

Allegato 1

Nota metodologica¹

Base integrata di microdati amministrativi “Persons and Places 2018”

PSN 2661 - Popolazioni che usano un territorio e mobilità per studio e lavoro

Sommario

1	Definizioni, classificazioni	2
1.1	Definizioni	2
1.2	L’universo, le unità statistiche, le fonti e le principali variabili	3
2	L’equazione della popolazione insistente.....	5
3	Metodologia di costruzione della base informativa	6
4	Variabili rilasciate nella base finale	7

¹ A cura di Roberta Vivio vivio@istat.it e Gaia Rocchetti grocchetti@istat.it

1 Definizioni, classificazioni

1.1 Definizioni

La Popolazione Insistente è costituita dall'unione degli individui italiani e stranieri che risultano iscritti in anagrafe in Italia alla data di riferimento - sia in famiglia che in convivenza - e da quelli (sia iscritti in anagrafe che non iscritti ma con domicilio fiscale) che hanno *segnali amministrativi di lavoro o studio* sul territorio italiano.

In particolare:

- a livello nazionale è costituita da: tutti gli individui residenti in Italia, tutti gli individui che lavorano o studiano in Italia (sia residenti, sia non residenti con domicilio fiscale), dagli stranieri presenti per motivi occasionali (es. turisti). Non comprende gli italiani residenti all'estero.
- a livello comunale è costituita da: tutti gli individui che risiedono nel comune senza segnali di lavoro o studio, tutti gli individui residenti nel comune con mobilità per studio o lavoro all'interno dello stesso comune di residenza, tutti gli individui non residenti nel comune in entrata da un altro comune per studio o lavoro. Infine, comprende i non residenti nel comune che sono presenti per motivi occasionali (es. turisti). Esclude i residenti/domiciliati nel comune che lavorano/studiano in altro comune.

Nell'ambito del progetto è stata studiata una tassonomia che classifica gli individui in cinque tipologie in funzione del *motivo*, della *frequenza*, e del *verso della mobilità casa/lavoro (casa/studio)*.

Tav. 1 - Popolazione insistente per tipologia di individuo e caratteristiche dello spostamento

Tipologia di individuo		Spostamento		
strato		Motivo(*)	Verso	Frequenza
1	Individui Statici senza mobilità per lavoro/studio	Nessuna attività di Studio, lavoro	-	--
2	Individui Dinamici con mobilità in entrata per studio o lavoro	Studio, lavoro	In entrata da altra area	Giornaliera o Periodica
3	Individui Dinamici con mobilità in uscita per studio o lavoro	Studio, lavoro	In uscita verso altra area	Giornaliera o Periodica
4	Individui Dinamici con mobilità all'interno per studio o lavoro	Studio, lavoro	All'interno dell'area	Giornaliera o Periodica
5	<i>City users</i>	Turismo, eventi sportivi, salute, ecc.	In entrata da altra area	Occasionale

(*) per studio si intende la formazione scolastica di qualsiasi livello (scuola dell'infanzia, scuola primaria, scuola secondaria, università).

Per il motivo e il verso si fanno le seguenti assunzioni: che gli spostamenti avvengano dal luogo di residenza (origine) al luogo di destinazione (studio/lavoro), che ogni lavoratore/studente si sposti per

svolgere la sua attività effettuando uno ed un solo spostamento, e che uno studente si sposti per studio, un lavoratore per lavoro².

Allo stato attuale non sono disponibili informazioni utili a distinguere tra frequenza giornaliera e periodica, né quelle per la individuazione dei *city users*. L'output riguarda quindi le prime quattro tipologie di individuo.

1.2 L'universo, le unità statistiche, le fonti e le principali variabili

L'universo target è costruito integrando gli individui iscritti all'Anagrafe della Popolazione residente (Liste Anagrafiche Comunali – LAC - allo 01/01/2019), con gli individui che hanno segnali di lavoro/studio in Italia nel mese di dicembre dell'anno 2018 (Tav. 2a) e che siano registrati nell'anagrafe fiscale.

Le unità:

Lavoratore: individuo con almeno un segnale di presenza negli archivi/registri seguenti:

- Istat – Registro ASIA DB Occupazione
- MIUR - Archivio del personale universitario
- MIUR - Archivio del personale delle scuole statali
- INPDAP - Posizioni degli assicurati iscritti
- MEF - NOIPA - Cedolini stipendiali
- INPS - Rapporti di lavoro domestico
- INPS - Lavoratori autonomi in agricoltura.

Tali archivi/registri contengono informazioni relative ad occupati di imprese, enti, aziende, ecc. localizzate in Italia e facenti parte del sistema economico italiano.

Si assume che il segnale sia indicativo della condizione di “lavoratore che si reca nella sede di lavoro”, ovvero che l'informazione sia riferibile ad una persona che effettua degli spostamenti casa/lavoro.

Studente: individui con un segnale di presenza nell'archivio:

- MIUR - Anagrafe degli studenti 2018/19

L'archivio contiene le iscrizioni scolastiche in strutture italiane sul territorio italiano. Dalle scuole dell'infanzia a quelle superiori di secondo grado. Sono considerati eleggibili soltanto gli studenti per i quali è indicato che “frequentano”. Si assume che l'informazione sia riferibile ad una persona che effettua degli spostamenti casa/scuola.

Studente universitario: individui con un segnale di presenza nell'archivio:

- MIUR - Archivio degli iscritti e delle iscrizioni universitarie 2018/19.

Si assume che tale presenza sia indicativa della condizione “universitario che frequenta corsi accademici”, ovvero che l'informazione sia riferibile ad una persona che effettua degli spostamenti casa/università.

² Per gli universitari lavoratori si assume predominante la motivazione “lavoro”.

Nota:

- Gli archivi di cui sopra non contengono gli individui con iscrizione anagrafica in Italia che lavorano o studiano all'estero, ovvero oltre confine o in territori fuori giurisdizione (es. organizzazioni o università internazionali su territorio italiano). Tali individui pertanto sono conteggiati nella popolazione insistente come residenti, ma non sono conteggiati nella popolazione con mobilità per lavoro/studio. Nella stessa fattispecie rientrano anche i pensionati, i disoccupati, le casalinghe.
- le duplicazioni di segnali sono state risolte durante il processo con regole deterministiche.

Le variabili della base intermedia:

Sesso, luogo nascita e data nascita

Le variabili sesso, luogo nascita e data nascita, sono estratte dalla base dati anagrafica dell'Anagrafe tributaria, se presenti, altrimenti per ciascuna di esse è stata selezionata la modalità più frequente presente in SIM INDIVIDUI.

Le modalità fuori range sono state imputate con "NULL".

Residenza e Domicilio fiscale

La residenza anagrafica è quella della LAC al 01/01/2019; il domicilio fiscale è quello registrato in Anagrafe Tributaria al 01/01/2019.

Cittadinanza e codice famiglia

Sono estratte dall'Anagrafe della popolazione residente.

Data di decesso

E' estratta dalla base dati anagrafica dell'Anagrafe tributaria.

Comune di Origine della Mobilità per studio/lavoro: si assume che per l'individuo con un segnale di lavoro/studio sia il comune di iscrizione anagrafica, se è iscritto nelle LAC riferite al 01/01/2019; e sia il comune di domicilio fiscale riportato nella base dati dell'Anagrafe Tributaria al 01/01/2019 per chi non lo è.

- Liste anagrafiche comunali (LAC) al 01/01/2019;
- Agenzia delle Entrate -Anagrafe delle Persone Fisiche al 01/01/2019

Per gli individui appartenenti ai comuni le cui liste anagrafiche comunali non sono entrate nel processo di produzione è stata imputato, se presente, il domicilio fiscale al 01/01/2019.

Comune di destinazione della Mobilità: si assume che per l'individuo con un segnale di studio sia il comune registrato in uno dei seguenti archivi (subordinatamente al tipo di unità di analisi):

- MIUR - Anagrafe delle scuole statali e non statali
- MIUR - Anagrafica delle strutture universitarie
- MIUR - Anagrafica dei corsi universitari

Si assume che per l'individuo con un segnale di lavoro sia il comune registrato in uno dei seguenti archivi (subordinatamente al tipo di unità di analisi):

- Istat - ASIA_Unità Locali
- MEF - NOIPA - Cedolini stipendiali
- MEF - NOIPA - Uffici di Servizio
- MIUR – Personale Universitario
- MIUR – Personale Scuole
- INAIL – Posizioni assicurative territoriali
- INPS_Rapporti di Lavoro domestico
- INPS_Lavoratori Autonomi Agricoltura

Per l'identificazione del luogo di lavoro è stata studiata una procedura ad hoc di tipo gerarchico.

2 L'equazione della popolazione insistente

Si definisce *Popolazione insistente* (PI) in una data area di riferimento j :

$$\begin{aligned}
 PI_j &= NO_LUS_j + LUS_{jj}^{\leftrightarrow} + LUS_{ij}^{\rightarrow} + CU_{ij}^{\rightarrow} = \\
 &= Pop. residente - LUS_{ji}^{\leftarrow} + LUS_{ij}^{\rightarrow} + CU_{ij}^{\rightarrow} \\
 &i=1,...,n; j=1,...,n; i \neq j \\
 &(n = \text{numero di comuni}; i = \text{comune di origine}; j = \text{comune di destinazione})
 \end{aligned}$$

dove:

- (1) NO_LUS_j = *Individui statici senza attività di lavoro/studio*: individui non lavoratori, non studenti e non universitari con residenza anagrafica nel comune j (es. bambini in età pre-scolare, pensionati, casalinghe, disoccupati).
- (2) LUS_{ij}^{\rightarrow} = *Individui Dinamici con mobilità in entrata (in j) per studio o lavoro*: lavoratori, studenti o universitari con residenza anagrafica/domicilio fiscale in un comune i diverso da quello di lavoro/studio. La mobilità è in entrata nel comune di studio/lavoro j , in uscita da quello di residenza/domicilio i .
- (3) LUS_{ij}^{\leftarrow} = *Individui Dinamici con mobilità in uscita (da j) per studio o lavoro*: lavoratori, studenti o universitari con residenza anagrafica/domicilio fiscale in un comune diverso da quello di lavoro/studio. La mobilità è in entrata nel comune di studio/lavoro i , in uscita da quello di residenza/domicilio j .
- (4) $LUS_{jj}^{\leftrightarrow}$ = *Individui Dinamici con mobilità all'interno (di j) per studio o lavoro*: lavoratori, studenti o universitari con residenza anagrafica/domicilio fiscale nello stesso comune j di lavoro/studio. La mobilità è interna al comune.
- (5) CU_{ij}^{\rightarrow} = *City users*: individui con residenza anagrafica/domicilio fiscale in un comune i diverso da quello di destinazione verso cui si spostano per turismo, salute, sport, ecc. La

mobilità è in entrata nel comune di destinazione, in uscita da quello di residenza/domicilio. La frequenza è occasionale.

3 Metodologia di costruzione della base informativa

Il processo di produzione è ingegnerizzato. Gli input sono unità e variabili di fonti amministrative e registri statistici, integrati con linkage deterministico basato sulla chiave *codice_individuo_SIM*.

Per la valutazione della qualità sono stati effettuati controlli di coerenza e copertura rispetto alle fonti ufficiali e controlli sugli errori non campionari nei microdati (missing, dati anomali, fuori dominio, errori di classificazione). Sono state effettuate esclusivamente correzioni deterministiche per il recupero di valori mancanti e la ricodifica di valori anomali.

Per la tutela della riservatezza sono state realizzate aggregazioni e soppressione di alcune modalità delle variabili, sono state limitate al minimo le classificazioni incrociate e, infine, sopprese le celle a rischio di violazione. I risultati della valutazione della qualità hanno guidato la scelta del livello territoriale e delle variabili per l'output presentato.

I passi del metodo seguito sono stati sinteticamente i seguenti:

1. Ricognizione tra le fonti amministrative disponibili in Istat nel Sistema Integrato di Microdati (SIM) volta all'individuazione di quelle contenenti le unità e le variabili di interesse: ciascuna fonte è stata esaminata sia per il possibile contributo in termini di unità di rilevazione e/o di analisi, sia per quello in termini di variabili dirette, derivate, di controllo. Si è constatato che non esistono informazioni amministrative utili a quantificare la categoria degli occasionali e a determinare la frequenza degli spostamenti che consentirebbe di distinguere tra *giornalieri* ed i *periodici*. Anche la categoria dei *city users* non è al momento quantificabile attraverso i dati amministrativi³. Nella selezione delle fonti è stato seguito il principio di pertinenza e non ridondanza che ha portato a concludere per una scelta limitata ad alcune variabili estratte da 21 archivi. Dal momento che le fonti amministrative, se considerate separatamente, non contengono tutte le variabili di interesse e, nella maggior parte dei casi, si riferiscono a una parte della popolazione target, si rende necessaria la loro integrazione. L'integrazione consente di ottenere sia una completezza delle variabili di interesse (per esempio, per arrivare alla copertura della variabile "comune di lavoro" occorre acquisire la stessa informazione da una decina di archivi, unificarla e armonizzarla) sia una copertura pressoché totale dell'universo di riferimento (ottenuta unificando il segnale di presenza sul territorio da 13 archivi).
2. Utilizzo combinato di archivi e variabili guidato dai risultati della ricognizione e dalla progettazione: alcuni entrano nella fase finalizzata all'individuazione della popolazione residente, altri in quella dei non residenti con segnali di attività in Italia (studenti, studenti universitari, lavoratori), altri nella fase di individuazione dei comuni di origine/destinazione degli spostamenti.
3. Validazione a livello di microdato: controllo e correzione sugli errori non campionari (fuori dominio, missing, ecc.).
4. Validazione a livello macro: controllo rispetto a fonti di benchmark ufficiali.

³ Come sviluppo futuro è possibile studiare modalità di integrazione di dati aggregati a livello comunale rilevati da indagine.

5. Realizzazione della base integrata finale in cui ogni record corrisponde ad uno strato di unità. Gli strati corrispondono alle combinazioni delle modalità delle variabili finali. In nessuno strato è presente una numerosità inferiore alla soglia che permette l'identificazione delle unità.

4 Variabili rilasciate nella base finale

Le variabili e le relative modalità della base finale sono:

Var.1. **Comune di origine della mobilità:** residenza, domicilio fiscale, null;

Var.2. **Comune di destinazione della mobilità:** comune di lavoro, di studio, null;

Var.3. **Tipo di attività individuo:** lavoratore, studente, universitario, altro;

Var.4. **Tipologia di insidenza dell'individuo:**

1. Individui Statici (senza mobilità per lavoro/studio),
2. Individui Dinamici con mobilità in entrata (per studio o lavoro),
3. Individui Dinamici con mobilità in uscita (per studio o lavoro),
4. Individui Dinamici con mobilità all'interno dell'area (per studio o lavoro).

Per ogni combinazione di modalità viene rilasciata la relativa numerosità di strato, ovvero il numero di individui che appartengono allo strato.

Per garantire una maggiore qualità del dato, i codici dei comuni di origine e destinazione che hanno meno di 5.000 abitanti sono posti a *null*, mantenendo solo il codice provincia.