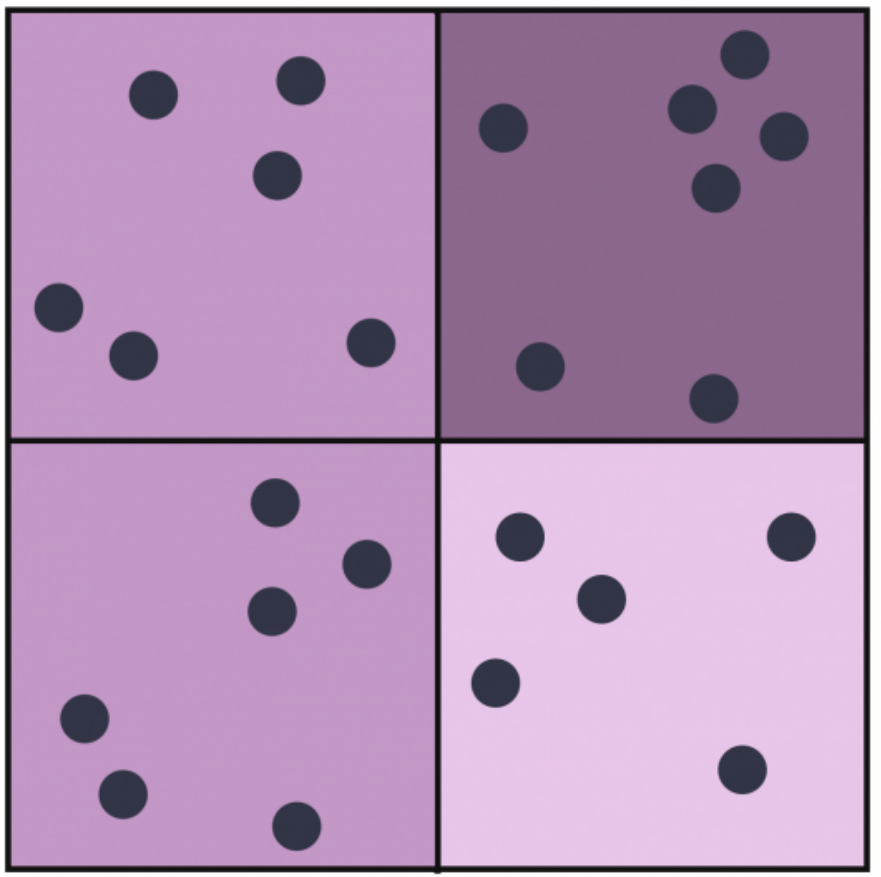
Area data

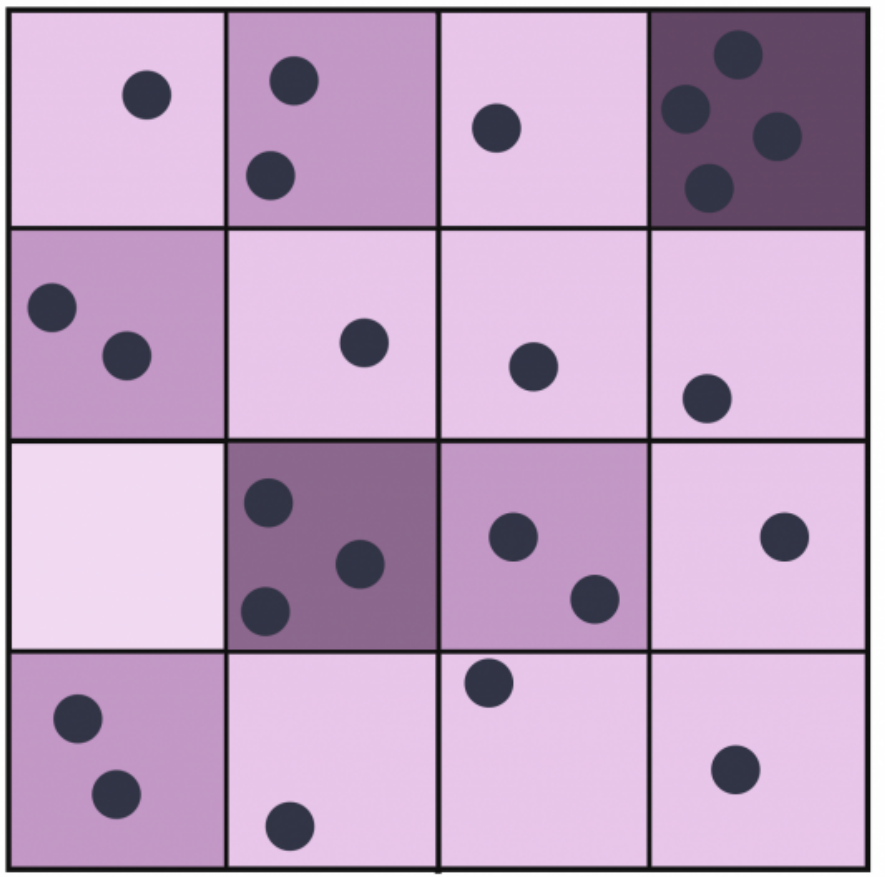
- We are starting to explore area data this week.
- Due to privacy concerns, especially in the social sciences, access to point-level data is rare.
- For instance, publicly available Canadian census data is aggregated to census geography.
  - Typical census geographies include Dissemination Areas, Census Tracts, and Census Subdivisions.
  - To explore the size of these geographies, refer to censusmapper.ca

Modifiable Areal Unit Problem (MAUP)

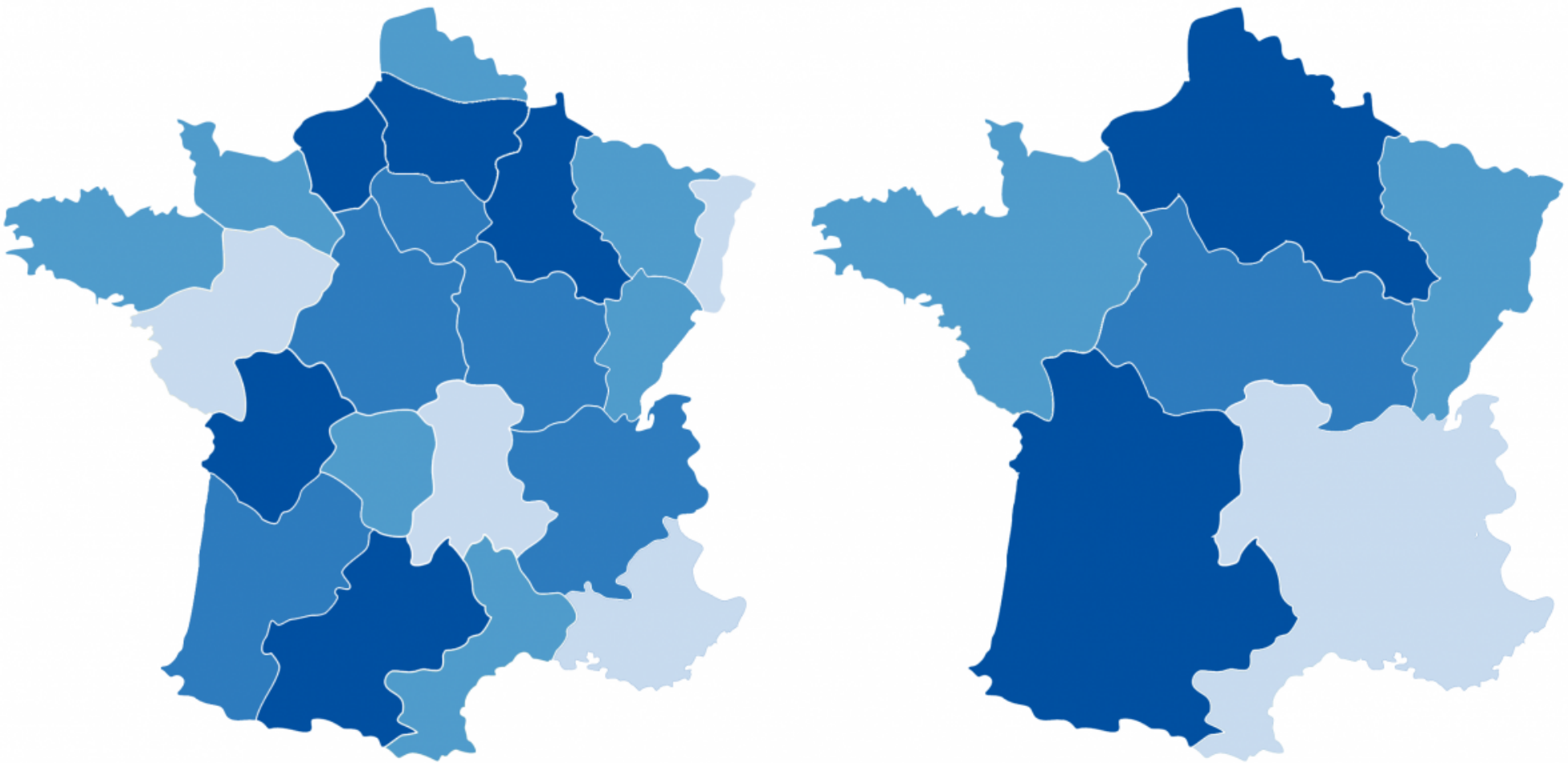
- When aggregating point data into different spatial units, the **choice of spatial units** impacts analysis results such as hypothesis tests.
- This is because any variations within the spatial unit are **lost** during aggregation.
  - Typically, we assume the variable is spatially **uniformly** distributed within each aggregation unit.
  - When working with aggregated data, additional assumptions must be made.

Scale effect



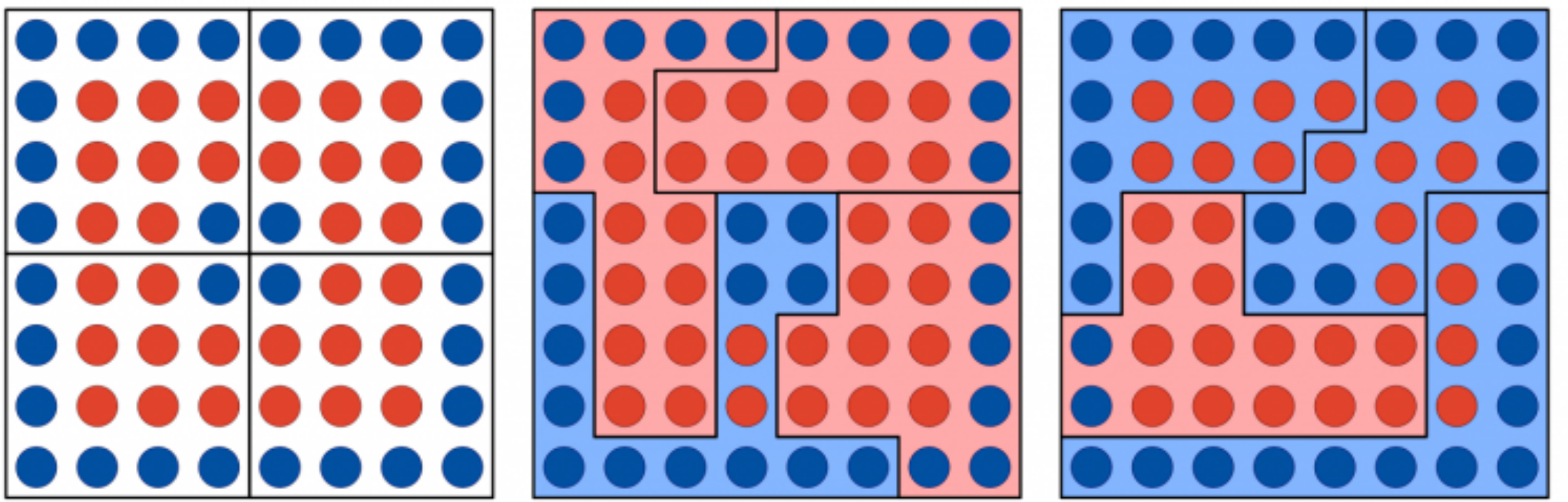


Zone effect

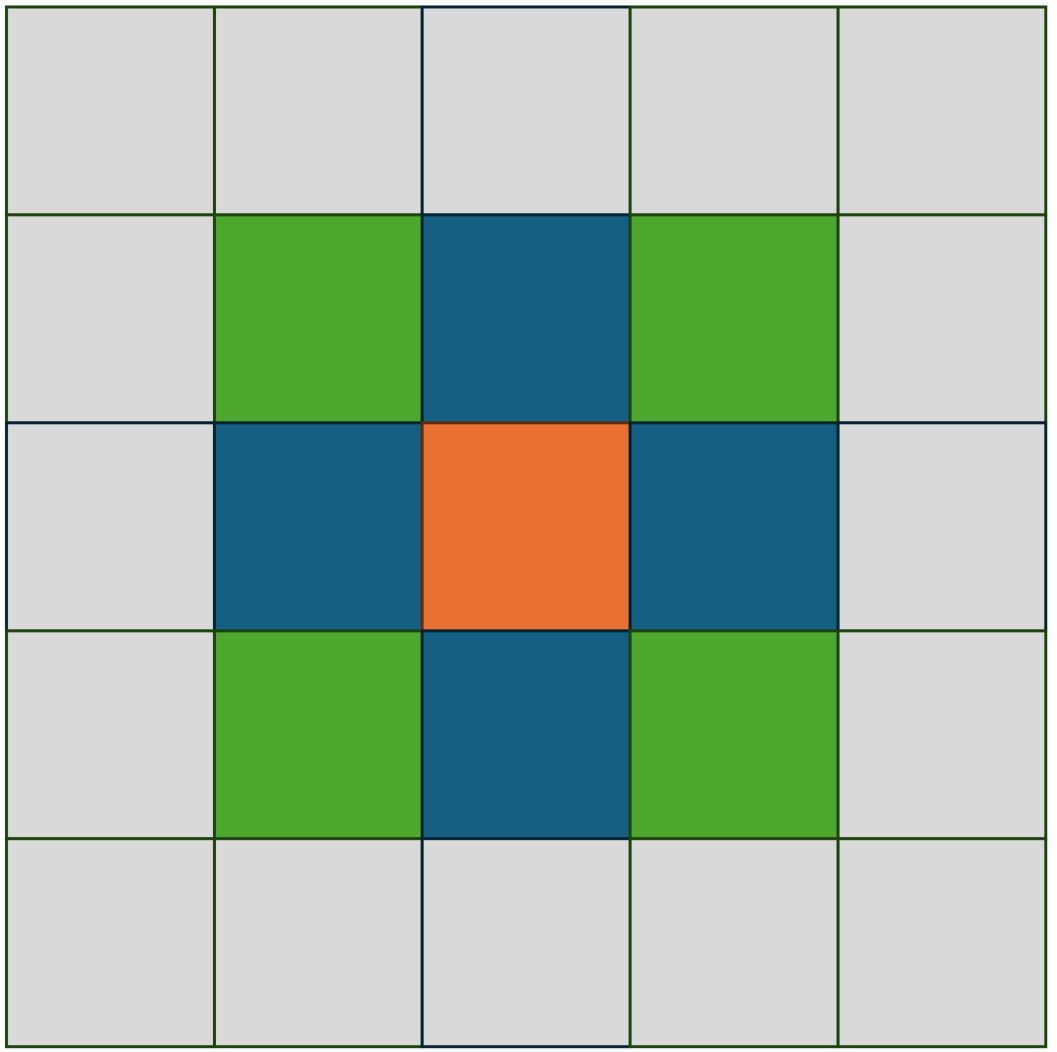


Gerrymandering

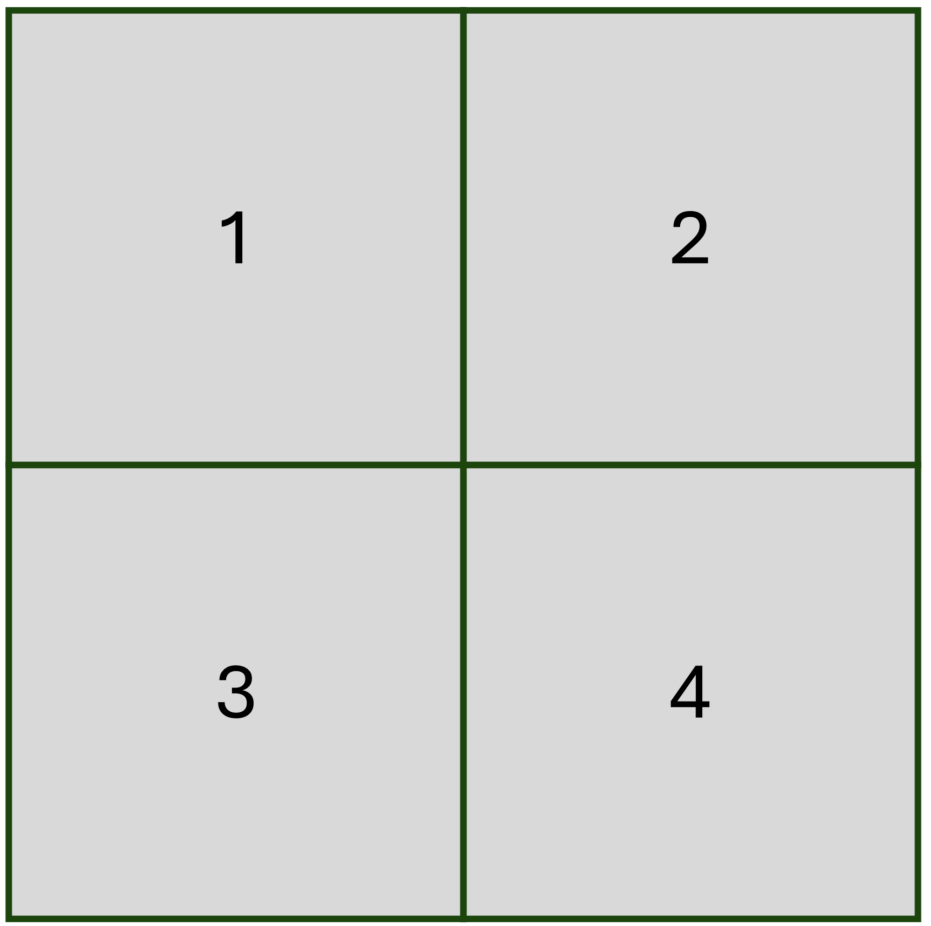
- In representative electoral systems, carefully modifying electoral district boundaries can provide advantages to a particular party.



Rock and queen contiguity



Spatial weight matrix



$$\mathbf{w} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

$$\mathbf{w}' = \begin{pmatrix} 0 & 0.5 & 0.5 & 0 \\ 0.5 & 0 & 0 & 0.5 \\ 0.5 & 0 & 0 & 0.5 \\ 0 & 0.5 & 0.5 & 0 \end{pmatrix}$$

Linear algebra recap

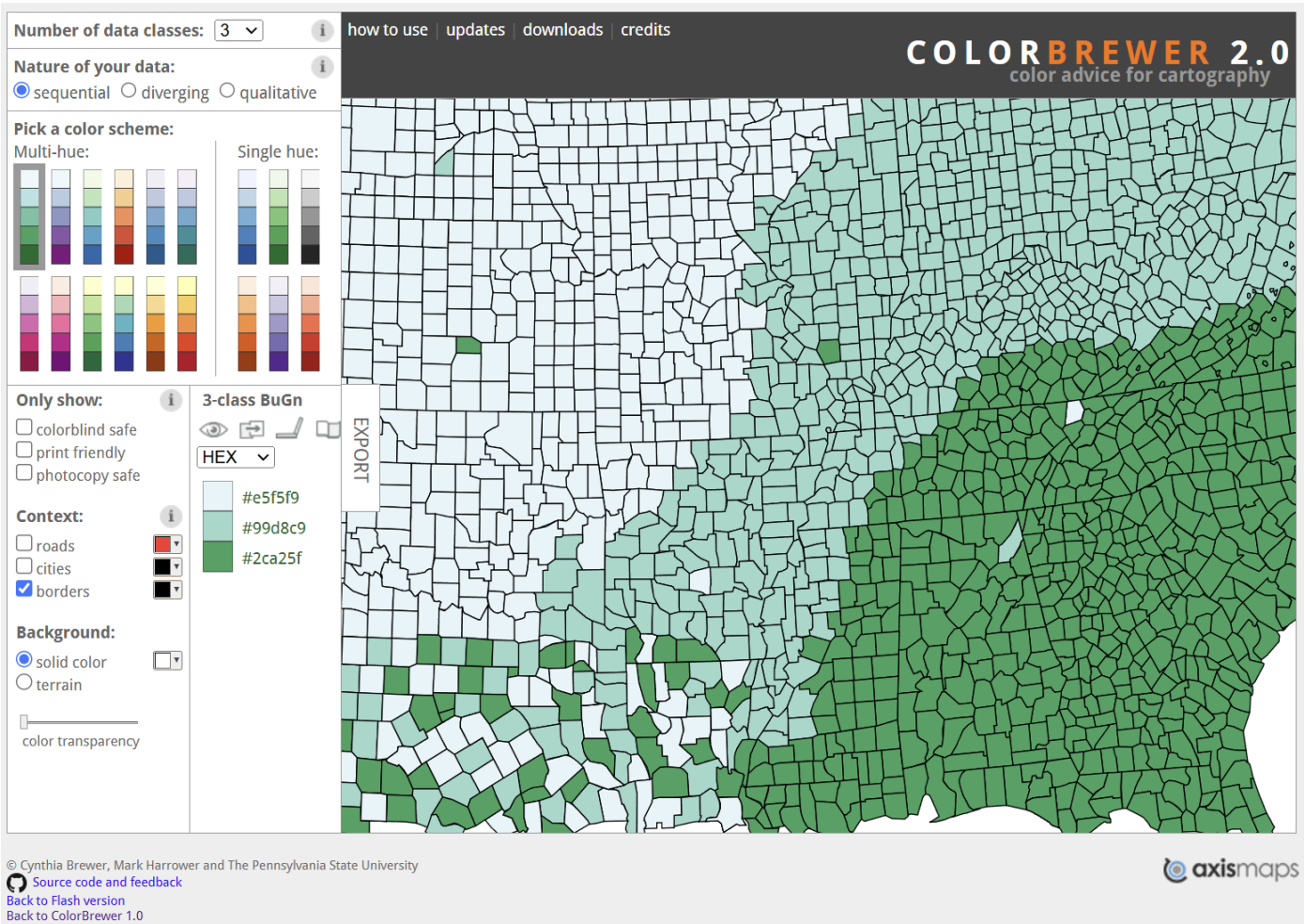
$$\mathbf{W} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$\mathbf{X} = \begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix}$$

$$\mathbf{W} \cdot \mathbf{X} = \begin{pmatrix} a_{11} \cdot b_{11} + a_{12} \cdot b_{21} \\ a_{21} \cdot b_{11} + a_{22} \cdot b_{21} \end{pmatrix}$$

Colour scheme

https://colorbrewer2.org/



Activities for today

- We will work on the following chapter from the textbook:
  - Chapter 20: Activity 9: Area Data I
  - Chapter 22: Activity 10: Area Data II
- The hard deadline is **Tuesday**, **February 24**.

Reference

- https://gisgeography.com/maup-modifiable-areal-unit-problem/