



Mentire con le statistiche

Gli imbrogli dei dati (non è solo questione di ironia)

Giancarlo Livraghi – maggio 2008

I pasticci che succedono nell'elaborazione dei dati, su cui spesso si basano informazioni o decisioni, possono essere comici. Ma hanno spesso conseguenze tutt'altro che divertenti. Nel fondamentale libro *Mentire con le statistiche*¹ di Darrell Huff, oltre alle molte illustrazioni di Irving Geis, ci sono parecchi esempi comici – e nel secondo capitolo del *post scriptum* all'edizione italiana c'è una raccolta di citazioni, di cui alcune vivacemente ironiche.² A tutto questo si aggiungono tre vignette pubblicate da Scott Adams, nella sua serie *Dilbert*, il 7, l'8 e il 9 maggio 2008.



copyright © Scott Adams 2008

Per chi non conosce l'inglese:

il direttore chiede di usare i dati "CRS" per valutare il mercato.

Dilbert gli risponde che quei dati sono sbagliati. La stessa cosa si ripete per i dati "SIBS" (il significato delle due sigle è irrilevante).

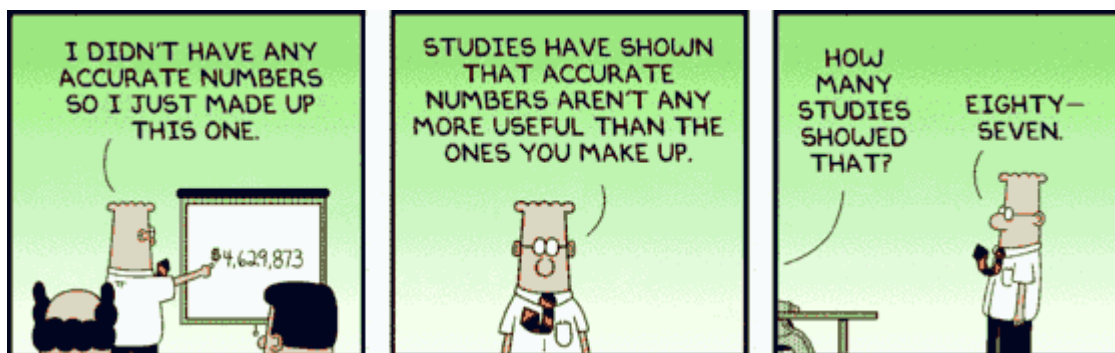
Allora il direttore chiede di farne una media – e Dilbert risponde:

«se vuoi li posso anche moltiplicare».

1 <http://htlws.it>

2 <http://gandalf.it/htlws/citazioni/>

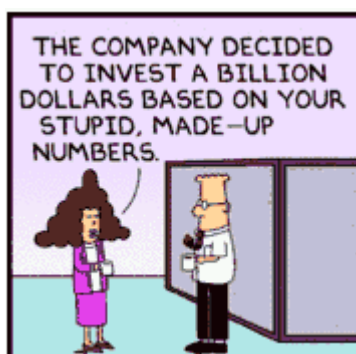
Questa è la seconda vignetta della serie, pubblicata l'8 maggio 2008.



copyright © Scott Adams 2008

Dilbert fa una presentazione usando un numero immaginario e osserva:
«*studi hanno dimostrato che i numeri precisi non sono più utili di quelli inventati*»
Alla domanda «*quanti studi?*» risponde «*ottantasette*»
(e anche quello, ovviamente, è un numero inventato).

In una terza vignetta, il 9 maggio 2008,
si scopre che in base ai dati “inventati” da Dilbert
l'impresa ha deciso di investire un miliardo di dollari.



copyright © Scott Adams 2008

Non è solo uno scherzo. Nella realtà pratica le cose non sono così ovvie, ma accade davvero che si usino metodi non meno assurdi. E non è raro il caso che “investimenti” insensati si basino su dati o proiezioni di scarsa o nulla attendibilità.

I sistemi di elaborazione hanno fatto grandi progressi. Con le tecnologie di oggi, quando sono usate bene, si possono ottenere risultati importanti non solo in tempi più brevi, ma anche con analisi più approfondite. Il problema è che le elaborazioni si fanno spesso con metodi standardizzati e inadatti a un caso specifico, con correlazioni fra dati diversi che ne confondono il significato, con un'infinità di elucubrazioni che mascherano l'inesattezza dei dati originari oppure, anche quando sono attendibili, ne distorcono il significato e l'interpretazione.

I rimedi sono concettualmente semplici. Il primo è la verifica *a priori* dell'attendibilità dei dati e del loro preciso significato. Il secondo è la scelta di un metodo di analisi che sia attinente al processo e non inquinato da inutili e pericolose complicazioni. Il terzo è la trasparenza: cioè la possibilità *a posteriori* di controllare tutte le fasi di analisi e di elaborazione, per poter correggere eventuali errori o elementi di dubbio e per poter verificare il risultato di ipotesi diverse in qualsiasi punto della procedura.

È lapalissiano? Sì. Ma l'esperienza pratica dimostra continuamente la mancanza di queste chiarezze, con conseguenze che possono variare da fastidiose incomprensioni a deduzioni o decisioni pericolosamente sbagliate.

Non si tratta solo di statistiche. Con ogni genere di dati (compresi quelli finanziari, organizzativi, demografici e di mercato) ci possono essere pericolosi errori (oltre che intenzionali inganni) nel significato dei numeri e nei metodi di elaborazione. Con l'aggiunta delle insidiose "false certezze" derivanti dalle tecnologie di *data processing* che non solo non "rendono vero" ciò che era dubbio all'origine, ma spesso confondono o deformano anche quei numeri che, prima di essere "frullati" in una confusa mescolanza, avevano un ragionevole livello di attendibilità.

Questa citazione di Scott Adams è interessante perché il tema dei suoi (seri) studi e delle sue (spesso divertenti) ironie non è la statistica o l'elaborazione numerica, ma la congerie di problemi che affliggono la gestione e il funzionamento delle imprese. Non è un caso che fra i fattori di dissesto ci segnali, ancora una volta, il cattivo uso dei dati e delle analisi che ne derivano.

Il rimedio? Oltre al rigore e alla trasparenza dei metodi, la risorsa più importante è sempre il buon senso. Se un dato mal si concilia con l'esperienza e con l'osservazione dei fatti, è sempre necessaria un'energica coltivazione del dubbio. Per capire se sono i dati a darci una percezione deformata – o se davvero ci siano fatti e circostanze inaspettate o difficilmente spiegabili, tali da richiedere una seria revisione di tutto il quadro di conoscenze e perciò delle strategie decisionali.