



reddit

An Analysis of the Political Landscape of Reddit

Project 1: July 23, 2019

Jarred Brouhard • Scott Fleishman • Jessica Lai • Hakeem Salami



Main Question

What is the political landscape of Reddit?

Additional Questions

- *Does Reddit lean in a particular direction in politics?*
- *Are left-leaning or right-leaning Redditors more politically engaged on Reddit?*
- *Are we seeing one side influence general political subreddits?*
- *How do the subreddits of Democratic presidential candidates compare to current polling data?*



What Is Reddit?

According to Wikipedia:

“Reddit is an American social news aggregation, web content rating, and discussion website. Registered members submit content to the site such as links, text posts, and images, which are then voted up or down by other members.”

Reddit is made up of many subreddits, each focusing on a specific topic.



Glossary | Terms to Know

Subreddit : *A forum on Reddit focusing on a specific topic*

Post : *A submission to Reddit*

Upvote Ratio : *The percentage of upvotes from all votes on the submission*

Score : *Number of upvotes minus downvotes for the submission*

Comments : *Comments made on the post*

Author : *User who created post*

Controversial : *Posts that are close to a 50% upvote / downvote ratio*



Part 1

Exploring the Engagement of Each Political Party's Reddit Activity

Overview:

- Focus on right-leaning & left-leaning
- Determine 5 subreddits to analyze for each side
- Pull top 100 posts (by score) for each
- Create bubble visualizations summarizing each subreddit's engagement level by Score, Upvote Ratio and Comments



First, How Do We Determine Top 5 Subreddits for Each Side?

Steps Taken:

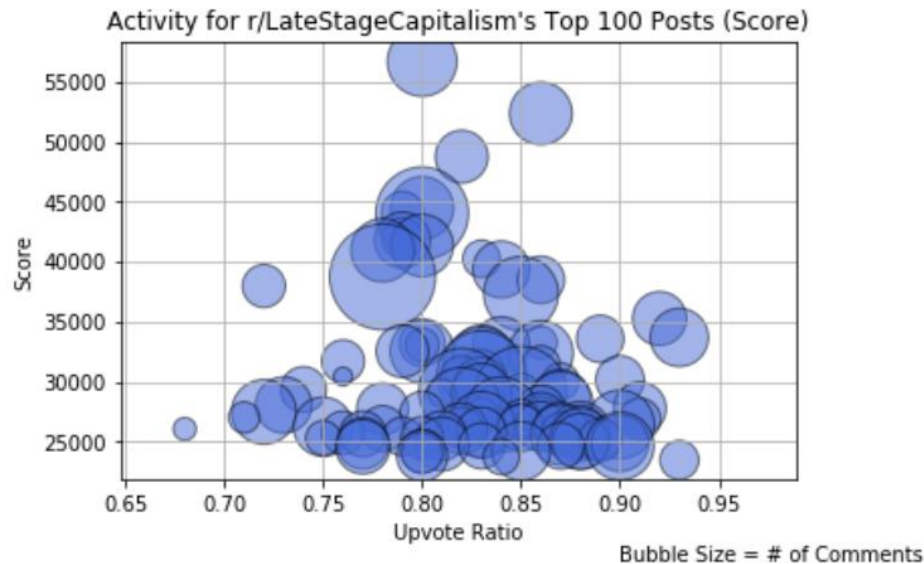
- Pull list of subreddits used by Reddit bot “userleansbot”
- Use pandas to pull subscriber count for each subreddit
- Sort by subscriber count and determine the top 5 subreddits for each party
- Supplemental research to reach final decision



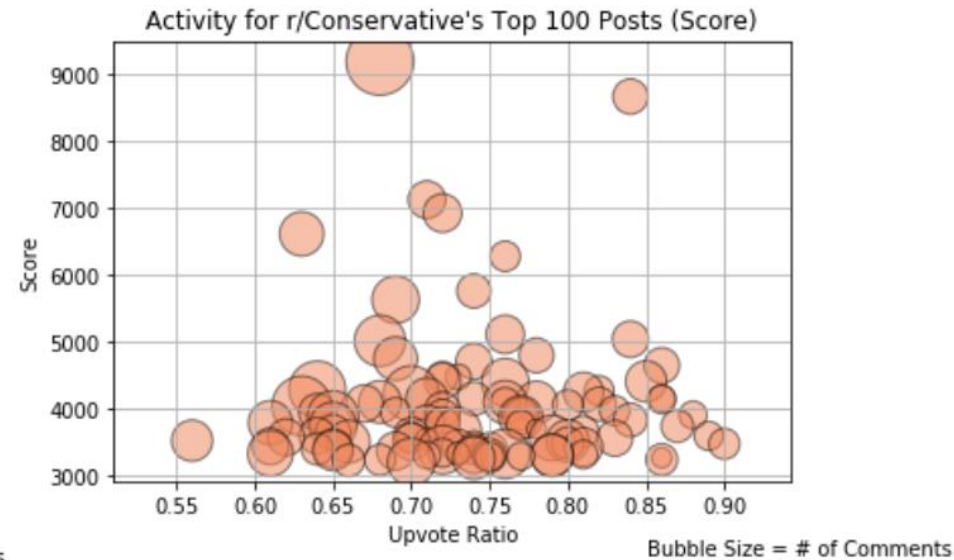
reddit

#1 Subreddit for Each Side (Subscribers)

r/LateStageCapitalism (463k) & r/Conservative (235k)



	Score	Upvote Ratio	Comments
mean	30170.5	0.8273	720.17

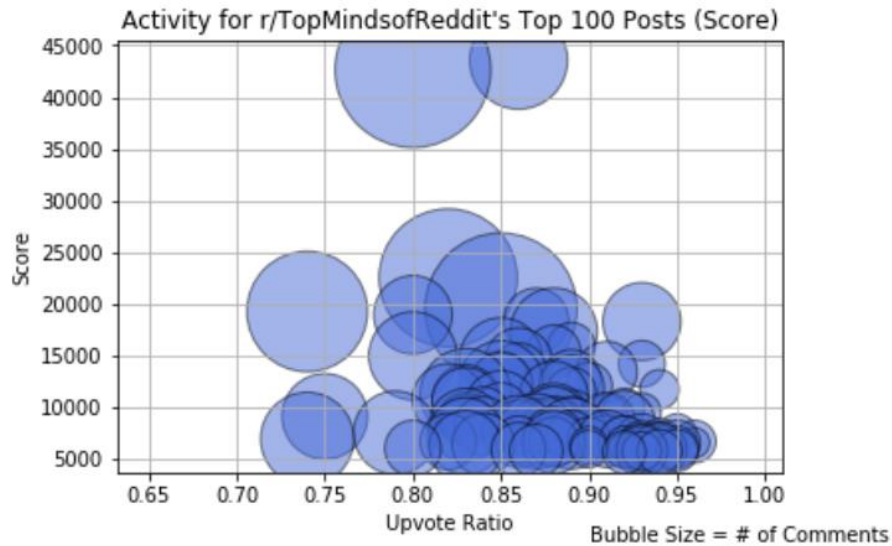


	Score	Upvote Ratio	Comments
mean	4054.78	0.7398	349.68

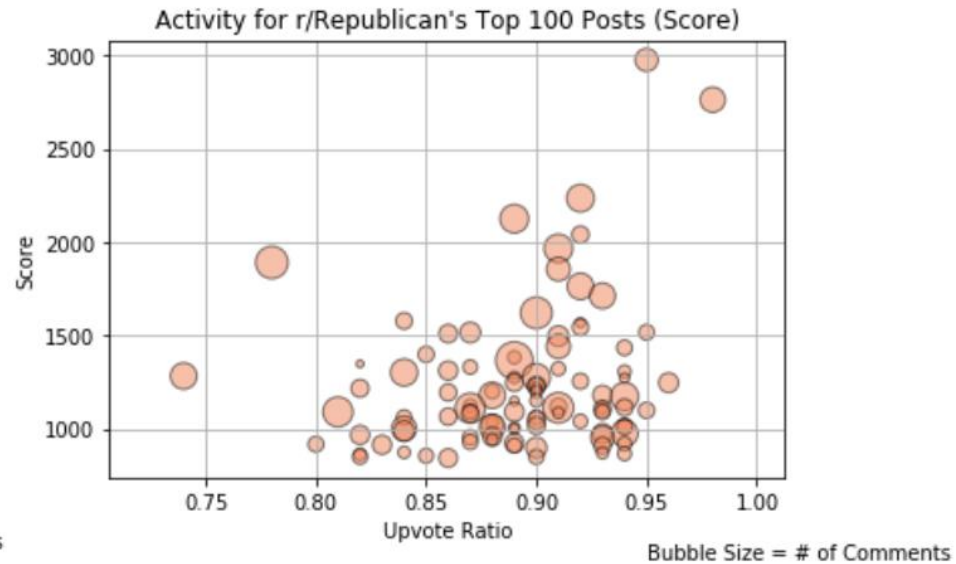


#2 Subreddit for Each Side (Subscribers)

r/TopMindsofReddit (207k) & r/Republican (86.8k)



	Score	Upvote Ratio	Comments
mean	10033.08	0.8799	1048.36

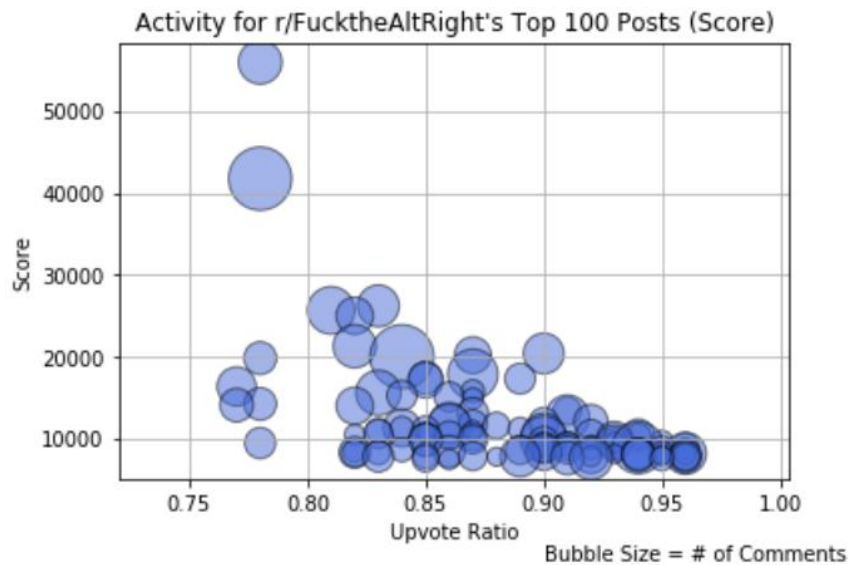


	Score	Upvote Ratio	Comments
mean	1237.51	0.8909	92.96

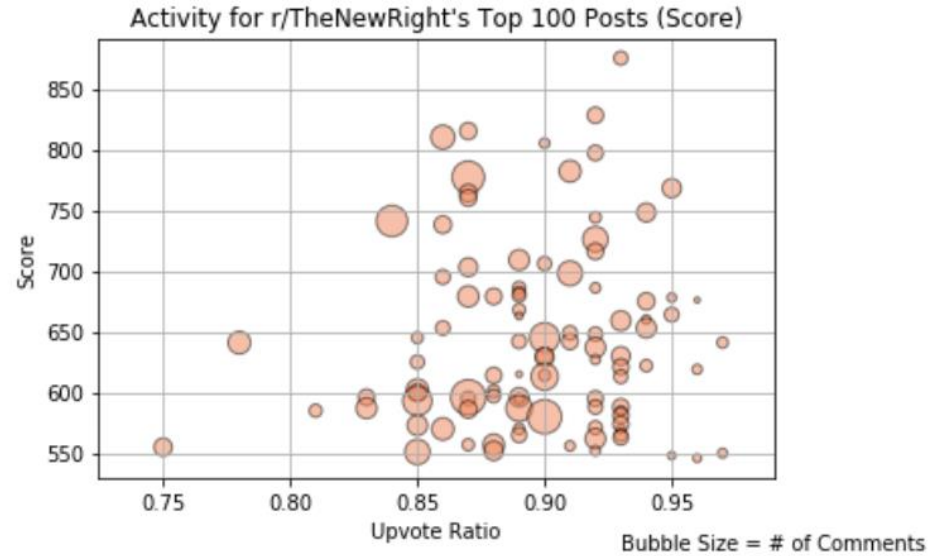


#3 Subreddit For Each Side (Subscribers)

r/FucktheAltRight (122k) & r/TheNewRight (30.6k)



	Score	Upvote Ratio	Comments
mean	12032.59	0.8795	265.16

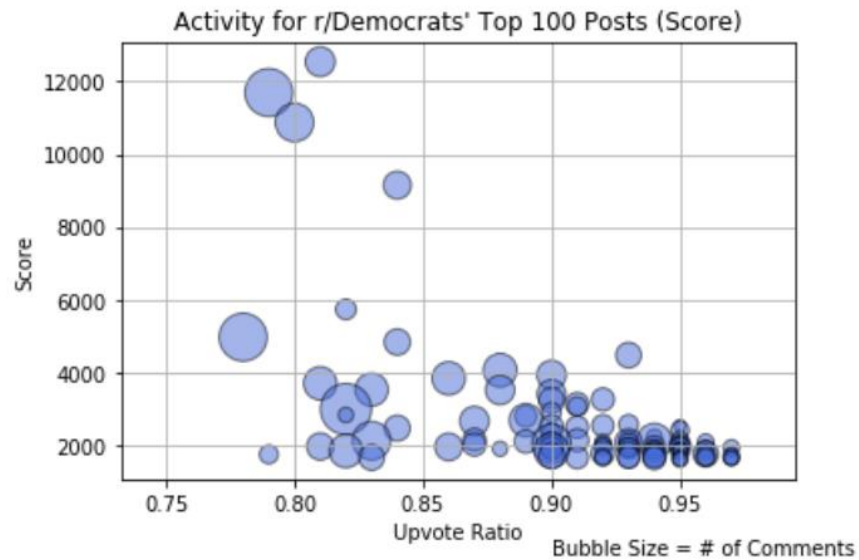


	Score	Upvote Ratio	Comments
mean	643.48	0.8975	76.47

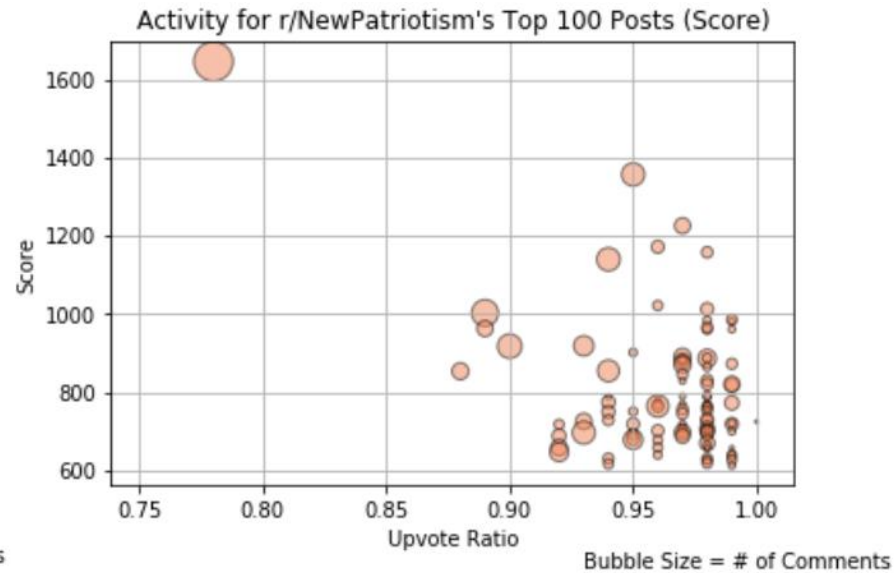


#4 Subreddit For Each Side (Subscribers)

r/Democrats (99.3k) & r/NewPatriotism (22k)



	Score	Upvote Ratio	Comments
mean	2677.33	0.9075	142.28

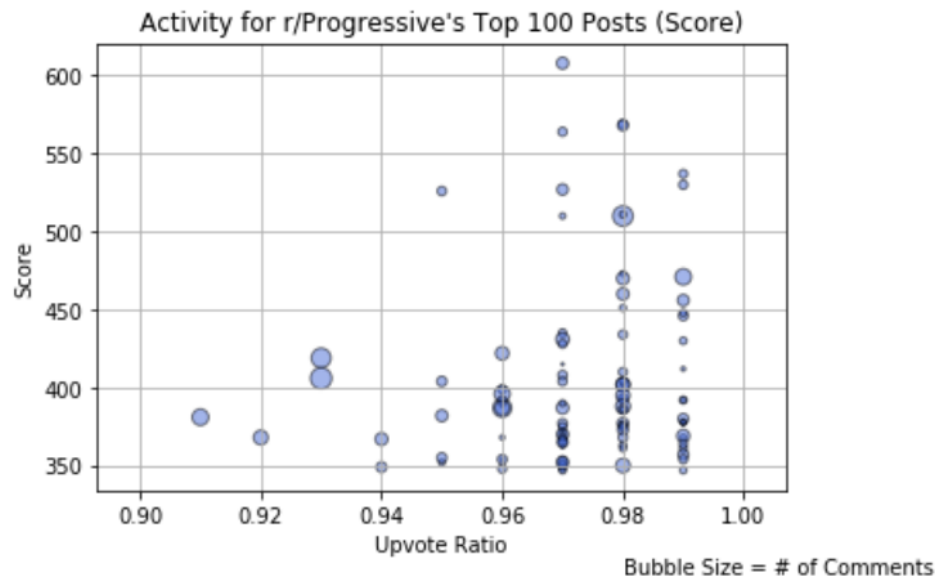


	Score	Upvote Ratio	Comments
mean	799.89	0.9633	42.57

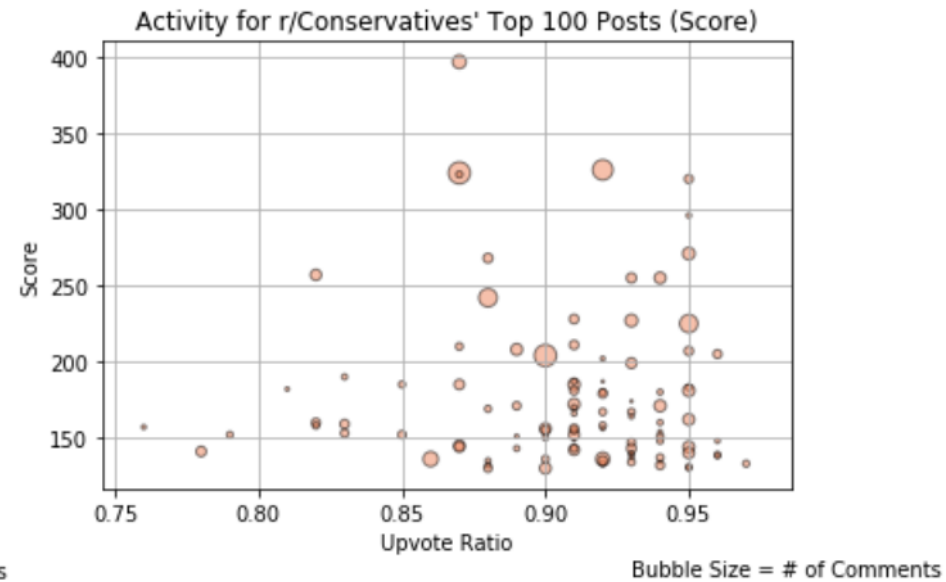


#5 Subreddit for Each Side (Subscribers)

r/Progressive (60.3k) & r/Conservatives (15.4k)



	Score	Upvote Ratio	Comments
mean	404.92	0.9725	24.73



	Score	Upvote Ratio	Comments
mean	177.46	0.9061	20.87



What Did We Discover About Engagement in Left-Leaning vs. Right-Leaning Subreddits?

Based on the posts, left-leaning Redditors were more active (scores & comments):

	Avg. Score	Avg. Upvote Ratio	Avg. Comments
Left-Leaning	11064.45	0.89	440.14
Right-Leaning	1382.62	0.88	116.51

	Subreddit	Score	Upvote Ratio	Comments		Subreddit	Score	Upvote Ratio	Comments
2	r/TopMindsofReddit	10034.47	0.8799	1048.38	0	r/Conservative	4054.78	0.7398	349.68
0	r/FucktheAltRight	12034.21	0.8795	265.16	1	r/Republican	1237.51	0.8909	92.96
1	r/Democrats	2677.89	0.9076	142.3	2	r/TheNewRight	643.48	0.8975	76.47
3	r/LateStageCapitalism	30170.9	0.8273	720.15	3	r/NewPatriotism	799.89	0.9633	42.57
4	r/Progressive	404.77	0.9725	24.73	4	r/Conservatives	177.46	0.9061	20.87



Part 2

Knowing Their Engagement, Are Left-Leaning and Right-Leaning Redditors Also Posting in “Neutral” Subreddits?

Overview:

- Identify top neutral subreddits
- Pull 1000 “Top” authors per subreddit in the last year
- Pull 1000 “Controversial” authors per subreddit in the last year
- Create venn diagrams to see overlap



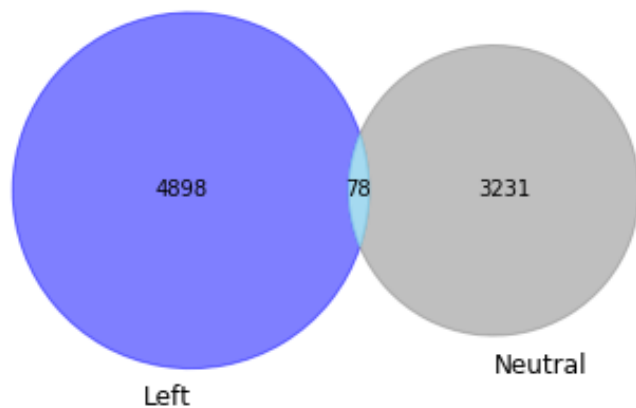
Why Did We Pull “Top” and “Controversial”?

Posts created by authors can be classified as “Top”, “Gilded”, “Hot”, “New” or “Controversial”. We noticed a small number of authors who often post on their specific party subreddits, get classified as “Controversial” (posts receiving equal upvotes and downvotes with a lot of activity) in subreddits outside of their political party.

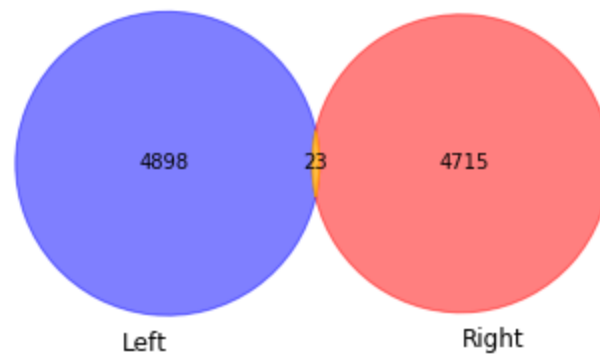


Where are Authors Posting?

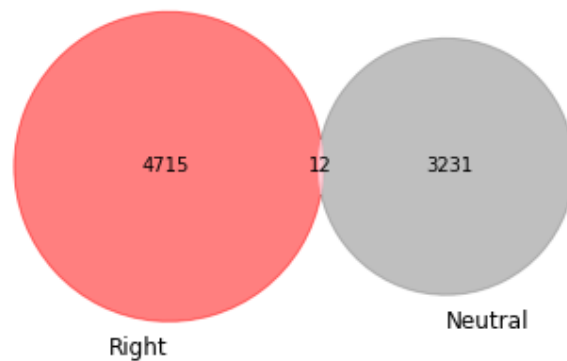
Shared Authors Between Subreddits



Shared Authors Between Subreddits



Shared Authors Between Subreddits





What Did We Discover?



Part 3

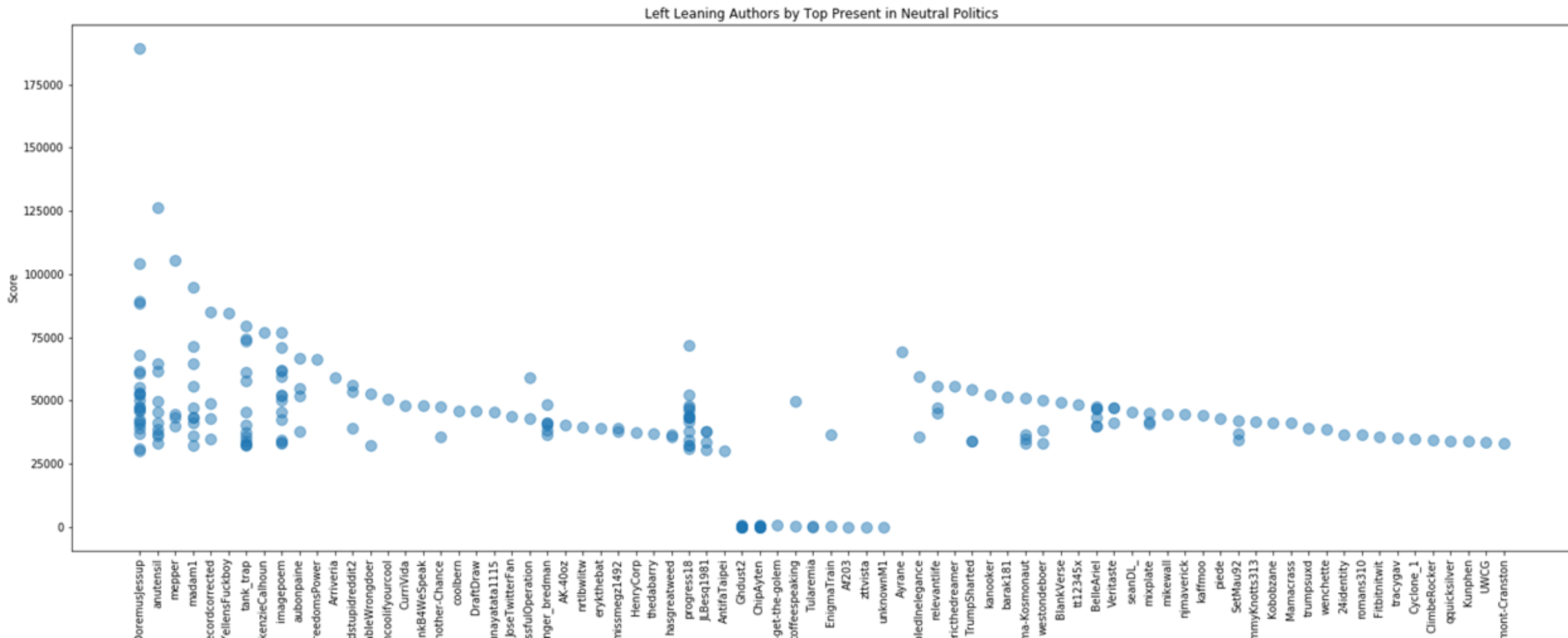
Knowing Part 2, How Are Left-Leaning and Right-Leaning Redditors' Posts Performing in “Neutral” Subreddits?

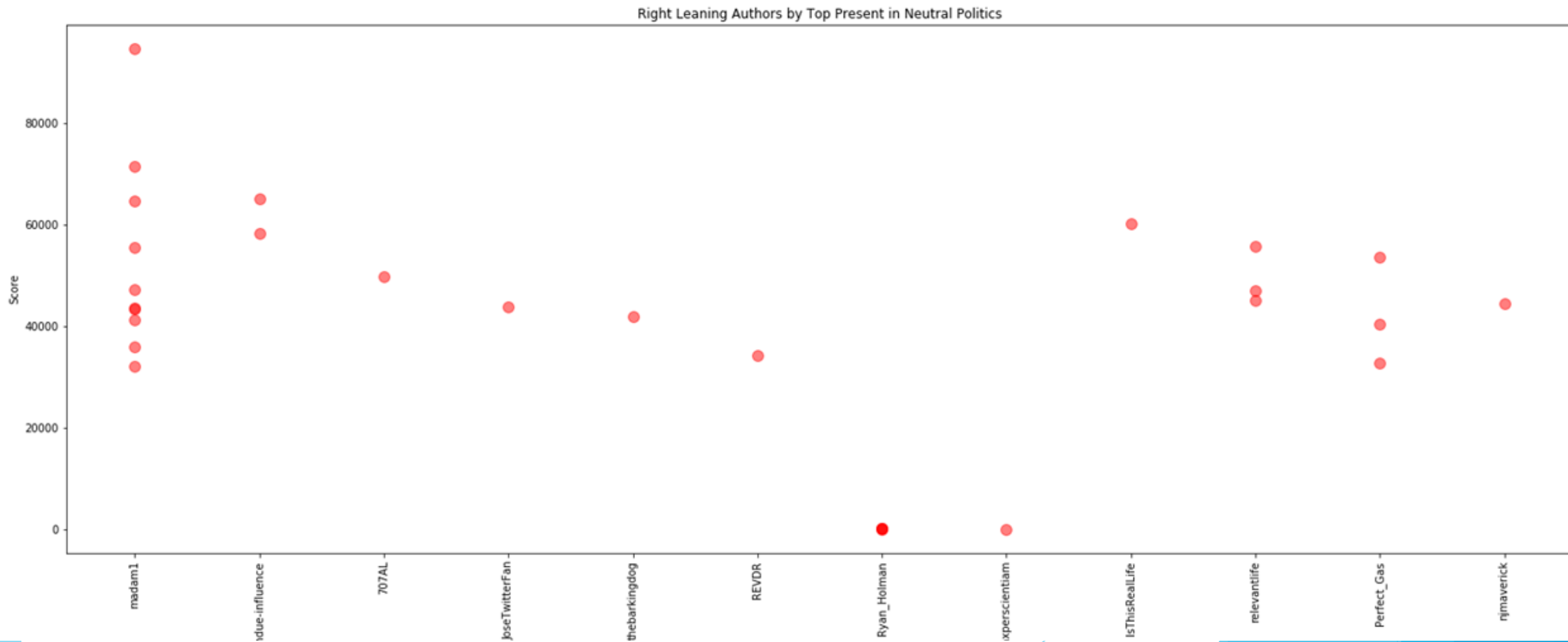
Overview:

- Create bubble scatters to see performance



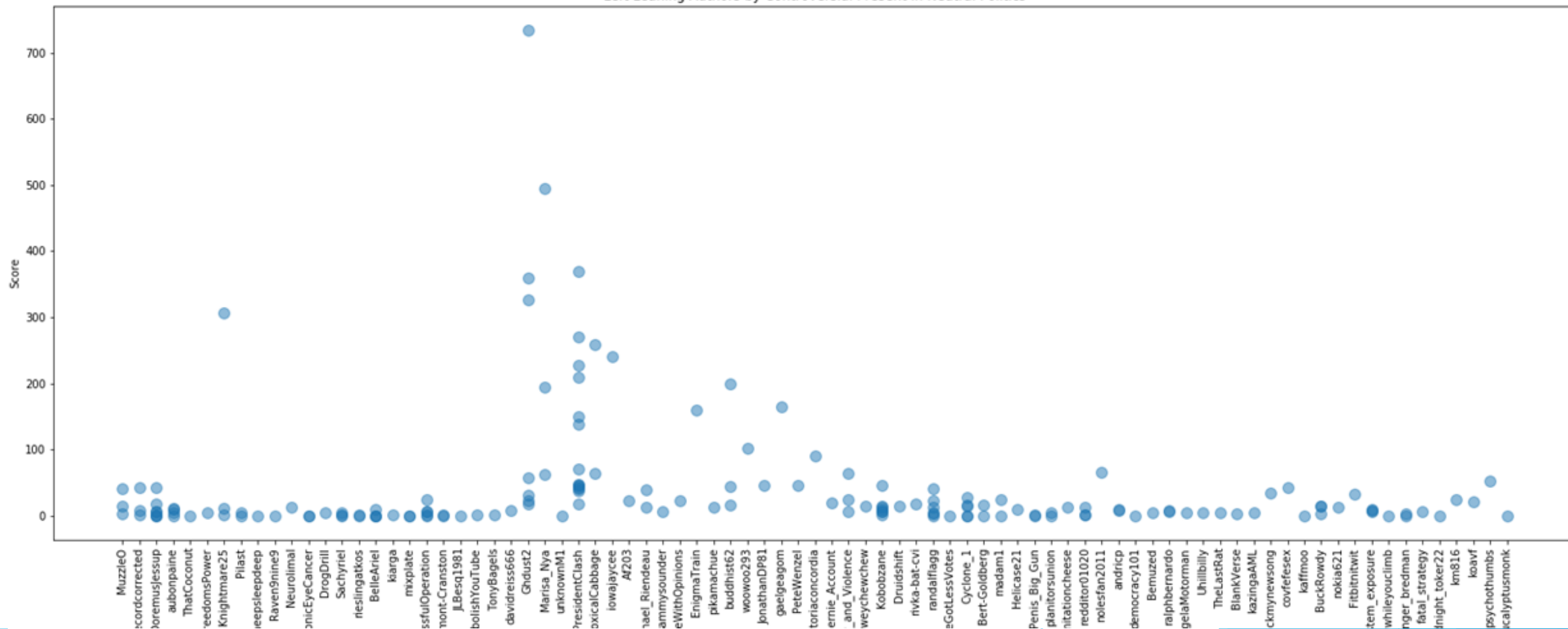
reddit





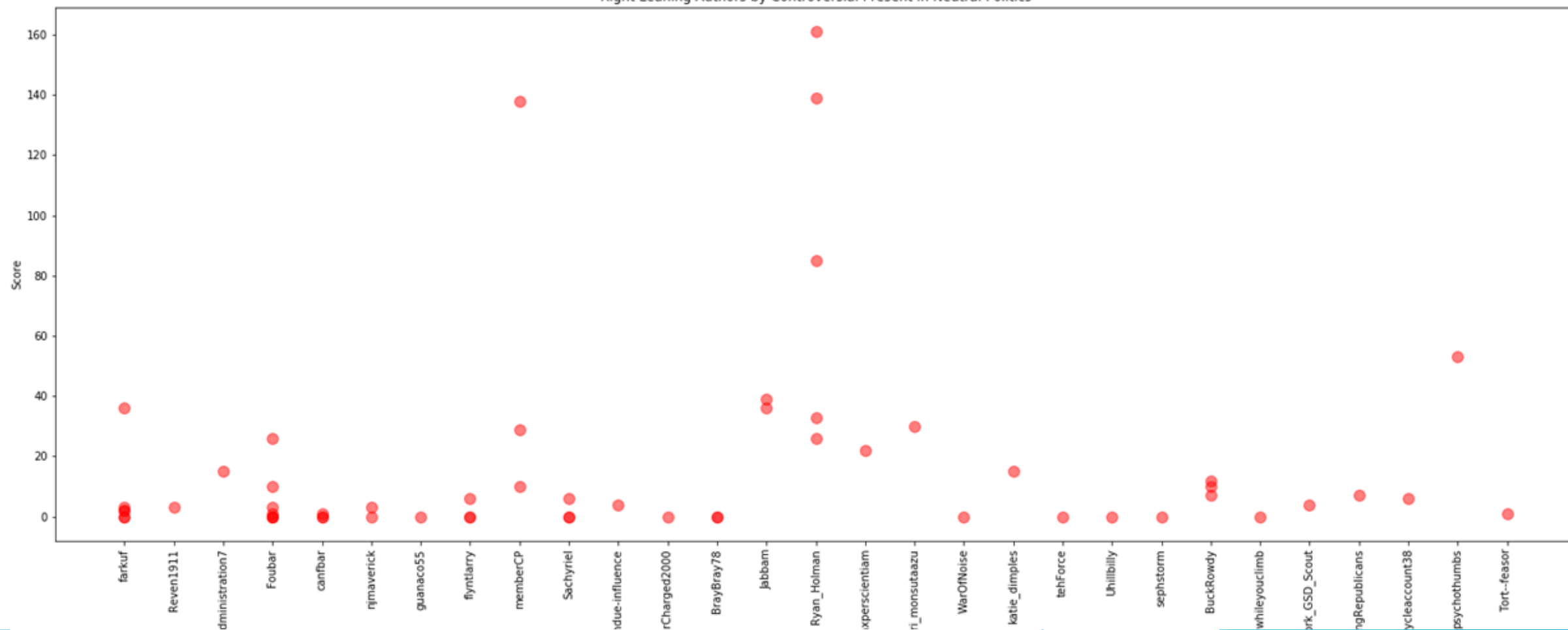


Left Leaning Authors by Controversial Present in Neutral Politics





Right Leaning Authors by Controversial Present in Neutral Politics





What Did We Discover?



Part 4

Now That We Conclude Reddit Leans Left, How Do the Subreddits of Democratic Presidential Candidates Compare to Current Polling Data?

Overview:

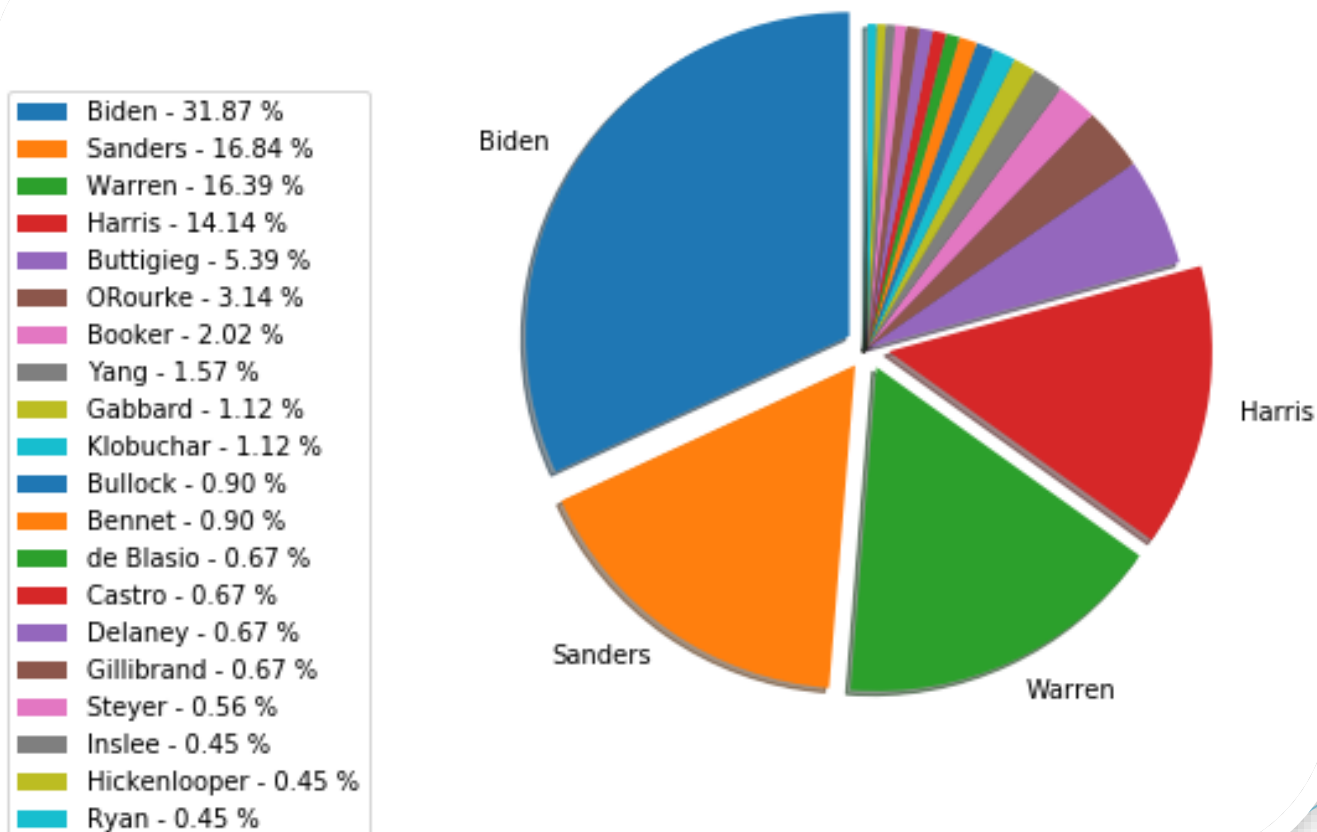
- Find current polling data for top candidates
- Research to find the top Subreddit for each of the top 20 candidates
- Use pie charts to illustrate and compare percentages



reddit

Current Polling Data

RealClear Politics Polling Data (7/19/19)

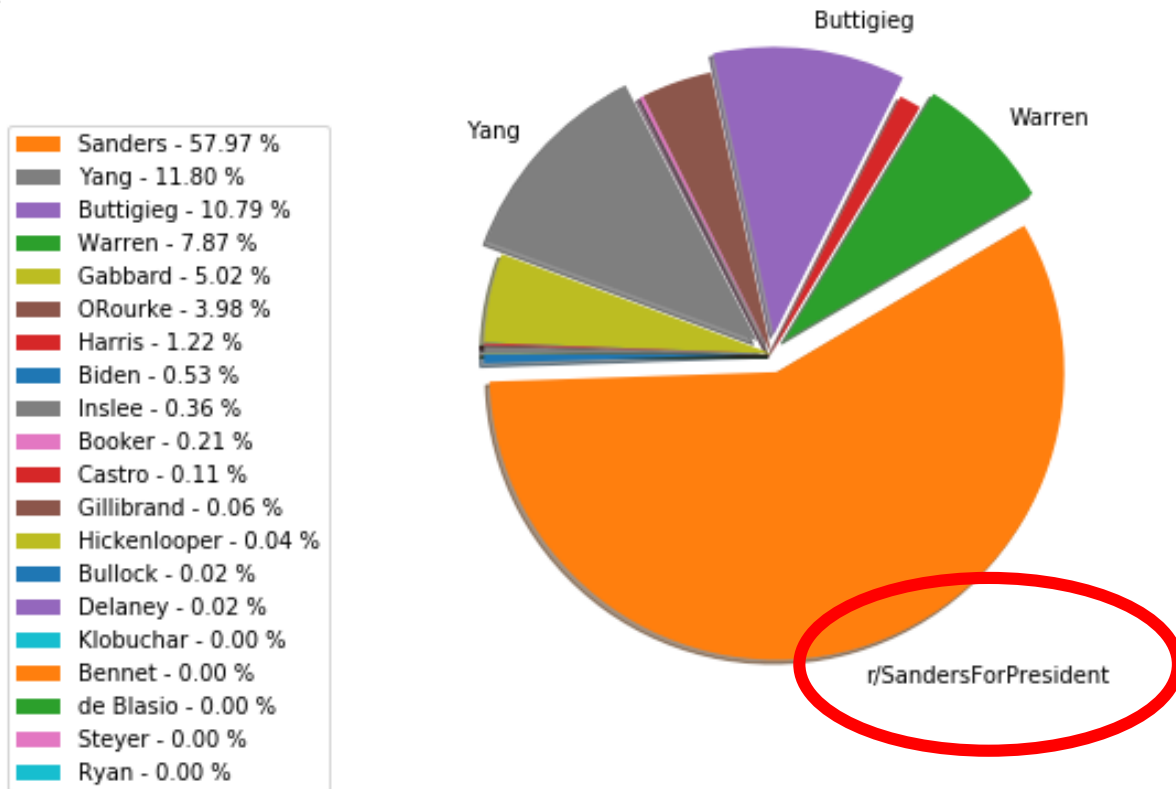




reddit

Candidates' Subreddit by Subscriber Count

Each Candidate's Subreddit Subscriber Count Percentage (r/SandersForPresident)



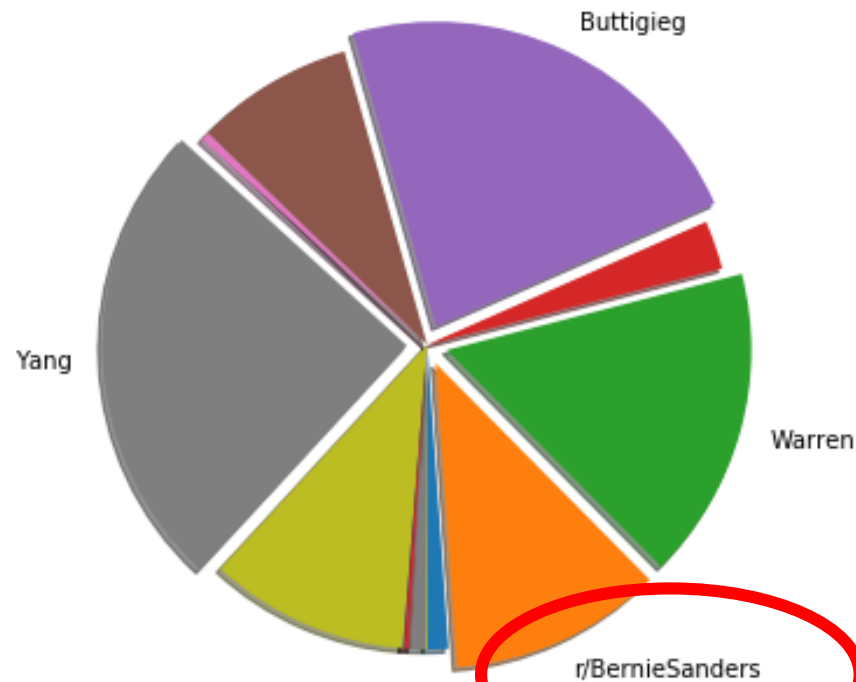


reddit

Candidates' Subreddit by Subscriber Count

Each Candidate's Subreddit Subscriber Count Percentage (r/BernieSanders)

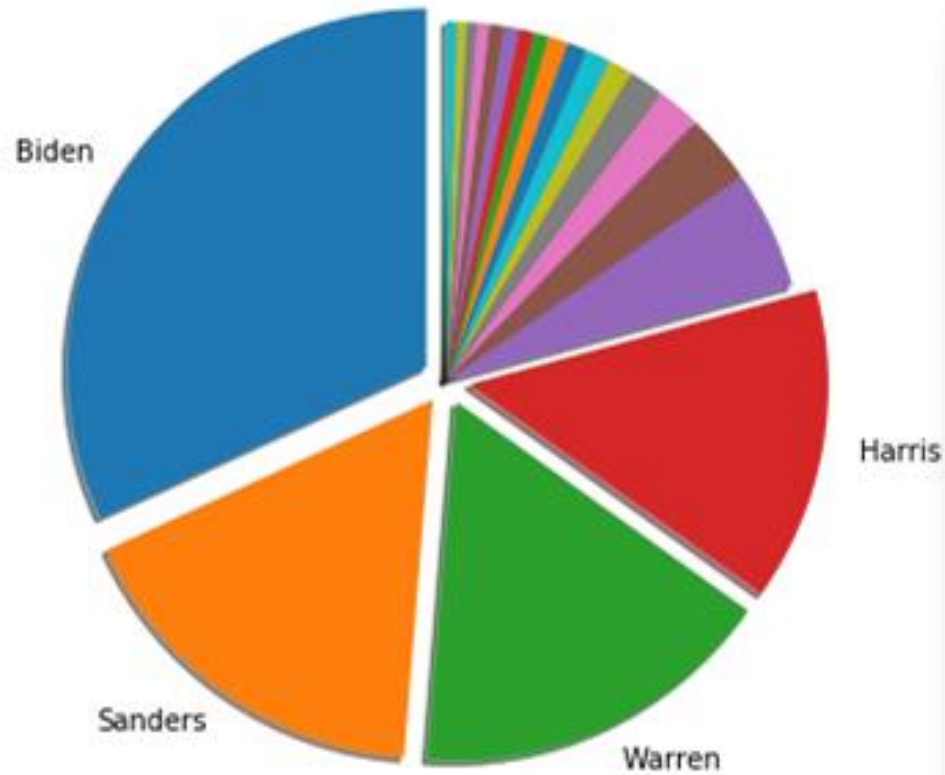
Yang	- 24.91 %
Buttigieg	- 22.78 %
Warren	- 16.61 %
Sanders	- 11.30 %
Gabbard	- 10.60 %
ORourke	- 8.41 %
Harris	- 2.58 %
Biden	- 1.12 %
Inslee	- 0.75 %
Booker	- 0.44 %
Castro	- 0.22 %
Gillibrand	- 0.13 %
Hickenlooper	- 0.08 %
Bullock	- 0.05 %
Delaney	- 0.03 %
Klobuchar	- 0.00 %
Bennet	- 0.00 %
de Blasio	- 0.00 %
Steyer	- 0.00 %
Ryan	- 0.00 %



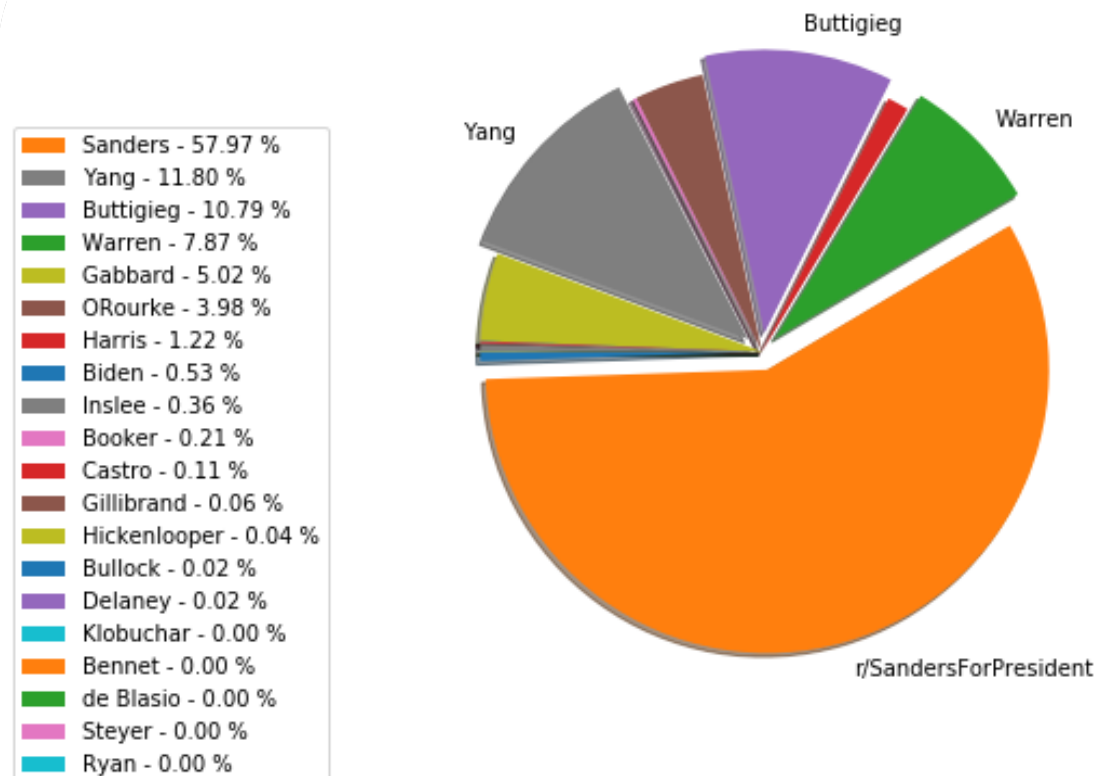


reddit

Comparison (Polling Data vs. Subreddit Subscriber Counts)



Each Candidate's Subreddit Subscriber Count Percentage (r/SandersForPresident)





What Did We Discover?



Things to Consider:

- Banned subreddits weren't included
Examples: r/The_Donald with its 771k subscribers (this was later discovered to be unbanned and was moved to quarantine) and r/Physical_Removal to name a couple
- Excluded from results are any posts classified as “Gilded”, “Hot”, or “New”
- We had to drop multiple “None” Authors (posts were deleted/banned/removed)
- Our limitations were based on what Praw was able to pull (one year of data and only 1000 posts)
- Smaller subreddits didn't have enough for us to pull 1000 posts



Our Main Question

What is the political landscape of Reddit?

Final Conclusion

Reddit is a left-leaning platform with a tendency to lean to the far left. There is more favorable overlapping activity amongst subscribers and authors in left-leaning subreddits and “neutral” subreddits.