

DICEMBRE 2024 | NUMERO 37

GLI APPROFONDIMENTI DEL SERVIZIO STATISTICO

La newsletter ufficiale del Servizio Statistico INVALSI



Gentilissimo/a,

la newsletter del mese di dicembre de “*Gli approfondimenti del Servizio Statistico*” offre alcune indicazioni sulle potenzialità dei dati rilasciati da INVALSI. Ricordiamo infatti (vedi newsletter n. 4, marzo 2022) che da qualche anno il patrimonio informativo realizzato da INVALSI viene reso pubblico in formato *Open Data*.

Il valore degli Open Data

Gli Open Data costituiscono un'opportunità, un territorio da esplorare. Da anni INVALSI è impegnato ad aumentare la quantità e la qualità dei dati che sono restituiti alle scuole, al mondo della ricerca e, più in generale, a tutti coloro che sono interessati allo stato di salute dell'istruzione e della formazione professionale nel nostro Paese.

Trasmettere conoscenze è difficile quanto realizzare un prodotto di alta pasticceria. Occorre saper interpretare e soddisfare le preferenze di chi legge ma anche, innanzitutto, fornire un prodotto sano, non sofisticato, atto a soddisfare i bisogni reali dei propri fruitori. Basta poco per rovinare il risultato finale.

È inevitabile che, per quanto ci si sforzi di elevare lo standard del lavoro di restituzione di risultati e informazioni, qualcosa rischi di venire a mancare. C'è sempre un elemento o un dettaglio che avrebbe potuto esser fatto in un altro modo, un ingrediente che avrebbe potuto essere aggiunto o omesso o combinato altrimenti. La messa a disposizione in formato aperto (Open) dei dati di partenza è un modo per riavvolgere il nastro e porre le esigenze di ogni singolo fruitore al centro di tutto.

In un bel film di animazione di qualche anno fa era detto che in ognuno di noi c'è potenzialmente un ottimo cuoco o un ottimo pasticcere. Gli Open Data consentono a ciascuno di provare a misurarsi con l'esperienza di "cucinare la propria torta". Gli Open Data sono infatti gli ingredienti dai quali si ottiene il prodotto informativo finale. Non si tratta qui di inventare il mondo da capo ma di provare a sperimentare a partire da ricette esistenti, in modo da adattare il più possibile il risultato ai propri bisogni conoscitivi. Perché, in ultima analisi, ciascuno conosce meglio di chiunque altro le proprie esigenze e, per quanto ci si possa sforzare, una risposta standardizzata fornita da un soggetto terzo è sempre incompleta sotto qualche aspetto o dettaglio.

Vediamo allora qualche esempio. Nel corso degli anni, l'INVALSI ha cercato di potenziare la produzione di statistiche territoriali. L'informazione territoriale è un tema sensibile. L'Italia è un paese di differenze profonde e un numero crescente di funzioni in materia di istruzione e formazioni tende ad essere decentrato a livello regionale e provinciale.

L'ultimo rapporto INVALSI contiene molti elementi che muovono nella direzione di ampliare l'informazione territoriale; ad esempio, inserendo grafici come questo, in cui la distribuzione percentuale degli studenti di 13° grado per livelli di competenza di italiano è disaggregata per regioni:

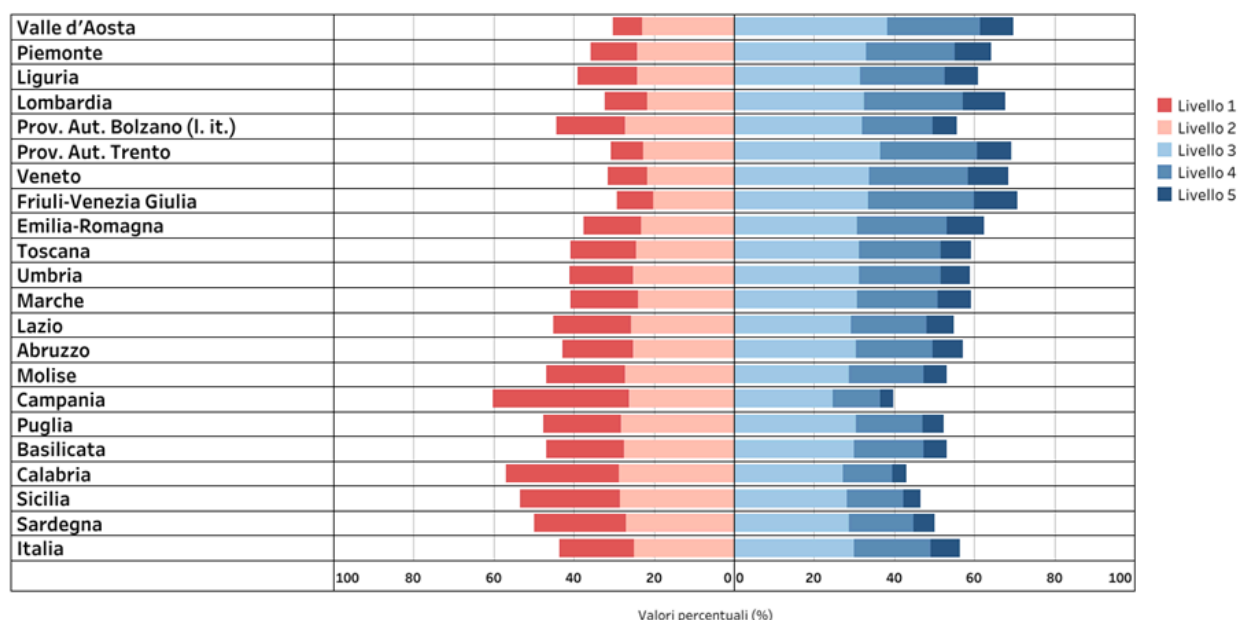


Grafico 1 – Distribuzione percentuale studenti 13° grado in base al livello conseguito nelle prove di italiano, anno 2023/2024 (fonte: INVALSI)

Il dettaglio regionale rappresenta un utile punto di riferimento. Tuttavia, molte importanti informazioni che hanno ad oggetto la programmazione e la gestione delle scuole sono prese ad un livello territoriale ancor più disaggregato, provinciale o addirittura comunale.

Attingendo agli *Open Data* INVALSI, chi è interessato ad approfondire i risvolti territoriali delle questioni che attengono al mondo della scuola può trovare molti elementi utili.

Per farlo è sufficiente infatti recarsi sulla pagina web del Servizio Statistica INVALSI all'URL: <https://serviziostatistico.invalsi.it/catalogo-dati/>



The screenshot shows the INVALSI Servizio Statistico website. The header includes the INVALSI logo and the text 'INVALSI Servizio Statistico'. There are social media icons for Facebook and YouTube, and a search bar labeled 'Cerca'. The navigation menu contains: Home, Eventi, Dati, Chi siamo, Approfondimenti, Progetti, FAQ, and Newsletter. Below the menu, the breadcrumb 'Home / Catalogo Dati' is visible. The main heading is 'Catalogo Dati'. Below this, there are four categories represented by icons and text: 'Open Data' (with an icon of a document and a pie chart), 'Popolazione' (with a globe icon), 'Campione' (with a magnifying glass over people icons), and 'Internazionali' (with a world map icon).

e cliccare sul pulsante *Open Data*. Questa operazione dà accesso a una pluralità di data base in formato .xls o .csv che possono essere liberamente scaricati ed elaborati (citando la fonte, in caso di diffusione).

Per questa piccola dimostrazione andremo a prendere il file che contiene i risultati delle prove INVALSI degli ultimi anni aggregati a livello provinciale (report_provincelivelli.csv), a cui si può accedere dalla scheda "Dati provinciali popolazione".

Aprire un file di dati è sempre un'esperienza emozionante. È un po' come scartare un regalo: un misto di aspettative e di curiosità.

In questo caso, quello che si materializza sotto i nostri occhi è una tabella con qualche migliaio di righe ma con un numero ragionevole di colonne, che corrispondono alla denominazione della provincia, all'anno ed al grado scolastico di riferimento, alle percentuali di studenti che si collocano nei vari livelli di competenze.

Ecco spiegato l'arcano del numero di righe: il file è così lungo perché contiene la "storia" dei risultati delle prove INVALSI da quando è iniziata la somministrazione, per tutti i gradi scolastici e per tutte le province del Paese. In questo caso, i risultati sono misurati in base a scale di livelli di competenze, che vanno da uno a cinque per italiano e matematica e da A1 a C2 (classificazione internazionale) per le prove di inglese.

Può sembrare tanta, troppa informazione tutta insieme, ma è semplice ridurla a una dimensione più gestibile in base alle proprie necessità. Tutti i fogli di calcolo dispongono infatti di funzioni che consentono di filtrare i dati in base ai criteri o alle dimensioni di interesse.

A	B	C	D	E	F	G	H	I	J	K	L
Codice_provinc	Sigla_provinc	Nome_provincia	Grado	Materia	Anno	LIVELLO_1	LIVELLO_2	LIVELLO_3	LIVELLO_4	LIVELLO_5	perc_copertura_s
1	TO	TORINO	13	Italiano	2023-24	12,3	23,7	32,2	22,4	9,4	98,1
2	VC	VERCELLI	13	Italiano	2023-24	15,0	28,2	33,0	17,6	6,1	98,8
3	NO	NOVARA	13	Italiano	2023-24	9,8	23,6	35,6	23,3	7,6	98,2
4	CN	CUNEO	13	Italiano	2023-24	10,8	24,1	33,1	22,1	9,9	98,2
5	AT	ASTI	13	Italiano	2023-24	13,1	25,3	30,4	20,7	10,5	98,1
6	AL	ALESSANDRIA	13	Italiano	2023-24	11,4	25,5	34,7	20,9	7,5	97,9
7	AO	AOSTA	13	Italiano	2023-24	7,3	22,8	38,2	23,4	8,3	95,1
8	IM	IMPERIA	13	Italiano	2023-24	14,1	27,9	34,1	18,2	5,7	97,9

Utilizzando i filtri appropriati si possono facilmente ottenere, ad esempio, una tabella e un grafico che mettono a confronto i risultati nelle prove di italiano del 13° grado per l'ultimo anno scolastico (2023/2024) degli studenti delle province del Lazio:

PROVINCIA	LIVELLO_1	LIVELLO_2	LIVELLO_3	LIVELLO_4	LIVELLO_5
VITERBO	21,0	27,6	29,0	17,4	5,1
RIETI	16,0	27,1	29,9	18,1	8,9
ROMA	18,9	25,1	29,3	19,4	7,3
LATINA	18,5	26,6	30,5	18,7	5,7
FROSINONE	24,1	28,2	27,8	15,2	4,8

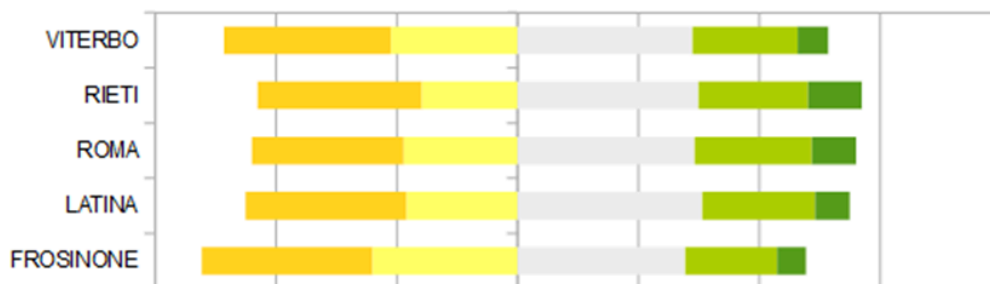


Grafico 2 - Distribuzione percentuale studenti 13° grado in base al livello conseguito nelle prove di italiano, anno 2023/2024, province del Lazio (fonte: elaborazione su [Open Data](#) INVALSI)

Questo tipo di risultato può essere comparato con i dati regionali forniti nel rapporto e rappresentati nel grafico 1, per ottenere un'idea del posizionamento relativo delle singole province ma anche della variabilità interna a una singola regione.

Ma negli *Open Data* c'è ancora di più. Ampliando la selezione degli anni si possono analizzare le tendenze dei fenomeni di interesse. Ad esempio, nel grafico 3 è rappresentato l'andamento dal 2019 a oggi della percentuale di studenti dell'ultimo anno delle scuole superiori che non raggiungono il livello minimo di competenze nelle prove INVALSI di Italiano. Le province riportate nell'esempio appena descritto sono quelle del Lazio ma nulla vieta di fare confronti fra territori di regioni diverse.

	2018-19	2020-21	2021-22	2022-23	2023-24
VITERBO	39,8	46,5	48,6	54,0	48,6
RIETI	35,3	49,7	57,5	50,6	43,1
ROMA	36,8	49,6	49,5	50,2	44,0
LATINA	40,6	52,7	54,5	53,5	45,1
FROSINONE	48,9	55,9	57,2	58,6	52,3

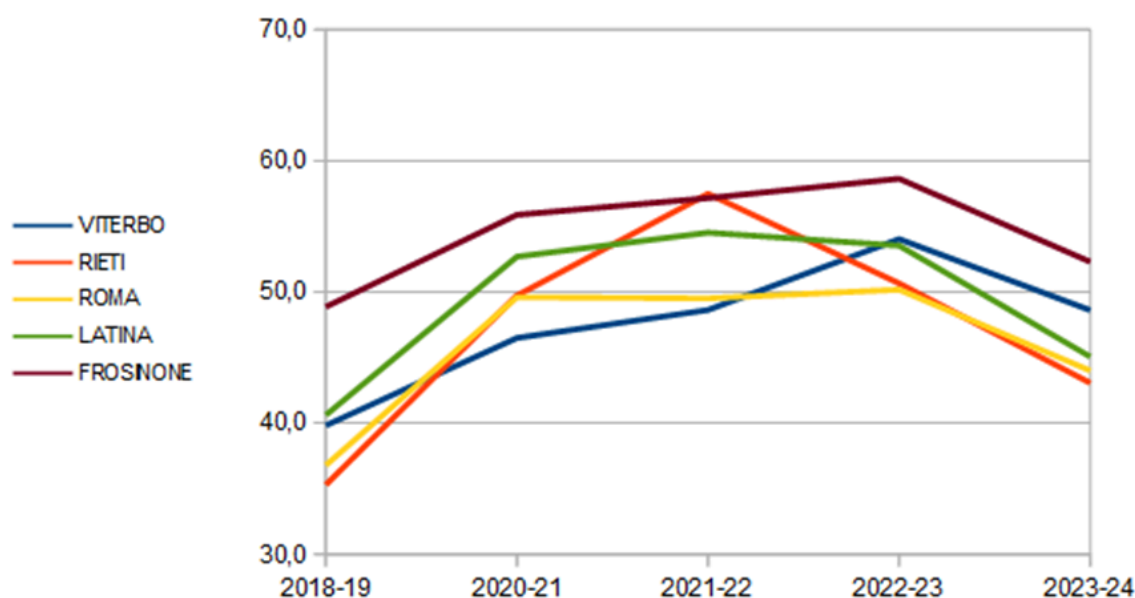


Grafico 3 - Trend percentuali studenti che non raggiungono il livello minimo nelle prove di italiano del 13° grado, province del Lazio (fonte: elaborazione su *Open Data* INVALSI)

Provare a lavorare con gli *Open Data* è non soltanto un modo per rispondere in modo più appropriato ai propri bisogni conoscitivi ma anche per ampliare l'orizzonte delle domande, per scavare nella crescente complessità dei fatti della scuola.

Nessun ente statistico, per quanto animato dalle migliori intenzioni, può soddisfare esigenze così dettagliate per tutti i fruitori delle informazioni che produce. Esso può, però, mettere a disposizione gli strumenti affinché ciascuno possa sviluppare il proprio percorso di conoscenza.

Di seguito il link per esplorare [il catalogo degli Open Data](#)

e il link [Newsletter - Servizio Statistico INVALSI](#) dove è possibile recuperare i precedenti numeri della newsletter.