Demographic

- Get from US census (map to precinct)
- Get 2010 is good enough
- Should be in precinct level

Voting Data

- Use original data first (census and state election data)
- Senate election data is not required
- Congressional result: Only US House of representative election data (2016 and 2018)
    - Each election district chose a single house of representative
    - Number of district is decided by census bureau
- And US presidential election data (2016)
- Precinct level election result (MIT or Office of Secretary of State)

Geographic Data

- Precinct boundaries
    - Harvard election data
    - Open elections
    - https://www.sos.state.mn.us/election-administration-campaigns/data-maps
- Congressional district boundaries : A congressional district is an electoral constituency that elects a single member of a congress.
- Us border
- State county
- National state park
- Water source

Data needed

- Geospatial boundary
    - Precincts
    - Existing congressional districts
    - Incumbent precincts
    - Cities / counties
    - Census tracts (including demographic data)
    - State boundary data
- Election result data
    - Congressional election districts
    - Congressional election result:
        - most recent 2 congressional elections
        - most recent presidential election
    - election district data
        - data by precinct
    - population demographic
        - from us census
    -

Preprocessing:

- Build ghost precincts
- Break out precinct boundary data if data source groups it together (into state level)
- Determine precinct neighbors (probably not all)
- Map some data identifiers to a canonical name (precinct name)
- Identify potential data problems for the user to correct
- Combine multiple data sources (census) to generate complete precinct data
- Precinct
    - Enclosed data
        - Set inner precinct to have only 1 neighbor (the enclosing precinct) OR
        - Simply merge the enclosed precincts into the enclosing precinct and drop the enclosed precincts
        - Merging precinct = merging hole of doughnut. (user case 25)
    - Data combining
        - Precinct object should contain:
            - Precinct identifier / county identifier
            - Boundary data
            - Election results
            - Demographic data
        - Same precinct identifier will make combining precinct easy
- Election and demographic data issues
    - Census bureau reports in various levels (blocks, groups, tracts, counties, and states), possibly not precincts
    - Need to identify a census block with a precinct, then accumulate demographic data into the precinct
    - Average precinct is about 60 times larger than average census block
    - Census attempts to coordinate with voting data through voting tabulation districts (VTDs)
    - If census block is located within 2 or more precincts, use a proportional approach to allocate
    - For RI use the county (not precinct) as the smallest unit if accurate precinct level data is not available.

Precinct graph

- graph
    - Each precinct is a node in the graph
    - Edges identify physically adjacent precincts (with distance <= 200 foot)
    - Use shaply python library to decide precinct neighbors. (this is fast) (str tree function)
- Constraints
    - Precinct border does not self-intersect
    - Precinct border is a single polygon

- Boundaries of precinct
    - Adjacent to another precinct OR
    - Part of the state border

Other:

- Mouse over show precinct data
- Show precincts, congressional districts, and density of multiple minority groups