

Changes in Online Political Advertising

Aaron Spielman, Paul Hwang, Fehmi Neffati, Nikolas Santamaria, Joshua Ingram

New College of Florida
IDC 5131: Distributed Computing





About the Data





Two Datasets: ProPublica and Ad Observer

- ProPublica collected ad data from mid-2017 to mid-2019.
- NYU's Ad Observer collected data from December 2020 to February 2022, with a gap between March 2021 to July 2021.
- Both tools rely on browser extensions installed by volunteers to gather data.

**Are there any trends
observed in these online
political advertisements?**

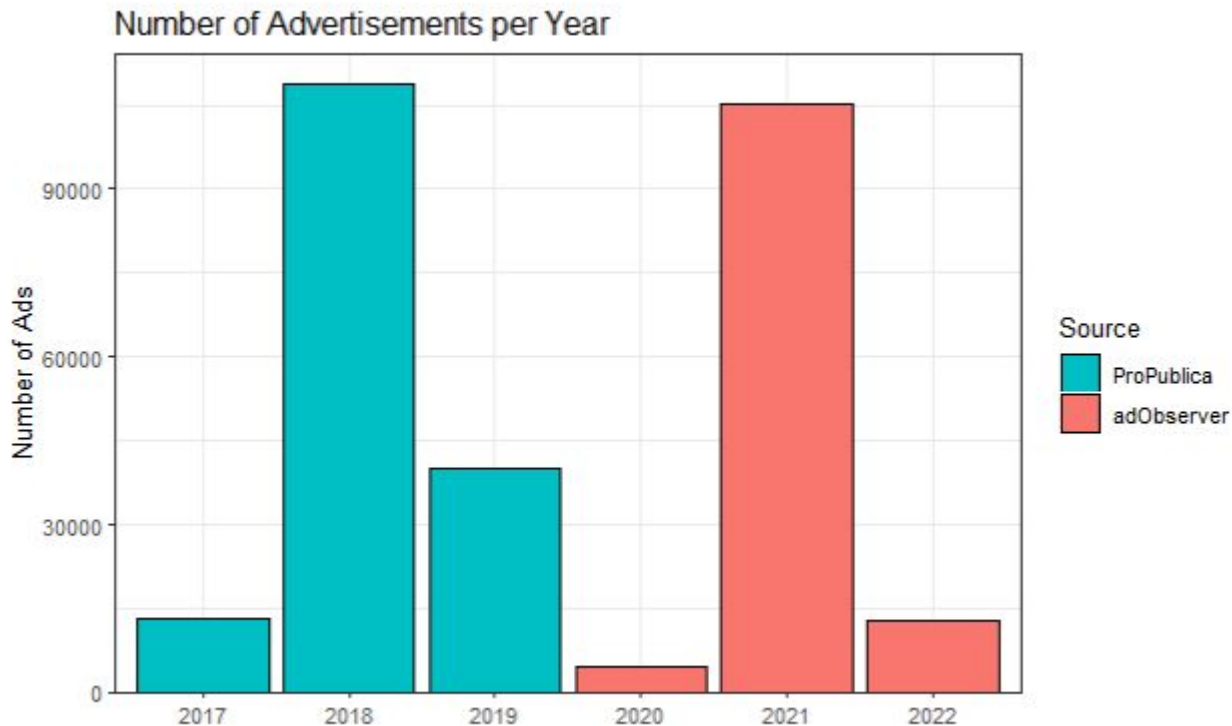


Analysis



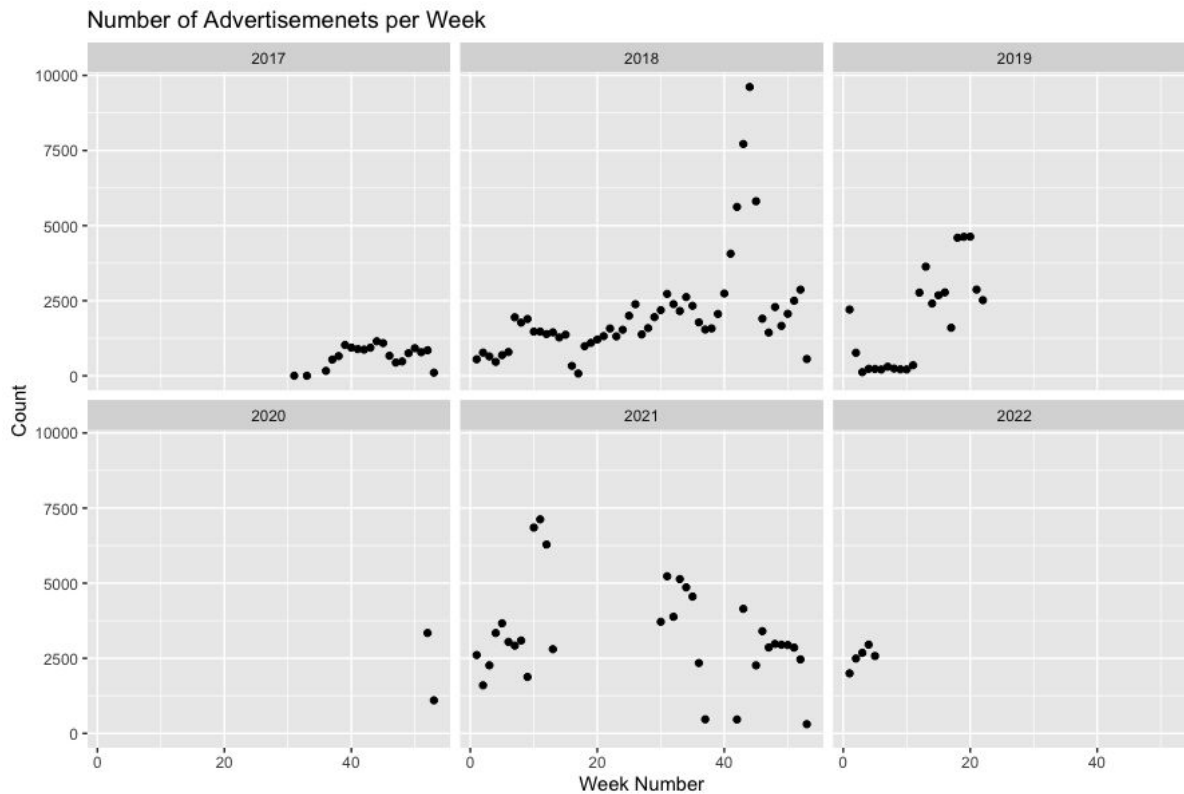


Frequency of Advertisements

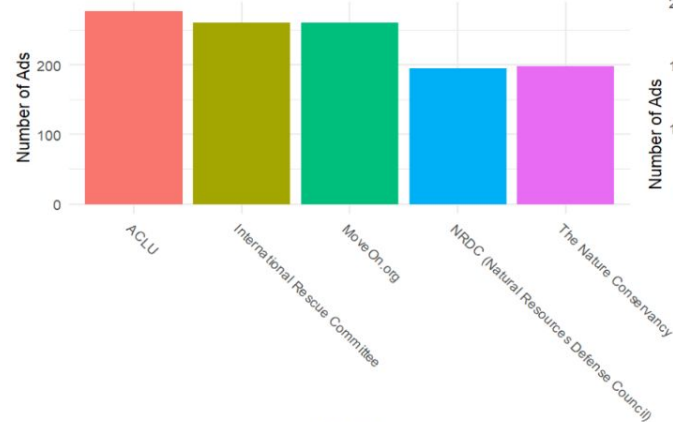




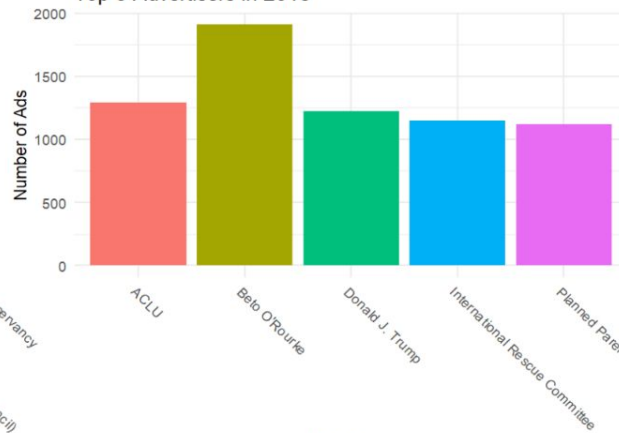
Frequency of Advertisements



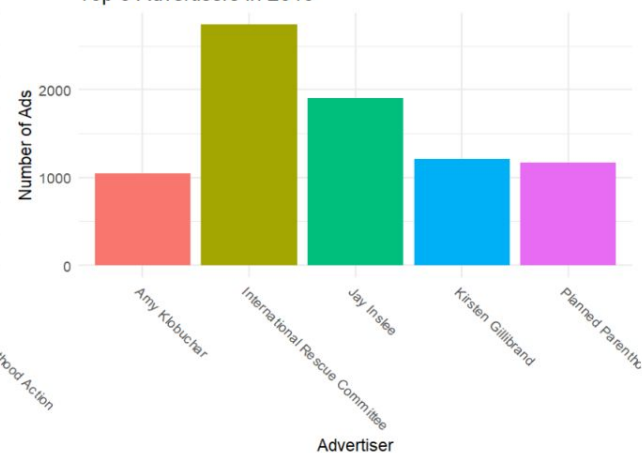
Top 5 Advertisers in 2017



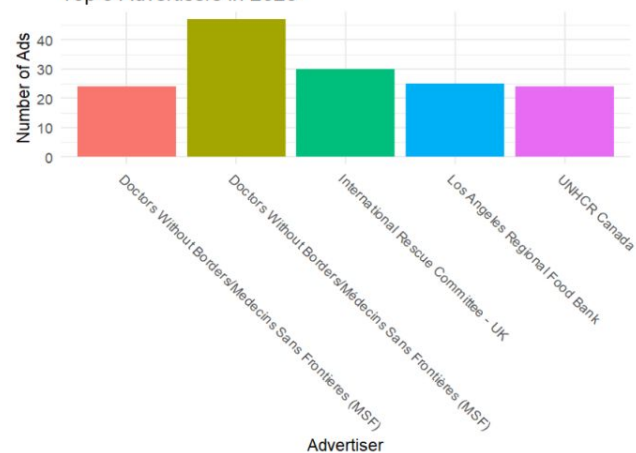
Top 5 Advertisers in 2018



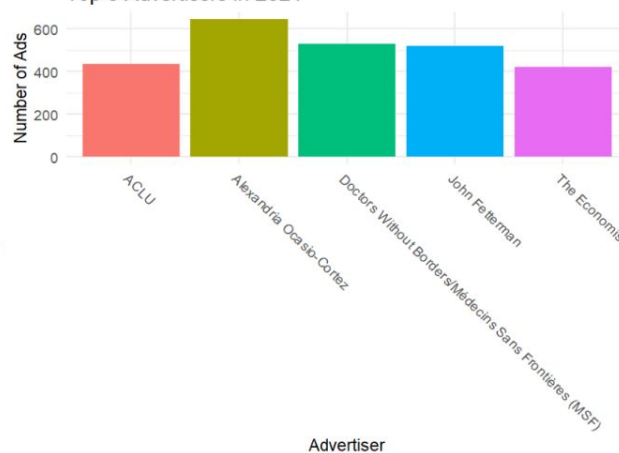
Top 5 Advertisers in 2019



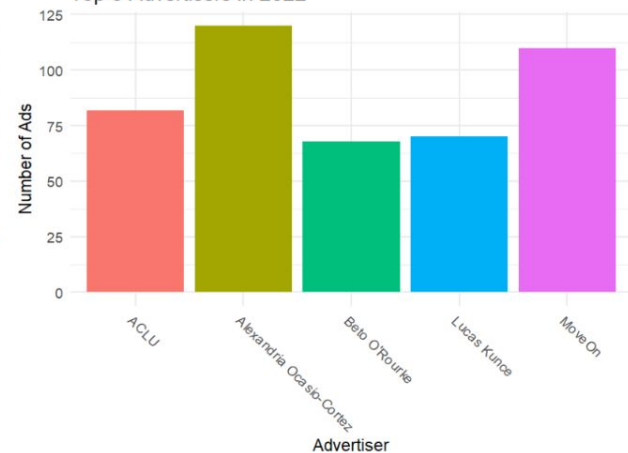
Top 5 Advertisers in 2020



Top 5 Advertisers in 2021

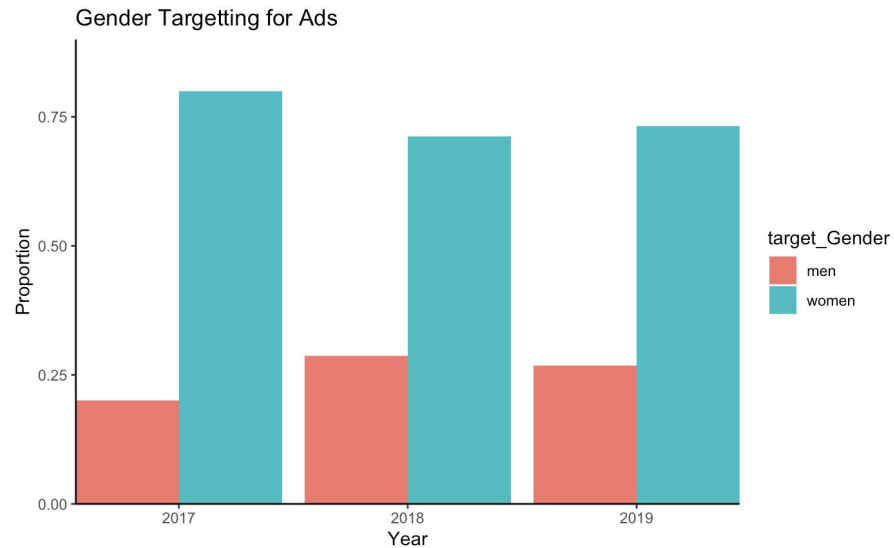
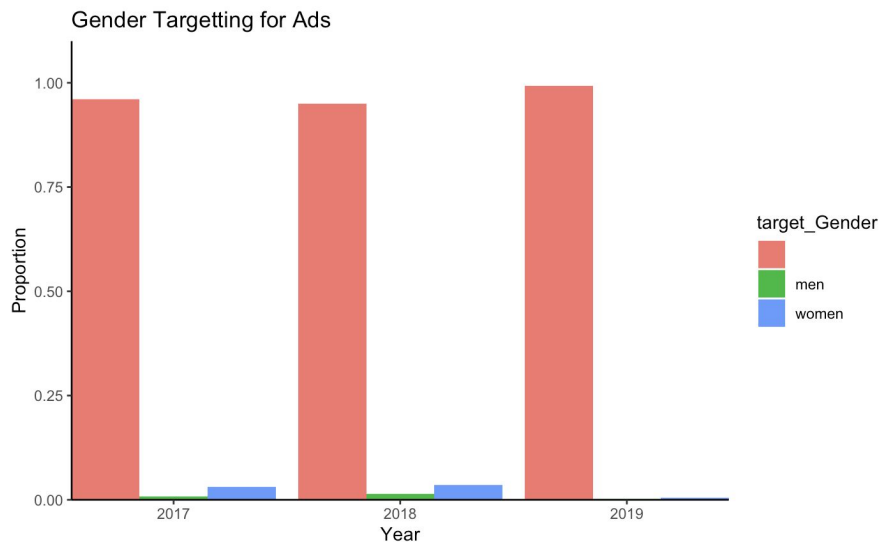


Top 5 Advertisers in 2022



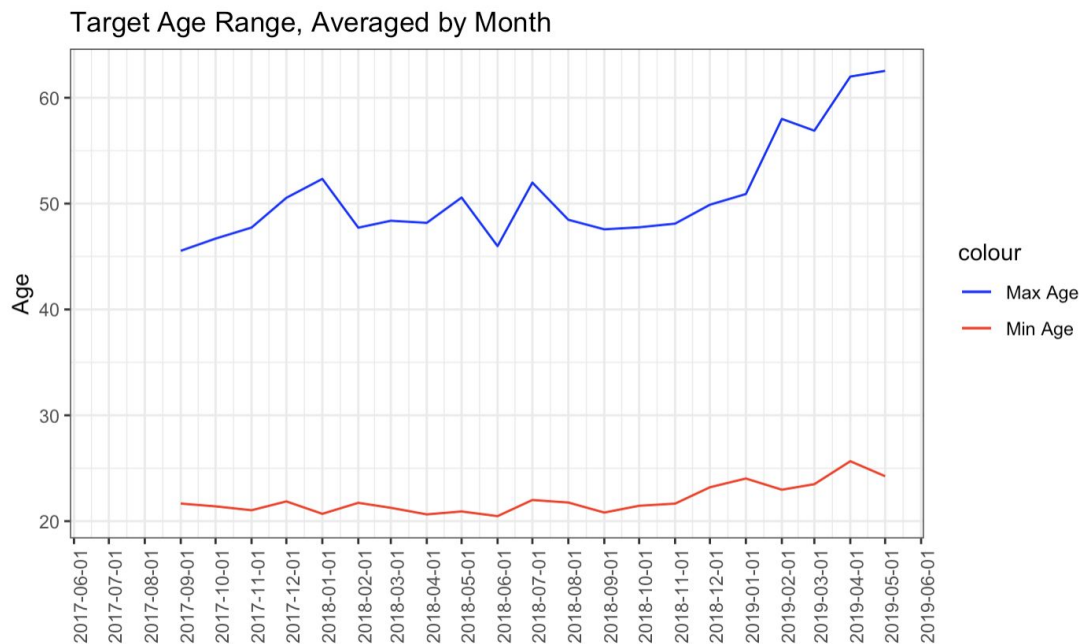


Advertisement Targets: Gender

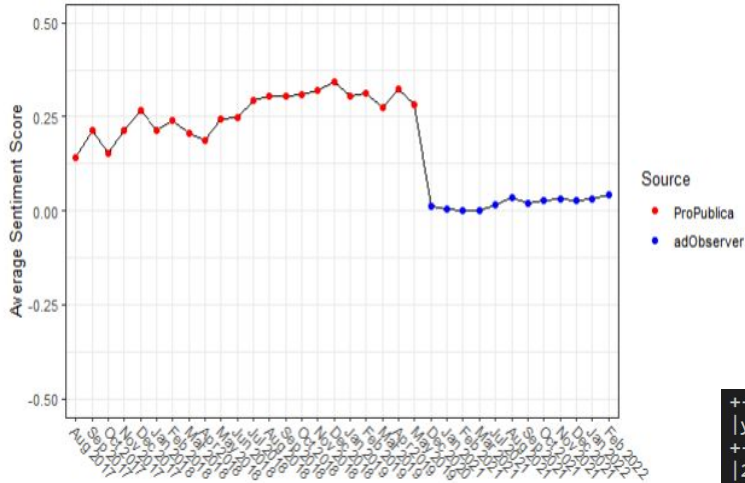




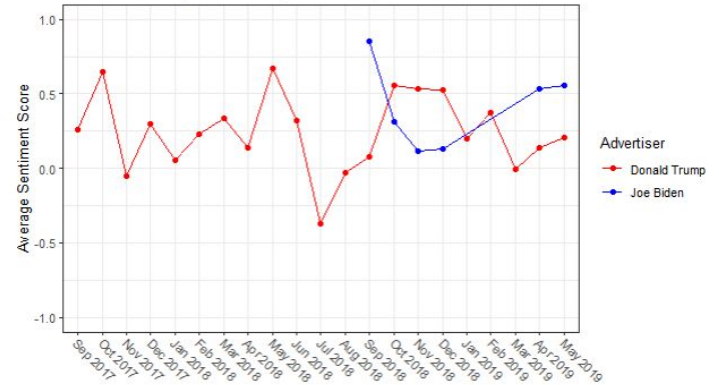
Advertisement Targets: Age Range



Average Monthly Sentiment Scores



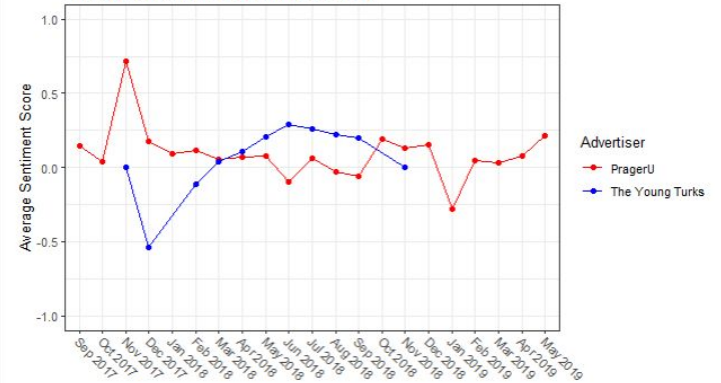
Sentiment Analysis Scores (ProPublica)



- Sentiment analysis over time using VADER (Valence Aware Dictionary and sEntiment Reasoner).
- Large discrepancy in scores between the two datasets.
 - adObserver dataset included many ads with urls, different languages, and non-standard sentence structure causing the readings to be inaccurate for adObserver.

year	avg_sentiment_score
2017	0.21
2018	0.289
2019	0.295
2020	0.011
2021	0.016
2022	0.03

Sentiment Analysis Scores (ProPublica)





Limitations





The data has many limitations...

Making it difficult to robustly identify any trends and make inferences. There are:

- Gaps in the data
 - Missing data in 2019, 2020, 2021, and 2022
- Inconsistencies in target labeling
- Self-selection bias
 - Any “trends” in targets, frequencies, etc. depends the people that install browser extensions
- Facebook hinders data collection process



Facebook Blocking Ad Data Collection

TECH / FACEBOOK / POLICY

Facebook bans academics who researched ad transparency and misinformation on Facebook

Facebook Moves to Block Ad Transparency Tools — Including Ours

Our tool had let the public see exactly how users were being targeted by advertisers. The social media giant urged us to shut it down last year.

by Jeremy B. Merrill, special to ProPublica, and Ariana Tobin, Jan. 28, 2019, 4:29 p.m. EST

TECH • FACEBOOK

Facebook Promised To Label Political Ads, But Ads For Biden, The Daily Wire, And Interest Groups Are Slipping Through

A failure in Facebook's transparency system means some election ads aren't disclosing who paid for them.



Recommendations





Our Recommendations

1. Seek out more and better data, or
2. Use other methods to identify trends in online political advertisements, or
3. Pursue other stories, or
4. Write about the limitations of data collection of advertisements on social media platforms.



Questions?



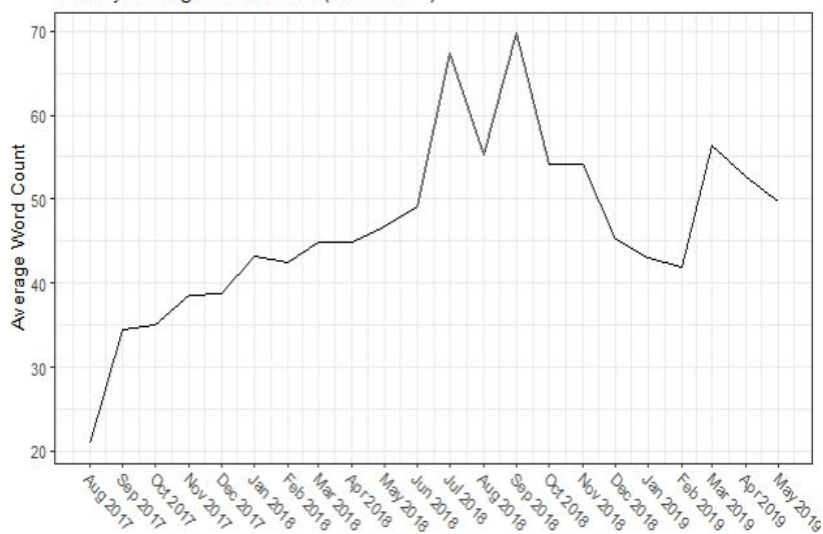
Bonus Slides



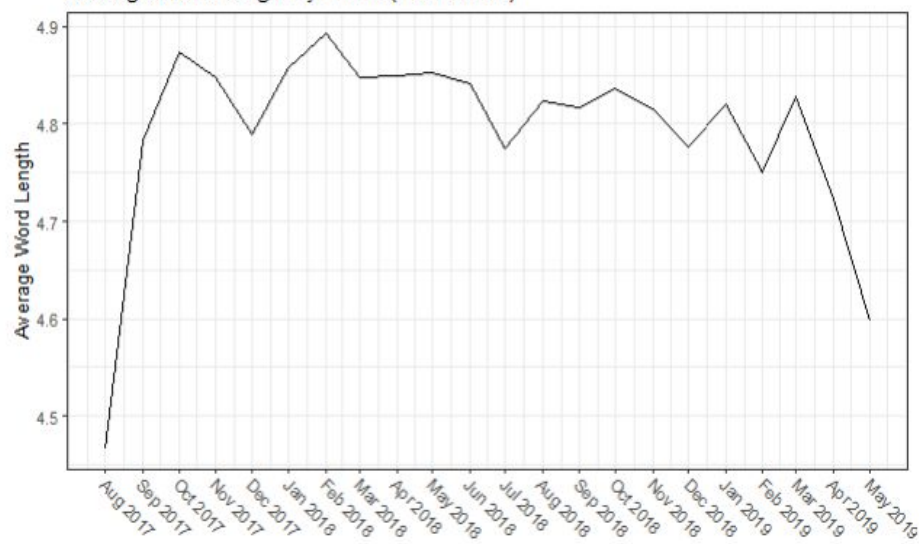


Advertisement Word Count

Monthly Average Word Count (ProPublica)



Average Word Length by Month (ProPublica)





2017

support
congress
span
trump
help
fight
need
make
people
today

2018

span
people
get
vote
time
like
help
need
trump
make

2019

campaign
get
span
class
need
help
make
today
people

2020

paid
shop
donate
sign
page
help
learn
like
apply
see

2021

send
confirmed
sign
learn
vaccine
donate
like
shop
see
get

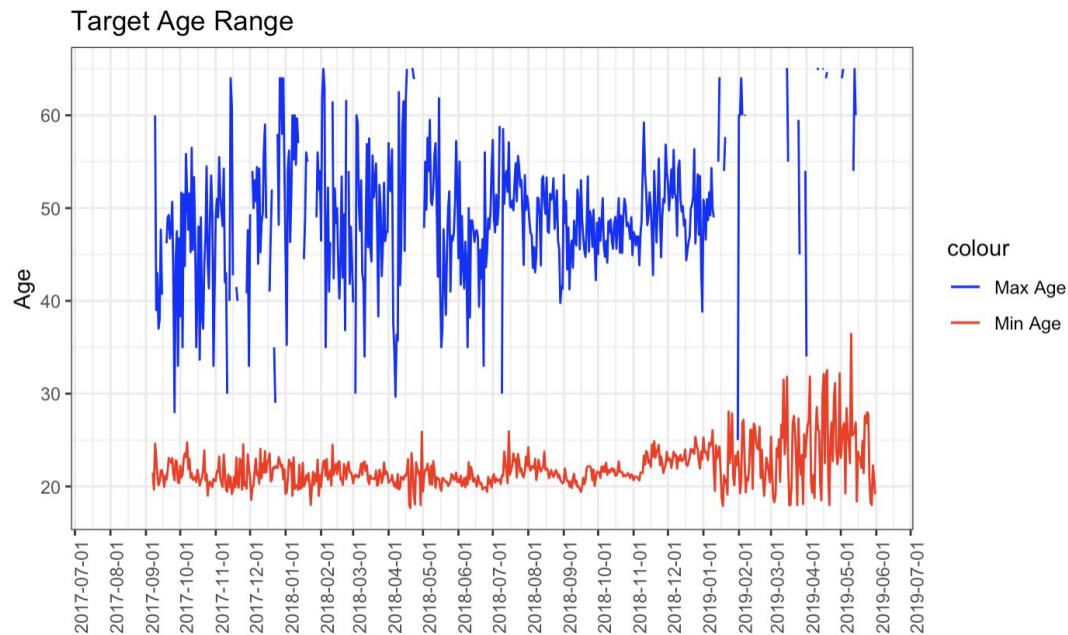
2022

new
send
sign
learn
covid
get
like
donate
confirmed
see

To investigate the text of the advertisements, we decided to examine the most frequent words per year to see if there were any obvious differences between years. The most notable was the appearance of words like “vaccine” from 2021 and “covid” from 2022. This sparked the idea to do sentiment analysis on the advertisement messages.



Advertisement Targets: Age Range





Other Advertisement Targets

target_Retargeting <chr>	n <int>
	136958
people who may be similar to their customers	24379
recently near their business	987

3 rows

target_Website <chr>	n <int>
	155991
people who have visited their website or used one of their apps	6333

target_Agency <chr>	n <int>
	161910
Acxiom	35
DLX	5
Epsilon	35
Experian	330
Oracle Data Cloud	9

6 rows



Advertisement Target Regions and States

Description: dt [30 x 3]

target_Region <chr>	n <int>	prop <dbl>
North Dakota	564	0.003
Ohio	948	0.006
Oregon	528	0.003
Pennsylvania	602	0.004
Tennessee	400	0.002
Texas	2051	0.013
Virginia	335	0.002
Washington	1045	0.006
Wisconsin	695	0.004
the United States	70290	0.433

21-30 of 30 rows

Previous 1 2 **3** Next

Description: dt [39 x 3]

target_State <chr>	n <int>	prop <dbl>
Hillsborough County, Florida	111	0.001
Illinois	760	0.005
Indiana	238	0.001
Iowa	278	0.002
Kansas	217	0.001
Kentucky	149	0.001
Maine	139	0.001
Maryland	551	0.003
Massachusetts	336	0.002
Michigan	663	0.004

11-20 of 39 rows

Previous 1 **2** 3 4 Next