
Roma, Forum P.A., 18 maggio 2012



**La statistica oggi:
a chi serve e perché**

Alberto Zuliani

Sommario

- **Statistiche e statistica**
 - **Campi di applicazione**
 - **Perché la statistica**
 - **Statistica e sondaggi**
 - **Per capire**
 - **Per valutare**
 - **Per scegliere/decidere**
 - **Per prevedere e operare**
 - **Per soddisfare curiosità**
 - **Per non farsi imbrogliare**
 - **A chi serve, dunque**
 - **Sei regole da seguire di fronte alle statistiche**
 - **Quattro domande da farsi di fronte a una tabella**
 - **Statistiche comparative internazionali: i pericoli**
 - **Importanza della statistica**
 - **Conclusioni**
 - **Statistica strumento di cittadinanza**
 - **Bibliografia**
-

Statistiche e statistica

- In generale **abbiamo a che fare con le statistiche**: il pil, il tasso di disoccupazione, l'indice dei prezzi e il tasso di inflazione, l'indice della produzione industriale....
- **Dietro quei dati c'è la statistica**: metodi di campionamento, costruzione dei questionari, tecniche di loro somministrazione, tecniche di analisi dei dati semplici e complesse, valutazioni probabilistiche.
- **Dietro la statistica ci sono i contenuti**, la sostanza dei problemi e l'esigenza di conoscere.
- **Di fronte a quei dati c'è ancora la statistica**: per capire, valutare, prevedere, decidere e per non farsi imbrogliare.

UTILIZZATA IN MODO PROFESSIONALE O STRUMENTALE, LA STATISTICA DEVE ESSERE CONOSCIUTA

(Alcuni) campi di applicazione

Studi di popolazione, demografia

Modelli econometrici

Modelli di finanza

Epidemiologia

Saggi biologici; **sperimentazione di farmaci**

Sperimentazione agraria e zootecnica

Controlli di qualità e di accettazione industriali

Marketing

Auditing

Sociologia (indagini su atteggiamenti, motivazioni eccetera)

Psicomelia (test psico-attitudinali)

Pedagogia (valutazione dell'apprendimento)

Legal statistics (analisi del DNA da tracce lasciate sui luoghi di delitti)

Riconoscimento di autori di scritti apocrifi (*text mining*)

Telerilevamento

Sondaggi

Perché la statistica (1)

- **Pensiamo statisticamente** => fin da bambini classifichiamo, sistemiamo, decidiamo.
 - **Gestione dell'incertezza** => i fenomeni fisici, naturalistici, sociali, economici, ... sono aleatori. Occorre orientarsi e prendere decisioni.
 - Transizione delle discipline sostanziali da qualitative a quantitative, **da deduttive a induttive** => salto di qualità, avanzamenti.
-

Perché la statistica (2)

- **Per capire**
 - **Per valutare**
 - **Per scegliere/decidere**
 - **Per prevedere e operare**
 - **Per soddisfare curiosità**
 - **Per non farsi imbrogliare**
-

Statistiche e sondaggi

- **Distinguiamo** fra statistiche e sondaggi.
 - Le statistiche rilevano principalmente **fatti**, i sondaggi rilevano principalmente **opinioni**.
 - **Buon senso e senso comune**.
 - Il **buon senso** si nutre di **fatti**, il **senso comune** di **miti**, stereotipi e luoghi comuni. Verificare la consistenza dei luoghi comuni con i dati è un atteggiamento intelligente.
-

Sondaggi politici

- Sondaggi sulle **intenzioni di voto**. Due sondaggi dello scorso novembre: (a) **Demopolis** per l'Espresso, Famiglia cristiana e altri quotidiani; (b) **Emg** per "la 7"
 - (a) contattate 2214 persone per avere **1000 risposte**; non fornisce i dati su quanti non voteranno o voteranno scheda bianca
 - (b) contattare 3933 persone per avere **1000 risposte**; 16% indecisi; 27% non andranno a votare; 2% voteranno scheda bianca; quindi la **vera espressione di voto per i partiti** riguarda poco più di **500 persone**
 - La **variazione** di intenzione di voto di **0,1%** (riferita ad esempio per la lista Bonino-Pannella) rappresenterebbe **mezza persona!**
 - **Gran parte delle variazioni** testimoniate rientrano nel **possibile errore statistico**
 - Le meta analisi (più sondaggi insieme) danno conto di alcuni movimenti nel corpo elettorale: da destra verso sinistra
 - Gli **exit polls**: un **esercizio completamente inutile**, con scopi esclusivamente mediatici
-

Per capire (e sfatare qualche luogo comune)

Nel nostro paese ci sono troppi immigrati

- **8%** dei residenti in Italia sono **stranieri** (come in Germania, Inghilterra e Francia)
- In Italia **100.000** bambini ogni anno **hanno almeno un genitore straniero (su 570.000 nati in un anno)**
- In Italia, senza immigrazioni, ogni 4 persone che compiono 65 anni, soltanto 3 ne compiranno 20 e **solo uno di loro** sarebbe disposto a fare l'**operaio**
- In Veneto, fra il 2004 e il 2008, gli occupati dipendenti esordienti nel settore privato con meno di 30 anni sono stati **65.000** ogni anno: **43.000** erano **italiani** e **22.000 stranieri**
- Se 25 anni prima, fra il 1979 e il 1983, i genitori veneti avessero avuto un numero di figli sufficienti a rimpiazzarli (2,1 per donna, anziché 1,4 come è avvenuto) sarebbero nati **65.000** bambini ogni anno; invece ne sono nati **43.000**. **I 21.000 mancanti sono stati sostituiti 25 anni dopo dagli stranieri entrati nel mercato del lavoro**
- L'**immigrazione** è una **risorsa**: se gli immigrati hanno affrontato tanti disagi è perché hanno speranza e voglia di crescere.

Per valutare (le politiche pubbliche): gli strumenti

- **Politiche educative basate sull'evidenza**
 - **Paese: USA**
 - **Spunto: Rapporto dell'Accademia delle scienze** sullo stato della valutazione in campo educativo (1999)
 - **Amministrazione Bush (2002), Education Science Reform Act** istituisce l'**Institute of Educational Sciences (IES)**, agenzia indipendente con il compito di promuovere l'uso rigoroso della valutazione in campo educativo e di promuovere nelle scuole l'uso delle pratiche risultate efficaci
 - **L'IES ha una dotazione iniziale di 500 milioni di \$ annui**, cresciuta nel 2011 a 740 milioni di \$ (da noi l'Invalsi nel 2011 disponeva di 15 milioni di euro)
 - **L'IES punta su metodi quantitativi rigorosi**
 - **Trova resistenze da parte delle Schools of Education**
 - **Ha finanziamenti aggiuntivi da parte della Bill e Melinda Gates Foundation**
-

Valutare le politiche pubbliche: Il programma *Reading first* negli USA

Politica: *No child left behind*

Intervento: *Reading first*: colmare entro la terza elementare i ritardi nella capacità di lettura dei bambini, soprattutto immigrati

Durata: 2002-2008

Finanziamento: **1 miliardo di \$ distribuiti alle scuole** che li hanno usati per attività a sostegno di insegnanti e alunni

Valutazione: campionate 248 scuole confrontate (**controfattualmente**) con un campione di controllo

Risultati: **Punteggio medio in comprensione della lettura** (su 1000) nelle scuole beneficiarie (**B**) e in quelle non beneficiarie (**NB**):

1[^] el: **B=544**; **NB=539**; 2[^] el: **B=584**; **NB=583**; 3[^] el: **B=609**; **NB=609**

Conclusione: **se non si fosse fatta la valutazione si sarebbe proseguito l'intervento sprecando soldi**; invece è stato sospeso. Apprendere il fallimento impone di cercare nuove soluzioni; ignorarlo porta soltanto a spendere male i soldi.

Valutare le politiche pubbliche: *il Perry Pre-school Project*

- **123 bambini di famiglie afro-americane** che vivevano nelle vicinanze della *Perry Elementary School* in Ypsilanti, Michigan nel 1960
 - **Due gruppi:** inserito e non inserito nel programma, **valutati da 3 a 11 anni e successivamente alle età 14-15, 19 e 27 anni**
 - **Intervento:** frequenza mattutina per 2 ore e mezza per un anno; visite degli insegnanti alle famiglie
 - **Risultati a 27 anni:** più alto livello di **reddito** (29% vs 7% redditi superiori a 2.000 \$ mensili, e per gli uomini 42% vs 6%); più alta frequenza di **proprietà dell'abitazione** 35% vs 13% e, per gli uomini 52% vs 21%) e di una **seconda automobile** (30% vs 13%); più alta probabilità di **completare gli studi** al 12° grado (71% vs 54%); minore probabilità di ricevere servizi di **sostegno sociale** nei precedenti 10 anni (59% vs 80%); minore probabilità di **essere arrestati** (7% vs 35% con 5 o più arresti e, per gli uomini 12% vs 49%) , in particolare per **produzione o spaccio di droga** (7% vs 25%)
 - **Conclusioni:** il programma ha sviluppato nei bambini la predisposizione a interagire positivamente con gli altri e con i propri compiti. **E' stato stimato un ritorno per la collettività di 7,16 \$ per ogni dollaro investito.**
-

Per scegliere (decidere): indicatori sanitari

Indicatori	valore minimo	valore massimo	Italia
Degenza media pre-operatoria (gg)	1,6 (Piemonte)	2,7 (Lazio)	1,9
Fratture di femore operate entro 2 gg (%)	15,8 (Campania)	59,4 (Marche)	33,6
Ricette per abitante	97,2 (Toscana)	154,5 (Puglia)	123,4
Parti cesarei (su 100)	23,1 (Friuli Venezia G.)	61,9 (Campania)	38,4

Parti cesarei (1)

Parti cesarei sul totale dei parti per regione. Anno 2008

**Parti cesarei su
100 in totale**

Regioni

< 25

Friuli Venezia Giulia, P.A. di Trento

25 – 30

Toscana, P.A. di Bolzano, Veneto, Lombardia

30 – 40

**Emilia Romagna, Umbria, Piemonte,
V. d'Aosta, Marche, Liguria, Sardegna**

40 – 50

Lazio, Abruzzo, Calabria, Basilicata, Molise, Puglia

> 50

Sicilia, Campania

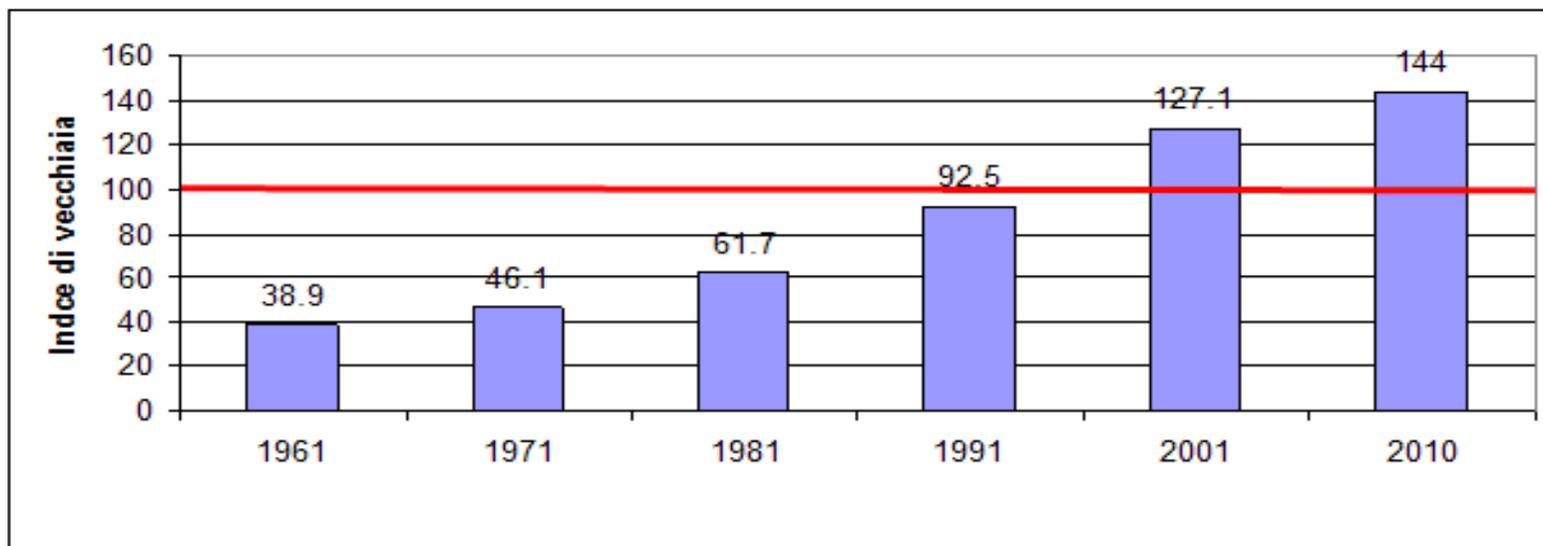
Parti cesarei (2)

Ricorso al parto cesareo secondo la percentuale di parti avvenuti in strutture convenzionate per regione.

Anno 2006 (*valori percentuali*)

PERCENTUALE DI PARTI AVVENUTI IN STRUTTURE CONVENZIONATE	Regioni	Frequenza di ricorso al parto cesareo (%)
< 10 %	Abruzzo, Basilicata, Emilia Romagna, Friuli Venezia Giulia, Liguria, Lombardia, Marche, Molise, Provincia autonoma di Bolzano, Provincia autonoma di Trento, Piemonte, Toscana, Umbria, Veneto	33,0
10 % - 25 %	Lazio, Puglia, Sicilia, Sardegna	45,1
> 25 %	Campania	61,9

Per prevedere e operare: invecchiamento della popolazione



Per prevedere e operare: disabilità e *long term care*

- Le **persone** che in Italia hanno problemi per almeno una delle attività quotidiane, quindi **disabili**, sono **3.000.000** (5% della popolazione); di queste 2.200.000 sono anziani e fra loro 1.250.000 hanno 80 anni e più. I gravi sono 950.000; **i gravissimi 810.000**.
 - Si stima che **nel 2035**, le persone disabili (essenzialmente per effetto dell'invecchiamento della popolazione) saranno **4.600.000**, delle quali 1.700.000 gravi e **1.150.000 gravissime**.
 - Si porranno forti problemi di assistenza (***long term care***)
-

**Per prevedere e operare:
diffusione delle droghe fra gli alunni
della scuola secondaria superiore nel 2011**

- **Dipartimento politiche antidroga della Presidenza del Consiglio dei ministri e Università di Roma Tor Vergata**
 - **Campione di 32.186 alunni fra 14 e 19 anni**
 - **Uso nel corso della vita: M 31%, F 23%**
 - **Uso nell'ultimo anno: M 25%, F 18%**
 - **Uso nell'ultimo mese: M 17%, F 12%**
-

Diffusione delle droghe nella scuola secondaria superiore

Uso di sostanze nel corso dell'intera via fra gli **studenti della scuola secondaria superiore**. Anno 2011 (distribuzioni percentuali)

Descrizione	Non ha		Ha usato		
	mai usato	almeno una sostanza	tranquillanti	cannabis	cocaina
Maschi	69,0	31,0	4,0	17,1	3,5
Femmine	77,0	23,0	6,0	18,4	1,6
<i>Totale</i>	<i>73,1</i>	<i>26,9</i>	<i>5,0</i>	<i>26,5</i>	<i>2,5</i>
meno di 16 anni	87,9	12,1	4,0	6,9	0,3
16 anni	79,3	20,7	4,4	15,5	0,6
17 anni	72,5	27,5	5,3	22,0	1,0
18 anni	65,5	34,5	5,3	29,6	1,4
19 anni	61,2	<i>38,8</i>	<i>6,2</i>	<i>33,4</i>	<i>4,6</i>

Predittori meno evidenti

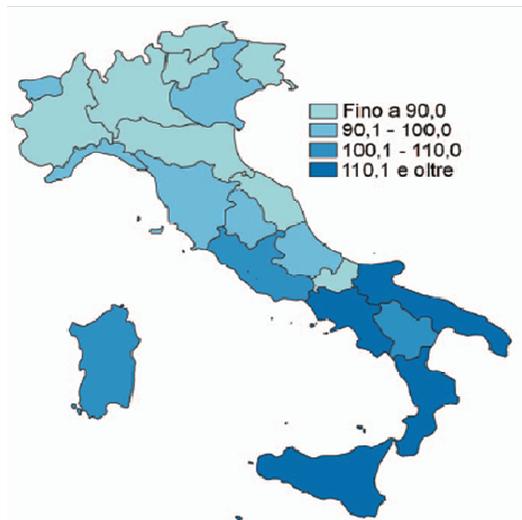
- Molte caratteristiche degli studenti hanno una stretta **prossimità con l'uso di sostanze psicotrope** (assunzione di alcol, fumo attivo, amici che usano droga **(10,70)**).
 - Altri fattori di rischio meno evidenti che **possono aiutare gli interventi a casa e a scuola** sono in ordine di importanza:
 - uscite frequenti da casa la sera **(2,28)**
 - numero elevato di giorni persi a scuola **(1,84)**
 - spese senza controllo **(1,77)**
 - elevata permissività in famiglia **(1,64)**
 - rapporto insoddisfacente con la madre **(1,30)**
 - rapporto insoddisfacente con il padre **(1,25)**
-

Per soddisfare curiosità

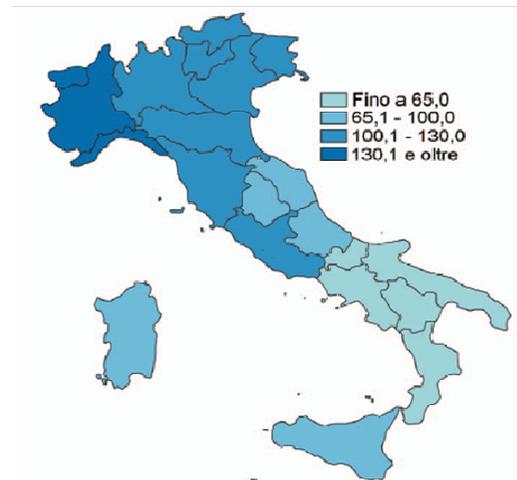
Noi Italia, 100 statistiche per capire il paese in cui viviamo

Tassi di **nuzialità** (matrimoni per 1000 abitanti) e tassi di **divorzialità** (divorzi per 10.000 abitanti) **per regione**.
Anno 2007 (*numeri indici: tasso per l'Italia=100*)

Tassi di nuzialità



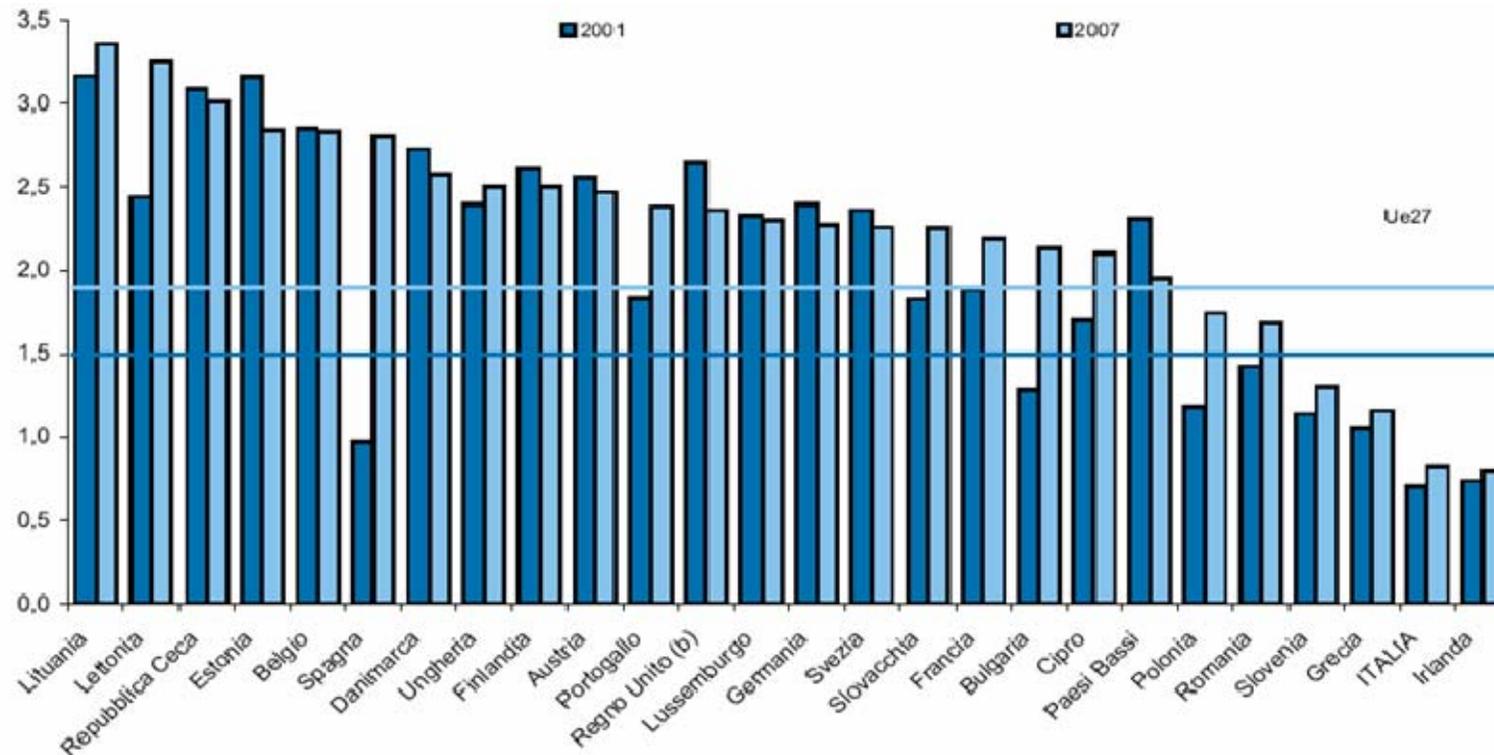
Tassi di divorzialità



Per soddisfare curiosità

Noi Italia, 100 statistiche per capire il paese in cui viviamo

Tassi di **divorzialità** in Europa (per 1000 abitanti)



Per non farsi imbrogliare

Il gusto di verificare: le norme cancellate

Il ministro per la semplificazione normativa ha affermato

- **375.000 norme cancellate**
 - Un nastro di **42 milioni di chilometri** (distanza terra-luna: 385.000 chilometri)
 - Una pira di 38 scatoloni
 - Due tir pieni di gazzette ufficiali
 - **21 miliardi di euro di risparmio !!!**
- invece**
- **32.500 norme certamente cancellate**
 - **Sostanzialmente quelle che già avevano perso di efficacia**
 - **Infatti le imprese si lamentano ancora dei “lacci e laccioli”**
 - **Una stima: il nastro sarebbe al massimo di 19.000 chilometri**
-

Per non farsi imbrogliare

Il lotto? Una tassa sugli imbecilli (e sui poveri)

Probabilità di vincita e vincite nel gioco del lotto

NUMERI GIOCATI	Probabilità di vincita	Vincita equa in euro per 1 euro giocato	Vincita lorda in euro	Vincita netta (a) in % di quella equa
1(ambata)	1/18	18	11,23	58,65
2 (ambo)	1/400,5	400,5	250	58,67
3 (terno)	1/11.748	11.748	4.500	36,00
4 (quaterna)	1/511.038	511.038	120.000	22,07
5 (cinquina)	1/43.949.268	43.949.268	6.000.000	12,83

(a) Al netto della tassazione del 6%.

A chi serve dunque?

Alle persone, alle famiglie, alle imprese, ai sindacati, alle amministrazioni pubbliche, ai governi a tutti i livelli, ai giornalisti (ma attenti, il giornalismo *evidence based* non si riduce a citare qualche dato statistico, ma deve fornire prove con canoni rigorosi), ai ricercatori, ai meteorologi, alle forze dell'ordine, agli agricoltori, agli allevatori, ai chimici, ai biologi,

.....

Sei regole da seguire di fronte alle statistiche

- **Riferirsi alle tabelle** piuttosto che ai grafici
 - **Leggere sempre le note**
 - **Avere in mente alcuni aggregati di riferimento**
 - **Porsi domande**, non farsi certezze
 - **Un problema complesso non può essere liquidato con un solo dato**
 - **Diffidare sempre dei dati che sostengono la promozione di un prodotto**
-

Quattro domande da farsi di fronte a una tabella

- Qual è la **fonte**?
 - **Come è stata prodotta**? Ampiezza e copertura del campione, modalità di raccolta - telefonica, postale, web, televisiva, personale - (approcci differenti => risultati differenti), questionario utilizzato, trattamento delle mancate risposte
 - Sono stati valutati **tutti gli aspetti rilevanti** del fenomeno?
 - Ci sono **altri dati disponibili**?
-

Approcci differenti, risultati differenti

- Una **identica domanda**: Nel corso della tua vita, sei mai stato attratto sessualmente, almeno una volta, da una persona del tuo stesso sesso?
 - **Due approcci**: a) indagine telefonica; b) questionario autocompilato
 - a) 11% rifiutano di rispondere; 4% ammettono l'attrazione, **maggiore fra gli uomini (6%)** che fra le donne (3%)
 - b) 1% rifiutano di rispondere; 6% ammettono l'attrazione, **maggiore fra le donne (8%)** che fra gli uomini (6%)
-

Statistiche comparative internazionali: i pericoli

- Per considerare molti paesi **si passa sopra all'affidabilità delle informazioni** (gli apparati statistici sono infatti molto diversi paese da paese)
 - Vengono introdotte **molte stime qualitative**, spesso arbitrarie, ottenute da c.d. esperti
 - Le **graduatorie** che emergono sono **poco convincenti, però colpiscono**
 - S&P ha messo recentemente **l'Italia insieme a Perù, Bolivia e Azerbaigian**
-

L'indagine Pisa

sui risultati scolastici dell'Ocse

Percentuali di studenti con elevato livello di performance - *top performers* (a) - in scienze, comprensione della lettura e matematica per genere, nei paesi dell'Ocse. Anno 2006 (per 100 studenti sottoposti ai test)

PAESI	Scienze		Comprensione della lettura		Matematica	
	Donne	Uomini	Donne	Uomini	Donne	Uomini
Corea	9,5	11,1	27,3	16,3	24,2	29,9
Finlandia	20,2	21,6	23,7	9,6	21,1	27,8
Francia	6,5	9,6	8,9	5,5	10,7	14,5
Germania	9,8	13,7	12,9	7,0	12,0	18,7
Italia	3,8	5,4	6,7	3,7	4,1	8,4
Nuova Zelanda	16,9	18,4	19,1	12,4	16,1	21,9
Spagna	4,1	5,6	2,4	1,1	5,4	9,0
Stati Uniti	8,2	10,0	ND	ND	6,6	8,6

(a) Si definiscono *top performers* gli studenti che arrivano almeno al livello 5 delle scale Pisa per ciascuno dei tre ambiti.

Il Global Gender Gap index del World Economic Forum

Tabella 2 - Valori dell'indice GGG per sette paesi, ottenuti utilizzando differenti sistemi di pesi

Pesi	Svezia	Filipp.	Spagna	Uganda	Francia	Tanzan.	Italia
Proporzionali a s							
- Valore dell'indice	0,851	0,769	0,882	0,657	0,742	0,646	0,680
- Numero indice (Svezia = 1)	1,000	0,904	0,967	0,772	0,872	0,759	0,800
Proporzionali a s/m							
- Valore dell'indice	0,843	0,771	0,805	0,706	0,730	0,634	0,675
- Numero indice (Svezia = 1)	1,000	0,914	0,955	0,837	0,866	0,752	0,801
Originali (proporzion. a l/s)							
- Valore dell'indice	0,802	0,765	0,756	0,717	0,703	0,683	0,675
- Numero indice (Svezia = 1)	1,000	0,954	0,943	0,893	0,875	0,851	0,843
- Posto nella graduatoria	4	9	11	33	46	66	74

Il Doing Business Indicators della Banca mondiale

Posizionamento di alcuni paesi dell'Unione europea rispetto ai "Doing business indicators" della Banca mondiale (su 183 economie). Graduatoria per il 2010 (dati riferiti a giugno 2009)

PAESI	Posizione complessiva	Avviare un'impresa	Assumere e licenziare	Accedere al credito	Pagare le tasse
Regno Unito	5	16	35	2	16
Germania	25	84	158	15	71
Francia	31	22	155	43	59
Spagna	62	146	157	43	78
Italia	78	75	99	87	136
Grecia	109	140	147	87	76

Fonte: Banca mondiale (2010), *Doing Business database*

Importanza della statistica

- Wells H.G. (1903): bisogna **promuovere un'educazione universale alla statistica**, perché enumerare e ragionare in termini di massimo e minimo divenga **comune come leggere e scrivere**.
- Benjamin A. (*mathemagician*) (2010): nella vita di tutti i giorni la statistica è più importante dell'analisi matematica; **ogni studente che esce dal liceo dovrebbe conoscere la statistica e il calcolo delle probabilità**.
- McKinsey (2011), *Big data: the next frontier for innovation, competition and productivity*. La stima del volume di **dati** prodotti nel 2009 nel mondo è di 800 *exabyte* (1 *exabyte* = 10^{18} byte); essi devono essere **elaborati per produrre informazione**. Il compito sarà svolto, in generale, **dagli statistici**.



Conclusioni

La **cultura della valutazione** è da noi **poco diffusa** a tutti i livelli: centrale, regionale, locale, delle unità operative

La cultura della parola e quella dell'**annuncio** prevalgono sulla cultura del dato e del **risultato**

Nelle occasioni di rendiconto **non** si riferisce su quanto si è **realizzato ma** spesso si illustrano i **programmi futuri**

I **politici sono refrattari** rispetto ai dati perché ritengono che possano limitarli anziché aiutarli nelle decisioni da prendere

I **dirigenti amministrativi e tecnici temono** che i dati consentano di valutare il loro operato

I cittadini sono **poco critici sui dati** che vengono proposti dai quotidiani e soprattutto dalle televisioni e, invece:

DEVONO RENDERSI PROTAGONISTI

La statistica strumento di cittadinanza

- **La statistica è “sovversiva”**; la verità è rivoluzionaria
 - **La statistica abitua** alla variabilità e quindi alla tolleranza
 - **La statistica è al servizio** principalmente **dei cittadini** e dell’opposizione (asimmetria informativa a favore dei governi; statistica pubblica finanziata dalla fiscalità generale); è uno **strumento fondamentale di cittadinanza**
-

Bibliografia essenziale

- Dalla Zuanna G. e Weber G., *Cose da non credere. Il senso comune alla prova dei numeri*, Bari, Laterza, 2011.
 - Istat, *Rapporto annuale. La situazione del Paese nel 2010*, Roma, 2011.
 - Istat, *Noi Italia, 100 statistiche per capire il Paese in cui viviamo*, Roma, 2011.
 - Rey G.M., Rossi C. e Zuliani A., *Il mercato delle droghe. Dimensione, protagonisti, politiche*, Venezia, Marsilio, 2011.
 - Trivellato U. e Martini A., *Sono soldi ben spesi? Perché e come valutare l'efficacia delle politiche pubbliche*, Venezia, Marsilio, 2011.
 - **Zuliani A., *Statistiche come e perché. A cosa servono, come si usano*, Roma, Donzelli, 2010.**
-