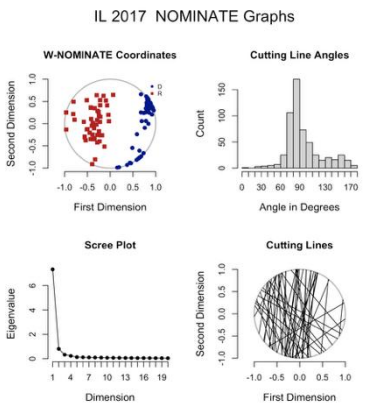
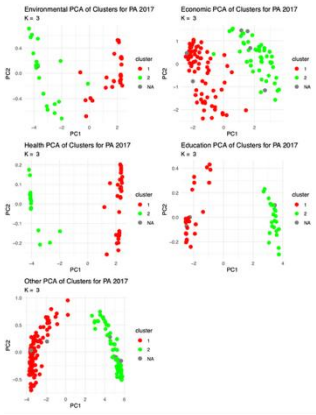
Statistical Investigation into Representations of Voting Bloc Structure in American State Legislatures

Lierin Peterson, Class of 2026

Research by political scientists and statisticians on the United States' legislative branch has shown an increasing divide between the ideological left and right in the US House and Senate, creating two groups who mainly vote along party lines. This polarization was clearly demonstrated in 1983 by Poole and Rosenthal, who introduced the DW-NOMINATE score, a spatial scoring method that tracked ideological shifts over time on the conservative-liberal and social issues scales. Although this score is technically two-dimensional, information from the first dimension predicts votes with 83% accuracy, and adding the second dimension only increases the accuracy to 85%. Therefore, the ideological position of members of the US House can be fairly accurately described using only a single number.



For our project, we investigated state legislatures, specifically state houses, which we predicted might have more nuance in their voting patterns, and thus be less well represented by the NOMINATE score. Based on the promising results from Lauren Traum's ('25) research, we suspected that latent patterns of ideology, not pure democratic/republican allegiance, would predict voting behavior. Our states of focus were Massachusetts, Pennsylvania, Maine, Minnesota, Maryland, and Oregon, but we also looked at several other states, including Illinois, New Mexico, and Alabama. We chose these states based on several criteria, including previous promising results, availability of data, and potential for interesting patterns from the states' known political structure.



To analyze each state, we created a data pipeline that organized each state's raw roll call data and created several representations of the ideological positions of the representatives, including the NOMINATE scores, NMDS graphs, PAM clustering, PCA Clustering, and weighted and unweighted Hamming Distances between legislators. Each of these methods visualized how "close" or "far" legislators were from each other ideologically, ultimately creating a picture of the legislature that showed the group's overall structure. Additionally, we examined bills by the committees they had been referred through, categorizing bills into Environment, Education, Health, Economic, and "Other" to determine if issue-specific patterns differed significantly from global voting patterns.

Overall, our results showed that the NOMINATE score is a strong representation of US State Legislatures, both at the issue-specific and at the global level, and most voting variation can be attributed to a representative's position on a left/right scale.

Faculty Mentor: Jack O'Brien
Funded by the Surdna Foundation Undergraduate Research Fellowship Program