

Political Dynasties in Brazil: Quantification through Text Mining and String Matching



Octavio E. Lima¹ and F. Daniel Hidalgo² (Ph.D.)

¹ Department of Economics, University of Oregon
 ² Department of Political Science, Massachusetts Institute of Technology

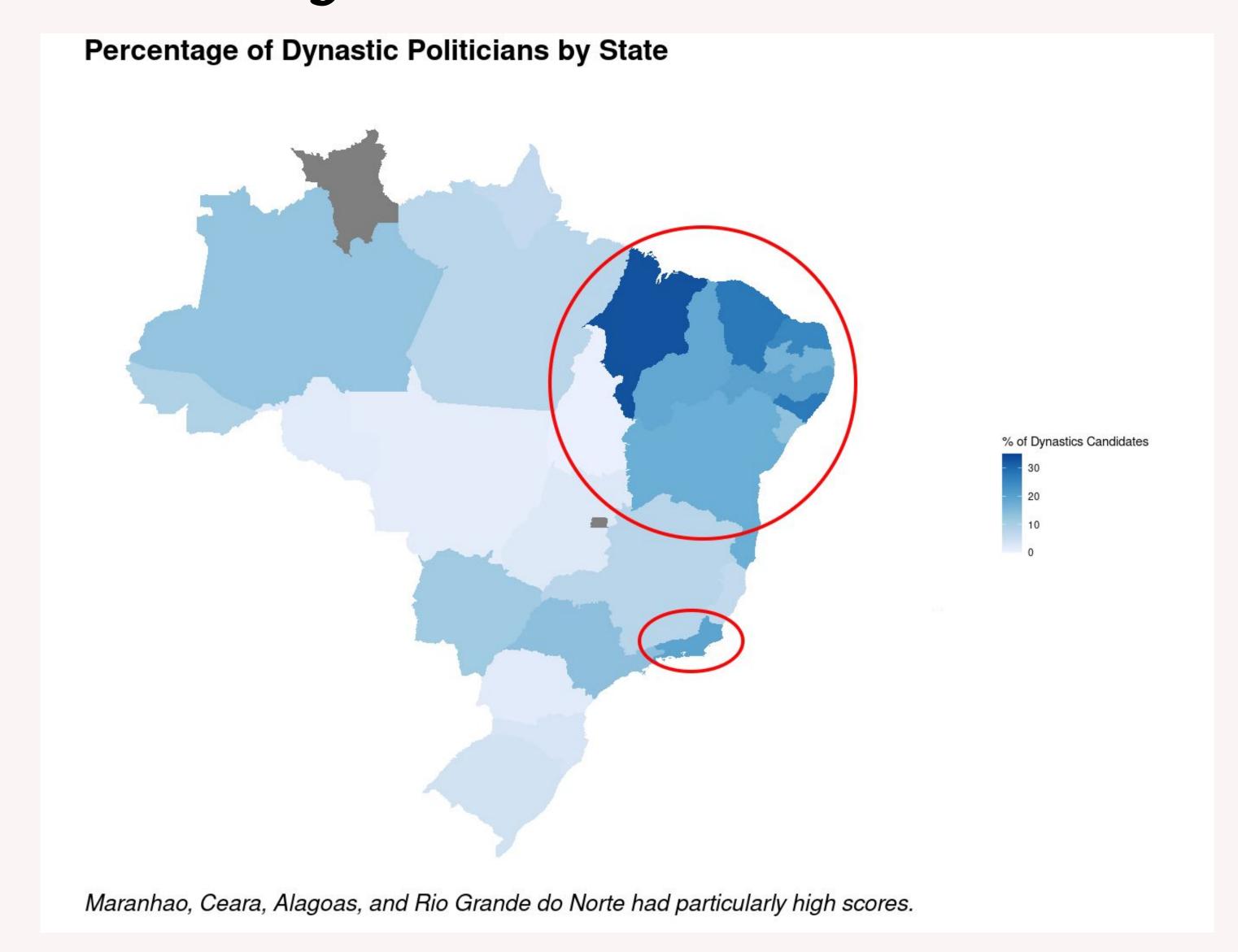
1. Background

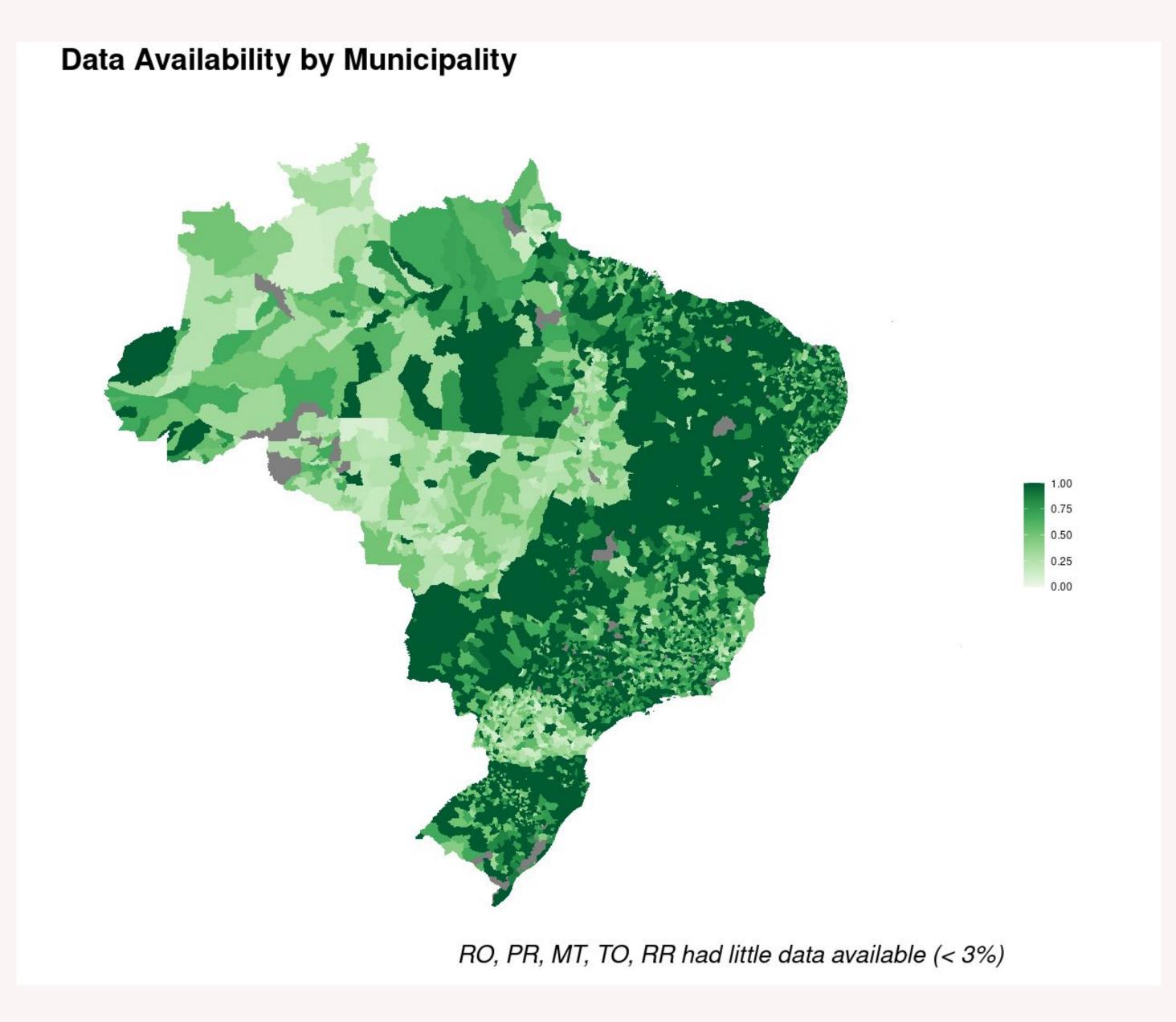
- Dynasties may affect democracies in unforeseeable ways.
- Quantifying political dynasties is challenging.
- Data in this field is complex and mostly inaccessible.
- Past studies are encouraging but limited.
- Previous approaches relied on last names alone as predictors; this may lead to biases when surnames are common.

2. Methodology

- We assessed the eligibility documents of the 2020 mayoral candidates (from Brazil's Superior Electoral Court).
- Several of these files contained the candidate's parents' information.
- We then converted these .png/.pdf files into .txt using O.C.R. and Cloud Computing.
- We used RStudio and Regular Expressions to loop through all these files (over 120,000).
- Based on specific keywords (eg, "son of," "parents," "father's name,"), our code grabbed the names of the candidates' parents.
- After that, we used R's fastLink and fuzzyjoin packages to run an *Approximate* ("Fuzzy") String Matching with an existing dataset of Brazilian politicians from 1994 to 2020 (with roughly 2.8 million observations).
- This approach enabled us to find not only exact matches but also approximate matches. For instance, our code managed to catch misspelling errors such as in "Bernardo Aguzzoli" vs "Bernardo Aguzzoli"."
- We were, then, able to find how certain states compare to others in terms of Dynasties within government positions, such as Mayoral and City Councilor.

3. Results





4. Conclusion

- We have found Northeastern States to have the highest rates of dynasties.
- Maranhão, Ceara, Alagoas, and Rio Grande do Norte had particularly high scores.
- On the other end, Parana, Santa Catarina, and Rio Grande do Sul had lower-than-average scores.
- It is important to note, however, that some states did not have much data available; e.g., **RO, PR, MT, TO, RR.**.

5. Future Work

- Future studies may investigate why Northeastern states have higher-than-normal rates of Dynasties.
- We suggest that interested scholars examine why Maranhão, in specific, had the highest score by our metrics.
- Additionally, scholars should further explore the influence of the **Sarney family** in the above mentioned state of **Maranhão**.

6. References

- [1] Ernesto Dal Bó, Pedro Dal Bó, and Jason Snyder. Political Dynasties. Rev Econ Stud, 76(1):115–142, January 2009. Publisher: Oxford Academic.
- [2] Julien Labonne, Sahar Parsa, and Pablo Querubín. Political Dynasties, Term Limits and Female Political Empowerment: Evidence from the Philippines. Working Paper 26431, National Bureau of Economic Research, November 2019. Series: Working Paper Series
- [3] Pablo Querubin. Family and Politics: Dynastic Persistence in the Philippines. QJPS, 11(2):151–181, July 2016. Publisher: Now Publishers, Inc.
- [4] Yasushi Asako, Takeshi Iida, Tetsuya Matsubayashi, and Michiko Ueda. Dynastic Politicians: Theory and Evidence from Japan. Japanese Journal of Political Science, 16(1):5–32, March 2015. Publisher: Cambridge University Press.
- [5] Arthur Bragança, Claudio Ferraz, and Juan Rios. Political Dynasties and the Quality of Government. page 53, 2015.