

# TP4

Claude Alie

## **Introduction**

## **Méthode**

2017.

## **Résultats**

Élections américaines de 2020

Nettoyage et scrapping de données

## **Wiki**

Clinton vs. Trump

Dark triad images

## **PDF**

Predict dark triad

Distinguishing different types of dark personalities

Tweets?

Obama vs. ddd

# Clinton vs. Trump

## Élections américaines de 2020

Nettoyage et scrapping de données

#Sauver les données scrappées pour pouvoir les relire sans scrapper à nouveau

	source
1	Ipsos/Reuters
2	YouGov/Economist
3	Research Co
4	IBD/TIPI
5	IBD/TIPI
6	USC Dornsife
7	USC Dornsife
8	Swayable
9	John Zogby Strategies/EMI Research
10	John Zogby Strategies/EMI Research
11	Redfield & Wilton Strategies
12	SurveyMonkey/Axios
13	YouGov/Yahoo News
14	Change Research/CNBC
15	Qriously/Brandwatch
16	Léger
17	Quinnipiac University
18	Pulse Opinion Research/Rasmussen Reports
19	AYTM/Aspiration
20	SurveyMonkey/Axios
21	Morning Consult
22	Swayable
23	RMG Research/Just the News
24	RMG Research/Just the News
25	RMG Research/Just the News
26	SurveyUSA/Cheddar
27	NBC/WSJ
28	IBD/TIPI
29	IBD/TIPI
30	Data for Progress
31	Gravis Marketing
32	Morning Consult
33	Fox News
34	Opinium/The Guardian

35		Swayable
36		Harvard-Harris
37		AtlasIntel
38		Pulse Opinion Research/Rasmussen Reports
39		SurveyMonkey/Axios
40		JL Partners/The Independent
41		HarrisX/The Hill
42		HarrisX/The Hill
43		Angus Reid Global
44		SurveyMonkey/Long Island University
45		YouGov/Economist
46	Global Marketing Research/FPU/Boston Herald[1]	Archived 2020-11-27 at the Wayback Machine
47		Ipsos/Reuters
48		Suffolk University/USA Today[2]
49		Suffolk University/USA Today[2]
50	YouGov/University of Massachusetts Amherst	[permanent dead link]
51		Redfield & Wilton Strategies
52		Emerson College
53		Morning Consult
54		Swayable
55	Winston Group (R)	Archived 2020-12-23 at the Wayback Machine
56	CNN/SSRS	Archived 2020-12-22 at the Wayback Machine
57		Qriously/Brandwatch
58		IBD/TIPI
59		IBD/TIPI
60		YouGov/Hofstra University
61		YouGov/GW Politics
62		Cometrends/University of Dallas
63		Lucid/Tufts University
64		Léger
65		SurveyMonkey/Axios
66		YouGov/Yahoo News
67		Pulse Opinion Research/Rasmussen Reports
68		Change Research/Crooked Media
69		RMG Research/Just the News
70		RMG Research/Just the News
71		RMG Research/Just the News
72	CNBC/Hart Research	Public Opinion Strategies
73		Morning Consult
74		Spry Strategies
75	Spry Strategies/Women's Liberation Front	
76		SurveyMonkey/Axios
77		Ipsos/Reuters

78		IBD/TIPI
79		IBD/TIPI
80	Rethink Priorities	Archived 2020-12-22 at the Wayback Machine
81		Data for Progress
82		YouGov/Economist
83		Morning Consult
84		Echelon Insights
85		Echelon Insights
86	Ipsos/Reuters	Archived 2020-11-26 at the Wayback Machine
87		Pulse Opinion Research/Rasmussen Reports
88		SurveyMonkey/Axios
89		SurveyUSA/Cheddar
90		Quinnipiac University
91		GSG/GBAO
92		Qriously/Brandwatch
93		GBAO/Omidyar Network
94		USC Dornsife
95		USC Dornsife
96		Change Research/CNBC
97		Redfield & Wilton Strategies
98		Research Co
99		Léger
100		YouGov/Yahoo News
101		Siena College/NYT Upshot
102		Morning Consult
103		RMG Research/Just the News
104		SurveyMonkey/Axios
105		IBD/TIPI
106		IBD/TIPI
107		HarrisX/The Hill
108		Ipsos/Reuters
109		Morning Consult
110		JL Partners/The Independent
111		SurveyMonkey/Axios
112		YouGov/Economist
113		HarrisX/The Hill
114	Ipsos/Reuters	Archived 2021-01-24 at the Wayback Machine
115		Marist College/NPR/PBS
116	Whitman Insight Strategies	Archived 2020-10-22 at the Wayback Machine
117		Pulse Opinion Research/Rasmussen Reports
118		Public Religion Research Institute
119		Public Religion Research Institute
120		NBC/WSJ

121	AP-NOR
122	GSG/GBA
123	Qriously/Brandwatch
124	Opinium/The Guardian
125	Kaiser Family Foundation
126	Public First
127	YouGov/UMass Lowell
128	Morning Consult
129	YouGov/Yahoo News
130	Léger
131	IBD/TIPI
132	IBD/TIPI
133	Redfield & Wilton Strategies
134	SurveyMonkey/Axios
135	RMG Research/Just the News
136	RMG Research/Just the News
137	RMG Research/Just the News
138	Morning Consult
139	YouGov/CCE
140	ABC/Washington Post
141	ABC/Washington Post
142	Ipsos/Reuters
143	Public Opinion Strategies/Conservative Energy Network[AN]
144	Edison Research
145	Ipsos/Reuters
146	SurveyMonkey/Axios
147	Data For Progress
148	Morning Consult
149	YouGov/Economist
150	HarrisX/The Hill
151	Fox News
152	Ipsos/Reuters
153	Ipsos/Reuters
154	Pulse Opinion Research/Rasmussen Reports
155	Innovative Research Group
156	GSG/GBA
157	Pew Research
158	USC Dornsife
159	USC Dornsife
160	Redfield & Wilton Strategies
161	Léger
162	SurveyMonkey/Axios
163	Change Research/CNBC

164 Qriously/Brandwatch  
 165 SurveyUSA  
 166 CNN/SSRS  
 167 Global Marketing Research/FPU/Boston Herald  
 168 NBC/WSJ  
 169 Ipsos/Reuters  
 170 YouGov/Yahoo News  
 171 RMG Research/Just the News  
 172 RMG Research/Just the News  
 173 RMG Research/Just the News  
 174 Morning Consult  
 175 Zogby Strategies/EMI Research  
 176 Zogby Strategies/EMI Research  
 177 YouGov/Yahoo News  
 178 St. Leo University  
 179 HarrisX/The Hill Archived 2020-12-10 at the Wayback Machine  
 180 Data for Progress  
 181 IBD/TIPI  
 182 SurveyMonkey/Long Island University  
 183 SurveyMonkey/Axiom  
 184 Ipsos/Reuters

	date	error_margin	Republican_cand	Democrate_cand
1	Oct 31	37	45	52
2	Oct 31	NA	43	53
3	Oct 31	30	42	50
4	Oct 29	32	46	50
5	Oct 29	32	46	51
6	Oct 20	NA	42	54
7	Oct 20	NA	43	54
8	Nov 1	17	46	52
9	Nov 1	32	43	48
10	Nov 1	32	45	52
11	Oct 30	NA	41	53
12	Oct 30	1	47	52
13	Oct 30	NA	43	53
14	Oct 29	226	42	52
15	Oct 29	NA	41	52
16	Oct 29	31	42	50
17	Oct 28	25	39	50
18	Oct 28	25	47	48
19	Oct 30-31	37	39	48
20	Oct 29-31	NA	46	52
21	Oct 29-31	1	44	52

22	Oct 29-31	24	46	52
23	Oct 29-31	28	44	51
24	Oct 29-31	28	42	53
25	Oct 29-31	28	45	50
26	Oct 29-31	32	44	52
27	Oct 29-31	34	42	52
28	Oct 27-31	32	45	49
29	Oct 27-31	32	45	50
30	Oct 28-29	26	44	54
31	Oct 27-29	27	44	50
32	Oct 27-29	1	43	52
33	Oct 27-29	25	44	52
34	Oct 26-29	NA	41	55
35	Oct 27-28	29	46	53
36	Oct 27-28	NA	46	54
37	Oct 26-28	2	46	51
38	Oct 26-28	25	47	48
39	Oct 26-28	NA	47	51
40	Oct 26-28	NA	41	55
41	Oct 25-28	2	45	49
42	Oct 25-28	2	47	53
43	Oct 23-28	21	45	53
44	Oct 26-27	35	36	47
45	Oct 25-27	NA	43	54
46	Oct 23-27	3	39	53
47	Oct 23-27	39	42	52
48	Oct 23-27	31	43	50
49	Oct 23-27	31	44	52
50	Oct 20-27	31	44	53
51	Oct 25-26	NA	41	51
52	Oct 25-26	28	47	51
53	Oct 24-26	1	43	52
54	Oct 23-26	12	46	51
55	Oct 23-26	NA	43	48
56	Oct 23-26	38	42	54
57	Oct 22-26	28	39	49
58	Oct 22-26	32	46	51
59	Oct 22-26	32	46	50
60	Oct 19-26	2	43	54
61	Oct 16-26	2	41	52
62	Oct 13-26	2	44	56
63	Oct 25	NA	45	52
64	Oct 23-25	31	41	49

65	Oct 23-25	NA	46	52
66	Oct 23-25	NA	42	54
67	Oct 21-22, Oct 25	25	48	47
68	Oct 23-24	3	43	51
69	Oct 23-24	28	44	51
70	Oct 23-24	28	43	53
71	Oct 23-24	28	46	50
72	Oct 21-24	35	40	51
73	Oct 21-23	1	43	52
74	Oct 20-23	31	48	46
75	Oct 20-23	31	46	48
76	Oct 20-22	NA	46	52
77	Oct 20-22	NA	43	51
78	Oct 17-21	32	45	50
79	Oct 17-21	32	46	50
80	Oct 20	2	42	51
81	Oct 20	NA	44	54
82	Oct 18-20	NA	43	52
83	Oct 18-20	1	43	52
84	Oct 16-20	NA	44	50
85	Oct 16-20	NA	44	51
86	Oct 16-20	36	42	51
87	Oct 14-15, Oct 18-20	2	46	49
88	Oct 17-19	NA	46	52
89	Oct 16-19	32	43	53
90	Oct 16-19	26	41	51
91	Oct 15-19	31	43	53
92	Oct 15-19	26	40	51
93	Oct 15-19	NA	40	53
94	Oct 6-19	NA	41	54
95	Oct 6-19	NA	42	54
96	Oct 17-18	19	42	52
97	Oct 17-18	NA	40	51
98	Oct 16-18	30	42	50
99	Oct 16-18	31	41	50
100	Oct 16-18	4	40	51
101	Oct 15-18	34	41	50
102	Oct 15-17	1	43	52
103	Oct 15-17	28	43	51
104	Oct 14-16	NA	45	53
105	Oct 12-16	32	43	50
106	Oct 12-16	32	43	50
107	Oct 13-15	225	42	46



108	Oct 13-15	NA	41	51
109	Oct 12-14	1	43	52
110	Oct 13	NA	42	52
111	Oct 11-13	NA	46	52
112	Oct 11-13	NA	42	52
113	Oct 10-13	183	40	47
114	Oct 9-13	38	41	51
115	Oct 8-13	38	43	54
116	Oct 8-13	29	42	54
117	Oct 7-8, Oct 11-13	2	45	50
118	Oct 9-12	NA	38	56
119	Oct 9-12	NA	40	54
120	Oct 9-12	31	42	53
121	Oct 8-12	4	36	51
122	Oct 8-12	31	43	53
123	Oct 8-12	28	38	52
124	Oct 8-12	NA	40	57
125	Oct 7-12	3	38	49
126	Oct 6-12	NA	34	47
127	Oct 5-12	43	43	53
128	Oct 9-11	1	43	51
129	Oct 9-11	NA	43	51
130	Oct 9-11	31	39	50
131	Oct 7-11	35	43	52
132	Oct 7-11	35	42	53
133	Oct 10	NA	41	49
134	Oct 8-10	NA	46	52
135	Oct 8-10	28	43	51
136	Oct 8-10	28	41	53
137	Oct 8-10	28	45	50
138	Oct 7-9	1	43	51
139	Sep 29	NA	43	51
140	Oct 6-9	4	42	54
141	Oct 6-9	4	43	55
142	Oct 6-8	NA	41	53
143	Oct 5-8	NA	41	55
144	Sep 25	NA	35	48
145	Sep 22	35	39	46
146	Oct 5-7	NA	45	53
147	Oct 6	NA	41	56
148	Oct 4-6	1	43	51
149	Oct 4-6	NA	42	51
150	Oct 3-6	184	40	45

151	Oct 3-6	3	43	53		
152	Oct 2-6	38	40	52		
153	Oct 2-6	38	40	52		
154	Sep 30	2	40	52		
155	Sep 29	NA	42	47		
156	Oct 2-5	31	44	52		
157	Sep 30	15	42	52		
158	Sep 22	NA	42	54		
159	Sep 22	NA	42	53		
160	Oct 3-4	NA	42	50		
161	Oct 2-4	31	40	49		
162	Oct 2-4	NA	46	52		
163	Oct 2-4	211	42	52		
164	Oct 1-4	27	38	51		
165	Oct 1-4	36	43	53		
166	Oct 1-4	36	41	57		
167	Sep 30	31	37	51		
168	Oct 2-3	346	39	53		
169	Oct 2-3	5	41	51		
170	Oct 2-3	NA	43	51		
171	Oct 1-3	35	43	51		
172	Oct 1-3	35	41	53		
173	Oct 1-3	35	45	49		
174	Oct 1-3	1	43	52		
175	Oct 2	32	45	47		
176	Oct 2	32	47	49		
177	Oct 1-2	NA	40	48		
178	Sep 27	3	38	52		
179	Sep 30	32	40	47		
180	Sep 30	29	41	51		
181	Sep 30	35	46	49		
182	Sep 30	35	31	48		
183	Sep 29	NA	46	52		
184	Sep 29	38	41	50		
	Libertarien_cand	Vert_cand	Autres_cand	Abstention_cand	Indécis_cand	lead
1	NA	NA	3	NA	NA	7
2	NA	NA	2	0	2	10
3	1	1	1	NA	5	8
4	2	1	1	NA	NA	4
5	NA	NA	NA	NA	NA	5
6	NA	NA	NA	NA	NA	12
7	NA	NA	NA	NA	NA	11
8	2	0	NA	NA	NA	6

9	4	2	2	NA	2	5
10	NA	NA	NA	NA	3	7
11	1	1	NA	NA	NA	12
12	NA	NA	NA	NA	NA	5
13	NA	NA	2	0	2	10
14	2	1	2	NA	1	10
15	2	1	NA	NA	NA	11
16	2	1	1	0	4	8
17	NA	NA	2	NA	9	11
18	NA	NA	3	NA	2	1
19	NA	NA	NA	NA	NA	9
20	NA	NA	NA	NA	NA	6
21	NA	NA	2	NA	3	8
22	2	0	NA	NA	NA	6
23	1	1	2	NA	1	7
24	1	1	2	NA	1	11
25	1	1	2	NA	1	5
26	NA	NA	3	NA	2	8
27	NA	NA	3	NA	3	10
28	3	1	0	NA	0	4
29	NA	NA	NA	NA	NA	5
30	1	1	NA	NA	NA	10
31	NA	NA	NA	NA	6	6
32	NA	NA	2	NA	3	9
33	NA	NA	2	0	2	8
34	NA	NA	2	NA	2	14
35	1	1	NA	NA	NA	7
36	NA	NA	NA	NA	NA	8
37	NA	NA	1	1	1	5
38	NA	NA	3	NA	2	1
39	NA	NA	NA	NA	NA	4
40	NA	NA	NA	NA	NA	14
41	NA	NA	3	3	3	4
42	NA	NA	NA	NA	NA	6
43	NA	NA	2	NA	NA	8
44	NA	NA	5	NA	9	11
45	NA	NA	2	0	2	11
46	NA	NA	6	NA	4	14
47	NA	NA	5	0	2	10
48	1	1	2	0	4	7
49	NA	NA	2	NA	2	8
50	NA	NA	3	0	1	9
51	1	1	1	NA	5	10

52	NA	NA	2	NA	NA	4
53	NA	NA	2	NA	3	9
54	2	1	NA	NA	NA	5
55	NA	NA	NA	NA	9	5
56	NA	NA	1	1	2	12
57	3	1	NA	4	4	10
58	1	1	0	NA	0	5
59	NA	NA	NA	NA	NA	4
60	NA	NA	4	NA	NA	11
61	NA	NA	2	1	4	11
62	NA	NA	NA	NA	NA	12
63	NA	NA	NA	NA	NA	7
64	4	1	1	NA	5	8
65	NA	NA	NA	NA	NA	6
66	NA	NA	2	0	2	12
67	NA	NA	3	NA	2	1
68	1	1	1	0	2	9
69	1	0	1	NA	2	7
70	1	0	1	NA	2	10
71	1	0	1	NA	2	4
72	NA	NA	NA	NA	NA	11
73	NA	NA	2	NA	3	9
74	NA	NA	2	NA	4	2
75	NA	NA	2	NA	4	2
76	NA	NA	NA	NA	NA	6
77	NA	NA	4	NA	2	8
78	3	1	0	NA	1	5
79	NA	NA	NA	NA	NA	4
80	NA	NA	4	NA	4	9
81	NA	NA	2	NA	NA	10
82	NA	NA	2	0	4	9
83	NA	NA	2	NA	3	9
84	1	1	0	NA	3	6
85	NA	NA	NA	NA	5	7
86	NA	NA	4	NA	3	9
87	NA	NA	2	NA	2	3
88	NA	NA	NA	NA	NA	6
89	NA	NA	2	NA	3	10
90	NA	NA	2	NA	4	10
91	NA	NA	NA	1	3	10
92	1	0	NA	3	5	11
93	NA	NA	3	1	4	13
94	NA	NA	NA	NA	NA	13

95	NA	NA	NA	NA	NA	12
96	NA	NA	NA	NA	NA	10
97	1	1	1	NA	6	11
98	1	1	7	NA	NA	8
99	2	2	1	0	5	9
100	NA	NA	3	0	5	11
101	2	0	1	0	6	9
102	NA	NA	2	NA	3	9
103	2	1	1	NA	2	8
104	NA	NA	NA	NA	NA	8
105	2	1	1	NA	1	7
106	NA	NA	NA	NA	NA	7
107	NA	NA	3	3	6	4
108	NA	NA	4	0	4	10
109	NA	NA	2	NA	4	9
110	NA	NA	NA	NA	NA	10
111	NA	NA	NA	NA	NA	6
112	NA	NA	1	0	4	10
113	NA	NA	3	3	7	7
114	NA	NA	4	NA	4	10
115	NA	NA	1	NA	2	11
116	NA	NA	1	NA	3	12
117	NA	NA	2	NA	3	5
118	NA	NA	NA	NA	NA	18
119	NA	NA	NA	NA	NA	14
120	NA	NA	3	NA	2	11
121	NA	NA	7	6	0	15
122	NA	NA	NA	1	3	10
123	1	1	NA	3	6	14
124	NA	NA	1	NA	2	17
125	NA	NA	5	NA	8	11
126	NA	NA	3	8	8	13
127	1	1	0	NA	3	10
128	NA	NA	2	NA	4	8
129	NA	NA	2	0	4	8
130	3	1	1	1	6	11
131	2	1	0	NA	0	9
132	NA	NA	NA	NA	NA	11
133	1	1	1	NA	7	8
134	NA	NA	NA	NA	NA	6
135	2	1	0	NA	2	8
136	2	1	0	NA	2	12
137	2	1	0	NA	2	5

138	NA	NA	2	NA	4	8
139	NA	NA	NA	NA	NA	8
140	2	1	0	0	2	12
141	NA	NA	0	1	1	12
142	NA	NA	2	0	3	12
143	NA	NA	NA	NA	4	14
144	NA	NA	NA	NA	NA	13
145	NA	NA	5	5	5	7
146	NA	NA	NA	NA	NA	8
147	NA	NA	3	NA	NA	15
148	NA	NA	2	NA	4	8
149	NA	NA	2	0	5	9
150	NA	NA	3	4	7	5
151	NA	NA	1	NA	3	10
152	1	1	3	NA	3	12
153	NA	NA	4	NA	4	12
154	NA	NA	4	NA	4	12
155	NA	NA	1	2	9	5
156	NA	NA	NA	1	4	8
157	4	1	1	NA	0	10
158	NA	NA	NA	NA	NA	12
159	NA	NA	NA	NA	NA	11
160	1	1	1	NA	6	8
161	2	1	1	1	6	9
162	NA	NA	NA	NA	NA	6
163	3	1	NA	1	2	10
164	1	0	NA	3	6	13
165	NA	NA	2	NA	3	10
166	NA	NA	1	0	1	16
167	NA	NA	3	NA	8	14
168	NA	NA	2	NA	6	14
169	NA	NA	4	NA	4	10
170	NA	NA	2	0	5	8
171	1	1	1	NA	3	8
172	1	1	1	NA	3	12
173	1	1	1	NA	3	4
174	NA	NA	2	NA	4	9
175	2	2	NA	NA	4	2
176	NA	NA	NA	NA	4	2
177	NA	NA	3	0	8	8
178	NA	NA	NA	NA	6	14
179	NA	NA	3	3	7	7
180	NA	NA	NA	NA	8	10

181	NA	NA	15	NA	4	3
182	NA	NA	7	5	9	17
183	NA	NA	NA	NA	NA	6
184	NA	NA	4	NA	5	9

# A tibble: 6 x 6

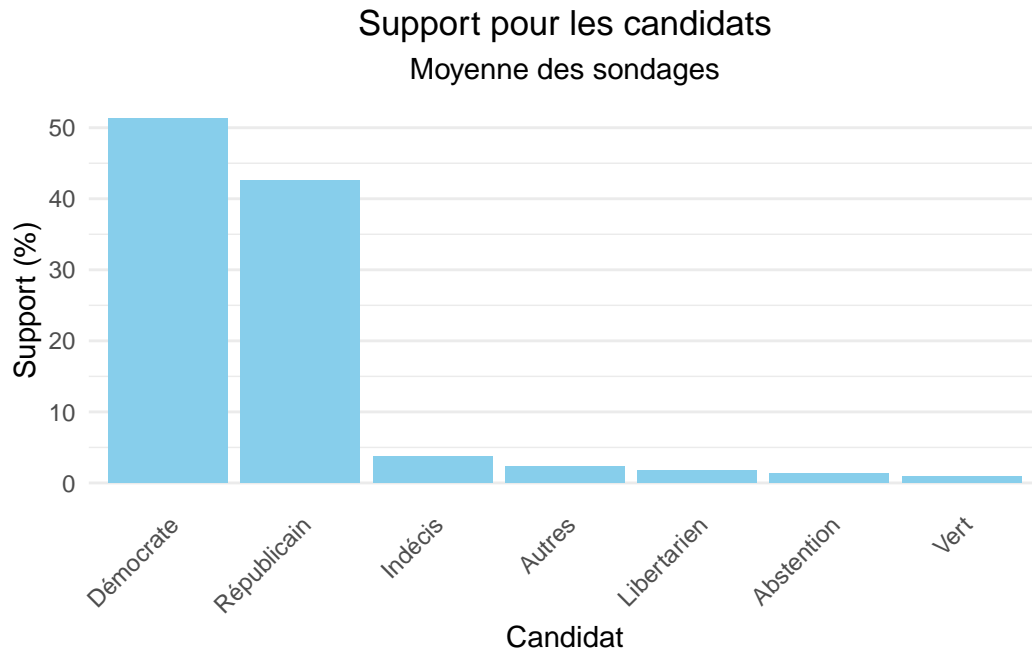
	source	date	error_margin	lead	candidats	appui
	<chr>	<chr>	<int>	<int>	<chr>	<int>
1	Ipsos/Reuters	Oct 31	37	7	Republicain_cand	45
2	Ipsos/Reuters	Oct 31	37	7	Democrate_cand	52
3	Ipsos/Reuters	Oct 31	37	7	Libertarien_cand	NA
4	Ipsos/Reuters	Oct 31	37	7	Vert_cand	NA
5	Ipsos/Reuters	Oct 31	37	7	Autres_cand	3
6	Ipsos/Reuters	Oct 31	37	7	Abstention_cand	NA

Préparation des données moyennes

# A tibble: 7 x 2

	candidats	appui_moy
	<chr>	<dbl>
1	Abstention	1.38
2	Autres	2.25
3	Démocrate	51.4
4	Indécis	3.68
5	Libertarien	1.70
6	Républicain	42.6
7	Vert	0.906

#Visualisation des données moyennes



## Biden vs. Trump

### Élections américaines de 2020

```
Table_results_2020 <- read_html(
  "https://en.wikipedia.org/wiki/2020_United_States_presidential_election"
) |>
  ### Chercher l'élément (ou node) "table" ###
  html_elements("table") |>
  ### Chercher le deuxième élément (ou table) de la liste ###
  pluck(28) |>
  ### Transformation en matrice de données ###
  html_table(fill = T)

view(Table_results_2020)
Table_results_2020
```

# A tibble: 8 x 9

	`Presidential candidate` <chr>	Party <chr>	`Home state` <chr>	`Popular vote` <chr>	`Popular vote` <chr>
1	Presidential candidate	Party	Home state	Count	Percentage



```

2 Joe Biden Democratic Delaware 81,283,501 51.31%
3 Donald Trump(incumbent) Republican Florida 74,223,975 46.85%
4 Jo Jorgensen Libertari~ South Carol~ 1,865,535 1.18%
5 Howie Hawkins Green New York 407,068 0.26%
6 Other Other Other 649,552 0.41%
7 Total Total Total 158,429,631 100%
8 Needed to win Needed to~ Needed to w~ Needed to win Needed to win
# i 4 more variables: Electoralvote <chr>, `Running mate` <chr>,
# `Running mate` <chr>, `Running mate` <chr>

```

```
names(Table_results_2020)
```

```

[1] "Presidential candidate" "Party" "Home state"
[4] "Popular vote" "Popular vote" "Electoralvote"
[7] "Running mate" "Running mate" "Running mate"

```

## A jeter?

```

dat_resultats_2020 <- Table_results_2020 %>% select(2, 5) |> rename(candidats = "Party",
resultat = "Popular vote") %>% slice(-1) %>% mutate(resultat = str_remove(resultat,
"%"))

```

```

# A tibble: 7 x 2
  candidats      resultat
  <chr>          <dbl>
1 Démocrate      51.3
2 Républicain    46.8
3 Libertarien    1.18
4 Vert           0.26
5 Autres         0.41
6 Total          100
7 Needed to win  NA

```

```

joined_data <- inner_join(dat_moy, dat_resultats_2020, by = "candidats")
joined_data

```

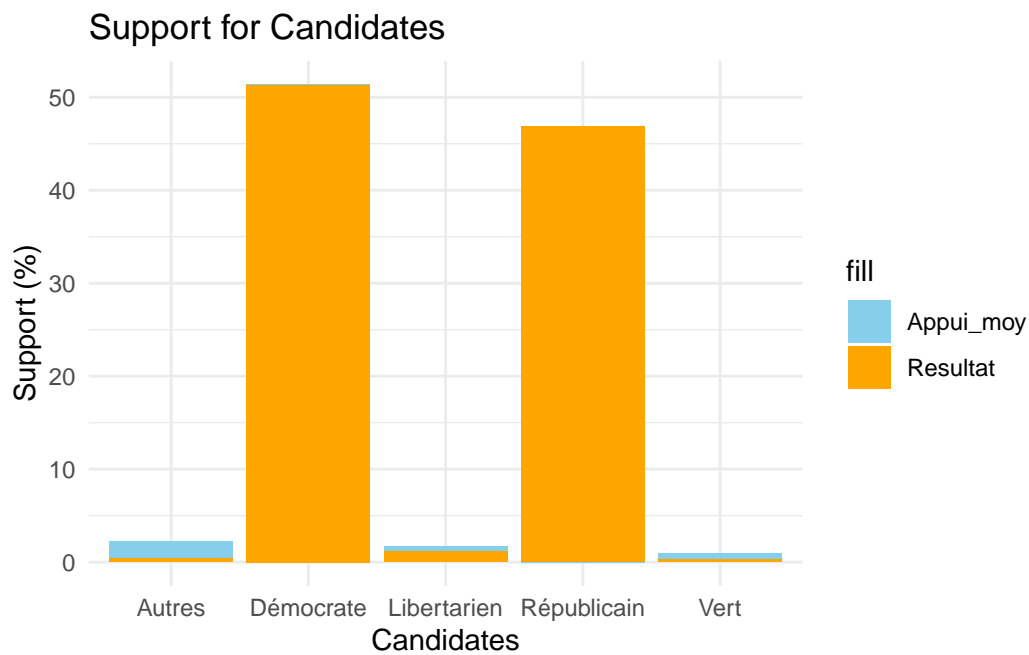
```

# A tibble: 5 x 3
  candidats appui_moy resultat

```

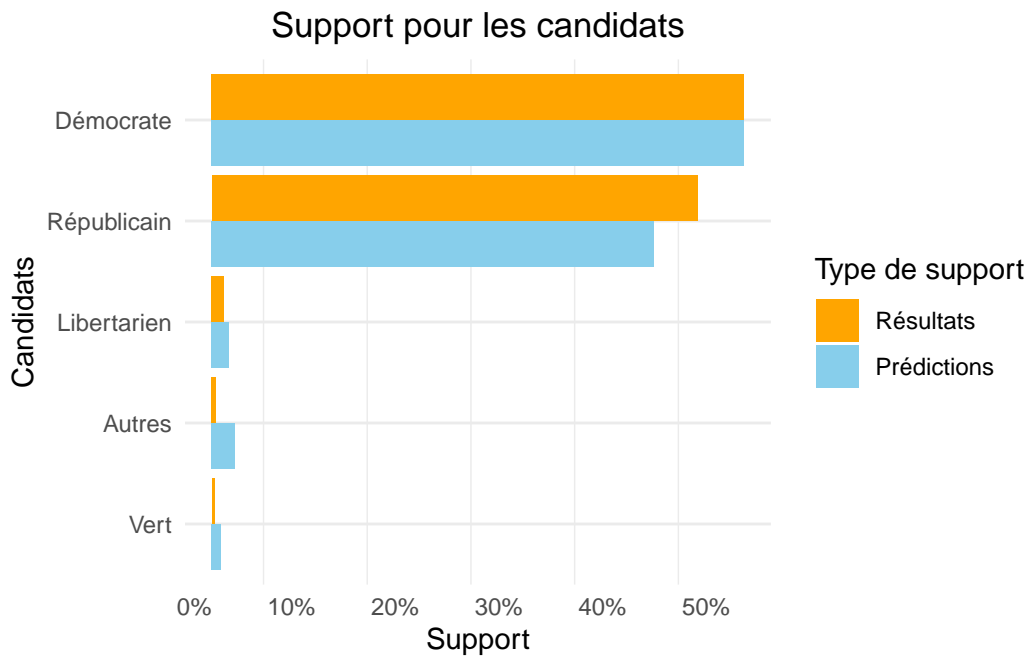
	<chr>	<dbl>	<dbl>
1	Autres	2.25	0.41
2	Démocrate	51.4	51.3
3	Libertarien	1.70	1.18
4	Républicain	42.6	46.8
5	Vert	0.906	0.26

## Plotting joined\_data with side-by-side bars



```
# A tibble: 10 x 3
  candidats Type      Support
  <chr>      <chr>      <dbl>
1 Autres    appui_moy    2.25
2 Autres    resultat     0.41
3 Démocrate appui_moy   51.4
4 Démocrate resultat  51.3
5 Libertarien appui_moy    1.70
6 Libertarien resultat    1.18
7 Républicain appui_moy   42.6
8 Républicain resultat   46.8
9 Vert      appui_moy    0.906
```

10 Vert      resultat      0.26



## Triage sombre

### Obtenir les urls des images

```
url <- "https://en.wikipedia.org/wiki/Dark_triad"
webpage <- read_html(url)
img_tags <- html_nodes(webpage, ".mw-file-element")
img_attr <- html_attr(img_tags, "src")
img_urls <- img_attr[grepl("^//", img_attr)]

# Extraire chaque image
for (url in img_urls) {
  url <- paste0("https:", url)
  filename <- basename(url)
  tryCatch({
    GET(url, write_disk(filename))
    cat("Downloaded:", filename, "\n")
  }, error = function(e) {
```

```

        cat("Failed to download:", filename, "\n")
    })
}

```

Failed to download: 300px-The\_Dark\_Triad.png

Failed to download: 138px-Dark\_Triad.svg.png

Failed to download: 16px-Symbol\_category\_class.svg.png

## Prédiction des traits de personnalité de la triade sombre

### Acquérir le tableau du .pdf

Lire le tableau à son lieu de storage

Localiser le tableau d'intérêt

Localiser la page. Il se trouve en page 6.

Segmenter la page en lignes

Localiser le tableau sur la page

Il termine avec la première ligne vide rencontrée. On cherche donc ""

Le début part avec le titre du tableau. On cherche donc le mot "Table"

On parcourt chaque ligne et on enregistre dans "table\_content".

### Formatter le tableau

Standardiser les éléments du tableaux pour avoir le même nombre de colonnes

Compléter les lignes des sous-titres pour obtenir le même nombre de rangées

```

for (i in seq_along(table_content)) {
  if (!grepl(" ", table_content[[i]])) {
    print(table_content[[i]])
    table_content[[i]] <- paste0(table_content[[i]], "-----")
  }
}

```

```
}
}
```

```
[1] " Linguistic Processes"
[1] " Psychological Processes"
[1] " Personal Concerns"
[1] " Spoken categories"
[1] " Punctuation"
```

```
### Remplacer les éléments susceptibles d'affecter le nombre de colonnes
```

```
replacements <- list(
  ", " = ";",
  "Words > 6 letters" = "Words_more_6_letters",
  ", " = ";",
  "AllPct" = "AllPct ---- ",
  "([A-Za-z])\\s([A-Za-z])" = "\\1_\\2",
  "----" = ", "
)

replacements <- unlist(replacements)
table_content[1] <- str_replace_all(table_content[1], "\\s+", " ")
for (i in 2:length(table_content)) {
  table_content[i] <- as.character(table_content[i])
  for (pattern in names(replacements)) {
    table_content[i] <- gsub(pattern, replacements[pattern], table_content[i])
  }
}
table_content
```

```
[1] " Category Abbrev Examples Na Ma Ps Op Co Ex Ag Ne"
[2] " Linguistic_Processes, , , , , , , , , , "
[3] " Words_more_6_letters          Sixltr          Words_more_6_letters          0.047
[4] " Dictionary_words              Dic              Dictionary_words              -0.088
[5] " Total_function_words          funct          Total_Function_words          -0.093
[6] " Total_pronouns                pronoun      I;them;itself                -0.043      0
[7] " Personal_pronouns             ppron       I;them;her                  -0.021      0
[8] " 1st_pers_singular             i           I;me;mine                   -0.017      0
[9] " 1st_pers_plural               we          We;us;our                   0.036      -0
[10] " Impersonal_pronouns           ipron       It;it's;those               -0.077      0
[11] " [Common_verbs]               verb        Walk;went;see               -0.084      0
```

[12]	"	Auxiliary_verbs	auxverb	Am;will;have	-0.078	0
[13]	"	Past_tense	past	Went;ran;had	-0.069	-0
[14]	"	Present_tense	present	Is;does;hear	-0.068	0
[15]	"	Adverbs	adverb	Very;really;quickly	-0.088	0
[16]	"	Prepositions	preps	To;with;above	-0.065	-0
[17]	"	Conjunctions	conj	And;but;whereas	-0.057	0
[18]	"	Negations	negate	No;not;never	-0.073	0
[19]	"	Quantifiers	quant	Few;many;much	-0.056	0
[20]	"	Numbers	number	Second;thousand	0.036	0
[21]	"	Swear_words	swear	Damn;piss;fuck	0.040	0
[22]	"	Psychological_Processes, , , , , , , , , ,		"		
[23]	"	Social_processes	social	Mate;talk;they;child	0.014	-0
[24]	"	Family	family	Daughter;husband;	0.008	-0
[25]	"	Friends	friend	Buddy;friend	0.073	-0
[26]	"	Affective_processes	affect	Happy;cried;abandon	-0.038	-0
[27]	"	Positive_emotion	posemo	Love;nice;sweet	-0.020	-0
[28]	"	Negative_emotion	negemo	Hurt;ugly;nasty	-0.023	0
[29]	"	Anxiety	anx	Worried;fearful	-0.054	-0
[30]	"	Anger	anger	Hate;kill;annoyed	0.004	0
[31]	"	Sadness	sad	Crying;grief;sad	-0.038	-0
[32]	"	Cognitive_processes	cogmech	cause;know;ought	-0.074	0
[33]	"	Insight	insight	think;know;consider	-0.048	-0
[34]	"	Discrepancy	discrep	should;would;could	-0.059	0
[35]	"	Tentative	tentat	maybe;perhaps;guess	-0.102	0
[36]	"	Inclusive	incl	And;with;include	0.004	-0
[37]	"	Exclusive	excl	But;without;exclude	-0.084	0
[38]	"	Perceptual_processes	percept	Heard;feeling	-0.052	-0
[39]	"	See	see	View;saw;seen	-0.049	-0
[40]	"	Biological_processes	bio	Eat;blood;pain	0.001	0
[41]	"	Body	body	Cheek;hands;spit	0.006	0
[42]	"	Health	health	Clinic;flu;pill	-0.005	-0
[43]	"	Sexual	sexual	Horny;love;incest	0.068	0
[44]	"	Relativity	relativ	Area;bend;exit;stop	-0.028	0
[45]	"	Motion	motion	Arrive;car;go	-0.020	-0
[46]	"	Time	time	End;until;season	-0.050	0
[47]	"	Personal_Concerns, , , , , , , , , ,				
[48]	"	Work	work	Job;majors;Xerox	-0.006	-0
[49]	"	Death	death	Bury;coffin;kill	0.021	0
[50]	"	Spoken_categories, , , , , , , , , ,				
[51]	"	Assent	assent	Agree;OK;yes	0.011	0
[52]	"	Nonfluencies	nonfl	Er;hm;umm	-0.023	-0
[53]	"	Fillers	filler	Blah;I mean;youknow	0.035	0
[54]	"	Punctuation, , , , , , , , , ,		"		

[55]	"	Total_Punctuation	AllPct	,	0.043
[56]	"	Exclam	Exclam	!	0.018
[57]	"	Dash	Dash	-	-0.006
[58]	"	Quote	Quote	"	0.057
[59]	"	Apostro	Apostro	'	-0.088
[60]	"	OtherP	OtherP	@#	0.070

## Éliminer les blancs en doublon

## Faire une dernière correction nécessaire pour la conversion en data.frame

```
[1] "Category Abbrev Examples Na Ma Ps Op Co Ex Ag Ne"
```

[2]	"Linguistic_Processes,	, , ,	, , , , , "
[3]	"Words_more_6_letters Sixltr Words_more_6_letters	0.047 -0.078 -0.047 0.035 0.042 0.042	
[4]	"Dictionary_words Dic Dictionary_words	-0.088 0.001 -0.050 -0.073 -0.051 -0.031 -0.007	
[5]	"Total_function_words funct Total_Function_words	-0.093 0.021 -0.033 -0.076 -0.094 -0.0	
[6]	"Total_pronouns pronoun I;them;itself	-0.043 0.016 -0.023 -0.027 -0.100 -0.004 -0.003	
[7]	"Personal_pronouns ppron I;them;her	-0.021 0.021 -0.017 -0.016 -0.092 0.022 0.013 -0.14	
[8]	"1st_pers_singular i I;me;mine	-0.017 0.050 -0.001 -0.011 -0.116 -0.028 -0.023 -0.171"	
[9]	"1st_pers_plural we We;us;our	0.036 -0.070 -0.071 -0.006 0.052 0.063 0.068 0.044"	
[10]	"Impersonal_pronouns ipron It;it's;those	-0.077 0.008 -0.033 -0.051 -0.099 -0.060 -0.03	
[11]	"[Common_verbs] verb Walk;went;see	-0.084 0.014 -0.028 -0.083 -0.071 -0.013 0.002 -0.13	
[12]	"Auxiliary_verbs auxverb Am;will;have	-0.078 0.025 -0.018 -0.074 -0.098 -0.021 -0.021 -	
[13]	"Past_tense past Went;ran;had	-0.069 -0.003 -0.040 -0.066 -0.051 0.001 0.000 -0.074"	
[14]	"Present_tense present Is;does;hear	-0.068 0.017 -0.024 -0.061 -0.071 -0.002 0.002 -0.1	
[15]	"Adverbs adverb Very;really;quickly	-0.088 0.026 -0.035 -0.077 -0.071 -0.034 -0.022 -0.	
[16]	"Prepositions preps To;with;above	-0.065 -0.030 -0.086 -0.059 0.062 -0.024 0.021 0.009"	
[17]	"Conjunctions conj And;but;whereas	-0.057 0.031 -0.030 -0.070 -0.085 -0.056 -0.014 -0.1	
[18]	"Negations negate No;not;never	-0.073 0.068 0.034 -0.080 -0.068 -0.034 -0.069 -0.124"	
[19]	"Quantifiers quant Few;many;much	-0.056 0.010 -0.012 -0.042 -0.064 -0.045 -0.032 -0.085	
[20]	"Numbers number Second;thousand	0.036 0.064 0.030 -0.043 0.082 0.038 -0.046 0.117"	
[21]	"Swear_words swear Damn;piss;fuck	0.040 0.129 0.187 -0.028 -0.171 0.025 -0.136 -0.097"	
[22]	"Psychological_Processes,	, , , , , , , , , , , "	
[23]	"Social_processes social Mate;talk;they;child	0.014 -0.060 -0.050 -0.007 -0.021 0.101 0	
[24]	"Family family Daughter;husband;	0.008 -0.036 -0.076 0.012 -0.007 0.041 0.102 -0.020"	
[25]	"Friends friend Buddy;friend	0.073 -0.030 0.001 0.048 -0.032 0.121 0.069 -0.049"	
[26]	"Affective_processes affect Happy;cried;abandon	-0.038 -0.062 -0.050 -0.017 -0.009 0.06	
[27]	"Positive_emotion posemo Love;nice;sweet	-0.020 -0.118 -0.124 0.006 0.077 0.108 0.183 -	
[28]	"Negative_emotion negemo Hurt;ugly;nasty	-0.023 0.073 0.083 -0.049 -0.135 -0.027 -0.109	
[29]	"Anxiety anx Worried;fearful	-0.054 -0.021 -0.042 -0.034 -0.056 -0.014 -0.007 -0.179"	
[30]	"Anger anger Hate;kill;annoyed	0.004 0.116 0.151 -0.034 -0.144 -0.016 -0.145 -0.103"	
[31]	"Sadness sad Crying:grief;sad	-0.038 -0.024 -0.036 -0.035 -0.078 -0.009 -0.024 -0.111"	

```
[32] "Cognitive_processes cogmch cause;know;ought -0.074 0.010 -0.029 -0.046 -0.090 -0.061 -
[33] "Insight insight think;know;consider -0.048 -0.022 -0.042 -0.001 -0.082 -0.060 -0.008 -
[34] "Discrepancy discrep should;would;could -0.059 0.012 -0.013 -0.053 -0.069 -0.024 -0.015 -
[35] "Tentative tentat maybe;perhaps;guess -0.102 0.020 -0.023 -0.036 -0.088 -0.081 -0.046 -
[36] "Inclusive incl And;with;include 0.004 -0.015 -0.054 -0.031 -0.001 0.032 0.027 -0.016"
[37] "Exclusive excl But;without;exclude -0.084 0.044 -0.002 -0.056 -0.100 -0.062 -0.068 -
[38] "Perceptual_processes percept Heard;feeling -0.052 -0.040 -0.092 -0.015 -0.020 -0.031 0
[39] "See see View;saw;seen -0.049 -0.036 -0.082 -0.043 0.038 -0.004 0.062 -0.039"
[40] "Biological_processes bio Eat;blood;pain 0.001 0.009 0.012 -0.026 -0.066 0.023 0.014 -
[41] "Body body Cheek;hands;spit 0.006 0.053 0.066 -0.021 -0.077 0.017 -0.029 -0.112"
[42] "Health health Clinic;flu;pill -0.005 -0.032 -0.019 0.007 -0.010 -0.005 0.032 -0.106"
[43] "Sexual sexual Horny;love;incest 0.068 0.030 0.051 0.029 -0.083 0.082 0.008 -0.156"
[44] "Relativity relativ Area;bend;exit;stop -0.028 0.021 -0.059 -0.060 0.088 0.019 0.021 0.
[45] "Motion motion Arrive;car;go -0.020 -0.008 -0.060 -0.029 0.075 0.046 0.055 0.000"
[46] "Time time End;until;season -0.050 0.041 -0.063 -0.070 0.061 -0.005 0.026 -0.054"
[47] "Personal_Concerns, , , , , , , , , ,"
[48] "Work work Job;majors;Xerox -0.006 -0.046 -0.077 -0.009 0.061 -0.015 -0.016 0.048"
[49] "Death death Bury;coffin;kill 0.021 0.039 0.094 0.050 -0.070 -0.051 -0.085 -0.003"
[50] "Spoken_categories, , , , , , , , , ,"
[51] "Assent assent Agree;OK;yes 0.011 0.013 0.038 0.024 -0.067 0.060 0.016 -0.125"
[52] "Nonfluencies nonfl Er;hm;umm -0.023 -0.027 0.005 -0.008 -0.041 0.051 -0.015 -0.041"
[53] "Fillers filler Blah;I mean;you know 0.035 0.102 0.073 0.000 -0.121 -0.024 -0.066 -0.127"
[54] "Punctuation, , , , , , , , , ,"
[55] "Total_Punctuation AllPct , 0.043 -0.039 -0.031 0.093 0.028 0.000 0.055 0.024"
[56] "Exclam Exclam ! 0.018 -0.003 -0.054 -0.016 0.050 0.107 0.104 -0.011"
[57] "Dash Dash - -0.006 -0.052 -0.025 0.046 0.042 -0.057 0.045 0.048"
[58] "Quote Quote "" 0.057 0.020 -0.035 0.050 -0.055 -0.098 -0.020 -0.034"
[59] "Apostro Apostro ‘ -0.088 0.012 -0.022 -0.028 -0.101 -0.085 -0.017 -0.129"
[60] "OtherP OtherP @# 0.070 -0.015 0.016 0.051 0.046 0.062 0.033 0.060"
```

## Vérifier le nombre de colonnes

```
[1] 11
[1] 11
[1] 11
[1] 11
[1] 11
[1] 11
[1] 11
[1] 11
[1] 11
```





```
[1] 11
[1] 11
[1] 11
[1] 11
[1] 11
[1] 11
[1] 11
[1] 11
```

## Conversion du tableau de chaînes de caractères en data.frame

```
df <- data.frame()
for (row in table_content) {
  words <- unlist(strsplit(row, " "))
  words_df <- data.frame(words)
  words_df <- t(words_df)
  df <- rbind(df, words_df)
}
df <- df[-nrow(df), ] # La dernière rangée est un doublon
```

## Effectuer les dernières corrections plus faciles à effectuer en data.frame

	V1	V2	V3	V4	V5	V6
	Category	Abbrev	Examples	Na	Ma	Ps
1	Linguistic_Processes					
2	Words > 6 letters	Sixltr	Words > 6 letters	0.047	-0.078	-0.047
3	Dictionary_words	Dic	Dictionary_words	-0.088	0.001	-0.050
4	Total_function_words	funct	Total_Function_words	-0.093	0.021	-0.033
5	Total_pronouns	pronoun	I, them, itself	-0.043	0.016	-0.023
6	Personal_pronouns	ppron	I, them, her	-0.021	0.021	-0.017
7	1st_pers_singular	i	I, me, mine	-0.017	0.050	-0.001
8	1st_pers_plural	we	We, us, our	0.036	-0.070	-0.071
9	Impersonal_pronouns	ipron	It, it's, those	-0.077	0.008	-0.033
10	[Common_verbs]	verb	Walk, went, see	-0.084	0.014	-0.028
11	Auxiliary_verbs	auxverb	Am, will, have	-0.078	0.025	-0.018
12	Past_tense	past	Went, ran, had	-0.069	-0.003	-0.040
13	Present_tense	present	Is, does, hear	-0.068	0.017	-0.024
14	Adverbs	adverb	Very, really, quickly	-0.088	0.026	-0.035
15	Prepositions	preps	To, with, above	-0.065	-0.030	-0.086
16	Conjunctions	conj	And, but, whereas	-0.057	0.031	-0.030

18	Negations	negate	No, not, never	-0.073	0.068	0.034
19	Quantifiers	quant	Few, many, much	-0.056	0.010	-0.012
20	Numbers	number	Second, thousand	0.036	0.064	0.030
21	Swear_words	swear	Damn, piss, fuck	0.040	0.129	0.187
22	Psychological_Processes					
23	Social_processes	social	Mate, talk, they, child	0.014	-0.060	-0.050
24	Family	family	Daughter, husband,	0.008	-0.036	-0.076
25	Friends	friend	Buddy, friend	0.073	-0.030	0.001
26	Affective_processes	affect	Happy, cried, abandon	-0.038	-0.062	-0.050
27	Positive_emotion	posemo	Love, nice, sweet	-0.020	-0.118	-0.124
28	Negative_emotion	negemo	Hurt, ugly, nasty	-0.023	0.073	0.083
29	Anxiety	anx	Worried, fearful	-0.054	-0.021	-0.042
30	Anger	anger	Hate, kill, annoyed	0.004	0.116	0.151
31	Sadness	sad	Crying, grief, sad	-0.038	-0.024	-0.036
32	Cognitive_processes	cogmech	cause, know, ought	-0.074	0.010	-0.029
33	Insight	insight	think, know, consider	-0.048	-0.022	-0.042
34	Discrepancy	discrep	should, would, could	-0.059	0.012	-0.013
35	Tentative	tentat	maybe, perhaps, guess	-0.102	0.020	-0.023
36	Inclusive	incl	And, with, include	0.004	-0.015	-0.054
37	Exclusive	excl	But, without, exclude	-0.084	0.044	-0.002
38	Perceptual_processes	percept	Heard, feeling	-0.052	-0.040	-0.092
39	See	see	View, saw, seen	-0.049	-0.036	-0.082
40	Biological_processes	bio	Eat, blood, pain	0.001	0.009	0.012
41	Body	body	Cheek, hands, spit	0.006	0.053	0.066
42	Health	health	Clinic, flu, pill	-0.005	-0.032	-0.019
43	Sexual	sexual	Horny, love, incest	0.068	0.030	0.051
44	Relativity	relativ	Area, bend, exit, stop	-0.028	0.021	-0.059
45	Motion	motion	Arrive, car, go	-0.020	-0.008	-0.060
46	Time	time	End, until, season	-0.050	0.041	-0.063
47	Personal_Concerns					
48	Work	work	Job, majors, Xerox	-0.006	-0.046	-0.077
49	Death	death	Bury, coffin, kill	0.021	0.039	0.094
50	Spoken_categories					
51	Assent	assent	Agree, OK, yes	0.011	0.013	0.038
52	Nonfluencies	nonfl	Er, hm, umm	-0.023	-0.027	0.005
53	Fillers	filler	Blah, Imean, youknow	0.035	0.102	0.073
54	Punctuation					
55	Total_Punctuation	AllPct		0.043	-0.039	-0.031
56	Exclam	Exclam	!	0.018	-0.003	-0.054
57	Dash	Dash	-	-0.006	-0.052	-0.025
58	Quote	Quote	"	0.057	0.020	-0.035
59	Apostro	Apostro	'	-0.088	0.012	-0.022
	V7	V8	V9	V10	V11	

1	Op	Co	Ex	Ag	Ne
2					
3	0.035	0.042	0.042	0.017	0.079
4	-0.073	-0.051	-0.031	-0.007	-0.113
5	-0.076	-0.094	-0.046	-0.033	-0.123
6	-0.027	-0.100	-0.004	-0.003	-0.142
7	-0.016	-0.092	0.022	0.013	-0.145
8	-0.011	-0.116	-0.028	-0.023	-0.171
9	-0.006	0.052	0.063	0.068	0.044
10	-0.051	-0.099	-0.060	-0.035	-0.115
11	-0.083	-0.071	-0.013	0.002	-0.132
12	-0.074	-0.098	-0.021	-0.021	-0.131
13	-0.066	-0.051	0.001	0.000	-0.074
14	-0.061	-0.071	-0.002	0.002	-0.140
15	-0.077	-0.071	-0.034	-0.022	-0.119
16	-0.059	0.062	-0.024	0.021	0.009
17	-0.070	-0.085	-0.056	-0.014	-0.109
18	-0.080	-0.068	-0.034	-0.069	-0.124
19	-0.042	-0.064	-0.045	-0.032	-0.085
20	-0.043	0.082	0.038	-0.046	0.117
21	-0.028	-0.171	0.025	-0.136	-0.097
22					
23	-0.007	-0.021	0.101	0.096	-0.036
24	0.012	-0.007	0.041	0.102	-0.020
25	0.048	-0.032	0.121	0.069	-0.049
26	-0.017	-0.009	0.069	0.091	-0.095
27	0.006	0.077	0.108	0.183	-0.042
28	-0.049	-0.135	-0.027	-0.109	-0.138
29	-0.034	-0.056	-0.014	-0.007	-0.179
30	-0.034	-0.144	-0.016	-0.145	-0.103
31	-0.035	-0.078	-0.009	-0.024	-0.111
32	-0.046	-0.090	-0.061	-0.030	-0.099
33	-0.001	-0.082	-0.060	-0.008	-0.094
34	-0.053	-0.069	-0.024	-0.015	-0.094
35	-0.036	-0.088	-0.081	-0.046	-0.078
36	-0.031	-0.001	0.032	0.027	-0.016
37	-0.056	-0.100	-0.062	-0.068	-0.110
38	-0.015	-0.020	-0.031	0.070	-0.089
39	-0.043	0.038	-0.004	0.062	-0.039
40	-0.026	-0.066	0.023	0.014	-0.136
41	-0.021	-0.077	0.017	-0.029	-0.112
42	0.007	-0.010	-0.005	0.032	-0.106
43	0.029	-0.083	0.082	0.008	-0.156

```

44 -0.060  0.088  0.019  0.021  0.006
45 -0.029  0.075  0.046  0.055  0.000
46 -0.070  0.061 -0.005  0.026 -0.054
47
48 -0.009  0.061 -0.015 -0.016  0.048
49  0.050 -0.070 -0.051 -0.085 -0.003
50
51  0.024 -0.067  0.060  0.016 -0.125
52 -0.008 -0.041  0.051 -0.015 -0.041
53  0.000 -0.121 -0.024 -0.066 -0.127
54
55  0.093  0.028  0.000  0.055  0.024
56 -0.016  0.050  0.107  0.104 -0.011
57  0.046  0.042 -0.057  0.045  0.048
58  0.050 -0.055 -0.098 -0.020 -0.034
59 -0.028 -0.101 -0.085 -0.017 -0.129

```

## Élections américaines de 2016

```

# A tibble: 5 x 2
  Candidats Résultat
  <chr>      <chr>
1 Clinton  48.18
2 Trump    46.09
3 Johnson  3.28
4 Stein    1.07
5 Others   1.38

```