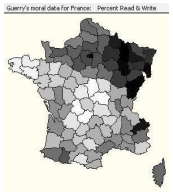
André-Michel Guerry and the Rise of Moral Statistics

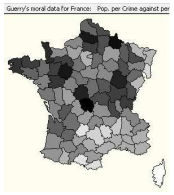
Challenges for Multivariable Spatial Analysis

Michael Friendly

Psychology Department
York University

May 2006
Joint Statistical Meetings



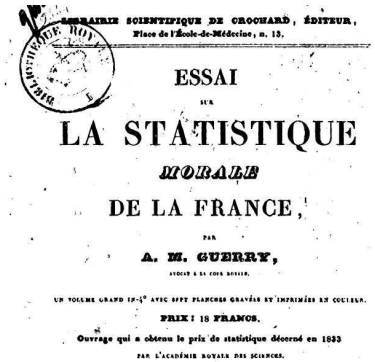


Outline

Essai sur la statistique moral de la France

The launching pad of modern social science

- Presented to Academie des Sciences Français July 2, 1832
- First systematic analysis of comprehensive data on crime, suicide, and other social variables.
- Along with Quetelet (1831, 1835), established the study of "moral statistics" $\mapsto$ modern social science, criminology, sociology


LIBRAIRIE SCIENTIFIQUE DE CROCHARD, ÉDITEUR,
Place de l'École-de-Médecine, n. 13.

ESSAI
SUR
LA STATISTIQUE
MORALE
DE LA FRANCE,
PAR
A. M. GUERRY,
AVOCAT À LA COUR ROYALE.

UN VOLUME GRAND IN-4° AVEC SEPT PLANCHES GRAVÉES ET IMPRIMÉES EN COULEUR.

PRIX : 18 FRANCS.

Ouvrage qui a obtenu le prix de statistique décerné en 1833
PAR L'ACADÉMIE ROYALE DES SCIENCES.

The discovery of "social facts"

Stability and Variation

Guerry's results were both compelling and startling:

- Rates of crime and suicide remained remarkably invariant over time, yet varied sytematically by region, sex of accused, type of crime, etc.
- In any given French city or department, almost the same number committed suicide, stole, gave birth out of wedlock, etc.

Year	1826	1827	1828	1829	1830	Avg
Sex			All accused (%)			
Male	79	79	78	77	78	78
Female	21	21	22	23	22	22
Age			Accused of Theft (%)			
16–25	37	35	38	37	37	37
25–25	31	32	30	31	32	31
Crime			Committed in summer (%)			
Indecent assault	.	36	36	35	38	36
Assault & battery	.	28	27	27	27	28

The discovery of “social facts”

Social laws á la physical laws

Do crime and other moral variables represent:

- structural, lawful characteristics of society, or are they
- simply indicants of individual behaviour?

Guerry argued:

> *Each year sees the same number of crimes of the same degree reproduced in the same regions.(Guerry, 1833, p.10)*

> *... We are forced to recognize that the facts of the moral order are subject, like those of the physical order, to invariable laws (Guerry, 1833, p14)*

Social context of crime

What to do about crime?

- Crime a serious concern: Explosive growth in Paris, widespread unemployment, emergence of the “dangerous classes.”
- Liberal (*philanthrope*) view: increase education, better prison conditions, religious instruction, better diet (bread *and* soup!)
- Conservative view: build more prisons, harsher treatment for recidivists!

Guerry’s results were startling

- Crimes against persons and crimes against property showed different distributions over departments of France
- Crimes against persons unrelated to literacy
- Crimes against property increased with literacy!

Guerry’s Life

Sources

- Primary: Larousse, *Grand Dictionnaire*, 1866; necrology and notices by A. Maury, H. Diard, E. Vinet, 1867
- Secondary: Whitt, translation of Guerry (1833), Beirne (1993, Ch 4), *The Social Cartography of Crime*, brief mentions, often in relation to Quetelet by criminologists (Radzinowicz), sociologists (Lazarsfeld), historians (Porter, Hacking), ...

Basic facts

- Born: Tours, 25 Dec 1802; Died: Paris, 9 Apr 1866
- Father: Michel Guerry, enterpreneur
- Studied law, literature and physiology at Univ. Poitiers
- Admitted to the bar in Paris, becomes *Advocat Royale*
- 1827: Assigned to work with the crime data collected by the Ministry of Justice
- 1830: Appointed Director of Criminal Statistics in Ministry of Justice

Guerry’s Life

Accomplishments

- Three major works on moral statistics:
  - 1829: *Statistique comparée de l’état de l’instruction...*;
  - 1833: *Essai sur la statistique morale...*;
  - 1864: *Statistique morale de l’Angleterre comparée...*
  - 1833 & 1864 awarded the Moynton Prize for work in statistics by the Academie des Sciences
- Invented the *ordonateur statistique*, to facilitate calculation and tabulation of these data (no details survive). [“*Ordonateur*” adopted by IBM France to avoid the franglais “*computeur*.”]
- First analysis of the motives for suicide (content analysis of *all* suicide notes found by police in Paris) [Adopted, with little credit, by Durkheim]
- Developed methodological ideas for defining indicators of moral variables
  - How to measure crime: # of accused? convictions?
  - National standards for statistics on literacy, education?
  - How to determine relations between variables?

Guerry's data

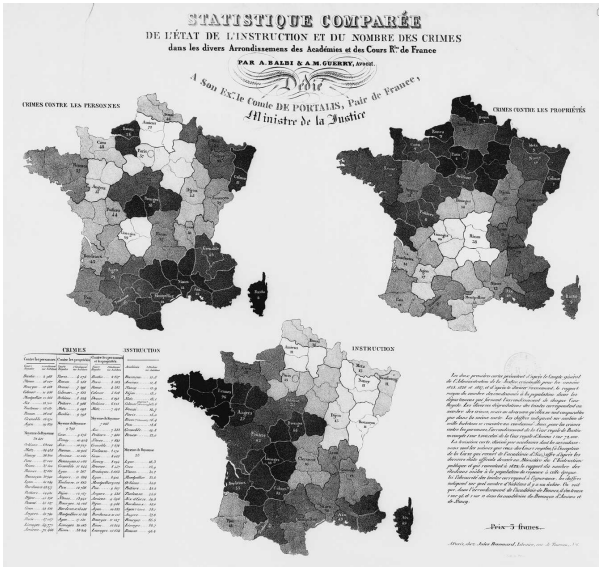
- **Compte général** de l'administration de la justice criminelle en France
  - The first national compilation of official justice data (1825)
    - detailed data on all charges and disposition
    - collected quarterly in all 86 departments.
  - Other sources: Bureau de Longitudes (illegitimate births); Parent-Duchâtelet (prostitutes in Paris); Compte du ministere du guerre (military desertions); ...
- **Moral variables**: Scaled so 'more' is 'better'

`Crime_pers`	Population per Crime against persons
`Crime_prop`	Population per Crime against property
`Donations`	Donations to the poor
`Infants`	Population per illegitimate birth
`Literacy`	Percent who can read & write
`Suicides`	Population per suicide

- **Other variables**: Ranks by department: wealth, commerce, ...

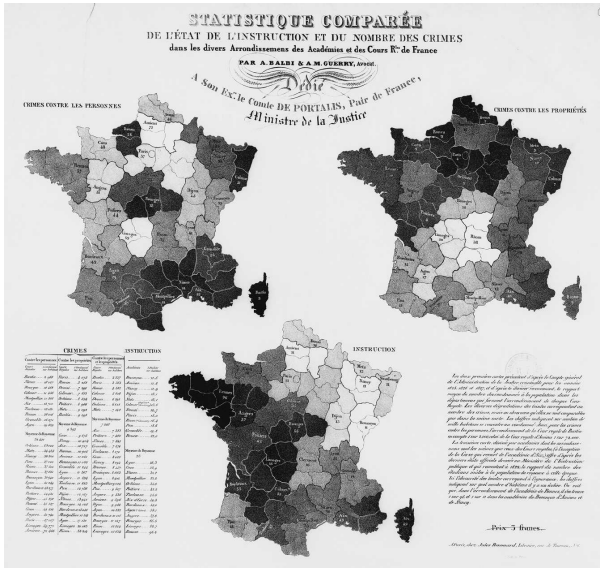
1829: *Statistique comparée de l'état de l'instruction...*

- Done with Adriano Balbi
- Single-sheet set of three shaded maps (darker = worse)
- Crime against persons, property (pop per crime)
- Instruction (# male school children)



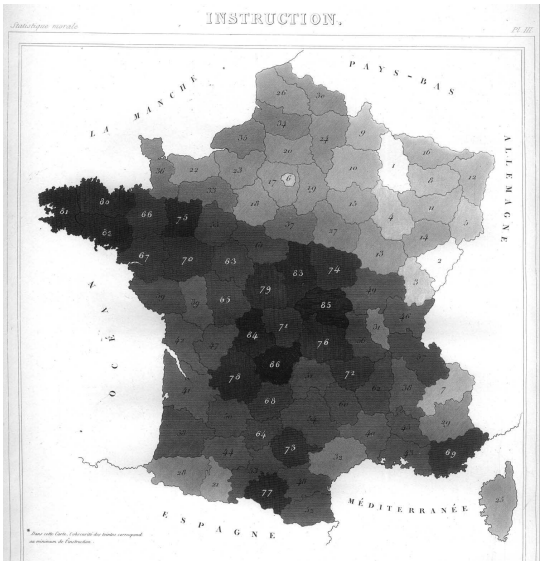
1829: *Statistique comparée de l'état de l'instruction...*

- First shaded thematic maps of crime data
- First comparative maps of social data
- $\mapsto$ crime against persons seemed inversely related to crime against property!
- Instruction: $\mapsto$ *France obscure* and *France éclairée* (Dupin, 1826)
- North of France highest in education, but also in property crime!



1833: *Essai sur la statistique morale de la France*

- Divided the 86 departments into 5 regions
- Supplemented data from the *Compte général* with:
  - Suicides in Paris, 1794–1832
  - Prostitutes in Paris (Parent-Duchâtelet)
  - Wealth (taxes per inhabitant)
  - Distribution of clergy
  - ...
- First study to use crime data to 'test' hypotheses
- Attracted widespread interest in Europe



Guerry's 1833 map of literacy in France

1833: *Essai sur la statistique morale de la France*

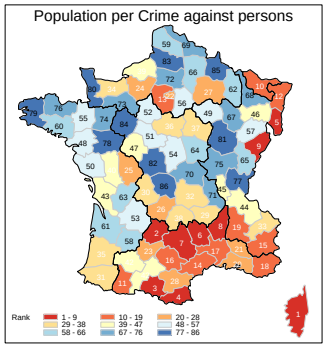
First comprehensive, national data on moral variables

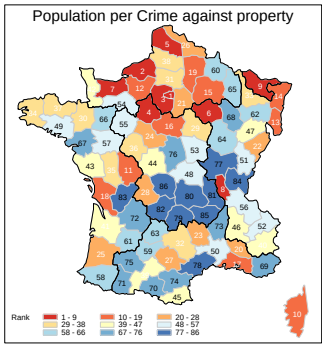
dept	Reg	Department	Crime_pers	Crime_prop	Liter	Donation	Infants	Suicides	Wealth	Commer	Clergy	Infanticid	Lottery	Desertio	Prostitut
1	E	Ain	28870	15890	37	5098	33120	35039	73	58	11	60	41	55	13
2	N	Aisne	26226	5521	51	8901	14572	12831	22	10	82	82	38	82	327
3	C	Allier	26747	7925	13	10973	17044	114121	61	66	68	42	66	16	34
4	E	Basses-Alpes	12935	7289	46	2733	23018	14238	76	49	5	12	80	32	2
5	E	Hautes-Alpes	17488	8174	69	6962	23076	16171	83	65	10	23	79	35	1
7	S	Ardeche	9474	10263	27	3188	42117	52547	84	1	28	47	70	19	1
8	N	Ardennes	35203	8847	67	6400	16106	26198	33	4	50	85	31	62	83
9	S	Ariege	6173	9597	18	3542	22916	123625	72	60	39	28	75	22	3
10	E	Aube	19602	4086	59	3608	18642	10989	14	3	42	54	28	86	207
11	S	Aude	15647	10431	34	2582	20225	66498	17	35	15	35	50	63	1
12	S	Aveyron	8236	6731	31	3211	21981	116671	50	70	3	5	81	10	4
13	S	Bouches-du-R	13409	5291	38	2314	9325	8107	2	26	30	74	3	23	25
14	N	Calvados	17577	4500	52	27830	8983	31807	10	48	7	56	13	12	194
15	C	Cantal	18070	11645	31	4093	15335	87338	59	7	6	83	82	1	20
16	W	Charente	24964	13018	36	13602	19454	25720	86	47	79	7	60	61	8
17	W	Charente-Inferi	18712	5357	39	13254	23999	16798	18	5	86	38	35	74	27
18	C	Cher	21934	10503	13	9561	23574	19497	63	56	83	11	44	51	26

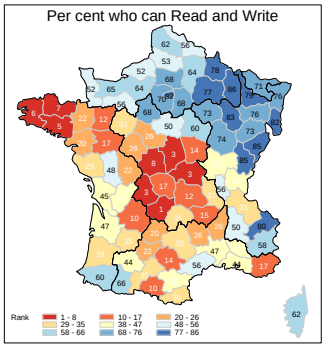
1833: *Essai sur la statistique morale de la France*

Reproduced maps (good=blue, middle=yellow, bad=red)

- Crimes against persons
- Crimes against property
- Literacy



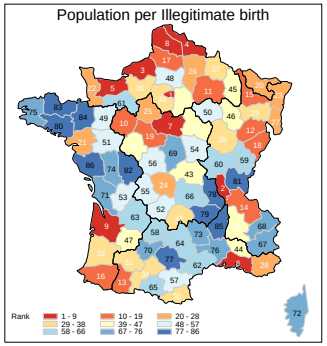


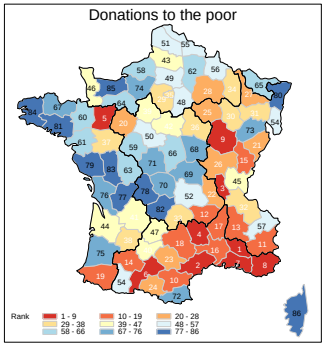


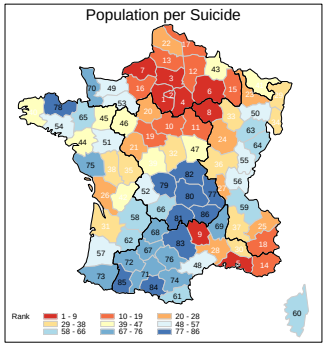
1833: *Essai sur la statistique morale de la France*

Reproduced maps (good=blue, middle=yellow, bad=red)

- Illigitimate births (*infants naturelles*)
- Donations to poor
- Suicide







1864: *Statistique morale de l'Angleterre comparée...*

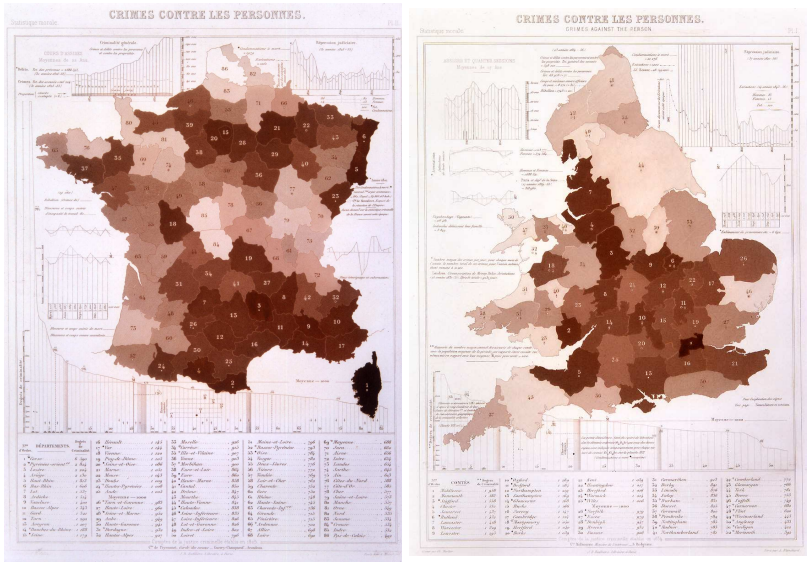
Dayenu!

- Proposes to replace simple "moral statistics" (tables) with "analytical statistics"
  - calculation, graphic display
  - ⟼ general, abstract results
- 17 large color plates (56 × 39 cm):
  - data for France (1825–1855), England (1834–1855)
  - crimes against persons and property decomposed in various ways
  - first attempt to delineate multivariate relations among moral variables
- Voluminous data:
  - 85,564 suicide records (1836–1860), classified by motive
  - 226,224 accused of personal crime
  - numbers, in a line → 1170 meters!


STATISTIQUE MORALE
DE
L'ANGLETERRE
COMPARÉE AVEC
LA STATISTIQUE MORALE
DE
LA FRANCE
PAR A.-M. GUERRY
ATLAS
PROSPECTUS
PARIS
J.-B. BAILLIÈRE ET FILS,
LONDRES, NEW-YORK,
MADRID,

1864: *Statistique morale de l'Angleterre comparée...*

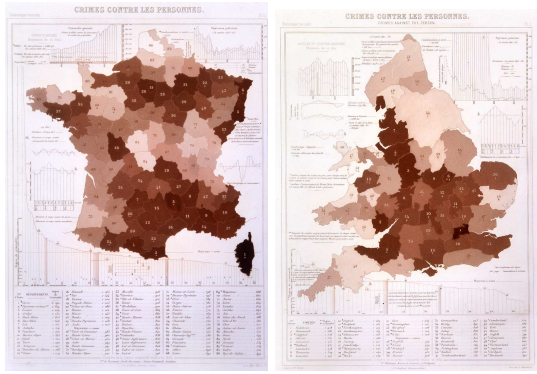
Comparing France and England



Comparing France and England

Maps for:

- crimes against persons,
- crimes against property,
- murder,
- rape,
- larceny by servants (*vol domestique*),
- arson,
- instruction,
- suicide (only for France)



1864: *Statistique morale de l'Angleterre comparée...*

Statistique analytique

- Special symbols & annotations used to mark noteworthy patterns, circumstances (↑, ↓ shows increase/decrease)

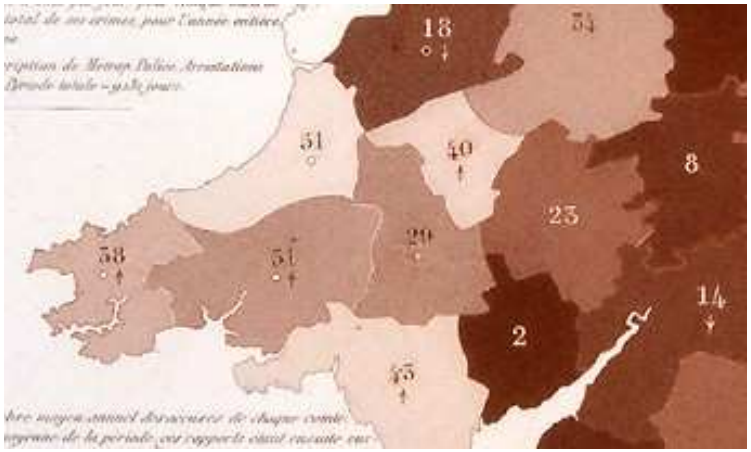


Figure: Detail from Plate II: Increase and Decrease in Crime

1864: *Statistique morale de l'Angleterre comparée...*

Statistique analytique

- Surrounding line graphs designed to decompose overall facts or relate to other facts

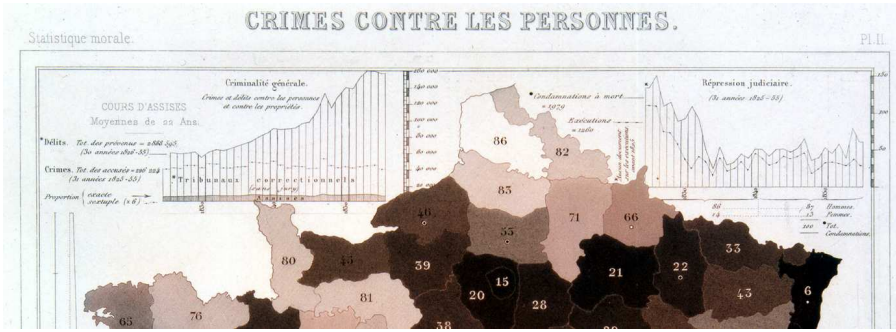
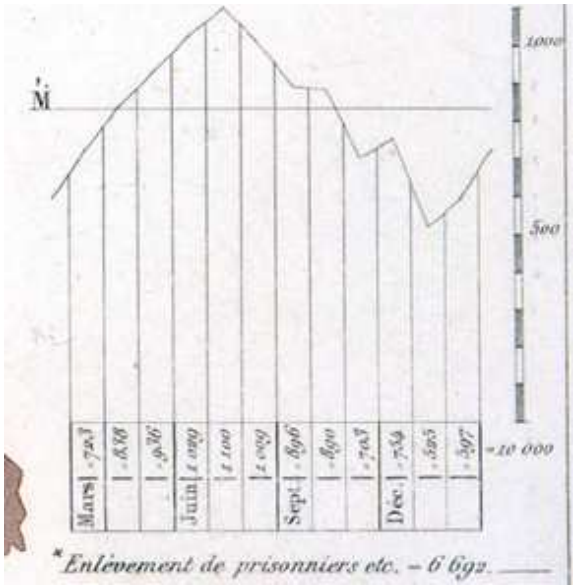


Figure: Detail from Plate I: Time series of (L) Crimes, (R) Condemned to death

1864: *Statistique morale de l'Angleterre comparée...*

Statistique analytique

Crimes against persons by month
(Detail from Plate II)

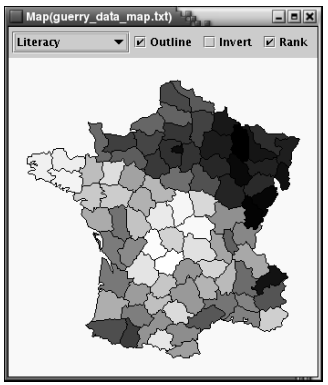


Methodological Questions

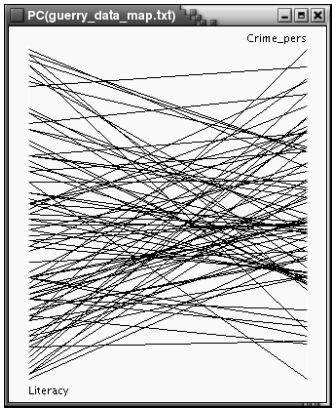
- How to define valid and reliable social indicators?
  - Education
    - Reported levels of instruction (# male children in primary school) were suspect.
    - More uniform data from Ministry of War: exams for new recruits ↦ % who could read and write.
  - Crimes
    - Number of convictions (*condamnés*) subject to factors that affect juries (severity of punishment, place where accused is judged)
    - Better to use number of indictments (*accusés*): Indictment doesn't necessarily ↦ guilt, but it reasonably ↦ a crime was committed.
- Migration and bias: Can one attribute crimes in a department to inhabitants (vs. strangers)?
  - Data from 1828+ ↦ 72% of accused either born or lived in each department
  - Only 3% committed by foreigners
- Interpretive issues:
  - Shows some awareness of issues related to ecological fallacy and spurious correlation: e.g., literacy of *actual* prisoners

Graphical Comparisons

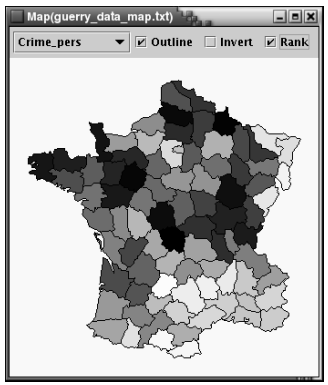
- Guerry worked before ideas of correlation, regression, scatterplots
- ↦ Used direct comparison of pairs of maps or ranked lists



(a) Literacy



(b) Ranked lists



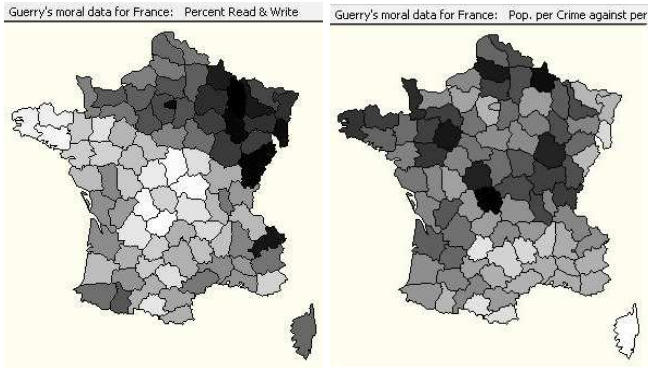
(c) Crimes against persons

Figure: Comparison of crimes against persons with literacy (% who can read and write)

Graphical Comparisons

↦ No evidence for relation between crime and literacy

- Corsica highest on crime, ∼ middle on literacy
- Literacy lowest in west and central France, but crime varies considerably
- *"Clearly the relationship people talk about does not exist"* (Guerry, 1833, p. 90)

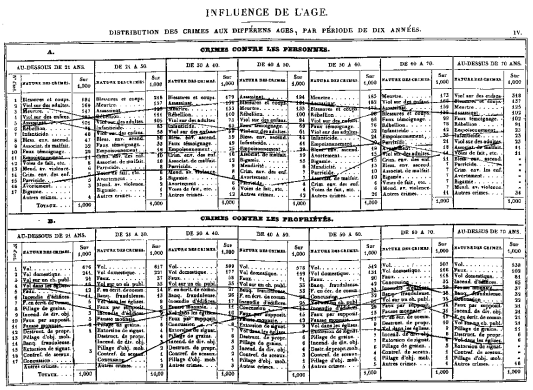


↦ similar analyses for other variables (suicides, illigitimate births, ...)

1833: Semi-graphic tables

How does type of crime vary with age?

- ↦ Used ranked tables of crime/1000 connected by colored lines
- First instance of modern parallel coordinates plot



INFLUENCE DE L'AGE.

DISTRIBUTION DES CRIMES AUX DIFFÉRENS AGES, PAR PÉRIODE DE DIX ANNÉES. IV.

A. CRIMES CONTRE LES PERSONNES.

AU-DESSOUS DE 21 ANS.		DE 21 A 30.		DE 30 A 40.		DE 40 A 50.		DE 50 A 60.		DE 60 A 70.		AU-DESSUS DE 70 ANS.	
	NATURE DES CRIMES.	Sur 1,000	NATURE DES CRIMES.	Sur 1,000	NATURE DES CRIMES.	Sur 1,000	NATURE DES CRIMES.	Sur 1,000	NATURE DES CRIMES.	Sur 1,000	NATURE DES CRIMES.	Sur 1,000	NATURE DES CRIMES.	Sur 1,000
1	Blessures et coups.	184	Blessures et coups.	216	Blessures et coups.	179	Assassinat.	194	Meurtre.	185	Meurtre.	173	Viol sur des enfans.	318
2	Viol sur des adultes.	168	Meurtre.	157	Assassinat.	154	Blessures et coups.	191	Assassinat.	182	Viol sur des enfans.	166	Blessures et coups.	157
3	Meurtre.	147	Assassinat.	150	Meurtre.	152	Meurtre.	153	Blessures et coups.	132	Assassinat.	159	Meurtre.	125
4	Viol sur des enfans.	82	Rébellion.	111	Rébellion.	110	Rébellion.	100	Rébellion.	100	Blessures et coups.	96	Assassinat.	102
5	Assassinat.	101	Viol sur des adultes	105	Viol sur des adultes	73	Viol sur des enfans.	94	Viol sur des enfans.	94	Faux témoignage.	68	Faux témoignage.	102
6	Rébellion.	78	Infanticide.	83	Infanticide.	65	Faux témoignage.	69	Faux témoignage.	76	Rébellion.	78	Rébellion.	91
7	Infanticide.	46	Viol sur des enfans.	58	Viol sur des enfans.	59	Violence des adultes.	61	Viol sur des adultes.	72	Infanticide.	43	Empoisonnement.	23
8	Bless. env. accend.	37	Bless. env. accend.	59	Bless. env. accend.	58	Infanticide.	44	Empoisonnement.	24	Empoisonnement.	35	Infanticide.	25
9	Associat. de malfait.	32	Faux témoignage.	49	Infanticide.	44	Empoisonnement.	29	Parricide.	20	Parricide.	21	Viol sur des adultes	23
10	Faux témoignage.	29	Empoisonnement.	25	Empoisonnement.	29	Bless. env. accend.	19	Avortement.	15	Viol sur des adultes.	14	Assassinat.	15
11	Empoisonnement.	11	Crim. env. des enf.	10	Assassinat.	18	Assassinat.	19	Crim. env. des enf.	11	Bless. env. accend.	11	Bless. env. accend.	11
12	Vols de fait, etc.	8	Associat. de malfait.	12	Bigamie.	15	Bigamie.	15	Bless. env. accend.	7	Vols de fait, etc.	11	Vols de fait, etc.	11
13	Mend. av. violence.	6	Parricide.	9	Mendicité.	8	Mendicité.	8	Associat. de malfait.	7	Crim. env. les enf.	7	Crim. env. les enf.	"
14	Crim. env. des enf.	5	Vols de fait, etc.	8	Crim. env. des enf.	8	Parricide.	7	Bigamie.	10	Bigamie.	7	Avortement.	"
15	Parricide.	4	Avortement.	5	Mend. av. violence.	7	Avortement.	7	Mendicité.	10	Mendicité.	7	Bigamie.	"
16	Avortement.	3	Mend. av. violence.	5	Avortement.	6	Parricide.	4	Crim. env. des enf.	6	Avortement.	7	Parricide.	"
17	Bigamie.	1	Bigamie.	4	Vols de fait, etc.	6	Vols de fait, etc.	4	Vols de fait, etc.	6	Bigamie.	7	Mend. av. violence.	"
	Autres crimes.	4	Autres crimes.	7	Autres crimes.	12	Autres crimes.	12	Autres crimes.	26	Mend. av. violence.	6	Autres crimes.	34
	TOTAUX.	1,000		1,000		1,000		1,000		1,000	Autres crimes.	11		1,000
												1,000		

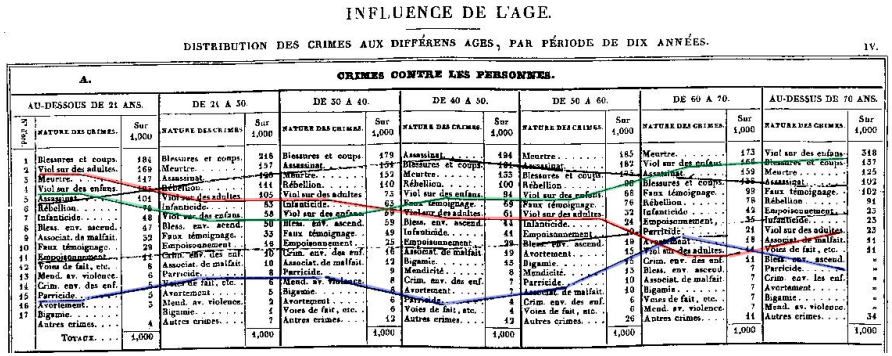
B. CRIMES CONTRE LES PROPRIÉTÉS.

Figure: Relative ranking of crimes at different ages

1833: Semi-graphic tables

Crimes against persons

- Indecent assault on adults (*viol sur des adultes*) decreases with age
- Indecent assault on children increases with age (top for 70+)
- Paricide rises to max at age 60–70



INFLUENCE DE L'AGE.

DISTRIBUTION DES CRIMES AUX DIFFÉRENS AGES, PAR PÉRIODE DE DIX ANNÉES. IV.

A. CRIMES CONTRE LES PERSONNES.

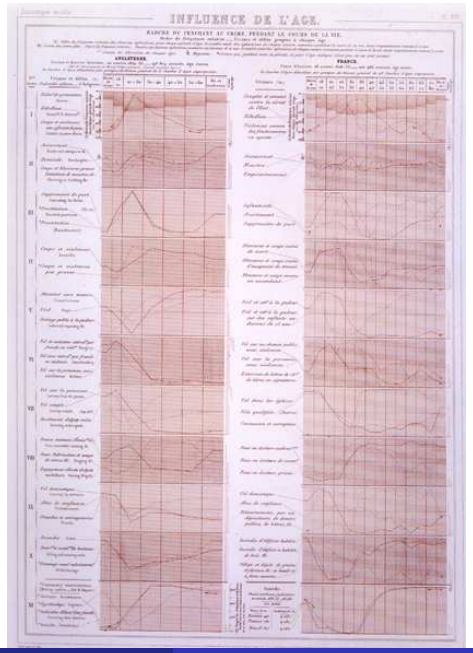
AU-DESSOUS DE 21 ANS.		DE 21 A 30.		DE 30 A 40.		DE 40 A 50.		DE 50 A 60.		DE 60 A 70.		AU-DESSUS DE 70 ANS.	
	NATURE DES CRIMES.	Sur 1,000	NATURE DES CRIMES.	Sur 1,000	NATURE DES CRIMES.	Sur 1,000	NATURE DES CRIMES.	Sur 1,000	NATURE DES CRIMES.	Sur 1,000	NATURE DES CRIMES.	Sur 1,000	NATURE DES CRIMES.	Sur 1,000
1	Blessures et coups.	184	Blessures et coups.	216	Blessures et coups.	179	Assassinat.	194	Meurtre.	185	Meurtre.	173	Viol sur des enfans.	318
2	Viol sur des adultes.	168	Meurtre.	157	Assassinat.	154	Blessures et coups.	191	Assassinat.	182	Viol sur des enfans.	166	Blessures et coups.	157
3	Meurtre.	147	Assassinat.	150	Meurtre.	152	Meurtre.	153	Blessures et coups.	132	Assassinat.	159	Meurtre.	125
4	Viol sur des enfans.	82	Rébellion.	111	Rébellion.	110	Rébellion.	100	Rébellion.	100	Blessures et coups.	96	Assassinat.	102
5	Assassinat.	101	Viol sur des adultes	105	Viol sur des adultes	73	Viol sur des enfans.	94	Viol sur des enfans.	94	Faux témoignage.	68	Faux témoignage.	102
6	Rébellion.	78	Infanticide.	83	Infanticide.	65	Faux témoignage.	69	Faux témoignage.	76	Rébellion.	78	Rébellion.	91
7	Infanticide.	46	Viol sur des enfans.	58	Viol sur des enfans.	59	Violence des adultes.	61	Viol sur des adultes.	72	Infanticide.	43	Empoisonnement.	23
8	Bless. env. accend.	37	Bless. env. accend.	59	Bless. env. accend.	58	Infanticide.	44	Empoisonnement.	24	Empoisonnement.	35	Infanticide.	25
9	Associat. de malfait.	32	Faux témoignage.	49	Infanticide.	44	Empoisonnement.	29	Parricide.	20	Parricide.	21	Viol sur des adultes	23
10	Faux témoignage.	29	Empoisonnement.	25	Empoisonnement.	29	Bless. env. accend.	19	Avortement.	15	Viol sur des adultes.	14	Assassinat.	15
11	Empoisonnement.	11	Crim. env. des enf.	10	Assassinat.	18	Assassinat.	19	Crim. env. des enf.	11	Bless. env. accend.	11	Bless. env. accend.	11
12	Vols de fait, etc.	8	Associat. de malfait.	12	Bigamie.	15	Bigamie.	15	Bless. env. accend.	7	Vols de fait, etc.	11	Vols de fait, etc.	11
13	Mend. av. violence.	6	Parricide.	9	Mendicité.	8	Mendicité.	8	Associat. de malfait.	7	Crim. env. les enf.	7	Crim. env. les enf.	"
14	Crim. env. des enf.	5	Vols de fait, etc.	8	Crim. env. des enf.	8	Parricide.	7	Bigamie.	10	Bigamie.	7	Avortement.	"
15	Parricide.	4	Avortement.	5	Mend. av. violence.	7	Avortement.	7	Mendicité.	10	Mendicité.	7	Bigamie.	"
16	Avortement.	3	Mend. av. violence.	5	Avortement.	6	Parricide.	4	Crim. env. des enf.	6	Avortement.	7	Parricide.	"
17	Bigamie.	1	Bigamie.	4	Vols de fait, etc.	6	Vols de fait, etc.	4	Vols de fait, etc.	6	Bigamie.	7	Mend. av. violence.	"
	Autres crimes.	4	Autres crimes.	7	Autres crimes.	12	Autres crimes.	12	Autres crimes.	26	Mend. av. violence.	6	Autres crimes.	34
	TOTAUX.	1,000		1,000		1,000		1,000		1,000	Autres crimes.	11		1,000
												1,000		

B. CRIMES CONTRE LES PROPRIÉTÉS.

Figure: Ranking of crimes against persons at different ages

1864: Statistique Analytique: Influence of Age

- Age distributions of criminals in England vs. France
- 10 categories of crime broken down by subtype
  - *Viol* → rape, indecent exposure
  - *Vol* → burglary, housebreaking, ...
  - *Assassinat* → murder, manslaughter, shooting, stabbing, ...

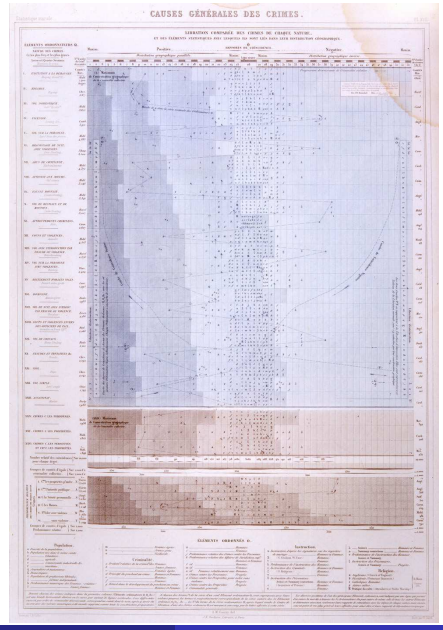


INFLUENCE DE L'AGE.

Statistique Analytique: General Causes of Crime

Plate XVII: M. Guerry's Magnum Opus

- Analysis of the factors associated with crimes and their geographic distribution
- Rows: 23 crimes, ordered by frequency and seriousness
  - keeping baudy house, bigamy, cattle stealing, ...
  - ... fraud, rape, murder
- Cols: Rank order of degree of criminality of English counties
- Entries: Symbols for associated moral aspects
  - Population (% Irish, agricultural, domestics, ...)
  - Criminality (Male, young, ...)
  - Religion (Anglicans, "dissenters", ...)
- Curves: positive and negative



CAUSES GÉNÉRALES DES CRIMES.

Guerry's Challenge

- What can we learn from reanalysis of Guerry's data?
- What could we do for Guerry as a consultant?

Statistical historiography

- Understanding through reproduction
- What was he thinking?

Statistical graphics

- How to visualize and understand relations among many variables?
- How to relate these to geographic information?

Summary

Graphical methods for multivariate data

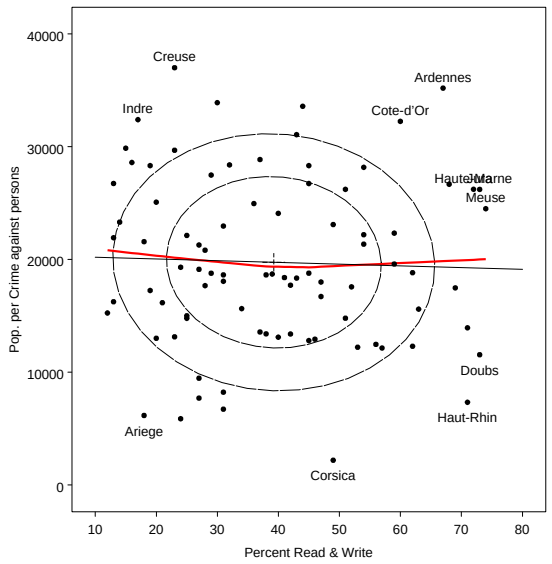
- **Bivariate displays**: Can be enhanced to show statistical relations more clearly and effectively
  - Scatterplots with data (concentration) ellipses and smoothed (loess) curves
  - Scatterplot matrices
- **Reduced-rank displays**: Multivariate visualization techniques can show the statistical data in simple ways, using dimension reduction techniques.
  - Biplots - show variables and observations in space accounting for greatest variance
  - Canonical discriminant plots - show variables and observations in space accounting for greatest between-group variation

Reduced rank

Bivariate plots: Data ellipse and smoothing

Scatterplot with 40% and 68% data ellipses, and smoothed (loess) curve

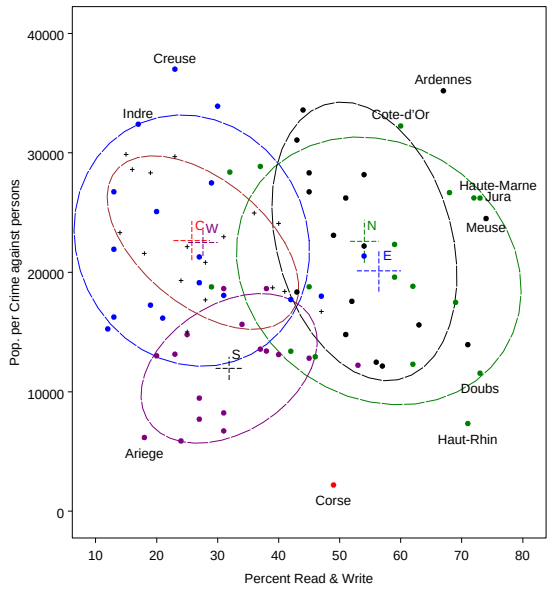
- No linear relation between crimes against persons and literacy
- No indication of non-linear relation
- Substantial number of unusual departments



Bivariate plots: Region differences

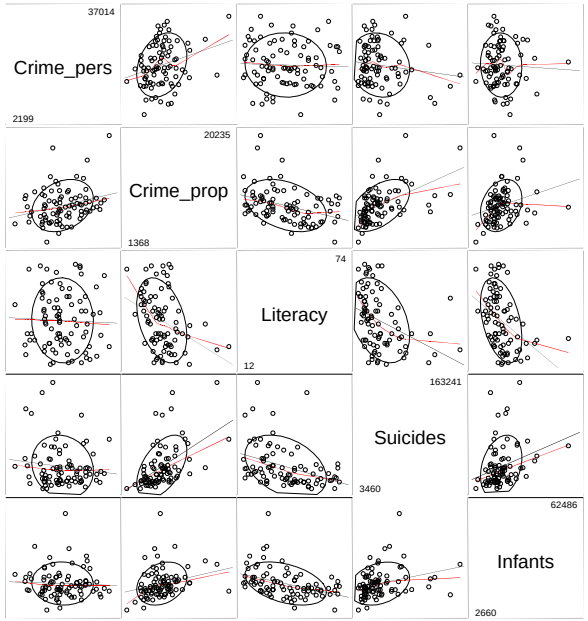
Summary ellipses for each region

- Biggest differences (C, W) vs. (N, E) on literacy
- South low on both crime and literacy



Bivariate plots: Scatterplot matrices

- Crime vs. persons and property: ↗!
- Literacy ↘ most others
- Suicide ↗ out of wedlock



Multivariate analyses: Reduced rank displays

- Multivariate visualization techniques can show the statistical data in simple ways, using dimension reduction techniques.
  - Biplots - show variables and departments in space accounting for greatest variance
  - Canonical discriminant plots - show variables and departments in space accounting for greatest between-region variation
- Show geographic location by color coding or other visual attributes.
  - Color code by region
  - Data ellipse to summarize regions
- → **Data-centric displays**: The multivariate data is shown directly; geographic relations indirectly

Biplots

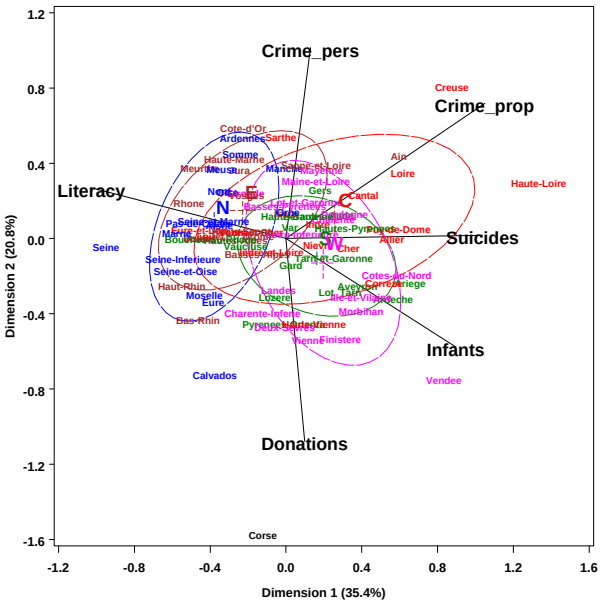
- Biplots represent both variables (attributes) and observations (departments) in the same plot— a low-rank (2D) approximation to a data matrix

$$\boldsymbol{Y}^\star \approx \boldsymbol{A}\boldsymbol{B}^\mathsf{T} = \sum_{k=1}^{d} \boldsymbol{a}_k \boldsymbol{b}_k^\mathsf{T}$$

  - Variables usually represented by vectors from origin (mean)
  - Observations usually represented by points
  - Can show clusters of observations by data ellipses
- Properties:
  - Angles between vectors show correlations ($r \approx \cos(\theta)$)
  - Length of variable vectors $\sim$ % variance accounted for
  - $y_{ij} \approx \boldsymbol{a}_i^\mathsf{T} \boldsymbol{b}_j$: projection of observation on variable vector
  - Dimensions are uncorrelated overall (but not necessarily within group)

Biplots: Guerry data

- Dim 1 (34.5%): *France obscure* vs. *France éclairée*
- Dim 2 (20.8%): Personal crime vs. Donations (benevolence?)

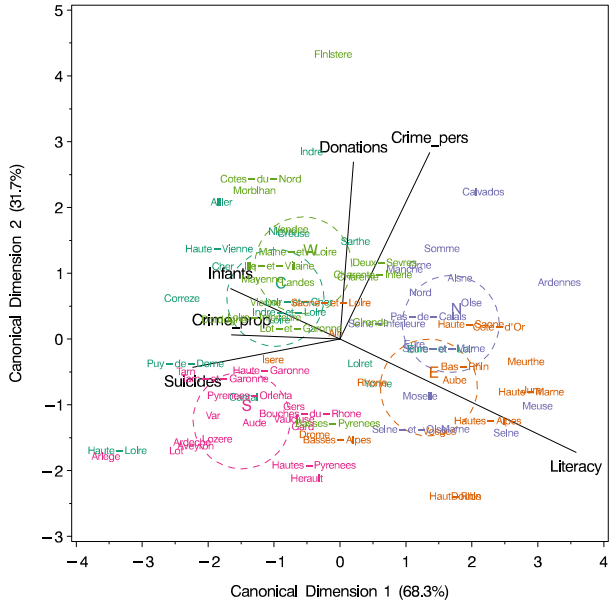


Canonical discriminant plots

- Project the variables into a low-rank (2D) space that maximally discrimates among regions
  - Visual summary of a MANOVA
  - Canonical dimensions are linear combinations of the variables with maximum univariate F-statistics.
  - Vectors from the origin (grand mean) for the observed variables show the correlations with the canonical dimensions
- Properties:
  - Canonical variates are uncorrelated
  - Circles of radius $\sqrt{\chi^2_2(1-\alpha)/n_i}$ give confidence regions for group means.
  - Variable vectors show how variables discriminate among groups
  - Lengths of variable vectors $\sim$ contribution to discrimination

Canonical discriminant plots: Guerry data, by Region

- Dim 1 (61.9%): *France obscure* vs. *France éclairée*
- Dim 2 (28.8%): Donations; (S, E) vs. (N, C, W)



Multivariate mapping: Map-centric displays

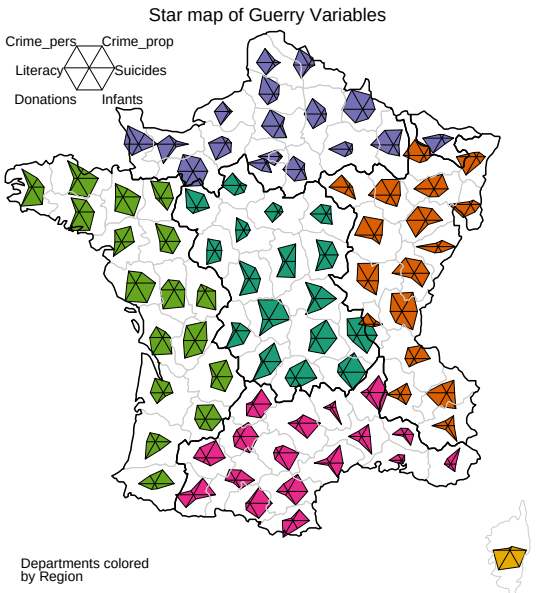
How to generalize choropleth maps to many variables?

- **Star maps**: Show multivariate data on the map using star icons, variable $\sim$ length of ray
- **Blended RGB displays**: (V1, V2, V3) $\mapsto$ (R, G, B) shading
- **Conditioned choropleth maps**: Stratify by two variables, show conditioned maps in a trellis-like display

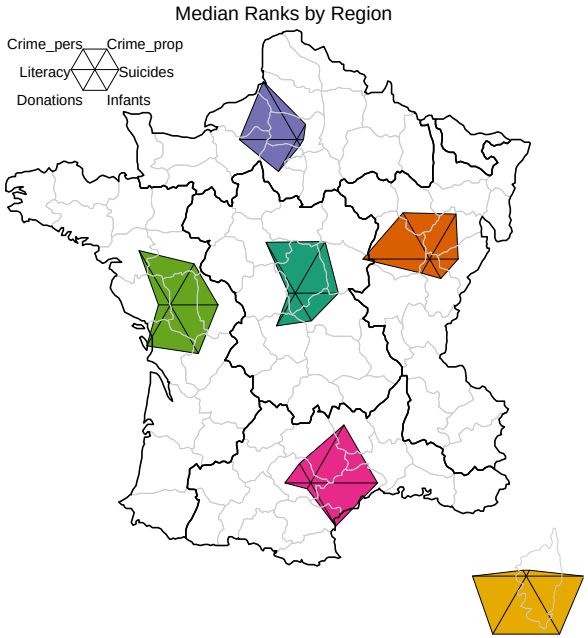
Summary

Star maps

- Bigger $\mapsto$ better
- Variables ordered as in biplot
- $\mapsto$ identify unusual depts.

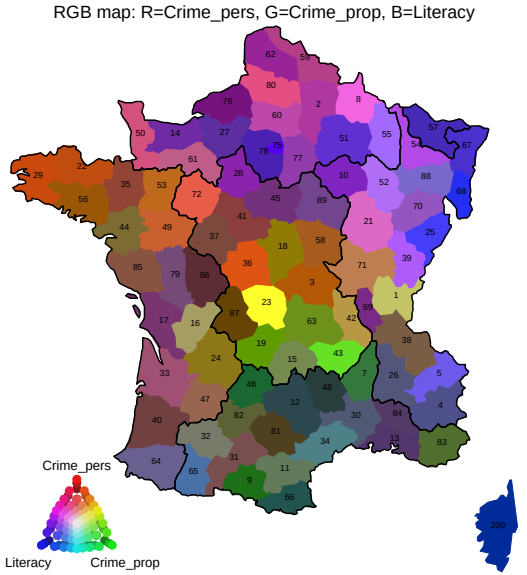


Star maps: Medians by region



RGB map

Crime against persons, Crime against property, Literacy

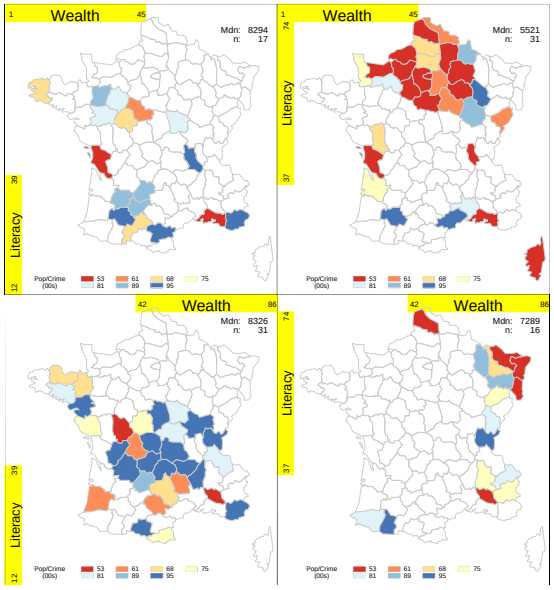


- N: Hi lit & personal crime
- Creuse (23), Ain (1): Hi pop/crime, low lit.

Conditioned choropleth maps

Crime against property | % Literacy and Wealth (rank)

- Crime: good = blue; bad = red
- Control for two background variables
- e.g., UR: High Lit, Wealth $\mapsto$ (N, High crime)
- Dynamic version: choose ranges with sliders



Summary

Guerry's place in history

- (Along with Quetelet) Guerry's work established the empirical study of moral statistics $\mapsto$ "Moral statistics movement" $\mapsto$ modern criminology, sociology, social science.
- The 1833 *Essai* broke new ground in thematic cartography and data visualization.
- The 1864 comparative study contemplated multivariate explanations beyond available theory and methods.
- Guerry's questions, methods and data still present challenges for multivariate spatial data visualization today.

References I

Balbi, A. and Guerry, A.-M. Statistique comparée de l'état de l'instruction et du nombre des crimes dans les divers arrondissements des académies et des cours royales de France. Jules Renouard, Paris, 1829. BL:Tab.597.b.(38); BNF: Ge C 9014 .

Beirne, P. *Inventing Criminology*. State University of Albany Press, Albany, NY, 1993.

Dupin, C. *Carte figurativ de l'instruction populaire de la France*. Jobard, 1826. BNF: Ge C 6588 (Funkhouser (1937, p. 300) incorrectly dates this as 1819).

Guerry, A.-M. *Essai Sur La Statistique Morale de la France*. Crochard, Paris, 1833. English translation: Hugh P. Whitt and Victor W. Reinking, Lewiston, N.Y. : Edwin Mellen Press, 2002.

Guerry, A.-M. *Statistique morale de l'Angleterre comparée avec la statistique morale de la France, d'aprés les comptes de l'administration de la justice criminelle en Angleterre et en France, etc.* J.-B. Baillière et fils, Paris, 1864. BNF: GR FOL-N-319; SG D/4330; BL: Maps 32.e.34; SBB: Fe 8586; LC: 11005911.