



Exploration,
selling pitch
and deliverable;
the promises of visualization

Séminaire
«Methods and Research»
26 mai 2015

andreas.perret@unil.ch

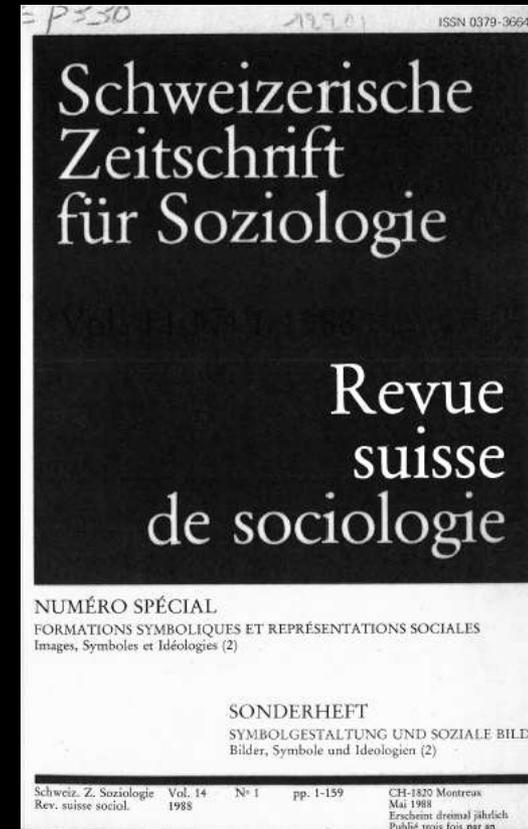
Büroangestellte(r)	-0.076 **	-0.095 ***
Handelsdiplomandin	-0.057 *	-0.054 **
Detaillhandelsangestellte(r)	-0.068 ***	-0.036 **
Pharma-AssistentIn	-0.059 ***	-0.081 ***
VerkäuferIn	-0.136 ***	-0.116 ***
Koch/Köchin	-0.049 *	0.010
Coiffeur(-euse)	-0.158 ***	-0.099 **
Andere (Lehrberufe)	0.066 ***	0.040
Andere (Tertiärabschlüsse)	-0.039 *	-0.029
Alter	-0.013 **	-0.015 **
Alter (quadratic)	0.000 **	0.000 ***
Migrationshintergrund	0.048 **	0.033 **
Partnerschaft	0.017	0.018
ISEI	0.000	0.000
Französischsprachige CH	-0.017	-0.014
Italienischsprachige CH	-0.129	-0.123 ***
Stadt	0.015	0.014
Wert intrinsisch	-0.010	-0.011
Wert extrinsisch	0.012	0.012
Wert Partnerschaft/Familie	0.000	0.000
Erwerbstätigkeit nach ISCO (Referenz: Administration)		
Direktion, Geschäftsleitung usw.		0.067 *
InformatikerIn		0.088
IngenieurIn		0.015

Tabelle 2 Loglineare Regressionsschätzungen und Lohndekomposition

	Alle		Frauen		Männer	
	Ausbildungen kontrolliert	+ Ausbildungen Erwerbstätigkeit kontrolliert	Ausbildungen kontrolliert	+ Ausbildungen Erwerbstätigkeit kontrolliert	Ausbildungen kontrolliert	+ Ausbildungen Erwerbstätigkeit kontrolliert
	(1)	(2)	(3)	(4)	(5)	(6)
Frau	-0.057 ***	-0.047 ***				
Erfahrung	0.002	0.000	0.005	0.000	0.000	-0.001
Erfahrung (quadratic)	0.000	0.000	0.000	0.000	0.000	0.000
Schultyp	0.029 **	0.031 **	0.019	0.009	0.027	0.037 **
PISA-Lesekompetenzen	0.000	0.000	0.000	0.000	0.000	0.000
Note	0.001	0.011	-0.013	-0.025	0.004	0.021
Höchster erworbener Ausbildungsabschluss (Referenz: beruf iche Grundbildung)						
Matur/Lehre mit Berufsmatur	0.008	0.006	0.008	-0.003	0.010	0.017
Höhere Fachschule, Lehrerseminar	0.314 ***	0.260 ***	0.340 ***	0.273 ***	0.297 ***	0.273 ***
Universität, Fachhochschule	0.366 ***	0.283 ***	0.395 ***	0.277 ***	0.357 ***	0.311 ***
20 am häufigsten erlernte Lehrberufe und 5 häufigste Tertiärabschlüsse (Referenz: Kaufmännische(r) Angestellte(r))						
SchreinerIn	0.049	0.107 *	-0.148 ***	0.029	0.127 **	0.152 **
AutomatikerIn	0.052	0.045	0.000	0.000	0.142 ***	0.094 **
PolymechanikerIn	0.053	0.055	-0.019	-0.012	0.130 **	0.097 *
Büroangestellte(r)	-0.076 **	-0.095 ***	-0.088 **	-0.099 ***	-0.116 **	-0.171 ***
Handelsdiplomandin	-0.057 **	-0.054 **	-0.093 ***	-0.078 **	0.088 **	0.084 *
Detaillhandelsangestellte(r)	-0.068 ***	-0.036	-0.078 ***	-0.007	-0.025	-0.018
Pharma-AssistentIn	-0.059 ***	-0.081 ***	-0.070 ***	-0.071 ***	0.000	0.000
VerkäuferIn	-0.136 ***	-0.116 ***	-0.163 ***	-0.112 **	-0.102	-0.082
Koch/Köchin	-0.049 *	0.010	-0.067	0.037	0.020	0.033
Coiffeur(-euse)	-0.158 ***	-0.099 **	-0.190 ***	-0.095 **	-0.136 ***	0.114
Andere (Lehrberufe)	0.066 ***	0.040	0.047	0.049	0.000	0.000
Andere (Tertiärabschlüsse)	-0.039 *	-0.029	-0.073 ***	-0.024	0.035	0.015
Alter	-0.013 **	-0.015 ***	-0.016 *	-0.014	0.000	-0.006
Alter (quadratic)	0.000 **	0.000 ***	0.000 *	0.000	0.000	0.000
Migrationshintergrund	0.048 **	0.033 **	0.047 **	0.030	0.047	0.036
Partnerschaft	0.017	0.018	0.010	0.005	0.031	0.041
ISEI	0.000	0.000	0.001	0.001 *	-0.001	-0.001
Französischsprachige CH	-0.017	-0.014	-0.034 ***	-0.027 **	-0.003	0.005
Italienischsprachige CH	-0.129	-0.123 ***	-0.144 ***	-0.139 ***	-0.122 ***	-0.117 ***
Stadt	0.015	0.014	0.019	0.025 **	0.006	0.001
Wert intrinsisch	-0.010	-0.011	0.011	-0.001	-0.024 *	-0.025 *
Wert extrinsisch	0.012	0.012	0.009	0.006	0.004	0.010
Wert Partnerschaft/Familie	0.000	0.000	-0.005	-0.007	0.008	0.009
Erwerbstätigkeit nach ISCO (Referenz: Administration)						
Direktion, Geschäftsleitung usw.			0.067 *	0.061		0.041
InformatikerIn			0.088	-0.044		0.103
IngenieurIn			0.015	0.043		-0.013

Fortsetzung der Tabelle 2 auf der nächsten Seite.

« On pourrait donc faire l'hypothèse que ce n'est pas sa "nature" médiate, étrangère aux faits tels que la sociologie les appréhende, qui de l'image fait problèmes en sciences sociales, mais au contraire, sa réalité trop proluxe, trop proximale, facteur inhibiteur principal, au point que nous ne savons plus qu'en faire » (Fisher et Zoller, 1987 : 313)



Revue Suisse de Sociologie, 1987 (cité par Michaël Meyer)

Phénomène
irréductible par
les graphes

**Visualisation
et sociologues**

Data Visualization in Sociology

Kieran Healy and James Moody

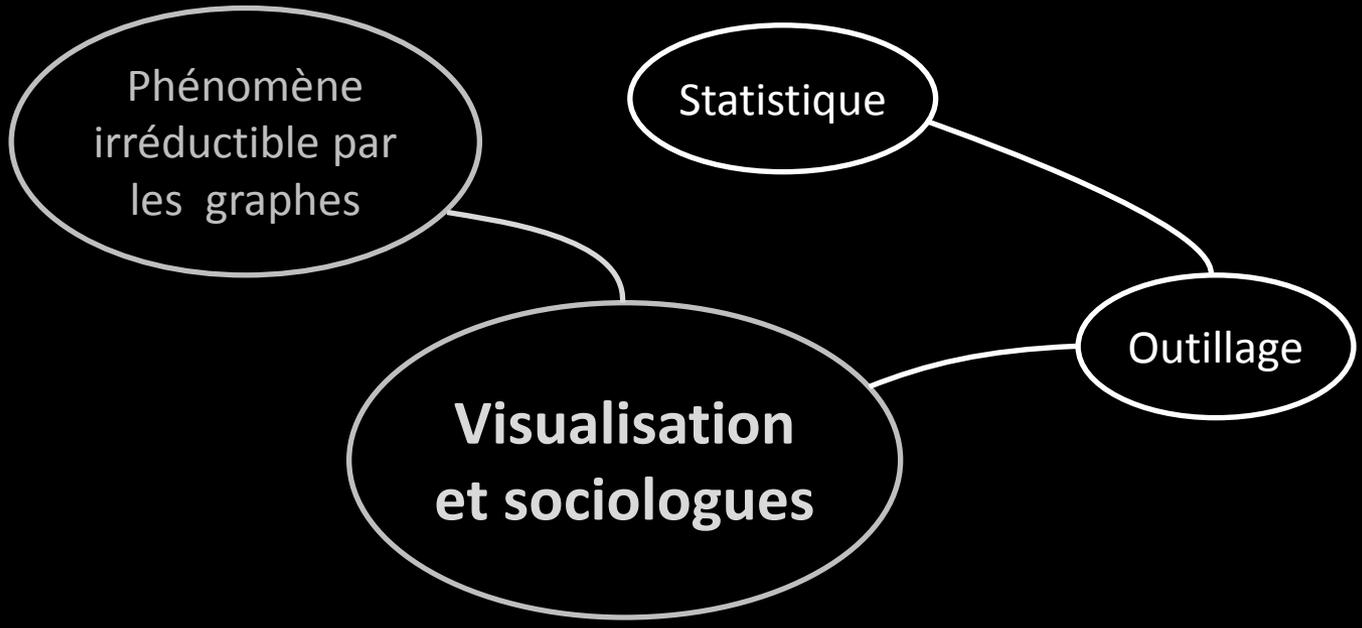
Sociology Department, Duke University, Durham, North Carolina 27708;
email: kjhealy@soc.duke.edu, jmoody77@soc.duke.edu

Keywords

visualization, statistics, methods, exploratory data analysis

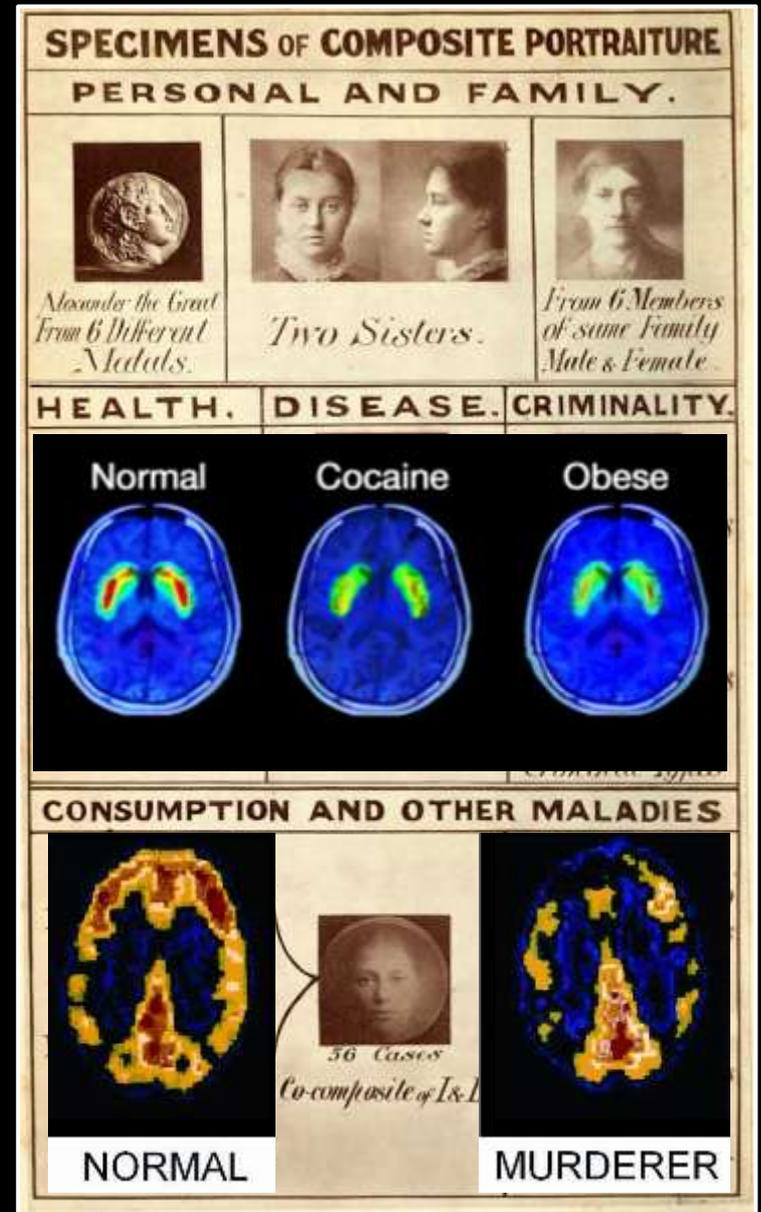
Abstract

Visualizing data is central to social scientific work. Despite a promising early beginning, sociology has lagged in the use of visual tools. We review the history and current state of visualization in sociology. Using examples throughout, we discuss recent developments in ways of seeing



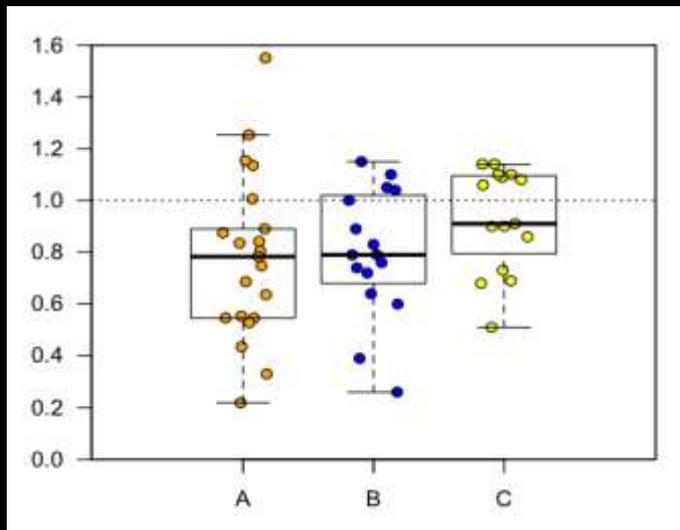
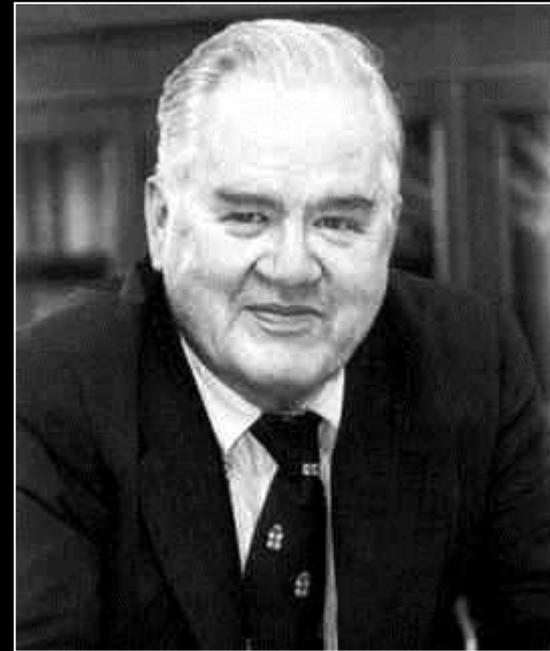
Francis Galton
 Karl Pearson
 Ronald Fisher

(Wikimedia commons)



Journal of the American
 Juge Association, 2014

*Anything to which a confirmatory procedure was not explicitly attached was decried as mere descriptive statistics, no matter how much we had learned from it.
(in «Exploratory Data Analysis» 1977)*



John Tukey (1915-2000)

Exploration ←————→ Confirmation

Détective

Juge

Recherche de structures et de motifs

Confirmer une hypothèse

Peu de certitudes

Phénomène bien connu

Peu de présupposés

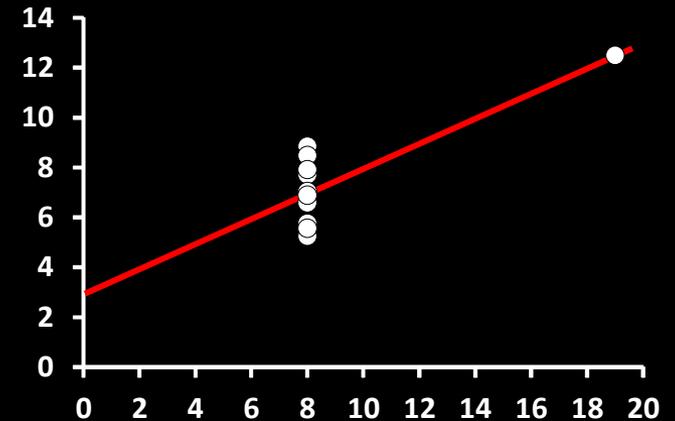
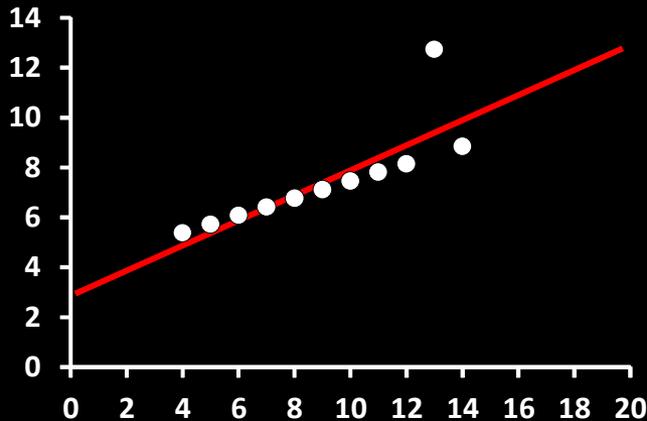
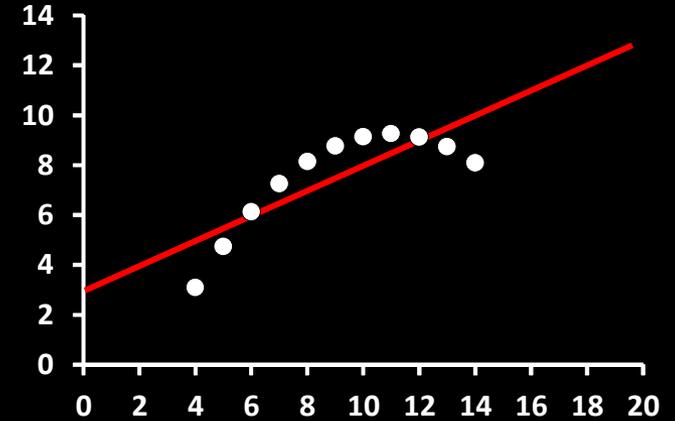
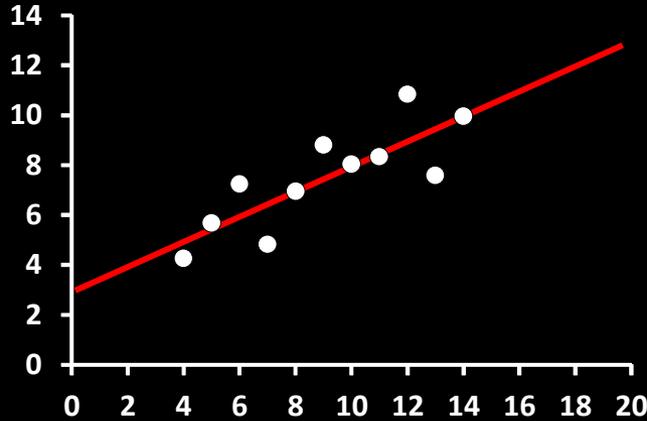
Hypothèses fortes

Induction

Déduction

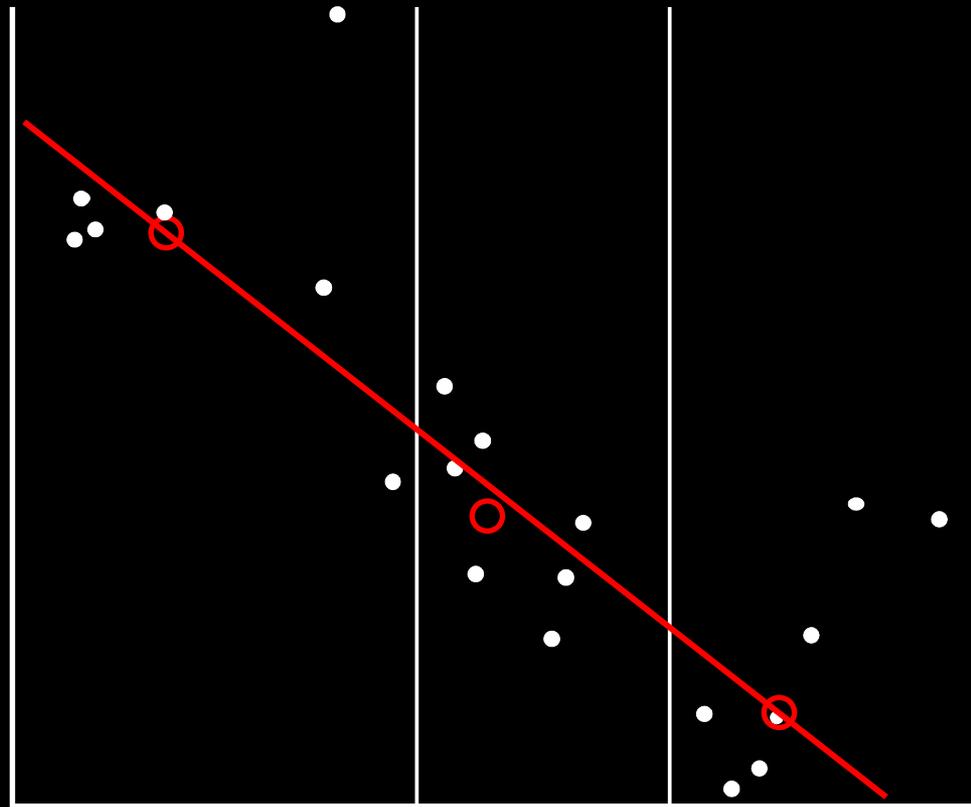
Quartet d'Anscombe

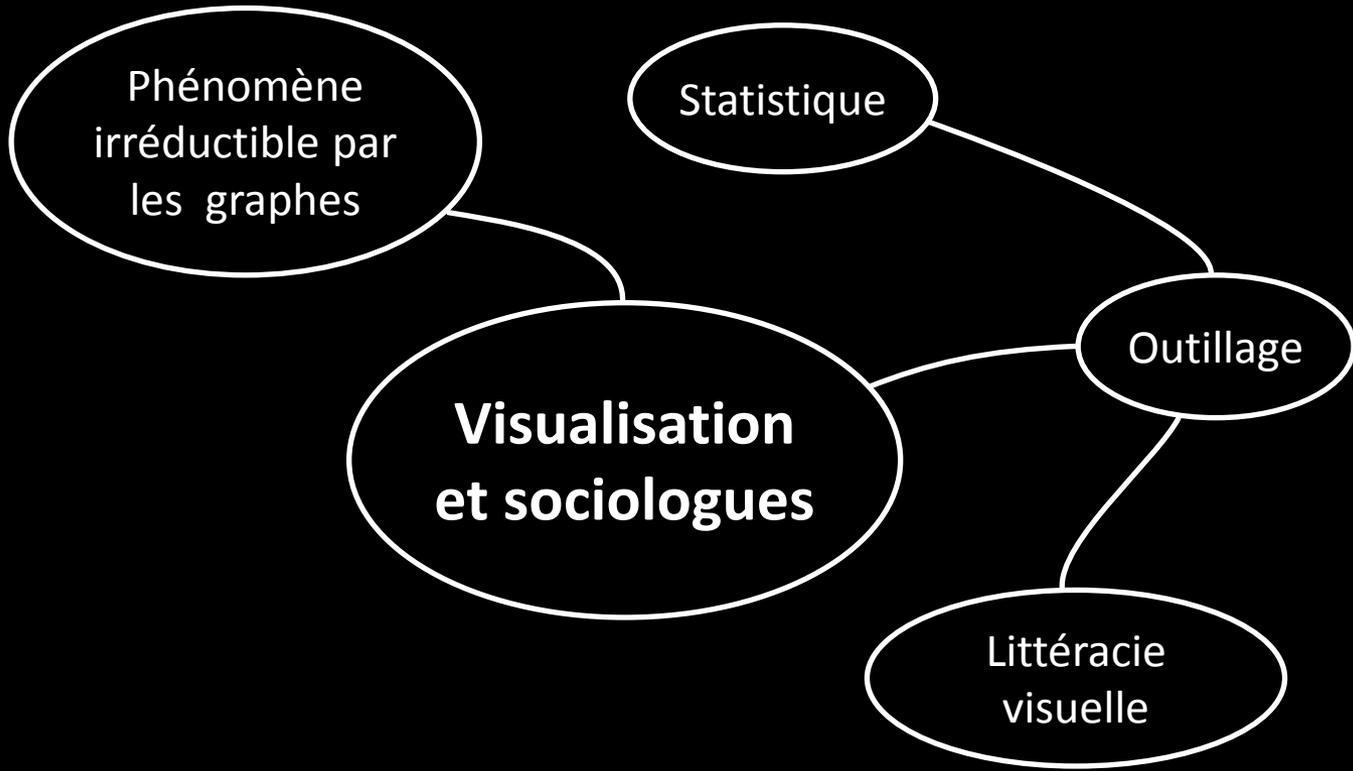
N=11
Moyenne X = 9
Variance X = 11
Moyenne Y = 7.5
Variance Y = 4.12
Correlation = 8.16
Reg: $Y = 0.5x+3$

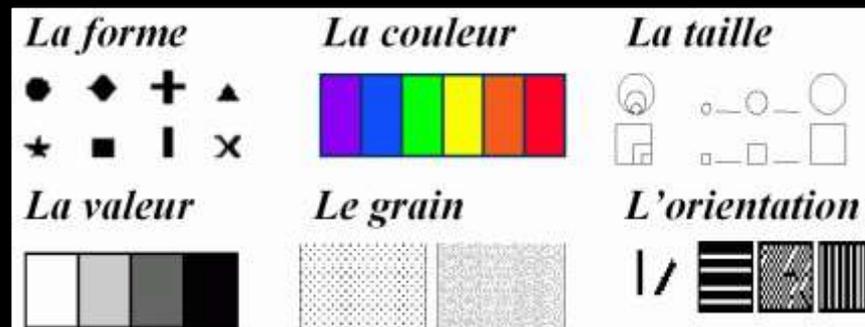
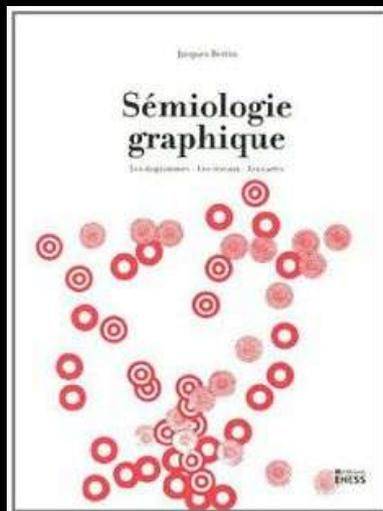
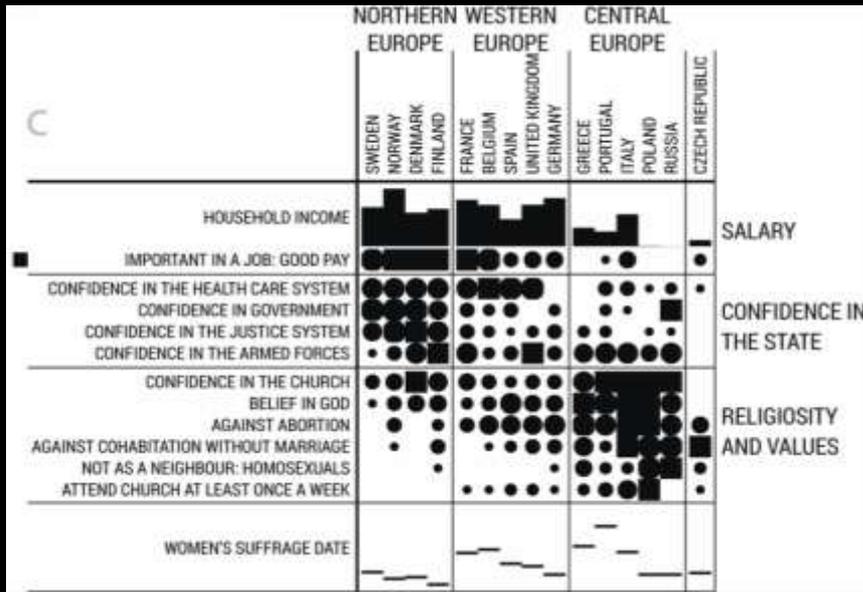


(Frank Anscombe)

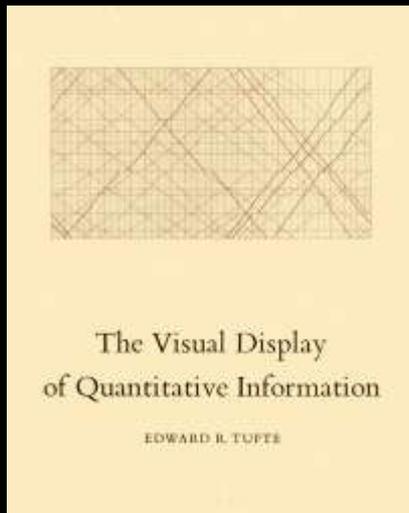
Droite robuste de Tukey



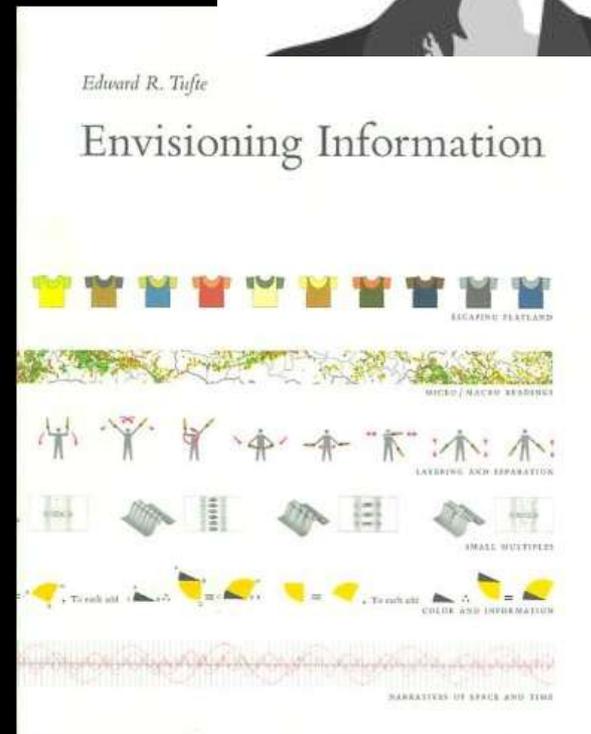




Jacques Bertin (1918-2010)



		Close	Max	Min
AT&T		40,28	41,34	33,30
Boeing		98,15	100,59	84,79
Citigroup		53,98	55,20	48,27
Exxon Mobil		85,94	85,94	69,56
General Electric		38,12	38,12	34,09
General Motors		34,66	36,20	28,85
Intel		24,24	24,24	18,76
Microsoft		30,49	31,11	26,63



Edward Tufte (1942-)

Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813.

Dessiné par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les longueurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Legur, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk et Mohilew et qui rejoignent Orscha et Witebsk, avaient toujours marché avec l'armée.

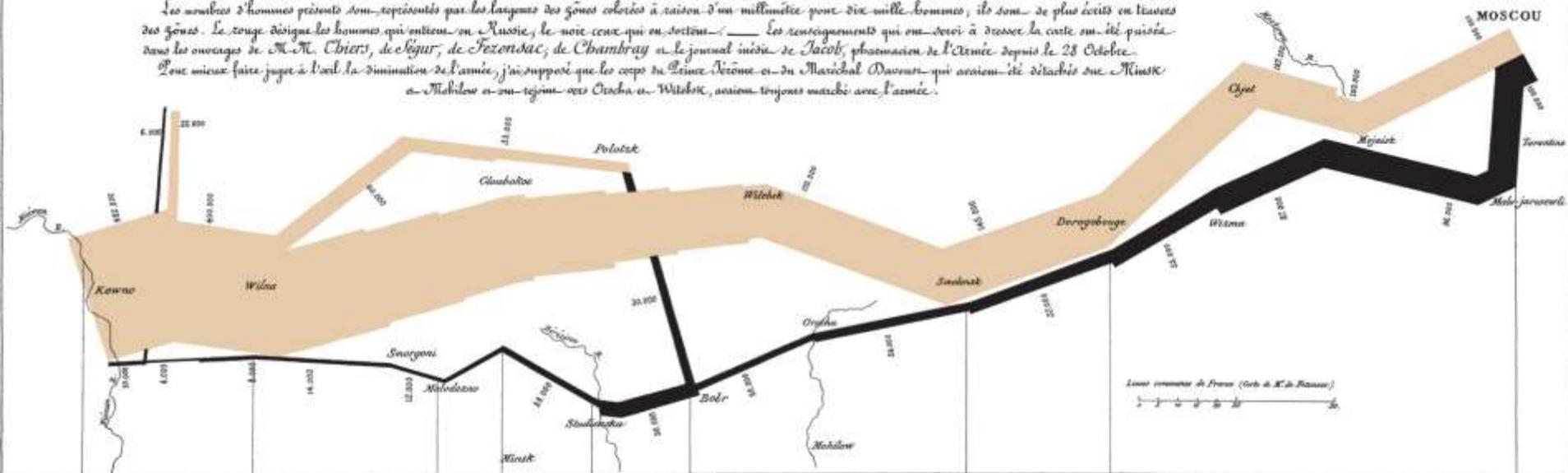
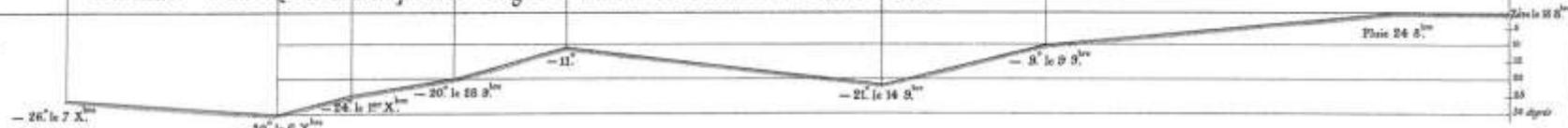


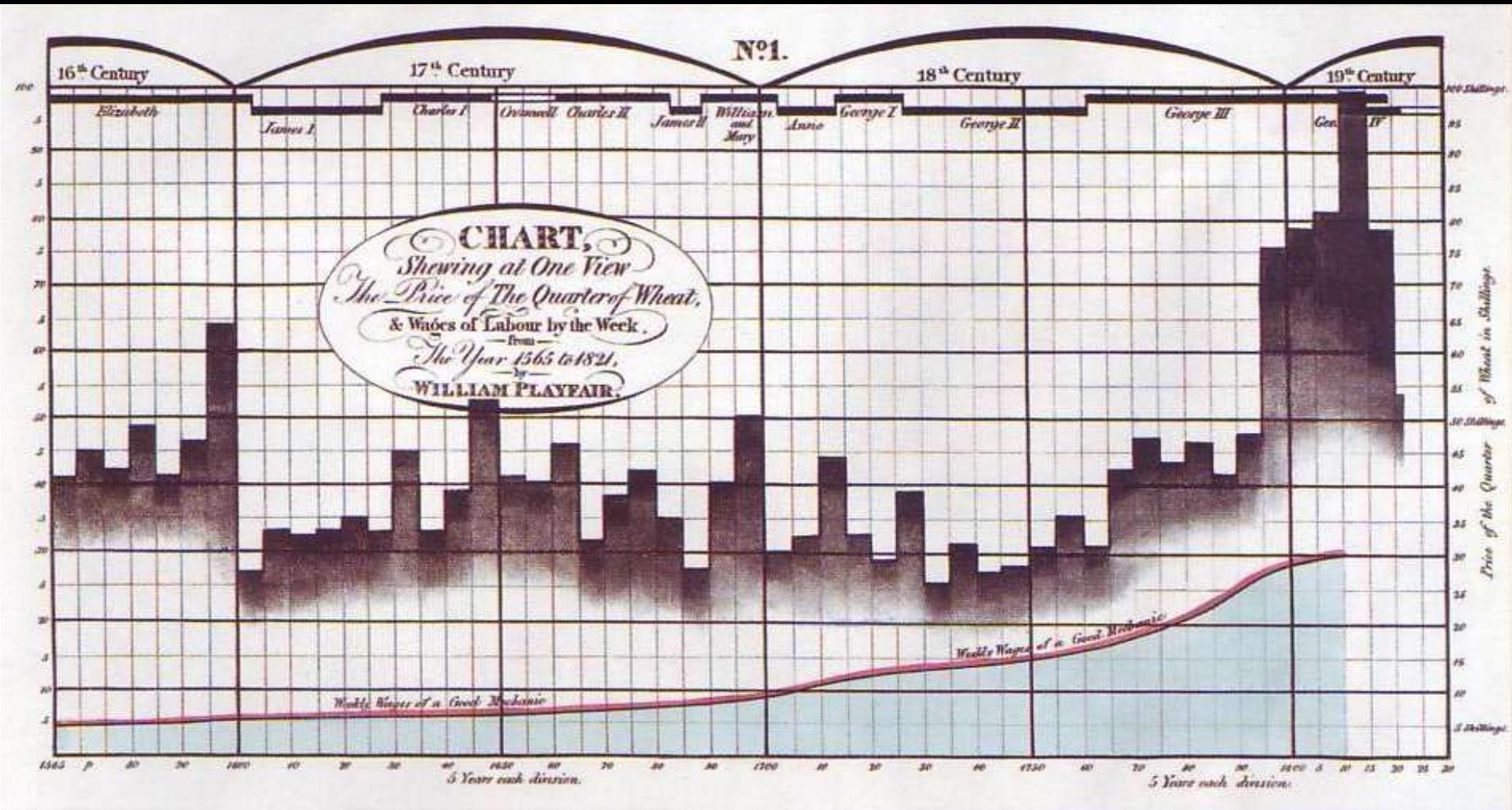
TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.



Dessiné par Regnier, à Paris. 3^e Rue Marie 23 24^e à Paris.

Imp. Lab. Regnier et Bourcier.

Charles Joseph Minard (1781-1870),



William Playfair (1795-1823),

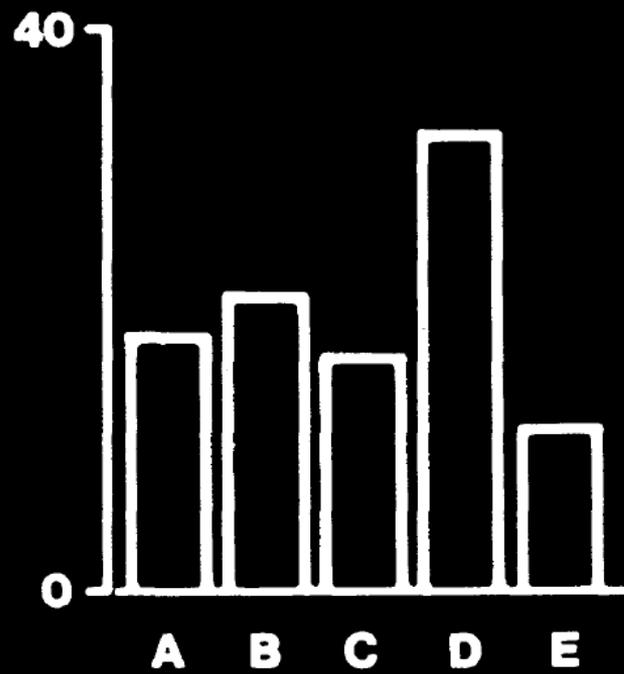
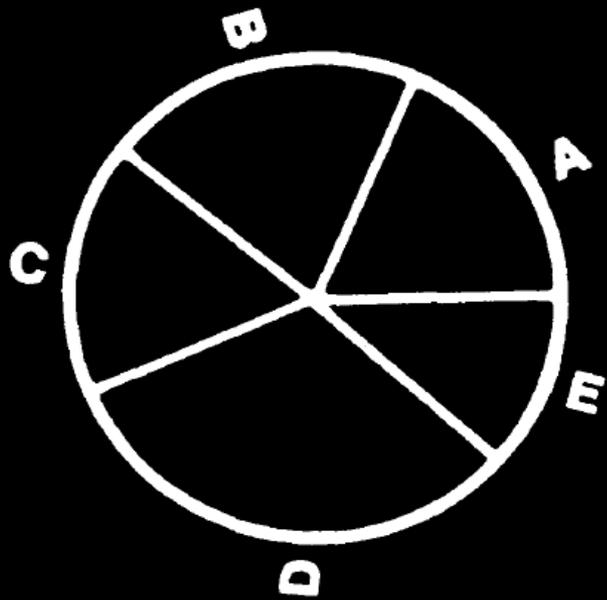
Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods

WILLIAM S. CLEVELAND and ROBERT McGILL*

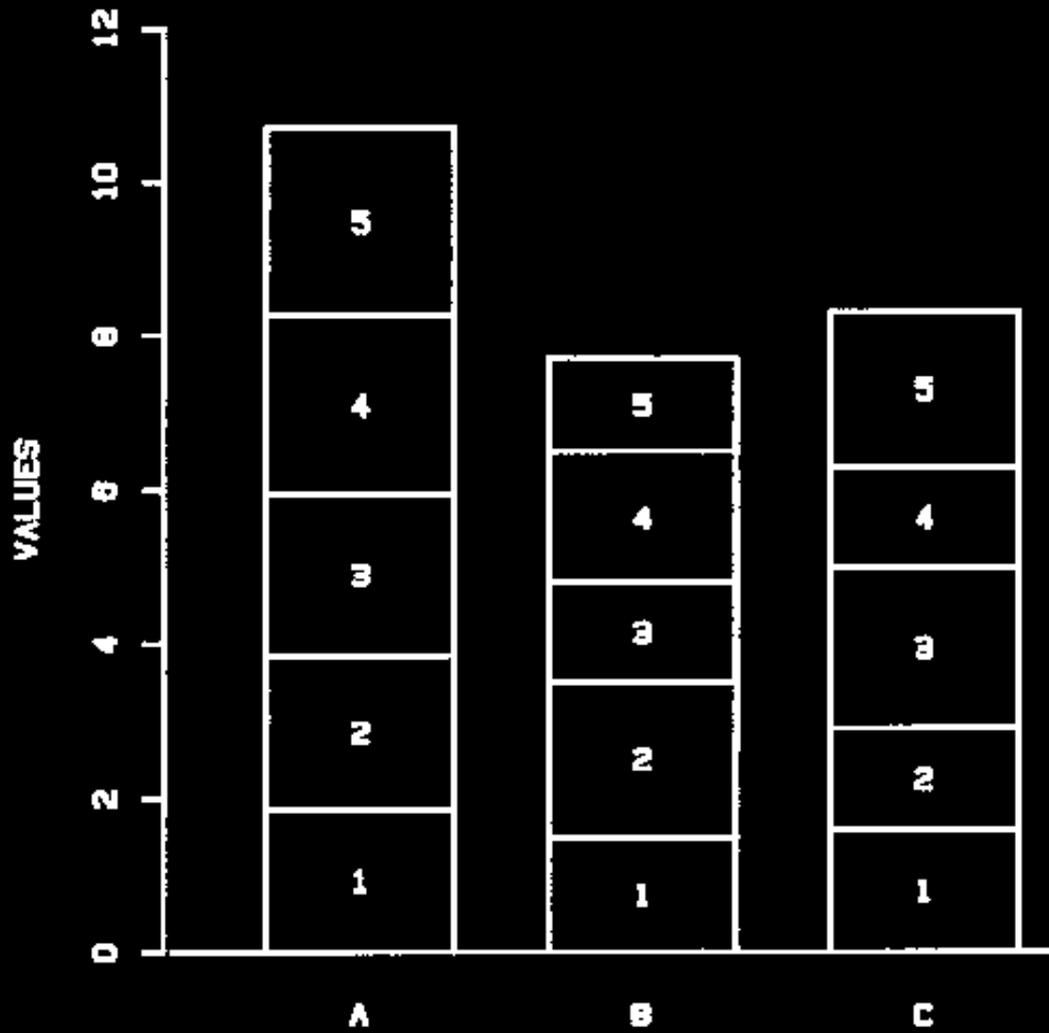
1. INTRODUCTION

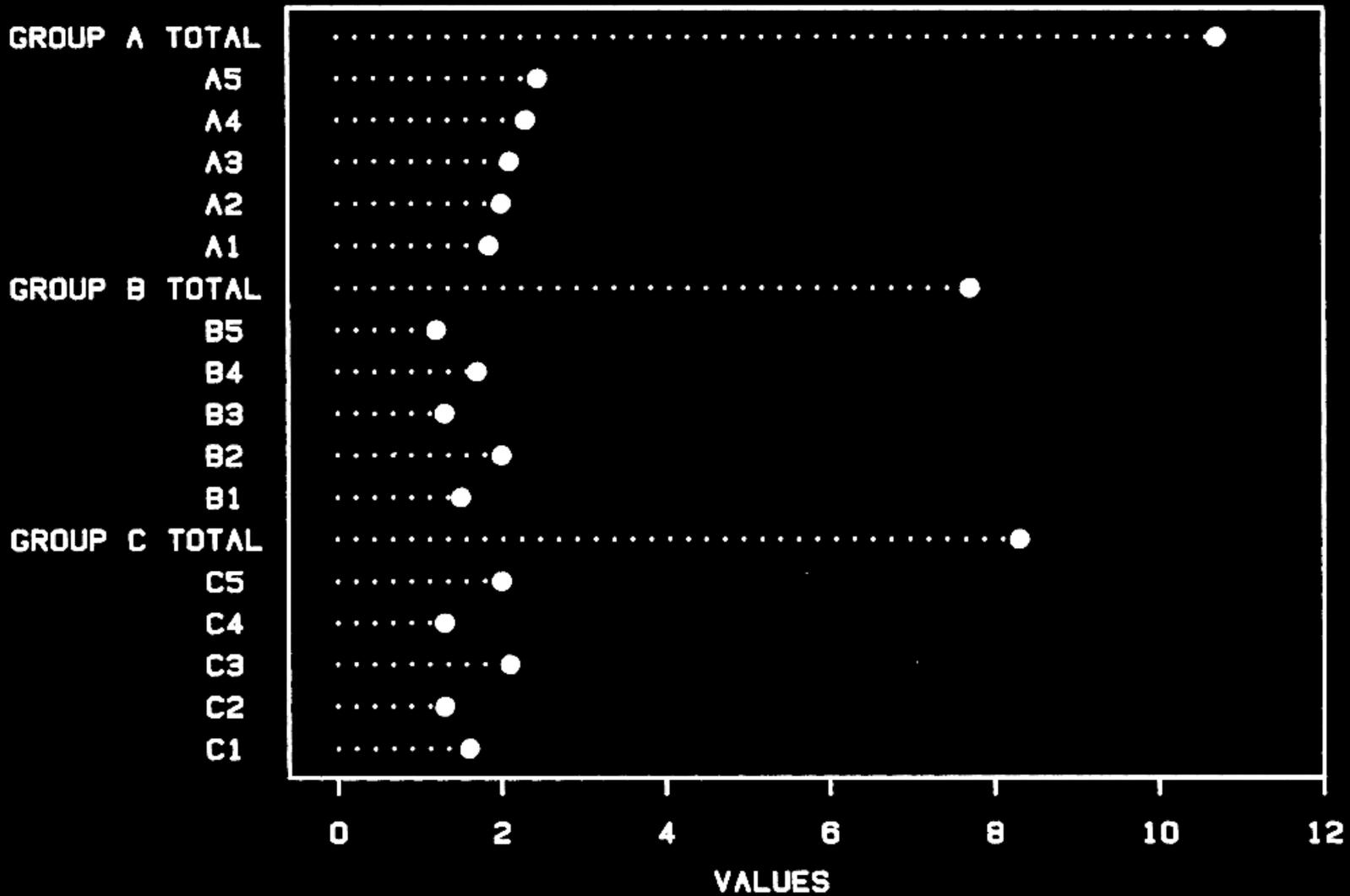
Nearly 200 years ago William Playfair (1786) began the serious use of graphs for looking at data. More than 50 years ago a battle raged on the pages of the *Journal of the American Statistical Association* about the relative merits of bar charts and pie charts (Eells 1926; Croxton 1927; Croxton and Stryker 1927; von Huhn 1927). Today graphs are a vital part of statistical data analysis and a vital part of communication in science and technology, business, education, and the mass media.

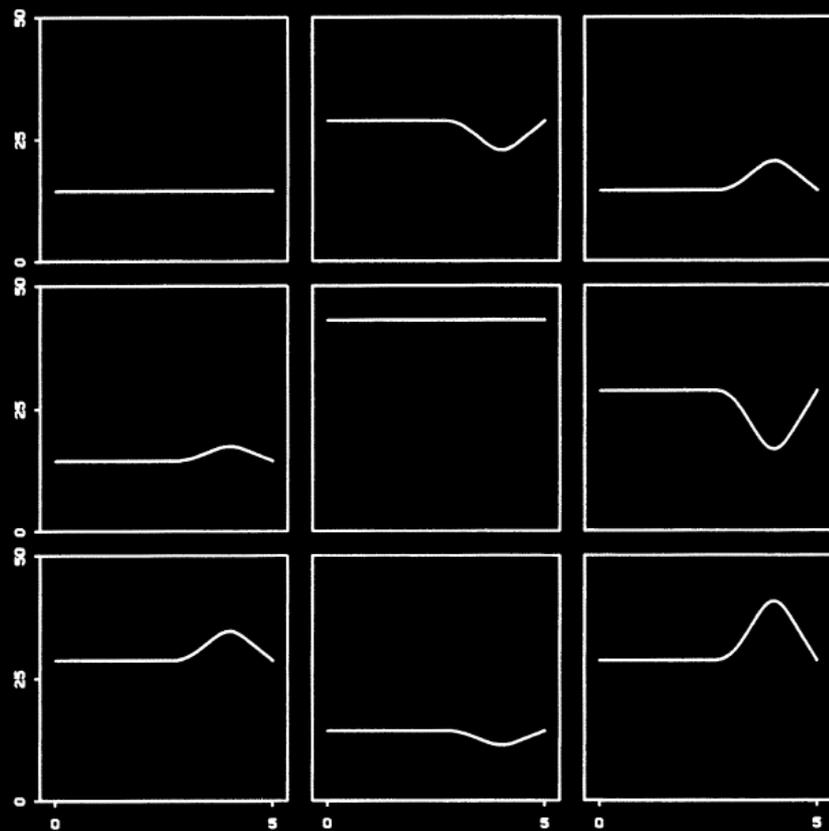
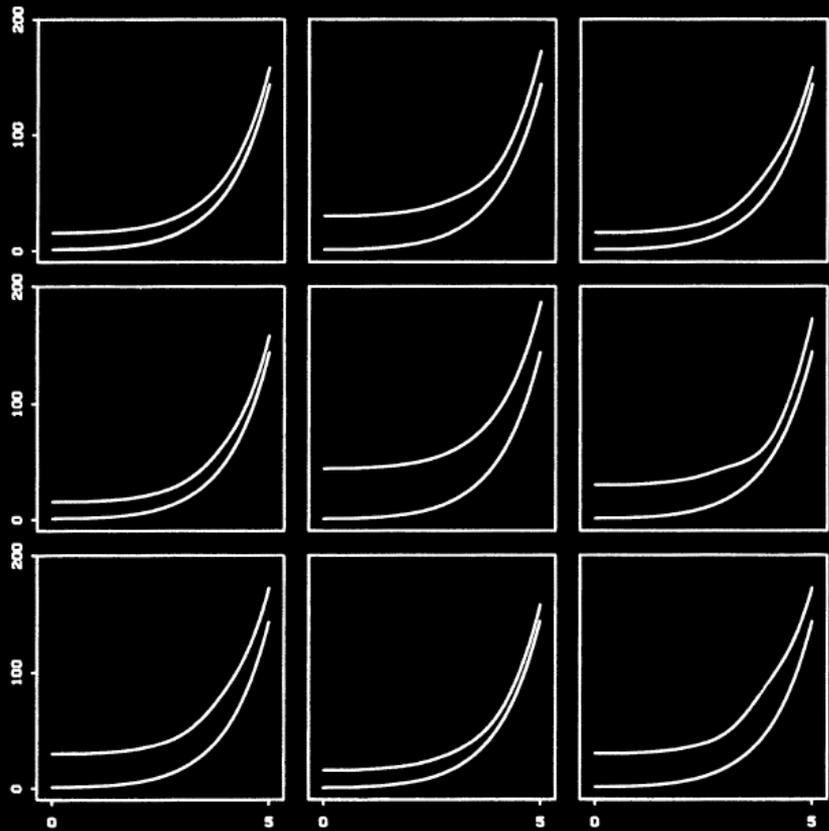
Journal of the American Statistical Association
September 1984, Volume 79, Number 387



Cleveland – McGill, 1984







Cleveland – McGill, 1984

CHART of EXPORTS and IMPORTS to and from the EAST INDIES

From the Year 1700 to 1780 by W. Playfair



1700

1710

1720

1730

1740

1750

1760

1770

1780

The Bottom Line is Divided into Years the Right hand Line into HUNDRED THOUSAND POUNDS

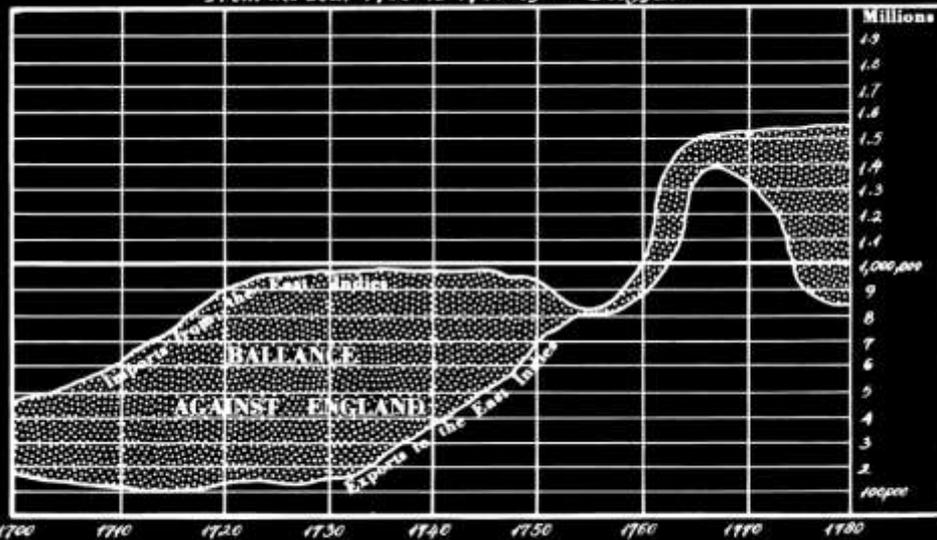
along the Height

is force Page 31st

Published in the *Stet* Review 16th Aug. 1785

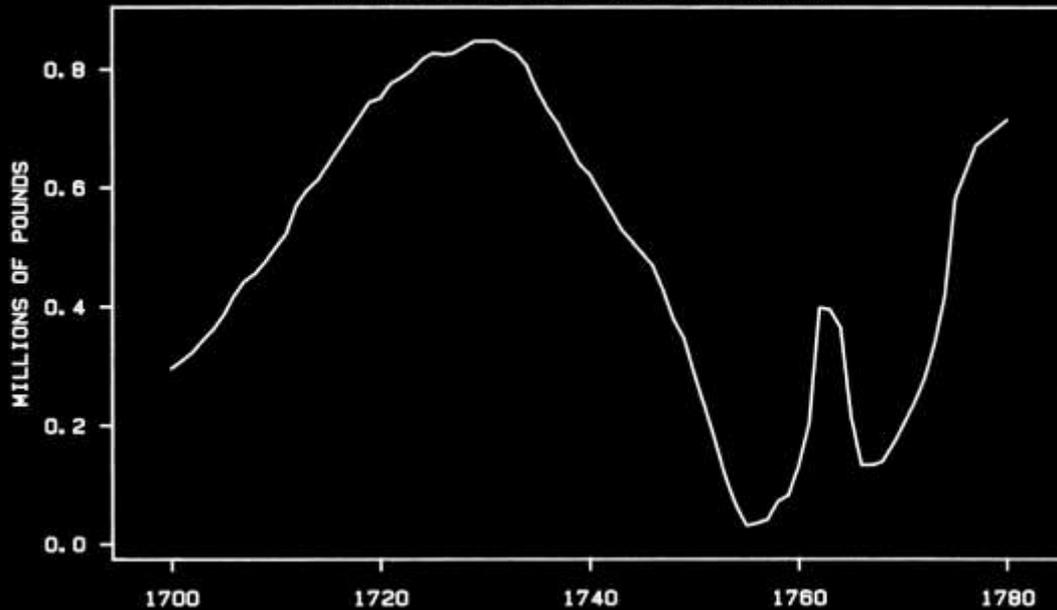
William Playfair

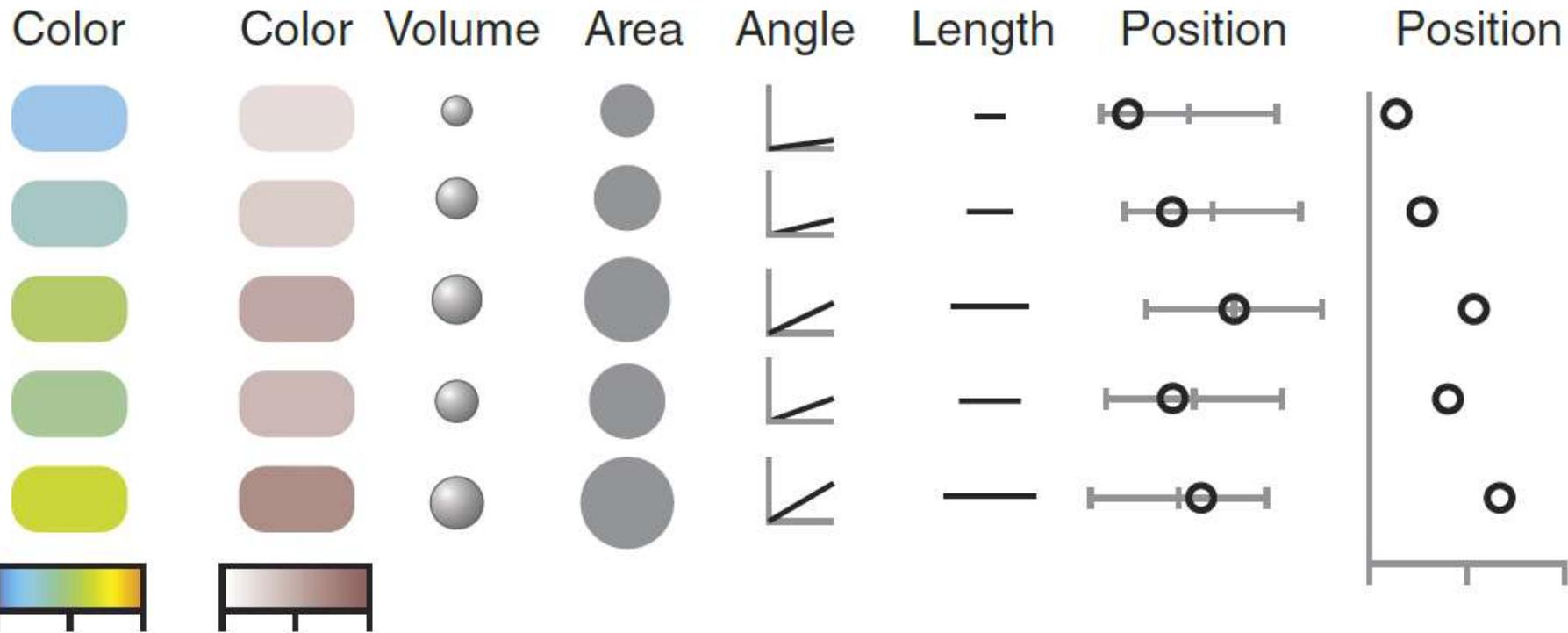
CHART of EXPORTS and IMPORTS to and from the EAST INDIES
From the Year 1700 to 1780 by W. Playfair



The Bottom Line is Divided into Years the Right hand Line into HUNDRED THOUSAND POUNDS
Angelicus Trogus *is from Page 51st* *Published in the Art Magazine 16th Aug. 1795*

CHART OF BALANCE AGAINST ENGLAND





Points of view: Design of data figures
 Bang Wong
 Nature Methods 7, 665 (2010)

How Deceptive are Deceptive Visualizations?: An Empirical Analysis of Common Distortion Techniques

Anshul Vikram Pandey
School of Engineering,
New York University
anshul.pandey@nyu.edu

Katharina Rall
School of Law,
New York University
kr1326@nyu.edu

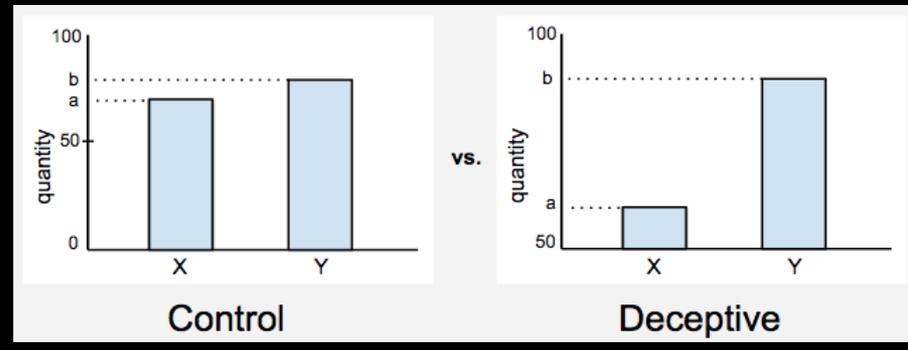
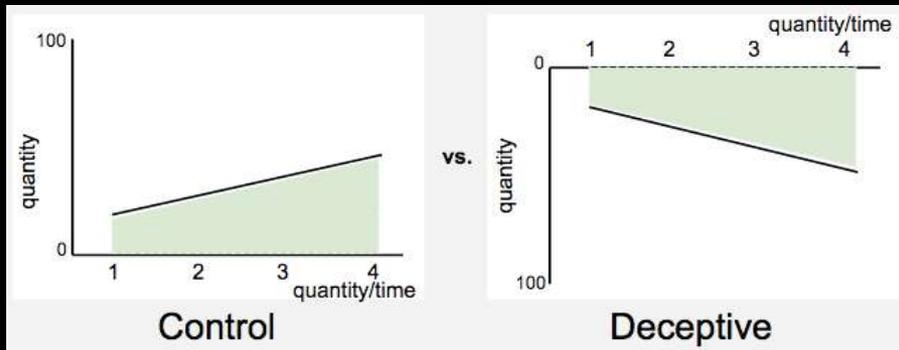
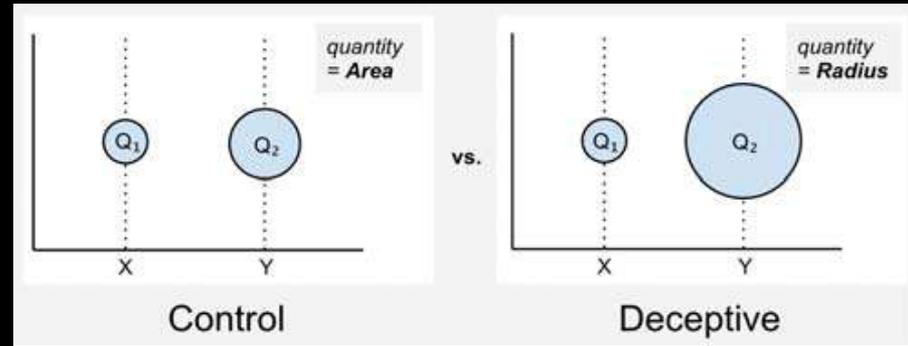
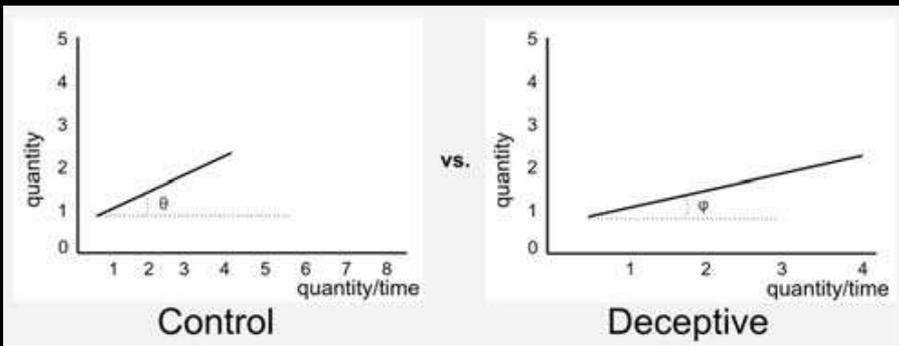
Margaret L. Satterthwaite
School of Law,
New York University
satterth@exchange.law.nyu.edu

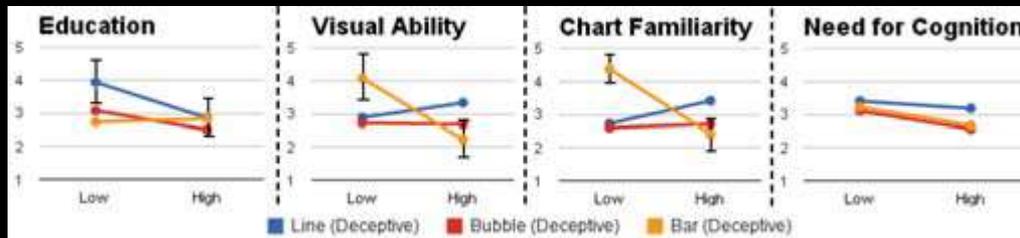
Oded Nov
School of Engineering,
New York University
onov@nyu.edu

School of Engineering,
New York University
enrico.bertini@nyu.edu

ABSTRACT

In this paper, we present an empirical analysis of deceptive visualizations. We start with an in-depth analysis of what deception means in the context of data visualization, and categorize deceptive visualizations based on the type of deception they lead to. We identify popular distortion techniques and the type of visualizations these distortions can be applied to.

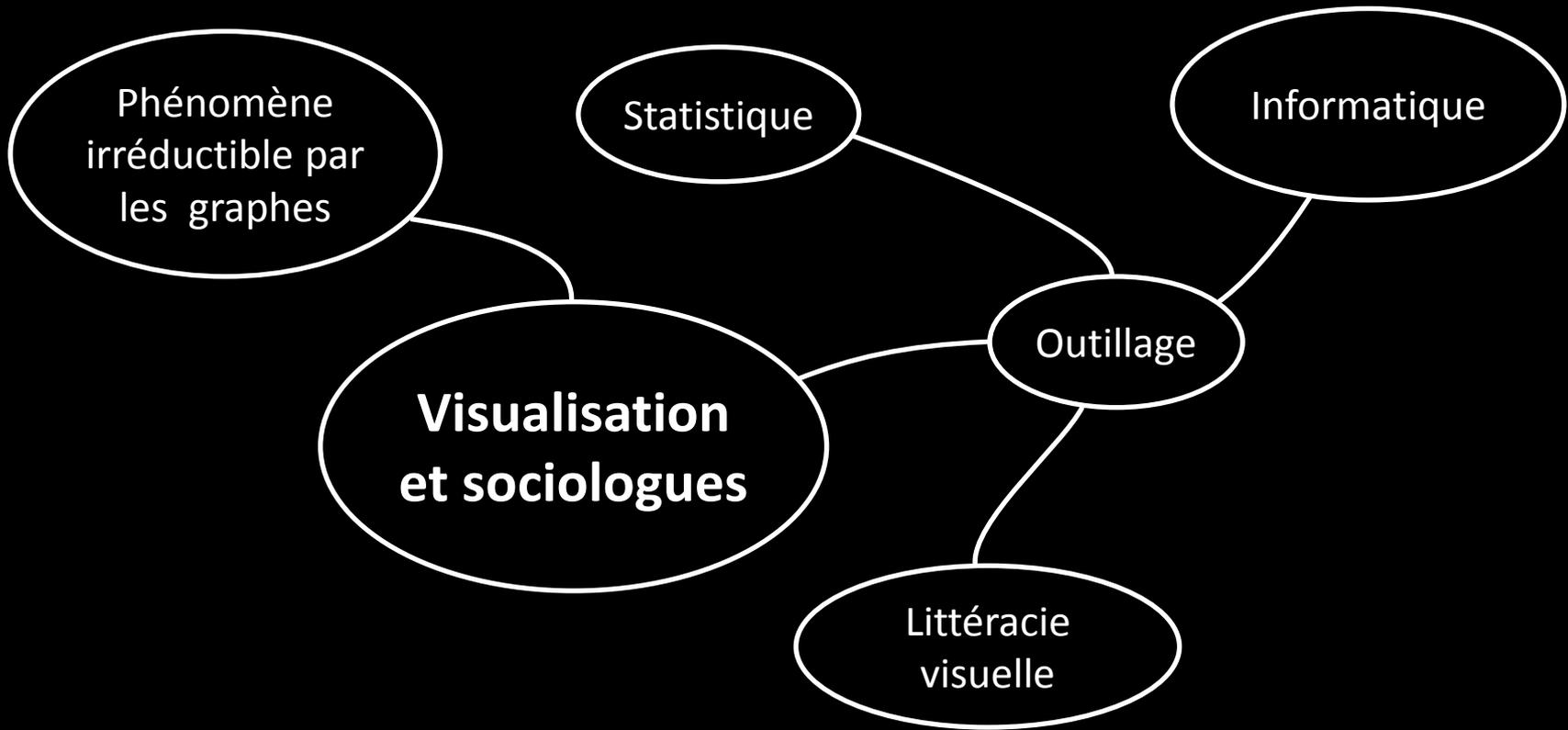


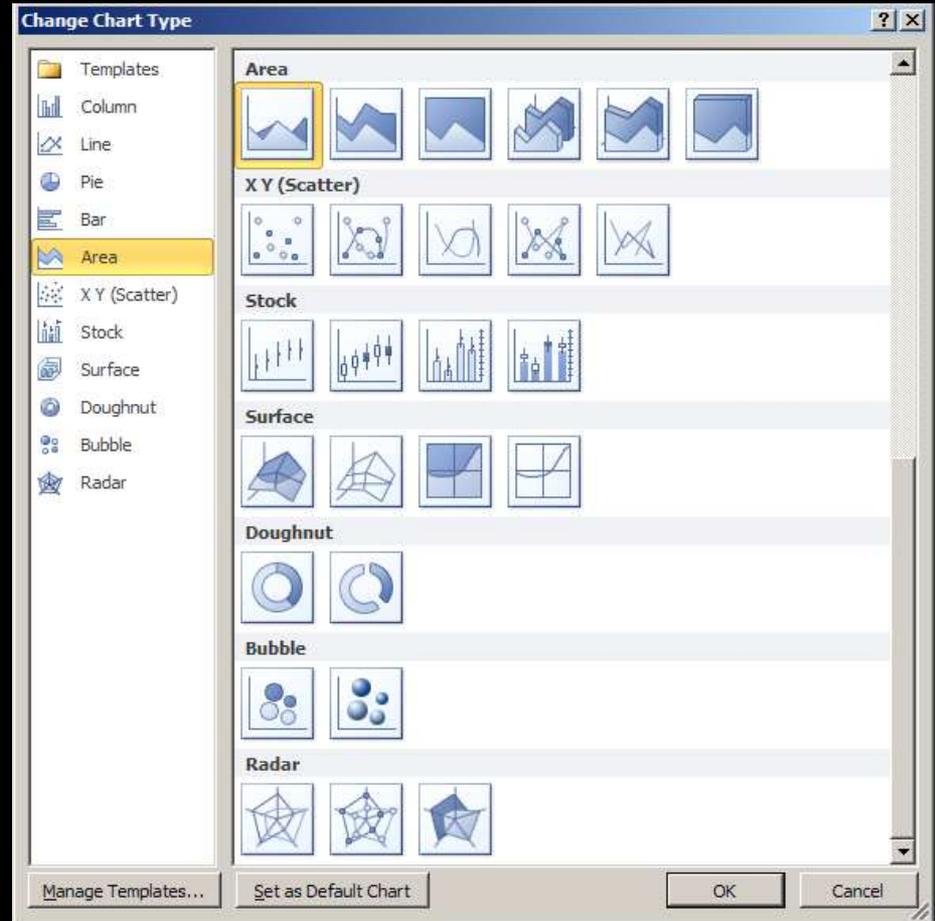
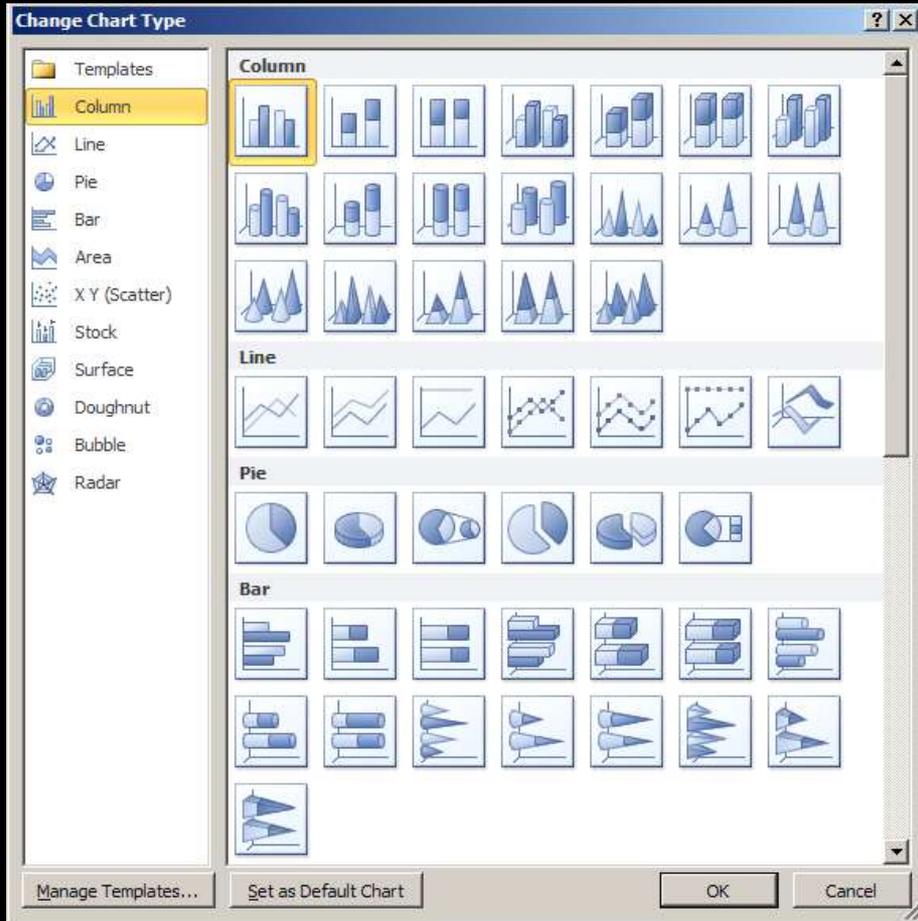


[..]

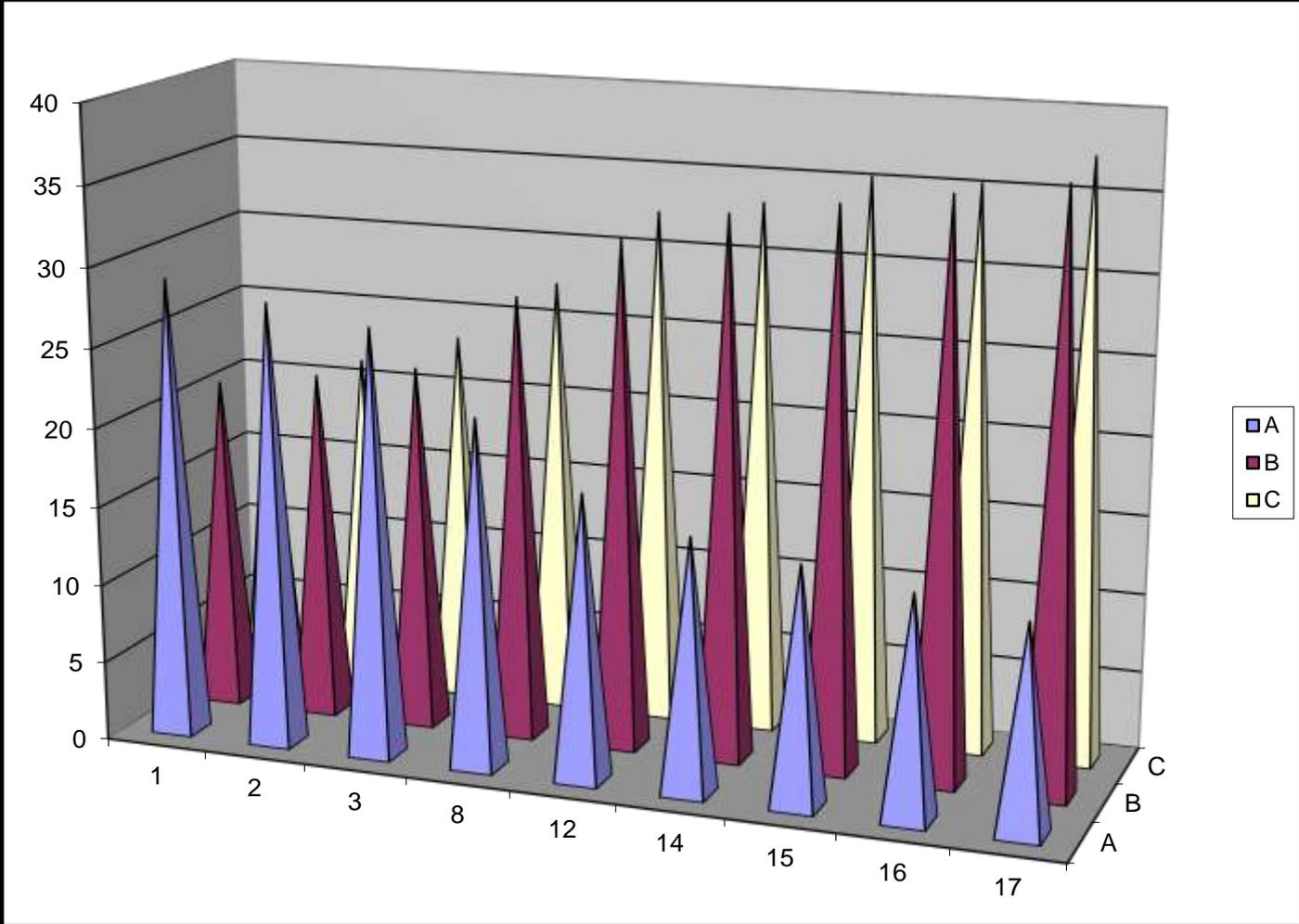
Another interesting direction is to take into account the general graphical literacy of the participants, and see whether it correlates with the response

[..]

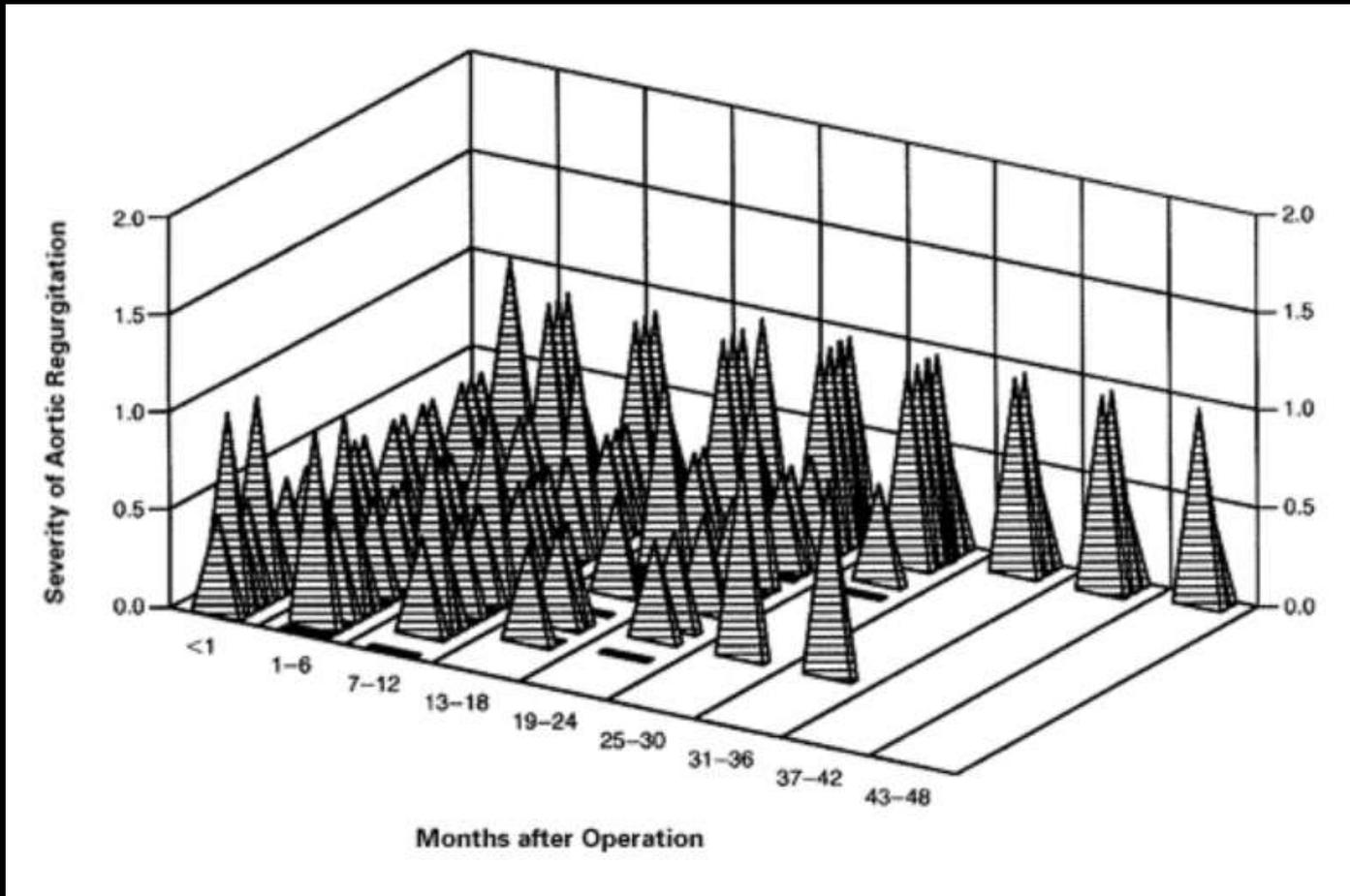




Courtesy Frederic Schütz

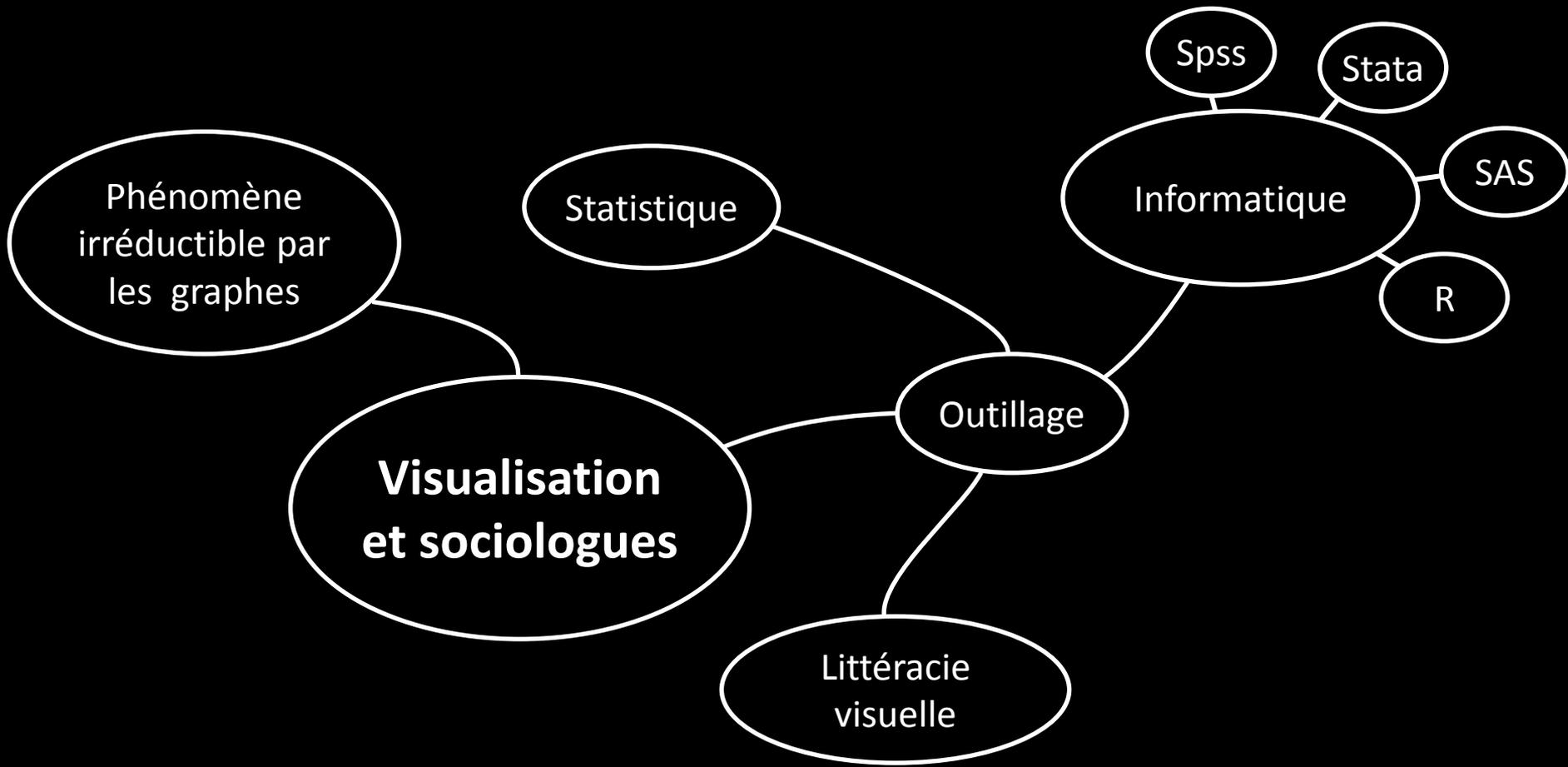


Courtesy: Frederic Schütz



N.T. Kouchoukos et al. "Replacement of the Aortic Root with a Pulmonary Autograft in Children and Young Adults with Aortic-Valve Disease." New England Journal of Medicine, 330 (1994), p.4.
 Cited by Edward Tufte, "The cognitive style of PowerPoint"

Courtesy: Frederic Schütz



Phénomène
irréductible par
les graphes

**Visualisation
et sociologues**

Statistique

Outillage

Littéracie
visuelle

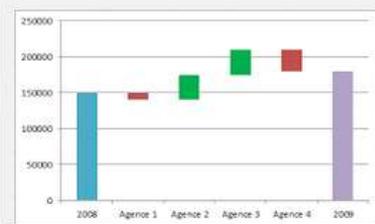
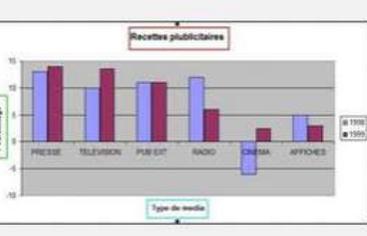
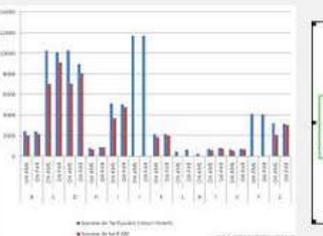
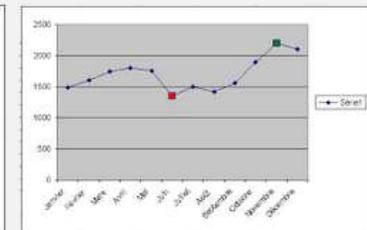
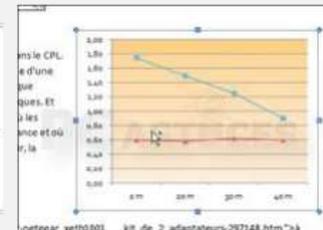
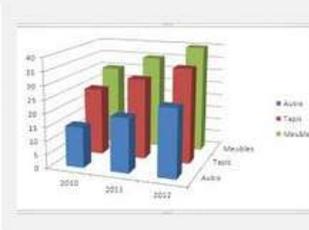
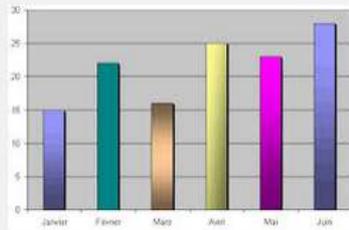
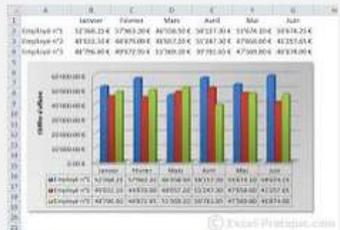
Informatique

Spss

Stata

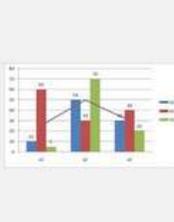
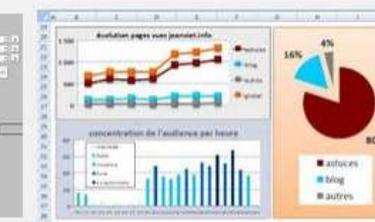
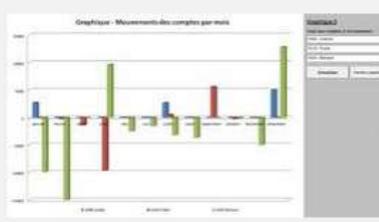
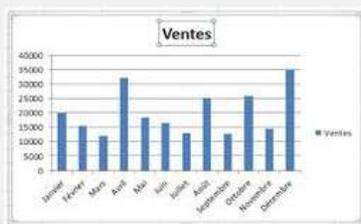
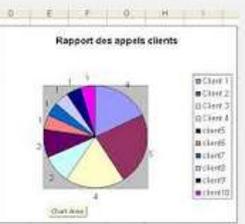
SAS

R



Excel spreadsheet showing a table of client calls. The table has columns for 'Client' and 'Appel'. The data is as follows:

Client	Appel
Client 1	5
Client 2	6
Client 3	8
Client 4	2
Client 5	1
Client 6	1
Client 7	1
Client 8	1
Client 9	1
Client 10	1



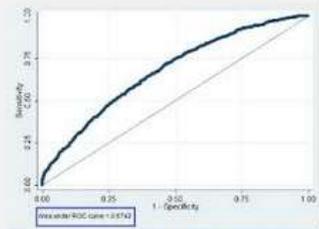
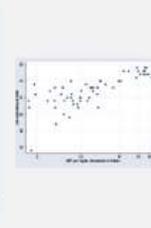
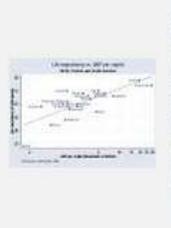
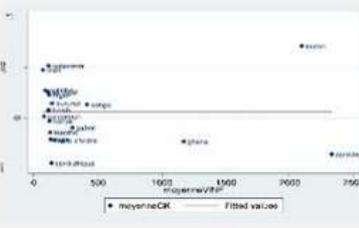
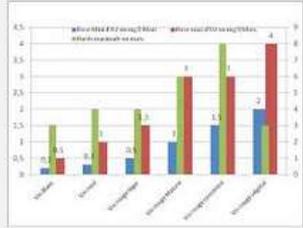
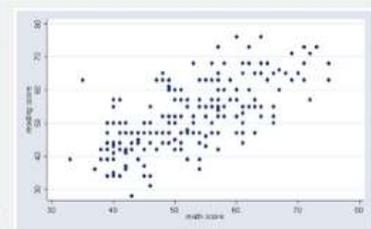
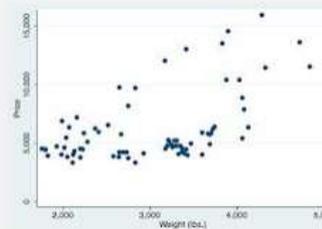
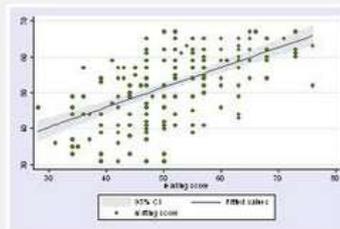
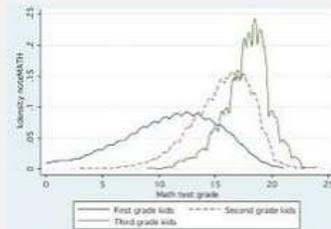
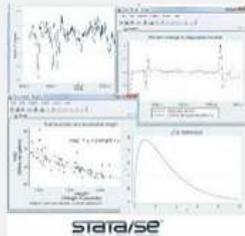


STATA graphique

Web **Images** Vidéos Actualités Shopping Plus Outils de recherche

Safe Search

Connexion



Stata par la pratique :
statistiques, graphiques et
éléments de programmation

```

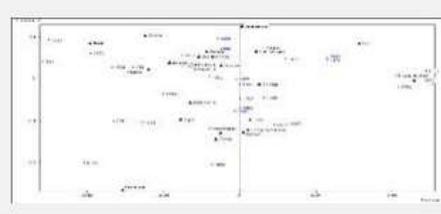
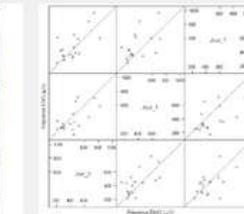
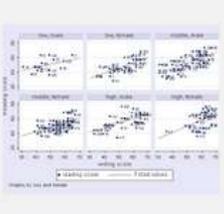
estatistics: Fitted regression
group variable: TI = 8995
-----+-----
_robust_  = 0.2065
Sarg-Jen  = 0.4628
Ljung-Box = 0.5006

df: F(1, 8995) = 0.5712
Prob > F = 0.4548

-----+-----
Number of obs = 9000
Number of groups = 20
Obs per group: min = 25.0
                max = 450
                mean = 450.0
                std. dev. = 10.0
                F(1, 8995) = 1.17
                Prob > F = 0.2782
    
```

Variables	Nombre d'observations	Moyenne	Écart-type	Minimum	Maximum
ENCRE	14	19.38071	12.31548	3.48	49.66
PADEN	14	0.45	0.1830186	0.2	0.7
E_PP	14	0.277857	0.0968866	0.1	0.45
MEAB	14	-0.579288	0.2487932	-1.06	-0.18
CEBR	14	2.55113	0.138842	2.3	2.6
QUAL	14	40.388378	0.1424975	-1.11	-0.59
INPLB	14	3.882143	0.9618287	2	5.88
PER_HAB	14	1855.837	386.7556	1467	1770
YMAR	14	4.614286	0.491242	3.8	5.1
EXALPH	14	92.54286	2.331145	88	94.3

Variable	Obs	Min	Max	Mean	Std. Dev.	Skewness	Kurtosis
ENCRE	14	3.48	49.66	19.38071	12.31548	0.11	2.81
PADEN	14	0.2	0.7	0.45	0.1830186	0.11	2.81
E_PP	14	0.1	0.45	0.277857	0.0968866	0.11	2.81
MEAB	14	-1.06	-0.18	-0.579288	0.2487932	0.11	2.81
CEBR	14	2.3	2.6	2.55113	0.138842	0.11	2.81
QUAL	14	-1.11	-0.59	40.388378	0.1424975	0.11	2.81
INPLB	14	2	5.88	3.882143	0.9618287	0.11	2.81
PER_HAB	14	1467	1770	1855.837	386.7556	0.11	2.81
YMAR	14	3.8	5.1	4.614286	0.491242	0.11	2.81
EXALPH	14	88	94.3	92.54286	2.331145	0.11	2.81



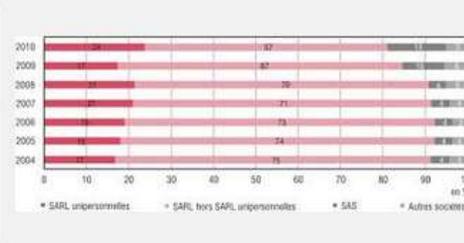
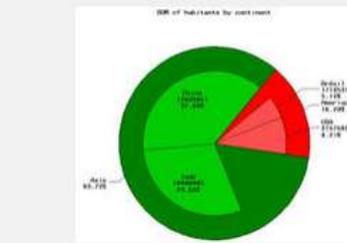
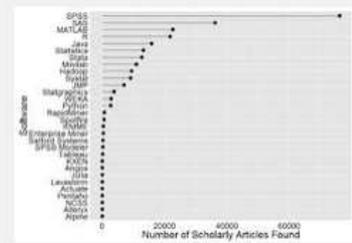
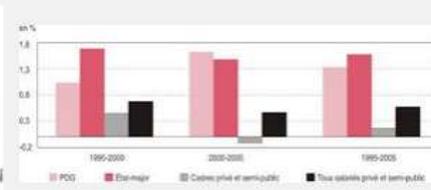
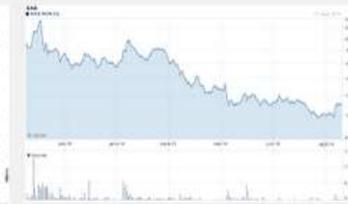
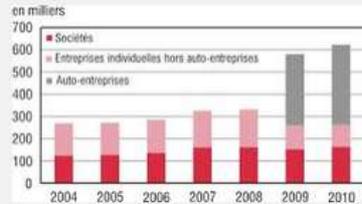
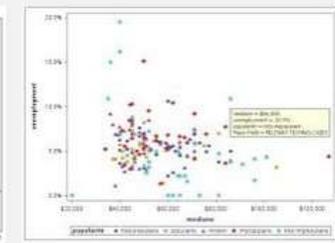
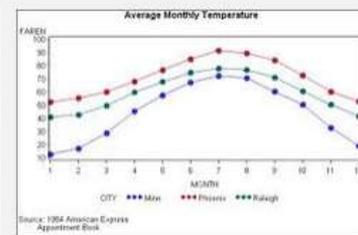
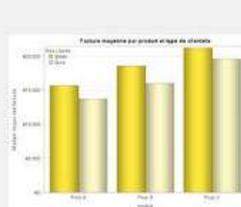
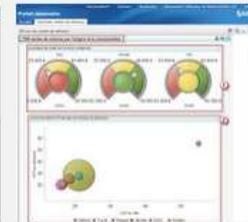
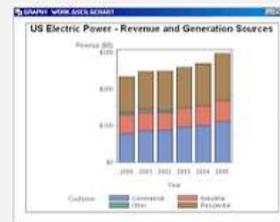
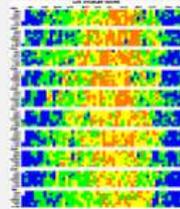
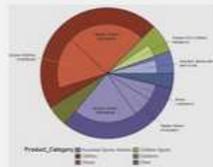
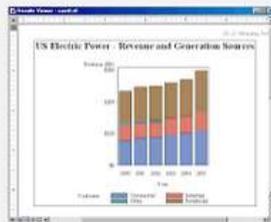
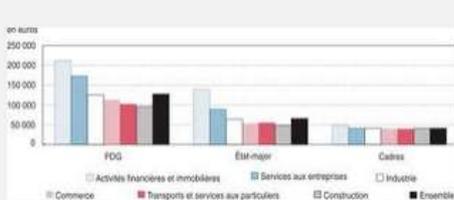


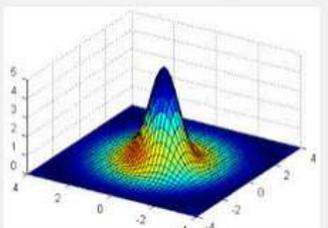
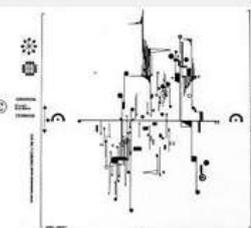
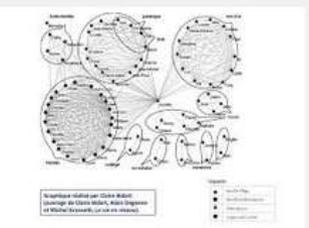
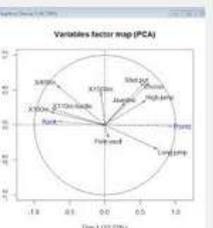
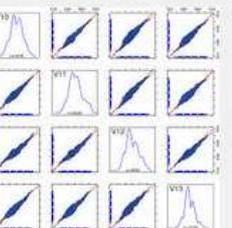
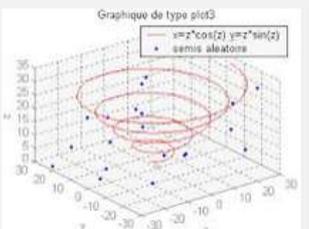
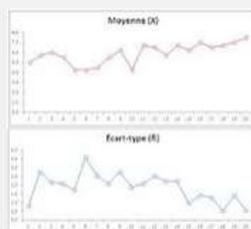
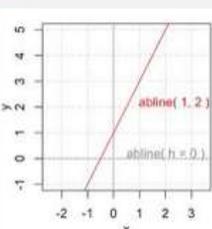
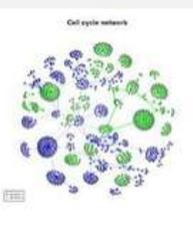
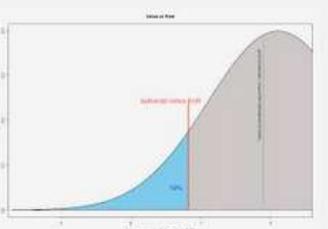
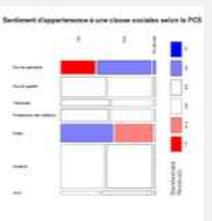
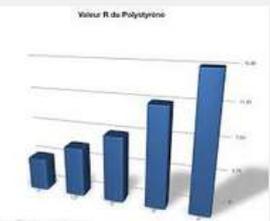
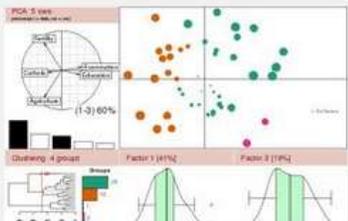
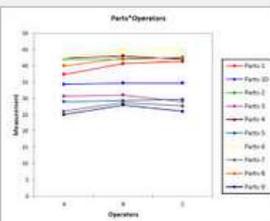
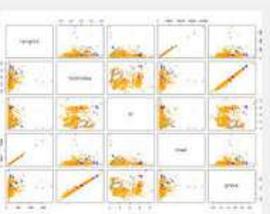
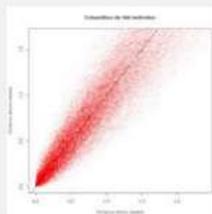
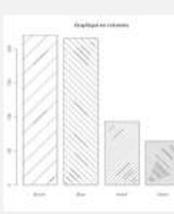
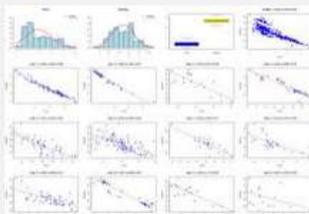
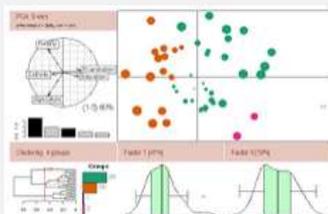
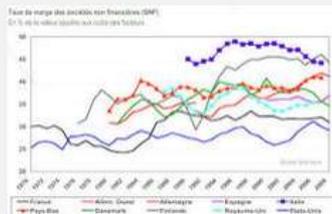
SAS graphique

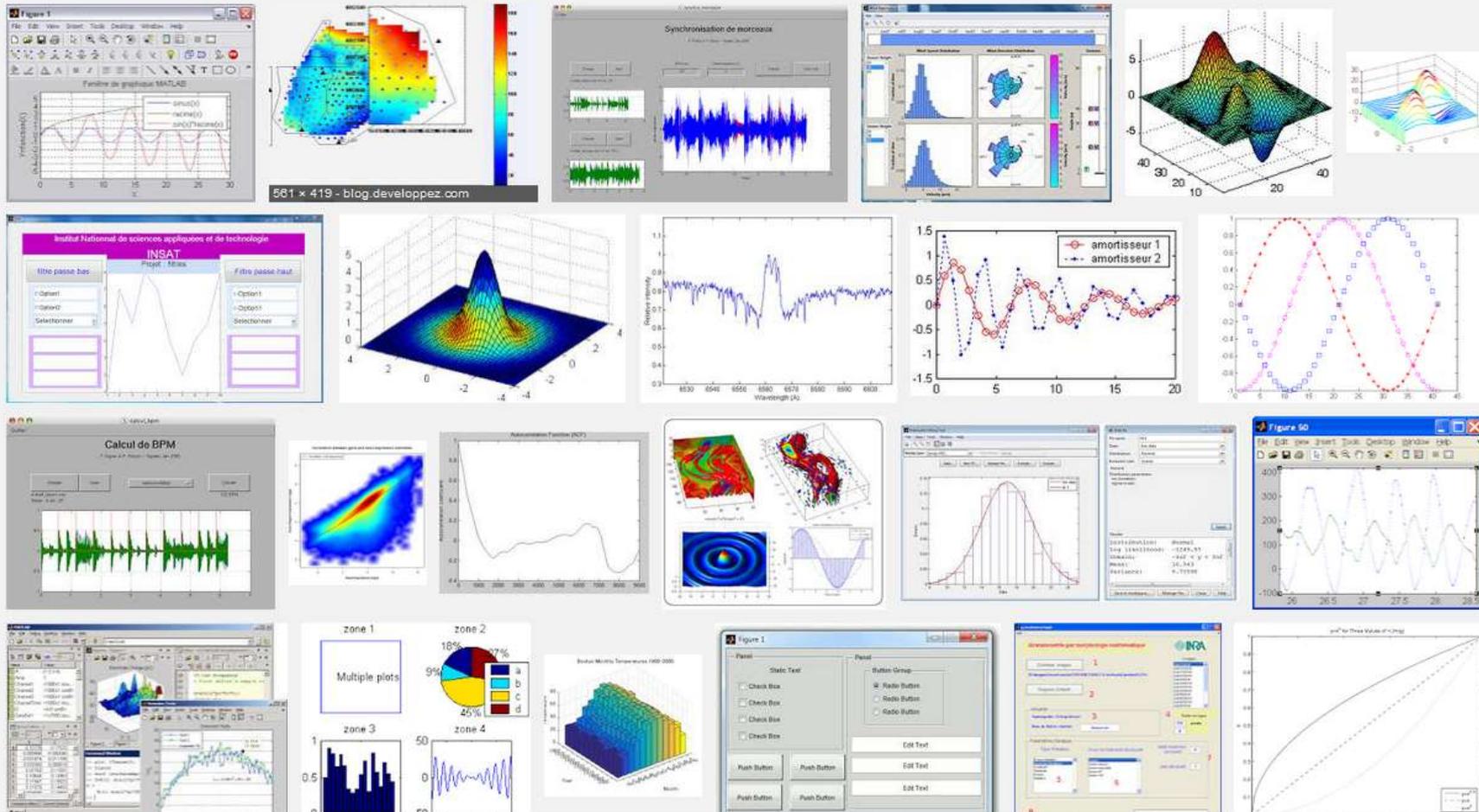
Connexion

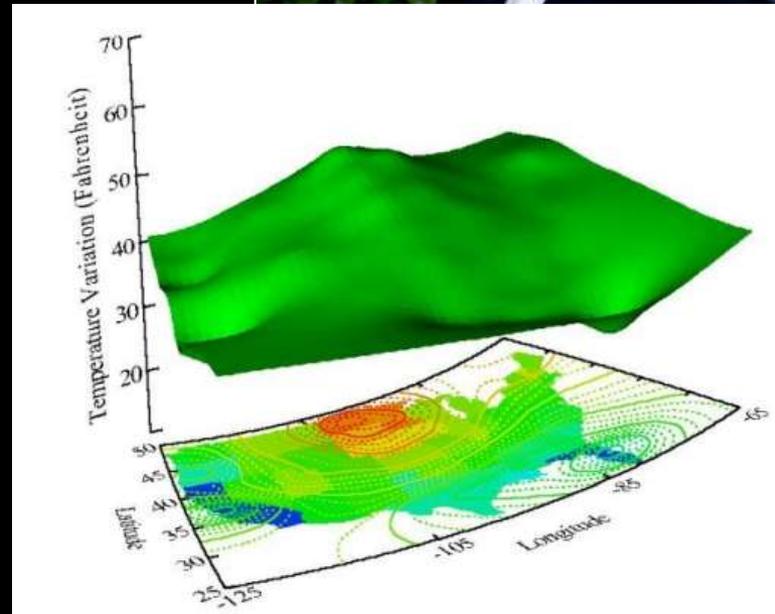
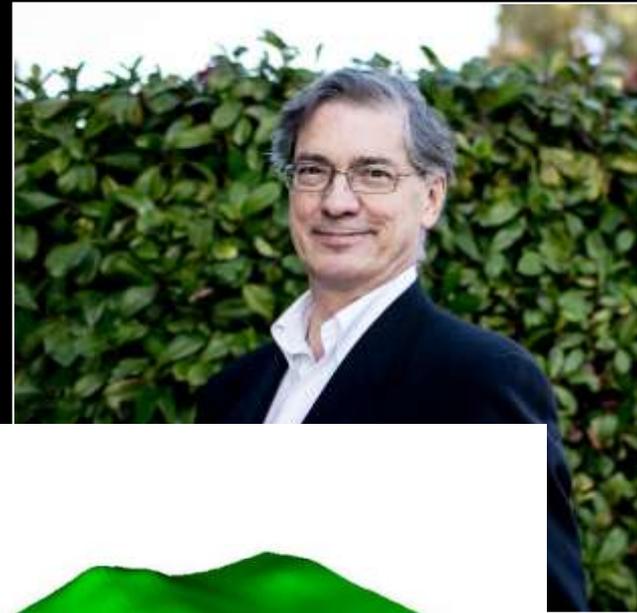
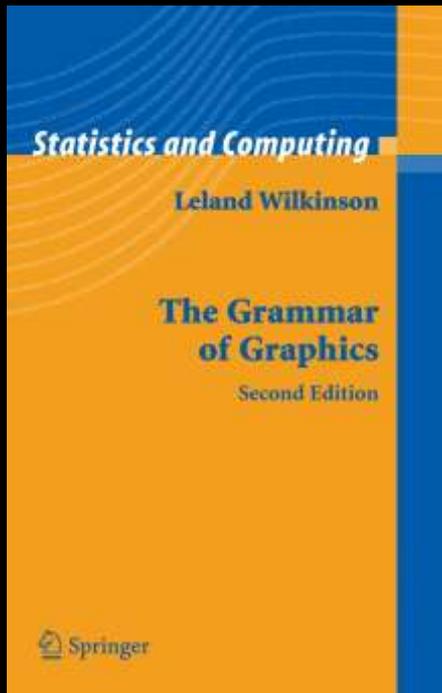
Web Images Vidéos Actualités Shopping Plus Outils de recherche

Safe Search









Leland Wilkinson (1945-)

Chrome File Edit View History Bookmarks Window Help

Workbook: Malte Spitz Call

public.tableausoftware.com/views/MalteSpitzCallData/MalteSpitzcalldatadashboard?:embed=yes&toolbar=yes&tabs=yes

Calls by weekday | Calls by service type by weekday | Calls by daytime | Calls by service type and daytime | Call duration by service type and daytime | Malte Spitz call data dashboard

Calls by service type and geo location

Map data © OpenStreetMap (and) contributors, CC-BY-SA

Calls by service type by weekday

Call Start

Day	GPRS/3G	Phone	SMS	MMS
Monday	~2800	~800	~500	~200
Tuesday	~2900	~800	~500	~200
Wednesday	~2500	~800	~500	~200
Thursday	~2400	~800	~500	~200
Friday	~2800	~800	~500	~200
Saturday	~3000	~800	~500	~200
Sunday	~3100	~800	~500	~200

Calls by service type

Service Type

Service Type	Percentage
GPRS/3G	64.00%
Phone	15.78%
SMS	10.57%
Unknown	9.48%
MMS	0.09%
MBox	0.07%
WLAN	0.01%

Calls by daytime and service type

Call Start

Hour	GPRS/3G	Phone	SMS
0	~800	~100	~100
1	~500	~100	~100
2	~300	~100	~100
3	~250	~100	~100
4	~250	~100	~100
5	~250	~100	~100
6	~250	~100	~100
7	~400	~100	~100
8	~1000	~200	~100
9	~1500	~300	~200
10	~1500	~300	~200
11	~1500	~300	~200
12	~1800	~400	~200
13	~1600	~400	~200
14	~1600	~400	~200
15	~1700	~400	~200
16	~1600	~400	~200
17	~1600	~400	~200
18	~1600	~400	~200
19	~1500	~400	~200
20	~1300	~400	~200
21	~1100	~400	~200
22	~1000	~400	~200
23	~1000	~400	~200

Service Type

- Unknown
- GPRS/3G
- Phone
- SMS
- MMS

Call Type

(Multiple values)

Call Start

8/31/2009 07: 2/27/2010 11:

Weekday of Call Start

- (All)
- Null
- Sunday
- Monday
- Tuesday
- Wednesday

Hour of Call Start

- (All)
- Null
- 0
- 1
- 2
- 3

Number of Calls

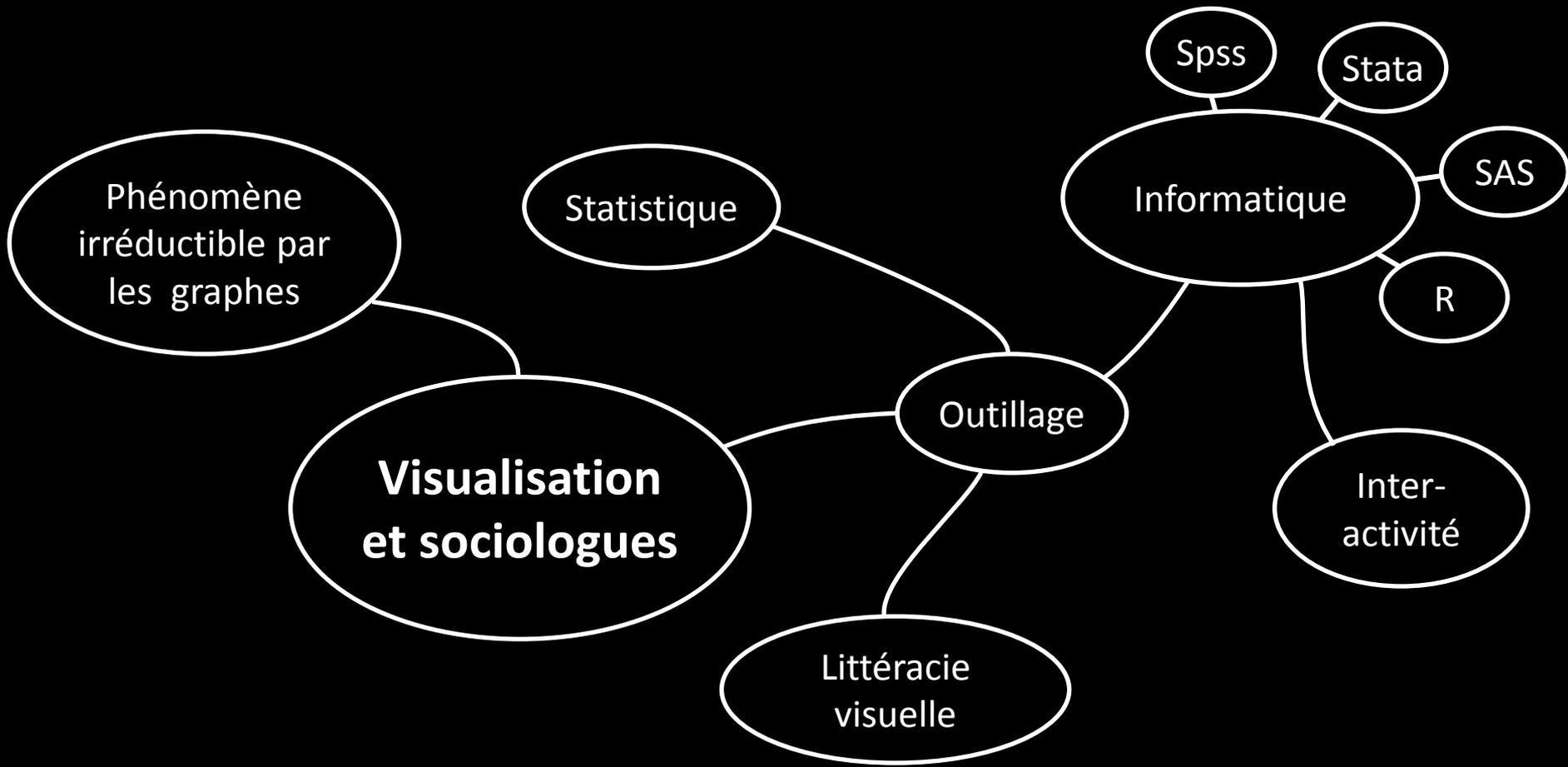
1, 1,000, 2,000, 3,000, 4,000

Share

Download

tableau

Die Zeit / Electronic Frontier Foundation,
29. Mars 2011



Data Desk® 7.0.2 Viewer - Untitled

Data Desk File Edit Data Special Modify

Regression: Weight by MPG Scatterplot: Weight by ...

Response variable is: **MPG**
 No Selector
 91 total cases of which 41 are missing
 R squared = 75.6% R squared (adjusted) = 75.1%
 s = 2.413 with 50 - 2 = 48 degrees of freedom

Source	Sum of Squares	df	Mean Square	F-ratio
Regression	865.41	1	865.41	149
Residual	279.57	48	5.82436	

Variable	Coefficient	SE(Coeff)	t-ratio	p-value
Intercept	48.7393	1.976	24.7	≤ 0.0001
Weight	-0.00821362	673.8e-6	-12.2	≤ 0.0001

MPG

Weight

Regression Guidance

A regression is a formal summary of the linear relationship between two quantitative variables. You should always look at a scatterplot of the two variables to ensure that the underlying relationship is linear.

The regression calculates a straight line equation of the form $y = a + bx$. The coefficients a and b are the constant coefficient and slope, respectively. The constant coefficient is reported in the second column at the bottom of the regression table.

The R-squared statistic, also called the coefficient of determination, measures the fraction of variability of y that is explained by linear regression on x . R-squared is a value between 0 and 1. A value close to 1 indicates that there is a strong linear relationship. R-squared is reported as a percentage in the regression table.

Home
 Glossary
 Exit to Excel

A/p/S/p Plot

Assp

S/p

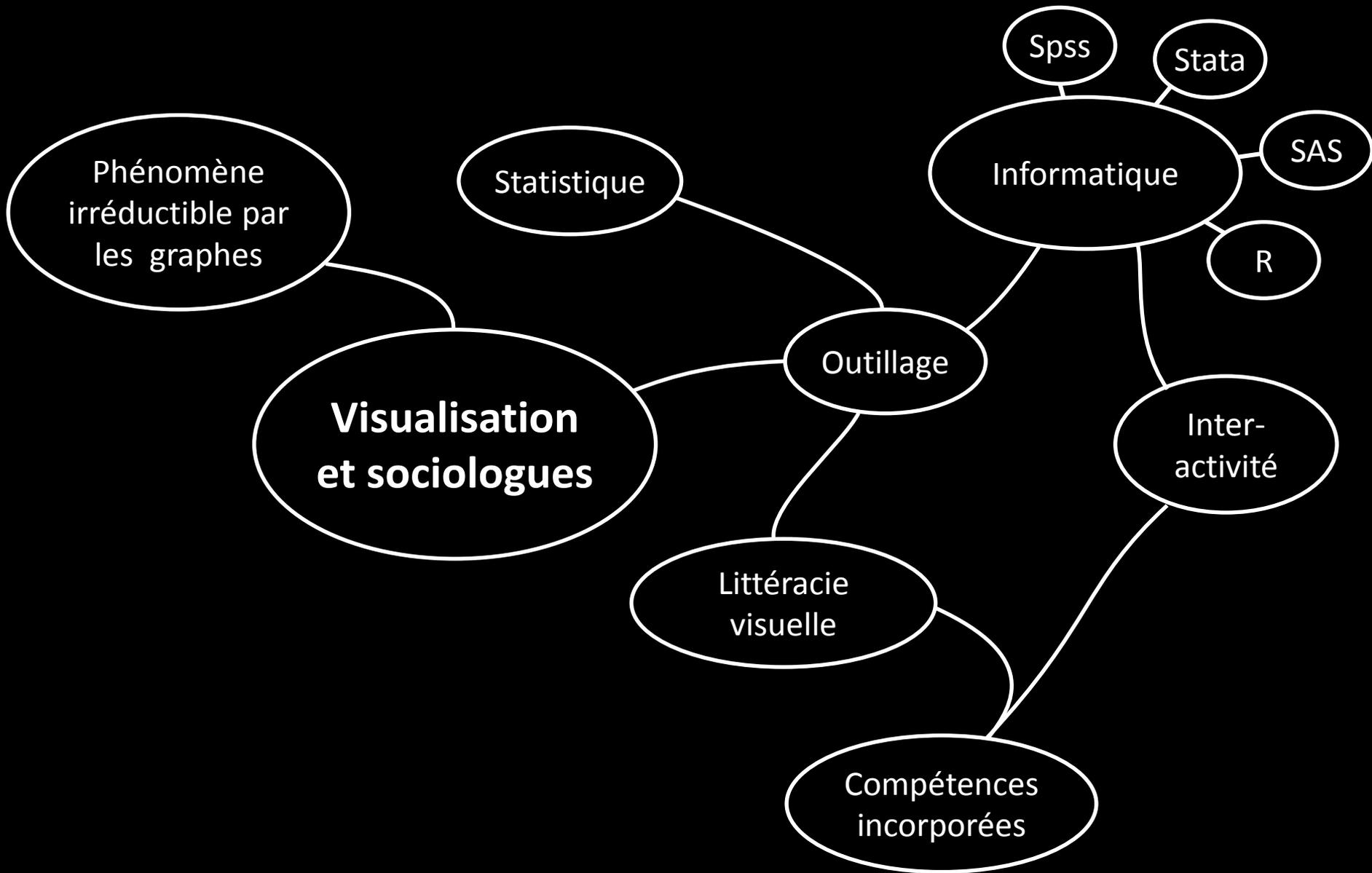
Assets^p Histogram

Sales^p Histogram

File

HELP

Trash



Phénomène
irréductible par
les graphes

Statistique

**Visualisation
et sociologues**

Outillage

Informatique

Spss

Stata

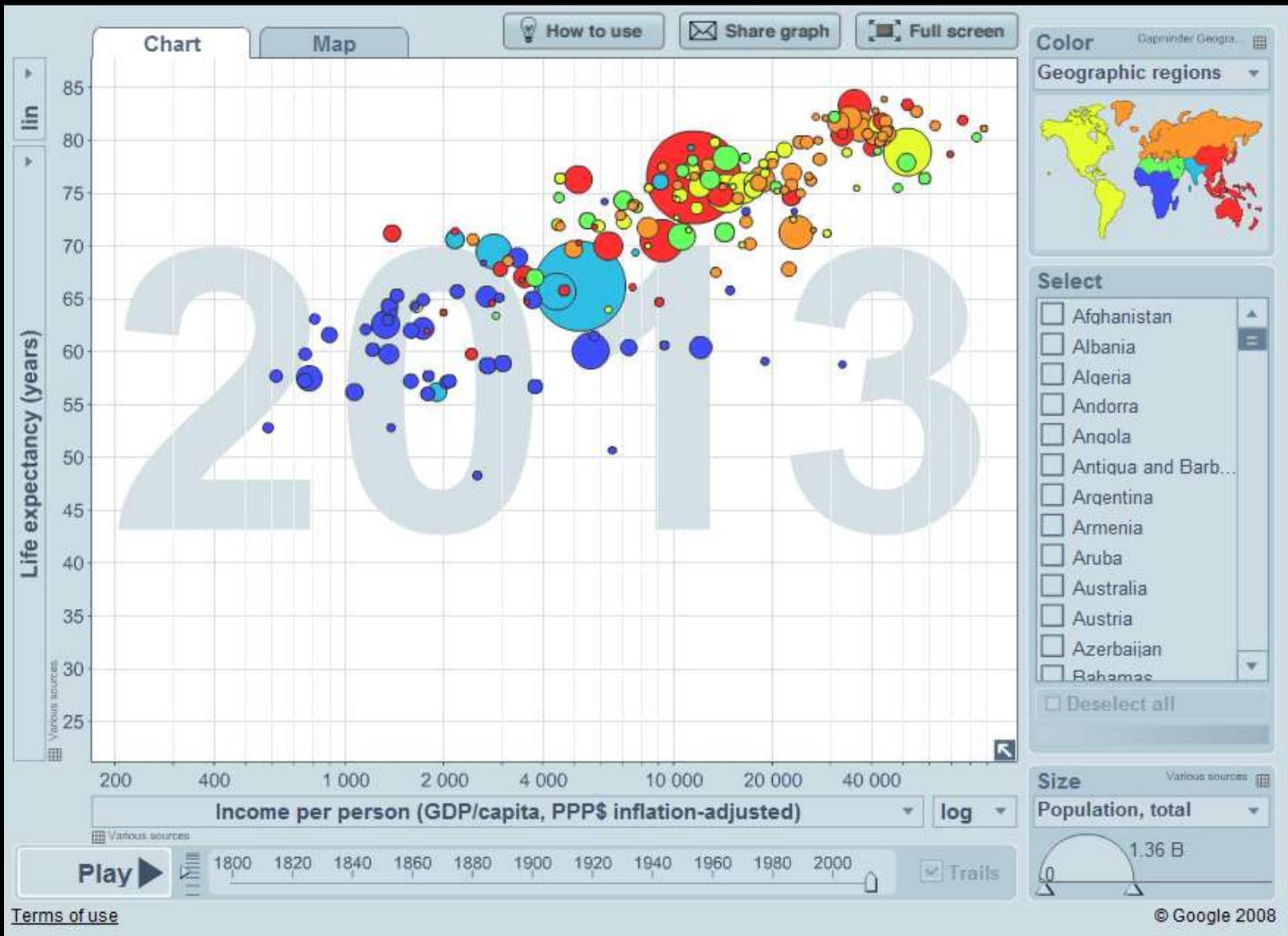
SAS

R

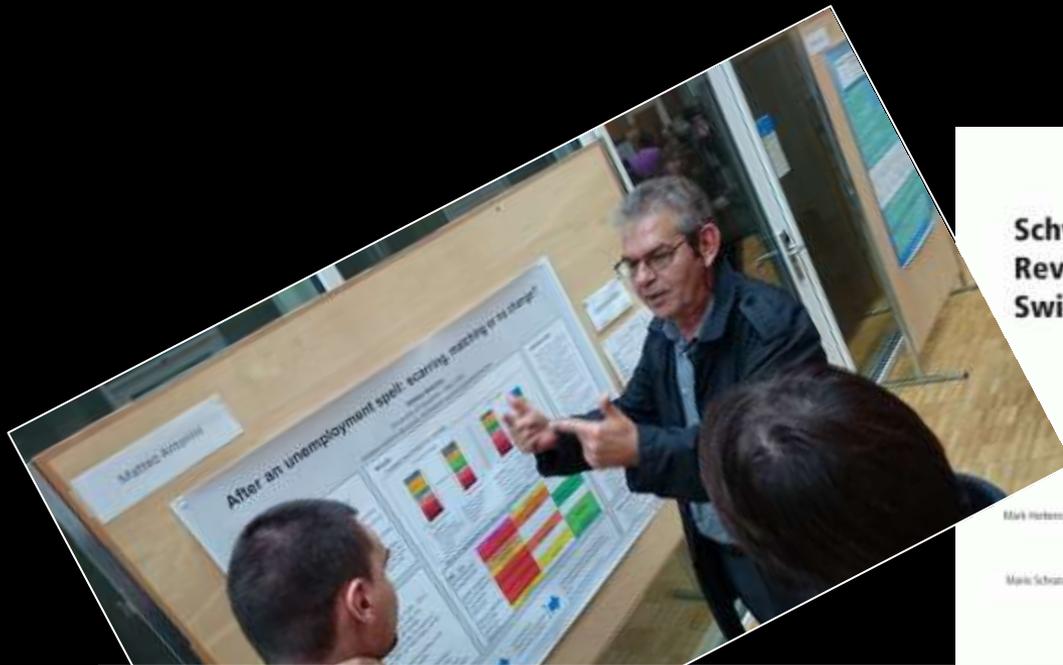
Littéracie
visuelle

Inter-
activité

Compétences
incorporées



Hans Rosling (1948-),



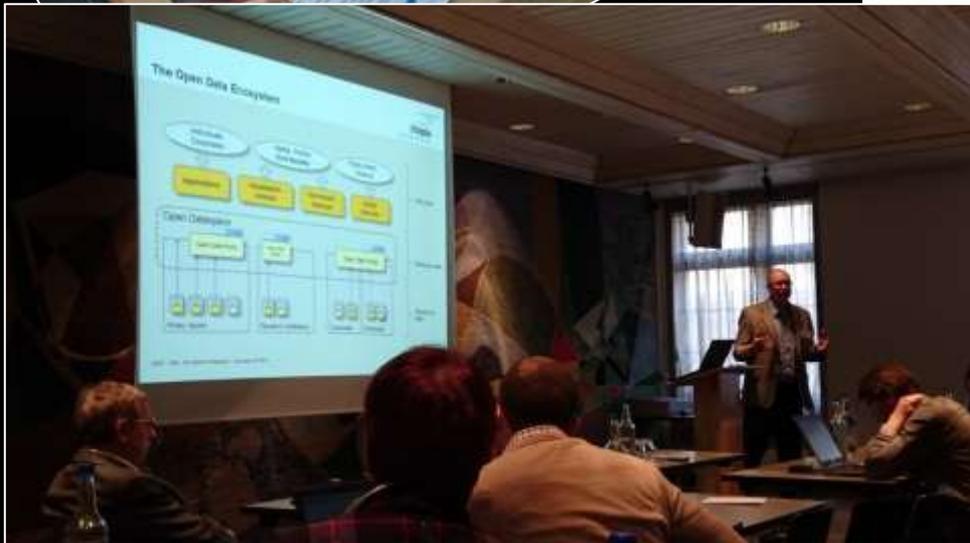
Schweizerische Zeitschrift für Soziologie
Revue suisse de sociologie
Swiss Journal of Sociology

Vol. 37 Issue 2

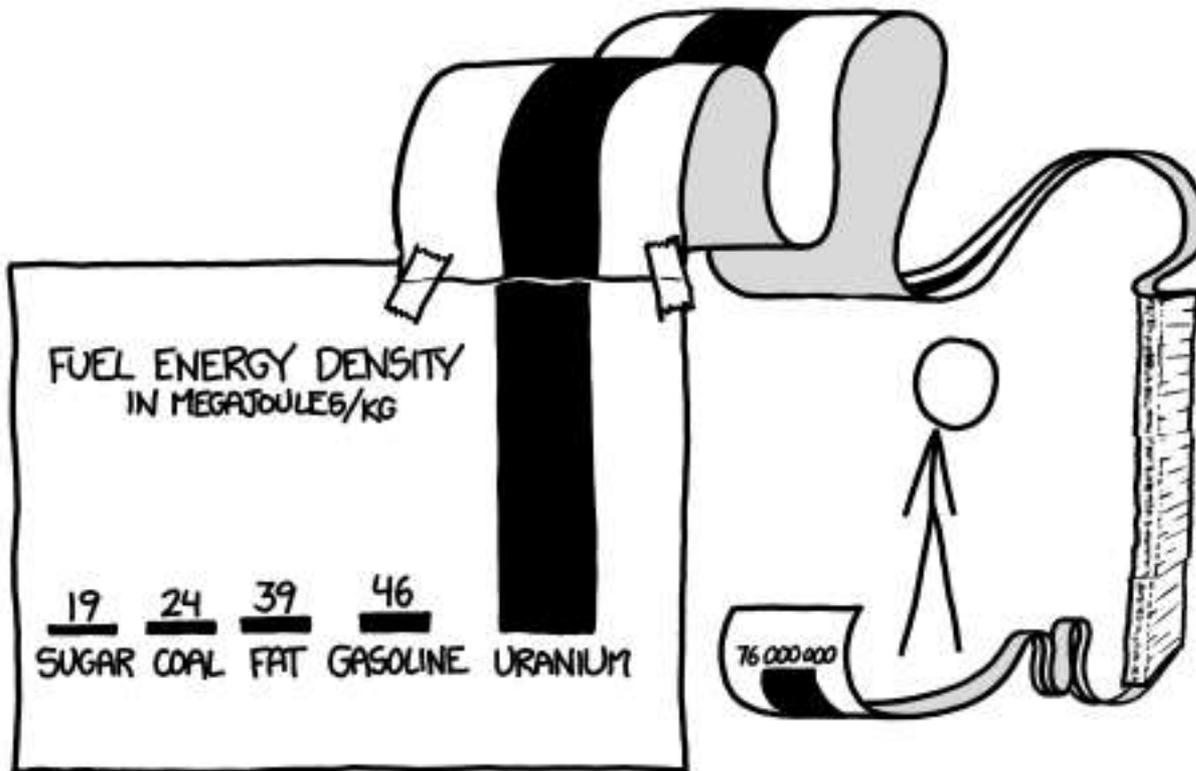
The Global Economic Crisis: Perceptions and Impacts
 Edited by Mark Henkeloth and Christian Söfer

- Mark Henkeloth and Christian Söfer: World Economic Crisis and World Society: Introduction to the Special Issue (I)
- Alberto Mattioli: Some Neglected Aspects of the Global Crisis (II)
- Marie Schanz and Mark Hansegger: The Media Construction of the Financial Crisis in a Comparative Perspective – An Analysis of Newspapers in the UK, the USA and Switzerland between 2007 and 2009 (I)
- Steffen Pöhl: Economics in Crisis: From Mass-Media Discourses towards a Sociology of Economics (I)
- Steffen Pöhl: The Impact of the Global Financial Crisis on Australia (II)
- Yu-long Wang: Global Crisis, China, and the Strange Demise of the East Asian Model (I)
- Guo Jitong: The Global Economic Crisis as Disclosure of Different Types of Capitalism in Latin America (I)
- Steffen Pöhl, Maria Rotstein and Sebastian Scheel: Financial Crisis in Chile and Costa Rica: Perceptions of Households in Precarious Prosperity (I)

Book Reviews / Abstracts



Articles, congrès, posters



SCIENCE TIP: LOG SCALES ARE FOR QUITTERS WHO CAN'T FIND ENOUGH PAPER TO MAKE THEIR POINT *PROPERLY*.

