

Visualizzazione dinamica dei dati statistici: esperienze e applicazioni in Istat

Stefano De Francisci
Michele Ferrara
Antonio Pitrone

Istat – Direzione centrale per le esigenze
degli utilizzatori, integrazione e territorio

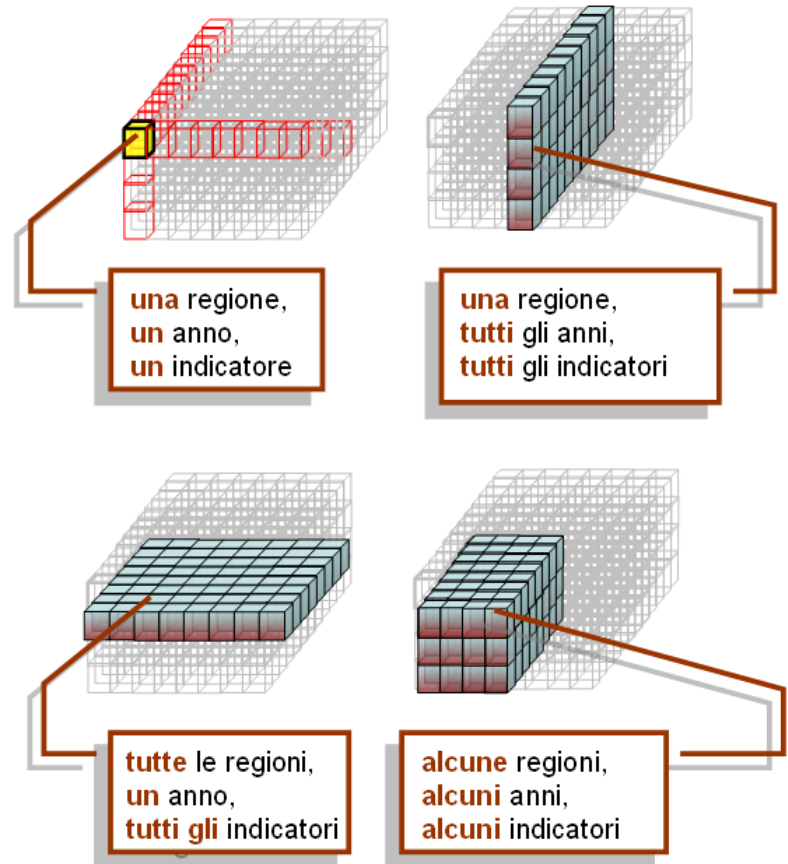
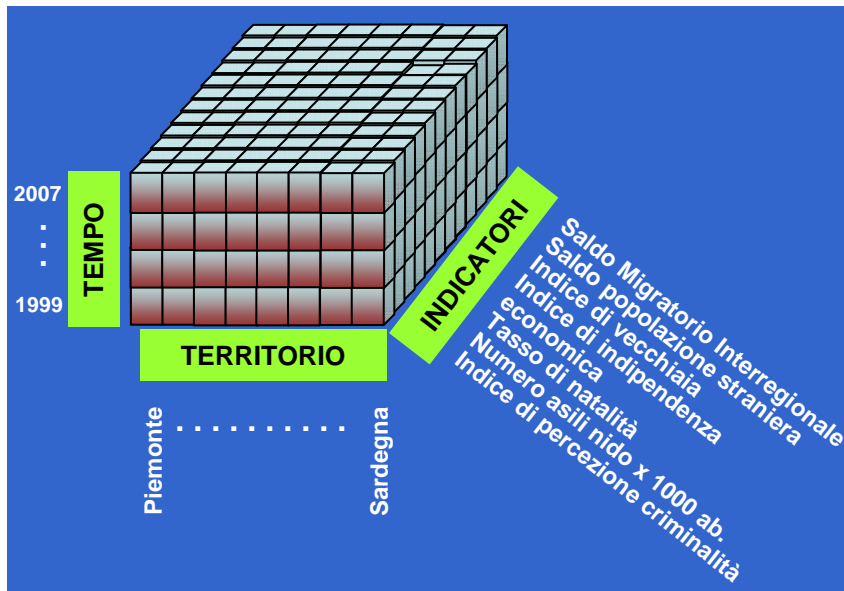
Numeri, dati, storie

“I numeri non sono intrinsecamente noiosi, né intrinsecamente interessanti” (Stephen Few)

Se questo è vero, possiamo dire che i numeri sono solo numeri e che – se hanno una storia da raccontare – dobbiamo trovare il modo di saperlo fare?

Pensiamo a un archivio di dati storici, organizzati per regione, contenente dati relativi a indicatori demografici e sociali.

La struttura che avremmo a disposizione può essere descritta tramite un “cubo di dati”



Il mantra dell'analista di dati statistici

**Come analizzare le serie
singolarmente e in complesso?
Come metterle in relazione spaziale
tra loro e nel tempo?
Come interpretare i fenomeni?
Come visualizzare in modo efficace i
dati?**

Il mantra dell'analista di dati statistici:

“Overview first, zoom and filter, then details-on-demand” (*Ben Shneiderman*)

Le tecniche classiche di analisi e visualizzazione non sembrano più rispondere ai requisiti di complessità che la vasta offerta di informazione statistica possiede.

Le nuove tecniche di visualizzazione dinamica di dati statistici, assieme all'uso di ambienti collaborativi e all'adozione di nuovi approcci al disegno dell'informazione e all'organizzazione della conoscenza aprono, per i produttori di informazione statistica e per gli utilizzatori, prospettive nuove nel campo dell'integrazione e della produzione di conoscenza.

Alla ricerca del dato

Avendo a disposizione un foglio Excel, i dati potrebbero essere organizzati più o meno così:

Ma potremmo usare anche un'altra rappresentazione:

Indicatori demografici e sociali per regione. Anno 1999

Regione	Saldo Migratorio Internazionale	Saldo popolazione straniera	Tasso di disoccupazione giovanile	Indice di vecchiaia	Indice di indipendenza economica	Tasso di natalità	Tasso di occupazione femminile	Indice di sviluppo umano	Indice di diffusione TIC (2005)	Indice di diffusione TIC (2006)
Piemonte	3042	9262	26,5	172,0	-4,0	9,8	48,0	8,1	8,6	94,3
Valle d'Aosta	182	382	19,9	20,4	348,2	2,2	83,2	9,1	7,9	55,3
Lombardia	14530	13945	14,4	124,0	1,0	10,0	48,0	10,0	10,0	100,0

Saldo Migratorio Interregionale per regione. Anni 1999-2007

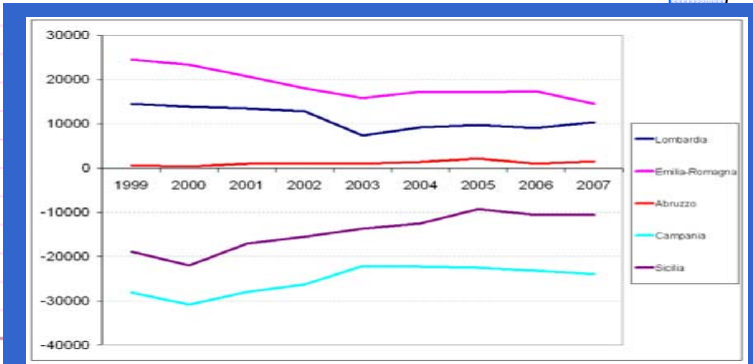
REGIONE	1999	2000	2001	2002	2003	2004	2005	2006	2007
Piemonte	3042	2212	1790	430	-178	-278	-1024	145	423
Valle d'Aosta	182	309	110	152	370	302	292	11	-287
Lombardia	14520	13945	13541	12819	7454	9217	9684	9044	10442
Trentino-Alto Adige	1862	1811	919	1184	1640	2058	1782	1759	1829
Veneto	8710	8969	6717	6990	7198	5376	4187	4235	5193
Friuli-Venezia Giulia	3888	4550	3843	3507	3244	2890	1998	2080	3095
Liguria	538	1007	961	1297	1013	1945	2218	3722	795
Emilia-Romagna	24534	23373	20819	18014	18014	18014	18014	18014	18014
Toscana	10675	11059	9656	8350	8394	7460	6991	7388	6026
Umbria	2360	3138	2215	2300	2300	2300	2300	2300	2300
Marche	4641	5146	5596	5197	5197	5197	5197	5197	5197
Lazio	2013	7071	3464	2350	2350	2350	2350	2350	2350
Abruzzo	600	371	1077	980	980	980	980	980	980
Molise	-318	-214	36	184	184	184	184	184	184
Campania	-28053	-30801	-28024	-26362	-26362	-26362	-26362	-26362	-26362
Puglia	-13518	-14667	-11659	-11453	-11453	-11453	-11453	-11453	-11453
Basilicata	-2207	-2044	-1765	-1922	-1922	-1922	-1922	-1922	-1922
Calabria	-11497	-9624	-9036	-7820	-7820	-7820	-7820	-7820	-7820
Sicilia	-18902	-22065	-17104	-15439	-15439	-15439	-15439	-15439	-15439
Sardegna	-3080	-3546	-3156	-758	-758	-758	-758	-758	-758

Indicatori demografici e sociali per regione. Anno 1999

Saldo Migratorio Interregionale per regione. Anni 1999-2007

REGIONE	1999	2000	2001	2002	2003	2004	2005	2006	2007
Piemonte	3042	2212	1790	430	-178	-278	-1024	145	423
Valle d'Aosta	182	309	110	152	370	302	292	11	-287
Lombardia	14530	13945	13541	12819	7454	9217	9684	9044	10442
Trentino-Alto Adige	1862	1811	919	1184	1640	2058	1782	1759	1829
Veneto	8710	8969	6717	6990	7198	5376	4187	4235	5193
Friuli-Venezia Giulia	3888	4550	3843	3507	3244	2890	1998	2080	3095
Liguria	538	1007	961	1297	1013	1945	2218	3722	795
Emilia-Romagna	24534	23373	20819	18014	18014	18014	18014	18014	18014
Toscana	10675	11059	9656	8350	8394	7460	6991	7388	6026
Umbria	2360	3138	2215	2300	2300	2300	2300	2300	2300
Marche	4641	5146	5596	5197	5197	5197	5197	5197	5197
Lazio	2013	7071	3464	2350	2350	2350	2350	2350	2350
Abruzzo	600	371	1077	980	980	980	980	980	980
Molise	-318	-214	36	184	184	184	184	184	184
Campania	-28053	-30801	-28024	-26362	-26362	-26362	-26362	-26362	-26362
Puglia	-13518	-14667	-11659	-11453	-11453	-11453	-11453	-11453	-11453
Basilicata	-2207	-2044	-1765	-1922	-1922	-1922	-1922	-1922	-1922
Calabria	-11497	-9624	-9036	-7820	-7820	-7820	-7820	-7820	-7820
Sicilia	-18902	-22065	-17104	-15439	-15439	-15439	-15439	-15439	-15439
Sardegna	-3080	-3546	-3156	-758	-758	-758	-758	-758	-758

Alla quale potremmo vedere associata una differente forma di visualizzazione grafica



E in qu
territori
(scatt

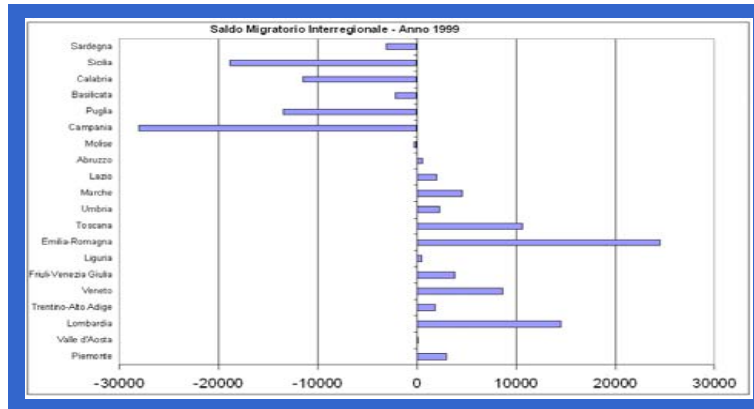
Saldo popolazione straniera

Saldo Migratorio Interregionale

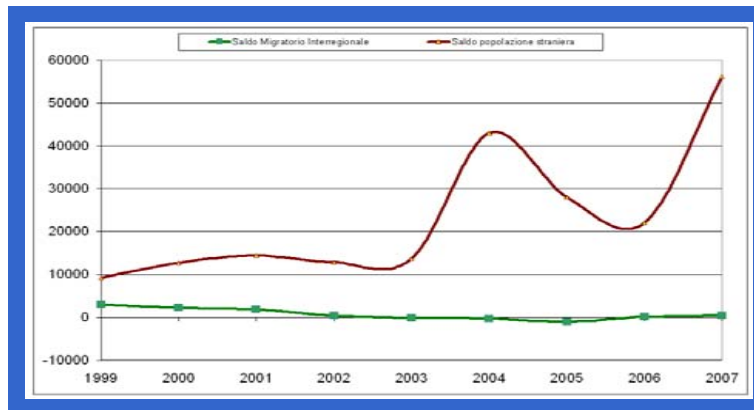
Saldo Migratorio Interregionale

Barre, onde, mappe, isole, tesori

Possiamo sistemare i dati in grafici a barre



Confrontare l'andamento nel tempo di due fenomeni per una stessa regione



Sapremmo costruire mappe tematiche



Ma sarà difficile scoprire i tesori informativi nascosti dentro i dati

E per quante volte?

dovremmo replicare le tabelle per anno

oppure per regione

Indicatore demografici e sociali per regione. Anno 1999

Regione	Saldo Migratorio Interregionale	Saldo popolazione straniera	Tasso di disoccupazione giovanile	Indice di vecchiaia	Indice di indipendenza economica	Tasso di natalità	Tasso di occupazione femminile	Numero asili nido x 1000 ab.	Indice di diffusione internet x 1000 ab.	Indice di percezione criminalità
Piemonte	3042	9262	20,5	172,0	-4,0	8,2	48,0	1	8,6	34,2
Valle d'Aosta	182	198	20,4	148,2	14,4	5,3	3	7,5	15,1	
Lombardia	14530	33925	15,4	135,5	15,6	49,3	2	8,4	36,4	
Trentino-Alto Adige	1862	2286	11,5	110,0	-1,4	51,4	2	8,8	13,0	
Veneto	8710	11280	13,5	117,9	-1,1	46,9	2	8,6	33,4	
Friuli-Venezia Giulia	3000	3000	14,8	118,0	-1,1	44,2	3	7,8	19,2	
Liguria	530	2000	25,1	139,0	-1,1	49,5	3	7,7	28,8	
Emilia-Romagna	24534	10000	11,2	136,0	-1,1	57,3	3	8,4	28,0	
Toscana	10475	10000	6,5	191,3	-2,6	7,8	3	9,8	28,1	
Umbria	2369	2312	18,2	182,3	3,8	8,1	50,7	2	9,5	28,6
Marche	4641	2489	18,2	185,8	3,3	8,4	48,4	2	8,9	22,3
Lazio	2013	13167	33,7	124,3	-7,9	9,9	38,1	1	11,2	41,0
Abruzzo	600	1966	34,8	139,4	6,7	8,5	45,1	0	5,7	13,7
Molise	-318	324	36,4	140,2	16,3	8,6	34,3	0	3,3	5,8
Campania	-28053	3716	48,3	72,3	19,0	12	26,8	1	4,9	53,3
Puglia	-13518	3175	40,2	89,4	16,2	10,5	26,5	0	5,1	32,8
Basilicata	-2207	338	49,0	110,3	13,6	9,4	32,5	1	2,9	14,8
Calabria	-11497	1847	48,5	94,9	29,5	9,7	24,5	0	3,6	19,6
Sicilia	-18902	4306	53,1	92,3	24,8	11	23,8	1	4,7	27,6
Sardegna	-9000	747	37,7	107,3	17,2	8,3	34,1	0	7,6	21,6

Indicatore demografici e sociali per regione e anno. Piemonte, anni 1999-2007

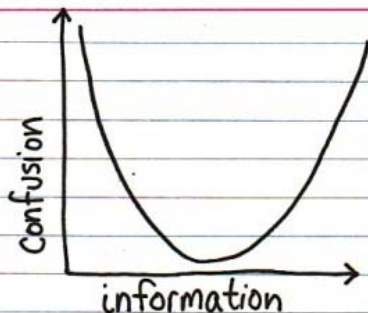
	1999	2000	2001	2002	2003	2004	2005	2006	2007
Saldo Migratorio Interregionale	3042	2212	1790	430	-178	-278	-1024	145	423
Saldo popolazione straniera	9262	12718	14431	12844	13704	42999	28076	22110	56244
Tasso di disoccupazione giovanile	20,5	18,7	12,8	12,6	13,4	15,8	16,9	15,5	14,3
Indice di vecchiaia	172,0	174,2	175,9	177,5	178,3	179,7	180,7	181,2	180,1
Indice di indipendenza economica	-4,0	-1,9	-2,9	-0,9	-1,5	-0,3	-0,8	0,1	-0,6
Tasso di natalità	8,2	8,5	6,8	8,4	8,5	8,6	8,6	8,7	8,8
Tasso di occupazione femminile	48,0	50,5	52,3	52,9	53,7	53,7	54,4	55,9	56,3
Numero asili nido x 1000 ab.	1	1,0	1,0	1	1	1	1	1	1
Grado di diffusione internet x 1000 ab.	8,6	16,9	25,1	30,2	27,8	-	-	-	-
Indice di percezione criminalità	34,2	34	33,7	32,7	27,7	-	-	30,9	33,1

Il diluvio informativo

“Ciò che succede qui è che si ha un'enorme massa di informazioni, anzi non soltanto una massa, un flusso di informazioni, ma un vero e proprio diluvio. Ho un amico, Royan Scott, che dice: stiamo vivendo il secondo diluvio. Il primo diluvio è stato di acqua, il secondo è il diluvio dell'informazione”

P. Lèvy, “Il diluvio informativo”

Jessica Hagy “Needles and haystacks and such.”

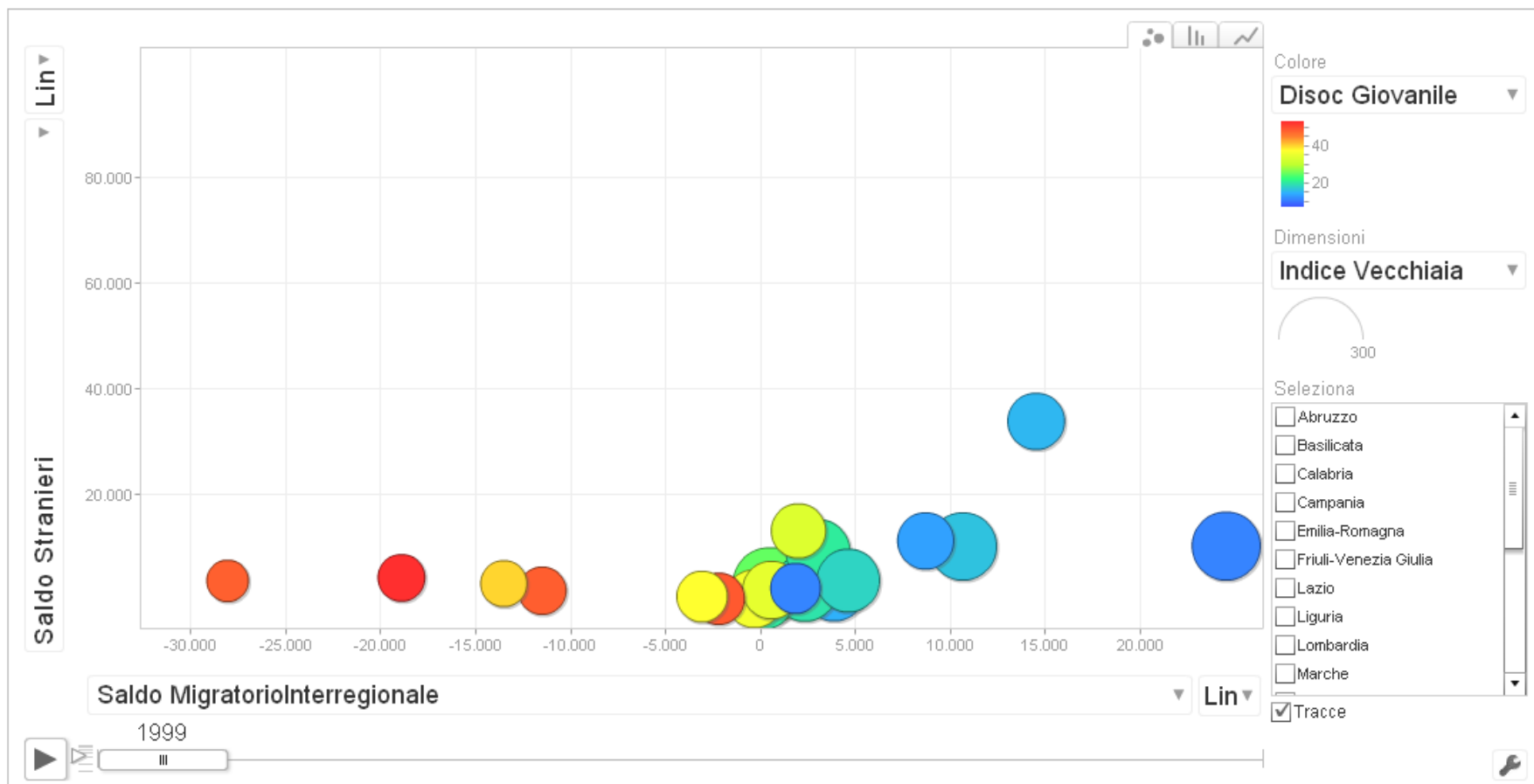


<http://thisisindexed.com/2009/10/needles-and-haystacks-and-such/>

Ma abbiamo un modo per vedere tutto insieme:

- *Molti* indicatori?
- Trend temporali?
- Tematizzazioni territoriali?
- Comportamenti per combinazioni *a piacere* di fenomeni?
- Confronti tra indicatori?
- Valori fuori soglia?

Trendalyzer



eXplorer

