Geographical variation of U.S presidential candidates' speeches in 2024 Presidential Election

EPPS 6356 Data Visualization

Students: Kelvin Chun Man Chan, Minh Dao, Sadman, Karim Iqbal

I. Research Question

Do United States presidential candidates adjust their speeches based on their audience at the presidential campaign rallies? Although past research shows that presidential candidates would hold more campaign events in the swing states, as these states' electoral results are more likely to determine who is going to win the elections (Snyder & Yousaf, 2020), whether the candidates would adjust their speech based on the interests of the audience who physically attend these events is less certain. The uncertainty results from the increased coverage of social media, which redefines the scope of the audience, as their rally speeches can reach every voter in real-time.

Would social media cause the candidates to deliver more similar rally speeches? We propose to conduct a text analysis on the presidential candidates of the major parties in the 2024 presidential election to test the between-candidate variations, to see whether Donald Trump and Kamala Harris would prioritize different issues in their speeches. Also, we would ask where to test within-candidate variations, to test whether the same candidate would say differently at rallies held in different states.

Current research on speech campaigns tends to be more single-issued and mostly focuses on word choice differences among candidates (BBC, 2024; Bonikowski & Gidron, 2016; Lerer & Haberman, 2024, Slatcher et al., 2007), our proposal hopes to extend the current research further to conduct candidate comparison from a different angle.

II. Data Collection & Methodology

Speech data of the Democratic and Republican presidential candidates is collected from the [American Presidency Project](#) and [https://www.rev.com/transcripts](#) website. Separate corpora is created for each of the candidates. A dictionary-based model has been employed to find out the relevant topics identified by the supervised machine learning method. Finally, the obtained results are matched to Google Trends data to compare whether the presidential candidates mention the topics in the Google Trends.

III. Visualization

We present four key components in the dashboard:

1. A **US map** displaying the geographical distribution of Trump's and Harris's speeches by topic.

2. A **US map** showing Google Trends' search interest for each topic from April 2024 to July 21, 2024 (the day Biden withdrew from the 2024 election and Harris became the presumptive nominee).

3. A **stacked percentage column chart** visualizing the proportional emphasis that presidential candidates Donald Trump and Kamala Harris place on various topics within the battleground states.

4. A **dynamic correlation heatmap** offering a sophisticated tool for exploring the correlations between different campaign topics discussed by Donald Trump and Kamala Harris.

3.1 Geographical distribution of Trump's and Harris' speeches

First, we collected Trump's and Harris' speeches online.

Second, after collecting all the speeches, we run a dictionary-based text analysis to analyze the speeches using the following steps:

A. We use the "tm" and "quanteda" R packages to clean the text and apply stemming to all the words.

B. We adopt the keyword list from Guo et al.'s (2016) paper, which focuses on the topics and words that are relevant to the United States presidential elections. We refined the topics based on the 2024 presidential election survey (Pew, 2024) and expanded the list of keywords from insights in the 2020 and 2024 elections by ChatGPT. We apply stemming to all the keywords.

C. The speeches were split into sentences, and each sentence was classified as relevant or irrelevant to each topic, creating a binary variable.

D. We calculate the frequency of sentences related to each topic.

E. We plot the state-level heatmaps to show the frequency of topics based on the location of the presidential rallies. We try to visualize the level of intensity each topic was mentioned relative to geographical locations.

F. Some states have null values indicating that no presidential rally was recorded in our corpus.
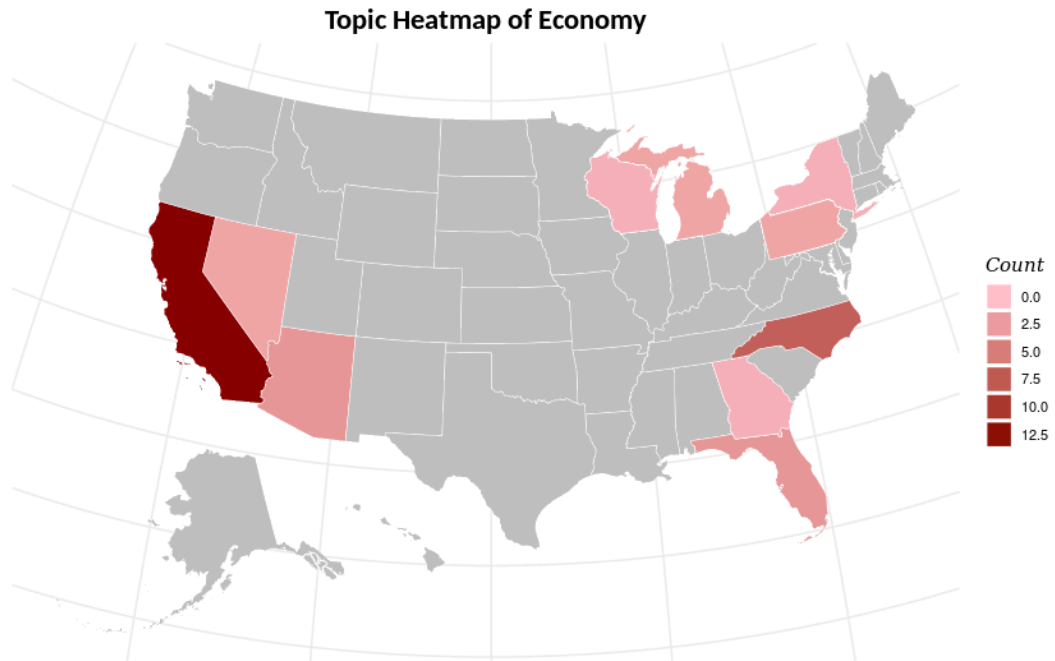
Topic Heatmap of Economy

Figure 1. Sample Screenshot of Speeches Heatmap

The following topics are included in the analysis. The full dictionary of keywords is included in the appendix:

1. Economy

2. Healthcare

3. Supreme Court appointments

4. Foreign affairs

5. Violent Crime

6. Immigration

7. Gun Policy

8. Abortion

9. Racial Inequality

10. Climate Change

3.2 Google Trends data

We use gtrendsR R package to obtain the data. We extract the interest over time for different topics at the state level. In the gtrends() command, we set the parameters (geo = "US", compared_breakdown = FALSE) to get the raw ratio of search queries across the United States. Because Google Trends data is highly sensitive to the study period, we fixed the timeframe from April 2024 to July 21 2024 to ensure the comparability of results across topics.

We visualize the interest data by creating the geographical heatmap with ggplot2 R package, where the state with the highest number of searches is assigned a value of 100. Since we are using the state-level data, there are some missing values across states. For instance, netizens in New England rarely searched for gun policy, therefore, the trend score has missing values for Maine, Vermont, and New Hampshire.
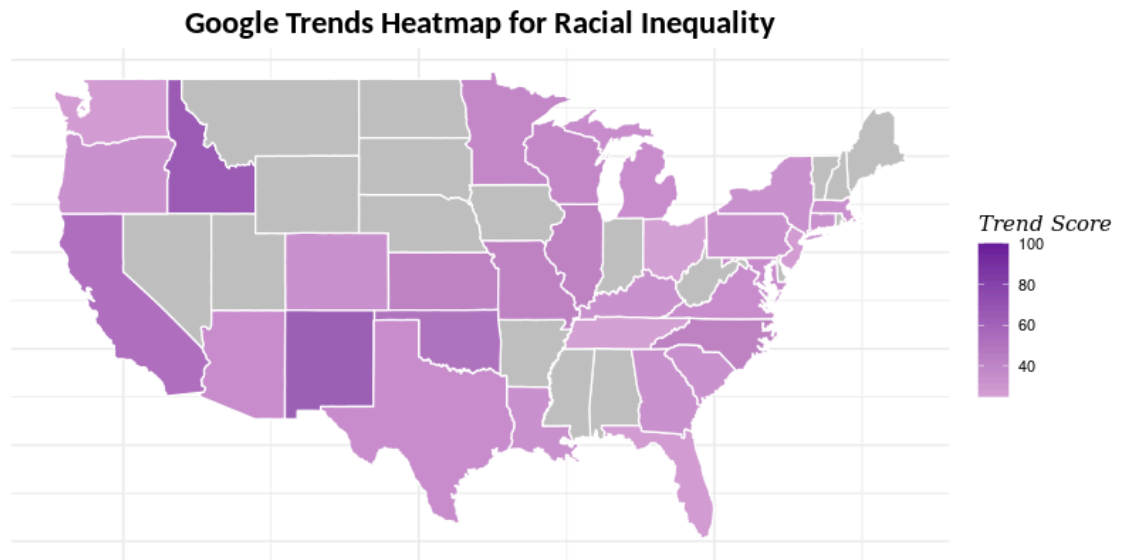
Google Trends Heatmap for Racial Inequality

Figure 2. Sample Screenshot of Google TrendsHeatmap

3.3 Stacked Percentage Column Chart

The stacked percentage column chart is designed to analyze and visualize the proportional emphasis that presidential candidates Donald Trump and Kamala Harris place on various campaign topics within the battleground states. These battleground states (also known as 'Swing States') include Arizona, Georgia, Michigan, Nevada, New Hampshire, North Carolina, Pennsylvania, Texas, and Wisconsin. It allows users to select a battleground state from a dropdown menu. Upon selection, the application displays a stacked percentage bar chart where each bar represents a candidate and each segment of the bar delineates the proportion of their speeches dedicated to ten topics listed in section 3.1.
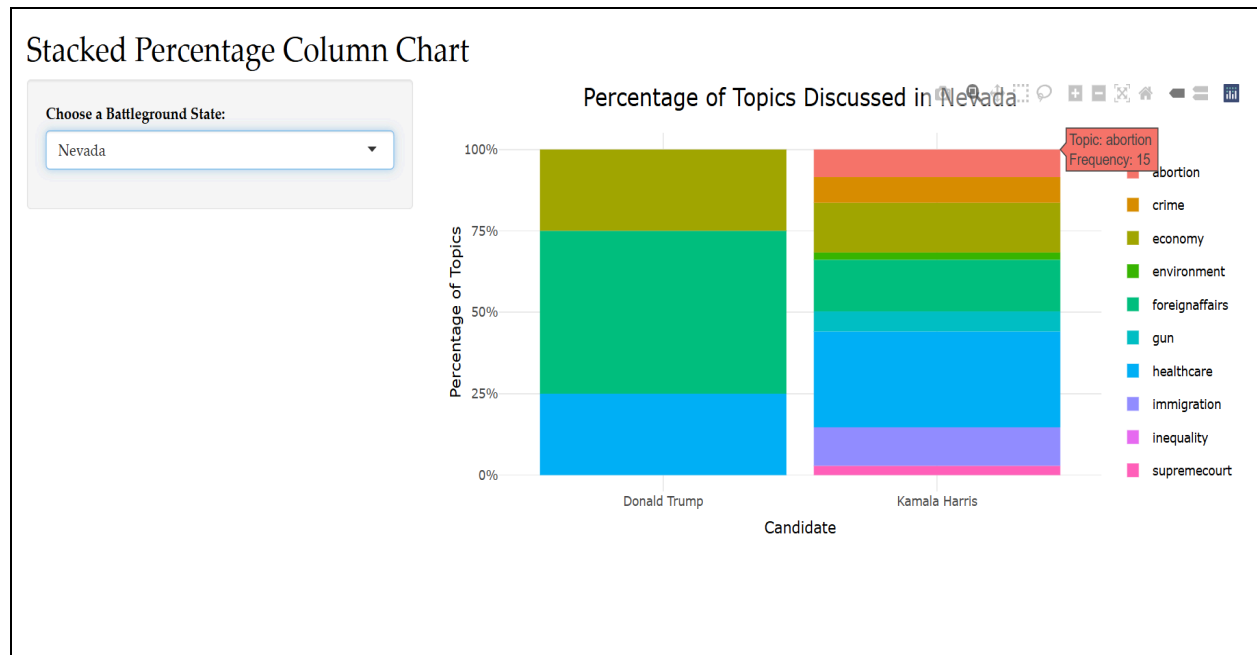
Figure 3: Stacked Percentage Column Chart

This visual representation helps users to directly compare the strategic differences in topic emphasis between the two candidates within the same state. For instance, users can observe how much of each candidate's speech was devoted to the economy in Pennsylvania as compared to North Carolina. The application uses dplyr for data manipulation - filtering and summarizing speech data by state - and ggplot2 combined with plotly for creating interactive, visually engaging stacked bar charts that update dynamically based on user input.

3.4 Dynamic Correlation Heatmap

The dynamic correlation heatmap explores the correlations between different campaign topics discussed by either Donald Trump or Kamala Harris. It allows users to select a candidate and various topics they are interested in exploring. The core feature of this application is a heatmap that illustrates the strength and direction of correlations

among the selected topics. This visualization helps in identifying patterns or clusters of topics that are frequently discussed together, suggesting a strategic linkage or common thematic focus in the candidates' speeches.
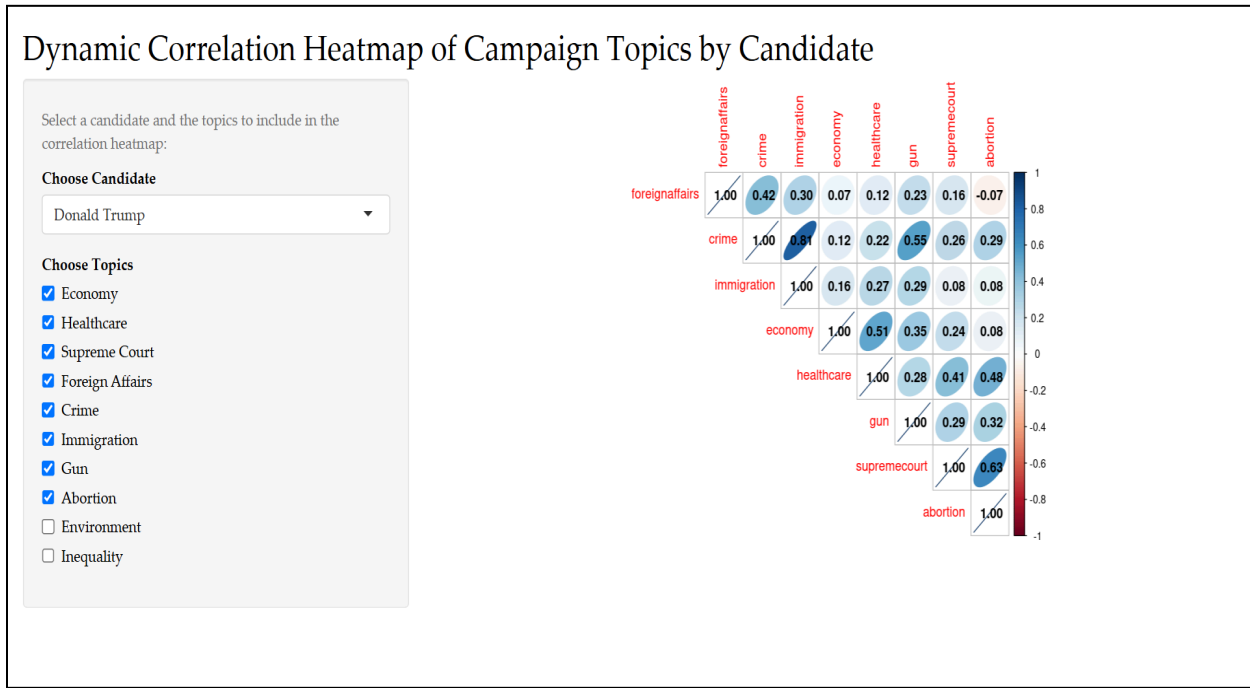
Figure 4: Dynamic Correlation Heatmap

Users interact with the dynamic heatmap through a simple interface that includes options to select a candidate and multiple topics. The heatmap updates in real time as selections are made, showing correlations that can range from strong positive (topics frequently discussed together) to strong negative (topics rarely discussed together). The correlation values are computed using the cor() function in R, handling missing data effectively, and visualized using the corrplot package, which provides a clear and detailed representation of the correlation matrix.

IV. Visualization techniques

4.1 Shiny

The heatmaps of speeches and Google Trends scores are displayed through two apps created using Shiny. Users can select presidential candidates and topics from the sidebar drop-down list. While Google Trends scores are a ratio variable, meaning they cannot be used to directly compare interest across different topics, they allow for comparison within the same topic. Viewers can easily examine how geographical variations in interest for specific topics differ or align between the two presidential candidates and across states. For example, Google Trends results indicate that netizens in Southern states show more interest in gun policy, which aligns with Trump's frequent discussions on gun rights in this region. If all states show a similar level of purple color, this suggests that there is no geographical variations in the Google Trends score.

Viewers can also visually compare the strategies of Candidate Donald Trump and Candidate Kamala Harris through the content of their speeches, as categorized and quantified in our analysis (please refer to the methodology section for details). For instance, Harris discussed abortion more frequently than Trump, while Trump emphasized gun rights more often, particularly in Southern states.

To enhance the user experience, we customized the font of the app's sidebars by adding a tag in the UI of the Shiny apps. The fonts for plot titles and legends were modified using the theme parameter in ggplot2, with fonts loaded via the extrafont R package. To better align with the themes of American politics, the heatmap for Trump is displayed using a red color scale, while the heatmap for Harris uses a blue color scale.

Additionally, we disabled both the titles and ticks on all axes of the heatmaps to make the visualizations more user-friendly.

4.2 Dashboard

The presentation layout is organized using a dashboard. Two heatmaps are displayed in the same column using the tabset command in R Markdown, allowing for a larger width for each heatmap and enhancing visibility for viewers. To improve the interface's aesthetic appeal, we added custom CSS to change the background color and font of the tabs, making the design less dull. The remaining space is used to display the frequencies of topics discussed by each presidential candidate.

References

Bonikowski, B., & Gidron, N. (2016). The populist style in American politics: Presidential campaign discourse, 1952–1996. *Social Forces*, *94*(4), 1593-1621.

BBC News. (2024, October 7). *China is part of the US election - but only from one candidate*. BBC. https://www.bbc.com/news/articles/ckgvr9kgkqy0

Guo, L., Vargo, C. J., Pan, Z., Ding, W., & Ishwar, P. (2016). Big social data analytics in journalism and mass communication: Comparing dictionary-based text analysis and unsupervised topic modeling. *Journalism & Mass Communication Quarterly*, *93*(2), 332-359.

Lerer, L., & Haberman, M. (2024, September 13). *Trump and Harris offer dueling speeches, with selective accuracy*. The New York Times. https://www.nytimes.com/2024/09/13/us/elections/trump-harris-speeches-accuracy.html

Pew Research Center. (2024, September 9). *Issues and the 2024 election*. Pew Research Center.

https://www.pewresearch.org/politics/2024/09/09/issues-and-the-2024-election/

Slatcher, R. B., Chung, C. K., Pennebaker, J. W., & Stone, L. D. (2007). Winning words: Individual differences in linguistic style among US presidential and vice presidential candidates. *Journal of Research in Personality*, *41*(1), 63-75.

Snyder Jr, J. M., & Yousaf, H. (2020). *Making rallies great again: The effects of presidential campaign rallies on voter behavior, 2008-2016* (No. w28043). National bureau of economic research.

Appendix

	Word	Stemmed word	Topic
1	Economic	econom	Economy
2	Recession	recess	Economy
3	Economy	economi	Economy
4	Recovery	recoveri	Economy
5	Recoveries	recoveri	Economy
6	Inflation	inflat	Economy
7	Stock market	stock market	Economy
8	Dow	dow	Economy
9	GDP	gdp	Economy
10	Gross domestic product	gross domest product	Economy
11	Stimulus checks	stimulus check	Economy
12	COVID relief	covid relief	Economy
13	Pandemic economy	pandem economi	Economy
14	Essential workers	essenti worker	Economy
15	Student loan forgiveness	student loan forgiv	Economy
16	Supply chain	suppli chain	Economy

17	Inflation reduction	inflat reduct	Economy
18	Cost of living	cost live	Economy
19	Minimum wage	minimum wage	Economy
20	Housing crisis	hous crisi	Economy
21	Wage gap	wage gap	Economy
22	Digital economy	digit economi	Economy
23	Health	health	Health Care
24	Healthcare	healthcar	Health Care
25	Medical	medic	Health Care
26	Obamacare	obamacar	Health Care
27	Affordable care	afford care	Health Care
28	Romneycare	romneycar	Health Care
29	Medicare	medicar	Health Care
30	Medicaid	medicaid	Health Care
31	COVID-19	covid	Health Care
32	Pandemic	pandem	Health Care
33	Vaccine	vaccin	Health Care
34	Vaccine mandates	vaccin mandat	Health Care

35	Health equity	health equiti	Health Care
36	Telemedicine	telemedicin	Health Care
37	Mental health	mental health	Health Care
38	Long COVID	long covid	Health Care
39	Public health infrastructure	public health infrastructur	Health Care
40	Medicare expansion	medicar expans	Health Care
41	Roe v. Wade	roe v wade	Supreme Court Appointments
42	Dobbs decision	dobb decis	Supreme Court Appointments
43	Court packing	court pack	Supreme Court Appointments
44	Judicial reform	judici reform	Supreme Court Appointments
45	Judiciary	judiciari	Supreme Court Appointments
46	Supreme Court	suprem court	Supreme Court Appointments
47	Originalism	origin	Supreme Court Appointments

48	Progressive judges	progress judg	Supreme Court Appointments
49	Conservative justices	conserv justic	Supreme Court Appointments
50	Lifetime appointments	lifetim appoint	Supreme Court Appointments
51	SCOTUS leaks	scotus leak	Supreme Court Appointments
52	Terrorist	terrorist	Foreign Affairs
53	Foreign	foreign	Foreign Affairs
54	Iraq	iraq	Foreign Affairs
55	Iran	iran	Foreign Affairs
56	Afghan	afghan	Foreign Affairs
57	Israel	israel	Foreign Affairs
58	Islam	islam	Foreign Affairs
59	Palestinian	palestinian	Foreign Affairs
60	Arab	arab	Foreign Affairs
61	Syria	syria	Foreign Affairs
62	Libya	libya	Foreign Affairs

63	Troop	troop	Foreign Affairs
64	Outsource	outsourc	Foreign Affairs
65	Insource	insourc	Foreign Affairs
66	Russia	russia	Foreign Affairs
67	Benghazi	benghazi	Foreign Affairs
68	United Nation	unit nation	Foreign Affairs
69	US embassy	us embassi	Foreign Affairs
70	U.S. embassy	us embassi	Foreign Affairs
71	Ahmadinejad	ahmadinejad	Foreign Affairs
72	Putin	putin	Foreign Affairs
73	Chavez	chavez	Foreign Affairs
74	Castro	castro	Foreign Affairs
75	Kim Jong-un	kim jongun	Foreign Affairs
76	North Korea	north korea	Foreign Affairs
77	North Korean	north korean	Foreign Affairs
78	World leaders	world leader	Foreign Affairs
79	Nations	nation	Foreign Affairs
80	Hamas	hama	Foreign Affairs

81	Terrorism	terror	Foreign Affairs
82	War on terror	war terror	Foreign Affairs
83	Osama	osama	Foreign Affairs
84	Bin Laden	bin laden	Foreign Affairs
85	Al Qaeda	al qaeda	Foreign Affairs
86	China	china	Foreign Affairs
87	Chinese	chines	Foreign Affairs
88	Trade	trade	Foreign Affairs
89	Cheap labor	cheap labor	Foreign Affairs
90	Currency manipulation	currenc manipul	Foreign Affairs
91	World Trade Organization	world trade organ	Foreign Affairs
92	Middle East	middl east	Foreign Affairs
93	Middle Eastern	middl eastern	Foreign Affairs
94	Saddam	saddam	Foreign Affairs
95	Persian Gulf	persian gulf	Foreign Affairs
96	Muslim	muslim	Foreign Affairs
97	Palestine	palestin	Foreign Affairs
98	North Africa	north africa	Foreign Affairs

99	North African	north african	Foreign Affairs
100	Asia	asia	Foreign Affairs
101	Overseas	oversea	Foreign Affairs
102	Taliban	taliban	Foreign Affairs
103	Yemen	yemen	Foreign Affairs
104	Homeland security	homeland secur	Foreign Affairs
105	National security	nation secur	Foreign Affairs
106	Pentagon	pentagon	Foreign Affairs
107	Military	militari	Foreign Affairs
108	Defense	defens	Foreign Affairs
109	CIA	cia	Foreign Affairs
110	Armed forces	arm forc	Foreign Affairs
111	Ukraine	ukrain	Foreign Affairs
112	Russia sanctions	russia sanction	Foreign Affairs
113	NATO	nato	Foreign Affairs
114	Withdrawal from Afghanistan	withdraw afghanistan	Foreign Affairs
115	Chinese influence	chines influenc	Foreign Affairs

116	Uyghurs	uyghur	Foreign Affairs
117	Taiwan	taiwan	Foreign Affairs
118	Belt and Road	belt road	Foreign Affairs
119	Cybersecurity	cybersecur	Foreign Affairs
120	TikTok	tiktok	Foreign Affairs
121	Zelenskyy	zelenskyy	Foreign Affairs
122	Xi Jinping	xi jinp	Foreign Affairs
123	Ukraine war	ukrain war	Foreign Affairs
124	Iran protests	iran protest	Foreign Affairs
125	AUKUS	aukus	Foreign Affairs
126	Quad alliance	quad allianc	Foreign Affairs
127	Nord Stream	nord stream	Foreign Affairs
128	Climate diplomacy	climat diplomaci	Foreign Affairs
129	Illegal drug	illeg drug	Violent Crime
130	Marijuana	marijuana	Violent Crime
131	Heroin	heroin	Violent Crime
132	Cocaine	cocain	Violent Crime
133	Methamphetamine	methamphetamin	Violent Crime

134	Drug trade	drug trade	Violent Crime
135	Drug addiction	drug addict	Violent Crime
136	Drug abuse	drug abus	Violent Crime
137	Alcoholism	alcohol	Violent Crime
138	Alcohol addition	alcohol addit	Violent Crime
139	Alcohol abuse	alcohol abus	Violent Crime
140	Crime rate	crime rate	Violent Crime
141	Prisons	prison	Violent Crime
142	Law enforcement	law enforc	Violent Crime
143	Death penalty	death penalti	Violent Crime
144	Gun violence	gun violenc	Violent Crime
145	Mass shootings	mass shoot	Violent Crime
146	Opioid crisis	opioid crisi	Violent Crime
147	Fentanyl	fentanyl	Violent Crime
148	Domestic terrorism	domest terror	Violent Crime
149	Hate crimes	hate crime	Violent Crime
150	Police reform	polic reform	Violent Crime
151	Criminal justice reform	crimin justic reform	Violent Crime

152	Black Lives Matter protests	black live matter protest	Violent Crime
153	Insurrection	insurrect	Violent Crime
154	Immigration	immigr	Immigration
155	Immigrant	immigr	Immigration
156	Immigrate	immigr	Immigration
157	DREAM Act	dream act	Immigration
158	Border issue	border issu	Immigration
159	Border issues	border issu	Immigration
160	Border safety	border safeti	Immigration
161	Border security	border secur	Immigration
162	Deportation	deport	Immigration
163	Family separation	famili separ	Immigration
164	Child detention	child detent	Immigration
165	Asylum seekers	asylum seeker	Immigration
166	Title 42	titl	Immigration
167	DACA	daca	Immigration
168	Border wall	border wall	Immigration
169	Migrant caravans	migrant caravan	Immigration

170	Sanctuary cities	sanctuari citi	Immigration
171	Immigration reform	immigr reform	Immigration
172	Gun control	gun control	Gun Control
173	Gun rights	gun right	Gun Control
174	Firearm	firearm	Gun Control
175	NRA	nra	Gun Control
176	Assault weapons ban	assault weapon ban	Gun Control
177	Red flag laws	red flag law	Gun Control
178	Background checks	background check	Gun Control
179	School shootings	school shoot	Gun Control
180	Second Amendment	second amend	Gun Control
181	Ghost guns	ghost gun	Gun Control
182	Gun violence prevention	gun violenc prevent	Gun Control
183	Planned Parenthood	plan parenthood	Abortion
184	Contraception	contracept	Abortion
185	Abortion	abort	Abortion
186	Pro-choice	prochoic	Abortion

187	Pro-life	prolif	Abortion
188	Wade	wade	Abortion
189	Reproductive rights	reproduct right	Abortion
190	Roe overturned	roe overturn	Abortion
191	Abortion bans	abort ban	Abortion
192	Fetal heartbeat laws	fetal heartbeat law	Abortion
193	Abortion access	abort access	Abortion
194	Abortion pills	abort pill	Abortion
195	Trigger laws	trigger law	Abortion
196	Pro-choice rallies	prochoic ralli	Abortion
197	Dobbs v. Jackson	dobb v jackson	Abortion
198	Renewable	renew	Environment/Climate
199	Environmental	environment	Environment/Climate
200	Pollution	pollut	Environment/Climate
201	Pollute	pollut	Environment/Climate
202	Pollutes	pollut	Environment/Climate
203	Clean air	clean air	Environment/Climate
204	Global warming	global warm	Environment/Climate

205	Climate change	climat chang	Environment/Climate
206	Wildlife	wildlif	Environment/Climate
207	Clean water	clean water	Environment/Climate
208	Natural resource	natur resourc	Environment/Climate
209	Sea levels	sea level	Environment/Climate
210	Sustainable development	sustain develop	Environment/Climate
211	Climate crisis	climat crisi	Environment/Climate
212	Net zero	net zero	Environment/Climate
213	Paris Agreement	pari agreement	Environment/Climate
214	Carbon emissions	carbon emiss	Environment/Climate
215	Green energy	green energi	Environment/Climate
216	Electric vehicles	electr vehicl	Environment/Climate
217	Climate migration	climat migrat	Environment/Climate
218	Heatwaves	heatwav	Environment/Climate
219	Wildfire mitigation	wildfir mitig	Environment/Climate
220	Solar panels	solar panel	Environment/Climate
221	Systemic racism	system racism	Ethnic and Racial Inequality

222	Police brutality	polic brutal	Ethnic and Racial Inequality
223	Racial justice	racial justic	Ethnic and Racial Inequality
224	Critical race theory	critic race theori	Ethnic and Racial Inequality
225	Juneteenth	juneteenth	Ethnic and Racial Inequality
226	Anti-Asian hate	antiasian hate	Ethnic and Racial Inequality
227	Racial wealth gap	racial wealth gap	Ethnic and Racial Inequality
228	Equity initiatives	equiti initi	Ethnic and Racial Inequality
229	Reparations debate	repar debat	Ethnic and Racial Inequality
230	Indigenous rights	indigen right	Ethnic and Racial Inequality
231	Islamophobia	islamophobia	Ethnic and Racial Inequality

232	Anti-Semitism	antisemit	Ethnic and Racial Inequality
233	Xenophobia	xenophobia	Ethnic and Racial Inequality
234	Hate speech	hate speech	Ethnic and Racial Inequality
235	Black Lives Matter	black live matter	Ethnic and Racial Inequality