

Tweet Corpus Description

Duration

01/01/2016-05/01/2016

Total Tweets (Election classified AND containing URL)

16,677,685 (~16.6M)

Avg 4M / month

Avg 140K / day

Total Stories in Electome DB (These are stories classified as related to the election using a machine learning classifier from my labmate, 1/1 - 5/1, with 0.8+ confidence)

22,959 (~ 23K) stories

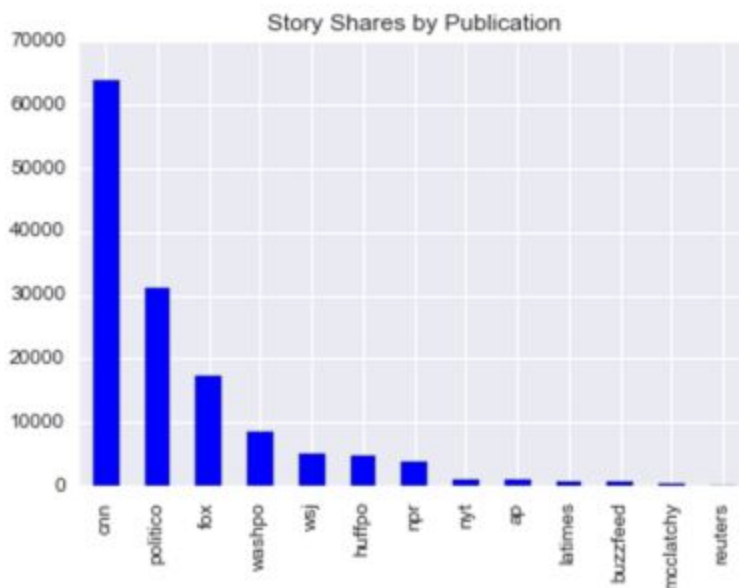
Avg 5.7 K / Month

191 / Day

Total Story Matches (This means the url in the tweet matched a story in our database.)

137,986 matches (This means ~ 140K tweets in our database contain a URL of a story in our election database.)

6,911 **unique** stories



Descriptive Stats

count	6911.000000
mean	19.940530
std	42.788286
min	1.000000
10%	1.000000
25%	2.000000
50%	6.000000
75%	19.000000
99%	194.900000
max	1262.000000

Top 25 Shared Stories

The One Weird Trait That Predicts Whether You're a Trump Supporter	1262
Donald Trump Is Shocking, Vulgar and Right	901
Biden praises Sanders on income inequality	563
Why I'm voting for Trump	554
Anne Frank's stepsister compares Donald Trump to Adolf Hitler	445
Trump basks in his spotlight	436
Rubio: Law-abiding undocumented immigrants could stay	413
Terrorists use Trump's 'Muslim ban' speech in recruitment video	398
Iowa caucuses: Donald Trump's moment of truth	364
GOP senators: If Cruz wins, we lose	357
Cruz, Kasich join forces to stop Trump	346
FOX NEWS POLL Trump gains in Iowa, dominates New Hampshire	331
Trump campaign silent on veterans' groups receiving funds	313
Poll: Trump dominates GOP field	312
National Review, conservative thinkers stand against Donald Trump	312
Republican elders ask 'who's worse' for the GOP brand: Trump or Cruz	310
Campaigns lower expectations in Iowa, brace for Trump win	308
TODD STARNES School caught trying to get students to work for Hillary	305
Donald Trump, Bernie Sanders hold solid leads in Iowa, CNN/ORC poll finds	296
NEXT ROUND: Fox News announces lineup for GOP prime time debate	292
GOP's Trump revolt: Too little too late?	287
FOX NEWS POLL Trump, Cruz top GOP race, few have 'birther' concerns	282
South Carolina heads to the polls after GOP brawl	281

Trump won't disavow KKK

275

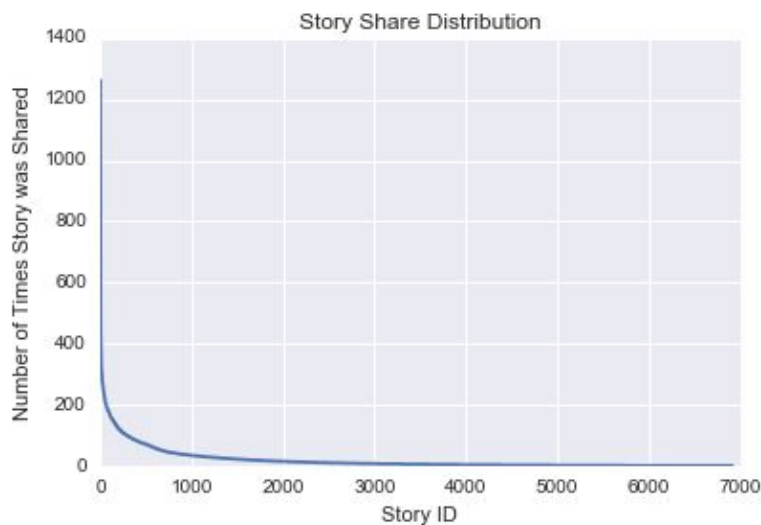
Glenn Beck: Donald Trump shooting remark 'dangerous'

274

As you can see, Trump stories gain a lot of twitter volume.

Here's a distribution of how many times stories are tweeted about:

It follows a power law and most stories are tweeted about just 1 time.



NOTES

Comparison to Milkman's Corpus:

6,956 articles total, timespan Aug 30-Nov 30 (2008)

Subset of 2,566 manually coded