

LE PAROLE PER CAPIRE I NUMERI I NUMERI PER CAPIRE IL MONDO

Giovanni A. Barbieri

Ex direttore Istat per lo sviluppo dell'informazione e della cultura statistica

Genova, I mestieri dello statistico 2019

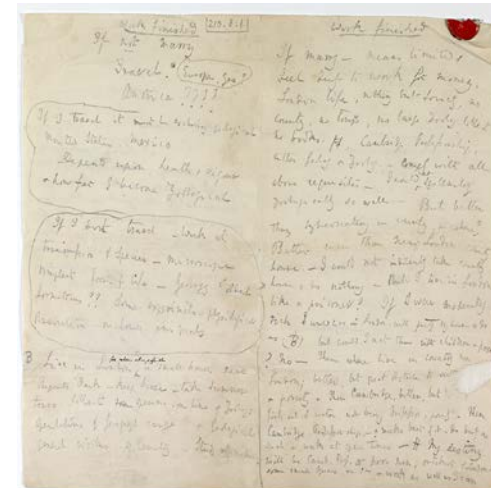
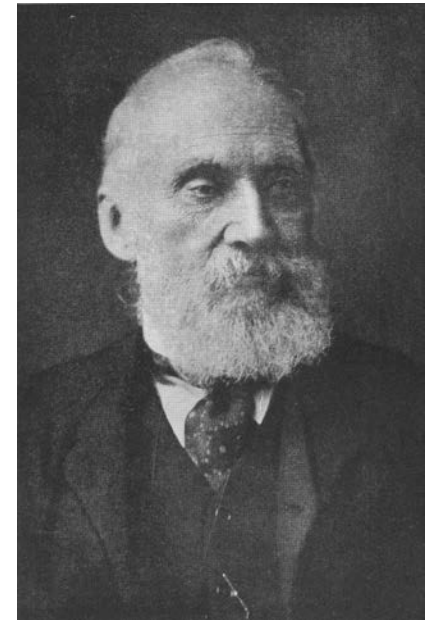
Palazzo della Borsa, Sala delle grida, 18 febbraio 2019

Sommario

- I numeri per capire il mondo
- Le parole per capire i numeri
- La statistica è facile o difficile?
- Statistica, fenomeni collettivi, modelli
- La statistica è di tutti
- Come possiamo la nostra comprensione delle statistiche

I numeri per capire il mondo

- Dell'esistenza di quello che non si può **misurare**, o quanto meno **quantificare**, conviene dubitare
 - «Dico spesso che quando si può misurare ciò di cui si sta parlando, ed esprimerlo in numeri, si sa qualcosa al riguardo; ma quando non si può misurare, quando non si può esprimere in numeri, la conoscenza è magra e insoddisfacente; può essere l'inizio della conoscenza, ma si è raramente avanzati allo stadio della scienza, quale che sia la materia» [Lord Kelvin, 1883]
- Darwin: i pro e i contro del matrimonio [1838]
 - If not marry | Travel. Europe? | America????
 - If marry – means limited, Feel duty to work for money. London life, no country, no tours



Sommario

- I numeri per capire il mondo
- **Le parole per capire i numeri**
- La statistica è facile o difficile?
- Statistica, fenomeni collettivi, modelli
- La statistica è di tutti
- Sviluppare la nostra comprensione delle statistiche

La risposta alla domanda fondamentale sulla vita, l'universo e tutto quanto

- «Quarantadue! – urlò Loonquawl – Questo è tutto ciò che sai dire dopo un lavoro di sette milioni e mezzo di anni?»
- «Ho controllato molto approfonditamente – disse il computer – e questa è sicuramente la risposta. A essere sinceri, penso che il problema sia che voi non abbiate mai saputo veramente qual è la domanda.»



Le parole per capire i numeri

- Un numero non dà alcun contributo alla conoscenza, se non sai che cosa questo numero significa: l'insieme delle sue ramificazioni semantiche
- Le parole sono importanti
 - «Chi parla male, pensa male e vive male. Bisogna trovare le parole giuste: le parole sono importanti!» [Michele Apicella in *Palombella rossa* di Nanni Moretti]
- Le parole della scienza
- Le parole della statistica

Parole da usare con cautela nel discorso scientifico

- Credere → pensare
- Prova → evidenza
- Dimostrare → mostrare
- Teoria
- Verificare → falsificare



Come usa le parole la statistica: un esempio, i disoccupati

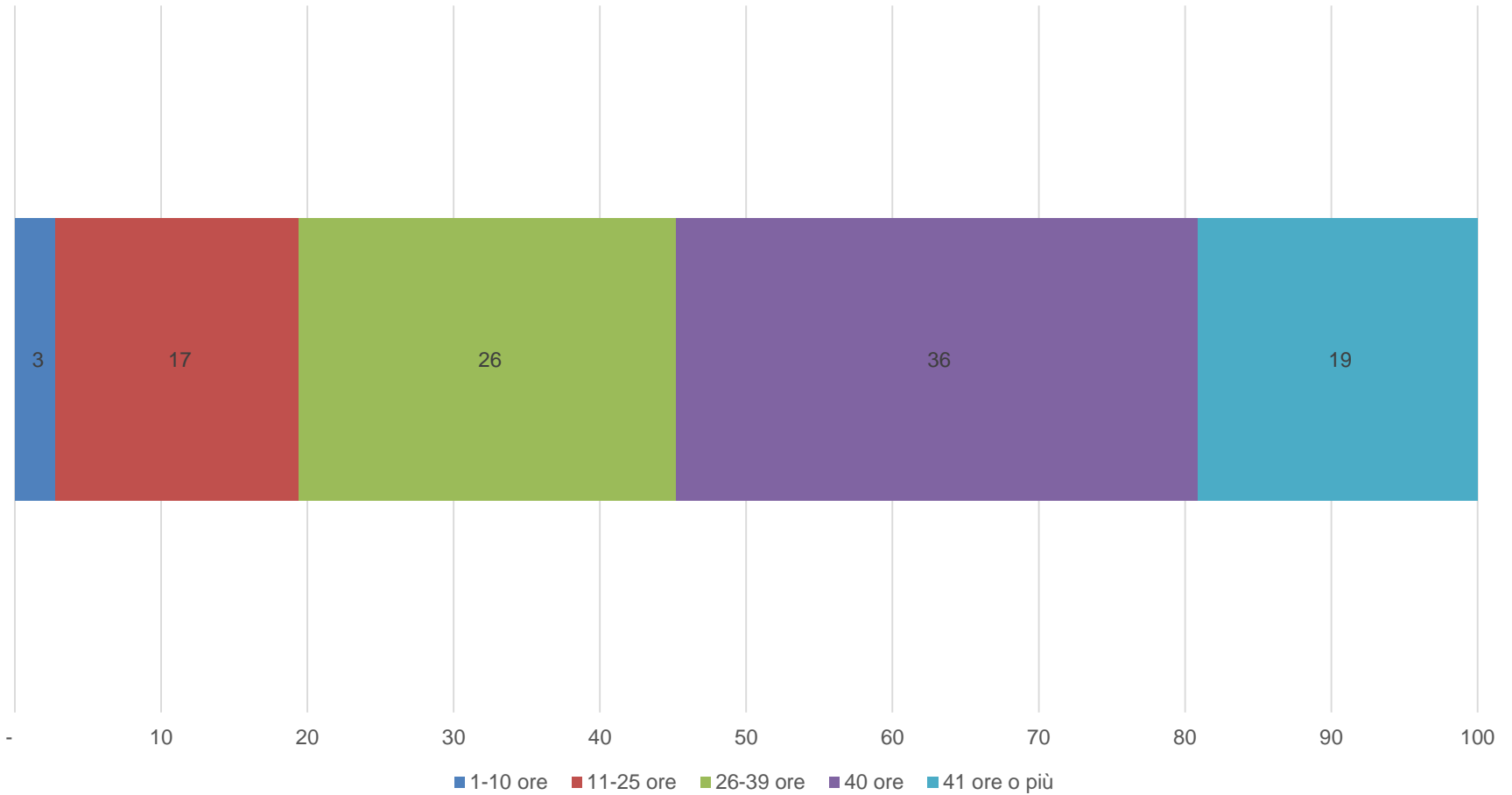
- In Italia i disoccupati sono 2.522.000
- Ma chi sono i disoccupati? Disoccupati (o in cerca di occupazione): comprendono le persone non occupate tra i 15 e i 74 anni che:
 - hanno effettuato almeno un'azione attiva di ricerca di lavoro nelle quattro settimane che precedono la settimana di riferimento
 - e sono disponibili a lavorare (o ad avviare un'attività autonoma) entro le due settimane successive
 - **oppure**, inizieranno un lavoro entro tre mesi dalla settimana di riferimento
 - e sarebbero disponibili a lavorare (o ad avviare un'attività autonoma) entro le due settimane successive, qualora fosse possibile anticipare l'inizio del lavoro
- Come verifico la disponibilità a lavorare nelle due settimane successive all'intervista?
- E poi, per individuare i disoccupati in questa accezione tecnica, ho necessità di distinguerli dagli occupati, da una parte, e dagli inattivi, cioè da coloro che non lavorano e non cercano lavoro, dall'altra



Come usa le parole la statistica: un esempio, gli occupati

- In prima battuta: gli occupati sono quelli che lavorano; gli inattivi sono quelli che non lavorano e non intendono lavorare perché studenti, o pensionati, o impegnati in attività domestiche ('casalinghe')
- Quante ore alla settimana occorre lavorare per essere statisticamente occupati? «almeno un'ora di lavoro in una qualsiasi attività che preveda un corrispettivo monetario o in natura»
- Scandalo? Non proprio, perché oltre alla durata media della settimana lavorativa (sugli oltre 23 milioni di occupati del 2017, la maggioranza relativa, il 36%, lavora 40 ore la settimana), ci sono anche delle 'code' di occupati che lavorano molto di più o molto di meno: il 19% dalle 41 ore in su, ma soltanto il 3% dieci ore o meno

Occupati per ore lavorate nella settimana di riferimento (composizione percentuale)

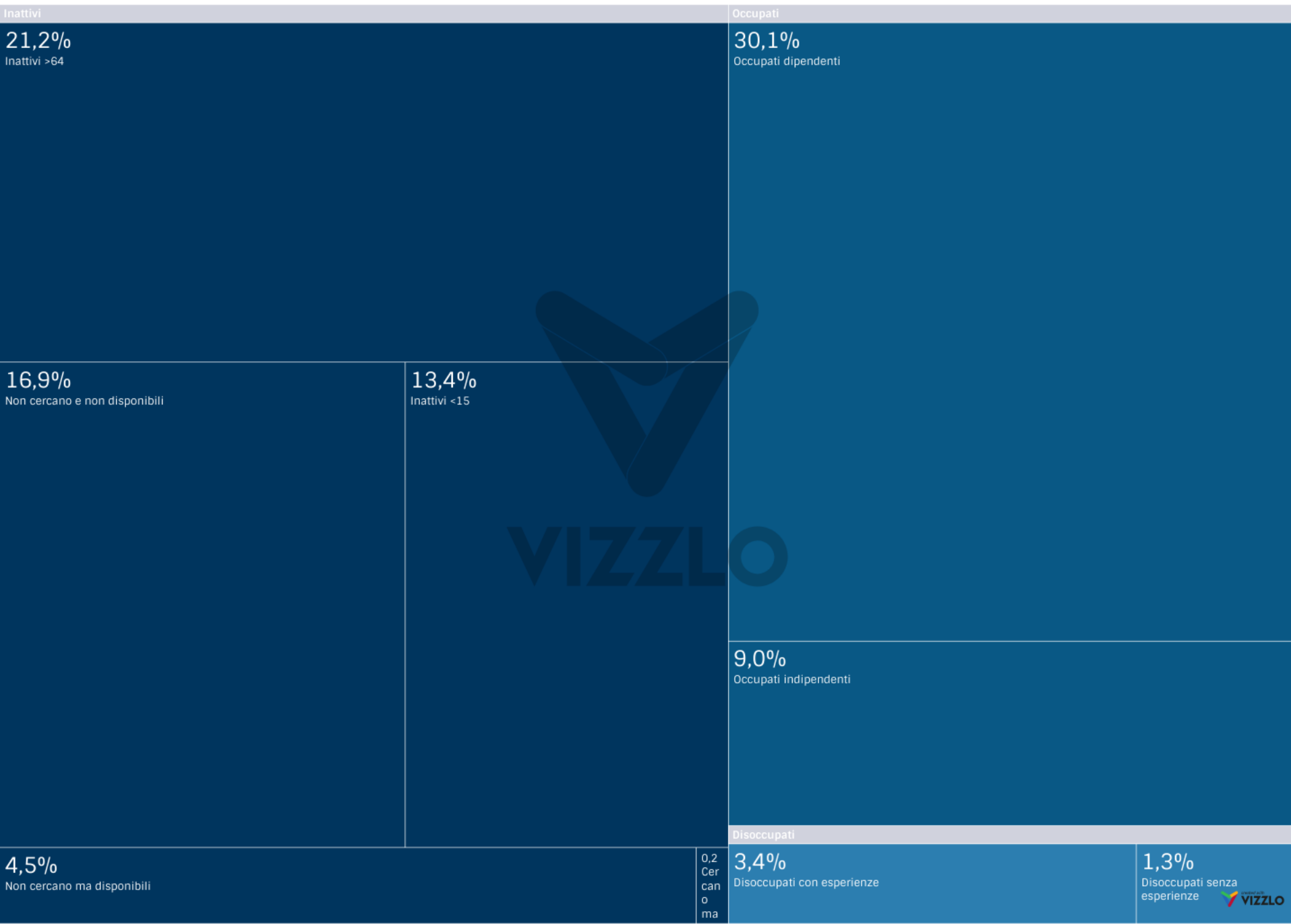




Come usa le parole la statistica: un esempio, gli inattivi

- Almeno questo non dovrebbe creare problemi, perché sono definiti in modo residuale, come la parte della popolazione che non lavora (altrimenti sarebbero occupati) e non cerca un'occupazione (altrimenti sarebbe disoccupati)
- Coerenza con le altre definizioni, complicata ma fruttuosa:
 - non hanno cercato lavoro nelle ultime 4 settimane e non sono disponibili a lavorare entro 2 settimane dall'intervista (per lo più di studenti, ritirati dal lavoro e 'casalinghe')
 - pur non avendo cercato un lavoro nelle ultime 4 settimane si sono dichiarati disponibili a iniziare un lavoro entro 2 settimane
 - coloro che hanno cercato un lavoro nelle ultime 4 settimane, ma che non sono disponibili a iniziare un lavoro entro 2 settimane
- Forze di lavoro potenziali
 - Confine fluido
 - La ricerca di lavoro ha un costo

Partecipazione al mercato del lavoro della popolazione residente



Sommario

- I numeri per capire il mondo
- Le parole per capire i numeri
- **La statistica è facile o difficile?**
- Statistica, fenomeni collettivi, modelli
- La statistica è di tutti
- Sviluppare la nostra comprensione delle statistiche

La statistica è facile o difficile?

- «Su 100 persone in età lavorative quante, secondo lei, sono disoccupate e cercano lavoro?»: 49 (invece all'epoca della rilevazione erano 12)
 - **Il Paese con le percezioni più sbagliate**, per questa domanda e per l'insieme delle domande della rilevazione Ipsos-MORI (30 quesiti, 13 Paesi, 50.000 interviste)
- L'istinto della statistica
 - Il minestrone è giusto di sale (Golini)?
 - Siamo cacciatori di configurazioni (*pattern*): [Pareidolia](#)
 - Modelli e metafore
 - Complessi, ma sempre meno della realtà
 - Linguaggio comune
 - Sotto controllo quello che è rilevante e quello che no
 - Fanno emergere nuove idee (procedere per metafore)



La statistica è istintiva ma ingannevole

- I nostri sensi non ingannano quasi mai
 - Ma ci sono illusioni [percettive](#)
- Anche l'istinto non inganna quasi mai
 - La credenza che il futuro sia influenzato dal passato è quasi sempre corretta (nei rapporti sociali, e persino nelle previsioni del tempo)
- Ci sono situazioni 'create apposta':
 - Le estrazioni del lotto sono congegnate in modo da essere indipendenti dall'estrazione precedente
 - Lo stesso accade per la roulette, che è una macchina di precisione progettata per ingannare il nostro 'intuito statistico'
- I nostri sensi e il nostro cervello funzionano bene soltanto alle scale che ci sono più vicine

Le leggi statistiche sono diverse

- La geometria e la matematica sono antiche (Pitagora, Euclide)
- La probabilità si fa risalire al 1654, con uno scambio di lettere tra Pascal e Fermat
- Perché così tardi? Ho due spiegazioni:
 - Le leggi statistiche sono diverse dalle altre leggi scientifiche
 - La somma degli angoli di un triangolo fa sempre 180°
 - Un sifone funziona sempre allo stesso modo, quale che sia il materiale di cui è fatto il tubo – gomma piombo o ceramica – e quale che sia il fluido – acqua olio o vino.
 - Invece, nel lancio di una moneta, la probabilità che venga testa è del 50%: ma che significa? Non ha senso, per un singolo lancio. E non mi garantisce che, su 100 lanci, in 50 venga testa.
 - Lo scandalo della filosofia [Ian Hacking, *The Emergence of Probability*, 1975]
 - Probabilità come il nesso tra ipotesi ed evidenza, ma il concetto di evidenza non esisteva
 - Quindi, niente induzione

Sommario

- I numeri per capire il mondo
- Le parole per capire i numeri
- La statistica è facile o difficile?
- **Statistica, fenomeni collettivi, modelli**
- La statistica è di tutti
- Sviluppare la nostra comprensione delle statistiche

La statistica e i fenomeni collettivi [1]

- Stai guidando nella nebbia e la macchina davanti a te inchioda
 - Ti attacchi al clacson e pensi: «Ma che fai? Levati di mezzo, pericolo pubblico»
- Poi la nebbia si solleva, e vedi che davanti a te, a perdita d'occhio, c'è una fila di macchine ferme
 - Di colpo, pensare che quello davanti a te sia un cretino o che gli si è rotta la macchina diventa un'idiozia
 - Non serve più la psicologia o la competenza di un meccanico
- Il livello a cui devi cercare la spiegazione diventa un altro: che cosa ha causato l'ingorgo? Un incidente? Lavori in corso più avanti? Un temporale?
- È necessario fare un'astrazione. Passare da una spiegazione causale individuale e locale, allo studio di un fenomeno collettivo

La statistica e i fenomeni collettivi [2]

- La macchina davanti a te è indubbiamente la causa immediata del tuo stop, ma al tempo stesso è irrilevante per il fenomeno che vuoi capire, è una pedina in un gioco più grande
- Capire la causa di quello che sta succedendo richiede di spostarsi di livello, da un livello micro a un livello macro
 - Il ruolo delle singole macchine resta reale, ma per capire davvero devi prendere in considerazione delle strutture più grandi e le loro interazioni più astratte
- Qui entra in gioco – anche storicamente – la statistica
 - Pensiamo – tanto per restare in tema – a quello che succede nel motore: scoppia la scintilla, il gas scaldandosi si dilata e quindi spinge il pistone
 - Quello che accade alle singole molecole del gas, al livello micro resta assolutamente vero, ma è irrilevante
- La statistica fornisce un modello per collegare le proprietà di atomi singoli e molecole alle proprietà macroscopiche dei materiali che vediamo nella vita quotidiana

Pensare per modelli

- Un modello astratto (o concettuale) è una costruzione teorica che rappresenta processi fisici, biologici o sociali, con un insieme di variabili e un insieme di relazioni logiche e quantitative tra loro
- In questa accezione, il modello consente di ragionare all'interno di uno schema logico astratto e semplificato:
 - Astratto (idealizzato) perché il modello può formulare ipotesi esplicite di cui è noto che – a un certo livello di dettaglio – sono false
 - Semplificato perché ciò consente di pervenire a soluzioni ragionevolmente accurate, trascurando la complessità implicita nel grande numero di variabili e attori del processo modellizzato



I modelli in statistica

- Albert Einstein 1933
 - «Models should be as simple as possible, but not simpler»
- George Box
 - 1976: «Since all models are wrong the scientist cannot obtain a 'correct' one by excessive elaboration»
 - 1978: «All models are wrong but some are useful» è il titolo di un paragrafo → «...parsimonious models often do provide remarkably useful approximations [...]. For such a model there is no need to ask the question 'Is the model true?'. If 'truth' is to be the 'whole truth' the answer must be 'No'. The only question of interest is 'Is the model illuminating and useful?'»
- I modelli sono costosi (Boltzmann, Shannon, Landauer)

Sommario

- I numeri per capire il mondo
- Le parole per capire i numeri
- La statistica è facile o difficile?
- Statistica, fenomeni collettivi, modelli
- **La statistica è di tutti**
- Sviluppare la nostra comprensione delle statistiche



La statistica per pochi

- Per tutto il Settecento e gran parte dell'Ottocento, l'informazione statistica era un segreto di Stato
 - Il potere insito nell'informazione statistica è noto da secoli ai governi
 - L'appetito di Napoleone per i dati del suo *bureau de statistique* è leggendario
 - Ancora allo scoppio del caso Dreyfus, alla fine dell'Ottocento, i servizi segreti militari si chiamano *bureau de statistique*
- La volontà di rendere pubblici i dati economici e demografici è abbastanza recente (circa 1830)



La statistica per tutti

- La crescente disseminazione di informazione statistica ai cittadini – durante l'Ottocento e il Novecento – è associata alla diffusione della democrazia
- Una caratteristica delle democrazie è di rendere disponibile una grande massa di informazioni statistiche ai cittadini
 - H. G. Wells, padre della fantascienza, scrive: «non appare lontano il tempo in cui si capirà che la piena cittadinanza [...] richiede la capacità di calcolare, di pensare in termini di medie, minimi e massimi, come oggi è necessario saper leggere e scrivere» [*Mankind in the Making*, 1911]
 - Ronald A. Fischer, un grande statistico: «la statistica è l'aspetto particolare del progresso umano che ha dato la sua impronta al XX secolo» [1954]
 - Hal Varian, Chief Economist di Google: «Continuo a dire che il lavoro sexy nei prossimi dieci anni sarà quello di statistico» [*The McKinsey Quarterly*, 2009]

Eppure...

- L'informazione che riceviamo è utile a noi?
 - Ci aiuta a fare scelte migliori, a prendere decisioni con maggiore cognizione di causa, a evitare qualche rischio?
 - Oppure è utile soltanto a chi ce la trasmette, e cerca più o meno consapevolmente di influenzarci, di manipolarci, di far pendere a suo favore il piatto della bilancia, come accade con la pubblicità, ma anche con la politica e gli organi d'informazione?
- Ci possiamo fidare dell'informazione che riceviamo?
 - Dato che sullo stesso argomento riceviamo più informazioni e da più fonti, come facciamo a sapere di quale ci possiamo fidare?
 - Qual è il dato giusto, o comunque più attendibile?
 - Quale – quanto meno – è condiviso da più persone?
 - Su quale c'è il consenso degli esperti?

Sommario

- I numeri per capire il mondo
- Le parole per capire i numeri
- La statistica è facile o difficile?
- Statistica, fenomeni collettivi, modelli
- La statistica è di tutti
- **Sviluppare la nostra comprensione delle statistiche**

Come possiamo sviluppare la nostra comprensione delle statistiche

- *Numeracy* → *data literacy*
- Diffidare dello *story-telling*
- Dialogare con gli esperti
- Fiducia e affidabilità
- Valutazione critica

Numeracy e data literacy

- *Numeracy*: la capacità di applicare al ragionamento concetti quantitative e numerici (ad esempio, il senso del numero, il senso dell'operazione, la quantificazione, la misurazione, la computazione, la probabilità, i modelli statistici)
 - La *numeracy* statistica ha aspetti specifici, come abbiamo visto
- *Data literacy*: la capacità di comprendere e ragionare con i dati, o argomentazioni che utilizzano i dati, e di applicare questa competenza nelle deliberazioni e decisioni che riguardano la vita personale, sociale e politica
- In questo contesto, la *data literacy* è fondamentale per la missione della statistica ufficiale e la *numeracy* è un insieme di strumenti e competenze, una *cassetta degli attrezzi*

Diffidare dello *story-telling*

- La falsificazione intenzionale (*fake news*) non è il problema principale perché si può sfatare abbastanza facilmente, ad esempio con il *fact checking*
- Un rischio molto più grande è quello della «manipolazione ed esagerazione attraverso un'interpretazione inappropriata di 'fatti' che possono essere tecnicamente corretti ma sono distorti da quelle che potremmo chiamare 'pratiche discutibili di interpretazione e comunicazione'» (Spiegelhalter)
- La narrazione è spesso il vettore di queste pratiche
- Se pensate a una tabella in cui nelle righe sono rappresentate le unità e nelle colonne le variabili, il metodo della narrazione è leggere per righe, quello della statistica leggere per colonne (Desrosières)

Dialogare con gli esperti

- Il punto di forza degli istituti di statistica è il loro capitale umano
 - Un indizio di questo potenziale è il divario tra i cittadini britannici che si fidano dell'accuratezza delle cifre ufficiali (78%) e quelli che si fidano dell'ONS, l'istituto nazionale di statistica (90%) (NCSR 2017)
- Gli esperti in grado di mettere insieme dati e metadati, dando informazioni sul significato e l'utilità delle statistiche ufficiali, possono innescare un circolo virtuoso, in cui produttori e utenti co-evolvono nella loro capacità di utilizzare i dati
- L'accesso a queste competenze deve essere libero e facile

Fiducia e affidabilità

- *Statistics: a matter of trust* (UK Green Paper 1998)
- Se vuoi la fiducia dei tuoi utenti, devi essere tu stesso affidabile: «La fiducia è la risposta, l'affidabilità è ciò che dobbiamo giudicare» (O'Neill):
 - Dimostrare competenza, onestà e affidabilità
 - Rendersi vulnerabili, fornendo agli altri i mezzi per verificare se si è affidabili
 - Diverso dalla *peer review*, più vicino ai modi di creare reputazione online
 - Piattaforme per gli utenti

Grazie

Giovanni A. Barbieri
gabarbieri[at]gmail.com

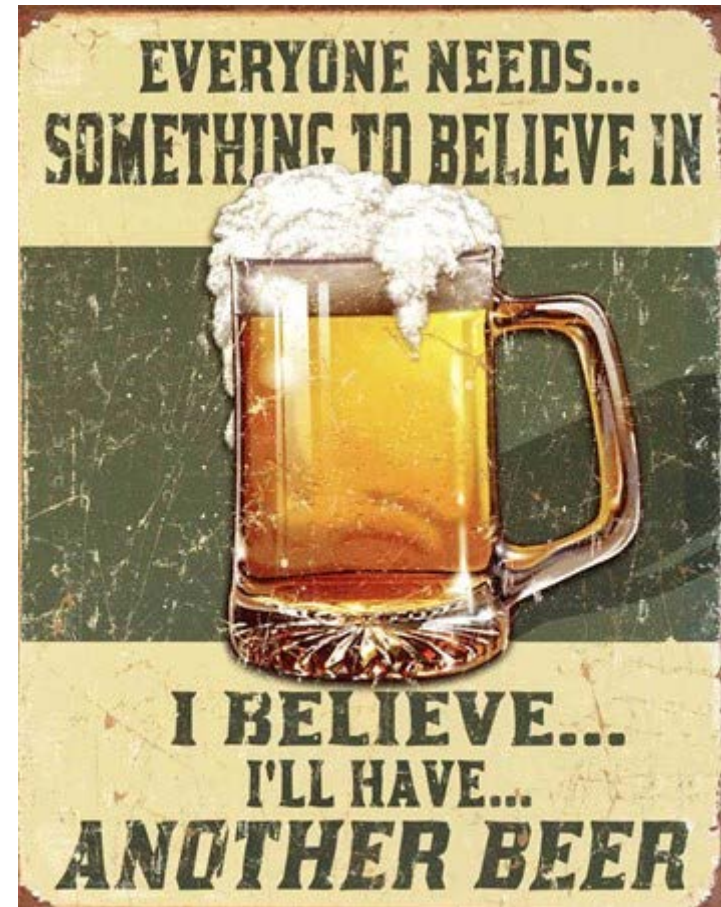
Quantificare e misurare

- Quantificare: introdurre precisi elementi di valutazione quantitativa; tradurre in termini di quantità, tradurre in cifre, in valori numerici
- Misurare: determinare la misura di una grandezza (lunghezza, distanza, superficie, peso...), confrontandolo con uno standard predefinito
- Il ‘buon senso quantitativo’ (Laplace) aiuta a prendere decisioni meglio informate e – soprattutto – consente di vedere se e quanto ci stiamo avvicinando o allontanando da un obiettivo



Parole da usare con cautela nel discorso scientifico: credere

- Dire 'io credo che' nel senso di 'io sono dell'opinione che' è del tutto corretto
- Però 'credere' ha anche il significato di 'essere certo dell'esistenza di qualcuno o di qualcosa, con riferimento a convinzioni religiose' [Vocabolario Treccani]
- Questo significato influenza per contiguità le altre accezioni
- Il massimo rispetto per i credenti; scarsa considerazione per i creduloni; nessuna per chi esprime con prepotenza opinioni non suffragate da sufficienti argomentazioni e spacciate per evidenti



Parole da usare con cautela nel discorso scientifico: prova

- Anche ‘prova’ è una parola dai molti significati: il Vocabolario Treccani ne elenca più di dieci
- In un Paese intriso di cultura giuridica richiama immediatamente un contesto giudiziario
- In inglese c’è una differenza tra *proof* and *evidence*
 - La scienza non prova
 - La scienza adduce *evidenze* pro o contro una particolare ipotesi
 - Nella scienza, le *prove* sono possibili solo nel mondo altamente astratto della matematica
- «Nella scienza, ‘fatto’ può significare solo ‘confermato a tal punto che sarebbe perverso negare il consenso provvisorio’» [Stephen Jay Gould, «Evolution as Fact and Theory», pp. 254–55 (originariamente su *Discover*, maggio 1981)]

Parole da usare con cautela nel discorso scientifico: dimostrare

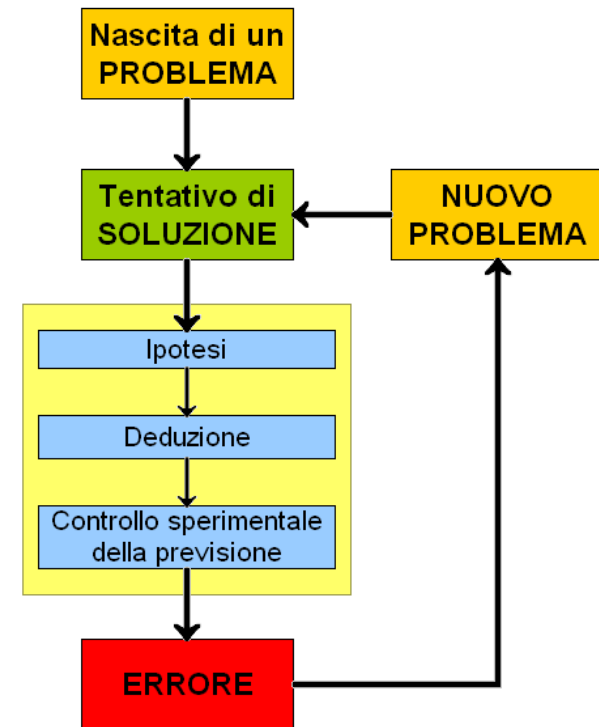
- Dovrebbe essere chiaro dalla discussione su ‘prova’ ed ‘evidenza’: soltanto la matematica *dimostra*, gli esperimenti scientifici, che raccolgono evidenze, *mostrano*
 - Anche in questo caso la cultura giuridica può trarci in inganno, perché diciamo ‘dimostrare la propria innocenza’
 - Dimostrare: «Provare con ragionamenti, con esperimenti o in altro modo la verità di un principio, di una tesi, l’esattezza di quanto si afferma...» [Vocabolario Treccani]
 - Mostrare: «Far vedere, presentare ad altri perché veda, esamini, osservi...» [Vocabolario Treccani]

Parole da usare con cautela nel discorso scientifico: teoria

- Nel linguaggio quotidiano usiamo 'teoria' per riferirci a un'opinione non suffragata: si riferisce a un modo soggettivo di pensare, a una possibilità astratta [Vocabolario Treccani]
 - «La mia teoria è che le chiavi della macchina le abbia prese il monacello»
- Nel linguaggio scientifico, questa è un'ipotesi: «prima formulazione di una legge, non ancora sperimentata o sperimentabile in sé, ma tale da rendere ragione di fatti sperimentali o sperimentabili» [Vocabolario Treccani]
- «Alcune spiegazioni scientifiche sono così ben stabilite che verosimilmente nessuna nuova prova può alterarle. La spiegazione diventa una teoria scientifica. Nel linguaggio comune una teoria significa un'intuizione o una speculazione. Non è così nella scienza. Nella scienza, la parola teoria si riferisce ad una spiegazione completa di una caratteristica importante della natura suffragata da fatti raccolti nel tempo. Le teorie permettono anche agli scienziati di fare previsioni su fenomeni non ancora osservati» [National Academies of Science]
- «Una teoria scientifica è una spiegazione ben documentata di alcuni aspetti del mondo naturale, basata su un insieme di fatti che sono stati ripetutamente confermati attraverso l'osservazione e l'esperimento. Tali teorie sostenute dai fatti non sono "congetture", ma resoconti affidabili del mondo reale» [American Association for the Advancement of Science]

Parole da usare con cautela nel discorso scientifico: verificare

- La scienza non procede per verifiche, ma per confutazioni
 - «Nessuna quantità di esperimenti potrà dimostrare che ho ragione; un unico esperimento potrà dimostrare che ho sbagliato.» [Albert Einstein, lettera a Max Born del 4 dicembre 1926]
- Il *criterio di falsificabilità* – introdotto da Karl Popper proprio a partire dalla frase di Einstein – afferma che una teoria, per essere scientifica, deve essere confutabile
 - Dalle sue premesse di base devo poter dedurre almeno un esperimento che, qualora sia errata, la possa confutare alla prova dei fatti



Illusioni ottiche

- 3 esempi
 - [Il cubo di Necker](#)
 - [Ragazza o befana?](#)
 - [Lo vedete?](#)
- Come funziona?
 - Nel primo caso:
 - La profondità è ambigua per la mancanza voluta degli indizi consueti
 - Non percepiamo un'ambiguità, ma un'oscillazione: *Gestalt switch* (mai contemporaneamente!)
 - Nel terzo:
 - Una volta 'vista' la figura nascosta, non puoi smettere di vederla
 - Il tuo sistema percettivo è cambiato permanentemente

Illusioni cognitive

■ Crescita esponenziale

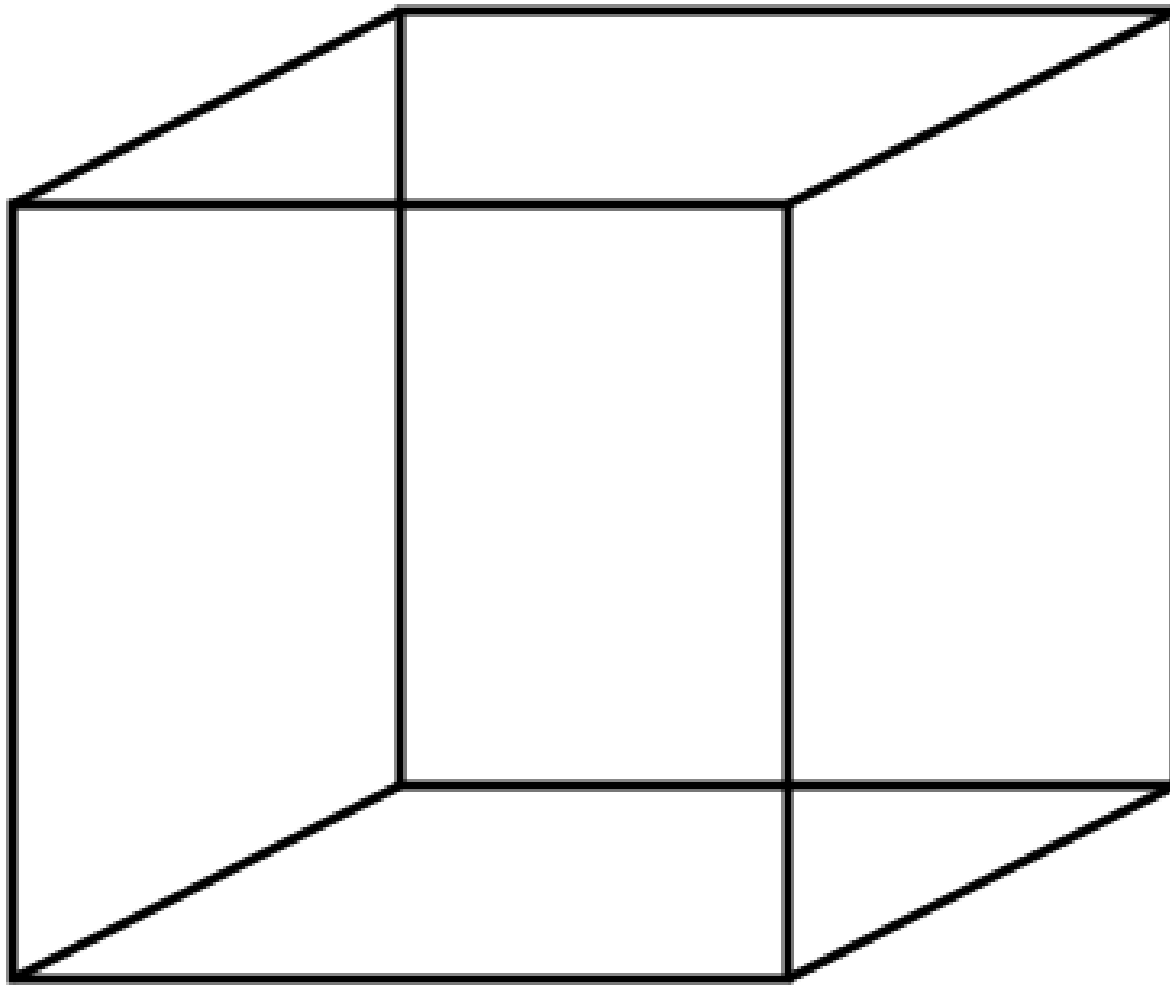
- Manchiamo di un modulo mentale per capire la crescita esponenziale, e l'approssimiamo con la crescita lineare
- Un modo di ragionare 'intuitivamente lineare': diamo per scontato che il tasso di sviluppo corrente continuerà nel futuro
- Per la crescita lineare, il passato è una buona guida per il futuro; per quella esponenziale, è una guida ingannevole

■ Decisioni razionali e buon senso

- Daniel Kahneman. *Pensieri lenti e veloci* (trad. Laura Serra). Milano: Mondadori. 2012.
- Non siamo massimizzatori dell'utilità probabilisticamente razionali
- Due modi di pensare:
 - Sistema 1: veloce, istintivo, emotivo
 - Sistema 2: lento, logico, deliberativo
- Sostituiamo un contesto di scelta difficile con uno che ci è più familiare

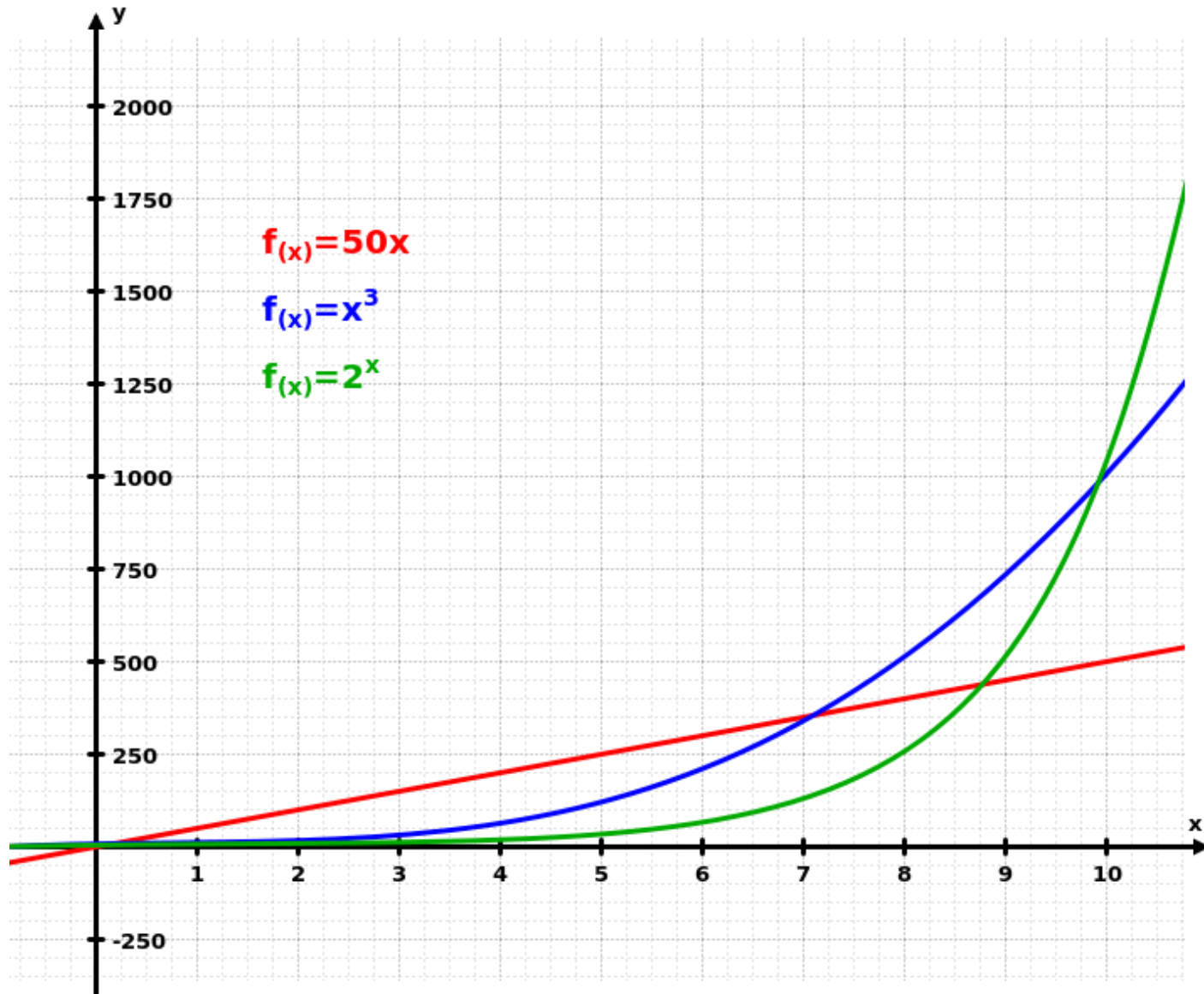












Statistica e democrazia

- La parola *statistica* viene dal latino *statisticum collegium* (consiglio di stato) ed è apparentata con *statista*
 - *Statistica* è il femminile sostantivato dell'aggettivo *statistico*, il cui significato originario era 'che riguarda lo Stato, la vita e i problemi dello Stato'
 - Il tedesco *Statistik* fu introdotto da Gottfried Achenwall (1749) e faceva riferimento all'analisi di dati sullo Stato
- Assume il significato di raccolta e classificazione dei dati all'inizio dell'Ottocento
 - In questo modo, il principale scopo originario della statistica riguarda dati a uso del governo e dell'amministrazione (spesso centralizzata)

