

Gender and Religion in Congressional Twitter Bios

Kenan Anderson and Rodrigo Quijano-Luna

DS 301: Advanced Topics in Data Science: NLP and Reinforcement Learning

## Table of Contents

Abstract .....	3
Introduction .....	4
Methods .....	5
Results .....	6
Discussion .....	7

### Abstract

Politicians and voters increasingly use Twitter to communicate and engage in political discourse. In this report we investigate what words and phrases politicians use to describe themselves in their Twitter bio. Specifically, we investigate the frequency of both a politician's preferred gender pronouns as well as religious references and look to see if a number of variables such as party, partisan lean of constituency, age, and number of followers. Looking at every member of Congress we find that while Democrats are more likely to include their preferred gender pronouns (PGP) than Republicans, and less likely than Republicans to include religious references, both are relatively uncommon among members of Congress and are not predicted by other variables.

### Author Responsibilities

Kenan was responsible for gathering and cleaning all of the data before the API process, as well as creating the binary variables for religion and pronouns. Kenan also wrote the majority of the data blitz presentation and this report. Rodrigo used the Twitter API to gather the actual text of the bios as well as performing the basic data cleaning process (removing punctuation, capitalization, etc.). Rodrigo also narrated the data blitz and conducted the larger analysis using linear regression and created the charts featured in the report. He also submitted suggestions and edits to the report after reading the first draft.

## Introduction

Politicians use Twitter more and more frequently, but the nature of the platform means you can often see tweets from a user who you are unfamiliar with or do not follow. In these situations, it is often a first reaction to click on that user and read how they describe themselves in their bio. This can make the Twitter bio a first or second impression for a politician to make on a voter who stumbles across their tweet. So, how are politicians describing or differentiating themselves in their descriptions? The existing literature on Twitter and politicians is extensive, including papers that analyze how political activity on Twitter clusters along partisan lines<sup>1</sup> as well as whether or not legislation drives social media or vice versa.<sup>2</sup> However, little has been done specifically analyzing the use of the Twitter bio, with work either being largely anecdotal<sup>3</sup> or focusing on other professions such as journalism.<sup>4</sup> Given the short length of Twitter bios, we chose to focus on two short sets of words to look for. First, we would look for politicians who indicated their preferred gender pronouns (or PGP) such as writing “(she/her)” or “(he/him)”. And second, we would create a glossary of religious terms and search for them. Our hypothesis was that Democratic politicians would be more likely to have PGP in their Twitter bio, while Republicans would be more likely to have religious mentions in theirs. Additionally, we thought these effects would be concentrated among members who represented “safe” constituencies.

---

<sup>1</sup> Barberá, P. (2015). Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation Using Twitter Data. *Political Analysis*, 23(1), 76-91. doi:10.1093/pan/mpu011

<sup>2</sup> BARBERÁ, P., CASAS, A., NAGLER, J., EGAN, P., BONNEAU, R., JOST, J., & TUCKER, J. (2019). Who Leads? Who Follows? Measuring Issue Attention and Agenda Setting by Legislators and the Mass Public Using Social Media Data. *American Political Science Review*, 113(4), 883-901. doi:10.1017/S0003055419000352

<sup>3</sup> D, Romy. Analyzing 7 World Leaders' Twitter Bios. 14 Sept. 2016, blog.banzee.net/en/analyzing-7-world-leaders-twitter-bios/

<sup>4</sup> Ottovordemgentschenfelde S. ‘Organizational, professional, personal’: An exploratory study of political journalists and their hybrid brand on Twitter. *Journalism*. 2017;18(1):64-80. doi:10.1177/1464884916657524

## Methods

We chose to limit our analysis to current members of Congress to take advantage of existing databases while giving us a large sample size of over 500 members. We began by merging data gathered from the *@unitedstates* project<sup>5</sup> which contained basic biographical information with data from *FiveThirtyEight.com*<sup>6</sup> on the partisan lean of every state and/or district in Congress. While the *@unitedstates* data included Twitter handles of a politician's official congressional office, we elected to gather and use their personal/campaign Twitter accounts instead, as in our estimation those tended to be more visible and faced no legal restriction the way official accounts did. We then used a Twitter API to gather each biography, basing our code on examples from this repository.<sup>7</sup> The methods we implemented include `create_csv()` and `get_userinfo()`, however, modifications were made to the originals since we were only interested in gathering the text of the bio and the follower count.

After cleaning and standardizing the bio it was added to the main dataset. We then created a function to look for uses of a pronoun but required at least two pronouns to appear in the bio in order to avoid false positives such as Senator Mitch McConnell's bio (which reads "For more information visit *his* website") and manually checked the remaining ones. We then created a list of religious terms<sup>8</sup> and checked whether any were present in each bio. These were turned into binary variables "has\_pronouns" and "has\_religion" respectively. The final results were 21

---

<sup>5</sup> <https://github.com/unitedstates/congress-legislators>

<sup>6</sup> <https://github.com/fivethirtyeight/data/tree/master/partisan-lean>

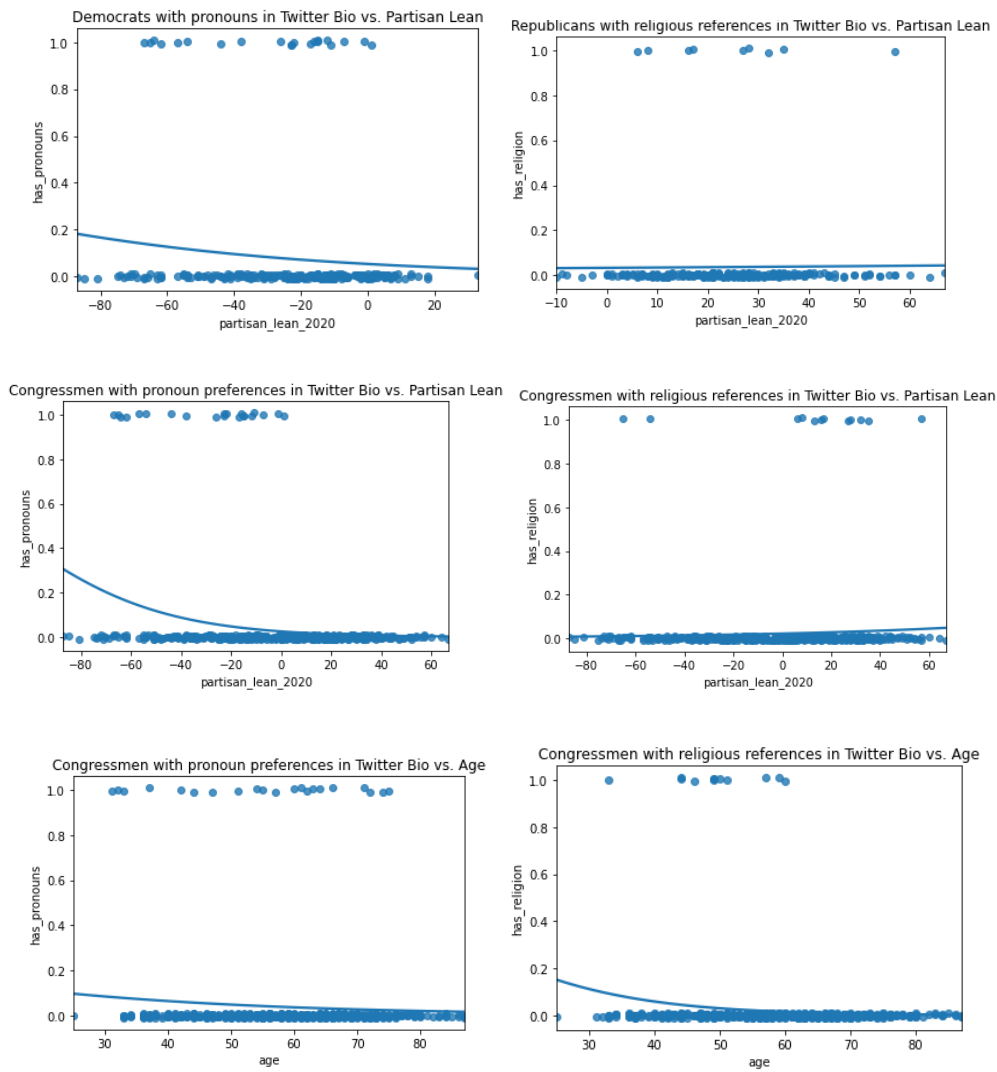
<sup>7</sup> <https://gist.github.com/macloo/5c69cdf5294fa97eb41d6ad950233cee>

<sup>8</sup> Religious Terms used: [christ, christian, jesus, god, holy spirit, bible, muslim, muslima, allah, islam, koran, quran, qur, imam, hindu, sikh, jew, jewish, Judaism, rabbi, preacher, pastor, chaplain, minister, priest]

Democrats and 0 Republicans with preferred gender pronouns and 9 Republicans and 3 Democrats with religious references. We then conducted several linear regressions to see if partisan lean predicted either the presence of pronouns or religious terms.

### Results

Overall, the results found that other than party, there was little predictive power of either variable. Linear regressions found that partisan lean, age, and number of followers all failed to effectively predict whether a politician mentioned either gender pronouns or religious terms. These results held even when the regression was limited to a single party.



## Discussion

Our conclusions were mixed in regard to our hypotheses. On the one hand, the core of our hypothesis was true. While nearly 10% of Democrats mention their pronouns in their Twitter bio, not a single Republican does. The gap is less stark on religious mentions, but there are still three times as many Republicans who mention religion as Democrats. Unfortunately, our belief that this gap would follow the partisan lean of districts was untrue. However, what was most confounding is how both of these variables were lacking from the overwhelming majority of politicians. Most members of Congress mention neither religion nor their gender pronouns in their bio. The latter could perhaps be explained by the frequency of other words which indicate gender identity (such as father, dad, wife, etc.) which were very common in the bios we analyzed. Finally, our assumption that these would be more common is perhaps evidence of the phenomenon for voters to overestimate the size of a demographic group in both coalitions.<sup>9</sup> Our research was limited to only members of Congress, and a more comprehensive search that included state legislatures, governors, as well as potentially political campaign staff or activists might find different conclusions.

---

<sup>9</sup> Ahler Assistant Professor of Political Science, Douglas J. *Something Democrats and Republicans Have in Common: Exaggerated Stereotypes about Both Parties*. 19 Jan. 2021, [theconversation.com/something-democrats-and-republicans-have-in-common-exaggerated-stereotypes-about-both-parties-130848](https://theconversation.com/something-democrats-and-republicans-have-in-common-exaggerated-stereotypes-about-both-parties-130848).