



Metodi e fonti: note metodologiche

9 gennaio 2024

Per informazioni: statistiche@bancaditalia.it
www.bancaditalia.it/statistiche/index.html

Indagini sull'alfabetizzazione finanziaria e le competenze di finanza digitale in Italia: giovani

Aspetti generali

La Banca d'Italia ha condotto nel primo trimestre del 2023 un'indagine sull'alfabetizzazione finanziaria e le competenze di finanza digitale dei giovani, compresi tra i 18 e 34 anni. Si rilevano la comprensione di concetti economico-finanziari di base, i comportamenti tenuti degli intervistati in campo finanziario e i servizi finanziari utilizzati, da quelli tradizionali a quelli più complessi. Nel questionario sono presenti domande utili a cogliere opinioni e aspirazioni dei giovani. Una sezione approfondisce l'interesse degli intervistati verso i temi economico-finanziari.

Questa nota fornisce una sintesi delle principali caratteristiche metodologiche dell'indagine. Si discutono l'universo di riferimento e le caratteristiche del campione e del questionario sottoposto agli intervistati; si illustrano quindi le modalità di raccolta dei dati e, infine, le analisi condotte a presidio della qualità e della rappresentatività delle informazioni raccolte.

L'universo di riferimento e il campione

L'universo di riferimento dell'indagine include gli individui di età tra 18 e 34 anni residenti in Italia. Il disegno di campionamento prevede una selezione *per quote* che rispetta le proporzioni della popolazione ripartita per età e area geografica di residenza degli individui. Le informazioni sulle caratteristiche dell'universo di riferimento sono di fonte ISTAT. Il disegno campionario considera 3 classi di età (18-23, 24-29 e 30-34 anni) e 5 aree geografiche (Nord-ovest, Nord-est, Centro, Sud e Isole). Il prodotto cartesiano delle classi di età e delle aree geografiche definisce il numero di classi (15) in cui le unità statistiche sono allocate, in proporzione all'importanza di ciascuna classe nella popolazione.

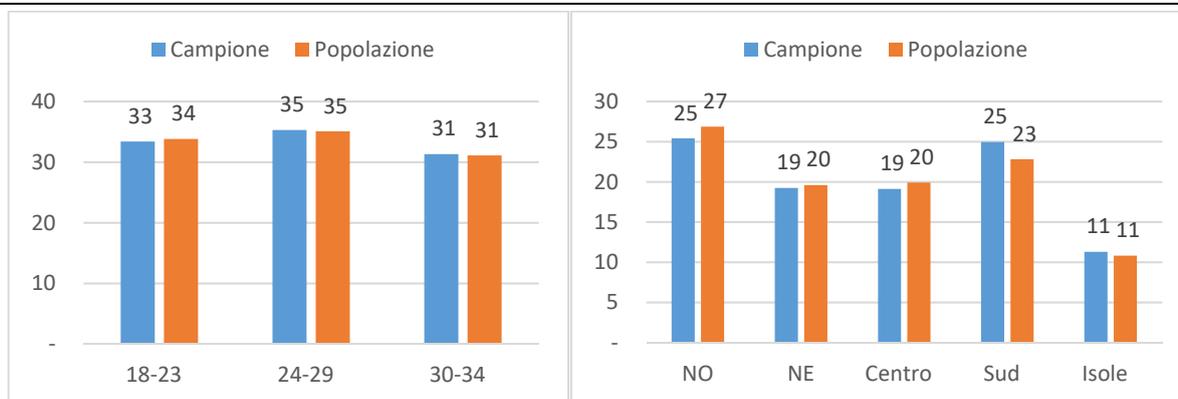
Il campione di individui intervistati comprende oltre 5000 giovani. Per ottenerlo sono stati contattati circa 42.000 individui e il tasso di risposta è pari al 12,7%¹. Le mancate risposte sono state più alte tra i meno istruiti e tra le fasce più giovani dell'universo di riferimento dell'indagine.

¹ I tassi di risposta delle indagini condotte online sono mediamente bassi. Per una discussione si veda Daikeler et al. (2020), "Web Versus Other Survey Modes: An Updated and Extended Meta-Analysis Comparing Response Rates", [Journal of Survey Statistics and Methodology](#).

Nella Figura 1 sono illustrate le composizioni percentuali della popolazione e del campione per età e area geografica di residenza.

Figura 1

Il campione e la popolazione di riferimento: composizione per età e residenza degli intervistati (valori percentuali)



Il questionario e la rilevazione

Il questionario include sezioni che rilevano informazioni anagrafiche e dell'intervistato, conoscenze e comportamenti in campo finanziario e servizi finanziari utilizzati, anche nell'ambiente finanziario digitale; altre si concentrano su opinioni, aspirazioni e interesse verso i temi economico-finanziari.

Versioni provvisorie del questionario sono state sottoposte a *pre-testing* (*expert review* e *respondent debriefing*) all'interno della Banca d'Italia. Affinamenti ulteriori hanno recepito i risultati di un esercizio pilota condotto su un campione di 50 unità dalla società che ha raccolto i dati.

Le informazioni sono state rilevate con la tecnica *Computer Assisted Web Interview* (CAWI). L'applicativo per la compilazione online del questionario è stato predisposto per dispositivi PC e laptop, tablet e mobile. I partecipanti all'indagine sono stati selezionati tra gli individui appartenenti al campione (*opt-in*) a disposizione della società, che ha messo a disposizione dei partecipanti un numero verde per eventuali richieste di assistenza.

Il tempo medio di compilazione del questionario è stato pari a 18 minuti. Nel 73 per cento dei casi è stato utilizzato uno smartphone per la compilazione; pc e tablet sono stati utilizzati nel 22 e nel 5 per cento dei casi, rispettivamente.

La qualità dei dati

Sono state condotte analisi sulla qualità delle informazioni raccolte. Le analisi hanno tenuto conto, tra l'altro, del tempo impiegato per la compilazione del questionario e della coerenza concettuale tra risposte formulate in parti diverse dell'indagine. Sono stati esclusi dai risultati i questionari la cui durata di compilazione risultava non verosimile e quelli che presentavano incoerenze che impedivano l'interpretazione dei dati. È risultato generalmente basso il tempo di compilazione dei questionari che hanno evidenziato casi di incoerenze concettuali. In alcuni casi la rilevazione ammetteva risposte in forma "aperta", formulabili alternativamente alle opzioni in forma "chiusa" (benchmark). Un'analisi del testo ha permesso di riclassificare – e utilizzare - le risposte rese in forma "aperta".

Nelle indagini campionarie, le informazioni che rilevano abitudini, comportamenti, opinioni degli intervistati possono essere esposte a rischi di distorsione legati a bias di memoria (*memory bias*) o ad atteggiamenti che favoriscono risposte ritenute dall'intervistato socialmente accettabili (*desiderability bias*). La tecnica Computer Assisted Web Interview (CAWI), utilizzata in questa indagine, mitiga questi rischi, principalmente a causa dell'assenza del ruolo dell'intervistatore nel processo di raccolta dei dati e della flessibilità nei tempi di compilazione².

Ponderazioni e stimatori

Per il riporto all'universo dei dati campionari sono stati applicati metodi di *post-stratificazione* e *raking* statistico. La post-stratificazione mitiga eventuali distorsioni legate alla mancata risposta degli individui selezionati per la partecipazione all'indagine³; inoltre, affina la distribuzione delle unità statistiche per tener conto delle eterogeneità che possono esservi lungo dimensioni rilevanti. Per questa indagine sono stati definiti 90 *post-strati*, ottenuti dal prodotto cartesiano tra: 15 classi definite in base all'area geografica di residenza e la classe di età, che è una delle variabili che rilevano nell'analisi delle mancate risposte; 2 classi definite in base al genere (uomo, donna); 3 classi definite in base alla dimensione del comune di residenza (fino a 30 mila; da 30 a 100 mila; oltre 100 mila abitanti). All'interno di ogni *post-strato*, i pesi sono definiti dalla formula:

$$w_{ij} = \frac{N_i}{Nn_i}$$

dove w_{ij} è il peso attribuito all'unità j nell' i -esimo *post-strato*; N è la numerosità della popolazione di riferimento; n_i il numero di interviste valide nel *post-strato* i -esimo e N_i la numerosità della popolazione nel *post-strato* i -esimo.

La tecnica del *raking* statistico è stata infine applicata per accrescere la rappresentatività del campione anche in base al titolo di studio, che costituisce un'ulteriore variabile rilevante nell'analisi delle mancate risposte. Sono considerati 3 livelli di istruzione: diploma di scuola secondaria di 1° grado o titolo inferiore; diploma di scuola secondaria di 2° grado; laurea⁴.

Per le variabili obiettivo si è utilizzato di stimatore di Horvitz-Thompson, definito a partire dai pesi campionari ricalibrati con le tecniche descritte. Per la stima della varianza degli stimatori è utilizzata l'espressione linearizzata in serie di Taylor. Analisi condotte tramite metodi di ricampionamento (*bootstrap*) restituiscono risultati simili.

² Cfr. Kreuter et al., (2008), "Social Desirability Bias in CATI, IVR, and Web Surveys: The Effects of Mode and Question Sensitivity", [Public Opinion Quarterly](#).

³ Le mancate risposte potrebbero comportare distorsioni nei risultati delle indagini campionarie se segmenti di popolazione meno disposti a partecipare sono sottorappresentati; l'entità della distorsione può dipendere dal divario tra il valore atteso della variabile misurata per i partecipanti e i non partecipanti all'indagine.

⁴ In base a tali livelli, le distribuzioni 'marginali' del campione sono allineate a quelle della popolazione di riferimento. La calibrazione dei pesi finali è eseguita allineando consecutivamente le distribuzioni marginali del campione lungo ognuna delle dimensioni considerate (area geografica classe di età, genere, dimensione del comune di residenza e titolo di studio).