



Note di stabilità finanziaria e vigilanza

N. 33

Novembre 2023

Sintesi	1
1. Introduzione	1
2. I dati e la metodologia di stima	2
3. Il metodo K-means	3
4. Il <i>clustering</i> gerarchico	5
5. Conclusioni	6

*I contributi pubblicati nella serie
"Note di stabilità finanziaria
e vigilanza" riflettono le opinioni
degli autori e non impegnano
la responsabilità della
Banca d'Italia*

Banche a rilevanza sistemica nazionale: analisi dei cluster per l'Italia

F. Francetti – M. Galardo*

Sintesi

La nota presenta esercizi di *cluster analysis* volti a identificare il numero ottimale di *cluster* per classificare le banche italiane in gruppi di rilevanza sistemica nazionale (O-SII) omogenei. La *cluster analysis* è una tecnica di statistica descrittiva ampiamente utilizzata a supporto della politica macroprudenziale, anche a livello internazionale: il Comitato di Basilea e il Financial Stability Board hanno sfruttato il *clustering* per identificare il gruppo delle istituzioni a rilevanza sistemica globale; la BCE ha applicato la *cluster analysis* per definire un livello minimo dei buffer per le banche identificate come O-SII dalle autorità macroprudenziali nazionali dei paesi facenti parte dell'SSM (*Single Supervisory Mechanism*).

1. Introduzione

L'analisi dei gruppi (*cluster analysis*) è una tecnica di statistica descrittiva ampiamente utilizzata per classificare i dati e individuare insiemi di osservazioni con caratteristiche simili¹. Queste tecniche sono state utilizzate dal Comitato di Basilea (BCBS) e dal Financial Stability Board (FSB) per identificare il gruppo delle istituzioni a rilevanza sistemica globale (G-SIBs) e definire quindi una soglia di rilevanza². Anche la BCE ha

¹ Si veda: Kaufman, L., and P. J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*, New York: Wiley (1990); Johnson, Richard Arnold, and Dean W. Wichern, *Applied multivariate statistical analysis* (2002); Rui Xu, Donald C. Wunsch II, *Clustering* (2008); Härdle, Wolfgang Karl, and Léopold Simar, *Applied multivariate statistical analysis* (2019).

² BCBS, *Global systemically important banks: assessment methodology and the additional loss absorbency requirement*, 2011.

* F. Francetti, M. Galardo, Dipartimento Economia e statistica.

applicato una *cluster analysis* ai dati di bilancio di un campione di banche appartenenti all'SSM (*Single Supervisory Mechanism*) per definire la struttura a scaglioni (*bucket*) utilizzata per associare un livello minimo di buffer (ECB O-SII *buffer floor*) al punteggio di identificazione degli intermediari a rilevanza sistemica domestica (O-SII)³. Questa nota presenta una serie di esercizi di analisi dei gruppi applicata ai dati di bilancio delle banche italiane per identificare il numero ottimale di *cluster* per il sistema bancario italiano e definire gruppi omogenei di intermediari per rilevanza sistemica nazionale. I punteggi di rilevanza sistemica domestica (*O-SII scores*) degli intermediari bancari italiani sono stati calcolati seguendo le indicazioni contenute negli orientamenti EBA (EBA/GL/2014/10).

Dall'analisi emergono i seguenti risultati principali:

- **Nel sistema bancario italiano esistono cinque classi di intermediari con caratteristiche simili:** una raggruppa quelli più piccoli; una gli intermediari più grandi con punteggio da 254 a 366; una include un solo intermediario (con punteggio 511); due classi comprendono gli intermediari con punteggi superiori a 2.700.
- **Per separare le banche a rilevanza sistemica nazionale dalle altre, la *cluster analysis* suggerirebbe una soglia intorno a 250 punti**, più bassa della soglia di identificazione standard definita dall'EBA (pari a 350 punti).

2. I dati e la metodologia di stima

Il calcolo degli *O-SII scores* per l'analisi dei *cluster* è stato effettuato per ciascuna banca o gruppo bancario italiano⁴, seguendo le indicazioni contenute negli orientamenti EBA per la valutazione degli enti a rilevanza sistemica (EBA/GL/2014/10). In particolare, gli orientamenti prevedono che gli intermediari vengano valutati secondo quattro profili (dimensione, importanza per l'economia domestica, complessità e interconnessione con il sistema finanziario) quantificati mediante 10 indicatori (Tavola 1); per il sistema bancario italiano, i punteggi degli indicatori – e, di conseguenza, dei singoli profili di rilevanza – sono calcolati sulla base delle segnalazioni di vigilanza armonizzate e dei dati Target2 riferiti alla fine di ogni anno.

I punteggi così calcolati – che descrivono sinteticamente le caratteristiche di ciascun intermediario bancario operante nel sistema italiano – sono stati quindi impiegati nell'analisi dei *cluster* secondo due metodologie: l'algoritmo K-means e il *clustering* gerarchico basato sull'approccio bottom-up.

³ ECB, [Macroprudential Bulletin, Issue n. 3](#), giugno 2017. Alla fine dello scorso anno, la BCE ha pubblicato un aggiornamento dei floor volto a ridurre l'elevata eterogeneità dei buffer fissati nei diversi paesi europei: ECB, [Governing Council statement on macroprudential policies](#), dicembre 2022.

⁴ Il calcolo viene effettuato al più alto livello di consolidamento. Secondo le Linee Guida EBA sulla valutazione delle O-SII "le autorità pertinenti dovrebbero calcolare un punteggio per ogni entità pertinente almeno al più alto livello di consolidamento della parte del gruppo che rientra sotto la loro giurisdizione", (EBA/GL/2014/10, Titolo II, paragrafo 5).

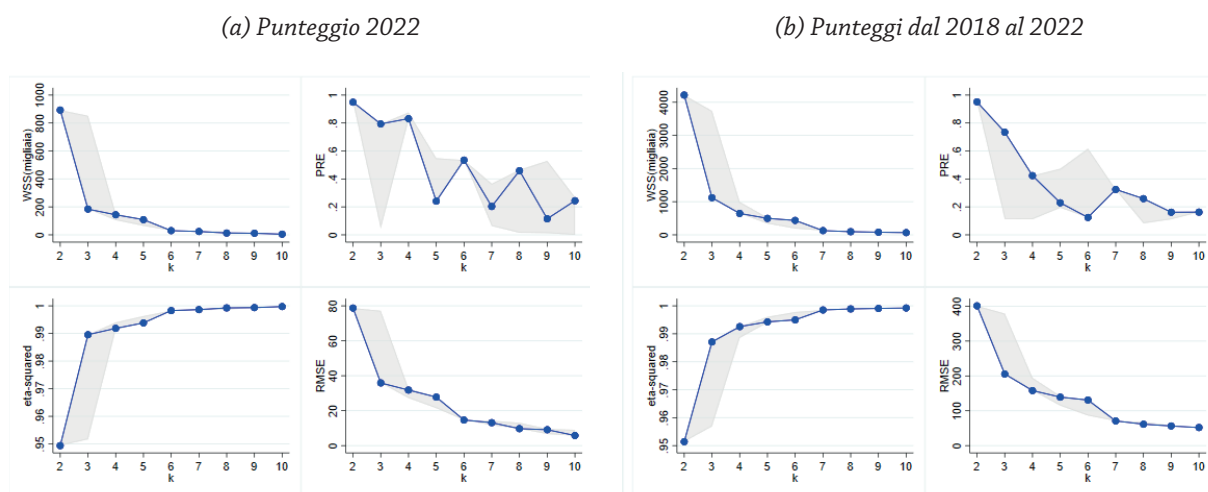
Tav. 1 – Metodologia EBA: indicatori obbligatori per l'assegnazione del punteggio

Criterio	Indicatori	Ponderazione
Dimensione	Totale attività	25,00%
Importanza (compresa sostituibilità/ infrastruttura del sistema finanziario)	Valore delle operazioni di pagamento nazionali	8,33%
	Depositi del settore privato da parte di depositanti nell'UE	8,33%
	Prestiti del settore privato in favore di destinatari nell'UE	8,33%
Complessità/attività transfrontaliera	Valore dei derivati OTC (nozionale)	8,33%
	Passività inter-giurisdizionali	8,33%
	Crediti inter-giurisdizionali	8,33%
Interconnessione	Passività verso altri enti finanziari	8,33%
	Attività verso altri enti finanziari	8,33%
	Titoli di debito in circolazione	8,33%

3. Il metodo K-means

L'algoritmo K-means richiede di impostare il numero dei *clusters* (k) a priori, l'esercizio è stato ripetuto ipotizzando un numero di *cluster* da 2 a 10 e confrontando diverse misure della qualità dei *clusters* per identificarne il numero ottimale. Vengono utilizzate quattro misure: (i) la devianza interna (*Within Sum of Square*, WSS), ossia la somma degli scarti di ciascun punteggio dalla media del proprio gruppo; (ii) la riduzione della WSS per l'analisi che assume k gruppi rispetto all'analisi basata su $(k-1)$ gruppi (*Proportional Reduction in Error*, PRE); (iii) l'indice eta-quadro dato dal rapporto tra la devianza esterna, ossia la somma degli scarti delle medie di categoria dalla media generale, e la devianza totale, ossia la somma dello scarto di ciascun punteggio dalla media generale; (iv) la radice quadrata della media degli errori al quadrato (*Root Mean Square Error*, RMSE). Inoltre, poiché i risultati dell'algoritmo K-means potrebbero dipendere dalla scelta del punto di inizializzazione (il centro del *cluster* iniziale), per ogni k l'esercizio è stato replicato 500 volte selezionando casualmente il punto di partenza e riportando come risultato quello che si verifica con la maggiore frequenza unitamente alle bande che identificano l'intervallo tra i risultati minimo e massimo. Le analisi sono state condotte sui punteggi calcolati in base ai dati dei bilanci dal 2018 al 2022. Le evidenze sono riportate nella Figura 1. Nel pannello (a) si presentano gli indicatori ottenuti considerando i punteggi del 2022 e nel pannello (b) quelli ottenuti considerando tutti i punteggi dal 2018 al 2022. In entrambi gli esercizi la WSS ha un forte calo per $k=3$, continua a scendere fino a $k=6$ per poi rimanere pressoché stabile. L'indicatore eta-quadro segnala una riduzione della WSS di oltre il 99 per cento già per $k=4$. Per l'esercizio che assume 4 e 5 gruppi, PRE segnala riduzioni tra l'80 e il 20 per cento, l'indicatore risale da $k=6$ ma la riduzione della devianza interna (WSS) è limitata per valori di k superiori a 5. Gli indicatori suggeriscono quindi un numero di *clusters* tra 3 e 5.

Fig. 1 - Indicatori di qualità dei clusters ottenuti con l'algoritmo K-means (1)



Fonte: Elaborazioni Banca d'Italia.

(1) Il campione include tutti gli intermediari bancari che segnalano il totale attivo a ciascuna data di rilevazione. L'algoritmo K-means utilizza come misura della dispersione del cluster la distanza euclidea.

La Tavola 2 riporta i risultati dell'analisi assumendo un numero di *clusters* da $k=3$ a $k=5$. I due intermediari maggiori vengono sempre assegnati a due gruppi diversi; allo stesso modo, si identifica chiaramente un gruppo di intermediari (area verde) più numeroso che viene confermato in tutte le simulazioni.

Tav. 2 - Clustering in base ai punteggi del 2022
(punti base)

Cluster	K=3				K=4				K=5			
	N	media	min	Max	N	Media	min	max	N	media	min	max
1	1	3085	3085	3085	1	3085	3085	3085	1	3085	3085	3085
2	1	2800	2800	2800	1	2800	2800	2800	1	2800	2800	2800
3	144	29	0	511	7	339	254	511	1	511	511	511
4					137	13	0	155	6	310	254	366
5									137	13	0	155

L'analisi che assume $k=3$ cattura solo in parte l'eterogeneità del sistema bancario italiano, tutti gli intermediari diversi dai due maggiori vengono infatti classificati in un unico gruppo. Aumentando il numero dei cluster la rappresentazione diventa più accurata e si isola il gruppo degli intermediari minori (con un punteggio compreso tra 0 e 155) da quello che, per $k=4$, raccoglie gli enti con punteggi compresi tra 254 e 511 punti base. Per $k=5$ si ha un'ulteriore separazione con la creazione di un gruppo a cui viene assegnato solo l'intermediario con score pari a 511 punti base che dista oltre 200 punti dalla media della classe subito inferiore (310 punti base). Per riflettere in maniera adeguata l'eterogeneità del sistema bancario italiano è consigliabile un *clustering* a 5 gruppi.

Per verificare la robustezza dei risultati mostrati nella Tavola 2 l'esercizio è stato ripetuto per gli anni dal 2018 al 2022. La Tavola 3 mostra che nel 2021 e 2022 si forma un nuovo *cluster* (area arancione) con un unico intermediario il cui punteggio aumenta rispetto agli anni precedenti, i due *cluster* inferiori si compattano in un unico gruppo di intermediari più piccoli (area bianca) e resta la classe intermedia a cui si associa un punteggio medio di 310 punti base (area gialla). I due intermediari maggiori vengono sempre identificati in due *cluster* diversi (area verde e azzurra).

Tav. 3 – Clustering per k=5 in base ai punteggi dal 2018 al 2022
(punti base)

Cluster	2018				2019				2020			
	N	Media	Min	Max	N	Media	Min	Max	N	Media	Min	Max
1	1	3053	3053	3053	1	3199	3199	3199	1	3199	3199	3199
2	1	2633	2633	2633	1	2557	2557	2557	1	2776	2776	2776
3	5	346	255	469	2	420	383	457	5	349	289	458
4	10	115	74	201	6	235	141	318	11	122	77	231
5	124	6	0	55	131	11	0	105	124	6	0	59

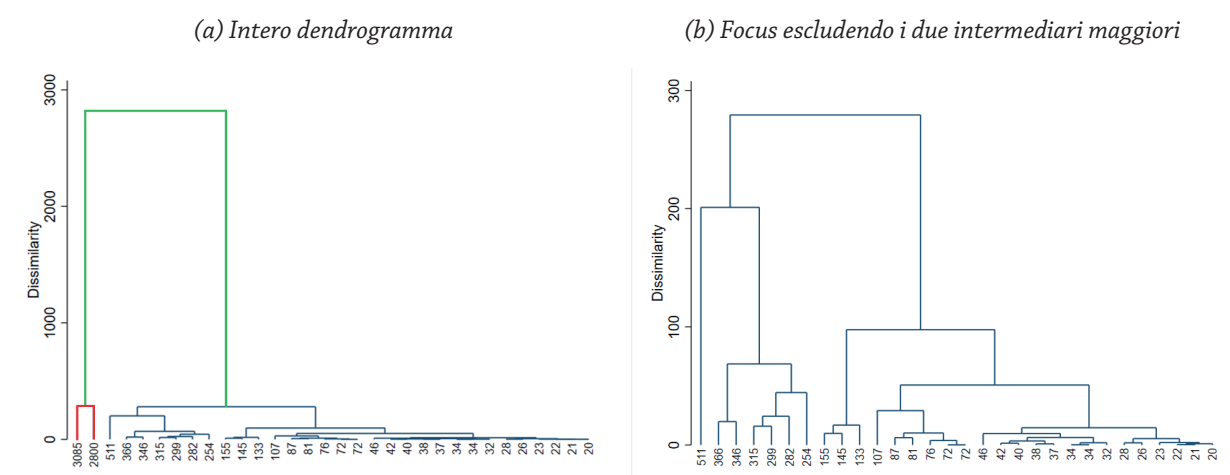
Cluster	2021				2022			
	N	Media	Min	Max	N	Media	Min	Max
1	1	3124	3124	3124	1	3085	3085	3085
2	1	2806	2806	2806	1	2800	2800	2800
3	1	484	484	484	1	511	511	511
4	6	310	245	332	6	310	254	366
5	137	13	0	148	137	13	0	155

4. Il clustering gerarchico

L'analisi dei *cluster* secondo il metodo gerarchico non richiede di scegliere il numero dei gruppi prima di classificare i dati. L'algoritmo parte considerando tutti gli intermediari separati in gruppi individuali, e quindi procede iterativamente con l'aggregazione di coppie di gruppi meno distanti tra loro, arrivando al termine a produrre un unico gruppo contenente tutte le unità analizzate. Il risultato non è una sola partizione dei dati, ma una serie di ripartizioni. La Figura 2 riporta i dendrogrammi ottenuti considerando i punteggi del 2022. Il *clustering* gerarchico, come l'algoritmo K-means, evidenzia chiaramente una prima forte distinzione tra i due intermediari maggiori e il resto del sistema, identificata nella Figura 2.a dai due rami verdi più lunghi a cui corrisponde la maggiore dissimilarità; gli altri due rami più lunghi del dendrogramma sono quelli rossi che collocano i due intermediari più grandi in due *cluster* separati. Per avere una visione

più chiara delle ramificazioni per il resto del sistema bancario italiano, i due intermediari maggiori sono stati esclusi dall'analisi nella Figura 2.b. La prima distinzione (ramo più lungo) riguarda gli intermediari con punteggio inferiore o uguale a 155 punti base; nel gruppo di quelli rimanenti (punteggio sopra i 254 punti base) si distingue l'estremo superiore a cui corrisponde un punteggio pari a 511 e restano appaiati in un unico gruppo gli intermediari con punteggio da 254 a 366 punti base. Il *clustering* gerarchico, in linea con l'algoritmo K-means, suggerisce un numero di *clusters* pari a 5, dove i tre punteggi più alti vengono assegnati a tre gruppi diversi, i punteggi da 254 a 366 a un unico gruppo e un ultimo *cluster* più numeroso di banche con punteggi più bassi.

Fig. 2 - Clustering gerarchico in base ai punteggi del 2022 (1)
(punti base)



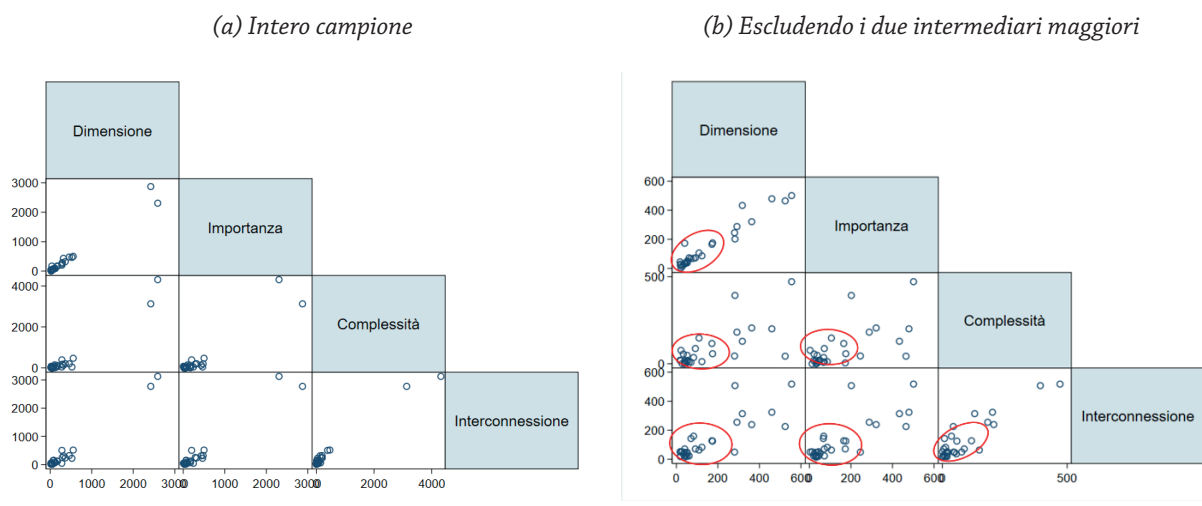
(1) Il campione include solo gli intermediari con un punteggio superiore ai 20 punti base. L'algoritmo associa i gruppi di osservazioni minimizzando l'aumento della somma degli errori al quadrato (metodo di Ward).

Gli orientamenti EBA prevedono che i punteggi complessivi siano calcolati come somma pesata dei punteggi che ogni intermediario ottiene in quattro categorie: dimensione, importanza, complessità e interconnessione (cfr. Tavola 1). La Figura 3 riporta le matrici dei grafici a dispersione dei punteggi per le quattro categorie. Anche in questo caso, i due intermediari più grandi si distinguono chiaramente dal resto del sistema per tutte le categorie (Figura 3.a) e per osservare più chiaramente la dispersione per il resto del sistema (Figura 3.b) è necessario escluderli dalla matrice. Il *cluster* di intermediari minori è visibile per tutte le quattro categorie (area cerchiata in rosso) mentre quelli delle banche con punteggi intermedi mostrano una maggiore dispersione nelle categorie importanza, complessità e interconnessione.

5. Conclusioni

Le analisi dei *cluster* suggeriscono per il sistema bancario italiano l'esistenza di cinque classi di intermediari con caratteristiche simili: una raggruppa quelli più piccoli; una gli intermediari più grandi, con punteggio da 254 a 366; una include un solo intermediario (con punteggio 511); due classi comprendono gli intermediari con punteggi superiori a 2.700.

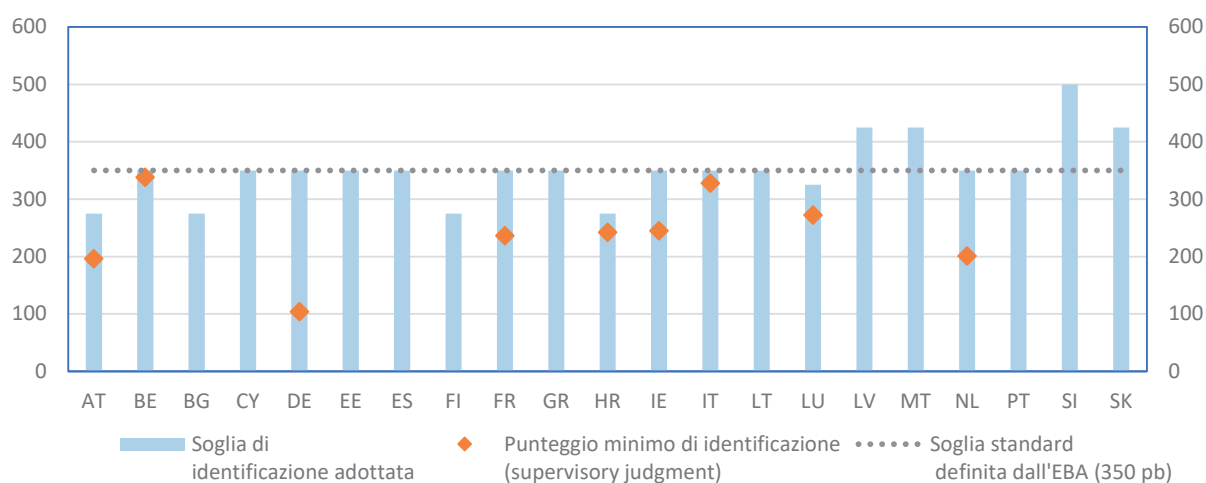
Fig. 3 - Matrici di dispersione in base ai punteggi del 2022 (1)



(1) Il campione include solo gli intermediari con un punteggio superiore ai 20 punti base.

Per separare le banche a rilevanza sistemica nazionale dalle altre, l'analisi dei *cluster* suggerirebbe una soglia intorno ai 250 punti base considerando il valore minimo e intorno a 310 punti considerando la media. Gli orientamenti EBA raccomandano per l'identificazione delle istituzioni a rilevanza sistemica domestica la soglia di 350 punti base che può essere aumentata o diminuita per un massimo di 75 punti base (i.e. fino a 425 o a 275 punti base). Alcuni dei paesi partecipanti al Meccanismo di vigilanza unico (*Single Supervisory Mechanism, SSM*) utilizzano soglie inferiori ai 350 punti base. In particolare: Austria, Belgio e Croazia hanno fissato la soglia a 275 punti base; il Lussemburgo a 325 punti base (Figura 4). Alcuni paesi hanno invece esercitato il *supervisory judgment* per ampliare la platea delle banche identificate come O-SII; in particolare, Belgio, Germania, Francia, Irlanda e Olanda. Austria, Croazia e Lussemburgo combinano la riduzione della soglia con l'esercizio del *supervisory judgment*. L'utilizzo di quest'ultimo ha portato, in alcuni casi, a identificare come O-SII anche intermediari con punteggi sotto la soglia minima prevista dall'EBA (275 punti base).

Fig. 4 - Soglie di identificazione adottate dai paesi SSM



Fonte: Elaborazioni Banca d'Italia su dati EBA (cfr. EBA, *List of O-SIIs notified to the EBA by year, 2022*) e ESRB (cfr. ESRB, *Systemically important institutions*).