



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

Geography of Campylobacteriosis in the Czech Republic

*Prostorová analýza výskytu kampylobakterií v
České Republice*

Lukáš MAREK

lukas.marek@upol.cz

StátGis Team

www.geoinformatics.upol.cz



KATEDRA GEOINFORMATIKY
Univerzita Palackého v Olomouci

Why Geographical Information Systems?

- ▶ Advanced methods for spatial analyses
- ▶ Exploration of spatial pattern
- ▶ Spatial statistics
- ▶ Visualization and presentation for non-geographers (doctors, specialist)

Spatial Analyses of Health Data - AOI

- ▶ Disease mapping
 - ▶ Visual description of spatial variability of the disease incidence
 - ▶ Maps of incidence risk, identification of areas with high risk
- ▶ Analyses of spatial pattern
 - ▶ Exploration of spatial and spatio-temporal patterns in data
 - ▶ Disease clusters, randomness, ...
- ▶ Geographic correlation studies
 - ▶ Analysis of associations among the incidence and environmental factors

Health and Medical Data

- ▶ require specific procedures because of their confidentiality
 - ▶ management, presentation and operations
- ▶ aggregated, anonymized or incomplete data sets
- ▶ usage of suitable analytical procedures, while the uncertainty and the inaccuracy of data characteristics need to be taken into account

Data Privacy

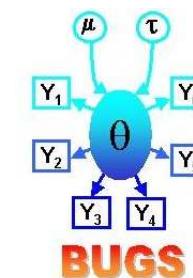
- ▶ Health and medical data = private, confidential and sensitive data
 - ▶ Keeping all available records but prevent their re-identification
 - ▶ Usefulness of the local scale analysis X privacy protection
 - ▶ Unlikely to explore the relations on the individual level (and not necessary)
 - ▶ Availability, accessibility and restrictions
-

Software



 **SpaceStat™**
 **ClusterSeer®**
 **BoundarySeer®**

 **OpenEpi**



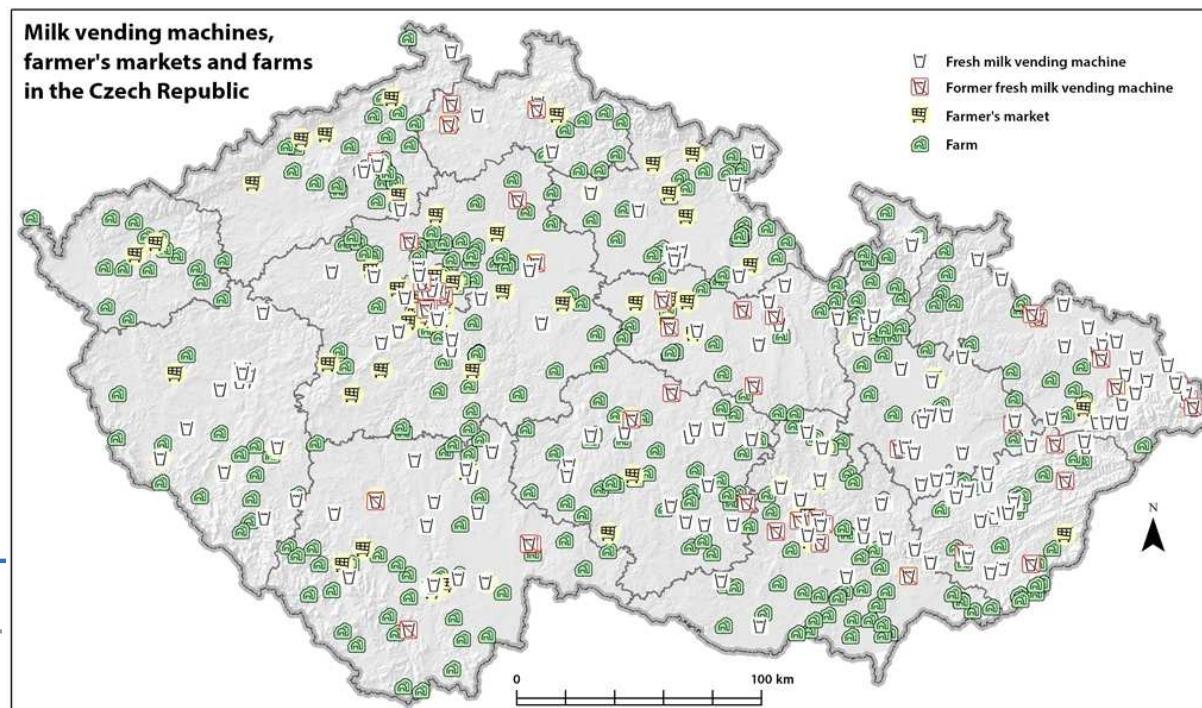
Objectives

- ▶ Disease occurrence as the event
 - ▶ Space, time, attributes
 - ▶ Spatial (point) pattern
- ▶ Spatial evaluation of infectious diseases in Olomouc Region, Czech Republic
 - ▶ Parotitis, Salmonella, Viral intestinal infections
- ▶ **Spatial evaluation of Campylobacter infection in the Czech Republic**

Campylobacteriosis



- ▶ Campylobacter bacterium (*C. jejuni*)
- ▶ Frequent
- ▶ Often foodborne
- ▶ Symptom are similar to salmonella
- ▶ Poultry or fresh milk products



Data

- ▶ Collaboration with Regional Public Health Service in Olomouc and the National Institute of Public Health
- ▶ EPIDAT database
 - ▶ Mandatory records about infectious diseases and patients, manually fulfilled
 - ▶ Age, Sex, Date, Profession, Place of residence, infection, isolation, ...
 - ▶ 2008 - 2012
 - ▶ ≈ 100 000 records
 - ▶ (weakly) Anonymized
- ▶ Aggregation
 - ▶ Hexagons with the size of average cadastral unit, administrative units
 - ▶ EUROSTAT population grid

Soubor	Úpravy	Format	Zobrazení	Nápověda
2011P89	1	11	P8	9
2011P810	1	11	P8	10
2011P11	10	11	P1	11
2011P811	1	11	P8	11
2011P12	11	11	P1	12
2011P512	2	11	P5	12
N				
2011P812	1	11	P8	12
2011P513	2	11	P5	13
N				
2011P114	12	11	P1	14
2011P514	2	11	P5	14
2011P614	3	11	P6	14
N				
2011P415	2	11	P4	15
2011P515	2	11	P5	15
2011P615	3	11	P6	15
2011P116	15	11	P1	16
2011P516	2	11	P5	16
2011P616	3	11	P6	16
2011P916	1	11	P9	16
N				
2011P517	2	11	P5	17
2011P617	3	11	P6	17
2011P917	1	11	P9	17
2011P518	2	11	P5	18
2011P618	3	11	P6	18
2011P918	1	11	P9	18
N				
2011P019	3	11	P0	19
N				
2011P919	1	11	P9	19
2011P020	3	11	P0	20
N				
2011P920	1	11	P9	20
N				
2011P221	4	11	P2	21
2011P321	6	11	P3	21
2011P921	1	11	P9	21
N				
2011P222	18	11	P1	22
2011P223	5	11	P2	22
2011P223	5	11	P2	23
A				
2011P024	4	11	P0	24
N				
2011P124	20	11	P1	24
2011P025	4	11	P0	25
N				
2011P125	20	11	P1	25
2011P026	4	11	P0	26
N				
2011P526	4	11	P5	26
2011P027	4	11	P0	27
N				
2011P527	3	11	P5	27
N				

RStudio

File Edit Code View Plots Session Project Build Tools Help

statistika.R x geokod_NA.R x time_series.R x data.pw x geokod.R x

Source on Save

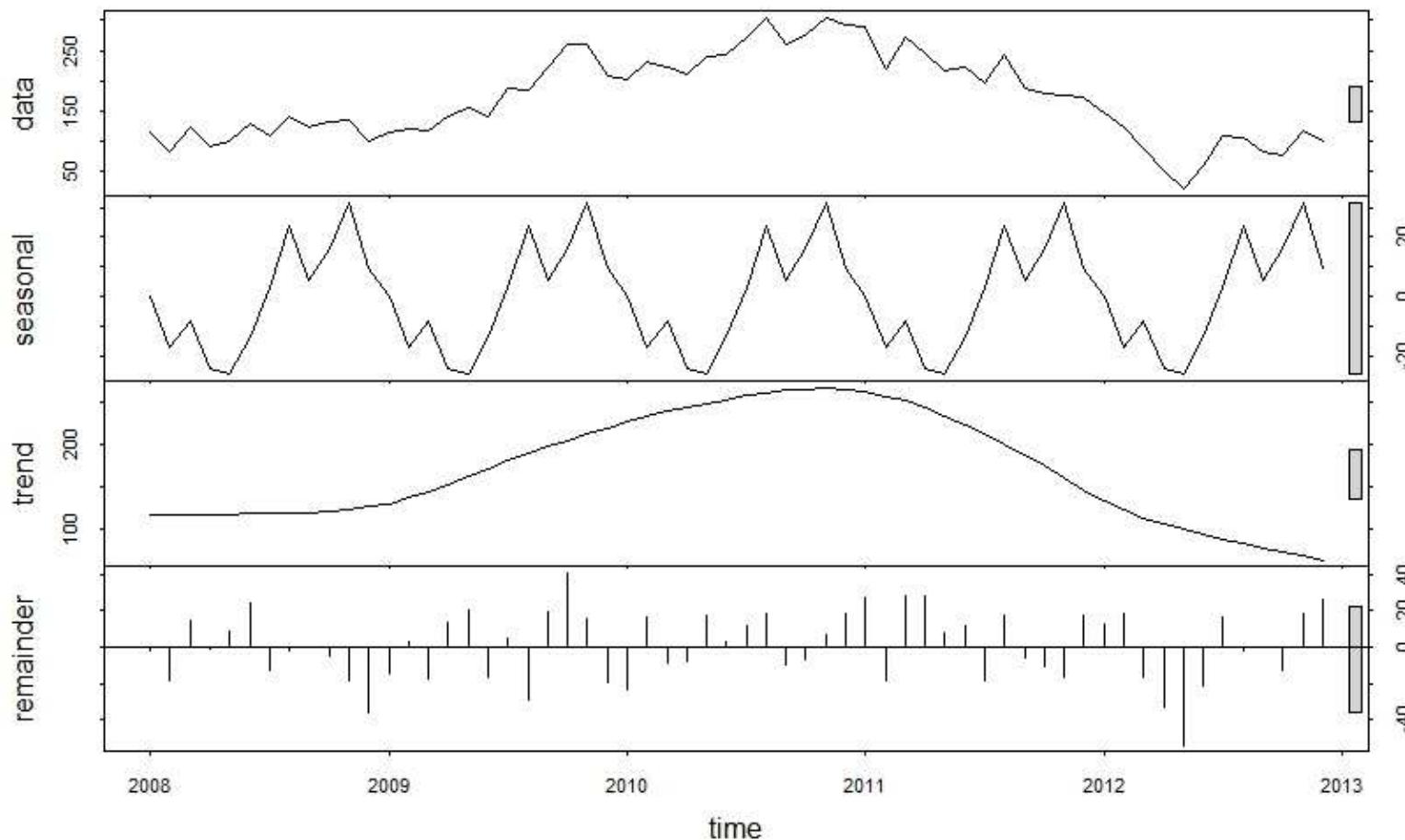
```

1 setwd("D:/skola/szu/txt_bez_diakritiky")
2 library(audio)
3 geokoduj <- function(adr) {
4   library(XML) # knihovna pro rozpracovani vysledneho xml
5   adr <- gsub(' ', '%20', adr) # generovani spravnych parametru do URL
6   #otevreni spojeni
7   connectStr <- paste("http://api4.mapy.cz/geocode?query=", adr, sep="")
8   con <- url(connectStr)
9   data.xml <- xmlTreeParse(paste(readLines(con), collapse=""))
10  close(con)
11  #zavreni spojeni a nasleduje reseni vyhovenych dat
12  koren<-xmlRoot(data.xml)
13  if (!is.null(koren[[["point"]]][[["item"]]])) {
14    x<-xmlGetAttr(koren[[["point"]]][[["item"]]], "x")
15    y<-xmlGetAttr(koren[[["point"]]][[["item"]]], "y")
16    return(c(x,y))
17  } # v koren se dostavam k hodnotam atributu x a y
18 }
19 }
20
21 data <- read.delim("D:/skola/SZU/txt_bez_diakritiky/bezd_2009.txt", dec=",")
22 data[,3] <- ""
23 data[,4] <- ""
24
25 # geokodovani zaznamu
26 # for (i in 1:10){
27 for (i in 1:nrow(data)){
28   adr <- geokoduj(data[i,2])
29   wait(0.5)
30   ifelse(is.null(adr) == FALSE, data[i,3:4] <- adr, data[i,3:4] == 0)
31 }
32 write.table(data, "sour2009.txt")
33
34 data2 <- read.delim("D:/skola/SZU/txt_bez_diakritiky/bezd_2011.txt", dec=",")
35 data2[,3] <- ""
36 data2[,4] <- ""
37
38 for (j in 1:nrow(data2)){
39   adr2 <- geokoduj(data2[j,2])
40   wait(0.5)
41   ifelse(is.null(adr2) == FALSE, data2[j,3:4] <- adr2, data2[j,3:4] == 0)
42 }
43 write.table(data2, "sour2011.txt")
44
45 data3 <- read.delim("D:/skola/SZU/txt_bez_diakritiky/bezd_2010.txt", dec=",")
46 data3[,3] <- ""
47 data3[,4] <- ""
48
49 for (k in 1:nrow(data3)){
50   adr3 <- geokoduj(data3[k,2])
51   wait(0.5)
52   ifelse(is.null(adr3) == FALSE, data3[k,3:4] <- adr3, data3[k,3:4] == 0)
53 }
54 write.table(data3, "sour2010.txt")

```

Time

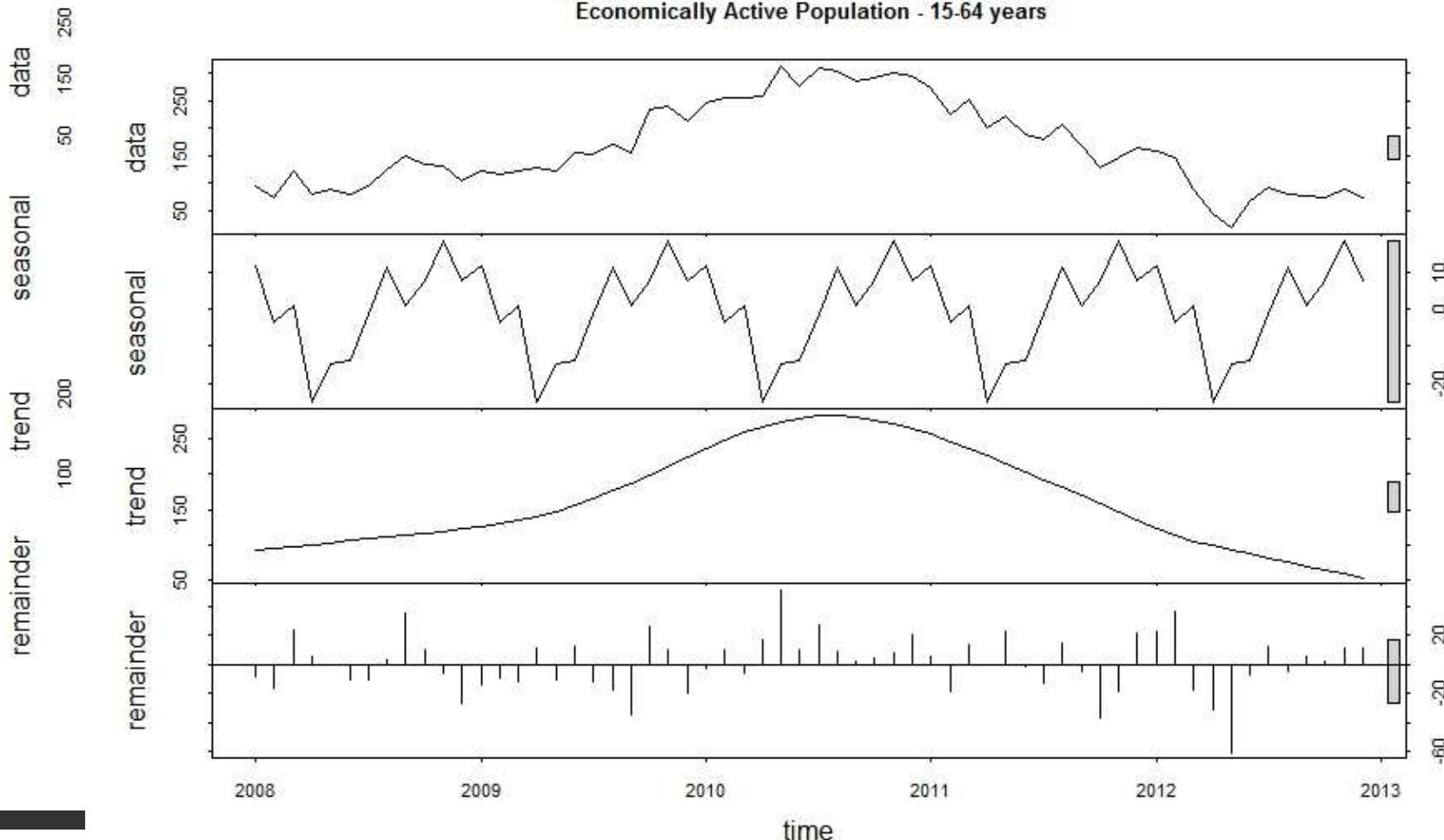
Occurrence of Campylobacteriosis in 2008 - 2012
Economically Inactive Population - 0-14 years



Time

Occurrence of Campylobacteriosis in 2008 - 2012
Economically Inactive Population - 0-14 years

Occurrence of Campylobacteriosis in 2008 - 2012
Economically Active Population - 15-64 years

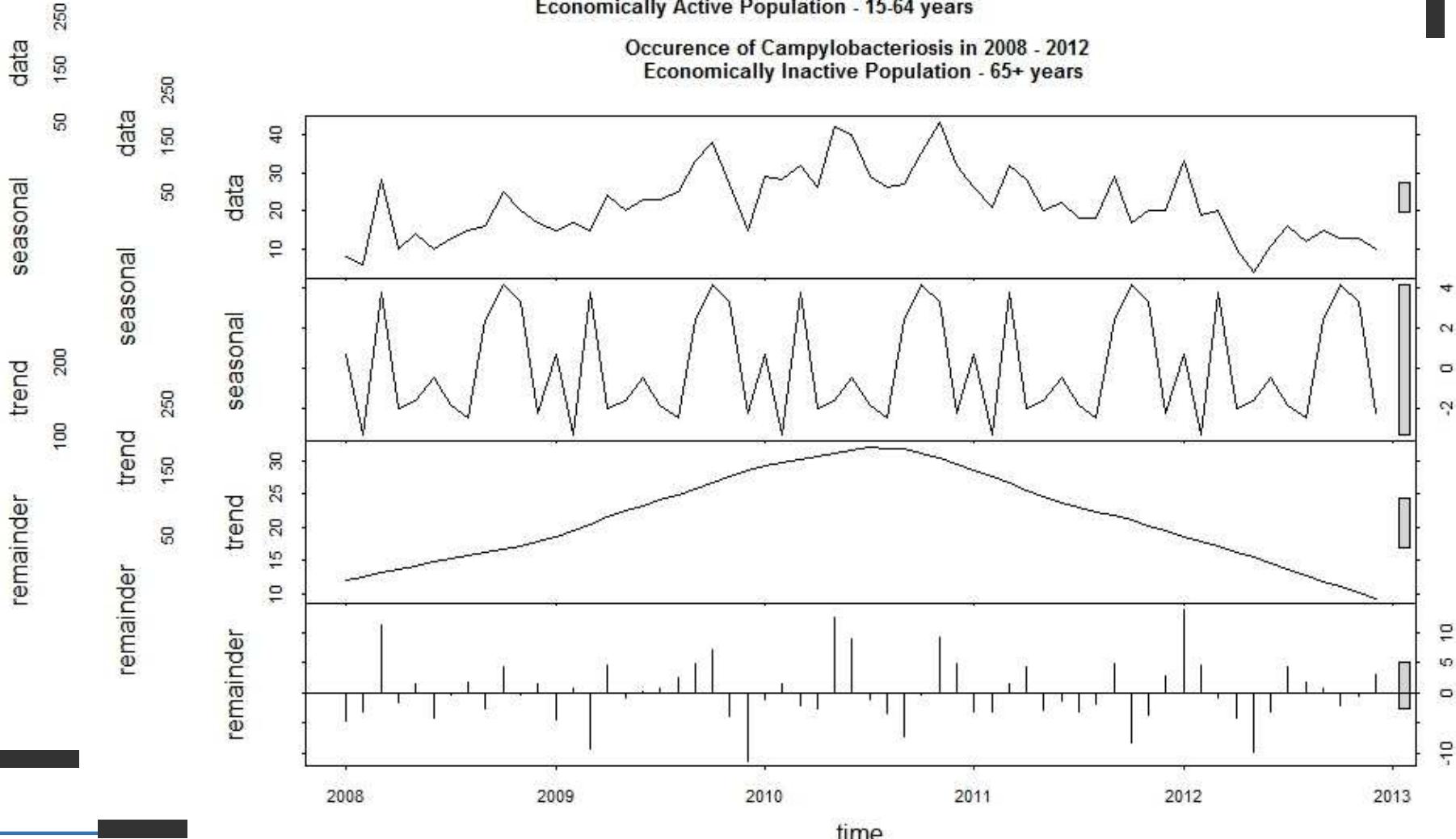


Time

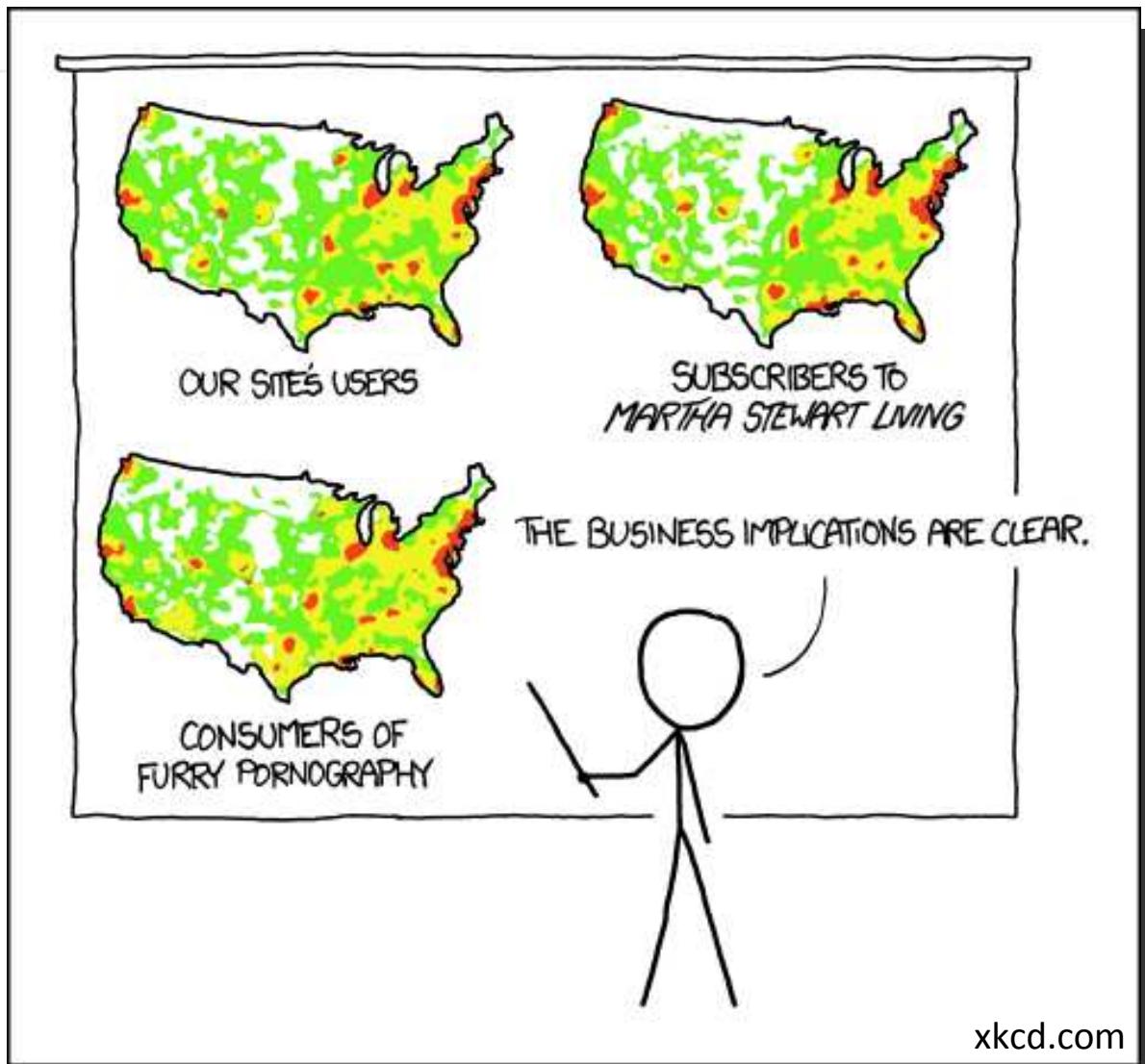
Occurrence of Campylobacteriosis in 2008 - 2012
Economically Inactive Population - 0-14 years

Occurrence of Campylobacteriosis in 2008 - 2012
Economically Active Population - 15-64 years

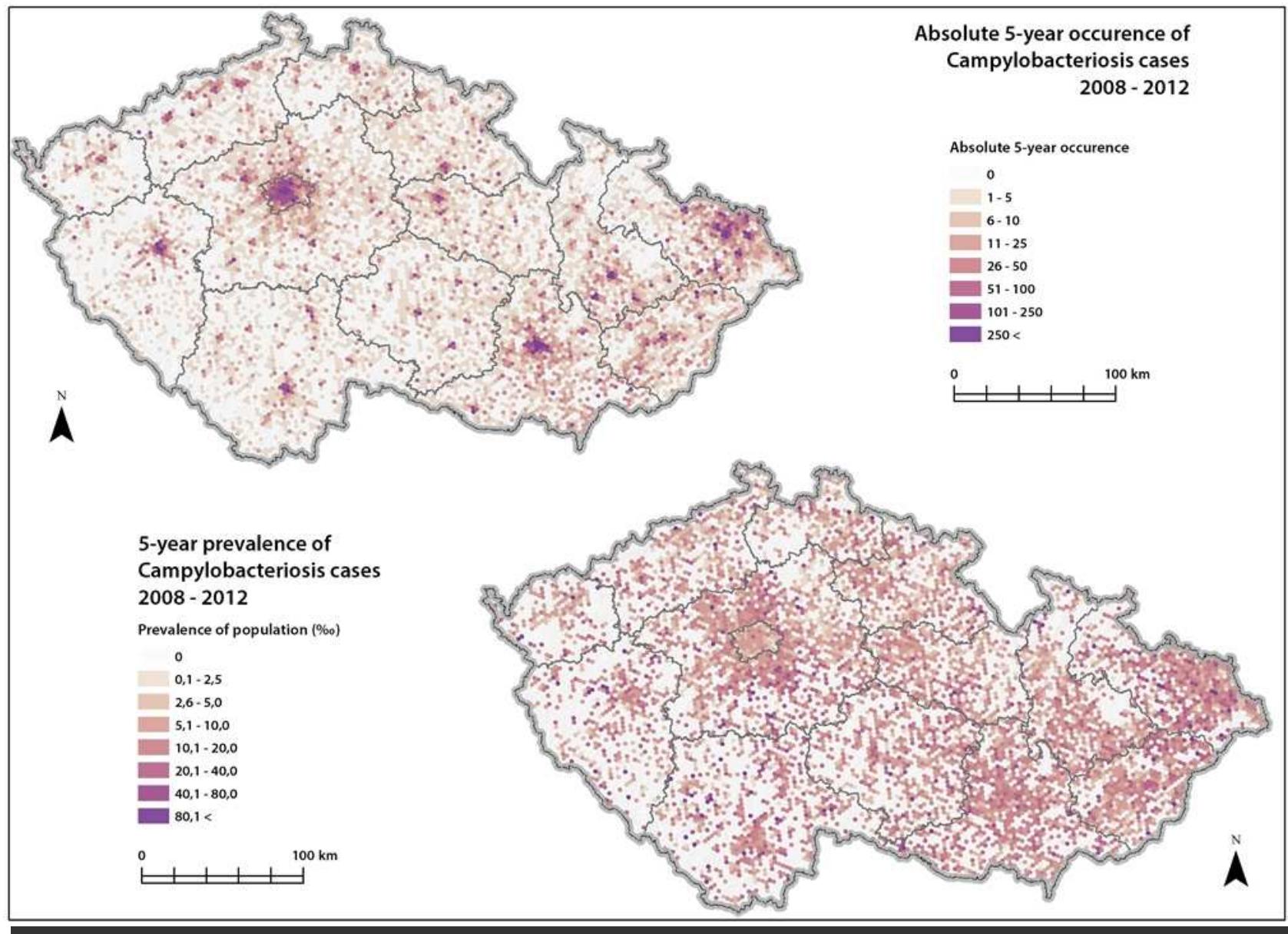
Occurrence of Campylobacteriosis in 2008 - 2012
Economically Inactive Population - 65+ years



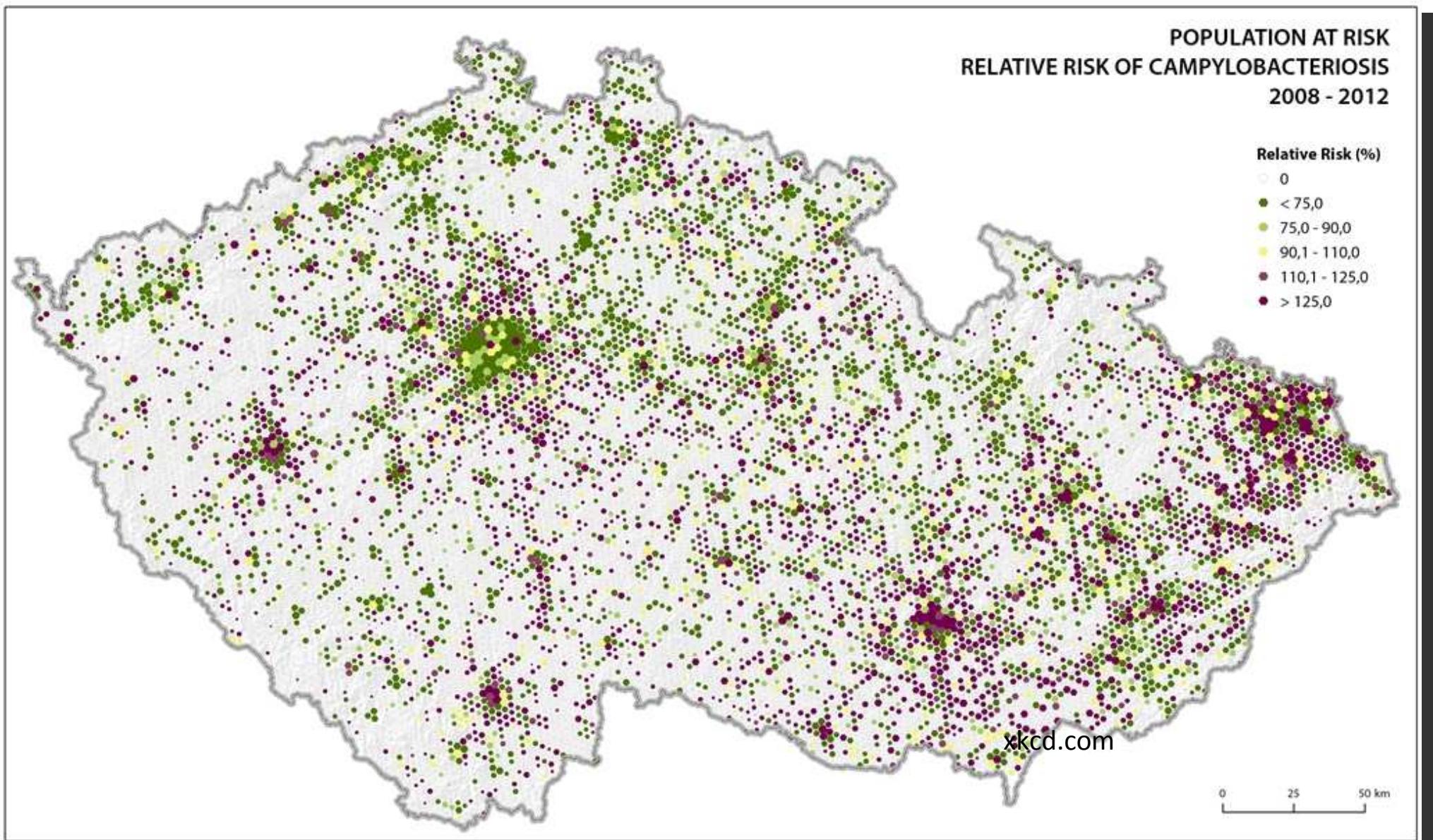
Disease mapping



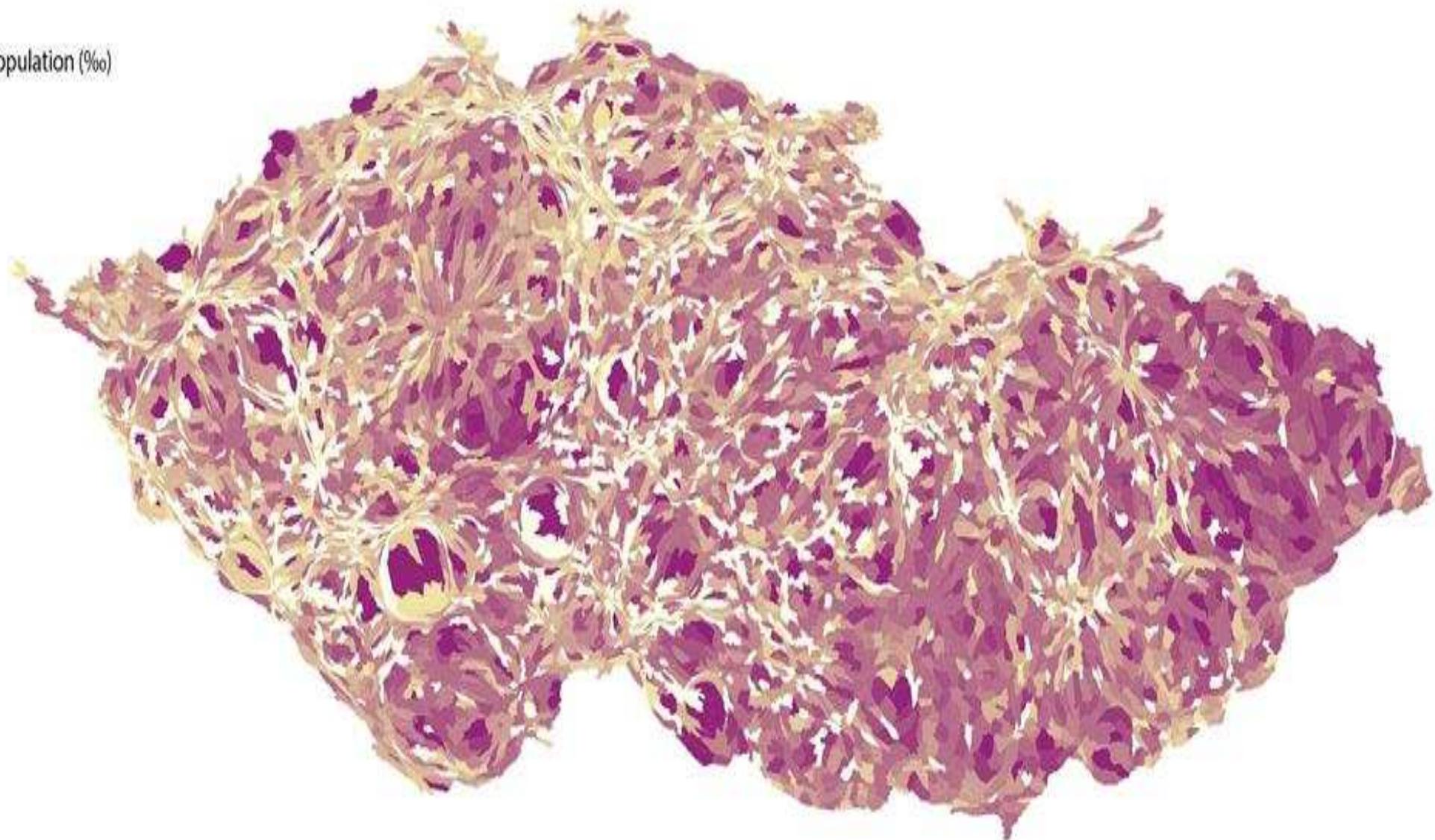
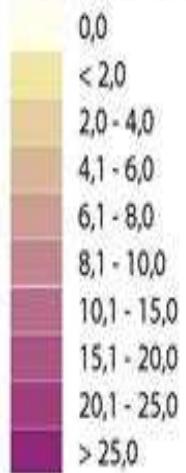
Disease mapping



Disease mapping

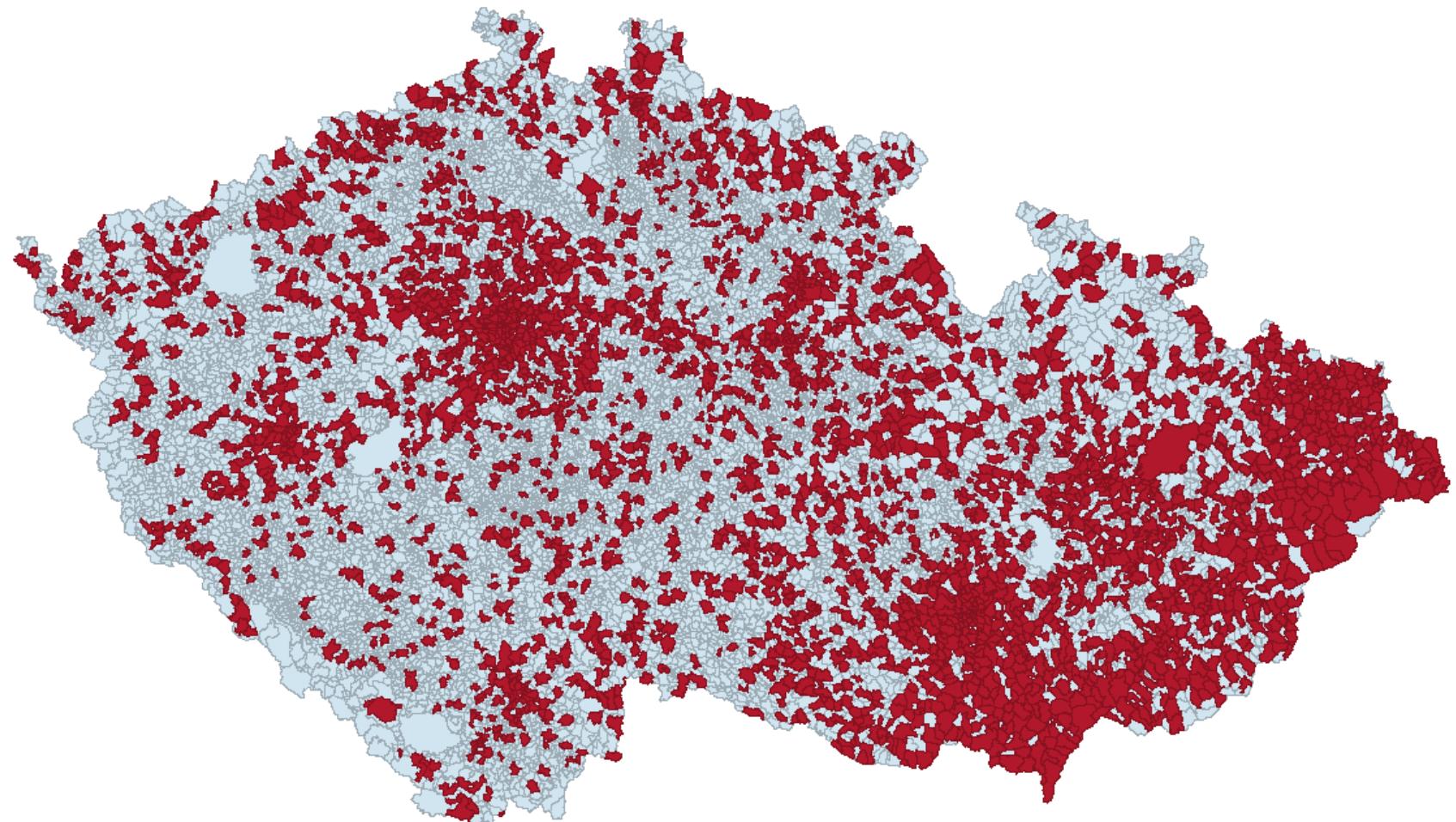


Prevalence of population (%)



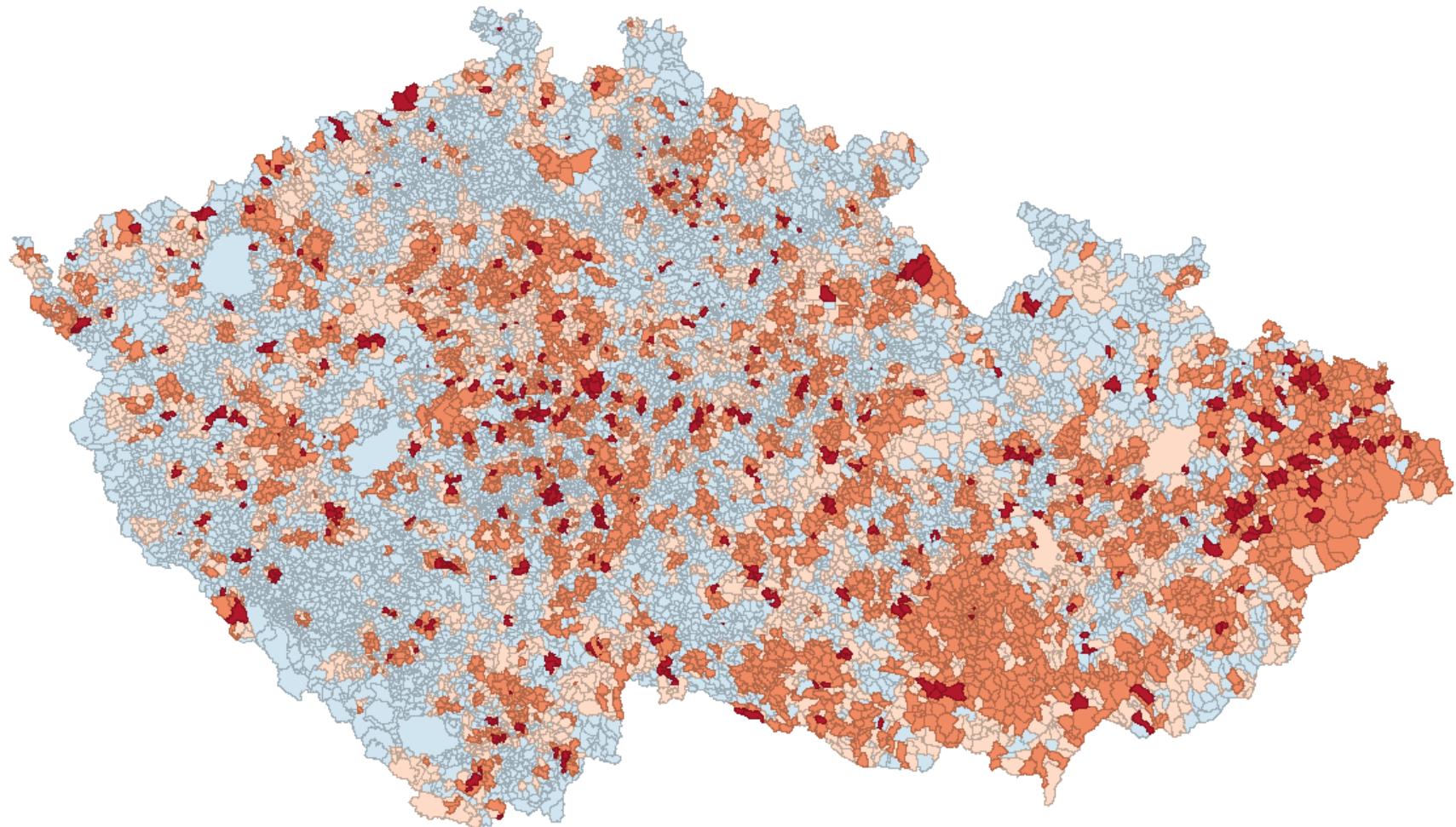
Hinge=1.5: Raw Rate P_ALL

- Lower outlier (0)
- < 25% (0)
- 25% - 50% (11692)
- 50% - 75% (0)
- > 75% (0)
- Upper outlier (3431)



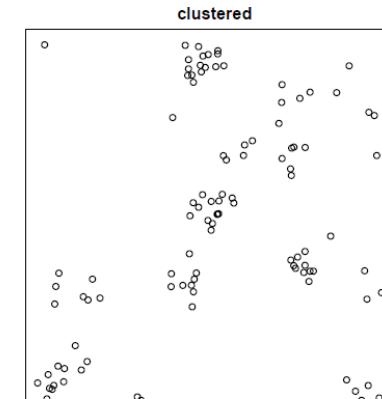
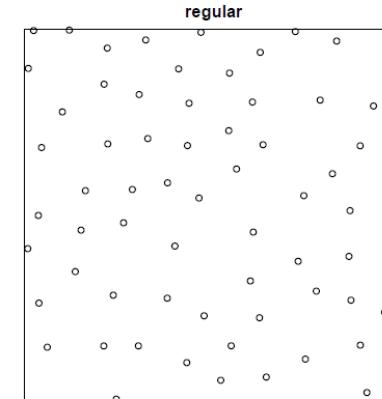
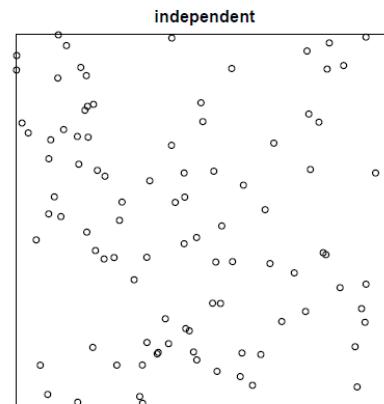
Hinge=1.5: SEBS-Smoothed

- Lower outlier (0)
- < 25% (0)
- 25% - 50% (7562)
- 50% - 75% (3780)
- > 75% (3270)
- Upper outlier (511)

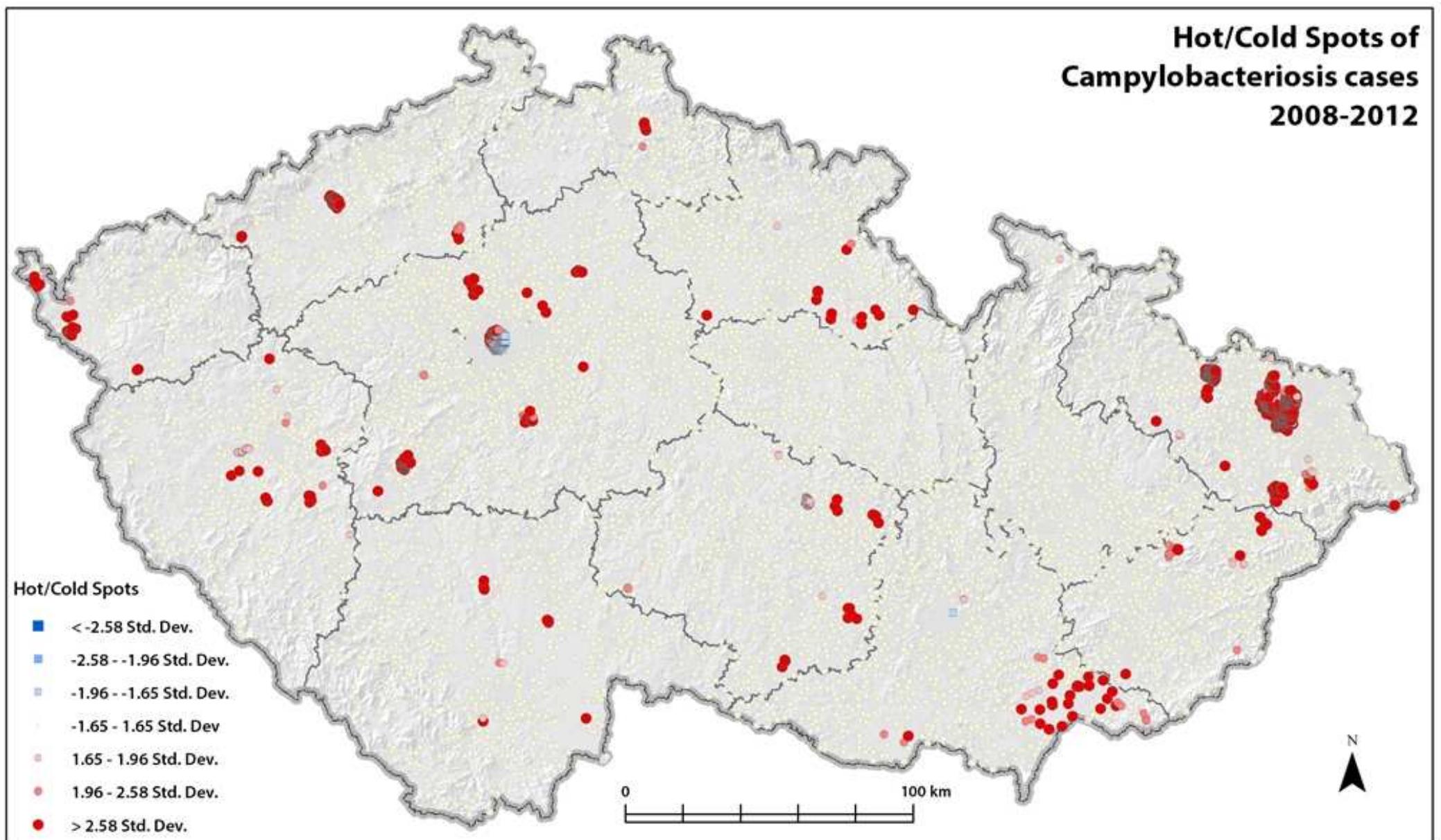


Identification of Spatial Processes

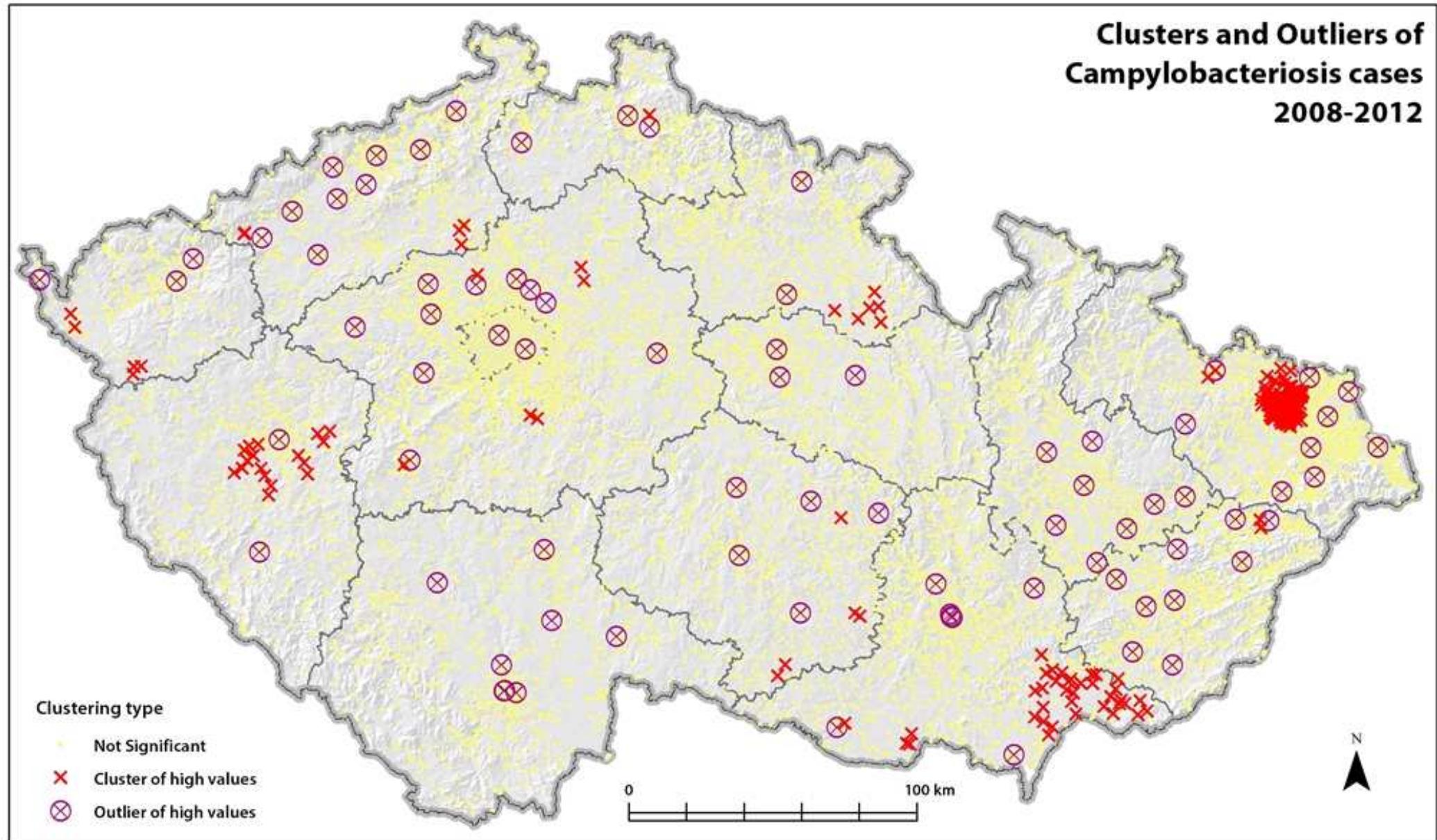
- ▶ Estimation of the nature of phenomenon
- ▶ Plenty of methods
 - ▶ Global vs. Local
 - ▶ Graphical vs. Numerical
- ▶ Testing of Complete Spatial Randomness
 - ▶ Or other spatial processes
- ▶ Scale dependency
- ▶ Bayesian modelling



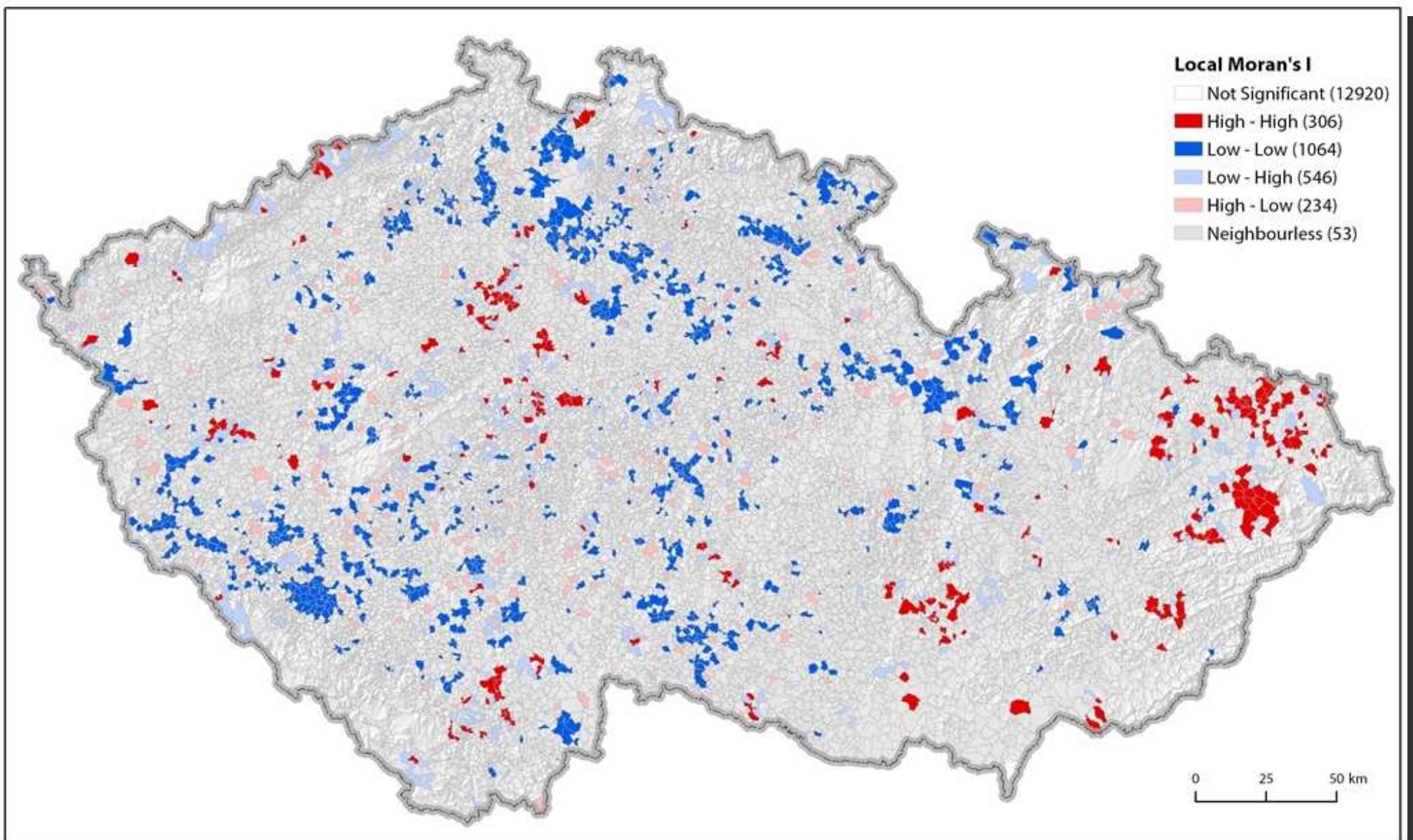
Spatial Pattern



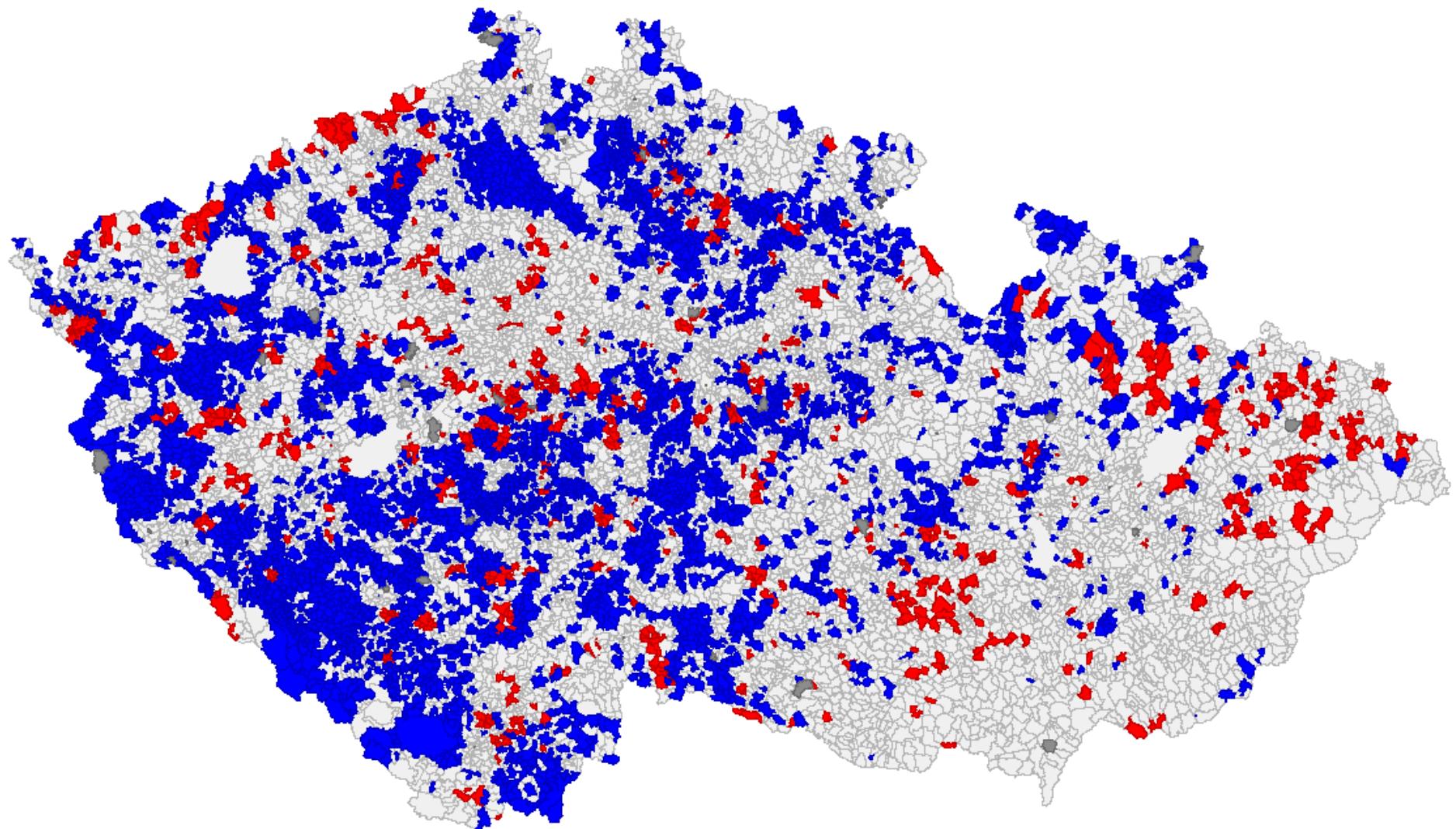
Spatial Pattern



Local Moran's I with EB Rate

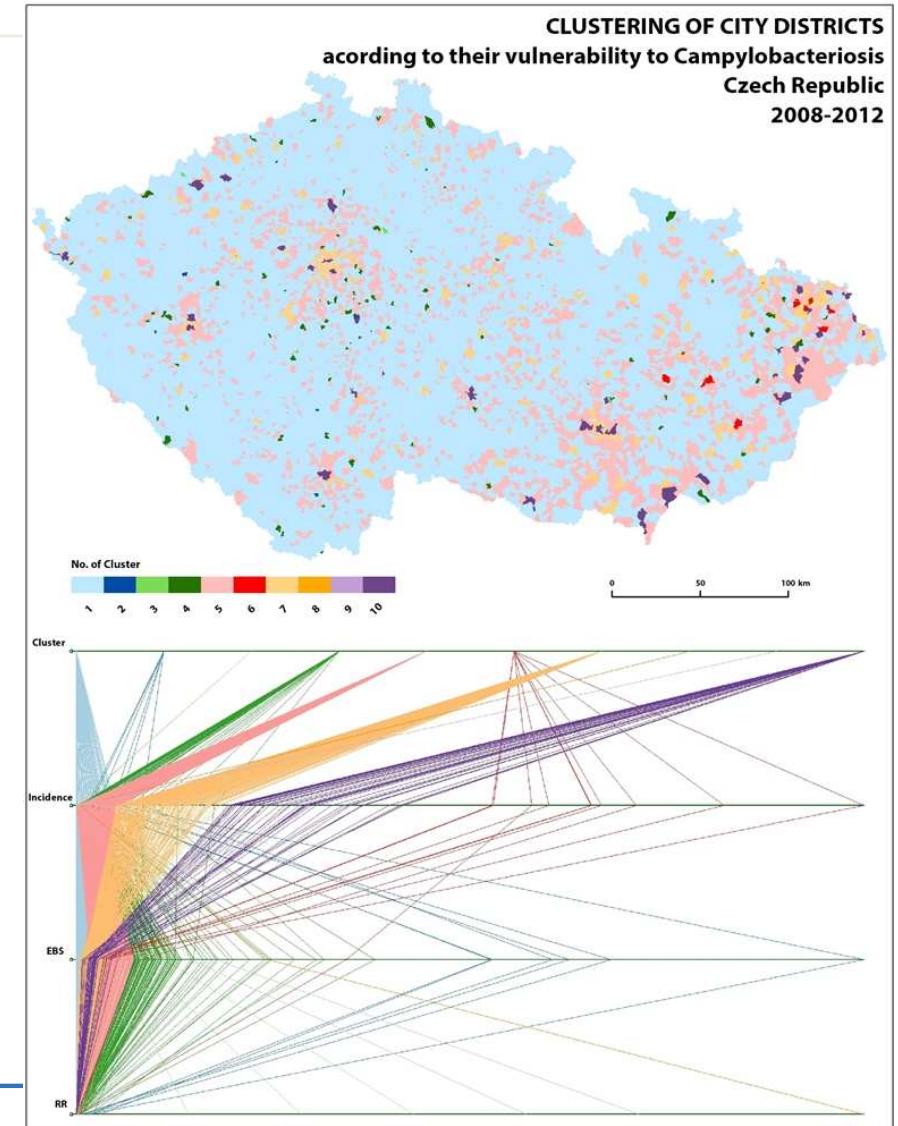


Local Gi^* in time

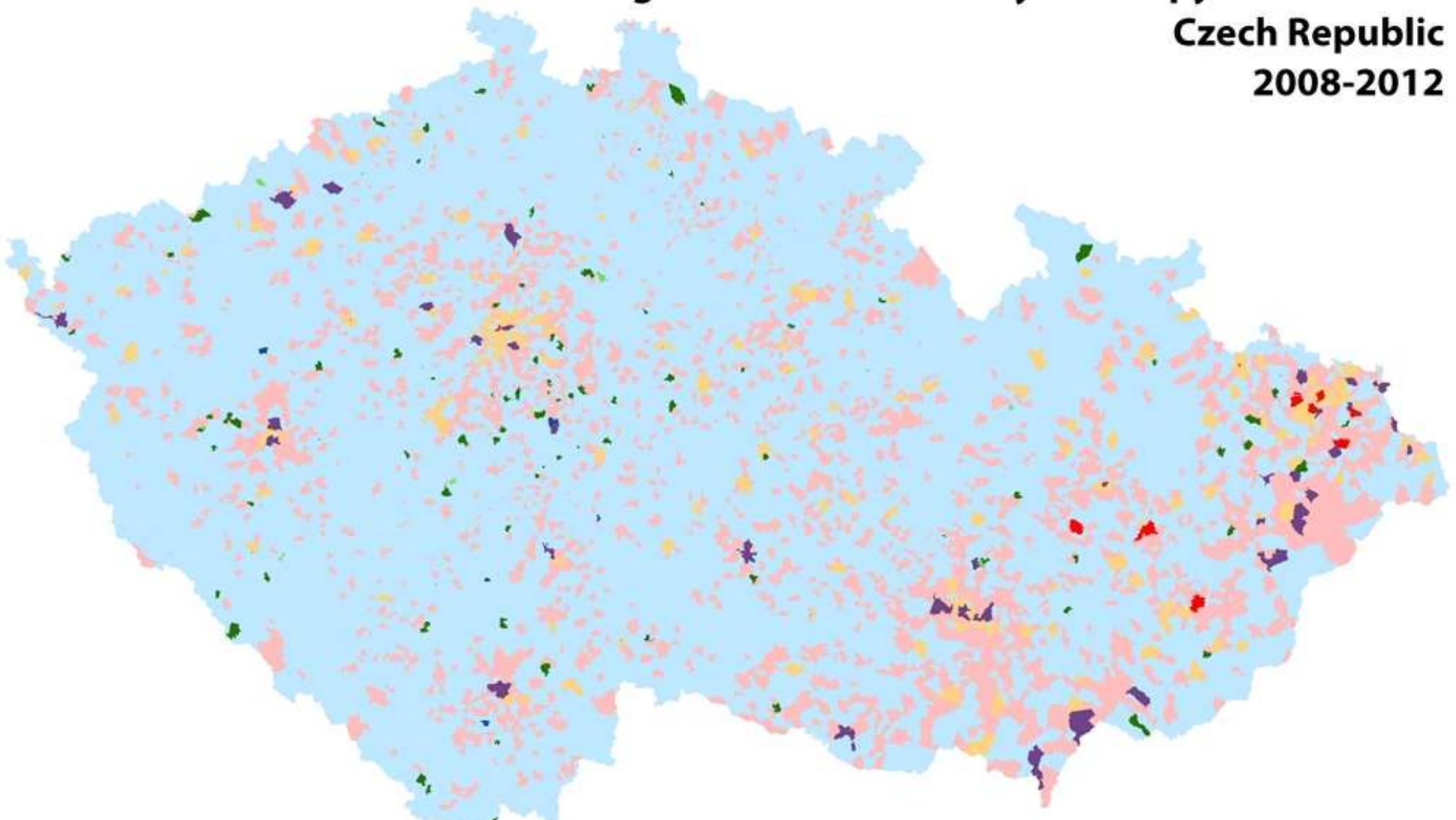


Multivariate Clustering

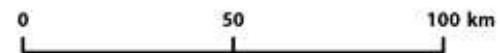
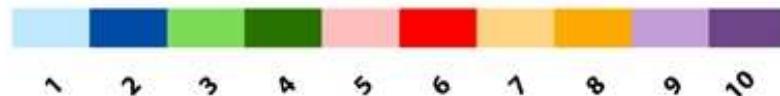
- ▶ Similarity searching in attribute space
- ▶ Classification of areas with related properties of occurring diseases and their parameters
- ▶ Classification of similar cases
- ▶ Without spatial relations



**CLUSTERING OF CITY DISTRICTS
according to their vulnerability to Campylobacteriosis
Czech Republic
2008-2012**



No. of Cluster

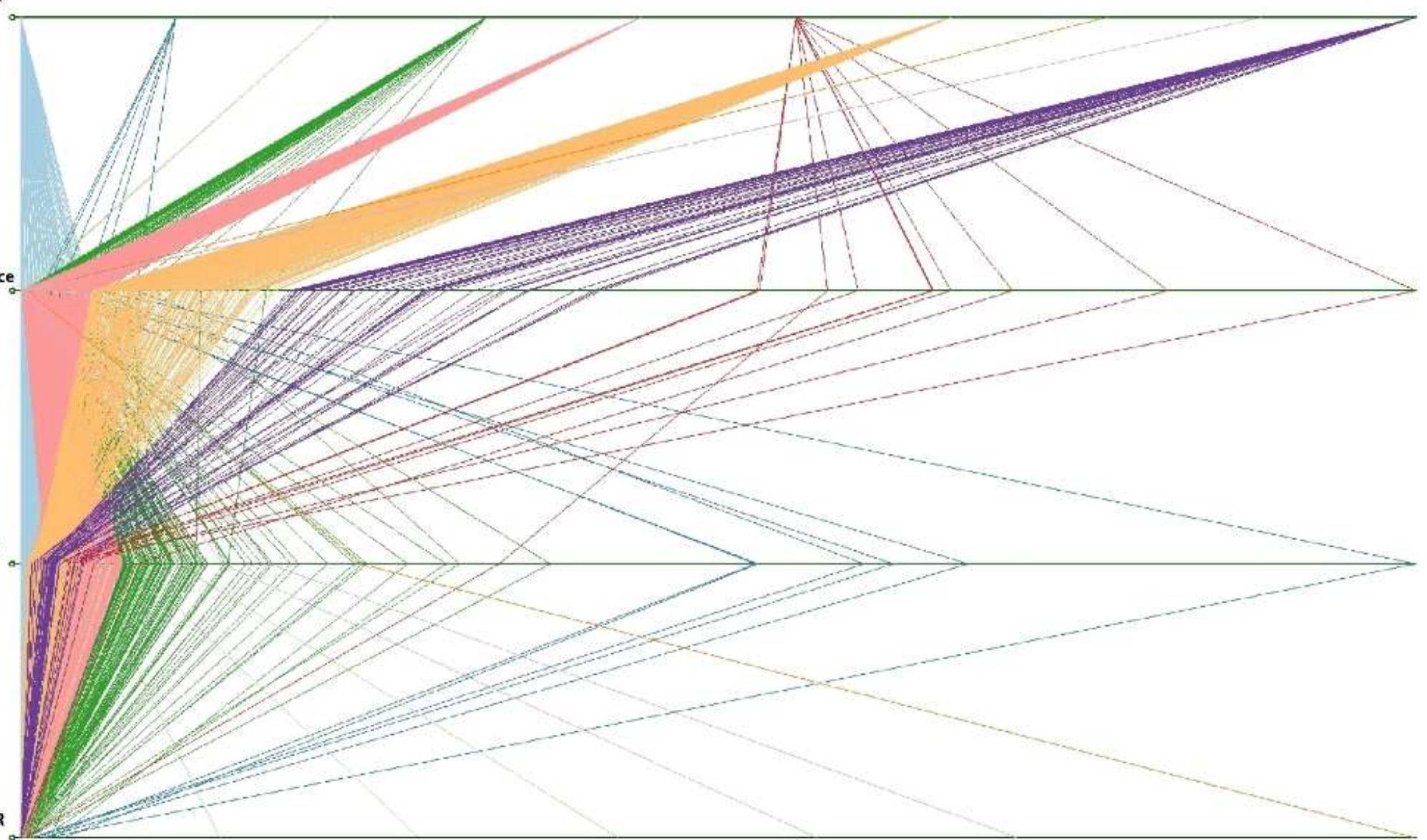


Cluster

Incidence

EBS

RR



Problems / Challenges

- ▶ Geocoding
 - ▶ Aggregation
 - ▶ Age / Population standardization
 - ▶ Neighbourhood estimation
 - ▶ Modifiable areal unit problem
 - ▶ Probability distribution of the disease occurrence
 - ▶ Underlying processes
 - ▶ Under / Overestimation of results leading to misinterpretation
-

►ACKNOWLEDGEMENT



INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

► The author gratefully acknowledge the support by the Operational Program Education for Competitiveness - European Social Fund (project CZ.1.07/2.3.00/20.0170 of the Ministry of Education, Youth and Sports of the Czech Republic)

StuGIS Team

► THANK YOU FOR YOUR ATTENTION

Geography of Campylobacteriosis in the Czech Republic

*Prostorová analýza výskytu kampylobakterií v České
Republike*

- ▶ Lukáš MAREK
- ▶ lukas.marek@upol.cz