

POL 212
Winter 2024
Assignment 5

Answer ONE of the following questions (A or B), unless you are a methods minor (in which case, do both!):

A. Simulation:

1. Simulate a DGP with a single dependent variable (y) that is the function of three independent variables (x_1 , x_2 , x_3). This DGP should have both systematic and stochastic components, but otherwise you are free to generate x_1 , x_2 , x_3 , and y as you choose.
2. Now modify the DGP so that any two of the independent variables (x_1 , x_2 , x_3) interact in some way to influence y .
3. Fit and present a single decision tree to model y using x_1 , x_2 , and x_3 as predictor variables.
4. Assess how well the tree performed in finding your interaction.

B. Data analysis:

1. Fit and present a single decision tree to the ANES 2022 Pilot Study data from the previous assignment (Assignment 4). Recall we are trying to predict the difference in Trump and Biden feeling thermometers as a function of race, gender, age, income, and education.
2. Interpret the tree's results. What do its decision rules (splits) mean? Are there any noteworthy candidate interactions in the tree? Explain.
3. What is the R^2 value for the tree? Compare this to the R^2 value for a linear regression model of the same data.