

Visualization of Data and Statistical Models Using R

Yihui XIE

School of Statistics, RUC

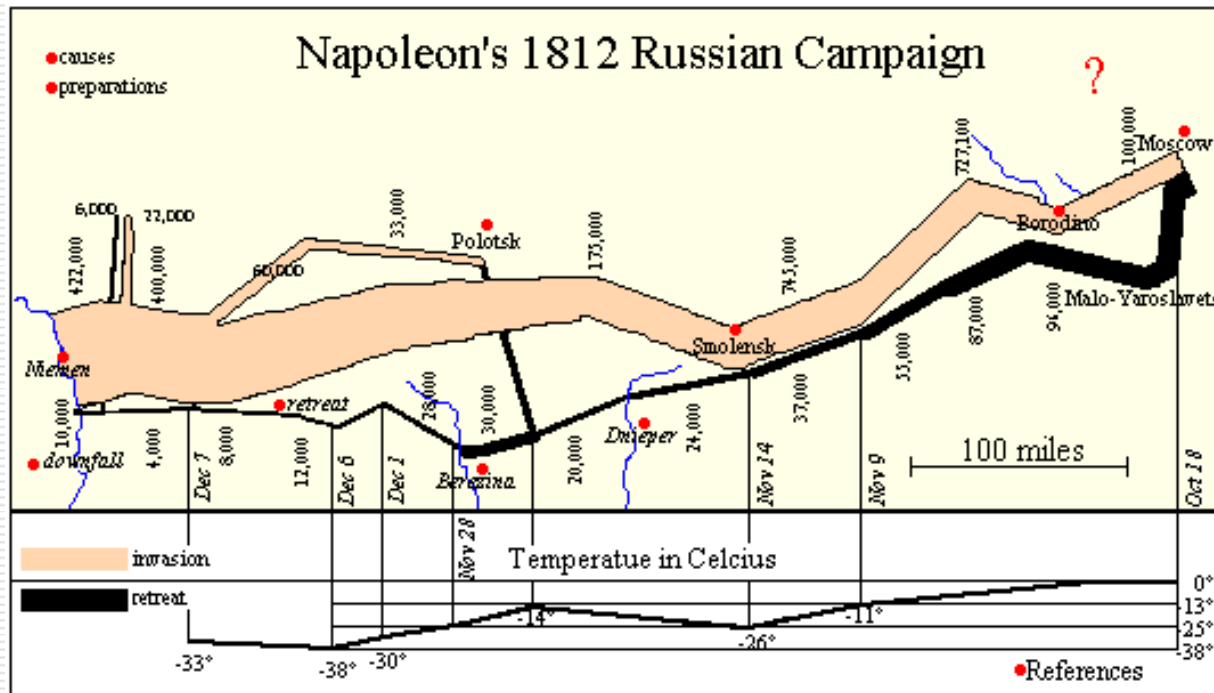
Outline

1. Get started by a map
2. Introduction to R
 - a) What's R
 - b) R graphics
3. Two examples
 - a) Regression tree
 - b) Geographical maps

1. Get started

- Why visualization?
 - Intuitional – data are complicated
- How?
 - Compute statistics – information
 - Technique: combine as many variables as possible in a single plot

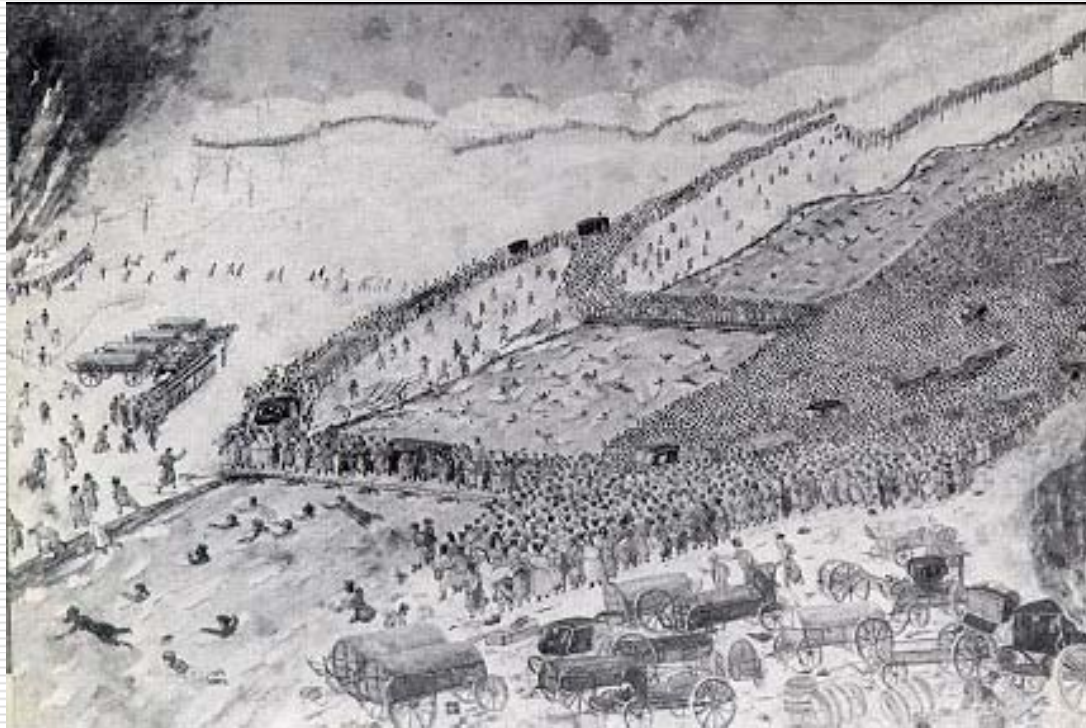
1.1 A Famous Map



C. J. Minard's Map: Napoleon's March of 1812

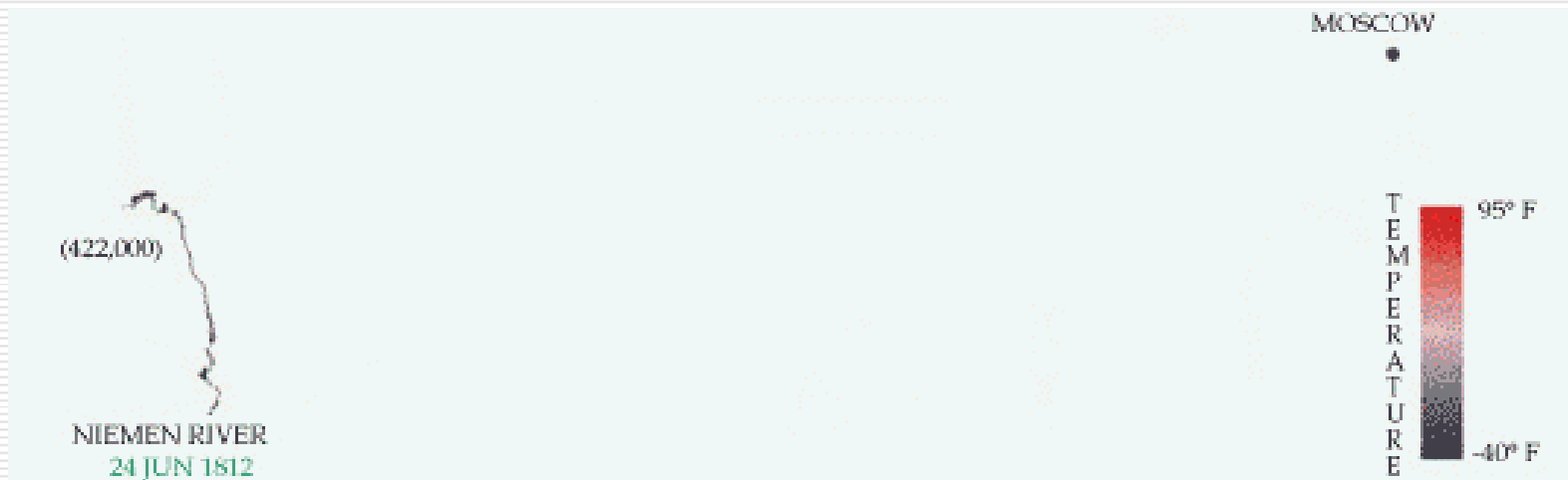
Source: <http://www.ddg.com/LIS/InfoDesignF96/Kelvin/Napoleon/map.html>

1.1 A Famous Map (cont'd)



Crossing the Bérézina, 26th November. Of 40,000 men, 25,000 were lost.

1.1 A Famous Map (cont'd)



Map of Napoleon's March of 1812 (animated)

Source: <http://www.ddg.com/LIS/InfoDesignF96/Kelvin/Napoleon/map.html>

1.2 Summary – six variables

- Size of the French army
 - Band width
- Location on a 2D surface
 - From Poland to Moscow
- Direction of invasion and retreat
 - Upper and lower band
- Temperature during the retreat
 - Cold

2. Introduction to R

- Free, open source
- For statistical computation & graphics
 - Flexible & powerful
 - Very quick implementations of advanced theories in statistics
- Many contributors (strong teamwork)
- Homepage: www.R-project.org

2. Introduction to R (cont'd)

- Demonstrations of R graphics
 - Contour plot, level plot
 - Conditioning plot
 - Scatterplot matrix
 - 3D surfaces, lines, dots, etc
 - Symbols plot
 - Maps
 - ...

3. Two examples

□ Regression tree

- Export competitiveness: how to select explanatory variables? And what information can we learn from the tree?

□ Geographical maps

- Combining agricultural export competitiveness with geographical information

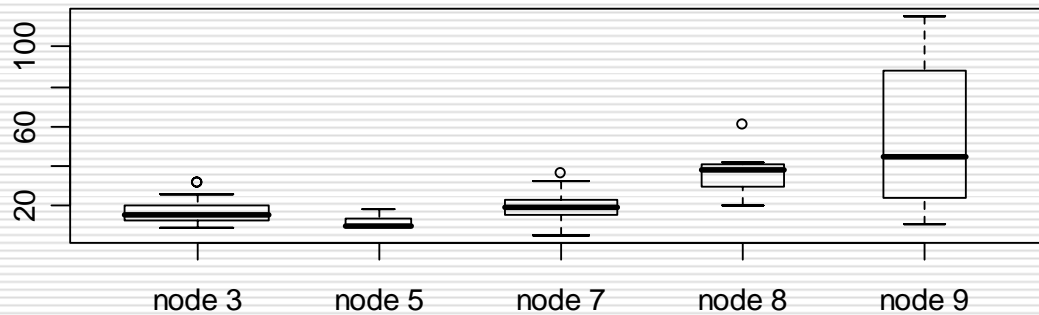
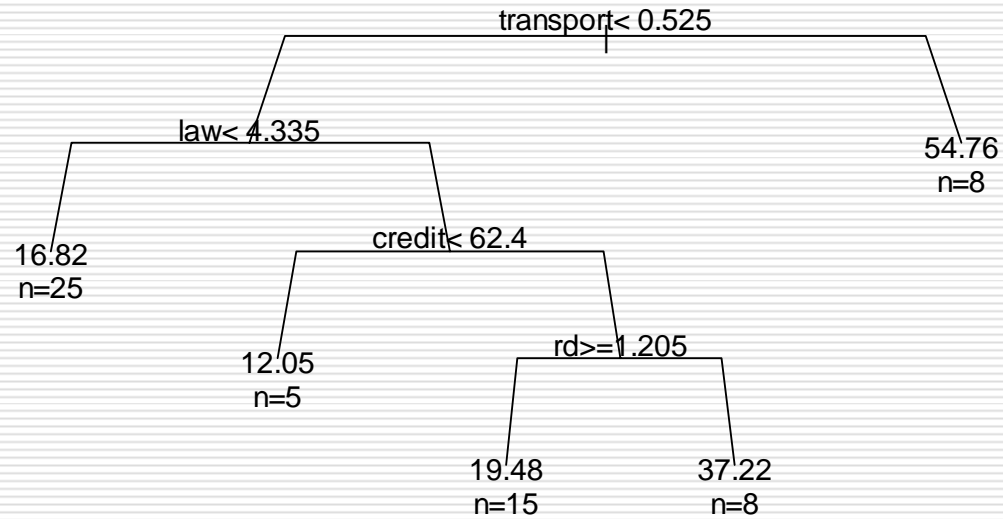
3.1 Examples: regression tree

- Classification And Regression Tree
 - CART, known as decision tree earlier
 - "Classify" the response variable by splitting explanatory variables
 - R package: `rpart`

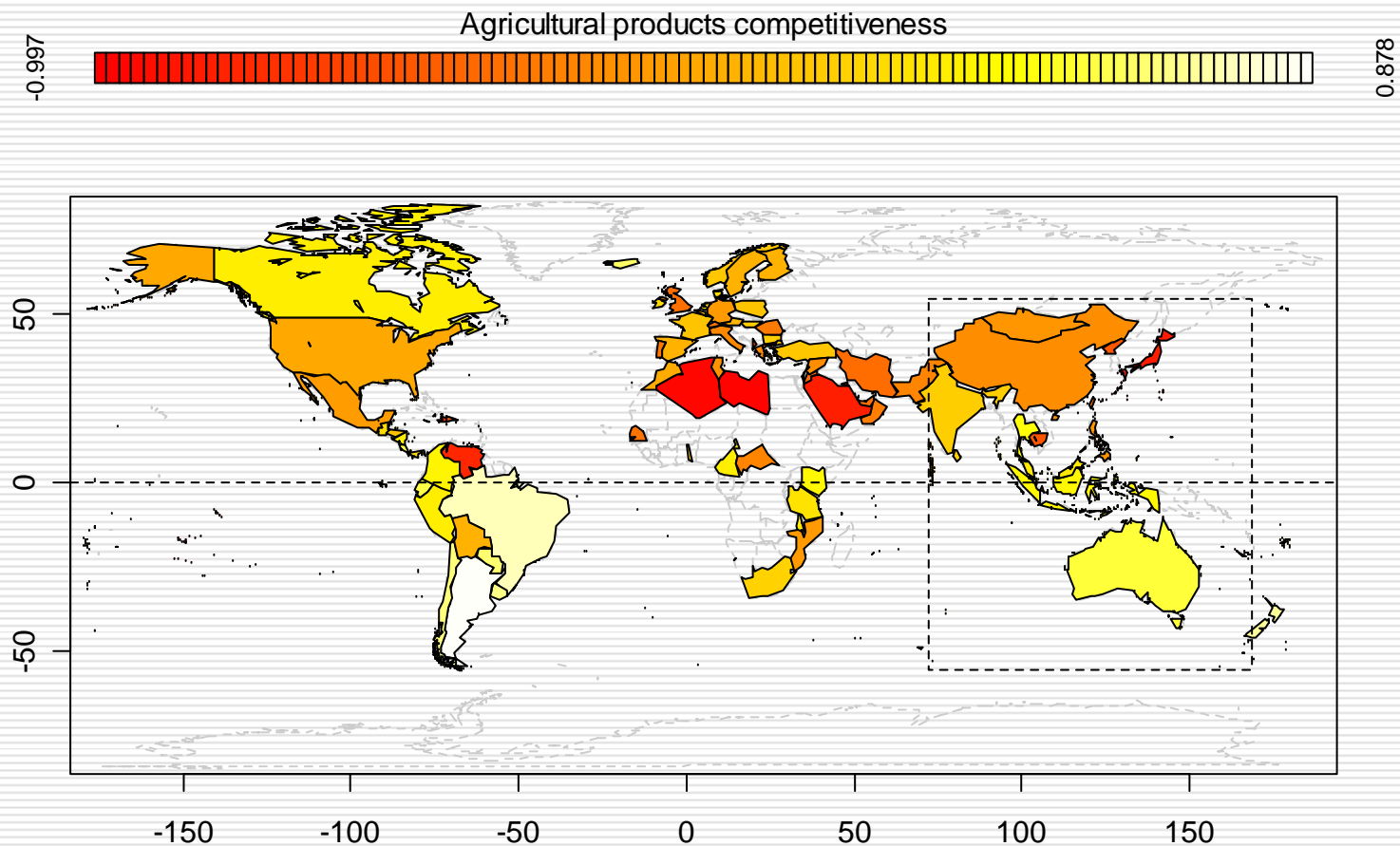
3.1 Examples: regression tree

Country / Region	export	protectionism	culture	law	politics	credit	transport	rd
ARGENTINA	12.83	4.13	6.34	2.40	4.47	26.40	0.21	0.44
AUSTRALIA	9.445	7.56	8.22	6.49	9.71	87.20	0.22	1.69
AUSTRIA	28.33	7.92	7.28	6.83	9.42	91.40	0.53	2.26
BAVARIA	18.285	6.84	6.58	4.26	9.23	91.40	0.54	2.90
BELGIUM	51.515	7.08	7.30	3.40	7.87	89.40	0.78	1.85
BRAZIL	8.55	4.95	7.59	2.90	6.30	48.20	0.06	0.93
BULGARIA	30.965	4.92	7.50	4.00	5.50	55.00	0.21	0.50
...

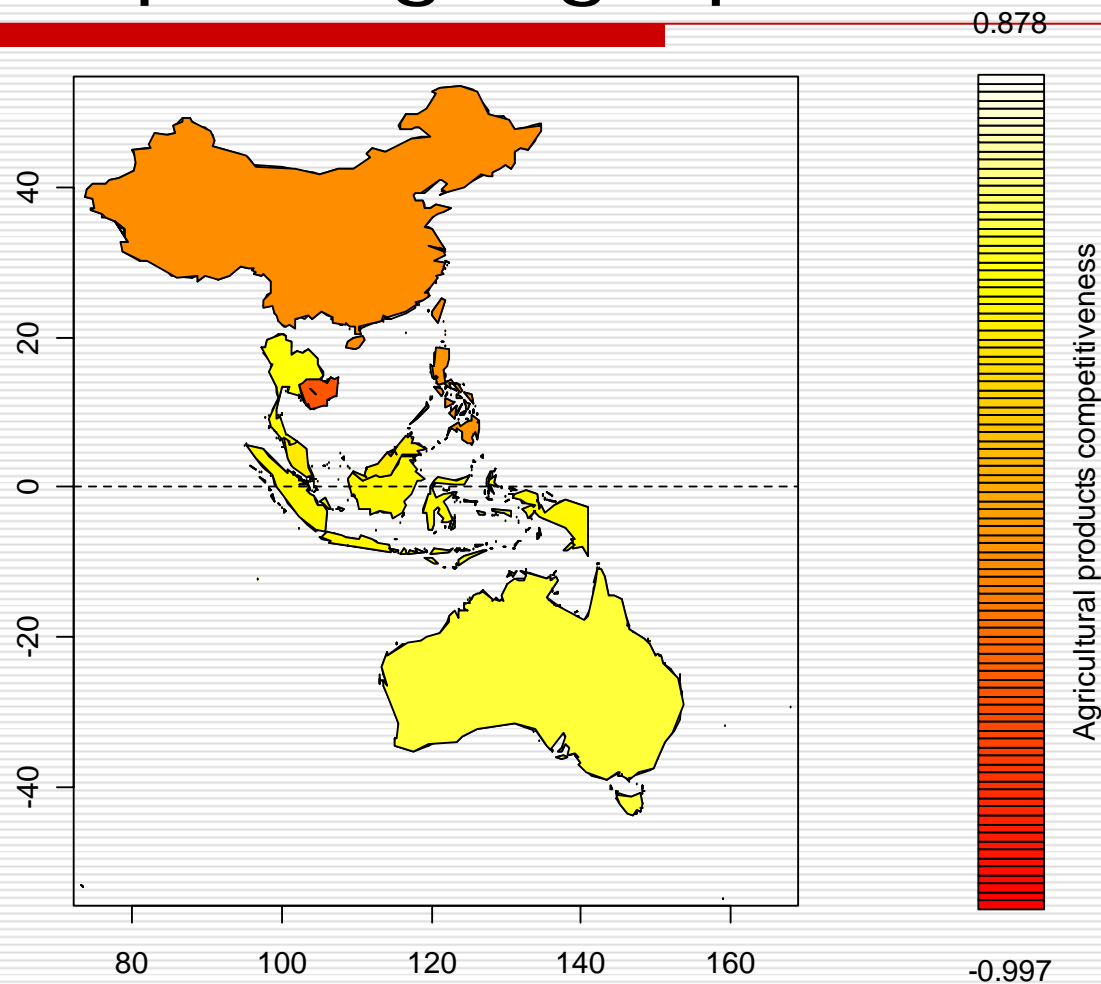
3.1 Examples: regression tree



3.2 Examples: geographical maps



3.2 Examples: geographical maps



Thanks!

- A "THANKS" drawn by hand 😊