

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

Luglio 2023

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). Successivamente sarà prodotta 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 2011¹. Tali elaborazioni sono infatti basate sul Registro Statistico di Base dei Luoghi costituito da diverse componenti: Situas, Registro degli indirizzi, Registro delle Basi Territoriali, Registro degli edifici e degli alloggi.

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 in cui è censita la popolazione di riferimento².

¹ Metodologia precedente (https://www.istat.it/it/files/2015/04/Nota_Tecnica.pdf)

² Con la nuova strategia del censimento permanente la geocodifica delle unità statistiche proviene dal Sistema Integrato dei Registri (SIR) implementato nell’Istituto, in particolare dall’integrazione tra il Registro Base degli Individui (RBI) e il Registro Statistico di Base dei Luoghi (RSBL).

La metodologia utilizzata, denominata *bottom-up*, 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.

Questo metodo prevede la sovrapposizione (*overlapping*) delle coordinate geografiche degli indirizzi, ai quali è associata la popolazione residente, con la griglia elaborata Eurostat³. Tale sovrapposizione è effettuata mediante il tool di analisi spaziale di *geoprocessing*, denominato *Spatial Join*⁴, 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. Come step successivo, si effettua il conteggio degli indirizzi all'interno di ogni cella e la somma degli individui presenti per ogni indirizzo.

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

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

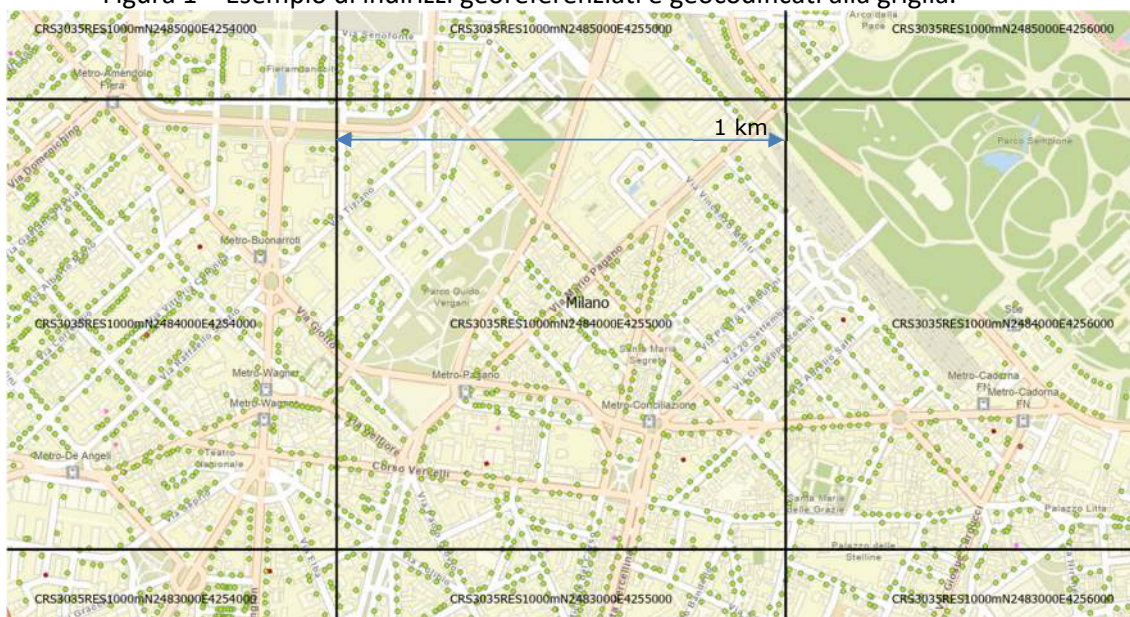
L'attuale griglia è stata prodotta, sia attraverso la georeferenziazione degli indirizzi di residenza, sia attraverso la geocodifica alla sezione di censimento degli indirizzi di residenza. L'infrastruttura ove sono elaborati e processati i dati relativi agli indirizzi e

³ Conformemente all'allegato IV, sezione 1.5, del regolamento (UE) n. 1089/2010, la griglia statistica di riferimento con celle di 1 km² per l'uso paneuropeo è la griglia di pari superficie «Grid_ETRS89-LAEA1000».

⁴ Per maggiori informazioni: <https://pro.arcgis.com/en/pro-app/latest/tool-reference/analysis/spatial-join.htm>

alle coordinate geografiche è la componente indirizzi del Registro Statistico di Base dei Luoghi (RSBL).

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



La popolazione residente coperta da coordinata geografica, in questa versione provvisoria, è pari a circa 54,5 milioni di individui (92,4%). Una quota parte pari a circa 2,8 milioni di individui (4,9%) è, invece, localizzabile all'interno del comune, attraverso la geocodifica dell'indirizzo alla sola sezione di censimento, senza l'utilizzo di coordinate geografiche. Quest'ultimo dato è stato ricavato principalmente dai risultati del Censimento della Popolazione del 2011 e dall'Archivio Nazionale dei Numeri Civici delle Strade Urbane (ANNCSU). Un ulteriore parte di popolazione, pari a 1,5 milioni di individui (2,5%), è stata geocodificata alla griglia sempre attraverso l'uso della sezione di censimento, ma è stata posizionata attraverso metodologie sia deterministiche, sia probabilistiche progettate ed elaborate a questo fine. Ciò è stato necessario, poiché si tratta di popolazione associata ad indirizzi per i quali non era possibile individuare la collocazione sub-comunale. Si tratta principalmente di indirizzi privi di numero civico.

Pur avendo una quota consistente di coordinate geografiche, si è stabilito di utilizzare, inoltre, una metodologia ibrida per poter geocodificare la popolazione alla griglia: si utilizzano sia le coordinate geografiche, sia il centroide della sezione di censimento nella quale è stato geocodificato l'indirizzo. In alcuni casi per garantire la migliore qualità possibile, poiché le coordinate geografiche non sempre hanno

un'accuratezza posizionale idonea per un'adeguata georeferenziazione degli indirizzi, si è optato di utilizzare il dato tabellare dell'associazione indirizzo e sezione di censimento proveniente da archivi pregressi, rispetto al dato geografico delle coordinate.

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 o come senza tetto.

La popolazione è stata geocodificata con le seguenti metodologie e relative percentuali di copertura:

- uso di coordinate geografiche degli indirizzi: circa 48,4 milioni di individui (82,0%);
- uso del centroide della sezione di censimento degli indirizzi: circa 9 milioni di individui (15,5%);
- uso del centroide della sezione di censimento degli individui, attribuita tramite metodologie deterministiche e probabilistiche: circa 1,5 milioni di individui (2,5%). Quest'ultima metodologia, come detto in precedenza, è stata necessaria poiché non si disponeva né di indirizzi georeferenziati alla coordinata geografica né di indirizzi geocodificati alla sezione di censimento.

Le coordinate geografiche utilizzate sono di tipo rilevate, interpolate e, infine, approssimate. Le coordinate geografiche rilevate dovrebbero indicare il punto esatto dove si trova l'accesso dall'area di circolazione. Pur avendo coordinate rilevate, è stato riscontrato che non sempre l'accuratezza posizionale rispecchia la realtà. La popolazione geocodificata alla griglia con questa tipologia di coordinate è di circa 41,83 milioni di individui, per un totale di coordinate pari a circa 7,7 milioni. La seconda tipologia di coordinate utilizzate sono quelle interpolate dal grafo stradale. La popolazione coperta da questa tipologia è pari a 2,47 milioni di individui, per un totale di coordinate pari a circa 600 mila. L'ultima tipologia di coordinate utilizzate sono quelle approssimate al numero civico più vicino. L'approssimazione massima utilizzata è quella a 10 numeri civici, indipendentemente dal lato della strada; in quest'ultima categoria sono presenti circa 4,11 milioni di individui, per un totale di coordinate pari a 1,1 milioni.

Per il prossimo rilascio definitivo della distribuzione della popolazione per griglia regolare, si terrà conto anche dell'integrazione dell'informazione geografica proveniente dal Registro degli Edifici e dalle Abitazioni, sempre nell'ambito del Registro Statistico di Base dei Luoghi.

