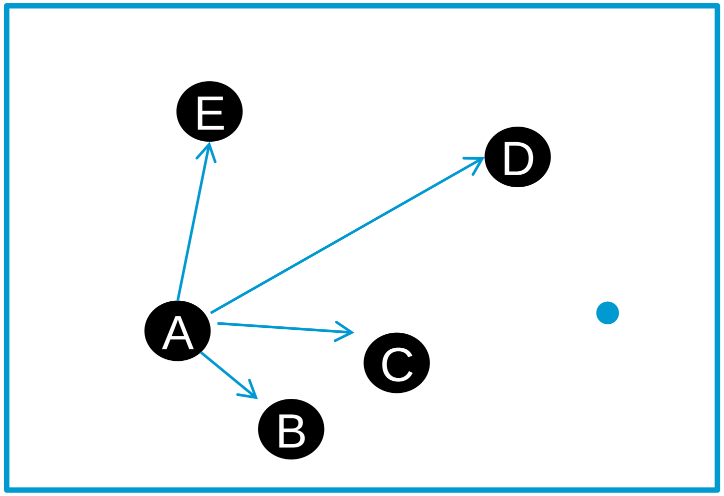
„Presentation and Analysis of Spatial Data“

(5) Analysis of spatial patterns



Thomas Wöhler, Universität Konstanz

Kiev, October, 2016

Agenda

- Motivation of spatial analyses
- Distances between observational units
- Examples of spatial methods
- Analysis of spatial dependence

Motivation: Potentials for the Social Sciences

The potential of better illustrations.

The potential of new data (more detailled and more objective).

The potential of more adequate statistical analyses.

The potential of new research questions.
On the spatial dimension of social processes. On mobility. On processes of diffusion. On the differences betwenn objective and subjective live conditions.

…

Motivation: Potentials for the Social Sciences

In what ways does space determine or influence social phenomena?

- **Actors influence each other and (spatial) distance reflects the potential.**
- **Tobler‘s first Law of geography is: „All things are related to each other, but near things are more strongly related.“**
- **In this regard can space influence social phenomena by several mechanisms:**
  - Contagion and imitations
  - Learning
  - Reference groups
  - Externalities
- **Thus it seems plausibel to collect spatial data and model the mechanisms.**

Motivation: Spatial analyses in the political sciences

556 O'Loughlin et al.

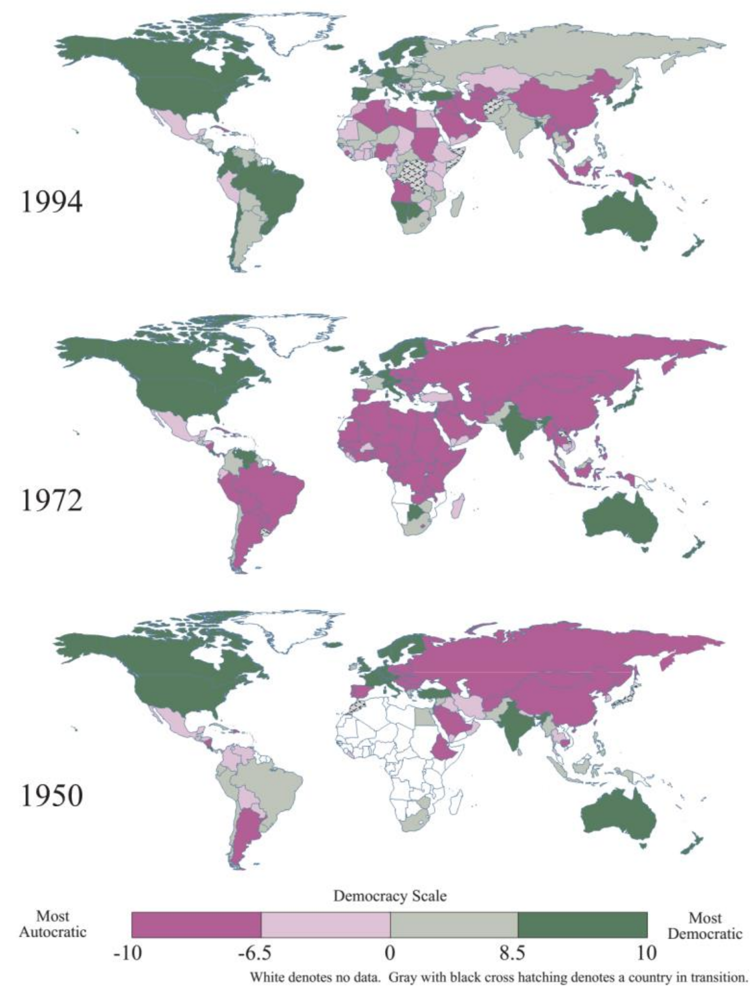


Figure 1. Geographic distribution of democracy scores, 1950, 1972, and 1994.

Motivation: Spatial analyses in economy

- **spatial econometrics**
- **Example: Housing (hedonic model)**

Abbildung 1. Schematische Darstellung zeitlicher und räumlicher Effekte.

Zeitreihe:	**Räumliche Daten:**
T1 → T2 → T3	P1
	↗↙ ↘↖
	P2 ↔ P3

Anmerkung: In Anlehung an Baller et al. (2001, S. 564)

Distances between observational units

Spatial analyses account for the relative position of all units with regard to all other units of observation, i.e. basis of all statistical method is a distance or adjacency matrix.

Spatial statistical analyses allow for:

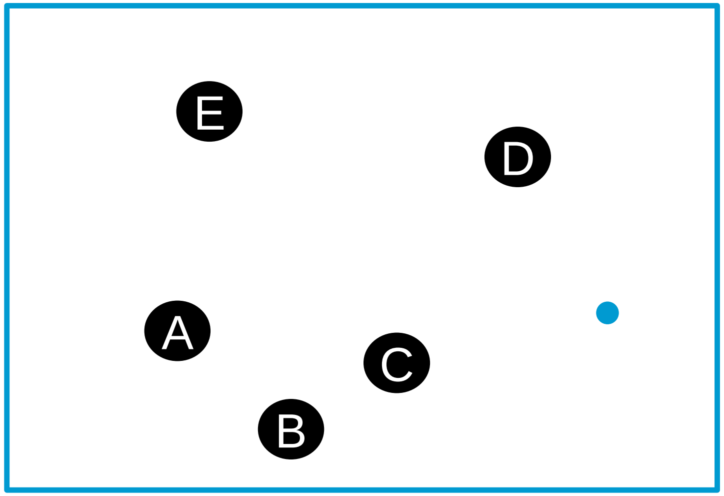
- the use of more fine grain analyses than the simple use of spatial variables as context effetcs (e.g. East- vs. West-Germany)
- the analysis of spatial units in relation to each other and not as isolated container (as in multilevel models)
- and they allow for the specific analysis of spatial processes.

Examples of spatial methods of analysis:

Spatial correlation (Moran‘s I), spatial interpolation, spatial regression, Geographically Weighted Regression (GWR), analysis of spatial heteroskedasticity, measures of segregation

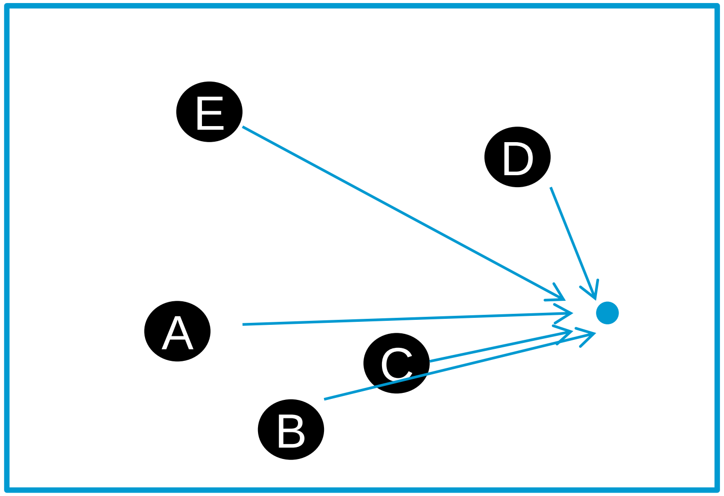
Distances between observational units

Arrangement of Features in Space



Distances between observational units

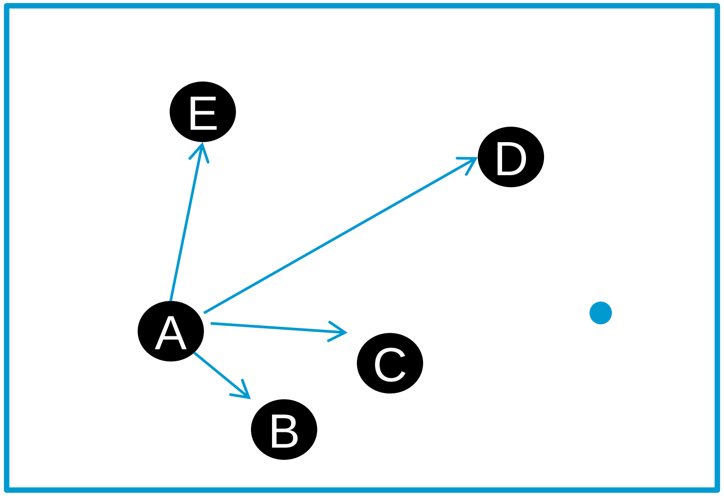
Distance to one feature



Feature	Distance
A	8
B	6
C	4
D	4
E	10

Distances between observational units

Distances between features of the same kind



ID / ID	A	B	C	D	E
A	-	2	3	10	7
B	2	-	1	12	10
C	3	1	-	8	9
D	10	12	8	-	6
E	7	10	9	6	-

→ **Distance matrix**

Spatial methods of analysis: Distribution in space

Are units clustered, i.e. is the occurrence more than random at specific places?

ID	A	B	C	D	E	Mean
A	-	2	3	10	7	5,5
B	2	-	1	12	10	6,3
C	3	1	-	8	9	5,3
D	10	12	8	-	6	9
E	7	10	9	6	-	8,8
						8,7

1. Mean distance of units to each other

2. Mean distance with random allocation in space

Spatial methods of analysis: Interpolation

Problem with spatial aggregates:

1.) The Modifiable Areal Unit Problem (MAUP):

a) Scaling effects

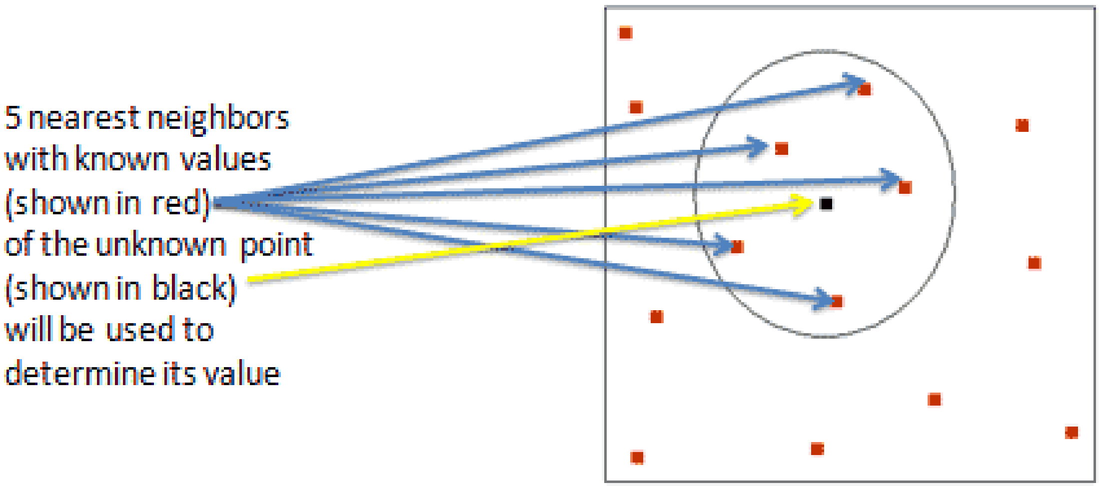
b) Zoning Effects

2.) Wrong level of aggregation for the analysis

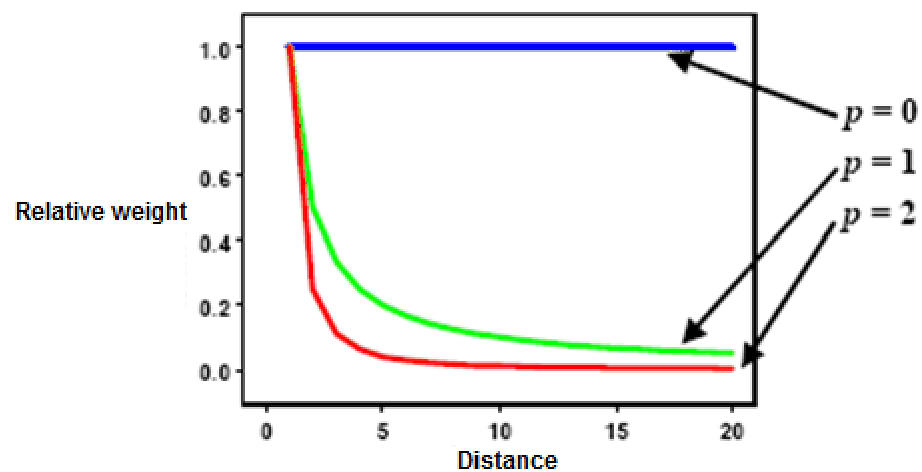
→ Interpolation.

The surface can be aggregated again in the desired spatial units.

Spatial methods of analysis: Interpolation

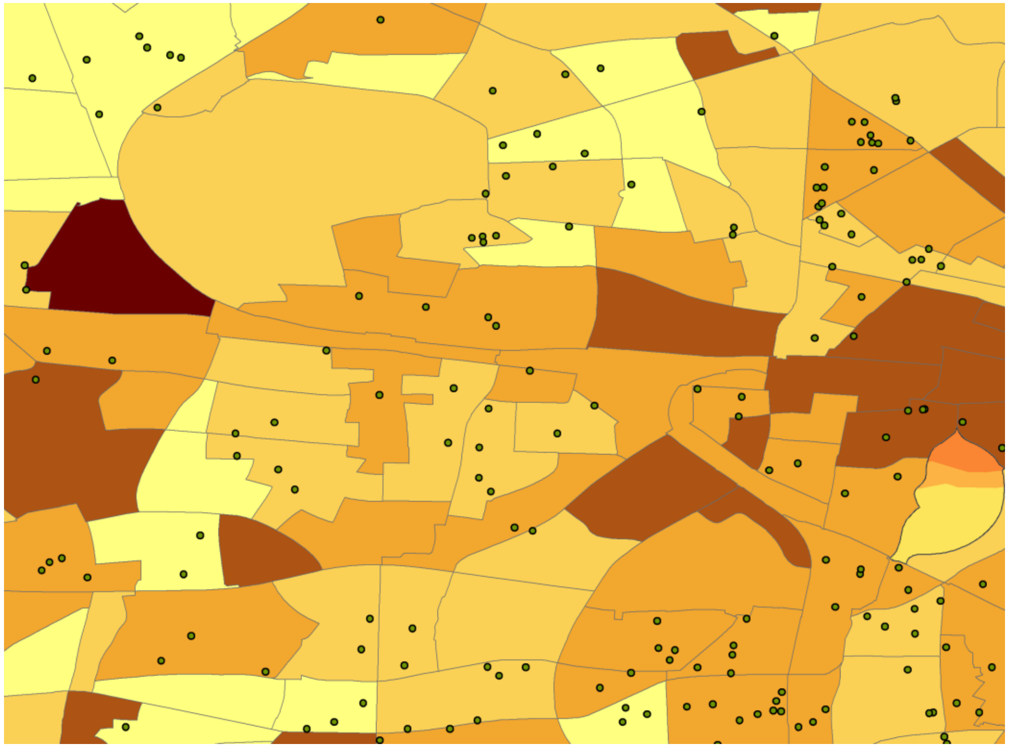


Interpolation as analysis of density or as weighted means of the Surrounding area (Inverse Distance Weighted-Interpolation)



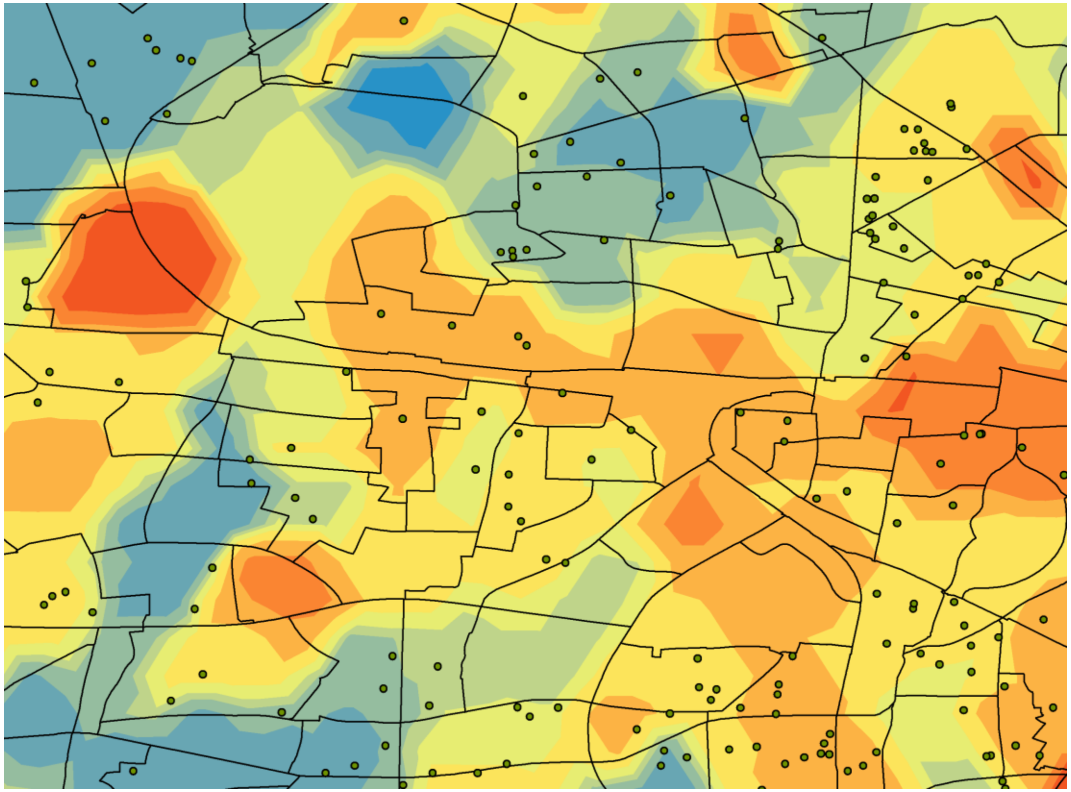
Spatial methods of analysis: Interpolation

Share of foreigners in city districts and places of living



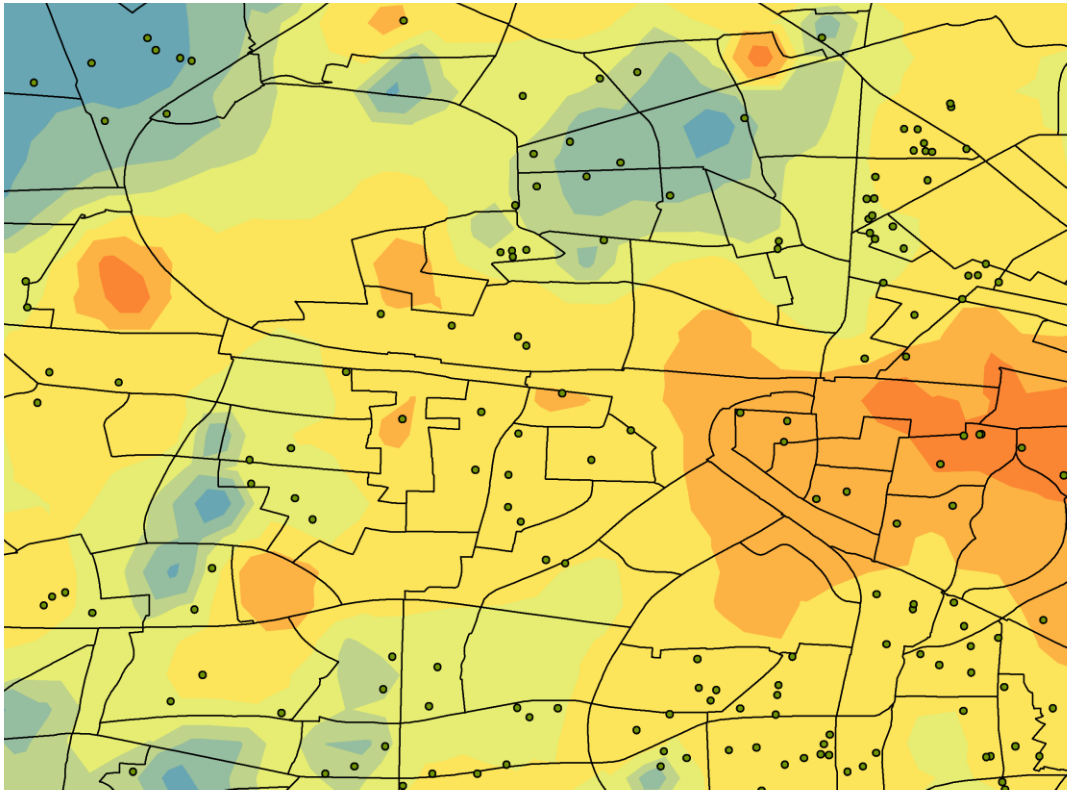
Spatial methods of analysis: Interpolation

Share of foreigners interpolated (narrow)



Spatial methods of analysis: Interpolation

Share of foreigners interpolated (wide)



Methods of Spatial Analysis: Residential Segregation

„The extent to which, within a given geographical area (e.g. a city), individuals belonging to different social groups live in neighborhoods characterized by different social compositions.“ (Reardon/O‘Sullivan 2004)

Dissimilarity index:

- Most used measurement of segregation (residential or otherwise)
- 2 groups, e.g. Germans and Swiss
- Interpretation: Share of one group that had to relocate in order to achieve even distribution erreichen (aspatial index)

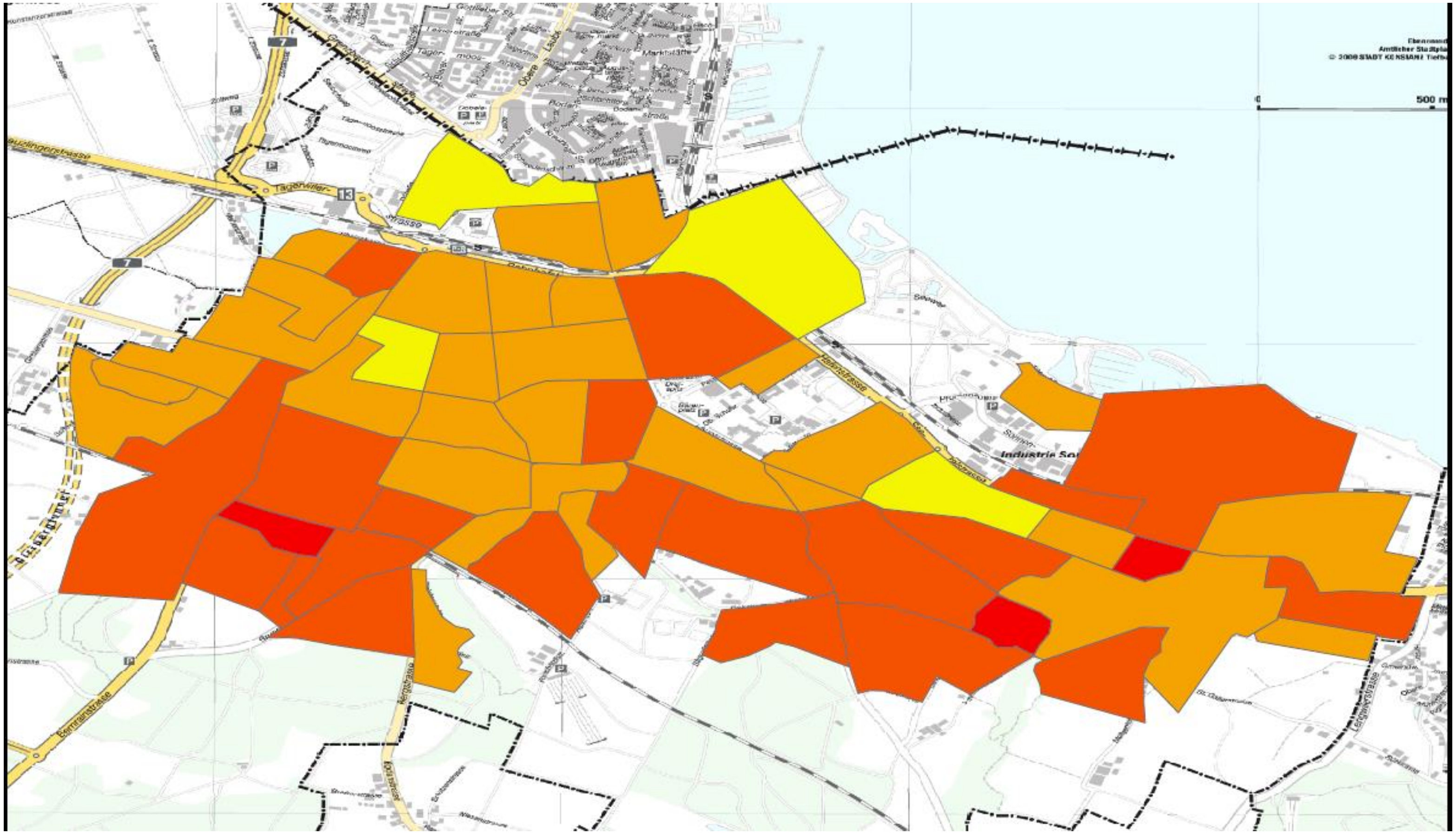
$$D = \sum_{i=1}^{n} \frac{t_i |\pi_i - \pi|}{2T\pi(1-\pi)}$$

aspatial index of dissimilarity (Duncan)

$$\widetilde{D} = \sum_{p \epsilon R} \frac{\tau_p |\tilde{\pi}_i - \pi|}{2T\pi(1-\pi)}$$

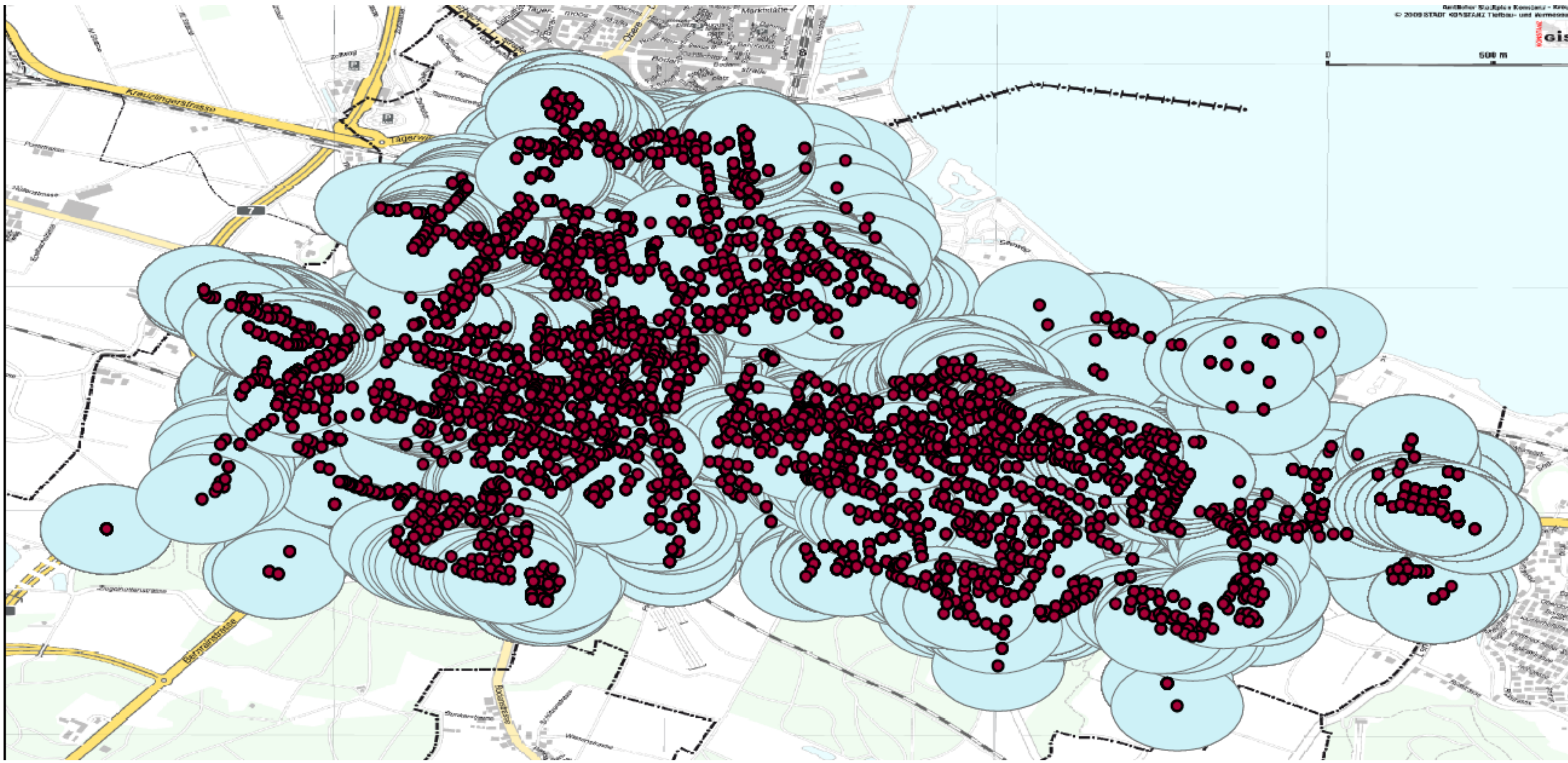
spatial index of dissimilarity

Methods of Spatial Analysis: Residential Segregation



Methods of Spatial Analysis: Residential Segregation

„Individual“ Neighborhoods (here: 200m)



Methods of Spatial Analysis: Residential Segregation

Neighbourhood (aspatial)

Year	1990	2000	2010
D	0.24	0.24	0.20

50m bandwidth

Year	1990	2000	2010
	0.36	0.37	0.32

200m bandwidth

Year	1990	2000	2010
	0.22	0.18	0.18

400m bandwidth

Year	1990	2000	2010
	0.13	0.12	0.11

Distances between units of observation

Spatial analyses account for the relative position of all units with regard to all other units of observation, i.e. basis of all statistical method is a distance or adjacency matrix.

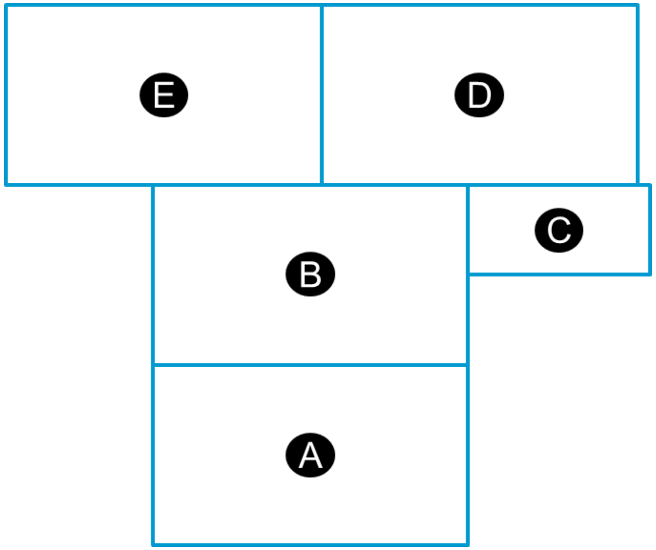
On this basis we can construct a Spatial-Weights-Matrix:

- Units with shared borders (contiguity)
- Distance
- Nearest neighbor
- Network

The spatial weights represent the distance between two units. We assume, that larger weights represent stronger correlation, influence of each other or common context.

What spatial weights? → THEORY

Spatial Weights Matrix: Contiguity

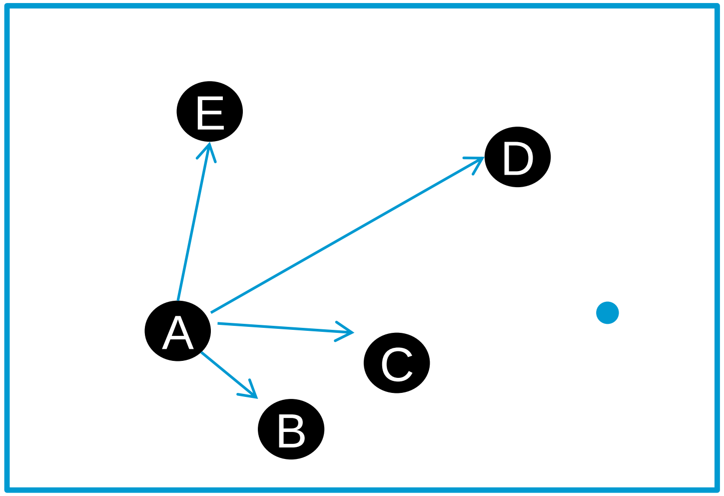


Spatial Weights Matrizen are usually row standardized.

ID / ID	A	B	C	D	E
A	-	1	0	0	0
B	1	-	1	1	1
C	0	1	-	1	0
D	0	1	1	-	1
E	0	1	0	1	-

ID / ID	A	B	C	D	E
A	-	1	0	0	0
B	0,25	-	0,25	0,25	0,25
C	0	0,5	-	0,5	0
D	0	1/3	1/3	-	1/3
E	0	0,5	0	0,5	-

Spatial Weights Matrix: Inverse Distance



ID	A	B	C	D	E
A	-	2	3	10	7
B	2	-	1	12	10
C	3	1	-	8	9
D	10	12	8	-	6
E	7	10	9	6	-

→

ID	A	B	C	D	E
A	-	1/2	1/3	1/10	1/7
B	1/2	-	1/1	1/12	1/10
C	1/3	1/1	-	1/8	1/9
D	1/10	1/12	1/8	-	1/6
E	1/7	1/10	1/9	1/6	-

Spatial Weights Matrix: Inverse Distance

There are only so much influence, therefore we standardize the spatial weights to on (also for computational reasons).
→ Weight / (G1 + G2 + G3 + G4)

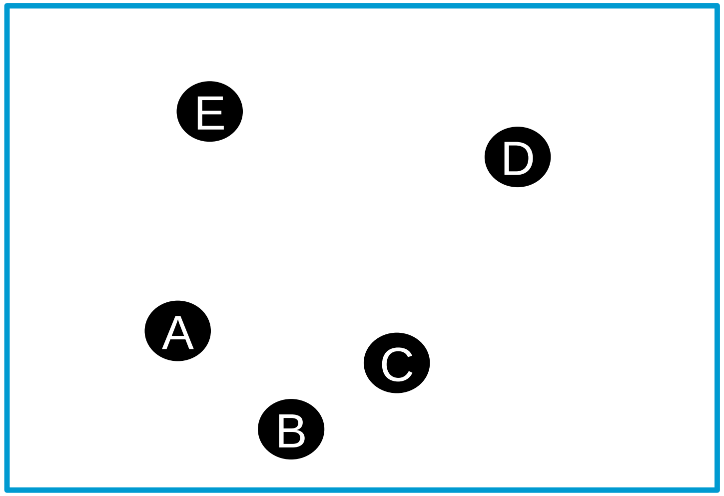
Further considerations:
- **Limit the influence at a maximal distance?**
- **Linear decay of spatial weight?**

ID	A	B	C	D	E
A	-	1/2	1/3	1/10	1/7
B	1/2	-	1/1	1/12	1/10
C	1/3	1/1	-	1/8	1/9
D	1/10	1/12	1/8	-	1/6
E	1/7	1/10	1/9	1/6	-

→

ID	A	B	C	D	E
A	-	0,5	0,3	0,1	0,1
B	0,3	-	0,6	0,05	0,05
C	0,2	0,6	-	0,1	0,1
D	0,2	0,2	0,25	-	0,35
E	0,3	0,2	0,2	0,3	-

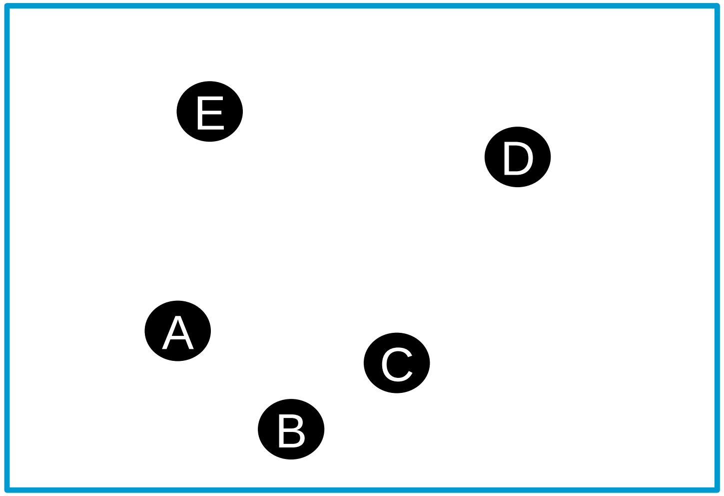
Methods of Spatial Analysis: Spatial Autocorrelation



ID	A	B	C	D	E
A	-	0,5	0,3	0,1	0,1
B	0,3	-	0,6	0,05	0,05
C	0,2	0,6	-	0,1	0,1
D	0,2	0,2	0,25	-	0,35
E	0,3	0,2	0,2	0,3	-

ID	Wert
A	4
B	4
C	5
D	9
E	2

Methods of Spatial Analysis: Spatial Autocorrelation



ID	A	B	C	D	E
A	-	0,5	0,3	0,1	0,1
B	0,3	-	0,6	0,05	0,05
C	0,2	0,6	-	0,1	0,1
D	0,2	0,2	0,25	-	0,35
E	0,3	0,2	0,2	0,3	-

ID	Value	Spatial weighted Mean	
A	4	0,5*4+0,3*5+0,1*9+0,1*2	4,6
B	4		4,75
C	5		4,3
D	9		3,55
E	2		5,7

Methods of Spatial Analysis: Spatial Autocorrelation

$$W = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix} \qquad W^* = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

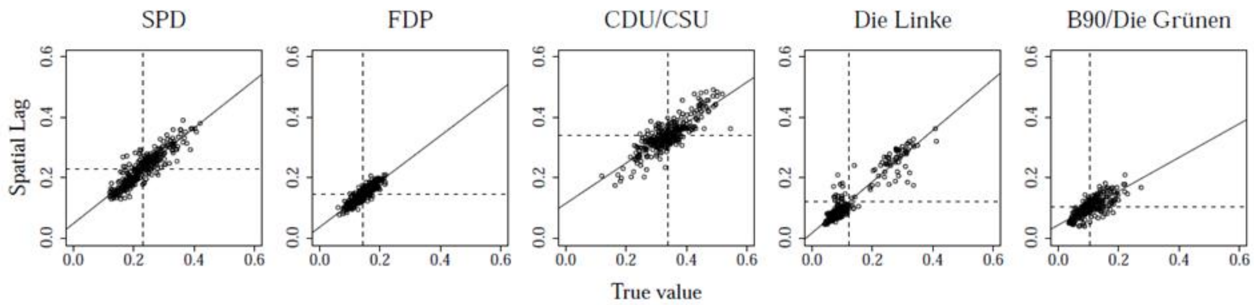
$$W^* y_i = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix} = \begin{pmatrix} y_2 \\ \frac{1}{2}y_1 + \frac{1}{2}y_3 \\ \frac{1}{3}y_2 + \frac{1}{3}y_4 + \frac{1}{3}y_5 \\ \frac{1}{2}y_3 + \frac{1}{2}y_5 \\ \frac{1}{2}y_3 + \frac{1}{2}y_4 \end{pmatrix}$$

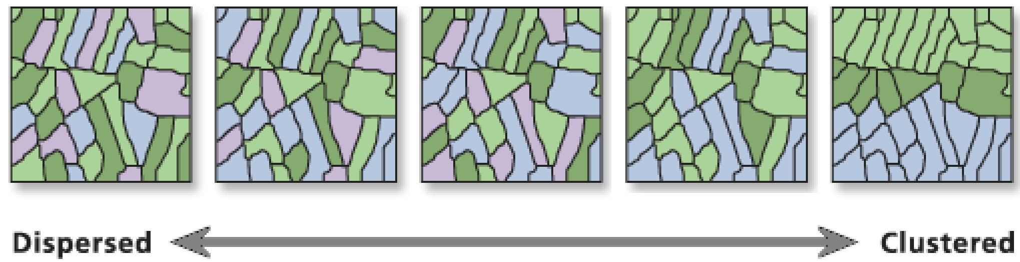
- **Calculate the correlation of values**
- **Significant?**
- **Positive? Negative?**

Methods of Spatial Analysis: Spatial Autocorrelation

Global Moran's I

$$I = \frac{N}{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}} \times \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}\,(y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^{N}(y_i - \bar{y})^2}$$





Local Moran's I

$$I_i = \frac{(y_i - \bar{y})\sum_{j=1}^{N} w_{ij}\,(y_j - \bar{y})}{\left(\sum_{i=1}^{N}(y_i - \bar{y})^2\right)/N}$$

Next steps in Stata

→ open the program

```
ssc install fre
ssc install shp2dta
ssc install spmap

findit spatwmat

sysuse auto
fre make
sum price
```

Hands-on 5

- Import spatial data in Stata
- Graphical presentation of spatial data in Stata
- Creation of Spatial Weights Matrix
- Calculation of Moran‘s I

Universität Konstanz

Thank You For Your Attention!

Thomas Wöhler
Fachbereich Geschichte und Soziologie, AG Hinz

Tel.: +49 (0) 75 31/88 - 3345
thomas.woehler@uni-konstanz.de