



A proposito di *Mentire con le statistiche* – htlws.it

Un articolo di Giancarlo Livraghi  
nella rivista *l'attimo fuggente*

ottobre 2009

# Il pollo di Trilussa e gli inganni delle statistiche

*“La matematica – si usa dire – non è un’opinione”.  
Ma molte valutazioni numeriche sono opinabili.  
La materia è complessa, ma alcuni criteri di buon senso  
possono aiutarci a capire un po’ meglio  
il significato di dati e statistiche.*

**S**iamo inondati dalle statistiche. Su ogni sorta di argomenti ci somministrano numeri di ogni specie. Raramente abbiamo il tempo di chiederci che cosa vogliamo dire (se non, talvolta, nei casi vistosi in cui sullo stesso argomento si pubblicano e si diffondono dati completamente diversi).

Anche molte notizie e valutazioni che non sono esplicitamente “numeriche” derivano da qualche specie di dato statistico che non è neppure citato, perciò si rischia di crederci senza sapere su che cosa è basata un’opinione troppo spesso presentata come se fosse un fatto o un’incontestabile verità.

Lo fanno apposta? Spesso sì. Dati e statistiche si possono manipolare per tentare di “dimostrare” questa o quella tesi, ipotesi o preconetto. Ma è ancora più frequente che si tratti di errori o superficialità – cose citate o riferite senza verificarne l’attendibilità o il significato. Non è raro che siano in gioco tutti e due i fattori (deformazioni consapevoli e stupidaggini involontarie) con conseguenze che sarebbero comiche se non fossero pericolose.

Lewis Carroll, “al secolo” Charles Dodgson, non scriveva solo le ambigue favole di Alice. Era un matematico. Questa è una sua osservazione. «*Se vuoi ispirare fiducia, dai molti dati statistici. Non importa che siano esatti, neppure che siano comprensibili. Basta che siano in quantità sufficiente*».

C’è chi si diverte a fare statistiche bizzarre, su argomenti futili e inutili. Sarebbe solo una curiosità se in quel modo, con ogni sorta di errori e approssimazioni, non si inquinassero anche i pettegolezzi e le mode.

C'è chi si diverte a propalare numeri completamente inventati, “per vedere l'effetto che fa” – e se ne sta incuriosito a vedere quanti ci cascano (come fanno gli inventori di “leggende metropolitane”, di cui alcune nascono per caso, ma altre sono seminate apposta per verificare la credulità non solo dell'umanità in generale, ma anche dei grandi sistemi di cosiddetta informazione).

Ma le statistiche sono una cosa seria. Usate bene, sono uno strumento di notevole utilità. Lo sa per esperienza chi, come me, si è trovato molte volte a doversene servire per motivi di studio e di lavoro. E poiché ne deve trarre conseguenze significative, è costretto a capire come funzionano – e perciò ad approfondire l'origine e la natura dei dati. Spesso scoprendo che il significato è molto diverso da quello che sembra (o che i numeri sono privi di significato ed è necessario non tenerne conto per non cadere in pericolosi errori).

Così ho imparato molto – e continuo a imparare. Non dal punto di vista di chi produce dati, ma da quello di chi li usa per trarne deduzioni concrete. E se non vuole sbagliare deve capire bene qual è l'origine dei numeri e come si possono interpretare.

## Il fascino dei numeri – pro e contro

Prima di parlar male dei numeri è necessario dirne bene. Sono uno strumento essenziale della ricerca scientifica. <sup>1</sup> Perfino uno come me, ignorante e impreparato nelle sottigliezze delle formule e dei calcoli, è affascinato dall'evoluzione degli studi sull'infinitamente grande e sull'infinitamente piccolo, dalla molteplicità di scoperte nella fisica, nella biologia, eccetera. Non è insensato dire che la vita è un algoritmo. Ma non bastano le formule matematiche per capire che cosa significa.

Naturalmente anche agli estremi confini della conoscenza ci sono incertezze, dubbi, diversità di interpretazione. Sono così sottili che possono essere capite solo da studiosi con una specifica competenza in materie difficili e complesse. Ma è chiara un'osservazione di Albert Einstein. «*Quando le regole della matematica si riferiscono alla realtà non sono certe – e quando sono certe non si riferiscono alla realtà*».

Se questo è vero alle frontiere estreme della ricerca scientifica, lo è ancora di più nel caso dei dati che imperversano su cose più semplici e facilmente comprensibili. Il fascino dei numeri può essere falso e bugiardo. Non è vero che qualsiasi baggianata diventa verità quando è espressa in termini numerici. Spesso è vero il contrario. E perciò è utile capire quali sono gli errori e gli inganni delle statistiche.

Si dice che le statistiche siano un'invenzione recente (nel diciassettesimo secolo – più diffusamente nel diciannovesimo). Ma, se fosse così, come si spiegherebbe il commento di Platone? «*Sappiamo bene che queste argomentazioni basate sulle probabilità sono imposture e se non abbiamo molta cautela nel loro uso possono essere ingannevoli*». Anche nel mondo antico si ragionava sulle probabilità. E anche allora si sapeva che non è facile.

## La statistica è un'arma?

Le statistiche non uccidono (se non in alcuni casi, come quando sono l'origine di un errore medico o della mancata prevenzione di un disastro). Ma l'informazione è un'arma, spesso usata come tale. Lo insegnava Sun Zu, nel suo trattato sull'arte della guerra. E lo sapevano bene, prima di lui, anche altri, lontani dalla Cina.

---

<sup>1</sup> Non è vero che, come alcuni affermano, la matematica sia usata dalla scienza solo nell'evo moderno. Benché in modo meno complesso, era uno strumento della conoscenza (e della tecnologia) anche nell'antichità. Vedi *Il computer di Archimede* [gandalf.it/arianna/olimpia.htm](http://gandalf.it/arianna/olimpia.htm) sull'uso di macchine da calcolo in epoca ellenistica.

Per esempio le potenti legioni romane vincevano non solo per la forza delle armi, l'efficacia delle strategie e l'addestramento dei militi, ma anche perché sapevano come avere informazioni e come confondere quelle dell'avversario.

E l'esito della guerra di Troia potrebbe essere stato diverso se i troiani avessero dato retta a Cassandra e a Laocoonte invece di cadere nell'inganno del cavallo.

Così è stato nei secoli e nei millenni. E così è ancora oggi. Le statistiche sono uno strumento dell'informazione, perciò sono utili in guerra, per organizzarsi in modo vincente se si hanno dati validi e anche per ingannare gli avversari con numeri falsi o manipolati. Diceva Winston Churchill: «*le sole statistiche di cui ci possiamo fidare sono quelle che abbiamo falsificato*».

Ma anche quando non si tratta di imprese militari sono un'arma nei conflitti o contrasti politici, economici, commerciali e culturali. E anche quando, anziché competere, si cerca un'intesa.

Si vince con le statistiche avendo informazioni migliori, più attendibili e perciò più utili. Ma non basta avere dati. Si tratta anche, o soprattutto, di saperli capire e gestire.

## Trilussa e il pollo

La più “proverbiale” osservazione a proposito delle medie statistiche è quella per cui se qualcuno mangia un pollo, e qualcun altro no, in media hanno mangiato mezzo pollo. (Ce ne sono altre di uguale significato, come il caso di una persona annegata in un fiume con un profondità media di mezzo metro).

L'osservazione non è così ovvia come può sembrare. Ma cominciamo col fatto che l'origine è un bel sonetto di Trilussa (che definisce il problema in modo un po' diverso).

### *La Statistica*

*Sai ched'è la statistica? È 'na cosa  
che serve pe' fa' un conto in generale  
de la gente che nasce, che sta male,  
che more, che va in carcere e che sposa.  
Ma pe' me la statistica curiosa  
è dove c'entra la percentuale,  
pe' via che, lì, la media è sempre eguale  
puro co' la persona bisognosa.  
Me spiego: da li conti che se fanno  
secondo le statistiche d'adesso  
risurta che te tocca un pollo all'anno:  
e, se nun entra ne le spese tue,  
t'entra ne la statistica lo stesso  
perché c'è un antro che ne magna due.<sup>2</sup>*

La fonte originale è meglio della “vulgata”. Non solo per lo stile e per l'umorismo, ma anche perché imposta il concetto in modo più chiaro.

Oggi, con una più evoluta cultura della statistica, nessuno cade più in quel genere di errore? La cosa non è così semplice. I dati possono essere generati e interpretati in modi diversi, spesso portando a risultati intenzionalmente ingannevoli o sbadatamente deformanti.

---

<sup>2</sup> Ai tempi di Trilussa mangiare pollo era considerata “una cosa da ricchi”.  
Oggi la situazione è diversa. Ma non cambia il significato del ragionamento.

Ci si dimentica, un po' troppo spesso, che la "media" è un dato poco significativo se non sappiamo a che cosa si riferisce, su quale base è calcolata, con quale criterio è definita. Come dice, beffardo, Des McHale, *«l'umano medio ha una mammella e un testicolo»*. Più seriamente Aldous Huxley ci ha insegnato che *«nella vita reale non c'è alcun uomo medio»*.

La media, comunque calcolata, è un concetto astratto. Una delle poche certezze assolute della statistica è che ciò che è "medio" non esiste. Ogni cosa si colloca necessariamente sopra o sotto il dato "medio". E non è solo una questione di aritmetica.

## L'uomo medio non esiste

Sarebbe lungo riassumere i molteplici problemi di complesse diversità che possono essere ingestibili se si basa a una media, comunque calcolata. Ma l'importante è diffidare di ogni concetto "standardizzato", anche quando non è espresso in forma di media numerica.

Per esempio accade spesso di leggere o sentir dire che "tutti" fanno o pensano qualcosa. Con un po' di approfondimento non è difficile scoprire che quei "tutti" in realtà sono pochi – o che una certa "media", chissà come calcolata, non riflette alcuna persona o situazione reale.

Se è facile capire che "tutti mangiano mezzo uovo al giorno" è un'affermazione priva di senso, perché c'è chi ne mangia di più e chi di meno, altre generalizzazioni altrettanto false sono più insidiose. È sempre bene ricordare che la media è solo un indice numerico, il cui significato può variare molto secondo i fattori di "dispersione" dei dati e secondo i criteri di analisi adottati in ciascun caso.<sup>3</sup>

Del resto se fossimo tutti uguali – e tutti facessimo le stesse cose – il mondo sarebbe terribilmente noioso. Come sono noiosi (oltre che pericolosi) i discorsi, i proclami e i ragionamenti di chi ha la pretesa di "standardizzarci" – magari sperando di poterci ridurre a "numeri" neutri e ripetitivi che soddisfino qualche suo arbitrario schema mentale.

Se alle origini dell'automobile poteva avere un temporaneo significato il primo prodotto fabbricato in serie, la mitica Ford T – uguale per tutti e di un solo colore (nero) – si scoprì presto che anche le linee di montaggio dovevano essere in grado di produrre modelli diversi. Perché le persone (e le loro esigenze) non sono uguali.

L'industria dell'abbigliamento sa che se producesse per una sola taglia, quella dell'immaginario "uomo medio", i capi sarebbero adatti solo a una piccola parte della popolazione (le persone poco sotto o poco sopra la media). Ma in altri campi sono accadute cose bizzarre.

Per esempio è un fatto vero, non una barzelletta, che un grosso progetto edilizio negli Stati Uniti fu impostato con l'idea di fare tutte le case per quattro persone, perché allora era quella la "famiglia media". Furono costretti a grossi e costosi cambiamenti "in corso d'opera" quando si accorsero di un fatto ovvio: molte famiglie erano più grandi o più piccole dello standard.

Può sembrare solo un episodio di insolita sconsideratezza. Ma si potrebbero citare parecchi altri esempi, anche recenti. Non sono pochi i casi in cui per badare alla "media" si dimentica la diversità. Sembra comodo pensare che esista un "uomo medio" o "uomo comune". Ma presto o tardi ci si accorge di avere a che fare con l'inesistente.

---

<sup>3</sup> A proposito di polli e di uova, c'è il caso di una pessima gestione delle statistiche e dei dati numerici nella tragicomica epidemia di infondate paure per una infezione "aviaria" nel 2005 – come è spiegato a pagina 195 di *Mentire con le statistiche*, un libro di cui si parlerà poco più avanti.

## Gli inganni delle parole

A peggiorare le cose c'è il linguaggio con cui si interpretano o si descrivono i dati. Spesso afflitto da manierismi che ne deformano il significato. Per esempio accade di leggere o sentir dire che “tutti” fanno o pensano qualcosa. Con un po' di attenzione non è difficile scoprire che quei “tutti” sono pochi – o che una certa “media”, chissà come calcolata o immaginata, non riflette alcuna persona o situazione reale.

Fra i manierismi c'è anche l'uso (e abuso) dell'aggettivo “esponenziale”.

A parte il fatto che la parola ha un significato preciso, raramente applicabile a una crescita o a una tendenza, la si usa a casaccio per ogni sorta di vicende il cui andamento non somiglia neppure remotamente a un “moto uniformemente accelerato”.

Un'altra terminologia superficiale (quando non è intenzionalmente ingannevole) è quella che dice “vola” quando qualcosa aumenta dell'uno o due per cento o “crolla” quando scende di una altrettanto modesta misura.

Un trucchetto largamente usato è il piccolo avverbio “ben”. È sostanzialmente inutile, perché i numeri, quando hanno un significato, sono chiari da soli. Ma comunque è un' enfasi spesso deformante.

Se qualcuno ci dice che un bosco ha “ben mille” alberi, vuol farci intendere che è un bosco importante e che altri (quali non si sa) ne hanno meno. Se invece dicesse che ne ha “solo mille”, vorrebbe farci pensare che sia un bosco piccolo (in confronto a cosa?) o che abbia pochi alberi di quella specie. Spesso quelle “paroline” sono inserite a caso, senza alcuna base che ne giustifichi il significato.

Questi sono solo alcuni di tanti possibili esempi. Fra manierismi e “modi di dire”, usi banali o deformanti del linguaggio, deduzioni arbitrarie o infondate, il problema non è soltanto nelle statistiche, ma anche nel modo di interpretarle e spiegarle. E quando (come succede spesso) le spiegazioni sbagliate si accumulano con le inesattezze dei dati il risultato è una moltiplicazione di insensatezza.

Vuol dire che siamo irrimediabilmente confusi, in balia del pressapochismo e della disinformazione? Per fortuna no. Vedremo alla fine come è possibile difendersi. Ma intanto, per cominciare, il passo fondamentale è una sana diffidenza. I dati, statistici o non, se usati bene, possono aiutarci a capire. Ma è necessario ricordare che nulla è mai più “certo” o credibile solo perché è espresso in numeri o accompagnato da dati numerici.

## La ripetizione non è una conferma

Possiamo immaginare che se le stesse statistiche, o le stesse osservazioni, sono largamente ripetute in diverse occasioni e da diversi autori, se ne può dedurre che sono credibili e ben fondate. Ma non è vero.

Accade spesso che un dato, una notizia o un'opinione, arbitrariamente o incautamente pubblicata da uno, sia ripresa acriticamente da altri e abbia un'enorme diffusione senza alcuna verifica sull'attendibilità della sua origine. Talvolta una “bufala” può sopravvivere per millenni (per esempio non c'è mai stata alcuna prova attendibile che Nerone avesse incendiato Roma).

Può essere esagerato affermare, parafrasando Proust, che la diffusione di una notizia è inversamente proporzionale alla sua credibilità. Ma è un fatto che molte cose considerate “vere” non hanno alcun fondamento se non il fatto che sono così diffuse da sembrarlo. E accade con dati e statistiche come con ogni altro genere di informazioni.

C'è un'efficace sintesi di questa sindrome in un'osservazione di Alessandro Manzoni. «*Il buon senso c'era, ma se ne stava nascosto, per paura del senso comune*».

Il problema non sta tanto nell'esistenza degli errori, che sono sempre possibili, quanto nella diffusa abitudine di accettare dati sballati e incoerenti come se fossero "certezze" indiscutibili – e di ripeterli *ad infinitum* senza mai verificarne la credibilità.

## La gatta frettolosa

Un'attivissima, imperversante fabbrica di errori è la fretta. Dati e statistiche (come ogni sorta di informazione) richiedono verifiche attente. Accade che l'ansia di essere veloci, di "dare la notizia" o di decidere senza approfondire, porti alla diffusione di affermazioni infondate, talvolta grottesche – o provochi scelte sbagliate che poi è faticoso correggere.

La stessa fretta induce a copiare, cioè a ripetere ciò che ha detto o fatto qualcun altro, senza avere il tempo o la voglia di controllare. Così gli errori si moltiplicano e continuano a riprodursi per giorni, o mesi, o anni, fino a diventare un "patrimonio culturale" di assurdità.

Non è vero che gli struzzi mettono la testa sotto la sabbia (se fossero così stupidi, a quest'ora sarebbero estinti). Ma lo fanno troppo spesso gli umani, per fretta, distrazione o superficialità.<sup>4</sup>

## Un libro utile e interessante

Un libro "unico nel suo genere" è stato pubblicato in America nel 1954. Si chiama *How to Lie with Statistics*, scritto da Darrell Huff – un giornalista, non un matematico. È un brillante esempio di buona divulgazione. Se ne sono vendute, in inglese, oltre mezzo milione di copie. È meritatamente il più diffuso fra tutti i libri che siano mai stati scritti sull'argomento.

Per il suo tono ironico e per la scrittura scorrevole, spesso divertente, comunque facile anche per i "non iniziati", è stato dapprima disprezzato dal mondo scientifico. Ma poi si è capito che da un punto di vista tecnico è molto preciso e ben documentato. Tanto è vero che è stato celebrato dall'*Institute of Mathematical Statistics* nel 2005 per il suo cinquantesimo anniversario.

Ne sono uscite traduzioni in varie lingue (compreso il cinese) ma non in italiano – fino a quando, due anni fa, è stato finalmente pubblicato *Mentire con le statistiche*.<sup>5</sup> C'è un sito online [htlws.it](http://htlws.it) che, oltre a presentare il libro, contiene anche parecchie altre osservazioni sull'argomento.<sup>6</sup>

Si presenta, a prima vista, quasi come se fosse un manuale per gli imbroglioni. Ma è esattamente il contrario. È uno strumento per capire come dati, statistiche e numeri possono essere deformati – e così praticare l'arte necessaria del dubbio e distinguere i dati utili e significativi da quelli che, per intenzione o per errore, ci danno percezioni sbagliate.

---

<sup>4</sup> Sulla sindrome della fretta vedi il capitolo 16 di *Il potere della stupidità* (anche online [gandalf.it/stupid/cap16.htm](http://gandalf.it/stupid/cap16.htm)). Sulle "cose che crediamo ma non sono vere" vedi *La stupidità dei luoghi comuni* [gandalf.it/stupid/panzane.htm](http://gandalf.it/stupid/panzane.htm)

<sup>5</sup> Monti & Ambrosini editori, 2007. L'edizione italiana contiene ampie annotazioni, aggiornamenti e approfondimenti che si aggiungono all'intrinseco valore (e alla sostanziale attualità) del testo originale (vedi il sito [htlws.it](http://htlws.it)). Ce n'è una recensione in [gandalf.it/arianna/duelibri.htm](http://gandalf.it/arianna/duelibri.htm) che, oltre a questo libro, ne commenta anche un altro, di diversa impostazione ma di interessante utilità: *Quando i numeri ingannano* di Gerd Gigerenzer – edizione italiana Raffaello Cortina, 2003.

<sup>6</sup> Varie osservazioni, utili anche per uno sviluppo un po' più ampio dei temi trattati in questo articolo, si trovano in [gandalf.it/htlws/premesse.htm](http://gandalf.it/htlws/premesse.htm) e [gandalf.it/htlws/aggiunte.htm](http://gandalf.it/htlws/aggiunte.htm)

Ricordo che un giorno, parecchi anni fa, mandai una copia del libro di Huff, in inglese, a un mio amico professore di fisica. Mi aspettavo una risposta un po' sdegnosa, come "divertente ma non abbastanza scientifico". Mi sorprese, con piacere, il suo commento. «*Dovrebbe essere adottato come testo in tutte le scuole di fisica e matematica*». E infatti si trova nelle biblioteche, e in evidenza nelle librerie interne, di alcune fra le più serie università americane.

È un po' diverso il percorso dell'edizione italiana. *Mentire con le statistiche* è apprezzato dagli studiosi e dagli "addetti ai lavori" (o almeno dai più seri, coscienti del fatto che non sono "infallibili"). Ma, almeno finora, è meno conosciuto fra le persone cui è destinato, i "normali esseri umani" che non hanno voglia di addentrarsi in astruse disquisizioni tecnico-matematiche e possono imparare molto da un testo divertente quanto istruttivo.

Sembra diffusa l'opinione che un libro sulle statistiche debba necessariamente essere complicato e difficile. O noioso. Darrell Huff ha brillantemente dimostrato che non è vero.

## Il "campione rappresentativo"

Sono rari i casi in cui è possibile avere tutti i dati su tutte le persone o cose che si sta cercando di misurare. Uno dei pochi esempi è il censimento, che dovrebbe raccogliere informazioni sulla totalità della popolazione (eppure anche in quel caso ci possono essere rilevanti errori o imperfezioni).

Ci sono metodi per cui si possono fare rilevazioni da fonti "indirette". Fra tanti, un esempio semplice. Come si fa a stimare quante persone sono rimaste a Milano o a Roma in agosto? Qualcuno riesce a contarle una per una? Ovviamente no. Ma si può avere una "ragionevole stima" conoscendo la quantità di spazzatura o il consumo di acqua o di energia elettrica per uso domestico.

Un sistema diffuso è il "campione". In teoria è semplice. Si scelgono in modo "casuale" persone in numero sufficiente per avere una "significatività statistica" precisamente definibile in base a un fattore matematico (che si chiama *sigma*). Tutto chiaro e tutto preciso? Ahimè no. Perché entrano in gioco parecchi fattori che possono influire sulla qualità dei risultati.

Sarebbe lungo analizzarli tutti. Ma uno, per esempio, sta nel fatto che è difficile (e costoso) usare un campione "assolutamente casuale". Si usano perciò campioni "rappresentativi", cioè ricostruiti in base a categorie demografiche. Il metodo è concettualmente ragionevole, ma in pratica può introdurre variabili deformanti.

Molto dipende anche dal comportamento degli intervistatori, dalla formulazione delle domande, dalla struttura dei questionari.

Sullo stesso argomento si possono avere risposte molto diverse secondo il modo in cui è impostata una domanda e il momento in cui è collocata. E può influire anche il rapporto fra chi chiede e chi risponde, il luogo e la situazione in cui avviene l'incontro, eccetera.

Mi sono trovato in molti casi a dover valutare con attenzione questi fattori per poter capire il vero significato di una ricerca, che spesso non è sufficientemente chiaro se ci si limita a guardare i numeri.

## *Post hoc o propter hoc*

Spero di essere perdonato per l'uso del latino, ma così il problema è definito in molte lingue diverse (compreso l'inglese). Naturalmente non si tratta solo di statistiche.

Se fra due dati (o fatti o comportamenti) c'è un nesso, vuol dire che uno è la causa dell'altro? Si può facilmente sbagliare.

La connessione causa-effetto può essere in senso contrario a quella che sembra la più evidente. O tutti e due possono essere conseguenze di un terzo fattore. Oppure (anche se una totale casualità è rara) si può trattare di una non significativa coincidenza.

Gli esempi potrebbero essere infiniti. Ce n'è uno curioso, a modo suo esemplare, citato da Darrell Huff.

Nelle Nuove Ebridi era diffusa la convinzione che i pidocchi facessero bene alla salute. Dove stava l'errore? Nella correlazione causa-effetto. Quasi tutti da quelle parti avevano quasi sempre i pidocchi. Se a qualcuno veniva la febbre il suo corpo diventava troppo caldo per essere un'abitazione confortevole. E i pidocchi se ne andavano.

Succede solo in qualche isola del Pacifico? No. Ci sono dovunque errori di quel genere, anche recenti, in analisi apparentemente serie su ogni sorta di argomenti.

## Le insidie dei grafici

Può essere molto utile, quando è ben fatto, spiegare dati e statistiche con grafici che danno una percezione immediata di ciò che significano i dati. Ma qui si nasconde un'altra trappola. Ci sono molti modi per deformare il senso dei numeri quando si traducono in immagini. E, anche in questo caso, si può trattare di involontari errori o di intenzionali manipolazioni.

Il problema c'è sempre stato, ma si è aggravato con l'elettronica. Con vari software è facile trasformare i numeri in istogrammi o altre sintesi visive. Fin che si usano le funzioni più semplici il risultato può essere corretto ed efficace. Ma è forte la tentazione di “abbellire” con effetti decorativi che, se non sono usati con estrema cura nel rappresentare correttamente ciò che significano, possono deformare la percezione e il significato dei dati.

Nel 2004 avevo scritto un breve articolo intitolato *Il morbo di powerpoint* – in cui, fra l'altro, si parla delle insidie dei grafici.<sup>7</sup> Ha fatto il giro del mondo, con poche polemiche e molti consensi.

Le tecniche di presentazione visiva non sono necessariamente diaboliche (come diceva un'efficace vignetta di Alex Gregory pubblicata dal *New Yorker* il 29 settembre 2003). Ma sono spesso usate male, con il risultato di addormentare gli astanti o ingannarli con mirabolanti travestimenti di affermazioni prive di sostanza.

I trucchi sono tanti. Uno dei più semplici è giocare con la “scala dei valori”. Per trasformare una modesta crescita in un esaltante sviluppo – o viceversa.

È quello che Darrell Huff chiama *the gee-whiz graph* – “il grafico fantasmagorico”.

Effetti deformanti sono provocati anche dall'uso di disegni invece di istogrammi. Un'immagine si forma nella nostra mente a due dimensioni. Perciò le differenze sembrano “al quadrato” – se una cosa è tre volte più grande (o più piccola) di un'altra, l'impressione è che lo sia nove volte. Se è aumentata di un terzo, sembra che sia quasi raddoppiata.

---

<sup>7</sup> Si trova in [gandalf.it/htlws/pwp.htm](http://gandalf.it/htlws/pwp.htm) – alla fine c'è anche la vignetta “demoniaca” del *New Yorker*. Il problema dei grafici è più ampiamente descritto, con vari esempi, nel capitolo 5 di *Mentire con le statistiche*.

È ancora peggio quando si rappresentano oggetti o persone. Così la percezione è tridimensionale – e la deformazione del dato non è al quadrato, è al cubo. (Darrell Huff cita due esempi: una grande mucca, accanto a una piccola, che ingigantisce la crescita della produzione di latte – e un minuscolo rinoceronte, confrontato con uno enorme, che esagera la diminuzione della specie in Africa).

Sarebbe lungo descrivere le varie possibilità di deformazione nei grafici o in altre rappresentazioni visive. Spero che pochi esempi bastino a chiarire il problema.

## Sondaggi di opinione e ricerche di mercato

Sono utili. Specialmente se (come è sempre stato nella mia esperienza) non sono fatte in base a qualche pregiudizio o al desiderio di “dimostrare” una tesi, ma con l’unico scopo (anzi necessità) di capire il meglio possibile la situazione reale. Il problema è che troppo spesso se ne traggono deduzioni superficiali o sbagliate, perché non si approfondisce abbastanza il reale significato dei dati.

Un esempio fra mille. Nel 2006 uno studio svolto da un serio istituto comprendeva alcune informazioni sulla crescita del numero di denunce di reati. In una “classifica” per percentuali di aumento, fra le città italiane risultava al primo posto Ferrara (la storia potrebbe essere uguale in qualsiasi altra città).

Si scatenò una sfrenata bagarre, a causa del modo in cui la “notizia” era diffusa dai giornali. Allarme collettivo, imbarazzo del sindaco e del prefetto, proposte di “misure straordinarie”... perfino la bizzarra idea di “ricontare” denunce e accertamenti come se fossero schede elettorali contestate... insomma un enorme fracasso dovuto a una interpretazione frettolosa, “scandalistica” e sbagliata di dati che non giustificavano in alcun modo tutto quel parapiglia.

L’istituto aveva correttamente spiegato che il dato riguardava le denunce, non i reati, perciò si poteva trattare non di un aumento di criminalità, ma di una migliore efficienza delle forze dell’ordine e di un maggiore impegno dei cittadini nel denunciare. E comunque la significatività statistica non era tale da poter giustificare alcuna “classifica” delle città. Ma non riuscì a frenare lo sconquasso, che continuò a imperversare fino a quando (come succede quasi sempre) si esaurì fino a cadere nel dimenticatoio.

Un esempio intenzionalmente irrealista è un sondaggio che ho inventato su Cesare e Pompeo – dove con un metodo apparentemente uguale si ottengono risultati molto diversi. Forse un po’ esagerato in quella impostazione “didattica”, ma sostanzialmente simile a cose che sono accadute davvero e che continuano a ripetersi.<sup>8</sup>

## Il problema delle “proiezioni”

Non si tratta di quelle “proiezioni” che imperversano dopo ogni elezione e costringono politici e “opinionisti” a fare inutili congetture quando basta aspettare un giorno, o forse due, per avere non solo risultati definitivi, ma anche il tempo di analizzarli in modo non troppo superficiale.

C’è un altro genere di proiezioni, il cui compito è indovinare il futuro. Non sempre sono esercitazioni prive di senso. Può essere utile, in alcune cose necessario, avere una “ragionevole stima” di quale potrà essere l’estensione di un’esigenza, lo sviluppo di un problema o la disponibilità di una risorsa, fra sei mesi o fra trent’anni. Ma è bizzarro che si pubblicino grafici, tabelle o anche semplicemente “notizie” che presentano come un fatto acquisito dati riguardanti il 2020 o il 2050.

---

<sup>8</sup> Vedi *Cesare e Pompeo* [gandalf.it/htlws/pompeo.htm](http://gandalf.it/htlws/pompeo.htm)

Nessuno è in grado di fare profezie. È vero che un astronomo può prevedere un'eclissi con notevole precisione, ma poche cose hanno una regolarità costante nel tempo come i movimenti dei pianeti e dei loro satelliti.

La soluzione è semplice. Prendere i dati per quello che sono. “Se la tale tendenza verificata finora in base a ... continuerà nello stesso modo, è probabile che...”. Dove “se”, “in base a” e “probabile” sono parole chiave, troppo spesso dimenticate da chi si sbizzarrisce con le profezie.

Sono molti (e talvolta fallimentari) gli errori provocati da “proiezioni” basate sul passato senza tener conto dei cambiamenti (spesso prevedibili, o almeno ipotizzabili) che possono interferire con l'evoluzione di una tendenza. Come dice Scott Adams. «È facile fare profezie. Quando si accorgeranno che le mie sono sbagliate, sarò morto». O Niels Bohr. «La predizione è molto difficile, specialmente sul futuro».

Ma è altrettanto sbagliato essere “imprevedenti”, cioè non tener conto di problemi (o di potenzialità) che erano chiaramente “proiettabili” e sono stati ignorati o sottovalutati, per poi dover affrontare le situazioni quando è troppo tardi – e la possibilità di rimediare ai danni (o approfittare di occasioni favorevoli) è perduta o è diventata molto più difficile.<sup>9</sup>

L'importante è capire che una proiezione non è una profezia. Se e quando è utile, deve essere gestita, ricontrollata, seguita nel tempo per capire gli sviluppi. Ed è meglio anche verificare se con criteri completamente diversi si trovano tendenze che confermano (o modificano) il significato di un'evoluzione.<sup>10</sup>

Se no, si rischia di cadere nella situazione descritta da John Galbraith. «L'unica funzione della previsione economica è far sembrare rispettabile l'astrologia».

## Il rimedio? Diffidenza e buon senso

Una percezione diffusa, quanto sbagliata, è che se qualcosa è espresso in numeri è “più vero”. E sembrano ancora più credibili quando sono molto precisi, anche con i decimali, mentre spesso quei dettagli sono al di sotto della significatività statistica. Quando leggiamo “52,14 per cento”, anche se l'analisi è ben fatta su dati validi, è probabile che l'interpretazione più corretta e utile sia “circa la metà”.<sup>11</sup>

Già nel Settecento Samuel Johnson diceva: «i numeri precisi sono sempre falsi». E, vent'anni fa, Alfred Sauvy: «in ogni statistica, l'inesattezza dei numeri è compensata dalla precisione dei decimali».

Perciò la prima difesa è non lasciarsi affascinare dall'apparente esattezza dei numeri. Ma non basta. Occorre anche capire se dati e statistiche hanno un senso – e se il modo in cui sono interpretate è ragionevole.

Solo nel caso che l'argomento ci interessi in modo particolare possiamo trovare il tempo e la pazienza di controllare, verificare altre fonti, risalire all'origine dei dati. Ma in generale può essere utile la soluzione che propone Darrell Huff: “dare una seconda occhiata”. Numeri e deduzioni sembrano ragionevoli? Come coincidono o divergono da altre cose che sappiamo sull'argomento?

---

<sup>9</sup> Vedi *Il prevedibile e l'imprevedibile*, capitolo 3 di *Il potere della stupidità* (online [gandalf.it/stupid/cap03.htm](http://gandalf.it/stupid/cap03.htm)). E anche *Il (dis)senso di poi* [gandalf.it/stupid/dissenno.htm](http://gandalf.it/stupid/dissenno.htm)

<sup>10</sup> Il metodo è noto, per esempio, nella “navigazione piana”, cioè quella con la carta nautica, la bussola e il compasso (che è meglio conoscere anche se esistono i navigatori satellitari e i radar). Per “fare il punto” con ragionevole approssimazione si usano tre riferimenti diversi. (Anche in tante altre cose è utile “triangolare” – e le prospettive possono essere più di tre).

<sup>11</sup> Vedi *L'utilità dei numeri imprecisi* [gandalf.it/htmlws/approx.htm](http://gandalf.it/htmlws/approx.htm)

Serve anche (e spesso è possibile) guardare dietro la superficie. Diceva Aaron Levenstein: *«le statistiche sono come i bikini – ciò che rivelano è suggestivo, ma ciò che nascondono è più importante»*.

È meno difficile di come può sembrare. Con un po' di esercizio si impara non solo ad avere una “sana diffidenza”, ma anche a sapere come orientarla. Si tratta di qualcosa che conferma ciò che sappiamo o che potevamo immaginare? O di un'interessante scoperta che può farci cambiare idea? Se teniamo gli occhi aperti, presto o tardi avremo un modo per saperlo.

Soprattutto è utile il buon senso. Che cosa sappiamo, su quell'argomento, dalla nostra esperienza? O dall'opinione di qualcuno che se ne intende e di cui ci fidiamo? Non sono poche le possibilità di verifica.

Il concetto era stato efficacemente riassunto da Mark Twain. *«Le statistiche sono come un lampione. Le possiamo usare per fare luce, ma non come l'ubriaco, che ci si appoggia»*.

Insomma dalle statistiche si può imparare. Ma ce ne sono troppe sbagliate, ingannevoli o male interpretate. Per evitare di essere confusi o imbrogliati non occorre avere una laurea in matematica o conoscere in profondità i complessi fattori della significatività statistica. Basta sapere che prima di fidarci dei numeri è meglio capire se hanno un senso – e, se ce l'hanno, quale può essere un credibile significato.