

TRA SCALA GEOGRAFICA E RISOLUZIONE DEI DATI: METODOLOGIE
STATISTICHE PER L'ANALISI TERRITORIALE. IL CASO DI ROMA.

Salvucci Gianluigi*, Sanna Francesco Maria*

SOMMARIO

Nel presente lavoro si intendono applicare metodi esplorativi di analisi multivariata per giungere a tipologizzazioni del territorio (alle diverse scale) significative a livello geografico. Dopo aver individuato un'opportuna 'batteria' di indicatori e aver definito la scala di risoluzione dei dati (problema-chiave in questo tipo di analisi), si è analizzato inizialmente l'intero territorio nazionale, giungendo dapprima ad una visione di estrema sintesi (che peraltro consente di isolare tutte le più significative concentrazioni urbane); con procedimento a ritroso si è quindi segmentato il gruppo volta a volta più numeroso, al fine di cogliere meglio le caratteristiche differenziali delle varie aree (caratteristiche, è bene sottolinearlo, che, con sfumature diverse valgono per tutto il territorio nazionale) e si è pervenuti, in definitiva, a 7 gruppi significativi. Per capire se e quanto una scala di lettura a livello così elevato potesse essere utile e riproducibile a scale più ridotte, analoga procedura è stata applicata al caso di Roma, dove si sono riscontrate, da un lato, la solidità e validità del percorso di analisi proposto, dall'altro significativi scostamenti rispetto a modi più 'ordinari' di classificazione ed analisi del territorio romano.

* Dipartimento di Metodi e Modelli per l'Economia, il Territorio e la Finanza (MEMOTEF), Università degli Studi di Roma La Sapienza.

Lavoro svolto nell'ambito del gruppo di ricerca ISTAT – Università degli studi di Roma La Sapienza “Analisi geo-statistica del territorio sulla base di dati per sezioni di censimento”.

1 Introduzione

L'individuazione di aree contigue che riescano ad esplicitare fenomeni socio economici è uno dei punti cruciali della geografia quantitativa. Per riuscire ad ottenere risultati validi, vale a dire capaci di esprimere regioni significative, è necessario integrare l'analisi geografica con quella statistica. In questo lavoro il dato a livello di sezione censuaria viene assunto come la migliore rappresentazione possibile della realtà a nostra disposizione al massimo livello di dettaglio, e quindi comunque una importante risorsa nonostante la difformità delle superfici, la variabilità dei confini nel tempo delle sezioni, e via dicendo. Restano però aperti due problemi: da un lato la scala alla quale rappresentare un fenomeno con risoluzione di sezione, dall'altro quella della più opportuna scelta degli indicatori più idonei ai fini della rappresentazione e della successiva individuazione – a partire dalla batteria di indicatori prescelti – di gruppi con caratteristiche al tempo stesso di omogeneità e di rappresentatività.

L'obiettivo è quello di individuare eventuali strutture latenti e di farle emergere nella cartografia. Queste strutture devono essere confrontabili con altre realtà e per questo motivo si è proceduto preliminarmente all'analisi dell'intero territorio nazionale. Il risultato ottenuto permette di individuare un nuovo modello urbano, una diversa struttura dell'Italia che, scendendo nei dettagli, mostra caratteristiche strutturali costanti. Il ricercatore che individua il singolo indicatore e tenta di spiegarne il comportamento in funzioni di una o più variabili ha già in mente un risultato: il suo intento è quello di dimostrare la validità della tesi. Affrontare il territorio con la metodologia proposta permette di osservare e quindi comprendere i fenomeni osservati senza partire da ipotesi 'pre-concette'.

La necessità di rappresentazioni tematiche (coloroplate) capaci di localizzare sul territorio i diversi fenomeni socio-economici è sempre più fortemente avvertita. Il cartogramma è quindi una fotografia del territorio che ne vuole "immortalare" gli atteggiamenti, le tendenze, le problematiche. La scala e l'estensione geografica diventano un problema fondamentale: come in una fotografia, si tratta di scegliere tra il dare risalto al sorriso dei protagonisti della foto o inquadrare il panorama alle loro spalle.

La fotografia deve riuscire a produrre suggestioni, curiosità, non è solo una riproduzione fedele, che al contrario potrebbe suscitare indifferenza. È noto che nel rappresentare i fenomeni riducendoli secondo la scala geografica si compie un'operazione di semplificazione, che fa perdere taluni tratti della fisionomia originaria del fenomeno studiato. La carta è quindi una sintesi del territorio, allo stesso modo di un indicatore statistico, nel quale si sfumano i caratteri dei singoli individui per fornire un'immagine la più fedele possibile del collettivo.

Per ottenere un'immagine capace di rappresentare i fenomeni sul territorio si fa ricorso all'interazione delle metodologie statistiche e geografiche, nel tentativo di creare una procedura di regionalizzazione esplorativa.

2 I dati di base

Si è proceduto perciò con una analisi preliminare delle variabili presenti nel file ISTAT sui risultati del 14° Censimento generale della popolazione e delle abitazioni (2001) che fornisce le misurazioni a livello di singola sezione di censimento. Per quanto riguarda gli indicatori più direttamente connessi alle attività economiche, si è fatto ricorso ai dati (sempre a livello di sezione censuaria) derivanti dall'8° Censimento generale dell'industria (2001). Alcune di tali variabili risultano ovviamente fortemente correlate tra loro; inoltre il loro elevato numero rischia di condurre ad una interpretazione del risultato finale assai problematica e discutibile. Più variabili vengono considerate e maggiori sono le possibilità di avere gruppi simili. Il primo momento d'analisi è stato perciò quello di individuare degli indicatori giudicati in grado di differenziare il territorio nelle sue caratteristiche fondamentali¹. (cfr. Tabella 1).

Si è quindi proceduto a clusterizzare l'intero territorio nazionale, per individuare quelle singolarità che determinano sì le differenze locali ma che, ritrovandosi ripetute in più città (o aree) determinano gruppi con caratteristiche consolidate di più facile comprensione.

Tabella 1 - Variabili e indicatori utilizzati nel processo di clusterizzazione (dati Istat 2001)

Nome variabile	Descrizione	classi ateco
A2	totale abitazioni	
AD	Adulti popolazione da 15 a 64 anni	
ANZ	Popolazione over 64	
DIP	Lavoratori dipendenti e coadiuvanti	
E24	Totale interni in edifici	
E3	edifici ad uso residenziale	
GIOV	Giovani popolazione da 0 a 14 anni	
IND	Imprenditori, liberi professionisti e lavoratori in proprio	
P1	popolazione residente	
P138	Popolazione che si sposta quotidianamente fuori dal comune di residenza per motivo di studio o lavoro	
P47	Popolazione con titolo di Laurea o diploma universitario o terziario di tipo non universitario	
P62	Forze lavoro-Disoccupati e altre persone in cerca di occupazione	
P68	Popolazione occupata in Agricoltura	A,B
P72	Popolazione occupata nell'Industria	C,D,E,F
P77	Popolazione occupata nella Pubblica Amministrazione	L
ST07	Stranieri	
Terz	Popolazione occupata nei servizi esclusa la Pubblica Amministrazione	G,H,I,J,M,N,O,P,Q
Nome indicatore	descrizione	formula
DPOP	densità popolazione per km	$P1/AREAKM$
DAGR	densità occupati del primario per km	$P68/AREAKM$
DSEC	densità del secondario	$P72/AREAKM$
DTERZ	densità del terziario	$TERZ/AREAKM$
DPA	densità della pubblica amministrazione	$P77/AREAKM$
LAUPOP	laureati su pop residente	$P47/P1$
DISPOP	disoccupati su popolazione	$P62/P1$
GIOVPOP	giovani su popolazione	$GIOV/P1$
ADPOP	adulti su popolazione	$ADULTI/P1$
ANZPOP	anziani su popolazione	$ANZ/P1$
DIPPOP	dipendenti su pop residente	$DIP/P1$
INDPOP	indpendenti su pop res	$IND/P1$
MINTERNI	numero medio di interni per edificio	$E24/E3$
DSTRA	densità degli stranieri	$ST07/AREAKM$
PENDPOP	persone che si spostano fuori dal comune su pop resi	$P138/P1$

¹ Si ringrazia il dott. Abbate dell'Istat per la collaborazione nella scelta degli indicatori.

DPEND	densità dei pendolari	P138/AREAKM
MABSTANZE	numero medio di abitanti per stanza in abitazione	P1/A2

3 Scelta della scala e risoluzione dei dati

L'estensione della scala costituisce uno dei punti di forza e, al tempo stesso, di debolezza dell'elaborazione. Aumentando l'estensione della carta si rischia infatti di "appiattire" le diversità, che risaltano sicuramente ad una scala meno estesa. Tuttavia solo ampliando la scala, aumentando l'estensione a livello nazionale, i fenomeni possono ricondursi a situazioni ripetute e meglio identificabili. Quali e quante sono queste situazioni? Ovvero, quale deve essere il numero dei gruppi che si devono individuare?

Per determinare il numero dei gruppi bisogna ricorrere al buon senso, per riuscire a classificare in maniera significativa i diversi tessuti, in particolare quello urbano, per cui si è ritenuto opportuno iniziare eliminando dall'analisi tutte le sezioni (circa 100.000) che non contengono popolazione e unità locali, individuando così la zona verde². A livello metodologico³, in una prima fase si procede ad individuare un elevato numero di gruppi utilizzando il metodo K-means, i cui centroidi vengono riaggregati con una procedura gerarchica. Questo agevola enormemente, perché riduce la numerosità dei casi (sezioni) aggregandoli in insiemi simili, mantenendo l'analisi della distribuzione della dissomiglianza dei gruppi ottenuti attraverso la clusterizzazione gerarchica.

Nel parlare di rappresentatività del territorio, si intende verificare la consistenza dei gruppi ottenuti rispetto alla superficie e alla popolazione da essi rappresentate. È chiaro che non si può considerare rappresentativo un gruppo che comprenda, ad esempio, l'80% della superficie e il 90% della popolazione; una situazione del genere non avrebbe alcuna utilità interpretativa, contenendo un solo gruppo quasi l'intera distribuzione dei casi. Si è ritenuto congruo effettuare un primo tentativo con 30 gruppi (cfr. Tab. 2). L'aggregazione mostra una struttura interna latente interessante: ordinando i cluster per numerosità si ottiene un gruppo estremamente ampio (il 7, che comprende il 68% dei casi e che copre oltre la metà del territorio nazionale).

Effettivamente, come evidenziato in Fig. 1, il gruppo 7 rappresenta la quasi totalità dell'abitato; accanto ad esso si individua il gruppo 16, che sembra in prima accezione ricondursi ai centri di rango maggiore (spiccano infatti le città di Milano, Roma, Napoli).

Come si è accennato il risultato non è utile ai fini rappresentativi per restituire le differenziazioni locali, coprendo una superficie troppo ampia del territorio nazionale. Aumentando i gruppi il risultato non muta: la struttura urbana fortemente coesa non subisce mutamenti significativi passando ad una aggregazione delle sezioni in 50 gruppi, nella quale si rileva ancora la presenza di un gruppo sopra il 60% e di un cluster per i centri di rango

² Si sottolinea che nei diversi tentativi questo cluster è stato sempre individuato; la sua eliminazione facilita le operazioni di calcolo senza togliere validità ai risultati

³ Si veda al riguardo Bolasco (1999).

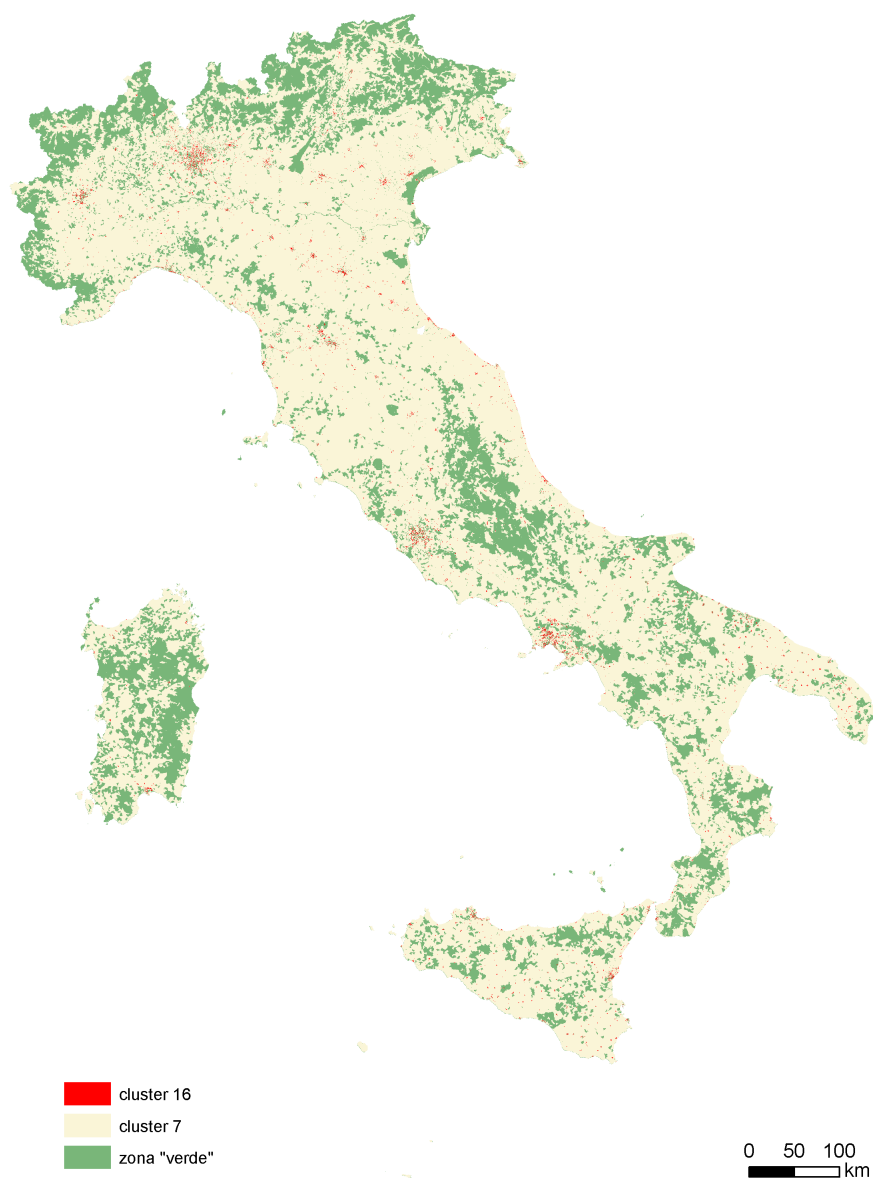
superiore (cfr, Tab. 3, che viene presentata limitatamente agli 8 cluster più numerosi⁴ - quelli con almeno 1000 sezioni ciascuno).

Tabella 2 - Risultato preliminare dell'aggregazione K means a 30 gruppi

cluster	numero di casi aggregati (sezioni)	% di casi su totale	% cumulate considerato n cluster
7	239655	68,06%	68,06%
16	66871	18,99%	87,05%
8	27825	7,90%	94,95%
14	8096	2,30%	97,25%
4	4000	1,14%	98,38%
27	3094	0,88%	99,26%
18	1003	0,28%	99,55%
10	832	0,24%	99,78%
21	279	0,08%	99,86%
15	266	0,08%	99,94%
28	76	0,02%	99,96%
9	53	0,02%	99,98%
17	24	0,01%	99,98%
12	16	0,00%	99,99%
25	10	0,00%	99,99%
26	9	0,00%	99,99%
3	7	0,00%	99,99%
22	4	0,00%	99,99%
29	4	0,00%	100,00%
11	2	0,00%	100,00%
20	2	0,00%	100,00%
30	2	0,00%	100,00%
1	1	0,00%	100,00%
2	1	0,00%	100,00%
5	1	0,00%	100,00%
6	1	0,00%	100,00%
13	1	0,00%	100,00%
19	1	0,00%	100,00%
23	1	0,00%	100,00%
24	1	0,00%	100,00%
totali	352138	100,00%	

⁴ Ben 30 dei 50 gruppi generati contengono meno di 10 casi (sezioni censuarie).

Figura 1 - Rappresentazione dei primi due cluster per numerosità di casi su cluster a 30 gruppi



È chiaro che si rilevano delle differenze, tuttavia la struttura è abbastanza simile, per cui si è preferito approfondire l'analisi a partire dal risultato a 30 gruppi, assunto come riferimento di base.

Tabella 3 - Numerosità dei cluster ottenuti con Kmeans a 50 gruppi (primi 8 gruppi)

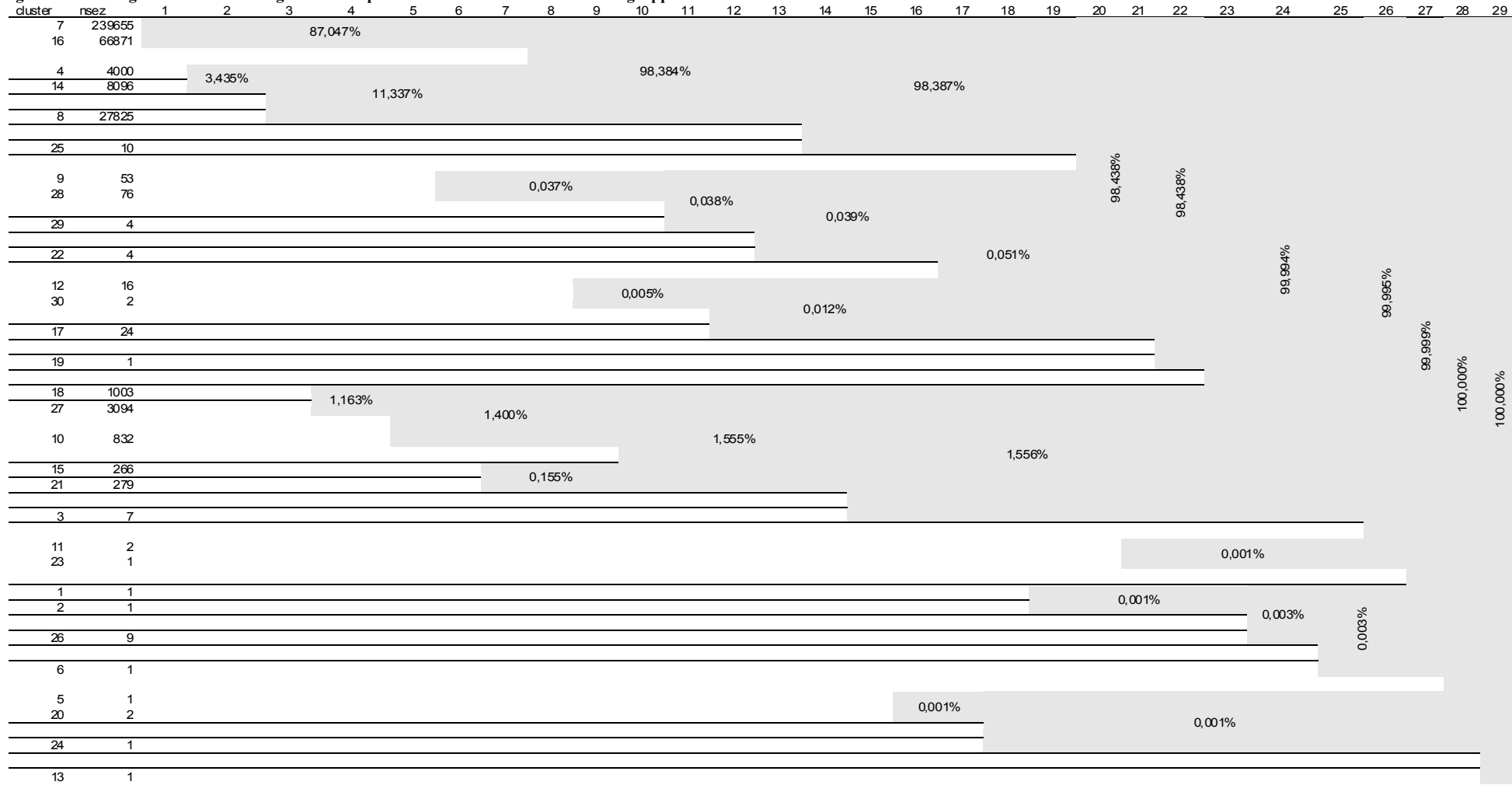
cluster	numero di casi aggregati (sezioni)	perc di casi su totale	percentuali cumulate considerato n cluster
2	224.774	63,83%	63,83%
1	67.442	19,15%	82,98%
47	31.921	9,06%	92,05%
7	9.919	2,82%	94,87%
41	7.452	2,12%	96,98%
5	4.006	1,14%	98,12%
39	2.194	0,62%	98,74%
50	1.538	0,44%	99,18%

A tal fine si è costruito un dendrogramma “cumulativo”, all’interno del quale ad ogni successiva aggregazione il nuovo gruppo rappresenta il totale delle sezioni aggregate. In questo modo è possibile rendersi conto, in valori relativi, dell’importanza del cluster in termini di numerosità.

Il dendrogramma cumulativo (Fig. 2) evidenzia come i gruppi più simili – e quindi primi ad aggregarsi – rappresentino la quasi totalità dei casi: al primo passaggio si aggrega in unico gruppo ben l’87%; trattasi di gruppo troppo grande per poter individuare specificità del territorio o, meglio, si può affermare che tale gruppo non è in grado di spiegare le differenze territoriali esistenti.

Tuttavia, l’analisi ha condotto ad un primo risultato su cui riflettere, perché non può essere un caso che tutte le sezioni urbane vengano aggregate in un unico cluster. È segno evidente che i cittadini “urbani” hanno comunque numerose caratteristiche simili che li accomunano e che fanno premio sulle caratteristiche differenzianti. In definitiva (utilizzando anche aggregazioni iniziali con un maggior numero di gruppi) la metodologia applicata conduce sempre ad una netta distinzione tra ciò che è urbano in senso lato (il *Central Business District*), e tutto quello che non è urbanizzato. Questa struttura latente emerge con chiarezza.

Figura 2 - Dendrogramma cumulativo gerarchico e percentuale dei cluster Kmeans a 30 gruppi



4 Il contenuto della fotografia: la rappresentatività

E' lecito chiedersi: le macro-categorie ottenute sono una valida rappresentazione? oppure ci restituiscono solo un'immagine sfocata? sono state d'aiuto per la costruzione di una realtà semplificata ma distinguibile?

In termini di rappresentatività, il risultato non è apparso soddisfacente, per cui si è deciso di procedere ad una serie di clusterizzazioni successive a catena, andando a scomporre piuttosto che ad aggregare. Questa procedura è solo in parte automatica, essa è infatti guidata dal ricercatore. Il criterio di scelta si basa sulla rappresentazione più dettagliata possibile delle differenze territoriali, per cui al passaggio successivo la procedura di clusterizzazione sarà ripetuta sul cluster contenente il maggior numero di casi. Si tratta, in definitiva, di bilanciare due esigenze opposte: migliorare la rappresentatività del risultato a prezzo di una riduzione nell'omogeneità del gruppo precedentemente individuato. Avendo considerato un così elevato numero di casi, sintetizzato in un numero così piccolo di gruppi, si sono fondamentalmente individuate solo le macro-categorie urbane. Si è palesata perciò la necessità di scendere nel dettaglio, dividendo ulteriormente le macro-categorie, che hanno assunto una funzione di struttura latente. Si è deciso di "pilotare" strategicamente il processo in un nuovo processo reiterato: individuata una macro-categoria compatta e diffusa, si è ritenuto opportuno che questa venisse sottoposta ad ulteriore analisi. La macro-categoria più diffusa è una sorta di territorio urbano indistinto, che è stato separato fino ad individuare il territorio più simile all'ambiente naturale.

L'esperienza ha mostrato la necessità di creare circa una decina di gruppi, numero rivelatosi congruo per classificare il territorio. Si è generato un processo a catena che ripeteva ad ogni passaggio la stessa operazione (clusterizzazione in 3 gruppi, metodo K-means).

Individuati i primi 3 gruppi, la procedura è stata applicata a quello più numeroso; teoricamente, la procedura può ripetersi fino al momento in cui i gruppi si esauriscono, ovvero il cluster più numeroso contiene un'unica unità. Evidentemente, una tale operazione risulterebbe di nulla utilità; anche in questo caso occorre determinare empiricamente un numero contenuto di cluster, capaci di individuare diversi tipi di tessuto territoriale.

I primi tre blocchi evidenziano subito una specificità: il cluster 2, sul quale si concentrerà l'attenzione. Avendo come obiettivo la rappresentatività della classificazione non avrebbe senso seguire pedissequamente la prassi statistica, che porterebbe a suddividere ulteriormente il gruppo più eterogeneo (il 2, composto però di appena 7 sezioni). È bene invece che queste eccezionalità rimangano unite, mentre è il cluster 1 quello che deve essere ulteriormente "spaccato", essendo il più numeroso. E così via: al secondo passaggio il cluster 1.1 presenta la numerosità più alta e la maggior similitudine e viene quindi successivamente scomposto.

4.1 Una comparazione tra risultati dei diversi metodi

Il confronto di appartenenza tra i risultati ottenuti con la metodologia di scomposizione proposta e il Kmeans a 30 gruppi è stato operato confrontando il primo passaggio di scomposizione e il risultato dei 30 gruppi precedentemente ottenuti, costruendo a tal fine una nuova matrice dove per ciascuna sezione censuaria si sono considerati il gruppo di appartenenza al primo passaggio della scomposizione (A) e quello di appartenenza nel K-means a 30 gruppi (B). La Tab. 4 mostra che al primo passaggio si rileva la quasi perfetta corrispondenza del cluster 1-A in quello 7-B; una buona corrispondenza si trova anche per il cluster 16-B, che ricade quasi interamente nel cluster 1-A.

Tabella 4 Confronto dei risultati nei diversi metodi proposti

Cluster di scomposizione (A)	Cluster kmeans a 30 gruppi (B)	numero casi
1	7	239.634
1	16	66.864
1	8	6.943
1	28	3
2	11	2
2	7	2
2	26	1
2	24	1
2	23	1
3	8	20.882
3	14	8.096
3	4	4.000
3	27	3.094
3	18	1.003
3	10	832
3	21	279
3	15	266
3	28	73
3	9	53
3	17	24
3	7	19
3	12	16
3	25	10
3	26	8
3	16	7
3	3	7
3	29	4
3	22	4
3	30	2
3	20	2
3	19	1
3	13	1
3	6	1
3	5	1
3	2	1
3	1	1

4.2 Un tentativo di interpretazione

L'analisi ha portato a riclassificare il territorio in sette cluster significativi per tipologia e dimensioni, tra loro diversi e riconducibili a tre aree. Il processo come precedentemente descritto ha permesso di classificare il territorio sulla base non solo della densità della popolazione, ma anche di altre caratteristiche strutturali quali, ad esempio, il numero dei laureati e l'intensità dell'abitato. Escludendo il cluster 2, che rappresenta dei casi assolutamente eccezionali (solamente 7 sezioni su tutto il territorio nazionale), i diversi raggruppamenti individuano 3 diverse tipologie di territorio (cfr. Tabella 5).

Tabella 5 - Riepilogo delle aree risultate dalla metodologia proposta

area	cluster
urbana	2 casi eccezionali (esclusi)
	3 città degli uffici
	1.3 città residenziale
	1.2 città periferica pseudo-terziaria
industriale	1.1.1 città degli operai
	1.1.2 transito industria-agricoltura
agricola	1.1.3.2 agricoltura urbanizzata
	1.1.3.3 agricoltura coltivata

In particolare l'analisi ha individuato un'area urbana suddivisa in 3 sottogruppi⁵, così caratterizzati:

- cluster 3. **città degli uffici**, detiene le caratteristiche peculiari dei Central Business District italiani: un'elevatissima densità dell'abitato, una forte specializzazione nella Pubblica Amministrazione e nel terziario in genere. La bassa incidenza della pendolarità indica che le persone residenti in queste aree lavorano prevalentemente nel loro quartiere; trattasi di aree fortemente specializzate nei diversi settori, che costituiscono la città "*ricca*";
- cluster 1.3. **città residenziale**, situata solitamente ai margini della "città degli uffici", riveste essenzialmente il ruolo di area residenziale di supporto alla specializzazione terziaria del centro, con minore dotazione di servizi, ha sicuramente un prezzo degli immobili più basso e più accessibile ai giovani, sensibilmente più presenti rispetto al centro (cluster 3). Offre però meno opportunità di lavoro e per questo motivo aumenta il numero dei pendolari.
- cluster 1.2. "**città periferica pseudo-terziaria**", collocata ai limiti della "città degli uffici" che costituisce la periferia e rappresenta la zona di espansione della nuova città,

⁵ Ai quali si va aggiungere il **cluster 2** che, come ricordato, ha carattere di eccezionalità: si tratta di un territorio limitatissimo, con un'alta percentuale di stranieri e di Pubblica Amministrazione (riconducibile ad ambasciate ed Enti sovranazionali presenti sul territorio nazionale).

intensamente edificata sfruttando la disponibilità di spazio, precedentemente vuoto, lontano dal centro, privo di attrattive storiche o paesaggistiche, che ne abbiano garantito una certa tutela. Si tratta di un'area dal profilo piuttosto variegato ma con tendenza verso il basso di molti indicatori demografici e sociali; la presenza di forza lavoro (non necessariamente occupata) è particolarmente elevata ed è orientata prevalentemente verso la Pubblica Amministrazione e l'industria. Costituisce il bacino di occupati a cui attinge quotidianamente la città per produrre.

Con il precedente cluster si completa l'analisi delle aree urbane, naturalmente circondate da aree industriali tipologie di attività storicamente escluse dal centro, anche per l'esigenza di ampi spazi a basso prezzo, che si dividono in due sottogruppi, così caratterizzati:

- cluster 1.1.1. **“città degli operai”**, un'area ancora caratterizzata da significativi insediamenti di popolazione, che si è andata via via spostando dall'estrema periferia dei centri urbani di rango superiori all'area industriale. Ha un basso indice di edificabilità per il fatto che ampio spazio è occupato da capannoni industriali e stabilimenti. È abitata prevalentemente da occupati nel secondario che si spostano nei diversi comuni sedi degli stabilimenti in cui prestano il proprio lavoro. Si tratta di quartieri-dormitorio, caratterizzati da una bassa presenza di servizi, che rende queste aree poco attraente sotto il profilo abitativo.
- cluster 1.1.2. **“periferia di transito industria-agricoltura”**, che completa l'area industriale; rappresenta la vera zona di espansione, il confine tra città e campagna, peraltro sempre più labile. E' caratterizzata da popolazione anziana, con poche occasioni di lavoro che costringono la forza lavoro a spostamenti quotidiani. A differenza della “città degli operai” ha una forte componente agricola, a segnalare il passaggio verso il territorio agricolo. La forte componente di lavoratori indipendenti è evidentemente legata al settore agricolo, non emergendo una forte componente industriale. È sostanzialmente un territorio di passaggio sia spaziale, in riferimento al pendolarismo, che generazionale, in riferimento all'inevitabile abbandono della campagna verso la città, meta del lavoro quotidiano dei giovani.

L'apparente vuoto residuo è una graduale scomparsa del tessuto urbano, come emerge analizzando gli ultimi due cluster che compongono l'Area agricola suddivisa in 2 sottogruppi, così caratterizzati:

- cluster 1.1.3.2, **“agricoltura urbanizzata”**. E' un'area non intensamente urbanizzata, posta ai limiti della città e costituisce la frontiera tra tessuti intensamente antropizzati e intensamente naturali. E' caratterizzata da una bassa densità di popolazione, nella quale sono sovra rappresentati i lavoratori indipendente e i disoccupati; la popolazione è relativamente anziana. L'area da un lato può essere vista come complementare alla

precedente (per il carattere di transizione verso il territorio agricolo), ma costituisce probabilmente la zona più prossima di espansione. Dove lo spazio abbonda il suo prezzo infatti è relativamente basso e conveniente per operazioni commerciali di vasta scala⁶.

- cluster 1.1.3.3, **“campagna coltivata”**. Si tratta di aree che vedono ancora l’agricoltura come attività prevalente; il territorio è abitato da occupati nell’agricoltura e lavoratori indipendenti. È la parte meno densamente abitata, ma offre lavoro ai residenti. Si tratta tuttavia di piccole attività contadine, visto il basso livello della forza lavoro; sono estremamente alti gli indipendenti, probabilmente coltivatori diretti. La popolazione giovane è indirizzata verso la città.

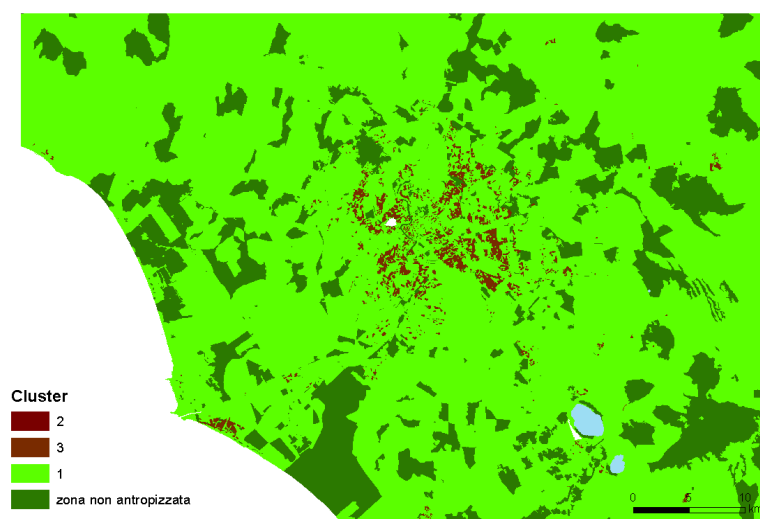
6 Un approfondimento mirato: la città di Roma

Una delle possibili applicazioni delle metodologie esposte è quella di confrontare i diversi livelli interpretativi del territorio. Per esempio la rappresentazione a scala comunale non permette di cogliere quel sottile e delicato rapporto tra società ed ecosistema, posto alla base delle scelte dello sviluppo sostenibile. È opportuna una risoluzione dei dati statici ad un livello più fine, capaci di cogliere meglio le differenze territoriali e loro specificità. Nel seguito, vengono presentati alcuni dei principali risultati ottenuti focalizzandosi al solo territorio della città di Roma e alla sua provincia. Il primo passaggio della clusterizzazione effettuata evidenzia la distribuzione dei primi tre gruppi individuati per la provincia di Roma (Figura 2).

L’apparente confusione della carta è in realtà un’istantanea di un territorio vivo e in continuo movimento. Analizzando i dati relativi alla città di Roma emerge una cartografia diversa da quella abituale, in cui la città non appare più come un pentagono ma come una stella. La città radiale emerge dalla carta mostrando il policentrismo che la caratterizza. Solo con queste rappresentazioni si può comprendere una città densamente abitata, che concentra le sue attività irradiandole dal centro lungo le antiche vie consolari. Dall’analisi emerge un modello di città concentrico con una specializzazione delle aree più ricche e produttive verso il centro, mentre spostandosi verso l’esterno la densità abitativa, la qualità dell’edificato e la capacità produttiva vanno diminuendo.

⁶ Ne sono un esempio la nascita dei centri commerciali intorno il comune di Roma in zone rurali fino a pochi anni fa, diffuse su tutto il Grande Raccordo Anulare.

Figura 2 Dettaglio sulla provincia di Roma del primo livello di clusterizzazione



Individuati i gruppi sarà interessante verificarne la loro localizzazione. Fattori storici hanno caratterizzato un genere di vita “romano”, spostando i diversi gruppi nel territorio della provincia. È opportuno allora verificare come questo processo abbia in realtà fatto emergere una diversa domanda di ambiente da parte dei soggetti. Mentre il gruppo sociale rappresenta la caratteristica della popolazione, l’uso del suolo individua la qualità ambientale; confrontando la loro localizzazione emerge la domanda di ambiente dei diversi gruppi, nonché il loro impatto sul territorio. Se si rapportasse la densità della popolazione della città con quella della campagna della provincia di Roma ci si potrebbe accorgere che la “città degli uffici” si trova ad un rapporto di 1 a 1000. Tale riscontro testimonia la percezione diversa del vivere in una città degli uffici invasa dal traffico piuttosto che la relativa tranquillità delle zone residenziali.

Tabella 6 Confronto delle densità della popolazione dei diversi cluster per la Provincia di Roma

cluster	nome	densità del cluster	equivalente al cluster 1.1.3.3.
3	la città degli uffici	29952,33	1030,129
1.3	città residenziale	12955,43	445,567
1.2	città pseudo terziaria	6197,961	213,1621
1.1.1	città degli operai	2945,714	101,3099
1.1.2	periferia di transito industria-agricoltura	1388,152	47,74174
1.1.3.2	l’agricoltura urbanizzata	669,0647	23,01067
1.1.3.1		347,2976	11,94437
1.1.3.3	la campagna coltivata	29,07627	1

È possibile classificare il suolo occupato dalla popolazione in diversi usi per poter analizzare la domanda di ambiente da parte dei diversi gruppi sociali. Utilizzando un raster satellitare è possibile dare forma alla città, individuarne i confini e le caratteristiche naturali. La popolazione occupa spazi urbani più o meno densi, ma esiste una preferenza per i diversi gruppi, (cfr. Tabella 7), dove sono evidenziate in grigio le modalità prevalenti di uso del suolo per cluster. La città degli uffici è evidentemente la città a tessuto continuo, mentre le altre forme urbane si caratterizzano per avere più spazi verdi in un tessuto discontinuo e rado. La campagna naturalmente si divide in zone ad agricoltura eterogenea e seminativo, mentre le aree protette sono prevalentemente utilizzate a seminativi.

7 Conclusioni

La metodologia e le applicazioni proposte mirano ad una analisi territoriale esplorativa antecedente alle analisi usuali. Normalmente il ricercatore ha un'idea di cosa cerca e la vuole dimostrare. La metodologia proposta parte invece dal confronto di indicatori, numeri puri, che solo in una visione comune riescono a far emergere le dinamicità locali. In questo però non differisce da molte altre analisi multivariate; il valore aggiunto di questa metodologia si rileva nella risoluzione spaziale, ad un livello di dettaglio tale da rendere un'immagine visiva molto realistica.

I gruppi socio-economici individuati costituiscono regioni dinamiche coese che occupano spazi contigui.

Le problematiche rilevate, e che saranno oggetto degli sviluppi futuri, riguardano l'utilizzo diacronico della metodologia proposta. Alla luce del prossimo Censimento è opportuno rilevare le difficoltà di un possibile confronto tra la situazione al 2001 e la futura al 2011. La cluster analysis non garantisce infatti l'individuazione degli stessi cluster nei prossimi censimenti; da qui un problema di confronto con i risultati attuali.

Tabella 7 Distribuzione delle superfici per i diversi usi del suolo e I cluster della popolazione

cluster	nome	Ambienti naturali a vegetazione e arbustiva, pascoli, praterie	Ambienti naturali a vegetazione e rada o assente e zone umide	Corpi idrici	Frutteti	Oliveti	Parchi urbani e aree ricreative	Seminativi	Superfici artificiali a tessuto continuo e aree industriali	Superfici artificiali a tessuto discontinuo e rado	Vigneti	Zone agricole eterogenee	Zone boscate	Totale di area km
3	la città degli uffici	0,21		0,30		0,03	0,41	1,60	25,84	9,90	0,02	1,83	0,00	40,15
1.03	città residenziale	0,04		0,14		0,01	1,08	1,70	15,89	17,51	0,13	1,00	0,29	37,79
	città pseudo													
1.02	terziaria	0,75		0,22	0,04	3,33	1,41	16,45	24,60	65,15	1,80	15,65	1,37	130,76
1.01.01	città degli operai	3,79		0,08	0,71	2,33	1,22	35,90	10,92	46,70	1,85	29,33	3,70	136,53
	periferia di transito													
	industria-													
1.01.02	agricoltura	8,57	1,12	20,68	0,39	7,83	2,43	59,08	19,00	85,27	9,03	80,62	52,67	346,69
	l'agricoltura													
1.1.3.2	urbanizzata	4,50		0,82	1,42	9,14	2,03	18,27	5,15	21,96	8,81	43,74	34,17	150,02
1.1.3.1		5,60	0,00	6,16	5,42	11,21	4,13	53,71	11,79	30,72	40,96	67,51	18,59	255,82
	la campagna													
1.1.3.3	coltivata	282,90	17,12	55,58	36,38	247,80	23,52	1.404,76	99,17	129,56	82,26	777,98	914,86	4.071,89
	zone disabitate	92,80	13,82	1,15	0,20	1,78	1,19	5,83	1,62	0,35	0,72	11,18	53,25	183,89
	di cui a parco	22,14		2,48	0,28	0,07	6,26	225,83	9,41	17,41	0,07	72,77	62,09	418,82

Bibliografia

- ALLEVA G., CELANT A. (1990), *Struttura urbana e terziario alle imprese*, CNR, Milano.
- ARBIA G., ESPA G. (1996), *Statistica economica territoriale*, CEDAM, Padova.
- BOLASCO S. (1999), *Analisi multidimensionale dei dati: metodi, strategie e criteri d'interpretazione*, Carocci, Roma.
- BONETTI E. (1976), *L'analisi fattoriale e le sue applicazioni nella ricerca geografica*, *Cultura e scuola*, no. 60.
- CORPOSANTO C. (2001), *La classificazione in sociologia: reti neurali, discriminant e cluster analysis*, Franco Angeli, Milano.
- KING C. A. M., COLE, J. P. (1969), *Quantitative geography: techniques and theories in geography*, J. Wiley & sons, London.
- LLOYD P. E., DICKEN P. (1992), *Spazio e localizzazione: un'interpretazione geografica dell'economia*, Franco Angeli, Milano.
- MERLIN P. (1981), *Analisi quantitativa e spazio urbano*, Franco Angeli, Milano.
- VALLEGA A. (1984), *Geografia regionale: avviamento metodologico*, Pàtron, Bologna.

ABSTRACT

The paper deals with an attempt to classify a territory in a few number of clusters (according to a set of socioeconomic and demographic indicators), at different levels of spatial resolution.

After an exploratory analysis of the whole Italian territory, the focus is on the city of Rome.

Original results have been reached using an automatic method to segment the main clusters in some others, better informative.

Statistical results supply a very effective spatial resolution, at different geographical scales.