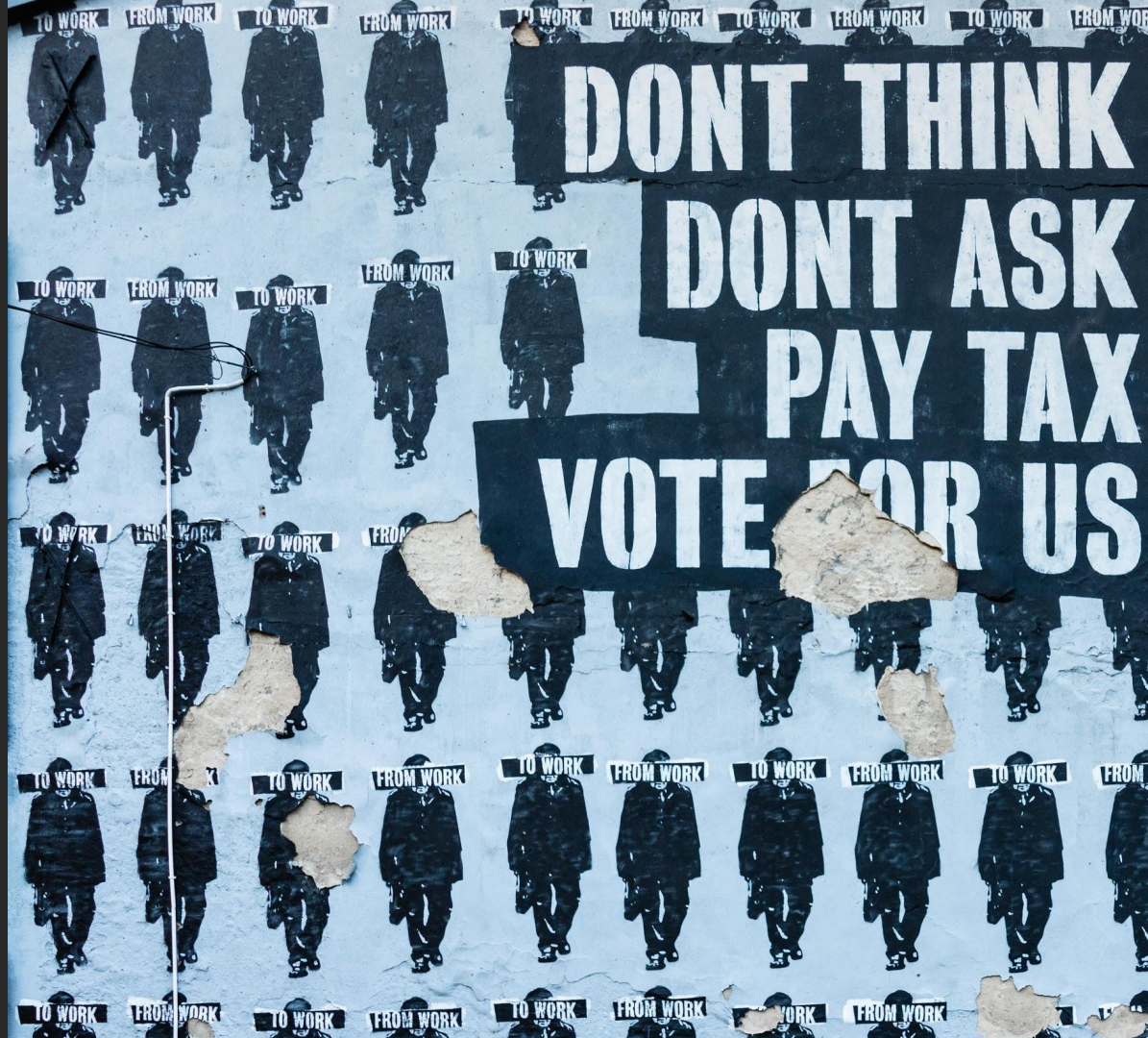
CORNELL TECH

Political Advertising's potential impact on Hate Crimes in the U.S

Erik Nestorovic | Steve Wong
Gabi Wolozin | Silvana Volio

Designing Data Products Fall 2020



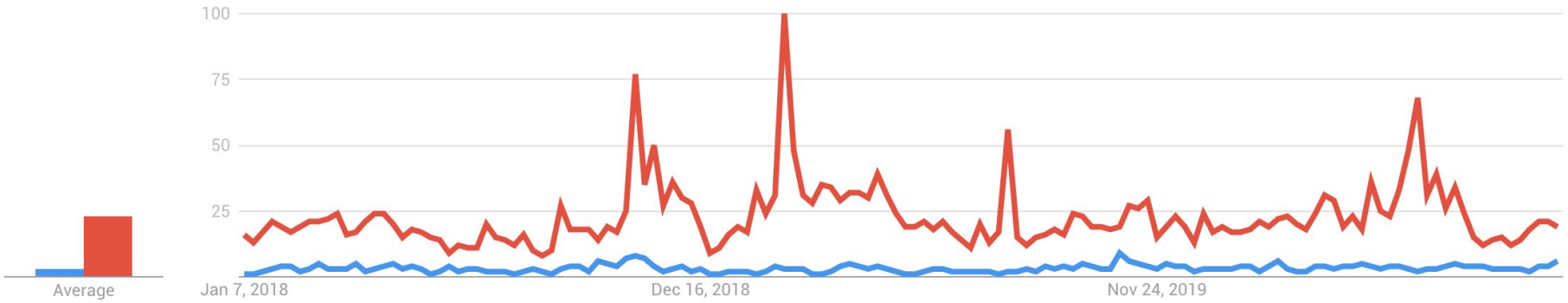
political advertising
Search term

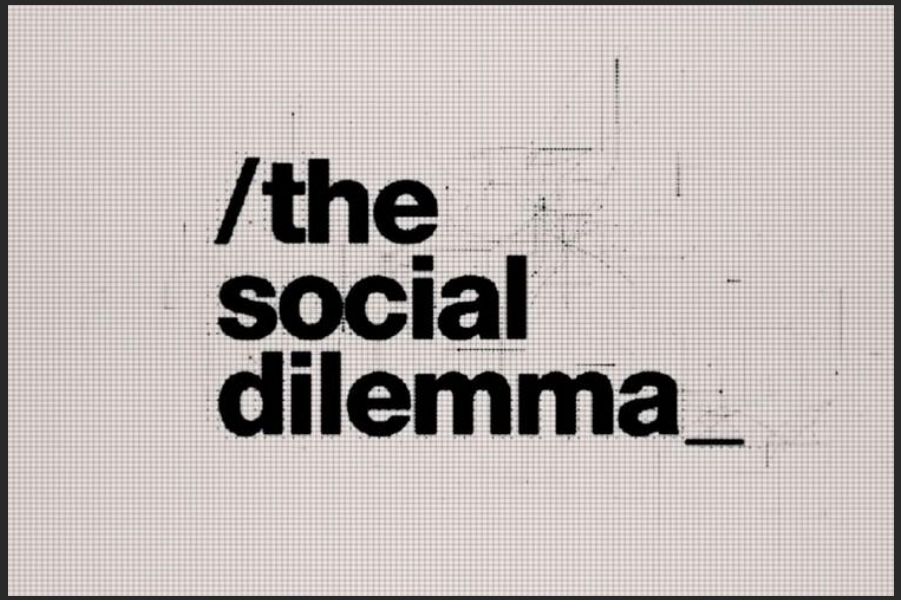
hate crimes
Search term

+ Add comparison

United States | 1/1/18 - 9/26/20 | All categories | Web Search

Interest over time





Observation

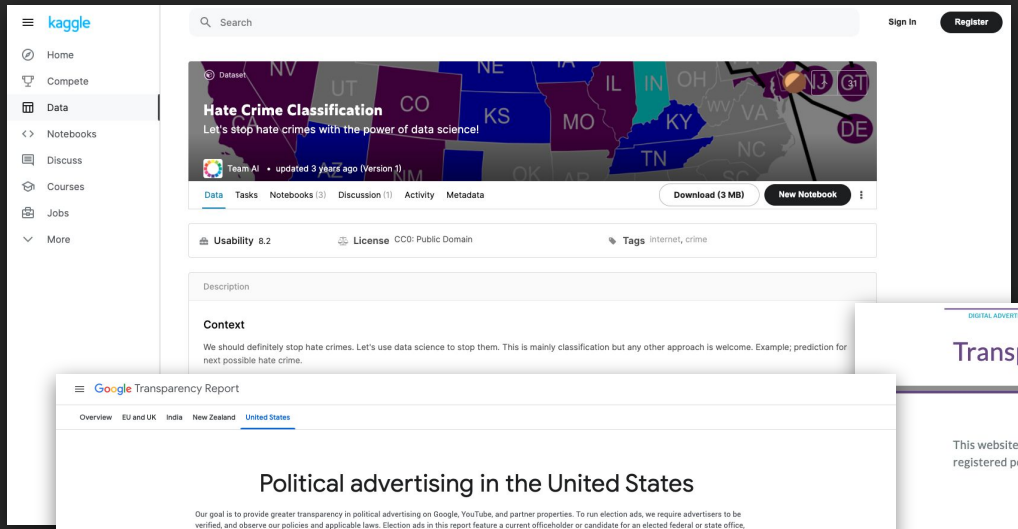
Targeted advertising can influence people's behavior but it is hate crimes that usually create more media attention. We also noticed high volumes of political advertising around election months.

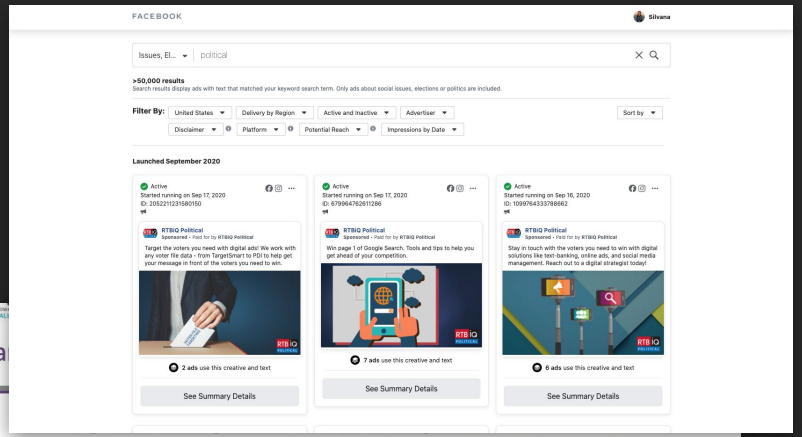
Business Question / Hypothesis

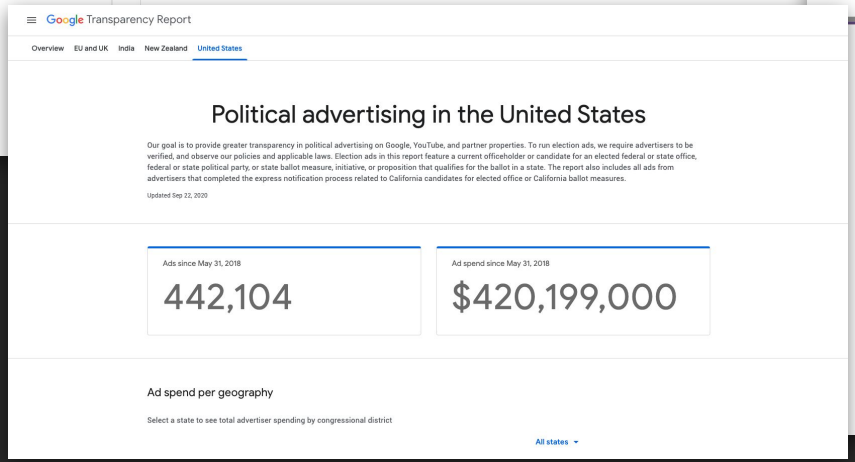
Is there a connection between the *seasonality* of political advertisements in the U.S and the amount of hate crimes reported in that location? Would a supervised model provide any indication of this trend?

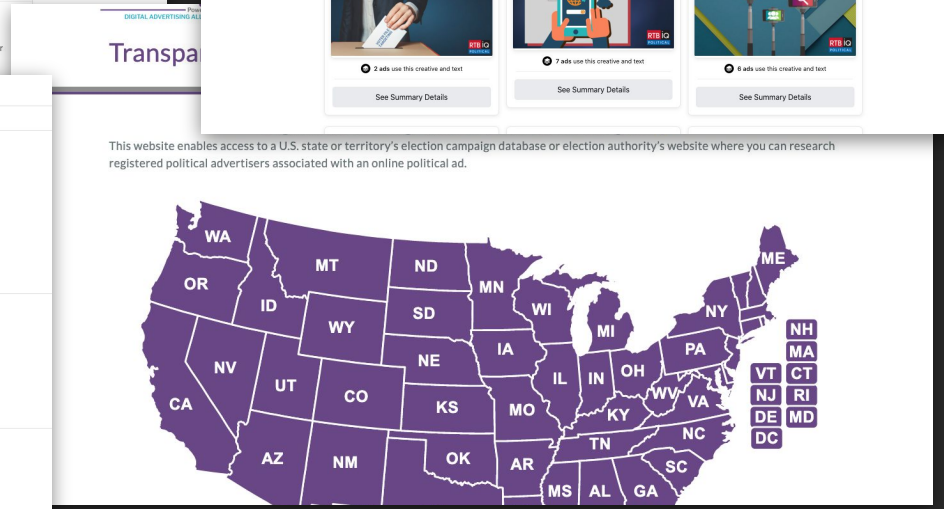
From a business perspective, could local and state governments better prepare crisis management planning around certain seasons? Is there enough data available to support a supervised learning approach?

Process: Data & Hypothesis Formulation

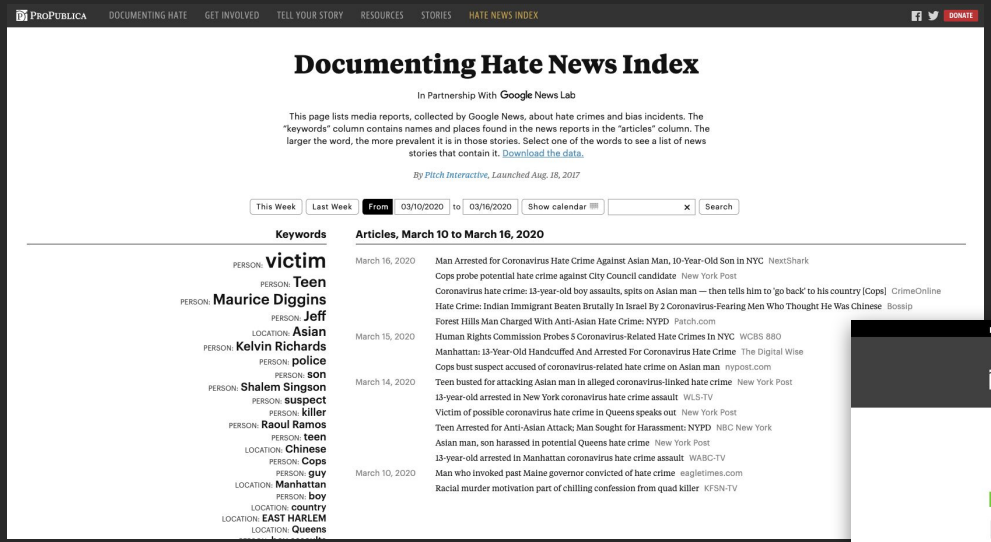








Data Scraping - Hate Crimes



ProPublica

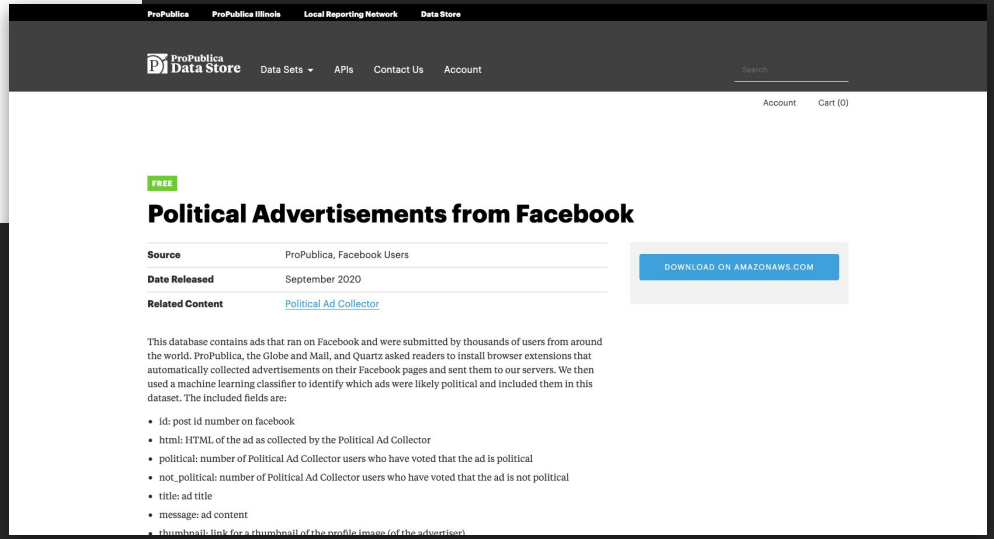
Scraped data to find:

Crime News reports over location / time

Political Advertisement over location / time

Political Advertising word usage

***See **hatecrime.rb** file for our scraping custom code*



ProPublica

US Retirement Bureau
Sponsored · Paid for by Senior Financial Services ·
FREE Retirement Newscast!

1

Discover The Little-Known Retirement Savings Secrets!

2

Protect Your Retirement From Potential Account Freezes

Discover The Closely Guarded Formula

What We Collect:

1: The advertiser's name and disclosure string

2: The ad's text, image, and link

3: The information Facebook provides about how the ad was targeted

4: When the ad was shown to you

5: Your browser language

What We Don't Collect:

Anything personally identifying, including:
Your Facebook ID number
Your name, birthday, friend list, etc.
How you interacted with ads or anything else

Why You're Seeing This Ad

3

Only you can see this

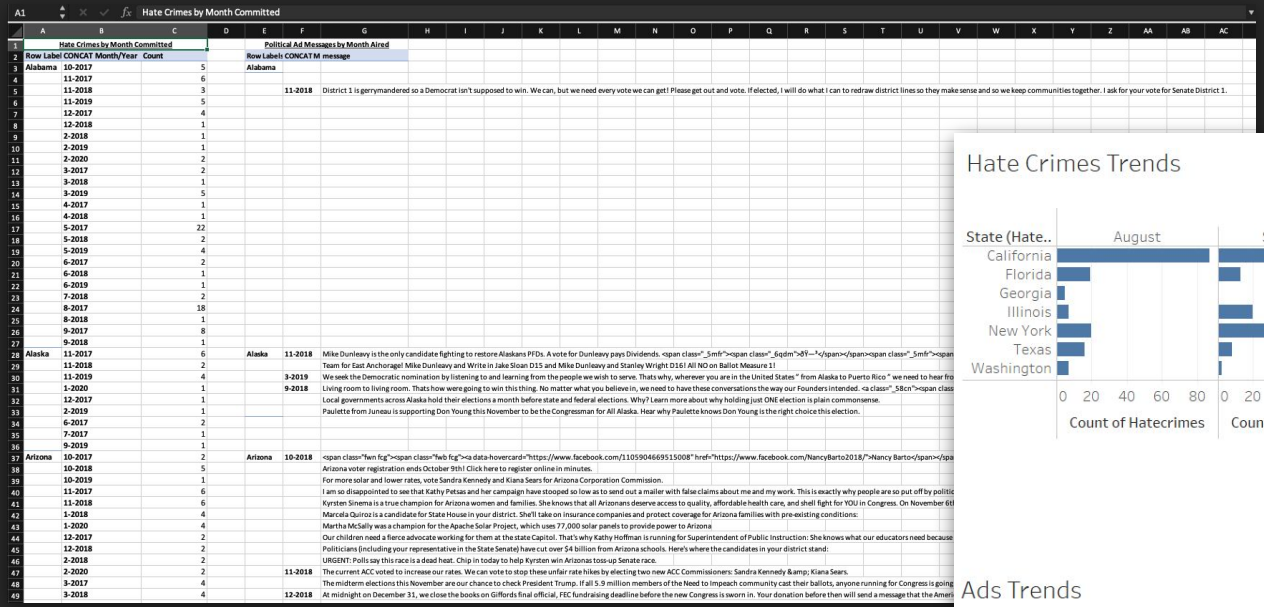
You're seeing this ad because your information matches **US Retirement Bureau's** advertising requests. There could also be more factors not listed here. **Learn More**

US Retirement Bureau wants to reach people who may be similar to their customers.

US Retirement Bureau is trying to reach people, ages 18 and older.

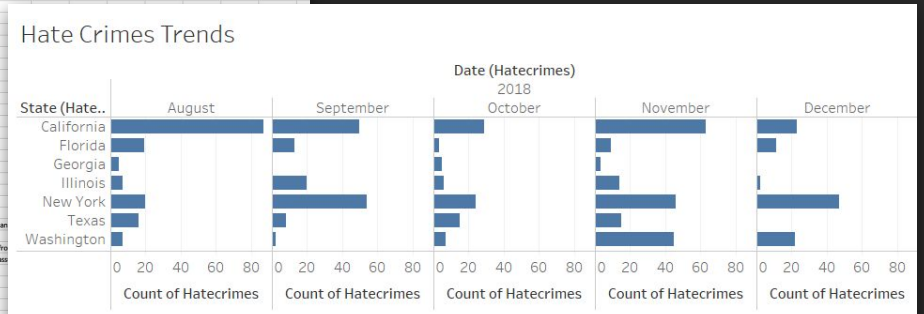
US Retirement Bureau is trying to reach people whose primary location is the United States.

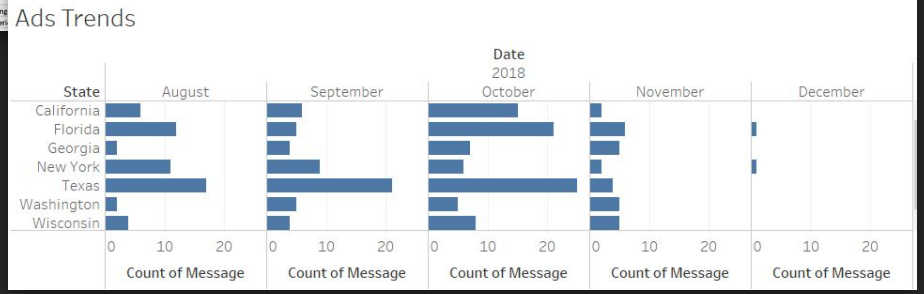
Data Clean Up - Political Ads



Hate Crimes by Month Committed		
Row Labels	CONCAT Month/Year	Count
Alabama	10-2017	5
	11-2017	6
	11-2018	3
	11-2019	5
	12-2017	4
	12-2018	1
	2-2018	1
	2-2019	1
	2-2020	2
	3-2017	2
	3-2018	1
	3-2019	5
	4-2017	1
	4-2018	1
	5-2017	22
	5-2018	2
	5-2019	4
	6-2017	2
	6-2018	1
	6-2019	1
	7-2018	2
	8-2017	18
	8-2018	1
	9-2017	8
	9-2018	1
Alaska	11-2017	6
	11-2018	2
	11-2019	4
	1-2020	1
	12-2017	1
	2-2019	1
	6-2017	2
	7-2017	1
	9-2019	1
Arizona	10-2017	2
	10-2018	5
	10-2019	1
	11-2017	6
	11-2018	6
	1-2018	4
	1-2020	4
	12-2017	2
	12-2018	2
	2-2018	2
	2-2020	2
	3-2017	4
	3-2018	4

Political Ad Messages by Month Aired		
Row Labels	CONCAT M	message
Alabama		
	11-2018	District 1 is gerrymandered so a Democrat isn't supposed to win. We can, but we need every vote we can get! Please get out and vote. If elected, I will do what I can to redraw district lines so they make sense and so we keep communities together. I ask for your vote for Senate District 1.
Alaska	11-2018	Mike Dunleavy is the only candidate fighting to restore Alaskans PFDs. A vote for Dunleavy pays Dividends. $9—³<span
		Team for East Anchorage! Mike Dunleavy and Write in Jake Sloan D15 and Mike Dunleavy and Stanley Wright D16! All NO on Ballot Measure 1!
	3-2019	We seek the Democratic nomination by listening to and learning from the people we wish to serve. Thats why, wherever you are in the United States " from Alaska to Puerto Rico " we need to hear fro
	9-2018	Living room to living room. Thats how were going to win this thing. No matter what you believe in, we need to have these conversations the way our Founders intended. <a class="_58cn"><span class=
		Local governments across Alaska hold their elections a month before state and federal elections. Why? Learn more about why holding just ONE election is plain commonsense.
		Paulette from Juneau is supporting Don Young this November to be the Congressman for All Alaska. Hear why Paulette knows Don Young is the right choice this election.
Arizona	10-2018	<a data-hovercard="https://www.facebook.com/1105904669515008" href="https://www.facebook.com/NancyBarto2018/">Nancy Barto</spa
		Arizona voter registration ends October 9th! Click here to register online in minutes.
		For more solar and lower rates, vote Sandra Kennedy and Kiana Sears for Arizona Corporation Commission.
		I am so disappointed to see that Kathy Petsas and her campaign have stooped so low as to send out a mailer with false claims about me and my work. This is exactly why people are so put off by politic
		Kyrsten Sinema is a true champion for Arizona women and families. She knows that all Arizonans deserve access to quality, affordable health care, and shell fight for YOU in Congress. On November 6t
		Marcela Quiroz is a candidate for State House in your district. She'll take on insurance companies and protect coverage for Arizona families with pre-existing conditions:
		Martha McSally was a champion for the Apache Solar Project, which uses 77,000 solar panels to provide power to Arizona
		Our children need a fierce advocate working for them at the state Capitol. That's why Kathy Hoffman is running for Superintendent of Public Instruction: She knows what our educators need because
		Politicians (including your representative in the State Senate) have cut over $4 billion from Arizona schools. Here's where the candidates in your district stand:
		URGENT: Polls say this race is a dead heat. Chip in today to help Kyrsten win Arizonas toss-up Senate race.
	11-2018	The current ACC voted to increase our rates. We can vote to stop these unfair rate hikes by electing two new ACC Commissioners: Sandra Kennedy & Kiana Sears.
		The midterm elections this November are our chance to check President Trump. If all 5.9 million members of the Need to Impeach community cast their ballots, anyone running for Congress is going
	12-2018	At midnight on December 31, we close the books on Giffords final official, FEC fundraising deadline before the new Congress is sworn in. Your donation before then will send a message that the Ameri





Over 16 hours collectively scraping, cleansing, sourcing, and coding the data.

This influenced which models we were able to consider

Models

We did a linear regression to assess whether there are more hate crimes during election months, as well as to identify what the ideal lag time between political ads and a spike in hate crimes using several dummy variables showing different lags before and after elections

Two models tested - by seasonality and state

Shortcomings of the models:

- R_2: The model explained roughly 25% of the observations
- Adjusted R_2: .23
- Explanatory Variables - not enough statistical relevance (p-values > 0.05%, high likelihood of Type I Error)
- Number of observations: 38 - Not enough publicly available data on both crimes and ad data

# Hate Crime Articles	Elec Month -2	Elec Month -1	Elec Month 0	Elec Month +1
214	0	0	0	0
236	0	0	0	0
118	0	0	0	0
499	0	0	0	0
462	0	1	0	0
80	0	0	0	0
849	0	0	0	0
666	0	0	1	0
252	0	0	0	0
306	0	0	0	0
317	0	0	0	1
144	0	0	0	0
80	0			
219	0			
258	0			
244	0			
410	0			
237	0			
384	0			
45	0			
359	0			
206	0			
262	0			
1073	0			

***see Ad and Crime Pivot Comparison by Month-Year.xlsx*

State	# Of Political Ads	# of hate crimes
Alabama	1	100
Alaska	6	19
Arizona	18	116
California	30	1510
Colorado	11	152
Connecticut	2	201
Delaware	2	46
Florida	47	476
Georgia	20	180
Hawaii	1	18
Idaho	5	51
Illinois	8	738
Iowa	4	144
Kansas	9	288
Kentuck	1	0
Kentucky	3	205
Louisiana	2	184

Predicting hate crimes by seasonality

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.23
R Square	0.05
Adjusted R Square	-0.10
Standard Error	280.30
Observations	38

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	5	139398.67	27879.73	0.35	0.88
Residual	32	2514125.88	78566.43		
Total	37	2653524.55			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	344.39	48.79	7.06	0.00	245.00	443.78	245.00	443.78
Elec Month -2	-103.39	284.51	-0.36	0.72	-682.93	476.14	-682.93	476.14
Elec Month -1	117.61	284.51	0.41	0.68	-461.93	697.14	-461.93	697.14
Elec Month 0	321.61	284.51	1.13	0.27	-257.93	901.14	-257.93	901.14
Elec Month +1	-27.39	284.51	-0.10	0.92	-606.93	552.14	-606.93	552.14
Elec Month +2	-108.39	284.51	-0.38	0.71	-687.93	471.14	-687.93	471.14

Linear Regression: # of Political Ads in state vs. Total Hate Crimes in state

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.499
R Square	0.249
Adjusted R Square	0.233
Standard Error	271.982
Observations	49

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	1154061.343	1154061.343	15.601	0.000
Residual	47	3476791.188	73974.281		
Total	48	4630852.531			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	125.518	49.519	2.535	0.015	25.898	225.138	25.898	225.138
# Of Political Ads	12.327	3.121	3.950	0.000	6.049	18.606	6.049	18.606

Findings

RESULTS

- We were not able to find advertising data that was relevant to the hate crime data we had, it was essentially impossible to cross reference the date and the location of the advertisements with the amount of hate crimes
  - Therefore, a Naive Bayes model was not feasible as we could not define the prior of political advertisements having an influence in the amount or nature of hate crimes

WHY THIS ISSUE MATTERS IN A BUSINESS CONTEXT

- This could create a digital business model by licensing for advertising platforms (Facebook, Twitter, TikTok) to local governments who have highly partisan residents, could be used for capacity planning of city and local employees (number of officers needed, etc.)

OTHER MODEL OPPORTUNITIES

- With greater data, one could apply an unsupervised learning approach to determine whether words used in political ads are likely to trigger a peak in Hate Crimes news reports

Appendix - Data Sources (URLs)

- Link: Group 6's Data Sets and Source Files
  - Combined Data + Regression Models: *Ad and Crime Pivot Comparison by Month-Year.xlsx*
  - FB Ad data set post-data wrangling: *Ad_Data - post text formatting.xlsx*
  - Ads and hate crime dashboard created from raw data: *Ads and crime dashboard.twbx*
  - Database created from raw FB ad file: *Facebookad.db*
  - Original raw FB ads data set file: *fbpac-ds-end-US.csv*
  - Custom Code for running data scraper on hate crime website: *hatecrime.rb*
  - Data Scraping Output: *hatecrimes_by_state.xlsx*

- Link: original ProPublica Facebook political ads data source

- Link: original ProPublica hate crimes data source