

# Exploratory Spatial Data Analysis

## Part II Dynamically Linked Views

1

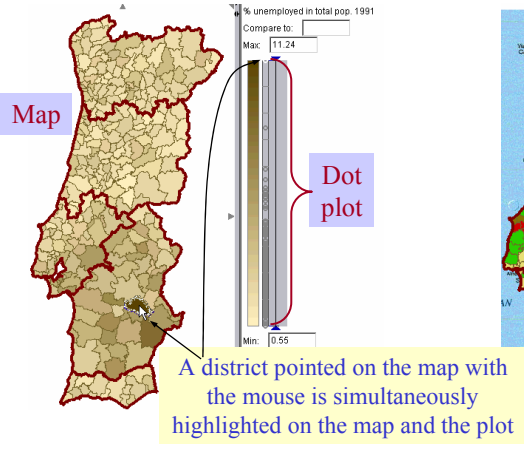
## Contents

- Introduction: why to use non-cartographic data displays
- Display linking by object highlighting
- Dynamic Query
- Object classification and class propagation
- Use of non-cartographic displays for classification: cumulative curves

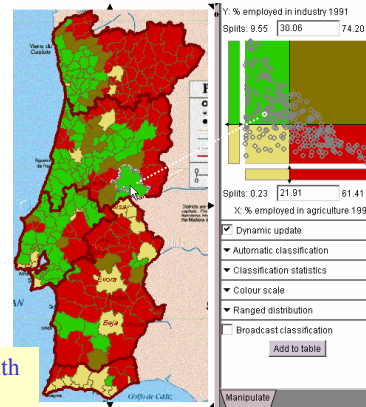
2

# Linked Displays We Already Used

Map and dot plot; each district shown on the map is also represented by a dot

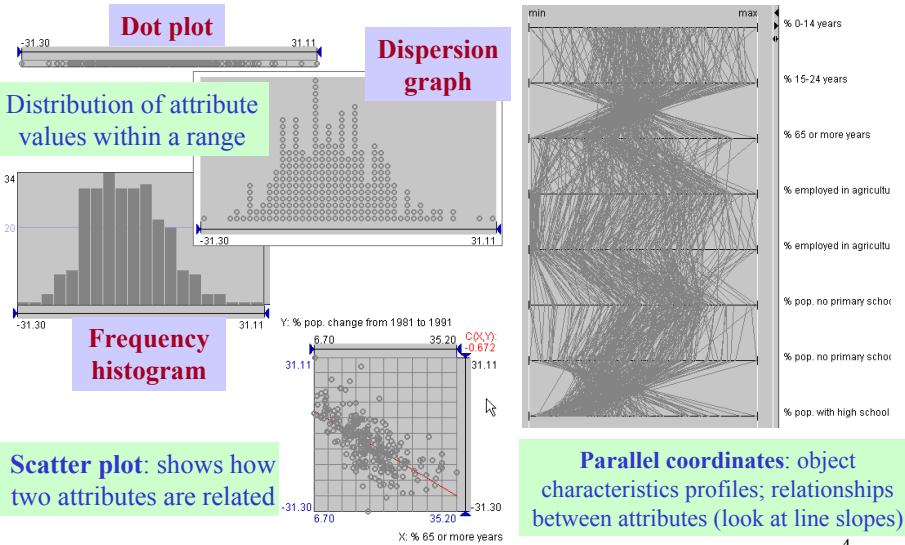


Map and scatter plot: the same technique



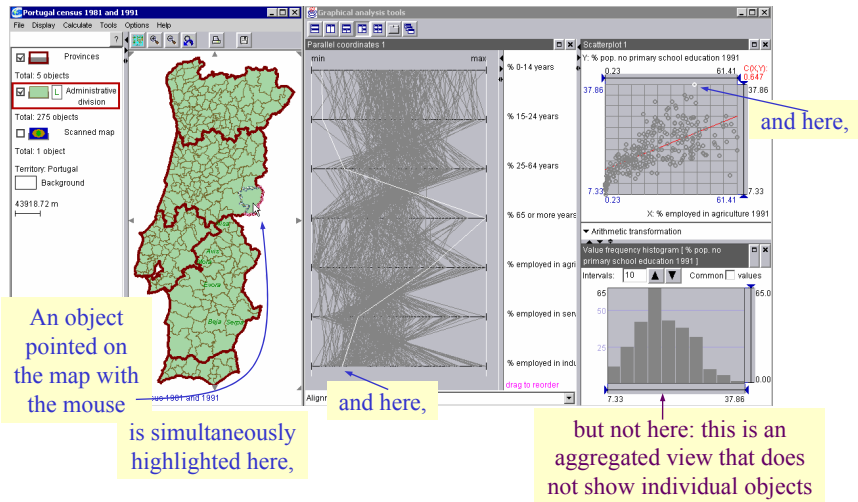
3

# Why to Use Multiple Displays?



4

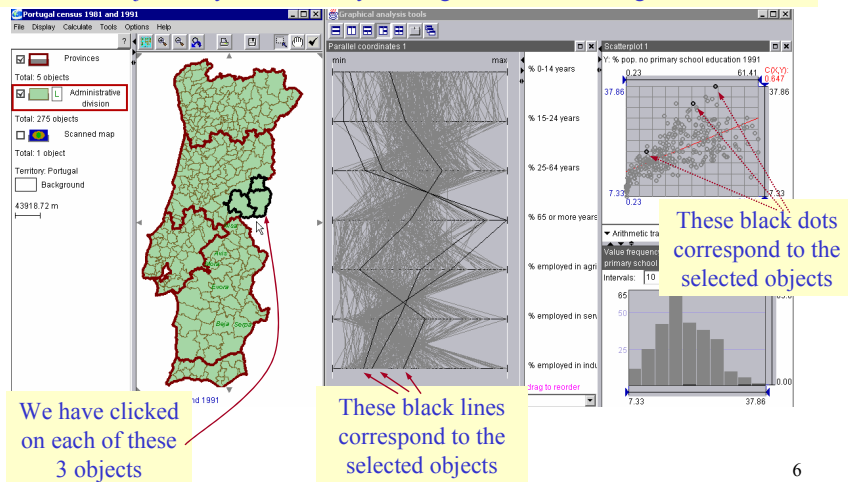
# Display Linking by Highlighting



5

# Display Linking by Selection

Selection (durable highlighting) does not disappear after the mouse is moved away. One or more objects may be selected by clicking on them or drawing a frame around.



6

# Using Display Linking (1)

Let us examine characteristics of districts in this area

This box must be "checked"

Enclose the area in a frame

The characteristics in terms of the upper 4 attributes are rather coherent

Two distinct clusters in the value space of these two attributes

The values of this attribute are split in two groups with a gap between

The values of this attribute greatly vary

The districts fit in the left half of the histogram, mostly in bars 1 and 4

7

# Using Display Linking (2)

Let us look at the districts with high % employed in industry:

Select high values by drawing a frame

The districts form 3 spatial clusters

The districts have average or high proportions of children and young

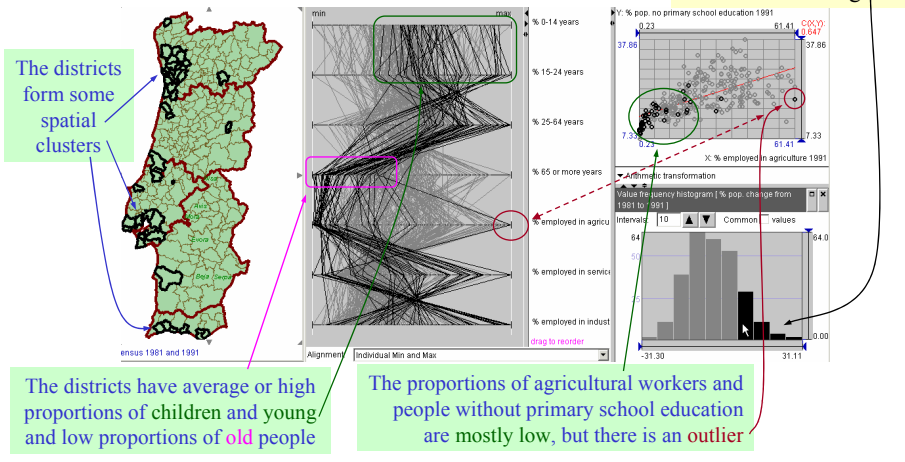
Low proportions of agricultural workers and people without primary school education

Population change: mostly between -0.1% and 12.4%

# Using Display Linking (3)

Let us look at the districts with the highest population growth:

Click on the rightmost bars in the histogram

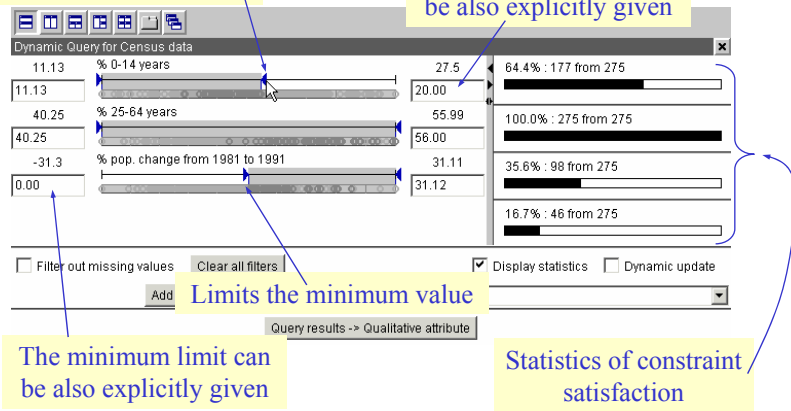


# Dynamic Query

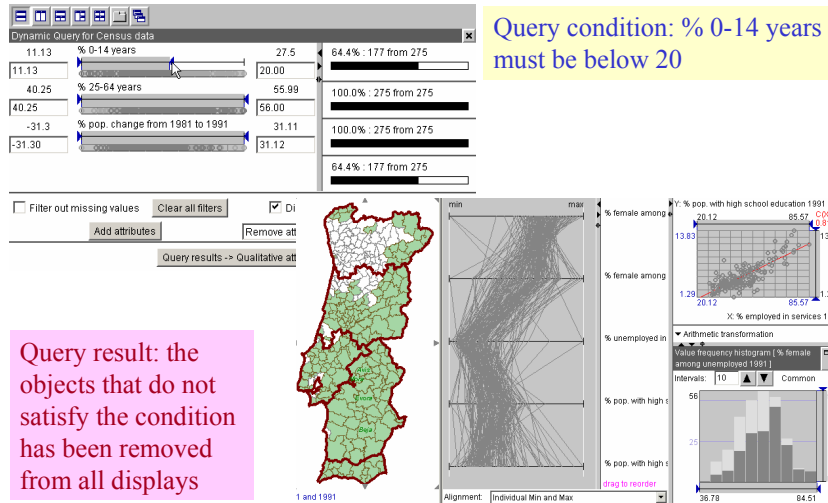
Dynamic query allows us to set constraints on attribute values

Limits the maximum value

The maximum limit can be also explicitly given

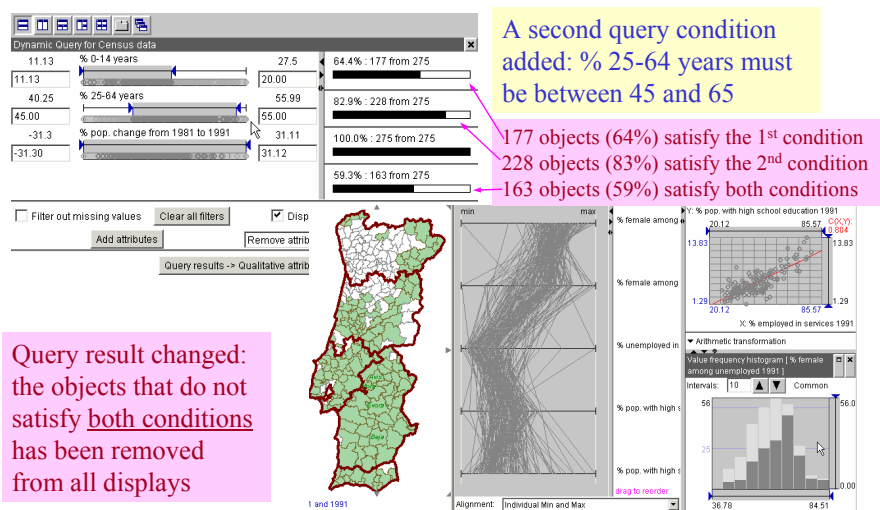


# Dynamic Query in Action (1)



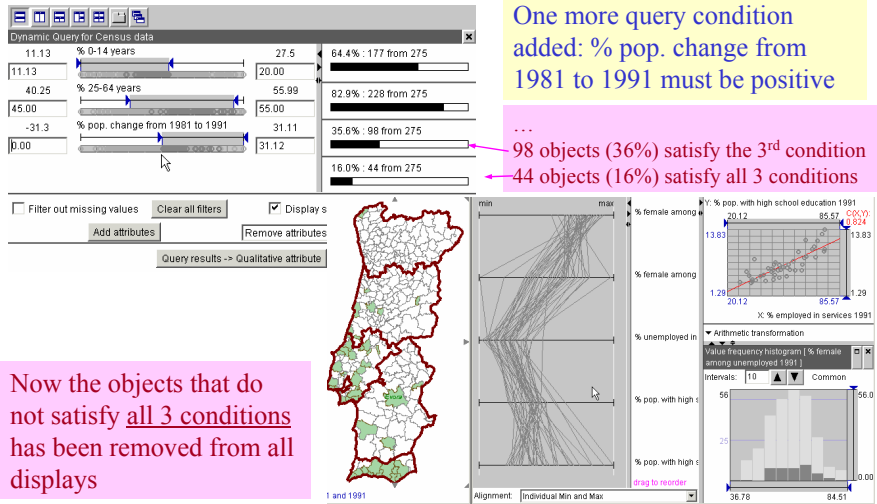
Query result: the objects that do not satisfy the condition has been removed from all displays

# Dynamic Query in Action (2)

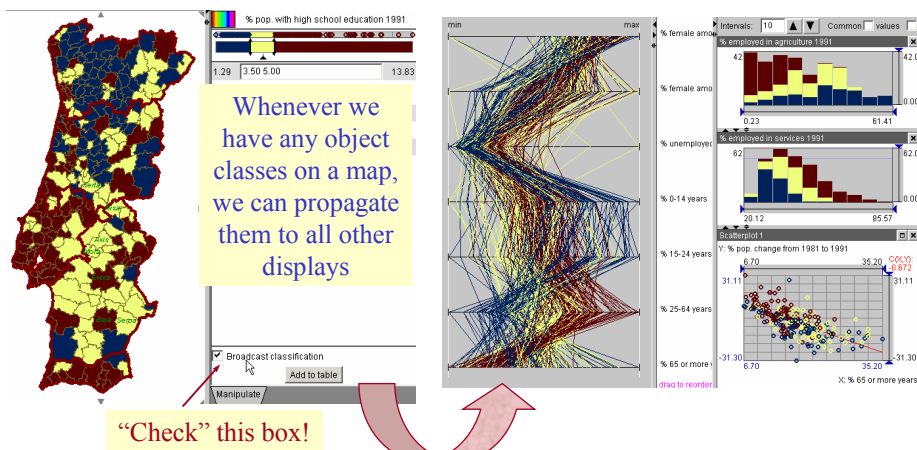


Query result changed: the objects that do not satisfy both conditions has been removed from all displays

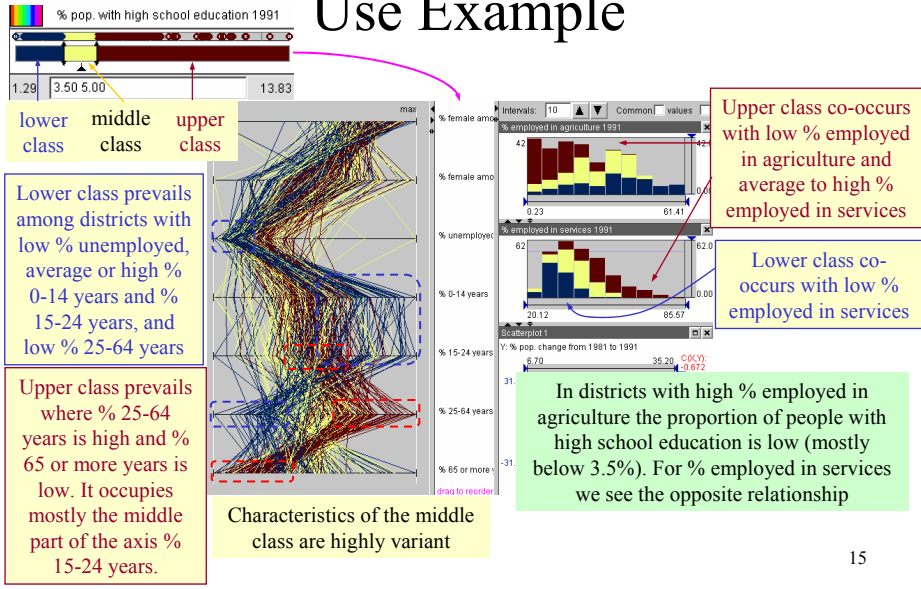
## Dynamic Query in Action (3)



## Propagation of Object Classes



# Propagation of Object Classes: Use Example



# Table View and Table Lens (1)

Click for sorting

Table cell shading shows the relative position of the values between the minimum and maximum values of the respective attributes

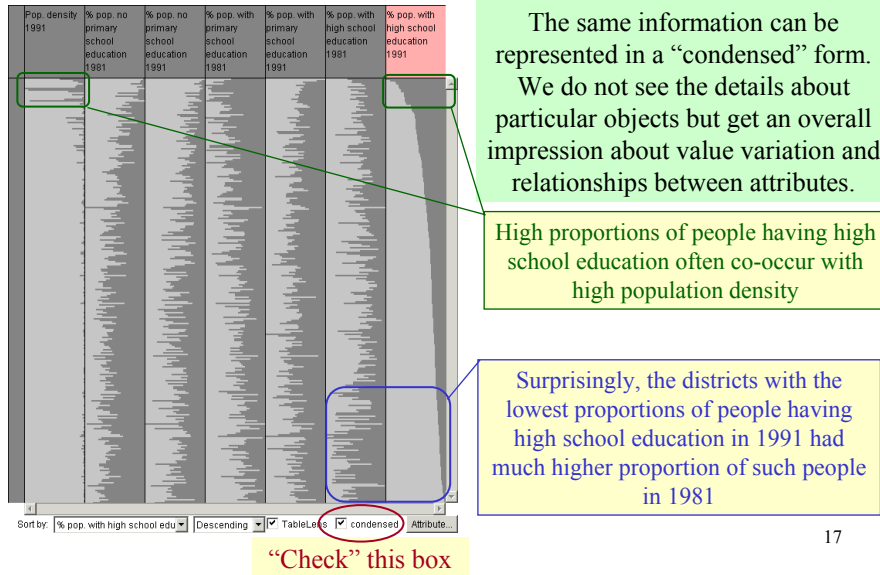
Identifiers	Pop. density 1981	Pop. density 1991	% pop. no primary school education 1981	% pop. no primary school education 1991	% pop. with primary school education 1981	% pop. with primary school education 1991	% pop. with high school education 1981	% pop. with high school education 1991
Lisboa	9636.423	7912.43	37.26	8.51	39.49	25.45	7.675	11.11
Porto	7858.089	7260.49	37.32	8.83	39.97	26.22	8.756	9.84
Amadora	6895.481	7454.641	40.41	9.68	38.66	25.18	9.941	10.80
Oeiras	3257.592	3301.527	37.89	7.33	32.80	18.69	10.444	13.83
Barreiro	2796.012	2723.485	40.64	10.01	39.09	28.16	10.089	9.85
Matosinhos	2190.979	2434.703	42.45	11.01	40.12	29.08	9.857	7.78
Almada	2110.581	2169.072	40.58	10.26	38.49	25.99	8.813	10.66
Sao Joao da Madeira	2027.62	2275.216	41.12	10.21	39.62	28.59	9.658	7.68
Espinho	1513.025	1631.933	44.68	11.37	35.82	27.50	10.919	7.56
Cascais	1457.75	1579.276	39.37	8.92	34.97	21.13	9.718	12.95
Loures	1419.716	1654.349	40.84	10.29	39.32	27.12	9.950	10.09
Vila Nova de Gaia	1324.968	1455.128	43.66	11.61	39.18	29.86	10.352	7.07
Gondomar	981.172	1074.426	43.20	12.31	39.93	29.72	10.493	8.93
Maia	975.854	1112.915	44.45	11.71	39.08	30.47	10.406	7.07
Moita	966.594	1181.863	46.27	14.50	37.64	26.61	9.123	7.71
Seixal	952.844	1249.3	41.18	10.15	38.33	24.60	10.375	11.15
Valongo	880.038	1016.194	42.80	11.90	39.94	30.70	11.482	6.40
Entroncamento	874.161	1038.394	38.51	9.10	40.18	26.73	7.952	11.22

Sort by: Pop. density 1981 Descending [TableLens] condensed Attribute...

“Check” this box

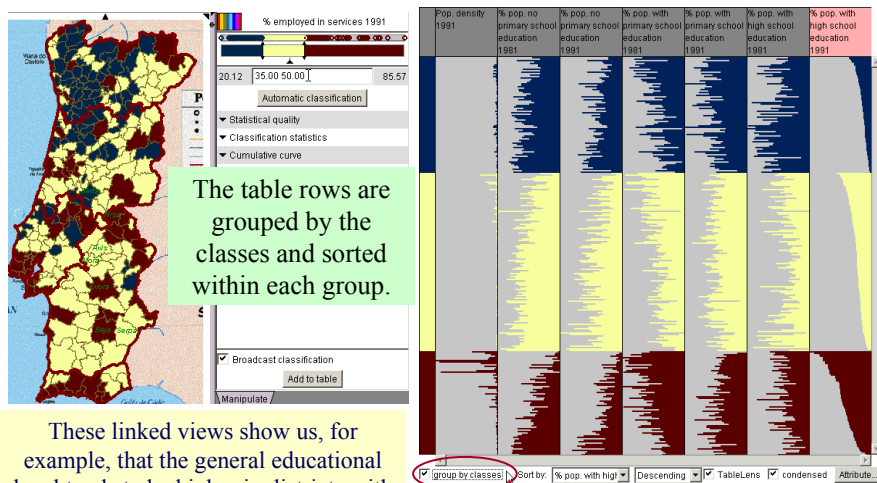


## Table View and Table Lens (2)



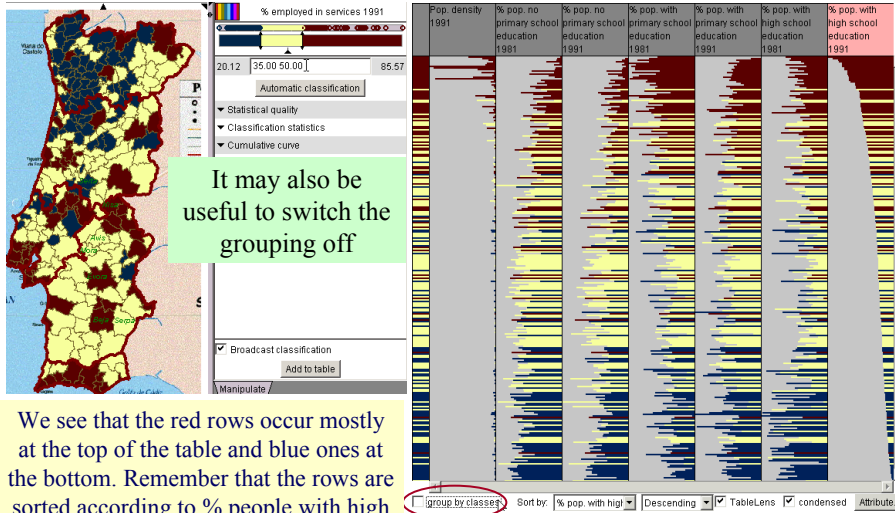
17

## Class Propagation to Table View (1)



18

## Class Propagation to Table View (2)

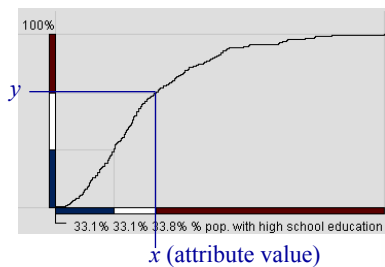


19

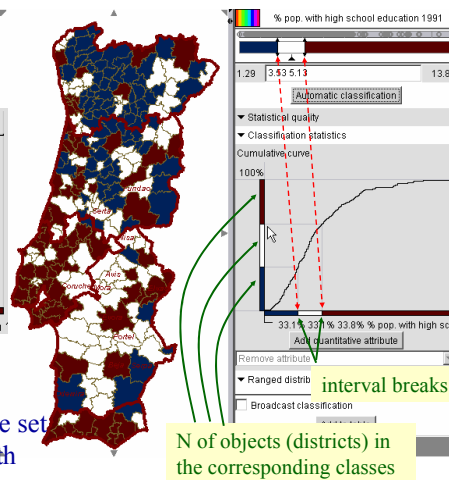
## Dynamic Classification: Additional Analytical Facilities

### Cumulative Frequency Curve

How it is built:



X-axis: attribute's value range  
Y-axis: object number or % of the whole set  
y is the number of objects (districts) with values less than or equal to x



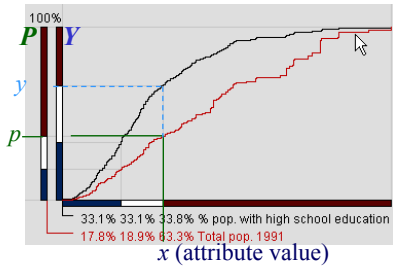
20

# Cumulative Curve: an Extension

## Cumulative Population Curve

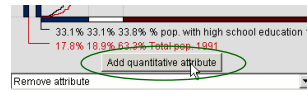
What we can learn about the distribution of the population over the classes:

How it is built:



% pop. with high school education (classes)	N of districts	Aggregate population (% of country's total)
up to 3.53	91 (33.1%)	17.8%
over 3.53 up to 5.13	91 (33.1%)	18.9%
over 5.13	93 (33.8%)	63.3%

X-axis: attribute's value range  
 Y-axis: object number or % of the whole set  
 P-axis: population number or % of the whole country's population  
 p is the **aggregate population** of districts with values less than or equal to x



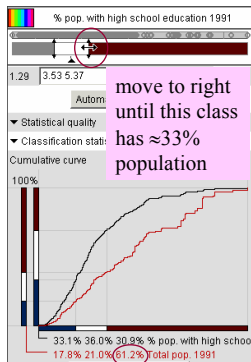
...and any other quantitative (summable) attribute can be analogously considered

21

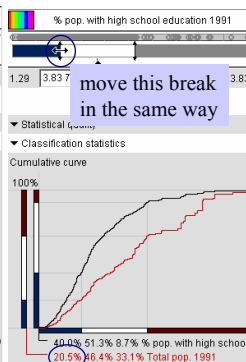
# Using Cumulative Curves (1)

Let us move the breaks so that the classes have approx. equal population

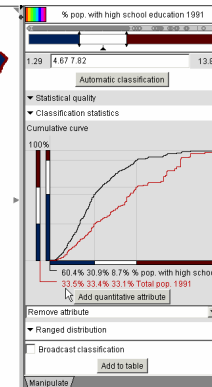
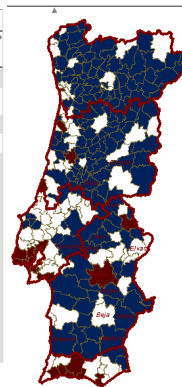
And here is the result:



Look here!

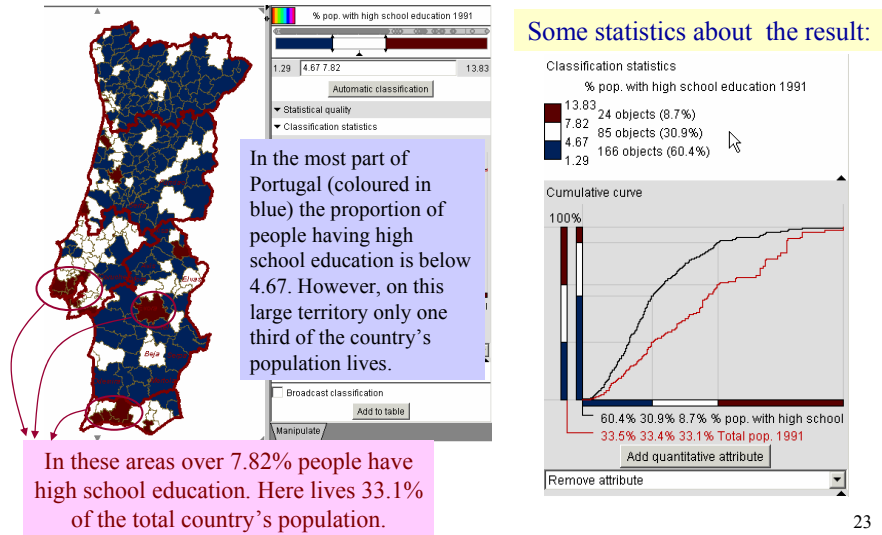


Look here!



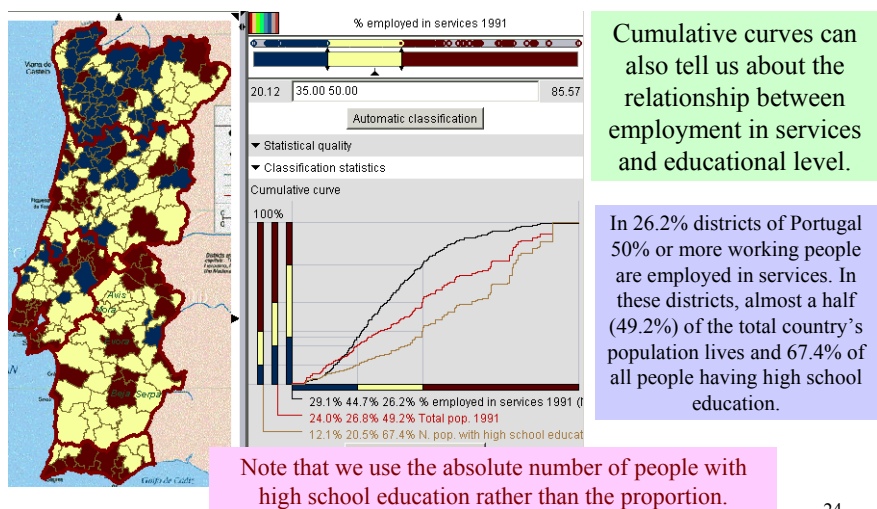
22

## Using Cumulative Curves (2)



23

## Some More Discoveries with Cumulative Curves



24

## Summary

This lection was supposed to

- explain the value of non-cartographical data displays
- stress the importance of exploring various aspects of data using multiple views
- demonstrate various techniques of display linking
- show how to use this in data analysis

25