

## Nota metodologica

### Gli obiettivi conoscitivi dell'indagine

L'indagine "Bambini e ragazzi: comportamenti, atteggiamenti e progetti futuri" è condotta dall'Istituto nazionale di statistica con l'obiettivo di raccogliere informazioni su alcuni aspetti fondamentali della vita quotidiana di ragazzi e ragazze tra gli 11 e i 19 anni residenti in Italia. Specifica attenzione viene dedicata ai ragazzi di cittadinanza straniera.

La rilevazione è inserita nel Programma statistico nazionale ed è stata attentamente seguita nella sua progettazione dall'Autorità garante per la protezione dei dati personali.

Ai giovanissimi che rientrano nel campione viene chiesto di compilare un breve questionario *online* accessibile anche attraverso *smartphone*. Attraverso domande semplici vengono raccolte informazioni sulle relazioni con gli amici e con la famiglia, sull'utilizzo dei *social media*, sulla povertà educativa, sulla cittadinanza e il senso di appartenenza e sui progetti futuri delle nuove generazioni.

### Definizione e target di indagine

All'indagine sono stati chiamati a rispondere circa 108mila ragazzi e ragazze - italiani e stranieri - tra gli 11 e i 19 anni residenti in Italia. I bambini e i ragazzi sono stati estratti casualmente dagli archivi sulla popolazione dell'Istat. La strategia di campionamento ha consentito di costruire un campione statisticamente rappresentativo della popolazione tra gli 11 e i 19 anni, sia di cittadinanza italiana sia di cittadinanza straniera (con particolare attenzione per le cittadinanze più numerose). Hanno risposto 39.214 giovani.

### La durata della rilevazione

La rilevazione si è svolta dal 1° ottobre al 20 dicembre 2023.

### Come sono stati raccolti i dati

I dati sono stati raccolti esclusivamente tramite un questionario *online*. I ragazzi e le loro famiglie hanno ricevuto, via posta presso il loro indirizzo di residenza, una lettera informativa a firma del Presidente dell'Istat contenente indicazioni sull'indagine e le informazioni necessarie per accedere al questionario. La lettera è stata indirizzata direttamente ai ragazzi, se maggiorenni, alla famiglia se minorenni. Le lettere rivolte ai ragazzi erano disponibili in 10 lingue oltre l'italiano: albanese, arabo, cinese, francese, inglese, romeno, russo, sloveno, spagnolo e tedesco.

Il questionario poteva essere compilato anche attraverso *smartphone* ed era disponibile per la compilazione nelle seguenti lingue: albanese, arabo, cinese, francese, inglese, romeno, spagnolo, tedesco, ucraino. Era possibile accedere al questionario direttamente attraverso un QR code riportato sulla comunicazione dell'Istat: ovviamente ciò, oltre ad aver facilitato notevolmente l'accesso al questionario, ha favorito la compilazione attraverso dispositivi, quali *smartphone* e *tablet*, più vicini all'utilizzo quotidiano di giovani e giovanissimi. Ai ragazzi maggiorenni e alle famiglie dei minorenni sono stati inviati anche dei promemoria. Un promemoria è stato inviato anche attraverso l'app IO. L'Istat ha prestato particolare attenzione alla salvaguardia della *privacy* e all'attuazione di protocolli di sicurezza nella raccolta dati.

Il modello di rilevazione è stato il risultato di studi condotti durante gli anni precedenti, coinvolgendo in gruppi di discussione e altre esperienze partecipative giovani e studenti. Il questionario, attraverso poche domande, copre in otto sezioni molti aspetti diversi della vita quotidiana dei giovani: a- chi sei?; b- lo studio; c- cittadinanza e identità; d- relazioni sociali; e - tempo libero; f- il tuo futuro; g- opinioni su uomini e donne; h- altre informazioni.

Il fac-simile del questionario somministrato ai ragazzi e le lettere inviate alle famiglie sono disponibili *online* al seguente link: <https://www.istat.it/it/archivio/287601>

### **Disegno di campionamento**

La popolazione di interesse dell'indagine "Bambini e ragazzi, comportamenti, atteggiamenti e progetti futuri", è costituita dagli individui residenti in Italia di età compresa tra gli 11 e i 19 anni. L'archivio di selezione è il Registro Base degli Individui (RBI) e contiene informazioni a livello individuale quali l'età, il sesso, la cittadinanza.

I domini di studio, ossia gli ambiti rispetto ai quali sono riferiti i parametri di popolazione oggetto di stima, sono definiti sulla base delle seguenti variabili di stratificazione:

- regione (21 modalità);
- cittadinanza (sette modalità): italiana, le prime cinque cittadinanze straniere presenti in Italia (albanese, cinese, marocchina, rumena e ucraina), altre cittadinanze straniere;
- classi di età (11-13, 14-19).

Il disegno campionario è di tipo stratificato a uno stadio ed è stato progettato per garantire la precisione delle stime a livello dei seguenti domini:

- cittadinanza per classi d'età;
- ripartizione geografica per classi d'età;
- ripartizione geografica per cittadinanza;
- regione per cittadinanza a due modalità (italiano, straniero) per classi d'età.

La stratificazione è stata definita dall'incrocio delle modalità della regione, delle classi di età, delle cittadinanze e del sesso, ottenendo complessivamente 588 strati. L'allocazione del campione è stata ottenuta sulla base degli errori campionari attesi di una generica stima di interesse, una prevalenza del 10% e fissando i vincoli sugli errori in modo differenziato nei domini di stima sopra definiti. Sono stati così allocati tra gli strati 39.496 bambini e ragazzi. Il numero di giovani da selezionare sulla base dell'allocazione ottima è stato incrementato in ottica di sovra campionamento a 107.961, per sopperire preventivamente alle mancate risposte, sulla base dei tassi di risposta osservati nella precedente edizione dell'indagine.

Nella rilevazione sono state utilizzate le seguenti classificazioni Istat:

- la classificazione dei codici comunali (codici Istat a 6 cifre con codice provincia e codice comune): <https://www.istat.it/it/archivio/6789>;
- la classificazione degli Stati esteri (codici Istat e 3 cifre) per la codifica univoca delle cittadinanze straniere: <https://www.istat.it/it/archivio/6747>.

### **Metodologia di calcolo dei pesi campionari**

Le stime prodotte dall'indagine sono principalmente stime di frequenze assolute.

Il principio su cui è basato ogni metodo di stima campionaria è che le unità appartenenti al campione rappresentino anche le unità della popolazione che non sono incluse nel campione. Questo principio viene realizzato attribuendo a ogni unità campionaria un peso che denota il numero di unità della popolazione rappresentate dalla unità medesima.

La procedura di costruzione dei pesi finali da assegnare alle unità campionarie consta di più fasi:

1. la prima fase, in cui si calcola il peso base (o peso diretto) come inverso della probabilità di inclusione delle unità selezionate nel campione;

2. la seconda fase, in cui si calcola un fattore correttivo di mancata risposta per fare in modo che i rispondenti all'indagine rappresentino anche le unità statistiche che non hanno risposto; tale fattore è stato calcolato come l'inverso del tasso di risposta osservato a livello di ripartizione (Nord, Centro, Sud) e cittadinanza (italiana o straniera);

3. nella fase finale, per ogni unità campionaria rispondente, si calcola un fattore correttivo, detto fattore di "calibrazione", che consente di soddisfare la condizione di uguaglianza tra i totali noti della popolazione e le corrispondenti stime campionarie.

I totali noti della popolazione, declinati a livello regionale, sono stati:

- il numero degli 11-19enni di ognuna delle "principali" cittadinanze considerate come variabili di stratificazione (7 vincoli);
- il numero degli 11-19enni distinti per sesso e tipo di scuola frequentata (medie, superiori e non iscritti), con la tipologia di scuola così come risultante dagli archivi del MIM (6 vincoli);
- il numero degli 11-19enni distinti per sesso, età suddivisa in due classi (11-13enni, 14-19enni) e cittadinanza suddivisa tra italiana e "non italiana" (8 vincoli).

Il peso finale della generica unità campionaria è dato dal prodotto del suo peso base per il fattore correttivo di mancata risposta e per il fattore di calibrazione.

### Valutazione del livello di precisione delle stime

Le principali statistiche di interesse per valutare la variabilità campionaria delle stime prodotte da un'indagine sono l'errore di campionamento assoluto e l'errore di campionamento relativo, definite dalle seguenti espressioni:

$$\hat{\sigma}(\hat{Y}) = \sqrt{\hat{V}(\hat{Y})}$$

$$\hat{\varepsilon}(\hat{Y}) = \frac{\hat{\sigma}(\hat{Y})}{\hat{Y}}$$

Le stime prodotte dall'indagine sono state ottenute mediante uno stimatore di calibrazione in due passi sulla base di una funzione di distanza di tipo lineare. Poiché lo stimatore adottato non è funzione lineare dei dati campionari non è possibile ottenere una espressione analitica per la stima della varianza. Pertanto si è utilizzato il metodo proposto da Woodruff che, ricorrendo all'espressione linearizzata in serie di Taylor, consente di ottenere la varianza di ogni stimatore non lineare calcolando la varianza dell'espressione linearizzata ottenuta. Tale metodologia di stima della varianza è implementata nel *software* generalizzato ReGenesees, che è stato utilizzato per la stima della varianza delle stime.

Poiché le stime prodotte dall'indagine in oggetto sono in numero molto elevato, si è fatto ricorso ad una presentazione sintetica degli errori campionari. A tal fine si utilizza il metodo dei modelli regressivi che si basa sulla determinazione di una funzione matematica che mette in relazione ciascuna stima con il proprio errore campionario relativo stimato.

Il modello utilizzato per le stime di frequenze assolute e relative è il seguente:

$$\log \hat{\varepsilon}^2(\hat{Y}) = a + b * \log(\hat{Y})$$

dove i parametri a e b sono stimati, per un certo dominio di stima, con il metodo dei minimi quadrati su un insieme di stime ottenute dall'indagine (con i rispettivi errori relativi) che coprono approssimativamente l'intervallo di variazione delle stime di frequenze che vengono pubblicate. I parametri dei modelli descritti, che permettono la presentazione sintetica degli errori di campionamento, sono stati stimati tramite il *software* ReGenesees.

Utilizzando gli opportuni coefficienti è possibile calcolare una stima dell'errore campionario relativo di una generica stima di una frequenza assoluta  $\hat{Y}$  applicando la seguente formula:

$$\hat{\varepsilon}(\hat{Y}) = \sqrt{\exp(a + b * \log(\hat{Y}))}$$

I modelli di presentazione sintetica degli errori di campionamento sono stati stimati per diversi domini di stima. I modelli proposti sono: il totale generale a livello Italia, i ragazzi distinti per cittadinanza (a 2 e a 7 modalità), la distinzione per ripartizione geografica e incrociando la ripartizione con la cittadinanza (a 2 modalità).

Nel seguito sono riportati, per tutti i domini di stima appena indicati, i prospetti relativi ai valori dei coefficienti a e b e dell'indice di determinazione  $R^2$  dei modelli d'interpolazione tra stime ed i relativi errori (Prospetto A). A questo prospetto segue il corrispettivo recante, per ogni modello, i valori interpolati indicativi degli errori campionari relativi alle corrispondenti stime assolute (Prospetto B).

**Prospetto A.** Valori dei coefficienti a e b e dell'indice di determinazione  $R^2$  del modello per l'interpolazione degli errori campionari delle stime riferite a diversi domini di stima

	A	B	R <sup>2</sup>
<b>ITALIA</b>			
	9,263	-1,339	0,874
<b>Italia per cittadinanza a due modalità</b>			
Italiani	5,093	-0,947	0,909
Stranieri	2,1	-0,815	0,803
<b>Ripartizione</b>			
Nord Ovest	4,064	-0,838	0,789
Nord Est	3,547	-0,86	0,815
Centro	4,121	-0,913	0,841
Sud e isole	3,727	-0,86	0,814
<b>Ripartizione per cittadinanza a due modalità</b>			
Italiani - Nord Ovest	4,45	-0,845	0,838
Stranieri - Nord Ovest	2,013	-0,742	0,703
Italiani - Nord Est	3,797	-0,852	0,867
Stranieri - Nord Est	1,917	-0,806	0,794
Italiani - Centro	4,333	-0,903	0,877
Stranieri - Centro	1,855	-0,791	0,755
Italiani - Sud e isole	4,324	-0,891	0,885
Stranieri - Sud e isole	1,752	-0,804	0,841
<b>Cittadinanza a 7 modalità</b>			
Italia	5,093	-0,947	0,909
Albania	2,958	-0,983	0,908
Cina	2,744	-0,971	0,93
Marocco	2,977	-0,969	0,941
Romania	3,989	-1,011	0,938
Ucraina	2,225	-1,021	0,94
Altro	4,148	-0,944	0,894

**Prospetto B.** Valori interpolati degli errori campionari di alcune prevalenze tipiche per diversi domini di stima, stime assolute

	1.000	10.000	25.000	50.000	100.000	200.000	500.000	1.000.000
<b>ITALIA</b>								
	42,17	14,07	9,09	6,53	4,69	3,37	2,18	1,57
<b>Italia per cittadinanza a due modalità</b>								
Italiani	48,38	16,25	10,53	7,58	5,46	3,93	2,55	1,83
Stranieri	17,1	6,69	4,6	3,47	2,61	1,97	1,36	1,02
<b>Ripartizione</b>								
Nord Ovest	42,26	16,11	10,97	8,21	6,14	4,59	3,13	2,34
Nord Est	30,26	11,25	7,58	5,63	4,18	3,1	2,09	1,55
Centro	33,53	11,72	7,72	5,62	4,1	2,99	1,97	1,43
Sud e isole	33,02	12,26	8,27	6,14	4,55	3,38	2,28	1,69
<b>Ripartizione per cittadinanza a due modalità</b>								
Italiani - Nord Ovest	49,91	18,85	12,8	9,55	7,12	5,31	3,61	2,69
Stranieri - Nord Ovest	21,07	8,96	6,38	4,93	3,81	2,95	2,1	1,62
Italiani - Nord Est	35,18	13,19	8,93	6,64	4,94	3,68	2,49	1,85
Stranieri - Nord Est	16,13	6,38	4,41	3,34	2,52	1,91	1,32	1
Italiani - Centro	38,62	13,66	9,03	6,61	4,83	3,53	2,34	1,71
Stranieri - Centro	16,43	6,61	4,6	3,49	2,66	2,02	1,41	1,07
Italiani - Sud e isole	40,1	14,39	9,57	7,03	5,16	3,79	2,52	1,85
Stranieri - Sud e isole	14,96	5,93	4,1	3,1	2,35	1,78	1,23	0,93
<b>Cittadinanza a 7 modalità</b>								
Italia	48,38	16,25	10,53	7,58	5,46	3,93	2,55	1,83
Albania	14,72	4,75	3,03	2,15	1,53	1,09	0,69	0,49
Cina	13,77	4,5	2,88	2,06	1,47	1,05	0,67	0,48
Marocco	15,61	5,12	3,28	2,35	1,68	1,2	0,77	0,55
Romania	22,37	6,98	4,39	3,1	2,18	1,54	0,97	0,68
Ucraina	8,95	2,76	1,73	1,22	0,85	0,6	0,38	0,26
Altro	30,57	10,32	6,7	4,83	3,48	2,51	1,63	1,17

#### Riferimenti ad altri comunicati e note metodologiche

- ISTAT (2024), Nuove generazioni sempre più digitali e multiculturali, Statistiche report, 20 maggio 2024: <https://www.istat.it/comunicato-stampa/indagine-bambini-e-ragazzi-anno-2023>
- ISTAT (2025), File ad uso pubblico mIcro.STAT su “Bambini e ragazzi: comportamenti, atteggiamenti e progetti futuri”, Aspetti metodologici dell’indagine: <https://www.istat.it/microdati/integrazione-delle-seconde-generazioni-microdati-ad-uso-pubblico>