

## A. Appendice al Settore Residenziale: un esempio di cluster analysis

Il mercato immobiliare nazionale delle abitazioni risulta molto variegato e frammentato. Con questo studio si è provato ad individuare delle classi di comuni che presentassero andamenti simili rispetto ad alcune variabili cruciali che indicano lo 'stato' del mercato nell'attuale fase, quali le variazioni del volume di compravendite, il livello dei prezzi, la presenza di nuove costruzioni, la mobilità del mercato.

Il metodo di aggregazione usato è stato quello della 'cluster analysis', limitando, per il momento, tale analisi ai soli comuni capoluogo.

Le variabili prese in considerazione sono: il tasso medio semestrale 04-06 di variazione delle quotazioni medie comunali (TM quotazioni), tratte dalla banca dati OMI, il tasso medio semestrale 04-06 di variazione del NTN (TM NTN), la quotazione media comunale al I sem 04, il differenziale delle quotazioni al I semestre 2004 (riferito alla quotazione media minima), l'IMI media registrata nel periodo 2004-06, l'incidenza media di nuove costruzioni rispetto al volume di compravendite 2005-06 (INC/NC).

La matrice delle correlazioni pone in evidenza come alcune variabili presentano una correlazione significativa. Ad esempio, oltre alla ovvia correlazione tra differenziale dei valori ed i valori medesimi, esiste una correlazione positiva tra l'INC/NTN rispetto al TM NTN, il che vuol dire che le NC influenzano l'incremento delle compravendite, come era logico aspettarsi. Interessante è, invece, notare che la correlazione tra INC/NTN sia con l'IMI sia con le quotazioni risulti negativa. Questo fenomeno è collegato al fatto che le NC incidono maggiormente nei mercati delle città più piccole, con indice IMI e quotazioni meno elevate. Infatti, per un motivo analogo, si nota che anche tra IMI e TM quotazioni esiste una correlazione negativa.

Tabella 2-6: matrice delle correlazioni

		TM quotazioni	TM NTN	INC/NTN media05-06	IMI Media	val I 04	diff valori 104
TM quotazioni	Correlazione di Pearson	1	,024	-,104	-,177 *	-,186 *	-,186 *
	Sig. (1-coda)		,406	,152	,040	,033	,033
	N	99	99	99	99	99	99
TM NTN	Correlazione di Pearson	,024	1	,410 **	-,125	-,135	-,135
	Sig. (1-coda)	,406		,000	,109	,092	,092
	N	99	99	99	99	99	99
INC/NTN media05-06	Correlazione di Pearson	-,104	,410 **	1	-,268 **	-,399 **	-,399 **
	Sig. (1-coda)	,152	,000		,004	,000	,000
	N	99	99	99	99	99	99
IMI Media	Correlazione di Pearson	-,177 *	-,125	-,268 **	1	,255 **	,255 **
	Sig. (1-coda)	,040	,109	,004		,005	,005
	N	99	99	99	99	99	99
val I 04	Correlazione di Pearson	-,186 *	-,135	-,399 **	,255 **	1	1,000 **
	Sig. (1-coda)	,033	,092	,000	,005		,000
	N	99	99	99	99	99	99
diff valori 104	Correlazione di Pearson	-,186 *	-,135	-,399 **	,255 **	1,000 **	1
	Sig. (1-coda)	,033	,092	,000	,005	,000	
	N	99	99	99	99	99	99

\*. La correlazione è significativa al livello 0,05 (1-coda).

\*\* . La correlazione è significativa al livello 0,01 (1-coda).

Solitamente, uno dei presupposti perché una elaborazione cluster sia efficace è che le variabili risultino tra loro indipendenti. Nello specifico, pertanto, si possono eseguire cluster utilizzando separatamente le variabili IMI, INC/NTN e Differenziale o tasso medio delle quotazioni.

L' algoritmo utilizzato per eseguire la 'cluster analysis', tra i diversi possibili, è stato quello 'two step' fornito dal software SPSS, poiché è quello che sembra restituire delle aggregazioni mag-



giormente significative. Questo algoritmo può definire automaticamente il numero di classi in base alle caratteristiche delle variabili esaminate, utilizza il criterio bayesano di Schwarz per il raggruppamento con misura della distanza data dalla verosimiglianza.

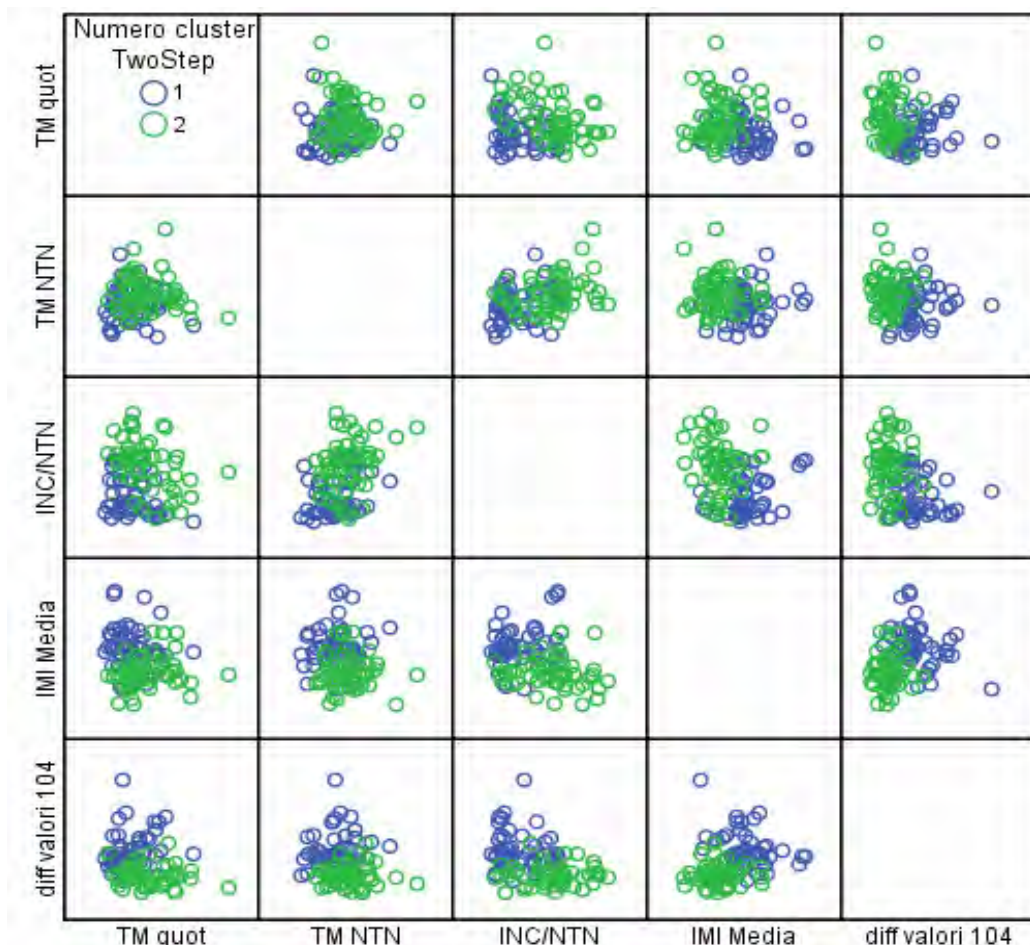
Sono state così utilizzate 5 delle sei variabili sopra indicate (è stata esclusa la quotazione media comunale al I semestre 2006 perché perfettamente analoga al differenziale delle quotazioni medesime), l'algoritmo propone una classificazione in due classi, di popolazione analoga, 48 e 51 città, con caratteristiche che, a livello medio, sono evidentemente differenti. La prima classe individua quei capoluoghi che hanno il TM quotazioni meno elevato, il TM NTN vicino a 0 e negativo, l'INC/NTN è relativamente bassa mentre sia l'IMI sia le quotazioni sono più elevati, Tabella 2-7. In questa classe ricadono capoluoghi del Nord e del Centro e solo due del Sud, Napoli e Salerno. Nella seconda classe, dove mediamente risultano più elevati il TM quotazioni ed NTN e l'INC/NTN, mentre sono più bassi l'IMI e le quotazioni, ricadono tutti gli altri capoluoghi del Sud, otto città del Nord e nove del Centro.

Tabella 2-7: centroidi delle 2 classi di capoluoghi derivanti da *cluster two step* con 5 variabili.

Cluster	TM quotazioni		TM NTN		INC/NTN		IMI Media		diff valori 104	
	Media	Dev. std.	Media	Dev. std.	Media	Dev. std.	Media	Dev. std.	Media	Dev. std.
1	0,025	0,024	- 0,008	0,055	0,164	0,093	0,031	0,005	3,012	0,891
2	0,045	0,031	0,028	0,057	0,311	0,123	0,024	0,004	1,807	0,481
Combinati	0,035	0,029	0,011	0,059	0,240	0,131	0,027	0,005	2,391	0,930

In realtà, poiché la variabilità dei valori è molto alta, come si evince dall'elevato valore della deviazione standard (vedi Tabella 2-7), l'assegnazione di una città ad una classe invece che all'altra può essere solo indicativo, soprattutto in quei casi in cui la distanza dai centroidi che individuano le due cluster è molto elevata.

Figura 2.60: matrice delle corrispondenze tra variabili per le due classi di capoluoghi.

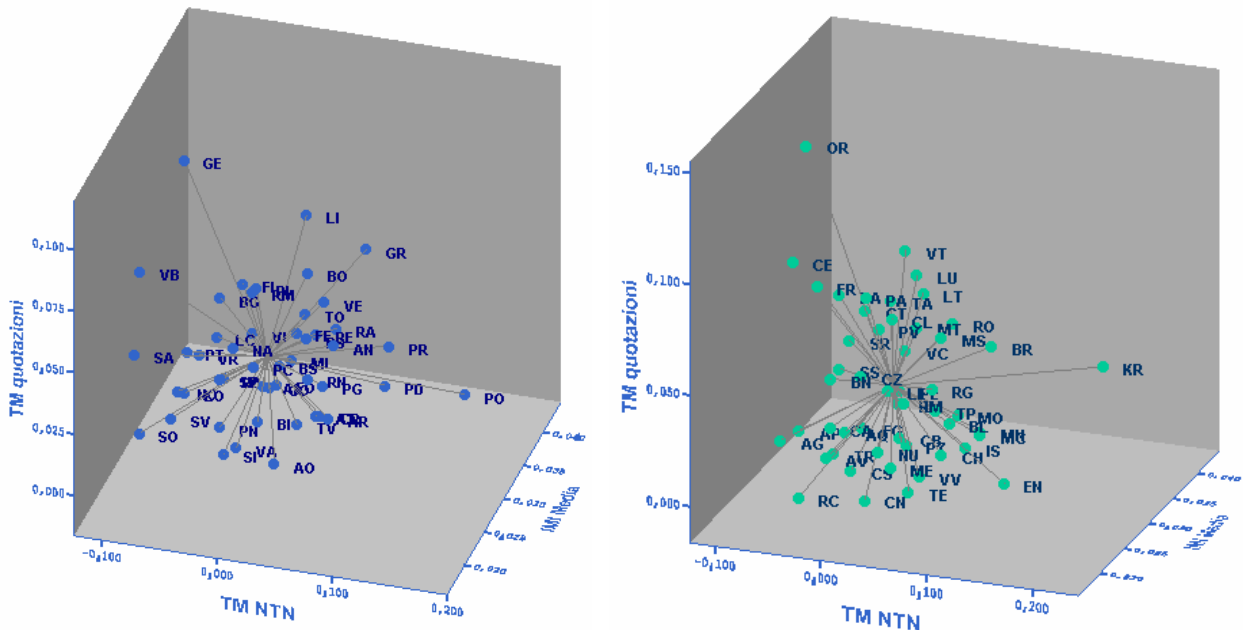




Anche osservando come si disperdono i valori considerando le variabili oggetto dell'analisi a due a due, Figura 2.60, si nota che per alcune coppie di variabili le due classi non si differenziano particolarmente

La figura seguente rappresenta le due classi di città separatamente, la classe uno in blu, la due in verde in relazione a tre variabili, TM quotazioni, TM NTN ed IMI media e risulta evidente oltre alla grande dispersioni dei valori, la caratterizzazione delle due classi tra Centro-Nord e Sud, a parte poche eccezioni.

**Figura 2.61: le due classi di città in relazione alle variabili TM quotazioni, TM NTN ed IMI media**



La suddivisione in queste due classi, pur risultando abbastanza indicativa in alcuni casi, individua, comunque, due tipologie di mercato caratterizzate da una parte (classe 1), da una stabilità o decremento del volume di compravendite collegato ad un incremento dei prezzi meno accentuato, dovuto anche all'elevato livello raggiunto dalle quotazioni, con scarsità di nuove costruzioni, ma, comunque, con un'elevata mobilità di mercato (IMI alto), e dall'altra (classe 2) da una crescita dei prezzi e dei volumi di compravendita, superiore alla media. La prima collegata anche al basso livello delle quotazioni, la seconda aiutata da una maggiore presenza di nuove costruzioni. L'IMI di questa classe in cui ricadono molte città del Sud è invece mediamente minore alla media nazionale.

Poiché le uniche variabili, tra quelle scelte, tra loro completamente indipendenti sono il TM quotazioni ed il TM NTN si è verificato come cambiasse la classificazione dei capoluoghi, utilizzando nuovamente la cluster two step, con le variabili suddette. La soluzione proposta dall' algoritmo di calcolo suddivide le città in tre classi, di cui fanno parte rispettivamente 34, 23 e 42 capoluoghi, caratterizzati, per la prima classe, da quotazioni in crescita moderata (tasso medio pari a 1,8% mediamente) ed NTN in sensibile calo (tasso medio pari a -3,2%), per la seconda da forte crescita delle quotazioni (tasso medio +7,6% mediamente) e stabilità delle compravendite (tasso medio 0,4%), per la terza da un' inflazione dei prezzi abbastanza contenuta (tasso medio 2,8%) e da una crescita consistente delle compravendite (tasso medio +6,9%).

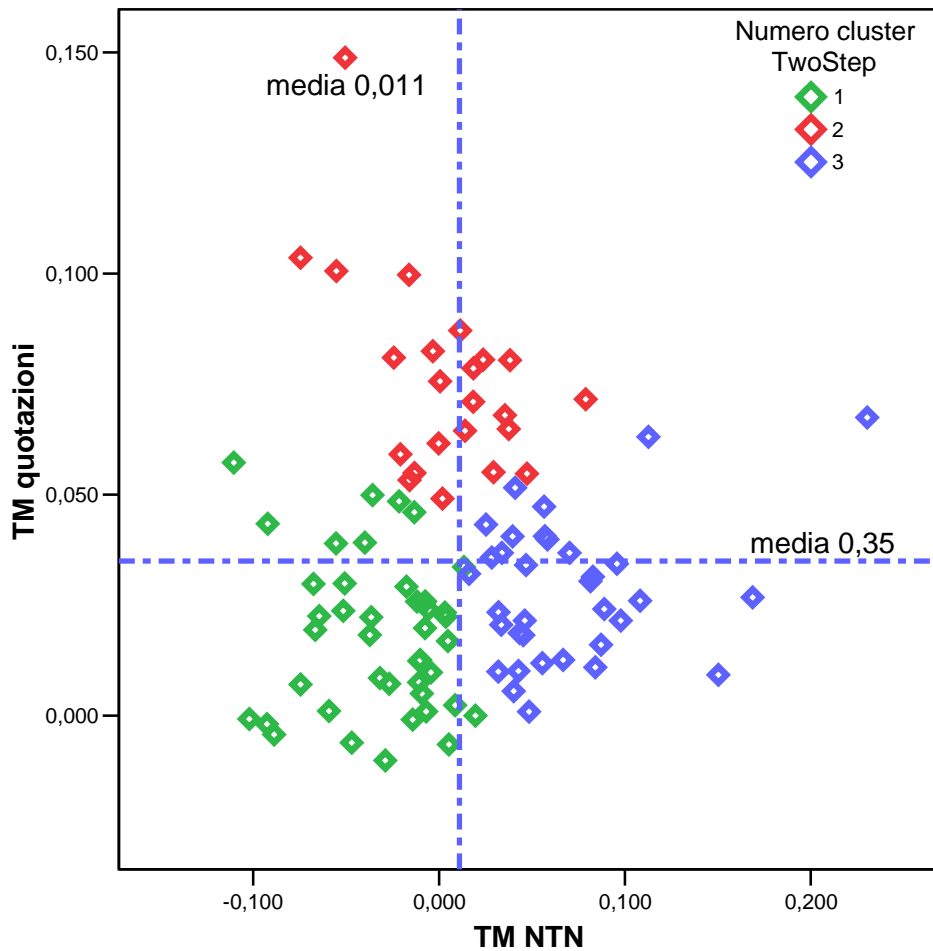
Anche questa classificazione, pur facendo riferimento a due sole variabili, individua dei gruppi di città che hanno una grande dispersione dei valori al loro interno, ed è pertanto possibile che in alcuni casi l'assegnazione ad una classe piuttosto che ad un'altra risulti solamente indicativa.



Tabella 2-8: centroidi delle 3 classi di capoluoghi derivanti da *cluster two step* con 2 variabili

Centroidi	TM quotazioni		TM NTN	
	Media	Deviazione std.	Media	Deviazion e std.
1	1,8%	1,7%	-3,2%	3,4%
2	7,6%	2,2%	0,4%	3,5%
3	2,8%	1,6%	6,9%	4,5%
Combinati	3,5%	2,9%	1,1%	5,9%

Figura 2.62: individuazione delle tre classi di capoluoghi per TM quotazioni e TM NTN



Il grafico della Figura 2.62 mostra chiaramente la distribuzione delle tre classi, di cui la prima individua le città con bassa inflazione dei prezzi a cui non segue un incremento significativo delle compravendite, la seconda individua quei casi in cui l'inflazione dei prezzi è elevata, con compravendite stabili, la terza individua la classe 'più ragionevole', ovvero quella dove ad un aumento contenuto dei prezzi segue un aumento consistente delle compravendite.

Dai grafici della Figura 2.63 che distinguono le tre classi di comuni, si nota, come la distribuzione in queste classi risulti piuttosto diversa dalla precedente. Infatti, in questo caso, le città si spartiscono molto più uniformemente senza una netta distinzione geografica, dettata, nell'analisi precedente dalla marcata difformità sia dei differenziali delle quotazioni sia dell'IMI (entrambi mediamente più bassi nel Sud), che portava a distinguere le città del Sud da quelle del Centro-Nord.



Figura 2.63: tre classi di capoluoghi per TM quotazioni e TM NTN

