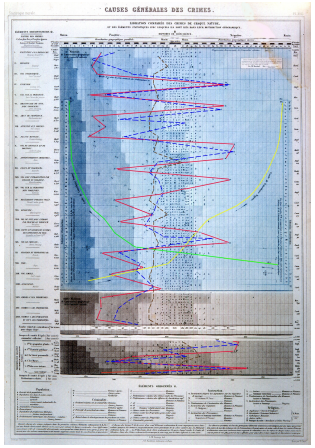
André-Michel Guerry's *Ordonnateur Statistique* The First Statistical Calculator?

Michael Friendly Nicolas de Saint Agathe

QM Brownbag Talk, York University

December 10, 2012

Outline



Guerry's 1864 Plate 17: *General Causes of Crimes*

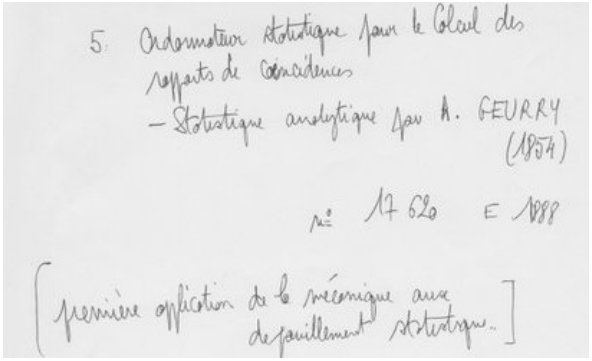
Further information:

- http://datavis.ca/papers/ordonnateur/
- http://datavis.ca/gallery/guerry/

How this all began...

- Pascale to Nicolas: It's time to clean up your study!
- Nicolas to me: email, 4/21/2012, re: guerry ordonnateur statistique

  it will be a pleasure to speak about this document



Who was Andre-Michel Guerry?

- Born: Tours, 24 Dec 1802; Died: Paris, 9 Apr 1866
- Studied law, literature and physiology at Univ. Poitiers
- First systematic analysis of comprehensive data on crime, suicide and other *moral statistics*
  - *Essai sur la statistique morale de la France* (1832)
- Along with Quetelet, the inventor of modern social science
  - the idea that data on "moral statistics" (rates of crimes, suicides, out-of-wedlock births, ...) could be used to argue for social laws, akin to physical laws.
  - My pet peeve: I have to say *along with Quetelet*, because Quetelet gets most of the credit, but Guerry was far more incisive.

The discovery of "social laws"

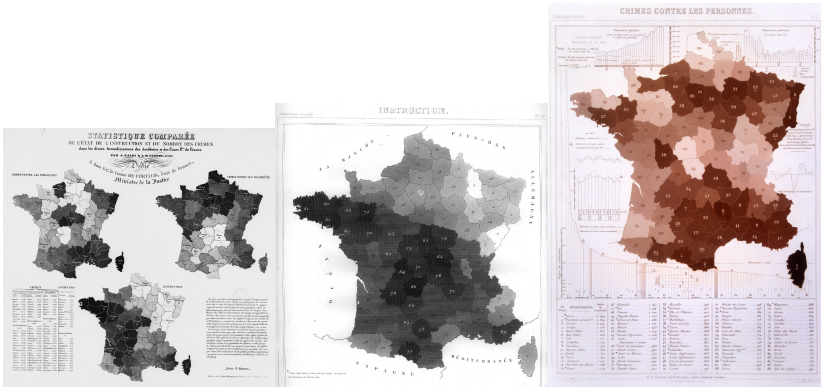
Guerry's (1832) results on rates of crime and suicide were both compelling and startling:

- Rates of crime and suicide remained remarkably invariant over time
- ... yet they varied systematically by region, sex of accused, type of crime, etc.
- In any given French department, nearly constant numbers of suicides, out-of-wedlock birth, 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

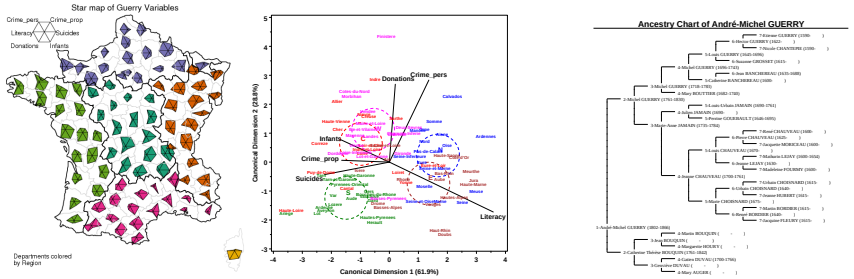
Guerry's work

- Three major works on moral statistics:
  - 1829: *Statistique comparée de l'etat de l'instruction ...*
  - 1833: *Essai sur la statistique morale de la France*
  - 1864: *Statistique morale de l'Angleterre comparée avec ... France*
  - 1833 & 1864 awarded the Moynton prize for work in statistics by the Académie des Sciences



My previous work on Guerry

- In Friendly (2007), I describe Guerry's contributions to the analysis of multivariate data in a spatial context, along with some challenges for modern methods.
- Another paper (Friendly, 2008) traces as much of his personal and family history as could be determined.



The Guerry Mysteries

There were two unanswered questions, which I called the *Guerry Mysteries*:

1. Guerry's unpublished papers and work-in-progress

- In Guerry (1864), he makes clear that this is just an introduction to more detailed studies.
- He died, somewhat suddenly in 1866, leaving behind a large collection of work-in-progress.
- I traced these works to his friend, Hippolyte Diard, and a proposed project (1867–1868) to publish these in Tours (Societe d'Agriculture, Science, Arts et Belles Lettres) — never completed.
- Extensive search in the archives in Tours: → these papers may have vanished in the upheaval of the Franco-Prussian war (July, 1870 – May, 1871).

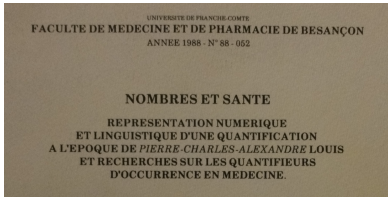
Side note: my other French hero, Charles Joseph Minard fled from Paris to Bordeaux at the same time, carrying with him some unpublished work, that also vanished. I still blame Germany!

The Guerry Mysteries (cont'd)

2. The *ordonnateur statistique*

- In Guerry (1864), he presents data on over 226,000 cases of crime in England and France over a 25-year period and over 85,000 suicide records classified by motive and other factors.
- For the first time, he presents graphic maps and charts designed to show relationships between moral variables.
- He estimated that his numbers, written down in a line, would stretch over 1170 meters!
- The nagging question: how did he do this work?

Nicolas' thesis


UNIVERSITE DE FRANCHE-COMTE
FACULTE DE MEDECINE ET DE PHARMACIE DE BESANÇON
ANNEE 1988 - N° 88 - 052

NOMBRES ET SANTE

REPRESENTATION NUMERIQUE
ET LINGUISTIQUE D'UNE QUANTIFICATION
A L'EPOQUE DE *PIERRE-CHARLES-ALEXANDRE* LOUIS
ET RECHERCHES SUR LES QUANTIFIEURS
D'OCCURRENCE EN MEDECINE.

- In 1984, Nicolas was working on his thesis on the history of the use of numbers and quantification of health and medicine in 19^{th} century France (Faculté de Médecine et de Pharmacie de Besançon).
- Visited the CNAM to inquire about Guerry's *ordonnateur statistique*
- Search by an archivist:
  - the *ordonnateur statistique* could not be found
  - however, a two-page typewritten description was discovered
- **This document provides the first, concrete information about the goal, design, and operation of Guerry's device**

The CNAM document

ORDONNATEUR STATISTIQUE
pour le calcul des rapports de concordance
par A.M. Guerry(1).

Cette machine a pour objet de classer dans un ordre déterminé des fiches qui portent des nombres à additionner par groupes. Dans les travaux d'analyse statistique, on compare des grandeurs ordonnées d'après un certain caractère afin d'apprécier l'influence de ce caractère sur les grandeurs étudiées. Par exemple on aura recours à de semblables comparaisons pour analyser les relations qui peuvent exister entre la criminalité d'un pays et la densité de la population, le degré d'instructio

L'inventeur de la machine a poursuivi pendant de longues années de laborieuses recherches sur la criminalité en Angleterre et en France. Dans une de ses notes il indique un exemple qui fera comprendre le but et le fonctionnement de l'ordonnateur mécanique.

Pour chacun des 52 Comtés anglais et pour une longue période de temps, il a relevé le nombre des incendies,et le nombre des assassinats, le nombre des habitants par kilomètre carré, etc. etc.

Ces différents nombres ont été inscrits sur des tableaux comportant une ligne par Comté. Puis on a découpé chaq tableau en bandes étroites à raison d'une bande par ligne ou par comté. Enfin,sur chaque bande,on a inscrit un numéro d'ordre suivant l'ordre de grandeur du nombre qui y figure. D'autre numéros d'ordre caractérisent l'ordre alphabétique, l'ordre de juridiction des comtés, etc.

bandes qui portent des nombres d'assassinats (ordre nouveau) dans l'ordre de grandeur des nombres d'incendies (ordre ancien): la machine opérera mécaniquement ce classement après une certaine préparation.

Elle a pour organe essentiel un cylindre portant à sa circonférence 52 rangées de 58 trous dans lesquels on peut piquer des chevilles. Supposons que le classement des comtés anglais, d'abord suivant le nombre des incendies, puis suivant le nombre des assassinats, soit conforme aux indications suivantes:

Comtés	Rang suivant le nombre des incendies	Rang suivant le nombre des assassinats
Derby	1	42
Stafford	2	27
Brecon	3	25
Norfolk	4	6
"	"	"

On fera tourner le cylindre de manière que les rangées de trous se présentent successivement. Dans la première rangée on piquera une cheville au 42e trou, une autre au 27e trou de la seconde rangée, puis au 25e trou de la troisième rangée; au 6e trou de la 4e rangée, et ainsi de suite.

Dans un cadre spécial disposé devant la machine les bandes sont rangées séparément suivant l'ordre ancien (incendies). Si alors on fait tourner le cylindre à l'aide d'une manivelle, les chevilles qui se présentent successivement sur la même ligne viennent à tour de rôle décrocher des brimballes qui chassent les bandes à classer dans l'ordre nouveau, c'est-à-dire les bandes numérotées 42, 27, 25, 6 -- dans l'ordre de classement suivant l'ordre des assassinats.

Un cadran compteur avec sonnerie permet de s'arrêter après chaque série de dix bandes. Les bandes ainsi chassées sont disposées sur une tablette, ce qui permet d'additionner commodément les nombre placés en colonne. A l'aide des totaux formés on calcule ensuite les moyennes et les coefficients utiles pour l'analyse des résultats.

L'ordonnateur de M. Guerry constitue probablement la première application de la mécanique aux dépouillements statistiques.

Our translation

ORDONNATEUR STATISTIQUE
pour le calcul des rapports de concordance
par A. M. Guerry

Anonymous*

undated

Cette machine a pour objet de classer dans ordre déterminé des fiches qui portent des nombres a additionner par groupes. Dans les travaux d'analyse statistique, on compare des grandeurs ordonnées d'après un certain caractère afin d'apprécier l'influence de ce caractère sur les grandeurs étudiées. Par exemple on aura recours a de semblables comparaisons pour analyser les relations qui peuvent exister entre la criminalité d'un pays et la densité de la population, le degré d'instruction.

L'inventeur de la machine a poursuivi pendant de longues années de laborieuses recherches sur la criminalité en Angleterre et en France. Dans une de ses notes il indique un exemple qui fera comprendre le but et le fonctionnement de l'ordonnateur mécanique.

This machine has the aim of classifying in a given order the records which list numbers to be added by groups. In the work of statistical analysis, one compares values ordered according to a certain character/trait in order to understand the influence of this character on the studied values. For example we may have similar comparisons when analyzing the relations which may exist between criminality in a country, density of the population and the level of education.

The inventor of the machine pursued for long years laborious research on criminality in England and France. In one of his notes, he indicates an example which will help us understand the goal and the operation of the mechanical organizer.

Description I

- Purpose:

  The aim of this machine is to classify in a given order the records which list numbers to be added by groups. . . . For example we may have similar comparisons when analyzing the relations which may exist between criminality in a country, density of the population and the level of education.

- Design:
  - An early spreadsheet organizer for sorting and relating variables!
  - The machine was designed to work from a table, listing English counties or French departments in rows, with social, moral and crime variables in columns (converted to ranks)
  - Consisted of a cylinder with rotatable bands
  - Each band contained holes into which one could push pegs and then rotate the bands to show the concordance between variables

Description II

- Example:

  Let us suppose that the classification of the English counties, initially according to the number of fires, then according to the number of murders, is in conformity with the following indications:

Counties	*Rank according to number*	
	Fires	*Murders*
Derby	*1*	*42*
Stafford	*2*	*27*
Brecon	*3*	*25*
Norfolk	*4*	*6*
. . .	. . .	. . .

Description III

- Operation:

  We will turn the cylinder so that the lines of holes show up successively. In the first line one will push the peg at the 42nd hole, another with the 27th hole of the second line, then at the 25th hole of the third line; at the 6th hole of the 4th line, and so on.
  In a special frame put in front of the machine the rows are arranged separately according to the old order (fires). If then we turn the cylinder using a crank, the pegs which show up successively on the same line in turn come to take down the levers that drive out the rows to be classified in the new order, i.e. the rows numbered 42, 27, 25, 6 — in the order of classification following the order of the murders.

Summary of this description

- The primary function of this machine was for sorting one target variable in relation to another focal variable, as one might do today in a spreadsheet.
- Having done this, one could then manually calculate summaries (averages) for the ranks of the target variable, or others carried along by sorting.

The author of this document concludes:

The organizer of Mr. Guerry probably constitutes the first application of mechanics to statistical evaluations.

Another minor mystery:

- Who wrote this document? When?
- Repeated requests to the CNAM has turned up no information.

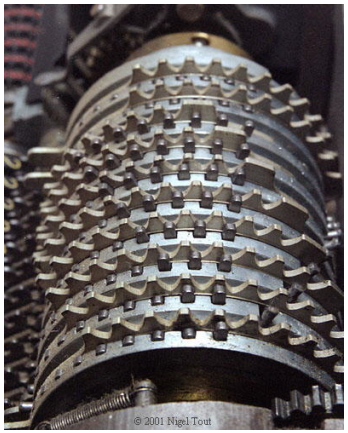
But, what did it look like?

Nicolas suggested this:

But, what did it look like?

Or, maybe something like these:

Curta calculator



Odhner pinwheel calculator

How did Guerry use it?

Goal of use:

- The ordonnateur was designed as a special purpose device to summarize the relation between crimes and other possible explanatory factors
- One variable was arranged by ranks; turning the crank arranged another variable in correspondence
- The goal was thus to assess the relation between two rank-ordered series.

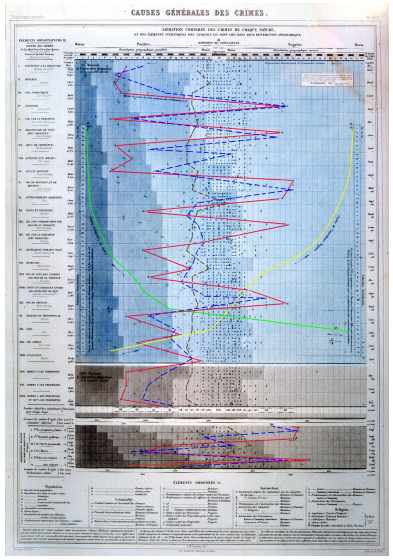
Why not a scatterplot?

- The scatterplot was only invented in 1833, by Herschel and had not entered common practice.
- The idea of correlation would not be invented until after 1886 (Galton, Pearson)

How did Guerry use it?

Guerry's magnum opus: General causes of crime (Plate 17)

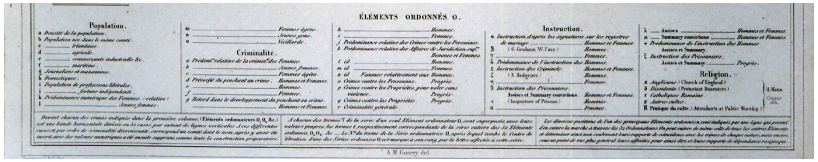
- Analysis of factors associated with crime
- Rows: 23 crimes, ordered by seriousness:
- Cols: rank order of criminality in English counties
- Entries: symbols for moral/pop^n factors
  - % Irish, agricultural, domestics, ...
  - % Male, young, ...
  - Religion (Anglicans, "dissenters," ...
- Lines: Centres de libration



How did Guerry use it?

The legend at the bottom lists the following kinds of symbols, shown in the graphic table in relation to various crimes:

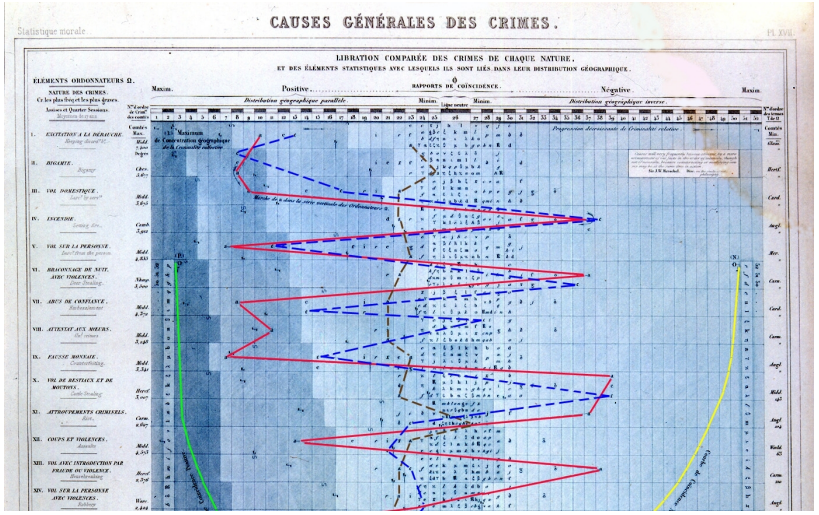
- Aspects of population: population density, % Irish, agricultural, domestics, ...
- Aspects of criminality: predominance of male, female, young, old, relative to average, ...
- Degree of instruction of males, criminals, prisoners
- Religion: Anglicans, Catholic, ...



Guerry (1864) Plate 17

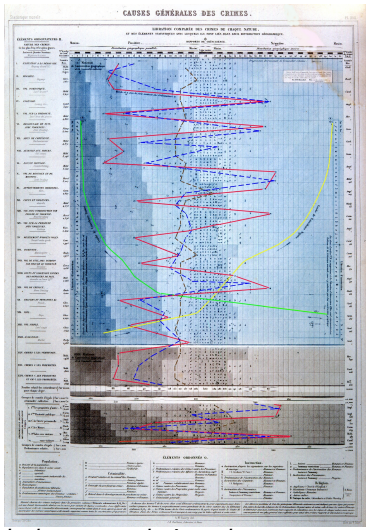
How did Guerry use it?

Lines and curves in the chart trace the center of libration (points of equilibrium ∼ "average") of the various social indicators



How did Guerry use it?

- Red lines (symbol **a**): population density
  - positively related: bigamy, domestic violence
  - negatively related: arson, cattle theft
- Blue lines (**c**): % Catholic
  - Similar to population density
  - Associated with the same crimes
- Brown lines: % Anglican
  - Does not vary much across crimes



Inset quote: "*Causes will very frequently become obvious by a mere arrangement of facts in the order of intensity...*" (Herschel, 1831)

When did he invent it?

- Nicolas' note suggests 1854 as the date, and this corresponds to the CNAM catalog entry
- There is now suggestive evidence that the actual date was before 1851, when Guerry presented his work in England
- Fletcher's 1851 report to the Statistics section of the BAAS:

  M. Guerry exhibited eighteen coloured Maps illustrating some important conclusions respecting the Criminal Statistics of England for 16 years, ending 1850. ...

  Besides the maps, he showed a series of tables, exhibiting by curved lines for each county, the degrees of positive and negative criminality corresponding with the coloured maps.

- Thus, it is clear that Guerry produced a set of charts similar to his *General Causes of Crime* Plate 17 by 1851.
- Most likely, these were done using his ordonnateur statistique.

So, what happened to it?

- Guerry's *ordonnateur statistique* was listed in the CNAM catalog of 1906
- As late as 1941, it was still registered in their collection.
- It could not be found (except for the document) in 1984 when Nicolas asked.
- In 1991, the CNAM underwent a large renovation; many items were moved to their warehouse in the north of Paris (St. Denis)
- They now say either it was lost or stolen, or may yet surface when they complete the inventory and cataloging of the collection in 3–4 years.

Guerry's Perfect Storm

Guerry's work (1829–1864) occurred at a time when three main areas coalesced:

- **data**: the explosive growth of systematic, centralized records by state agencies
- **visualization**: development of shaded maps (Dupin, 1826) to *see* geographic distributions and compare variables
- **technology**: machines for mechanical calculation

For the first time, scientific methods could be applied to important social, medical and legal issues

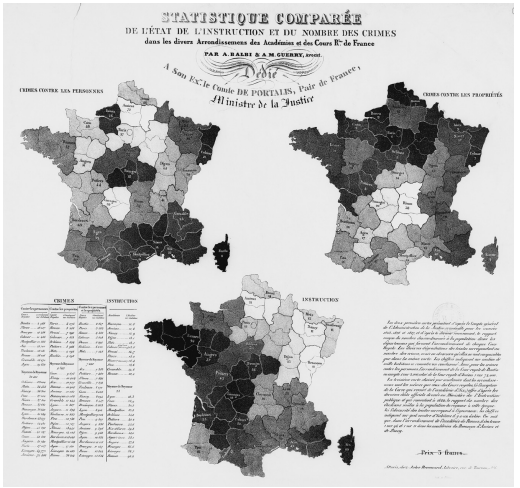
An avalanche of data

- In France, the first national compilation of official justice data began in 1825 (*Compte général de l'adiministration de la justice criminelle*)
  - detailed data on all charges and their disposition (gender, date, guilty? length of sentence, ...)
  - collected quarterly in all 86 departments
- Other systematic data was collected by other state agencies and individuals
  - Bureau de Longitudes: illigitimate births
  - Ministère du guerre: desertions, literacy of recruits, ...
  - Parent-Duchatâlet: prostitutes in Paris
  - Guerry: Suicides in Paris (every suicide note analyzed for motive, etc.)
- Data became the new tool for social policy and planning!
- Similar efforts began in other European countries

Visualization: Comparative maps to explore relations

Balbi & Guerry (1829): Statistique comparée de l'état de l'instruction ...

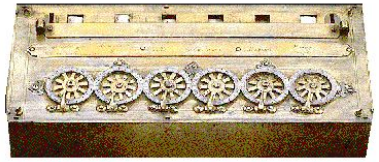
- First shaded thematic maps of crime data
- First comparative maps of social data
- $\mapsto$ crime against persons seemed inversely related to crime against property!
- Instruction: $\mapsto$ *France obscure* and *France éclairée* (Dupin, 1826)
- North of France highest in education, but also in property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



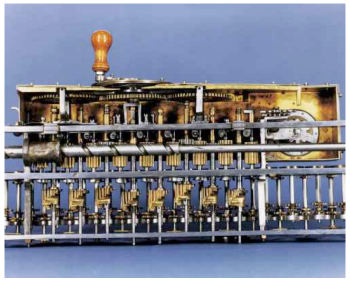
Technology: Early modern calculating devices

The final component was the emergence of machines for mechanical calculation to aid in processing large volumes of raw data.
Earliest modern developments:

- 1642: The Pascaline by Blaise Pascal— geared dials for performing addition and subtraction
- 1672: Leibnitz' Stepped Reckoner— a stepped cylindrical drum and gears; could perform multiplication and division by successive addition and subtraction.



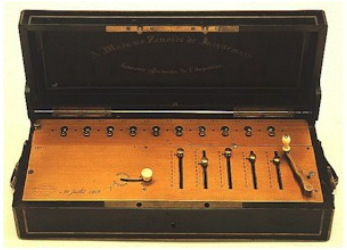
Pascaline



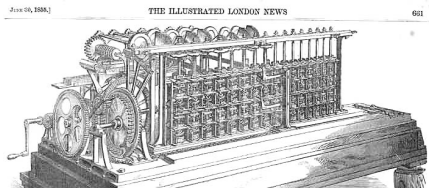
Leibnitz Stepped Reckoner

Technology: Calculating devices in Guerry's time

- 1820: Four-function calculators become popular, beginning with the Arithmometer (Thomas de Colmar)
- 1834: Charles Babbage develops the initial prototype of the difference engine— could be mechanically programmed programmed to calculate approximations to functions (log, trig) by mechanical analog of divided differences
- 1853: Georges & Edvard Scheutz construct the first working difference engine.



Arithmometer



Scheutz calculator

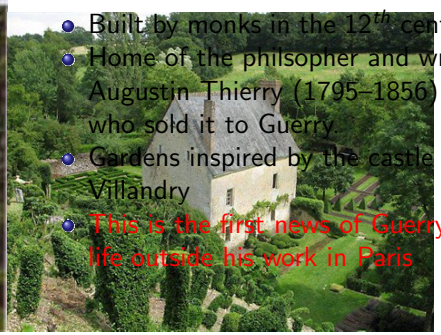
A Babbage connection?

- Guerry corresponded extensively with English colleagues— notably William Farr and Charles Babbage
- We recently obtained Babbage's correspondence on microfilm from the British Library (*Papers of Charles Babbage 1791–1871*)
- We found 9 letters from Guerry, dated 22 Feb 1856 – 6 Mar 1861; only a few are legible, and 3 relate to Guerry's works.
- In several letters from 1860 (#103, 7 Jul; #101, 7 Aug.), Guerry says he has sent Babbage color proof copies of some of the plates for his *Statistique morale de l'Angleterre comparée ...*
- We have not yet found any mention of the *ordonnateur statistique*
- However, we now have Guerry's signature

A country home where he served as mayor

Another email: Michel Moser, 1 Sep. 2012:

J'habite á Beaumont sur dême en France, et je voulais vous dire que André Michel Guerry á été maire de ce village de 1846 a 1855.



- Built by monks in the 12th century
- Home of the philsopher and writer Augustin Thierry (1795–1856) who sold it to Guerry.
- Gardens inspired by the castle of Villandry
- This is the first news of Guerry's life outside his work in Paris

Where Guerry lived in Paris

- Guerry's house in Beaumont was sold by his heirs, Charles & René Poisson in the year after his death
- The present owner supplied the bill of sale, which lists Guerry as living at 123 Boulevard St. Michel in Paris
- My other French hero, Charles Joseph Minard lived nearby (23 Rue du Bac), so they were nearly neighbors!

Conclusions

- Guerry's invention of the *ordonnateur statistique* took place in a time when the necessary components for modern social science were first available
- It served his program of statistique analytique— going beyond mere description, to reason about relationships among moral variables and potential causes.
- It was surely a special-purpose device for organizing statistical tables, and assessing the relationships among variables.
- We can find no earlier example of a mechanical device for aiding statistical calculation.
- It is all the more remarkable because even the idea of correlation would not be invented for another 30 years.

Further information:

- http://datavis.ca/papers/ordonnateur/
- http://datavis.ca/gallery/guerry/

References

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

Friendly, M. (2007). A.-M. Guerry's Moral Statistics of France: Challenges for multivariable spatial analysis. *Statistical Science*, 22(3), 368–399.

Friendly, M. (2008). La vie et l'oeuvre d'André-Michael Guerry (1802-1866). *Mémoires de l'Académie de Touraine*, 20. Read: Feb. 8, 2008, Académie de Touraine.

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