

Abstract

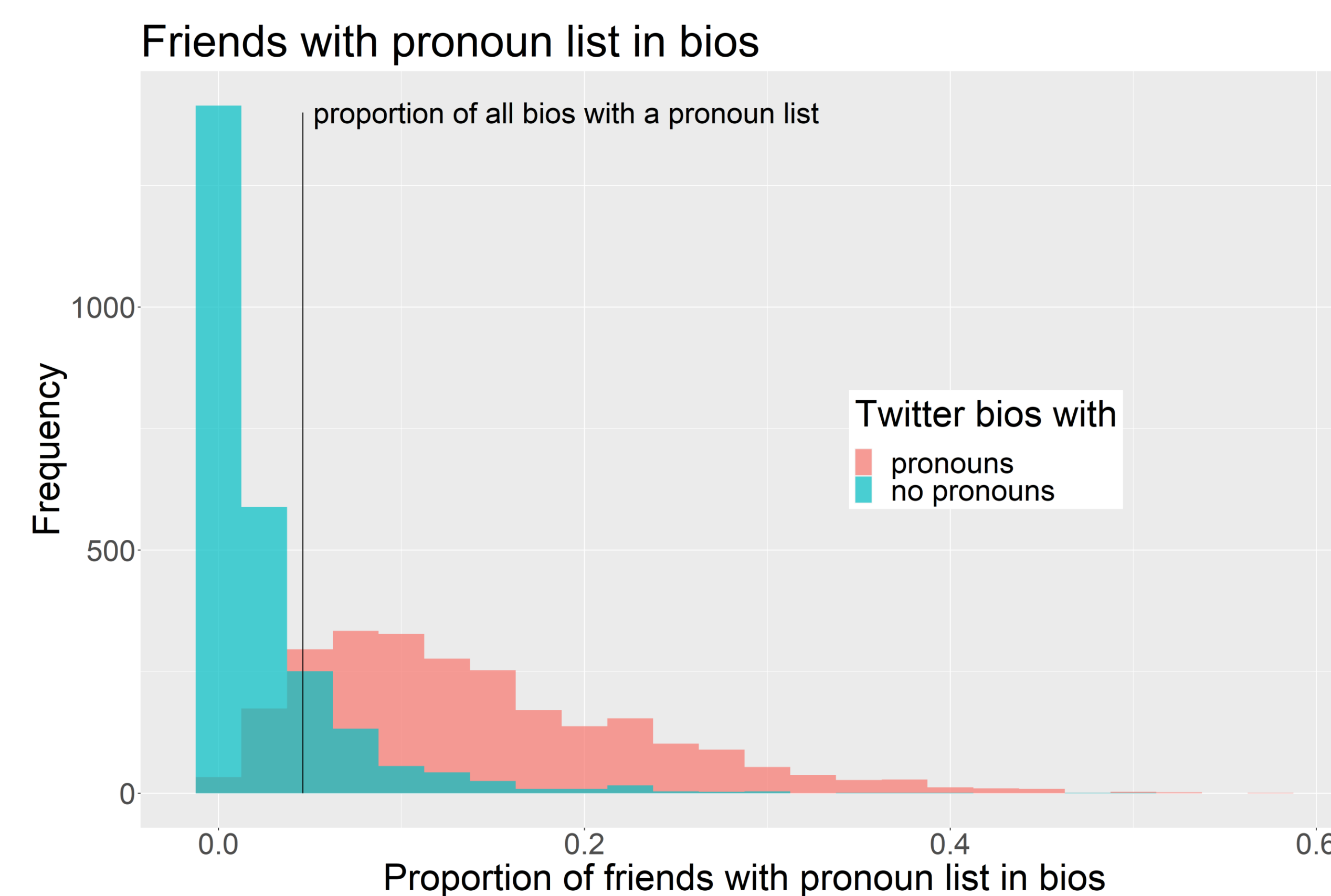
There has been a large increase in the number of people publicly sharing their preferred pronouns in the past few years. This research explores which people comprise the **0.0461**, or **4.61%**, of people with a pronoun list in their Twitter bio in 2022. We find left-wing political tokens appear disproportionately often alongside pronoun lists, while religious and financial terms appear disproportionately infrequently. Additionally, we showed that users with pronoun lists are more strongly connected than random users and predicted which users would add pronouns to their bios with 70% accuracy with a Multinomial Naïve Bayes Model.

Definitions

Word	Definition
Pronoun List	A list specifying one's preferred pronouns (he/him, they/them, etc.)
Token	A keyword that appears alongside a pronoun list
Prevalence	The number of times a token appears per 10,000 Twitter bios
Relative prevalence	The ratio of a token's prevalence in all bios to its prevalence in bios with a certain other token
Friend	A Twitter friend is someone you follow

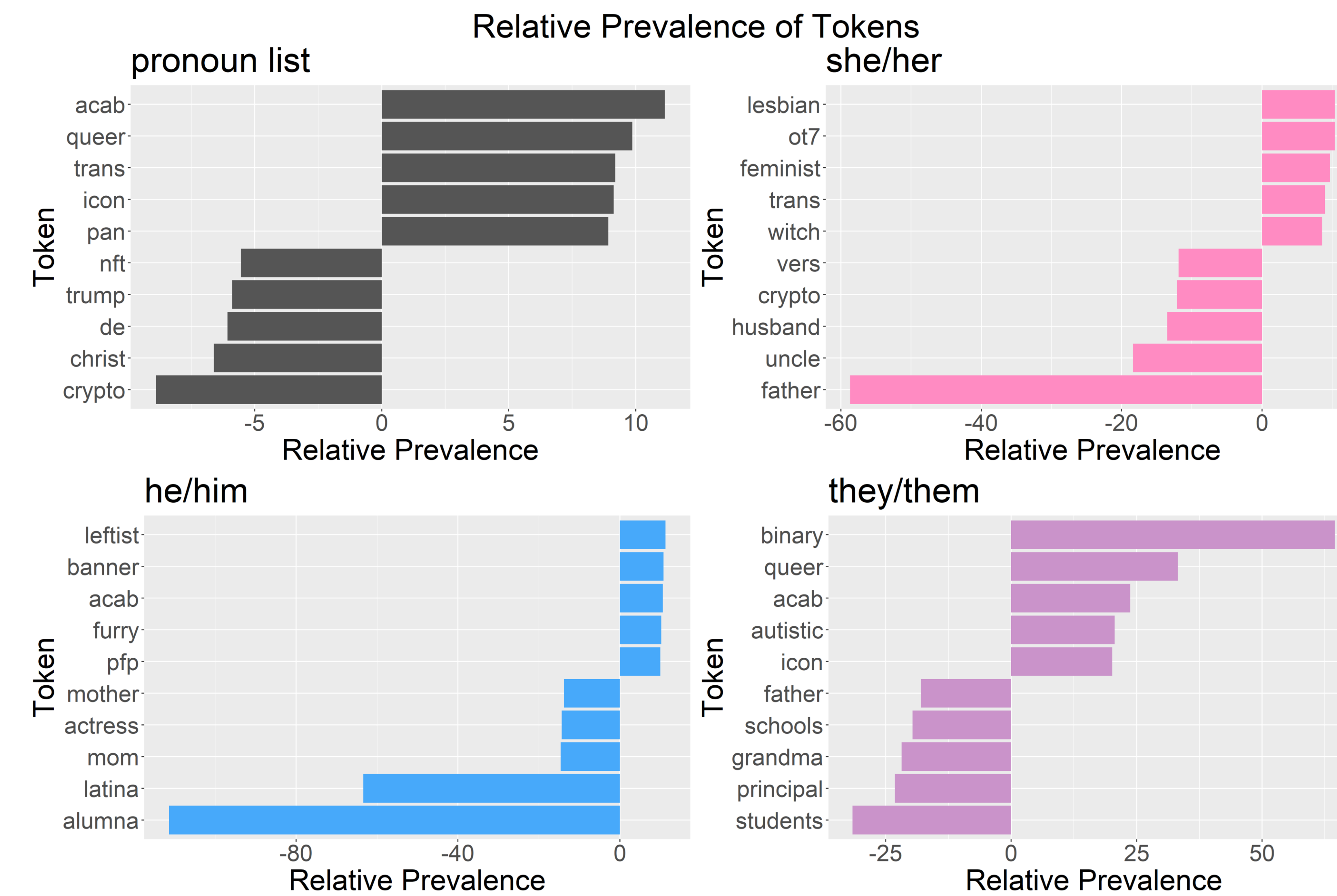
Clustering

14.1% of friends of a user with a pronoun list have a pronoun list, compared to **2.6%** of friends of users without a pronoun list



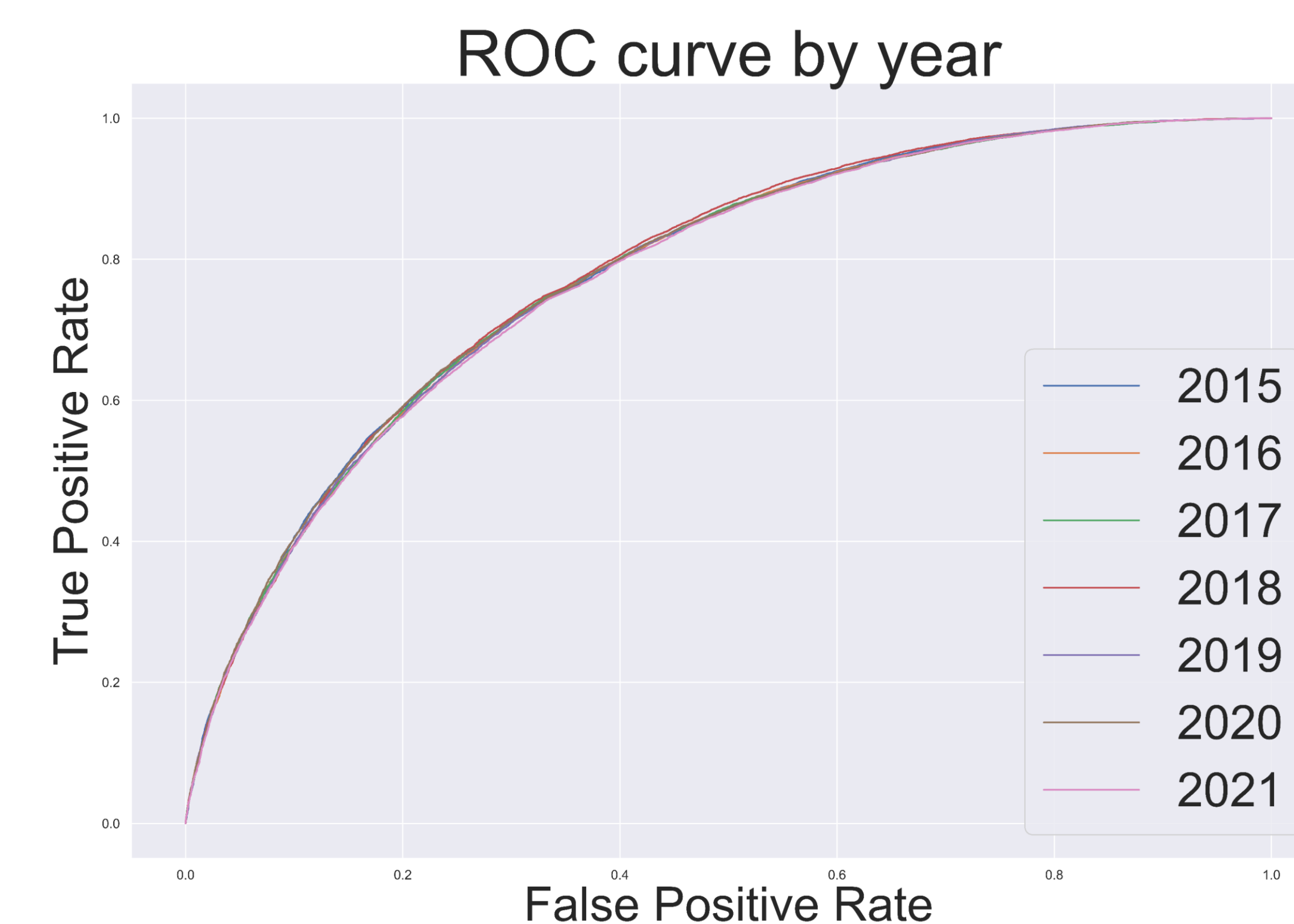
Tokens Co-occurring with Pronoun Lists

- Dataset: 5.4 million US Twitter bios from Jan 1 - June 30, 2022
- Any given user appears at most once
- Methodology: split bios into tokens, count appearances of each token, calculate prevalence and relative prevalence



Multinomial Naïve Bayes Model

- Predicts whether a user will add a pronoun list to their Twitter bio between X and 2022
- Two classes: added a pronoun list to their bio, and did not have pronouns either year



Predicting adding a pronoun list from 2021 to 2022

Accuracy 0.693

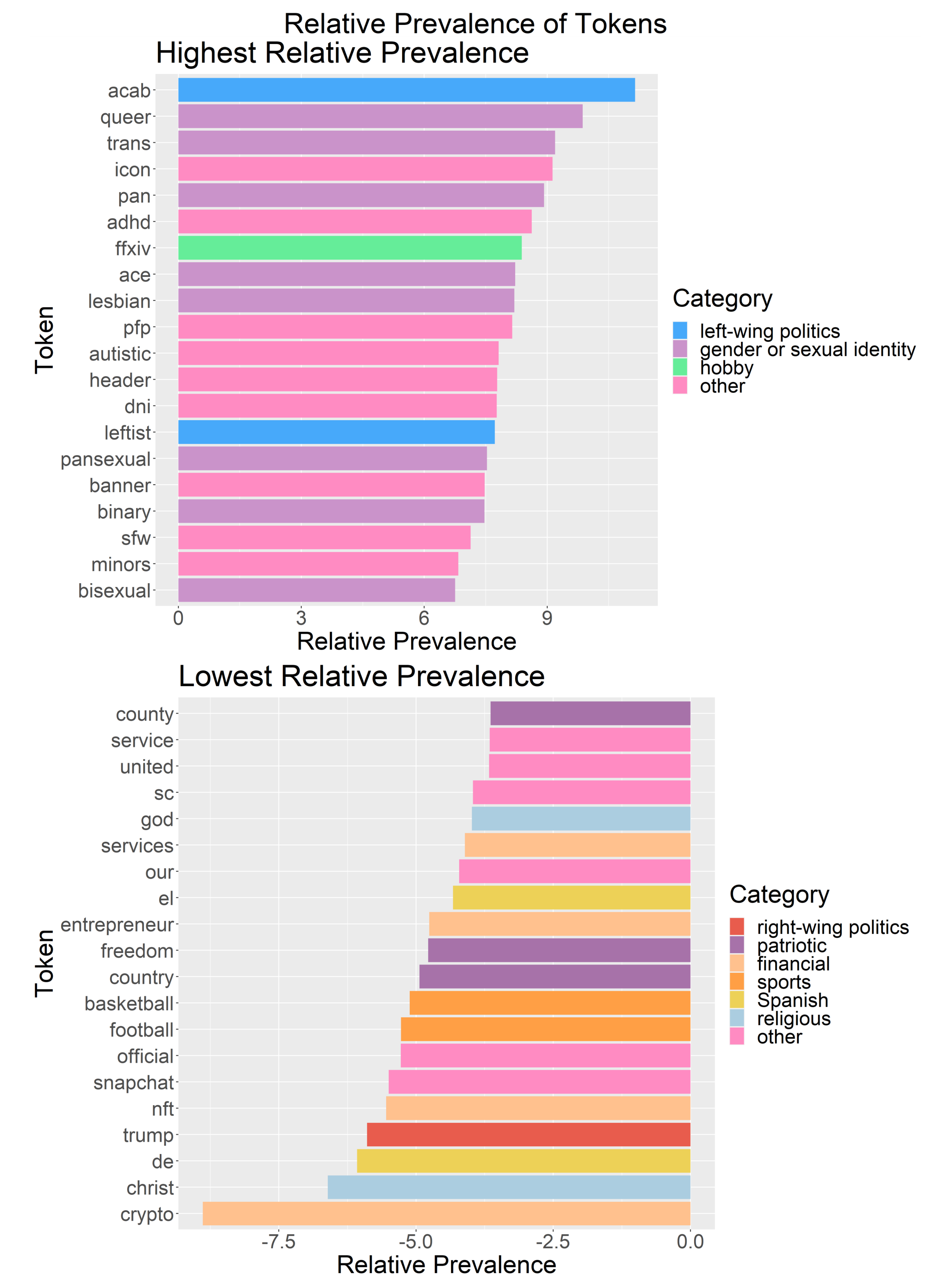
F1 Score 0.692

ROC AUC 0.778

Bios per class 32,664

Categorizing Relative Prevalence

- Relative prevalence is among bios with any pronoun list
- Tokens manually grouped into categories



References

Rogers, N., & Jones, J. J. (2021). Using Twitter BIOS to measure changes in self-identity: Are Americans defining themselves more politically over time? *Journal of Social Computing*, 2(1), 1–13. <https://doi.org/10.23919/jsc.2021.0002>

Wamsley, L. (2021, June 2). *A guide to gender identity terms*. NPR. Retrieved July 5, 2022, from <https://www.npr.org/2021/06/02/996319297/gender-identity-pronouns-expression-guide-lgbtq>

Jones, J. J. (2021). A dataset for the study of identity at scale: Annual prevalence of American twitter users with specified token in their profile bio 2015–2020. *PLOS ONE*, 16(11). <https://doi.org/10.1371/journal.pone.0260185>

This material is based upon work supported by the National Science Foundation under grants IIS-1927227 and OAC-1950052.
Address correspondence to: Liam E. Tucker letucker@crimson.ua.edu
Special thanks to Margot A. Hare