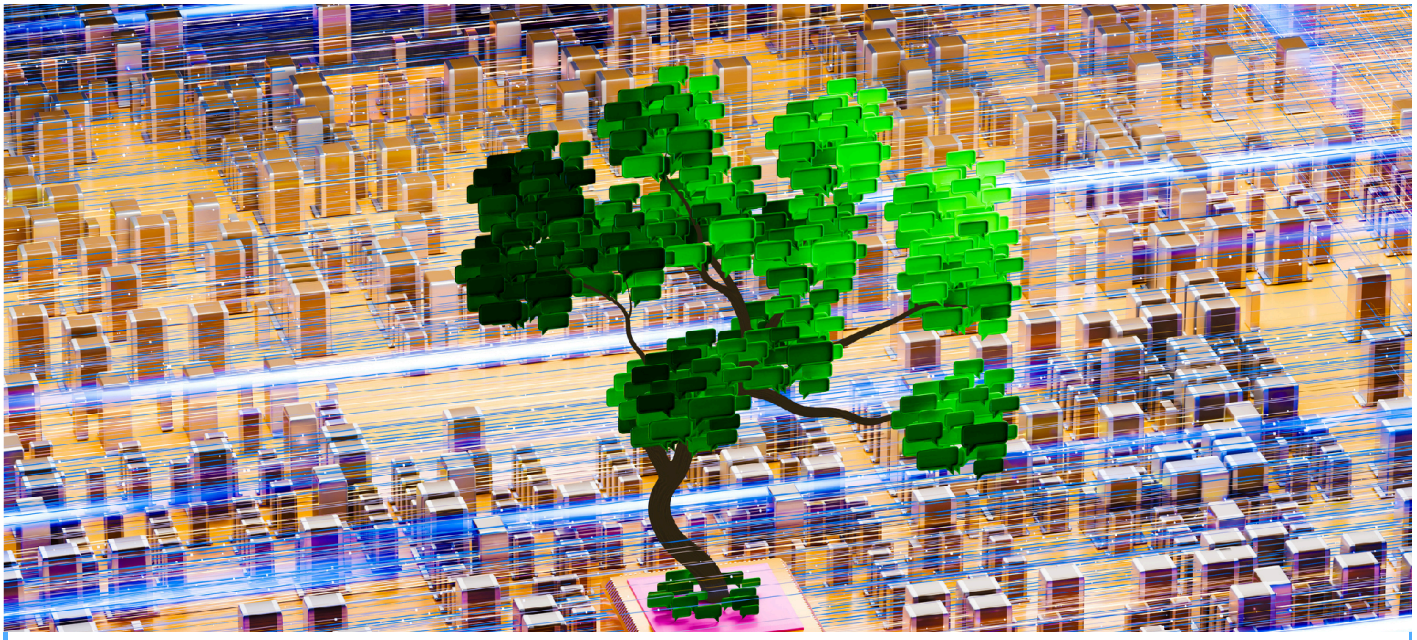
I dati pubblici: una risorsa per capire e migliorare il mondo

DI MASSIMO VAILATI

12/10/2026

INTELLIGENZA ARTIFICIALE

ITE TURISMO, LICEO, TECNICO, TT INFORMATICA



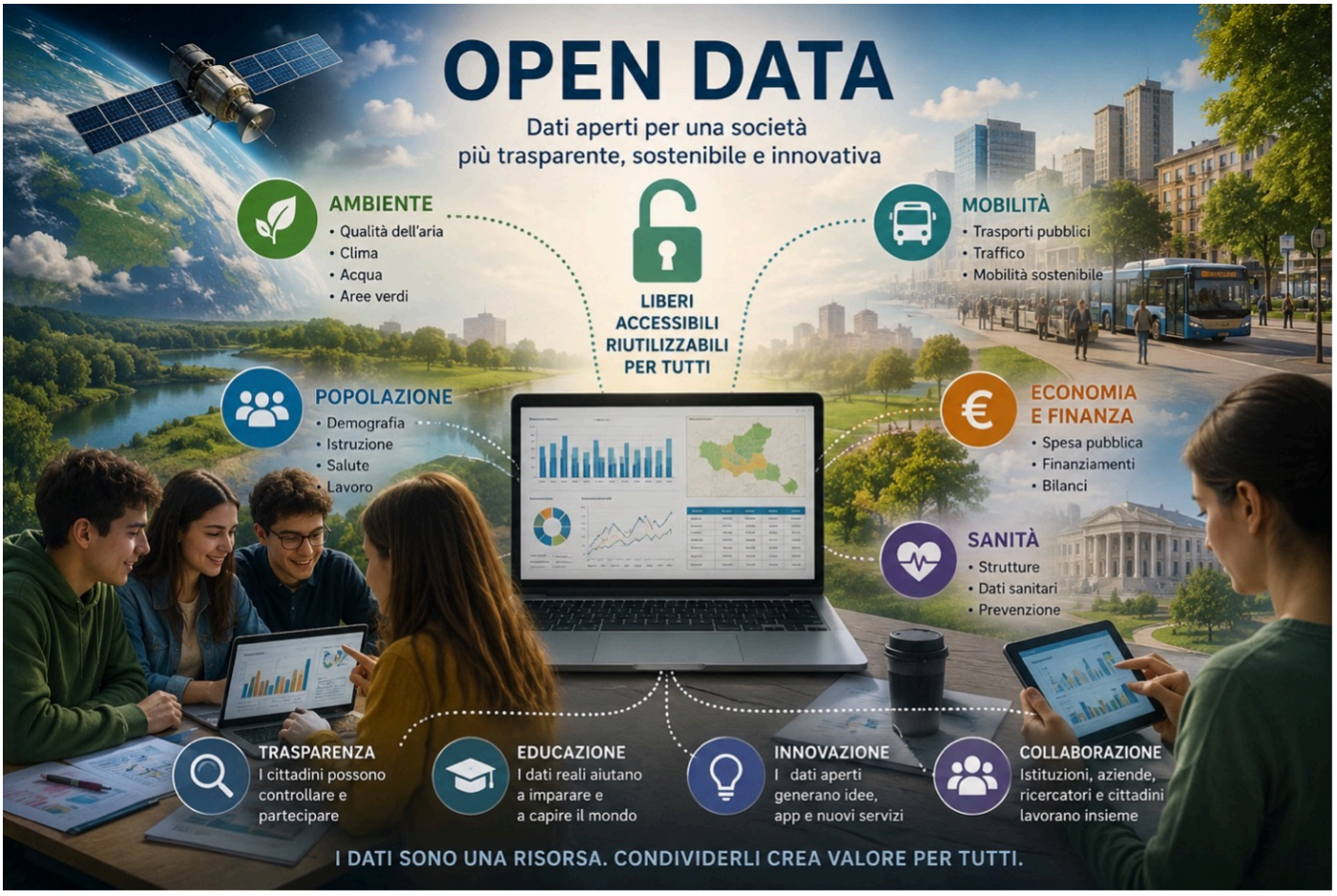
I dati pubblici, o open data, sono informazioni raccolte dagli enti pubblici e rese disponibili a tutti. Essi permettono di comprendere meglio fenomeni sociali, ambientali ed economici e possono essere utilizzati per sviluppare nuovi servizi e ricerche. Il testo evidenzia anche il loro valore educativo: analizzare dati reali aiuta gli studenti a sviluppare competenze digitali, matematiche e critiche. Come esempio viene analizzata la qualità dell'aria a Milano attraverso i dati della Regione Lombardia relativi alla concentrazione di PM10. I dati, elaborati con un foglio di calcolo, mostrano valori più elevati nei mesi invernali e alcuni picchi il primo giorno dell'anno, probabilmente legati ai fuochi d'artificio e alle condizioni meteorologiche della Pianura Padana.

L'attività dimostra come i dati pubblici possano aiutare a comprendere la realtà e a sviluppare una maggiore consapevolezza sui problemi ambientali.

Ogni giorno produciamo enormi quantità di informazioni: spostamenti, dati meteorologici, statistiche sanitarie, risultati scolastici, traffico urbano, consumi energetici. Una parte di queste informazioni viene raccolta da enti pubblici e resa disponibile a tutti sotto forma di **public data** (o dati pubblici aperti).

I public data rappresentano una delle risorse più importanti dell'era digitale perché permettono ai cittadini, alle scuole, alle aziende e ai ricercatori di analizzare fenomeni reali e prendere decisioni più consapevoli. Non si tratta soltanto di numeri o tabelle: dietro ogni dato ci sono fenomeni reali, problemi concreti e possibilità di miglioramento. I dati pubblici possono raccontare come cambia il clima, come si muovono le persone in una città, quali quartieri sono più popolosi oppure quanto inquina il traffico urbano.

Molti governi e organizzazioni internazionali hanno scelto di pubblicare grandi quantità di dati online in modo gratuito e aperto. In questo caso si parla spesso anche di **open data**, cioè dati che possono essere scaricati, riutilizzati e analizzati liberamente. Questa scelta favorisce la trasparenza e rende possibile la nascita di nuovi servizi digitali, applicazioni e ricerche scientifiche.



Un esempio molto importante è il programma europeo **Copernicus**, che raccoglie dati satellitari sulla Terra grazie ai satelliti **Sentinel**. Attraverso queste informazioni è possibile osservare l'andamento delle temperature, monitorare incendi boschivi, studiare l'inquinamento atmosferico o controllare lo stato dei ghiacciai. I dati di Copernicus vengono utilizzati da scienziati, meteorologi, università e anche da studenti che vogliono approfondire temi legati all'ambiente e al cambiamento climatico.

Anche in Italia esistono numerosi portali di dati pubblici. Il portale nazionale **dati.gov.it** raccoglie dataset provenienti da ministeri, regioni e comuni. Altri siti, come quelli dell'**ISTAT** o delle **ARPA** regionali, permettono di consultare statistiche sulla popolazione, sull'economia o sulla qualità dell'aria. Questi dati possono aiutare a comprendere meglio il territorio in cui viviamo e a osservare come cambiano nel tempo fenomeni sociali e ambientali.

I public data hanno inoltre un forte valore educativo. Utilizzare dati reali durante le attività scolastiche permette agli studenti di sviluppare competenze digitali, statistiche e critiche. Analizzare informazioni autentiche rende lo studio più concreto e vicino alla realtà quotidiana. Per esempio, osservare i dati sull'inquinamento della propria città può aiutare a riflettere sul traffico urbano e sulle scelte ambientali, mentre studiare i dati demografici permette di capire come cambia la popolazione nel tempo.

Naturalmente è importante utilizzare i dati in modo corretto. I numeri, da soli, non bastano: devono essere interpretati e contestualizzati. Inoltre i dati pubblici devono rispettare la privacy delle persone e non contenere informazioni sensibili. Per questo motivo è necessario imparare non solo a leggere i dati, ma anche a verificarne l'origine, la qualità e l'affidabilità.

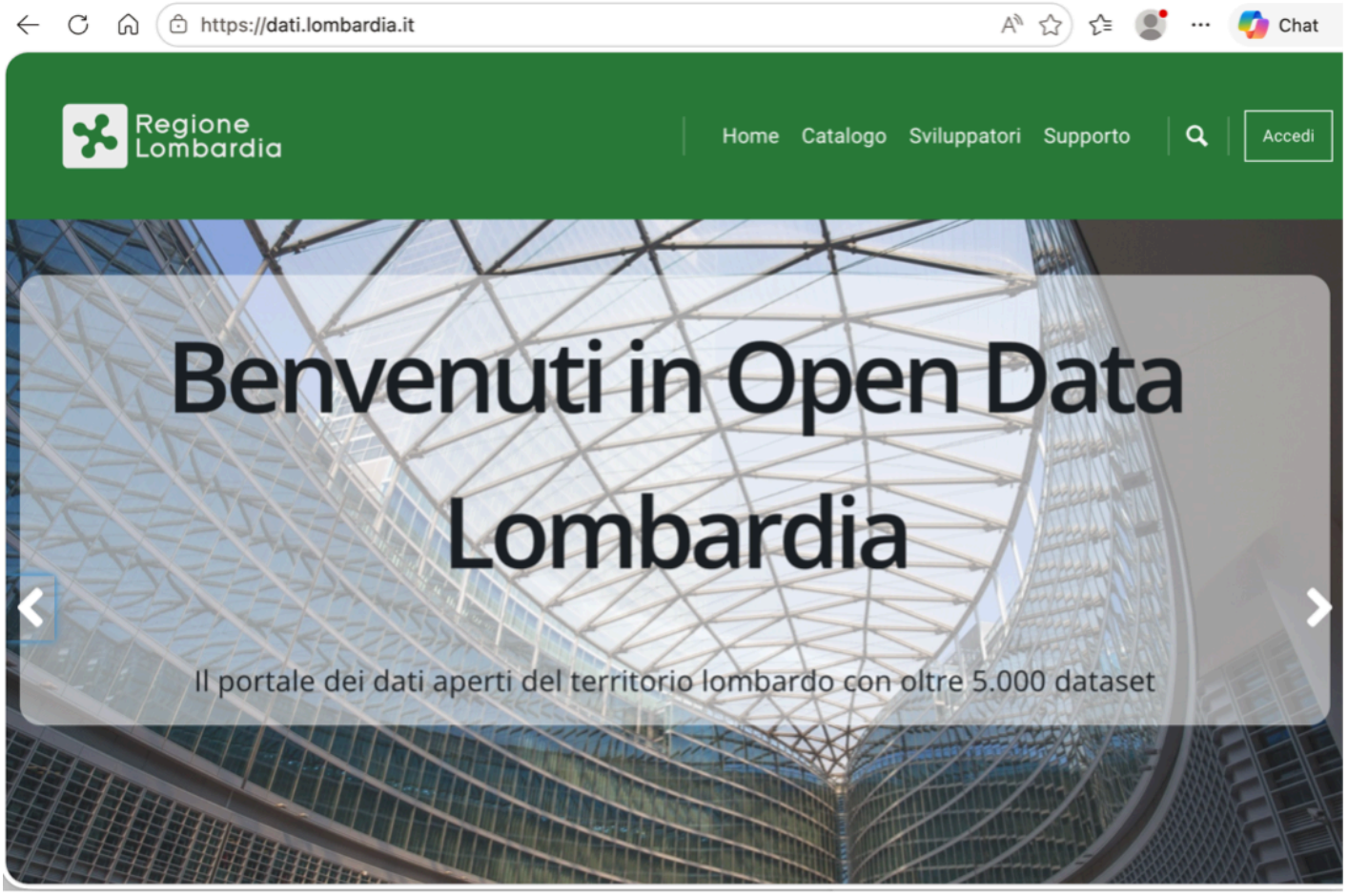
I **public data** rappresentano una delle risorse più preziose della società digitale. Grazie ai dati aperti è possibile osservare il mondo in modo più consapevole, comprendere problemi complessi e partecipare attivamente alla vita pubblica. Imparare a leggere e utilizzare i dati significa sviluppare una competenza fondamentale per il presente e per il futuro.

Attività pratica

L'attività didattica proposta consiste nell'analizzare un dataset reale proveniente dal proprio comune. Gli studenti possono scaricare dati sulla qualità dell'aria, sulle piste ciclabili o sui trasporti pubblici e costruire grafici o mappe per individuare tendenze e differenze tra quartieri o periodi dell'anno. Un lavoro di questo tipo unisce informatica,

matematica, geografia ed educazione civica, mostrando come discipline diverse possano collaborare per interpretare la realtà.

Come esempio pratico accediamo al sito della **Regione Lombardia** all'indirizzo **dati.lombardia.it**



Da qui è possibile accedere a:

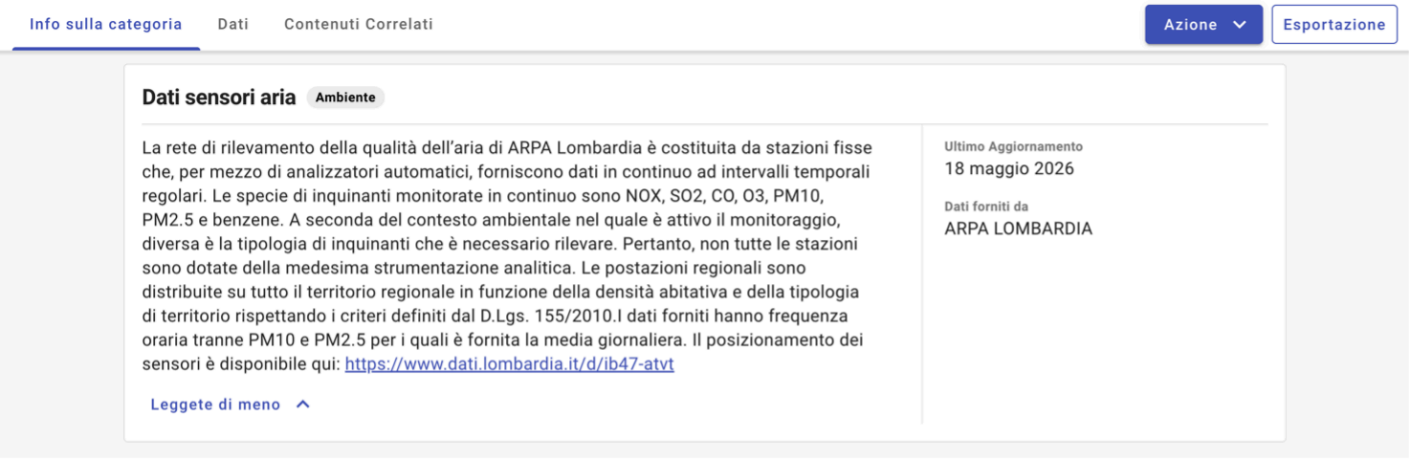
- **Catalogo**, elenco completo degli open data pubblicati suddivisi per categoria,
- **Mappe**, elenco di mappe che utilizzano dati geospaziali,
- **Grafici**, elenco di grafici (istogrammi, a torta, etc.) in grado di rappresentare in modo visuale i dati,
- **Datalens**, strumento interattivo di analisi dei dati attraverso grafici e mappe sincronizzati tra loro.

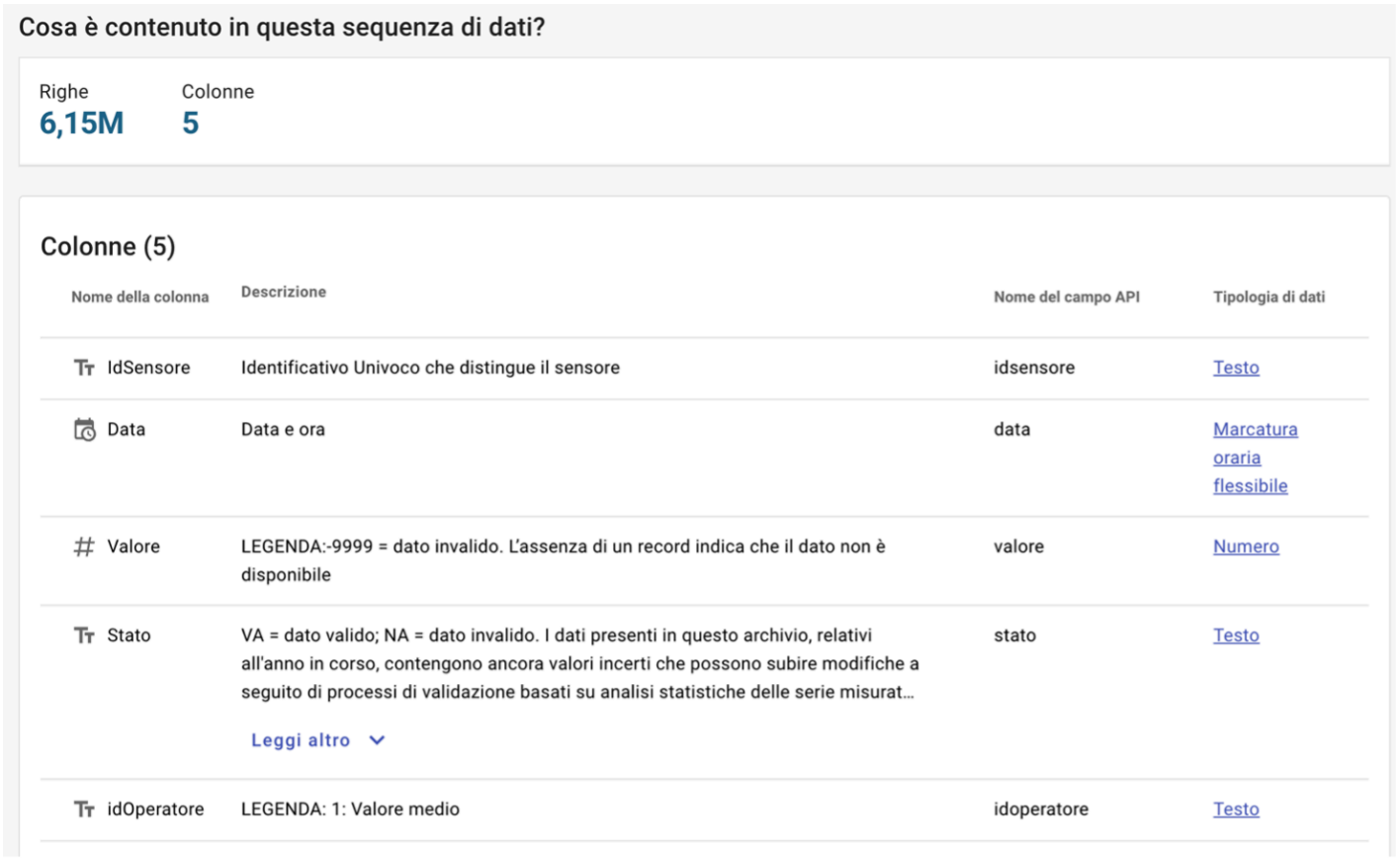
Vogliamo analizzare la qualità dell'aria, ad esempio la concentrazione di PM10 in una zona di Milano. Scegliamo tra le categorie del catalogo la voce **Ambiente** per accedere

all'archivio **Dati sensori aria**.



Nella pagina successiva è possibile eseguire diverse azioni, come ad esempio visualizzare i dati attraverso strumenti online, interrogare i dati mediante API, ed altro. Noi siamo interessati a scaricare i dati sul nostro computer quindi sceglieremo **Dati della Query**. Prima però leggiamo come sono organizzati i dati e quale è il loro significato. Leggere la struttura dei dati è importante per capire cosa stiamo esaminando.





Cosa è contenuto in questa sequenza di dati?

Righe	Colonne
6,15M	5

Colonne (5)

Nome della colonna	Descrizione	Nome del campo API	Tipologia di dati
IdSensore	Identificativo Univoco che distingue il sensore	idsensore	Testo
Data	Data e ora	data	Marcatura oraria flessibile
Valore	LEGENDA:-9999 = dato invalido. L'assenza di un record indica che il dato non è disponibile	valore	Numero
Stato	VA = dato valido; NA = dato invalido. I dati presenti in questo archivio, relativi all'anno in corso, contengono ancora valori incerti che possono subire modifiche a seguito di processi di validazione basati su analisi statistiche delle serie misurat… Leggi altro	stato	Testo
idOperatore	LEGENDA: 1: Valore medio	idoperatore	Testo

Poiché l'archivio in questione contiene oltre 6 milioni di righe possiamo filtrare i dati prima di scaricarli. Ci interessa un particolare sensore e possiamo quindi analizzare l'archivio dei sensori indicato nella descrizione. Andiamo all'archivio cliccando su link riportato nella descrizione: **https://www.dati.lombardia.it/d/ib47-atvt**

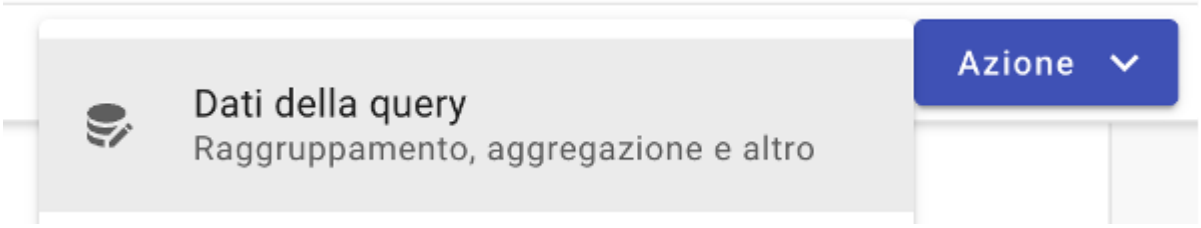
Cosa è contenuto in questa sequenza di dati?

Righe	Colonne	Identificatore di riga
998	16	IdSensore

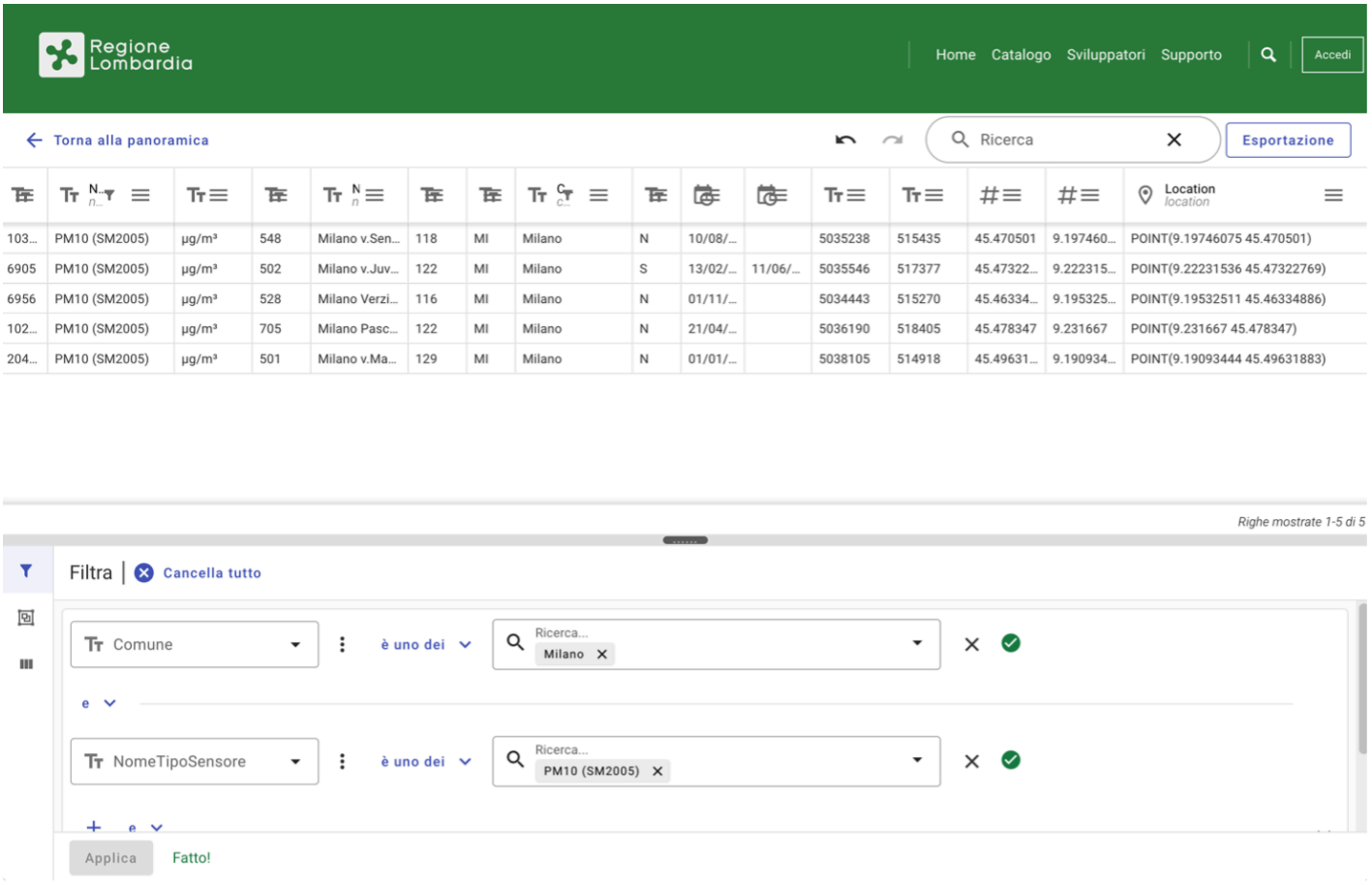
Colonne (16)

Nome della colonna	Descrizione	Nome del campo API	Tipologia di dati
IdSensore		idsensore	Testo
NomeTipoSensore		nometiposensore	Testo
UnitaMisura		unitamisura	Testo
Idstazione		idstazione	Testo
NomeStazione		nomestazione	Testo
Quota		quota	Testo
Provincia		provincia	Testo
Comune		comune	Testo

Da qui possiamo vedere l'elenco dei sensori e la loro ubicazione. Anziché scaricare l'archivio in questo caso possiamo limitarci a visualizzare i dati per individuare il sensore che ci interessa. Scegliamo **Dati della query** dal menu **Azione**.

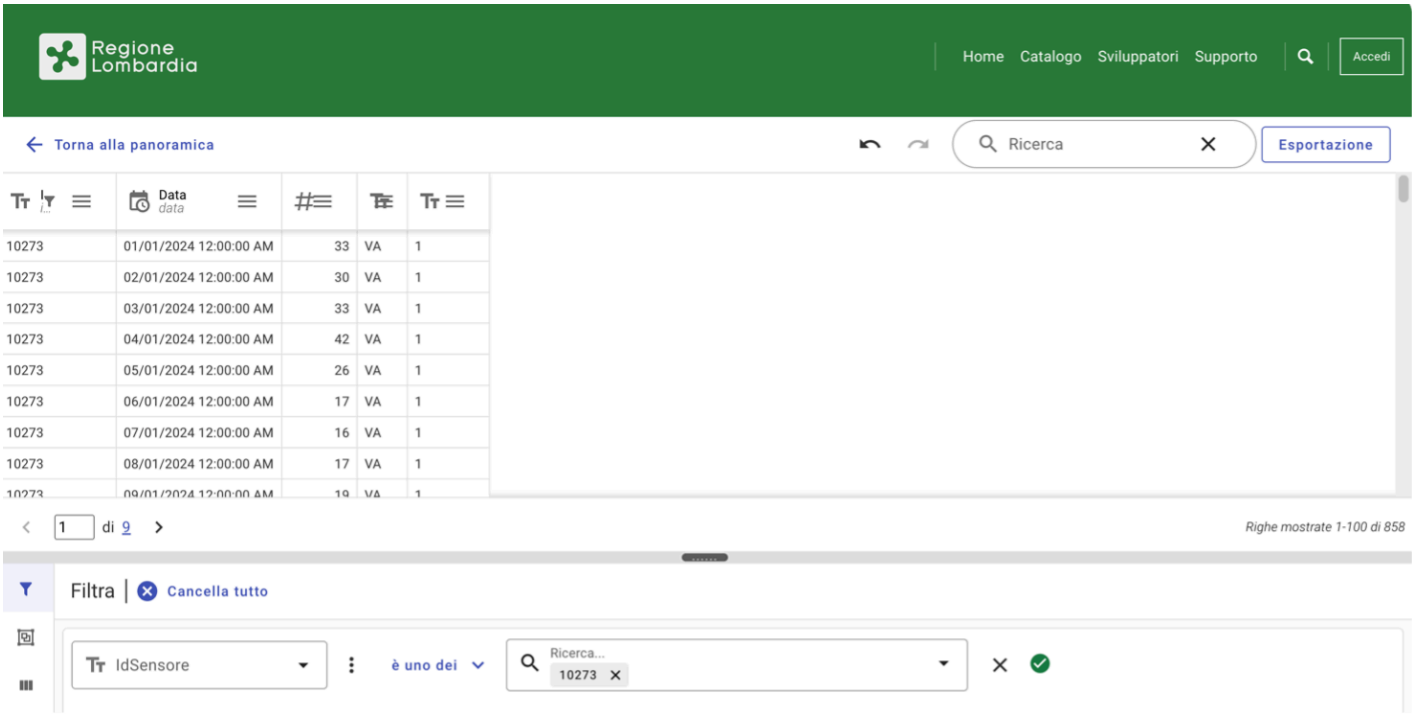


La pagina successiva mostra i dati in anteprima e possiamo applicare alcuni filtri per cercare quello che ci interessa, ad esempio sensori di PM10 nel comune di Milano.



Vediamo che ci sono 5 sensori di PM10 situati in varie zone della città. Prendiamo come esempio il sensore **10273** situato in **Viale Pascal Città Studi** (il penultimo dell'elenco).

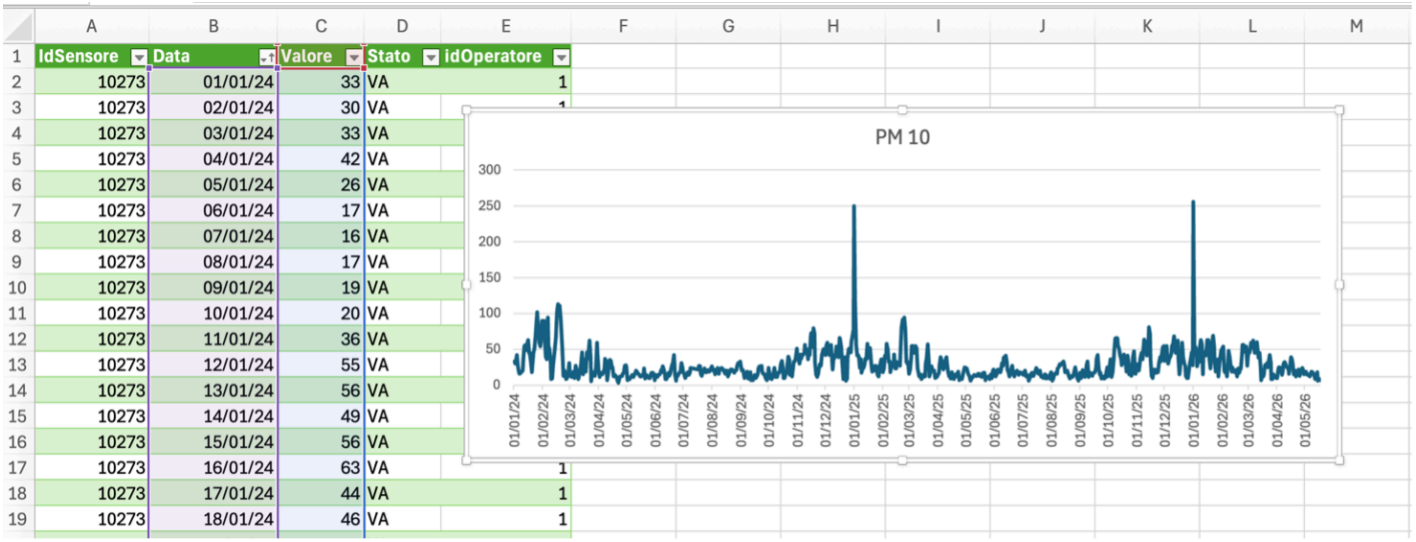
Torniamo all'archivio dei dati iniziale (sensori aria) e filtriamo i dati in modo da limitare le informazioni relative al solo sensore di interesse.



Possiamo ora scaricare i dati cliccando su **Esportazione**. Scegliamo il formato CSV più pratico da elaborare.



A questo punto possiamo elaborare i dati scaricati con il nostro foglio di calcolo realizzando grafici e svolgendo le analisi che desideriamo. Per prima cosa ripuliamo i dati dai valori non validi (come indicato nella descrizione il dato -9999 con il tag NA). Poi ordiniamo i dati per Data crescente e realizziamo un grafico a linea. Otteniamo il risultato in figura.



IdSensore	Data	Valore	Stato	idOperatore
10273	01/01/24	33	VA	1
10273	02/01/24	30	VA	
10273	03/01/24	33	VA	
10273	04/01/24	42	VA	
10273	05/01/24	26	VA	
10273	06/01/24	17	VA	
10273	07/01/24	16	VA	
10273	08/01/24	17	VA	
10273	09/01/24	19	VA	
10273	10/01/24	20	VA	
10273	11/01/24	36	VA	
10273	12/01/24	55	VA	
10273	13/01/24	56	VA	
10273	14/01/24	49	VA	
10273	15/01/24	56	VA	
10273	16/01/24	63	VA	1
10273	17/01/24	44	VA	1
10273	18/01/24	46	VA	1

Dall'analisi osserviamo come la concentrazione di PM10 nella zona Città Studi di Milano sia più alta nei mesi invernali e molto più basse nei mesi estivi, com'è ovvio aspettarsi. Inoltre, osserviamo anche due picchi in corrispondenza del primo giorno di ogni anno. Anomalia? Chiedendo all'IA possibili spiegazioni è emerso che la più comune è una combinazione di fattori:

1. **Fuochi d'artificio e petardi nella notte di San Silvestro**
I botti rilasciano in atmosfera grandi quantità di **particolato fine (PM10 e PM2.5)**, metalli e residui della combustione. In molte città europee si osservano aumenti molto forti proprio tra la notte del 31 dicembre e il 1° gennaio.

2. **Condizioni meteo invernali della Pianura Padana**
La Pianura Padana tende a trattenere gli inquinanti:

- poco vento,
- alta pressione,
- inversione termica (aria fredda intrappolata vicino al suolo),
- nebbia/umidità.

Quindi il materiale emesso dai fuochi resta “schiacciato” vicino al terreno invece di disperdersi.

3. **Contributi aggiuntivi stagionali**
A fine dicembre c’è già un fondo di PM10 elevato dovuto a:

- riscaldamento domestico,
- traffico residuo,
- combustioni varie.

Se hai visto valori tipo **250 µg/m³ il 1° gennaio** e poi **30-40 µg/m³ negli altri giorni**, il pattern è coerente con ciò che molte centraline mostrano dopo mezzanotte di Capodanno. A volte il picco dura solo alcune ore, altre volte resta elevato per un giorno intero se il meteo è stagnante.

Interessante!

