Tools of the Trade

HES 505 Fall 2025: Session 2

Matt Williamson

Checking in

1. What can I clarify about the course?

Today's Plan

- What do we mean by reproducibility?
- What is version control and why is it helpful?
- What is a (spatial) data workflow?
- Why (not) R?

Why (not) R?

Why R?

- Open Source
- Huge useR community
- Integrated analysis pipelines
- Reproducible workflows

Code | Plot

```
library(maps)
library(socviz)
library(tidyverse)
party_colors <- c("#2E74C0", "#CB454A")
us_states <- map_data("state")
election$region <- tolower(election$state)
us_states_elec <- left_join(us_states, election)
p0 <- ggplot(data = us_states_elec,
             mapping = aes(x = long, y = lat,
                           group = group,
                           fill = party))
p1 <- p0 + geom_polygon(color = "gray90",
                        size = 0.1) +
    coord_map(projection = "albers",
              lat0 = 39, lat1 = 45)
p2 <- p1 + scale_fill_manual(values = party_colors
    labs(title = "Election Results 2016",
         fill = NULL)
```

Why not R?

```
## ---
## Error: could not find function "performance"
## ---
##  [1] "Error in if (str_count(string = f[[j]],
##  pattern = \"\\\\S+\") == 1)
##  { : \n  argument is of length zero"
## ---
## Error in eval(expr, envir, enclos) : object 'x' not found
## ---
## Error in file(file, "rt") : cannot open the connection
## ---
```

- Coding can be hard…
- Memory challenges
- Speed
- Decision fatigue

Getting Help

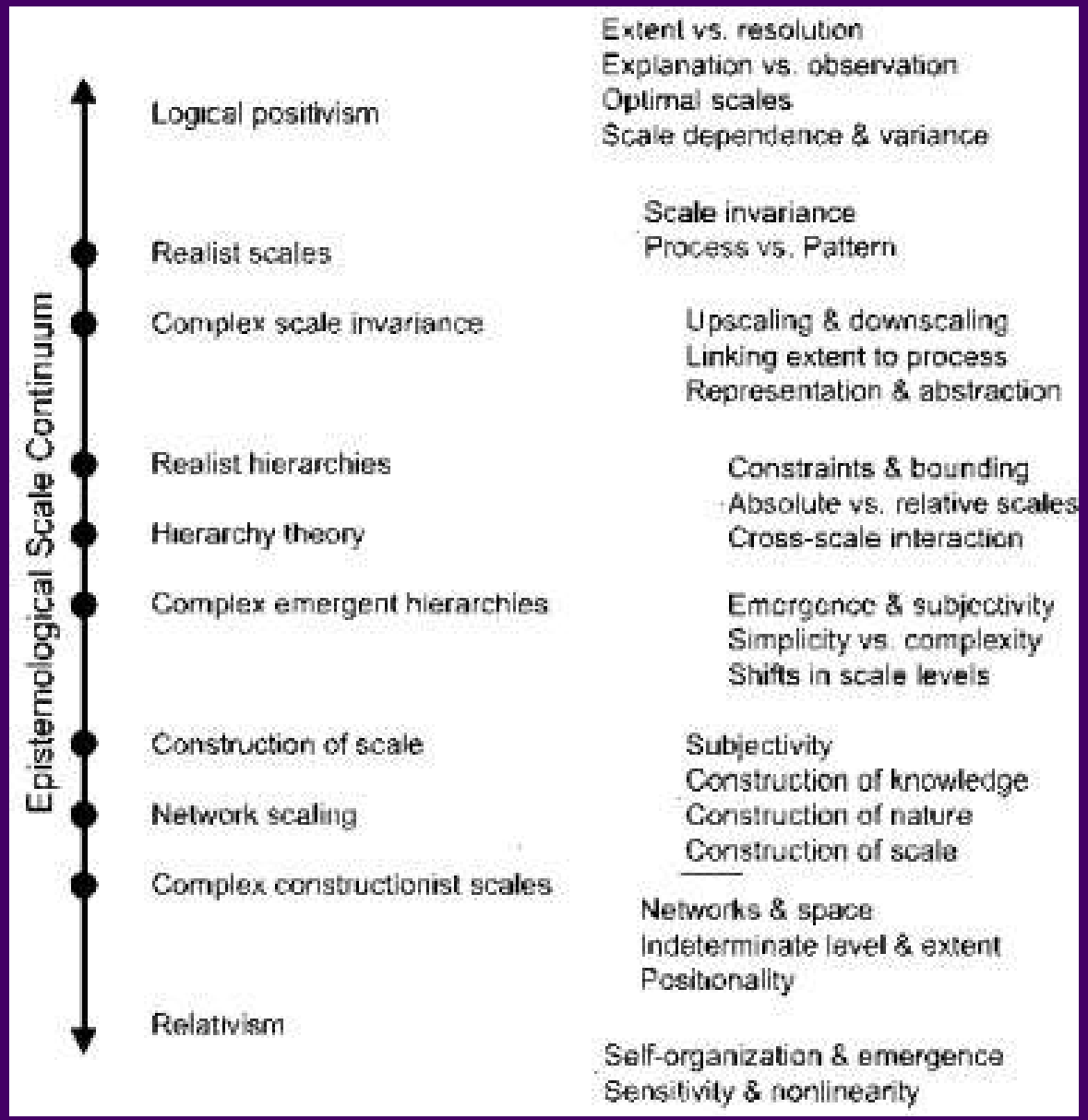
- Google it!!
  - Use the exact error message
  - Include the package name
  - include “R” in the search
- Stack Overflow
  - Reproducible examples
- Package “issue” pages
- r_spatial slack channel
- Common errors

Ask Me

Conceptual challenges

Scale

What do we even mean?

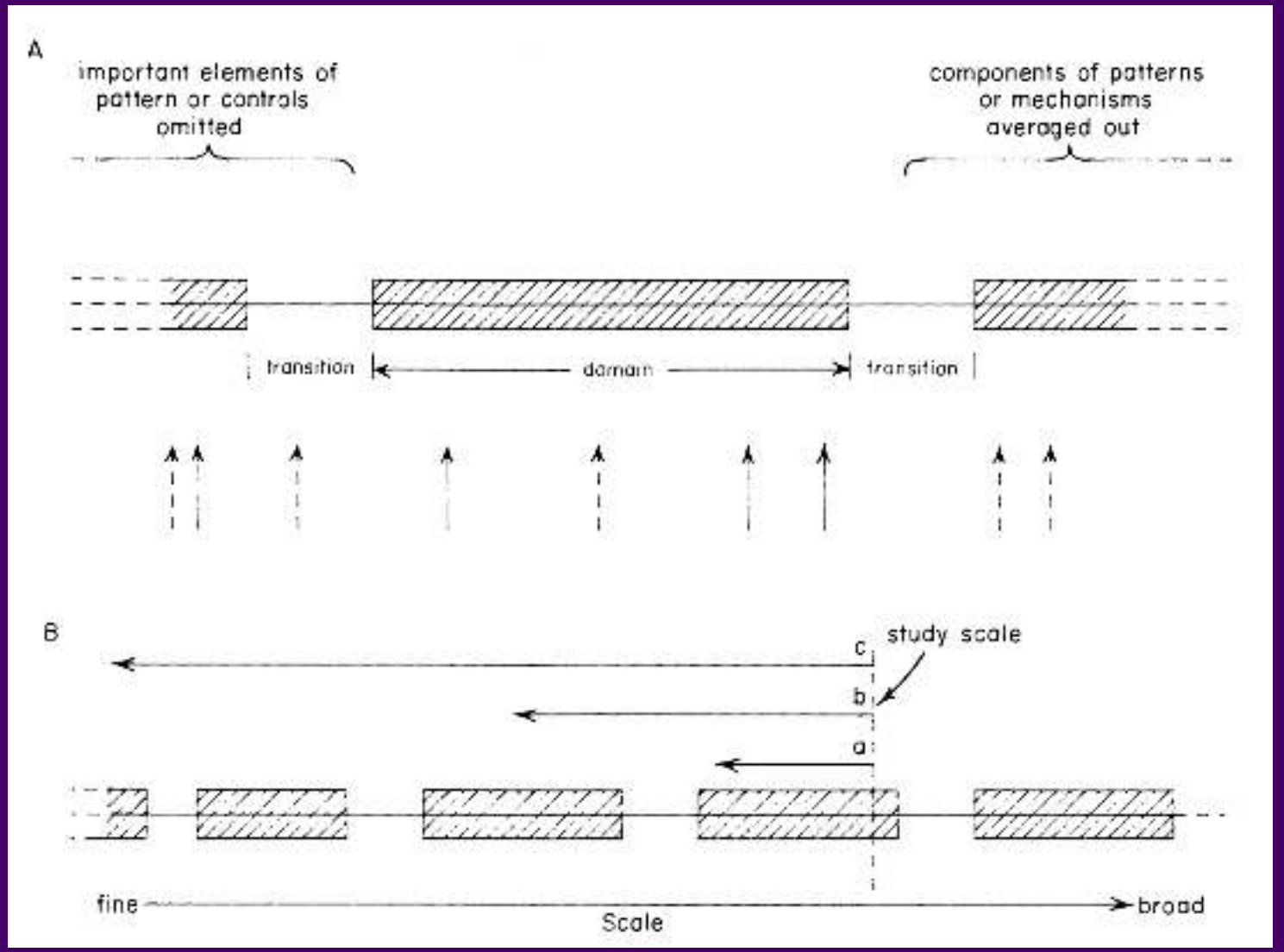


From Manson 2008

- **Grain**: the smallest unit of measurement
- **Extent**: the areal coverage of the measurement

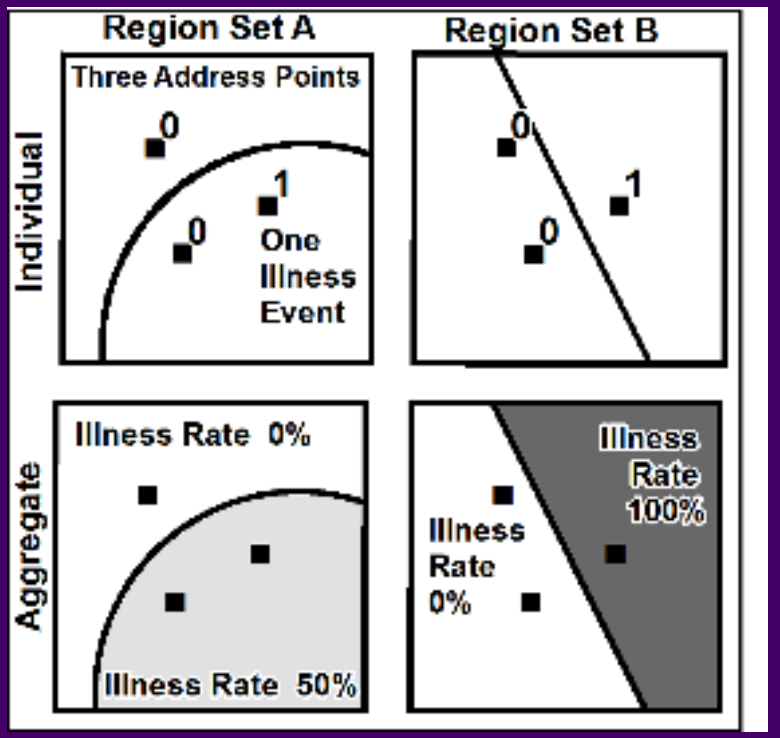
Scale

Even if it exists, how do we know we are measuring at the *right* scale?

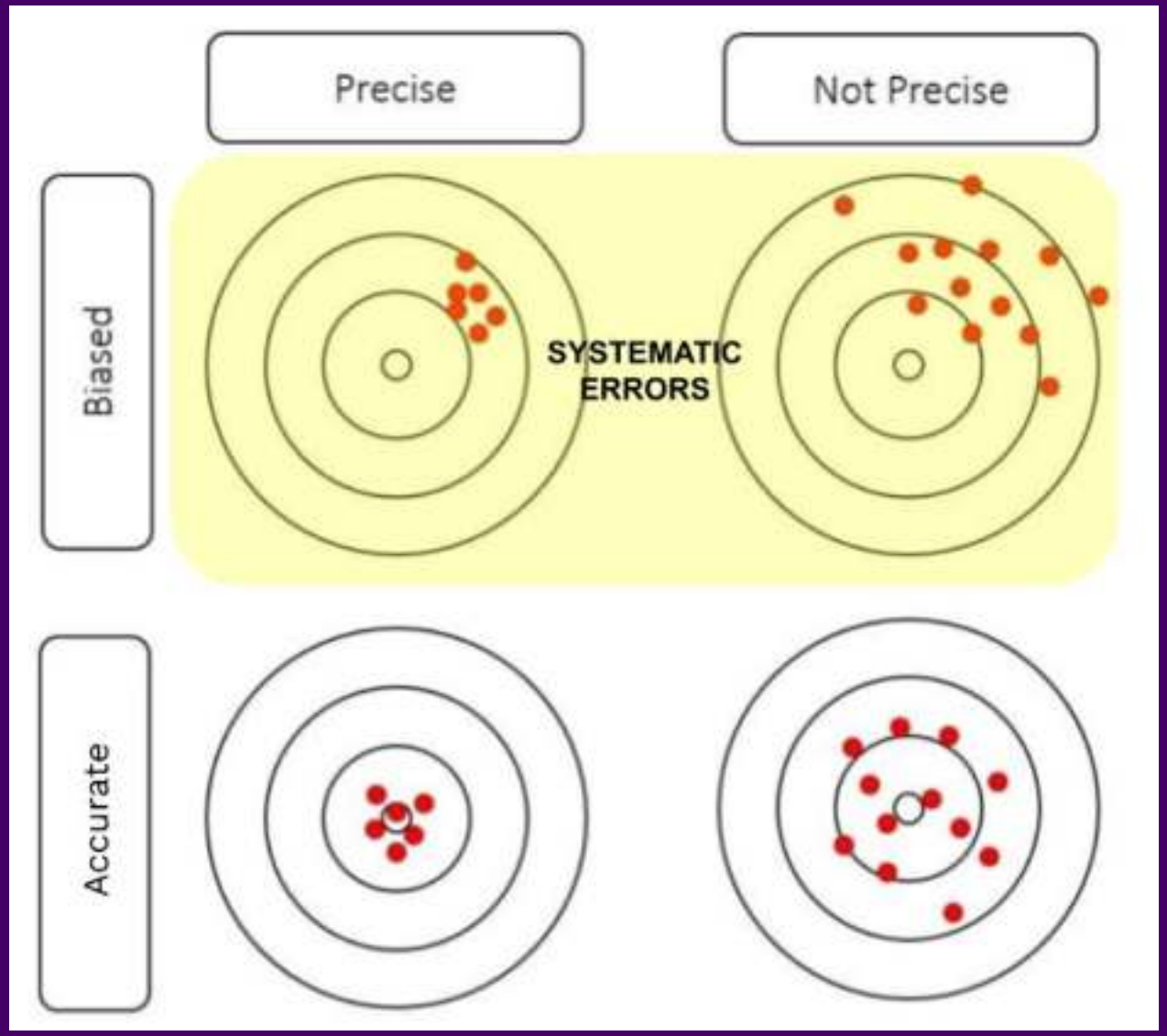


Fallacies

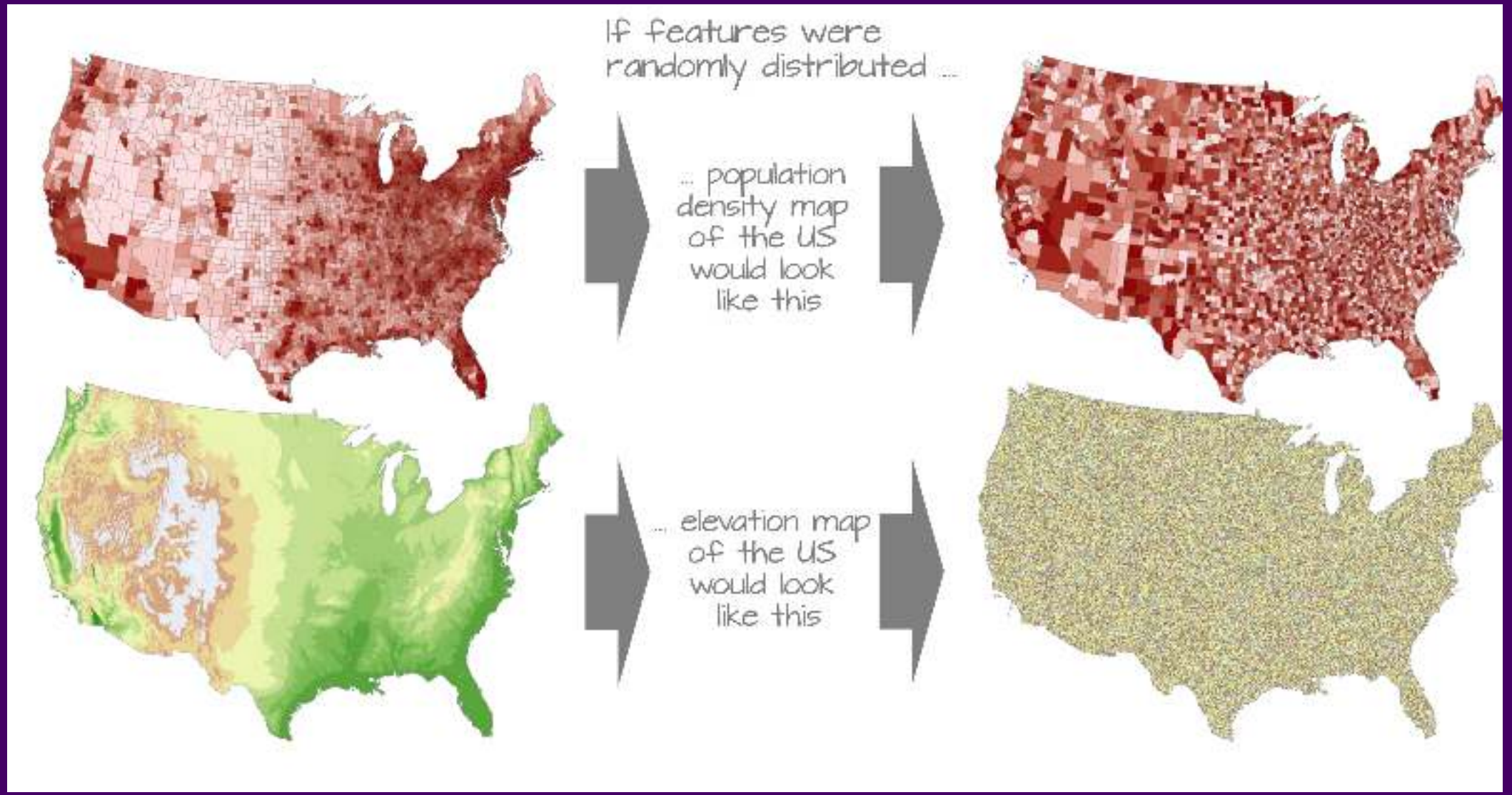
- **Locational Fallacy:** Error due to the spatial characterization chosen for elements of study
- **Atomic Fallacy:** Applying conclusions from individuals to entire spatial units
- **Ecological Fallacy:** Applying conclusions from aggregated information to individuals



Measurement Error and Mismatch



Spatial Autocorrelation

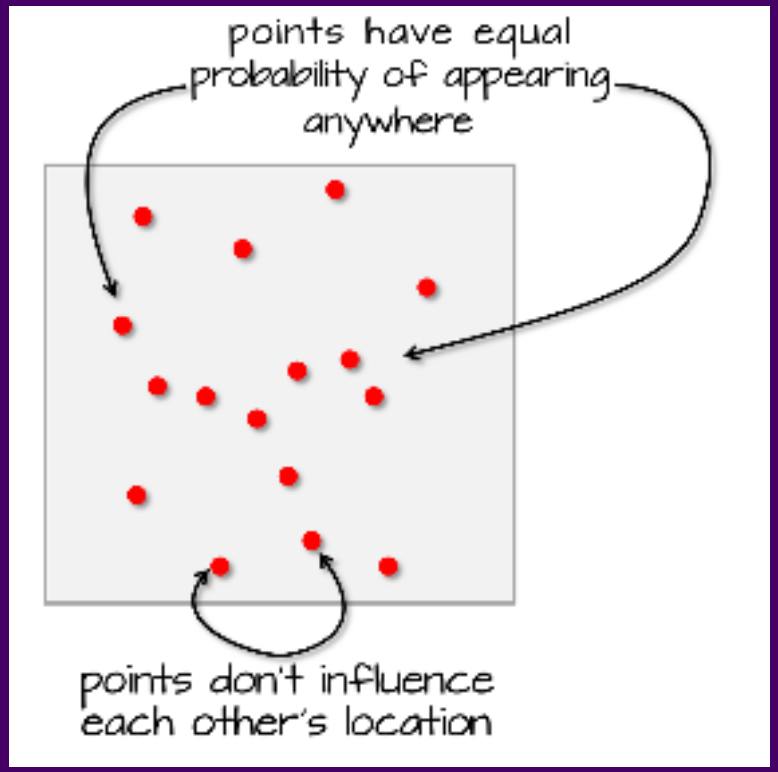


From Manuel Gimond

Stationarity

The rules governing a process do not *drift* over space-time

- **First Order** effects: any event has an equal probability of occurring in a location
- **Second Order** effects: the location of one event is independent of the other events



From Manuel Gimond

Key Critiques

Not all geography needs to be quantitative

1. Abstraction removes the interesting part
2. What “is” may require assumptions we don’t want to accept
3. Wholly dependent on the military-industrial complex

Wrapping Up

1. Themes in geography
2. Description, explanation, prediction
3. Key challenges and critiques