

□

ANÀlisi DE DADES ESPACIALS EN L'ÀMBIT DE L'EPIDEMIOLOGIA

Prof. Dr. Maria A Barceló i Prof. Dr. Marc Saez

8, 10, 14 i 16 de setembre de 2021

Grup de Recerca en Estadística, Econometria i Salut (GRECS), Universitat de Girona
CIBER d'Epidemiologia i Salut Pública(CIBERESP)

INTRODUCCIÓ AL CURS

1. Introducció al curs
2. **Introducció a l'epidemiologia i l'estadística espacial**
3. Panoràmica del models mixtos
4. Panoràmica del models mixtos - Pràctiques
5. Introducció a INLA i R INLA
6. R INLA - Pràctiques

Dimecres 8

Divendres 10

INTRODUCCIÓ AL CURS

- 7. Mapes de malalties. Estandardització de raons d'incidència i mortalitat
- 8. Mapes de malalties. Suavització de raons d'incidència i de mortalitat estandarditzades
- 9. Mapes de malalties – Pràctiques
- 10. Estudis d'associació geogràfica. Regressió ecològica espacial
- 11. Regressió ecològica espacial - Pràctiques

Dimarts 14

INTRODUCCIÓ AL CURS

- 12. Agrupació de casos
- 13. Extensions: BYM2, processos puntuals, leaflet, pc priors
- 14. Extensions – Pràctiques

} Dijous 16

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

Epidemiologia

- L'**epidemiologia** es dedica a l'estudi de la distribució, freqüència, causes i control dels factors relacionats amb la salut i la malaltia en poblacions humanes ben definides i a l'aplicació d'aquest estudi per defensar i millorar la salut de la població.
- L'**epidemiologia** estudia, sobretot, la relació entre exposició i malaltia.
- L'**epidemiologia** es considera la ciència bàsica per a la medicina preventiva i una font d'informació per a la formulació de polítiques de salut pública.

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

- Les malalties no es produeixen de forma aleatòria. Tenen causes, moltes d'elles d'origen humà, que poden evitar-se.
- Els mètodes epidemiològics han estat crucials per identificar nombrosos factors etiològics que, al seu moment, han justificat la formulació de polítiques sanitàries encaminades a la prevenció de malalties.

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

- L'epidemiologia va sorgir de l'estudi de les epidèmies de les malalties infeccioses.
- En l'actualitat, l'epidemiologia s'interesa per l'estudi demogràfic de qualsevol malaltia amb l'ajuda de l'estadística.

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

Epidemiologia espacial

“Spatial epidemiology is the description and analysis of geographic variations in disease with respect to demographic, environmental, behavioural, socioeconomic, genetic, and infectious risk factors”.

Elliot P, Wartenberg D. Spatial epidemiology: current approaches and futures challenges. *Environ. Health Perspect.* 2004; 112(9):998-1006. doi: 10.1289/ehp.6735.

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

- **L'epidemiologia espacial** està relacionada amb la descripció de les malalties i l'estudi de les causes i prevenció de les mateixes, utilitzant diferents perspectives d'anàlisi en què la localització dels esdeveniments és una component fonamental, ja que estudia també les variacions geogràfiques de les malalties.
- **L'epidemiologia espacial** és un subcamp de l'epidemiologia centrat en l'estudi de la distribució espacial dels resultats en salut.
- **L'epidemiologia espacial** es basa en un concepte de salut, en el qual els individus són vistos en el seu context sociocultural-ambiental.

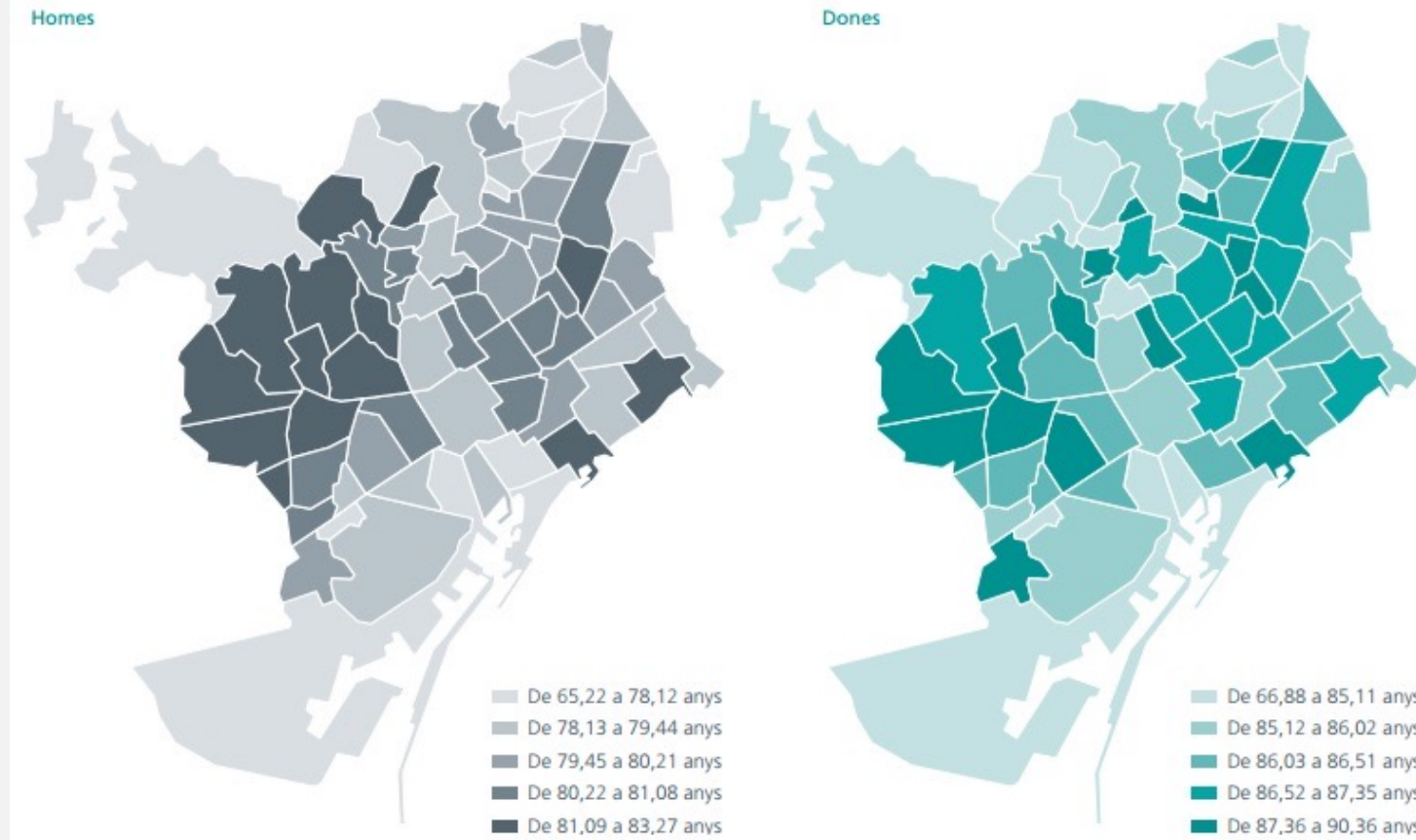
EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

Principal objectiu de l'epidemiologia espacial:

- Mostrar quina part de la variació espacial de la distribució de l'ocurrència d'un esdeveniment de salut no està explicada ni per la distribució espacial de factors explicatius coneguts, ni per una variació aleatòria.
- De fet, sovint ens interessa trobar pistes sobre algun factor de risc desconegut d'una determinada malaltia.

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

Figura 1. Esperança de vida en homes i dones als barris. Barcelona, 2008-2012.



Font: Registre de Mortalitat de Barcelona. Agència de Salut Pública de Barcelona.

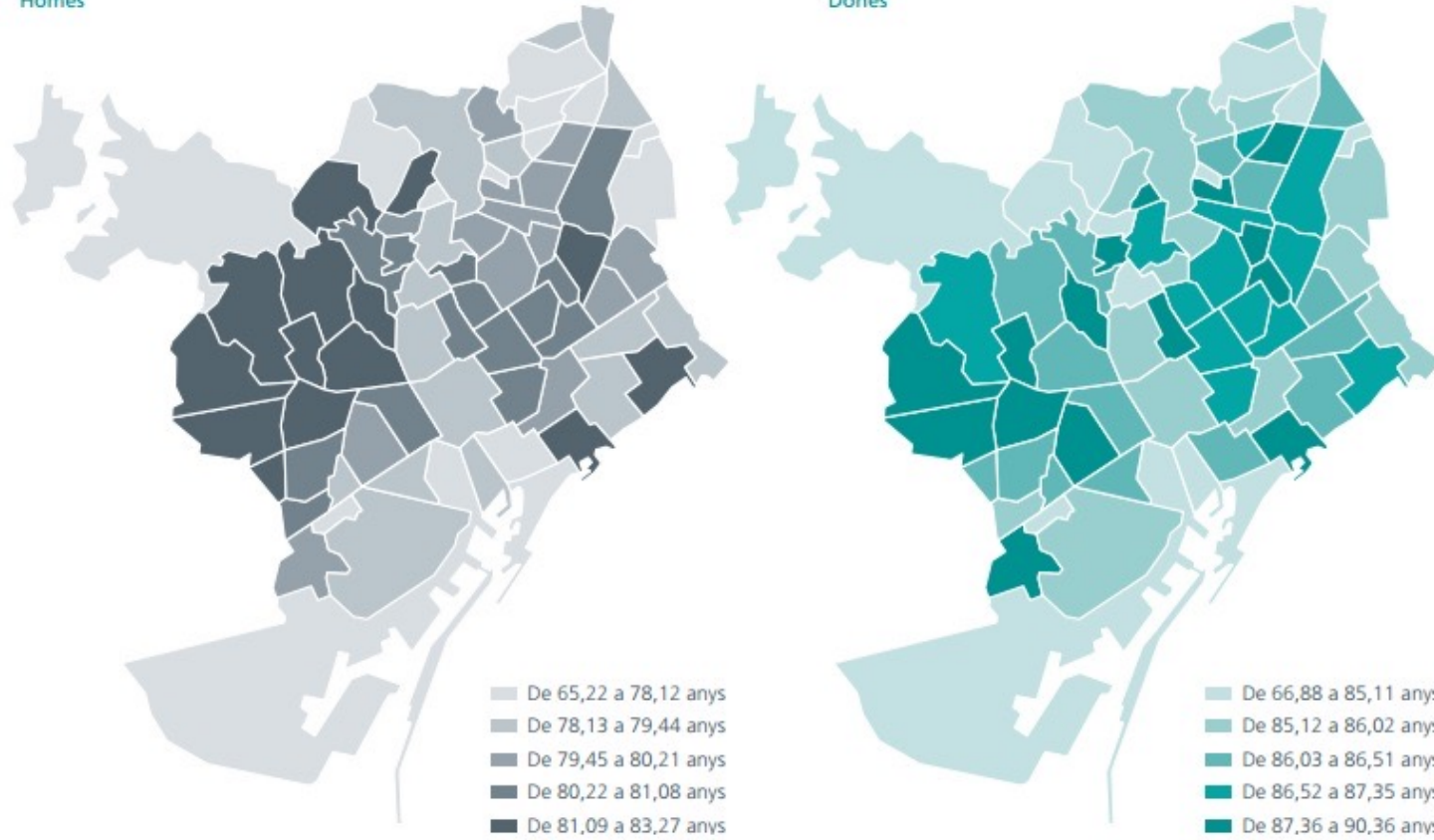
2. Introducció a l'epidemiologia i l'estadística espacial

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

Figura 1. Esperança de vida en homes i dones als barris. Barcelona, 2008-2012.

Homes

Dones



Font: Registre de Mortalitat de Barcelona. Agència de Salut Pública de Barcelona.

Mapa Guia Barris



2. Introducció a l'epidemiologia i l'estadística espacial

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

- L'estudi de la distribució geogràfica dels esdeveniments de salut ha estat de gran interès entre els epidemiòlegs durant les últimes dècades, encara que com veurem a continuació, té una història de més de 200 anys.
- Els **primers mapes de malalties** es van fer per representar la ubicació de casos de malalties infeccioses:
 - Febre groga a Nova York (Seaman, 1798)
 - Còlera a Londres (Snow, 1854)

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

Mapa de Filippo Arrieta

Visualització de l'estratègia per contenir la propagació de la pesta en la regió de Bari, Itàlia 1690-92.

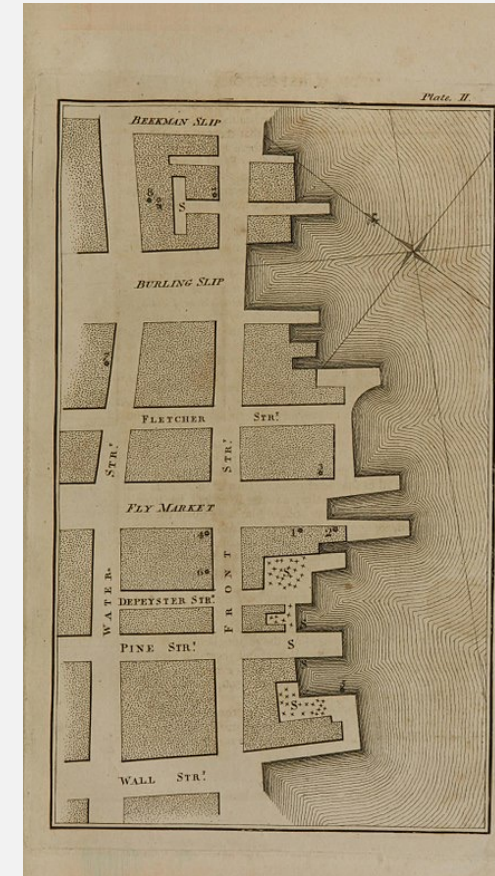


2. Introducció a l'epidemiologia i l'estadística espacial

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

Mapa de Seaman de casos de febre groga

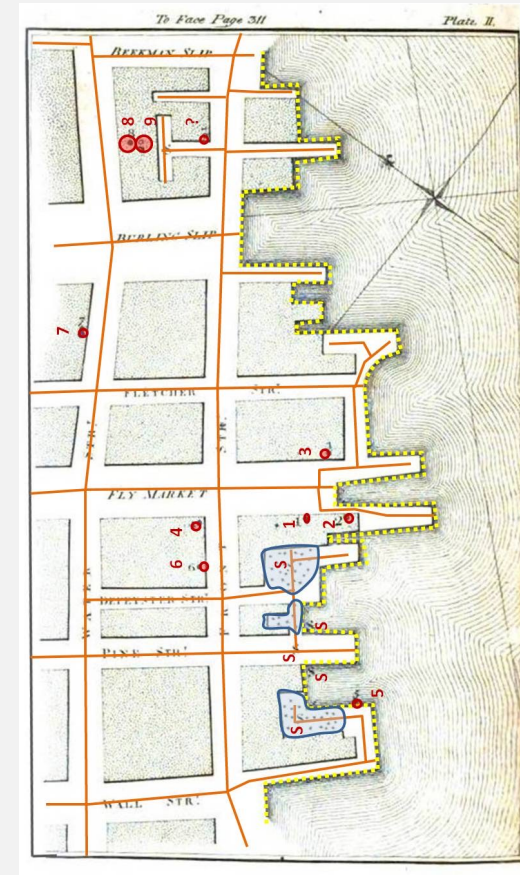
Els punts representen les morts per febre groga i les "S", els abocadors.



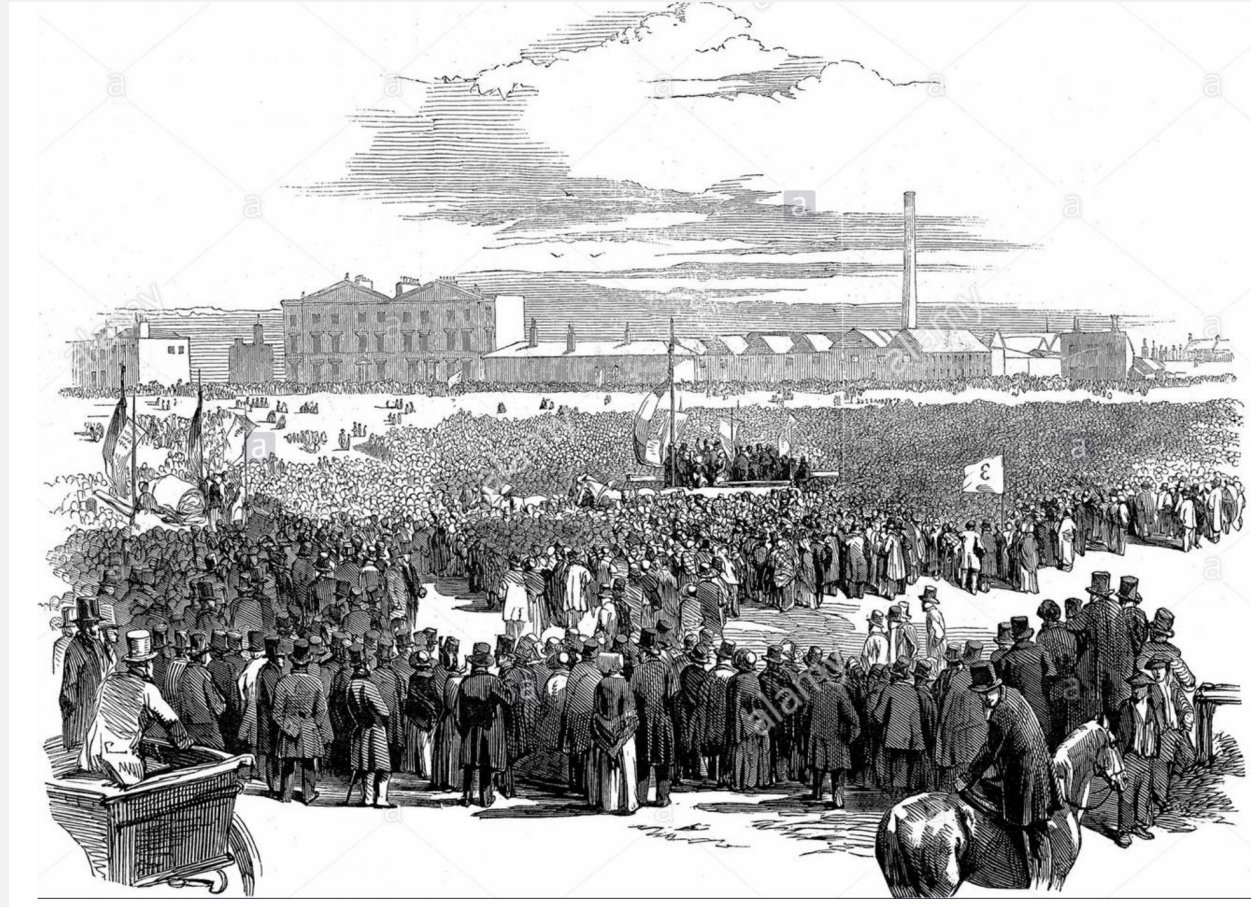
EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

Mapa de Seaman de casos de febre groga

Els punts representen les morts per febre groga i les creus o les "S", els abocadors.



EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL



2. Introducció a l'epidemiologia i l'estadística espacial

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

- Durant la tardor de 1848, es va produir una segona epidèmia de còlera a Anglaterra, causant una gran mortalitat.
- No es coneixia amb certesa l'etiologia, ni la manera de transmissió d'aquesta malaltia.
- Dos corrents teòrics:
 - Contagi per contacte
 - Miasmes

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

Snow va observar:

- **Miasmes no podien causar la malaltia:** els pacients haurien de presentar símptomes respiratoris producte de la inhalació dels “miasmes” i no la síndrome diarreic agut present al còlera.
- **Defuncions per còlera** ocorregudes entre 1848-49 (segona epidèmia) estaven **concentrades** en els districtes de la **zona sud de Londres**. Les taxes de mortalitat observades en aquesta zona eren molt superiors a la de la resta de la ciutat (8,0 i 2,4 defuncions per 1.000 habitants, respectivament).

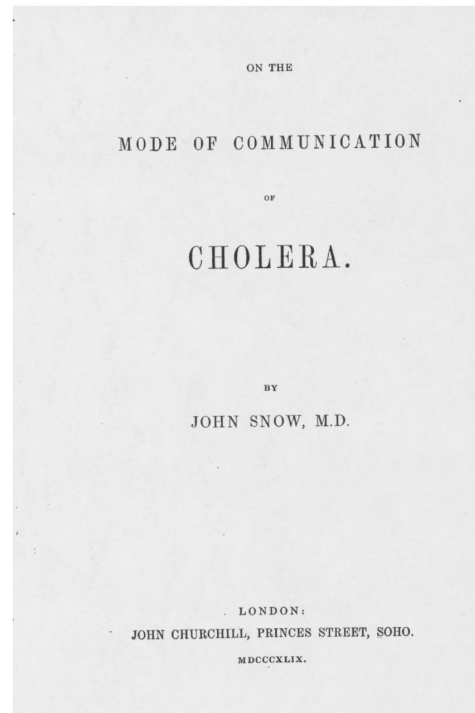
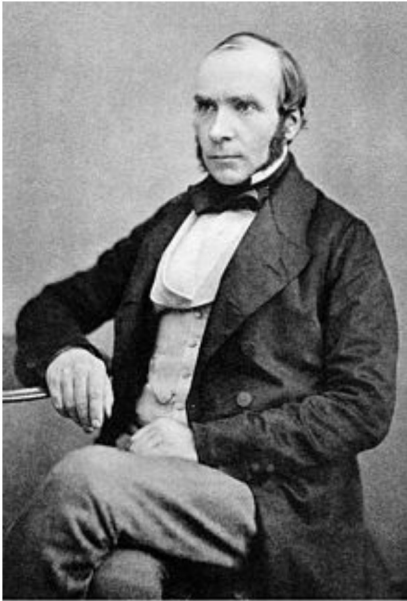
EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

A més, Snow va observar que:

- Els habitants de la zona sud de Londres obtenien aigua per beure riu avall del Tàmesi (aigües altament contaminades).
- Les altres zones de Londres obtenien l'aigua per beure de sectors menys contaminats del riu (aigües amunt del Tàmesi o d'afluents).

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

John Snow
1813-1858



Deaths from Cholera in London, registered from September 23d, 1848, to August 25th, 1849.

Districts of London.	Population in 1841.	Deaths from Cholera.	Deaths to each 1,000 inhabitants.
West . .	300,711	533	1·77
North . .	375,971	415	1·10
Central . .	373,605	920	2·48
East . . .	392,444	1,597	4·06
South . .	502,548	4,001	7·95
Total . .	1,948,369	7,466	3·83

On the Mode of Communication of Cholera. John Snow. 1849

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

- En 1853-1854, es dóna la tercera epidèmia de còlera a Londres.
- Els habitants de certs districtes del sud de la ciutat extreien l'aigua de petits afluents del riu Tàmesi o bé a partir de nombroses bombes d'aigüa d'ús públic.
- Les deixalles humanes eren abocades en improvisades clavagueres o directament al riu.

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

- Hi havia dues companyies responsables de l'abastament: Sothwark and Vauxhall Company i Lambeth Water Company.
- Durant la segona epidèmia de còlera de 1848-49, ambdues companyies extreien l'aigua de sectors contaminats del Tàmesi, presentant similar nombre de morts els districtes abastits per una i altra companyia.

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

- En 1853:
 - Lambeth Water Company havia traslladat les seves instal·lacions riu amunt (aigües no contaminades).
 - Southwark and Vauxhall Water Company seguia aigües avall.
- Snow va observar que la taxa de mortalitat per còlera a llars abastits per la segona companyia era 8,5 vegades més gran que la de llars abastits per la primera.

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

- A principis de setembre de 1854:
 - En el sector de Londres “Golden Square” (al Soho), es va produir un brot epidèmic de còlera d’inusual intensitat (500 morts en tan sols 10 dies).
 - La majoria dels residents del sector extreien l’aigua a partir d’una bomba d’ús públic situada a Broad Street.

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL



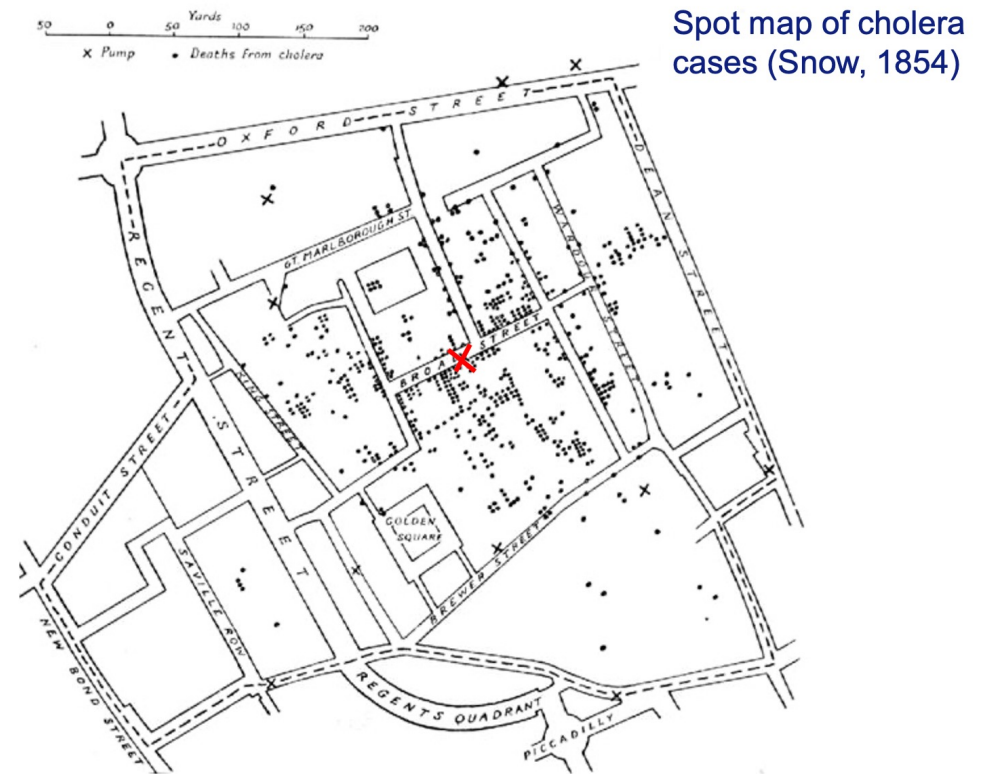
2. Introducció a l'epidemiologia i l'estadística espacial

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

- Va realitzar un **mapa del sector**, marcant els punts corresponents a defuncions per còlera i les diferents bombes d'aigua potable existents.
- Va plantejar que aquest brot es devia a la ingestió d'aigües contaminades procedents de la bomba de Broad Street.

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

Snow va identificar l'agregació espacial dels casos de còlera a Londres en 1857.



EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

- Va agafar mostres de la bomba de Broad Street i altres 4 bombes properes (hi havia diferències en relació amb la claredat de l'aigua).
- Va calcular la distància entre la residència de cada difunt i la bomba d'aigua més propera.
- Va observar que en 73 de 83 casos la bomba de Broad Street era la més propera i va comunicar això a les autoritats sanitàries i aquestes van tancar la bomba.

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

Evolució dels mapes en epidemiologia espacial

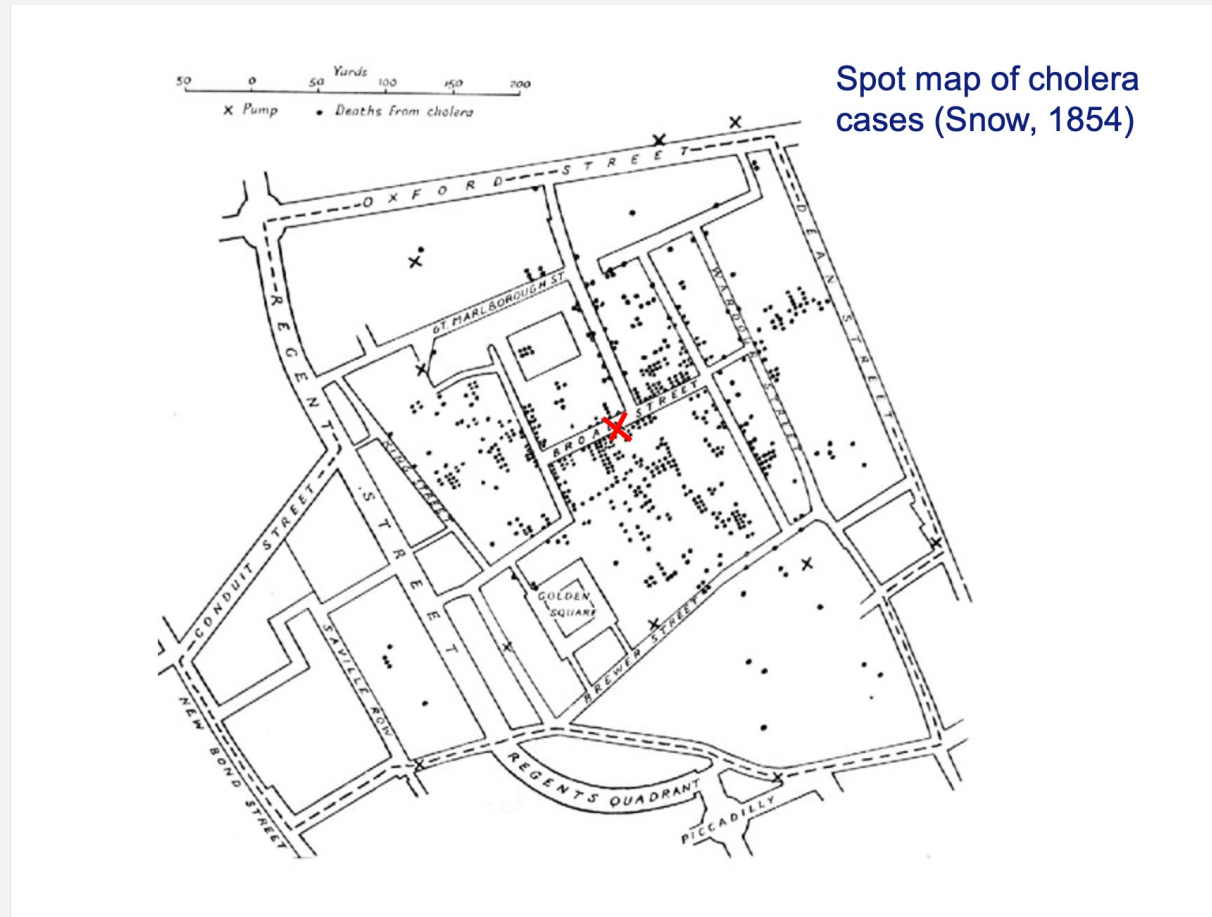
- Mapes de punts
- Mapes de coropletes
- Atles de malalties, a nivell nacional e internacional

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

Evolució dels mapes en epidemiologia espacial

- Mapes de punts
 - Febre groga a Nova York (Seaman, 1798)
 - Còlera a Londres (Snow, 1854)

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL



2. Introducció a l'epidemiologia i l'estadística espacial

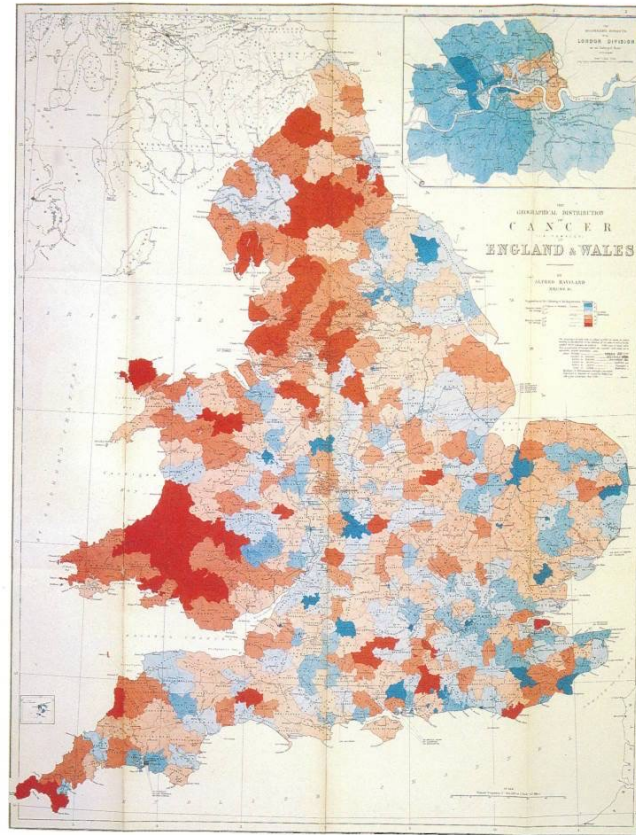
EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

Evolució dels mapes en epidemiologia espacial

- Mapes de coropleta
 - Distribució geogràfica de la mortalitat per malaltia cardíaca, càncer i tuberculosi a Anglaterra i Gal·les (Haviland, 1878)
 - Mortalitat per càncer per comtat a Anglaterra i Gal·les, ajustada per edat i sexe (Stocks, 1936,1937,1939)

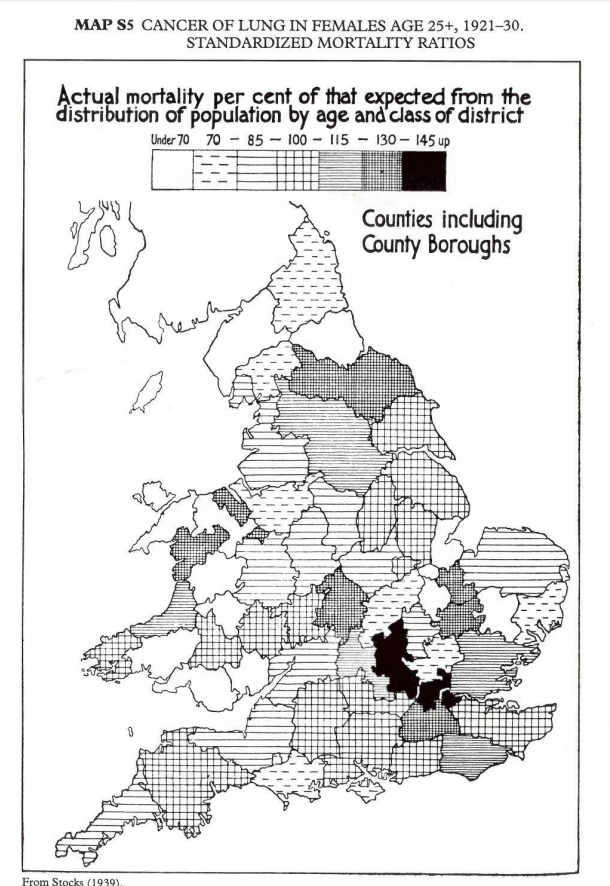
EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

Càncer en dones,
1851-1860
(Haviland, 1878)



The geographical distribution of cancer in females in England and Wales, 1851-60. From Haviland (1878).

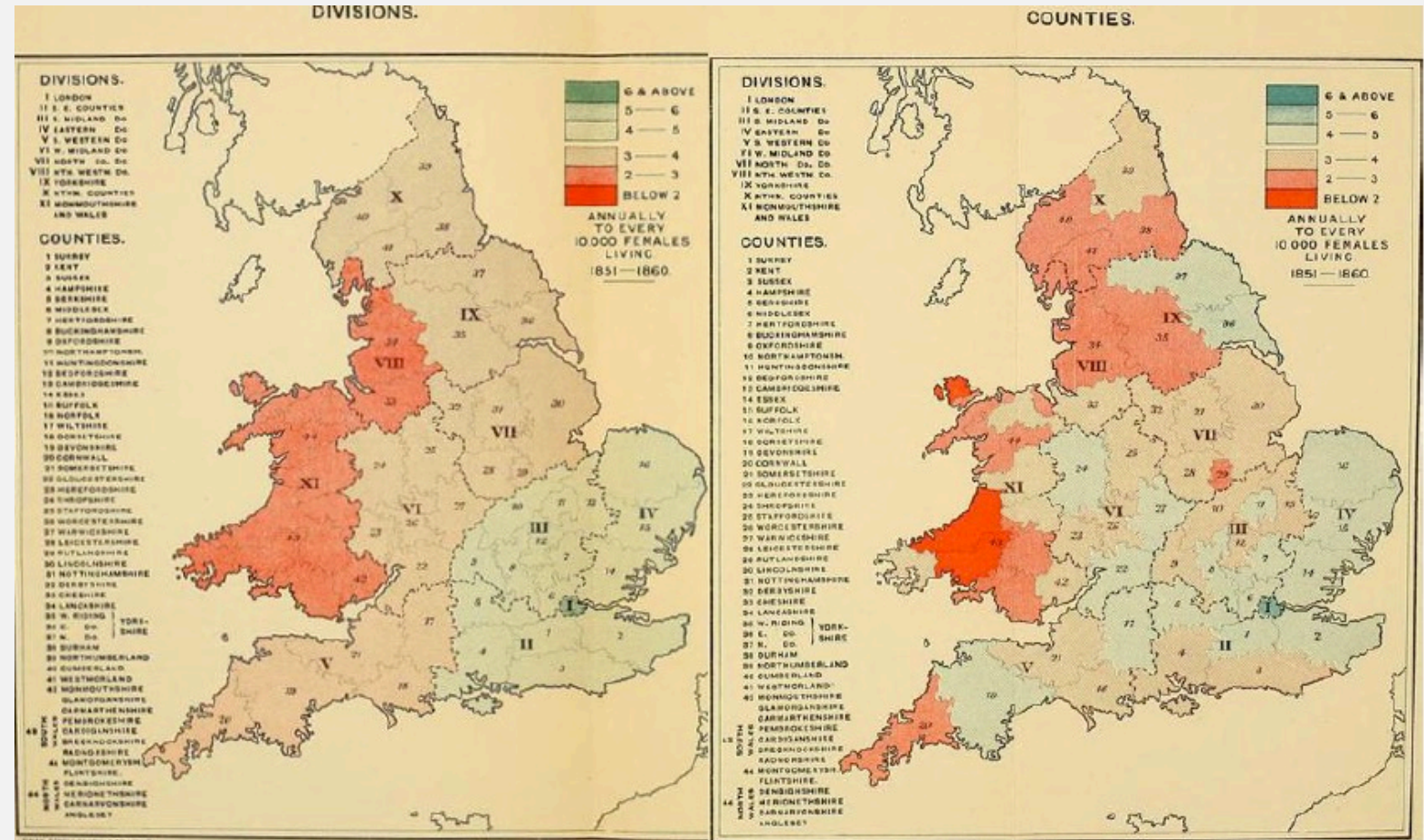
Càncer de pulmó en dones,
1921-1930
(Stocks, 1939)



2. Introducció a l'epidemiologia i l'estadística espacial

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

Càncer en dones, 1851-1860
(Haviland, 1878)



2. Introducció a l'epidemiologia i l'estadística espacial

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

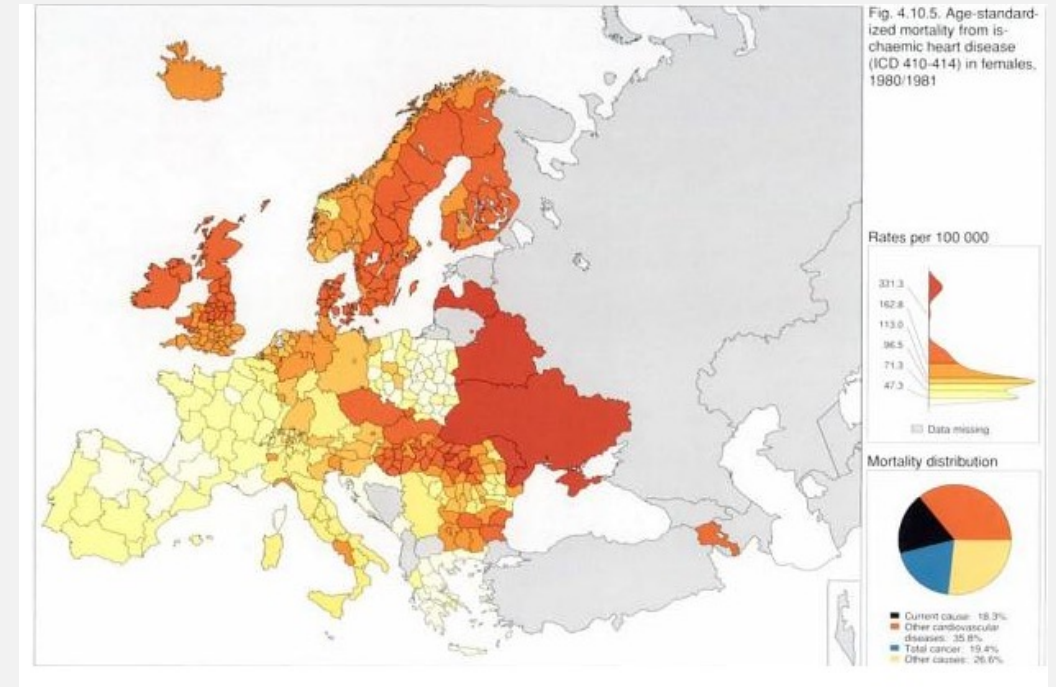
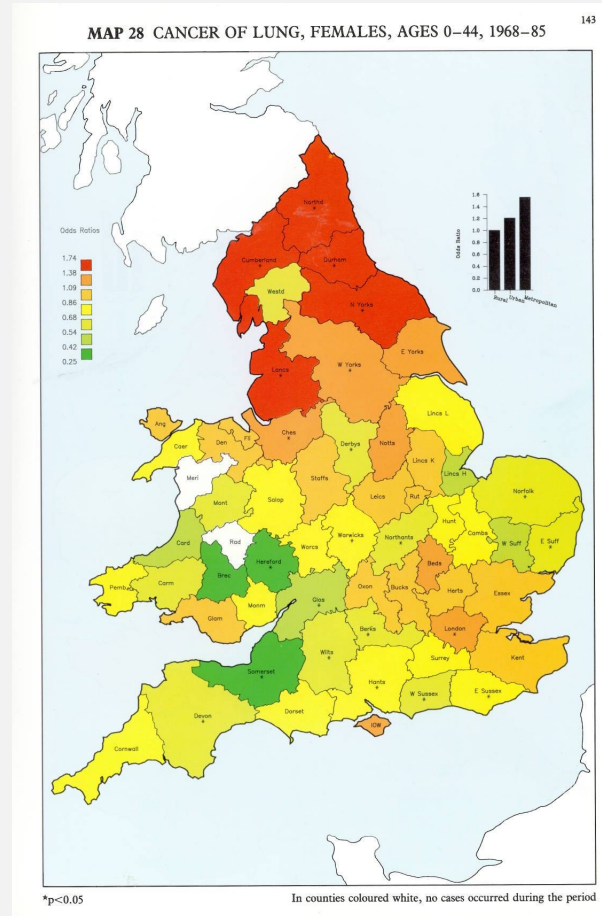
Evolució dels mapes en epidemiologia espacial

- Atles de malalties, a nivell nacional e internacional
 - Atles d'incidència del càncer a Anglaterra i Gal·les 1968-85 (Swerdlow i dos Santos Silva, 1993)
 - Atles de mortalitat a Europa 1980-81 i 1990-91 (OMS, 1997)

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

Incidència de càncer de pulmó
en dones 1968-1985

(Swerdlow and dos Santos
Silva, 1993)



Mortalitat per Infart de miocardi
estandarditzada per edat, 1980-81

(OMS)

2. Introducció a l'epidemiologia i l'estadística espacial

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

- Com podem veure, l'epidemiologia espacial va partir de la representació de la distribució espacial d'esdeveniments en salut amb el doble objectiu de caracteritzar la seva extensió i establir hipòtesis de les possibles causes.
- Posteriorment, l'epidemiologia espacial va créixer molt en complexitat (mètodes d'anàlisi, unitats petites) i utilitat.
- La confluència entre l'epidemiologia, l'estadística i la informàtica, juntament amb el gran desenvolupament tecnològic, han permès que l'epidemiologia espacial es desenvolupi a passos de gegant.

2. Introducció a l'epidemiologia i l'estadística espacial

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

Desenvolupaments a partir de 1996

- Sistemes d'Informació Geogràfica (SIG)
- Software per representar mapes i analitzar dades espacials
- Major disponibilitat de dades georeferenciades (GPS, etc.)

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

Desenvolupaments a partir de 1996

- Desenvolupament de mètodes estadístics especialitzats
 - Tècniques sofisticades per separar el senyal de soroll
 - Mètodes per controlar la dependència espacial i temporal
 - Mètodes per detectar aglomeracions de casos (clústers)

EPIDEMIOLOGIA I EPIDEMIOLOGIA ESPACIAL

Problemes en epidemiologia espacial:

- Diferents escales (per exemple, comunitats autònomes, província, municipi, districte, barri, codi postal, secció censal, ...)
- Canvis en els límits d'algunes d'aquestes unitats
- Poden existir errors en la georeferenciació (a causa de direccions errònies o inexistents, etc.)
- Desalineació (misalignment)

ESTADÍSTICA ESPACIAL

- **L'estadística espacial** s'ocupa de l'exploració, descripció, visualització i anàlisi de les dades, considerant les seves característiques de distribució en l'espai, que solen expressar-se a través de l'ús de coordenades geogràfiques.
- **L'estadística espacial** és la branca de l'estadística que analitza dades georeferenciades, és a dir, dades de què es disposa de les seves coordenades espacials (dades espacials).

ESTADÍSTICA ESPACIAL

- Per **dades espacials** s'entenen aquelles mesures o observacions realitzades en localitzacions o en àrees específiques. A més del valor del mesurament, incorporen la localització/posició dels valors observats.

ESTADÍSTICA ESPACIAL

Característiques de les dades espacials:

1. Heterogeneïtat espacial

- Les observacions no son homogènies en l'espai

2. Dependència espacial

- Les observacions d'una localització depenen d'altres observacions en altres localitzacions (generalment pròximes).

ESTADÍSTICA ESPACIAL

Característiques de les dades espacials:

1. Heterogeneïtat espacial

- Les observacions no son homogènies en l'espai

ESTADÍSTICA ESPACIAL

Hi ha dos tipus d'**heterogeneïtat espacial**:

➤ ***Heteroscedasticitat:***

- *Causa estructural*: a causa de la utilització de dades procedents d'unitats espacials arbitràries
- *Causa mostral*: existència d'outliers, omissió de variables rellevants, errors de mesura i altres errors d'especificació.

➤ ***Canvi-inestabilitat estructural***: els individus no són homogenis en l'espai (per exemple, nord-sud; centre-perifèria, ...).

ESTADÍSTICA ESPACIAL

Característiques de les dades espacials:

2. Dependència espacial

- Les observacions d'una localització depenen d'altres observacions en altres localitzacions (generalment pròximes).

ESTADÍSTICA ESPACIAL

Existeixen dos tipus de **dependència espacial**:

- ***Interacció espacial (dependència espacial substantiva)***: és un efecte spillover d'un individu sobre un altre individu.
- ***Autocorrelació en els residus (dependència espacial residual)***: causada per errors de mesura espacialment relacionats.

ESTADÍSTICA ESPACIAL

Tipus dades espacials:

Les dades espacials s'han classificat històricament en tres grans grups (Cressie, 1993):

1. Dades en graella o **dades d'àrea** ('lattice data' o 'areal data')
2. Dades referenciades puntualment: **Processos puntuals** ('point processes')
3. Dades referenciades puntualment: **Dades geoestadístiques** ('geoestatistical data')

ESTADÍSTICA ESPACIAL

1. Dades en graella o dades d'àrea ('lattice data' o 'areal data')
 - Les **dades d'àrea** o en graella es corresponen amb variables aleatòries discretes (recomptes) ("count data").
 - En les dades d'àrea ***no es coneix la localització exacta del cas.***
 - ***Les localitzacions són àrees*** amb límits geogràfics ben definits, habitualment unitats administratives (ciutats, barris, seccions censals, etc.).

ESTADÍSTICA ESPACIAL

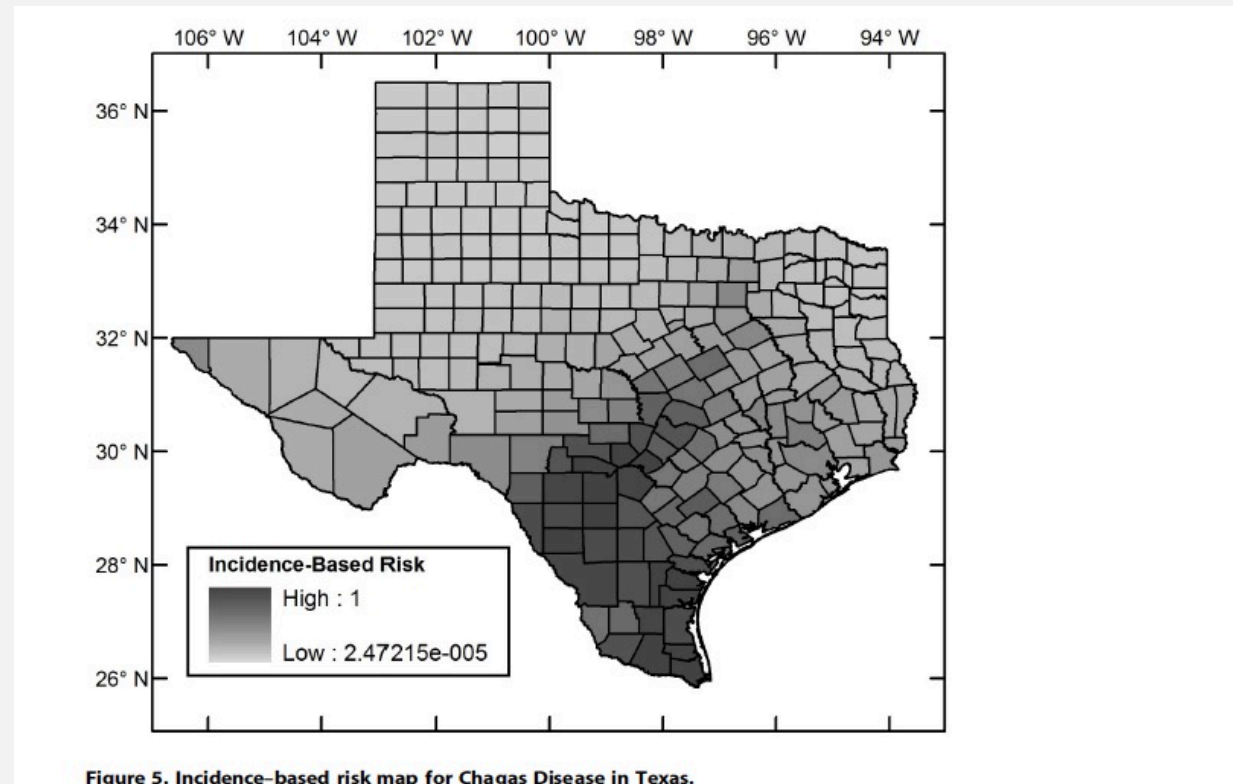
1. Dades en graella o dades d'àrea ('lattice data' o 'areal data')
 - La ***variable resposta és el nombre agregat de casos en aquesta àrea*** en un període de temps determinat.
 - Aquest era el tipus de dada més comunament utilitzat al començament de l'estadística espacial.

ESTADÍSTICA ESPACIAL

1. **Dades en graella o dades d'àrea ('lattice data' o 'areal data')**
 - Observacions procedents d'un procés aleatori sobre un conjunt de regions espacials distribuïdes espacialment de forma regular o irregular.
 - Queden definides matemàticament com un conjunt d'índexs de localitzacions amb un conjunt associat de veïns.
 - **Veïns:** àrees veïnes (proximitat, contigüitat, etc.) a una regió concreta.

ESTADÍSTICA ESPACIAL

1. Dades en graella o dades d'àrea ('lattice data' o 'areal data')



2. Introducció a l'epidemiologia i l'estadística espacial

ESTADÍSTICA ESPACIAL

2. Dades referenciades puntualment: Processos puntuals ('point processes')

- Els **processos puntuals** es corresponen amb variables aleatòries Bernouilli.
- En els processos puntuals ***es coneix la localització exacta del cas i aquesta és aleatòria.***
- Les ***localitzacions*** de l'esdeveniment d'interés ***s'observen en una regió determinada.***
Per exemple, les coordenades dels domicilis dels casos d'ELA a Catalunya.

ESTADÍSTICA ESPACIAL

2. Dades referenciades puntualment: Processos puntuals ('point processes')

- La **informació** sobre aquestes dades **no és pública**. Es recullen en estudis de casos i controls o en estudis de cohorts.
- Es poden afegir per unitats espacials, podent-se crear dades d'àrea.

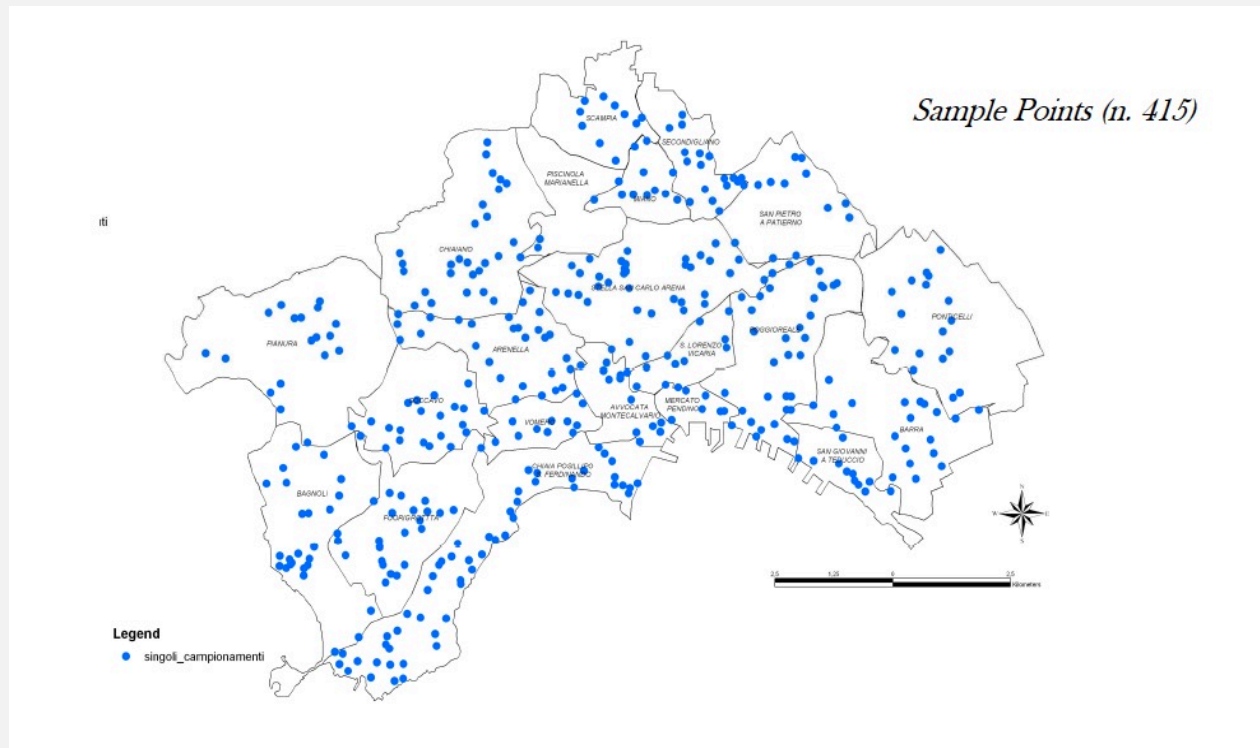
ESTADÍSTICA ESPACIAL

2. Dades referenciades puntualment: Processos puntuals ('point processes')

- En aquest cas, ens pot interessar:
 - Valorar si els esdeveniments segueixen un patró espacial determinat (agregació, forma regular, etc.)
 - Estudiar si un patró observat s'associa amb alguna variable (exposició a alguna variable ambiental, com la contaminació atmosfèrica; proximitat a focus contaminants; context socioeconòmic, etc.)

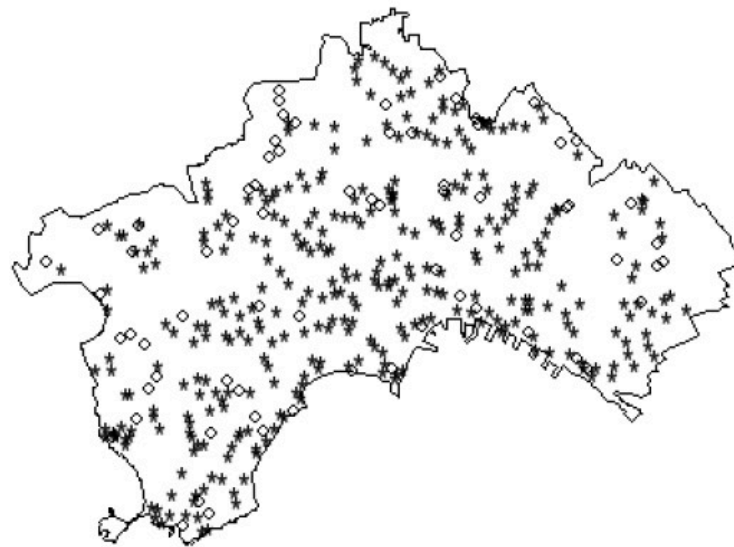
ESTADÍSTICA ESPACIAL

2. Dades referenciades puntualment: Processos puntuals ('point processes')



ESTADÍSTICA ESPACIAL

2. Dades referenciades puntualment: Processos puntuals ('point processes')



Distribution of cases (positives: circle) and controls (negatives: stars). Naples, February - May 2005

ESTADÍSTICA ESPACIAL

3. Dades referenciades puntualment: Dades geoestadístiques ('geoestatistical data')

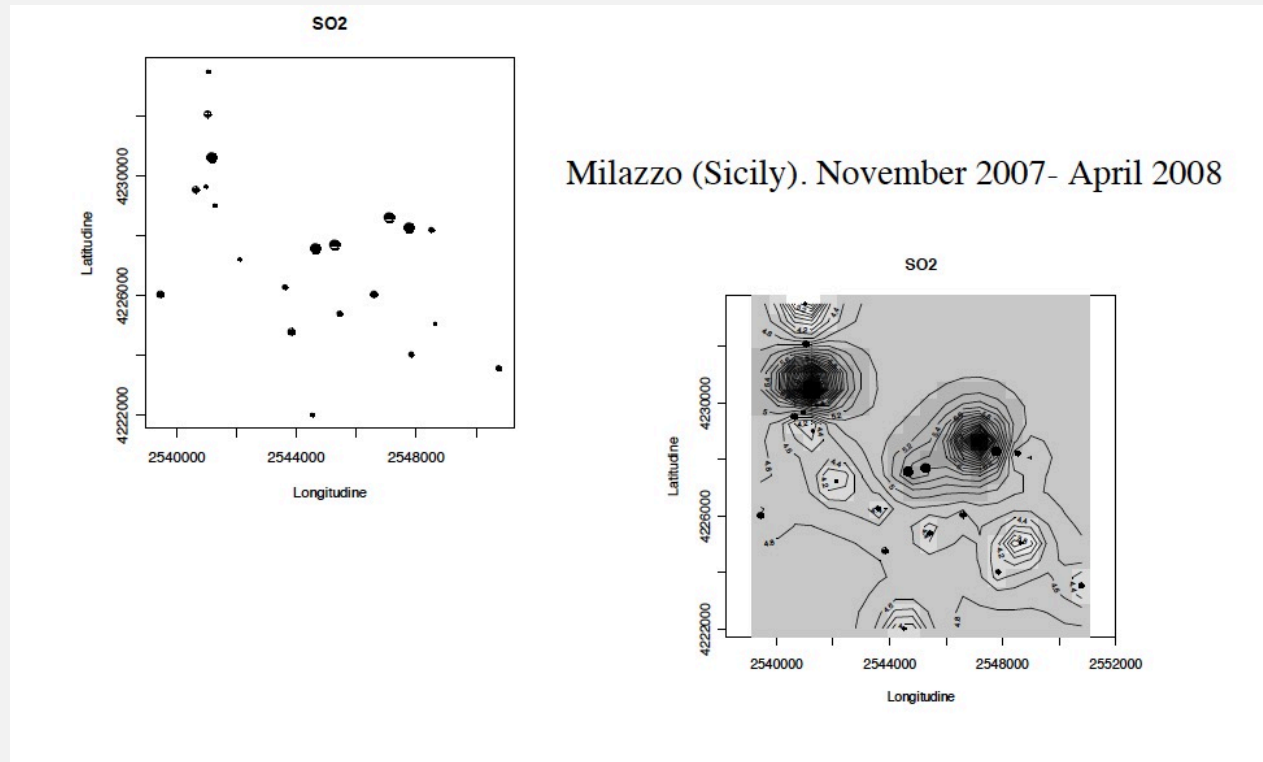
- Les **dades geoestadístiques** es corresponen amb variables aleatòries contínues.
- En les dades geoestadístiques ***es coneix la localització exacta del cas i aquesta és fixa*** (estacions captadores de contaminació atmosfèrica, vetes d'un determinat metall, etc.)
- ***La variable resposta es mesura en cada localització*** (per exemple, mesures de contaminació atmosfèrica, mesures de composició química del sòl, temperatura, etc.).

ESTADÍSTICA ESPACIAL

3. Dades referenciades puntualment: Dades geoestadístiques ('geoestatistical data')
- És a dir, disposem de mesures preses en punts fixos, definides en qualsevol lloc de l'espai, de manera que les seves **localitzacions defineixen una superfície espacialment contínua**.
 - Se sol **estendre la distribució espacial** dels valors d'un atribut **sobre la regió d'estudi** mitjançant models matemàtics (per exemple, kriging).

ESTADÍSTICA ESPACIAL

3. Dades referenciades puntualment: Dades geoestadístiques ('geoestatistical data')



ESTADÍSTICA ESPACIAL

Però, els tipus de disseny espacial, què són en realitat? processos? models? mètodes?

- El 2012, Diggle proposa un canvi de paradigma i re-defineix l'estadística espacial com *'un conjunt de models i mètodes estadístics que pretenen ajudar els científics a entendre fenòmens espacials, que no poden ser observats directament, sinó de forma indirecta, amb informació incompleta'*, en forma de dades en graella, processos puntuals i dades geoestadístiques. El model estadístic que va proposar, com a model únic base, és el **model log Cox**.

ESTADÍSTICA ESPACIAL

- És a dir, Diggle (2012) va unificar l'estadística espacial, tal i com McCullagh i Nelder (1989) van unificar els models lineals generalitzats.

