

# Homophily and Segregation



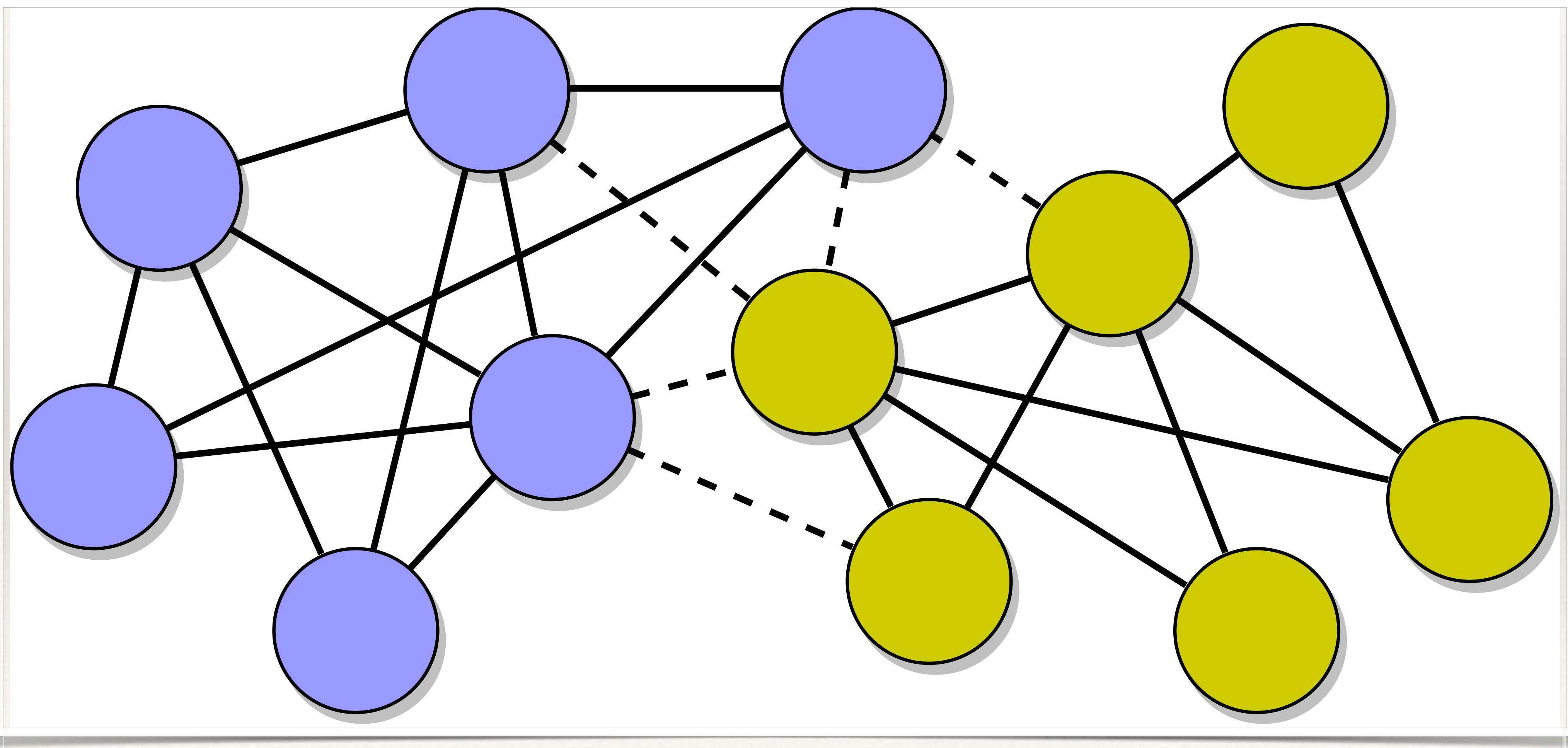
---

# Homophily

---

- ❖ The principle that we tend to be similar to our friends
- ❖ This makes your friends not statistically significant as a random sample of the population
- ❖ Similarities:
  - ❖ immutable characteristics
  - ❖ mutable characteristics

# "Birds of a feather flock together"









fraction of white nodes:  $p = \frac{2}{3}$

fraction of pink nodes:  $q = \frac{1}{3}$



$$p \cdot p = p^2 = \frac{4}{9}$$

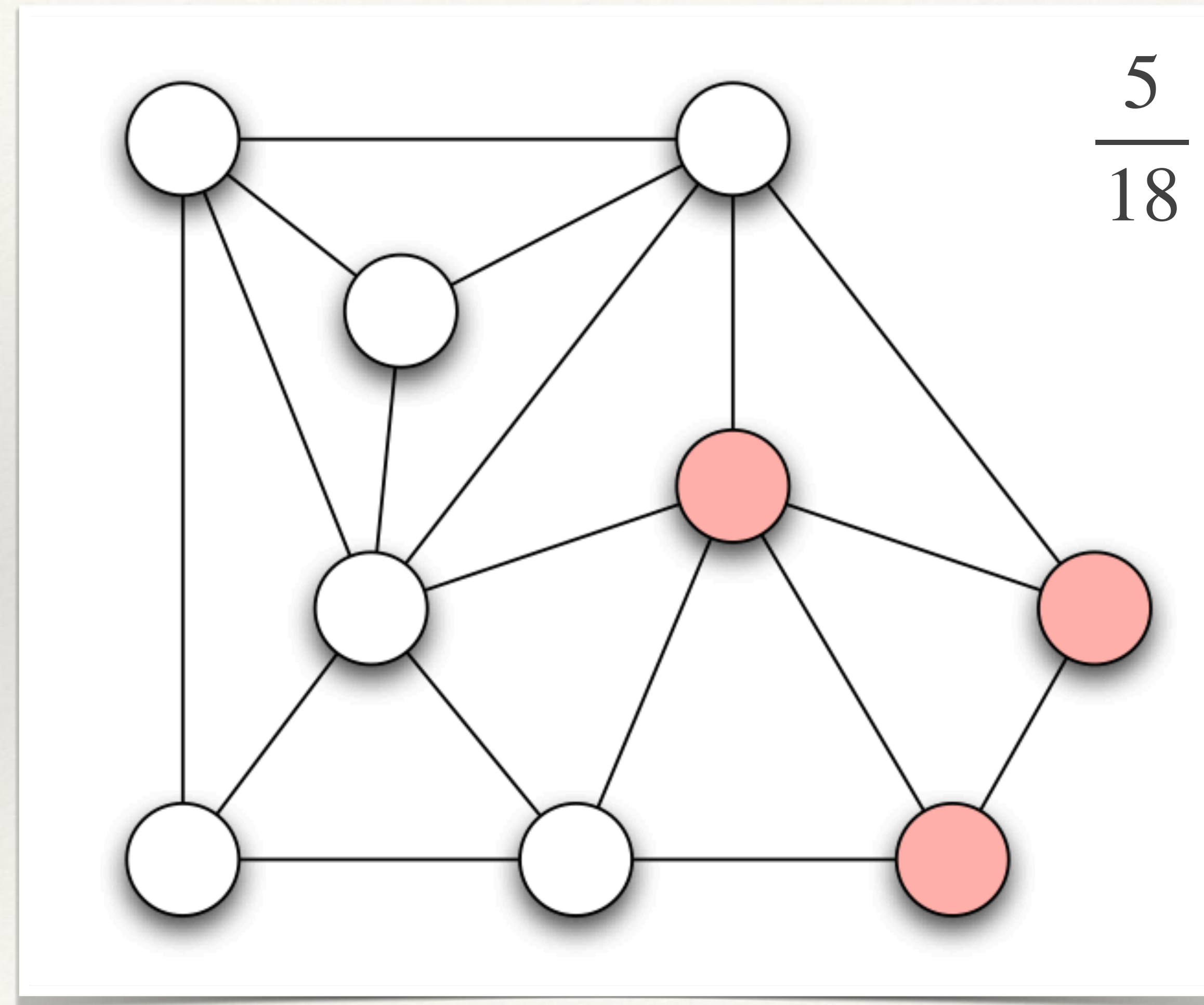


$$q \cdot q = q^2 = \frac{1}{9}$$



$$2 \cdot p \cdot q = 2 \frac{2}{3} \frac{1}{3} = \frac{4}{9}$$

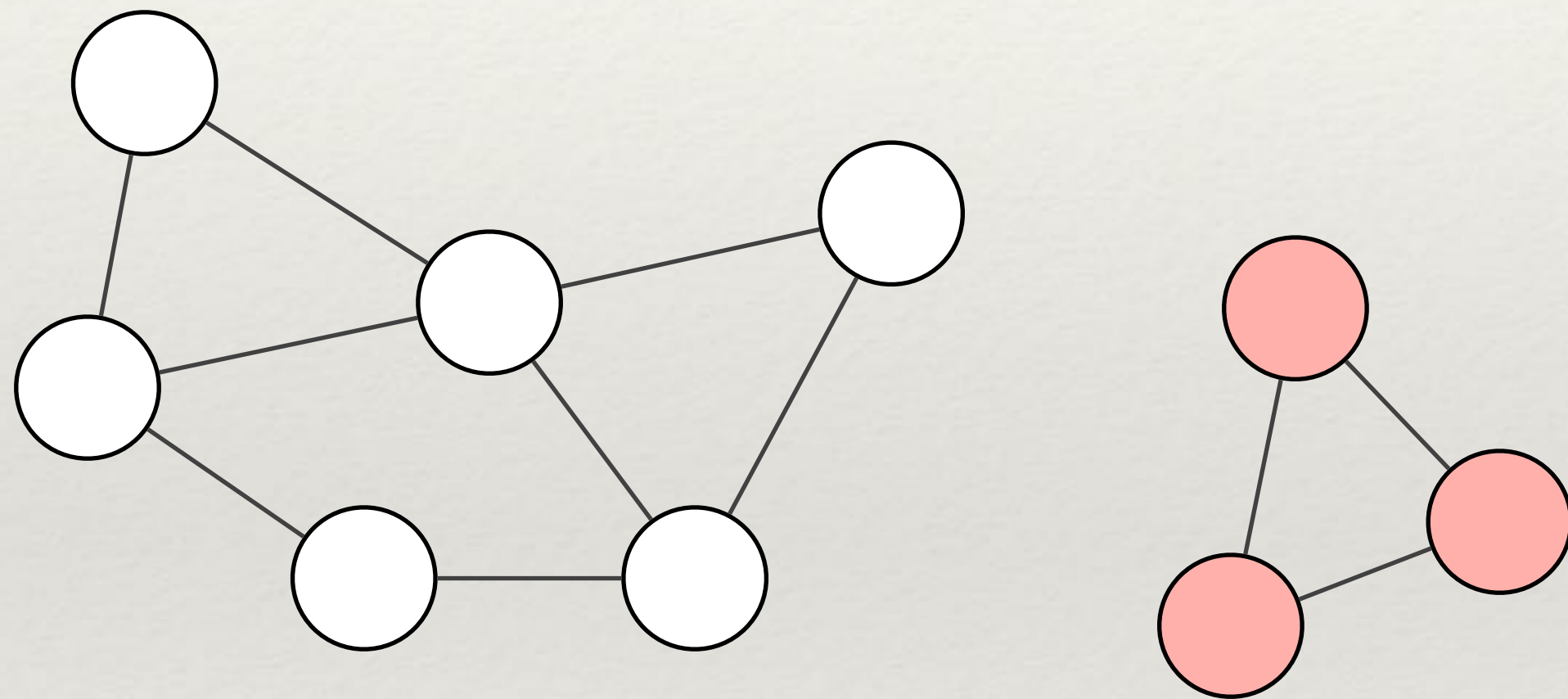
$$\frac{5}{18} < \frac{4}{9}$$





**homophily test:** check if *# actual cross groups edges*  $< 2pq$

$$\frac{5}{18} < \frac{4}{9} \Rightarrow \text{homophily!}$$



perfect homophily:  $0 < \frac{4}{9}$

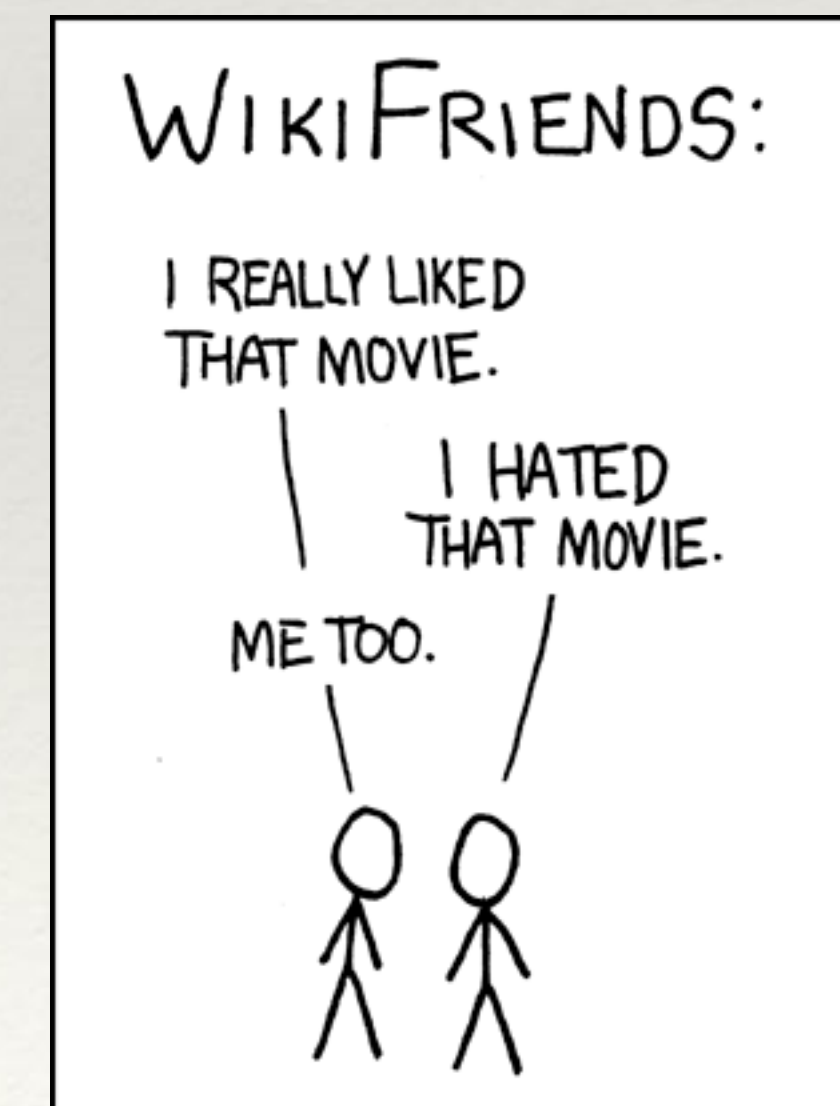
More precisely

**homophily test:** if the fraction of cross-types edges is *significantly less* than  $2pq$ , then there is a *signal* of homophily



# Underlying mechanisms of homophily

- ❖ Two possible mechanisms by which homophily (also: assortativity) emerges naturally:
  1. **Selection:** similar nodes become connected
  2. (Social) **influence:** connected nodes become more





---

# The interplay of selection and social influence

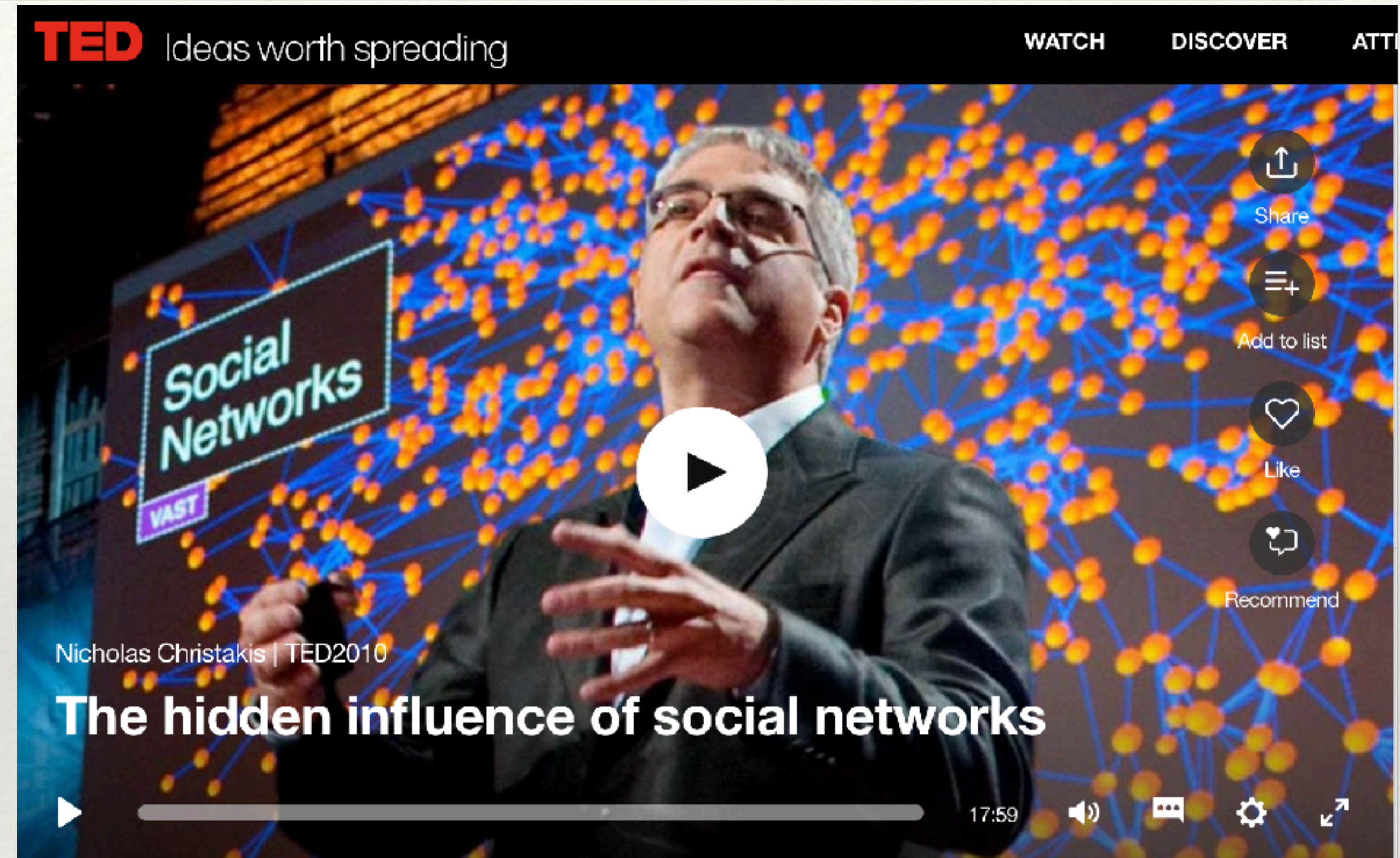
---

- ❖ longitudinal methodology:
  - ❖ observe a network for a long period of time
  - ❖ observe both factors in action
  - ❖ how do we quantify the impact?
- ❖ example: obesity as a social contagion phenomenon



# obesity "contagion"

- ❖ dataset: 12,000 people
- ❖ obesity status
- ❖ social network structure
- ❖ obese vs non obese: there is a tendency toward clustering
- ❖ homophily test: passed
- ❖ why?
  - ❖ selection?
  - ❖ homophily that correlates with something else?
  - ❖ social influence? —> contagion!

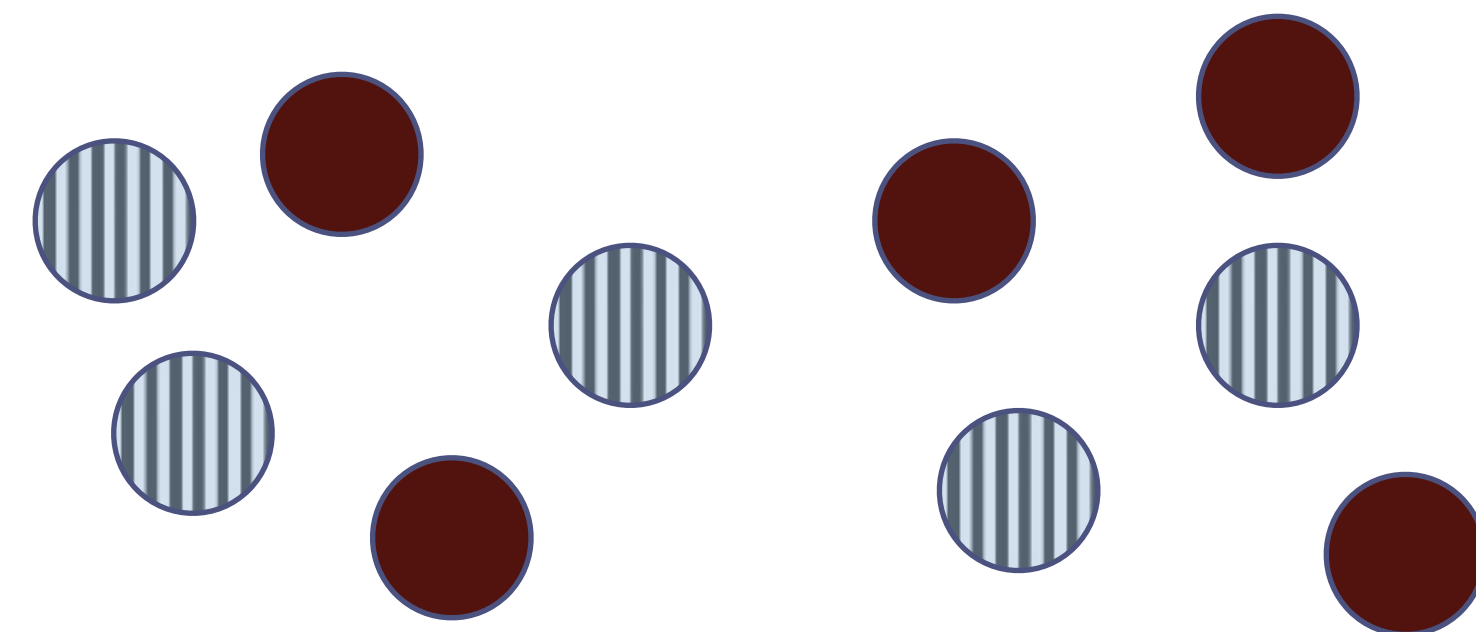
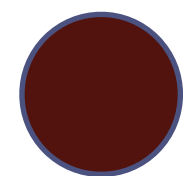
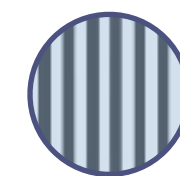


[https://www.ted.com/talks/nicholas\\_christakis\\_the\\_hidden\\_influence\\_of\\_social\\_networks](https://www.ted.com/talks/nicholas_christakis_the_hidden_influence_of_social_networks)



# The emergence of segregation

- ❖ Society's structure is shaped in function of **immutable characteristics** of individuals
  - ❖ ethnic group
  - ❖ age
  - ❖ religious belief
  - ❖ ...



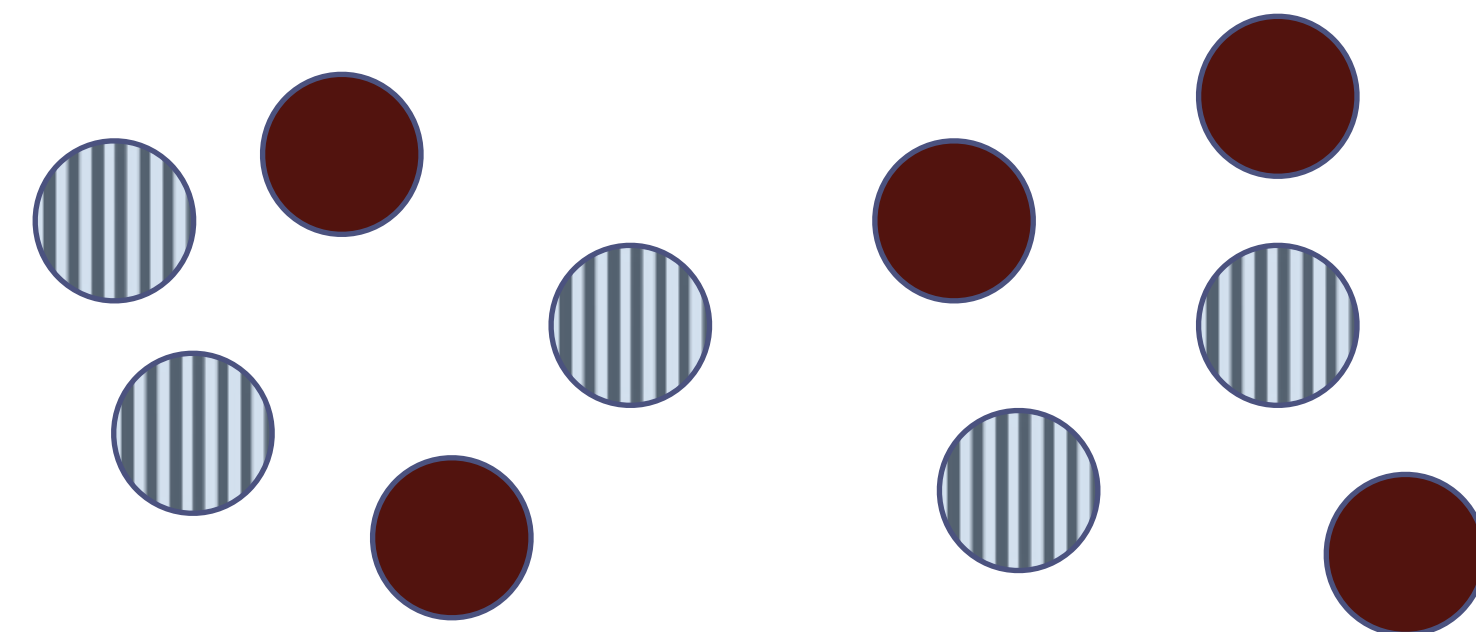
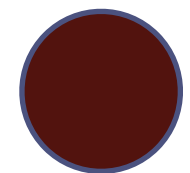
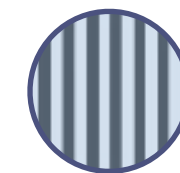


---

# Segregation

---

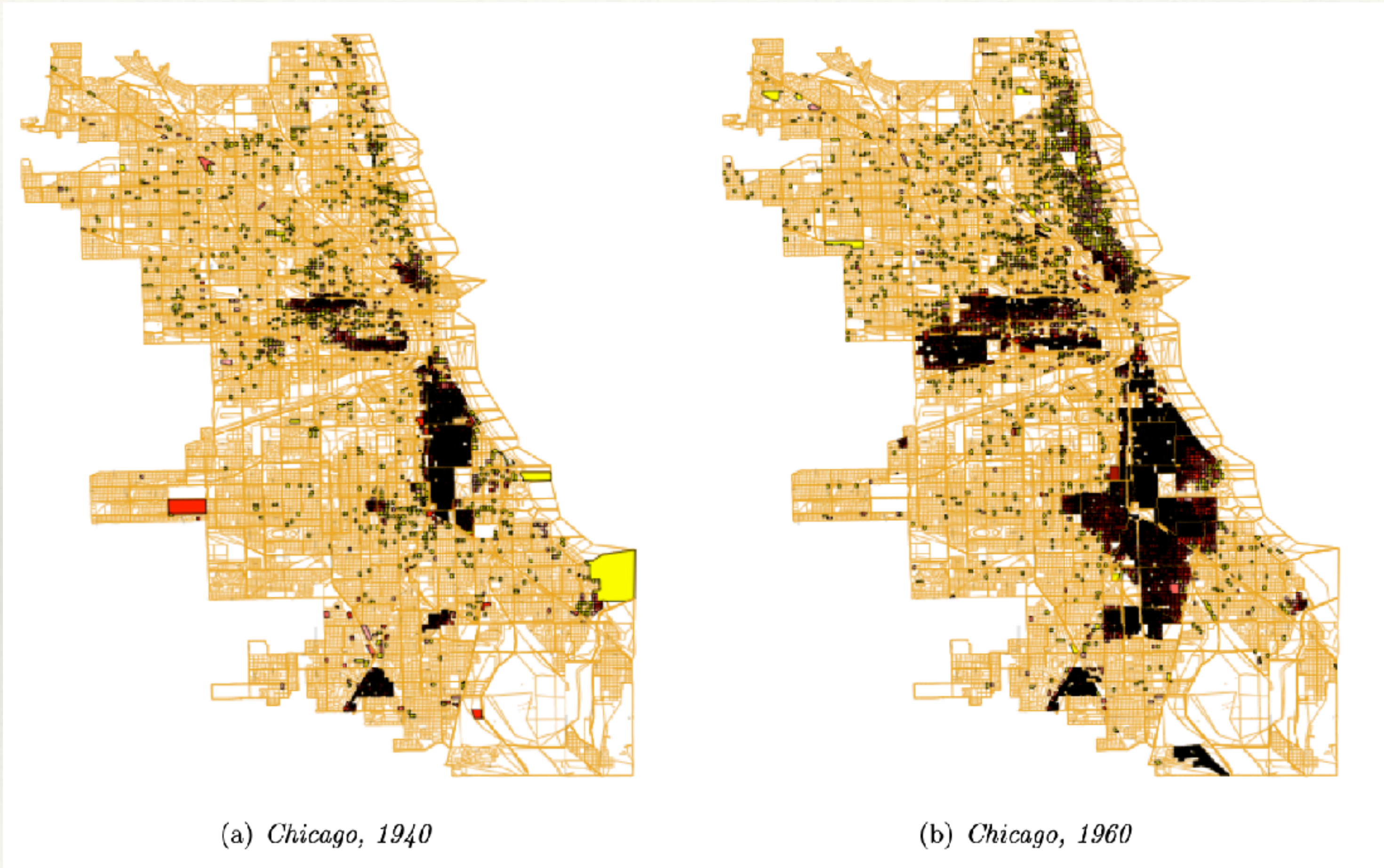
- ❖ Society's structure is shaped in function of **immutable characteristics** of individuals
  - ❖ ethnic group
  - ❖ age
  - ❖ religious belief
  - ❖ ...






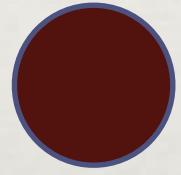
# Natural spatial "signature" in cities

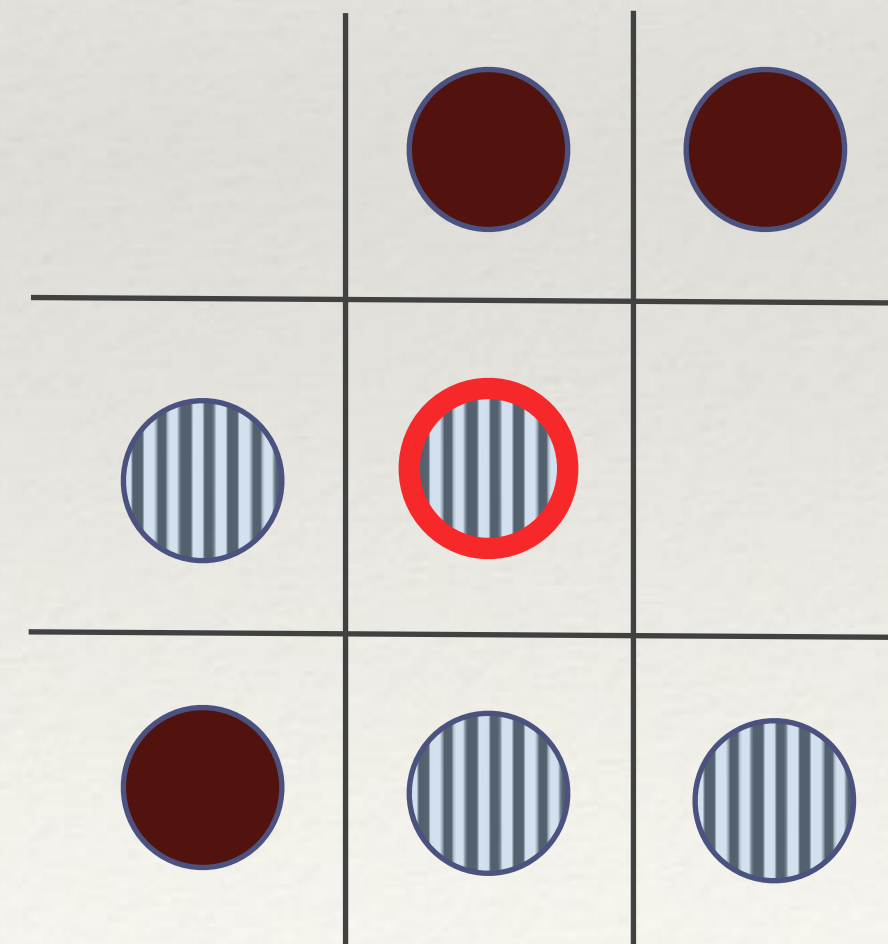
- ❖ Formation of homogeneous (according to some "type" or "class") neighbors in cities
- ❖ Which are the causes of "ghettization"?



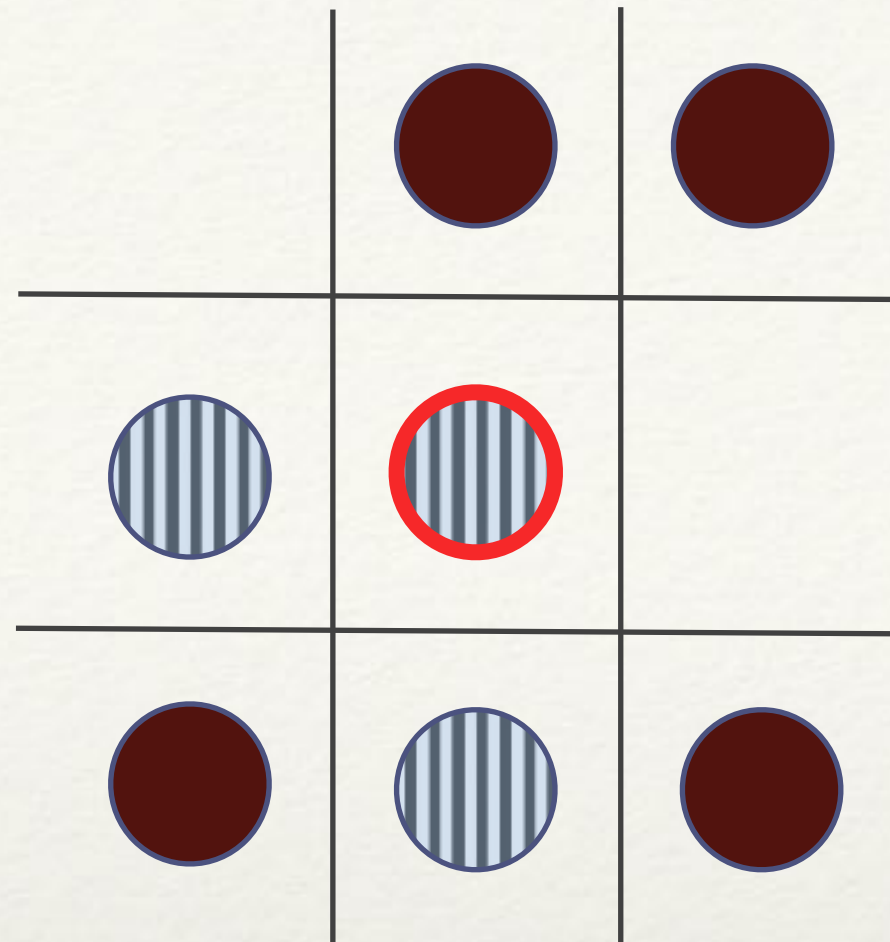


# The Schelling model

- ❖ Can spatial segregation arise from the effect of homophily operating at a local level?
- ❖ Assumption: no individual want segregation explicitly
- ❖ Agents:
  - ❖ two types:  
  - ❖ immutable characteristics
- ❖ Agents reside in a cell of a grid
  - ❖ some cells contain agents
  - ❖ some other cells are unpopulated
- ❖ Neighbors: 8 other cells "touching" an agent







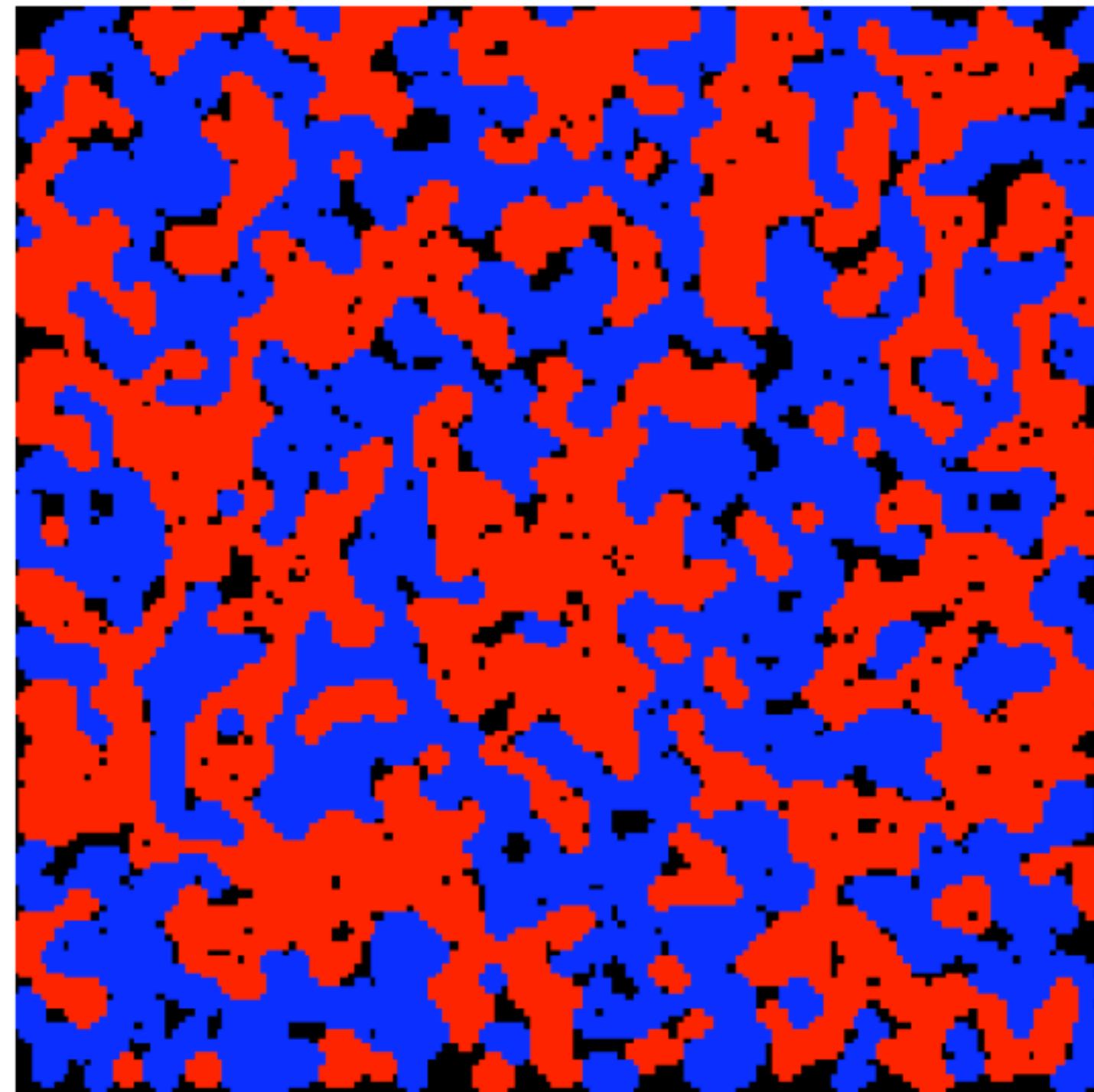
$t = 3 \Rightarrow :-()$

- ❖ Each agent wants to have at least  $t$  neighbors of their own type
- ❖ If an agent find  $< t$  neighbors of the same type, then they are **unsatisfied**
- ❖ If unsatisfied, they want to **move**

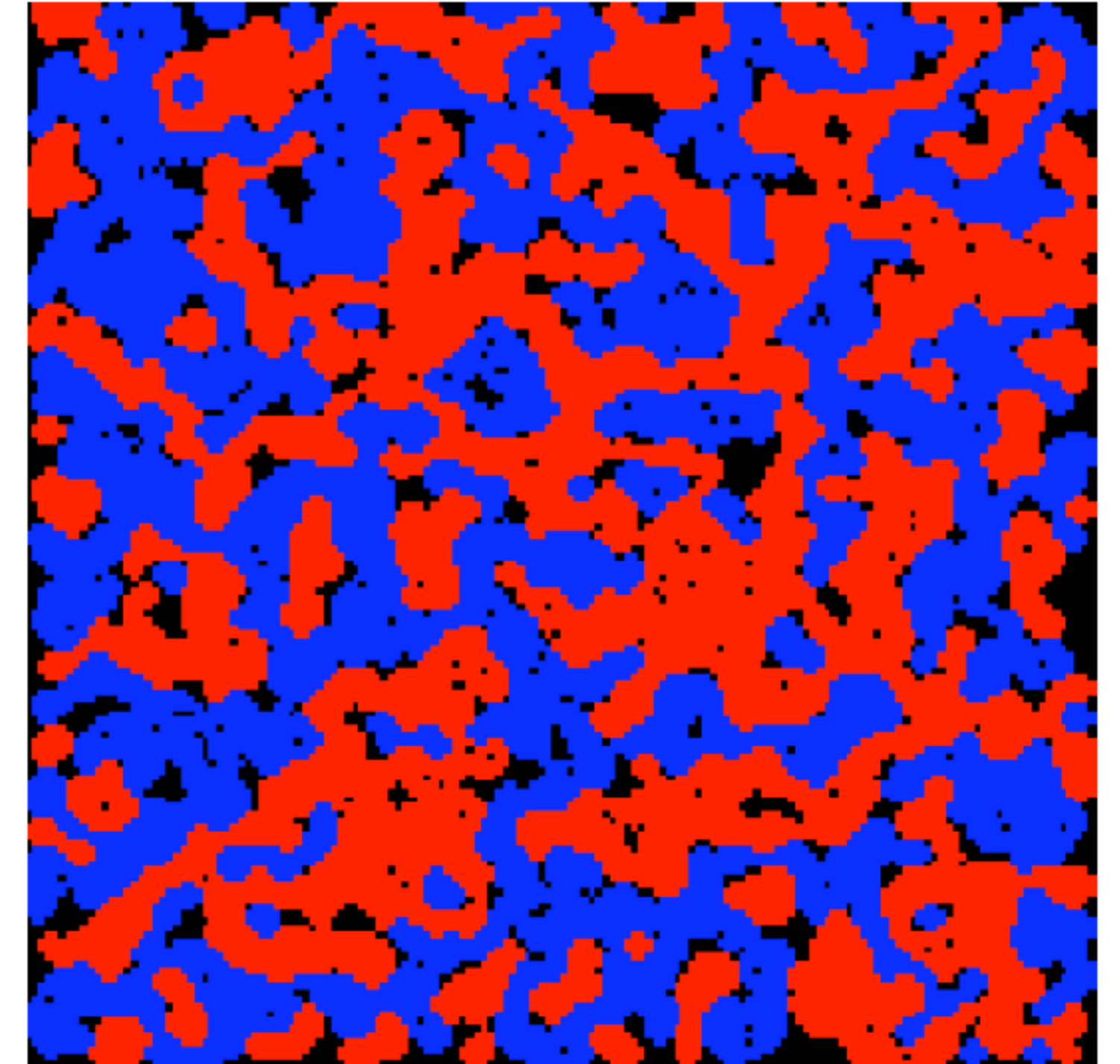


# Larger examples

- ❖ Computer simulations to look for patterns at larger scale
- ❖ We want to run different simulations and make some comparisons  
=> integrated pattern?
- ❖ on the right: two runs of a simulation of the Schelling model with a threshold  $t$  of 3
  - ❖ 150x150 grid
  - ❖ 10,000 agents



(a) A simulation with threshold 3.



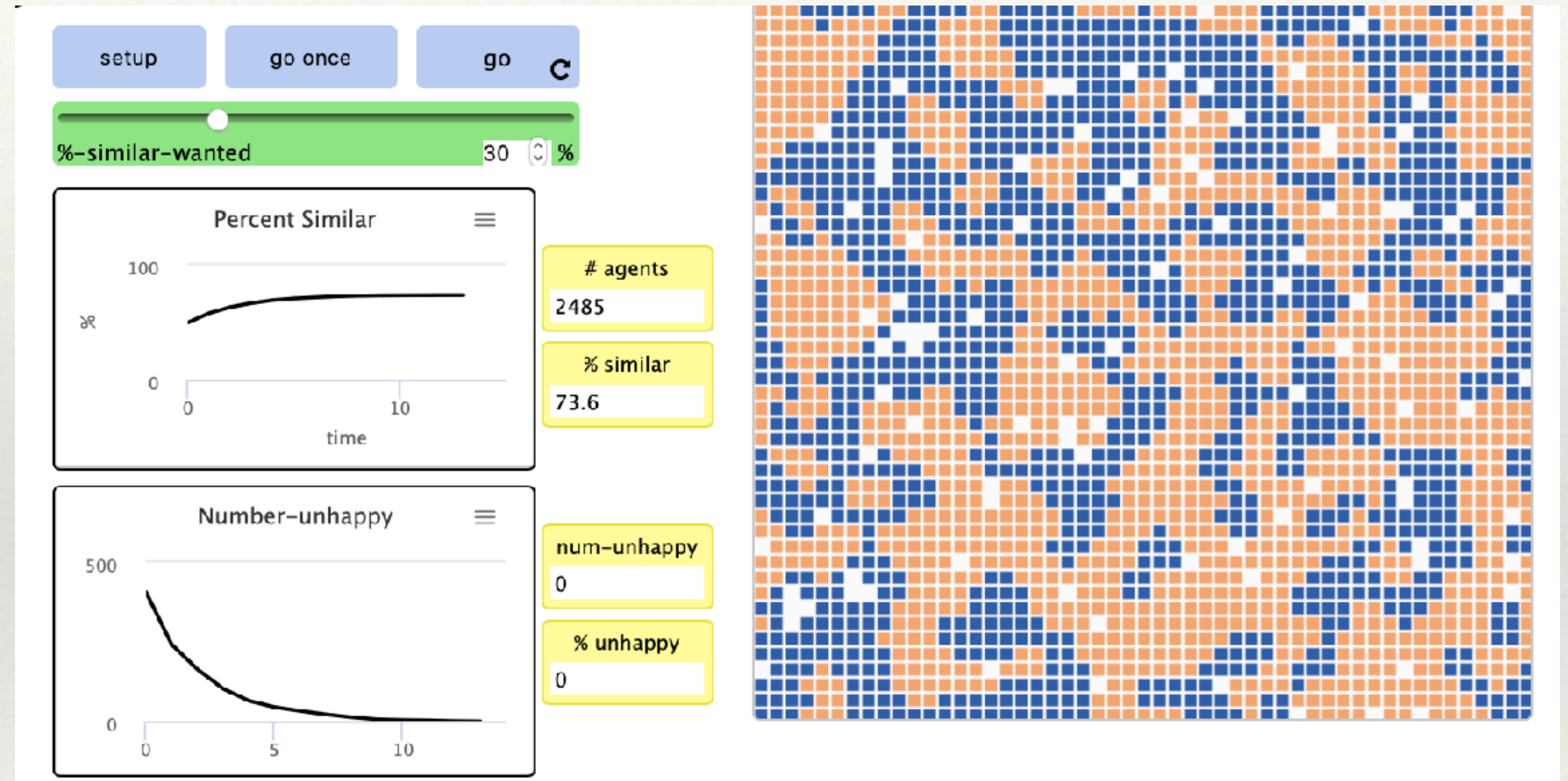
(b) Another simulation with threshold 3.

**Segregation emerges even when agents accept to be a minority!**



# Example with NetLogo

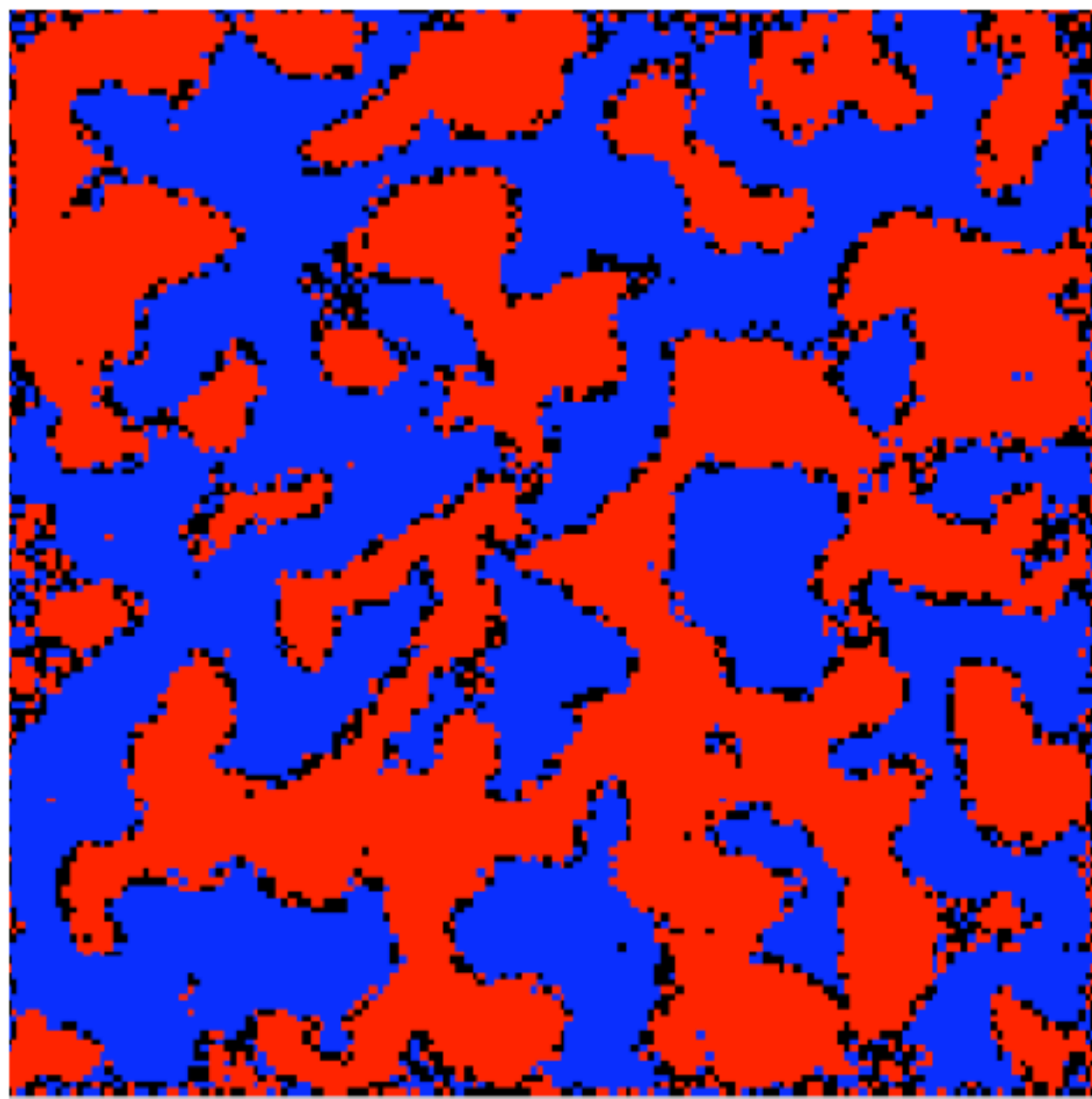
## Agent based simulations



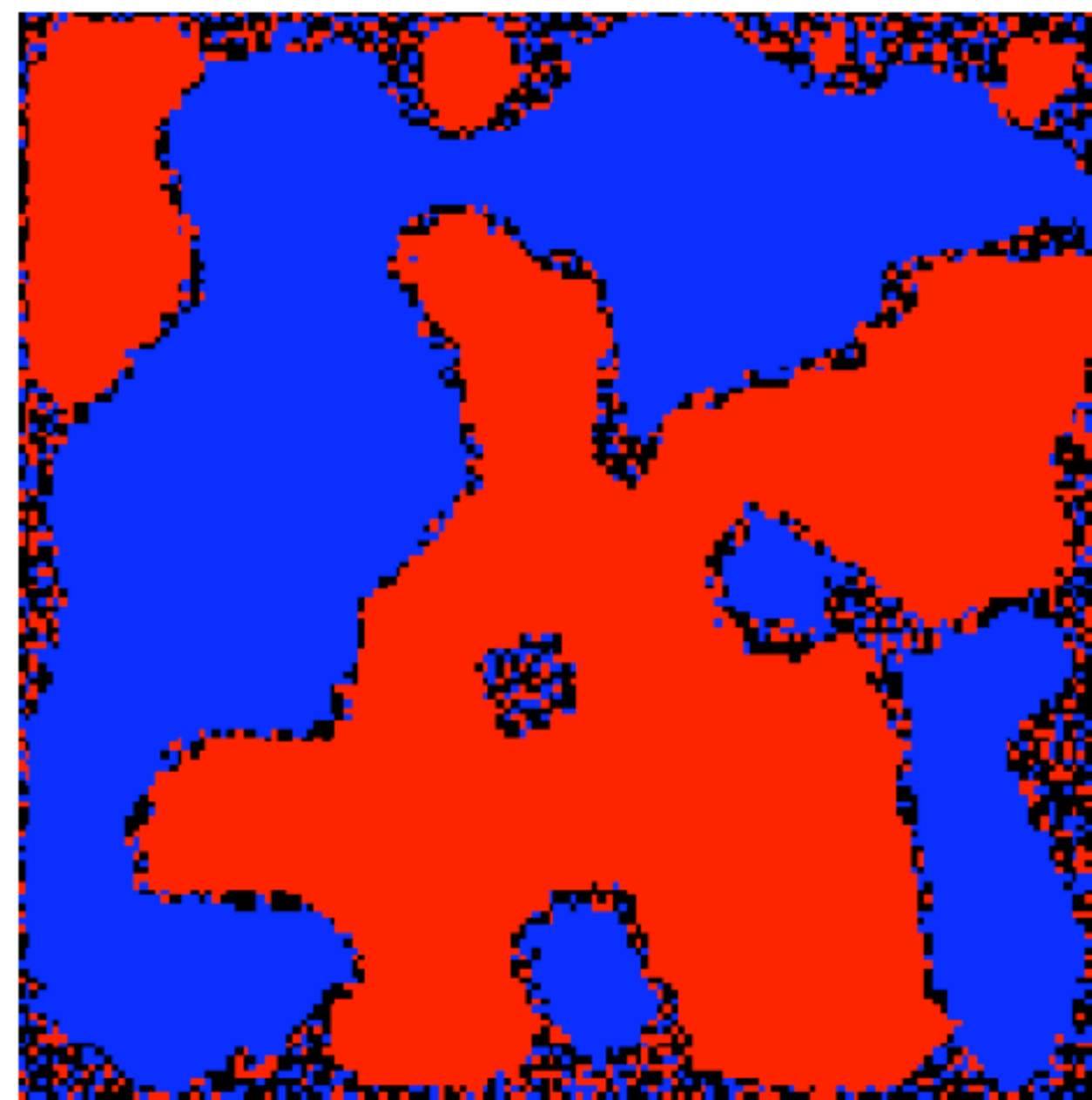
<http://www.netlogoweb.org/launch#http://www.netlogoweb.org/assets/modelslib/Sample%20Models/Social%20Science/Segregation.nlogo>



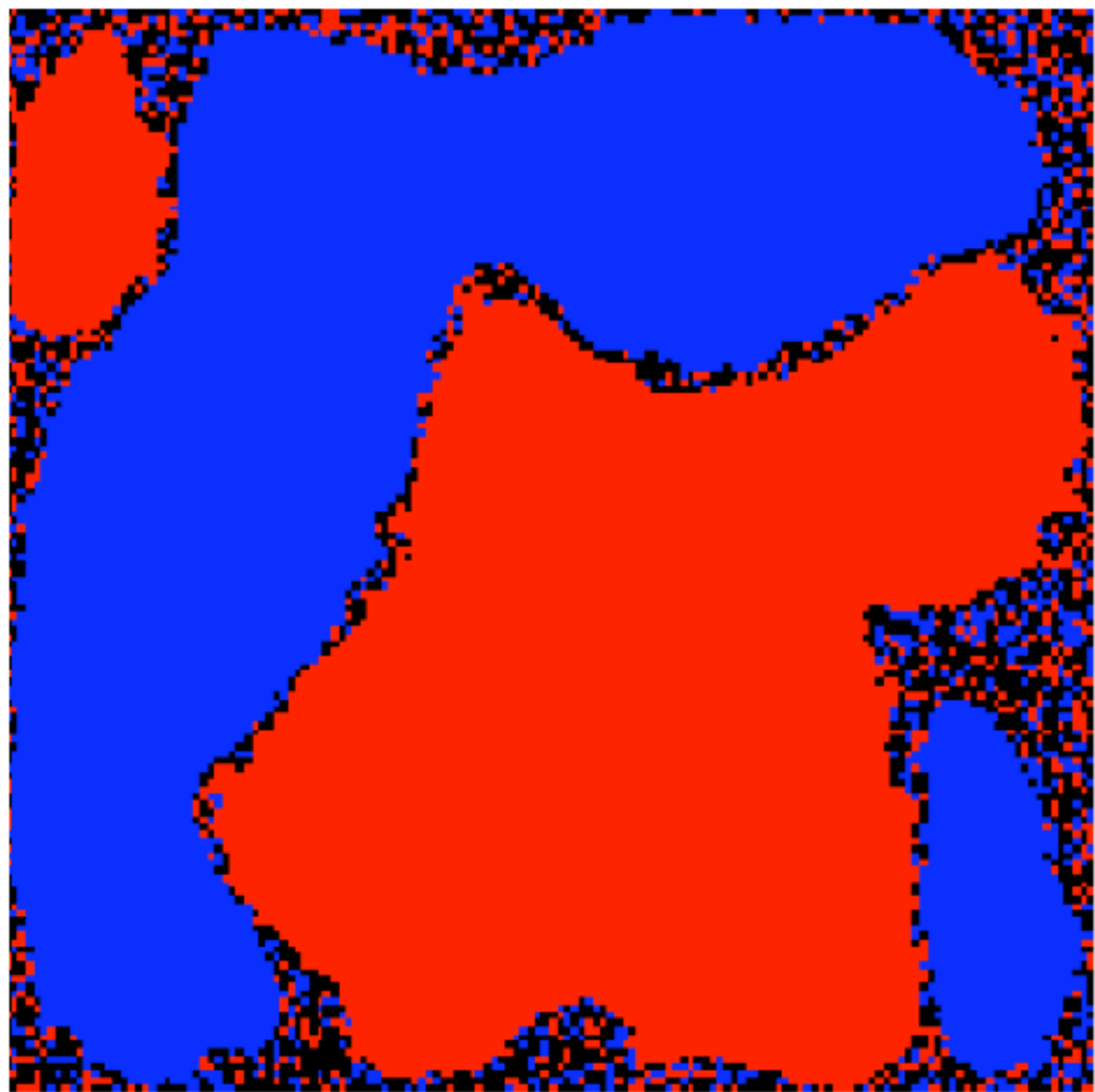
$t > 3 \Rightarrow$



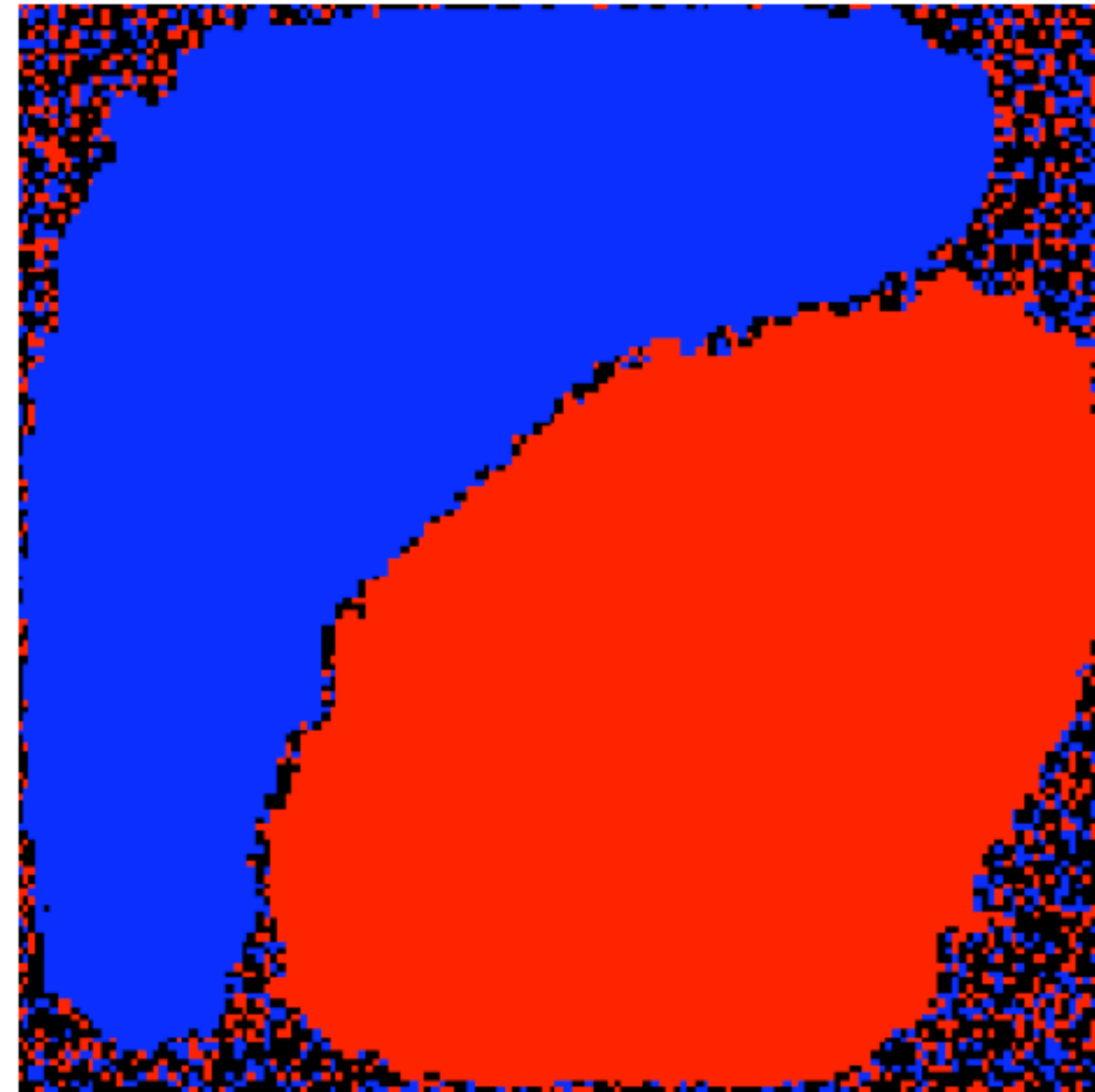
(a) *After 20 steps*



(b) *After 150 steps*



(c) *After 350 steps*



(d) *After 800 steps*

**Segregation is  
(trivially) amplified in  
an intolerant society**



---

# Impacts of segregation

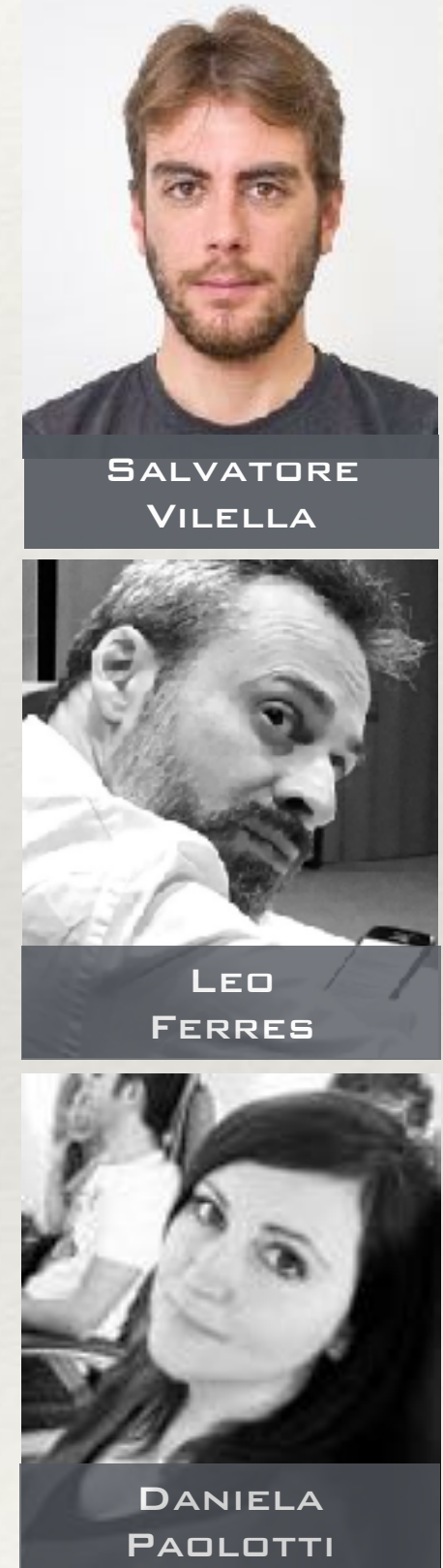
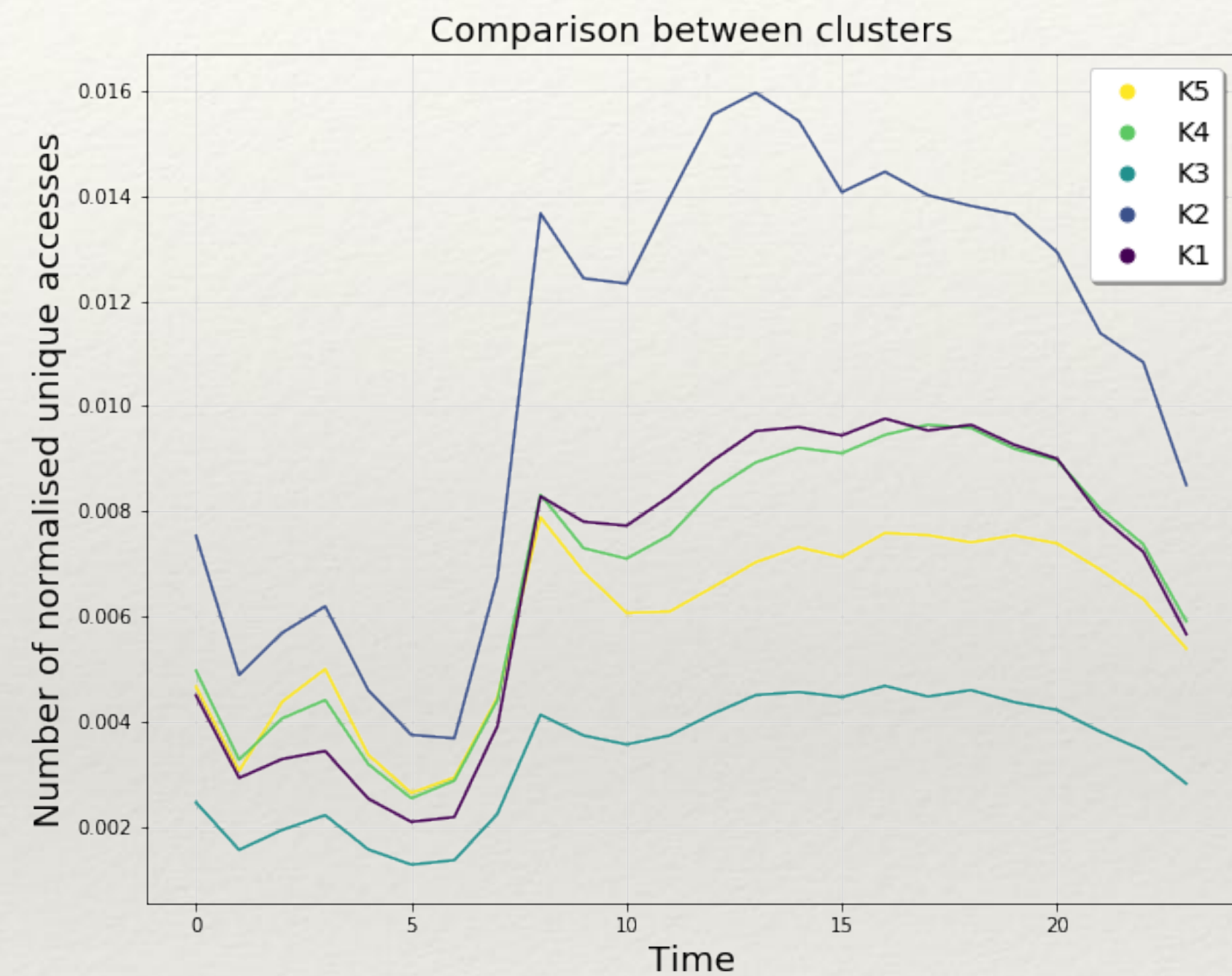
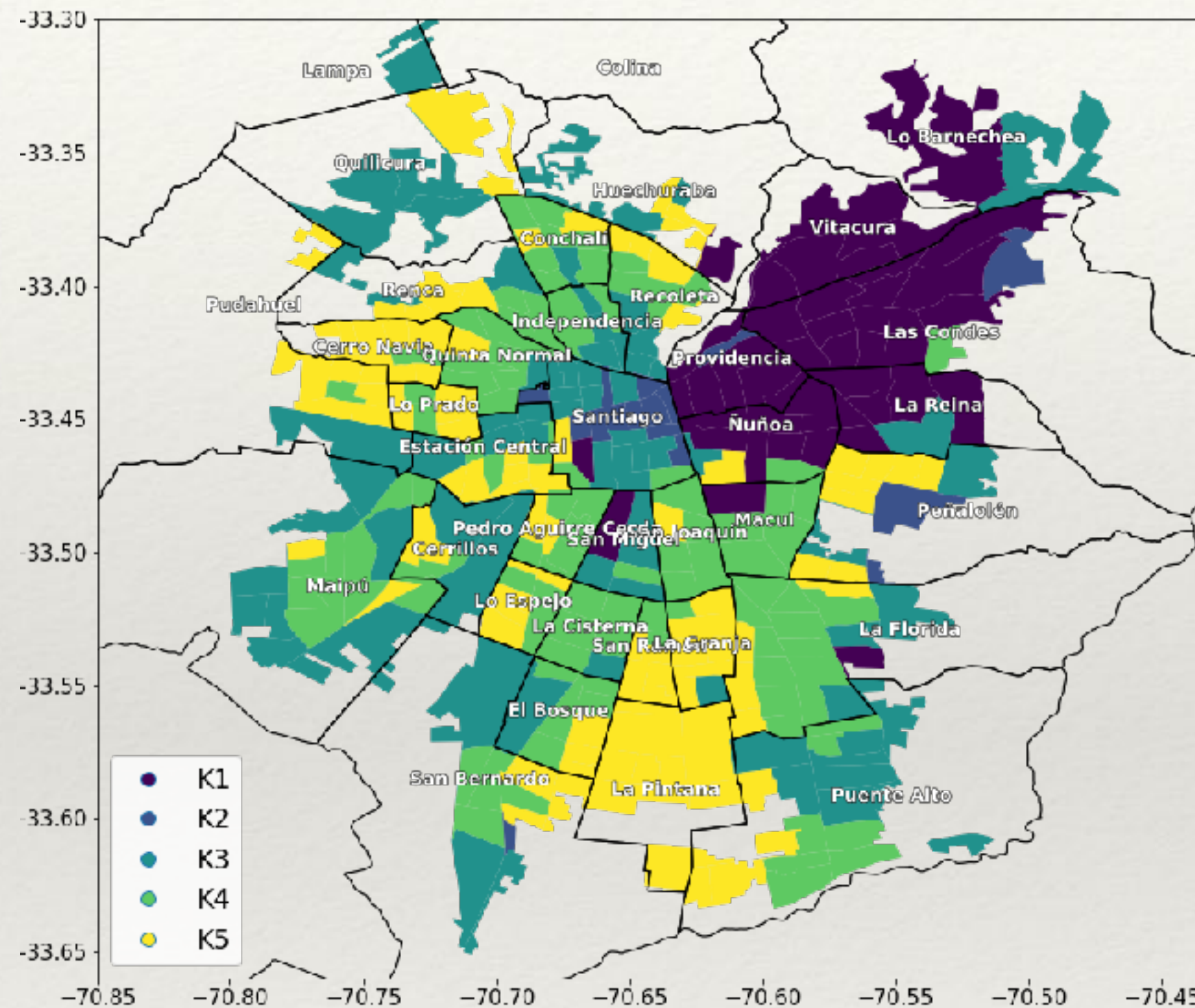
---

- ❖ Let's accept that segregation emerges naturally even in the most tolerant society (unless we do not design our 'societies' properly)
- ❖ Segregation has consequences (not necessarily bad...)
- ❖ Examples:
  - ❖ on news consumption
  - ❖ on outbreaks diffusion



# Segregation vs information consumption

Study of geo-located accesses to websites of news media revealed strong differences between different “classes” of the population of SCL.



Vilella, S., Paolotti, D., Ruffo, G. and Ferrer, L.. [News and the city: understanding online press consumption patterns through mobile data](#). EPJ Data Sci. 9, 10 (2020)



# Segregation by age and virus transmission



Crowds take in the the cherry blossoms as visitors from holding *sakura*-viewing parties.

COMMENTARY / JAPAN

## Why is Japan still a coronavirus hot spot?

BY OSCAR BOYD  
STAFF WRITER

At the time of writing, Japan has just recorded its first coronavirus case. That's 900 cases recorded in the first person — a man who had traveled from Italy and returned to Japan on June 23. Shortly after, 50,000 people were quarantined in a handful of towns in

In Italy, the first case was recorded there on March 23. Shortly after, 50,000 people were quarantined in a handful of towns in



*hypothesis not supported by scientific evidences, yet!*

CLICK TO ENLARGE

thought: that Japan is spreading in the way it has: relatively less social to wear masks when us, already high e voluntary self- at Japan is flattening



# Dynamical Processes in Information Networks



---

# Overview of network dynamics

---

- ❖ Social contagion
- ❖ Emergence of polarization
- ❖ Consequences: confirmation biases, echo chambers
- ❖ Intro to epidemic spreading
- ❖ Impact of verification and fact checking: SBFC model and what-if analysis



# Emergence of polarization



*“Polarization is both a state and a process. Polarization as a **state** refers to the extent to which opinions on an issue are opposed in relation to some theoretical maximum. Polarization as a **process** refers to the increase in such opposition over time.”*

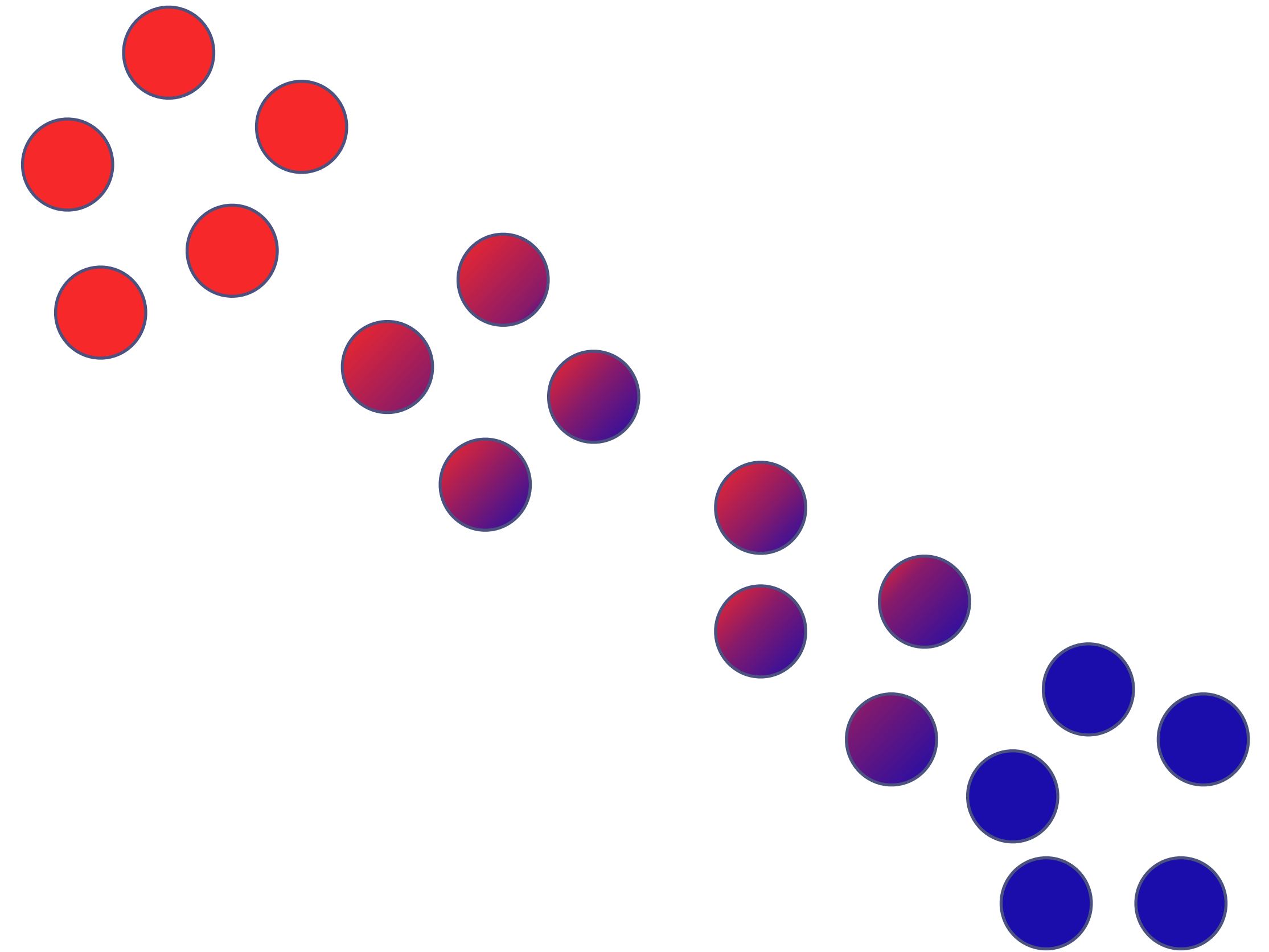
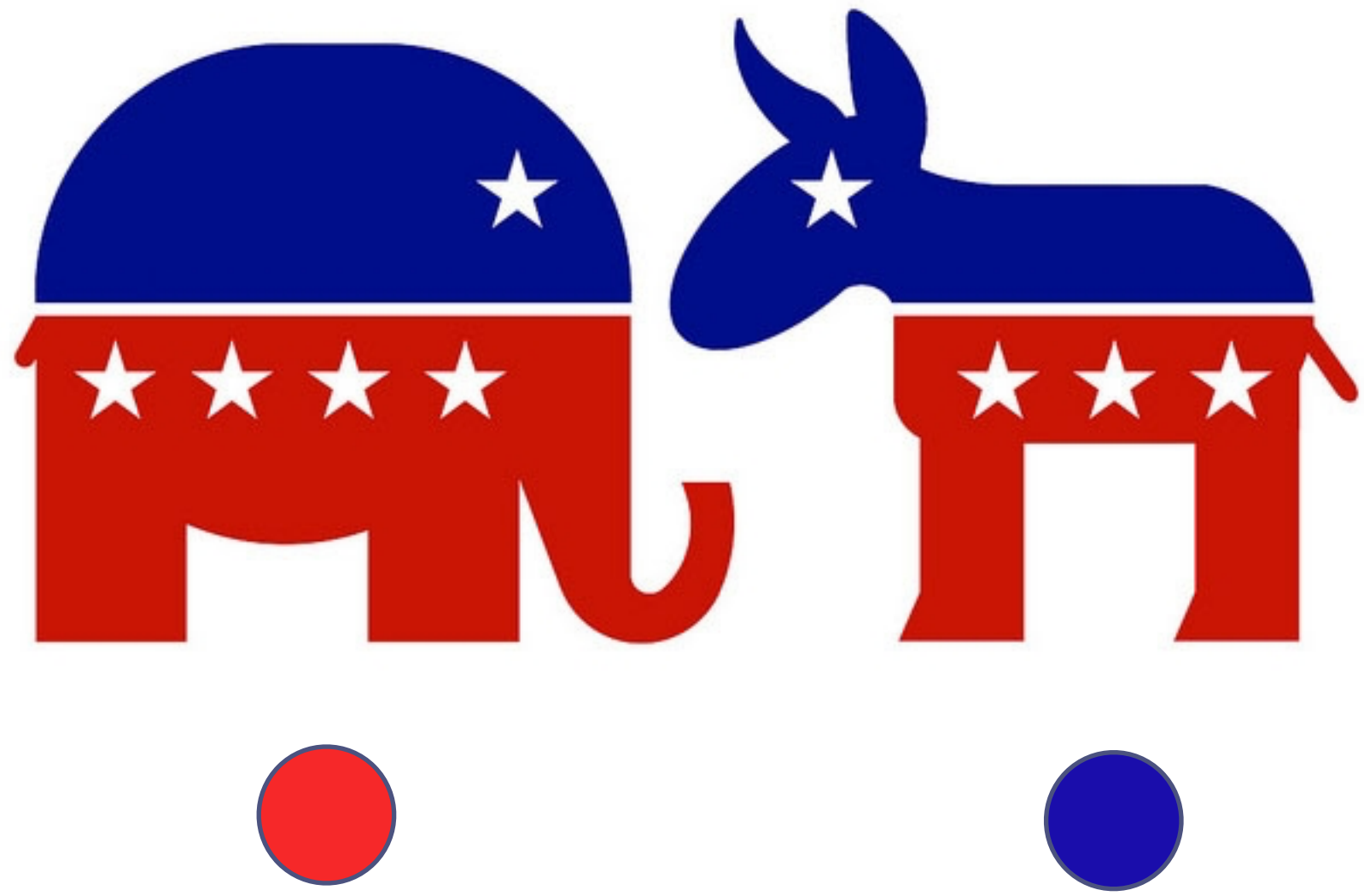
*– DiMaggio et. al, American Journal of Sociology, 1996*



---

# Polarization

---

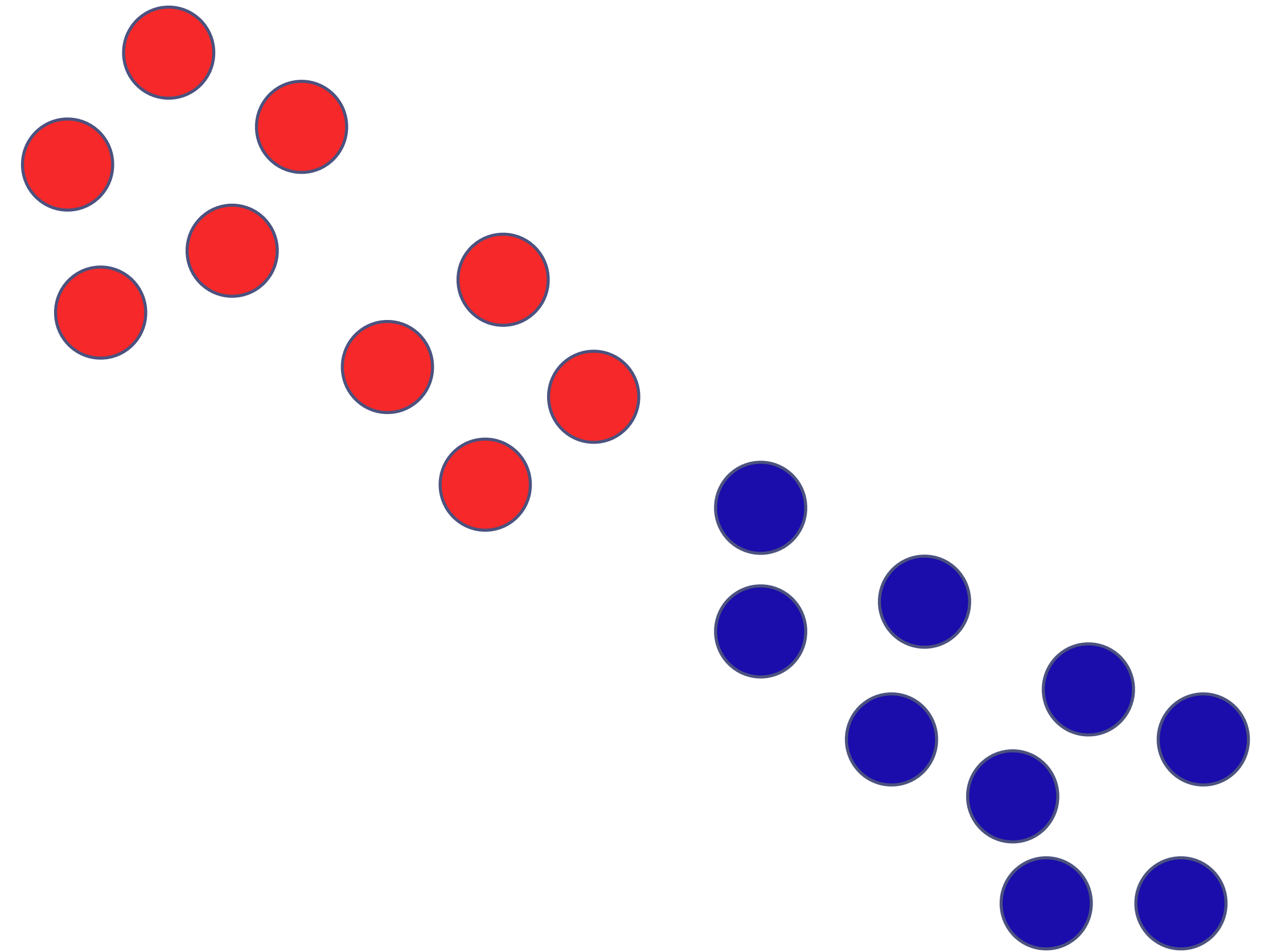
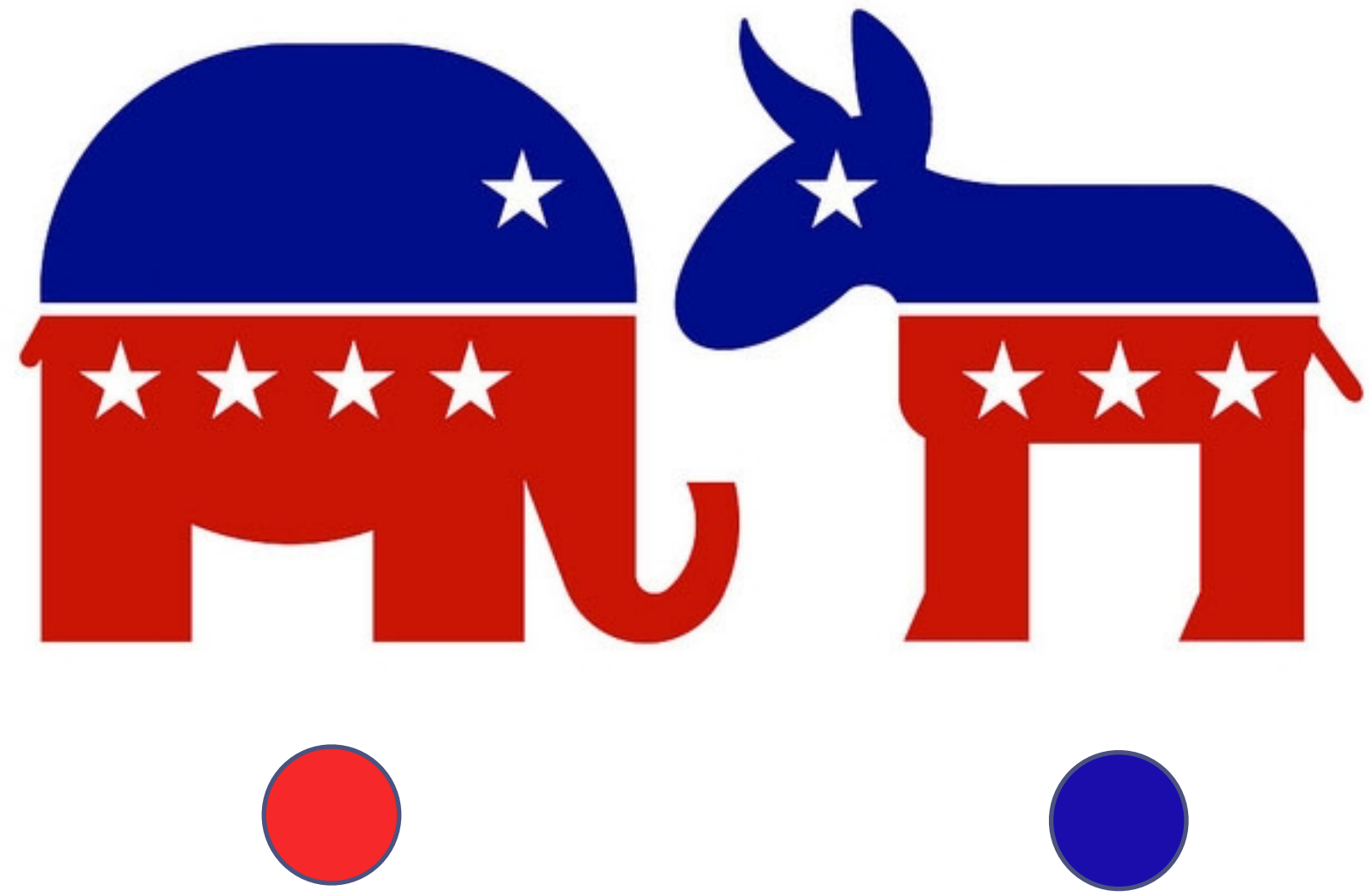




---

# Polarization

---





---

# Issues with studying polarization

---

- ❖ **State:** difficult to detect
  - ❖ e.g., NLP based techniques as "*stance detection*" are great, but errors prone
- ❖ **Process:** difficult to observe
  - ❖ e.g., opinions can mitigate or polarize over time, but people do not necessarily express them
- ❖ Polarization by **selection** and by **influence**
  - ❖ do I get along with people that share my opinion, or I am influenced by people with whom I get along? or both processes are at interplay?
- ❖ "**Social contagion**" is more rational than we may think...