes, this data included only the headline language of news items and was not augmented by Facebook metadata. This means that unlike the previous analysis, synopsis language was not included for each headline, but could be acquired in the future.

The same dataset was analyzed for basic volume bias and sentiment bias under algorithm 1.0. Differential candidate treatment was again analyzed using topic sets *t1* and *t2*, and differential partisan bias was analyzed using topics *t3* and *t4*, and the input sentiment lexicon was once again from [BING CITE]. The results are shown below.

|  |  |
| --- | --- |
|  |  |
|  |  |

The results are consistent with the previous results and provide a much more all-encompassing view of media bias and misinformation patterns. Notably, the same anti-conservative pattern emerges of higher frequency coverage which was also substantially more negative overall for topics t3/t4. Likewise, a definite pattern of anti-Trump coverage that was both higher frequency and more negative. Moreover, the *t1/t2* sentiment plots bear out a significant degree of cross correlation once the general election phase was underway, around ISO-week 20-2016. This once again reaffirms that 2016 election media content overall, not simply on an organizational basis, became highly polarized by an adversarial head-to-head polarization which consistently favored Clinton and relentlessly undermined Trump with negative coverage. The negative spikes for Trump exemplify such sustained negative coverage. For example, the dramatic negative spike in Trump coverage centered around ISO week 41-2016 occurred because of the release of character damaging audio tapes.

It cannot be understated that the overall patterns of more negative anti-Trump, anti-conservative, and moreover the temporal patterns these distributions obey, directly refute key claims of (Faris, et al., 2017), whose authors provided no quantitative evidence of a tonally anti-Clinton information distribution in the 2016 election. In terms of both volume and sentiment, the distribution of information was one that covered Clinton less and more positively, whereas Trump was covered far more often and more negatively. Additionally, this coverage obeyed patterns of strategic misinformation, which became increasingly anti-Trump in the approach to election day. Lastly, the negative spikes above occurred because of personal, character-damaging stories, not information relevant to policy discussion or horserace content. These characteristics directly refute the authors’ claims that mainstream media did not engage in “partisan-disciplined messaging”. The results above demonstrate the exact opposite: an extraordinarily deterministic partisan-discipline of pro-left sentiment signaling, in terms of overall bias, its development over time, and a fixation on personalized non-policy stories.

In summary, characteristics of agenda-setting in favor of Trump or against Clinton are not evident in this data and under these quantitative analyses of the distribution of information and sentiment. Moreover, the deterministic anti-conservative bias of 2016 election media potentially reframes the emergence of fragmented, partisan right-leaning media as a product of consumer disaffection with measurable leftwing bias in mainstream media, and consumers gravitating towards a more pluralistic information environment. In short, a news media that has moved far to the left of its audience. This is a fruitful subject of future analyses using sharing statistics to analyze the divergence between the distribution of information disseminated by news media organizations versus the distributions which audiences shared on social media in 2016. While the authors of [] portray the emergence of a fragmented right-leaning network structures as sources of misinformation, the results above suggest that it is both marginal right-leaning organizations and the leftwing mainstream media which are engaging in a mutual interplay of misinformation and bias within a highly polarized information environment.

Discussion

These results make a firm quantitative statement that anti-conservative strategic bias in news media indeed exists, and that this bias has increased in recent years. Additionally, the temporal reporting sentiment patterns foreground partisan strategic logic demonstrating how different topics and events of equal objective merit are reported, as well as the dependencies between topical sentiment for opposing topics and . The results for foreign organizations RT, Sputnik, and BBC, provide a groundbreaking view into their strategic, soft-power influence strategies. This demonstrates the soundness of using lexical analysis methods to analyze geopolitical topics across a range of international and regional media sources for democratic security and research purposes.

Since PPMI-based sentiment bias is normalized by topic probability, the sentiment results imply that not only does the American media cover right-leaning candidates and topics far more often, this content is also substantially more negative and behaves according to patterns of misinformation and influence. The compounding effect of higher coverage volume and sustained negative tenor effectively demonstrate that leading media organizations in the U.S. are far from impartial, exerting partisan bias in terms of the tenor and frequency of content they disseminate. The Google News 2016 source histogram further demonstrate structural insularity and algorithmic bias in terms of how content is propagated, providing evidence of a partisan and left-leaning “echo chamber” among mainstream social, media, and tech platforms.

Conclusions and Future Work

This work set forth methods and metrics for analyzing bias and misinformation across domains, languages, and sources to establish reproducible relationships between geopolitical events and communication dynamics. Justifying strategic patterns of partisan bias is beyond its scope, however, the results for RT and leading American media outlets provide patterns indicative of political activity and demonstrate the importance of further developing these methods. The methods could easily be extended to larger datasets, other regional organizations, social communication, different types of signal lexica, and lastly, examining structural characteristics of misinformation and bias. Other possible applications include modeling information patterns with respect to specific policy issues or events, financial data, as well as measuring how bias influences policy outcomes and public knowledge. Although this work illuminates temporal strategic information patterns, distributional approaches offer greater precision and firmer connections between patterns and theoretical statements.

There is also a large potential body of work analyzing audience responsiveness to polarized content. Media structure changed dramatically as news platforms have pivoted away from independent research and toward content propagation, and it would be illuminating to evaluate whether the polarity of organizations correlates with these structural trends in terms of ownership, market structure, audience engagement statistics, or the audience tracking technologies installed on their sites. Game theoretic economic models of bias were examined by (Morris, 2001), and it would be interesting to re-evaluate apply these econometric models by incorporating lexical analyses to understand the potential incentive structures behind apparently volatile media bias.

Overall, this work demonstrates the use large-scale textual analysis to uncover hidden social and political patterns for many applications, at a moment when such tools are critically needed. All data used in this work can be obtained on request, and all technical feedback and comments are welcome.

References

Facebook Graph API, version 2.11.

Liu, Bing. (2012). Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, *5*(1), 1-167.

Elbaz, Gil. Common Crawl.

Defense Advanced Research Projects (DARPA). “Memex”. [https://www.darpa.mil/program/memex. Accessed 12/9/2017](https://www.darpa.mil/program/memex.%20Accessed%2012/9/2017). Earliest reference, September, 2014.

Bessi, A., Petroni, F., Del Vicario, M., Zollo, F., Anagnostopoulos, A., Scala, A., ... & Quattrociocchi, W. (2016). Homophily and polarization in the age of misinformation. The European Physical Journal Special Topics, 225(10), 2047-2059.

Hutto, C. J., & Gilbert, E. (2014, May). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In Eighth international AAAI conference on weblogs and social media.

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, *5*(4), 1093-1113.

Manning, Christopher D., and Schütze, Hinrich. (1999). *Foundations of statistical natural language processing* (Vol. 999). Cambridge: MIT press.

Jurafsky, D. (2000). Speech and language processing: An introduction to natural language processing. *Computational linguistics, and speech recognition*.

Background to “Assessing Russian Activities and Intentions in Recent US Elections”: The Analytic Process and Cyber Incident Attribution. Office of the Director of National Intelligence. January 6, 2017. Accessed November 11, 2017, from https://www.dni.gov/files/documents/ICA\_2017\_01.pdf.

Lin, H. S., & Kerr, J. (2017). On Cyber-Enabled Information/Influence Warfare and Manipulation.

Baron, J. (2000). *Thinking and deciding*. Cambridge University Press.

Hu, M. and Liu, B. (2004). Mining and summarizing customer reviews. In SIGKDD-04.

Liu, B. (2015). Sentiment Analysis: Mining Opinions, Sentiments, and Emotions. Cambridge University Press.

Monroe, B. L., Colaresi, M. P., and Quinn, K. M. (2008). Fightin’words: Lexical feature selection and evaluation for identifying the content of political conflict. Political Analysis, 16 (4), 372–403.

Gentzkow, M., & Shapiro, J. M. (2006). Media bias and reputation. *Journal of political Economy*, *114*(2), 280-316.

Church, K. W., & Hanks, P. (1990). Word association norms, mutual information, and lexicography. *Computational linguistics*, *16*(1), 22-29.

Church, K. W. and Hanks, P. (1989). Word association norms, mutual information, and lexicography. In ACL-89, Vancouver, B.C., pp. 76–83

Van de Sompel, H., Nelson, M. L., Sanderson, R., Balakireva, L. L., Ainsworth, S., & Shankar, H. (2009). Memento: Time travel for the web. arXiv preprint arXiv:0911.1112

Fano, R. (1961) Transmission of Information: A Statistical Theory of Communications. MIT Press, Cambridge, MA.

[SENTIMENT LEXICON]

Hu, M., & Liu, B. (2004, August). Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 168-177). ACM.

Gentzkow, M., Shapiro, J. M., & Stone, D. F. (2015). Media bias in the marketplace: Theory. In *Handbook of media economics* (Vol. 1, pp. 623-645). North-Holland.

Lyons, J. (1995). *Linguistic semantics: An introduction*. Cambridge University Press.

Morris, S. (2001). Political Correctness. *Journal of Political Economy*, *109*(2), 231-265.

LEETARU, Kalev. Culturomics 2.0: Forecasting large-scale human behavior using global news media tone in time and space. First Monday, [S.l.], aug. 2011. ISSN 13960466. Available at: <http://firstmonday.org/ojs/index.php/fm/article/view/3663/3040>. Date accessed: 17 dec. 2017. doi:http://dx.doi.org/10.5210/fm.v16i9.3663.

HTTP Framework for Time-Based Access to Resource States – Memento. Retrieved from https://mementoweb.org/guide/rfc/. Accessed on 1/2/2018.

Recasens, M., Danescu-Niculescu-Mizil, C., & Jurafsky, D. (2013, August). Linguistic Models for Analyzing and Detecting Biased Language. In ACL (1) (pp. 1650-1659).

Mishne, G., & Glance, N. S. (2006, March). Predicting Movie Sales from Blogger Sentiment. In *AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs* (pp. 155-158).

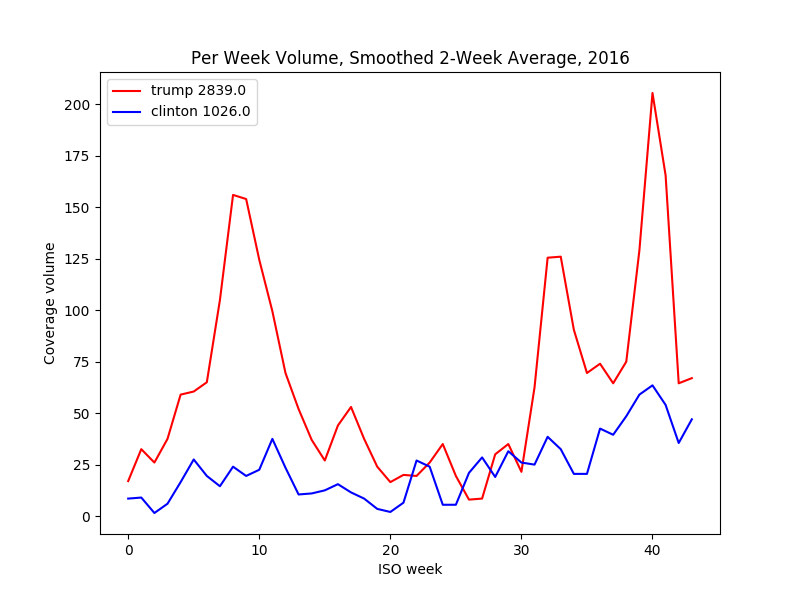
Zubair, S., & Cios, K. J. (2015, January). Extracting News Sentiment and Establishing Its Relationship with the S&P 500 Index. In *System Sciences (HICSS), 2015 48th Hawaii International Conference on* (pp. 969-975). IEEE.

Stone, P. J., Dunphy, D. C., & Smith, M. S. (1966). The general inquirer: A computer approach to content analysis.

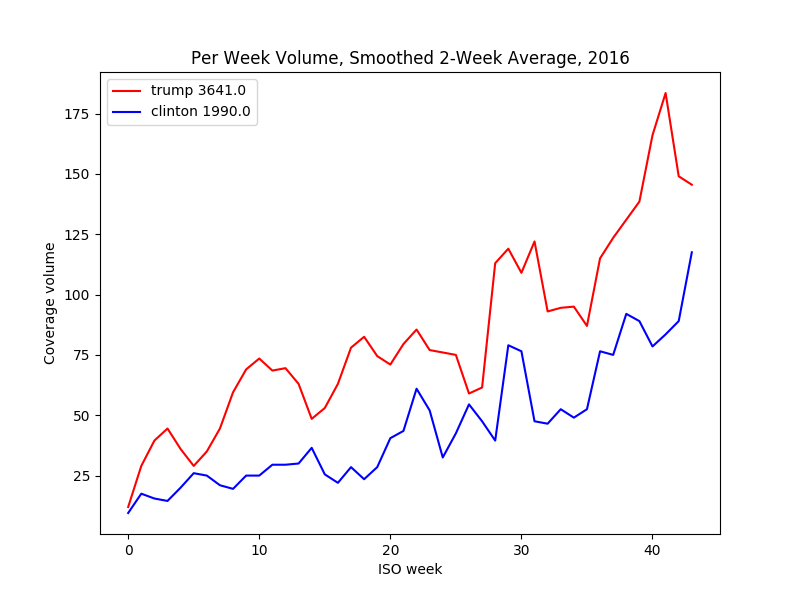
Faris, Robert and Roberts, Hal and Etling, Bruce and Bourassa, Nikki and Zuckerman, Ethan and Benkler, Yochai,

Partisanship, Propaganda, and Disinformation: Online Media and the 2016 U.S. Presidential Election (August 2017). Berkman Klein Center Research Publication 2017-6. Available at SSRN: <https://ssrn.com/abstract=3019414>

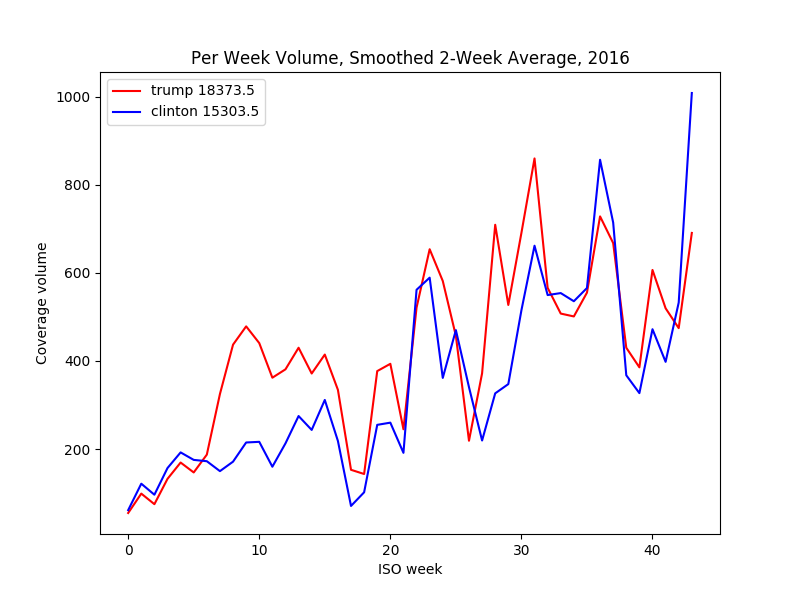
**Appendix I: Presidential Candidate Coverage Volume**



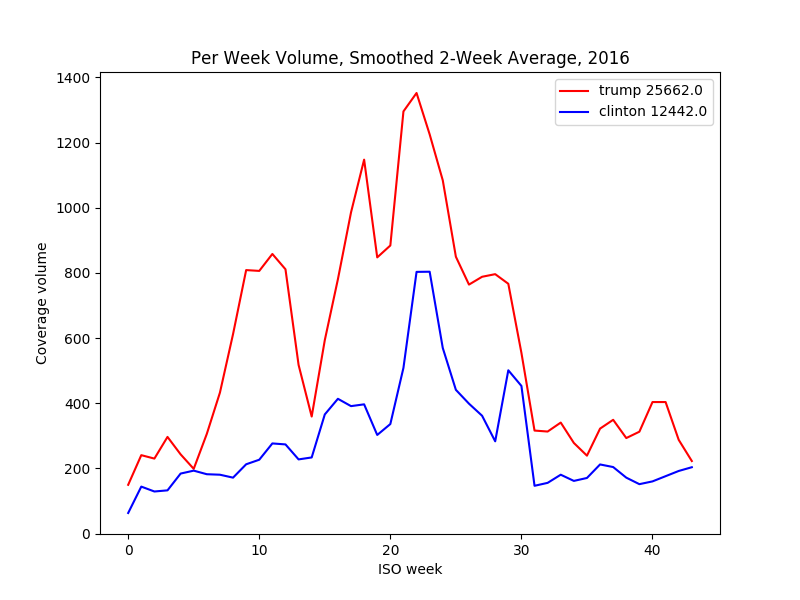
BBC



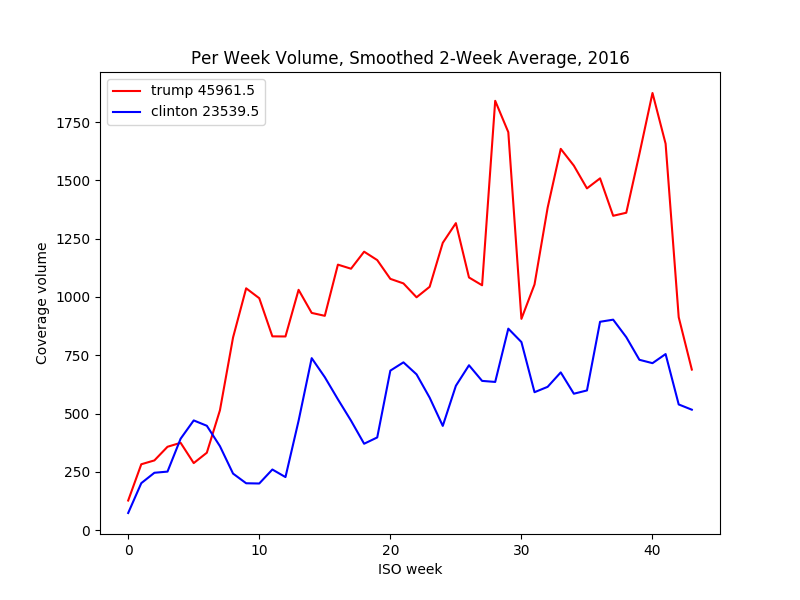
CNN



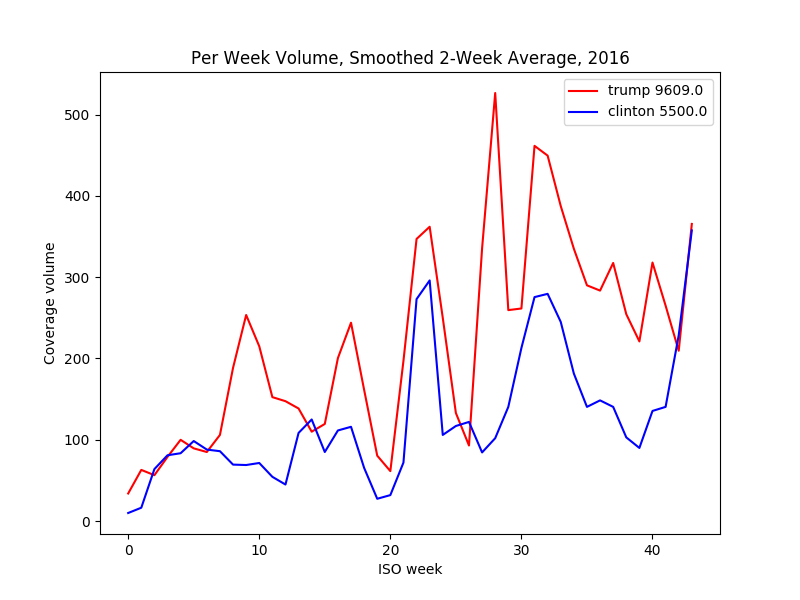
Fox News



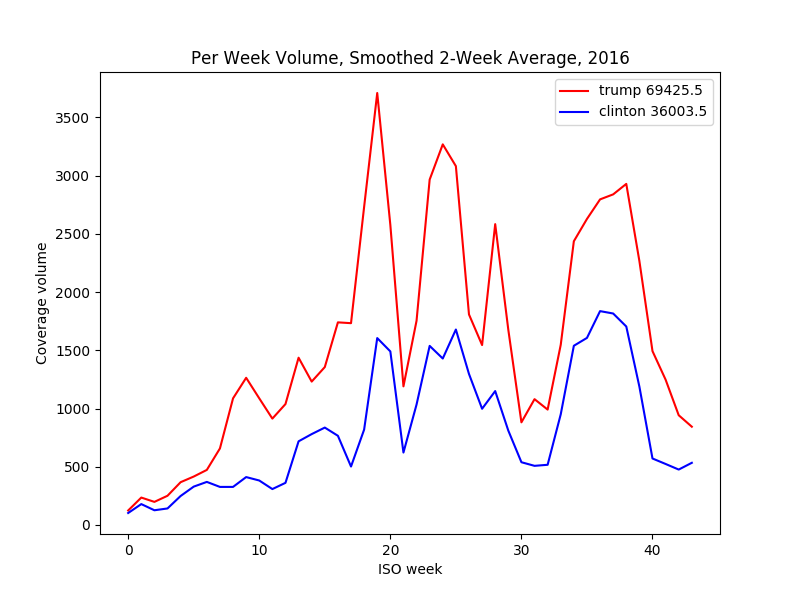
Google News



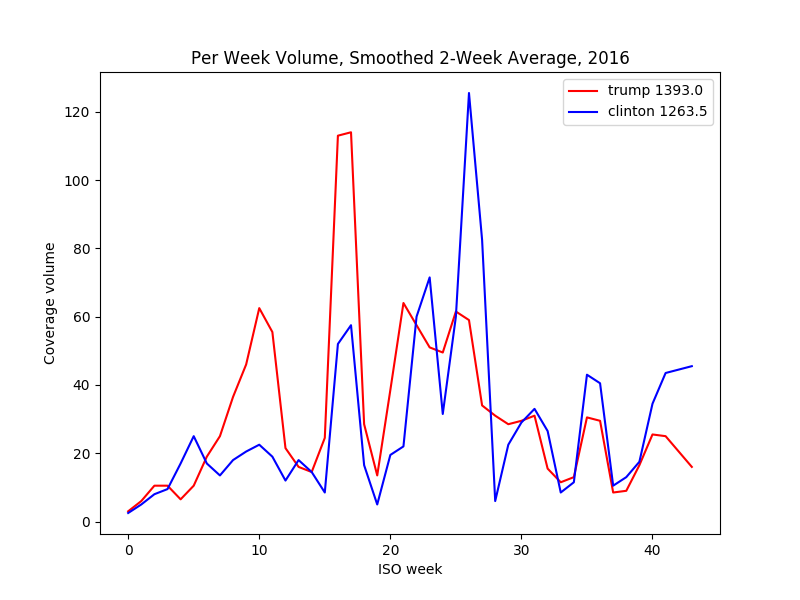
NBC News



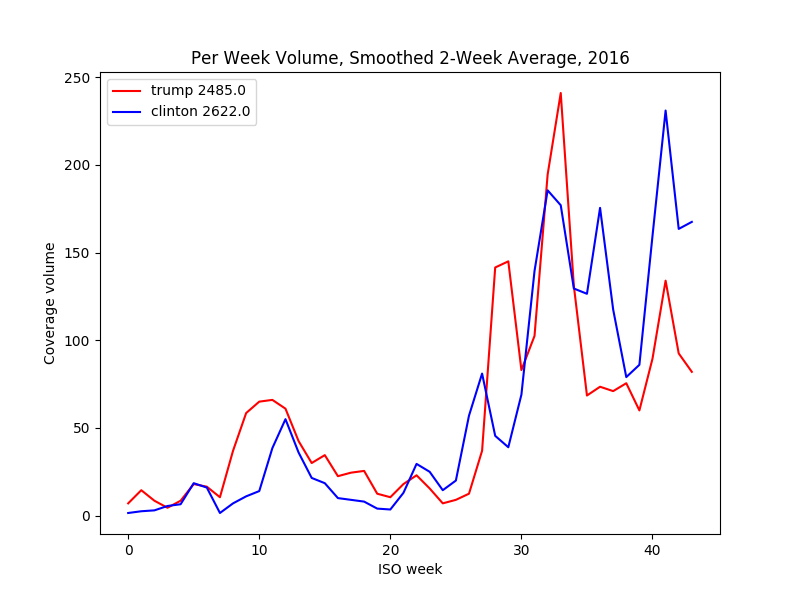
NPR



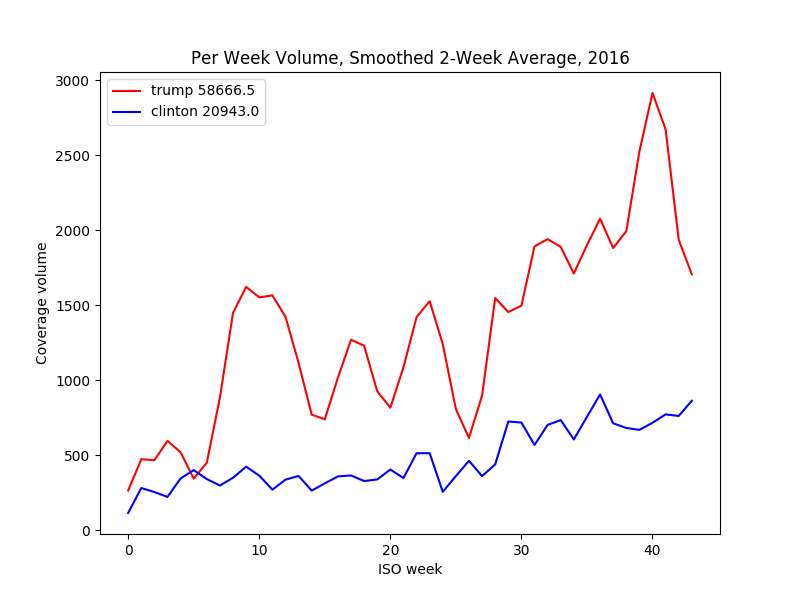
Politico



Russia Today (RT)

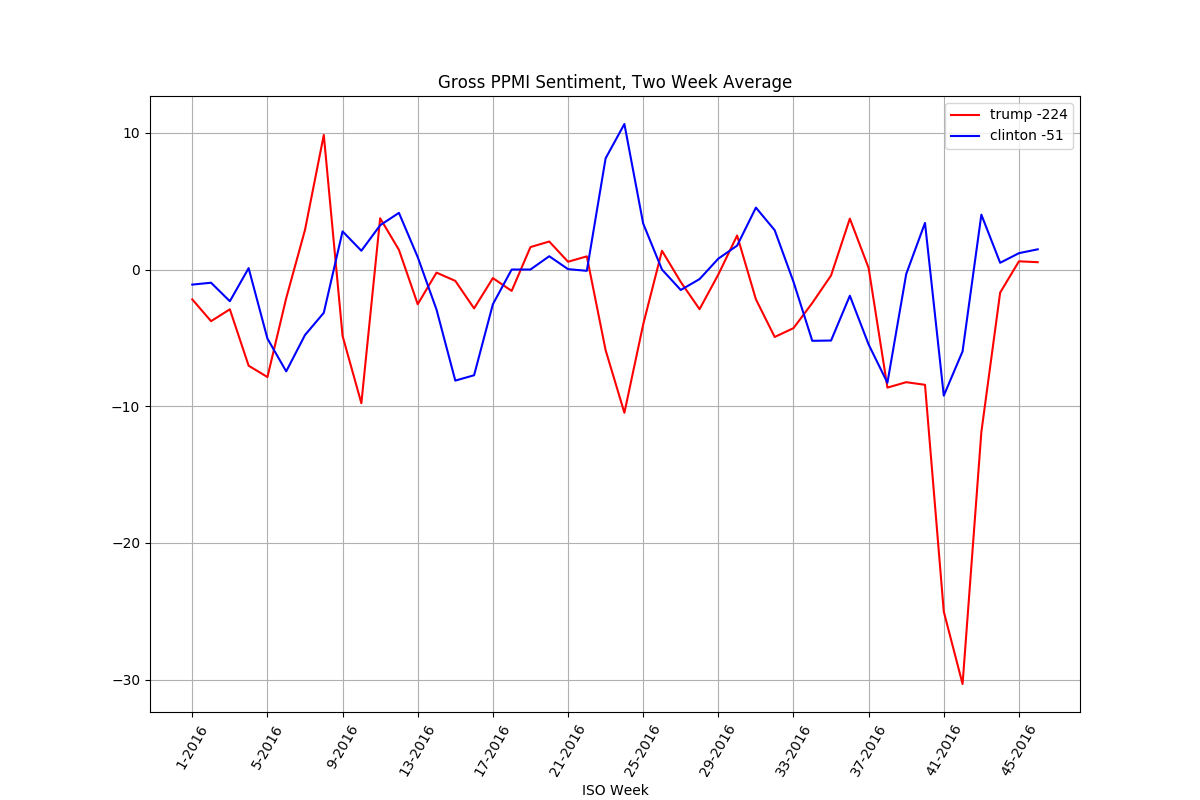


Sputnik News

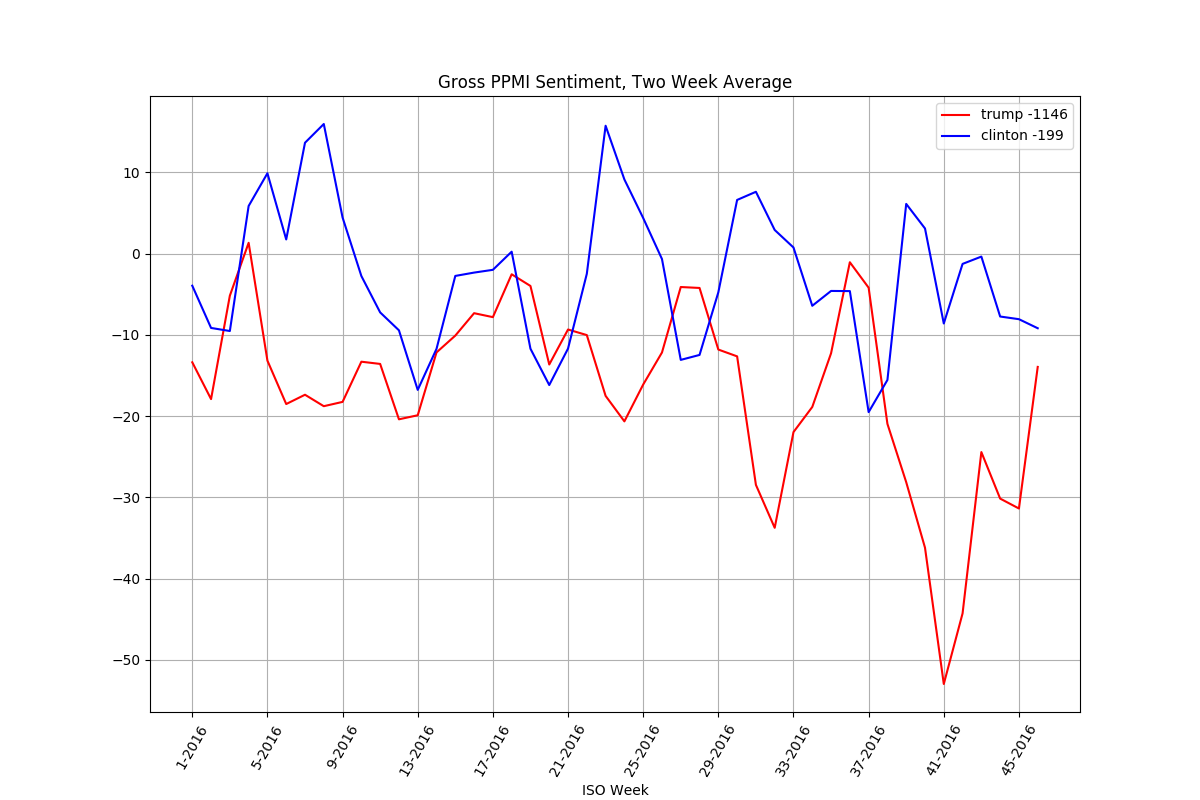


Washington Post

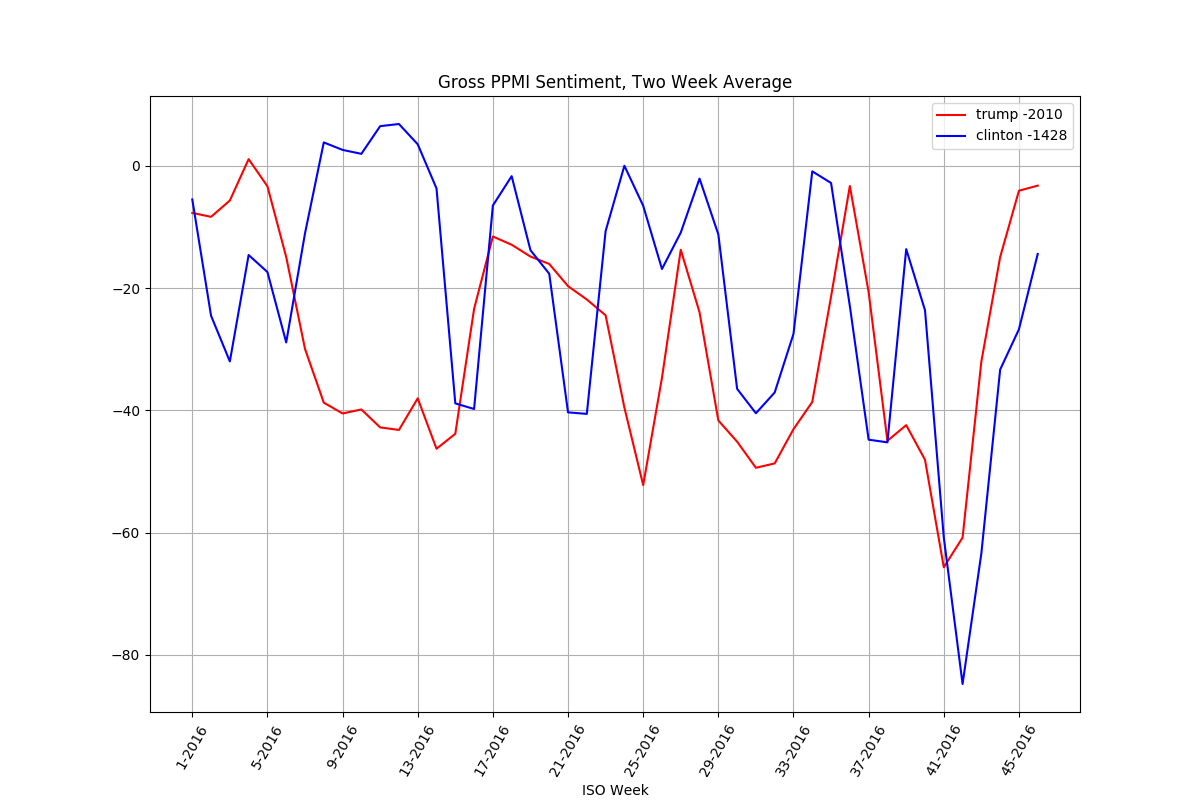
**Appendix II: Presidential Candidate Sentiment Analysis**



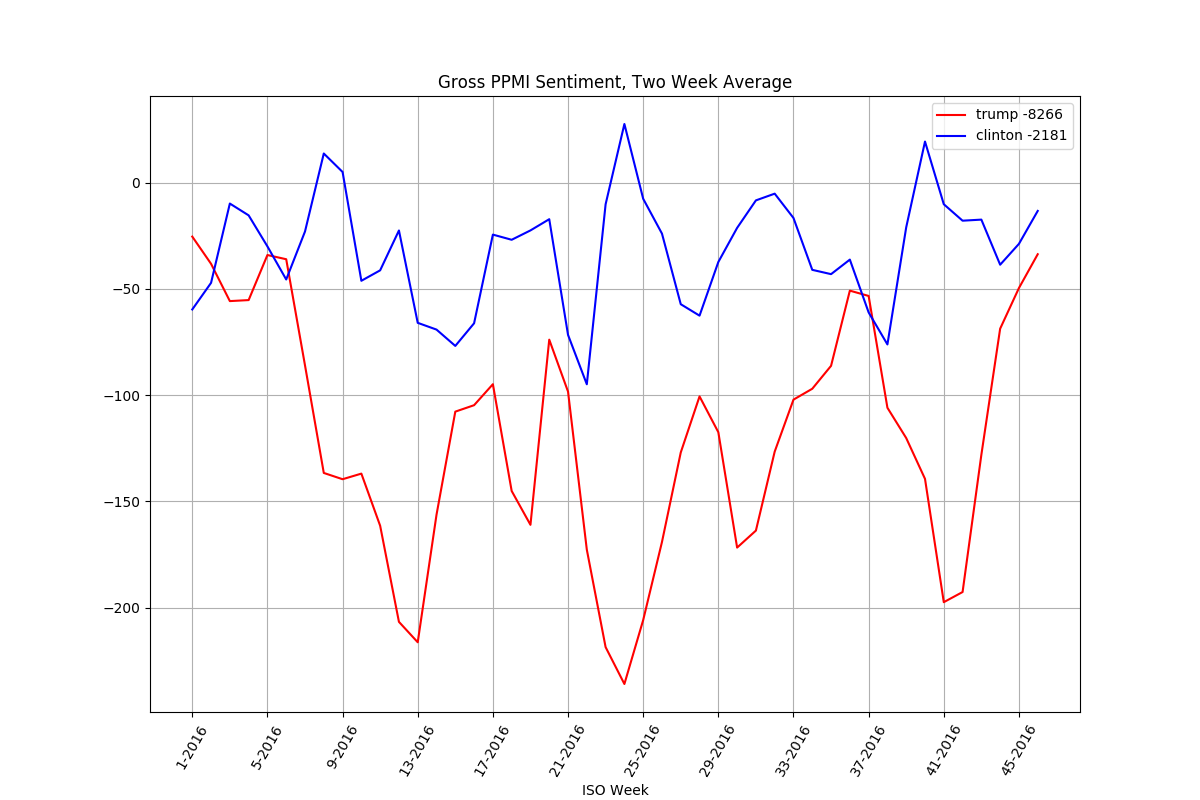
BBC Candidate Sentiment



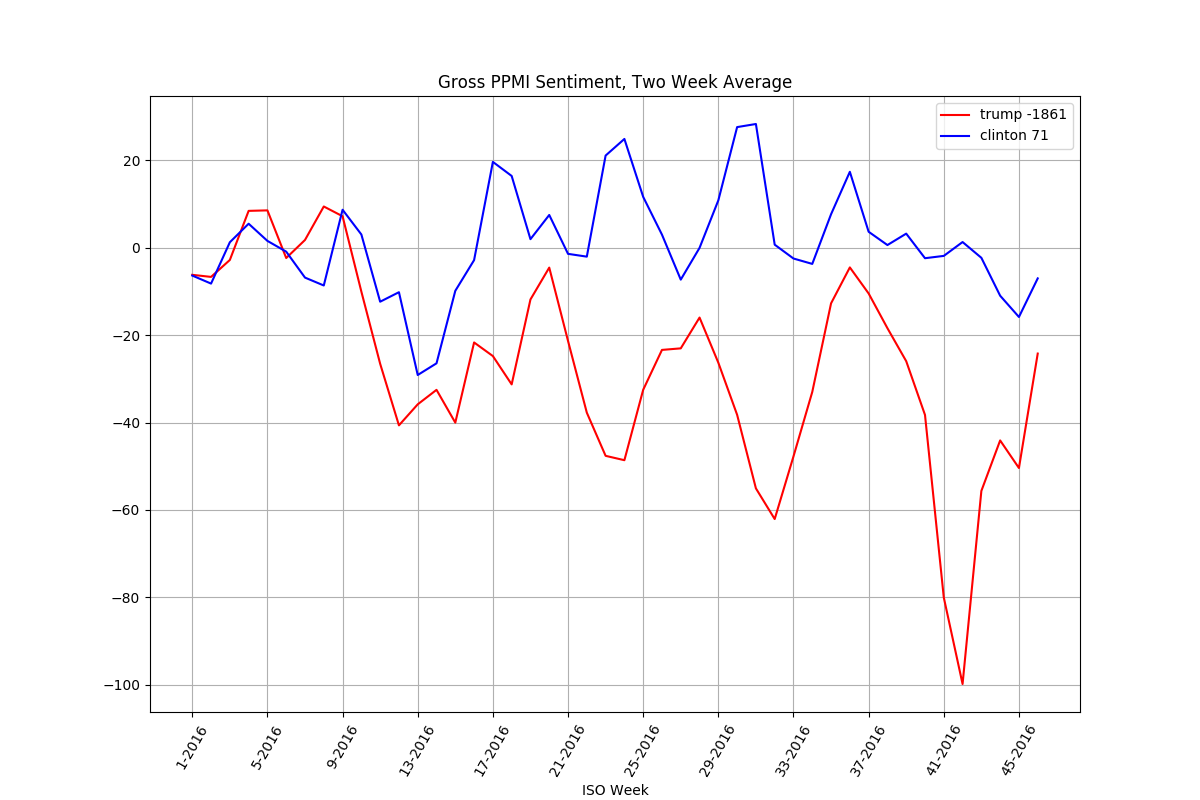
CNN



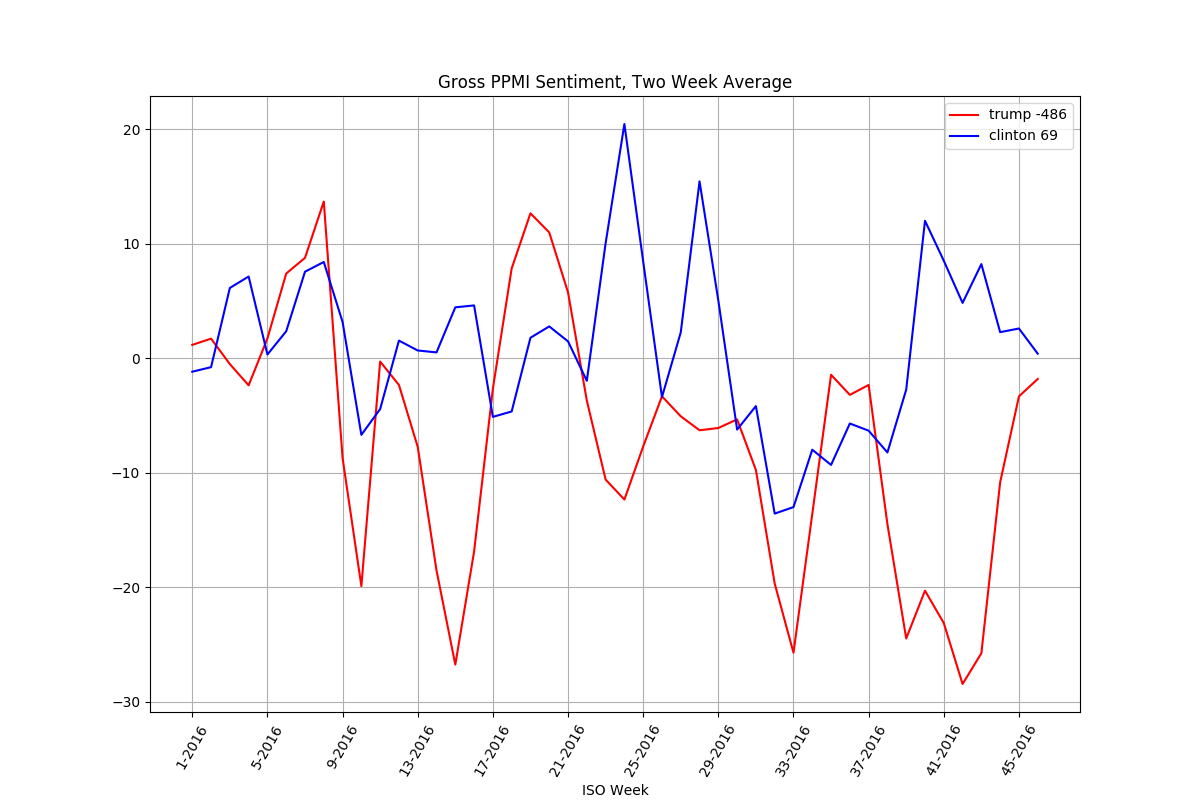
Fox News



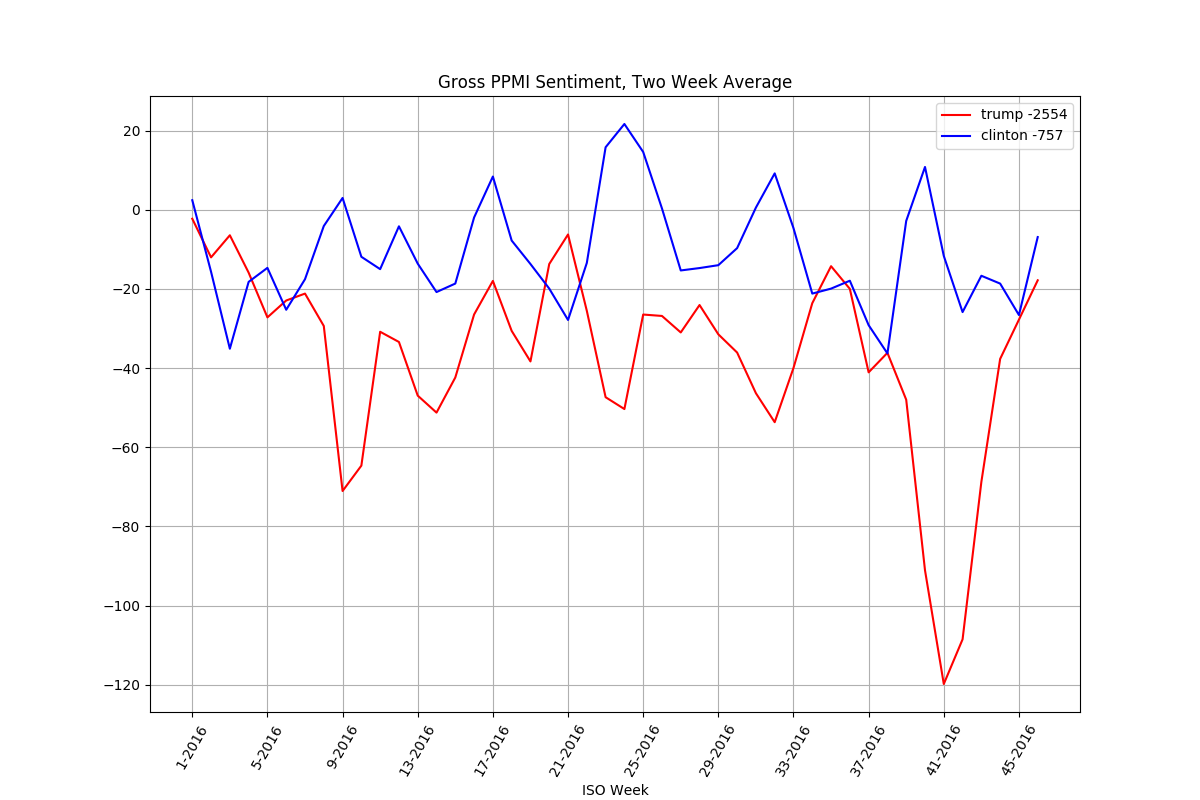
Google News



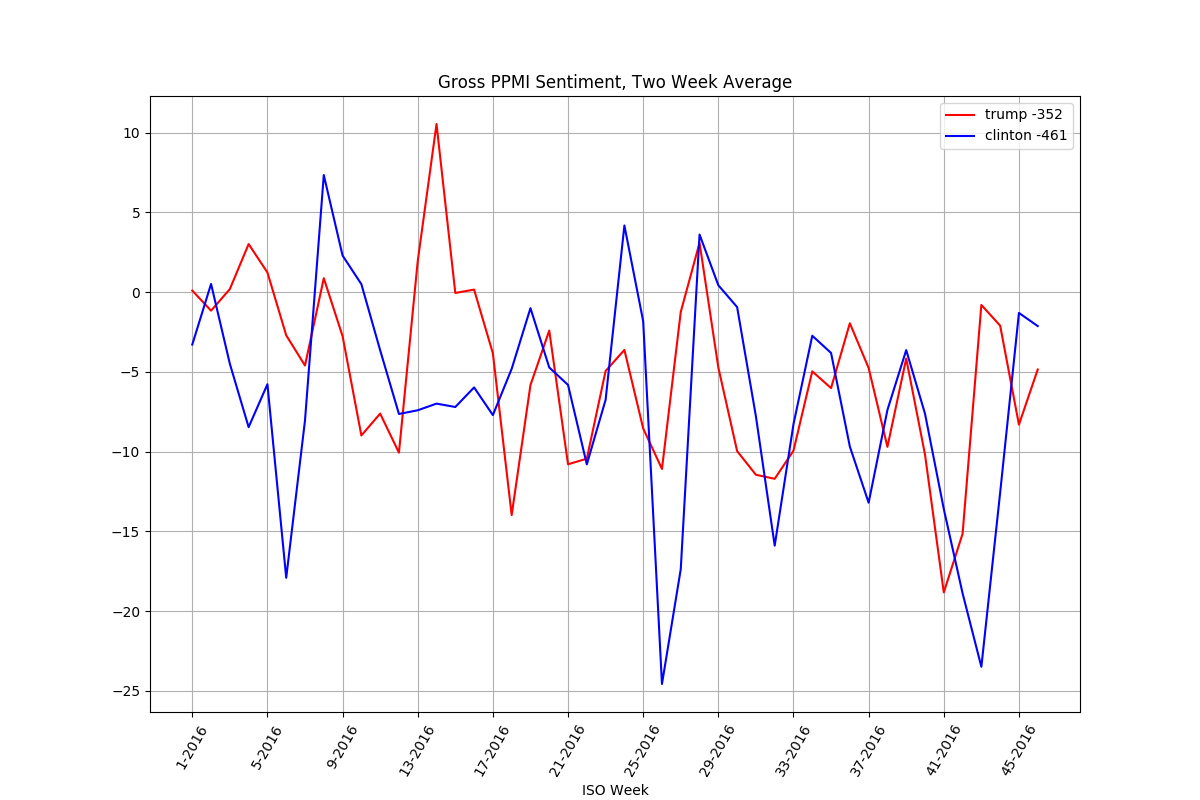
NBC News



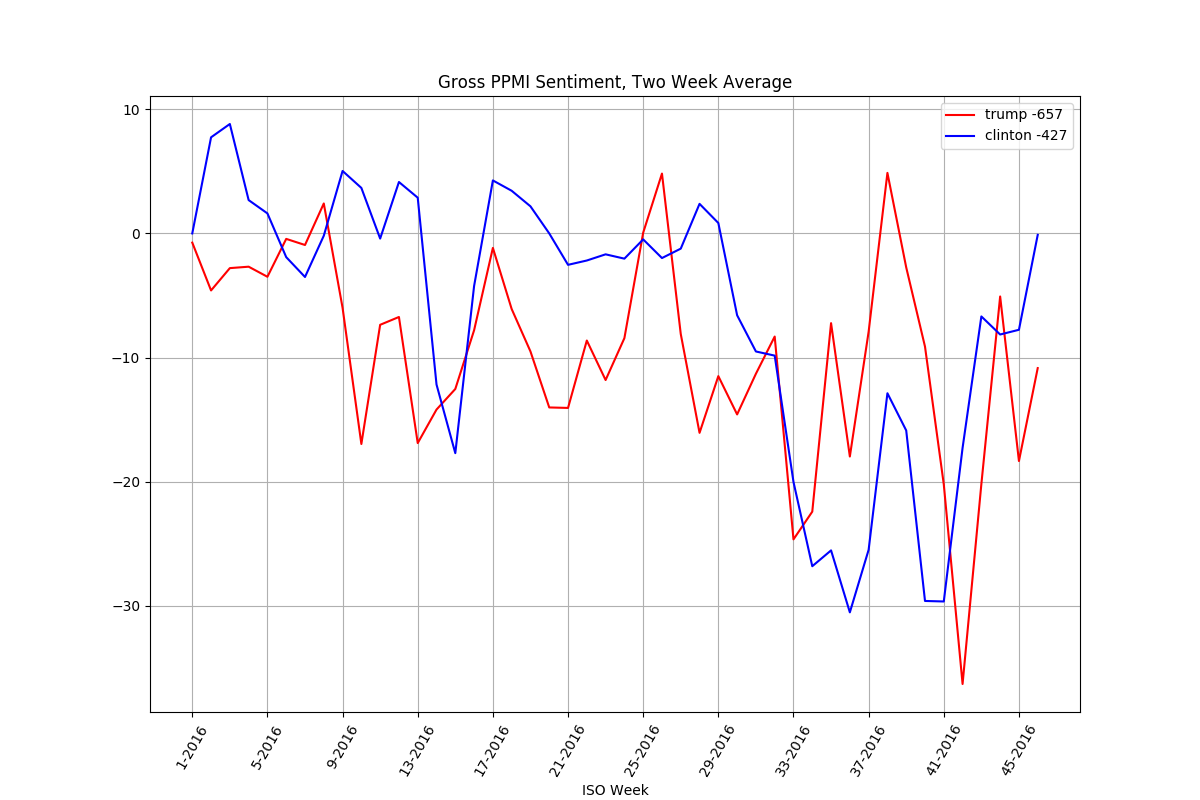
NPR



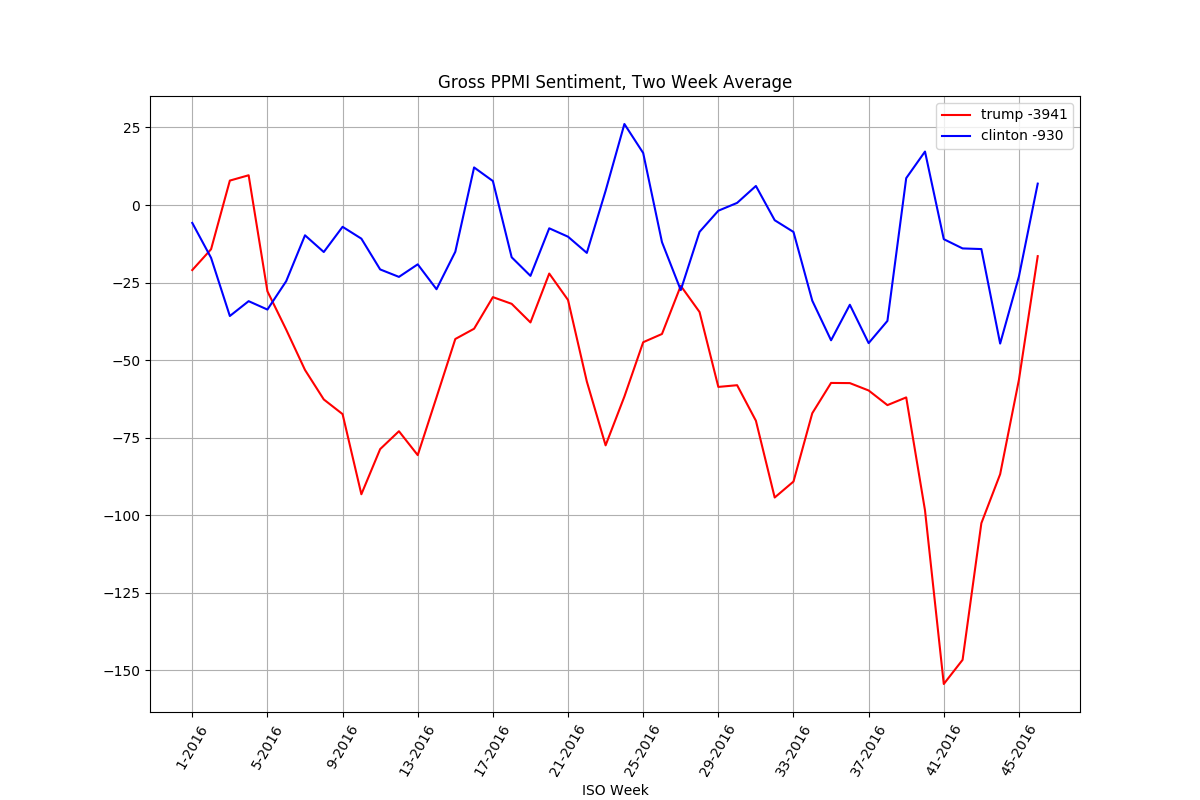
Politico



Russia Today (RT)

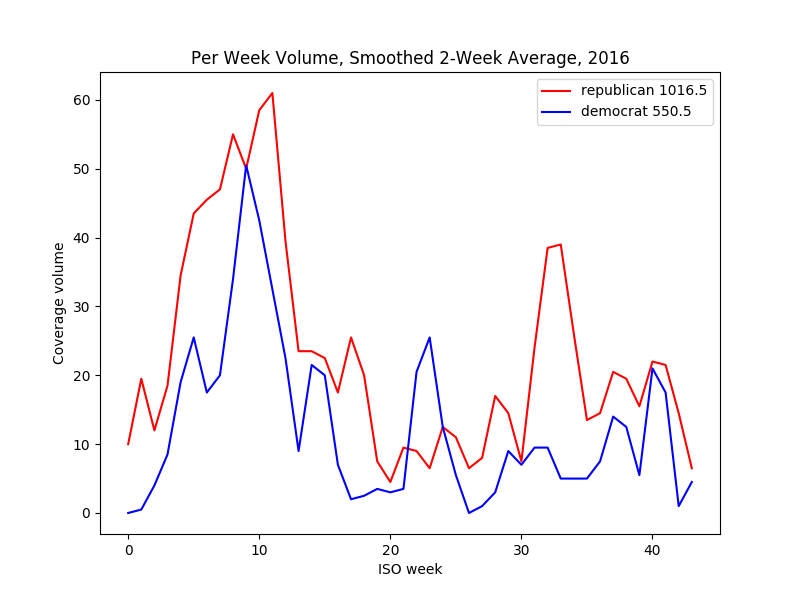


Sputnik News

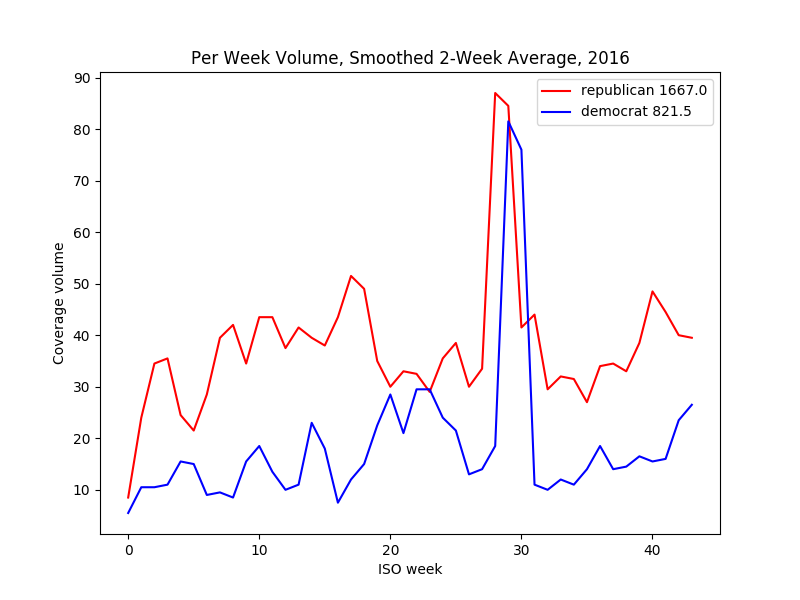


Washington Post

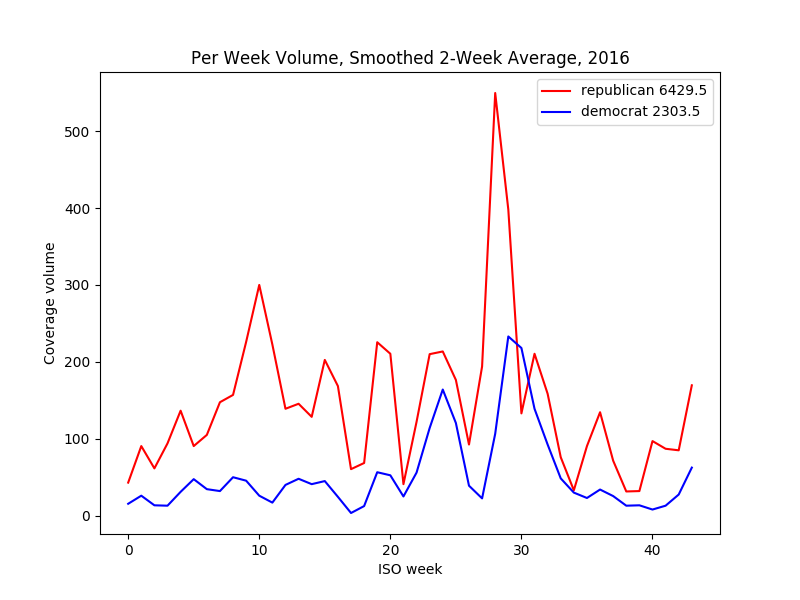
**Appendix III: Partisan Topic Coverage Volume**



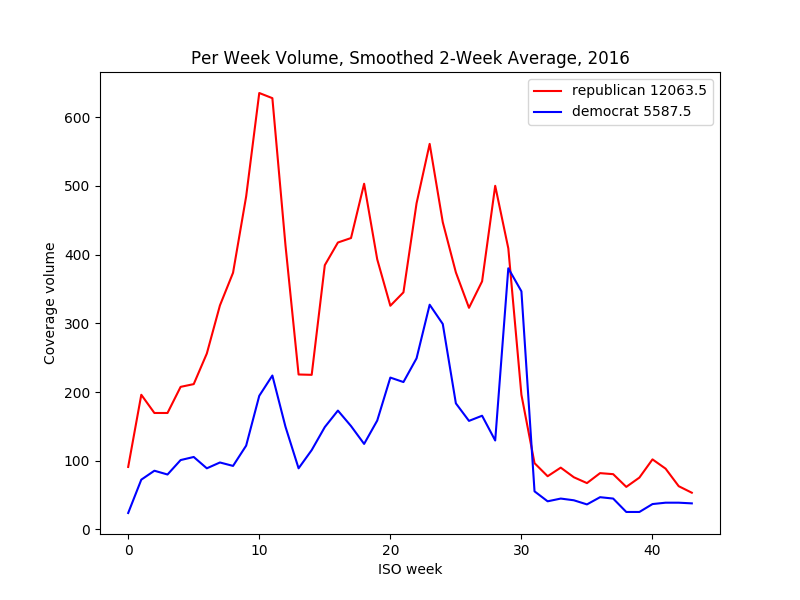
BBC



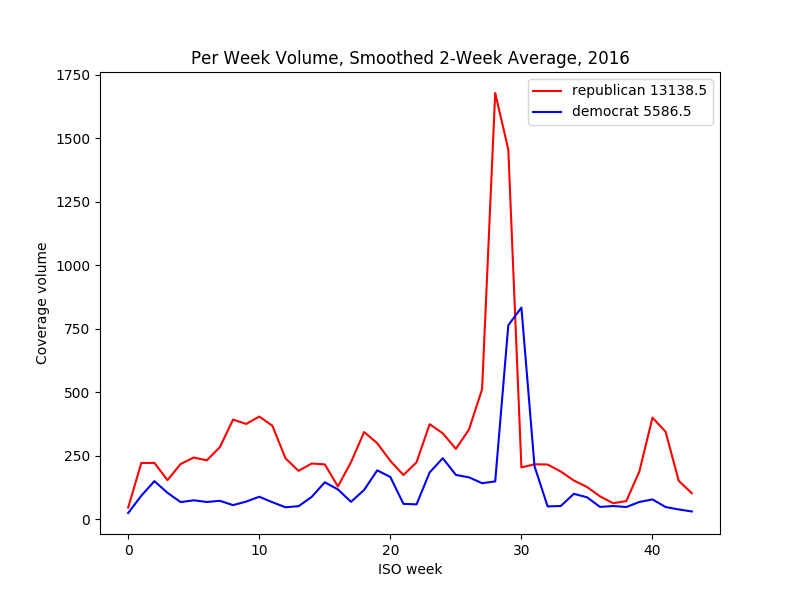
CNN



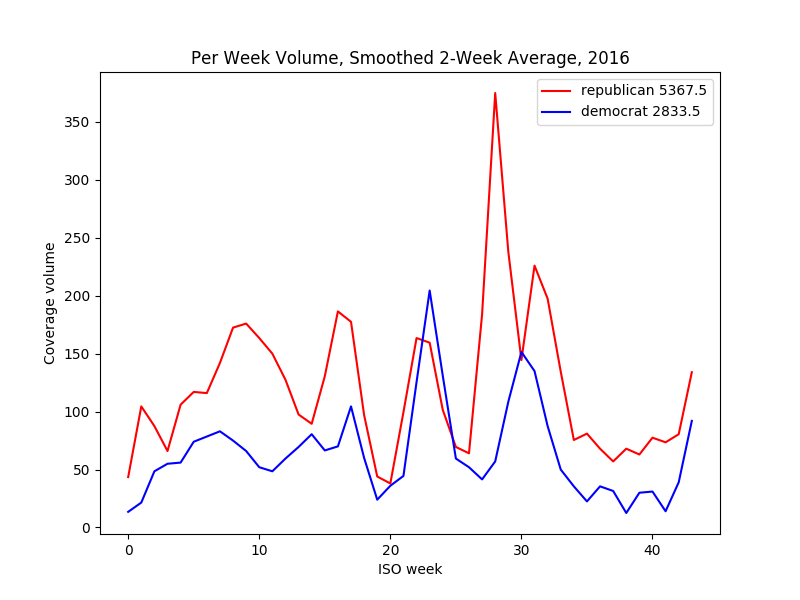
Fox News



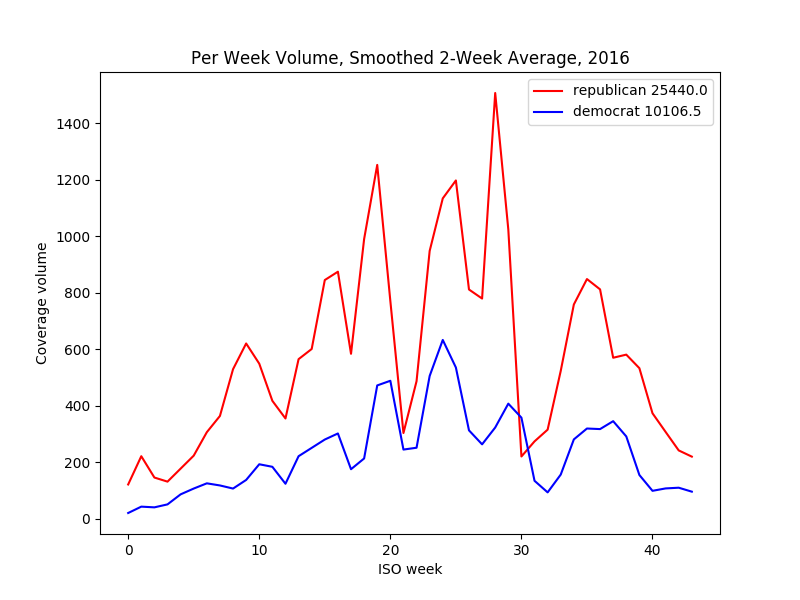
Google News



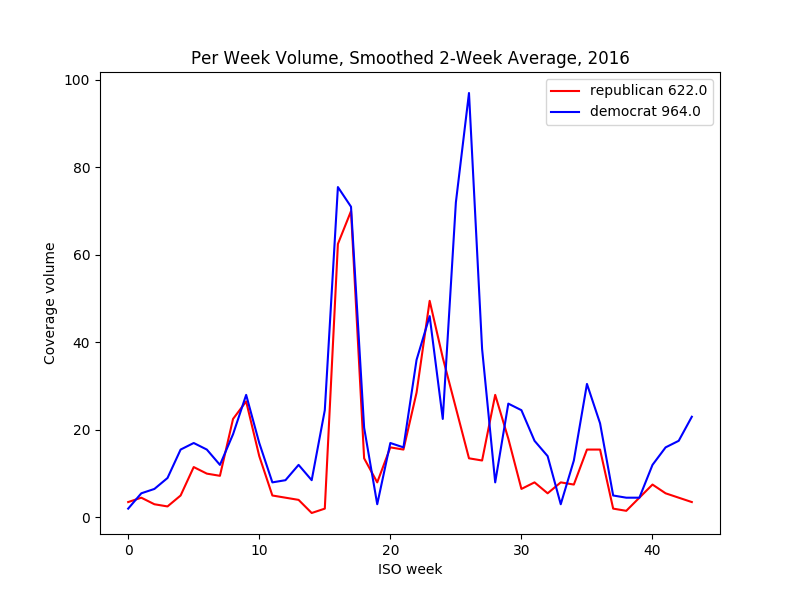
NBC News



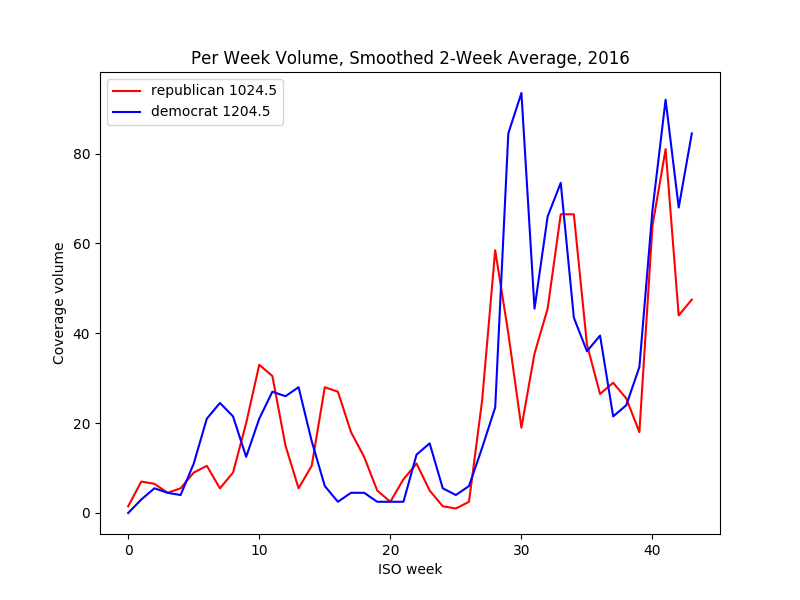
NPR



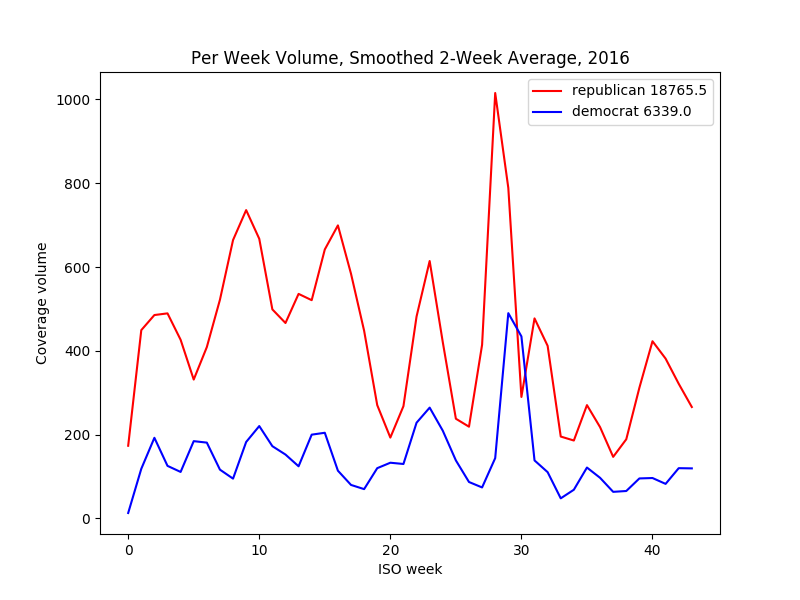
Politico



Russia Today (RT)

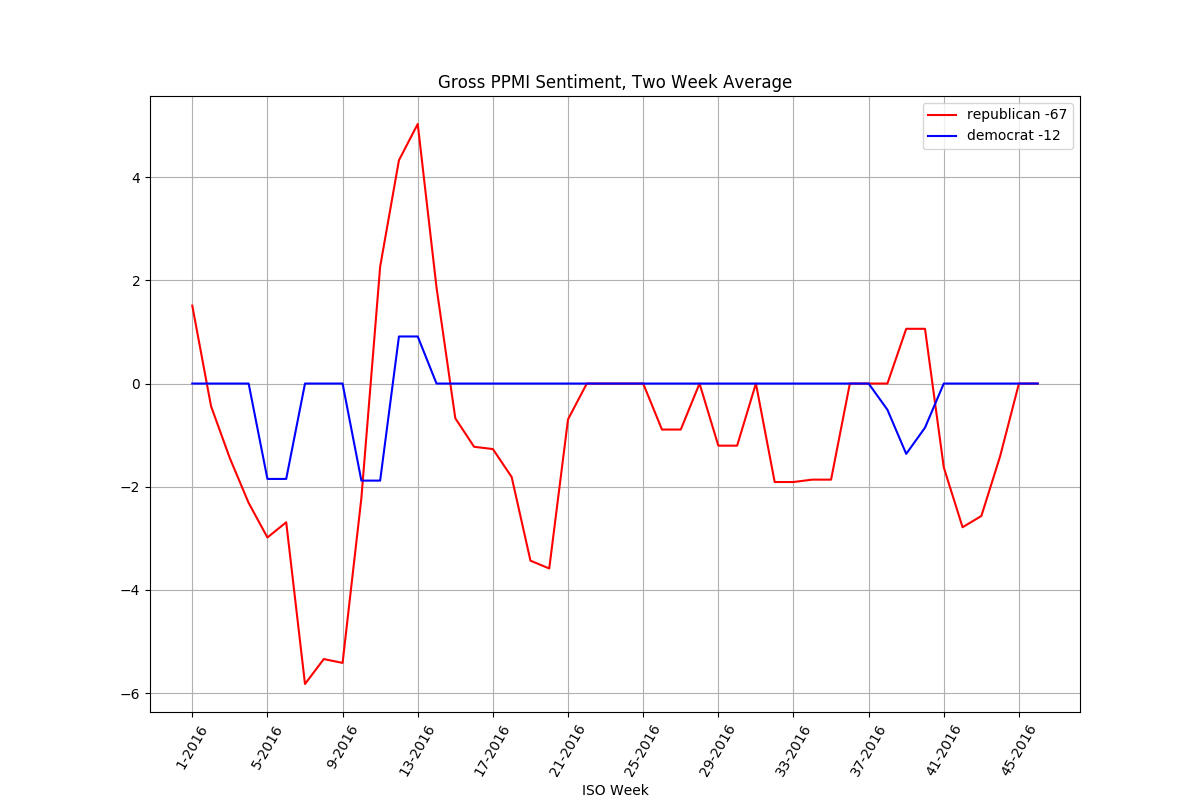


Sputnik News

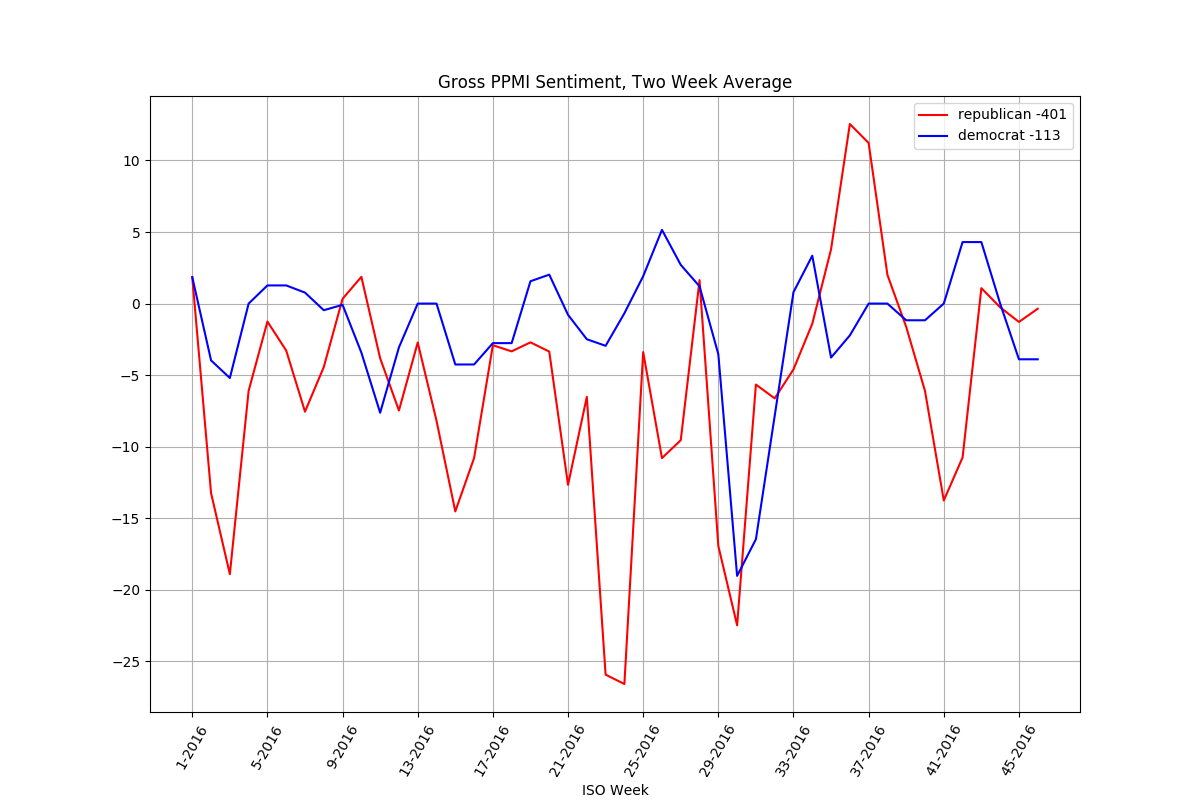


Washington Post

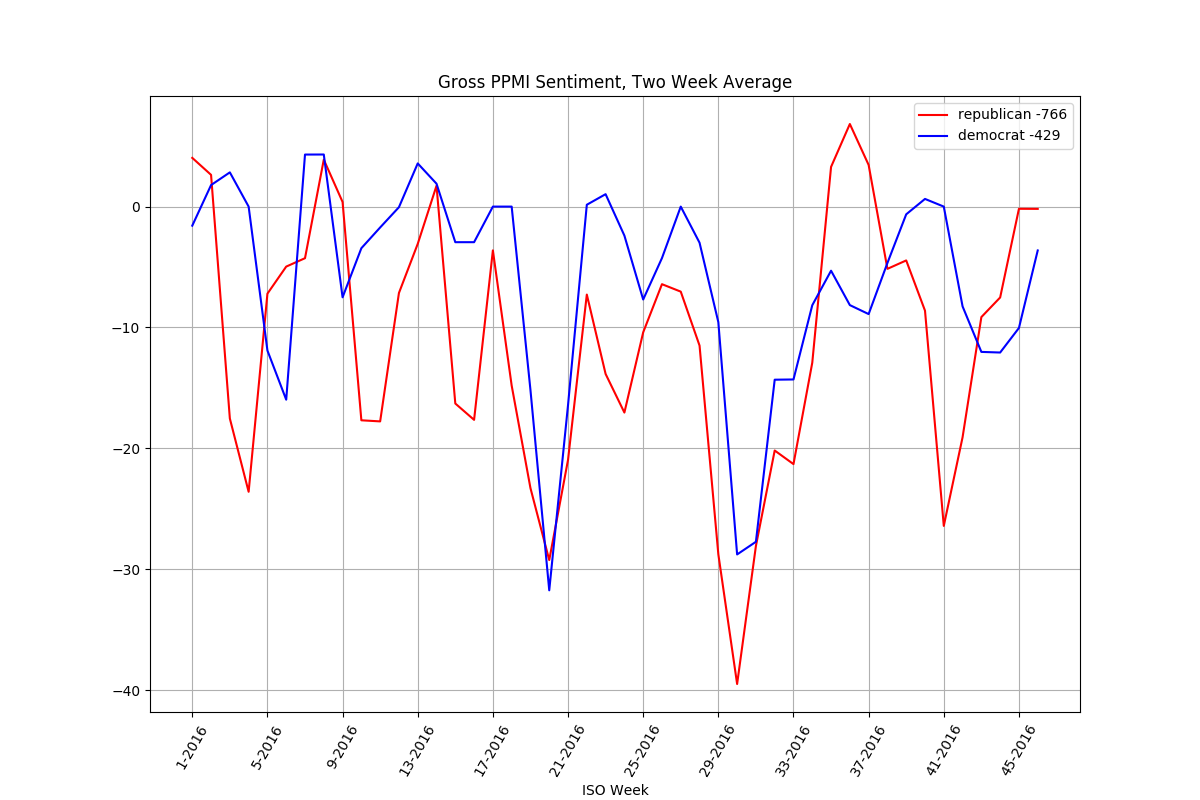
**Appendix IV: Partisan Topic Sentiment**



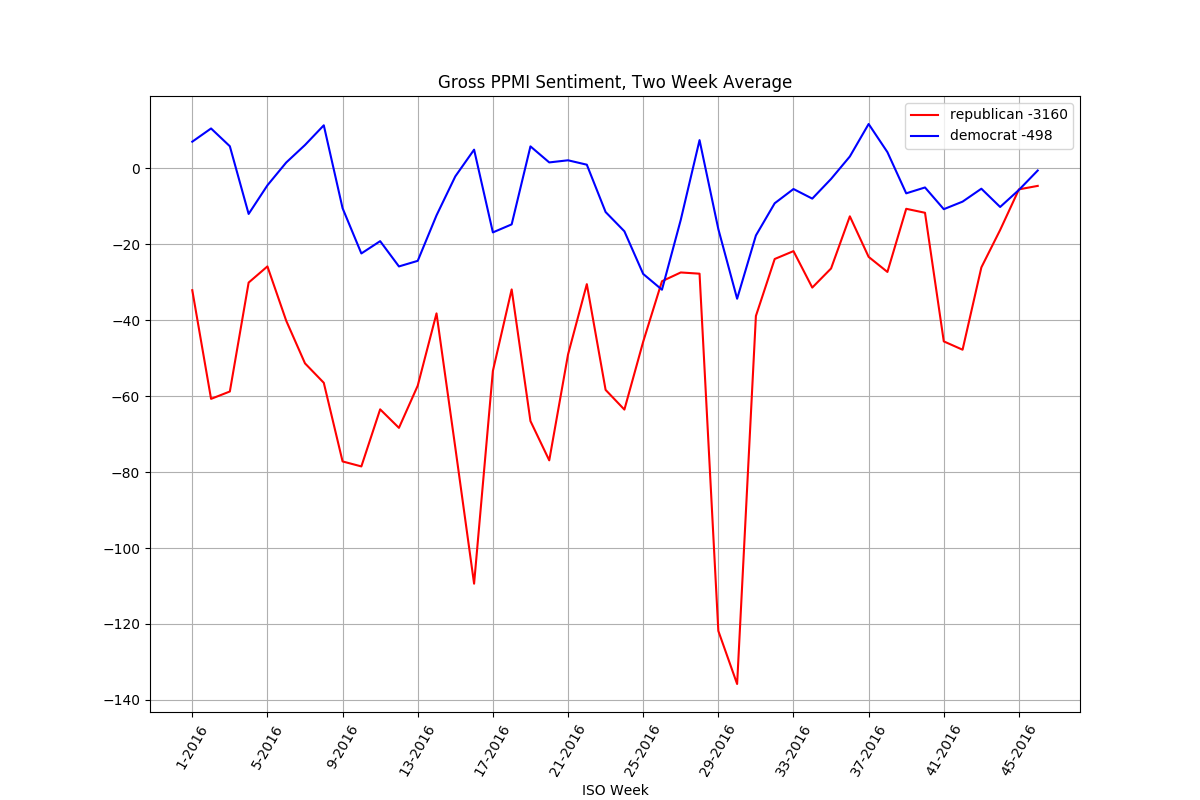
BBC



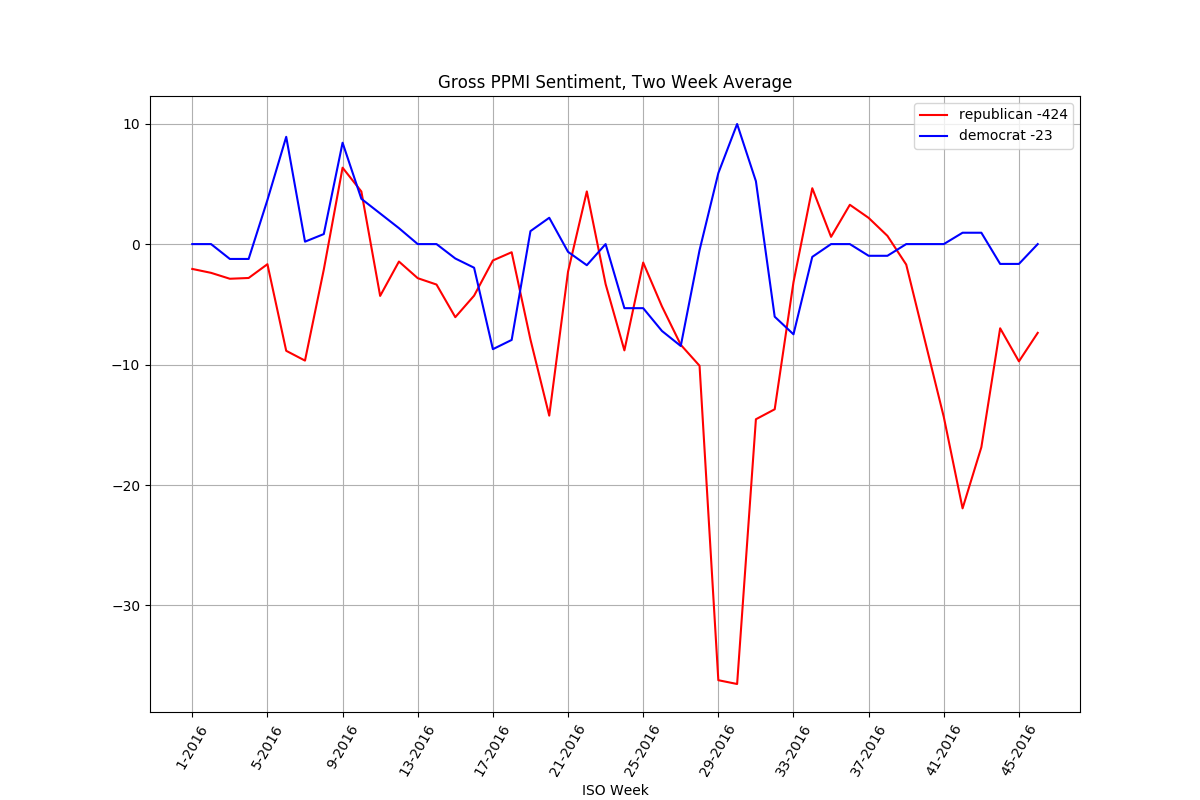
CNN



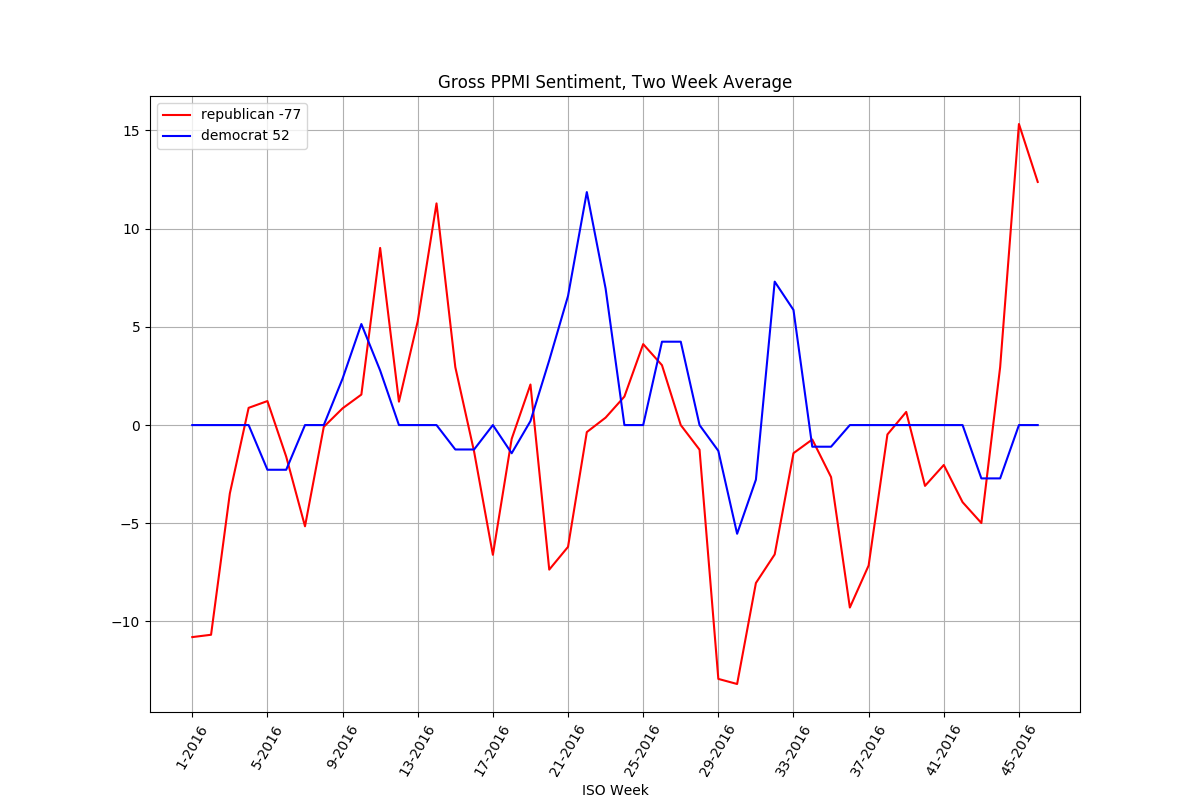
Fox News



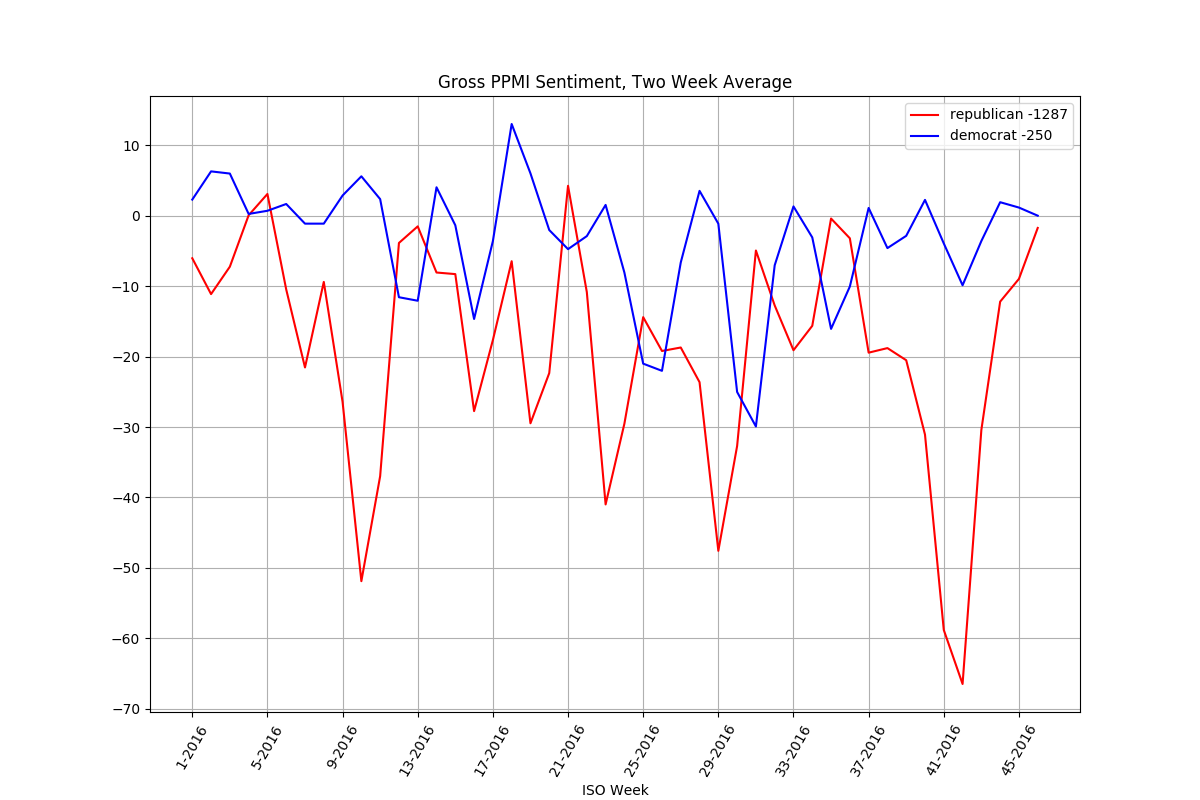
Google News



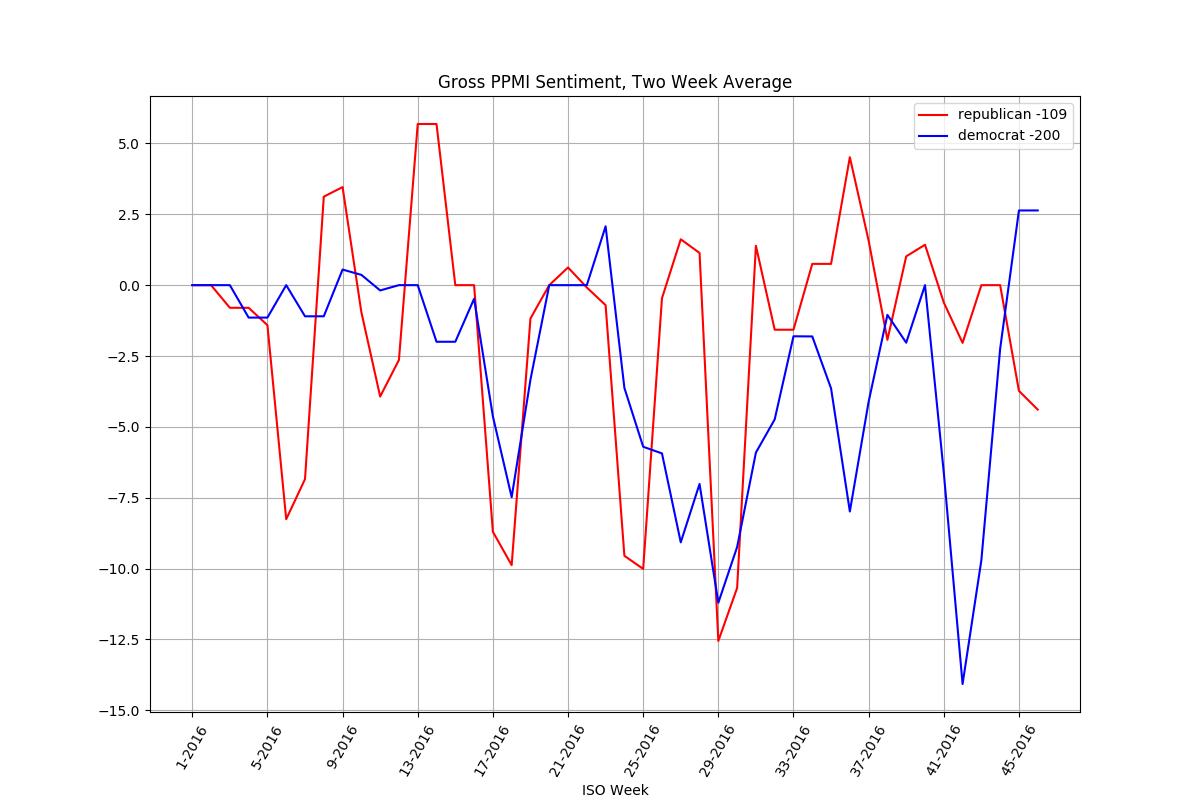
NBC News



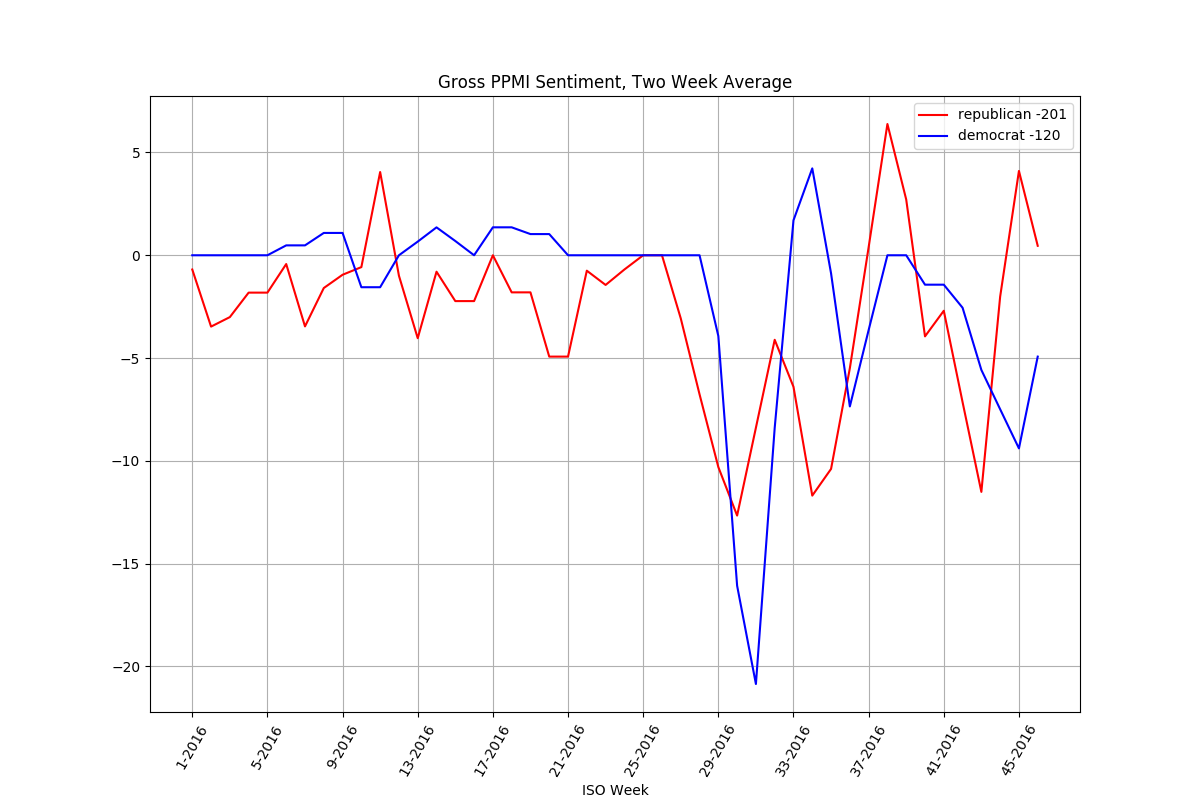
NPR



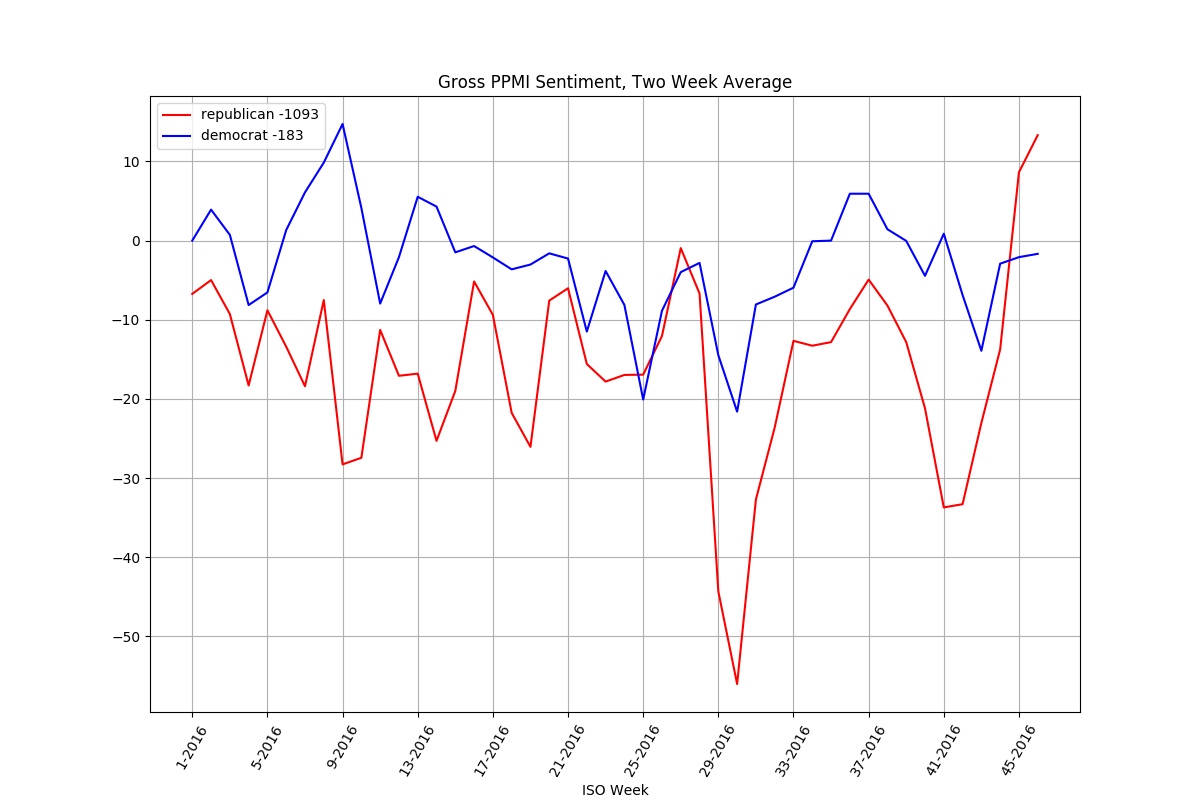
Politico



Russia Today (RT)

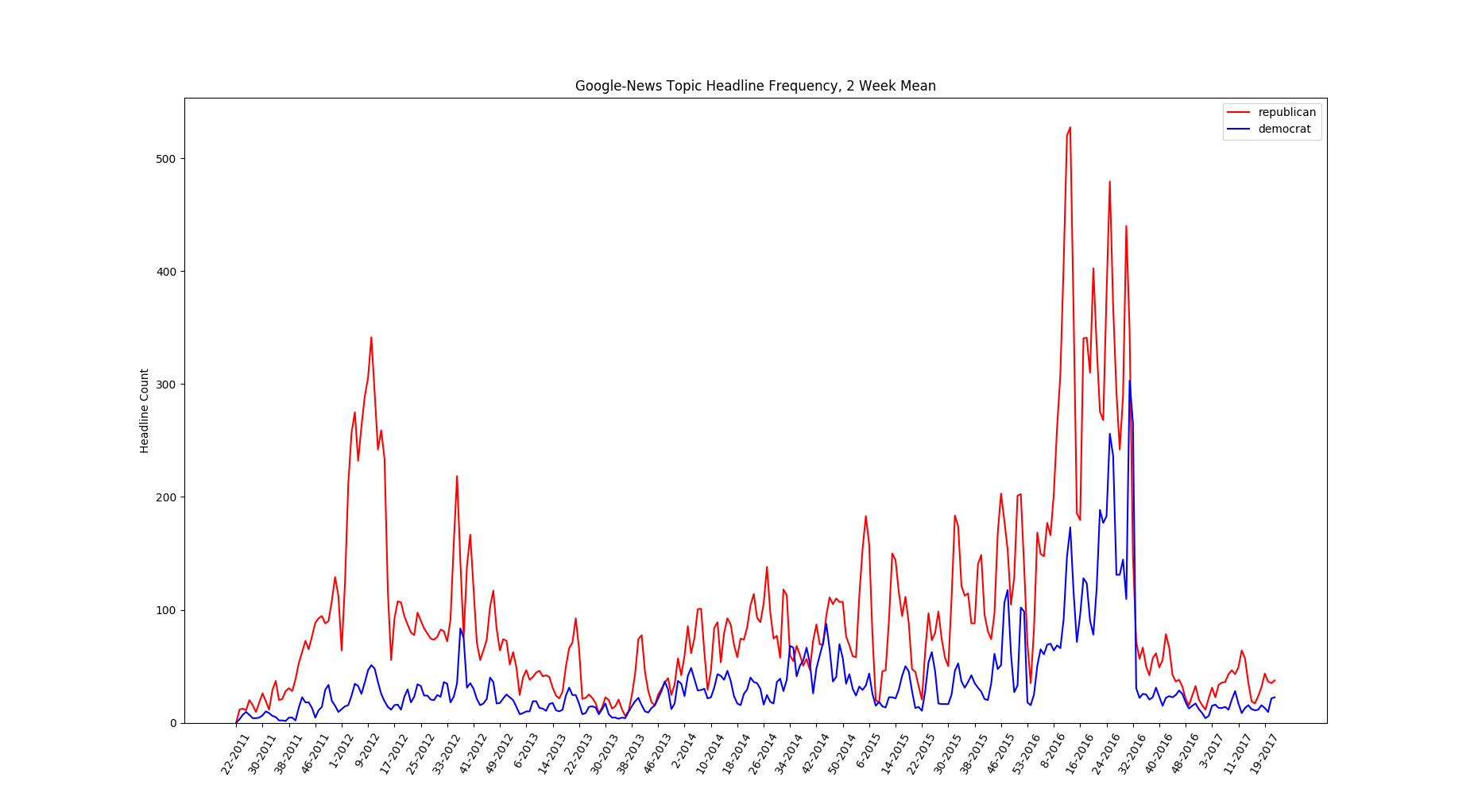


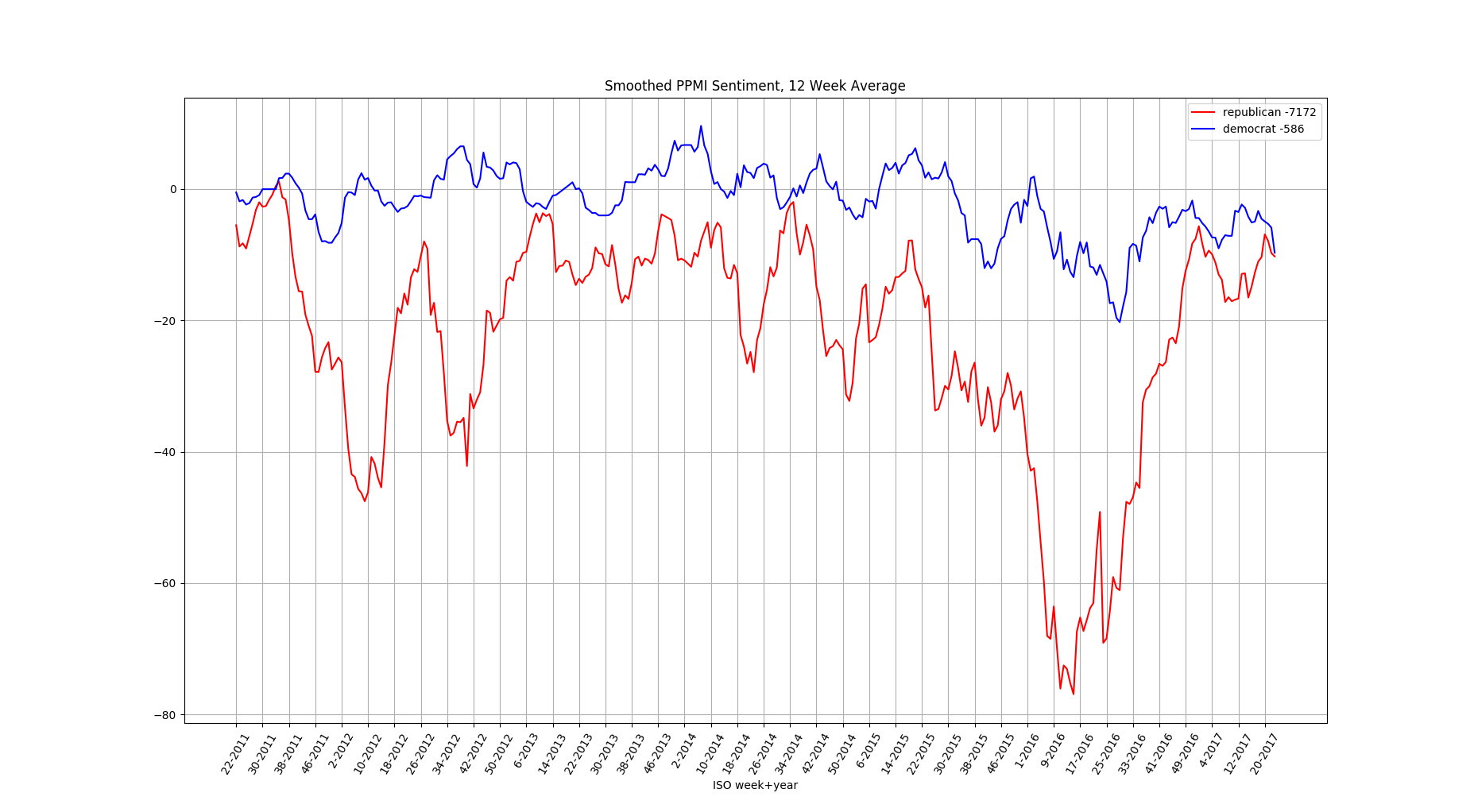
Sputnik News



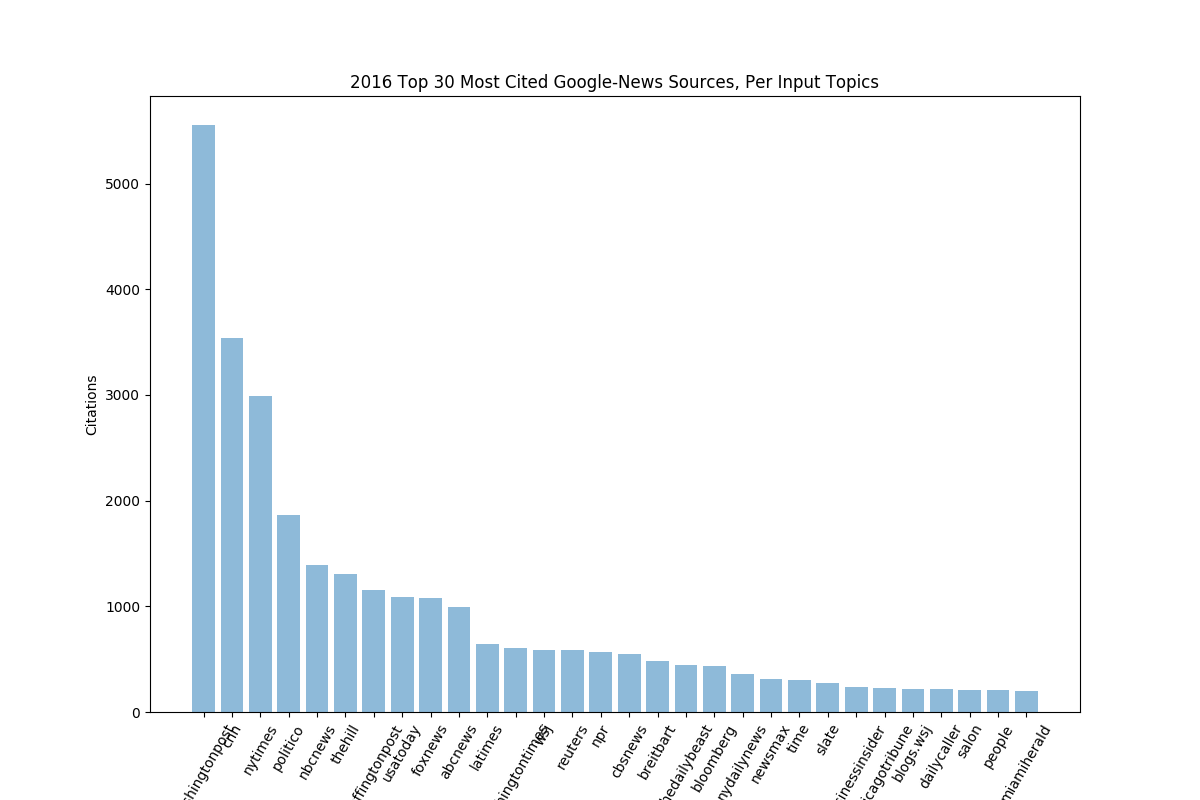
Washington Post

**Appendix V: Google News Partisan Volume, Topic Sentiment, and 2016 Topical Link Histograms**

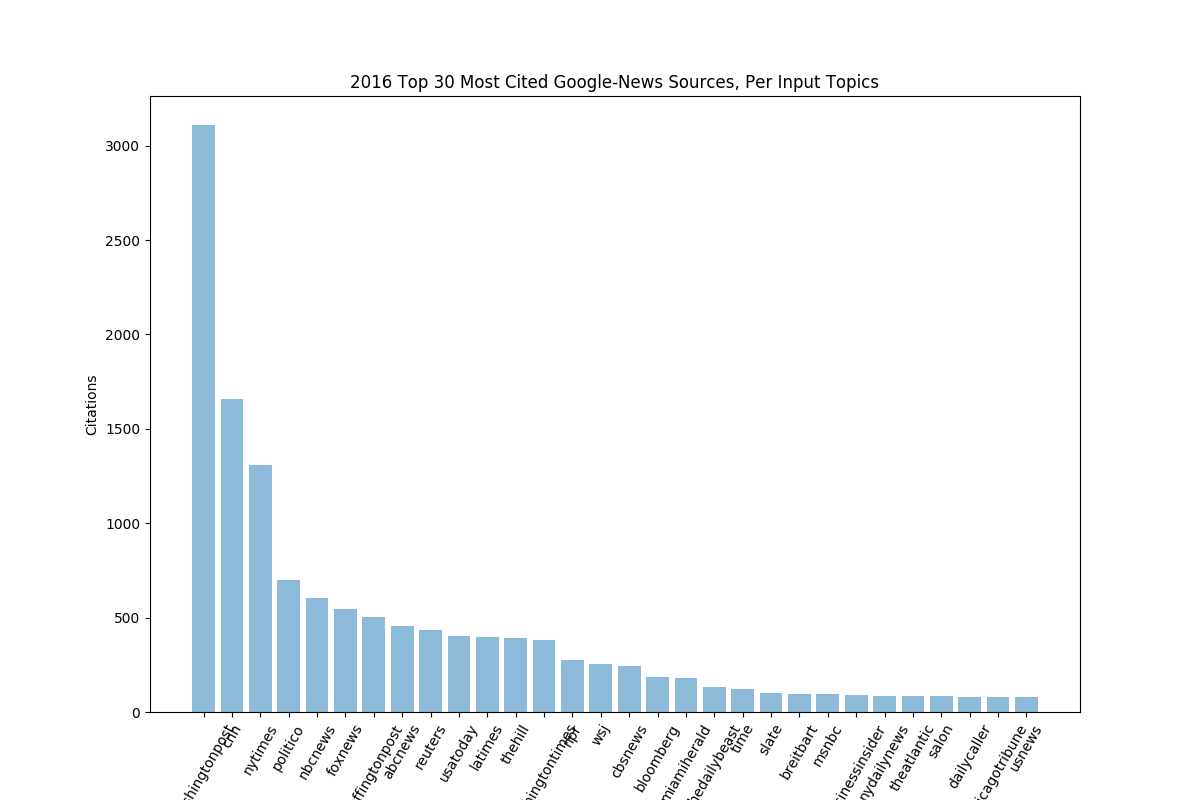


Google News partisan coverage volume, per topic sets T3 and T4, 6-2011 through 6-2017.

Google News extended partisan sentiment, per topic sets T3 and T4. Notably, coverage of Republican topics is far more negative in aggregate. Negativity increases before elections, which occur around ISO weeks 45 for 2012, 2014, and 2016.



Google New link histogram for 2016 coverage of presidential candidates, topic sets T1 and T2.



Google News link histogram per partisan coverage, for topic sets T3 and T4.