

Exploratory Spatial Data Analysis

Part IV
Computationally enhanced visualisation

1

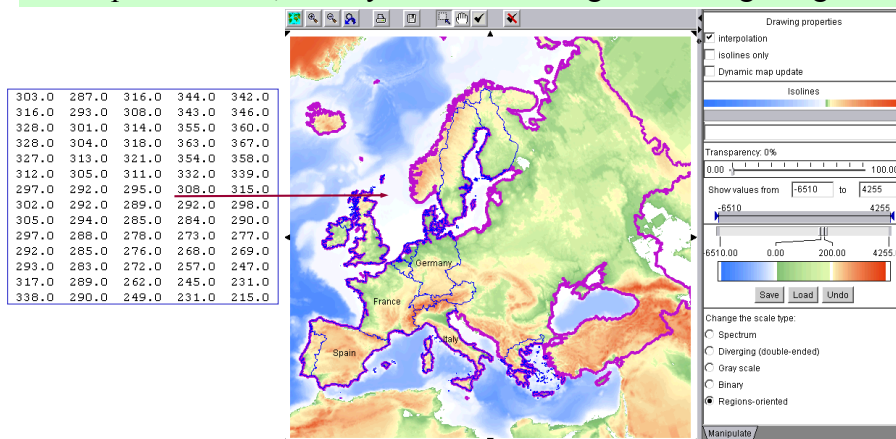
Contents

- Tools for analysis of grid (raster) data
 - Grid visualisation, interactive manipulation
 - Grid-based computations
 - Grid-to-table transformation
- Data mining
 - Introduction to data mining
 - Clustering
 - Classification trees

2

Grid Data

Typically, represent continuous phenomena. Values are measured at sample locations, usually in nodes of a regular rectangular grid.



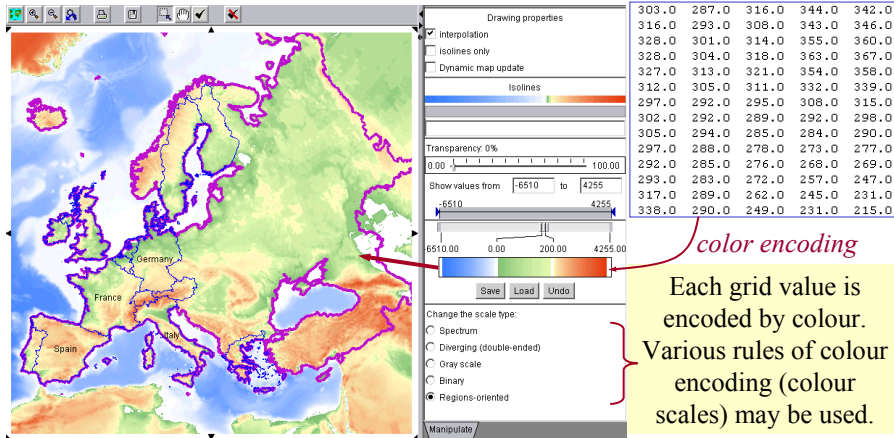
3

Grid Data Properties

- One grid contains values of a single attribute.
- The values refer to regularly spaced locations or uniform rectangular cells rather than to any real spatial objects or to administrative units.
- Values between grid nodes are derived by means of interpolation.

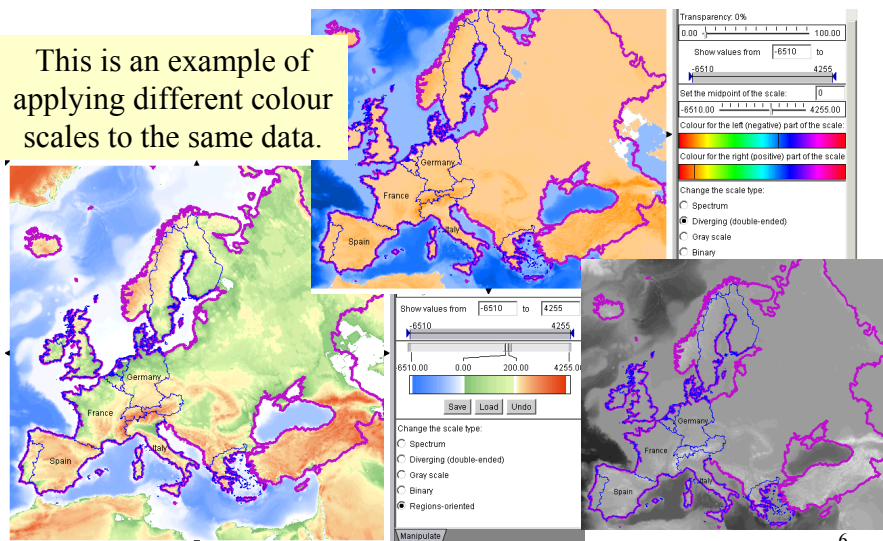
4

Grid Data Visualisation (1)



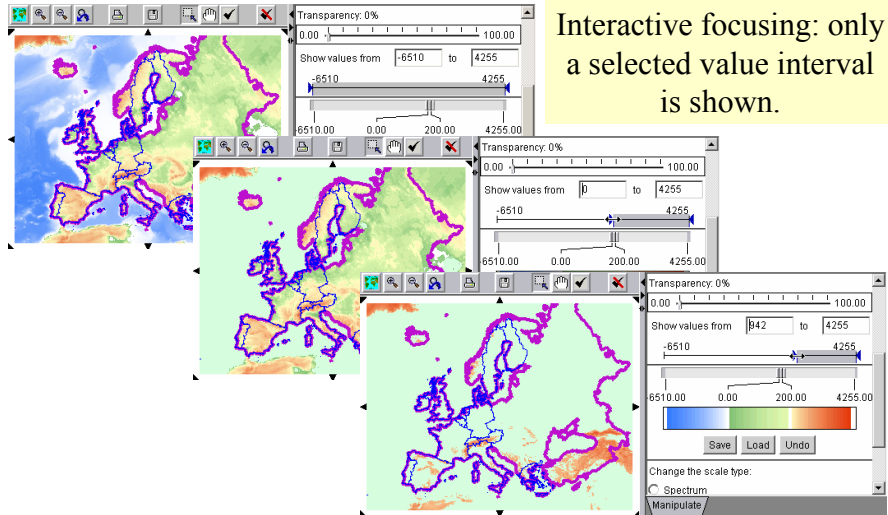
5

Grid Data Visualisation (2)



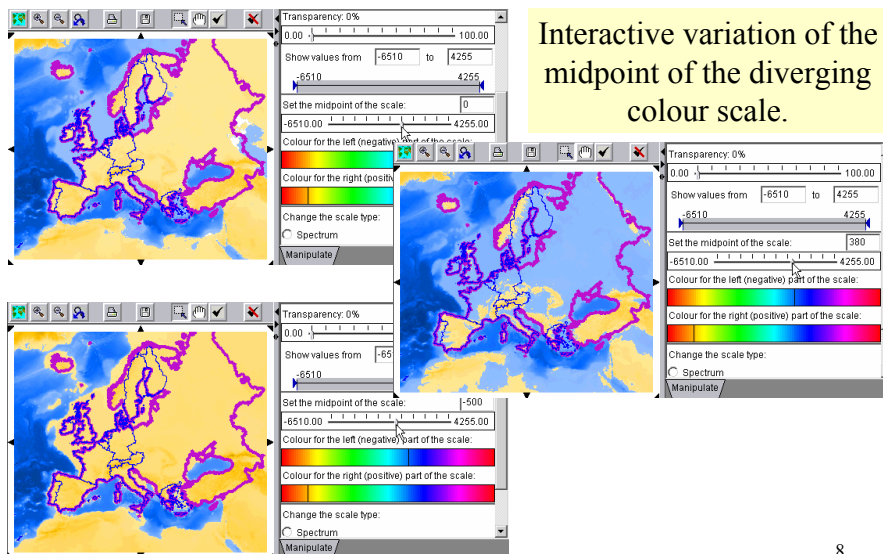
6

Manipulating Grid Visualisation 1



7

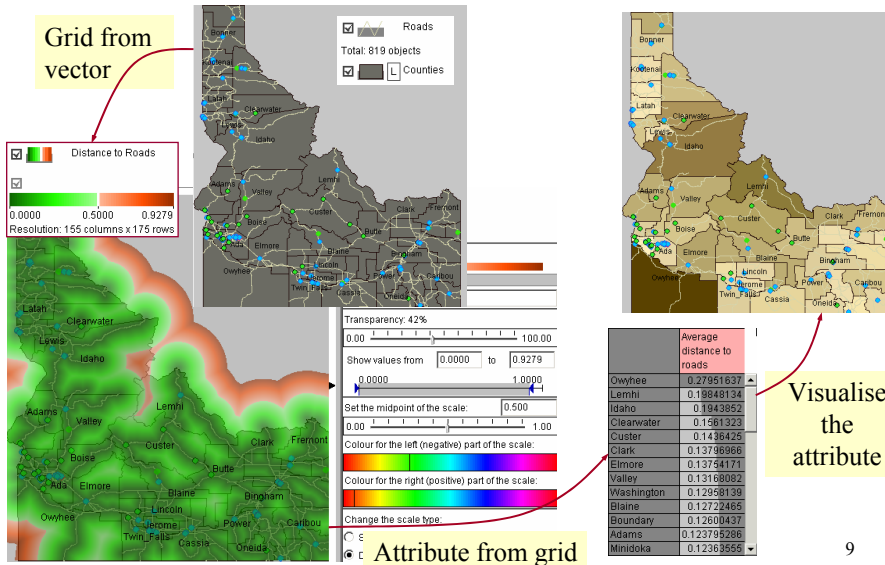
Manipulating Grid Visualisation 2



8

Use of Grids for Data Derivation

Grid from vector



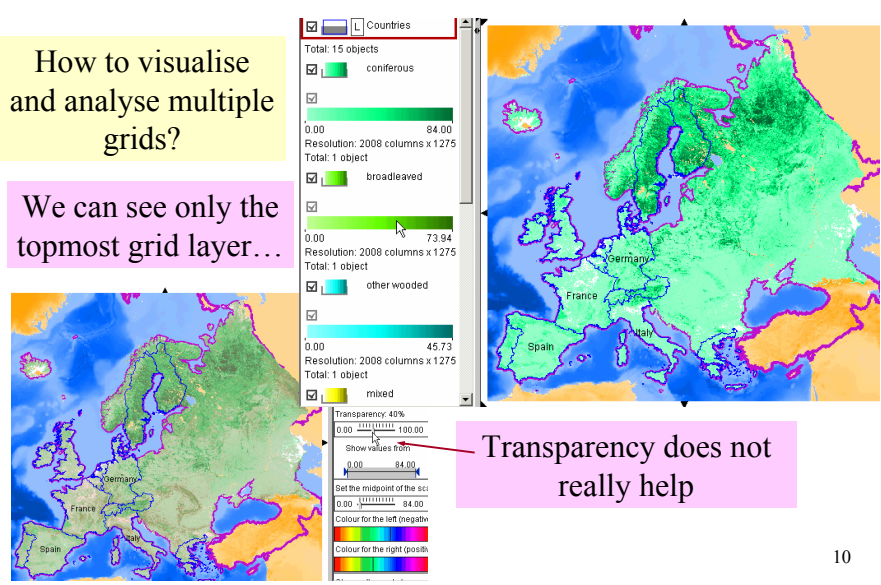
Visualise the attribute

Attribute from grid

Some Problems with Grids...

How to visualise and analyse multiple grids?

We can see only the topmost grid layer...



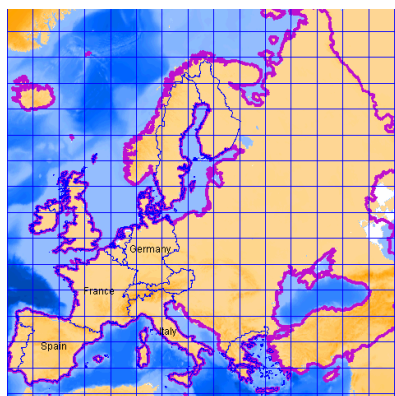
Transparency does not really help

Interactive Grid-Table Bridge (1)

Step 1. Introduce an artificial territory division.
It is convenient to have uniform areas, e.g. equal-size rectangles.

11

Interactive Grid-Table Bridge (2)



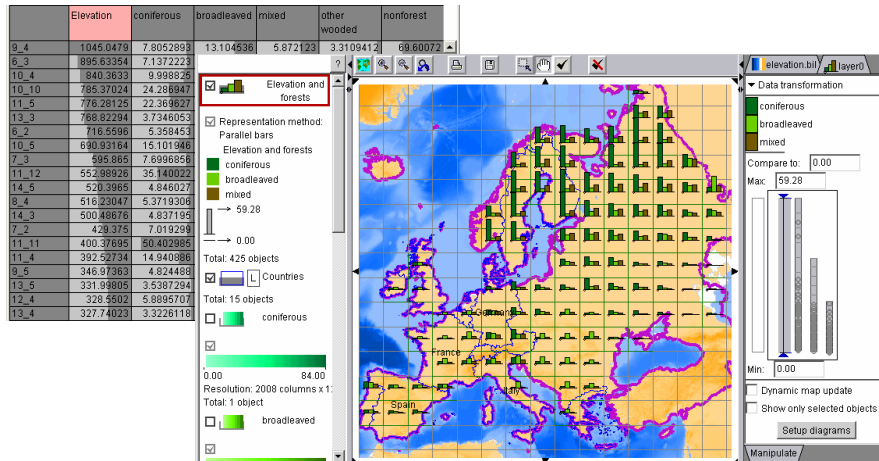
| | Elevation | coniferous | broadleaved | mixed | other wooded | nonforest |
|-------|-----------|------------|-------------|-----------|--------------|-----------|
| 9_4 | 1045.0479 | 7.8052893 | 13.104536 | 5.872123 | 3.3109412 | 68.80072 |
| 8_3 | 895.63354 | 7.1372223 | 7.4486217 | 2.7511675 | 20.408121 | 61.964834 |
| 10_4 | 840.3633 | 9.998826 | 12.937923 | 5.360093 | 8.05673 | 63.849795 |
| 10_10 | 785.37024 | 24.288947 | 6.970478 | 7.457732 | 2.552092 | 58.571743 |
| 11_5 | 776.28126 | 22.369827 | 11.365161 | 10.576998 | 2.7555962 | 52.743294 |
| 13_3 | 768.82294 | 3.7346053 | 14.309546 | 4.090813 | 9.760096 | 67.81471 |
| 8_2 | 716.5596 | 5.359453 | 4.5545139 | 1.521516 | 28.371569 | 59.982135 |
| 10_5 | 690.93164 | 15.101946 | 12.330395 | 8.648095 | 2.7672572 | 60.91224 |
| 7_3 | 595.865 | 7.6996956 | 8.176775 | 2.6910169 | 17.018682 | 63.10536 |
| 11_12 | 552.98926 | 35.140022 | 6.285668 | 8.712789 | 2.0118082 | 47.764194 |
| 14_5 | 520.3966 | 4.846027 | 17.05882 | 6.5686135 | 2.3491848 | 68.97509 |
| 8_4 | 516.23047 | 5.3719306 | 13.094806 | 2.9343758 | 2.6425462 | 75.4853 |
| 14_3 | 500.48676 | 4.837195 | 11.139068 | 4.2039137 | 6.8322487 | 72.857166 |
| 7_2 | 428.375 | 7.019299 | 3.918073 | 1.3403729 | 30.82854 | 56.656293 |
| 11_11 | 400.37695 | 50.402965 | 7.506363 | 1.3856096 | 2.081328 | 26.23074 |
| 11_4 | 392.52734 | 14.940388 | 18.83855 | 3.606678 | 6.1169369 | 50.473978 |
| 9_5 | 346.97363 | 4.824483 | 16.084341 | 5.6440325 | 2.5189521 | 70.59787 |
| 13_5 | 331.99806 | 3.5387294 | 13.52229 | 4.317274 | 2.6509849 | 75.72688 |
| 12_4 | 328.5502 | 5.8895707 | 15.99362 | 6.998373 | 3.126134 | 68.179375 |
| 13_4 | 327.74023 | 3.3226118 | 14.132109 | 4.1035585 | 3.0310817 | 75.17449 |

Step 2. Derive attributes of the rectangular areas from the grids

...like we did before for the counties of Idaho

12

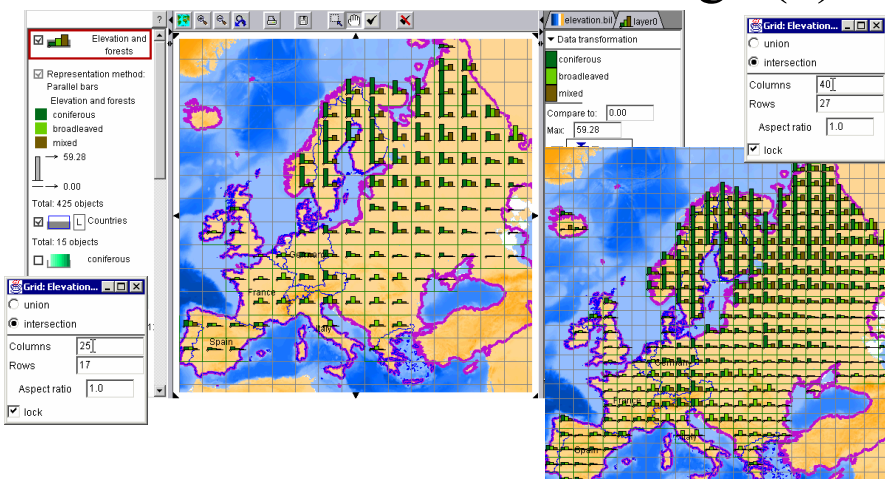
Interactive Grid-Table Bridge (3)



Step 3. Visualise and analyse the resulting attributes as usual

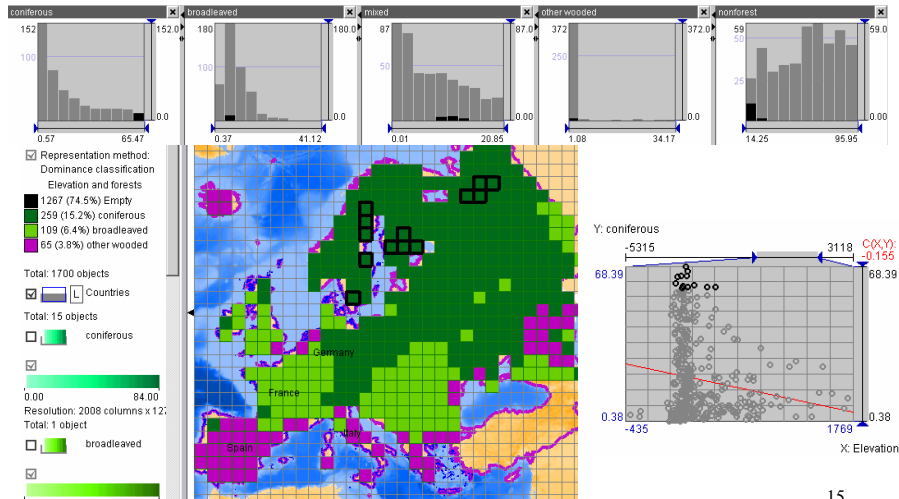
13

Interactive Grid-Table Bridge (4)



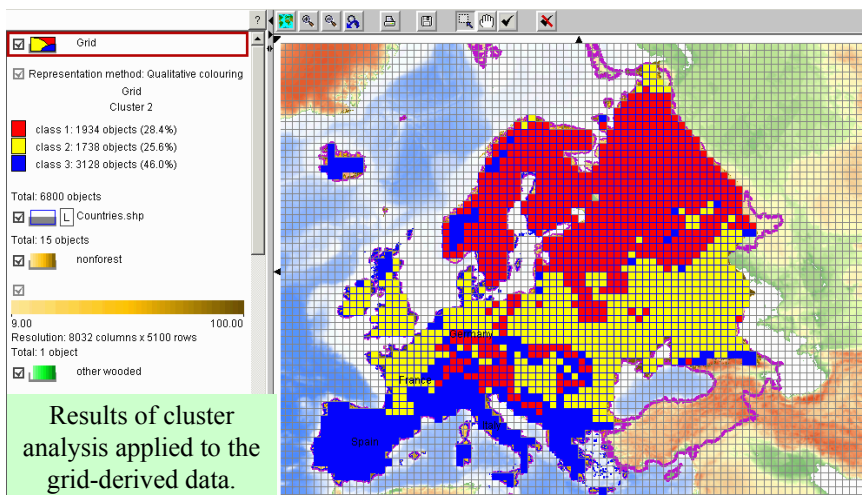
The sizes of the rectangles (grid-table bridge resolution) can be interactively changed. The attributes are automatically re-computed and the displays updated.

Further Possibilities for Analysis of Grid-Derived Attributes



15

Yet More Analysis: Data Mining



16

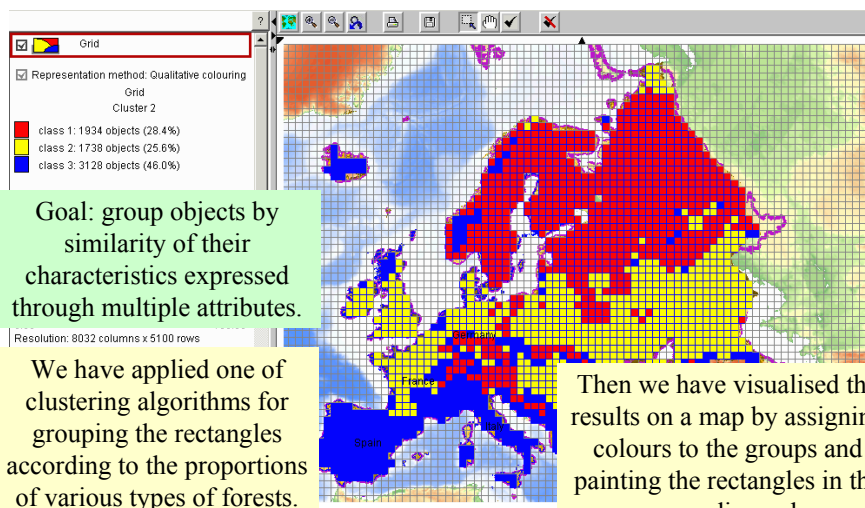
EDA vs. Data Mining

- Goal: revealing unknowns in data
- Approach: help a human analyst by means of
 - Data visualisation
 - High user interaction

- Goal: revealing unknowns in data
 - Approach: automatic extraction of patterns, rules, dependencies by means of
 - Mathematics-based methods
- But: a human analyst needs to interpret the results

17

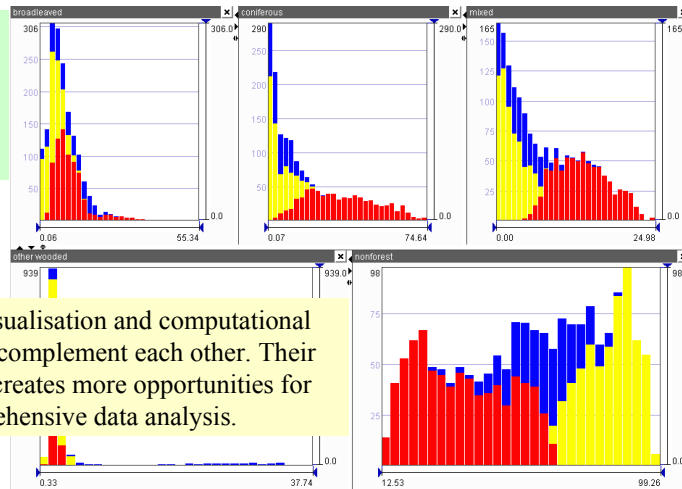
Cluster Analysis



18

Visualisation and Interpretation of Cluster Analysis Results

Visualisation helps us to understand how the groups are built.



In general, visualisation and computational methods well complement each other. Their combination creates more opportunities for comprehensive data analysis.

19

Summary

This lecture was supposed to

- introduce some computational techniques that can effectively complement data visualisation
- show that data transformation into another form (e.g. grid → table) may bring new opportunities for analysis
- acquaint with the discipline of data mining – automated exploration of data by means of mathematical methods

20