

Sophie Thesis Update 05/31 Meeting

Twitter Users Description

Number of Unique Users in Dataset:

22,540

Number of Unique Stories They've Shared:

6,911

Users that follow at least one candidate:

14,382

Users that only follow one candidate:

7,003

Number of Clinton-only followers: 2008

Total tweets: 7528

Number of Trump-only followers: 2131

Total tweets: 10177

Number of Sanders-only followers: 1728

Total tweets: 5266

Number of Cruz-only followers: 962

Total tweets: 3785

Number of Kasich-only followers: 174

Total tweets: 1400

Frequency of Tweeting: Most Authors just Tweet Once

Tweets per Author

count 22540.000000

mean 6.121473

std 40.330685

min 1.000000

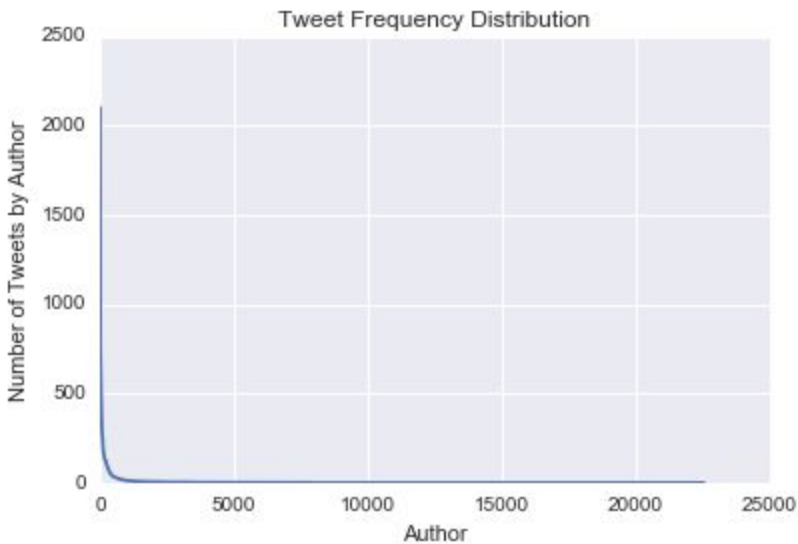
10% 1.000000

25% 1.000000

50% 1.000000

75% 2.000000

99% 120.610000
max 2103.000000



Cutoff of at least 10 Tweets for a User

Number of Authors with More than 10 Tweets

1,339

Number of Tweets by those Authors

101,055

Number of Unique Stories They've Shared

6,496

Number of Tweepers Who Follow a Candidate

650

Number of Users Who Follow a Single Candidate (AND > 10 Tweets)

258

Number of Tweets By Those Authors

16,890

Number of Unique Stories Shared by Them

4,336

Number of Clinton-only followers: 60

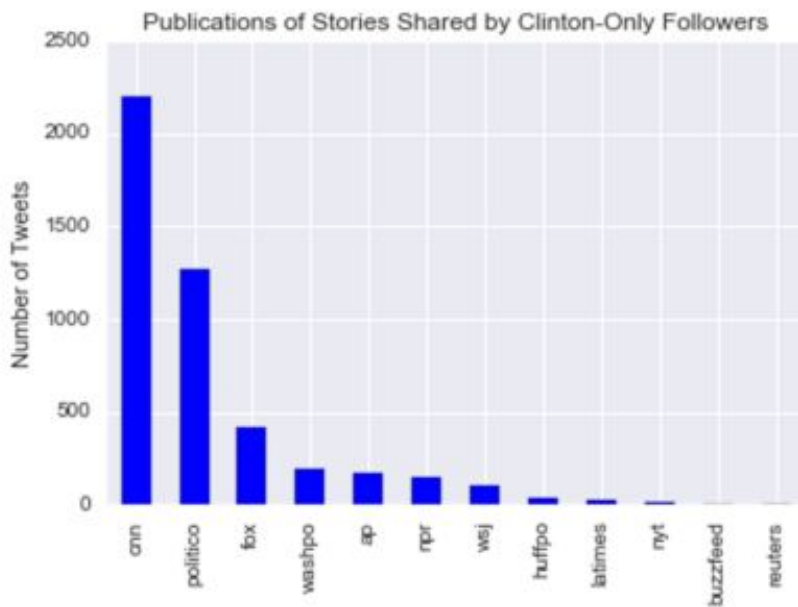
Total tweets: 4558

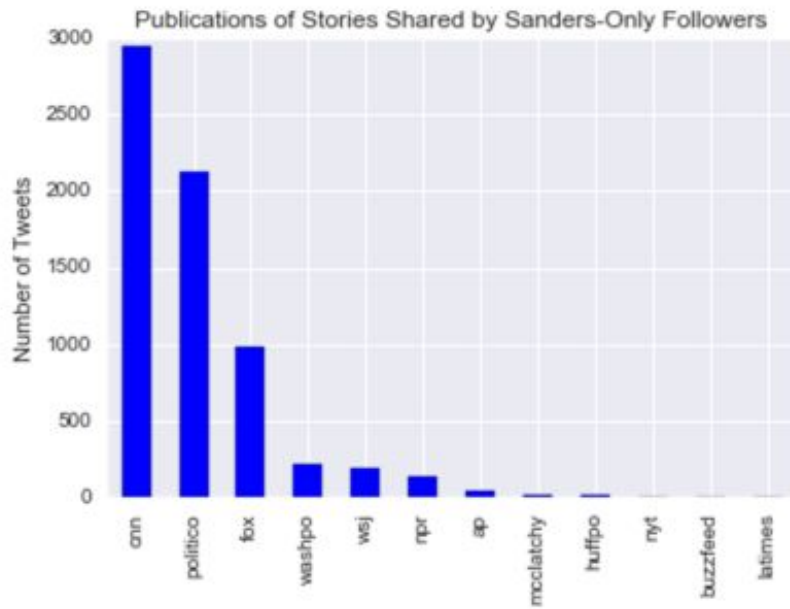
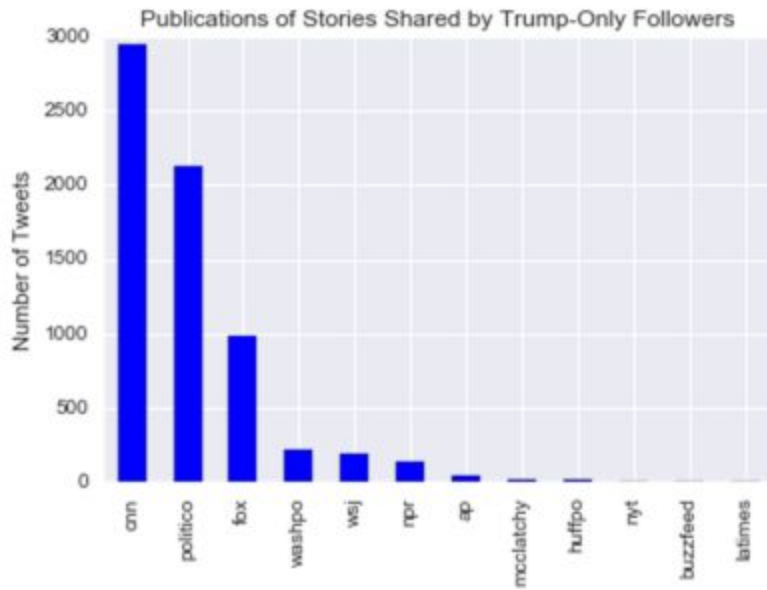
Number of Trump-only followers: 107
Total tweets: 6678

Number of Sanders-only followers: 47
Total tweets: 2465

Number of Cruz-only followers: 40
Total tweets: 2087

Number of Kasich-only followers: 4
Total tweets: 1102





Top 10 Shared Stories by Clinton-Only Followers

Judge dismisses lawsuits over Clinton's emails

29

Ben Carson just called his new plan a 'genuine flat tax.' He's wrong.

16

GOP senators: If Cruz wins, we lose	13	
White House won't commit to supporting Sanders over guns		13
Dont be fooled — Donald Trump's 'silent majority' doesn't exist		12
GOP's Trump revolt: Too little too late?	12	
Republican elders ask 'who's worse' for the GOP brand: Trump or Cruz		12
Pope vs. Trump: 'Not Christian' to only build border walls	12	
I Get Sanders' Appeal. But He's Not a Credible President.		11
Diplomats, national security officials blast Sanders on Iran normalization, ISIL		11

Top 10 Shared Stories by Trump-Only Followers

Trump basks in his spotlight	43	
FOX NEWS POLL: Trump, Cruz lead Republican field nationally		32
MEDIA BUZZ: Conservative elite, sounding a bit panicked, says Trump will destroy GOP - Trump says Cruz should resolve citizenship issue - Carson prays with Clemson coach ahead of title game - CAMPAIGN 2016 LIVE BLOG		31
Terrorists use Trump's 'Muslim ban' speech in recruitment video	30	
Why I'm voting for Trump	29	
Cruz: Hillary Clinton's backers are 'very eager to support' Donald Trump		26
Donald Trump Is Shocking, Vulgar and Right	25	
Campaigns lower expectations in Iowa, brace for Trump win		25
Ted Cruz's misleading attack that Donald Trump 'convinced' the government to 'bulldoze' a widow's home	24	
Poll: Evangelicals flocking to Trump	23	

Top 10 Shared Stories by Sanders-Only Followers

State Department asks for deadline extension on Hillary Clinton emails	70	
Bernie Sanders Will Win Iowa, New Hampshire, Nevada and South Carolina. Here's Why.		50
Hillary's Corporate Democrats Taking Down Bernie Sanders		44
Bernie Sanders Is a Once in a Lifetime Presidential Candidate. The Time Is Now.	25	
Bernie Sanders Is Now the 'Inevitable' Democratic Nominee and Presidential Winner		22
The Nation endorses Bernie Sanders	21	
Sanders cracks Clinton's Nevada firewall	18	
On January 20, 2017 Bernie Sanders Will Be Sworn In as America's 45th President		18
5 Reasons Bernie Sanders Will Win the Presidency in a Landslide		17
Donald Trump Is Shocking, Vulgar and Right	15	

User Features

Number of users with deleted/suspended accounts:

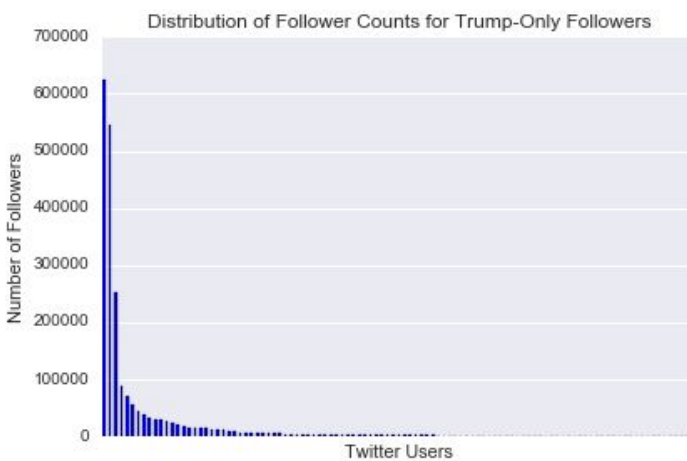
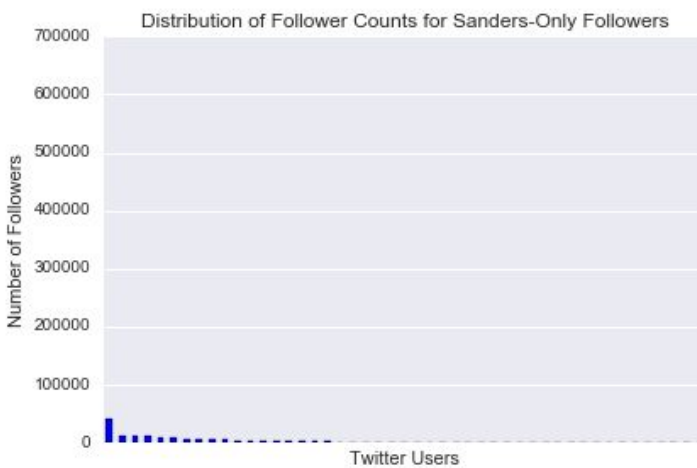
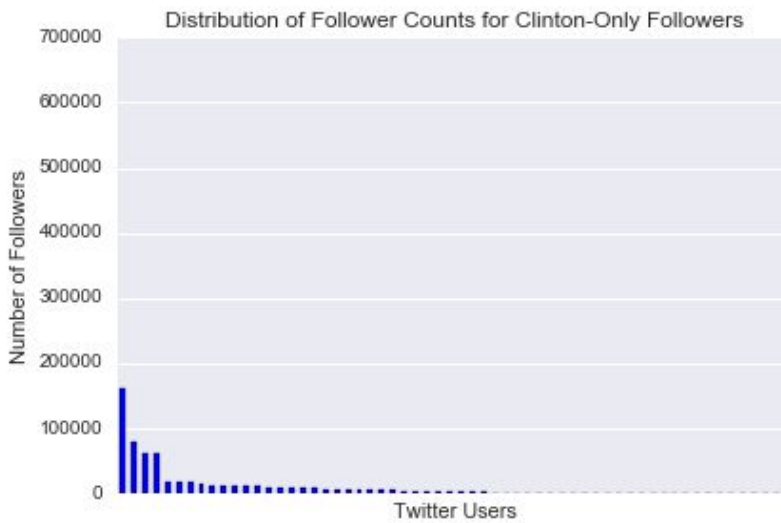
59 out of 1,339 (1282 remaining --- off by 2?)

Features we have:

- Follower count
- Friend count
- Favorites count (participation)
- Verified
- Candidates
- Sentiment of Tweet

How loud is your megaphone (Follower Count)?

For each camp of single-candidate followers:



List of all Features so Far:

Story Context

- Org
- Candidate

Story Context

- Emotional Content
- Reading Level
- Topic

User

- Credibility (verified?)
- Role (# followers : # followees -- Soroush)
- Camp (What candidate)

Y axis:

- Volume of tweets
- Controversiality (via Tweets / Soroush thesis) :
 - $(p + n) ^ \min(p,n)$

Features currently up in the air:

- Gender: Prashanth's model not yet ready
- Replies: talked to Martin + Soroush (impossible with API, potentially possible by crawling page -- talking to Prashanth)

NOTES

Comparison to Milkman's Corpus:

6,956 articles total, timespan Aug 30-Nov 30 (2008) [**Us**: 6,496]

Subset of 2,566 manually coded

Quality Checks:

- I sampled 100 random unmatched articles to look for any clear patterns-- do not see anything sticking out (I have attached those stories in a CSV).
- Will include caveats of this pipeline in methodology section.

Questions:

- Should we include Kasich followers?
- Should we look at those who follow more than 1 candidate?