RUTGERS

The Spread of Facebook Misinformation

Samantha Lee - Fall 2021

DATA SCIENCE FOR POLITICAL SCIENCE

ABSTRACT

Social media serve as primary sources of news for many people, and researchers study them to examine political literacy, polarization, and the factors that contribute to the spread of misinformation. Evidence suggests that eye-catching posts create bias in favor of fake news, clickbait, and graphics, and that conservative people demonstrate high amounts of in-group bias. A dataset of Facebook posts shows that the distribution of activity leans toward articles that are false, conservative, or visual.

BACKGROUND

Research Question. What factors correlate to the spread of misinformation?

Importance. Misinformation has the power to create contention, influence public opinion, undermine democracy, and uplift malicious or ignorant parties. Its recent momentum has given reason to investigate the quality and correlations of news on popular sites like Facebook.

Background. Previous research suggests:

- News sources report falsehoods to increase engagement (Zannettou et al., 2019).
- Conservatives more frequently conform to in-group beliefs, even if false (Jost et al., 2018).
- Graphics add visual bias to media and can signal importance (Grabe and Bucy, 2009).

Objective. Considering background information, the following relationships can reveal the qualities of misinformation and its momentum:

- Activity vs. truth by partisanship
- Activity vs. truth by media type

DATA AND METHODS

DATA

BuzzFeed News (2016) collected posts from a variety of political Facebook pages during the 2016 US presidential election.

- The dataset has 2282 rows and 14 columns.
- 'Category' indicates partisanship: liberal (left), nonpartisan (mainstream), and conservative (right).
- 'Post.Type' indicates the medium of the post: link, photo, text, or video.
- 'Rating' indicates the truthfulness of the content: mostly true (3), a mixture of true and false (2), mostly false (1), and no factual content (0).
- 'share_count', 'reaction_count', and 'comment_count' (fig. 1) measure activity.

```
 share_count      reaction_count     comment_count
Min.   :      1   Min.   :     2.0   Min.   :     0.0
1st Qu.:     24   1st Qu.:   149.0   1st Qu.:    37.0
Median :     96   Median :   545.5   Median :   131.5
Mean   :   4045   Mean   :  5364.3   Mean   :   516.1
3rd Qu.:    739   3rd Qu.:  2416.8   3rd Qu.:   390.2
Max.   :1088995   Max.   :456458.0   Max.   :159047.0
NA's   :70        NA's   :2          NA's   :2
```

Fig. 1. Metrics of activity (and therefore spread).

METHODS

- A scatter plot can reveal correlations between truth and activity to determine if the quality of truthfulness itself is associated with spread.
  - Dependent variable: activity (measured as the sum of shares, reactions, and comments)
  - Independent variable: truth rating
  - Confounding variables: partisanship, media
- Subsetting data and observing mean and median truth ratings and activity by partisanship and media can additionally reveal how the other factors may act as confounding variables.
- Note: a scatter plot can handle the wide range of activity values; a regression would not translate into a constructive visual.

RESULTS

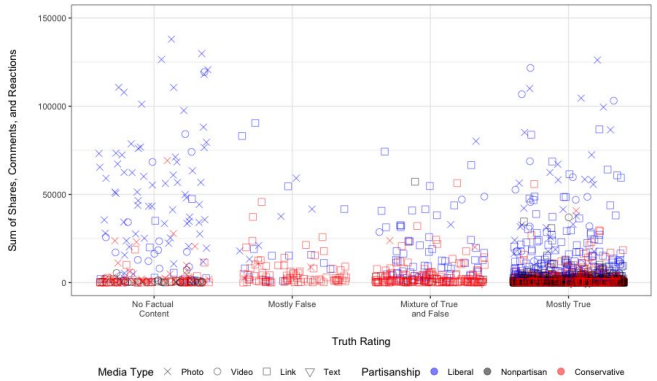


Fig. 2. A plot of activity vs. truth rating by partisanship and media (activity above 15,000 not pictured)

Findings:

- Low truth ratings correlate most to high activity
  - Mean, median for mostly false: (10014, 4006)
  - Mean. median for no factual information: (42771, 3116)
- High activity correlates most to…
  - Liberal partisanship (mean = 42287, median = 16380)
  - Photos (mean = 42614, median = 23820)
- Low truth ratings correlate most to…
  - Conservative partisanship (mean = 2.068, median = 2)
  - Liberal partisanship (mean = 2.023, median = 3)
  - Photos (mean = 1.014, median = 0)

Brief Discussion:

- Fake news is indeed associated with increased engagement
- People tended to interact with visual posts that had images
- Overall, instances of misinformation was relatively infrequent
- People of either partisanship engaged with fake news, conservatives slightly more often

CONCLUSION

Facebook data shows that falsehood itself corresponds to high user engagement, supporting prior research. Visual forms of news additionally garner more attention, as do liberal sources. Generally, partisans are more heavily associated with fake news than nonpartisans. These findings can help experts predict and mitigate the spread of misinformation. More research can be done to further explore the four-dimensional relationship between truth, activity, partisanship, and media and determine causation rather than mere correlation.

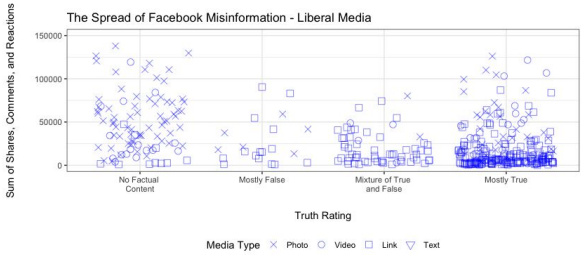
The Spread of Facebook Misinformation - Appendix

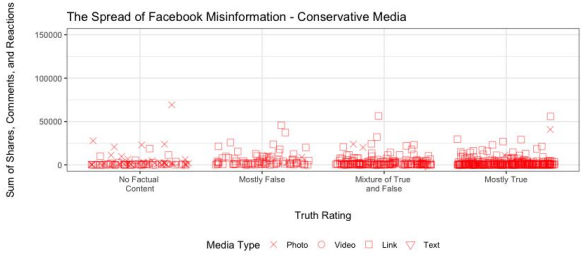
Samantha Lee - Fall 2021

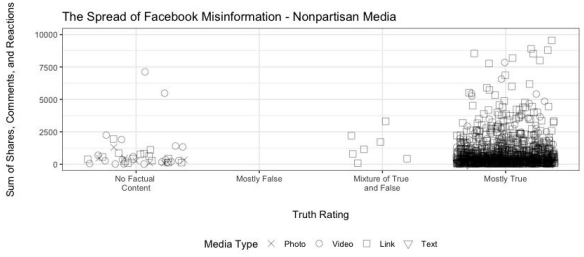
RUTGERS

DATA SCIENCE FOR POLITICAL SCIENCE

ADDITIONAL PLOTS







The plots to the left are present the same information as the main scatter plot but break down the data into partisanship. The following patterns are made more clear:

- Liberal content yields higher activity overall.
- Truth amongst conservatives sources is the most dispersed.
- Truth amongst liberal sources is generally dispersed, but most content is mostly true.
- Nonpartisan content is rarely false.

WORKS CITED

Jost, J. T., van der Linden, S., Panagopoulos, C., & Hardin, C. D. (2018). Ideological asymmetries in conformity, desire for shared reality, and the spread of misinformation. *Current Opinion in Psychology*, *23*, 77–83. https://doi.org/10.1016/j.copsyc.2018.01.003

Grabe, M. E., & Bucy, E. P. (2009). Visual Bias. In *Image Bite Politics*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195372076.003.0005

Silverman, Craig. *Hyperpartisan Facebook Pages Are Publishing False And Misleading Information At An Alarming Rate*. (2016). BuzzFeed News. Retrieved December 16, 2021, from https://www.buzzfeednews.com/article/craigsilverman/partisan-fb-pages-analysis

Zannettou, S., Sirivianos, M., Blackburn, J., & Kourtellis, N. (2019). The Web of False Information: Rumors, Fake News, Hoaxes, Clickbait, and Various Other Shenanigans. *Journal of Data and Information Quality*, *11*(3), 10:1-10:37. https://doi.org/10.1145/3309699

LEARN MORE

For more details, as well as interactive versions of these plots, I highly recommend you visit the following RMarkdown report: https://rpubs.com/samanthallee/facebook-misinformation

A Github repository is also available: https://github.com/SamanthaLLee/facebook-misinformation