Indicators Valuation using FCA and Moebius Inversion Function

Petr Gajdoš, Daniela Ďuráková, Lubor Tvrdý, George Lev

Department of Computer Science, Department of Regional Economics, VŠB-Technical University of Ostrava Interconcult A.S.A.

Contents

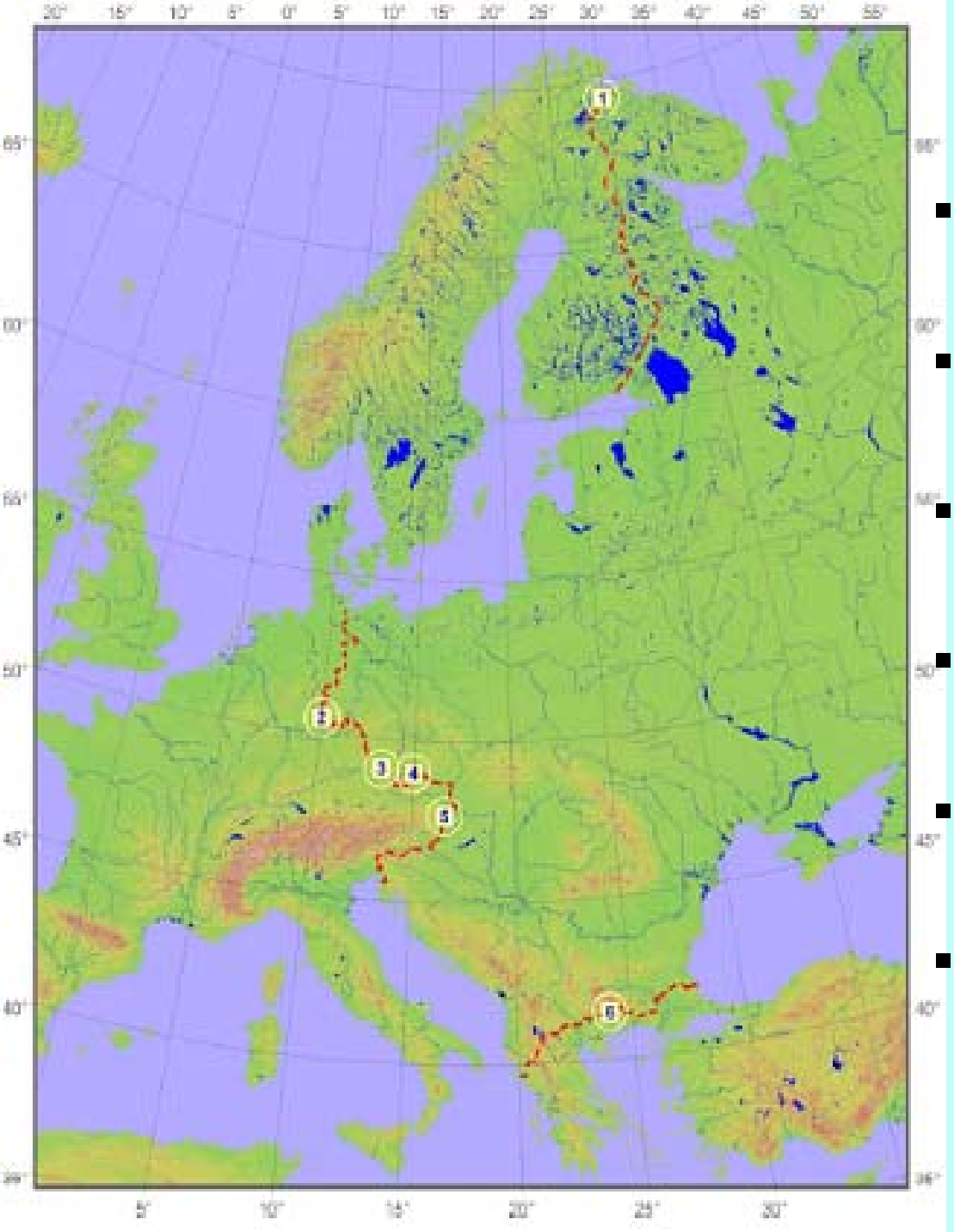
- Motivation of collaboration
- Theoretical Background
  - Context, Concept and Concept Lattice
  - Diversity Function
  - Conjugate Möebius Inverse (CMI)
- Applications
  - FCA, Diversity, CMI
  - Order hierarchy
  - Basic interpretation of numbers
- Conclusion

The Iron Curtain Project

Innovative models of critical key indicators as planning and decision support for sustainable rural development and integrated cross border regional management in former Iron Curtain areas based on north to south European reference studies

http://www.ironcurtainproject.com

Reference Area



- **RA1**
  NORWAY/RUSSIA
  Pasvik Valley
- **RA2**
  GERMANY/GERMAN
  Biosphere Reserve Rhön
- **RA3**
  GERMANY/CZECH REPUBLIC
  Bavarian Forest/ Šumava
- **RA4**
  CZECH REPUBLIC/AUSTRIA
  Waldviertel/Trebon Basin
- **RA5**
  AUSTRIA/HUNGARY
  Kekfrankos Area
- **RA6**
  GREECE/BULGARY
  Nestos/Mesta River

Motivation of collaboration

- Set of indicators – influence upon the area evolution
- Four indicators categories – aspects of regional competition

  N - environmental (natural),

  E - economical,

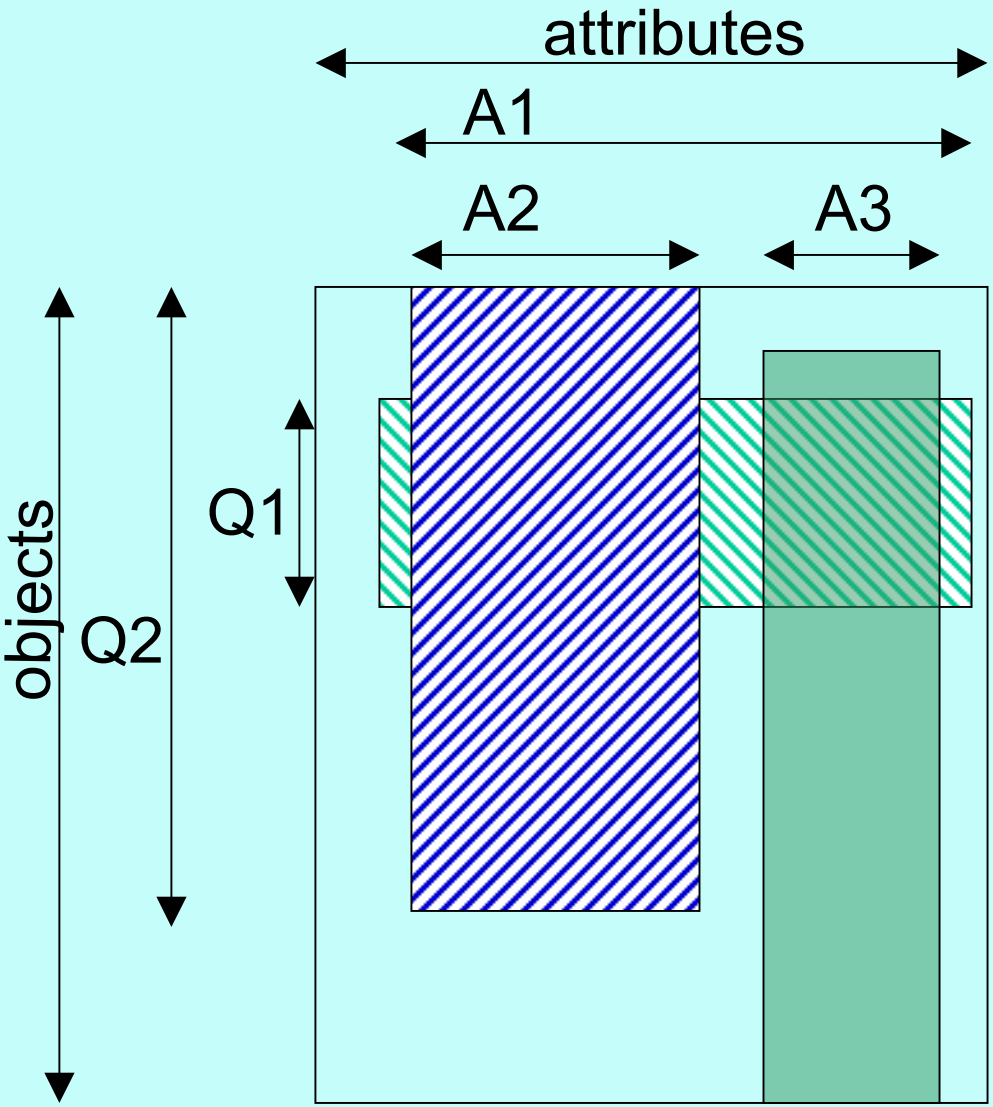
  S - social and

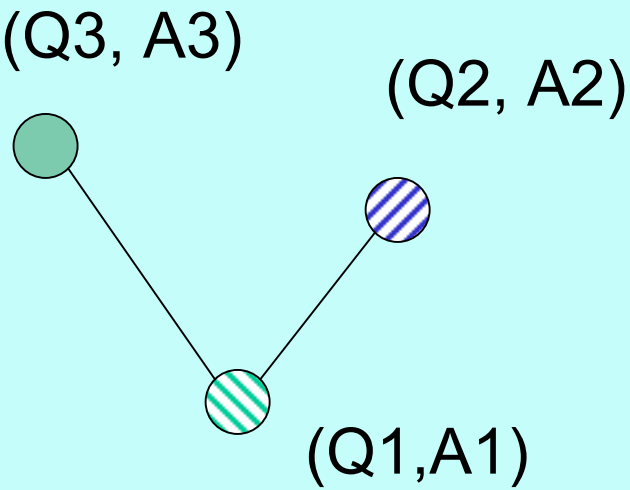
  G - local position in global context
- Searching of most important indicators
- Ten points scale of weight for every aspect indicator
- Order indicators of every area – FCA, CMI

Indicators example

Frequency	Indicators	Index	E	G	N	S
6	Unemployment rate	1	7	0	0	4
5	Population change in 10 years (1990-2000)	2	8	5	0	0
4	Share of protected areas in RA	3	8	3	0	2
4	Average Overnight Stays	4	5	0	0	5
4	Net migration rate	5	7	4	0	0
4	Spatial share of forest	6	3	0	0	7
4	Spatial share of organic farming	7	4	0	0	7
4	Wage /average income	8	5	0	0	5
4	Accommodation capacity	9	9	0	5	2
4	GDP per capita	10	9	0	4	2
4	Use of fertilizers (kg/ha of arable land)	11	9	0	0	2
4	Share of University graduated	12	7	4	5	0
4	Population age structure	13	4	0	0	2
3	Abundance of Selected Key Species	14	4	4	0	0
3	Share of employment in Services	15	0	0	9	0
3	Share of employment in Industry	16	4	2	6	0
3	Share of agricultural land	17	7	0	0	3

Basic notions of Formal Concept Analysis





$(Q1,A1) \leq (Q2,A2)$

$Q1 \subseteq Q2 \Leftrightarrow A2 \subseteq A1$

Diversity Function

Definition 5. *Let F be the totality of all features deemed relevant in the specific context, and denote by $R \subseteq X \times F$ the "incidence" relation that describes the features possessed by each object, i.e. $(x, f) \in R$ whenever object $x \in X$ possesses feature $f \in F$. For each relevant feature $f \in F$, let $\lambda_f \geq 0$ quantify the value of realization of f. Upon normalization, λ_f can thus be thought of as the relevant importance, or weight of feature f. The diversity value of a set S is defined as*

$$v(S) = \sum_{f \in F : (x,f) \in R \text{ for some } x \in S} \lambda_f$$

$$v(S) = \sum_{A \cap S \neq \emptyset} \sum_{f \in F_A} \lambda_f$$

Then, for each subset $A \subseteq X$ denote by $\lambda_A := \sum_{f \in F_A} \lambda_f$ the total weight of all features with extension A, with the convention that $\lambda_A = 0$ whenever $F_A = \emptyset$. With this notation we write

$$v(S) = \sum_{A \cap S \neq \emptyset} \lambda_A$$

Conjugate Möebius Inverse

Theorem 1. *For any function $v : 2^X \to R$ with $v(\emptyset) = 0$ there exists unique function $\lambda : 2^X \to R$, the Conjugate Möebius Inverse, such that $\lambda_\emptyset = 0$ and, for all S,*

$$v(S) = \sum_{A : A \cap S \neq \emptyset} \lambda_A$$

Furthermore, the Conjugate Möebius Inverse λ is given by the following formula. For all $A \neq \emptyset$,

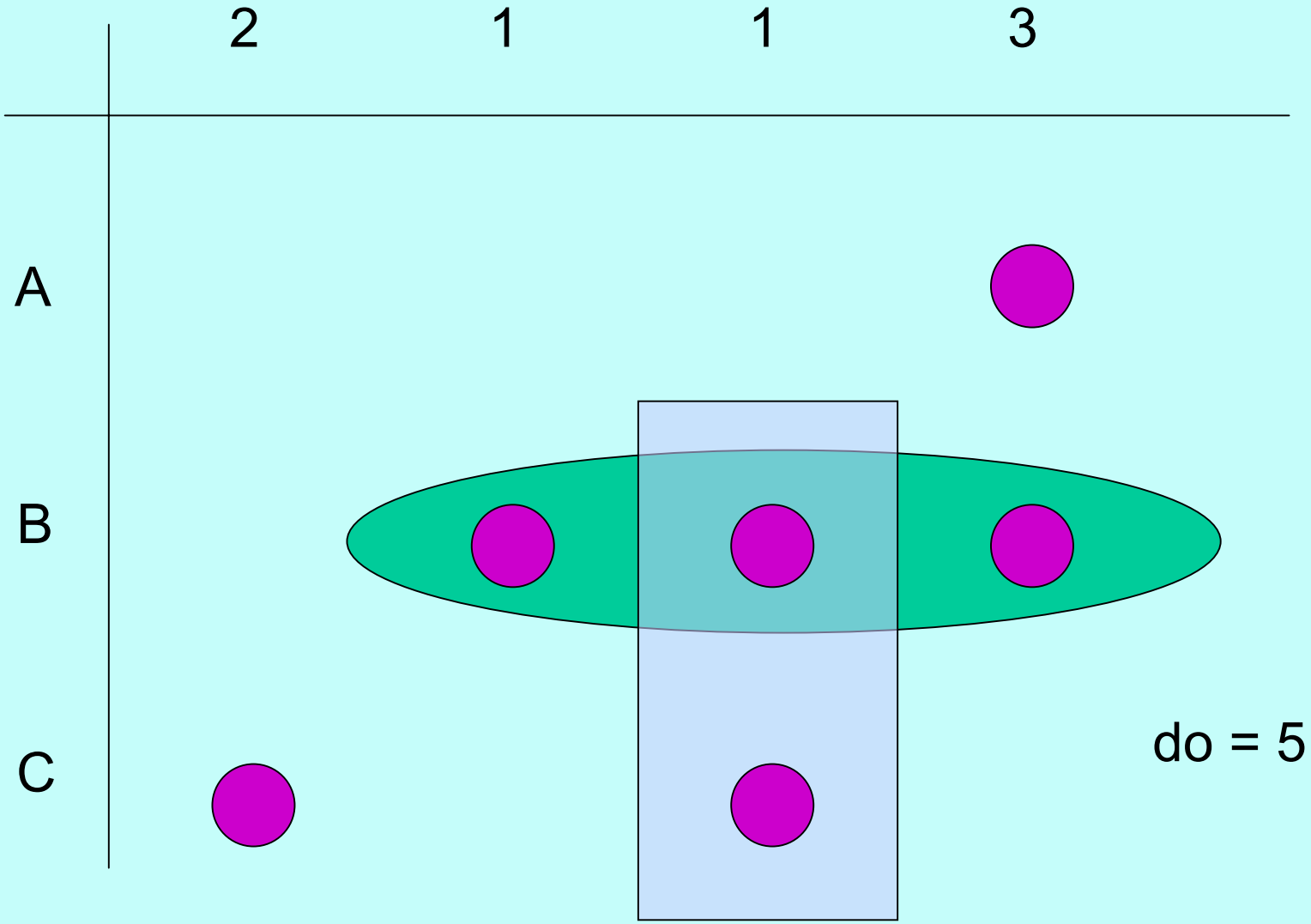
$$\lambda_A = \sum_{A : A \cap S \neq \emptyset} (-1)^{|A|-|S|+1} * v(S^c),$$

where S^c denotes the complement of S in X.

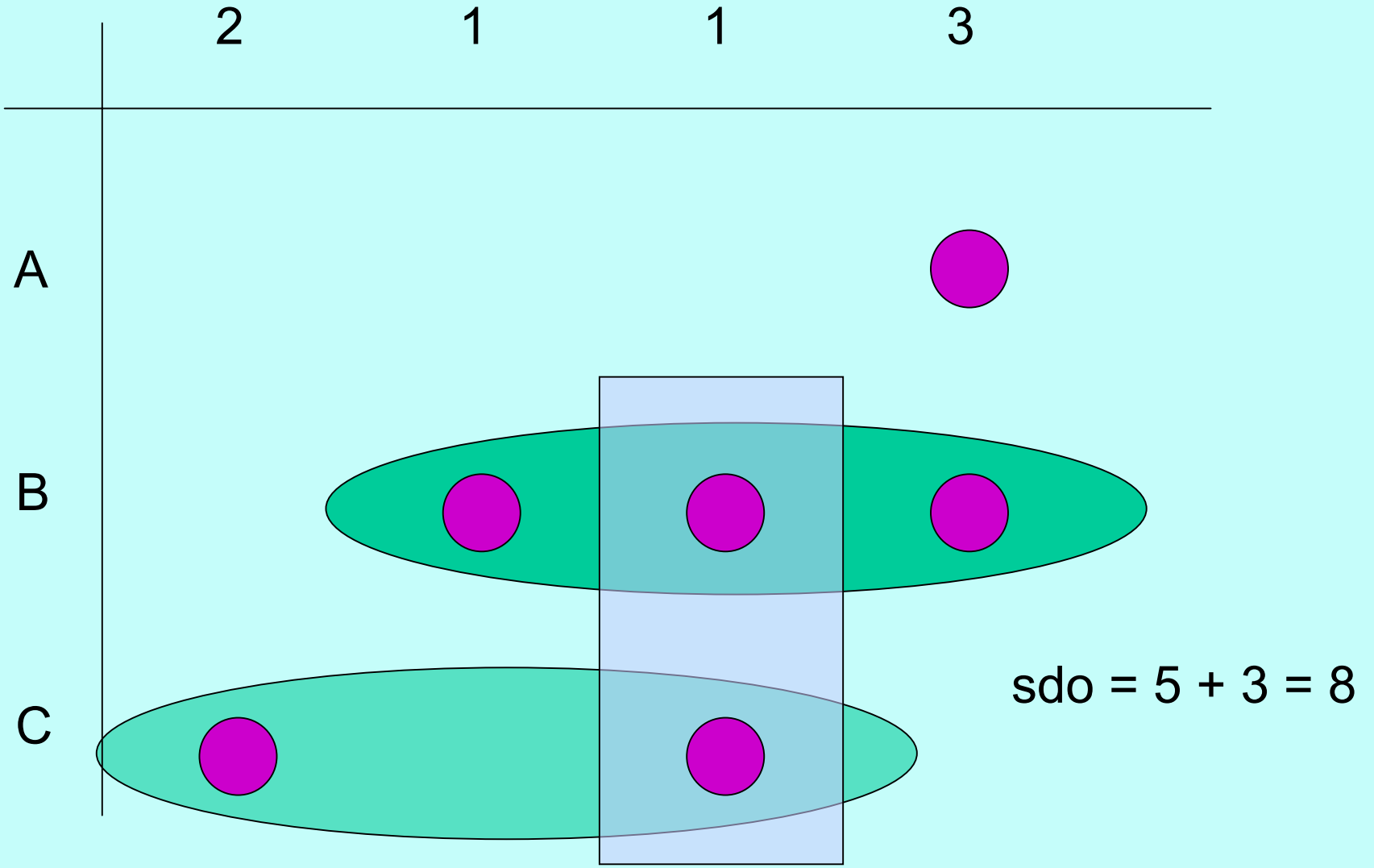
Diversity, Möebius function

sign	Variable description
$v(S)$	diversity of the set S. The elements of S are all objects in selected concept
λ_A	weight of selected concept - Möebius function
do_g	diversity of one specific object g
sdo	sum of diversities of objects
wca	weight of common attributes of object in selected concept
$impo_g$	the importance of object

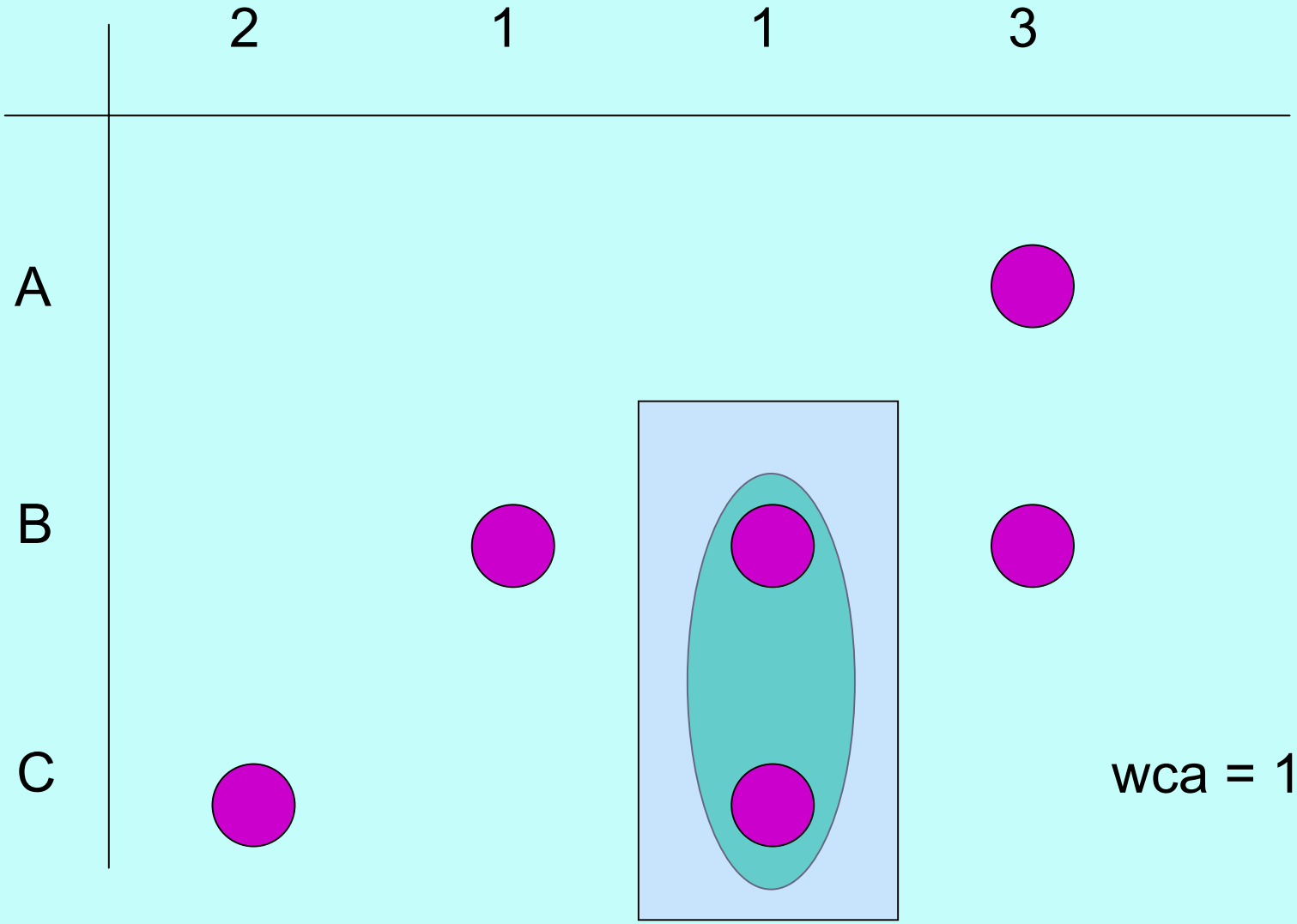
Diversity of object (B)



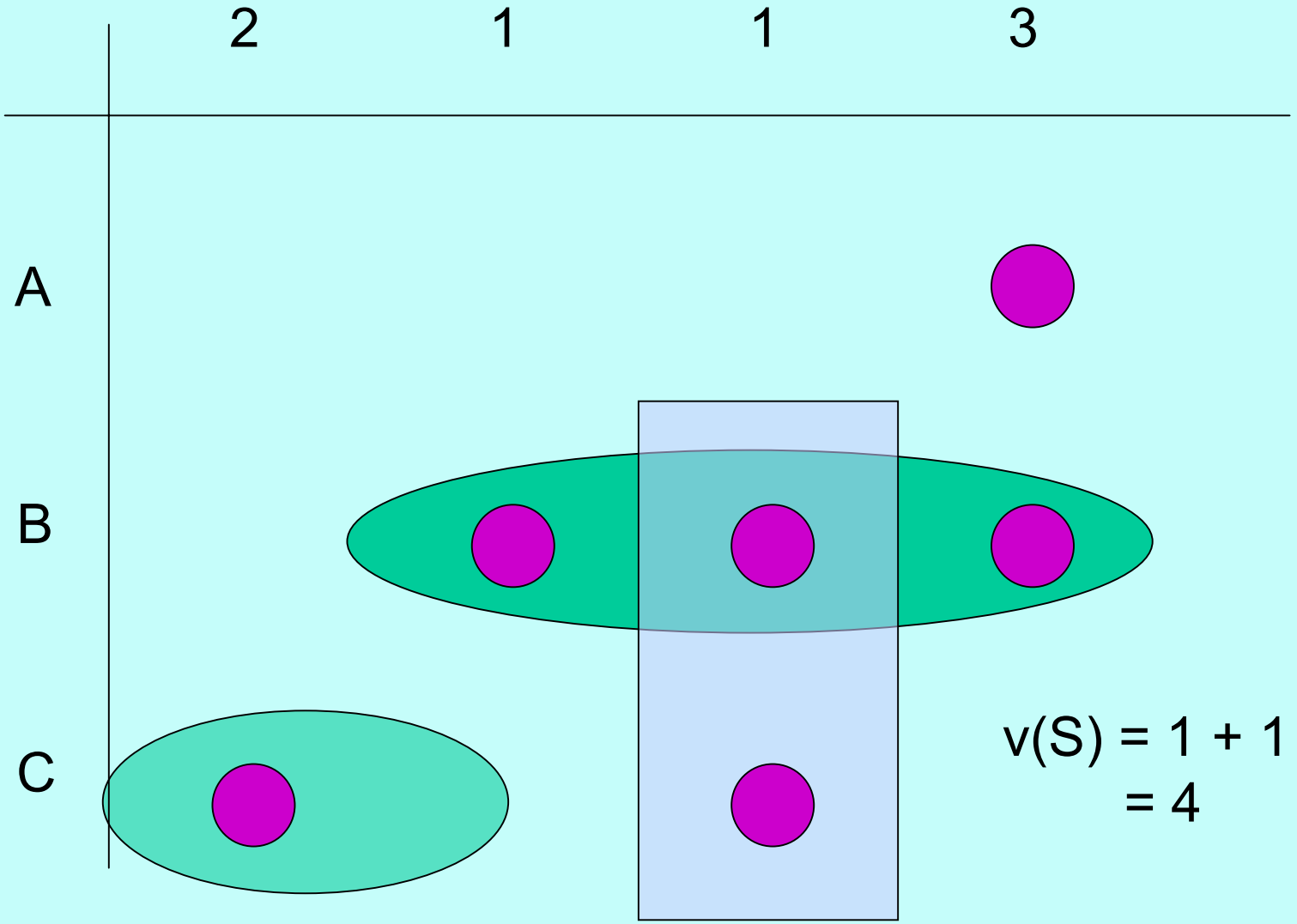
Sum of diversities of concept's objects



Weight of common attributes

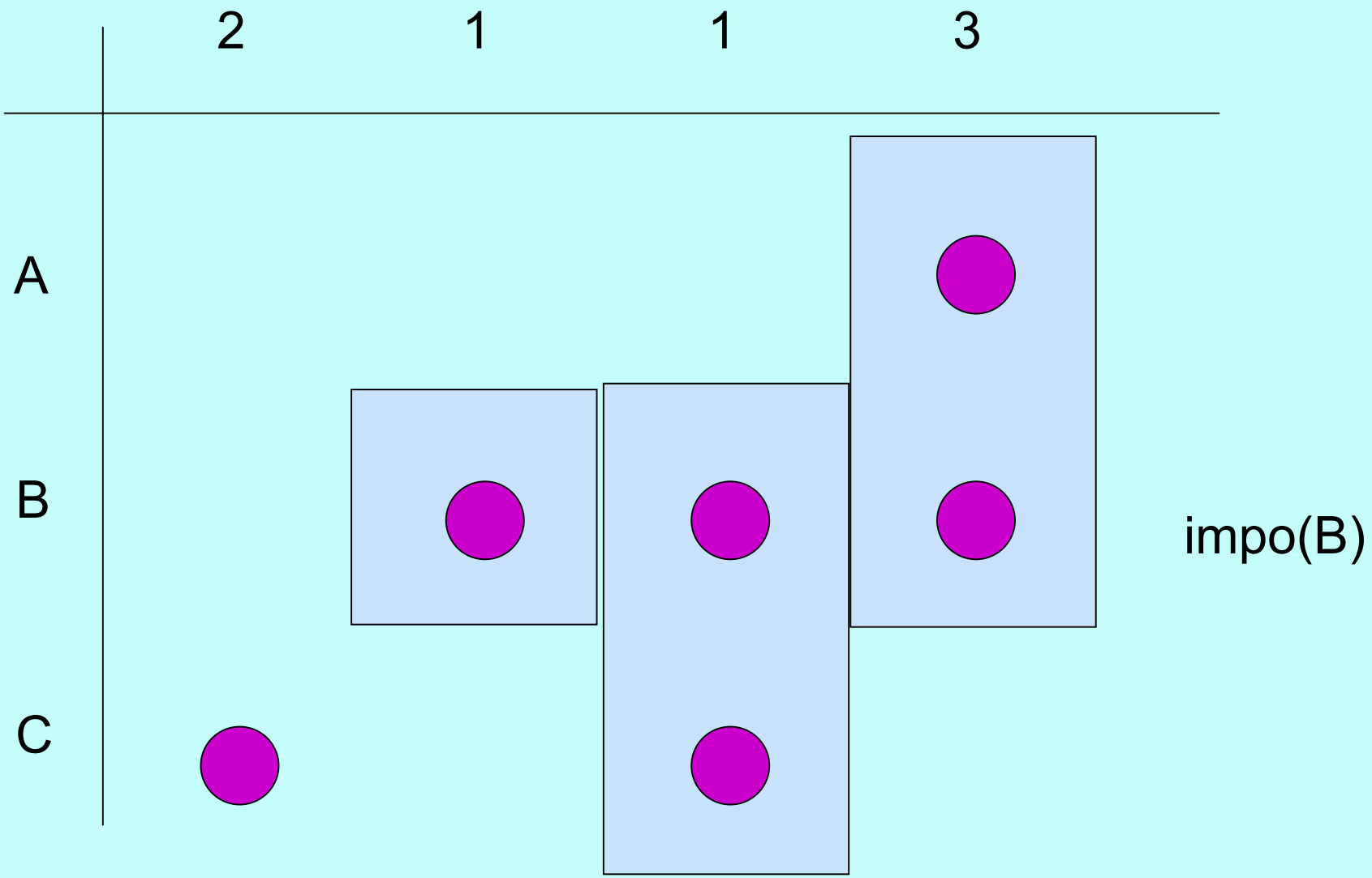


Diversity of set of concept objects

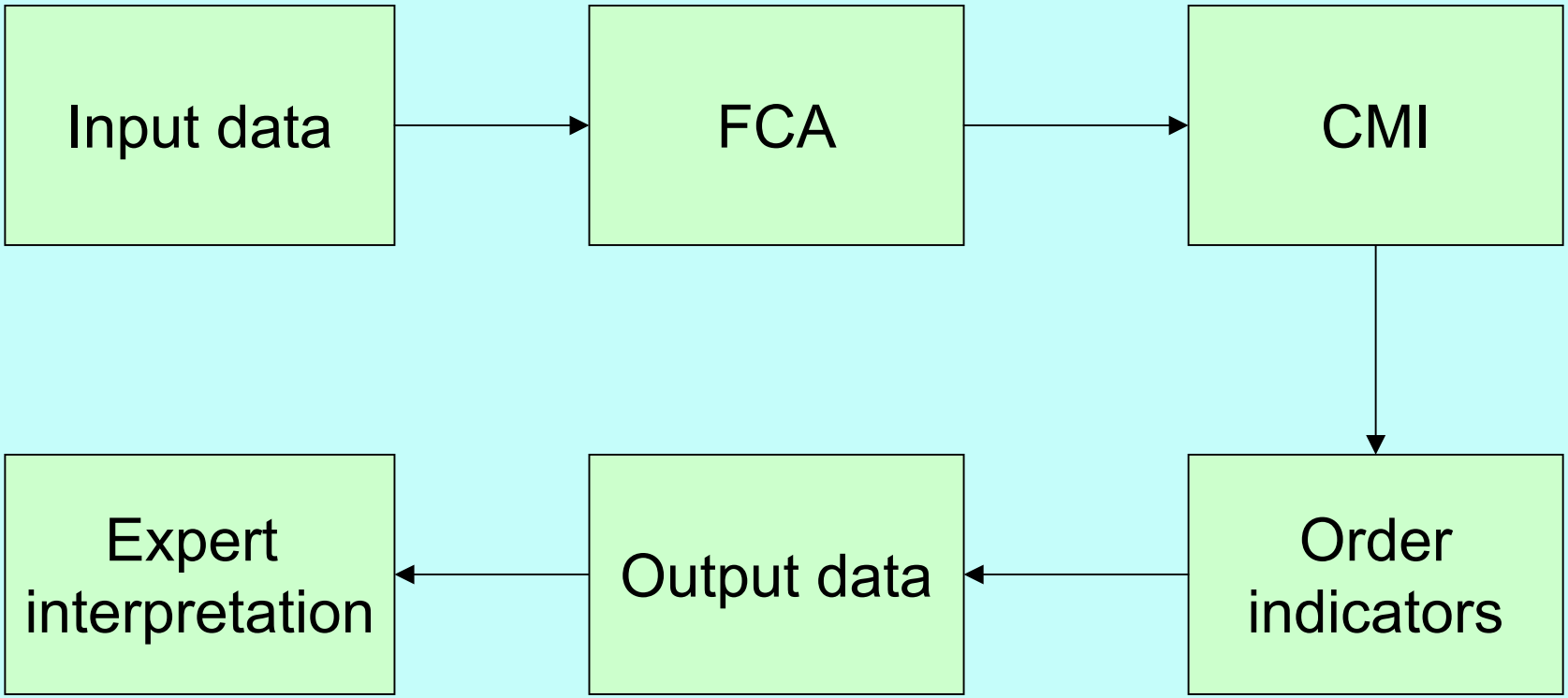


$v(S) = 1 + 1 + 2 + 3$
$= 4$

Importance of each object



Applications



Input data (RA1)

	E	G	N	S
1	7	0	0	4
2	8	5	0	0
3	8	3	0	2
4	5	0	0	5
5	7	4	0	0
6	3	0	0	7
7	4	0	0	7
8	5	0	0	5
9	9	0	5	2
10	9	0	4	2
11	9	0	0	2
12	7	4	5	0
13	4	0	0	2
14	4	4	0	0
15	0	0	9	0
16	4	2	6	0
17	7	0	0	3
18	3	0	4	0
19	2	0	4	0

Scaling data

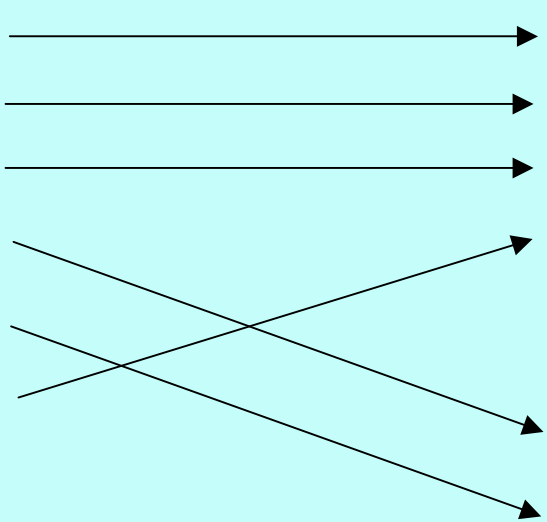
	e1	e2	e3	e4	e5	e6	e7	e8	e9	e10	g1	g2	g3	g4	g5	g6	g7	g8	g9	g10	n1	n2	n3	n4	n5	n6	n7	n8	n9	n10	s1	s2	s3	s4	s
1							1																											1	
2								1							1																				
3								1					1																			1			
4					1																														
5							1							1																					
6			1																																
7				1																															
8					1																														
9									1																1							1			
10									1															1								1			
11									1																							1			
12							1							1											1										
13				1																												1			
14				1										1																					
15																													1						
16				1								1														1									

CMI calculation

concepts	w	div	sdo	wca
{ ADB AOI AOO BCP BCT CCC CCI	{ }	0	166	535	0
{ ADB BCT ELH FMB HAD TSF }	{ e7 }	7	42	81	7
{ ADB GWD HMR HPH MVI NTA UF	{ g4 }	4	62	98	4
{ ADB }	{ e7 g4	0	11	11	11
{ AOI AOO GWD HMR UAL }	{ e8 }	8	37	73	8
{ AOI }	{ e8 g5	5	13	13	13
{ AOO EEA EEI EES ELP GDC NPA	{ g2 }	2	51	93	2
{ AOO GAT HAD HEU NPA RLC }	{ g3 }	3	61	83	3
{ AOO NPA }	{ g3 g2	0	25	30	5
{ AOO }	{ e8 g3	0	13	13	13
{ BCP CCE HAI }	{ g5 }	5	12	27	5
{ BCP CCE HEU HPC MVI RLC }	{ e5 }	5	36	69	5
{ BCP CCE }	{ e5 g5	0	10	20	10
{ BCT ELH EXP GDC HLE HPD NBI	{ g4 }	4	62	113	4

Order indicators

Order after application of our method		Sum of weights	
9127.135	NTA	24	NTA
8445.38	HMR	21	HMR
6589.766	HPD	19	HPD
5996.009	EEA	17	TSF
5719.889	HAD	17	HEU
5491.328	TSF	17	NPA
5342.943	ELH	16	EEA
5319.748	EEI	16	HAD
5316.073	GDC	16	ELH
5279.444	HEU	16	HLE
5041.946	UAL	15	EEI
4554.579	NPA	15	GDC
4439.619	HPC	15	HPC
4427.185	URT	15	RTT



The most important indicators in each area

	RA1		RA2		RA3	
	Indicator	**Weight**	**Indicator**	**Weight**	**Indicator**	**Weight**
1	Temperature	9127,135	spatial share of forest	5363,821	Unemployment rate	18129,64
2	net migration rate	8445,38	Unemployment rate	4632,171	Share of University graduated	15733,83
3	Population density	6589,766	net migration rate	4416,977	wage /avarge income	13943,61
4	Share of employment in Agriculture, forest and fishing	5996,009	population age structure	3756,424	Share of commuting population	13417,97
5	population age structure	5719,889	wage /avarge income	3376,268	population age structure	9241,468

	RA4		RA5		RA6	
	Indicator	**Weight**	**Indicator**	**Weight**	**Indicator**	**Weight**
1	Unemployment rate	21982,13	Standard gross margin	5664,613	Annual withdrawal of ground and surface water	5833,757
2	wage /avarge income	21974,98	Investment share in GDP	4162,936	Sustainable development strategy	5388,211
3	Share of University graduated	20198,65	population change in 10 years (1990-2000)	4139,425	Abundance of Selected Key Species	4300,805
4	population age structure	20155,66	Accommodation capacity	3696,442	GDP per capita	3519,612
5	Average salary	18740,85	Share of protected areas	3670,267	Unemployment rate	3477,18

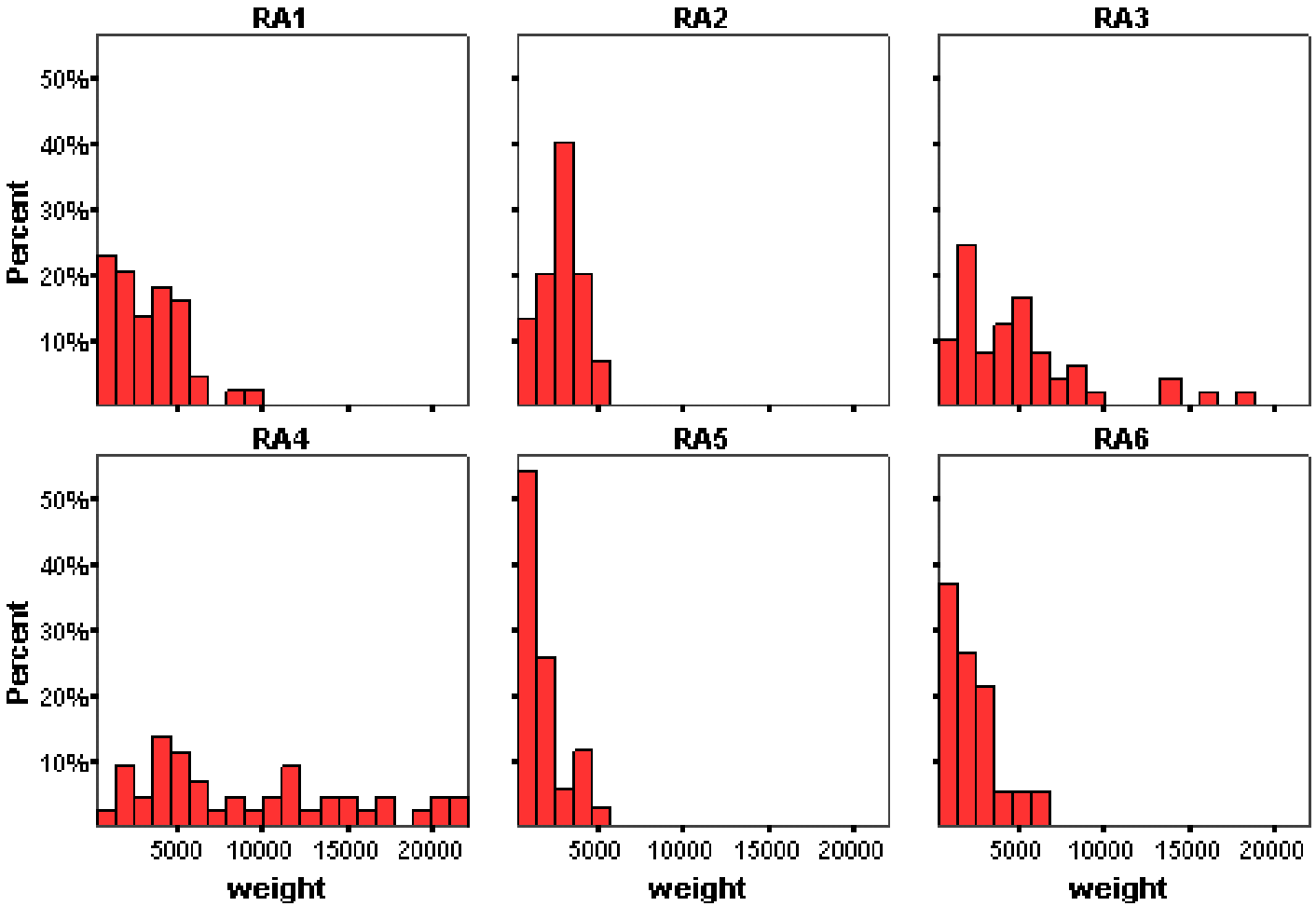
Conclusion

- Order based on concepts, weights of attributes and diversity of objects
- Help in methodology of indicators selection
- Interest in next projects
- Checking and comparison with other statistical methods

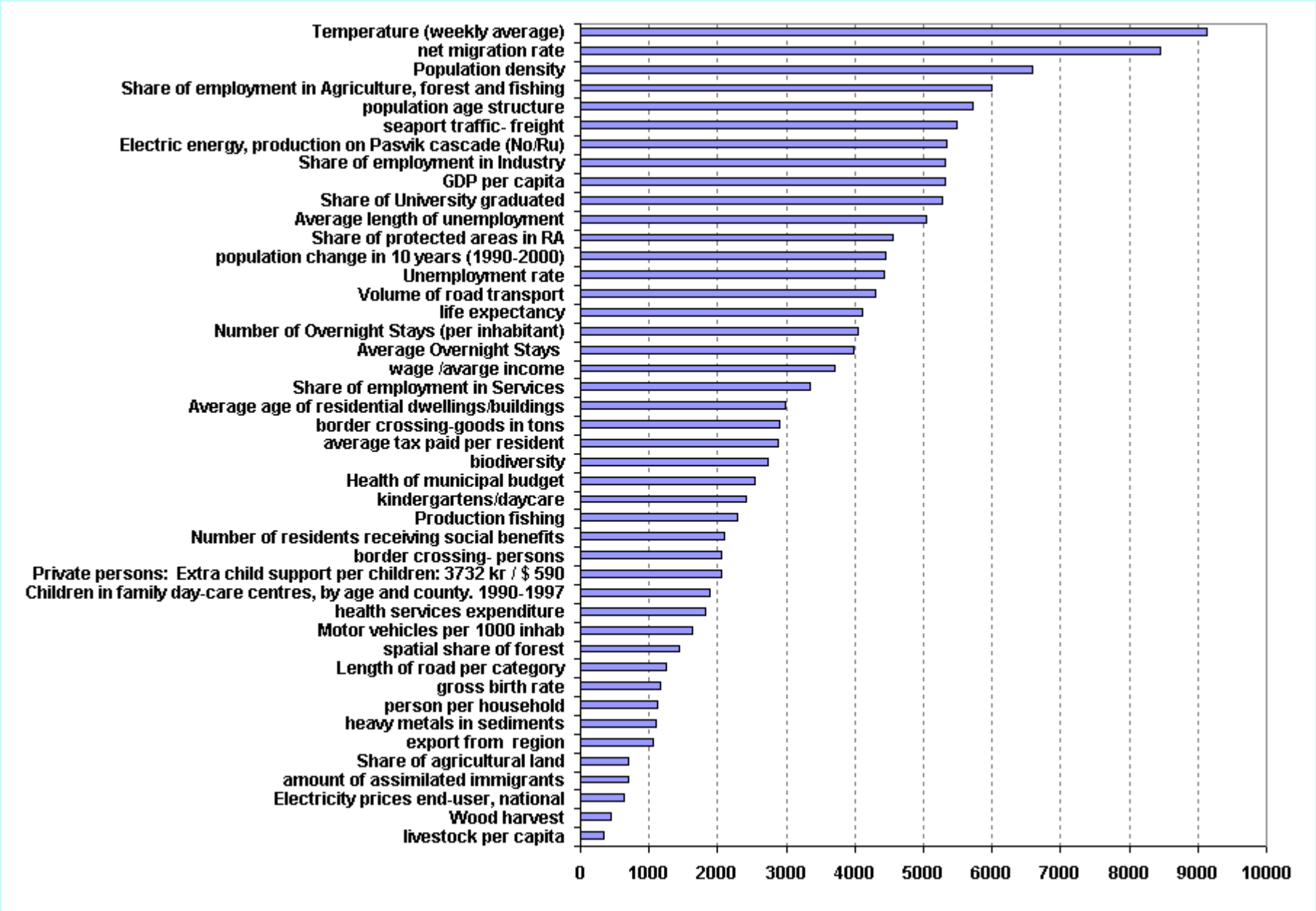
References

1. B. Ganter and R. Wille. *Formal Concept Analysis.* Springer-Verlag Berlin Heidelberg, 1999.
2. K. Nehring. A theory of diversity. *Econometrica 70*, pages 1155–1198, 2002.
3. K. Nehring and C. Puppe. Modelling phylogenetic diversity. *Resource and Energy Economics*, 2002.
4. K. Nehring and C. Puppe. Diversity and dissimilarity in lines and hierarchies. *Mathematical Social Sciences 45*, pages 167–183, 2003.
5. I. Vondrak and V. Snasel. Using concept lattices for organizational structure analysis. *In Proceedings of European Concurrent Engineering Conference ECEC '02 (Modena, Italy)*, 2002.

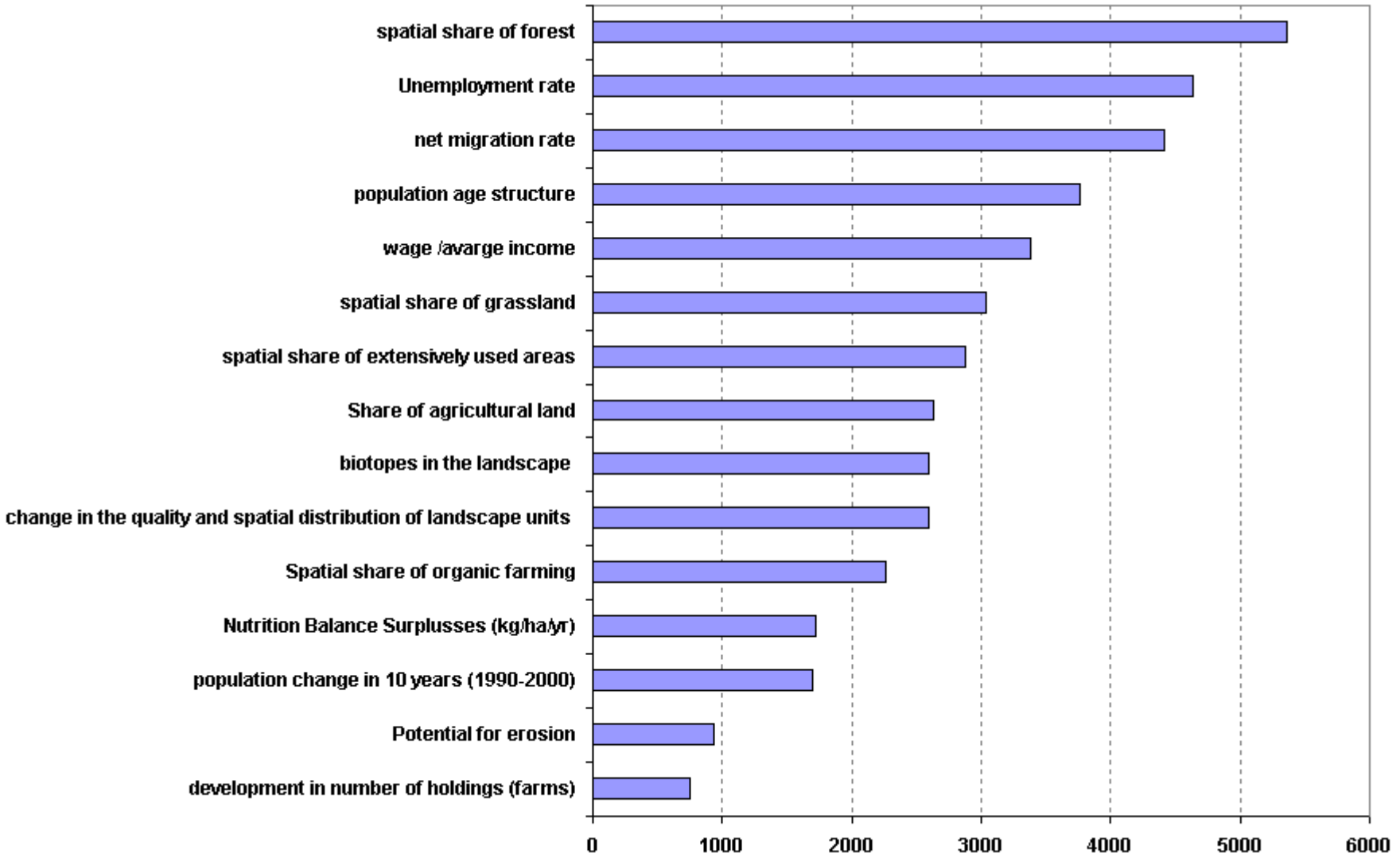
Histogram


RA1
RA2
RA3
RA4
RA5
RA6
Percent
50%
40%
30%
20%
10%
5000
10000
15000
20000
weight

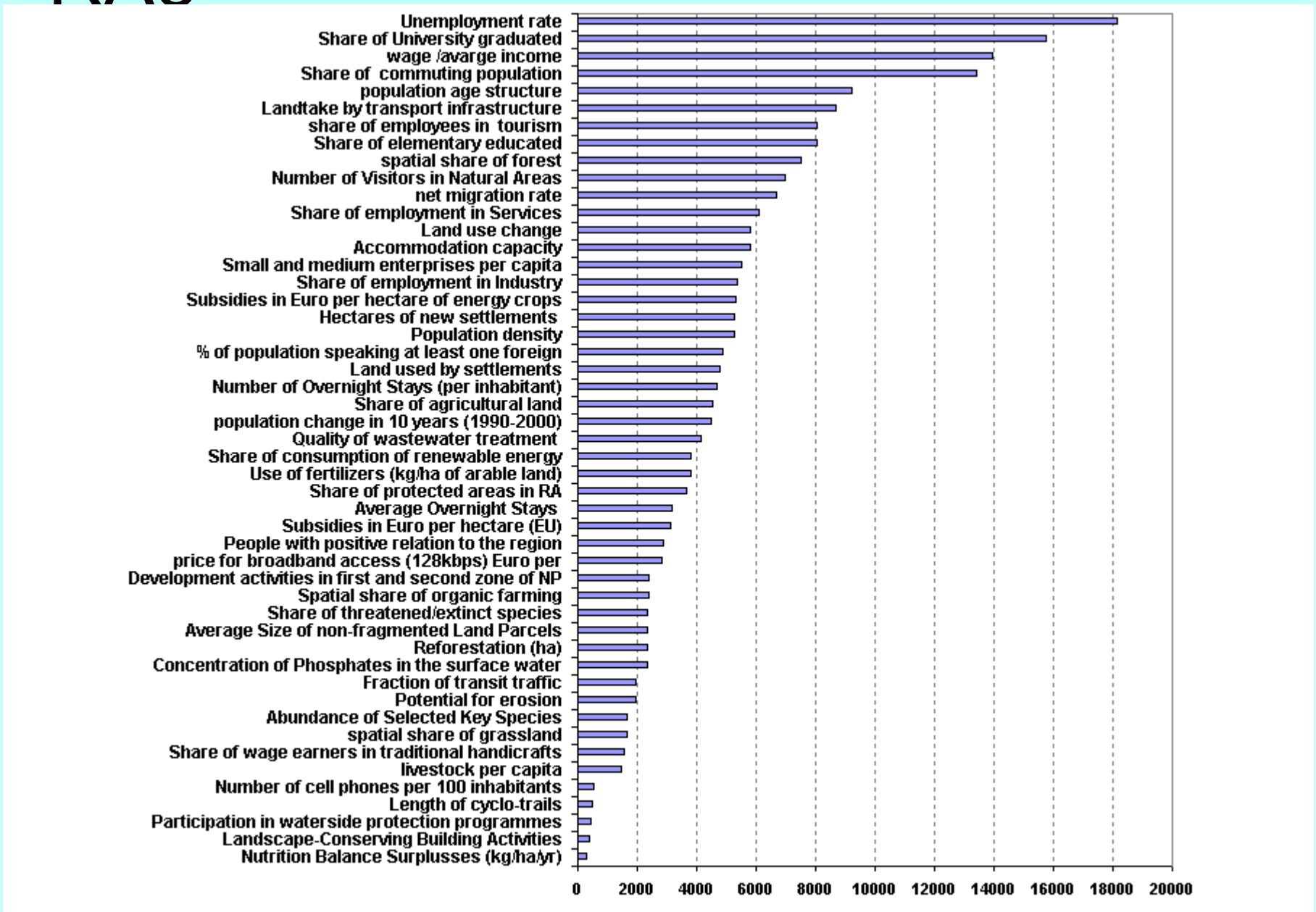
RA1


Temperature (weekly average)
net migration rate
Population density
Share of employment in Agriculture, forest and fishing
population age structure
seaport traffic- freight
Electric energy, production on Pasvik cascade (No/Ru)
Share of employment in Industry
GDP per capita
Share of University graduated
Average length of unemployment
Share of protected areas in RA
population change in 10 years (1990-2000)
Unemployment rate
Volume of road transport
life expectancy
Number of Overnight Stays (per inhabitant)
Average Overnight Stays
wage /avarge income
Share of employment in Services
Average age of residential dwellings/buildings
border crossing-goods in tons
average tax paid per resident
biodiversity
Health of municipal budget
kindergartens/daycare
Production fishing
Number of residents receiving social benefits
border crossing- persons
Private persons: Extra child support per children: 3732 kr / $ 590
Children in family day-care centres, by age and county. 1990-1997
health services expenditure
Motor vehicles per 1000 inhab
spatial share of forest
Length of road per category
gross birth rate
person per household
heavy metals in sediments
export from region
Share of agricultural land
amount of assimilated immigrants
Electricity prices end-user, national
Wood harvest
livestock per capita
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
10000

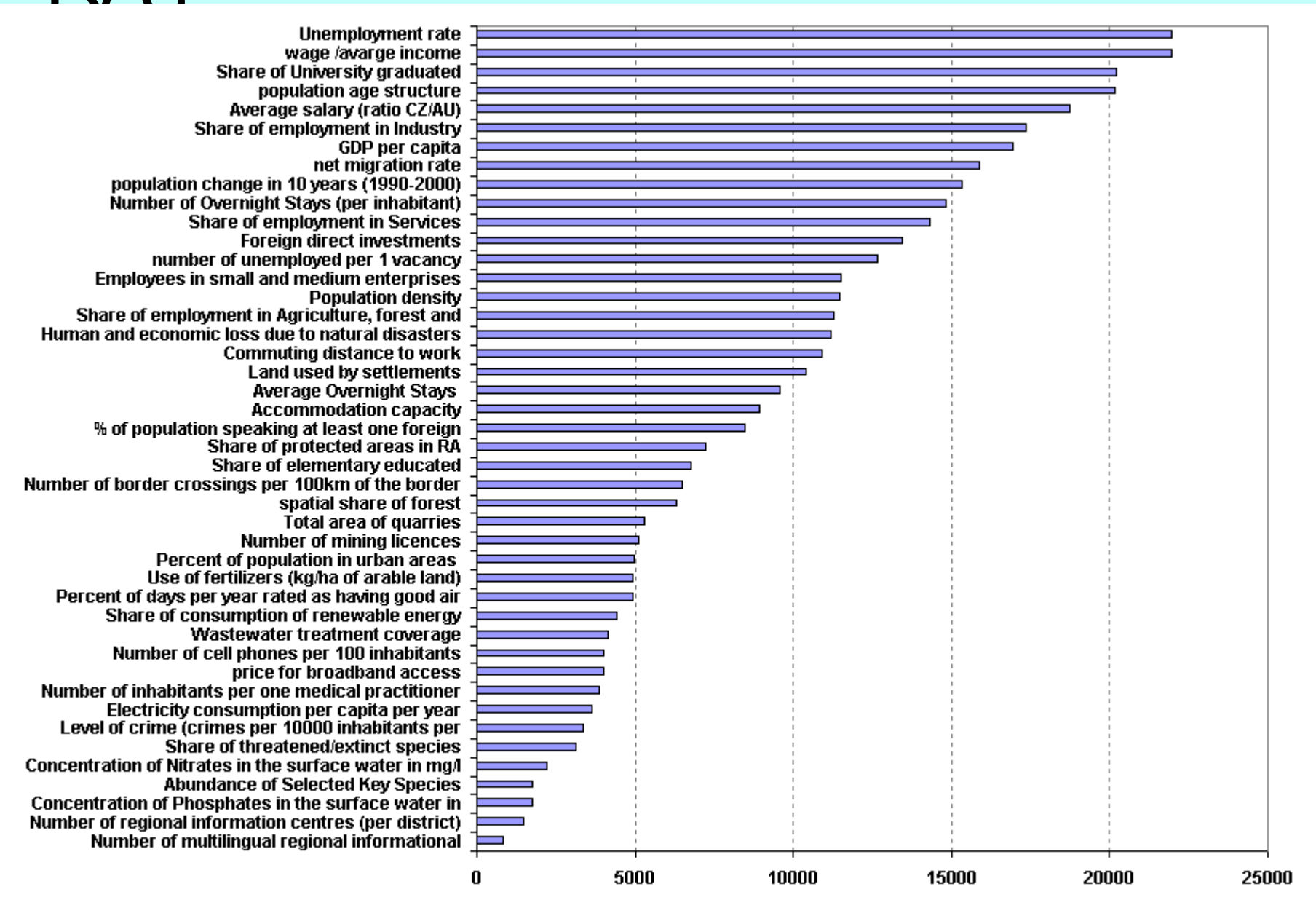
RA2


spatial share of forest
Unemployment rate
net migration rate
population age structure
wage /avarge income
spatial share of grassland
spatial share of extensively used areas
Share of agricultural land
biotopes in the landscape
change in the quality and spatial distribution of landscape units
Spatial share of organic farming
Nutrition Balance Surplusses (kg/ha/yr)
population change in 10 years (1990-2000)
Potential for erosion
development in number of holdings (farms)
0
1000
2000
3000
4000
5000
6000

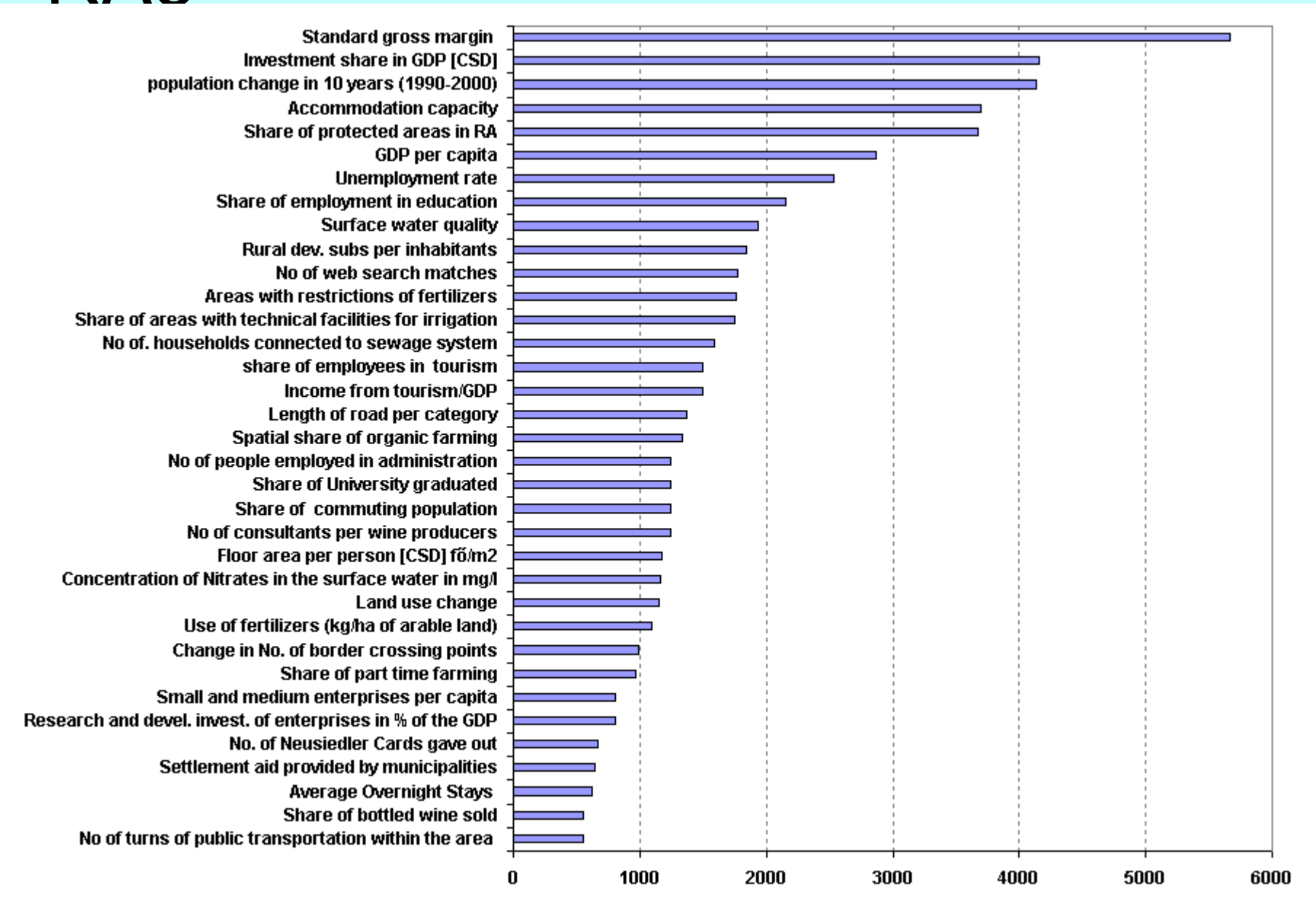
RA3


Unemployment rate
Share of University graduated
wage /avarge income
Share of commuting population
population age structure
Landtake by transport infrastructure
share of employees in tourism
Share of elementary educated
spatial share of forest
Number of Visitors in Natural Areas
net migration rate
Share of employment in Services
Land use change
Accommodation capacity
Small and medium enterprises per capita
Share of employment in Industry
Subsidies in Euro per hectare of energy crops
Hectares of new settlements
Population density
% of population speaking at least one foreign
Land used by settlements
Number of Overnight Stays (per inhabitant)
Share of agricultural land
population change in 10 years (1990-2000)
Quality of wastewater treatment
Share of consumption of renewable energy
Use of fertilizers (kg/ha of arable land)
Share of protected areas in RA
Average Overnight Stays
Subsidies in Euro per hectare (EU)
People with positive relation to the region
price for broadband access (128kbps) Euro per
Development activities in first and second zone of NP
Spatial share of organic farming
Share of threatened/extinct species
Average Size of non-fragmented Land Parcels
Reforestation (ha)
Concentration of Phosphates in the surface water
Fraction of transit traffic
Potential for erosion
Abundance of Selected Key Species
spatial share of grassland
Share of wage earners in traditional handicrafts
livestock per capita
Number of cell phones per 100 inhabitants
Length of cyclo-trails
Participation in waterside protection programmes
Landscape-Conserving Building Activities
Nutrition Balance Surplusses (kg/ha/yr)
0
2000
4000
6000
8000
10000
12000
14000
16000
18000
20000

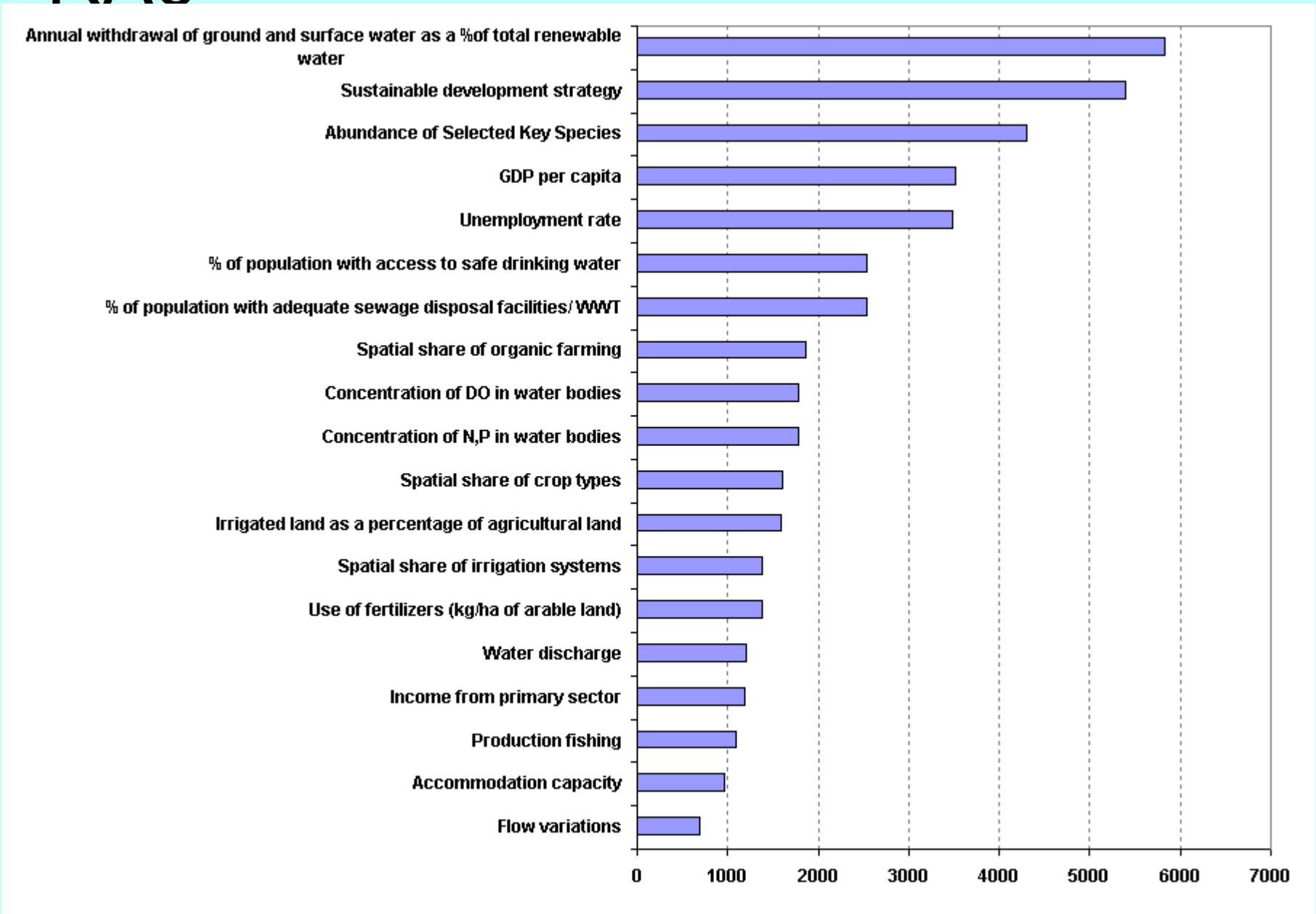
RA4


Unemployment rate
wage /avarge income
Share of University graduated
population age structure
Average salary (ratio CZ/AU)
Share of employment in Industry
GDP per capita
net migration rate
population change in 10 years (1990-2000)
Number of Overnight Stays (per inhabitant)
Share of employment in Services
Foreign direct investments
number of unemployed per 1 vacancy
Employees in small and medium enterprises
Population density
Share of employment in Agriculture, forest and
Human and economic loss due to natural disasters
Commuting distance to work
Land used by settlements
Average Overnight Stays
Accommodation capacity
% of population speaking at least one foreign
Share of protected areas in RA
Share of elementary educated
Number of border crossings per 100km of the border
spatial share of forest
Total area of quarries
Number of mining licences
Percent of population in urban areas
Use of fertilizers (kg/ha of arable land)
Percent of days per year rated as having good air
Share of consumption of renewable energy
Wastewater treatment coverage
Number of cell phones per 100 inhabitants
price for broadband access
Number of inhabitants per one medical practitioner
Electricity consumption per capita per year
Level of crime (crimes per 10000 inhabitants per
Share of threatened/extinct species
Concentration of Nitrates in the surface water in mg/l
Abundance of Selected Key Species
Concentration of Phosphates in the surface water in
Number of regional information centres (per district)
Number of multilingual regional informational
0
5000
10000
15000
20000
25000

RA5


Standard gross margin
Investment share in GDP [CSD]
population change in 10 years (1990-2000)
Accommodation capacity
Share of protected areas in RA
GDP per capita
Unemployment rate
Share of employment in education
Surface water quality
Rural dev. subs per inhabitants
No of web search matches
Areas with restrictions of fertilizers
Share of areas with technical facilities for irrigation
No of. households connected to sewage system
share of employees in tourism
Income from tourism/GDP
Length of road per category
Spatial share of organic farming
No of people employed in administration
Share of University graduated
Share of commuting population
No of consultants per wine producers
Floor area per person [CSD] fő/m2
Concentration of Nitrates in the surface water in mg/l
Land use change
Use of fertilizers (kg/ha of arable land)
Change in No. of border crossing points
Share of part time farming
Small and medium enterprises per capita
Research and devel. invest. of enterprises in % of the GDP
No. of Neusiedler Cards gave out
Settlement aid provided by municipalities
Average Overnight Stays
Share of bottled wine sold
No of turns of public transportation within the area
0
1000
2000
3000
4000
5000
6000

RA6


Annual withdrawal of ground and surface water as a %of total renewable water
Sustainable development strategy
Abundance of Selected Key Species
GDP per capita
Unemployment rate
% of population with access to safe drinking water
% of population with adequate sewage disposal facilities/ WWT
Spatial share of organic farming
Concentration of DO in water bodies
Concentration of N,P in water bodies
Spatial share of crop types
Irrigated land as a percentage of agricultural land
Spatial share of irrigation systems
Use of fertilizers (kg/ha of arable land)
Water discharge
Income from primary sector
Production fishing
Accommodation capacity
Flow variations
0
1000
2000
3000
4000
5000
6000
7000