

QUANTIFICATION IN HISTORY

Francisco J. Beltrán Tapia, Norwegian University of Science & Technology



WHY QUANTIFYING?

What are the advantages and disadvantages of using quantitative data to study history?

QUANTIFICATION IN HISTORY

- ❑ Historians often loosely ‘quantify’

Fogel (1975) famously suggested to choose a historical text at random, open it to any page, and count up the quantitative statements found there

“One of the most important developments in twentieth-century society has been the growth of leisure and recreation. Already by the Edwardian era, many pastimes and pursuits had been fashioned or transformed to meet the needs of a primarily urban and industrial society. In sport, entertainment and private recreations, one of the major driving forces was commercialization, drawing upon the increased spending power of a mass consumer market. Another was the increased leisure time available as a result of shorter working hours, paid holidays, longer life expectancy after retirement, smaller families and, for some, enforced idleness through unemployment. [...]”.

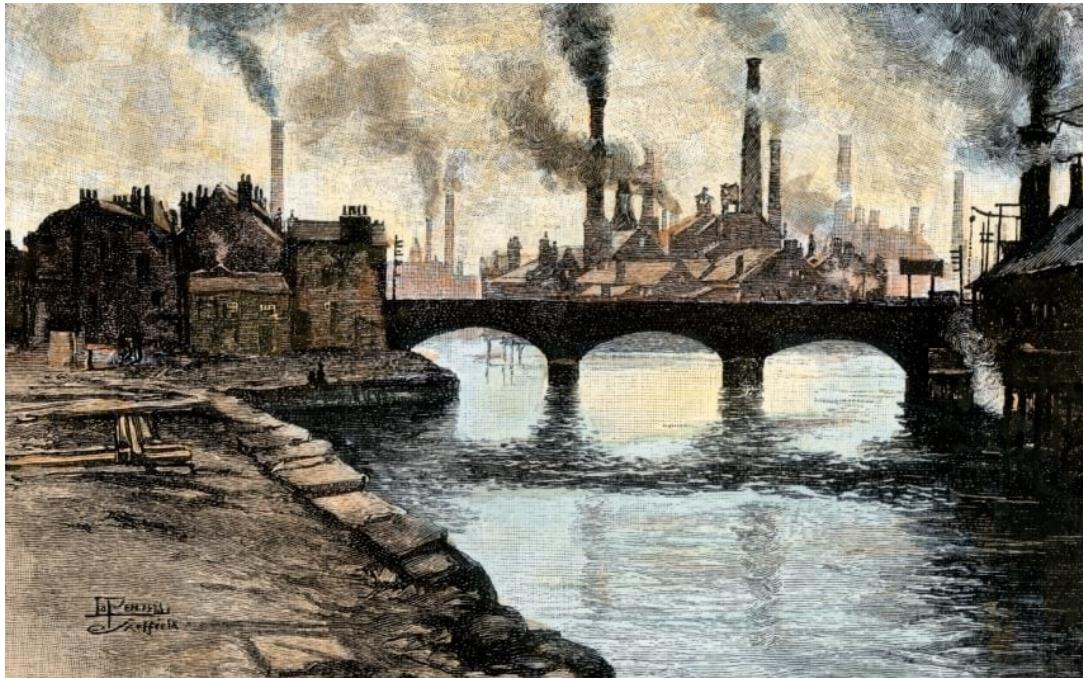
John Stevenson, ‘British Society 1914-1945’, in *The Pelican Social History of Britain* (1984, p. 381).

[...] But the **growth** of leisure illustrates more than commercialism and **more** free time from work. With the **growth** of the media, it was part of the development of **a more uniform and homogenous** society, partaking of an **increasingly** common culture. Notwithstanding regional and class differences, by 1945 **only the remotest** parts of Britain were insulated against the pervasive influences of the latest **popular** tune or **major** sporting event. In contrast, some aspects of leisure, particularly those centred around hobbies and domestic life, reflected an **increasing** home-centredness. The two themes of an **increasingly** common culture, balanced by the cult of domesticity and individual choice, dominated the development of leisure in this period”.

John Stevenson, ‘British Society 1914-1945’, in *The Pelican Social History of Britain* (1984, p. 381).

QUANTIFICATION IN HISTORY - EXAMPLE

Did the condition of the working classes improve or deteriorate during the industrial revolution?



Qualitative sources:

- Contemporary reports
- Newspapers
- Institutional minutes
- Private diaries & letters
- Travellers' accounts
- ...

Optimists vs pessimists

(comparisons with the past and other countries)

QUANTIFICATION IN HISTORY - EXAMPLE

The pessimists



“[Before the Industrial Revolution] the workers enjoyed a comfortable and peaceful existence. ... Their standard of life was much better than that of the factory worker today”.

Friedrich Engels, *The Condition of the Working Class in England*, 1845.

QUANTIFICATION IN HISTORY - EXAMPLE

The pessimists



“[...] the various forms of epidemic, endemic, and other diseases caused, or aggravated, or propagated chiefly amongst the labouring classes by atmospheric impurities ..., by damp and filth, and close and overcrowded dwellings prevail amongst the population in every part of the kingdom”.

Edwin Chadwick, *Report on the Sanitary Condition of the Labouring Population, 1843.*

QUANTIFICATION IN HISTORY - EXAMPLE

The optimists



“The labouring classes of this island, though they have their grievances and distresses ... are on the whole better off as to physical comforts than the inhabitants of any equally extensive district of the old world. For this very reason, suffering is more acutely felt and more loudly bewailed here than elsewhere”.

Lord Macaulay, Review of Southey's *Colloquies*, 1830.

QUANTIFICATION IN HISTORY - EXAMPLE

The optimists



“If we look back to the condition of the mass of the people as it existed in this country ... and then look around us at the indications of greater comfort and respectability that meet us on every side, it is hardly possible to doubt that here ... the elements of social improvement have been successfully at work, and that they have been and are producing an increased amount of comfort to the great bulk of the people”.

G. R. Porter, *The Progress of the Nation*, 1847.

QUANTIFICATION IN HISTORY – EXAMPLE

Negative accounts were more common but...



- How much subjectivity these opinions involve?
- How representative are they of the typical population?
- Can we actually be more precise? Is it possible to make comparisons over time and across regions?

QUANTIFICATION IN HISTORY – EXAMPLE

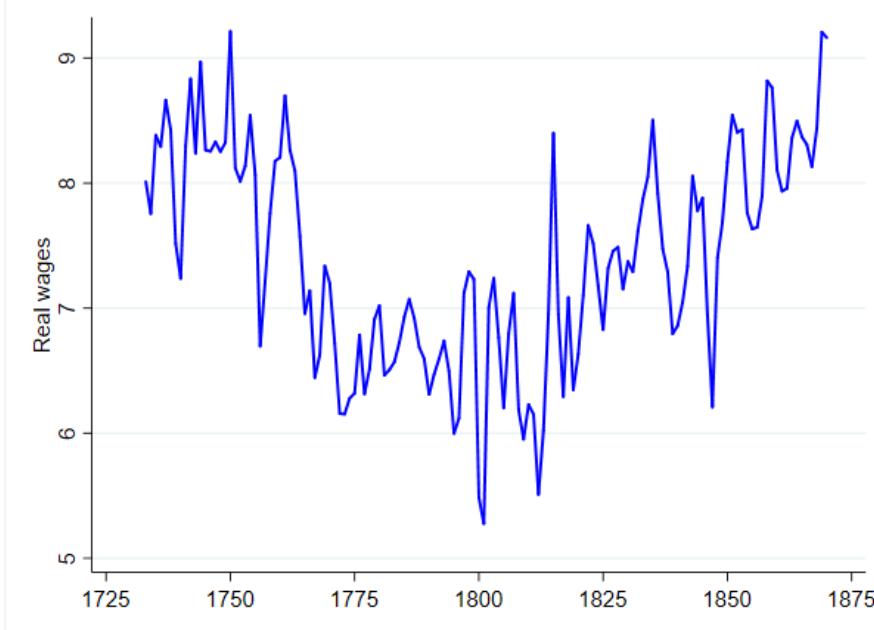
- Historians have employed data on wages to reflect changes in living standards

Wages of unskilled labourers working on London construction sites.

Source: Stephenson (2018)

QUANTIFICATION IN HISTORY – EXAMPLE

Real Wages, 1725-1875



- Wages of unskilled workers in London
- Qualified by the price level, thus a proxy of their purchasing power
- It first declines and then increases from 1815 onwards

Sources: [Allen et al. \(2011\)](#)

QUANTIFICATION IN HISTORY – EXAMPLE

- Historians have also turned to military and prison records to trace biological living standards
- **Heights** reflect net nutrition
 - nutrition, working hours, living conditions, disease environment...

253
4

Name, No. John Hearn 5722 14 June 73.

and Aliases.

Description

Age (on discharge)	12
Height	4 ft 6 $\frac{1}{2}$
Hair	Br. Brown
Eyes	Blue
Complexion	Fair
Where born	Lambeth
Married or Single	Single
Trade or occupation	Errand boy
Distinguishing marks.	None



Address at time of apprehension 151 Peagut St. Lambeth Wart

Place and date of conviction Lambeth 29 May 73.

Offence for which convicted Simple Larceny - 3.0 a. & 1/2 pieces of leather = 2/-

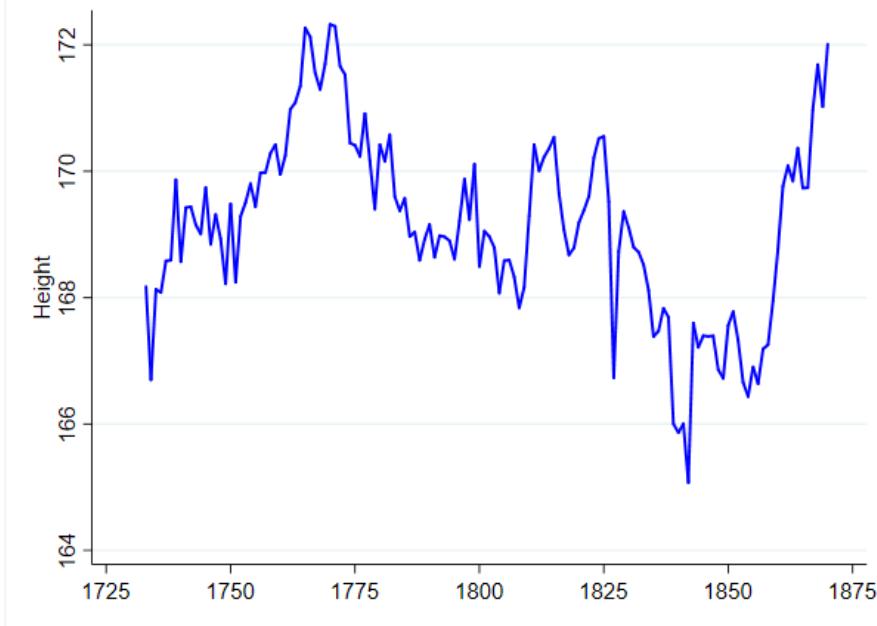
Sentence 1 Cal. Mth. & F.

Date to be liberated 20 June 73

Intended residence after liberation As above

QUANTIFICATION IN HISTORY – EXAMPLE

Heights, 1725-1875



- The heights of the military only show improvements after 1850 (even later than with wages)

Sources: [Meredith and Oxley \(2014\)](#)

WHY QUANTIFICATION IS IMPORTANT?

- (1) Assess the importance of historical processes
 - Be more precise regarding:
 - changes over time
 - differences across space or populations
- (2) Test historical hypotheses
- (3) Address questions that are less visible using other sources
- (4) Discerning pattern from chance

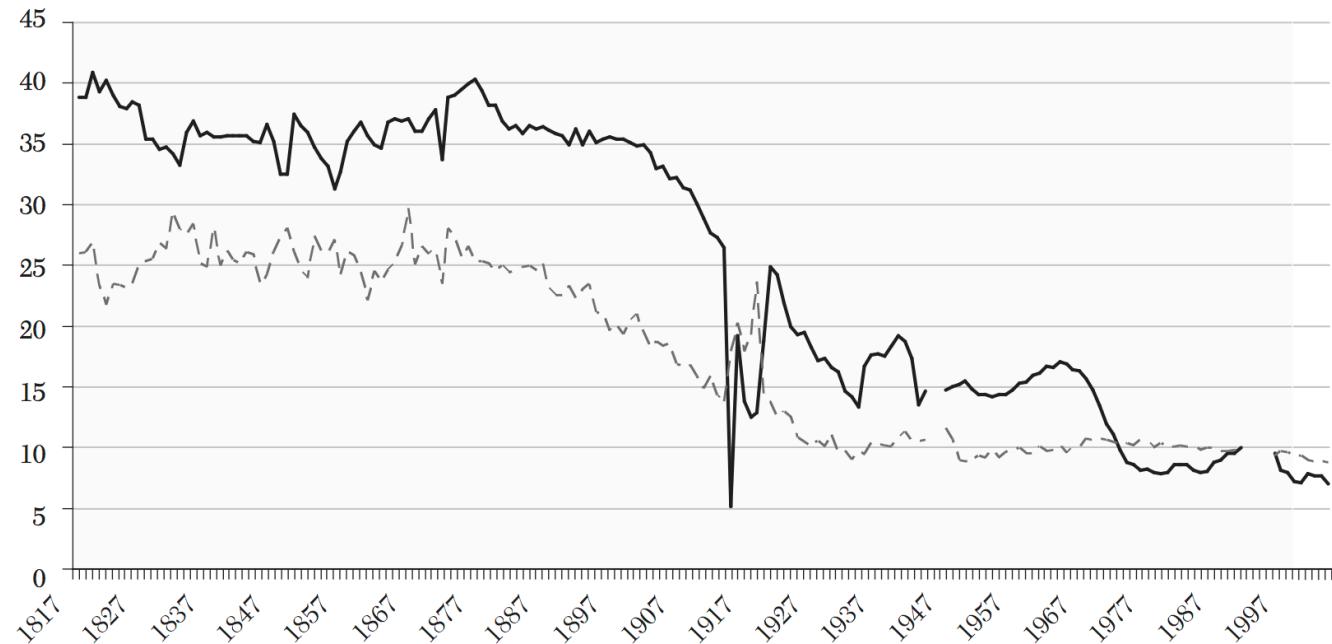
It can be applied to both numerical and qualitative information

MEASURE CHANGES OVER TIME / ACROSS POPULATIONS

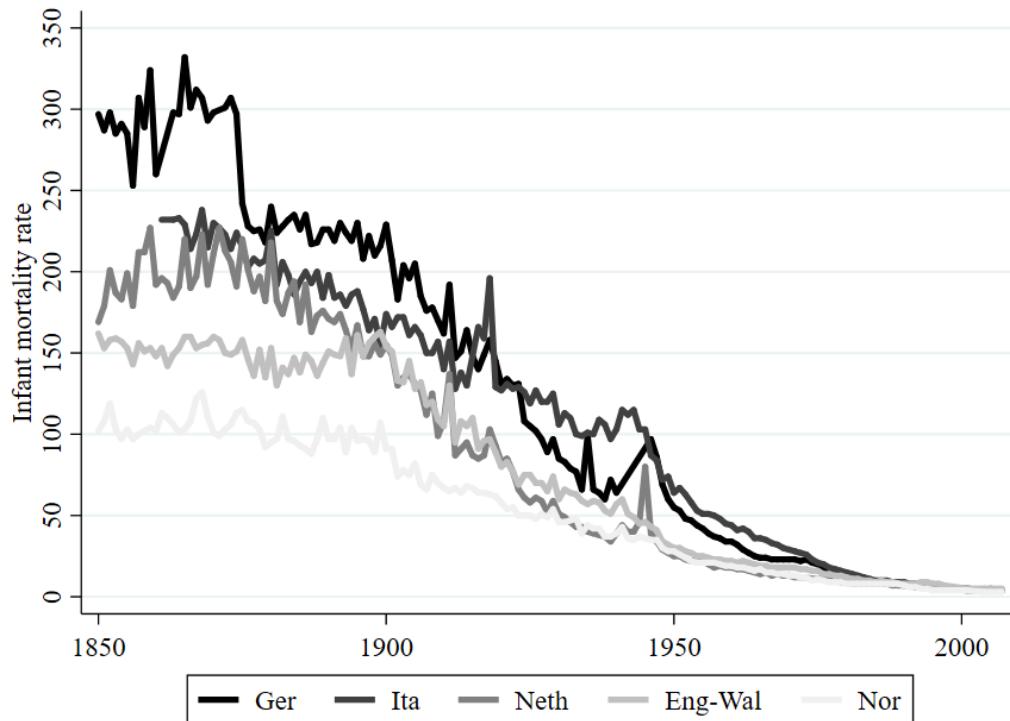
(1)

- Assess the importance of historical processes
 - The Industrial Revolution, the Demographic Transition, etc

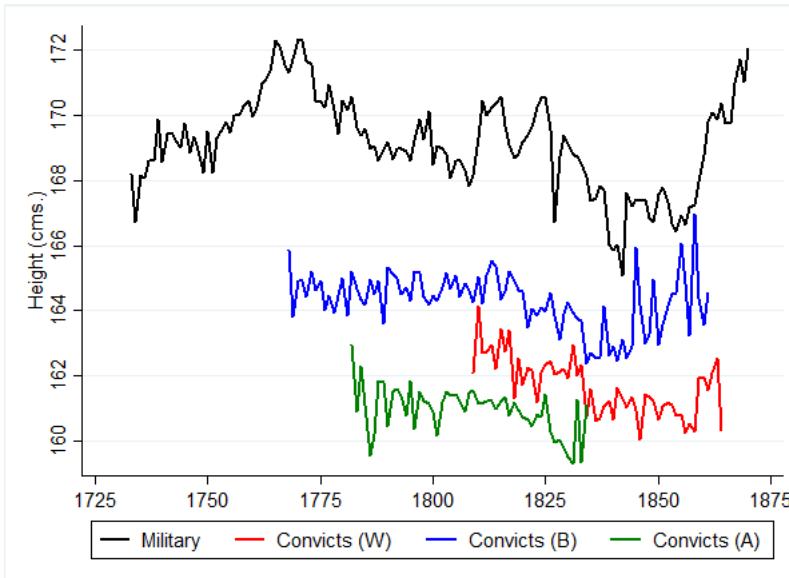
Birth and death rates in Germany, 1817-2005
[\(Guinnane 2011\)](#)



Infant mortality rates in Europe, 1850-2010



Heights for different groups



Source: [Meredith and Oxley \(2014\)](#)

- Heights in the:
 - Black: Military
 - Red: Wandsworth Prison (Surrey)
 - Blue: Bedford Prison
 - Green: Convicts sent to Australia

Large socio-economic differences

- Formally assess the relationship between variables, causation, etc.

Did inequality hampered or promoted education?

Two examples using:

- regional-level data
- individual-level data

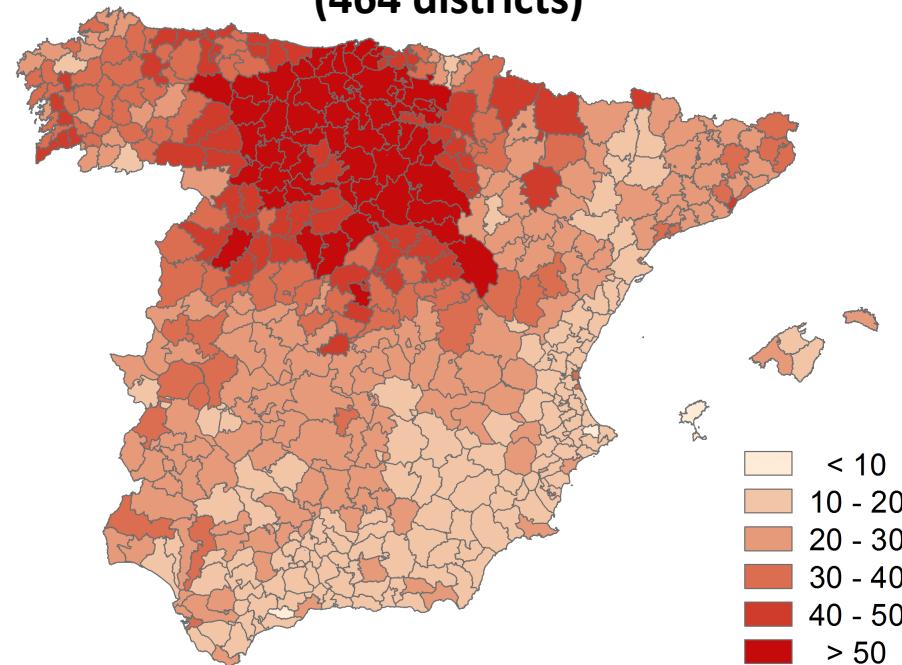
TESTING HISTORICAL HYPOTHESES

(2)

- Regional-level information from the 1860 Spanish Population Census:

PARTIDOS JUDICIALES.	NÚMERO DE		HABITANTES QUE			TOTAL
	Ayunta- mientos.	Cédulas de inscripción.	Saben leer y no escribir.	Saben leer y escribir.	No saben leer.	
ALBA DE TÓRREZ.....	47	4.893	734 851	4.240 817	5.359 8.189	20.493
BÉJAR.....	40	8.524	1.212 1.083	5.781 1.618	10.392 15.200	35.316
CIUDAD RODRIGO.....	63	10.493	1.465 1.293	7.391 1.295	12.750 18.281	42.483
LEMONTE.....	52	6.105	920 1.212	5.986 1.069	6.818 10.603	26.670
PESCARADA DE BRACAMONTE.....	36	6.062	808 1.212	5.569 1.110	6.868 10.515	26.412
SALAMANCA.....	61	9.970	1.058 2.411	11.422 3.751	8.628 15.009	42.272
SEGOVIA.....	67	7.658	902 767	5.892 956	8.137 13.275	29.829
TITIGÜINOS.....	46	9.826	1.286 1.776	7.233 1.685	10.911 15.931	38.818
	390	63.533	8.405 10.670	53.117 12.661	70.196 107.331	262.383
TOTAL GENERAL...	390	63.533	19.075	65.781	177.527	262.383

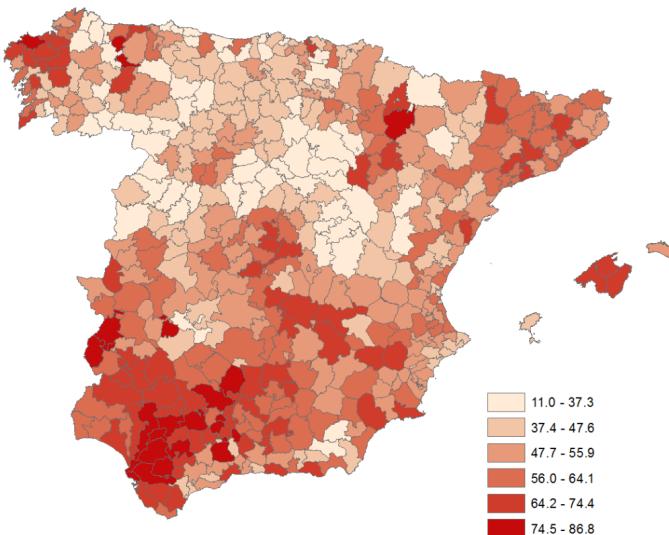
Male literacy levels, 1860
(464 districts)



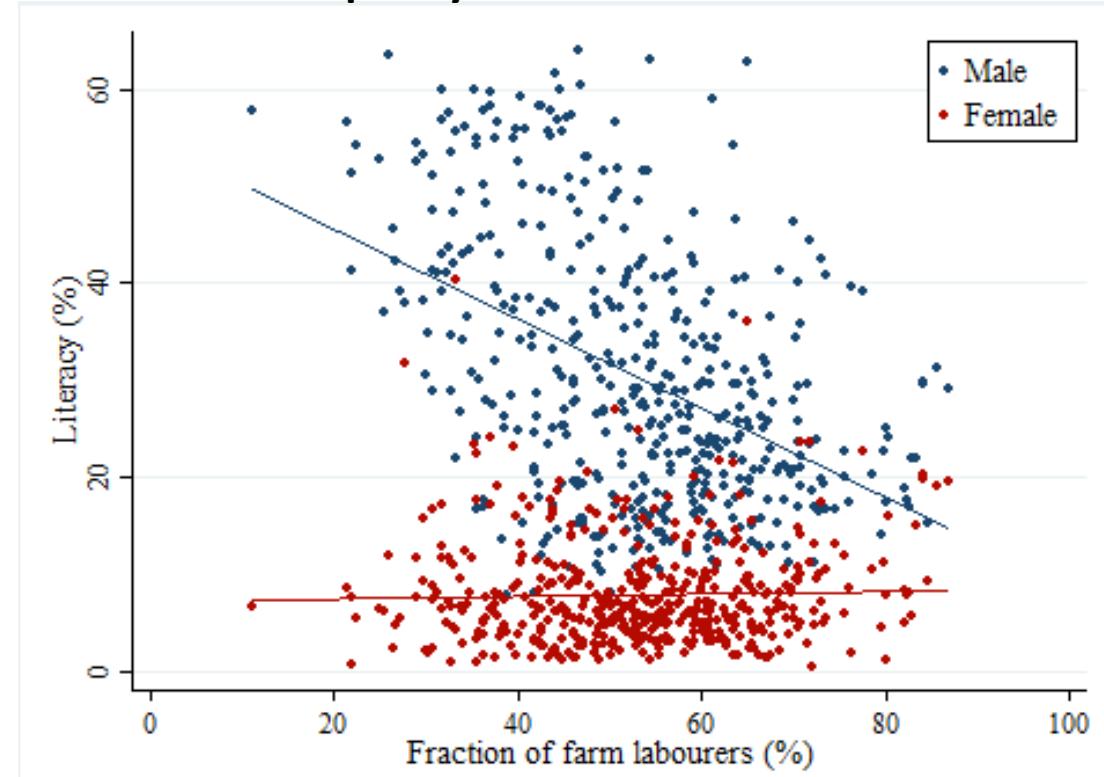
TESTING HISTORICAL HYPOTHESES

(2)

Landless labourers (%), 1860



Inequality and education in 1860



TESTING HISTORICAL HYPOTHESES

(2)

- Individual-level information from the Population Register of Madrid

AYUNTAMIENTO DE MADRID PADRÓN MUNICIPAL QUINQUENAL												DISTRITO DE <u>HAMBERI</u> DISTRIBUICIÓN DE <u>MALASAÑA</u>											
Diciembre de 1905. (Artículos 17, 18 y 20 de la ley Municipal).												(v) calle de <u>Malasaña</u> n.º <u>19</u> cuarto <u>5º</u> piso											
SE RECOMIENDA MUY ESPECIALMENTE LA LECTURA DE LA INSTRUCCIÓN QUE VA A LA FLETA												Denominación o destino del edificio, si fuera público _____											
HOJA DECLARATORIA N.º <u>4103</u>												Industria o comercio que se ejerce en la habitación _____											
SOBRENOMBRE		APELLIDO PATERNO		APELLIDO MATERO		PESO Y LUGAR DEL NACIMIENTO				ESTADO	PROVINCIA sobre el escudo de armas	OPCIÓN Militar, mercantil, industrial, desde que se practica, oficio o profesión.	CLASIFICACIÓN 4 bandas y punto	IMPRESO EN EL ESTADO DE ESTADOUNIDENSE	SACADO en	TIPO de habitación	TIPO de habitación	TIPO de habitación	TIPO de habitación				
SI	NO	SI	NO	SI	NO	DÍA	MES	AÑO	PAÍS											PROVINCIA	DOMICILIO	DOMICILIO	DOMICILIO
<u>Gregorio</u>	<u>Martínez</u>	<u>López</u>		<u>6 Marzo 1891</u>	<u>Madrid</u>	<u>Madrid</u>	<u>Casa Estadounidense</u>		<u>—</u>	<u>Gregorio Martínez López</u>	<u>1905</u>	<u>Si</u>	<u>Si</u>	<u>Si</u>	<u>Si</u>	<u>Si</u>	<u>Si</u>						
<u>Maria del Rosario</u>	<u>Lejárraga</u>	<u>García</u>		<u>19 Enero 1895</u>	<u>San Blas de Legorreta</u>	<u>Legorreta</u>	<u>Casa Estadounidense</u>	<u>1905</u>	<u>—</u>	<u>Maria del Rosario Lejárraga García</u>	<u>1905</u>	<u>Si</u>	<u>Si</u>	<u>Si</u>	<u>Si</u>	<u>Si</u>	<u>Si</u>						
												<i>F. Martínez Pérez</i>											

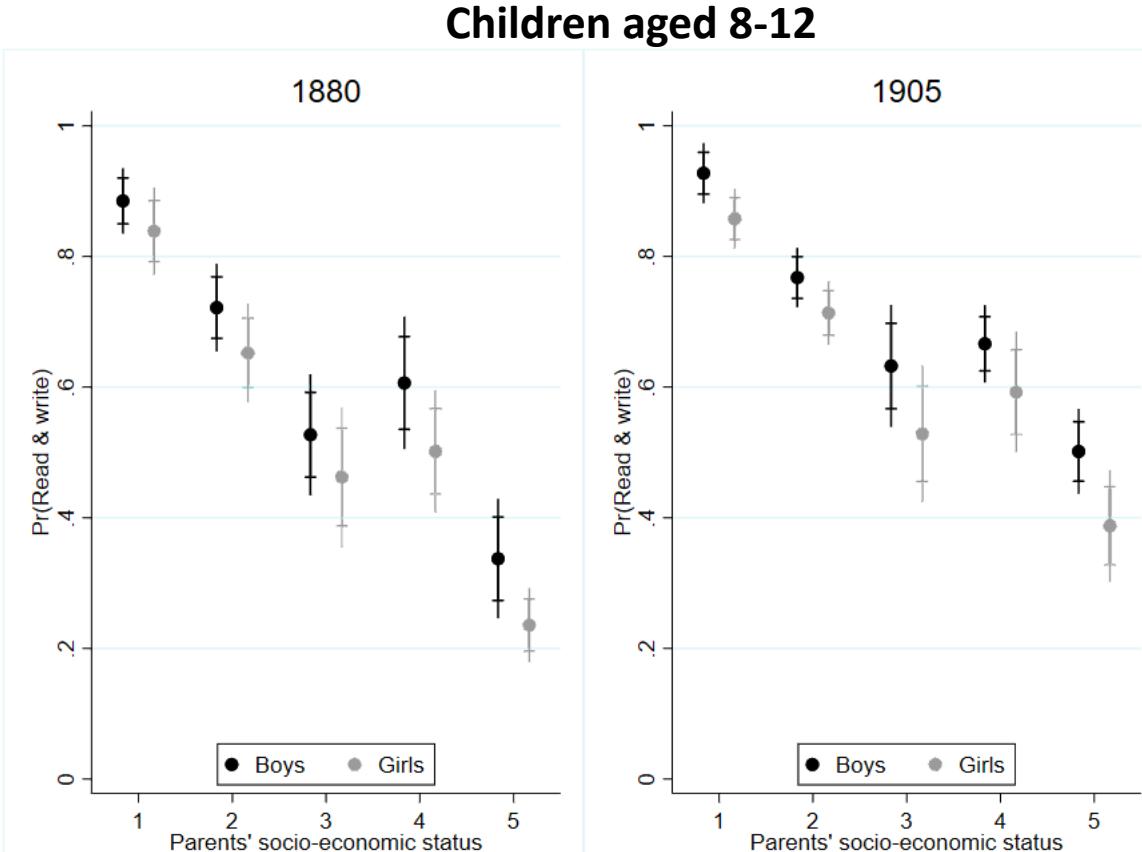
Name, address, birthdate, birthplace, marital status, children, occupation, literacy, etc.

TESTING HISTORICAL HYPOTHESES

(2)

Probability of being literate according to parents' socio-economic status

- 1. Higher managers and professionals
- 2. Lower managers, professionals, clericals & sales personnel
- 3. Foremen & medium-skilled workers
- 4. Low-skilled workers
- 5. Unskilled workers



ADDRESS ISSUES THAT ARE LESS VISIBLE

(3)

- Explore questions that are:
not visible using other
sources (or difficult to
assess their importance)
 - About particular populations
i.e. lower social groups
 - About particular behaviour
i.e. gender discrimination,
contraception, etc.



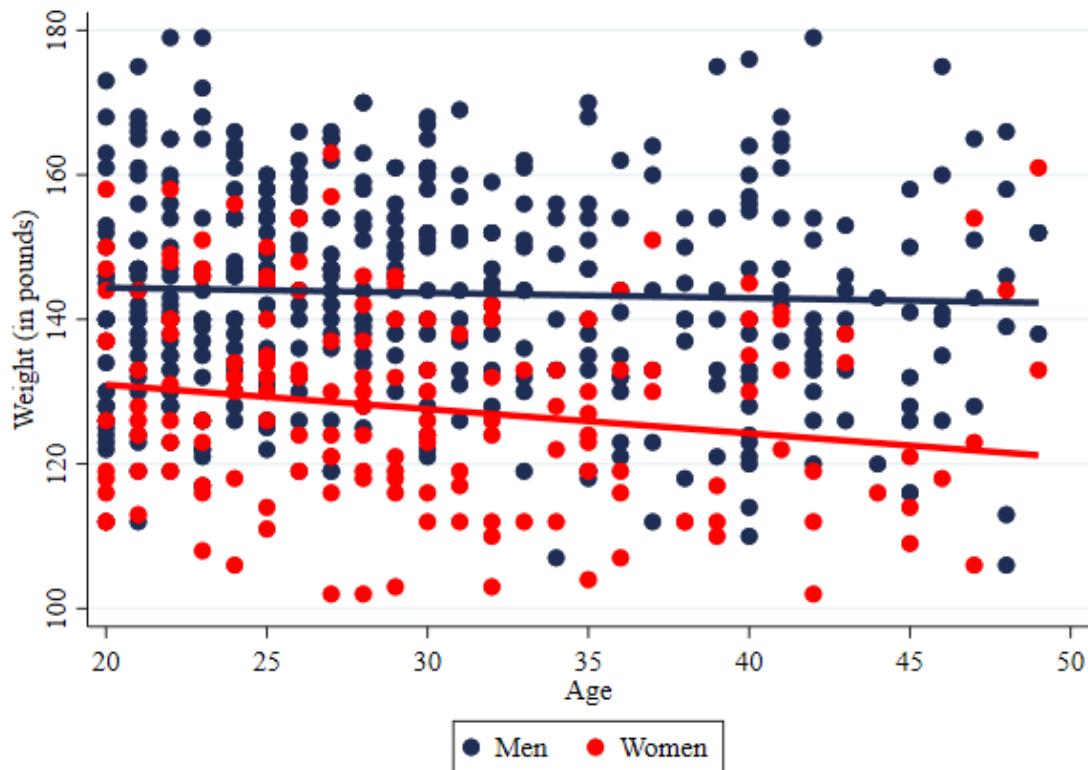
© Hulton-Deutsch Collection/CORBIS

ADDRESS ISSUES THAT ARE LESS VISIBLE

(3)

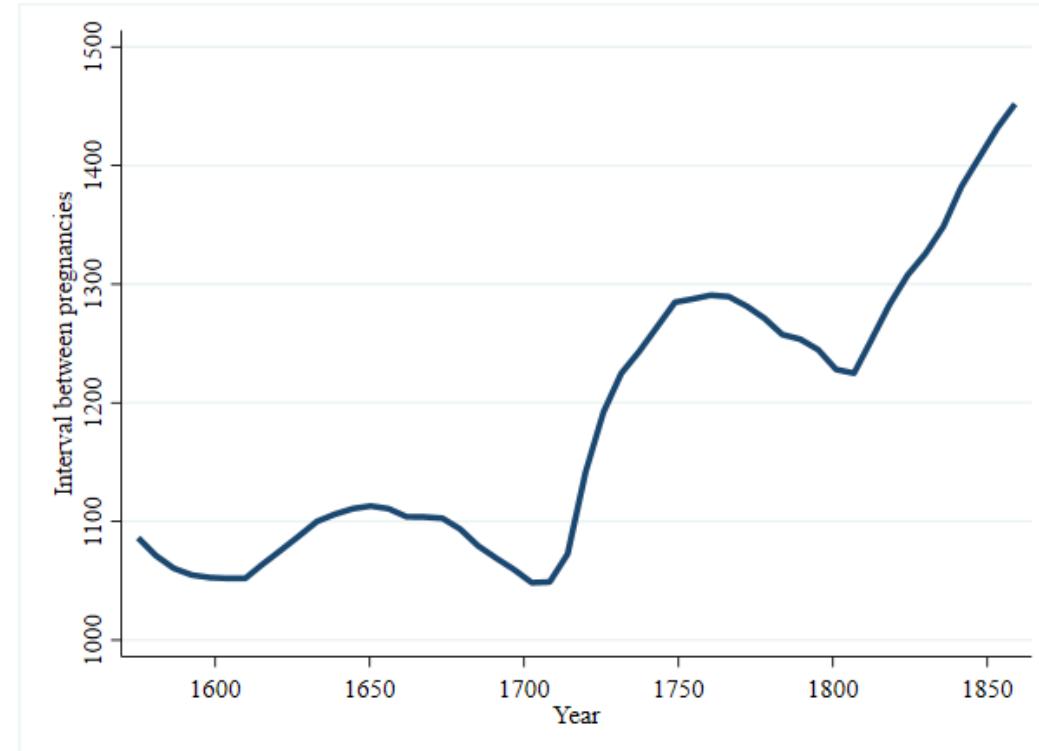
Contrary to men's, women's **weight** in 19th century Britain seems to have declined as they grew older (according to info on prisoners)

Unequal allocation of resources within families?



Parish registers from 13 villages in NE Spain between 1575 and 1869 show that **the interval between the first and second birth** increased over time

Fertility control?



- Measure the accuracy of estimations based on historical samples
 - How well does the observed result fit the overall population?

If a different sample would have been drawn, we would have obtained a different result but... how different?

- The ideal solution is to obtain more samples but this is costly and often impossible if no further evidence is available

Statistical theory is used instead

- We draw conclusions about the characteristics of the entire **population** based on a **sample** of observations
 - The observed fact will however deviate from the true value in the population

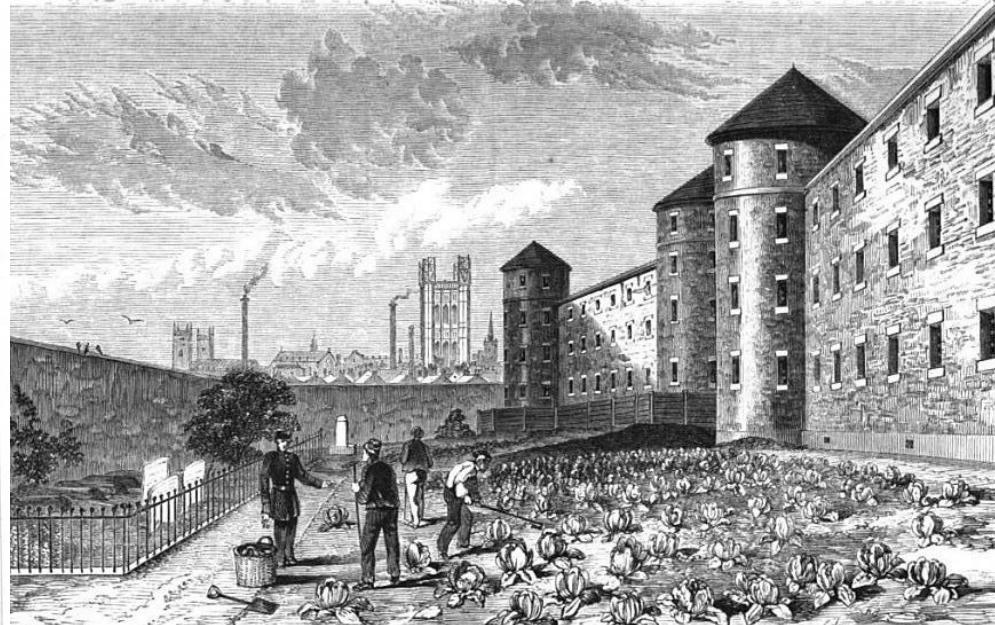
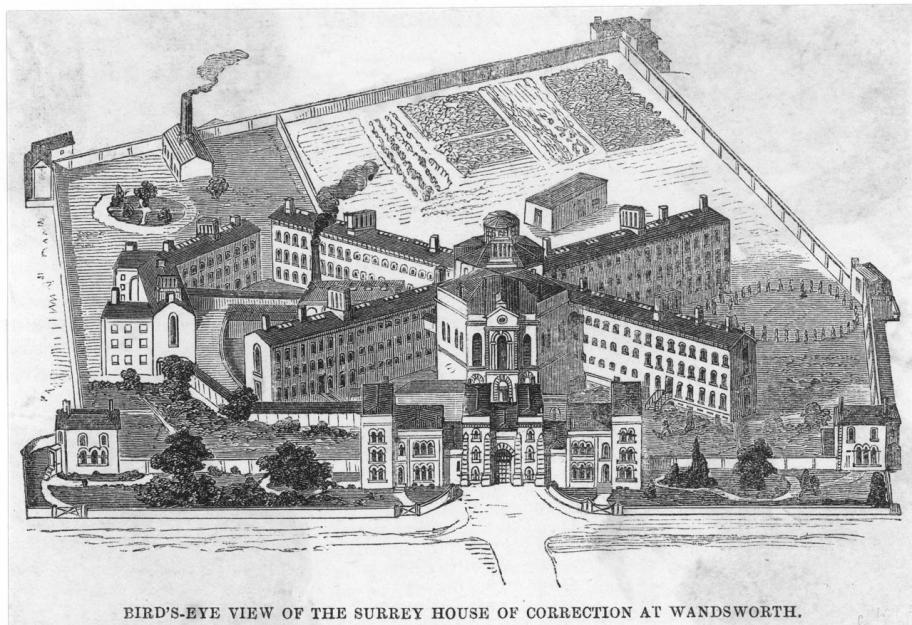
How the mean height of a sample of students here will approximate the mean height of all the class?

The **sampling error** depends on:

- the variability within the population
- the sample size



- Let's illustrate this with a historical example taken from a sample of prisoners from the Wandsworth House of Correction (nearby London)



DISCERN PATTERN FROM CHANCE

(4)

National Archives

- Individual Prison Registers

Education, occupations, ages and types of crimes (petty crimes) suggest that these prisoners were drawn from the low working classes

- a large proportion of the population at the time

Which prisoners were taller?

253
4

Name, No. John Hearn 5722 14 June 73.

and Aliases.

Description

Age (on discharge)	12
Height	4 ft 6 1/2
Hair	Br. Brown
Eyes	Blue
Complexion	Flesh
Where born	Lambeth
Married or Single	Single
Trade or occupation	Errand boy
Distinguishing marks	None



Address at time of apprehension 151 Regent St. Lambeth Walk

Place and date of conviction Lambeth 29 May 73.

Offence for which convicted Simple Larceny - 100s. Stg 11 pence leather = 2/-

Sentence 1 Cal. M.R. & T.

Date to be liberated 20 June 73

Intended residence after liberation As above

DISCERN PATTERN FROM CHANCE

(4)

casen	no	month	year	forename	surname	sex	age	born	countryb	reside	feet	inches	weight	occup	employed	literacy	
1	1	17	january	1841	AGNES	M' INTYRE	female	24	paisley	scotland	colinslee	5	5	-	prostitute	unemployed	read a little
2	2	45	january	1841	CATHERINE	CARLIN OR WILSON	female	30	irvine	scotland	paisley	4	5	-	shewing	employed	illiterate
3	3	68	january	1841	JEAN	WRIGHT	female	17	paisley	scotland	paisley	5	1	-	prostitute	unemployed	read a little
4	4	91	february	1841	MARGRET	M'HAFFERTY	female	18	glasgow	scotland	paisley	5	11.5	-	prostitute	unemployed	read tolerably
5	5	93	february	1841	JANET	M'LEAN	female	25	cathcart	scotland	strabungo	5	6	-	servant	employed	read a little
6	6	263	april	1841	ELIZA	DUNCAN	female	34	belfast	ireland	paisley	5	8	-	prostitute	unemployed	read a little
7	7	280	april	1841	ANN	RYLEY	female	45	sligo	ireland	glasgow	5	5	-	hawking	unemployed	illiterate
8	8	299	april	1841	MARGRET	M' LEOD	female	40	greenock	scotland	greenock	5	.5	-	hawking	employed	read & write tolerably
9	9	300	april	1841	MARY	MILLAR OR LINTON	female	19	greenock	scotland	greenock	5	1.25	-	mill girl	unemployed	read a little
10	10	310	april	1841	JEAN	M' KINLAY OR POLLOCK	female	29	glasgow	scotland	pollockshaws	5	3.5	-	hawking	employed	read a little
11	11	343	may	1841	AGNES	CURRIE OR LACHLAN	female	35	islay	scotland	paisley	4	11.5	-	hawking	employed	read a little
12	12	382	june	1841	ELLIZA	MUNN	female	18	johnston	scotland	johnston	5	0	-	shewing shawls	unemployed	read a little
13	13	425	june	1841	SARAH	BLACK OR M' PHERSON	female	36	glasgow	scotland	kelvindock	4	11.25	-	millgirl	unemployed	read a little
14	14	41	january	1844	JANET	McCULLOCK	female	34	derry	ireland	pollockshaws	4	9.75	-	dressmaker	employed	cannot write
15	15	43	january	1844	AGNES	GIBBON	female	15	paisley	scotland	paisley	5	2	-	vagrant	unemployed	illiterate
16	16	122	february	1844	MARY	KEITH	female	14	paisley	scotland	paisley	4	10	-	mill girl	unemployed	illiterate
17	17	133	march	1844	AGNES	CRAIG	female	16	johnston	scotland	johnston	5	2	-	millgirl	unemployed	cannot write
18	18	236	may	1844	DIANNA	KELLY	female	16	paisley	scotland	paisley	5	1.5	-	draw girl	unemployed	read a little
19	19	237	may	1844	MARGARET	JAAP	female	17	paisley	scotland	paisley	5	1	-	sewer	unemployed	illiterate
20	20	303	june	1844	HELEN	McCREADY OR YOUNG	female	42	ireland	ireland	pollockshaws	5	4.5	-	housekeper	unemployed	illiterate
21	21	311	june	1844	BIDDY	MURRAY	female	32	sligo	ireland	barrhead	5	2.75	-	hawker	employed	illiterate
22	22	313	june	1844	ROSE	DONELLY	female	28	glasgow	scotland	greenock	4	10	-	prostitute	unemployed	illiterate
23	23	335	june	1844	MARY ANN	GILLESPIE	female	30	ireland	ireland	barrehad	5	0	-	bleacher	unemployed	illiterate
24	24	350	june	1844	AGNES	MARTIN OR MUNRO	female	40	greenock	scotland	erskine	5	.5	-	hawker	employed	illiterate
25	25	351	june	1844	JANET	MITCHELL	female	25	linthigow	scotland	paisley	5	1.5	-	prostitute	unemployed	cannot write
26	26	360	june	1844	JEAN	DYKES	female	22	rathgen	scotland	paisley	5	1.5	-	house keeper	unemployed	read a little
27	27	370	july	1844	HELEN	DOLLAN	female	30	crosslie	scotland	paisley	5	3	-	sewer	unemployed	illiterate
28	28	447	august	1844	MARY ANNE	MOORE OR ALLISON	female	33	ireland	ireland	paisley	5	2.5	-	housekeper	employed	cannot write
29	29	455	august	1844	ISABELLA	McARTHUR	female	50	highlands	scotland	barrhead	5	5	-	hawker	employed	illiterate
30	30	501	september	1844	MARY	McTAGGART GIBB	female	43	prestick	scotland	catrine	5	1	-	housekeeper	employed	cannot write
31	31	505	september	1844	JANET	McKAY	female	24	bridgeton	scotland	johnston	5	5	-	labourer	unemployed	cannot write
32	32	581	october	1844	MARGT	DOUGALL	female	28	port glasgow	scotland	port glasgow	5	1.5	-	muslin sewer	employed	cannot write
33	33	591	november	1844	MARY	MCDONALD	female	17	manchester	england	paisley	5	1	-	housekeeper	employed	cannot write
34	34	640	november	1844	ELIZA	McLACHLAN	female	52	paisley	scotland	paisley	5	1	-	housekeeper	employed	read & write a little

DISCERN PATTERN FROM CHANCE

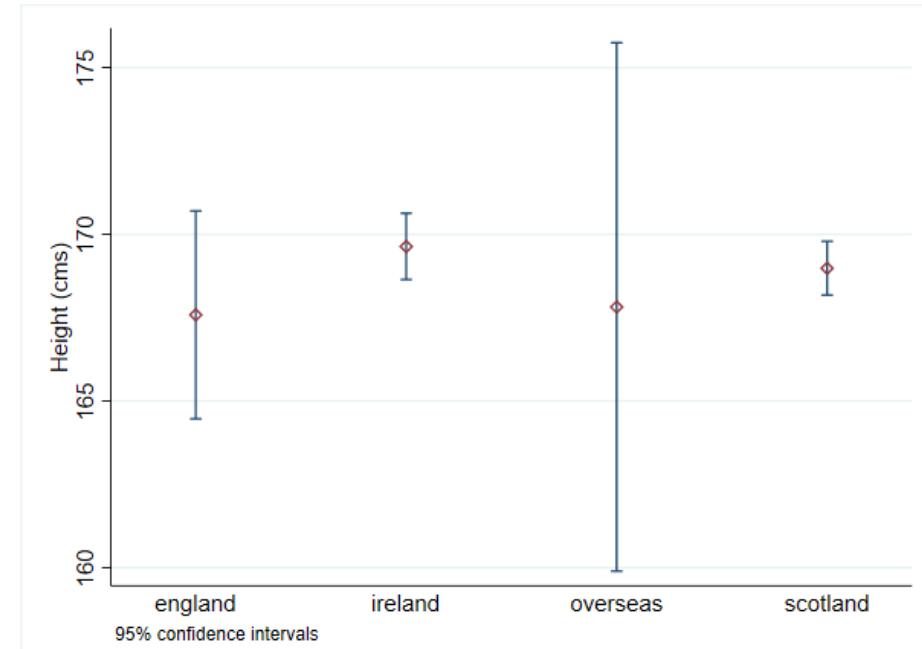
(4)

Average height, men aged 20-50 (1841-1883)

Country of birth	Height	Obs.
England	167.6	23
Ireland	169.6	161
Scotland	169.0	324
Overseas	167.8	7

A confidence interval indicates that the true population parameter will lie within that range with a 95% confidence

Are these groups different in terms of heights?

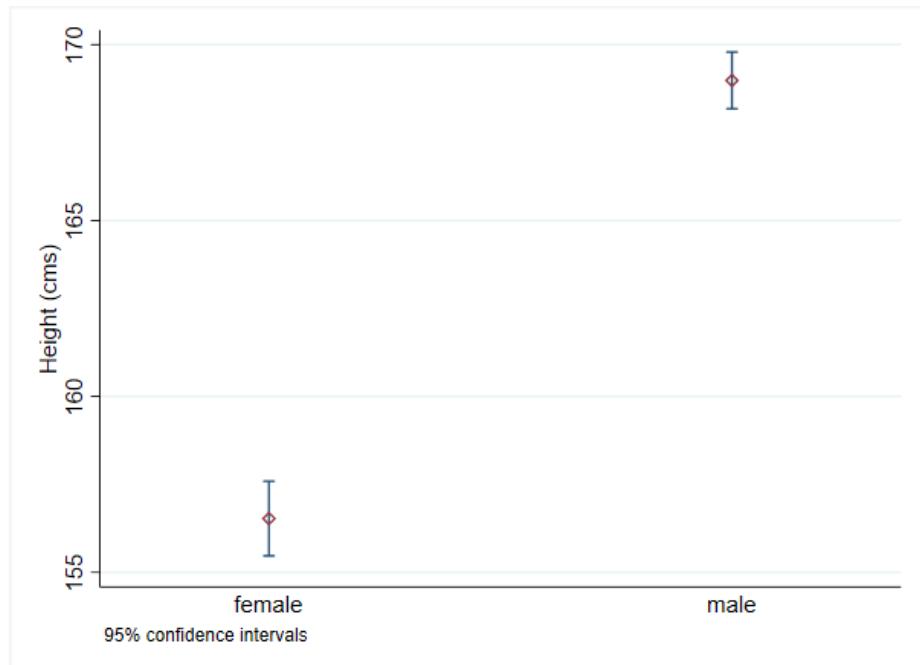


Average height, Scottish prisoners aged 20-50 (1841-1883)

Country of birth	Height	Obs.
Men	169.0	324
Women	156.5	143

Scottish men's are women's height (169.0 vs 156.5) are statistically different (at least with a 95% confidence)

Male and female average height



DISCERN PATTERN FROM CHANCE

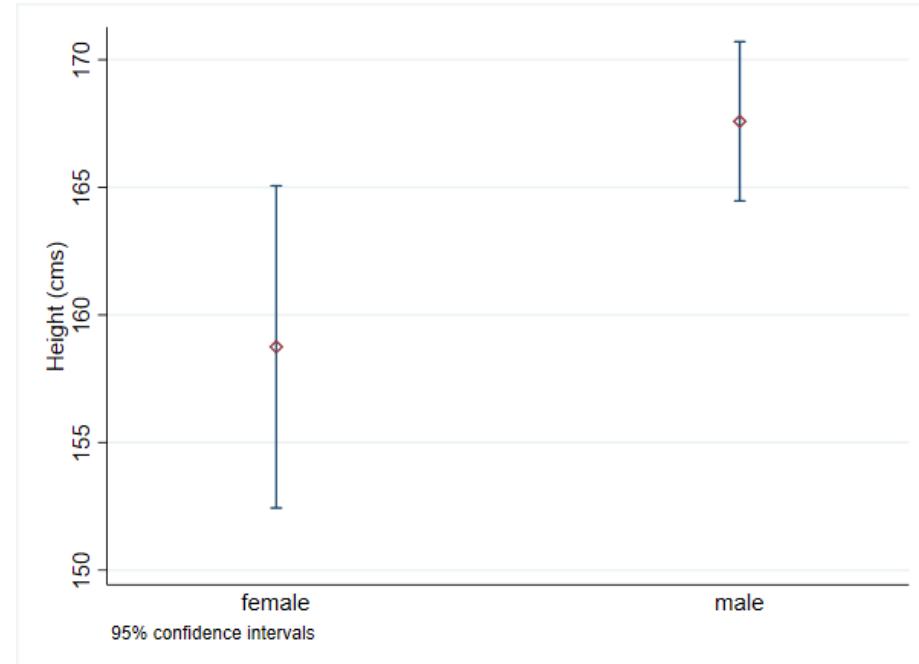
(4)

Average height, English prisoners
aged 20-50 (1841-1883)

Country of birth	Height	Obs.
Men	167.6	23
Women	158.7	3

We cannot be sure that English men
and women had different statures
(at least with a 95% confidence)

Male and female average height



ADVANTAGES OF QUANTIFICATION

- Simplify, summarise and display raw data
- Rigorous testing of hypothesis:
 - Assess the relative importance of historical processes
 - Discern pattern from chance (statistical significance - sample size)
 - Relationship between variables, causation, etc.
 - Explore questions that are less visible using other sources

Less subject to the
narrator's subjectivity

We will address some of these dimensions in the next session

WHY QUANTIFICATION IS IMPORTANT?

- To answer questions such as:
 - What impact did industrialization have on living standards?
 - Who migrated? Were migrants healthier and more literate than the ones who remained behind?
 - How important was gender discrimination for girls' education?
 - What was the effect of enclosure on agricultural productivity?
 - When did fertility rates begin decreasing in Europe? Is this process related to the fall in mortality?
 - ...

WHY QUANTIFICATION IS (INCREASINGLY) IMPORTANT?

- Strengthen by availability of digitalised sources and computing power
 - For both numeric and textual (content) analyses
- Increasing use of statistics as a language of enquiry in the social sciences and public debates (why not in history?)
- Statistical techniques are actually increasingly applied to historical research
 - Importance of being able to critically engage with this literature

Valuable research tool
(to complement other
sources & methods)

Graphs, tables and maps
can be very persuasive

- Data is never perfect, especially in a historical setting (precision illusion)
 - Reliability, representativeness & biases (missing data: selection / survivorship)
 - Comparability problems (over time and across regions): definitions vary
- Manipulating data may produce distortions
- Correlation is not causation: omitted variables, reverse causality
- Statistical vs historical significance
- Being driven by techniques instead of historical questions
- Simplistic representation of reality (economic/demographic determinism)

Be aware of the problems with both the data and the techniques applied

Historical & quantitative skills are required to evaluate and analyse them

NOT EVERYTHING IS SUBJECT TO QUANTIFICATION

- ❑ Not appropriate for all questions: cultural / intellectual / political ... history

“[...] almost all important questions are important precisely because they are **not** susceptible to quantitative answers.”

Arthur Schlesinger, Jr., *The Humanist Looks at Empirical Social Research* (1962), p. 770.

NOT EVERYTHING IS SUBJECT TO QUANTIFICATION

"[...] the grievances felt by working people as to changes in the character of capitalist exploitation: the rise of a master-class without traditional authority or obligations: the growing distance between master and man: the transparency of the exploitation at the source of their new wealth and power: the loss of status and above all of independence for the worker, his reduction to total dependence on the master's instruments of production: the partiality of the law: the disruption of the traditional family economy: the discipline, monotony, hours and conditions of work: loss of leisure and amenities: the reduction of the man to the status of an "instrument".

That working people felt these grievances at all -and felt them passionately- is itself a sufficient fact to merit our attention. And it reminds us forcibly that some of the most bitter conflicts of these years turned on issues which are not encompassed by cost-of-living series. The issues which provoked the most intensity of feeling were very often ones in which such values as traditional customs, "justice", "independence", security, or family-economy were at stake, rather than straightforward "bread-and-butter" issues. The early years of the 1830s are aflame with agitations which turned on issues in which wages were of secondary importance [...].

E. P. Thomson, *The Making of the English Working Class* (1963), p. 202-203.

NOT EVERYTHING IS SUBJECT TO QUANTIFICATION BUT...

□ This might be true to some extent but...

- Qualitative information can be counted and/or classified in different categories, which makes them amenable for quantitative analyses:
 - locations, marital status, race, sex, occupation, names, etc.
- Textual (content/topic) analysis can provide insights into what past actors thought about the world they were in
 - culture, values, motivations, etc.

→ turning unstructured texts into structured databases

CLASSIFYING QUALITATIVE INFO INTO CATEGORIES

NAMES

□ 1801 Census

	no1801a_namefrst	no1801a_namelast	no1801a_status	no1801a_marstrstr	no1801a_occstr
500468	Thorine	Eliædtr	Døttre	Ugift	
500469	Olene	Eliædtr	Døttre		
500470	Elen Marie	Eliædtr	Døttre		
500471	Bent	Davidsen	Huusbonde	Begge 1 gang givte	Bonde,gaardbruger,lods
500472	Anna Malena	Olsdtr	Hans kone	Begge 1 gang givte	
500473	Anders	Knudsen	Tienestefolk	Ugift	
500474	Anna	Pedersdtr	Tienestefolk		
500475	Ananias	Holgersen	Huusbonde	Begge i 1ste ægteskab	Huusmand med jord og lods
500476	Ingeborg	Gundersdtr	Hans kone	Begge i 1ste ægteskab	
500477	Gunder	Ananiasen	Børn		
500478	Holger	Ananiasen	Børn		
500479	Ole	Ananiasen	Børn		
500480	Agnete	Ananiasdtr	Børn		
500481	Anders	Gundersen	Huusbonde		
500482	Johanne Marie	Larsdtr	Hans kone		
500483	Johane Marie	Andersdtr	Døttre	Ugifte	
500484	Ingeborg Karine	Andersdtr	Døttre	Ugifte	
500485	Gunder	Thomæs	Tienestedrång		
500486	Hans	Didrichsen	Huusbonde	Begge i 1ste ægteskab	Bonde med lidet brug, reparerer stue-uhr, smed
500487	Magreta	Andersdtr	Hans kone	Begge i 1ste ægteskab	
500488	Gertrud	Hansdtr	Datter		
500489	Guri	Andersdtr	Tienestepige	Ugift	

Highly symbolical
Names may convey info about parents' preferences and cultural values

CLASSIFYING QUALITATIVE INFO INTO CATEGORIES

NAMES

- Most popular names in the following population censuses

1801

forename	obs
Ole	56506
Anne	31848
Anders	21319
Hans	21312
Peder	20518
Lars	19909
Kari	17527
Niels	14750
Karen	11565
Marthe	11405

1865

forename	obs
Ole	79521
Anne	41121
Hans	32519
Anders	25653
Lars	25536
Peder	24370
Ingeborg	22873
Karen	21787
Johan	20518
Johannes	19137

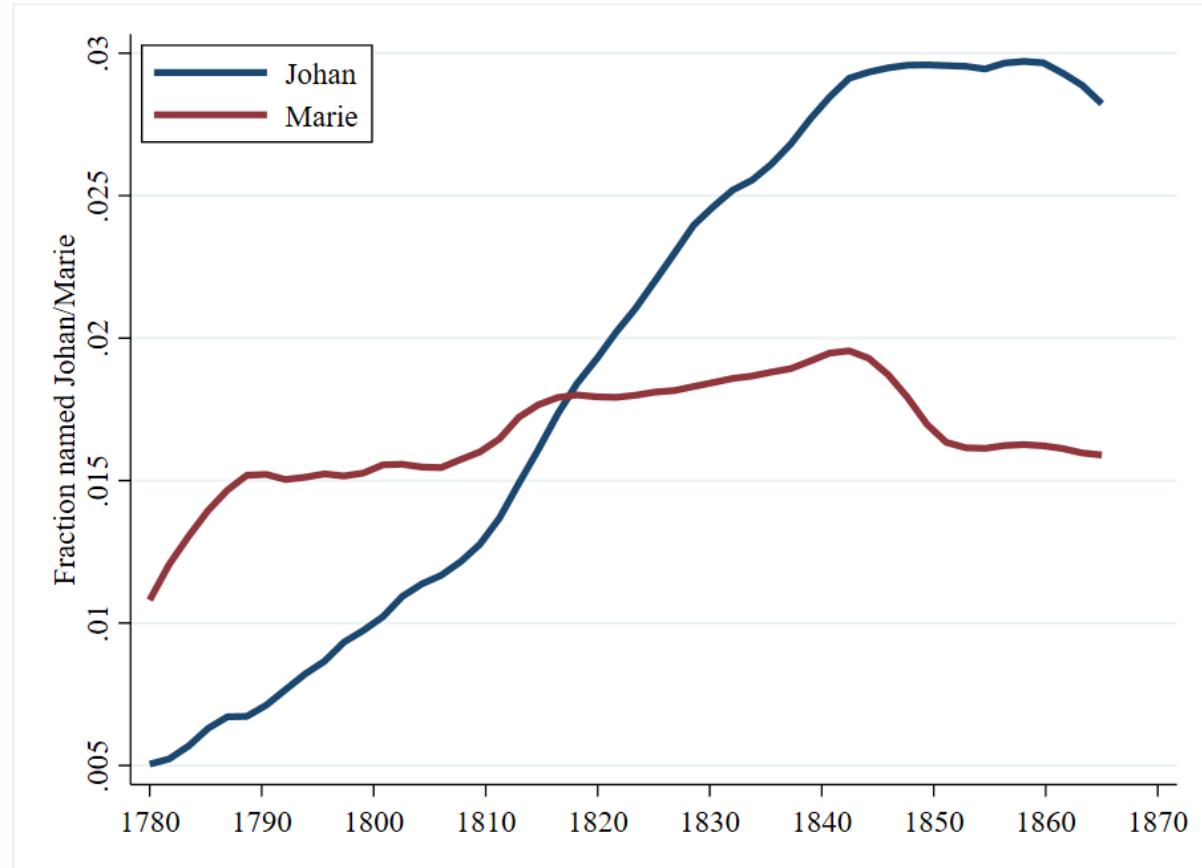
1900

forename	obs
Ole	61726
Anna	53750
Marie	36946
Johan	36581
Hans	33610
Ingeborg	25330
Anne	24272
Peder	23105
Karen	22660
Karl	22373

CLASSIFYING QUALITATIVE INFO INTO CATEGORIES

NAMES

- 1865 census
- Fraction of the (male/female) population named Johan or Marie (by year of birth)

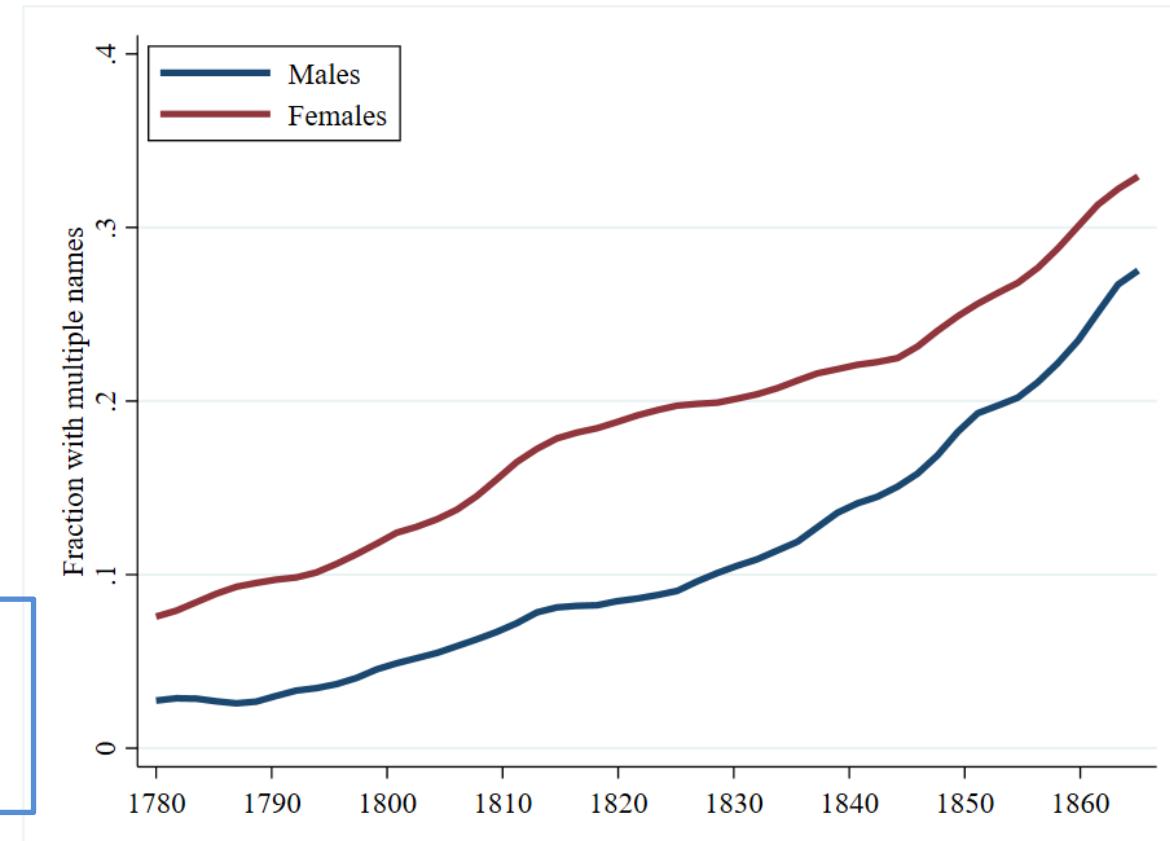


CLASSIFYING QUALITATIVE INFO INTO CATEGORIES

NAMES

- 1865 census
- Fraction of the population given 2 or more names (by year of birth)

Research on Italy shows that children with more names had lower mortality rates



CLASSIFYING QUALITATIVE INFO INTO CATEGORIES

SEX

Child abandonment

- Foundling hospital
- 1880/90
- 780 registers

- SR = 96 boys per hundred girls
- SR = 86
(if we only count those left being <1-2 days old)

	Name	Age	Origin	Date	Type	Clothes	Other
--	------	-----	--------	------	------	---------	-------

NÚM. DE ÓRDEN.	NOMBRES.	EDAD.	NATURALEZA Ó PROCEDENCIA.	FECHA DEL INGRESO AÑO. MES. DÍA	FORMA DEL MISMO CON EL NOMBRE DE LA PERSONA QUE LE HA ENTREGADO SI CONSTA	ROPAS CON QUE HA INGRESADO.	SEÑAS PARTICULARES.
3762	Alejandra Camacho Caballero Iberia	Isabel	Recién nacido	1880 Febrero 22	Recién nacido	Camisa blanca - pantalón negro Vestido blanco - 1 pañuelo pañuelo - sombrero blanco - 1 pañuelo blanco - 2 pañuelos blancos sombrero blanco - 1 abanico de raso	"
3763	José Chádres Robles	Isabel	Gabrielle	1880 Marzo 1	Recién nacido	2 pañuelos blancos - 1 pañuelo blanco - pañuelo blanco pañuelo blanco - 1 pañuelo blanco pañuelo blanco - 2 pañuelos blancos 2 pañuelos blancos - 1 pañuelo blanco sombrero blanco	"
3764	Eulalia Román y Asensio	Iberia	Leticia	1880 Marzo 11	Recién nacido en la Maternidad	"	"

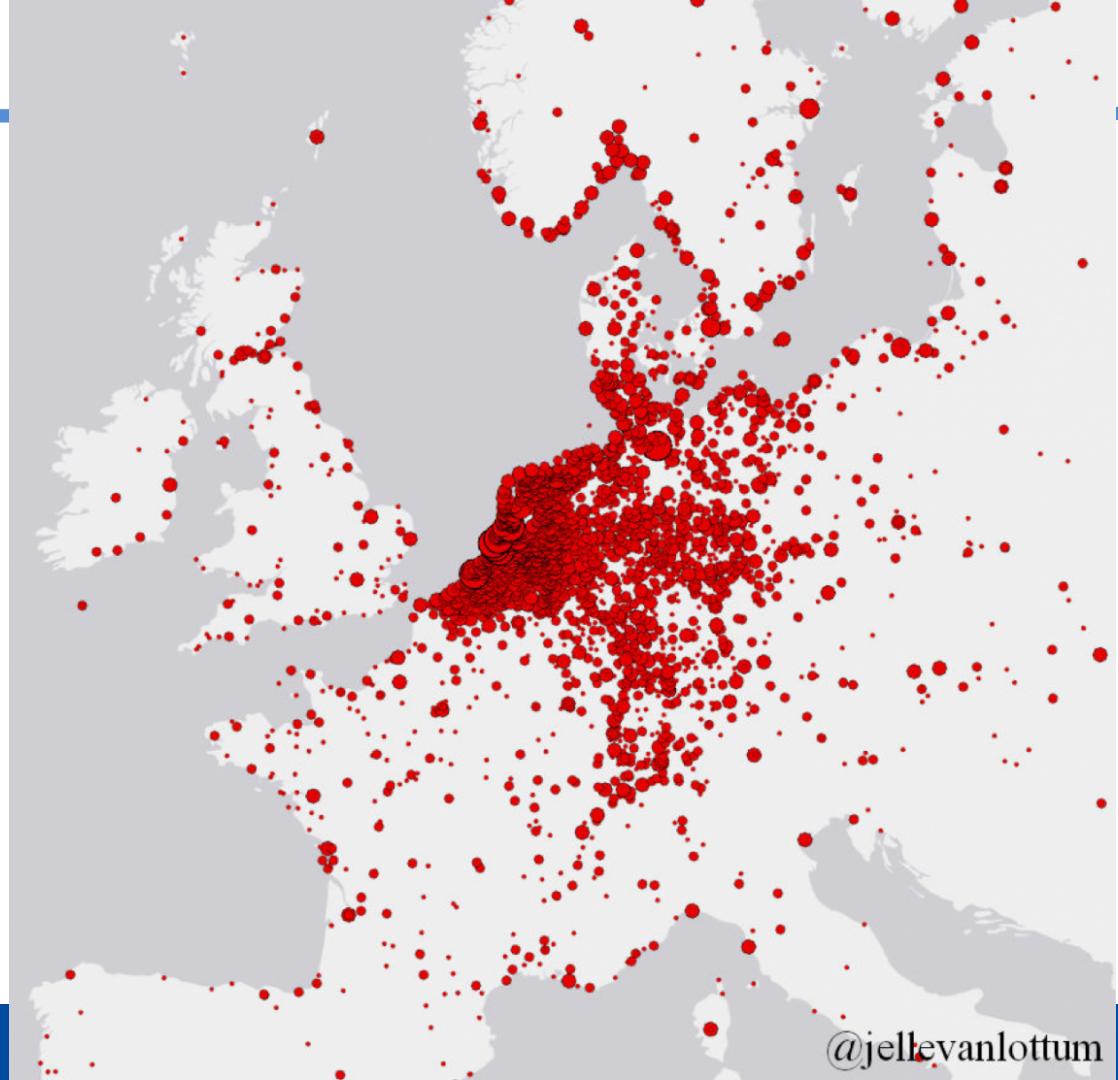
Where the Dutch East India Company crews came from?

- Pay ledgers (1602-1795)
- 774,200 records (from almost 5,000 ships)
- Personal info, salaries...

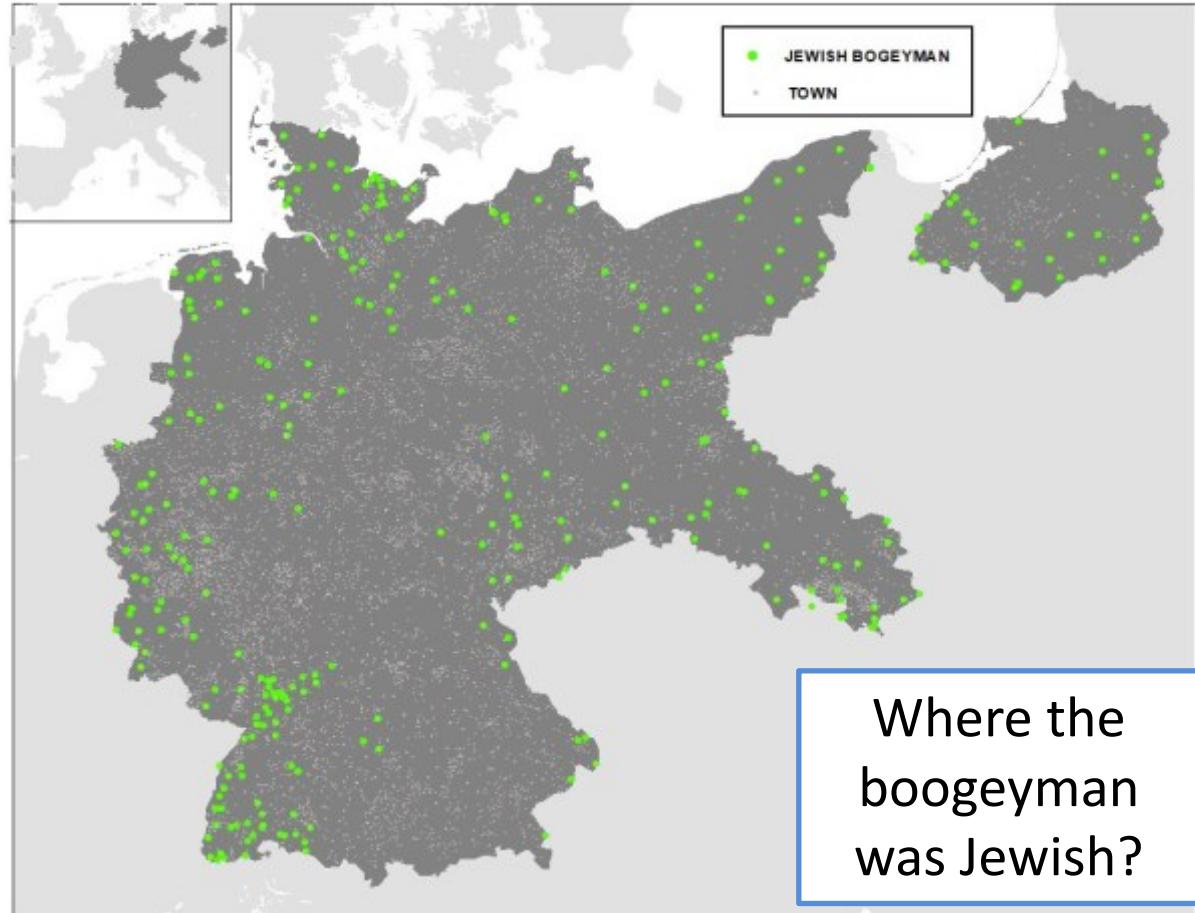
LOCATIONS I

- ❑ Where the Dutch East India Company crews came from?

More international than expected



- Children stories and antisemitism
 - Collection of German folktales: Atlas Der Deutsche Volkskunde (ADV)
 - Based on a survey on experts living in over 20,000 localities in 1930



QUALITATIVE CATEGORIES AS NUMBERS

CENSUSES

□ 1865 Norwegian Population Census

Specialliste over Folketallet m. m. i				Tøle		Stoeddistrikt	
1.	2.	3.	4.	5.	6.	7.	
Gaardens eller Stedets	Boldest. Sif.	Boldest. Sif. Gudsbohusint.	Personernes Navne (Fornavn og Bivavn).	Gud. Enkeh. er i Familien, saafom Huslader, Søn, Son, Datter, Foralder, Ejendeselvende eller Logerende, jamt Enkvens Stand efter Næringssoc.	Ugift, Gift, Ulemand, Kone eller Brætke (som sædanne an- føres brætke med Hensyn til Verd. og Sang).	Moder, det løbende År var bærtne	
Navn.	Wa- terful Lege- No.					Mand- Hjon.	Ma- dr. Hj.
Hagen			Karen Lærddatter	Jørgen Lærde	bukke		50
	11		Sonata Eriksen	Husmænd og jærl	gaff.	59	
			Povel Pidarddatter	hans Korn	do		5
			Johanna Lærddatter	dans Døller	ugift		2
Aanes	14		Andrea Lærddatter	Husmænd og ke mæ	Eru		4
			Marii Olsdatter	hendes Døller	ugift		1
			Povel Olsdatter	hendes Døller	do		1
			Karen Olsdatter	hendes Døller	do		2
			Pauline Olsdatter	hendes Døller	do		3
			Peder Halsdækkar	Hans ejendeselvende og jærl	gaff.	71	
			Marii Olsdatter	Konte, hennemod Gudsbohusint.	do		6
Karlens	35m	11	Ole Olsen	Husmænd og Gudsbohusint.	do	40	
			Eli Olsdatter	hans Korn	do		4
			Ole Olsen	dans Søn	ugift	12	
			Ole Olsen	dans Søn	do	9	

Municipality, region, name, sex, age, occupation, place of birth, marital status, number of children, religious affiliation, ethnicity, etc.

- Fertility: Number of children per married women aged 30-35
- Early marriage: Percentage of women married at age 16-20
- Family type: Prevalence of extended kin living in the same household
- Celibacy: Probability of being single (aged 45-50)
- Economic structure: Percentage of men employed in manufacturing
- Education: Probability of being able to read/write (aged 8-10)
- ...

→ Comparisons

- Across families' socio-economic status (father's occupation)
- By sex
- Across locations (municipalities, regions)
- Over time (different censuses)

CUALITATIVE CATEGORIES AS NUMBERS

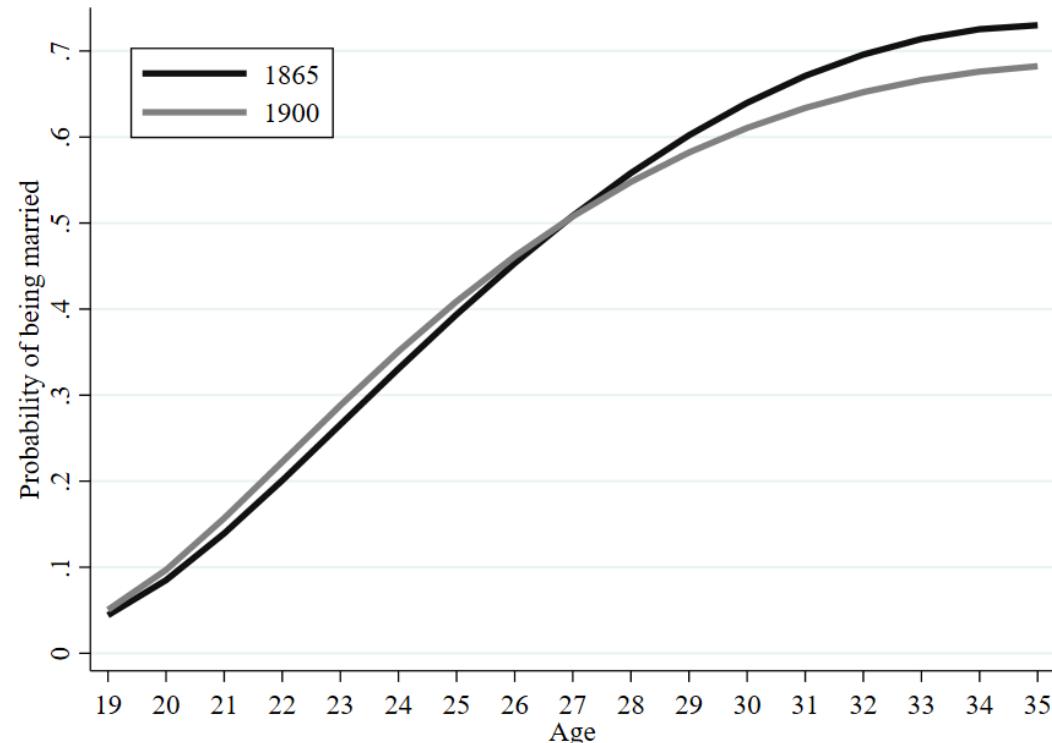
CENSUSES

Marital Status in 1865 (>15 years old)

	Obs.	%
Single	394,805	37.69
Married	544,281	51.97
Separated/divorced	85	0.01
Widow	92,006	8.78
Unknown	16,198	1.55
Total	1,047,375	100

Each individual categorised as
1 (married) or 0 (otherwise)

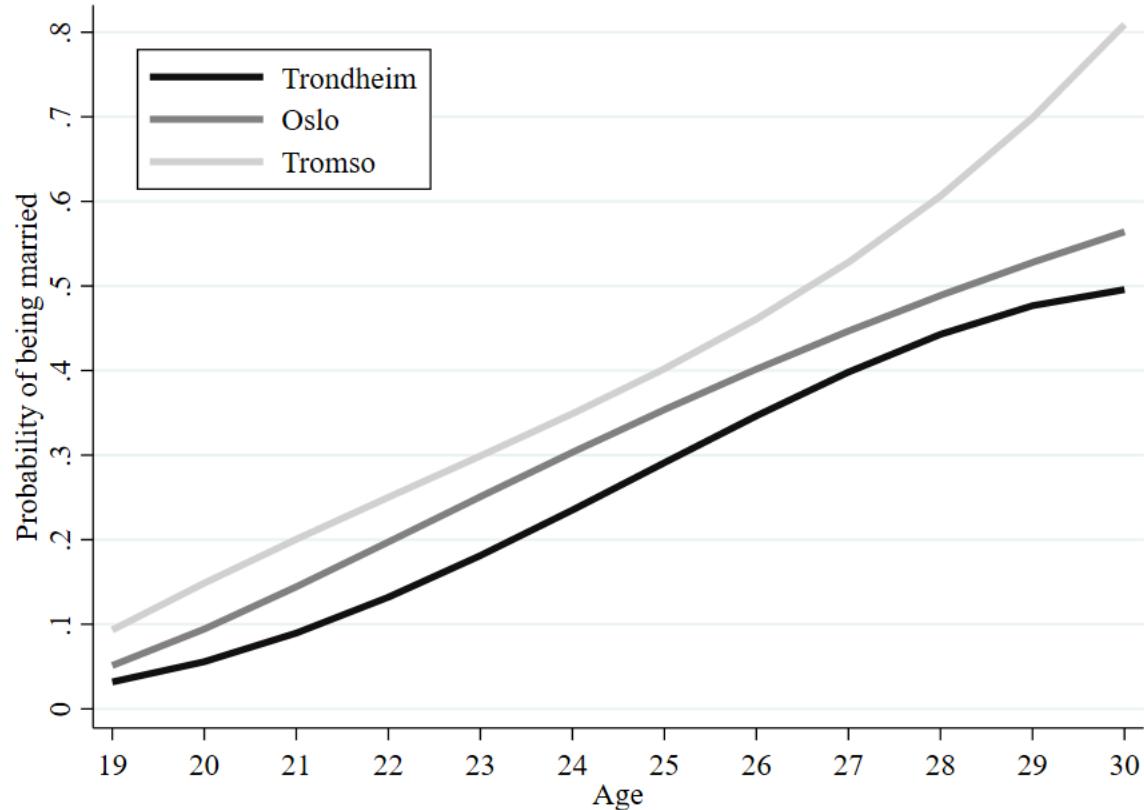
Probability of being married (women)



Probability of being married:

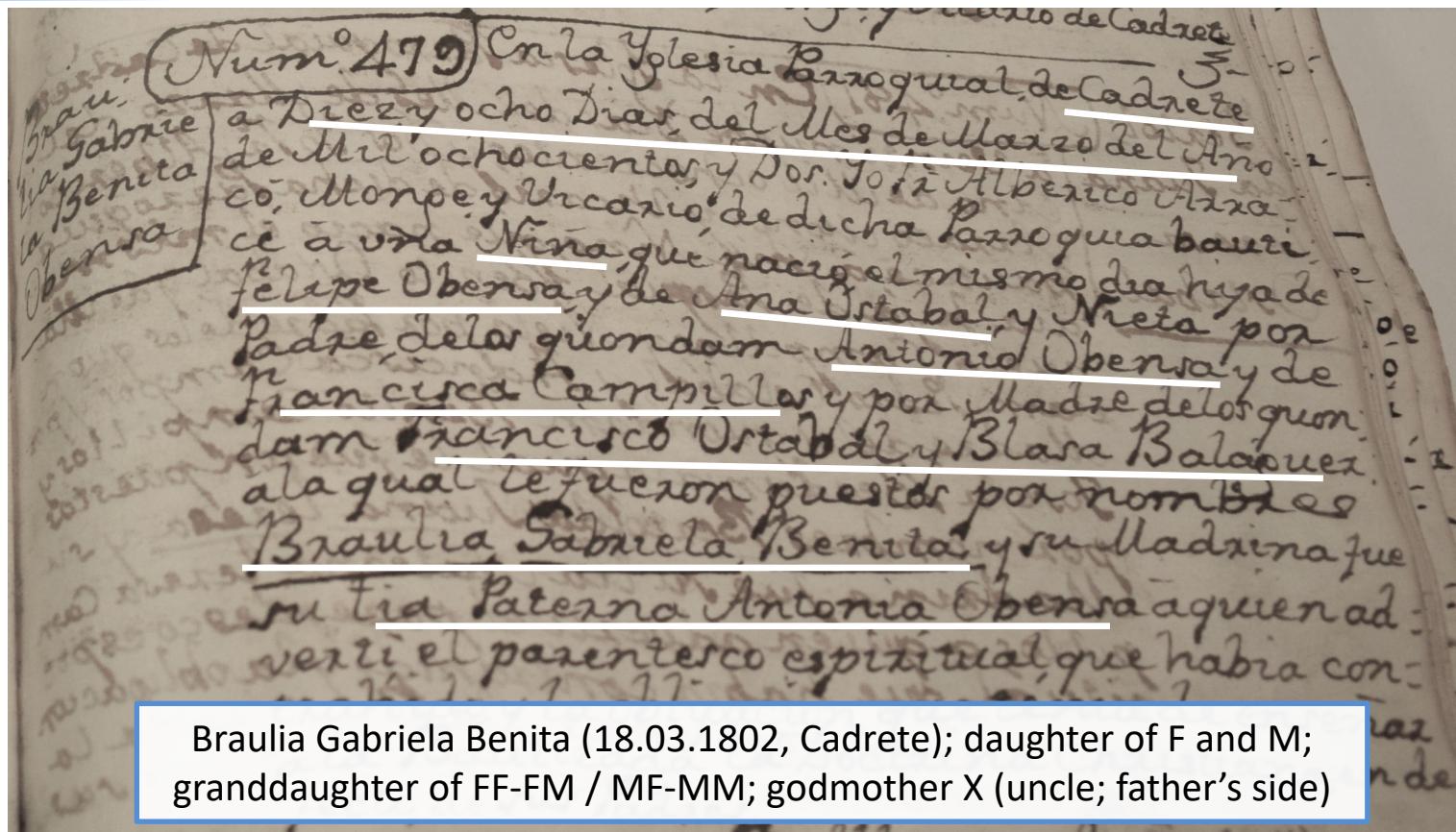
-- Women in 1865

- Trondheim
- Oslo
- Tromso



Baptisms

- Name
- Sex
- Village
- Date
- Legitimate
- Parents' name
- Grandparents
- Godparents
(relationship)

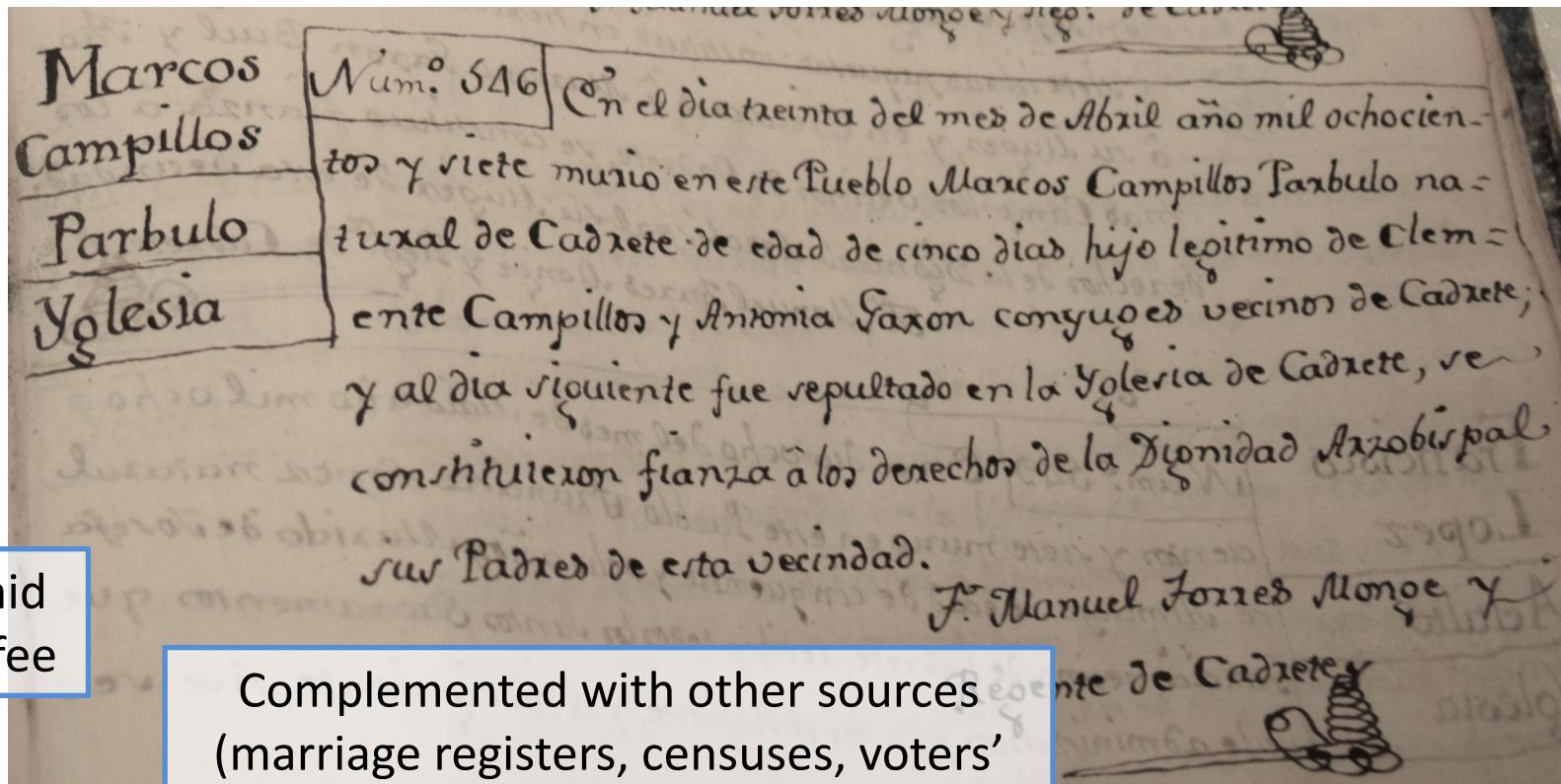


Braulia Gabriela Benita (18.03.1802, Cadrete); daughter of F and M;
granddaughter of FF-FM / MF-MM; godmother X (uncle; father's side)

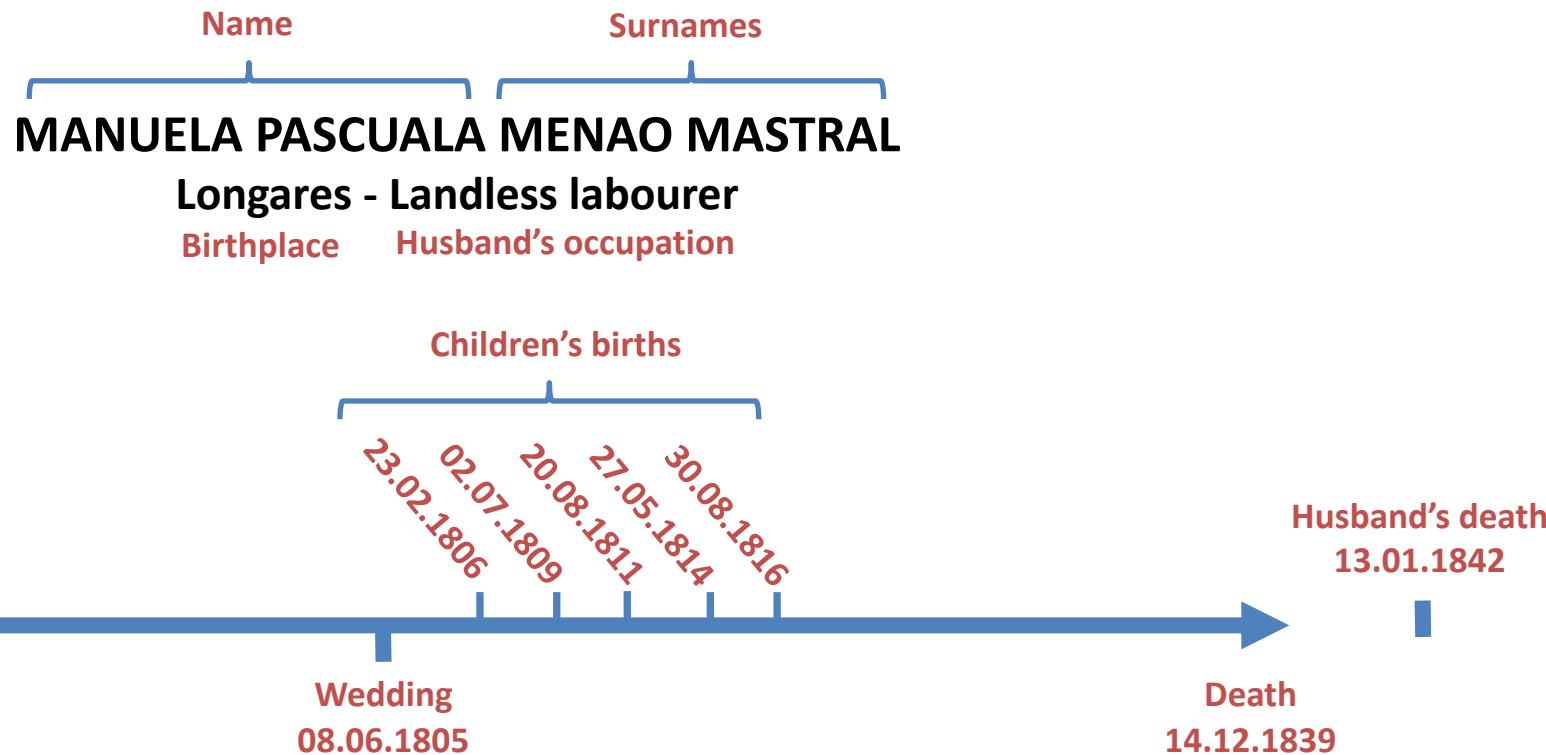
Deaths

- Name
- Date
- Village
- Age of death
- Parents

And who paid the funeral fee



Complemented with other sources
(marriage registers, censuses, voters' registers, ...): occupation, literacy...



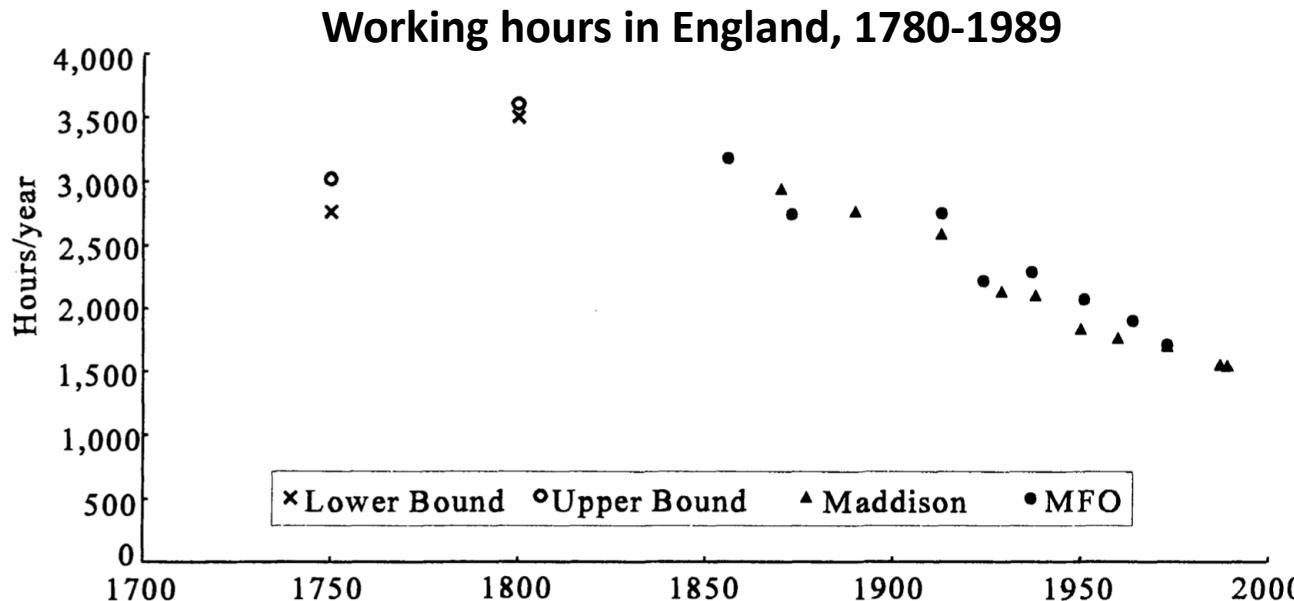
- Age at marriage
- Age at death
- Number of children
- Intervals between pregnancies (fertility control?)
- Probability of death according to:
 - Mother's age, number and sex-composition of siblings, death of mother/father, etc.
- ...

→ Comparisons

- Across families' socio-economic status (father's occupation)
- By sex
- Across locations (municipalities, regions)
- Over time (different censuses)

□ Voth (1998): Time and work in eighteenth-century London

- Relying on 7,650 court cases, it analyses what witnesses were doing, and when, at the time of the crime (over 2,000 observations)



He can estimate when they were working and when they weren't and compute working hours

- Impact of the Spanish inquisition

- More than 67,000 trials between 1540 and 1700
- Intensity measured as the proportion of the years during which a trial was initiated against at least one member of the community

