Codebook for Congressional Elections and Ideology Dataset

October 23, 2025

Description:
This dataset combines fundraising and candidate positioning data with candidate-level vote shares in congressional primary and general elections. The dataset features a composite ideology measure derived from 14 separate input measures, including campaign finance, roll-call votes, survey responses, and text analysis.

Variable Descriptions:

1) `Year`: Four-digit number that indicates the two-year election cycle.

2) `Chamber`: Chamber of Congress (House or Senate).

3) `distcyc`: Unique District-Cycle identifier (e.g., 2024_AL_01).

4) `Democrat`: Name of the Democratic candidate.

5) `Republican`: Name of the Republican candidate.

6) `dem_rid`: Unique DIME ID assigned to the Democratic candidate. Constant across cycles.

7) `dem_fec_cand_id`: The candidate ID assigned by the FEC to the Democratic candidate.

8) `dem_fec_comm_id`: The ID assigned by the FEC to the Democratic candidate's campaign committee.

9) `rep_rid`: Unique DIME ID assigned to the Republican candidate. Constant across cycles.

10) `rep_fec_cand_id`: The candidate ID assigned by the FEC to the Republican candidate.

11) `rep_fec_comm_id`: The ID assigned by the FEC to the Republican candidate's campaign committee.

12) `dem_cand_gender`: Gender of the Democratic candidate.

13) `rep_cand_gender`: Gender of the Republican candidate.

14) `dem_vote_share`: Democratic candidate's vote share in the general election.

15) `dem_pres_vs`: District-level percentage of the two-party vote share won by the Democratic presidential nominee in the most recent presidential election cycle.

16) `dem_diff`: Democratic candidate's vote share minus the presidential vote share.

17) `incumbency_status`: Incumbency status of the district (e.g., Dem Incumbent, Rep Incumbent, Open Seat).

18) `dem_tenure`: Tenure of the Democratic candidate.

19) `rep_tenure`: Tenure of the Republican candidate.

20) `turnout_rate_dem`: Turnout rate for registered Democrats in district.

21) `turnout_rate_rep`: Turnout rate for registered Republicans in district.

22) `turnout_rate_ind`: Turnout rate for registered Independents in district.

23) `dem_composite`: Composite measure of the Democratic candidate's ideology, derived from 14 input measures using multiple over-imputation and principal component analysis.

24) `rep_composite`: Composite measure of the Republican candidate's ideology, derived from 14 input measures using multiple over-imputation and principal component analysis.

25) `dem_cces_voter_perceptions`: Survey-based measure. Respondents rate where they perceive their congressional candidate on a liberal–conservative scale, with positions estimated using a Bayesian Aldrich–McKelvey model.

26) `rep_cces_voter_perceptions`: Survey-based measure. Respondents rate where they perceive their congressional candidate on a liberal–conservative scale, with positions estimated using a Bayesian Aldrich–McKelvey model.

27) `dem_cfscore`: Recipient DIME (CFscores). Ideal point based on the ideologies of a candidate's contributors.

28) `rep_cfscore`: Recipient DIME (CFscores). Ideal point based on the ideologies of a candidate's contributors.

29) `dem_dwdime`: DW-DIME. Predicts a candidate's future DW-NOMINATE score using their campaign finance history.

30) `rep_dwdime`: DW-DIME. Predicts a candidate's future DW-NOMINATE score using their campaign finance history.

31) `dem_dwnom`: DW-NOMINATE. First-dimension scores for members of Congress based on their voting records.

32) `rep_dwnom`: DW-NOMINATE. First-dimension scores for members of Congress based on their voting records.

33) `dem_contrib_cfscore`: Contributor DIME. Ideal point based on a candidate's own donations to other political actors.

34) `rep_contrib_cfscore`: Contributor DIME. Ideal point based on a candidate's own donations to other political actors.

35) `dem_contrib_dwnom`: Contributor DW-NOMINATE. Similar to Contributor DIME, but scaled based on the DW-NOMINATE scores of donation recipients.

36) `rep_contrib_dwnom`: Contributor DW-NOMINATE. Similar to Contributor DIME, but scaled based on the DW-NOMINATE scores of donation recipients.

37) `dem_hall_snyder`: Hall and Snyder Scores. Supervised measure based on candidates' own donations and roll-call votes.

38) `rep_hall_snyder`: Hall and Snyder Scores. Supervised measure based on candidates' own donations and roll-call votes.

39) `dem_ggum`: GGUM. "Ends-against-the-middle" scores for non-monotonic voting.

40) `rep_ggum`: GGUM. "Ends-against-the-middle" scores for non-monotonic voting.

41) `dem_shor_mcarty_state_leg`: Shor-McCarty. Common-space scores for state legislators based on state legislative roll-call votes.

42) `rep_shor_mcarty_state_leg`: Shor-McCarty. Common-space scores for state legislators based on state legislative roll-call votes.

43) `dem_bailey_bridge`: Bailey Bridge Scores. Roll-call scores that are "bridged" across congresses to ensure comparability over time.

44) `rep_bailey_bridge`: Bailey Bridge Scores. Roll-call scores that are "bridged" across congresses to ensure comparability over time.

45) `dem_web_text_score`: WEB scores. Scores estimated using a word embedding model (Doc2Vec) on campaign issue statements.

46) `rep_web_text_score`: WEB scores. Scores estimated using a word embedding model (Doc2Vec) on campaign issue statements.

47) `dem_meisels_text_score`: Meisels Wordfish. Unsupervised scaling of text from campaign platforms using a Poisson IRT model.

48) `rep_meisels_text_score`: Meisels Wordfish. Unsupervised scaling of text from campaign platforms using a Poisson IRT model.

49) `dem_twitter_score`: Twitter Scores. Bayesian ideal points estimated from the follower networks of political actors.

50) `rep_twitter_score`: Twitter Scores. Bayesian ideal points estimated from the follower networks of political actors.

51) `dem_total_receipts`: Total dollars raised by the Democratic candidate during the election cycle.

52) `rep_total_receipts`: Total dollars raised by the Republican candidate during the election cycle.

53) `dem_pac_mod_nyt`: Democratic moderate per NYT/PAC rules. Indicator of NYT codings: 'mod' or 'nonmod' based on PAC support.

54) `rep_pac_mod_nyt`: Republican 'moderate' per NYT/PAC rules. Indicator of NYT codings: 'mod' or 'nonmod' based on PAC support.

55) `dem_pac_mod_nyt`$_2$4 : $Democratic moderate per NYT/PAC rules. Indicator of NYT codings:' mod' or' nonmod' based on PAC support limited to the 2024 election cycle.$

55) `rep_pac_mod_nyt`$_2$4 : $Republican' moderate' per NYT/PAC rules. Indicator of NYT codings:' mod' or' nonmod' based on PAC support limited to the 2024 election cycle.$

55) `dem_total_receipts.1`: Total dollars raised by the Democratic candidate (duplicate of `dem_total_receipts`).

56) `dem_total_pac_contribs`: Total PAC receipts for the Democratic candidate.

57) `dem_total_indiv_unitemized`: Total unitemized individual receipts for the Democratic candidate.

58) `dem_total_indiv_contrib`: Total itemized individual receipts for the Democratic candidate.

59) `dem_total_contribs_from_self`: Total receipts from the Democratic candidate's own contributions.

60) `dem_total_disbursements`: Total campaign disbursements (in dollars) for the Democratic candidate.

61) `rep_total_receipts.1`: Total dollars raised by the Republican candidate (duplicate of `rep_total_receipts`).

62) `rep_total_pac_contribs`: Total PAC receipts for the Republican candidate.

63) `rep_total_indiv_unitemized`: Total unitemized individual receipts for the Republican candidate.

64) `rep_total_indiv_contrib`: Total itemized individual receipts for the Republican candidate.

65) `rep_total_contribs_from_self`: Total receipts from the Republican candidate's own contributions.

66) `rep_total_disbursements`: Total campaign disbursements (in dollars) for the Republican candidate.