



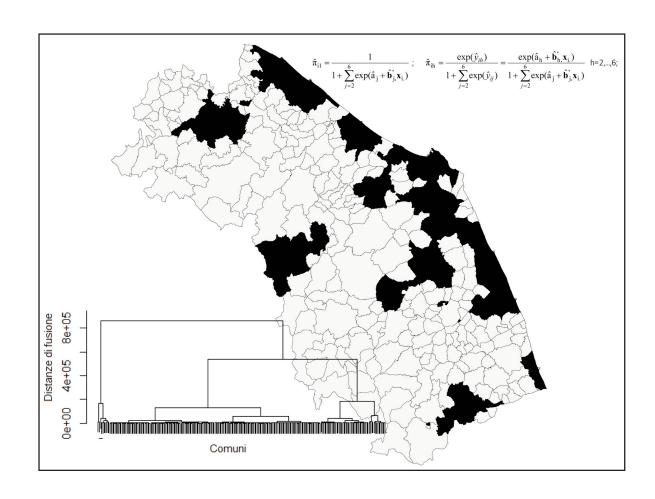






METODOLOGIA PER L'IDENTIFICAZIONE DELLA TIPOLOGIA DEI PICCOLI COMUNI

Elvio Mattioli, Gabriele Morettini





Rapporto di Ricerca realizzato nell'ambito del progetto PRIN "Piccoli comuni e coesione sociale: politiche e pratiche urbane per l'inclusione sociale e spaziale degli immigrati", finanziato dal MIUR - programmi di ricerca di interesse nazionale (Prin) 2010/2011

© Cattedra UNESCO SSIIM, 2014 Università luav di Venezia Ca' Tron, Santa Croce 1957 30135 Venezia info@unescochair-iuav.it www.unescochair-iuav.it

ISBN 978-88-87697-98-8













METODOLOGIA PER L'IDENTIFICAZIONE DELLA TIPOLOGIA DEI PICCOLI COMUNI

Elvio Mattioli, Gabriele Morettini

Università Politecnica delle Marche



CAPITOLO 1 - Premessa

1.1. I Piccoli Comuni

CAPITOLO 2 – La metodologia proposta

- 2.1 Le variabili considerate
- 2.2 La tecnica di identificazione dei Piccoli Comuni
 - 2.2.1. Variabili da considerare
 - 2.2.2. Distanze tra le singole unità
 - 2.2.3. Metodo di raggruppamento
 - 2.2.4. Numero dei gruppi da ritenere

CAPITOLO 3 - I risultati ottenuti

- 3.1. Risultati della cluster analysis
- 3.2 Classificazione parametrica della tipologia dimensionale

CAPITOLO 4 - Conclusioni

4.1 Conclusioni e proposte

Bibliografia di riferimento

Appendice

CAPITOLO 1 - PREMESSA

1.1. I Piccoli Comuni

Il rapporto con il territorio è sempre stato centrale in un paese come l'Italia, caratterizzato da grandi differenze regionali. Il dibattito economico e sociale ha però trascurato i Piccoli Comuni (PC). Il tema era affrontato con un'ottica stereotipata che alternativamente riproponeva le tradizionali virtù di una periferia sovente mitizzata o sottolineava l'irreversibile declino di luoghi isolati e marginali.

Recentemente si è però registrata una parziale inversione di rotta, nella consapevolezza che lo sviluppo regionale dipende da un'eterogenea molteplicità di elementi locali.

Questo rinnovato interesse si scontra con una ridotta conoscenza dei PC. Gran parte delle riflessioni si riduce infatti a considerazioni manieristiche, generiche e spesso obsolete, incapaci di descrivere compiutamente un universo articolato e in profondo mutamento.

L'evoluzione dei PC è intimamente connessa al fenomeno migratorio, principale responsabile dello spopolamento e delle criticità sorte nel Secondo Dopoguerra. A partire dagli anni '90 il quadro è stato ulteriormente modificato dai consistenti flussi di stranieri giunti in Italia, che hanno esercitato un variegato impatto su di un territorio già profondamente diversificato.

Le ricerche sull'immigrazione si sono focalizzate soprattutto sui sistemi urbani e sulle aree metropolitane, ambiti esaurientemente esaminati da una vasta letteratura internazionale. I PC sono stati ancora una volta relegati in secondo piano. In Italia, i pochi casi di studio si riferiscono in genere a situazioni peculiari e anomale (Prato, Rosarno), assurte agli onori della cronaca per la presenza di ampie comunità straniere o per episodi di turbolenza e tensione sociale.

Si avverte l'assenza di un'analisi sistematica e dettagliata dell'immigrazione nei PC. Tale lacuna è tanto più significativa se si considera il peculiare peso assunto dai PC nel modello migratorio nazionale. L'Italia si distingue dalla maggior parte delle esperienze europee e internazionali per la presenza di un'immigrazione diffusa sul territorio e nei PC (dove si registra un'incidenza di popolazione straniera significativa e talvolta superiore agli stessi livelli raggiunti nei contesti urbani).

Gli studi sono stati però frenati da scarsità di dati, problemi metodologici e dalla difficoltà di identificare i PC. Questi ultimi scontano infatti l'indeterminatezza dei confini e una certa vaghezza insita nella loro definizione.

Piccolo è un aggettivo qualificativo, da declinare nello spazio, nel tempo e nella specifica problematica esaminata. Bisogna infatti definire rispetto a chi, in quale epoca e quale parametro si utilizza l'aggettivo *piccolo*. Non sembra quindi opportuno continuare ad adottare la corrente metodologia, che tende a delimitare i confini del PC in base a soglie dimensionali assolute. In ambito amministrativo la definizione di PC attualmente più diffusa è quella riportata nel disegno di legge approvato dalla Camera dei Deputati il 18 aprile 2007 ("Misure per il sostegno e la valorizzazione dei piccoli comuni") secondo la quale: "*Per piccoli comuni si intendono i comuni con popolazione pari o inferiore a 5000 abitanti*". Già nel 2003 l'ANCI (Associazione Nazionale Comuni Italiani) utilizzava tale criterio, classificando come "piccoli" i comuni con popolazione inferiore o uquale a 5000 abitanti, pari al 72% dei comuni italiani.

Questa definizione appare fortemente riduttiva se si considera la grande eterogeneità dei caratteri demografici, economici e sociali dei comuni italiani e infatti, ad esempio, in una pubblicazione della Regione Marche ("I piccoli comuni delle Marche") il discrimine viene spostato da 5000 a 3000 residenti.

L'utilizzo di una soglia dimensionale univoca non permette di individuare i reali confini di un universo eterogeneo e complesso, in cui convivono periferie urbane e realtà isolate, aree industriali e località marginali. Tale contesto richiede una metodologia che sostituisca le rigide e discrezionali soglie assolute con valori elastici e differenziati per i vari ambiti territoriali, determinati prendendo in considerazione qualificazioni ulteriori rispetto al numero di abitanti.

La presente ricerca abbandona quindi la scelta *a priori* della sola dimensione demografica e cerca di individuare i PC in modo induttivo, partendo da dati empirici oggettivi e provenienti da fonti ufficiali, prontamente reperibili e riproducibili, in modo da assicurare facili comparazioni e tempestivi aggiornamenti. A tal fine sono stati utilizzati dati censuari.

CAPITOLO 2 LA METODOLOGIA PROPOSTA

2.1. Le variabili considerate

La metodologia proposta poggia su una pluralità di indicatori di diversa natura, che rispecchia le varie dimensioni in cui si può declinare l'aggettivo *piccolo*. La consistenza demografica rappresenta la principale discriminante ma vengono considerati (come suggerito dalla legge del 2007 sopra citata) anche elementi territoriali, ambientali, economici e sociali.

Per ogni comune sono state utilizzate le seguenti tipologie di indicatori:

- Livelli delle entità rilevanti;
- Indici che misurano la dinamica;
- Indici di intensità relativa.

La letteratura esistente suggerisce varie dimensioni di analisi. Un comune può essere *piccolo*.

- a livello di abitanti (popolazione residente)
- rispetto ai gruppi sociali (numero di famiglie)
- nella struttura fisica (numero complessivo di abitazioni)
- a livello economico (numero di addetti)
- riquardo al fenomeno osservato (entità e incidenza popolazione straniera)

L'aggettivo *piccolo* va inoltre esaminato nel tempo e nello spazio. L'ottica diacronica viene soddisfatta tramite l'inserimento delle variazioni nel periodo 2001-2011 mentre l'aspetto spaziale è approfondito attraverso lo studio delle 6 regioni cui fanno riferimento le Unità di Ricerca del Progetto, tutte accomunate da una forte presenza di PC. La regione sembra il contesto più opportuno in cui condurre, distintamente, la classifica dimensionale dei comuni. L'ambito regionale permette di eliminare opinabili scelte di altri aggregati territoriali e allo stesso tempo consente di tener conto delle diverse legislazioni amministrative, realtà storiche e peculiarità geografiche delle varie aree.

Nell'indagine svolta sono state prese in considerazione le sequenti variabili:

- 1) Popolazione residente al censimento del 2011 (POP11);
- 2) Numero di famiglie residenti al censimento del 2011 (FAM11);
- 3) Numero di abitazioni al censimento del 2011 (ABIT11);
- 4) Variazione percentuale 2001-2011 della popolazione residente (IPR01_11);
- 5) Variazione percentuale 2001-2011 delle famiglie residenti (IFM01_11);
- 6) Densità demografica al censimento del 2011 (DENS11).

È opportuno ricordare che la scelta delle 6 variabili è il risultato di una serie di sperimentazioni condotte sui comuni della Calabria, delle Marche e del Veneto nonché dal recepimento di indicazioni provenienti da diversi partecipanti del PRIN. Riguardo alle variabili utilizzate, sembra utile rilevare quanto segue:

- a. è stata esclusa la popolazione straniera, il cui studio è procrastinato ad approfondimenti successivi;
- b. non è stato incluso l'indice di attività, sia per il suo marginale effetto sulla classificazione sperimentata per la regione Marche sia per la difficoltà di reperire dati simili per altre regioni;
- c. è stata esclusa la variazione percentuale 2001-2011 delle abitazioni. Tale variabile presentava deboli correlazioni negative con le variabili di livello e, soprattutto, esercitava un impatto irrilevante sui risultati della classificazione;
- d. non è stata inserita l'altitudine, a causa della sua netta correlazione negativa con le variabili di livello. Per valutare la dimensione dei comuni è infatti opportuno utilizzare un insieme di variabili tra loro correlate positivamente ("consistenza interna"). A tale proposito, un'utile indicazione sulla consistenza interna delle variabili considerate per valutare l'entità "dimensione dei comuni" è fornita dall'indice α di Cronbach, che nella Tabella 1 assume valori confortanti:

Tabella 1– Indice α di Cronbach nelle 6 regioni considerate

REGIONI	CALABRIA	EMILIA	LAZIO	LOMBARDIA	MARCHE	VENETO
α	0.82	0.81	0.79	0.75	0.82	0.81

Ricordiamo che in letteratura la consistenza interna viene considerata *accettabile* per $0.7 \le \alpha < 0.8$ e *buona* per $\alpha > 0.8$.

2.2. La tecnica di identificazione dei Piccoli Comuni

Per individuare gruppi di comuni omogenei in base alle caratteristiche sopra ricordate è naturale ricorrere alla cluster analysis, cioè a quel vasto insieme di algoritmi che hanno l'obiettivo di suddividere un collettivo di unità statistiche in gruppi omogenei.

Si tratta di un'analisi statistica dei dati eminentemente esplorativa: non si conosce infatti il numero dei gruppi omogenei che possono essere individuati né se nella realtà osservata esistono gruppi ragionevolmente omogenei.

Le diverse tecniche di classificazione automatica che costituiscono la cluster analysis si prefiggono di dividere il collettivo in sottoinsiemi comprendenti elementi il più possibile omogenei (vicini) all' interno dei gruppi ed il più possibile diversi (distanti) se appartengono a gruppi distinti: questo è il criterio della omogeneità interna e diversità esterna.

L'applicazione di questa tecnica multivariata richiede scelte inerenti:

- a) Le variabili da considerare;
- b) Le distanze (o misure di diversità) tra le singole unità da adottare;
- c) Il metodo di raggruppamento da usare;
- d) Il numero di gruppi da ritenere.

2.2.1 Variabili da considerare

Nei dati raccolti è stato assegnato un congruo peso ai livelli delle variabili demografiche ed abitative, al fine di cogliere le più immediate caratteristiche dimensionali dei comuni.

Per caratterizzare ulteriormente le tipologie individuabili con il procedimento seguito sono stati, comunque, presi in considerazione anche gli indici che misurano la dinamica e quelli di intensità relativa.

2.2.2 Distanze tra le singole unità

Abbiamo adottato la semplice distanza euclidea al quadrato, che è stata scelta per il suo immediato significato e soprattutto per la notevole proprietà di scomposizione della variabilità di cui gode. Formalizzando, se indichiamo con:

$$\mathbf{X_i} = [x_{i1}, ..., x_{ij}, ..., x_{ik}]$$

il vettore le cui componenti rappresentano il valore delle k variabili (nel nostro caso k=6) rilevate nell'i-mo comune (ad esempio per le Marche i=1, ..., 246), la misura della diversità tra due generici comuni r-mo ed s-mo è:

$$d(\boldsymbol{x}_{r.}, \boldsymbol{x}_{s.}) = \sum_{i=1}^{k} (x_{rj} - x_{sj})^{2}.$$

2.2.3 Metodo di raggruppamento

Per quanto riguarda il metodo di raggruppamento si è scelto uno degli usuali metodi aggregativi gerarchici, considerando inizialmente ogni comune come un gruppo composto da una sola unità. I due gruppi più

vicini secondo il criterio adottato sono fusi in nuovo insieme che sostituisce i due aggregati originari. Tale procedimento viene più volte ripetuto sino a che si ottiene un solo gruppo.

Tra i vari algoritmi aggregativi disponibili abbiamo scelto il metodo di Ward poichè definisce esplicitamente una funzione obiettivo. Dato che lo scopo della classificazione è ottenere gruppi con la maggiore coesione interna, questo metodo richiede di riunire ad ogni passo del processo di aggregazione i due gruppi dalla cui fusione deriva il minimo incremento possibile della variabilità all'interno dei gruppi.

Definiamo la variabilità totale:

[1]
$$SQT = \sum_{i=1}^{k} \sum_{i=1}^{n} (x_{ij} - \mu_j)^2$$

dove μ_j è la media della j-ma variabile per l'intero collettivo. Data una partizione in g gruppi, tale variabilità può essere suddivisa nelle due componenti costituite dalla variabilità entro i gruppi SQW_g e dalla variabilità tra gruppi SQB_g :

[2]
$$SQT = SQW_q + SQB_q$$

Dove:

[3]
$$SQW_g = \sum_{j=1}^k \sum_{h=1}^g \sum_{i \in I_h} (x_{ij} - \mu_{hj})^2$$
 ,

essendo I_h l'insieme di unità del gruppo h-mo, μ_{hj} la media della variabile j-ma nel gruppo h-mo. La variabilità tra i gruppi è invece:

[4]
$$SQB_g = \sum_{i=1}^k \sum_{h=1}^g (\mu_{hj} - \mu_j)^2 n_h$$
,

essendo n_h il numero di unità del gruppo h-mo.

Nel passare da g a g-1 gruppi la variabilità all'interno dei gruppi SQW aumenta mentre, per la [2], la variabilità tra gruppi SQB diminuisce.

Nel metodo di Ward, ad ogni passaggio si aggregano tra loro i gruppi per cui vi è minore incremento della variabilità interna. Ciò avviene in quanto una partizione si considera tanto migliore quanto più i gruppi sono omogenei al loro interno e differenti l'uno dall'altro.

2.2.4 Numero di gruppi da ritenere

Nel caso di algoritmi gerarchici un utile strumento di aiuto alla decisione sul numero di gruppi da ritenere è costituito da un diagramma ad albero (dendrogramma) che permette di visualizzare le distanze alle quali si fondono i diversi gruppi.

Se tale distanza registra un salto notevole nel passare da h+1 ad h gruppi si registra un alto valore della differenza $d_h - d_{h+1}$; ciò indica che si stanno fondendo gruppi piuttosto eterogenei e quindi conviene ritenere h+1 gruppi.

Per valutare ulteriormente l'entità di tale incremento può essere utile considerare anche l'incremento

$$\text{relativo:} \ \ r_{h+1} = \frac{d_h - d_{h+1}}{d_{h+1}} \ .$$

Esistono, inoltre, diversi indici che aiutano a prendere questa decisione. In questa sede è sufficiente ricordare l'indice di determinazione $R_{\rm g}^2$ associato ad una partizione in g gruppi:

[5]
$$R_g^2 = \frac{SQB_g}{SOT}$$
 g = 1, ..., n,

dove SQT ed SQB_q hanno il significato e le espressioni illustrate nelle [2] e [4] sopra riportate.

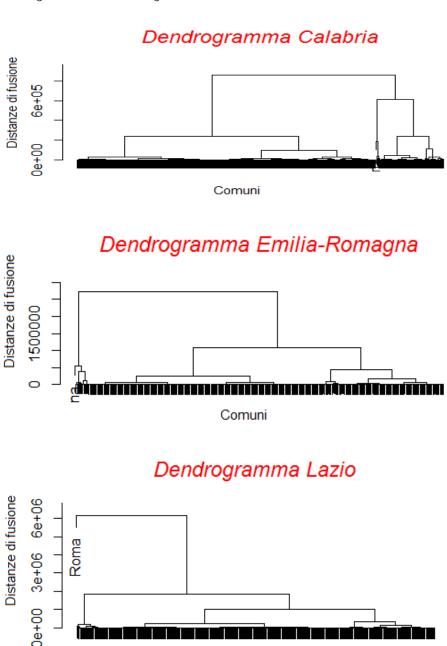
CAPITOLO 3 I RISULTATI OTTENUTI

3.1. Risultati della cluster analysis

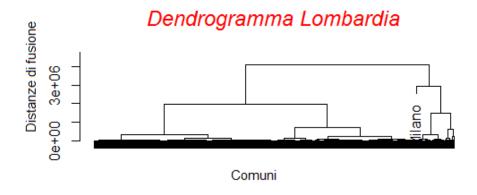
Nella selezione del numero dei gruppi da ritenere normalmente si cerca un compromesso tra il potere di sintesi della classificazione, tanto più elevato quanto più basso è il numero di gruppi scelto, ed il potere esplicativo dei gruppi, che aumenta quanto più omogenei sono i gruppi, cioè quanto più bassa è la variabilità al loro interno.

Un'importante indicazione su questa scelta è fornita dall'esame del dendrogramma che illustra il processo di fusione delle unità in gruppi (Figura 1.1).

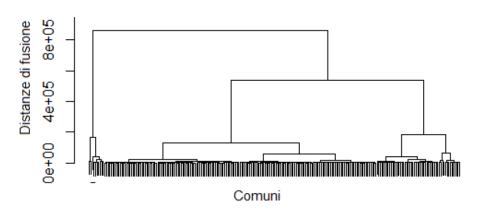
Figura 1.1 – Dendrogramma relativo alle regioni considerate



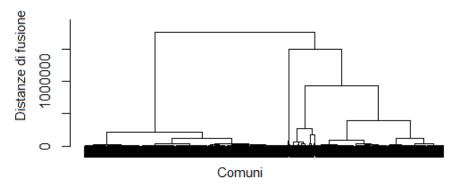
Comuni



Dendrogramma Marche



Dendrogramma Veneto



I dendrogrammi sopra riportati suggeriscono, con lievi discrepanze tra una regione e l'altra, la scelta di un numero di gruppi compreso tra 3 e 6. Un'ulteriore e più chiara indicazione è fornita dai valori assunti dall'indice di determinazione in corrispondenza della scelta di 3, 4 e 6 gruppi (Tabella 2).

Tabella 2– Indice di determinazione $R_{\ g}^2$ per diversi gruppi nelle 6 regioni considerate

n. di gruppi	CALABRIA	EMILIA	LAZIO	LOMBARDIA	MARCHE	VENETO
3	0.76	0.71	0.98	0.95	0.80	0.85
4	0.79	0.91	0.98	0.95	0.84	0.91
6	0.96	0.98	0.99	0.98	0.96	0.98

In virtù di quanto appena illustrato si è ritenuto opportuno suddividere i comuni in sei gruppi dimensionali, definiti come:

- gruppo 1: comuni piccolissimi,
- qruppo 2: comuni piccoli,
- gruppo 3: comuni medio piccoli,
- gruppo 4: comuni medi,
- gruppo 5: comuni medio grandi,
- gruppo 6: comuni grandi.

Le principali statistiche descrittive relative a ciascuna delle sei variabili considerate, all'interno dei sei gruppi e per ciascuna regione, sono esposte in Appendice dove si può rilevare come in tutti i casi esaminati la media delle fondamentali variabili di livello (POP11, FAM11 e ABIT11) è sempre crescente passando dal gruppo 1 al gruppo 6; un andamento analogo, con lievi eccezioni (Calabria gruppo 6 ed Emilia-Romagna gruppo 4) caratterizza la variabile di intensità relativa (DENS11), strettamente correlata con le prime. L'Appendice contiene anche una rappresentazione cartografica delle regioni esaminate. Le mappe illustrano la suddivisione del territorio (distinto in sei tipologie di comuni) effettuata attraverso la metodologia proposta.

I risultati evidenziano come attraverso una metodologia statistica multivariata si sia ottenuto un modello interpretativo affidabile e significativo. I Piccoli Comuni ricadono all'interno dei primi tre gruppi selezionati. Tale classificazione ha il merito di evitare rigide quanto sterili soglie dimensionali assolute, di semplice implementazione ma a volte poco aderenti alla realtà.

La metodologia proposta si adatta invece con flessibilità e pragmatismo a tutti i casi, dalle aree a popolazione diffusa delle Marche e della Calabria a regioni con significativi centri metropolitani come Lombardia e Lazio. L'approccio induttivo e multidisciplinare proposto permette di descrivere efficacemente contesti eterogenei, in cui la variabile popolazione residente è elemento importante ma non esclusivo.

In tale ottica occorre ribadire che, poiché la classificazione utilizza una pluralità di variabili, si verificano inevitabilmente situazioni di non ordinabilità: nel confronto tra due comuni alcuni indici possono risultare più elevati in un contesto che però registra valori inferiori per un'altra variabile.

Queste apparenti contraddizioni vengono risolte dal metodo adottato. Ad esempio, considerando esclusivamente la popolazione o soltanto il numero di famiglie, alcuni comuni verrebbero classificati nel gruppo 3 (comuni medio piccoli) ma con l'approccio proposto vengono invece inclusi nel gruppo 4 perché l'azione delle altre variabili li rende complessivamente più simili ai comuni del gruppo superiore.

3.2. Classificazione parametrica della tipologia dimensionale.

La classificazione dimensionale sinora effettuata si è basata sulla cluster analysis, una tecnica non parametrica che richiede l'esame congiunto di tutti i comuni della regione ma non consente immediate estensioni ai comuni di altre aree, nazionali o estere.

Occorre inoltre osservare che quando l'insieme da classificare è composto da poche unità i risultati ottenuti usando la cluster analysis sono scarsamente affidabili.

Per superare questi inconvenienti si è fatto ricorso alla regressione logistica multinomiale, una tecnica parametrica che consente di valutare un set di parametri utili per classificare immediatamente ogni singolo comune, anche al di fuori del territorio osservato precedentemente con la cluster analysis.

Indichiamo con $\hat{\mathbf{a}}$ e $\hat{\mathbf{B}}$ i parametri che la regressione logistica stima attraverso i dati di una regione. Per ciascun comune i-mo per il quale sono state rilevate le consuete variabili censuarie (raccolte nel vettore $\mathbf{x_{i.}}$) e considerata come variabile di risposta (vettore $\hat{\mathbf{y}}_{i.}$) la classificazione ottenuta con la cluster analysis, è possibile calcolare i logit:

[6]
$$\hat{\mathbf{y}}_{\mathbf{i}} = \hat{\mathbf{a}} + \hat{\mathbf{B}}\mathbf{x}_{\mathbf{i}} = \begin{bmatrix} \hat{\mathbf{a}}_2 + \hat{\mathbf{b}}_2' \mathbf{x}_{\mathbf{i}} & \hat{\mathbf{a}}_3 + \hat{\mathbf{b}}_3' \mathbf{x}_{\mathbf{i}} & \dots & \hat{\mathbf{a}}_6 + \hat{\mathbf{b}}_6' \mathbf{x}_{\mathbf{i}} \end{bmatrix}' = \begin{bmatrix} \ln \frac{\hat{\pi}_{i2}}{\hat{\pi}_{i1}} & \ln \frac{\hat{\pi}_{i3}}{\hat{\pi}_{i1}} & \dots & \ln \frac{\hat{\pi}_{i6}}{\hat{\pi}_{i1}} \end{bmatrix}'$$

e quindi le probabilità:

$$[7] \quad \hat{\boldsymbol{\pi}}_{i1} = \frac{1}{1 + \sum_{j=2}^{6} \exp(\hat{\boldsymbol{a}}_j + \hat{\boldsymbol{b}}_{j.}^{'} \boldsymbol{x}_{i.})}; \quad \hat{\boldsymbol{\pi}}_{ih} = \frac{\exp(\hat{y}_{ih})}{1 + \sum_{j=2}^{6} \exp(\hat{y}_{ij})} = \frac{\exp(\hat{\boldsymbol{a}}_h + \hat{\boldsymbol{b}}_{h.}^{'} \boldsymbol{x}_{i.})}{1 + \sum_{j=2}^{6} \exp(\hat{\boldsymbol{a}}_j + \hat{\boldsymbol{b}}_{j.}^{'} \boldsymbol{x}_{i.})} \quad h=2,...,6;$$

Naturalmente, il comune i-mo è assegnato al gruppo che presenta la probabilità più alta.

La sperimentazione condotta su ogni singola regione fornisce una nuova classificazione che mostra elevate similitudini con le ripartizioni prodotte mediante la cluster analysis, a testimonianza che i gruppi in precedenza individuati possono essere considerati come "gruppi naturali".

Successivamente i parametri $\hat{\mathbf{a}}$ e $\hat{\mathbf{B}}$, sinora esaminati per ciascuna regione sono stati calcolati per l'insieme dei 3498 comuni appartenenti all'area sinora considerata.

La riclassificazione svolta in questo caso, caratterizzato dalla maggiore eterogeneità esistente tra le sei regioni, non presenta l'elevato grado di accostamento osservato nei casi precedenti ma i risultati ottenuti sembrano comunque d'indubbio interesse.

L'affidabilità del modello, testata attraverso il rapporto di verosimiglianza, evidenzia un livello di confidenza molto elevato, superiore al 99%.

Le stime dei parametri $\hat{\mathbf{a}}$ e $\hat{\mathbf{B}}$, che mediante la [7] consentono di classificare i comuni, sono riportate nella Tabella 3.

Tabella 3- Parametri stimati mediante la regressione logistica multinomiale

h	Intercetta	Pop11	Fam11	Abit11	IPR01_11	IFM01_11	DENS11
2	-3.42523	0.00162	-0.00311	0.00068	-0.04981	0.04725	0.00333
3	-6.87744	0.00243	-0.00462	0.00109	-0.00918	0.00300	0.00564
4	-12.50568	0.00350	-0.00658	0.00122	0.02682	-0.06909	0.00691
5	-18.92759	0.00369	-0.00665	0.00119	0.02417	-0.03206	0.00729
6	-30.78750	0.00385	-0.00624	0.00062	-0.70530	0.34749	0.00460

Fonte: nostre elaborazioni su dati Istat

Ci limitiamo a osservare il diminuire dell'importanza dell'intercetta a_h , che determina la stima della probabilità con un peso pari a $exp(a_h)$, all'aumentare della dimensione del comune e la correzione apportata dalla variabile "famiglie residenti" all'effetto delle altre due variabili dimensionali popolazione e abitazioni. La classificazione dei comuni ottenuta con il metodo parametrico, messa a confronto con quella non parametrica che tiene conto delle specificità regionali, è sinteticamente riportata nella tabella che segue.

Tabella 4- Distribuzioni dei comuni in base alle differenti metodologie proposte

	Gruppo 1	Gruppo 2	Gruppo 3	Gruppo 4	Gruppo 5	Gruppo 6	Totale
Cluster Analysis	1561	972	646	217	93	9	3498
Regressione logistica	1735	881	603	188	84	7	3498
Δ	174	-91	-43	-29	-9	-2	0
Δ%	4.97	-2.60	-1.22	-0.82	-0.25	-0.05	0

La regressione logistica produce un incremento di 174 comuni del primo gruppo, pari a circa il 5% dei 3498 comuni considerati. Corrispondentemente si ha una riduzione nel numero di unità appartenenti agli altri gruppi; tale flessione diminuisce all'aumentare della dimensione dei comuni.

Tenendo conto dell'eterogeneità delle regioni considerate, localizzate in ambiti geografici assai differenti (una nel Nord-ovest, due nel Nord-est, due nel Centro ed una nel Sud), i risultati ottenuti nella Tabella 4 sembrano abbastanza confortanti e corroborano la scelta di continuare a utilizzare la regressione logistica per effettuare estrapolazioni anche su ambiti territoriali più vasti.

Conclusioni

4.1 Conclusioni e proposte

La metodologia proposta è funzionale ad ulteriori, articolati approfondimenti.

In particolare si sta classificando, mediante la cluster analysis, i comuni di tutte le regioni italiane. I risultati così ottenuti, oltre ad essere di per sé interessanti, consentiranno di stimare nuovamente i parametri $\hat{\mathbf{a}}$ e

 $\hat{\mathbf{B}}$ in modo da renderli maggiormente affidabili e abbastanza robusti per poter classificare comuni di altri paesi.

I dati raccolti servono a evidenziare le specificità dei PC rispetto ai centri medio-grandi e ad effettuare confronti tra gli stessi PC, distinti per dimensione o per regione di appartenenza. Tale analisi verrà poi proficuamente estesa alla scala europea; le eventuali divergenze o punti di contatto permetteranno di verificare se la presenza di immigrati nei PC è una peculiarità italiana, per enfatizzare eventuali diversità nelle loro caratteristiche demografiche e per individuare le determinanti delle scelte insediative degli stranieri in alcune nazioni europee.

Il lavoro svolto è inoltre propedeutico a una seconda fase di ricerca in cui verranno analizzate le variabili economiche, sociali, geografiche e demografiche sinora non utilizzate. Lo studio di tali aspetti consentirà di delineare specifiche tipologie di piccoli comuni, da esaminare tramite un' accurata indagine sul campo.

Riferimenti bibliografici

- Camera dei Deputati della Repubblica Italiana (2007) "Disegno di legge n. 1516" Misure per il sostegno e la valorizzazione dei piccoli comuni" approvato il 18 aprile 2007", consultabile al sito http://www.piccolicomuni.anci.it/
- Cittàltalia, Fondazione Anci Ricerche (2013), "Atlante dei piccoli comuni" (http://www.anci.it/Contenuti/Allegati/Atlante%20dei%20Piccoli%20Comuni%202013.pdf)
- Crescimanno, A., Ferlaino, F., Rota F.S. (2009) "Classificazione della marginalità dei piccoli comuni del Piemonte", IRES Piemonte Working Paper n.235, consultabile al sito http://www.ires.piemonte.it/
- De Lillo, A., Argentin, G., Lucchini, M., Sarti, S. e M. Terraneo (2007). *Analisi multivariata per le Scienze sociali*, Milano, Pearson Italia.
- DPS-Istat (2013), "Le aree interne: di quale territori parliamo? Nota esplicativa sul metodo di classificazione delle aree interne", consultabile al sito (http://www.dps.tesoro.it/Aree_interne/doc/Nota%20Territorializzazione%20Al_03%20marzo_2013.pdf)
- Gallucci, M. e L. Leone (2012). *Modelli statistici per le scienze sociali*, Milano, Pearson Italia.
- Regione Marche, Sistema Informativo Statistico (2003), "I piccoli comuni delle Marche", consultabile al sito http://www.sistar.marche.it
- Zani, S., e A. Cerioli (2007). *Analisi dei dati e data mining per le decisioni aziendali*, Milano, Giuffrè.

Appendice

A- Gruppi dimensionali

A.1- Gruppi dimensionali dei comuni della Calabria (Indice di determinazione $R^2 = 0.96$)

		Grup	ро 1		
Variabile	N	media	Dev Standard	Minimo	Massimo
POP11	161	1037.52	377.24	279.00	1833.00
FAM11	161	446.12	155.14	158.00	814.00
ABIT11	161	740.57	323.74	257.00	2144.00
IPR01_11	161	-9.91	9.35	-41.21	20.38
IFM01_11	161	1.26	11.04	-22.58	60.24
DENS11	161	61.80	60.26	8.20	606.42
	•	Grup	ро 2		1
Variabile	N	Media	Dev Standard	Minimo	Massimo
POP11	172	3057.12	1057.41	1405.00	6492.00
FAM11	172	1225.88	402.72	637.00	2446.00
ABIT11	172	1900.67	791.67	825.00	4981.00
IPR01_11	172	-3.41	11.19	-33.59	43.58
IFM01_11	172	8.80	14.03	-20.79	64.67
DENS11	172	127.98 Gru p	119.59	16.39	772.25
Variabile	N	media Grup	Dev Standard	Minimo	Massimo
POP11	49	7484.45	1814.86	4490.00	11115.00
FAM11	49	2947.92	679.54	1870.00	4272.00
ABIT11	49	4852.22	1384.08	2939.00	8453.00
IPRO1_11	49	-0.61	6.44	-11.89	21.22
IFM01_11	49	13.20	10.38	-2.80	49.92
DENS11	49	281.91	308.86	55.92	1790.79
Variabile	l Ni		po 4 Dev Standard	Minimo	Magginga
	N 22	Media			Massimo
POP11		19571.18	8300.75	10065.00	38501.00
FAM11	22	7474.50	3180.94	4202.00	14406.00
ABIT11	22	10887.23	5207.65	4976.00	23921.00
IPRO1_11	22	0.13	4.57	-7.06	11.20
IFM01_11	22	11.45	6.44	0.04	21.01
DENS11	22	371.09	249.34	63.40	1100.21
Manifest -	T NI	Grup		A4::	
Variabile	N	Media	Dev Standard	Minimo	Massimo
POP11	4	72016.25	12684.83	58881.00	89364.00
FAM11	4	27932.25	5220.89	21526.00	33967.00
ABIT11	4	31224.75	5866.26	24650.00	38541.00
IPR01_11	4	-3.28	2.71	-6.18	-0.23
IFM01_11	4	7.37	2.29	4.05	9.00
DENS11	4	846.14	689.19	323.53	1835.23
			po 6		
Maniahil:	N I		Dev Standard	Minimo	Massimo
Variabile	N	Media		100017.00	10001700
POP11	1	180817.00	0.00	180817.00	180817.00
POP11 FAM11	1	180817.00 68320.00	0.00 0.00	68320.00	68320.00
POP11 FAM11 ABIT11	1 1 1	180817.00 68320.00 75992.00	0.00 0.00 0.00	68320.00 75992.00	68320.00 75992.00
POP11 FAM11 ABIT11 IPR01_11	1 1 1	180817.00 68320.00 75992.00 0.26	0.00 0.00 0.00 0.00	68320.00 75992.00 0.26	68320.00 75992.00 0.26
POP11 FAM11 ABIT11	1 1 1	180817.00 68320.00 75992.00	0.00 0.00 0.00	68320.00 75992.00	68320.00 75992.00

A.2- Gruppi dimensionali dei comuni dell'Emilia Romagna (Indice di determinazione $R^2 = 0.98$)

		Grup	ро 1		
Variabile	N	media	Dev Standard	Minimo	Massimo
POP11	219	3670.46	2023.85	92.00	7773.00
FAM11	219	1593.11	819.86	68.00	3210.00
ABIT11	219	2177.87	949.72	328.00	4555.00
IPR01_11	219	5.30	14.39	-34.29	71.96
IFM01_11	219	12.18	14.97	-25.35	95.14
DENS11	219	104.90	116.49	3.81	1266.89
	l .		ро 2	-	-
Variabile	N	Media	Dev Standard	Minimo	Massimo
POP11	94	11878.65	3008.64	7275.00	18635.00
FAM11	94	4983.26	1322.55	3377.00	8167.00
ABIT11	94	5466.10	1480.22	3507.00	9281.00
IPR01_11	94	14.15	8.37	-5.76	32.31
IFM01_11	94	21.40	8.82	3.83	42.38
DENS11	94	309.30	310.54	39.44	2671.08
V ' 1 '1		Grup			
Variabile	N	media	Dev Standard	Minimo	Massimo
POP11	22	27350.86	5933.68	18591.00	39885.00
FAM11	22	11751.91	2624.09	7718.00	16756.00
ABIT11	22	14787.36	6925.30	8777.00	41210.00
IPR01_11	22	10.68	6.83	0.08	26.14
IFM01_11	22	18.36	7.15	5.79	33.67
DENS11	22	593.76	544.31	71.01	2029.18
			ро 4		
Variabile	N	Media	Dev Standard	Minimo	Massimo
POP11	6	84273.83	23194.47	57748.00	116434.00
FAM11	6	36637.83	10333.49	25143.00	50444.00
ABIT11	6	39454.17	11254.97	27623.00	53457.00
IPR01_11	6	6.76	1.72	4.93	9.42
IFM01_11	6	13.54	2.18	9.31	15.66
DENS11	6	475.60	206.79	267.65	848.40
	•	Grup	ро 5		
Variabile	N	Media	Dev Standard	Minimo	Massimo
POP11	6	157168.67	18892.62	132545.00	179149.00
FAM11	6	71214.33	8670.16	60181.00	81743.00
ABIT11	6	79986.83	10585.18	69901.00	94055.00
IPR01_11	6	7.97	5.64	1.19	14.24
IFM01_11	6	15.44	5.84	7.86	23.27
DENS11	6	657.76	325.68	235.14	1028.70
	l .	Grup	ро 6		
Variabile	N	Media	Dev Standard	Minimo	Massimo
POP11	1	371337.00	0.00	371337.00	371337.00
FAM11	1	194042.00	0.00	194042.00	194042.00
	1	206582.00	0.00	206582.00	206582.00
ABIT11					
ABII11 IPRO1 11	1	0.03	0.00	0.03	0.03
	1	9.21	0.00	0.03 9.21	9.21

A.3- Gruppi dimensionali dei comuni del Lazio (Indice di determinazione $R^2 = 0.99$)

		Grup			
Variabile	N	media	Dev Standard	Minimo	Massimo
POP11	165	1019.35	563.68	97.00	2256.00
FAM11	165	458.62	233.45	73.00	968.00
ABIT11	165	643.89	315.11	88.00	1514.00
IPR01_11	165	0.90	9.81	-23.71	28.33
IFM01_11	165	9.41	10.31	-15.82	40.30
DENS11	165	61.28	52.69	3.55	395.90
			ро 2		
Variabile	N	Media	Dev Standard	Minimo	Massimo
POP11	96	3437.19	985.74	551.00	5346.00
FAM11	96	1439.68	389.57	310.00	2453.00
ABIT11	96	1824.01	676.82	1007.00	5201.00
IPR01_11 IFM01_11	96 96	5.52 15.00	12.32 12.95	-13.40 -3.88	62.45 74.59
DENS11	96	129.42	124.53	7.06	1127.58
DENSTI	90		ppo 3	7.00	1127.30
Variabile	N	media	Dev Standard	Minimo	Massimo
POP11	68	9490.09	2784.12	5655.00	15596.00
FAM11	68	3790.60	1035.73	2227.00	6236.00
ABIT11	68	3984.54	1162.67	2197.00	8909.00
IPR01_11	68	15.84	15.78	-4.08	64.80
IFM01_11	68	25.70	17.06	2.79	77.71
DENS11	68	259.71	170.28	39.03	944.31
DENSTI	1 00		ppo 4	33.03	J44.51
Variabile	N	Media	Dev Standard	Minimo	Massimo
POP11	23	21181.17	3587.70	16016.00	30572.00
FAM11	23	8327.43	1124.62	6378.00	11252.00
ABIT11	23	9247.70	2145.67	6331.00	15023.00
IPR01_11	23	12.99	11.84	-1.97	38.05
IFM01_11	23	22.82	12.07	8.47	45.87
DENS11	23	543.38	407.52	57.34	1533.15
	•	Grup	ро 5	•	
Variabile	N	Media	Dev Standard	Minimo	Massimo
POP11	25	50202.24	18587.95	33658.00	117892.00
FAM11	25	20139.84	7101.99	12158.00	45308.00
ABIT11	25	22203.76	7407.24	12506.00	44605.00
IPR01_11	25	17.00	14.75	-4.09	65.48
IFM01_11	25	28.78	15.51	7.13	70.47
DENS11	25	756.37	613.09	155.60	2864.61
	The state of the s	Grur	ро 6		•
Variabile	N	Media	Dev Standard	Minimo	Massimo
Variabile POP11	N		Dev Standard 0.00	Minimo 2617175.00	Massimo 2617175.00
		Media			
POP11	1	Media 2617175.00	0.00	2617175.00	2617175.00
POP11 FAM11	1 1	Media 2617175.00 1184641.00	0.00 0.00	2617175.00 1184641.00	2617175.00 1184641.00
POP11 FAM11 ABIT11	1 1 1	Media 2617175.00 1184641.00 1151762.00	0.00 0.00 0.00	2617175.00 1184641.00 1151762.00	2617175.00 1184641.00 1151762.00

		Grupp	o 1		
Variabile	N		Dev Standard	Minimo	Massimo
POP11	780	1275.01	766.42	30.00	3100.00
FAM11	780	533.86	304.95	17.00	1287.00
ABIT11	780	742.80	424.90	24.00	2736.00
IPR01_11	780	8.14	18.11	-27.75	219.07
IFM01_11	780	13.96	19.21	-19.85	238.96
DENS11	780	187.66	211.68	2.48	1597.17
BENSTI	700	Grupp		2.10	1337.17
Variabile	N		Dev Standard	Minimo	Massimo
POP11	437	4369.08	1205.23	540.00	7426.00
FAM11	437	1774.69	479.88	271.00	3161.00
ABIT11	437	2141.98	770.97	1086.00	7853.00
IPR01_11	437	13.77	12.78	-7.06	103.78
IFM01_11	437	20.90	14.04	-0.31	114.32
DENS11	437	619.68	534.79	6.30	3751.83
	T	Grupp			T
Variabile	N		Dev Standard	Minimo	Massimo
POP11	164	8550.75	1186.51	6703.00	11270.00
FAM11	164	3502.59	506.47	2685.00	4822.00
ABIT11	164	3711.13	694.65	2684.00	8072.00
IPR01_11	164	12.04	9.17	-9.23	57.30
IFM01_11	164	19.28	10.00	-4.42	65.78
DENS11	164	854.41	586.50	76.99	2636.14
		Grupp	o 4		
Variabile	N	Media	Dev Standard	Minimo	Massimo
POP11	121	16576.50	4240.67	10350.00	26617.00
FAM11	121	6848.13	1782.87	4531.00	11398.00
ABIT11	121	7126.69	1870.42	4597.00	12148.00
IPR01_11	121	10.04	8.00	-5.65	36.05
IFM01_11	121	17.35	8.32	1.72	47.31
DENS11	121	1494.32	1247.93	184.48	7601.70
	•	Grupp	o 5		
Variabile	N	Media	Dev Standard	Minimo	Massimo
POP11	40	53753.93	31151.95	26793.00	189902.00
FAM11	40	23660.68	14548.16	12245.00	87459.00
ABIT11	40	24679.18	14996.32	12700.00	87656.00
IPR01_11	40	4.77	6.40	-5.13	22.55
 IFM01_11	40	12.11	7.41	-0.66	34.43
DENS11	40	2516.53	1556.32	452.10	6540.05
	I	Grupp			
Variabile	N	Media	Dev Standard	Minimo	Massimo
POP11	1	1242123.00	0.00	1242123.00	1242123.00
FAM11	1	627117.00	0.00	627117.00	627117.00
ABIT11	1	645833.00	0.00	645833.00	645833.00
IPR01_11	1	-1.12	0.00	-1.12	-1.12
					=
IFM01_11	1	6.62	0.00	6.62	6.62

		Grup	ро 1		
Variabile	N	media	Dev Standard	Minimo	Massimo
POP11	87	944.38	430.89	122.00	1749.00
FAM11	87	391.23	170.88	55.00	707.00
ABIT11	87	625.21	389.71	75.00	2392.00
IPR01_11	87	-0.69	8.76	-16.22	30.09
IFM01_11	87	7.32	9.47	-17.58	43.14
DENS11	87	55.22	39.17	3.87	179.83
			ро 2	<u> </u>	-
Variabile	N	Media	Dev Standard	Minimo	Massimo
POP11	88	3110.55	1018.03	1839.00	5378.00
FAM11	88	1229.53	404.53	667.00	2175.00
ABIT11	88	1529.92	538.49	722.00	3068.00
IPR01_11	88	8.17	12.69	-8.85	43.52
IFM01_11	88	17.33	14.48	0.36	61.30
DENS11	88	150.15	130.63	19.52	787.64
	1		po 3	T	T
Variabile	N	media	Dev Standard	Minimo	Massimo
POP11	40	8486.30	2190.77	3716.00	13153.00
FAM11	40	3294.63	841.49	1973.00	5358.00
ABIT11	40	4018.43	1279.92	2459.00	7618.00
IPR01_11	40	12.47	12.41	-3.74	63.88
IFM01_11	40	21.76	13.35	2.95	74.39
DENS11	40	334.49	256.80	39.80	1184.37
	•	Grup	ро 4		
Variabile	N	Media	Dev Standard	Minimo	Massimo
POP11	14	20536.14	6331.24	14858.00	33991.00
FAM11	14	8198.57	2625.92	5605.00	13381.00
ABIT11	14	9318.21	2681.78	5980.00	14568.00
IPR01_11	14	7.09	5.56	-5.78	15.49
IFM01_11	14	16.27	6.16	2.83	28.22
DENS11	14	596.27	526.47	68.44	1814.68
	•	Grup	ро 5	•	
Variabile	N	Media	Dev Standard	Minimo	Massimo
POP11	8	45467.25	8151.36	37016.00	62901.00
FAM11	8	18923.38	3705.89	14994.00	26931.00
ABIT11	8	21992.25	5844.96	15598.00	34075.00
IPRO1 11	8	4.05	3.50	-2.76	9.34
IFM01 11	8	13.97	4.97	5.64	21.26
DENS11	8	631.50	524.88	297.24	1848.45
DENSTI			po 6	251.24	1040.43
Variabile	N	Media	Dev Standard	Minimo	Massimo
POP11	2	97367.00	4426.49	94237.00	100497.00
FAM11	2	42195.00	3570.89	39670.00	44720.00
ABIT11	2	41702.50	3507.96	39222.00	44183.00
IPRO1_11	2	1.72	2.45	-0.01	3.46
IFM01_11					
DENS11	2 2	10.75	3.03	8.61	12.90
DENSII	1 /	774.18	43.58	743.37	805.00

		Grun	ро 1		
Variabile	N	media	Dev Standard	Minimo	Massimo
POP11	329	2676.54	1357.34	123.00	5504.00
FAM11	329	1056.11	514.78	61.00	2145.00
ABIT11	329	1436.52	723.12	202.00	5557.00
IPR01_11	329	4.09	9.99	-20.63	55.48
IFM01_11	329	12.00	11.44	-17.30	71.96
DENS11	329	149.93	113.17	5.53	631.38
BENSTI	323		ppo 2	3.33	031.30
Variabile	N	Media	Dev Standard	Minimo	Massimo
POP11	118	7325.51	1217.28	4317.00	9797.00
FAM11	118	2827.25	479.33	2046.00	3862.00
ABIT11	118	3327.59	1118.40	2146.00	8869.00
IPR01_11	118	12.96	8.80	-6.91	41.54
IFM01_11	118	22.47	10.39	-2.44	55.40
DENS11	118	359.00	196.49	23.30	1224.63
	1		ро 3		T •
Variabile	N	media	Dev Standard	Minimo	Massimo
POP11	90	12811.49	2224.95	9598.00	17883.00
FAM11	90	5028.78	910.49	3713.00	7358.00
ABIT11	90	5496.93	1041.00	3998.00	9371.00
IPR01_11	90	13.64	9.53	-5.70	41.65
IFM01_11	90	23.95	11.11	4.87	54.87
DENS11	90	506.59	275.35	39.15	1536.06
			ро 4		
Variabile	N	Media	Dev Standard	Minimo	Massimo
POP11	31	23076.32	4883.83	11793.00	32894.00
FAM11	31	9438.19	2066.37	5046.00	13094.00
ABIT11	31	11865.16	5009.28	7642.00	28695.00
IPR01_11	31	8.45	5.87	-2.27	22.55
IFM01_11	31	18.76	6.93	2.43	41.04
DENS11	31	644.21	381.47	76.66	1795.70
		Grup	ро 5		
Variabile	N	Media	Dev Standard	Minimo	Massimo
POP11	10	52374.50	24774.21	34428.00	111500.00
FAM11	10	22799.00	11441.90	14984.00	50017.00
ABIT11	10	24947.40	12426.79	16447.00	54417.00
IPR01_11	10	3.43	5.49	-3.95	14.76
IFM01_11	10	12.73	6.75	4.90	26.33
DENS11	10	716.38	440.27	241.75	1457.62
			ро 6		
Variabile	N	Media	Dev Standard	Minimo	Massimo
POP11	3	240024.67	29631.61	206192.00	261362.00
FAM11	3	111257.33	15217.93	94378.00	123928.00
ABIT11	3	118546.00	16661.52	100420.00	133194.00
IPR01_11	3	-1.07	2.22	-3.58	0.65
IFM01_11	3	6.75	1.64	5.17	8.45
DENS11	3	1371.47	798.93	628.43	2216.50
ווטוו	J	13/1.4/	/ 30.93	020.43	2210.30

