

Nota Metodologica per l'elaborazione della distribuzione della popolazione censuaria per griglia regolare rilasciata ad EUROSTAT

Roma, 9 Settembre 2024

DCAT/ATA

Davide FARDELLI

Stefano MUGNOLI

Il Regolamento di esecuzione (UE) 2018/1799 della Commissione Europea prevede la diffusione di una selezione di variabili del censimento della popolazione e delle abitazioni del 2021 geo-codificate in base ad una griglia con celle di 1 km² («dati della griglia da 1 km²»). Come da regolamento entro il 31 dicembre 2022 gli Stati membri hanno fornito ad Eurostat il totale della popolazione per singola cella (in formato provvisorio). Inoltre, entro il 31 marzo 2024, sempre gli Stati membri hanno prodotto la versione definitiva e saranno diffuse informazioni statistiche relative alla popolazione.

Di seguito si descrive la nuova metodologia utilizzata da ISTAT per la produzione dello strato informativo per la rappresentazione della Popolazione censuaria al 2021 per maglia regolare di 1 Km². Tale metodologia, che considera le metodologie adottate a livello internazionale e l'arricchimento delle fonti informative attualmente disponibili, si differenzia notevolmente dall'approccio utilizzato per la costruzione della precedente griglia, relativa alla popolazione del 2011. Tali elaborazioni sono infatti basate sul Registro Statistico di Base dei Luoghi costituito da diverse componenti:

- ✓ *Il Sistema Informativo Territoriale delle Unità Amministrative e Statistiche (Situas¹)*, che permette di documentare le variazioni amministrative occorse al territorio e di stabilire in ciascun momento l'esatta composizione del Paese in termini di Comuni, con la possibilità di ricostruire diverse geografie tematiche (Sistemi Locali del Lavoro, Aree Interne, etc.) e di rappresentarne di nuove (Bacini idrografici, Ecoregioni, etc.).
- ✓ *Le Basi Territoriali (BT) 2021*, costituite dalle sezioni di censimento: si tratta di circa 756mila unità territoriali che coprono tutto il territorio nazionale, perimetrare in funzione di una copertura omogenea della porzione di territorio identificata (ad esempio: territorio edificato, infrastrutture di trasporto, aree verdi, corsi d'acqua, ecc.).
- ✓ *Il Registro degli Indirizzi e relative coordinate geografiche*, costruito dall'integrazione di numerosi archivi amministrativi, trattati in modo da definire un codice univoco di identificazione dell'indirizzo con le rispettive coordinate geografiche e i corrispondenti indicatori di qualità.
- ✓ *Il Registro degli edifici e delle unità immobiliari*, costruito sulla base delle informazioni presenti negli archivi amministrativi catastali, di agenzie territoriali e di fonti aperte (o *open source*); le informazioni sono trattate in modo da generare un identificativo unico del fabbricato con le relative coordinate geografiche.

Per la determinazione del dato su griglia regolare è stato necessario geocodificare la popolazione censuaria ad ogni "cella" di 1 km². Tale operazione si è basata principalmente sulla disponibilità informativa riferita agli indirizzi e agli edifici in cui è censita la popolazione di riferimento.

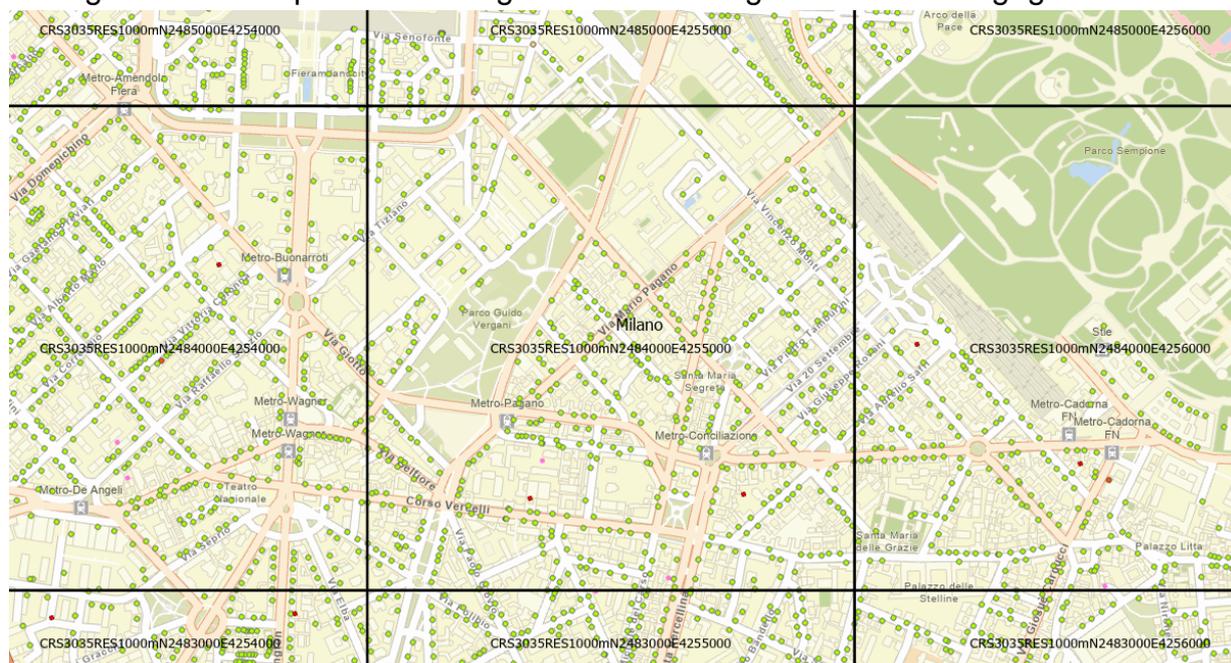
La metodologia utilizzata, denominata *bottom-up*², prevede l'aggregazione di microdati georeferenziati: questa, per poter essere attuata, richiede la disponibilità di dati geocodificati in una posizione geografica che viene, a sua volta aggregata nella cella per chilometro quadrato.

¹ Si veda il sito ufficiale <https://www.istat.it/it/archivio/296512>.

² https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Population_grids#Methodology

Questo metodo prevede la sovrapposizione (*overlapping*) delle coordinate geografiche degli indirizzi e degli edifici, ai quali è associata la popolazione residente, con la griglia elaborata Eurostat (Figura 1). Tale sovrapposizione è effettuata mediante il tool di analisi spaziale di *geoprocessing*, denominato *Spatial Join*, comune a tutti i software GIS. Attraverso tale strumento si associa ad ogni coordinata geografica il corrispondente codice di griglia in cui ricade la coordinata geografica. Il codice della griglia è fornito direttamente dalla griglia rilasciata da Eurostat. Tale associazione è univoca, ovvero per ogni coordinata si avrà esclusivamente una cella di appartenenza.

Figura 1 – Esempio di indirizzi georeferenziati e geocodificati alla griglia.



Come step successivo, si effettua il conteggio degli indirizzi e/o degli edifici all'interno di ogni cella e la somma degli individui presenti per ogni indirizzo e/o edificio a cui la popolazione risulta associata.

Questa tecnica è fortemente condizionata ed è sensibile principalmente a tre fattori:

- copertura e disponibilità di coordinate geografiche relative agli indirizzi e agli edifici;
- accuratezza posizionale delle coordinate geografiche dell'indirizzo, ovvero il grado di precisione nel posizionamento dei punti sulla carta rispetto alla realtà;
- qualità del linkage tra indirizzo, edificio e popolazione.

Laddove i tre indicatori sopradescritti si presentano con valori elevati, la qualità dell'output è alta e il prodotto finale risulta molto più accurato. La generalizzazione del dato puntuale all'interno della cella di 1 Km², porta comunque ad ottimi risultati anche per quei punti la qualità dei quali non è quella ottimale.

L'attuale griglia è stata prodotta principalmente sia attraverso la georeferenziazione degli indirizzi di residenza, sia attraverso la georeferenziazione degli edifici a cui la popolazione è associata. A supporto delle precedenti informazioni geografiche si è utilizzato anche la geocodifica alla sezione di censimento (BT2011)

degli indirizzi di residenza. Tale informazione è risultata utile sia quando le informazioni geografiche erano assenti, sia come benchmark per la valutazione qualitativa della geocodifica alla griglia. L'infrastruttura ove sono elaborati e processati i dati relativi agli indirizzi, compresi quelli degli edifici e alle coordinate geografiche è la componente indirizzi del Registro Statistico di Base dei Luoghi (RSBL).

La popolazione residente coperta da coordinata geografica è pari a circa 56,6 milioni di individui (96,1%). Quota rilevante è la copertura della popolazione anche attraverso il centroide dell'edificio pari a circa 51,5 milioni di individui (87,5%)³. Invece, una quota marginale pari a circa 500 mila individui (0,8%) è localizzabile all'interno del comune, attraverso la geocodifica alla sola sezione di censimento del 2011, senza l'utilizzo di coordinate geografiche. Quest'ultima quota di popolazione è stata geocodificata alla griglia attraverso l'uso della sezione di censimento, e, quindi, attraverso metodologie di recupero, sia deterministiche, sia probabilistiche. Ciò è stato necessario, poiché quest'ultima parte di popolazione era associata ad indirizzi per i quali era sconosciuta la collocazione sub-comunale. Si tratta principalmente di indirizzi privi di numero civico.

Pur avendo una quota consistente di coordinate geografiche, considerando le analisi spaziali effettuate, si è optato per una metodologia ibrida finalizzata alla geocodifica della popolazione alla griglia. Il dato delle due coppie di coordinate geocodificate alla griglia (indirizzi ed edifici) è stato comparato con la geocodifica alla griglia delle coordinate geografiche del centroide della sezione di censimento 2011 nella quale è stato geocodificato l'indirizzo in RSBL. In base alla convergenza o divergenza delle tre variabili geografiche predette si è scelto la fonte della geocodifica a cui protendere. Le divergenze sono la conseguenza di una non ottimale accuratezza posizionale, delle unità statistiche.

La popolazione residente geocodificata alla griglia regolare è pari a 58.933.908, con una percentuale pari a con una percentuale pari al 99,8%. La popolazione inserita nella cella virtuale è pari a 96.225. Tale popolazione è stata classificata come senza fissa dimora.

In sintesi, la popolazione è stata geocodificata con le seguenti metodologie e relative percentuali di copertura:

- uso di coordinate geografiche degli indirizzi e/o degli edifici: circa 58,4 milioni di individui (99,2%);
- uso del centroide della sezione di censimento del 2011 degli individui: circa 0,5 milioni di individui (0,8%); Quest'ultima metodologia, come detto in precedenza, è stata necessaria poiché non si disponeva di indirizzi georeferenziati alla coordinata geografica o di indirizzi geocodificati alla sezione di censimento.

³ Questo dato è riferito alle elaborazioni di marzo 2023.