

Data Cleaning Notes:

age buckets:

- 1 = 18-34
- 2 = 35-54
- 3 = 55+

vote choice:

- 1 = donald trump
- 0 = joe biden

political party:

- 1 = republican
- 2 = democrat
- 3 = unknown

race

- 1 = white
- 2 = black
- 3 = asian
- 4 = latino
- 9 = other/unknown

gender / male

- 1 = male
- 0 = female / not male

education

- 1 = high school or less
- 2 = beyond highschool

political leaning

- 1 = liberal
- 2 = conservative
- 3 = moderate
- 4 = unknown

states split into regions based on:

https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf

region:

- 1 = new england
- 2 = midwest
- 3 = south
- 4 = west

new_england = [9.0, 23.0, 25.0, 33.0, 44.0, 50.0, 34.0, 36.0, 42.0]

midwest = [18.0, 17.0, 26.0, 39.0, 55.0, 19.0, 20.0, 27.0, 29.0, 31.0, 38.0, 46.0]

south = [10.0, 11.0, 12.0, 13.0, 24.0, 37.0, 45.0, 51.0, 54.0, 1.0, 21.0, 28.0, 47.0, 5.0, 22.0, 40.0, 48.0]

west = [4.0, 8.0, 16.0, 35.0, 30.0, 49.0, 32.0, 56.0, 2.0, 6.0, 15.0, 41.0, 53.0]