

Come capire la realtà che ci circonda in tempi di pandemia

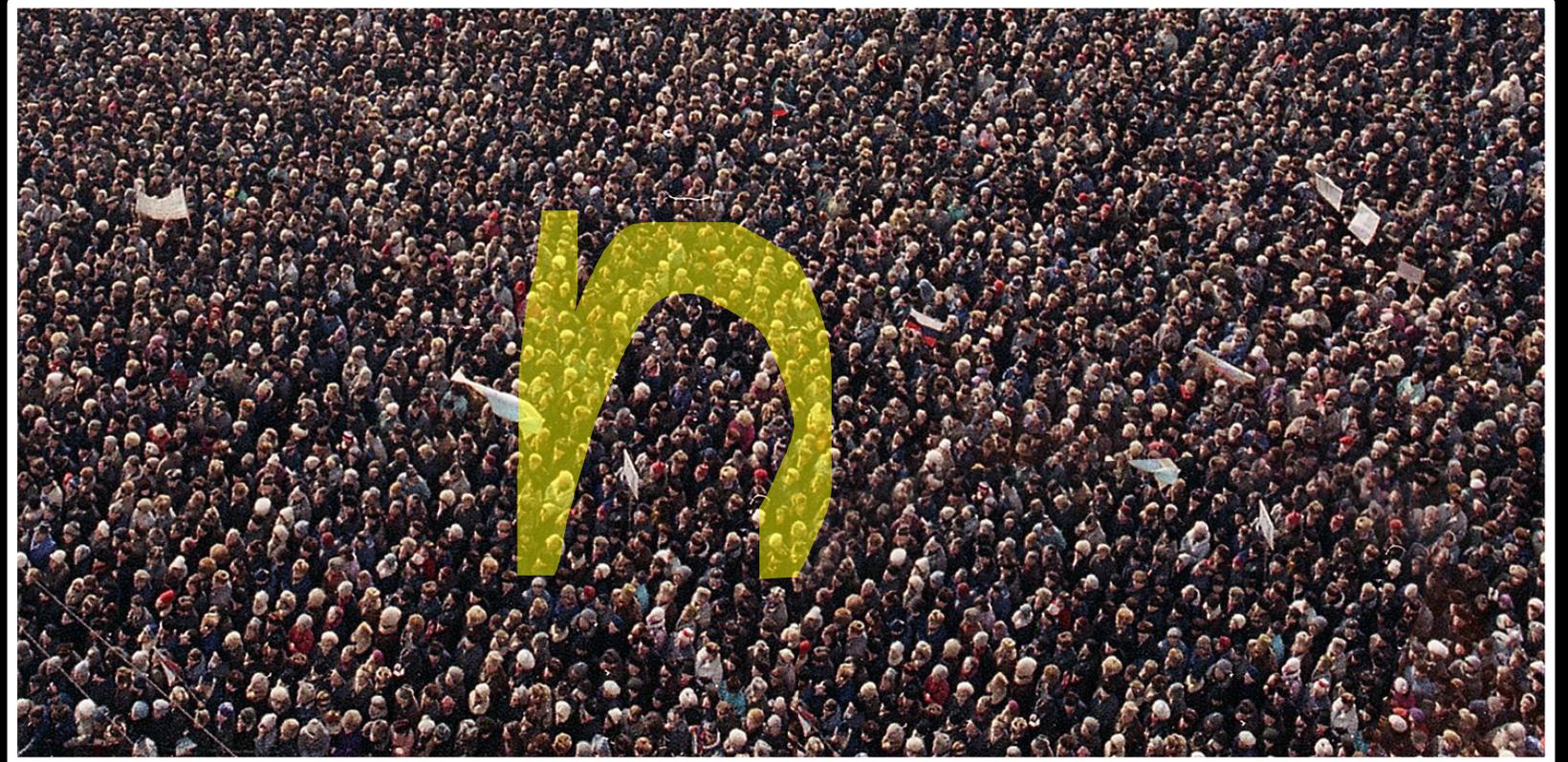
Dario Menasce
INFN - Milano Bicocca



E perché questo ci può aiutare a ridurre gli effetti di un lockdown generalizzato

(e tante altre cose ancora..)

La realtà che percepiamo presenta due punti di vista possibili

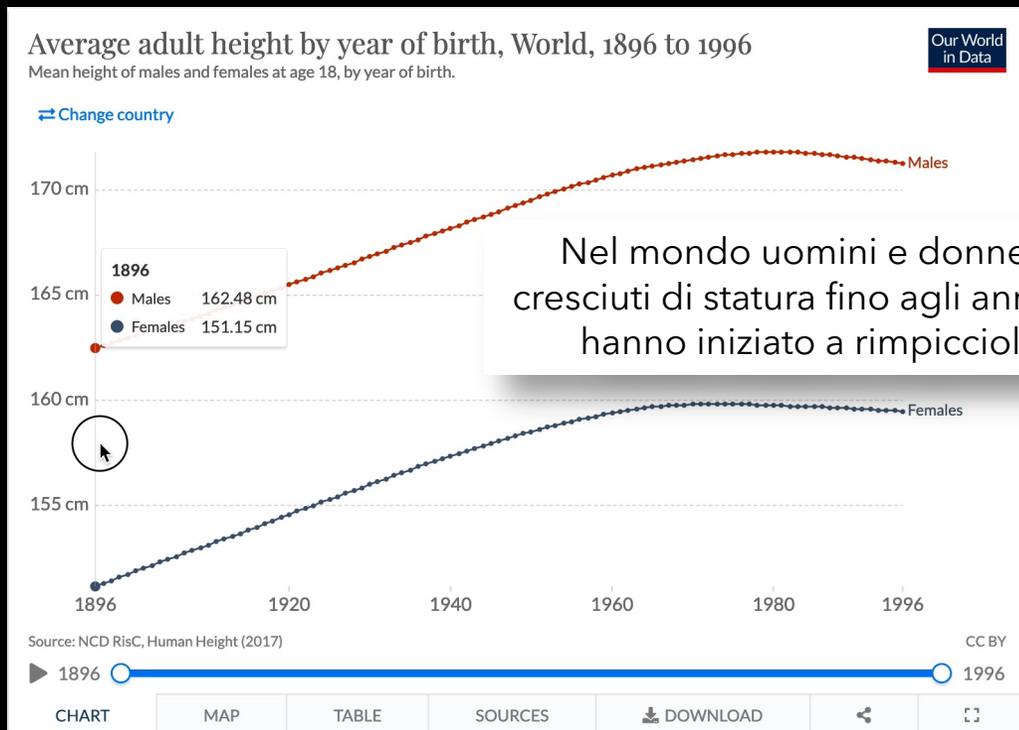


Noi e la nostra visione del mondo (il nostro caso singolo dal nostro punto di vista di individuo)

Tutto il resto sono casistiche, spesso davvero molto variegate

Potremmo, ad esempio, essere interessati all'altezza delle persone: tutti conosciamo persone piccole o gigantesche ma se ci venisse chiesto qual è l'altezza **media** delle persone oppure quanto sia **cambiata** la statura di uomini e donne nell'ultimo secolo potremmo solo citare le nostre sensazioni personali o i casi che conosciamo direttamente.

Diversamente potremmo consultare tabelle numeriche o grafici con i dati statistici, rimanendo magari sorpresi....



Nel mondo uomini e donne sono cresciuti di statura fino agli anni 70, poi hanno iniziato a rimpicciolire...

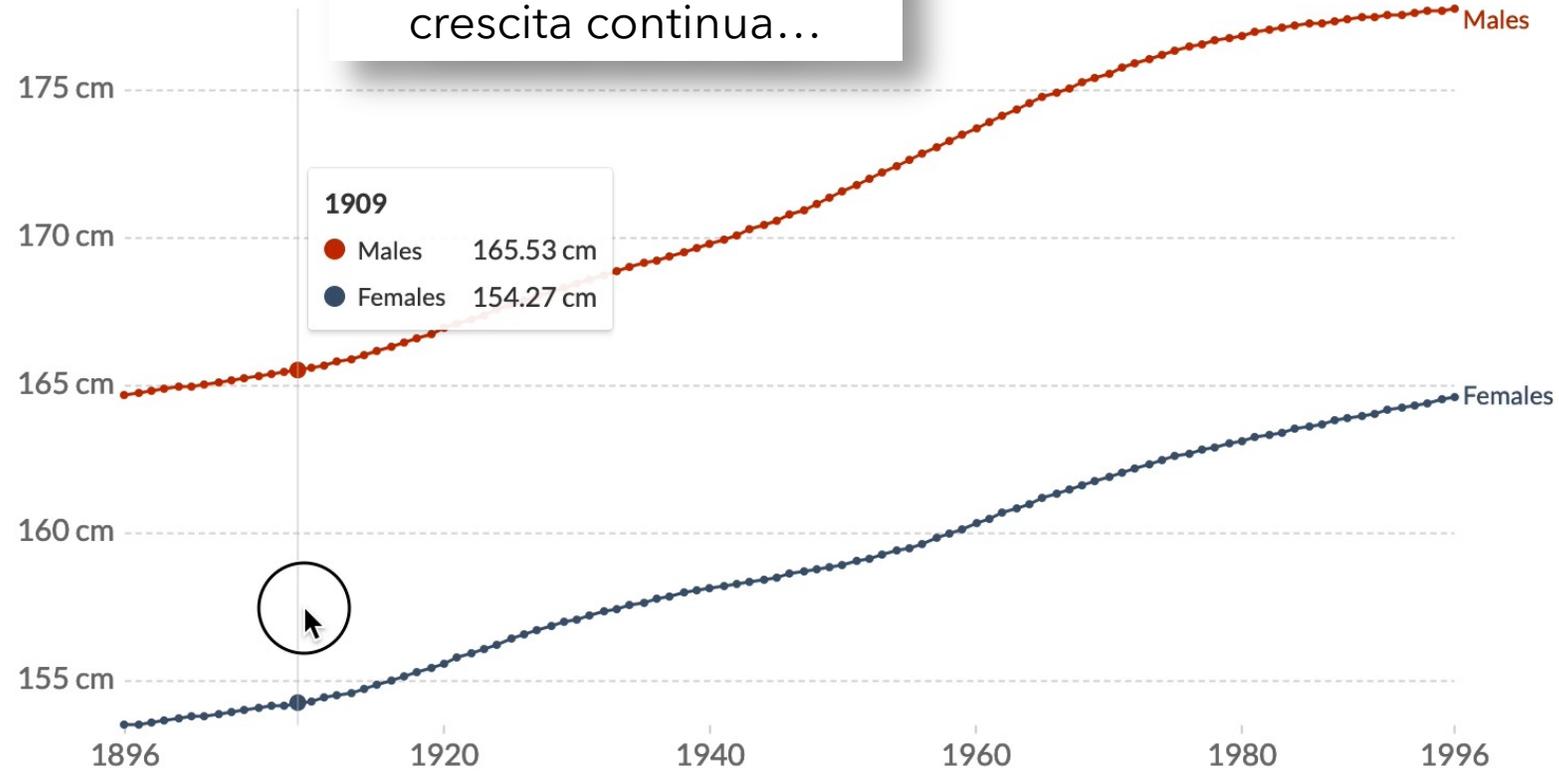


Average adult height by year of birth, Italy, 1896 to 1996

Mean height of males and females at age 18, by year of birth.

↔ Change country

Ma non in Italia, dove la
crescita continua...



Source: NCD RisC, Human Height (2017)

CC BY



CHART

MAP

TABLE

SOURCES

↓ DOWNLOAD



Altro punto importante: siamo abituati a pensare che cose che appaiono identiche lo siano in modo assoluto



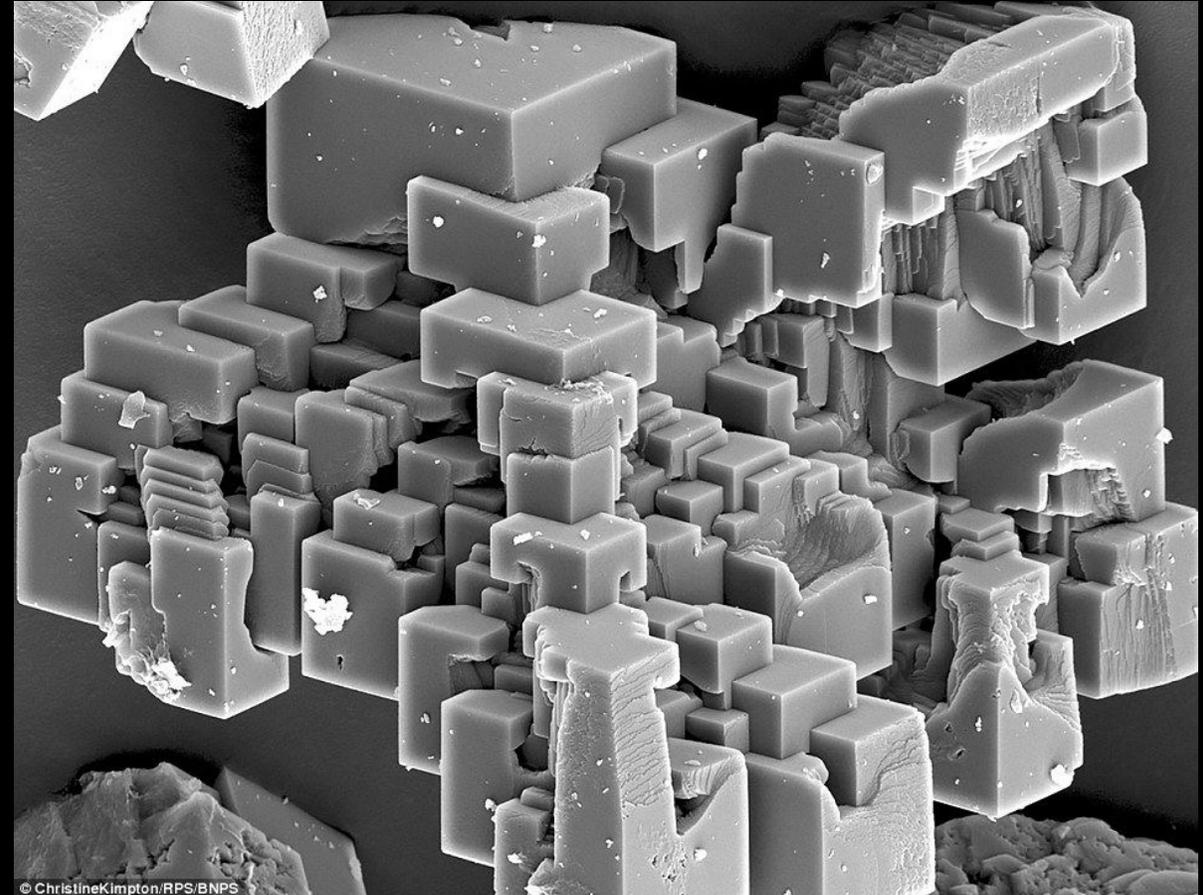
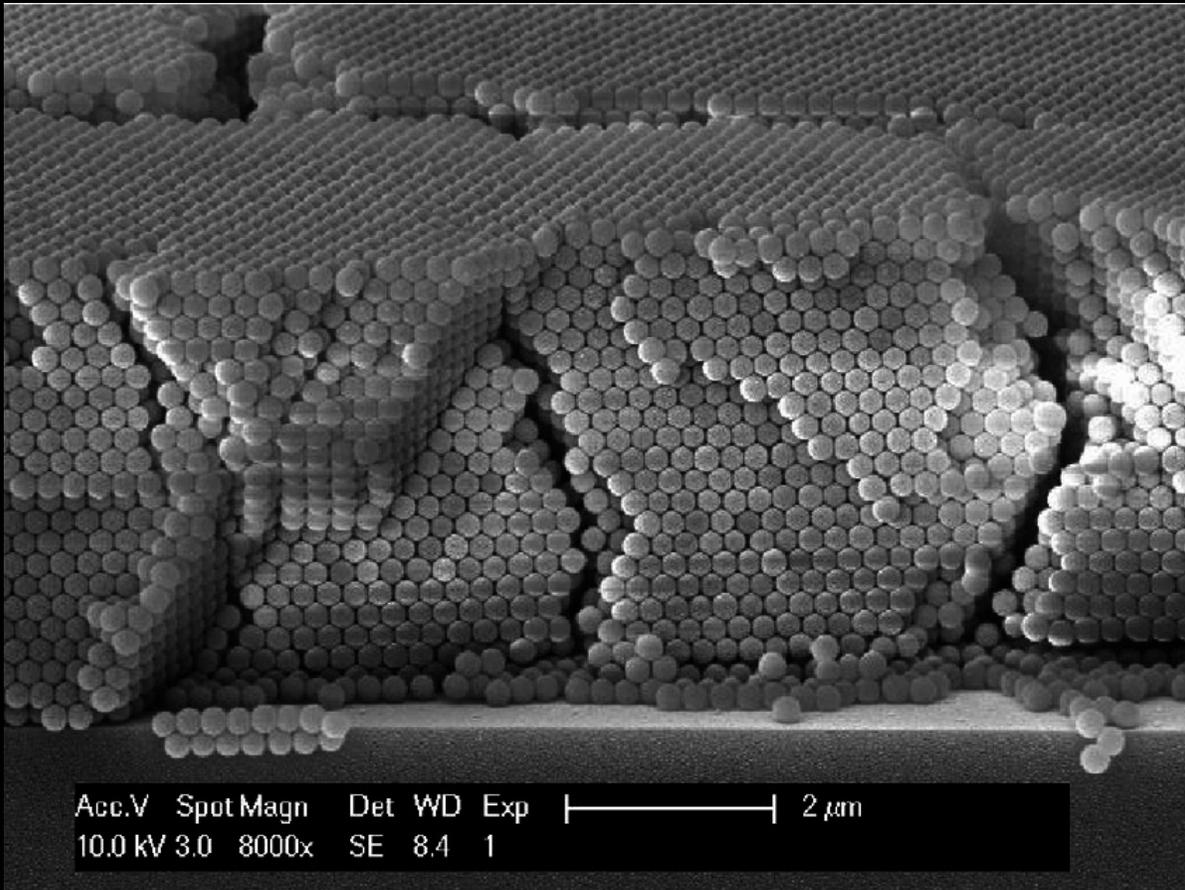
La perfezione assoluta in natura non esiste



Anche quando sembrerebbe invece di sì...



Basta guardare con maggiore attenzione, più in profondità e si scoprono sempre imperfezioni ed errori: il dettaglio conta, eccome!



Occorre possedere un metodo per tener conto e quantificare opportunamente queste sottili differenze

Tutto ciò si traduce in un fatto inevitabile: ogni misura che si possa mai sperare di fare avrà sempre un errore associato, un valore che rappresenti la sua incertezza.

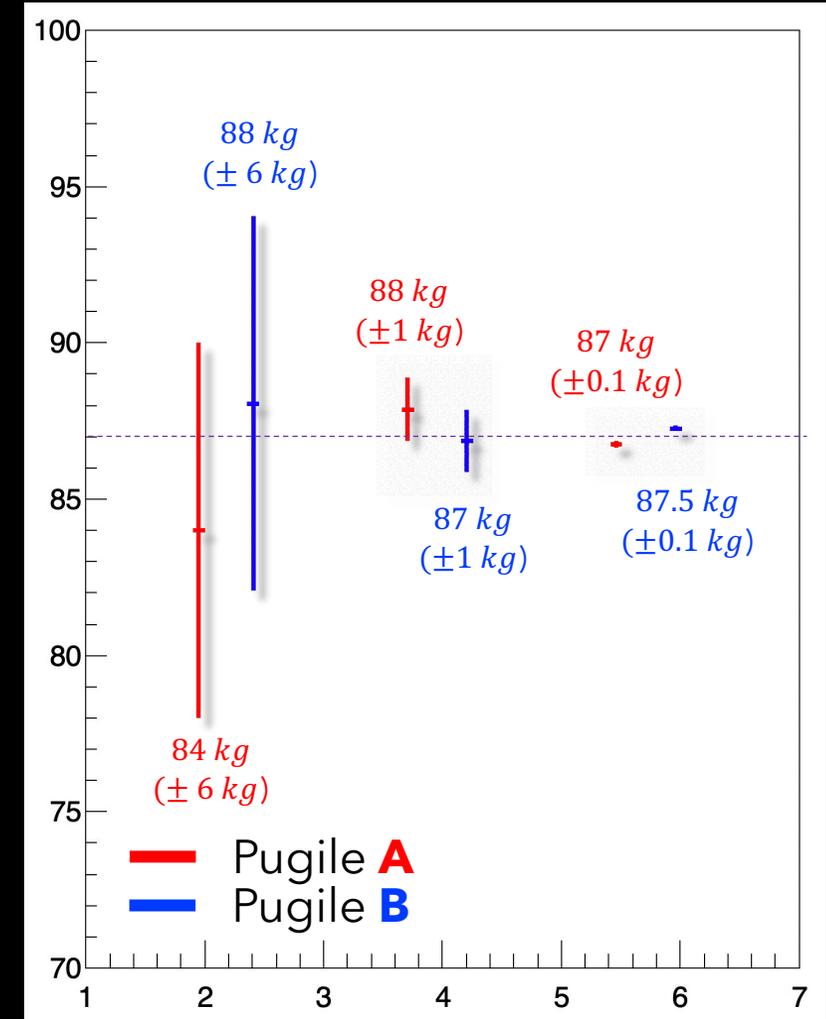
Dovessimo pesare due pugili per determinare la loro classe di gara potremmo farlo in diversi modi:

- A occhio (con una approssimazione di circa 6 kg)
- Con una bilancia da trasportatore (precisione di circa 1 kg)
- Con una bilancia da farmacista (precisione di circa 1 hg)

Supponiamo che la soglia per non far gareggiare i due pugili sia posta a 87.25 kg: misurando a occhio o con la bilancia da trasportatore, date le incertezze delle relative misure, entrambi i pugili potrebbero essere considerati pesare allo stesso modo.

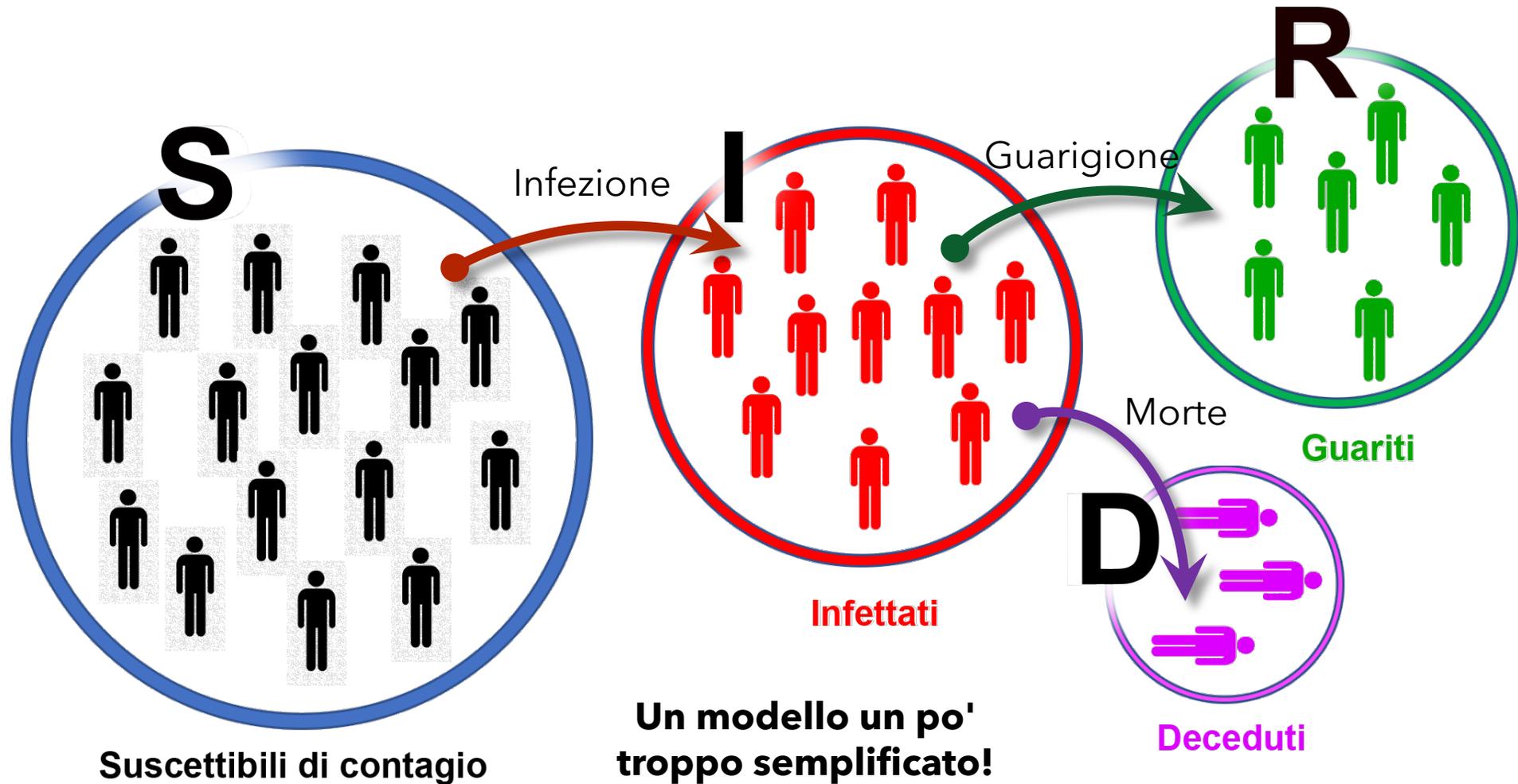
È solo la misura di precisione che ci permette di osservare che i due valori sono tra loro più lontani delle rispettive incertezze e che uno dei due supera la soglia di peso ammessa (il blue) mentre il rosso risulta sotto soglia.

Peso (in kg)

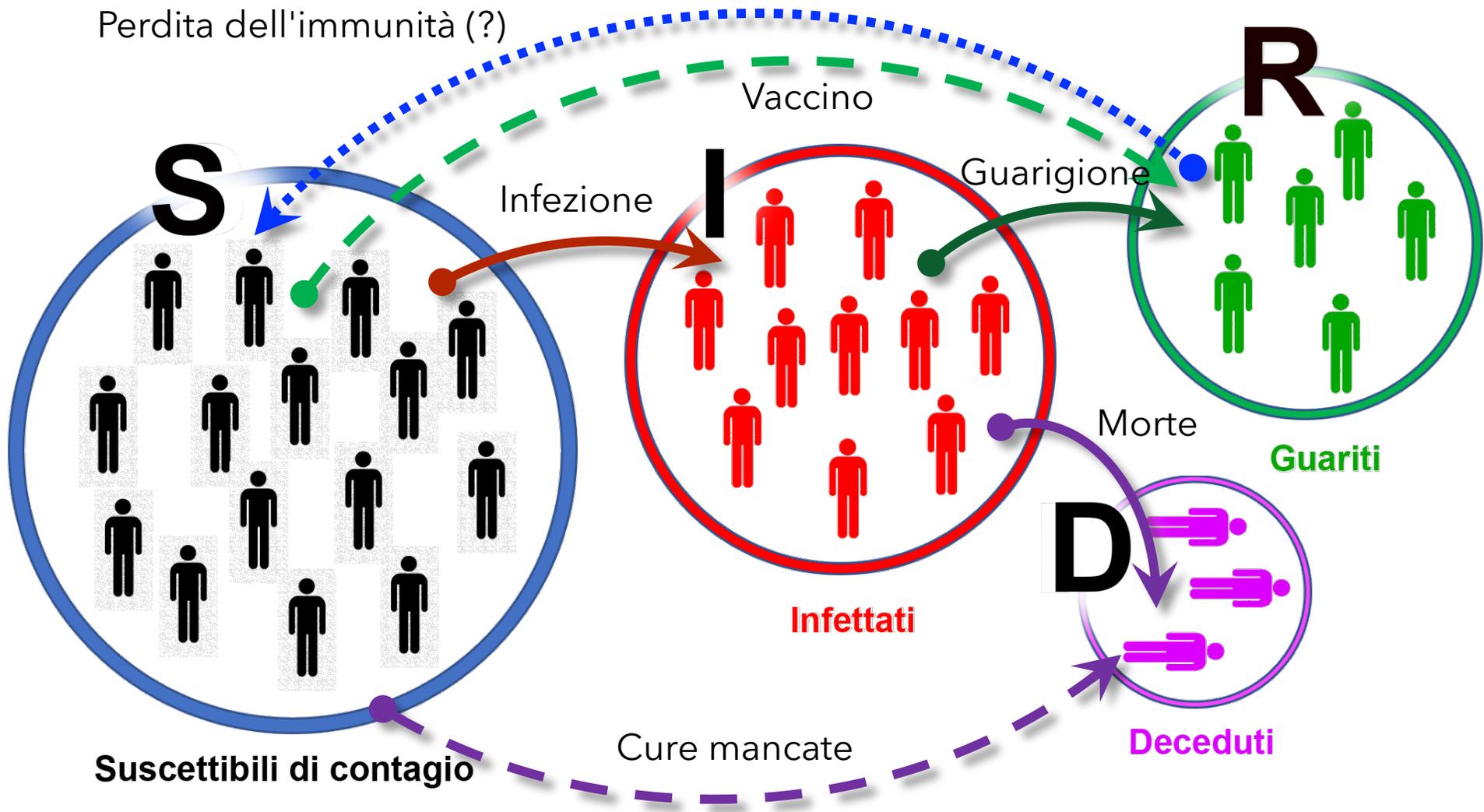


Veniamo ora ad un aspetto della realtà che da un anno a questa parte assilla tutto il mondo: la pandemia di COVID19

Un rudimentale modello di sviluppo di un'epidemia



Un rudimentale modello di sviluppo di un'epidemia

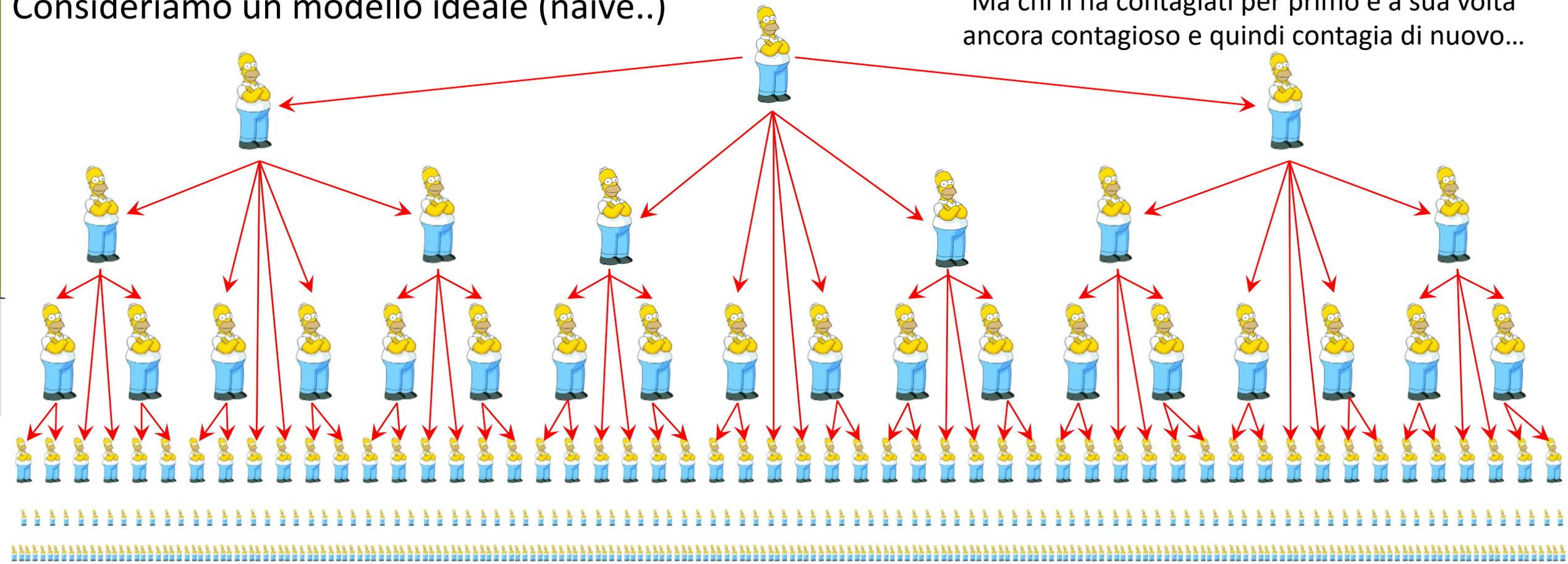


Ma come si sviluppa, *nel tempo*, il contagio?

Perché sono tutti allarmati quando personalmente non ci accorgiamo dell'aumento dei malati e dei ricoverati ?

Consideriamo un modello ideale (naive..)

Ma chi li ha contagiati per primo è a sua volta ancora contagioso e quindi contagia di nuovo...

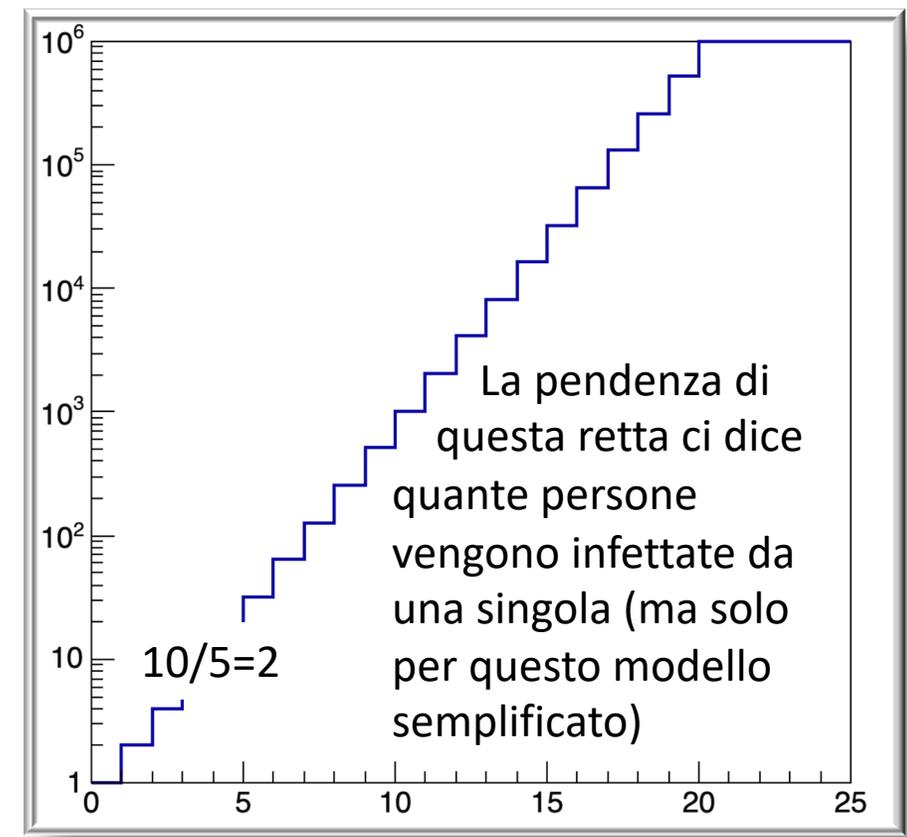
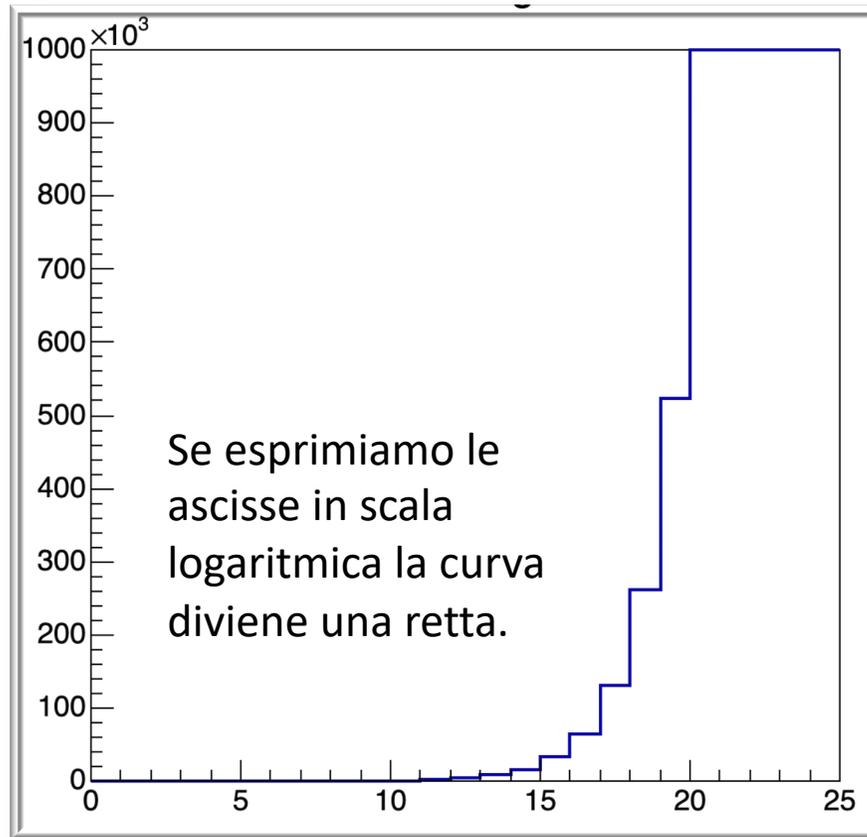


Questo è un modello realistico ma non accurato: non tiene conto, per esempio, del fatto che dopo essere stati infettati o si muore (pochi, per fortuna) o si guarisce: comunque sia, dopo un certo tempo si cessa di essere infettivi e si viene rimossi dalla catena di contagio. Lo sviluppo reale può essere quindi, in genere, un po' diverso da quanto illustrato qui e la matematica che lo descrive più complicata di conseguenza (serve un modello più raffinato...).

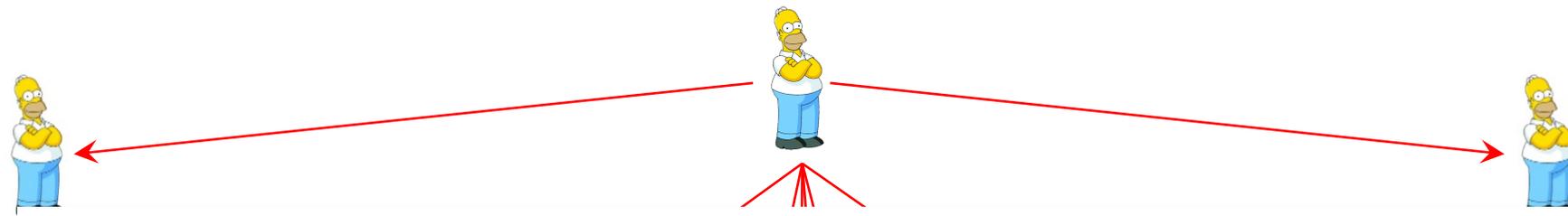
Giorno 7
Contagiati 486

Supponiamo adesso, per semplicità, che solo i nuovi contagiati siano a loro volta in grado di infettare altre 2 persone: la formula che esprime questo modello naïve, meno esplosivo del precedente ma più irrealistico è data da $y = 2^x$, dove x sono i giorni che passano. Come si vede, questo semplice modello, partendo da un solo infetto, arriva a 1 milione di contagiati in soli 21 giorni (considerando 2 infettati alla volta per persona...)

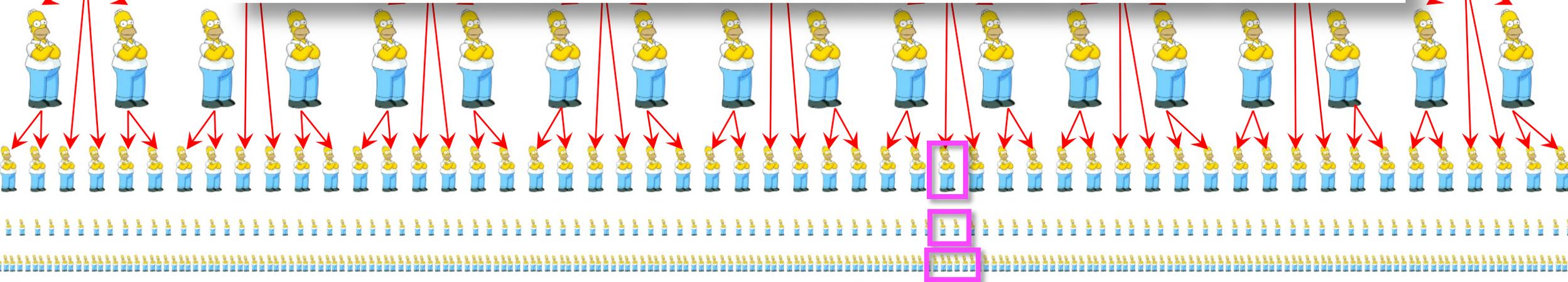
1	1
2	2
3	4
4	8
5	16
6	32
7	64
8	128
9	256
10	512
11	1024
12	2048
13	4096
14	8192
15	16384
16	32768
17	65536
18	131072
19	262144
20	524288
21	1.04858e+06
22	2.09715e+06



Cosa si può fare per frenare questa esplosione esponenziale?



Agendo su una persona **solo al quinto giorno** si riesce quindi a ridurre il numero dei contagiati totali, al settimo giorno, da 486 a 477. Una diminuzione del solo 1.85% del totale, ben poca cosa.

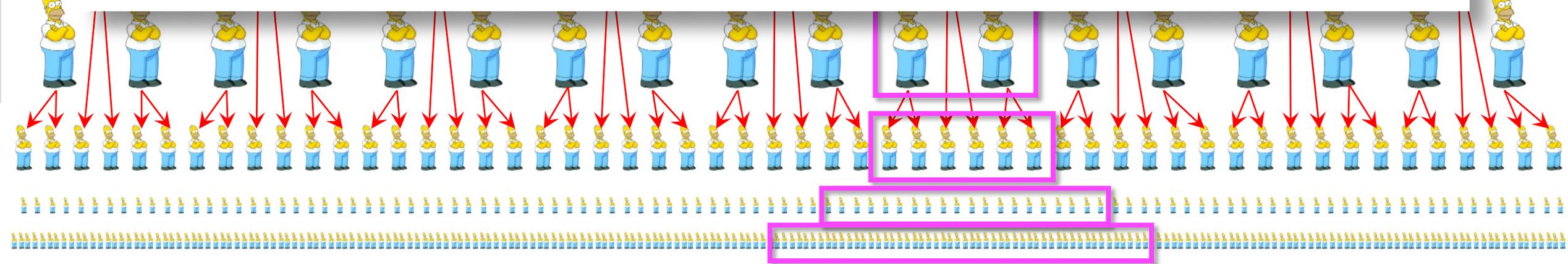


Supponiamo che un contagiato, che si renda conto di esserlo stato tre giorni prima, decida, per senso civico, di mettersi spontaneamente in quarantena a partire solo **dal quinto giorno**. Questo significa che le persone che avrebbe potuto contagiare (e quelle a loro volta altre) non lo faranno più. Da questo schema quelle persone spariscono dal gruppo dei contagiati infettivi! Quante ne spariscono? Secondo questo modello sparisce lui il quinto giorno, lui più i due che avrebbe potuto contagiare il sesto giorno fino a un totale di **nove** al **settimo giorno**.



Quante ne spariscono grazie a questa saggia decisione di anticipare la quarantena al secondo giorno? Secondo questo modello il terzo giorno sparisce lui dalla lista dei contagiosi, il quarto giorno lui più i due che avrebbe potuto contagiare fino a un totale, al settimo giorno, di 54 persone in meno.

← Abbiamo migliorato la situazione: siamo passati all'11% dei contagi in meno



Se la persona contagiata decidesse invece di mettersi in quarantena **non appena ha scoperto di esserlo**, ossia già al **terzo giorno** dall'inizio della pandemia, cosa accadrebbe in questo scenario?

Si vede subito che, raggiunto il settimo giorno sarebbero molte di più le persone cui viene evitato il contagio e che quindi sparirebbero a loro volta dal gruppo dei potenziali contagiati/contagiosi.

Capiamo quindi bene che non basta dire
"*ho predisposto un lock-down*" per ridurre l'entità del problema,
bisogna farlo il prima possibile!

La variabile tempo è la più cruciale in questa lotta.

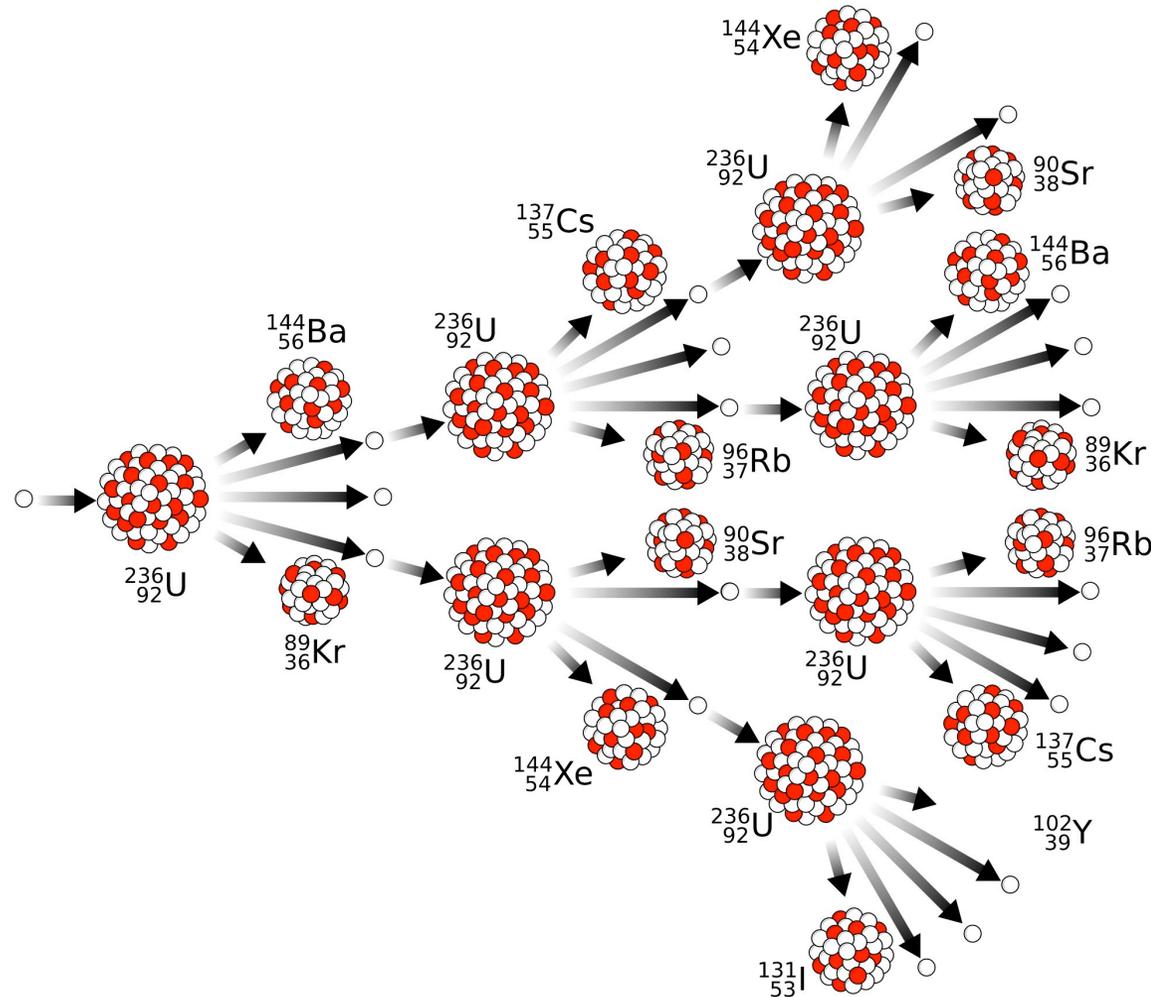


Per essere davvero efficaci in questo, bisogna però bloccare le possibilità di contagio per quante ***più persone*** possibili ma, cosa ben più importante, occorre farlo il ***prima possibile***.
Quando è il "***prima possibile***"? Dipende dalla bontà degli indicatori che abbiamo a disposizione sul progresso della pandemia e dalla qualità e affidabilità dei dati resi disponibili localmente.
Prima raccogliamo i dati e prima siamo avvisati che sta succedendo qualcosa: se agiamo immediatamente a quel punto abbiamo la possibilità di bloccare tutto sul nascere!

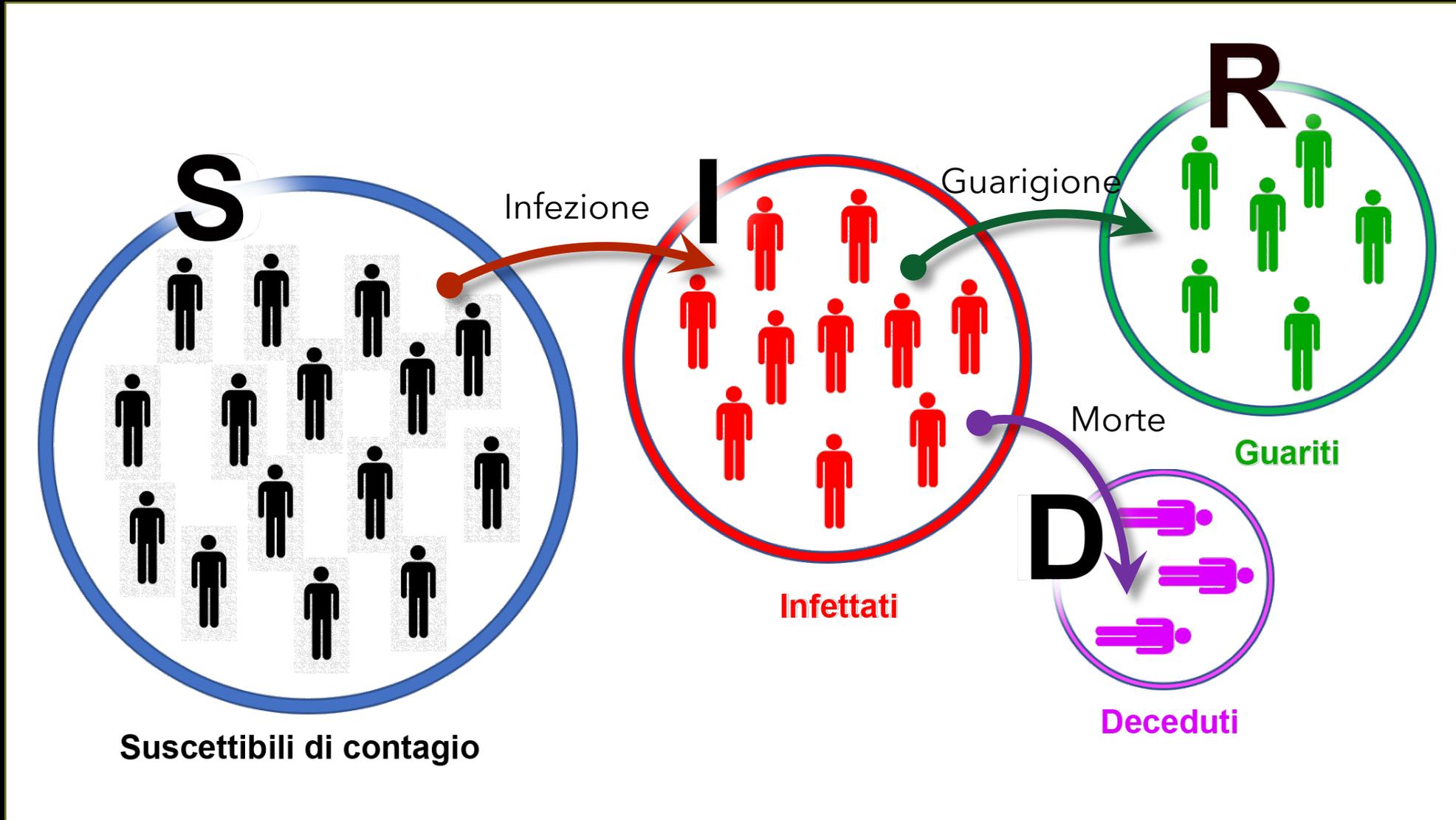
Prima, però, per farci un'idea di come una legge esponenziale abbia una progressione così veloce che, una volta innescata sia quasi impossibile fermarla vediamo un esperimento davvero impressionante:

Immaginiamo che questa pallina da ping pong sia il paziente zero che, ad un certo momento entra in contatto con altre persone suscettibili di contagio (e immaginiamo il tutto al rallentatore)

Questo è ciò che accade, in modo analogo, in una reazione a catena nucleare, solo che questa si sviluppa in tempi dell'ordine dei microsecondi. Il contagio da COVID è più lento, certo, ma la rapidità con la quale cresce il numero di infetti è comunque regolata da un'esponenziale (un contagio ogni 10 secondi)



Come possiamo modellare matematicamente questo fenomeno?



Le relazioni matematiche che descrivono questo modello

Il passaggio da **Suscettibili** a **Infetti** è una **variazione nel tempo** (quindi una **derivata** rispetto a t)

Questa stessa variazione sarà **proporzionale** (ossia moltiplicata per un numero, $\alpha(t)$) al prodotto dei suscettibili e degli infetti diviso per N , la popolazione totale, tramite un coefficiente, chiamato **tasso di incremento degli infetti**, tasso che dipende dal tempo. Il segno meno indica che, nel tempo, il numero dei suscettibili **diminuisce** a favore di quello degli infetti.

$$\frac{dS}{dt} = -\alpha(t) \frac{IS}{N}$$

$$\frac{dI}{dt} = \alpha(t) \frac{IS}{N} - \beta(t)I - \gamma(t)I$$

$$\frac{dR}{dt} = \beta(t)I$$

$$\frac{dD}{dt} = \gamma(t)I$$

Potremo allora uguagliare questi due termini, ottenendo un'equazione differenziale che descrive l'andamento del contagio, equazione la cui incognita è una funzione, $\alpha(t)$:

considerazioni analoghe ci permettono di scrivere le equazioni che descrivono l'evoluzione temporale anche del numero di infettati, guariti e deceduti:

a questo punto abbiamo un sistema di quattro equazioni con tre incognite (α , β e γ) che possiamo risolvere:

Cosa ce ne facciamo delle soluzioni? Costruiamo un'importante variabile derivata

Poiché si tratta di equazioni differenziali le incognite non sono numeri, bensì funzioni a loro volta (del tempo, in particolare). Senza scrivere esplicitamente le forme che queste tre equazioni assumono, (α, β e γ) possiamo però definire con esse un nuovo parametro, chiamato R_t :

$$R_t(t) = \frac{\text{Frazione di persone che **entra** nel gruppo degli infetti}}{\text{Frazione di persone che **esce** dal gruppo degli infetti}} = \frac{\text{Tasso infettività } \alpha(t)}{\text{Tasso di guarigione } \beta(t) + \text{Tasso di decesso } \gamma(t)}$$

Questa definizione è **costruita** in modo tale per cui il suo valore diventa $R_t(t) < 1$ quando il numero di persone che esce dalla condizione di infetto (denominatore) diventa maggiore di quella che ci entra (numeratore), indice che l'epidemia tende a spegnersi.

La quantità $R_t(t)$ è chiamata indice di contagio ed è la principale variabile che viene tenuta in considerazione dai responsabili del governo per il controllo della pandemia.

La forma che ha $R_t(t)$ al variare del tempo dipende dal modello considerato: le equazioni differenziali che abbiamo usato (il cosiddetto modello SIRD) sono **una** tra le **tante** scelte possibili che si possono fare. Ognuna fornirà una curva un po' diversa (anche se l'andamento sarà simile)

Come si calcola, in pratica R_t ?

Esistono vari metodi in letteratura, per semplicità citerò quello più rapido sviluppato dal nostro gruppo (metodo covidStat).

Assumiamo che I_t sia il numero di persone infettate al tempo t a partire dal giorno di ipotetico inizio dell'epidemia.

Ogni persona può contagiare altre durante il suo periodo di contagiosità: si può ipotizzare che il giorno d infetti un certo numero di altre persone che diventeranno contagiose a loro volta il giorno $t > d$ con una distribuzione di probabilità discreta w_s con $s = t - d$ (il tempo necessario perché l'infezione si sviluppi dal momento del contagio, chiamato *generation time*). Per $d > 20$ giorni la probabilità, approssimabile a una gamma distribution, diventa trascurabile.

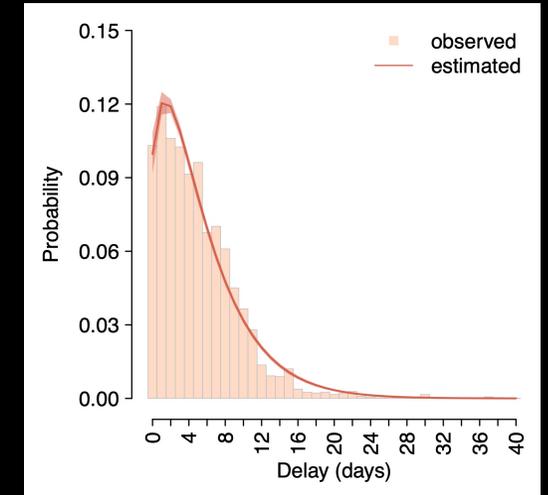
Per w_s vale una condizione di normalizzazione:
$$\sum_{s=1}^{\infty} w_s = 1$$

Potremo allora affermare che al tempo t il valore di aspettazione del numero di infetti, $\mathbb{E}[I_t]$, possa essere determinato da I_d , con $d = 0, \dots, t - 1$ mediante la relazione

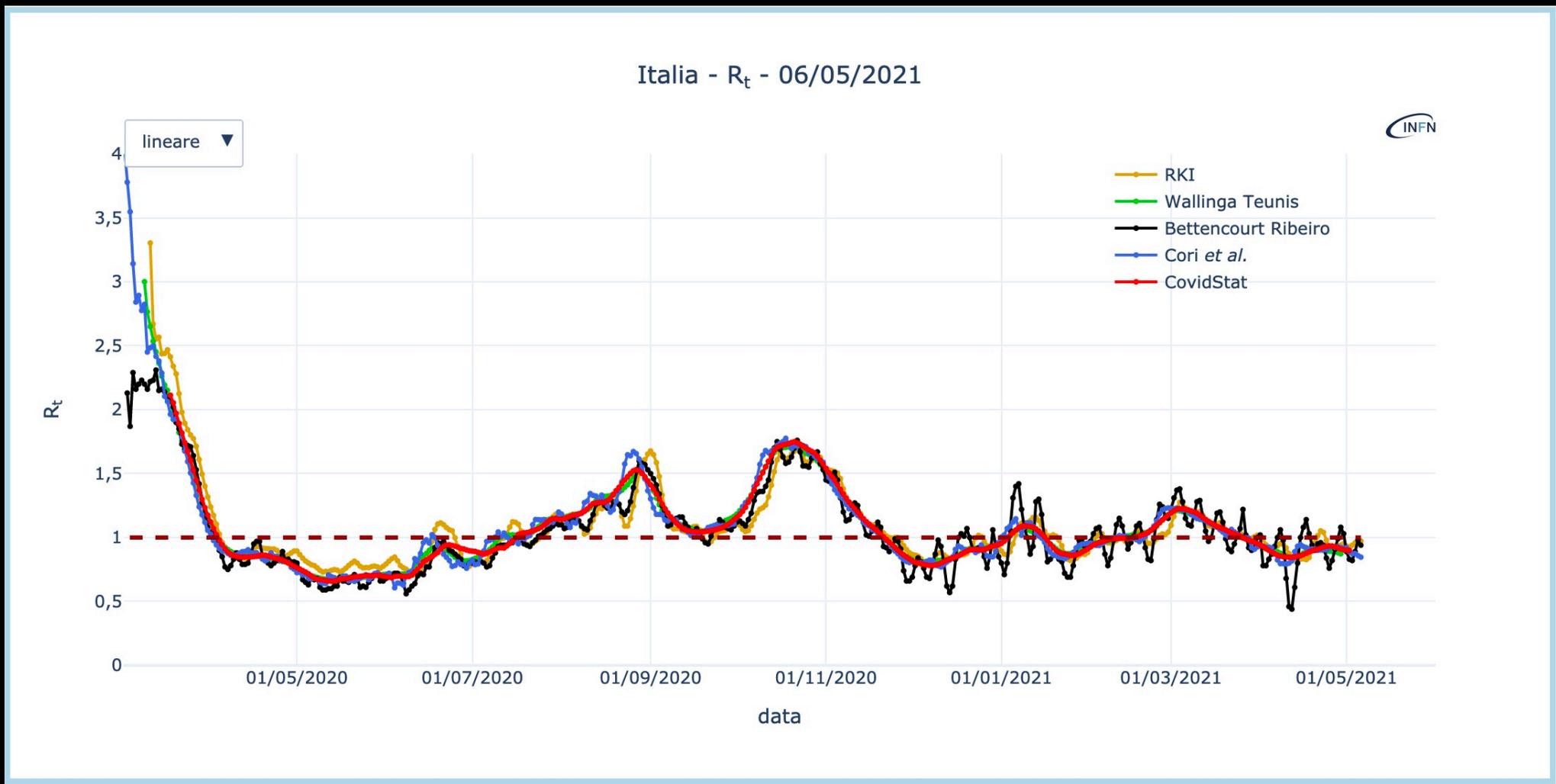
$$\mathbb{E}[I_t] = R_t \sum_{d=0}^{t-1} I_d w_{t-d}$$

Un punto critico è l'opportuna scelta della funzione peso w_s : assumendo per essa il valor medio della gamma distribution ($w_g = 4$) si può dimostrare che

$$\hat{R}_t = \frac{I_t}{I_{t-g}}$$

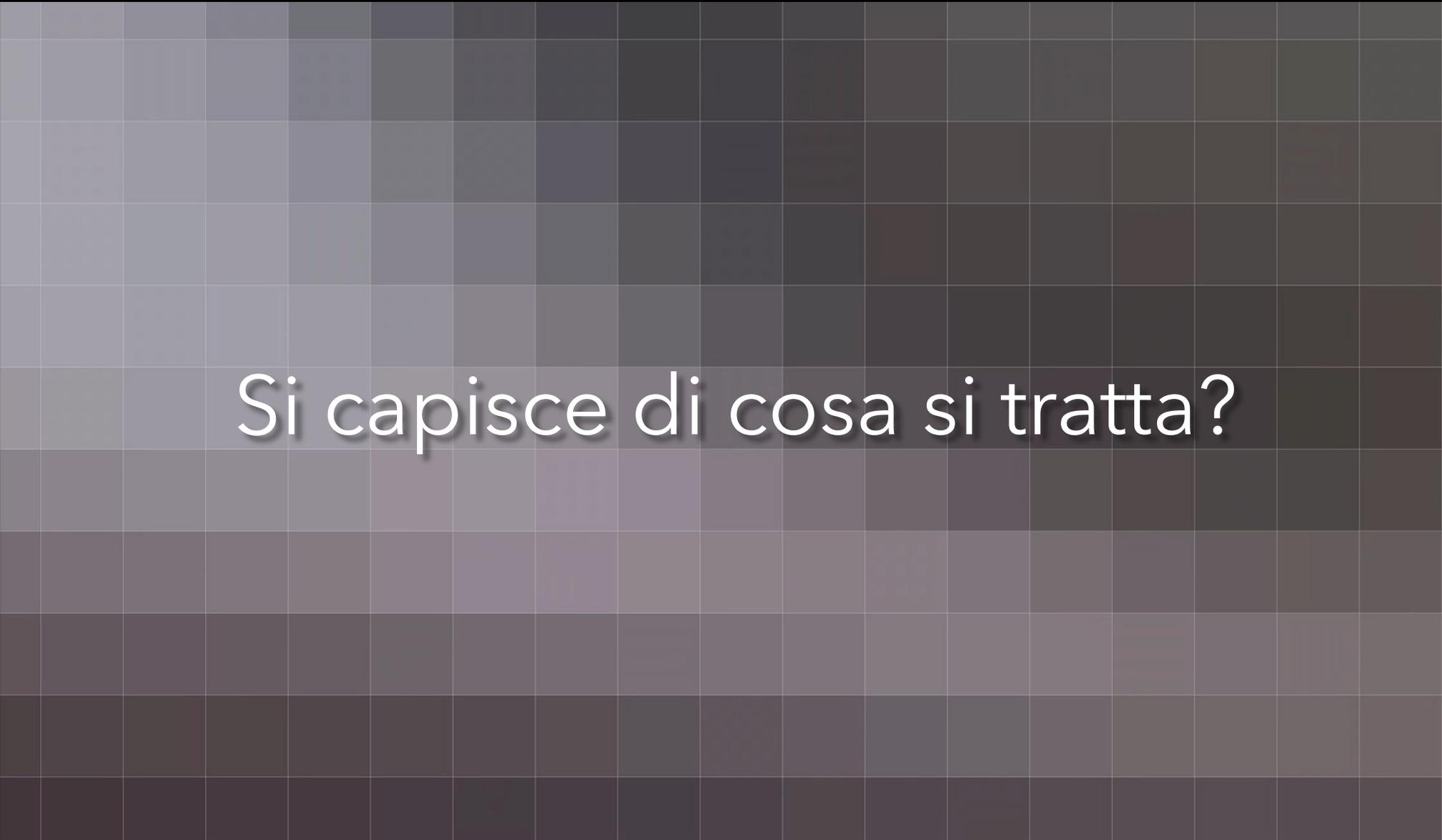


L'andamento nel tempo di R_t , calcolato in questo modo basato sul numero di casi quotidiani, ha questo andamento:



Perché si afferma che per poter dire qualcosa su un fenomeno occorre avere molti dati?

Vediamo un esempio banale: una griglia di circa 17x11 quadrati colorati, 187 pixels

A 17x11 grid of 187 pixels with a grayscale gradient. The grid is composed of 17 columns and 11 rows of squares. The color transitions from light gray on the left to dark gray on the right, and from light gray at the top to dark gray at the bottom. The text "Si capisce di cosa si tratta?" is centered in the grid.

Si capisce di cosa si tratta?

Credo siate d'accordo nel dire di no. Proviamo allora ad acquisire più dati, più pixels

Forse una sfera che rotola?...

$$61 \times 36 = 2196$$



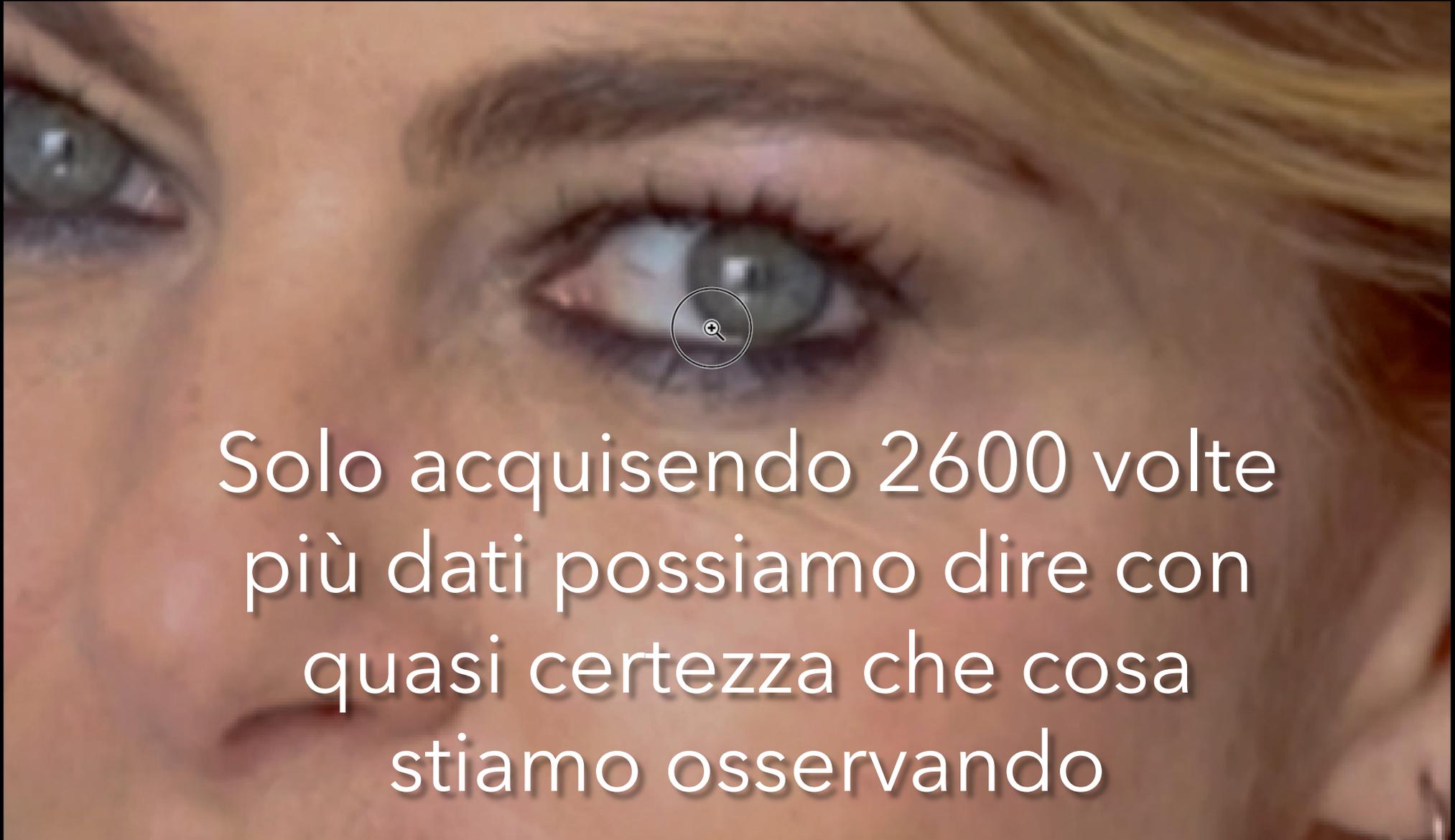
Si inizia a intravedere un
viso, in particolare un occhio.

215x127=27305



Il viso pare quello di una
donna, ma chi è?...

847x573=485331 (2595 volte più dati che non la prima immagine)



Solo acquisendo 2600 volte
più dati possiamo dire con
quasi certezza che cosa
stiamo osservando

Un esempio più vicino alle scienze fisiche: il dispositivo di Galton e la campana di Gauss

Da un foro (in alto) facciamo cadere delle palline su dei pioli che sporgono dallo schermo: dopo molti rimbalzi cadranno entro delle specie di cassettoni a compartimento posti in basso al dispositivo:

50%

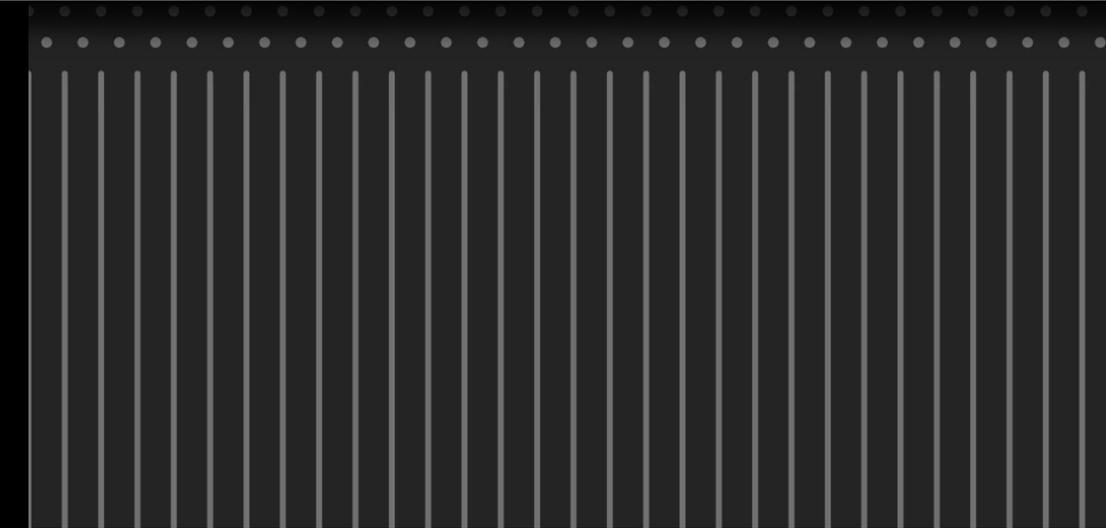


Ad ogni rimbalzo contro un piolo avremo la stessa probabilità che una pallina vada a destra piuttosto che non a sinistra



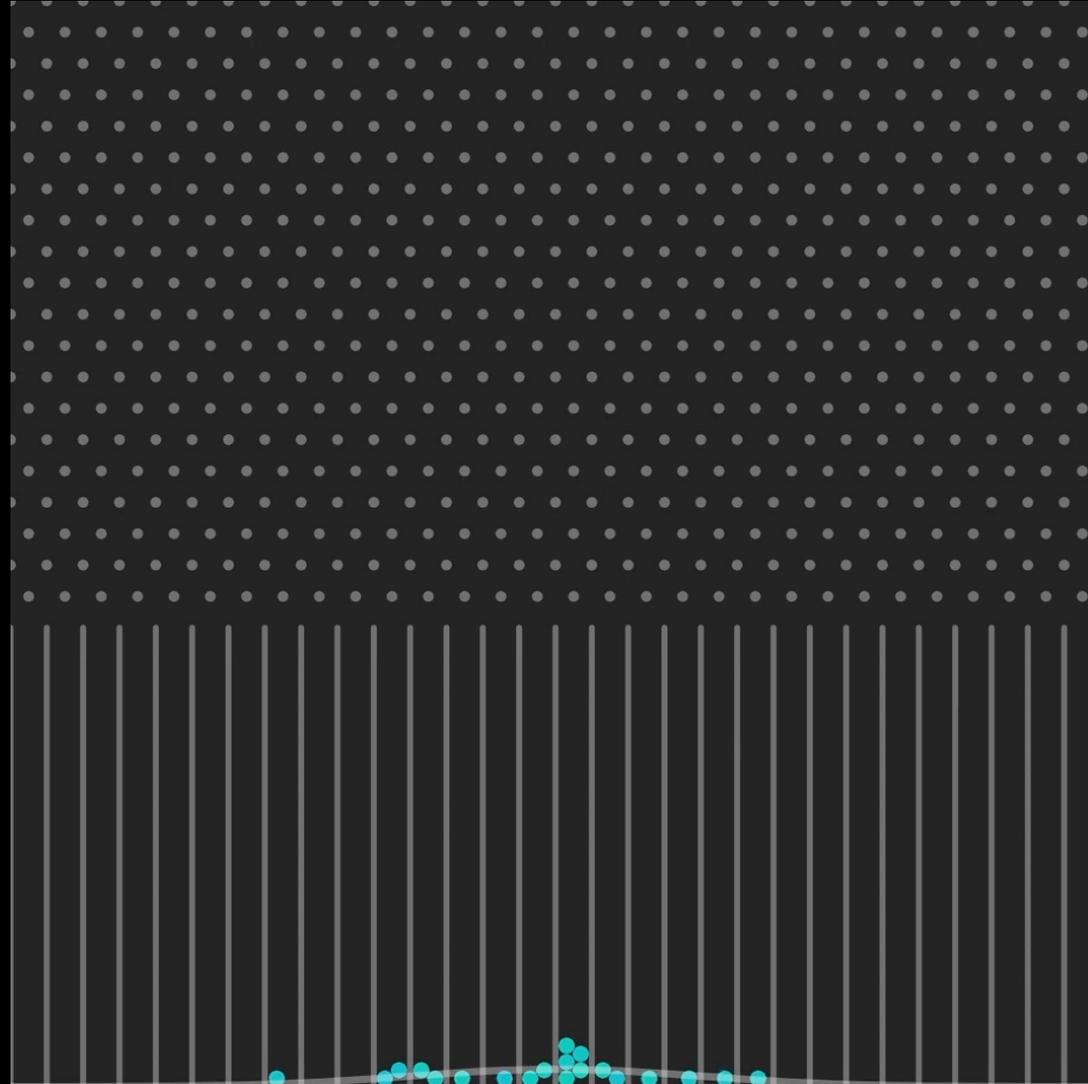
50%

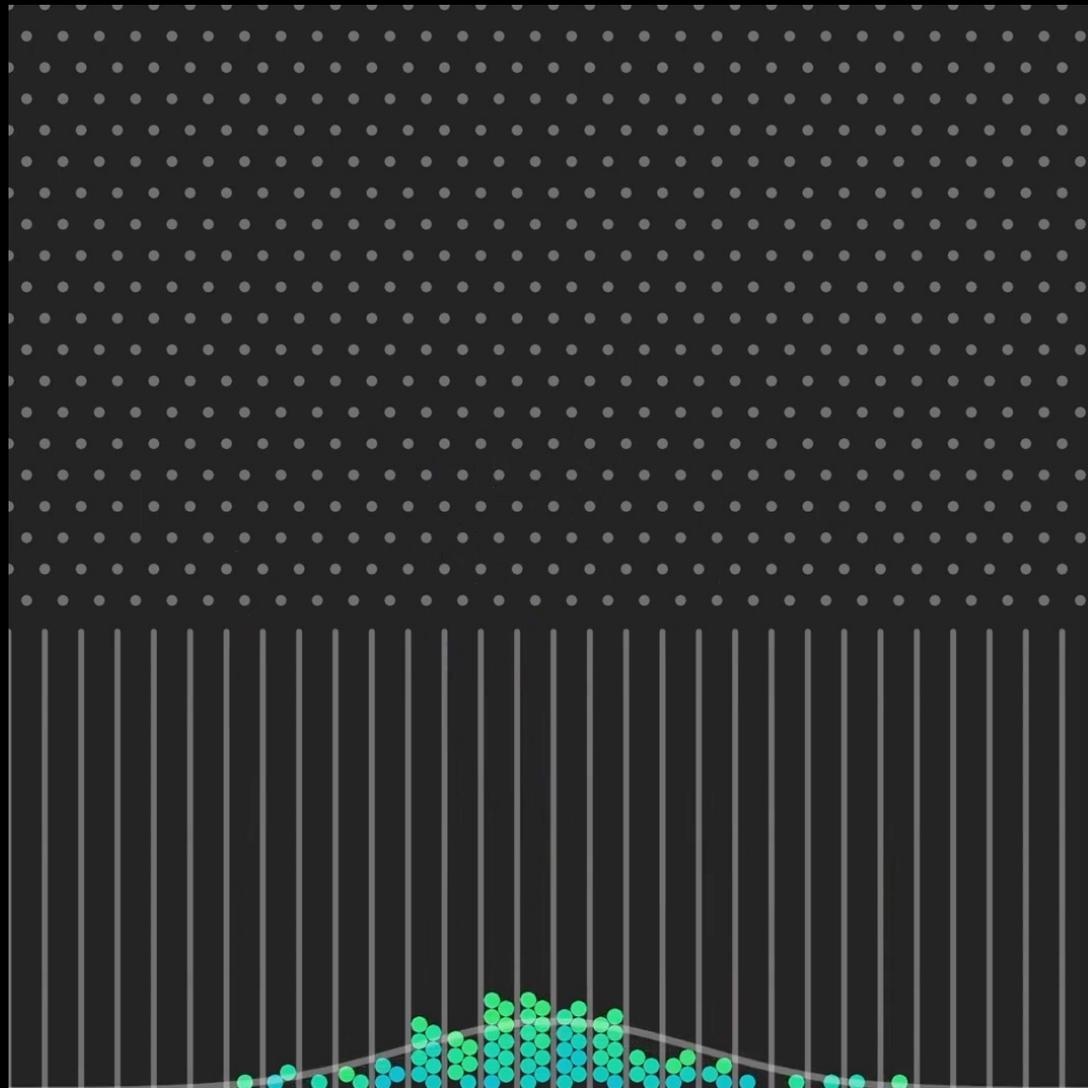
Con 20 lanci intravediamo il fatto che le palline tendono ad accumularsi nel compartimento che sta in mezzo, ma il numero di quelle che cadono in una cella rispetto a quella adiacente oscilla tanto. Troppo pochi lanci (pochi dati)



Aumentiamo allora il numero di lanci...

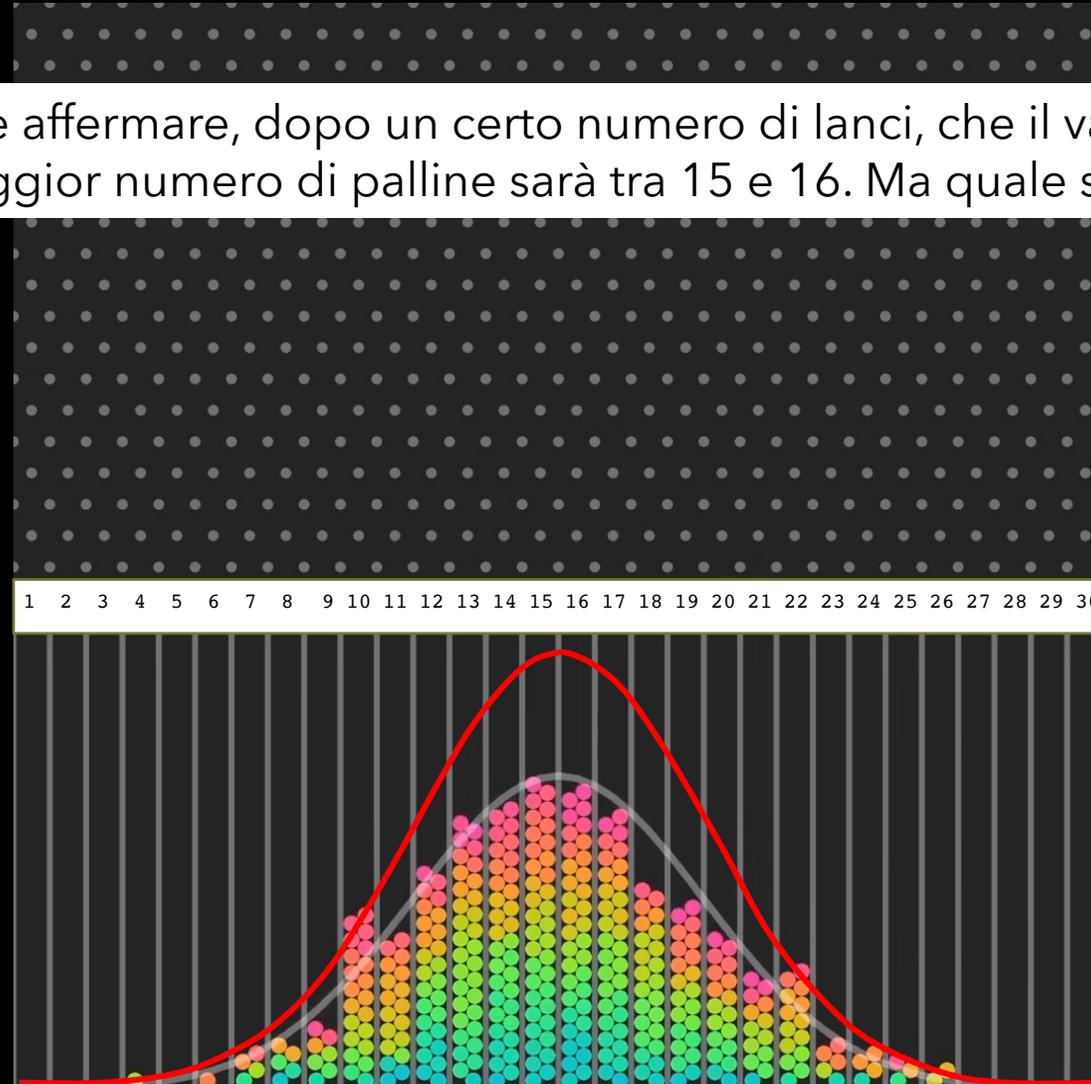
La situazione migliora, nel senso che diventa più chiaro dove vadano ad accumularsi le palline



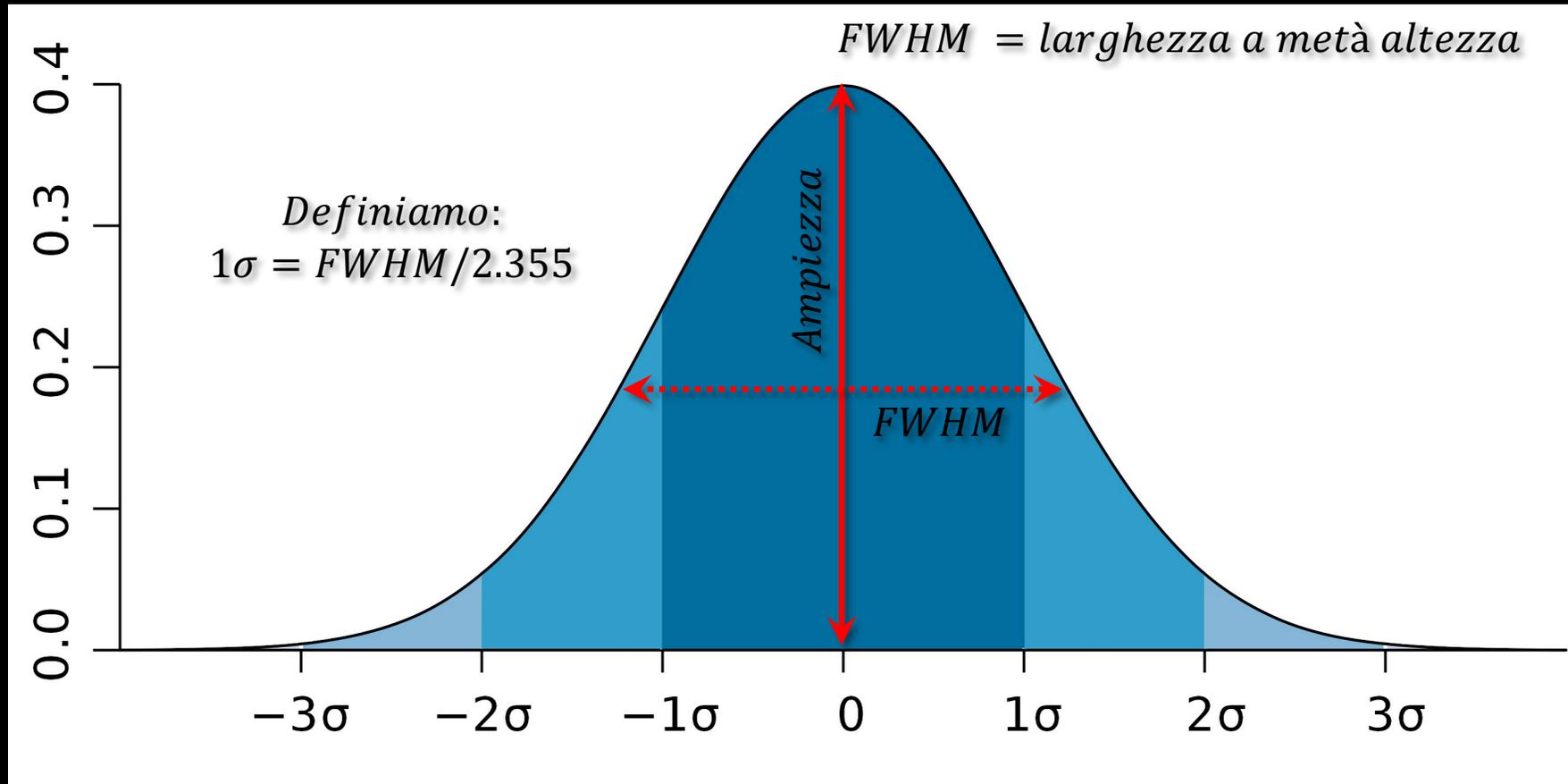


Per quanto migliorata la situazione, resta però evidente che ci sono ancora delle discrepanze significative tra il numero di palline in ogni celletta e il numero che si avrebbe dopo un numero di lanci infinito rappresentato dalla linea rossa continua (immaginando un numero infinito di lanci di palline infinitamente piccole)

Potremo comunque affermare, dopo un certo numero di lanci, che il valore più probabile per la celletta con il maggior numero di palline sarà tra 15 e 16. Ma quale sarà questa probabilità?

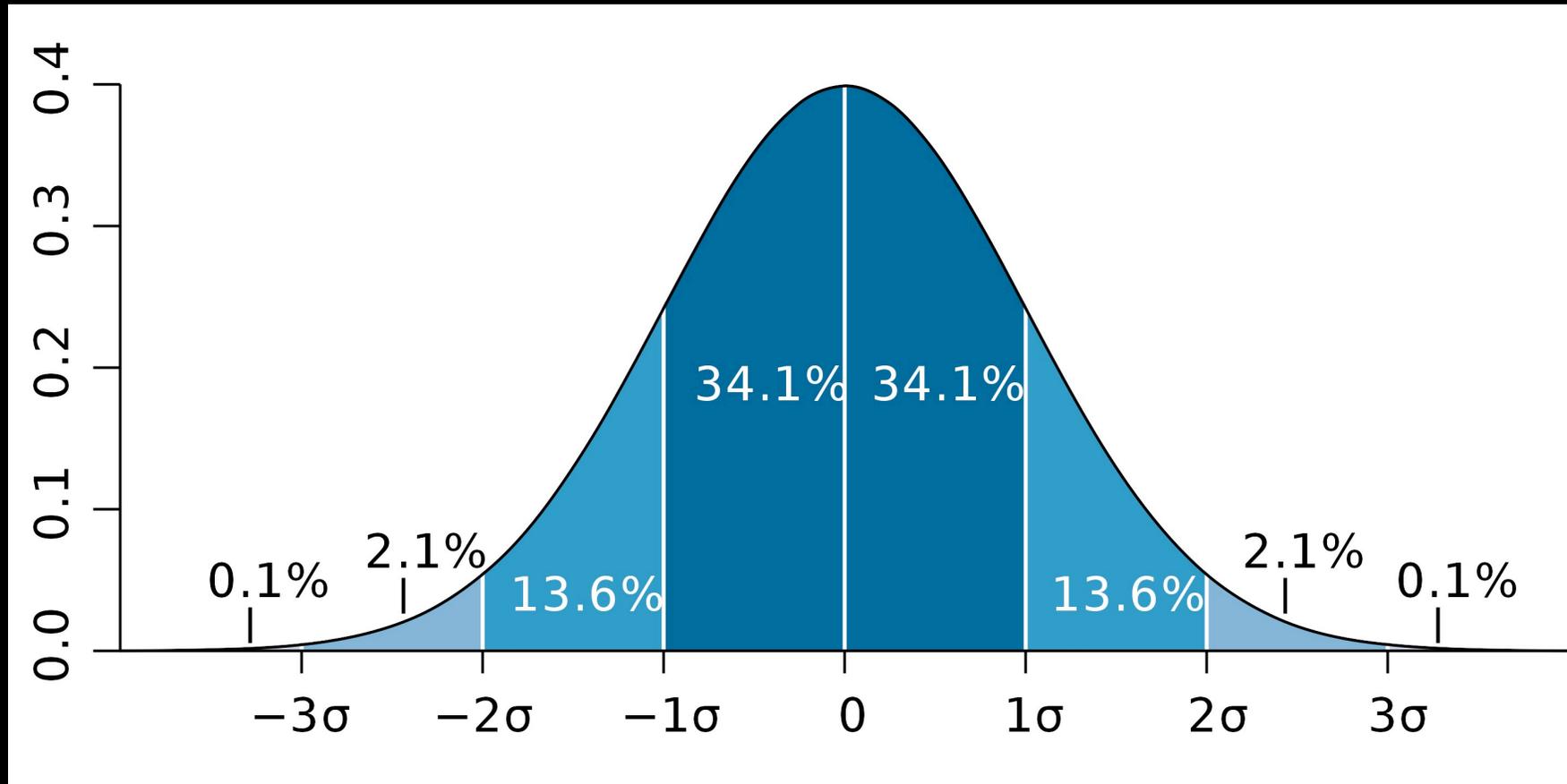


Non voglio tediarvi qui con i dettagli matematici, ma posso dirvi che Gauss fu in grado di ricavare un'equazione (detta gaussiana) che descrive la probabilità che una pallina finisca in una certa celletta. Si tratta di un problema di carattere generale, valido in tutti i casi nei quali il risultato di un fenomeno fisico ripetuto n volte sia indipendente da quello ottenuto in ognuno dei singoli casi



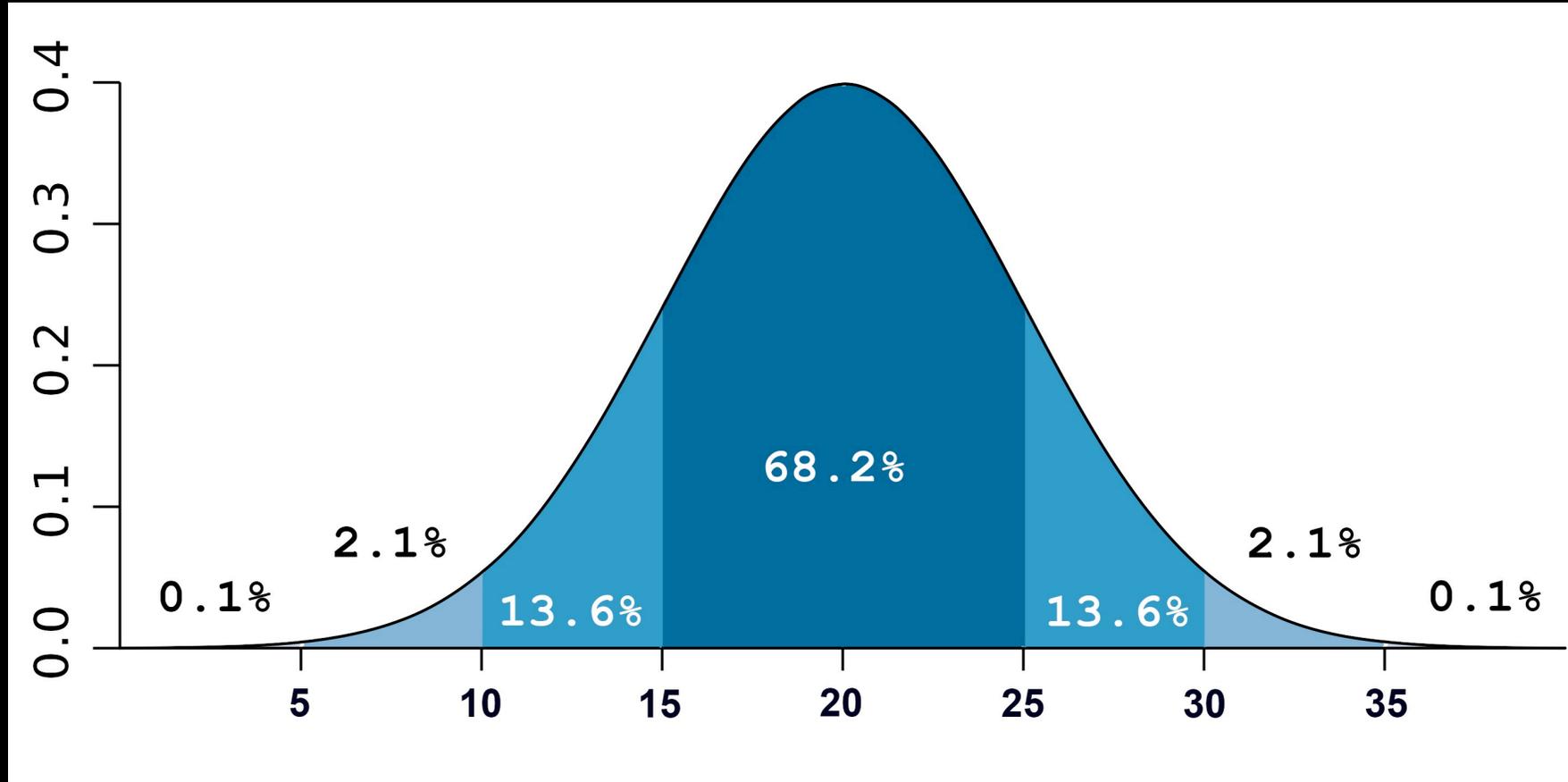
Si dimostra matematicamente che vi è una probabilità del 34.1% che una pallina (il valore di una misura) cada entro 1 sigma, del 13.6 entro due sigma 2 solo del 2.1 % entro tre sigma. Oltre tre sigma la probabilità scende a un valore trascurabile (circa l'uno per mille).

Più sarà stretta questa curva, quindi, più preciso sarà il valore col quale potremo determinare la posizione del valor medio, ossia del valore della quantità che stiamo cercando di misurare



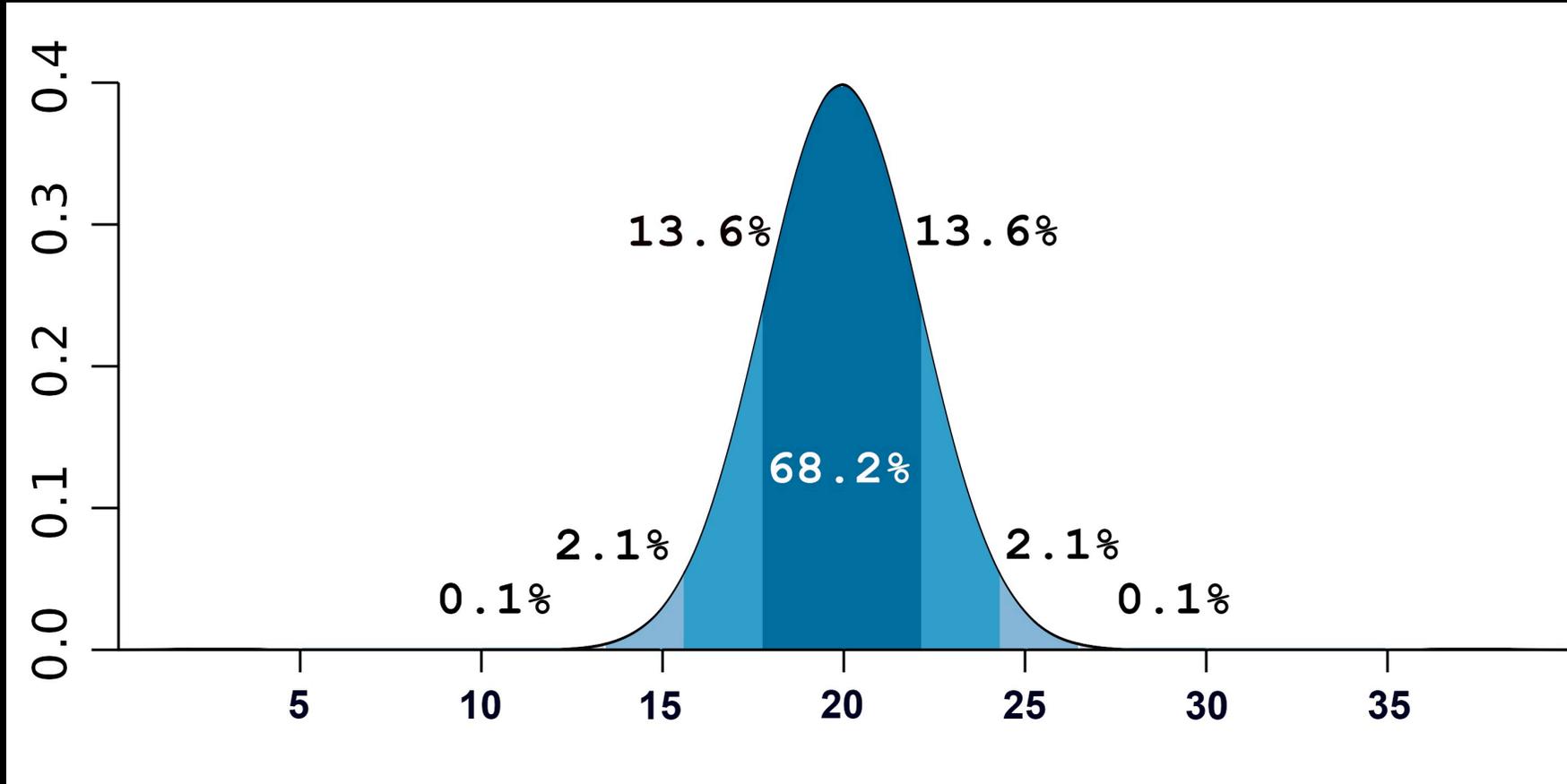
Un esempio: vogliamo misurare la distanza tra due luoghi che sospettiamo siano a circa 20 km uno dall'altro. Se facciamo la misura usando come stima i nostri passi (un sistema davvero molto approssimativo) ripetendo il percorso migliaia di volte troveremo probabilmente una distribuzione di questo tipo.

L'unica cosa che potremo dire, anche con un numero infinito di viaggi avanti e indietro, è che al 68.2% la distanza tra i due luoghi è tra 15 e 25 km. A causa della scarsa precisione dello strumento, non potremo dire di più.



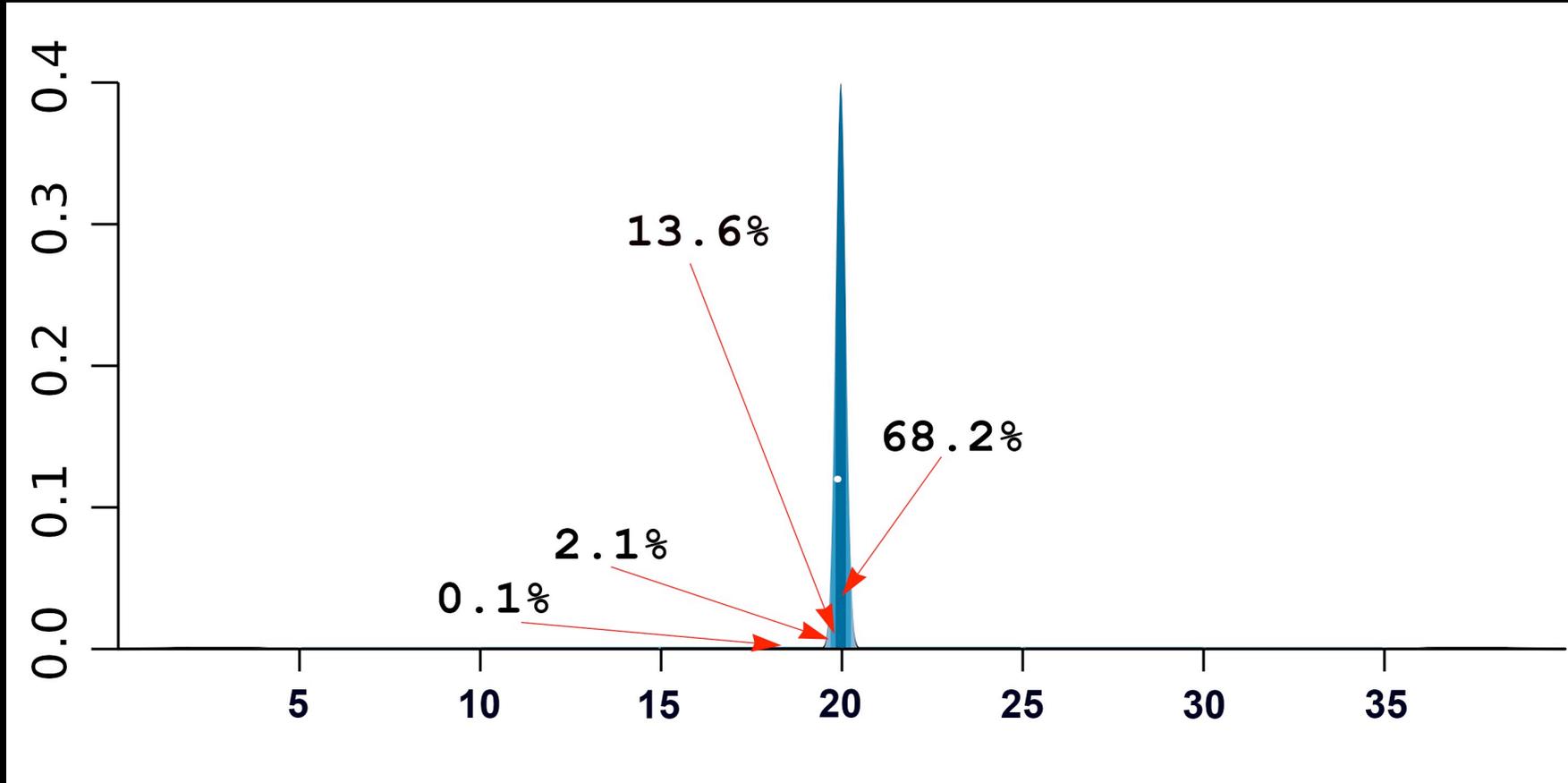
Ma se invece usassimo una ruota che fa un giro ogni tot metri, i valori che misureremmo ad ogni nuovo percorso si distribuirebbero in modo meno sparpagliato: una volta saremmo passati attraverso della sabbia, sbagliando il numero di giri, un'altra attraverso un piccolo dosso e i valori ottenuti sarebbero quindi tra loro leggermente diversi

Comunque, essendo lo strumento un poco più preciso di un piede, potremmo migliorare la stima della distanza dicendo che stavolta abbiamo il 68.2 % di probabilità che essa stia tra 18 e 22 km



Se usassimo invece un metro campione, restringeremmo ancor di più la dispersione delle misure, ottenendo il 99.9% di probabilità che la distanza sia tra 19.5 e 20.5 km.

Non sarà mai possibile misurare con precisione infinita la distanza, per quanto preciso sia lo strumento: ogni minima imprecisione nel prendere la misura si comporrà sommandosi, portando il valore totale ad essere sparpagliato in insieme di possibili valori con un ben precisa probabilità associata ad esso.



Come si rimedia a questo stato di cose, a questa **impossibilità** di effettuare misure che siano del tutto prive di errore?

Lo si fa associando, a ogni misura effettuata, una stima (calcolata tramite un precisa formula matematica chiamata "**deviazione standard**") dell'errore stimato.

Potremo a quel punto fare un'affermazione del tipo:

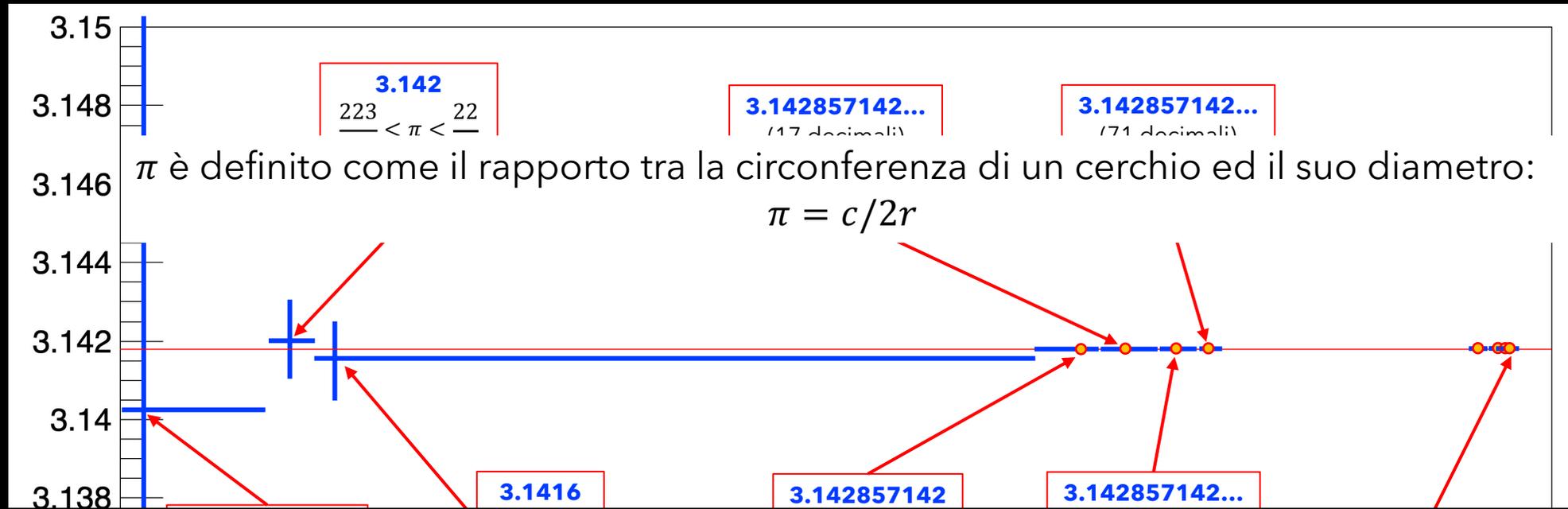
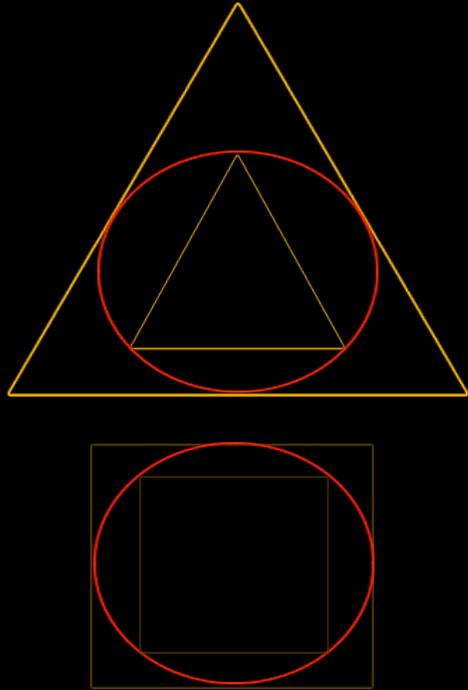
La distanza tra X e Y misurata a piedi ci risulta essere $20 \text{ km} \pm 300 \text{ m}$

La distanza tra X e Y misurata con una ruota ci risulta essere $20 \text{ km} \pm 2 \text{ m}$

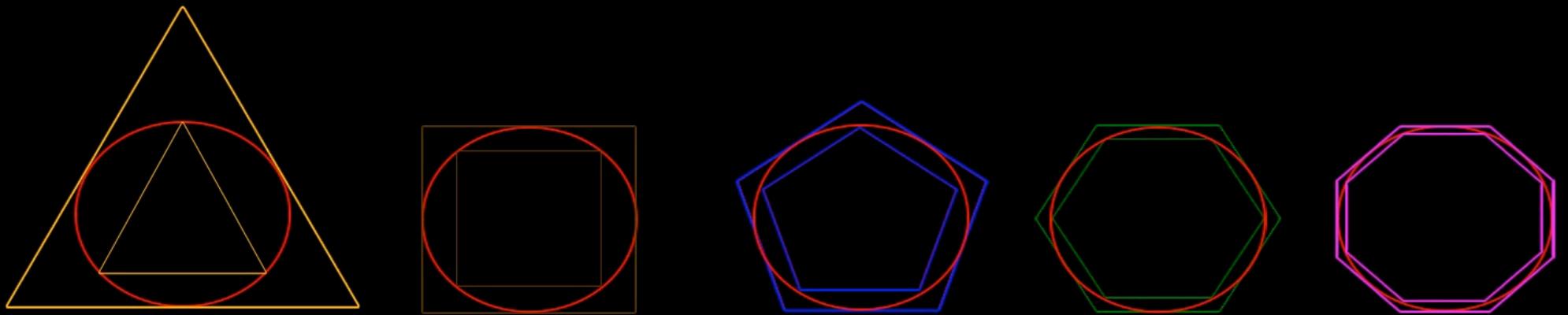
La distanza tra X e Y misurata con un metro campione ci risulta essere $20 \text{ km} \pm 0.05 \text{ m}$

Nessuna affermazione su un fenomeno fisico (pandemie incluse) può essere misurato e valutato con **precisione assoluta**: maggiori le indeterminazioni sul metodo con cui si prendono dati (misure) maggiore l'imprecisione che avremo sul risultato.

Questo discorso vale anche per π : ne conosciamo l'esatta definizione, ma se vogliamo scriverlo su un foglio di carta dobbiamo calcolarne i valori e questo lo si sa fare con un precisione che nel tempo è aumentata vertiginosamente. Come si può calcolare (potremmo dire "misurare sperimentalmente") π ?



π può essere calcolato come il limite della differenza tra i perimetri dei poligoni esterni ed interni alla circonferenza



Ci sono anche altri modi, algebrici e non geometrici per calcolare π

Per quanto ci si possa avvicinare al vero valore di π ,
resterà sempre un'impresione residua, ineliminabile

$$\pi = \frac{4}{1 + \frac{1^2}{2 + \frac{3^2}{2 + \frac{5^2}{2 + \frac{7^2}{2 + \frac{9^2}{2 + \dots}}}}}} = 3 + \frac{1^2}{6 + \frac{3^2}{6 + \frac{5^2}{6 + \frac{7^2}{6 + \frac{9^2}{6 + \dots}}}}$$

$$= \frac{4}{1 + \frac{1^2}{3 + \frac{2^2}{5 + \frac{3^2}{7 + \frac{4^2}{9 + \dots}}}}}$$

3.141592653589793238462643383279502884197169399375105820974944592307816406286208998628034825342117
0679821480865132823066470938446095505822317253594081284811174502841027019385211055596446229489549
3038196442881097566593344612847564823378678316527120190914564856692346034861045432664821339360726
0249141273724587006606315588174881520920962829254091715364367892590360011330530548820466521384146
9519415116094330572703657595919530921861173819326117931051185480744623799627495673518857527248912
2793818301194912983367336244065664308602139494639522473719070217986094370277053921717629317675238
4674818467669405132000568127145263560827785771342757789609173637178721468440901224953430146549585
3710507922796892589235420199561121290219608640344181598136297747713099605187072113499999983729780
4995105973173281609631859502445945534690830264252230825334468503526193118817101000313783875288658
7533208381420617177669147303598253490428755468731159562863882353787593751957781857780532171226806
6130019278766111959092164201989380952572010654858632788659361533818279682303019520353018529689957
7362259941389124972177528347913151557485724245415069595082953311686172785588907509838175463746493
9319255060400927701671139009848824012858361603563707660104710181942955596198946767837449448255379
7747268471040475346462080466842590694912933136770289891521047521620569660240580381501935112533824
3003558764024749647326391419927260426992279678235478163600934172164121992458631503028618297455570
6749838505494588586926995690927210797509302955321165344987202755960236480665499119881834797753566
369807426542527862551818417574672890977727938000816470600161452491921732172147723501414419735685
4816136115735255213347574184946843852332390739414333454776241686251898356948556209921922218427255
0254256887671790494601653466804988627232791786085784383827967976681454100953883786360950680064225
1252051173929848960841284886269456042419652850222106611863067442786220391949450471237137869609563
6437191728746776465757396241389086583264599581339047802759009946576407895126946839835259570982582
2620522489407726719478268482601476990902640136394437455305068203496252451749399651431429809190659
2509372216964615157098583874105978859597729754989301617539284681382686838689427741559918559252459
5395943104997252468084598727364469584865383673622262609912460805124388439045124413654976278079771
5691435997700129616089441694868555848406353422072225828488648158456028506016842739452267467678895
2521385225499546667278239864565961163548862305774564980355936345681743241125150760694794510965960
9402522887971089314566913686722874894056010150330861792868092087476091782493858900971490967598526
1365549781893129784821682998948722658804857564014270477555132379641451523746234364542858444795265

Come gestire i risultati di misure usando la matematica

La prima cosa che occorre fare per capire cosa succede è raccogliere i dati: a questo ci pensano enti come l'Istituto Superiore di Sanità (ISS), il Dipartimento della Protezione Civile (DPC) e altre istituzioni preposte con un mandato.

Questi dati devono essere accessibili in formato elettronico da un repository reso pubblico dal quale possano essere prelevati e analizzati.

Una volta prelevati però cosa possiamo farcene? In genere si tratta di documenti di milioni di righe, dai quali è, in pratica, impossibile trarre qualsiasi conclusione, anche se i dati sono perfettamente leggibili da un essere umano.

Occorre una altro approccio, un'opportuna miscela di matematica, di grafica e di multidisciplinarietà.

Cosa sia un virus lo sanno i virologi, come si propaghi lo sanno gli epidemiologi, come reagiscono gli essere umani lo sanno i medici e i sociologi. Come trattare grandi moli di dati usando metodologie statistiche corrette lo sanno invece i matematici e i fisici.

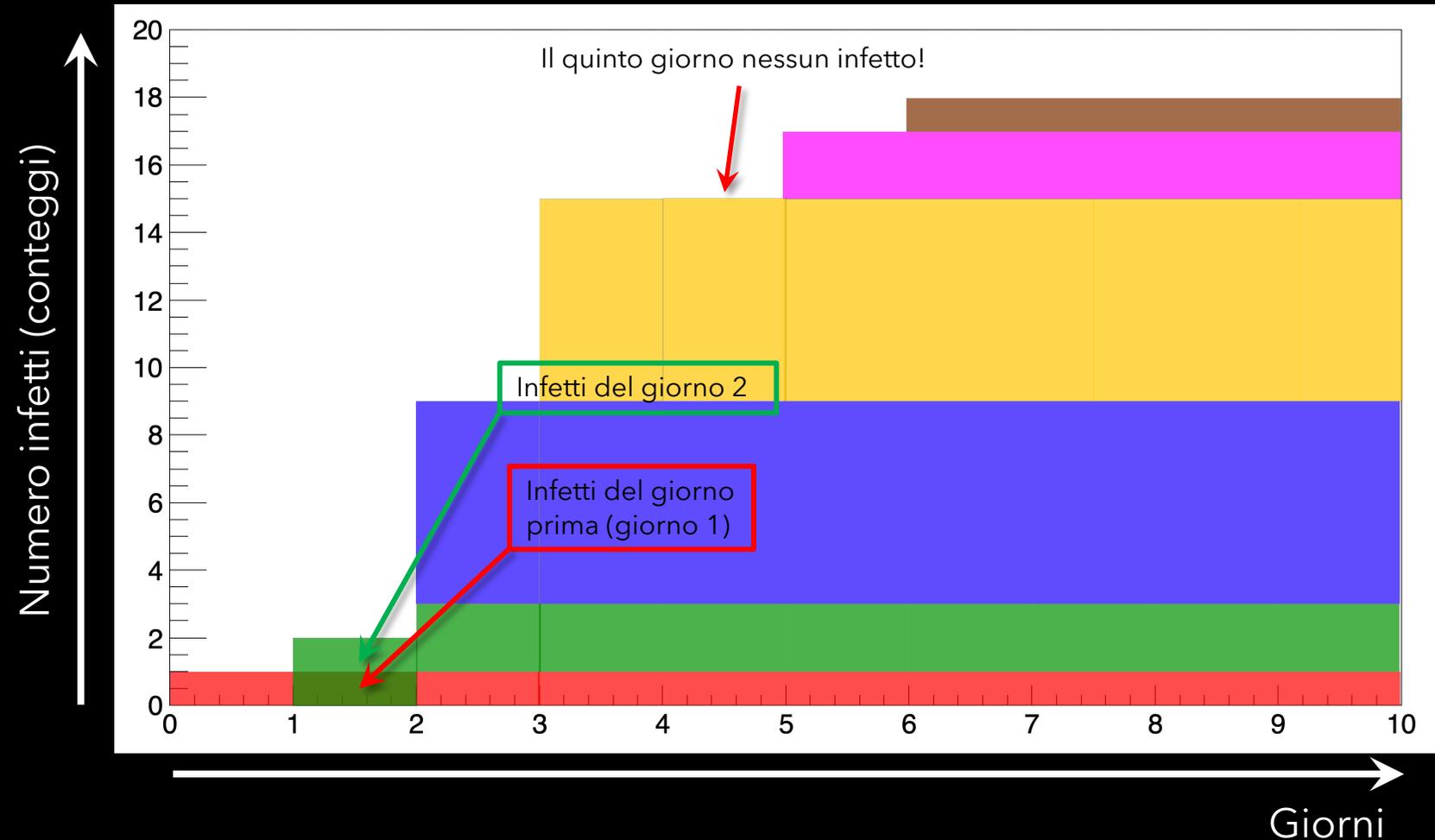
È dall'insieme di queste competenze multidisciplinari che si possono distillare modelli da studiare e verificare se i dati ne confermano la validità.

```
2 2021-02-16T17:00:00,ITA,13,Abruzzo,066,L'Aquila,AQ,42.35122196,13.39843823,12739,,ITF,ITF1,ITF11
3 2021-02-16T17:00:00,ITA,13,Abruzzo,067,Teramo,TE,42.6589177,13.70439971,11887,,ITF,ITF1,ITF12
4 2021-02-16T17:00:00,ITA,13,Abruzzo,068,Pescara,PE,42.46458398,14.21364822,11900,,ITF,ITF1,ITF13
5 2021-02-16T17:00:00,ITA,13,Abruzzo,069,Chieti,CH,42.35103167,14.16754574,11683,,ITF,ITF1,ITF14
6 2021-02-16T17:00:00,ITA,13,Abruzzo,879,Fuori Regione / Provincia Autonoma,,,,,407,,ITF,ITF1
7 2021-02-16T17:00:00,ITA,13,Abruzzo,979,In fase di definizione/aggiornamento,,,,,175,,ITF,ITF1
8 2021-02-16T17:00:00,ITA,17,Basilicata,076,Potenza,PZ,40.63947052,15.80514834,9527,,ITF,ITF5,ITF51
9 2021-02-16T17:00:00,ITA,17,Basilicata,077,Matera,MT,40.66751177,16.59792442,4410,,ITF,ITF5,ITF52
10 2021-02-16T17:00:00,ITA,17,Basilicata,880,Fuori Regione / Provincia Autonoma,,,,,423,,ITF,ITF5
11 2021-02-16T17:00:00,ITA,17,Basilicata,980,In fase di definizione/aggiornamento,,,,,0,,ITF,ITF5
12 2021-02-16T17:00:00,ITA,18,Calabria,078,Cosenza,CS,39.29308681,16.25609692,10425,,ITF,ITF6,ITF61
13 2021-02-16T17:00:00,ITA,18,Calabria,079,Catanzaro,CZ,38.90597598,16.59440194,5097,,ITF,ITF6,ITF63
14 2021-02-16T17:00:00,ITA,18,Calabria,080,Reggio di Calabria,RC,38.10922769,15.643452700000001,13895,,ITF,ITF6,ITF65
15 2021-02-16T17:00:00,ITA,18,Calabria,101,Crotone,CR,39.0806878,17.12538864,2719,,ITF,ITF6,ITF62
16 2021-02-16T17:00:00,ITA,18,Calabria,102,Vibo Valentia,VV,38.67624147,16.10157414,3190,,ITF,ITF6,ITF64
17 2021-02-16T17:00:00,ITA,18,Calabria,882,Fuori Regione / Provincia Autonoma,,,,,359,,ITF,ITF6
18 2021-02-16T17:00:00,ITA,18,Calabria,982,In fase di definizione/aggiornamento,,,,,0,,ITF,ITF6
19 2021-02-16T17:00:00,ITA,15,Campania,061,Caserta,CE,41.07465878,14.53240464,41579,,ITF,ITF3
20 2021-02-16T17:00:00,ITA,15,Campania,062,Benevento,BN,41.12969987,14.78151683,6562,,ITF,ITF3,ITF32
21 2021-02-16T17:00:00,ITA,15,Campania,063,Napoli,NA,40.83956555,14.25084984,147771,,ITF,ITF3,ITF33
22 2021-02-16T17:00:00,ITA,15,Campania,064,Avellino,AV,40.91404699,14.79528803,10792,,ITF,ITF3,ITF34
23 2021-02-16T17:00:00,ITA,15,Campania,065,Salerno,SA,40.67821961,14.759402599999999,36382,,ITF,ITF3,ITF35
24 2021-02-16T17:00:00,ITA,15,Campania,883,Fuori Regione / Provincia Autonoma,,,,,0,,ITF,ITF3
25 2021-02-16T17:00:00,ITA,15,Campania,983,In fase di definizione/aggiornamento,,,,,1931,,ITF,ITF3
26 2021-02-16T17:00:00,ITA,08,Emilia-Romagna,033,Piacenza,PC,45.05193462,9.692632596000001,18665,,ITH,ITH5,ITH51
27 2021-02-16T17:00:00,ITA,08,Emilia-Romagna,034,Parma,PR,44.60107394,10.32834985,16083,,ITH,ITH5,ITH52
28 2021-02-16T17:00:00,ITA,08,Emilia-Romagna,035,Reggio nell'Emilia,RE,44.69735289,10.63007973,31091,,ITH,ITH5,ITH53
29 2021-02-16T17:00:00,ITA,08,Emilia-Romagna,036,Modena,MO,44.64600009,10.92615487,41079,,ITH,ITH5,ITH54
30 2021-02-16T17:00:00,ITA,08,Emilia-Romagna,037,Bologna,BO,44.49436681,11.341720800000001,54308,,ITH,ITH5,ITH55
31 2021-02-16T17:00:00,ITA,08,Emilia-Romagna,038,Ferrara,FE,44.83599085,11.61866934,13858,,ITH,ITH5,ITH56
32 2021-02-16T17:00:00,ITA,08,Emilia-Romagna,039,Ravenna,RA,44.41722493,12.19913936,17780,,ITH,ITH5,ITH57
33 2021-02-16T17:00:00,ITA,08,Emilia-Romagna,040,Fiorlana-Cesena,FC,44.22268559,12.04068608,19639,,ITH,ITH5,ITH58
34 2021-02-16T17:00:00,ITA,08,Emilia-Romagna,099,Rimini,RN,44.06090087,12.5656295,21941,,ITH,ITH5,ITH59
35 2021-02-16T17:00:00,ITA,08,Emilia-Romagna,884,Fuori Regione / Provincia Autonoma,,,,,2578,,ITH,ITH5
36 2021-02-16T17:00:00,ITA,08,Emilia-Romagna,984,In fase di definizione/aggiornamento,,,,,1052,,ITH,ITH5
37 2021-02-16T17:00:00,ITA,06,Friuli Venezia Giulia,030,Udine,UD,46.06255516,13.2348383,32465,,ITH,ITH4,ITH42
38 2021-02-16T17:00:00,ITA,06,Friuli Venezia Giulia,031,Gorizia,GO,45.94149817,13.62212502,8596,,ITH,ITH4,ITH43
39 2021-02-16T17:00:00,ITA,06,Friuli Venezia Giulia,032,Trieste,TS,45.6494354,13.76813649,14229,,ITH,ITH4,ITH44
40 2021-02-16T17:00:00,ITA,06,Friuli Venezia Giulia,033,Pordenone,PN,45.95443546,12.66002909,16301,,ITH,ITH4,ITH41
41 2021-02-16T17:00:00,ITA,06,Friuli Venezia Giulia,885,Fuori Regione / Provincia Autonoma,,,,,857,,ITH,ITH4
42 2021-02-16T17:00:00,ITA,06,Friuli Venezia Giulia,985,In fase di definizione/aggiornamento,,,,,0,,ITH,ITH4
43 2021-02-16T17:00:00,ITA,12,Lazio,056,Viterbo,VT,42.4173828,12.10473416,11468,,ITI,ITI4,ITI41
44 2021-02-16T17:00:00,ITA,12,Lazio,057,Rieti,RI,42.40488444,12.86205939,7001,,ITI,ITI4,ITI42
45 2021-02-16T17:00:00,ITA,12,Lazio,058,Roma,RM,41.89277044,12.48366722,157447,,ITI,ITI4,ITI43
46 2021-02-16T17:00:00,ITA,12,Lazio,059,Latina,LT,41.46759465,12.90368482,22194,,ITI,ITI4,ITI44
47 2021-02-16T17:00:00,ITA,12,Lazio,060,Frosinone,FR,41.63964569,13.35117161,18594,,ITI,ITI4,ITI45
48 2021-02-16T17:00:00,ITA,12,Lazio,886,Fuori Regione / Provincia Autonoma,,,,,4234,,ITI,ITI4
49 2021-02-16T17:00:00,ITA,12,Lazio,986,In fase di definizione/aggiornamento,,,,,21,,ITI,ITI4
50 2021-02-16T17:00:00,ITA,07,Liguria,008,Imperia,IM,43.88570648,8.027850297999999,9074,,ITC,ITC3,ITC31
51 2021-02-16T17:00:00,ITA,07,Liguria,009,Savona,SV,44.30750461,8.481108654,10818,,ITC,ITC3,ITC32
52 2021-02-16T17:00:00,ITA,07,Liguria,010,Genova,GE,44.41149314,8.9326992,40570,,ITC,ITC3,ITC33
53 2021-02-16T17:00:00,ITA,07,Liguria,011,La Spezia,SP,44.10704991,9.8281897,10175,,ITC,ITC3,ITC34
54 2021-02-16T17:00:00,ITA,07,Liguria,887,Fuori Regione / Provincia Autonoma,,,,,1364,,ITC,ITC3
55 2021-02-16T17:00:00,ITA,07,Liguria,987,In fase di definizione/aggiornamento,,,,,2229,,ITC,ITC3
56 2021-02-16T17:00:00,ITA,03,Lombardia,012,Varese,VA,45.81701677,8.822868344,60349,,ITC,ITC4,ITC41
57 2021-02-16T17:00:00,ITA,03,Lombardia,013,Como,CO,45.8099912,9.085159546,40922,,ITC,ITC4,ITC42
58 2021-02-16T17:00:00,ITA,03,Lombardia,014,Sondrio,SO,46.17099261,9.87147489,10596,,ITC,ITC4,ITC44
59 2021-02-16T17:00:00,ITA,03,Lombardia,015,Milano,MI,45.46679409,9.190347404,196245,,ITC,ITC4,ITC4C
60 2021-02-16T17:00:00,ITA,03,Lombardia,016,Bergamo,BG,45.69441368,9.668424528,32518,,ITC,ITC4,ITC46
61 2021-02-16T17:00:00,ITA,03,Lombardia,017,Brescia,BS,45.53993052,10.21910323,58709,,ITC,ITC4,ITC47
62 2021-02-16T17:00:00,ITA,03,Lombardia,018,Pavia,PA,45.18509264,9.160157191,29918,,ITC,ITC4,ITC48
63 2021-02-16T17:00:00,ITA,03,Lombardia,019,Cremona,CR,45.13336675,10.02420865,16866,,ITC,ITC4,ITC4A
64 2021-02-16T17:00:00,ITA,03,Lombardia,020,Mantova,MN,45.15726772,10.79277363,21849,,ITC,ITC4,ITC4B
65 2021-02-16T17:00:00,ITA,03,Lombardia,097,Lecco,LC,45.85575781,9.393392246,16214,,ITC,ITC4,ITC43
66 2021-02-16T17:00:00,ITA,03,Lombardia,098,Lodi,LO,45.31440693,9.503720769,12416,,ITC,ITC4,ITC49
67 2021-02-16T17:00:00,ITA,03,Lombardia,108,Monza e della Brianza,MB,45.58439043,9.273582472000001,54082,,ITC,ITC4,ITC4
68 2021-02-16T17:00:00,ITA,03,Lombardia,888,Fuori Regione / Provincia Autonoma,,,,,6667,,ITC,ITC4
69 2021-02-16T17:00:00,ITA,03,Lombardia,988,In fase di definizione/aggiornamento,,,,,8368,,ITC,ITC4
70 2021-02-16T17:00:00,ITA,11,Marche,041,Pesaro e Urbino,PU,43.91014021,12.91345989,13984,,ITI,ITI3,ITI31
```

Vediamo anzitutto cosa significa "rappresentare i dati mediante un istogramma":

cominciamo con quello che si chiama "**istogramma cumulativo**"

In un istogramma di questo genere la curva dei dati sale sempre, i valori si accumulano nel tempo



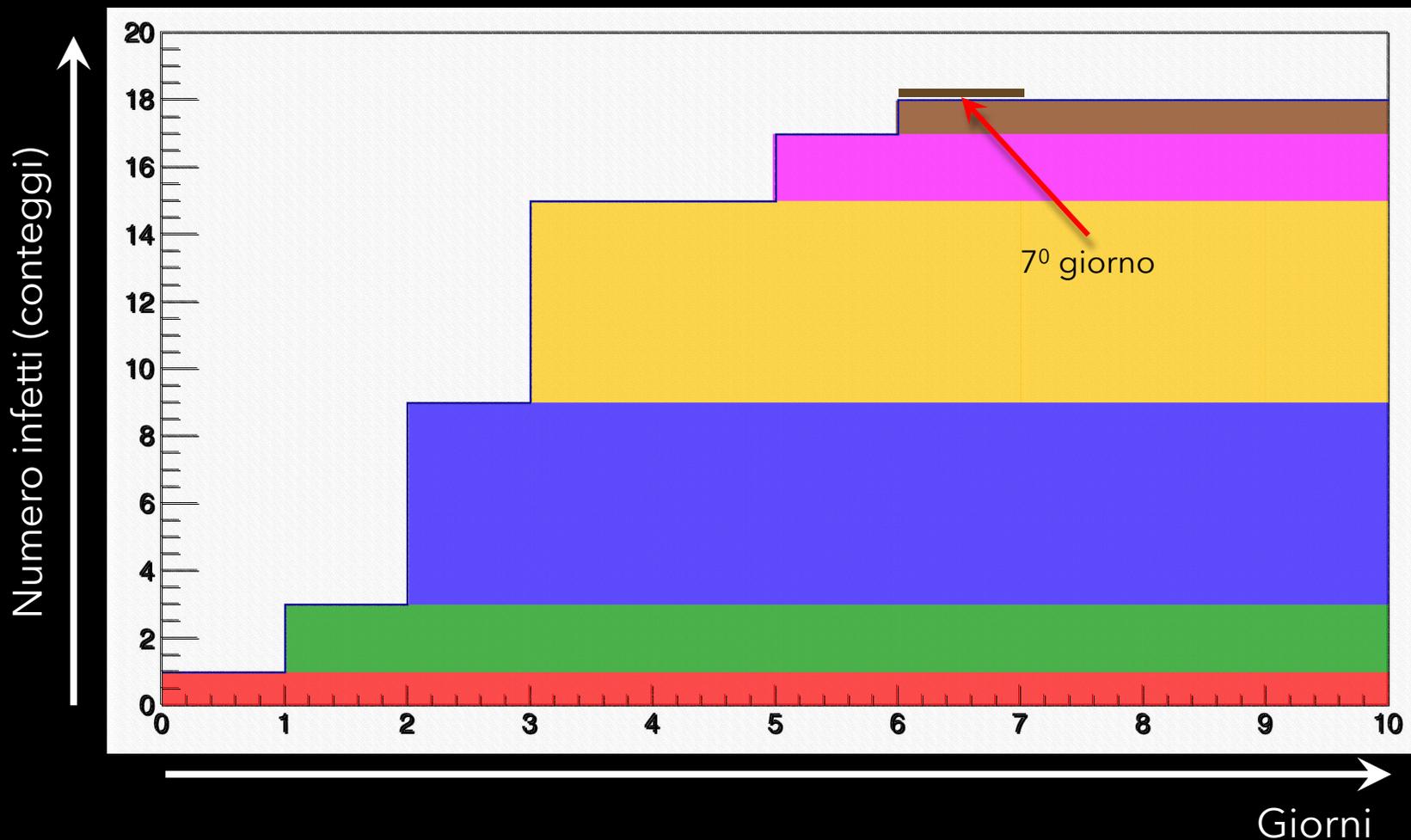
Dieci giorni di raccolta dei dati potrebbero quindi mostrare un andamento di questo genere

Dal settimo giorno in avanti i dati non registrano nuovi infetti: all'istogramma non si aggiungono nuovi contenuti e si sarebbe tentati di pensare che la curva dovrebbe, di conseguenza, mostrare una *discesa*.

Trattandosi di una curva cumulativa, invece, la cessazione di nuovi infetti NON si manifesta come una discesa dei dati, bensì mediante un *andare a regime*, un *approssimarsi ad un plateau* di 18 casi.

Esiste però un modo alternativo di rappresentare questi stessi identici dati: tramite un grafico di tipo *differenziale*.

Vediamo un esempio esplicativo



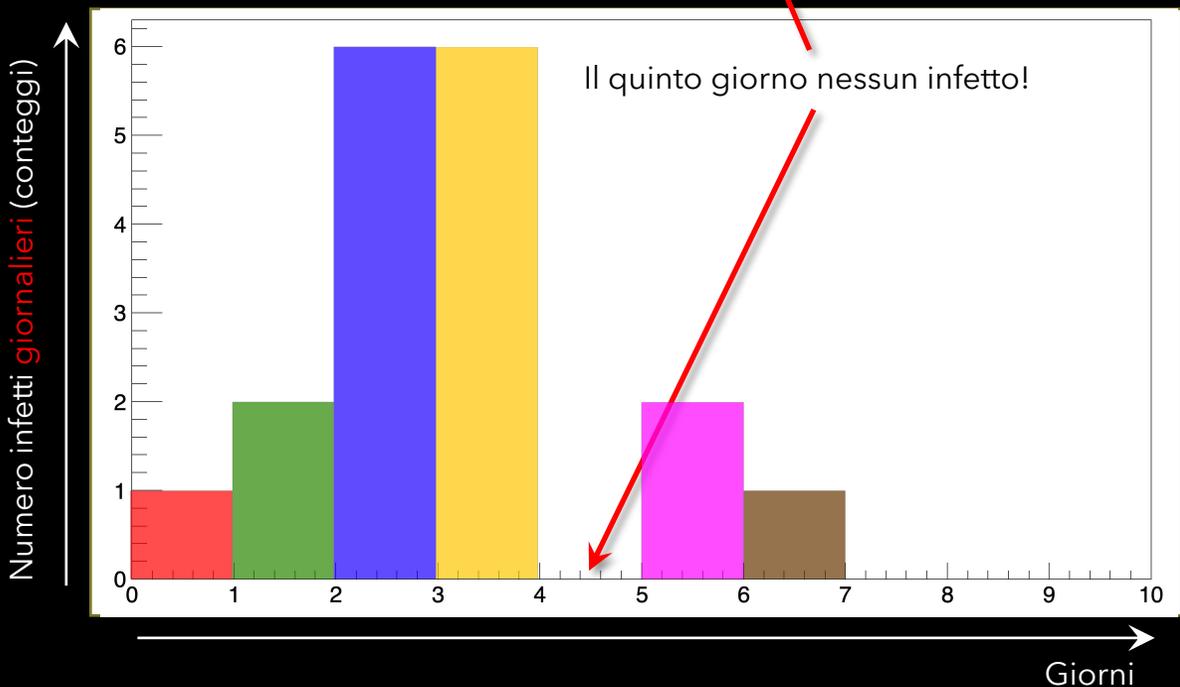
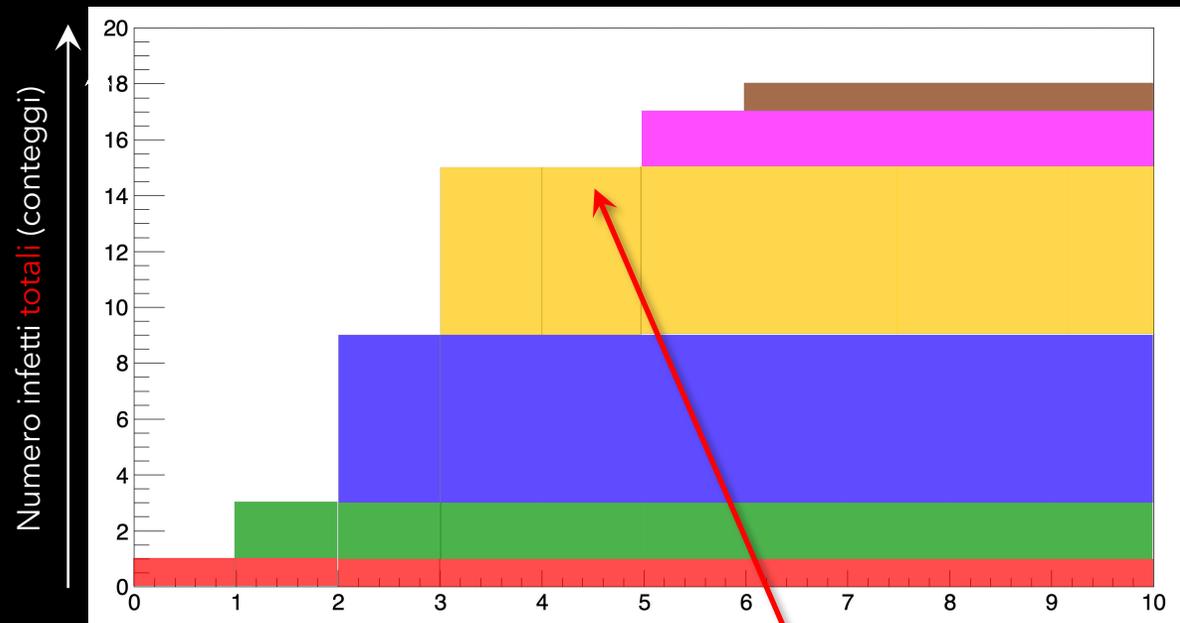
Il grafico cumulativo aumenta sempre, mentre quello differenziale può salire ma anche scendere.

Quando la situazione non cambia da un giorno all'altro, nel grafico cumulativo la curva si fa orizzontale mentre in quello differenziale si ha un *buco*, con un crollo verticale dal giorno prima.

Il grafico cumulativo ci fornisce quindi la situazione accumulata nel tempo di quel che è accaduto in quei giorni, mentre il differenziale ci mostra il vero andamento giornaliero, con aumenti e decrescite.

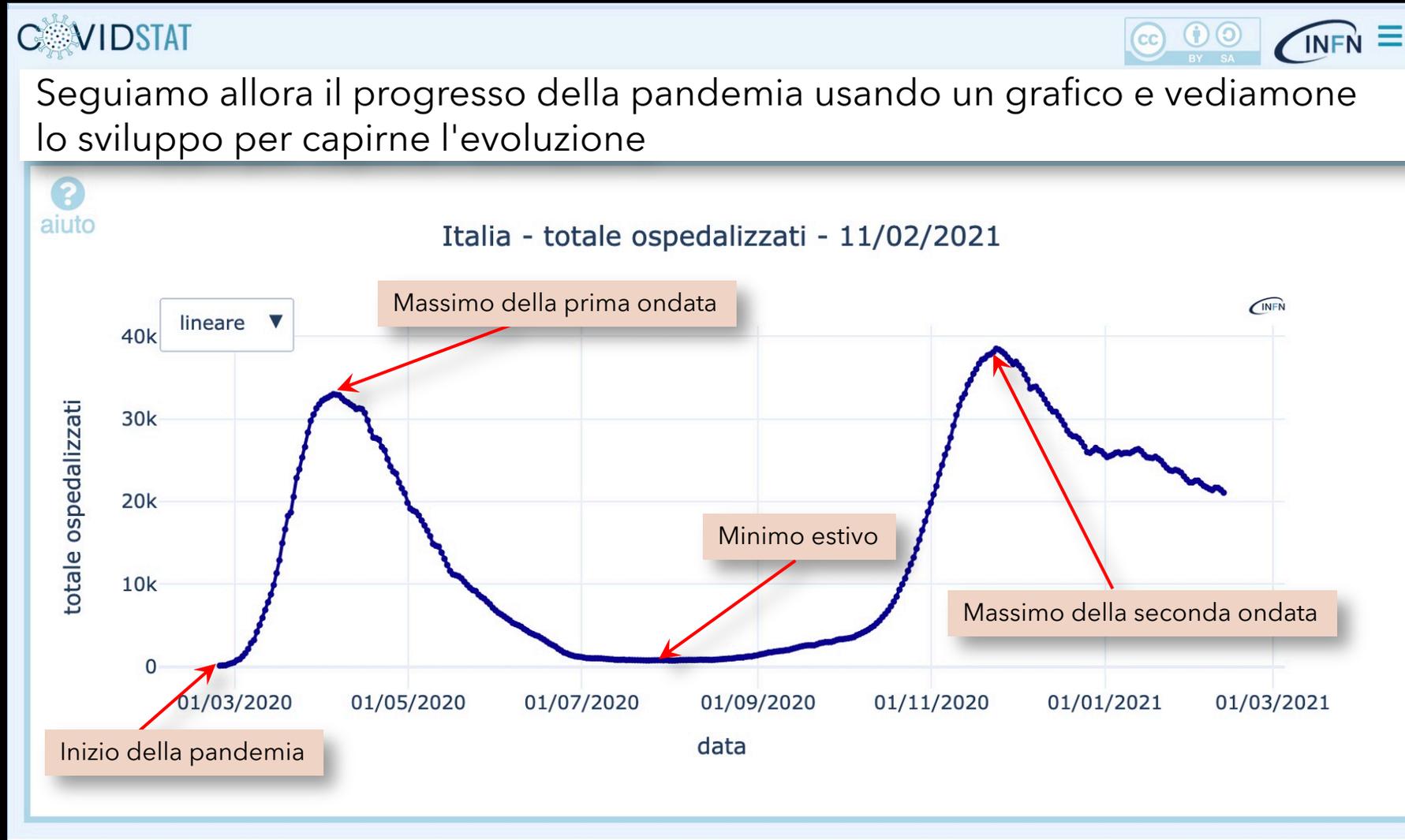
Dalle nostre reminiscenze alle superiori: il grafico di sopra corrisponde all'**integrale** di quello di sotto.

Quello di sotto mostra invece la derivata di quello di sopra. La derivata ci mostra **la variazione nel tempo** di un fenomeno e verrà utilizzata quando questo è ciò che si vuole evidenziare. L'integrale invece mostra **l'accumulo** finale del fenomeno, sull'asse verticale (quanti infetti ci sono stati in quel periodo): questo verrà usato nei contesti in cui è importante il totale.



Un modo più efficiente di analizzare i dati è quello di trasformare i numeri in un grafico (in istogrammi, in gergo).

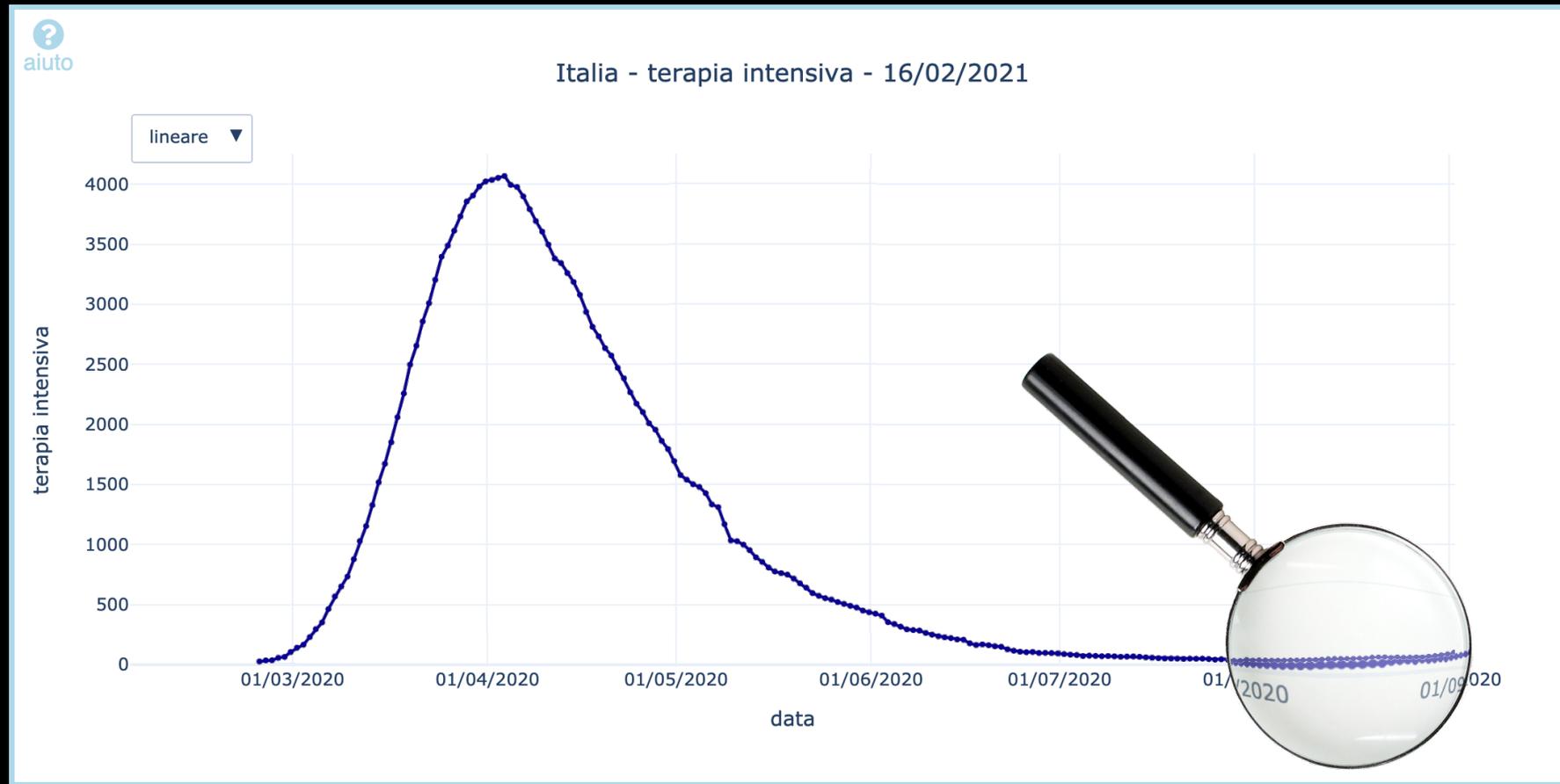
Questo grafico mostra, ad esempio, il numero di ospedalizzati per ogni dato giorno: il tempo sta sull'asse orizzontale mentre invece il numero di casi, ossia i conteggi quotidiani viene posto sull'asse verticale.



Un certo giorno iniziano ad giungere sempre più malati in ospedale che manifestano i sintomi di una malattia sconosciuta. I medici iniziano a registrare questi casi contandoli e qualcuno inizia a osservare il loro andamento...

A partire dal 24 Febbraio 2020 hanno iniziato ad aumentare sempre più velocemente i ricoveri in terapia intensiva.

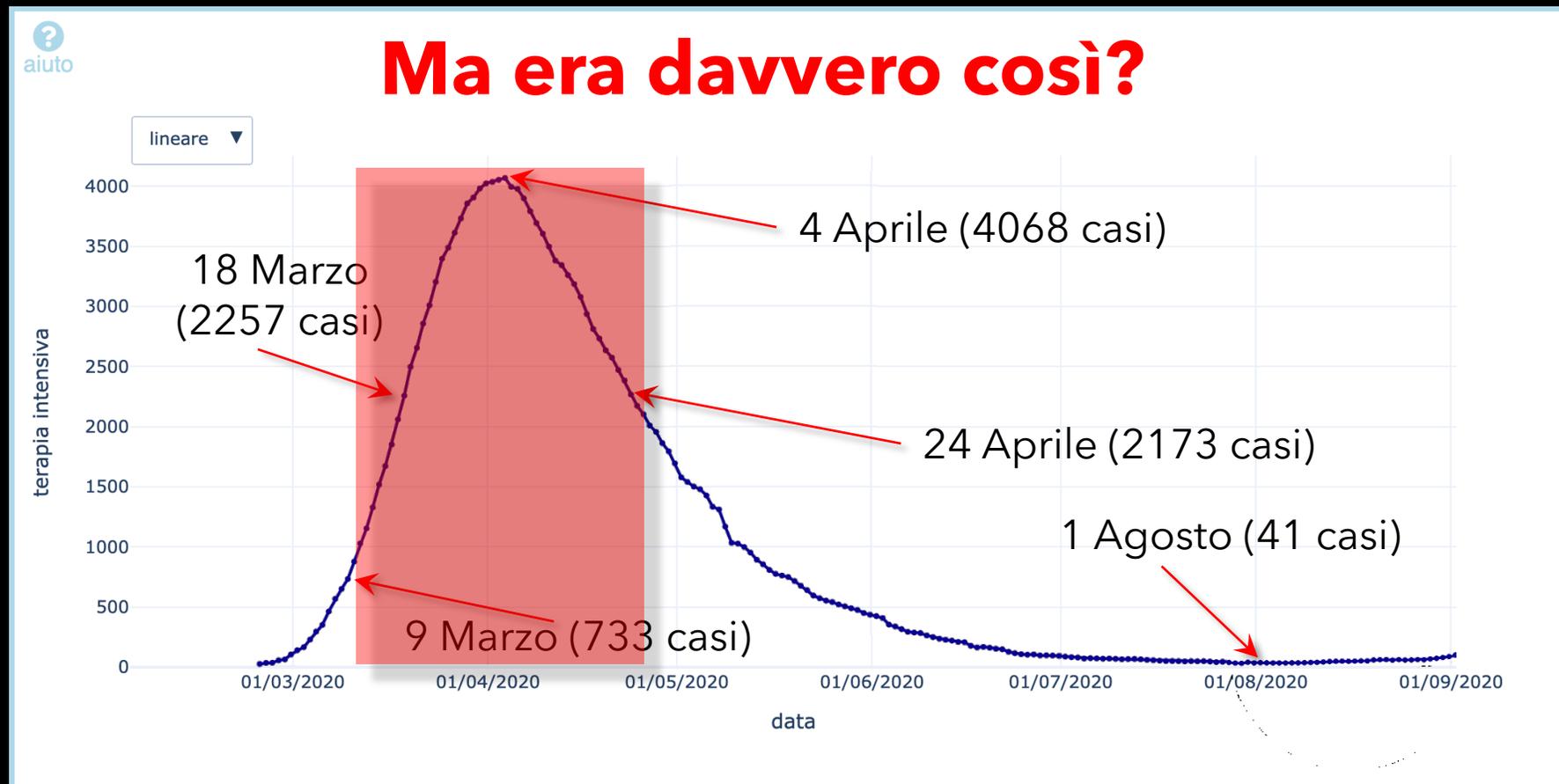
Dalla seconda metà di Agosto, guardando questo grafico, molti dissero "il virus è virtualmente morto!..."



Iniziamo col dire che fortunatamente il Governo aveva imposto il lock-down generale il 9 Marzo: ci vollero 10 giorni (fino al 18 Marzo per osservare un cambio di pendenza della crescita, segno che iniziava ad avere effetto il blocco) Cambiato ritmo, l'inerzia della progressione di ricoveri impiegò fino al 4 Aprile per raggiungere il massimo di 4068.

Da quel giorno iniziò una decrescita, all'inizio veloce e poi, dal 24 Aprile, più lenta.

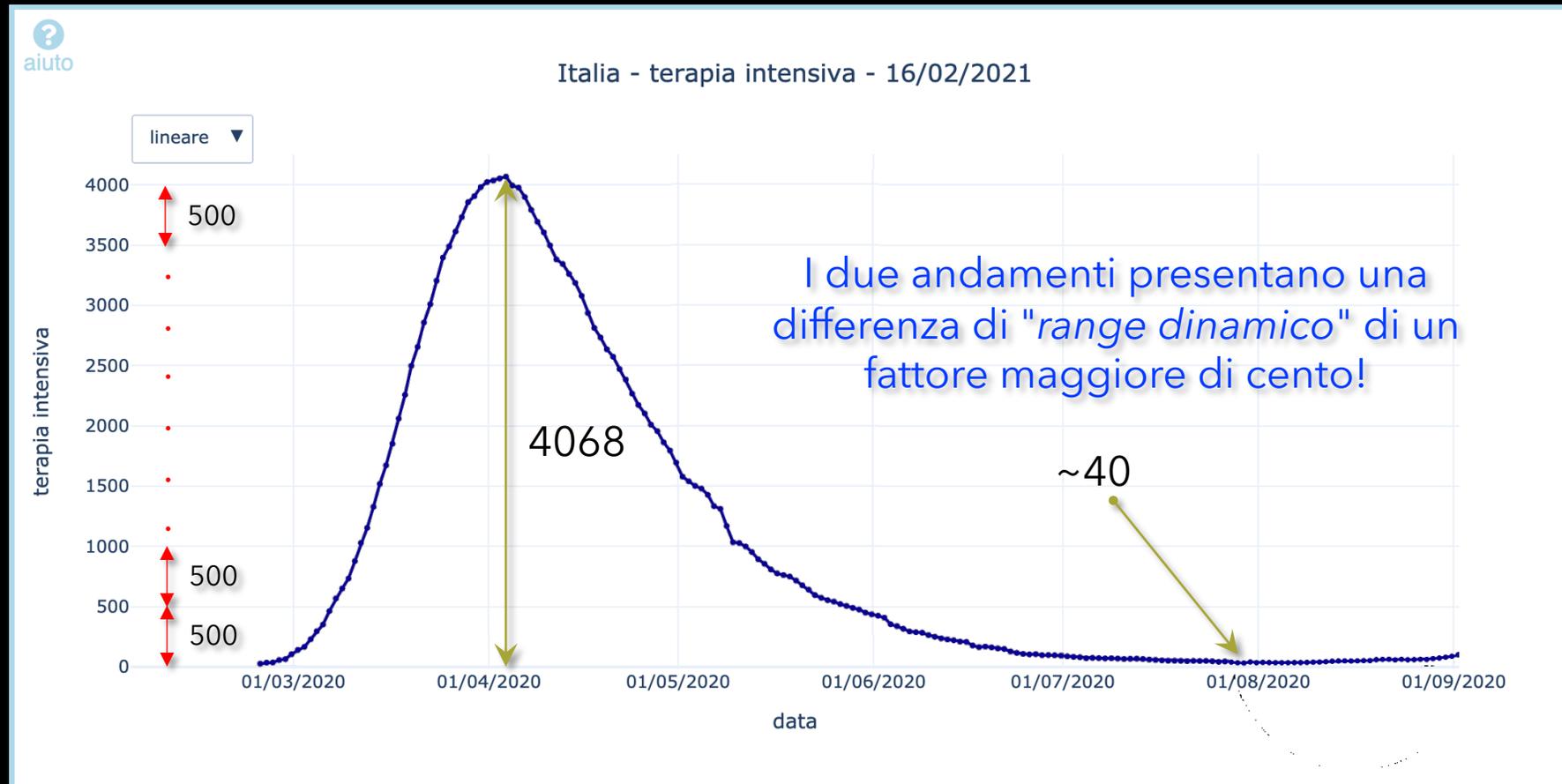
Intorno al 1 Agosto si registravano circa 40 casi al giorno e, in effetti, pareva davvero finita, i ricoveri in terapia intensiva (quelli preoccupanti per la tenuta del sistema sanitario nazionale) non aumentavano più.



Qualcuno, a conoscenza delle proprietà delle funzioni matematiche, fece notare che, osservando il grafico su una scala verticale logaritmica, la prospettiva cambiava totalmente...

Il punto importante qui è che in scala lineare, ogni tacca successiva dell'asse verticale corrisponde a esattamente 500 conteggi e questo valore si mantiene costante fino in cima.

Di conseguenza la grandezza del picco a sinistra rende minuscola, in proporzione, la variazione dei conteggi in basso a destra. Semplicemente non si notano perché schiacciati verso il basso



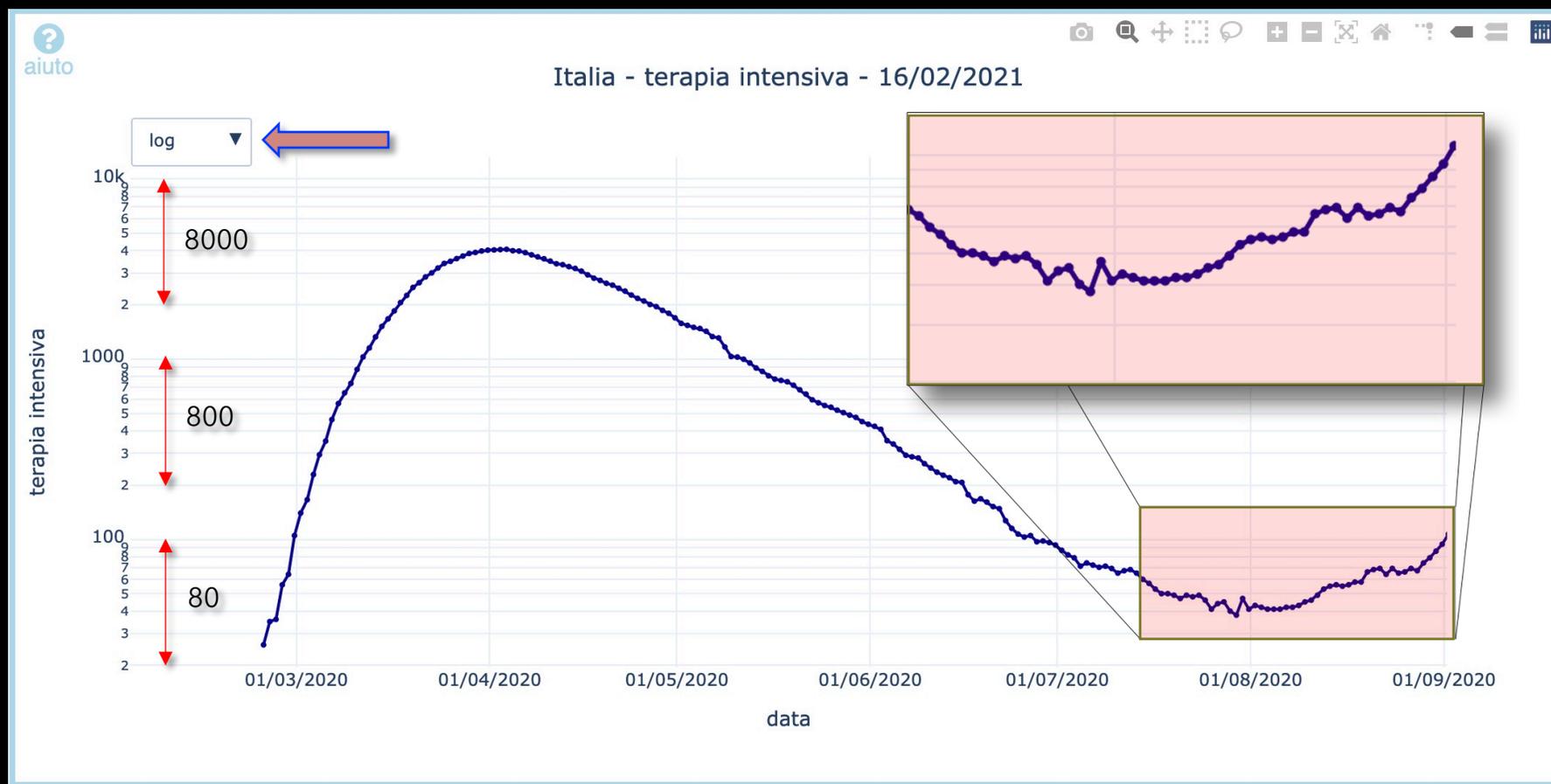
È chiaro a questo punto che vorremmo poter correggere questa grande differenza numerica tra valori grandi e valori piccoli. Vorremmo poter "zoomare" nella zona dei valori piccoli, come avevamo fatto con la lente.

Ma una lente ingrandisce tutto, anche il picco, e così non cambia nulla e così l'operazione risulta inutile.



Per fortuna in matematica esiste una funzione che opera una deformazione adatta al nostro scopo: si tratta di un'operazione che trasforma una serie di numeri reali nel loro logaritmo (parola che respinge molti perché suona impegnativa ma in realtà è molto semplice se se ne capisce l'utilità...)

Con questa trasformazione alteriamo i numeri in modo che ogni tacca verticale, salendo, diventa 10 volte più piccola: in questo modo i numeri grandi vengono compressi progressivamente di più che non i numeri piccoli. In questo modo si evidenziano strutture che sarebbero troppo schiacciate per essere visibili.



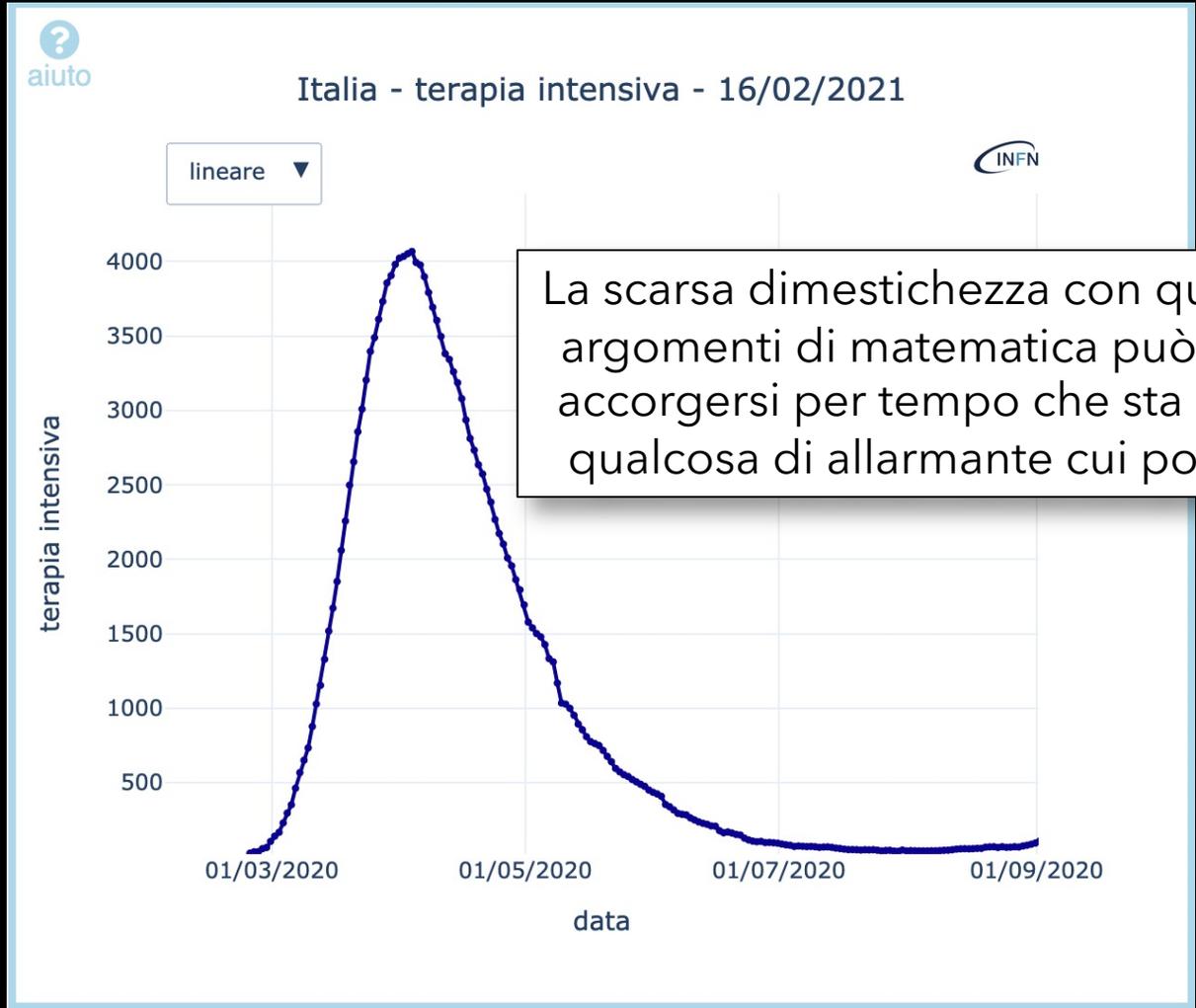
Scopriamo così che quella che sembra una linea piatta, attestata a circa 40 casi al giorno in realtà è una curva in piena risalita! Grazie alla deformazione di ingrandimento introdotta dai logaritmi, ci accorgiamo che il minimo era stato raggiunto il 29 Luglio e non il 1 Agosto (in scala lineare questi due giorni parevano alla stessa identica altezza)

Ma il punto vero è che da un minimo di 38 casi il 29 Luglio si raggiunge in un mese il valore di 107 casi il 1 Settembre. Mica vero che il virus era scomparso!!!!

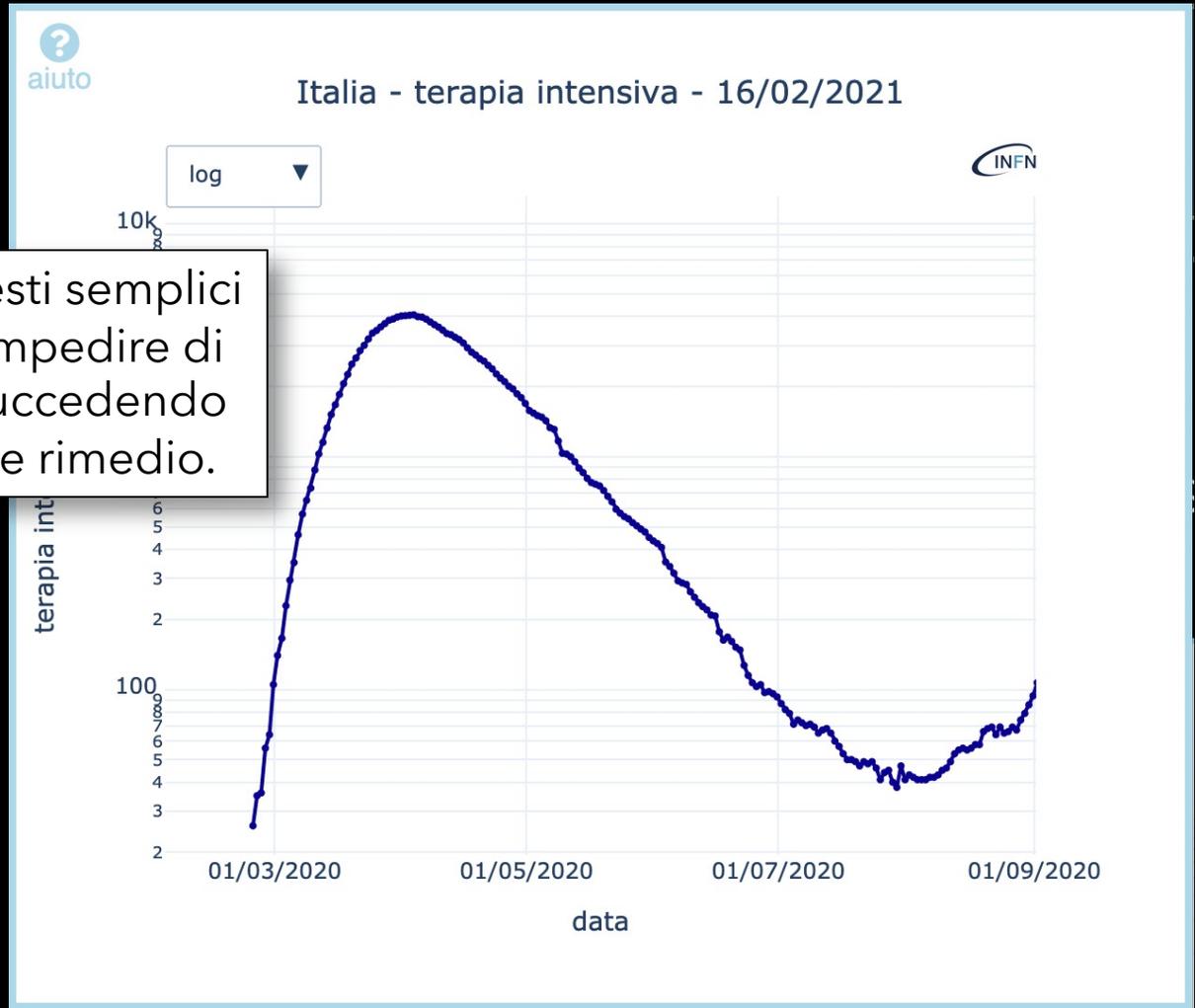


Diventa a questo punto chiaro che i logaritmi, lungi dall'essere una pignoleria astrusa per matematici hanno invece applicazioni utilissime anche in altri campi (molte, molte di più che non solo questa, naturalmente)

Abbiamo anche capito così, che semplicemente osservando gli stessi identici dati in due modi diversi possiamo facilmente mettere in evidenza fatti che in un modo risultano praticamente invisibili e nell'altro diventano eclatanti.



La scarsa dimestichezza con questi semplici argomenti di matematica può impedire di accorgersi per tempo che sta succedendo qualcosa di allarmante cui porre rimedio.



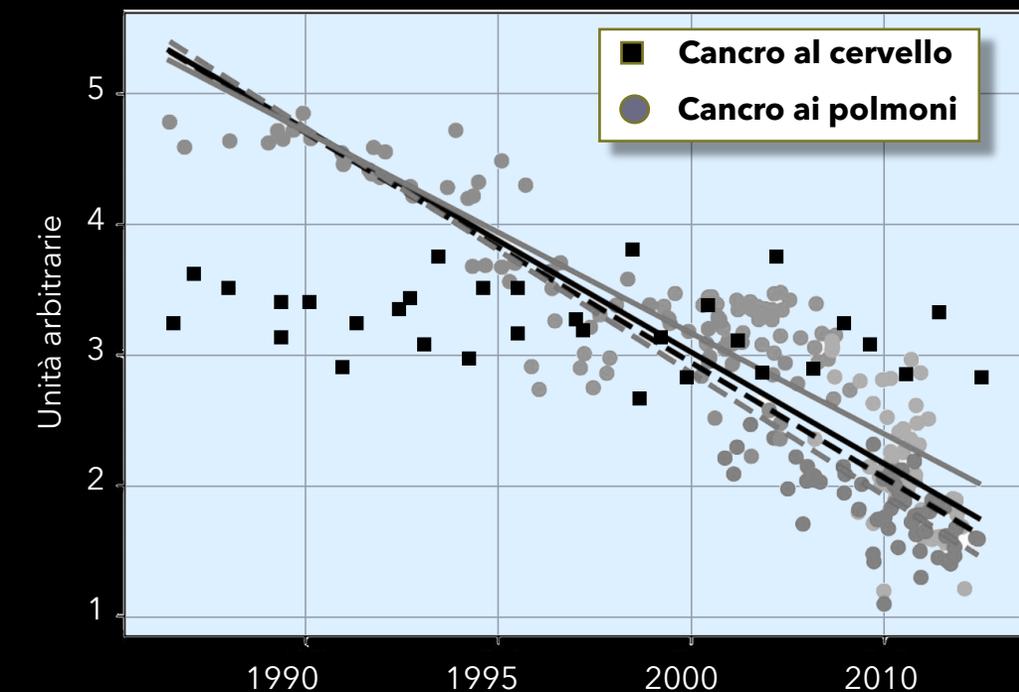
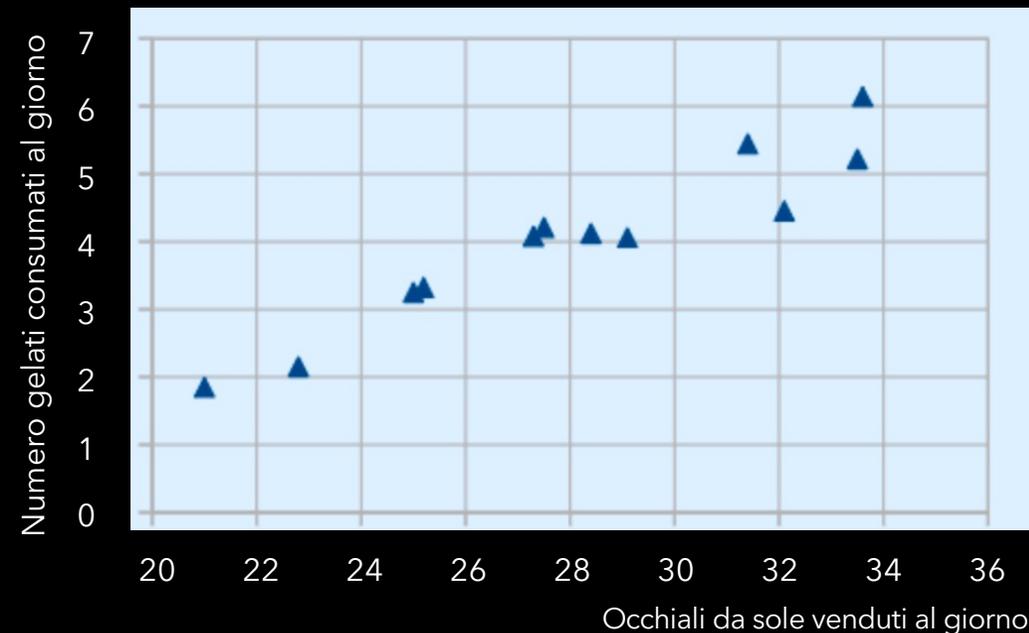
Correlation is not causation



Molte persone si sono spinte in un territorio scivoloso cercando di individuare correlazioni tra l'espandersi della pandemia e cose quali, ad esempio, l'inquinamento atmosferico.

Se una correlazione esiste, comunque, questo non garantisce che gli agenti atmosferici possano essere una concausa dell'epidemia (**potrebbero** esserlo ma non è garantito)

Se osserviamo un grafico che mette in correlazione il numero di occhiali da sole e di gelati venduti al giorno saremmo tentati di dire che la vendita di occhiali sia la causa dell'aumento di gelati



La reazione c'è ma è dovuta ad una terza causa scatenante: la stagione estiva quando c'è molto sole e fa quindi molto caldo.

Anni fa un'inchiesta in Svezia trovò che al passare degli anni i possessori di cellulari avevano una probabilità costante (bassa) di cancro al cervello ma una diminuzione di quella ai polmoni.

La spiegazione di questo paradosso fu che la fetta di popolazione che poteva permettersi un cellulare era la più benestante, con un tenore di vita alto e composto prevalentemente non fumatori. Anche qui la correlazione era sottile: dipendeva dal particolare campioni di casi selezionati.

Quando si analizzano dati occorre un modello per interpretarli e la scelta del modello adatto è un punto critico.

Se l'andamento dei dati conferma il modello questo non significa che il modello descriva perfettamente la realtà: significa solo che il modello ne coglie un particolare aspetto in modo più o meno fedele. Il modello è valido ma solo come approssimazione. Più i dati sono precisi, poco affetti da incertezze, in grande numero e più il modello potrà essere validato come rappresentante di quell'aspetto della realtà.

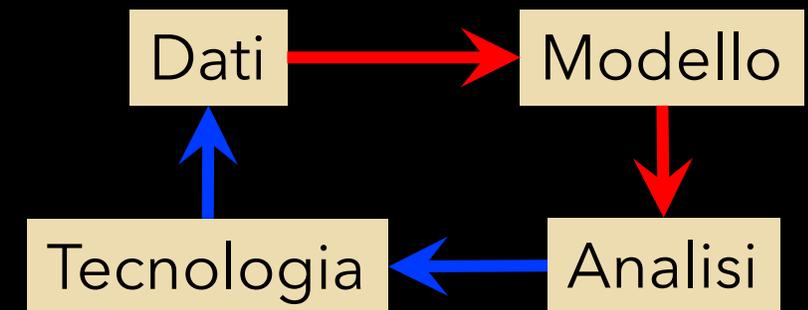
Un altro punto fondamentale è la scelta dei dati: se per stimare quanti infetti ci sono ogni giorno ci basassimo, come criterio, sul semplice contare le persone risultate positive a un tampone potremmo sbagliare di grosso.

Basterebbe infatti ridurre il numero di tamponi fatti e magicamente diminuirebbe il numero di infetti...

Questo si traduce nel fatto che i criteri di selezione dei nostri dati, in fase di raccolta e analisi, devono essere privi di condizionamenti rispetto al risultato che ci si aspetta.

Se nel fare un sondaggio per verificare il livello di gradimento di un personaggio politico intervistassi solamente gli iscritti al suo partito non ricaverai un'immagine reale della situazione.

Il progresso scientifico si basa su un meccanismo retroattivo di analisi dati basati su un modello. Migliori i dati, più accurato il modello, più tecnologia avanzata riusciamo a generare e questo ci porta a migliorare la capacità di raffinare i nostri modelli.



Una breve digressione sulle trasformazioni di un sistema di coordinate

I grafici che abbiamo visto sono dotati di un sistema di coordinate: l'asse orizzontale rappresenta il passare del tempo mentre l'asse verticale indica i conteggi registrati in ogni particolare giorno.

Per mettere in evidenza quel che ci interessava osservare, ossia se ci fosse o meno una significativa ripresa dei ricoveri in terapia intensiva dopo l'estate, siamo ricorsi ad una opportuna deformazione del sistema di coordinate.

Abbiamo stiracchiato l'asse verticale usando la notazione logaritmica invece di quella lineare: questo trucco però non è nuovo. Lo utilizziamo anche noi, senza minimamente rendercene conto, tutte le volte che consultiamo una cartina geografica del mondo



A cosa serve un sistema di coordinate?

In alcuni casi ci serve per poter comunicare a qualcun altro dove ci troviamo e come, eventualmente ci si possa raggiungere da un altro luogo o si possa noi andare in un particolare posto. Si tratta quindi di un concetto legato alla "geometria".

Una mappa è un buon esempio di sistema di coordinate. C'è un'origine, un orientamento (un verso, diremmo in geometria) e delle unità di misura che ci permettono di localizzare luoghi e valutarne le distanze.

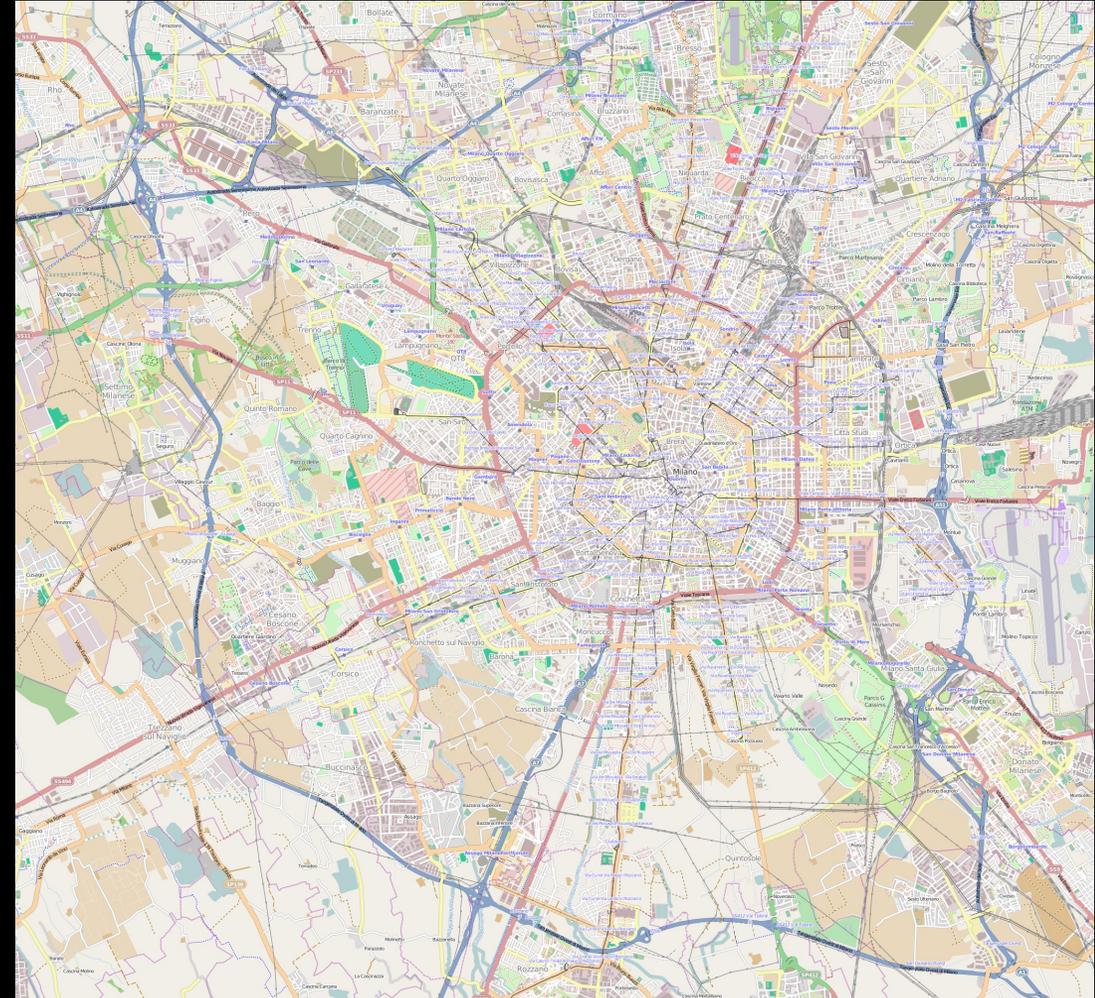
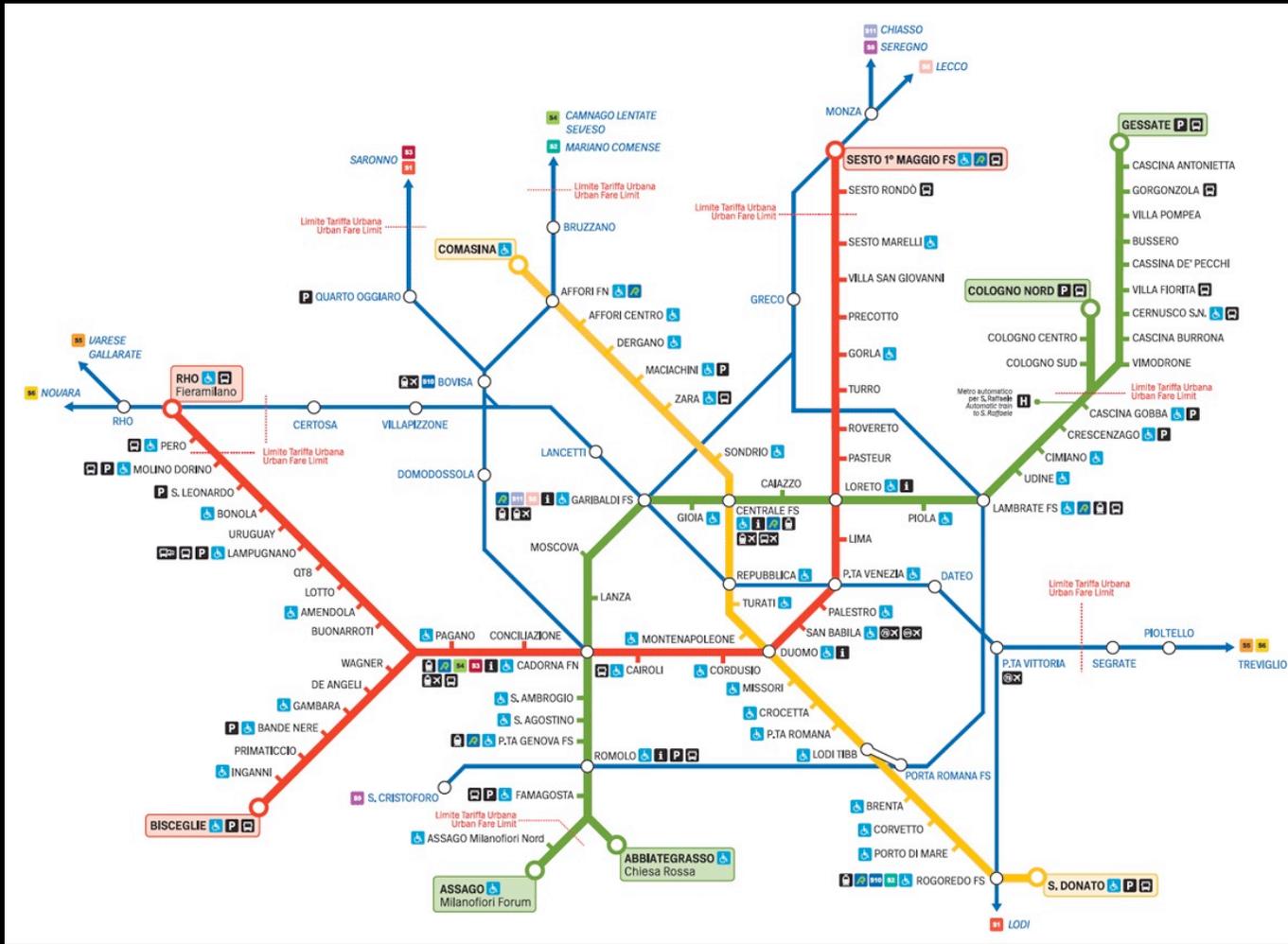
Come realizzare una mappa dipende moltissimo da che tipo di informazione ci aspettiamo che essa possa fornirci. La mappa Peutingeriana, ad esempio, riscoperta da Konrad Peutinger nel 1507 e copiata da un monaco a Colmar nel 1265 da un originale di epoca romana, intendeva informare, chi la consultava, dell'esistenza di strade percorribili all'epoca dalle armate dell'Impero.



A cosa ci serve un sistema di riferimento?

La carta Peutingeringiana somiglia quindi più a una mappa descrittiva dei collegamenti logici tra punti diversi, come può esserlo la rete dei trasporti pubblici di Milano, piuttosto che non a una cartina geografica della città.

In quella a sinistra contano le relazioni logiche, relative, tra i luoghi rappresentati, in quella a destra contano le loro posizioni assolute. Sono entrambe rappresentazioni spaziali, ma dotate di quella che si chiama una "*metrica*" differente.



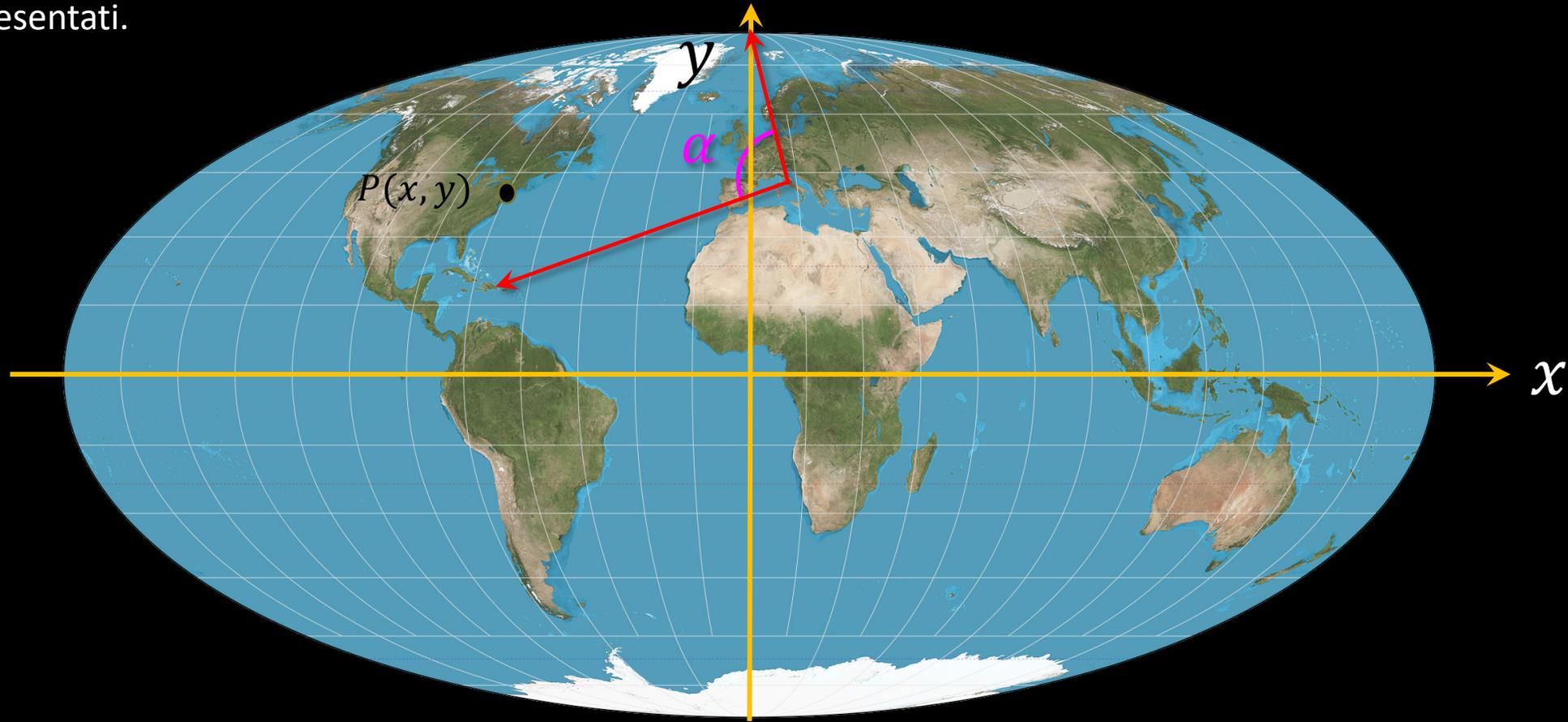
Torniamo alle carte geografiche che siamo abituati a considerare come tali, ossia qualcosa che ci permette di localizzare con precisione i luoghi e le distanze tra essi. Qui vediamo rappresentata una mappa del mondo che riconosciamo facilmente.



Guardandola con attenzione non si può evitare di notare come Russia, Canada e Groenlandia sembrano molto più grandi di tutti gli altri paesi. Ma si tratta di un fatto vero o di una voluta distorsione della mappa? Vediamo di capire questo concetto, molto importante nel contesto della geometria in generale (e della Teoria della Relatività in particolare).

Cercando in rete troverete che ci sono molti modi di rappresentare la Terra su una mappa bidimensionale.

Ecco ad esempio una mappa chiamata Proiezione di Mollwide (in seguito raffinata da Babinet) (1805): si tratta di una mappa che cerca di riprodurre con un buon compromesso sia gli angoli esistenti tra i luoghi che, contemporaneamente, le corrette aree dei territori rappresentati.



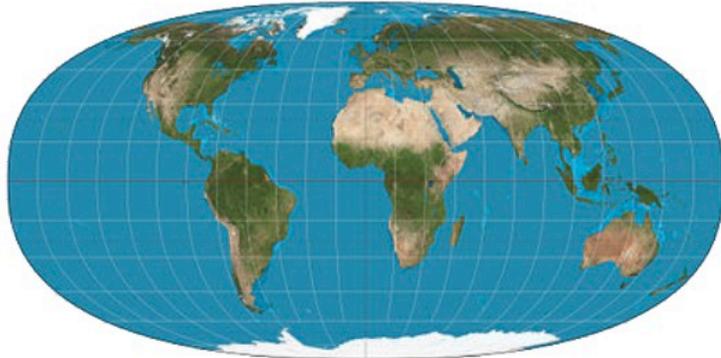
Un punto $P(x, y)$ sulla mappa è calcolato secondo queste equazioni, dove λ è la longitudine del punto (in coordinate polari), λ_0 è l'origine delle longitudini (Greenwich), R il raggio terrestre e ϕ la latitudine

$$x = R \frac{2\sqrt{2}}{\pi} (\lambda - \lambda_0) \cos\theta \quad y = R\sqrt{2} \sin\theta \quad 2\theta + 2\sin\theta = \pi \sin\phi$$

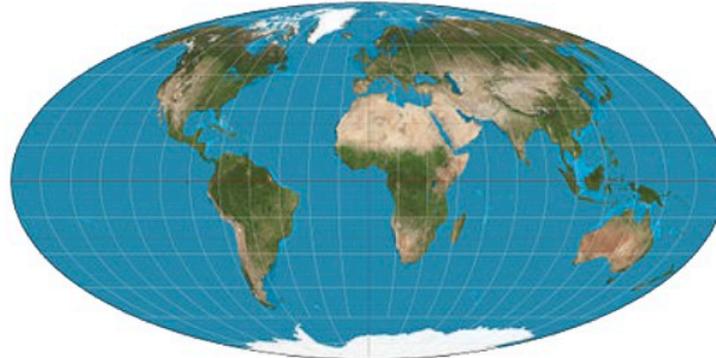
Per ottenere quindi un mappa che riproduca ciò che ci serve, occorre trasformare le coordinate secondo certe regole (arbitrarie)

Di criteri per produrre mappe dotate di particolari caratteristiche utili se ne possono elencare davvero tanti: una mappa può rappresentare fedelmente quanto è grande un territorio rispetto ad un altro, un'altra quanto distano i luoghi tra loro e un'altra ancora le relazioni angolari tra i luoghi (utile per la navigazione col sestante e la bussola): qui ne vediamo una selezione tra quelle possibili ed esistenti.

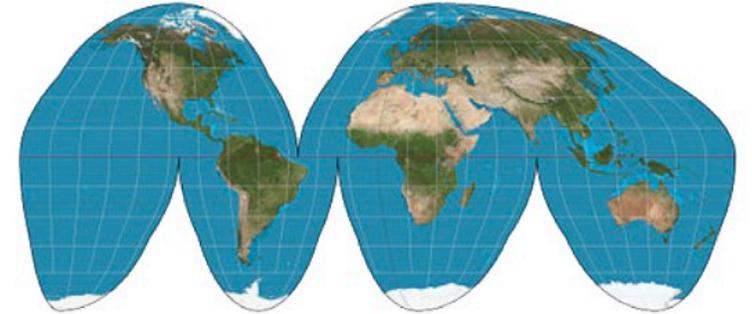
- Tobler hyperelliptical



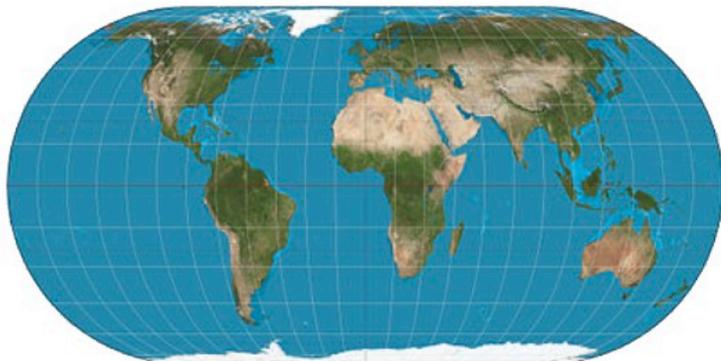
- Mollweide



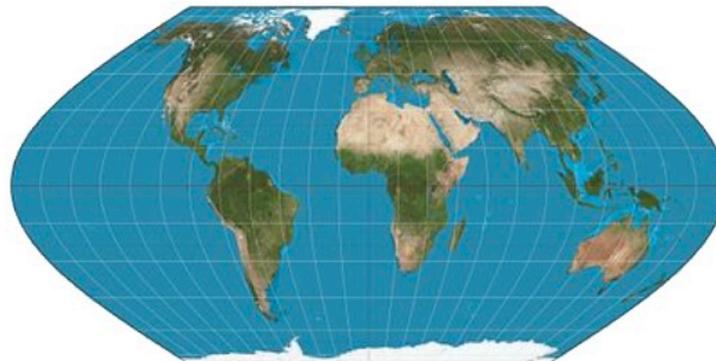
- Goode homolosine



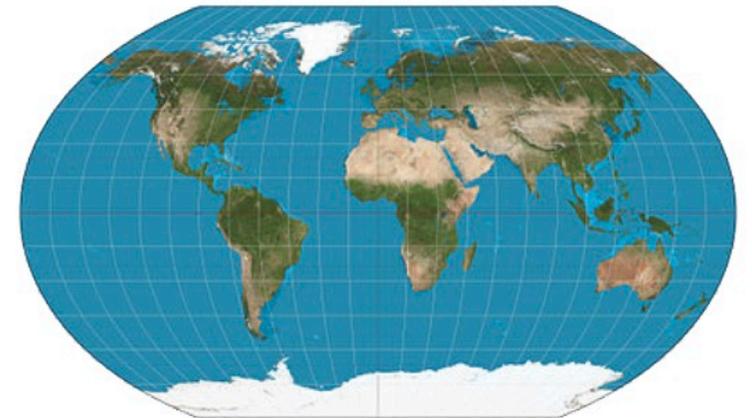
- Eckert IV



- Eckert VI



- Kavrayskiy VII



Ma perché esistono tutte queste varianti? Non si potrebbe semplicemente rappresentare la superficie terrestre su una carta in modo corretto come la vediamo su un mappamondo?

Purtroppo no, non è possibile farlo, come vedremo tra un istante: questo piccolo excursus nella geometria tridimensionale ci darà occasione di introdurre concetti di geometria avanzata che torneranno utili quando parleremo di Teoria della Relatività Generale.

Per ora accontentiamoci di capire perché una mappa bidimensionale non può *mai* essere una rappresentazione fedele di una mappa sferica.





Ricordiamo cosa vogliamo fare: vogliamo rappresentare l'aspetto della superficie terrestre esattamente come si trova sulla superficie di una sfera anche su un foglio di carta (per produrre una mappa portatile): vediamo se questo sia possibile e come fare per stabilirlo.



Potremmo, ad esempio, avvolgere il foglio di carta attorno alla sfera creando un cilindro che la avvolga tangenzialmente e "*ricalcare*" su di esso la posizione dei vari luoghi.

Ci accontentiamo di un cilindro perché sappiamo tutti per esperienza che avvolgere completamente un foglio di carta attorno ad una palla senza creare delle antipatiche pieghe è praticamente impossibile.

Ci dovremo allora accontentare di avvolgere il foglio su un cilindro che a sua volta circonda la Terra e vedere se riusciamo in questo modo a "*ricalcare*" l'immagine sul foglio così arrotolato

Ricalcare i luoghi sulla mappa significa immaginare che il foglio di carta sia trasparente e si possano scorgere attraverso di esso i punti che si vogliono riprodurre sulla mappa.

Proviamo allora a ricalcare sul foglio la posizione di due luoghi a caso, ad esempio Madrid e Napoli





Madrid

1550 km km

Napoli

Madrid

1550 km

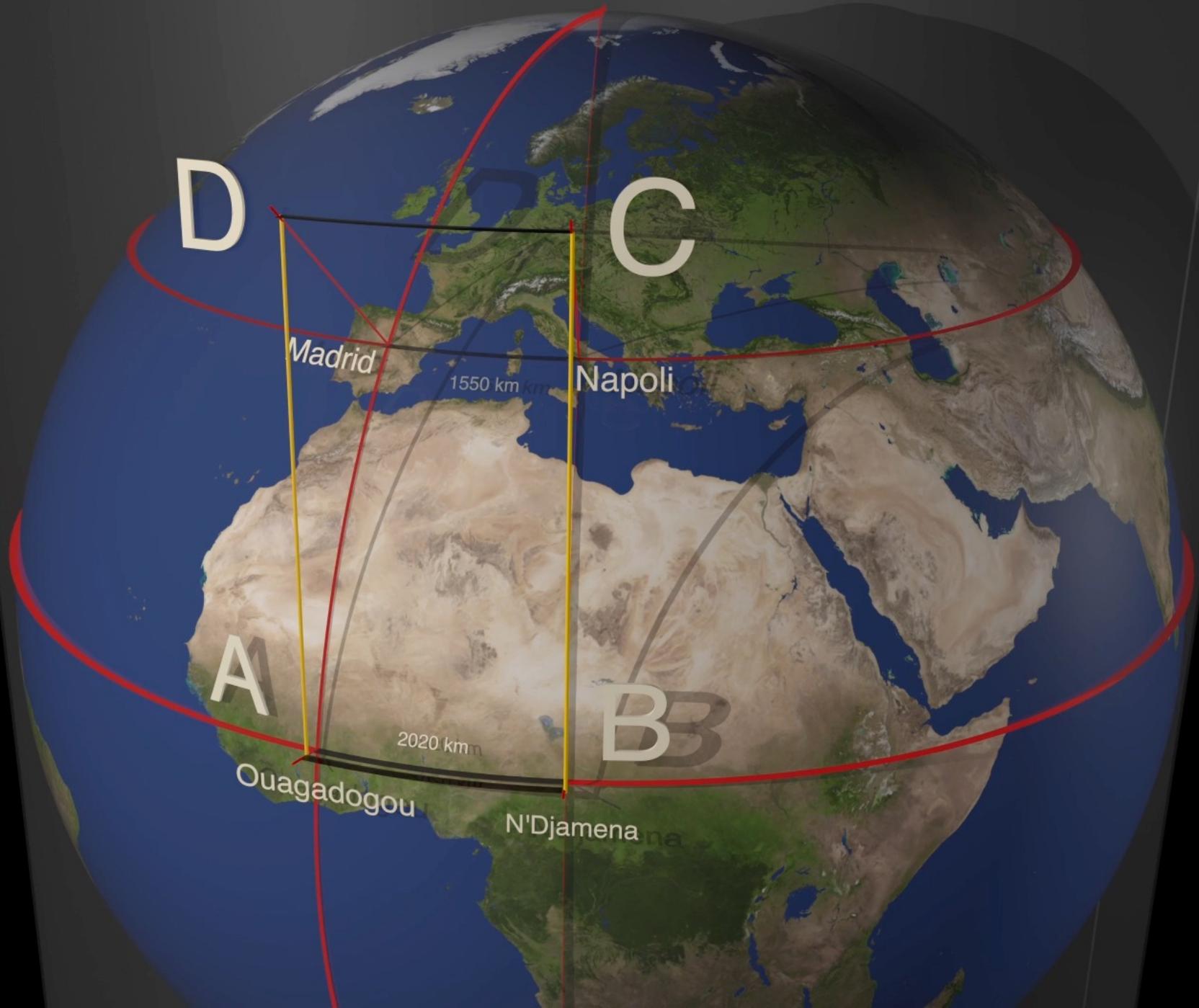
Napoli

2020 km

Ouagadougou

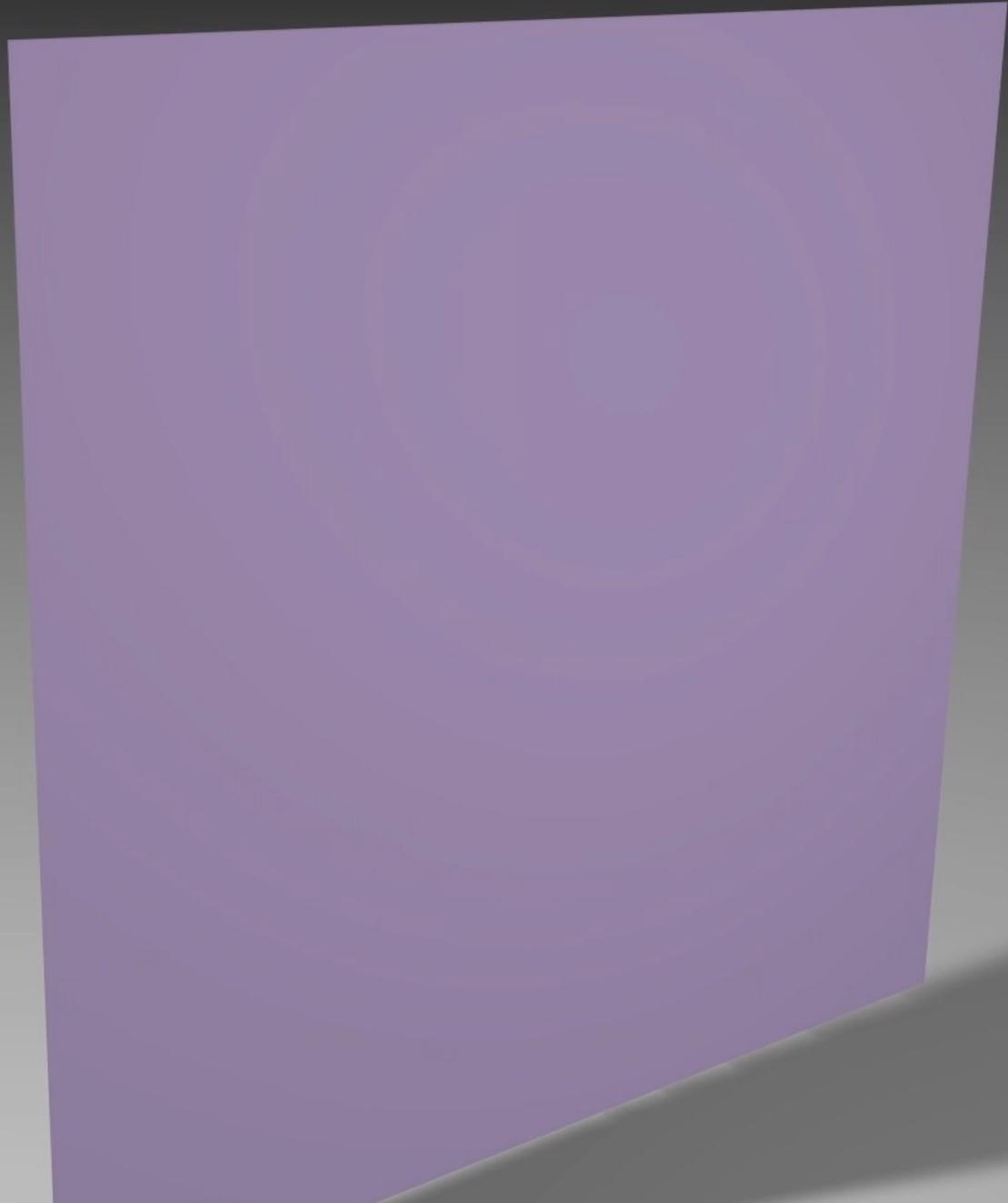
N'Djamena





Torniamo al caso tridimensionale e osserviamo questa distorsione in altro modo:

Suddividiamo ora questa mappa in un reticolo di 16x16 quadrati con gli assi ortogonali tra loro

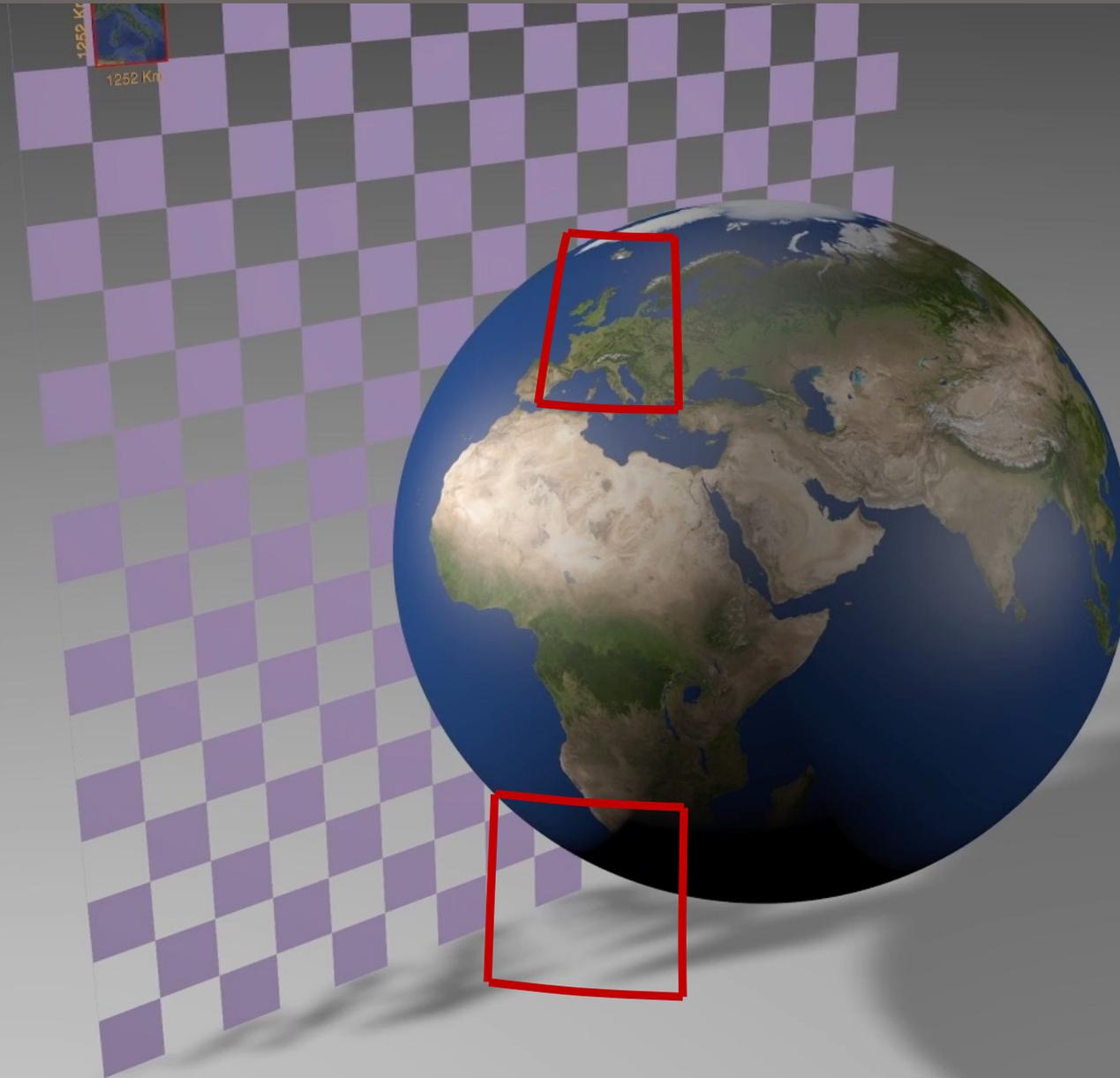


1252 Km

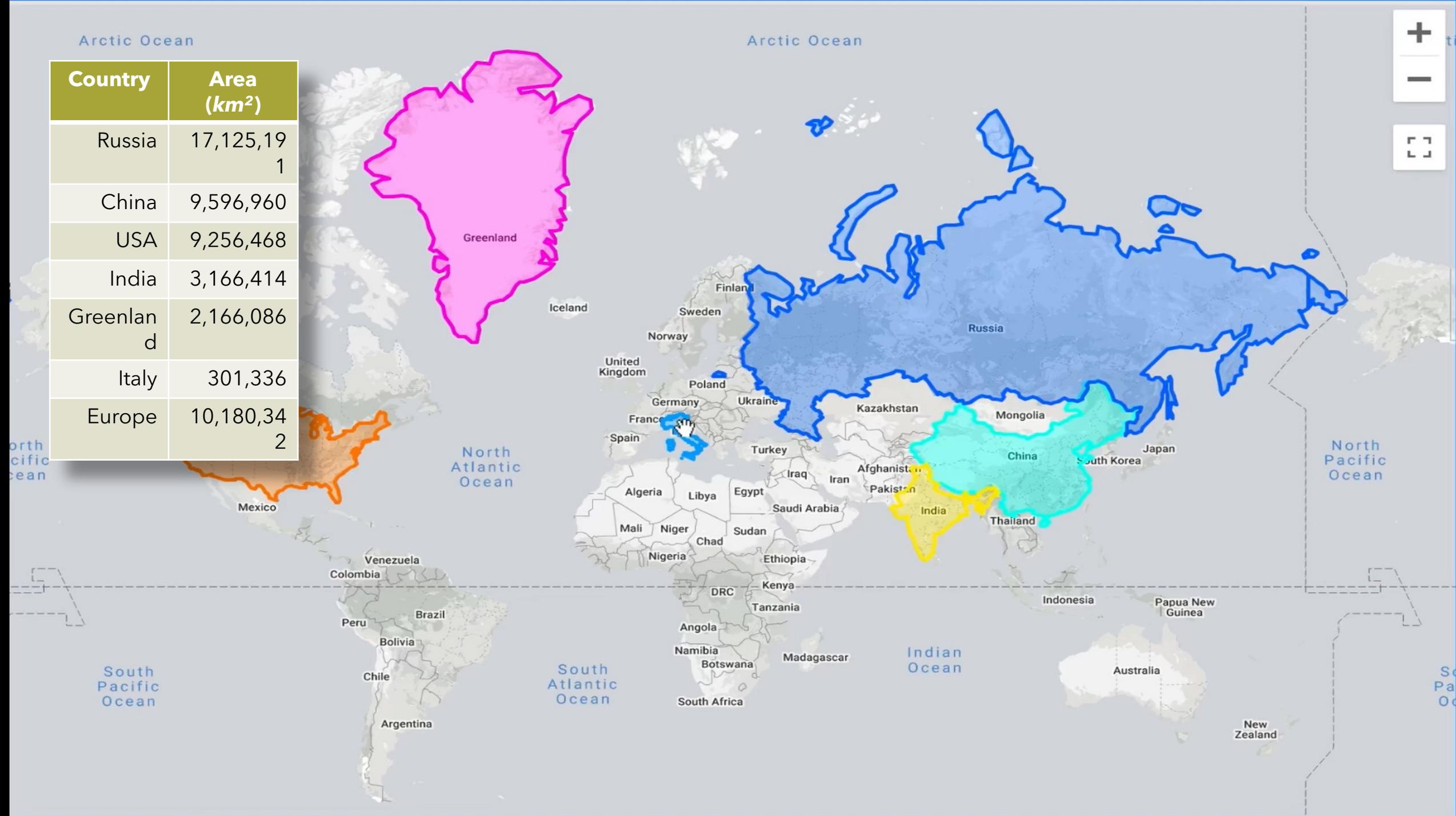


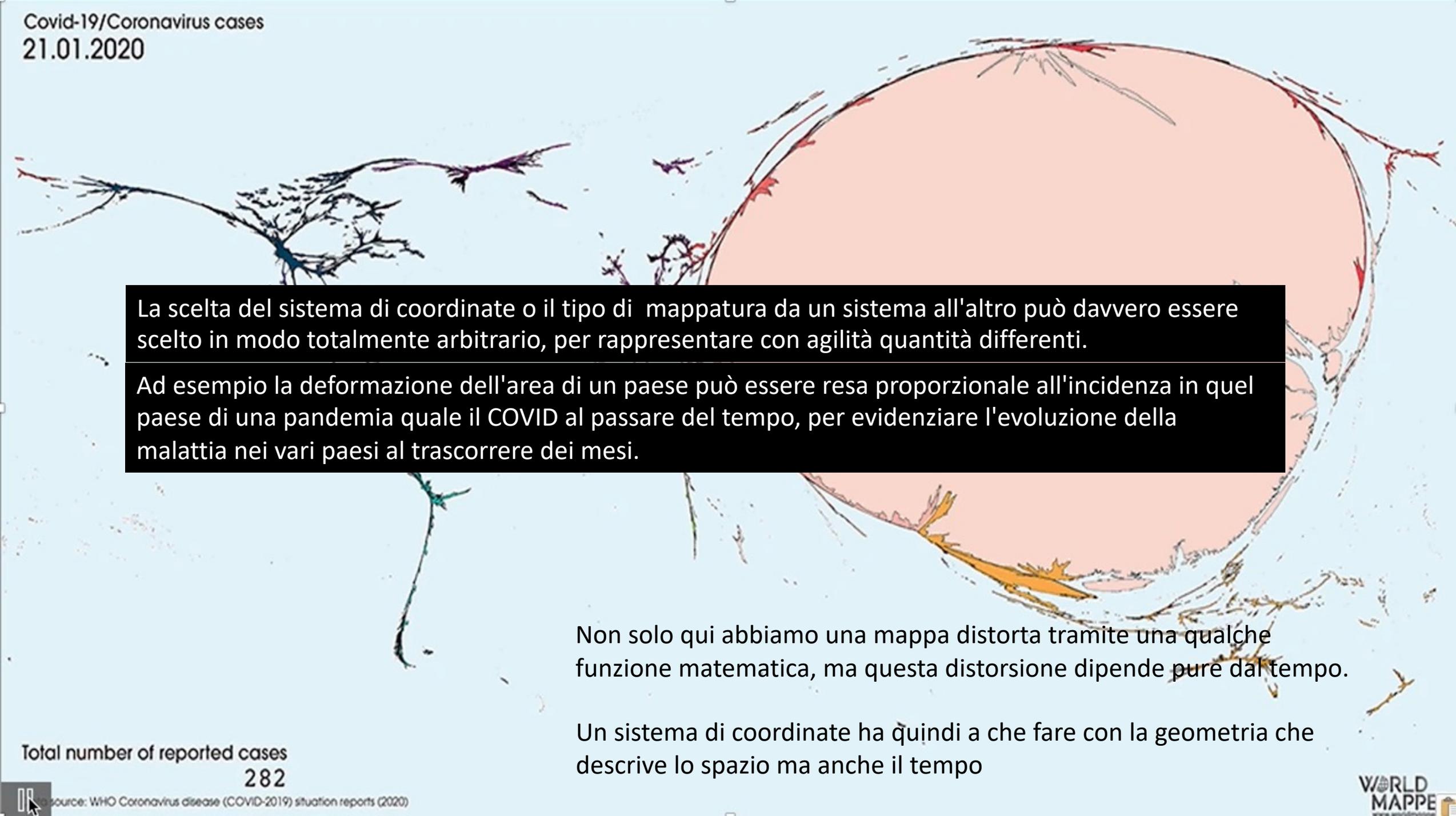
1252 Km

Vogliamo ora "avvolgere" questo piano sulla sfera e vedere cosa accade ai 256 quadrati in cui lo abbiamo suddiviso. Questa volta non consideriamo quindi un "avvolgimento" cilindrico, bensì uno direttamente sferico.



Country	Area (km ²)
Russia	17,125,191
China	9,596,960
USA	9,256,468
India	3,166,414
Greenland	2,166,086
Italy	301,336
Europe	10,180,342





La scelta del sistema di coordinate o il tipo di mappatura da un sistema all'altro può davvero essere scelto in modo totalmente arbitrario, per rappresentare con agilità quantità differenti.

Ad esempio la deformazione dell'area di un paese può essere resa proporzionale all'incidenza in quel paese di una pandemia quale il COVID al passare del tempo, per evidenziare l'evoluzione della malattia nei vari paesi al trascorrere dei mesi.

Non solo qui abbiamo una mappa distorta tramite una qualche funzione matematica, ma questa distorsione dipende pure dal tempo.

Un sistema di coordinate ha quindi a che fare con la geometria che descrive lo spazio ma anche il tempo

Conclusione (I)

Il modo col quale vogliamo rappresentare la localizzazione spaziale dei punti su una superficie dipende solo dalle nostre necessità, siamo liberi di scegliere quella che vogliamo. Questa rappresentazione prende il nome di sistema di coordinate.

Una volta fissata la scelta di un sistema di coordinate possiamo quindi cambiare la sua rappresentazione mediante una trasformazione matematica che ci porti, in modo opportunamente scelto per soddisfare una nuova necessità, ad una nuova rappresentazione. Questa trasformazione matematica prende il nome di "*mappatura*".

Mappare i dati serve a mettere in evidenza fatti che nei medesimi sono presenti ma non in primo piano, sono un utile ausilio per la loro analisi. Naturalmente occorre saperli usare con saggezza, conoscendone le proprietà matematiche. Ecco uno dei motivi per cui la matematica non solo è utile, ma spesso ci salva la vita...

Conclusione (II)



La realtà del mondo ci appare ben esemplificata dal Mito della Caverna di platonica memoria: i dati sono solo il distorto riflesso della realtà sulla parete della caverna e noi osservatori da quei segni deformati ne cogliamo l'immagine.

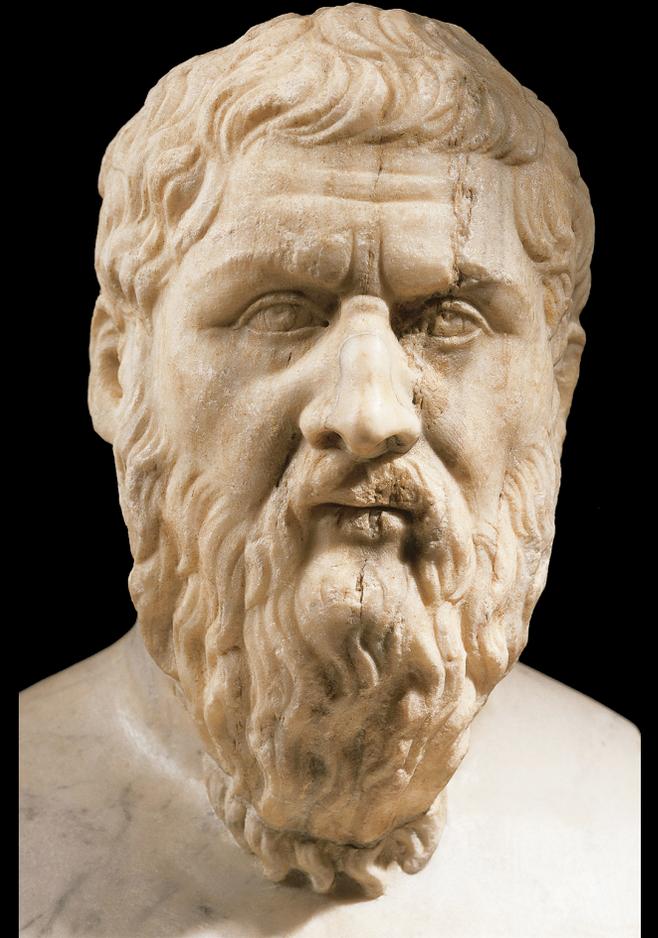
I dati dunque NON SONO la realtà, sono *una parte* di essa: di conseguenza occorre un criterio interpretativo per farne un uso corretto, un uso che ne riduca la distorsione di cui sono affetti.

Branche della matematica, come la statistica, hanno proprio lo scopo di fornire i corretti strumenti per ridurre in modo massimale possibile le distorsioni presenti nei dati.

Più i dati sono ricchi di informazione più sarà possibile essere accurati nel valutare e correggere queste distorsioni e più ci approssimeremo ad una ragionevolmente accurata distorsione della realtà. Ma questa in assoluto temo ci sarà per sempre preclusa.

Grazie

(e grazie Platone!)



<http://covid19.infn.it>