Abstract
As people increasingly use social media to learn about the world, how do algorithmic newsfeeds and one’s social graph influence viewpoints on important political issues? We investigate how these viewpoints cluster into echo chambers based on how people consume and respond to news on the world’s most popular social network, Facebook. Using a dataset of almost 200,000 Facebook posts across 15 of the top mainstream news sources, we create a visualization tool to model trendlines of the 5 Facebook Reactions (Love, Angry, Wow, Haha, Sad) over time for any search query across these posts. Users can filter by news source and explore contextual information along the trendlines. In a case study on the 2016 U.S. presidential election, we discover significant polarization between left-leaning and right-leaning sources and observe clustering trendlines across left-leaning sources.

Index Terms: Visualization, Chart.js, Echo Chambers, Social Media, Visualization Design

1 Introduction
With the rise of social media and personalized feeds, people see content that is increasingly tailored to their tastes, preferences, and friend circles. Therefore, they can be isolated from information that disagrees with their viewpoints. Eli Pariser coined the phrase “filter bubble” in 2011 to describe the negative effects of Internet content personalization (for example, via search results customized to one’s browsing history or social media feeds) - namely, that personalization reinforces a state of intellectual isolation. Because users see content that is increasingly tailored to their tastes, preferences, and friend circles, they are isolated from information that disagrees with their viewpoints. In 2017, Bill Gates echoed this sentiment, stating that “[Technologies such as social media] let you go off with like-minded people, so you’re not mixing and sharing and understanding other points of view."

Consuming content in Pariser’s “filter bubble” results in echo chambers where users only see opinions that reflect their own and thus become oblivious or close-minded to contrary, dissenting viewpoints. A prime example of social media echo chambers rose to national attention during the 2016 presidential election, when Donald Trump’s victory surprised many citizens who were potentially consuming news in their own echo chambers and unaware of Trump’s resonance with vital groups of supporters who ultimately gave him the presidency. Thus, as social media only continues to dominate our attention, it’s crucial to understand and curate awareness of its biases we may encounter.

2 Background & Related Work
Prior work such as the Wall Street Journal’s “Blue Feed, Red Feed” seeks to address these echo chambers by providing a qualitative, side-by-side look at liberal and conservative Facebook. The project has many thoughtful features:

- Dynamic updates via hourly fetches from the Facebook API
- Filtering by 8 topics (President Trump, Health Care, Guns, Abortion, Isis, Budget, Executive Order, Immigration)

It also has a few weaknesses:

- Lack of a bird’s-eye view of what’s happening - the burden is on the user to carefully read through the posts to assess each perspective
- Limited search scope of 8 pre-selected topics

Our work seeks to go beyond qualitative comparisons that depend on users to make evaluations and instead produce a bird’s eye visualization of different echo chambers that is both interactive and quantitative. In general, quantitative approaches are historically challenging because they often require sentiment analysis, which remains an active field of research on its own. However, the February 2016 launch of Facebook Reactions allows people to interact with posts through five additional dimensions beyond the original Like: Love, Angry, Wow, Haha, Sad. Reactions offer a lightweight mode of interaction while simultaneously capturing more nuance over how readers feel. We use these as a proxy for sentiment analysis to investigate different readerships’ perspectives towards contemporary issues.

Another study by Garijella, Morales, and Giones investigates political echo chambers in social media. The researchers focus on Twitter and maps clusters of users based on the other users they follow and the content they share. The study observes highly polarized left and right clusters using a graph visualization of nodes and edges. Unlike the WSJ’s visualization, the research does provide quantitative measures of polarity through custom content and ideology scores. However, though the polarization is striking, the study does not present a way for users to explore and interact. Additionally, when considering datasets, Twitter has one advantage over Facebook in that Twitter content is often pre-classified by hashtag, which can be helpful in the disambiguation of search queries. However, sentiment analysis still remains a challenge as Likes and Retweets remain the primary modes of interaction on Twitter.

In a study by Gullani, Yuan, and Saveski, the researchers investigate political echo chambers on Twitter through a visualization tool called Social Mirror. They recruit politically active Twitter users to experiment with their tool, which visually depicts a user’s social network and its ideological fragmentations through a graph of nodes and edges. While the tool does allow interactivity at the granular level by allowing users to explore different clusters, it focuses on the users participating in an echo chamber rather than the content in an echo chamber. We seek to build out a tool that visualizes the latter.

3 Methods
3.1 Dataset
We start with a dataset of over 4 million Facebook posts from 2012-2016 across 15 top mainstream news sources. The dataset was compiled by Patrick Martinchek on data.world and for each news source, includes all posts the source’s Facebook page published
Figure 1: Number of posts by news source. Biases measured by AllSides.

during the timeframe. As Facebook Reactions launched globally in February 2016, we take the subset from then onwards, giving us 197,747 posts [Figures 1, 2].

3.2 User Interactivity
Users have the option to specify 3 fields: search query, date range, and news sources [Figures 3, 4]. Because we allow users to search for any subject rather than limiting them to a preset selection, we want to ensure we capture as many relevant posts as possible. To do so, we improve upon simple string matching by tabulating the most frequent words appearing across all titles in the dataset. We create a mapping of common entities to perform naïve entity recognition on search queries. For example, searching for “Donald Trump” will also surface posts with “Trump” and “President Trump.” Finally, the dataset is indexed by these tags [Figure 5].

3.3 Calculating Reaction Trendlines
Given a user’s specification of a query, dates, and news sources, we aggregate the monthly reaction count across all posts for each publication. Because each publication has varying post volume, comparing raw reaction counts would be misleading. Additionally, from experimentation, we found that daily and weekly aggregations were too noisy to see significant trends and monthly aggregations were the most insightful. Therefore, we calculate the monthly ratio of each reaction. For example, searching for “Donald Trump” will also surface posts with “Trump” and “President Trump.” Finally, the dataset is indexed by these tags [Figure 5].

3.4 Technical Details
We use Chart.js to produce visualizations and d3 to parse our dataset. We experimented with parsing and filtering data on the server-side using Google Cloud Functions and the Firebase Database, but found that queries took upwards of 30 seconds to return. While server-side computation (with the necessary optimizations) is more suitable for the larger-scale vision of this project once the dataset grows, we use client-side parsing for the approximately 200,000 posts we currently have. In practice, client-side operations have produced no noticeable
delay in responding to user queries for our current dataset.

### 3.5 Design Tradeoffs

We use color to represent each echo chamber on the political spectrum (blue for left, gray for center, and red for right). We experimented with giving each news source its own color, but found it became too visually noisy. Having distinct color buckets of left, right, and center was more intuitive to grasp the echo chambers. However, we still individually distinguish news sources by plotting their trendlines with different point shapes.

### 4 RESULTS

#### 4.1 Case Study: 2016 Presidential Election

As a case study, we look at the 2016 U.S. Presidential Election. This election prompted the surge of discussion on social media’s role in delivering news, especially with regards to Facebook’s Ad platform which was manipulated by Russian interest groups to sway the election in Trump’s favor. It also prompted the rise of “fake news” - interestingly, searching for fake news in our visualization produces a relatively flat graph until the latter half of 2016 when Trump won the election [Figure 8]. Persons of interest in this election include Donald Trump (the Republican nominee), Hillary Clinton (the Democratic nominee), Bernie Sanders (Democratic candidate), and Barack Obama (President in office).

##### 4.1.1 Searching for “Donald Trump”

Looking at reaction trendlines for a left-leaning source (NY Times) and a right-leaning source (Fox), the diverging sentiment towards Trump, particularly in November 2016 when he wins the election, is immediately striking. From the time period of February to November 2016, the Fox readership has generally interacted with high “Anger” ratios towards posts, but the ratio reaches an all-time low when Trump wins. Similarly, the “Love” ratio from the Fox readership reaches an all-time high. In contrast, we see this trend virtually reversed for the New York Times. In November 2016, the “Angry” ratio of the New York Times readership reaches an all-time high while the ‘Love’ ratio dips [Figure 9].

##### 4.1.2 Searching for “Barack Obama”

We also see polarization when examining reactions towards then-president Obama. The “Love” and “Angry” trendlines across the two sources are nearly mirror opposites with zero points of intersection. The separation is visually striking and illuminates these separate channels of communication [Figure 10].

##### 4.1.3 Searching for “Hillary Clinton”

We see similar polarization for nominee Hillary Clinton. Notably, the only irregularity is in April 2016 of the New York Times “Angry” trendline which surges above Fox trendline. Hovering over the data point, we observe that in April, Clinton won the New York Primary. Thus the surge could have been a result of the internal division of Democratic constituents between Clinton and Sanders.

It is additionally striking to stack multiple left-leaning news sources. These sources are independent and publish their own posts, yet there is remarkable similarity in the trendlines. They share the same dips, surges, and patterns throughout, while the red-leaning source remains distinctly separate [Figure 11].

For comparison, we graph the Wall Street Journal’s trendline, a center-leaning source, and discover that it falls approximately in the middle between the red and blue trendlines [Figure 12].

##### 4.1.4 Searching for “Bernie Sanders”

Here, the polarization between red and blue is not as striking as it is for Obama and Clinton. For example, we observe internal blue polarization from the New York Times “Angry” trendline in July 2016. Hovering over the data point, we see that in July, Wikileaks released DNC emails showing bias against Sanders and the party’s favoritism towards Hillary, a scandal which prompted the chair to resign [Figure 13].

#### 4.2 Interactive Visualization

In addition to these example graphs, users can search for anything on their own by specifying a search query, date range, and news sources [Figure 6].

#### 4.3 Custom Echo Chamber

We also devise a methodology for calculating any given user’s echo chamber from their Facebook social graph. To compute this trendline, we look at each publication’s Facebook page and tabulate the number of the user’s friends who liked that page. Each page thus has a weight biased by the user’s friend circle on Facebook. We assume a user’s friends are more likely to read, comment on, and share posts from pages they follow, therefore causing these posts...
to appear in the user’s feed. Finally, we use these friend-influenced page weights to compute a weighted average of the reaction ratios across the pages. We demonstrate the author’s own echo chamber pulled from Facebook’s social graph and find that the New York Times (left), NPR (center), and WSJ (center) influence the echo chamber most prominently [Figure 7].

5 Discussion

By visualizing trendlines across politically-biased news sources, we achieve a holistic pulse across any given search query segmentable by news source. While the trendlines are not directly mappable to sentiment (e.g. the “Love” trendline cannot be taken at face value for the ratio of people who respond with “Love” towards a search query), the resulting visualizations from comparing these trendlines across news sources yields insightful results. We see that left-leaning readerships respond similarly to politically charged topics (such as “Hillary Clinton”) while right-leaning readerships hold distinctly separate trendlines from the aggregated echo chamber of the left. Stacking the left-leaning trendlines (in blue) on top of each other helps visualize the echo chamber as seen in Figure 11.

It is additionally striking that left-leaning mainstream media sources outweigh right-leaning sources (9 vs. 2 respectively in our dataset of top 15 media sources). With their biases (as highlighted by these visualizations), there is a risk of misrepresenting how the entire nation feels towards a subject. As a result, trending posts and the echo chambers they bounce around in may reflect different perspectives from the ground-truth, which could have contributed to the surprise felt by some demographics when Trump won the election in 2016. By providing a tool to visualize echo chambers in products people use for hours each day, we hope users can be more cognizant of the biases, however subtle, present in the content they consume.

6 Future Work

Looking forward, we can yield a more rich and dynamic dataset across a wider time range by periodically scraping Facebook to update the posts. We can also build a smarter query system to disambiguate search terms and surface more accurate pulses on how people are feeling. Another promising avenue is incorporating additional sources popular amongst Gen Z (such as Buzzfeed) and other social media data from Twitter and Instagram.

We also hope to add richer contextual information beyond the number of posts and top reacted articles. Instead of a presenting a text blurb upon a user’s hover, we can render the Facebook post adjacent to the data point, akin to the qualitative information found in WSJ’s “Blue Feed, Red Feed” visualization.

Through our work, our goal is to continue raising awareness about homogenity in the content people consume on social media. We hope that exploring different Facebook Reaction trendlines adds context to each publication’s readership demographic, making us more cognizant of the biases surrounding us and the echo chambers in our feeds.

<table>
<thead>
<tr>
<th>Facebook Page</th>
<th>Author’s Friend Likes of Page</th>
<th>Page Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEFT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABC</td>
<td>8</td>
<td>0.0097</td>
</tr>
<tr>
<td>CBS</td>
<td>3</td>
<td>0.0036</td>
</tr>
<tr>
<td>CNN</td>
<td>55</td>
<td>0.067</td>
</tr>
<tr>
<td>Huffington Post</td>
<td>42</td>
<td>0.051</td>
</tr>
<tr>
<td>LA Times</td>
<td>41</td>
<td>0.050</td>
</tr>
<tr>
<td>NBC</td>
<td>12</td>
<td>0.015</td>
</tr>
<tr>
<td>NY Times</td>
<td>243</td>
<td>0.29</td>
</tr>
<tr>
<td>Time</td>
<td>59</td>
<td>0.072</td>
</tr>
<tr>
<td>Washington Post</td>
<td>45</td>
<td>0.055</td>
</tr>
<tr>
<td>CENTER</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BBC</td>
<td>68</td>
<td>0.082</td>
</tr>
<tr>
<td>NPR</td>
<td>129</td>
<td>0.16</td>
</tr>
<tr>
<td>USA Today</td>
<td>6</td>
<td>0.0073</td>
</tr>
<tr>
<td>WSJ</td>
<td>107</td>
<td>0.13</td>
</tr>
<tr>
<td>RIGHT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fox News</td>
<td>6</td>
<td>0.0073</td>
</tr>
<tr>
<td>Fox and Friends</td>
<td>1</td>
<td>0.0012</td>
</tr>
</tbody>
</table>

Figure 7: Author’s own echo chamber.
Figure 8: Searching for fake news produces few results until the latter half of 2016.

Figure 9: Searching for Donald Trump. We observe the opposing reactions of left (blue) and right (red) when he wins the election in November 2016.
Figure 10: Searching for Barack Obama. We observe nearly mirror-opposite polarity between left (blue) and right (red).

Figure 11: Searching for Hillary Clinton, we see the homogeneity of the left-leaning sources (in blue).
Figure 12: Again searching for Hillary Clinton, we observe center sources fall approximately between left and right sources.

Figure 13: Searching for Bernie Sanders, we observe less striking polarization and more blue internal polarization.
REFERENCES